Décodage en liste des codes géométriques

Lancelot Pecquet

Directrice de thèse: Pascale Charpin
Co-Directeur de thèse: Daniel Augot
Projet CODES / *INRIA-Rocquencourt*

Soutenance de Thèse de Doctorat
Université Pierre et Marie Curie, Paris 6

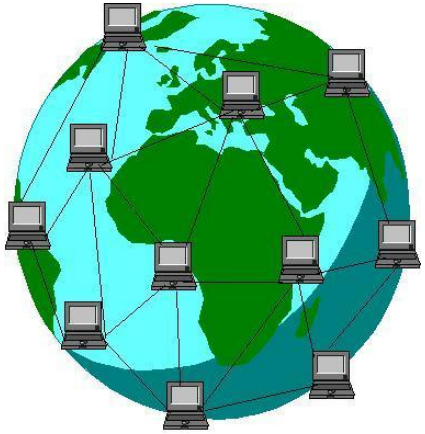
18 décembre 2001

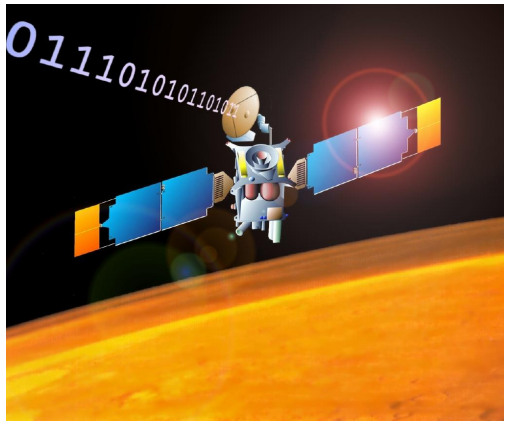
Plan

1. Introduction:
   (a) définitions et historique du décodage en liste;
   (b) contribution de cette thèse.
2. Principe du décodage souple métrique.
3. Décodage en liste des codes de Reed-Solomon:
   (a) principe fondamental et comparatif des méthodes existantes;
   (b) théorème général de décodage en liste et décodage souple algébrique ML;
   (c) algorithmique associée au décodage en liste.
4. Décodage en liste des codes géométriques: motivation et généralisation.
5. Conclusion: limites, autres codes, connexions et perspectives.

Introduction

Fiabilisation algorithmique des télécommunications





COMPACT disc DIGITAL AUDIO

DVD VIDEO

Codes en blocs et décodages

La distance de HAMMING entre deux mots x, y de de même longueur est le nombre de lettres $\mathrm{d}(x, y)$ dont ils diffèrent.

Sur l'alphabet $A = \{\mathtt{a}, \ldots, \mathtt{z}\}$, l'ensemble:

$$C = \{\mathtt{cassis}, \mathtt{goyave}, \mathtt{mangue}, \mathtt{banane}\}$$

est un code en blocs de longueur $n = 6$.
Sa distance minimale est $d = 3$.

$\mathrm{d}(\mathtt{cassis}, \mathtt{goyave})$	6
$\mathrm{d}(\mathtt{cassis}, \mathtt{mangue})$	5
$\mathrm{d}(\mathtt{cassis}, \mathtt{banane})$	5
$\mathrm{d}(\mathtt{goyave}, \mathtt{mangue})$	5
$\mathrm{d}(\mathtt{goyave}, \mathtt{banane})$	4
$\mathrm{d}(\mathtt{mangue}, \mathtt{banane})$	**3**

Décoder le mot $y = \mathtt{bangue}$, c'est trouver `mangue`: son plus proche voisin dans C.

$\mathrm{d}(\mathtt{bangue}, \mathtt{cassis})$	5
$\mathrm{d}(\mathtt{bangue}, \mathtt{goyave})$	5
$\mathrm{d}(\mathtt{bangue}, \mathtt{mangue})$	**1**
$\mathrm{d}(\mathtt{bangue}, \mathtt{banane})$	2

Décoder en liste $y = \mathtt{bangue}$ à distance $\tau = 2$, c'est trouver les deux mots `mangue`, `banane` qui sont à distance au plus τ de y.

Le codage en pratique

Un code servant à coder des mots binaires de 256 bits a un nombre d'éléments au moins égal à:

$$2^{256} \simeq 1.15792 \times 10^{77} \simeq \text{nombre d'atomes estimé de l'Univers}.$$

Stockage explicite du code impossible:
On prend un code linéaire.

Décodage par force brute impossible:
On exploite de la structure algébrique du code.

Historique du décodage en liste

1957 Elias, Wozencraft: concept de décodage en liste.

1996 Sudan: décodage en liste des codes de Reed-Solomon (RS) de faible taux de transmission (*i.e.* de grande distance relative).

1997 Shokrollahi-Wasserman: adaptation aux codes fortement géométriques à un point (SAG1) de faible taux.

1999 Guruswami-Sudan: amélioration pour tous les codes RS et SAG1 jusqu'au rayon de Johnson.

2000 Koetter-Vardy: décodage à entrée souple mais ne maximise pas la vraisemblance en général.

Contributions principales de cette thèse

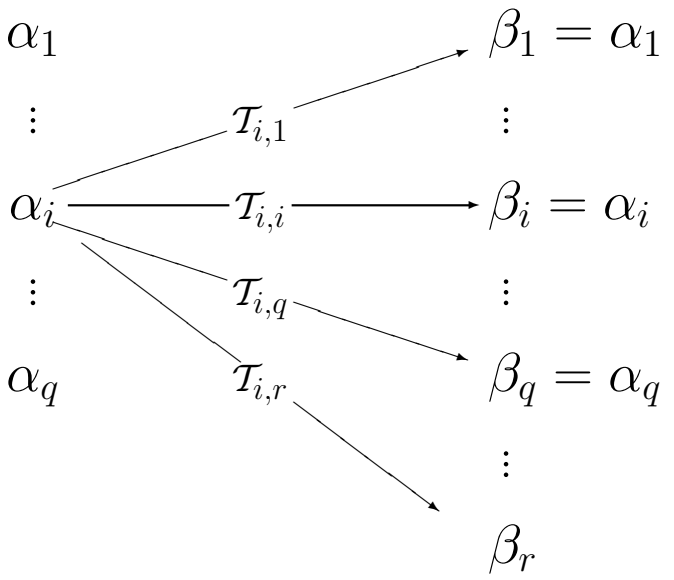
1. Décodage en liste de tous les codes géométriques pour la distance généralisée d_λ.
2. Choix de λ permettant de maximiser la vraisemblance dans un décodage souple algébrique, sur tout canal discret sans mémoire.
3. Étude globale des méthodes de géométrie algébrique effective pour les codes géométriques:
   (a) construction: algorithmique des courbes, désingularisation, genre, diviseurs, espaces de Riemann-Roch, sur la base des travaux de Gaétan Haché et en collaboration avec Paweł Wocjan;
   (b) décodage en liste: méthodes p-adiques, algorithmes de Newton-Hensel (collaboration avec Daniel Augot) et de Newton-Puiseux sur les corps de fonctions de courbes.
4. Implantation en Magma .

Principe du décodage souple métrique

Canaux discrets sans mémoire

Alphabets d'entrée et de sortie:

$$A = \{\alpha_1, \ldots, \alpha_q\} \subseteq B = \{\beta_1, \ldots, \beta_r\}$$



Probabilités conditionnelles:

$$\mathcal{T}_{i,j} = \mathsf{pr}[\beta_j \mid \alpha_i]$$

Pour tout $x \in A^n$ émis et tout $y \in B^n$ reçu:

$$\mathsf{pr}[y \mid x] = \prod_{j=1}^{n} \mathsf{pr}[y_j \mid x_j] \ .$$

Dans le canal q-aire symétrique de probabilité de transition p:

$$\log\left(\mathsf{pr}[y \mid x]\right) = \mathrm{d}(x,y) \cdot \Bigg(\underbrace{\log\left(\frac{p}{q-1}\right) - \log(1-p)}_{<0 \text{ ssi } p<\frac{q-1}{q}} \Bigg) + n\log(1-p) \ .$$

Distance généralisée et décodage en liste

$$\begin{array}{ccccc} \text{probabiliste} & \longrightarrow & \text{métrique} & \longrightarrow & \text{algébrique} \\ \mathsf{pr}[y \mid x] & \longrightarrow & \mathrm{d}_\lambda(x,y) & \longrightarrow & \mathrm{s}_\lambda(x,y) \end{array}$$

Pour tout $\lambda \in \mathbb{R}_+^n$, on définit λ-distance et la λ-similarité entre x et y dans A^n:

$$\mathrm{d}_\lambda(x,y) = \sum_{\substack{1 \le j \le n \\ x_j \neq y_j}} \lambda_j \quad \text{et} \quad \mathrm{s}_\lambda(x,y) = \sum_{\substack{1 \le j \le n \\ x_j = y_j}} \lambda_j \ .$$

Pour $\tau + s = \|\lambda\|_1$, on cherche $\mathrm{B}_\lambda(y,\tau) \cap C = \mathrm{S}_\lambda(y,s) \cap C$ où:

$$\mathrm{B}_\lambda(y,\tau) \stackrel{\text{def}}{=} \{x \in A^n \mid \mathrm{d}_\lambda(x,y) < \tau\} = \mathrm{S}_\lambda(y,s) \stackrel{\text{def}}{=} \{x \in A^n \mid \mathrm{s}_\lambda(x,y) > s\} \ .$$

Cas triviaux: $\tau = 0$ ou $s = 0$.

Similarité et vraisemblance

Pas de comparaison directe entre $x \in A^n$ et $y \in B^n$; on les déploie dans A^{qn}:

$$\begin{aligned}
\mathrm{span}(x) &= \Big(\boxed{x_1 \quad \cdots \quad x_1} \quad \cdots \quad \boxed{x_n \quad \cdots \quad x_n} \Big) \in A^{qn} \\
\mathsf{y} &= \Big(\boxed{\alpha_1 \quad \cdots \quad \alpha_q} \quad \cdots \quad \boxed{\alpha_1 \quad \cdots \quad \alpha_q} \Big) \in A^{qn} \\
\lambda(y) &= \Big(\boxed{\lambda(y)_{1,1} \quad \cdots \quad \lambda(y)_{1,q}} \quad \cdots \quad \boxed{\lambda(y)_{n,1} \quad \cdots \quad \lambda(y)_{n,q}} \Big) \in \mathbb{R}_+^{qn}
\end{aligned}$$

Nous choisissons $\lambda(y)_{i,j} = \log \mathsf{pr}[y_j \mid \alpha_i] - \xi_j$ où $\xi_j \stackrel{\text{def}}{=} \min\limits_{1 \le i \le q} \log \mathsf{pr}[y_j \mid \alpha_i]$:

$$\mathrm{s}_{\lambda(y)}\big(\mathrm{span}(x), \mathsf{y}\big) = \log \mathsf{pr}[y \mid x] - \log \min_{z \in A^n} \mathsf{pr}[y \mid z] \ .$$

Koetter et Vardy choisissent $\lambda(y)_{i,j} = \mathsf{pr}[y_j \mid \alpha_i]$:

$$\mathrm{s}_{\lambda(y)}\big(\mathrm{span}(x), \mathsf{y}\big) = \sum_{j=1}^{n} \mathsf{pr}[y_j \mid x_j] \ .$$

Décodage en liste des codes de Reed-Solomon

Codes de Reed-Solomon

Déf: Soient $p = (p_1, \ldots, p_n) \in \mathbb{F}_q^n$ et $k \in \mathbb{N}$, on a:

$$\begin{array}{ccc} \mathbb{F}_q[x] & \underset{\text{interpolation}}{\overset{\text{évaluation}}{\rightleftarrows}} & \mathbb{F}_q^n \\ \uparrow & & \uparrow \\ \langle 1, x, \ldots, x^{k-1} \rangle = L(k) & \longleftrightarrow & C = \mathrm{ev}_p\big(L(k)\big) \\ f_c & \longleftrightarrow & c = \big(f_c(p_1), \ldots, f_c(p_n)\big) \end{array}$$

1. C est linéaire;
2. C a une forte structure algébrique ($\mathbb{F}_q[x]$ est une algèbre factorielle);
3. si les p_j sont distincts alors $n \leq q$ et si $k < n$:

$$\dim(C) = k \quad \text{et} \quad \mathrm{d}(C) = n - k + 1$$

Principe du décodage en liste effectif

Pour tout $y \in \mathbb{F}_q^n$ et $m \in \mathbb{N}^*$, on construit un réel s_m, et un polynôme s_m-reconstructeur $G_m(T) \in \mathbb{F}_q[x][T]$ de degré au plus b_m tels que, pour tout $c \in C$:

$$\mathrm{s}_\lambda(c, y) > s_m \quad \Longrightarrow \quad G_m(f_c) = 0 \text{ dans } \mathbb{F}_q[x] \ .$$

Idée: on cherchera $G_m \in \mathbb{F}_q[x][T]$ tel que, pour tout $f_c \in L(k)$:

1. $G_m(f_c)$ a un degré au plus $m \cdot s_m$
2. $G_m(f_c)$ a au moins $m \cdot \mathrm{s}_\lambda(c, y)$ zéros

donc $\mathrm{s}_\lambda(c, y) > s_m \Rightarrow G_m(f_c) = 0$.

Cas triviaux exclus:

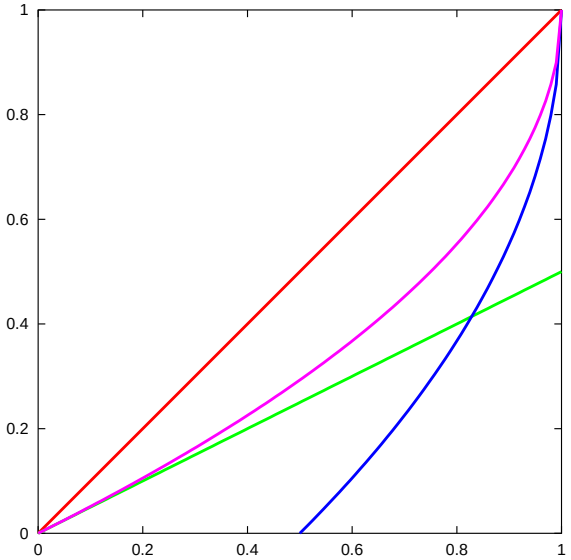
$$G_m(T) = 0 \quad \text{ou} \quad G_m(T) = \prod_{f_c \in L(k)} (T - f_c) \ .$$

NB: Il y a au plus b_m mots de C dans la boule $\mathrm{S}_\lambda(y, s_m)$.

Performance du décodage en liste

Jusqu'à quel rayon τ_{lim} sait-on décoder en liste un $[n, k, d]_q$-code de RS?
(distance de HAMMING)

Rayon τ_{lim}	Combinatoire	Effectif
$\frac{d}{2}$	empilement	BERLEKAMP $m = 1, \lambda_j = 1, b_1 = 1$
$n\left(1 - \sqrt{2\left(1 - \frac{d}{n}\right)}\right)$	?	SUDAN $m = 1, \lambda_j = 1$
$n\left(1 - \sqrt{1 - \frac{d}{n}}\right)$	JONHSON	GURUSWAMI-SUDAN $\lambda_j = 1$ (voire $\lambda_j \in \mathbb{N}$)
$n\frac{q-1}{q}\left(1 - \sqrt{1 - \frac{d}{n\frac{q-1}{q}}}\right)$	JOHNSON q-aire	KOETTER-VARDY $\lambda_j \in [0, 1]$
d	recouvrement	non



τ_{lim}/n en fonction de d/n.

Début
Intro.
d_λ
RS
AG
Concl.
Fin

Décodage en liste et décodage souple algébrique ML

Th [L.P. 2001]: Pour tout $\lambda \in \mathbb{R}_+^n$, on peut décoder en liste jusqu'à:

$$\tau_{\lim} = \|\lambda\|_1 - \|\lambda\|_2\sqrt{n-d} \ :$$

pour tout m, il existe s_m et un polynôme s_m-reconstructeur $G_m(T)$ tels que:

$$s_m = \underbrace{\|\lambda\|_2\sqrt{n-d}}_{s_{\lim}} + \mathsf{O}(1/\sqrt{m}) \quad \text{et} \quad \deg G_m \leq b_m = \frac{\|\lambda\|_2}{\sqrt{n-d}} \cdot m + \mathsf{O}(m^{3/4}) \ .$$

Le code déployé issu de C est un $[nq, k, dq]_q$-code RS; avec notre choix de $\lambda(y)$:
Th [L.P. 2001]: On peut calculer tous les mots $c \in C$ tels que:

$$\mathsf{pr}[y \mid c] > \min_{x \in A^n} \mathsf{pr}[y \mid x] \cdot \mathrm{e}^{\|\lambda(y)\|_2\sqrt{q(n-d)}} \ .$$

Algorithmique du décodage en liste

Soient y et $s > s_{\lim}$, on procède en trois phases pour calculer $\mathrm{S}_\lambda(y, s) \cap C$:

0. Calculer m assez grand pour que $s_m \leq s$, en déduire b_m.

1. Trouver un polynôme reconstructeur $G_m(T)$.

Méthode proposée par	Principe
Sudan, ... L.P.	Gauss
Olshevsky-Shorkollahi	structure de déplacement
Nielsen-Høholdt	interpolation itérative

2. Trouver les racines $f_c \in L(k)$ de $G_m(T)$ telles que $\mathrm{s}_\lambda(c, y) > s$.

Méthode proposée par	Principe
L.P.	Newton-Hensel, Newton-Puiseux
Gao-Shokrollahi	autre méthode à polygone de Newton.
Guruswami-Sudan	calcul dans une extension de $\mathbb{F}_q$.

Recherche des racines: Newton-Hensel

Th [Relèvement de Newton-Hensel]: Si $G_m(p_j, y_j) = 0$ et $G'_m(p_j, y_j) \neq 0$ (avec $(p_j, y_j) = (0, 0)$, sinon il suffit de translater) alors l'unique racine f de G_m dans $\mathbb{F}_q[\![x]\!]$ telle que $f(p_j) = y_j$ satisfait $f \equiv f_l \bmod x^{2^l}$ où:

$$f_0 = y_j \quad \text{et} \quad f_l = f_{l-1} - G_m(f_{l-1})/G'_m(f_{l-1}) \text{ rem } x^{2^l} \ , \ \text{pour } l \geq 1 \ .$$

Th [Augot, L.P. 1999]: Soient $\lambda = (1, \ldots, 1)$ et $G_1(T)$ de degré minimal. Pour toute racine $f \in L(k)$ de G_1, il existe j tel que f est relevable en p_j, *i.e.*:

$$f(p_j) = y_j \quad \text{et simultanément} \quad G'_m(p_j, y_j) \neq 0 \ .$$

Th [Augot, L.P. 1999]: L'algorithme de Sudan peut décoder en liste un code de longueur n en $O(n^2 \log n)$ opérations arithmétiques sur $\mathbb{F}_q$.

Pour λ et m quelconques $G'_m(p_j, y_j)$ peut être nul sur $\mathbb{F}_q^2$: Newton-Puiseux.

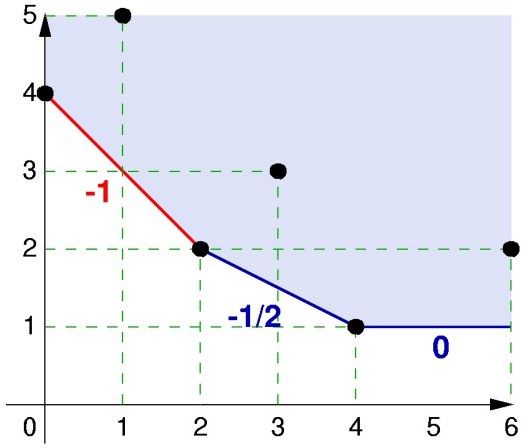
Recherche des racines: NEWTON-PUISEUX

Th [NEWTON-PUISEUX] Si $G_m(T) = x^{\varepsilon_0}(\alpha_0 + \cdots)T^0 + \cdots + x^{\varepsilon_{b_m}}(\alpha_{b_m} + \cdots)T^{b_m}$, soit $f = cx^e + \cdots \in \mathbb{F}_q[x]$ tel que $G_m(f) = 0$, alors $-e$ est une pente du polygone de NEWTON de G_m et c est une racine du polynôme de $\mathbb{F}_q[z]$:

$$\chi_e(z) \stackrel{\text{def}}{=} \frac{1}{z^{\min \Xi_e}} \sum_{i \in \Xi_e} \alpha_i z^i \quad \text{où} \quad \Xi_e = \{i \in \operatorname{Supp} G_m \mid \varepsilon_i + e_i = \sigma_e\}$$
$$\text{et} \quad \sigma_e = \min_{i \in \operatorname{Supp} G} (\varepsilon_i + ei) \ .$$

L'enveloppe convexe inférieure de $\{(0,4),(1,5),(2,2),(3,3),(4,1),(6,2)\}$ est le polygone de NEWTON de:

$$\begin{aligned} G_m(T) = {} & x^4 \cdot (2+x) \cdot T^0 \\ & + x^5 \cdot T^1 \\ & + x^2 \cdot (1+8x^5) \cdot T^2 \\ & + x^3 \cdot 7 \cdot T^3 \\ & + x^1 \cdot 3 \cdot T^4 \\ & + x^2 \cdot 2 \cdot T^6 \ . \end{aligned}$$



Décodage en liste des codes géométriques

Avantages et inconvénients des codes géométriques

Avantages principaux

1. Une structure algébrique analogue à celle des codes RS.
2. Des constructions de codes AG de longueur arbitraire sur $\mathbb{F}_q$ fixé aux performances pouvant parfois dépasser la borne de Gilbert-Varshamov.
3. Des méthodes algorithmiques issues de la géométrie algébrique effective et de la théorie algorithmique des nombres.
4. L'algorithme de décodage en liste peut être généralisé à ces codes.

Inconvénients principaux

1. Une sophistication *beaucoup* plus grande que celle des codes RS.
2. Une algorithmique associée *beaucoup* plus lourde, par conséquent.

Codes de GOPPA géométriques

Déf: Soit une courbe $X/\mathbb{F}_q$ projective, géométriquement irréductible, non-singulière de genre g, $p = (p_1, \ldots, p_n) \in X(\mathbb{F}_q)^n$, $D \in \mathrm{Div}(X)$ avec $p_j \notin D$:

$$\begin{array}{ccccc} \bigcap_{j=1}^{n} \mathscr{O}_{X,p_j} = \mathscr{O}_p & & \xrightarrow{\text{évaluation}} & & \mathbb{F}_q^n \\ & & \xleftarrow{\text{interpolation}} & & \\ \langle f_1, f_2, \ldots, f_k \rangle = \mathscr{L}(D) & & \longrightarrow & & C = \mathrm{ev}_p\big(\mathscr{L}(D)\big) \\ f_c & & \longrightarrow & & c = \big(f_c(p_1), \ldots, f_c(p_n)\big) \end{array}$$

1. C est linéaire;
2. C a une forte structure algébrique (BÉZOUT et RIEMANN-ROCH);
3. si les p_j sont distincts alors $n \leq \mathrm{N}_q(g)$ et si $\deg D < n$:

$$\dim(C) = k \geq k' = \deg D + 1 - g \quad \text{et} \quad \mathrm{d}(C) \geq d' = n - k + 1 - g \ .$$

avec $k = k'$ si $2g - 2 < \deg D$ (fortement géométrique: SAG).

Principe du décodage en liste des codes géométriques

Pour tout $y \in \mathbb{F}_q^n$ et $m \in \mathbb{N}^*$, on construit un réel s_m et un polynôme s_m-reconstructeur $G_m(T) \in \mathscr{O}_p[T]$ tels que, pour tout $c \in C$:

$$\mathrm{s}_\lambda(c, y) > s_m \quad \Longrightarrow \quad G_m(f_c) = 0 \; .$$

Idée: on construira un diviseur auxilliaire Δ_m de degré au moins $m \cdot s_m$ et on cherchera $G_m(T) \in \mathscr{L}(\Delta_m)[T]$ tel que, pour tout $f_c \in \mathscr{L}(D)$:

1. $\deg \big(G_m(f_c)\big)_\infty \leq m \cdot s_m$
2. $\deg \big(G_m(f_c)\big)_0 \geq m \cdot \mathrm{s}_\lambda(c, y)$

$$\text{donc} \quad \mathrm{s}_\lambda(c, y) > s_m \Rightarrow G_m(f_c) = 0 \; .$$

$$\uparrow$$

(Bézout)

On généralise avec le Théorème de Riemann-Roch:

$$s_m = \underbrace{\|\lambda\|_2 \sqrt{n - d'}}_{s_{\mathrm{lim}}} + \mathsf{O}(1/\sqrt{m}) \quad \text{et} \quad \deg G_m \leq b_m = \frac{\|\lambda\|_2}{\sqrt{n - d'}} \cdot m + \mathsf{O}(m^{3/4}) \; .$$

Remarques

Fonctionnement analogue du décodage souple.

$K \longrightarrow \widehat{K} \simeq \mathbb{F}_q(\!(x)\!)$: Newton-Hensel et Newton-Puiseux s'appliquent.

Précalcul de la phase « Riemann-Roch » $\longrightarrow$ implantation bas niveau.

Coût du décodage $\simeq$ coût de construction (*e.g.* codes de Shum *et al.*)

Conclusion

Au delà du rayon de JOHNSON?

Th [GOLDREICH, RUBINFELD, SUDAN 1998]: Soit $d < n\frac{q-1}{q}$, pour tout réel $\varepsilon \in]0,1[$, il existe un code C de longueur n et de distance minimale d sur $\mathbb{F}_q$, tel et $y \in \mathbb{F}_q^n$ tels que:

$$\tau = n\frac{q-1}{q}(1+\varepsilon)\left(1-\sqrt{1-\frac{d}{n\frac{q-1}{q}}+\varepsilon}\right) \quad \Longrightarrow \quad |\mathrm{B}(y,\tau) \cap C| = \mathrm{e}^{\Omega(\varepsilon^2 n)} \ .$$

Probleme ouvert: Peut-on choisir C AG? RS? linéaire?

HØHOLDT et JUSTESEN: le rayon de JOHNSON est optimal pour le décodage en liste de taille fixée des codes MDS.

Décodage en liste d'autres codes et connexions

Sous-codes dans les sous-corps (BCH, alternants, *etc.*)

Codes concaténés *via* le décodage souple.

Codes de Reed-Muller.

Codes CRT et factorisation des entiers.

Cryptanalyse par interpolation probabiliste de Jakobsen.

Interactions avec la théorie de la complexité.

Fin