



**HAL**  
open science

## Identification de sommets dans les graphes

Julien Moncel

► **To cite this version:**

Julien Moncel. Identification de sommets dans les graphes. Automatique / Robotique. Université Paul Sabatier - Toulouse III, 2012. tel-00803528

**HAL Id: tel-00803528**

**<https://theses.hal.science/tel-00803528>**

Submitted on 22 Mar 2013

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

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

UNIVERSITÉ PAUL SABATIER  
ÉCOLE DOCTORALE SYSTÈMES  
**HABILITATION À DIRIGER DES  
RECHERCHES**

Présentée et soutenue par  
Julien MONCEL

**Identification de sommets dans les  
graphes**

Habilitation dirigée par Pierre LOPEZ  
préparée au Laboratoire d'Architecture et  
d'Analyse des Systèmes  
soutenue le 3 juillet 2012

**Composition du jury :**

<i>Rapporteurs :</i>	Victor CHEPOI	- Université de la Méditerranée, LIF
	Oriol SERRA	- Universitat Politècnica de Catalunya
	Éric SOPENA	- Université Bordeaux 1, LaBRI
<i>Directeur :</i>	Pierre LOPEZ	- LAAS
<i>Présidente :</i>	Marie-Christine LAGASQUIE-SCHIEX	- Université Paul Sabatier, IRIT
<i>Examineurs :</i>	Olivier HUDRY	- Télécom ParisTech
	Bernard PENZ	- Grenoble INP, G-SCOP



Ce n'est qu'en essayant continuellement que l'on finit par réussir. Autrement dit : plus ça rate, plus on a de chances que ça marche.

*Devise Shadok*



# Remerciements

Je remercie chaleureusement tous les membres de mon jury, qui m'ont fait l'honneur et le plaisir d'assister à ma soutenance d'habilitation, le 3 juillet 2012 à Toulouse. Je remercie également toute l'équipe du département QLIO de l'IUT de Rodez, qui m'a apporté tout son soutien lors de la phase de rédaction du manuscrit. Merci également à toute l'équipe MOGISA du LAAS, et en particulier à Pierre Lopez et Christian Artigues pour leur confiance, leur soutien moral, logistique et leurs encouragements. Merci à ma famille, mes amis et aux relecteurs du présent manuscrit.

L'origine de ce travail remonte à bien des années en arrière. Le fait que ma trajectoire personnelle se soit dirigée vers les mathématiques discrètes doit beaucoup à quelques enseignants extraordinaires, au lycée et à l'université ; aux chroniques de Ian Stewart et de Jean-Paul Delahaye dans *Scientific American / Pour la Science* ; et enfin au cours « Jeux combinatoires et raisonnement mathématique » de l'université Joseph Fourier à Grenoble (alors dispensé par Denise Grenier et Charles Payan).

Les résultats présentés dans le présent document sont issus, pour la plupart, de travaux en collaboration. Merci à tous mes coauteurs, à leur bonne humeur, leur curiosité, leur rigueur, leur patience (qui hélas ! était parfois nécessaire) et pour la masse de travail accomplie. J'ai appris beaucoup de vous, de vos façons de penser, de chercher, de conjecturer, de réfuter et de prouver.

Enfin, je tenais à souligner que l'ensemble des résultats présentés dans ce document sont aussi le fruit d'un nombre incalculable d'idées fausses, de conjectures réfutées, de pistes qui se sont révélées stériles, d'inspirations fulgurantes restées sans suite. Merci à elles aussi !

Chercher, c'est se tromper, beaucoup, et trouver, parfois.



# Table des matières

<b>Table des matières</b>	<b>vi</b>
<b>Introduction</b>	<b>1</b>
<b>1 Curriculum vitæ</b>	<b>3</b>
1.1 Cursus . . . . .	3
1.2 Enseignement . . . . .	5
1.3 Responsabilités collectives . . . . .	6
1.4 Recherche . . . . .	7
1.4.1 Thématiques de recherche . . . . .	7
1.4.2 Direction de travaux de recherche . . . . .	12
1.4.3 Expertise . . . . .	14
1.4.4 Collaborations internationales . . . . .	15
1.5 Liste de publications . . . . .	16
1.5.1 Publications parues ou à paraître dans des revues in- ternationales à comité de lecture . . . . .	16
1.5.2 Participation à des conférences internationales . . . . .	18
1.5.3 Manuscrits soumis . . . . .	19
<b>2 Synthèse des travaux sur les codes identifiants</b>	<b>21</b>
2.1 Définitions et généralités sur les codes identifiants . . . . .	22
2.1.1 Définitions . . . . .	22
2.1.2 Trois exemples d'applications . . . . .	26
2.1.3 Dynamique de la thématique de recherche . . . . .	29
2.2 Structures régulières . . . . .	31
2.2.1 Grilles et bandes . . . . .	31
2.2.2 Cycles et puissances de cycles . . . . .	40
2.2.3 Hypercubes . . . . .	44
2.2.4 Produits de cliques . . . . .	48
2.2.5 Graphes de Sierpiński . . . . .	50
2.3 Aspects algorithmiques . . . . .	53
2.3.1 État de l'art . . . . .	53
2.3.2 Arbres orientés . . . . .	57
2.3.3 Fasciagraphes et rotagraphes . . . . .	60
2.3.4 Généralisation à d'autres problèmes de l'algorithme pour les fasciagraphes et les rotagraphes . . . . .	63
2.3.5 Approximabilité du problème . . . . .	66

---

2.4	Questions structurelles . . . . .	70
2.4.1	Graphes extrémaux pour l'identification d'un seul sommet . . . . .	71
2.4.2	Construction de graphes admettant de petits codes pour l'identification d'ensembles de sommets . . . . .	79
2.4.3	Structures des graphes admettant un code . . . . .	89
2.5	Codes adaptatifs . . . . .	96
2.5.1	Vers des dispositifs d'identification dynamiques . . . . .	96
2.5.2	Les codes identifiants vus comme un problème de recherche binaire . . . . .	97
2.5.3	Définition et intérêt du concept . . . . .	98
2.5.4	Identification adaptative dans les tores . . . . .	100
2.5.5	Autres structures . . . . .	105
<b>3</b>	<b>Conclusion et perspectives</b>	<b>107</b>
3.1	Bilan des principaux résultats . . . . .	107
3.2	Perspectives de recherche . . . . .	108
3.2.1	Questions ouvertes de la littérature . . . . .	108
3.2.2	Thématiques de recherche émergentes . . . . .	117
3.3	Projet personnel . . . . .	124
3.3.1	Codes identifiants . . . . .	124
3.3.2	Ordonnancement . . . . .	125
	<b>Index</b>	<b>129</b>
	<b>Bibliographie</b>	<b>131</b>
<b>A</b>	<b>CV détaillé</b>	<b>145</b>
<b>B</b>	<b>Échantillon de manuscrits</b>	<b>157</b>

# Introduction

Ce document est structuré en trois parties. La première partie est un curriculum vitæ détaillé, qui décrit mon parcours, les enseignements dans lesquels je me suis impliqué, les responsabilités collectives que j'ai exercées, ainsi que mes thèmes de recherches. On y trouvera en particulier une liste de publications.

Dans la deuxième partie, je fais une synthèse de l'ensemble des travaux auxquels j'ai contribué, sur le thème des codes identifiants dans les graphes. L'intérêt et l'impact de ces travaux est replacé dans le contexte de ce qui était connu en 2012 sur les codes identifiants. Dans cette partie, je me suis essentiellement attaché à donner et à expliquer les idées sous-jacentes aux travaux auxquels j'ai contribué. Les preuves détaillées, quant à elles, sont bien entendu disponibles dans les articles publiés en revue relatifs à cette partie.

Enfin, la troisième partie dresse un bilan de ces résultats, et propose quelques perspectives de recherche sur les codes identifiants dans les graphes. J'y ajoute quelques éléments au sujet de mes projets personnels de recherche pour les années à venir.



# CHAPITRE 1

## Curriculum vitæ

---

### Sommaire

---

<b>1.1 Cursus</b> . . . . .	<b>3</b>
<b>1.2 Enseignement</b> . . . . .	<b>5</b>
<b>1.3 Responsabilités collectives</b> . . . . .	<b>6</b>
<b>1.4 Recherche</b> . . . . .	<b>7</b>
1.4.1 Thématiques de recherche . . . . .	7
1.4.2 Direction de travaux de recherche . . . . .	12
1.4.3 Expertise . . . . .	14
1.4.4 Collaborations internationales . . . . .	15
<b>1.5 Liste de publications</b> . . . . .	<b>16</b>
1.5.1 Publications parues ou à paraître dans des revues internationales à comité de lecture . . . . .	16
1.5.2 Participation à des conférences internationales . . . . .	18
1.5.3 Manuscrits soumis . . . . .	19

---

## 1.1 Cursus

Je suis actuellement maître de conférences, en poste à l'IUT de Rodez, qui dépend de l'université Toulouse 1 Capitole. Je mène mes activités de recherche au sein de l'équipe MOGISA (Modélisation et Gestion Intégrée de Systèmes d'Activités) du LAAS-CNRS (Laboratoire d'Analyse et d'Architecture des Systèmes).

Ma principale thématique de recherche concerne les codes identifiants dans les graphes. Ce sujet est rattaché à la théorie des graphes ainsi qu'à la théorie des codes. Il a été introduit à la fin des années 1990 pour modéliser des problèmes de détection de défaillances dans les réseaux. Il fait l'objet depuis d'un grand nombre de travaux de recherche, et intéresse un nombre croissant de chercheurs.

Après des études de mathématiques et d'informatique à l'université Joseph Fourier puis à Grenoble INP- ENSIMAG, j'entreprends un travail de

thèse en 2002, sous la direction de Sylvain Gravier. Soutenue en juin 2005 à l'université Joseph Fourier (Grenoble), ma thèse porte sur les codes identifiants dans les graphes. Après une année en tant qu'attaché temporaire d'enseignement et de recherche à Grenoble INP – ENSIMAG entre 2005 et 2006, je suis recruté à Grenoble INP – Génie Industriel (anciennement École Nationale Supérieure de Génie Industriel) sur un poste de maître de conférences en 2006. Ayant alors intégré le laboratoire G-SCOP, j'y mène jusqu'à septembre 2010 des recherches au sein de l'équipe ROSP (Recherche Opérationnelle pour les Systèmes de Production). Je demande alors ma mutation à l'université Toulouse 1 Capitole et intègre le LAAS à la rentrée 2010.

Je suis actuellement membre de la structure fédérative de recherche « maths à modeler ». Cette équipe — composée de didacticiens, de mathématiciens discrets, et de chercheurs en sciences de l'éducation — mène des recherches transversales visant à la construction de situations-recherche pour la classe. Ces situations-recherche sont des problèmes de recherche présentés sous la forme de casse-têtes ou de jeux combinatoires, qui permettent l'appropriation du problème et la prise en charge de sa résolution. Elles constituent une initiation au raisonnement et à la démarche de recherche qui favorise le développement de l'esprit critique, et sont utilisées tant dans des contextes scolaires ou universitaires que dans des contextes de vulgarisation scientifique (typiquement des manifestations comme la « fête de la science »). J'ai été également, entre 2009 et 2011, membre du projet ANR IDEA (« Identifying coDes in Evolving grAphs »), porté par le Laboratoire Bordelais de Recherche en Informatique (LaBRI). Ce projet concernait les versions dynamiques et adaptatives des codes identifiants.

J'ai eu le plaisir de co-encadrer avec Nadia Brauner (G-SCOP) le travail de thèse de Julien Darlay, qui porte sur l'analyse combinatoire de données (thèse soutenue à Grenoble en décembre 2011). Je co-encadre par ailleurs depuis septembre 2008 avec Myriam Preissmann (G-SCOP) la thèse de Marwane Bouznif, qui porte sur l'étude de problèmes combinatoires dans les grilles.

J'ai participé à deux accords de coopération scientifique internationaux (avec l'Algérie et la Hongrie), et mes collaborateurs internationaux sont Yael Ben-Haim (Israël), Mostafa Blidia (Algérie), Mustapha Chellali (Algérie), Alan Frieze (États-Unis), Tero Laihonen (Finlande), Ryan Martin (États-Unis), Miklós Ruszinkó (Hongrie), Ahmed Semri (Algérie), et Clifford Smyth (États-Unis). J'ai effectué plusieurs séjours de recherche à l'étranger, notamment en Israël, en Finlande, et en Hongrie où j'ai passé huit mois au sein du laboratoire d'automatique et d'informatique de l'Académie des Sciences de Hongrie, dans l'équipe « Structures discrètes »,

avec Miklós Ruszinkó et András Gyárfás.

Si depuis ma thèse j'ai quelque peu diversifié mes activités de recherche, notamment *via* mes co-encadrements de thèse et la fréquentation de laboratoires de recherches ouverts sur d'autres thématiques de mathématiques discrètes, la grande majorité de mes publications concerne les codes identifiants dans les graphes. Le présent manuscrit s'attachera donc à décrire l'ensemble de mes travaux sur cette thématique depuis le début de ma vie de chercheur en 2002.

## 1.2 Enseignement

Ayant été moniteur pendant ma thèse, puis ATER pendant un an, et enfin maître de conférences depuis mon premier recrutement en 2006, j'enseigne dans le supérieur depuis une dizaine d'années.

À l'IUT de Rodez, je suis actuellement responsable et chargé de cours, travaux dirigés, et travaux pratiques pour les matières suivantes :

- Systèmes de gestion de bases de données (niveau L1)
- Conception des systèmes d'information (niveau L1)
- Bases de la planification de projet (niveau L1)
- Structure et organisation des entreprises (niveau L1)
- Typologie des produits et des procédés (niveau L1)
- Gestion physique des stocks (niveau L2)
- Aménagement du poste de travail (niveau L2)
- Informatique avancée (niveau L3)

Lorsque j'étais à Grenoble, je me suis impliqué dans les cours suivants :

- Probabilités et statistiques (niveau L3)
- Recherche opérationnelle (niveau M1)
- Management de la production et des services (niveau M1)
- Outils formels – complexité et cryptographie (niveau M1)
- Ordonnancement de la production (niveau M1)
- Théorie des jeux et décision (niveau M2)
- Graphes et structures discrètes (niveau M2)

J'ai également été tuteur de nombre de stages, projets, et études de terrain. La liste détaillée de mes activités d'enseignement se trouve dans mon CV, qui est fourni en Annexe A.

### 1.3 Responsabilités collectives

Entre 2008 et 2010, j'ai été responsable des admis sur titres à Grenoble INP – Génie Industriel. Les admis sur titres sont des étudiants qui intègrent l'école après un cursus autre que celui d'une classe préparatoire. Ils sont généralement titulaires d'un DUT (diplôme universitaire de technologie) ou d'une licence générale. En tant que responsable, j'étais chargé du recrutement et de l'accueil de ces étudiants.

La sélection des étudiants se déroulait selon le processus suivant. Tout d'abord, il fallait choisir parmi les quelques 150 dossiers reçus une quarantaine de « bons » dossiers d'étudiants que nous allions convoquer à un entretien individuel. Ces entretiens se déroulaient à Grenoble ou à Paris, en même temps que les entretiens de recrutement pour les étudiants des classes préparatoires. À l'issue de ces entretiens, une dizaine d'étudiants étaient alors reçus à l'école.

Le deuxième volet de mon action en tant que responsable des admis sur titres concernait alors l'accueil de ces étudiants. Afin de faciliter leur intégration à l'école, j'ai tout d'abord contribué à la mise en place d'une semaine de pré-rentrée pour les admis sur titres. Pendant cette semaine, les étudiants pouvaient suivre des séances de remise à niveau sur certaines thématiques, notamment en mathématiques et en mécanique. En effet, nous avons observé que, si les admis sur titre réussissaient tout aussi bien que les autres étudiants, ils avaient cependant de grandes difficultés au semestre 1. Ce semestre étant de loin le plus théorique, avec notamment de gros modules de mathématiques appliquées et de mécanique, cette semaine de pré-rentrée avait pour but de les préparer au « choc » et de leur transmettre les prérequis pour aborder ces cours. J'ai également contribué à instaurer un système de tutorat. Ce système consistait à mettre à disposition des admis sur titre, une fois par semaine sur un créneau bloqué pendant tout le premier semestre, un doctorant en mathématiques-informatique et un doctorant en mécanique, qui avaient pour mission de répondre aux questions des étudiants sur ces thématiques. L'objectif de ce tutorat était de soutenir ces étudiants tout au long du semestre, et de leur fournir un interlocuteur qui serait à la fois « proche » d'eux et compétent sur les matières en question.

Au laboratoire G-SCOP, j'ai été responsable du séminaire de travail de l'équipe ROSP (le SÉTRO), entre 2007 et 2010. Ce séminaire a accueilli une vingtaine d'exposés, réalisés pour moitié par des membres du laboratoire, et pour moitié par des experts extérieurs. Les thématiques privilégiées étaient celles en lien avec la modélisation et l'optimisation des systèmes

de production industriels, et en particulier la recherche opérationnelle, au sens large (programmation mathématique, complexité, théorie des graphes, automates...). Les chercheurs étrangers ayant exposé dans le cadre de ce séminaire sont Valery Gordon (Académie des Science du Bélarus, décédé), Fatiha Kacher (université de Montréal, Canada), Jihane Alami Chentoufi (université Ibn Tofail, Maroc), et Jacek Błażewicz (université de Poznań, Pologne).

J'ai également participé à l'organisation de congrès et workshops (voir la liste dans mon CV en Annexe A), ainsi qu'à un comité de sélection à Grenoble en 2008.

## 1.4 Recherche

Je présente brièvement en Section 1.4.1 l'ensemble de mes thématiques de recherche, mes travaux sur les codes identifiants étant discutés de façon plus approfondie dans le Chapitre 2 de ce document. Mes encadrements de travaux de recherche sont exposés en Section 1.4.2. Les Sections 1.4.3 et 1.4.4 présentent, respectivement, mes activités d'expertise de travaux de recherche ainsi que mes collaborations internationales. La liste complète de mes publications est quant à elle donnée en Section 1.5.

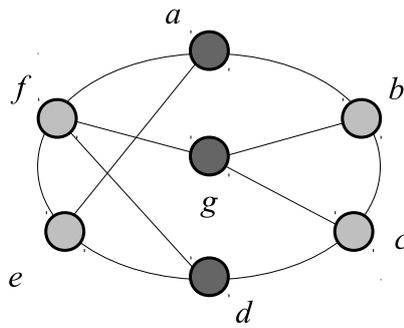
### 1.4.1 Thématiques de recherche

#### 1.4.1.1 Codes identifiants dans les graphes

Les codes identifiants ont été introduits en 1998 par Karpovsky *et al* dans un article des *IEEE Transactions on Information Theory* [120]. À l'origine introduits dans le cadre de la théorie des codes, ils modélisent un problème de détection de défaillance dans les réseaux. La communauté « graphes » s'étant depuis emparée de la thématique, ils sont aujourd'hui un thème à la frontière entre la théorie des graphes et la théorie des codes.

Considérons un graphe  $G = (V, E)$ , et soit  $C$  un sous-ensemble de ses sommets :  $C \subseteq V$ . On note le voisinage fermé d'un sommet  $v \in V$  par  $N[v]$  (le voisinage fermé de  $v$  est l'ensemble des voisins de  $v$  plus le sommet  $v$  lui-même). On dit que  $C$  est un *code identifiant* de  $G$  si, d'une part, pour tout sommet  $v \in V$  on a  $C \cap N[v] \neq \emptyset$  (contrainte de couverture), et, d'autre part, pour tout couple de sommets  $u, v \in V$ ,  $u \neq v$ , on a  $C \cap N[u] \neq C \cap N[v]$  (contrainte de séparation). Autrement dit, un sommet de  $G$  est identifié de façon unique par l'intersection de son voisinage fermé avec le code, d'où

la terminologie code *identifiant* (voir un exemple Figure 1.1 ci-dessous). On peut voir un code identifiant comme un code couvrant (ou ensemble dominant) vérifiant une propriété additionnelle de séparation, ce qui explique pourquoi la communauté « graphes » s'est rapidement intéressée à ces codes.



**FIGURE 1.1 :** Un exemple de code identifiant. Les sommets foncés sont les sommets du code. Il est facile de vérifier que tout sommet est identifié de façon unique par l'intersection de son voisinage fermé avec le code. Par exemple, le sommet  $b$  est l'unique sommet voisin à la fois de  $a$  et de  $g$ . De même, le seul sommet à n'être voisin que de  $g$  est le sommet  $g$  lui-même.

Il est facile de vérifier que  $C = V$  est toujours un code identifiant, pourvu toutefois que le graphe soit sans jumeaux (deux sommets  $u, v$ ,  $u \neq v$ , sont dits jumeaux si  $N[u] = N[v]$ ). En terme d'optimisation combinatoire, l'enjeu est donc, étant donné un graphe  $G$ , de trouver un code identifiant de  $G$  de cardinalité minimum. Par analogie avec la domination, la cardinalité minimum d'un tel code, lorsqu'il existe, est notée  $\gamma^{\text{ID}}(G)$ . Les définitions et généralités sur ces codes sont données en Section 2.1 du Chapitre 2.

Le problème du calcul de  $\gamma^{\text{ID}}(G)$  est NP-difficile dans le cas général [49]. Ceci peut sembler naturel si l'on voit un code identifiant comme un code couvrant, les problèmes de couverture de graphes ou d'hypergraphes par des sommets faisant partie du catalogue standard de problèmes NP-complets (voir notamment les problèmes GT1, GT2, GT61 ; ainsi que les problèmes SP5, SP6, et SP8 dans le Garey-Johnson [88]).

La discussion de l'ensemble de mes résultats sur cette thématique constitue le Chapitre 2 de ce document.

Une partie importante de mes travaux de recherche concerne l'étude de ces codes dans des classes de graphes particulières (cycles, puissances de cycles, grilles, hypercubes, arbres, etc.). Ces travaux, décrits en Section 2.2, s'insèrent dans une littérature relativement fournie sur le sujet. Je me suis également intéressé à des questions algorithmiques sur ce

problème (Section 2.3). Une autre partie de mes travaux concerne des questions extrémales, en particulier l'étude des bornes inférieures et supérieures sur la cardinalité d'un tel code en fonction du nombre de sommets du graphe, ainsi que la caractérisation des graphes atteignant ces bornes (Section 2.4). Finalement, j'ai introduit une nouvelle variante de ces codes, appelée *codes identifiants adaptatifs*, qui concerne des procédés d'identification dynamiques des sommets (Section 2.5).

#### 1.4.1.2 Ordonnancement de la production

Recruté au laboratoire G-SCOP puis au LAAS-CNRS, j'ai eu l'opportunité de diversifier mes thématiques de recherche, autour notamment de problématiques plus industrielles telles que l'ordonnancement de la production. Je m'intéresse actuellement à des problèmes sur une machine, et ce autour de deux problématiques.

La première concerne l'ordonnancement de tâches lorsque la machine doit subir une période de maintenance à une date fixée et connue à l'avance.

On suppose disposer d'une machine, disponible sur l'intervalle de temps  $[0, T] \cup [T + D, +\infty[$ , sur laquelle on doit exécuter  $n$  tâches de durées quelconques  $p_1, \dots, p_n$ . La machine ne peut exécuter aucune tâche pendant l'intervalle de temps  $]T, T + D[$ , et la préemption n'est pas autorisée (c'est-à-dire qu'une tâche de durée  $p$  doit être impérativement exécutée par la machine sur un intervalle de temps  $[t, t + p]$ ).

En terme de géométrie discrète, ce problème s'apparente à l'empilement d'intervalles de longueurs  $p_1, \dots, p_n$  dans l'espace  $[0, T] \cup [T + D, +\infty[$  (problème de *packing*). Dans ce contexte, le problème de la minimisation de la date de fin de la dernière tâche est clairement NP-complet, puisque directement réductible au problème de partition (problème SP12 dans le Garey-Johnson [88]).

Un autre critère important en ordonnancement est le *flot*, défini comme la somme des dates de fin de toutes les tâches. Le problème de la minimisation du flot est lui aussi NP-complet [1]. Nous donnons dans un manuscrit actuellement soumis [150] un schéma d'approximation polynomial pour ce problème, qui généralise [1, 162] et améliore [106]. Nous démontrons de plus que notre résultat est le meilleur possible en donnant une famille de problèmes atteignant asymptotiquement la borne sur la garantie calculée. Enfin, des tests expérimentaux mettent en évidence le fait que, en pratique, ce schéma d'approximation est très performant, et fournit des ordonnancements qui sont en moyenne d'une qualité bien supérieure à la garantie de l'algorithme, et très proches de l'optimum.

Ce travail est le fruit d'une collaboration avec Ariel Waserhole (docteur au laboratoire G-SCOP) et Jérémie Thiery (ingénieur chez DIAGMA Supply Chain Services) alors qu'ils étaient tous deux étudiants à Grenoble INP – ENSIMAG. Une version préliminaire de ces travaux a été présentée au congrès de la ROADEF en 2008 [170]. La version complète a été présentée dans un congrès international en avril 2012 [151].

La seconde problématique à laquelle je me suis récemment intéressé est celle de l'ordonnancement de tâches sujettes à des phénomènes d'usure ou d'apprentissage. On parle d'usure ou d'apprentissage lorsque la durée d'une tâche dépend de sa position (temps et/ou rang) dans l'ordonnancement. Ces phénomènes, connus dans le milieu de la gestion industrielle depuis les années 1930 [175, 178], n'ont été introduits en ordonnancement que récemment [3, 27] (voir un exemple ci-dessous en Figure 1.2).

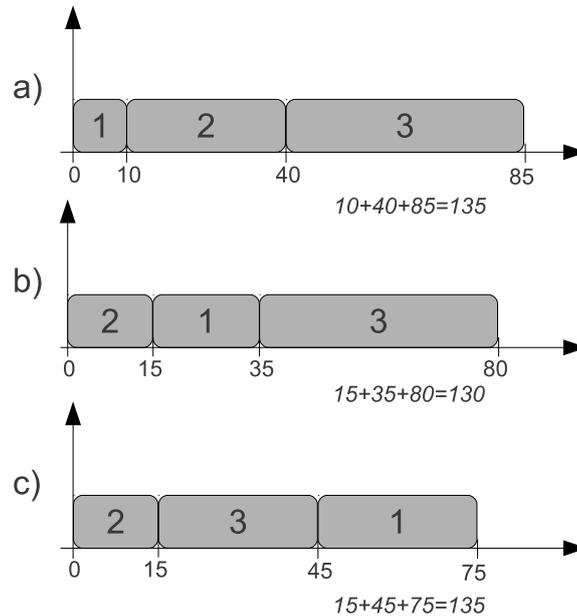
Dans un article actuellement soumis [148], nous donnons un cadre théorique permettant de généraliser et d'unifier de nombreux résultats de la littérature concernant les problèmes sur une machine où le temps opératoire dépend du rang. Notre approche permet en particulier d'obtenir des algorithmes polynomiaux et efficaces pour l'ordonnancement de telles tâches sur une machine. Les critères d'optimisation pour lesquels nos résultats s'appliquent contiennent le flot et la date de fin de la dernière tâche.

Nous démontrons qu'une large classe de problèmes d'ordonnancement de ce type se réduisent à un problème de couplage de poids maximum dans un graphe biparti (problème dit d'affectation, solvable en temps cubique). Pour une importante sous-classe de ces problèmes, nous montrons que ce problème de couplage peut de plus être résolu en temps  $O(n \log n)$ , avec  $n$  le nombre de tâches du problème. Nous donnons de plus une caractérisation de cette sous-classe de problème, en montrant un lien avec les matrices dites de Monge [39].

Ce travail est poursuivi en collaboration avec Gerd Finke (G-SCOP) et Vincent Jost (LIX), et a fait l'objet d'une communication au congrès de la ROADEF en 2010 [147]. Nos résultats ont été présentés dans un congrès international en avril 2012 [149].

### 1.4.1.3 Autres thématiques de recherche

De façon ponctuelle, j'ai également travaillé sur d'autres thématiques, en particulier relevant de la théorie des jeux. Dans un travail en collaboration avec Nadia Brauner et Roland Grappe, nous avons donné une modélisation d'un problème de surveillance (qui nous avait à l'origine était



**FIGURE 1.2 :** Dans cet exemple, on considère trois tâches 1, 2, 3 de durées de base  $p_1 = 10, p_2 = p_3 = 15$ , et on suppose que la durée réelle d'une tâche  $i$  est égale à  $r \times p_i$  si celle-ci est exécutée au rang  $r$  dans l'ordonnancement. Il s'agit donc d'un problème d'usure de la machine (plus le temps passe, plus la machine est lente). Dans le cas a) ci-dessus, la durée réelle de la tâche 1 est donc sa durée de base car elle est réalisée en premier. La durée réelle de la tâche 2 est  $2 \times 15$ , celle de la tâche 3 est  $3 \times 15$ . Dans ce cas le flot est égal à  $10 + 40 + 85 = 135$ . Dans le cas b), on a permuté le rang des tâches 1 et 2. Si la durée réelle de la tâche 2 a diminué, celle de la tâche 1 a elle augmenté. Mais les gains compensent les pertes, car le flot a diminué de 5. La troisième possibilité est donnée dans le cas c). En ce cas, la date de fin de la dernière tâche est à son minimum : 75. Cependant le flot est quant à lui égal à 135. Cet exemple nous permet de plus d'illustrer le fait que, en toute généralité, minimiser la date de fin de la dernière tâche n'est pas équivalent à minimiser le flot.

proposé par la Marine Nationale) par un jeu à deux joueurs. Ce travail a été présenté au congrès de la ROADEF en 2008 [92], et a fait l'objet par la suite de deux travaux d'étudiants de l'ENSIMAG [18, 80]. Nous disposons aujourd'hui de deux algorithmes de simulation du jeu, l'un utilisant une approche « minimax » par la programmation linéaire, et l'autre utilisant un algorithme d'inférence qui « étudie » et « apprend » comment l'autre joueur joue afin de pouvoir mieux le contrer.

J'ai proposé de nombreux sujets relatifs aux jeux aux étudiants de l'ENSIMAG, et l'un de ces travaux a d'ailleurs donné lieu à l'implémentation

d'un générateur de cartes aléatoires pour le jeu FreeDroidRPG [81].

En collaboration avec Éric Duchêne (laboratoire GAMA, Lyon) et Sylvain Gravier, j'ai également travaillé sur le Clobber [68], qui est un jeu combinatoire introduit récemment. Un de mes premiers travaux de recherche (effectué en fait avant ma thèse) concernait le problème d'exclusion des pentominos de Golomb. Ces travaux ont été publiés dans *Discrete Mathematics* en 2007 [97].

Dans le cadre de ma collaboration à la fédération de recherche « maths à modeler », je participe également à des activités de valorisation en didactique et enseignement des mathématiques, visant à donner une visibilité à notre approche dans la littérature du domaine. Ceci m'a amené à participer à la rédaction d'articles destinés à être publiés dans des revues ou des conférences de didactique des mathématiques [40, 41].

Les thématiques de recherche spécifiques de mes deux étudiants en thèse sont détaillées ci-après.

## 1.4.2 Direction de travaux de recherche

### 1.4.2.1 Thèse de Julien Darlay – analyse combinatoire de données

Dans le cadre du travail de thèse de Julien Darlay [63], j'ai également travaillé sur l'application de méthodes de la recherche opérationnelle à l'analyse combinatoire de données, en vue notamment de l'aide à la décision dans le milieu médical.

L'analyse combinatoire de données se distingue de l'analyse de données « classique » en ce sens qu'elle vise non pas à quantifier la corrélation entre un paramètre et un autre (par exemple un symptôme et une pathologie), mais plutôt à donner une « explication » de la pathologie sous la forme d'une formule logique combinant différents symptômes. L'idée est donc de fournir une « logique » de la pathologie aux médecins, elle a donc à la fois une fonction prédictive et une fonction explicative des phénomènes mis en œuvre. Ces travaux se rattachent à des problèmes combinatoires de couverture d'ensembles (voir Figure 1.3 ci-dessous), et sont dans le prolongement des travaux initiés par Peter Hammer et son équipe à RUTGERS [32].

À ce jour nous avons un manuscrit soumis avec Nadia Brauner (G-SCOP, Grenoble) et Julien Darlay [38], concernant un problème de partition de graphes en sous-graphes denses. Nos résultats sur ce problème concernent sa complexité et le développement d'un algorithme polynomial pour les arbres. L'algorithme pour les arbres utilise les résultats classiques sur les couplages maximum dans les graphes bipartis. Le contenu en a été présenté au congrès de la ROADEF en 2010 [64]. Julien est ingénieur de

$$A = \begin{pmatrix} 0 & * & * & 1 & 0 & * & 0 & 0 & 0 & 3 & 1 \\ 1 & 1 & * & 1 & 5 & 1 & * & 2 & 0 & 2 & 1 \\ 4 & 2 & 1 & 3 & 1 & 1 & 1 & 1 & 1 & * & 0 \\ 0 & 1 & 1 & * & 0 & * & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 2 & 3 & 5 & * & * & 1 & 2 & 1 & 0 \\ 3 & * & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & * & * & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 2 & 5 & 1 & * & 0 & 4 & 2 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = b.$$

**FIGURE 1.3 :** Dans cet exemple, les lignes représentent des patients, les colonnes de  $A$  représentent des caractéristiques des patients (taille, poids, pression artérielle, sexe... – le  $*$  dénote l'absence d'information). Le vecteur  $b$  distingue les patients « positifs » des patients « négatifs » vis-à-vis d'une certaine pathologie. L'enjeu est de trouver des patterns, c'est-à-dire des combinaisons de conditions logiques sur les valeurs des caractéristiques (formules booléennes), distinguant les patients positifs des patients négatifs. L'aspect optimisation du problème réside dans la sélection de « meilleurs » ensembles de patterns, vis-à-vis d'un ou plusieurs critères donnés (typiquement on considère le nombre de patterns et la « complexité » des patterns, par exemple le nombre de termes de la formule booléenne).

recherche au Bouygues e-lab depuis novembre 2011 .

#### 1.4.2.2 Thèse de Marwane Bouznif – combinatoire dans les grilles

Marwane Bouznif, quant à lui, réalise un travail de thèse (depuis septembre 2008) sur l'étude de problèmes combinatoires dans les grilles et dans les structures généralisant les grilles. Dans cette thèse, nous nous efforçons de dégager des résultats généraux permettant de résoudre de façon optimale une large palette de problèmes combinatoires dans les grilles et, plus généralement, dans des graphes qui sont structurés en grille. L'idée est de construire un cadre théorique permettant de donner des résultats génériques, en utilisant les outils de la programmation dynamique et de l'algèbre  $(\max, +)$ . Ces résultats génériques permettraient d'unifier de nombreux travaux antérieurs, tels par exemple [119, 126, 127, 139]. Ils nous donnent en particulier des temps d'exécution spectaculaires dès lors qu'une des dimensions des grilles est fixée, et permettent de dériver des résultats de périodicité ou de pseudo-périodicité des solutions optimales.

Marwane a présenté son travail au congrès de la ROADEF en 2010 [37], et notre premier manuscrit est paru à *Discrete Mathematics* [35]. Ce ma-

nuscrit, écrit avec Marwane Bouznif et Myriam Preissmann (G-SCOP, Grenoble), concerne la partie « décision » de notre approche, et un deuxième manuscrit concernant la partie « optimisation » est actuellement soumis [36].

### 1.4.2.3 Mémoires de M2

Mes encadrements de stages de Master 2 Recherche m'ont également donné l'occasion de découvrir d'autres thématiques de recherche. En 2007, j'ai encadré le stage d'Onur Çelebi [42], portant sur un problème de localisation de sites de concentration proposé par France Télécom R & D (encadrant industriel : Olivier Klopfenstein). Ce travail consistait en l'étude d'une version stochastique du problème, où l'on a supposé que les coûts et les demandes n'étaient pas déterministes mais susceptibles de fluctuations. Le mémoire a porté sur l'adaptation de méthodes d'optimisation stochastique de la littérature à ce problème. Onur est actuellement développeur au sein de la société Murex à Paris, spécialisée dans le développement de logiciels et d'outils d'aide à la décision dans le domaine financier.

En 2009, j'ai encadré le stage de Christophe-Marie Duquesne [69], portant sur le *fleet assignment problem*, proposé par Amadeus (encadrant industriel : Semi Gabteni), et qui faisait suite au Challenge 2009 de la ROADEF. Le travail de Christophe-Marie a consisté en l'extension des travaux de Barnhart *et al* [17] sur le sujet. Christophe-Marie poursuit actuellement ses travaux de recherche dans le cadre d'une thèse de doctorat, co-encadrée par Denis Naddef et Olivier Briant, en partenariat avec la société Amadeus.

### 1.4.3 Expertise

On fait régulièrement appel à mes services pour arbitrer des manuscrits, qui portent en grande majorité sur les codes identifiants. Les journaux faisant appel à moi pour ce travail de relecture sont les suivants :

- *Discrete Mathematics*
- *Discrete Applied Mathematics*
- *Parallel Computing*
- *European Journal of Combinatorics*
- *Electronic Journal of Combinatorics*
- *Journal of Graph Theory*
- *Discrete Mathematics and Theoretical Computer Science*
- 4OR
- *IIE Transactions*

Sans tenir un compte précis de ces activités, je me donne comme objectif personnel de relire au moins autant de manuscrits que je n'en soumetts. À titre indicatif, j'ai arbitré cinq manuscrits au premier semestre de l'année 2011-2012.

En 2011, j'ai été sollicité pour être rapporteur de la thèse de Ville Junnila [116]. En Finlande, il n'est pas requis que le rapporteur ait un statut ou un diplôme particulier (si ce n'est une thèse) pour être lui-même rapporteur d'une thèse. Le système finlandais demande à deux rapporteurs d'évaluer le travail de thèse et de se prononcer si le candidat peut ou non aller jusqu'à la soutenance. La soutenance est alors réellement une « défense » des travaux, puisque le candidat doit alors répondre aux questions d'un troisième expert, l'*opposant*. Cet opposant est le seul membre du jury, il doit nécessairement être une personne distincte des deux rapporteurs. Il a une heure de temps pour poser ses questions et interagir avec le candidat.

#### 1.4.4 Collaborations internationales

Pendant ma thèse, j'ai eu l'opportunité de passer huit mois à Budapest, au laboratoire d'automatique et d'informatique de l'Académie des Sciences de Hongrie (SZTAKI), au sein l'équipe « Structures discrètes », avec Miklós Ruzinkó et András Gyárfás. Ce séjour a eu lieu en 2004, grâce à un financement de la région Rhône-Alpes qui s'appelait à l'époque « EURODOC ». Ce séjour a donné lieu à une collaboration avec Miklós Ruzinkó, au sujet d'une approche probabiliste des codes identifiants, publiée dans *Discrete Mathematics* [83]. Lors de ce séjour, j'ai également rencontré Ryan Martin (université d'Iowa), qui est l'un des coauteurs de ce travail.

Suite à ce séjour de recherche, j'ai contribué à monter un accord de coopération entre le laboratoire Leibniz (Grenoble) où j'effectuais ma thèse et le SZTAKI. Cet accord « Balaton », était alors appelé Plan d'Action Intégrée; il correspondrait aujourd'hui à un programme Hubert Curien. Cet accord a duré deux années (2005 et 2006). Il a contribué à maintenir des liens avec la Hongrie, et notamment à d'autres membres des deux laboratoires d'effectuer des séjours de recherche chez l'autre partenaire.

J'ai passé une dizaine de jours à Tel Aviv en 2006, au sein du *Department of Electrical Engineering-Systems*, où j'ai travaillé avec Simon Litsyn et Yale Ben-Haim. Ce séjour faisait suite à un travail préliminaire fait avec Yal Ben-Haim, Antoine Lobstein (Télécom ParisTech) et Sylvain Gravier sur les codes identifiants adaptatifs. Il nous a permis de développer nos résultats. Le fruit de ce travail a été publié dans deux revues [20, 21]. Cette

thématique, que je proposais dans ma thèse, est maintenant introduite dans la littérature, et a, par la suite, fait l'objet du projet ANR IDEA (2009–2011).

Cette même année, j'ai été invité par Tero Laihonon à passer deux semaines à l'université de Turku en Finlande. Nous avons travaillé sur les codes identifiants des ensembles de sommets, et cette collaboration a débouché sur un manuscrit publié dans l'*Australasian Journal of Combinatorics* [134].

J'ai rencontré Ahmed Semri (université Houari Boumédiène, Algérie) lors de ses nombreux séjours à Grenoble, qui avaient lieu dans le cadre de l'accord de coopération franco-algérien du Comité Mixte d'Évaluation et de Prospective dont nous étions membres tous les deux. Nos collaborations régulières ont donné lieu à deux publications. Notre travail sur les cycles a été publié dans l'*European Journal of Combinatorics* en 2006 [98]. En 2008, nous avons de plus publié dans l'*Electronic Journal of Combinatorics* une note sur les codes identifiants dans les produits cartésiens de cliques [100]. Ces travaux ont été obtenus en collaboration avec Sylvain Gravier.

J'ai par ailleurs rencontré Matjaž Kovše (postdoctorant au LaBRI, Bordeaux) à l'occasion d'une de ses visites à Grenoble, dans le cadre du programme d'action intégrée Proteus avec la Slovénie. Nous avons travaillé sur une structure qui est une spécialité slovène, les graphes de Sierpiński. Ce travail, en collaboration avec Sylvain Gravier, Michel Mollard, et Aline Parreau (tous à l'Institut Fourier, Grenoble), a donné lieu à un article paru dans *Designs, Codes, and Cryptography* [94].

## 1.5 Liste de publications

### 1.5.1 Publications parues ou à paraître dans des revues internationales à comité de lecture

- [R20] N. Brauner, J. Darlay, J. Moncel, *Dense & sparse graph partition*, à paraître dans *Discrete Applied Mathematics*.
- [R19] M. Bouznif, J. Moncel, M. Preissmann, *Generic algorithms for some decision problems on fasciagraphs and rotagraphs*, à paraître dans *Discrete Mathematics*.
- [R18] S. Gravier, M. Kovše, M. Mollard, J. Moncel, A. Parreau, *New results on variants of covering codes in Sierpiński graphs*, à paraître dans *Designs, Codes and Cryptography*.

- [R17] Y. Ben-Haim, S. Gravier, A. Lobstein, J. Moncel, *Adaptive identification in torii in the king lattice*, *Electronic Journal of Combinatorics* **18(1)** (2011), P116.
- [R16] E. Duchêne, S. Gravier, J. Moncel, *New results about Solitaire Clobber*, *RAIRO—Operations Research*, **43** (2009), 463–482.
- [R15] Y. Ben-Haim, S. Gravier, A. Lobstein, J. Moncel, *Adaptive identification in graphs*, *Journal of Combinatorial Theory Series A* **115(7)** (2008), 1114–1126.
- [R14] S. Gravier, J. Moncel, A. Semri, *Identifying codes of cartesian product of two cliques of the same size*, *Electronic Journal of Combinatorics* **15(1)** (2008), N4.
- [R13] S. Gravier, R. Klasing, J. Moncel, *Hardness results and approximation algorithms for identifying codes and locating-dominating codes in graphs*, *Algorithmic Operations Research* **3(1)** (2008), 43–50.
- [R12] T. Laihonen, J. Moncel, *On graphs admitting codes identifying sets of vertices*, *Australasian Journal of Combinatorics* **41** (2008), 81–91.
- [R11] M. Blidia, M. Chellali, F. Maffray, J. Moncel, A. Semri, *Locating-dominating and identifying codes in trees*, *Australasian Journal of Combinatorics* **39** (2007), 219–232.
- [R10] A. Frieze, R. Martin, J. Moncel, M. Ruszinkó, C. Smyth, *Codes identifying sets of vertices in random networks*, *Discrete Mathematics* **307(9-10)** (2007), 1094–1107.
- [R9] S. Gravier, J. Moncel, *On graphs having a  $V \setminus \{x\}$  set as an identifying code*, *Discrete Mathematics* **307(3-5)** (2007), 432–434.
- [R8] S. Gravier, J. Moncel, C. Payan, *A generalization of the pentomino exclusion problem : Dislocation of graphs*, *Discrete Mathematics* **307(3-5)** (2007), 435–444.
- [R7] J. Moncel, *Constructing codes identifying sets of vertices*, *Designs, Codes and Cryptography* **41(1)** (2006), 23–31.
- [R6] J. Moncel, *On graphs on  $n$  vertices having an identifying code of cardinality  $\lceil \log_2(n+1) \rceil$* , *Discrete Applied Mathematics* **154(14)** (2006), 2032–2039.
- [R5] I. Charon, S. Gravier, O. Hudry, A. Lobstein, M. Mollard, J. Moncel, *A linear algorithm for minimum 1-identifying codes in oriented trees*, *Discrete Applied Mathematics*, **154(8)** (2006), 1246–1253.
- [R4] J. Moncel, *Monotonicity of the minimum cardinality of an identifying code in the hypercube*, *Discrete Applied Mathematics* **154(6)** (2006), 898–899.

- [R3] S. Gravier, J. Moncel, A. Semri, *Identifying codes of cycles*, European Journal of Combinatorics **27(5)** (2006), 767–776.
- [R2] S. Gravier, J. Moncel, *Construction of codes identifying sets of vertices*, Electronic Journal of Combinatorics **12(1)** (2005), R13.
- [R1] M. Daniel, S. Gravier, J. Moncel, *Identifying codes in some subgraphs of the square lattice*, Theoretical Computer Science **319(1-3)** (2004), 411–421.

### 1.5.2 Participation à des conférences internationales

- [C9] J. Moncel, G. Finke, V. Jost, *Single-machine scheduling problems with position-dependent processing times*, 13th International Conference on Project Management and Scheduling (1–4 avril 2012, Louvain – Belgique).
- [C8] J. Moncel, J. Thiery, A. Waserhole, *Computational performances of a simple interchange heuristic for a scheduling problem with an availability constraint*, 13th International Conference on Project Management and Scheduling (1–4 avril 2012, Louvain – Belgique).
- [C7] J. Darlay, N. Brauner, J. Moncel, *Dense and Sparse Graph partition*, 24rd EURO (European Conference On Operational Research, Lisbonne 10–14 juillet 2010).
- [C6] M. Bouznif, J. Moncel, M. Preissmann, *Algorithms for combinatorial problems on fasciagraphs and rotagraphs*, 8th French Combinatorial Conference (Paris, 28 juin – 2 juillet 2010).
- [C5] L. Cartier, J. Moncel, *Learner’s conceptions in different situations around Königsberg’s bridges problem*, ICME 11 (International Colloquium on Mathematical Education, Monterrey 6–13 juillet 2008), Topic Study Group 15.
- [C4] S. Gravier, J. Moncel, A. Semri, *Codes identifiants : une approche combinatoire d’un problème de détection de défaillances dans les réseaux*, COSI’06 (International Colloquium on Optimisation and Information Systems, Alger 11–13 juin 2006).
- [C3] J. Moncel, *Codes Identifying Vertices in Graphs*, ICGT’05 (7th International Colloquium on Graph Theory, Hyères 12–16 septembre 2005).
- [C2] A. Frieze, R. Martin, J. Moncel, M. Ruszinkó, C. Smyth, *Identifying codes in random networks*, ISIT 2005 (2005 IEEE International Symposium on Information Theory 4–9 septembre 2005).

- [C1] S. Gravier, J. Moncel, Ch. Payan, *A generalization of the pentomino exclusion problem : Dislocation of graphs*, 5th International Slovenian Conference on Graph Theory (Bled 22–27 juin 2003).

### 1.5.3 Manuscrits soumis

- [S3] M. Bouznif, J. Moncel, M. Preissmann, *Optimization algorithms for rotagraphs and fasciagraphs*, soumis.
- [S2] J. Moncel, G. Finke, V. Jost, *A general result on single machine scheduling problems subject to learning effects or deteriorating jobs*, soumis à OR Letters, disponible en ligne à l'url <http://hal.archives-ouvertes.fr/hal-00647789/fr/>.
- [S1] J. Moncel, J. Thiery, A. Waserhole, *Computational performances of a simple interchange heuristic for a scheduling problem with an availability constraint*, en révision à Computers and Industrial Engineering. Une version préliminaire est disponible en ligne sur HAL : <http://hal.archives-ouvertes.fr/hal-00181166/fr/>



# Synthèse des travaux sur les codes identifiants

---

## Sommaire

---

<b>2.1 Définitions et généralités sur les codes identifiants</b>	<b>22</b>
2.1.1 Définitions	22
2.1.2 Trois exemples d'applications	26
2.1.3 Dynamique de la thématique de recherche	29
<b>2.2 Structures régulières</b>	<b>31</b>
2.2.1 Grilles et bandes	31
2.2.2 Cycles et puissances de cycles	40
2.2.3 Hypercubes	44
2.2.4 Produits de cliques	48
2.2.5 Graphes de Sierpiński	50
<b>2.3 Aspects algorithmiques</b>	<b>53</b>
2.3.1 État de l'art	53
2.3.2 Arbres orientés	57
2.3.3 Fasciagraphes et rotagraphes	60
2.3.4 Généralisation à d'autres problèmes de l'algorithme pour les fasciagraphes et les rotagraphes	63
2.3.5 Approximabilité du problème	66
<b>2.4 Questions structurelles</b>	<b>70</b>
2.4.1 Graphes extrémaux pour l'identification d'un seul sommet	71
2.4.2 Construction de graphes admettant de petits codes pour l'identification d'ensembles de sommets	79
2.4.3 Structures des graphes admettant un code	89
<b>2.5 Codes adaptatifs</b>	<b>96</b>
2.5.1 Vers des dispositifs d'identification dynamiques	96
2.5.2 Les codes identifiants vus comme un problème de recherche binaire	97
2.5.3 Définition et intérêt du concept	98
2.5.4 Identification adaptative dans les tores	100
2.5.5 Autres structures	105

---

Dans ce chapitre, le lecteur est supposé familier avec les concepts et la terminologie de base en théorie des graphes et en optimisation discrète en général. Sous cette hypothèse, cette partie se veut autosuffisante.

Pour la terminologie et les concepts de base en théorie des graphes, nous renvoyons le lecteur à l'ouvrage *Graph theory* de Reinhard Diestel [66]. La complexité des problèmes n'étant pas traitée dans cet ouvrage, nous renvoyons le lecteur à l'ouvrage *Computers and Intractability : A Guide to the Theory of NP-Completeness* de Garey et Johnson [88], qui sera souvent désigné par « le Garey-Johnson » dans ce document.

Ceci étant dit, quelques définitions seront toutefois rappelées au fil du manuscrit, ne serait-ce que pour fixer les notations.

## 2.1 Définitions et généralités sur les codes identifiants

### 2.1.1 Définitions

Un *graphe* (non orienté et sans boucle)  $G = (V, E)$  est constitué d'*arêtes* (éléments de  $E$ ) liant deux *sommets* distincts (éléments de  $V$ ). Lorsque deux sommets  $u, v$  de  $V$  sont reliés par une arête, ceux-ci sont dits *voisins*. En d'autres termes, un graphe est un couple  $(V, E)$  avec  $V$  un ensemble quelconque et  $E$  un sous-ensemble des parties à deux éléments de  $V$ . Un sommet  $u$  est alors voisin de  $v$  si  $\{u, v\} \in E$ . Afin de simplifier les notations, on désignera par la suite une arête  $\{u, v\}$  par  $uv$ . On considère en général dans ce chapitre des graphes finis, c'est-à-dire des graphes  $(V, E)$  tels que  $V$  (et par conséquent  $E$ ) est fini. On note en général  $n = |V|$  et  $m = |E|$ .

Un *sous-graphe* d'un graphe  $G = (V, E)$  est un graphe  $H = (V_H, E_H)$ , tel que  $V_H \subseteq V$  et  $E_H \subseteq E$ . Celui-ci est dit *induit* (par le sous-ensemble de sommets  $V_H$ ) si l'on a de plus

$$u \in V_H, v \in V_H, uv \in E \quad \Rightarrow \quad uv \in E_H.$$

Il existe plusieurs variantes à cette notion de graphe. L'une d'entre elles est celle de *graphe orienté*. Dans cette variante, on a  $(V, A)$  avec  $A$  un sous-ensemble de  $A \times A$ . Les éléments de  $A$  sont appelés des *arcs*. Les *hypergraphes* sont quant à eux définis par des couples  $(V, E)$  avec  $E$  un sous-ensemble de l'ensemble des parties non vides de  $V$ . Un élément de  $E$  est

alors appelé *hyperarête*. Dans ce document, le sens par défaut du terme « graphe » sera celui d'un graphe non orienté et fini.

L'ensemble de tous les voisins d'un sommet  $u$  est son *voisinage*, noté  $N(u)$ . Le *degré* d'un sommet  $u$  est égal à la cardinalité de  $N(u)$ , il est noté  $d(u)$ . Un graphe est dit *d-régulier* (avec  $d \geq 0$ ) si on a  $d(v) = d$  pour tout  $v$  de  $V$ . L'ensemble  $N[u]$  désigne le *voisinage fermé* de  $u$ , c'est-à-dire  $N(u) \cup \{u\}$ . Un sommet  $v \in N[u]$  est dit couvert (ou dominé) par le sommet  $u$ .

**Définition 1** (Dominant, code couvrant). *Un sous-ensemble de sommets  $D \subseteq V$  est un code couvrant (ou dominant) d'un graphe  $G = (V, E)$  si tout sommet de  $V$  est couvert par au moins un sommet de  $D$ .*

Les codes couvrants sont bien connus dans la littérature (voir par exemple [58, 104]), et font partie de la boîte à outil standard en optimisation discrète (voir le problème GT2 du Garey-Johnson [88]).

Deux sommets  $u$  et  $v$  sont dits séparés par  $w$  si celui-ci couvre l'un des sommets  $u$  ou  $v$  mais pas l'autre. En d'autres termes  $u$  et  $v$  sont séparés par  $w$  si  $w$  appartient à  $N[u] \Delta N[v]$ , la différence symétrique de  $N[u]$  et  $N[v]$  ( $A \Delta B$  étant défini comme  $(A \cup B) \setminus (A \cap B)$ ).

**Définition 2** (Code séparateur). *Un sous-ensemble  $S \subseteq V$  est un code séparateur d'un graphe  $G = (V, E)$  si deux sommets distincts de  $V$  sont séparés par au moins un sommet de  $S$ .*

**Définition 3** (Code identifiant). *Un sous-ensemble  $C \subseteq V$  qui est à la fois un code couvrant et un code séparateur d'un graphe  $G = (V, E)$  est appelé code identifiant de  $G$ .*

Les codes identifiants ont été introduits par Karpovsky, Chakrabary et Levitin en 1998 dans un article de *IEEE Transactions on Information Theory* [120].

Un graphe quelconque n'admet pas nécessairement de code identifiant. Il est facile de voir qu'un graphe admet un code identifiant si, et seulement si, celui-ci ne contient pas deux sommets distincts ayant le même voisinage fermé (de tels sommets sont appelés *sommets jumeaux*).

Le terme « identifiant » provient du fait que tout sommet du graphe est déterminé de façon unique par l'intersection de son voisinage fermé avec le code. En d'autres termes, la connaissance de  $N[u] \cap C$  permet de déterminer qui est  $u$ .

Ces codes admettent de nombreuses variantes dans la littérature, dont deux principales.

La première concerne l'identification à distance  $r$ , avec  $r \geq 1$  entier.

Dans ce cadre, un sommet  $u$   $r$ -couvre un sommet  $v$  s'il existe un chemin entre  $u$  et  $v$  empruntant au plus  $r$  arêtes (on rappelle qu'un *chemin* entre  $u$  et  $v$  est une séquence finie de sommets  $u = u_1, u_2, \dots, u_{k-1}, u_k = v$  telle que  $u_i u_{i+1}$  est une arête du graphe pour tout  $i = 1, \dots, k-1$ ).

Deux sommets  $u, v$  sont  $r$ -séparés par un sommet  $w$  si celui-ci  $r$ -couvre l'un des sommets  $u, v$  mais pas l'autre.

Un *code  $r$ -couvrant* est un sous-ensemble de sommets qui  $r$ -couvre tous les sommets, et un *code  $r$ -séparateur* est un sous-ensemble de sommets qui  $r$ -sépare toutes les paires de sommets distincts.

Un *code  $r$ -identifiant* est alors un sous-ensemble de sommets qui est à la fois un code  $r$ -couvrant et un code  $r$ -séparateur du graphe considéré.

Identifier des sommets à distance  $r$  revient donc à identifier des sommets au sens classique (c'est-à-dire à distance 1) dans la *fermeture  $r$ -transitive* du graphe considéré. En effet, soit  $G^r$  la fermeture  $r$ -transitive de  $G$ , c'est-à-dire le graphe dont les sommets sont ceux de  $G$ , et dont les arêtes sont définies par :

$u$  adjacent à  $v$  dans  $G^r$  si, et seulement si, il existe un chemin empruntant au plus  $r$  arêtes entre  $u$  et  $v$  dans  $G$ .

En ce cas, un code identifiant de  $G^r$  correspond à un code  $r$ -identifiant de  $G$ .

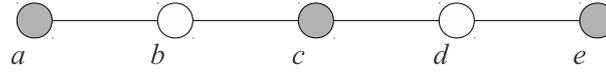
L'autre variante concerne l'identification de sous-ensembles de sommets. Pour un sous-ensemble  $X \subseteq V$ , le voisinage fermé de  $X$ , noté  $N[X]$ , est défini comme

$$N[X] = \bigcup_{u \in X} N[u].$$

Deux sous-ensembles de sommets  $X, Y$  sont alors dits séparés par  $C \subseteq V$  s'il existe un sommet de  $C$  appartenant à la différence symétrique de  $N[X]$  et  $N[Y]$ .

On peut alors combiner l'identification à distance  $r$  avec l'identification de sous-ensembles de sommets. Un *code  $(r, \leq \ell)$ -identifiant* est un sous-ensemble de sommets  $C \subseteq V$  qui couvre tous les sommets et sépare toutes les paires de sous-ensembles distincts d'au plus  $\ell$  sommets dans la fermeture  $r$ -transitive du graphe considéré.

Ainsi, un code identifiant tel que défini en Définition 3 est un code  $(1, \leq 1)$ -identifiant (voir Figure 2.1). Dans la suite on utilisera le terme « code identifiant » pour désigner un code  $(1, \leq 1)$ -identifiant. Le terme *code  $r$ -identifiant* désignera un code  $(r, \leq 1)$ -identifiant, et *code  $\ell$ -set-identifiant* désignera un code  $(1, \leq \ell)$ -identifiant.



**FIGURE 2.1 :** L'ensemble des sommets grisés forme un code identifiant du graphe. En effet,  $b$  est par exemple l'unique sommet à ne voir que  $a$  et  $c$  parmi les sommets du code. De même,  $a$  est l'unique sommet dont le voisinage fermé ne contient que  $a$  parmi les sommets du code, et ainsi de suite. Ainsi l'ensemble des sommets grisés forme un code  $(1, \leq 1)$ -identifiant. Ce code n'est par contre pas un code  $(2, \leq 1)$ -identifiant, car  $a$  et  $b$ , par exemple, sont tous deux 2-couverts par exactement les mêmes sommets du code ( $a$  et  $c$ ). Ce n'est pas non plus un code  $(1, \leq 2)$ -identifiant, car les ensembles  $\{a, c\}$  et  $\{b\}$ , par exemple, ne sont pas séparés.

Lorsqu'un graphe  $G$  admet un code identifiant, chercher la cardinalité minimum d'un tel code est un problème NP-difficile [47, 49]. Par analogie avec le problème de domination, la cardinalité minimum d'un code  $(r, \leq \ell)$ -identifiant d'un graphe  $G$  est notée  $\gamma_\ell^{\text{ID}}(G^r)$ . Lorsque  $\ell = 1$  cette quantité sera notée  $\gamma^{\text{ID}}(G^r)$ .

Si l'on s'intéresse à donner des bornes en fonction du nombre de sommets  $n$  d'un graphe, la cardinalité minimum d'un code identifiant peut varier dans un intervalle de large amplitude. Tout d'abord, si le graphe n'a aucune arête alors on a  $\gamma^{\text{ID}}(G) = n$ , car tout sommet de  $G$  doit être dans le code identifiant, ne serait-ce que pour se couvrir lui-même. Dans ce premier exemple extrême le graphe n'admet qu'un unique code identifiant, l'ensemble de tous ses sommets. D'autre part, regardons d'un peu plus près la contrainte de séparation. Celle-ci implique en particulier l'existence d'une injection de  $V$  dans l'ensemble des parties non vides de  $C$ , car tout sommet  $v$  est uniquement déterminé par le sous-ensemble de  $C$  dans son voisinage fermé  $N[v]$ . Ceci implique que  $n = |V| \geq 2^{|C|} - 1$ , ce qui équivaut à  $|C| \geq \lceil \log_2(n + 1) \rceil$ . Ainsi on a les inégalités suivantes (déjà connues dans l'article de Karpovsky *et al* [120]) :

**Proposition 1** (Bornes générales). *Soit  $G$  un graphe sans jumeaux à  $n$  sommets,  $n \geq 1$ . Alors on a*

$$\lceil \log_2(n + 1) \rceil \leq \gamma^{\text{ID}}(G) \leq n.$$

Ces bornes illustrent la différence fondamentale entre les codes identifiants et les codes couvrants (ou dominants). En effet, dans un dominant la borne inférieure est 1, car tout graphe admettant un sommet  $u$  voisin de tous les autres admet  $\{u\}$  comme ensemble dominant. Un graphe extrémal possédant cette propriété est par exemple l'étoile  $K_{1,n}$ , dont les sommets sont  $u, v_1, \dots, v_n$  et les arêtes  $\{uv_i \mid i = 1, \dots, n\}$ . Or le graphe

$K_{1,n}$  n'admet aucun code identifiant de cardinalité 1, ni même de cardinalité  $\lceil \log_2(n+1) \rceil$ . En effet, comme chaque sommet  $v_i$  doit pouvoir être séparé d'un autre sommet  $v_j$ , alors il ne peut pas exister plus d'un sommet  $v_i$  qui ne soit pas dans le code identifiant. Comme par ailleurs chaque sommet doit être couvert, alors on a en fait  $\gamma^{\text{ID}}(K_{1,n}) = n$ . Dans le cadre des codes identifiants, l'étoile est donc plutôt un graphe extrémal pour la borne supérieure ! Cet exemple montre que cette propriété additionnelle de séparation introduit des contraintes qui peuvent rendre très différentes les structures des codes couvrants et des codes identifiants.

## 2.1.2 Trois exemples d'applications

Nous allons donner ici trois exemples d'applications des codes identifiants, dans divers contextes. La première application concerne la détection de défaillances dans les réseaux multiprocesseurs. C'est cette application qui a, à l'origine, motivé l'introduction de ces codes [120]. Les deux autres applications présentées ici sont, dans l'ordre chronologique d'apparition dans la littérature, la surveillance de bâtiments par des réseaux de capteurs, et l'analyse de structures secondaires d'ARN (acide ribonucléique).

### 2.1.2.1 Détection de défaillances dans les réseaux multiprocesseurs

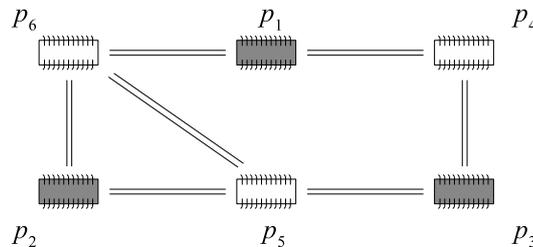
Les codes identifiants ont été introduits à l'origine pour modéliser un problème de détection de défaillances dans des réseaux multiprocesseurs [120]. Considérons en effet un tel réseau, dans lequel on suppose que chaque processeur puisse exécuter une certaine procédure test sur lui-même et l'ensemble de ses voisins, dont le résultat serait binaire (« test OK » ou « problème détecté »).

Supposons dans un premier temps que le réseau contient au plus un processeur défectueux. On suppose également que la réponse « test OK » nous assure que le processeur fonctionne correctement ainsi que l'ensemble de ses voisins, alors que la réponse « problème détecté » nous indique la présence d'un défaut, sans pour autant nous dire si c'est le processeur ayant lancé le test qui a lui-même un problème, ou si c'est l'un de ses voisins (et lequel).

Si l'ensemble des processeurs ayant exécuté la procédure test forme un code identifiant du graphe sous-jacent au réseau, alors les choses se passent de façon relativement agréable pour le superviseur ayant lancé les tests.

En effet, si tous les processeurs répondent « test OK » alors on sait qu'il n'y a aucun processeur défectueux dans le réseau. Si certains des processeurs interrogés répondent « problème détecté », alors l'ensemble des sommets

ayant fourni cette réponse permet de déterminer de façon unique le processeur défectueux. Il suffit en effet de déterminer l'unique sommet dont l'intersection du voisinage fermé avec le code est égale à l'ensemble de sommets ayant répondu « problème détecté ».



**FIGURE 2.2 :** Les processeurs  $p_1, p_2, p_3$  permettent d'identifier de façon unique chaque processeur du réseau. Les processeurs  $p_4, p_5, p_6$ , quant à eux, couvrent bien le réseau mais ne permettent pas d'identifier de façon unique ses processeurs en cas de défaillance de l'un de ceux-ci. En effet, il serait impossible de distinguer le cas  $p_5$  défectueux du cas  $p_6$  défectueux si l'on prenait comme code identifiant l'ensemble  $\{p_4, p_5, p_6\}$ .

Si l'on suppose maintenant qu'au plus  $\ell$  processeurs peuvent être défectueux dans le réseau ( $\ell \geq 1$  entier), alors le modèle adéquat est naturellement celui de code  $(1, \leq \ell)$ -identifiant.

Cette application a motivé l'introduction de procédés dynamiques d'identification, où l'on s'autorise à interroger les processeurs les uns après les autres, et à adapter la séquence des processeurs interrogés en fonction des réponses obtenues (voir Section 2.5).

### 2.1.2.2 Surveillance de bâtiments par des réseaux de capteurs

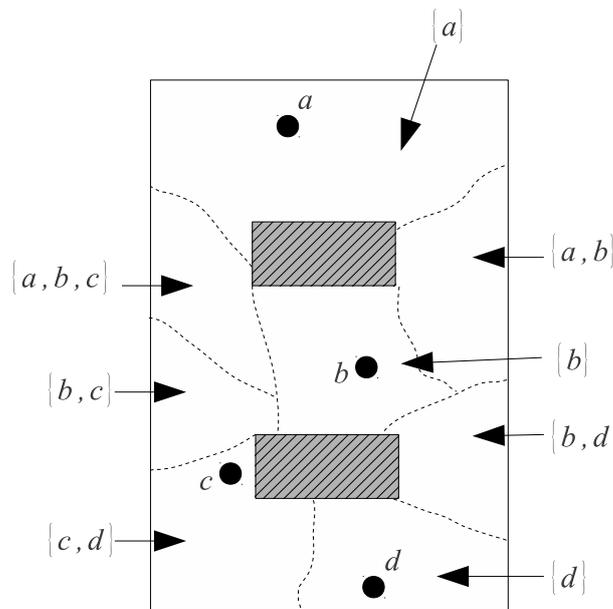
En 2003–2004, une équipe de l'université de Boston, autour d'Ari Trachtenberg, a développé une application des codes identifiants dans le cadre de la surveillance de bâtiments munis de réseaux de capteurs [157, 171].

Une implémentation possible de cette application est par exemple l'implantation de détecteurs de fumée, afin de pouvoir localiser un éventuel incendie (cette application est décrite en détail dans la thèse de David Auger [8, Chapitre 1]). Une autre implémentation concerne la dissémination d'un réseau de capteurs au sein d'un bâtiment en vue de pouvoir localiser en temps réel les personnes présentes.

C'est cette dernière application que nous allons détailler ci-après.

On suppose les personnes munies de badges émettant en continu un signal radio susceptible d'être détecté par les capteurs. Chaque capteur

possède une *zone de contrôle*, qui est une partie du bâtiment dont les emplacements sont tous à portée du capteur. Cette zone de contrôle dépend de nombre de paramètres, dont l'intensité du signal émis par le capteur et la configuration du bâtiment (présence d'obstacles ou de réflecteurs). L'ensemble des zones de contrôle des capteurs discrétise le bâtiment en *secteurs* (voir Figure 2.3), et l'objectif est de pouvoir à tout moment connaître le secteur dans lequel se trouve un individu donné.



**FIGURE 2.3 :** Les recouvrements des zones de contrôle des capteurs  $a, b, c, d$  discrétisent le bâtiment en huit secteurs. Chaque secteur est déterminé de façon unique par l'ensemble des capteurs à portée. Le secteur en bas à gauche est par exemple défini comme étant l'ensemble des points à portée des capteurs  $c$  et  $d$  seulement.

Les applications classiques des réseaux de capteurs pour la localisation de personnes correspondent à des codes couvrants du graphe sous-jacent, dans un bâtiment discrétisé en autant de secteurs que de capteurs. La condition de couverture assure en effet la couverture de tous les secteurs du bâtiment, et la position d'un individu correspond au secteur du capteur le plus proche.

L'utilisation des codes identifiants permet d'obtenir une configuration du réseau tirant parti des éventuels recouvrements de zones de contrôle. La condition de séparation permet en effet de localiser les personnes avec une précision plus fine qu'en utilisant un simple code couvrant, et le bâtiment se retrouve discrétisé en un plus grand nombre de secteurs qu'avec un tel code.

Cette application soulève des questions de robustesse [157], qui ont donné lieu à des développements théoriques sur les codes disjoints [129] et les codes connexes [73]. Il est également à noter que, dans un tel contexte, les codes identifiants fournissent de façon naturelle des protocoles de routage compacts dans le réseau de capteurs sous-jacents [131].

### 2.1.2.3 Analyse de structures secondaires d'ARN

Les graphes sont un modèle standard de représentation de structures moléculaires. L'équipe de Tamar Schlick (Université de New York) développe en particulier depuis quelques années le modèle RAG (« RNA As Graphs »), qui utilise une représentation de structures secondaires d'ARN sous forme d'arbres [75, 85, 86]. Cette représentation n'est pas surjective, dans le sens où il existe des arbres ne correspondant à aucune structure d'ARN. Sachant que l'ensemble des structures d'ARN existants n'est pas encore connu de façon complète, les arbres à  $n \leq 8$  sommets sont actuellement classés en trois catégories :

- (1) arbre correspondant à une structure d'ARN existante dans la nature et identifiée dans la littérature
- (2) arbre ressemblant à une structure d'ARN (structure « candidate »)
- (3) arbre ne ressemblant pas à une structure d'ARN

Cette classification est le fruit des travaux de Schlick *et al* [86]. L'interprétation usuelle est que la catégorie (2) contient des structures existantes dans la nature mais non encore identifiées, ou des structures qu'il serait possible de synthétiser. La catégorie (3) quant à elle contient des arbres ne correspondant pas à des structures possibles (pour des raisons qui tiennent à la biologie).

Dans [105], Haynes *et al* proposent des techniques permettant de prédire si un arbre donné peut correspondre à une structure d'ARN. Ces techniques utilisent les codes identifiants, ainsi que d'autres variantes de domination. Il s'avère que pour les arbres à 7 et 8 sommets (qui sont le périmètre de l'étude de [105]), Haynes *et al* parviennent à définir des modèles de prédiction classifiant *tous* les arbres des catégories (1) et (3) dans la bonne catégorie. Leur interprétation est que les paramètres de domination sont une bonne mesure de l'efficacité de l'arbre en tant que réseau de communication.

### 2.1.3 Dynamique de la thématique de recherche

Depuis leur introduction en 1998 [120], les codes identifiants ont fait l'objet de nombreux travaux de recherche. Antoine Lobstein (Télécom Pa-

risTech) maintient une bibliographie en ligne sur le sujet (ainsi que ses variantes), comportant à ce jour plus de 200 références [140].

Outre ma thèse [146], six autres thèses de doctorat soutenues à ce jour ont concerné les codes identifiants. Il s'agit des thèses des personnes suivantes :

- Sanna Ranto [155] (dirigée par Iiro Honkala, université de Turku, soutenue en 2007)
- Robert Skaggs [164] (dirigée par Marietjie Frick et Gerd Fricke, université d'Afrique du Sud, soutenue en 2007)
- Moshe Laifenfeld [128] (dirigée par Ari Trachtenberg, université de Boston, soutenue en 2008)
- David Auger [8] (dirigée par Olivier Hudry, Télécom ParisTech, soutenue en 2010)
- Ville Junnila [116] (dirigée par Tero Laihonen, université de Turku, soutenue en 2011)
- Brendon Stanton [166] (dirigée par Ryan Martin, université d'Iowa, soutenue en 2011)

Il existe de plus, à ma connaissance, actuellement quatre thèses en cours sur le sujet, celle de Florent Foucaud (dirigée par André Raspaud et Ralf Klasing, université de Bordeaux), celle d'Aline Parreau (dirigée par Sylvain Gravier, université de Grenoble), celle de Mikko Pelto (dirigée par Iiro Honkala, université de Turku en Finlande) et celle de Marwane Bouznif (dirigée par Myriam Preissmann et moi-même, université de Grenoble).

Le réseau de recherche sur ces thématiques est donc de dimension internationale. Les principaux groupes de recherche se situent en France (Télécom ParisTech, université de Bordeaux, université de Grenoble), en Finlande (université de Turku), et aux États-Unis (université de Boston). J'ai participé à l'organisation du premier workshop international sur le sujet, « Codes and discrete probability », qui a eu lieu à Grenoble du 3 au 6 décembre 2007. L'ANR a financé le projet IDEA (« Identifying coDes in Evolving grAphs », <https://idea.labri.fr/>) de 2009 à 2011, dont j'ai été membre permanent. Ce projet avait pour objectif de développer de nouvelles approches sur ces codes. Le projet IDEA a organisé dans ce cadre le deuxième workshop international sur le sujet, le « Bordeaux Workshop on Identifying Codes », du 21 au 25 novembre 2011. Cet évènement a réuni plus de 60 participants venus de toute l'Europe, des États-Unis, et d'Australie.

## 2.2 Structures régulières

Si  $G$  est un graphe quelconque sans jumeaux, on sait que calculer la valeur de  $\gamma^{\text{ID}}(G)$  est un problème NP-difficile [47, 49].

Une démarche naturelle consiste alors à étudier ces codes dans des classes de graphes particulières. Parmi les classes de graphes classiques, nombreuses sont celles présentant une forte régularité de structure. Les structures régulières sont des domaines d'étude naturels de ces codes, dans la mesure où certaines de ces structures (par exemple les grilles ou les hypercubes) sont des modèles fréquents de réseaux d'interconnexion. De plus, ces structures possèdent de nombreuses symétries, ce qui facilite parfois l'étude des problèmes.

Je présente ci-dessous l'essentiel des résultats obtenus concernant les grilles (Section 2.2.1), les cycles (Section 2.2.2), les hypercubes (Section 2.2.3), les produits de cliques (Section 2.2.4), et les graphes de Sierpiński (Section 2.2.5), ainsi que le contexte scientifique dans lequel s'insèrent ces résultats. Le cas échéant, je donne également les derniers développements de la littérature sur ces thématiques.

### 2.2.1 Grilles et bandes

Les premiers travaux sur les codes identifiants succédant à l'article initial de Karpovsky, Chakrabarty et Levitin [120] ont majoritairement concerné les hypercubes et les grilles infinies, autour notamment du noyau dur Charon-Hudry-Lobstein de Télécom ParisTech (anciennement École Nationale Supérieure des Télécommunications). La motivation pour étudier les grilles infinies provient de l'inégalité suivante, démontrée dans [120] :

**Théorème 1** (Borne inférieure générale pour les graphes réguliers). *Soit  $G = (V, E)$  un graphe et  $r \geq 1$  tel que  $G^r$  est  $\Delta$ -régulier (où  $G^r$  désigne le graphe de sommets  $V$  tel que deux sommets  $u$  et  $v$  sont voisins dans  $G^r$  si, et seulement si, il existe dans  $G$  un chemin empruntant au plus  $r$  arêtes entre  $u$  et  $v$ ). Si  $G^r$  admet un code identifiant, alors on a :*

$$\gamma^{\text{ID}}(G^r) \geq \frac{2|V|}{\Delta + 2}.$$

Ce résultat est une conséquence de l'observation suivante : si  $C$  est un code identifiant de  $G^r$ , alors il y a au plus  $|C|$  sommets de  $G^r$  qui ne sont couverts que par un unique sommet du code (conséquence de la condition de séparation). Ainsi, il y a au moins  $|V| - |C|$  sommets de  $G^r$  qui sont couverts

par au moins deux sommets de  $C$ . En comptant de deux façons le nombre de paires  $(u, v)$  avec  $u$  sommet de  $G^r$  et  $v$  un sommet de  $C$  qui couvre  $u$ , on obtient l'inégalité

$$2(|V| - |C|) \leq (\Delta + 1)|C|,$$

ce qui est l'inégalité désirée.

Cette inégalité est serrée, dans le sens où il existe des graphes pour lesquels la cardinalité du code est égale à la borne. C'est par exemple le cas pour les cycles pairs. Un cycle pair  $C_{2n}$  est un graphe dont les sommets sont  $\{v_1, \dots, v_{2n}\}$ , et tel que  $v_i v_{i+1}$  est une arête pour tout  $i = 1, \dots, 2n$ , les indices étant pris modulo  $2n$ . En ce cas, il est facile de voir que le sous-ensemble de sommets  $\{v_i \mid i \text{ pair}\}$  est un code identifiant de  $C_{2n}$ . Or, la cardinalité de ce code est de  $n$ , ce qui est égal à la borne du Théorème 1.

On dénombre plus d'une trentaine d'articles au sujet de codes identifiants dans les grilles. Les structures les plus étudiées sont les grilles carrées, triangulaires, hexagonales (structure en nids d'abeille), ainsi que la grille dite royale, qui est obtenue à partir de la grille carrée en rajoutant les deux diagonales du carré (c'est la grille des déplacements légaux du roi au jeu d'échec). Le plus souvent, les structures étudiées sont les grilles infinies ou les grilles toroïdales, ce qui permet d'appliquer le Théorème 1, ou tout du moins de considérer une structure régulière sans se soucier des effets de bord (qui compliquent en général considérablement les choses). C'est notamment le cas dans [22, 44, 45, 48, 59, 61, 108, 109, 111, 141].

### 2.2.1.1 Bandes infinies

La grille carrée infinie, dénotée communément (et abusivement)  $\mathbb{Z}^2$ , est le graphe infini dont les sommets sont les points de  $\mathbb{Z}^2$ , et tel qu'il existe une arête entre  $(i, j)$  et  $(i', j')$  si et seulement si  $|i - i'| + |j - j'| = 1$ .

Dans un article paru en 2004 dans *Theoretical Computer Science* [62], nous abordons avec Marc Daniel et Sylvain Gravier l'étude de bandes infinies, qui sont des sous-graphes de la grille carrée infinie.

Une *bande* (infinie) de hauteur  $k \geq 1$ , notée  $B_k$ , a pour sommets l'ensemble  $\{1, \dots, k\} \times \mathbb{Z}$ . Deux sommets  $(i, j)$  et  $(i', j')$  de la bande sont voisins si et seulement si  $|i - i'| + |j - j'| = 1$ .

Dans le cas des graphes infinis, on parle de la densité d'un code plutôt que de sa cardinalité.

Pour tout sommet  $v$  et pour tout entier  $r \geq 1$ , soit  $B_r(v)$  l'ensemble des sommets reliés par un chemin d'au plus  $r$  arêtes à  $v$ . L'ensemble  $B_r(v)$  est appelé la *boule* de rayon  $r$  centrée en  $v$ . Noter que la boule de rayon 1 centrée

en un sommet  $v$  n'est autre que son voisinage fermé : on a  $B_1(v) = N[v]$  pour tout sommet  $v$ .

**Définition 4** (Densité d'un code). *Soit  $C$  un sous-ensemble de sommets de  $\mathbb{Z}^2$  (ou de  $B_k$ ,  $k \geq 1$ ), et soit  $v_0$  un sommet quelconque. La densité de  $C$ , notée  $d(C)$ , est définie comme*

$$d(C) = \limsup \frac{|B_r(v_0) \cap C|}{|B_r(v_0)|}.$$

Noter que la notion de densité d'un code est indépendante du choix du sommet initial. En effet, pour deux sommets  $v$  et  $v'$  de  $\mathbb{Z}^2$ , la cardinalité de  $B_r(v) \cap B_r(v')$  est en  $\Theta(r^2)$ , alors que celle de  $B_r(v) \Delta B_r(v')$  est en  $\Theta(r)$ . De même, dans  $B_k$ ,  $k \geq 1$ , on a  $|B_r(v) \cap B_r(v')| = \Theta(r)$ , alors que  $|B_r(v) \Delta B_r(v')| = \Theta(1)$ .

Après plusieurs articles sur le sujet, qui ont peu à peu rapproché des bornes inférieures et des bornes supérieures établies par différents auteurs [57, 59, 120], Yael Ben-Haim et Simon Litsyn, de l'université de Tel-Aviv, sont parvenus en 2005 à déterminer la densité minimum d'un code identifiant de la grille infinie [22] :

**Théorème 2** (Densité optimale dans  $\mathbb{Z}^2$ ). *Soit  $d^*(\mathbb{Z}^2)$  la densité minimum d'un code identifiant dans la grille infinie  $\mathbb{Z}^2$ . Alors on a*

$$d^*(\mathbb{Z}^2) = \frac{7}{20}.$$

Une différence importante entre les bandes infinies et les grilles infinies (ou les grilles toroïdales finies) réside précisément dans la présence d'effets de bords, liés à la gestion des sommets de degré inférieur ou égal à 3. Ces effets sont d'autant plus importants que  $k$  est petit, la densité optimale d'un code identifiant de  $B_k$  étant naturellement asymptotiquement égale (lorsque  $k$  tend vers l'infini) à celle de la grille carrée infinie. Dans [62] nous montrons :

**Théorème 3** (Densité optimale dans  $B_2$ ). *Soit  $d^*(B_2)$  la densité minimum d'un code identifiant dans la bande infinie de hauteur 2 dans la grille carrée. Alors on a*

$$d^*(B_2) = \frac{3}{7}.$$

Notons que cette densité de  $\frac{3}{7} \simeq 0,43$  est bien loin de  $\frac{7}{20} = 0,35$ , la densité optimale d'un code identifiant dans la grille carrée infinie. Ceci illustre l'importance des effets de bord pour ce problème.

Ce résultat est obtenu à l'issue d'une décomposition de  $B_2$  en blocs correspondant à certaines configurations locales d'un code identifiant optimal.

L'étude des successions possibles entre ces blocs nous conduit à une borne inférieure égale à  $\frac{3}{7}$ . Nous concluons en observant que la translation du motif 7-périodique suivant est un code identifiant de  $B_2$ , de densité  $\frac{6}{14} = \frac{3}{7}$  (voir Figure 2.4).

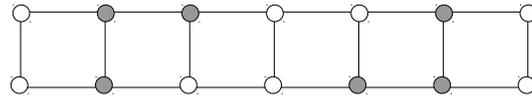


FIGURE 2.4 : La translation de ce motif 7-périodique est un code identifiant de la bande infinie  $B_2$  (les sommets du code sont les sommets grisés).

Cette preuve, quoique élémentaire, est relativement fastidieuse. Nous avons récemment trouvé un argument plus simple utilisant une technique de déchargement avec Marwane Bouznif (laboratoire G-SCOP, Grenoble) et Myriam Preissmann (laboratoire G-SCOP, Grenoble).

De façon générale, une méthode de *déchargement* fonctionne de la façon suivante. On suppose tout d'abord disposer d'un sous-ensemble de sommets  $C$ , dont on souhaite borner inférieurement la cardinalité. Chaque sommet de  $C$  va recevoir une *charge* initiale égale à 1, et chaque sommet qui n'est pas dans  $C$  va recevoir une charge initiale égale à 0. Les sommets de  $C$  vont alors devoir donner une partie de leur charge aux autres sommets, selon un ensemble fini de règles  $R_1, R_2, \dots, R_k$ . On démontre ensuite qu'après le partage des charges selon les règles définies, tous les sommets du graphe ont une charge finale qui est supérieure ou égale à une certaine quantité  $\delta > 0$ . Comme il n'y a que des transferts de charge entre les sommets, alors la somme des charges initiales est égale à la somme des charges finales. Ceci implique que  $|C| \geq \delta|V|$ , et nous fournit donc une borne inférieure sur la cardinalité de  $C$ . Dans le cas où le graphe est infini, cela nous fournit une borne inférieure sur la densité du code.

Considérons maintenant  $B_2$ , la bande infinie de hauteur 2. On décompose  $C$  en  $C_0 \cup C_1 \cup C_2 \cup C_3$ , et  $V \setminus C$  en  $U_1 \cup U_2 \cup U_3$ , avec :

- $C_i$  = ensemble des sommets de  $C$  qui ont exactement  $i$  voisins qui ne sont pas dans  $C$
- $U_i$  = ensemble des sommets de  $V \setminus C$  qui ont exactement  $i$  voisins qui sont dans  $C$

On considère alors les règles de déchargement suivantes :

- R1 : un sommet de  $v \in C$  donne  $\frac{3}{7i}$  de charge à tout sommet  $u \in U_i$  dont il est voisin
- R2 : un sommet de  $v \in C_0 \cup C_1$  donne  $\frac{1}{14}$  de charge à tout sommet  $u \in C_2$  dont il est voisin

Montrons alors que, après application des règles R1 et R2, tout sommet a une charge supérieure ou égale à  $\frac{3}{7}$ .

Par définition, un sommet  $v \in U_i$  reçoit donc exactement  $i \times \frac{3}{7i} = \frac{3}{7}$  de charge. Tout sommet de  $V \setminus C$  a donc une charge finale égale à  $\frac{3}{7}$ . Maintenant il nous faut montrer que tout sommet de  $C$  ne perd pas trop de charge.

Notons qu'un sommet de  $C$  a au plus un voisin dans  $U_1$ , car dans le cas contraire ces deux voisins ne seraient pas séparés.

Clairement, un sommet  $v \in C_0$  ne donne de la charge que dans la règle R2, car il n'a, par définition, pas de voisin dans  $V \setminus C$ . Un tel sommet ayant au plus trois voisins dans  $C$  (voir Figure 2.5(a)), alors sa charge finale est d'au moins  $1 - 3 \frac{1}{14} = \frac{11}{14} > \frac{3}{7}$ .

Le cas où un sommet de  $C_1$  perd le plus de charge est celui où il est voisin d'un sommet de  $U_1$  et de deux sommets de  $C_2$  (voir Figure 2.5(b)). En ce cas, sa charge finale est alors de  $1 - \frac{3}{7} - 2 \frac{1}{14} = \frac{6}{14} = \frac{3}{7}$ .

Le cas où un sommet  $u$  de  $C_2$  perd le plus de charge à la règle R1 est celui où il est voisin d'un sommet de  $U_1$  et d'un sommet de  $U_2$  (voir Figure 2.5(c)). En ce cas, sa charge est alors de  $1 - \frac{3}{7} - \frac{3}{14} = \frac{5}{14}$  après application de la règle R1. Mais un tel sommet est adjacent à un sommet  $v \in C$ , qui ne peut appartenir ni à  $C_2$  (sinon  $u$  et  $v$  ne seraient pas séparés), ni à  $C_3$  (par définition de  $C_3$ ). Ce sommet  $v$  est donc dans  $C_0 \cup C_1$ , et donne donc  $\frac{1}{14}$  à  $u$ . Au final  $u$  a donc une charge égale à  $\frac{5}{14} + \frac{1}{14} = \frac{3}{7}$ . Dans les autres cas, tout sommet de  $C_2$  finit avec au moins  $\frac{3}{7}$  de charge.

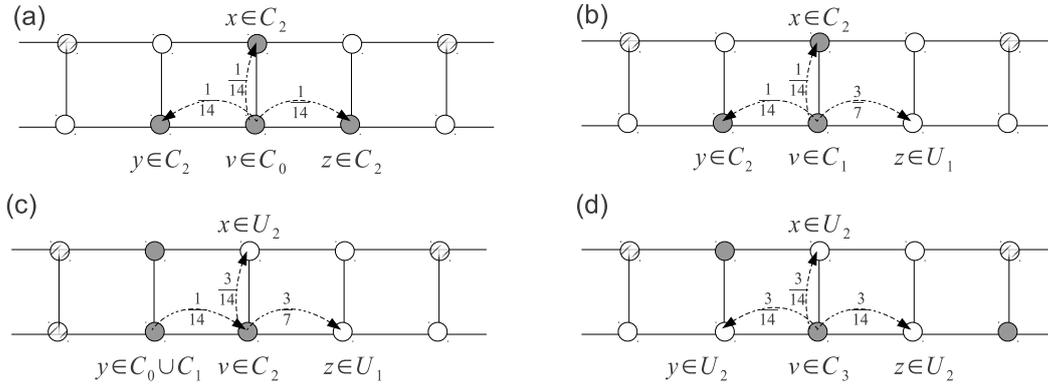
Enfin, considérons un sommet  $u$  de  $C_3$ . Celui-ci n'a aucun voisin dans  $U_1$  (car un voisin  $v \in U_1$  de  $u$  ne serait pas séparé de  $u$ ). Il ne peut également pas avoir ses trois voisins dans  $U_2$ , car en ce cas deux de ces voisins ne seraient pas séparés. Le cas où il donne le plus de charge est donc le cas où il a deux voisins dans  $U_2$  et un dans  $U_3$  (voir Figure 2.5(d)). En ce cas, sa charge finale est de  $1 - 2 \frac{3}{14} - \frac{3}{21} = \frac{3}{7}$ .

Au final tout sommet de  $C$  a une charge supérieure ou égale à  $\frac{3}{7}$ , et tout sommet de  $V \setminus C$  a une charge de  $\frac{3}{7}$ . La densité d'un code identifiant de  $B_2$  est donc supérieure ou égale à  $\frac{3}{7}$ .

Cette analyse nous permet de plus d'obtenir l'unicité du code optimal, à symétries près, et à altération sur un nombre fini de sommets près.

Nous avons également donné dans [62] des bornes générales sur la densité d'un code identifiant dans  $B_k$ , avec  $k \geq 3$  quelconque. Nous montrons en particulier les inégalités suivantes :

**Théorème 4** (Bornes générales sur les bandes infinies). *Soit  $d^*(B_k)$  la densité*



**FIGURE 2.5 :** Les sommets grisés sont dans le code, les sommets blancs ne sont pas dans le code, et les sommets hachurés ont un statut indéterminé. Les cas représentés sont donnés sans perte de généralité.

optimale d'un code identifiant dans la bande de hauteur  $k$ , avec  $k \geq 3$ . Alors on a

$$\frac{7}{20} - \frac{1}{2k} \leq d^*(B_k) \leq \min \left( \frac{2}{5}, \frac{7}{20} + \frac{2}{k} \right).$$

Ces résultats sont obtenus grâce aux arguments suivants.

Considérons un code identifiant optimal de  $\mathbb{Z}^2$ . Coupons alors une bande de hauteur  $k$  dans un code optimal de  $\mathbb{Z}^2$ , et forçons ensuite chaque sommet sur les premières et dernières lignes de  $B_k$  à être dans le code. On obtient alors un code identifiant de  $B_k$ . Or, il existe nécessairement une façon de découper cette bande de hauteur  $k$  de sorte à ce que la densité du code ainsi obtenu soit au plus de  $\frac{7}{20} + \frac{2}{k}$ . Ceci démontre la borne  $d^*(B_k) \leq \frac{7}{20} + \frac{2}{k}$ .

Considérons le code identifiant tel que  $(i, j)$  est dans le code si, et seulement si,  $j \equiv 1 \pmod{5}$  ou  $j \equiv 3 \pmod{5}$ . Ce code est de densité  $\frac{2}{5}$ ; il permet de démontrer la borne  $d^*(B_k) \leq \frac{2}{5}$ .

Enfin, considérons un code identifiant optimal de  $B_k$ . Sa translation pave  $\mathbb{Z}^2$  et en est un code couvrant. Les paires de sommets sur deux lignes à la frontière entre deux copies de  $B_k$  ne sont alors peut-être pas toutes séparées. En ajoutant au plus un sommet dans le code pour chaque telle paire, on obtient un code identifiant de  $\mathbb{Z}^2$ , ce qui démontre que  $d^*(B_k) \geq \frac{7}{20} - \frac{1}{2k}$ .

Frédéric Havet a récemment amélioré ces bornes en utilisant des techniques de déchargement [103] :

**Théorème 5** (Bornes inférieures améliorées dans  $B_k$ ). Soit  $d^*(B_k)$  la densité

optimale d'un code identifiant de la bande  $B_k$ , avec  $k \geq 3$ . Alors on a

$$d^*(B_k) \geq \max \left( \frac{10}{29}, \frac{7}{20} - \frac{1}{2k} \right).$$

Ces bornes améliorent celles du Théorème 4 pour  $k \leq 96$ . Il parvient également, au terme d'une application particulièrement fastidieuse de la méthode de déchargement (avec notamment 11 règles de déchargement), à calculer la densité optimale d'un code identifiant de  $B_3$  :

**Théorème 6** (Densité optimale dans  $B_3$ ). *Soit  $d^*(B_3)$  la densité minimum d'un code identifiant dans la bande infinie de hauteur 3 dans la grille carrée. Alors on a*

$$d^*(B_3) = \frac{7}{18}.$$

Avec Marwane Bouznif, Julien Darlay, et Myriam Preissmann, nous avons également obtenu ce résultat par un calcul par ordinateur détaillé dans la section ci-après. Ce résultat a été présenté par Marwane lors des Journées Graphes et Algorithmes en 2010, ce qui a d'ailleurs été à l'origine du travail de Frédéric Havet sur ces codes.

### 2.2.1.2 Programmation dynamique

Le cas sans doute le plus difficile lorsque l'on étudie les codes identifiants dans les grilles est celui des grilles finies. Une *grille* carrée (finie) de dimensions  $k, n$  est le sous-graphe de la grille infinie induit par les sommets  $(i, j)$  tels que  $1 \leq i \leq k$  et  $1 \leq j \leq n$ . Par analogie avec les bandes, une telle grille sera notée  $B_{k,n}$ .

Dans l'article de *Theoretical Computer Science* écrit avec Marc Daniel et Sylvain Gravier [62], nous avons également développé une méthode algorithmique permettant de calculer par ordinateur la cardinalité minimum d'un code identifiant dans des grilles finies  $B_{k,n}$ . Cette méthode est l'adaptation d'un algorithme de programmation dynamique qui avait à l'origine été développé pour le problème de domination dans les grilles [127, 139, 165, 179]. Elle est basée sur la recherche de chemins ayant un nombre d'arcs prescrits dans un graphe auxiliaire orienté, dont le nombre de sommets ne dépend que de  $k$ . Une caractéristique importante de cette méthode est l'obtention de formules en temps constant.

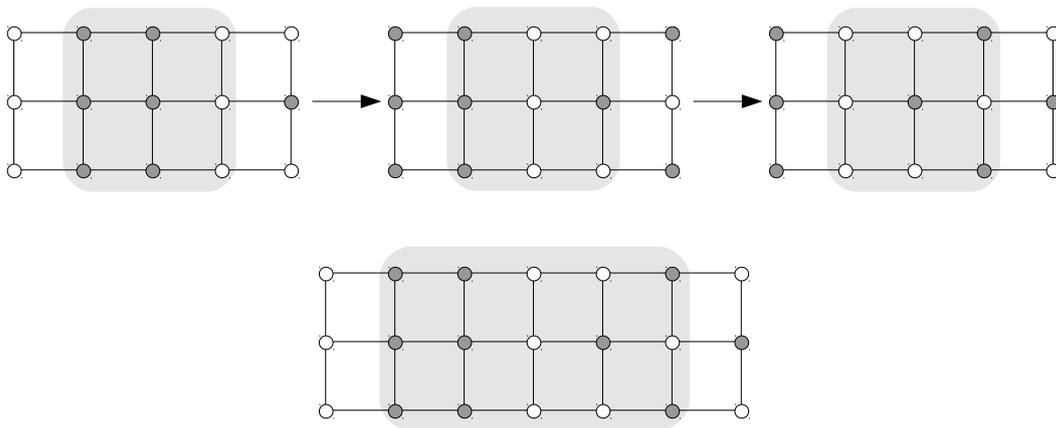
Ce graphe auxiliaire est basé sur la notion de précode d'une grille  $B_{k,n}$ . Un sous-ensemble de sommets  $C$  d'une telle grille est un *précode* si, et seulement si, tous les sommets de  $\{(i, j) \in V(B_{k,n}) \mid 2 \leq j \leq k - 1\}$  sont couverts

et séparés par les sommets de  $C$ . Un précode d'une grille est donc un sous-ensemble de sommets qui identifie tous les sommets, sauf éventuellement ceux de la première et de la dernière colonne de  $B_{k,n}$ .

Le graphe auxiliaire orienté  $\mathcal{G}_k$  associé à une grille  $B_{k,n}$  est alors le graphe dont les sommets sont l'ensemble des grilles  $B_{k,5}$  munies d'un précode (un tel sommet sera noté  $(B_{k,5}, C)$ ), plus deux sommets additionnels début et fin. Un arc sort d'un sommet  $(B_{k,5}, C)$  pour aller vers un sommet  $(B_{k,5}, C')$  dès lors qu'une condition de superposabilité de  $(B_{k,5}, C)$  et  $(B_{k,5}, C')$  est remplie (voir Figure 2.6). Intuitivement, cette condition de superposabilité est telle que deux sommets  $u$  et  $v$  reliés par un arc  $(u, v)$  correspondent alors à une grille  $B_{n,6}$  munie d'un précode. Un chemin orienté de  $\mathcal{G}_k$  empruntant  $t$  arcs correspond alors à une grille  $B_{k,5+t}$  munie d'un précode.

Il y a, de plus, un arc depuis début vers tout sommet  $(B_{k,5}, C)$  tel que les sommets de la première colonne de  $B_{k,5}$  soient également couverts et identifiés par  $C$ . De même, il y a un arc depuis  $(B_{k,5}, C)$  vers fin si les sommets de la dernière colonne de  $B_{k,5}$  sont également couverts et identifiés par  $C$ .

La recherche d'un code identifiant dans  $B_{k,n}$  se ramène alors à la recherche d'un chemin empruntant  $n - 3$  arcs dans  $\mathcal{G}_k$ , partant de début et allant vers fin. En pondérant de façon appropriée les arcs de  $\mathcal{G}_k$ , on peut faire en sorte que la longueur d'un tel chemin corresponde à la cardinalité du code.



**FIGURE 2.6 :** Ces grilles  $B_{3,5}$  munies des précodes représentés (les sommets grisés sont les sommets du code) sont des sommets de  $\mathcal{G}_3$ . Ces trois sommets de  $\mathcal{G}_3$  forment par ailleurs un chemin dans  $\mathcal{G}_3$ . En effet, les quatre dernières colonnes d'un sommet sont superposables aux quatre premières colonnes du sommet suivant. Le chemin constitué de ces trois sommets correspond à une grille  $B_{3,7}$  munie d'un précode.

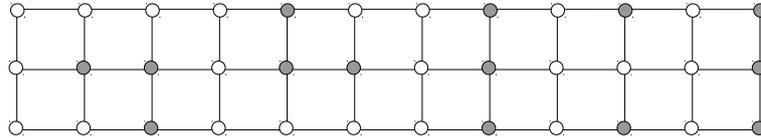
En utilisant la régularité de la structure de  $B_{k,n}$  et la « localité » du problème de code identifiant, nous parvenons :

- d’une part, à démontrer qu’il existe des solutions optimales périodiques ou pseudo-périodiques ;
- d’autre part, à faire détecter cette périodicité ou pseudo-périodicité par l’algorithme.

Comme nous montrons par ailleurs que la périodicité ou la pseudo-périodicité est bornée par une quantité dépendant uniquement de  $k$ , nous obtenons alors un algorithme qui, à  $k$  fixé, détermine en temps constant des paramètres  $T, \nu, \alpha_0, \dots, \alpha_{T-1}$ , tels que :

$$\gamma^{\text{ID}}(B_{k,n}) = \alpha_i + \nu \left\lceil \frac{n}{T} \right\rceil \quad \text{si } n \equiv i \pmod{T} \quad (2.1)$$

Nous avons récemment implémenté cet algorithme pour étudier la bande  $B_3$  [34], et avons eu la surprise de voir apparaître un motif optimal de période 12 (voir Figure 2.7). Ce motif, de densité  $\frac{14}{36} = \frac{7}{18}$ , permet en particulier d’infirmar la conjecture que nous avons faite dans [99] avec Sylvain Gravier et Ahmed Semri. Nous pensions en effet que la densité optimale d’un code dans  $B_3$  était de  $\frac{2}{5}$ . Ce résultat a ensuite été démontré par Frédéric Havet en utilisant une méthode de déchargement [103] (voir Théorème 6 dans la section précédente).



**FIGURE 2.7 :** La translation de ce motif 12-périodique est un code identifiant de la bande infinie  $B_3$  (les sommets du code sont les sommets grisés). Ce motif a été obtenu par l’implémentation d’un calcul sur ordinateur [34], basé sur l’algorithme de programmation dynamique décrit dans [36]. Il a ensuite été redémontré comme étant effectivement optimal par Frédéric Havet [103].

L’intérêt de cet algorithme est de nous permettre d’obtenir les valeurs exactes de  $\gamma^{\text{ID}}(B_{n,k})$ , pour tout  $n$ , et pour de petites valeurs de  $k$ . Ces valeurs sont exprimées par une formule close du type de (2.1).

Par ailleurs, cet algorithme est applicable pour d’autres types de grilles, en particulier les grilles royales, triangulaires, hexagonales qui sont étudiées dans la littérature. Il est même également applicable à d’autres types de problèmes, qui incluent en particulier toutes les variantes des codes identifiants. Avec Marwane Bouznif et Myriam Preissmann, nous

avons donné une classe générale de problèmes pour lesquels ce type d'algorithme peut s'appliquer [35, 36]. Ces travaux sont présentés en Section 2.3.

## 2.2.2 Cycles et puissances de cycles

Dans cette section, nous allons discuter des codes identifiants dans les cycles, qui est une question n'ayant été complètement résolue que très récemment, et ayant fait l'objet de six publications par divers auteurs [25, 53, 98, 118, 159, 177].

Le cycle à  $n$  sommets est le graphe, noté  $C_n$ , dont les sommets sont  $\{v_1, \dots, v_n\}$ , et tel que pour tout  $i$ , le sommet  $v_i$  soit voisin de  $v_{i-1}$  et  $v_{i+1}$  (dans toute cette section les indices sont considérés modulo  $n$ ).

Pour un entier  $r \geq 1$  donné, la puissance  $r$  du cycle à  $n$  sommets est le graphe, noté  $C_n^r$ , dont les sommets sont  $\{v_1, \dots, v_n\}$ , et tel que pour tout  $-r \leq j \leq r$ ,  $j \neq i$ , le sommet  $v_i$  est voisin de  $v_{i+j}$  (c'est la fermeture  $r$ -transitive du cycle  $C_n$ ).

### 2.2.2.1 Cas des cycles pairs

Bertrand *et al* s'intéressent aux cycles pairs dans [25] en 2004. Ils montrent que pour tout  $n$  pair,  $n \geq 2r + 4$ , on a :

**Théorème 7** (Cycles pairs). *Soit  $r \geq 1$  et soit  $C_n$  le cycle à  $n \geq 2r + 4$  sommets, avec  $n$  pair. Alors on a*

$$\gamma^{\text{ID}}(C_n^r) = \frac{n}{2}. \quad (2.2)$$

De plus, lorsque  $n = 2r + 2$ , on a

$$\gamma^{\text{ID}}(C_{2r+2}^r) = 2r + 1. \quad (2.3)$$

La preuve de (2.2) est basée sur l'observation suivante, qui est valide pour tout  $r \geq 1$  et pour tout  $n \geq 2r + 2$  ( $n$  pair ou impair).

Pour tout  $i$ , on a  $N[v_i] \Delta N[v_{i+1}] = \{v_{i-r}, v_{i+r+1}\}$  dans  $C_n^r$ . Ainsi, pour tout  $i$ , au moins l'un des deux sommets  $v_{i-r}, v_{i+r+1}$  doit appartenir au code  $r$ -identifiant. En d'autres termes, dès lors qu'un sommet  $v_j$  n'est pas dans le code, alors le sommet  $v_{j+2r+1}$  est nécessairement dans le code. Il y a donc une injection de  $V \setminus C$  dans  $C$  pour tout code  $r$ -identifiant  $C$  de  $C_n$ . Ceci implique en particulier que  $n - |C| \leq |C|$ , c'est-à-dire

$$|C| \geq \frac{n}{2}. \quad (2.4)$$

Dans le cas  $n$  pair, il est par ailleurs facile de montrer que  $\{v_{2k} \mid 1 \leq k \leq \frac{n}{2}\}$  est un code identifiant de  $C_n^r$ , d'où (2.2).

On peut aussi voir (2.4) comme une conséquence de la borne inférieure générale pour les graphes réguliers (Théorème 1). Le Théorème 7 a été également donné par Gimbel *et al* [90].

Pour (2.3), il suffit d'observer que dans le cas  $n = 2r + 1$ , le voisinage fermé de  $v_i$  dans  $C_n^r$  est égal à l'ensemble de tous les sommets de  $C_n^r$  sauf le sommet « antipodal »  $v_{i+r+1}$ . Ceci implique qu'au plus un sommet de  $C_n^r$  peut ne pas appartenir au code (si l'on a deux sommets  $v_i$  et  $v_j$  qui ne sont pas dans le code, alors les sommets  $v_{i+r+1}$  et  $v_{j+r+1}$  ne seraient pas  $r$ -séparés).

Il est par ailleurs aisé de vérifier que n'importe quel sous-ensemble de sommets de cardinalité  $2r + 1$  est un code  $r$ -identifiant de  $C_n$ .

### 2.2.2.2 Cas des cycles impairs et introduction du graphe auxiliaire

En 2006, dans un article publié dans *European Journal of Combinatorics* [98], nous introduisons avec Sylvain Gravier et Ahmed Semri (USTHB, Alger) un graphe auxiliaire  $\tilde{C}_n^r$ , dont les sommets sont  $\{v_1, \dots, v_n\}$ , et tel que, pour tout  $i$ , il y a une arête entre  $v_{i-r}$  et  $v_{i+r+1}$ . Nous avons mentionné au-dessus qu'un code identifiant de  $C_n^r$  doit en particulier séparer les sommets  $v_{i-r}$  et  $v_{i+r+1}$ , pour tout  $i$ . Du point de vue du graphe auxiliaire  $\tilde{C}_n^r$ , ceci signifie qu'un code identifiant de  $C_n^r$  doit être également un transversal de  $\tilde{C}_n^r$ . Rappelons qu'un *transversal* d'un graphe  $G = (V, E)$  est un sous-ensemble de sommets  $T \subseteq V$  tel que, pour toute arête  $uv \in E$ , on a  $T \cap \{u, v\} \neq \emptyset$ .

Cette observation conduit à une borne inférieure améliorant (2.2). En effet, l'analyse de la condition

pour tout  $i$ , au moins l'un des sommets  $v_{i-r}$  et  $v_{i+r+1}$  est dans le code,

en tant que graphe, nous permet d'exploiter la structure de celui-ci et de dériver des résultats plus précis. Plus précisément, le graphe  $\tilde{C}_n^r$  consiste en l'union disjointe de  $\text{pgcd}(n, 2r + 1)$  cycles de  $\frac{n}{\text{pgcd}(n, 2r + 1)}$  sommets. Comme la cardinalité minimum d'un transversal d'un cycle de  $k$  sommets est  $\lceil \frac{k}{2} \rceil$ , alors la cardinalité minimum d'un code identifiant de  $C_n^r$  vérifie :

**Théorème 8** (Borne inférieure pour les cycles). *Soit  $r \geq 1$  et  $n \geq 2r + 2$ . Alors on a*

$$\gamma^{\text{ID}}(C_n^r) \geq \text{pgcd}(n, 2r + 1) \left\lceil \frac{n}{2 \text{pgcd}(n, 2r + 1)} \right\rceil. \quad (2.5)$$

Si, dans le cas pair, cette borne coïncide avec (2.2), elle permet d'avoir une vision beaucoup plus fine du cas  $n$  impair, qui est l'objet principal de notre publication [98]. Nous parvenons, dans ce cas, à confiner la valeur de  $\gamma^{\text{ID}}(C_n^r)$  dans un intervalle d'amplitude au plus  $r$ , en montrant la borne supérieure générale :

**Théorème 9** (Borne supérieure pour les cycles). *Soit  $r \geq 1$  et  $n \geq 2r + 3$ . Alors on a*

$$\gamma^{\text{ID}}(C_n^r) \leq \frac{n+1}{2} + r.$$

La suite de notre article concerne l'étude de cas particuliers. Notons que pour  $n \equiv 0 \pmod{3}$  alors les deux bornes des Théorèmes 8 et 9 coïncident. Pour  $n \not\equiv 0 \pmod{3}$ , alors les deux bornes diffèrent d'une unité. Les valeurs de  $\gamma^{\text{ID}}(C_n)$  sont toutes déterminées par le résultat suivant :

**Théorème 10** (Cycles impairs). *Pour tout  $n \geq 7$  impair, on a*

$$\gamma^{\text{ID}}(C_n) = \frac{n+1}{2} + 1.$$

Comme les deux bornes des Théorèmes 8 et 9 diffèrent d'une unité, ce résultat est obtenu en démontrant qu'il ne peut pas y avoir de code identifiant de cardinalité  $\frac{n+1}{2}$ , et en exhibant un code de cardinalité  $\frac{n+1}{2} + 1$ .

Nous avons découvert récemment que ce résultat avait été également donné par Gimbel *et al* [90].

Un des points délicats de ce graphe auxiliaire  $\tilde{C}_n^r$  que nous avons introduit est la non-inversibilité de la transformation

$$\Phi : \{\text{codes identifiants de } C_n^r\} \rightarrow \{\text{transversaux de } \tilde{C}_n^r\}.$$

En effet, si tout code identifiant de  $C_n^r$  induit par définition un transversal de  $\tilde{C}_n^r$ , il n'est pas vrai qu'un transversal de  $\tilde{C}_n^r$  est toujours un code identifiant de  $C_n^r$ . Dans la suite de notre article, nous donnons alors des conditions sur  $n$  et  $r$  pour que cette transformation  $\Phi$  soit inversible, ou, tout au moins, pour qu'il existe un transversal optimum de  $\tilde{C}_n^r$  qui soit également un code identifiant de  $C_n^r$ .

Nous obtenons, par exemple, dans le cas  $n \geq 3r + 2$  et  $2r + 1$  non premier avec  $n$  :

**Théorème 11.** *Soit  $r \geq 1$  et  $n \geq 3r + 2$  impair tel que  $2r + 1$  non premier avec  $n$ . Alors on a*

$$\gamma^{\text{ID}}(C_n^r) = \text{pgcd}(n, 2r + 1) \left\lceil \frac{n}{2 \text{pgcd}(n, 2r + 1)} \right\rceil.$$

Nous parvenons de plus à déterminer  $\gamma^{\text{ID}}(C_n^r)$  pour certaines valeurs de  $n$  dans le cas où  $2r + 1$  est premier avec  $n$ , comme par exemple :

**Théorème 12.** *Soit  $r \geq 1$  et soit  $4r + 5 \leq n \leq 8r + 1$  impair tel que  $2r + 1$  est premier avec  $n$ . Alors on a*

$$\gamma^{\text{ID}}(C_n^r) = \frac{n+1}{2}.$$

### 2.2.2.3 Fermeture des derniers cas

En 2008, Roberts et Roberts font une analyse détaillée du cas  $r = 2$  et parviennent à déterminer toutes les valeurs de  $\gamma^{\text{ID}}(C_n^2)$  manquantes [159]. Leur analyse est, elle aussi, basée sur l'observation qu'un code identifiant doit en particulier séparer toutes les paires de sommets consécutifs sur le cycle.

Ils définissent et étudient ce qu'ils appellent le *flot de contraintes* associé, et parviennent à disséquer la structure des codes optimaux de  $C_n^2$ . Cette approche est une reformulation de l'idée du transversal sous la forme de contraintes logiques dans le cas particulier  $r = 2$ . Roberts et Roberts citent d'ailleurs notre article et précisent :

*« A similar condition for arbitrary  $r$  corresponds to the idea of a transversal in an auxiliary graph »*

Leur résultat fait en particulier apparaître une régularité de structure modulo 5.

Xu *et al* s'attaquent en 2008 [177], eux aussi, aux cas laissés ouverts dans notre article [98], en particulier le cas où  $2r + 1$  est premier avec  $n$ . Ils parviennent à déterminer la valeur de  $\gamma^{\text{ID}}(C_n^r)$  dans les cas  $n \geq 3r + 2$  et où  $n$  s'écrit  $n = 2m(2r + 1) + 1$  ou  $n = (2m + 1)(2r + 1) + 2r$  pour un entier  $m \geq 1$ . Leur approche réutilise explicitement le transversal du graphe auxiliaire que nous avons défini dans [98].

Les derniers cas sont résolus simultanément et de façon indépendante en 2011, par Junnila et Laihonen d'une part [118], et Chen *et al* d'autre part [53]. Junnila et Laihonen réutilisent l'idée du transversal du graphe auxiliaire, et parviennent à donner des conditions sur ce transversal pour que celui-ci corresponde à un code identifiant pour des valeurs de  $n$  petites par rapport à  $r$ .

Chen *et al* quant à eux reprennent l'approche des flots de contraintes de Roberts et Roberts [159].

### 2.2.3 Hypercubes

Rappelons ici que l'*hypercube* de dimension  $n \geq 1$ , noté  $Q_n$ , est le graphe dont les sommets sont les mots binaires de longueur  $n$ , et tel qu'il existe une arête entre deux mots  $u$  et  $v$  si ceux-ci diffèrent en exactement un bit. La distance de Hamming entre deux mots  $u$  et  $v$ , notée  $d(u, v)$ , est définie comme étant le nombre de bits sur lesquels diffèrent  $u$  et  $v$ . Il est facile de voir que  $d(u, v)$  est égal à la longueur (en nombre d'arêtes) d'un plus court chemin entre  $u$  et  $v$  dans  $Q_n$ .

Dans  $Q_5$ , le mot  $u = 00000$  est par exemple relié au mot  $v = 01011$  par un chemin de trois arêtes, un tel chemin passant par exemple par les mots  $01000$  et  $01001$ .

Notons que l'hypercube ne contient pas de triangles (un *triangle* étant un cycle à trois sommets), et que la cardinalité de l'intersection de deux voisinages fermés de sommets distincts est de 0 ou 2. Ceci implique que la cardinalité de l'intersection de trois voisinages fermés de sommets distincts est égale à 0 ou 1. Ces propriétés sont importantes pour l'identification de sommets.

Les hypercubes sont une structure naturelle pour l'étude des codes identifiants, et ont fait l'objet d'au moins une vingtaine de publications. Les spécialistes de ces questions sont Antoine Lobstein (Télécom ParisTech) et les finlandais de l'université de Turku, autour d'Iiro Honkala et de Tero Laihonen. La thèse de Sanna Ranto [155] concerne d'ailleurs exclusivement les hypercubes. Cette focalisation sur les hypercubes (et les espaces de Hamming en général) provient de leur expertise sur les codes couvrants (rappelons qu'Antoine Lobstein et Iiro Honkala sont deux co-auteurs de l'ouvrage de référence sur les codes couvrants [58]), les hypercubes étant les espaces « naturels » d'investigation pour ces codes.

Historiquement, les premiers travaux sur les codes identifiants ont, en majorité, concerné les hypercubes, ainsi que les grilles infinies. Dès l'article initial de Karpovsky *et al* [120], il est remarqué que, dans le cas des hypercubes, on peut assez facilement construire des codes identifiants à partir de codes couvrants :

**Théorème 13** (Construction de code identifiant à partir de code 2-couvrant).  
Soit  $K$  un code 2-couvrant de  $Q_n$ . Alors l'ensemble des sommets  $C := \{v \mid \exists u \in K \text{ tel que } d(u, v) = 1\}$  est un code identifiant de  $Q_n$ .

En effet, un sommet  $v \in K$  est couvert par l'ensemble de ses voisins, qui, par définition, sont tous dans  $C$ . De plus, ces sommets identifient  $v$  de

façon unique, car deux sommets distincts de  $Q_n$  ont nécessairement deux voisinages distincts.

Un sommet  $v \in C \setminus K$  est clairement couvert (au moins par lui-même). Supposons que  $v$  ait au moins un voisin  $v' \in C$ . Comme il n'y a pas de triangles dans  $Q_n$ , alors  $v$  est séparé de tout sommet distinct de  $v'$ . De plus,  $v$  est séparé de  $v'$ .

En effet, soit  $x$  un voisin de  $v$  appartenant à  $K$ . Clairement, il existe un unique sommet  $w$  à distance 1 de  $v'$  et de  $x$ , et ce sommet sépare  $v$  et  $v'$ . Si  $v$  n'est couvert que par lui-même, alors il est séparé de tout autre sommet de  $C$ .

Reste à vérifier que les sommets qui ne sont ni dans  $C$ , ni dans  $K$ , sont bien identifiés. Soit  $v \notin C \cup K$  un tel sommet. Comme  $K$  est un code 2-couvrant, alors  $v$  est couvert. Si  $v$  est couvert par au moins trois sommets de  $C$ , alors  $v$  est clairement identifié, car deux sommets distincts de  $Q_n$  ont, soit exactement 2, soit exactement 0, voisins communs. Pour la même raison, il est impossible que  $v$  soit couvert par un seul sommet de  $C$ . Ainsi, reste à considérer le cas où  $v$  est couvert par exactement deux sommets de  $x, y \in C$ . Soit  $u$  l'unique sommet de  $Q_n$  distinct de  $v$  à distance 1 de  $x$  et  $y$ , il nous faut montrer que  $u$  et  $v$  sont séparés. Si  $u \in K$ , alors nous avons déjà montré que  $u$  était identifié, donc  $u$  et  $v$  sont séparés. Supposons maintenant par l'absurde que  $u \notin K$ . En ce cas, soit  $w$  un sommet de  $K$  voisin de  $x$ . Les sommets  $w$  et  $v$  ont alors un deuxième voisin commun, qui appartient donc à  $C$  et est un troisième sommet de  $C$  couvrant  $v$ , contradiction.

Il existe également des liens forts avec les codes couvrants multiples.

Un code  $r$ -couvrant est dit  $t$ -multiple,  $t \geq 1$ , si tout sommet est  $r$ -couvert par au moins  $t$  sommets distincts du code. Blass *et al* démontrent par exemple dans [29] qu'un code 1-couvrant 3-multiple de  $Q_n$  est automatiquement un code identifiant de  $Q_n$ .

En effet, comme il n'y a pas de triangle dans  $Q_n$ , alors l'intersection des voisinages de trois sommets distincts est de cardinalité 0 ou 1. Ceci implique en particulier que l'ensemble des sommets couvrant un sommet donné  $u$  le détermine de façon unique.

La littérature sur le sujet comporte de nombreuses constructions de codes identifiants à partir de codes couvrants, voir par exemple [28, 29, 110, 120, 121].

Il existe cependant de nombreuses différences entre la couverture et l'identification de sommets dans  $Q_n$ . En particulier des propriétés de duplication ou de monotonie, qui semblent naturelles pour ces codes, sont triviales pour les codes couvrants et difficiles à montrer pour les codes identifiants — certaines restant d'ailleurs à l'état de conjecture pour ces codes.

Ces propriétés sont basées sur l'observation qu'un hypercube de dimension  $n + 1$  peut être vu comme deux hypercubes de dimension  $n$  reliés par  $n$  arêtes disjointes. Plus précisément,  $Q_{n+1}$  peut être vu comme le produit cartésien du graphe à deux sommets et une arête par  $Q_n$ .

Rappelons ici que le *produit cartésien* de deux graphes  $G_1 = (V_1, E_1)$  et  $G_2 = (V_2, E_2)$ , noté  $G_1 \square G_2$ , est le graphe dont les sommets sont  $V_1 \times V_2$ , et tel qu'il y a une arête entre  $(u_1, u_2)$  et  $(v_1, v_2)$  si, et seulement si, l'une des deux conditions suivantes est remplie :

- $u_1 v_1 \in E_1$  et  $u_2 = v_2$
- $u_2 v_2 \in E_2$  et  $u_1 = v_1$

Pour l'hypercube, d'un point de vue algébrique, cela revient à dire que les mots binaires de longueur  $n + 1$  peuvent être obtenus en concaténant un 0 ou un 1 aux mots binaires de longueur  $n$ .

### 2.2.3.1 Duplication de codes

Dans ce contexte, il est par exemple facile de voir que si  $C$  est un code couvrant de  $Q_n$ , alors en dupliquant  $C$  dans deux copies de  $Q_n$  on obtient un code couvrant de  $Q_{n+1}$ . De la même façon, on peut s'attendre à ce que la duplication d'un code identifiant de  $Q_n$  donne un code identifiant de  $Q_{n+1}$ , ce qui entraînerait l'inégalité

$$\gamma^{\text{ID}}(Q_{n+1}) \leq 2\gamma^{\text{ID}}(Q_n). \quad (2.6)$$

Cet argument de duplication est d'ailleurs utilisé (sans preuve tellement le résultat semble naturel) dans [120] et [121]. Or, l'argument n'est pas valide! En effet, s'il existe dans le code identifiant un sommet couvert uniquement par lui-même, alors on se retrouve avec au moins deux sommets indistinguables dans  $Q_{n+1}$ . Ceci a été observé dans [29], où il est en particulier démontré que cette construction fonctionne seulement si tout sommet du code identifiant  $C$  de  $Q_n$  a au moins un voisin dans  $C$ . Ils démontrent également le résultat plus faible  $\gamma^{\text{ID}}(Q_{n+2}) \leq 4\gamma^{\text{ID}}(Q_n)$ , issu de la quadruplication d'un code identifiant (qui, elle, fonctionne très bien).

À ce jour, l'inégalité (2.6) reste à l'état de conjecture, le résultat le plus fort concernant cette conjecture ayant été démontré récemment par Exoo *et al* [72] :

**Théorème 14** (Quasi-duplication). *Pour tout  $n \geq 1$ , on a*

$$\gamma^{\text{ID}}(Q_{n+1}) \leq \left(2 + \frac{1}{n+1}\right) \gamma^{\text{ID}}(Q_n). \quad (2.7)$$

### 2.2.3.2 Monotonie

Une autre propriété, que l'on pourrait s'attendre à obtenir immédiatement, concerne la monotonie de la fonction  $n \mapsto \gamma^{\text{ID}}(Q_n^r)$  à  $r$  fixé. En effet, il semble naturel de dire que la cardinalité minimum d'un code augmente avec la dimension de l'hypercube. Pour les codes couvrants, c'est trivial, par un simple argument de projection. En effet, soit  $C$  un code  $r$ -couvrant de  $Q_{n+1}$ ,  $r \geq 1$ . Clairement,  $\tilde{C}$  obtenu en tronquant le dernier bit des mots de  $C$ ,  $r$ -couvre  $Q_n$ , et est de cardinalité inférieure ou égale à  $C$ . La cardinalité minimum d'un code  $r$ -couvrant croît donc avec la dimension d'un hypercube.

En ce qui concerne les codes  $r$ -identifiants, tout se passe très différemment. En effet, tronquer le dernier bit d'un code  $r$ -identifiant nous permet d'obtenir un code  $r$ -couvrant de  $Q_n$ . Cependant, il n'y a aucune raison que ce code  $r$ -couvrant soit également un code  $r$ -identifiant, et il est très facile de trouver des exemples où le code obtenu n'est pas identifiant. Blass *et al* conjecturent en 2000 [28] que la fonction  $n \mapsto \gamma^{\text{ID}}(Q_n^r)$  est monotone pour tout  $r \geq 1$  :

**Conjecture 1** (Monotonie dans l'hypercube pour tout  $r$ ). *Soit  $r \geq 1$ , et soit  $n_r$  le plus petit  $n$  tel que  $Q_n$  admette un code  $r$ -identifiant. Alors pour tout  $n \geq n_r$  on a*

$$\gamma^{\text{ID}}(Q_n^r) \leq \gamma^{\text{ID}}(Q_{n+1}^r).$$

De façon assez surprenante, cette conjecture reste une question ouverte aujourd'hui — sauf pour le cas  $r = 1$ , que j'ai démontré dans une note parue dans *Discrete Applied Mathematics* en 2006 [143].

**Théorème 15** (Monotonie dans l'hypercube pour  $r = 1$ ). *Soit  $Q_n$  l'hypercube de dimension  $n \geq 1$ . Alors on a*

$$\gamma^{\text{ID}}(Q_n) \leq \gamma^{\text{ID}}(Q_{n+1}).$$

L'argument est assez simple et consiste à montrer que, si tronquer un code identifiant de  $Q_{n+1}$  ne fournit pas toujours un code identifiant de  $Q_n$ , on peut toujours *altérer* ce code projeté  $\tilde{C}$  de sorte à obtenir un code identifiant de  $Q_n$ , de cardinalité inférieure ou égale à celle de  $\tilde{C}$ . En effet, supposons  $n$  assez grand et soit  $w \in Q_n$  un mot du code tronqué  $\tilde{C}$ . Si  $w$  est séparé de tous les autres sommets de  $Q_n$ , alors nous conservons  $w$  tel quel. Sinon, il existe un unique sommet  $w' \in Q_n$  qui est couvert par exactement les mêmes sommets de  $\tilde{C}$  que  $w$ . L'idée est alors de remplacer  $w$  par un voisin de  $w$  dans  $Q_n$  qui ne soit pas voisin de  $w'$ . Comme  $Q_n$  ne contient pas de triangle, alors c'est toujours possible. Ceci nous permet de séparer  $w$

de  $w'$ , qui restent par ailleurs couverts, car nous travaillons à partir d'une projection d'un code identifiant de  $Q_{n+1}$ .

Pour le cas  $r \geq 2$ , toutes mes tentatives pour adapter cette altération de la projection sont restées infructueuses, et la communauté n'a toujours pas trouvé d'argument permettant de démontrer la Conjecture 1.

## 2.2.4 Produits de cliques

Il existe diverses notions de produits de graphes, permettant de construire un graphe  $G_1 \star G_2$  à partir de deux graphes  $G_1, G_2$  (voir par exemple le *Handbook of Product Graphs* de Hammack, Imrich, et Klavžar, qui est sorti en 2011 [102]). Un des produits classiques est le cas où  $\star = \square$ , le produit cartésien, dont nous avons déjà discuté en Section 2.2.3. En effet, on peut définir de façon récursive l'hypercube de la façon suivante :

$$Q_n = K_2 \square K_2 \square \dots \square K_2 \quad (n - 1 \text{ produits})$$

où  $K_2$  est le graphe complet à deux sommets. On rappelle ici qu'un *graphe complet* (ou *clique*) est un graphe à  $n$  sommets et  $\binom{n}{2}$  arêtes. En d'autres termes, c'est un graphe tel que  $uv$  est une arête pour toute paire de sommets distincts  $u, v$ . Il est en général noté  $K_n$ .

Le produit de cliques de taille  $d \geq 2$  quelconque est appelé *graphe de Hamming*, il est noté  $H(n, d)$  :

$$H(n, d) = K_d \square K_d \square \dots \square K_d \quad (n - 1 \text{ produits}).$$

Notons  $C_n$  le cycle à  $n$  sommets, et  $P_n$  le chemin à  $n$  sommets (obtenu en enlevant une arête à  $C_n$ ). Alors le produit cartésien  $P_k \square P_n$  n'est autre que la grille carrée de dimensions  $k, n$  (voir Section 2.2.1). Le produit cartésien  $C_k \square C_n$  est, quant à lui, le tore de la grille carrée de dimensions  $k, n$ . C'est également l'espace de la métrique dite de Lee en théorie des codes correcteurs [138].

Ainsi le produit cartésien est une opération de graphes importante, qui est à la base de nombre de structures classiques (grilles, tores, graphes de Hamming).

Dans [100], nous étudions avec Sylvain Gravier et Ahmed Semri le produit cartésien de deux cliques de la même taille. Nous déterminons la valeur exacte de  $\gamma^{\text{ID}}(K_n \square K_n)$  pour tout  $n$ . Par ailleurs, nous montrons que pour tout  $n \geq 5$  impair, il y a unicité du code optimal à permutation près.

**Théorème 16** (Produits de cliques). Soit  $H(2, n) = K_n \square K_n$  le graphe de Hamming de dimension 2, avec  $n \geq 2$ . Alors on a

$$\gamma^{\text{ID}}(H(2, n)) = \left\lfloor \frac{3n}{2} \right\rfloor.$$

De plus, lorsque  $n \geq 5$ , alors il y a unicité du code optimal de  $H(2, n)$  (à permutations près).

Pour tout  $n \geq 4$  nous démontrons de plus qu'il existe au moins deux codes optimaux structurellement différents.

On peut remarquer que  $K_n$  n'admet pas de code identifiant, car deux sommets quelconques de  $K_n$  sont jumeaux (ils ont le même voisinage et sont voisins). Le produit cartésien est donc une opération intéressante pour obtenir des graphes admettant des codes identifiants, dans la mesure où il nous fournit un graphe admettant un code identifiant à partir de deux graphes n'en admettant aucun. Cet aspect du produit cartésien sera discuté en Section 2.4.3.3, où l'on utilisera cette propriété pour construire des graphes admettant des codes identifiant des ensembles de sommets (codes  $\ell$ -set-identifiants).

Le produit cartésien de graphes est une occasion supplémentaire d'observer que les codes identifiants sont très différents des codes couvrants, ou dominants. En effet, une conjecture célèbre en domination est la conjecture de Vizing, qui date de la fin des années 1960 [173] :

**Conjecture 2** (Vizing). Soient  $G$  et  $H$  deux graphes quelconques. Alors on a

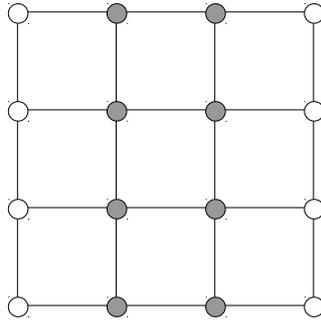
$$\gamma(G \square H) \geq \gamma(G)\gamma(H)$$

où  $\gamma(\cdot)$  dénote la cardinalité minimum d'un dominant d'un graphe.

À ce jour, cette conjecture est toujours ouverte, même si l'on sait que l'inégalité est vraie pour de nombreuses classes de graphes particulières. Si l'on s'intéresse à cette conjecture à une constante multiplicative près, la meilleure inégalité connue à ce jour est celle de Clark et Suen, établie en 2000 [55] :

**Théorème 17** (Clark et Suen). Soient  $G$  et  $H$  deux graphes quelconques. Alors on a

$$\gamma(G \square H) \geq \frac{1}{2}\gamma(G)\gamma(H).$$



**FIGURE 2.8 :** L'ensemble des sommets grisés forme un code identifiant de  $P_4 \square P_4$ , de cardinalité  $8 < 3 \times 3 = \gamma^{\text{ID}}(P_4)\gamma^{\text{ID}}(P_4)$ . Ceci montre que l'équivalent de la conjecture de Vizing pour les codes identifiants n'est pas vraie.

En ce qui concerne l'identification de sommets, l'inégalité de la conjecture de Vizing est clairement fautive. Pour s'en convaincre, considérer  $P_4$ , le chemins de longueur 4. Il est facile de voir que  $\gamma^{\text{ID}}(P_4) = 3$ . Par ailleurs, le code de la Figure 2.8 ci-dessous est un code identifiant de  $P_4 \square P_4$ , de cardinalité  $8 < 3 \times 3 = \gamma^{\text{ID}}(P_4)\gamma^{\text{ID}}(P_4)$ .

À ma connaissance, aucune inégalité du type de celle de Clark et Suen n'est connue pour le cas des codes identifiants. Ceci est une question qui mériterait certainement d'être étudiée.

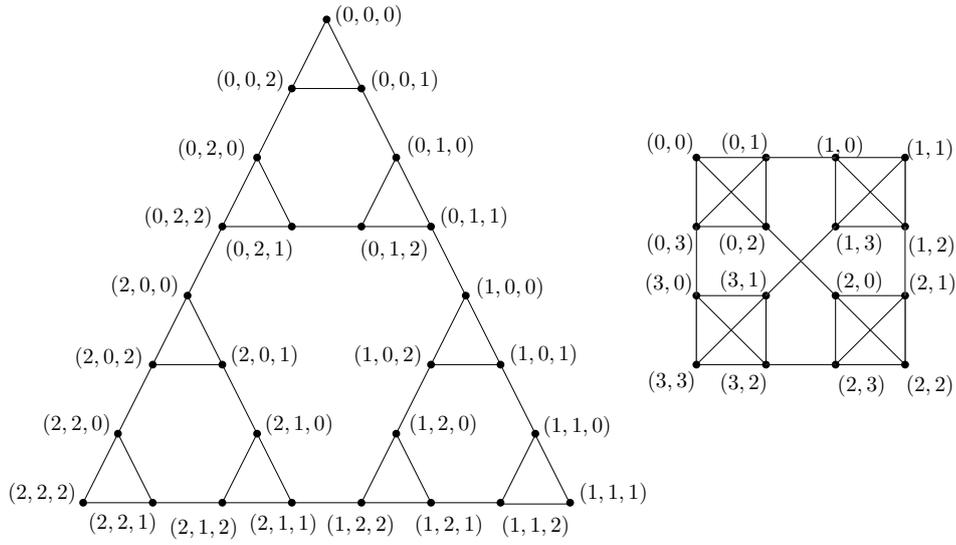
## 2.2.5 Graphes de Sierpiński

Avec Sylvain Gravier, Matjaž Kovše, Michel Mollard, et Aline Parreau [94], nous avons étudié les codes identifiants dans des graphes récursifs particuliers, appelés graphes de Sierpiński.

Un *graphe de Sierpiński*  $S(n, k)$  est obtenu à partir de  $n$  copies d'un graphe de Sierpiński  $S(n-1, k)$ , et en reliant les sommets de ces graphes par un schéma correspondant à celui d'une clique de taille  $k$ , le graphe de Sierpiński  $S(1, k)$  étant défini comme la clique  $K_k$  (voir Figures 2.9 pour une représentation graphique de  $S(3, 3)$  et de  $S(2, 4)$ ).

Formellement, les sommets de  $S(n, k)$  sont les éléments de  $\{0, 1, 2, \dots, k-1\}^n$ , et deux sommets distincts  $(i_1, i_2, \dots, i_n)$  et  $(j_1, j_2, \dots, j_n)$  sont adjacents si, et seulement si, il existe un indice  $h$  dans  $\{1, 2, \dots, n\}$  tel que :

- (i)  $i_t = j_t$  pour tout  $t = 1, \dots, h-1$  ;
- (ii)  $i_h \neq j_h$
- (iii)  $i_t = j_h$  et  $j_t = i_h$  pour tout  $t = h+1, \dots, n$ .



**FIGURE 2.9 :** Les graphes de Sierpiński  $S(3,3)$  et  $S(2,4)$ . Les sommets extrêmes de  $S(3,3)$  sont  $(0,0,0)$ ,  $(1,1,1)$  et  $(2,2,2)$ , et ceux de  $S(2,4)$  sont  $(0,0)$ ,  $(1,1)$ ,  $(2,2)$ , et  $(3,3)$ . Noter que chaque sommet interne  $u$  a exactement un voisin hors de la clique  $K(u)$ , et que l'ensemble d'arêtes  $\{uv \text{ arête de } S(n,k) \mid K(u) \neq K(v)\}$  est un couplage.

Cette structure est définie de façon récursive, mais ne fournit pas un graphe régulier (au sens des degrés). En effet, les sommets de la forme  $(i, i, \dots, i)$  (qui sont appelés *sommets extrêmes*) sont de degré  $k - 1$ , alors que les autres sommets (appelés *sommets internes*) sont de degré  $k$ . Le graphe  $S(n, k)$  a donc  $k$  sommets extrêmes, de degré  $k - 1$ , et  $k^n - k$  sommets internes, de degré  $k$ .

Pour un sommet quelconque  $u = (i_1, i_2, \dots, i_n)$ , les sommets de la forme  $(i_1, i_2, \dots, i_{n-1}, j)$ ,  $0 \leq j \leq k - 1$  induisent une clique à  $k$  sommets, notée  $K(u)$ . On peut observer que le voisinage d'un sommet extrême  $u$  est égal à  $K(u) \setminus \{u\}$ , et qu'un sommet interne  $u$  a un unique voisin qui n'appartient pas à  $K(u)$ . De plus, l'ensemble  $\{uv \text{ arête de } S(n, k) \mid K(u) \neq K(v)\}$  est un couplage.

On rappelle ici qu'un *couplage* d'un graphe  $G = (V, E)$  est un sous-ensemble d'arêtes  $M \subseteq E$  tel que deux arêtes distinctes de  $M$  soient disjointes, c'est-à-dire tel que

$$uv \in M, xy \in M, uv \neq xy \quad \Rightarrow \quad \{u, v\} \cap \{x, y\} = \emptyset.$$

Les graphes de Sierpiński  $S(n, 3)$  correspondent aux graphes des mouvements légaux dans le problème des tours de Hanoï (voir [107]). Introduits par Klavžar et Mulitinović en 1997 [124], ils sont appelés graphes de

Sierpiński dans [125]. Dans l'ouvrage de Lipscomb [136], ils sont appelés graphes de Klavžar-Milutinović. Ces graphes, ainsi que quelques variantes, ont été étudiés dans [114, 123, 169]. Récemment, Beaudou *et al* ont étudié les codes couvrants dans ces graphes [19].

Dans notre manuscrit (actuellement en révision à *Designs Codes and Cryptography*), nous déterminons la cardinalité minimum d'un code identifiant dans ces graphes, ainsi que plusieurs variantes de ces codes. Notre résultat principal est le suivant :

**Théorème 18** (Graphes de Sierpiński). *Soit  $S(n, k)$  le graphe de Sierpiński. Alors on a*

$$\gamma^{\text{ID}}(S(n, k)) = k^{n-1}(k-1).$$

L'idée de la preuve est la suivante.

Tout d'abord, nous montrons que tout code identifiant de  $S(n, k)$  a une cardinalité bornée supérieurement par  $k^{n-1}(k-1)$ . En effet, pour tout sommet interne  $u$  de  $S(n, k)$ , soit  $m(u)$  l'unique voisin de  $u$  qui n'est pas dans  $K(u)$ . Pour chaque clique  $K$  de taille  $k$  de  $S(n, k)$ , considérons alors l'ensemble

$$M(K) = \{m(u) \mid u \text{ est un sommet interne de } K\}.$$

Ces ensembles  $M(K)$  sont disjoints, et il y a exactement  $k^{n-1}$  tels ensembles.

Supposons maintenant qu'il existe deux sommets internes  $u, v$  dans une clique  $K$  de taille  $k$  tels que ni  $m(u)$  ni  $m(v)$  n'appartiennent au code. En ce cas,  $u$  et  $v$  ne seraient pas séparés, ce qui est une contradiction. De plus, si  $K$  contient un sommet extrême, alors tous les sommets de  $M(K)$  doivent être dans le code, afin de séparer ce sommet extrême de tous les autres sommets de  $K$ . Ainsi, pour toute clique  $K$  de taille  $k$ , au moins  $k-1$  sommets de  $M(K)$  sont dans le code.

Comme il y a exactement  $k^{n-1}$  ensembles  $M(K)$ , et que ceux-ci sont disjoints, ceci implique que la cardinalité du code est d'au moins  $k^{n-1}(k-1)$ .

Maintenant, si  $n = 2$ , alors l'ensemble de tous les sommets internes de  $S(2, k)$  est un code identifiant de  $S(2, k)$ , de cardinalité  $k(k-1)$ . On a donc  $\gamma^{\text{ID}}(S(2, k)) = k(k-1)$ .

Dans le cas général  $S(n, k)$ ,  $n \geq 2$ , soit  $C$  l'ensemble des sommets de  $S(n, k)$  tels que, pour tout sous-graphe  $S$  de  $S(n, k)$  isomorphe à  $S(2, k)$ , les sommets de  $S$  qui sont dans  $C$  sont exactement les sommets internes de  $S$ . Il est aisé de vérifier que  $C$  est un code identifiant de  $S(n, k)$ , de cardinalité  $k^{n-1}(k-1)$ . Ainsi on a bien  $\gamma^{\text{ID}}(S(n, k)) = k^{n-1}(k-1)$  pour tout  $k, n$ .

Une caractéristique intéressante de ces graphes est qu'ils sont reliés à une conjecture de 2010 de Foucaud *et al* [79].

**Conjecture 3** (Borne supérieure utilisant le degré maximum). *Soit  $G$  un graphe connexe sans jumeau à  $n$  sommets de degré maximum  $\Delta \geq 3$ . Alors la cardinalité minimum d'un code identifiant de  $G$  vérifie l'inégalité suivante*

$$\gamma^{\text{ID}}(G) \leq \left\lceil n - \frac{n}{\Delta} \right\rceil.$$

Comme  $S(n, k)$  a  $k^n$  sommets, et que le degré maximum de  $S(n, k)$  est  $\Delta = k$ , alors on a

$$\gamma^{\text{ID}}(S(n, k)) = k^{n-1}(k-1) = \frac{n}{\Delta}(\Delta-1) = n - \frac{n}{\Delta}$$

et les graphes de Sierpiński constituent donc une famille infinie de graphes atteignant la borne de la conjecture de Foucaud *et al.*

## 2.3 Aspects algorithmiques

### 2.3.1 État de l'art

Nous avons déjà mentionné que le problème de la détermination de la cardinalité minimum d'un code identifiant dans un graphe était un problème NP-difficile [49]. Plus précisément, considérons le problème de décision suivant, défini pour une certaine classe de graphes  $\mathcal{C}$  :

$(r, \ell)$ -ID CODE DANS  $\mathcal{C}$

Instance : Un graphe  $G \in \mathcal{C}$  et un entier  $k \geq 1$ .

Question : Existe-t-il un code  $(r, \ell)$ -identifiant de  $G$  de cardinalité inférieure ou égale à  $k$  ?

En transposant les conventions de notation définies en Section 2.1, on écrira  $r$ -ID CODE au lieu de  $(r, \leq 1)$ -ID CODE,  $\ell$ -set-ID CODE au lieu de  $(1, \leq \ell)$ -ID CODE, et ID CODE au lieu de  $(1, \leq 1)$ -ID CODE. En cas d'absence de la classe  $\mathcal{C}$  dans l'intitulé du problème de décision, cela signifiera que celle-ci sera simplement égale à celle de tous les graphes.

Charon, Hudry et Lobstein ont montré en 2003 dans un article de *Theoretical Computer Science* [49] :

**Théorème 19** (Complexité en général). *Le problème ID CODE est NP-complet. De plus, pour tout  $r \geq 1$ , le problème  $r$ -ID CODE est NP-complet.*

Leur preuve de complexité utilise une réduction du problème de satisfiabilité 3-SAT vers  $r$ -ID CODE, avec 3-SAT défini comme suit :

## 3-SAT

Instance : Une formule logique  $\Phi(x_1, \dots, x_n)$ , exprimée comme une disjonction de clauses sur l'ensemble des variables  $\{x_1, \dots, x_n\}$ , chaque clause comportant au plus 3 littéraux.

Question : Existe-t-il une assignation des variables  $x_1, \dots, x_n$  telle que la formule  $\Phi(x_1, \dots, x_n)$  soit vraie ?

Notons  $\wedge$  le « et » logique et  $\vee$  le « ou » logique. On rappelle qu'une *clause* est une conjonction de littéraux  $l_1 \wedge l_2 \wedge \dots \wedge l_k$ , un *littéral* étant égal à une variable ou à sa négation (c'est-à-dire que, pour tout  $i$ , il existe  $j$  tel que  $l_i = x_j$  ou  $l_i = \bar{x}_j$ ). Une disjonction de clauses  $c_1, c_2, \dots, c_t$  est par définition  $c_1 \vee c_2 \vee \dots \vee c_t$ .

Leur réduction utilisant un graphe biparti de degré borné, elle démontre en fait que le problème reste NP-complet lorsque restreint aux graphes bipartis de degré borné :

**Théorème 20** (Complexité pour les bipartis). *Soit  $\mathcal{B}_{\Delta \leq 3}$  la classe des graphes bipartis de degré maximum inférieur ou égal à 3. Alors le problème ID CODE DANS  $\mathcal{B}_{\Delta \leq 3}$  est NP-complet.*

On rappelle qu'un graphe  $G = (V, E)$  est dit *biparti* si  $V$  peut être partitionné en  $V_1 \cup V_2$ , de sorte que  $uv \notin E$  pour tout  $u \in V_1$  et  $v \in V_2$ .

C'est une famille de graphes élémentaire, mais importante, dans laquelle nombre de problèmes difficiles en général deviennent polynomiaux. C'est, par exemple, le cas du problème de coloration, ou du problème du transversal minimum. Tout comme le problème de domination, qui est NP-complet en général et reste NP-complet dans les graphes bipartis, y compris de degré borné (c'est un résultat d'Alan Bertossi qui date des années 1980 [23]), le problème de code identifiant est donc lui aussi NP-complet dans les bipartis de degré borné.

Ce résultat a été récemment raffiné par David Auger [7] (Théorème 21 ci-dessous) et Foucaud *et al* [78] (Théorème 22 ci-dessous) :

**Théorème 21** (Graphes planaires de maille quelconque). *Soit  $\mathcal{P}_{\Delta \leq 4}^{m \geq k}$  la classe des graphes planaires de degré maximum inférieur ou égal à 4 et de maille supérieure ou égale à  $k$ . Alors le problème ID CODE DANS  $\mathcal{P}_{\Delta \leq 4}^{m \geq k}$  est NP-complet pour tout  $k \geq 3$ .*

**Théorème 22** (Lines-graphs planaires parfaits 3-colorables). *Soit  $\mathcal{P}_{\Delta \leq 4, LG}^{\text{perf. 3-col}}$  la classe des graphes planaires parfaits 3-colorables de degré maximum inférieur ou égal à 4 qui sont aussi des line-graphs. Alors le problème ID CODE DANS  $\mathcal{P}_{\Delta \leq 4, LG}^{\text{perf. 3-col}}$  est NP-complet.*

Un graphe est dit *planaire* s'il peut être représenté dans le plan  $\mathbb{R}^2$  sans que deux arêtes ne se croisent. Un *line-graph* est un graphe  $G = (V, E)$  tel qu'il existe un graphe  $H = (V_H, E_H)$  tel que  $V = E_H$  et  $e_1 e_2 \in E$  si, et seulement si, il existe  $u \in V_H$  tel que  $e_1 = uv$  et  $e_2 = uv'$ . Un graphe  $G$  est dit *parfait* s'il existe un code couvrant  $C$  de  $G$  tel que tout sommet de  $G$  est couvert par exactement un sommet de  $C$ . La *maille* d'un graphe  $G$  est le plus petit entier  $k$  tel que tout cycle de  $G$  contient au moins  $k$  sommets.

Müller et Sereni ont également étudié le problème dans le cas des graphes d'intersection de disques (appelés également *unit disk graphs*) [154]. Un *graphe d'intersection de disques* est un graphe  $G = (V, E)$  dont les sommets sont  $n$  points du plan  $(x_i, y_i) \in \mathbb{R}^2$ , tels qu'il existe  $r \in \mathbb{R}^+$  de sorte que

$$(x_i, y_i)(x_j, y_j) \in E \iff \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \leq 2r.$$

Autrement dit, le graphe est obtenu à partir de  $n$  disques du plan, de rayons  $r$ , centrés en des points  $(x_i, y_i)$ , de sorte qu'il y a une arête entre deux points distincts si les disques centrés en ces points ont une intersection non vide. Noter que la donnée de  $V$  et de  $E$  suffit à déterminer un rayon  $r$  tel que  $(V, E)$  est un graphe d'intersection de disques de rayon  $r$ . En effet, il suffit de prendre un  $r$  quelconque dans l'intervalle  $[r_{\min}, r_{\max}]$ , avec

$$r_{\min} = \max_{(x_i, y_i)(x_j, y_j) \in E} \frac{1}{2} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$$

et

$$r_{\max} = \min_{(x_i, y_i)(x_j, y_j) \notin E} \frac{1}{2} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}.$$

Ces graphes sont particulièrement pertinents dans le cadre d'applications en télécommunications. On peut, par exemple, imaginer que chaque point représente un émetteur radio, et que le disque représente la portée de cet émetteur.

Leur résultat est que le problème reste NP-complet lorsque restreint aux graphes d'intersection de disques, et ce même dans le cas où ceux-ci sont bipartis et planaires [154].

**Théorème 23** (Graphes d'intersection de disques). *Soit  $\mathcal{U}_{\text{bip.,plan.}}$  la classe des graphes d'intersection de disques qui sont à la fois bipartis et planaires. Alors le problème ID CODE DANS  $\mathcal{U}_{\text{bip.,plan.}}$  est NP-complet.*

La complexité des problèmes identifiant des ensembles de sommets est très mal connue. Auger *et al* ont récemment montré [11] que l'identification de sous-ensembles d'au plus 2 sommets était NP-complète.

**Théorème 24** (Identification d'ensembles d'au plus 2 sommets). *Le problème  $(r, \leq 2)$ -ID CODE est NP-complet pour tout  $r \geq 1$ .*

Leur preuve utilisant un graphe planaire de degré maximum borné par 3, le problème  $(r, \leq 2)$ -ID CODE DANS  $\mathcal{P}_{\Delta \leq 3}$  est également NP-complet (la classe  $\mathcal{P}_{\Delta \leq 3}$  étant celle des graphes planaires de degré maximum borné par 3) :

**Théorème 25** (Identification d'ensembles d'au plus 2 sommets). *Soit  $\mathcal{P}_{\Delta \leq 3}$  la classe des graphes planaires de degré maximum borné par 3. Alors le problème  $(r, \leq 2)$ -ID CODE DANS  $\mathcal{P}_{\Delta \leq 3}$  est NP-complet pour tout  $r \geq 1$ .*

À ce jour, aucun autre résultat n'est connu pour l'identification d'ensembles de sommets. David Auger conjecture dans sa thèse que :

**Conjecture 4** (Complexité de l'identification d'ensembles de sommets). *Le problème  $(r, \leq \ell)$ -ID CODE est NP-complet pour tout  $r \geq 1$  et tout  $\ell \geq 1$ .*

Je ne peux qu'abonder dans le sens de cette conjecture.

Les résultats de Auger *et al* [11] sont basés sur le problème TRANSVERSAL, qui consiste à déterminer un transversal de cardinalité minimum dans un graphe. C'est un problème classique en théorie des graphes (c'est d'ailleurs le problème GT1 du Garey-Johnson [88]).

TRANSVERSAL

Instance : Un graphe  $G = (V, E)$  et un entier  $k \geq 1$ .

Question : Existe-t-il un sous-ensemble  $T \subseteq V$ , de cardinalité  $k$ , tel que  $uv \in E \Rightarrow \{u, v\} \cap T \neq \emptyset$ ?

Étendre les résultats du Théorème 25 en utilisant une réduction de TRANSVERSAL à  $(r, \leq \ell)$ -ID CODE pour tout  $r$  et  $\ell$  nécessiterait beaucoup de travail. En effet, les réductions sont obtenues en substituant un gadget à chaque arête du graphe. Pour  $r = 1$  et  $\ell = 1$ , le gadget inclut un cycle à 9 sommets. Pour  $r = 2, \ell = 1$ , c'est un autre gadget qui est utilisé (nécessitant donc une autre preuve), qui a lui 11 sommets. Pour  $r \geq 3$  et  $\ell = 1$ , un gadget générique est fourni, comportant  $10r^2 + 35r + 18$  sommets. Ce cas nécessite bien entendu lui aussi une autre preuve. Enfin, dans la quatrième preuve, concernant le cas  $r \geq 1$  et  $\ell = 2$ , Auger *et al* utilisent un gadget générique à  $6r + 4$  sommets. La perspective d'une preuve générale utilisant le problème TRANSVERSAL pour  $r \geq 1$  et  $\ell \geq 1$  quelconques semble donc pour l'instant assez lointaine.

L'approximabilité du problème est une question naturelle – quoique récente dans la littérature – que nous avons contribué à introduire avec Ralf Klasing (LaBRI, Bordeaux) et Sylvain Gravier [93]. Cette question est discutée en détail en Section 2.3.5.

## 2.3.2 Arbres orientés

### 2.3.2.1 Généralités sur les arbres

Un *arbre* est un graphe à  $n$  sommets et  $n - 1$  arêtes n'admettant aucun cycle. Un arbre admet nécessairement au moins deux sommets de degré 1, qui sont appelés des *feuilles*. L'unique voisin d'une feuille  $v$  est appelé le *père* de  $v$ .

Un *arbre orienté* est quant à lui un graphe orienté  $G$  tel que, si l'on fait abstraction des orientations des arcs, le graphe non orienté obtenu  $\tilde{G}$  est un arbre. Les arbres font partie des structures les plus simples en théorie des graphes. Un très grand nombre de problèmes combinatoires NP-complets en général sont polynomiaux sur les arbres. C'est par exemple le cas de la domination [56] ou du transversal.

Une caractéristique importante de la classe des arbres est que celle-ci peut être obtenue par ajouts successifs de feuilles à un arbre. Plus précisément, pour tout arbre, il existe une façon de numéroter les sommets de l'arbre  $v_1, \dots, v_n$  de sorte à ce que pour tout  $i \leq n$ , on ait

$$|N(v_i) \cap \{v_1, \dots, v_{i-1}\}| = 1.$$

La technique de preuve standard sur les arbres consiste alors à raisonner localement sur les feuilles de l'arbre, de traiter puis couper celles-ci et de propager l'information par récurrence sur l'arbre plus petit obtenu. C'est ce que Mitchell, Cockayne, et Hedetniemi présentent par exemple à la fin des années 1970 dans leur article, intitulé *Linear algorithms on recursive representations of trees* [142].

On peut voir le fait qu'un grand nombre de problèmes combinatoires NP-complets en général soient polynomiaux sur les arbres comme une conséquence d'un résultat plus récent d'Arnborg, Lagergren, et Seese [5], basé sur une approche utilisant la logique monadique du second ordre. Ne souhaitant pas rentrer dans les détails de cette approche et de la logique monadique du second ordre, on renvoie le lecteur au livre de Courcelle et Engelfriet sur le sujet [60]. Un des résultats forts de cette approche pourrait être formulé ainsi. Il existe un certain langage (au sens de la théorie des langages), appelé **MSOL**, tel que le résultat suivant s'applique :

**Théorème 26** (Problèmes faciles sur les arbres). *Soit  $P$  un problème combinatoire tel que  $P$  puisse être exprimé dans le langage **MSOL**. Alors  $P$  est polynomial lorsque restreint à la classe des arbres.*

Lorsque l'on utilise ce théorème, le fait que le problème de domination, de transversal, ou même de code identifiant, soit polynomial dans les arbres

est immédiat, car il est très facile d'exprimer ces problèmes dans le langage **MSOL** (voir par exemple ma thèse pour le cas des codes identifiants [146, Chapitre 2, page 39]).

Un des inconvénients de l'approche du Théorème 26 est que l'algorithme polynomial associé est, en général, inutilisable en pratique. En effet, la complexité de cet algorithme est de l'ordre de  $O(f(k)g(n))$ , avec :

- $n$  la taille de l'arbre
- $k$  la taille de la formule logique exprimant le problème  $P$  dans le langage **MSOL**
- $f$  une fonction « affreuse » ; typiquement une tour d'exponentielle (c'est-à-dire une exponentielle d'exponentielles d'exponentielles etc., voir par exemple l'article de Frick et Grohe [82])
- $g$  un polynôme (très souvent de degré 1)

Si, en théorie, la complexité d'un tel algorithme est polynomiale en la taille de l'arbre, en pratique, la constante multiplicative cachée  $O(f(k))$  a une valeur telle qu'elle annule l'intérêt pratique de l'algorithme.

Ainsi, si cette approche a le mérite de pouvoir nous assurer qu'un problème donné est polynomial dans les arbres, elle ne rend toutefois pas caduque la recherche d'un algorithme polynomial explicite ayant une complexité ne cachant aucune constante multiplicative « monstrueuse ».

### 2.3.2.2 Codes identifiants dans les arbres

Peter Slater annonçait ainsi à la communauté en 2001 disposer d'un algorithme explicite polynomial pour le problème ID CODE DANS  $\mathcal{A}$ , avec  $\mathcal{A}$  la classe des arbres. Nous avons alors décidé, avec Irène Charon, Antoine Lobstein et Olivier Hudry d'une part (tous de Télécom ParisTech), et avec Sylvain Gravier et Michel Mollard d'autre part (Institut Fourier, Grenoble), de nous intéresser au problème ID CODE DANS  $\vec{\mathcal{A}}$ , où  $\vec{\mathcal{A}}$  est la classe des arbres orientés.

Dans le cas d'un graphe orienté  $G = (V, A)$ , on étend la notion de code identifiant comme suit. On dit qu'un sommet  $u$  couvre un sommet  $v$  si  $(u, v) \in A$ . La notion de couverture est donc dans ce cas asymétrique, car l'on peut avoir  $u$  qui couvre  $v$  sans que  $v$  ne couvre  $u$ . Deux sommets  $u$  et  $v$  sont séparés par  $w$  si celui-ci couvre un seul des deux sommets  $u, v$ . Un code identifiant de  $G$  est alors un sous-ensemble de sommets  $C \subseteq V$  qui couvre tous les sommets et sépare toutes les paires de sommets distincts de  $G$ .

On peut observer que le fait de considérer des graphes orientés ne simplifie pas les choses par rapport au problème de code identifiant. Charon,

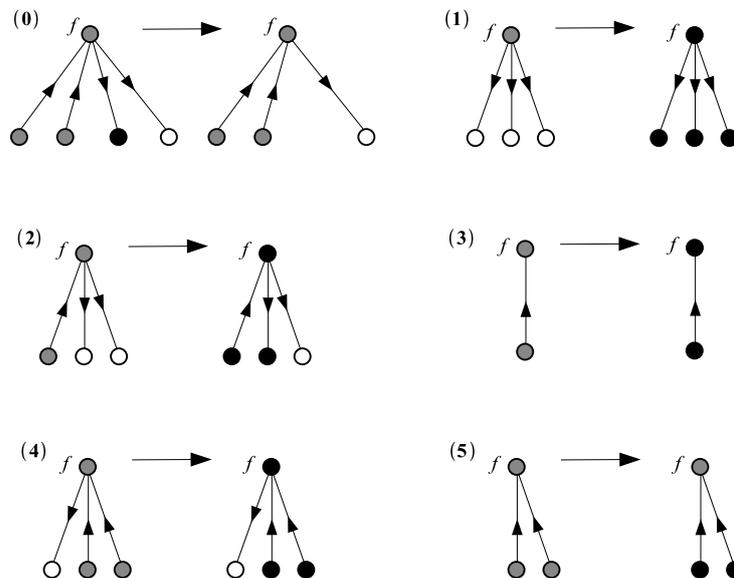
Hudry et Lobstein ont montré [50] que le problème restait NP-complet dans ce cas :

**Théorème 27** (Complexité dans les graphes orientés). *Soit  $\vec{\mathcal{G}}$  la classe des graphes orientés. Alors le problème ID CODE DANS  $\vec{\mathcal{G}}$  est NP-complet.*

Nous donnons dans [43] la description d'un algorithme linéaire déterminant un code identifiant minimum dans un arbre orienté.

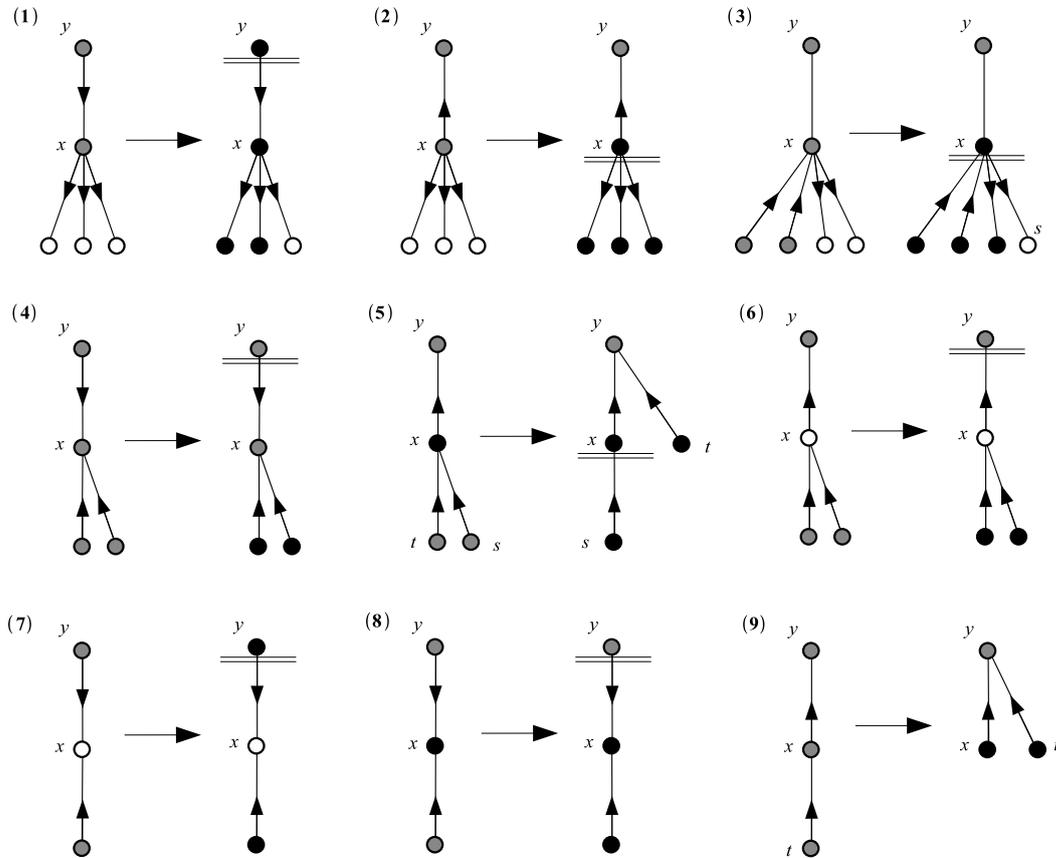
Cet algorithme est basé sur des techniques classiques de récursion dans les arbres, et a nécessité une classification des configurations possibles et des règles de réduction permettant de raisonner par induction. Les Figures 2.10 et 2.11 ci-dessous illustrent l'ensemble des règles de réduction que nous avons étudiées. Selon la configuration des sommets de l'arbre, ces règles nous disent comment étiqueter les sommets de sorte à propager l'information, et comment découper l'arbre pour obtenir un arbre plus petit sur lequel appliquer la récursion.

Quoique élémentaire, la preuve complète de la validité de notre algorithme est relativement fastidieuse, car nous devons considérer un grand nombre de cas et de sous-cas (voir Figures 2.10 et 2.11).



**FIGURE 2.10** : Ensemble des règles de réduction étudiées dans le cas où l'arbre orienté considéré est petit (cas de base de la récursion). Les sommets en noir sont dans le code, ceux en blanc ne sont pas dans le code. Les sommets gris ont un statut indéterminé.

Peter Slater n'ayant finalement jamais publié son algorithme pour les arbres non orientés, David Auger a publié un tel algorithme en 2010 [7].



**FIGURE 2.11 :** Ensemble des règles de réduction étudiées dans le cas où l'arbre orienté considéré est grand (cas général de la récursion). Les sommets en noir sont dans le code, ceux en blanc ne sont pas dans le code. Les sommets gris ont un statut indéterminé. Les doubles barres  $\equiv$  indiquent les points de coupe permettant de raisonner sur un arbre orienté plus petit.

Tout comme notre algorithme pour les arbres orientés, l'algorithme de David Auger utilise une récursion, qui nécessite l'étude d'une vingtaine de sous-cas.

### 2.3.3 Fasciagraphes et rotagraphes

Nous avons déjà discuté en Section 2.2.1 d'un algorithme de programmation dynamique permettant de calculer, en temps constant, les paramètres  $\alpha_0, \dots, \alpha_{T-1}, \nu, T$  d'une formule de la forme

$$\gamma^{\text{ID}}(B_{k,n}) = \alpha_i + \nu \left\lceil \frac{n}{T} \right\rceil \quad \text{si } n \equiv i \pmod{T} \quad (2.8)$$

pourvu que le paramètre  $k$  soit fixé. En pratique, cet algorithme est utili-

sable pour de petites valeurs de  $k$ , et nous a permis de déterminer des valeurs de  $\gamma^{\text{ID}}(B_{k,n})$  et de  $\gamma^{\text{ID}}(B_k)$  qui étaient inconnues jusqu'à présent (voir Section 2.2.1).

Dans l'article où nous présentons ce résultat [62], avec Marc Daniel et Sylvain Gravier, nous montrons que notre algorithme est valide pour une large classe de graphes, qui sont les fasciagraphes et les rotagraphes. Les fasciagraphes et les rotagraphes, respectivement, généralisent les notions de grilles et de grilles toroïdales. Ils englobent notamment toutes les notions de grilles classiques dans les réseaux carrés, triangulaires, hexagonaux, et royaux.

Un *fasciagraphe*  $\phi(M, n)$  est défini à partir d'un graphe mixte  $M$  et d'un entier  $n \geq 2$ . La notion de graphe mixte unifie celles de graphe orienté et de graphe non orienté. Un *graphe mixte*  $M$  est défini sur un ensemble de sommets  $V$ , il possède à la fois des arêtes (éléments de  $E$ ) et des arcs (éléments de  $A$ ). On le note  $M = (V, E, A)$ .

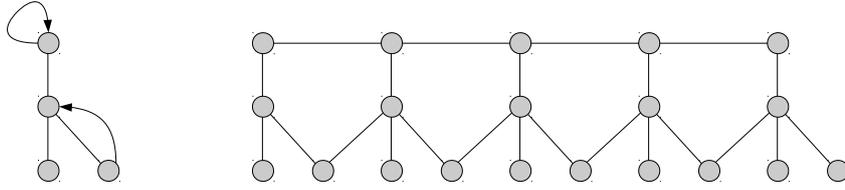
Étant donné un graphe mixte  $M = (V, E, A)$ , le fasciagraphe  $\phi(M, n)$  a pour ensemble de sommets  $V \times \{1, \dots, n\}$ . Pour toute arête  $uv \in E$  de  $M$ , et pour tout  $i = 1, \dots, n$ ,  $(u, i)(v, i)$  est une arête de  $\phi(M, n)$ . De plus, pour tout arc  $(u, v) \in A$  de  $M$ , et pour tout  $i = 1, \dots, n - 1$ ,  $(u, i)(v, i + 1)$  est une arête de  $\phi(M, n)$ .

Le sous-graphe de  $\phi(M, n)$  induit par l'ensemble  $\{(u, i) \mid u \in V\}$  est appelé la  $i$ -ème *fibre* de  $\phi(M, n)$ , notée  $M_i$ , qui n'est autre que le graphe  $(V, E)$ . On peut voir le fasciagraphe  $\phi(M, n)$  comme  $n$  copies ordonnées de  $(V, E)$ , telles que deux copies consécutives sont reliées par un ensemble d'arêtes défini par  $A$  (voir Figure 2.12). C'est en ce sens que les fasciagraphes généralisent les grilles. On peut par exemple voir la grille carrée finie  $B_{k,n}$  comme le fasciagraphe  $\phi(M, n)$ , avec  $M = (V, E, A)$  défini comme suit :

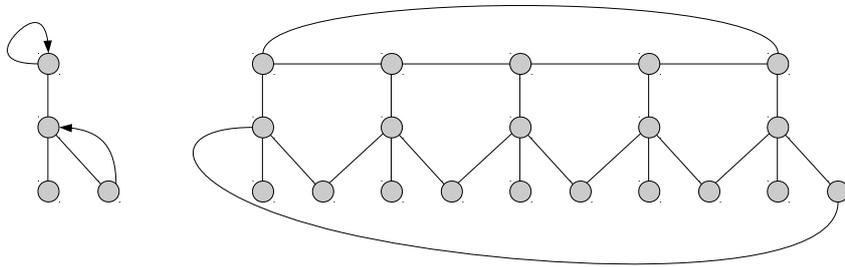
- $V = \{v_1, \dots, v_k\}$
- $E = \{v_i v_{i+1} \mid i = 1, \dots, n - 1\}$
- $A = \{(v_i, v_i) \mid i = 1, \dots, n\}$

Un *rotagraphe*  $\rho(M, n)$  est quant à lui défini comme le fasciagraphe  $\phi(M, n)$  auquel on ajoute l'ensemble d'arêtes  $\{(u, n)(v, 1) \mid (u, v) \in A\}$ . Autrement dit, c'est un fasciagraphe auquel on a ajouté les arêtes qui font reboucler la  $n$ -ème fibre de  $\phi(M, n)$  sur sa première fibre (voir Figure 2.13). Les rotagraphes généralisent donc les grilles toroïdales.

Pour l'algorithme décrit en Section 2.2.1, il n'est pas plus difficile de considérer le problème dans un fasciagraphe ou un rotagraphe quelconque que dans une grille carrée finie  $B_{k,n}$ . Considérons par exemple le cas des fasciagraphes.



**FIGURE 2.12 :** Exemple de fasciagraphe. À gauche, un graphe mixte  $M = (V, E, A)$ . À droite, le fasciagraphe  $\phi(M, 5)$ . On peut observer que ce fasciagraphe est obtenu en prenant 5 copies de  $(V, E)$ , qui sont ordonnées, et telles que deux copies consécutives sont reliées par un schéma de connexion correspondant aux arcs de  $M$ .



**FIGURE 2.13 :** Exemple de rotagraph. À gauche, un graphe mixte  $M = (V, E, A)$ . À droite, le rotagraph  $\rho(M, 5)$ .

On construit de façon analogue un graphe orienté auxiliaire  $\mathcal{G}(V, E)$ , dont les sommets sont l'ensemble des fasciagraphe  $\phi(M, 5)$  munis d'un précode. Un précode d'un fasciagraphe  $\phi(M, 5)$  est défini comme un sous-ensemble de sommets de  $\phi(M, 5)$  qui couvre et sépare tous les sommets appartenant aux deuxième, troisième, et quatrième fibres de  $\phi(M, 5)$ . Il y a également deux sommets additionnels début et fin. Un arc sort d'un sommet  $(\phi(M, 5), C)$  pour aller vers un sommet  $(\phi(M, 5), C')$  dès lors qu'une condition de superposabilité de  $(\phi(M, 5), C)$  et  $(\phi(M, 5), C')$  est remplie. Un chemin orienté de  $\mathcal{G}(V, E)$  empruntant  $t$  arcs correspond alors à un fasciagraphe  $\phi(M, 5 + t)$  muni d'un précode.

Par un raisonnement similaire au cas particulier des grilles carrées finies, on parvient alors à dériver en temps constant (pourvu que  $|V|$  soit borné par une constante) une formule analogue à (2.8), qui permet de déterminer  $\gamma^{\text{ID}}(\phi(M, n))$  pour tout  $n \geq 2$ .

Dans le cas des rotagraphes, on construit un graphe auxiliaire n'ayant pour sommets que les fasciagraphe  $\phi(M, 5)$  munis d'un précode (il n'y a pas besoin de sommet début et fin). On se ramène alors à la recherche d'un cycle orienté à  $n$  sommets dans ce graphe auxiliaire. On obtient en temps constant (pourvu que  $|V|$  soit borné par une constante) une formule

analogue à (2.8), qui permet alors de déterminer  $\gamma^{\text{ID}}(\rho(M, n))$  pour tout  $n \geq 2$ .

### 2.3.4 Généralisation à d'autres problèmes de l'algorithme pour les fasciagraphes et les rotagraphes

La généralisation de cet algorithme à une plus large classe de problèmes fait l'objet de la thèse de Marwane Bouznif, que je co-encadre avec Myriam Preissmann. Nous avons donné dans [35] (accepté à *Discrete Mathematics*) et [36] (soumis) un cadre théorique permettant de démontrer que pour une large classe de problèmes, qui inclut nombre de problèmes classiques d'optimisation combinatoire (coloration, domination, stable, etc.), il est possible de construire un algorithme de programmation dynamique permettant d'obtenir en temps constant une formule analogue à (2.8) décrivant la valeur optimale de la fonction à optimiser pour le problème considéré.

L'algorithme est basé sur la notion de pseudo-localité, introduite dans [35], et inspirée d'une notion de localité définie par Klavžar et Vesel en 2003 [126]. Informellement, une propriété de graphe est dite pseudo-d-locale si la vérification de cette propriété peut se réduire à la vérification de la propriété sur des sous-graphes de « taille »  $d$ .

#### 2.3.4.1 Propriétés locales

Un  $q$ -étiquetage d'un graphe est une fonction de ses sommets (et éventuellement de ses arêtes) dans  $\{1, \dots, q\}$ . L'ensemble de tous les  $q$ -étiquetages d'un graphe donné  $G$  est noté  $\mathcal{F}_q(G)$ . On dit qu'une propriété  $P$  de graphes est une  $q$ -propriété si elle peut être définie par l'existence d'un  $q$ -étiquetage du graphe vérifiant un certain ensemble de conditions  $C$ .

Le problème de  $k$ -coloration est par exemple clairement une  $k$ -propriété, l'étiquette d'un sommet étant égal à sa couleur, et l'ensemble de conditions  $C$  correspondant à l'ensemble des contraintes

$$uv \text{ voisins} \Rightarrow \text{les étiquettes de } u \text{ et } v \text{ sont différentes.}$$

Il en est de même pour le problème de domination, et, de façon générale, pour un grand nombre de problèmes combinatoires correspondant à l'existence d'un sous-ensemble  $X$  de sommets (ou d'arêtes) vérifiant certaines propriétés. En effet, en ce cas l'étiquetage consiste, par exemple, à attribuer l'étiquette 1 aux sommets de  $X$ , et l'étiquette 2 aux autres sommets.

Pour une  $q$ -propriété  $P$  de graphes, notons  $P(\mathcal{F}_q(G))$  l'ensemble des  $q$ -étiquetages de  $G$  vérifiant l'ensemble de conditions  $C$ . Pour simplifier

l'écriture, on dira dans la suite qu'un  $q$ -étiquetage vérifie une propriété  $P$  si celui-ci vérifie l'ensemble des conditions  $\mathbf{C}$  associé à  $P$ .

Étant donné un  $q$ -étiquetage  $f \in \mathcal{F}_q(\phi(M, n)) \cup \mathcal{F}_q(\rho(M, n))$ , et deux entiers  $i \geq 1$  et  $k \geq 1$ , on définit  $f_{i,k}$  comme la trace de  $f$  sur le sous-graphe de  $\rho(M, n)$  induit par les  $k$  fibres consécutives  $M_i, M_{i+1}, \dots, M_{i+k-1}$  (les indices étant pris modulo  $n$  dans le cas d'un rotagraphe). Le  $q$ -étiquetage  $f_{i,k}$  est donc en particulier un  $q$ -étiquetage de  $\mathcal{F}_q(\phi(M, k))$ , il est appelé *étiquetage partiel*.

### 2.3.4.2 Cas des problèmes de décision dans les rotagraphes

Intéressons-nous ici au cas le plus simple, qui est le problème général de déterminer si un rotagraphe satisfait une  $q$ -propriété  $P$  donnée (problème de décision).

**Définition 5** (Pseudo-localité dans les rotagraphes). *Une  $q$ -propriété  $P$  de rotagraphe est dite pseudo-d-locale s'il existe une propriété  $P^{\text{loc}}$  sur les fasciagraphes  $\phi(M, d)$  telles que :*

$$f \in P(\mathcal{F}_q(\rho(M, n))) \Leftrightarrow f_{i,d} \in P^{\text{loc}}(\mathcal{F}_q(\phi(M, d))) \quad \forall i = 1, \dots, n$$

La propriété  $P^{\text{loc}}$  est donc une propriété locale, telle qu'un  $q$ -étiquetage d'un rotagraphe  $\rho(M, n)$  vérifie  $P$  si et seulement la propriété locale  $P^{\text{loc}}$  est vérifiée sur tous étiquetages partiels  $f_{i,k}$  des sous-graphes de  $\rho(M, n)$  induits par les  $k$  fibres consécutives  $M_i, M_{i+1}, \dots, M_{i+k-1}$ .

Ceci nous permet de construire de nouveau un graphe auxiliaire orienté  $\mathcal{G}(V, E, A)$ , dont les sommets sont l'ensemble des fasciagraphes  $\phi(M, d)$  munis d'un  $q$ -étiquetage partiel  $\tilde{f}$  vérifiant la propriété locale  $P^{\text{loc}}$ . Un arc sort d'un sommet  $(\phi(M, d), \tilde{f})$  pour aller vers un sommet  $(\phi(M, d), \tilde{f}')$  dès lors qu'une condition de superposabilité de  $(\phi(M, d), \tilde{f})$  et  $(\phi(M, d), \tilde{f}')$  est remplie. Un cycle orienté de  $\mathcal{G}(V, E, A)$  empruntant  $n$  arcs correspond alors à un rotagraphe  $\rho(M, n)$  muni d'un  $q$ -étiquetage vérifiant la propriété  $P$ .

On se ramène donc à déterminer l'existence d'un cycle orienté ayant un nombre d'arcs prescrit. Si la taille du graphe mixte  $M$  est bornée, ceci peut être fait en temps constant, et nous permet de déterminer une formule qui nous dit, pour tout  $n$ , si  $\rho(M, n)$  vérifie la propriété  $P$ .

### 2.3.4.3 Cas général

Lorsque nous nous intéressons à des problèmes d'optimisation, il nous faut intégrer dans le modèle une fonction  $\omega$  à optimiser. Lorsque cette fonction  $\omega$  présente des bonnes propriétés de décomposabilité (qui ne seront

pas présentées ici, voir [36]), alors il est possible de pondérer les arcs du graphe auxiliaire orienté  $\mathcal{G}(V, E, A)$  d'une façon telle que la longueur d'un cycle orienté de  $\mathcal{G}(V, E, A)$  corresponde à la valeur de la fonction  $\omega$ .

Typiquement, toutes les fonctions objectif du type « cardinalité de  $X$  », avec  $X$  sous-ensemble de sommets et/ou arêtes, vérifient de bonnes propriétés de décomposabilité, la cardinalité d'un sous-ensemble  $X$  d'un ensemble  $\Omega$  pouvant en effet être calculée comme la somme des cardinalités de la trace de  $X$  sur une partition quelconque de l'ensemble  $\Omega$ .

Dans le cas des fasciagraphes on doit tenir compte des effets de bord. Au lieu de réduire la vérification de la propriété  $P$  à la vérification d'une seule propriété locale, nous avons besoin de trois propriétés locales : une pour le début, une pour le milieu, et une pour la fin du fasciagraphe (voir Figure 2.14).

**Définition 6** (Pseudo-localité dans les fasciagraphes). *Une  $q$ -propriété  $P$  de fasciagraphe est dite pseudo- $d$ -locale s'il existe trois propriétés  $P_1, P_2, P_3$  sur les fasciagraphes  $\phi(M, d)$  telles que :*

$$f \in P(\mathcal{F}_q(\phi(M, n))) \Leftrightarrow \begin{cases} f_{1,d} \in P_1(\mathcal{F}_q(\phi(M, d))) \\ f_{i,d} \in P_2(\mathcal{F}_q(\phi(M, d))) \quad \forall i = 2, \dots, n-d \\ f_{n-d+1,d} \in P_3(\mathcal{F}_q(\phi(M, d))) \end{cases}$$

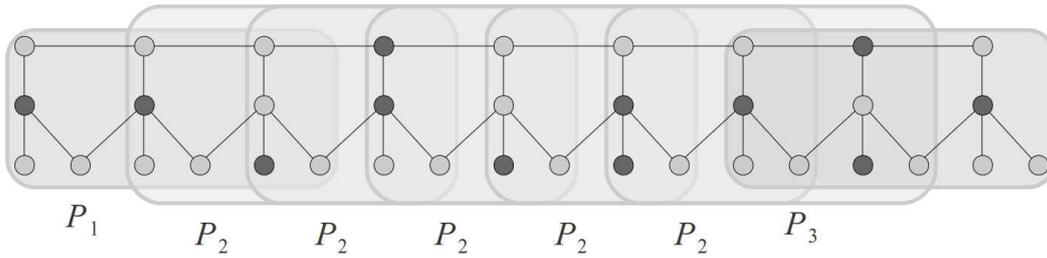
De même nous pouvons intégrer à ce modèle l'optimisation de fonctions objectif satisfaisant de bonnes propriétés de décomposabilité.

Dans ce cadre théorique, nous pouvons alors énoncer des théorèmes de ce type :

**Théorème 28** (Problèmes faciles dans les fasciagraphes et les rotagraphes). *Soit  $P$  un problème d'optimisation pseudo- $d$ -local dans la classe des fasciagraphes (resp. des rotagraphes), dont la fonction objectif  $\omega$  admet de bonnes propriétés de décomposabilité. Soit  $M$  un graphe mixte donné, et soit  $\mathbf{F}(M) = \{\phi(M, n) \mid n \geq 2\}$  l'ensemble des fasciagraphes sur  $M$  (resp.  $\mathbf{R}(M) = \{\rho(M, n) \mid n \geq 3\}$  l'ensemble des rotagraphes sur  $M$ ). Alors on peut déterminer en temps constant les paramètres  $T, \nu, \alpha_0, \dots, \alpha_{T-1}$  d'une formule  $\psi$ , de la forme*

$$\psi(n) = \alpha_i + \nu \left\lceil \frac{n}{T} \right\rceil \quad \text{si } n \equiv i \pmod{T}$$

qui est telle que pour tout  $n$  on a  $\psi(n) = \omega(\phi(M, n))$  (resp.  $\psi(n) = \omega(\rho(M, n))$ ).



**FIGURE 2.14 :** Exemple de 2-propriété pseudo-3-locale dans un fasciagraphe  $\phi(M, 9)$ . Les sommets foncés sont ceux étiquetés 1, les autres sont étiquetés 2. L'étiquetage partiel du sous-graphe  $\phi_1$  de  $\phi(M, 9)$  induit par ses trois premières fibres doit vérifier une certaine propriété  $P_1$ , qui s'énonce comme suit : tout sommet étiqueté 2 de la première ou deuxième fibre de  $\phi_1$  est voisin d'au moins un sommet étiqueté 1. L'étiquetage partiel du sous-graphe  $\phi_7$  de  $\phi(M, 9)$  induit par ses trois dernières fibres doit vérifier une certaine propriété  $P_3$ , qui s'énonce comme suit : tout sommet étiqueté 2 de la deuxième ou troisième fibre de  $\phi_7$  est voisin d'au moins un sommet étiqueté 1. Enfin, pour  $k = 2, 3, 4, 5, 6$ , soit  $\phi_k$  le sous-graphe de  $\phi(M, 9)$  induit par les fibres  $k, k + 1$  et  $k + 2$ . Alors l'étiquetage partiel de  $\phi_k$  doit vérifier une propriété  $P_2$ , qui s'énonce comme suit : tout sommet étiqueté 2 de la deuxième fibre de  $\phi_k$  est voisin d'au moins un sommet étiqueté 1. La propriété  $P$  que doit vérifier cet étiquetage peut donc être décrite comme suit : tout sommet étiqueté 2 est voisin d'au moins un sommet étiqueté 1. En d'autres termes, l'ensemble des sommets étiquetés 1 est un dominant (ou code couvrant) du fasciagraphe  $\phi(M, 9)$ . La propriété de domination est donc une 2-propriété pseudo-3-locale.

### 2.3.5 Approximabilité du problème

Dans cette section, on travaille dans la classe NPO des problèmes d'optimisation. Le problème de la détermination de la cardinalité minimum d'un code identifiant dans un graphe appartenant à une classe de graphes  $\mathcal{C}$  est noté comme suit :

MIN ID CODE DANS  $\mathcal{C}$

Instance : Un graphe  $G \in \mathcal{C}$  admettant un code identifiant.

Solution : Un code identifiant  $C$  de  $G$ .

Objectif : La cardinalité minimum de  $C$ .

#### 2.3.5.1 Généralités sur les algorithmes d'approximation

Soit  $P$  un problème d'optimisation dont la fonction objectif est  $\omega$ . Sans perte de généralité, supposons que l'on souhaite minimiser  $\omega$  (ce sera

d'ailleurs le cas pour tous les problèmes dont nous allons discuter dans cette section).

Un algorithme  $\mathcal{A}$  est dit un *algorithme d'approximation*, de *garantie*  $c(I)$ , pour le problème  $P$ , si c'est un algorithme qui retourne, pour toute instance  $I$  de  $P$ , et en temps polynomial, une solution réalisable de  $P$  pour laquelle la valeur  $\omega_{\mathcal{A}}(I)$  de la fonction objectif vérifie

$$\omega_{\mathcal{A}}(I) \leq c(I)\omega^*(I),$$

avec  $\omega^*(I)$  l'optimum de la fonction objectif pour l'instance  $I$ . En particulier, la garantie vérifie l'inégalité  $c(I) \geq 1$  pour tout  $I$ . Lorsque celle-ci est une constante, alors on parle d'algorithme d'approximation à *facteur constant*.

La question de l'approximabilité d'un problème d'optimisation NP-difficile est fort légitime, les algorithmes d'approximation fournissant des classes d'algorithmes utiles pour la résolution par ordinateur de problèmes NP-difficiles (surtout lorsque le facteur d'approximation est constant et « petit »). Leur caractéristique essentielle est de fournir des solutions quasi-optimales en un temps raisonnable.

La sous-classe  $APX \subseteq NPO$  est la classe des problèmes admettant un algorithme d'approximation à facteur constant. De la même façon que pour la classe NP, on peut définir des notions de réduction polynomiale de problèmes dans NPO, qui préservent l'approximabilité. Un problème donné est alors dit *APX-difficile* si tout problème de la classe APX peut s'y réduire. De façon analogue à la classe NP, un problème de la classe APX qui est APX-difficile est dit *APX-complet*.

Lorsqu'un problème est APX-complet, alors il existe une constante  $c_{\text{lim}}$  telle qu'aucun algorithme d'approximation de garantie  $c < c_{\text{lim}}$  ne peut exister pour ce problème, à moins d'avoir  $P=NP$ .

On renvoie le lecteur à l'ouvrage d'Ausiello *et al* [16] pour les détails sur ces questions.

### 2.3.5.2 Cas des codes identifiants

Soit  $\mathcal{C}$  une classe de graphes. Notons le problème de domination restreint à la classe  $\mathcal{C}$  comme suit :

MIN DOM DANS  $\mathcal{C}$

Instance : Un graphe  $G \in \mathcal{C}$ .

Solution : Un dominant  $D$  de  $G$ .

Objectif : La cardinalité minimum de  $D$ .

Alimonti et Kann ont montré en 2000 [4] que le problème MIN DOM DANS  $\mathbf{G}_{\Delta \leq B}$  était APX-complet pour  $B = 3$ , avec  $\mathbf{G}_{\Delta \leq B}$  la classe des graphes de degré maximum borné par  $B$ . Avec Sylvain Gravier et Ralf Klasing (LaBRI, Bordeaux), nous avons démontré en 2008 [93] que l'on pouvait utiliser ce résultat pour montrer que le problème MIN ID CODE DANS  $\mathbf{G}_{\Delta \leq 8}$  était lui aussi APX-difficile. Ce résultat utilise le concept de L-réduction introduit par Papadimitriou et Yannakakis [153]. Il utilise une réduction intermédiaire utilisant un troisième problème, dont nous montrons également qu'il est APX-difficile au passage.

**Théorème 29** (Résultat négatif d'approximabilité). *Soit  $\mathbf{G}_{\Delta \leq B}$  la classe des graphes de degré maximum borné par  $B$ . Pour tout  $B \geq 8$ , le problème MIN ID CODE DANS  $\mathbf{G}_{\Delta \leq 8}$  est APX-difficile.*

Une conséquence de ce résultat est que le problème général MIN ID CODE est APX-difficile.

Par ailleurs, nous montrons que le problème MIN ID CODE DANS  $\mathbf{G}_{\Delta \leq B}$  est dans APX pour tout  $B$ . Ce résultat est obtenu en raisonnant par rapport au problème suivant :

MIN B-SET COVER

Instance : Un ensemble  $S$ , une famille  $\mathcal{F}$  de sous-ensembles de  $S$  telle que chaque élément de  $\mathcal{F}$  contient au plus  $B$  éléments de  $S$ .

Solution : Une sous-famille  $C \subseteq \mathcal{F}$ , telle que chaque élément de  $S$  soit contenu dans au moins un élément de  $C$ .

Objectif : La cardinalité minimum de  $C$ .

Le cas général où les éléments de  $\mathcal{F}$  sont de cardinalité quelconque est noté MIN SET COVER. C'est un problème classique de couverture d'ensemble (c'est d'ailleurs le problème SP5 dans le Garey-Johnson [88]), qui est NP-difficile, et admet un algorithme d'approximation de garantie non constante  $(1 + \ln |S|)$  [115]. Dans le cas où les sous-ensembles sont de cardinalité bornée, le problème MIN B-SET COVER admet un algorithme d'approximation de garantie  $(1 + \ln B)$  [115].

Le lien avec le problème MIN ID CODE est le suivant. Considérons un graphe  $G = (V, E)$  dans lequel on cherche un code identifiant. Associons à ce graphe une instance de MIN SET COVER en prenant  $S$  comme l'union disjointe de  $S_1$  et  $S_2$ , avec  $S_1 = V$  et

$$S_2 = \{\{u, v\} \subseteq V \mid u \neq v \text{ et } N[u] \cap N[v] \neq \emptyset\}.$$

La famille  $\mathcal{F}$  de sous-ensembles de  $S$  considérée est alors construite de la façon suivante. Pour tout sommet  $z \in V$ , le sous-ensemble  $A_z \subseteq S$  appar-

tient à  $\mathcal{F}$ , avec

$$v \in A_z \Leftrightarrow v \in S_1 \text{ et } z \in N[v]$$

et

$$\{u, v\} \in A_z \Leftrightarrow \{u, v\} \in S_2 \text{ et } z \in N[u] \Delta N[v].$$

Ainsi, un sommet  $v \in V$  est dans  $A_z$  si et seulement si  $z$  couvre  $v$ . De même, une paire de sommets  $\{u, v\}$  est dans  $A_z$  si, et seulement si,  $z$  sépare les sommets  $u$  et  $v$ . Seules les paires de sommets « proches » sont à considérer. En effet, dès lors que  $u$  et  $v$  sont tels que  $N[u] \cap N[v] = \emptyset$ , alors  $u$  et  $v$  sont séparés par un code si, et seulement si, ceux-ci sont couverts par celui-ci.

Par définition, un sous-ensemble de sommets  $C \subseteq V$  est donc un code identifiant de  $G$  si, et seulement si, l'ensemble  $\tilde{C} = \{A_z \mid z \in C\}$  est une solution du problème MIN SET COVER. Par ailleurs, on a  $|C| = |\tilde{C}|$ . Comme on a  $|S| \leq |V|^2$ , ceci montre que le problème MIN ID CODE admet un algorithme d'approximation de facteur  $(1 + 2 \ln |V|)$ .

Supposons maintenant que les sommets du graphe  $G$  soient de degré borné par  $B$ . En ce cas, chaque élément de  $\mathcal{F}$  contient au plus  $B + 1$  éléments de  $S_1$ , et  $B^2(B - 1)$  éléments de  $S_2$ . Ainsi, l'instance associée de MIN SET COVER est telle que chaque élément de  $\mathcal{F}$  contient au plus  $B^3 - B^2 + B + 1$  éléments de  $S$ . Autrement dit, c'est une instance de MIN  $(B^3 - B^2 + B + 1)$ -SET COVER, qui admet donc un algorithme d'approximation de garantie constante  $1 + \ln(B^3 - B^2 + B + 1) \leq 1 + 3 \ln B$ .

Certains des résultats de notre article [93] ont été obtenus indépendamment par Jukka Suomela [168] ainsi que Moshe Laifenfeld et Ari Trachtenberg [130].

Laifenfeld et Trachtenberg ont, de plus, démontré un résultat complémentaire dans leur article. Une des conséquences du Théorème 29 est que, si l'on fait l'hypothèse  $P \neq NP$ , alors il existe une constante  $c_{\text{lim}} > 1$  telle que le problème MIN ID CODE n'admet pas d'algorithme d'approximation de garantie  $c < c_{\text{lim}}$ .

Sous une hypothèse plus forte que  $P \neq NP$ , ils montrent en effet que pour tout  $\varepsilon > 0$ , le problème MIN ID CODE n'admet pas d'algorithme d'approximation de garantie  $c(n) = (1 - \varepsilon) \ln n$  (avec  $n$  le nombre de sommets du graphe).

L'hypothèse plus forte utilisée est  $NP \not\subseteq \text{DTIME}(n^{O(\log \log n)})$ , où  $\text{DTIME}(n^{O(\log \log n)})$  est la classe de problèmes de décision solvables en temps borné par  $n^{O(\log \log n)}$ .

Ils obtiennent ce résultat en utilisant le lien avec le problème MIN SET COVER et un résultat similaire de la littérature sur ce problème [74]. Ce

résultat a été également présenté de façon indépendante par Xiao, Hadjicostis, et Thulasiraman [176].

## 2.4 Questions structurelles

Les questions structurelles qui sont présentées dans cette section tournent essentiellement autour des bornes générales sur la cardinalité d'un code (voir aussi la Proposition 1) :

$$\lceil \log_2(n+1) \rceil \leq \gamma_\ell^{\text{ID}}(G^r) \leq n. \quad (2.9)$$

Ces inégalités sont valides pour tout graphe  $G$  admettant un code  $(r, \leq \ell)$ -identifiant.

La borne supérieure est clairement serrée pour tout  $r$  et pour tout  $\ell$ . En effet, pour le graphe à  $n$  sommets n'ayant aucune arête, il est nécessaire et suffisant de prendre tous les sommets du graphe pour  $(r, \leq \ell)$ -identifier ceux-ci. Dans le cas  $\ell = 1$ , on va voir que cette borne peut être réduite d'une unité si l'on considère des graphes ayant au moins une arête (voir Section 2.4.1).

Dans des perspectives applicatives, c'est plus particulièrement la borne inférieure qui nous intéresse. Dans le cas  $\ell = 1$ , il était déjà montré dans l'article initial de Karpovsky *et al* [120] que celle-ci était serrée. Dans la Section 2.4.1, on va s'intéresser à l'ensemble des graphes atteignant cette borne pour  $r = 1$ .

On donnera quelques éléments sur la structure de ces graphes, et en particulier une caractérisation de ceux-ci.

Le cas de l'identification de sous-ensembles de sommets (cas  $\ell \geq 1$  quelconque) est traité dans la Section 2.4.2. L'une des raisons pour traiter ce cas à part est que celui-ci est considérablement plus complexe que le cas  $\ell = 1$ . En effet, nous ne connaissons à l'heure actuelle toujours pas la borne inférieure serrée de  $\gamma_\ell^{\text{ID}}(G)$ . Dans la Section 2.4.2, nous améliorons la borne inférieure de (2.9). Nous donnons de plus diverses constructions de familles de graphes dont la cardinalité minimum d'un code est proche de cette nouvelle borne inférieure.

Enfin, dans la Section 2.4.3, nous discutons quelques propriétés structurelles des graphes admettant un code. Nous nous intéressons, en particulier, à déterminer le nombre minimum de sommets d'un graphe admettant un code, ainsi qu'à la construction de graphes admettant un code à l'aide du produit cartésien de graphes.

## 2.4.1 Graphes extrémaux pour l'identification d'un seul sommet

### 2.4.1.1 Borne supérieure

La borne supérieure de (2.9) est clairement serrée, et ce pour tout  $r$  et  $\ell$ , à cause du graphe sans arête. Ce graphe étant un cas particulier sans grand intérêt par ailleurs, nous pouvons nous demander ce qu'il se passe lorsque nous considérons des graphes admettant au moins une arête.

Il s'avère que, dans ce cas, la borne supérieure descend d'une unité. En effet, nous avons démontré dans [96] avec Sylvain Gravier que tout graphe à  $n$  sommets et au moins une arête admettant un code identifiant admettait nécessairement un code identifiant de cardinalité  $n - 1$ . Plus généralement, nous montrons :

**Théorème 30** (Borne supérieure pour un graphe non vide de degré borné). *Soit  $r \geq 1$  et  $\Delta \geq 1$ , et soit  $G = (V, E)$  un graphe (fini ou infini) admettant un code  $r$ -identifiant, tel que tout sommet de  $G$  est de degré au plus  $\Delta$ . Alors il existe un sommet  $x$  de  $G$  tel que  $V \setminus \{x\}$  est un code  $r$ -identifiant de  $G$ .*

Ceci généralise un résultat de Charon, Hudry, et Lobstein [52], qui dit que tout graphe (fini) à  $n$  sommets ayant au moins une arête admet un code identifiant de cardinalité  $n - 1$ .

La preuve de ce théorème est basée sur l'observation suivante. Si  $x$  est un sommet ayant au moins un voisin, il faut et il suffit de montrer que tout sommet de  $V \setminus N[x]$  est séparé de tout sommet de  $N[x]$  pour montrer que  $V \setminus \{x\}$  est un code identifiant. En effet, comme on suppose que  $V$  est un code identifiant, alors les sommets de  $N[x]$  sont couverts et séparés entre eux par d'autres sommets que  $x$ , de même que les sommets de  $V \setminus N[x]$ .

L'astuce consiste alors à prendre un sommet  $x$  de degré maximum. Si  $V \setminus \{x\}$  est un code, alors c'est gagné. Sinon, cela signifie qu'il existe exactement une paire de sommets distincts  $u, v$  tels que  $N[u] \Delta N[v] = \{x\}$ . En particulier, l'un de ces sommets (disons  $u$ ) est voisin de  $x$ , tandis que l'autre (disons  $v$ ) ne l'est pas. Dans ce cas, nous montrons que  $V \setminus \{v\}$  est un code. En effet,  $x$  est déjà séparé de  $v'$  par lui-même, et  $v$  joue le même rôle que  $v'$  vis-à-vis des sommets distincts de  $x$ . S'ils étaient couverts et séparés par  $V$ , alors ils le sont également par  $V \setminus \{v\}$ . La seule possibilité serait alors que  $x$  ne soit pas séparé d'un voisin  $y$  de  $v'$ . En ce cas, on aurait  $N[y] = N[x] \cup \{v'\}$ , ce qui contredirait le fait que  $x$  est de degré maximum.

Le cas  $r \geq 1$  quelconque est obtenu en prenant la fermeture  $r$ -transitive du graphe. En effet, si  $G$  est de degré borné par  $\Delta$ , alors  $G^r$  est de degré borné par  $\Delta(\Delta - 1)^{r-1}$ .

Robert Skaggs montre également dans sa thèse l'inégalité  $\gamma^{\text{ID}}(G^r) \leq n - 1$  pour un graphe  $G$  à  $n$  sommets et au moins une arête [164, Théorème 3.2]. Sa preuve est relativement complexe, elle est basée sur une preuve fautive de cette inégalité publiée par Gimbel *et al* [90]. Nathalie Bertrand a également démontré cette inégalité dans son mémoire de M2 [24].

Ce résultat est le meilleur possible pour  $r = 1$ , comme le montre par exemple le cas de l'étoile  $K_{1,n}$ , qui comporte  $n$  sommets  $v_1, \dots, v_n$  de degré 1 et un sommet  $u$  de degré  $n$  (ce cas a déjà été discuté en Section 2.1). En effet, il ne peut y avoir, au plus, qu'un des sommets  $v_i$  qui n'appartienne pas au code (sinon deux tels sommets ne seraient pas séparés), et si  $v_i$  n'est pas dans le code alors cela implique que  $u$  est dans le code (pour couvrir  $v_i$ ). Ainsi la cardinalité minimum d'un code identifiant d'une étoile à  $n$  sommets est  $n - 1$ .

Un graphe particulièrement utile pour la suite est le graphe  $H_\infty(X)$  suivant, défini pour un sous-ensemble  $X \subseteq \mathbb{Z}$  quelconque (fini ou infini). Ses sommets sont les éléments de

$$\{x_i\}_{i \in X} \cup \{y_j\}_{j \in X},$$

et ses arêtes sont définies comme

$$\{x_i x_j \mid i \neq j\} \cup \{y_i y_j \mid i \neq j\} \cup \{x_i y_j \mid i < j\}.$$

On peut remarquer que le graphe  $H_\infty(\mathbb{Z})$  n'admet que l'ensemble de tous ses sommets comme code identifiant (voir Figure 2.15).

En effet, pour tout sommet  $x_i$ , on a  $N[y_{i+1}] = N[y_i] \cup \{x_i\}$ , ce qui force tout sommet  $x_i$  à être dans le code. Par ailleurs, on a par symétrie  $N[x_i] = N[x_{i+1}] \cup \{y_{i+1}\}$ , ce qui force tout sommet  $y_i$  à être dans le code. Cette observation était faite dans [52, 96]. Elle démontre que le résultat du Théorème 30 ne peut pas être étendu au cas des graphes infinis quelconques.

Le cas de ce graphe  $H_\infty(\mathbb{Z})$  est, en un sens, « pathologique », dans la mesure où tout graphe infini  $G = (V, E)$  n'admettant que  $V$  comme code identifiant est tel que tout sommet est contenu dans un sous-graphe de  $G$  isomorphe à  $H_\infty(\mathbb{Z})$ . Cette observation, que j'avais formulée dans ma thèse [146, Section 4.3], a été par la suite précisée par Foucaud *et al* (voir Théorème 32 ci-après).

Une version améliorée du Théorème 30 a été donnée par Bertrand [24] et Foucaud *et al* [77]. L'amélioration réside dans la possibilité d'avoir  $V \setminus \{x\}$  comme code identifiant avec  $x$  appartenant au voisinage fermé d'un sommet quelconque.

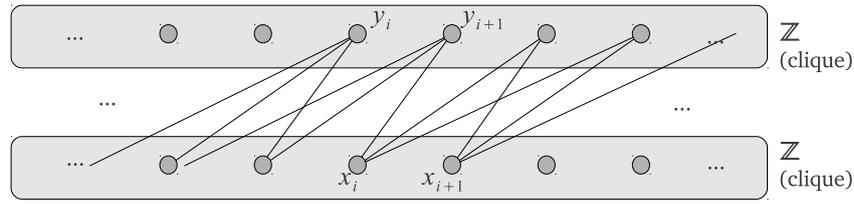


FIGURE 2.15 : Le graphe  $H_\infty(\mathbb{Z})$ . Ce graphe a la propriété de n'admettre que l'ensemble de tous ses sommets comme code identifiant.

**Théorème 31** (Borne supérieure avec choix local du sommet). Soit  $G = (V, E)$  un graphe (fini ou infini) admettant un code identifiant. Si  $G$  ne contient pas  $H_\infty(\mathbb{N})$  comme sous-graphe induit, alors pour tout sommet  $v$  de  $G$ , il existe  $u \in N[v]$  tel que  $V \setminus \{u\}$  est un code identifiant de  $G$ .

#### 2.4.1.2 Caractérisations de graphes extrémaux

Foucaud *et al* [77] ont caractérisé les graphes infinis n'admettant que leur ensemble de sommets comme code identifiant. Pour un graphe  $G = (V, E)$  (fini ou infini) muni d'un couplage *parfait*  $M \subseteq E$  (c'est-à-dire un sous-ensemble d'arêtes disjointes qui couvre l'ensemble des sommets), ils définissent une opération  $\Gamma(G, M)$  consistant à :

- substituer à chaque arête  $uv \in M$  un graphe infini  $H_u (= H_v)$  qui est isomorphe à  $H_\infty(\mathbb{Z})$
- pour chaque arête  $uv \in E \setminus M$ , joindre tous les sommets de  $H_u$  avec tous ceux de  $H_v$

L'opération  $\Gamma$  transforme alors un graphe (fini ou non)  $G$  muni d'un couplage parfait  $M$  en un graphe infini  $\Gamma(G, M)$  tel que tout sommet de  $\Gamma(G, M)$  appartient à un sous-graphe isomorphe à  $H_\infty(\mathbb{Z})$  (voir Figure 2.16).

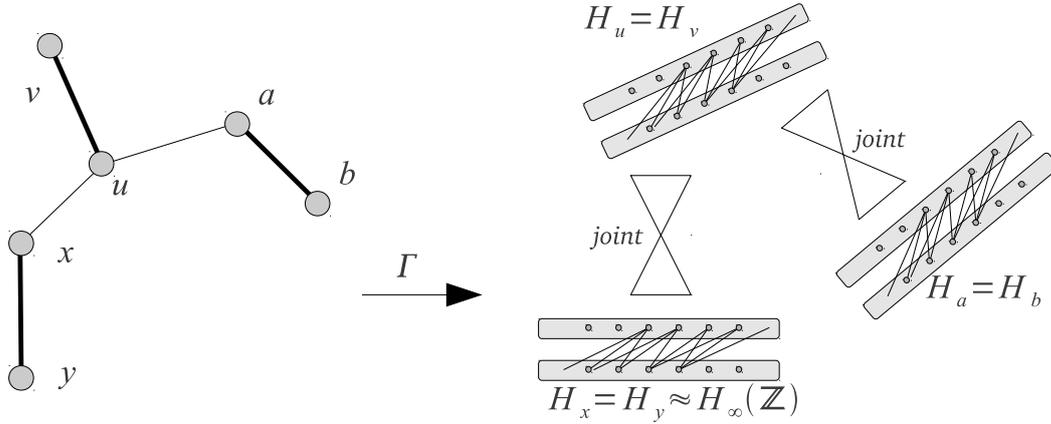
On dira dans la suite qu'un graphe est un  $\Gamma$ -*graphe* s'il peut être vu comme  $\Gamma(G, M)$ , où  $M$  est un couplage parfait d'un graphe  $G$  donné.

La caractérisation de Foucaud *et al* est alors la suivante :

**Théorème 32** (Caractérisation des graphes n'admettant que  $V$  comme code). Soit  $G = (V, E)$  un graphe infini n'admettant que  $V$  comme code identifiant. Alors  $G$  est l'union disjointe d'un nombre fini ou infini de  $\Gamma$ -graphes.

Foucaud *et al* ont également caractérisé les graphes à  $n$  sommets dont la cardinalité minimum d'un code identifiant est  $n - 1$  [77].

Cette caractérisation est basée sur le graphe  $H_k$ , dont les sommets sont  $\{v_1, \dots, v_{2k}\}$ , et tel que  $v_i v_j$  est une arête de  $H_k$ ,  $i \neq j$ , si, et seulement si, on a  $|i - j| \leq k - 1$ .



**FIGURE 2.16 :** Illustration de l'opération  $\Gamma$  de Foucaud et al. Les arêtes en gras sont celles du couplage parfait  $M$ . Chacune de ces arêtes est remplacée par le graphe  $H_\infty(\mathbb{Z})$ . Une arête du graphe initial, qui n'était pas dans le couplage, induit un joint entre deux copies, c'est-à-dire l'ensemble de toutes les arêtes possibles entre deux copies de  $H_\infty(\mathbb{Z})$ .

En d'autres termes, le graphe  $H_k$  est la fermeture  $(k-1)$ -transitive du chemin à  $2k$  sommets  $P_{2k} : H_k = P_{2k}^{k-1}$ .

Il est aisé de vérifier que  $\gamma^{\text{ID}}(H_k) = 2k-1$  pour  $k \geq 2$ . En effet, pour tout  $i \in \{1, \dots, k-1\}$ , on a  $N[x_{i+1}] = N[x_i] \cup \{x_{i+k}\}$ , ce qui implique que  $x_{k+1}, \dots, x_{2k-1}$  doivent être dans le code ; et pour tout  $i \in \{k+1, \dots, 2k\}$ , on a  $N[x_i] = N[x_{i+1}] \cup \{x_{i-k+1}\}$ , ce qui implique que  $x_1, \dots, x_{k-1}$  doivent être dans le code. De plus, pour séparer  $x_k$  de  $x_{k+1}$ , il est nécessaire que l'un des deux sommets  $x_1$  et  $x_{2k}$  soit dans le code.

Par ailleurs,  $\{v_2, \dots, v_{2k}\}$  et  $\{v_1, \dots, v_{2k-1}\}$  sont des codes identifiants de  $H_k$ .

Le deuxième ingrédient de la caractérisation de Foucaud *et al* est l'opérateur de joint. Étant donnés deux graphes  $G = (V_G, E_G)$  et  $H = (V_H, E_H)$ , le joint de  $G$  et  $H$ , noté  $G \bowtie H$ , est le graphe dont les sommets sont  $V_G \cup V_H$ , et dont les arêtes sont

$$E_G \cup E_H \cup \{uv \mid u \in V_G \text{ et } v \in V_H\}.$$

Autrement dit,  $G \bowtie H$  est obtenu en joignant par une arête tout sommet de  $G$  à tout sommet de  $H$ .

Cet opérateur de jointure est utilisé dans la construction de  $\Gamma(G, M)$  décrite ci-dessus (voir Figure 2.16).

Soit alors  $\mathcal{A}$  la classe de graphes définie comme :

$$\mathcal{A} = \{H_k \mid k \geq 1\},$$

et soit  $\mathcal{A}^*$  la classe de graphes définie comme :

$$\mathcal{A}^* = \{G_1 \bowtie G_2 \bowtie \dots \bowtie G_t \mid t \geq 1 \text{ et } G_1, \dots, G_t \in \mathcal{A}\}.$$

La caractérisation de Foucaud *et al* est alors la suivante :

**Théorème 33** (Caractérisation des graphes tels que  $\gamma^{\text{ID}} = n - 1$ ). *Soit  $G$  un graphe à  $n$  sommets admettant un code identifiant. Alors  $\gamma^{\text{ID}}(G) = n - 1$  si, et seulement si, on a*

$$G \in \{K_{1,t} \mid t \geq 2\} \cup \mathcal{A}^* \cup \mathcal{A}^* \bowtie K_1 \setminus \{H_1\}.$$

Robert Skaggs avait donné un sous-ensemble de ces graphes dans sa thèse [164, Propositions 3.5, 3.6 et 3.7].

Pour le cas général  $r \geq 1$ , la caractérisation des graphes tels que  $\gamma^{\text{ID}}(G^r) = |V(G)| - 1$  est laissée comme une question ouverte par Foucaud *et al* [77]. Charon, Hudry et Lobstein ont cependant montré que la borne  $\gamma^{\text{ID}}(G^r) \leq |V(G)| - 1$  était serrée, pour  $n$  assez grand [52] :

**Théorème 34.** *Soit  $r \geq 1$  et  $n \geq 3t^2$ . Alors il existe un graphe  $G$  à  $n$  sommets, admettant un code  $r$ -identifiant, tel que  $\gamma^{\text{ID}}(G^r) = n - 1$ .*

### 2.4.1.3 Borne inférieure

Dans cette section, on s'intéresse aux graphes à  $n$  sommets admettant un code identifiant de cardinalité  $\lceil \log_2(n + 1) \rceil$ , que l'on appellera *graphes optimaux*.

Des graphes optimaux étaient déjà donnés dans l'article de Karpovsky *et al* [120]. La construction de Karpovsky *et al* est la suivante.

Soit  $n \geq 1$ , et considérons le graphe  $G_n$  dont les sommets sont les  $N = 2^n - 1$  mots binaires de longueur  $n$  distincts du mot  $000 \dots 0$ . Soit  $e_i$  le mot binaire comportant un unique 1 sur le bit  $i$ , et soit  $C = \{e_1, \dots, e_n\}$ . Pour tout sommet  $x$  de  $G_n$  n'appartenant pas à  $C$ , il y a une arête entre  $x$  et  $e_i$  si, et seulement si, le bit  $i$  de  $x$  vaut 1. Clairement,  $C$  couvre l'ensemble des sommets de  $G_n$ , et le voisinage d'un sommet  $x \notin C$  identifie  $x$ . Par ailleurs, les sommets de  $C$  sont couverts et identifiés par eux-mêmes, et  $C$  est un code identifiant de  $G_n$ , de cardinalité  $n$ .

On obtient alors un graphe biparti

$$G_n = (C \cup (\{0, 1\}^n \setminus (\{000 \dots 0\} \cup C)), E),$$

tel que  $C$  est un code identifiant du graphe. En ce cas, le sous-graphe de  $G_n$  induit par les sommets du code n'a pas d'arête, de même que

le sous-graphe de  $G_n$  induit par les sommets qui ne sont pas dans le code. On peut modifier  $G_n$  en ajoutant des arêtes entre les sommets de  $\{0, 1\}^n \setminus (\{000\dots 0\} \cup C)$  sans changer la cardinalité minimum d'un code identifiant de  $G_n$ .

En fait, cette construction peut être étendue au cas où le sous-graphe induit par les sommets du code a des arêtes. C'est l'un des résultats que je présente dans une note parue dans *Discrete Applied Mathematics* [144].

La construction est la suivante. Pour un entier  $n \geq 1$  donné, soit  $p = \lceil \log_2(n+1) \rceil$  et soit  $H$  un graphe à  $p$  sommets  $x_1, \dots, x_p$  admettant un code identifiant. Soit alors  $\mathcal{X} = \{N[x_i] \mid i = 1, \dots, p\}$ , et soit  $\mathcal{Y}$  un sous-ensemble de l'ensemble des parties non vides de  $\{x_1, \dots, x_p\}$ , de cardinalité  $n - p$ , et tel que  $X \cap Y = \emptyset$ .

Pour tout sous-ensemble  $Y \in \mathcal{Y}$ , ajoutons alors à  $H$  un sommet  $v_Y$  qui est voisin de  $x_i$  si, et seulement si,  $x_i \in Y$ . Autrement dit, ce sommet  $v_Y$  est tel que  $N[v_Y] = Y$ , avec  $Y \notin \mathcal{X}$ .

Ajoutons enfin des arêtes quelconques entre des sommets du type  $v_Y$ .

Par construction, le graphe obtenu admet  $\{x_1, \dots, x_p\}$  comme code identifiant. Inversement, tout graphe à  $n$  sommets admettant un code identifiant de cardinalité  $\lceil \log_2(n+1) \rceil$  peut clairement être construit de cette façon.

**Théorème 35** (Caractérisation des graphes tels que  $\gamma^{\text{ID}} = \lceil \log_2(n+1) \rceil$ ). Soit  $G$  un graphe à  $n \geq 1$  sommets. Alors on a  $\gamma^{\text{ID}}(G) = \lceil \log_2(n+1) \rceil$  si, et seulement si,  $G$  peut être obtenu par la construction décrite ci-dessus.

Dans [144], je discute d'un certain nombre de paramètres concernant les graphes optimaux. En particulier, je montre que lorsque le graphe de la construction ci-dessus a  $n = 2^p - 1$  sommets, alors celui-ci a au minimum

$$p \left( 2^{p-1} - 1 \right) - \binom{p}{2} + \left\lfloor \frac{p}{2} \right\rfloor$$

arêtes. Je donne, de plus, une façon de construire un graphe optimal ayant ce nombre d'arêtes. Ce résultat est basé sur la proposition suivante, qui donne une borne générale sur le nombre d'arêtes d'un graphe admettant un code identifiant :

**Proposition 2** (Nombre maximum d'arêtes d'un graphe admettant un code). Soit  $G$  un graphe à  $n \geq 1$  sommets admettant un code identifiant. Alors  $G$  a au plus  $\binom{n}{2} - \lfloor \frac{n}{2} \rfloor$  arêtes.

Ce résultat est le meilleur possible. Pour s'en convaincre, il suffit de considérer un graphe complet à  $n$  sommets dont on enlève un couplage

maximum (un *couplage maximum* est tel qu'au plus un des sommets du graphe n'appartient à aucune arête du couplage). Je montre dans ma thèse que ce graphe est l'unique graphe extrémal de la Proposition 2.

Robert Skaggs a démontré plus tard ce résultat dans sa thèse [164, Corollaire 2.5], de façon indépendante.

La question du nombre minimum d'arêtes d'un graphe optimal a été reprise récemment par Raspaud et Tong [156], qui parviennent à déterminer le nombre minimum d'arêtes d'un graphe optimal ayant un nombre de sommets quelconque.

Pour un ensemble  $S$  de cardinalité  $n$ , soient  $A_1, \dots, A_{2^n-1}$  les sous-ensembles non vides de  $S$ , ordonnés d'une façon telle que

$$|A_1| \leq |A_2| \leq \dots \leq |A_{2^n-1}|.$$

Pour tout  $i$ , notons  $a_i = |A_i|$ .

Le résultat principal de Raspaud et Tong est le suivant.

**Théorème 36** (Nombre minimum d'arêtes d'un graphe optimal). *Soit  $n \geq 11$  et  $G$  un graphe optimal à  $n$  sommets, et soit  $p = \lceil \log_2(n+1) \rceil$ . Alors  $G$  a au moins*

$$\left\lceil \sum_{i=1}^n a_i - \left( \sum_{j=n-p+1}^n \frac{a_j}{2} \right) - \frac{p}{2} \right\rceil$$

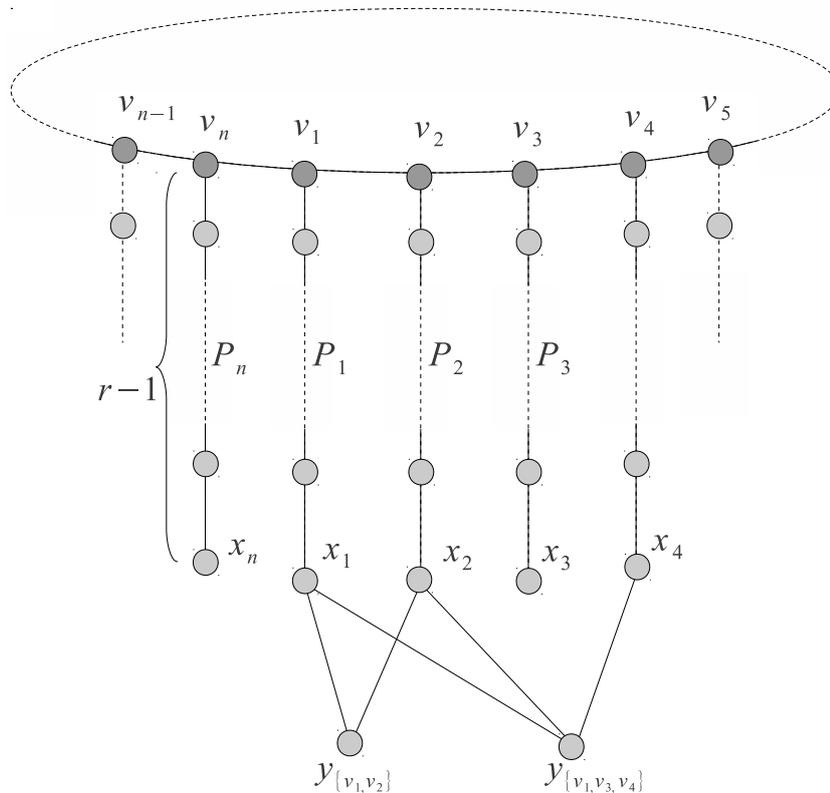
arêtes, et cette borne est serrée.

Pour  $r > 1$ , la borne  $\gamma^{\text{ID}}(G^r) \geq \lceil \log_2(n+1) \rceil$  est serrée. Charon, Hudry et Lobstein donnent en effet la construction d'un graphe à  $2^n - 1$  sommets admettant un code identifiant de cardinalité  $n$  pour tout  $n \geq 2r + 2$  [52].

Ce graphe est obtenu en attachant  $n$  chemins  $P_1, \dots, P_n$  de longueur  $r - 1$  aux sommets d'un cycle de  $n$  sommets  $v_1, \dots, v_n$ . Les extrémités de ces chemins sont alors des points d'attache  $x_1, \dots, x_n$ . Pour chaque sous-ensemble  $Y \subseteq \{v_1, \dots, v_n\}$  qui est distinct de tout  $B_r(v_i)$  et de tout  $B_r(x)$ , avec  $x$  appartenant à l'un des chemins  $P_1, \dots, P_n$ , on ajoute alors un sommet  $y_Y$  qui est voisin d'un point d'attache  $x_i$  si, et seulement si, on a  $v_i \in Y$ . Par construction, ce sommet  $y_Y$  est tel que  $B_r(y_Y) = Y$ , et le graphe admet donc  $\{v_1, \dots, v_n\}$  comme code  $r$ -identifiant (voir Figure 2.17).

Nous ne disposons pas à l'heure actuelle d'une caractérisation de l'ensemble des graphes extrémaux pour la borne inférieure dans le cas général  $r \geq 1$ .

En fait, dans le cas général  $r \geq 1$ , le problème s'apparente à la recherche de la racine d'un graphe, où l'on considère qu'un graphe  $H$  est la *racine*  $r$ -ème d'un graphe  $G$  si on a  $H^r = G$ . En effet, par définition,  $C$  est un code



**FIGURE 2.17 :** Un graphe extrémal pour la borne inférieure dans le cas général  $r \geq 1$ . Pour tout sous-ensemble  $Y \subseteq \{v_1, \dots, v_n\}$  qui est distinct de tout  $B_r(v_i)$  et de tout  $B_r(x)$ , avec  $x$  appartenant à l'un des chemins  $P_1, \dots, P_n$ , on ajoute alors un sommet  $y_Y$  qui est voisin d'un point d'attache  $x_i$  si, et seulement si, on a  $v_i \in Y$ . Ici, deux tels sommets  $y_Y$  sont représentés.

identifiant de  $G = H^r$  si, et seulement si,  $C$  est un code  $r$ -identifiant de  $H$ . Ayant caractérisé l'ensemble des graphes optimaux dans le cas des codes identifiants, il nous suffirait de considérer l'ensemble des racines  $r$ -èmes de ces graphes pour obtenir l'ensemble des graphes optimaux dans le cas des codes  $r$ -identifiants.

Malheureusement, cette approche a un inconvénient majeur, qui réside dans la complexité (au sens algorithmique) du calcul d'une racine d'un graphe. Motwani et Sudan ont montré dans les années 1990 que le problème était NP-difficile dans le cas  $r = 2$  [152].

**Théorème 37** (RACINE CARRÉE DE GRAPHE est NP-complet). Soit RACINE CARRÉE DE GRAPHE le problème suivant :

RACINE CARRÉE DE GRAPHE

Instance : Un graphe  $G$ .

Question : Existe-t-il un graphe  $H$  tel que  $H^2 = G$  ?

Alors RACINE CARRÉE DE GRAPHE est NP-complet.

Cette question fait, en partie, l'objet de la thèse de David Auger [8], ainsi que des articles récents suivants de Auger, Charon, Hudry et Lobstein [13, 14, 15].

Lorsque l'on considère le cas général des codes  $\ell$ -set-identifiants, alors la borne inférieure  $\gamma_\ell^{\text{ID}}(G) \geq \lceil \log_2(n+1) \rceil$  n'est pas serrée, et la situation est beaucoup plus complexe. La construction de graphes admettant de petits codes  $\ell$ -set-identifiants fait l'objet de la section suivante.

### 2.4.2 Construction de graphes admettant de petits codes pour l'identification d'ensembles de sommets

Dans une perspective applicative, une question naturelle — et, en un certain sens, « duale » à celle de la recherche d'un code minimum dans un graphe donné — concerne la recherche de graphes admettant des codes de faible cardinalité.

En effet, si l'on considère par exemple l'application concernant la détection de défaillance dans des réseaux multiprocesseurs (voir Section 2.1.2.1), il paraît légitime de se demander quelles sont les topologies de réseau qui sont efficaces vis-à-vis de la détection de défaillances. En admettant que le concepteur du réseau ait le choix, il aurait intérêt à opter pour une structure dans laquelle il existe un code identifiant de faible cardinalité.

De même, pour l'application de surveillance de bâtiments par des réseaux de capteurs (voir Section 2.1.2.2), on peut penser que la connaissance de structures efficaces vis-à-vis de la détection d'incidents (par exemple les incendies) pourra avoir un impact sur la conception du réseau en lui-même (typiquement l'ensemble des emplacements possibles pour l'implantation de capteurs).

Or, il s'avère que la cardinalité minimum de ce code peut être sujette à d'importantes variations. En effet, l'écart entre la borne supérieure et la borne inférieure générales est exponentiel, ce qui rend d'autant plus légitime la recherche de graphes admettant de petits codes. En particulier, il est à noter que la borne  $\gamma_\ell^{\text{ID}}(G) \geq \lceil \log_2(n+1) \rceil$  n'est pas serrée, et qu'on ne connaît d'ailleurs toujours pas de borne inférieure serrée sur la cardinalité minimum d'un code  $(r, \leq \ell)$ -identifiant d'un graphe à  $n$  sommets, même pour  $r = 1$ .

Dans cette section on va s'attacher à améliorer la borne inférieure sur la cardinalité d'un code  $\ell$ -set-identifiant, ainsi qu'à construire des familles de graphes admettant des codes  $\ell$ -set-identifiants de faible cardinalité.

### 2.4.2.1 Approche probabiliste

Une des possibilités pour construire des familles de graphes admettant des codes  $\ell$ -set-identifiants de faible cardinalité est d'utiliser des techniques probabilistes.

On va s'intéresser ici au modèle de *graphe aléatoire* dit « de Erdős-Rényi », dans lequel un graphe aléatoire  $G(n, p)$  a pour ensemble de sommets  $\{v_1, \dots, v_n\}$ , et est tel que l'arête  $v_i v_j$  existe avec une probabilité  $p$ . Notons que l'on a  $p \in [0, 1]$ , et que l'on s'autorise éventuellement à ce que  $p$  soit fonction de  $n$ . L'existence d'une arête  $v_i v_j$  est indépendante de l'existence d'une autre arête  $v_i' v_j'$ . L'espace probabilisé auquel appartient  $G(n, p)$  est noté  $\mathcal{G}(n, p)$ .

On dira alors que presque tout graphe de  $\mathcal{G}(n, p)$  a une certaine propriété  $Q$ , si l'on a

$$P(G(n, p) \text{ a la propriété } Q) \xrightarrow[n \rightarrow +\infty]{} 1.$$

Lorsque  $p$  est constant, il est très fréquent que l'on ait soit presque tout graphe qui vérifie  $Q$ , soit presque tout graphe qui ne vérifie pas  $Q$ , avec  $Q$  une propriété donnée. C'est notamment le cas pour les propriétés exprimables dans la logique du premier ordre [174].

Pour plus d'informations sur les graphes aléatoires, on renvoie le lecteur à l'ouvrage *Random graphs* de Bollobás [31].

En ce qui concerne les codes identifiants, Gimbel *et al* remarquent dans leur article de 2001 [90] que presque tout graphe admet un code identifiant.

**Proposition 3** (Presque tout graphe a un code identifiant). *Presque tout graphe de  $\mathcal{G}(n, \frac{1}{2})$  a un code identifiant.*

Ce résultat peut s'interpréter comme le fait que les graphes n'admettant pas de code identifiant sont « rares ». En effet, si l'on considère le cas particulier où  $p = \frac{1}{2}$ , l'espace  $\mathcal{G}(n, \frac{1}{2})$  contient tous les graphes étiquetés à  $n$  sommets, chacun de ces graphes ayant la même probabilité d'apparaître.

Dans un article de *Discrete Mathematics* de 2007 [83], issu d'un travail en collaboration avec Alan Frieze, Ryan Martin, Miklós Ruszinkó, et Clifford Smyth, nous étudions les codes  $\ell$ -set-identifiants dans les graphes aléatoires.

Pour  $\ell = 1$ , nous raffinons le résultat de Gimbel *et al* ci-dessus :

**Théorème 38** (Existence d'un code identifiant dans un graphe aléatoire). *Pour tout  $\varepsilon > 0$ , on a*

- (i) lorsque  $p = o(n^{-2})$  : presque tout graphe de  $\mathcal{G}(n, p)$  admet un code identifiant,
- (ii) lorsque  $pn^2 \rightarrow +\infty$  et  $p \leq \frac{1}{2n}(\ln n + (1 - \varepsilon) \ln \ln n)$  : presque aucun graphe de  $\mathcal{G}(n, p)$  n'admet de code identifiant,
- (iii) lorsque  $\frac{1}{2n}(\ln n + (1 + \varepsilon) \ln \ln n) \leq p \leq 1 - \frac{1}{n}(\ln n + \varepsilon \ln \ln n)$  : presque tout graphe de  $\mathcal{G}(n, p)$  admet un code identifiant,
- (iv) lorsque  $p \geq 1 - \frac{1}{n}(\ln n - \varepsilon \ln \ln n)$  : presque aucun graphe de  $\mathcal{G}(n, p)$  n'admet de code identifiant.

Ce résultat est dû au fait que, fondamentalement, il y a deux obstacles à surmonter pour qu'un graphe quelconque admette un code identifiant.

Tout d'abord, celui-ci ne doit pas avoir trop d'arêtes. En effet, un trop grand nombre d'arêtes nous fait prendre le risque d'avoir des jumeaux dans le graphe, et c'est presque sûrement le cas lorsque  $p \geq 1 - \frac{1}{n}(\ln n - \varepsilon \ln \ln n)$ .

Par ailleurs, lorsque le graphe a trop peu d'arêtes, on s'expose à ce que celui-ci possède des arêtes isolées, c'est-à-dire des paires de sommets voisins de degré 1. C'est le cas lorsque  $pn^2 \rightarrow +\infty$  et  $p \leq \frac{1}{2n}(\ln n + (1 - \varepsilon) \ln \ln n)$ . Noter que lorsque le graphe n'a presque sûrement aucune arête (cas où  $p = o(n^{-2})$ ), alors on est sauvé car, en ce cas, l'ensemble de tous les sommets du graphe est l'unique code identifiant de celui-ci.

Nous avons, par ailleurs, pu déterminer que presque tout graphe avait un code identifiant de cardinalité logarithmique :

**Théorème 39** (Presque tout graphe a un petit code). *Soit  $p$  tel que  $p \geq 4 \frac{\ln \ln n}{\ln}$  et  $1 - p \geq 4 \frac{\ln \ln n}{\ln}$ , et soit  $q = p^2 + (1 - p)^2$ . Pour tout  $\varepsilon > 0$ , presque tout graphe  $G(n, p) \in \mathcal{G}(n, p)$  est tel que*

$$\left| \frac{\gamma^{\text{ID}}(G(n, p)) \ln(q^{-1})}{2 \ln n} \right| \leq \varepsilon.$$

La preuve de ce résultat consiste, tout d'abord, à voir que presque tout sous-ensemble de sommets de cardinalité  $\frac{(2+\varepsilon) \ln n}{\ln(q^{-1})}$  est un code identifiant, ce qui peut être obtenu par l'application d'estimations élémentaires sur la probabilité que deux sommets distincts ne soient pas séparés.

Pour montrer que presque aucun sous-ensemble de sommets de cardinalité  $\frac{(2-\varepsilon) \ln n}{\ln(q^{-1})}$  n'est un code identifiant, la preuve utilise des arguments plus raffinés. Nous utilisons en particulier l'inégalité de Suen [167], qui utilise un graphe de dépendance associé aux événements que nous souhaitons éviter. Nous ne rentrons pas ici dans les détails de cette notion de graphe de dépendance, ni de cette inégalité.

Ce résultat a été redémontré, avec les mêmes techniques, par Xiao, Hadjicostis, et Thulasiraman [176].

Si l'on compare le Théorème 39 aux bornes générales  $\lceil \log_2(n+1) \rceil \leq \gamma^{\text{ID}}(G) \leq n$ , on en déduit que les graphes  $G$  tels que  $\gamma^{\text{ID}}(G) = \Omega(n)$  sont « rares », et que l'écrasante majorité des graphes à  $n$  sommets a un code identifiant de cardinalité  $O(\ln n)$ .

Dans le cas général  $\ell \geq 1$ , nous obtenons des résultats similaires à ceux des Théorèmes 38 et 39.

Pour ce qui concerne la cardinalité d'un code, notre résultat est le suivant :

**Théorème 40** (Presque tout graphe a un petit code, version  $\ell \geq 1$  quelconque). *Soit  $\varepsilon = \varepsilon(n) > 0$  tel que  $n^\varepsilon \rightarrow +\infty$ , et soit  $p \in ]0, 1[$  constant. Alors presque tout graphe  $G(n, p) \in \mathcal{G}(n, p)$  est tel quel*

$$\gamma_\ell^{\text{ID}}(G(n, p)) \leq \frac{2(\ell + \varepsilon) \ln n}{\ln(q_\ell^{-1})},$$

avec  $q_\ell = 1 - \min\{p, 2p(1-p)\}(1-p)^{\ell-1}$ .

Ceci implique que presque tout graphe admet un code  $\ell$ -set-identifiant de cardinalité  $O(\ell^2 \ln n)$ . Dans le cas particulier où  $p = \frac{1}{\ell}$ , nous donnons une meilleure estimation de la cardinalité d'un code :

**Théorème 41.** *Soit  $\varepsilon = \varepsilon(n) > 0$  tel que  $n^\varepsilon \rightarrow +\infty$ . Alors presque tout graphe  $G(n, p) \in \mathcal{G}(n, \ell^{-1})$  est tel quel*

$$\gamma_\ell^{\text{ID}}(G(n, p)) \leq \sqrt{2}(\ell^2 + \varepsilon) \ln n.$$

Une conséquence de ce résultat est l'existence d'une famille de graphes admettant un petit code  $\ell$ -set-identifiant :

**Corollaire 1** (Famille infinie de graphes ayant un petit code). *Il existe  $\mathcal{F}$ , une famille infinie de graphes admettant tous un code  $\ell$ -set-identifiant, telle que*

$$\gamma_\ell^{\text{ID}}(G) \leq O(\ell^2 \ln |V(G)|)$$

pour tout graphe  $G$  de  $\mathcal{F}$  (où  $|V(G)|$  désigne le nombre de sommets de  $G$ ).

Müller et Sereni ont obtenu des résultats similaires pour le modèle de graphe aléatoire toroïdal d'intersection de disques [154].

Dans ce modèle, noté  $\mathcal{U}(n, r)$ , on considère le carré  $[0, 1] \times [0, 1]$  dont les bords sont identifiés, c'est-à-dire que l'on considère que  $(x, 0) \equiv (x, 1)$  pour tout  $x$  et  $(0, y) \equiv (1, y)$  pour tout  $y$ .

On tire alors  $n$  points de façon indépendante selon une loi uniforme sur  $[0, 1]^2$ , et il y a une arête entre deux points dès lors que leur distance sur le tore est inférieure ou égale à  $2r$ . Le paramètre  $r$  peut éventuellement dépendre de  $n$ .

Le comportement de la probabilité d'existence d'un code dans un graphe aléatoire de  $\mathcal{U}(n, r)$  est similaire à celui de la probabilité d'existence d'un code dans un graphe aléatoire de  $\mathcal{G}(n, p)$ , dans le sens où, lorsqu'il y a trop d'arêtes, cela empêche le graphe d'avoir un code, de même que dans une zone où  $r$  est tel que des arêtes isolées existent presque sûrement.

Par ailleurs, lorsqu'il n'y a presque sûrement aucune arête, alors l'existence d'un code est assurée presque sûrement.

Cependant, pour d'autres valeurs de  $r$ , l'existence d'un code n'est jamais assurée presque sûrement, la probabilité tend alors vers une expression fonction de  $r$ .

**Théorème 42** (Existence de code dans les graphes aléatoires toroïdaux d'intersection de disques). *Soit  $G(n, r) \in \mathcal{U}(n, r)$ . Alors on a*

$$\lim_{n \rightarrow +\infty} P(G(n, r) \text{ a un code identifiant}) = \begin{cases} 1 & \text{si } nr^3 \xrightarrow[n \rightarrow +\infty]{} 0, \\ \exp\left(-\frac{\pi\lambda}{2}\right) & \text{si } nr^3 \xrightarrow[n \rightarrow +\infty]{} \lambda > 0, \\ 0 & \text{si } n^{-1} \ll nr^2 \ll n, \\ \exp(-\mu(r)) & \text{si } r \in ]0, \frac{\sqrt{2}}{2}[, \\ 0 & \text{si } r \geq \frac{\sqrt{2}}{2}, \end{cases}$$

avec  $f(n) \ll g(n)$  si, et seulement si,  $\frac{f(n)}{g(n)} \xrightarrow[n \rightarrow +\infty]{} 0$ , et  $\mu(r)$  une fonction « af-freuse » à écrire.

#### 2.4.2.2 Relation avec les codes superposés

Les codes identifiant des ensembles de sommets sont reliés aux codes dits *superposés*, introduits par Kautz et Singleton dans les années 1960 [122].

**Définition 7** (Code superposé). *Une famille  $\mathcal{F}$  de sous-ensembles de  $\{1, \dots, n\}$  est un code  $\ell$ -superposé (de dimension  $n$ ) si et seulement si :*

$$\bigcup_{X \in \mathcal{X}} X \neq \bigcup_{Y \in \mathcal{Y}} Y$$

pour tout  $\mathcal{X}, \mathcal{Y} \subseteq \mathcal{F}$ ,  $\mathcal{X} \neq \mathcal{Y}$ ,  $|\mathcal{X}| \leq \ell$ ,  $|\mathcal{Y}| \leq \ell$ .

Autrement dit, l'union d'au plus  $\ell$  éléments de  $\mathcal{F}$  est un sous-ensemble de  $\{1, \dots, n\}$  identifiant de façon unique les éléments de  $\mathcal{F}$  composant cette union.

Le lien entre les codes superposés et les codes identifiants est le suivant :

**Proposition 4** (Un code identifiant nous donne un code superposé). *Soit  $C$  un code  $\ell$ -set-identifiant d'un graphe  $G = (V, E)$ . Alors la famille*

$$\{N[v] \cap C \mid v \in V\}$$

*est un code  $\ell$ -superposé, de dimension  $|C|$ .*

En effet, par définition, l'intersection du code avec les voisinages fermés des sommets du graphe, constitue une famille de sous-ensembles de  $C$  telle que l'union d'au plus  $\ell$  de ces sous-ensembles détermine les sous-ensembles constituant cette union.

Le problème d'optimisation sous-jacent consiste à déterminer la cardinalité maximum d'un code  $\ell$ -superposé. La littérature contient de nombreux résultats sur les codes superposés et ses variantes. En particulier :

**Théorème 43** (Bornes générales sur les codes superposés). *Il existe deux constantes  $c_1, c_2$  telles que la cardinalité maximum d'un code  $\ell$ -superposé de dimension  $n$ , notée  $K(n, \ell)$ , vérifie :*

$$2^{c_1 n / \ell^2} \leq K(n, \ell) \leq 2^{c_2 n \ln \ell / \ell^2} \quad (2.10)$$

La borne inférieure est connue depuis les années 1960 [122], et un algorithme construisant un code  $\ell$ -superposé de cette cardinalité a été donné par Hwang et Sós dans les années 1980 [113]. En ce qui concerne la borne supérieure, une preuve utilisant des arguments probabilistes a été donnée par D'yachkov et Rykov [70], et des preuves purement combinatoires ont été données ultérieurement par Füredi et Ruzinkó [84, 161].

Une conséquence élémentaire de (2.10), que nous donnons dans [83], est que la cardinalité minimum d'un code  $\ell$ -set-identifiant est bornée inférieurement par un  $\Omega(\frac{\ell^2}{\ln \ell} \ln n)$ . Ceci améliore, en particulier, la borne originale de Karpovsky *et al*, qui était de  $\Omega(\ell \ln n)$ .

**Proposition 5** (Borne inférieure améliorée dans le cas général  $\ell \geq 1$ ). *Il existe une constante  $c > 0$  telle que*

$$\gamma_{\ell}^{\text{ID}}(G) \geq c \frac{\ell^2}{\ln \ell} \ln n$$

*pour tout graphe  $G$  admettant un code  $\ell$ -set-identifiant.*

Si un code  $\ell$ -set-identifiant nous donne toujours un code  $\ell$ -superposé (Proposition 4), on peut se poser la question sur la correspondance inverse. En effet, si une telle correspondance existait, on pourrait imaginer utiliser la borne inférieure de (2.10) afin d'obtenir une famille de graphes admettant un code  $\ell$ -set-identifiant de cardinalité  $O(\ell^2 \ln n)$ .

Karpovsky *et al* annoncent dans [120] qu'il est possible de construire des graphes admettant un code  $\ell$ -set-identifiant de faible cardinalité à partir de codes  $\ell$ -superposés :

« *Near-optimal constructions [of the graph] for sets of vertices can be obtained using superimposed codes.* »

Cette correspondance n'est cependant pas claire. En effet, la différence essentielle entre un code identifiant et un code superposé est que ces derniers évoluent dans un espace très peu structuré. En effet, un code superposé est une famille de sous-ensembles quelconques, alors que les voisinages fermés des sommets d'un graphe induisent de fortes contraintes structurelles sur la famille de sous-ensembles obtenue.

Si l'on voit ces codes comme des vecteurs binaires, on peut voir un code  $\ell$ -superposé comme une matrice dont les lignes sont telles que l'union d'au plus  $\ell$  lignes est distincte de l'union d'au plus  $\ell$  autres lignes. Un code  $\ell$ -set-identifiant correspond, quant à lui, à une matrice vérifiant cette même condition, à laquelle s'ajoutent des conditions supplémentaires de symétrie. En effet, on a  $u \in N[v]$  si, et seulement si,  $v \in N[u]$ . Par ailleurs, on a  $u \in N[u]$  pour tout sommet  $u$ .

Ainsi, si l'on souhaite construire un graphe admettant un code à partir de la matrice binaire  $M$  associée à un code superposé, il nous faut trouver une permutation des lignes et des colonnes de  $M$  de sorte à obtenir une matrice vérifiant des conditions de symétrie particulières. Or, il n'existe, à ma connaissance, aucune procédure nous permettant de trouver une permutation de lignes et de colonnes de  $M$  telle que la matrice obtenue puisse être vue comme la matrice de l'intersection des voisinages fermés des sommets d'un graphe.

Tout au plus, nous avons montré avec Alan Frieze, Ryan Martin, Miklós Ruszinkó, et Clifford Smyth que ceci était possible dans le cas des graphes orientés, où un sommet  $u$  est considéré couvrir un sommet  $v$  si  $(u, v)$  est un arc du graphe orienté (nous avons déjà considéré la notion de code identifiant dans les graphes orientés dans ce document, en Section 2.3.2).

Notre résultat utilise un code superposé *maximal*, c'est-à-dire qui est tel qu'aucun vecteur supplémentaire ne peut lui être ajouté pour obtenir un nouveau code superposé.

**Proposition 6** (Correspondance codes superposés – codes identifiants dans le cas des graphes orientés). *Soit  $\mathcal{F}$  un code  $\ell$ -superposé maximal de dimension  $n$ . Alors il est possible de trouver un ensemble de permutations des lignes et des colonnes de la matrice binaire  $M$  associée à  $\mathcal{F}$  de sorte que  $M$  puisse être vue comme la matrice d'incidence entre  $C$  et  $V$  d'un graphe orienté à  $|\mathcal{F}|$  sommets, muni d'un code  $\ell$ -set-identifiant  $C$  de cardinalité  $|C| = n$ .*

Dans la suite nous allons montrer comment, quitte à obtenir un code  $\ell$ -set-identifiant  $C$  de cardinalité strictement supérieure à la dimension d'un code  $\ell$ -superposé  $\mathcal{F}$ , nous pouvons construire des graphes (non orientés) munis d'un code identifiant à partir de codes superposés. Ces constructions nous permettront de construire des familles de graphes admettant des codes  $\ell$ -set-identifiants de faible cardinalité.

### 2.4.2.3 Constructions utilisant des codes superposés

Avec Sylvain Gravier, nous avons abordé la question de la construction de graphes ayant un code de faible cardinalité à partir de codes superposés, dans un article paru en 2005 à *Electronic Journal of Combinatorics* [95].

En utilisant les bornes connues sur les codes superposés (Théorème 43), nous parvenons à construire une famille infinie de graphes admettant un code  $\ell$ -set-identifiant de cardinalité  $\Theta(\ell^4 \log n)$ . La technique que nous développons établit une première passerelle depuis les codes superposés vers les codes identifiants. Nous utilisons comme gadgets additionnels des graphes connexes à  $\Theta(\ell^2)$  sommets admettant un code  $\ell$ -set-identifiant, l'existence de tels graphes étant assurée par la théorie des plans projectifs

**Théorème 44** (Plan projectif). *Soit  $A$  la matrice d'incidence d'un plan projectif de dimension  $q$ , avec  $q$  puissance d'un nombre premier. Alors le graphe ayant pour matrice d'adjacence la matrice*

$$B = \begin{pmatrix} 0 & A \\ A^T & 0 \end{pmatrix}$$

*est un graphe connexe à  $2(q^2 + q + 1)$  sommets admettant un code  $q$ -set-identifiant.*

On rappelle ici qu'un *plan projectif* d'ordre  $n$  est un hypergraphe ayant  $n^2 + n + 1$  sommets, tel que :

- (i) toute paire de sommets distincts est contenue dans une unique hyperarête,
- (ii) deux hyperarêtes s'intersectent en un unique sommet,
- (iii) tout sommet est contenu dans exactement  $n + 1$  hyperarêtes,

(iv) toute hyperarête contient exactement  $n + 1$  sommets.

L'idée de notre construction est la suivante. On considère tout d'abord la matrice binaire  $M$  dont les lignes correspondent aux vecteurs d'un code  $\ell$ -superposé maximal de cardinalité  $\Omega(2^{n/\ell^2})$  et de dimension  $n$  (grâce au Théorème 43 on sait qu'un tel code existe). Cette matrice a donc  $\Omega(2^{n/\ell^2})$  lignes et  $n$  colonnes. Nous montrons alors qu'il est possible de trouver une sous-matrice  $N$  de  $M$  telle que chaque ligne de  $N$  contient au moins un 1, et telle que  $N$  soit une sous-matrice carrée  $n \times n$ .

Si cette sous-matrice  $N$  était, de plus, symétrique et avait sa diagonale égale à 1, alors on pourrait considérer que celle-ci serait la matrice d'adjacence des sommets d'un code  $\ell$ -set-identifiant  $C$ , et chaque ligne de  $M \setminus N$  correspondrait alors à un sommet du graphe qui n'est pas dans le code. On obtiendrait alors un graphe à  $\Omega(2^{n/\ell^2})$  sommets, admettant un code  $\ell$ -set-identifiant de cardinalité  $n$ .

Malheureusement  $N$  n'a pas nécessairement ces propriétés. Nous allons donc altérer  $M$  en remplaçant chacune de ses colonnes par  $\Theta(\ell^2)$  colonnes, de sorte à obtenir une matrice  $\widetilde{M}$  ayant  $\Omega(2^{n/\ell^2})$  lignes et  $\Theta(\ell^2 n)$  colonnes. Cette matrice  $\widetilde{M}$  admet alors une sous-matrice carrée de taille  $\Theta(\ell^2 n) \times \Theta(\ell^2 n)$  qui a les propriétés de symétrie désirées. On peut alors voir  $\widetilde{M}$  comme une matrice correspondant à un graphe à  $N = \Omega(2^{n/\ell^2})$  sommets qui admet un code  $\ell$ -set-identifiant de cardinalité  $\Theta(\ell^2 n)$ .

Comme on a  $n = O(\ell^2 \ln N)$ , alors la cardinalité du code  $\ell$ -set-identifiant est de  $O(\ell^4 \ln N)$ .

**Théorème 45** (Famille de graphes ayant un code en  $O(\ell^4 \ln n)$ ). *Soit  $\ell \geq 2$ . Il existe une famille infinie de graphes  $\mathcal{F}$ , telle que  $G$  admet un code  $\ell$ -set-identifiant pour tout  $G \in \mathcal{F}$ , et telle que l'on ait*

$$\gamma_\ell^{\text{ID}}(G) \leq O(\ell^4 \ln |V(G)|)$$

pour tout graphe  $G \in \mathcal{F}$  (avec  $|V(G)|$  le nombre de sommets de  $G$ ).

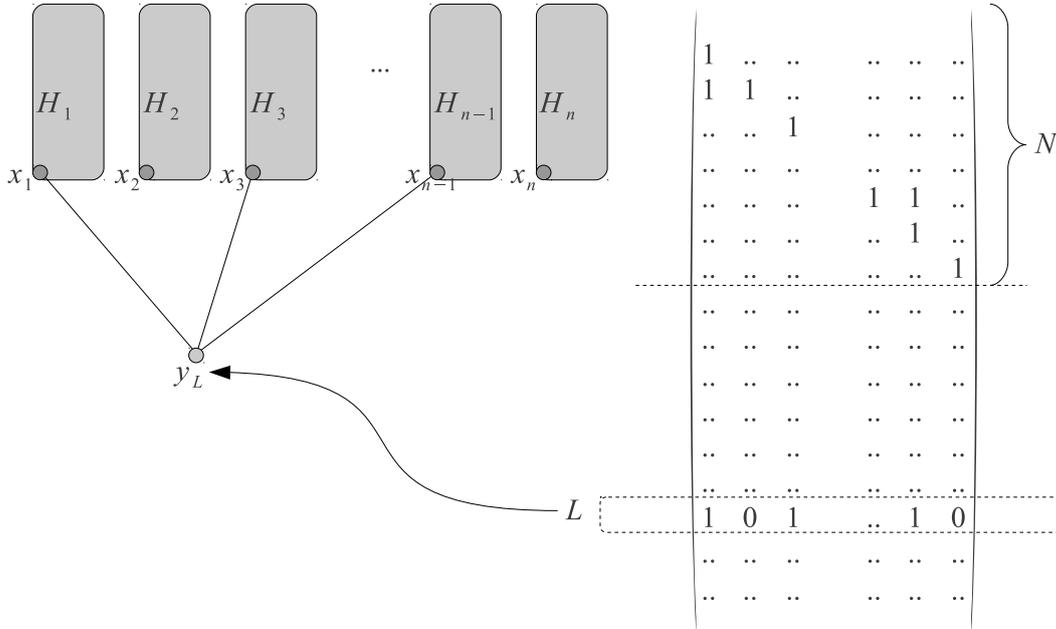
La technique utilisée pour altérer  $M$  est d'utiliser la matrice d'adjacence des graphes connexes à  $\Theta(\ell^2)$  sommets admettant un code  $\ell$ -set-identifiant du Théorème 44.

En terme de graphes, cela correspond à prendre  $n$  copies  $H_1, \dots, H_n$  du graphe du Théorème 44, et à sélectionner dans chacune de ces copies un sommet  $x_i$ ,  $i = 1, \dots, n$ . Ces sommets joueront le rôle de points d'attache, dans le sens où les sommets additionnels du graphe auront tous leur voisinage inclus dans  $\{x_1, \dots, x_n\}$ .

Pour chaque ligne  $L \notin N$  de la matrice  $M$ , on ajoute alors un sommet  $y_L$ , de sorte que

$$y_L \text{ est voisin de } x_i \Leftrightarrow L_i = 1,$$

où  $L_i$  désigne la  $i$ -ème coordonnée du vecteur ligne  $L$  (voir Figure 2.18).



**FIGURE 2.18 :** Procédé permettant de construire un graphe admettant un code  $\ell$ -set-identifiant de faible cardinalité, à partir d'un code  $\ell$ -superposé maximal (au sens de l'inclusion). Les graphes  $H_1, \dots, H_n$  sont  $n$  copies du graphe du Théorème 44.

Je suis par la suite parvenu à adapter l'algorithme de Hwang et Sós [113] — qui avait à l'origine été défini pour les codes superposés — au cas des codes identifiants [145] (article paru en 2006 dans *Designs Codes and Cryptography*).

L'idée de l'algorithme est d'ajouter de façon gloutonne des sommets à un graphe de base  $G_0$  obtenu par le Théorème 44. Les sommets qui vont être ajoutés auront un faible degré et partageront peu de sommets entre eux. En particulier, si  $x$  et  $y$  sont des sommets ajoutés à deux étapes consécutives de l'algorithme, alors la proportion des voisins de  $x$  qui sont aussi voisins de  $y$  sera toujours strictement inférieure à  $\frac{1}{\ell}$ . Ceci garantit qu'à chaque étape de l'algorithme les sommets de  $G_0$  forment un code  $\ell$ -set-identifiant du graphe obtenu. En effet, si deux ensembles de sommets  $X$  et  $Y$  ont le même voisinage fermé, alors il existe nécessairement un sommet de  $X$  dont la proportion de voisins, qui sont aussi voisins d'un sommet de  $Y$ , est supérieure ou égale à  $\frac{1}{\ell}$ .

Par un choix judicieux du degré des sommets ajoutés à chaque étape, on peut montrer que le nombre d'étapes de cet algorithme est au moins exponentiel. J'obtiens alors une procédure constructive permettant d'obtenir des familles infinies de graphes ayant les mêmes caractéristiques que dans le Corollaire 1.

**Théorème 46** (Famille de graphes ayant un code en  $O(\ell^2 \ln n)$ ). *Soit  $\ell \geq 2$ . Il existe une famille infinie de graphes  $\mathcal{F}$ , telle que  $G$  admet un code  $\ell$ -set-identifiant pour tout  $G \in \mathcal{F}$ , et telle que l'on ait*

$$\gamma_\ell^{\text{ID}}(G) \leq O(\ell^2 \ln |V(G)|)$$

pour tout graphe  $G \in \mathcal{F}$  (avec  $|V(G)|$  le nombre de sommets de  $G$ ).

L'algorithme est le suivant :

---

**Algorithme 1** Construction d'un code identifiant à partir d'un code superposé

---

- 1:  $G_0 \leftarrow$  graphe du Théorème 44
- 2: soit  $v_1, \dots, v_n$  les sommets de  $G_0$
- 3:  $\mathcal{F} \leftarrow \{A \subseteq \{v_1, \dots, v_n\} \mid |A| = \frac{n}{3\ell}\}$
- 4: **pour**  $i = 1 \dots n$  **faire**
- 5:    $\mathcal{F} \leftarrow \mathcal{F} \setminus \{A \in \mathcal{F} \mid |A \cap N(v_i)| \geq \frac{1}{\ell} |N(v_i)|\}$
- 6: **tant que**  $\mathcal{F} \neq \emptyset$  **faire**
- 7:   choisir un sous-ensemble  $A \in \mathcal{F}$
- 8:   ajouter  $y_A$  à  $G_0$ , de sorte que

$$y_A \text{ est voisin de } v_i \iff v_i \in A \quad \forall i = 1, \dots, n$$

- 9:    $\mathcal{F} \leftarrow \mathcal{F} \setminus \{B \in \mathcal{F} \mid |A \cap B| \geq \frac{1}{\ell} |A|\}$
- 

L'analyse de cet algorithme, et notamment la preuve que la dernière boucle tant que est exécutée un nombre de fois exponentiel, fait l'objet de la publication [145].

## 2.4.3 Structures des graphes admettant un code

### 2.4.3.1 Motivations

Un graphe donné  $G = (V, E)$  admet un code  $(r, \leq \ell)$ -identifiant si, et seulement si, celui-ci ne contient pas deux sous-ensembles distincts de sommets  $X, Y \subseteq V, X \neq Y$  tels que  $\bigcup_{x \in X} B_r(x) = \bigcup_{y \in Y} B_r(y)$ . Ceci montre que

la question de la reconnaissance des graphes admettant des codes est un problème facile (c'est-à-dire polynomial), dès lors que les paramètres  $r$  et  $\ell$  du problème sont donnés.

Ceci étant dit, il est légitime d'étudier la *structure* de ces graphes : quels sous-graphes ont-ils nécessairement ? Quel peut être leur nombre d'arêtes ou de sommets ? etc. Répondre à ces questions est très certainement utile pour l'étude des codes identifiants. Comme nous l'avons vu avec la Proposition 2, connaître, par exemple, le nombre maximum d'arêtes que peut avoir un graphe admettant un code, nous a permis de déterminer le nombre minimum d'arêtes d'un graphe optimal (voir Section 2.4.1.3).

Pour prendre un autre exemple, considérons la question de l'existence de procédures récursives de construction de tels graphes. Dans le cas  $r = \ell = 1$ , cette question a été traitée dans la thèse de David Auger, où il montre, en particulier [8, Théorème 2.1] :

**Théorème 47** (Construction récursive de graphes admettant un code). *Soit  $G = (V, E)$  un graphe, ayant au moins 4 sommets, connexe, et admettant un code identifiant. Alors, il existe  $v \in V$ , tel que le sous-graphe de  $G$  induit par  $V \setminus \{v\}$  est lui aussi connexe et admet un code identifiant.*

On rappelle qu'un graphe  $G = (V, E)$  est *connexe* s'il est impossible de partitionner  $V$  en  $V_1 \cup V_2$  (avec  $V_1, V_2 \neq \emptyset$ ) de sorte à avoir :

$$uv \in E \Rightarrow (u, v) \in V_1 \times V_1 \cup V_2 \times V_2.$$

En d'autres termes, un graphe est connexe si, et seulement si, il existe un chemin entre toute paire de sommets distincts du graphe.

Ce résultat permet de construire les graphes connexes admettant des codes de façon récursive en partant du chemin à trois sommets. Couplé au lemme suivant (de Charon, Hudry, Lobstein, mentionné dans la thèse de David Auger [8, Lemme 2.3]), il permet d'ailleurs de fournir une preuve alternative au fait que tout graphe fini ayant au moins une arête admet un code de cardinalité au plus  $n - 1$  (Théorème 30).

**Lemme 1.** *Soit  $G = (V, E)$  un graphe connexe, et admettant un code identifiant, et soit  $v \in V$  tel que le sous-graphe de  $G$  induit par  $V \setminus \{v\}$ , noté  $G[V \setminus \{v\}]$ , soit lui aussi connexe et admette un code identifiant. Alors on a*

$$\gamma^{\text{ID}}(G) \leq \gamma^{\text{ID}}(G[V \setminus \{v\}]) + 1.$$

En effet, si l'on enlève  $v$  à  $G$ , alors le seul problème que l'on peut rencontrer est que  $v$  soit non couvert (en ce cas ajouter  $v$  à un code identifiant

de  $G[V \setminus \{v\}]$  nous fournit un code identifiant de  $G$ ), ou que  $v$  ne soit pas séparé d'un sommet  $x$ . Dans ce cas, il suffit d'ajouter n'importe quel sommet  $y \in N[v] \Delta N[x]$  à un code identifiant de  $G[V \setminus \{v\}]$  pour obtenir un code identifiant de  $G$ .

La preuve que  $\gamma^{\text{ID}}(G) \leq |V(G)| - 1$ , pour tout graphe admettant un code ayant au moins une arête, découle alors du fait que  $\gamma^{\text{ID}}(P_3) = 2 (= 3 - 1)$ .

Dans la suite, nous allons discuter de questions ayant trait à la structure des graphes admettant des codes. En Section 2.4.3.2, nous discuterons du nombre de sommets et d'arêtes de graphes admettant des codes, et notamment des sous-graphes que ceux-ci ont nécessairement. Dans la Section 2.4.3.3, nous étudierons le produit cartésien en tant qu'opération permettant d'obtenir des graphes admettant des codes. Cette problématique a déjà été abordée en Section 2.2.4 (voir le Théorème 16).

#### 2.4.3.2 Sous-graphes et nombre de sommets

Pour  $r \geq 1$ , un graphe admettant un code  $r$ -identifiant doit avoir ses sommets relativement « éloignés » les uns des autres. Ceci implique naturellement l'existence d'un chemin de longueur minimum dans un tel graphe.

**Proposition 7** (Nombre minimum de sommets). *Soit  $r \geq 1$ , et soit  $G$  un graphe connexe admettant un code  $r$ -identifiant ayant au moins une arête. Si  $n$  est le nombre de sommets de  $G$ , alors on a*

$$n \geq 2r + 1.$$

*De plus, le chemin à  $2r + 1$  sommets est l'unique graphe admettant un code  $r$ -identifiant ayant  $2r + 1$  sommets.*

La preuve de ce résultat, que je donnais dans ma thèse [146, Proposition 4.1], consiste à montrer que tout graphe admettant un code  $r$ -identifiant possède nécessairement  $P_{2r+1}$ , le chemin à  $2r + 1$  sommets, comme sous-graphe. Ce résultat a été par la suite redémontré par Charon, Honkala, Hudry, et Lobstein [46], qui ont alors conjecturé qu'un tel graphe devait nécessairement contenir un chemin  $P_{2r+1}$  comme sous-graphe induit.

Cette conjecture a été démontrée l'année suivante par David Auger [6].

**Théorème 48** (Longs chemins induits). *Soit  $r \geq 1$  et soit  $G$  un graphe connexe admettant un code  $r$ -identifiant ayant au moins une arête. Alors  $G$  contient  $P_{2r+1}$  comme sous-graphe induit.*

Les questions connexes du diamètre et du rayon des graphes admettant des codes  $r$ -identifiants ont été abordées par Auger, Charon, Honkala, Hudry, et Lobstein [9].

Une conséquence de la Proposition 7 est que le nombre minimum d'arêtes d'un graphe connexe admettant un code  $r$ -identifiant est de  $2r$ . En ce qui concerne le nombre maximum d'arêtes que peut avoir un tel graphe, nous avons déjà vu (Proposition 2) qu'il était de

$$\binom{n}{2} - \lfloor \frac{n}{2} \rfloor$$

dans le cas où  $r = 1$ .

Le cas où  $r \geq 1$  est étudié par Auger, Charon, Honkala, Hudry, et Lobstein [9], qui n'obtiennent pas le nombre minimum d'arêtes d'un graphe admettant un code  $r$ -identifiant, mais une estimation relativement précise de celui-ci. Dans le cas  $r \geq 3$ , ils montrent que le nombre d'arêtes que l'on doit enlever à un graphe complet pour obtenir un graphe admettant un code  $r$ -identifiant, est de l'ordre de  $O(rn \log_2 n)$ .

**Théorème 49** (Nombre maximum d'arêtes, cas  $r \geq 3$ ). *Soit  $r \geq 3$ , et soit  $m_{\min}^r$  le nombre minimum d'arêtes d'un graphe admettant un code  $r$ -identifiant ayant au moins une arête. Alors on a*

$$\left| m_{\min}^r - \binom{n}{2} \right| \leq O(rn \log_2 n).$$

Dans le cas de l'identification de sous-ensembles de sommets, la littérature est beaucoup moins diserte que dans le cas  $\ell = 1$ , et même si l'on considère l'identification à distance  $r = 1$  (c'est-à-dire les codes  $\ell$ -set-identifiants).

Clairement, un graphe admettant un code 2-set-identifiant ne peut avoir de sommet de degré 1. En effet, si  $x$  est un tel sommet, et  $y$  son unique voisin, alors on aurait  $N[x] \cup N[y] = N[y]$ , ce qui implique que les sous-ensembles  $\{x, y\}$  et  $\{y\}$  ne sont pas séparés, ce qui est une contradiction.

De façon générale, on a d'ailleurs :

**Proposition 8** (Degré minimum d'un graphe admettant un code). *Soit  $G$  un graphe admettant un code  $\ell$ -set-identifiant, qui ne contient aucun sommet de degré 0. Alors le degré de tout sommet de  $G$  est au moins  $\ell$ .*

En effet, si  $x$  est un sommet de degré  $\leq \ell - 1$ , alors on aurait  $N[x] \cup N[N(x)] = N[N(x)]$ , et les sous-ensembles  $N(x)$  et  $N[x]$  ne seraient pas séparés. Cette observation a été formulée par Laihonon et Ranto [135].

Ainsi, un graphe connexe admettant un code *2-set-identifiant* contient nécessairement un cycle. Intuitivement, le plus long de ces cycles ne doit pas être trop petit, au risque de trouver des sous-ensembles de sommets ne pouvant pas être séparés. Par exemple, dans le cycle à 6 sommets  $v_1, \dots, v_6$ , on a  $N[v_1] \cup N[v_4] = N[v_2] \cup N[v_5]$ , ce qui implique que le cycle à 6 sommets ne peut admettre de code *2-set-identifiant*.

Si l'on considère maintenant un cycle à 7 sommets, il est facile de voir que  $N[u] \cup N[v] \neq N[u'] \cup N[v']$  pour tout  $u, v, u', v'$  tels que  $\{u, v\} \neq \{u', v'\}$ . Ceci montre que le cycle à 7 sommets admet un code *2-set-identifiant*.

La question de savoir si tout graphe ayant un code *2-set-identifiant* possède un cycle de longueur au moins 7 comme sous-graphe avait été posée lors de la session de problèmes du *Workshop on Codes and Discrete Probability* à Grenoble en 2007. Auger, Charon, Hudry et Lobstein y ont répondu par l'affirmative dans un article récent de *Discussiones Mathematicae-Graph Theory* [13].

**Proposition 9** (Existence d'un cycle de longueur au moins 7 dans un graphe admettant un code *2-set-identifiant*). *Soit  $r \geq 1$  et soit  $G$  un graphe connexe admettant un code *2-set-identifiant* ayant au moins une arête. Alors  $G$  contient un cycle de longueur au moins 7 comme sous-graphe.*

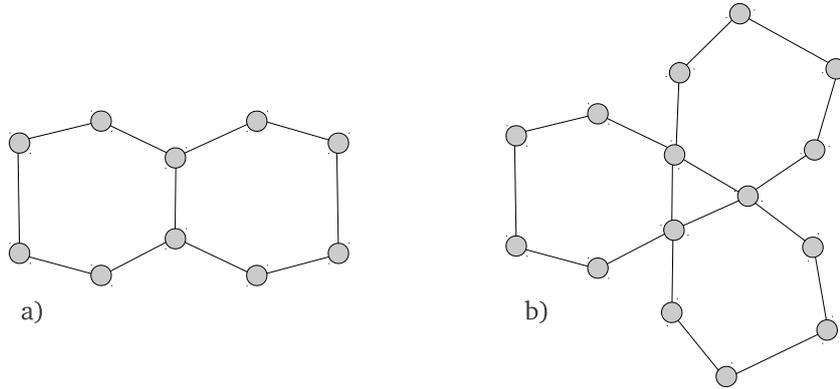
On ne peut étendre ce résultat au cas d'un cycle induit, comme le montre la Figure 2.19 ci-dessous. Notons que des cycles  $C_k$ , avec  $k < 7$ , peuvent cependant être présents en tant que sous-graphes induits dans un graphe admettant un code *2-set-identifiant* (voir Figure 2.19).

Dans le cas  $r \geq 1$ , Auger, Charon, Hudry et Lobstein conjecturent dans [13] qu'un graphe admettant un code  $(r, \leq 2)$ -identifiant a un cycle de longueur au moins  $4r + 3$  comme sous-graphe.

Dans le cas général  $\ell \geq 2$ , peu de choses sont connues à l'heure actuelle. Tero Laihonon a donné un graphe admettant un code *3-set-identifiant* à 16 sommets [133]. On ignore aujourd'hui le nombre minimum de sommets d'un graphe admettant un code *3-set-identifiant*.

Nous avons montré avec Sylvain Gravier que pour tout  $\ell \geq 1$ , il existe des graphes admettant un code  $\ell$ -set-identifiant de degré minimum  $\ell$ . Notre construction utilise les graphes issus des plans projectifs du Théorème 44.

Tero Laihonon a par la suite montré qu'il existe des graphes admettant un code  $\ell$ -set-identifiant qui sont aussi  $\ell$ -réguliers (c'est-à-dire que le degré



**FIGURE 2.19 :** Deux graphes admettant un code 2-set-identifiant sans cycle  $C_k$  induit pour  $k \geq 7$ . Notons que ces graphes possèdent des  $C_6$  induits, et que le graphe b) possède même un  $C_3$  induit. Ces deux graphes sont issus de [13].

de chaque sommet est égal à  $\ell$ ). Sa construction utilise des cages, une  $(k, g)$ -cage étant un graphe ayant un nombre minimum de sommets parmi les graphes  $k$ -réguliers de maille au moins  $g$  (avec  $k \geq 1, g \geq 3$ ).

En effet, il montre le résultat suivant :

**Théorème 50** (Graphes de grande maille). *Soit  $\ell \geq 2$  et  $G$  un graphe. Si  $G$  est  $\ell$ -régulier et a une maille d'au moins 7, alors  $G$  admet un code  $\ell$ -set-identifiant. Si  $G$  est  $(\ell + 1)$ -régulier et a une maille d'au moins 5, alors  $G$  admet un code  $\ell$ -set-identifiant.*

Comme pour tout  $\ell$ , tel que  $\ell - 1$  est la puissance d'un nombre premier, il existe des  $(\ell, 8)$ -cages ayant  $\Theta(\ell^3)$  sommets (voir par exemple [163]), alors l'existence de graphes  $\ell$ -réguliers à  $\Theta(\ell^3)$  sommets admettant un code  $\ell$ -set-identifiant est assurée.

Ce résultat améliore le Théorème 44 dans le sens où celui-ci nous fournit des graphes  $(\ell + 1)$ -réguliers à  $\Theta(\ell^2)$  sommets admettant un code  $\ell$ -set-identifiant.

Dans [134], rédigé lors d'un séjour de recherche invité à l'université de Turku, nous montrons avec Tero Laihonon qu'il existe des graphes  $\ell$ -réguliers à  $\Theta(\ell^2)$  sommets admettant un code  $\ell$ -set-identifiant. Nous utilisons de nouveau des plans projectifs pour construire des graphes distance-réguliers, selon une méthode inspirée par Gardiner [87].

Un graphe est dit *distance-régulier* si, pour tout  $i$ , il existe  $b_i$  et  $c_i$  tels que pour deux sommets  $u$  et  $v$  à distance  $i$ , on a exactement  $b_i$  sommets qui sont à la fois voisins de  $u$  et à distance exactement  $i + 1$  de  $v$ , et on a exactement  $c_i$  sommets qui sont à la fois voisins de  $u$  et à distance exactement  $i - 1$  de  $v$ .

Nous démontrons, de plus, en utilisant le lien entre codes identifiants et codes superposés (voir la Proposition 4 et le Théorème 43), que ce résultat est le meilleur possible.

### 2.4.3.3 Produit cartésien

Les constructions de graphes admettant des codes  $\ell$ -set-identifiants décrites dans la section ci-dessus invoquent des objets combinatoires très spécifiques (cages, plans projectifs, etc.), et nous fournissent des graphes admettant des codes ayant des propriétés particulières.

Il serait cependant utile que la boîte à outil du théoricien des graphes dispose d'outils génériques plus simples pour construire des graphes admettant des codes. Dans cette section, on va s'intéresser au produit cartésien de graphes, déjà discuté en Section 2.2.3.

Nous avons déjà vu que  $K_2$  n'admet pas de code identifiant (ses deux sommets étant jumeaux), mais que l'hypercube  $Q_n$ , qui peut être vu comme

$$Q_n = K_2 \square K_2 \square \dots \square K_2 \quad (n - 1 \text{ produits}),$$

admet, lui, un code identifiant.

Dans le cas des codes  $\ell$ -set-identifiants, cette opération peut également être utile. Je montre, par exemple, dans ma thèse [146, Proposition 4.2] que le produit cartésien de cycles fournit un graphe admettant un code identifiant.

**Proposition 10** (Produits de cycles). *Soit  $\ell \geq 1$  et soient  $p_1, \dots, p_\ell$  tels que  $p_i \geq 4$  pour tout  $i = 1, \dots, \ell$ . Alors le graphe  $C_1 \square C_2 \square \dots \square C_\ell$  admet un code  $\ell$ -set-identifiant (où  $C_k$  désigne le cycle à  $k$  sommets).*

Notons l'intérêt du produit cartésien, puisqu'un cycle admet, au mieux, un code 2-set-identifiant (à cause du degré minimum, voir Proposition 8), mais que le produit cartésien de  $\ell$  cycles admet un code  $\ell$ -set-identifiant.

Avec Tero Laihonen, nous avons abordé la question de l'existence d'une fonction  $f : \mathbb{N}^2 \rightarrow \mathbb{N}$  telle que, pour tout graphe  $G_1$  admettant un code  $\ell_1$ -set-identifiant, et tout graphe  $G_2$  admettant un code  $\ell_2$ -set-identifiant, le graphe  $G_1 \square G_2$  admet un code  $f(\ell_1, \ell_2)$ -set-identifiant.

Notre résultat principal [134] est le suivant :

**Théorème 51** (Produit cartésien de graphes en général). *Soient  $\ell_1, \ell_2 \geq 1$ , et soit  $G_1$  admettant un code  $\ell_1$ -set-identifiant,  $G_2$  admettant un code  $\ell_2$ -set-identifiant, tels que  $G_1$  et  $G_2$  sont connexes et ont au moins deux sommets. Alors, le produit cartésien  $G_1 \square G_2$  admet un code  $\max(\ell_1, \ell_2)$ -set-identifiant.*

Ce résultat est le meilleur possible, dans la mesure où nous avons également montré [134] :

**Théorème 52.** *Soient  $\ell_1, \ell_2 \geq 1$ , et soit  $\delta_2$  tel que  $\delta_2 \geq \ell_2 + 1$  et  $\ell_1 \geq \delta_2 + 1$ . Soient  $G_1$  de degré minimum  $\ell_1$  admettant un code  $\ell_1$ -set-identifiant,  $G_2$  de degré minimum  $\delta_2$  admettant un code  $\ell_2$ -set-identifiant, tels que  $G_1$  et  $G_2$  sont connexes et ont au moins deux sommets. Alors, le produit cartésien  $G_1 \square G_2$  n'admet pas de code  $(\ell_1 + 1)$ -set-identifiant. De plus, il existe de tels graphes  $G_1, G_2$ .*

Ainsi, le cas des cycles et des hypercubes doit être considéré comme une « bonne surprise » de la nature. Ce sont, en effet, des graphes tels que  $G_1 \square G_2$  admet un code  $f(\ell_1, \ell_2)$ -set-identifiant, avec  $f(\ell_1, \ell_2) > \max(\ell_1, \ell_2)$ . Dans le cas de l'hypercube, on a même  $f(\ell_1, \ell_2) = \ell_1 + \ell_2$ . En effet, il est connu que pour  $n$  impair, l'hypercube  $Q_n$  admet un code  $\frac{n+1}{2}$ -set-identifiant, mais pas de code  $\frac{n+3}{2}$ -set-identifiant [132]. Or, le produit cartésien  $Q_n \square Q_n$  n'est autre que l'hypercube  $Q_{2n}$ , qui admet un code  $(n + 1)$ -set-identifiant.

De même, pour les cycles,  $C_4$  n'admet qu'un code 1-set-identifiant, alors que  $\square^\ell C_4$  admet un code  $\ell$ -set-identifiant.

Il serait intéressant de connaître les conditions structurelles sur  $G_1$  et  $G_2$  qui font que  $G_1 \square G_2$  admet un code  $f(\ell_1, \ell_2)$ -set-identifiant, avec  $f(\ell_1, \ell_2) > \max(\ell_1, \ell_2)$ .

Il serait également intéressant d'étudier les autres produits classiques de graphes, et de déterminer ceux qui peuvent présenter un intérêt pour la construction de graphes admettant des codes  $\ell$ -set-identifiants.

L'exemple de l'opération de joint, utilisé dans la caractérisation de Foucaud *et al* des graphes admettant un code identifiant minimum de cardinalité  $n - 1$  (voir Théorème 33), montre qu'un produit de graphes quelconques n'a pas nécessairement d'intérêt. En effet, pour tout graphe  $G_1, G_2$ , le graphe  $G_1 \bowtie G_2$  n'admet pas de code  $\ell$ -set-identifiant pour  $\ell \geq 2$ .

Il y a cependant d'autres produits de graphes à étudier [102], qui pourraient éventuellement fournir de bons outils pour la construction de graphes admettant des codes  $\ell$ -set-identifiants.

## 2.5 Codes adaptatifs

### 2.5.1 Vers des dispositifs d'identification dynamiques

D'un point de vue applicatif, les codes identifiants tels que définis par Karpovsky *et al* [120] forment un dispositif de localisation et d'identification essentiellement statique.

Considérons par exemple l'application à la détection de défaillances dans des réseaux de processeurs (voir Section 2.1.2.1). Tout se passe comme si les processeurs du code interrogeaient *simultanément* leur voisinage étendu pour détecter l'éventuelle présence d'un processeur défectueux dans le réseau. On peut alors imaginer que ces processeurs envoient leurs réponses à un contrôleur central qui déduit, de l'ensemble des réponses collectées, la présence ou non d'un processeur défectueux dans le réseau. Ainsi, un code identifiant est un sous-ensemble de sommets *fixé une fois pour toutes*.

Cette vision des choses est certainement pertinente pour un grand nombre d'applications pratiques, comme par exemple dans le cas de la surveillance de bâtiments pour la détection d'incendies (voir Section 2.1.2.2).

On peut néanmoins imaginer que, dans certains contextes, il soit pertinent et efficace de ne pas fixer un code *a priori*, mais de *construire dynamiquement celui-ci* au fur et à mesure des réponses collectées.

Ainsi un code identifiant n'est plus un sous-ensemble de sommets fixé à l'avance mais une stratégie adaptative, que l'on peut voir comme un *arbre de décision*, permettant d'adapter l'ordre d'interrogation des sommets en fonction des réponses déjà collectées. L'intérêt d'une telle approche est de pouvoir dénicher l'éventuel sommet défaillant en interrogeant un nombre de sommets moindre que dans le cas d'un code identifiant non-adaptatif. Dans le cas de réseaux multiprocesseurs, on peut typiquement supposer que le contrôleur central sera capable d'ordonnancer ses requêtes de sorte à suivre une stratégie de détection basée sur un arbre de décision plutôt que sur un sous-ensemble fixé. C'est la proposition que je fais dans ma thèse [146], et que nous développons dans un article publié dans le *Journal of Combinatorial Theory Series A* [20] avec Yael Ben-Haim, Sylvain Gravier, et Antoine Lobstein, sous la terminologie *code identifiant adaptatif*.

### 2.5.2 Les codes identifiants vus comme un problème de recherche binaire

Les problèmes de recherche binaire, tels que définis par Rényi [158], peuvent être vus comme des jeux à deux joueurs.

L'un des joueurs, Paul, doit deviner une information connue du deuxième joueur, Carole (une justification du choix de ces prénoms peut être trouvée dans l'article de Christian Deppe [65]). Paul n'a le droit de poser que des questions admettant une réponse binaire (OUI/NON). Selon le contexte, l'information détenue par Carole peut être encadrée par un ensemble de contraintes, et on peut donner des règles plus précises sur le

type de questions que Paul a le droit de poser. L'objectif pour Paul est de déterminer l'information détenue par Carole en un nombre minimum de questions. Dans certains cas, on considère que Carole a le droit de mentir jusqu'à  $k$  fois, avec  $k$  constante fixée à l'avance (et connue de Paul et Carole) — dans ce cas on parle plutôt de jeux de Rényi-Ulam.

Les codes identifiants adaptatifs dans un graphe peuvent être vus comme un jeu de Rényi, où Carole sait si, oui ou non, il existe un sommet défectueux (et si oui, lequel est défectueux) dans un graphe donné  $G$ , dont la structure est connue de Paul et Carole. Paul n'a le droit que de poser des questions du type

« y a-t-il un sommet défectueux dans le voisinage fermé du sommet  $v$  ? »

avec  $v$  sommet quelconque de  $G$ .

Si l'ensemble des questions que Paul va poser doit être fixé à l'avance, alors celui-ci peut au mieux proposer un nombre de questions égal à la cardinalité d'un code identifiant dans le graphe  $G$ .

On parle de code *adaptatif* lorsque Paul a le droit — comme dans la plupart des vrais jeux de devinettes — de poser ses questions les unes après les autres, et de décider, à la volée, quelle question poser en fonction des réponses déjà reçues. En choisissant soigneusement les questions qu'il va poser à Carole, Paul pourra alors envisager de déterminer si, oui ou non, il existe un sommet défectueux (et si oui, lequel) en un nombre de questions inférieur à la cardinalité minimum d'un code identifiant dans  $G$ .

### 2.5.3 Définition et intérêt du concept

Un code identifiant adaptatif d'un graphe  $G = (V, E)$  admettant un code identifiant est un arbre binaire de décision, dont les nœuds qui ne sont pas des feuilles, sont des questions du type

« y a-t-il un sommet défectueux dans le voisinage fermé du sommet  $v$  ? »

et dont les arcs correspondent à la réponse (OUI/NON) à ces questions (voir Figure 2.20).

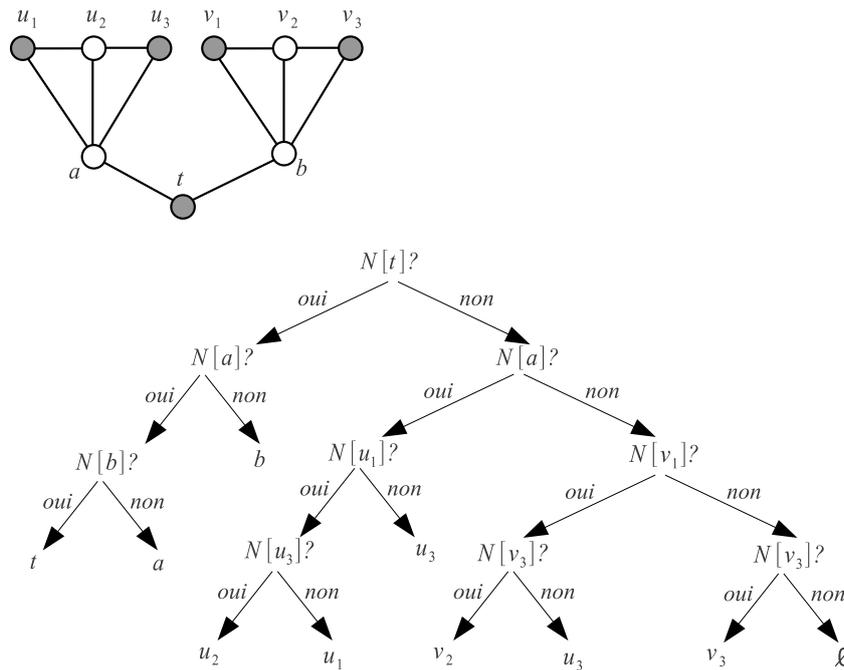
Cet arbre de décision doit avoir la propriété que toutes les réponses conduisant à une feuille  $u$  correspondent à un système de contraintes du type

$$C_u(x) = \begin{cases} x \in N[v] & \text{pour tout nœud } v \text{ conduisant à } u \text{ ayant répondu OUI} \\ x \notin N[v] & \text{pour tout nœud } v \text{ conduisant à } u \text{ ayant répondu NON} \end{cases} \quad (2.11)$$

La variable  $x$  de la contrainte  $C_u(x)$  correspond au sommet défectueux que l'on a à chercher dans le graphe.

Par ailleurs,  $C_u(x)$  doit avoir au plus une solution pour toute feuille  $u$  de l'arbre, et l'on a :

- pour tout sommet  $z$  du graphe, il existe au moins un système de contraintes de type (2.11) admettant  $\{z\}$  comme unique solution,
- il existe au moins un système de contraintes de type (2.11) n'admettant aucune solution.



**FIGURE 2.20 :** Exemple d'arbre de décision correspondant à un code adaptatif. Les nœuds étiquetés «  $N[x]?$  » correspondent aux questions « a-t-on un sommet défectueux dans  $N[x]?$  ». Une feuille est étiquetée avec la solution du système (2.11) correspondant aux nœuds permettant d'accéder à cette feuille. Comme la hauteur de cet arbre est 4, alors on peut toujours déterminer s'il existe un sommet défectueux dans le graphe (et si oui, lequel) en au plus 4 questions. Il est par ailleurs facile de voir que la cardinalité minimum d'un code identifiant (non-adaptatif) dans ce graphe est de 5. Les sommets grisés forment un tel code.

La hauteur minimum d'un tel arbre est notée  $a^{ID}(G)$ , c'est le nombre maximum de questions à poser pour pouvoir déterminer s'il existe un sommet défectueux dans  $G$  (et si oui, lequel).

Clairement, si  $C = \{v_1, \dots, v_k\}$  est un code identifiant de  $G$  (au sens usuel du terme), alors l'arbre binaire complet construit de telle sorte que tous les

nœuds de niveau  $i$  correspondent à la question

« y a-t-il un sommet défectueux dans le voisinage fermé du sommet  $v_i$  ? »

est un code identifiant adaptatif de  $G$ . Ceci montre que l'on a toujours :

**Proposition 11** (Borne générale sur les codes adaptatifs). *Soit  $G$  un graphe admettant un code identifiant. Alors on a :*

$$\alpha^{\text{ID}}(G) \leq \gamma^{\text{ID}}(G). \quad (2.12)$$

Cette inégalité peut être stricte. Par exemple, la famille de graphes décrite en Figure 2.21 ci-dessous admet un code identifiant adaptatif utilisant un nombre maximum de questions logarithmique en le nombre de sommets.

Par ailleurs, pour identifier les sommets d'un graphe de cette famille avec un code identifiant statique, le nombre de sommets nécessaires est linéaire en le nombre de sommets du graphe. On obtient donc une famille de graphes où l'écart entre  $\alpha^{\text{ID}}(G)$  et  $\gamma^{\text{ID}}(G)$  est exponentiel.

Ainsi l'étude des codes identifiants adaptatifs, justifiée par des considérations pratiques, promet des possibilités intéressantes en terme de cardinalité, basées sur l'inégalité (2.12) et sur le fait que celle-ci peut être stricte (voire présenter un écart important, cf. Figure 2.21).

## 2.5.4 Identification adaptative dans les tores

### 2.5.4.1 Torres étudiés

On considère dans cette section des tores dans les grilles carrées et royales.

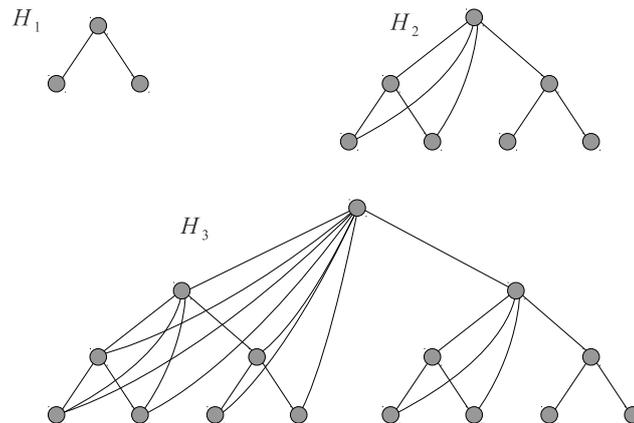
Le tore de dimensions  $p, q$  de la grille carrée est le graphe, noté  $T_{p,q}$ , dont les sommets sont  $\{0, \dots, p-1\} \times \{0, \dots, q-1\}$  et tel qu'il y a une arête entre  $(i, j)$  et  $(i', j')$  si, et seulement si, l'une des deux conditions suivantes est remplie :

- (i)  $i' = i + 1$  et  $j = j'$ ,
- (ii)  $j' = j + 1$  et  $i' = i$ ,

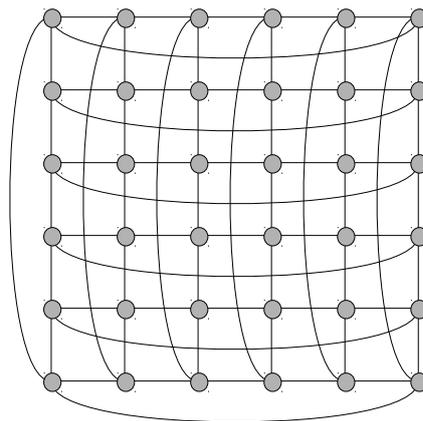
les indices sur  $i, i'$  et sur  $j, j'$  étant considérés modulo  $p$  et modulo  $q$ , respectivement.

Le tore  $T_{p,q}$  peut aussi être vu comme le rotagraphe  $\rho(C_q, p)$  (ou  $\rho(C_p, q)$ ), avec  $C_k$  le cycle à  $k$  sommets (voir Section 2.3.3), ou encore le produit cartésien  $C_p \square C_q$  (voir Figure 2.22).

Le tore de la grille royale, noté  $KT_{p,q}$ , est obtenu à partir du tore de la grille carrée  $T_{p,q}$  en y ajoutant toutes les arêtes croisées du type  $(i, j)(i +$



**FIGURE 2.21 :** Le graphe  $H_1$  est l'arbre à trois sommets, dont la « racine » est le sommet de degré 2. Pour tout  $i \geq 2$ , le graphe  $H_i$  est obtenu en prenant deux copies de  $H_{i-1}$ . Un sommet est alors ajouté, qui fera office de « racine » de  $H_i$ . La « racine » de  $H_i$  est alors reliée à la « racine » d'une des copies de  $H_{i-1}$ , ainsi qu'à tous les sommets de l'autre copie de  $H_{i-1}$ . Comme l'interrogation de la « racine » d'un arbre  $H_k$  sépare les sommets de  $H_k$  en deux sous-ensembles de taille à peu près égale, alors il est facile de construire un code identifiant adaptatif utilisant un nombre maximum de questions logarithmique en le nombre de sommets. Par ailleurs, considérons les « feuilles » d'un graphe  $H_k$ . Deux telles « feuilles » ayant le même « père » ne peuvent clairement être séparées l'une de l'autre que par elles-mêmes, donc tout code identifiant doit nécessairement contenir au moins une de ces « feuilles ». Ainsi, il y approximativement la moitié des « feuilles » de  $H_k$  dans tout code identifiant. Or, celles-ci représentent une fraction linéaire du nombre total de sommets de  $H_k$ .



**FIGURE 2.22 :** Le tore de dimensions 6,6 de la grille carrée, noté  $T_{6,6}$ . Celui-ci peut être vu comme le rotographe  $\rho(C_6, 6)$ , ou comme le produit cartésien  $C_6 \square C_6$ .

$1, j + 1)$  ou  $(i, j)(i + 1, j - 1)$ , les indices étant toujours pris modulo  $p$  et  $q$ , respectivement.

Ce graphe ne peut pas être vu comme un produit cartésien, il peut toutefois lui aussi être vu comme un rotagraphe.

### 2.5.4.2 Bornes générales dans les graphes réguliers

Nous établissons, dans notre premier article avec Yael Ben-Haim, Sylvain Gravier, et Antoine Lobstein [20], des bornes générales sur la cardinalité d'un code  $r$ -identifiant adaptatif dans les graphes réguliers.

Cette borne est basée sur la notion d'empilement. Un *empilement* (*packing* en anglais) dans un graphe  $G = (V, E)$  est un sous-ensemble de sommets  $P \subseteq V$  tels que la distance entre deux points distincts quelconques de  $P$  est d'au moins trois. La cardinalité maximum d'un empilement dans  $G$  est noté  $c(G)$ .

On a toujours  $c(G) \leq \gamma(G)$ , car tout empilement  $P$  montre qu'il faut au moins  $|P|$  sommets pour couvrir  $G$ , un sommet d'un dominant de  $G$  ne pouvant couvrir au plus qu'un des sommets de  $P$ .

Notre borne s'applique aux graphes réguliers tels que, pour tout sommet  $x$ , la cardinalité du voisinage fermé de  $x$  dans  $G^r$  est invariante. Cette quantité est notée  $v(G^r)$ . On a donc par hypothèse  $v(G^r) = |N[x]|$  pour tout sommet  $x$  de  $G^r$ , c'est le volume d'une boule de rayon  $r$  dans  $G$ .

Par ailleurs, on définit  $d_r(G)$  comme étant le nombre minimum de questions (adaptatives) nécessaires pour traiter une boule  $B_r(x)$ , avec  $x$  sommet quelconque du graphe, en supposant qu'il n'y a pas de sommet défectueux en-dehors de  $B_r(x)$  (ainsi, il peut y avoir 0 ou 1 sommet défectueux dans  $B_r(x)$ ).

Notre résultat est le suivant :

**Théorème 53** (Bornes générales dans les graphes réguliers). *Soit  $r \geq 1$  et soit  $G$  un graphe admettant un code  $r$ -identifiant. On suppose, de plus, que  $G$  est régulier, et que pour tout sommet  $x$ , la cardinalité du voisinage fermé de  $x$  dans  $G^r$  est une constante, notée  $v(G^r)$ . Notons de plus  $d_r(G)$  comme ci-dessus. Alors on a*

$$c(G^r) - 1 + \lceil \log_2(v(G^r) + 1) \rceil \leq \alpha^{\text{ID}}(G^r) \leq \gamma(G^r) - 1 + d_r(G).$$

Pour la borne inférieure, on montre que, si les réponses aux  $c(G^r) - 1$  premières questions sont toutes NON, alors il nous reste au moins  $v(G^r)$  sommets qui n'ont pas été couverts par ces  $c(G^r) - 1$  questions. Ainsi, il y a  $v(G^r) + 1$  possibilités : soit l'un de ces  $v(G^r)$  sommets est défectueux, soit il n'y a pas du tout de sommet défectueux dans le graphe. Ces  $v(G^r) + 1$  possibilités nécessitent au moins  $\lceil \log_2(v(G^r) + 1) \rceil$  questions.

Par ailleurs, si  $C$  est un code  $r$ -couvrant de  $G$ , de cardinalité  $\gamma(G^r)$ , alors on peut lui associer une stratégie adaptative comme suit. On interroge les sommets de  $C$  dans un ordre quelconque. Dès lors qu'un de ces sommets  $x$  répond OUI, on sait qu'il existe un sommet défectueux dans  $B_r(x)$ , et il nous faut par définition au plus  $d_r(x)$  questions pour trouver celui-ci. Ainsi, si l'un des  $\gamma(G^r) - 1$  premiers sommets interrogés répond OUI, alors on peut conclure en au plus  $\gamma(G^r) - 1 + d_r(G)$  questions. Si ceux-ci répondent tous NON, alors il nous reste à considérer l'ensemble des sommets non couverts par ces  $\gamma(G^r) - 1$  premières questions, qui est un sous-ensemble de  $B_r(v)$ , avec  $v$  le sommet de  $C$  que nous n'avons pas encore interrogé. Par définition, traiter un sous-ensemble de cette boule  $B_r(v)$  nécessite au plus  $d_r(v)$  questions, et nous avons, dans ce cas, utilisé au plus  $\gamma(G^r) - 1 + d_r(G)$  questions.

### 2.5.4.3 Résultats sur les tores

L'utilité du Théorème 53 réside dans le fait que, lorsque  $p$  et  $q$  sont tous deux multiples de  $2r^2 + 2r + 1$ , alors la puissance du tore  $T_{p,q}^r$  admet un *code parfait*, c'est-à-dire un empilement  $P$  qui est aussi un dominant [91]. En ce cas on a  $c(T_{p,q}^r) = \gamma(T_{p,q}^r)$ , et des bornes plus précises sur  $a^{\text{ID}}(T_{p,q}^r)$  peuvent être obtenues en étudiant  $d_r(T_{p,q}^r)$ .

Nous obtenons alors les valeurs exactes de  $a^{\text{ID}}(T_{p,q}^r)$  pour de nombreuses valeurs de  $p$ , et  $q$ .

De façon générale, on montre que l'on a [20] :

**Théorème 54** (Tores de la grille carrée). *Soit  $T_{p,q}$  le tore de la grille carrée de dimensions  $p, q$ , et soit  $r \geq 1$ . Alors on a :*

$$a^{\text{ID}}(T_{p,q}^r) = \frac{pq}{2r^2 + 2r + 1} + \Theta(p + q).$$

En particulier, la densité d'un code adaptatif dans la grille carrée infinie est égal à  $\frac{1}{2r^2 + 2r + 1}$ , ce qui est meilleur que la densité d'un code identifiant (non-adaptatif). En effet, on sait par exemple pour  $r = 1$  que la densité optimale d'un code (non-adaptatif) est 0,35 [22], à comparer à 0,2 pour un code adaptatif. De même, pour  $r > 1$ , on sait que la densité d'un code (non-adaptatif) est supérieure à  $\frac{3}{8r+4}$  [44].

Le cas du tore dans la grille royale est analogue. En effet, lorsque  $p$  et  $q$  sont tous deux multiples de  $2r + 1$ , alors la puissance du tore  $KT_{p,q}^r$  admet un code parfait. Ceci nous permet d'obtenir des bornes générales plus fines que celles du Théorème 53 dans le cas du tore de la grille royale.

Par ailleurs, le problème est apparenté à un jeu de Rényi étudié par Miklós Ruzinkó [160]. Le jeu étudié par Ruzinkó est le suivant. Le joueur Carole choisit un sommet mystère dans le rectangle

$$R_{a,b} = \{(x, y) \mid 1 \leq x \leq a, 1 \leq y \leq b\} \subset \mathbb{R}^2.$$

Le joueur Paul ne peut poser que des questions du type

« le sommet mystère est-il dans le rectangle  $R_{r,s}$  ? »

avec  $r$  et  $s$  paramètres entiers de son choix.

Ce jeu s'apparente à notre problème d'identification adaptative dans le sens où, lorsque  $a = b = 2r + 1$ , le rectangle  $R_{a,b}$  n'est autre que la boule de rayon  $r$  centrée en  $(r, r)$ . Les questions autorisées à Paul forment alors un sous-ensemble des questions légales pour le problème d'identification adaptative dans la grille royale. Les résultats de Ruzinkó sur ce jeu nous permettent d'obtenir des bornes très fines sur  $d_r(KT_{p,q})$ .

Pour un grand nombre de valeurs de  $p$  et  $q$ , nous obtenons la valeur exacte de  $\alpha^{\text{ID}}(KT_{p,q}^r)$ .

De façon générale, nous montrons que l'on a [21] :

**Théorème 55** (Tores de la grille royale). *Soit  $KT_{p,q}$  le tore de la grille royale de dimensions  $p, q$ , et soit  $r \geq 1$ . Alors on a :*

$$\alpha^{\text{ID}}(KT_{p,q}^r) = \frac{pq}{(2r+1)^2} + \Theta(p+q).$$

Par conséquent, la densité d'un code adaptatif dans la grille carrée infinie est égal à  $\frac{1}{(2r+1)^2}$ , à comparer à  $\frac{2}{9}$  pour le cas d'un code identifiant non adaptatif avec  $r = 1$  [48], et à  $\frac{1}{4r}$  dans le cas  $r > 1$  [45].

Ville Junnila étend ces résultats dans un manuscrit soumis à *Discrete Mathematics and Theoretical Computer Science* [117], où il étudie l'identification de sous-ensembles de sommets, c'est-à-dire les codes  $(r, \leq \ell)$ -identifiants adaptatifs.

Si l'on note  $\alpha_\ell^{\text{ID}}$  l'identification adaptative de sous-ensembles d'au plus  $\ell$  sommets, alors Ville Junnila montre :

**Théorème 56** (Tores pour  $\ell > 1$ ). *Soit  $T_{p,q}$  (resp.  $KT_{p,q}$ ) le tore de la grille carrée (resp. royale) de dimensions  $p, q$ , et soit  $r \geq 1$ . Alors on a :*

$$\alpha_2^{\text{ID}}(KT_{p,q}^r) = \frac{pq}{(2r+1)^2} + \Theta(\ln r).$$

Si  $p$  et  $q$  sont tous deux multiples de 5, on a :

$$\alpha_2^{\text{ID}}(\mathbb{T}_{p,q}) = \frac{pq}{5} + F(p, q),$$

et

$$\alpha_3^{\text{ID}}(\mathbb{T}_{p,q}) = \frac{pq}{5} + G(p, q),$$

avec  $F(p, q) \in [3, 6]$  et  $G(p, q) \in [4, 9]$  pour tout  $p, q$ .

Dans le cas  $\ell \geq 3$ , l'identification adaptative est impossible dans les tores de la grille royale. De même, le cas  $\ell \geq 4$  est trivial pour les tores de la grille carrée.

### 2.5.5 Autres structures

Ville Junnila étudie de plus dans sa thèse [116] et dans son manuscrit soumis à *Discrete Mathematics and Theoretical Computer Science* [117] l'identification adaptative dans les hypercubes ainsi que dans les cycles.

Pour les hypercubes, il donne en particulier les bornes suivantes :

**Théorème 57** (Identification adaptative dans l'hypercube). *Soit  $Q_n$  l'hypercube de dimension  $n$ . Alors on a :*

$$c(Q_n) + \left\lceil \frac{n+1}{8} \right\rceil \leq \alpha^{\text{ID}}(Q_n) \leq \gamma(Q_n) + \left\lfloor \frac{n+1}{2} \right\rfloor.$$

Dans le cas  $\ell > 1$ , il donne de plus :

**Théorème 58** (Identification adaptative dans l'hypercube, cas  $\ell > 1$ ). *Soit  $Q_n$  l'hypercube de dimension  $n = 2^s - 1$ , avec  $s \geq 3$ . Soit de plus  $1 < \ell < \frac{n}{6} + 1$ . Alors on a :*

$$\alpha_\ell^{\text{ID}}(Q_n) \leq c(Q_n) + \ell \frac{n+3}{2}.$$

Ce dernier résultat est obtenu en considérant le fait que, lorsque  $n = 2^s - 1$ , alors  $Q_n$  admet un code parfait [58].

En ce qui concerne les cycles, seule l'identification de sous-ensembles d'au plus  $\ell \leq 2$  sommets est pertinente. Ville Junnila obtient les résultats suivants :

**Théorème 59** (Identification adaptative dans les cycles). *Soit  $C_n$  le cycle à  $n$  sommets, et soit  $r \geq 1$ . Si  $n = 2r + 1 + k$ , avec  $1 \leq k \leq 2r$ , alors on a :*

$$\alpha^{\text{ID}}(C_n^r) = \left\lfloor \frac{2r+1}{k} \right\rfloor + \left\lceil \log_2 \left( 2r+1 - k \left( \left\lfloor \frac{2r+1}{k} \right\rfloor - 1 \right) \right) \right\rceil.$$

Si  $n \geq 2(2r + 1)$ , alors on a :

$$\alpha^{\text{ID}}(C_n^r) = \left\lfloor \frac{n}{2r+1} \right\rfloor - 1 + \left\lceil \log_2 \left( n - \left( \left\lfloor \frac{n}{2r+1} \right\rfloor - 1 \right) (2r+1) + 1 \right) \right\rceil.$$

Si  $n \geq 3(2r + 1)$ , alors on a :

$$\alpha_2^{\text{ID}}(C_n^r) \leq \left\lfloor \frac{n}{2r+1} \right\rfloor + 1 + 2 \lceil \log_2(2r+1) \rceil.$$

# Conclusion et perspectives

---

## Sommaire

---

<b>3.1 Bilan des principaux résultats</b>	<b>107</b>
<b>3.2 Perspectives de recherche</b>	<b>108</b>
3.2.1 Questions ouvertes de la littérature	108
3.2.2 Thématiques de recherche émergentes	117
<b>3.3 Projet personnel</b>	<b>124</b>
3.3.1 Codes identifiants	124
3.3.2 Ordonnancement	125

---

## 3.1 Bilan des principaux résultats

Ce manuscrit s'appuie sur un ensemble d'une quinzaine d'articles publiés entre 2004 et 2012 sur les codes identifiants dans les graphes. Ces publications sont parues dans neuf journaux de référence en mathématiques discrètes et en combinatoire.

Les résultats présentés s'articulent autour de quatre thématiques.

La première thématique développée en Section 2.2 concerne l'étude des codes identifiants dans des *structures régulières*. Les résultats forts sont ceux de l'article sur les cycles [98] (voir la Section 2.2.2), dans lequel est introduit un graphe auxiliaire qui est à la base des derniers développements de la littérature sur le sujet, ainsi que la résolution de la conjecture de monotonie dans les hypercubes de Blass, Honkala et Litsyn (voir la Conjecture 1) pour le cas  $r = 1$  [143].

En Section 2.3 sont présentés des résultats concernant des *questions algorithmiques*. Les résultats forts sont ceux concernant l'approximabilité du problème [93] (voir la Section 2.3.5), ainsi que les développements récents de méthodes génériques de résolution de problèmes combinatoires dans les grilles, qui va au-delà du cas des codes identifiants [35] (voir la Section 2.3.4).

Les *questions structurelles* sur le sujet font l'objet de la Section 2.4. Les résultats forts concernent la construction de familles de graphes admettant des codes  $\ell$ -set-identifiants de faible cardinalité, et ont été publiés dans trois articles [83, 95, 145]. Ces articles s'appuient, respectivement, sur des méthodes probabilistes [83], des outils provenant de la théorie des plans projectifs [95], ainsi que sur un algorithme de Hwang et Sós concernant les codes superposés [145].

Enfin, l'introduction de la problématique d'*identification adaptative* est présentée en Section 2.5. Cette thématique de recherche émergente a fait l'objet de deux publications récentes [20, 21], dans lesquels sont, entre autres, posées les bases de l'identification adaptative dans les graphes réguliers (Théorème 53). Un premier lien avec les jeux de Rényi est, de plus, établi en [21].

## 3.2 Perspectives de recherche

Je présente ici quelques pistes possibles de recherche sur les codes identifiants dans les graphes. Ces pistes concernent, d'une part, l'extension de résultats existants de la littérature et la résolution (ou l'infirmité) de quelques conjectures qui résistent encore (voir la Section 3.2.1), et d'autre part, le développement de nouvelles thématiques de recherche qui me semblent prometteuses (voir la Section 3.2.2).

### 3.2.1 Questions ouvertes de la littérature

#### 3.2.1.1 La conjecture de monotonie dans les hypercubes

Rappelons ici la conjecture de Blass, Honkala et Litsyn [28] au sujet de la monotonie de la cardinalité minimum d'un code  $r$ -identifiant dans l'hypercube :

**Conjecture** (Monotonie dans l'hypercube). *Soit  $r \geq 1$ , et soit  $n_r$  le plus petit  $n$  tel que  $Q_n$  admette un code  $r$ -identifiant. Alors, pour tout  $n \geq n_r$ , on a*

$$\gamma^{\text{ID}}(Q_n^r) \leq \gamma^{\text{ID}}(Q_{n+1}^r).$$

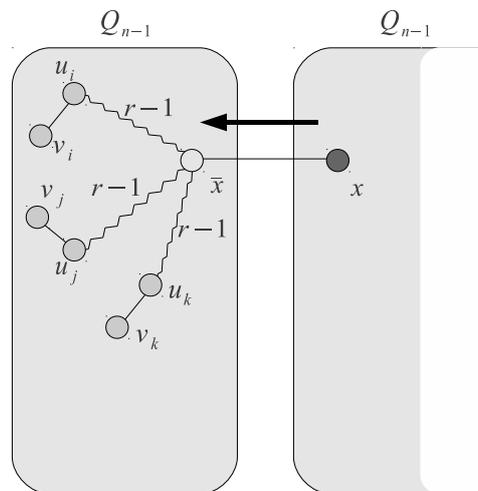
À ce jour, seul le cas  $r = 1$  a été traité, cas dans lequel la conjecture est vraie [143]. L'argument réside dans le fait que, si l'on voit l'hypercube  $Q_n$  comme deux copies de  $Q_{n-1}$  reliées par un couplage parfait, alors on peut projeter un code identifiant  $C$  de  $Q_n$  sur  $Q_{n-1}$ .

Formellement, pour un sommet  $u = x_1 \dots x_n \in \{0, 1\}^n$  de  $Q_n$ , on définit le sommet  $\bar{u}$  de  $Q_{n-1}$  comme  $x_1 \dots x_{n-1}$ . Ainsi, le projeté d'un code identifiant  $C$  de  $Q_n$  est  $\bar{C} = \{\bar{u} \mid u \in C\}$ .

Si le code  $\bar{C}$  obtenu sur  $Q_{n-1}$  n'est pas nécessairement un code identifiant, on peut cependant montrer qu'à chaque sommet  $\bar{u}$  de  $\bar{C}$  correspond, au plus, une paire  $(x, y)$  de sommets de  $Q_{n-1}$  qui ne sont pas séparés.

L'une des raisons pour lesquelles le cas général  $r \geq 2$  est plus difficile est que, lorsque l'on projette un code  $r$ -identifiant  $C$  de  $Q_n$  sur  $Q_{n-1}$ , chaque sommet de  $\bar{C}$  peut correspondre à *plusieurs* paires de sommets de  $Q_{n-1}$  qui ne sont pas séparés (voir la Figure 3.1). En ce cas, l'argument du cas  $r = 1$ , qui consiste à altérer localement le code  $\bar{C}$  pour régler les cas des sommets non séparés, ne fonctionne plus, car on est confronté à la difficile question de résoudre simultanément plusieurs problèmes avec un même sommet.

Par ailleurs, la voie de la projection est celle qui semble la plus naturelle vis-à-vis de cette conjecture.



**FIGURE 3.1 :** Tentative de projection de code dans le cas général  $r \geq 2$ . Les traits pleins représentent des arêtes, et les « zig-zag » des chemins de longueur  $r - 1$ . Un sommet  $x$  du code est projeté sur un sommet  $\bar{x}$  dans  $Q_{n-1}$ . Le problème est que, dans  $Q_n$ , ce sommet  $x$  pouvait servir à séparer plusieurs paires de sommets (ici  $u_i v_i$ ,  $u_j v_j$ , et  $u_k v_k$ ). Lorsque  $x$  est projeté dans  $Q_{n-1}$ , le sommet  $\bar{x}$  obtenu est alors à distance au plus  $r$  de tous les sommets que  $x$  séparait. Le casse-tête consiste à trouver comment déplacer  $\bar{x}$  de sorte à séparer toutes ces paires de sommets qui ne sont plus séparés.

### 3.2.1.2 La conjecture de duplication dans les hypercubes

La duplication de code consiste à dire que, en terme de cardinalité, un code  $r$ -identifiant minimum de  $Q_{n+1}$  est de cardinalité au plus deux fois celle d'un code de  $Q_n$  :

**Conjecture 5** (Duplication de code). *Pour tout  $n \geq 2$ , on a*

$$\gamma^{\text{ID}}(Q_{n+1}^r) \leq 2\gamma^{\text{ID}}(Q_n^r).$$

À ce jour, le meilleur résultat connu à ce sujet est celui d'Exoo, Junnila, Laihonen et Ranto [72] :

**Théorème** (Quasi-duplication). *Pour tout  $n \geq 1$ , on a*

$$\gamma^{\text{ID}}(Q_{n+1}) \leq \left(2 + \frac{1}{n+1}\right) \gamma^{\text{ID}}(Q_n).$$

Une première piste pour démontrer cette conjecture provient de l'observation de Blass, Honkala et Litsyn [29], qui consiste à dire que dupliquer un code identifiant  $C$  de  $Q_n$  donne un code identifiant de  $Q_{n+1}$  si, et seulement si, le code  $C$  ne contient pas de *sommet isolé* (un sommet isolé étant un sommet de  $C$  n'ayant aucun voisin dans  $C$ ). La question qui se pose est alors la suivante :

**Question 1** (Existence de sommets isolés). *Pour tout  $n \geq 4$ , existe-t-il un code identifiant optimum de  $Q_n$  n'admettant aucun sommet isolé ?*

La deuxième piste qui permettrait de contourner le problème des sommets isolés (par exemple si la réponse à la question ci-dessus était négative), pourrait consister à étudier les « permutations » de codes. L'idée serait, au lieu de *dupliquer* un code optimal  $C$  de  $Q_n$ , d'*altérer* celui-ci en un code identifiant optimal  $\tilde{C}$  de  $Q_n$ , et d'obtenir un code identifiant de  $Q_{n+1}$  comme la juxtaposition de  $C$  et de  $\tilde{C}$  dans deux copies de  $Q_n$ .

Une question possible pourrait être, par exemple :

**Question 2** (Permutation de code). *Soit  $C$  un code identifiant optimum de  $Q_n$ . Existe-t-il un cycle hamiltonien  $\mathcal{C}$  de  $Q_n$ , tel que, si l'on numérote  $v_1, \dots, v_{2^n}$  les sommets de ce cycle hamiltonien  $\mathcal{C}$  (dans un ordre tel que  $v_i$  et  $v_{i+1}$  sont adjacents pour tout  $i$ , modulo  $2^n$ ) alors le code  $C(\mathcal{C}^{(+1)})$  est, lui aussi, un code identifiant optimum de  $Q_n$  (où  $C(\mathcal{C}^{(+1)})$  est défini comme  $\{v_{i+1} \mid v_i \in C\}$ ) ?*

### 3.2.1.3 Complexité de l'identification d'ensembles de sommets

Considérons le problème général suivant :

$(r, \ell)$ -ID CODE

Instance : Un graphe  $G$  et un entier  $k \geq 1$ .

Question : Existe-t-il un code  $(r, \ell)$ -identifiant de  $G$  de cardinalité inférieure ou égale à  $k$  ?

Pour  $\ell = 1$ , on sait que le problème est NP-complet pour tout  $r \geq 1$  [49]. Lorsque  $\ell \geq 2$ , seuls des résultats partiels sont connus. En particulier, lorsque  $\ell = 2$ , on sait que le problème est NP-complet, pour tout  $r \geq 1$  [11] :

**Théorème** (Identification d'ensembles d'au plus 2 sommets). *Le problème  $(r, \leq 2)$ -ID CODE est NP-complet pour tout  $r \geq 1$ .*

À ma connaissance aucun autre résultat similaire n'est connu pour  $\ell \geq 3$ . Il me semblerait utile de montrer que  $(r, \ell)$ -ID CODE est NP-complet pour tout  $r, \ell$ .

**Question 3** (Complexité de l'identification d'ensembles de sommets). *Est-il vrai que le problème  $(r, \leq 2)$ -ID CODE est NP-complet pour tout  $r, \ell \geq 1$  ?*

Si l'on adopte la voie d'Auger *et al*, consistant à effectuer une réduction depuis le problème TRANSVERSAL, la difficulté réside dans le fait de trouver, pour tout  $r, \ell$ , un gadget  $g_{r, \ell}$  que l'on va substituer à toute arête du graphe. En un sens, ce gadget  $g_{r, \ell}$  devra être tel que l'essentiel de la complexité du problème  $(r, \leq \ell)$ -ID CODE sera concentrée en  $g_{r, \ell}$ . Le problème principal est que la structure même des graphes admettant des codes  $(r, \leq \ell)$ -identifiants est encore relativement mal connue (voir en particulier la section suivante).

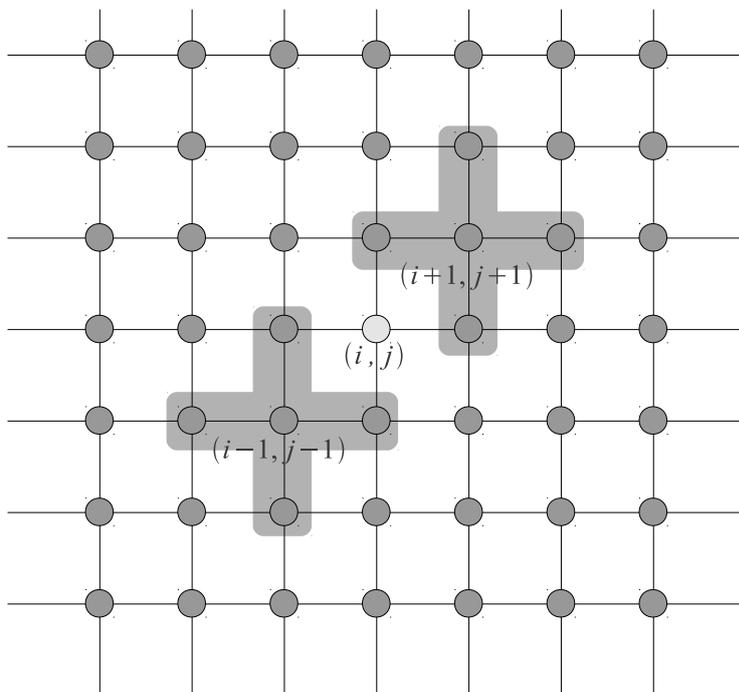
### 3.2.1.4 Borne inférieure serrée pour $\gamma_{\ell}^{\text{ID}}(G)$

Étant donné un graphe  $G$  à  $n$  sommets admettant un code  $(r, \leq \ell)$ -identifiant, on dispose aujourd'hui des bornes suivantes sur la cardinalité d'un code :

$$\Theta\left(\frac{\ell^2}{\ln \ell} \ln n\right) \leq \gamma_{\ell}^{\text{ID}}(G^r) \leq n.$$

La borne supérieure est serrée, ne serait-ce qu'à cause du graphe sans arêtes. Il existe, par ailleurs, des graphes finis  $G = (V, E)$  ayant des arêtes qui n'admettent que  $V$  comme code  $(r, \leq \ell)$ -identifiant — et c'est là une première différence essentielle avec le cas  $\ell = 1$ . C'est, par exemple, le cas du tore de la grille carrée, qui n'admet que l'ensemble de tous ses

sommets comme code 3-set-identifiant. En effet, si un sommet  $(i, j)$  était absent du code, alors les sous-ensembles  $\{(i+1, j+1), (i-1, j-1)\}$  et  $\{(i+1, j+1), (i-1, j-1), (i, j)\}$  ne seraient pas séparés (voir Figure 3.2).



**FIGURE 3.2 :** Pour  $k$  et  $n$  assez grands, le tore de la grille carrée  $T_{k,n}$  n'admet que l'ensemble de ses sommets comme code 3-set-identifiant. En effet, si un sommet  $(i, j)$  était absent du code, alors les sous-ensembles  $\{(i+1, j+1), (i-1, j-1)\}$  et  $\{(i+1, j+1), (i-1, j-1), (i, j)\}$  ne seraient pas séparés. De même, les sous-ensembles  $\{(i+1, j-1), (i-1, j+1)\}$  et  $\{(i+1, j-1), (i-1, j+1), (i, j)\}$  ne seraient pas séparés. La raison en est que le voisinage (ouvert) du sommet  $(i, j)$  est inclus dans l'union des voisinages fermés de  $(i+1, j+1)$  et  $(i-1, j-1)$  (respectivement  $(i+1, j-1)$  et  $(i-1, j+1)$ ). Pour cette même raison, le tore  $T_{k,n}$  n'admet d'ailleurs pas de code  $\ell$ -set-identifiant pour  $\ell \geq 4$ .

On ignore, à l'heure actuelle, ce qu'il en est pour la borne inférieure. Les meilleures constructions de familles de graphes admettant des codes  $\ell$ -set-identifiants de faible cardinalité [83, 145] fournissent des graphes admettant un code  $\ell$ -set-identifiant de cardinalité  $O(\ell^2 \ln n)$ , avec  $n$  le nombre de sommets du graphe. Il serait intéressant de pouvoir connaître des bornes inférieures serrées, ne serait-ce que dans des cas particuliers.

**Question 4 (Borne inférieure serrée).** Est-il possible de déterminer, même pour quelques valeurs particulières de  $\ell \geq 2$ , une borne inférieure serrée de  $\gamma_\ell^{\text{ID}}(G)$ , en fonction du nombre de sommets de  $G$  ?

Le cas des codes identifiants est, en ce sens, analogue à celui des codes superposés, où il persiste depuis plusieurs dizaines d'années un écart entre la meilleure borne et les meilleures constructions connues (voir Section suivante).

### 3.2.1.5 Extension des liens avec les codes superposés

Karpovsky *et al* annoncent que les codes identifiant des ensembles de sommets sont reliés aux codes superposés.

Nous l'avons vu, ce lien, quoique intuitif, est loin d'être évident à établir de façon formelle (voir Section 2.4.2). Tout au plus, nous savons que ce lien est fort dans le cas des codes identifiants dans les graphes orientés (voir la Proposition 6). Par ailleurs, nous avons vu comment utiliser des codes  $\ell$ -superposés maximaux pour construire des codes  $\ell$ -set-identifiants quasi-optimaux.

Il serait cependant intéressant de pouvoir développer des résultats généraux sur cette correspondance entre les codes identifiants et les codes superposés.

**Question 5** (Correspondance codes identifiants – codes superposés dans le cas non orienté). *Peut-on établir une correspondance formelle, à l'image de celle de la Proposition 6, entre les codes  $\ell$ -superposés et les codes  $\ell$ -set-identifiants dans les graphes non orientés ?*

### 3.2.1.6 Liens avec d'autres notions de codes

On peut, de plus, se poser la question de savoir si la notion de code superposé est la « bonne » notion à mettre en regard avec celle de code identifiant. Si l'on regarde les définitions, la notion de code superposé correspond clairement à celle de code identifiant. Cependant, nous l'avons vu, en terme de preuve, nous peinons à établir une correspondance formelle « propre » entre les deux notions.

Or, il existe d'autres notions gravitant autour de celle de code superposé. C'est, par exemple, le cas de la notion de famille  $\ell$ -part-intersecting, introduite par Hwang et Sós. Une famille  $\mathcal{F}$  de sous-ensembles d'un ensemble  $S$  est dite  $\ell$ -part-intersecting si l'on a

$$|A \cap B| < \frac{1}{\ell} \min(|A|, |B|)$$

pour tout  $A, B \in \mathcal{F}$ ,  $A \neq B$ . La dimension de cette famille est alors par définition la cardinalité de l'ensemble  $S$ .

Si la terminologie *l-part-intersecting* est donnée par Ruszinkó [161], la notion semble avoir été introduite par Hwang et Sós [113].

Une autre notion intéressante est celle de famille *l-cover-free*. Une famille  $\mathcal{F}$  de sous-ensembles d'un ensemble  $S$  est dite *l-cover-free* si l'on a

$$A_0 \not\subseteq (A_1 \cup A_2 \cup \dots \cup A_\ell)$$

pour tout  $A_0, A_1, \dots, A_\ell \in \mathcal{F}$ ,  $A_i \neq A_j \forall i \neq j$ . La dimension d'une telle famille est égale à la cardinalité de l'ensemble  $S$ .

Cette notion, tout comme celle de code superposé, a été introduite par Kautz et Singleton [122].

Une famille qui est *l-part-intersecting* est aussi *l-cover-free*, car, dès lors qu'un sous-ensemble  $A_0$  est inclus dans une union  $A_1 \cup \dots \cup A_\ell$ , cela implique qu'il existe au moins un  $A_i$ ,  $i \geq 1$ , tel que  $A_i$  contienne au moins une fraction  $\frac{1}{\ell}$  des éléments de  $A_0$ .

De même, une famille qui est *l-cover-free* est aussi un code *l-superposé*. Par ailleurs, un code *l-superposé* est nécessairement aussi une famille  $(\ell - 1)$ -*cover-free*. En effet, si l'on a  $A_0 \subseteq A_1 \cup \dots \cup A_{\ell-1}$ , alors cela implique  $A_0 \cup A_1 \cup \dots \cup A_{\ell-1} = A_1 \cup \dots \cup A_{\ell-1}$ .

Ainsi, si l'on note  $T'(\ell, n)$  (respectivement,  $T(\ell, n)$  et  $T^*(\ell, n)$ ) la cardinalité maximum d'un code *l-superposé* (respectivement, d'une famille *l-cover-free*, d'une famille *l-part-intersecting*) de dimension  $n$ , alors on a :

$$T^*(\ell, n) \leq T(\ell, n) \leq T'(\ell, n) \leq T(\ell - 1, n). \quad (3.1)$$

Ce qui peut nous faire douter du fait que la notion de code superposé soit la « bonne » notion à considérer provient du fait que, si l'on regarde la littérature, les résultats déjà mentionnés sur les codes superposés (voir le Théorème 43 au sujet des bornes générales sur la cardinalité d'un tel code) ont été en fait obtenus *via* d'autres notions.

Le résultat de Ruszinkó [161], par exemple, a été obtenu grâce à un résultat sur les familles *l-cover-free* (et *via* l'équation (3.1)). En ce qui concerne l'algorithme de Hwang et Sós [113], la notion considérée est celle de famille *l-part-intersecting*. C'est d'ailleurs cette même notion que j'utilise dans [145] pour obtenir un algorithme similaire pour les codes *l-set-identifiants*.

Ainsi, il semble que des notions connexes à celle de code superposé soient utiles pour dériver des résultats sur ces codes. D'un certain point de vue, la notion de famille *cover-free* est d'ailleurs équivalente à celle de code superposé. En effet, si l'on reprend (3.1) à la lueur du Théorème 43,

on constate que, selon la dépendance en  $\ell$  de  $T'(\ell, n)$  et  $T(\ell, n)$ , on peut considérer que l'on a

$$T'(\ell, n) \simeq T(\ell, n),$$

dès lors que  $T(\ell, n) \simeq T(\ell - 1, n)$ . Asymptotiquement, la notion de code superposé est donc équivalente à celle de famille *cover-free*.

Un exemple ultime de notion asymptotiquement équivalente à celle de code superposé, est la notion de famille  $\ell$ -*parity-check*, définie comme suit. Une famille  $\mathcal{F}$  de sous-ensembles de  $S$  est dite  $\ell$ -*parity-check*, si l'on a

$$A_1 \cup \dots \cup A_s = B_1 \cup \dots \cup B_t \quad \Rightarrow \quad s \equiv t \pmod{2}$$

pour tout  $s, t \leq \ell$ , avec les  $A_i$  tous distincts et les  $B_j$  tous distincts. Autrement dit, si l'on connaît l'union d'au plus  $\ell$  éléments de  $\mathcal{F}$ , alors on peut déterminer la *parité* du nombre d'éléments considérés. La dimension d'une telle famille est la cardinalité de  $S$ .

D'un certain point de vue, cette notion est plus faible que celle de code  $\ell$ -superposé, dans le sens où, pour un code  $\ell$ -superposé, la connaissance de l'union d'au plus  $\ell$  éléments de  $\mathcal{F}$  nous permet de déterminer qui sont les éléments considérés dans cette union. Ainsi, un code  $\ell$ -superposé est nécessairement aussi une famille  $\ell$ -*parity-check*.

Par ailleurs, cette notion est également, en un sens, plus forte que celle de code superposé. En effet, un famille  $\ell$ -*parity-check* est aussi toujours un code  $(\ell - 1)$ -superposé. Pour s'en convaincre, considérons  $\{A_1, \dots, A_s\}$  et  $\{B_1, \dots, B_t\}$  tels que  $s, t \leq \ell - 1$  et

$$A_1 \cup \dots \cup A_s = B_1 \cup \dots \cup B_t.$$

Si  $s \equiv t \pmod{2}$ , alors on a fini, et la paire  $\{A_1, \dots, A_s\}, \{B_1, \dots, B_t\}$  constitue un obstacle à ce que la famille considérée soit  $\ell$ -*parity-check*. Si  $s \not\equiv t \pmod{2}$ , alors il existe  $j$  tel que la paire  $\{B_j, A_1, \dots, A_s\}, \{B_1, \dots, B_t\}$  constitue un obstacle à ce que la famille considérée soit  $\ell$ -*parity-check*.

Ainsi, dès lors que l'on a  $T(\ell, n) \simeq T(\ell - 1, n)$ , alors on peut considérer que

$$T^\bullet(\ell, n) \simeq T(\ell, n),$$

avec  $T^\bullet(\ell, n)$  la cardinalité maximum d'une famille  $\ell$ -*parity-check* de dimension  $n$ .

La notion de famille  $\ell$ -*parity-check* et les observations ci-dessus ont été formulées par Miklós Ruszinkó et moi-même en 2006 (travail non publié).

**Question 6.** *Les notions de familles part-intersecting, cover-free, ou parity-check, sont-elles plus appropriées que la notion de code  $\ell$ -superposé pour dériver des bornes générales sur la cardinalité maximum d'un tel code ?*

### 3.2.1.7 Unification de ces différentes notions de codes

Une première piste pour l'unification de toutes ces notions au sein d'un même cadre théorique est celle développée par Sylvain Gravier et Bernard Ycart (Laboratoire Jean Kuntzmann, Grenoble), dans leur article de 2006 publié à *Discrete Mathematics and Theoretical Computer Science Proceedings* [101].

Dans cet article, ils introduisent la notion de matrice  $S$ -contrainte. Soit  $q, d, s \geq 1$ , et soit  $S = \{\eta_1, \dots, \eta_s\}$  un ensemble de vecteurs de  $\{0, \dots, q-1\}^d$ . Soient  $m \geq d$  et  $n \geq s$ , et soit  $M$  une matrice  $m \times n$  dont les éléments sont dans  $\{0, \dots, q-1\}$ . La matrice  $M$  est alors dite  $S$ -contrainte si, pour tout  $J \subseteq \{1, \dots, n\}$ ,  $|J| = d$ , il existe  $(i_1, \dots, i_s) \in \{1, \dots, m\}^s$  tels que, pour tout  $h = 1, \dots, s$  et tout  $j \in J$ , on a

$$M_{i_h, j} = \eta_h(j).$$

En d'autres termes, tout sous-ensemble de  $d$  colonnes de  $M$  contient, à permutation près, une sous-matrice dont les lignes sont les vecteurs de  $S$ .

Les colonnes de  $M$  peuvent être considérées comme les éléments d'un code  $q$ -aire, et l'ensemble  $S$  modélise des contraintes qui s'appliquent sur tout sous-ensemble de  $d$  mots du code.

Lorsque  $q = 2$ , ces éléments d'un code binaire peuvent être également vus comme les vecteurs caractéristiques des éléments d'une famille de sous-ensembles d'un ensemble à  $m$  éléments.

Considérons, par exemple,  $q = 2, d = \ell + 1$ , et  $S$  l'ensemble de tous les vecteurs binaires comportant un seul 1. En ce cas, tout sous-ensemble  $X$  de  $\ell + 1$  colonnes de  $M$  doit contenir  $S$  comme sous-matrice (à permutation près). Ceci implique que, pour toute colonne  $c$  de  $X$ , il existe un indice  $i$  tel que  $c_i = 1$  et  $c'_i = 0$  pour toute colonne  $c' \in X, c' \neq c$ . Si l'on voit les colonnes de  $X$  comme les vecteurs caractéristiques des éléments d'une famille de sous-ensembles d'un ensemble à  $m$  éléments, alors ceci implique que le sous-ensemble correspondant à  $c$  n'est pas inclus dans l'union des sous-ensembles correspondant aux  $c'$ . Autrement dit, les vecteurs colonne de  $M$  forment une famille  $\ell$ -cover-free.

Inversement, les vecteurs caractéristiques de toute famille  $\ell$ -cover-free forment une matrice  $S$ -contrainte, avec  $S$  comme ci-dessus. Il existe donc  $S$  tel qu'une famille  $\mathcal{F}$  est  $\ell$ -cover-free si, et seulement, la matrice des vecteurs caractéristiques des éléments de  $\mathcal{F}$  est  $S$ -contrainte.

Cette notion de matrice  $S$ -contrainte constitue une unification partielle des notions évoquées dans la Section précédente. En effet, pour ce qui concerne les codes superposés, il n'existe pas de  $S$  tel que  $\mathcal{F}$  est *cover-free*

si, et seulement si, la matrice des vecteurs caractéristiques des éléments de  $\mathcal{F}$  est  $S$ -contrainte (voir la Proposition 1 de [101]).

Comme indiqué dans la conclusion de l'article de Gravier et Ycart, une extension possible de cette notion serait de considérer des matrices  $M$  dont tout sous-ensemble de  $d$  colonnes contiendrait une copie d'une certaine sous-matrice  $N$ . L'intérêt de cette extension serait de nous permettre d'exiger que  $M$  contienne un nombre prescrit de copies de vecteurs donnés, ce qui n'est pas possible pour les matrices  $S$ -contraintes. Ceci pourrait permettre de capturer la notion de famille *part-intersecting*.

De même, ils suggèrent de considérer une variante de cette notion où l'on fixerait plusieurs ensembles  $S_1, \dots, S_k$ , et où l'on exigerait d'une matrice  $M$  de contenir au moins une copie (à permutation près) d'un des ensemble  $S_1, \dots, S_k$ .

Il est à noter que ces notions de matrices contraintes capturent également d'autres notions de la littérature, comme celle de la dimension de Vapnik-Chervonenkis, et de *testing dimension* d'une famille de fonctions (voir l'introduction de l'article de Gravier et Ycart [101]).

**Question 7** (Unification de différentes notions de codes). *Peut-on définir une notion similaire à celle de matrice  $S$ -contrainte, qui permettrait de capturer et d'unifier les notions de codes identifiant, superposé, et de famille cover-free, part-intersecting, etc. dans un cadre théorique permettant de dériver des résultats similaires à ceux du Théorème 43 ?*

## 3.2.2 Thématiques de recherche émergentes

### 3.2.2.1 Identification adaptative

Introduit dans ma thèse, et étudié par la suite dans trois articles [20, 21, 117], l'identification adaptative avait partiellement fait l'objet du projet ANR IDEA.

La théorie n'en est qu'à ses balbutiements. En effet, l'identification adaptative est considérablement plus difficile à appréhender que l'identification non-adaptative. En effet, un code identifiant non-adaptatif est un simple sous-ensemble de sommets, alors qu'un code adaptatif est une stratégie complète de jeu, décrite par un arbre de décision. Ainsi, les raisonnements habituels sur les codes en tant que sous-ensembles se retrouvent inopérants.

Par exemple, on sait qu'un sommet forcé  $x$  apparaît nécessairement dans tout code (rappelons que  $x$  est dit *forcé* s'il existe  $u, v$  tels que  $N[u] =$

$N[v] \cup \{x\}$ ). Pour un code identifiant non-adaptatif, on sait que ce sommet forcé va nous « coûter » 1, dans le sens où l'on a toujours

$$\gamma^{\text{ID}}(G) \geq f(g),$$

avec  $f(G)$  le nombre de sommets forcés de  $G$ .

Pour les codes non-adaptatifs, cette inégalité n'est plus valide, et il est bien difficile d'estimer ce que nous « coûte » un sommet forcé. Considérons par exemple le cas de la puissance  $k$  d'un chemin  $P_{2k}$  à  $2k$  sommets, avec  $k \geq 2$ . Si l'on numérote  $v_1, \dots, v_k$  ses sommets, de sorte qu'il y a une arête entre  $v_i$  et  $v_j$  si, et seulement si,  $|i - j| \leq k + 1$ , alors il est élémentaire de vérifier que les  $2k - 2$  sommets  $v_2, \dots, v_{2k-1}$  sont forcés. En effet, pour tout  $i \in \{1, \dots, k - 1\}$ , on a  $N[v_{i+1}] = N[v_i] \cup \{v_{i+k}\}$ , et pour tout  $i \in \{k + 1, \dots, 2k\}$ , on a  $N[v_i] = N[v_{i+1}] \cup \{v_{i-k+1}\}$ . Cette observation avait déjà été formulée en Section 2.4.1.2.

Cependant, il existe des codes adaptatifs de  $P_{2k}^k$  nécessitant, dans le pire cas, un nombre logarithmique de questions. En effet, il suffit de commencer par interroger le sommet  $v_k$  (par exemple), puis, selon la réponse obtenue, d'interroger  $v_{k+\lfloor \frac{k}{2} \rfloor}$  (si la réponse était OUI), ou d'interroger  $v_{\lfloor \frac{k}{2} \rfloor}$  (si la réponse était NON). Ainsi de suite, on peut appliquer une stratégie de recherche dichotomique nécessitant, dans le pire cas, un nombre de questions de l'ordre de  $\log_2 k$ , strictement inférieur à  $2k - 2$ . Le nombre de sommets forcés d'un graphe  $G$  n'est donc pas une borne inférieure sur  $\alpha^{\text{ID}}(G)$ .

Comme nous l'avons exposé dans [21] et [146], ces codes correspondent donc à des jeux de Rényi. Combinés à l'identification de sous-ensembles de sommets, ils offrent un sujet d'étude vaste et complexe, et très largement inexploré.

### 3.2.2.2 Produits de graphes

Le produit cartésien de graphes a été étudié par Laihonon et moi-même dans un article de l'*Australasian Journal of Combinatorics* [134]. L'intérêt de cet opérateur, comme discuté en Section 2.4.3.3, réside dans la possibilité de construire des graphes admettant un code  $\ell$ -set-identifiant à partir de graphes n'admettant pas nécessairement de code  $\ell$ -set-identifiant.

Une première série de travaux qu'il me semblerait utile d'entreprendre concernerait l'étude plus approfondie du comportement de l'identification d'ensembles de sommets vis-à-vis du produit cartésien.

**Question 8** (Conditions suffisantes pour l'additivité). *Quelles conditions structurelles sur  $G_1$  et  $G_2$  (admettant, respectivement, un code  $\ell_1$ -identifiant*

et un code  $\ell_2$ -identifiant) assurent que  $G_1 \square G_2$  admet un code  $(\ell_1 + \ell_2)$ -identifiant ?

**Question 9** (Au-delà de l'additivité). Pour tout  $\ell_1, \ell_2$ , soit  $f(\ell_1, \ell_2)$  le plus grand  $\ell$  tel qu'il existe  $G_1$  et  $G_2$  (admettant, respectivement, un code  $\ell_1$ -identifiant et un code  $\ell_2$ -identifiant) qui sont tels que  $G_1 \square G_2$  admet un code  $\ell$ -identifiant. Que peut-on dire sur cette fonction  $f$  ? Quels sont les graphes pertinents pour atteindre cette valeur maximum ?

Par ailleurs, il existe d'autres notions de produits de graphes, telles celles de joint (étudiée en Section 2.4.1.2), de produit tensoriel, lexicographique, etc. (voir par exemple l'ouvrage de Hammack, Imrich et Klavžar [102]). Pour chacun de ces produits, il serait utile de connaître son comportement vis-à-vis de l'identification de sommets, et de dériver des résultats analogues à ceux du Théorème 51.

Enfin, il est à noter que l'équivalent de la conjecture de Vizing n'est pas vraie pour l'identification de sommets (voir la Section 2.2.4). Cependant, on peut se demander si l'analogue du résultat de Clark et Suen [55] est vrai pour les codes identifiants :

**Question 10** (Inégalité à la Clark et Suen pour les codes identifiants). Existe-t-il une constante  $\alpha > 0$  telle que  $\gamma^{\text{ID}}(G_1 \square G_2) \geq \alpha \gamma^{\text{ID}}(G_1) \gamma^{\text{ID}}(G_2)$  pour tous graphes  $G_1, G_2$  admettant un code identifiant ?

### 3.2.2.3 Systèmes de surveillance

Les systèmes de surveillance (*watching systems*) ont été introduits dans la thèse de David Auger [8] et dans un article de Auger, Charon, Hudry et Lobstein à paraître dans *Discrete Applied Mathematics* [12].

**Définition 8** (Système de surveillance). Soit  $G = (V, E)$  un graphe et soit  $r, \ell \geq 1$ . Considérons un ensemble  $W = \{w_1, \dots, w_k\}$ , tel que  $w_i$  est un couple  $(v_i, Z_i)$  pour tout  $i$ , où  $v_i \in V$  et  $Z_i \subseteq B_r(v_i)$ . Pour tout sommet  $v \in V$ , notons  $L_W(v)$  l'ensemble  $\{w_i = (v_i, Z_i) \mid v \in Z_i\}$ . Pour un sous-ensemble de sommets  $A \subseteq V$ , notons  $I_W(A)$  l'ensemble  $\bigcup_{v \in A} L_W(v)$ . L'ensemble  $W$  est un  $(r, \leq \ell)$ -système de surveillance de  $G$  si les  $I_W(A)$  sont tous distincts, pour tous les sous-ensembles  $A$  tels que  $|A| \leq \ell$ .

Noter que si l'on impose que  $Z_i = B_r(w_i)$  pour tout  $i$ , alors on retrouve la définition de code  $(r, \leq \ell)$ -identifiant. Les systèmes de surveillance sont donc une extension des codes identifiants dans le sens où, pour chaque sommet  $v$ , on va choisir un sous-ensemble de sommets de  $B_r(v)$  que  $v$

va être chargé de surveiller. L'un des intérêts est de pouvoir écarter le problème des sommets jumeaux, puisqu'un système de surveillance existe dans tout graphe (considérer  $\{(v, \{v\}) \mid v \in V\}$ ).

À l'instar de la convention de notation utilisée dans ce document, on désignera un  $(1, \leq \ell)$ -système de surveillance par le terme  $\ell$ -set-système de surveillance.

Pour un graphe  $G$ , on note  $w_\ell(G^r)$  la cardinalité minimum d'un  $(r, \leq \ell)$ -système de surveillance de  $G$ . Dans le cas  $r = \ell = 1$ , on notera cette quantité  $w(G)$ .

Une autre différence importante entre les codes identifiants et les systèmes de surveillance est illustrée par les bornes générales suivantes (démontrées dans [12]) :

**Proposition 12.** *Soit  $G$  un graphe à  $n \geq 5$  sommets. Alors on a*

$$\lceil \log_2(n+1) \rceil \leq w(G) \leq \frac{2n}{3}.$$

La borne supérieure provient d'une propriété de monotonie, que vérifient les systèmes de surveillance, et non les codes identifiants.

**Proposition 13** (Systèmes de surveillance dans les sous-graphes). *Soient  $G = (V, E)$  un graphe et  $H = (V, E')$  un sous-graphe de  $G$  sur le même ensemble de sommet. Alors on a*

$$w(G) \leq w(H).$$

En effet, comme on peut choisir un sous-ensemble  $Z_i \subseteq N[v_i]$  pour tout sommet  $v_i$  du système de surveillance, alors un système de surveillance de  $H$  est également un système de surveillance de  $G$ . La borne supérieure  $w(G) \leq \frac{2n}{3}$  est alors obtenue en considérant une forêt couvrante de  $G$  et en raisonnant sur les arbres.

Outre le fait que cette notion me semble intéressante en soi, et qu'elle est susceptible de déboucher sur nombre de résultats intéressants, les systèmes de surveillance permettent d'établir un lien formel avec les codes superposés. En effet, le principal obstacle à la construction de codes identifiants à partir de codes superposés provient de contraintes structurelles assez fortes que l'on exige des codes superposés pour que ceux-ci puissent être vus comme correspondant à une structure de graphe.

Les systèmes de surveillance offrent eux une plus grande souplesse, liée au fait que l'on peut choisir de quels sommets de son voisinage un sommet du code donné va s'occuper. Cette souplesse supplémentaire permet alors d'établir un lien formel entre les deux notions.

**Théorème 60** (Équivalence entre les codes superposés et les systèmes de surveillance). Soit  $G = (V, E)$  un graphe à  $n$  sommets et soit  $W = \{(v_i, Z_i)\}_{i=1, \dots, k}$  un système de surveillance de  $G$  de cardinalité  $k$ . Soit  $\mathcal{F}(W)$  la famille de sous-ensembles de  $\{v_i \mid i = 1, \dots, k\}$  définie comme suit :

$$\mathcal{F}(W) = \{L_W(v) \mid v \in V\}.$$

Alors  $\mathcal{F}(W)$  est un code  $\ell$ -superposé de dimension  $k$ , et de cardinalité  $n$ . Réciproquement, soit  $\mathcal{F}$  un code  $\ell$ -superposé maximal de dimension  $n$ . Alors il est possible de trouver un ensemble de permutations des lignes et des colonnes de la matrice binaire  $M$  associée à  $\mathcal{F}$  de sorte que  $M$  puisse être vue comme la matrice d'incidence entre  $W$  et  $V$  d'un graphe à  $|\mathcal{F}|$  sommets, muni d'un  $\ell$ -set-système de surveillance  $W$  de cardinalité  $|w| = n$ .

La première implication est une simple conséquence de la définition d'un système de surveillance, et était déjà observée par Auger *et al* [12, Section 6].

En ce qui concerne la construction d'un graphe muni d'un  $\ell$ -set-système de surveillance à partir d'un code  $\ell$ -superposé, c'est une conséquence de la Proposition 6. En effet, on sait par cette proposition qu'il est possible de trouver un ensemble de permutations des lignes et des colonnes de la matrice binaire  $M$  associée à  $\mathcal{F}$  de sorte que  $M$  puisse être vue comme la matrice d'incidence entre  $C$  et  $V$  d'un graphe orienté  $\vec{G}$  à  $|\mathcal{F}|$  sommets, muni d'un code  $\ell$ -set-identifiant  $C$  de cardinalité  $|C| = n$ .

Maintenant, construisons un graphe non-orienté  $G$  à partir de  $\vec{G}$  comme suit. Les sommets de  $G$  sont ceux de  $\vec{G}$ , et  $uv$  est une arête de  $G$  si, et seulement si, on a  $(u, v)$  ou  $(v, u)$  qui est une arête de  $\vec{G}$ . Considérons alors le  $\ell$ -système de surveillance suivant :

$$W(C) = \{(v, \Gamma(v)) \mid v \in C\},$$

où  $\Gamma(v)$  désigne l'ensemble  $\{u \mid (v, u) \text{ est un arc de } \vec{G}\}$ . Il est alors immédiat de vérifier que  $W$  est un  $\ell$ -set-système de surveillance de  $G$ .

En un sens, ce résultat peut être vu comme une conséquence de la propriété de monotonie, qui dit essentiellement qu'ajouter des arêtes à un graphe ne peut qu'améliorer les choses du point de vue de la surveillance de ses sommets.

La notion de système de surveillance unifie donc dans un même cadre les graphes orientés et les graphes non-orientés.

Ainsi, les raisons exposées ci-dessus me semblent constituer un argument indiquant que la notion de système de surveillance offre un potentiel intéressant pour des recherches futures sur les codes identifiants et ses généralisations.

### 3.2.2.4 Colorations identifiantes

Les colorations identifiantes offrent un cadre d'étude très général pour l'identification de sommets dans les graphes.

**Définition 9** (Coloration identifiante). *Une  $k$ -coloration identifiante d'un graphe  $G = (V, E)$  est une fonction  $c : V \rightarrow \{1, \dots, k\}$  qui assigne une couleur à chaque sommet, de sorte que les ensembles  $I_c(v) = \{c(u) \mid u \in N[v]\}$  sont tous deux à deux distincts.*

Cette notion a été suggérée par Éric Duchêne à l'occasion des Journées Maths à Modeler (octobre 2009), et étudiée lors des Journées Discrètes à Grenoble en octobre 2009. Les premiers résultats sur cette notion ont été par la suite réunis dans un manuscrit non publié.

Esperet *et al* l'introduisent dans un manuscrit soumis [71], qui traite plus spécifiquement de colorations *localement identifiantes*, c'est-à-dire telles que  $uv \in E \Rightarrow I_c(u) \neq I_c(v)$  (mais on s'autorise éventuellement à avoir  $I_c(u) = I_c(v)$  pour deux sommets  $u, v$  non voisins). Les colorations localement identifiantes ont de plus été étudiées par Foucaud *et al* [76].

Cette notion ressemble à la coloration d'arêtes sommet-distinguantes, introduite par Horňák et Soták [112], qui consiste à colorier les arêtes d'un graphe de sorte à ce que deux arêtes incidentes aient des couleurs différentes, et telles que l'ensemble des couleurs des arêtes incidentes à un sommet identifie celui-ci de façon unique.

Si l'on note  $\chi^{\text{ID}}(G)$  le  $k$  minimum tel qu'il existe une  $k$ -coloration identifiante de  $G$ , alors on a toujours  $\chi^{\text{ID}} \leq \gamma^{\text{ID}}(G) + 1$ . En effet, soit  $C$  un code identifiant de  $G$  et soit  $\phi$  une bijection entre  $C$  et  $\{1, \dots, |C|\}$ . Si l'on colorie les sommets de  $G$  comme suit :

$$c(v) = \begin{cases} \phi(v) & \text{si } v \in C, \\ |C| + 1 & \text{si } v \notin C \end{cases}$$

alors on obtient une  $(|C| + 1)$ -coloration identifiante de  $G$ .

Dans sa version de coloration par liste, cette notion permet d'unifier les colorations identifiantes et les codes identifiants.

**Définition 10** (Coloration par liste identifiante). *Soit  $G = (V, E)$  un graphe. Pour tout sommet  $v \in V$ , considérons  $L(v)$  un ensemble (la liste des couleurs légales pour  $v$ ). Une  $k$ -coloration par liste identifiante est un étiquetage  $c$ , qui assigne une couleur  $c(v) \in L(v)$  à chaque sommet  $v$ , de sorte que les ensembles  $I_c(v) = \{c(u) \mid u \in N[v]\}$  sont tous deux à deux distincts, et tel que  $|\{c(v) \mid v \in V\}| = k$ .*

Clairement, une  $k$ -coloration identifiante de  $G$  est une  $k$ -coloration par liste identifiante de  $G$ , où la liste  $L(v)$  de couleurs légales pour chaque sommet  $v$  est simplement égale à  $\{1, \dots, k\}$ .

Par ailleurs, pour chaque sommet  $v \in V$  d'un graphe  $G = (V, E)$ , considérons la liste  $L_i$  constituée de  $v$  lui-même et d'un élément abstrait  $\omega \notin V$ . En ce cas, chaque sommet  $v$  est colorié soit avec  $v$ , soit avec ce symbole  $\omega$ . Considérons alors une  $k$ -coloration par liste identifiante  $c$  de  $G$ . Si l'on considère l'ensemble  $C$  des sommets coloriés par eux-mêmes,  $C = \{v \mid c(v) = v\}$ , alors il est clair que  $C$  est un code identifiant de  $G$ , de cardinalité  $k - 1$ .

Inversement, si  $C$  est un code identifiant de  $G$ , alors la coloration  $c$  définie par

$$c(v) = \begin{cases} v & \text{si } v \in C \\ \omega & \text{sinon} \end{cases}$$

est une  $(|C| + 1)$ -coloration par liste identifiante de  $G$ .

Ainsi, la notion de code identifiant est un cas particulier de coloration par liste identifiante.

Si l'on généralise maintenant au cas de la multicoloration par liste, où chaque sommet  $v$  reçoit un *sous-ensemble* de couleurs de sa liste  $L(v)$  de couleurs légales, alors la notion de multicoloration par liste identifiante généralise naturellement les notions précédentes, ainsi que la notion de système de surveillance.

**Définition 11** (Multicoloration par liste identifiante). *Soit  $G = (V, E)$  un graphe. Pour tout sommet  $v \in V$ , considérons  $L(v)$  un ensemble (la liste des couleurs légales pour  $v$ ). Une  $k$ -multicoloration par liste identifiante est un étiquetage  $c$ , qui assigne un sous-ensemble non vide de couleurs  $c(v) \subseteq L(v)$  à chaque sommet  $v$ , de sorte que les ensembles  $I_c(v) = \bigcup_{u \in N[v]} c(u)$  sont tous deux à deux distincts, et tel que  $|\bigcup_{v \in V} c(v)| = k$ .*

En effet, considérons, pour chaque sommet  $v \in V$  d'un graphe  $G = (V, E)$ , la liste  $L_w$  constituée de tous les éléments de  $N[v]$ . En ce cas, chaque sommet  $v$  est colorié avec un sous-ensemble de  $N[v]$ . Considérons alors une  $k$ -multicoloration par liste identifiante  $c$  de  $G$ . Pour tout sommet  $u$  de  $G$  tel qu'il existe  $v \in V$  avec  $u \in c(v)$ , posons alors  $Z_u = \{v \mid u \in c(v)\}$ . Clairement, l'ensemble des paires  $(u, Z_u)$  ainsi obtenues forme un système de surveillance de  $G$ , de cardinalité  $k$ .

Inversement, si  $W = \{(v_1, Z_1), \dots, (v_k, Z_k)\}$  est un système de surveillance de  $G$ , alors la multicoloration  $c$  définie par

$$c(v) = \{v_i \mid v \in Z_i\}$$

est une  $k$ -multicoloration par liste identifiante de  $G$ .

Ainsi, la notion de système de surveillance est un cas particulier de multicoloration par liste identifiante.

### 3.3 Projet personnel

Mes projets personnels de recherche concernent, d'une part, certaines des perspectives de recherche énoncées ci-dessus, et, d'autre part, de nouvelles thématiques que je souhaite développer sur des problèmes d'ordonnement. Les perspectives de recherche sur les codes identifiants étant détaillées dans la section précédente, la partie que je développe le plus ci-après concerne les nouvelles thématiques en ordonnancement.

#### 3.3.1 Codes identifiants

Parmi les questions ouvertes de la littérature, deux questions me tiennent particulièrement à cœur.

La première concerne la monotonie dans les hypercubes (voir la Conjecture 1 de Blass, Honkala, et Litsyn – voir aussi la Section 3.2.1.1). Une voie possible pour contourner le fait que chaque sommet du code projeté peut correspondre à plusieurs paires de sommets non séparés, serait de projeter simultanément plusieurs sommets, puis d'altérer localement le code projeté pour régler le cas de ces sommets non séparés.

L'intérêt de considérer la projection de plusieurs sommets du code en simultané serait de pouvoir disposer de *plusieurs sommets pour régler plusieurs problèmes*. On pourrait espérer trouver une façon astucieuse de projeter ces sommets, de sorte à ce que les voisinages de ceux-ci se « combinent » harmonieusement et nous permettent de régler tous les problèmes. En effet, si un seul sommet peut régler potentiellement un seul problème,  $k$  sommets peuvent, eux, potentiellement régler  $2^k - 1$  problèmes...

L'autre question qui me tient particulièrement à cœur concerne l'unification des différentes notions de codes (codes identifiants, codes superposés, familles *cover-free*, etc. – voir la Section 3.2.1.7). En effet, il me semble que, d'un certain point de vue, de nombreux résultats sur ces différents types de codes utilisent « moralement » les mêmes techniques et les mêmes méthodes. Le cas le plus flagrant est sans doute celui du parallèle qui existe entre l'algorithme de Hwang et Sós pour le cas des familles *part-intersecting* [113], et mon propre algorithme pour le cas des codes  $\ell$ -set-identifiants [145]. Un premier objectif pourrait être de dégager des résultats

généraux, unifiant au sein d'un même cadre théorique une série de résultats sur tous les problèmes d'identification.

Pour ces problèmes, il semble de plus toujours y avoir un lien fort entre ce que l'on peut obtenir à l'aide de méthodes probabilistes, et ce que l'on peut obtenir à l'aide de constructions explicites. La plupart du temps, les constructions explicites peuvent en effet être vues comme des « dérandomisations » de constructions probabilistes. Si le concept de « dérandomisation » est classique dans la littérature, il pourrait être utile de dégager des résultats théoriques génériques sur la dérivation d'algorithmes explicites à partir de constructions probabilistes, pour le cas des problèmes d'identification.

Enfin, la thématique de recherche émergente à laquelle je souhaiterais apporter des contributions en priorité est celle des codes identifiants adaptatifs. Dans un travail en cours [67], nous introduisons avec Paul Dorbec, Aline Parreau et Gabriel Renault la notion de *noyau* d'un code adaptatif, qui est défini comme la branche « NON » de l'arbre de décision. Dans le cas des arbres, il s'avère que le noyau d'un code identifiant adaptatif permet de décrire de façon compacte le code (aux symétries près). Une question sous-jacente serait alors de savoir si cette notion de noyau nous permet de calculer en temps polynomial un code identifiant adaptatif optimum dans un arbre. On peut également se demander ce que l'on peut dire du noyau d'un code adaptatif pour un graphe quelconque. Il serait en particulier intéressant de pouvoir dégager des bornes générales sur la cardinalité d'un code adaptatif, en fonction de paramètres liés au noyau (par exemple, sa cardinalité).

### 3.3.2 Ordonnancement

Le deuxième volet de mon projet personnel de recherche concerne des thématiques relativement orthogonales à celle des codes identifiants, qui ont été peu évoquées dans ce document. J'ai en effet initié depuis quelques années des travaux sur des problèmes d'ordonnancement de la production [147, 148, 149, 150, 151, 170], et je souhaite continuer à travailler dans cette direction.

Je présente ci-après deux familles de problèmes sur lesquels je souhaiterais travailler, en collaboration avec mes collègues de l'équipe MOGISA du LAAS.

### 3.3.2.1 Décomposition de graphes et ordonnancement

De nombreux modèles d'ordonnancement invoquent des graphes. Le modèle le plus classique est celui où les tâches sont soumises à des *contraintes de précédence*, qui contraignent typiquement une tâche  $j$  à ne pouvoir démarrer que lorsque la tâche  $i$  est terminée. Le graphe de précédence — qui est un graphe auxiliaire associé au problème d'ordonnancement — possède dans ce cas un arc depuis  $i$  vers  $j$ .

Lorsque l'on étudie la littérature, les résultats d'ordonnancement pour des problèmes avec des contraintes de précédence invoquent naturellement des raisonnements et des résultats de nature résolument « graphes ». Dans ce cas, les problèmes considérés concernent en général des structures de chemins dans le graphe de précédence. Typiquement, dans le problème de gestion de projet — la « méthode PERT » — la date de fin au plus tôt du projet correspond à la longueur d'un plus long chemin dans le graphe de précédence.

Or, nombre de problèmes d'ordonnancement qui sont NP-complets sous contrainte de précédences quelconques, deviennent polynomiaux lorsque le graphe de précédence a une structure particulière — typiquement, lorsque celui-ci a une structure arborescente, ou est un graphe série-parallèle. C'est par exemple le cas du problème  $1 \mid \text{prec} \mid \sum w_j C_j$ , qui est fortement NP-complet, alors que les problèmes  $1 \mid \text{chains} \mid \sum w_j C_j$ ,  $1 \mid \text{intree} \mid \sum w_j C_j$ ,  $1 \mid \text{outtree} \mid \sum w_j C_j$  et  $1 \mid \text{series} - \text{parallel} \mid \sum w_j C_j$  sont tous polynomiaux (voir par exemple le *survey* de Chen *et al* [54]).

Les résultats de polynomialité ci-dessus ont été démontrés de façon indépendante, et l'on est en droit de se demander s'il est possible de donner un théorème général de polynomialité. La thématique que je souhaiterais étudier concerne le rapprochement de résultats de théorie des graphes sur la décomposition de graphes avec les modèles d'ordonnancement sous contraintes de précédence.

Par exemple, la théorie de la largeur d'arborescence (ou *tree-width*) fournit des résultats du type « si  $\mathcal{P}$  est un problème exprimable dans un certain langage  $\mathcal{L}$ , alors  $\mathcal{P}$  est polynomial lorsque restreint à une classe de graphes de largeur d'arborescence bornée » (voir par exemple le livre de Courcelle et Engelfriet [60]). Sans rentrer dans les détails, les arbres et les graphes série-parallèle sont des graphes de largeur d'arborescence bornée.

La question générale que je souhaiterais explorer est alors la suivante :

**Question 11.** *Peut-on, à l'aide de résultats généraux sur la décomposition de graphes (par exemple la tree-width, la décomposition modulaire, etc.), obtenir des résultats généraux sur des problèmes d'ordonnancement sous contraintes*

*de précédences lorsque le graphe de précedence possède de bonnes propriétés de décomposabilité (par exemple lorsque celui-ci est une union disjointe de chaînes, d'arbres, etc.) ?*

Il est à noter qu'au-delà des problèmes d'ordonnancement où la structure de graphe est explicite, par exemple sous la forme d'un graphe de précedence, il existe également des problèmes où la résolution du problème peut passer par l'étude de la structure d'un graphe auxiliaire implicite. Par exemple, dans les problèmes de type RCPSP (*resource-constrained project scheduling problem*, ou problème d'ordonnancement de projet sous contraintes de ressources), lorsque deux tâches  $i$  et  $j$  nécessitent toutes deux la même ressource pour être exécutée, alors ces deux tâches ne peuvent clairement pas être exécutées simultanément. En ce cas, on peut construire un graphe auxiliaire tel qu'il existe une arête entre  $i$  et  $j$  dès lors que  $i$  et  $j$  nécessitent toutes deux la même ressource pour être exécutée. Le problème de graphes auquel nous sommes confrontés s'apparente alors en ce cas à un problème de coloration.

De même, lorsque l'on travaille avec des fenêtres de temps, c'est-à-dire lorsqu'une tâche  $i$  de durée  $p_i$  doit être exécutée entre les temps  $r_i$  et  $d_i$  (avec  $r_i + p_i \leq d_i$ ), on peut définir un graphe auxiliaire comme suit. Si l'on a simultanément  $r_i \leq r_j$  et  $\max(d_i, d_j) - r_i < p_i + p_j$ , alors les tâches  $i$  et  $j$  ne pourront clairement pas être exécutées sur la même machine. En ce cas, on met une arête entre  $i$  et  $j$ , et le problème de graphe sous-jacent correspond de nouveau à un problème de coloration.

Le nombre de problèmes d'ordonnancement invoquant (implicitement ou explicitement) des problèmes de graphes « classiques » (cheminements, coloration, etc.) est donc très important. Dans ce contexte, il me semble que les résultats généraux sur la décomposition de graphes devraient pouvoir trouver des applications fructueuses dans le champ de l'ordonnancement de la production.

### 3.3.2.2 Ordonnancement de tâches dont le temps opératoire dépend du temps et/ou du rang

Le deuxième type de problèmes sur lequel je souhaite travailler concerne les problèmes d'ordonnancement de tâches dont le temps opératoire dépend du temps et/ou du rang.

Dans la théorie « classique » de l'ordonnancement, une tâche  $i$  a une durée exprimée sous la forme d'un paramètre  $p_i$ . Si l'on s'intéresse à la modélisation de problèmes concrets, il existe de nombreuses situations où cette formulation est une hypothèse simplificatrice, et où la durée d'une

tâche va essentiellement dépendre du *moment* où celle-ci sera réalisée. Par exemple, si la tâche est un feu de forêt à éteindre, alors le temps nécessaire pour éteindre l'incendie va clairement dépendre du moment où les sapeurs-pompiers vont commencer à éteindre celui-ci. En effet, plus le temps passe, plus le feu se propage, et plus il va falloir de temps (et de ressources) pour éteindre celui-ci. Inversement, d'autres exemples concrets incluent, par exemple, des situations où les opérateurs sont en situation d'apprentissage, et où ceux-ci de plus en plus efficaces au fur et à mesure que le temps passe.

Le cas des problèmes d'ordonnement de tâches dont le temps opératoire dépend du temps et/ou du rang modélise une large classe de phénomènes concrets. Ces modèles sont apparus dans la littérature depuis la fin des années 1990, et font depuis l'objet de beaucoup d'attention de la part de la communauté (voir par exemple le *survey* de Biskup [27] ou l'ouvrage de Gawiejnowicz [89]). Dans un tel contexte, la durée d'une tâche  $i$  s'exprime alors comme une fonction  $p_i(r, t)$ , avec  $r$  le rang de la tâche (c'est-à-dire sa position dans l'ordonnement), et  $t$  le temps auquel cette tâche va démarrer.

Sous ces hypothèses, il semble que l'ensemble des résultats connus en ordonnancement doivent être revus. Il existe en particulier des problèmes classiques qui sont polynomiaux, mais deviennent NP-difficiles lorsque plongés dans un contexte où le temps opératoire dépend du temps et/ou du rang. Par exemple, le problème  $1 \mid r_j \mid C_{\max}$  est polynomial, alors que le problème  $1 \mid r_j, p_j(r, t) = r^a p_j \mid C_{\max}$  est NP-difficile (avec  $a$  constante).

De façon surprenante, il existe d'ailleurs des problèmes simples dont la complexité n'est toujours pas connue. C'est notamment le cas du problème  $1 \mid p_j(r, t) = a_j t + b_j \mid \sum C_j$ , qui est toujours ouvert à ce jour (et ce même si l'on a  $a_j = 1$  pour tout  $j$ ).

Nous avons récemment obtenu des résultats en collaboration avec Vincent Jost et Gerd Finke sur le sujet, dans le cas particulier de problèmes sur une machine, où la durée des tâches dépend uniquement du rang (et non du temps) [147, 148, 149]. Notre approche utilise un point de vue « graphes », et nous montrons qu'une large classe de problèmes se réduit à la recherche d'un couplage maximum de poids minimum dans un graphe biparti particulier. Nous unifions ainsi de nombreux résultats de la littérature, et répondons à certaines questions ouvertes.

Les perspectives de ces travaux concernent l'adaptation de ces résultats au cas où la durée d'une tâche dépend du temps, et constituent la prochaine question que je souhaite étudier sur le sujet.

# Index

- algorithme d'approximation, 63
- algorithme d'approximation à facteur constant, 64
- APX-complet, 64
- APX-difficile, 64
- arête, 19
- arbre, 53
- arbre orienté, 53
- arc, 20
  
- $B_r(v)$ , 29
- bande, 29
- boule, 29
  
- $c(G)$ , 99
- $C_n$ , 37
- cage, 91
- charge, 31
- chemin, 21
- clause, 50
- clique, 45
- code couvrant, 20
- code identifiant, 20
- code  $l$ -set-identifiant, 22
- code parfait, 100
- code  $r$ -couvrant, 21
- code  $r$ -couvrant  $t$ -multiple, 42
- code  $r$ -identifiant, 21, 22
- code  $(r, \leq \ell)$ -identifiant, 21
- code  $r$ -séparateur, 21
- code séparateur, 20
- code superposé, 80
- coloration identifiante, 117
- coloration par liste identifiante, 118
- couplage, 47, 70
- couplage maximum, 73
  
- couplage parfait, 70
- cover-free*, 109
- cycle, 37
  
- $d_r(G)$ , 99
- $d$ -régulier, 20
- déchargement, 31
- degré, 20
- densité, 30
- distance-régulier, 91
- dominant, 20
  
- empilement, 99
- étiquetage, 60
- étiquetage partiel, 60
- étoile, 23, 69
  
- $\mathcal{F}_q(G)$ , 60
- famille  $l$ -cover-free, 109
- famille  $l$ -part-intersecting, 109
- fasciagraphe, 58
- fermeture transitive, 21
- feuille, 53
- fibre, 58
- flot, 9
  
- $\gamma^{\text{ID}}(G^r)$ , 22
- $\Gamma(G, M)$ , 70
- $\Gamma$ -graphe, 70
- $\gamma_\ell^{\text{ID}}(G)$ , 22
- garantie, 63
- graphe, 19
- graphe aléatoire, 77
- graphe biparti, 51
- graphe complet, 45
- graphe connexe, 87

- graphe d'intersection de disques, 51  
 graphe de Hamming, 45  
 graphe de Sierpiński, 47  
 graphe mixte, 58  
 graphe optimal, 72  
 graphe orienté, 20  
 graphe parfait, 51  
 graphe planaire, 51  
 grille, 34  
  
 $H(n, d)$ , 45  
 hyperarête, 20  
 hypercube, 40  
 hypergraphe, 20  
  
 joint, 71  
 jumeau, 21  
  
 k-coloration identifiante, 117  
 k-coloration par liste identifiante, 118  
 $K_{m,n}$ , 23  
 k-multicoloration identifiante, 119  
 $K_n$ , 45  
  
 *$\ell$ -parity-check*, 110  
*line-graph*, 51  
 littéral, 50  
  
 maille, 51  
 matrice S-contrainte, 111  
 multicoloration par liste identifiante, 119  
  
 noyau, 120  
  
 $P(\mathcal{F}_q(G))$ , 60  
*parity-check*, 110  
*part-intersecting*, 109  
 père, 53  
 plan projectif, 83  
 précode, 34  
 produit cartésien, 42  
 pseudo-d-local, 61, 62  
  
 q-étiquetage, 60  
 $Q_n$ , 40  
 q-propriété, 60  
  
 *$\ell$ -système de surveillance*, 115  
 régulier, 20  
 racine, 74  
 rotagraphe, 58  
  
 S-contrainte, 111  
 $S(n, k)$ , 47  
 sommet, 19  
 sommet extrême, 47  
 sommet forcé, 113  
 sommet interne, 47  
 sous-graphe, 19  
 sous-graphe induit, 19  
 système de surveillance, 115  
  
 transversal, 38  
 triangle, 41  
  
*unit disk graph*, 51  
  
 $v(G)$ , 99  
 voisin, 19  
 voisinage, 20  
 voisinage fermé, 20  
  
*watching system*, 115

# Bibliographie

- [1] I. Adiri, J. Bruno, E. Frostig, A. H. G. Rinnooy Kan, *Single machine flow-shop scheduling with a single breakdown*, Acta Informatica **26** (1989), 679–696.
- [2] S. Alanko, S. Crevals, A. Isopoussu, P. Östergård, V. Pettersson, *Computing the Domination Number of Grid Graphs*, Electronic Journal of Combinatorics **18(1)** (2011), P141.
- [3] B. Alidaee, N. Womer, *Scheduling with time dependent processing times : Review and extensions*, Journal of the Operational Research Society **50** (1999), 711–720.
- [4] P. Alimonti, V. Kann, *Some APX-completeness results for cubic graphs*, Theoretical Computer Science **237(1-2)** (2000), 123–134.
- [5] S. Arnborg, J. Lagergren, D. Seese, *Easy problems for tree-decomposable graphs*, Journal of Algorithms **12(2)** (1991), 308–340.
- [6] D. Auger, *Induced paths in twin-free graphs*, Electronic Journal of Combinatorics **15(1)** (2008), R17.
- [7] D. Auger, *Minimal identifying codes in trees and planar graphs with large girth*, European Journal of Combinatorics **31(5)** (2010), 1372–1384.
- [8] D. Auger, *Problèmes d’identification combinatoire et puissances de graphes*, thèse de doctorat Télécom ParisTech (2010). Disponible en ligne à l’url <http://www.lri.fr/~dauger/thesisDA.pdf>.
- [9] D. Auger, I. Charon, I. Honkala, O. Hudry, A. Lobstein, *Edge number, minimum degree, maximum independent set, radius and diameter in twin-free graphs*, Advances in Mathematics of Communications **3(1)** (2009), 97–114.
- [10] D. Auger, I. Charon, O. Hudry, A. Lobstein, *On the existence of a cycle of length at least 7 in a  $(1, \leq 2)$ -twin-free graph*, Discussiones Mathematicae-Graph Theory **30(4)** (2010), 591–609.
- [11] D. Auger, I. Charon, O. Hudry, A. Lobstein, *Complexity results for identifying codes in planar graphs*, International Transactions in Operational Research **17(6)** (2010), 691–710.
- [12] D. Auger, I. Charon, O. Hudry, A. Lobstein, *Watching systems in graphs : an extension of identifying codes*, à paraître dans Discrete Applied Mathematics.

- [13] D. Auger, I. Charon, O. Hudry, A. Lobstein, *On the sizes of the graphs  $G$ ,  $G^r$ ,  $G^r \setminus G$  : the directed case*, Australasian Journal of Combinatorics, **48** (2010), 87–109.
- [14] D. Auger, I. Charon, O. Hudry, A. Lobstein, *On the sizes of the graphs  $G$ ,  $G^r$ ,  $G^r \setminus G$  : the undirected case*, Discrete Applied Mathematics **159**(16) (2011), 1666–1675.
- [15] D. Auger, I. Charon, O. Hudry, A. Lobstein, *On the square roots of graphs*, soumis.
- [16] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, M. Protasi, *Complexity and Approximation : Combinatorial optimization problems and their approximability properties*, Springer-Verlag, 1999.
- [17] C. Barnhart, C. Kniker, M. Lohatepanout, *Itinerary-based airline fleet assignment*, Transportation Science **36**(2) (2002), 199–217.
- [18] B. Baumann, *Modélisation d'un problème de surveillance par un jeu de tirs aux buts – calculs de stratégies optimales*, mémoire de travail d'étude et de recherche ENSIMAG (mai 2008).
- [19] L. Beaudou, S. Gravier, S. Klavžar, M. Kovše, M. Mollard, *Covering codes in Sierpiński graphs*, Discrete Mathematics and Theoretical Computer Science **12**(3) (2010), 63–74.
- [20] Y. Ben-Haim, S. Gravier, A. Lobstein, J. Moncel, *Adaptive identification in graphs*, Journal of Combinatorial Theory Series A **115**(7) (2008), 1114–1126.
- [21] Y. Ben-Haim, S. Gravier, A. Lobstein, J. Moncel, *Adaptive identification in torii in the king lattice*, Electronic Journal of Combinatorics **18**(1) (2011), P116.
- [22] Y. Ben-Haim, S. Litsyn, *Exact minimum density of codes identifying vertices in the square grid*, SIAM Journal on Discrete Mathematics **19** (2005), 69–82.
- [23] A. A. Bertossi, *Dominating sets for split and bipartite graphs*, Information Processing Letters **19** (1984), 37–40.
- [24] N. Bertrand, *Codes identifiants et codes localisateurs-dominateurs sur certains graphes*, mémoire de master 2, Télécom ParisTech (2001).
- [25] N. Bertrand, I. Charon, O. Hudry, A. Lobstein, *Identifying and locating-dominating codes on chains and cycles*, European Journal of Combinatorics **25** (2004), 969–987.
- [26] N. Bertrand, I. Charon, O. Hudry, A. Lobstein, *1-identifying codes on trees*, Australasian Journal of Combinatorics **31** (2005), 21–35.

- [27] D. Biskup, *A state-of-the-art review on scheduling with learning effects*, European Journal of Operational Research **188** (2008), 315–329.
- [28] U. Blass, I. Honkala, S. Litsyn, *On binary codes for identification*, Journal of Combinatorial Designs **8(2)** (2000), 151–156.
- [29] U. Blass, I. Honkala, S. Litsyn, *Bounds on identifying codes*, Discrete Mathematics **241** (2001), 119–128.
- [30] M. Blidia, M. Chellali, F. Maffray, J. Moncel, A. Semri, *Locating-dominance and identifying codes in trees*, Australasian Journal of Combinatorics **39** (2007), 219–232.
- [31] B. Bollobás, *Random Graphs*, Academic Press (1985).
- [32] T. Bonates, P. Hammer, *Logical Analysis of Data : From Combinatorial Optimization to Medical Applications*, Annals of Operations Research **148** (2006), 203–225.
- [33] M. Bouznif, J. Darlay, J. Moncel, A. Parreau, *Identifying codes in generalized Sierpiński graphs*, en préparation.
- [34] M. Bouznif, J. Darlay, J. Moncel, M. Preissmann, *New results on identifying codes in square grids of small height*, en préparation.
- [35] M. Bouznif, J. Moncel, M. Preissmann, *Generic algorithms for some decision problems on fasciagraphs and rotagraphs*, Discrete Mathematics **312(17)** (2012), 2707–2719.
- [36] M. Bouznif, J. Moncel, M. Preissmann, *Optimization algorithms for rotagraphs and fasciagraphs*, soumis.
- [37] M. Bouznif, J. Moncel, M. Preissmann, *Étude de problèmes combinatoires sur les fasciagraphes et les rotagraphes*, congrès ROADEF 2010, 24–26 février 2010, Toulouse.
- [38] N. Brauner, J. Darlay, J. Moncel, *Dense & sparse graph partition*, accepté à Discrete Applied Mathematics. Disponible en ligne à l’url <http://hal.archives-ouvertes.fr/hal-00454999/fr/>.
- [39] R. E. Burkard, B. Klinz, R. Rudolf, *Perspectives of Monge properties in optimization*, Discrete Applied Mathematics **70** (1996), 95–161.
- [40] L. Cartier, P. Dorbec, É. Duchêne, N. Giroud, J. Moncel, C. Ouvrier-Buffet, *Towards a characterization of research situations for the classroom - Analysis of a case study*, soumis.
- [41] L. Cartier, J. Moncel, *Learners’ conceptions in different class situations around Königsberg’s bridges problem*, International Congress on Mathematical Education, 6–13 juillet 2008, Monterrey (Mexique). Disponible en ligne à l’url <http://tsg.icme11.org/tsg/show/16>.

- [42] O. Çelebi, *Problème de localisation et de dimensionnement de site de concentration en contexte incertain*, mémoire de M2R recherche opérationnelle, combinatoire et optimisation, Grenoble (2007).
- [43] I. Charon, S. Gravier, O. Hudry, A. Lobstein, M. Mollard, J. Moncel, *A linear algorithm for minimum 1-identifying codes in oriented trees*, *Discrete Applied Mathematics*, **154(8)** (2006), 1246–1253.
- [44] I. Charon, I. Honkala, O. Hudry, A. Lobstein, *General bounds for identifying codes in some infinite regular graphs*, *Electronic Journal of Combinatorics* **8** (2001), R39.
- [45] I. Charon, I. Honkala, O. Hudry, A. Lobstein, *The minimum density of an identifying code in the king lattice*, *Discrete Mathematics* **276** (2004), 95–109.
- [46] I. Charon, I. Honkala, O. Hudry, A. Lobstein, *Structural properties of twin-free graphs*, *Electronic Journal of Combinatorics* **14** (2007), R16.
- [47] G. Cohen, I. Honkala, A. Lobstein, G. Zémor, *On identifying codes*, *Proceedings of DIMACS Workshop on Codes and Association Schemes 1999, Piscataway (USA)* **56** (2001), 97–109.
- [48] I. Charon, O. Hudry, A. Lobstein, *Identifying codes with small radius in some infinite regular graphs*, *Electronic Journal of Combinatorics* **9** (2002), R11.
- [49] I. Charon, O. Hudry, A. Lobstein, *Minimizing the Size of an Identifying or Locating-Dominating Code in a Graph is NP-hard*, *Theoretical Computer Science* **290(3)** (2003), 2109–2120.
- [50] I. Charon, O. Hudry, A. Lobstein, *Identifying and locating-dominating codes : NP-completeness results for directed graphs*, *IEEE Transactions on Information Theory* **48(8)**, 2192–2200.
- [51] I. Charon, O. Hudry, A. Lobstein, *Possible cardinalities for identifying codes in graphs*, *Australasian Journal of Combinatorics* **32** (2005), 177–195.
- [52] I. Charon, O. Hudry, A. Lobstein, *Extremal cardinalities for identifying and locating-dominating codes in graphs*, *Discrete Mathematics* **307** (2007), 356–366.
- [53] C. Chen, C. Lua, Z. Miao, *Identifying codes and locating-dominating sets on paths and cycles*, *Discrete Applied Mathematics* **159** (2011), 1540–1547.
- [54] B. Chen, C. Potts, G. Woeginger, *A Review of Machine Scheduling : Complexity, Algorithms and Approximability*, 21–169 dans *Handbook of Combinatorial Optimization*, édité par D.-Z. Du et P.M. Pardalos, Kluwer (1998).

- [55] W. E. Clark, S. Suen, *An inequality related to Vizing's conjecture*, Electronic Journal of Combinatorics **7(1)** (2000), N4.
- [56] E. J. Cockayne, S. E. Goodman, S. T. Hedetniemi, *A linear algorithm for the domination number of a tree*, Information Processing Letters **4(2)** (1975), 41–44.
- [57] G. Cohen, S. Gravier, I. Honkala, A. Lobstein, M. Mollard, C. Payan, G. Zémor, *Improved Identifying Codes for the Grid*, Electronic Journal of Combinatorics, comments to **6(1)** R19 (1999).
- [58] G. Cohen, I. Honkala, S. Litsyn, A. Lobstein, *Covering Codes*, North-Holland Mathematical Library Series **54**, Elsevier (1997).
- [59] G. Cohen, I. Honkala, A. Lobstein, G. Zémor, *New bounds for identifying vertices in graphs*, Electronic Journal of Combinatorics **6** (1999), R19.
- [60] B. Courcelle, J. Engelfriet, *Graph structure and monadic second-order logic, a language theoretic approach*, à paraître chez Cambridge University Press. Disponible en ligne à l'url <http://hal.archives-ouvertes.fr/hal-00646514>.
- [61] D. Cranston, G. Yu, *A new lower bound on the density of vertex identifying codes for the infinite hexagonal grid*, Electronic Journal of Combinatorics **16** (2009), R113.
- [62] M. Daniel, S. Gravier, J. Moncel, *Identifying codes in some subgraphs of the square lattice*, Theoretical Computer Science **319** (2004), 411–421.
- [63] J. Darlay, *Analyse combinatoire de données – structures et optimisation*, thèse de doctorat de l'université de Grenoble (2011).
- [64] J. Darlay, N. Brauner, J. Moncel, *Partition en sous-graphes denses pour la détection de communautés*, congrès ROADEF 2010, 24–26 février 2010, Toulouse.
- [65] Ch. Deppe, *Strategies for the Renyi-Ulam Game*, Theoretical Computer Science **314** (2004), 45–55.
- [66] R. Diestel, *Graph theory*, Springer-Verlag, 4ème édition (2010). Disponible en ligne à l'url <http://diestel-graph-theory.com/>.
- [67] P. Dorbec, J. Moncel, A. Parreau, G. Renault, *Adaptive identifying codes in trees*, en préparation.
- [68] E. Duchêne, S. Gravier, J. Moncel, *New results about Solitaire Clobber*, RAIRO–Operations Research, **43** (2009), 463–482.
- [69] Ch.-M. Duquesne, *The fleet assignment problem*, mémoire de M2R recherche opérationnelle, combinatoire et optimisation, Grenoble (2009).

- [70] A. D'yachkov, V. Rykov, *Bounds on the length of disjunctive codes*, Problems of Information Transmission **18** (1983), 166–171.
- [71] L. Esperet, S. Gravier, M. Montassier, P. Ochem, A. Parreau, *Locally identifying coloring of graphs*, soumis. Disponible en ligne à l'url <http://arxiv.org/abs/1010.5624>.
- [72] G. Exoo, V. Junnila, T. Laihonen, S. Ranto, *Upper bounds for binary identifying codes*, Advances in Applied Mathematics **42** (2009), 277–289.
- [73] N. Fazlollahi, D. Starobinski, A. Trachtenberg, *Connected identifying codes for sensor network monitoring*, WCNC 2011, 2011 IEEE Wireless Communications and Networking Conference, 28–31 Mars 2011, Cancún (Mexique), publié dans les proceedings pp. 1026–1031.
- [74] U. Feige, *A threshold of  $\ln n$  for approximating set cover*, Journal of the ACM **45**(4) (1998), 634–652.
- [75] D. Fera, N. Kim, N. Shiffeldrim, J. Zorn, U. Laserson, H. Gan, T. Schlick, *RAG : RNA-As-Graphs web resource*, BMC Bioinformatics **5** 2004, 88.
- [76] F. Foucaud, I. Honkala, T. Laihonen, A. Parreau, G. Perarnau, *Locally identifying colouring of graphs with given maximum degree*, soumis. Disponible en ligne à l'url <http://arxiv.org/abs/1108.2769>.
- [77] F. Foucaud, E. Guerrini, M. Kovše, R. Naserasr, A. Parreau, P. Valicov, *Extremal graphs for the identifying code problem*, European Journal of Combinatorics **32** (2011), 628–638.
- [78] F. Foucaud, S. Gravier, R. Naserasr, A. Parreau, P. Valicov, *Identifying codes in line graphs*, soumis. Disponible en ligne à l'url <http://hal.archives-ouvertes.fr/hal-00605406>.
- [79] F. Foucaud, R. Klasing, A. Kosowski, A. Raspaud, *Bounds on the size of identifying codes in triangle-free graphs*, soumis (2010). Disponible en ligne à l'url <http://hal.archives-ouvertes.fr/hal-00530172>.
- [80] L. Frachet, G. Massonnet, *Modélisation d'un problème de surveillance par un jeu stochastique*, mémoire de projet de spécialité ENSIMAG (juin 2008).
- [81] FREEDROIDRPG, jeu de donjons sous licence GPL  
<http://freedroid.sourceforge.net/>
- [82] M. Frick, M. Grohe, *The complexity of first-order and monadic second-order logic revisited*, Annals of Pure and Applied Logic **130**(1-3) (2004), 3–31.

- [83] A. Frieze, R. Martin, J. Moncel, M. Ruszinkó, C. Smyth, *Codes identifying sets of vertices in random networks*, *Discrete Mathematics* **307(9-10)** (2007), 1094–1107.
- [84] Z. Füredi, *On  $r$ -cover-free families*, *Journal of Combinatorial Theory Series A* **73(1)** (1996), 172–173.
- [85] H. Gan, D. Fera, J. Zorn, N. Shiffeldrim, U. Laserson, N. Kim, T. Schlick, *RAG : RNA-As-Graphs database – concepts, analysis, and features*, *Bioinformatics* **20** (2004), 1285–1291.
- [86] H. Gan, S. Pasquali, T. Schlick, *Exploring the repertoire of RNA secondary motifs using graph theory; implications for RNA design*, *Nucleic Acids Research* **31(11)** (2003), 2926–2943.
- [87] A. Gardiner, *Imprimitive distance-regular graphs and projective planes*, *Journal of Combinatorial Theory Series B* **16** (1974), 274–281.
- [88] M. R. Garey, D. S. Johnson, *Computers and Intractability : A Guide to the Theory of NP-Completeness*, Freeman (1979).
- [89] S. Gawiejnowicz, *Time-Dependent Scheduling*, *Monographs in Theoretical Computer Science – An EATCS Series*, Springer (2008).
- [90] J. Gimbel, B. van Gorden, M. Nicolescu, Ch. Umstead, N. Vaiana, *Location with dominating sets*, *Congressus Numerantium* **151** (2001), 129–144.
- [91] S.W. Golomb, L.R. Welch, *Perfect codes in the Lee metric and the packing of polyominoes*, *SIAM Journal on Applied Mathematics* **18** (1970) 302–317.
- [92] R. Grappe, N. Brauner, J. Moncel, *Un problème de surveillance : modélisation et simulation avec la théorie des jeux*, congrès ROADEF 2008, 25–27 février 2008, Clermont-Ferrand.
- [93] S. Gravier, R. Klasing, J. Moncel, *Hardness results and approximation algorithms for identifying codes and locating-dominating codes in graphs*, *Algorithmic Operations Research* **3(1)** (2008), 43–50.
- [94] S. Gravier, M. Kovše, M. Mollard, J. Moncel, A. Parreau, *New results on variants of covering codes in Sierpiński graphs*, à paraître dans *Designs, Codes and Cryptography*. Disponible en ligne à l'url <http://www-fourier.ujf-grenoble.fr/~mollard/soumis/IDcodes-SierpinskiGraphs-V2.pdf>.
- [95] S. Gravier, J. Moncel, *Construction of codes identifying sets of vertices*, *Electronic Journal of Combinatorics* **12(1)** (2005), R13.
- [96] S. Gravier, J. Moncel, *On graphs having a  $V \setminus \{x\}$  set as an identifying code*, *Discrete Mathematics* **307(3-5)** (2007), 432–434.

- [97] S. Gravier, J. Moncel, C. Payan, *A generalization of the pentomino exclusion problem : Dislocation of graphs*, *Discrete Mathematics* **307(3-5)** (2007), 435–444.
- [98] S. Gravier, J. Moncel, A. Semri, *Identifying codes of cycles*, *European Journal of Combinatorics* **27** (2006), 767–776.
- [99] S. Gravier, J. Moncel, A. Semri, *Codes identifiants : une approche combinatoire d'un problème de détection de défaillances dans les réseaux*, CO-SI'06 (International Colloquium on Optimisation and Information Systems, Alger 11–13 juin 2006), publié dans les proceedings pp. 186–196.
- [100] S. Gravier, J. Moncel, A. Semri, *Identifying codes of cartesian product of two cliques of the same size*, *Electronic Journal of Combinatorics* **15(1)** (2008), N4.
- [101] S. Gravier, B. Ycart, *S-constrained random matrices*, Fourth Colloquium on Mathematics and Computer Science – Algorithms, Trees, Combinatorics and Probabilities, Nancy (septembre 2006), *Discrete Mathematics and Theoretical Computer Science Proceedings* (2006) **357–364**.
- [102] R. Hammack, W. Imrich, S. Klavžar, *Handbook of Product Graphs*, CRC Press (2011).
- [103] F. Havet, *Minimum-density identifying codes in grids*, manuscrit (2010).
- [104] T. W. Haynes, S. T. Hedetniemi, P. J. Slater, *Fundamentals of Domination in Graphs*, Marcel Dekker (1998).
- [105] T. Haynes, D. Knisley, E. Seier, Y. Zou, *A quantitative analysis of secondary RNA structure using domination based parameters on trees*, *BMC Bioinformatics* **7** (2006), 108.
- [106] Y. He, W. Zhong, H. Gu, *Improved algorithms for two single machine scheduling problems*, *Theoretical Computer Science* **363** (2006), 257–265.
- [107] A. Hinz, S. Klavžar, U. Milutinović, C. Petr, *The Tower of Hanoi : Myths and Maths*, à paraître.
- [108] I. Honkala, T. Laihonen, *On identifying codes in the hexagonal mesh*, *Information Processing Letters* **89** (2004), 9–14.
- [109] I. Honkala, T. Laihonen, *On identification in the triangular grid*, *Journal of Combinatorial Theory Series B* **91** (2004), 67–86.
- [110] I. Honkala, A. Lobstein, *On identifying codes in binary Hamming spaces*, *Journal of Combinatorial Theory Series A* **99** (2002), 232–243.
- [111] I. Honkala, A. Lobstein, *On the density of identifying codes in the square lattice*, *Journal of Combinatorial Theory Series B* **85** (2002), 297–306.

- [112] M. Horňák, R. Soták, *Observability of complete multipartite graphs with equipotent parts*, *Ars Combinatoria* **41** (1995), 289–301.
- [113] F. Hwang, V. Sós, *Non adaptive hypergeometric group testing*, *Studia Scientiarum Mathematicarum Hungaricae*, **22(1-4)** (1987), 257–263.
- [114] M. Jakovac, *A 2-parametric generalization of Sierpiński gasket graphs*, à paraître dans *Ars Combinatoria*.
- [115] D. S. Johnson, *Approximation algorithms for combinatorial problems*, *Journal of Computer and System Sciences* **9** (1974), 256–278.
- [116] V. Junnila, *On identifying and locating-dominating codes*, thèse de doctorat de l’université de Turku (2011). Disponible en ligne à l’url <http://www.doria.fi/handle/10024/69771>.
- [117] V. Junnila, *Adaptive identification of sets of vertices in graphs*, soumis à *Discrete Mathematics and Theoretical Computer Science*.
- [118] V. Junnila, T. Laihonen, *Optimal identifying codes in cycles and paths*, à paraître dans *Graphs and Combinatorics*.
- [119] M. Juvan, B. Mohar, A. Graovac, S. Klavžar, J. Žerovnik, *Fast computation of the Wiener index of fasciagraphs and rotagraphs*, *Journal of Chemical Information and Computer Science* **35** (1995), 834–840.
- [120] M. Karpovsky, K. Chakrabarty, L. Levitin, *On a New Class of Codes for Identifying Vertices in Graphs*, *IEEE Transactions on Information Theory* **44(2)** (1998), 599–611.
- [121] M. Karpovsky, K. Chakrabarty, L. Levitin, D. Avresky, *On the covering of vertices for fault diagnosis in hypercubes*, *Information Processing Letters* **69** (1999), 99–103.
- [122] W. Kautz, R. Singleton, *Nonrandom binary superimposed codes*, *IEEE Transactions on Information Theory* **10(4)**, (1964), 363–377.
- [123] S. Klavžar, B. Mohar, *Crossing numbers of Sierpiński-like graphs*, *Journal of Graph Theory* **50** (2005), 186–198.
- [124] S. Klavžar, U. Milutinović, *Graphs  $S(n, k)$  and a variant of the Tower of Hanoi problem*, *Czechoslovak Mathematical Journal* **47(122)** (1997), 95–104.
- [125] S. Klavžar, U. Milutinović, C. Petr, *1-perfect codes in Sierpiński graphs*, *Bulletin of the Australian Mathematical Society* **66** (2002), 369–384.
- [126] S. Klavžar, A. Vesel, *Computing graph invariants on rotagraphs using dynamic algorithm approach : the case of (2,1)-colorings and independence numbers*, *Discrete Applied Mathematics* **129(2-3)** (2003), 449–460.

- [127] S. Klavžar, J. Žerovnik, *Algebraic approach to fasciagraphs and rotagraphs*, Discrete Applied Mathematics **68(1-2)** (1996), 93–100.
- [128] M. Laifenfeld, *Coding for network applications : Robust identification and distributed resource allocation*, thèse de doctorat de l'université de Boston (2008).
- [129] M. Laifenfeld, A. Trachtenberg, *Disjoint Identifying-Codes for Arbitrary Graphs*, ISIT 05, 2005 IEEE International Symposium on Information Theory, 4–9 septembre 2005, Adelaïde (Australie), publié dans les proceedings pp. 244–248.
- [130] M. Laifenfeld, A. Trachtenberg, *Identifying codes and covering problems*, IEEE Transactions on Information Theory **54(9)** (2008), 3929–3950.
- [131] M. Laifenfeld, A. Trachtenberg, R. Cohen, D. Starobinski, *Joint Monitoring and Routing in Wireless Sensor Networks Using Robust Identifying Codes*, Mobile networks and applications **14(4)** (2009), 415–432.
- [132] T. Laihonen, *Sequences of optimal identifying codes*, IEEE Transactions on Information Theory **48** (2002), 774–776.
- [133] T. Laihonen, *On cages admitting identifying codes*, European Journal of Combinatorics **29(3)** (2008), 737–741.
- [134] T. Laihonen, J. Moncel, *On graphs admitting codes identifying sets of vertices*, Australasian Journal of Combinatorics **41** (2008), 81–91.
- [135] T. Laihonen, S. Ranto, *Codes identifying sets of vertices*, Lecture Notes in Computer Science **2227** (2001), 82–91.
- [136] S. L. Lipscomb, *Fractals and universal spaces in dimension theory*, Springer-Verlag, Berlin (2009).
- [137] S. L. Lipscomb, J. C. Perry, *Lipscomb's  $L(A)$  space fractalized in Hilbert's  $l^2(A)$  space*, Proceedings of the American Mathematical Society **115** (1992), 1157–1165.
- [138] C. Y. Lee, *Some properties of nonbinary error-correcting codes*, IRE Transactions on Information Theory **4(4)** (1958), 77–82.
- [139] M. Livingston, Q. F. Stout, *Constant time computation of minimum dominating sets*, Congressus Numerantium **105** (1994), 116–128.
- [140] A. Lobstein, *Watching systems, identifying, locating-dominating and discriminating codes in graphs*, bibliographie. Disponible en ligne à l'url <http://www.infres.enst.fr/~lobstein/debutBIBidetlocdom.pdf>.
- [141] R. Martin, B. Stanton, *Lower bounds for identifying codes in some infinite grids*, Electronic Journal of Combinatorics **17** (2010), R122.

- [142] S. L. Mitchell, E. J. Cockayne, S. T. Hedetniemi, *Linear algorithms on recursive representations of trees*, Journal of Computer and System Sciences **18(1)** (1979), 76–85.
- [143] J. Moncel, *Monotonicity of the minimum cardinality of an identifying code in the hypercube*, Discrete Applied Mathematics **154(6)** (2006), 898–899.
- [144] J. Moncel, *On graphs on  $n$  vertices having an identifying code of cardinality  $\lceil \log_2(n+1) \rceil$* , Discrete Applied Mathematics **154(14)** (2006), 2032–2039.
- [145] J. Moncel, *Constructing codes identifying sets of vertices*, Designs, Codes and Cryptography **41(1)** (2006), 23–31.
- [146] J. Moncel, *Codes identifiants dans les graphes*, thèse de doctorat de l’université de Grenoble (2005). Disponible en ligne à l’url <http://tel.archives-ouvertes.fr/tel-00010293/>.
- [147] J. Moncel, *Minimisation des encours sur une machine pour des tâches sujettes à des phénomènes d’usure ou d’apprentissage*, congrès ROADEF 2010, 24–26 février 2010, Toulouse.
- [148] J. Moncel, G. Finke, V. Jost, *A general result on single machine scheduling problems subject to learning effects or deteriorating jobs*, soumis à OR Letters. Disponible en ligne à l’url <http://hal.archives-ouvertes.fr/hal-00647789/fr/>.
- [149] J. Moncel, G. Finke, V. Jost, *Single-machine scheduling problems with position-dependent processing times*, 13th International Conference on Project Management and Scheduling (1–4 avril 2012, Louvain – Belgique).
- [150] J. Moncel, J. Thiery, A. Wasserhole, *Computational performances of a simple interchange heuristic for a scheduling problem with an availability constraint*, soumis. Une version préliminaire est disponible en ligne sur HAL : <http://hal.archives-ouvertes.fr/hal-00181166/fr/>.
- [151] J. Moncel, J. Thiery, A. Wasserhole, *Computational performances of a simple interchange heuristic for a scheduling problem with an availability constraint*, 13th International Conference on Project Management and Scheduling, 1–4 avril 2012, Louvain (Belgique).
- [152] R. Motwanin M. Sudan, *Computing roots of graphs is hard*, Discrete Applied Mathematics **54(1)** (1994), 81–88.
- [153] Ch. Papadimitriou, M. Yannakakis, *Optimization, approximation, and complexity classes*, Journal of Computer and System Sciences **43** (1991), 425–440.

- [154] T. Müller, J.-S. Sereni, *Identifying and locating-dominating codes in (random) geometric networks*, *Combinatorics Probability and Computing* **18(6)** (2009), 925–952.
- [155] S. Ranto, *Identifying and locating-dominating codes in binary Hamming spaces*, thèse de doctorat de l'université de Turku (2007). Disponible en ligne à l'url <http://users.utu.fi/samano/>.
- [156] A. Raspaud, L.-D. Tong, *The minimum identifying code graphs*, à paraître dans *Discrete Applied Mathematics*.
- [157] S. Ray, D. Starobinski, A. Trachtenberg, R. Ungrangsi, *Robust location detection with sensor networks*, *IEEE Journal on Selected Areas in Communications* **22(6)** (2004), 1016–1025.
- [158] A. Rényi, *On a problem of information theory*, *Publications of the Mathematical Institute of the Hungarian Academy of Sciences* **6B** (1962), 505–516.
- [159] D. L. Roberts, F. S. Roberts, *Locating sensors in paths and cycles : The case of 2-identifying codes*, *European Journal of Combinatorics* **29** (2008), 72–82.
- [160] M. Ruszinkó, *On a 2-dimensional search problem*, *Journal of Statistical Planning and Inference* **37(3)** (1993), 371–383.
- [161] M. Ruszinkó, *On the upper bound of the size of the  $r$ -cover-free families*, *Journal of Combinatorial Theory Series A* **66(2)** (1994), 302–310.
- [162] C. Sadfi, B. Penz, C. Rapine, J. Błażewicz, P. Formanowicz, *An improved approximation algorithm for the single machine total completion time scheduling problem with availability constraints*, *European Journal of Operations Research* **161** (2005), 3–10.
- [163] R.R. Singleton, *On minimal graphs of maximum even girth*, *Journal of Combinatorial Theory* **1** (1966), 306–332.
- [164] R. Skaggs, *Identifying vertices in graphs and digraphs*, thèse de doctorat de l'université d'Afrique du Sud (2007). Disponible en ligne à l'url [uir.unisa.ac.za/bitstream/handle/10500/2226/](http://uir.unisa.ac.za/bitstream/handle/10500/2226/).
- [165] A. Spalding, *Min-plus algebra and graph domination*, thèse de doctorat de l'université du Colorado (1998).
- [166] B. Stanton, *On vertex identifying codes for infinite lattices*, thèse de doctorat de l'université d'Iowa (2011). Disponible en ligne à l'url [http://orion.math.iastate.edu/rymartin/students/StantonB\\_PhD\\_S11.pdf](http://orion.math.iastate.edu/rymartin/students/StantonB_PhD_S11.pdf).

- [167] W.-C. S. Suen, *A correlation inequality and a Poisson limit theorem for nonoverlapping balanced subgraphs of a random graph*, *Random Structures and Algorithms* **1(2)** (1990), 231–242.
- [168] J. Suomela, *Approximability of identifying codes and locating–dominating codes*, *Information Processing Letters* **103(1)** (2007), 28–33.
- [169] A. M. Teguia, A. P. Godbole, *Sierpiński gasket graphs and some of their properties*, *Australasian Journal of Combinatorics* **35** (2006), 181–192.
- [170] J. Thiery, A. Waserhole, J. Moncel, *Schéma d’approximation polynomial pour la minimisation d’encours sur une machine avec période de maintenance*, congrès ROADEF 2008, 25–27 février 2008, Clermont-Ferrand.
- [171] R. Ungrangsi, A. Trachtenberg, D. Starobinski, *An Implementation of Indoor Location Detection Systems Based on Identifying Codes*, *Lecture Notes in Computer Science* **3283** (2004), 175–189.
- [172] V. Vazirani, *Approximation algorithms*, Springer-Verlag (2001).
- [173] V. G. Vizing, *Some unsolved problems in graph theory*, *Russian Mathematical Surveys* **23(6)** (1968), 125–141.
- [174] P. Winkler, *Random Structures and Zero-One Laws*, *Proceedings of the NATO Advanced Study Institute, NATO ASI Series C* **411** (1993), 399–420.
- [175] T. Wright, *Factors affecting the cost of airplanes*, *Journal of Aeronautical Sciences* **3** (1936), 122–128.
- [176] Y. Xiao, C. Hadjicostis, K. Thulasiraman, *The d-identifying codes problem for vertex identification in graphs : probabilistic analysis and an approximation algorithm*, *COCOON 2006, 12th Annual International Computing and Combinatorics Conference*, 15–18 Août 2006, Taipei (Taiwan).
- [177] M. Xu, K. Thulasiraman, X.-D. Hu, *Identifying codes of cycles with odd orders*, *European Journal of Combinatorics* **29** (2008) 1717–1720.
- [178] L. Yelle, *The Learning Curve : Historical Review and Comprehensive Survey*, *Decision Sciences* **10** (1979), 302–328.
- [179] J. Žerovnik, *Deriving formulas for domination numbers of fasciagraphs and rotagraphs*, *Lecture Notes in Computer Science* **1684** (1999), 559–568.



ANNEXE A

# CV détaillé

---

## État civil

---

---

### **Julien Moncel**

Bureau A25

IUT de Rodez

33 avenue du 8 mai 45

12000 Rodez

Courriel : [julien.moncel@iut-rodez.fr](mailto:julien.moncel@iut-rodez.fr)

33 ans

Marié, un enfant

Nationalité française

## Cursus

---

---

- 2010 – ... **Maître de conférences (section 61)**  
Université Toulouse 1 Capitole  
Département QLIO (Qualité, Logistique Industrielle et Organisation) de l'IUT de Rodez  
  
Laboratoire LAAS-CNRS (Laboratoire d'Analyse et d'Architecture des Systèmes – UPR 8001)  
Équipe de recherche MOGISA (Modélisation Optimisation et Gestion Intégrée de Systèmes d'Activités)
- 2006 – 2010 **Maître de conférences (section 27)**  
Grenoble INP – Génie Industriel  
  
Laboratoire G-SCOP (Laboratoire Grenoblois des Sciences pour la Conception, l'Optimisation et la Production – UMR 5272)  
Équipe ROSP (Recherche Opérationnelle pour les Systèmes de Production)

- 2005 **Qualification aux fonctions de maître de conférences**  
En sections 25, 26, 27, et 61
- 2005 – 2006 **ATER**  
ENSIMAG (École Nationale Supérieure d'Informatique et de Mathématiques Appliquées de Grenoble)
- 2002 – 2005 **Doctorat d'informatique et mathématiques appliquées**  
Université Joseph Fourier – Grenoble  
Laboratoire Leibniz  

	<i>Titre de la thèse : Codes identifiants dans les graphes</i>
	<i>Directeur de thèse : Sylvain Gravier</i>
	<i>Date de soutenance : 27 juin 2005</i>
- 2002 – 2005 **Moniteur en mathématiques**  
Université Pierre Mendès France – Grenoble
- 2001 – 2002 **DEA de Recherche Opérationnelle, Combinatoire et Optimisation**  
Institut National Polytechnique de Grenoble
- 2000 – 2002 **Diplôme d'ingénieur**  
École Nationale Supérieure d'Informatique et de Mathématiques Appliquées de Grenoble (ENSIMAG)
- 1999 – 2002 **Magistère de mathématiques et applications**  
Université Joseph Fourier – Grenoble

## Recherche

---

### *Résumé*

---

Thématiques Recherche opérationnelle, théorie des graphes, ordonnancement, optimisation et géométrie discrète

Valorisation 20 publications dans des revues internationales  
9 conférences internationales avec comité de lecture

Encadrement Deux thèses (en co-direction)  
Deux mémoires de Master 2 Recherche

---

## Direction et expertise de travaux de recherche

---

► Direction de travaux de recherche

- 2008 – ...    **Thèse de doctorat de Marwane Bouznif (co-direction)**  
                   | *Thématique : Notion de localité de problèmes combinatoires sur les fasciagraphes et les rotagraphes — applications aux codes identifiants*  
                   | *Lieu de la thèse : Laboratoire G-SCOP*  
                   | *Directrice de thèse : Myriam Preissmann*
- 2008 – 2011    **Thèse de doctorat de Julien Darlay (co-direction)**  
                   | *Thématique : Analyse combinatoire de données — structures et optimisation*  
                   | *Lieu de la thèse : Laboratoire G-SCOP*  
                   | *Directrice de thèse : Nadia Brauner*
- 2009            **Mémoire de Master 2 « Recherche Opérationnelle Combinatoire et Optimisation » de Christophe-Marie Duquesne**  
                   | *Titre : The fleet assignment problem*  
                   | *Entreprise partenaire : Amadeus (Sofia-Antipolis)*
- 2007            **Mémoire de Master 2 « Recherche Opérationnelle Combinatoire et Optimisation » d’Onur Çelebi**  
                   | *Titre : Problème de localisation et de dimensionnement de site de concentration en contexte incertain*  
                   | *Entreprise partenaire : France Télécom R & D (Paris)*
- 2007 – 2010    **Encadrement d’une dizaine de projets de spécialité « Recherche Opérationnelle » à l’ENSIMAG**  
                   | *Deux communications dans des conférences nationales et un manuscrit soumis issus de ces travaux*

► Expertise

- Arbitrage      Arbitre pour les journaux suivants : *Discrete Mathematics, Discrete Applied Mathematics, Parallel Computing, European Journal of Combinatorics, Electronic Journal of Combinatorics, Journal of Graph Theory, Discrete Mathematics and Theoretical Computer Science, 4OR, et IIE Transactions*

Évaluation Rapporteur de la thèse de Ville Junnila  
*Titre de la thèse : On identifying and locating-dominating codes*  
*Lieu de la thèse : Université de Turku (Finlande)*  
*Directeur de thèse : Tero Laihonen*  
*Thèse soutenue le 22 juin 2011*

## *Animation scientifique*

---

### ► Organisation d'évènements

- 2003 – ... **Co-organisateur d'une douzaine d'actions de vulgarisation et de diffusion de la culture scientifique auprès des scolaires ou du grand public**  
*Exemples : fêtes de la science, festivals « remue-méninges », animations dans des établissements scolaires, animations auprès d'associations*
- 26 – 28 mars 2010 **Co-organisateur du congrès « maths en jeans » à Grenoble**
- 2007 – 2010 **Responsable du SÉTRO**  
 Séminaire de travail de l'équipe ROSP du laboratoire G-SCOP
- 3 – 6 déc. 2007 **Co-organisateur du workshop « Codes and discrete probability »**  
 Institut Fourier
- 12 novembre 2004 **Co-organisateur de la 2ème journée de l'ERTé « maths à modeler »**  
 Laboratoire Leibniz
- 29 sept. – 1er oct. 2004 **Co-organisateur des journées « graphes et algorithmes »**  
 Maison Jean Kuntzmann – Grenoble
- 12 décembre 2003 **Co-organisateur de la journée de l'action spécifique CNRS « algorithmique des grands graphes »**  
 Laboratoire Leibniz

11 sep- **Co-organisateur de la 1ère journée de l'ERTé « maths à mo-**  
 tembre **deler »**  
 2003 Laboratoire Leibniz

► Projets scientifiques nationaux

2012 – 2013 **Membre du Projet PEPS « MISÈRE » de l'INS2I**  
 Graphes pour les jeux combinatoires en convention Misère  
 | Partenaires : LaBRI (Bordeaux), GAMA (Lyon), Institut Fourier  
 | (Grenoble), LIMOS (Clermont-Ferrand)

2010 – ... **Membre de la fédération de recherche « maths à modeler »**  
 Équipe pluridisciplinaire composée de chercheurs en  
 mathématiques discrètes, en didactique, et en sciences de  
 l'éducation  
 Prolongement de l'ERTé « maths à modeler » (2003–2010)  
 | Site web du projet : <http://mathsamodeler.ujf-grenoble.fr/>

2009 – 2011 **Membre du projet ANR IDEA**  
 Concerne les approches dynamiques et adaptatives des codes  
 identifiants  
 | Partenaires : LaBRI (Bordeaux), Institut Fourier (Grenoble)  
 | Page web : <https://idea.labri.fr/>

2003 – 2010 **Membre de l'Équipe Recherche Technologie éducation**  
**« maths à modeler »**  
 Équipe pluridisciplinaire composée de chercheurs en  
 mathématiques discrètes, en didactique, et en sciences de  
 l'éducation  
 | Site web du projet : <http://mathsamodeler.ujf-grenoble.fr/>

► Responsabilités collectives

2012 – ... **Chef de département**  
 Département QLIO IUT Rodez

- 2008 – 2010 **Responsable des admis sur titre**  
Grenoble INP – Génie Industriel  
| *Organisation du concours de recrutement*  
| *Suivi et accompagnement des étudiants*  
| *Mise en place d'un système de tutorat et d'une semaine de pré-  
rentrée pour la remise à niveau*
- 2008 **Membre de la commission de spécialistes de l'INPG**  
Informatique et mathématiques appliquées (sections 25, 26 et 27)  
| *Participation au concours de recrutement pour 3 postes*
- 2006 – 2008 **Membre du conseil de l'école**  
Grenoble INP – Génie Industriel
- 2004 – 2005 **Responsable de l'atelier « maths à modeler »**  
Centre d'initiation à l'enseignement supérieur de Grenoble
- 2003 **Membre du comité d'organisation du séminaire des doct-  
rants**  
Laboratoire Leibniz
- 2003 – 2005 **Membre du conseil de laboratoire et du conseil du personnel**  
Laboratoire Leibniz

### *Rayonnement scientifique*

---

► Prix et distinctions divers

- 2011 – 2015 **Titulaire de la prime d'excellence scientifique**  
Université Toulouse 1 Capitole
- Août 2010 **2ème prix ex-æquo du concours de la SCM (Société de Calcul  
Mathématique)**  
| *Thème du concours : Conception d'un réseau de transport  
d'électricité*  
| *Concours organisé conjointement avec la FFJM (Fédération  
Française de Jeux Mathématiques) en partenariat avec RTE  
France (Réseau de Transport d'Électricité)*

---

2009 – 2010 **Titulaire de la prime d'excellence scientifique**  
Grenoble INP

► Mobilité internationale

Juillet **Séjour de recherche**  
– août 2006 Institut des mathématiques Alfréd Rényi à Budapest (Hongrie)

Juin 2006 **Séjour de recherche invité**  
Université de Turku (Finlande)

Mars 2006 **Séjour de recherche**  
Université de Tel-Aviv (Israël)

Janvier – **Séjour de recherche**  
septembre 2006 Laboratoire d'informatique et d'automatique de l'académie des sciences de Hongrie à Budapest

► Projets scientifiques internationaux

2005–2008 **Membre de l'accord de coopération franco-algérien du Comité Mixte d'Évaluation et de Prospective de l'Université des Sciences et des Techniques Houari Boumédiène**

*Titre : « mathématiques discrètes »*

*Partenaire : Université des Sciences et de la Technologie Houari Boumédiène – Alger*

2005–2006 **Membre du Programme d'Action Intégrée BALATON**

*Titre : « combinatoire des problèmes de communication dans les réseaux »*

*Partenaire : Laboratoire d'informatique et d'automatique de l'académie des sciences de Hongrie*

Enseignement

---

*Cours*

---

2011 – ... **Informatique avancée**  
IUT de Rodez – Licence Professionnelle Pilotage des Activités Logistiques (niveau L3)

- 2010 – ... **Systèmes de gestion de bases de données**  
IUT de Rodez, département QLIO – Qualité, Logistique Industrielle et Organisation (niveau L1)
- 2010 – ... **Conception des systèmes d'information**  
IUT de Rodez, département QLIO (niveau L1)
- 2010 – ... **Bases de la planification de projet**  
IUT de Rodez, département QLIO (niveau L1)
- 2010 – ... **Aménagement du poste de travail**  
IUT de Rodez, département QLIO (niveau L2)
- 2010 – ... **Gestion physique des stocks**  
IUT de Rodez, département QLIO (niveau L2)
- 2010 – ... **Structure et organisation des entreprises**  
IUT de Rodez, département QLIO (niveau L1)
- 2010 – ... **Typologie des produits et des procédés**  
IUT de Rodez, département QLIO (niveau L1)
- 2010 – ... **Mathématiques pour la gestion**  
IUT de Rodez, département GEA – Gestion des entreprises et administration (niveau L1)
- 2009 – 2010 **Théorie des jeux et décision**  
Grenoble INP – Génie Industriel (niveau M2 et M2 recherche)
- 2009 – 2010 **Bureautique avancée**  
Polytech'Grenoble (niveau L1)
- 2009 – 2010 **Bases de données**  
IUT de Grenoble, département GEA – Gestion des entreprises et administration (niveau L1)
- 2009 – 2010 **Bases de données avancées**  
UFR d'informatique, université Joseph Fourier – Grenoble (niveau M1)

- 
- 2009 – 2010 **Graphes et optimisation : fondements théoriques, modèles et algorithmes**  
Formation continue Grenoble INP
- 2008 – 2010 **Recherche opérationnelle et planification de la production**  
Grenoble INP – Génie Industriel (niveau M1)
- 2008 – 2009 **Probabilités et statistiques**  
Grenoble INP – Génie Industriel (niveau L3)
- 2008 – 2010 **Graphes et structures discrètes**  
Grenoble INP (niveau M2 recherche)
- 2007 – 2010 **Management de la production et des services**  
Grenoble INP – ENSIMAG (niveau M1)
- 2007 – 2010 **Introduction à la recherche opérationnelle**  
Formation continue Grenoble INP
- 2007 – 2010 **Programmation par contraintes et recherche opérationnelle**  
Université Joseph Fourier – Grenoble (niveau M1)
- 2007 – 2010 **Outils formels : complexité et cryptographie**  
Polytech'Grenoble (niveau M1)
- 2007 – 2008 **Optimisation combinatoire et jeux**  
Grenoble INP (niveau M2 recherche)
- 2006 – 2008 **Aide à la décision multicritère**  
Grenoble INP – Génie Industriel (niveau M1)
- 2006 – 2008 **Modèles à événements discrets**  
Grenoble INP – Génie Industriel (niveau M1)
- 2006 – 2008 **Ordonnancement de la production**  
Grenoble INP – Génie Industriel (niveau M1)
- 2006 – 2008 **Probabilités**  
Grenoble INP – Génie Industriel (niveau L3)

- 2006 – 2008 **Recherche opérationnelle**  
Grenoble INP – Génie Industriel (niveau L3)
- 2005 – 2007 **Compilation et langages**  
Grenoble INP – ENSIMAG (niveau M1)
- 2005 – 2007 **Théorie des langages**  
Grenoble INP – ENSIMAG (niveau L3)
- 2005 – 2006 **Recherche opérationnelle**  
Grenoble INP – ENSIMAG (niveau M1)
- 2004 – 2005 **Bureautique**  
Université Pierre Mendès France – Grenoble (niveau L1)
- 2002 – 2005 **Statistiques descriptives**  
Université Pierre Mendès France – Grenoble (niveau L1)

### *Projets encadrés*

---

Plus d'une quarantaine de projets encadrés en tout, parmi lesquels :

- 2011– ... **Stages PILAL**  
IUT Rodez (niveau L2)  
| *Stages en entreprise de 3 mois des étudiants en Licence Professionnelle PILAL (pilotage des activités logistique)*
- 2010– ... **Stages QLIO2**  
IUT Rodez (niveau L2)  
| *Stages en entreprise de 3 mois des étudiants en 2ème année de DUT QLIO*
- 2010– ... **Projets tuteurés**  
IUT Rodez (niveaux L1-L2)  
| *Projets en groupe sur toute la durée de la scolarité des étudiants*
- 2006 – 2010 **Projets de fin d'étude**  
Grenoble INP – Génie Industriel (niveau M2)  
| *Stages en entreprise de 5 mois minimum*

- 2006 – 2010 **Stages ingénieur adjoint**  
Grenoble INP – Génie Industriel (niveau M1)  
| *Stages en entreprise de 3 mois*
- 2006 – 2010 **Études de terrain**  
Grenoble INP – Génie Industriel (niveau M1)  
| *Stages en entreprise de 100 heures en trinôme*
- 2006 – 2010 **Projets de spécialité**  
Grenoble INP – ENSIMAG (niveau M1)  
| *Stages en laboratoire d'un mois en binômes ou trinômes*
- 2006 – 2010 **Travaux d'étude et de recherche**  
Grenoble INP – ENSIMAG (niveau M1)  
| *Stages en laboratoire de 50 heures réparties sur 4 mois*



# Échantillon de manuscrits

---

Les manuscrits joints au présent document d'habilitation constituent un échantillon représentatif des travaux réalisés sur les codes identifiants dans les graphes.

► Le manuscrit

S. Gravier, J. Moncel, A. Semri, *Identifying codes of cycles*, European Journal of Combinatorics **27** (2006), 767–776

concerne l'étude des cycles. Il contient des résultats améliorant ceux de Bertrand, Charon, Hudry et Lobstein [25] ainsi que ceux de Gimbel, van Gorden, Nicolescu, Umstead, et Vaiana [90]. Y est introduit, en particulier, un graphe auxiliaire dont l'étude des transversaux permet de dériver des bornes sur les codes identifiants dans les cycles.

► Le manuscrit

A. Frieze, R. Martin, J. Moncel, M. Ruzinkó, C. Smyth, *Codes identifying sets of vertices in random networks*, Discrete Mathematics **307(9-10)** (2007), 1094–1107

introduit l'étude de la construction de graphes admettant des codes  $\ell$ -set-identifiants de faible cardinalité. Les outils utilisés relèvent de méthodes probabilistes. Un premier lien avec les codes superposés y est de plus établi.

► La question de la construction explicite de familles de graphes admettant des codes de faible cardinalité est abordée dans le manuscrit

S. Gravier, J. Moncel, *Construction of codes identifying sets of vertices*, Electronic Journal of Combinatorics **12(1)** (2005), R13

Le résultat principal de cet article est la construction de familles de graphes en utilisant des codes superposés et des graphes obtenus par la théorie des plans projectifs.

► Une construction explicite de familles de graphes dont les caractéristiques sont identiques à celles obtenues par des constructions probabilistes est donnée dans

J. Moncel, *Constructing codes identifying sets of vertices*, *Designs, Codes and Cryptography* **41(1)** (2006), 23–31

Cette construction est inspirée d'un algorithme similaire développé par Hwang et Sós dans le cas des codes superposés.

► Dans le manuscrit

S. Gravier, R. Klasing, J. Moncel, *Hardness results and approximation algorithms for identifying codes and locating-dominating codes in graphs*, *Algorithmic Operations Research* **3(1)** (2008), 43–50

sont établis des résultats d'approximabilité sur les codes identifiants. Ces résultats utilisent les liens avec SET COVER et DOMINATION.

► Dans le manuscrit

T. Laihonen, J. Moncel, *On graphs admitting codes identifying sets of vertices*, *Australasian Journal of Combinatorics* **41** (2008), 81–91

nous étudions la question de la construction de graphes admettant des codes  $\ell$ -set-identifiants *via* le produit cartésien de graphes. Nous donnons de plus la réponse à une question sur l'existence de graphes  $\ell$ -réguliers à  $\Theta(\ell^2)$  sommets admettant un code  $\ell$ -set-identifiant.

► Enfin, le manuscrit

Y. Ben-Haim, S. Gravier, A. Lobstein, J. Moncel, *Adaptive identification in graphs*, *Journal of Combinatorial Theory Series A* **115(7)** (2008), 1114–1126

établit les premiers résultats généraux sur les codes identifiants adaptatifs dans les graphes réguliers, ainsi qu'une série de résultats sur les tores dans la grille carrée.