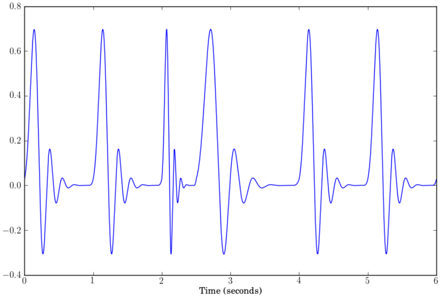
Reconstruction parcimonieuse des signaux

Leo Davy

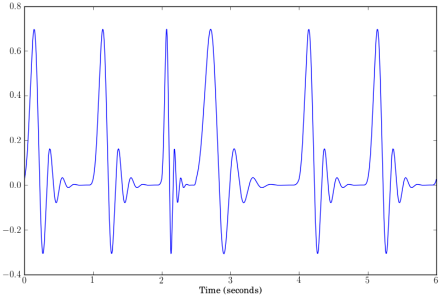
Université de Montpellier

Montpellier, 1 Juin 2021

Signaux



Signaux



Exemple (familles de signaux)

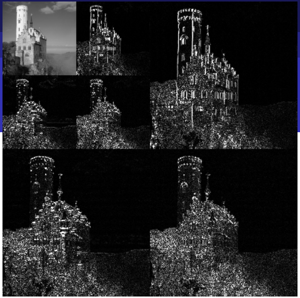
- $\mathbb{R}^N$
- $L^2(\mathbb{R})$

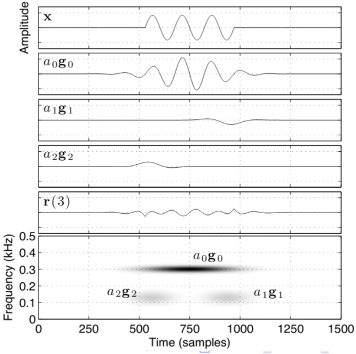
Définition

Un signal est une application $f : X \to Y$

Exemples de problèmes

- Compresser un signal
- Décomposer un signal en éléments simples
- Reconstruire un signal en utilisant un minimum de mesures





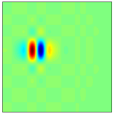
Traitement du signal

Mesure du signal

Opérateur d'analyse

$$A : f \mapsto (c_i(f))_I$$

avec $c_i(f) = \langle f, \varphi_i \rangle$.

$y_m = \Big\langle$  ,  $\Big\rangle$

Traitement du signal

Mesure du signal

Opérateur d'analyse

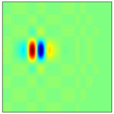
$$A : f \mapsto (c_i(f))_I$$

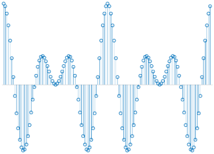
avec $c_i(f) = \langle f, \varphi_i \rangle$.

Reconstruction

Opérateur de synthèse

$$S : (c_i(f))_I \mapsto f$$

$y_m = \Big\langle$  ,  $\Big\rangle$



Traitement du signal

Mesure du signal

Opérateur d'analyse

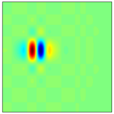
$$A : f \mapsto (c_i(f))_I$$

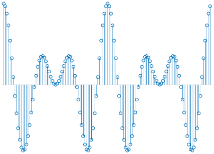
avec $c_i(f) = \langle f, \varphi_i \rangle$.

Reconstruction

Opérateur de synthèse

$$S : (c_i(f))_I \mapsto f$$

$y_m = \Big\langle$  ,  $\Big\rangle$



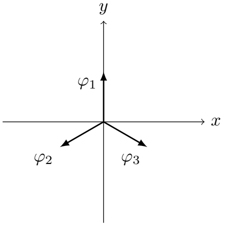
Quelles conditions pour reconstruire n'importe quel signal f de $\mathcal{F}$ à partir de $A(f)$?

Théorie des frames

Définition

On appelle $F = (\varphi_i)_I$ un frame de H si $\exists m, M > 0$ tels que

$$m||f||_2^2 \leq \sum_I |\langle f, \varphi_i \rangle|^2 \leq M||f||_2^2, \quad \forall f \in H$$

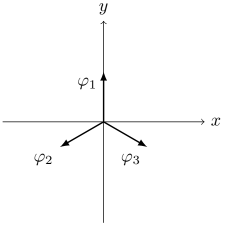


Théorie des frames

Définition

On appelle $F = (\varphi_i)_I$ un frame de H si $\exists m, M > 0$ tels que

$$m||f||_2^2 \leq \sum_I |\langle f, \varphi_i \rangle|^2 \leq M||f||_2^2, \quad \forall f \in H$$

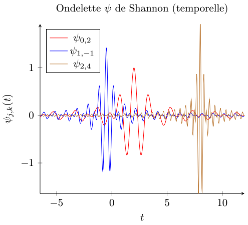


Frame serré

Si $m = M$, formule de reconstruction :

$$f = \frac{1}{M} \sum_I \langle f, f_i \rangle f_i.$$

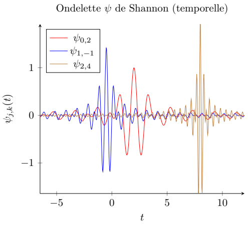
Décomposition en ondelettes



$$\psi_{j,k} := 2^{j/2}\psi(2^j \cdot - k).$$

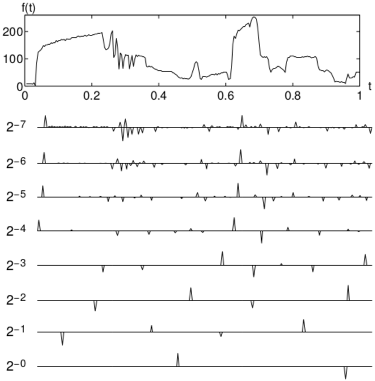
j : échelle (fréquence), k : localisation

Décomposition en ondelettes



$$\psi_{j,k} := 2^{j/2}\psi(2^j \cdot - k).$$

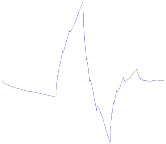
j : échelle (fréquence), k : localisation

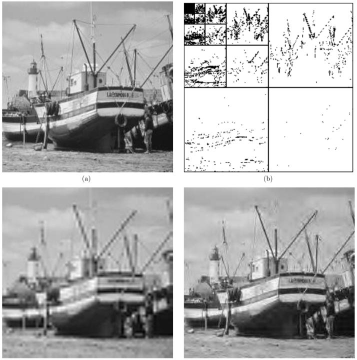


$$f = \sum_{k=-\infty}^{+\infty} \langle f, \varphi_k \rangle \varphi_k + \sum_{j=1}^{+\infty} \left(\sum_{k=-\infty}^{+\infty} \langle f, \psi_{j,k} \rangle \psi_{j,k} \right)$$

Compression par les ondelettes

Moments nuls : $\langle P, \psi_{j,k} \rangle = 0$ si $\deg(P) \leq m$

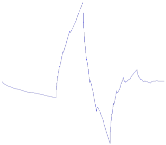


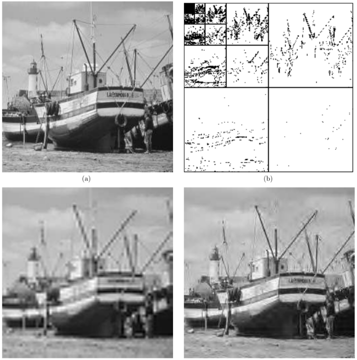


(a) (b)

Compression par les ondelettes

Moments nuls : $\langle P, \psi_{j,k} \rangle = 0$ si $\deg(P) \leq m$



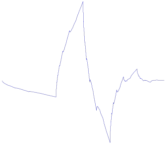


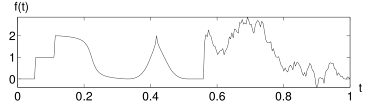
Compression par les ondelettes

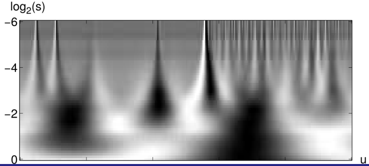
Moments nuls : $\langle P, \psi_{j,k} \rangle = 0$ si $\deg(P) \leq m$

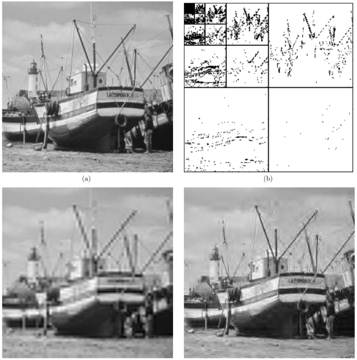
Jaffard 91 : $f \in C^\alpha$ et ψ avec $m \geq \alpha$ moments nuls, alors

$$|\langle f, \psi_{j,k} \rangle| \leq C_\alpha 2^{-j(\alpha + \frac{1}{2})}$$









Une représentation parcimonieuse "optimale" ?

Soit $D = (\varphi_i)_I$ une famille de vecteurs (dictionnaire d'atomes), $f \in Vect(D)$:

1. Reconstruction : Quels $(c_i)_J$ pour que $f = \sum_J c_i \varphi_i$?

Une représentation parcimonieuse "optimale" ?

Soit $D = (\varphi_i)_I$ une famille de vecteurs (dictionnaire d'atomes), $f \in Vect(D)$:

1. Reconstruction : Quels $(c_i)_J$ pour que $f = \sum_J c_i \varphi_i$?
2. Reconstruction parcimonieuse : Quel $(c_i)_J$ minimise le nombre de coefficients non nuls et reconstruit f ?

Une représentation parcimonieuse "optimale" ?

Soit $D = (\varphi_i)_I$ une famille de vecteurs (dictionnaire d'atomes), $f \in Vect(D)$:

1. Reconstruction : Quels $(c_i)_J$ pour que $f = \sum_J c_i \varphi_i$?
2. Reconstruction parcimonieuse : Quel $(c_i)_J$ minimise le nombre de coefficients non nuls et reconstruit f ?

Problème P0 (décomposition atomique) :

$$\min ||x||_0 : f = F_D x$$

avec $x = (c_i)_J$ et F_D matrice dont les colonnes sont les φ_i.

Une représentation parcimonieuse "optimale" ?

Soit $D = (\varphi_i)_I$ une famille de vecteurs (dictionnaire d'atomes), $f \in Vect(D)$:

1. Reconstruction : Quels $(c_i)_J$ pour que $f = \sum_J c_i \varphi_i$?
2. Reconstruction parcimonieuse : Quel $(c_i)_J$ minimise le nombre de coefficients non nuls et reconstruit f ?

Problème P0 (décomposition atomique) :

$$\min ||x||_0 : f = F_D x$$

avec $x = (c_i)_J$ et F_D matrice dont les colonnes sont les φ_i.
Est-ce bien posé ? Est-ce calculable ?

Exemple (Fourier-Dirac)

$D = (T, W)$ avec $T = (\delta_k)_{0 \leq k \leq N-1}$ et $W = (\frac{e^{-\frac{2i\pi k \cdot}{N}}}{\sqrt{N}})_{0 \leq k \leq N-1}$

Théorème (Donoho-Stark 89)

Soit $f \in \mathbb{R}^N$ un signal non nul, alors

$$||F_T f||_0 + ||F_W f||_0 \geq 2\sqrt{N} \tag{1}$$

Théorème (Donoho-Stark 89)

Soit $f \in \mathbb{R}^N$ et $x = (x_T, x_W)$ tel que $f = F_T x_T + F_W x_W$ et $||x||_0 < \sqrt{N}$. Alors x est l'unique solution de P0.

Principe d'incertitude $\rightarrow$ Unicité de la décomposition atomique
Cas particulier du dictionnaire Fourier-Dirac ?

Principe d'incertitude $\rightarrow$ Unicité de la décomposition atomique
Cas particulier du dictionnaire Fourier-Dirac ?
Non, la matrice clef dans la preuve est :

$$F_W^t F_T = \begin{bmatrix} \langle \psi_1, \varphi_1 \rangle & \langle \psi_1, \varphi_2 \rangle & \cdots & \langle \psi_1, \varphi_N \rangle \\ \langle \psi_2, \varphi_1 \rangle & \ddots & \vdots & \langle \psi_2, \varphi_N \rangle \\ \vdots & \cdots & \ddots & \vdots \\ \langle \psi_N, \varphi_1 \rangle & \cdots & \cdots & \langle \psi_N, \varphi_N \rangle . \end{bmatrix}$$

Principe d'incertitude $\rightarrow$ Unicité de la décomposition atomique
Cas particulier du dictionnaire Fourier-Dirac ?
Non, la matrice clef dans la preuve est :

$$F_W^t F_T = \begin{bmatrix} \langle \psi_1, \varphi_1 \rangle & \langle \psi_1, \varphi_2 \rangle & \cdots & \langle \psi_1, \varphi_N \rangle \\ \langle \psi_2, \varphi_1 \rangle & \ddots & \vdots & \langle \psi_2, \varphi_N \rangle \\ \vdots & \cdots & \ddots & \vdots \\ \langle \psi_N, \varphi_1 \rangle & \cdots & \cdots & \langle \psi_N, \varphi_N \rangle . \end{bmatrix}$$

Quantité clef : $M = max_{i,j} |\langle \psi_i, \varphi_j \rangle|$ (incohérence)

Soit $D = (\Psi, \Phi)$ une paire de bases arbitraires de $\mathbb{R}^N$ et $M = max_{i,j} |\langle \psi_i, \varphi_j \rangle|$.

Théorème (Principe d'incertitude généralisé)

Soit $f \in \mathbb{R}^N$ un signal non nul. Alors

$$||F_\Psi f||_0 + ||F_\Phi f||_0 \geq \frac{2}{M}$$

Dictionnaire Fourier-Dirac : $M = \frac{1}{\sqrt{N}}$.

Soit $D = (\Psi, \Phi)$ une paire de bases arbitraires de $\mathbb{R}^N$ et $M = max_{i,j}|\langle \psi_i, \varphi_j \rangle|$.

Théorème (Principe d'incertitude généralisé)

Soit $f \in \mathbb{R}^N$ un signal non nul. Alors

$$||F_\Psi f||_0 + ||F_\Phi f||_0 \geq \frac{2}{M}$$

Théorème (Unicité de la solution de P0 généralisé)

Soit $f \in \mathbb{R}^N$ alors la décomposition atomique x de f est unique si $||x||_0 < \frac{1}{M}$.

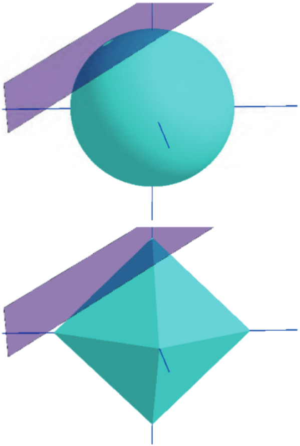
Dictionnaire Fourier-Dirac : $M = \frac{1}{\sqrt{N}}$.

P1 :

$$\min ||x||_1 : f = F_D x \qquad (2)$$

Avantages :

1. Sous des hypothèses de parcimonie, la solution de P1 coincide avec celle de P0.

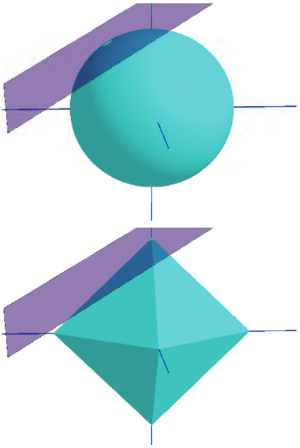


P1 :

$$\min ||x||_1 : f = F_D x \qquad (2)$$

Avantages :

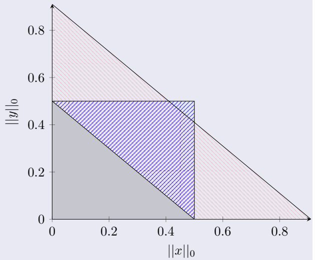
1. Sous des hypothèses de parcimonie, la solution de P1 coincide avec celle de P0.
2. Beaucoup d'algorithmes efficaces de résolution (Programmation linéaire)



Théorème

Soit $D = (\Phi, \Psi)$ *une paire de bases orthonormales et soit* $f \in \mathbb{R}^N$ *un signal,* x_0 *la solution de P0 pour* $f = F_D x$*, si l'une des hypothèses suivantes est vérifiée :*

- *(Donoho-Huo 98)*
  $||x_{0,\Phi}||_0 + ||x_{0,\Psi}|| < \frac{1}{2M}$

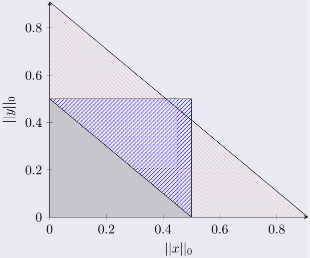


alors x_0 *est l'unique solution de P1.*

Théorème

Soit $D = (\Phi, \Psi)$ *une paire de bases orthonormales et soit* $f \in \mathbb{R}^N$ *un signal,* x_0 *la solution de P0 pour* $f = F_D x$*, si l'une des hypothèses suivantes est vérifiée :*

- *(Donoho-Huo 98)*
  $||x_{0,\Phi}||_0 + ||x_{0,\Psi}|| \; < \frac{1}{2M}$
- *(Elad-Bruckstein 02)*
  $||x_{0,\Phi}||_0 + ||x_{0,\Psi}||_0 < \frac{\sqrt{2}-0.5}{M} \; \sim \frac{0.92}{M}$

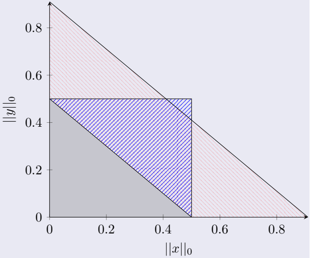


alors x_0 *est l'unique solution de P1.*

Théorème

Soit $D = (\Phi, \Psi)$ *une paire de bases orthonormales et soit* $f \in \mathbb{R}^N$ *un signal,* x_0 *la solution de P0 pour* $f = F_D x$*, si l'une des hypothèses suivantes est vérifiée :*

- *(Donoho-Huo 98)*
  $||x_{0,\Phi}||_0 + ||x_{0,\Psi}|| \; < \frac{1}{2M}$
- *(Elad-Bruckstein 02)*
  $||x_{0,\Phi}||_0 + ||x_{0,\Psi}||_0 < \frac{\sqrt{2}-0.5}{M} \sim \frac{0.92}{M}$
- $||x_{0,\Phi}||_0 < \frac{1}{2M}$ *et* $||x_{0,\Psi}||_0 < \frac{1}{2M}$

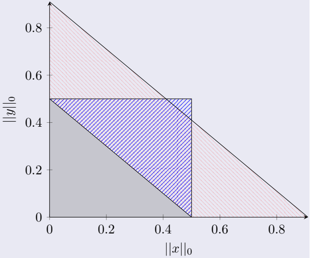


alors x_0 *est l'unique solution de P1.*

Théorème

Soit $D = (\Phi, \Psi)$ une paire de bases orthonormales et soit $f \in \mathbb{R}^N$ un signal, x_0 la solution de P0 pour $f = F_D x$, si l'une des hypothèses suivantes est vérifiée :

- *(Donoho-Huo 98)* $||x_{0,\Phi}||_0 + ||x_{0,\Psi}|| \; < \frac{1}{2M}$
- *(Elad-Bruckstein 02)* $||x_{0,\Phi}||_0 + ||x_{0,\Psi}||_0 < \frac{\sqrt{2}-0.5}{M} \sim \frac{0.92}{M}$
- $||x_{0,\Phi}||_0 < \frac{1}{2M}$ *et* $||x_{0,\Psi}||_0 < \frac{1}{2M}$



alors x_0 est l'unique solution de P1.

Généralisations : Dictionnaire de bases arbitraire, frames (Torrésani 15).

Comment démontrer ces théorèmes ?
Avec un principe d'incertitude.

Comment démontrer ces théorèmes ?
Avec un principe d'incertitude.
$\mathcal{N} = \{\delta : F_D\delta = 0\}$. On pose

$$\mu(\Gamma_\Phi, \Gamma_\Psi) = \sup_{\delta \in \mathcal{N}} \frac{\sum_{i\in\Gamma_\Phi} |\delta_{\Phi,i}| + \sum_{i\in\Gamma_\Psi} |\delta_{\Psi,i}|}{||\delta_\Phi||_1 + ||\delta_\Psi||_1} \quad (3)$$

Comment démontrer ces théorèmes ?
Avec un principe d'incertitude.
$\mathcal{N} = \{\delta : F_D\delta = 0\}$. On pose

$$\mu(\Gamma_\Phi, \Gamma_\Psi) = \sup_{\delta \in \mathcal{N}} \frac{\sum_{i \in \Gamma_\Phi} |\delta_{\Phi,i}| + \sum_{i \in \Gamma_\Psi} |\delta_{\Psi,i}|}{||\delta_\Phi||_1 + ||\delta_\Psi||_1} \tag{3}$$

Lemme (Donoho-Huo 98)

Si $\mu(\Gamma_\Phi, \Gamma_\Psi) < \frac{1}{2}$ alors n'importe quelle solution de $f = F_D x$ supportée sur $\Gamma_\Phi \cup \Gamma_\Psi$ est l'unique solution de P1.

On sait :

- Trouver la décomposition atomique d'un signal parcimonieux

On sait :

- Trouver la décomposition atomique d'un signal parcimonieux

Difficultés :

- Hypothèses de parcimonie fortes

On sait :

- Trouver la décomposition atomique d'un signal parcimonieux

Difficultés :

- Hypothèses de parcimonie fortes

Solutions :

- Candes-Romberg 06 : Pour l'équivalence P0-P1, avec probabilité tendant vers 1, il suffit d'avoir $||x||_0 < \frac{C}{M^2 \log(N)}$.

On sait :

- Trouver la décomposition atomique d'un signal parcimonieux
- Représenter un signal régulier avec peu de coefficients (ondelettes)

Difficultés :

- Hypothèses de parcimonie fortes

Solutions :

- Candes-Romberg 06 : Pour l'équivalence P0-P1, avec probabilité tendant vers 1, il suffit d'avoir $||x||_0 < \frac{C}{M^2 \log(N)}$.

On sait :

- Trouver la décomposition atomique d'un signal parcimonieux
- Représenter un signal régulier avec peu de coefficients (ondelettes)

Difficultés :

- Hypothèses de parcimonie fortes
- Il faut faire toutes les mesures puis enlever la plupart des coefficients

Solutions :

- Candes-Romberg 06 : Pour l'équivalence P0-P1, avec probabilité tendant vers 1, il suffit d'avoir $||x||_0 < \frac{C}{M^2 \log(N)}$.

On sait :

- Trouver la décomposition atomique d'un signal parcimonieux
- Représenter un signal régulier avec peu de coefficients (ondelettes)

Difficultés :

- Hypothèses de parcimonie fortes
- Il faut faire toutes les mesures puis enlever la plupart des coefficients

Solutions :

- Candes-Romberg 06 : Pour l'équivalence P0-P1, avec probabilité tendant vers 1, il suffit d'avoir $||x||_0 < \frac{C}{M^2 \log(N)}$.
- Candes-Romberg 06 : Pour reconstruire un signal avec S coefficients non nuls, avec probabilité tendant vers 1, il suffit de faire $S \log(N)$ mesures et résoudre P1.

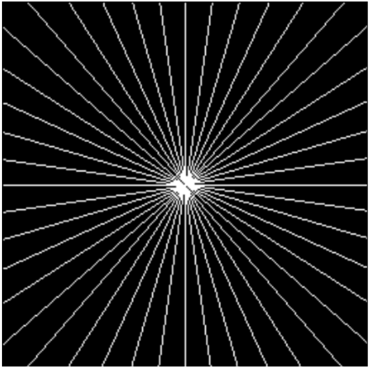
Compressed sensing

Idée : Exploiter la parcimonie pour faire moins de mesures

Compressed sensing

<u>Idée :</u> Exploiter la parcimonie pour faire moins de mesures

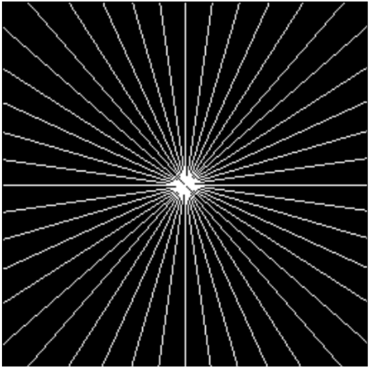
Cadre du problème : F_Φ matrice d'analyse, $\Omega \subset \{0, \cdots, N-1\}$. On mesure $y_\Omega = F_{\Phi,\Omega} s$, on veut retrouver s.



Compressed sensing

<u>Idée :</u> Exploiter la parcimonie pour faire moins de mesures

Cadre du problème : F_Φ matrice d'analyse, $\Omega \subset \{0, \cdots, N-1\}$. On mesure $y_\Omega = F_{\Phi,\Omega} s$, on veut retrouver s.

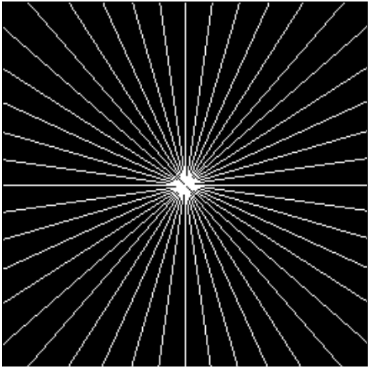


Difficulté : système sous déterminé $K := \mathbb{E}(|\Omega|) << N$

Compressed sensing

Idée : Exploiter la parcimonie pour faire moins de mesures

Cadre du problème : F_Φ matrice d'analyse, $\Omega \subset \{0, \cdots, N-1\}$. On mesure $y_\Omega = F_{\Phi,\Omega} s$, on veut retrouver s.

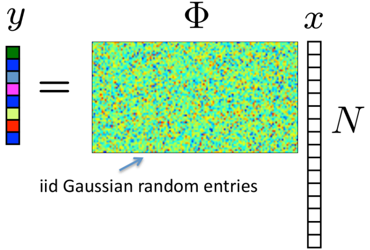


Difficulté : système sous déterminé $K := \mathbb{E}(|\Omega|) << N$

Solution : s est parcimonieux dans une base

P1 (Compressed Sensing) :

$$\min_{x\in\mathbb{R}^N} ||x||_1 : F_{\Phi,\Omega}x = F_{\Phi,\Omega}s$$



- Matrice gaussienne $F_{G,\Omega} = (c_{i,j})_{(i,j)\in\Omega\times\{0,\cdots,N-1\}}$ avec $c_{i,j} \sim N(0, \frac{1}{\sqrt{N}})$
- Matrice de Fourier $F_{F,\Omega}$ restriction aux lignes indexées par Ω de la matrice de Fourier.

Propriétés à vérifier pour la matrice d'analyse

Définition

On dit que F_Ω vérifie

- λ-UUP si
$$\frac{K}{2N}||x||_2^2 \leq ||F_\Omega x||_2^2 \leq \frac{3K}{2N}||x||_2^2$$

est vrai avec probabilité au moins $1 - \mathcal{O}(N^{-\rho/\alpha})$ pour tout signal $x \in \mathbb{R}^N$ avec moins de $\alpha\frac{K}{\lambda}$ coefficients non nuls.

Propriétés à vérifier pour la matrice d'analyse

Définition

On dit que F_Ω vérifie

- λ-UUP si
$$\frac{K}{2N}||x||_2^2 \leq ||F_\Omega x||_2^2 \leq \frac{3K}{2N}||x||_2^2$$

- λ-ERP si, en notant $\sigma = sign(x)$, il existe $P \in \mathbb{R}^N$ tel que
  1. $P(t) = \sigma(t), \forall t \in supp(x)$
  2. $\exists Q \in \mathbb{R}^\Omega$ tel que $P = F_\Omega^* Q$
  3. $P(t) < \frac{1}{2}, \forall t \notin supp(x)$

est vrai avec probabilité au moins $1 - \mathcal{O}(N^{-\rho/\alpha})$ pour tout signal $x \in \mathbb{R}^N$ avec moins de $\alpha\frac{K}{\lambda}$ coefficients non nuls.

Exemple

- La matrice Gaussienne $F_{G,\Omega}$ vérifie $\log(N)$-UUP et $\log(N)$-ERP
- La matrice de Fourier $F_{F,\Omega}$ vérifie $\log(N)^5$-UUP et $\log(N)$-ERP

Théorème (Candes-Tao 04)

Soit F_Ω qui vérifie λ-ERP et λ-UUP. Soit $K \geq \lambda$.
Soit s un signal dans $\mathbb{R}^N$ tel que ses coefficients dans une base de référence décroissent comme :

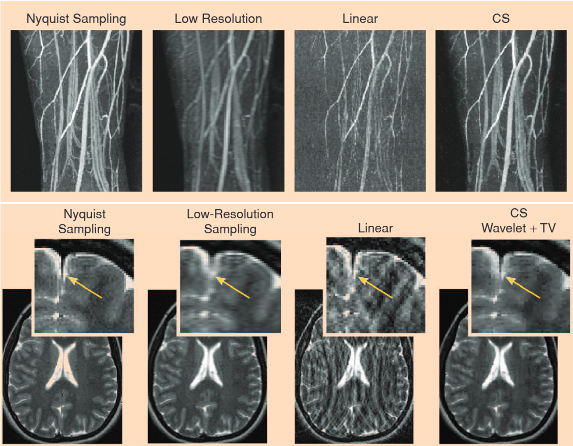
$$|\theta_{(n)}| \ \leq Cn^{-\frac{1}{p}}$$

pour un certain $C > 0$ et $0 < p \leq 1$.
On pose $r = \frac{1}{p} - \frac{1}{2}$, alors n'importe quelle solution x de (P1) vérifie :

$$||s - x||_2 \leq C_r \left(\frac{K}{\lambda}\right)^{-r}$$

avec probabilité au moins $1 - \mathcal{O}(N^{-\frac{\rho}{\alpha}})$, pour certains ρ et α.

Conséquences des théorèmes



Conclusion

- On peut représenter de façon parcimonieuse des signaux réguliers (ondelettes)
- On peut résoudre le problème de décomposition atomique (P1)
- Pour reconstruire un signal S-parcimonieux, $S \log(N)$ mesures suffisent (Compressed sensing)