École Normale Supérieure
Département de Mathématiques et Applications
Mémoire de Magistère MMFAI

Lacunarité et incohérence

Yohann de Castro

19 octobre 2007

Table des matières

1

Un modèle déterministe comme première approche

Au début du vingt et unième siècle, E. Candès, J. Romberg et T. Tao ont résolu un problème tomographique classique en imagerie médicale : *Reconstruire une image* $2D$ *à partir de la connaissance de sa transformée de Fourier sur un domaine étoilé* Ω. Dans ce problème on connaît un petit nombre de valeurs de la transformée de Fourier, et on cherche à reconstruire exactement le signal échantillonné. Dans le cas de *signaux lacunaires*, c'est-à-dire avec un grand nombre de coefficients nuls, E. Candès, J. Romberg et T. Tao se sont aperçus que l'on avait reconstruction exacte en considérant le signal qui minimise la norme ℓ_1 sous la contrainte imposée par l'échantillonnage. Cette approche est basée sur un algorithme bien établi, le *Basis Pursuit*, de minimisation de la norme ℓ_1 d'un signal sous une contrainte linéaire.

1.1 Un problème de reconstruction

On suppose que le signal que l'on cherche à reconstruire est lacunaire, c'est-à-dire qu'il a un grand nombre de coefficients nuls. Cette hypothèse est primordiale ; elle est la seule et unique hypothèse faite dans ce chapitre. Fixer cette hypothèse amène à considérer la classe des signaux ayant soit une représentation lacunaire dans une base donnée (base d'ondelettes par exemple), soit un gradient lacunaire dans une base donnée (signaux BV par exemple). Le théorème qui suit s'applique donc à une large classe de signaux.

Enoncé du problème

On s'intéresse au problème suivant :

> *Est-il possible de reconstruire un signal* x *connaissant seulement ses coefficients de Fourier sur un ensemble aléatoire* Ω *?*

On considère un signal x appartenant à $\mathbb{C}^N$ que l'on représente comme une fonction de l'ensemble des indices $\mathbb{N}_N := \{0, 1, \ldots, N-1\}$ à valeurs dans $\mathbb{C}$. On désigne par $T := \{t \in \mathbb{N}_N, x(t) \neq 0\}$ le support du signal x, et par S le cardinal de cet ensemble. On définit *la transformée de Fourier discrète* à la fréquence Ω de x par

$$\widehat{x}(\omega) = \frac{1}{\sqrt{N}} \sum_{t \in \mathbb{N}_N} x(t) e^{-i2\pi\omega t/N} .$$

On cherche à savoir s'il est possible de reconstruire le signal x connaissant K de ses coefficients de Fourier. Dans ce chapitre on note U la matrice de Fourier. Soit Ω un sous-ensemble aléatoire choisi de manière uniforme parmi les sous-ensembles de cardinal K. On note U_Ω la matrice de Fourier dont on n'a gardé que les colonnes d'indice appartenant à l'ensemble Ω.

$$U_\Omega\, x = \left(\widehat{x}(\omega)\right)_{\omega \in \Omega} .$$

Enoncé du théorème

Dans leur article [4], Emmanuel Candès, Justin Romberg et Terence Tao ont montré qu'un signal suffisamment lacunaire peut être reconstruit exactement, avec grande probabilité, comme *la* solution du problème (P_1) de minimisation convexe

$$\min_{g \in \mathbb{C}^N} \|g\|_{\ell_1} \quad \text{sachant que } U_\Omega\, x = U_\Omega\, g,$$

où la norme $\|g\|_{\ell_1}$ vaut $\sum_{t \in \mathbb{N}_N} |g(t)|$.

Théorème 1.1 (Théorème de reconstruction de Candès-Romberg-Tao). *Soit Ω un sous-ensemble aléatoire de taille K choisi de manière uniforme parmi les sous-ensembles de taille K. Soit un paramètre de fiabilité M donné. Alors tout signal $x \in \mathbb{R}^N$ dont on ne connaît qu'une borne sur la taille de son support :*

$$S \leq C_M \ (K/\log N), \tag{1}$$

*où C_M est une constante qui ne dépend que de M, peut être reconstruit **exactement** comme l'**unique** solution du problème (P_1) :*

$$\min_{g \in \mathbb{C}^N} \|g\|_{\ell_1} \quad \text{sachant que } U_\Omega x = U_\Omega g\,,$$

avec une probabilité d'au moins $1 - O(N^{-M})$.

La suite de ce chapitre apportera les principales étapes de la démonstration de ce théorème.

Remarque 1.2. — La constante implicite dans la notation $O(\,.\,)$ peut dépendre de M. Le paramètre M dépend de la taille N du signal et de la taille K de l'échantillon de fréquence.
— Ce théorème montre qu'il suffit d'un échantillonnage de la taille du signal modulo un facteur logarithmique pour avoir reconstruction exacte dans la majorité des cas.

1.2 Le vecteur dual π

Le problème (P_1) est un problème d'optimisation convexe, on sait qu'il existe une solution à ce problème mais il n'est pas clair que ce minimum soit unique, ni même qu'il permette de restituer le signal x. Une condition *nécessaire et suffisante* pour que le signal x que l'on cherche à reconstruire soit l'unique solution du problème (P_1) est qu'il existe un *vecteur dual* π qui vérifie trois conditions énoncées dans le lemme suivant.

Lemme 1.3. — *Soit Ω un sous-ensemble de $\mathbb{N}_N$ et x un signal de support T. On suppose qu'il existe un vecteur π de $\mathbb{C}^N$ tel que*

(i) La transformée de Fourier $U\pi$ du vecteur π ait pour support Ω.

(ii) Le vecteur π coïncide avec le vecteur[1] $\mathrm{sgn}(x)$ sur l'ensemble T.

(iii) L'amplitude de π est strictement inférieure à 1 en dehors de l'ensemble T.

Si de plus l'application $U_{\Omega T} : \ell_2(T) \mapsto \ell_2(\Omega)$ (définie comme la projection sur le sous-espace vectoriel des vecteurs de support Ω de l'application linéaire hermitienne U restreinte au sous-espace vectoriel des vecteurs de support T) est injective, le problème (P_1) admet alors un unique minimum égal à x. Réciproquement, si x est l'unique minimum du problème (P_1) alors il existe un vecteur π de $\mathbb{C}^N$ vérifiant les trois propriétés ci-dessus.

La preuve du théorème de reconstruction de Candès-Romberg-Tao (théorème 1.1) est articulée autour du vecteur dual π. Les auteurs prennent le vecteur de plus petite norme ℓ_2 vérifiant les deux premières hypothèses du lemme précédent et démontrent qu'avec une grande probabilité ce vecteur vérifie la troisième et dernière hypothèse. Mais quel est précisément ce vecteur ?

Un bon candidat

On choisit en espérant que son amplitude sera petite en dehors de l'ensemble T, le *vecteur d'énergie minimale* qui coïncide avec $\mathrm{sgn}(x)$ sur l'ensemble T et dont le support de sa transformée de Fourier est inclus dans l'ensemble Ω. Pour définir ce vecteur π, qui on l'espère sera un bon candidat pour répondre aux critères imposés par le lemme précédent, on introduit quelques notations. Soit E un ensemble inclus dans $\mathbb{N}_N$, notons $\iota_E : \ell_2(E) \mapsto \ell_2(\mathbb{N}_N)$, l'opérateur d'injection qui étend une application définie sur E en une application définie sur $\mathbb{N}_N$ (l'ensemble des applications définies sur $\mathbb{N}_N$ peut être canoniquement identifié à $\mathbb{C}^N$) en la complétant par des zéros en dehors de E. On note $U_{\Omega T} : \ell_2(T) \mapsto \ell_2(\Omega)$ la projection sur le sous-espace vectoriel des vecteurs de support Ω de l'application linéaire hermitienne U restreinte au sous-espace vectoriel des vecteurs de support T. On définit le vecteur π par :

$$\pi := U_\Omega^* U_{\Omega T} \left(U_{\Omega T}^* U_{\Omega T}\right)^{-1} \iota_T^* \mathrm{sgn}(x)\,. \tag{2}$$

[1] Le vecteur $\mathrm{sgn}(x)$ est le signal qui vaut $+1$ lorsque x est strictement positif, -1 lorsque x est strictement négatif et 0 sinon.

Le vecteur π vérifie-t-il les deux propriétés annoncées ?

On montre tout d'abord que π satisfait la condition (i) du lemme 1.3. On peut en effet remarquer[2] que

$$UU_\Omega^* = U(\iota_\Omega^* U)^* = UU^*\iota_\Omega = \iota_\Omega \,,$$

on en déduit que $U\pi$ a pour support l'ensemble Ω. On montre ensuite que π satisfait la condition (ii) du lemme 1.3. En effet,

$$\iota_T^* U_\Omega^* = \iota_T^* U^* \iota_\Omega = (U\iota_T)^*\iota_\Omega = (\iota_\Omega^* U \iota_T)^* = U_{\Omega T}^* \,,$$

on a bien $\iota_T^*\pi = \iota_T^* \operatorname{sgn}(x)$ comme attendu. La suite de la démonstration du théorème de reconstruction de Candès-Romberg-Tao (théorème 1.1) consiste à montrer que π est bien défini, c'est-à-dire que l'opérateur $U_{T\Omega}^* U_{T\Omega}$ est bien inversible, et qu'il est suffisamment petit en dehors de l'ensemble T.

1.3 Stabilité en norme ℓ_1

On ouvre dans cette section une parenthèse sur le lemme précédent et la place du *vecteur dual* au sein des travaux de Candès sur le *compressed sensing*. On élargit le cadre de notre exposé en considérant un signal x quelconque (on fait abstraction de l'hypothèse de support fini). On considère une matrice de mesure U_Ω quelconque et on note $x^\sharp$ la solution du problème (P_1) qui s'écrit ici :

$$\min_{g\in\mathbb{C}^N} \|g\|_{\ell_1} \quad \text{sachant que } U_\Omega\, x = U_\Omega\, g \,.$$

On présente un lemme de localisation des solutions $x^\sharp$ du problème (P_1).

Lemme 1.4. — *Soit x un vecteur de $\mathbb{C}^N$ et x_0 un vecteur de $\mathbb{R}^N$ de support un ensemble T. On pose h le vecteur de $\mathbb{C}^N$ tel que $x = x_0 + h$. On suppose qu'il existe un vecteur π appartenant à $\mathbb{C}^N$ tel que*

(i) *Le vecteur π appartient à l'image de U_Ω^*, c'est-à-dire qu'il existe un vecteur V pour lequel $\pi = U_\Omega^* V$,*

(ii) *Le vecteur π coïncide avec le vecteur $\operatorname{sgn}(x_0)$ sur l'ensemble T,*

(iii) *Le vecteur π est de module inférieur à $1/2$ sur le complémentaire de l'ensemble T.*

Alors on a l'inégalité

$$\left\|x^\sharp\right\|_{\ell_1(T^c)} \leq 3\,\|h\|_{\ell_1} \,,$$

où $\|\,.\,\|_{\ell_1(T^c)}$ est la norme induite par projection de la norme ℓ_1 sur l'espace vectoriel des signaux de support le complémentaire de l'ensemble T.

Remarque 1.5. — Le vecteur π sera appelé *vecteur dual associé au vecteur x_0*.

Démonstration. — Par définition de $x^\sharp$,

$$\left\|x^\sharp\right\|_{\ell_1} \leq \|x\|_{\ell_1} \leq \|x_0\|_{\ell_1} + \|h\|_{\ell_1} \,.$$

D'après l'hypothèse (i) le vecteur π appartient à l'image de U_Ω^*, il est donc orthogonal au noyau de U_Ω :

$$\langle x^\sharp, \pi\rangle = \langle x^\sharp, U_\Omega^* V\rangle = \langle U_\Omega x^\sharp, V\rangle = \langle U_\Omega x, V\rangle = \langle x, \pi\rangle \,.$$

On utilise les propriétés (ii) et (iii) pour en déduire que :

$$\begin{aligned}
\langle x, \pi\rangle &= \langle x_0 + h, \pi\rangle \\
&\geq \langle x_0, \pi\rangle - 1/2\,\|h\|_{\ell_1} \\
&= \|x_0\|_{\ell_1} - 1/2\,\|h\|_{\ell_1}
\end{aligned}$$

D'autre part,

$$\begin{aligned}
\left|\langle x^\sharp, \pi\rangle\right| &\leq \sum_T \left|x^\sharp(t)\pi(t)\right| + \sum_{T^c} \left|x^\sharp(t)\pi(t)\right| \\
&\leq \sum_T \left|x^\sharp(t)\pi(t)\right| + 1/2 \sum_{T^c} \left|x^\sharp(t)\right| \\
&= \left\|x^\sharp\right\|_{\ell_1} - 1/2\left\|x^\sharp\right\|_{\ell_1(T^c)}
\end{aligned}$$

[2]On utilise ici le fait que $U^{-1} = U^*$.

On en déduit que

$$\|x_0\|_{\ell_1} - 1/2\,\|h\|_{\ell_1} \leq \langle x, \pi\rangle \;= \langle x^\sharp, \pi\rangle \leq \left|\langle x^\sharp, \pi\rangle\right| \leq \left\|x^\sharp\right\|_{\ell_1} - 1/2\left\|x^\sharp\right\|_{\ell_1(T^c)} \leq \|x_0\|_{\ell_1} + \|h\|_{\ell_1} - 1/2\left\|x^\sharp\right\|_{\ell_1(T^c)}.$$

En simplifiant par $\|x_0\|_{\ell_1}$ on trouve le résultat attendu.

Remarque 1.6. — Ce lemme est un formidable outil pour localiser la solution $x^\sharp$ du problème (P_1) en fonction de n'importe quel élément de support fini x_0 appartenant à l'espace affine $x + \ker U_\Omega$ défini par l'information a priori, c'est-à-dire n'importe quel vecteur de support fini x_0 tel que $U_\Omega x_0 = U_\Omega x$.

1.4 Le modèle de Bernoulli

Il est difficile de travailler avec un sous-ensemble Ω choisi aléatoirement de manière uniforme, car la probabilité pour une fréquence donnée d'appartenir au sous-ensemble Ω est conditionnée par le fait que l'on sache si les autres fréquences y appartiennent ou pas. On introduit un modèle probabiliste plus simple, dit de *Bernoulli*, pour générer le sous-ensemble Ω ; et on montre dans cette section comment les résultats issus de ce modèle peuvent être translatés en résultats pour le modèle aléatoire uniforme. Le *modèle de Bernoulli* est un modèle simple où l'on considère que les fonctions indicatrices des fréquences mesurées sont des variables aléatoires *indépendantes identiquement distribuées* selon la *loi de Bernoulli* de paramètre $\tau \in [0, 1]$. On génère alors un sous-ensemble aléatoire Ω' de la manière suivante

$$\Omega' := \{\omega \in \mathbb{N}_N,\ \delta_\omega = 1\},$$

où $\delta_0, \ldots, \delta_{N-1}$ sont des variables aléatoires indépendantes identiquement distribuées de loi de Bernoulli de paramètre τ. La taille du sous-ensemble Ω' est une variable aléatoire de loi binomiale de paramètre τ et d'espérance

$$\mathbb{E}(|\Omega'|) = \tau N.$$

Remarque 1.7. — On remarque que $|\Omega'|/N \approx \tau$ avec une grande probabilité. En effet, on peut montrer à l'aide de l'*inégalité de Hoeffding* que pour tout réel ε strictement positif,

$$\mathbb{P}\left(\left|\frac{|\Omega'|}{N} - \tau\right| \geq \varepsilon\right) \leq 2\exp(-2\varepsilon^2).$$

On définit $\text{échec}(\Omega_0)$ comme l'événement où il est impossible de trouver un vecteur dual π dont la transformée de Fourier est supportée par Ω_0 et vérifiant les trois conditions du lemme 1.3. Soit Ω un sous-ensemble de taille K généré en utilisant le modèle aléatoire uniforme, et soit Ω' un sous-ensemble généré en utilisant le modèle de Bernoulli avec le paramètre $\tau = K/N$. On a alors

$$\begin{aligned}\mathbb{P}\left(\text{échec}\left(\Omega'\right)\right) &= \sum_{k=0}^{N} \mathbb{P}\left(\text{échec}\left(\Omega'\right) \mid |\Omega'| = k\right)\mathbb{P}\left(|\Omega'| = k\right)\\ &= \sum_{k=0}^{N} \mathbb{P}\left(\text{échec}\left(\Omega_k\right)\right)\mathbb{P}\left(|\Omega'| = k\right)\end{aligned}$$

où Ω_k est choisi aléatoirement de manière uniforme parmi les sous-ensembles de taille k. On fait alors deux remarques :

— La fonction $k \mapsto \mathbb{P}\left(\text{échec}\left(\Omega_k\right)\right)$ est une fonction décroissante de la variable k. Cela vient du fait que plus Ω est grand plus il est facile de construire un bon vecteur dual π.

— Lorsque τN est un entier, il correspond à la valeur médiane de $|\Omega'|$,

$$\mathbb{P}\left(|\Omega'| \leq \tau N - 1\right) < 1/2 < \mathbb{P}\left(|\Omega'| \leq \tau N\right).$$

Ces remarques permettent d'écrire que

$$\begin{aligned}\mathbb{P}\left(\text{échec}\left(\Omega'\right)\right) &\geq \sum_{k=0}^{K} \mathbb{P}\left(\text{échec}\left(\Omega_k\right)\right)\mathbb{P}\left(|\Omega'| = k\right)\\ &\geq \mathbb{P}\left(\text{échec}\left(\Omega\right)\right)\sum_{k=0}^{K}\mathbb{P}\left(|\Omega'| = k\right)\\ &\geq \frac{1}{2}\mathbb{P}\left(\text{échec}\left(\Omega\right)\right).\end{aligned}$$

On a ainsi montrer que si l'on pouvait borner la probabilité d'échec pour le modèle de Bernoulli, on sait alors que le taux d'échec pour le modèle uniforme ne sera pas deux fois plus grand.

2

L'incohérence, levier de la reconstruction

On va présenter et démontrer un résultat dû à Candès et Romberg qui généralise le travail précédent aux matrices N-hermitiennes. Soit U une matrice $N \times N$, on dira que U est N-hermitienne si et seulement si

$$U^*U = N\,\mathrm{Id}\,.$$

On fait le choix de cette définition pour simplifier les notations. La matrice U est une matrice de passage entre deux bases orthogonales. Soit T un sous-ensemble d'indices fixé de taille S et Ω un sous-ensemble aléatoire construit suivant le modèle de Bernoulli de paramètre $\tau = K/N$. On rappelle que $U_\Omega : \ell_2(\mathbb{N}_N) \mapsto \ell_2(\Omega)$ est la projection sur le sous-espace vectoriel des vecteurs de support Ω de l'application U ; que $U_T : \ell_2(T) \mapsto \ell_2(\mathbb{N}_N)$ désigne l'application U restreinte au sous-espace vectoriel des vecteurs de support T ; et que $U_{\Omega T} : \ell_2(T) \mapsto \ell_2(\Omega)$ est la projection sur le sous-espace vectoriel des vecteurs de support Ω de l'application U restreinte au sous espace vectoriel des vecteurs de support T. On note u^k la k-ième ligne de la matrice U_T et $\|\,.\,\|$ la norme standard d'opérateur au sens de la norme ℓ_2.

2.1 Principe d'incertitude et géométrie

Notre approche reste la même que précédemment, on cherche à reconstruire un signal x (ayant un grand nombre de coefficients nuls dans la base canonique) à partir d'une mesure $U_\Omega x$ faite dans une base orthogonale $\mathcal{B}$. Le vecteur Ux est la lecture du signal x dans la base $\mathcal{B}$ définie par les vecteurs colonne de la matrice U^*. La base $\mathcal{B}$ joue le rôle de la base de Fourier du chapitre précédent ; et intuitivement on voit que si elle présente de bonnes propriétés, comme ce fut le cas avec la matrice de Fourier, on pourra reconstruire le signal x dans de bonnes conditions. Mais quelle propriété doit vérifier la matrice U ?

Incohérence mutuelle

On appellera *incohérence mutuelle* d'une matrice N-hermitienne U, le réel $\mu(U)$ défini par :

$$\mu(U) := \max_{i,j \in \mathbb{N}_N} |U_{i,j}|\ .$$

La matrice U étant N-hermitienne, ses vecteurs colonne sont de norme ℓ_2 égale à $\sqrt{N}$. Par équivalence de la norme ℓ_2 et de la norme ℓ_∞, on en déduit que

$$1 \leq \mu(U) \leq \sqrt{N}\,.$$

L'incohérence μ est grossièrement une mesure de la localisation des vecteurs colonne de la matrice U dans la base canonique. Dans le cas de la matrice de Fourier, l'incohérence vaut 1 et les vecteurs de la base de Fourier sont non localisés dans le domaine temporel. C'est dans ce sens que doit être entendu le terme d'*incohérence* : plus μ est proche de 1, plus les bases associées (la base canonique que l'on compare à la base $\mathcal{B}$ associée à la matrice U) vérifient le principe d'incertitude, c'est à dire que tout signal fortement localisé dans la base canonique sera fortement non localisé dans la base $\mathcal{B}$.

Un problème au caractère géométrique

Le problème de reconstruction à l'aide du basis pursuit est intimement lié à la géométrie de la sphère unité de la norme ℓ_1. On rappelle que l'information a priori est donnée par le vecteur $y = U_\Omega\, x$, c'est à dire par un plan affine $x + \ker U_\Omega$ de plan directeur le noyau de l'application linéaire U_Ω. On peut dégager deux facteurs importants dans la réussite du programme (P_1). Tout d'abord, la matrice de mesure U_Ω qui détermine la direction du plan affine. Jusqu'ici tous nos modèles probabilistes portaient sur elle et uniquement elle. Celle-ci va de pair avec un deuxième élément, la *facette* de la sphère unité de la norme ℓ_1 sur laquelle repose le vecteur x. Une facette est un élément intime de la géométrie de la sphère unité ℓ_1. On appelle S-facette de la sphère unité ℓ_1, l'ensemble des barycentres à coefficients positifs de S points extrémaux non diamétralement opposés 2 à 2 de la sphère unité de la norme ℓ_1. Ces points extrémaux sont exactement les vecteurs avec un seul coefficient non nul qui vaut ± 1. Le deuxième élément important de la réussite du problème (P_1) est la S-facette sur laquelle repose le vecteur x. De manière condensée on représente cette S-facette par le vecteur $Z := \mathrm{sgn}(x)$. Toute l'information géométrique de la S-facette est contenue dans ce vecteur Z. Le problème (P_1) est entièrement déterminé par ces deux éléments géométriques. De manière intuitive, le problème de reconstruction consiste à savoir si le plan affine donné par U_Ω se présente bien par rapport à la S-facette déterminée par le vecteur x.

Un modèle probabiliste pour les S-facettes

On introduit un modèle aléatoire pour les S-facettes. On rappelle que, depuis le début, le support T du vecteur x est fixé. On introduit un modèle de Bernoulli pour construire une suite de signes ± 1 de support l'ensemble T. On construit un vecteur $Z = (Z_0, \ldots, Z_{N-1})$ de taille N et de support T de la manière suivante :

$$Z_t := \delta_T(t)(2\, b_{1/2}(t) - 1)\,,$$

où $b_{1/2}$ est la variable aléatoire canonique de loi de Bernoulli de paramètre $1/2$, qui vaut 1 avec probabilité $1/2$ et 0 avec probabilité $1/2$. Ainsi la variable aléatoire Z_t vaut 0 si t n'appartient pas à T ; elle vaut 1 avec probabilité $1/2$ et -1 avec probabilité $1/2$, si t appartient à l'ensemble T.

Enoncé du théorème de Candès-Romberg

On donne l'énoncé du théorème de Candès-Romberg que l'on démontre dans toute la suite de ce chapitre. Ce théorème peut être trouvé dans leur article [3].

Théorème 2.1 (Théorème de reconstruction de Candès-Romberg). *Soient N, K, S trois entiers naturels, θ un réel strictement compris entre 0 et 1, U une matrice N-hermitienne ($U^*U = N\mathrm{Id}$) telle que*

$$|U_{i,j}| \leq \mu(U)\,,$$

et T un sous-ensemble fixé de $\mathbb{N}_N$ de taille S. On se donne un sous-ensemble Ω de taille K et une suite de signes Z de taille N et de support T en appliquant le modèle de Bernoulli. On suppose que

$$K \geq C_0\, S\, \mu^2(U)\, \log(N/\theta)\,,$$

et

$$K \geq C_0'\, \log^2(N/\theta)\,,$$

où C_0 et C_0' sont deux constantes universelles strictement positives. Alors, avec une probabilité plus grande que $1 - \theta$, tout signal x de support T tel que $\mathrm{sgn}(x) = Z$, peut être reconstruit de manière exacte comme la solution du problème (P_1) sous la contrainte définie par la matrice de mesure U_Ω.

On a vu au chapitre précédent que le vecteur dual π que l'on prend tel que

$$\pi := U_\Omega^* U_{\Omega T}\, (U_{\Omega T}^* U_{\Omega T})^{-1}\, \iota_T^* \mathrm{sgn}(x)\,.$$

avait un rôle primordial dans la preuve du premier théorème de reconstruction. Mais, bien au delà de l'existence du vecteur dual π, le comportement de la matrice $U_{\Omega T}^* U_{\Omega T}$ est une pièce bien plus profonde dans la preuve. La démonstration du théorème de Candès-Romberg repose elle aussi sur l'aptitude, avec forte probabilité, qu'à la matrice $U_{\Omega T}^* U_{\Omega T}$ à être proche d'une isométrie. La preuve se déroule en deux temps. Tout d'abord on montre que l'espérance de $U_{\Omega T}^* U_{\Omega T}$ est proche de l'identité. Dans un deuxième et dernier temps on montre qu'avec forte probabilité la matrice $U_{\Omega T}^* U_{\Omega T}$ est proche de son espérance. On démontre la première étape avec une inégalité dû à Rudelson.

2.2 Une inégalité de Rudelson

Théorème 2.2 (Rudelson). *Dès lors que le membre de droite est plus petit que deux, on a*

$$\mathbb{E}\left\|\frac{1}{K}\,U_{\Omega T}^{*}U_{\Omega T}-\mathrm{Id}\right\|\leq C_R\,\frac{\sqrt{\log S}}{\sqrt{K}}\,\max_{1\leq k\leq N}\left\|u^k\right\|\,, \tag{3}$$

où C_R est une constante universelle strictement positive.

Démonstration. — On remarque que

$$U_{\Omega T}^{*}U_{\Omega T}=\sum_{k=0}^{N-1}\delta_k\,u^k\otimes u^k\,,$$

où (δ_k) est défini selon le modèle de Bernoulli (voir chapitre précédent), et la matrice $u^k\otimes u^k$ est définie par :

$$\left(u^k\otimes u^k\right)_{i,j}:=\overline{u}_{k,i}\,u_{k,j}\,.$$

On remarque que Ω est un sous-ensemble aléatoire généré par le modèle de Bernoulli, ainsi (δ_k) peut être vue comme une suite à N éléments de variables aléatoires indépendantes identiquement distribuées de loi de Bernoulli de paramètre τ. On s'interesse à l'espérance $\mathbb{E}\,\|Y\|$ où Y est la variable aléatoire définie par :

$$Y:=\frac{1}{K}\sum_{k=0}^{N-1}\delta_k\,u^k\otimes u^k-\mathrm{Id}\,.$$

On remarque que

$$\mathbb{E}\,Y=\frac{1}{K}\sum_{k=0}^{N-1}\tau\,u^k\otimes u^k-\mathrm{Id}=\frac{1}{N}\sum_{k=0}^{N-1}u^k\otimes u^k-\mathrm{Id}=0\,.$$

Par symétrie ("symmetrization trick", on peut avantageusement lire la page 5 de l'article [2]), on majore l'espérance de la norme de la variable aléatoire Y. Soit Y' une copie indépendante identiquement distribuée de Y :

$$Y':=\frac{1}{K}\sum_{k=0}^{N-1}\delta_k'\,u^k\otimes u^k-\mathrm{Id}\,,$$

où δ_k' sont des variables aléatoires indépendantes identiquement distribuées de loi de Bernoulli de paramètre τ. La fonction $M\mapsto\|Y-M\|$ est convexe sur $\mathbb{C}^{N\times N}$. D'après l'inégalité de Jensen,

$$\|Y-\mathbb{E}\,Y'\|\leq\mathbb{E}\,\|Y-Y'\|\,.$$

Comme l'espérance de la variable aléatoire Y est nulle, il vient que

$$\|Y\|\leq\mathbb{E}\,\|Y-Y'\|\,.$$

Le théorème de Fubini permet de conclure que :

$$\mathbb{E}\,\|Y\|\leq\mathbb{E}\,\|Y-Y'\|\,. \tag{4}$$

Soit $\varepsilon_0,\varepsilon_1,\ldots,\varepsilon_{N-1}$ une suite de variables aléatoires indépendantes identiquement distribuées de loi de Bernoulli de paramètre $1/2$ prenant les valeurs ±1. D'après l'inégalité 4 on a,

$$\mathbb{E}\,\|Y\|\leq\mathbb{E}_{\delta,\delta'}\left\|\frac{1}{K}\sum_{k=0}^{N-1}(\delta_k-\delta_k')\,u^k\otimes u^k\right\|\,.$$

Les variables aléatoires $(\delta_k-\delta_k')$ et $-(\delta_k-\delta_k')$ ont même loi, le théorème de Fubini donne

$$\mathbb{E}_{\delta,\delta'}\left\|\frac{1}{K}\sum_{k=0}^{N-1}(\delta_k-\delta_k')\,u^k\otimes u^k\right\|=\mathbb{E}_{\varepsilon}\,\mathbb{E}_{\delta,\delta'}\left\|\frac{1}{K}\sum_{k=0}^{N-1}\varepsilon_k(\delta_k-\delta_k')\,u^k\otimes u^k\right\|\,.$$

On obtient un premier résultat à l'aide de l'inégalité triangulaire :

$$\mathbb{E}\,\|Y\| \leq 2\,\mathbb{E}_\varepsilon\,\mathbb{E}_\delta\,\left\|\frac{1}{K}\sum_{k=0}^{N-1}\varepsilon_k\delta_k\,u^k\otimes u^k\right\|\,. \tag{5}$$

Un lemme dû à Rudelson dans son article [10] montre que :

$$\mathbb{E}_\varepsilon\,\left\|\sum_{k=0}^{N-1}\varepsilon_k\delta_k\,u^k\otimes u^k\right\| \leq \frac{C_R}{4}\sqrt{\log S}\,\max_{k|\delta_k=1}\left\|u^k\right\|\sqrt{\left\|\sum_{k=0}^{N-1}\delta_k\,u^k\otimes u^k\right\|}\,, \tag{6}$$

où C_R est une constante universelle strictement positive. Les inégalités 5 et 6, et l'inégalité de Jensen donnent :

$$\begin{aligned}\mathbb{E}\,\|Y\| &\leq \frac{C_R}{2}\frac{\sqrt{\log S}}{K}\max_{0\leq k\leq N-1}\left\|u^k\right\|\mathbb{E}\sqrt{\left\|\sum_{k=0}^{N-1}\delta_k\,u^k\otimes u^k\right\|}\\ &\leq \frac{C_R}{2}\frac{\sqrt{\log S}}{K}\max_{0\leq k\leq N-1}\left\|u^k\right\|\sqrt{\mathbb{E}\left\|\sum_{k=0}^{N-1}\delta_k\,u^k\otimes u^k\right\|}\end{aligned}$$

L'inégalité triangulaire montre que

$$\mathbb{E}\left\|\sum_{k=0}^{N-1}\delta_k\,u^k\otimes u^k\right\| = \mathbb{E}\,\|K\,Y + K\,\mathrm{Id}\| \leq K\,(\mathbb{E}\,\|Y\| + 1)\,.$$

On a l'inégalité

$$\mathbb{E}\,\|Y\| \leq a\,\sqrt{\mathbb{E}\,\|Y\| + 1}\,,\ \text{ où } a := \frac{C_R}{2}\frac{\sqrt{\log S}}{K}\max_{0\leq k\leq N-1}\left\|u^k\right\|\,.$$

En résolvant cette inéquation du second degré, on trouve que si le réel positif a est plus petit que 1 alors l'espérance de la norme de la variable aléatoire Y est plus petite que $2a$, comme attendu.

Remarque 2.3. — On a la majoration

$$\left\|u^k\right\|_{\ell_2} \leq \mu(U)\,\sqrt{S}\,.$$

L'inégalité 3 s'écrit alors de manière plus faible comme

$$\mathbb{E}\left\|\frac{1}{K}\,U_{\Omega T}^*U_{\Omega T} - \mathrm{Id}\right\| \leq C_R\,\frac{\sqrt{S\log S}}{\sqrt{K}}\,\mu(U)\,. \tag{7}$$

2.3 Un théorème de Talagrand

On démontre ensuite que la matrice $U_{\Omega T}^*U_{\Omega T}$ est proche de son espérance. On utilise un puissant théorème dû à Talagrand.

Théorème 2.4 (Talagrand). *Soit $\mathcal{F}$ une famille dénombrable de fonctions d'un espace de banach E à valeurs dans $\mathbb{R}$. Supposons qu'il existe un réel strictement positif B tel que pour toute fonction f appartenant à $\mathcal{F}$, la valeur absolue de f soit bornée par B. Soit n un entier naturel et $Y_1, Y_2, \ldots, Y_n$ une suite de variables aléatoires indépendantes identiquement distribuées à valeurs dans E. Soit Z la variable aléatoire définie par :*

$$Z := \sup_{f\in\mathcal{F}}\sum_{i=1}^{N} f(Y_i)\,.$$

On suppose que pour tout indice $i = 1, \ldots, n$ et tout f appartenant à $\mathcal{F}$,

$$\mathbb{E}\,f(Y_i) = 0\,.$$

Alors pour tout réel t positif ou nul,

$$\mathbb{P}\left(|Z - \mathbb{E}\,Z| > t\right) \leq 3\exp\left(-\frac{t}{KB}\log\left(1 + \frac{Bt}{\sigma^2 + B\,\mathbb{E}\,\overline{Z}}\right)\right)\,,$$

où

$$\sigma^2 := \sup_{f\in\mathcal{F}} \sum_{i=1}^{N} \mathbb{E}\, f^2(Y_i)\,,$$

$$\overline{Z} := \sup_{f\in\mathcal{F}} \left| \sum_{i=1}^{N} f(Y_i) \right|\,,$$

et K est une constante universelle strictement positive.

Ce théorème permet de mettre en place la seconde étape de la démonstration.

Théorème 2.5 (Candès). *Soit θ un réel strictement positif. On suppose que l'ensemble Ω est suffisamment grand au sens où*

$$K \geq S\mu(U)^2 \max\left(C_1 \log S,\ C_2 \log(3/\theta)\right)\,, \tag{8}$$

où C_1, C_2 sont deux constantes strictement positives. Alors,

$$\mathbb{P}\left(\left\| \frac{1}{K}\, U^*_{\Omega T} U_{\Omega T} - \mathrm{Id} \right\| \geq \frac{1}{2} \right) \leq \theta\,.$$

Démonstration. — On rappelle que

$$Y := \frac{1}{K} \sum_{k=0}^{N-1} \delta_k\, u^k \otimes u^k - \mathrm{Id}\,.$$

Ce que l'on écrit comme la somme de variables aléatoires Y_k :

$$Y = \sum_{k=0}^{N-1} \left(\delta_k - \frac{K}{N}\right) \frac{u^k \otimes u^k}{K} := \sum_{k=0}^{N-1} Y_k\,,$$

où

$$Y_k := \left(\delta_k - \frac{K}{N}\right) \frac{u^k \otimes u^k}{K}\,.$$

On s'intéresse à la norme spectrale de Y définie par :

$$\|Y\| := \sup_{f_1, f_2} \langle f_1, Y f_2\rangle = \sup_{f_1, f_2} \sum_{k=0}^{N-1} \langle f_1, Y_k f_2\rangle\,,$$

où le supremum est pris sur une famille dénombrable dense de la sphère unité pour la norme ℓ_2. On applique le théorème de Talagrand pour prouver le résultat (théorème 2.4). Tout d'abord, on remarque que

$$\mathbb{E}\, Y_k = 0\,.$$

Soit un couple $f = (f_1, f_2)$ de vecteurs de la sphère unité ℓ_2, on note abusivement $f : \mathbb{C}^{N\times N} \to \mathbb{R}$, l'application qui à une matrice M associe $\langle f_1, M f_2\rangle$. Soit $k = 0, \ldots, N-1$,

$$|f(Y_k)| \leq \frac{1}{K} \left|\langle f_1, u^k\rangle \langle f_2, u^k\rangle\right| \leq \frac{\left\|u^k\right\|^2}{K}\,.$$

On pose

$$B := \max_{0\leq k\leq N-1} \frac{\left\|u^k\right\|^2}{K}\,.$$

On majore l'espérance de $f^2(Y_k)$,

$$\begin{aligned}
\mathbb{E}\, f^2(Y_k) &= \mathbb{P}(\delta_k = 1)\left(1 - \frac{K}{N}\right)^2 \frac{\left\langle f_1,\, u^k \otimes u^k f_2\right\rangle^2}{K^2} \ +\ \mathbb{P}(\delta_k = 0)\left(\frac{K}{N}\right)^2 \frac{\left\langle f_1,\, u^k \otimes u^k f_2\right\rangle^2}{K^2} \\
&= \frac{K}{N}\left(1 - \frac{K}{N}\right) \frac{\left|\left\langle f_1,\, u^k\right\rangle \left\langle f_2,\, u^k\right\rangle\right|^2}{K^2} \\
&\leq \frac{K}{N}\left(1 - \frac{K}{N}\right) \frac{\left\|u^k\right\|^2}{K^2} \left|\left\langle f_2,\, u^k\right\rangle\right|^2
\end{aligned}$$

On remarque d'abord que

$$\sum_{k=0}^{N-1} \left|\langle u^k,\, f_2\rangle\right|^2 = N \left\|f_2\right\|^2 = N\,,$$

on a alors la majoration attendue,

$$\sum_{k=0}^{N-1} \mathbb{E}\, f^2(Y_k) \leq \left(1-\frac{K}{N}\right)\frac{1}{K}\max_{0\leq k\leq N-1}\left\|u^k\right\|^2 \leq B\,.$$

On applique le théorème de Talagrand (théorème 2.4). En remarquant que $Z = \|Y\| = \overline{Z}$, on obtient pour tout réel t strictement positif :

$$\mathbb{P}\left(\left|\,\|Y\| - \mathbb{E}\,\|Y\|\,\right| > t\right) \leq 3\exp\left(-\frac{t}{KB}\log\left(1+\frac{t}{1+\mathbb{E}\,\|Y\|}\right)\right)\,, \tag{9}$$

avec $Y = \frac{1}{K}U^*_{\Omega T}U_{\Omega T} - \mathrm{Id}$. Le théorème 2.2 démontre l'inégalité 7 :

$$\mathbb{E}\,\|Y\| = \mathbb{E}\left\|\frac{1}{K}\,U^*_{\Omega T}U_{\Omega T} - \mathrm{Id}\right\| \leq C_R\,\frac{\sqrt{S\log S}}{\sqrt{K}}\,\mu(U)\,.$$

En prenant K suffisamment grand,

$$K \geq 4\,C_R^2\,\mu^2(U)\,S\log S\,,$$

on a

$$\mathbb{E}\,\|Y\| \leq 1/4\,.$$

Il suffit alors de prendre $t = 1/4$ dans l'inégalité 9 et de remarquer que $B \leq \mu^2(U)\,S/K$ pour obtenir :

$$\mathbb{P}\left(\|Y\| > 1/2\right) \leq 3\exp\left(-\frac{K}{C_T\,\mu^2(U)\,S}\right)\,,$$

où $C_T = 4K/\log(6/5)$. La preuve se termine en posant $C_1 = 16C_R$ et $C_2 = C_T$.

2.4 Preuve du théorème de Candès-Romberg

Tous les éléments sont en place pour achever la preuve du théorème.

Démonstration. — On se place dans le cadre des hypothèses du théorème, en particulier on se donne un vecteur de Bernoulli Z représentant une S-facette. Soit π le vecteur dual associé au vecteur Z. On renvoie à la section 1.3 pour la définition du vecteur dual. On rappelle que[1]

$$\pi := U^*_{\Omega}U_{\Omega T}\left(U^*_{\Omega T}U_{\Omega T}\right)^{-1} Z\,.$$

Soit t_0 un élément du complémentaire du sous-ensemble T, la matrice $\left(U^*_{\Omega T}U_{\Omega T}\right)^{-1}$ étant N-hermitienne,

$$\pi(t_0) = \left\langle v^0,\, \left(U^*_{\Omega T}U_{\Omega T}\right)^{-1} Z\right\rangle = \left\langle w^0,\, Z\right\rangle\,,$$

où $w^0 := \left(U^*_{\Omega T}U_{\Omega T}\right)^{-1} v^0$ et v^0 est le t_0-ième vecteur de ligne de $U^*_{\Omega}U_{\Omega T}$.

Lemme 2.6. — *Le moment d'ordre 2 de $Z_0 := \left\|v^0\right\|$ vérifie*

$$\mathbb{E}\,Z_0^2 \leq \mu^2(U)KS\,. \tag{10}$$

Démonstration. — On pose $\lambda_k^0 = \overline{U}_{t_0,k}$. On rappelle que v^0 est le t_0-ième vecteur de ligne de $U^*_{\Omega}U_{\Omega T}$, le calcul du produit matriciel donne

$$v^0 = \sum_{k=0}^{N-1} \delta_k \lambda_k^0\, u^k\,.$$

[1] On renvoie à l'équation 2 pour plus de détails.

Soit t un indice appartenant à T. On désigne par $u^k(t)$ l'entrée d'indice t du vecteur ligne u^k, soit encore le coefficient $U_{k,t}$. On note u_t le t-ème vecteur colonne de la matrice U. On a

$$\sum_{k=0}^{N-1} \lambda_k^0 u^k(t) = \sum_{k=0}^{N-1} \overline{U}_{t_0,k} U_{k,t} = \langle u_{t_0},\, u_t \rangle = 0\,,$$

par orthogonalité des vecteurs u_{t_0} et u_t. On a

$$v^0 = \sum_{k=0}^{N-1} (\delta_k - \mathbb{E}\,\delta_k)\, \lambda_k^0\, u^k\,. \tag{11}$$

On écrit v^0 comme la somme de variables aléatoires d'espérance nulle,

$$v^0 = \sum_{k=0}^{N-1} Y_k\,,$$

avec

$$Y_k := (\delta_k - \mathbb{E}\,\delta_k)\, \lambda_k^0\, u^k\,.$$

On rappelle que la variable aléatoire Y_k est d'espérance nulle et on calcule le moment d'ordre 2 de Z_0

$$\begin{aligned}\mathbb{E}\, Z_0^2 &= \sum_k \mathbb{E}\, \langle Y_k,\, Y_k \rangle + \sum_{k\neq k'} \mathbb{E}\, \langle Y_k,\, Y_{k'} \rangle\,,\\ &= \sum_k \mathbb{E}\, \langle Y_k,\, Y_k \rangle\,.\end{aligned}$$

On majore le moment d'ordre 2 de $\|Y_k\|$,

$$\begin{aligned}\mathbb{E}\, \|Y_k\|^2 &= \frac{K}{N}\left(1 - \frac{K}{N}\right) \left|\lambda_k^0\right|^2 \left\|u^k\right\|^2\,,\\ &\leq \frac{K}{N}\left(1 - \frac{K}{N}\right) \left|\lambda_k^0\right|^2 \mu^2(U) S\end{aligned}$$

On remarque que

$$\sum_{k=0}^{N-1} \left|\lambda_k^0\right|^2 = N\,,$$

ce qui permet de conclure la preuve du lemme.

Lemme 2.7. — *Soit t_0 un élément du complémentaire de T. On pose*

$$\sigma^2 := \mu^2(U) K \max\left(1,\, \mu(U) S \sqrt{K}\right)\,.$$

Soit un réel a strictement positif tel que

$$\begin{cases} a \leq \left(K/\mu^2(U)\right)^{1/4} & \text{si } \mu(U)\, S/\sqrt{K} > 1\\ a \leq \left(K/\mu^2(U) S\right)^{1/2} & \text{sinon}\end{cases}$$

Alors,

$$\mathbb{P}\left(Z_0 \geq \mu(U)\sqrt{KS} + a\sigma^2\right) \leq \exp\left(-\gamma a^2\right)\,,$$

où γ est une constante universelle strictement positive.

Démonstration. — Par définition,

$$Z_0 := \sup_{\|f\|=1} \left\langle v^0,\, f \right\rangle = \sup_{\|f\|=1} \sum_{k=0}^{N-1} \langle Y_k,\, f \rangle\,,$$

avec (voir la preuve précédente)

$$Y_k := (\delta_k - \mathbb{E}\,\delta_k)\, \lambda_k^0\, u^k\,.$$

On prouve que l'on est bien dans le cadre des hypothèses du théorème 2.4. Soit f un vecteur unité fixé appartenant à un sous ensemble dense fixé de la sphère unité de la norme ℓ_2. On note $f(Y_k)$ l'application $\langle Y_k,\, f\rangle$. La variable aléatoire Y_k est d'espérance nulle, et donc, par linéarité,

$$\mathbb{E}\, f(Y_k) = 0\,.$$

On majore la valeur absolue de $f(Y_k)$,

$$|f(Y_k)| \le \left|\lambda_k^0\right| \left|\left\langle u^k,\, f\right\rangle\right| \le \left|\lambda_k^0\right| \left\|u^k\right\| \le \mu^2(U)\, S^{1/2} = B\,,$$

où on a posé

$$B := \mu^2(U)\, S^{1/2}\,.$$

On majore l'espérance de Z_0 à l'aide du lemme 2.6 sur le moment d'ordre 2 de Z_0 et de l'inégalité de Jensen,

$$\mathbb{E}\, Z_0 \le \sqrt{\mathbb{E}\, Z_0^2} \le \mu(U)\sqrt{KS}\,.$$

On majore ensuite σ^2, on a

$$\mathbb{E}\, f^2(Y_k) = \frac{K}{N}\left(1-\frac{K}{N}\right)\left|\lambda_k^0\right|^2 \left|\left\langle u^k,\, f\right\rangle\right|^2 \le \frac{K}{N}\left(1-\frac{K}{N}\right)\mu^2(U)\left|\left\langle u^k,\, f\right\rangle\right|^2\,.$$

Comme $\sum_{1\le k\le N} \left|\left\langle u^k,\, f\right\rangle\right|^2 = N$, on a

$$\sum_{k=0}^{N-1} \mathbb{E}\, f^2(Y_k) \le K\left(1-\frac{K}{N}\right)\mu^2(U)\,.$$

On applique le théorème 2.4, pour tout réel t positif ou nul,

$$\mathbb{P}\left(|Z_0 - \mathbb{E}\, Z_0| > t\right) \le 3\exp\left(-\frac{t}{KB}\,\log\left(1+\frac{Bt}{\mu^2(U)K + B\mu(U)\sqrt{KS}}\right)\right)\,.$$

On suppose que $\overline{\sigma}^2 = B\mu(U)\sqrt{KS} \ge \mu^2(U)K$ et on pose $t = a\overline{\sigma}$, alors l'inégalité ci-dessus montre que

$$\mathbb{P}\left(|Z_0 - \mathbb{E}\, Z_0| > t\right) \le 3\exp\left(-\gamma a^2\right)\,,$$

dès lors que $Bt \le \overline{\sigma}^2$. On montre le même résultat lorsque $\overline{\sigma}^2 = \mu^2(U)K \ge B\mu(U)\sqrt{KS}$ et $Bt \le \mu^2(U)K$, ce qui prouve le lemme.

Lemme 2.8. — *On rappelle que $w^0 := \left(U_{\Omega T}^* U_{\Omega T}\right)^{-1} v^0$. Avec les même notations que précédemment,*

$$\mathbb{P}\left(\sup_{t_0\in T^c}\left\|w^0\right\| \ge 2\mu(U)\sqrt{S/K} + 2a\overline{\sigma}^2/K\right) \le N\exp\left(-\gamma a^2\right) + \mathbb{P}\left(\left\|U_{\Omega T}^* U_{\Omega T}\right\| \le 1/2\right)\,. \tag{12}$$

Démonstration. — Soient A et B les événements $\{\|U_{\Omega T}^* U_{\Omega T}\| \ge K/2\}$ et $\left\{\sup_{t_0\in T^c}\left\|v^0\right\| \le \mu(U)\sqrt{KS} + a\overline{\sigma}\right\}$. Le lemme 2.7 montre que $\mathbb{P}(B^c) \le \exp\left(-\gamma a^2\right)$. Sur $A\cap B$, on a

$$\sup_{t_0\in T^c}\left\|w^0\right\| \le 2\mu(U)\sqrt{S/K} + 2a\overline{\sigma}^2/K\,.$$

Ce qui démontre le lemme.

Lemme 2.9. — *Soit λ un réel strictement positif,*

$$\mathbb{P}\left(\sup_{t\in T^c}|\pi(t)| > 1\right) \le 2N\exp\left(-\frac{1}{2\lambda^2}\right) + \mathbb{P}\left(\sup_{t_0\in T^c}\left\|w^0\right\| > \lambda\right)\,. \tag{13}$$

Démonstration. — On rappelle que Z est la S-facette associée à x. La preuve de ce lemme repose sur l'inégalité de *Hoeffding*, on pourra avantageusement lire l'article [1]. On écrit :

$$\left\langle w^0,\, Z\right\rangle = \sum_{k=0}^{N-1} w^0(k)\overline{\delta}_k = \sum_{k=0}^{N-1} X_k\,,$$

où $X_k = w^0(k)\overline{\delta}_k$. Les variables aléatoires δ_k sont indépendantes et les variables aléatoires X_k sont encadrées par 0 et $w^0(k)$, l'inégalité de Hoeffding donne alors :

$$\mathbb{P}\left(\left|\langle w^0,\, Z\rangle\right| > 1 \;\middle|\; w^0\right) \leq 2\exp\left(-2/\left\|w^0\right\|^2\right) .$$

On rappelle que $\pi(t_0) = \langle w^0,\, Z\rangle$, on en déduit que

$$\mathbb{P}\left(\sup_{t\in T^c} |\pi(t)| > 1 \;\middle|\; \sup_{t_0\in T^c} \left\|w^0\right\| \leq \lambda\right) \leq 2N\exp\left(-2/\lambda^2\right) ,$$

ce qui conclut la preuve du lemme.

On pose

$$\lambda = 2\mu(U)\sqrt{S/K} + 2a\overline{\sigma}/K .$$

A l'aide des lemmes 2.8 et 2.9, on sait que pour tout réel a strictement positif obéissant aux conditions imposées par le lemme 2.7,

$$\mathbb{P}\left(\sup_{t\in T^c} |\pi(t)| > 1\right) \leq 2N\exp\left(-\frac{1}{2\lambda^2}\right) + N\exp\left(-\gamma a^2\right) + \mathbb{P}\left(\|U_{\Omega T}^* U_{\Omega T}\| \leq 1/2\right) . \tag{14}$$

Pour rendre le second terme plus petit que δ, on choisit le réel a tel que

$$a^2 := \gamma^{-1}\log(N/\delta) .$$

Le premier terme est plus petit que δ si

$$\frac{1}{\lambda^2} \geq 2\log(2N/\delta) . \tag{15}$$

On suppose que $\mu(U)S \geq \sqrt{K}$. L'hypothèse du lemme 2.7 est $a \leq \left(K/\mu^2(U)\right)^{1/4}$, ou de manière équivalente

$$K \geq \mu^2(U)\frac{\log(N/\delta)^2}{\gamma^2} .$$

Dans ces conditions, $a\overline{\sigma} \leq \mu(U)\sqrt{KS}$, ou encore

$$\frac{1}{\lambda^2} \geq \frac{1}{16}\frac{K}{\mu^2(U)S} . \tag{16}$$

On suppose que $\mu(U)S \leq \sqrt{K}$. Si de plus $S \geq a^2$, l'inéquation 16 reste vraie. D'autre part, si $S \leq a$ alors $\lambda \leq 4a\overline{\sigma}/K$ et

$$\frac{1}{\lambda^2} \geq \frac{1}{16}\frac{K}{a^2(U)S} .$$

On veut que l'inéquation 15 est vraie. De l'analyse précédente, on voit qu'il suffit que le nombre de mesures K vérifie :

$$\frac{K}{16\mu^2(U)}\min\left(\frac{1}{S},\, \frac{1}{a^2}\right) \geq 2\log\left(\frac{2N}{K}\right) .$$

Le second terme du second membre de l'inéquation 14 est plus petit que δ dès lors que

$$K \geq K_1\mu^2(U)\max\left(S,\, \log(N/\delta)\right)\log(N/\delta) ,$$

où K_1 est une constante strictement positive. Le théorème 2.5 permet de majorer par δ le troisième et dernier terme du second membre de l'inéquation 14 dès lors que

$$K \geq K_2\mu^2(U)S\log(N/\delta) ,$$

où K_2 est une constante strictement positive. Ce qui prouve le théorème.

Références

1. O. Bousqet S. Boucheron and G. Lugosi. *Advanced Lectures in Machine Learning*, pages 208–240. Springer, 2004.
2. O. Bousqet S. Boucheron and G. Lugosi. Theory of classification : A survey of some recent advances. 23 Septembre 2005.
3. E.J Candès and J. Romberg. Sparsity and incoherence in compressive sampling. Novembre 2006.
4. E.J Candès, J. Romberg, and T. Tao. Robust uncertainty principles : Exact signal reconstruction from highly incomplete frequency information. Juin 2004.
5. E.J Candès and T. Tao. Near optimal signal recovery from random projections : Universal encoding strategies ? Octobre 2004.
6. David L. Donoho. Compressed sensing. Septembre 2004.
7. D.L Donoho and P.B Stark. Uncertainty principles and signal recovery. *SIAM J. Applied Math*, 49(3) :906–931, Juin 1989.
8. Yves Meyer. *Ondelettes et opérateurs I, Ondelettes.* Hermann, 1990.
9. Yves Meyer. Oscillating patterns in image processing and some nonlinear evolution equations. *The fifteenth Dean Jaqueline B. Lewis, Memorial Lectures.*, 2002.
10. Mark Rudelson. Random vectors in the isotropic position. *Journal of Functional Analysis*, 1999.
11. T. Tao. An uncertainty principle for cyclic groups of prime order. *Mathematical Research Letters*, (12) :121–127, 2005.