Applications of semidefinite programming to coding theory

Christine Bachoc, Université Bordeaux 1

Semidefinite programming is a subfield of convex optimization which contains lienar programming as a special case. Thanks to the so-called interior point methods, semidefinite programs can be solved to a given approximation in polynomial time. Many practical problems in areas as different as combinatorial and polynomial optimization, statistics, control theory, and of course information theory can be modeled or approximated as semidefinite programming problems. In this talk we will review some of these applications concerned with information theory, namely sensor network localization, maximum likelihood decoding in MIMO communication model, and bounds on performance of binary codes.

Les bons sous-codes des codes de Reed–Muller.

Étant donnés un corps fini $\mathbb{F}_q$ et deux entiers r et m, le *code de Reed–Muller* q–aire $RM_q(r,m)$ est l'image de l'application d'évaluation

$$ev : \left\{ \begin{array}{ccc} \mathbb{F}_q[X_1,\ldots,X_m]_{\leq r} & \to & \mathbb{F}_q^{q^m} \\ P & \mapsto & (P(x_1,\ldots,x_m) \mid (x_1,\ldots,x_m) \in \mathbb{F}_q^m) \end{array} \right. .$$

Le code de Reed–Muller *projectif* $PRM_q(r,m)$ est une extension du code $RM_q(r,m)$ obtenue par "évaluation" de formes homogènes de degré r en les points de l'espace projectif $\mathbb{P}^m$.

Il est bien connu que les mots de poids minimal d'un code de Reed–Muller (ou de Reed–Muller projectif) proviennent de l'évaluation de polynômes qui sont produit de formes de degré 1. Plus généralement, les mots petits poids proviennent des polynômes très friables.

Partant de cette constatation, on se focalise sur les codes de Reed–Muller en deux variables $RM_q(r,2)$ et recherche de bons sous-codes linéaires de ces dernier. La méthode consiste à rechercher des sous-espaces vectoriels de polynômes peu friables dans l'espace $\mathbb{F}_q[X,Y]_{\leq r}$. Pour ce faire, on se concentre sur le problème géométrique équivalent de la recherche de systèmes linéaires (familles paramétrées par un espace projectif) de courbes planes ayant peu de composantes irréductibles.

On peut ensuite estimer les paramètres des codes construits par des méthodes géométriques. En particulier leur distance minimale est obtenue par des méthodes de comptage de points rationnels sur des courbes planes. Nous présenterons ainsi plusieurs constructions de très bons sous-codes des codes $PRM_q(3,2)$, $PRM_q(4,2)$ et $PRM_q(5,2)$, certains battant les paramètres des meilleurs codes connus jusque là.

Nous terminerons en présentant les extensions possibles de cette approche à un nombre supérieur de variables.

Une construction de codes LDPC quantiques basée sur des graphes de Cayley

Alain Couvreur, Nicolas Delfosse et Gilles Zémor
Institut de Mathématiques de Bordeaux, Université Bordeaux 1,
351, cours de la Libération, F-33405 Talence Cedex, France
Email : {Alain.Couvreur, Nicolas.Delfosse, Gilles.Zemor}@math.u-bordeaux1.fr

3 février 2011

Les codes LDPC permettent une correction des erreurs avec un décodage rapide. Cette idée nous encourage à étudier leur analogue quantique. Les constructions aléatoires qui marchent très bien dans le cas classique se généralisent difficilement à cause de certaines contraintes sur la matrice de parité notamment les relations d'orthogonalité. Depuis le début des années 2000 plusieurs constructions ont été proposées mais les distances minimales sont souvent petites ou inconnues. On s'intéresse à une construction proposée par D. Mackay, G. Mitchison et A. Shokrollahi dans l'article «More Sparse-Graph Codes for Quantum Error-Correction» en 2007. On obtient une borne inférieure sur la distance minimale.

L'idée est de construire la matrice de parité du code quantique $\mathbf{H} = \mathbf{H}_X = \mathbf{H}_Z$ comme la matrice d'adjacence d'un graphe de Cayley. D. Mackay, G. Mitchison et A. Shokrollahi ont donné une condition suffisante sur le groupe utilisé pour obtenir une matrice $\mathbf{H}$ qui a les bonnes propriétés et définit un code LDPC quantique. Ils ont calculé les paramètres obtenus pour de courtes longueurs. Ces exemples semblent donner de bons résultats pour les groupes $\mathbb{F}_2^r$ en faisant varier la partie génératrice.

Cette dernière construction permet d'associer un code LDPC quantique à un code classique. En effet, étant donné H une matrice de parité $r \times n$ d'un code binaire classique avec n pair, on considère le groupe $\mathbb{F}_2^r$ et l'ensemble de générateurs composé des colonnes de H. On en déduit une matrice d'adjacence dont les lignes sont de poids n et qui permet de construire un code quantique de longueur $N = 2^r$. En travaillant sur le graphe de Cayley nous avons démontré une borne inférieure sur la distance minimale d'un tel code quantique en fonction de celle du code classique. Nous avons aussi calculé explicitement la dimension et la distance minimale du code quantique associé à certains exemples. Avec le code de parité de longueur n, on obtient ainsi un code LDPC quantique de paramètres :

$$[[N = 2^{n-1}, K = 2^{\frac{n}{2}}, D = 2^{\frac{n}{2}-1}]],$$

lorsque n est un entier pair et $n \geq 4$; ceci était une question laissé ouverte par Shokrollahi *et al.*

Utilisation du groupe de permutations d'un code pour améliorer le décodage

Matthieu Legeay

IRMAR - Université de Rennes 1 - FRANCE
matthieu.legeay@univ-rennes1.fr

Dans la théorie des codes correcteurs d'erreurs linéaires et binaires, il n'est pas rare de trouver certains codes avec un groupe de permutations non trivial (codes de Reed-Solomon, codes de Reed-Muller, codes quasi-cycliques, codes BCH, etc). Généralement, ces codes ont un algorithme de décodage en temps polynomial.

Dans la cryptographie à clé publique basée sur les codes, également, on peut remarquer l'utilisation de codes dont le groupe de permutations est non trivial :

- ◇ Les codes de Goppa sont utilisés dans le cryptosystème original de McEliece. Lorsque ces codes sont construits à partir d'un polynôme générateur à coefficients dans un sous-corps du support, leurs groupes de permutations sont non triviaux, puisqu'ils possèdent au moins le Frobénius de l'extension.
- ◇ Plus récemment, les codes quasi-cycliques ont été utilisés dans le cryptosystème de McEliece. Ces codes ont également un groupe de permutations non trivial.
- ◇ Quelques fonctions de compression utilisées dans les fonctions de hachage utilisent aussi les codes quasi-cycliques.

Tous ces systèmes peuvent être attaqués par des algorithlmes de décodage génériques. Cependant, aucun de ces algorithmes ne prend en compte le fait que le groupe de permutations est non trivial, alors qu'il est possible de retrouver quelques informations sur ce groupe en utilisant le *support splitting algorithm*.

Une première approche pour améliorer les algorithmes de décodage génériques en utilisant le groupe de permutations a été introduite par MacWilliams en 1964. Elle y présentait un algorithme de décodage par ensembles d'information pour les codes cyliques, en utilisant la permutation cyclique et la multiplication par 2 modulo la longueur du code.

Dans cet exposé, nous présenterons une façon d'utiliser le groupe de permutation pour améliorer l'efficacité des algorithmes de décodage génériques. Dans le cas cyclique, nous avons une formule explicite et une évaluation théorique de la complexité de ces nouveaux algorithmes qui se révèle meilleure que celle connue de l'algorithme de décodage par ensembles d'information.

Identification aveugle des paramètres des codes convolutifs non-binaires

Y. Zrelli[1-2], R. Gautier[1-2], M. Marazin[1-2], E. Radoi[1-2] and E. Rannou[1-3]
[1]Université Européenne de Bretagne, France
[2]Université de Brest ; CNRS, UMR 3192 Lab-STICC,
[3]Université de Brest ; CNRS, UMR 6205 Laboratoire de Mathématiques,
6 avenue Victor Le Gorgeu, CS 93837, 29238 Brest cedex 3, France
yasamine.zrelli@univ-brest.fr
http://www.labsticc.fr/cacs/equipe/comint/

Depuis de nombreuses années, tout système de communication fiable se voit dans l'obligation d'intégrer dans sa chaîne de traitement un bloc de codage canal comprenant au moins un code correcteur d'erreurs. En effet, les codes correcteurs d'erreurs permettent de protéger les bits ou symboles informatifs par ajout de bits ou symboles redondants, ceci afin de pallier la présence d'erreurs de transmission, du côté récepteur, engendrées par le canal de transmission. La majorité des travaux de recherche et des implémentations pratiques dans des systèmes embarqués réels, se sont souvent restreints à des codes manipulant des données binaires, c'est-à-dire travaillant dans le corps de Galois GF(2). Pour les codes correcteurs non binaires, les implémentations et recherches associées se sont très longtemps limitées aux codes de Reed-Solomon pour des raisons de complexité, que ce soit au niveau du codage à l'émission, mais surtout au niveau du décodage à la réception. Cependant, depuis quelques années, des algorithmes de décodage à faible complexité pour les codes LDPC non binaires ont vu le jour [1]. Il en va de même pour les turbocodes non binaires, qui ont de bonnes propriétés comme codeurs externes pour le codage de symboles non binaires correspondant à des modulations numériques à grand nombre d'états, suscitant ainsi un grand intérêt [2]. Dans le cadre de cette étude, nous nous sommes intéressés au cas des codes convolutifs construits sur des corps de Galois non binaires de cardinal q (GF(q)) et tout particulièrement sur ceux correspondants aux extensions de corps GF(2), c'est-à-dire les corps GF(2^m). Le problème à résoudre est de savoir s'il serait possible d'identifier en aveugle les paramètres de ces codeurs comme cela est possible dans le cas des codes travaillant dans GF(2) [3].

Les codes convolutifs classiques les plus utilisés à l'heure actuelle travaillent dans GF(2) et traitent des informations binaires (i.e. un flux ou train binaire). Ils sont définis par trois paramètres $C(k, n, K)$ où k est la taille d'un mot d'information en entrée du codeur, n est la taille d'un mot de code en sortie du codeur et K la longueur de contrainte du codeur. Par contre, pour construire et surtout implémenter des codes convolutifs dans GF(q), il est nécessaire de prendre en compte les paramètres intrinsèques du corps dans lequel ils vont travailler. Dans le cas des extensions de corps GF(p^m) où p est premier et correspond à la caractéristique du corps, le cardinal du corps est p^m c'est-à-dire qu'il possède p^m éléments a_i et que tout élément non nul du corps peut-être construit à partir des puissances d'une racine α d'un polynôme irréductible primitif de degré m à coefficients dans GF(p), $a_i = \alpha^i, \forall i \in \{1, ..., p^m - 1\}$. Un tel polynôme irréductible dont les racines permettent d'engendrer tous les éléments du corps est dit primitif. D'un point de vue purement mathématique, les éléments du corps ainsi que les calculs effectués dans celui-ci ne dépendent que du cardinal du corps, puisque quel que soit le polynôme primitif choisi pour engendrer les éléments du corps, les calculs et donc les résultats sont équivalents à un isomorphisme de corps près. Par contre, au niveau implémentation sur une machine de traitement informatisé, les calculs et résultats sont directement liés à la représentation des éléments du corps et au choix du polynôme primitif permettant de les générer. Afin d'être en mesure d'identifier les paramètres du codeur utilisé, il est donc nécessaire au préalable ou en parallèle d'identifier le corps dans lequel un éventuel codeur travaille et également le polynôme primitif utilisé à l'émission lors de l'implémentation, et tout cela à partir d'un train binaire ou de symboles récupérés en réception. Une fois le corps utilisé à l'émission et son polynôme primitif identifiés, il est alors possible de mettre en œuvre un algorithme d'identification aveugle des paramètres du code utilisé.

Dans un contexte militaire, l'identification aveugle des paramètres de codage fait généralement partie intégrante d'une chaîne d'interception. Par contre dans un contexte civil, elle peut être considérée comme une brique logicielle ou matérielle d'un récepteur dans un contexte de Radio-Cognitive où ce récepteur devra estimer ces paramètres avec la seule connaissance des données reçues. Dans [3], nous avons développé un algorithme de reconnaissance aveugle de codes convolutifs binaires (GF(2)) s'appuyant sur un critère de calcul dans GF(2) de rang d'une matrice particulière constituée à partir de données reçues. Cette étude prenait en compte des propriétés des codes convolutifs dans GF(2), tout particulièrement celles des codes dits optimaux et leur construction. Pour le passage aux codes convolutifs dans GF(2^m), nous avons en premier lieu généralisé les travaux proposés par Ryan et Wilson dans [4] sur la recherche de codes optimaux de rendement $1/n$ dans GF(2^m) au cas des codes de rendement k/n. Puis, nous avons adapté la méthode d'identification des paramètres au cas GF(2) sous différentes hypothèses de bonne ou mauvaise identification des paramètres du corps, c'est-à-dire 2^m(ou tout simplement m) et du polynôme primitif réellement utilisé à l'émission. Tout cela dans le cas théorique d'une transmission non entachée d'erreurs afin d'évaluer la faisabilité d'une telle identification avant d'envisager dans un futur proche le cas où il y a des erreurs.

Les tests effectués pour $m \in \{1, 2, 3, 4\}$ ont montré que dans le cas d'une hypothèse de mauvaise identification du cardinal du corps ($2^{\tilde{m}} \neq 2^m$) ou celle d'une mauvaise identification du polynôme primitif $\tilde{p}(x) \neq p(x)$, alors la détection dans GF($2^{\tilde{m}}$) échoue (ce qui est normal) sauf évidemment pour le cas $\tilde{m} = 1$ (GF(2)) puisque tous ces corps sont des extensions de GF(2) (caractéristique 2). Dans ce cas comme dans [3], tous les paramètres sont identifiés à un facteur multiplicatif près qui est m (i.e. $\tilde{k} = m.k$, $\tilde{n} = m.n$ et $\tilde{K} = m.K$. En réalité ce sont les paramètres du codeur équivalent dans GF(2). Par contre, dans l'hypothèse où les paramètres du corps sont supposés être identifiés correctement (2^m et $p(x)$), alors le codeur est détecté et ses paramètres k, n et K dans GF(2^m) sont retrouvés.

Cette étude a donc démontré la pertinence de la généralisation de notre algorithme de détection et identification de paramètres de codeurs convolutifs au cas des codes non binaires dans des corps de caractéristique 2, sous l'hypothèse que ces mêmes corps soient connus ou correctement identifiés, ainsi que le polynôme primitif utilisé à l'émission. Cette méthode peut être généralisée pour des codes dans GF(p^m) de caractéristique p. Les travaux en cours ont pour but la reconnaissance complète du codeur en allant jusqu'à l'identification de sa matrice génératrice et également de s'attaquer au problème plus ardu d'un flux en réception de données codées entachées d'erreurs.

RÉFÉRENCES

[1] L. Barnault and D.Declercq, "Fast decoding algorithm for ldpc over gf(2^q)," in *IEEE Proceedings ITW2003*.

[2] J. Briffa and H. Schaathun, "Non-binary turbo codes and applications," in *5th International Symposium on Turbo Codes and Related Topics, 2008*.

[3] M. Marazin, R. Gautier, and G. Burel, "Dual code method for blind identification of convolutional encoder for cognitive radio receiver design," in *IEEE GLOBECOM Workshops, 2009*.

[4] W. E. Ryan and S. G. Wilson, "Convolutional codes over gf(q) with applications to frequency-hopping channels," *Military Communications Conference-Crisis Communications : The Promise and Reality*, 1987.

Accélération FPGA pour les codes et la cryptographie : Applications et limites

Jérémie Detrey, LORIA, Nancy

Circuits intégrés reconfigurables composés de dizaines (voire centaines) de milliers de cellules logiques très simples reliées entre elles par une interconnexion programmable, les FPGA sont aujourd'hui, de par leur flexibilité et leur parallélisme à grain fin, une cible de choix pour l'accélération de calculs spécifiques, tels que l'on peut en rencontrer en codage ou en cryptographie. Au cours de cet exposé, après une rapide présentation de l'architecture générale de ces FPGA, nous étudierons donc les possibilités d'accélération qu'ils offrent à travers plusieurs exemples, ainsi que leurs limites face aux CPU et GPU.

La multiplication multipartite

Pascal Giorgi Laurent Imbert Thomas Izard

La multiplication modulaire est au cœur de nombreux protocoles de cryptographie asymétrique, dans lesquels elle constitue l'opération de bas-niveau la plus coûteuse en temps de calcul. La taille des opérandes est contrainte par les normes de sécurité requises : pour des cryptosystèmes comme RSA ou ElGamal par exemple, elle doit être supérieure au millier de bits. Il est donc indispensable de pouvoir calculer $XY \mod N$ de façon efficace.

De nombreux algorithmes de multiplication modulaire ont été proposés, les plus connus étant les algorithmes de Montgomery et de Barrett qui opèrent la réduction modulaire séquentiellement, respectivement sur les bits de poids faibles et sur les bits de poids forts.
En 2008, Kaihara et Tagaki ont présenté la multiplication *bipartite* qui sépare la réduction modulaire en deux calculs indépendants, l'un agissant sur les bits de poids faibles, le second sur les bits de poids forts. Sur une architecture parallèle à deux unités de calcul, le coût global de la multiplication modulaire est ainsi réduit.

Nous proposons une version généralisée de cette multiplication bipartite, la multiplication *multipartite*, dans laquelle la réduction modulaire peut être exécutée de manière indépendante sur un nombre arbitraire d'unités de calcul.
Nous présenterons une analyse détaillée de sa complexité théorique, ainsi que les résultats expérimentaux obtenus sur une architecture multi-cœur à mémoire partagée.

Représentations redondantes des nombres pour recodage aléatoire pour ECC

Thomas Chabrier, Arnaud Tisserand, Danuta Pamula
IRISA–CAIRN, CNRS–INRIA–Université Rennes 1
6 rue de Kérampont, BP 80518, F-22305 Lannion
Contact : prénom.nom@irisa.fr

Proposition d'exposé court pour les journées « *Codage et Cryptographie* »2011

L'équipe CAIRN de l'IRISA travaille sur l'étude et la conception en matériel d'algorithmes et d'architectures pour des opérateurs arithmétiques de cryptographie sur les courbes elliptiques. En particulier, elle travaille sur des techniques permettant de protéger les cryptosystèmes ECC contre certaines attaques par canaux cachés. Ces attaques exploitent l'information éventuellement observable par le biais de canaux auxiliaires, comme la mesure de la consommation d'énergie ou du rayonnement électromagnétique d'un circuit intégré. L'attaquant cherche ainsi à récupérer de l'information, comme la clé secrète utilisée dans un système cryptographique. Dans le cas d'ECC, nous cherchons à protéger l'entier k utilisé lors de la multiplication scalaire $k[P]$ par un point de base P appartenant à une courbe elliptique donnée.

Par exemple, Coron [1] a introduit trois concepts comme contremesures contres des attaques par canaux cachés. Nous utilisons la première contremesure de Coron qui consiste à rendre aléatoire la séquence d'opérations de $[k]P$. Ceci est effectué en utilisant des représentations redondantes des nombres, qui serviront à rendre aléatoire la représentation du scalaire k. A chaque nouvelle exécution de l'algorithme de chiffrement, la clé secrète pourra avoir une représentation différente. Tout en garantissant le résultat du calcul de $[k]P$, nous pourrons ainsi avoir une séquence d'opérations différente lors d'une nouvelle exécution de l'algorithme. La signature en courant de celui-ci ne pourra être mise en corrélation avec une exécution antérieure. Cette méthode de représentation de la clé secrète peut donc constituer une contremesure face certaines attaques par canaux auxiliaires.

Nous avons étudié la représentation basée sur le système de base double (DBNS), et une représentation signée entière. Ces deux systèmes présentent l'avantage de posséder de nombreuses représentations possibles pour un même entier. La représentation DBNS consiste à représenter les entiers comme une somme de puissances combinées de deux nombres premiers deux et trois. L'entier x est représenté en DBNS par : $x = \sum_{i=0}^{n-1} s_i 2^{a_i} 3^{b_i}$, où $s_i \in \{-1, 1\}$ et $(a_i, b_i) \in \mathbb{N}^+$. L'avantage de ce système est que la représentation d'un nombre est creuse. La représentation en chiffres signés d'un entier (*borrow-save*) x est représentée par : $x = \sum_0^n k_i 2^i$, où $k_i \in \{-1, 0, 1\}$. De plus, il est possible de combiner ce système avec la forme non-adjacente (NAF) afin d'avoir un nombre de zéros autour de $\frac{2}{3}n$.

L'utilisation de telles représentations permettent le calcul de $[k]P$ indépendant et aléatoire vis à vis de la clé secrète et peut donc constituer une contremesure contre certaines attaques par canaux auxiliaires. De plus, les solutions choisies devront de plus allier vitesse et basse consommation d'énergie.

Références

[1] J.-S. Coron, "Resistance against differential power analysis for elliptic curve cryptosystems," in *Proc. Cryptographic Hardware and Embedded Systems (CHES)*, ser. LNCS, vol. 1717. Springer, Aug. 1999, pp. 292–302.

Accélérateur matériel compact pour le calcul du couplage de Tate sur des courbes supersingulières de 128 bits de sécurité

Nicolas Estibals

Les couplages interviennent dans des protocoles cryptographiques de plus en plus nombreux tels que le chiffrement basé sur l'identité ou la signature courte. Dès lors, fournir une implémentation efficace des couplages cryptographiques pour un grand nombre de supports, et plus spécialement les systèmes embarqués, devient un challenge intéressant.

Nous présentons une nouvelle méthode pour concevoir des accélérateurs matériels compacts pour le cacul du couplage de Tate sur des courbes supersingulières de petite caractéristique. Du fait de leur degré de plongement (embedding degree) limité, ces courbes ne sont généralement pas utilisées pour atteindre la sécurité standard de 128 bits. En effet, la taille du corps fini de définition d'une courbe à ce niveau de sécurité est très grande. Afin de pallier cet effet, nous considérons des courbes supersingulières définies sur des corps finis de degré d'extension modérément composé ($\mathbb{F}_{p^{nm}}$ avec n "petit" et m premier). Ces courbes deviennent alors vulnérables aux attaques basées sur la descente de Weil mais une analyse fine de celles-ci nous permet de montrer que leur impact reste limité et que nous pouvons maintenir la sécurité au-dessus de 128 bits.

Nous appliquons alors cette méthode à une courbe supersingulière définie sur $\mathbb{F}_{3^{5\cdot 97}}$ et décrivons ainsi une implémentation FPGA d'un accélérateur pour le calcul de couplage à 128 bits de sécurité. Sur un FPGA moyenne gamme (Xilinx Virtex-4), cet accélérateur cacule le couplage en 2.2ms sur une surface très modeste (4755 slices).