

**REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE**  
**MINISTERE DE L'ENSEIGNEMENT SUPERIEURE ET DE LE RECHERCHE SCIENTIFIQUE**  
**UNIVERSITE FERHAT ABBAS SETIF 1**  
**FACULTE DES SCIENCES**  
**DEPARTEMENT DE MATHEMATIQUES**

**THESE**

Présentée par :

**GUERRA Loubna**

Pour obtenir le titre de **Doctorat en Sciences**

**OPTION**

**MATHEMATIQUES APPLIQUEES**

**THEME**

**Méthodes de points intérieurs et fonctions noyaux pour  
l'optimisation quadratique semi-définie convexe**

**Soutenue le : 15 / 11 / 2018**

**Devant le jury composé de :**

<b>Président :</b>	<b>Mr. D. Benterki</b>	<b>Prof.</b>	<b>U.F.A. Sétif 1</b>
<b>Rapporteur :</b>	<b>Mr. M. Achache</b>	<b>Prof.</b>	<b>U.F.A. Sétif 1</b>
<b>Examineurs :</b>	<b>Mr. N. Benhamidouche</b>	<b>Prof.</b>	<b>U.M.B. M'sila</b>
	<b>Mr. A. Merzougui</b>	<b>M.C.A</b>	<b>U.M.B. M'sila</b>

**Année universitaire : 2018 / 2019**

# Remerciements

Avant tout; je tiens à remercier **DIEU** tout puissant pour la volonté, la patience qu'il m'a données toutes ces longues années d'études afin que je puisse arriver à ce stade.

Je voudrais tout d'abord remercier grandement mon directeur de Thèse; Monsieur **MOHAMED ACHACHE** Professeur à l'université Ferhat Abbas Sétif 1; pour tout son aide, je suis ravie d'avoir travaillé en sa compagnie car outre son appui scientifique, il a toujours été là pour m'encourager et me conseiller au cours de l'élaboration de cette Thèse.

Je tiens à exprimer mes plus vifs remerciements à Monsieur **DJAMEL BENTERKI** Professeur à l'université Ferhat Abbas Sétif 1; de l'honneur qu'il m'a fait en acceptant d'être président de jury, et je le remercie aussi pour ses conseils scientifiques.

J'exprime ma gratitude à Messieurs : **NOUREDDINE BENHAMIDOUCHE** ; Professeur; et **ABDELKRIM MERZOUGUI** ; maître de conférence classe A ; à l'université Mohamed Boudiaf, M'Sila ; d'être membres de jury et je les remercie également pour le temps consacré à la lecture de ce travail.

Enfin, les mots les plus simples étant les plus forts, j'adresse toute mon affection à ma famille et en particulier à mes parents, leur présence et leur encouragement sont pour moi les piliers fondateurs de ce que je suis et de ce que je fais, merci pour avoir fait de moi ce que je suis aujourd'hui ; est-ce un bon endroit pour dire ce genre de chose ? Je n'en connais en tout cas pas de mauvais.

Je remercie également mes frères et ma chère sœur ; merci pour votre support et vos encouragements.

Mention spéciale à ma belle mère et à mon cher époux pour leurs encouragements et leur soutien morale.

Je remercie tous ceux qui m'ont aidé de près ou de loin pour établir cette Thèse.

J'oublie sûrement certaines personnes qui par un geste ou une discussion ont apporté bien plus que l'acte, qu'ils reçoivent encore une fois mes sincères remerciements.

# Table des matières

<b>1</b>	<b>Calcul matriciel, analyse convexe et programmation mathématique</b>	<b>10</b>
1.1	Rappel sur les matrices . . . . .	10
1.2	Rappel d'analyse convexe . . . . .	14
1.3	Programmation mathématique . . . . .	15
<b>2</b>	<b>Programmation quadratique convexe semi-définie et méthodes de points intérieurs</b>	<b>20</b>
2.1	Position du problème . . . . .	20
2.1.1	Problème primal . . . . .	20
2.1.2	Problème dual . . . . .	21
2.2	L'importance de la programmation quadratique semi-définie . . . . .	22
2.2.1	Modélisation de quelques problèmes d'optimisations convexe en un problème CQSDP . . . . .	23
2.3	Dualité en CQSDP . . . . .	25
2.3.1	Dualité faible . . . . .	25
2.3.2	Conditions d'optimalité nécessaires et suffisantes de la solution optimale . . . . .	26
2.3.3	Complémentarité en CQSDP . . . . .	27
2.4	Méthodes de points intérieurs pour résoudre le problème CQSDP . . . . .	28
2.4.1	Méthode de la trajectoire centrale de type primal-dual pour CQSDP . . . . .	28
2.4.2	Les directions classiques . . . . .	31
2.4.3	La mesure de proximité . . . . .	34
2.4.4	Algorithme . . . . .	34
2.4.5	La convergence de l'algorithme et l'analyse de la complexité . . . . .	35
<b>3</b>	<b>Méthodes de points intérieurs basée sur une nouvelle fonction noyau pour CQSDP</b>	<b>48</b>
3.1	Les fonctions noyaux et leurs propriétés . . . . .	48

3.1.1	Notion d'une fonction noyau . . . . .	49
3.1.2	Qualification d'une fonction noyau . . . . .	50
3.1.3	Propriétés d'une fonction noyau éligible . . . . .	54
3.2	Nouvelles directions cherchées . . . . .	55
3.2.1	Algorithme . . . . .	56
3.2.2	Une borne supérieure de $\Psi(V)$ après chaque itération externe . .	58
3.2.3	La décroissance de la fonction barrière durant une itération interne	59
3.2.4	Borne de $\delta(V)$ en terme de $\Psi(V)$ . . . . .	62
3.2.5	Borne d'itérations . . . . .	62
3.3	La convergence de l'algorithme et l'analyse de la complexité . . . . .	64
3.3.1	Schéma pour analyser un algorithme basé sur une fonction noyau éligible . . . . .	64
3.3.2	Analyse de la complexité . . . . .	66
<b>4</b>	<b>Tests numériques</b>	<b>72</b>
4.1	Calcul de la matrice $V$ . . . . .	72
4.2	Calcul du pas de déplacement . . . . .	74
4.3	Problèmes à tester . . . . .	74
4.3.1	Problèmes à taille variable . . . . .	75
4.3.2	Problèmes à taille fixe . . . . .	76
4.3.3	Problème de $(\mathcal{NCM})$ . . . . .	80
4.4	Résultats numériques . . . . .	83
4.4.1	Résultats numériques pour l'Algorithme 2.4.4 à petit pas . . . . .	84
4.4.2	Résultats comparatifs pour l'Algorithme 3.2.1 à grand pas . . . . .	85
	<b>Conclusion générale et perspectives</b>	<b>98</b>
	<b>Bibliographie</b>	<b>100</b>
	<b>Annexe</b>	<b>104</b>

# Introduction générale

Ces dernières années, le problème semi-défini a fait l'objet de nombreuses études de recherche dans le domaine d'optimisation. Dans ce travail, on s'intéresse au problème quadratique convexe semi-défini (CQSDP) qui est d'une part, une généralisation d'un problème linéaire (LP), quadratique (QP) et ainsi un problème linéaire semi-défini (SDP) et d'autre part, il couvre plusieurs problèmes en architecture et en génie électrique.

Le problème quadratique semi-défini est un problème d'optimisation récent et qui acte sur l'ensemble des matrices symétriques semi définies positives, donc l'ensemble des contraintes est un cône non polyédrique ce qui fait les méthodes classiques simplicielles ne sont pas valables. Pour cela, on utilise les méthodes de points intérieurs qui ont été initialisées pour la première fois par Karmarkar en 1984. Depuis cette date, de nombreux algorithmes ont été développés pour la programmation linéaire [22, 31], la programmation quadratique [2] et la programmation semi-définie (SDP) [16-18, 24, 27, 32-34, 37] et ils ont ensuite été généralisés pour la programmation quadratique semi-définie [6, 25, 26, 36, 38]. Les méthodes de points intérieurs se distinguent en quatre catégories :

- **Méthodes affines.**
- **Méthodes projectives.**
- **Méthodes de réduction de potentiel.**
- **Méthodes de trajectoire centrale.**

On note que l'efficacité de quelques méthodes de points intérieurs a été étudiée pour résoudre CQSDP. Nie et Yuan [25], ont développé un algorithme de la réduction de potentiel pour CQSDP et ils ont utilisé les directions de NT. Dans [26], ils ont présenté un algorithme de points intérieurs de type prédicteur-correcteur en combinant le pas de Dikin-type et le pas de Newton et ils ont utilisé les directions de Nesterov Todd (NT) où ces dernières sont calculées par la méthode du gradient conjugué. Dans [35, 36], Toh et al ont présenté un algorithme de trajectoire centrale de type prédicteur-correcteur pour des classes spéciales du problème convexe CQSDP et ils ont montré l'efficacité numérique de

ses algorithmes. Wang et Bai [38], ont utilisé les fonctions noyaux pour résoudre le problème CQSDP par une méthode de points intérieurs primal-dual. En utilisant les fonctions noyaux, Wang et Zhu [41] ont présenté un algorithme de trajectoire centrale (TC) de type primal-dual pour résoudre le problème CQSDP basé sur le schéma de symétrisation de NT.

Actuellement, les méthodes de trajectoire centrale de type primal-dual sont les meilleures méthodes de points intérieurs car ces algorithmes admettent un bon comportement théorique telle que la complexité polynomiale et d'autre part, ils sont de type Newton ce qui conduit à une efficacité numérique. Dans [17, 18], De Klerk et al; ont proposé un algorithme de TC de type primal-dual à petit pas pour résoudre un problème semi-défini SDP basé sur la direction de Nesterov-Todd, ils ont montré que l'algorithme à une meilleure complexité polynomiale qui est de l'ordre  $O(\sqrt{n} \log \frac{n}{\epsilon})$ .

La majorité des méthodes de points intérieurs de type primal-dual qui ont la complexité polynomiale pour LP utilisent la fonction logarithmique classique comme étant une fonction barrière. Mais, actuellement il existe un écart entre le comportement pratique de ces algorithmes et les résultats théoriques de performance, surtout pour les algorithmes à grand pas pour lesquels la borne théorique d'itérations est  $O(n \log \frac{n}{\epsilon})$ . En pratique, les algorithmes à grand pas sont plus efficaces que les algorithmes à petit pas pour lesquels la borne théorique d'itérations n'est que  $O(\sqrt{n} \log \frac{n}{\epsilon})$ . Cet écart important entre la théorie et la pratique se rapporte à l'inconvénient des méthodes de points intérieurs (voir [9, 28]).

Les fonctions noyaux jouent un rôle important dans la conception de nouveaux algorithmes de points intérieurs de type primal-dual. Premièrement, Peng et al. [28] ont présenté des méthodes de points intérieurs pour LP et SDP basées sur les fonctions barrières auto-régulière. Puis, Bai et al. [9-11], ont proposé une classe des méthodes de points intérieurs de type primal-dual pour LP basé sur une variété de fonctions noyaux non-auto-régulière et ils ont obtenu la même complexité favorable pour les algorithmes à grand et petit- pas que [28]. De plus, Bai et ses co-auteurs ont fait l'extension des résultats ci-dessus obtenus pour LP à SDP [39] et CQSDP [38]. On note qu'ils y a des algorithmes similaires qui sont prolongés avec succès à un problème quadratique convexe sur un cône symétrique (CQSCO) (voir [15, 40]). Pour d'autre algorithmes de points intérieurs basé sur les fonctions noyaux on cite [1, 5, 14, 19, 37, 41, 44, 46, 47]. Notons que la meilleure borne d'itérations pour les algorithmes à grand pas qui est d'ordre  $O(\sqrt{n} \log n \log \frac{n}{\epsilon})$ , a été obtenue jusqu'à présent par quelques fonctions noyaux barrières

paramétrées dans le tableau suivant :

La fonction noyau	Paramètre	Références
$\frac{t^2-1}{2} - \frac{t^{1-q}-1}{q-1}, q > 1$	$q = \frac{1}{2} \log n$	[28, 29]
$\frac{t^2-1}{2} - \frac{t^{1-q}-1}{q-1} - \frac{q-1}{q}(t-1), q > 1$	$q = \frac{1}{2} \log n$	[28, 29]
$\frac{t^{p+1}-1}{p+1} - \frac{t^{1-q}-1}{q-1}, p \in [0, 1], q > 1$	$p = 1, q = \log n$	[11]
$\frac{t^2-1}{2} - \int_1^t e^{q(\frac{1}{\xi}-1)} d\xi, q \geq 1$	$q = O(\log n)$	[10]
$\frac{t^2-1}{2} + \frac{1}{\log q} \frac{t^{\frac{1}{q}-1}-1}{\log q}, q > 1$	$q = 1 + O(n)$	[1]

Tableau 1. Cinq fonctions noyaux paramétrées avec une meilleure borne d'itérations pour les algorithmes à grand pas

Dernièrement, M.W. Zhang [44], a présenté un algorithme de points intérieurs de type primal-dual pour résoudre un problème CQSDP, basé sur la fonction noyau suivante :

$$\psi(t) = \frac{t^2 - 1}{2} - (t - 1) e^{(\frac{1}{t}-1)}, t > 0,$$

il a montré que la complexité correspondant à cet algorithme à grand pas est d'ordre  $O(\sqrt{n}(\log n)^2 \log \frac{n}{\epsilon})$ .

L'objectif de cette thèse est :

- La résolution d'un problème quadratique convexe semi-défini (CQSDP) par la méthode de trajectoire centrale dont les directions ne sont pas orthogonales (contrairement au cas SDP), cela conduit à une difficulté de l'analyse de la complexité. On note que sous un choix spécial du paramètre  $\theta$  ( $\theta = O(\frac{1}{n})$ ), on montre que notre algorithme de trajectoire centrale admet une meilleure complexité polynomiale qui est de l'ordre  $O(\sqrt{n} \log \frac{n}{\epsilon})$  (voir [6]).

- L'introduction d'une nouvelle fonction noyau paramétrée définie par :

$$\psi(t) = \frac{t^2 - 1}{2} - \frac{t - q}{q^2 - q + 1} e^{q(\frac{1}{t}-1)} + \frac{1 - q}{q^2 - q + 1}, t > 0, q \geq 1,$$

où  $q$  est un paramètre, on note que si  $q = 1$ , la fonction se rapporte au fonction noyau de M.W. Zhang. Sous un choix spécial du paramètre  $q$ , on montre que la complexité d'un algorithme à grand pas qui est basée sur cette fonction noyau se réduit par un facteur de  $O(\log n)$  par rapport à la complexité obtenue par M.W. Zhang en [44] c'est-à-dire on obtient la meilleure complexité connue jusqu'à présent pour ces algorithmes à grand pas qui est de l'ordre  $O(\sqrt{n} \log n \log \frac{n}{\epsilon})$ . De plus, notre analyse est basée sur la programmation semi-définie (SDP) et on prend en considération que les directions de Nesterov-Todd

ne sont pas orthogonales.

Finalement, quelques résultats numériques comparatifs sont présentés pour voir le comportement numérique de ces deux algorithmes.

La thèse est composée de quatre chapitres :

- Dans le premier chapitre, on présente les notions fondamentales de calcul matriciel et d'analyse convexe ainsi la programmation mathématique qui seront utiles par la suite de la thèse.

- Le deuxième chapitre est consacré à la définition des deux problèmes primal et dual de la programmation quadratique semi-définie et le développement d'un algorithme de TC de type primal-dual à petit pas basé sur les directions classiques. Afin de déterminer ces directions de déplacement, on utilise le schéma de symétrisation de NT et on termine ce chapitre par l'analyse de la complexité de cet algorithme.

- Dans le troisième chapitre, on développe un algorithme de points intérieurs de type primal-dual basé sur une nouvelle fonction noyau qui doit être éligible. On suit le schéma présenté par Bai et al en 2004 [9] pour analyser la complexité de cet algorithme proposé.

- Dans le dernier chapitre, on présente des résultats numériques issus de l'application de deux algorithmes; un algorithme de trajectoire centrale de type primal-dual et un algorithme de points intérieurs de type primal-dual basé sur une nouvelle fonction noyau; sur des problèmes CQSDP.

On terminera cette thèse par une conclusion générale et perspectives.



# Notations

$\mathbb{R}^n$	:	L'espace des vecteurs réels de dimension $n$ ,
$\mathbb{R}_+^n$	:	L'orthant positif de $\mathbb{R}^n$ ,
$\mathbb{R}^{m \times n}$	:	L'espace vectoriel des matrices réelles de taille $(m \times n)$ ,
$A^\top$	:	Le transposé d'une matrice $A \in \mathbb{R}^{m \times n}$ ,
$a_{ij}$	:	Élément de la matrice $A \in \mathbb{R}^{m \times n}$ ,
$\mathbb{S}^n$	=	$\{X : X \in \mathbb{R}^{n \times n}, X = X^\top\}$ ,
$A \succeq 0$ ( $A \succ 0$ )	:	$A$ est une matrice semi définie positive (définie positive),
$A \preceq 0$ ( $A \prec 0$ )	:	$A$ est une matrice semi définie négative (définie négative),
$\mathbb{S}_+^n$	=	$\{X : X \in \mathbb{S}^n, X \succeq 0\}$ ,
$\mathbb{S}_{++}^n$	=	$\{X : X \in \mathbb{S}^n, X \succ 0\}$ ,
$\lambda_i(A)$	:	La $i^{\text{ème}}$ valeur propre de $A \in \mathbb{R}^{n \times n}$ ,
$sp(A)$	:	Le spectre de la matrice $A$ ,
$\lambda_{\max}(A)$	=	$\max_i \lambda_i(A)$ , si $\lambda_i(A) \in \mathbb{R}, \forall i$ ,
$\lambda_{\min}(A)$	=	$\min_i \lambda_i(A)$ , si $\lambda_i(A) \in \mathbb{R}, \forall i$ ,
$\mathbf{Tr}(A)$	=	$\sum_i a_{ii} = \sum_i \lambda_i(A)$ , (la trace d'une matrice $A \in \mathbb{R}^{n \times n}$ ),
$\det(A)$	=	$\prod_i \lambda_i(A)$ , (déterminant de $A \in \mathbb{R}^{n \times n}$ ),
$\rho(A)$	=	$\max_i  \lambda_i(A) $ , (le rayon spectral de $A$ ),
$\ A\ _F^2$	=	$\mathbf{Tr}(AA^\top) = \sum_i \sum_j a_{ij}^2$ ,
$\ A\ _2$	=	$\sqrt{\rho(A^\top A)}$ , la norme spectrale de $A$ ,
$A \bullet B$	=	$\mathbf{Tr}(A^\top B)$ , $\forall A, B \in \mathbb{R}^{n \times n}$ ,
$A^{\frac{1}{2}}$	:	La racine carrée de la matrice $A \succ 0$ ,
$A^{-1}$	:	L'inverse d'une matrice régulière $A$ ,
$A \sim B$	:	Il existe une matrice inversible $P$ telle que $A = PBP^{-1}$ ; (semblable),
$I$	:	La matrice identité d'ordre $n$ ,
$0_{m \times n}$	:	La matrice zéro $(m \times n)$ ,
$\text{diag}(x)$	=	$X$ la matrice diagonale avec $X_{ii} = x_i$ ,
$X^{(k)}$	:	Le $k^{\text{ème}}$ terme d'une suite des matrices,
$e$	:	Un vecteur de $\mathbb{R}^n$ tel que $e_i = 1, \forall i = 1, \dots, n$ ,
$f(x) = \Theta(g(x))$	$\Leftrightarrow$	$\exists k_1, k_2 > 0 : k_1 g(x) \leq f(x) \leq k_2 g(x)$ , pour tout $x > 0$ ,
$f(x) = O(g(x))$	$\Leftrightarrow$	$\exists k > 0 : f(x) \leq kg(x)$ , pour tout $x > 0$ .

# Terminologie

<b>CQSDP</b>	:	Programmation quadratique semi-définie,
<b>SDP</b>	:	Programmation linéaire semi-définie,
<b>(P)</b>	:	Le problème quadratique semi-défini sous forme primal,
<b>(D)</b>	:	Le problème dual de <b>(P)</b> ,
<b>TC</b>	:	La trajectoire centrale,
<b>CPI</b>	:	Condition de points intérieurs,
<b>K.K.T</b>	:	Karush-Kuhn-Tucker,
<b>NT</b>	:	Nesterov et Todd,
<b>SVD</b>	:	La décomposition en valeurs singulières,
<b>c-à-d</b>	:	C'est à dire,
$\nabla f(X)$	:	Le gradient de $f$ au point $X$ , où $f : \mathbb{S}^n \rightarrow \mathbb{R}$ ,
$\nabla^2 f(X)$	:	Le Hessien de $f$ au point $X$ , où $f : \mathbb{S}^n \rightarrow \mathbb{R}$ ,
$\frac{\partial f(X)}{\partial x_{ij}}$	:	Les dérivées partielles de $f$ au point $x_{ij}$ , où $X \in \mathbb{S}^n$ ,
$\mathcal{C}^k$	:	Espace des fonctions $k$ fois continûment différentiables.

# Chapitre 1

## Calcul matriciel, analyse convexe et programmation mathématique

Ce chapitre introductif contient un bagage mathématique élémentaire, qui permet de faciliter la compréhension de cette thèse, notamment les normes, les matrices symétriques, les matrices semi-définies positives et ses propriétés, ainsi quelques notions de base de l'analyse convexe et de la programmation mathématique qui seront utilisées par la suite.

### 1.1 Rappel sur les matrices

Dans cette partie, on présente quelques résultats connus sur les normes, les matrices symétriques ainsi les matrices semi-définies.

#### 1.1.1 Produit scalaire et normes

On commence par la définition du produit scalaire de deux vecteurs.

**Définition 1.1.1** *Le produit scalaire usuel de deux vecteurs  $x$  et  $y$  de  $\mathbb{R}^n$  est défini par :*

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i = x^\top y.$$

De même, on définit un produit scalaire sur l'ensemble des matrices carrées réelles :

**Définition 1.1.2** *Soient  $A, B \in \mathbb{R}^{n \times n}$ , le produit scalaire de  $A$  et  $B$  noté  $A \bullet B$  est défini par :*

$$A \bullet B = \mathbf{Tr}(A^\top B) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} b_{ij} = B \bullet A.$$

On rappelle que  $\mathbf{Tr}(\cdot)$  désigne la trace d'une matrice de  $\mathbb{R}^{n \times n}$ . Il est important de signaler que la trace est une fonction linéaire, de plus elle vérifie les propriétés suivantes :

1.  $\forall A, B \in \mathbb{R}^{n \times n} : \mathbf{Tr}(AB) = \mathbf{Tr}(BA)$ ,
2.  $\forall A \in \mathbb{R}^{n \times n} : \mathbf{Tr}(A) = \mathbf{Tr}(A^\top)$ ,
3.  $\forall A, B \in \mathbb{R}^{n \times n} : A \sim B \Rightarrow \mathbf{Tr}(A) = \mathbf{Tr}(B)$ .

Enonçons maintenant la notion de norme vectorielle.

**Définition 1.1.3** La norme vectorielle est une application de  $\mathbb{R}^n$  dans  $\mathbb{R}_+$ , notée par  $\|\cdot\|$  et vérifie les conditions suivantes :

1.  $\forall x \in \mathbb{R}^n : \|x\| = 0 \Leftrightarrow x = 0$ ,
2.  $\forall \lambda \in \mathbb{R}, \forall x \in \mathbb{R}^n : \|\lambda x\| = |\lambda| \|x\|$ ,
3.  $\forall x, y \in \mathbb{R}^n : \|x + y\| \leq \|x\| + \|y\|$ .

La norme matricielle associée au produit scalaire de deux matrices est définie par :

**Définition 1.1.4** Soient  $A, B \in \mathbb{R}^{n \times n}$ , l'application  $\|\cdot\| : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}_+$  est appelée norme matricielle si elle vérifie les conditions suivantes :

1.  $\|A\| = 0 \Leftrightarrow A = 0, \forall A \in \mathbb{R}^{n \times n}$ ,
2.  $\|\alpha A\| = |\alpha| \|A\|, \forall \alpha \in \mathbb{R}$ ,
3.  $\|A + B\| \leq \|A\| + \|B\|, \forall A, B \in \mathbb{R}^{n \times n}$ ,
4.  $\|AB\| \leq \|A\| \|B\|, \forall A, B \in \mathbb{R}^{n \times n}$ .

On note que les normes matricielles usuelles pour toute  $A \in \mathbb{R}^{n \times n}$  sont :

$$\|A\|_1 = \max_{j=1}^n \left( \sum_{i=1}^n |a_{ij}| \right), \|A\|_\infty = \max_{i=1}^n \left( \sum_{j=1}^n |a_{ij}| \right) \text{ et } \|A\|_2 = \sqrt{\rho(A^\top A)},$$

la dernière norme est appelée la norme spectrale. On utilisera également la norme de Frobenius :

$$\|A\|_F = \sqrt{A \bullet A} = \sqrt{\sum_{i=1}^n \sum_{j=1}^n (a_{ij})^2}.$$

Si  $A = A^\top$ , alors on obtient facilement les résultats suivants :

$$\|A\|_F = \sqrt{\mathbf{Tr}(A^\top A)} = \sqrt{\mathbf{Tr}(A^2)} = \sqrt{\sum_{i=1}^n \lambda_i^2(A)},$$

et

$$\|A\|_2 = \sqrt{\rho(A^2)} = \max_{i=1}^n |\lambda_i(A)|,$$

de plus,

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{n} \|A\|_2,$$

et pour toute norme matricielle on a :

$$\rho(A) \leq \|A\|.$$

## 1.1.2 Matrices (semi-) définies positives

Enonçons maintenant la notion d'une matrice (semi-) définie positive. Puis, on présente certaines propriétés qui seront utilisées par la suite.

**Définition 1.1.5** Soit  $A \in \mathbb{R}^{n \times n}$ .

$A$  est dite semi-définie positive si et seulement si :  $x^\top A x \geq 0, \forall x \in \mathbb{R}^n$ .

$A$  est dite définie positive si et seulement si :  $x^\top A x > 0, \forall x \in \mathbb{R}^n, x \neq 0$ .

**Théorème 1.1.1** Soit  $A \in \mathbb{S}^n$ , les conditions suivantes sont équivalentes :

1.  $A \in \mathbb{S}_+^n$  (resp  $A \in \mathbb{S}_{++}^n$ ),
2.  $\lambda_{\min}(A) \geq 0$  ( $\lambda_{\min}(A) > 0$ ),
3.  $\exists P \in \mathbb{R}^{n \times n} : A = P^\top P$  (resp  $\exists P \in \mathbb{R}^{n \times n} : \text{rg}(P) = n, A = P^\top P$ ).

**Proposition 1.1.1** : Soit  $A \in \mathbb{S}_{++}^n$ , alors il existe une matrice unique  $B \in \mathbb{S}_{++}^n$  telle que  $A = B^2$ , et on la note souvent par  $B = A^{\frac{1}{2}}$ . De plus,  $B$  est appelée la racine carrée de  $A$ .

Dans la suite de cette partie, on s'intéresse aux matrices symétriques semi-définies positives. On définit sur  $\mathbb{S}^n$  l'ordre de Löwner par :

$$A \succcurlyeq B \Leftrightarrow A - B \succcurlyeq 0, \text{ (resp } A \succ B \Leftrightarrow A - B \succ 0),$$

et on a la proposition suivante :

**Proposition 1.1.2** : Soient  $A, B \in \mathbb{S}_+^n$ , alors :

1.  $A + B \succcurlyeq B$ .
2.  $A^{\frac{1}{2}} B A^{\frac{1}{2}} \succcurlyeq 0$ .
3.  $\text{Tr}(AB) \leq \text{Tr}(A)\text{Tr}(B)$ .

4.  $\mathbf{Tr} (AB) \geq 0$ .

Les prochains lemmes seront utiles par la suite de la thèse.

**Lemme 1.1.1** : Soient  $A, B \in \mathbb{S}_+^n$ , les assertions suivantes sont équivalentes :

1.  $A \bullet B = 0$ ,
2.  $AB = 0$ ,
3.  $\frac{1}{2}(AB + BA) = 0$ .

**Lemme 1.1.2** : Soient  $A, B \in \mathbb{S}_+^n$ . Alors :

$$\lambda_{\min}(A)\lambda_{\max}(B) \leq \lambda_{\min}(A) \mathbf{Tr} (B) \leq A \bullet B \leq \lambda_{\max}(A) \mathbf{Tr} (B) \leq n\lambda_{\max}(A)\lambda_{\max}(B).$$

### 1.1.3 Théorème de décomposition en valeurs singulières (en anglais SVD)

La décomposition en valeurs singulières (SVD) permet la généralisation de la notion de valeurs propres aux matrices rectangulaires de taille  $(m \times n)$  et la construction d'inverses généralisés ([13], [30]).

**Lemme 1.1.3** Soit  $A \in \mathbb{R}^{m \times n}$ , alors la matrice  $A^\top A \in \mathbb{S}_+^n$ .

**Définition 1.1.6** Soit  $A \in \mathbb{R}^{m \times n}$ . Les valeurs singulières d'une matrice  $A$  sont les racines carrées des valeurs propres de  $A^\top A$ .

**Théorème 1.1.2** Soit  $A \in \mathbb{R}^{m \times n}$  avec les valeurs propres de  $A^\top A$  non toutes nulles. Il existe deux matrices unitaires  $U \in \mathbb{R}^{m \times m}$  et  $V \in \mathbb{R}^{n \times n}$  et une matrice  $\Sigma \in \mathbb{R}^{m \times n}$  telles que :

1.  $U^\top AV = \Sigma$ .
2.  $\Sigma = \begin{bmatrix} D & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix}$  où  $D = \text{diag}(\sigma_1, \dots, \sigma_r)$  avec  $r = \min(m, n)$  et  $\sigma_1 \geq \sigma_2 > \dots > \sigma_r > 0$  sont les valeurs singulières non nulles de  $A$ .
3. Les vecteurs colonnes  $v_1, \dots, v_r$  de  $V$  sont des vecteurs propres de  $A^\top A$  associés à  $\sigma_1^2, \dots, \sigma_r^2$ .  
Les vecteurs colonnes  $u_1, \dots, u_r$  de  $U$  sont des vecteurs propres de  $AA^\top$  associés à  $\sigma_1^2, \dots, \sigma_r^2$ .

## 1.2 Rappel d'analyse convexe

La notion de convexité est un outil mathématique d'importance capitale pour l'étude théorique et numérique des problèmes d'optimisation. A ce propos, on présente dans ce paragraphe quelques notions de base d'usage courant.

### 1.2.1 Ensembles et applications affines

**Définition 1.2.1** *Un sous-ensemble  $F$  de  $\mathbb{R}^n$  est dite affine si :*

$$\forall x, y \in F, \text{ et } \forall \lambda \in \mathbb{R} : (1 - \lambda)x + \lambda y \in F.$$

*On dit aussi que  $F$  est une variété affine (ou linéaire).*

**Définition 1.2.2** *Une application  $f$  est dite affine sur  $F \subset \mathbb{R}^n$  si :*

$$\forall x, y \in F, \text{ et } \forall \lambda \in \mathbb{R} : f[(1 - \lambda)x + \lambda y] = (1 - \lambda)f(x) + \lambda f(y).$$

*Si de plus  $f(0) = 0$  alors  $f$  est linéaire. Plus généralement tout fonction affine  $f$  définie sur  $F \subset \mathbb{R}^n$  s'écrit sous la forme :*

$$f(x) = c^\top x + d, \text{ où } c \in \mathbb{R}^n \text{ et } d \in \mathbb{R}.$$

### 1.2.2 Ensembles, Cônes et Fonctions convexes

**Définition 1.2.3** *Un sous-ensemble  $C$  de  $\mathbb{R}^n$  est dit convexe si :*

$$\forall x, y \in C : [x, y] = \{(1 - \lambda)x + \lambda y, \lambda \in [0, 1]\} \subset C.$$

*Autrement dit, si le segment de droite joignant deux points quelconques  $x, y$  est entièrement inclus dans  $C$ .*

Les opérations algébriques suivantes conservent la convexité :

- L'intersection quelconque.
- Le produit cartésien.
- Les transformations affines.
- Les combinaisons linéaires  $\sum_{i=1}^m \alpha_i C_i$ , ( $\alpha_i \in \mathbb{R}$  et  $m \in \mathbb{N}$ ).
- La translation  $C + a$  avec  $a \in \mathbb{R}^n$ .

**Définition 1.2.4** Un sous-ensemble  $K$  de  $\mathbb{R}^n$  est appelé cône si

$$\mathbb{R}_+^* K \subseteq K,$$

c'est à dire

$$\forall x \in K, \forall \lambda > 0 : \lambda x \in K.$$

Si  $K \cap (-K) = \{0\}$ ,  $K$  est dit cône pointé, avec  $-K = \{-x : x \in K\}$ . Si  $K$  est convexe, on l'appelle cône convexe.

**Exemple 1**  $\mathbf{S}_+^n$  est un cône convexe pointé de  $\mathbf{S}^n$ .

**Définition 1.2.5** Soit  $C$  un ensemble convexe, on dit que  $f : C \subset \mathbb{R}^n \rightarrow \mathbb{R}$  est convexe sur  $C$  si l'une des deux inégalités équivalentes suivantes est vérifiée :

1.  $f[(1 - \lambda)x + \lambda y] \leq (1 - \lambda)f(x) + \lambda f(y), \forall \lambda \in [0, 1], \forall x, y \in C.$
2.  $f(\sum_{i=1}^m \lambda_i x_i) \leq \sum_{i=1}^m \lambda_i f(x_i), \forall m \in \mathbb{N}^*, \forall \lambda_i > 0 / \sum_{i=1}^m \lambda_i = 1, \forall x_i \in C.$

Si la première inégalité est stricte pour  $x, y \in C, x \neq y$  et  $\lambda \in ]0, 1[$ ,  $f$  est dite strictement convexe sur  $C$  c-à-d

$$f[(1 - \lambda)x + \lambda y] < (1 - \lambda)f(x) + \lambda f(y).$$

La deuxième inégalité est appelée inégalité de Jensen.

**Définition 1.2.6** Une fonction  $f$  est dite concave si et seulement si  $(-f)$  est convexe.

**Théorème 1.2.1** Soit  $f \in \mathcal{C}^2(C)$ ,  $C$  convexe ouvert de  $\mathbb{R}^n$  (ou  $C = \mathbb{R}^n$ ). Alors les conditions suivantes sont équivalentes :

1.  $f$  est convexe sur  $C$ .
2.  $f(y) - f(x) \geq \langle \nabla f(x), y - x \rangle, \forall x, y \in C.$
3. La matrice Hessienne  $H(x) = \nabla^2 f(x)$  est semi-définie positive dans  $C$ .

## 1.3 Programmation mathématique

### 1.3.1 Définitions

- Problème d'optimisation



Sous sa forme générale, un problème d'optimisation s'écrit comme suit :

$$\begin{cases} \min_x f(x) \\ x \in C, \end{cases}$$

où  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  continue,  $\emptyset \neq C \subseteq \mathbb{R}^n$  est l'ensemble des contraintes.

Si  $C = \mathbb{R}^n$ , ce problème est appelé problème d'optimisation sans contraintes.

• **Programme mathématique**

Un programme mathématique est un problème d'optimisation qui consiste à trouver une solution du problème qui maximise ou minimise une fonction donnée sous un ensemble des contraintes.

En général, un programme mathématique est défini par :

$$(PM) \quad \begin{cases} \min_x f(x) \\ x \in D, \end{cases}$$

où  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  une fonction continue est appelée fonction objective,  $D \subseteq \mathbb{R}^n$  est l'ensemble des contraintes. Souvent  $D$  est présenté comme suit :

$$D = \{x \in \mathbb{R}^n / g_i(x) \leq 0, i = 1, \dots, m, h_j(x) = 0, j = 1, \dots, p\},$$

avec  $g_i, h_j$  sont des fonctions de  $\mathbb{R}^n \rightarrow \mathbb{R}$ .

• **Contrainte saturée (active) :** Une contrainte d'inégalité  $g_i(x) \leq 0, \forall i$ , est dite saturée (ou active) en  $\bar{x} \in D$  si  $g_i(\bar{x}) = 0$ , pour tout  $i = 1, \dots, m$ .

Une contrainte d'égalité  $h_j(x) = 0$  est par définition saturée en tout point  $x$  de  $D$ .

• **Solution réalisable (faisable) :** Un point  $x^0$  est dit solution réalisable de  $(PM)$  s'il vérifié les contraintes (c-à-d  $x^0 \in D$ ).

• **Solution optimale globale :** Tout point  $x^* \in D$  satisfaisant

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

est appelé solution optimale globale.

L'ensemble des solutions optimales globales de  $(PM)$  est noté par :

$$\arg \min_{x \in D} f(x).$$

- **Solution optimale locale** : Tout point  $x^* \in D$  satisfaisant

$$f(x^*) \leq f(x), \forall x \in D \cap \vartheta,$$

où  $\vartheta$  est un voisinage de  $x^*$ , est appelé solution optimale locale.

L'ensemble des solutions optimales locales de  $(PM)$  est noté par :

$$\text{loc min}_{x \in D} f(x).$$

On a toujours :

$$\arg \min_{x \in D} f(x) \subseteq \text{loc min}_{x \in D} f(x).$$

### 1.3.2 Classifications

On classifie le problème  $(PM)$  à partir de deux propriétés principales à savoir la convexité et la différentiabilité de la fonction objective et les contraintes.

- $(PM)$  est différentiable si les fonctions  $f, g_i, h_j$  sont différentiables.
- $(PM)$  est convexe si  $f, g_i$  sont convexes et  $h_j$  sont affines.

On note que si  $(PM)$  est convexe, alors tout optimum local est un optimum global.

### 1.3.3 Qualification des contraintes

La Qualification des contraintes est satisfaite pour tout  $\bar{x} \in D$  dans l'un des trois cas suivants :

- Les contraintes sont affines ( $D$  est un polyèdre convexe).
- La condition de Slater : Si  $D$  est convexe (c-à-d  $g_i$  sont convexes et  $h_j$  sont affines) et  $\text{int}(D) \neq \emptyset$  c-à-d

$$\exists x^0 : g_i(x^0) < 0 \text{ et } h_j(x^0) = 0, \forall i, j.$$

- La condition de Magazarian-Fromowitz : Si les gradients des contraintes saturées en  $\bar{x} \in D$  sont linéairements indépendants.

### 1.3.4 Principeaux résultats d'existence et d'unicité

**Théorème 1.3.1 (Weirstrass)** Si  $f$  est une fonction continue sur  $D \subseteq \mathbb{R}^n$ ,  $D$  est compact (fermé et borné), alors  $(PM)$  admet au moins une solution optimale  $x^* \in D$ .

**Corollaire 1.3.1** Si  $D \subseteq \mathbb{R}^n$  est non vide et fermé et si  $f$  est continue et coercive sur  $D$  (c-à-d  $f(x) \rightarrow +\infty$  lorsque  $\|x\| \rightarrow +\infty$ ), alors  $(PM)$  admet au moins une solution optimale.

Si  $f$  est strictement convexe et l'ensemble  $D$  est convexe, alors si (PM) admet une solution optimale, la solution est unique.

**Remarque 1.3.1** La stricte convexité assure l'unicité de la solution et non l'existence de ce dernier.

### 1.3.5 Conditions d'optimalité

Le théorème suivant dit de Karush-Kuhn-Tucker-Lagrange donne une condition nécessaire d'optimalité (condition du premier ordre).

**Théorème 1.3.2** Supposons que les fonctions  $f, g_i, h_j$  sont  $\mathcal{C}^1$  dans un voisinage de  $\bar{x} \in D$  et que les contraintes vérifient une des trois conditions de qualification ci-dessus. Si  $f$  a un minimum local en  $\bar{x}$  sur  $D$  alors il existe  $\lambda \in \mathbb{R}_+^m$  et  $\mu \in \mathbb{R}^p$  tels que :

$$(K.K.T) \quad \begin{cases} \nabla f(\bar{x}) + \sum_{i=1}^m \lambda_i \nabla g_i(\bar{x}) + \sum_{j=1}^p \mu_j \nabla h_j(\bar{x}) = 0, \\ \lambda_i g_i(\bar{x}) = 0, \forall i = 1, \dots, m \\ h_j(\bar{x}) = 0, \forall j = 1, \dots, p. \end{cases}$$

Les quantités  $\lambda_i$  et  $\mu_j$  sont appelées multiplicateurs de Karush-Kuhn-Tucker-Lagrange.

**Remarque 1.3.2** - Si les contraintes de (PM) ne sont pas qualifiées en  $\bar{x}$ , les conditions de (K.K.T) ne s'appliquent pas.

- Si (PM) est convexe, les conditions de (K.K.T) sont à la fois nécessaires et suffisantes pour que  $\bar{x}$  soit un minimum global.

### 1.3.6 Dualité d'un programme mathématique

Considérons le problème (PM), on associe à ce problème la fonction Lagrangienne suivante :

$$L(x, \lambda, \gamma) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{j=1}^p \gamma_j h_j(x), \quad x \in \mathbb{R}^n, \quad \lambda_i \geq 0, \quad \gamma_j \in \mathbb{R}.$$

Le dual de (PM) au sens de Lagrange est donné par :

$$(DPM) \quad \left\{ \begin{array}{l} \max_{\lambda \geq 0, \gamma \in \mathbb{R}^p} \min_{x \in \mathbb{R}^n} L(x, \lambda, \gamma) = \max_{\lambda \geq 0, \gamma \in \mathbb{R}^p} H(\lambda, \gamma), \end{array} \right.$$

sachant que  $H$  désigne la fonction duale de  $f$  et vérifie :  $H(\lambda, \gamma) = \min_{x \in \mathbb{R}^n} L(x, \lambda, \gamma)$ ,  $\forall \lambda \in \mathbb{R}_+^m$ ,  $\gamma \in \mathbb{R}^p$ .

### Principaux résultats de dualité

- On a toujours :  $\forall x \in \mathbb{R}^n$ ,  $\forall \lambda \in \mathbb{R}_+^m$ ,  $\gamma \in \mathbb{R}^p$  :  $H(\lambda, \gamma) \leq f(x)$ .
- La fonction  $H(\lambda, \gamma)$  est concave sur  $\mathbb{R}_+^m \times \mathbb{R}^p$ .
- La valeur  $s = f(x) - H(\lambda, \gamma) \geq 0$  s'appelle le saut de dualité.
- S'il existe  $x^* \in \mathbb{R}^n$  et  $\lambda^* \geq 0$ ,  $\gamma^* \in \mathbb{R}^p$  sachant que  $s = 0$ , alors  $x^* \in \arg \min_{\mathbb{R}^n} f(x)$  et  $(\lambda^*, \gamma^*) \in \arg \max_{\mathbb{R}_+^m \times \mathbb{R}^p} H(\lambda, \gamma)$  c'est à dire  $f(x^*) = H(\lambda^*, \gamma^*)$ , et dans ce cas on dit que  $(x^*, \lambda^*, \gamma^*)$  est un point stationnaire pour  $L(x, \lambda, \gamma)$ .
- $\inf_D f(x) = -\infty \Rightarrow H(\lambda, \gamma) = -\infty$ ,  $\forall \lambda \in \mathbb{R}_+^m$ ,  $\gamma \in \mathbb{R}^p$ .
- $\sup_{\lambda \geq 0, \gamma \in \mathbb{R}^p} H(\lambda, \gamma) = +\infty \Rightarrow D = \emptyset$ .
- Si  $(PM)$  est convexe et la condition de Slater est satisfaite, alors  $L(x, \lambda, \gamma)$  admet au moins un point stationnaire et on a :
  - $x^* \in \arg \min_{\mathbb{R}^n} f(x) \iff \exists (\lambda^*, \gamma^*) \in \mathbb{R}_+^m \times \mathbb{R}^p$  tel que  $(x^*, \lambda^*, \gamma^*)$  est un point stationnaire pour  $L(x, \lambda, \gamma)$ .
  - Si  $(PM)$  est différentiable,  $L(\cdot, \lambda, \gamma)$  est convexe par rapport à  $x$  et pour tout  $\lambda \in \mathbb{R}_+^m$ ,  $\gamma \in \mathbb{R}^p$ , alors  $(x^*, \lambda^*, \gamma^*)$  est un point stationnaire pour  $L(x, \lambda, \gamma) \iff$  les conditions de **(K.K.T)** sont satisfaites en  $(x^*, \lambda^*, \gamma^*)$ .

**Remarque 1.3.3** - La fonction Lagrangienne n'est pas unique et par conséquent le problème dual (DPM) n'est pas unique.

- Si le problème  $\min_{x \in \mathbb{R}^n} L(x, \lambda^*, \gamma^*)$  n'a pas de solution, alors le dual (DPM) ne sera pas définie et dans ce cas on utilise le dual au sens de Wolfe :

$$(DPM)_W \quad \begin{cases} \max_{\lambda \geq 0, \gamma \in \mathbb{R}^p} L(x, \lambda, \gamma) \\ \nabla_x L(x, \lambda, \gamma) = 0. \end{cases}$$

# Chapitre 2

## Programmation quadratique convexe semi-définie et méthodes de points intérieurs

Dans ce chapitre, on s'intéresse à l'étude des problèmes de minimisation de fonction quadratique sous l'ensemble des contraintes linéaires dont l'inconnu  $X$  est une matrice symétrique semi-définie positive ( $X \in \mathbb{S}_+^n$ ). Puis, on développe un algorithme de TC de type primal-dual pour résoudre un problème CQSDP et on termine par l'étude de la convergence de cet algorithme et l'analyse de la complexité.

### 2.1 Position du problème

#### 2.1.1 Problème primal

Soient deux entiers naturels  $m, n$  avec  $m \leq n$ , un vecteur  $b \in \mathbb{R}^m$ , des matrices  $C, A_i \in \mathbb{S}^n$ ,  $i = 1, \dots, m$  et  $\mathcal{Q} : \mathbb{S}^n \rightarrow \mathbb{S}^n$  un opérateur linéaire auto-adjoint c'est à dire :

$$\forall A, B \in \mathbb{S}^n : A \bullet \mathcal{Q}(B) = \mathcal{Q}(A) \bullet B.$$

Un programme quadratique semi-défini sous forme primal est un problème d'optimisation donné par :

$$(P) \quad \begin{array}{ll} \min_X & C \bullet X + \frac{1}{2} X \bullet \mathcal{Q}(X) \\ \text{sujet à :} & A_i \bullet X = b_i, \quad i = 1, \dots, m, \quad X \in \mathbb{S}_+^n. \end{array}$$

On note par la suite

$$\begin{aligned}\mathcal{F}_P &= \{X \in \mathbb{S}_+^n : A_i \bullet X = b_i, i = 1, \dots, m\}, \\ \mathcal{F}_P^+ &= \{X \in \mathbb{S}_{++}^n : A_i \bullet X = b_i, i = 1, \dots, m\},\end{aligned}$$

l'ensemble des solutions réalisables (et strictement réalisables, respectivement) pour  $(P)$ . La valeur optimale primale du problème  $(P)$  est définie par :

$$\text{val}(P) = \inf_X \left\{ C \bullet X + \frac{1}{2} X \bullet \mathcal{Q}(X) : A_i \bullet X = b_i, i = 1, \dots, m; X \in \mathbb{S}_+^n \right\}.$$

On dit que  $X^*$  est une solution optimale primale de  $(P)$  si :

$$X^* \in \mathcal{F}_P \text{ et } \text{val}(P) = C \bullet X^* + \frac{1}{2} X^* \bullet \mathcal{Q}(X^*).$$

## 2.1.2 Problème dual

Pour obtenir le problème dual de  $(P)$ , on considère la fonction Lagrangienne  $L : \mathbb{S}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  définie par :

$$L(X, y) = C \bullet X + \frac{1}{2} X \bullet \mathcal{Q}(X) + \sum_{i=1}^m (b_i - A_i \bullet X) y_i,$$

le dual de  $(P)$  est donné par :

$$\left\{ \max_{y \in \mathbb{R}^m} \min_{X \in \mathbb{S}_+^n} L(X, y) = \max_{y \in \mathbb{R}^m} H(y), \right.$$

où

$$\begin{aligned}H(y) &= \min_{X \in \mathbb{S}_+^n} L(X, y), y \in \mathbb{R}^m, \\ &= \min_{X \in \mathbb{S}_+^n} \left\{ C \bullet X + \frac{1}{2} X \bullet \mathcal{Q}(X) + \sum_{i=1}^m (b_i - A_i \bullet X) y_i \right\}, y \in \mathbb{R}^m, \\ &= \min_{X \in \mathbb{S}_+^n} \left\{ \left( C - \sum_{i=1}^m y_i A_i + \mathcal{Q}(X) \right) \bullet X + b^\top y - \frac{1}{2} X \bullet \mathcal{Q}(X) \right\}, y \in \mathbb{R}^m.\end{aligned}$$

Finalement, on obtient :

$$H(y) = \begin{cases} b^\top y - \frac{1}{2}X \bullet \mathcal{Q}(X) & \text{si } \left( C - \sum_{i=1}^m y_i A_i + \mathcal{Q}(X) \right) \succeq 0 \\ -\infty & \text{sinon} \end{cases} .$$

Donc le dual est un programme quadratique semi-défini donné par :

$$(D) \begin{cases} \max_{(X,y,Z)} b^\top y - \frac{1}{2}X \bullet \mathcal{Q}(X) \\ \sum_{i=1}^m y_i A_i - \mathcal{Q}(X) + Z = C \\ X, Z \in \mathbb{S}_+^n \text{ et } y \in \mathbb{R}^m \end{cases} ,$$

où l'inconnu est  $(X, y, Z) \in \mathbb{S}_+^n \times \mathbb{R}^m \times \mathbb{S}_+^n$ .

On note par la suite

$$\begin{aligned} \mathcal{F}_D &= \{X, Z \in \mathbb{S}_+^n \text{ et } y \in \mathbb{R}^m : \sum_{i=1}^m y_i A_i - \mathcal{Q}(X) + Z = C\}, \\ \mathcal{F}_D^+ &= \{X, Z \in \mathbb{S}_{++}^n \text{ et } y \in \mathbb{R}^m : \sum_{i=1}^m y_i A_i - \mathcal{Q}(X) + Z = C\}, \end{aligned}$$

l'ensemble des solutions réalisables (et strictement réalisables, respectivement) pour  $(D)$ . La valeur optimale duale du problème  $(D)$  est définie par :

$$val(D) = \sup_{(X,y,Z)} \left\{ b^\top y - \frac{1}{2}X \bullet \mathcal{Q}(X) : \sum_{i=1}^m y_i A_i + Z - \mathcal{Q}(X) = C; y \in \mathbb{R}^m \text{ et } X, Z \in \mathbb{S}_+^n \right\} .$$

On dit que  $(X^*, y^*, Z^*)$  est une solution optimale duale de  $(D)$  si :

$$(X^*, y^*, Z^*) \in \mathcal{F}_D \text{ et } val(D) = b^\top y^* - \frac{1}{2}X^* \bullet \mathcal{Q}(X^*) .$$

## 2.2 L'importance de la programmation quadratique semi-définie

La programmation quadratique semi-définie (CQSDP) est l'un des problèmes d'optimisation les plus étudiés ces dernières années.

L'intérêt pour ce type de problème est justifié par les facteurs suivants :

- Le problème CQSDP généralise plusieurs problèmes de la programmation mathé-

matique dans le sens où les variables sont des matrices symétriques semi-définies positives au lieu d'être des vecteurs. Parmi ces problèmes mathématiques on cite : la programmation linéaire semi-définie (SDP), la programmation linéaire (LP) et la programmation quadratique (QP)...

- L'existence de plusieurs applications du problème CQSDP en différents domaines comme : l'architecture et la génie électrique...

On note que si  $\mathcal{Q}(X) = 0_{n \times n}$ , CQSDP devient un problème de programmation semi-définie linéaire (SDP).

### 2.2.1 Modélisation de quelques problèmes d'optimisations convexe en un problème CQSDP

Plusieurs problèmes d'optimisation convexe peuvent être transformé au problème CQSDP. En particulier, les problèmes "The nearest correlation matrix ( $\mathcal{NCM}$ ) problem" et "The nearest Euclidean distance matrix ( $\mathcal{EDM}$ ) problem" sont deux exemples de CQSDP (pour plus de détails voir [20], [35], [36]).

#### The nearest correlation matrix ( $\mathcal{NCM}$ ) problem

Le premier exemple de la programmation quadratique semi-définie convexe est le problème ( $\mathcal{NCM}$ ) qui est défini par :

$$(\mathcal{NCM}) \quad \min_X \left\{ \frac{1}{2} \|\mathcal{L}(X - K)\|_F^2 : \text{diag}(X) = e, X \in \mathbb{S}_+^n \right\},$$

où  $K \in \mathbb{S}^n$  une matrice donnée et  $\mathcal{L}$  un opérateur linéaire auto-adjoint ( $\mathcal{L} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ ). D'une part, le problème ( $\mathcal{NCM}$ ) est un cas particulier d'un problème des moindres carrés semi-défini (SDLS) qui est donné par :

$$\min_X \left\{ \left\| \mathcal{L}(X) - \hat{K} \right\|_F^2 : \text{diag}(X) = b, X \in \mathbb{S}_+^n \right\},$$

où  $\mathcal{L} : \mathbb{S}^n \rightarrow \mathbb{S}^p$  un opérateur linéaire auto-adjoint et  $\hat{K} \in \mathbb{S}^p$  une matrice donnée, on prend souvent  $\hat{K} = \mathcal{L}(K)$  avec  $K \in \mathbb{S}^n$  une matrice donnée. Afin d'obtenir le problème ( $\mathcal{NCM}$ ), on prend  $p = n$ ,  $\mathcal{L}(X) = U^{\frac{1}{2}} X U^{\frac{1}{2}}$  avec  $U \in \mathbb{S}_{++}^n$  et  $b = e$ .

D'autre part, le problème CQSDP résulte du problème ( $\mathcal{NCM}$ ) avec  $C = -\mathcal{L}^2(K) = -\mathcal{L}(\mathcal{L}(K))$ , et par conséquent  $\mathcal{Q}(X) = \mathcal{L}^2(X)$  et pour  $i = 1, \dots, n$ , la matrice  $A_i$  est



définie par :

$$A_i[j, k] = \begin{cases} 1 & \text{si } i = j = k, \\ 0 & \text{ailleurs,} \end{cases}$$

et  $b = e$ . Donc, selon les choix de l'opérateur  $\mathcal{L}$  on obtient :

1. Si  $\mathcal{L}(X) = U^{\frac{1}{2}}XU^{\frac{1}{2}}$ , de même  $\mathcal{Q}(X) = UXU$  où  $U \in \mathbb{S}_{++}^n$  une matrice donnée. L'un des premiers travaux (avec cet choix) est réalisé par Higham [19] qui modélise un problème de finance sous forme du problème ( $\mathcal{NCM}$ ).
2. Si  $\mathcal{L}(X) = X$ , de même  $\mathcal{Q}(X) = X$ . Il existe plusieurs travaux qui sont basés sur ce choix spéciale qui est un cas particulier du choix précédent c-à-d  $\mathcal{L}(X) = U^{\frac{1}{2}}XU^{\frac{1}{2}}$  et  $\mathcal{Q}(X) = UXU$  avec  $U = I \in \mathbb{S}_{++}^n$ .
3. Si  $\mathcal{L}(X) = \Sigma \circ X$  pour une matrice donnée  $\Sigma \in \mathbb{S}^n \cap \mathbb{R}_+^{n \times n}$ , de même  $\mathcal{Q}(X) = U \circ X$  avec  $U = \Sigma \circ \Sigma$  où  $\circ$  désigne le produit de Hadamard qui est défini par :  $(\Sigma \circ X)_{ij} = \Sigma_{ij}X_{ij}$ , avec  $\Sigma$  et  $X$  deux matrices de même taille. On note que avec cet choix Higham [20] a proposé "Modified alternating projection solution method" pour traiter le problème ( $\mathcal{NCM}$ ).

### The nearest Euclidean distance matrix ( $\mathcal{EDM}$ ) problem

Le deuxième exemple de la programmation quadratique semi-définie convexe est le problème ( $\mathcal{EDM}$ ). Donc, on cherche à résoudre le problème suivant :

$$(\mathcal{EDM}) \quad \min_X \left\{ \frac{1}{2} \|H \circ (B - \mathcal{L}_E(X))\|_F^2 : (\mathcal{L}_E(X))_{ij} = B_{ij}, \forall (ij) \in w, X \in \mathbb{S}_+^n \right\},$$

où  $B \in \mathbb{S}^n$  une matrice donnée,  $H$  la matrice du poids et  $w$  un ensemble donné (désigne l'ensemble des indices). Le problème CQSDP sous forme primal standard résulte du problème ( $\mathcal{EDM}$ ) est obtenu, si on prend :

$$\mathcal{L}_E(X) = \text{diag}(X_V)e^\top + e\text{diag}(X_V)^\top - 2X_V,$$

où  $X_V = VXV^\top$  et  $V \in \mathbb{R}^{n \times (n-1)}$  une matrice donnée dont les colonnes de  $V$  sont orthonormées et  $V^\top e = 0$ . De plus, l'opérateur  $\mathcal{Q}$  associé au problème ( $\mathcal{EDM}$ ) est semi-défini positive.

## 2.3 Dualité en CQSDP

### 2.3.1 Dualité faible

**Définition 2.3.1** (*Saut de dualité*) Soient  $X \in \mathcal{F}_P$  et  $(X, y, Z) \in \mathcal{F}_D$ , alors la différence

$$\begin{aligned} \text{val}(P) - \text{val}(D) &= C \bullet X + \frac{1}{2}X \bullet \mathcal{Q}(X) - b^\top y + \frac{1}{2}X \bullet \mathcal{Q}(X) \\ &= C \bullet X + X \bullet \mathcal{Q}(X) - b^\top y, \\ &= X \bullet Z. \end{aligned}$$

est appelée le saut de dualité des problèmes  $(P)$  et  $(D)$ .

Maintenant, on a le résultat suivant :

**Théorème 2.3.1** Pour toute  $X \in \mathcal{F}_P$  et  $(X, y, Z) \in \mathcal{F}_D$  on a :

$$Z \bullet X = C \bullet X - b^\top y + X \bullet \mathcal{Q}(X) \geq 0.$$

**Preuve :** Soient  $X \in \mathcal{F}_P$  et  $(X, y, Z) \in \mathcal{F}_D$ , alors :

$$\begin{aligned} C \bullet X + X \bullet \mathcal{Q}(X) - b^\top y &= (C + \mathcal{Q}(X)) \bullet X - \sum_{i=1}^m b_i y_i, \\ &= (C + \mathcal{Q}(X)) \bullet X - \sum_{i=1}^m y_i \langle A_i, X \rangle, \\ &= \left( C + \mathcal{Q}(X) - \sum_{i=1}^m y_i A_i \right) \bullet X, \\ &= Z \bullet X \geq 0, \text{ car } X \text{ et } Z \in \mathbb{S}_+^n, \end{aligned}$$

la dernière inégalité découle de la Proposition [1.1.2]. Ce qui achève la preuve. ■

Pour résoudre les deux problèmes  $(P)$  et  $(D)$ , on suppose qu'ils vérifient les hypothèses suivantes :

**Hypothèse 1.** Condition d'indépendance : Les matrices  $[A_1, A_2, \dots, A_m]$  sont linéairement indépendantes.

**Hypothèse 2.** Condition de points intérieurs : **(CPI)** On suppose qu'il existe un point

primal-dual strictement réalisable  $(X^0, y^0, Z^0)$ , en autre terme ce point vérifie

$$\begin{cases} A_i \bullet X^0 & = b_i, i = 1, \dots, m \\ \sum_{i=1}^m y_i^0 A_i - \mathcal{Q}(X^0) + Z^0 & = C, \\ X^0, Z^0 \in \mathbb{S}_{++}^n \text{ et } y^0 \in \mathbb{R}^m. \end{cases}$$

**Hypothèse 3.** La monotonie. L'opérateur linéaire  $\mathcal{Q}$  est monotone c-à-d :

$$X \bullet \mathcal{Q}(X) \geq 0, \forall X \in \mathbb{S}^n.$$

### 2.3.2 Conditions d'optimalité nécessaires et suffisantes de la solution optimale

Pour étudier la convexité et la différentiabilité de CQSDP, on a besoin du lemme suivant :

**Lemme 2.3.1** *Soit  $\mathcal{Q}$  un opérateur linéaire auto-adjoint et monotone, alors :  $\nabla \mathcal{Q}(X)$  est aussi un opérateur monotone  $\forall X \in \mathbb{S}^n$ .*

**Preuve :** Soient  $X \in \mathbb{S}^n$  et  $\mathcal{Q}$  un opérateur linéaire auto-adjoint et monotone, alors  $\mathcal{Q}$  est différentiable et sa dérivée est lui même c-à-d

$$\nabla \mathcal{Q}(X)Y = \mathcal{Q}(Y), \forall Y \in \mathbb{S}^n,$$

car : si on prend  $Y \in \mathbb{S}^n$  telle que  $|\lambda_i(Y)|$  est suffisamment petite (voir [37, Lemme 2.5]) alors :

$$\mathcal{Q}(X + Y) \simeq \mathcal{Q}(X) + \nabla \mathcal{Q}(X)(Y), \forall X, Y \in \mathbb{S}^n,$$

qui est équivalent à

$$\mathcal{Q}(X + Y) - \mathcal{Q}(X) = \mathcal{Q}(Y) \simeq \nabla \mathcal{Q}(X)(Y), \forall X, Y \in \mathbb{S}^n,$$

donc

$$\nabla \mathcal{Q}(X)Y \bullet Y = \mathcal{Q}(Y) \bullet Y \geq 0, \forall Y \in \mathbb{S}^n \text{ car } \mathcal{Q} \text{ est monotone,}$$

et par conséquent  $\nabla \mathcal{Q}(X)$  est un opérateur monotone  $\forall X \in \mathbb{S}^n$ . Ce qui achève la preuve. ■

On rappelle que le problème ( $P$ ) consiste à minimiser la fonction :

$$f(X) = C \bullet X + \frac{1}{2} X \bullet \mathcal{Q}(X),$$

sous l'ensemble des contraintes

$$\mathcal{F}_P = \{A_i \bullet X = b_i, i = 1, \dots, m, X \in \mathbb{S}_+^n\},$$

qui s'écrit comme l'intersection d'un cône convexe non polyédrique noté par  $(X \in \mathbb{S}_+^n)$  et des contraintes linéaires  $\{A_i \bullet X = b_i, i = 1, \dots, m\}$ . Donc  $\mathcal{F}_P$  est convexe et comme  $\mathcal{F}_P$  est d'intérieur relatif non vide, alors la condition de Slater est satisfaite (c-à-d  $\mathcal{F}_P$  est convexe et  $\mathcal{F}_P^+ \neq \emptyset$ ).

De plus, on peut montrer que la fonction objective  $f(X)$  est deux fois différentiable et convexe c-à-d  $\nabla^2 f(X)H \bullet H = \nabla \mathcal{Q}(X)H \bullet H \geq 0, \forall X, H \in \mathbb{S}^n$ , où la dernière inégalité est déduite du Lemme [2.3.1].

Par conséquent, le problème d'optimisation  $(P)$  est convexe et différentiable, à contraintes qualifiées. Alors les conditions d'optimalités de **(K.K.T)** sont nécessaires et suffisantes et s'écrivent comme suit :

$$\text{(K.K.T)} \quad \begin{cases} A_i \bullet X & = b_i, i = 1, \dots, m, \\ \sum_{i=1}^m y_i A_i + Z - \mathcal{Q}(X) & = C, \\ Z \bullet X & = 0, X, Z \in \mathbb{S}_+^n. \end{cases}$$

### 2.3.3 Complémentarité en CQSDP

On peut exprimer la condition pour  $X^*$  et  $(X^*, y^*, Z^*)$  d'être des solutions optimales de  $(P)$  et  $(D)$ , sous forme d'une condition dite de complémentarité. Cela est donnée dans le cadre du théorème suivant :

**Théorème 2.3.2** [25, Théorème 2.2] *Si les matrices  $[A_1, A_2, \dots, A_m]$  sont linéairement indépendantes, la condition de points intérieurs est satisfaite et l'opérateur  $\mathcal{Q}$  est monotone alors  $X^* \in \mathcal{F}_P$  est une solution pour  $(P)$  si et seulement s'il existe  $y^* \in \mathbb{R}^m$  et  $Z^* \in \mathbb{S}_+^n$  tel que  $(X^*, y^*, Z^*) \in \mathcal{F}_D$  et  $X^*Z^* = 0$ .*

Sous les hypothèses  $H_1, H_2, H_3$ , il est connu que ([16], [25], [35]) trouver une solution optimale de  $(P)$  et  $(D)$  est équivalent à trouver une solution du système non linéaire suivant :

$$(S) \quad \begin{cases} A_i \bullet X & = b_i, i = 1, \dots, m, \\ \sum_{i=1}^m y_i A_i + Z - \mathcal{Q}(X) & = C, \\ XZ & = 0, X, Z \in \mathbb{S}_+^n, y \in \mathbb{R}^m. \end{cases}$$

Où la dernière équation présente la condition de complémentarité et le système  $(S)$  est appelé problème de complémentarité.

**Remarque 2.3.1** *Puisque  $(P)$  est un problème convexe, la solution du problème de complémentarité est un optimum global pour  $(P)$ .*

## 2.4 Méthodes de points intérieurs pour résoudre le problème CQSDP

Dans le cadre de la programmation mathématique avec contraintes, on désigne par les méthodes de points intérieurs, toute procédure de résolution générant une suite de points appartenant à l'intérieur relatif du domaine réalisable et convergeant vers une solution optimale.

De même, on appelle fonction barrière toute fonction  $f$  qui vérifie :

- $f$  est à valeurs finies à l'intérieur relatif du domaine réalisable.
- $f(x) \rightarrow \infty$  quand  $x$  s'approche de la frontière.

Ces méthodes sont réputées pour leur convergence polynomiale, leur rapidité et efficacité et se sont révélées comme de véritables concurrentes des méthodes classiques simpliciales. La littérature sur ces méthodes a connu une grande expansion et s'est enrichie de plusieurs classes de variantes dans le but de réduire la complexité et améliorer la convergence et l'efficacité. vous pouvez consulter les ouvrages ([42], [43]) qui retracent l'évolution des méthodes de points intérieurs.

Les méthodes de points intérieurs se répartissent en trois catégories : les méthodes affines, les méthodes projectives de réduction de potentiel et les méthodes de trajectoire centrale. Dans cette thèse, on s'intéresse seulement aux méthodes de trajectoire centrale pour résoudre le problème CQSDP. Ce qui concerne les autres catégories voir ([12]).

### 2.4.1 Méthode de la trajectoire centrale de type primal-dual pour CQSDP

Afin d'introduire une méthode de points intérieurs au système  $(S)$ , on associe au problème  $(P)$  le problème de minimisation barrière non linéaire suivant :

$$(P_\mu) \begin{cases} \min f_\mu(X), \\ A_i \bullet X = b_i, i = 1, \dots, m, \\ X \in \mathbb{S}_{++}^n, \mu > 0, \end{cases}$$

où

$$f_\mu(X) = C \bullet X + \frac{1}{2} X \bullet Q(X) - \mu \log \det X.$$

Les deux lemmes suivants permettent de calculer le gradient de la fonction barrière logarithmique primale  $f_\mu(X)$ . Pour la démonstration voir l'Annexe.

**Lemme 2.4.1** [16] Soit  $f : \mathbb{S}_{++}^n \rightarrow \mathbb{R}$  une fonction donnée par  $f(X) = \log \det X$ , et

$$\nabla f(X) = \begin{bmatrix} \frac{\partial f(X)}{\partial x_{11}} & \dots & \frac{\partial f(X)}{\partial x_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(X)}{\partial x_{n1}} & \dots & \frac{\partial f(X)}{\partial x_{nn}} \end{bmatrix},$$

alors  $\nabla f(X) = X^{-1}$ .

**Lemme 2.4.2** [16] Soit  $f : \mathbb{S}_{++}^n \rightarrow \mathbb{R}$  une fonction donnée par  $f(X) = \mathbf{Tr}(CX) = C \bullet X$ , où  $C \in \mathbb{S}^n$ , alors  $\nabla f(X) = C$ .

**Lemme 2.4.3** [16] Soient  $f : \mathbb{S}_{++}^n \rightarrow \mathbb{R}$  une fonction donnée par  $f(X) = \log \det X$  et  $\nabla f(X) = X^{-1}$ , alors  $\nabla^2 f(X)$  est un opérateur linéaire qui vérifie

$$\nabla^2 f(X)H = -X^{-1}HX^{-1}, \quad \forall H \in \mathbb{S}^n.$$

Le résultat suivant montre la stricte convexité de la fonction barrière logarithmique primale.

**Proposition 2.4.1** Soient  $\mu > 0$  et  $f_\mu : \mathbb{S}_{++}^n \rightarrow \mathbb{R}$  une fonction définie par :

$$f_\mu(X) = C \bullet X + \frac{1}{2}X \bullet \mathcal{Q}(X) - \mu \log \det X.$$

Alors, la fonction  $f_\mu(X)$  est strictement convexe pour tout  $\mu > 0$  (fixé).

**Preuve :** Soient  $X \in \mathbb{S}_{++}^n$ ,  $H \in \mathbb{S}^n$ . D'après les Lemmes précédents on a :

$$\nabla f_\mu(X) = C + \mathcal{Q}(X) - \mu X^{-1},$$

et

$$\begin{aligned} \nabla^2 f_\mu(X)H &= \nabla \mathcal{Q}(X)H + \mu X^{-1}HX^{-1}, \\ &= \mathcal{Q}(H) + \mu X^{-1}HX^{-1}. \end{aligned}$$

Par conséquent

$$\nabla^2 f_\mu(X)H \bullet H = \mathcal{Q}(H) \bullet H + \mu X^{-1}HX^{-1} \bullet H,$$

puisque  $\mathcal{Q}$  est monotone et  $\mu > 0$ , il suffit de montrer que :

$$X^{-1}HX^{-1} \bullet H > 0, \quad \forall X \in \mathbb{S}_{++}^n \text{ et } H \in \mathbb{S}^n.$$

On a :

$$\begin{aligned} X^{-1}HX^{-1} \bullet H &= \mathbf{Tr} (X^{-1}HX^{-1}H), \\ &= \mathbf{Tr} ((X^{-\frac{1}{2}}HX^{-\frac{1}{2}})^2), \\ &= \left\| X^{-\frac{1}{2}}HX^{-\frac{1}{2}} \right\|_F^2 > 0, \quad \forall H \neq 0 \text{ et } H \in \mathbb{S}^n. \end{aligned}$$

Donc  $f_\mu(X)$  est strictement convexe. Ce qui achève la preuve. ■

Notons que la démonstration de ce Lemme est faite en [26] avec un choix spécial de l'opérateur  $\mathcal{Q}(X)$ .

Enonçons maintenant les propriétés suivantes :

- $-\log \det X$  est appelée la fonction barrière logarithmique.
- Le paramètre  $\mu > 0$  est appelé le paramètre de barrière.
- $(P_\mu)$  est un problème d'optimisation convexe et différentiable.
- Les conditions d'optimalité nécessaires d'ordre 1 pour que  $X(\mu)$  soit une solution pour  $(P_\mu)$ , est l'existence d'un vecteur  $y(\mu)$  tel que :

$$\begin{cases} \nabla f_\mu(X) - \sum_{i=1}^m y_i(\mu)A_i &= 0, \\ A_i \bullet X(\mu) &= b_i, i = 1, \dots, m, \end{cases}$$

avec

$$\nabla f_\mu(X) = C + \mathcal{Q}(X) - \mu X^{-1}.$$

On pose

$$Z(\mu) = \mu X^{-1},$$

alors

$$X(\mu)Z(\mu) = \mu I.$$

Donc les conditions nécessaires et suffisantes de **(K.K.T)** s'écrivent :

$$\begin{cases} \sum_{i=1}^m y_i(\mu)A_i + Z(\mu) - \mathcal{Q}(X(\mu)) &= C, \\ A_i \bullet X(\mu) &= b_i, i = 1, \dots, m, \\ X(\mu)Z(\mu) &= \mu I. \end{cases} \quad (2.1)$$

- L'existence et l'unicité de la solution  $(X(\mu), y(\mu), Z(\mu))$  pour tout  $\mu > 0$  ( $\mu$  fixé) reviennent à l'existence et l'unicité de la solution du problème  $(P_\mu)$  (voir [25]).
- L'ensemble  $\mathcal{C} = \{(X(\mu), y(\mu), Z(\mu)); \mu > 0\}$  est appelé la trajectoire centrale, si  $\mu$  tend vers zéro alors  $(X(\mu), y(\mu), Z(\mu))$  devient une solution du système  $(S)$  c-à-d des solutions optimales de deux problèmes  $(P)$  et  $(D)$ , (voir [16], [25]).
- La courbe  $\mathcal{C}$  est lisse pour tout  $\mu > 0$ .
- Le système d'équations non linéaires (2.1) caractérise la solution optimale du problème  $(P_\mu)$ , donc la résolution de ce problème est équivalente à la résolution de (2.1), et pour résoudre ce dernier, on utilise la méthode de trajectoire centrale de type primal-dual.

### Stratégie de l'algorithme

- On commence par  $(X^0, y^0, Z^0)$  un point strictement réalisable primal-dual et appartenant à un voisinage de trajectoire centrale, avec  $\mu^{(0)}$  connu ( $\mu^{(0)} = \frac{X^0 \bullet Z^0}{n}$ ).
- On construit le nouveau triple

$$\begin{cases} X^{(k+1)} &= X^{(k)} + \Delta X^{(k)}, \\ y^{(k+1)} &= y^{(k)} + \Delta y^{(k)}, \\ Z^{(k+1)} &= Z^{(k)} + \Delta Z^{(k)}, \end{cases}$$

où  $(\Delta X^{(k)}, \Delta y^{(k)}, \Delta Z^{(k)})$  sont les directions de Newton.

- On vérifie que  $X^{(k+1)}, Z^{(k+1)}$  sont strictement réalisables pour tout  $\mu > 0$ , et on distingue deux cas :

Si  $n\mu^{(k)} < \epsilon$  où  $\epsilon$  est la précision donnée, alors  $X^{(k+1)}, y^{(k+1)}$  et  $Z^{(k+1)}$  sont des solutions approchées du système (2.1).

Sinon ( $n\mu^{(k)} \geq \epsilon$ )  $X^{(k+1)}, y^{(k+1)}$  et  $Z^{(k+1)}$  ne sont pas les solutions optimales approximatives de (2.1), dans ce cas, le paramètre  $\mu$  se réduit à  $\mu^+ = (1 - \theta)\mu$  avec  $0 < \theta < 1$ , et restons proche de la trajectoire centrale. Généralement, si  $\theta$  dépend de  $n$ , notamment  $\theta = \Theta(\frac{1}{\sqrt{n}})$ , alors l'algorithme est appelé : à petit pas "short-step", et si  $\theta = O(1)$  est un constant, alors l'algorithme est appelé : à grand pas "long-step".

### 2.4.2 Les directions classiques

Soient  $\mu > 0$  et  $(X, y, Z)$  un point strictement réalisable primal-dual (mais  $XZ \neq \mu I$ ), la direction de Newton  $(\Delta X, \Delta y, \Delta Z)$  en ce point est la solution unique du système



linéaire suivant :

$$\begin{cases} A_i \bullet \Delta X & = 0, i = 1, \dots, m, \\ \sum_{i=1}^m \Delta y_i A_i + \Delta Z - \mathcal{Q}(\Delta X) & = 0, \\ \Delta X Z + X \Delta Z & = \mu I - XZ, \end{cases} \quad (2.2)$$

qui est équivalent à :

$$\begin{cases} A_i \bullet \Delta X & = 0, i = 1, \dots, m, \\ \sum_{i=1}^m \Delta y_i A_i + \Delta Z - \mathcal{Q}(\Delta X) & = 0, \\ \Delta X + X \Delta Z Z^{-1} & = \mu Z^{-1} - X. \end{cases} \quad (2.3)$$

Sous les hypothèses précédentes ce système linéaire admet une solution unique  $(\Delta X, \Delta y, \Delta Z)$ . On note que  $\Delta Z$  est symétrique d'après la deuxième équation du système (2.3) mais  $\Delta X$  n'est pas forcément symétrique. Alors pour traiter ce problème, Y. Zhang [45] a proposé de remplacer la dernière équation en (2.2) par l'équation suivante :

$$H_P(\Delta X Z + X \Delta Z) = \mu I - H_P(XZ),$$

où  $H_P$  est une transformation linéaire donnée par :

$$H_P(M) = \frac{1}{2}[PMP^{-1} + P^{-t}M^tP^t],$$

avec  $P$  une matrice inversible. Tandis que Todd et al [34], ont montré que pour symétriser le système (2.3), on introduit directement une matrice  $P$  inversible comme suit :

$$\begin{cases} A_i \bullet \Delta X & = 0, i = 1, \dots, m, \\ \sum_{i=1}^m \Delta y_i A_i + \Delta Z - \mathcal{Q}(\Delta X) & = 0, \\ \Delta X + P \Delta Z P^\top & = \mu Z^{-1} - X. \end{cases} \quad (2.4)$$

### Les directions de Nesterov Todd

L'un des choix de la matrice  $P$  est :

$$P := X^{\frac{1}{2}}(X^{\frac{1}{2}}Z X^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}}, \text{ (Nesterov Todd scaling matrix),}$$

où  $P \in \mathbb{S}_{++}^n$ .

Maintenant, on définit la matrice symétrique et définie positive  $D = P^{\frac{1}{2}}$ . Le rôle de  $D$

est de mettre à l'échelle les deux matrices  $X$  et  $Z$  au même matrice symétrique et définie positive  $V$  qui vérifie :

$$V = \frac{1}{\sqrt{\mu}}D^{-1}XD^{-1} = \frac{1}{\sqrt{\mu}}DZD = \frac{1}{\sqrt{\mu}}(D^{-1}XZD)^{1/2}. \quad (2.5)$$

De plus, les nouvelles directions (scaling directions)  $D_X$  et  $D_Z$  sont définies par :

$$D_X := \frac{1}{\sqrt{\mu}}D^{-1}\Delta XD^{-1}, \quad D_Z := \frac{1}{\sqrt{\mu}}D\Delta ZD. \quad (2.6)$$

Posons  $(D_y)_i = \Delta y_i$ , alors le système (2.4) devient :

$$\begin{cases} \bar{A}_i \bullet D_X & = 0, i = 1, \dots, m, \\ \sum_{i=1}^m (D_y)_i \bar{A}_i + D_Z - \bar{Q}(D_X) & = 0, \\ D_X + D_Z & = V^{-1} - V, \end{cases} \quad (2.7)$$

où

$$\bar{A}_i = \frac{1}{\sqrt{\mu}}DA_iD, i = 1, \dots, m, \text{ et } \bar{Q}(D_X) = DQ(DD_XD)D. \quad (2.8)$$

On note que ce dernier système est équivalent au système (2.4), les matrices  $[\bar{A}_1, \bar{A}_2, \dots, \bar{A}_m]$  sont aussi linéairement indépendantes et l'opérateur  $\bar{Q}$  est aussi auto-adjoint et monotone. De plus, les directions  $D_X, D_Z$  sont symétriques et vérifiant ([25], [26], [41]) :

$$D_X \bullet D_Z = \frac{1}{\mu}\Delta X \bullet Q(\Delta X) \geq 0, \quad (2.9)$$

car

$$\begin{aligned} D_X \bullet D_Z &= \frac{1}{\sqrt{\mu}}(D^{-1}\Delta XD^{-1}) \bullet \frac{1}{\sqrt{\mu}}(D\Delta ZD), \quad D = D^\top, \\ &= \frac{1}{\mu}(DD^{-1}\Delta XD^{-1}D) \bullet \Delta Z, \\ &= \frac{1}{\mu}\Delta X \bullet \Delta Z, \end{aligned}$$

et on a :

$$\sum_{i=1}^m \Delta y_i A_i + \Delta Z - Q(\Delta X) = 0 \Leftrightarrow \Delta Z = - \sum_{i=1}^m \Delta y_i A_i + Q(\Delta X),$$

alors

$$\begin{aligned}
D_X \bullet D_Z &= \frac{1}{\mu} \langle \Delta X, -\sum_{i=1}^m \Delta y_i A_i + \mathcal{Q}(\Delta X) \rangle, \\
&= \frac{1}{\mu} \left[ -\sum_{i=1}^m \Delta y_i (A_i \bullet \Delta X) + \Delta X \bullet \mathcal{Q}(\Delta X) \right], \text{ où } A_i \bullet \Delta X = 0, \\
&= \frac{1}{\mu} \Delta X \bullet \mathcal{Q}(\Delta X) \geq 0, \text{ car } \mathcal{Q} \text{ est monotone.}
\end{aligned}$$

Notons que les directions cherchées ne sont pas orthogonales contrairement au cas SDP linéaire.

### 2.4.3 La mesure de proximité

Pour que les itérations produites par l'algorithme restent réalisables et proches de la TC, on est obligé d'introduire une mesure de proximité qui est défini par :

$$\delta := \delta(X, Z; \mu) = \frac{1}{2} \|V^{-1} - V\|_F.$$

C'est facile de vérifier que :

$$\delta(X, Z; \mu) = 0 \Leftrightarrow V^{-1} = V \Leftrightarrow XZ = \mu I,$$

qui justifie l'appartenance des points à la trajectoire centrale.

Dans cet algorithme, on utilise une valeur seuil  $\beta$  de la proximité et on suppose que  $(X^0, y^0, Z^0)$  est un point strictement réalisable avec  $\delta(X^0, Z^0; \mu^{(0)}) \leq \beta$  pour certain  $\mu^{(0)}$  connu. On note que la recherche d'une solution optimale primale-duale est équivalente à mettre le saut de dualité  $X \bullet Z$  vers le zéro, cela est exprimé par la mise à jour du paramètre  $\mu$  via :

$$\mu_+ = (1 - \theta)\mu, \quad 0 < \theta < 1.$$

Maintenant, on présente un algorithme primal-dual à petit pas pour tracer approximativement la trajectoire centrale.

### 2.4.4 Algorithme

Dans cette partie, on présente un algorithme de trajectoire centrale de type primal-dual pour CQSDP :

## Algorithme primal-dual à petit pas pour CQSDP

### Début d'algorithme

#### Données :

un paramètre de précision  $\epsilon > 0$ ;

un paramètre de proximité  $0 < \beta < 1$  (défaut  $\beta = \frac{1}{\sqrt{2}}$ );

un paramètre  $0 < \theta < 1$  (défaut  $\theta = \frac{1}{2\sqrt{n}}$ );

un paramètre de barrière  $\mu^{(0)} = \frac{X^0 \bullet Z^0}{n}$ .

**Initialisation :** soit  $(X^0, y^0, Z^0)$  vérifiant la **CPI** tel que

$\delta(X^0, Z^0; \mu^{(0)}) \leq \beta$  et  $k = 0$ ;

**Tant que**  $n\mu^{(k)} \geq \epsilon$  faire :

- $\mu^{(k+1)} = (1 - \theta)\mu^{(k)}$ .
- calculer  $(\Delta X^{(k)}, \Delta y^{(k)}, \Delta Z^{(k)})$  via le système (2.7);
- mise à jour  $X^{(k+1)} = X^{(k)} + \Delta X^{(k)}$ ,  $y^{(k+1)} = y^{(k)} + \Delta y^{(k)}$ ,  
 $Z^{(k+1)} = Z^{(k)} + \Delta Z^{(k)}$  et  $k = k + 1$ ;

**Fin tant que.**

**Fin algorithme.**

Algorithme 2.4.4

### 2.4.5 La convergence de l'algorithme et l'analyse de la complexité

On considère la matrice symétrique suivante :

$$D_{XZ} = \frac{1}{2} [D_X D_Z + D_Z D_X],$$

qui sera utile par la suite et  $0 \leq \alpha \leq 1$ .

Afin d'étudier la convergence de l'Algorithme 2.4.4 et d'analyser sa complexité, on a besoin de quelques résultats techniques qui sont utilisés pour montrer la stricte faisabilité du pas de Newton complet.

**Lemme 2.4.4** [16, Lemme 3.3.1] Soient  $X(\alpha) := X + \alpha\Delta X$ ,  $Z(\alpha) := Z + \alpha\Delta Z$  telles que  $X \succ 0$  et  $Z \succ 0$ . Si

$$\det(X(\alpha)Z(\alpha)) > 0, \forall 0 \leq \alpha \leq \bar{\alpha}.$$

Alors :  $X(\alpha) \succ 0$  et  $Z(\alpha) \succ 0, \forall 0 \leq \alpha \leq \bar{\alpha}$ .

**Lemme 2.4.5** [16, Lemme 3.3.3] Soient  $N \in \mathbf{S}^n$  et  $M \in \mathfrak{R}^{n \times n}$  une matrice anti-symétrique. Si  $N \succ 0$  alors  $\det(N + M) > 0$ . De plus, si  $\lambda_i(N + M) \in \mathfrak{R}$  et  $i = 1, \dots, n$ , alors

$$0 < \lambda_{\min}(N) \leq \lambda_{\min}(N + M) \leq \lambda_{\max}(N + M) \leq \lambda_{\max}(N).$$

**Lemme 2.4.6** Soit  $(D_X, D_Z)$  une solution du système (2.7) et  $\mu > 0$ . Si  $\delta := \delta(X, Z; \mu)$  alors

$$0 \leq D_X \bullet D_Z \leq 2\delta^2. \quad (2.10)$$

De plus, la norme spectrale de  $D_{XZ}$  vérifie :

$$\|D_{XZ}\|_2 \leq \frac{1}{4} \|D_X + D_Z\|_F^2 = \delta^2. \quad (2.11)$$

Pour montrer ce lemme, on présente la remarque suivante :

**Remarque 2.4.1** En CQSDP, on sait que les directions  $D_X$  et  $D_Z$  ne sont pas orthogonales c-à-d  $D_X \bullet D_Z \geq 0$  ce qui implique :

$$\|D_X + D_Z\|_F \geq \|D_X - D_Z\|_F.$$

**Preuve :** (du lemme [2.4.6]) D'une part on a :

$$D_X \bullet D_Z = \frac{1}{\mu} \Delta X \bullet Q(\Delta X) \geq 0. \quad (2.12)$$

D'autre part on a :

$$\begin{aligned} 2\delta^2 &= \frac{1}{2} \|D_X + D_Z\|_F^2, \\ &= \frac{1}{2} \mathbf{Tr}((D_X + D_Z)^2), \\ &= \frac{1}{2} [\|D_X\|_F^2 + \|D_Z\|_F^2 + 2D_X \bullet D_Z], \\ &= \frac{1}{2} [\|D_X\|_F^2 + \|D_Z\|_F^2] + D_X \bullet D_Z \geq D_X \bullet D_Z, \end{aligned} \quad (2.13)$$

de (2.12) et (2.13) on obtient :

$$0 \leq D_X \bullet D_Z \leq 2\delta^2.$$

Maintenant, on montre que

$$\|D_{XZ}\|_2 \leq \frac{1}{4} \|D_X + D_Z\|_F^2 = \delta^2.$$

Par définition on a :

$$\delta = \frac{1}{2} \|D_X + D_Z\|_F,$$

alors

$$\frac{1}{4} \|D_X + D_Z\|_F^2 = \delta^2, \quad (2.14)$$

il nous reste à montrer que :

$$\|D_{XZ}\|_2 \leq \frac{1}{4} \|D_X + D_Z\|_F^2.$$

On a

$$D_X D_Z + D_Z D_X = \frac{1}{2} [(D_X + D_Z)^2 - (D_X - D_Z)^2],$$

ce qui implique que

$$-\frac{1}{4} (D_X - D_Z)^2 \preceq D_{XZ} \preceq \frac{1}{4} (D_X + D_Z)^2,$$

et que :

$$-\frac{1}{4} \|D_X - D_Z\|_F^2 I \preceq D_{XZ} \preceq \frac{1}{4} \|D_X + D_Z\|_F^2 I.$$

D'après la Remarque [2.4.1], on a :

$$\|D_X + D_Z\|_F \geq \|D_X - D_Z\|_F,$$

qui est équivalent à :

$$-\frac{1}{4} \|D_X + D_Z\|_F^2 \leq -\frac{1}{4} \|D_X - D_Z\|_F^2,$$

donc

$$-\frac{1}{4} \|D_X + D_Z\|_F^2 I \preceq D_{XZ} \preceq \frac{1}{4} \|D_X + D_Z\|_F^2 I,$$

en utilisant les valeurs propres, on obtient :

$$-\frac{1}{4} \|D_X + D_Z\|_F^2 \leq \lambda_i(D_{XZ}) \leq \frac{1}{4} \|D_X + D_Z\|_F^2, \quad \forall i = 1, \dots, n.$$

Alors

$$|\lambda_i(D_{XZ})| \leq \frac{1}{4} \|D_X + D_Z\|_F^2, \quad \forall i = 1, \dots, n.$$

Et par conséquent

$$\max_{i=1}^n |\lambda_i(D_{XZ})| = \rho(D_{XZ}) \leq \frac{1}{4} \|D_X + D_Z\|_F^2,$$

puisque  $D_{XZ}$  est symétrique, alors :

$$\rho(D_{XZ}) = \|D_{XZ}\|_2 \leq \frac{1}{4} \|D_X + D_Z\|_F^2 \quad (2.15)$$

De (2.14) et (2.15) on déduit que :

$$\|D_{XZ}\|_2 \leq \frac{1}{4} \|D_X + D_Z\|_F^2 = \delta^2.$$

Ce qui achève la preuve. ■

Afin de simplifier les calculs, on pose

$$P_V = D_X + D_Z, \text{ et } Q_V = D_X - D_Z.$$

**Lemme 2.4.7** *On a :*

$$\|D_{XZ}\|_F^2 \leq \frac{1}{8} \|P_V\|_F^4.$$

**Preuve :** On sait que :  $D_X D_Z + D_Z D_X = \frac{1}{2} [(D_X + D_Z)^2 - (D_X - D_Z)^2]$ , alors

$$D_{XZ} = \frac{1}{4} [P_V^2 - Q_V^2],$$

en utilisant la norme de Frobenius, on obtient :

$$\begin{aligned} \|D_{XZ}\|_F^2 &= \left\| \frac{1}{4} (P_V^2 - Q_V^2) \right\|_F^2 \\ &= \frac{1}{16} \|P_V^2 - Q_V^2\|_F^2, \\ &\leq \frac{1}{16} [\|P_V^2\|_F^2 + \|Q_V^2\|_F^2], \end{aligned}$$

puisque

$$\|P_V^2\|_F^2 \leq \|P_V\|_F^2 \|P_V\|_F^2 = \|P_V\|_F^4,$$

alors

$$\|D_{XZ}\|_F^2 \leq \frac{1}{16} [\|P_V\|_F^4 + \|Q_V\|_F^4],$$

D'après la Remarque [2.4.1], on a :

$$\|D_X + D_Z\|_F \geq \|D_X - D_Z\|_F \Leftrightarrow \|P_V\|_F \geq \|Q_V\|_F.$$

Donc

$$\|D_{XZ}\|_F^2 \leq \frac{1}{16} [\|P_V\|_F^4 + \|P_V\|_F^4] = \frac{1}{8} \|P_V\|_F^4.$$

Ce qui achève la preuve. ■

Dans les lemmes suivants, on va analyser la complexité de l'algorithme, on commence de donner quelques conditions pour assurer la stricte faisabilité du pas de Newton complet.

### Analyse de la stricte faisabilité

**Lemme 2.4.8** (*Condition suffisante*) Si

$$\delta := \delta(X, Z; \mu) < 1,$$

alors le pas de Newton complet est strictement réalisable c-à-d :

$$X(1) \succ 0 \text{ et } Z(1) \succ 0.$$

**Preuve :** D'après le Lemme [2.4.4], il suffit de montrer que  $\det(X(\alpha)Z(\alpha))$  est positif pour tout  $0 \leq \alpha \leq 1$ , alors  $X(1) \succ 0$  et  $Z(1) \succ 0$ . Pour obtenir ce résultat, on note que :

$$\begin{aligned} X(\alpha)Z(\alpha) &= (X + \alpha\Delta X)(Z + \alpha\Delta Z), \\ &= XZ + \alpha(\Delta XZ + X\Delta Z) + \alpha^2\Delta X\Delta Z. \end{aligned}$$

D'après (2.5) et (2.6), on obtient :



$$\begin{aligned}
X(\alpha)Z(\alpha) &= \mu DV^2 D^{-1} + \alpha(\mu DD_X V D^{-1} + \mu DV D_Z D^{-1}) + \alpha^2 \mu DD_X D_Z D^{-1}, \\
&= \mu D [V^2 + \alpha(D_X V + V D_Z) + \alpha^2 D_X D_Z] D^{-1}, \\
&\sim \mu [V^2 + \alpha(D_X V + V D_Z) + \alpha^2 D_X D_Z], \\
&= \mu [V^2 + \alpha(D_X V + V D_Z) + \alpha^2 D_X D_Z] + \frac{1}{2} \mu \alpha^2 D_Z D_X - \frac{1}{2} \mu \alpha^2 D_Z D_X \\
&\quad + \frac{1}{2} \mu \alpha (V D_X + D_Z V) - \frac{1}{2} \mu \alpha (V D_X + D_Z V), \\
&= B(\alpha) + M(\alpha), \tag{2.16}
\end{aligned}$$

avec

$$B(\alpha) = \mu \left[ V^2 + \frac{1}{2} \alpha (D_X V + V D_Z + V D_X + D_Z V) + \frac{1}{2} \alpha^2 (D_X D_Z + D_Z D_X) \right],$$

et

$$M(\alpha) = \mu \left[ \frac{1}{2} \alpha (D_X V + V D_Z - V D_X - D_Z V) + \frac{1}{2} \alpha^2 (D_X D_Z - D_Z D_X) \right],$$

où  $B(\alpha)$  est une matrice symétrique et  $M(\alpha)$  est anti-symétrique. Alors d'après le Lemme [2.4.5], on a :

$$\det(X(\alpha)Z(\alpha)) > 0,$$

si la matrice  $B(\alpha)$  est définie positive,  $\forall 0 \leq \alpha \leq 1$ . Pour cela, on écrit la matrice  $B(\alpha)$  sous la forme :

$$\begin{aligned}
B(\alpha) &= \mu \left[ V^2 + \frac{1}{2} \alpha ((D_X + D_Z)V + V(D_Z + D_X)) + \alpha^2 D_{XZ} \right]; \text{ où } D_X + D_Z = P_V, \\
&= \mu \left[ V^2 + \frac{1}{2} \alpha (P_V V + V P_V) + \alpha^2 D_{XZ} \right],
\end{aligned}$$

donc d'après (2.8),  $B(\alpha)$  devient :

$$\begin{aligned}
B(\alpha) &= \mu \left[ V^2 + \frac{1}{2} \alpha (V^{-1} V - V^2 + V V^{-1} - V^2) + \alpha^2 D_{XZ} \right], \\
&= \mu [(1 - \alpha) V^2 + \alpha (I + \alpha D_{XZ})], \tag{2.17}
\end{aligned}$$

et par hypothèse on a  $\delta < 1$ , alors d'après le Lemme [2.4.6], on déduit que

$$\|D_{XZ}\|_2 < 1,$$

et par conséquent  $\|\alpha D_{XZ}\|_2 < 1$ , pour tout  $0 \leq \alpha \leq 1$ . Cela implique que

$$(I + \alpha D_{XZ}) \in \mathbf{S}_{++}^n, \forall 0 \leq \alpha \leq 1.$$

D'autre part on sait que  $V \in \mathbf{S}_{++}^n$  et  $0 \leq \alpha \leq 1$ , donc

$$(1 - \alpha)V^2 \in \mathbf{S}_{++}^n.$$

Alors,  $B(\alpha)$  est définie positive pour tout  $0 \leq \alpha \leq 1$  si  $\delta < 1$ . De plus, puisque  $X(0) \succ 0$  et  $Z(0) \succ 0$ , le Lemme [2.4.4] implique que  $X(1) \succ 0$  et  $Z(1) \succ 0$ , ce qui achève la preuve. ■

Pour simplifier, on peut écrire  $X_+ = X(1)$  et  $Z_+ = Z(1)$ . Alors :

$$X_+ Z_+ \sim \mu V_+^2.$$

### La convergence quadratique de la mesure de proximité

Pour montrer la convergence quadratique de la mesure de proximité, on a besoin d'un résultat technique concernant le  $sp(V_+^2)$ .

**Lemme 2.4.9** *On a :*

$$\lambda_{\min}(V_+^2) \geq (1 - \delta^2).$$

Où  $\lambda_{\min}$  est la plus petite valeur propre de  $V_+^2$ .

**Preuve :** D'après (2.16) dans le Lemme [2.4.8], on pose  $\alpha = 1$ , on obtient :

$$\begin{aligned} X(1)Z(1) &\sim B(1) + M(1), \\ &\sim \mu I + \mu D_{XZ} + M(1) = \mu(I + D_{XZ}) + M(1), \end{aligned}$$

Comme  $X_+ Z_+ \sim \mu V_+^2$ , alors :

$$\lambda_{\min}(V_+^2) = \lambda_{\min}\left((I + D_{XZ}) + \frac{1}{\mu}M(1)\right),$$

où  $(I + D_{XZ}) \in \mathbf{S}^n$  et  $M(1) = -M(1)^\top$ . Alors d'après le Lemme [2.4.5], on obtient :

$$\begin{aligned}
\lambda_{\min}(V_+^2) &\geq \lambda_{\min}(I + D_{XZ}), \\
&\geq 1 - |\lambda_{\min}(D_{XZ})|, \\
&\geq 1 - \rho(D_{XZ}) = 1 - \|D_{XZ}\|_2, \\
&\geq 1 - \frac{1}{4} \|P_V\|_F^2, \text{ d'après le Lemme [2.4.6].} \\
&= (1 - \delta^2).
\end{aligned}$$

Ce qui achève la preuve. ■

Maintenant, on montre la convergence quadratique de la mesure de proximité.

**Lemme 2.4.10** *Si  $\delta := \delta(X, Z; \mu) < 1$ , alors :*

$$\delta_+ := \delta(X_+, Z_+; \mu) \leq \frac{\delta^2}{\sqrt{2(1 - \delta^2)}}.$$

**Preuve :** On a :

$$\begin{aligned}
\delta_+^2 &= \frac{1}{4} \|V_+^{-1} - V_+\|_F^2 = \frac{1}{4} \|V_+^{-1}(I - V_+^2)\|_F^2, \\
&\leq \frac{1}{4} \lambda_{\max}^2(V_+^{-1}) \|I - V_+^2\|_F^2, \text{ d'après le Lemme [1.2.2],} \\
&= \frac{1}{4} \frac{1}{\lambda_{\min}^2(V_+)} \|I - V_+^2\|_F^2, \\
&= \frac{1}{4} \frac{1}{\lambda_{\min}(V_+^2)} \|I - V_+^2\|_F^2.
\end{aligned}$$

D'après le Lemme [2.4.9], on a :

$$\delta_+^2 \leq \frac{1}{4(1 - \delta^2)} \|I - V_+^2\|_F^2.$$

Pour compléter la démonstration, on montre que :

$$\|I - V_+^2\|_F^2 \leq \|D_{XZ}\|_F^2.$$

Puisque  $X(1)Z(1) \sim \mu V_+^2$  et d'après le Lemme [2.4.8], on déduit que :

$$\begin{aligned}
I - V_+^2 &= I - \frac{1}{\mu} D^{-1} X(1) Z(1) D, \\
&= I - \frac{1}{\mu} D^{-1} X_+ Z_+ D, \\
&= I - \frac{1}{\mu} D^{-1} [B(1) + M(1)] D, \\
&\sim I - \frac{1}{\mu} B(1) - \frac{1}{\mu} M(1),
\end{aligned}$$

d'après l'expression (2.17) de Lemme [2.4.8] avec  $\alpha = 1$ , on obtient :

$$B(1) = \mu(I + D_{XZ}),$$

alors

$$I - V_+^2 \sim -D_{XZ} - \frac{1}{\mu} M(1).$$

En utilisant la norme de Frobenius, on obtient :

$$\begin{aligned}
\|I - V_+^2\|_F^2 &= \sum_i [\lambda_i(-D_{XZ} - \frac{1}{\mu} M(1))]^2, \\
&= \mathbf{Tr} (D_{XZ} + \frac{1}{\mu} M(1))^2, \\
&= \mathbf{Tr} (D_{XZ}^2 - \frac{1}{\mu^2} M(1)M(1)^\top + \frac{2}{\mu} (D_{XZ}M(1))),
\end{aligned}$$

puisque  $M(1) = -M(1)^\top$  et  $D_{XZ} \in \mathbf{S}^n$ , alors  $(D_{XZ}M(1))$  est anti-symétrique ce qui implique que

$$\begin{aligned}
\|I - V_+^2\|_F^2 &= \mathbf{Tr} (D_{XZ}^2 - \frac{1}{\mu^2} M(1)M(1)^\top), \text{ car } \mathbf{Tr} (D_{XZ}M(1)) = 0, \\
&\leq \mathbf{Tr} (D_{XZ}^2) = \|D_{XZ}\|_F^2.
\end{aligned}$$

Et par conséquent,

$$\|I - V_+^2\|_F^2 \leq \|D_{XZ}\|_F^2,$$

d'après le Lemme [2.4.7], on déduit que :

$$\delta_+^2 \leq \frac{1}{4(1 - \delta^2)} \|D_{XZ}\|_F^2 \leq \frac{1}{4(1 - \delta^2)} \frac{1}{8} \|P_V\|_F^4,$$

c'est équivalent à :

$$\delta_+^2 \leq \frac{1}{2(1-\delta^2)}\delta^4, \text{ où } \delta^4 = \frac{1}{16} \|P_V\|_F^4.$$

Finalement,

$$\delta_+ := \delta(X_+, Z_+; \mu) \leq \frac{\delta^2}{\sqrt{2(1-\delta^2)}}.$$

Ce qui achève la preuve. ■

Le corollaire suivant montre la convergence quadratique de la mesure de proximité avec le pas de Newton complet.

**Corollaire 2.4.1** *Si  $\delta(X, Z; \mu) < \frac{1}{\sqrt{2}}$ , alors  $\delta_+ < \delta^2(X, Z; \mu)$ .*

### L'influence du nouveau itéré en saut de dualité

Dans le lemme suivant, on analyse l'influence du pas de Newton complet sur le saut de dualité.

**Lemme 2.4.11** *Après le pas de Newton complet et si  $\delta(X, Z; \mu) < \frac{1}{\sqrt{2}}$ , alors le saut de dualité vérifie :*

$$X_+ \bullet Z_+ \leq \mu(n+1). \quad (2.18)$$

**Preuve :** on a :

$$\begin{aligned} X_+ \bullet Z_+ &= \mu \mathbf{Tr} (V_+^2) = \mathbf{Tr} (B(1) + M(1)) = \mathbf{Tr} (\mu(I + D_{XZ}) + M(1)), \\ &= \mu \mathbf{Tr} \left( (I + D_{XZ}) + \frac{1}{\mu} M(1) \right), \\ &= \mu \mathbf{Tr} (I + D_{XZ}), \text{ car si } M(1) = -M(1)^\top \Rightarrow \mathbf{Tr} (M(1)) = 0. \\ &= \mu[n + \mathbf{Tr} (D_{XZ})], \\ &= \mu \left[ n + \frac{1}{2} \mathbf{Tr} (D_X D_Z + D_Z D_X) \right], \\ &= \mu \left[ n + \mathbf{Tr} (D_X D_Z) \right], \text{ car } D_X \bullet D_Z = D_Z \bullet D_X; \\ &= \mu[n + D_X \bullet D_Z]. \end{aligned}$$

D'après l'expression (2.10) du Lemme [2.4.6], et  $\delta < \frac{1}{\sqrt{2}}$ , on obtient :

$$D_X \bullet D_Z \leq 2 \left( \frac{1}{\sqrt{2}} \right)^2 = 1,$$

par conséquent,

$$X_+ \bullet Z_+ \leq \mu(n+1).$$

Ce qui achève la preuve. ■

### La mise à jour du paramètre barrière $\mu$

La recherche d'une solution optimale primale-duale est équivalente à mettre le saut de dualité  $X \bullet Z$  vers le zéro, cela est exprimé par la mise à jour du paramètre  $\mu$  via :

$$\mu_+ = (1 - \theta)\mu, \quad 0 < \theta < 1.$$

Par la suite on montre l'influence de la mise à jour du  $\mu$  sur la nouvelle proximité après un pas de Newton complet.

**Lemme 2.4.12** *Si  $\delta(X, Z; \mu) < \frac{1}{\sqrt{2}}$  et  $\mu_+ = (1 - \theta)\mu$ , où  $0 < \theta < 1$ . Alors*

$$\delta^2(X_+, Z_+; \mu_+) \leq (1 - \theta)\delta_+^2 + \frac{\theta^2(n + 1)}{4(1 - \theta)} + \frac{\theta}{2}. \quad (2.19)$$

De plus, si  $\delta := \delta(X, Z; \mu) < \frac{1}{\sqrt{2}}$ ,  $\theta = \frac{1}{2\sqrt{n}}$  et  $n \geq 2$ , alors :

$$\delta^2(X_+, Z_+; \mu_+) < \frac{1}{2}.$$

**Preuve :** Soient  $V_+^2 = \frac{1}{\mu}D^{-1}X_+Z_+D$ ,  $\mu_+ = (1 - \theta)\mu$  et  $\delta < \frac{1}{\sqrt{2}}$ . Alors,

$$\begin{aligned} 4\delta^2(X_+, Z_+; \mu_+) &= \left\| \sqrt{1 - \theta}V_+^{-1} - \frac{1}{\sqrt{1 - \theta}}V_+ \right\|_F^2, \\ &= \left\| \sqrt{1 - \theta}V_+^{-1} - \frac{(1 - \theta) + \theta}{\sqrt{1 - \theta}}V_+ \right\|_F^2, \\ &= \left\| \sqrt{1 - \theta}(V_+^{-1} - V_+) - \frac{\theta}{\sqrt{1 - \theta}}V_+ \right\|_F^2, \\ &= (1 - \theta) \|V_+^{-1} - V_+\|_F^2 + \frac{\theta^2}{1 - \theta} \|V_+\|_F^2 - 2\theta \mathbf{Tr}((V_+^{-1} - V_+)V_+), \\ &= (1 - \theta) \|V_+^{-1} - V_+\|_F^2 + \frac{\theta^2}{1 - \theta} \|V_+\|_F^2 - 2\theta \mathbf{Tr}(I - V_+^2), \\ &= (1 - \theta)4\delta_+^2 + \frac{\theta^2}{1 - \theta} \|V_+\|_F^2 - 2\theta n + 2\theta \|V_+\|_F^2, \end{aligned}$$

d'après le Lemme [2.4.11], on a :

$$\|V_+\|_F^2 = \mathbf{Tr}(V_+^2) = \frac{1}{\mu}X_+ \bullet Z_+ \leq (n + 1).$$

Donc

$$4\delta^2(X_+, Z_+; \mu_+) \leq (1 - \theta)4\delta_+^2 + \frac{\theta^2(n+1)}{1 - \theta} - 2\theta n + 2\theta(n+1),$$

par conséquent

$$\delta^2(X_+, Z_+; \mu_+) \leq (1 - \theta)\delta_+^2 + \frac{\theta^2(n+1)}{4(1 - \theta)} + \frac{\theta}{2}.$$

Comme  $\delta < \frac{1}{\sqrt{2}}$ , d'après le Corollaire [2.4.1] on obtient  $\delta_+^2 < \delta^4 < \frac{1}{4}$ , et par conséquent

$$\delta^2(X_+, Z_+; \mu_+) \leq \frac{(1 - \theta)}{4} + \frac{\theta^2(n+1)}{4(1 - \theta)} + \frac{\theta}{2}.$$

Maintenant, on considère  $\theta = \frac{1}{2\sqrt{n}}$ , alors

$$\begin{aligned} \delta^2(X_+, Z_+; \mu_+) &\leq \frac{(1 - \theta)}{4} + \frac{\frac{1}{4n}(n+1)}{4(1 - \theta)} + \frac{\theta}{2}, \quad \frac{n+1}{4n} \leq \frac{3}{8}, \quad \forall n \geq 2, \\ &\leq \frac{(1 - \theta)}{4} + \frac{\theta}{2} + \frac{3}{32(1 - \theta)}, \end{aligned}$$

pour  $n \geq 2$ , on a :  $0 \leq \theta \leq \frac{1}{\sqrt{8}}$  et la fonction

$$f(\theta) = \frac{(1 - \theta)}{4} + \frac{\theta}{2} + \frac{3}{32(1 - \theta)},$$

est strictement croissante sur l'intervalle  $[0, \frac{1}{\sqrt{8}}]$ . Par conséquent,

$$f(\theta) \leq f\left(\frac{1}{\sqrt{8}}\right) \simeq 0.48341 < \frac{1}{2}, \quad \forall \theta \in [0, \frac{1}{\sqrt{8}}],$$

finalement  $\delta^2(X_+, Z_+; \mu_+) < \frac{1}{2}$ . Ce qui achève la preuve. ■

### Analyse de la complexité

Dans cette partie, on présente le nombre d'itérations à partir duquel l'Algorithme 2.4.4 converge.

**Lemme 2.4.13** Soient  $X^{(k+1)}$  et  $Z^{(k+1)}$  la  $(k+1)$  ième itération produite par l'Algorithme 2.4.4, avec  $\mu := \mu^{(k)}$ , alors

$$X^{(k+1)} \bullet Z^{(k+1)} \leq \epsilon,$$

est satisfaite si

$$k \geq \frac{1}{\theta} \log \left( \frac{n}{\epsilon} \right).$$

**Preuve :** D'après le Lemme [2.4.11], on a :

$$X^{(k+1)} \bullet Z^{(k+1)} \leq \mu^{(k)}(n+1), \quad (2.20)$$

avec

$$\mu^{(k)} = (1 - \theta)\mu^{(k-1)} = (1 - \theta)^k \mu^{(0)}.$$

Alors (2.20) devient :

$$X^{(k+1)} \bullet Z^{(k+1)} \leq (1 - \theta)^k \mu^{(0)}(n+1),$$

où l'inégalité  $X^{(k+1)} \bullet Z^{(k+1)} \leq \epsilon$  est satisfaite si :

$$(1 - \theta)^k \mu^{(0)}(n+1) \leq \epsilon.$$

En utilisant la fonction croissante  $\log$  on obtient :

$$k \log(1 - \theta) \leq \log \epsilon - \log(\mu^{(0)}(n+1)) \leq \log \epsilon - \log(\mu^{(0)}n),$$

puisque  $\log(1 - \theta) \leq -\theta$ , pour  $0 < \theta < 1$ , alors

$$k\theta \geq \log \left( \frac{\mu^{(0)}n}{\epsilon} \right).$$

Par conséquent

$$k \geq \frac{1}{\theta} \log \left( \frac{\mu^{(0)}n}{\epsilon} \right).$$

Ce qui achève la preuve. ■

On note que pour  $\theta = \frac{1}{2\sqrt{n}}$ , on obtient le théorème suivant :

**Théorème 2.4.1** Soient  $\theta = \frac{1}{2\sqrt{n}}$ ,  $\beta = \frac{1}{\sqrt{2}}$ . Alors, l'algorithme de la trajectoire centrale de type primal-dual avec le pas de Newton complet nécessite de  $O(\sqrt{n} \log(\frac{n\mu^{(0)}}{\epsilon}))$  itérations.

**Preuve :** Soient  $\theta = \frac{1}{2\sqrt{n}}$ ,  $\beta = \frac{1}{\sqrt{2}}$ . Alors le résultat découle immédiatement du Lemme [2.4.13]. Ce qui achève la preuve. ■



# Chapitre 3

## Méthodes de points intérieurs basée sur une nouvelle fonction noyau pour CQSDP

Dans ce chapitre, on développe un algorithme d'une méthode de points intérieurs de type primal-dual basé sur une nouvelle fonction noyau pour résoudre le problème CQSDP. Premièrement, on commence par la notion d'une fonction matricielle qui sera utilisée par la suite de ce chapitre et l'introduction d'une nouvelle fonction noyau qui doit vérifier les conditions d'éligibilité. Puis, on établit la convergence de l'algorithme et l'analyse de la complexité suivant un schéma donné.

### 3.1 Les fonctions noyaux et leurs propriétés

Dans cette section, on définit une fonction matricielle à partir d'une fonction différentiable  $\psi(t)$  et à valeurs réelles. Puis, on introduit la notion d'une fonction noyau et leurs propriétés pour CQSDP.

**Définition 3.1.1** Soient  $X \in \mathbb{S}^n$  et

$$X = Q_X^{-1} \text{diag}(\lambda_1(X), \dots, \lambda_n(X)) Q_X,$$

la décomposition spectrale de  $X$ , où  $Q_X$  est orthogonale. La fonction matricielle  $\psi(X)$  est définie par :

$$\psi(X) = Q_X^{-1} \text{diag}(\psi(\lambda_1(X)), \dots, \psi(\lambda_n(X))) Q_X.$$

On note que la fonction  $\psi(X)$  est bien définie lorsque  $\psi(t)$  est bien définie à chaque

valeur propre de  $X$  et

$$\psi'(X) = Q_X^{-1} \text{diag}(\psi'(\lambda_1(X)), \dots, \psi'(\lambda_n(X))) Q_X.$$

La Définition 3.1.1 est appelée Théorème de Décomposition Spectrale d'une matrice symétrique. Elle nous permet de faire l'extension de la définition d'une fonction  $\psi : \mathbb{R} \rightarrow \mathbb{R}$  à une fonction de  $\mathbb{S}^n$  dans  $\mathbb{S}^n$  (voir [21]).

Maintenant, nous avons besoin de quelques concepts liés à la fonction matricielle. Pour plus de détails, consulter les ouvrages ([21], [28]).

**Définition 3.1.2** Une matrice  $M(t)$  est dite matrice de fonction si chaque élément de  $M(t)$  est une fonction c-à-d  $M(t) = [M_{ij}(t)]$ .

Par la suite, on présente les trois inégalités suivantes : si  $M, N \in \mathbb{S}^n$ , on a

$$|\mathbf{Tr}(MN)| \leq |\lambda_{\max}(M)| \sum_{i=1}^n |\lambda_i(N)|. \quad (3.1)$$

De plus, si  $M_1 \preceq M_2$  et  $N \succeq 0$ , on a

$$\mathbf{Tr}(M_1 N) \leq \mathbf{Tr}(M_2 N). \quad (3.2)$$

Soient  $M, M + N \in \mathbb{S}_+^n$ . Alors on a :

$$\lambda_i(M + N) \geq \lambda_{\min}(M) - |\lambda_{\max}(N)|, \quad i = 1, 2, \dots, n. \quad (3.3)$$

Notons que la dernière inégalité se découle du Lemme [5.18] en ([39]).

Les concepts habituels de la continuité, la différentiabilité et l'intégrabilité peuvent naturellement être étendus aux fonctions matricielles comme suit :

Soient  $M(t)$  et  $N(t)$  deux fonctions matricielles différentiables. Alors on a :

$$\begin{cases} \frac{d}{dt} M(t) &= [\frac{d}{dt} M_{ij}(t)] = M'(t), \\ \frac{d}{dt} \mathbf{Tr}(M(t)) &= \mathbf{Tr}(M'(t)), \\ \frac{d}{dt} \mathbf{Tr}(\psi(M(t))) &= \mathbf{Tr}(\psi'(M(t))M'(t)), \\ \frac{d}{dt} (M(t)N(t)) &= M'(t)N(t) + M(t)N'(t). \end{cases}$$

### 3.1.1 Notion d'une fonction noyau

**Définition 3.1.3** Soit  $\psi : ]0, +\infty[ \rightarrow ]0, +\infty[$ , on dit que  $\psi$  est une fonction noyau barrière si elle est différentiable et vérifie les conditions suivantes :

1.  $\psi(1) = \psi'(1) = 0$ ;
2.  $\psi''(t) > 0$ , pour tout  $t > 0$ ;
3.  $\lim_{t \rightarrow 0} \psi(t) = \lim_{t \rightarrow +\infty} \psi(t) = +\infty$ .

Les deux premières conditions désignent que  $\psi(t)$  est strictement convexe et admet une valeur minimale, si  $t = 1$ . On définit  $\psi(t)$  à partir de sa dérivée seconde comme suit :

$$\psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi. \quad (3.4)$$

De plus, la troisième condition indique la propriété de barrière pour la fonction  $\psi(t)$ .

Dans ce travail, on considère la fonction  $\psi$  définie de  $\mathbb{R}_{++}$  dans  $\mathbb{R}_+$  par :

$$\psi(t) = \frac{t^2 - 1}{2} - \frac{t - q}{q^2 - q + 1} e^{q(\frac{1}{t}-1)} + \frac{1 - q}{q^2 - q + 1}, \quad t > 0, \quad q \geq 1, \quad (3.5)$$

Les trois premières dérivées de  $\psi(t)$  sont :

$$\begin{aligned} \psi'(t) &= t - \frac{e^{q(\frac{1}{t}-1)}(t^2 - qt + q^2)}{t^2(q^2 - q + 1)}, \\ \psi''(t) &= 1 + \frac{q^2 e^{q(\frac{1}{t}-1)}(t + q)}{t^4(q^2 - q + 1)} > 0, \quad \text{pour tout } t > 0 \text{ et } q \geq 1, \\ \psi'''(t) &= -\frac{e^{q(\frac{1}{t}-1)}(3q^2t^2 + 5q^3t + q^4)}{t^6(q^2 - q + 1)} < 0, \quad \text{pour tout } t > 0 \text{ et } q \geq 1. \end{aligned}$$

Il est clair que  $\psi(1) = \psi'(1) = 0$  de plus

$$\lim_{t \rightarrow 0} \psi(t) = \lim_{t \rightarrow +\infty} \psi(t) = +\infty.$$

Donc,  $\psi(t)$  est une fonction noyau barrière.

### 3.1.2 Qualification d'une fonction noyau

Dans cette partie, on présente les conditions de qualification d'une fonction noyau barrière  $\psi(t)$  et quelques propriétés concernant ces conditions. On commence par le lemme technique suivant :

**Lemme 3.1.1** *Soit  $\psi(t)$  une fonction noyau barrière donnée. Si elle vérifie les deux propriétés suivantes :*

1.  $t\psi''(t) - \psi'(t) > 0, t > 1,$
2.  $\psi'''(t) < 0, t > 0.$

Alors  $\psi(t)$  vérifie

$$\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) > 0, t > 1, \beta > 1.$$

**Preuve :** Supposons que  $\psi(t)$  vérifié (1) et (2). Soit  $t > 1$ , on considère :

$$f(\beta) = \psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t), \beta > 1.$$

On note que  $f(1) = 0$ . De plus,

$$\begin{aligned} f'(\beta) &= t\psi''(t)\psi'''(\beta t) - \psi'(t)\psi'''(\beta t) - \beta t\psi'(t)\psi''''(\beta t), \\ &= \psi'''(\beta t)[t\psi''(t) - \psi'(t)] - \beta t\psi'(t)\psi''''(\beta t) > 0, \end{aligned}$$

la dernière inégalité est satisfaite du fait que  $\psi'''(\beta t) > 0, t\psi''(t) - \psi'(t) > 0$  d'après (1) et  $-\beta t\psi'(t)\psi''''(\beta t) > 0$ , car  $t > 1, \psi'(t) > 0$  et  $\psi''''(\beta t) < 0$  d'après (2). Ce qui implique que  $f(\beta) > 0$ , pour tout  $\beta > 1$ . Ce qui achève la preuve. ■

**Lemme 3.1.2** Soit  $\psi(t)$  la fonction définie en (3.5), on dit que  $\psi(t)$  est éligible (qualifiée) si elle vérifie les conditions suivantes :

$$t\psi''(t) + \psi'(t) > 0, t < 1, \quad (3.6)$$

$$t\psi''(t) - \psi'(t) > 0, t > 1, \quad (3.7)$$

$$\psi'''(t) < 0, t > 0, \quad (3.8)$$

$$2\psi''(t)^2 - \psi'''(t)\psi'(t) > 0, t < 1, \quad (3.9)$$

$$\psi''(t)\psi'(\beta t) - \beta\psi'(t)\psi''(\beta t) > 0, t > 1, \beta > 1 \quad (3.10)$$

**Remarque 3.1.1** D'après le Lemme [3.1.1], il est montré que (3.7) et (3.8) impliquent (3.10). Donc, pour montrer la qualification de  $\psi(t)$ , on vérifie seulement les conditions de (3.6) à (3.9).

**Preuve :** Pour (3.6), on remplace  $\psi'(t)$  et  $\psi''(t)$ , on obtient :

$$t\psi''(t) + \psi'(t) = 2t + \frac{e^{q(\frac{1}{t}-1)}}{t^3(q^2 - q + 1)}(-t^3 + qt^2 + q^3).$$

Puisque  $0 < t < 1$  et  $q \geq 1$ , alors  $q^2 - q + 1 > 0$  et

$$-t^3 + qt^2 + q^3 \geq (q-1)t^3 + q^3 > 0,$$

donc la condition (3.6) est satisfaite. De plus, on a pour tout  $t > 0$  et  $q \geq 1$ ,

$$t\psi''(t) - \psi'(t) = \frac{e^{q(\frac{1}{t}-1)}}{t^3(q^2 - q + 1)}(t^3 - qt^2 + 2q^2t + q^3).$$

Alors, pour montrer la condition (3.7), il suffit de prouver que  $t^3 - qt^2 + 2q^2t + q^3 > 0$  pour tout  $t > 1$  et  $q \geq 1$ . Soit  $h(t) = t^3 - qt^2 + 2q^2t + q^3$ . On a

$$\begin{aligned} h'(t) &= 3t^2 - 2qt + 2q^2, \\ &= 3\left(t - \frac{q}{3}\right)^2 + \frac{5}{3}q^2 > 0 \end{aligned}$$

pour tout  $t > 1$  et  $q \geq 1$ . Ce qui implique que  $h(t)$  est strictement croissante pour tout  $t > 1$ ,  $q \geq 1$  et

$$h(t) \geq h(1) = q(q^2 - 1) + 2q^2 + 1 > 0.$$

Cela montre que la condition (3.7) est satisfaite. La condition (3.8) est vérifiée à partir de l'expression du  $\psi'''(t)$ . Finalement,

$$\begin{aligned} 2\psi''(t)^2 - \psi'''(t)\psi'(t) &= 2 + \frac{q^2 e^{2q(\frac{1}{t}-1)}}{(q^2 - q + 1)^2 t^8} (-3t^4 - 2qt^3 + 3q^2t^2 + q^4) \\ &\quad + \frac{e^{q(\frac{1}{t}-1)}}{(q^2 - q + 1)t^5} (7t^2 + 9qt + q^2). \end{aligned}$$

Afin de simplifier l'écriture, on pose  $B = e^{q(\frac{1}{t}-1)}$  et  $C = q^2 - q + 1$ , puisque  $q \geq 1$  et  $0 < t < 1$  alors on obtient  $B^2 \geq B > 1$  de même,  $C^2 \geq C \geq 1$ .

En utilisant ces deux résultats, on trouve

$$\begin{aligned} 2\psi''(t)^2 - \psi'''(t)\psi'(t) &= 2 + \frac{q^2 B^2}{C^2 t^8} (-3t^4 - 2qt^3 + 3q^2t^2 + q^4) + \frac{B}{C t^5} (7t^2 + 9qt + q^2), \\ &\geq 2 + \frac{1}{C^2 t^8} (-3t^4 - 2qt^3 + 3q^2t^2 + q^4) + \frac{1}{C^2 t^5} (7t^2 + 9qt + q^2), \\ &= 2 + \frac{1}{C^2 t^8} (7t^5 + (9q - 3)t^4 + q(q - 2)t^3 + 3q^2t^2 + q^4). \end{aligned}$$

On a deux cas à discuter :

Si  $q \in [2, +\infty[$ , alors

$$7t^5 + (9q - 3)t^4 + q(q - 2)t^3 + 3q^2t^2 + q^4 > 0,$$

et par conséquent

$$2\psi''(t)^2 - \psi'''(t)\psi'(t) > 0, \text{ pour tout } 0 < t < 1.$$

Si  $q \in [1, 2[$ , on considère

$$g(t) = q(q - 2)t^3 + 3q^2t^2 + q^4,$$

alors

$$7t^5 + (9q - 3)t^4 + q(q - 2)t^3 + 3q^2t^2 + q^4 = 7t^5 + (9q - 3)t^4 + g(t) > 0,$$

si et seulement si  $g(t) > 0$  pour tout  $0 < t < 1$  et  $q \in [1, 2[$ . On a

$$g'(t) = 3q(q - 2)t^2 + 6q^2t = 0 \Leftrightarrow t = 0 \text{ ou } t = \frac{2q}{2 - q} > 0.$$

On note que si  $t = \frac{2q}{2 - q}$  avec  $q \in [1, 2[$  alors  $t \geq 2$ , donc cette racine est rejetée. On s'intéresse seulement au  $t = 0$ , pour laquelle on a

$$g''(t) = 6q(q - 2)t + 6q^2,$$

qui est nulle en  $t = \frac{q}{2 - q} \geq 1$ , pour tout  $q \in [1, 2[$ , ce qui implique que

$$g''(t) > 6q(q - 2)t + 6q^2t = 12q(q - 1)t \geq 0, \text{ pour tout } 0 < t < 1.$$

Et par conséquent,  $g'(t)$  est strictement croissante pour tout  $0 < t < 1$  c-à-d  $g'(t) > g'(0) = 0$ . Donc, la fonction  $g(t)$  est aussi strictement croissante pour tout  $0 < t < 1$  c-à-d  $g(t) \geq g(0) = q^4 > 0$ , et par conséquent,  $2\psi''(t)^2 - \psi'''(t)\psi'(t) > 0$ , pour tout  $0 < t < 1$  et  $q \in [1, 2[$ . Et la condition (3.9) est satisfaite pour tout  $0 < t < 1$  et  $q \geq 1$ . Finalement,  $\psi(t)$  est une fonction noyau éligible (qualifiée). Ce qui achève la preuve. ■

maintenant, on définit la fonction barrière  $\Psi(V) : \mathbb{S}_{++}^n \rightarrow \mathbb{R}_+$  associée à une fonction noyau éligible par :

$$\Psi(X, Z, \mu) := \Psi(V) := \sum_{i=1}^n \psi(\lambda_i(V)) = \mathbf{Tr} (\psi(V)), \lambda_i(V) > 0. \quad (3.11)$$

**Remarque 3.1.2** *Par la suite, on utilise la fonction  $\psi(\cdot)$  et ses trois premières dérivées  $\psi'(\cdot), \psi''(\cdot), \psi'''(\cdot)$  qui désignent une fonction matricielle si l'argument est une matrice et une fonction réelle si l'argument est un réel.*

D'après les deux premières conditions de la Définition 3.1.3, on a le lemme suivant, qui montre la propriété importante de la fonction barrière  $\Psi(V)$ , dont la preuve de ce lemme est identique à celle de la proposition 3 (I) en [28]. Par conséquent, on omet sa preuve.

**Lemme 3.1.3** [28, Proposition 3 (I)] *Soit  $\Psi(V)$  la fonction barrière associée à une fonction noyau  $\psi(t)$ , alors  $\Psi(V)$  est strictement convexe par rapport à  $V \succ 0$ , et vérifie  $\Psi(I) = \Psi'(I) = 0$ .*

### 3.1.3 Propriétés d'une fonction noyau éligible

Maintenant, on présente les lemmes techniques les plus utilisés dans ce chapitre.

**Lemme 3.1.4** *On a :*

1.  $\psi(t) \leq \frac{1}{2}\psi''(1)(t-1)^2$ , pour tout  $t \geq 1$ ,
2.  $\psi(t) \leq \frac{t^2-1}{2}$ , pour tout  $t \geq 1$ .

**Preuve :** Pour (1), en utilisant le Théorème de Taylor avec  $\psi(1) = \psi'(1) = 0$ ,  $\psi''(1) = 1 + \frac{q^2(1+q)}{q^2-q+1}$ ,  $q \geq 1$  et  $\psi'''(t) < 0$ , on obtient :

$$\begin{aligned} \psi(t) &= \psi(1) + \psi'(1)(t-1) + \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{3!}\psi'''(c)(c-1)^3, \\ &= \frac{1}{2}\psi''(1)(t-1)^2 + \frac{1}{3!}\psi'''(c)(c-1)^3, \\ &< \frac{1}{2}\psi''(1)(t-1)^2 = \left(1 + \frac{q^2(1+q)}{q^2-q+1}\right) \frac{(t-1)^2}{2}, \text{ où } 1 < c < t. \end{aligned}$$

Pour (2), on a :

$$\psi(t) = \frac{t^2-1}{2} + \psi_b(t), \quad t > 0,$$

où

$$\psi_b(t) = -\frac{t-q}{q^2-q+1} e^{q(\frac{1}{t}-1)} + \frac{1-q}{q^2-q+1},$$

désigne le terme barrière de  $\psi(t)$ . Alors, on obtient :

$$\psi(t) \leq \frac{t^2-1}{2},$$

car  $\psi_b(1) = 0$ , et  $\psi_b(t)$  est strictement décroissante pour tout  $t \geq 1$  et  $q \geq 1$ . Ce qui achève la preuve. ■

Le lemme suivant montre que toute fonction noyau éligible  $\psi(t)$  est exponentiellement convexe, dont l'importance de ce lemme est justifié par la suite (voir [9], [29]).

**Lemme 3.1.5** [29, Lemme 2.1.2] *Les propriétés suivantes sont équivalentes :*

- $\psi(\sqrt{t_1 t_2}) \leq \frac{1}{2}(\psi(t_1) + \psi(t_2))$ , pour tout  $t_1, t_2 > 0$ ,
- $t\psi''(t) + \psi'(t) > 0$ ,  $t > 0$ ,
- $\psi(e^\xi)$  est convexe.

On s'appuyant sur le Lemme [3.1.5], on obtient le lemme suivant qui est crucial pour analyser un algorithme de points intérieurs basé sur une fonction noyau pour CQSDP.

**Lemme 3.1.6** [29, Proposition 5.2.6] *Pour tout  $V_1, V_2 \succ 0$ , alors*

$$\Psi \left( \left[ V_1^{\frac{1}{2}} V_2 V_1^{\frac{1}{2}} \right]^{\frac{1}{2}} \right) \leq \frac{1}{2}(\Psi(V_1) + \Psi(V_2)).$$

## 3.2 Nouvelles directions cherchées

L'idée principale des fonctions noyaux, basée sur les directions de Nesterov Todd est de remplacer le membre à droite  $V^{-1} - V$  de la troisième équation du système (2.7) par  $-\psi'(V)$ , alors on obtient le système linéaire suivant :

$$\begin{cases} \bar{A}_i \bullet D_X & = 0, \quad i = 1, \dots, m, \\ \sum_{i=1}^m (D_y)_i \bar{A}_i + D_Z - \bar{Q}(D_X) & = 0, \\ D_X + D_Z & = -\psi'(V), \end{cases} \quad (3.12)$$

où  $\psi(t)$  est une fonction noyau donnée et  $\psi(V), \psi'(V)$  sont les fonctions matricielles associées. On note que si  $\psi(t)$  est la fonction noyau classique (la fonction barrière logarithmique)

$$\psi(t) = \frac{t^2-1}{2} - \log t, \quad \forall t > 0,$$



alors le système (3.12) coïncide avec les directions classiques de NT en (2.7). De plus, puisque le système (3.12) admet la même matrice de coefficients que le système (2.7) et sous nos hypothèses, le système (3.12) admet une solution unique. Ainsi, notons que c'est facile de vérifier que les directions cherchées ne sont pas orthogonales contrairement au cas SDP linéaire c-à-d  $D_X \bullet D_Z = \frac{1}{\mu} \Delta X \bullet \mathcal{Q}(\Delta X) \geq 0$ .

Les nouvelles directions cherchées  $(D_X, D_y, D_Z)$  sont obtenues par la résolution de système (3.12). Puis, on utilise (2.6) pour obtenir  $(\Delta X, \Delta y, \Delta Z)$ . Durant la recherche d'une direction, en prenant une taille de pas  $\alpha$  convenable, on obtient le nouvel itéré donné par :

$$X_+ = X + \alpha \Delta X, \quad y_+ = y + \alpha \Delta y, \quad Z_+ = Z + \alpha \Delta Z. \quad (3.13)$$

Pour analyser un algorithme de points intérieurs, on définit la mesure de proximité  $\delta(V) : \mathbb{S}_{++}^n \rightarrow \mathbb{R}_+$  par :

$$\delta := \delta(V) = \frac{1}{2} \|D_X + D_Z\|_F = \frac{1}{2} \|\psi'(V)\|_F = \frac{1}{2} \sqrt{\mathbf{Tr}(\psi'(V))^2}. \quad (3.14)$$

D'après le Lemme [3.1.3]  $\Psi(V)$  est strictement convexe et atteint sa valeur minimale zéro à  $V = I$ , alors on a :

$$\delta(V) = 0 \Leftrightarrow V = I \Leftrightarrow \Psi(V) = 0.$$

Par conséquent, la valeur de  $\Psi(V)$  peut être considérée comme une mesure de la distance entre un itéré donné  $(X, y, Z)$  et  $\mu$ -centre  $(X(\mu), y(\mu), Z(\mu))$ .

### 3.2.1 Algorithme

Dans cette partie, on présente un algorithme générique de trajectoire centrale de type primal-dual basé sur une fonction noyau pour CQSDP :

## Algorithme générique primal-dual pour CQSDP

**Début d'algorithme**

**Données :**

un paramètre de précision  $\epsilon > 0$ ;

un paramètre de proximité  $\tau, \tau \geq 1$ ;

un paramètre  $0 < \theta < 1$ ;

un paramètre de barrière  $\mu^{(0)} = \frac{X^0 \bullet Z^0}{n}$ .

**Initialisation :** soit  $(X^0, y^0, Z^0)$  vérifiant la **CPI** tel que

$\Psi(X^0, Z^0; \mu^{(0)}) \leq \tau$  et  $k = 0$ ;

**Tant que**  $n\mu^{(k)} \geq \epsilon$  faire : (itération externe)

**Début**

$\mu^{(k+1)} = (1 - \theta)\mu^{(k)}$ ;

**Tant que**  $\Psi(X^{(k)}, Z^{(k)}; \mu^{(k)}) > \tau$  faire : (itération interne)

**Début**

- calculer  $(\Delta X^{(k)}, \Delta y^{(k)}, \Delta Z^{(k)})$  via le système (3.12);
- déterminer un pas  $\alpha$ ;
- mise à jour  $X^{(k+1)} = X^{(k)} + \alpha\Delta X^{(k)}, y^{(k+1)} = y^{(k)} + \alpha\Delta y^{(k)}, Z^{(k+1)} = Z^{(k)} + \alpha\Delta Z^{(k)}$  et  $k = k + 1$ ;

**Fin tant que.**

**Fin tant que.**

**Fin algorithme.**

### Algorithme 3.2.1

D'après la description ci-dessus on sait que l'appartenance d'un point  $(X, y, Z)$  à  $(X(\mu), y(\mu), Z(\mu))$  est mesurée par la valeur de  $\Psi(V)$ , avec  $\tau \geq 1$  la valeur seuil : Si  $\Psi(V) < \tau$ , on commence une nouvelle itération externe en effectuant une mise à jour du paramètre  $\mu$  via :

$$\mu^+ = (1 - \theta)\mu, \quad 0 < \theta < 1.$$

Sinon ( $\Psi(V) \geq \tau$ ), nous entrons dans une itération interne par le calcul des directions cherchées à l'itération en cours par rapport à la valeur actuelle de  $\mu$  et on applique (3.13) pour obtenir des nouveaux itérés. Si nécessaire, on va répéter la procédure jusqu'à ce qu'on trouve un itéré au voisinage de  $(X(\mu), y(\mu), Z(\mu))$ . Alors  $\mu$  est encore réduit par un facteur de  $(1 - \theta)$  avec  $0 < \theta < 1$ , et on applique la méthode de Newton pour trouver les nouveaux  $\mu$ -centres, et ainsi de suite. Ce processus se répète jusqu'à ce que  $n\mu \leq \epsilon$ . A ce stade, on trouve une solution  $\epsilon$ -approximative pour CQSDP.

Les paramètres  $\tau, \theta$  et  $\alpha$  décrits dans l'algorithme sont choisis tel que le nombre d'itérations produites par l'algorithme sera minimal.

Par la suite on montre l'influence de la mise à jour du  $\mu$  sur la valeur de la fonction barrière  $\Psi(X, Z, \mu) := \Psi(V)$  après chaque itération externe.

### 3.2.2 Une borne supérieure de $\Psi(V)$ après chaque itération externe

Notons qu'au début de chaque itération externe du l'Algorithme 3.2.1, juste avant la mise à jour de  $\mu$ , on a  $\Psi(V) \leq \tau$ . Par la mise à jour de  $\mu$ , la matrice  $V$  est divisée par  $\sqrt{1-\theta}$ , ce qui conduit généralement à une augmentation des valeurs de  $\Psi(V)$ . Alors, pendant les itérations internes ultérieures,  $\Psi(V)$  diminue jusqu'à ce qu'il dépasse à nouveau le seuil  $\tau$ . Par conséquent, durant l'algorithme, la plus grande valeur de  $\Psi(V)$  se produit juste après la mise à jour de  $\mu$ . Pour cela, dans cette sous section, on donne une estimation de l'effet de la mise à jour du  $\mu$  sur la valeur de  $\Psi(V)$ . En d'autre terme, on trouve une borne supérieure de  $\Psi\left(\frac{1}{\sqrt{1-\theta}}V\right)$  en fonction de  $\Psi(V)$ .

Il sera clair que certaines fonctions inverses liées à la fonction noyau et sa première dérivée jouent un rôle important pour analyser l'Algorithme 3.2.1. On introduit ces fonctions inverses ici.

On note par  $\varrho : [0, +\infty[ \rightarrow [1, +\infty[$  et  $\rho : [0, \infty[ \rightarrow ]0, 1]$  les fonctions inverses de  $\psi(t)$  pour  $t \geq 1$ , et  $-\frac{1}{2}\psi'(t)$  pour  $t \in ]0, 1]$ , respectivement. En d'autre terme :

$$s = \psi(t) \Leftrightarrow t = \varrho(s), t \geq 1, \quad (3.15)$$

$$s = -\frac{1}{2}\psi'(t) \Leftrightarrow t = \rho(s), t \leq 1. \quad (3.16)$$

On a le théorème suivant qui est représenté en [19] et dont la démonstration dépend de la condition (3.10).

**Théorème 3.2.1** [19, Théorème 3.3.2] *Pour toute matrice définie positive  $V$  et tout  $\beta > 1$  on a :*

$$\Psi(\beta V) \leq n\psi\left(\beta\varrho\left(\frac{\Psi(V)}{n}\right)\right).$$

Par conséquence du Théorème [3.2.1], on a : si  $\Psi(V) \leq \tau$  et  $\beta = \frac{1}{\sqrt{1-\theta}}$  alors après la mise à jour de  $\mu$ ,  $\Psi\left(\frac{1}{\sqrt{1-\theta}}V\right)$  est bornée supérieurement par  $n\psi\left(\beta\varrho\left(\frac{\Psi(V)}{n}\right)\right)$ . Comme  $\varrho$  est

croissante et  $\Psi(V) \leq \tau$ , alors  $\varrho\left(\frac{\Psi(V)}{n}\right) \leq \varrho\left(\frac{\tau}{n}\right)$ . Et donc

$$\Psi_0 \leq L_\psi(n, \theta, \tau) := n\psi\left(\frac{\varrho\left(\frac{\tau}{n}\right)}{\sqrt{1-\theta}}\right),$$

où  $\Psi_0$  désigne la valeur de  $\Psi(V)$  après la mise à jour de  $\mu$ .

### 3.2.3 La décroissance de la fonction barrière durant une itération interne

Dans cette partie, on va calculer le pas de déplacement  $\alpha$  qui conserve la stricte faisabilité de nouveau itéré et réduire la valeur de proximité durant les itérations internes. Pour  $\mu$  fixé, si on prend  $\alpha$  le pas de déplacement, alors le nouveau itéré est donné par :

$$X_+ = X + \alpha\Delta X, \quad y_+ = y + \alpha\Delta y, \quad Z_+ = Z + \alpha\Delta Z.$$

D'après (2.5) et (2.6), on peut écrire

$$X_+ = X + \alpha\Delta X = X + \alpha\sqrt{\mu}DD_XD = \sqrt{\mu}D(V + \alpha D_X)D,$$

et

$$Z_+ = Z + \alpha\Delta Z = Z + \alpha\sqrt{\mu}D^{-1}D_ZD^{-1} = \sqrt{\mu}D^{-1}(V + \alpha D_Z)D^{-1}.$$

Alors

$$V_+^2 := (V + \alpha D_X)(V + \alpha D_Z),$$

et

$$V_+^2 = (V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_Z)(V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_X)^{-\frac{1}{2}},$$

On suppose que  $V + \alpha D_X \succ 0$  et  $V + \alpha D_Z \succ 0$  pour une telle taille de pas  $\alpha$  qui assure la stricte faisabilité. Alors il est clair que  $V_+^2$  est similaire à la matrice suivante :

$$(V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_Z)(V + \alpha D_X)^{\frac{1}{2}}.$$

Et par conséquent, on déduit que les valeurs propres de  $V_+$  sont les mêmes que la matrice

$$\left[ (V + \alpha D_X)^{\frac{1}{2}}(V + \alpha D_Z)(V + \alpha D_X)^{\frac{1}{2}} \right]^{\frac{1}{2}},$$

donc on a :

$$\Psi(V_+) = \Psi \left( \left[ (V + \alpha D_X)^{\frac{1}{2}} (V + \alpha D_Z) (V + \alpha D_X)^{\frac{1}{2}} \right]^{\frac{1}{2}} \right).$$

Pour tout  $\alpha > 0$ , on considère la fonction suivante :

$$f(\alpha) = \Psi(V_+) - \Psi(V).$$

Alors  $f(\alpha)$  est la différence de la proximité entre le nouveau itéré et l'ancien itéré, pour  $\mu$  fixé.

En utilisant le Lemme [3.1.6], on obtient :

$$\Psi(V_+) \leq \frac{1}{2} [\Psi(V + \alpha D_X) + \Psi(V + \alpha D_Z)].$$

Et par conséquent  $f(\alpha) \leq f_1(\alpha)$ , où

$$f_1(\alpha) = \frac{1}{2} [\Psi(V + \alpha D_X) + \Psi(V + \alpha D_Z)] - \Psi(V)$$

est une fonction convexe par rapport à  $\alpha$  car  $\Psi$  est convexe. Evidemment,  $f(0) = f_1(0) = 0$ . Maintenant, pour estimer la décroissance de la fonction de proximité durant un pas, on utilise les deux dérivées successive de  $f_1(\alpha)$  par rapport à  $\alpha$ . En utilisant la règle de différentiabilité pour les fonctions matricielles, on obtient

$$f_1'(\alpha) = \frac{1}{2} \mathbf{Tr} [\psi'(V + \alpha D_X) D_X + \psi'(V + \alpha D_Z) D_Z],$$

et

$$f_1''(\alpha) = \frac{1}{2} \mathbf{Tr} [\psi''(V + \alpha D_X) D_X^2 + \psi''(V + \alpha D_Z) D_Z^2].$$

D'après la définition de  $\delta$ , et puisque  $D_X + D_Z = -\psi'(V)$ , on trouve

$$\begin{aligned} f_1'(0) &= \frac{1}{2} \mathbf{Tr} [\psi'(V) D_X + \psi'(V) D_Z], \\ &= \frac{1}{2} \mathbf{Tr} [\psi'(V) (D_X + D_Z)], \\ &= -2\delta^2. \end{aligned}$$

Le lemme suivant présente une borne supérieure de  $f_1''(\alpha)$  pour CQSDP (voir [41, Théorème 9], [44, Lemme 4.1]) qui est identique à celle pour LP et SDP où les directions sont orthogonales (voir [9, Lemma 4.1], [19, Lemme 3.4.4]). Pour la démonstration voir l'Annexe.

**Lemme 3.2.1** [44, Lemme 4.1] On a :

$$f_1''(\alpha) \leq 2\delta^2 \psi''(\lambda_{\min}(V) - 2\alpha\delta),$$

**Lemme 3.2.2** [9, Lemma 4.2] On a  $f_1'(\alpha) \leq 0$ , si  $\alpha$  vérifie l'inégalité suivante :

$$-\psi'(\lambda_{\min}(V) - 2\alpha\delta) + \psi'(\lambda_{\min}(V)) \leq 2\delta.$$

**Lemme 3.2.3** [9, Lemma 4.3] La plus grande valeur de  $\alpha$  qui vérifie le Lemme [3.2.2] est donnée par :

$$\bar{\alpha} := \frac{1}{2\delta}(\rho(\delta) - \rho(2\delta)).$$

**Lemme 3.2.4** [9, Lemma 4.4] Soient  $\rho$  et  $\bar{\alpha}$  définis dans le Lemme [3.2.3]. Alors

$$\bar{\alpha} \geq \frac{1}{\psi''(\rho(2\delta))}.$$

Par la suite, on note par :

$$\tilde{\alpha} := \frac{1}{\psi''(\rho(2\delta))}, \tag{3.17}$$

et on utilise  $\tilde{\alpha}$  comme taille de pas par défaut. d'après le Lemme [3.2.4] on a  $\tilde{\alpha} \leq \bar{\alpha}$ .

Le lemme suivant montre que la proximité  $\Psi$  est décroissante. Pour plus de détails voir ([28]).

**Lemme 3.2.5** [28, Lemma 3.12] Soit  $h(t)$  une fonction convexe est deux fois différentiable telle que  $h(0) = 0$ ,  $h'(0) < 0$  et atteint son minimum globale en  $t^* > 0$ , et  $h''(t)$  est croissante sur l'intervalle  $[0, t^*]$ . Alors on a :

$$h(t) \leq \frac{th'(0)}{2}, \text{ pour tout } t \in [0, t^*].$$

Et par conséquent, on a le lemme suivant qui présente une borne supérieure pour la valeur décroissance de la proximité à une itération interne.

**Lemme 3.2.6** [9, Lemma 4.5] Si le pas de déplacement  $\alpha$  vérifie  $\alpha \leq \bar{\alpha}$ . Alors, on a :

$$f(\alpha) \leq -\alpha\delta^2.$$

En combinant les résultats de Lemme [3.2.6] et (3.17) on obtient :

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))}, \tag{3.18}$$

avec  $\tilde{\alpha}$  étant la taille de pas par défaut, comme déjà indiqué en (3.17).

**Corollaire 3.2.1** *Le membre droit de l'expression (3.18) est strictement décroissante par rapport à  $\delta$ .*

### 3.2.4 Borne de $\delta(V)$ en terme de $\Psi(V)$

Le théorème suivant présente une borne inférieure pour  $\delta(V)$  en terme de  $\Psi(V)$ . Notons que la preuve est identique au cas SDP (voir [19, Théorème 3.3.3]).

**Théorème 3.2.2** [19, Théorème 3.3.3] *Soit  $\varrho$  la fonction inverse de  $\psi(t)$  pour  $t \geq 1$ . Pour toute  $V \in \mathbb{S}_{++}^n$ , on a :*

$$\delta(V) \geq \frac{1}{2} \psi'(\varrho(\Psi(V))).$$

### 3.2.5 Borne d'itérations

On doit compter le nombre d'itérations internes nécessaire pour revenir à la situation où  $\Psi(V) \leq \tau$ . On note que  $\Psi_0$  désigne la valeur de  $\Psi(V)$  après la mise à jour de  $\mu$ , les valeurs suivantes dans la même itération externe sont notées par  $\Psi_k$ ,  $k = 1, 2, \dots, K$ , avec  $K$  désigne le nombre total d'itérations internes produites durant une itération externe. En utilisant le lemme suivant, pour la démonstration voir l'Annexe.

**Lemme 3.2.7** [28, Proposition 2.2] *Soit  $t_0, t_1, \dots, t_K$  une suite des nombres positifs qui vérifie :*

$$t_{k+1} \leq t_k - \bar{\beta} (t_k)^{1-\gamma}, \quad k = 0, 1, \dots, K-1,$$

tels que  $\bar{\beta} > 0$  et  $\gamma \in (0, 1]$ ; alors :

$$K \leq \left\lceil \frac{t_0^\gamma}{\bar{\beta}\gamma} \right\rceil.$$

**Lemme 3.2.8** *Soit  $K$  le nombre d'itérations internes, alors on a :*

$$K \leq \frac{\Psi_0^\gamma}{\gamma\bar{\beta}}.$$

**Preuve :** La définition de  $K$  implique que  $\Psi_{K-1} > \tau$  et  $\Psi_K \leq \tau$  et (d'après la décroissance de  $f(\tilde{\alpha})$  on obtient la valeur de  $\gamma$  et  $\bar{\beta}$ )

$$\Psi_{k+1} \leq \Psi_k - \bar{\beta} (\Psi_k)^{1-\gamma}, \quad k = 0, 1, \dots, K-1.$$

En utilisant le Lemme [3.2.7] avec  $t_k = \Psi_k$ , on obtient :

$$K \leq \frac{\Psi_0^\gamma}{\gamma\beta}.$$

Ce qui achève la preuve. ■

Le lemme suivant présente le nombre d'itérations externes à partir duquel l'Algorithme 3.2.1 converge.

**Lemme 3.2.9** *Soit  $k$  le nombre d'itérations externes nécessaire pour que l'Algorithme 3.2.1 converge vers une solution optimale primale-duale. Alors on a :*

$$k \geq \frac{1}{\theta} \log \left( \frac{n}{\epsilon} \right).$$

**Preuve :** D'après l'Algorithme 3.2.1, l'obtention d'une solution optimale primale-duale nécessite que

$$n\mu^{(k)} \leq \epsilon.$$

Notons que à chaque itération externe  $\mu_k$  se réduit par un facteur de  $(1 - \theta)$  c'est à dire :

$$\mu^{(k)} = (1 - \theta)\mu^{(k-1)} = (1 - \theta)^k \mu^{(0)},$$

on prend  $\mu^{(0)} = 1$ , alors

$$n\mu^{(k)} \leq n(1 - \theta)^k \leq \epsilon.$$

En utilisant la fonction croissante  $\log$  on obtient :

$$k \log(1 - \theta) \leq \log \epsilon - \log n,$$

puisque  $\log(1 - \theta) \leq -\theta$ , pour  $0 < \theta < 1$ , alors

$$k\theta \geq \log \left( \frac{n}{\epsilon} \right).$$

Par conséquent

$$k \geq \frac{1}{\theta} \log \left( \frac{n}{\epsilon} \right).$$

Ce qui achève la preuve. ■

En multipliant le nombre d'itérations externes par le nombre d'itérations internes, on



obtient le nombre total d'itérations produites par l'Algorithme 3.2.1 qui est donnée par :

$$\frac{\Psi_0^\gamma}{\theta\gamma\beta} \log \frac{n}{\epsilon}. \quad (3.19)$$

Pour conclure cette section, on présente dans le tableau suivant la dépendance des résultats obtenus jusqu'à présent sur les conditions d'éligibilité.

Conditions	Théorème (TH), Lemme (L) ou Corollaire (C)
3.6	L 3.1.5
3.7	L 3.1.1
3.8	L 3.1.1, L 3.1.4, L 3.2.4, L 3.2.1, TH 3.2.2
3.9	C 3.2.1
3.10	TH 3.2.1

Tableau 2. Utilisation des conditions d'éligibilité.

### 3.3 La convergence de l'algorithme et l'analyse de la complexité

#### 3.3.1 Schéma pour analyser un algorithme basé sur une fonction noyau éligible

L'analyse d'un algorithme générique basé sur une fonction noyau éligible nécessite de suivre le schéma suivant qui est déjà présenté par Bai et al en [9].

- Etape 1.** Résoudre l'équation  $-\frac{1}{2}\psi'(t) = s$  pour obtenir  $\rho(s)$ , la fonction inverse de  $-\frac{1}{2}\psi'(t)$  pour  $t \leq 1$ . Si l'équation est difficile à résoudre, trouver une borne inférieure pour  $\rho(s)$ .
- Etape 2.** Calculer la décroissance de  $\Psi(V)$  en terme de  $\delta = \delta(V)$  pour la taille de pas par défaut  $\tilde{\alpha}$ , tel que :  $f(\tilde{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))}$ .
- Etape 3.** Résoudre l'équation  $\psi(t) = s$  pour obtenir  $\varrho(s)$ , la fonction inverse de  $\psi(t)$  pour  $t \geq 1$ . Si l'équation est difficile à résoudre, trouver une borne supérieure pour  $\varrho(s)$ .
- Etape 4.** Dédire une borne inférieure pour  $\delta(V)$  en terme de  $\Psi(V)$  de la forme :  $\delta(V) \geq \frac{1}{2}\psi'(\varrho(\Psi(V)))$ .
- Etape 5.** Utiliser les résultats des étapes **3** et **4** pour obtenir les constantes positives  $\gamma$  et  $\bar{\beta}$ , avec  $\gamma \in (0, 1]$  et telle que :  $f(\tilde{\alpha}) \leq -\bar{\beta}\Psi(V)^{1-\gamma}$ .
- Etape 6.** Calculer une borne supérieure  $\Psi_0$  pour  $\Psi(V)$  sachant que :  $\Psi_0 \leq L_\psi(n, \theta, \tau) = n\psi\left(\frac{\varrho(\frac{\tau}{n})}{\sqrt{1-\theta}}\right)$ .
- Etape 7.** Dédire une borne supérieure pour le nombre d'itérations totales sous la forme :  $\frac{\Psi_0^\gamma}{\theta\gamma\bar{\beta}} \log \frac{n}{\epsilon}$ .
- Etape 8.** Si  $\tau = \Theta(n)$  et  $\theta = O(1)$  on obtient une borne d'itérations pour les algorithmes à grand pas. Si  $\tau = O(1)$  et  $\theta = \Theta(n)$  on obtient une borne d'itérations pour les algorithmes à petit pas.

Figure 1. Schéma pour analyser un algorithme basé sur une fonction noyau éligible.

Dans ce schéma, on utilise la fonction barrière  $\Psi(V)$  définie en (3.11) et la mesure de proximité  $\delta(V)$  définie en (3.14) pour analyser l'Algorithme 3.2.1.

Puisque la fonction noyau paramétré qui est présenté en (3.5) vérifie les conditions d'éligibilité, on peut simplement appliquer le schéma de la Figure 1. à cette fonction noyau. Les étapes suivantes de ce schéma sont justifiées par des résultats antérieurs comme indiqué

dans le tableau suivant :

Etape	Basée sur
1	Equation 3.16
2	Equation 3.18
3	Equation 3.15
4	Théorème 3.2.2
5	Les étapes 3 et 4
6	Théorème 3.2.1
7	Equation 3.19
8	Les algorithmes à grand et petit pas

Tableau 3. Justification de la validité du régime à la Figure 1.

### 3.3.2 Analyse de la complexité

Dans cette partie, on présente le nombre d'itérations à partir duquel l'Algorithme 3.2.1 converge. Pour cela, on suit le schéma précédent, on note que dans les étapes 1 et 3 on a besoin à des fonctions  $\rho(s)$  et  $\varrho(s)$ . Puisque on peut pas trouver les expressions explicites de ces deux fonctions inverses, on présente les deux lemmes suivants (voir [9]) qui nous donne quelques bornes de  $\rho(s)$  et  $\varrho(s)$ . Pour la démonstration voir l'Annexe.

**Lemme 3.3.1** [9, Lemme 6.2] Soit  $\varrho : [0, +\infty[ \longrightarrow [1, +\infty[$  la fonction inverse de  $\psi(t)$  pour  $t \geq 1$ . Alors on a :

$$\sqrt{1+2s} \leq \varrho(s) \leq 1 + \sqrt{2s}, \quad s \geq 0.$$

**Lemme 3.3.2** [9, Lemme 6.1] Soit  $\psi_b(t)$  le terme barrière de  $\psi(t)$  et soit  $\bar{\rho} : [0, \infty[ \rightarrow ]0, 1]$  la fonction inverse de la restriction  $-\psi'_b(t)$  sur l'intervalle  $]0, 1]$ . Alors on a :

$$\rho(s) \geq \bar{\rho}(1+2s).$$

On rappelle que la fonction noyau paramétré qui est présentée en (3.5) est éligible, donc on peut appliquer le schéma ci-dessus.

**Etape 1.** d'après l'équation  $-\psi'_b(t) = s$ , on a :

$$-\psi'_b(t) = \frac{e^{q(\frac{1}{t}-1)}(t^2 - qt + q^2)}{t^2(q^2 - q + 1)} = s.$$

Puisque  $t \in ]0, 1]$ , on obtient

$$e^{q(\frac{1}{t}-1)} = \frac{t^2(q^2 - q + 1)s}{(t^2 - qt + q^2)} \leq \frac{(q^2 - q + 1)s}{(t^2 - qt + q^2)}.$$

Pour tout  $t \in ]0, 1]$ , on a :

$$\begin{cases} t^2 - qt + q^2 \geq \frac{3q^2}{4} & \text{si } q \in [1, 2) \\ t^2 - qt + q^2 \geq q^2 - q + 1 & \text{si } q \geq 2. \end{cases}$$

Maintenant, c'est facile d'avoir que  $q^2 - q + 1 \geq \frac{3q^2}{4}$ , pour tout  $q \geq 1$ , alors

$$t^2 - qt + q^2 \geq \frac{3q^2}{4}, \text{ pour tout } q \geq 1.$$

Ce qui nous conduit à

$$\begin{aligned} e^{q(\frac{1}{t}-1)} &\leq \frac{4(q^2 - q + 1)s}{3q^2}, \text{ pour tout } q \geq 1. \\ &\leq \frac{4s}{3}, \text{ pour tout } q \geq 1, \end{aligned}$$

dont la dernière inégalité est vérifiée car  $\frac{q^2 - q + 1}{q^2} \leq 1$ , pour tout  $q \geq 1$ , et par conséquent on déduit que

$$t = \bar{\rho}(s) \geq \frac{1}{1 + \frac{1}{q} \log(\frac{4s}{3})}, \text{ pour tout } q \geq 1.$$

En utilisant le Lemme [3.3.2], puisque  $\rho(s) \geq \bar{\rho}(1 + 2s)$ , on trouve une borne inférieure pour  $\rho(s)$  comme suit :

$$\rho(s) \geq \frac{1}{1 + \frac{1}{q} \log(\frac{4}{3}(1 + 2s))}, \text{ pour tout } q \geq 1.$$

**Etape 2.** La fonction  $\psi''(t)$  est strictement décroissante, alors

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{\psi''(\rho(2\delta))} \leq -\frac{\delta^2}{\psi''(\bar{\rho}(1 + 4\delta))},$$

On pose  $t = \bar{\rho}(1 + 4\delta)$  et puisque  $t \leq 1$ , on peut écrire

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{\psi''(t)} \leq -\frac{\delta^2}{1 + \frac{q^2(t+q)e^{q(\frac{1}{t}-1)}}{t^4(q^2-q+1)}}. \quad (3.20)$$

On note que

$$t = \bar{\rho}(1 + 4\delta) \Leftrightarrow 1 + 4\delta = -\psi'_b(t) = \frac{e^{q(\frac{1}{t}-1)}(t^2 - qt + q^2)}{t^2(q^2 - q + 1)}.$$

En remplaçant cette dernière dans (3.20), on obtient :

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{1 + (1+q)(1+4\delta)\frac{q^2}{(t^2-qt+q^2)t^2}}.$$

Autre fois, en utilisant cette inégalité

$$t^2 - qt + q^2 \geq \frac{3q^2}{4}, \text{ pour tout } q \geq 1,$$

on obtient

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{1 + (1+q)(1+4\delta)\frac{4}{3t^2}}, \text{ pour tout } q \geq 1.$$

D'autre part, on a

$$\frac{1}{t^2} = \frac{1}{(\bar{\rho}(1+4\delta))^2} \leq \left(1 + \frac{1}{q} \log \left(\frac{4(1+4\delta)}{3}\right)\right)^2, \text{ pour tout } q \geq 1,$$

donc

$$f(\tilde{\alpha}) \leq -\frac{\delta^2}{1 + \frac{4}{3}(1+q)(1+4\delta) \left(1 + \frac{1}{q} \log \left(\frac{4}{3}(1+4\delta)\right)\right)^2}, \text{ pour tout } q \geq 1. \quad (3.21)$$

**Etape 3.** D'après le Lemme [3.3.1], la fonction inverse de  $\psi(t)$  pour  $t \geq 1$  vérifiée :

$$\sqrt{1 + 2\psi(t)} \leq \varrho(\psi(t)) \leq 1 + \sqrt{2\psi(t)}. \quad (3.22)$$

**Etape 4.** En utilisant  $\delta(V) \geq \frac{1}{2}\psi'(\varrho(\Psi(V)))$ , et puisque  $\psi'_b(t)$  est strictement croissante pour tout  $q \geq 1$  et  $t \geq 1$ , où  $\psi_b(t)$  désigne le terme barrière de  $\psi(t)$ , on obtient :

$$\begin{aligned} \delta(V) &\geq \frac{1}{2}\psi'(\varrho(\Psi(V))), \\ &\geq \frac{1}{2} \left( \sqrt{1+2\Psi} - \frac{e^{q(\frac{1}{\sqrt{1+2\Psi}}-1)}}{(1+2\Psi)(q^2-q+1)} (1+2\Psi - q\sqrt{1+2\Psi} + q^2) \right), \\ &\geq \frac{1}{2}(\sqrt{1+2\Psi} - 1) = \frac{\Psi}{1+\sqrt{1+2\Psi}} \geq \frac{1}{2}\sqrt{\frac{\Psi}{3}}. \end{aligned} \quad (3.23)$$

**Etape 5.** Soit  $\Psi_0 \geq \Psi \geq \tau \geq 1$ . Du l'expression (3.23), on déduit que  $\sqrt{\Psi} \leq 2\sqrt{3}\delta \leq 4\delta$ . Donc, d'après (3.21), on a :

$$\begin{aligned} f(\tilde{\alpha}) &\leq -\frac{\delta^2}{1 + \frac{4}{3}(2\sqrt{3}\delta + 4\delta)(1+q) \left(1 + \frac{1}{q} \log\left(\frac{4}{3}(1+4\delta)\right)\right)^2}, \text{ pour tout } q \geq 1, \\ &\leq -\frac{\delta}{\frac{60}{3}(1+q) \left(1 + \frac{1}{q} \log\left(\frac{4}{3}(1+4\delta)\right)\right)^2}, \text{ pour tout } q \geq 1, \end{aligned}$$

et par conséquent

$$f(\tilde{\alpha}) \leq -\frac{\sqrt{\Psi}}{80(1+q) \left(1 + \frac{1}{q} \log\left(\frac{4}{3}(1+\sqrt{\Psi_0})\right)\right)^2}, \text{ pour tout } q \geq 1. \quad (3.24)$$

Donc

$$\gamma = \frac{1}{2} \text{ et } \bar{\beta} = \frac{1}{80(1+q) \left(1 + \frac{1}{q} \log\left(\frac{4}{3}(1+\sqrt{\Psi_0})\right)\right)^2}, \text{ pour tout } q \geq 1,$$

**Etape 6.** D'après le Lemme [3.3.1], on a  $\varrho(\frac{\tau}{n}) \leq 1 + \sqrt{\frac{2\tau}{n}}$ . Et par conséquent

$$\Psi_0 \leq L_\psi(n, \theta, \tau) = n\psi\left(\frac{\varrho(\frac{\tau}{n})}{\sqrt{1-\theta}}\right) \leq n\psi\left(\frac{1 + \sqrt{\frac{2\tau}{n}}}{\sqrt{1-\theta}}\right).$$

En utilisant la deuxième propriété du Lemme [3.1.4], on trouve :

$$\Psi_0 \leq \frac{n}{2} \left( \frac{(1 + \sqrt{\frac{2\tau}{n}})^2}{1 - \theta} - 1 \right) = \frac{n\theta + 2\sqrt{2n\tau} + 2\tau}{2(1 - \theta)}. \quad (3.25)$$

**Etape 7.** D'après l'inégalité (3.24), le nombre total d'itérations est borné par :

$$\frac{\Psi_0^\gamma}{\gamma\beta} \frac{1}{\theta} \log \frac{n}{\epsilon} = 160(1 + q) \left( 1 + \frac{1}{q} \log \left( \frac{4}{3}(1 + \sqrt{\Psi_0}) \right) \right)^2 \Psi_0^{\frac{1}{2}} \frac{1}{\theta} \log \frac{n}{\epsilon}, \quad q \geq 1. \quad (3.26)$$

En remplaçant l'expression (3.25) dans (3.26), on obtient :

$$\frac{\Psi_0^\gamma}{\theta\gamma\beta} \log \frac{n}{\epsilon} \leq 114(1 + q) \left( 1 + \frac{1}{q} \log \left( \frac{4}{3} \left( 1 + \frac{\chi}{\sqrt{2}} \right) \right) \right)^2 \chi \frac{1}{\theta} \log \frac{n}{\epsilon}, \quad \text{pour tout } q \geq 1, \quad (3.27)$$

avec  $\chi = \left( \frac{n\theta + 2\sqrt{2n\tau} + 2\tau}{1 - \theta} \right)^{\frac{1}{2}}$ .

**Etape 8.** Pour les algorithmes à grand pas, on considère  $\tau = \Theta(n)$  et  $\theta = O(1)$ . Et par conséquent,

$$\Psi_0 \leq \frac{n\theta + 2\sqrt{2n\tau} + 2\tau}{2(1 - \theta)} = O(n).$$

Si on prend

$$q = \log \left( \frac{4}{3}(1 + O(n)) \right) = O(\log n),$$

la borne d'itérations en (3.27) devienne

$$O(\sqrt{n} \log n \log \frac{n}{\epsilon}),$$

qui est la meilleure complexité pour les algorithmes à grand pas jusqu'à maintenant.

Pour les algorithmes à petit pas, on considère  $\tau = O(1)$  et  $\theta = \Theta(\frac{1}{\sqrt{n}})$ . En utilisant la

première propriété du Lemme [3.1.4] et le Lemme [3.3.1], on obtient :

$$\begin{aligned}
\Psi_0 &\leq n\psi\left(\frac{\varrho(\frac{\tau}{n})}{\sqrt{1-\theta}}\right), \\
&\leq \frac{n}{2}\psi''(1)\left(\frac{\varrho(\frac{\tau}{n})}{\sqrt{1-\theta}}-1\right)^2, \\
&= \frac{n}{2}\left(1+\frac{q^2(1+q)}{q^2-q+1}\right)\left(\frac{\varrho(\frac{\tau}{n})}{\sqrt{1-\theta}}-1\right)^2, \\
&= \frac{n}{2}\left(1+\frac{q^2(1+q)}{q^2-q+1}\right)\left(\frac{\varrho(\frac{\tau}{n})-\sqrt{1-\theta}}{\sqrt{1-\theta}}\right)^2, \\
&\leq \frac{n}{2}\left(1+\frac{q^2(1+q)}{q^2-q+1}\right)\left(\frac{1+\sqrt{\frac{2\tau}{n}}-\sqrt{1-\theta}}{\sqrt{1-\theta}}\right)^2,
\end{aligned}$$

puisque  $1-\sqrt{1-\theta}\leq\theta$ , pour tout  $0<\theta<1$ , alors

$$\Psi_0\leq\frac{1}{2}\left(1+\frac{q^2(1+q)}{q^2-q+1}\right)\left(\frac{\sqrt{n\theta}+\sqrt{2\tau}}{\sqrt{1-\theta}}\right)^2.$$

Maintenant, en remplaçant cette dernière expression dans (3.26), on obtient une borne supérieure du nombre total d'itérations donné par :

$$114(1+q)\left(1+\frac{1}{q}\log\left(\frac{4}{3}\left(1+\frac{\eta}{\sqrt{2}}\right)\right)\right)^2\eta\frac{1}{\theta}\log\frac{n}{\epsilon},\quad q\geq 1. \quad (3.28)$$

avec  $\eta=\left(1+\frac{q^2(1+q)}{q^2-q+1}\right)^{\frac{1}{2}}\left(\frac{\sqrt{n\theta}+\sqrt{2\tau}}{\sqrt{1-\theta}}\right)$ .

Puisque  $\tau=O(1)$  et  $\theta=\Theta\left(\frac{1}{\sqrt{n}}\right)$ , alors

$$\Psi_0\leq O\left(1+\frac{q^2(1+q)}{q^2-q+1}\right).$$

Si on prend  $q=\text{constant}$ , alors la borne d'itérations en (3.28) devienne

$$O(\sqrt{n}\log\frac{n}{\epsilon}),$$

qui est la meilleure complexité pour les algorithmes à petit pas.



# Chapitre 4

## Tests numériques

Dans ce chapitre, on s'intéresse aux expériences numériques issues de l'application des Algorithmes 2.4.4 et 3.2.1 sur des problèmes quadratiques convexes semi-définis. Notre implémentation est faite sous Matlab 8.1 (R2013a) sur intel CORE i3.

On note par :

- $\delta(X, Z; \mu)$  : la mesure de proximité qui est associée à l'Algorithme 2.4.4.
- $\Psi(X, Z; \mu)$  : la fonction barrière qui est associée à l'Algorithme 3.2.1.
- $(X^0, y^0, Z^0)$  : un point initial strictement réalisable et vérifie  $\delta(X^0, Z^0; \mu^{(0)}) \leq \frac{1}{\sqrt{2}}$  pour l'Algorithme 2.4.4 et  $\Psi(X^0, Z^0; \mu^{(0)}) \leq \tau$  pour l'Algorithme 3.2.1 avec  $\mu^{(0)} > 0$  et  $\tau$  est précisé par la suite.
- $(X^*, y^*, Z^*)$  : la solution optimale du problème primal ( $P$ ) et dual ( $D$ ), respectivement.
- **iter** : le nombre des itérations produites par l'Algorithme 2.4.4.
- **inner** : le nombre des itérations internes produites par l'Algorithme 3.2.1.
- **outer** : le nombre des itérations extérieures produites par l'Algorithme 3.2.1.
- **CPU** : le temps d'exécution d'un algorithme en seconde.

### 4.1 Calcul de la matrice $V$

On rappelle que les deux algorithmes proposés en Chapitre 2 et Chapitre 3 nécessitent le calcul de la matrice  $V$  via l'expression (2.5) qui dépend de la matrice  $P$ . Pour cela, on utilise les techniques introduite par Todd, Toh et Tütüncü pour calculer la matrice  $P$  (voir [26], [34]). Tout d'abord on a besoin du lemme suivant :

**Lemme 4.1.1** *Soit  $B \in \mathbf{S}_{++}^n$  telle que  $B = CC^\top$ . Alors, la matrice  $W = C^{-1}B^{\frac{1}{2}}$  est orthogonale.*

**Preuve :** On a  $WW^\top = C^{-1}BC^{-\top} = C^{-1}CC^\top C^{-\top} = I$ . Ce qui achève la preuve. ■  
Maintenant, on considère la factorisation de Cholesky des deux matrices  $X$  et  $Z$  :

$$X = LL^\top, Z = RR^\top,$$

où  $L$  (respectivement,  $R$ ) sont des matrices triangulaires inférieures et on considère maintenant la matrice  $R^\top L$  où sa décomposition SVD est donnée par :

$$R^\top L = U\Sigma\Lambda^\top, \text{ t.q } \Sigma = \text{diag}(\sigma_i) \text{ et } \sigma_i \text{ sont les valeurs singulières non nulles de } R^\top L.$$

Posons  $W = L^{-1}X^{\frac{1}{2}}$ , d'après le Lemme [4.1.1]  $W$  est une matrice orthogonale, alors on obtient :

$$\begin{aligned} X^{\frac{1}{2}}ZX^{\frac{1}{2}} &= X^{\frac{1}{2}}L^{-\top}(L^\top RR^\top L)L^{-1}X^{\frac{1}{2}}, \\ &= W^\top(L^\top RR^\top L)W, \\ &= W^\top(\Lambda\Sigma U^\top)(U\Sigma\Lambda^\top)W, \\ &= (W^\top\Lambda)\Sigma^2(\Lambda^\top W), \end{aligned}$$

telle que  $\Lambda\Sigma^2\Lambda^\top$  est la décomposition en valeurs propres de  $L^\top ZL$ . Comme  $W^\top\Lambda$  est orthogonale, on obtient :

$$(X^{\frac{1}{2}}ZX^{\frac{1}{2}})^{-\frac{1}{2}} = (W^\top\Lambda)\Sigma^{-1}(\Lambda^\top W),$$

alors on a :

$$\begin{aligned} P &= X^{\frac{1}{2}}(X^{\frac{1}{2}}ZX^{\frac{1}{2}})^{-\frac{1}{2}}X^{\frac{1}{2}}, \\ &= X^{\frac{1}{2}}(W^\top\Lambda)\Sigma^{-1}(\Lambda^\top W)X^{\frac{1}{2}}, \\ &= X^{\frac{1}{2}}X^{\frac{1}{2}}L^{-\top}(\Lambda\Sigma^{-1}\Lambda^\top)L^{-1}X^{\frac{1}{2}}X^{\frac{1}{2}}, \\ &= XL^{-\top}(\Lambda\Sigma^{-1}\Lambda^\top)L^{-1}X, \\ &= LL^\top L^{-\top}(\Lambda\Sigma^{-1}\Lambda^\top)L^{-1}LL^\top, \\ &= L\Lambda\Sigma^{-1}\Lambda^\top L^\top = GG^\top. \end{aligned}$$

Où  $G = L\Lambda\Sigma^{-\frac{1}{2}}$ .

**Remarque 4.1.1** Notons que  $P \in \mathbf{S}_{++}^n$ , alors  $P$  est diagonalisable et ses valeurs propres sont réelles positives c'est à dire, il existe une matrice orthogonale  $U$  telle que

$$P = U \text{diag}(\lambda_1, \dots, \lambda_n)U^\top, \text{ avec } \lambda_i \in \text{sp}(P).$$

Puisque  $D = P^{\frac{1}{2}}$ , alors

$$D = U \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}) U^\top.$$

Par conséquent, la matrice  $V$  se calcule via la formule suivante :  $m$

$$V = \frac{1}{\sqrt{\mu}} D Z D.$$

## 4.2 Calcul du pas de déplacement

Malgré les expériences numériques avec le choix théorique sélectionné du pas de déplacement  $\alpha$  durant chaque itération interne garantie la convergence de la méthode, il nous conduit à des résultats non favorable c'est-à-dire la convergence de l'algorithme correspondant est lente. Pour cela, on utilise un choix pratique basé sur la stratégie suivante qui été déjà présenté par Achache [1] pour LP. A chaque itération interne on calcule le pas de déplacement maximal  $\alpha_{\max}$  tels que  $X + \xi \alpha_{\max} \Delta X \succ 0$  et  $Z + \xi \alpha_{\max} \Delta Z \succ 0$  avec  $\alpha_{\max} = \min(\alpha_X, \alpha_Z)$  et  $\xi \in (0, 1)$ , où  $\alpha_X$  et  $\alpha_Z$  sont les pas de déplacement réalisable primal et dual donnés par :  $\alpha_X = \min_{i=1}^n \alpha'_X[i]$ , avec

$$\alpha'_X[i] = \begin{cases} \frac{-1}{\lambda_i \left( X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}} \right)} & , \text{ si } \lambda_i \left( X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}} \right) < 0 \\ 1 & , \text{ si } \lambda_i \left( X^{-\frac{1}{2}} \Delta X X^{-\frac{1}{2}} \right) \geq 0, \end{cases}$$

de même  $\alpha_Z = \min_{i=1}^n \alpha'_Z[i]$ , avec

$$\alpha'_Z[i] = \begin{cases} \frac{-1}{\lambda_i \left( Z^{-\frac{1}{2}} \Delta Z Z^{-\frac{1}{2}} \right)} & , \text{ si } \lambda_i \left( Z^{-\frac{1}{2}} \Delta Z Z^{-\frac{1}{2}} \right) < 0 \\ 1 & , \text{ si } \lambda_i \left( Z^{-\frac{1}{2}} \Delta Z Z^{-\frac{1}{2}} \right) \geq 0. \end{cases}$$

## 4.3 Problèmes à tester

Dans cette partie, on présente les différents problèmes CQSDP à tester dont les matrices  $A_i$  sont linéairement indépendantes et l'opérateur  $\mathcal{Q} : \mathbf{S}^n \rightarrow \mathbf{S}^n$  est auto-adjoint et monotone.

### 4.3.1 Problèmes à taille variable

#### Problème 1.

Soient  $n = 2m$ ,  $\mathcal{Q}(X) = 0_{n \times n}$  et

$$C[i, j] = \begin{cases} -1 & \text{si } i = j \text{ et } i \leq m, \\ 0 & \text{ailleurs,} \end{cases}$$

$$A_i[j, k] = \begin{cases} 1 & \text{si } j = k = i \text{ ou } j = k = i + m, \\ 0 & \text{ailleurs,} \end{cases}$$

et

$$b[i] = 2, \quad i = 1, \dots, m,$$

la solution initiale primale-duale et strictement réalisable est donnée par :

$$X^0[i, j] = \begin{cases} 2 - \gamma & \text{si } i = j = 1, \dots, m, \\ \gamma & \text{si } i = j = m + 1, \dots, n, \\ 0 & \text{ailleurs,} \end{cases}$$

$$Z^0[i, j] = \begin{cases} \frac{1}{\gamma} - 1 & \text{si } i = j = 1, \dots, m, \\ \frac{1}{\gamma} & \text{si } i = j = m + 1, \dots, n, \\ 0 & \text{ailleurs,} \end{cases}$$

$$y^0[i] = -\frac{1}{\gamma}, \quad i = 1, \dots, m,$$

avec  $\gamma = (4 - \sqrt{8})/2$ . Pour cet exemple, on prend  $\mu^{(0)} = \frac{X^0 \bullet Z^0}{n} = 1$  et  $\epsilon = 10^{-5}$  avec  $\Psi(X^0, Z^0; \mu^{(0)}) \simeq 0 \leq \tau$ , où  $\tau = n$  et  $\delta(X^0, Z^0; \mu^{(0)}) \simeq 0 \leq \frac{1}{\sqrt{2}}$ .

Une solution optimale primale-duale est donnée par :

$$X^*[i, j] = \begin{cases} 2 & \text{si } i = j = 1, \dots, m, \\ 0 & \text{ailleurs,} \end{cases}$$

$$Z^*[i, j] = \begin{cases} 1 & \text{si } i = j = m + 1, \dots, n, \\ 0 & \text{ailleurs,} \end{cases}$$

et

$$y^*[i] = -1, \quad i = 1, \dots, m.$$

La valeur optimale pour les deux problèmes est égale à  $-2m = -n$ .

**Problème 2.**

Soient  $n = 2m$ ,  $\mathcal{Q}(X) = X$  et

$$C[i, j] = \begin{cases} -1 & \text{si } i = j \text{ et } i \leq m, \\ 0 & \text{ailleurs,} \end{cases}$$

$$A_i[j, k] = \begin{cases} 1 & \text{si } j = k = i \text{ ou } j = k = i + m, \\ 0 & \text{ailleurs,} \end{cases}$$

et

$$b[i] = 2, \quad i = 1, \dots, m,$$

la solution initiale primale-duale et strictement réalisable est donnée par :

$$X^0[i, j] = \begin{cases} \frac{5}{4} & \text{si } i = j = 1, \dots, m, \\ \frac{3}{4} & \text{si } i = j = m + 1, \dots, n, \\ 0 & \text{ailleurs,} \end{cases}$$

$$Z^0[i, j] = \begin{cases} \frac{10}{12} & \text{si } i = j = 1, \dots, m, \\ \frac{4}{3} & \text{si } i = j = m + 1, \dots, n, \\ 0 & \text{ailleurs,} \end{cases}$$

et

$$y^0[i] = \frac{-7}{12}, \quad i = 1, \dots, m.$$

Pour cet exemple, on prend  $\mu^{(0)} = \frac{X^0 \bullet Z^0}{n} = 1.0208$  et  $\epsilon = 10^{-5}$  avec  $\Psi(X^0, Z^0; \mu^{(0)}) \leq \tau$ , où  $\tau = n$  et  $\delta(X^0, Z^0; \mu^{(0)}) \leq \frac{1}{\sqrt{2}}$ .

Une solution optimale est donnée par :

$$X^*[i, j] = \begin{cases} 1.5 & \text{si } i = j = 1, \dots, m, \\ 0.5 & \text{si } i = j = m + 1, \dots, n, \\ 0 & \text{ailleurs,} \end{cases}$$

$$Z^* = 0_{n \times n} \text{ et } y^*[i] = 0.5, \quad i = 1, \dots, m.$$

La valeur optimale pour les deux problèmes est égale à  $\frac{-m}{4} = \frac{-n}{8}$ .

**4.3.2 Problèmes à taille fixe****Problème 1.**

On considère le problème SDP présenté en [39], avec  $m = 3$ ,  $n = 5$ ,  $\mathcal{Q}(X) = 0_{n \times n}$ ,

$$b = [-2, 2, -2]^\top,$$

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -2 & -1 \\ 0 & -1 & 1 & -1 & -2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & -2 & 2 & 0 \\ 0 & 2 & 1 & 0 & 2 \\ -2 & 1 & -2 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 2 \end{bmatrix}$$

$$A_3 = \begin{bmatrix} 2 & 2 & -1 & -1 & 1 \\ 2 & 0 & 2 & 1 & 1 \\ -1 & 2 & 0 & 1 & 0 \\ -1 & 1 & 1 & -2 & 0 \\ 1 & 1 & 0 & 0 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 3 & -3 & 1 & 1 \\ 3 & 5 & 3 & 1 & 2 \\ -3 & 3 & -1 & 1 & 2 \\ 1 & 1 & 1 & -3 & -1 \\ 1 & 2 & 2 & -1 & -1 \end{bmatrix}.$$

On prend

$$X^0 = Z^0 = I \text{ et } y^0 = [1, 1, 1]^\top$$

comme un point réalisable initial. Aussi, on prend  $\mu^{(0)} = 1$  et  $\epsilon = 10^{-8}$  avec  $\Psi(X^0, Z^0; \mu^{(0)}) = 0 \leq \tau$ , où  $\tau = 3$  et  $\delta(X^0, Z^0; \mu^{(0)}) = 0 \leq \frac{1}{\sqrt{2}}$ .

Une solution optimale approximative est donnée par :

$$X^* = \begin{bmatrix} 0.0914 & -0.0718 & 0.0169 & 0.0649 & -0.1583 \\ -0.0718 & 0.0724 & -0.0183 & -0.0602 & 0.1676 \\ 0.0169 & -0.0183 & 0.0103 & -0.0084 & -0.0772 \\ 0.0649 & -0.0602 & -0.0084 & 0.1481 & 0.0056 \\ -0.1583 & 0.1676 & -0.0772 & 0.0056 & 0.6022 \end{bmatrix},$$

$$Z^* = \begin{bmatrix} 1.4338 & 0.5754 & -0.0295 & -0.4043 & 0.2169 \\ 0.5754 & 1.0965 & 0.3401 & 0.2169 & -0.1120 \\ -0.0295 & 0.3401 & 1.1874 & 0.2169 & 0.0478 \\ -0.4043 & 0.2169 & 0.2169 & 0.2831 & -0.1415 \\ 0.2169 & -0.1120 & 0.0478 & -0.1415 & 0.0957 \end{bmatrix}$$

et

$$y^* = \begin{bmatrix} 0.8585 & 1.0937 & 0.7831 \end{bmatrix}^\top,$$

la valeur optimale pour les deux problèmes est égale à  $-1.0957$ .

**Problème 2.**

Soient  $m = 3$ ,  $n = 4$ ,  $\mathcal{Q}(X) = X$ ,  $b = [-2, 2, 0]^\top$ ,

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 2 & 0 & -1 \\ 0 & 0 & -2 & -1 \\ 0 & -1 & -1 & -2 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 & -2 & 0 \\ 0 & 2 & 1 & 2 \\ -2 & 1 & -2 & 1 \\ 0 & 2 & 1 & 2 \end{bmatrix}, A_3 = \begin{bmatrix} 2 & 2 & -1 & -1 \\ 2 & 0 & 2 & 1 \\ -1 & 2 & 0 & 1 \\ -1 & 1 & 1 & -2 \end{bmatrix},$$

et

$$C = \begin{bmatrix} 2 & 3 & -3 & -1 \\ 3 & 4 & 3 & 2 \\ -3 & 3 & -4 & 1 \\ -1 & 2 & 1 & -2 \end{bmatrix}.$$

On prend

$$X^0 = Z^0 = I \text{ et } y^0 = [1, 1, 1]^\top$$

comme un point réalisable initial. Pour cet exemple, on prend  $\mu^{(0)} = 1$  et  $\epsilon = 10^{-6}$  avec  $\Psi(X^0, Z^0; \mu^{(0)}) = 0 \leq \tau$ , où  $\tau = 3$  et  $\delta(X^0, Z^0; \mu^{(0)}) = 0 \leq \frac{1}{\sqrt{2}}$ .

Une solution optimale approximative est donnée par :

$$X^* = \begin{bmatrix} 0.0574 & -0.0368 & -0.0554 & -0.0304 \\ -0.0368 & 0.0648 & 0.0536 & 0.1540 \\ -0.0554 & 0.0536 & 0.2056 & 0.1688 \\ -0.0304 & 0.1540 & 0.1688 & 0.4996 \end{bmatrix},$$

$$Z^* = \begin{bmatrix} 0.1081 & 0.1681 & 0.0311 & -0.0557 \\ 0.1681 & 0.2615 & 0.0483 & -0.0867 \\ 0.0311 & 0.0483 & 0.0089 & -0.0160 \\ -0.0557 & -0.0867 & -0.0160 & 0.0287 \end{bmatrix},$$

et

$$y^* = \begin{bmatrix} 0.8458 & 1.0559 & 0.9747 \end{bmatrix}^\top,$$

la valeur optimale pour les deux problèmes est égale à 0.2101.

**Problème 3.**

Soient  $m = 4$ ,  $n = 8$ ,  $\mathcal{Q}(X) = X$ ,  $b = [8, 4, 12, 8]^\top$ ,

$$\begin{aligned}
 A_1 &= \begin{bmatrix} 2 & 1 & 0 & 1 & 1 & 1 & 2 & 1 \\ 1 & 0 & 0 & 2 & 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 2 & 1 & 0 \\ 1 & 2 & 1 & 0 & 1 & 1 & 2 & 2 \\ 1 & 1 & 0 & 1 & 2 & 0 & 1 & 1 \\ 1 & 2 & 2 & 1 & 0 & 0 & 2 & 1 \\ 2 & 1 & 1 & 2 & 1 & 2 & 2 & 2 \\ 1 & 0 & 0 & 2 & 1 & 1 & 2 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 & 1 & 2 & 2 & 0 & 1 & 1 \\ 1 & 0 & 2 & 1 & 2 & 0 & 1 & 1 \\ 1 & 2 & 1 & 2 & 2 & 1 & 1 & 1 \\ 2 & 1 & 2 & 0 & 1 & 1 & 0 & 1 \\ 2 & 2 & 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 2 & 2 \\ 1 & 1 & 1 & 0 & 0 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 & 0 & 2 & 1 & 0 \end{bmatrix}, \\
 A_3 &= \begin{bmatrix} 2 & 1 & 2 & 1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 2 & 1 & 2 & 1 & 1 & 1 \\ 2 & 2 & 2 & 0 & 1 & 1 & 1 & 2 \\ 1 & 1 & 0 & 0 & 1 & 2 & 0 & 1 \\ 1 & 2 & 1 & 1 & 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 2 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 2 & 1 & 2 & 1 \\ 0 & 1 & 2 & 1 & 0 & 1 & 1 & 2 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 2 & 1 & 1 & 2 & 1 & 0 & 1 & 1 \\ 1 & 2 & 2 & 1 & 1 & 1 & 0 & 2 \\ 1 & 2 & 1 & 0 & 1 & 2 & 2 & 2 \\ 2 & 1 & 0 & 0 & 2 & 1 & 1 & 0 \\ 1 & 1 & 1 & 2 & 0 & 1 & 0 & 1 \\ 0 & 1 & 2 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 & 0 & 1 & 2 & 2 \\ 1 & 2 & 2 & 0 & 1 & 0 & 2 & 0 \end{bmatrix},
 \end{aligned}$$

et

$$C = \begin{bmatrix} 6 & 4 & 4 & 6 & 5 & 1 & 5 & 3 \\ 4 & 4 & 6 & 5 & 6 & 4 & 3 & 4 \\ 4 & 6 & 6 & 3 & 4 & 6 & 5 & 5 \\ 6 & 5 & 3 & 0 & 5 & 5 & 3 & 4 \\ 5 & 6 & 4 & 5 & 4 & 2 & 3 & 2 \\ 1 & 4 & 6 & 5 & 2 & 2 & 6 & 4 \\ 5 & 3 & 5 & 3 & 3 & 6 & 8 & 6 \\ 3 & 4 & 5 & 4 & 2 & 4 & 6 & 2 \end{bmatrix}.$$

On prend

$$X^0 = Z^0 = I \text{ et } y^0 = [1, 1, 1]^\top$$

comme un point réalisable initial. Pour cet exemple, on prend  $\mu^{(0)} = 1$  et  $\epsilon = 10^{-6}$  avec  $\Psi(X^0, Z^0; \mu^{(0)}) = 0 \leq \tau$ , où  $\tau = 5$  et  $\delta(X^0, Z^0; \mu^{(0)}) = 0 \leq \frac{1}{\sqrt{2}}$ .



Une solution optimale approximative est donnée par :

$$X^* = \begin{bmatrix} 0.5143 & 0.0935 & 0.2653 & -0.0822 & -0.0149 & 0.0667 & 0.1155 & -0.0981 \\ 0.0935 & 0.444 & 0.1279 & 0.1464 & 0.1198 & 0.2706 & 0.1422 & 0.0587 \\ 0.2653 & 0.1279 & 0.4490 & -0.1267 & -0.0689 & 0.0553 & 0.0632 & 0.2285 \\ -0.0822 & 0.1464 & -0.1267 & 0.2143 & 0.0729 & 0.1745 & 0.0587 & 0.0516 \\ -0.0149 & 0.1198 & -0.0689 & 0.0729 & 0.3292 & -0.0809 & 0.3082 & 0.0623 \\ 0.0667 & 0.2706 & 0.0553 & 0.1745 & -0.0809 & 0.3255 & -0.0461 & 0.0138 \\ 0.1155 & 0.1422 & 0.0632 & 0.0587 & 0.3082 & -0.0461 & 0.3455 & 0.1314 \\ -0.0981 & 0.0587 & 0.2285 & 0.0516 & 0.0623 & 0.0138 & 0.1314 & 0.4684 \end{bmatrix},$$

$$Z^* = \begin{bmatrix} 0.0098 & -0.0085 & -0.0164 & -0.0338 & 0.0334 & 0.0304 & -0.0228 & 0.0159 \\ -0.0085 & 0.0121 & -0.034 & 0.0081 & -0.0480 & -0.0180 & 0.0401 & -0.0069 \\ -0.0164 & -0.034 & 0.0946 & 0.1375 & 0.0157 & -0.0829 & -0.0387 & -0.0531 \\ -0.0338 & 0.0081 & 0.1375 & 0.2143 & -0.0290 & -0.1435 & -0.0138 & -0.0868 \\ 0.0334 & -0.0480 & 0.0157 & -0.0290 & 0.1908 & 0.0695 & -0.1602 & 0.0261 \\ 0.0304 & -0.0180 & -0.0829 & -0.1435 & 0.0695 & 0.1096 & -0.0342 & 0.0620 \\ -0.0228 & 0.0401 & -0.0387 & -0.0138 & -0.1602 & -0.0342 & 0.1415 & -0.0068 \\ 0.0159 & -0.0069 & -0.0531 & -0.0868 & 0.0261 & 0.0620 & -0.0068 & 0.0363 \end{bmatrix}$$

et

$$y^* = [1.0363, 0.8497, 1.2161, 0.9999]^\top,$$

la valeur optimale pour les deux problèmes est égale à 33.1406.

### 4.3.3 Problème de $(\mathcal{NCM})$

On rappelle que le problème  $(\mathcal{NCM})$  est donné par :

$$(\mathcal{NCM}) \quad \min_X \left\{ \frac{1}{2} \|\mathcal{L}(X - K)\|_F^2 : \text{diag}(X) = e, X \in \mathbb{S}_+^n \right\},$$

où  $K \in \mathbb{S}^n$  et  $\mathcal{L}$  est un opérateur linéaire auto-adjoint ( $\mathcal{L} : \mathbb{S}^n \rightarrow \mathbb{S}^n$ ).

Soient  $m = n = 3$  et

$$K = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

Le problème CQSDP résulte du problème  $(\mathcal{NCM})$  avec  $C = -\mathcal{L}^2(K) = -\mathcal{L}(\mathcal{L}(K))$ , par

conséquent  $\mathcal{Q}(X) = \mathcal{L}^2(X)$ ,  $b = [1, 1, 1]^\top$  et

$$A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Donc, selon les choix de l'opérateur  $\mathcal{L}$  on obtient :

1. Si  $\mathcal{L}(X) = X$ , de même  $\mathcal{Q}(X) = X$ . Alors,

$$C = -K = \begin{bmatrix} 0 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & 0 \end{bmatrix},$$

la solution initiale est :

$$X^0 = I, Z^0 = \begin{bmatrix} 2.5 & 1 & 1 \\ 1 & 2.5 & 1 \\ 1 & 1 & 2.5 \end{bmatrix}$$

et

$$y^0 = [-1.5, -0.5, -1.5]^\top.$$

Pour cet exemple, on prend  $\mu^{(0)} = \frac{X^0 \bullet Z^0}{n}$  et  $\epsilon = 10^{-6}$  avec  $\Psi(X^0, Z^0; \mu^{(0)}) \leq \tau$ , où  $\tau = \sqrt{3} \simeq 1.7321$  et  $\delta(X^0, Z^0; \mu^{(0)}) \leq \frac{1}{\sqrt{2}}$ .

Une solution optimale approximative est donnée par :

$$X^* = \begin{bmatrix} 1 & -0.5 & -0.5 \\ -0.5 & 1 & -0.5 \\ -0.5 & -0.5 & 1 \end{bmatrix},$$

$$Z^* = \begin{bmatrix} 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0.5 \end{bmatrix}$$

et

$$y^* = [0.5, 1.5, 0.5]^\top,$$

la valeur optimale pour les deux problèmes est égale à 3.75.

2. Si  $\mathcal{L}(X) = U^{\frac{1}{2}}XU^{\frac{1}{2}}$ , de même  $\mathcal{Q}(X) = UXU$  avec

$$U = \begin{bmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 2 \end{bmatrix} \in \mathbb{S}_{+++}^n.$$

Alors,

$$C = -\mathcal{L}^2(K) = -\mathcal{Q}(K) = -UKU = \begin{bmatrix} -4 & 1 & 5 \\ 1 & 1 & 1 \\ 5 & 1 & -4 \end{bmatrix},$$

la solution initiale est :

$$X^0 = I, Z^0 = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 3 \end{bmatrix}$$

et

$$y^0 = [-2, -1, -2]^{\top}.$$

Pour cet exemple, on prend  $\mu^{(0)} = \frac{X^0 \bullet Z^0}{n}$  et  $\epsilon = 10^{-6}$  avec  $\Psi(X^0, Z^0; \mu^{(0)}) \leq \tau$ , où  $\tau = \sqrt{3} \simeq 1.7321$  et  $\delta(X^0, Z^0; \mu^{(0)}) \leq \frac{1}{\sqrt{2}}$ .

Une solution optimale approximative est donnée par :

$$X^* = \begin{bmatrix} 1 & -0.6531 & -0.1469 \\ -0.6531 & 1 & -0.6531 \\ -0.1469 & -0.6531 & 1 \end{bmatrix},$$

$$Z^* = \begin{bmatrix} 0.2656 & 0.3469 & 0.2656 \\ 0.3469 & 0.4530 & 0.3469 \\ 0.2656 & 0.3469 & 0.2656 \end{bmatrix}$$

et

$$y^* = [1.3219, 1.5470, 1.3219]^{\top},$$

la valeur optimale pour les deux problèmes est égale à 4.0548.

3. Si  $\mathcal{L}(X) = \Sigma \circ X$  où

$$\Sigma = \begin{bmatrix} 1 & \sqrt{2} & \sqrt{3} \\ \sqrt{2} & \sqrt{3} & 1 \\ \sqrt{3} & 1 & \sqrt{2} \end{bmatrix} \in \mathbb{S}^n \cap \mathbb{R}_+^{n \times n},$$

de même  $\mathcal{Q}(X) = U \circ X$  avec

$$U = \Sigma \circ \Sigma = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}.$$

Alors,

$$C = -\mathcal{L}^2(K) = -\mathcal{Q}(K) = -U \circ K = \begin{bmatrix} 0 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 0 \end{bmatrix},$$

la solution initiale est :

$$X^0 = \begin{bmatrix} 1 & 0 & -0.5 \\ 0 & 1 & 0 \\ -0.5 & 0 & 1 \end{bmatrix}, Z^0 = \begin{bmatrix} 4 & 2 & 1.5 \\ 2 & 4 & 1 \\ 1.5 & 1 & 3 \end{bmatrix}$$

et

$$y^0 = [-3, 2, -1]^\top.$$

Pour cet exemple, on prend  $\mu^{(0)} = \frac{X^0 \bullet Z^0}{n}$  et  $\epsilon = 10^{-6}$  avec  $\Psi(X^0, Z^0; \mu^{(0)}) \leq \tau$ , où  $\tau = \sqrt{3} \simeq 1.7321$  et  $\delta(X^0, Z^0; \mu^{(0)}) \leq \frac{1}{\sqrt{2}}$ .

Une solution optimale approximative est donnée par :

$$X^* = \begin{bmatrix} 1 & -0.5574 & -0.6701 \\ -0.5574 & 1 & -0.2427 \\ -0.6701 & -0.2427 & 1 \end{bmatrix},$$

$$Z^* = \begin{bmatrix} 1.1567 & 0.8851 & 0.9897 \\ 0.8851 & 0.6773 & 0.7573 \\ 0.9897 & 0.7573 & 0.8467 \end{bmatrix}$$

et

$$y^* = [-0.1567, 5.3227, 1.1533]^\top,$$

la valeur optimale pour les deux problèmes est égale à 8.7917.

## 4.4 Résultats numériques

Dans cette section, on présente les différents résultats numériques obtenus par l'application des Algorithmes 2.4.4 et 3.2.1 sur les problèmes quadratiques convexes semi-définis

mentionnés ci-dessus. Notons que :

- Pour le petit pas : on utilise l’Algorithme 2.4.4 avec le choix théorique de  $\theta$ , c’est à dire  $\theta = \Theta(\frac{1}{\sqrt{n}}) = \frac{1}{2\sqrt{n}}$ .
- Pour le grand pas : on utilise l’Algorithme 3.2.1 avec  $\theta = O(1)$  et  $\tau = \Theta(n)$ , on présente deux études comparatives basées sur les fonctions noyaux :
  - La première est basée sur la fonction noyau présentée en (3.5), avec

$$q \in \{1, 1.5, 2, 2.5, 3, 3.5, \log(\frac{4}{3}(1+n))\},$$

c’est à dire comparaison entre  $\psi_{Zhang}(t)$  et  $\psi(t)$  avec le choix théorique de  $q$ .

- La deuxième est entre quelques fonctions noyaux paramétrées, durant l’implémentation numérique on prend le choix théorique du paramètre associé à chaque fonction noyau dont la complexité pour ce type d’algorithme à grand pas est d’ordre  $O(\sqrt{n} \log n \log \frac{n}{\epsilon})$ .

#### 4.4.1 Résultats numériques pour l’Algorithme 2.4.4 à petit pas

Dans cette partie notre implémentation numérique est basée sur l’Algorithme 2.4.4 à petit pas, c’est à dire  $\theta = \Theta(\frac{1}{\sqrt{n}}) = \frac{1}{2\sqrt{n}}$ .

##### Problèmes à taille variable

###### Problème 1.

Le tableau suivant présente les résultats numériques obtenus pour les différentes tailles testées.

$(m, n)$	$\delta(X^0, Z^0; \mu^{(0)})$	iter	CPU
(5, 10)	$1.1102e^{-15}$	82	0.1320
(25, 50)	$2.4825e^{-15}$	212	0.9342
(50, 100)	$3.5108e^{-15}$	316	3.68
(100, 200)	$4.9651e^{-15}$	469	18.3968
(200, 400)	$7.0217e^{-15}$	693	143.9502

Tableau 4. Résultats numériques pour l’Algorithme 2.4.4 à petit pas.

###### Problème 2.

Le tableau suivant présente les résultats numériques obtenus pour les différentes tailles testées.

$(m, n)$	$\delta(X^0, Z^0; \mu^{(0)})$	iter	CPU
(5, 10)	0.0323	82	0.1579
(25, 50)	0.0722	212	1.0501
(50, 100)	0.1021	316	4.1027
(100, 200)	0.1443	469	18.7305
(200, 400)	0.2041	694	219.7951

Tableau 5. Résultats numériques pour l’Algorithme 2.4.4 à petit pas.

### Problèmes à taille fixe

Le tableau suivant présente les résultats numériques obtenus pour les trois problèmes.

Problème	iter	CPU
<b>Problème 1</b>	81	0.3743
<b>Problème 2</b>	54	0.2034
<b>Problème 3</b>	83	0.7253

Tableau 6. Résultats numériques pour l’Algorithme 2.4.4 à petit pas.

**Problème de  $(\mathcal{NCM})$**  Le tableau suivant présente les résultats numériques obtenus pour le problème  $(\mathcal{NCM})$  selon les différents choix de l’opérateur  $\mathcal{L}(X)$ .

$\mathcal{L}(X)$	$\delta(X^0, Z^0; \mu^{(0)})$	iter	CPU
$X$	0.4714	48	0.2194
$U^{\frac{1}{2}} X U^{\frac{1}{2}}$	0.3873	49	0.1784
$\Sigma \circ X$	0.4558	49	0.1854

Tableau 7. Résultats numériques pour l’Algorithme 2.4.4 à petit pas.

### 4.4.2 Résultats comparatifs pour l’Algorithme 3.2.1 à grand pas

Par la suite, on s’intéresse seulement aux algorithmes à grand pas basés sur les fonctions noyaux. Autrement dit, on utilise l’Algorithme 3.2.1 avec  $\theta = O(1)$  et  $\tau = \Theta(n)$ . Les différentes valeurs des paramètres  $q$  et  $\theta$  seront présentées pour montrer leur influence sur le nombre d’itérations produites par l’Algorithme 3.2.1. On distingue deux types de comparaison :

a) Comparaison basée sur la fonction noyau paramétrée de M.W. Zhang avec les différents choix de  $q$

On considère la fonction noyau mentionnée en (3.5) qui dépend d'un paramètre  $q$ . Selon les choix de ce paramètre, on établit numériquement l'influence de ce dernier sur le comportement de l'Algorithme 3.2.1 dont le paramètre  $\theta$  utilisé durant l'implémentation est comme suit :

$$\theta \in \{0.1, 0.5, 0.9\}.$$

**Problèmes à taille variable Problème 1.**

Les tableaux suivants présentent les résultats numériques pour différentes tailles du problème, où le paramètre  $q$  utilisé durant l'implémentation est comme suit :

$$q \in \{1, \log(\frac{4}{3}(1+n))\},$$

$(m, n) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
(5, 10)	14	133	0.2182	13	21	0.0522	12	7	0.0345
(25, 50)	16	148	0.8137	15	24	0.2248	14	8	0.1109
(50, 100)	17	154	2.8105	16	25	0.8785	14	8	0.5819
(100, 200)	18	161	12.3929	17	26	2.7772	16	9	1.9003
(200, 400)	19	168	54.9642	18	27	19.9239	16	9	13.0811
(400, 800)	<b>19</b>	174	<b>457.5409</b>	18	28	181.5622	16	9	118.6372

Tableau 8.1. Résultats numériques obtenus pour les différents choix de  $\theta$  avec  $q = 1$  (Zhang).

$(m, n) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
(5, 10)	<b>11</b>	133	<b>0.1721</b>	<b>11</b>	21	<b>0.0439</b>	<b>11</b>	7	<b>0.0293</b>
(25, 50)	<b>12</b>	148	<b>0.7800</b>	<b>12</b>	24	<b>0.2037</b>	<b>12</b>	8	<b>0.1059</b>
(50, 100)	<b>12</b>	154	<b>2.8015</b>	<b>12</b>	25	<b>0.6098</b>	<b>12</b>	8	<b>0.3669</b>
(100, 200)	<b>13</b>	161	<b>9.9347</b>	<b>13</b>	26	<b>2.2984</b>	<b>14</b>	9	<b>1.6556</b>
(200, 400)	<b>13</b>	168	<b>51.6072</b>	<b>14</b>	27	<b>18.8091</b>	<b>14</b>	9	<b>12.7826</b>
(400, 800)	20	174	629.2981	<b>14</b>	28	<b>155.4118</b>	<b>14</b>	9	<b>103.6936</b>

Tableau 8.2. Résultats numériques obtenus pour les différents choix de  $\theta$  avec  $q = \log(\frac{4}{3}(1+n))$ .

**Problème 2.**

Les tableaux suivants présentent les résultats numériques pour différentes tailles du problème, où le paramètre  $q$  utilisé durant l'implémentation est comme suit :

$$q \in \{1, \log(\frac{4}{3}(1+n))\},$$

$(m, n) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
(5, 10)	15	133	0.2922	13	21	0.0466	14	8	0.0360
(25, 50)	16	148	0.8353	16	24	0.2146	14	8	0.1366
(50, 100)	17	155	3.5089	16	25	1.1101	16	9	0.8455
(100, 200)	18	161	9.8203	17	26	2.4024	16	9	1.7221
(200, 400)	19	168	54.1711	18	27	16.6635	16	9	10.9580
(400, 800)	19	174	470.2712	18	28	141.5589	16	9	92.1385

Tableau 9.1. Résultats numériques obtenus pour les différents choix de  $\theta$  avec  $q = 1$  (Zhang).

$(m, n) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
(5, 10)	<b>11</b>	133	<b>0.1936</b>	<b>11</b>	21	<b>0.0441</b>	<b>12</b>	8	<b>0.0308</b>
(25, 50)	<b>12</b>	148	<b>0.7421</b>	<b>12</b>	24	<b>0.1851</b>	<b>12</b>	8	<b>0.1273</b>
(50, 100)	<b>12</b>	155	<b>2.8834</b>	<b>12</b>	25	<b>0.8091</b>	<b>14</b>	9	<b>0.6128</b>
(100, 200)	<b>13</b>	161	<b>8.8553</b>	<b>13</b>	26	<b>2.2645</b>	<b>14</b>	9	<b>1.6347</b>
(200, 400)	<b>13</b>	168	<b>52.7976</b>	<b>13</b>	27	<b>14.3179</b>	<b>14</b>	9	<b>10.1711</b>
(400, 800)	<b>14</b>	174	<b>451.1436</b>	<b>14</b>	28	<b>124.0735</b>	<b>14</b>	9	<b>82.7292</b>

Tableau 9.2. Résultats numériques obtenus pour les différents choix de  $\theta$  avec  $q = \log(\frac{4}{3}(1+n))$ .

Le tableau suivant présente la valeur de  $\Psi(X^0, Z^0; \mu^{(0)})$  selon le choix de  $q$  et pour diffé-



rentes tailles du problème.

$(m, n) \setminus q$	$\Psi(X^0, Z^0; \mu^{(0)})$	
	1	$\log(\frac{4}{3}(1+n))$
(5, 10)	0.0016	0.003
(25, 50)	0.0078	0.0192
(50, 100)	0.0156	0.0419
(100, 200)	0.0313	0.0910
(200, 400)	0.0625	0.1962
(400, 800)	0.1250	0.4210

Tableau 9.3. La valeur de  $\Psi(X^0, Z^0; \mu^{(0)})$  selon le choix de  $q$  et pour différentes tailles.

### Problèmes à taille fixe

#### Problème 1.

Les résultats numériques sont présentés dans le tableau suivant, où le paramètre  $q$  utilisé durant l'implémentation est comme suit :

$$q \in \{1, 1.5, 3, \log(\frac{4}{3}(1+n))\},$$

$q \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
1 (Zhang)	20	192	0.5087	18	30	0.0852	17	10	0.0682
1.5	16	192	0.3350	15	30	0.0768	15	10	0.0596
$\log(\frac{4}{3}(1+n))$	<b>15</b>	192	<b>0.3040</b>	<b>15</b>	30	<b>0.0722</b>	<b>15</b>	10	<b>0.0539</b>
3	39	192	0.3990	24	30	0.1000	17	10	0.0606

Tableau 10.1. Résultats numériques obtenus pour les différents choix de  $q$  et  $\theta$ .

#### Problème 2.

Le tableau suivant contient les résultats numériques obtenus, dont le paramètre  $q$  utilisé durant l'implémentation est comme suit :

$$q \in \{1, 1.5, 3, \log(\frac{4}{3}(1+n))\},$$

$q \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
1 (Zhang)	12	146	0.2316	12	23	0.0515	11	8	0.0377
1.5	11	146	0.2275	11	23	0.0463	11	8	0.0402
$\log(\frac{4}{3}(1+n))$	<b>10</b>	146	<b>0.2097</b>	<b>10</b>	23	<b>0.0397</b>	<b>10</b>	8	<b>0.0340</b>
3	22	146	0.2633	10	23	0.0432	10	8	0.0341

Tableau 10.2. Résultats numériques obtenus pour les différents choix de  $q$  et  $\theta$ .

### Problème 3.

On présente les résultats numériques dans le tableau suivant, où le paramètre  $q$  utilisé durant l'implémentation est comme suit :

$$q \in \{1, 1.5, 2, 3, 3.5, \log(\frac{4}{3}(1+n))\},$$

$q \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
1 (Zhang)	16	152	0.6044	15	24	0.2271	13	8	0.1634
1.5	14	152	0.5814	13	24	0.1659	12	8	0.1298
2	13	152	0.4531	12	24	0.1421	12	8	0.1264
$\log(\frac{4}{3}(1+n))$	<b>12</b>	152	<b>0.4334</b>	<b>12</b>	24	<b>0.1407</b>	<b>12</b>	8	<b>0.1090</b>
3	12	152	0.4512	13	24	0.1515	12	8	0.1104
3.5	14	152	0.4893	13	24	0.1523	12	8	0.1228

Tableau 10.3. Résultats numériques obtenus pour les différents choix de  $q$  et  $\theta$ .

### Problème de $(\mathcal{NCM})$

1. Soit  $\mathcal{L}(X) = X$ . On présente les résultats numériques dans le tableau ci-dessous, où le paramètre  $q$  utilisé durant l'implémentation est comme suit :

$$q \in \{1, 2, 2.5, 3, \log(\frac{4}{3}(1+n))\},$$

$q \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
1 (Zhang)	16	152	0.2256	15	24	0.0519	13	8	0.0405
$\log(\frac{4}{3}(1+n))$	<b>12</b>	152	<b>0.2028</b>	<b>14</b>	24	<b>0.0466</b>	<b>11</b>	8	<b>0.0251</b>
2	15	152	0.2179	14	24	0.0502	11	8	0.0305
2.5	17	152	0.2297	19	24	0.0600	11	8	0.0297
3	24	152	0.2318	43	24	0.0959	11	8	0.0255

Tableau 11.1. Résultats numériques obtenus pour les différents choix de  $q$  et  $\theta$ .

2. Soit  $\mathcal{L}(X) = U^{\frac{1}{2}} X U^{\frac{1}{2}}$  où  $U \in \mathbb{S}_{++}^n$ . On présente les résultats numériques dans le tableau suivant, où le paramètre  $q$  utilisé durant l'implémentation est comme suit :

$$q \in \{1, 2.5, 3.5, \log(\frac{4}{3}(1+n))\},$$

$q \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
1 (Zhang)	16	153	0.2195	15	25	0.0520	13	8	0.0383
$\log(\frac{4}{3}(1+n))$	<b>13</b>	153	<b>0.1952</b>	13	25	0.0424	<b>12</b>	8	<b>0.0362</b>
2.5	17	153	0.2479	<b>12</b>	25	<b>0.0406</b>	14	8	0.0402
3.5	109	153	0.3885	33	25	0.0775	27	8	0.0444

Tableau 11.2. Résultats numériques obtenus pour les différents choix de  $q$  et  $\theta$ .

3. Soit  $\mathcal{L}(X) = \Sigma \circ X$  pour une matrice donnée  $\Sigma \in \mathbb{S}^n \cap \mathbb{R}_+^{n \times n}$ . On présente les résultats numériques dans le tableau suivant, où le paramètre  $q$  utilisé durant l'implémentation est comme suit :

$$q \in \{1, 2.5, 3, 3.5, \log(\frac{4}{3}(1+n))\}.$$

$q \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
1 (Zhang)	16	154	0.2370	15	25	0.0496	13	8	0.0378
$\log(\frac{4}{3}(1+n))$	<b>13</b>	154	<b>0.2123</b>	<b>12</b>	25	<b>0.0375</b>	12	8	0.0254
2.5	25	154	0.2434	12	25	0.0387	<b>11</b>	8	<b>0.0244</b>
3	35	154	0.2630	43	25	0.0829	12	8	0.0304
3.5	168	154	0.3991	72	25	0.2111	48	8	0.0995

Tableau 11.3. Résultats numériques obtenus pour les différents choix de  $q$  et  $\theta$ .

Le tableau suivant présente la valeur de  $\Psi(X^0, Z^0; \mu^{(0)})$  selon le choix de  $q$  et l'opérateur  $\mathcal{L}(X)$ .

$q / \mathcal{L}(X)$	$\Psi(X^0, Z^0; \mu^{(0)})$		
	$X$	$U^{\frac{1}{2}} X U^{\frac{1}{2}}$	$\Sigma \circ X$
1(Zhang)	0.3382	0.2272	0.3189
$\log(\frac{4}{3}(1+n))$	0.5276	0.3472	0.5223
2	0.5939	/	/
2.5	0.6854	0.4408	0.7143
3	0.7754	/	0.8346
3.5	/	0.5425	0.9680

Tableau 11.4. La valeur de  $\Psi(X^0, Z^0; \mu^{(0)})$  selon le choix de  $q$  et l'opérateur  $\mathcal{L}(X)$ .

**Commentaire :**

Les résultats numériques obtenus par l'Algorithme 3.2.1 à grand pas montrent que le choix théorique du paramètre  $q$  joue un rôle important dans l'aspect théorique et numérique où le nombre d'itérations internes (ainsi le CPU) est meilleur que celui de M.W. Zhang ( $q = 1$ ). Ceci consolide et confirme notre objectif théorique d'améliorer la complexité algorithmique de l'Algorithme 3.2.1.

**b) Comparaison entre la fonction noyau paramétrée de M.W. Zhang et d'autre fonction noyau**

Afin de voir le comportement de la fonction mentionnée en (3.5), on va la comparer à d'autres fonctions noyaux paramétrées dont leur complexité pour les algorithmes à grand pas est la même selon un choix théorique convenable du paramètre associé. Le tableau suivant présente les fonctions noyaux testées :

La fonction noyau ( $\psi_i(t)$ )	Choix théorique du paramètre
$\psi_1(t) = \frac{t^2-1}{2} + \frac{q^{\frac{1}{t}-1}-1}{\log q}, q > 1,$	$q = 1 + O(n),$
$\psi_2(t) = \frac{t^2-1}{2} - \frac{t-q}{q^2-q+1} e^{q(\frac{1}{t}-1)} + \frac{1-q}{q^2-q+1}, q \geq 1,$	$q = \log\left(\frac{4}{3}(1+O(n))\right),$
$\psi_3(t) = \frac{t^{p+1}-1}{p+1} + \frac{t^{1-q}-1}{q-1}, p \in [0, 1], q > 1,$	$p = 1, q = \log(n),$
$\psi_4(t) = \frac{t^2-1}{2} - \int_1^t e^{q(\frac{1}{\xi}-1)} d\xi, q \geq 1,$	$q = \log(1+n),$

Tableau 12. Quatre fonctions noyaux paramétrées testées.

durant l'implémentation le paramètre  $\theta$  utilisé est comme suit :

$$\theta \in \{0.1, 0.5, 0.9\}.$$

Pour le paramètre  $q$  on prend le choix théorique associé à chaque fonction noyau comme il est mentionné dans le Tableau 12.

### Problèmes à taille variable

#### Problème 1.

Les tableaux suivants présentent les résultats numériques pour différentes tailles du problème.

$(5, 10) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	11	133	0.1764	11	21	<b>0.0438</b>	10	7	<b>0.0247</b>
$\psi_2(t)$	11	133	<b>0.1721</b>	11	21	0.0439	11	7	0.0293
$\psi_3(t)$	12	133	0.2213	11	21	0.0369	11	7	0.0254
$\psi_4(t)$	13	133	1.8325	12	21	0.4183	12	7	0.2245

$(25, 50) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	12	148	0.7161	12	24	<b>0.1729</b>	12	8	0.1187
$\psi_2(t)$	12	148	0.7800	12	24	0.2037	12	8	<b>0.1059</b>
$\psi_3(t)$	12	148	<b>0.6956</b>	12	24	0.1824	12	8	0.1242
$\psi_4(t)$	13	148	6.5637	13	24	0.9998	12	8	0.5362

$(50, 100) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	12	154	3.1240	12	25	0.8135	12	8	0.7208
$\psi_2(t)$	12	154	<b>2.8015</b>	12	25	<b>0.6098</b>	12	8	<b>0.3669</b>
$\psi_3(t)$	12	154	3.6692	12	25	0.9051	12	8	0.6368
$\psi_4(t)$	13	154	13.4153	13	25	2.6187	12	8	1.4102

$(100, 200) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	13	161	<b>8.8241</b>	13	26	2.6095	14	9	1.7900
$\psi_2(t)$	13	161	9.9347	13	26	2.2984	14	9	1.6556
$\psi_3(t)$	13	161	8.8987	13	26	<b>2.1948</b>	14	9	<b>1.5949</b>
$\psi_4(t)$	13	161	31.0667	13	26	7.0426	14	9	4.4507

$(200, 400) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	13	168	71.7334	14	27	<b>15.7112</b>	14	9	15.4713
$\psi_2(t)$	13	168	<b>51.6072</b>	14	27	18.8091	14	9	12.7826
$\psi_3(t)$	13	168	63.3051	13	27	18.1843	14	9	<b>12.6646</b>
$\psi_4(t)$	14	168	103.8902	14	27	24.0129	14	9	15.2260

$(400, 800) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	14	174	<b>477.9675</b>	14	28	<b>118.6905</b>	14	9	<b>78.5835</b>
$\psi_2(t)$	20	174	629.2981	14	28	155.4118	14	9	103.6936
$\psi_3(t)$	14	174	545.1800	14	28	147.7460	14	9	106.3961
$\psi_4(t)$	14	174	664.3248	14	28	179.9046	14	9	105.2804

Tableau 13.1. Résultats numériques obtenus pour les différents choix de  $\theta$ .

Le tableau suivant présente la valeur de  $\Psi(X^0, Z^0; \mu^{(0)})$  selon la fonction noyau  $\psi_i(t)$  et pour différentes tailles du problème.

$\psi_i(t) / (m, n)$	$\Psi(X^0, Z^0; \mu^{(0)})$					
	(5, 10)	(25, 50)	(50, 100)	(100, 200)	(200, 400)	(400, 800)
$\psi_1(t)$	$1.4182e^{-16}$	$2.8875e^{-16}$	$8.3081e^{-16}$	$3.7718e^{-16}$	$1.3422e^{-16}$	$7.6496e^{-16}$
$\psi_2(t)$	0	$6.9389e^{-16}$	0	$2.7756e^{-15}$	0	0
$\psi_3(t)$	$7.8622e^{-17}$	$5.0312e^{-16}$	$9.7162e^{-16}$	$5.4266e^{-16}$	$2.2782e^{-16}$	$8.4113e^{-16}$
$\psi_4(t)$	$1.3559e^{-30}$	$1.1093e^{-29}$	$2.2187e^{-29}$	$5.4234e^{-29}$	$1.2819e^{-28}$	$2.3666e^{-28}$

Tableau 13.2. La valeur de  $\Psi(X^0, Z^0; \mu^{(0)})$  selon la fonction noyau  $\psi_i(t)$  et pour différentes tailles.

### Problème 2.

Les tableaux suivants présentent les résultats numériques pour différentes tailles du problème.

$(5, 10) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	11	133	<b>0.1877</b>	11	21	<b>0.0383</b>	12	8	<b>0.0304</b>
$\psi_2(t)$	11	133	0.1936	11	21	0.0441	12	8	0.0308
$\psi_3(t)$	12	133	0.1933	11	21	0.0461	12	8	0.0361
$\psi_4(t)$	13	133	2.0482	12	21	0.4305	13	8	0.2651

$(25, 50) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	12	148	0.8308	12	24	0.1910	12	8	0.1325
$\psi_2(t)$	12	148	0.7421	12	24	0.1851	12	8	0.1273
$\psi_3(t)$	12	148	<b>0.7411</b>	12	24	<b>0.1834</b>	12	8	<b>0.1148</b>
$\psi_4(t)$	13	148	5.8598	13	24	0.9495	12	8	0.5186

$(50, 100) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	12	155	3.5288	12	25	0.8171	14	9	<b>0.5589</b>
$\psi_2(t)$	12	155	<b>2.8834</b>	12	25	<b>0.8091</b>	14	9	0.6128
$\psi_3(t)$	12	155	3.1932	12	25	0.8436	14	9	0.7562
$\psi_4(t)$	13	155	12.6597	13	25	2.7492	14	9	1.6891

$(100, 200) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	13	161	<b>8.6817</b>	13	26	2.5903	14	9	1.8763
$\psi_2(t)$	13	161	8.8553	13	26	2.2645	14	9	<b>1.6347</b>
$\psi_3(t)$	13	161	8.9051	13	26	<b>2.2401</b>	14	9	1.6693
$\psi_4(t)$	13	161	28.4137	13	26	5.8761	14	9	3.7466

$(200, 400) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	13	168	70.2503	13	27	20.5115	14	9	11.8764
$\psi_2(t)$	13	168	<b>52.7976</b>	13	27	<b>14.3179</b>	14	9	<b>10.1711</b>
$\psi_3(t)$	13	168	69.8250	13	27	16.2215	14	9	13.2611
$\psi_4(t)$	14	168	99.3479	14	27	28.2621	14	9	14.2496

$(400, 800) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	14	174	564.2270	14	28	146.0199	14	9	96.5220
$\psi_2(t)$	14	174	<b>451.1436</b>	14	28	<b>124.0735</b>	14	9	<b>82.7292</b>
$\psi_3(t)$	14	174	478.8081	14	28	133.1471	14	9	87.5587
$\psi_4(t)$	14	174	632.7938	14	28	165.5416	14	9	102.2211

Tableau 14.1. Résultats numériques obtenus pour les différents choix de  $\theta$ .

Le tableau suivant présente la valeur de  $\Psi(X^0, Z^0; \mu^{(0)})$  selon la fonction noyau  $\psi_i(t)$  et pour différentes tailles du problème.

$\psi_i(t) / (m, n)$	$\Psi(X^0, Z^0; \mu^{(0)})$					
	(5, 10)	(25, 50)	(50, 100)	(100, 200)	(200, 400)	(400, 800)
$\psi_1(t)$	0.0028	0.0181	0.0397	0.0866	0.1876	0.4041
$\psi_2(t)$	0.003	0.0192	0.0419	0.0910	0.1962	0.4210
$\psi_3(t)$	0.0017	0.0128	0.0292	0.0656	0.1457	0.3204
$\psi_4(t)$	0.0018	0.0128	0.0293	0.0657	0.1458	0.3205

Tableau 14.2. La valeur de  $\Psi(X^0, Z^0; \mu^{(0)})$  selon la fonction noyau  $\psi_i(t)$  et pour différentes tailles.

### Problèmes à taille fixe Problème 1.

On présente les résultats numériques dans le tableau suivant :

$\psi_i(t) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	15	192	0.4258	15	30	<b>0.0667</b>	15	10	0.0576
$\psi_2(t)$	15	192	<b>0.3040</b>	15	30	0.0722	15	10	<b>0.0539</b>
$\psi_3(t)$	18	192	0.5612	15	30	0.0823	15	10	0.0748
$\psi_4(t)$	19	192	1.5495	17	30	0.3293	16	10	0.1650

Tableau 15.1. Résultats numériques obtenus pour les différents choix de  $\theta$ .

### Problème 2.

On présente les résultats numériques dans le tableau ci-dessous.

$\psi_i(t) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	10	146	0.2218	10	23	0.0401	10	8	0.0341
$\psi_2(t)$	10	146	<b>0.2097</b>	10	23	<b>0.0397</b>	10	8	<b>0.0340</b>
$\psi_3(t)$	12	146	0.2692	10	23	0.0497	11	8	0.0437
$\psi_4(t)$	12	146	1.0560	11	23	0.2064	11	8	0.1170

Tableau 15.2. Résultats numériques obtenus pour les différents choix de  $\theta$ .



### Problème 3.

On présente les résultats numériques dans le tableau suivant :

$\psi_i(t) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	12	152	0.4385	12	24	<b>0.1292</b>	12	8	0.1131
$\psi_2(t)$	12	152	<b>0.4334</b>	12	24	0.1407	12	8	<b>0.1090</b>
$\psi_3(t)$	14	152	0.5699	13	24	0.1666	12	8	0.1197
$\psi_4(t)$	15	152	2.1199	14	24	0.5048	12	8	0.2836

Tableau 15.3. Résultats numériques obtenus pour les différents choix de  $\theta$ .

### Problème de $(\mathcal{NCM})$

1. Soit  $\mathcal{L}(X) = X$ . On présente les résultats numériques dans le tableau suivant :

$\psi_i(t) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	12	152	0.2238	12	24	<b>0.0369</b>	11	8	0.0294
$\psi_2(t)$	12	152	<b>0.2028</b>	14	24	0.0466	11	8	<b>0.0251</b>
$\psi_3(t)$	16	152	0.8156	14	24	0.0430	12	8	0.0358
$\psi_4(t)$	17	152	0.8411	15	24	0.1915	13	8	0.1168

Tableau 16.1. Résultats numériques obtenus pour les différents choix de  $\theta$ .

2. Soit  $\mathcal{L}(X) = U^{\frac{1}{2}} X U^{\frac{1}{2}}$  où  $U \in \mathbb{S}_{++}^n$ . On présente les résultats numériques dans le tableau ci-dessous.

$\psi_i(t) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	13	153	<b>0.1837</b>	13	25	0.0440	11	8	<b>0.0313</b>
$\psi_2(t)$	13	153	0.1952	13	25	<b>0.0424</b>	12	8	0.0362
$\psi_3(t)$	17	153	0.1982	15	25	0.0444	12	8	0.0342
$\psi_4(t)$	17	153	0.8528	16	25	0.2013	13	8	0.1259

Tableau 16.2. Résultats numériques obtenus pour les différents choix de  $\theta$ .

3. Soit  $\mathcal{L}(X) = \Sigma \circ X$  pour une matrice donnée  $\Sigma \in \mathbb{S}^n \cap \mathbb{R}_+^{n \times n}$ . On présente les

résultats numériques dans le tableau suivant :

$\psi_i(t) \setminus \theta$	0.1			0.5			0.9		
	inner	outer	CPU	inner	outer	CPU	inner	outer	CPU
$\psi_1(t)$	12	154	<b>0.1838</b>	12	25	0.0385	11	8	0.0305
$\psi_2(t)$	13	154	0.2123	12	25	<b>0.0375</b>	12	8	<b>0.0254</b>
$\psi_3(t)$	16	154	0.2276	14	25	0.0545	12	8	0.0438
$\psi_4(t)$	17	154	1.2424	16	25	0.2549	13	8	0.1425

Tableau 16.3. Résultats numériques obtenus pour les différents choix de  $\theta$ .

Le tableau suivant présente la valeur de  $\Psi(X^0, Z^0; \mu^{(0)})$  selon le choix de l'opérateur  $\mathcal{L}(X)$  et la fonction noyau  $\psi_i(t)$ .

$\psi_i(t) / \mathcal{L}(X)$	$\Psi(X^0, Z^0; \mu^{(0)})$		
	$X$	$U^{\frac{1}{2}} X U^{\frac{1}{2}}$	$\Sigma \circ X$
$\psi_1(t)$	0.5023	0.3338	0.4868
$\psi_2(t)$	0.5276	0.3472	0.5223
$\psi_3(t)$	0.2276	0.1573	0.5573
$\psi_4(t)$	0.2629	0.1795	0.6305

Tableau 16.4. La valeur de  $\Psi(X^0, Z^0; \mu^{(0)})$  selon l'opérateur  $\mathcal{L}(X)$  et la fonction noyau  $\psi_i(t)$ .

**Commentaire :**

Ces résultats numériques présentent le comportement de notre fonction noyau par rapport à d'autres fonctions noyaux qui ont la même complexité pour les algorithmes à grand pas et selon un choix théorique convenable du paramètre associé. Cette étude numérique montre et confirme nos résultats théoriques pour l'Algorithme 3.2.1 à grand pas.

# Conclusion générale et perspectives

Dans cette thèse, on a développé deux algorithmes primal-dual de points intérieurs pour résoudre un problème quadratique convexe semi-défini CQSDP :

- Le premier est de trajectoire centrale de type primal-dual à petit pas qui est efficace pour résoudre plusieurs problèmes d'optimisations. Notamment CQSDP, car de l'aspect théorique, on obtient une meilleure complexité polynomiale qui est de l'ordre  $O(\sqrt{n} \log(\frac{n}{\epsilon}))$ . Mais, l'absence de la symétrie des directions et l'obtention d'un point strictement réalisable avec la mesure de proximité  $\delta(X, Z; \mu) < \frac{1}{\sqrt{2}}$  justifient la difficulté de cette méthode.
- Le deuxième est basé sur une nouvelle fonction noyau paramétrée qui est la version générale de la fonction noyau proposée par M.W. Zhang [44]. Sous un choix spécial du paramètre  $q$ , on a montré que la complexité de cet algorithme à grand pas est réduite par un facteur de  $O(\log n)$  par rapport à la complexité obtenue par M.W. Zhang [44] et on obtient la meilleure complexité connue jusqu'à présent pour ces algorithmes à grand pas qui est de l'ordre  $O(\sqrt{n} \log n \log \frac{n}{\epsilon})$ . De plus, notre analyse est basée sur la programmation semi-définie (SDP) et on prend en considération que les directions de Nesterov-Todd ne sont pas orthogonales. Finalement, les différents tests numériques confirment et montrent l'efficacité de nos résultats théoriques.

Notons que notre étude a réalisé les contributions suivantes :

- Pour l'Algorithme 2.4.4 :
  - Etablir une condition suffisante pour assurer la stricte faisabilité du pas de Newton complet.
  - Etablir l'influence du pas de Newton complet en saut de dualité.
  - Etablir le taux de la réduction de  $\mu$  pour rester proche de la trajectoire centrale dans le cas CQSDP où les directions ne sont pas orthogonales.
- Pour l'Algorithme 3.2.1 :
  - Etablir l'influence du paramètre  $q$  sur l'Algorithme 3.2.1, dont théoriquement la complexité de cet algorithme avec  $\theta = O(1)$  et  $\tau = \Theta(n)$  est diminuée par un facteur de  $O(\log n)$  par rapport à celui du M.W. Zhang.
  - Etablir numériquement l'influence du paramètre  $q$  sur le comportement de l'Algorithme

3.2.1 avec  $\theta = O(1)$  et  $\tau = \Theta(n)$ .

- Etablir le comportement de la fonction mentionnée en (3.5) sur l'Algorithme 3.2.1 par rapport à d'autres fonctions noyaux paramétrées qui ont la même complexité pour les algorithmes à grand pas et selon un choix théorique convenable pour le paramètre  $q$ .

Ces contributions sont d'ordre algorithmique, théorique et numérique.

Notre prochain travail est consacré au développement d'un algorithme primal-dual de points intérieurs basé sur une nouvelle fonction noyau avec d'autre schéma de symétrisation dont le but est d'améliorer la complexité.

# Bibliographie

- [1] M. Achache. A new parameterized kernel function for LO yielding the best known iteration bound for a large-update interior point algorithm. *Afrika Matematika*, 27 (3) : 591-601, (2016).
- [2] M. Achache. A new primal-dual path-following method for convex quadratic programming. *Computational and Applied Mathematics*, (25) pp. 97-110, (2006).
- [3] M. Achache. A weighted path-following method for the linear complementarity problem. *Universitatis Babeş. Bolyai. Series Informatica*, (49) (1) : 61-73, (2004).
- [4] M. Achache. Complexity analysis and numerical implementation of a short-step primal-dual algorithm for linear complementarity problems. *Computational and Applied Mathematics*, (216), 1889-1895, (2010).
- [5] M. Achache. Complexity analysis of an interior point algorithm for the semidefinite optimization based on a kernel function with a double barrier term. *Acta Mathematica Sinica. English Series*, 31 (3) : 543-556, (2015).
- [6] M. Achache, L. Guerra. A full-Nesterov-Todd-step primal-dual interior point algorithm for convex quadratic semidefinite optimization. *Computational and Applied Mathematics*, (**231**) : 581-590, (2014).
- [7] L. Adler, R.D.C. Monteiro. Interior path-following primal-dual algorithm . Part II : convex quadratic programming. *Mathematical programming*, (44), 43-66, (1989).
- [8] F. Alizadeh. Interior point methods in semidefinite programming with application to combinatorial optimization. *SIAM Journal on Optimization*, (5), 13-51, (1995).
- [9] Y.Q. Bai, M. El Ghami, C. Roos. A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization. *SIAM Journal on Optimization*, **15** (1) : 101-128, (2004).
- [10] Y.Q. Bai, J.L. Guo, C. Roos. A new kernel function yielding the best known iteration bounds for primal-dual interior-point algorithms. *Acta Mathematica Sinica. English Series*, **25** (12) : 2169-2178, (2009).

- [11] Y.Q. Bai, G. Lesaja, C. Roos, G.Q. Wang, and M. El Ghami. A class of large-and small-update primal-dual interior-point algorithm for linear optimization. *Journal of optimization theory and application*, 138 (3) : 341-359, (2008).
- [12] D. Benterki, J.P. Crouzeix, B. Merikhi. A nematical implementation of an interior point method for semidefinite programming. *Pesquisa Operacional*, 23 (1) : 49-59, (2003).
- [13] C. Besse. Introduction à l'analyse matriciel. Laboratoire MIP. Université Paul Sabatier, Septembre (2003).
- [14] M. Bouafia, D. Benterki, A. Yassine. Complexity analysis of interior point methods for linear programming based on a parameterized kernel function. *RAIRO - Operations Research*, 50 (4-5) : 935-949, (2016).
- [15] X.Z. Cai, L. Wu, Y.J. Yue, M.M. Li, G.Q. Wang. Kernel-function based primal-dual interior-point methods for convex quadratic optimization over symmetric cone. *Journal of inequalities and applications*, 308, 1-22, (2014).
- [16] E. De Klerk. Interior point methods for semidefinite programming. Master of Science in the Faculty of Engineering University of Pretoria, (1997).
- [17] E. De Klerk, C. Roos, T. Terlaky. A short survey on semidefinite programming. Faculty of Technical Mathematics and Informatics, Delft University of Technology. Netherlands, (1997).
- [18] E. De Klerk, C. Roos, T. Terlaky. On primal-dual path-following algorithms for semidefinite programming. *Mathematical programming*, 137-157, (1998).
- [19] M. El Ghami. New primal-dual interior-point methods based on kernel functions. Phd thesis. Delft University, Netherland, (2005).
- [20] N.J. Higham. Computing the nearest correlation matrix—a problem from finance. *IMA journal of numerical analysis*, 22, 329-343, (2002).
- [21] R.A. Horn, R.J. Charles. *Matrix analysis*, Compridge University Press, UK, (1986).
- [22] B. Jansen, C. Roos, T. Terlaky, J.Ph. Vial, Primal-dual algorithm for linear programming based on the logarithmic barrier method. *Journal of optimization theory and applications*, 83 : 1-26, (1994).
- [23] A. Keraghel. *Analyse convexe. Theorie fondamentale et exercices*. Université de Sétif. Faculté des sciences. Laboratoire de mathématique fondamentale et numérique. (2001).
- [24] M. Kojima, M. Shida, S. Shindoh. Search directions in the SDP and monotone SDLCP : Generalization and inexact computation. *Mathematical Programming*, 85 : 51-80, (1999).

- [25] J.W. Nie, Y.X. Yuan. A potential reduction algorithm for an extended SDP problem. *Science in China (Series A)*, 43 (1) : 35-46, (2000).
- [26] J.W. Nie, Y.X. Yuan. A predictor-corrector algorithm for QSDP combining Dikin-type and Newton centering steps. *Annals of Operations Research*, 103, 155-133, (2001).
- [27] J. Peng, C. Roos, T. Terlaky. New complexity analysis of the primal-dual method for semidefinite optimization based on the Nesterov-Todd direction. *Journal of optimization theory and applications*; vol.109, no.2,pp. 327-343, (2001).
- [28] J. Peng, C. Roos, T. Terlaky. Self-regular functions and new search directions for linear and semidefinite optimization. *Mathematical Programming*, **93** : 129-171, (2002).
- [29] J. Peng, C. Roos, T. Terlaky. Self-regularity. A new paradigm for Primal-Dual Interior Point Algorithm. Princeton University Press. Princeton, (2002).
- [30] W.H. Press, B.P. Flannery, S.A. Tenkolsky, W.T. Vetterling. *Numerical Recipes in Pascal. The Art of Scientific computing.* Combridge University Prees, (1987).
- [31] C. Roos, T. Terlaky, J.Ph. Vial, *Theory and algorithms for linear optimization. An interior point approach.* John-Wiley. Sons, Chichester, UK, (1997).
- [32] J.F. Sturm, S. Zhang. Symmetric primal-dual path following algorithms for semidefinite programming. Technical Report 9554/A, Tinbergen Institute, Erasmus University Rotterdam, (1995).
- [33] M.J. Todd. A study of search directions in primal-dual interior-point methods for semidefinite programming. Working paper, school of Operations Research and Industrial Engineering, Cornell University, Ithaca. New-York 14853, (1998).
- [34] M.J. Todd, K.C. Toh, R.H. Tütüncü. On the Nesterov-Todd direction in semidefinite programming. *Mathematical programming.* Copyright (c) by the Society for Industrial and Applied Mathematics. *SIAM Journal on Optimization*, 8 : 769-796, (1998).
- [35] K.C. Toh. An inexact primal-dual path following algorithm for convex quadratic semidefinite programming. *Mathematical Programming*, 112 (1) : 221-254, (2008).
- [36] K.C. Toh, R.H. Tütüncü, M.J. Todd. Inexact primal-dual path-following algorithms for a special class of convex quadratic SDP and related problems. *Pacific journal of optimization.* (3), N<sup>o</sup>1, pp. 135-164, (2007).
- [37] G.Q. Wang, Y.Q. Bai. A new primal-dual path-following interior-point algorithm for semidefinite optimization. *Journal of Mathematical Analysis and Applications*, 353 (1) : 339-349, (2009).

- [38] G.Q. Wang, Y.Q. Bai. Primal-dual interior-point algorithms for convex quadratic semidefinite optimization. *Nonlinear analysis : theory, methods & applications*, 71 (7-8) : 3389-3402, (2009).
- [39] G.Q. Wang, Y.Q. Bai, C. Roos. Primal-dual interior-point algorithms for semidefinite optimization based on a simple kernel function. *Journal of mathematical modelling and algorithms*, 4 (4) : 409-433, (2005).
- [40] G.Q. Wang, Z.H. Zhang, D.T. Zhu. On extending primal-dual interior-point method for linear optimization to convex quadratic cone optimization. *Numerical functional analysis and optimization*, 34 (5) : 576-603, (2013).
- [41] G.Q. Wang, D.T. Zhu. A unified kernel function approach to primal-dual interior-point algorithms for convex quadratic SDO. *Numerical Algorithms*, 57 (4) : 537-558, (2011).
- [42] S.J. Wright. *Primal-dual interior point methods*. SIAM Publication, Philadelphia, (1997).
- [43] Y. Ye. *Interior point algorithms : Theory and Analysis*. John-Wiley. Sons, Chichester, UK, (1997).
- [44] M.W. Zhang. A large-update interior-point algorithm for convex quadratic semi-definite optimization based on a new kernel function. *Acta Mathematica Sinica. English Series*, 28 (11) : 2313-2328, (2012).
- [45] Y. Zhang. On extending some primal-dual interior-point algorithms from linear programming to semidefinite programming. *SIAM Journal on Optimization*, 8 : 365-386, (1998).
- [46] L. Zhang, Y. Xu, Z. Jin. An efficient algorithm for convex quadratic semi-definite optimization. *Numerical Algebra. Control and Optimization*, 2 (1) : 129-144. (2012).
- [47] D. Zhao, M.W. Zhang. A primal-dual large-update interior point algorithm for semi-definite optimization based on a new parametric kernel function. *Statistics, optimization and information computing*, (1) : 41-61, (2013).



# Annexe

On commence de présenter la preuve des Lemmes qui ont été utilisés en Chapitre 2 comme des lemmes techniques introduits par De klerk en SDP [16].

Notons que les deux premiers lemmes nous conduit à obtenir le gradient de la fonction  $f_\mu(X)$ , où

$$f_\mu(X) = C \bullet X + \frac{1}{2}X \bullet \mathcal{Q}(X) - \mu \log \det X, \mu > 0.$$

**Lemme I** Soit  $f : \mathbf{S}_{++}^n \rightarrow \Re$  une fonction donnée par  $f(X) = \log \det X$ , et

$$\nabla f(X) = \begin{bmatrix} \frac{\partial f(X)}{\partial x_{11}} & \dots & \frac{\partial f(X)}{\partial x_{1n}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f(X)}{\partial x_{n1}} & \dots & \frac{\partial f(X)}{\partial x_{nn}} \end{bmatrix},$$

alors  $\nabla f(X) = X^{-1}$ .

**Preuve :** Soient  $X \in \mathbf{S}_{++}^n$  et  $H \in \mathbf{S}^n$  avec  $(X + H) \in \mathbf{S}_{++}^n$ . On a :

$$\begin{aligned} f(X + H) - f(X) &= \log \det(X + H) - \log \det X, \\ &= \log \left[ \frac{\det(X + H)}{\det(X)} \right], \\ &= \log [\det(X^{-1}) \det(X + H)], \\ &= \log [\det((X^{-1})(X + H))], \\ &= \log \det(I + X^{-1}H), \\ &= \log \det(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}), \text{ car } X^{-1}H \sim X^{-\frac{1}{2}}HX^{-\frac{1}{2}}. \end{aligned}$$

Comme  $X \in \mathbf{S}_{++}^n$  et  $H \in \mathbf{S}^n$  telle que  $(X + H) \in \mathbf{S}_{++}^n$  alors :

$$\left[ \prod_{i=1}^n \lambda_i(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) \right]^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^n \lambda_i(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}), \forall \lambda_i(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) > 0,$$

vous pouvez consulter [23] pour voir la démonstration de cette inégalité, qui est équivalente à :

$$\det(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) \leq \left[ \frac{1}{n} \mathbf{Tr} (I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) \right]^n .$$

En prenant le logarithme pour cette dernière, on obtient :

$$\begin{aligned} \log \det(I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) &\leq \log \left[ \frac{1}{n} \mathbf{Tr} (I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) \right]^n , \\ &= n \log \left[ \frac{1}{n} \mathbf{Tr} (I + X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) \right] , \\ &= n \log \left[ 1 + \frac{1}{n} \mathbf{Tr} (X^{-\frac{1}{2}}HX^{-\frac{1}{2}}) \right] . \end{aligned}$$

On sait que  $\log(1+t) \leq t, \forall t > -1$ , alors :

$$\begin{aligned} f(X+H) - f(X) &\leq \mathbf{Tr} (X^{-\frac{1}{2}}HX^{-\frac{1}{2}}), \\ &= \mathbf{Tr} (X^{-1}H), \\ &= X^{-1} \bullet H, \end{aligned}$$

et par conséquent  $\nabla f(X) = X^{-1}$ , ce qui achève la preuve. ■

**Lemme II** Soit  $f : \mathbf{S}_{++}^n \rightarrow \Re$  une fonction donnée par  $f(X) = \mathbf{Tr} (CX) = C \bullet X$ , où  $C \in \mathbf{S}^n$ , alors  $\nabla f(X) = C$ .

**Preuve :** Soient  $X \in \mathbf{S}_{++}^n$  et  $H \in \mathbf{S}^n$  avec  $(X+H) \in \mathbf{S}_{++}^n$ . On a :

$$\begin{aligned} f(X+H) - f(X) &= \mathbf{Tr} (C(X+H)) - \mathbf{Tr} (CX), \\ &= \mathbf{Tr} (C(X+H) - CX), \\ &= \mathbf{Tr} (CX + CH - CX), \\ &= \mathbf{Tr} (CH) = C \bullet H, \end{aligned}$$

et par conséquent  $\nabla f(X) = C$ . Ce qui achève la preuve. ■

Le résultat suivant conduit à obtenir le Hessien de la fonction  $f_\mu(X)$ .

**Lemme III** Soient  $f : \mathbf{S}_{++}^n \rightarrow \Re$  une fonction donnée par  $f(X) = \log \det X$  et  $\nabla f(X) = X^{-1}$ , alors  $\nabla^2 f(X)$  est un opérateur linéaire qui vérifie

$$\nabla^2 f(X)H = -X^{-1}HX^{-1}, \forall H \in \mathbf{S}^n .$$

**Preuve :** Soit  $\mathcal{L}(\mathbf{S}^n, \mathbf{S}^n)$ , l'espace des opérateurs linéaires de  $\mathbf{S}^n$  dans  $\mathbf{S}^n$ . La dérivée de

$\nabla f$  au sens de Frechet est définie par une fonction

$$\nabla^2 f : \mathbf{S}^n \rightarrow \mathcal{L}(\mathbf{S}^n, \mathbf{S}^n)$$

avec

$$\lim_{\|H\| \rightarrow 0} \frac{\|\nabla f(X + H) - \nabla f(X) - \nabla^2 f(X)H\|}{\|H\|} = 0. \quad (*.1)$$

On montre que  $\nabla^2 f(X)H = -X^{-1}HX^{-1}$  vérifie (\*.1).

Soit  $H \in \mathbf{S}^n$  telle que  $(X + H)$  est inversible, on considère :

$$\begin{aligned} \|\nabla f(X + H) - \nabla f(X) - \nabla^2 f(X)H\| &= \|(X + H)^{-1} - X^{-1} + X^{-1}HX^{-1}\|, \\ &= \|(X + H)^{-1}(I - (X + H)X^{-1} + (X + H)X^{-1}HX^{-1})\|, \\ &= \|(X + H)^{-1}(HX^{-1}HX^{-1})\|, \\ &\leq \|(X + H)^{-1}\| \|HX^{-1}HX^{-1}\|, \\ &\leq \|(X + H)^{-1}\| \|H\| \|X^{-1}HX^{-1}\|, \end{aligned}$$

c-à-d  $\nabla^2 f(X)H$  vérifie (\*.1), et par conséquent, on obtient :  $\nabla^2 f(X)H = -X^{-1}HX^{-1}$ .

Ce qui achève la preuve. ■

Maintenant, on présente la preuve des Lemmes qui ont été utilisés en Chapitre 3 comme des lemmes techniques.

On commence par le lemme suivant qui nous donne une borne supérieure de  $f_1''(\alpha)$  pour un problème CQSDP où les directions ne sont pas orthogonales.

**Lemme IV** [44, Lemme 4.1] On a :

$$f_1''(\alpha) \leq 2\delta^2 \psi''(\lambda_{\min}(V) - 2\alpha\delta),$$

**Preuve :** La troisième équation du système (3.11) et (3.12) implique que

$$\|D_X + D_Z\|_F = 2\delta.$$

Puisque  $D_X \bullet D_Z \geq 0$ , on en déduit facilement que

$$\|D_X\|_F^2 + \|D_Z\|_F^2 \leq 4\delta^2. \quad (*.2)$$

Et par conséquent, on obtient

$$|\lambda_{\max}(D_X)| \leq 2\delta, \quad |\lambda_{\max}(D_Z)| \leq 2\delta.$$

En utilisant (3.3) et  $V + \alpha D_X \succeq 0$ , on obtient :

$$\lambda_i(V + \alpha D_X) \geq \lambda_{\min}(V) - \alpha |\lambda_{\max}(D_X)| \geq \lambda_{\min}(V) - 2\alpha\delta, \quad i = 1, \dots, n.$$

D'après la dérivée troisième de  $\psi(t)$ ,  $\psi''(t)$  est strictement décroissante pour tout  $t \in ]0, +\infty[$ , et on a :

$$\psi''(\lambda_i(V + \alpha D_X)) \leq \psi''(\lambda_{\min}(V) - 2\alpha\delta).$$

De plus, on obtient

$$\psi''(V + \alpha D_X) \preceq \psi''(\lambda_{\min}(V) - 2\alpha\delta)I.$$

Puisque  $D_X^2 \in \mathbb{S}_+^n$ , en utilisant (3.2) et (3.1), on obtient :

$$\begin{aligned} \mathbf{Tr} [\psi''(V + \alpha D_X) D_X^2] &\leq \mathbf{Tr} [\psi''(\lambda_{\min}(V) - 2\alpha\delta) I D_X^2], \\ &= \psi''(\lambda_{\min}(V) - 2\alpha\delta) \mathbf{Tr} (I D_X^2), \\ &= \psi''(\lambda_{\min}(V) - 2\alpha\delta) \mathbf{Tr} (D_X^2), \\ &= \psi''(\lambda_{\min}(V) - 2\alpha\delta) \sum_{i=1}^n \lambda_i(D_X^2). \end{aligned} \quad (*.3)$$

De même,

$$\mathbf{Tr} [\psi''(V + \alpha D_Z) D_Z^2] \leq \psi''(\lambda_{\min}(V) - 2\alpha\delta) \sum_{i=1}^n \lambda_i(D_Z^2). \quad (*.4)$$

En utilisant

$$f_1''(\alpha) = \frac{1}{2} \mathbf{Tr} [\psi''(V + \alpha D_X) D_X^2 + \psi''(V + \alpha D_Z) D_Z^2],$$

d'après (\*.2), (\*.3) et (\*.4), on obtient :

$$\begin{aligned} f_1''(\alpha) &\leq \frac{1}{2} \psi''(\lambda_{\min}(V) - 2\alpha\delta) \sum_{i=1}^n (\lambda_i(D_X^2) + \lambda_i(D_Z^2)), \\ &\leq 2\delta^2 \psi''(\lambda_{\min}(V) - 2\alpha\delta). \end{aligned}$$

Ce qui achève la preuve. ■

Les deux lemmes suivants sont très importants pour l'analyse de l'Algorithme 3.2.1.

**Lemme V** [9, Lemme 6.2] Soit  $\varrho : [0, +\infty[ \rightarrow [1, +\infty[$  la fonction inverse de  $\psi(t)$  pour  $t \geq 1$ . Alors on a :

$$\sqrt{1 + 2s} \leq \varrho(s) \leq 1 + \sqrt{2s}, \quad s \geq 0.$$

**Preuve :** Soit  $s = \psi(t)$ , pour tout  $t \geq 1$ , c'est à dire  $\varrho(s) = t$ ,  $t \geq 1$ . En utilisant la deuxième propriété du Lemme [3.1.4], on obtient :

$$s = \psi(t) \leq \frac{t^2 - 1}{2},$$

et par conséquent

$$t = \varrho(s) \geq \sqrt{1 + 2s}, \quad \text{pour tout } s \geq 0.$$

D'autre part, puisque  $\psi''(t) \geq 1$  pour tout  $t > 0$  et d'après la définition de  $\psi(t)$  en (3.4) on a :

$$s = \psi(t) = \int_1^t \int_1^\xi \psi''(\zeta) d\zeta d\xi \geq \int_1^t \int_1^\xi d\zeta d\xi = \frac{1}{2}(t-1)^2, \quad t \geq 1.$$

Alors

$$t = \varrho(s) \leq 1 + \sqrt{2s}, \quad \text{pour tout } s \geq 0.$$

Ce qui achève la preuve. ■

**Lemme VI** [9, Lemme 6.1] Soit  $\psi_b(t)$  le terme barrière de  $\psi(t)$  et soit  $\bar{\rho} : [0, \infty[ \rightarrow ]0, 1]$  la fonction inverse de la restriction  $-\psi_b'(t)$  sur l'intervalle  $]0, 1]$ . Alors on a :

$$\rho(s) \geq \bar{\rho}(1 + 2s).$$

**Preuve :** Soit  $t = \rho(s)$ , puisque  $\rho$  est la fonction inverse de  $-\frac{1}{2}\psi'(t)$  pour tout  $t \leq 1$ , alors

$$-2s = \psi'(t) = t + \psi'_b(t), \quad t \leq 1.$$

Comme  $t \leq 1$  alors

$$-\psi'_b(t) = t + 2s \leq 1 + 2s,$$

on note que  $-\psi'_b(t)$  est strictement décroissante et par conséquent

$$t = \rho(s) \geq \bar{\rho}(1 + 2s).$$

Ce qui achève la preuve. ■

On présente le lemme technique suivant qui sera utilisé par la suite.

**Lemme VII**[28, Lemme 2.1] Si  $\alpha \in [0, 1]$ , alors

$$(1 + t)^\alpha \leq 1 + \alpha t, \quad \text{pour tout } t \geq -1.$$

**Preuve :** Soit la fonction

$$f(t) = (1 + t)^\alpha - 1 - \alpha t, \quad t \geq -1.$$

Alors

$$f'(t) = \alpha(1 + t)^{\alpha-1} - \alpha$$

et

$$f''(t) = \alpha(\alpha - 1)(1 + t)^{\alpha-2}.$$

Puisque  $\alpha \in [0, 1]$ , alors  $f''(t) \leq 0$ , ce qui implique que  $f(t)$  est concave. De plus,  $f'(0) = 0$  alors la fonction  $f(t)$  est maximal au point  $t = 0$ . Finalement, puisque  $f(0) = 0$ , alors on obtient

$$(1 + t)^\alpha \leq 1 + \alpha t, \quad \text{pour tout } t \geq -1 \text{ et } \alpha \in [0, 1].$$

Ce qui achève la preuve. ■

**Lemme VIII** [28, Proposition 2.2] Soit  $t_0, t_1, \dots, t_K$  une suite des nombres positifs qui vérifie :

$$t_{k+1} \leq t_k - \bar{\beta}(t_k)^{1-\gamma}, \quad k = 0, 1, \dots, K - 1,$$

tels que  $\bar{\beta} > 0$  et  $\gamma \in ]0, 1]$ ; alors :

$$K \leq \left\lceil \frac{t_0^\gamma}{\bar{\beta}^\gamma} \right\rceil.$$

**Preuve :** Considérant l'inégalité suivante :

$$t_{k+1} \leq t_k - \bar{\beta} (t_k)^{1-\gamma}.$$

D'après le Lemme précédent, on peut écrire

$$(t_{k+1})^\gamma \leq (t_k - \bar{\beta} (t_k)^{1-\gamma})^\gamma = t_k^\gamma (1 - \bar{\beta} (t_k)^{-\gamma})^\gamma \leq t_k^\gamma (1 - \gamma \bar{\beta} (t_k)^{-\gamma}) = t_k^\gamma - \gamma \bar{\beta}.$$

Alors, pour tout  $k$ , on a :

$$t_k^\gamma \leq t_0^\gamma - k\gamma\bar{\beta}.$$

En remplaçant  $k$  par  $K$ , on obtient  $t_0^\gamma - K\gamma\bar{\beta} \geq 0$ . Ce qui achève la preuve. ■

## ملخص

في هذه الأطروحة، اقترحنا خوارزميتين أولية – ثنوية من النقاط الداخلية للبرمجة التربيعية المحدبة نصف المعرفة. الأولى ذات المسار المركزي حيث في كل التكرار نستعمل الخطوة الكاملة لنيوتن ومقياس القرب للحصول على حل تقريبي. وتستند الخوارزمية الثانية على دالة نواة جديدة بحيث أن هذه الدالة متعلقة بوسيط و تعميم تلك التي قدمها M. W. Zhang في عام 2012. دراسة هذه الدالة تقودنا إلى أفضل تكلفة معروفة حتى الآن لهذا النوع من الخوارزمية مع خطوة كبيرة وصغيرة. نتبع هذه الدراسة بالنتائج العددية لإظهار فعالية هذين الخوارزميتين المقترحتين. جلبت هذه المقترحات مساهمات جديدة من النظام الخوارزمي والنظري والعددي.

**الكلمات المفتاحية :** طرق النقاط الداخلية ; البرمجة التربيعية النصف معرفة ; دالة نواة ; خوارزمية أولية – ثنوية ; تكلفة خوارزمية.

## Résumé :

Dans cette thèse, on a proposé deux algorithmes primal-dual de points intérieurs pour la programmation quadratique convexe semi-définie (CQSDP). Le premier est de trajectoire centrale tel que à chaque itération on utilise le pas de Newton complet et une mesure de proximité pour obtenir une solution approximative du (CQSDP). Le deuxième algorithme est basé sur une nouvelle fonction noyau telle que cette fonction est la version paramétrée de celle qui est introduite par de M. W. Zhang en 2012. L'étude de cette fonction nous conduit à une meilleure complexité connue jusqu'à maintenant pour ce type d'algorithme à grand et petit pas.

On suit cette étude par des résultats numériques pour montrer l'efficacité de ces deux algorithmes proposés. Ces propositions ont apporté de nouvelles contributions d'ordre algorithmique, théorique et numérique.

**Mots clés :** Méthodes de points intérieurs; programmation quadratique convexe semi-définie; fonction noyau ; algorithme primal-dual ; complexité algorithmique.

## Abstract :

In this thesis, we have proposed two primal-dual interior point algorithms for convex quadratic semidefinite programming (CQSDP). The first one is of the central path where we use at each iteration the full Newton step and a suitable proximity measure to obtain an approximate solution for (CQSDP). The second algorithm is based on a new kernel function such that this function is the parameterized version of the one introduced by M. W. Zhang in 2012. The study of this function leads us to a better complexity known until now of this type of large and small update algorithm.

This study was followed by numerical results to show the efficiency of these two proposed algorithms. These proposals brought new contributions of algorithmic, theoretical and numerical order.

**Key words:** Interior point methods; convex quadratic semidefinite programming; kernel function; primal-dual algorithm; algorithmic complexity.