Quelques problèmes ... de taille ...
pour le traitement du signal et de l'image de demain

Rémi Gribonval
Directeur de Recherche INRIA
Equipe METISS
INRIA Rennes - Bretagne Atlantique
remi.gribonval@inria.fr
http://www.irisa.fr/metiss/members/remi

Signaux et images sont omniprésents

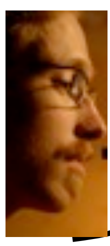
- Signaux
  - Son, parole, musique
  - Audiovisuel
  - Son 3D
  - Compression MP3
  - Biomédical ECG/EEG/MEG
  - Traffic réseau
  - Communications ultra-large bande
  - Radar
  - Géologie

Equipe METISS

- Images
  - Imagerie médicale
  - Imagerie satellite
  - Astronomie, cosmologie
  - Compression JPEG
  - Réalité virtuelle
- Mais aussi:
  - Neuro-chips
  - Puces ADN
  - Test non-destructif
  - Interfaces cerveau - machine

Du phénomène à l'information

$s_1(t)$



$s_2(t)$

$p(x,y,z,t)$

Réverberation
Bruit de fond
Mélange de sources

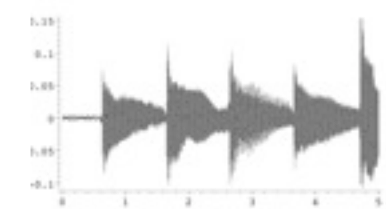
Domaine analogique

$x(t)$

conversion A/N

Domaine numérique

$x[n]$



Stockage

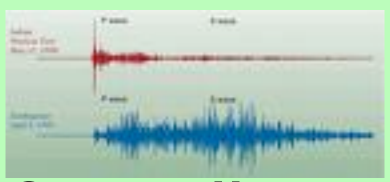
Transmission

Reconnaissance

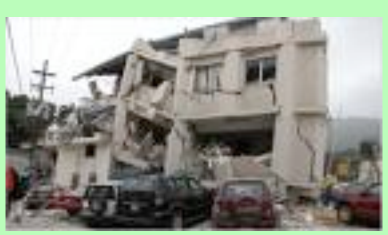
Ex : signaux acoustiques

20 Hertz *Fréquence* 20 000 Hertz

Infrasons



Surveillance sismique



Domaine audible = audio

MP3

Téléphonie

Aides auditives

Home cinema

Ultrasons

Echographie

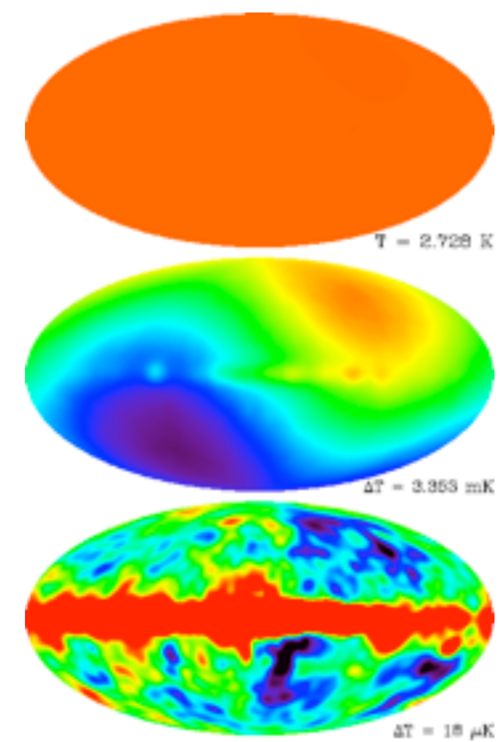
Mais ausssi ...

Interfaces cerveau machine

Imagerie médicale

Sécurité dans les transports ?

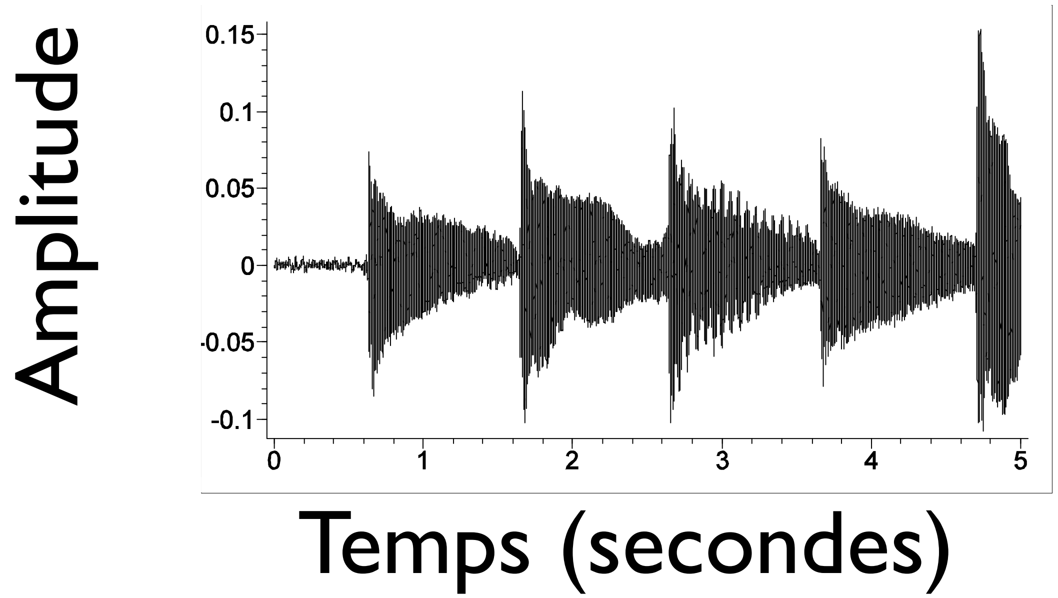
Astronomie



Analyser

Représentation temporelle

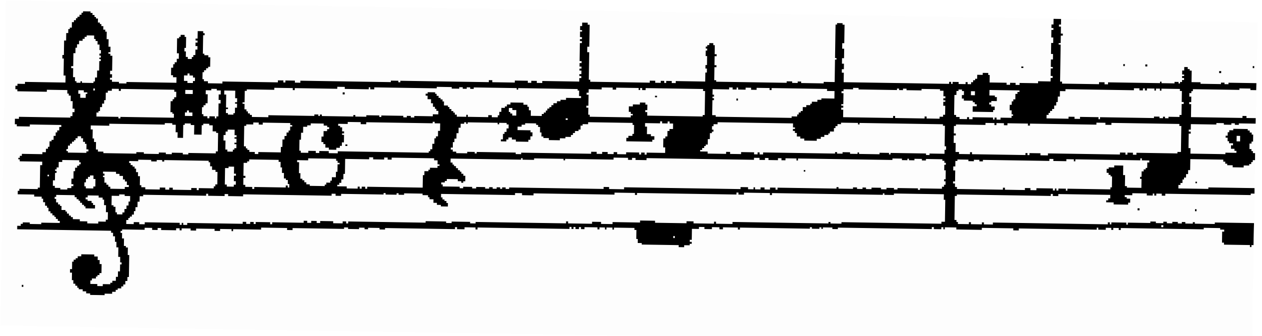
- Un son = une onde = une vibration, plus ou moins rapide et intense au cours du temps



- «Forme d'onde» = peu pratique ...

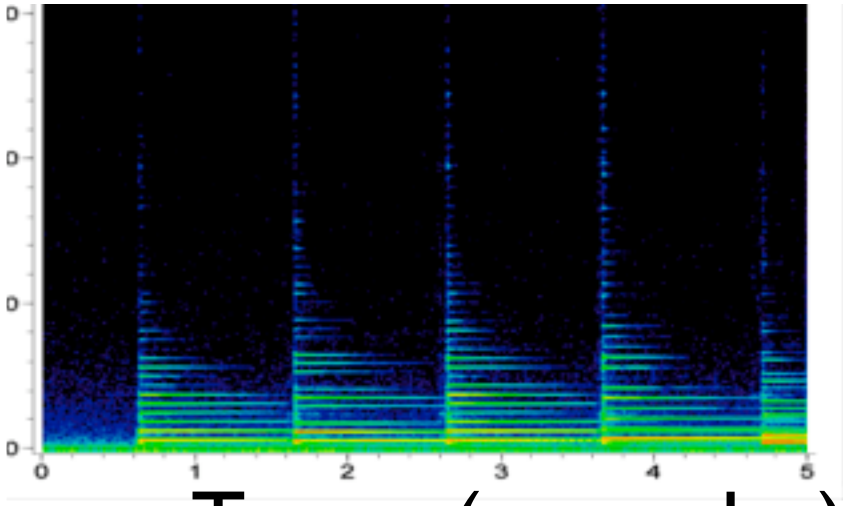
Représentation temps-fréquence

- Une visualisation plus familière aux musiciens ...

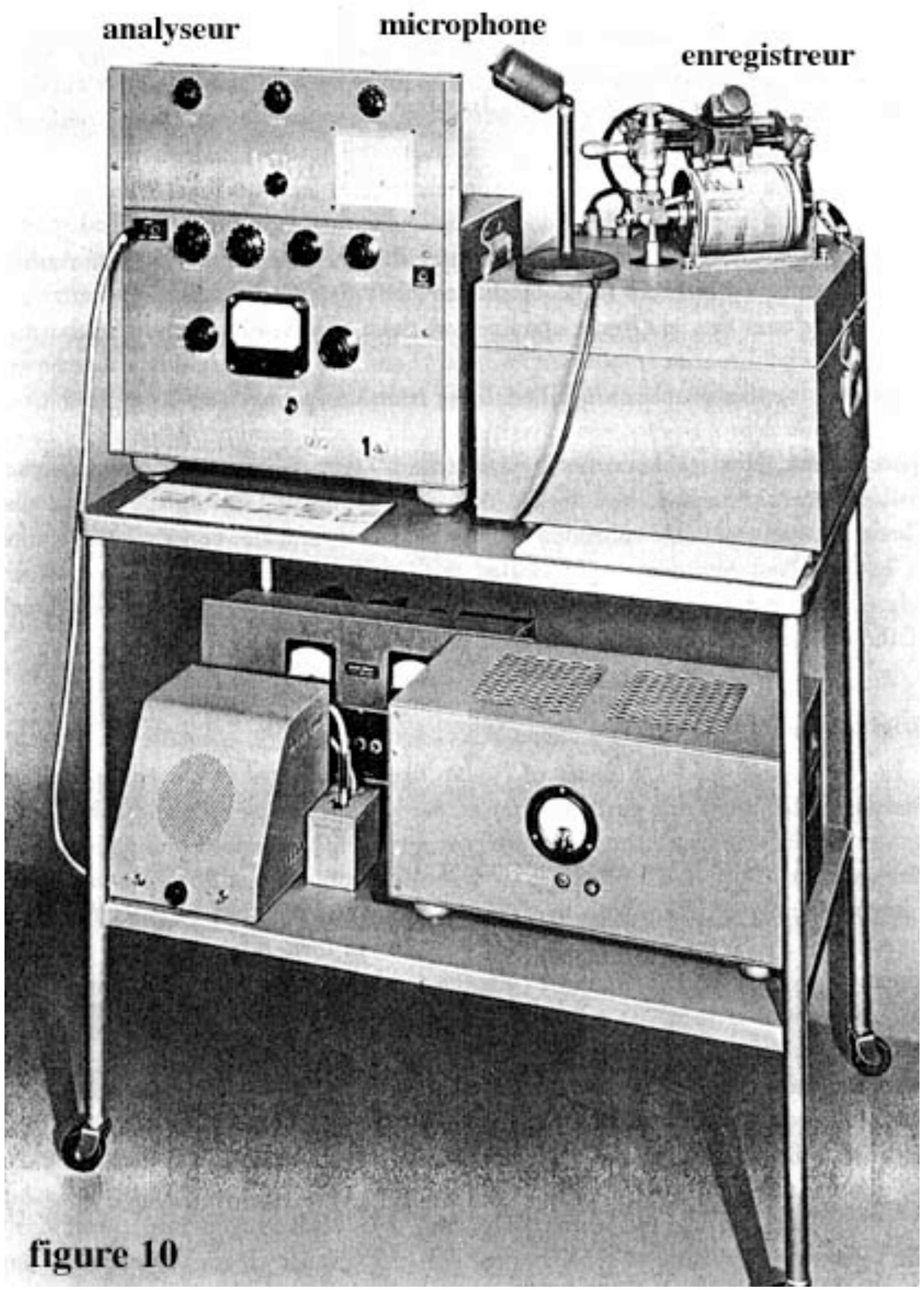


- Outil mathématique : le spectrogramme

Fréquence (Hertz)

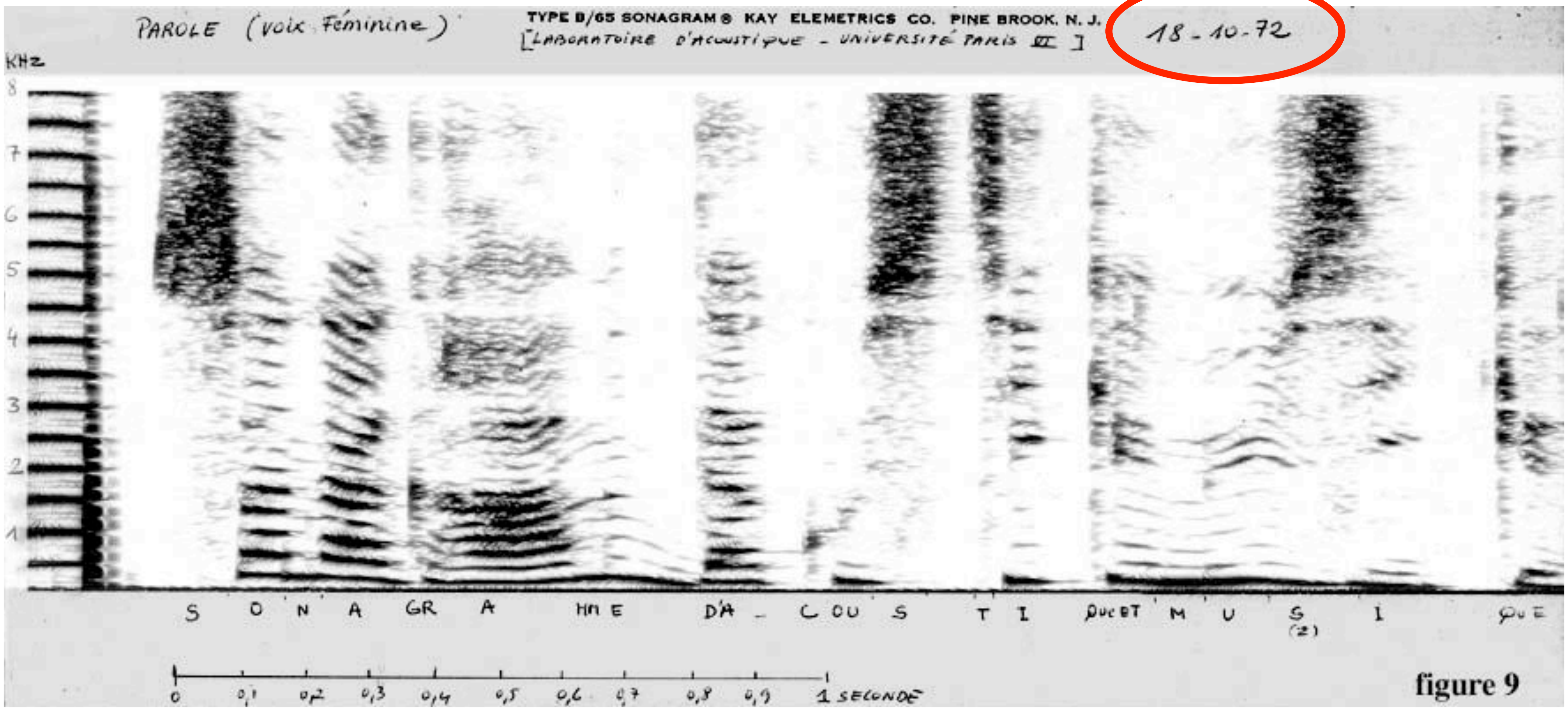


Temps (secondes)



Bell, vers 1940

Sonagramme mécanique

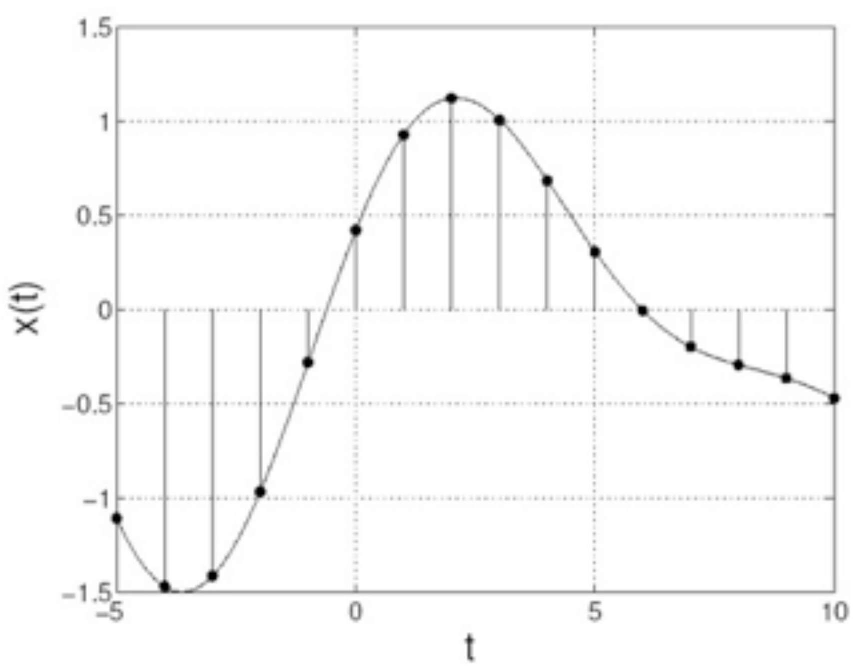


Reconnaissance de la parole

Numériser ...
et compresser

Numérisation

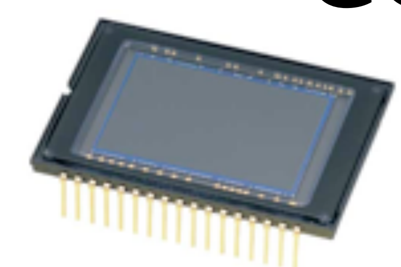
- Signaux et **échantillons**

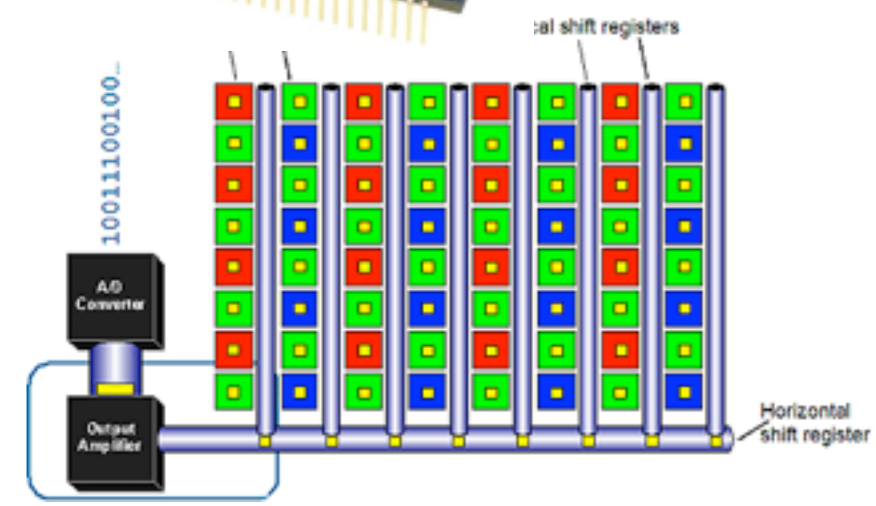


CD = 44100 échantillons
par seconde par canal

- Images et **pixels**

CCD





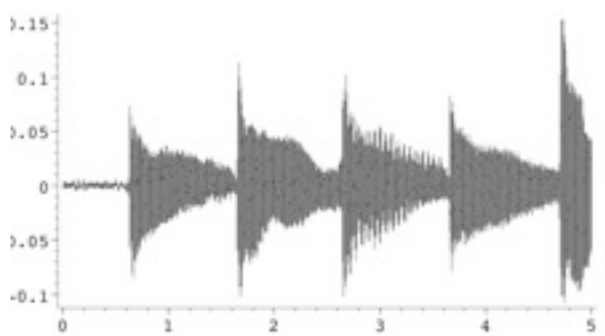
standard ~ 10 Mégapixels

A données volumineuses ...

- **Fait** : données numériques = grands volumes
  - ✦ 1 seconde d'audio stéréo qualité CD = 1,4 Mbits
  - ✦ 1 photo de10 Mpixels non compressée = 240 Mbits
- **Besoin** : représentations «concises»
  - ✦ stockage et transmission (volume / bande passante) ...
  - ✦ manipulation et traitements (complexité algorithmique)

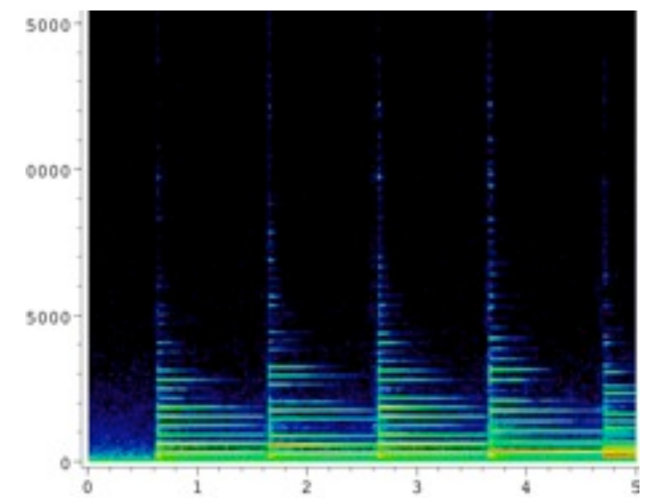
... représentations parcimonieuses

- Audio : représentations temps-fréquence (MP3)



ANALYSE

SYNTHESE



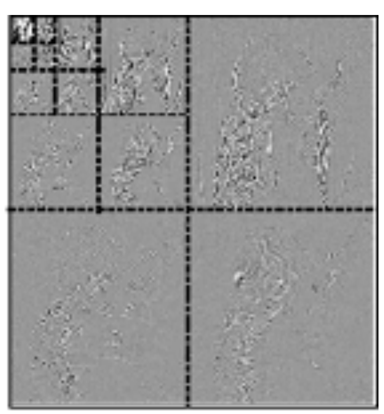
Noir
= zero

- Images : transformée en ondelettes (JPEG2000)

ORIGINAL
128, 129, 125, 64, 65,

ANALYSE

SYNTHESE



TRANSFORM COEFFICIENTS
4123, -12.4, -96.7, 4.5,

Gris
= zero

Expression mathématique

- Signal / image = vecteur de grande dimension

$$y \in \mathbb{R}^N$$

- **Modèle** = superposition linéaire de fonctions de base (ex: *atomes temps-fréquence, ondelettes*)

$$y \approx \sum_k x_k \varphi_k = \mathbf{\Phi} x$$

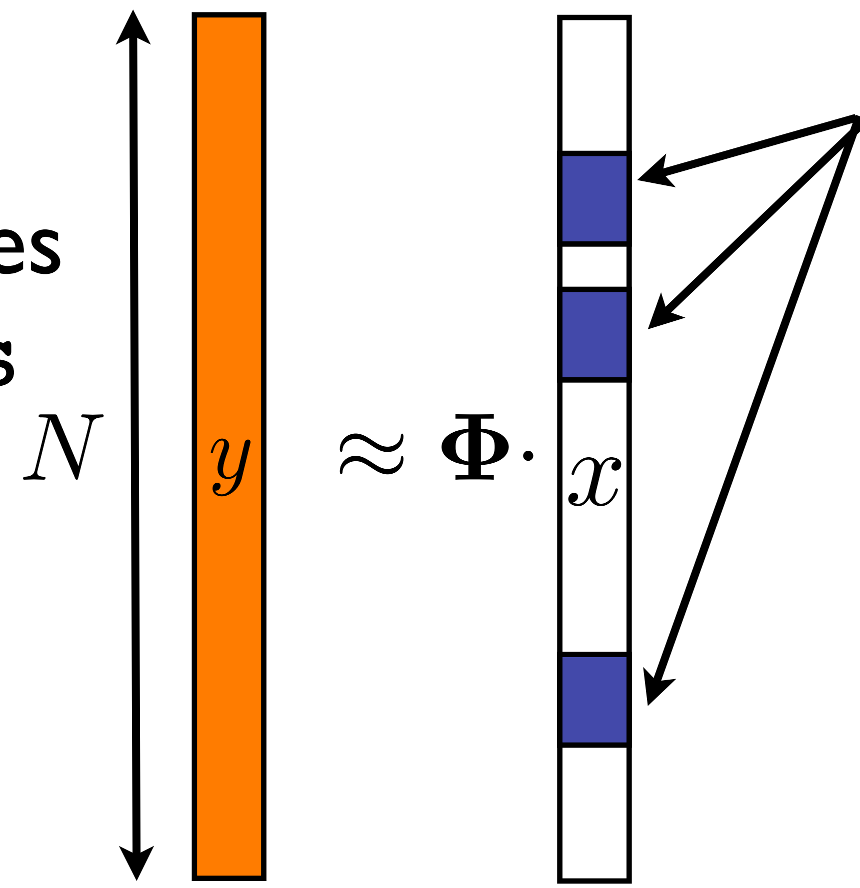
- **Parcimonie** = faible (quasi)-norme L0

$$\|x\|_0 = \sum_k |x_k|^0 = \mathrm{card}\{k, x_k \neq 0\}$$

Parcimonie et compression

- Vecteur plein
- Vecteur parcimonieux

N coordonnées
= N `floats`



$k \ll N$ coordonnées non nulles
= k `floats`

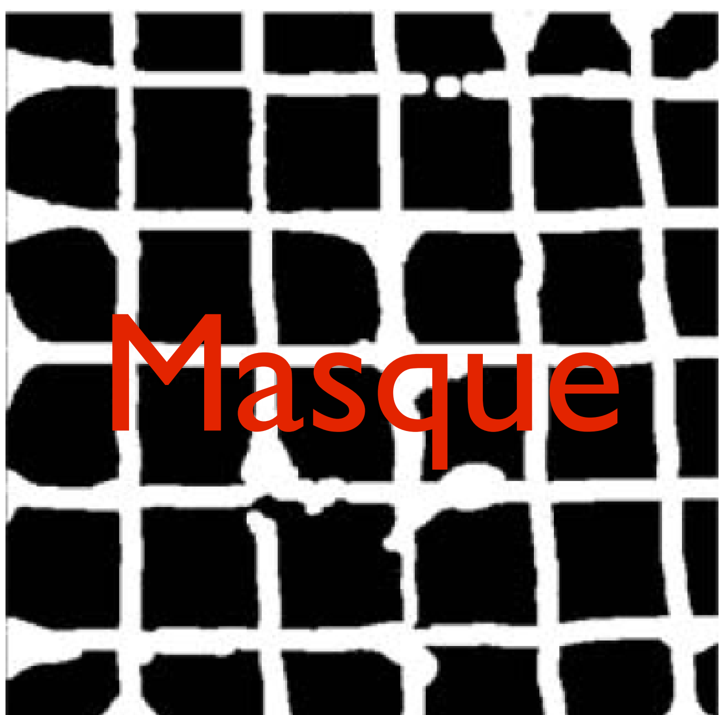
+ k positions parmi N

$$= \log_2 \binom{N}{k} \approx k \log_2 \frac{N}{k} \text{ bits}$$

Extraire de l'information de données incomplètes

Exemple: «inpainting»

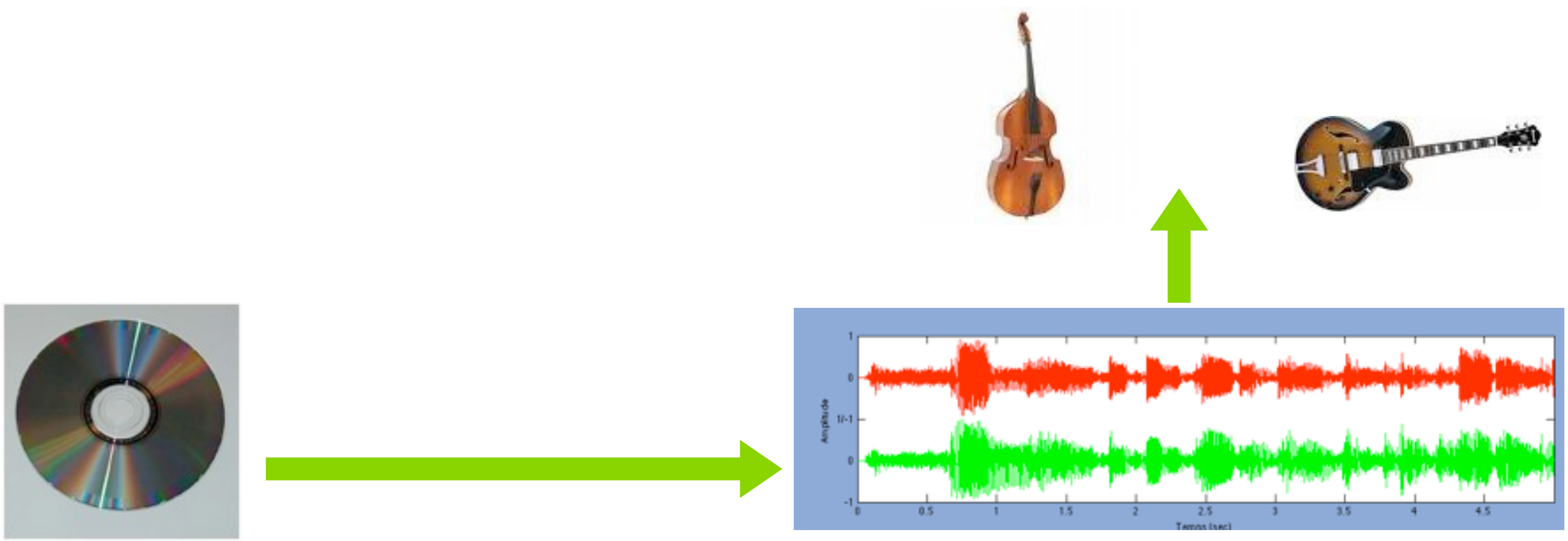
Merci à: G. Peyré, Ceremade, Université Paris 9 Dauphine



Inpainting

Exemple : séparation de sources audio

- « Softly as in a morning sunrise »

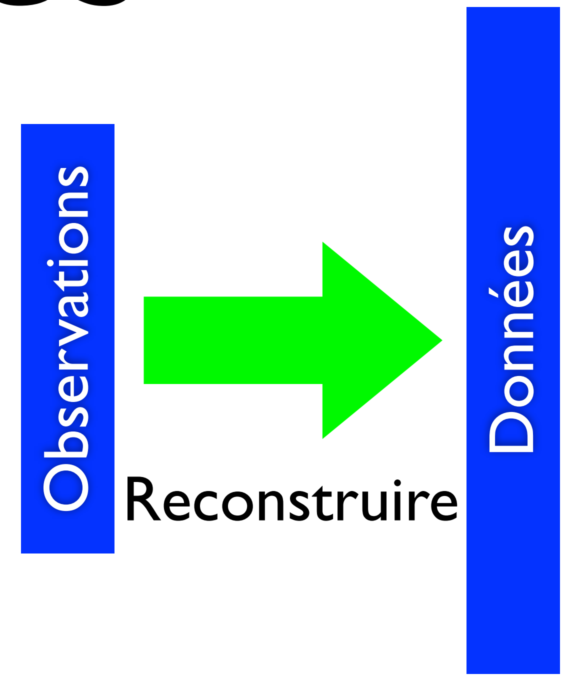


Problèmes inverses

- **Problème inverse** : utiliser une observation indirecte / incomplète des données pour les reconstruire

$$z = \mathbf{M}\mathbf{y}$$

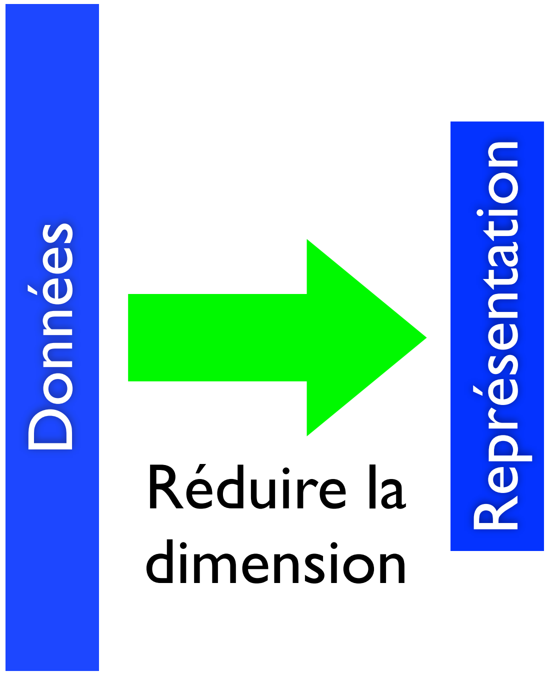
moins d'équations que d'inconnues



- **Parcimonie** : représenter / approcher des données complexes et de grande dimension à l'aide de peu de paramètres

$$\mathbf{y} \approx \mathbf{\Phi} x$$

peu de composantes non nulles



Fondements mathématiques

- Verrou 1990-2000 : si plus d'inconnues que d'équations

$$\mathbf{A}x_0 = \mathbf{A}x_1 \not\Rightarrow x_0 = x_1$$

- Innovations 2001-2006 :
  - ✦ Unicité de la solution parcimonieuse:
    - si x_0, x_1 sont "suffisament parcimonieux",
    - alors $\mathbf{A}x_0 = \mathbf{A}x_1 \Rightarrow x_0 = x_1$
  - ✦ Algorithmes d'identification de x_0
    - Seuillage, Matching Pursuits, Minimisation de normes Lp p<=1,...

Principes algorithmiques pour l'approximation parcimonieuse

Compromis recherché

- Qualité d'approximation

$$\|\mathbf{A}x - \mathbf{b}\|_2$$

- Mesure idéale de parcimonie : "norme" ℓ^0

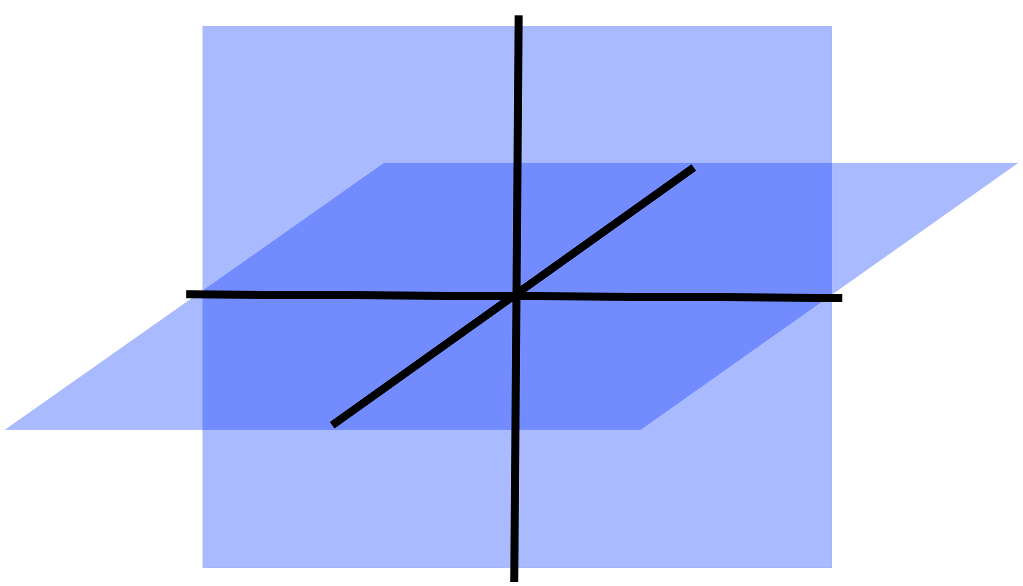
$$\|x\|_0 := \sharp\{n,\ x_n \neq 0\} = \sum_n |x_n|^0$$

- Mesures de parcimonie "relaxées"

$$0 < p < \infty,\ \|x\|_p := \big(\sum_n |x_n|^p\big)^{1/p}$$

Interprétation géométrique et complexité

- Espace des coefficients $\mathbb{R}^N$:
  - ensemble Σ_k de vecteurs parcimonieux $\|x\|_0 \leq k$



$\binom{N}{k}$ sous-espaces

- Ensemble $\mathbf{A}\Sigma_k$ = $\binom{N}{k}$ sous-espaces de l'espace signal
- Parcimonie idéale = plus proche sous-espace parmi $\binom{N}{k}$

Combinatoire !

Théorème: NP-complet

(Quasi)-normes L*p*

- **Normes** lorsque $1 \leq p < \infty$ = convexe

$$\|x\|_p = 0 \Leftrightarrow x = 0$$

$$\|\lambda x\|_p = |\lambda| \|x\|_p, \forall \lambda, x$$

Inégalité triangulaire

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p, \forall x, y$$

- **Quasi-normes** lorsque $0 < p < 1$ = non-convexe

Inégalité *quasi*-triangulaire

$$\|x + y\|_p \leq 2^{1/p} \left(\|x\|_p + \|y\|_p \right), \forall x, y$$

$$\|x + y\|_p^p \leq \|x\|_p^p + \|y\|_p^p, \forall x, y$$

- "*Pseudo*"-norme pour p=0

$$\|x + y\|_0 \leq \|x\|_0 + \|y\|_0, \forall x, y$$

Problèmes d'optimisation

- Approximation

$$\min_x \|\mathbf{b} - \mathbf{A}x\|_2 \text{ s.t. } \|x\|_p \leq \tau$$

- «Parcimonisation» sous contrainte

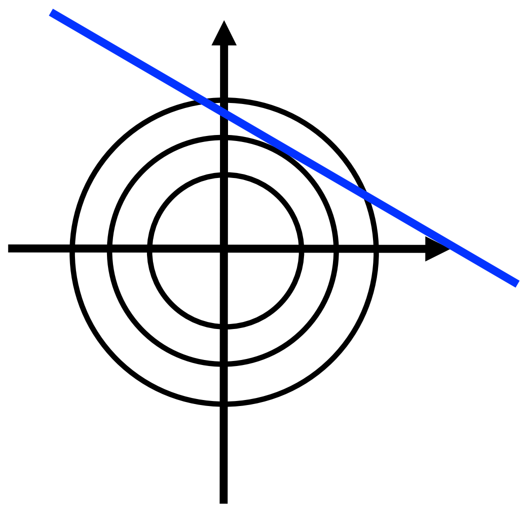
$$\min_x \|x\|_p \text{ s.t. } \|\mathbf{b} - \mathbf{A}x\|_2 \leq \epsilon$$

- Régularisation

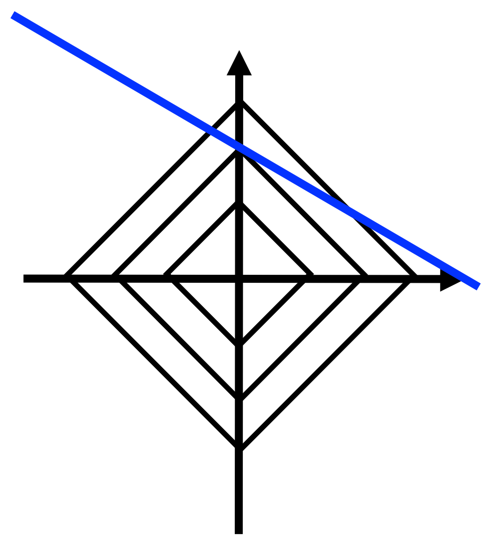
$$\min_x \frac{1}{2}\|\mathbf{b} - \mathbf{A}x\|_2 + \lambda\|x\|_p$$

Courbes de niveau des “normes” L*p*

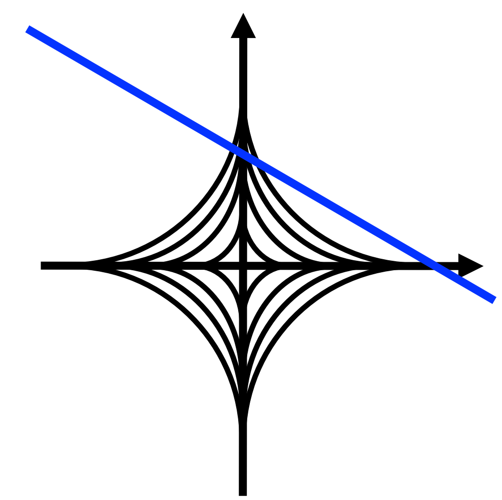
- Strictement convexes lorsque $p>1$



- Convexe $p=1$



- Non-convexe $p<1$



$\{x \text{ s.t.} \mathbf{b} = \mathbf{A}x\}$

Observation: *le minimiseur est parcimonieux*

Algorithmes

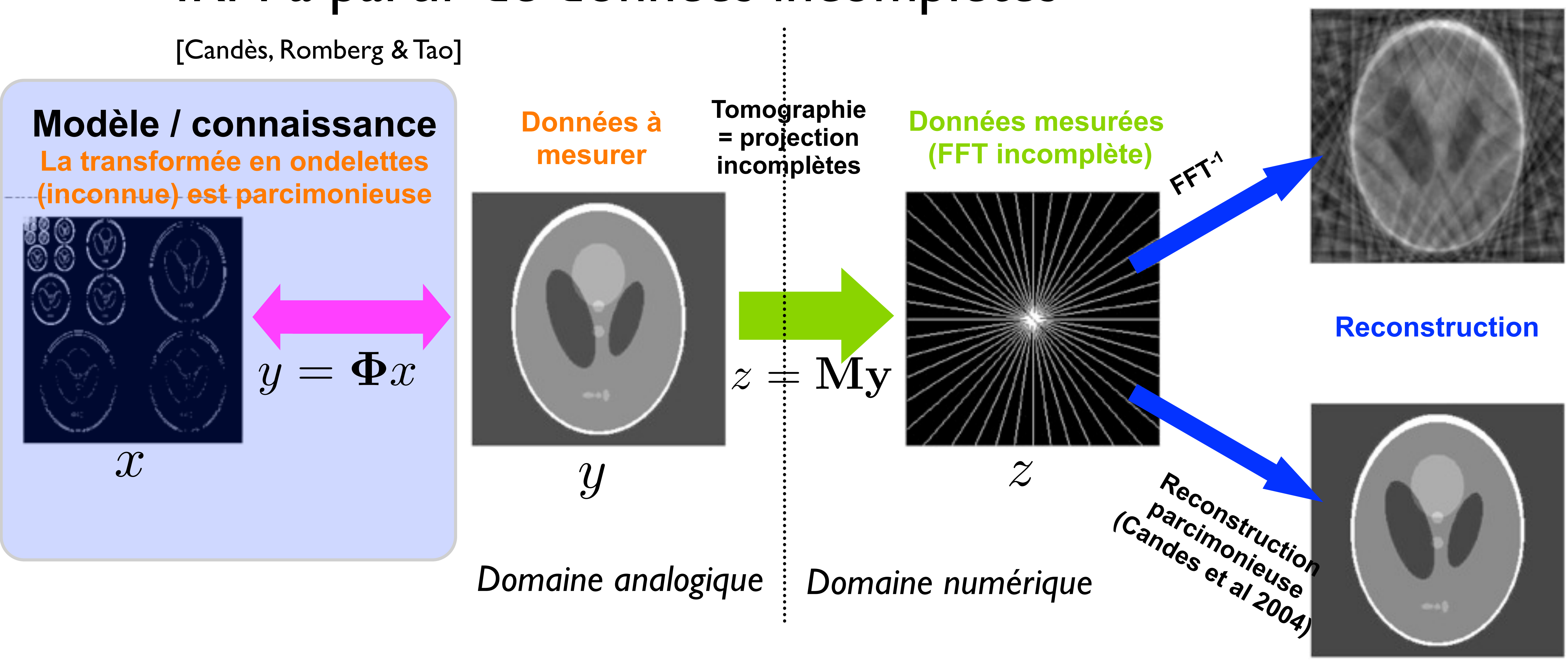
- Approximation parcimonieuse = NP-complet
- Deux grandes approches algorithmiques:
  - optimisation convexe
  - algorithme «greedy»
- Tendances
  - études empiriques + analyse théorique des conditions de succès (géométrie des espaces vectoriels de grande dimension, grandes matrices aléatoires)
  - approches prometteuses combinant
    - optimisation numérique
    - outils issus de l'informatique (hashage, etc.)

Compresser dès l'acquisition

Exemple: tomographie

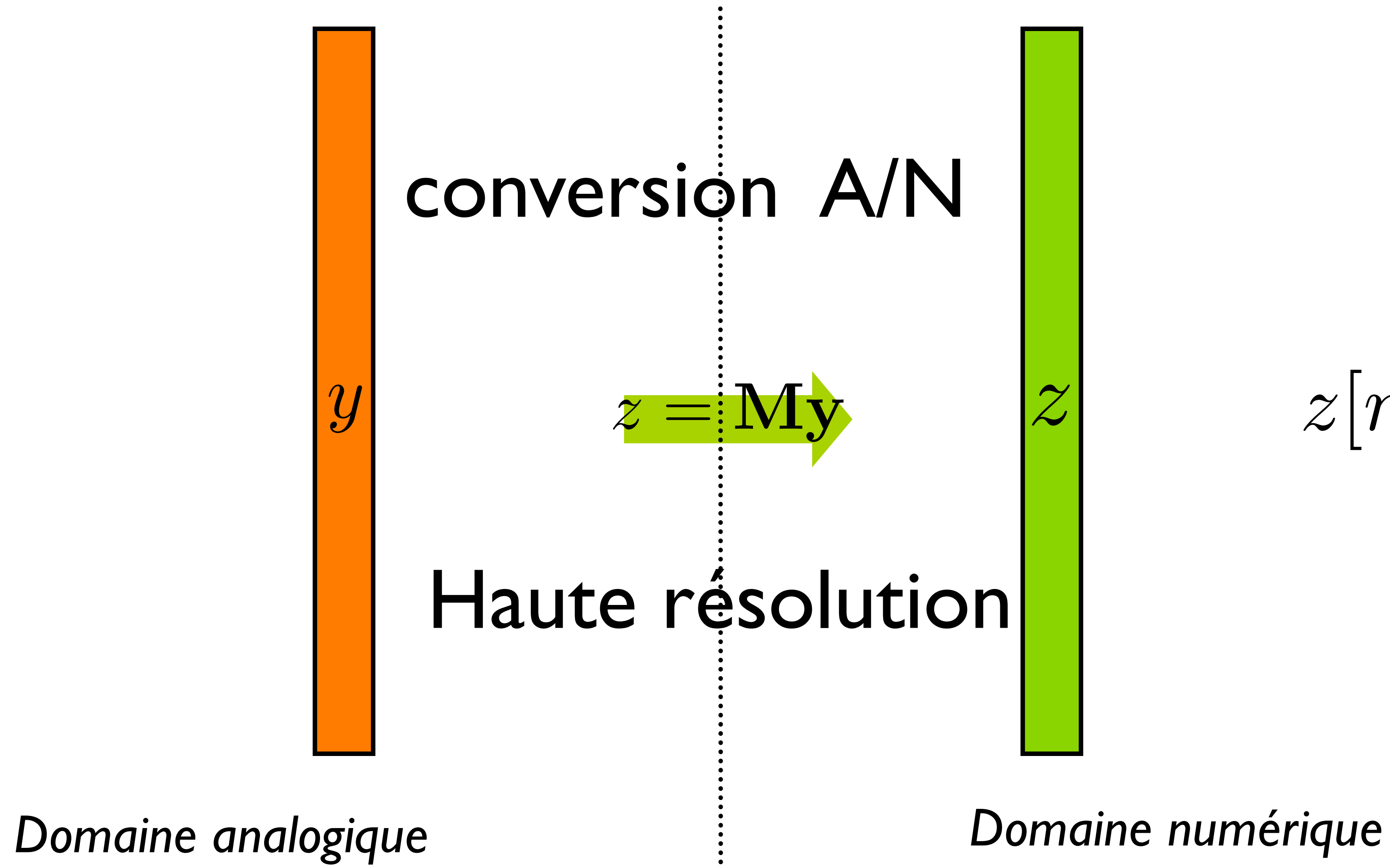
- IRM à partir de données incomplètes

[Candès, Romberg & Tao]



Paradigme standard

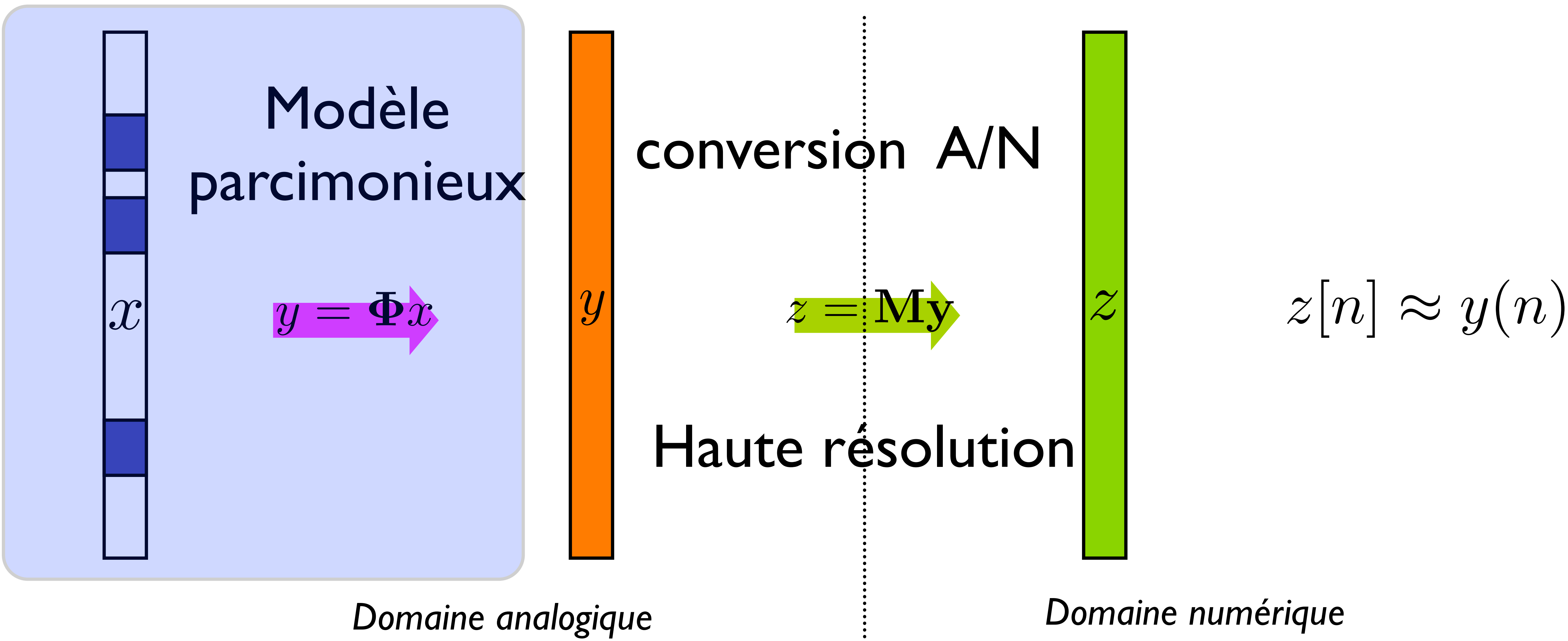
- « Numériser d'abord, ... traiter / compresser ensuite »



$$z[n] \approx y(n)$$

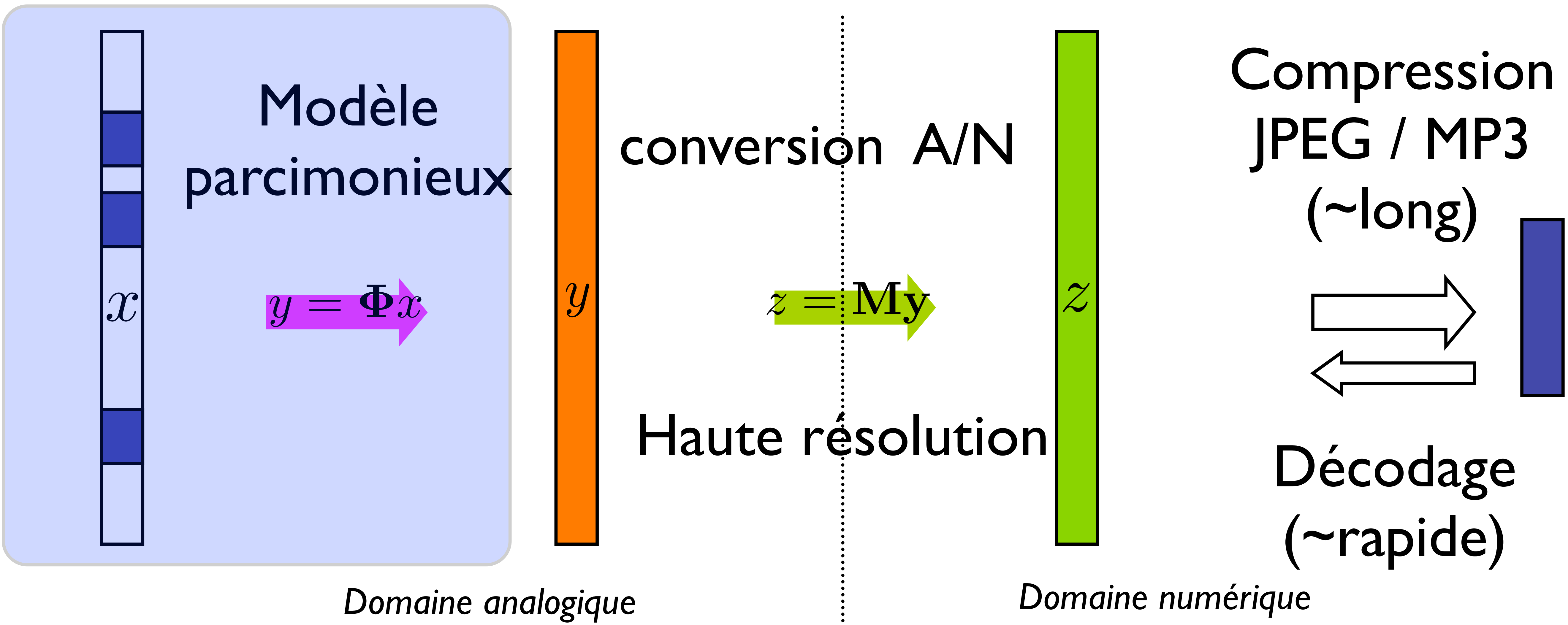
Paradigme standard

- « Numériser d'abord, ... traiter / compresser ensuite »



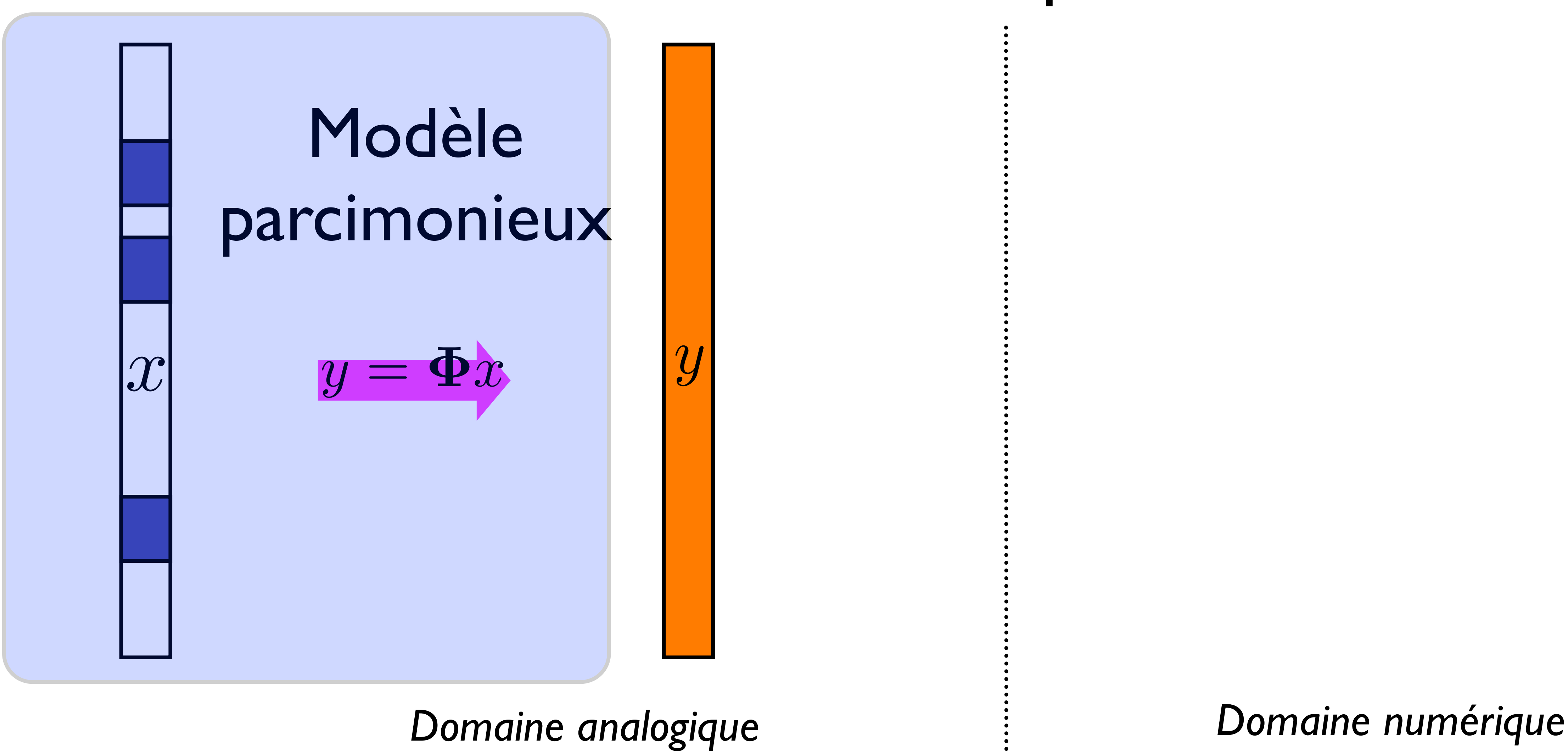
Paradigme standard

- « Numériser d'abord, ... traiter / compresser ensuite »



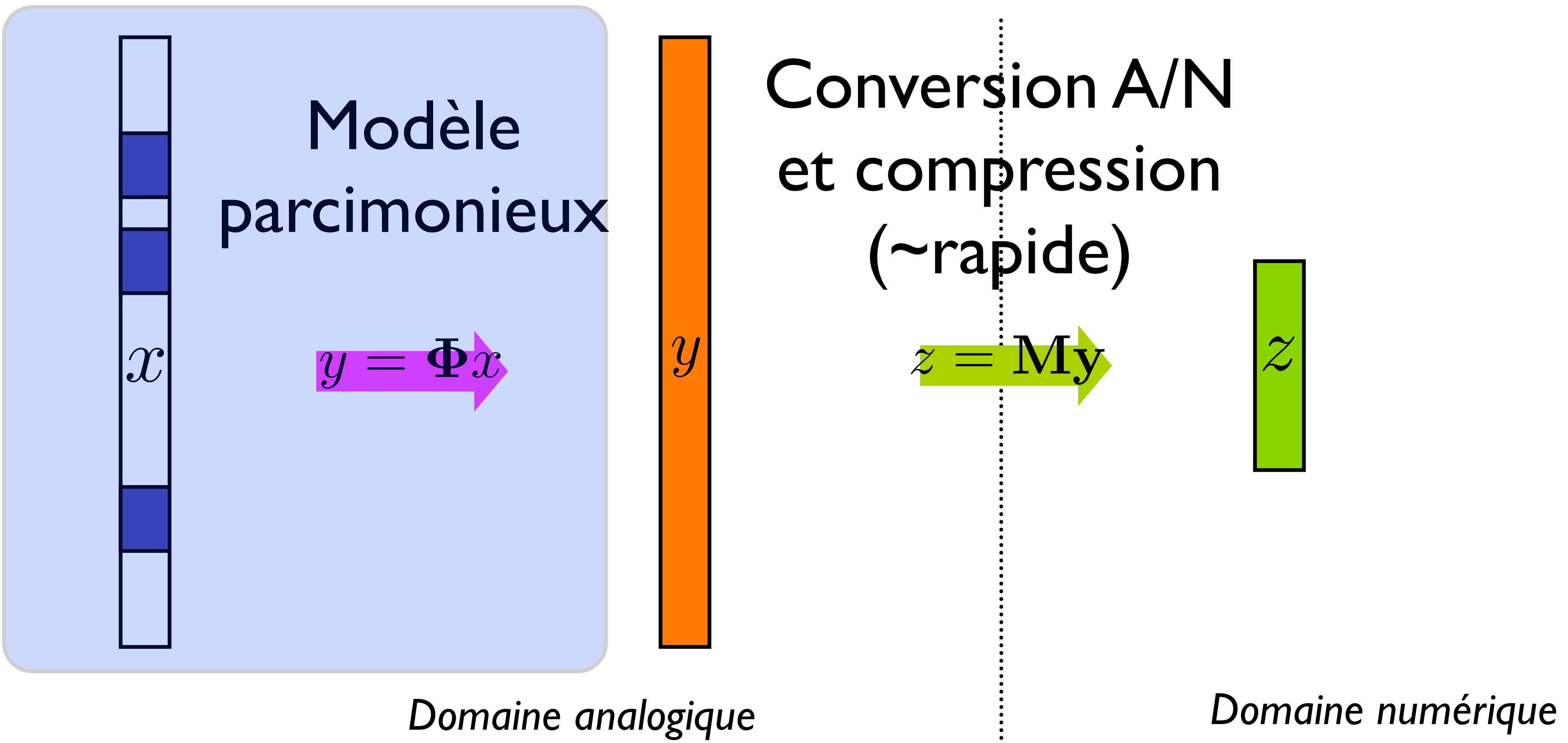
Compressed sensing

- « **Modéliser** d’abord, ... acquérir ensuite »



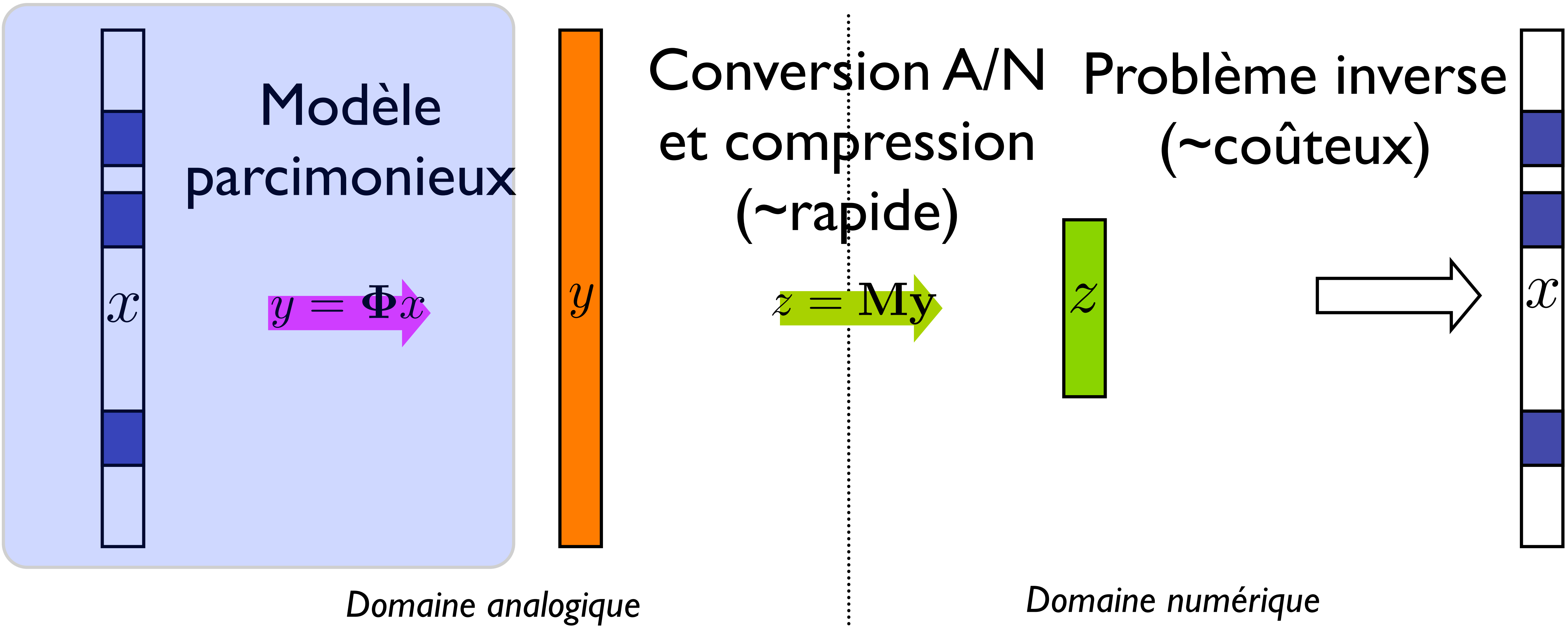
Compressed sensing

- « **Modéliser** d'abord, ... acquérir ensuite »



Compressed sensing

- « **Modéliser** d'abord, ... acquérir ensuite »

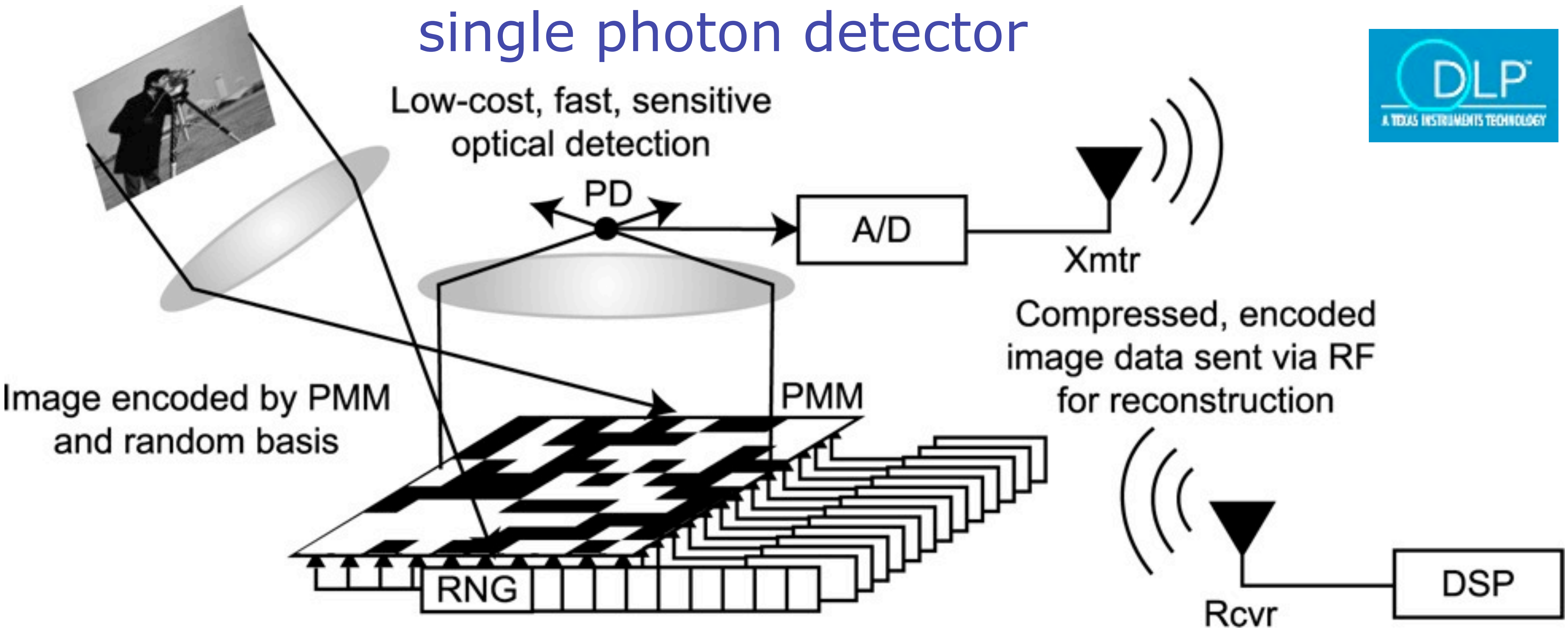


Conditions de succès

- Connaître domaine où données sont parcimonieuses
- «Incohérence» entre domaine de mesure et domaine parcimonieux (*principe d'incertitude* type Heisenberg), par exemple:
  - domaine temporel / domaine fréquentiel
  - domaine spatial / domaine fréquentiel
  - **mesures aléatoires!**
- Nombre de mesure suffisant élevé $m \geq Ck \log_2 \frac{N}{k}$

Exemple : *Caméra à un pixel* de Rice University

single photon detector



Random pattern on DMD array

image reconstruction

Bilan d'étape

Compression
Représentation
Description
Classification

Notion de parcimonie
(Fourier, ondelettes, ...)

Parcimonie

Débruitage
Séparation aveugle de sources
Compressed sensing
...

Rôle naturel / traditionnel

Parcimonie = faible coût (bits, computations, ...)
Objectif direct

Nouveau rôle indirect

Parcimonie = connaissance a priori
Outil pour les problèmes inverses

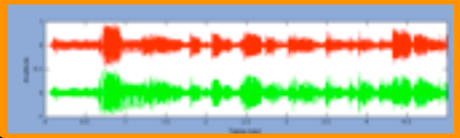
Défis et perspectives

Nouveaux verrous

- Parcimonie : historiquement pour signaux et images
  - ✦ verrou = **passage à l'échelle** des algorithmes

Modèles parcimonieux = Fourier, ondelettes, ...

Multimodal



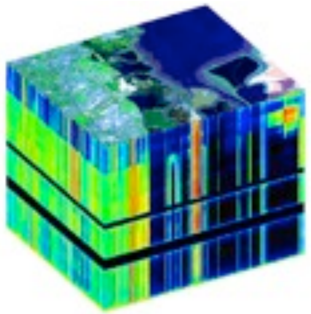
Signaux

Images

Données sur des graphes

Réseaux sociaux

Connectivité du cerveau

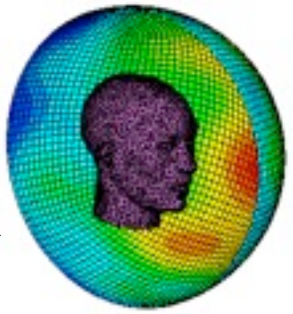


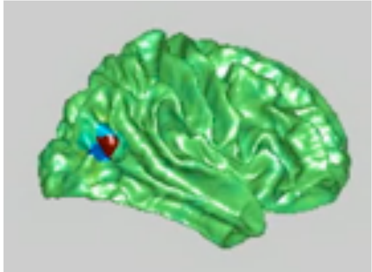
Hyperspectral

Imagerie satellite

Géométrie sphérique

Cosmologie, HRTF (son 3D)





Valeurs vectorielles

Tenseur de diffusion

- Données "exotiques" ou composites
  - ✦ verrou = **généricité des modèles**

Généricité des modèles

- *Grandes approches*
  1. Comprendre le **couplage** entre la parcimonie et les **phénomènes physiques** sous-jacents
  2. Exploiter les **structures** et **invariances** sous-jacentes (graphes, variétés géométriques,...)
  3. Construire des **formalismes génériques** pour modéliser des **données composites** (multicanal, multimodal, ...)
- *Un défi clé :*
  - ✦ **Adapter les modèles** aux données par **apprentissage** à partir de corpus

Passage à l'échelle

- *Approches*
  1. **Comprimer dès l'acquisition** (paradigme du Compressed Sensing)
  2. **Traiter l'information** dans le domaine compressé
  3. Contrôler le **compromis fidélité** des modèles / **efficacité** algorithmique

- Un défi : concevoir **conjointement capteurs** et **algorithmes**