

AUTOUR DE NOUVELLES NOTIONS POUR L'ANALYSE
DES ALGORITHMES D'APPROXIMATION :
FORMALISME UNIFIÉ ET CLASSES D'APPROXIMATION

MARC DEMANGE¹ ET VANGELIS PASCHOS²

Communiqué par Pierre Tolla

Abstract. The main objective of the polynomial approximation is the development of polynomial time algorithms for **NP**-hard problems, these algorithms guaranteeing feasible solutions lying “as near as possible” to the optimal ones. This work is the first part of a couple of papers where we introduce the key-concepts of the polynomial approximation and present the main lines of a new formalism. Our purposes are, on the one hand, to present this theory and its objectives and, on the other hand, to discuss the appropriateness and the pertinence of its constitutive elements, as people knew them until now, and to propose their enrichment. Henceforth, these papers are addressed to both domain researchers and non-specialist readers. We particularly quote the great theoretical and operational interest in constructing an internal structure for the class **NPO** (of the optimization problems in **NP**). In this first part, we focus on some basic tools allowing the individual evaluation of the approximability properties of any **NP**-hard problem. We present and discuss notions as algorithmic chain, approximation level, hardness threshold and two notions of limits (with respect to algorithmic chains and with respect to problems instances). The notions dealt in the paper are presented together with several illustrative examples.

Reçu en novembre 2001.

¹ ESSEC, Cergy-Pontoise, France ; e-mail : demange@essec.fr

² LAMSADE, Université Paris-Dauphine, France ; e-mail : paschos@lamsade.dauphine.fr

© EDP Sciences 2003

Résumé. Cet article est le premier d'une série de deux articles où nous présentons les principales caractéristiques d'un nouveau formalisme pour l'approximation polynomiale (algorithmique polynomiale à garanties de performances pour les problèmes **NP**-difficiles). Ce travail est l'occasion d'un regard critique sur ce domaine et de discussions sur la pertinence des notions usuelles. Il est aussi l'occasion de se familiariser avec l'approximation polynomiale, de comprendre ses enjeux et ses méthodes. Ces deux articles s'adressent donc autant aux spécialistes qu'aux non spécialistes de ce domaine. Nous insistons tout particulièrement sur l'intérêt, tant théorique qu'opérationnel, de mettre en évidence une structure au sein de la classe **NPO** des problèmes d'optimisation de **NP**. Dans ce premier article, nous nous intéressons aux outils qui permettent d'évaluer, dans l'absolu, les propriétés d'approximation de problèmes difficiles. Nous discutons notamment les notions de chaînes d'approximation, de niveau d'approximation, d'ordre de difficulté ainsi que deux notions de limites (par rapport à une suite d'algorithmes et par rapport aux instances). Chaque notion est largement discutée et illustrée par de nombreux exemples choisis essentiellement pour leur valeur pédagogique.

Mots Clés. Complexité, difficulté intrinsèque, analyse des algorithmes et des problèmes, algorithmes d'approximation.

Classification Mathématique. 68Q15, 68Q17, 68Q25, 68W25.

1. INTRODUCTION

L'approximation polynomiale a pour objet la conception et l'analyse d'algorithmes polynomiaux avec garanties de performances pour des problèmes difficiles n'admettant pas d'algorithme exact polynomial à moins que $\mathbf{P} = \mathbf{NP}$. Nous proposons, dans cet article, une promenade au sein d'un nouveau formalisme pour l'algorithmique à garanties de performances. Elle est l'occasion de découvrir certains aspects de cette théorie et d'engager une réflexion sur ses enjeux et sur la pertinence de ses concepts. La démarche et les raisons qui nous ont conduit à proposer ce formalisme nous semblent intéressants, non seulement pour tout spécialiste du domaine, mais aussi pour quiconque désirant en avoir un bref aperçu. Aux premiers nous proposons les rudiments d'un débat auquel nous les invitons à se joindre. Ce débat nous paraît important pour le développement cohérent, la bonne utilisation et la promotion de l'approximation à garanties de performances. Aux seconds nous voulons montrer un aspect souvent ignoré de ce champs de recherche qui, comme beaucoup d'autres, se montre souvent sous ses traits les plus techniques réservant aux spécialistes l'élaboration des concepts et les travaux sur ses fondements. Pourtant, ces aspects nous semblent au moins aussi importants que les techniques de démonstration pour aborder ce domaine et comprendre ses enjeux. C'est pourquoi nous invitons l'œil extérieur à partager ces premiers éléments de réflexion. Les concepts décrits dans cet article sont accessibles sans connaissance

préalable particulière et sont empreints d'une certaine vision de l'algorithmique à garanties de performances. Notre souci a en effet été de mieux répondre aux besoins et aux enjeux de l'approximation polynomiale ; ils en deviennent ainsi une bonne illustration. Enfin, la démarche vers ce formalisme, ses justifications et les besoins auxquels elle essaie de répondre, ont leur intérêt, même placés hors du contexte de l'approximation. En effet, comme nous le montrons dans la suite, l'approximation à garanties de performances est à un point de son développement où un besoin conceptuel se fait sentir. Des situations analogues peuvent se présenter dans moult domaines scientifiques et il serait intéressant de comparer l'expérience que nous relatons ici à des travaux du même genre dans d'autres thématiques.

Notre objet n'est pas de développer les détails de ce formalisme mais plutôt d'en retirer la logique générale, sa pertinence, ses conséquences pratiques et les raisons qui nous ont conduit à le concevoir. Dans cette optique, les résultats d'approximation cités dans ce document ont valeur d'exemples. Certains ont été spécialement conçus pour illustrer notre propos ; d'autres, par contre, existent dans la littérature mais sont repris ici (parfois même juste mentionnés dans le texte) pour motiver notre démarche. Pour ces derniers, seules les preuves les plus accessibles illustrant un point particulier sont reprises. Par souci de lisibilité et de continuité du document, nous avons sélectionné deux problèmes – *le stable maximum* (noté STABLE dans ce qui suit) et le *bin-packing* (BINPACKING) – auxquels se rapportent la majorité de nos exemples. D'autres problèmes, notamment *la clique maximum* (noté CLIQUE dans ce qui suit), *la coloration minimum* (COLORATION) et *la couverture d'ensembles* (noté H-TRANSVERSAL) sont mentionnés ponctuellement pour des exemples significatifs n'ayant pas d'équivalent pour nos deux problèmes de référence. Tous les problèmes cités dans le document sont définis en annexe dans la section A. Pour plus de renseignements sur l'état de l'art de l'approximation polynomiale le(la) lecteur(ice) intéressé(e) est invité(e) à consulter les ouvrages [2, 28, 41].

Nous rappelons, dans la section 2, les notions de base sur lesquelles repose la suite de ce travail. Nous situons l'approximation polynomiale par rapport à la théorie de la complexité et par rapport aux différentes méthodes de résolution de problèmes difficiles (Sect. 2.1) et détaillons ses enjeux (Sect. 2.3). Nous présentons notamment (Sect. 2.2) les notions de problème d'optimisation, d'algorithme approché et de rapport d'approximation. Dans le paragraphe 2.4, nous proposons alors une synthèse des principales motivations qui ont conduit à compléter ce formalisme usuel.

Dans la section 3, nous discutons de premières extensions des notions usuelles permettant d'exprimer les possibilités de résolution efficace d'un problème. Un petit historique concernant l'approximation du problème STABLE (Paragr. 3.1) motive ces extensions. Nous présentons alors les notions de *chaînes d'approximation* (Paragr. 3.2) et de *niveau d'approximation* (Paragr. 3.3). Ces concepts ouvrent la voie à différentes notions de limites présentées dans la section 4. En particulier, nous discutons la définition de *rapport asymptotique* (Sect. 4.2) et développons, à cet effet, plusieurs exemples. Le concept d'*ordre de difficulté* (Paragr. 4.2.3) est

central pour cette question. Enfin, la section 5 regroupe des exemples de synthèse qui illustrent ces notions et les motivent.

2. NOTIONS DE BASE – LE FORMALISME USUEL

2.1. LE CADRE

La classe **NP** désigne l'ensemble des problèmes solubles par une machine de TURING polynomiale non déterministe ; il s'agit, au moins dans la théorie classique de la complexité, de l'univers de travail. La classe **P** (des *problèmes polynomiaux*) est l'ensemble des problèmes de **NP** résolus par une machine de TURING polynomiale déterministe. Dans ce type d'étude, la complexité algorithmique polynomiale correspond à ce qui est communément considéré comme une *bonne résolution*. À l'opposé, la classe **NP** contient les problèmes **NP-complets** (classe **NP-C**) qui sont, dans un certain sens, les problèmes les plus difficiles de **NP** : la résolution de l'un d'entre eux par une machine de TURING polynomiale déterministe permettrait la résolution de n'importe quel problème de **NP** par une machine de ce type. Parmi les problèmes connus et entrant dans le cadre de nombreux modèles industriels ou en sciences humaines, relativement peu sont polynomiaux : les principaux problèmes de **P** sont *les problèmes de tris, de parcours d'arbres et de graphes, certains problèmes de flots, d'affectation, des problèmes de cheminements optimaux (avec notamment, comme application, les problèmes d'ordonnement les plus simples et, plus généralement, l'optimisation séquentielle), le problème de couplage maximum d'un graphe et la programmation linéaire en variables continues*. Ces problèmes suscitent un intérêt particulier, non seulement parce qu'ils ont de nombreuses applications, mais aussi parce qu'ils constituent les outils de base de l'algorithmique polynomiale. Si les problèmes polynomiaux sont relativement rares, de très nombreux modèles réels conduisent par contre à des problèmes **NP-complets**. Parmi eux, citons bien sûr le fameux problème de *satisfiabilité* qui fut le premier problème **NP-complet** à être mis en évidence par Cook [7]. Mentionnons aussi *le problème du voyageur de commerce* (TSP), STABLE, *l'arbre de STEINER*, COLORATION, les principaux problèmes de partage, de découpe et de couverture minimale, la plupart des problèmes d'ordonnement, *le problème du sac à dos* (KS) et, plus généralement, *la programmation linéaire en variables entières*. La découverte, depuis les travaux de Cook et de Karp [7, 31], d'une multitude de problèmes **NP-complets** issus d'applications réelles a mis en évidence les limites de la résolution informatique de problèmes concrets. Mais surtout, ce constat justifie les efforts actuels en vue de résoudre de tels problèmes de manière efficace, au moins partiellement.

Dans cette optique, plusieurs directions de recherches ont été développées. Parmi celles-ci, mentionnons d'abord l'identification de cas particuliers (sous-problèmes) polynomiaux d'un problème général difficile et la mise au point de méthodes exactes. La première est particulièrement intéressante pour comprendre ce qui fait la difficulté d'un problème et pour mettre en évidence des cas simples qui interviennent parfois de manière naturelle dans les modèles généraux. Les méthodes

exactes, quant à elles, permettent de traiter des instances de taille modérée (ce terme prenant des significations très différentes d'un modèle à un autre); elles sont fondées sur une recherche exhaustive arborescente pour laquelle on réussit à s'affranchir, par des arguments de bornes, d'une partie de l'espace de recherche.

Un autre point de vue pour résoudre efficacement des problèmes intrinsèquement difficiles est d'envisager des méthodes approchées. L'idée sous-jacente est qu'une bonne solution, éventuellement non optimale mais obtenue en temps raisonnable, vaut parfois mieux qu'une solution optimale nécessitant un temps de calcul trop élevé. D'ailleurs, lorsque la taille de l'instance ne permet pas sa résolution complète ou lorsque les sous-problèmes identifiés comme faciles n'ont aucune pertinence du point de vue des besoins, les deux premiers points de vue sont inopérants. L'idée de gagner en temps de calcul (ou de rendre le calcul possible) au prix de la qualité de la solution peut alors séduire, voire même s'imposer. Là encore, cette issue n'est envisageable que si le problème s'y prête. En effet, dans certains cas, une solution non exacte n'est pas exploitable et n'a même parfois pas de sens. C'est en particulier le cas pour un problème de décision dont la réponse ne peut être qu'alternative. Les méthodes approchées ne sont envisagées que pour des problèmes d'optimisation pour lesquels on considère qu'une solution satisfaisant les contraintes peut être exploitée sachant qu'elle le sera d'autant mieux que sa valeur objective est bonne. Les contraintes du problème représentent ce que doit absolument vérifier une solution, l'objectif détermine sa qualité.

Les méthodes approchées diffèrent les unes des autres par le type de compromis qualité/complexité qu'elles offrent. Nous en distinguons deux classes : les heuristiques³, dont la qualité de la solution qu'elles calculent ne peut être estimée qu'*a posteriori* (de façon expérimentale) et les algorithmes polynomiaux à garanties de performances pour lesquels on peut estimer *a priori* la qualité de la solution qu'ils fournissent. Les premières, très utilisées en pratique, n'offrent aucune garantie *a priori*. Il s'agit de stratégies, de types très variés, dont le principe repose sur des arguments intuitifs, voire sur des paradigmes de phénomènes naturels mais qui ne sont justifiées que par les résultats qu'elles offrent. Aussi, le recours à des plans d'expériences est souvent retenu pour tester des heuristiques. Dans certains cas, ces méthodes donnent lieu à des analyses complétant les observations. De telles analyses *a priori* sont systématiques (et même imposées dans les définitions) dans le cadre de l'approximation polynomiale. En effet, ce domaine se définit par des règles strictes que doivent satisfaire les algorithmes qu'il étudie. Le principe est de se limiter à l'emploi de *méthodes polynomiales* en exigeant de plus des *garanties absolues* sur la qualité des solutions obtenues. Jusqu'à présent, ce choix quant aux garanties de complexité et de qualité des solutions obtenues distingue l'approximation polynomiale des autres méthodes approchées; mais il serait tout à fait envisageable et particulièrement intéressant d'étudier le comportement des problèmes difficiles par rapport à un autre compromis complexité/qualité que le couple "*polynomial/garanties absolues*".

³Dans ce qui suit, par abus de langage, nous appellerons heuristiques des algorithmes non-exactes qui soit ne sont pas polynomiaux, soit n'ont pas de garanties de performance *a priori*.

D'un point de vue opérationnel, la pertinence de ces approches dépend des besoins spécifiques de chaque application. Dans la pratique, le recours à des heuristiques reste majoritaire, ce qui peut être attribué non seulement à leur bon comportement en pratique mais aussi à une relative méconnaissance des particularités de l'approximation polynomiale. L'autre raison est le pessimisme de certains résultats d'approximation qui provient essentiellement des restrictions que ce cadre impose. Soulignons néanmoins que les méthodes polynomiales peuvent avoir, dans la pratique, des comportements bien meilleurs que les bornes qu'elles garantissent. D'autant plus qu'en combinant une heuristique et un algorithme polynomial à garanties de performances, on peut bénéficier du bon comportement en pratique de la première et des garanties d'approximation du second.

2.2. GÉNÉRALITÉS SUR LES ALGORITHMES D'APPROXIMATION

2.2.1. La classe *NPO*

L'approximation polynomiale restreint son champs d'étude à certains problèmes pouvant s'exprimer en termes d'optimisation. Les problèmes **NP** sont définis à l'origine [19] comme des problèmes de décision exprimés sous forme d'une question ; la *résolution* du problème consiste alors à apporter une réponse (*oui* ou *non*). Pour de tels problèmes, il est difficile de concevoir une notion de résolution approchée. Par contre, pour certains d'entre eux, il est possible d'associer une *version optimisation* exprimée par un programme mathématique (optimisation d'une fonction réelle sous contraintes). À chaque instance on associe une question portant sur l'existence ou non d'une solution meilleure qu'un seuil fixé. C'est par ce biais qu'on peut faire le lien entre les problèmes d'optimisation et les problèmes de décision de la classe **NP**. Plus précisément, nous rappelons ici la définition formelle des problèmes d'optimisation.

Définition 1. *Un problème élémentaire ou "instance" (notée I) est un programme mathématique de la forme*

$$I : \begin{cases} \text{opt} & v(\vec{x}) \\ & \vec{x} \in \mathcal{C} \\ & \vec{x} \in \{0, 1\}^{|\vec{x}|} \end{cases}$$

où \mathcal{C} désigne ce qui est communément appelé "contraintes" et $|\vec{x}|$ désigne la dimension de \vec{x} . Un "problème d'optimisation combinatoire **NP** (resp., **NP-complet**)" est un ensemble d'instances tel que l'ensemble correspondant des questions "pour un M donné, existe-t-il $\vec{x} \in \mathcal{C}$ tel que $v(\vec{x})\theta M$?" (où θ vaut \geq (resp., \leq) si opt vaut \max (resp., \min)) est exactement l'ensemble des instances d'un problème **NP** (resp., **NP-complet**) au sens usuel de la théorie des langages [19, 33]. L'ensemble des problème d'optimisation **NP** est noté **NPO**.

Un problème **NPO** est au moins aussi difficile que sa version décision puisque pour résoudre cette dernière il suffit de connaître la valeur optimale du problème d'optimisation. Lorsque sa version décision est **NP-complète**, le problème **NPO**,

au moins aussi difficile qu'un problème **NP**-complet, est dit **NP**-difficile. Parmi les problèmes **NPO** qui sont **NP**-difficiles, citons **STABLE**, **COLORATION**, **TSP**, ou encore la programmation linéaire en variables entières. Nous proposons en annexe (Annexe A) une liste de problèmes **NPO** utilisés dans ce document.

Le principal intérêt d'exprimer un problème en termes d'optimisation est d'avoir à sa disposition, non seulement une notion de solution acceptable⁴ (ou réalisable) du problème, mais aussi une notion de qualité d'une solution réalisable par référence à la valeur de l'objectif.

2.2.2. Algorithmes approchés et mesures d'approximation

Pour un tel problème, on appelle *algorithme approché* un algorithme polynomial par rapport à la taille des instances fournissant, pour toute instance I , une solution réalisable de valeur $\lambda(I)$. Un résultat d'approximation permet de caractériser la qualité de la solution approchée à l'aide d'une *mesure d'approximation* définie pour chaque instance. La mesure la plus utilisée est le rapport

$$\gamma(I) = \min \left\{ \frac{\lambda(I)}{\beta(I)}, \frac{\beta(I)}{\lambda(I)} \right\}$$

où $\beta(I)$ désigne la valeur optimale de l'instance I . D'autres mesures d'approximation peuvent être employées, notamment le *rapport différentiel* que nous avons défini, justifié et discuté dans [8, 11, 12]. Pour une instance I , il s'exprime par

$$\delta(I) = \frac{|\lambda(I) - \omega(I)|}{|\beta(I) - \omega(I)|}$$

où $\omega(I)$ désigne la *pire valeur* de l'instance I ; $\delta(I)$ mesure donc la position de la valeur garantie par l'algorithme dans l'intervalle entre la pire valeur et la meilleure.

2.2.2.1. Pourquoi un rapport différentiel? Revenons en quelques lignes sur les raisons qui nous ont amené à définir ce rapport. Il s'agissait de répondre à quelques "paradoxes" connus du cadre classique qui sont directement imputables à la mesure utilisée. Tous reviennent à la remarque suivante : du point de vue du rapport classique, des transformations très naturelles et intuitivement anodines de problèmes **NPO** changent radicalement leur comportement du point de vue de l'approximation. Un exemple significatif de telles transformations est l'ajout d'une constante à la fonction objectif; ceci ne change rien du point de vue de l'optimisation et peut pourtant modifier le comportement du problème si on considère le rapport d'approximation classique.

Examinons l'exemple de **BINPACKING** qui consiste à ranger une liste de nombres positifs inférieurs à 1 en un nombre minimum de boîtes (appelées *bin*) de capacité totale 1 (la somme des nombres placés dans chaque boîte ne doit pas excéder 1). Ajouter à une instance un nombre k fixé de "1" ne modifie intuitivement pas

⁴Elle exprime, par le biais des contraintes, ce qu'on désire imposer à n'importe quelle solution envisageable.

l'instance puisque chacun de ces "1" ajoutés devra, dans toute solution, utiliser une boîte à lui tout seul. Ceci revient à ajouter k à la valeur objective de toute solution réalisable de l'instance initiale. Cette transformation, quasiment insignifiante, a pourtant une conséquence importante sur le rapport d'approximation classique : il suffit d'augmenter à souhait k pour rendre toute solution "bonne" (rapport proche de 1) pour le rapport classique. Un autre exemple célèbre est le lien entre le problème STABLE et la couverture de sommets minimum dans un graphe (noté G-TRANSVERSAL dans ce qui suit). Un stable d'un graphe est un ensemble de sommets deux à deux non liés par une arête; une couverture de sommets est un ensemble de sommets qui "touche" toutes les arêtes. Il est très facile de se rendre compte que le complémentaire de toute couverture de sommets est un stable et réciproquement. Ainsi, maximiser le cardinal d'un ensemble stable devrait exactement revenir à minimiser celui du complémentaire d'un stable, *i.e.*, d'une couverture de sommets. Pourtant, du point de vue de l'approximation classique (rapport γ), le premier de ces problèmes est réputé très difficile à approximer⁵ alors que le second admet des algorithmes offrant des garanties communément jugées comme bonnes. Contrairement au cas précédent, les valeurs des solutions ne subissent pas seulement une translation au cours de la transformation, mais une transformation affine tout aussi naturelle. Dans l'exemple choisi, la transformation affine ($x \mapsto n - x$ où n désigne le nombre de sommets) est décroissante, ce qui explique la correspondance entre un problème de maximisation et un problème de minimisation. Sur la base d'exemples de ce type, une réflexion sur la mesure d'approximation à utiliser a été menée dans [8, 11, 12]. Le rapport différentiel a alors été défini par une approche axiomatique imposant notamment la stabilité du rapport vis-à-vis de transformations affines de l'objectif. En outre, cette démarche nous a conduit à réfléchir à la notion de pire valeur qui s'est avérée particulièrement riche. Aujourd'hui, l'approximation de plusieurs problèmes **NPO** a été étudiée sous les rapports classique et différentiel. Les résultats montrent la complémentarité de ces mesures et l'intérêt qu'il y a à mener de front des études sous chacune d'elles. Précisons que, non seulement les résultats d'approximation, mais aussi les méthodes déployées diffèrent en fonction du rapport. L'un des points forts de cette étude a été d'engager une réflexion sur le concept de rapport d'approximation et de montrer ce qu'un tel débat pouvait apporter. Les réflexions sur d'autres concepts que nous relatons ici se situent dans la continuité logique de cette étude.

2.2.3. Résultats d'approximation

Étant donnée une mesure d'approximation⁶ $\rho(I)$, un résultat d'approximation s'énonce comme une borne valable pour toute instance I . Comme pour la plupart des travaux de théorie de la complexité, on mesure donc le comportement d'un

⁵Nous nous autorisons ce néologisme pour désigner la démarche de l'approximation ; de même, l'adjectif *approximable* désigne la faculté d'un problème à être approximé.

⁶Ici, nous prendrons la convention d'une mesure dans l'intervalle $[0, 1]$ avec la valeur 1 pour la résolution exacte.

algorithme par rapport au pire cas, ce qui permet d'obtenir des garanties absolues valables pour toutes les instances.

Les résultats d'approximation mettent en évidence des différences de difficulté entre les problèmes **NP**-difficiles alors que leurs versions décision sont équivalentes par rapport aux réductions polynomiales. Certains problèmes **NP**-difficiles sont très bien approchés par un algorithme polynomial ; le problème de sac à dos (KS) est l'exemple le plus favorable puisqu'il admet un *schéma complet d'approximation polynomial* [29], c'est-à-dire, une famille (indicée par $\epsilon \in]0, 1[$) d'algorithmes polynomiaux (à la fois en n et en $1/\epsilon$) A_ϵ telle que, pour tout ϵ , l'algorithme A_ϵ garantit un rapport γ supérieur à $1 - \epsilon$. Pour d'autres problèmes au contraire, on sait que de tels résultats ne sont pas possibles sous une hypothèse de complexité. Par exemple, pour STABLE, si $\mathbf{P} \neq \mathbf{NP}$, aucun algorithme polynomial ne peut garantir un rapport γ borné inférieurement par une constante $c > 0$ [1]. Les *résultats positifs* correspondent à la conception d'un algorithme polynomial dont l'analyse met en évidence un certain niveau d'approximation. Les *résultats négatifs* correspondent à un seuil de qualité qu'on ne peut atteindre, pour un problème, sous une hypothèse fortement probable du type $\mathbf{P} \neq \mathbf{NP}$. L'association des deux permet d'établir une hiérarchie des problèmes d'optimisation **NP**-complets par rapport à l'approximation. De ce point de vue, l'approximation polynomiale apparaît comme un outil fondamental pour la compréhension de la classe **NPO** et pour l'étude de sa structure.

L'un des aspects de cette théorie est donc l'étude de la difficulté intrinsèque des problèmes en offrant, par rapport à la théorie de la complexité, plusieurs niveaux de difficulté. La définition de la **NP**-complétude offre deux niveaux de difficulté, les classes **P** et **NP-C**. Par contre, en approximation, l'ensemble des problèmes d'optimisation **NP**-complets se décompose en couches d'équi-approximabilité.

2.3. ENJEUX DE L'APPROXIMATION À GARANTIES DE PERFORMANCES

Les types de résultats obtenus et leur grande diversité mettent en évidence deux enjeux complémentaires de l'approximation à garanties de performances qui distinguent ce domaine de celui des heuristiques. Le plus évident est la résolution efficace de problèmes difficiles, l'efficacité correspondant ici au caractère polynomial des algorithmes et aux garanties de qualité de la solution qu'ils offrent. Comme nous l'avons déjà signalé, cet *intérêt opérationnel* de l'approximation complète celui des heuristiques pour la résolution de problèmes réels. Mais l'approximation polynomiale se distingue par *la possibilité d'étudier une structure de la classe des problèmes difficiles fondée sur les propriétés d'approximation*. Ce second enjeu s'avère particulièrement important. Il permet de mieux connaître la classe des problèmes difficiles avec pour principal objectif une meilleure compréhension du phénomène de difficulté, donc sa meilleure gestion.

2.4. MOTIVATIONS POUR ENRICHIR LE FORMALISME DE L'APPROXIMATION POLYNOMIALE

Notre travail visant à enrichir le formalisme classique de l'approximation polynomiale se situe dans la continuité de la réflexion que nous avons engagée – et qui reste d'actualité – sur le rapport d'approximation.

Avant d'en aborder d'autres aspects, nous exposons dans cette section nos principales motivations à engager ce travail. Dans la continuité directe de la théorie de la complexité [19], les développements de l'approximation à garanties de performances ont surtout consisté, dans un premier temps, à établir des résultats sur la base de quelques définitions fondatrices sans que se fasse sentir le besoin de les modifier ou de les étendre. Aujourd'hui un tel besoin apparaît aux yeux de nombreux chercheurs si l'on en juge par le nombre d'articles récents présentant des variantes de définitions en vue d'intégrer de nouveaux résultats. Le besoin d'unifier des définitions aux variantes multiples et d'augmenter l'expressivité du formalisme pour intégrer de manière satisfaisante des résultats de plus en plus fins ou pour envisager des problématiques nouvelles, la volonté de simplifier l'expression des résultats pour mieux en rendre compte et enfin l'intérêt pour engager une réflexion sur la pertinence des notions et leur adéquation avec les objectifs de l'approximation sont les maîtres mots de nos motivations. Nous en donnons ici quelques exemples.

L'explosion du nombre de définitions et le degré de technicité auquel elles aboutissent motive la volonté de les unifier. L'unification des notions permet, non seulement de les simplifier, mais aussi de mieux les comparer. Ce travail permet aussi de répondre aux besoins actuels du domaine et d'envisager des problématiques qui ne trouvent pas leur place dans le formalisme usuel. En effet, beaucoup de concepts d'approximation polynomiale ont été conçus, à l'origine, dans un cadre relativement restrictif correspondant aux besoins et aux résultats de l'époque. Un exemple significatif est la définition originelle [19] du rapport d'approximation ; avec les notations que nous avons adoptées, il était défini comme la borne inférieure de $\gamma(I)$ pour I décrivant l'ensemble des instances. Cette définition est adaptée au cas d'un rapport d'approximation indépendant de l'instance mais devient totalement inopérante pour des rapports d'approximation non minorés pour lesquels l'infimum est nul. Or, le besoin de prendre en compte des rapports dépendant de l'instance s'est fait très vite sentir puisque pour certains problèmes tels que STABLE, TSP, COLORATION, etc., un rapport d'approximation constant ne peut être garanti⁷ par un algorithme polynomial. Cependant, la simple adaptation des définitions (ce qui a souvent été utilisé) n'est, selon nous, pas entièrement satisfaisante. En effet, le cas des rapports dépendant de l'instance soulève des problèmes spécifiques tels que le type de paramètres à prendre en compte dans le rapport d'approximation et la manière dont évolue la difficulté des instances en fonction de ces paramètres. L'exemple de STABLE, développé au paragraphe 3.1, illustre tout à fait cette question. Deux familles de résultats d'approximation ont été établies

⁷Sous une hypothèse de complexité du type $\mathbf{P} \neq \mathbf{NP}$.

en parallèle : ceux exprimés en fonction du nombre de sommets du graphe et ceux exprimés en fonction de son degré maximum. Les formes de rapports d'approximation mis en jeu pour chaque paramètre sont très différentes. Le formalisme que nous proposons intègre le point de vue des paramètres à l'ensemble des concepts ; il permet ainsi d'envisager ces différentes questions de manière systématique, ne serait-ce déjà que par la possibilité de les poser. Par ailleurs, des travaux mettant en œuvre des analyses de plus en plus fines ne peuvent être directement exposés dans le cadre usuel. Il en résulte une difficulté d'exprimer les résultats, voire même de les identifier et les répertorier de manière précise. Là encore, comme nous le verrons, STABLE fournit un exemple où le manque de formalisme a bloqué les possibilités de rendre compte de résultats dans leur globalité et de les améliorer. Les travaux sur le rapport différentiel ont été directement motivés par la question de la pertinence de la mesure usuelle. Un autre exemple que nous mentionnons dans cet article est celui de la définition d'un *rapport asymptotique*. Ce terme recouvre des analyses qui ne sont valables que lorsque "l'instance tend vers l'infini", en un sens à préciser. Nous montrerons dans la section 4.2 pourquoi la définition usuelle ne nous semble pas toujours satisfaisante et en proposerons une autre.

3. APPROXIMABILITÉ – VERS UNE CLASSIFICATION DE NPO

La notion de résolution est implicite dans la définition même d'un problème. On ne peut en effet définir ce dernier ou, plus généralement, une question que par rapport au concept de réponse qu'on adopte. D'un point de vue historique d'ailleurs, le moteur (mais aussi le frein) des travaux sur la décidabilité, la complexité ou plus généralement de toute démarche visant à étudier les limites de l'algorithmique a été la modélisation de la notion d'algorithme⁸ représentant une forme de résolution sur laquelle repose alors la définition des classes de complexité. De même, dans le cadre de l'approximation polynomiale, la notion de problème d'optimisation ne s'impose qu'au terme d'une réflexion sur les buts de cette théorie, notamment à propos de la signification à donner à la *résolution approchée*. Dans l'article [12], nous avons insisté sur le choix du "tout dans la formulation" qui consiste à intégrer dans le problème même les caractéristiques de sa résolution. Nous avons alors montré à quel point, non seulement la fonction objective et les contraintes, mais aussi les valeurs extrémales (optimale et pire) de l'instance sont indispensables pour clarifier le concept de *bonne résolution* en vue d'une évaluation pertinente des algorithmes.

Le choix d'un rapport d'approximation conditionne la manière dont la qualité de résolution d'une instance est prise en compte. Pour étudier la qualité d'un algorithme pour un problème donné, il faut définir un moyen d'agrèger les informations pour l'ensemble des instances. Pour cela, la traditionnelle notion d'algorithme approché garantissant une proportion constante de la valeur optimale ne suffit plus comme l'illustrent les exemples détaillés dans ce document. L'intérêt des résultats obtenus avec chaque rapport et les possibilités d'analyses qu'ils offrent

⁸Notamment par l'outil des machines de TURING.

nous convainquent de concevoir un formalisme unique permettant de travailler avec l'un ou l'autre des rapports d'approximation. De même, nous avons déjà évoqué le besoin de prendre en compte de manière systématique des résultats d'approximation fonction de l'instance. Cette idée n'est pas neuve en elle-même mais n'a été que très partiellement intégrée au formalisme existant avec les conséquences que nous avons déjà signalées, en particulier sur la mise en retrait de la question des paramètres de l'instance et de leur rôle dans les résultats d'approximation. Une troisième nouveauté concerne des généralisations des algorithmes d'approximation : dans la lignée des schémas d'approximation qui jusque là se singularisaient dans l'outillage de l'approximation polynomiale, nous introduisons l'idée d'étudier le comportement de suites d'algorithmes (appelées *chaînes d'approximation*) qui permettent alors de rendre compte de résultats de convergence. Un bref historique de l'approximation de STABLE (Paragr. 3.1) permet d'illustrer les besoins de recourir à ce type d'outils. Dans la suite de cette section, nous discutons les notions de chaîne d'approximation et de niveau d'approximation. Dans la partie 4, nous creusons alors l'idée de résultats limites en approximation. Nous discutons en particulier les notions asymptotiques présentées comme une forme de "limite par rapport aux instances".

3.1. UN EXEMPLE : UNE BRÈVE HISTOIRE DE L'APPROXIMATION DU PROBLÈME DE STABLE MAXIMUM

Un bref survol des résultats d'approximation pour STABLE met en évidence l'intérêt de recourir à de nouvelles notions et de poser de nouvelles problématiques. Certains seront repris ou précisés ultérieurement afin d'illustrer un concept ou une problématique. Rappelons en effet que le but de ce travail est de mener une discussion sur les concepts, les résultats ne servant que d'exemples.

Une première famille de résultats concerne des rapports d'approximation fonction du nombre de sommets n de l'instance. Il est évident que, pour le problème non-pondéré, le rapport d'approximation k/n peut être garanti en temps polynomial pour toute constante k : il suffit de chercher de manière exhaustive un meilleur stable parmi les ensembles d'au plus k sommets. Le principal résultat fonction de n pour le cas non-pondéré met en jeu [5] un rapport d'approximation de la forme $O(\log^2 n/n)$. Pour le cas pondéré, nous montrons dans [15] comment déduire de ce résultat un algorithme garantissant le rapport $O(\log^2 n/(n \log \log n))$.

Parallèlement à ces résultats, de nombreux travaux ont porté sur des analyses fonction du degré maximum Δ ou du degré moyen μ ($\mu = 2|E|/n$, où E désigne l'ensemble d'arêtes du graphe). La comparaison de ces deux types de résultats motive une réflexion de fond sur les paramètres en fonction desquels exprimer les résultats d'approximation.

Les premiers résultats d'approximation pour STABLE fonction du degré proposent le rapport $1/\Delta$ [19] garanti par un algorithme glouton. Nous reprenons son analyse dans le paragraphe 4.2. Cependant, un résultat de Turán [40] bien antérieur fournit le rapport $1/(\mu + 1)$ pour ce même algorithme, ouvrant ainsi la voie des rapports fonction du degré moyen. Bien que ces résultats fassent désormais

partie du lieu commun, ils ont longtemps été les seuls connus et restent relativement compétitifs. La première amélioration significative a été le rapport $2/\Delta$ [27] valide également pour le cas pondéré *WSTABLE*. À partir de 1994, l'approximation de *STABLE* en termes de degré a connu une évolution rapide. Une analyse fine de l'algorithme glouton a permis d'obtenir [23], pour le cas non-pondéré, les rapports $3/(\Delta + 2)$ et $2/(\mu + 2)$ (voir également la Prop. 4.4 du Paragr. 4.2) ; ce dernier rapport peut facilement être amélioré en $5/(2\mu + 3)$ en utilisant, en *preprocessing*, la très élégante méthode due à Nemhauser et Trotter [36] qui permet de déterminer une solution semi-intégrale du programme linéaire associé à *STABLE*. Pour atteindre une approximation similaire pour le cas pondéré, il a fallu développer une toute autre démarche ; le rapport $3/(\Delta + 2)$ est établi dans [21] par une technique de partitionnement du graphe. Dès lors, la version pondérée s'est fortement démarquée : pour *WSTABLE*, le rapport $3/(\Delta + 2)$ est longtemps resté le meilleur au moment où le cas non-pondéré bénéficiait de plusieurs améliorations. Mentionnons en particulier le rapport $5/(\Delta + 3) - \epsilon$ [4] et même le rapport $O(\log \log \Delta/\Delta)$ [24]⁹. Toutefois, les algorithmes correspondants ne sont polynomiaux que pour des instances de degré borné et deviennent exponentiels pour des graphes quelconques. Pour le cas général, l'article [24] a été une étape cruciale : son principal résultat est un algorithme de complexité $O(n\Delta^{k-1})$ garantissant, pour une constante k , un rapport de la forme $(6/\Delta) - \epsilon(k) - \eta(\Delta)$ avec $\epsilon(k) \rightarrow 0$ quand $k \rightarrow \infty$ et $\eta(\Delta) \rightarrow 0$ quand $\Delta \rightarrow \infty$. Cependant, le résultat est exprimé de manière très confuse et même imprécise alors que nous verrons qu'il a, dans le formalisme que nous présentons, une forme très simple (Prop. 5.3). Il s'agit d'un exemple significatif où l'inadéquation du cadre classique pour exprimer ces résultats a contraint les auteurs à n'en présenter qu'une forme partielle.

Ces différentes approximations correspondent à des améliorations d'un facteur multiplicatif constant qui sont obtenues par des analyses de plus en plus fines. Elles posent implicitement la question de la "meilleure constante" devant le terme $1/\Delta$ ou $1/\mu$. Nous y avons répondu dans [14] en proposant, pour toute constante k , un algorithme de complexité $O(n^{k/2})$ garantissant un rapport de la forme $k/\Delta - o(1/\Delta)$ ainsi qu'un algorithme de complexité $O(n^k)$ garantissant le rapport $\min\{k/\mu, k' \log \log \Delta/\Delta\} - o(1/\mu)$ où k' est une valeur dépendant de k . Là encore, nous verrons comment exprimer ces résultats plus simplement (Prop. 5.3). La question naturelle émanant de ces travaux concerne la possibilité de dépasser l'ordre de grandeur $O(1/\Delta)$. Elle reste à l'étude même si une réponse positive a déjà été établie (Th. 2). En faisant abstraction des détails dans l'expression de ces derniers résultats, leur forme générale se distingue par le rôle que joue k . Il apparaît comme un paramètre de l'algorithme en fonction duquel on peut moduler la qualité de l'approximation avec une incidence directe sur la complexité. Il ne s'agit plus de l'analyse d'un algorithme mais plutôt d'une suite d'algorithmes à l'image d'un schéma d'approximation polynomial. L'idée des chaînes d'approximation est de formaliser ce type de situation, même pour des cas où les rapports

⁹Ce résultat est l'une des premières référence à un rapport du type $\Omega(1/\Delta)$.

d'approximation correspondant ne sont pas du type $1 - \epsilon(k)$ avec $\epsilon \rightarrow 0$ quand $k \rightarrow +\infty$.

3.2. ALGORITHMES ET CHAÎNES D'APPROXIMATION

Étant donné un problème d'optimisation **NP**-complet, nous avons déjà précisé dans la première section la notion d'*algorithme approché* qui, pour toute instance I de taille n , construit en temps polynomial par rapport à n une solution réalisable. La valeur objective $\lambda(I)$ de cette solution ainsi que les valeurs optimale ($\beta(I)$) et (éventuellement) pire ($\omega(I)$) de I sont essentielles pour l'évaluation (par le biais d'un rapport d'approximation) de la qualité de l'algorithme. Parmi les rapports utilisés, nous avons déjà mentionné $\gamma(I)$ et $\delta(I)$, mais l'étude qui suit est applicable à n'importe quelle mesure d'évaluation. Notre seule restriction est de considérer une mesure $\rho : I \mapsto \rho(I)$ à valeurs entre 0 et 1 d'autant plus proche de 1 que I est bien résolue. Ainsi, un résultat d'approximation consiste à garantir, pour toute instance du problème, un minorant de $\rho(I)$.

La notion de *chaîne algorithmique* généralise celle d'algorithme approché. Le principe est d'étudier globalement le comportement d'une famille d'algorithmes.

Définition 2. Soit Π un problème d'optimisation combinatoire (Déf. 1); on note \mathcal{I} l'ensemble des instances de Π . Soit $f : \mathcal{I} \times \mathbb{N} \rightarrow]0, 1[; (I, k) \mapsto f(I, k)$ une application croissante en k . Une chaîne (algorithmique) d'approximation polynomiale¹⁰ de rapport f pour Π est une suite d'algorithmes $(A_k)_{k \in \mathbb{N}}$ telle que, pour tout $k \in \mathbb{N}$, A_k est un algorithme approché polynomial pour Π garantissant le rapport d'approximation $\rho_k(I) \geq f(I, k)$, $\forall I \in \mathcal{I}$.

L'intérêt de cette notion est double. D'abord, elle rend compte globalement d'une famille d'algorithmes et à ce titre permettra de simplifier certains énoncés. Si on se limitait à considérer, pour chaque valeur de k , l'algorithme correspondant et ses garanties, cet intérêt resterait mineur puisqu'un énoncé paramétré par k permet d'exprimer le même résultat de manière aussi simple. Par contre, la simplification devient sensible pour des résultats décrivant le comportement global de la suite d'algorithmes et de la garantie associée qui peut être vue comme une suite de fonctions $k \mapsto f(I, k)$. Cette présentation permet en particulier de définir une notion de convergence par rapport à k rendant compte des garanties associés aux grandes valeurs de k . C'est exactement le type de gain que l'on a en utilisant la notion de schéma d'approximation polynomial au lieu d'un algorithme paramétré dont le rapport d'approximation se rapproche de 1 pour certaines valeurs limites du paramètre. Ce type de simplification permet, non seulement de mieux rendre compte des résultats, mais surtout de mieux comparer différentes approximations. Le second intérêt, illustré notamment par l'exemple de STABLE, est de laisser à l'utilisateur le choix du niveau de qualité et surtout d'explicitier, pour un type de résolution fixé, le coût en complexité d'une amélioration de la qualité.

¹⁰Ou simplement *chaîne* lorsqu'aucune confusion n'est possible.

Dans les prochains paragraphes, nous précisons les notions de niveau d'approximation, à la base de la conception de classes d'approximation, et celle de convergence.

3.3. NIVEAUX D'APPROXIMATION

Nous insistons, une fois de plus, sur le fait qu'un des principaux enjeux de l'approximation polynomiale est de mettre en évidence, au sein de la classe **NPO**, une structure fondée sur les propriétés d'approximation des problèmes. La grande variété de résultats d'approximation suggère de manière très naturelle de hiérarchiser les problèmes en fonction du type de rapport d'approximation qu'on peut garantir. Les classes les plus courantes définies jusqu'à présent sont **APX** et **PTAS**. La première regroupe les problèmes admettant un algorithme à rapport constant et la seconde ceux admettant un schéma d'approximation polynomial; elles vérifient l'inclusion **PTAS** \subset **APX**. On rencontre aussi de manière un peu plus anecdotique les classes **Log-APX** et **Poly-APX** qui recouvrent les problèmes admettant un algorithme à rapport de la forme $g(1/n)$ où n désigne la taille de l'instance et g est respectivement une fonction logarithmique et une fonction polynomiale. Ces deux dernières classes ont été introduites pour pallier au manque d'outils adaptés pour hiérarchiser la région **NPO** \setminus **APX** des problèmes n'admettant pas d'algorithme polynomial à rapport constant. Un premier essai de formalisation est donné dans [32] : la classe **F-APX** comprend les problèmes approximables à un rapport de la forme $g(1/n)$ où n désigne la taille de l'instance et $g \in F$, où F représente une classe de fonctions vérifiant,

$$g \in F \implies \forall(c, h) \in \mathbb{N} \times \mathbb{R}, x \mapsto hg(x^c) \in F. \tag{1}$$

De telles restrictions sur la famille de fonctions définissant le niveau sont vérifiées par les classes des fonctions logarithmiques et polynomiales; cependant ces hypothèses excluent certaines classes intéressantes de rapports d'approximation. Mentionnons par exemple le cas de fonctions de la forme $g(x) \in O(x^{1-\epsilon})$, $\epsilon \in]0, 1[$. Pour plusieurs problèmes, notamment pour **STABLE**, le rapport $1/n$ est trivial à garantir alors que sous une hypothèse de complexité, aucun algorithme polynomial ne peut garantir un rapport de la forme $O(1/n^{1-\epsilon})$, $\epsilon \in]0, 1[$ [26]. Ce type de résultat justifie qu'on ait besoin, entre autres, de différencier des sous-classes de **Poly-APX** en fonction du degré des polynômes, ce qui n'est pas compatible avec l'expression (1). Nous suggérons d'assouplir les règles permettant à une classe de fonctions de définir un niveau d'approximation. La seule règle que nous choisissons d'imposer est la stabilité par passage au minimum (si f et g sont dans F , $\min\{f, g\}$ est dans F). Ceci donne la possibilité, pour garantir un certain niveau d'approximation, de partitionner l'ensemble des instances et garantir ce niveau sur chaque partie; il s'agit d'une technique à la base de nombreuses démonstrations.

Un autre inconvénient de la définition de **F-APX** est qu'elle exclut la classe **PTAS** qui est pourtant l'une des plus utilisées en approximation polynomiale.

Notons enfin que les classes existantes ne sont conçues que pour le rapport d'approximation classique γ ; pour concevoir un cadre unifié, nous devons préserver la possibilité de moduler les résultats en fonction du rapport.

Pour répondre à ces besoins, nous envisageons la notion de *niveau d'approximation* comme un ensemble de résultats d'approximation. On lui associe alors l'ensemble des problèmes pour lesquels un de ces résultats au moins peut être garanti. Étant donné un problème Π d'ensemble d'instances \mathcal{I} , on note \mathcal{F}_Π (ou seulement \mathcal{F} lorsqu'aucune confusion n'est possible) l'ensemble des rapports possibles pour les chaînes d'approximation polynomiale pour Π , *i.e.*,

$$\mathcal{F}_\Pi = \{f : \mathcal{I} \times \mathbb{N} \rightarrow]0, 1[: (I, k) \mapsto f(I, k)\}.$$

Nous utilisons également la notation \mathcal{F}'_Π pour désigner les rapports indépendants de k , *i.e.*,

$$\mathcal{F}'_\Pi = \{f : \mathcal{I} \rightarrow]0, 1[: I \mapsto f(I)\}.$$

Étant donné un rapport d'approximation, un *niveau d'approximation* est associé à une famille de rapports $F \subset \mathcal{F}_\Pi$ vérifiant $\forall (f, g) \in F \times F, \min\{f, g\} \in F$. Toutes les chaînes d'approximation polynomiale garantissant un rapport dans F appartiennent à ce niveau.

4. NOTIONS DE LIMITES

Nous évoquons, dans ce paragraphe, deux notions de limites de chaînes d'approximation (l'une par rapport au paramètre k et l'autre par rapport aux instances) qui permettent d'affiner la définition de niveau d'approximation.

Certaines analyses ont recours à des notions de limite pour exprimer de manière plus simple un résultat assez technique. L'esprit est toujours le même : il s'agit d'extraire du résultat complet une partie significative plus maniable. Dans le cadre que nous venons d'ébaucher, deux types de limites apparaissent naturellement : l'une se rapportant à une suite d'algorithmes (nous parlons de convergence), et l'autre attachée à l'idée d' "instances limites" (notion asymptotique). La première est directement liée à la notion de chaîne d'approximation ; la seconde a été maintes fois utilisée mais sous des définitions qui méritent, selon nous, commentaires et débats.

4.1. CONVERGENCE D'UNE CHAÎNE D'APPROXIMATION POLYNOMIALE

Dans la littérature, la convergence par rapport à une suite d'algorithmes n'est formalisée que dans le cas de schémas d'approximation mais elle est implicite dans plusieurs résultats récents tels que ceux que nous avons évoqués pour STABLE.

Définition 3. *Une chaîne d'approximation polynomiale $(\mathbf{A}_k)_{k \in \mathbb{N}}$ converge vers un rapport $\tilde{\rho} \in \mathcal{F}'_\Pi$ si : $\forall \epsilon > 0, \exists \kappa$ tel que $\forall k \geq \kappa$ et $\forall I \in \mathcal{I}, \rho(I, k) = \rho_{\mathbf{A}_k}(I) \geq \tilde{\rho}(I)(1 - \epsilon)$.*

En d'autres termes $(A_k)_{k \in \mathbb{N}}$ garantit une suite de rapports qui converge uniformément vers $\tilde{\rho}$. Le cas du schéma d'approximation correspond à une chaîne convergeant vers le rapport constant 1 associé à la résolution exacte.

4.2. NOTIONS ASYMPTOTIQUES

Le terme *asymptotique* représente l'idée d'*instances limites* qui mérite d'être discutée et approfondie. De nombreux résultats utilisent ce qualificatif sans référence à une définition précise; d'ailleurs, la seule définition explicite de rapport asymptotique [19] exclut la plupart de ces résultats. Cependant, le principe consiste toujours à restreindre l'analyse à un groupe d'instances pour lesquelles un paramètre de référence (la valeur optimale, la taille, le degré (pour un graphe), ...) dépasse un certain seuil. La philosophie sous-jacente est d'associer la valeur de ce paramètre à une mesure de difficulté de l'instance et ainsi, d'une certaine façon, de ne tenir compte que des instances les plus difficiles. Le paramètre joue donc un rôle primordial; or, il n'est souvent dicté que par des besoins techniques sans justification particulière. Nous proposons, dans cette partie, une analyse critique de la définition usuelle et des résultats qualifiés d'asymptotiques dans la littérature. Cette discussion nous conduit à une nouvelle définition; d'un coté elle limite le champ de la définition usuelle pour des questions de pertinence et de l'autre offre de nombreuses possibilités nouvelles, permet d'unifier différents résultats de la littérature et les intègre à un cadre précis.

4.2.1. Définition usuelle

Jusqu'ici, la seule notion asymptotique explicitement définie pour un algorithme A est la suivante [19] : $\gamma_A^\infty = \sup_{\kappa \geq 0} \inf\{\gamma_A(I), I \text{ telle que } \beta(I) \geq \kappa\}$. Nous ferons référence à cette définition en parlant de *notion asymptotique par rapport à la valeur optimale*. Elle a été adaptée au cas d'un schéma d'approximation polynomial [35] : un schéma d'approximation polynomial asymptotique est une famille A_ϵ d'algorithmes tels que, pour chaque $\epsilon \in]0, 1[$, A_ϵ garantit le rapport asymptotique $1 - \epsilon$, i.e., $\forall \epsilon \in]0, 1[, \gamma_{A_\epsilon}^\infty \geq 1 - \epsilon$.

De manière informelle, le rapport asymptotique (meilleur que le rapport constant garanti) décrit le comportement de l'algorithme sur les instances I dont la valeur optimale $\beta(I)$ est élevée. Il revient donc, pour toute constante K , à retirer de l'analyse le comportement des instances de valeur optimale inférieure à K . Si on cherche à justifier intuitivement cette définition, ne s'intéresser qu'aux instances de grande valeur optimale revient à considérer que ce sont les instances les plus "intéressantes".

4.2.2. Commentaires

La définition de [19] n'est adaptée qu'au cas des rapports constants; toutefois la généralisation au cas général n'est pas difficile. On dit que le rapport asymptotique ρ est garanti par A (par rapport à la valeur optimale), si $\forall \epsilon, \exists \kappa$ tel que $\forall I$ tel

que $\beta(I) \geq \kappa$, $\gamma_A(I) \geq \rho(I)(1 - \epsilon)$. Ce type de définition s'adapte immédiatement au cas d'autres rapports d'approximation.

Hormis ces ajustements, cette définition ne nous semble pas satisfaisante pour deux raisons. D'une part, elle exclut de nombreuses situations analogues pour lesquelles $\beta(I)$ est remplacé par un autre paramètre tel que la taille, le degré, ... De tels résultats, relativement courants, sont d'ailleurs souvent qualifiés d'asymptotiques dans la littérature; c'est notamment le cas de certains résultats du paragraphe 3.1 qui distinguent les instances pour lesquelles $\Delta \rightarrow \infty$. Mais d'un autre côté, le critère de la valeur optimale utilisé dans la définition usuelle n'a pas de justification particulière et s'avère parfois discutable. Nous illustrons ces deux inconvénients sur la base d'exemples de résultats et d'analyses concernant BINPACKING et STABLE. Ils nous conduiront à poser la question des paramètres en fonction desquels définir un rapport asymptotique. Au-delà de l'exactitude mathématique d'un résultat, il nous semble en effet important de poser la question de sa pertinence. Comme l'illustrent ces exemples, il ne faut pas minimiser l'effet psychologique de la formulation car elle peut impliquer un mauvais usage du résultat ou contribuer à sa mauvaise compréhension.

4.2.2.1. Exemple du *bin-packing*. Rappelons le résultat de Fernandez de la Vega et Lueker [18] : pour tout $k \in \mathbb{N}$, il existe un algorithme polynomial \mathbf{A}_k pour BINPACKING garantissant, pour toute instance I , le rapport $\gamma_{\mathbf{A}_k}(I) \geq 1 - 1/\beta(I) - 1/k$. Bien qu'il ait été amélioré depuis, ce résultat reste l'un des plus connus pour ce problème. Un tel rapport d'approximation entre dans le cadre de la définition usuelle d'un schéma d'approximation asymptotique et c'est d'ailleurs ainsi qu'il est le plus souvent référencé. Pourtant, malgré cette terminologie, cet algorithme ne se rapproche en rien d'un schéma d'approximation polynomial. Pour nous en convaincre, rappelons un résultat mentionné sans preuve dans [19]. Il fournit un seuil de difficulté qui montre qu'un *schéma d'approximation ne peut être garanti* et que, par conséquent, le rapport garanti par [18] n'est pas toujours proche de 1. Par ailleurs, la preuve que nous proposons met en évidence que les instances les plus difficiles à approximer sont celles qui ont une petite valeur optimale.

Proposition 4.1. [19] *Si $\mathbf{P} \neq \mathbf{NP}$, aucun algorithme polynomial ne peut garantir un rapport strictement meilleur que $2/3$ pour BINPACKING.*

Preuve. La preuve consiste en une réduction du problème de PARTITION ([19], p. 61) qui est \mathbf{NP} -complet. Pour une liste de x_1, \dots, x_n de n rationnels, ce problème consiste à décider si on peut la partager en deux listes disjointes $\{x_i, i \in I\}$ et $\{x_j, j \in J\}$ (avec $\{1, \dots, n\} = I \cup J$) telles que les sommes des nombres de chaque liste coïncident : $\sum_{i \in I} x_i = \sum_{j \in J} x_j$.

En divisant chaque nombre par la demi-somme totale $\tilde{x}_i := x_i(1/2) \sum_{j=1}^n x_j$, le problème revient à décider si les nombres \tilde{x}_i , $i \in \{1, \dots, n\}$ peuvent être rangés en deux boîtes de capacité 1. Il s'agit donc de décider si la liste $\tilde{L} = (\tilde{x}_i, i \in \{1, \dots, n\})$, vue comme instance de BINPACKING, a pour valeur optimale 2 ou strictement plus. Mais alors, un algorithme d'approximation pour BINPACKING garantissant un rapport $\rho > 2/3$ permettrait de répondre à cette question : si la

valeur optimale de \tilde{L} est égale à 2, un tel algorithme retournerait une solution de valeur strictement inférieure à 3, donc de valeur 2. Par contre, dans le cas inverse, toute solution réalisable, donc en particulier celle retournée par l'algorithme, serait de valeur au moins égale à 3. \square

Remarquons que pour toute instance I de BINPACKING de valeur optimale $\beta(I) \geq 4$, le résultat de [18] permet d'obtenir un rapport d'approximation strictement meilleur que $2/3$. Les instances bloquant les possibilités d'approximation sont donc celles de valeur optimale inférieure à 3. Énoncer le résultat de [18] comme un schéma d'approximation asymptotique nous paraît alors réducteur puisque cette formulation ne rend pas compte des instances difficiles (de faible valeur optimale).

Pour renforcer encore notre propos, considérons la classe \mathcal{C} d'instances de BINPACKING correspondant aux listes croissantes de rationnels de la forme $(x_1, \dots, x_k, a, b, 1, \dots, 1)$ avec $\sum_{i=1}^k x_i \leq 1$ et $a + b > 1$; \mathcal{C} est bien sûr reconnaissable en temps polynomial. Appelons BINPACKING $_{\mathcal{C}}$ le problème de BINPACKING restreint aux listes de \mathcal{C} . On a alors le résultat suivant.

Proposition 4.2. BINPACKING $_{\mathcal{C}}$ est NP-difficile.

Preuve. Reprenons la démonstration de la NP-difficulté du problème de PARTITION [19] qui résulte d'une réduction polynomiale du problème 3-DM (couplage tri-dimensionnel défini dans la Sect. A). Connaître cette réduction n'est pas nécessaire pour la compréhension de la preuve; son principe, que nous rappelons ici, suffit pour la suite. À partir d'une instance I_1 de 3-DM comprenant k triplets on construit polynomialement $k + 1$ entiers (en ordre croissant) a_1, \dots, a_k, B (avec $B \leq \sum_{i=1}^k a_i$) tels que I_1 est admissible si et seulement si il existe une partie $A' \subset \{1, \dots, k\}$ telle que $B = \sum_{i \in A'} x_i$. Posons alors :

$$\left\{ \begin{array}{l} a = 2 \sum_{i=1}^k a_i - B \\ b = \sum_{i=1}^k a_i + B \\ K = 2 \sum_{i=1}^k a_i \\ \tilde{a}_i = \frac{a_i}{K} \quad \forall i \in \{1, \dots, k\} \\ \tilde{a} = \frac{K}{K} \\ \tilde{b} = \frac{b}{K} \end{array} \right.$$

Il est clair qu'étant donnés $A' \subset \{1, \dots, k\}$ et $\bar{A}' = \{1, \dots, k\} \setminus A'$, le fait que $\sum_{i \in A'} x_i = B$ implique

$$\tilde{a} + \sum_{i \in A'} \tilde{x}_i = \tilde{b} + \sum_{i \in \bar{A}'} \tilde{x}_i = 1.$$

Par ailleurs, $\tilde{a} + \tilde{b} > 1$. La liste $(\tilde{x}_1, \dots, \tilde{x}_k, \min\{\tilde{a}, \tilde{b}\}, \max\{\tilde{a}, \tilde{b}\})$ est donc une instance de $\text{BINPACKING}_{\mathcal{C}}$. Sa valeur optimale est 2 si et seulement si l'instance initiale de 3-DM est admissible et 3 sinon. Résoudre cette instance de $\text{BINPACKING}_{\mathcal{C}}$ permettrait de résoudre l'instance initiale de 3-DM. Ce dernier problème étant NP -complet, $\text{BINPACKING}_{\mathcal{C}}$ est NP -difficile. \square

Par le même argument que dans la proposition 4.1 on montre que sous l'hypothèse $\mathbf{P} \neq \mathbf{NP}$, $\text{BINPACKING}_{\mathcal{C}}$ n'admet pas d'algorithme d'approximation polynomial garantissant un rapport strictement meilleur que $2/3$.

Par contre, étant donnée une instance I de $\text{BINPACKING}_{\mathcal{C}}$, si on note p le nombre de "1", la valeur optimale de I est soit $2 + p$, soit $3 + p$ selon que les nombres x_1, \dots, x_k, a, b peuvent tenir dans 2 ou 3 boîtes. Ajouter des "1" ne rend l'instance ni plus simple, ni plus difficile puisque chacun constituera une boîte à lui seul dans toute solution réalisable. Ici, la valeur de l'optimum n'est donc nullement liée à la difficulté de l'instance. Par contre, tout algorithme mettant les k premiers nombres dans la même boîte garantit le rapport $(p + 2)/(p + 3)$, *i.e.*, le rapport asymptotique 1. On en déduit le résultat suivant.

Proposition 4.3. *Si $\mathbf{P} \neq \mathbf{NP}$, $\text{BINPACKING}_{\mathcal{C}}$ n'admet pas d'algorithme polynomial d'approximation garantissant un rapport strictement meilleur que $2/3$. Par contre, il admet un algorithme polynomial asymptotiquement optimal (par rapport à la valeur optimale).*

Pour les problèmes BINPACKING et $\text{BINPACKING}_{\mathcal{C}}$, l'aspect asymptotique par rapport à la valeur optimale revient donc à éliminer les instances bloquant les possibilités d'approximation pour obtenir des résultats asymptotiques dépassant radicalement les niveaux de qualité qu'on peut obtenir pour les instances générales.

4.2.2.2. Exemple du problème de stable maximum. Pour d'autres problèmes, au contraire, les instances de grande valeur optimale sont les plus difficiles et les résultats asymptotiques qui s'y rapportent permettent d'exprimer des résultats qui représentent assez fidèlement le comportement pour toutes les instances. Considérons en particulier la classe très large des problèmes consistant à sélectionner de manière optimale une partie (satisfaisant une propriété) dans un ensemble \mathcal{E} de taille n : les problèmes STABLE ou CLIQUE ou encore H-TRANSVERSAL et G-TRANSVERSAL sont de ce type. Pour les problèmes de minimisation de cette forme, les instances dont la valeur optimale est majorée par une constante k sont reconnaissables et résolubles en temps polynomial : il suffit, pour les reconnaître, d'énumérer toutes les parties de taille k de \mathcal{E} pour s'assurer qu'une au moins est réalisable. Pour les résoudre, il suffit d'énumérer toutes les parties de \mathcal{E} de taille au plus k . Un raisonnement analogue s'applique à certains problèmes de maximisation, en particuliers ceux pour lesquels toute partie d'une solution réalisable est réalisable : il suffit alors, pour vérifier que la valeur optimale est au plus k , de vérifier qu'aucune partie de taille $k + 1$ n'est réalisable. La résolution de telles instances est, quant à elle, identique au cas de la minimisation.

Le problème STABLE est de ce type. Nous proposons ci-dessous une analyse de l'algorithme glouton pour ce problème qui montre à quoi correspond, dans ce cas,

un résultat asymptotique (par rapport à la valeur optimale). Dans ce qui suit, pour $v \in V$, nous notons $\Gamma(v)$ l'ensemble de sommets adjacents à v (voisins de v) et par $\delta(v)$ la quantité $|\Gamma(v)|$ (le degré du sommet v).

```

BEGIN *GLOUTON*
  S ← ∅ ;
  REPEAT
    v ← argminvi ∈ V {δ(vi)};
    S ← S ∪ {v};
    V ← V \ ({v} ∪ Γ(v));
    G ← G[V];
  UNTIL V = ∅ ;
END. *GLOUTON*
```

De nombreuses analyses existent pour mettre en évidence des garanties de performances pour cet algorithme (voir notamment [23]). Parmi celles-ci, nous en repreneons ici une relativement naïve mais qui illustre bien notre propos. Remarquons d'abord qu'il suffit de faire l'analyse sur une composante connexe : si un rapport ρ s'applique à chaque composante, il est clair qu'il s'applique alors au graphe entier en remarquant que les valeurs approchées et optimales des différentes composantes s'ajoutent pour constituer respectivement la valeur approchée et la valeur optimale du graphe entier. Considérons donc un graphe connexe. De même, en notant S^* un stable optimal, on peut se ramener au cas où $S \cap S^* = \emptyset$ en faisant la remarque suivante :

$$\frac{|S| - |S \cap S^*|}{|S^*| - |S \cap S^*|} \leq \frac{|S|}{|S^*|}.$$

Afin de majorer le nombre d'arêtes adjacentes au stable S remarquons qu'au plus Δ arêtes sont adjacentes au premier sommet mis dans la solution et au plus $\Delta - 1$ aux suivants (d'après la connexité et le critère glouton). Au plus $|S|(\Delta - 1) + 1$ arêtes sont donc adjacentes à S . Par ailleurs, chaque sommet de S^* est lié à au moins un sommet de S en utilisant, d'une part le fait que $S \cap S^* = \emptyset$ et d'autre part le fait que S est maximal pour l'inclusion. On en déduit :

$$|S|(\Delta - 1) + 1 \geq |S^*| \iff \frac{|S|}{|S^*|} \geq \frac{1}{\Delta - 1} \left(1 - \frac{1}{|S^*|}\right). \quad (2)$$

Cette expression correspond au rapport asymptotique (au sens de la valeur optimale) étendu au cas d'un rapport non constant. Il a alors une interprétation complètement différente que pour BINPACKING. En effet, sachant que les graphes pour lesquels $|S^*| \leq k$ (où k est une constante arbitraire) peuvent être résolus en temps polynomial, l'algorithme peut être amélioré de la manière suivante. On commence par rechercher un stable de taille k ; s'il n'en existe pas on recherche exhaustivement un stable optimal, sinon ($|S^*| \geq k + 1$) on détermine une solution gloutonne. L'analyse précédente fournit alors, pour cet algorithme, la garantie :

$$\frac{1}{\Delta - 1} \left(1 - \frac{1}{k + 1}\right). \quad (3)$$

En d'autres termes, la garantie asymptotique fournie par l'expression (2) peut artificiellement être modifiée (modulo une plus forte complexité) en une chaîne à rapport convergeant vers $1/(\Delta - 1)$ (expression (3)). Dire que l'algorithme garantit le rapport asymptotique $1/(\Delta - 1)$ est alors essentiellement une simplification technique permettant d'exprimer la partie la plus significative du rapport garanti.

4.2.2.3. Autres paramètres. Avant même de prendre parti sur la signification à accorder à un résultat asymptotique, il nous semble gênant d'avoir, sous une même terminologie, des situations aussi différentes. Dans le cas de BINPACKING, les résultats asymptotiques consistent à considérer des classes d'instances " plus faciles " pour obtenir des résultats meilleurs que ce qui est possible pour le cas général. Pour le cas de STABLE, ils ne servent qu'à exprimer de manière plus simple un résultat en extrayant " sa partie prépondérante ". Nous suggérons de ne retenir que ce second type de résultats asymptotiques sachant que dans l'autre cas, l'expression complète du rapport d'approximation fournit une information intéressante, notamment sur la manière dont il évolue en fonction de la valeur optimale. En se restreignant aux instances difficiles, un résultat asymptotique exploite alors les simplifications techniques que procurent ces instances sans affecter la teneur du résultat.

Mais pourquoi se restreindre alors aux expressions fondées sur la valeur optimale et ne pas prendre en compte d'autres types de " limites d'instances " ? Dans le cas de BINPACKING, un rapport du type $\rho(1 - 1/n)$ pourrait tout à fait être lu comme un *rapport asymptotique* ρ . On pourrait notamment lui appliquer la remarque faite pour STABLE : les instances comportant un nombre borné d'objets sont résolubles en temps polynomial de sorte qu'un résultat du type $\rho(1 - 1/n)$ pourrait être transformé en une chaîne d'approximation à rapport convergeant vers ρ .

Revenons maintenant à l'analyse de l'algorithme glouton pour STABLE. Par un raisonnement similaire n'exploitant pas le critère glouton on peut établir, pour tout stable \tilde{S} maximal pour l'inclusion la garantie

$$\frac{|\tilde{S}|}{|S^*|} \geq \frac{1}{\Delta} = \frac{1}{\Delta - 1} \left(1 - \frac{1}{\Delta}\right). \quad (4)$$

Le cas des graphes tels que $|S^*| < \Delta$ peut être écarté de telles analyses car un raisonnement très simple permet alors d'établir que l'algorithme glouton garantit le rapport $2/(\Delta - 1)$. Si $|S^*| \geq \Delta$, l'analyse que nous avons faite pour l'algorithme glouton fournit logiquement un meilleur rapport que celle valable pour tout stable maximal. Une différence importante entre ces deux résultats réside dans le fait qu'on ne peut pas résoudre de manière optimale le cas des graphes pour lesquels $\Delta \leq k$. Le problème STABLE pour de tels graphes est **NP**-complet dès que $k \geq 3$ [19]. Toutefois, considérer ce résultat comme asymptotique " par rapport à Δ " ne serait pas absurde sachant que, si STABLE reste difficile pour les graphes de degré borné, il peut quand même être approché avec un rapport constant pour ces graphes alors qu'une telle garantie est exclue pour le problème général. Bien sûr, pour cet exemple, considérer le rapport asymptotique n'a pas d'intérêt puisque

l'expression exacte $(1/\Delta)$ du rapport – qui est toujours préférable – est plus simple que le rapport asymptotique (expression (4)) qui en résulte.

Considérons maintenant l'algorithme de recherche locale 2-OPT pour STABLE. Dans le cadre de la recherche d'un stable maximum, les améliorations élémentaires usuelles sont les *a-améliorations* qui consistent, pour $a \in \mathbb{N}$, à retirer $a' < a$ sommets d'un stable S et ajouter $a' + 1$ sommets de $V \setminus S$ en préservant la structure de stable. Lorsqu'aucune 1-amélioration n'est possible, le stable est maximal. Une 2-amélioration consiste, soit en une 1-amélioration ($a' = 0$), soit en le retrait d'un sommet et l'ajout de deux. Dans ce qui suit, nous supposons qu'un stable initial est calculé par application du GLOUTON.

```
BEGIN *2-OPT*
    S ← GLOUTON(G) ;
    WHILE ∃ 2-amélioration (v1, v2, u) de S DO S ← S ∪ {v1, v2} \ {u} OD
    OUTPUT S ;
END. *2-OPT*
```

La proposition suivante est établie dans [24]. Nous mentionnons néanmoins la preuve qui montre bien la différence avec les analyses précédentes.

Proposition 4.4. [24] *2-OPT* garantit le rapport $2/(\Delta + 1)$. Si $|S^*| \leq n/2$, ce rapport peut être amélioré en $3/(\Delta + 2)$.

Preuve. Appelons S la solution construite et S^* une solution optimale. Comme pour l'analyse de l'algorithme glouton, on peut se ramener au cas où $S^* \cap S = \emptyset$. On décompose alors S^* en $S_1^* \cup S_2^*$ où $S_1^* = \{s \in S^* : |\Gamma(s) \cap S| = 1\}$ (les sommets de S^* n'ayant qu'un voisin dans S). Un sommet de S ne peut avoir deux voisins dans S_1^* , sinon l'algorithme aurait eu une étape de plus. On en déduit $|S| \geq |S_1^*|$. L'ensemble S étant maximal pour l'inclusion, le nombre d'arêtes issues de S est au moins égal à $|S_1^*| + 2|S_2^*| + n - |S^*| - |S|$. On en déduit :

$$\Delta|S| \geq 2|S^*| - |S| + n - |S^*| - |S|. \tag{5}$$

Le rapport $2/(\Delta + 1)$ se déduit immédiatement de l'expression (5) en posant $n - |S^*| - |S| \geq 0$. Le rapport $3/(\Delta + 2)$ se déduit de l'expression (5) en posant $n \geq 2|S^*|$ (nous verrons, dans un prochain article, [13], comment transférer ce résultat à tout graphe). □

Dans ce cas, parler de “ rapport asymptotique $2/\Delta$ et $3/\Delta$ ” a déjà plus de sens et permet de mieux comparer ces deux résultats. Enfin, les approximations que nous avons évoquées dans la partie 3.1 qui mettent en jeu des rapports de la forme $k/\Delta - o(1/\Delta)$ illustrent également l'intérêt de s'intéresser à des notions asymptotiques fondées sur le degré, au moins pour ce problème. Nous montrerons, dans la section 5, les énoncés simplifiés de ces résultats.

La définition que nous proposons dans le paragraphe 4.2.4 donne la possibilité de considérer différents types de paramètres et fixe un cadre d'interprétation des résultats asymptotiques permettant de les comparer. Le paramètre intervenant dans la définition doit mesurer, d'une manière ou d'une autre, la difficulté des

instances de sorte que considérer les situations limites où ce paramètre tend vers l'infini réponde à l'objet que nous nous sommes fixé.

4.2.3. Ordre de difficulté

La notion d'ordre de difficulté a cette vocation. Elle fait référence à un niveau d'approximation car il est naturel de considérer que la difficulté (au sens de l'approximation) dépend du type d'approximation pris comme référence.

Définition 4. Soient un problème Π , \mathcal{I} son ensemble d'instances et un niveau d'approximation \mathcal{F} pour Π tel que Π n'appartient pas à ce niveau (i.e., il n'existe pas, sous une hypothèse du type $P \neq NP$, de chaîne polynomiale pour Π garantissant un rapport dans \mathcal{F}). Alors une application $d : \mathcal{I} \rightarrow \mathbb{N}$ est un ordre de difficulté pour Π par rapport au niveau \mathcal{F} si, $\forall M \in \mathbb{N}$, la restriction de Π aux instances $I \in \mathcal{I}$ telles que $d(I) \leq M$ appartient au niveau \mathcal{F} .

Les instances de faible degré sont donc mieux approximables que les générales, ce qui permet de voir d comme un indicateur de difficulté pour ce niveau d'approximation.

Nous donnons ci-après des exemples immédiats illustrant cette notion et montrant que plusieurs concepts très classiques de théorie de la complexité peuvent s'exprimer en termes d'ordre de difficulté. Dans le prochain paragraphe, nous l'utilisons pour définir une limite par rapport aux instances. Mais l'intérêt de cet outil dépasse largement la définition d'un cadre asymptotique ; en particulier, il permet de structurer les instances d'un problème en vue de mieux comprendre ce qui conditionne leur difficulté. Nous développerons alors des exemples plus techniques qui risqueraient, dans ce paragraphe, de nuire à la fluidité du document.

Plusieurs notions traditionnelles de complexité sont liées à la définition 4. Pour tout problème d'optimisation \mathbf{NP} -difficile, la taille de l'instance constitue un ordre de difficulté pour le niveau de la résolution exacte (i.e., pour le niveau d'approximation correspondant au rapport constant égal à 1). De même, les problèmes pseudo-polynomiaux sont définis dans [19] à partir d'une fonction $\max : I \mapsto \max\{I\}$ représentant la valeur du plus grand nombre intervenant dans I . Par exemple, dans le cas de problèmes pondérés, \max est la valeur du plus grand poids. Les problèmes \mathbf{NP} -complets pseudo-polynomiaux s'interprètent comme les problèmes pour lesquels $\log \max$ est un ordre de difficulté pour la résolution exacte. Pour les problèmes STABLE et COLORATION, le degré maximum est un ordre de difficulté par rapport au niveau d'approximation des rapports (classiques) constants. Il en est de même pour la taille du plus grand ensemble dans H-TRANSVERSAL.

Avant d'aller plus loin dans l'analyse de cette notion, nous nous intéressons, dans le prochain paragraphe, à son utilisation pour définir un cadre cohérent de notions asymptotiques.

4.2.4. *Proposition de définition*

Le concept d'ordre de difficulté nous permet de définir une notion de limite des instances. Remarquons que dans la définition 4, le fait que Π n'appartient pas¹¹ au niveau \mathcal{F} permet d'établir le résultat suivant.

Proposition 4.5. *Soit un problème Π d'ensemble d'instances \mathcal{I} et d'un ordre de difficulté pour Π par rapport à un niveau d'approximation \mathcal{F} . On a : $\sup_{I \in \mathcal{I}} d(I) = +\infty$.*

Nous proposons de définir des notions asymptotiques en considérant les instances ayant un grand degré de difficulté. Étant donné un problème Π et un degré de difficulté d , un résultat d'approximation asymptotique se rapporte aux instances I vérifiant $d(I) \geq D$, où D est une constante arbitrairement grande. Cette notion dépend du degré de difficulté choisi mais permet une définition générique pour l'ensemble des problèmes.

La démarche consiste donc à exclure les instances dont le degré de difficulté ne dépasse pas D . Par définition, de telles instances sont plus faciles à approximer que le problème général, ce qui justifie la pertinence de ce point de vue dans le cadre de l'approximation. Intuitivement, le résultat, qualifié d'asymptotique, peut s'interpréter comme la "partie prépondérante" d'un résultat plus lourd à exprimer pour l'ensemble des instances. Sous ce point de vue, la partie "négligée", ou plutôt non spécifiée, correspond à des instances mieux approximables, ce qui justifie qu'on s'autorise, dans la formulation du résultat, à se focaliser sur les instances de grand degré de difficulté. La proposition 4.6 formalise cette intuition.

Définition 5. *Étant donné un ordre de difficulté (associé à un niveau d'approximation \mathcal{F}) $d : \mathcal{I} \rightarrow \mathbb{N}$ des instances d'un problème Π , on dit qu'une chaîne algorithmique $(A_k)_{k \in \mathbb{N}}$ pour Π a un rapport asymptotique $f \in \mathcal{F}_\Pi$ si : $\forall \epsilon > 0$ et $\forall k \in \mathbb{N}$, $\exists D$ tel que $\forall I \in \mathcal{I}$ avec $d(I) \geq D$, $\rho_k(I) \geq f(I, k)(1 - \epsilon)$.*

Dans le cadre plus restreint des algorithmes approchés, la définition équivalente devient ($f \in \mathcal{F}_\Pi$) : $\forall \epsilon > 0$, $\exists D$ tel que $\forall I \in \mathcal{I}$ avec $d(I) \geq D$, $\rho(I) \geq f(I)(1 - \epsilon)$, si bien que la définition 5 correspond à une suite de rapports asymptotiques.

Les notions de rapport asymptotique d'un algorithme et d'un schéma d'approximation asymptotique définies dans [19, 35] correspondent au cas où le degré de difficulté est la valeur optimale. L'intérêt de notre définition est d'abord d'offrir un cadre permettant l'usage d'autres degrés de difficulté. Nous verrons dans les prochains paragraphes qu'elle permet de rendre compte de nombreux résultats qualifiés jusque là d'asymptotiques sans définition précise de ce terme. Mais elle a aussi le mérite de mettre l'accent sur le débat concernant la pertinence des résultats asymptotiques. Conformément au choix que nous avons fait, cette définition exclut du cadre asymptotique certains résultats qualifiés jusqu'ici comme tels mais qu'il ne nous semble pas opportun d'évoquer en ces termes afin d'en éviter de mauvaises interprétations. C'est le cas de BINPACKING et du résultat à rapport $1 - 1/\beta(I) - 1/k$ que nous avons déjà discuté. Du point de vue de notre définition, on ne peut plus

¹¹Sous une hypothèse du type $\mathbf{P} \neq \mathbf{NP}$.

parler de *schéma d'approximation asymptotique*. Nous avons délibérément choisi de limiter l'usage de résultats asymptotiques à la simplification d'énoncés complexes par la mise en évidence de leur partie prépondérante. La proposition suivante découle immédiatement de la définition et généralise une remarque que nous avons faite au paragraphe 4.2.2. Elle illustre qu'un résultat asymptotique fournit bien une représentation, simplifiée mais fidèle, de résultats plus lourds.

Proposition 4.6. *En adoptant les notations de la définition 5 :*

- 1° *si Π admet une chaîne d'approximation avec un rapport asymptotique $f \in \mathcal{F}'$ par rapport à un ordre de difficulté associé au niveau \mathcal{F} , alors pour tout $\epsilon > 0$, il existe une chaîne à rapport dans $\mathcal{F} \cup (1 - \epsilon)\mathcal{F}'$, où $(1 - \epsilon)\mathcal{F}' = \{(1 - \epsilon)f' : f' \in \mathcal{F}'\}$;*
- 2° *si Π admet un algorithme d'approximation avec un rapport asymptotique f , et si $\forall g \in \mathcal{F}, g \geq f$, alors Π admet une chaîne d'approximation convergeant vers f .*

Le principal cas où la condition $\forall g \in \mathcal{F}, g \geq f$ est satisfaite est $\mathcal{F} = \{1\}$ (correspondant à la résolution exacte). Cependant, dans l'expression $\mathcal{F} \cup (1 - \epsilon)\mathcal{F}'$, \mathcal{F}' constitue la composante prépondérante sachant que \mathcal{F} est, dans ce type de résultat, un niveau d'approximation " meilleur " que \mathcal{F}' .

4.3. CONVERGENCE ASYMPTOTIQUE

Mentionnons enfin la possibilité de mixer les deux notions de limite. Parmi les différentes manières de le faire, la *convergence asymptotique*, présentée dans la définition suivante, est la seule notion mixte qui nous a jusqu'ici été utile.

Définition 6. *Étant donné un ordre de difficulté (associé à un niveau d'approximation \mathcal{F}) $d : \mathcal{I} \rightarrow \mathbb{N}$ des instances d'un problème Π , on dit qu'une chaîne d'approximation polynomiale $(\mathbf{A}_k)_{k \in \mathbb{N}}$ converge asymptotiquement vers un rapport $\tilde{\rho} \in \mathcal{F}'_{\Pi}$ si : $\forall \epsilon > 0, \exists \kappa$ tel que $\forall k \geq \kappa, \exists D$ tel que $\forall I \in \mathcal{I}$ avec $d(I) \geq D, \rho(I, k) = \rho_{\mathbf{A}_k}(I) \geq \tilde{\rho}(I)(1 - \epsilon)$.*

Cette définition est équivalente à dire que $(\mathbf{A}_k)_{k \geq 0}$ garantit une suite de rapports asymptotiques qui convergent uniformément vers $\tilde{\rho}$. Dans le cas où $\tilde{\rho}$ est la fonction constante égale à 1 et où $d = \beta$, on retrouve la définition usuelle d'un schéma d'approximation asymptotique (voir Paragr. 4.2.1).

5. EXEMPLES

De nombreuses connexions existent entre les concepts que nous avons présentés dans les deux précédentes sections. C'est pourquoi nous regroupons ici quelques exemples significatifs de leur utilisation qui illustrent leur intérêt. Nous avons sélectionné ces résultats de manière à mettre en évidence leur lien avec le formalisme que nous avons décrit et les différentes questions qu'ils soulèvent. Cet objectif

justifie leur classement par type, par référence aux notions auxquelles ils se rapportent. Ce plan nous amène donc à évoquer les résultats relatifs à un problème donné sans chercher à respecter leur chronologie ni un ordre dans leur qualité.

Comme nous l'avons déjà signalé, l'intérêt de la définition de *niveau d'approximation* est qu'elle permet d'associer, de manière unifiée, la plupart des résultats d'approximation à des classes de problèmes.

5.1. ALGORITHMES D'APPROXIMATION

Dans le formalisme des chaînes d'approximation, les algorithmes d'approximation correspondent au cas des rapports f indépendants de k :

$$\exists \tilde{f} : \mathcal{I} \rightarrow]0, 1[, \forall I \in \mathcal{I}, \forall k \in \mathbb{N}, f(I, k) = \tilde{f}(I).$$

Le rapport f peut alors être assimilée à un élément de \mathcal{F}'_{Π} et on peut, sans perte de généralité, restreindre la chaîne d'approximation à un unique algorithme A en sélectionnant l'un des A_k .

5.1.1. Classe APX

NPO peut être structuré en fonction de la dépendance de \tilde{f} par rapport au paramètre I . Si \tilde{f} est une constante universelle, on retrouve le niveau **APX** des algorithmes à rapport γ constant.

Citons par exemple l'algorithme de couplage maximal pour le problème **G-TRANSVERSAL** qui consiste à prendre tous les sommets couverts par un couplage maximal pour l'inclusion. Il s'agit bien d'une couverture de sommets : le couplage étant maximal, toute arête du graphe est adjacente à au moins une arête de ce couplage. Par ailleurs, une couverture doit nécessairement contenir au moins une extrémité de chaque arête du couplage (pour couvrir ces arêtes). Cet algorithme polynomial garantit donc le rapport $\gamma = 1/2$, c'est-à-dire $g - transversal \in \mathbf{APX}$.

Le niveau **APX** peut lui-même être divisé en sous-niveaux en fonction des valeurs de la constante.

5.1.2. Rapports dépendant de l'instance

De nombreux problèmes sont connus comme ne faisant pas partie de **APX**¹². En particulier, si $\mathbf{P} \neq \mathbf{NP}$, ni **STABLE**, ni **CLIQUE** [1], ni **COLORATION** [34] ne sont dans **APX**. De tels résultats justifient qu'on s'intéresse à des rapports \tilde{f} pour lesquels $\inf_{I \in \mathcal{I}} \tilde{f}(I) = 0$. On retrouve la notion de rapport d'approximation dépendant de l'instance. L'intérêt d'avoir une gamme riche de niveaux d'approximation est de pouvoir distinguer différentes couches d'approximabilité dans $\mathbf{NPO} \setminus \mathbf{APX}$.

Les deux principales couches usuelles sont **Log-APX** et **Poly-APX** qui correspondent respectivement au cas où \tilde{f} est une fonction logarithmique et polynomiale

¹²Sous une hypothèse de complexité.

par rapport à la taille du problème. L'un des premiers problèmes à avoir été identifié dans **Log-APX** est **H-TRANSVERSAL** (cf. [30] pour sa version non-pondérée et [6] pour sa version pondérée). Il montre l'intérêt de pouvoir subdiviser ce niveau d'approximation puisque le rapport $1/(1+\log n)$ peut être garanti alors qu'il existe une constante $0 < c < 1$ telle que si $\mathbf{P} \neq \mathbf{NP}$, aucun algorithme polynomial ne peut garantir le rapport $1/c \log n$ [38].

La classe **Poly-APX** contient de très nombreux problèmes et en particulier **COLORATION**, **STABLE**, **CLIQUE**. Là encore, comme nous l'avons déjà signalé dans le paragraphe 3.3, on a besoin de différencier cette classe en fonction du degré du polynôme intervenant dans l'expression du rapport d'approximation. En effet, les classes usuelles ne permettent pas cette distinction ; pourtant les trois problèmes que nous venons de citer ne peuvent, sous l'hypothèse $\mathbf{NP} \neq \mathbf{ZPP}^{13}$, être approchés avec un rapport du type $O(1/n^{1-\epsilon})$ [17, 26] mais admettent par contre trivialement le rapport $1/n$. Ceci justifie aussi qu'on s'intéresse, pour ces problèmes, à des rapports du type $f(n)/n$ où $\forall \epsilon > 0, f(n) = o(n^\epsilon)$.

Les résultats actuels pour ces trois problèmes correspondent à des cas où f est une fonction poly-logarithmique de sorte de tels niveaux d'approximation semblent de plus en plus intéressants à considérer.

Théorème 1. • **STABLE** et **CLIQUE** sont approximables à rapport $O(\log^2 n/n)$ [5].
• **COLORATION** est approximable à rapport $O(\log^3 n/(n \log^2 \log n))$ [20].

5.1.3. Rapports fonction d'autres paramètres

Une question importante que notre travail permet de considérer de manière systématique est la *problématique des paramètres*. Elle consiste à envisager différents types de résultats non-constants en fonction des paramètres de l'instance qui interviennent dans l'expression du rapport. En effet, si les outils classiques n'envisagent que des rapports fonction de la taille du problème, de nombreux résultats d'approximation (voir par exemple le Paragr. 4.2.2) font intervenir d'autres paramètres de l'instance, notamment le degré maximum pour des problèmes de graphes. Enrichir le formalisme n'est pas absolument indispensable pour établir de tels résultats ; toutefois, une telle extension a l'intérêt de mettre en évidence l'existence de cette question et surtout offre un cadre adéquat pour envisager des études autour de cette problématique. Dans ce cadre, il nous semble en particulier intéressant d'étudier dans quelle mesure, pour un problème donné, le choix du paramètre conditionne le résultat. C'est notamment une manière d'étudier quel paramètre conditionne la difficulté d'une instance. Le fait que le formalisme classique n'intègre pas la possibilité d'exprimer les rapports d'approximation en fonction de différents paramètres est autant le signe que cette question n'a pas été envisagée de manière systématique que le frein à son étude.

Pour illustrer le type de résultats auquel cette question peut conduire, revenons au problème de **STABLE**. Souvenons-nous (cf. Paragr. 3.1) que deux familles

¹³**ZPP** désigne la classe des problèmes résolus sans erreur par un algorithme non-déterministe de complexité moyenne polynomiale ; $\mathbf{P} \subset \mathbf{ZPP} \subset \mathbf{NP}$.

de résultats d'approximation ont été développées pour STABLE, l'une fonction du paramètre n et l'autre fonction des paramètres Δ (degré maximum) et μ (degré moyen). Elles sont radicalement différentes de deux points de vue. Elles diffèrent d'abord dans la forme des rapports d'approximation : au résultat du théorème 1 a longtemps été opposée une suite de résultats du type k/Δ ou k/μ pour des valeurs particulières de k . Mais le plus surprenant est que ces deux familles de résultats diffèrent surtout par l'étagement des résultats et des améliorations successives. Du point de vue du paramètre n , on peut considérer que le résultat de [5] apporte un gain d'un facteur $\log^2 n$ par rapport au résultat le plus trivial alors que du point de vue du paramètre Δ , les nombreux travaux relatifs à ce problème n'ont longtemps permis de n'apporter qu'une amélioration d'un facteur constant, l'enjeu de chaque amélioration étant d'améliorer la constante. Les résultats de la forme $k/\Delta - o(1/\Delta)$ pour tout k [14] ont clos cette question pour poser celle du dépassement de l'ordre de grandeur $O(1/\Delta)$; nous l'avons déjà évoquée dans le paragraphe 3.1. Une première réponse positive réside dans le théorème suivant, où nous notons WSTABLE le problème du stable pondéré (*cf.* annexe A).

Théorème 2. [15] WSTABLE est approximable au niveau

$$\min \left\{ \frac{\log n}{3(\Delta + 1) \log \log n}, O \left(n^{-4/5} \right) \right\}. \tag{6}$$

Ce résultat est intéressant par sa forme : dans l'expression (6) le terme de droite est un rapport qu'on ne peut garantir en temps polynomial pour toute instance [26], ce qui rend le terme de gauche prépondérant. Or, celui-ci est justement une expression qui domine l'ordre de grandeur $O(1/\Delta)$. L'autre remarque intéressante sur ce théorème est qu'il fait conjointement intervenir les paramètres n et Δ . Notons enfin que le résultat est directement établi pour le problème pondéré.

Cette question des paramètres se pose aussi pour COLORATION. On peut chercher à opposer au théorème 1 des résultats fonction du degré Δ . On peut notamment obtenir des énoncés du même type que ceux sur STABLE, ce qui montre l'intérêt de cette question qui n'est pas le fait d'un comportement exceptionnel du problème STABLE.

Théorème 3. [15] $\forall \epsilon$, COLORATION est approximable au niveau

$$\min \left\{ \frac{\log n}{\Delta \log \log n}, O \left(n^{-\epsilon} \right) \right\}. \tag{7}$$

Pour les problèmes pondérés, les poids maximum et minimum constituent des paramètres intéressants. Dans de nombreux cas, notamment pour les problèmes STABLE et CLIQUE, les analyses d'algorithmes pour les versions pondérées peuvent aboutir à des rapports d'approximation indépendants des poids, comme l'illustre le théorème 2. Cette situation est considérée comme favorable vu que les poids peuvent prendre de grandes valeurs, même pour des instances petites.

Pour d'autres problèmes par contre, les poids interviennent dans les analyses. Arrêtons-nous un instant sur le cas, très particulier, du problème de *stable maximal*

de poids minimum (noté WSTABLEMAXMIN dans ce qui suit). Nous établissons le résultat suivant pour la version PWSTABLEMAXMIN correspondant au cas où les poids sont bornés par un polynôme; le résultat est donc *a fortiori* vrai pour WSTABLEMAXMIN .

Proposition 5.1. *Si $P \neq NP$, alors aucun algorithme ne peut garantir le rapport $n/(4w_{\max})$ pour PWSTABLEMAXMIN .*

Preuve. La preuve se fait par réduction polynomiale [31] du problème de 3-COLORABILITÉ qui est \mathbf{NP} -complet [19]. Une instance est un graphe et la question est de savoir s'il peut être coloré avec trois couleurs de sorte que deux sommets adjacents n'ont pas la même couleur.

Soit $G = (V, E)$ une instance de 3-COLORABILITÉ. On construit alors le graphe pondéré ($G' = (V', E'), w$), instance de WSTABLEMAXMIN , de la manière suivante :

$$\begin{cases} V' = V \times \{1, 2, 3, 4\} \\ ((v, i); (v', j)) \in E' \Leftrightarrow [(i = j) \wedge vv' \in E] \vee [i \neq j \wedge v = v']. \end{cases}$$

Le graphe G' est donc constitué de quatre copies de G , les copies de chaque sommet formant des cliques de taille 4. On affecte aux sommets des trois premières copies le poids 1 : pour $i = 1, 2, 3$, $W_{(v,i)} = 1$. On affecte aux sommets de la quatrième copie le même poids $W > n = |V|$. Ici $W = w_{\max}$. Pour garantir au graphe pondéré construit d'être instance de PWSTABLEMAXMIN , il suffit de veiller à ce que $W \leq n^2$. Il est clair que la construction peut être effectuée en temps polynomial. En particulier, G' a $4n$ sommets.

Si G est 3-colorable, soit S_1, S_2, S_3 trois ensembles stables de G disjoints deux à deux et couvrant V . L'ensemble $S' = S_1 \times \{1\} \cup S_2 \times \{2\} \cup S_3 \times \{3\}$ constitue un stable de G' . De plus, il est maximal pour l'inclusion car pour tout sommet $(v, j) \notin S'$, il existe $i \in \{1, 2, 3\}$ tel que $i \neq j$ et $(v, i) \in S'$. Alors $((v, i), (v, j)) \in E'$. Par ailleurs, le poids de S' est $w(S') = |S_1| + |S_2| + |S_3| = n$. La valeur minimum d'un stable maximal est donc au plus n .

Si le nombre chromatique de G vérifie $\chi(G) > 3$, considérons une coloration en $\chi(G)$ couleurs et soit S_i les sommets colorés avec la couleur i , $1 \leq i \leq \chi(G)$. Puisque $\chi(G) > 3$, il existe $v \in S_4$. Soit $S'_i = S_i \times \{i\}$, $1 \leq i \leq 3$; l'ensemble $S'_1 \cup S'_2 \cup S'_3 \cup \{(v, 4)\}$ est stable donc $w(S') > W$.

Par conséquent, supposons qu'un algorithme polynomial pour WSTABLEMAXMIN garantisse le rapport $n/(4w_{\max})$ pour tout graphe d'ordre n avec pour poids maximum w_{\max} . Pour G' d'ordre $4n$ et de poids maximum W , le rapport garanti serait n/W . Par conséquent, si G est 3-colorable, la solution retournée par l'algorithme serait de valeur au plus W alors que dans le cas contraire elle serait de valeur strictement supérieure à W . En conclusion cet algorithme permettrait de décider, en temps polynomial, si G est 3-colorable ou non, ce qui conclut la preuve. \square

En particulier, il suffit de poser, pour w_{\max} , une valeur polynomiale en n pour interdire un rapport du type $P(1/n)$ où P est un polynôme. Ceci signifie que le problème PWSTABLEMAXMIN pour lequel les poids sont polynomialement bornés

est hors de la classe **Poly-APX**. Ce résultat montre une différence notoire entre **WSTABLEMAXMIN** et **WSTABLE**. Ce dernier problème (donc en particulier sa version **PWSTABLE**) appartient évidemment à **Poly-APX** : il suffit de sélectionner un stable contenant le poids maximum pour garantir le rapport $1/n$. Si on s'intéresse maintenant à la version générale de **WSTABLEMAXMIN**, remarquons que pour toute fonction $f(G)$ associant à un graphe non pondéré une valeur dans $]0, 1]$, aucun algorithme pour **WSTABLEMAXMIN** ne peut garantir le rapport $f(G)$. En d'autres termes, tout rapport d'approximation pour ce problème dépend nécessairement des poids. Il s'agit d'une situation tout à fait remarquable pour laquelle un paramètre (plus exactement une famille de paramètres) est incontournable.

Ce résultat illustre la diversité des situations pouvant intervenir dans le domaine des rapports non-constants et apporte une justification de plus à la question des paramètres et à leur prise en compte dans le formalisme de l'approximation.

5.2. CHAÎNES D'APPROXIMATION

La possibilité d'étudier globalement le comportement d'une famille d'algorithmes constitue l'un des apports de ce travail. Il s'agit également d'une forme de paramétrage, par rapport à l'algorithme cette fois.

L'ensemble des fonctions $f \in \mathcal{F}_\Pi$ constantes en I , *i.e.*, des fonctions pour lesquelles il existe une suite $(x_k)_{k \in \mathbb{N}}$ de $]0, 1[$ telle que $\forall (I, k) \in \mathcal{I} \times \mathbb{N}, f(I, k) = x_k$, définit le niveau des *chaînes indépendantes de l'instance*. Remarquons que le cas où (x_k) est une suite convergente vers 1 correspond à un schéma d'approximation polynomial. Cette notion se distinguait jusqu'ici parmi les outils d'approximation. Le formalisme unifié permet de la rapprocher d'autres situations et de définir la classe **PTAS** comme un niveau d'approximation.

La convergence vers 1 a un statut particulier puisque que le rapport limite correspond à la résolution exacte. La première extension est la convergence d'une chaîne vers un autre rapport constant. Afin d'illustrer cette situation, citons un résultat pour le problème **H-3-TRANSVERSAL** correspondant aux instances de **H-TRANSVERSAL** pour lesquelles les ensembles contiennent au plus 3 éléments. Un résultat établi dans [22] s'interprète comme une convergence.

Théorème 5.2. [22] **H-3-TRANSVERSAL** admet une chaîne d'approximation polynomiale à rapport (γ) convergeant vers $5/7$.

Le cadre de l'approximation de **STABLE** en fonction du paramètre Δ fut la première motivation à considérer des chaînes d'approximation. Nous reformulons ci-dessous (Prop. 5.3) certains résultats évoqués dans le paragraphe 3.1 afin de mieux les comparer. Cela montre aussi comment le formalisme que nous avons introduit permet de simplifier leur énoncé (le point 5.3 de la Prop. 5.3 et démontré dans [24], tandis que les points 5.3 et 5.3 sont démontrés dans [14]).

Proposition 5.3. 1° **STABLE** admet une chaîne d'approximation polynomiale de complexité $O(n^k)$ convergeant asymptotiquement (par rapport à l'ordre de difficulté Δ) vers $(6/\Delta)$.

- 2° Il existe une chaîne d'approximation polynomiale de complexité $O(n^{k/2})$ pour STABLE garantissant le rapport d'approximation asymptotique (par rapport à Δ) k/Δ .
- 3° Il existe une chaîne d'approximation polynomiale de complexité $O(n^k)$ pour STABLE garantissant le rapport d'approximation asymptotique (par rapport à μ)

$$\min \left\{ \frac{k}{\mu}, k' \frac{\log \log \Delta}{\Delta} \right\}. \quad (8)$$

Le résultat du point 5.3 est plus compréhensible que sous sa forme originelle. Par ailleurs, cette présentation permet de le comparer au résultat du point 5.3 : pour $k = 6$, celui-ci correspond à un algorithme de complexité $O(n^3)$ garantissant le rapport asymptotique $6/\Delta$. Le gain par rapport au point 5.3 réside donc, d'une part dans la complexité, et d'autre part dans le type de garantie : un rapport asymptotique contre une convergence asymptotique.

Bien entendu, ces résultats ont été supplantés par le théorème 2. Cependant, ce dernier ainsi que le théorème 3 (expression (7)) ont une version en termes de chaînes. Dans le théorème 5.4, les termes fonction de Δ sont moins bons que leur équivalent dans les théorèmes 2 et 3 ; par contre, les termes auxquels ils sont opposés sont encore plus improbables que dans ces deux théorèmes.

Théorème 5.4. [15]

- 1° WSTABLE admet une chaîne d'approximation garantissant comme rapport le minimum entre $k/(\Delta + 1)$ et $O(\log^{-3(k+1)/4} n)$;
- 2° $\exists k' > 0$ tel que COLORATION admet une chaîne d'approximation garantissant comme rapport (γ) le minimum entre $k'k/\Delta$ et $k' \log^{-3k/4} n$.

5.3. APPROXIMATION DIFFÉRENTIELLE

Le cadre du rapport différentiel δ donne lieu à différents types de résultats. Nous en citons ici quelques uns pour mémoire.

Un résultat assez remarquable est que COLORATION, particulièrement difficile à approximer du point de vue classique (γ), se prête bien à l'approximation différentielle. En effet, il est dans ce cadre approximable à rapport constant : les différents rapports différentiels obtenus ont été $1/2$ [8], $2/3$ [25], $5/7$ et $3/4$ [22] et enfin $289/360$ [16] ; par contre, si $\mathbf{P} \neq \mathbf{NP}$, il n'admet pas de schéma d'approximation polynomial [22].

Le problème de BINPACKING, de son côté, a donné lieu à plusieurs analyses différentielles. Dans [10] par exemple, nous analysons les différents algorithmes gloutons pour ce problème. Parmi eux, citons par exemple l'algorithme FFD (*first-fit-decreasing*) qui consiste à prendre les nombres par ordre décroissant et à les placer dans la première boîte pouvant les contenir. Du point de vue du rapport classique, cet algorithme garantit le rapport $2/3$ [39]. Du point de vue différentiel, nous montrons qu'il garantit le rapport $3/4$, cette borne étant atteinte. Par contre, son analyse asymptotique met en évidence le rapport $7/9$ par rapport à l'ordre de difficulté $\sigma(L)$ (le support de l'instance défini au Paragr. 4.2.3). Notons d'ailleurs

que la notion de support a initialement été utile pour définir un rapport différentiel asymptotique. Ici, $\sigma(L)$ est un ordre de difficulté par rapport à la résolution exacte. D'après la proposition 4.6, FFD peut donc être facilement converti en une chaîne d'approximation convergeant vers $7/9$. Plus généralement, l'étude différentielle des différents algorithmes gloutons pour BINPACKING mettent en évidence des phénomènes relativement similaires au cas classique.

Par contre, du point de vue général de l'approximation, BINPACKING s'avère mieux approximable sous le rapport différentiel que sous le rapport classique. En effet, un schéma d'approximation différentiel est établi dans [9], ce qu'on ne peut obtenir du point de vue classique (Prop. 4.1).

6. CONCLUSION

Cet article se veut une tribune sur l'approximation polynomiale dans laquelle nous présentons, avec un regard critique, différents concepts de ce domaine, des exemples les illustrant et les débats qu'ils soulèvent. La plupart de ces concepts sont relativement récents, certains sont même présentés ici dans leur première mouture. Nous relatons donc, en quelque sorte, la naissance (encore en cours) d'un cadre conceptuel pour l'approximation. Pour cela, nous avons choisi d'insister sur la démarche qui a conduit à modifier des concepts à partir de leur analyse critique et sur les possibilités qu'apportent ces nouveautés. Mais au-delà de l'évocation d'un nouveau cadre, la lecture de cet article est aussi l'occasion de se familiariser avec l'approximation, ses outils, ses enjeux et ses points de vue.

Ainsi, cette discussion s'adresse autant aux non spécialistes de l'approximation qu'aux chercheurs travaillant sur ce thème. Les premiers y trouveront l'occasion d'aborder ce domaine par une approche critique à laquelle ils peuvent accéder grâce aux exemples simples que nous avons sélectionnés. Les seconds y trouveront quelques aspects d'un débat sur la pertinence des concepts de l'approximation, débat auquel nous les convions car il reste d'actualité. Un intérêt de ce travail réside dans ce débat, peut-être même plus que dans les réponses que nous y apportons. En effet, alors que cette réflexion n'a jamais été menée, nous montrons à quel point elle conditionne les résultats et leur interprétation. Ils pourront également y trouver une présentation de nouveaux concepts assortie de quelques exemples significatifs.

La section 2 est une première introduction à l'approximation polynomiale. Elle présente des concepts clef qui sembleront presque banals à quiconque à déjà travaillé dans ce cadre. Cependant, ces rappels sont déjà l'occasion de poser un regard critique : nous évoquons en particulier la question de la mesure d'approximation qui a conduit à l'étude du rapport différentiel. Notre approche privilégie quelques idées phares mettant en évidence les enjeux de l'approximation ; ceux-ci nous semblent essentiels pour comprendre les concepts et les motivations qui nous conduisent à les affiner.

La section 3 s'intéresse à une classification " absolue " des problèmes d'optimisation **NPO**. La richesse de cette structure est illustrée par la grande diversité des résultats possibles. L'exemple de STABLE justifie le besoin d'enrichir la gamme usuelle des résultats en introduisant l'idée des chaînes d'approximation qui consistent à étudier globalement le comportement d'une suite d'algorithmes. Un autre intérêt de cette notion tient dans la possibilité d'exploiter le paramètre k des termes de la suite pour mettre en exergue une dépendance entre la complexité de l'algorithme d'approximation et la qualité qu'il garantit.

La section 4 complète la précédente en élargissant la gamme des outils permettant de décrire, dans l'absolu, les possibilités de résolution d'un problème difficile. Plus précisément elle s'intéresse à deux notions de limite en approximation qui offrent de nouvelles possibilités pour l'analyse fine d'algorithmes d'approximation. La première, directement associée au concept de chaîne d'approximation, permet de rendre compte de manière concise du comportement des algorithmes A_k de la chaîne lorsque le paramètre k tend vers l'infini. La qualité de l'approximation augmentant en général avec k , ce comportement limite permet d'évaluer le meilleur type de rapport d'approximation qu'on peut déduire de l'analyse de la chaîne. Cette information est intéressante à confronter à l'analyse en complexité de la chaîne algorithmique. La seconde notion, quant à elle, est celle de *comportement asymptotique*, terme très souvent employé mais rarement défini. Il correspond à une certaine idée de limite des instances. Nous ne prétendons pas apporter une réponse absolue à cette question ; notre objet est plutôt de convaincre de sa pertinence. Il s'agit d'un exemple particulièrement significatif de l'importance qu'il y a, selon nous, à conserver un regard critique sur le type de résultats obtenus indépendamment de leur exactitude mathématique. Ce point nous semble essentiel pour permettre une exploitation des résultats et pour préserver l'intérêt de l'approximation comme outil de compréhension des problèmes. La réponse que nous proposons est clairement orientée et relève d'un choix que nous avons explicité et justifié ; elle consiste à voir le résultat limite comme une simplification du résultat général permettant de mieux appréhender sa teneur sur les instances les plus difficiles. Un autre intérêt de cette étude est justement de soulever la question de la difficulté des instances du point de vue de l'approximation. En effet, la notion d'*ordre de difficulté*, qui nous a permis de définir de manière systématique l'idée d'" instances tendant vers l'infini ", représente une notion de difficulté.

Les sections 3 et 4 forment un tout puisqu'elles traitent d'outils permettant de décrire directement (et dans l'absolu) la possibilité d'approximer un problème **NPO**. La section 5 fournit alors quelques exemples significatifs de résultats relatifs à ces deux sections.

Alors que la plupart des outils usuels ont été mis au point pour l'étude de **APX**, ces exemples mettent en évidence de nouvelles problématiques qui se posent au delà de ce niveau d'approximation. Mentionnons en particulier la question des paramètres en fonction desquels exprimer un résultat d'approximation ainsi que leurs liens. Cette question rejoint d'ailleurs celle de la difficulté intrinsèque des instances qui est sous-jacente à la définition d'un rapport asymptotique. En fait la notion de paramètre pourrait être érigée comme point charnière de toute notre

étude : les chaînes d'approximations, les degrés de difficulté et même la conception d'un formalisme unique pour les différentes mesures d'approximation sont des types différents de paramétrages qui illustrent, d'une part la grande souplesse et d'autre part la " fibre unificatrice de ce travail " .

Toutefois, ces outils que nous avons présentés ne permettent de répondre que partiellement aux enjeux de l'approximation tels que nous les avons identifiés. Les possibilités qu'ils ouvrent se limitent à une " description statique et opérationnelle " des propriétés d'approximation en fonction des résultats d'approximation possibles. Statique en ce sens que la classification en niveaux d'approximation ne permet que de positionner les différents problèmes d'approximation dans la hiérarchie indépendamment les uns des autres. Opérationnel parce que ce point de vue est tout à fait adapté à la conception et à l'analyse d'algorithmes approchés pour des problèmes spécifiques. Il est en particulier très utile pour visualiser les possibilités d'approximation et les améliorations successives pour chaque problème. De ce fait, les notions présentées dans ce travail répondent pleinement à son objectif en tant qu'introduction à l'approximation polynomiale. Par contre, une étude reposant exclusivement sur ces outils et les possibilités qu'ils offrent ne nous renseignerait que très partiellement sur les liens entre problèmes et leur difficulté relative. En particulier, deux résultats d'approximation du même type pour différents problèmes ne sont pas nécessairement comparables.

La notion de *réduction en approximation* est justement adaptée pour comparer différents problèmes du point de vue de l'approximation, indépendamment des résultats connus pour ceux-ci, ce qui est particulièrement intéressant pour étudier une structure de **NPO**. Les réductions s'avèrent même être particulièrement riches, même au delà de résultats structurels. Elles offrent par exemple la possibilité de transférer des résultats d'approximation (positifs ou négatifs) d'un problème à un autre; de nombreux résultats d'approximation sont obtenus par ce biais. Enfin, ce concept est également très souple et peut être étendu au cadre que nous avons présenté dans cet article. En particulier, de nouveaux types de réductions permettent d'envisager toute question de difficulté relative entre problèmes, entre différents paramètres pour un problème donné, entre cadres classique et différentiel, entre versions pondérée et non pondérée, . . .

Dans l'article [13], nous proposons un prolongement de ce travail qui inclut les réductions et permet d'approfondir les questions structurelles au sein de **NPO**. Nous montrons aussi comment les outils de l'approximation permettent d'étudier, pour un problème donné, la structure de l'ensemble de ses instances.

Remerciements. Les remarques, commentaires et suggestions des deux lecteurs/lectrices anonymes ont très largement contribué à l'amélioration de la qualité scientifique et de la lisibilité de cet article. Qu'ils/elles en soient remercié(e)s.

ANNEXES

A. QUELQUES PROBLÈMES NPO

- **Stable maximum** (STABLE)
 INSTANCE : graphe $G = (V, E)$;
 SOLUTION RÉALISABLE : un ensemble stable, *i.e.*, un ensemble $V' \subseteq V$ tel que $\forall (v, v') \in V' \times V', vv' \notin E$;
 OBJECTIF : maximiser le cardinal de V' , *i.e.* $|V'|$;
 VERSION PONDÉRÉE (WSTABLE) : chaque sommet $v \in V$ a une valeur $w_v > 0$ et l'objectif consiste à maximiser $w(V') = \sum_{v \in V'} w_v$; PWSTABLE et BWSTABLE désignent les restrictions de WSTABLE, respectivement aux cas de poids polynomiaux et à celui de poids bornés par une constante.
- **Clique maximum** (CLIQUE)
 INSTANCE : graphe $G = (V, E)$;
 SOLUTION RÉALISABLE : une clique *i.e.*, un ensemble $V' \subseteq V$ tel que $\forall (v, v') \in V' \times V', vv' \in E$;
 OBJECTIF : maximiser le cardinal de V' ;
 VERSION PONDÉRÉE (WCLIQUE) : chaque sommet $v \in V$ a une valeur $w_v > 0$; l'objectif consiste à maximiser $w(V') = \sum_{v \in V'} w_v$.
- **Couverture minimum de sommets**¹⁴ (G-TRANSVERSAL)
 INSTANCE : graphe $G = (V, E)$;
 SOLUTION RÉALISABLE : une couverture de sommets, *i.e.*, un ensemble $V' \subseteq V$ tel que $\forall vv' \in E, \{v, v'\} \cap V' \neq \emptyset$;
 OBJECTIF : minimiser le cardinal de V' .
- **Coloration minimum des sommets** (COLORATION)
 INSTANCE : graphe $G = (V, E)$;
 SOLUTION RÉALISABLE : une coloration de G , *i.e.*, une partition de V en ensembles stables disjoints non vides (V_1, \dots, V_p) ;
 OBJECTIF : minimiser le nombre p de stables V_i ;
 COMMENTAIRE : chaque stable correspond à une couleur, il s'agit alors de colorer les sommets avec un nombre minimum de couleurs de sorte que deux sommets adjacents n'ont pas la même couleur.
- **Stable maximal minimum** ou **Stable dominant minimum** (STABLE-MAXMIN)
 INSTANCE : graphe $G = (V, E)$;
 SOLUTION RÉALISABLE : un stable maximal pour l'inclusion V' (*i.e.*, pour tout sommet $v \notin V', V' \cup \{v\}$ non stable) ;
 OBJECTIF : minimiser le cardinal de V' ;
 VERSION PONDÉRÉE (WSTABLEMAXMIN) : chaque sommet $v \in V$ a une valeur $w_v > 0$ et l'objectif consiste à minimiser $w(V') = \sum_{v \in V'} w_v$; la version PWSTABLEMAXMIN est définie comme celle de STABLE.

¹⁴Ou **Transversal minimum d'un graphe** [3].

- **Arbre de Steiner**

INSTANCE : graphe complet $G = (V, E)$ dont chaque arête $e \in E$ a une valeur $w_e > 0$, un ensemble $R \subset V$;

SOLUTION RÉALISABLE : un sous-graphe partiel (V', E') de G qui est un arbre et tel que $R \subset V'$;

OBJECTIF : minimiser la valeur de l'arbre, *i.e.*, $\sum_{e \in E'} w_e$.

- **Voyageur de commerce** (TSP)

INSTANCE : graphe complet $G = (V, E)$ dont chaque arête $e \in E$ a une valeur $w_e > 0$;

SOLUTION RÉALISABLE : un cycle hamiltonien $H \subset E$ (graphe partiel connexe tel que tout sommet a un degré 2) ;

OBJECTIF : minimiser la valeur du cycle, *i.e.*, $\sum_{e \in H} w_e$;

COMMENTAIRE : Δ -TSP est la version de TSP où les distances sur les arêtes vérifient les inégalités triangulaires.

- **Couverture d'ensemble**¹⁵ (H-TRANSVERSAL)

INSTANCE : un ensemble de base E et une collection \mathcal{S} de parties de E couvrant E , *i.e.*, $\mathcal{S} \subset 2^E$, $\cup_{S \in \mathcal{S}} S = E$;

SOLUTION RÉALISABLE : une partie $S' \subseteq \mathcal{S}$ couvrant E , *i.e.*, $\cup_{S \in S'} S = E$;

OBJECTIF : minimiser le cardinal de S' ;

COMMENTAIRE : H-3-TRANSVERSAL est la version où $\forall S \in \mathcal{S}$, $|S| \leq 3$.

- **Bin-packing** (BINPACKING)

INSTANCE : une liste de rationnels $L = (x_1, \dots, x_n)$, $\forall i \in \{1, \dots, n\}$, $0 < x_i \leq 1$;

SOLUTION RÉALISABLE : une partition de L en ensembles (appelés *bins*) B_1, \dots, B_p de sorte que la somme des nombres de chaque *bin* ne dépasse pas 1 ;

OBJECTIF : minimiser le nombre p de *bins* ;

COMMENTAIRE : il s'agit donc de ranger les nombres en un nombre minimum de boîtes de capacité 1.

- **Sac à dos** (KS)

INSTANCE : deux ensembles de n rationnels $\{a_1, \dots, a_n\}$ et $\{b_1, \dots, b_n\}$, un rationnel B ;

SOLUTION RÉALISABLE : ensemble d'indices $S \subseteq \{1, \dots, n\}$ tel que $\sum_{i \in S} b_i \leq B$;

OBJECTIF : maximiser $\sum_{i \in S} a_i$;

COMMENTAIRE : il s'agit de la programmation linéaire en variables bivalentes et à une contrainte ; le problème existe aussi en version minimisation.

- **Couplage maximum d'un graphe**

INSTANCE : graphe $G = (V, E)$;

SOLUTION RÉALISABLE : un couplage, *i.e.*, $E' \subseteq E$ tel que les arêtes de E' sont 2 à 2 non adjacentes ;

OBJECTIF : maximiser $|E'|$;

COMMENTAIRE : le problème est polynomial [37].

¹⁵Ou **Transversal d'un hypergraphe** [3].

B. RAPPEL DE QUELQUES DÉFINITIONS DE BASE

Un algorithme polynomial approché est un algorithme qui détermine, pour chaque instance I d'un problème $\Pi \in \mathbf{NPO}$, une solution réalisable avec une complexité polynomiale en la taille de I . Il garantit un rapport $\rho(I)$ si cette solution vérifie, pour toute instance I de Π :

- 1° si on travaille avec le rapport γ (approximation classique ou standard)
 - 1° $\lambda(I)/\beta(I) \geq \rho(I)$, si Π est un problème de maximization, ou
 - 2° $\beta(I)/\lambda(I) \geq \rho(I)$, si Π est un problème de minimization
- 2° $|\omega(I) - \lambda(I)|/|\omega(I) - \beta(I)| \geq \rho(I)$ pour le cas de l'approximation différentielle (rapport δ).

Nous pouvons classer les algorithmes en fonction du meilleur rapport (connu) qu'ils garantissent ; c'est ce que nous appelons dans l'article niveau d'approximation. Nous donnons ci-après une liste des principales classes mentionnées dans cet article et dans la littérature du domaine :

- **Algorithme à rapport constant ρ** :
quand $\rho(I)$ est une constante ne dépendant pas de I , *i.e.*, ne dépendant d'aucun paramètre de l'instance du problème ;
- **Schéma d'approximation polynomial** :
suite d'algorithmes indicée par $\epsilon > 0$ garantissant le rapport $1 - \epsilon$; chaque algorithme est polynomial ;
- **Schéma complet d'approximation polynomial** :
schéma d'approximation dont la complexité est polynomial en $|I|$ et en $1/\epsilon$;
- **Algorithmes à rapport logarithmique** :
cas où $\rho(I) \geq O(1/\log |I|)$;
- **Algorithmes à rapport $O(n^{\epsilon-1})$** :
cas où $\exists \epsilon > 0$ tel que $\rho(I) \geq O(n^{\epsilon-1})$.

La classification considérée pour les algorithmes approchés peut être étendue aux problèmes si on considère que le rapport d'approximation d'un problème Π est le rapport du meilleur algorithme résolvant Π . Ainsi, on peut considérer une structure pour \mathbf{NPO} par rapport à l'approximabilité de ses problèmes. La figure 1 représente les classes d'approximabilité les plus usuelles, évidemment sous la conjecture $\mathbf{P} \neq \mathbf{NP}$. Par analogie, nous notons \mathbf{PO} la classe des problèmes d'optimisation polynomiaux.

Nous donnons ci-après les principales classes d'approximabilité pour l'approximation classique. Évidemment cette classification reste valable pour l'approximation différentielle.

- **APX** : la classe des problèmes \mathbf{NPO} admettant un algorithme à rapport constant ;
- **PTAS** : la classe des problèmes \mathbf{NPO} admettant un schéma d'approximation polynomial ;
- **FPTAS** : la classe des problèmes \mathbf{NPO} admettant un schéma complet d'approximation polynomial ;

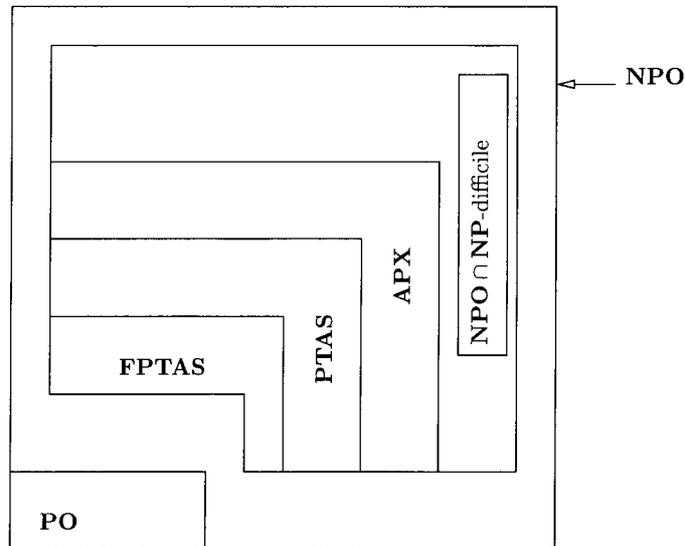


FIGURE 1. Le monde de l'approximabilité selon les conjectures des chercheurs.

Log-APX : la classe des problèmes **NPO** admettant un algorithme approché à rapport logarithmique en I ;

Poly-APX : la classe des problèmes **NPO** admettant un algorithme approché garantissant un rapport qui est un polynôme en la taille de l'instance.

Remarquons enfin que, comme la valeur d'un rapport d'approximation appartient à \mathbb{R}^+ , il existe tout un *continuum* de classes d'approximabilité.

RÉFÉRENCES

- [1] S. Arora, C. Lund, R. Motwani, M. Sudan et M. Szegedy, Proof verification and intractability of approximation problems, in *Proc. FOCS'92* (1992) 14-23.
- [2] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela et M. Protasi, *Complexity and approximation. Combinatorial optimization problems and their approximability properties*. Springer, Heidelberg (1999).
- [3] C. Berge, *Graphs and hypergraphs*. North Holland, Amsterdam (1973).
- [4] P. Berman et M. Fürer, Approximating maximum independent set in bounded degree graphs, in *Proc. Symposium on Discrete Algorithms* (1994) 365-371.
- [5] B.B. Boppana et M.M. Halldórsson, Approximating maximum independent sets by excluding subgraphs. *BIT* **32** (1992) 180-196.
- [6] V. Chvátal, A greedy-heuristic for the set covering problem. *Math. Oper. Res.* **4** (1979) 233-235.
- [7] S.A. Cook, The complexity of theorem-proving procedures, in *Proc. STOC'71* (1971) 151-158.

- [8] M. Demange, P. Grisoni et V.T. Paschos, Differential approximation algorithms for some combinatorial optimization problems. *Theoret. Comput. Sci.* **209** (1998) 107-122.
- [9] M. Demange, J. Monnot et V.T. Paschos, Bridging gap between standard and differential polynomial approximation : The case of bin-packing. *Appl. Math. Lett.* **12** (1999) 127-133.
- [10] ———, Maximizing the number of unused bins. *Found. Comput. Decision Sci.* **26** (2001) 169-186.
- [11] M. Demange et V.T. Paschos, On an approximation measure founded on the links between optimization and polynomial approximation theory. *Theoret. Comput. Sci.* **158** (1996) 117-141.
- [12] ———, Valeurs extrémales d'un problème d'optimisation combinatoire et approximation polynomiale. *Math. Inf. Sci. Humaines* **135** (1996) 51-66.
- [13] ———, Autour de nouvelles notions pour l'analyse des algorithmes d'approximation : de la structure de NPO à la structure des instances. *RAIRO : Oper. Res.* (à paraître).
- [14] ———, Improved approximations for maximum independent set *via* approximation chains. *Appl. Math. Lett.* **10** (1997) 105-110.
- [15] ———, *Towards a general formal framework for polynomial approximation*. LAMSADE, Université Paris-Dauphine, *Cahier du LAMSADE* **177** (2001).
- [16] R. Duh et M. Fürer, Approximation of k -set cover by semi-local optimization, in *Proc. STOC'97* (1997) 256-265.
- [17] U. Feige et J. Kilian, Zero knowledge and the chromatic number, in *Proc. Conference on Computational Complexity* (1996) 278-287.
- [18] W. Fernandez de la Vega, Sur la cardinalité maximum des couplages d'hypergraphes aléatoires uniformes. *Discrete Math.* **40** (1982) 315-318.
- [19] M.R. Garey et D.S. Johnson, *Computers and intractability. A guide to the theory of NP-completeness*. W.H. Freeman, San Francisco (1979).
- [20] M.M. Halldórsson, A still better performance guarantee for approximate graph coloring. *Inform. Process. Lett.* **45** (1993) 19-23.
- [21] ———, *Approximations via partitioning*. JAIST Research Report IS-RR-95-0003F, Japan Advanced Institute of Science and Technology, Japan (1995).
- [22] ———, Approximating k -set cover and complementary graph coloring, in *Proc. International Integer Programming and Combinatorial Optimization Conference*. Springer Verlag, *Lecture Notes in Comput. Sci.* **1084** (1996) 118-131.
- [23] M.M. Halldórsson et J. Radhakrishnan, Greed is good : Approximating independent sets in sparse and bounded-degree graphs, in *Proc. STOC'94* (1994) 439-448.
- [24] ———, Improved approximations of independent sets in bounded-degree graphs *via* subgraph removal. *Nordic J. Comput.* **1** (1994) 475-492.
- [25] R. Hassin et S. Lahav, Maximizing the number of unused colors in the vertex coloring problem. *Inform. Process. Lett.* **52** (1994) 87-90.
- [26] J. Håstad, Clique is hard to approximate within $n^{1-\epsilon}$. *Acta Math.* **182** (1999) 105-142.
- [27] D.S. Hochbaum, Efficient bounds for the stable set, vertex cover and set packing problems. *Discrete Appl. Math.* **6** (1983) 243-254.
- [28] ———, *Approximation algorithms for NP-hard problems*. PWS, Boston (1997).
- [29] O.H. Ibarra et C.E. Kim, Fast approximation algorithms for the knapsack and sum of subset problems. *J. Assoc. Comput. Mach.* **22** (1975) 463-468.
- [30] D.S. Johnson, Approximation algorithms for combinatorial problems. *J. Comput. System Sci.* **9** (1974) 256-278.
- [31] R.M. Karp, Reducibility among combinatorial problems, dans *Complexity of computer computations*, édité par R.E. Miller et J.W. Thatcher. Plenum Press, New York (1972) 85-103.
- [32] S. Khanna, R. Motwani, M. Sudan et U. Vazirani, On syntactic *versus* computational views of approximability. *SIAM J. Comput.* **28** (1998) 164-191.
- [33] H.R. Lewis et C.H. Papadimitriou, *Elements of the theory of computation*. Prentice-Hall (1981).
- [34] C. Lund et M. Yannakakis, On the hardness of approximating minimization problems. *J. Assoc. Comput. Mach.* **41** (1994) 960-981.

- [35] R. Motwani, *Lecture notes on approximation algorithms*, Vol. I. Stanford University (1993).
- [36] G.L. Nemhauser, L.A. Wolsey et M.L. Fischer, An analysis of approximations for maximizing submodular set functions. *Math. Programming* **14** (1978) 265-294.
- [37] C.H. Papadimitriou et K. Steiglitz, *Combinatorial optimization : Algorithms and complexity*. Prentice Hall, New Jersey (1981).
- [38] R. Raz et S. Safra, A sub-constant error probability low-degree test and a sub-constant error probability PCP characterization of NP, in *Proc. STOC'97* (1997) 475-484.
- [39] D. Simchi-Levi, New worst-case results for the bin-packing problem. *Naval Res. Logistics* **41** (1994) 579-585.
- [40] P. Turán, On an extremal problem in graph theory (in Hungarian). *Mat. Fiz. Lapok* **48** (1941) 436-452.
- [41] V. Vazirani, *Approximation algorithms*. Springer, Heidelberg (2001).