Compressive Sensing

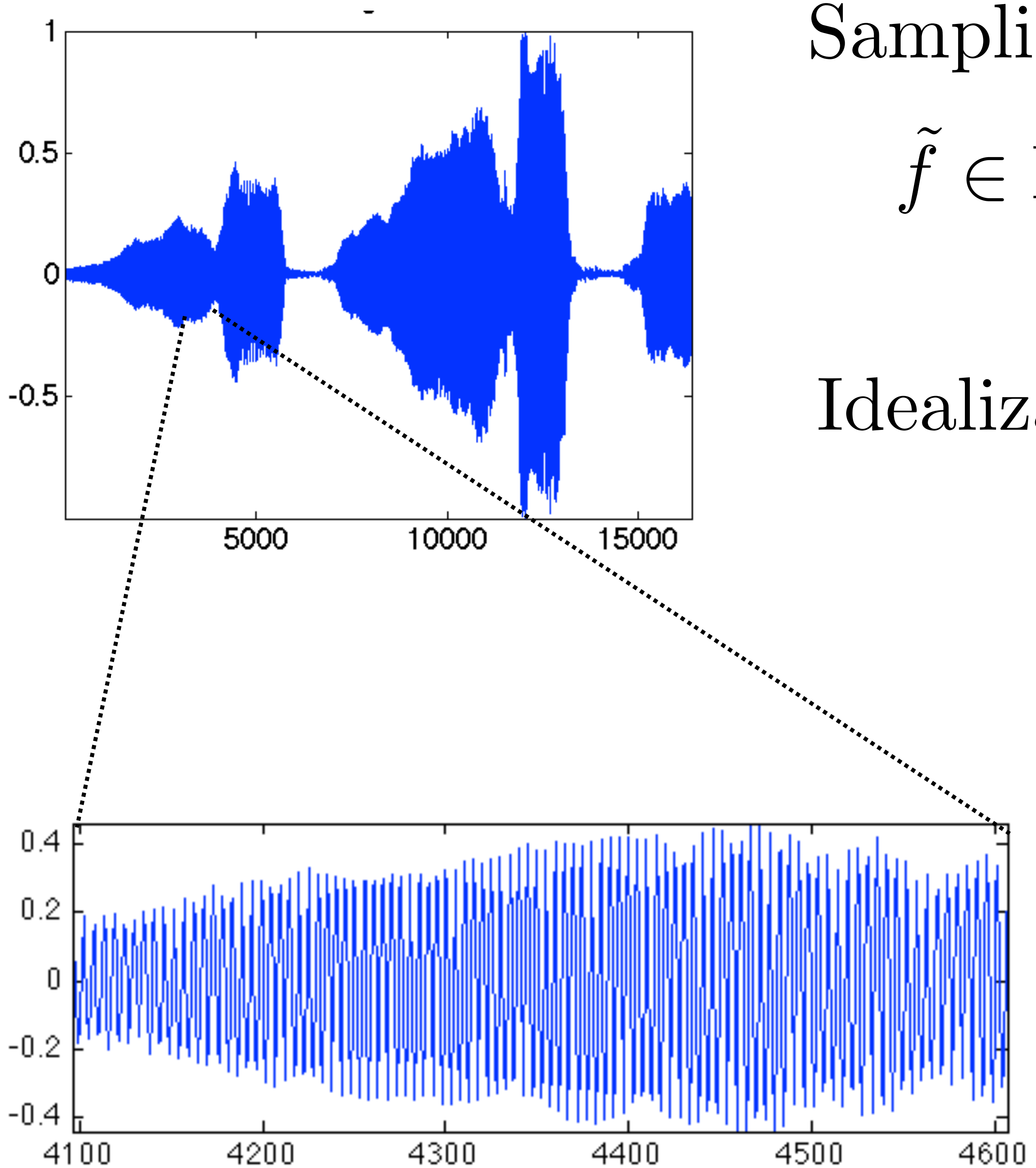
Gabriel Peyré

www.numerical-tours.com

Overview

- Shannon's World
- Compressive Sensing Acquisition
- Compressive Sensing Recovery
- Theoretical Guarantees
- Fourier Domain Measurements

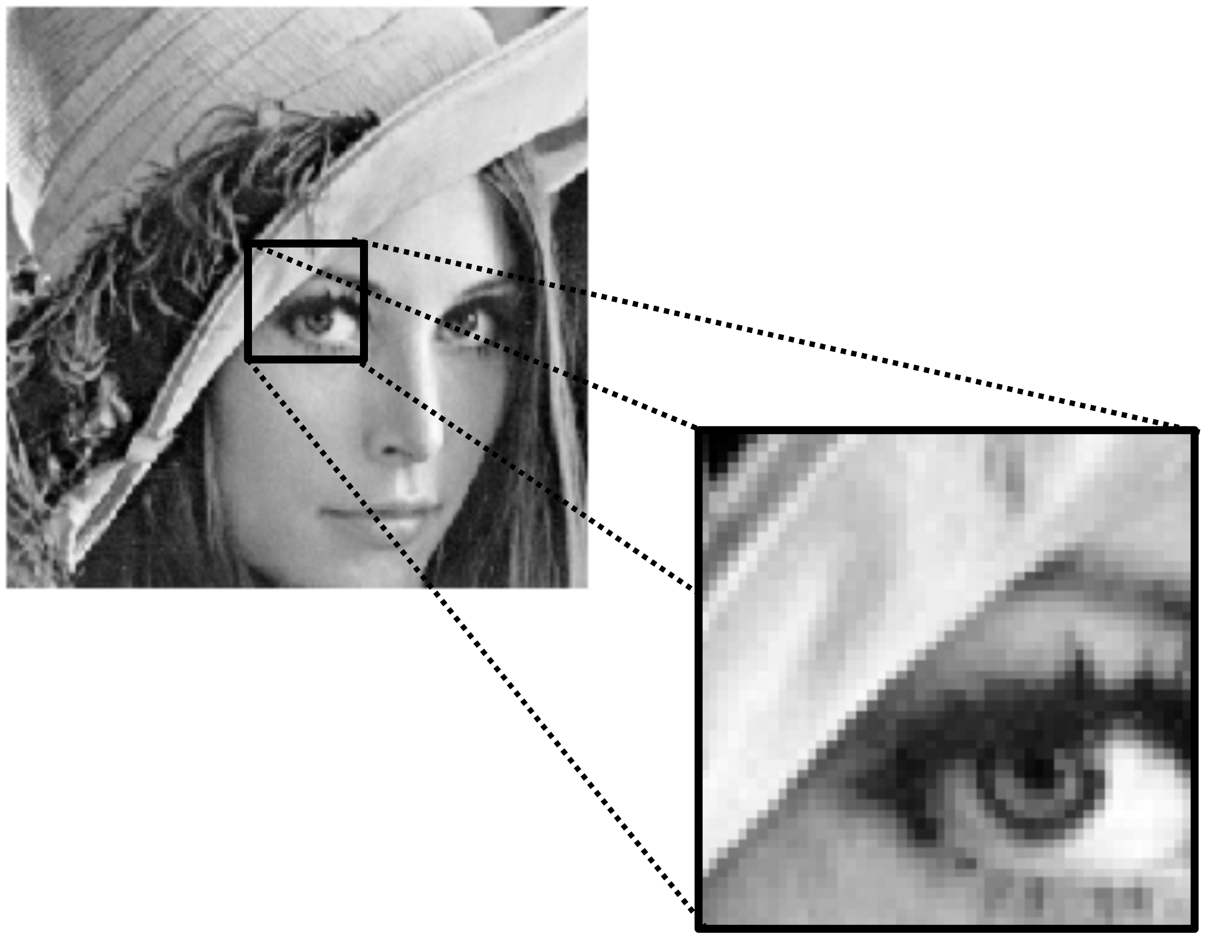
Discretization



Sampling:

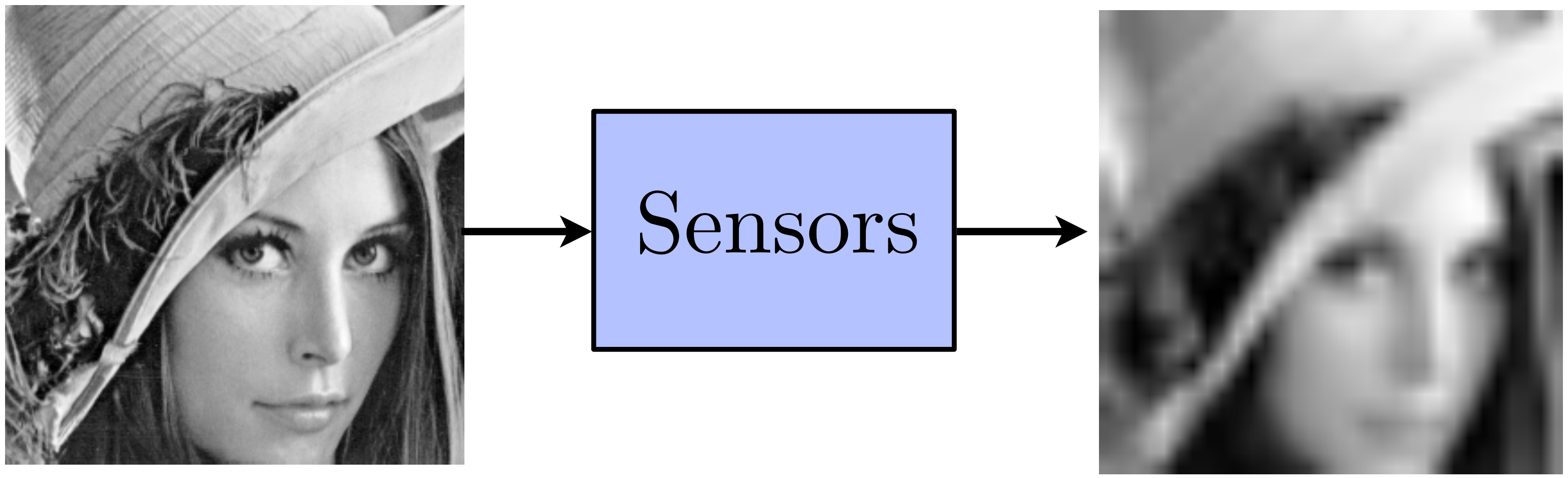
$$\tilde{f} \in \mathrm{L}^2([0,1]^d) \xrightarrow[\text{device}]{\text{acquisition}} f \in \mathbb{R}^N$$

Idealization: $f[n] \approx \tilde{f}(n/N)$



Pointwise Sampling and Smoothness

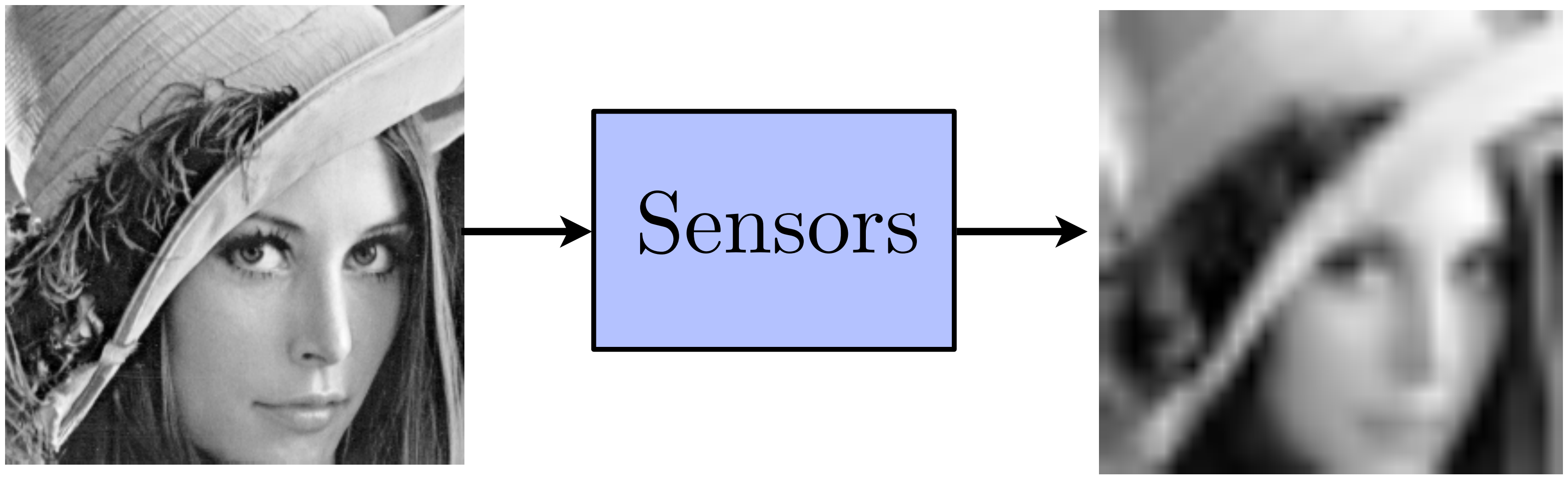
Data aquisition: $f[i] = \tilde{f}(i/N)$



$\tilde{f} \in \mathrm{L}^2$ $\qquad$ $f \in \mathbb{R}^N$

Pointwise Sampling and Smoothness

Data aquisition: $f[i] = \tilde{f}(i/N)$



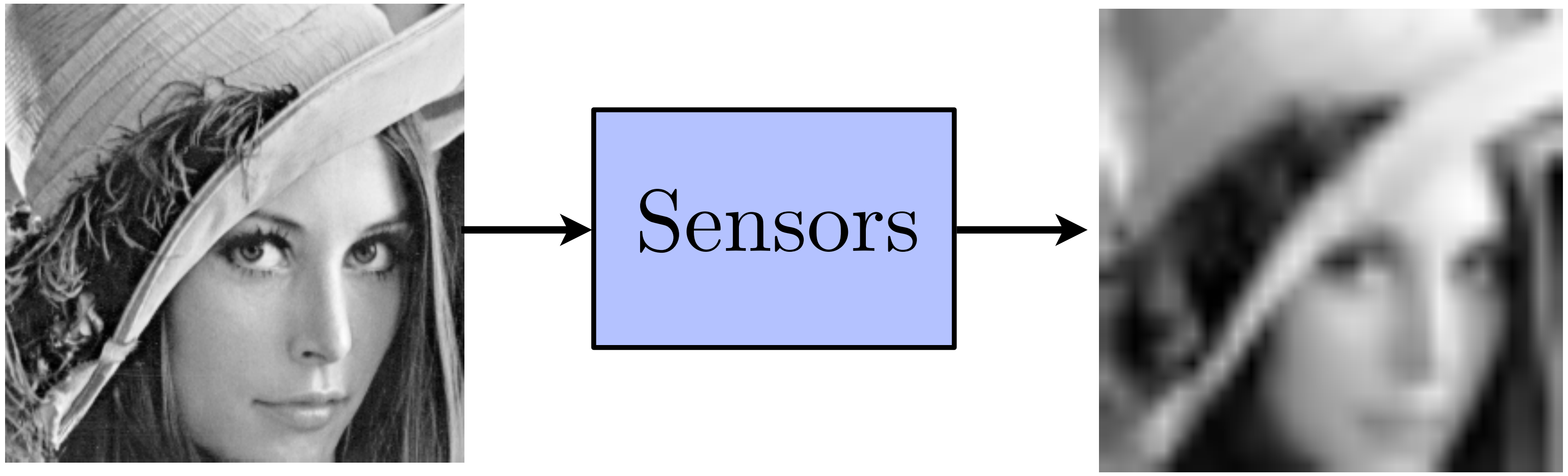
$\tilde{f} \in \mathrm{L}^2$ $\qquad$ $f \in \mathbb{R}^N$

Shannon interpolation: if $\text{Supp}(\hat{\tilde{f}}) \subset [-N\pi, N\pi]$

$$\tilde{f}(t) = \sum_i f[i] h(Nt - i) \qquad h(t) = \frac{\sin(\pi t)}{\pi t}$$

Pointwise Sampling and Smoothness

Data aquisition: $f[i] = \tilde{f}(i/N)$

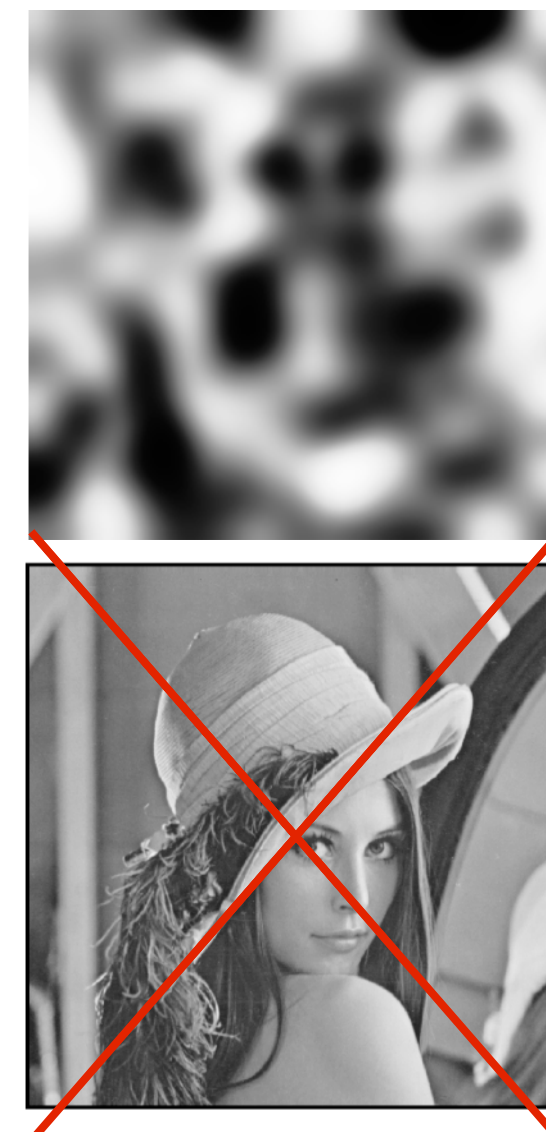


$\tilde{f} \in \mathrm{L}^2$ $\qquad$ $f \in \mathbb{R}^N$

Shannon interpolation: if $\text{Supp}(\hat{\tilde{f}}) \subset [-N\pi, N\pi]$

$$\tilde{f}(t) = \sum_i f[i] h(Nt - i) \qquad h(t) = \frac{\sin(\pi t)}{\pi t}$$

$\longrightarrow$ Natural images are not smooth.



Pointwise Sampling and Smoothness

Data aquisition: $f[i] = \tilde{f}(i/N)$

$\tilde{f} \in \mathrm{L}^2$

→ Sensors →

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

→ JPEG-2k → 0,1,0,...

Shannon interpolation: if $\mathrm{Supp}(\hat{\tilde{f}}) \subset [-N\pi, N\pi]$

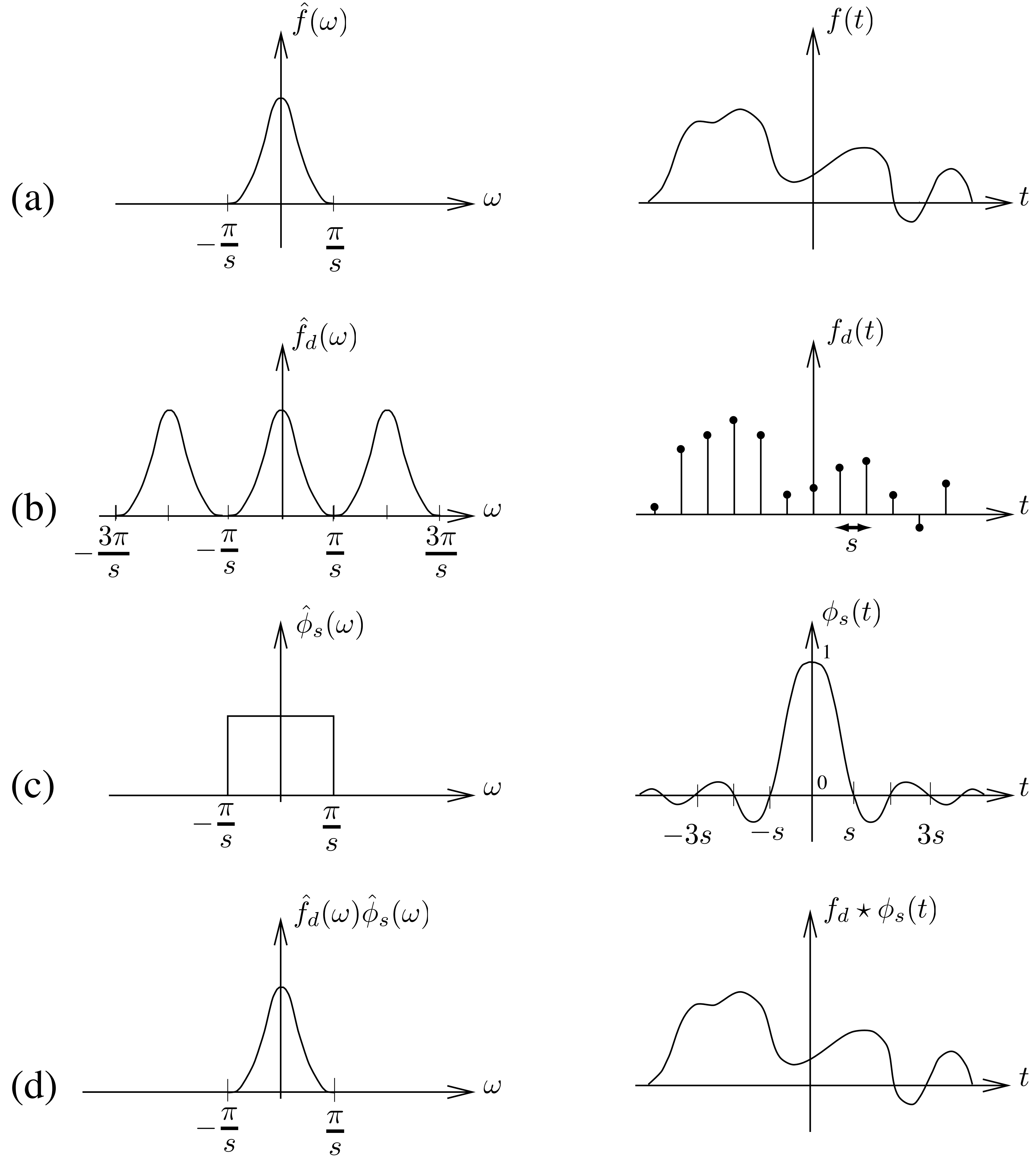
$$\tilde{f}(t) = \sum_i f[i] h(Nt - i) \qquad h(t) = \frac{\sin(\pi t)}{\pi t}$$

⟶ Natural images are not smooth.

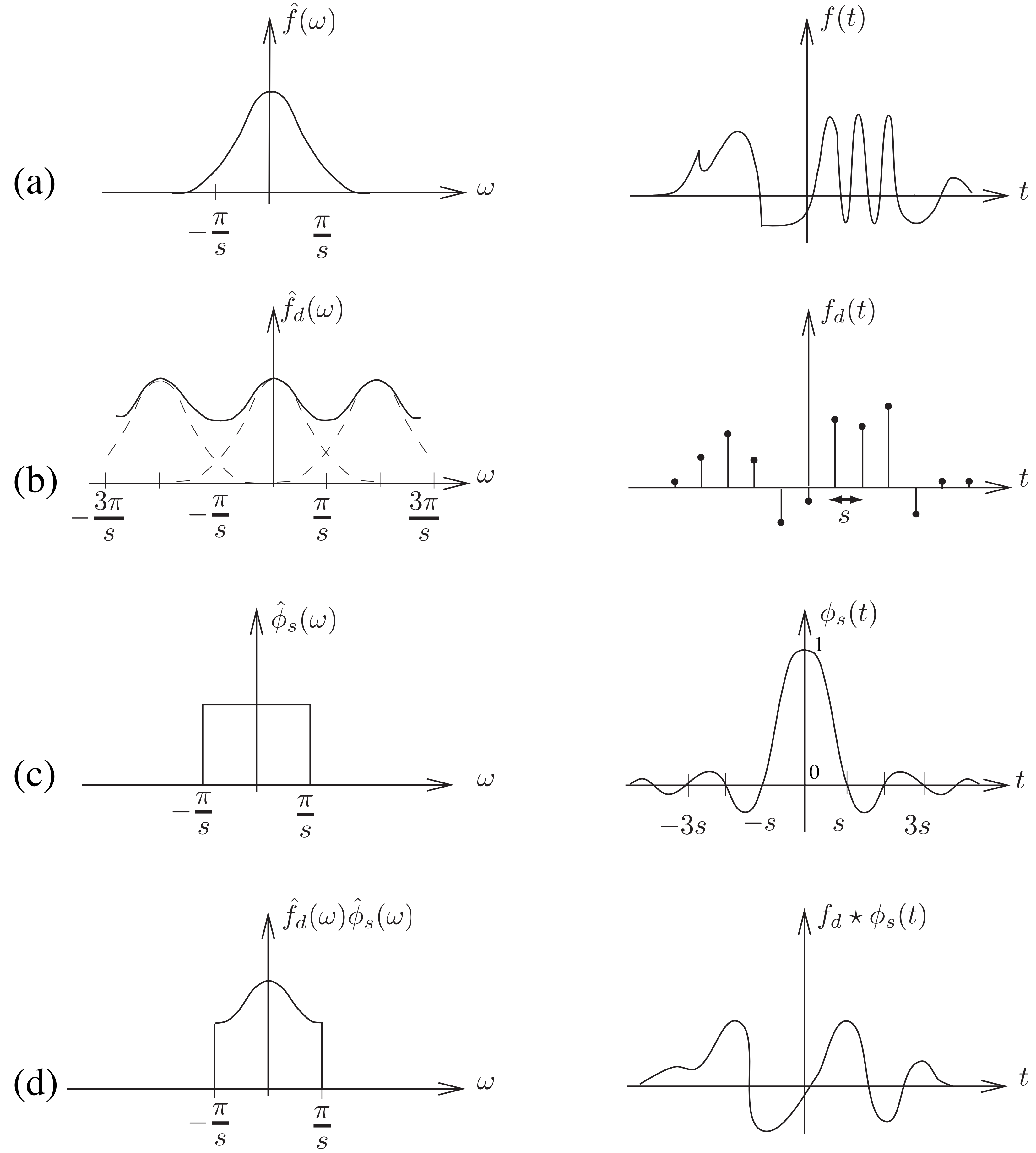
⟶ But can be compressed efficiently.

⟶ Sample *and* compress simultaneously?

Sampling and Periodization



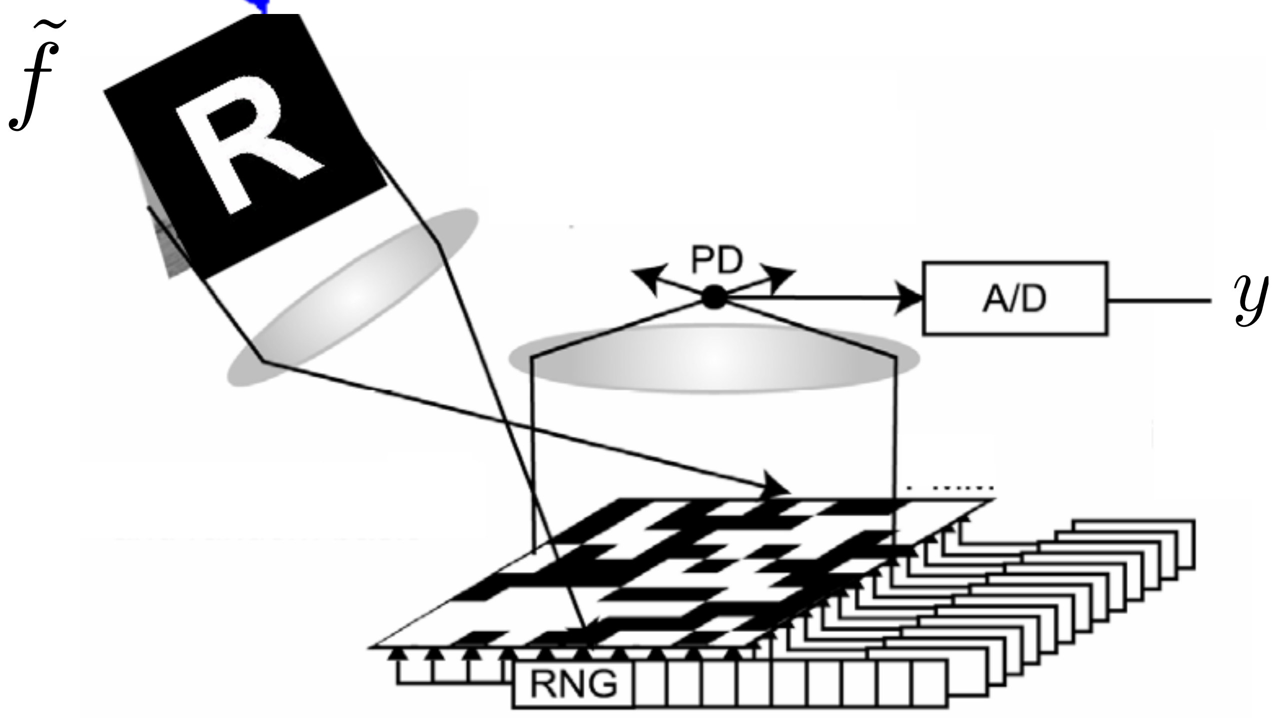
Sampling and Periodization: Aliasing

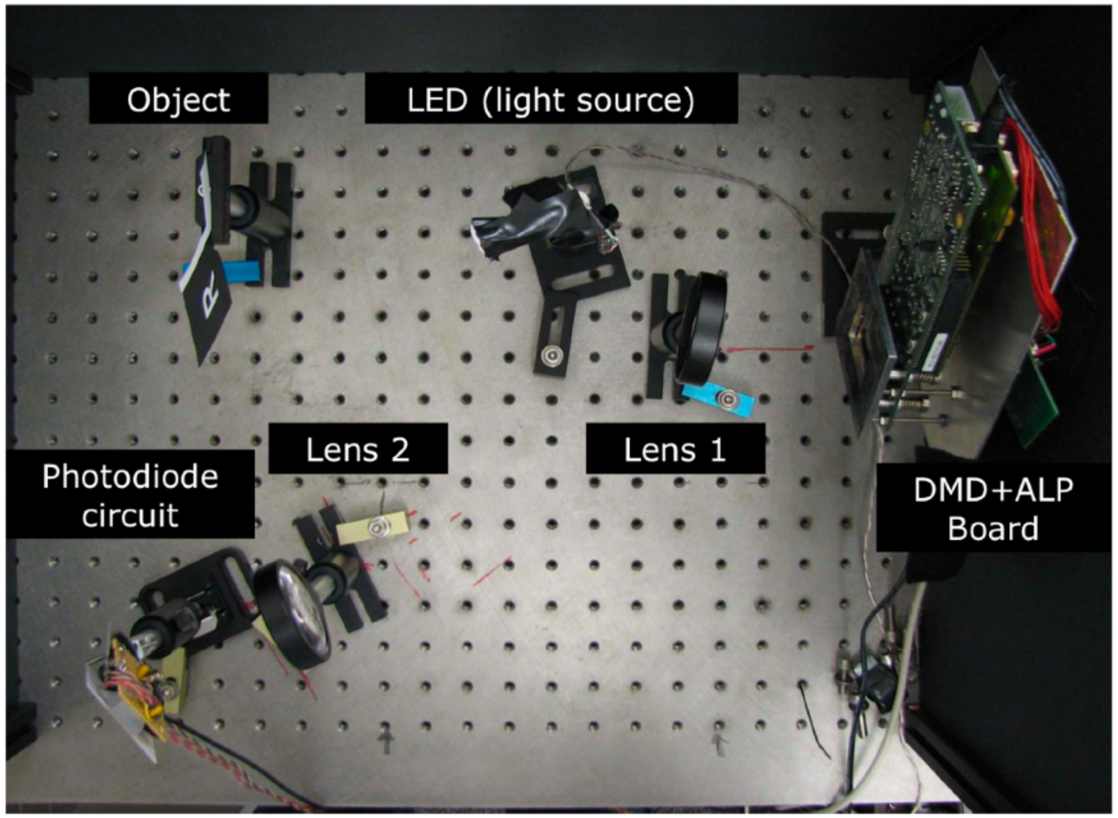


Overview

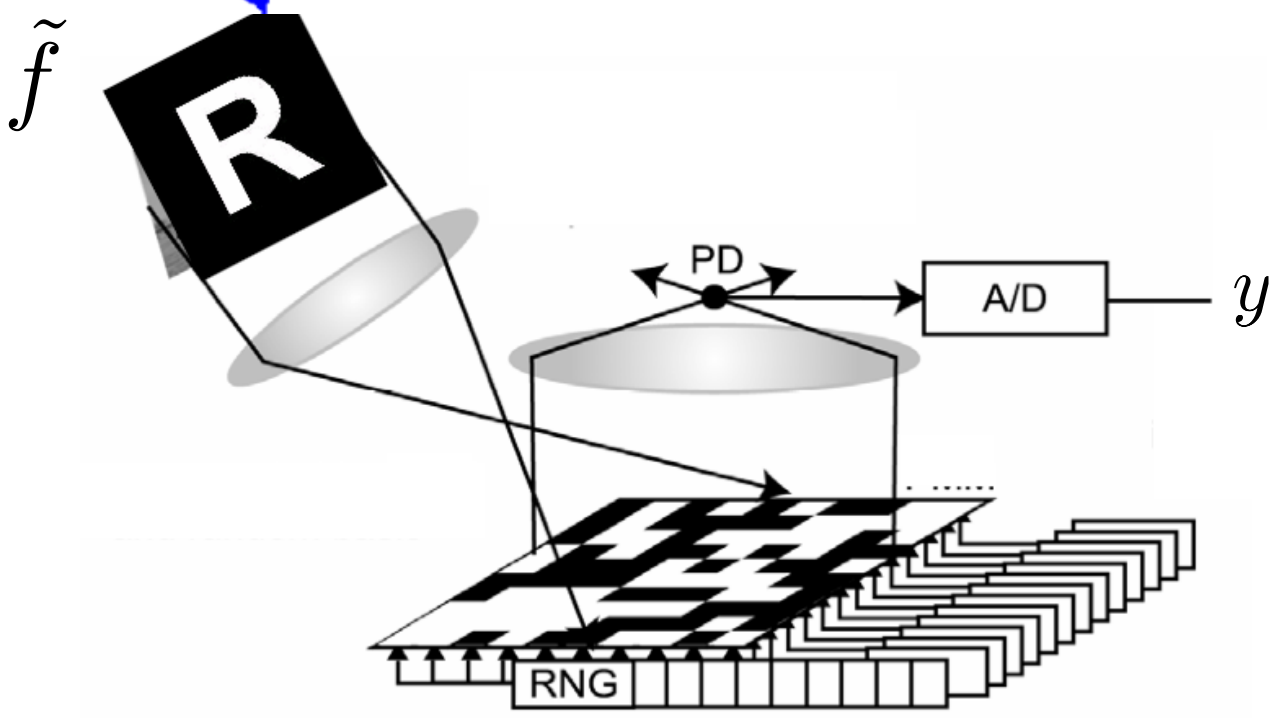
- Shannon's World
- **Compressive Sensing Acquisition**
- Compressive Sensing Recovery
- Theoretical Guarantees
- Fourier Domain Measurements

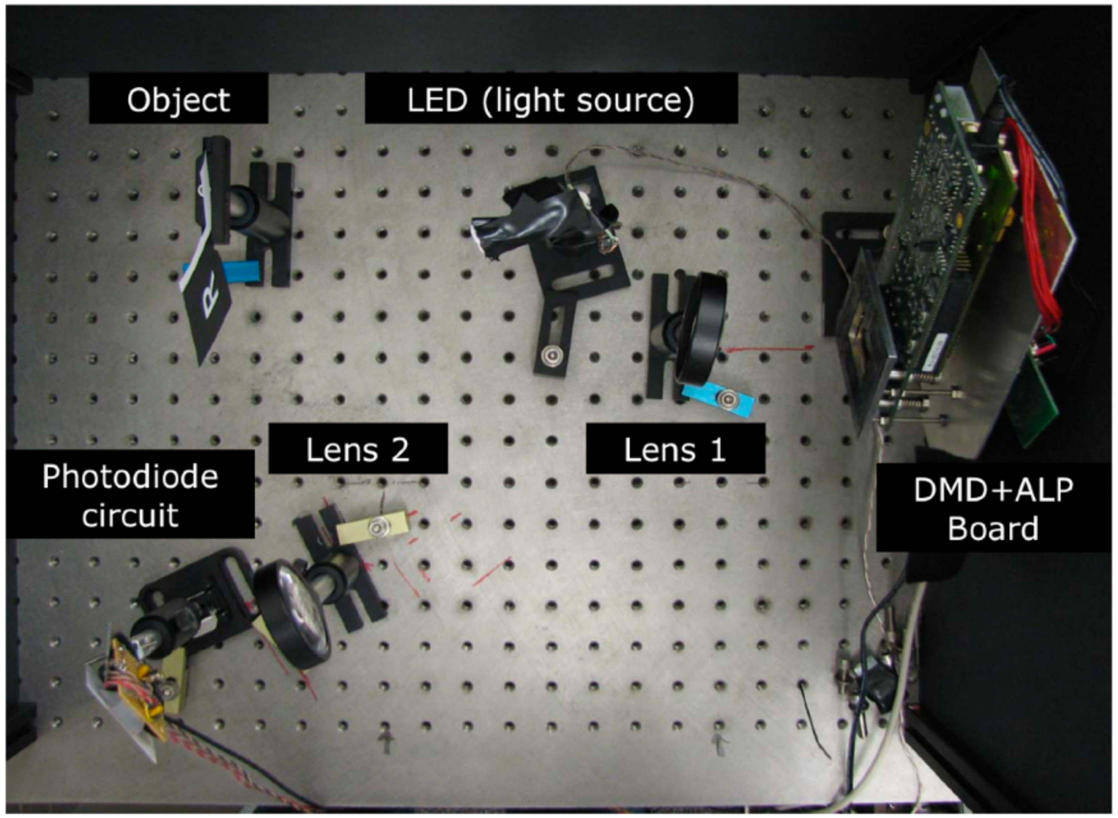
Single Pixel Camera (Rice)



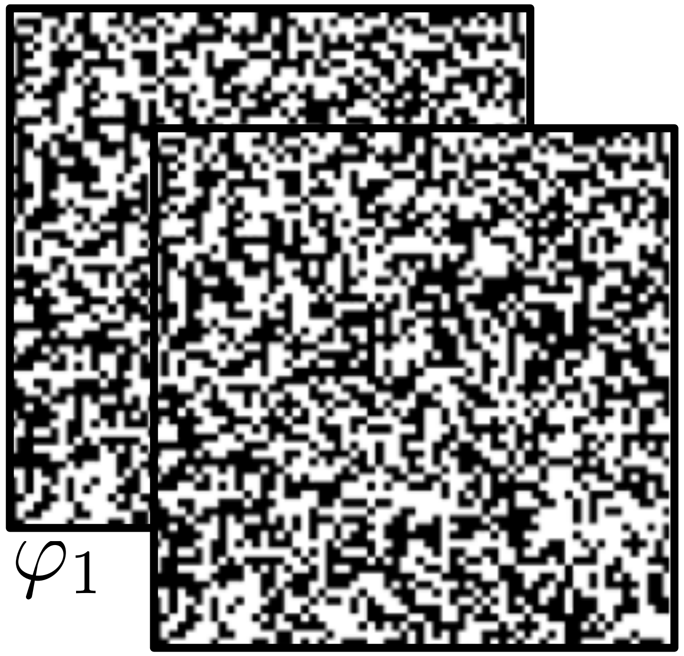


Single Pixel Camera (Rice)





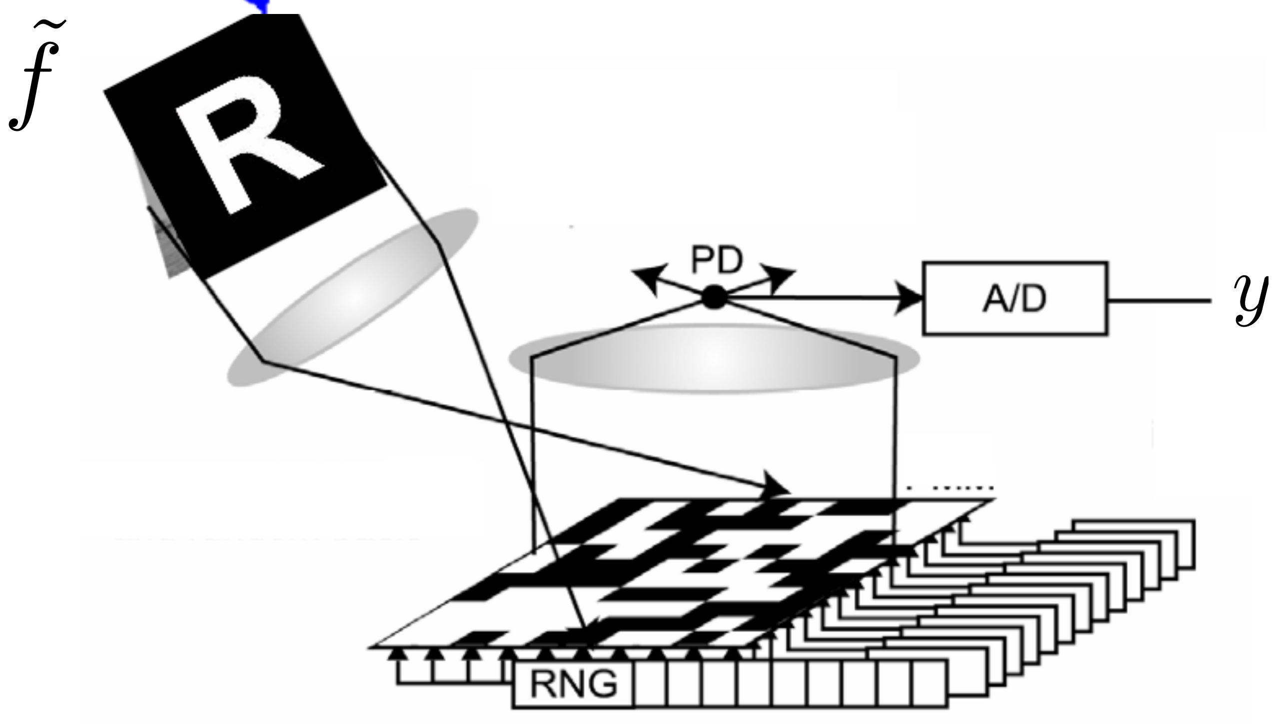
$$y[i] = \langle f, \varphi_i \rangle$$

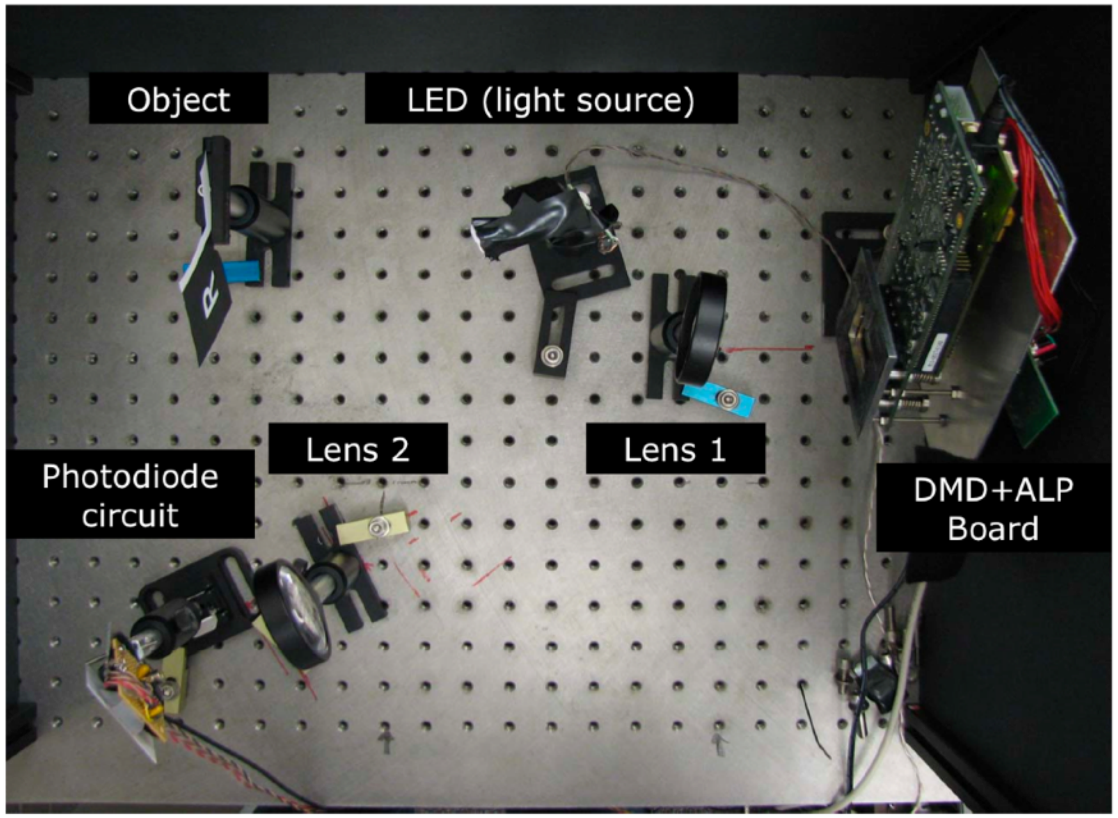


φ_1 φ_2

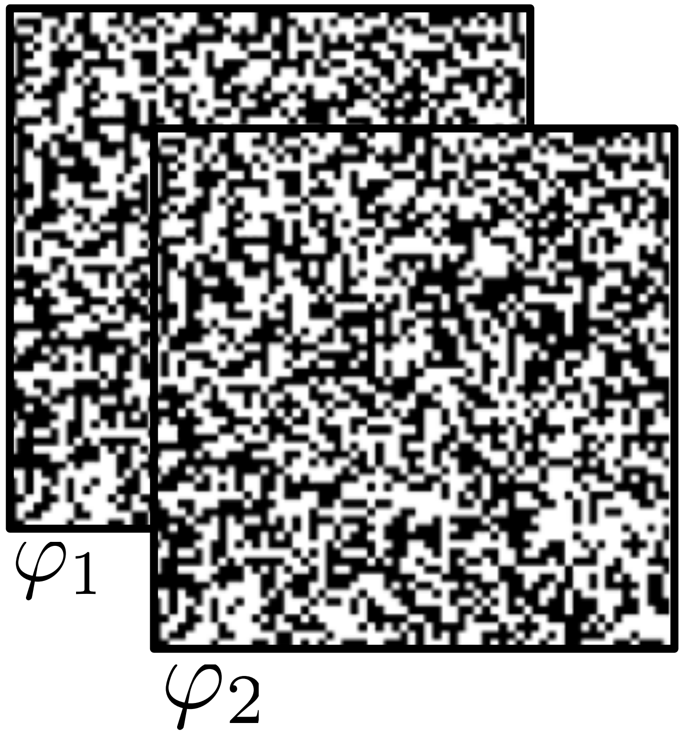
P measures $\ll N$ micro-mirrors

Single Pixel Camera (Rice)





$$y[i] = \langle f, \varphi_i \rangle$$



P measures $\ll N$ micro-mirrors

$P/N = 1$

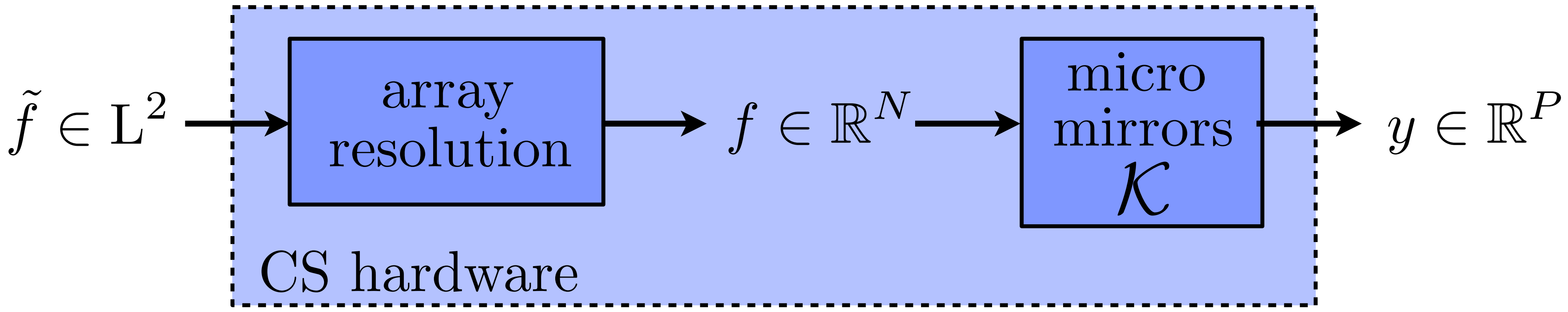
$P/N = 0.16$

$P/N = 0.02$

CS Hardware Model

CS is about designing hardware: input signals $\tilde{f} \in \mathrm{L}^2(\mathbb{R}^2)$.

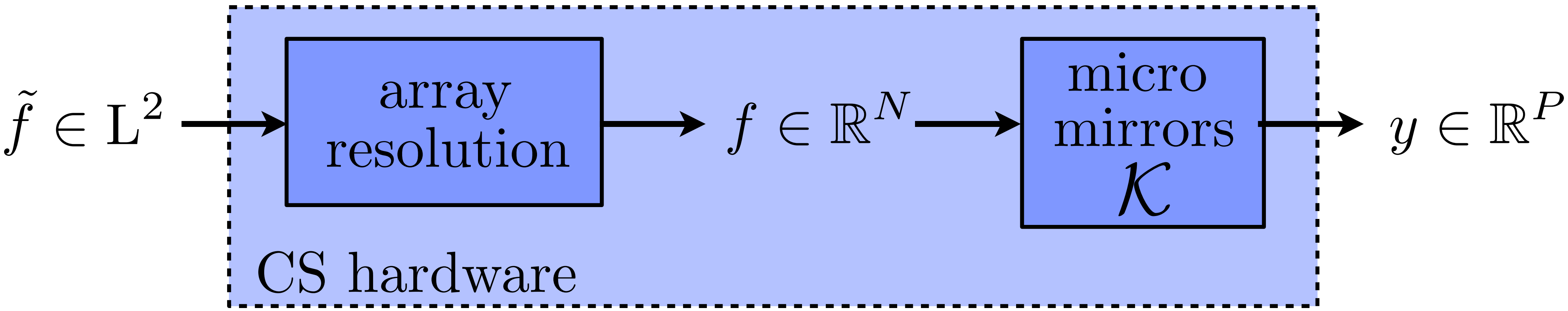
Physical hardware resolution limit: target resolution $f \in \mathbb{R}^N$.

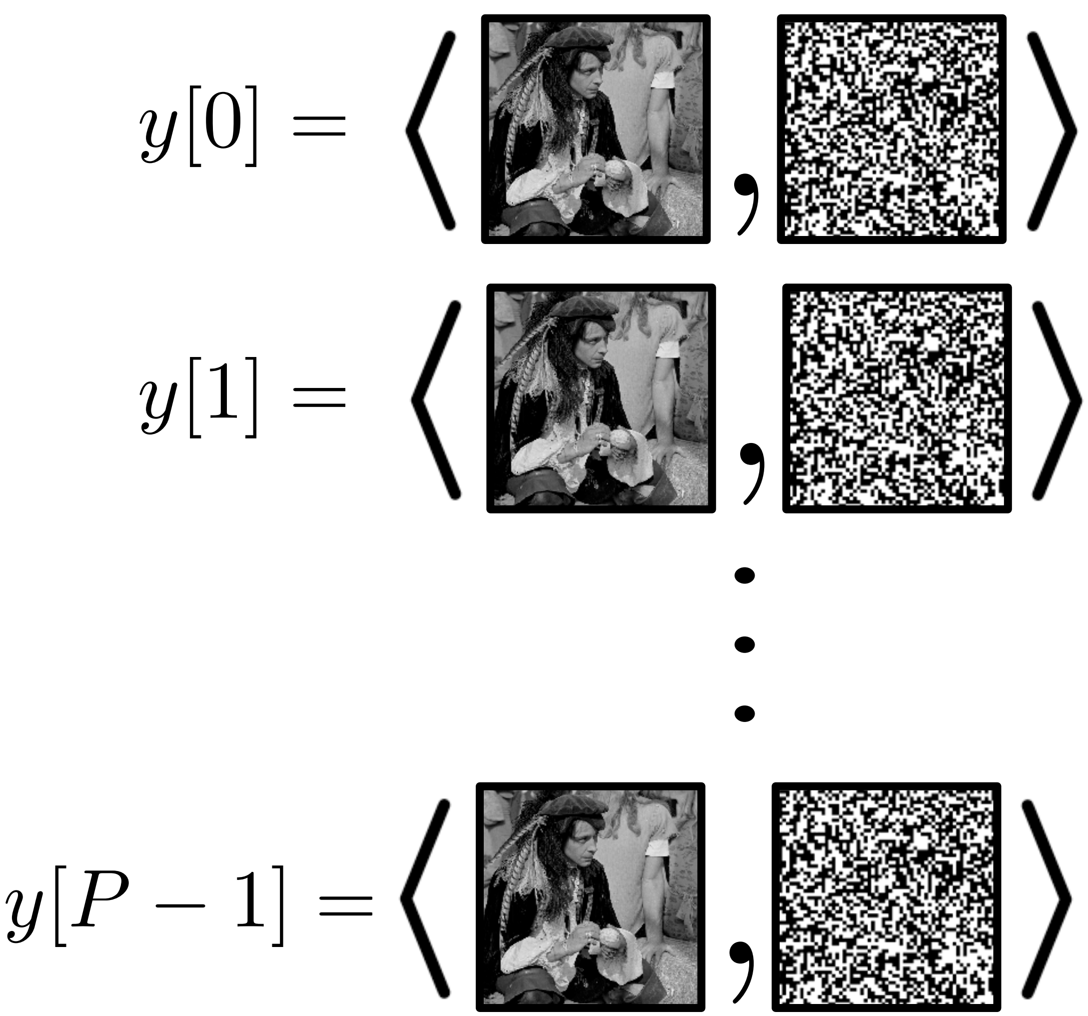


CS Hardware Model

CS is about designing hardware: input signals $\tilde{f} \in \mathrm{L}^2(\mathbb{R}^2)$.

Physical hardware resolution limit: target resolution $f \in \mathbb{R}^N$.

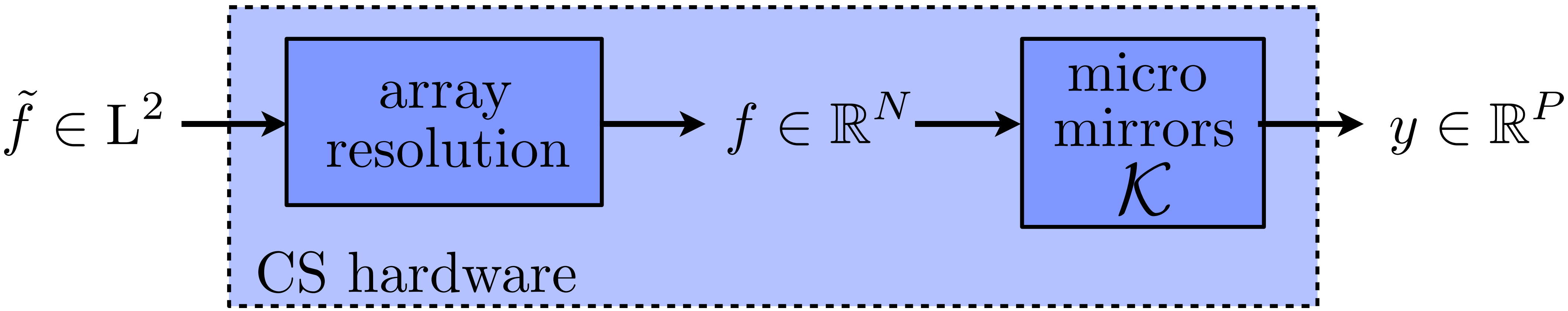


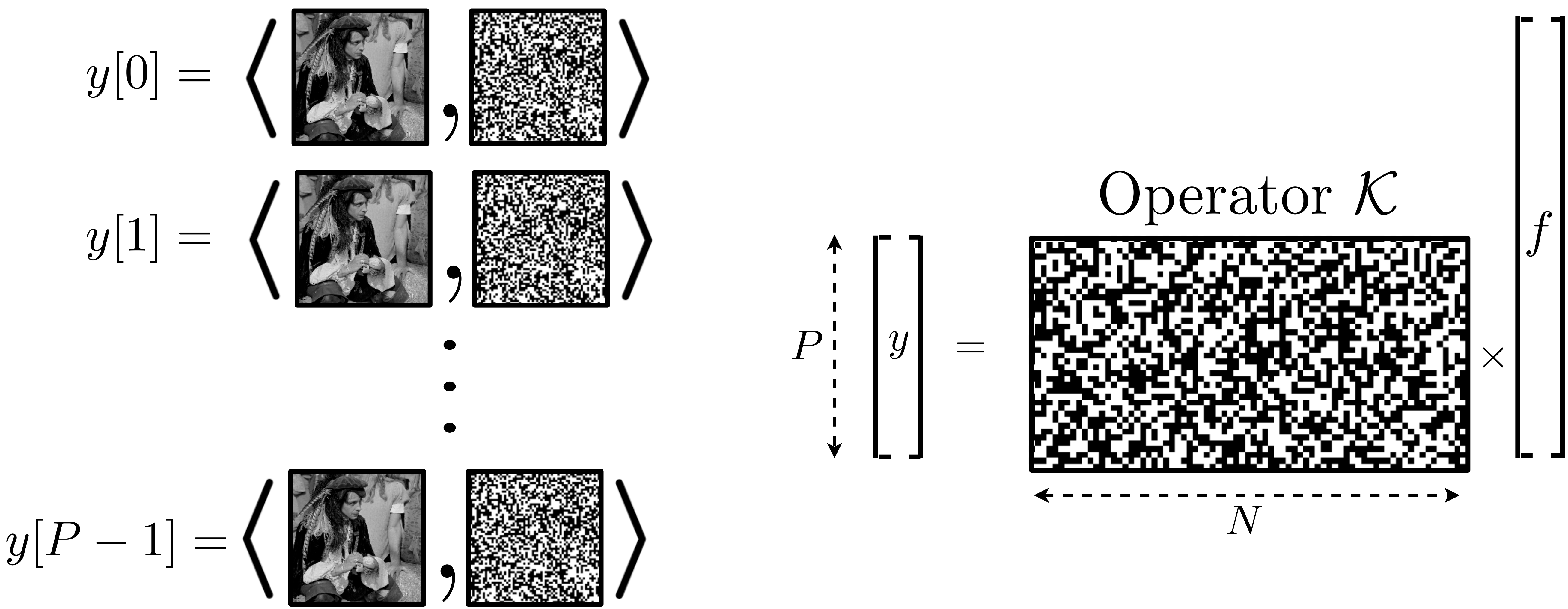


CS Hardware Model

CS is about designing hardware: input signals $\tilde{f} \in \mathrm{L}^2(\mathbb{R}^2)$.

Physical hardware resolution limit: target resolution $f \in \mathbb{R}^N$.





Overview

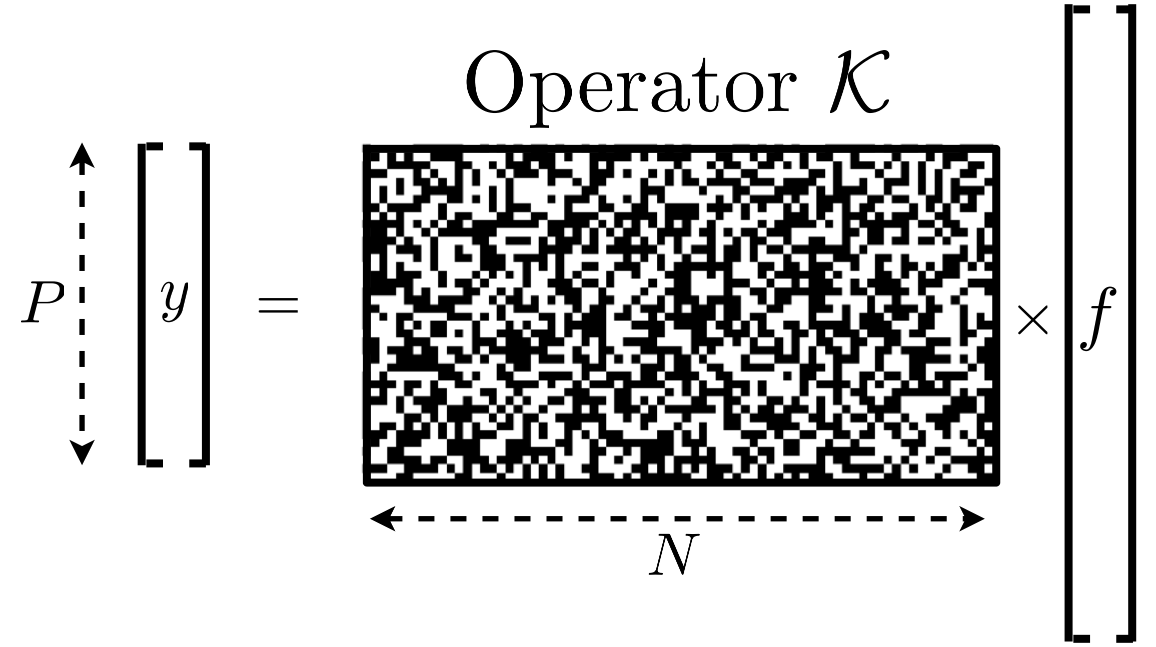
- Shannon's World
- Compressive Sensing Acquisition
- **Compressive Sensing Recovery**
- Theoretical Guarantees
- Fourier Domain Measurements

Inversion and Sparsity

Need to solve $y = \mathcal{K} f$.

$\rightarrow$ More unknown than equations.

$\dim(\ker(\mathcal{K})) = N - P$ is huge.

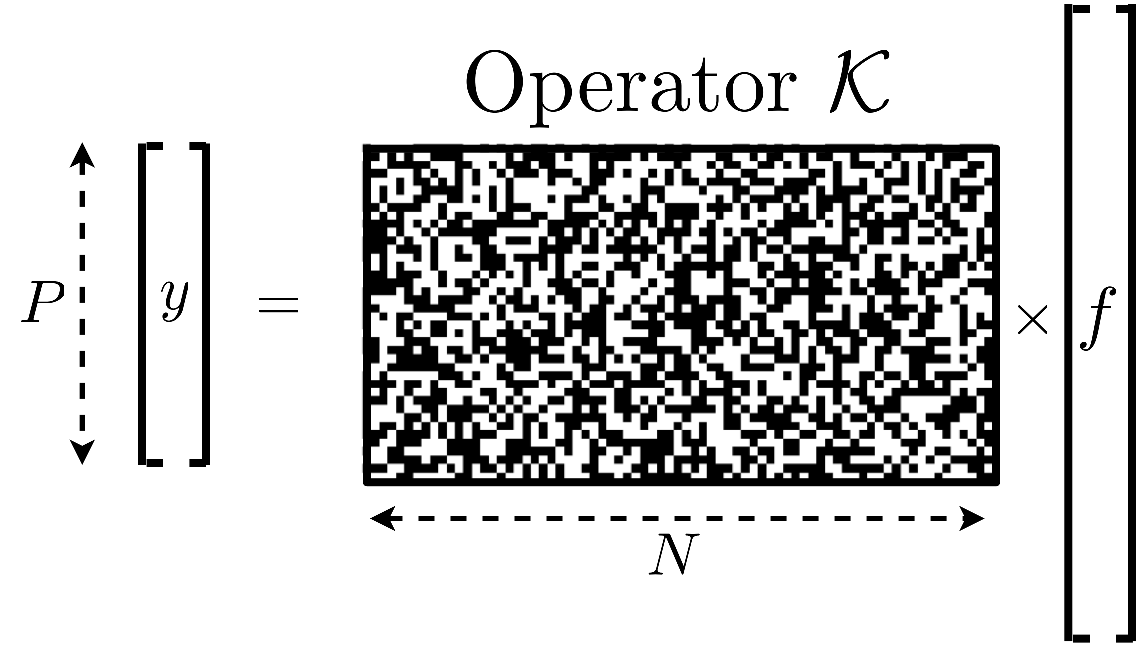


Inversion and Sparsity

Need to solve $y = \mathcal{K} f$.

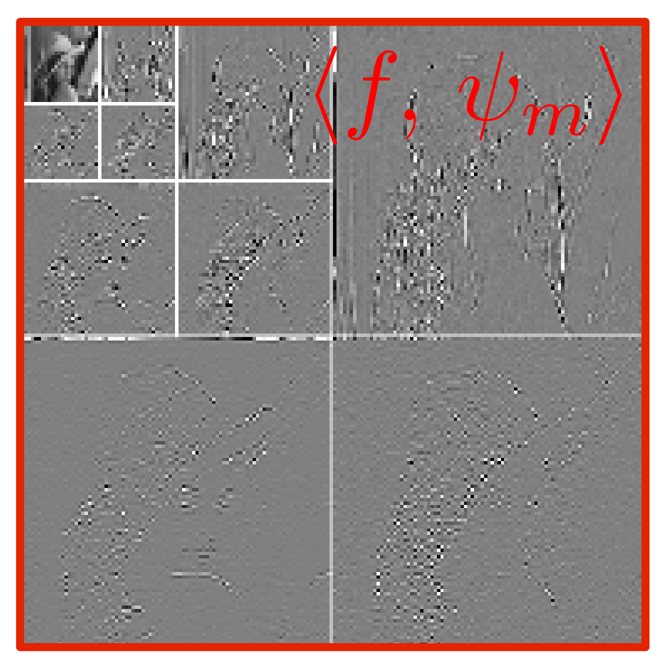
$\rightarrow$ More unknown than equations.

$\dim(\ker(\mathcal{K})) = N - P$ is huge.



Prior information: f is sparse in a basis $\{\psi_m\}_m$.

$$J_\varepsilon(f) = \text{Card}\,\{m \setminus |\langle f,\, \psi_m \rangle| > \varepsilon\} \quad \text{is small.}$$



Convex Relaxation: L1 Prior

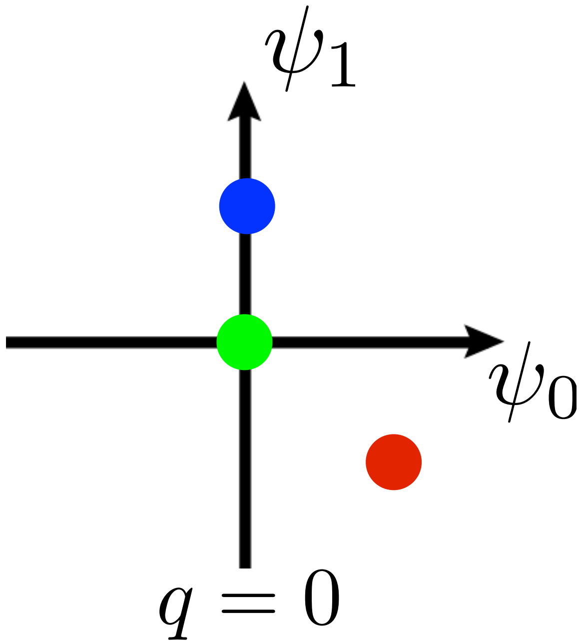
"Ideal" sparsity prior: $J_0(f) = \#\{m \setminus \langle f, \psi_m \rangle \neq 0\}$

Image with 2 pixels:

$J_0(f) = 0 \longrightarrow$ null image.

$J_0(f) = 1 \longrightarrow$ sparse image.

$J_0(f) = 2 \longrightarrow$ non-sparse image.



Convex Relaxation: L1 Prior

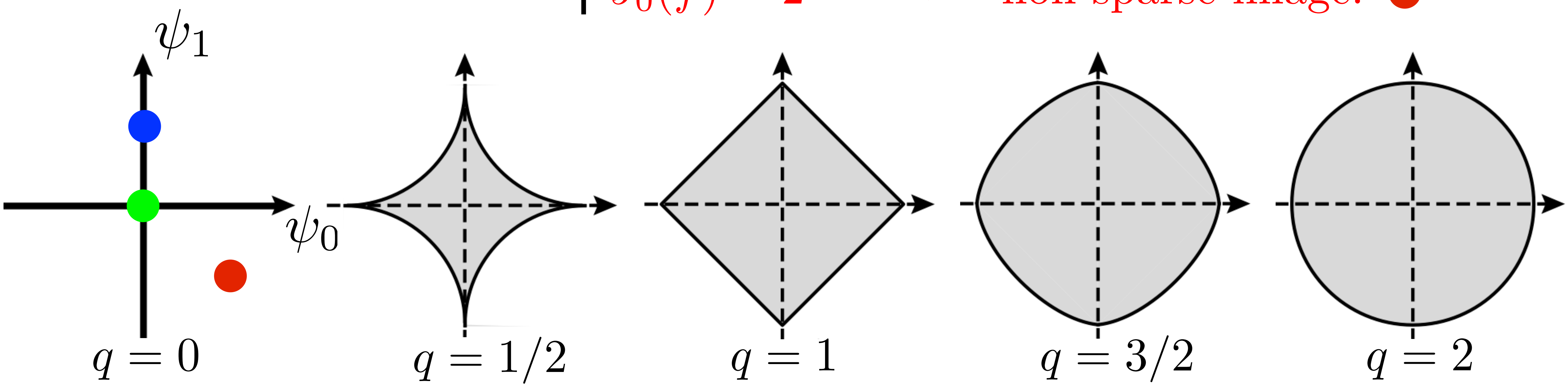
"Ideal" sparsity prior: $J_0(f) = \# \{ m \setminus \langle f,\, \psi_m \rangle \neq 0 \}$

Image with 2 pixels:

$J_0(f) = 0 \longrightarrow$ null image.

$J_0(f) = 1 \longrightarrow$ sparse image.

$J_0(f) = 2 \longrightarrow$ non-sparse image.



ℓ^q priors: $J_q(f) = \sum_m |\langle f,\, \psi_m \rangle|^q$ (convex for $q \geqslant 1$)

Convex Relaxation: L1 Prior

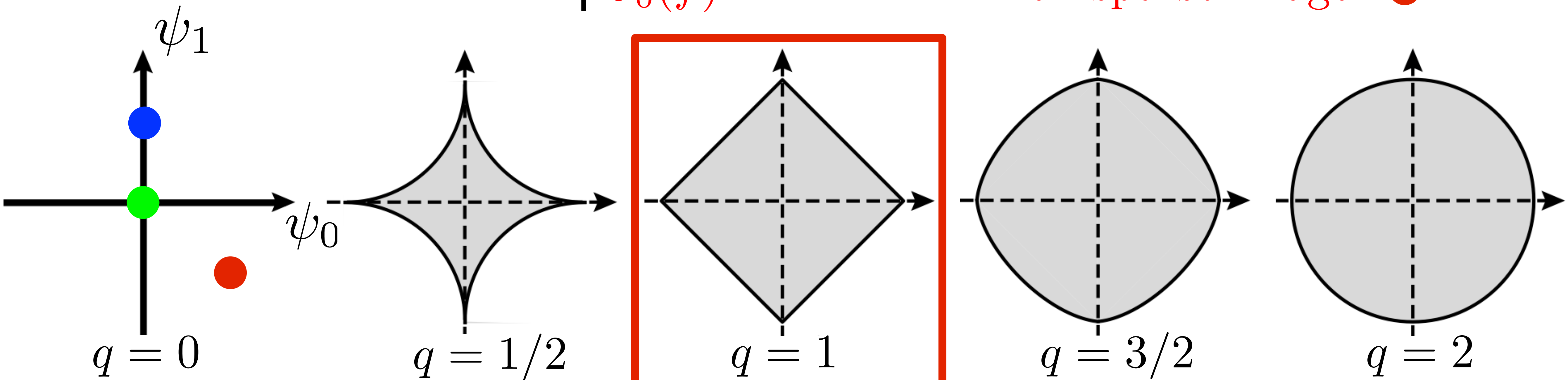
"Ideal" sparsity prior: $J_0(f) = \#\{m \backslash \langle f,\, \psi_m \rangle \neq 0\}$

Image with 2 pixels:

$J_0(f) = 0 \longrightarrow$ null image.

$J_0(f) = 1 \longrightarrow$ sparse image.

$J_0(f) = 2 \longrightarrow$ non-sparse image.



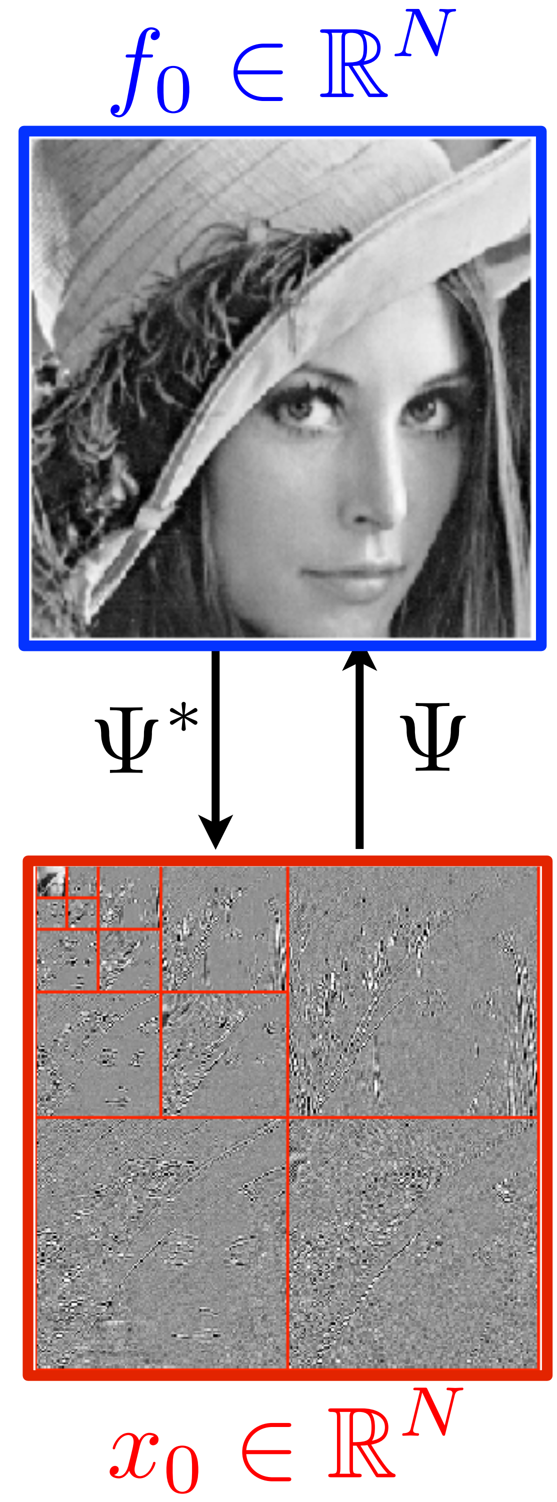
ℓ^q priors: $J_q(f) = \sum_m |\langle f,\, \psi_m \rangle|^q$ (convex for $q \geqslant 1$)

ℓ^1 norm: ℓ^q norm the "closest" to the ℓ^0 ideal sparsity.

Sparse ℓ^1 prior: $J_1(f) = \sum_m |\langle f,\, \psi_m \rangle|$

Sparse CS Recovery

$f_0 \in \mathbb{R}^N$ sparse in ortho-basis Ψ

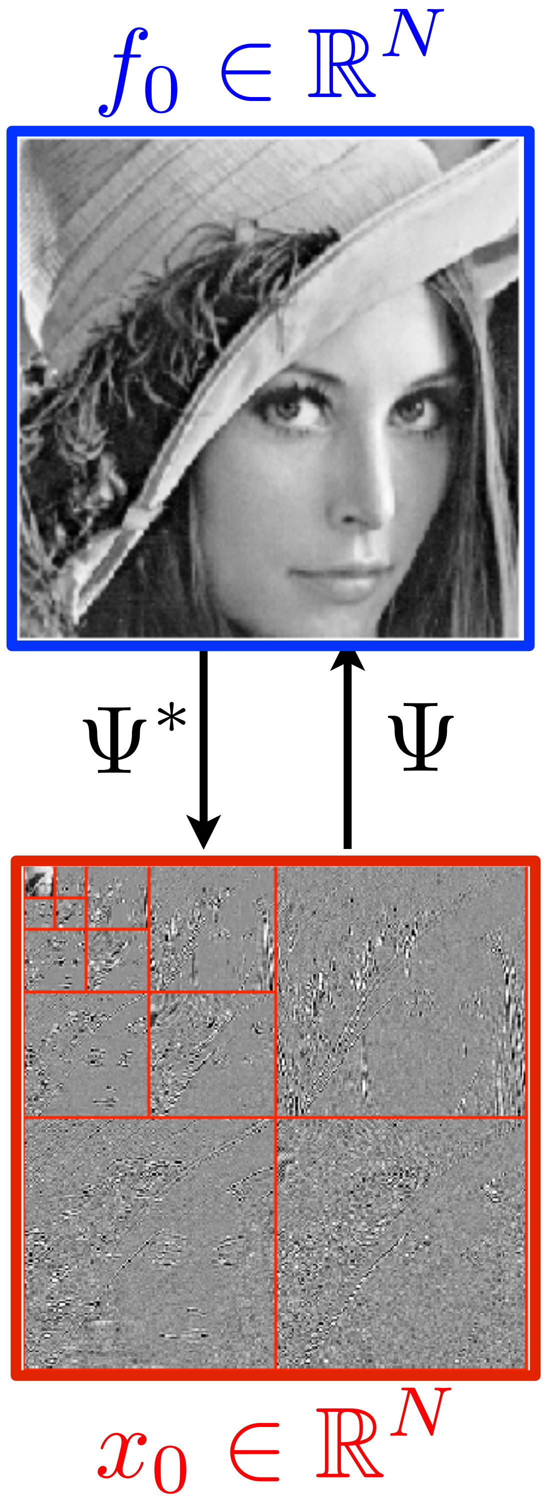


Sparse CS Recovery

$f_0 \in \mathbb{R}^N$ sparse in ortho-basis Ψ

(Discretized) sampling acquisition:

$$y = \mathcal{K} f_0 + w = \underbrace{\mathcal{K} \circ \Psi}_{= \Phi}(x_0) + w$$



Sparse CS Recovery

$f_0 \in \mathbb{R}^N$ sparse in ortho-basis Ψ

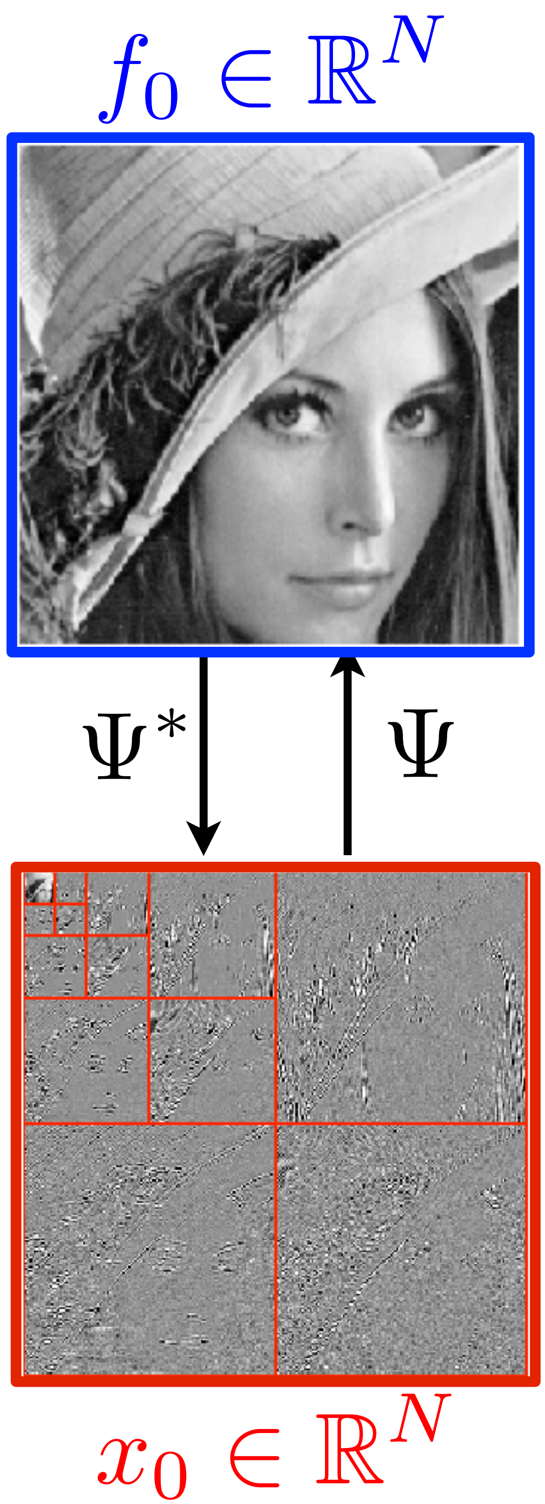
(Discretized) sampling acquisition:

$$y = \mathcal{K} f_0 + w = \mathcal{K} \circ \Psi(x_0) + w$$
$$= \Phi$$

$\mathcal{K}$ drawn from the Gaussian matrix ensemble

$$\mathcal{K}_{i,j} \sim \mathcal{N}(0, P^{-1/2}) \text{ i.i.d.}$$

$\Rightarrow \Phi$ drawn from the Gaussian matrix ensemble



Sparse CS Recovery

$f_0 \in \mathbb{R}^N$ sparse in ortho-basis Ψ

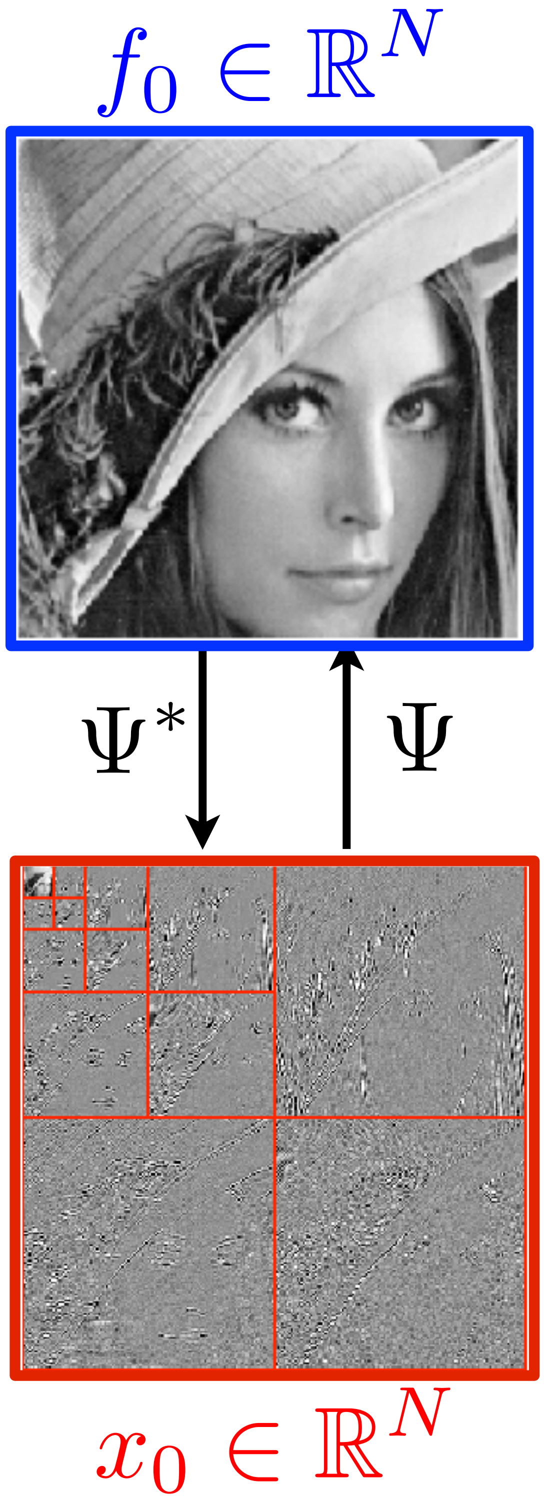
(Discretized) sampling acquisition:

$$y = \mathcal{K} f_0 + w = \underbrace{\mathcal{K} \circ \Psi}_{= \Phi}(x_0) + w$$

$\mathcal{K}$ drawn from the Gaussian matrix ensemble

$$\mathcal{K}_{i,j} \sim \mathcal{N}(0, P^{-1/2}) \text{ i.i.d.}$$

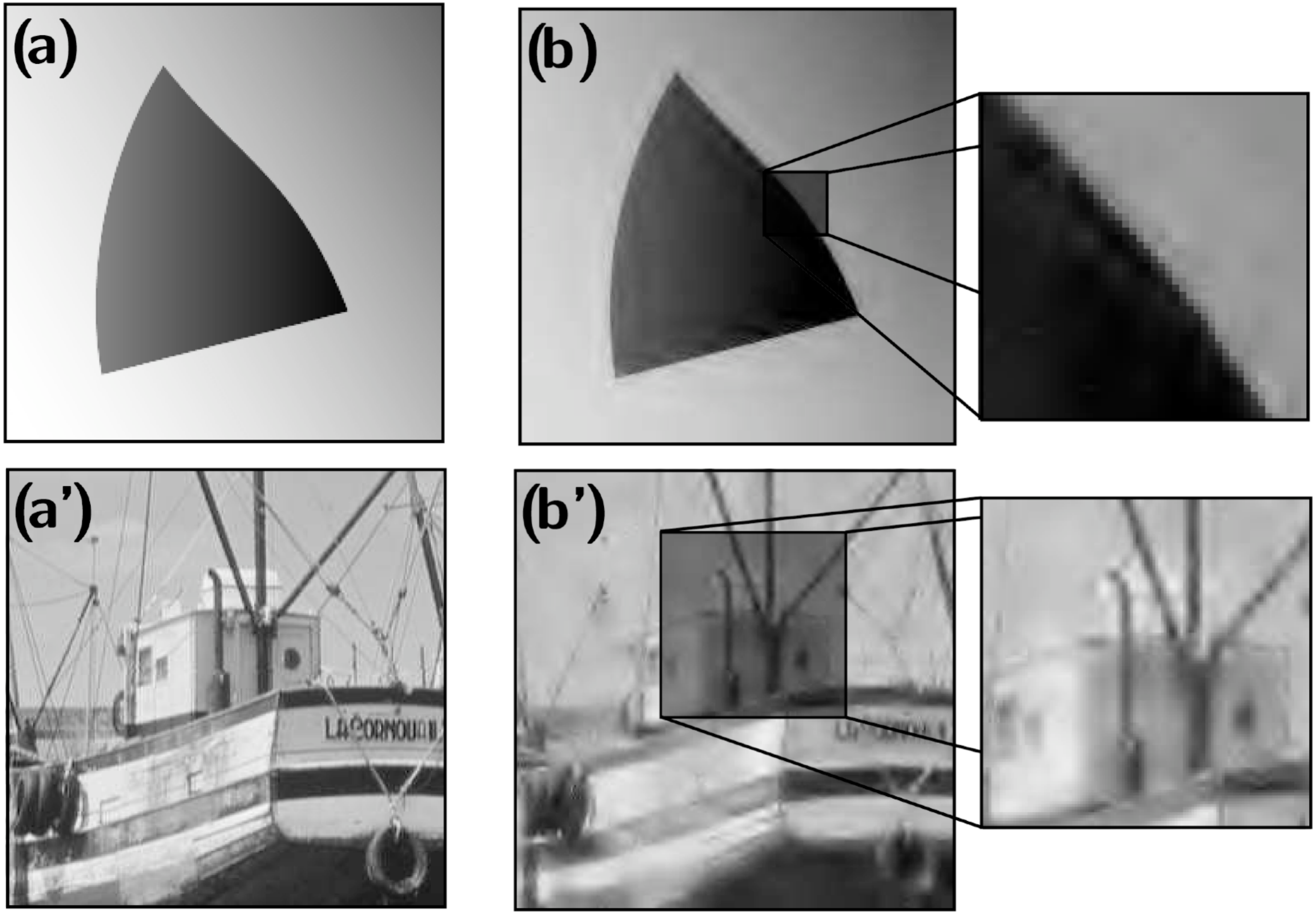
$\Rightarrow$ Φ drawn from the Gaussian matrix ensemble



Sparse recovery:

$$\min_{\|\Phi x - y\| \leqslant \|w\|} \|x\|_1 \quad \overset{\|w\| \longleftrightarrow \lambda}{\longleftrightarrow} \quad \min_x \frac{1}{2}\|\Phi x - y\|^2 + \lambda \|x\|_1$$

CS Simulation Example



Original f_0

Recovery $f^\star$, $P = N/6$

Ψ = translation invariant wavelet frame

Overview

- Shannon's World
- Compressive Sensing Acquisition
- Compressive Sensing Recovery
- **Theoretical Guarantees**
- Fourier Domain Measurements

CS with RIP

ℓ^1 *recovery:*

$$x^\star \in \underset{\|\Phi x - y\| \leqslant \varepsilon}{\operatorname{argmin}} \; \|x\|_1 \quad \text{where} \quad \begin{cases} y = \Phi x_0 + w \\ \|w\| \leqslant \varepsilon \end{cases}$$

Restricted Isometry Constants:

$$\forall \, \|x\|_0 \leqslant k, \quad (1-\delta_k)\|x\|^2 \leqslant \|\Phi x\|^2 \leqslant (1+\delta_k)\|x\|^2$$

CS with RIP

ℓ^1 recovery:

$$x^\star \in \underset{\|\Phi x - y\| \leqslant \varepsilon}{\operatorname{argmin}} \|x\|_1 \quad \text{where} \quad \begin{cases} y = \Phi x_0 + w \\ \|w\| \leqslant \varepsilon \end{cases}$$

Restricted Isometry Constants:

$$\forall \, \|x\|_0 \leqslant k, \quad (1 - \delta_k)\|x\|^2 \leqslant \|\Phi x\|^2 \leqslant (1 + \delta_k)\|x\|^2$$

Theorem: If $\delta_{2k} \leqslant \sqrt{2} - 1$, then [Candes 2009]

$$\|x_0 - x^\star\| \leqslant \frac{C_0}{\sqrt{k}} \|x_0 - x_k\|_1 + C_1 \varepsilon$$

where x_k is the best k-term approximation of x_0.

RIP for Gaussian Matrices

Link with coherence: $$\mu(\Phi) = \max_{i \neq j} |\langle \varphi_i, \, \varphi_j \rangle|$$

$$\delta_2 = \mu(\Phi)$$

$$\delta_k \leqslant (k-1)\mu(\Phi)$$

RIP for Gaussian Matrices

Link with coherence: $\mu(\Phi) = \max_{i \neq j} |\langle \varphi_i, \varphi_j \rangle|$

$$\delta_2 = \mu(\Phi)$$

$$\delta_k \leqslant (k-1)\mu(\Phi)$$

For Gaussian matrices:

$$\mu(\Phi) \sim \sqrt{\log(PN)/P}$$

RIP for Gaussian Matrices

Link with coherence: $\mu(\Phi) = \max_{i \neq j} |\langle \varphi_i, \varphi_j \rangle|$

$$\delta_2 = \mu(\Phi)$$

$$\delta_k \leqslant (k-1)\mu(\Phi)$$

For Gaussian matrices:

$$\mu(\Phi) \sim \sqrt{\log(PN)/P}$$

Stronger result:

Theorem: If $k \leqslant \frac{C}{\log(N/P)} P$ [Candès et al, 2004]

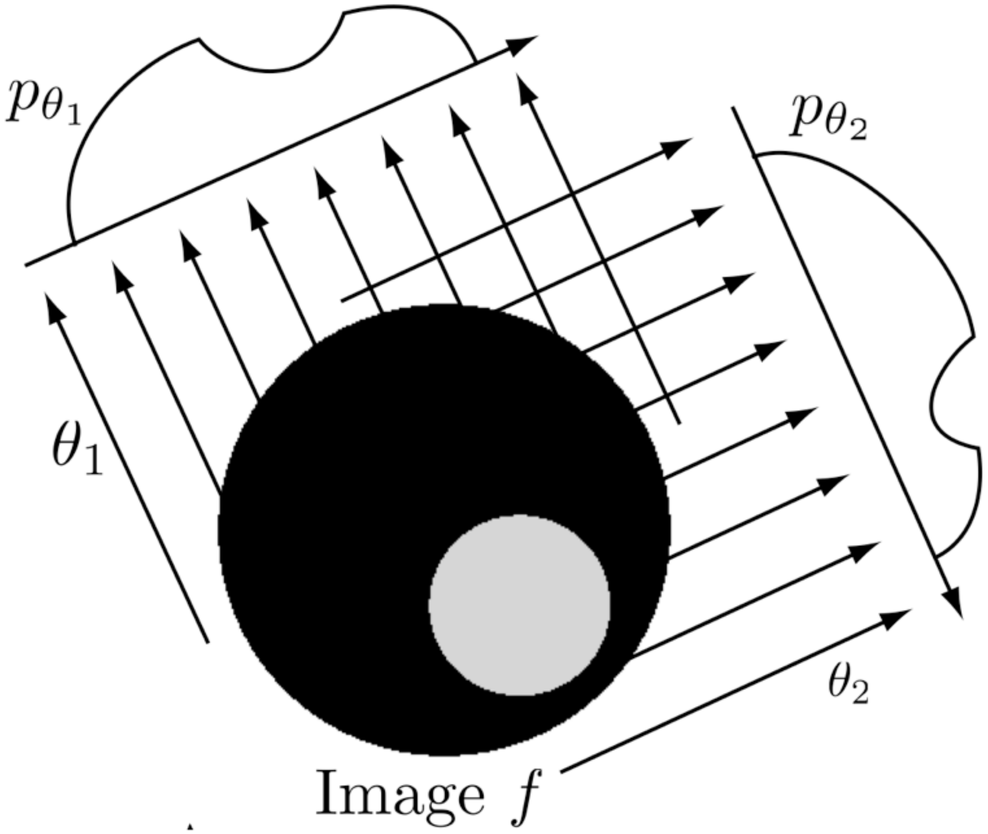
then $\delta_{2k} \leqslant \sqrt{2} - 1$ with high probability.

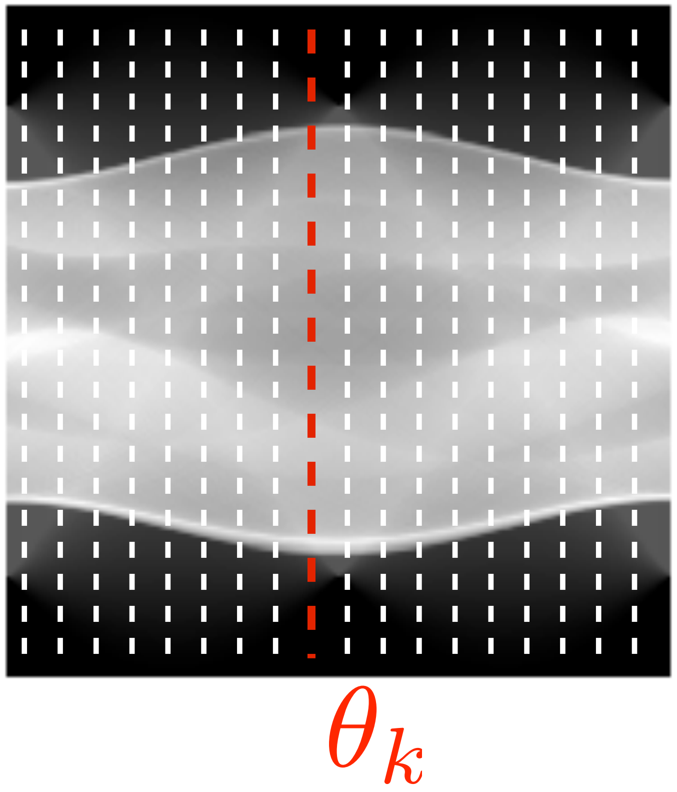
Overview

- Shannon's World
- Compressive Sensing Acquisition
- Compressive Sensing Recovery
- Theoretical Guarantees
- Fourier Domain Measurements

Tomography and Fourier Measures

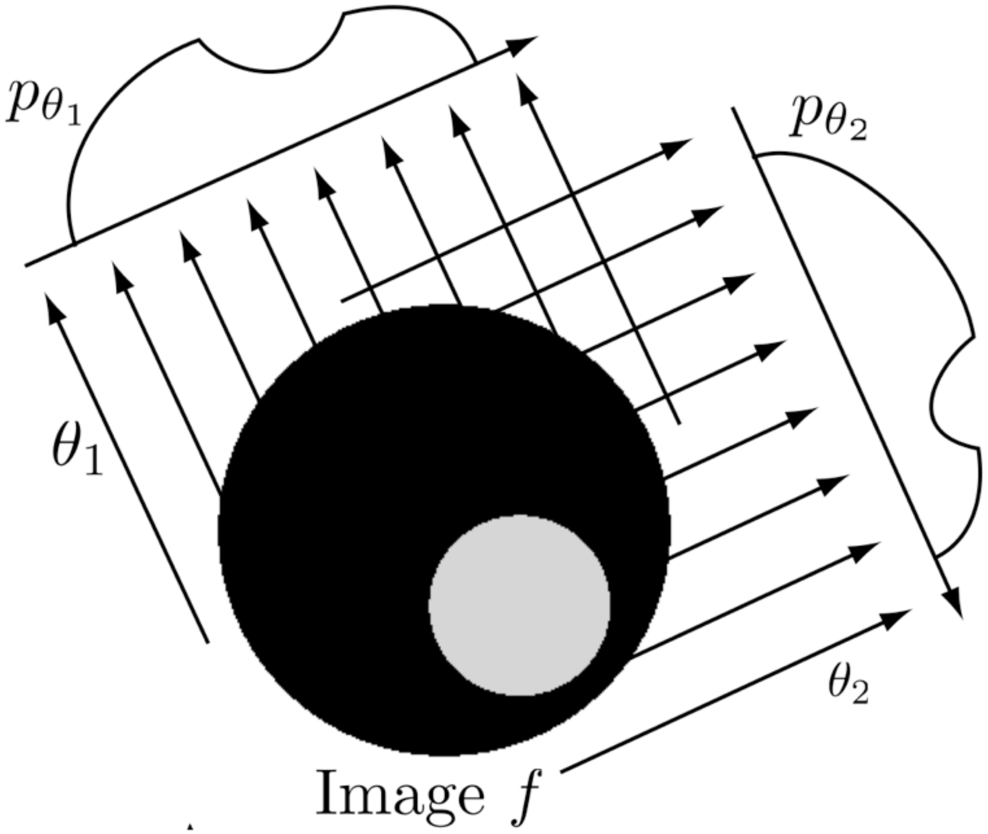
Tomography projection:

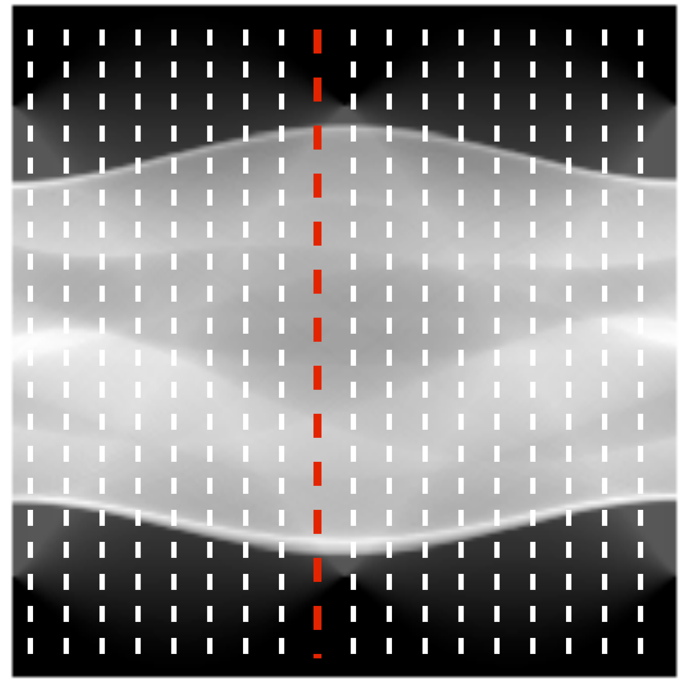




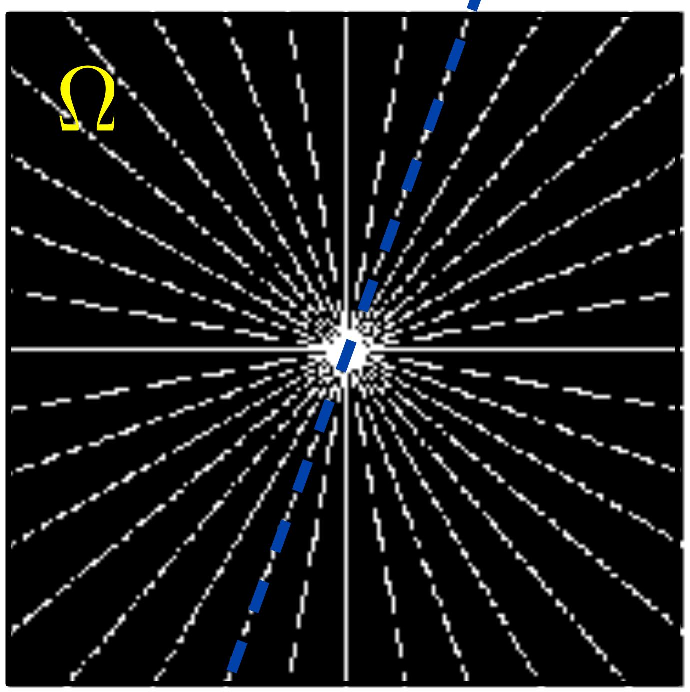
Tomography and Fourier Measures

Tomography projection:





$\hat{f} = \text{FFT2}(f)$



Fourier slice theorem: $\underbrace{\hat{p}_\theta(\rho)}_{\text{1D}} = \underbrace{\hat{f}(\rho\cos(\theta), \rho\sin(\theta))}_{\text{2D Fourier}}$

Partial Fourier measurements: $\{p_{\theta_k}(t)\}_{0 \leqslant k < K}^{t \in \mathbb{R}}$

Equivalent to: $\mathcal{K}f = (\hat{f}[\omega])_{\omega \in \Omega}$

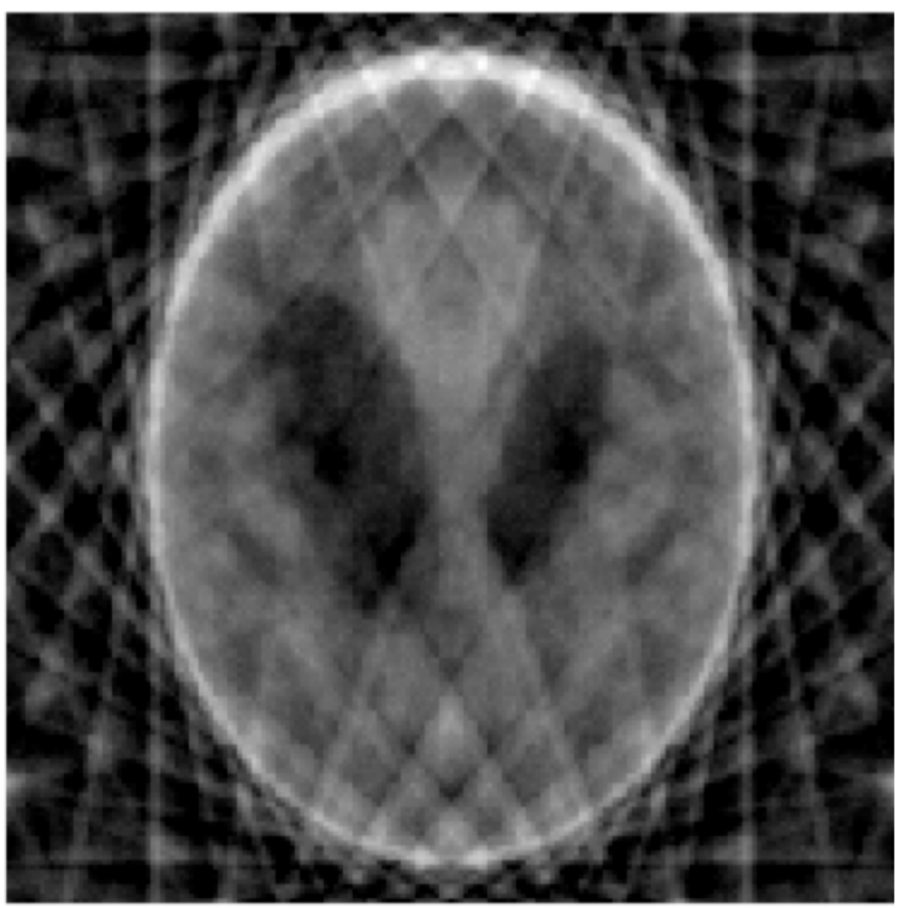
Regularized Inversion

Noisy measurements: $\forall\, \omega \in \Omega,\ y[\omega] = \hat{f}_0[\omega] + w[\omega].$

Noise: $w[\omega] \sim \mathcal{N}(0, \sigma)$, white noise.

ℓ^1 regularization:

$$f^\star = \underset{f}{\text{argmin}}\ \frac{1}{2} \sum_{\omega \in \Omega} |y[\omega] - \hat{f}[\omega]|^2 + \lambda \sum_m |\langle f,\ \psi_m \rangle|.$$

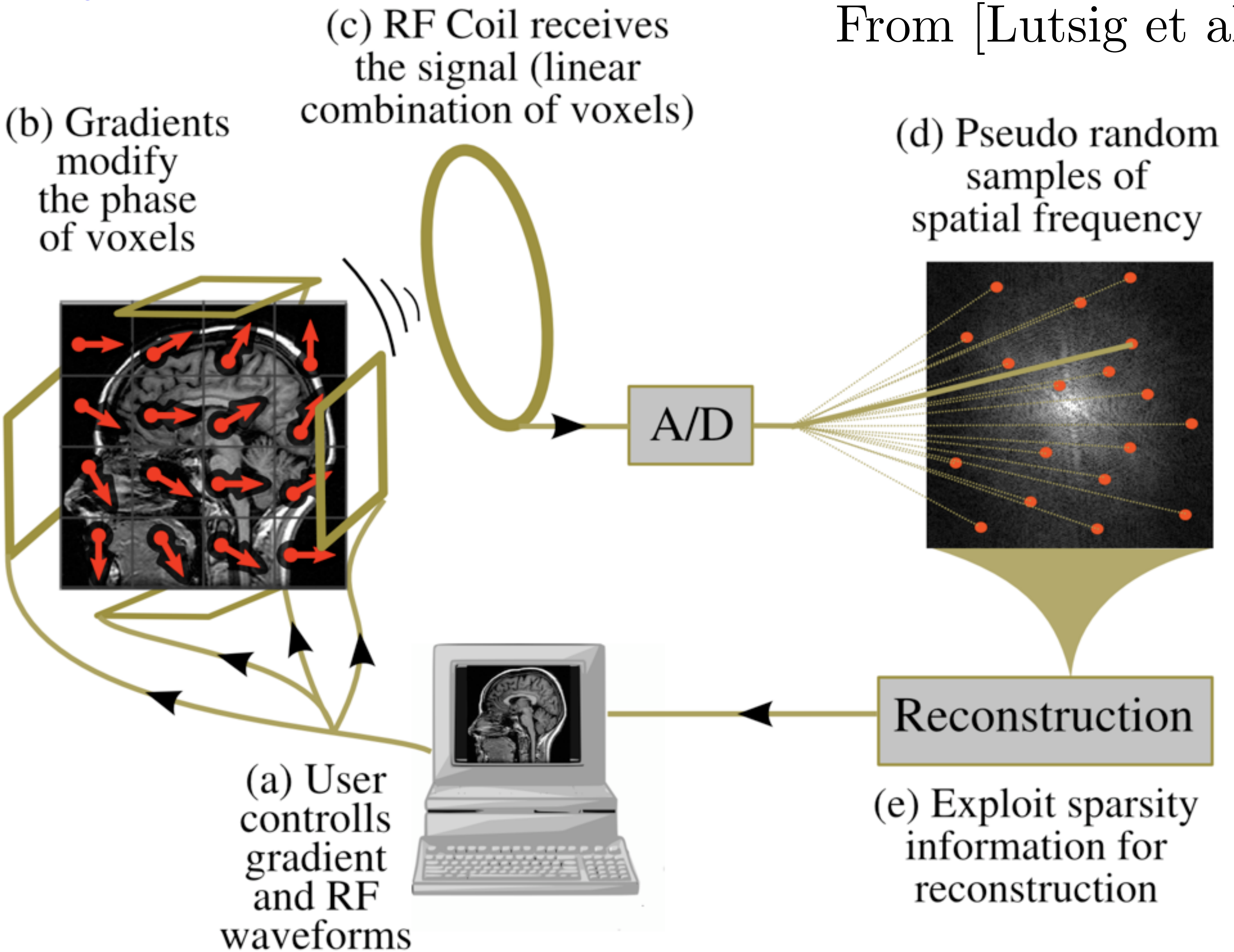


f^+

$f^\star$

MRI Imaging

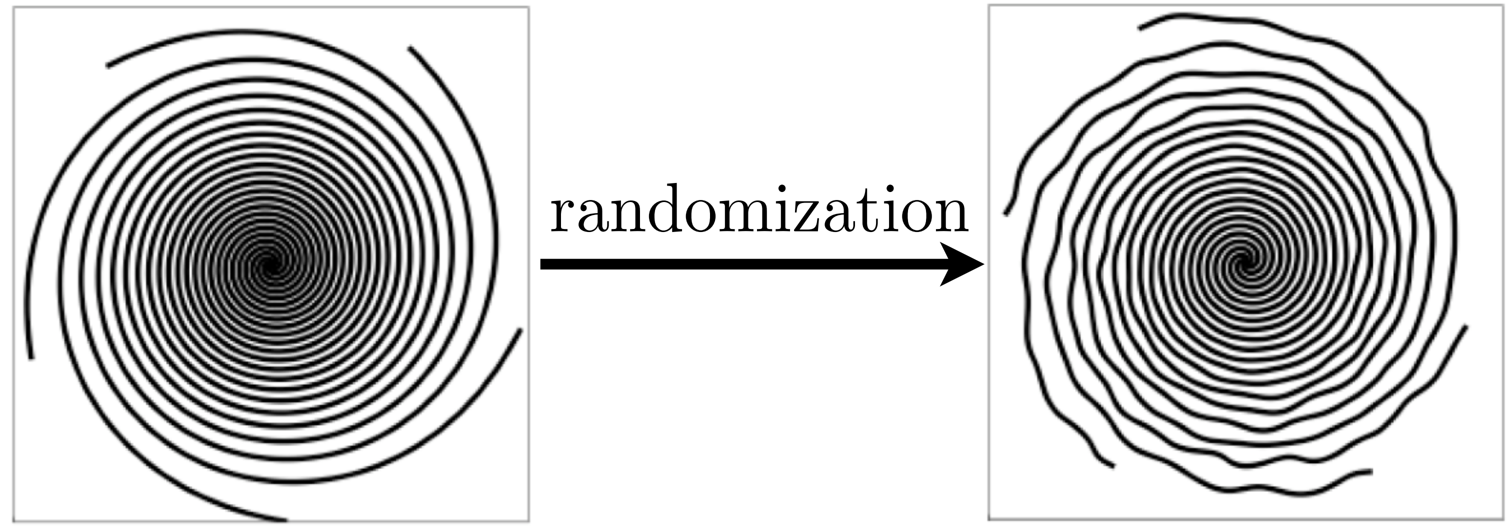
From [Lutsig et al.]

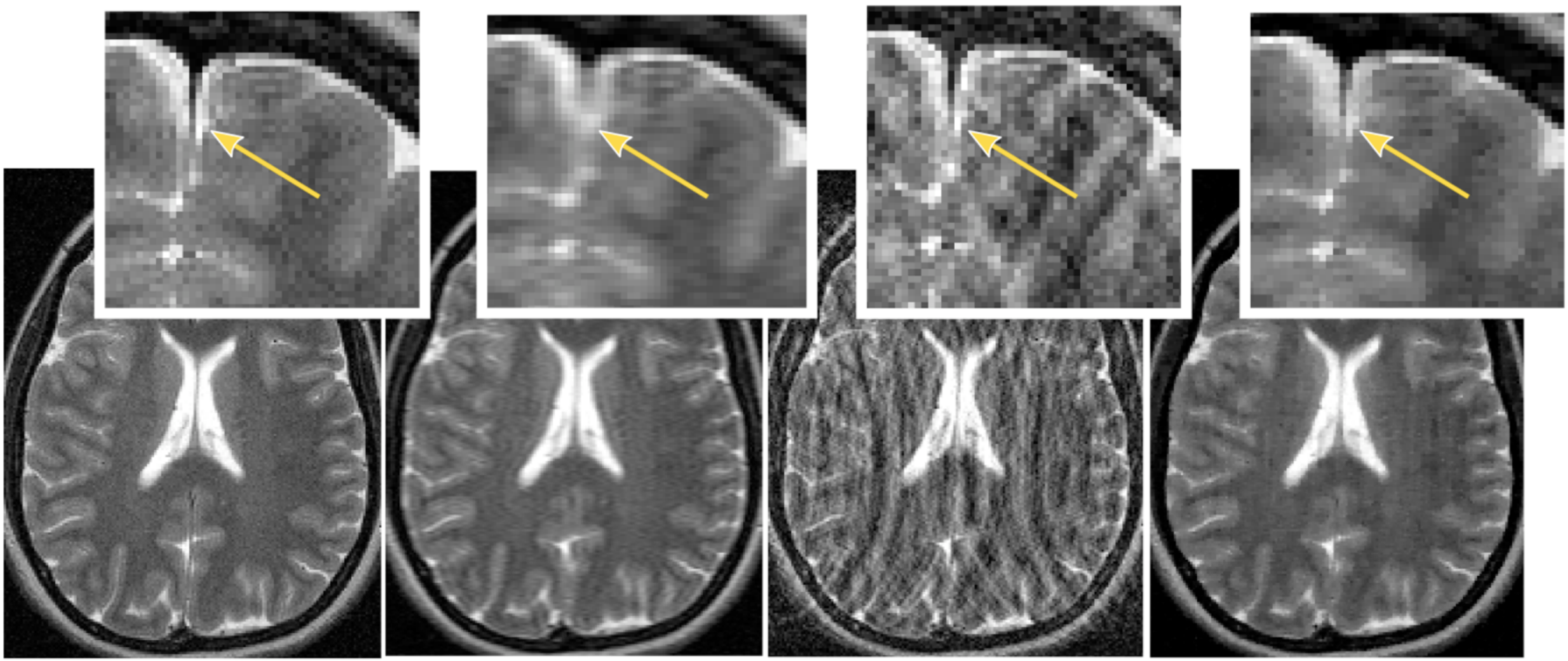


MRI Reconstruction

From [Lutsig et al.]

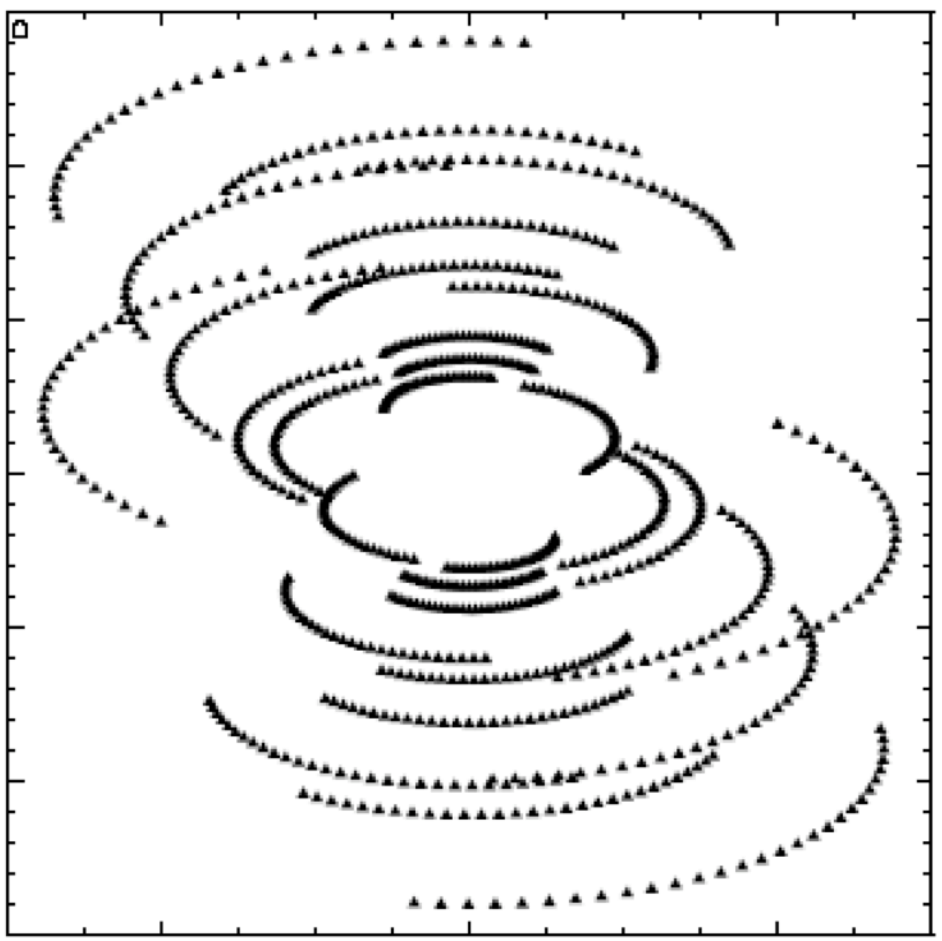
Fourier sub-sampling pattern:



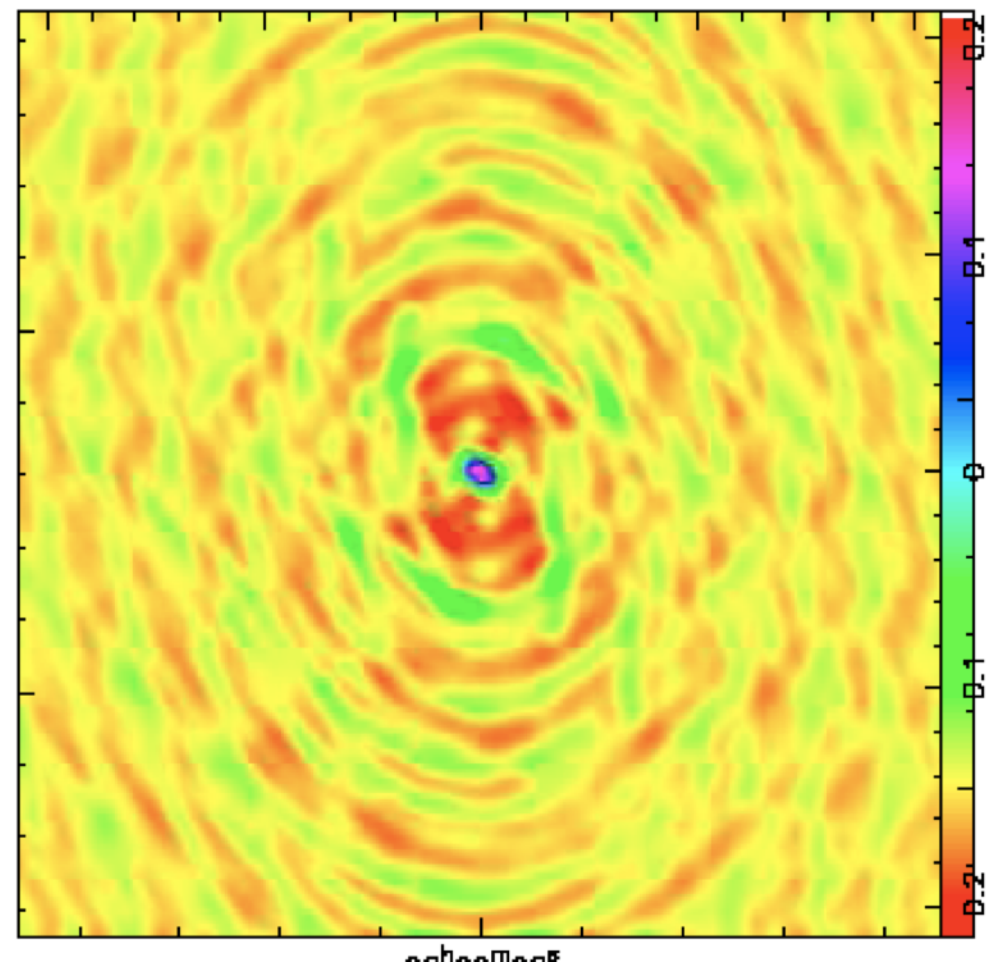


High resolution | Low resolution | Linear | Sparsity

Radar Interferometry



Fourier sampling
(Earth's rotation)



Linear
reconstruction

Structured Measurements

Gaussian matrices: intractable for large N.

Random partial orthogonal matrix: $\{\varphi_\omega\}_\omega$ orthogonal basis.

$$\mathcal{K}f = (\langle \varphi_\omega, f \rangle)_{\omega \in \Omega} \quad \text{where } |\Omega| = P \text{ uniformly random.}$$

Fast measurements: (e.g. Fourier basis)

Structured Measurements

Gaussian matrices: intractable for large N.

Random partial orthogonal matrix: $\{\varphi_\omega\}_\omega$ orthogonal basis.

$$\mathcal{K}f = (\langle \varphi_\omega, f \rangle)_{\omega \in \Omega} \quad \text{where } |\Omega| = P \text{ uniformly random.}$$

Fast measurements: (e.g. Fourier basis)

Mutual incoherence: $\mu = \sqrt{N} \max_{\omega, m} |\langle \varphi_\omega, \psi_m \rangle| \in [1, \sqrt{N}]$

Fourier/Diracs: $\mu = 1$. Wavelets/noiselets: $\mu \approx 1$.

Structured Measurements

Gaussian matrices: intractable for large N.

Random partial orthogonal matrix: $\{\varphi_\omega\}_\omega$ orthogonal basis.

$$\mathcal{K}f = (\langle \varphi_\omega,\, f\rangle)_{\omega\in\Omega} \quad \text{where } |\Omega| = P \text{ uniformly random.}$$

Fast measurements: (e.g. Fourier basis)

Mutual incoherence: $\mu = \sqrt{N} \max_{\omega,m} |\langle \varphi_\omega,\, \psi_m\rangle| \in [1, \sqrt{N}]$

Fourier/Diracs: $\mu = 1$. Wavelets/noiselets: $\mu \approx 1$.

Theorem: with high probability on Ω, $\Phi = \mathcal{K}\Psi$

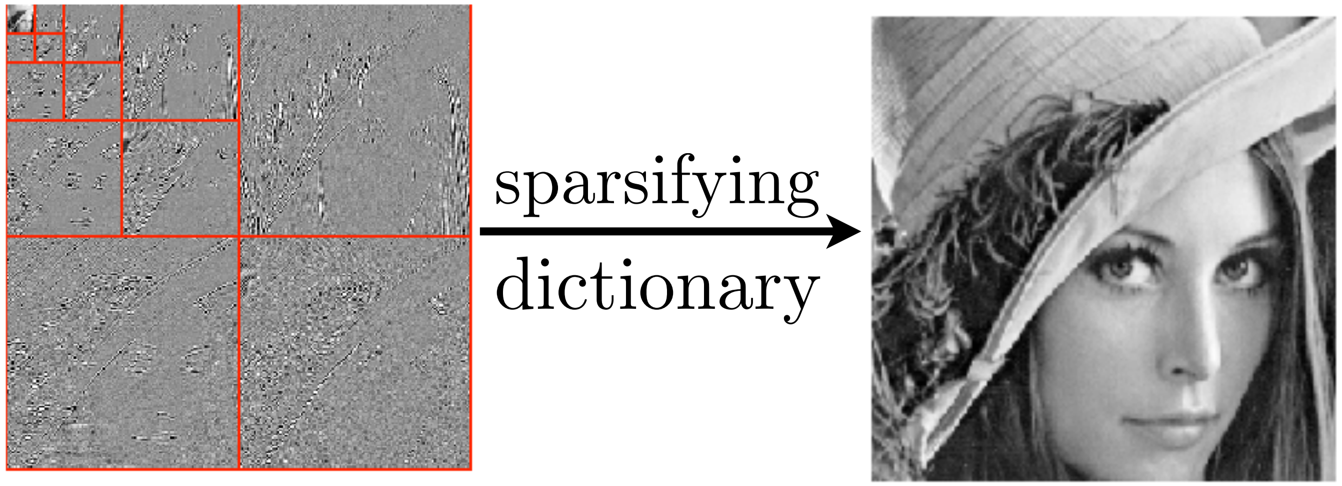
$$\text{If } M \leqslant \frac{CP}{\mu^2 \log(N)^4}, \quad \text{then } \delta_{2M} \leqslant \sqrt{2} - 1$$

[Rudelson, Vershynin, 2006]

$\longrightarrow$ not universal: requires incoherence.

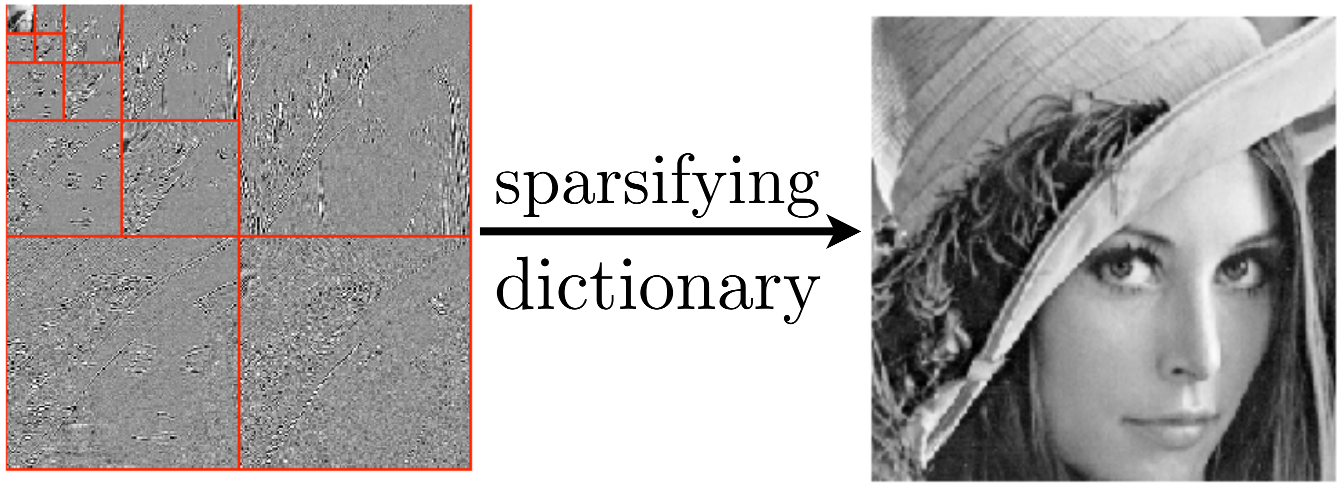
Conclusion

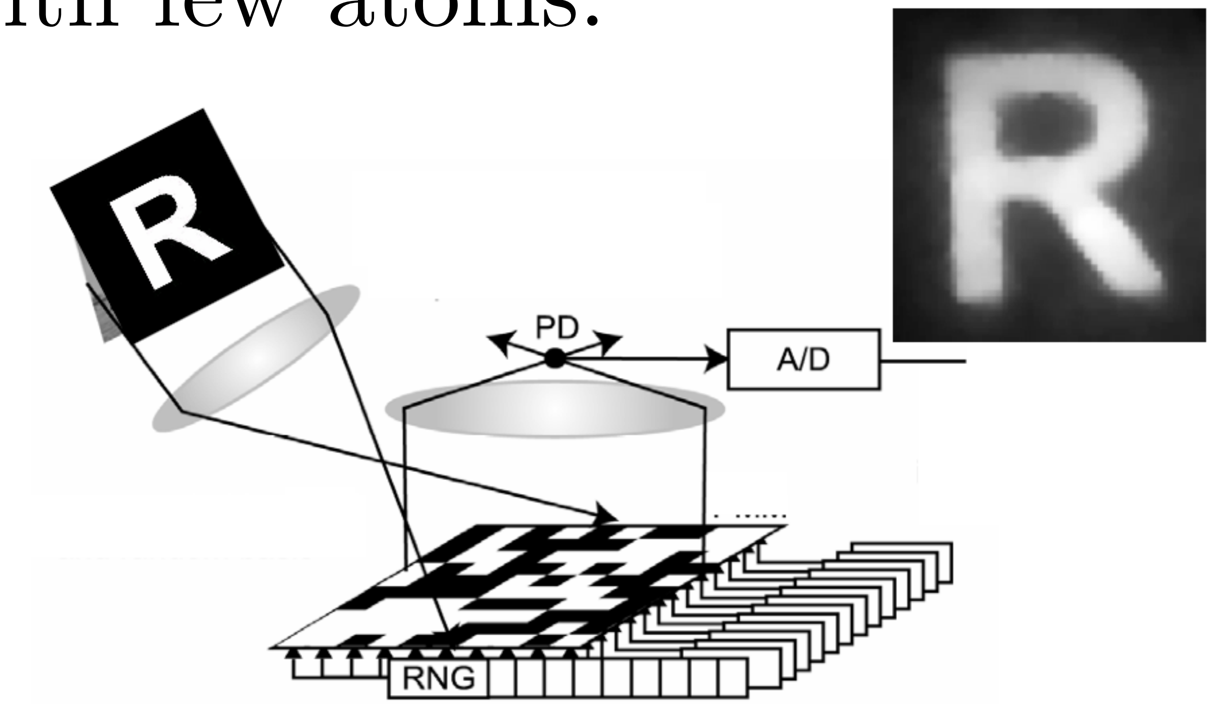
Sparsity: approximate signals with few atoms.



Conclusion

Sparsity: approximate signals with few atoms.





Compressed sensing ideas:

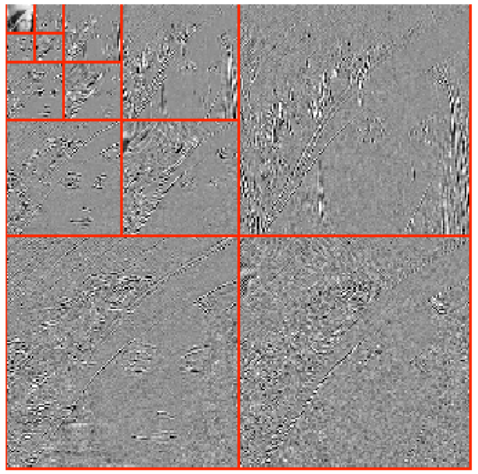
$\longrightarrow$ Randomized sensors + sparse recovery.

$\longrightarrow$ Number of measurements $\approx$ signal complexity.

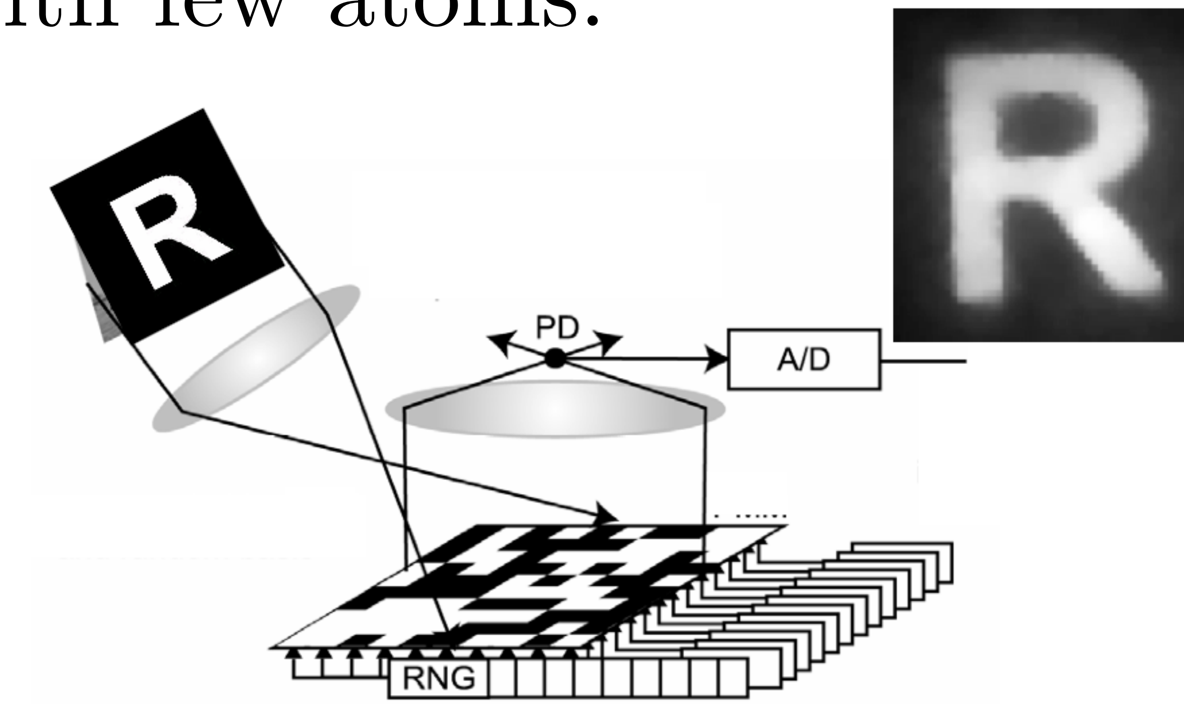
$\longrightarrow$ CS is about designing new hardware.

Conclusion

Sparsity: approximate signals with few atoms.



sparsifying
dictionary



Compressed sensing ideas:

⟶ Randomized sensors + sparse recovery.

⟶ Number of measurements ≈ signal complexity.

⟶ CS is about designing new hardware.

The devil is in the constants:

⟶ Worse case analysis is problematic.

⟶ Designing good signal models.