



Semidefinite Programming

Methods and algorithms for energy management

Agnès Gorge

September 2013

Cette thèse a pour objet d'évaluer l'apport de la programmation semi-définie positive (SDP), méthode prometteuse de l'optimisation conique, pour la résolution pratique des problèmes d'optimisation rencontrés en management d'énergie. Elle est motivée par les résultats récents reportés dans la littérature exhibant un fort potentiel de la SDP pour le traitement des problèmes à caractère combinatoire et/ou aléatoire marqués tels que ceux fréquemment rencontrés en management d'énergie, et pour lesquels les méthodes classiques de résolution sont limitées. La thèse a pour ambition d'évaluer ce qu'une méthode alternative comme la SDP pourrait apporter quant à la résolution de ces problèmes en général, avec une attention particulière portée au problème de la planification des arrêts des centrales nucléaires pour rechargement du combustible et maintenance.

Le travail a consisté à identifier les problèmes concernés par notre démarche, à les modéliser de façon appropriée et à expérimenter la mise en oeuvre numérique de leur résolution à l'aide de la SDP. Il est décliné selon deux axes. Concernant le premier, nous investiguons, d'un point de vue théorique et numérique, les potentialités de la SDP pour l'élaboration de relaxations performantes de problèmes NP-difficiles présentant un caractère combinatoire et/ou quadratique. Concernant le second, nous exploitons la puissance de modélisation de la SDP pour la prise en compte de la nature aléatoire des problèmes d'optimisation. Afin de préciser le contexte scientifique de la thèse, nous rappelons préalablement quelques éléments d'introduction à la programmation conique et plus particulièrement à la SDP. Puis nous présentons les caractéristiques des problèmes d'optimisation rencontrés en management d'énergie et développons différentes approches pour leur traitement par la SDP.

Introduction à l'optimisation conique et à la SDP

L'optimisation conique peut être vue comme une extension naturelle de la programmation linéaire dans laquelle l'orthant positif de \mathbb{R}^n est remplacé par un cône convexe \mathcal{K} . Ce formalisme présente de nombreux avantages, en particulier le problème dual admet également une formulation conique faisant intervenir \mathcal{K}^* , le cône dual de \mathcal{K} , ce qui confère au problème d'intéressantes propriétés de symétrie et une extension des théorèmes de dualité faible et forte de la programmation linéaire.

Si $\mathcal{K} \subset \mathbb{R}^n$, alors un problème conique (P_P) et son dual (P_D) sont définis par la donnée d'un vecteur $c \in \mathbb{R}^n$ et d'une matrice $(A, b) \in \mathbb{R}^{m, n+1}$, de la façon suivante :

$$(P_P) \begin{cases} \inf & c^T x \\ \text{s.t.} & Ax = b \\ & x \in \mathcal{K} \end{cases} \quad (P_D) \begin{cases} \sup & b^T z \\ \text{s.t.} & y = c - A^T z \\ & y \in \mathcal{K}^* \end{cases}$$

L'optimisation conique bénéficie des bonnes propriétés de l'optimisation convexe, notamment concernant sa complexité. Ainsi, par une application directe du résultat de Grötschel et al. [12], la solution optimale peut être approchée aussi finement que voulu en temps polynomial à condition qu'il existe un oracle de séparation polynomial pour \mathcal{K} . La méthode utilisée, dite *des ellipsoïdes*, est donc fondamentale sur le plan théorique. En pratique, elle s'est révélée peu performante et a rapidement été supplantée par d'autres méthodes polynomiales plus efficaces. En particulier, selon le résultat fondamental de Nesterov et Nemirovski [20], il est possible d'étendre les méthodes de points-intérieurs, initialement conçues pour la programmation linéaire, à n'importe quel problème convexe du moment qu'il existe une fonction barrière pour l'ensemble réalisable exhibant une propriété de régularité spécifique dite *d'auto-concordance*.

La programmation semi-définie positive (SDP) est le cas particulier de la programmation conique dans lequel le cône \mathcal{K} est \mathbb{S}_+^n , c'est-à-dire l'ensemble des matrices semi-définies positives. Rappelons qu'une matrice X est semi-définie positive, ce qui est noté $X \succcurlyeq 0$, si elle est symétrique et si toutes ses valeurs propres sont positives ou nulles. Une autre définition fréquemment utilisée est la suivante : $X \in \mathbb{S}_+^n \Leftrightarrow u^T X u \geq 0, \forall u \in \mathbb{R}^n$.

Un SDP primal (SDP_P) et son dual (SDP_D) se présentent donc de la façon suivante :

$$(SDP_P) \begin{cases} \inf & A_0 \bullet X \\ \text{s.t.} & A_j \bullet X = b_j, \quad j = 1, \dots, m \\ & X \succcurlyeq 0 \end{cases} \quad (SDP_D) \begin{cases} \sup & b^T z \\ \text{s.t.} & A_0 - \sum_{j=1}^m A_j z_j \succcurlyeq 0 \end{cases}$$

où \bullet désigne le produit scalaire sur les matrices symétriques, à savoir $M \bullet N = \sum_{i=1}^n \sum_{j=1}^n M_{ij} N_{ij}$, pour $M, N \in \mathbb{S}^n$.

Parmi les méthodes de l'optimisation conique, la SDP se positionne comme la méthode polynomiale possédant la plus grande puissance de modélisation. En particulier, elle subsume d'autres méthodes connues de l'optimisation conique telles que la programmation linéaire (LP) et l'optimisation conique du second-ordre (SOCP). Ces caractéristiques attrayantes ont suscité un grand intérêt parmi la communauté scientifique ces dernières années. Il en résulte de nombreux travaux portant aussi bien sur la théorie sous-jacente à ces problèmes que sur leurs méthodes de résolution et leur applicabilité.

Concernant l'aspect théorique, de nombreux résultats ont été établis sur la géométrie des SDP, c'est-à-dire sur la caractérisation de leurs ensembles réalisables primal et dual. Il en ressort des résultats sur la caractérisation des points extrêmes de ces ensembles (en particulier, leur rang) [21], sur l'unicité des solutions optimales [2] ou encore sur la dimension de leurs faces et facettes [4]. Des travaux poussés ont également été menés sur la dualité des problèmes SDP [5], débouchant sur l'identification de 11 configurations possibles, différant par la présence ou non de la dualité forte et par le fait que les valeurs optimales primale et duale existent ou non, et sont atteintes ou non.

Les méthodes de résolution ont également fait l'objet de nombreux travaux. Les algorithmes de points-intérieurs restent à ce jour les plus étudiés et les plus usités pour leur efficacité et leur applicabilité à n'importe quel SDP. Initiés par Alizadeh en 1991 [1], ces méthodes ont donné lieu aux deux solveurs SDP les plus réputés, à savoir CSDP [9] et DSDP [6]. D'autres méthodes issues de la programmation non linéaire ont également été testées. Parmi elles, citons les méthodes de relaxation lagrangienne [15] et les méthodes de faisceaux [14].

L'intérêt pour la SDP s'est encore accru ces dernières années lorsque de nombreuses applications ont été identifiées dans des domaines variés tels que le contrôle, les statistiques, la finance, la localisation, l'optimisation robuste et l'ingénierie. Parmi toutes ces applications, l'utilisation de la SDP pour approximer le célèbre problème des moments généralisés (GPM), a particulièrement attiré l'attention de la communauté scientifique, de par la généralité de ce problème et son applicabilité à l'optimisation quadratique et combinatoire. Considérons par exemple le problème quadratique suivant :

$$(QCQP) \begin{cases} \min & x^T P_0 x + 2p_0^T x + \pi_0 \\ \text{s.t.} & x^T P_j x + 2p_j^T x + \pi_j \leq 0, \quad j = 1, \dots, m \end{cases} \quad \text{avec } P_j \in \mathbb{S}^n, p_j \in \mathbb{R}^n, \pi_j \in \mathbb{R}, \quad j = 0, \dots, m \quad (1)$$

Ce problème à objectif et contraintes quadratiques (QCQP) est convexe si et seulement $P_j \succcurlyeq 0$, $j = 0, \dots, m$. Autrement, il appartient à la classe des problèmes NP-difficiles. Il suffit pour le comprendre de remarquer que de nombreux problèmes difficiles peuvent se mettre sous cette forme, en particulier les problèmes combinatoires à variables binaires puisque $x_i \in \{0, 1\}$ est équivalent à la contrainte quadratique $x_i^2 - x_i = 0$. On peut établir très simplement une relaxation SDP de ce problème, dite *standard*, de la façon suivante :

$$\begin{cases} \inf & Q_0 \bullet Y \\ \text{s.t.} & Q_j \bullet Y \leq 0, \quad i = 1, \dots, m \\ & Y_{1,1} = 1 \\ & Y \succcurlyeq 0 \end{cases} \quad \text{avec } Q_j = \begin{pmatrix} \pi_j & p_j^T \\ p_j & P_j \end{pmatrix}, \quad j = 0, \dots, m \quad (2)$$

Cette relaxation peut être obtenue et interprétée de nombreuses façons, la plus simple d'entre elles consistant à introduire une nouvelle variable $Y = \tilde{x}\tilde{x}^T$, à remplacer les formes quadratiques $x^T P_j x + 2p_j^T x + \pi_j$ par leur équivalent $Q_j \bullet Y$, puis à relaxer la contrainte $Y = \tilde{x}\tilde{x}^T$ en $Y_{1,1} = 1$ et $Y \succcurlyeq 0$.

L'idée d'une relaxation SDP est attribuée à Lovász [19] et à Shor [26], mais ce sont les travaux de Goemans et Williamson [11] offrant une garantie sur l'optimalité de la borne ainsi obtenue dans le cas d'un $\{-1, 1\}$ -QP, qui ont déclenché le véritable engouement que l'on connaît pour la SDP.

Cependant, dans le cas d'un QCQP quelconque, il est généralement nécessaire de renforcer cette relaxation standard pour la rendre véritablement efficace. Pour cela, il suffit d'ajouter des contraintes valides au problème quadratique initial, puis d'appliquer la relaxation SDP standard à ce nouveau QCQP. Toute la difficulté réside donc dans l'identification des contraintes valides (ou coupes) les plus efficaces. Cette façon de voir unifie de nombreux travaux recensés dans la littérature, voir par exemple [1, 13, 17, 18, 24].

Une autre source importante d'application pour la SDP provient de sa capacité à approximer le problème des moments généralisés (GPM). Ce problème (GPM_P) et son dual (GPM_D) se définissent de la façon suivante :

$$(GPM_P) \left\{ \begin{array}{l} \min \int_{\mathcal{S}} h(\omega) P(\omega) d\omega \\ \text{s.t.} \int_{\mathcal{S}} f_i(\omega) P(\omega) d\omega = b_i, \quad i = 1, \dots, m \\ P \in \mathcal{M}(\mathcal{S}) \end{array} \right. \quad (GPM_D) \left\{ \begin{array}{l} \max \sum_{i=1}^m b_i z_i \\ \text{s.t.} \sum_{i=1}^m f_i(\omega) z_i \leq h(\omega), \quad \forall \omega \in \mathcal{S} \end{array} \right.$$

Dans le primal, la variable d'optimisation n'est pas un vecteur euclidien comme c'est le cas généralement, mais P , une mesure positive ou nulle sur $\mathcal{B}(\mathcal{S})$, la σ -algèbre de Borel de $\mathcal{S} \subset \mathbb{R}^n$. Il est cependant possible de la considérer comme un vecteur de dimension infinie, dans lequel chaque composante correspond à une valeur de $P(\omega)$ pour tout $\omega \in \mathcal{S}$. Le problème devient alors linéaire et le dual se déduit simplement comme le dual d'un programme linéaire, le nombre infini de variables induisant un nombre infini de contraintes. Prises ensemble, ces contraintes prennent la forme de la positivité sur \mathcal{S} de la fonction $f_z(\omega) = h(\omega) - \sum_{i=1}^m f_i(\omega) z_i$ dont les coefficients dépendent de la variable duale $z \in \mathbb{R}^m$.

De nombreux problèmes peuvent se modéliser comme des instances particulières de ce problème, mais l'intérêt n'en est que purement théorique car il n'existe pas de méthode de résolution générale connue pour ce problème. Toutefois, deux restrictions se révèlent très intéressantes puisqu'il existe alors une suite d'approximations SDP dont la valeur optimale tend vers la valeur initiale du (GPM). Ces restrictions sont les suivantes : on suppose premièrement que $P(\mathcal{S})$ est borné et on va prendre $P(\mathcal{S}) = 1$, ce qui revient à supposer que P est une mesure de probabilité. Deuxièmement, on se place dans un cadre polynomial, où les fonctions $h, f_i, i = 1, \dots, m$ sont des polynômes et où \mathcal{S} est un ensemble semi-algébrique.

C'est de ces restrictions que le problème tire son nom, puisque le primal se formule alors via des combinaisons linéaires de moments de P et toute la difficulté du problème se retrouve dans la dernière contrainte de la formulation ci-dessous, imposant à y d'être le vecteur des moments associés à P :

$$\left\{ \begin{array}{l} \min \sum_{\kappa \in \mathbb{N}_d^n} h_{\kappa} y_{\kappa} \\ \text{s.t.} \sum_{\kappa \in \mathbb{N}_d^n} f_{i\kappa} y_{\kappa} = b_i, \quad i = 1, \dots, m \\ y_{\kappa} = E_P(\omega^{\kappa}), \quad \kappa \in \mathbb{N}_d^n \quad (\text{signifie que } y \text{ est le vecteur des moments de } P) \\ P \in \mathcal{M}(\mathcal{S}) \end{array} \right.$$

où $\mathbb{N}_d^n = \{\kappa \in \mathbb{N}^n : \sum_{i=1}^n \kappa_i \leq d\}$ et pour tout polynôme f de degré d sur \mathbb{R}^n , f est le vecteur de ses coefficients, c'est-à-dire $f(x) = \sum_{\kappa \in \mathbb{N}_d^n} f_{\kappa} x^{\kappa}$ où $x^{\kappa} = \prod_{i=1}^n x_i^{\kappa_i}$ est un monôme.

Or il existe une relaxation SDP de la contrainte imposant à y d'être un vecteur de moment sur \mathcal{S} , via la semi-définie positivité d'une matrice dépendant de y et de \mathcal{S} . Il s'avère que le dual de ce SDP peut également être interprété comme une approximation (conservative cette fois) de (GPM_D), en appliquant le théorème de Putinar. En effet, ce dernier établit des conditions suffisantes pour la positivité d'un polynôme sur un ensemble semi-algébrique, qui font appel à des matrices semi-définies positives. De plus,

ces relaxations dépendent d'un paramètre entier r , dit le rang, qui lorsqu'il tend vers l'infini, fait tendre la valeur optimale des SDP ainsi obtenus vers la valeur optimale du problème initial. Cette séquence de SDP est connue sous le nom de *hiérarchie de Lasserre* [16].

Parmi les instances que nous pourrions ainsi approximer, se trouve le problème classique des moments (CPM), dans lequel les fonctions f_i sont les monômes de degré inférieur à d et h est la fonction indicatrice d'un certain ensemble \mathcal{K} . Ce problème revient à minimiser la probabilité $P[\omega \in \mathcal{K}]$ connaissant les moments d'ordre inférieur à d et le support de P . De nombreuses variantes de ce problème ont été étudiées par le passé, donnant lieu à des inégalités célèbres telles que les inégalités de Chebyshev ou de Markov.

En conclusion, même si les applications de la SDP sont légions, elles sont difficiles à identifier car la contrainte d'imposer à une matrice d'être semi-définie positive n'apparaît pas naturellement. Afin de prendre du recul sur la façon dont ces applications émergent, nous proposons une classification de ces processus en 3 catégories :

- identification d'une matrice SDP : lorsque l'une des définitions possibles d'une matrice SDP se retrouve clairement dans le problème. C'est le cas par exemple si l'on requiert d'une fonction quadratique en x d'être positive ou nulle pour tout $x : f(x, y) = \tilde{x}^T P(y) \tilde{x} \geq 0, \forall x \in \mathbb{R}^n \Leftrightarrow P(y) \succcurlyeq 0$;
- relaxation : lorsqu'on relaxe la contrainte $Y \in \mathcal{S}$, avec $\mathcal{S} \subset \mathbb{S}_+^n$, par $Y \succcurlyeq 0$. Exemple : $\mathcal{S} = \{xx^T : x \in \mathbb{R}^n\}$;
- exploitation d'un résultat reposant sur l'existence d'une matrice SDP. Exemple : f convexe $\Leftrightarrow \nabla^2 f \succcurlyeq 0$.

Parmi les résultats reposant sur l'existence d'une matrice SDP se trouve le S-Lemma [22], d'importance cruciale pour notre étude. Ce lemme donne en effet une condition suffisante, parfois nécessaire, pour qu'une contrainte quadratique soit valide sur un ensemble défini par des contraintes quadratiques. Plus concrètement, si l'on cherche une matrice Q telle que la contrainte $\tilde{x}Q\tilde{x} \geq 0$ soit valide sur $\mathcal{K} = \{x \in \mathbb{R}^n : \tilde{x}Q_j\tilde{x} \geq 0, j = 1, \dots, m\}$, on peut approximer cette condition de façon conservative par l'application du S-Lemma, c'est-à-dire par la condition qu'il existe $\lambda_j \geq 0, j = 1, \dots, m$ tels que $Q - \sum_{j=1}^m \lambda_j Q_j \succcurlyeq 0$. Cette approximation conservative est au coeur de la relaxation SDP de problèmes quadratiques et d'approximation des problèmes de moments d'ordre 2.

Présentation des problèmes de management d'énergie

Ce paragraphe vise à introduire le management d'énergie, avec une attention particulière portée au problème de planification des arrêts des centrales nucléaires.

Le management d'énergie regroupe l'ensemble des problèmes relatifs à la production, à l'approvisionnement, au transport et à la consommation d'énergie, plus particulièrement l'électricité et le gaz naturel. De façon très simplifiée, on s'attache à satisfaire l'équilibre offre-demande à tout moment et sur tout point du réseau, à moindre coût. Du fait de l'importance stratégique de l'énergie dans notre société, les enjeux associés à ces problèmes sont colossaux, aussi bien sur le plan économique qu'industriel, social et écologique. De par sa taille, le problème d'optimisation sous-jacent est un problème difficile. A ceci s'ajoute un contexte géo-économico-politique très changeant, pour ne pas dire aléatoire, la nécessité de prendre en compte de nombreuses subtilités technologiques sur le fonctionnement des moyens de production et de transport, la modélisation de mécanismes de marché complexes, soumis à une demande très inélastique, et l'importance de l'impact climatique à la fois sur la consommation et sur les moyens de productions. Concernant la gestion de l'électricité, à laquelle nous nous limiterons désormais, le levier de gestion que constitue habituellement la constitution de réserve est très contraint. En effet, il n'est pas rentable de stocker l'énergie directement sous forme électrique (batteries, ...). En revanche, les stocks d'eau des barrages constituent une réserve d'électricité puisqu'on peut facilement (et gratuitement) les convertir sous cette forme. La difficulté est que la majeure partie de cette réserve provient des apports climatiques et est donc soumise à des aléas.

Le respect de l'équilibre offre-demande donne lieu à un gigantesque problème d'optimisation. Afin de le pouvoir traiter, une première décomposition consiste à regrouper les variables de décisions par horizon de temps :

- Sur le long-terme (de 10 à 20 ans) sont prises les décisions d'investissement qui déterminent la structure du portefeuille ;
- Au moyen-terme (de 1 à 5 ans), il est nécessaire de planifier l'utilisation de certains actifs, comme les centrales nucléaires, et en particulier leurs arrêts pour rechargement et maintenance (ce problème sera détaillé par la suite), les stratégies d'utilisation des barrages, ou encore de souscrire certains contrats d'approvisionnement ;
- A court-terme se réalise l'équilibre offre-demande proprement dit, via l'élaboration des programmes de fonctionnement de chaque centrale, complétés par des achats/ventes sur les marchés de l'électricité.

Cette décomposition donne lieu à un grand nombre de sous-problèmes d'optimisation, chacun se distinguant par une finesse de modélisation différente et des difficultés variées.

Parmi elles, et ce d'autant plus que le problème se situe à long-terme, se trouve la prise en compte des incertitudes pouvant affecter les données d'entrées, comme les aléas climatiques, les indisponibilités des moyens de production ou les prix sur les différents marchés de l'énergie. Ces données sont autant de variables aléatoires à prendre en compte dans les modèles, dont il est difficile de déterminer précisément la distribution de probabilité du fait de la complexité des processus impliqués. Cependant, les observations historiques de ces processus nous fournissent une connaissance partielle de ces lois de distributions, menant aux représentations suivantes :

- Une approximation déterministe utilisant la valeur moyenne ou la valeur dans le pire cas ;
- Une représentation robuste établissant que la variable évolue dans un ensemble donné ;
- Une représentation "distributionnellement robuste" dans laquelle on suppose connus le support et les k premiers moments de la distribution de probabilité ;
- Une représentation stochastique dans laquelle on suppose connue la distribution de probabilité. En particulier, c'est le cas lorsqu'on utilise des scénarios de réalisations de la variable, issus par exemple des observations historiques.

Parmi les différents problèmes de management d'énergie, nous nous intéressons particulièrement au problème de la planification des arrêts des centrales nucléaires.

L'objectif de ce problème est de déterminer, sur un horizon de temps à moyen-terme (1 à 5 ans), le meilleur moment pour arrêter les réacteurs afin d'y effectuer les opérations nécessaires de rechargement en combustible et de maintenance, de façon à perturber le moins possible la satisfaction de l'équilibre offre-demande. De par l'importance de la production nucléaire en France, ce problème présente des enjeux financiers importants. Il est difficile à résoudre du fait de sa nature combinatoire, liée à la modélisation de l'état "en marche" ou "en arrêt" des centrales.

L'horizon de temps considéré comporte N_t pas de temps et un parc nucléaire de N_ν réacteurs. La vie d'un réacteur $i \in \mathcal{N}_\nu$, avec $\mathcal{N}_\nu = \{1, \dots, N_\nu\}$ se décompose en cycles indicés par $j = 1, \dots, J_i$ sur l'horizon de temps, chaque cycle étant constitué d'une campagne de production suivi d'un arrêt de durée $\delta_{i,j}$. L'arrêt du cycle j peut débuter à toute date de l'ensemble $\mathcal{E}_{(i,j)} \subset \{1, \dots, N_t\}$ et pour chaque date possible, on définit une variable binaire $x_{i,j,t}$, valant 1 si et seulement si l'arrêt débute effectivement à cette date. On doit alors satisfaire une contrainte dite *d'affectation* imposant à chaque cycle (i, j) une et une seule date d'arrêt : $\sum_{t \in \mathcal{E}_{(i,j)}} x_{i,j,t} = 1$. On déduit de ces variables deux grandeurs essentielles pour notre modèle. Tout d'abord la date effective du début de l'arrêt du cycle (i, j) qui vaut $t_{i,j} = \sum_{t \in \mathcal{E}_{(i,j)}} tx_{i,j,t}$. Puis la *disponibilité nucléaire*, c'est-à-dire la puissance totale des tranches n'étant pas arrêtée au pas de temps t , qui s'exprime également comme une fonction affine de x .

La contrainte d'équilibre offre-demande n'est pas prise en compte explicitement dans le modèle. En fait, on suppose qu'il existe toujours suffisamment de moyens de productions alternatifs et d'offre de vente sur le marché pour satisfaire la demande, mais que leur utilisation se traduit par un coût croissant de

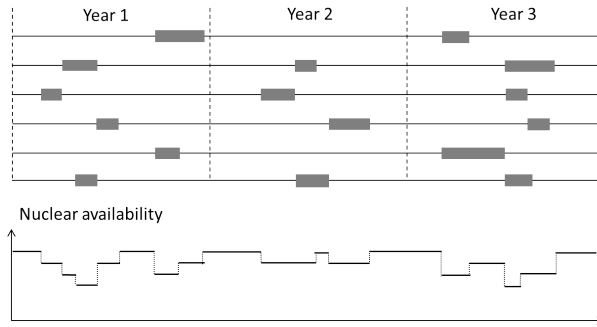


FIGURE 1 – Un exemple de planning et de disponibilité nucléaire

satisfaction de l'équilibre offre-demande. Ce coût est à l'objectif que l'on cherche à minimiser. Il prend la forme d'une fonction convexe, linéaire par morceaux, de la disponibilité nucléaire. Dans un objectif de concision du modèle, cette fonction pourra être approximée par une fonction convexe quadratique.

Les nombreuses contraintes du problème sont liées principalement aux exigences de sûreté et à la disponibilité des ressources utilisées pendant les arrêts. Ces contraintes peuvent nécessiter l'introduction de variables continues, modélisant par exemple la production de la tranche pendant le cycle. Elles sont généralement linéaires, sauf la contrainte dite de *recouvrement maximal*, qui impose à certaines paires d'arrêts (i, j) et (i', j') de ne pas dépasser une certaine valeur N de recouvrement. Cette contrainte peut être vue comme une disjonction : $t_{i,j} - t_{i',j'} < N - \delta_{i,j}$ ou $t_{i,j} - t_{i',j'} > -N + \delta_{i',j'}$. Elle admet 3 modélisations, dont une l'une est quadratique :

- l'exclusion 2 à 2, qui interdit toutes les paires d'instanciation menant à une violation de cette contrainte : $x_{i,j,t} + x_{i',j',t'} \leq 1$ pour tout t, t' tels que $t - t' \in]N - \delta_{i,j}, -N + \delta_{i',j'}[$;
- la formulation "bigM", qui repose sur l'introduction d'une variable binaire $z \in \{0, 1\}$, valant 0 ou 1 selon la partie de la disjonction qui est satisfaite : $t_{i,j} - t_{i',j'} \leq N - \delta_{i',j'} + M_1 z$ et $t_{i,j} - t_{i',j'} \geq -N + \delta_{i,j} - M_2(1 - z)$
- la formulation quadratique $(t_{i,j} - t_{i',j'} - N + \delta_{i,j})(t_{i,j} - t_{i',j'} + N - \delta_{i',j'}) \geq 0$.

Enfin, nous verrons que les différentes variantes de ce problème étudiées dans cette thèse varient également dans la façon dont sont pris en compte les aléas qui affectent la demande et la disponibilité des centrales nucléaires.

Ce problème est donc une application de choix pour notre étude, puisqu'il présente un caractère combinatoire et est soumis à des incertitudes. Nous commencerons par prendre en compte l'aspect combinatoire dans l'axe 1, avant de traiter l'incertitude dans l'axe 2.

Axe 1 : la SDP pour la relaxation de problèmes combinatoires et quadratiques

Le premier axe de cette thèse vise à utiliser la SDP pour produire des relaxations de problèmes combinatoires et quadratiques, comme suggéré par les nombreux résultats prometteurs dans ce domaine. L'obtention de ces relaxations commence généralement par l'implémentation de la relaxation SDP standard, qui doit ensuite être renforcée au moyen de coupes. Celles-ci peuvent être déterminées par l'étude de la structure du problème ou à l'aide de méthodes plus systématiques. Cette approche sera expérimentée sur différentes versions du problème de planification des arrêts nucléaires décrit ci-dessus, choisi pour sa nature combinatoire et ses possibles composantes quadratiques.

Nous commençons par utiliser une modélisation relativement complète, proche de celle utilisée en exploitation, donnant lieu à un problème quadratique à variables mixtes (MIQP pour Mixed Integer Quadratic Program) de grande taille auquel nous allons appliquer une relaxation SDP. Plus précisément, le modèle étudié correspond à une version déterministe du problème, pour laquelle l'objectif est une fonction quadratique de la disponibilité nucléaire et la formulation de la contrainte de recouvrement maximale est

linéaire, correspondant à la formulation "bigM". Parmi les contraintes, la contrainte d'affectation est utilisée pour générer des contraintes quadratiques valides permettant de renforcer la relaxation SDP. Suivant le principe de Sherali-Adams [25], on la multiplie par l'une des variables binaires impliquées, ce qui donne lieu à la contrainte quadratique suivante : $\sum_{t' \in \mathcal{E}_{i,j}, t' \neq t} x_{i,j,t} x_{i,j,t'} = 0$, que l'on ajoute à la relaxation SDP.

On fait suivre la relaxation SDP d'une procédure d'arrondi randomisé permettant d'obtenir une solution entière. Le principe de cet arrondi est d'interpréter la valeur d'une variable dans la solution relaxée comme la probabilité que la variable vaille 1. On tire alors des solutions entières suivant la loi de probabilité ainsi obtenue, jusqu'à obtenir une solution qui satisfasse toutes les contraintes.

Supposons que la relaxation produise une valeur optimale p_r et que la valeur optimale du problème initial soit p^* . Alors la qualité de la relaxation se mesure à son gap, égal à $(p^* - p_r)/p_r$, qui doit être le plus faible possible.

Sur la figure 2, on reporte le gap et la valeur de la solution arrondie obtenue, pour la relaxation SDP standard (*SDP*), la relaxation SDP renforcée (*SDP-R*) et la relaxation obtenue en relaxant la contrainte d'intégrité des variables binaires (*QP*), pour des jeux de données avec $N_t = 156$ et $N_\nu \leq 20$, menant à des problèmes d'environ 200 à 1300 variables binaires. On observe donc que les relaxations SDP donnent des gaps plus faibles, donc meilleurs, que la relaxation QP.

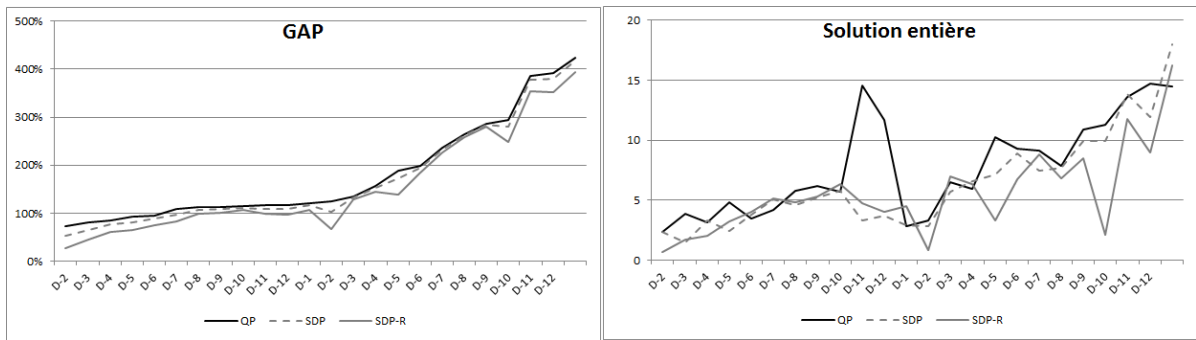


FIGURE 2 – Résultats de la relaxation SDP et de l'arrondi randomisé sur le problème des arrêts nucléaires

En conclusion, le gap moyen est réduit de 1.80% à 1.71% pour la relaxation SDP standard, et jusqu'à 1.56% pour la relaxation SDP renforcée, tout en conservant des temps de calculs raisonnables ($\leq 1030s$ pour SDP et $\leq 2231s$ pour SDP-R). Ces améliorations peuvent sembler faibles, mais elles sont à comparer aux gaps, eux-mêmes faibles, du fait d'une importante part constante dans la fonction objectif. Le gain sur l'arrondi randomisé est également significatif puisqu'il permet de ramener la perte d'optimalité de 7.75% à 6.41% et 5.59% pour les relaxations QP, SDP et SDP-R respectivement.

Suite à cette première expérimentation, nous proposons une méthode systématique permettant de déterminer une relaxation SDP renforcée d'un QCQP à variables bornées. Puis nous appliquons cette méthode à une version allégée du problème de planification des arrêts nucléaires, donnant lieu à un 0/1-QCQP, incluant un objectif quadratique, des contraintes quadratiques correspondant à la formulation quadratique de la contrainte de recouvrement maximal et des contraintes linéaires d'égalité et d'inégalité.

La méthode utilisée pour obtenir la relaxation SDP repose sur le constat que toutes les contraintes quadratiques valides proposées dans la littérature pour renforcer la relaxation SDP peuvent être vues comme des combinaisons linéaires de contraintes quadratiques et de produits deux à deux des contraintes linéaires du problème, y compris les contraintes de borne. Alors, parmi toutes ces contraintes, il reste à sélectionner la contrainte la plus violée par la relaxation SDP courante, dans l'esprit d'un algorithme de séparation classique. Les différentes étapes de la méthode sont représentées sur la Figure 3.

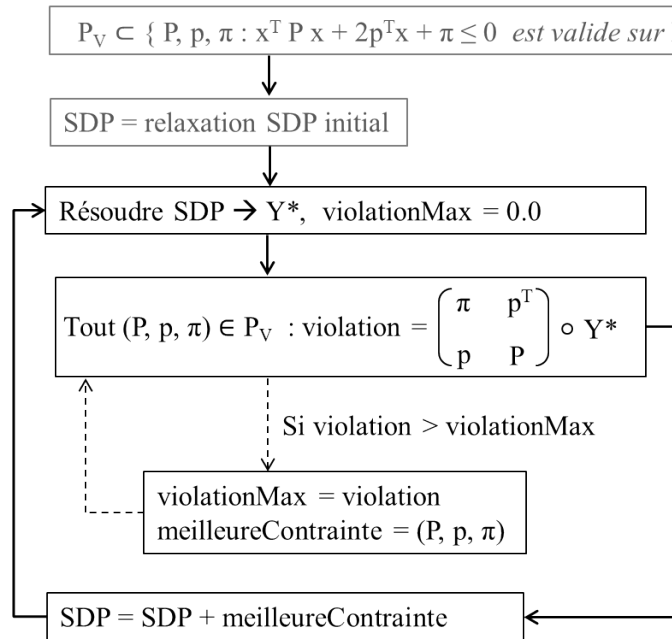


FIGURE 3 – Algorithme de renforcement automatique d’une relaxation SDP

Nous testons également une version alternative de cette phase de séparation, visant à imposer à la contrainte obtenue d’être convexe. Dans ce cas, au lieu de sélectionner une contrainte valide parmi tous les produits deux à deux de contraintes linéaires, on considère toutes les combinaisons positives de ces contraintes et des contraintes initiales du problème. Le problème de séparation étant alors un SDP, cette méthode se révèle trop coûteuse en temps de calcul par rapport au gain obtenu sur les relaxations.

Au-delà du principe même de la méthode, notre contribution sur ce point a consisté à établir un certain nombre de preuves visant à écarter d’office des contraintes valides qui n’apportent rien pour le renforcement de la relaxation SDP. L’exemple le plus typique est la contrainte $(a^T x + b)^2 \geq 0$, valide pour n’importe quelles valeurs de a et b , mais inutile pour la relaxation SDP. Nous avons également montré que notre approche unifie un grand nombre de contraintes valides proposées dans la littérature pour le renforcement de la relaxation SDP [3, 13, 18, 23, 24].

Cette approche est également mise en oeuvre numériquement sur des jeux de données du problème des arrêts nucléaires possédant entre 200 et 1000 variables binaires et de 100 à 500 contraintes, en limitant à 100 le nombre de contraintes quadratiques valides ajoutées. Il en ressort que le gap moyen obtenu est de 6.88% pour la relaxation SDP renforcée, contre 6.97% pour la relaxation SDP standard et 25.76% pour la relaxation linéaire. L’avantage de la méthode est qu’elle permet également d’identifier les contraintes quadratiques les plus intéressantes pour le renforcement de la relaxation SDP. Ainsi on observe que les contraintes choisies en priorité sont celles issues du produit de deux contraintes linéaires qui partagent un grand nombre de variables. Plus précisément, pour une contrainte construite comme le produit des deux contraintes linéaires C_1 et C_2 , on définit son ratio de recouvrement comme le rapport entre le nombre de variables impliqué dans C_1 et C_2 et le nombre de variables impliqué dans C_1 ou C_2 . La décroissance des ratios des contraintes sélectionnées au cours des itérations est représentée sur la Figure 4.

Pour poursuivre sur cet axe d’étude, nous nous attachons à comparer les différentes modélisations de la contrainte de recouvrement maximal du point de vue de la performance des relaxations obtenues. Il nous a également semblé important d’expérimenter la hiérarchie de Lasserre, une théorie très puissante considérant un problème d’optimisation polynomial quelconque et construisant une suite de SDP dont

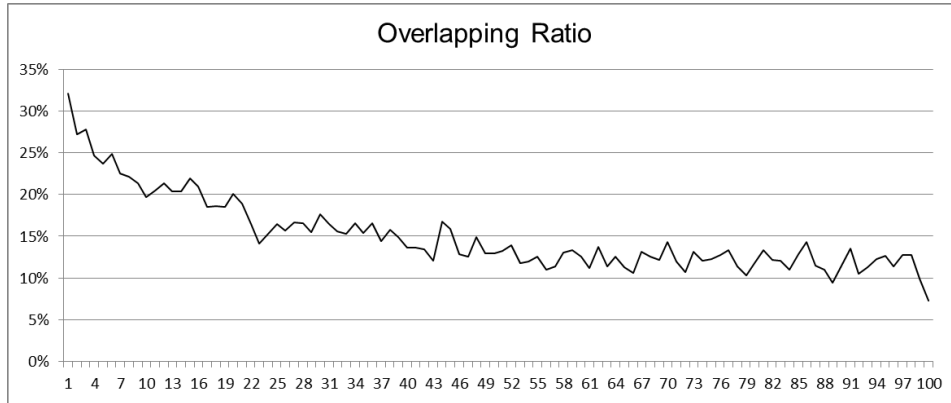


FIGURE 4 – Ratio de recouvrement des contraintes sélectionnées

la valeur optimale tend vers la solution du problème initial. Pour ce faire, nous travaillons sur une version simplifiée du problème de planification des arrêts, dans laquelle ne sont considérées que les contraintes de recouvrement maximal et d'affectation, et dans laquelle l'objectif est une fonction linéaire de la disponibilité nucléaire. 12 relaxations SDP sont testées, différant par :

- la mise au carré des contraintes linéaires ;
- l'ajout des contraintes dite de type *RLT* (Reformulation Linearization Technique) ($x_i x_j \geq 0$, $x_i x_j \leq x_i, x_i x_j \leq x_j$ et $x_i x_j \geq 1 - x_i - x_j$) ;
- l'ajout de contraintes de type Sherali-Adams ;
- l'ajout de contraintes triangulaires.

Pour chacune de ces relaxations SDP, nous construisons la relaxation linéaire équivalente, en linéarisant le QCQP intermédiaire ayant mené à l'obtention de la relaxation SDP. Puis une expérimentation est réalisée sur un ensemble de jeux de données regroupés en classe "i-j", dans lequel i correspond à la taille de l'instance et j à la modélisation utilisées pour la contrainte de recouvrement maximal ($j = 1$ pour la formulation "bigM", $j = 2$ pour l'exclusion 2 à 2 et $j = 3$ pour la formulation quadratique). Les résultats sont calculés en moyenne sur les 100 instances de chaque taille. Parmi toutes les relaxations SDP testées, les trois ci-dessous se démarquent, dont les gaps sont reportés sur la figure 5 :

- SDP-4, obtenues en élevant au carré toutes les contraintes linéaires ;
- SDP-7, obtenues en ajoutant à SDP-4 le produit de toutes les contraintes linéaires par toutes les contraintes de borne, dans l'esprit de Sherali-Adams ;
- SDP-10, obtenues en ajoutant à SDP-7 les contraintes RLT.

Il ressort de cette expérimentation les constats suivants : pour les 3 modélisations, la meilleure relaxation SDP est obtenue tout d'abord en élevant les contraintes linéaires au carré, puis en ajoutant les produits de toutes les contraintes linéaires par toutes les contraintes de borne, puis les contraintes RLT. Parmi les 3 modélisations, l'exclusion 2 à 2 offre le meilleur potentiel pour la relaxation SDP par rapport à la relaxation LP, ce qui s'explique par le fait que le nombre de contraintes linéaires est très grand et que le renforcement est d'autant plus efficace que le nombre de contraintes linéaires est élevé.

Une dernière expérimentation consiste à mettre en oeuvre la hiérarchie de Lasserre sur ce problème. Le lecteur est renvoyé à l'article [16] pour davantage de détails sur cette séquence de problème SDP menant à la solution optimale. Le rang 1 de cette hiérarchie correspond à la relaxation SDP standard et on s'attelle donc à résoudre le rang 2 pour les instances de taille 1. On observe alors que sur toutes les instances, le gap vaut 0.0%, ou autrement dit, le SDP fournit la valeur optimale du problème entier, et ceci avec des temps de calculs inférieurs à 10 s. Cette hiérarchie tient donc ses promesses en ce qui concerne la qualité de la borne mais il demeure une difficulté majeure quant à son applicabilité à des

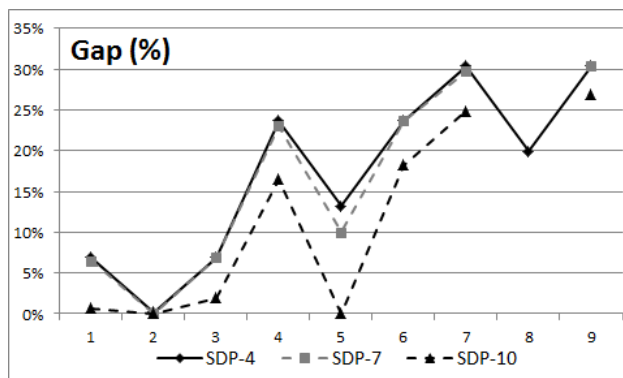


FIGURE 5 – Comparaison des gaps des relaxations SDP-4, SDP-7 et SDP-10

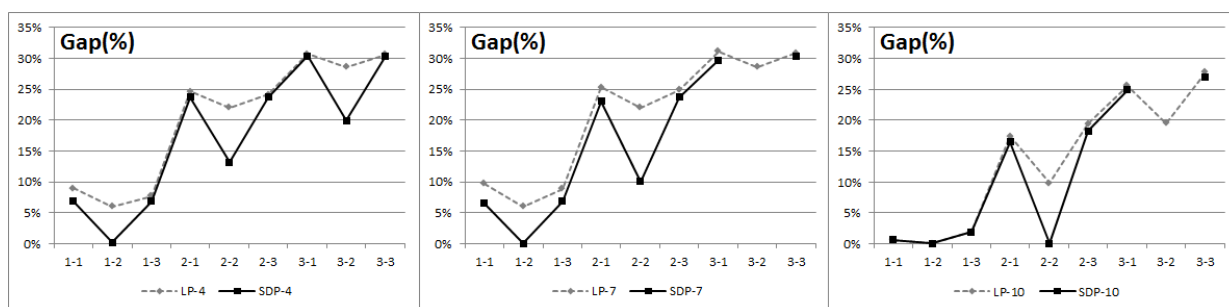


FIGURE 6 – Comparaison des gaps SDP et LP pour les relaxations 4, 7 et 10

problèmes de plus grande taille, puisqu'en l'état des solveurs, il n'est pas possible de dépasser un nombre de variables de l'ordre de 30.

Axe 2 : la SDP pour la prise en compte de l'incertitude

Le second axe de la thèse, portant sur l'application de la SDP à la prise en compte de l'incertitude, se décline en trois études. Dans la première étude, nous travaillons sur une version du problème des arrêts nucléaires dans laquelle l'incertitude se présente sous la forme de scénarios équiprobables et les contraintes concernées par les incertitudes sont à satisfaire en probabilité. Il est alors classique de formuler un équivalent déterministe de ce problème comme un problème combinatoire de grande taille en ajoutant une variable binaire par contrainte et par scénario. Nous appliquons alors la relaxation SDP standard à ce problème et la faisons suivre d'un arrondi randomisé.

Le problème considéré correspond au modèle décrit ci-dessus, dans lequel l'objectif et les contraintes de recouvrement maximal sont quadratiques. Les expérimentations sont menées avec 10 scénarios, sur des jeux de données comportant de 700 à 2400 variables binaires. Les résultats sont illustrés sur la courbe de la figure 7, dans laquelle la solution entière représente l'écart entre la solution de l'arrondi randomisé et la meilleure solution entière trouvée par CPLEX en 1800 s. Ces résultats sont très prometteurs, avec un gap moyen de 2.76% avec la relaxation SDP contre 53.35% pour la relaxation linéaire, et ceci sans le moindre renforcement.

La deuxième étude met en oeuvre une méthode plus ambitieuse pour la prise en compte de l'incertitude, permettant de reformuler le problème, ou d'en donner une approximation conservative, sous la forme d'un SDP. Cette méthode, qui a fait l'objet de nombreux travaux récemment, est connue sous le nom d'optimisation distributionnellement robuste. Il s'agit en fait d'un compromis entre l'optimisation

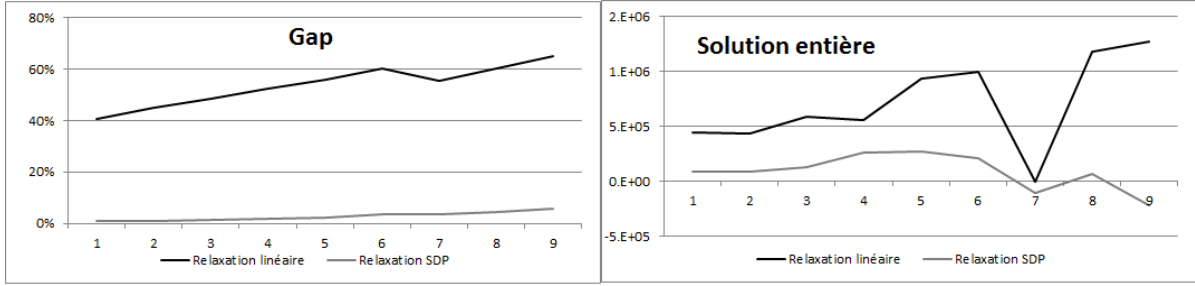


FIGURE 7 – Résultats de la relaxation SDP et de l'arrondi randomisé sur le problème des arrêts des centrales nucléaire stochastique

stochastique, qui nécessite une connaissance parfaite des lois de probabilités utilisées, et l'optimisation robuste, qui ne requiert que la connaissance du support des variables aléatoires. En effet, l'optimisation distributionnellement robuste ne nécessite pas de connaître la distribution de probabilité, mais uniquement son support et certains de ses moments. Nous appliquons cette méthode à un problème d'équilibre offre-demande dans lequel la demande et la disponibilité des moyens de production sont soumis à des aléas, dont on connaît le support, l'espérance et la covariance. Notre objectif est d'estimer l'apport de cette méthode par rapport à une méthode de type robuste basée sur la connaissance du support et de l'espérance, permettant d'obtenir une approximation conservative du problème sous forme d'un SOCP.

Le problème d'équilibre offre-demande considéré peut se mettre sous la forme d'un problème à contrainte en probabilité jointe linéaire :

$$\min_{x \in F} \{c^T x : \mathbb{P}[g(x, \xi) \leq 0] \geq 1 - \varepsilon\}$$

avec :

- c un vecteur de \mathbb{R} . et $F \subset \mathbb{R}^n$ un polyèdre;
- $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^T$ une fonction affine de x et de ξ telle que $g(x, \xi)_t = \tilde{x}^T A_t \tilde{\xi}$;
- ξ un vecteur aléatoire de \mathbb{R}^m dont on connaît le support $\mathcal{S} = \{\xi \in \mathbb{R}^m : a_i \leq \xi_i \leq b_i\}$, l'espérance μ et la covariance Σ ;
- ε le niveau de probabilité requis quant à la satisfaction de la contrainte.

On définit $\mathcal{P}(\mathcal{S})$ comme l'ensemble des distributions de probabilités de support \mathcal{S} , d'espérance μ et de covariance Σ . Plus précisément $\mathcal{P}(\mathcal{S}) = \{P \in \mathcal{M}(\mathcal{S}) : \Omega_P(\xi) = \Omega\}$ avec $\Omega = \begin{pmatrix} 1 & \mu^T \\ \mu & \mu\mu^T + \Sigma \end{pmatrix}$. On cherche alors à satisfaire la contrainte pour toutes les distributions de $\mathcal{P}(\mathcal{S})$, c'est-à-dire :

$$(C_1) \min_{x \in F} \{c^T x : \mathbb{P}[g(x, \xi) \leq 0] \geq 1 - \varepsilon, \forall P \in \mathcal{P}\} \quad \text{ou} \quad (C_2) \min_{x \in F} \{c^T x : \inf_{P \in \mathcal{P}(\mathcal{S})} \mathbb{P}[g(x, \xi) \leq 0] \geq 1 - \varepsilon\}$$

La formulation (C_2) fait intervenir le problème de moments $\inf_{P \in \mathcal{P}(\mathcal{S})} \mathbb{P}[g(x, \xi) \leq 0]$, auquel s'applique la hiérarchie de Lasserre. Notre contribution se décline alors en trois points. Tout d'abord, nous montrons que le SDP obtenu en appliquant la hiérarchie de Lasserre peut également être obtenu très simplement en utilisant le S-Lemma et qu'il unifie un grand nombre d'approximations SDP proposées dans la littérature, en particulier celle de Zymler et al. [29] mais aussi [7, 8, 10, 27], ainsi que d'inégalités célèbres sur les probabilités, à savoir les inégalités de Markov et de Cantelli. Puis nous donnons des conditions suffisantes pour que cette approximation soit exacte. Enfin nous comparons les bornes obtenues à des méthodes plus classiques, sur des problèmes académiques et d'équilibre offre-demande.

Nous commençons par montrer grâce au S-Lemma qu'une condition suffisante pour que la contrainte en probabilité $\mathbb{P}[g(x, \xi) \leq 0] \geq 1 - \varepsilon$ soit respectée pour tout $P \in \mathcal{P}(\mathcal{S})$ est qu'il existe une matrice M

qui satisfasse le système SDP suivant :

$$\begin{cases} \Omega \bullet M \leq \varepsilon \\ M - \sum_{i=1}^m \lambda_{0,s} W^i \succcurlyeq 0 \\ M - M_0 - \sum_{i=1}^m \lambda_{t,s} W^i + \tau_t Y^t(x) \succcurlyeq 0, \quad t = 1, \dots, T \\ \lambda \geq 0, \quad \tau \geq 0 \end{cases} \quad \text{où} \quad \begin{cases} W^i = \frac{1}{2} \begin{pmatrix} -2a_i b_i & a_i + b_i \\ a_i + b_i & -2 \end{pmatrix} \\ Y^t(x) = \frac{1}{2} \begin{pmatrix} 2\tilde{x}^T A_{*,1} & \tilde{x}^T A_{*,2\dots m+1} \\ A_{*,2\dots m+1}^T \tilde{x} & 0 \end{pmatrix} \\ M_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \end{cases}$$

Nous montrons que cette approximation est exacte dès lors que $T = 1$, c'est-à-dire lorsque la contrainte en probabilité est individuelle. Puis nous procédons à des comparaisons numériques pour le cas particulier où aucune variable de commande x n'est présente ($n = 0$). Il s'agit alors de déterminer une borne inférieure d'une probabilité, avec ou non prise en compte de la covariance, sachant qu'une borne inférieure est meilleure lorsqu'elle est élevée puisque l'approximation associée sera d'autant moins conservative.

Nous nous intéressons en premier lieu au cas $m = 1$ et $g(x, \xi) = \xi$. Alors, sans la prise en compte de la covariance, l'optimisation distributionnellement robuste mène à la borne de Markov, à savoir μ/a , si $[a, b]$ est le support de ξ . Cette borne est comparée aux valeurs obtenues par optimisation distributionnellement robuste avec covariance et à la borne dite "robuste", basée sur la connaissance du support et de l'espérance, permettant de formuler une approximation conservative du problème sous la forme d'un SOCP via l'application de l'inégalité de Hoeffding [10, 28]. Les résultats sont transcrit sur la courbe de la figure 8. Pour chaque jeu de données sont indiquées la borne de Markov, la borne robuste valant $1 - \exp(-2\mu^2/\|b - a\|^2)$, ainsi que la plus petite (*Min var*) et la plus grande (*Max var*) borne obtenue par optimisation distributionnellement robuste avec covariance.

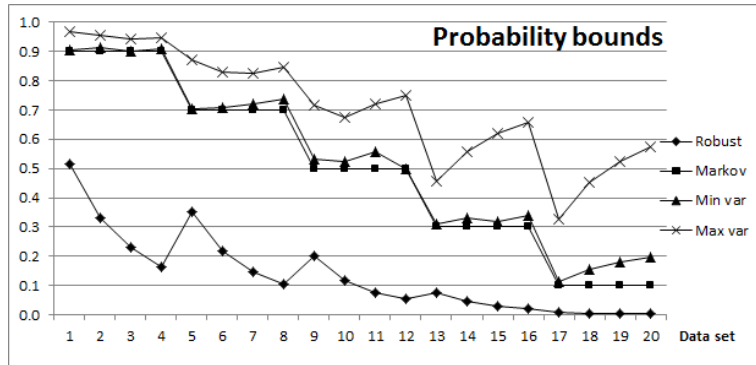


FIGURE 8 – Comparaison de différentes bornes inférieurs de $\min_{P \in \mathcal{P}} P[\xi \leq 0]$

Nous considérons ensuite le cas où $m \geq 1$ et $g(x, \xi) = e^T \xi$. Les jeux de données sont construits aléatoirement, en tirant les valeurs de a_i, b_i, μ_i et de la matrice de covariance. La figure 9 illustre la comparaison entre la borne robuste ($1 - \exp(-2(e^T \mu)^2 / \|b - a\|^2)$), la borne de Markov ($e^T \mu / e^T a$) obtenue en considérant $e^T \xi$ comme une variable aléatoire de moyenne $e^T \mu$ et de support $[e^T a, e^T b]$, et les résultats de l'optimisation distributionnellement robuste avec et sans covariance.

Enfin, il reste à étudier le cas où $T \geq 1$, avec $g_t(x, \xi) = d_t^T \tilde{\xi}$. Pour ce problème, on compare la borne robuste, qui s'avère être négative donc inutile, et les bornes distributionnellement robustes. Les résultats sont représentés sur la courbe 10.

Il ressort de ces comparaisons que les bornes distributionnellement robuste avec covariance sont clairement les plus performantes. Il reste à les appliquer à un véritable problème d'optimisation avec des variables de commandes. Pour cela, on considère le problème d'équilibre offre-demande suivant, dans lequel on minimise un coût de production linéaire, tout en satisfaisant en probabilité l'équilibre offre-

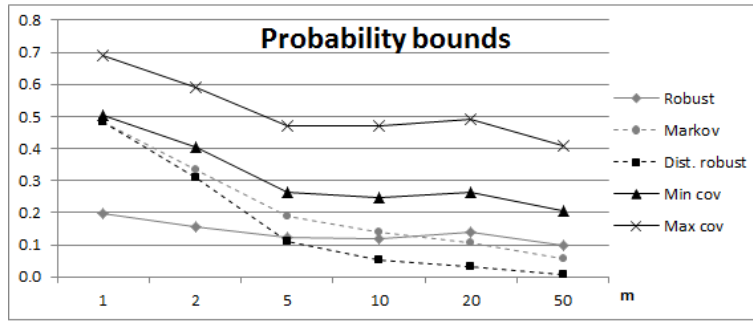


FIGURE 9 – Comparaison de différentes bornes inférieures de $\min_{P \in \mathcal{P}} P[e^T \xi \leq 0]$ pour $m \geq 1$

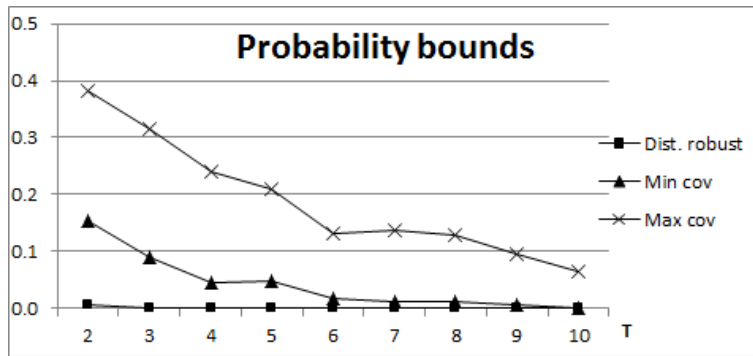


FIGURE 10 – Comparaison de différentes bornes inférieures de $\min_{P \in \mathcal{P}} P[d_t^T \tilde{\xi} \leq 0, t = 1, \dots, T]$

demande à chaque pas de temps :

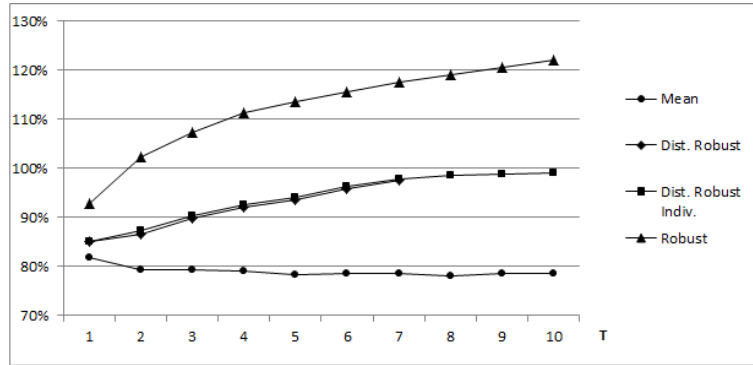
$$\begin{cases} \min & c^T x \\ \text{s.t.} & P \left[\sum_{i=1}^N D_{i,t} x_{t,i} \geq D_{0,t}, t = 1, \dots, T \right] \geq 1 - \varepsilon \\ & \sum_{t=1}^T x_{t,i} \leq r_i, i = 1, \dots, N \\ & x_{t,i} \in [0, 1], i = 1, \dots, N, t = 1, \dots, T \end{cases}$$

où :

- T est le nombre de pas de temps considérés ;
- N est le nombre d'unité de production ;
- $c_{t,i}$ est le coût unitaire de production de l'unité de production i au pas de temps t ;
- $x_{t,i}$ représente la production de l'unité de production i au pas de temps t ;
- $D_{0,t}$ est une variable aléatoire représentant la demande au pas de temps t ;
- $D_{i,t}$ est une variable aléatoire représentant le coefficient de disponibilité de l'unité de production i au pas de temps t ;
- r_i représente la production maximale de l'unité de production i sur l'horizon de temps.

Les données sont calculées sur la base des observations historiques. Pour chaque jeu de données sont comparées :

- p_m^* , la valeur optimale du LP obtenu en remplaçant l'aléa par sa valeur moyenne ;
- p_{dr}^* , la valeur optimale obtenue par optimisation distributionnellement robuste ;

FIGURE 11 – Comparaison du ratio p_w^*/p^* pour $\varepsilon = 0.8$

T	$\varepsilon = 0.5$			$\varepsilon = 0.4$			$\varepsilon = 0.3$			$\varepsilon = 0.2$		
	p_s^*	p_{dri}^*	loss	p_s^*	p_{dri}^*	loss	p_s^*	p_{dri}^*	loss	p_s^*	p_{dri}^*	loss
1	262.9	261.3	-0.59%	267.0	265.6	-0.53%	271.8	271.2	-0.23%	277.5	279.4	0.69%
2	512.9	538.6	5.02%	521.2	548.6	5.26%	530.8	562.5	5.98%	542.5	582.1	7.31%
3	782.3	848.6	8.47%	794.5	866.5	9.06%	808.9	886.8	9.63%	826.5	895.2	8.32%
4	1040.1	1162.7	11.79%	1056.6	1182.8	11.95%	1076.0	1191.1	10.70%	1099.9	1192.7	8.44%
5	1290.6	1473.8	14.19%	1311.4	1483.9	13.15%	1336.0	1484.3	11.10%	1366.3	1489.1	8.99%
6	1538.3	1762.4	14.57%	1563.3	-	-	1592.8	-	-	1629.4	-	-
7	1780.2	2040.6	14.63%	1810.5	-	-	1846.0	-	-	1889.4	-	-
8	2021.3	2325.5	15.05%	2057.1	-	-	2098.9	-	-	2149.9	-	-
9	2263.9	-	-	2305.7	-	-	2354.1	-	-	2412.9	-	-
10	2513.0	-	-	2560.1	-	-	2614.8	-	-	2681.1	-	-

TABLE 1 – Comparaison avec l'approche stochastique

- p_{dr}^* , la valeur optimale obtenue par optimisation distributionnellement robuste, avec approximation de la probabilité jointe par une somme de probabilités individuelles, via l'inégalité de Boole ;
- p_r^* , la valeur optimale obtenue par approche robuste ;
- p_w^* , la valeur optimale du LP obtenu en remplaçant l'aléa par sa valeur dans le pire cas.

Pour $\varepsilon = 0.8$, les résultats obtenus sont reportés sur la figure 11.

Nous constatons que la perte liée à l'approximation de la probabilité jointe par une somme de probabilité individuelle est faible. De plus, la borne distributionnellement robuste est nettement meilleure que la borne robuste, ce qui peut s'expliquer par le fait qu'elle exploite une information supplémentaire, à savoir la covariance.

Nous procédons finalement à une comparaison avec une approche stochastique, basée sur l'hypothèse que $g(x, \xi)$ suit une loi normale. La probabilité jointe est approximée par une probabilité individuelle, comme déjà évoqué. Nous observons que la perte d'optimalité liée à l'approche distributionnellement robuste n'est pas aussi grande qu'on aurait pu le craindre. En particulier, pour $T = 1$, cette approche est même meilleure que l'approche stochastique, du fait sans doute qu'on y exploite une information supplémentaire : le support.

Cet axe d'étude se termine par une troisième étude combinant les axes combinatoires et stochastiques explorés préalablement. Il s'avère que de nombreux problèmes aléatoires admettent un équivalent déterministe, ou une approximation, sous forme d'un SOCP et dans le cas où le problème initial comporte des variables entières, le problème obtenu est alors un MISOCP, pour lesquels il n'existe pas de méthodes de résolution de référence. Nous nous intéressons ici à l'utilisation de relaxations SDP pour ce type de problème. Le principe est de convertir le MISOCP en MIQCQP, puis d'appliquer la relaxation SDP standard, qui est ensuite renforcée par l'ajout de contraintes du problème initial mises au format

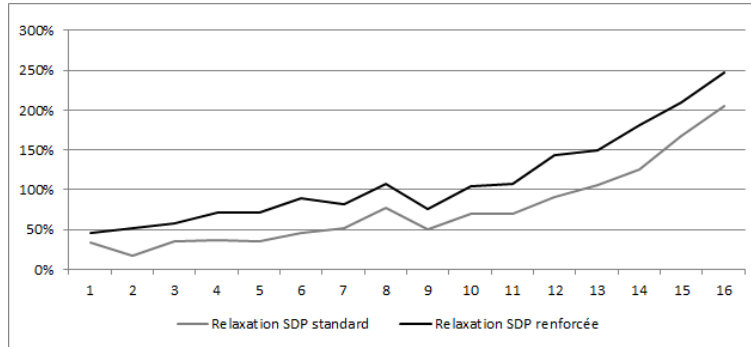


FIGURE 12 – Amélioration de la relaxation SDP par rapport à la relaxation continue

SDP.

Cet ajout de contraintes est rendu nécessaire par le fait que la formulation d'un SOCP comme un QCQP est généralement synonyme d'une perte de convexité. En effet, cette reformulation suit le principe suivant :

$$\|Ax + b\| \leq c^T x + d \Leftrightarrow \begin{cases} x^T(A^T A - cc^T)x + 2(b^T Ax - dc^T x) + b^T b - d^2 \leq 0 \\ c^T x + d \geq 0 \end{cases}$$

Or la matrice $A^T A - cc^T$ n'est généralement pas semi-définie positive. On montre le résultat suivant :

Proposition Soit $A \in \mathbb{R}^{m,n}$ une matrice de rang plein et $c \in \mathbb{R}^{n,1}$. Alors la matrice symétrique $A^T A - cc^T$ est semi-définie positive si et seulement si il existe $u \in \mathbb{R}^{m,1}$, avec $\|u\| \leq 1$, tel que $c = A^T u$.

Cette conversion d'un SOCP en QCQP illustre parfaitement la différence entre convexité et convexité abstraite. En effet, l'ensemble réalisable du QCQP est nécessairement convexe, puisqu'il est équivalent à l'ensemble réalisable du SOCP donc le QCQP est convexe au sens abstrait. Cependant, cet ensemble n'est pas décrit à l'aide de contraintes convexes, donc le problème n'est pas convexe.

Afin d'exploiter la structure particulière du QCQP ainsi obtenue, et de restaurer sa convexité, nous renforçons la relaxation SDP standard au moyen des contraintes SDP obtenues en écrivant les contraintes SOCP directement sous forme SDP, suivant l'équivalence bien connue :

$$\|Ax + b\| \leq c^T x + d \Leftrightarrow \begin{pmatrix} (c^T x + d)I & Ax + b \\ (Ax + b)^T & c^T x + d \end{pmatrix} \succcurlyeq 0$$

On obtient ainsi la borne une relaxation SDP standard et une relaxation SDP renforcée. Pour chacune d'entre elle, on calcule l'amélioration par rapport à la relaxation continue, comme suit : $r = \frac{p_s - p_c}{p_c}$, où p_s est la borne SDP et p_c la borne continue. On trace ces deux indicateurs sur la figure 12.

Cette approche donne des résultats encourageants, avec une relaxation SDP nettement plus performante que la relaxation continue. Les solutions obtenues sont même très proches de l'optimal sur de nombreuses instances tout en conservant un temps de calcul raisonnable, de l'ordre de quelques minutes pour des instances à quelques centaines de variables binaires.

Conclusions et perspectives

Cette thèse avait pour objet d'évaluer les potentialités de la programmation semi-définie positive (SDP) pour les problèmes d'optimisation issus du management d'énergie. Nous avons montré qu'il existe de nombreuses opportunités d'innovation pour le traitement de ces problèmes par la SDP, en particulier pour les problèmes présentant un caractère combinatoire ou quadratique, ou pour la prise en compte de l'aléa.

Concernant l'axe quadratique et combinatoire, la SDP fournit des relaxations convexes de ces problèmes NP-difficiles. La difficulté est alors de déterminer le bon niveau de compromis entre taille du SDP obtenu et qualité de la relaxation. A un extrême se situe la relaxation SDP standard, simple à obtenir et présentant quasiment le même nombre de contraintes que le problème initial. Cette relaxation n'est généralement pas très performante. A l'autre extrême se trouve les relaxations SDP obtenues en appliquant la hiérarchie de Lasserre. Celles-ci sont extrêmement performantes mais sont malheureusement de trop grande taille pour être résolues pour des problèmes de plus de 30 variables binaires. Enfin, des relaxations intermédiaires peuvent être construites en ajoutant des contraintes quadratiques valides au problème initial, obtenues par exemple en multipliant des contraintes linéaires valides entre elles.

Du côté de la prise en compte de l'aléa, la SDP se prête plutôt aux approches robustes et distributionnellement robuste. En effet, certains résultats bien connus de la littérature montrent que des problèmes robustes peuvent s'écrire sous forme de problème conique du second-ordre, qui est un cas particulier de SDP. Nous avons étendu cette équivalence à des problèmes possédant des variables binaires, pour lesquels l'équivalent robuste est donc un problème conique du second-ordre à variables binaires et nous avons montré que la SDP produit de très bonnes relaxations de ces problèmes difficiles. Dans le cadre de l'optimisation distributionnellement robuste, de nombreux travaux proposent des approximations conservatrices de contraintes en probabilité sous forme d'inégalités linéaires matricielles. Afin d'offrir une vision claire sur ces approches, nous les avons unifié et présentons un certain nombre d'expérimentations numériques illustrant la pertinence de cette façon de prendre en compte l'aléa.

Ainsi, en dépit de nombreuses difficultés pratiques imputables au fait que cette technologie n'est pas encore tout à fait mature, la SDP n'en reste pas moins une méthode extrêmement prometteuse, combinant toute la puissance de l'optimisation conique. De nombreuses applications méritent encore d'être explorées, aussi aussi bien en management d'énergie que dans d'autres domaines tels que l'ingénierie ou la gestion de portefeuille d'actifs financiers.

Bibliographie

- [1] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5 :13–51, 1993.
- [2] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton. Complementarity and nondegeneracy in semidefinite programming. *Mathematical Programming*, 77(2) :111–128, 1997.
- [3] K. M. Anstreicher. Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. *Journal of Global Optimization*, 43(2–3) :471–484, 2009.
- [4] A. Barvinok. Problems of distance geometry and convex properties of quadratic maps. *Discrete and Computational Geometry*, 13 :189–202, 1995.
- [5] A. Ben-Israel, A. Charnes, and K. Kortanek. Duality and asymptotic solvability over cones. *Bulletin of American Mathematical Society*, 75(2) :318–324, 1969.
- [6] S. J. Benson, Y. Ye, and X. Zhang. Solving large-scale sparse semidefinite programs for combinatorial optimization. *SIAM Journal on Optimization*, 10(2) :443–461, 2000.
- [7] D. Bertsimas and I. Popescu. On the relation between option and stock prices : A convex optimization approach. *Operations Research*, 50(2) :358–374, 2002.
- [8] D. Bertsimas and I. Popescu. Optimal inequalities in probability theory : a convex optimization approach. *SIAM Journal on Optimization*, 15(3) :780–804, 2005.
- [9] B. Borchers. CSDP, a C library for semidefinite programming. *Optimization Methods and Software*, 11, 1999.
- [10] G. Calafiore and L. El Ghaoui. On distributionally robust chance-constrained linear programs. *Journal of Optimization Theory and Applications*, 130(1) :1–22, 2006.
- [11] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6) :1115–1145, 1995.
- [12] L. S. A. Grötschel, Martin; Lovász. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2) :169–197, 1981.
- [13] C. Helmberg. Semidefinite programming for combinatorial optimization. Habilitationsschrift, TU Berlin, Konrad-Zuse-Zentrum, 2000.
- [14] C. Helmberg and F. Rendl. A spectral bundle method for semidefinite programming. *SIAM Journal on Optimization*, 10 :673–696, 2000.
- [15] M. Kocvara and M. Stingl. On the solution of large-scale SDP problems by the modified barrier method using iterative solvers. *Mathematical Programming, Series B*, 109(2–3) :413–444, 2007.

- [16] J.-B. Lasserre. A semidefinite programming approach to the generalized problem of moments. *Mathematical Programming*, 112 :65–92, 2008.
- [17] M. Laurent and F. Rendl. *Handbook on Discrete Optimization*, chapter Semidefinite Programming and Integer Programming, pages 393–514. Elsevier B.V., 2005.
- [18] C. Lemaréchal and F. Oustry. Semidefinite relaxations and lagrangian duality with application to combinatorial optimization. Technical report, INRIA, 1999.
- [19] L. Lovász. On the shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(1) :1–7, 1979.
- [20] Y. Nesterov and A. Nemirovski. *Interior-point polynomial methods in convex programming*. SIAM, 1994.
- [21] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. *Mathematics of operations research*, 23(2), 1998.
- [22] I. Polik and T. Terlaky. A survey of the s-lemma. *SIAM Review*, 49(3) :371–418, 2007.
- [23] S. Poljak, F. Rendl, and H. Wolkowicz. A recipe for semidefinite relaxation for (0,1)-quadratic programming. *Journal of Global Optimization*, 7 :51–73, 1995.
- [24] F. Roupin. From linear to semidefinite programming : an algorithm to obtain semidefinite relaxations for bivalent quadratic problems. *Journal of Combinatorial Optimization*, 8(4) :469–493, 2004.
- [25] H. D. Sherali and W. P. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3(3) :411–430, 1990.
- [26] N. Shor. Quadratic optimization problems. *Soviet Journal of computer Systems Sciences*, 25 :1–11, 1987.
- [27] L. Vandenberghe, S. Boyd, and K. Comanor. Generalized chebyshev bounds via semidefinite programming. *SIAM Review*, 49(1) :52–64, 2007.
- [28] R. Zorgati, W. Van Ackooij, and A. Gorge. Uncertainties on power systems. probabilistic approach and conic approximation. *IEEE 11th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 2010.
- [29] S. Zymler, D. Kuhn, and B. Rustem. Distributionnally robust joint chance constraint with second-order moment information. *Mathematical Programming*, 2011.