

# Compressed Sensing

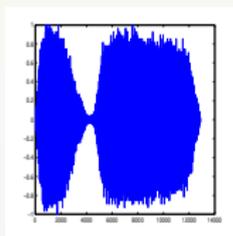
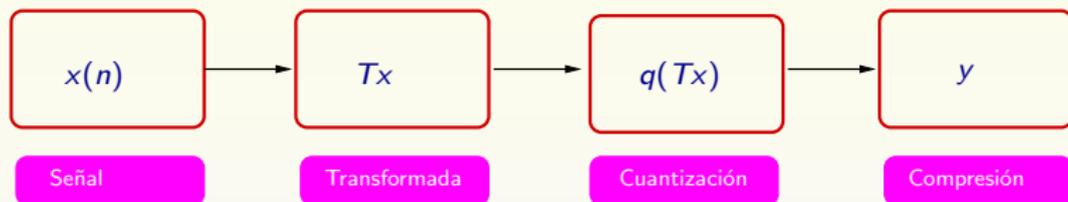
Miguel A. Hernández-Medina

<sup>1</sup>Departamento de Matemática Aplicada  
a las Tecnologías de la Información  
ETSIT, Universidad Politécnica de Madrid

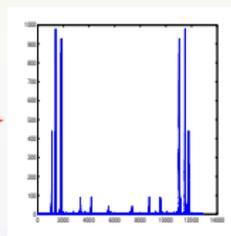
Abril 2010

- 1 ¿Qué es el Compressed Sensing?
- 2 Ejemplo
- 3 Medidas. Propiedad RIP
- 4 Aproximación

# Compresión



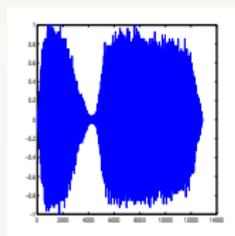
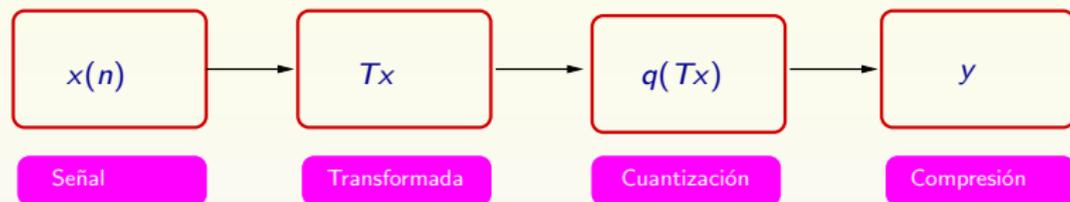
*DFT*



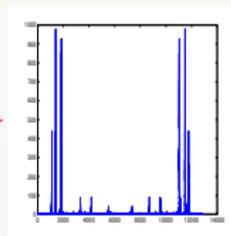
$K \ll N$   
 Coeficientes  
 significativos

$T$  revela la “estructura” de la señal

# Compresión



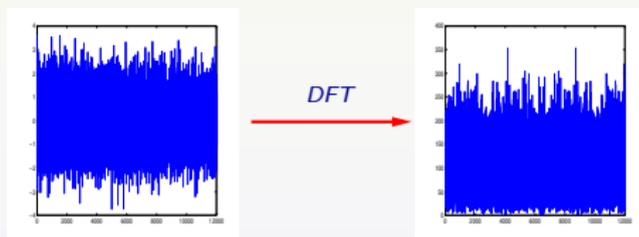
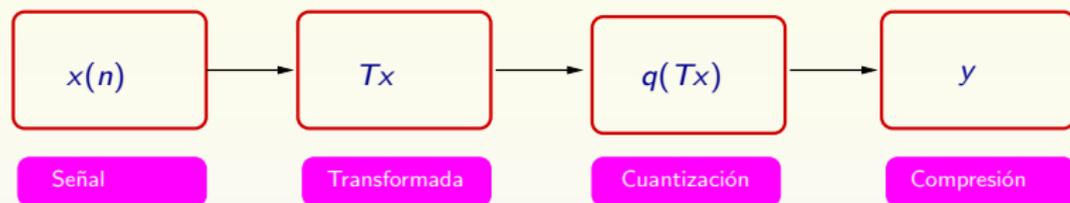
$DFT$



$K \ll N$   
 Coeficientes  
 significativos

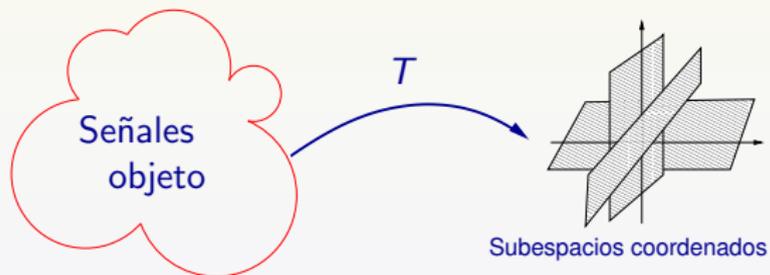
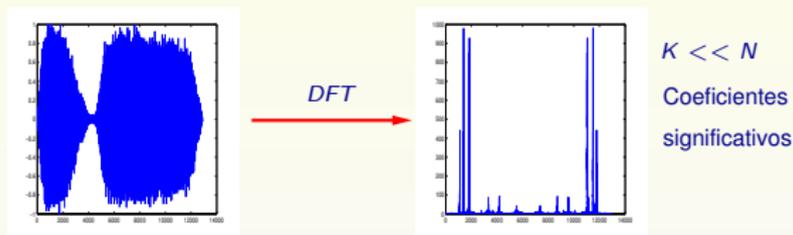
$T$  revela la “estructura” de la señal

# Compresión



$T$  revela la “estructura” de la señal

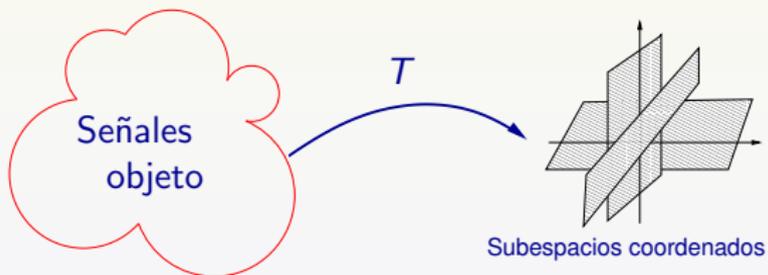
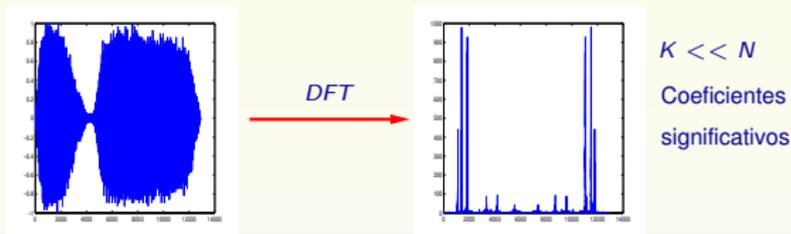
# Compresión. Transformada



Compressed Sensing

¿Podemos medir directamente la parte significativa de la señal?

# Compresión. Transformada



## Compressed Sensing

**¿Podemos medir directamente la parte significativa de la señal?**

# Compressed Sensing

## Adquisición de la señal (SENSING)

### Muestreo

- $\delta_j = [0, \dots, 0, 1, 0 \dots, 0]$   
 $\langle \delta_j, \mathbf{x} \rangle = \mathbf{x}[j]$

# Compressed Sensing

## Adquisición de la señal (SENSING)

### Muestreo

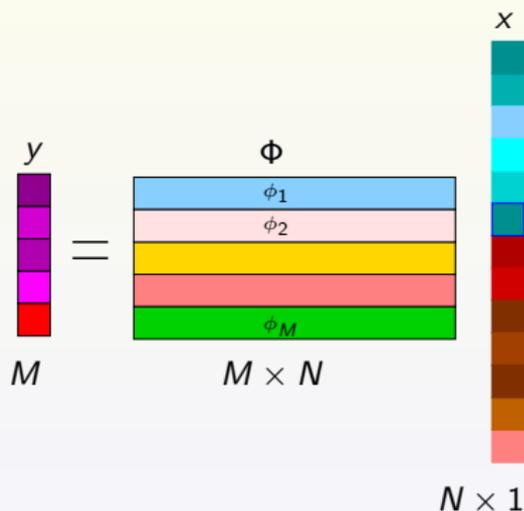
- $\delta_j = [0, \dots, 0, 1, 0 \dots, 0]$   
 $\langle \delta_j, \mathbf{x} \rangle = \mathbf{x}[j]$

### Medidas generalizadas

- $\phi = [\phi_1, \phi_2, \dots, \phi_n]$   
 $\langle \phi, \mathbf{x} \rangle = m$

# Compressed Sensing

Adquisición de la señal (SENSING)



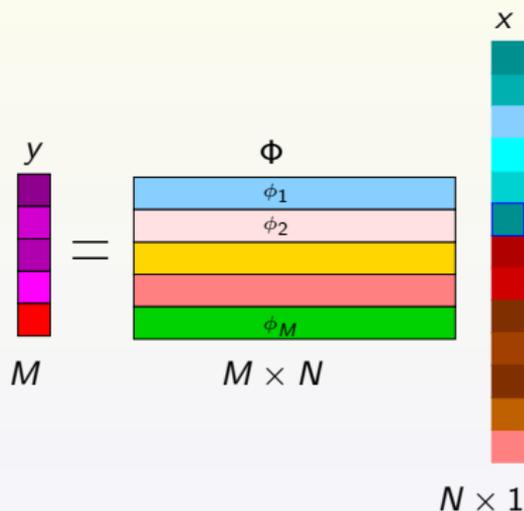
- $y_k = \langle \phi_k, x \rangle$ ,  $M \ll N$ .
- $\Phi$  matriz de medidas
- El vector dato es  $y$

## Cuestiones

- ¿Cómo reconstruir la señal a partir de  $y$ ?
- ¿ $M$ ? ¿ $\Phi$ ?

# Compressed Sensing

Adquisición de la señal (SENSING)



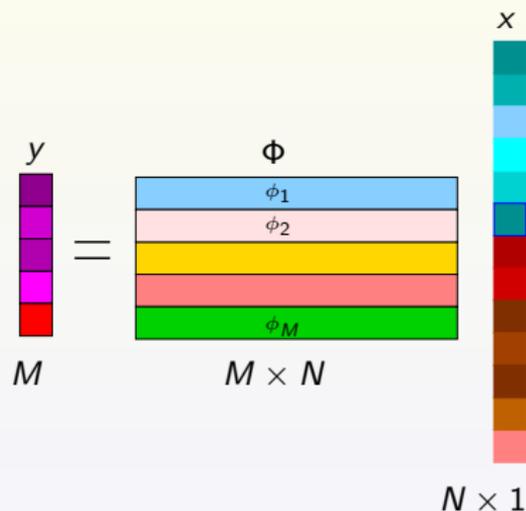
- $y_k = \langle \phi_k, x \rangle$ ,  $M \ll N$ .
- $\Phi$  matriz de medidas
- El vector dato es  $y$

## Cuestiones

- ¿Cómo reconstruir la señal a partir de  $y$ ?
- ¿ $M$ ?, ¿ $\Phi$ ?

# Compressed Sensing

Adquisición de la señal (SENSING)

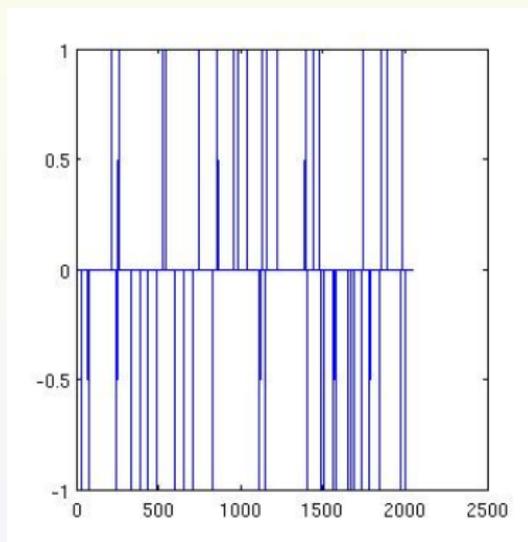


- $y_k = \langle \phi_k, x \rangle$ ,  $M \ll N$ .
- $\Phi$  matriz de medidas
- El vector dato es  $y$

## Cuestiones

- ¿Cómo reconstruir la señal a partir de  $y$ ?
- ¿ $M$ ?, ¿ $\Phi$ ?

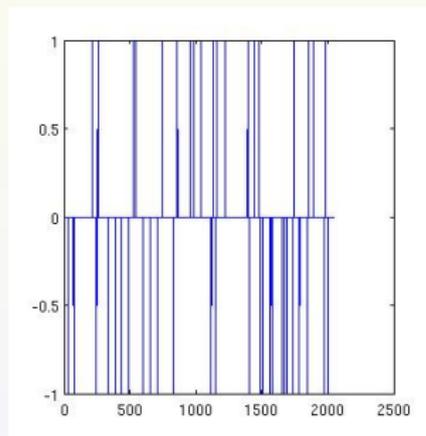
# Ejemplo



# Ejemplo: recuperación señal discreta dispersa

## Construcción de la señal

```
% longitud de la señal  
N = 2048;  
% Señal K-dispersa  
K = 50;  
  
% señal aleatoria +/- 1  
x = zeros(N,1);  
q = randperm(N);  
x(q(1:K)) = sign(randn(K,1));
```



# Número mínimo de medidas

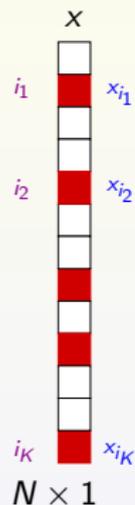


- $x$  es  $K$ -disperso

$$\|x\|_0 = |\{j \mid |x_j| \neq 0\}| \leq K$$

- $K + K = 2K$  incógnitas
- $M \geq 2K$

# Número mínimo de medidas

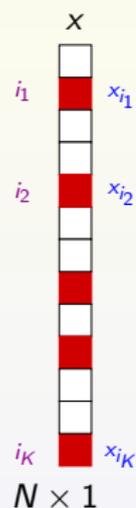


- $x$  es  $K$ -disperso

$$\|x\|_0 = |\{j \mid |x_j| \neq 0\}| \leq K$$

- $K + K = 2K$  incógnitas
- $M \geq 2K$

# Número mínimo de medidas



- $x$  es  $K$ -disperso

$$\|x\|_0 = |\{j \mid |x_j| \neq 0\}| \leq K$$

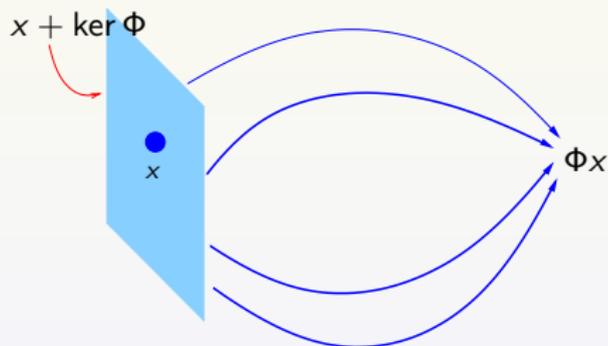
- $K + K = 2K$  incógnitas
- $M \geq 2K$

# Condición necesaria sobre la matriz de medidas

- $x \in \Sigma_K = \{u \in \mathbb{R}^N \mid \|u\|_0 \leq K\}$ ,  $x$  es  $K$ -dispersa
- $\Phi$  matriz de medidas.  $\Phi x = y$  vector información

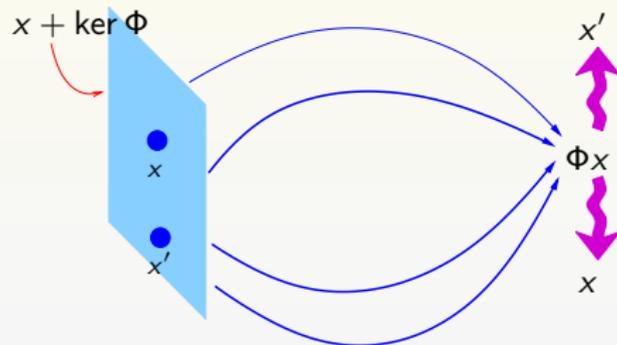
# Condición necesaria sobre la matriz de medidas

- $x \in \Sigma_K = \{u \in \mathbb{R}^N \mid \|u\|_0 \leq K\}$ ,  $x$  es  $K$ -dispersa
- $\Phi$  matriz de medidas.  $\Phi x = y$  vector información



# Condición necesaria sobre la matriz de medidas

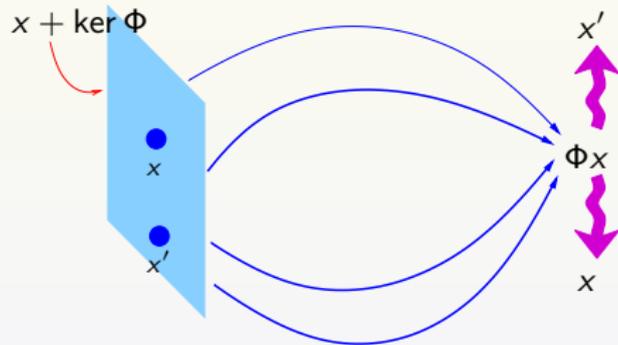
- $x', x \in \Sigma_K = \{u \in \mathbb{R}^N \mid \|u\|_0 \leq K\}$ ,  $x', x$  son  $K$ -dispersas
- $\Phi$  matriz de medidas.  $\Phi x = y = \Phi x'$



$x' - x \in \ker \Phi \cap \Sigma_{2K}$   
**No se puede reconstruir  $x$**

# Condición necesaria sobre la matriz de medidas

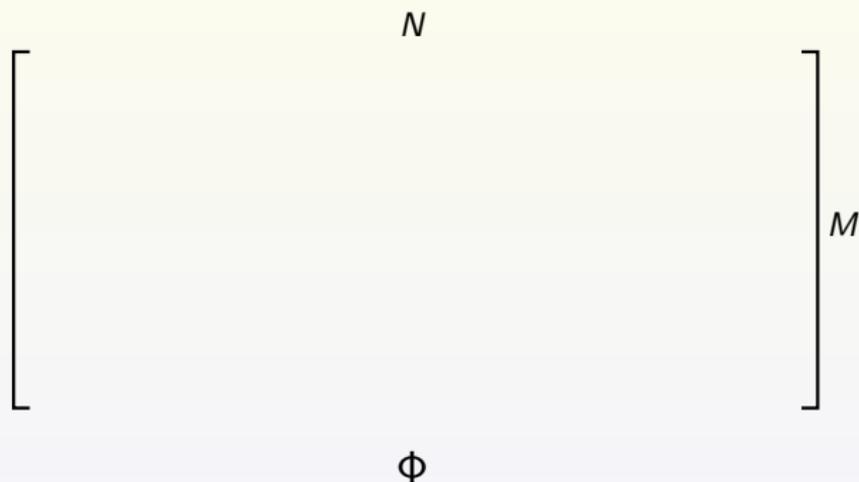
- $x \in \Sigma_K = \{u \in \mathbb{R}^N \mid \|u\|_0 \leq K\}$ ,  $x$  son  $K$ -dispersa
- $\Phi$  matriz de medidas.  $\Phi x = y = \Phi x'$



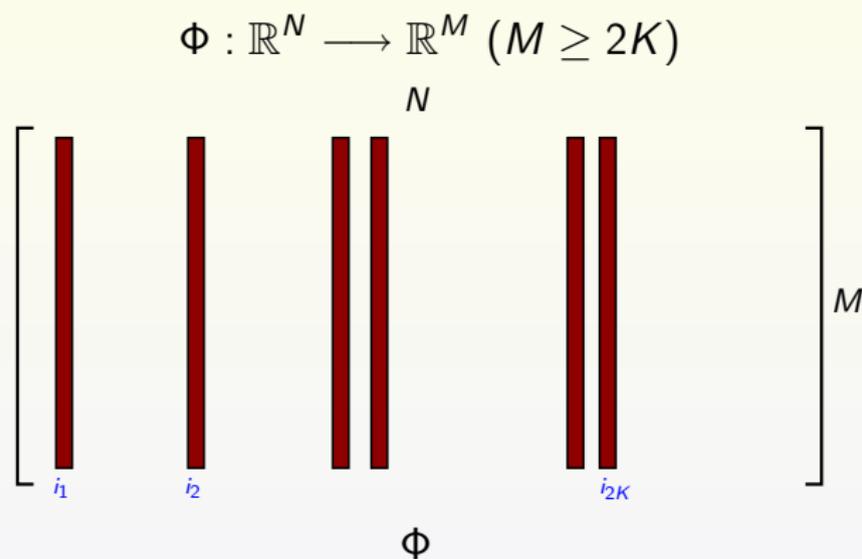
Reconstrucción  $\implies \ker \Phi \cap \Sigma_{2K} = \{0\}$

# Condición necesaria sobre la matriz de medidas

$$\Phi : \mathbb{R}^N \longrightarrow \mathbb{R}^M \quad (M \geq 2K)$$



## Condición necesaria sobre la matriz de medidas

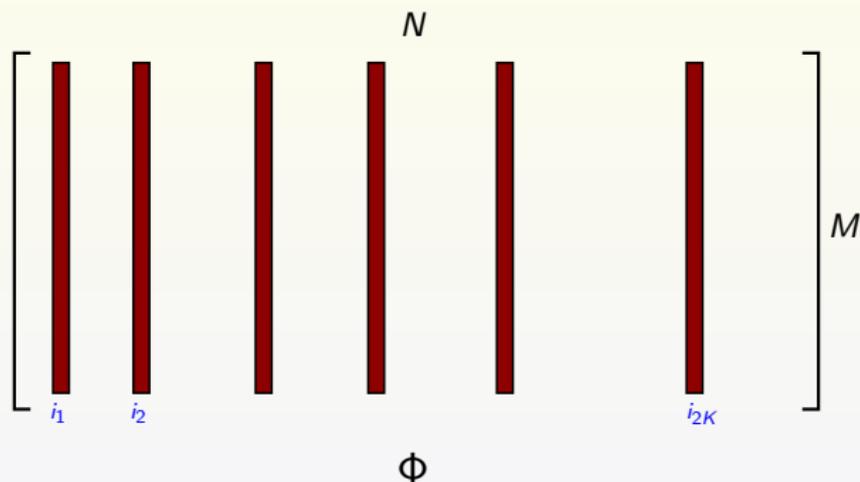


Condición necesaria

 $\ker \Phi \cap \Sigma_{2K} = \{0\} \iff$  todo conjunto de  $2K$  columnas es l.i.

## Condición necesaria sobre la matriz de medidas

$$\Phi : \mathbb{R}^N \longrightarrow \mathbb{R}^M \quad (M \geq 2K)$$

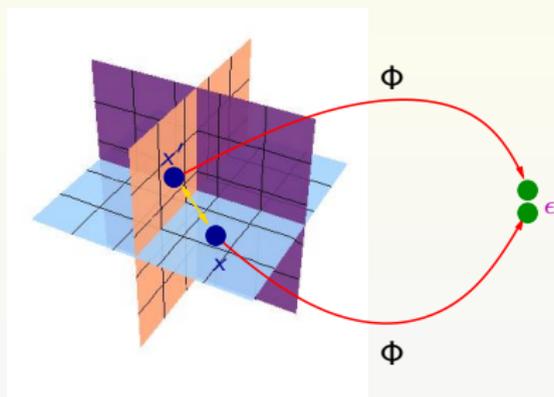


Condición necesaria

$\ker \Phi \cap \Sigma_{2K} = \{0\} \iff$  todo conjunto de  $2K$  columnas es l.i.

# Condición necesaria sobre la matriz de medidas

$$\Phi : \mathbb{R}^N \longrightarrow \mathbb{R}^M \quad (M \geq 2K)$$



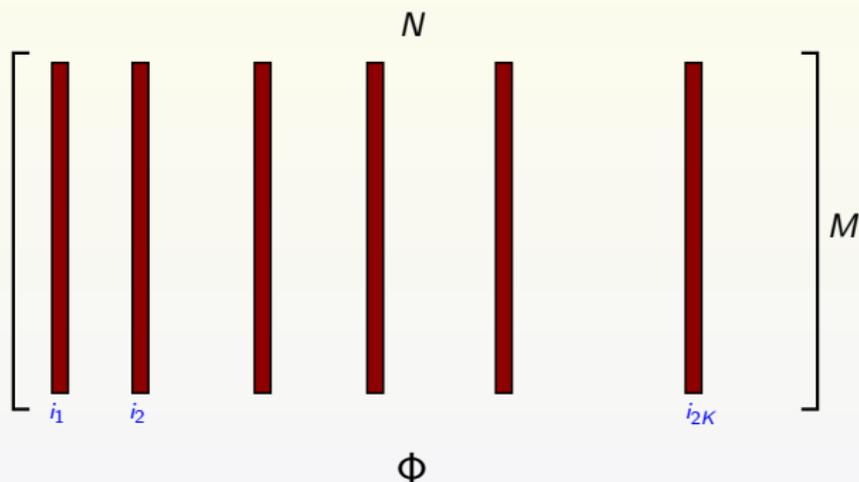
## Condición necesaria

$\ker \Phi \cap \Sigma_{2K} = \{0\} \iff$  todo conjunto de  $2K$  columnas es l.i.

**Para una reconstrucción efectiva no es suficiente**

## Condición necesaria sobre la matriz de medidas

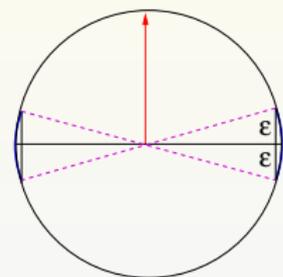
$$\Phi : \mathbb{R}^N \longrightarrow \mathbb{R}^M \quad (M \geq 2K)$$



Posible solución

**Todo conjunto de  $2K$  columnas de  $\Phi$  sea “casi” ortonormal**

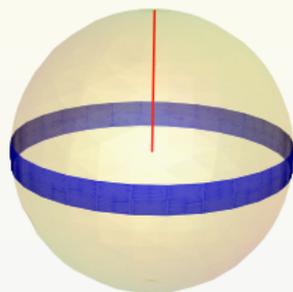
## Esfera M-dimensional



- Sea  $\mathbf{v} = [0, \dots, 0, 1] \in \mathbb{R}^M$
- $f(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle$  para  $\mathbf{w} \in \mathcal{S}^{M-1}$  (continua)

$$P(|f(\mathbf{w})| < \epsilon)?$$

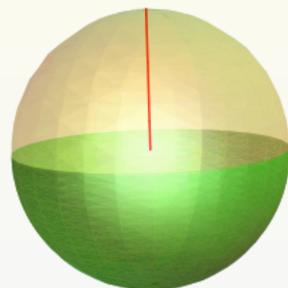
# Esfera M-dimensional



- Sea  $\mathbf{v} = [0, \dots, 0, 1] \in \mathbb{R}^M$
- $f(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle$  para  $\mathbf{w} \in \mathcal{S}^{M-1}$  (continua)

$$P(|f(\mathbf{w})| < \epsilon) = \frac{\text{Área}_\epsilon}{\text{Área Total}}$$

# Esfera M-dimensional

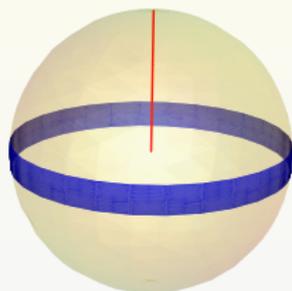


- Sea  $\mathbf{v} = [0, \dots, 0, 1] \in \mathbb{R}^M$
- $f(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle$  para  $\mathbf{w} \in \mathcal{S}^{M-1}$  (continua)

Ap. desigualdad de Brunn-Minkowsky

- $P(f(\mathbf{w}) > 0) = P(f(\mathbf{w}) < 0) = \frac{1}{2}$

# Esfera M-dimensional



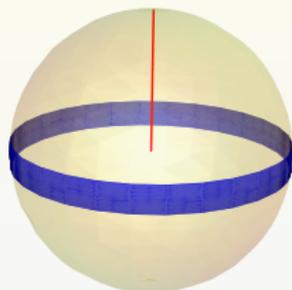
- Sea  $\mathbf{v} = [0, \dots, 0, 1] \in \mathbb{R}^M$
- $f(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle$  para  $\mathbf{w} \in \mathcal{S}^{M-1}$  (continua)

Ap. desigualdad de Brunn-Minkowsky

- $P(f(\mathbf{w}) > 0) = P(f(\mathbf{w}) < 0) = \frac{1}{2}$

Entonces  $P(|f(\mathbf{w})| < \epsilon) \approx 1 - 2e^{-M\epsilon^2/4}$

## Esfera M-dimensional



- Sea  $\mathbf{v} = [0, \dots, 0, 1] \in \mathbb{R}^M$
- $f(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle$  para  $\mathbf{w} \in \mathcal{S}^{M-1}$  (continua)

Ap. desigualdad de Brunn-Minkowsky

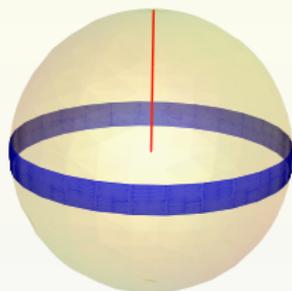
- $P(f(\mathbf{w}) > 0) = P(f(\mathbf{w}) < 0) = \frac{1}{2}$

Entonces  $P(|f(\mathbf{w})| < \epsilon) \approx 1 - 2e^{-M\epsilon^2/4}$

Si  $\epsilon = \frac{\delta}{\sqrt{M}}$  ( $\delta > 1$ ),

$$P(|f(\mathbf{w})| < \epsilon) \approx 1 - 2e^{-\delta^2/4}$$

# Esfera M-dimensional



- Sea  $\mathbf{v} = [0, \dots, 0, 1] \in \mathbb{R}^M$
- $f(\mathbf{w}) = \langle \mathbf{v}, \mathbf{w} \rangle$  para  $\mathbf{w} \in \mathcal{S}^{M-1}$  (continua)

Ap. desigualdad de Brunn-Minkowsky

- $P(f(\mathbf{w}) > 0) = P(f(\mathbf{w}) < 0) = \frac{1}{2}$

Entonces  $P(|f(\mathbf{w})| < \epsilon) \approx 1 - 2e^{-M\epsilon^2/4}$

**Utilizar el azar para que todo conjunto de  $2K$  columnas de  $\Phi$  sea “casi” ortonormal**

# Ejemplo: recuperación señal discreta dispersa

## Medidas

- Matriz de medidas

```
% Numero de medidas
```

```
M = 300;
```

```
% A matriz gaussiana
```

```
A = randn(M,N);
```

```
% Ortogonalización
```

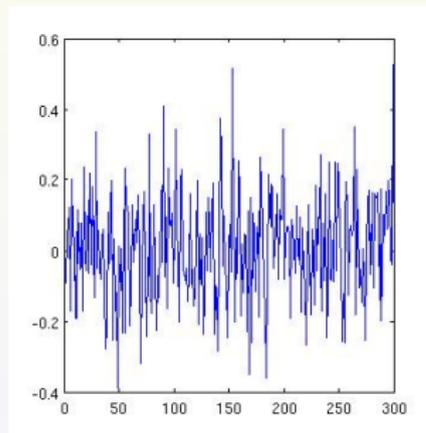
```
% filas de A
```

```
A = orth(A')';
```

- Vector de información

```
% medidas
```

```
y = A*x;
```

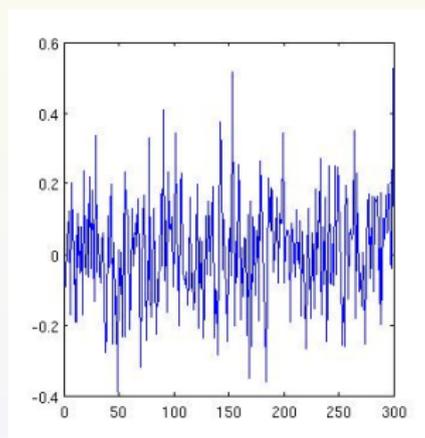


# Ejemplo: recuperación señal discreta dispersa

## Reconstrucción

### Datos

- $y = \Phi x$  vector información
- $x$  es  $K$ -dispersa

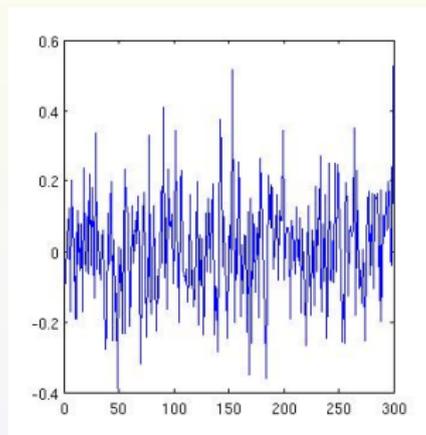


# Ejemplo: recuperación señal discreta dispersa

## Reconstrucción

### Datos

- $y = \Phi x$  vector información
- $x$  es  $K$ -dispersa



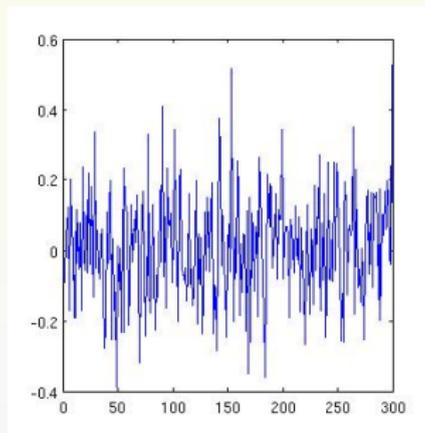
Si  $\ker \Phi \cap \Sigma_{2K} = \{0\}$   $x$  es la única señal  $K$ -dispersa tal que  $\Phi x = y$

# Ejemplo: recuperación señal discreta dispersa

## Reconstrucción

### Datos

- $y = \Phi x$  vector información
- $x$  es  $K$ -dispersa



Si  $\ker \Phi \cap \Sigma_{2K} = \{0\}$   $x$  es la única señal  $K$ -dispersa tal que  $\Phi x = y$

### Reconstrucción

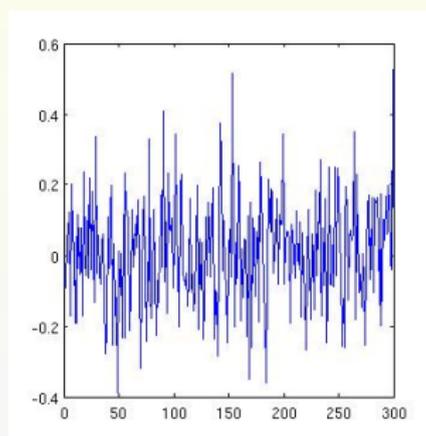
$$\operatorname{argmin}_{z \in \mathbb{R}^N} \|z\|_0 \text{ tal que } \Phi z = y$$

# Ejemplo: recuperación señal discreta dispersa

## Reconstrucción

### Datos

- $y = \Phi x$  vector información
- $x$  es  $K$ -dispersa



Si  $\ker \Phi \cap \Sigma_{2K} = \{0\}$   $x$  es la única señal  $K$ -dispersa tal que  $\Phi x = y$

### Reconstrucción

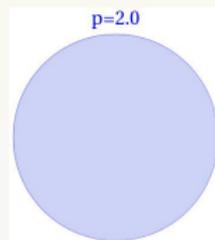
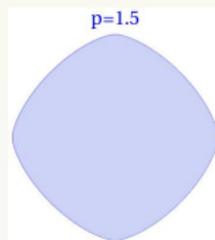
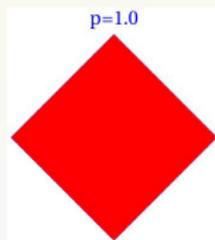
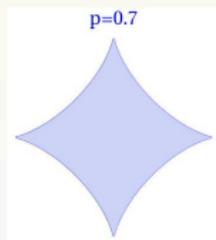
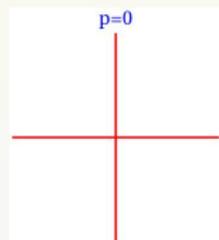
$$\operatorname{argmin}_{z \in \mathbb{R}^N} \|z\|_0 \text{ tal que } \Phi z = y \quad (*)$$

**¡El problema (\*) es NP-hard!**

# Ejemplo: recuperación señal discreta dispersa

Normas  $\|\cdot\|_p$

$$\|x\|_p^p = |x_1|^p + \dots + |x_N|^p \leq 1$$

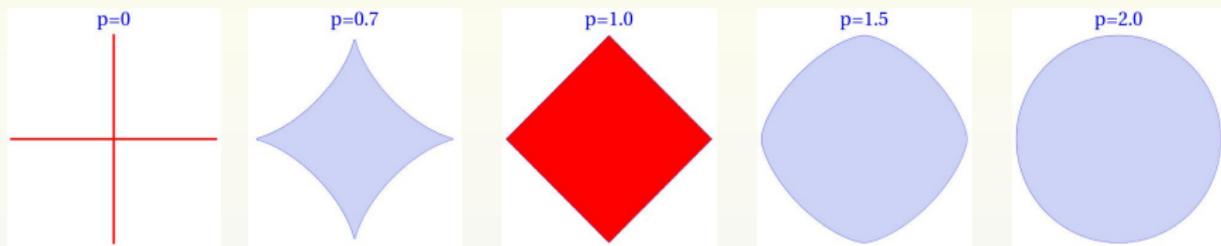


- Las normas  $\|\cdot\|_p$  con  $0 \leq p \leq 1$  “favorecen” soluciones dispersas

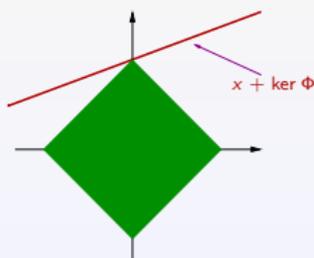
## Ejemplo: recuperación señal discreta dispersa

Normas  $\|\cdot\|_p$ 

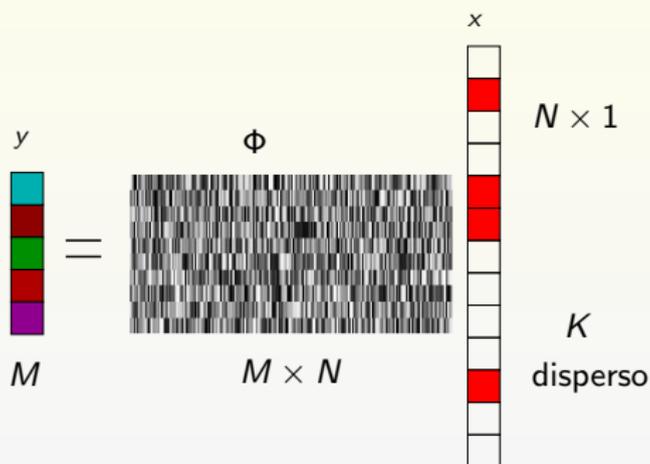
$$\|x\|_p^p = |x_1|^p + \dots + |x_N|^p \leq 1$$



- Las normas  $\|\cdot\|_p$  con  $0 \leq p \leq 1$  “favorecen” soluciones dispersas



## Ejemplo: recuperación señal discreta dispersa

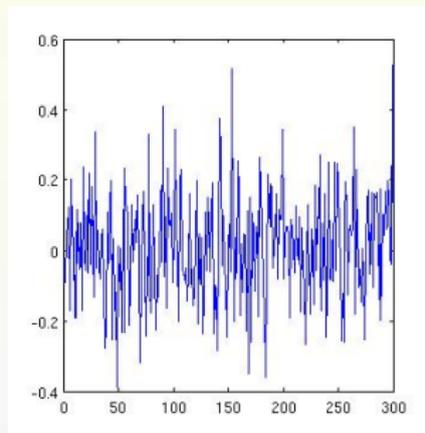
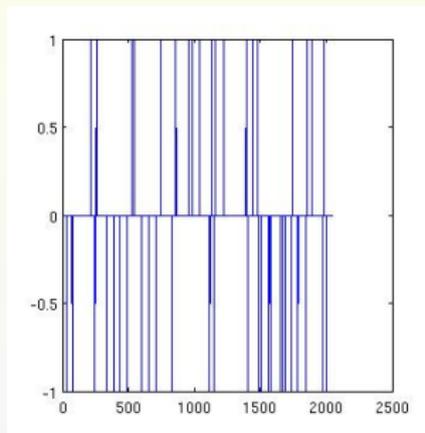
Minimización  $\ell^1$ 

$$x \in \Sigma_K$$

$$\tilde{x} = \operatorname{argmin}_{z \in \mathbb{R}^N} \|z\|_1 \text{ tal que } \Phi z = \Phi x$$

# Ejemplo: recuperación señal discreta dispersa

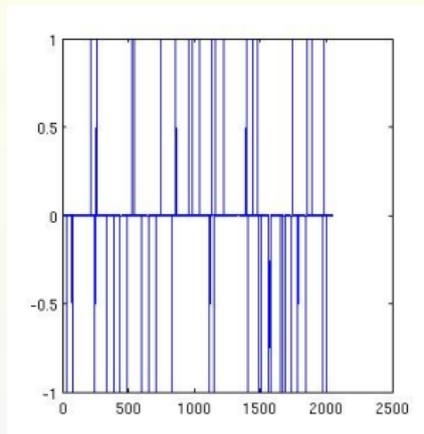
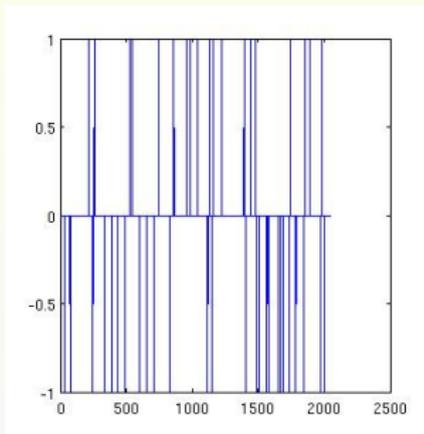
Minimización  $\ell^1$



$$x \in \Sigma_K$$

$$\tilde{x} = \operatorname{argmin}_{z \in \mathbb{R}^M} \|z\|_1 \text{ tal que } \Phi z = \Phi x$$

## Ejemplo: recuperación señal discreta dispersa

Minimización  $\ell^1$ 

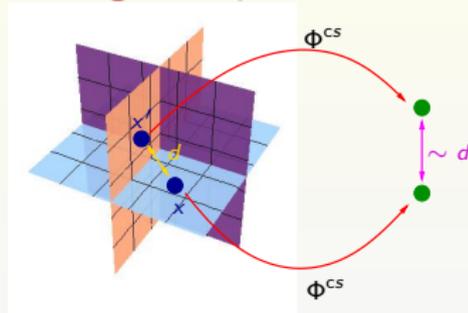
$$x \in \Sigma_K$$

$$\tilde{x} = \operatorname{argmin}_{z \in \mathbb{R}^N} \|z\|_1 \text{ tal que } \Phi z = \Phi x$$

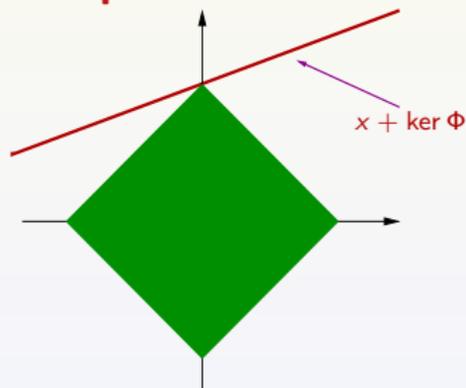
$$\tilde{x} = x$$

# Un poco de Teoría

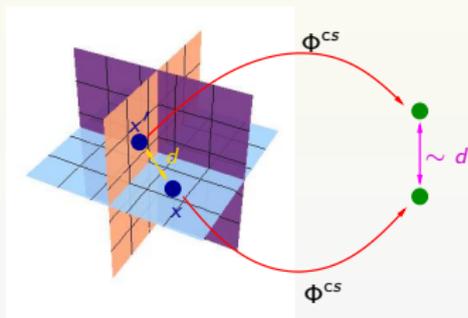
## Sensing. Propiedad RIP



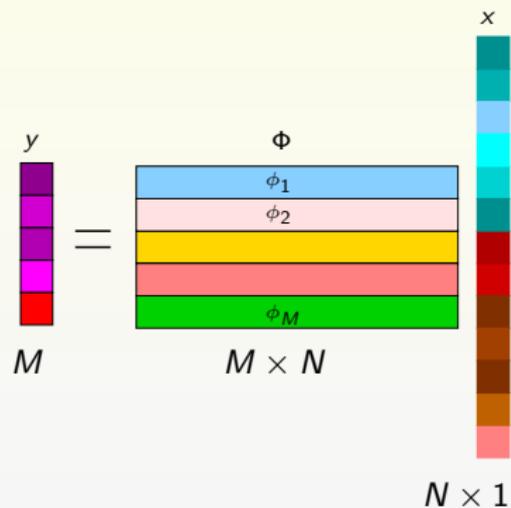
## Aproximación



## Sensing. Propiedad RIP



# Medidas. Propiedad RIP



## Objetivo

**Contruir matrices  $\Phi$  capaces de codificar la información del mayor conjunto posible de señales**

# Medidas. Propiedad RIP

- $\Psi = \{\Psi_i\}_{i=1}^N$  base ortonormal de  $\mathbb{R}^N$
- Para todo vector  $x \in \mathbb{R}^N$  se tiene

$$x = \sum_{i=1}^N \theta_i \Psi_i, \quad \theta_i = \langle x, \Psi_i \rangle$$



# Medidas. Propiedad RIP

- $\Psi = \{\Psi_i\}_{i=1}^N$  base ortonormal de  $\mathbb{R}^N$
- $\Psi$  no tiene necesariamente que ser una base ortonormal
  - Base,
  - *Frame*
- Para todo vector  $x \in \mathbb{R}^N$  se tiene

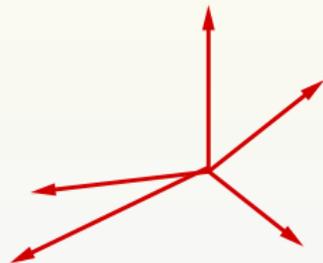
$$x = \sum_{i=1}^N \theta_i \Psi_i, \quad \theta_i = \langle x, \Psi_i \rangle$$



# Medidas. Propiedad RIP

- $\Psi = \{\psi_i\}_{i=1}^N$  base ortonormal de  $\mathbb{R}^N$
- $\Psi$  no tiene necesariamente que ser una base ortonormal
  - Base,
  - *Frame*
- Para todo vector  $x \in \mathbb{R}^N$  se tiene

$$x = \sum_{i=1}^N \theta_i \psi_i, \quad \theta_i = \langle x, \psi_i \rangle$$



# Medidas. Propiedad RIP

- $\Psi = \{\Psi_i\}_{i=1}^N$  base ortonormal de  $\mathbb{R}^N$
- Para todo vector  $x \in \mathbb{R}^N$  se tiene

$$x = \sum_{i=1}^N \theta_i \Psi_i, \quad \theta_i = \langle x, \Psi_i \rangle$$

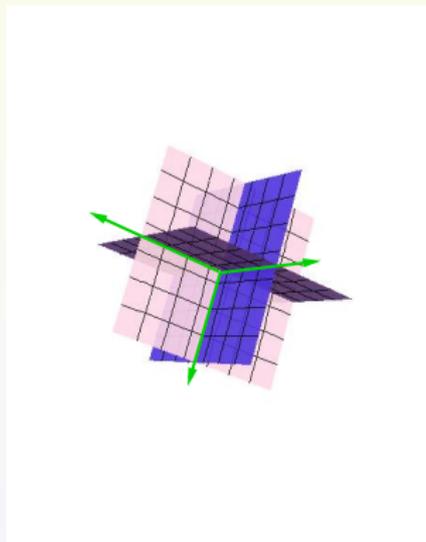
# Medidas. Propiedad RIP

- $\Psi = \{\Psi_i\}_{i=1}^N$  base ortonormal de  $\mathbb{R}^N$
- Para todo vector  $x \in \mathbb{R}^N$  se tiene

$$x = \sum_{i=1}^N \theta_i \Psi_i, \quad \theta_i = \langle x, \Psi_i \rangle$$

- $\Sigma_K(\Psi)$  señales  $K$ -dispersas respecto de  $\Psi$
- Si  $x \in \Sigma_K(\Psi)$ ,

$$[\Psi_1 \quad \Psi_2 \quad \cdots \quad \Psi_N] \begin{bmatrix} 0 \\ \theta_{i_1} \\ \vdots \\ \theta_{i_K} \\ 0 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$



# Medidas. Propiedad RIP

- Supongamos que disponemos de  $M < N$  medidas del vector  $x \in \mathbb{R}^N$

$$\langle \phi_i, x \rangle, i = 1, 2, \dots, M.$$

- Matricialmente lo anterior se puede expresar como

$$y = \Phi x$$

- $\Phi$  matriz de medidas,  $\Psi$  base de dispersión

$$\Phi^{CS} = \Phi \Psi \text{ matriz CS}$$

# Medidas. Propiedad RIP

- Supongamos que disponemos de  $M < N$  medidas del vector  $x \in \mathbb{R}^N$

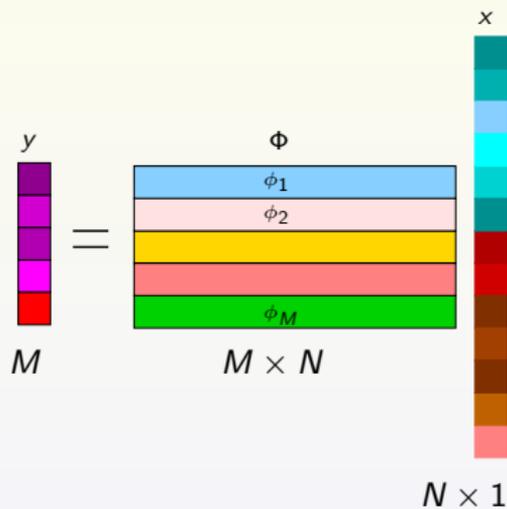
$$\langle \phi_i, x \rangle, i = 1, 2, \dots, M.$$

- Matricialmente lo anterior se puede expresar como

$$y = \Phi x$$

- $\Phi$  matriz de medidas,  $\Psi$  base de dispersión

$$\Phi^{CS} = \Phi \Psi \text{ matriz CS}$$



# Medidas. Propiedad RIP

- Supongamos que disponemos de  $M < N$  medidas del vector  $x \in \mathbb{R}^N$

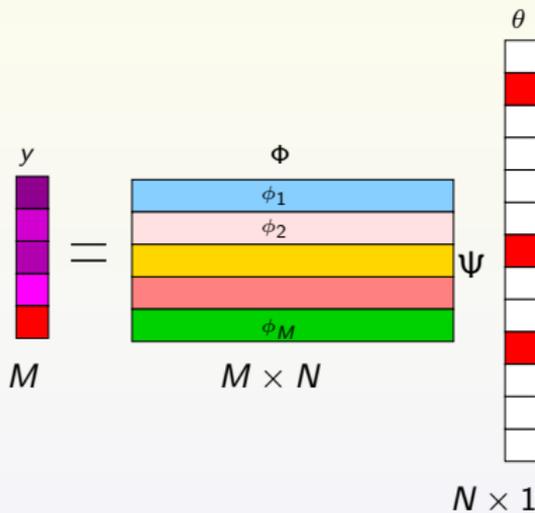
$$\langle \phi_i, x \rangle, i = 1, 2, \dots, M.$$

- Matricialmente lo anterior se puede expresar como

$$y = \Phi x = \Phi \Psi \theta = \Phi^{CS} \theta$$

- $\Phi$  matriz de medidas,  $\Psi$  base de dispersión

$$\Phi^{CS} = \Phi \Psi \text{ matriz CS}$$



# Propiedad RIP. Isometrías restringidas

Sea  $\Phi^{CS} \in \mathbb{C}^{M \times N}$ . (Matriz CS,  $M \ll N$ )

- Para  $T \subset \{1, 2, \dots, N\}$  con  $|T| = S$   $\Phi_T^{CS}$  es la sección de  $\Phi^{CS}$  formada por las columnas  $T$ . Si  $T = \{i_1, i_2, \dots, i_S\}$

$$\left[ \begin{array}{c|c|c|c|c|} \color{red}{\rule{0.5em}{0.5em}} & \color{red}{\rule{0.5em}{0.5em}} & \color{red}{\rule{0.5em}{0.5em}} & \color{red}{\rule{0.5em}{0.5em}} & \color{red}{\rule{0.5em}{0.5em}} \\ \color{blue}{i_1} & \color{blue}{i_2} & & & \color{blue}{i_S} \end{array} \right]_M^N$$

$\Phi^{CS}$

- Para cada entero  $1 \leq S \leq M$  sea  $\delta_S$  el número más pequeño que verifica

$$(1 - \delta_S) \leq \frac{\|\Phi_T^{CS} z\|_2}{\|z\|_2} \leq (1 + \delta_S), \quad z \in \mathbb{R}^S$$

para cada subconjunto  $T \subset \{1, 2, \dots, N\}$  de cardinalidad  $S$ .

- Si  $0 < \delta_S < 1$

$\Phi^{CS}$  es una **isometría restringida de orden  $S$** .

# Propiedad RIP. Isometrías restringidas

Sea  $\Phi^{CS} \in \mathbb{C}^{M \times N}$ . (Matriz CS,  $M \ll N$ )

- Para  $T \subset \{1, 2, \dots, N\}$  con  $|T| = S$   $\Phi_T^{CS}$  es la sección de  $\Phi^{CS}$  formada por las columnas  $T$ .
- Para cada entero  $1 \leq S \leq M$  sea  $\delta_S$  el número más pequeño que verifica

$$(1 - \delta_S) \leq \frac{\|\Phi_T^{CS} z\|_2}{\|z\|_2} \leq (1 + \delta_S), \quad z \in \mathbb{R}^S$$

para cada subconjunto  $T \subset \{1, 2, \dots, N\}$  de cardinalidad  $S$ .

- Si  $0 < \delta_S < 1$

$\Phi^{CS}$  es una **isometría restringida de orden  $S$** .

# Propiedad RIP. Isometrías restringidas

Sea  $\Phi^{CS} \in \mathbb{C}^{M \times N}$ . (Matriz CS,  $M \ll N$ )

- Para  $T \subset \{1, 2, \dots, N\}$  con  $|T| = S$   $\Phi_T^{CS}$  es la sección de  $\Phi^{CS}$  formada por las columnas  $T$ .
- Para cada entero  $1 \leq S \leq M$  sea  $\delta_S$  el número más pequeño que verifica

$$(1 - \delta_S) \leq \frac{\|\Phi_T^{CS} z\|_2}{\|z\|_2} \leq (1 + \delta_S), \quad z \in \mathbb{R}^S$$

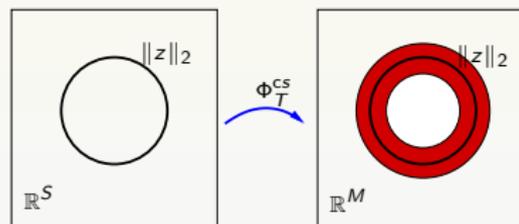
para cada subconjunto  $T \subset \{1, 2, \dots, N\}$  de cardinalidad  $S$ .

- Si  $0 < \delta_S < 1$   
 $\Phi^{CS}$  es una **isometría restringida de orden  $S$** .

# Propiedad RIP

## Geometría de la propiedad RIP

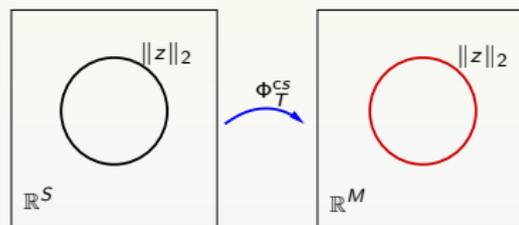
- Cuanto menor sea  $\delta_S$  más cerca estarán las secciones  $\Phi_T^{CS}$  de ser isometrías



# Propiedad RIP

## Geometría de la propiedad RIP

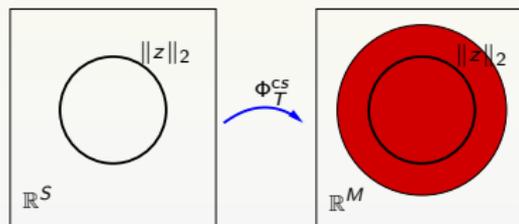
- Cuanto menor sea  $\delta_S$  más cerca estarán las secciones  $\Phi_T^{CS}$  de ser isometrías
- $\delta_S$  mide la “ortonormalidad” de las columnas de las secciones  $\Phi_T^{CS}$  con  $|T| = S$ 
  - Si  $\delta_S = 0$  todos los conjuntos de columnas de  $S$  elementos son ortonormales.
  - Si  $\delta_S = 1$  existe una submatriz tal que  $\text{rank } \Phi_T^{CS} < S$ .



# Propiedad RIP

## Geometría de la propiedad RIP

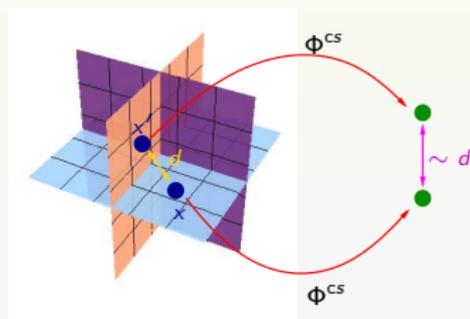
- Cuanto menor sea  $\delta_S$  más cerca estarán las secciones  $\Phi_T^{CS}$  de ser isometrías
- $\delta_S$  mide la “ortonormalidad” de las columnas de las secciones  $\Phi_T^{CS}$  con  $|T| = S$ 
  - Si  $\delta_S = 0$  todos los conjuntos de columnas de  $S$  elementos son ortonormales.
  - Si  $\delta_S = 1$  existe una submatriz tal que  $\text{rank } \Phi_T^{CS} < S$ .



# Propiedad RIP

## Geometría de la propiedad RIP

- Cuanto menor sea  $\delta_S$  más cerca estarán las secciones  $\Phi_T^{CS}$  de ser isometrías
- $\delta_S$  mide la “ortonormalidad” de las columnas de las secciones  $\Phi_T^{CS}$  con  $|T| = S$
- Una isometría restringida puede permitir una reconstrucción robusta



# CS y matrices RIP

$\Phi^{CS} \in \mathbb{R}^{M \times N}$  isometría restringida de orden  $S$

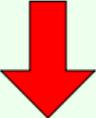
## Objetivos



## CS y matrices RIP

$\Phi^{CS} \in \mathbb{R}^{M \times N}$  isometría restringida de orden  $S$

## Objetivos

  $S$        $\delta_S$



  $K$

+ Señales codificadas

# Propiedad RIP

## Cálculo de $\delta_S$

$\forall T \subset \{1, 2, \dots, N\}$  con  $|T| = S$ .

$$\sigma((\Phi_T^{CS})^* \Phi_T^{CS}) \subset [(1 - \delta_S)^2, (1 + \delta_S)^2]$$

# Propiedad RIP

## Cálculo de $\delta_S$

$\forall T \subset \{1, 2, \dots, N\}$  con  $|T| = S$ .

$$\sigma((\Phi_T^{\text{CS}})^* \Phi_T^{\text{CS}}) \subset [(1 - \delta_S)^2, (1 + \delta_S)^2]$$

- Para calcular  $\delta_S$ : SVD de todas las submatrices  $\Phi_T^{\text{CS}}$  con  $|T| = S$ .

# Propiedad RIP

## Cálculo de $\delta_S$

$\forall T \subset \{1, 2, \dots, N\}$  con  $|T| = S$ .

$$\sigma((\Phi_T^{CS})^* \Phi_T^{CS}) \subset [(1 - \delta_S)^2, (1 + \delta_S)^2]$$

- Para calcular  $\delta_S$ : SVD de todas las submatrices  $\Phi_T^{CS}$  con  $|T| = S$ .

**Hay**  $\binom{N}{S} = \frac{N!}{S!(N-S)!}$  **matrices**  $\Phi_T^{CS}$  **con**  $|T| = S$

# Propiedad RIP

Cálculo de  $\delta_S$

- **No hay forma efectiva de comprobar si una matriz  $\Phi^{CS}$  verifica RIP**

# Propiedad RIP

## Cálculo de $\delta_S$

- No hay forma efectiva de comprobar si una matriz  $\Phi^{CS}$  verifica RIP
- No hay algoritmos deterministas que construyan matrices RIP con  $S$  grande (de Vore 07)

# Propiedad RIP

## Cálculo de $\delta_S$

- No hay forma efectiva de comprobar si una matriz  $\Phi^{CS}$  verifica RIP
- No hay algoritmos deterministas que construyan matrices RIP con  $S$  grande (de Vore 07)
- **Solución: hacer intervenir al azar**

# Matrices RIP

## Matrices gaussianas y de Rademacher

Si  $\Phi \in \mathbb{R}^{M \times N}$ ,

- **Matriz gaussiana**

$$\Phi_{ij} = X_{ij} \text{ i.i.d. } N(0, 1/M)$$

- **Matriz de Rademacher**

$$\Phi_{ij} = \frac{1}{\sqrt{N}} X_{ij}$$

$$X_{ij} \text{ i.i.d. } P(X_{ij} = \pm 1) = \frac{1}{2}$$



# Matrices RIP

## Matrices gaussianas y de Rademacher

Si  $\Phi \in \mathbb{R}^{M \times N}$ ,

- Matriz gaussiana

$$\Phi_{ij} = X_{ij} \text{ i.i.d. } N(0, 1/M)$$

- Matriz de Rademacher

$$\Phi_{ij} = \frac{1}{\sqrt{N}} X_{ij}$$

$$X_{ij} \text{ i.i.d. } P(X_{ij} = \pm 1) = \frac{1}{2}$$



# Matrices RIP

## Matrices gaussianas y de Rademacher

Si  $\Phi \in \mathbb{R}^{M \times N}$ ,

- Matriz gaussiana

$$\Phi_{ij} = X_{ij} \text{ i.i.d. } N(0, 1/M)$$

- Matriz de Rademacher

$$\Phi_{ij} = \frac{1}{\sqrt{N}} X_{ij}$$

$$X_{ij} \text{ i.i.d. } P(X_{ij} = \pm 1) = \frac{1}{2}$$

$$\forall \delta < 1 \exists c_1(\delta), c_2(\delta) > 0$$

$$\log(N/S)S \leq c_1 M$$

$$\downarrow$$

$$\delta_S(\Phi\Psi) < \delta$$

con probabilidad  $1 - 2e^{-c_2 M}$

# Matrices RIP

## Proyecciones aleatorias

- $\Theta = \{\theta_i\}_{i=1}^N$  base ortonormal
- $\{i_1, i_2, \dots, i_M\} \subset \{1, 2, \dots, N\}$   
elegido al azar con probabilidad  
uniforme en los subconjuntos de  
 $M$  elementos de  $\{1, 2, \dots, N\}$

$$\Phi = \begin{bmatrix} \theta_{i_1} \\ \theta_{i_2} \\ \vdots \\ \theta_{i_M} \end{bmatrix}$$

$$\Theta$$

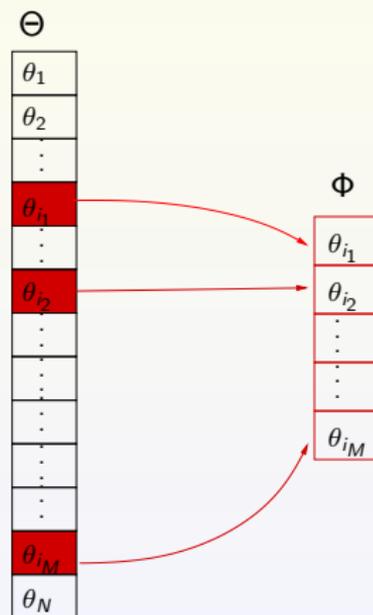
$\theta_1$
$\theta_2$
$\vdots$
$\theta_N$

# Matrices RIP

## Proyecciones aleatorias

- $\Theta = \{\theta_i\}_{i=1}^N$  base ortonormal
- $\{i_1, i_2, \dots, i_M\} \subset \{1, 2, \dots, N\}$  elegido al azar con probabilidad uniforme en los subconjuntos de  $M$  elementos de  $\{1, 2, \dots, N\}$

$$\Phi = \begin{bmatrix} \theta_{i_1} \\ \theta_{i_2} \\ \vdots \\ \theta_{i_M} \end{bmatrix}$$



# Matrices RIP

## Proyecciones aleatorias

- $\Theta = \{\theta_i\}_{i=1}^N$  base ortonormal
- $\{i_1, i_2, \dots, i_M\} \subset \{1, 2, \dots, N\}$  elegido al azar con probabilidad uniforme en los subconjuntos de  $M$  elementos de  $\{1, 2, \dots, N\}$

$$\Phi = \begin{bmatrix} \theta_{i_1} \\ \theta_{i_2} \\ \vdots \\ \theta_{i_M} \end{bmatrix}$$

$$\forall \delta < 1 \exists \beta(\delta) > 0$$

$$S \leq \frac{\beta M}{N \mu(\Theta \cup \Psi)^2 (\log(N))^5}$$

$$\Downarrow$$

$$\delta_S(\Phi \Psi) < \delta$$

con probabilidad  $1 - N^{-c}$

# Matrices RIP

## Proyecciones aleatorias

- $\Theta = \{\theta_i\}_{i=1}^N$  base ortonormal
- $\{i_1, i_2, \dots, i_M\} \subset \{1, 2, \dots, N\}$  elegido al azar con probabilidad uniforme en los subconjuntos de  $M$  elementos de  $\{1, 2, \dots, N\}$

$$\Phi = \begin{bmatrix} \theta_{i_1} \\ \theta_{i_2} \\ \vdots \\ \theta_{i_M} \end{bmatrix}$$

$$\forall \delta < 1 \exists \beta(\delta) > 0$$

$$S \leq \frac{\beta M}{N \mu(\Theta \cup \Psi)^2 (\log(N))^5}$$

$$\downarrow$$

$$\delta_S(\Phi \Psi) < \delta$$

con probabilidad  $1 - N^{-c}$

$$\mu(\Theta \cup \Psi) = \max_{1 \leq k, j \leq N} |\langle \theta_k, \psi_j \rangle|$$

$$\mu(\Theta \cup \Psi) \in \left[ \frac{1}{\sqrt{N}}, 1 \right]$$

# Matrices RIP

## Proyecciones aleatorias

### Ejemplo

- $\mathcal{F} \equiv$  base de Fourier
- $\Psi = Id_N$

$$\mu(\mathcal{F} \cup \Psi) = \frac{1}{\sqrt{N}}$$

- $\Phi$  proyección aleatoria obtenida de  $\mathcal{F}$  escogiendo al azar  $M$  filas.

$$\forall \delta < 1 \exists \beta(\delta) > 0$$

$$S \leq \frac{\beta M}{(\log(N))^5}$$

$$\downarrow$$

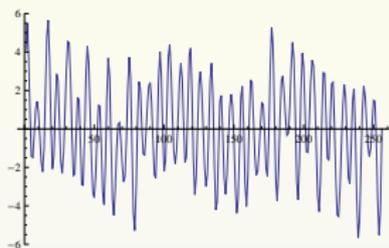
$$\delta_S(\Phi) < \delta$$

con probabilidad  $1 - N^{-c}$

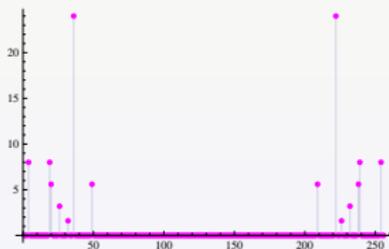
# Matrices RIP

## Proyecciones aleatorias

$$\mathbf{x} \in \mathbb{R}^{256}$$



$$\text{DFT}(\mathbf{x})$$

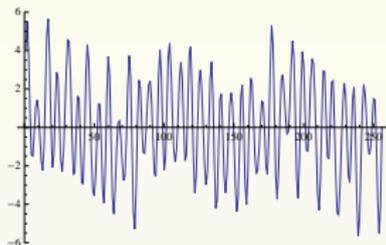


$\mathbf{x} \in \mathbb{R}^{256}$  suma de 7 sinusoides

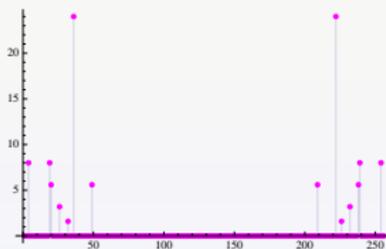
# Matrices RIP

## Proyecciones aleatorias

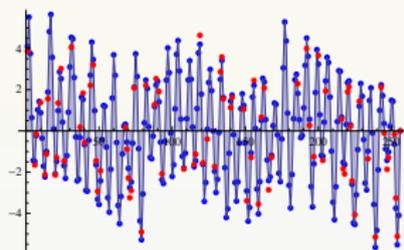
$$\mathbf{x} \in \mathbb{R}^{256}$$



$$\text{DFT}(\mathbf{x})$$



$$\mathbf{y} \in \mathbb{R}^{80}$$

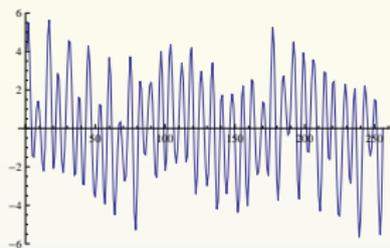


$$\mathbf{y} \in \mathbb{R}^{80} \text{ (80 muestras aleatorias de } \mathbf{x}\text{)}$$

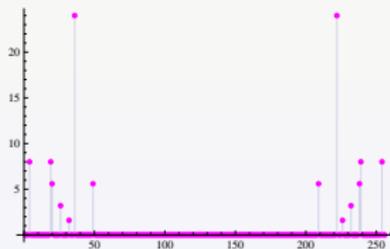
# Matrices RIP

## Proyecciones aleatorias

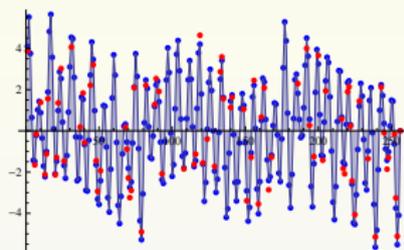
$\mathbf{x} \in \mathbb{R}^{256}$



DFT( $\mathbf{x}$ )



$\mathbf{y} \in \mathbb{R}^{80}$

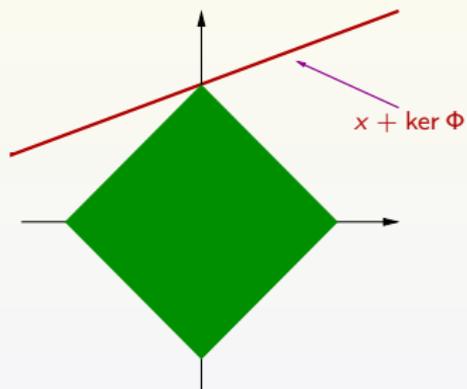


Reconstrucción



$$\theta = \operatorname{argmin} \|\mathbf{z}\|_1 \quad \mathbf{x}[\mathbf{ij}] = \mathbf{z}[\mathbf{ij}], \quad \theta = \mathbf{x}$$

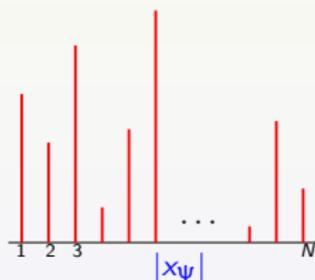
# Aproximación



# Aproximación por $K$ términos

$x \in \mathbb{R}^N$ ,  $\Psi = \{\Psi_i\}_{i=1}^N$  base ortonormal de  $\mathbb{R}^N$

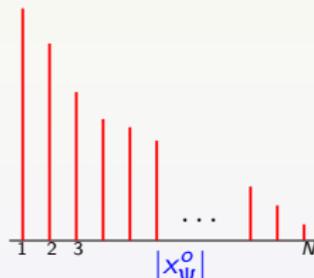
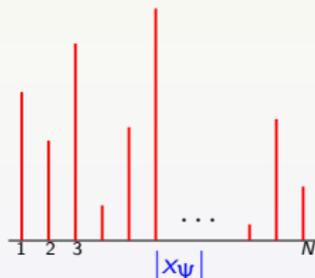
$$x = \sum_{i=1}^N x_{\Psi}[i] \Psi_i$$



# Aproximación por $K$ términos

$x \in \mathbb{R}^N$ ,  $\Psi = \{\Psi_i\}_{i=1}^N$  base ortonormal de  $\mathbb{R}^N$

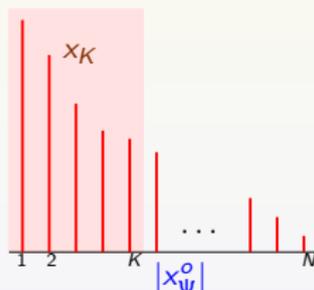
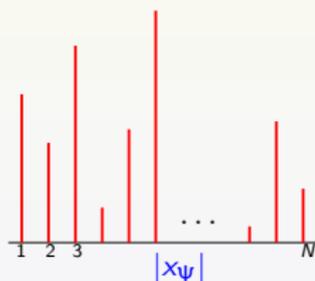
$$x = \sum_{i=1}^N x_{\Psi}[i] \Psi_i$$



# Aproximación por $K$ términos

$x \in \mathbb{R}^N$ ,  $\Psi = \{\Psi_i\}_{i=1}^N$  base ortonormal de  $\mathbb{R}^N$

$$x = \sum_{i=1}^N x_{\Psi}[i] \Psi_i$$

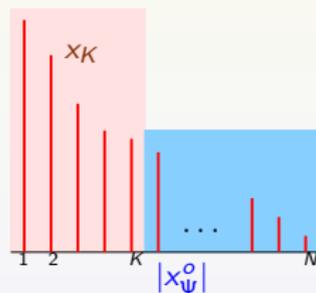
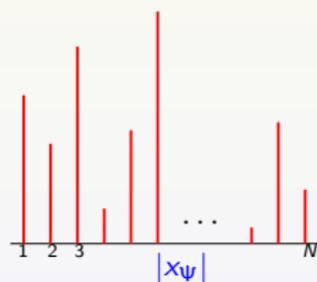


$$x_K = \sum_{k=1}^K x_{\Psi}^o[k] \Psi_{i_k}$$

# Aproximación por $K$ términos

$x \in \mathbb{R}^N$ ,  $\Psi = \{\Psi_i\}_{i=1}^N$  base ortonormal de  $\mathbb{R}^N$

$$x = \sum_{i=1}^N x_{\Psi}[i] \Psi_i$$



$$\|x - x_K\|^2 \leq \sum_{k>K} |x_{\Psi}^o[k]|^2$$

# Aproximación

## Teorema Fundamental

$$Y = \Phi X + W$$

### Teorema

Sea  $\Phi^{cs} = \Phi\Psi$ . Supongamos que  $K$  es tal que  $\delta_{3K}(\Phi^{cs}) \leq \delta < 1/3$ . Sea  $X = \Psi\theta$  donde

$$\theta = \operatorname{argmin}_{z \in \mathbb{R}^N} \|z\|_1 \text{ verificando } \|\Phi^{cs}z - Y\|_2 \leq \|W\|_2$$

Existe una constante  $C(\delta)$  tal que

$$\|x - X\|_2 \leq \frac{C}{\sqrt{K}} \sum_{k>K} |x_{\Psi}^o[k]| + C\|W\|_2$$

donde  $x_{\Psi}^o[k]$  son los coeficientes ordenados de  $x$  respecto de  $\Psi$ .

# Aproximación

## Comentarios al teorema

$$Y = \Phi x + W$$

### Reconstrucción exacta

$$W = 0, x \in \Sigma_K(\Psi)$$

$$\delta_{3K}(\Phi^{CS}) < 1/3$$



Reconstrucción perfecta de  $x$

### Aproximación sin ruido

$$W = 0$$

$$\delta_{3K}(\Phi^{CS}) < 1/3$$



$$\|x - X\|_2 \leq \frac{C}{\sqrt{K}} \sum_{k>K} |x_{\Psi}^o[k]|$$

# Razón $M/K$ para reconstrucción exacta

## Resumen teorema

Una señal  $K$  – dispersa puede ser reconstruida de forma exacta si

- $\delta_{3K}(\Phi^{CS}) < 1/3$
- $K = \beta M / (\log N/M)$

# Razón $M/K$ para reconstrucción exacta

## Resumen teorema

Una señal  $K$  – dispersa puede ser reconstruida de forma exacta si

- $\delta_{3K}(\Phi^{CS}) < 1/3$
- $K = \beta M / (\log N/M)$

Estimaciones de  $\beta$  suelen ser pesimistas

# Razón $M/K$ para reconstrucción exacta

## Resumen teorema

Una señal  $K$  – dispersa puede ser reconstruida de forma exacta si

- $\delta_{3K}(\Phi^{cs}) < 1/3$
- $K = \beta M / (\log N/M)$

Estimaciones de  $\beta$  suelen ser pesimistas

- Utilizando técnicas estadísticas

$M = 100$ ,  $\Phi$  gaussiana

<b>N</b>	<b>M/K</b>
512	6,2
1024	7,7
2048	11
4096	16



