UCLA David Geffen School of Medicine | UCLA Health | Radiology | UCLA Magnetic Resonance Research Labs

Principles of Image Reconstruction

Anthony G. Christodoulou, PhD
Department of Radiological Sciences
David Geffen School of Medicine
University of California, Los Angeles

M219
05 March 2025

Lecture goals

Goals for today:

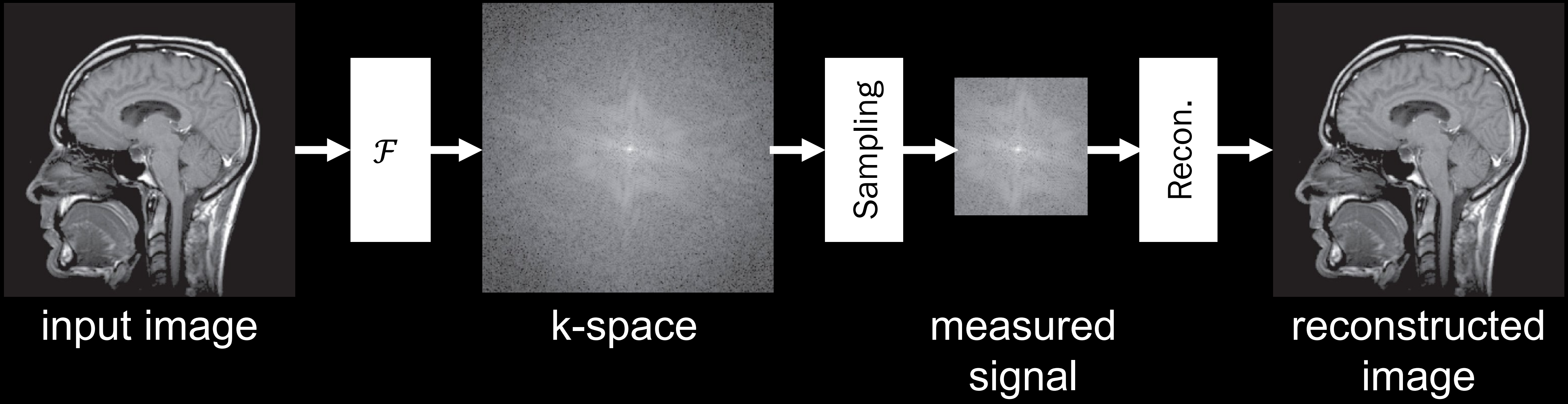
- Review and build upon image reconstruction methods you have previously seen
  - (Fourier reconstruction, parallel imaging)

- Introduce formal principles of image reconstruction
  - Conditions for solution existence
  - Uniqueness of solutions
  - Probabilistic interpretations of data and images

- Increase understanding of advanced techniques (e.g., compressed sensing)

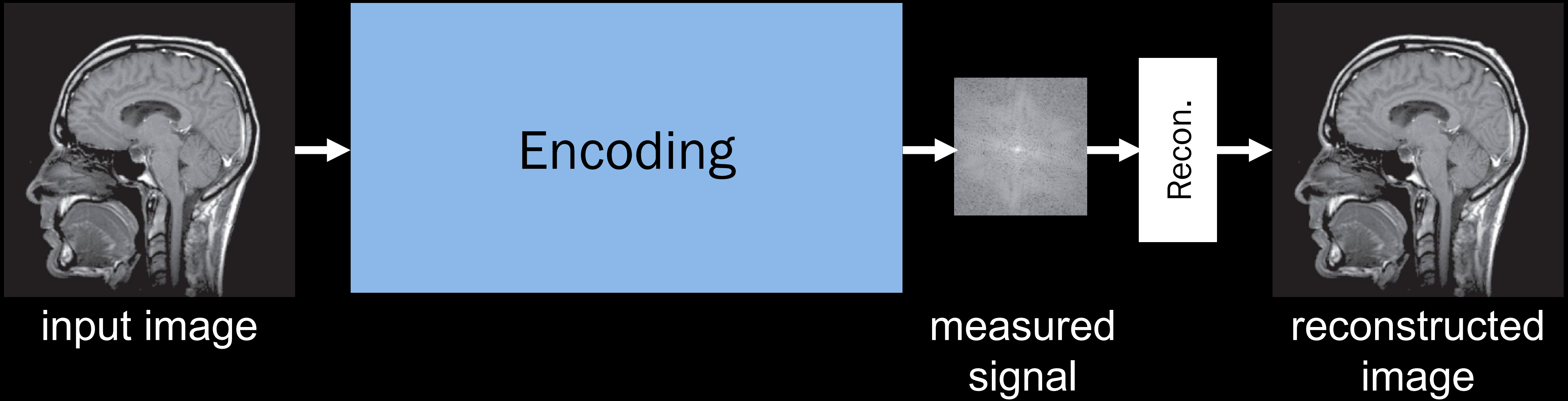
Continuous domain image reconstruction

System view

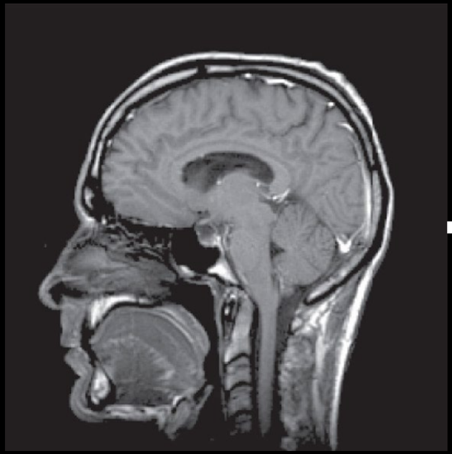


$\mathcal{F}$
Sampling
Recon.
input image
k-space
measured signal
reconstructed image

System view

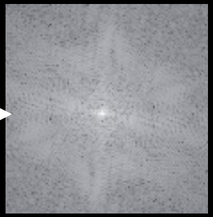

Encoding
Recon.
input image
measured signal
reconstructed image

Inverse problem view



input image

Encoding

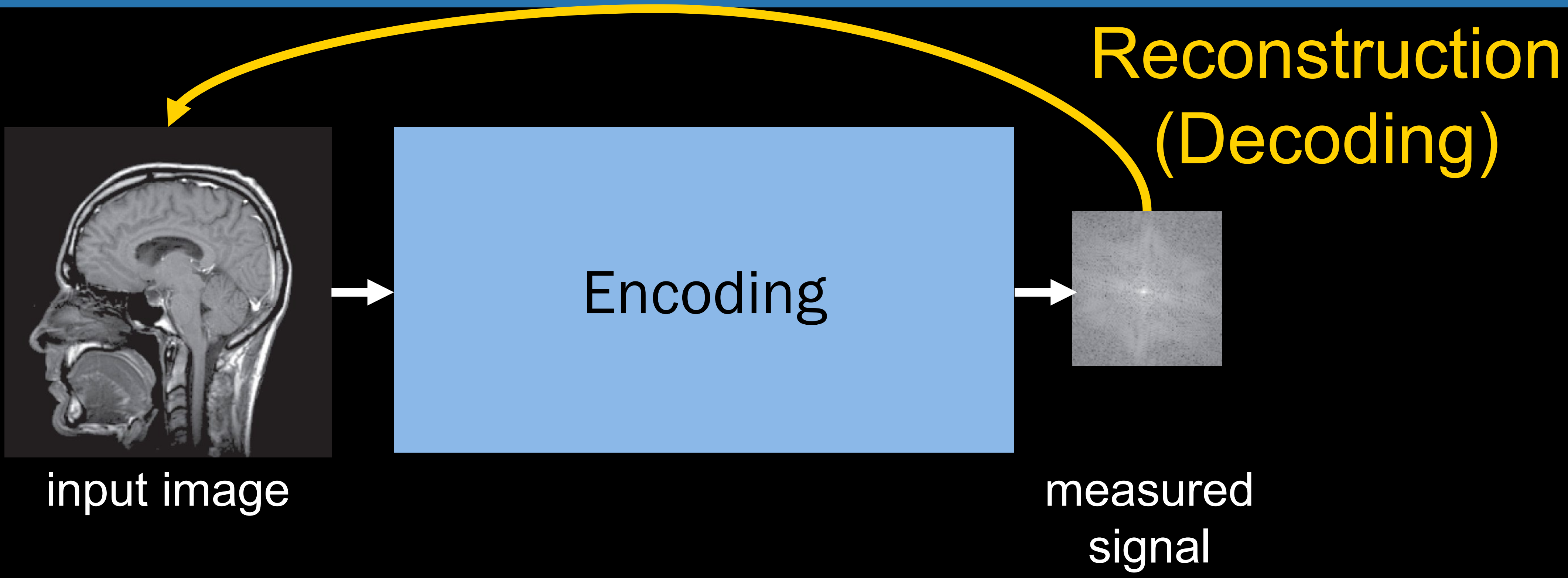


measured signal

Forward problem: what is signal given input?

Inverse problem: what was input given signal?

Inverse problem view



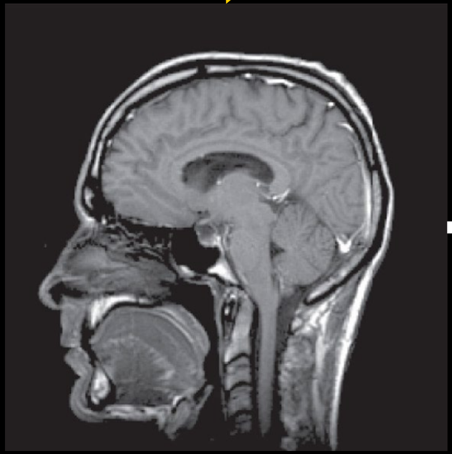
Forward problem: what is <u>signal</u> given <u>input</u>?

Inverse problem: what was <u>input</u> given <u>signal</u>?

Inverse problem view

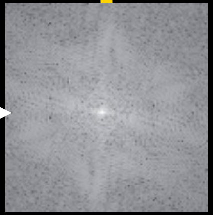
Reconstruction operator

$\mathcal{T}^{-1}\{\cdot\}$



Encoding transform

$\mathcal{T}\{\cdot\}$



input image

$I(\vec{r})$

signal

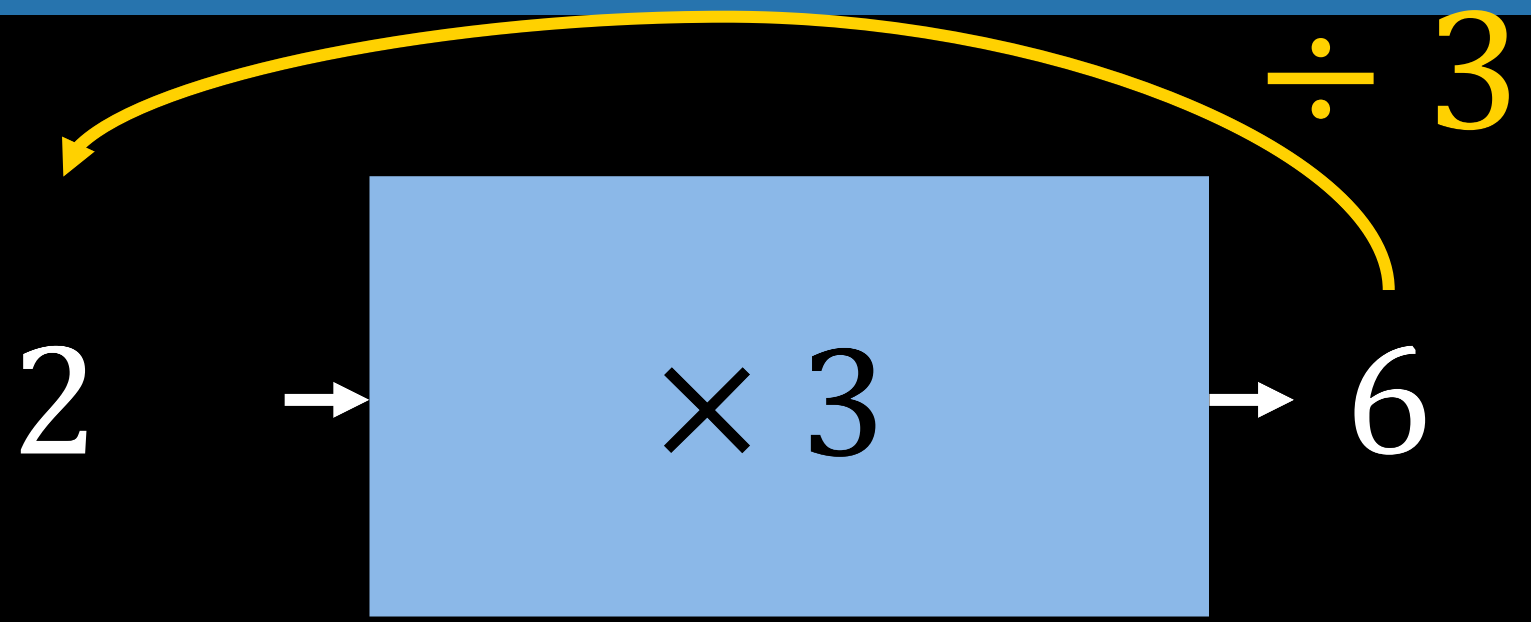
$S(\vec{k})$

Forward problem: $\mathcal{T}\{I\} = ?$

Inverse problem: $\mathcal{T}\{?\} = d$

$? = \mathcal{T}^{-1}\{d\}$

Going from arithmetic to algebra



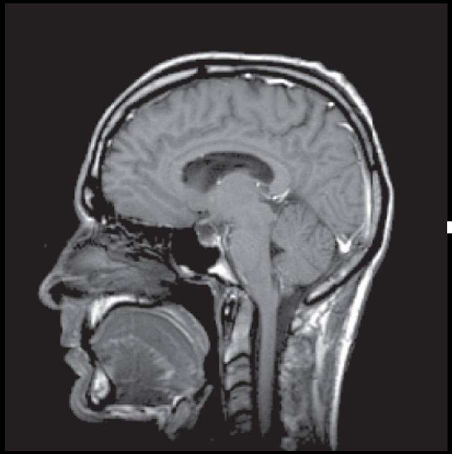
Forward problem: $3(2) = ?$

Inverse problem: $3x = 6$

$x = 6 \div 3 = 2$

Inverse problem view

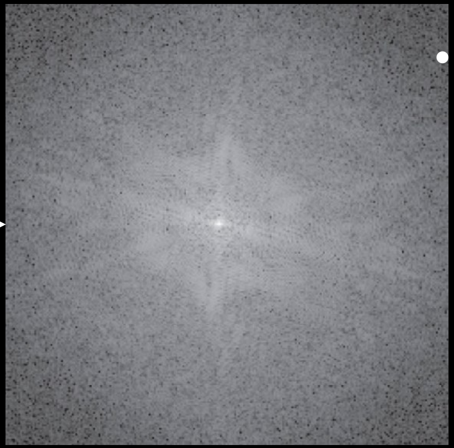
Can we do this in MRI?



input image

$I(\vec{r})$

→ $\mathcal{T}\{\cdot\}$ →



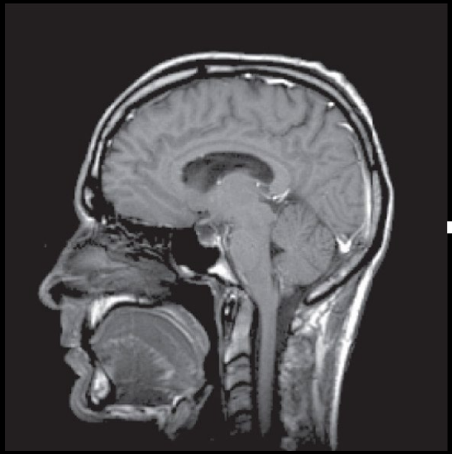
measured signal

$S(\vec{k})$

- Measured within finite window

Inverse problem view

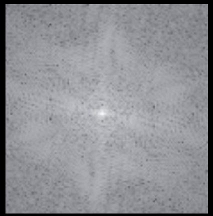
Can we do this in MRI?



input image

$I(\vec{r})$

$$\mathcal{T}\{\cdot\} = \mathcal{F}\{\cdot\}\mathrm{rect}(\vec{k})$$



measured signal

$S(\vec{k})$

- Measured within finite window
- (Full k-space) × (rect function)
- rect function has zeroes!

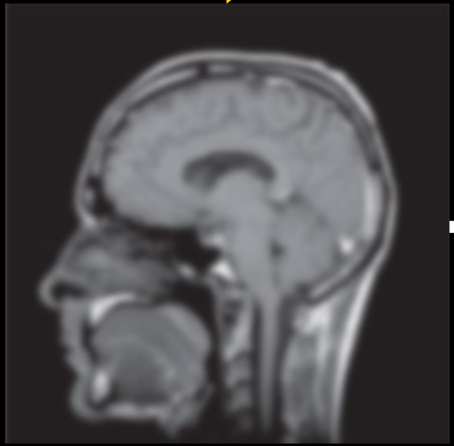
$$0x = 0$$
$$x = 0 \div 0$$

cannot recover original x

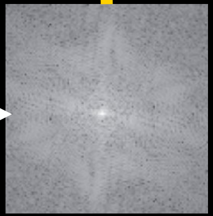
An inverse function $\mathcal{T}^{-1}\{\cdot\}$ doesn't always exist!

Inverse problem view

Reconstruction operator



$$\mathcal{T}\{\cdot\} = \mathcal{F}\{\cdot\}\mathrm{rect}(\vec{k})$$



reconstructed image

$\hat{I}(\vec{r}) = I(\vec{r}) * \mathrm{sinc}(\vec{r})$

measured signal

$S(\vec{k})$

Feasible solution

Encoding: $\mathcal{T}\{I\} \rightarrow S$

Reconstruction: $Recon\{S\} \rightarrow \hat{I}$

What if you re-encode $\hat{I}$? What does $\mathcal{T}\{\hat{I}\}$ equal?

Feasible solution

Fourier encoding: $\mathcal{F}\{I\} \cdot \mathrm{rect} \rightarrow S$

Fourier reconstruction: $\mathcal{F}^{-1}\{S\} \rightarrow \hat{I} = I * \mathrm{sinc}$

What if you re-encode $\hat{I}$? What does $\mathcal{T}\{\hat{I}\}$ equal?

$$\mathcal{T}\{\hat{I}\} = \mathcal{T}\{I * \mathrm{sinc}\} = \mathcal{F}\{I * \mathrm{sinc}\} \cdot \mathrm{rect} = \mathcal{F}\{I\} \cdot \mathrm{rect} = S$$

Our reconstruction $\hat{I}$ did not recover the original image I,

but $\hat{I}$ is exactly consistent with the measured signal: $\mathcal{T}\{\hat{I}\} = S$

$\hat{I}(\vec{r})$ is a <u>feasible solution</u>

Image reconstruction objectives

Objective of feasible image reconstruction:

- Reconstruct an image which is consistent with the data

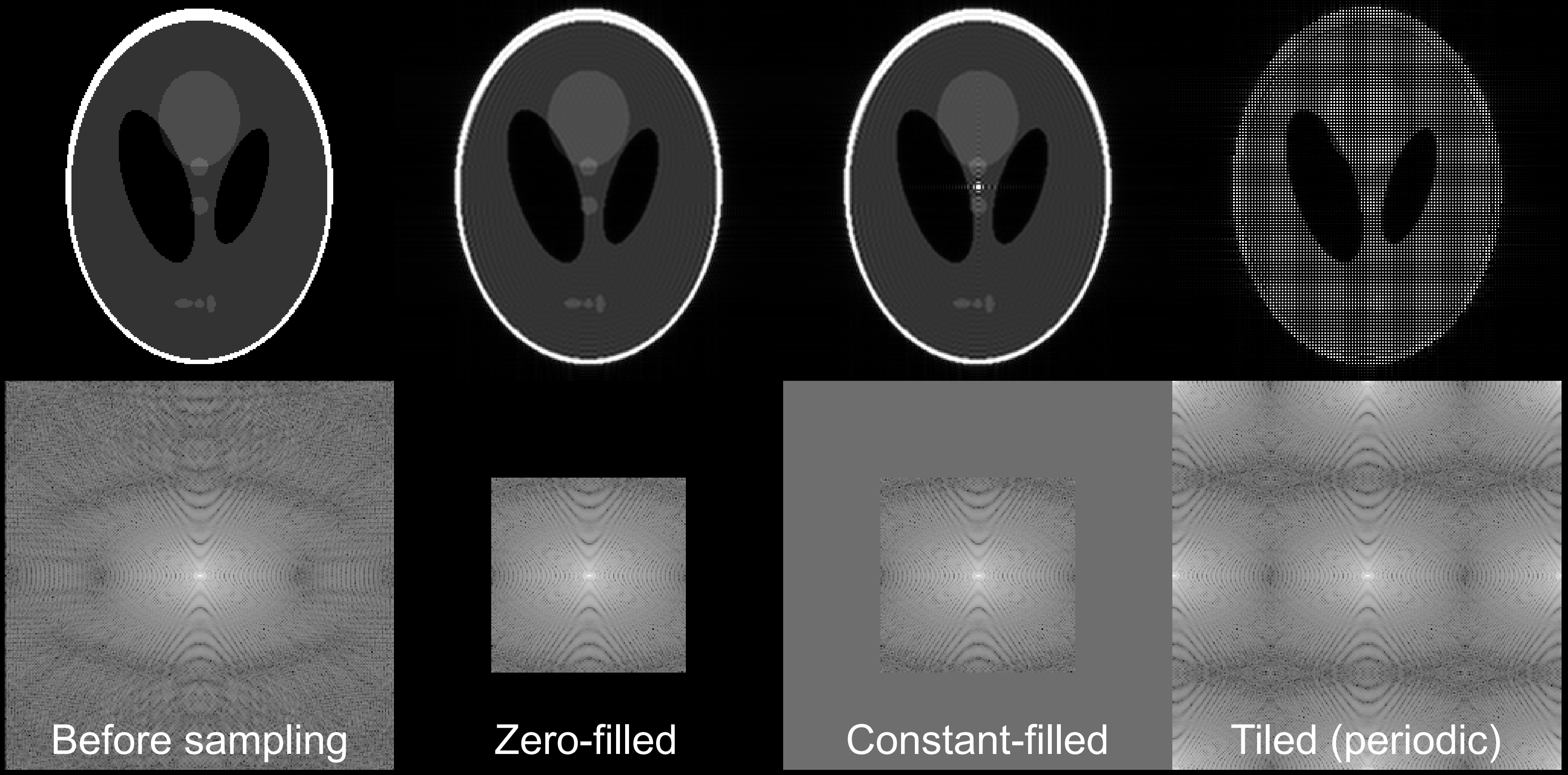
More formally:

- Find an image $\hat{I}$ such that $\mathcal{T}\{\hat{I}\} = S$

Feasible solution(s)

How many images satisfy $\mathcal{T}\{\hat{I}\} = S$?

- Infinite! We can put anything outside our measured region and retain feasibility



Feasible solution

How many images satisfy $\mathcal{T}\{\hat{I}\} = S$?

- Infinite! We can put anything outside our measured region and retain feasibility

In the continuous domain,

feasible solution is not unique

However! Some feasible solutions are "better" than others

- The solution assuming zeros outside measured region is the *minimum norm solution*

Image reconstruction objectives

Objective of feasible image reconstruction:

- Find an image $\hat{I}$ such that $\mathcal{T}\{\hat{I}\} = S$

With infinite solutions, we need a **<u>second objective</u>** as well, e.g.

- Of all the images $\hat{I}$ such that $\mathcal{T}\{\hat{I}\} = S$, choose the one with minimum norm $\|\hat{I}(\vec{r})\| = \sqrt{\int |I(\vec{r})|^2 \, d\vec{r}}$
- In other words, pick the "smallest" solution

$$\hat{I} = \arg\min_{I} \|I\| \ \text{ s.t. } \mathcal{T}\{I\} = S$$

"Solution = argument I which minimizes $||I||$ such that $\mathcal{T}\{I\} = S$"

i.e, keep the data you measured and fill the unknown values with zeros!

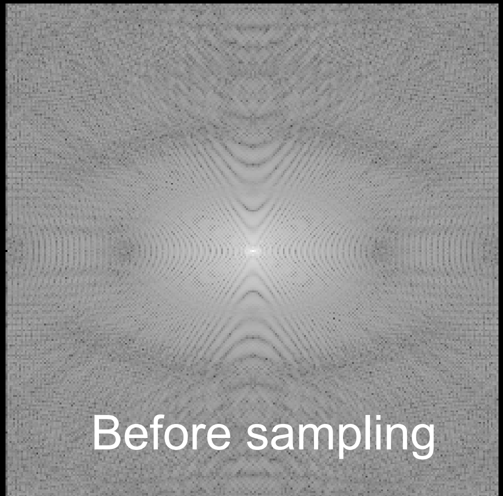
Feasibility is not everything! (e.g., ringing)

Sometimes a second objective is important

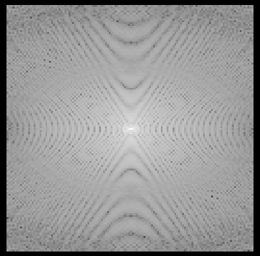
- Additional information/additional goal (e.g., minimize ringing)

$\mathcal{T}\{\hat{I}(\vec{r})\} = s(\vec{k})$

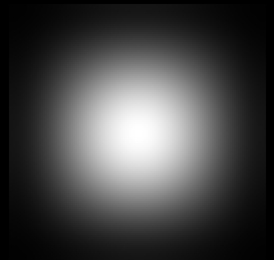
$\mathcal{T}\{\hat{I}(\vec{r})\} = w(\vec{k})s(\vec{k})$



Before sampling



Feasible, minimum-norm



Hamming windowed by $w(\vec{k})$

Feasibility is not everything! (e.g., noise)

Data are corrupted by noise

- "Perfect" noiseless reconstruction is not "feasible"

$$\mathcal{T}\{I\} + N = S, \qquad \text{where } N \text{ is noise distributed according to } \mathcal{N}(0, \sigma^2)$$

$$S - \mathcal{T}\{I\} = N$$

Can modify data consistency objective

- Noiseless: Find an image $\hat{I}$ such that $\mathcal{T}\{\hat{I}\} = S$
- Noisy: Find an image $\hat{I}$ which minimizes $\left\|S - \mathcal{T}\{\hat{I}\}\right\|^2 = \|N\|^2$
    - Has maximum likelihood interpretation for additive white Gaussian noise (AWGN)

$$\hat{I} = \arg\min_{I} \left\|S - \mathcal{T}\{\hat{I}\}\right\|^2$$

"Least-squares solution". This will still produce a feasible solution if one exists!

Maximum likelihood interpretation

Each measured data point is a Gaussian RV: $S \sim \mathcal{N}(\mathcal{T}\{I\}, \sigma^2)$ $\quad p(z) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|z-\mu|^2}{2\sigma^2}}$

Likelihood (probability of signal given image): $\mathcal{L}(I|S) = p(S|I) = \prod_{N_k} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|S-\mathcal{T}\{I\}|^2}{2\sigma^2}}$

Maximum likelihood: $\arg\max_I \mathcal{L}(I|d) = \arg\max_I \prod_{N_k} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|S-\mathcal{T}\{I\}|^2}{2\sigma^2}}$

Maximum log-likelihood: $\arg\max_I \mathcal{L}(I|d) = \arg\max_I \sum_{N_k} \left[\left(\log \frac{1}{\sqrt{2\pi\sigma^2}}\right) - \frac{|S - \mathcal{T}\{I\}|^2}{2\sigma^2}\right]$

$$= \arg\min_I \sum_{N_k} |S - \mathcal{T}\{I\}|^2 = \boxed{\arg\min_I \|S - \mathcal{T}\{I\}\|^2}$$

Discrete domain image reconstruction

Discrete-to-discrete inverse problem

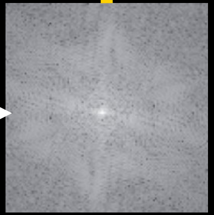
If we accept the resolution limit, we can re-frame the goal of image reconstruction:

- Recover discretized version of $\hat{I} = I * \mathrm{sinc}$ (instead of continuous I)

Reconstruction operator $\mathbf{E}^{-1}$

desired image $\hat{I}[\vec{r}]$

Discrete-to-discrete encoding operator $\mathbf{E}$



signal $s[\vec{k}]$

- $\hat{I}$ is feasible, so $\mathbf{E}$ still generates exact measured data
- $\mathbf{E}^{-1}$ may now exist, as it is not trying to undo resolution change

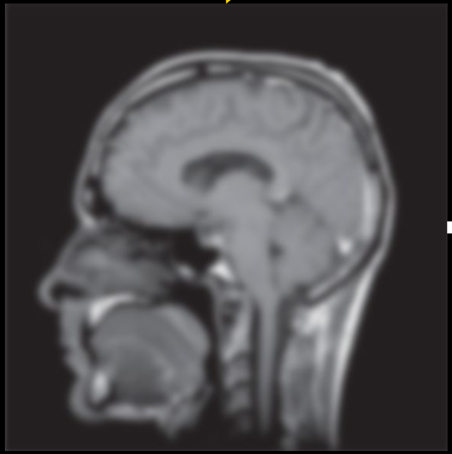
Matrix-vector inverse problem

When encoding is a linear operation (like Fourier encoding), it can be described by matrix multiplication…

…it does not have to be implemented by matrix multiplication (e.g. FFT implementation of DFT matrix operation)

Reconstruction operator

$\mathbf{E}^{-1}$

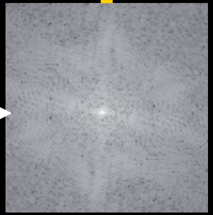


vectorized image

$\mathbf{m}$

Matrix operator

$\mathbf{E}$



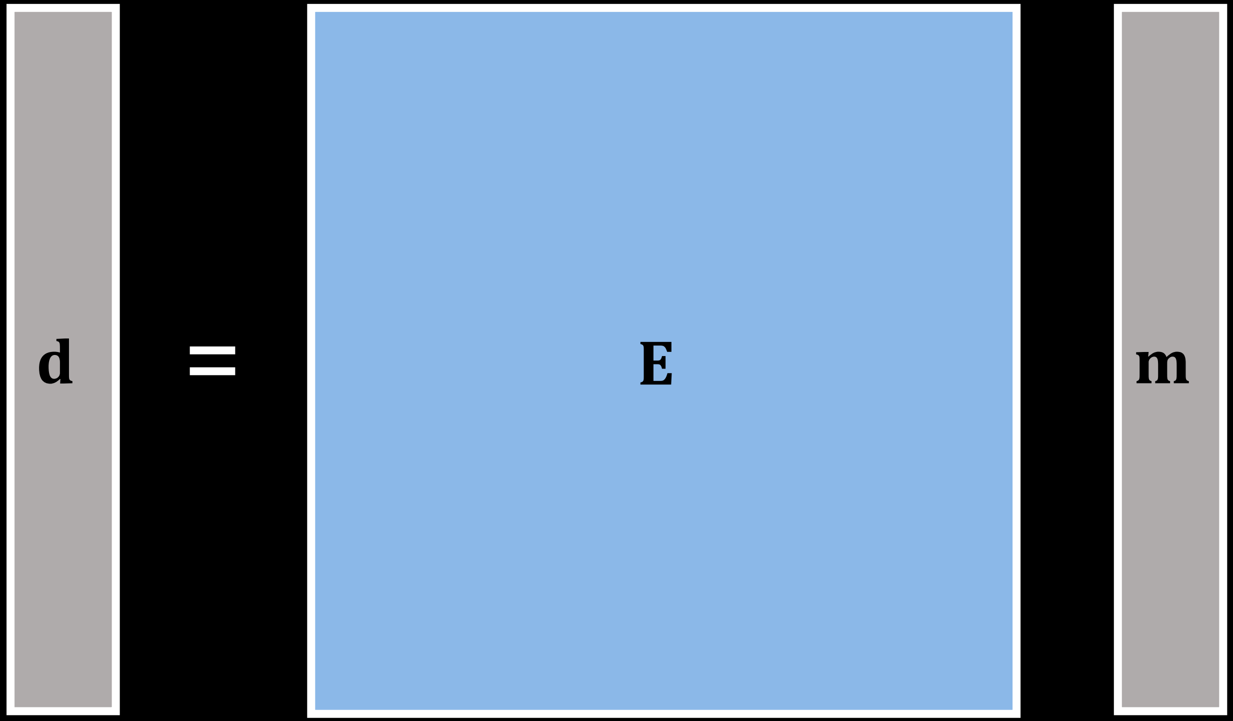
data vector

$\mathbf{d}$

$$\mathbf{d} = \mathbf{E}\mathbf{m}$$

$$\mathbf{m} = \mathbf{E}^{-1}\mathbf{d}$$

When does $\mathbf{E}^{-1}$ exist?

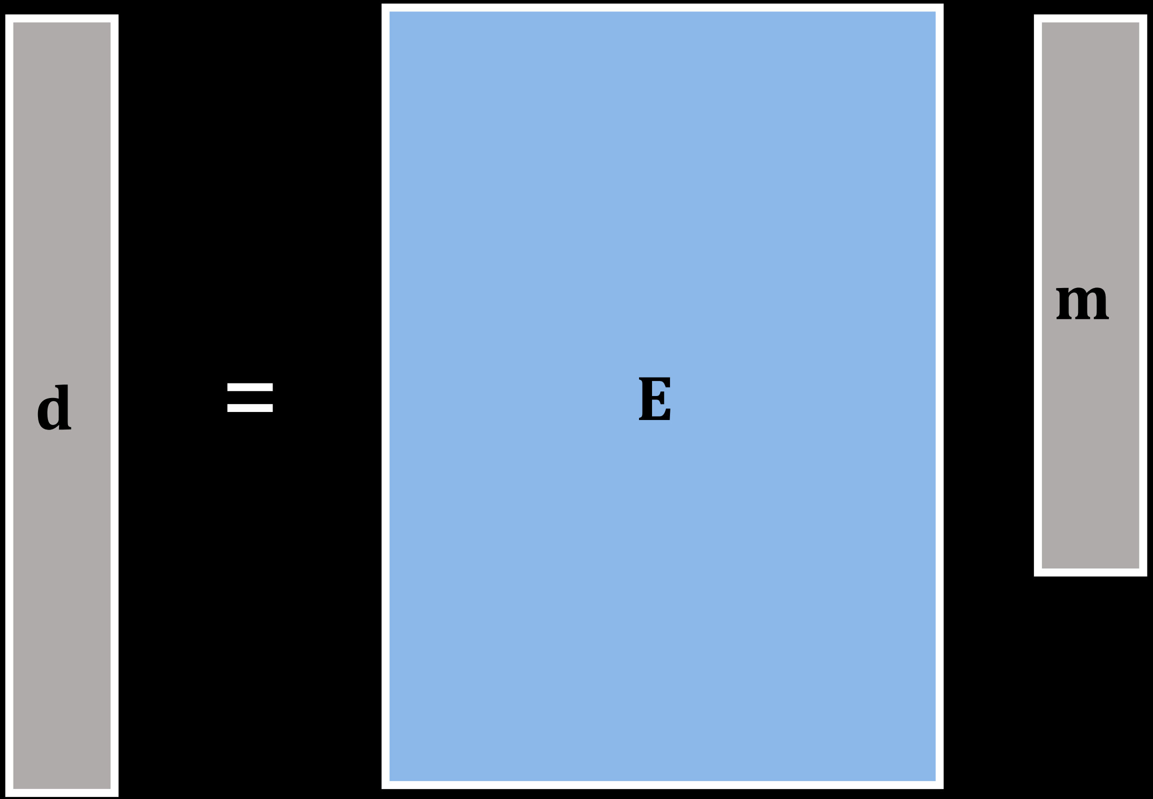


When $\mathbf{E}$ is square (as many data in $\mathbf{d}$ as unknowns in $\mathbf{m}$):

$\mathbf{E}^{-1}$ exists! Unique solution: $\mathbf{m} = \mathbf{E}^{-1}\mathbf{d}$

(Assuming linearly independent rows/columns)

When does $\mathbf{E}^{-1}$ exist?



When $\mathbf{E}$ is “tall” (more data than unknowns): Problem is overdetermined

No $\mathbf{E}^{-1}$ exists. Unique least-squares solution: $\arg\min_{\mathbf{m}} ||\mathbf{d} - \mathbf{Em}||^2$

(Assuming linearly independent columns)

Least-squares solution

$$\tilde{\mathbf{m}} = \arg\min_{\mathbf{m}} ||\mathbf{d} - \mathbf{E}\mathbf{m}||^2$$

$$\mathbf{E}\mathbf{m} = \mathbf{d}$$

$$\mathbf{E}^H\mathbf{E}\tilde{\mathbf{m}} = \mathbf{E}^H\mathbf{d}$$ ($\mathbf{E}^H$ is Hermitian/conjugate transpose)

$$(\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{E}\tilde{\mathbf{m}} = (\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{d}$$

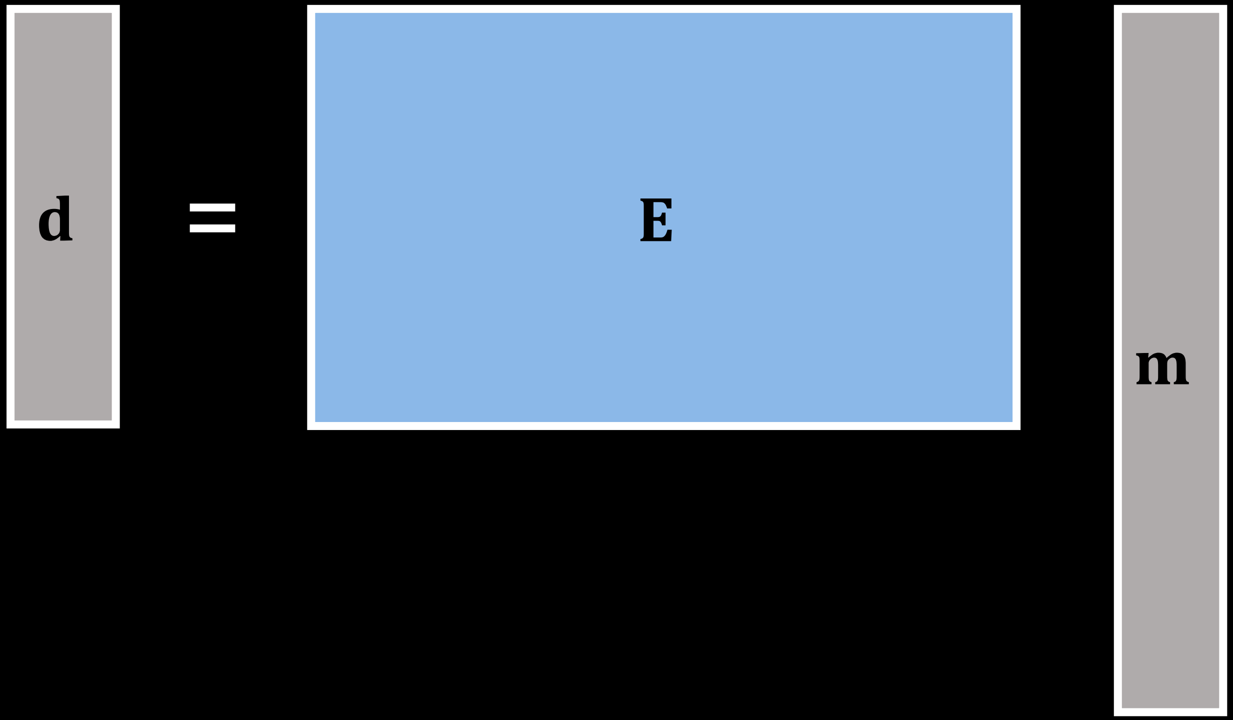
$$\tilde{\mathbf{m}} = (\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{d}$$

$\tilde{\mathbf{m}}$ will have least possible squared error $||\mathbf{d} - \mathbf{E}\tilde{\mathbf{m}}||^2$

$\tilde{\mathbf{m}}$ is unique (when $\mathbf{E}$ has linearly independent columns)

If there is a feasible solution, it is also the least-squares solution: $||\mathbf{d} - \mathbf{E}\tilde{\mathbf{m}}||^2 = 0$

When does $\mathbf{E}^{-1}$ exist?



When $\mathbf{E}$ is “wide” (fewer data than unknowns): Problem is underdetermined (ill-posed)

No $\mathbf{E}^{-1}$ exists. Infinite solutions: $\mathbf{m} \quad s.t. \quad \mathbf{Em} = \mathbf{d}$

(Assuming linearly independent rows)

Underdetermined problem

Some approaches

$$\hat{\mathbf{m}} = \arg\min_{\mathbf{m}} R(\mathbf{m}) \quad \text{s.t.} \quad \mathbf{Em} = \mathbf{d}$$ (force solution to be feasible)

$$\hat{\mathbf{m}} = \arg\min_{\mathbf{m}} ||\mathbf{d} - \mathbf{Em}||^2 + R(\mathbf{m})$$ (allow deviation from data)

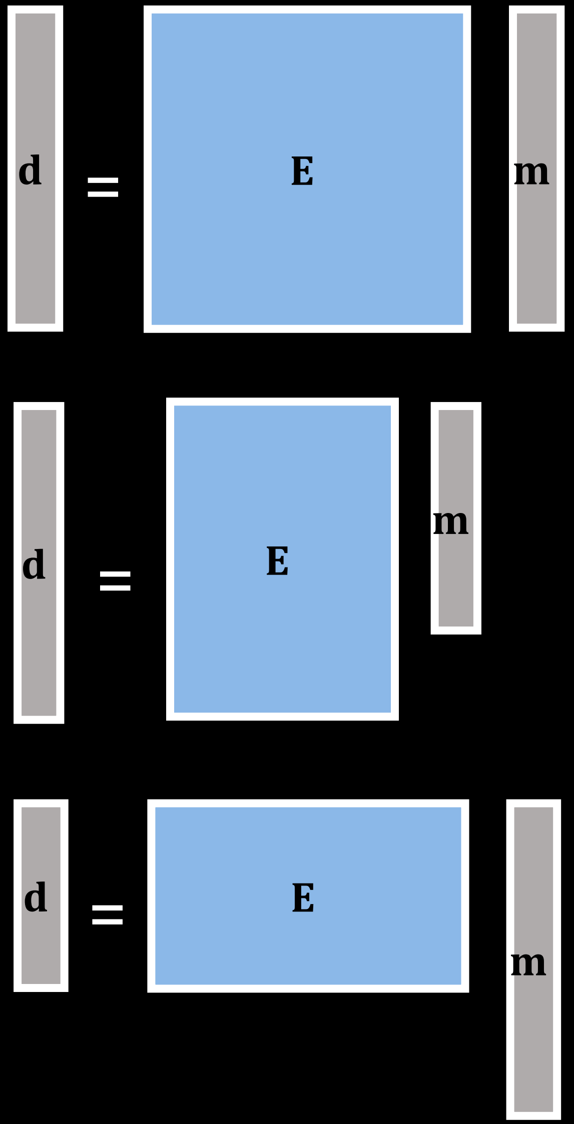
$R(\mathbf{m})$ "regularizes"/constrains the problem

Can enforce other image properties or encourage a probable solution $\mathbf{m}$

Example: $R(\mathbf{m}) = ||\mathbf{m}||^2$: prioritize minimum-norm solution

Foundation of regularized image reconstruction (e.g., compressed sensing)

When does $\mathbf{E}^{-1}$ exist?



Square	Exactly determined	$\mathbf{E}^{-1}$ exists	Feasible solution exists	Solution is unique
Tall	Overdetermined	No $\mathbf{E}^{-1}$	Feasibility not guaranteed	Least-squares solution is unique
Wide	Underdetermined	No $\mathbf{E}^{-1}$	Feasible solutions exist	Solution is not unique (infinite sols.)

(Still assuming E is not rank-deficient!)

Linear least-squares reconstruction of noisy data

$$\mathbf{d} = \mathbf{E}\mathbf{m} + \mathbf{n}$$

$$\tilde{\mathbf{m}} = (\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{d}$$

$$\tilde{\mathbf{m}} = (\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H(\mathbf{E}\mathbf{m} + \mathbf{n})$$

$$\tilde{\mathbf{m}} = \cancel{(\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{E}}\mathbf{m} + (\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{n}$$

$$\tilde{\mathbf{m}} = \mathbf{m} + (\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{n}$$

Reconstructed image = desired image + reconstruction of noise

$\mathbf{E}$ and $\mathbf{n}$ determine noise characteristics; $\mathbf{m}$ does not

Special case examples

DFT as a matrix operation

Discrete Fourier transform:

$$S[k] = \sum_{n=-N/2}^{N/2-1} I[n] e^{-j2\pi kn/N}$$

$$\xrightarrow{\quad} n$$

$$\begin{bmatrix} S[-N/2] \\ \vdots \\ S[0] \\ \vdots \\ S[N/2-1] \end{bmatrix} = \Big\downarrow k \begin{bmatrix} e^{-j2\pi kn/N} & \cdots & e^{-j2\pi kn/N} & \cdots & e^{-j2\pi kn/N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ e^{-j2\pi kn/N} & \cdots & e^{-j2\pi kn/N} & \cdots & e^{-j2\pi kn/N} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ e^{-j2\pi kn/N} & \cdots & e^{-j2\pi kn/N} & \cdots & e^{-j2\pi kn/N} \end{bmatrix} \begin{bmatrix} I[-N/2] \\ \vdots \\ I[0] \\ \vdots \\ I[N/2-1] \end{bmatrix}$$

DFT as a matrix operation

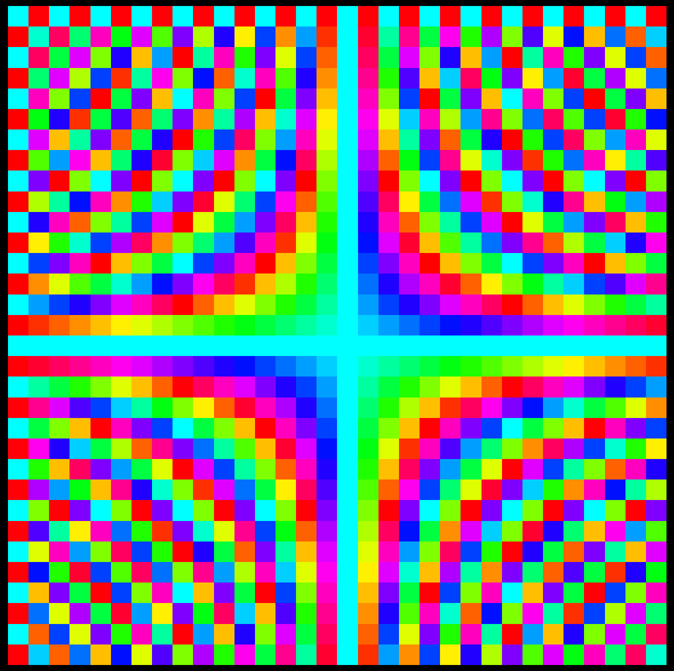
Discrete Fourier transform:

$$S[k] = \sum_{n=-N/2}^{N/2-1} I[n] e^{-j2\pi kn/N}$$

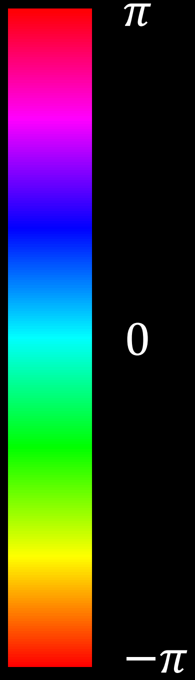
$$\begin{bmatrix} S[-N/2] \\ \vdots \\ S[0] \\ \vdots \\ S[N/2-1] \end{bmatrix} = \Bigg[\quad \Bigg] \begin{bmatrix} I[-N/2] \\ \vdots \\ I[0] \\ \vdots \\ I[N/2-1] \end{bmatrix}$$

$\rightarrow n$

$\downarrow k$



Phase $(-2\pi kn/N)$

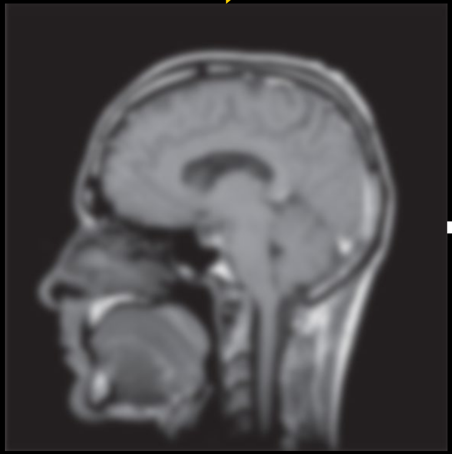


DFT matrix-vector inverse problem

DFT matrix is square and has linearly independent rows/columns, so an inverse exists

Inverse DFT (IDFT) matrix

$\mathbf{F}^{-1}$

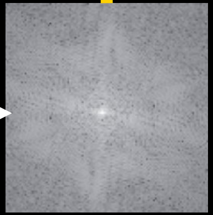


vectorized image

$\mathbf{m}$

DFT matrix

$\mathbf{F}$



data vector

$\mathbf{d}$

$$\mathbf{d} = \mathbf{F}\mathbf{m}$$

$$\mathbf{m} = \mathbf{F}^{-1}\mathbf{d}$$

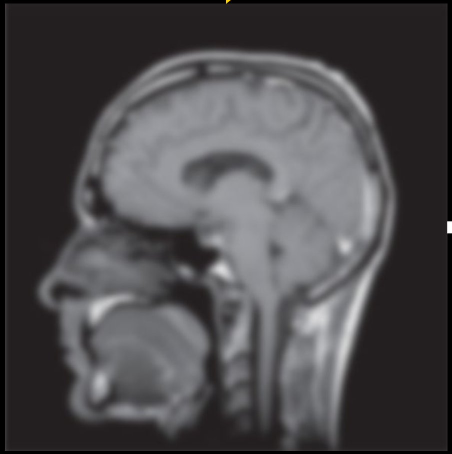
DFT matrix-vector inverse problem

DFT matrix is square and has linearly independent rows/columns, so an inverse exists

FFTs (Fast Fourier Transforms) used in implementation, not matrix multiplication

Inverse FFT

$\mathbf{F}^{-1}$

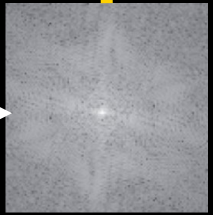


vectorized image

$\mathbf{m}$

FFT

$\mathbf{F}$



data vector

$\mathbf{d}$

$$\mathbf{d} = \mathbf{F}\mathbf{m}$$

$$\mathbf{m} = \mathbf{F}^{-1}\mathbf{d}$$

IDFT reconstruction of two averages

$$\mathbf{d} = \mathbf{F}\mathbf{m} + \mathbf{n}$$

$$\tilde{\mathbf{m}} = \mathbf{F}^{-1}\mathbf{d}$$

$$\tilde{\mathbf{m}} = \mathbf{F}^{-1}(\mathbf{F}\mathbf{m} + \mathbf{n})$$

$$\tilde{\mathbf{m}} = \cancel{\mathbf{F}^{-1}\mathbf{F}}\mathbf{m} + \mathbf{F}^{-1}\mathbf{n}$$

$$\tilde{\mathbf{m}} = \mathbf{m} + \mathbf{F}^{-1}\mathbf{n}$$

Reconstructed image = desired image + IDFT of noise

Effect of IDFT on additive white Gaussian noise (AWGN)

AWGN properties in k-space

$$\mathbf{n}$$

- Gaussian-distributed
- Zero-mean
- Variance σ_k^2 is constant throughout k-space
- Noise at different samples are independent

AWGN properties in image space after IDFT

$$\mathbf{F}^{-1}\mathbf{n}$$

- Gaussian-distributed
- Zero-mean
- Variance σ^2 is constant throughout k-space
- Noise at different voxels are independent

$\mathbf{F}^{-1}$ preserves the basic properties of our noise
Equally valid to consider AWGN in k-space or as AWGN in image space

IDFT reconstruction of multiple averages (rescans)

$$\mathbf{d} = \mathbf{E}\mathbf{m} + \mathbf{n}$$

$$\tilde{\mathbf{m}} = \mathbf{m} + (\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{n}$$

$$\begin{bmatrix}\mathbf{d}_1\\ \mathbf{d}_2\end{bmatrix} = \begin{bmatrix}\mathbf{F}\\ \mathbf{F}\end{bmatrix}\mathbf{m} + \begin{bmatrix}\mathbf{n}_1\\ \mathbf{n}_2\end{bmatrix}$$

$$\tilde{\mathbf{m}} = \mathbf{m} + \mathbf{F}^{-1}\left(\frac{\mathbf{n}_1+\mathbf{n}_2}{2}\right)$$

reduces noise std. dev. by $\sqrt{2}$

$$\begin{bmatrix}\mathbf{d}_1\\ \mathbf{d}_2\\ \vdots\\ \mathbf{d}_T\end{bmatrix} = \begin{bmatrix}\mathbf{F}\\ \mathbf{F}\\ \vdots\\ \mathbf{F}\end{bmatrix}\mathbf{m} + \begin{bmatrix}\mathbf{n}_1\\ \mathbf{n}_2\\ \vdots\\ \mathbf{n}_T\end{bmatrix}$$

$$\tilde{\mathbf{m}} = \mathbf{m} + \mathbf{F}^{-1}\frac{\sum_t \mathbf{n}_t}{T}$$

reduces noise std. dev. by $\sqrt{T}$

Complex coil combination (SENSE, R=1)

$$\mathbf{d} = \mathbf{E}\mathbf{m} + \mathbf{n}$$

$$\tilde{\mathbf{m}} = \mathbf{m} + (\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{n}$$

$$\begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_C \end{bmatrix} = \mathbf{F}\begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \vdots \\ \mathbf{C}_L \end{bmatrix}\mathbf{m} + \mathbf{n}$$

$$\tilde{\mathbf{m}} = \mathbf{m} + (\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H\mathbf{F}^{-1}\mathbf{n}$$

$(\mathbf{C}^H\mathbf{C})^{-1}\mathbf{C}^H\mathbf{x}$ is voxelwise phase correction & scaling

$\frac{1}{\sum_\ell |C_\ell(\vec{r})|^2} \sum_\ell C_\ell^*(\vec{r}) x_\ell(\vec{r})$ → Noise std. dev $\propto \frac{1}{\sqrt{\sum_\ell |C_\ell(\vec{r})|^2}}$

Noise still Gaussian and still independent from voxel to voxel,

but noise amplification $\propto^{-1}$ collective coil sensitivity $\sqrt{\sum_\ell |C_\ell(\vec{r})|^2}$

Parallel imaging (SENSE, R>1)

$$\mathbf{d} = \mathbf{E}\mathbf{m} + \mathbf{n}$$

$$\tilde{\mathbf{m}} = \mathbf{m} + (\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{n}$$

$$\begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_C \end{bmatrix} = \mathbf{\Omega}\mathbf{F} \begin{bmatrix} \mathbf{C}_1 \\ \mathbf{C}_2 \\ \vdots \\ \mathbf{C}_L \end{bmatrix} \mathbf{m} + \mathbf{n}$$

$$\tilde{\mathbf{m}} = \mathbf{m} + (\mathbf{C}^H\mathbf{F}^H\mathbf{\Omega}^H\mathbf{\Omega}\mathbf{F}\mathbf{C})^{-1}\mathbf{C}^H(\mathbf{\Omega}\mathbf{F})^H\mathbf{n}$$

$\mathbf{F}^H\mathbf{\Omega}^H\mathbf{n}$ is aliased noise “image” → noise pattern repeats in space!

Noise still Gaussian, but no longer independent

Noise amplification depends on $\mathbf{\Omega}$ and $\mathbf{C}$ *together* (g-factor)

Takeaways from specific examples

- For linear reconstructions,
  noise properties depend on the reconstruction operator, not on the image

$$\tilde{\mathbf{m}} = \mathbf{m} + (\mathbf{E}^H\mathbf{E})^{-1}\mathbf{E}^H\mathbf{n}$$

- Fourier reconstruction preserves i.i.d. properties of AWGN

$\mathbf{F}^{-1}\mathbf{n}$ is still i.i.d. AWGN

- For other reconstruction operators, image-space noise may not be i.i.d.

Be careful during post-processing!

Constrained image reconstruction

Recall: Underdetermined problem

Some approaches

$$\hat{\mathbf{m}} = \arg\min_{\mathbf{m}} R(\mathbf{m}) \quad \text{s.t.} \quad \mathbf{Em} = \mathbf{d}$$ (force solution to be feasible)

$$\hat{\mathbf{m}} = \arg\min_{\mathbf{m}} \|\mathbf{d} - \mathbf{Em}\|^2 + R(\mathbf{m})$$ (allow deviation from data)

Regularizer $R(\mathbf{m})$ provides a second objective beyond the data term

- Can "break the tie" between infinite feasible solutions
- Can denoise (when feasibility is not the goal)
- Constrains solution to leverage other knowledge about images

Probabilistic interpretation

Least-squares minimization gave *maximum likelihood* (ML) solution:

$$\hat{I}_{\text{ML}} = \arg\max_I p(S|I) = \arg\max_I \log p(S|I) = \arg\min_I \|S - \mathcal{T}\{I\}\|^2$$

What if we want the <u>most probable image given the data</u> (*maximum a posteriori* [MAP] estimate)

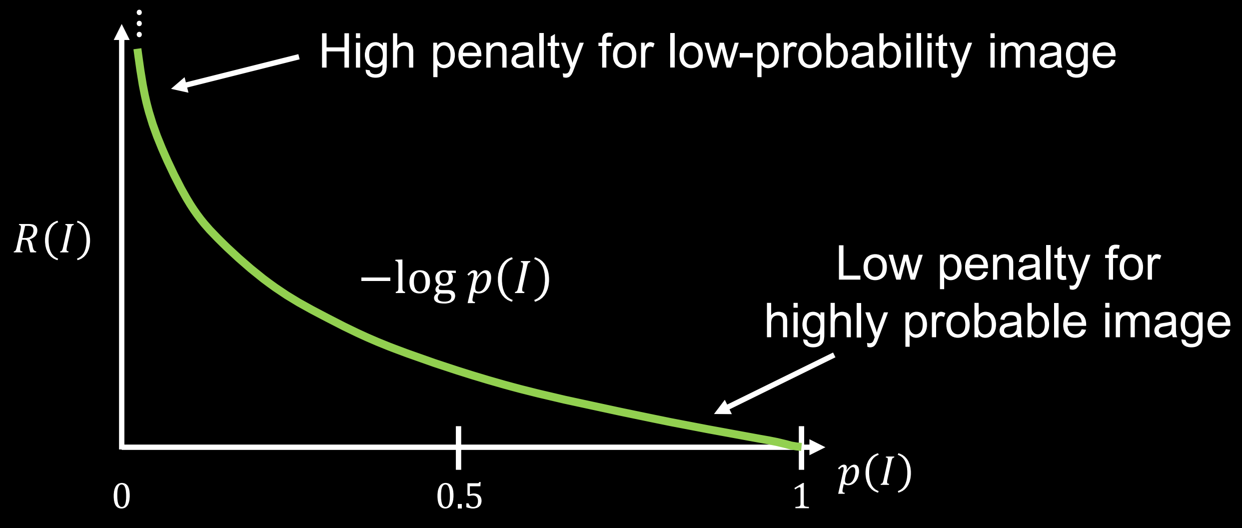
$$\hat{I}_{\text{MAP}} = \arg\max_I p(I|S)$$

$$\hat{I}_{\text{MAP}} = \arg\max_I \frac{p(S|I)\,p(I)}{p(S)}$$

$$\hat{I}_{\text{MAP}} = \arg\max_I p(S|I)p(I)$$

$$\hat{I}_{\text{MAP}} = \arg\max_I \log p(S|I) + \log p(I)$$

$$\hat{I}_{\text{MAP}} = \arg\min_I \frac{1}{2\sigma^2}\|S - \mathcal{T}\{I\}\|^2 - \log p(I)$$



Can define regularization term $R(I)$ to express the prior probability of an image, e.g., $R(I) = -\log p(I)$

If $p(I)$ is constant (uniform distribution; all images equally likely), then MAP solution reduces to ML solution

What makes an image “more probable”?

When it conforms to certain properties:

- Phase properties → partial Fourier imaging
- Sparsity properties → compressed sensing
- Rank properties → low-rank imaging
- Learned properties → artificial intelligence/machine learning
- *Et cetera*

$R(\cdot)$ expresses our <u>prior knowledge</u> about what images can/should look like
Squared-norm data term expresses <u>posterior knowledge</u> (observed data)

Regularized least squares is not the only way to constrain image reconstruction,
but it is still a useful framework for understanding other image reconstruction algorithms too

How to solve a regularized least-squares problem?

Best algorithm for solving $\hat{\mathbf{m}} = \arg\min_{\mathbf{m}} ||\mathbf{d} - \mathbf{E}\mathbf{m}||^2 + R(\mathbf{m})$ depends on both $\mathbf{E}$ and $R(\cdot)$

Many algorithms use variations of alternating minimization

Two forms of knowledge

1. Observations from data (small $||\mathbf{d} - \mathbf{E}\mathbf{m}||^2$)
2. Known image properties (small $R(\mathbf{m})$)

Iterate over two steps enforcing each objective:

1. Enforce data consistency ($\mathbf{E}\mathbf{m} \approx \mathbf{d}$)
2. Impose desired image properties (reduce $R(\mathbf{m})$)

These reconstruction operators are not necessarily linear!

- Image noise may not be i.i.d Gaussian
- Image noise may depend on the image itself

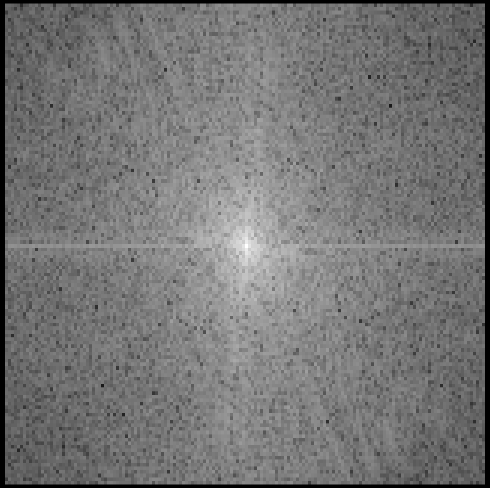
Exception: when $R(\mathbf{m})$ is a squared 2-norm, reconstruction operator is linear and produces Gaussian noise (but not necessarily i.i.d.)

Partial-Fourier imaging: Phase properties

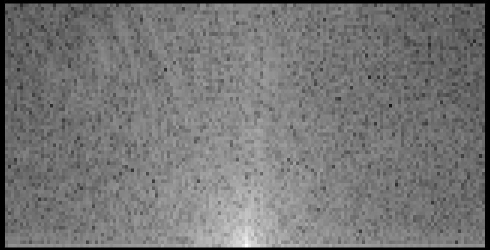
Fourier conjugate symmetry

Real images are conjugate symmetric ($S[\vec{k}] = S^*[-\vec{k}]$),

so only ½ k-space would need sampling



Full **k**-space



Conjugate synthesis
from ½ **k**-space

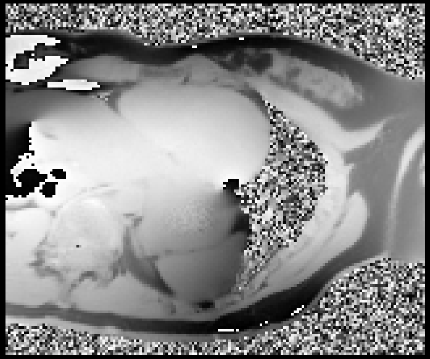
$$S[-\vec{k}] = S^*[\vec{k}]$$

Phase smoothness

MR images are not typically real-valued,

but we can exploit smooth or known phase

Magnitude

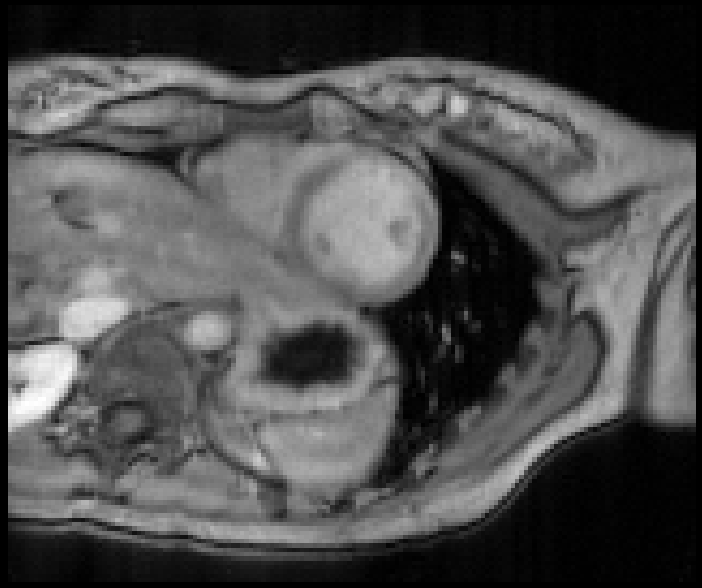


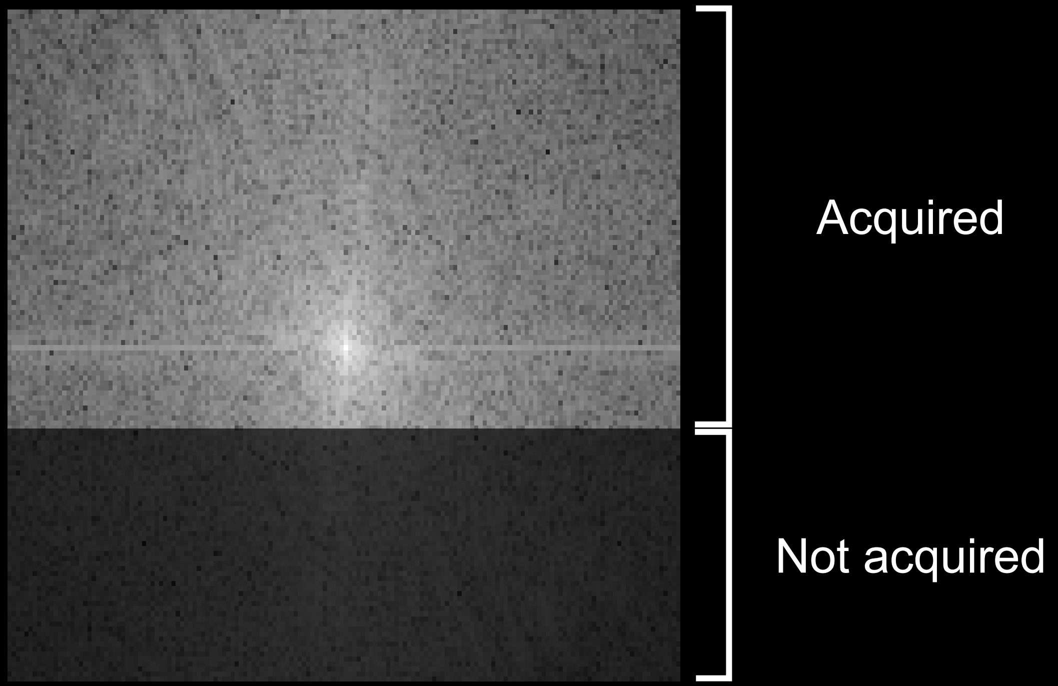
Phase

Partial-Fourier imaging: Sampling

Asymmetric coverage

with enough of central **k**-space to estimate smooth phase



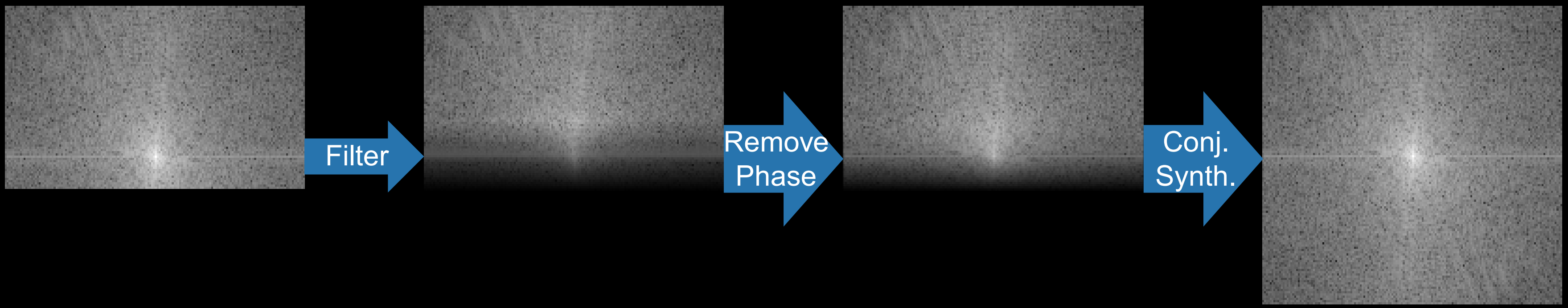


Partial-Fourier imaging: Reconstruction

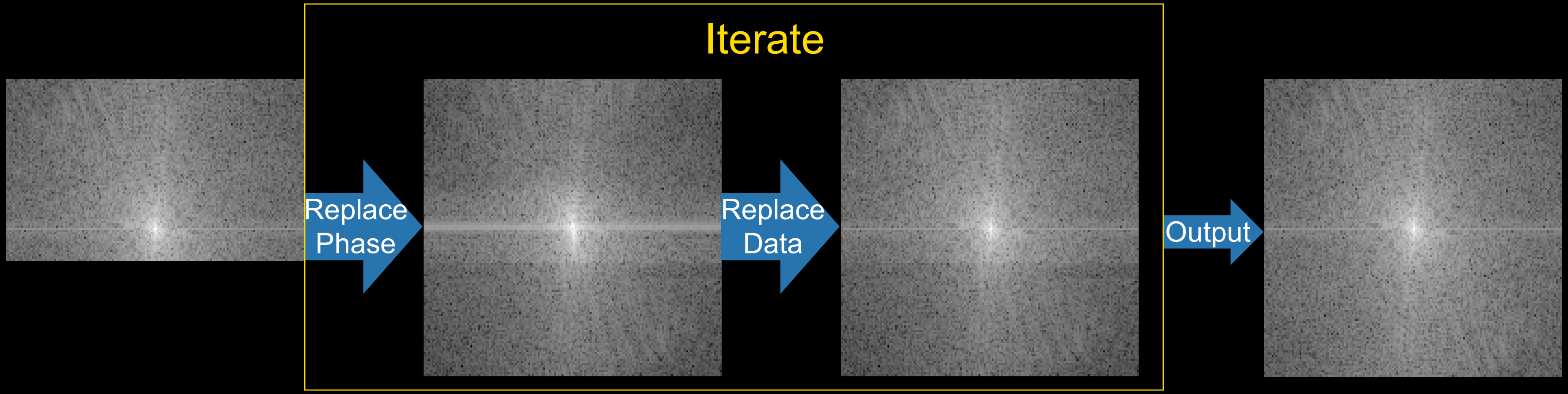
1. Estimate phase from symmetric portion of acquired **k**-space

2. Use phase estimate to synthesize missing data

- Margosian/Homodyne (Margosian et al. *SMRM* 1985; Noll et al. *IEEE-TMI* 1991)
  - Filter and divide image by estimated phase (make it real), then perform conjugate synthesis

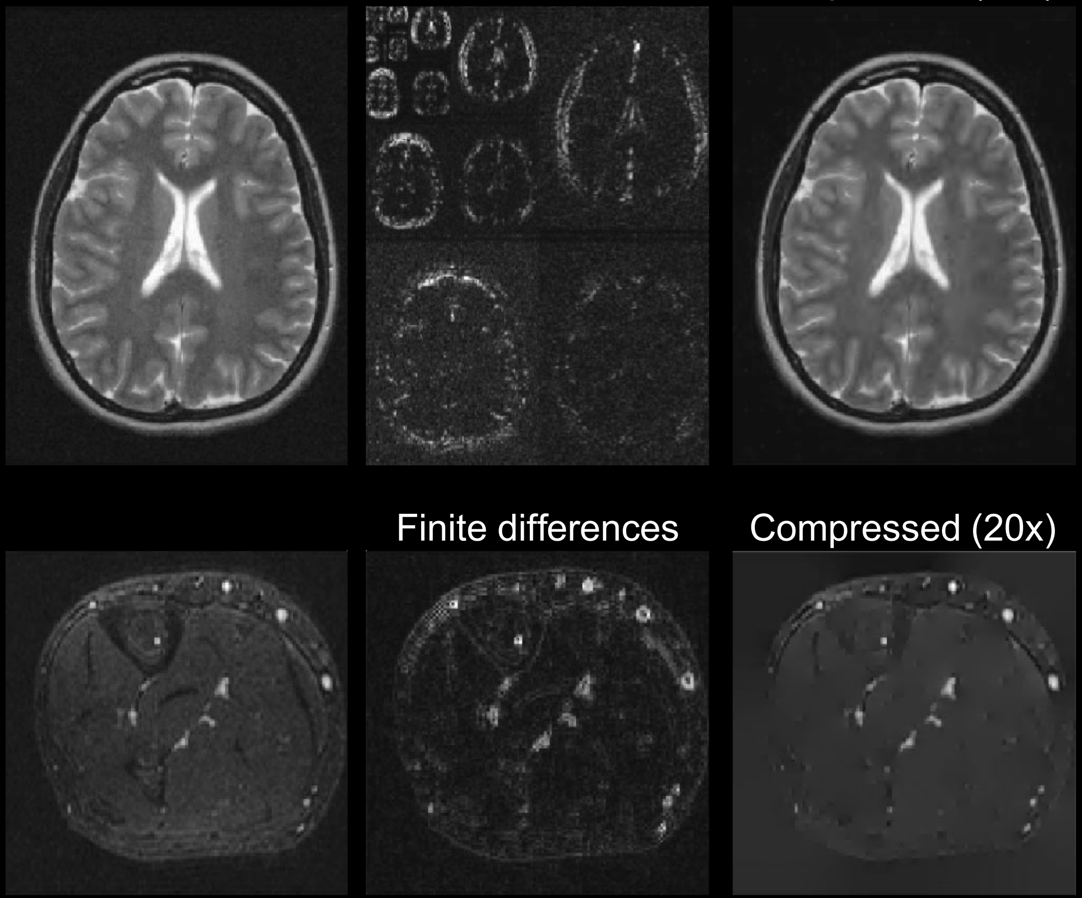


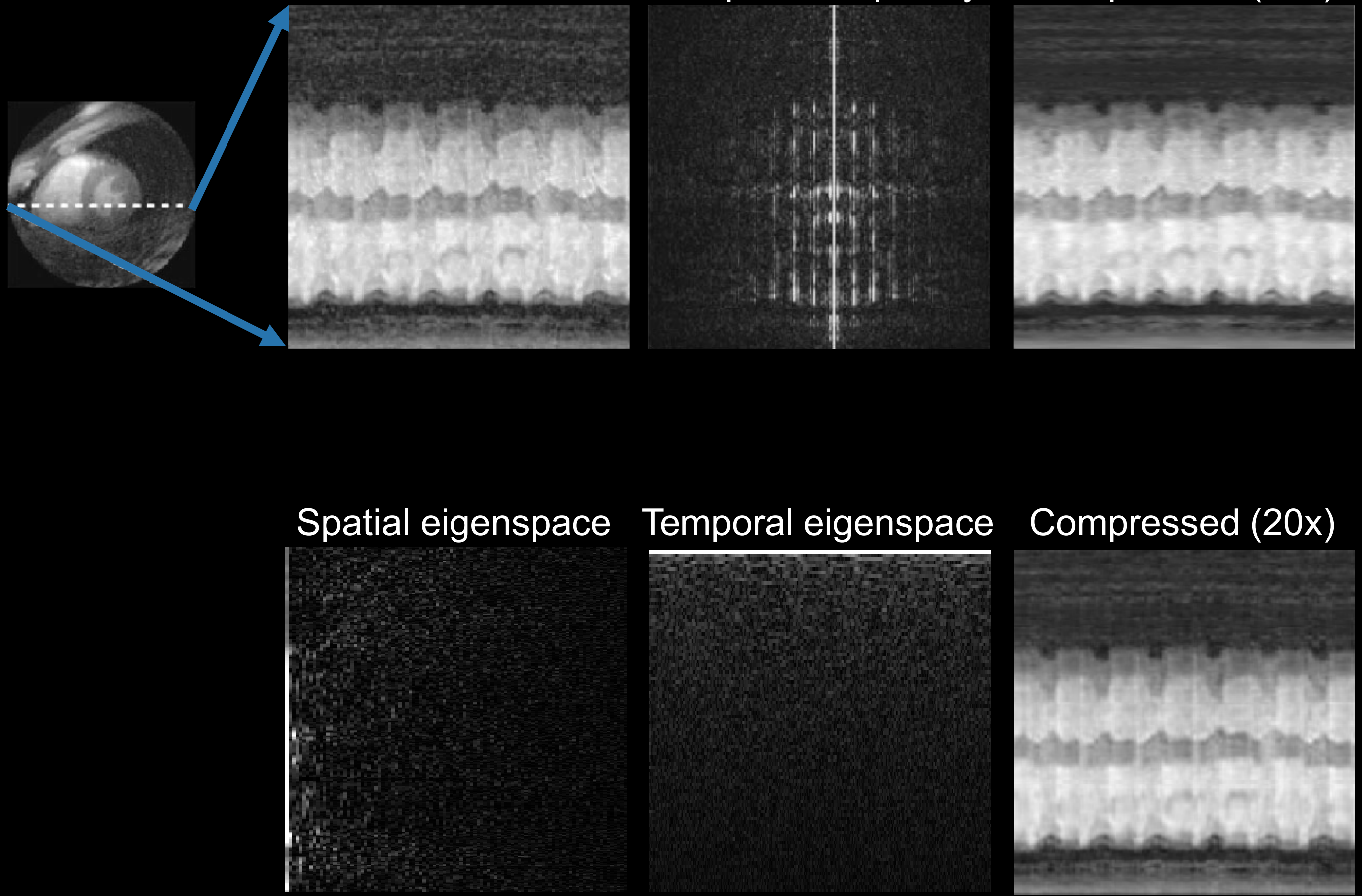
Partial-Fourier imaging: Reconstruction



- POCS: Projection onto convex sets (Lindskog et al. *SMRM* 1989)
  - Iteratively solve for an image which:
    - Matches acquired k-space samples
    - Matches estimated phase in image space
  - A simple form of alternating minimization

Compressed sensing: Transform sparsity





Images are sparse in these domains (many small/zero values)

Figures adapted from
Lustig M et al., *IEEE Signal Process Mag* 2008

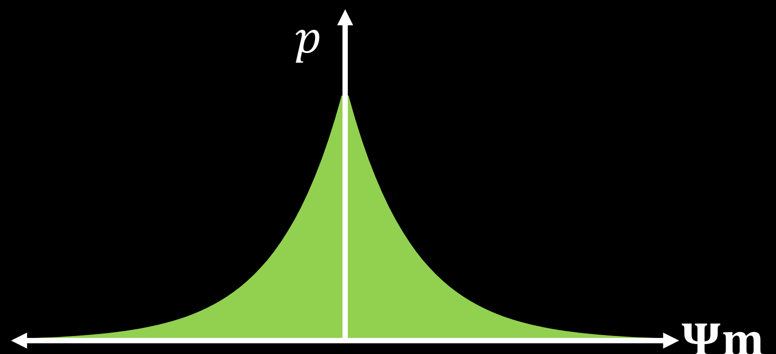
Compressed sensing: L1 regularization

$$R(\mathbf{m}) \propto \|\boldsymbol{\Psi}\mathbf{m}\|_1 = \sum_n |[\boldsymbol{\Psi}\mathbf{m}]_\mathrm{n}|$$

Probabilistic interpretation

$$p(\mathbf{m}) = \prod_n \frac{1}{2b} e^{-\frac{|(\boldsymbol{\Psi}\mathbf{m})_\mathrm{n}|}{b}}$$

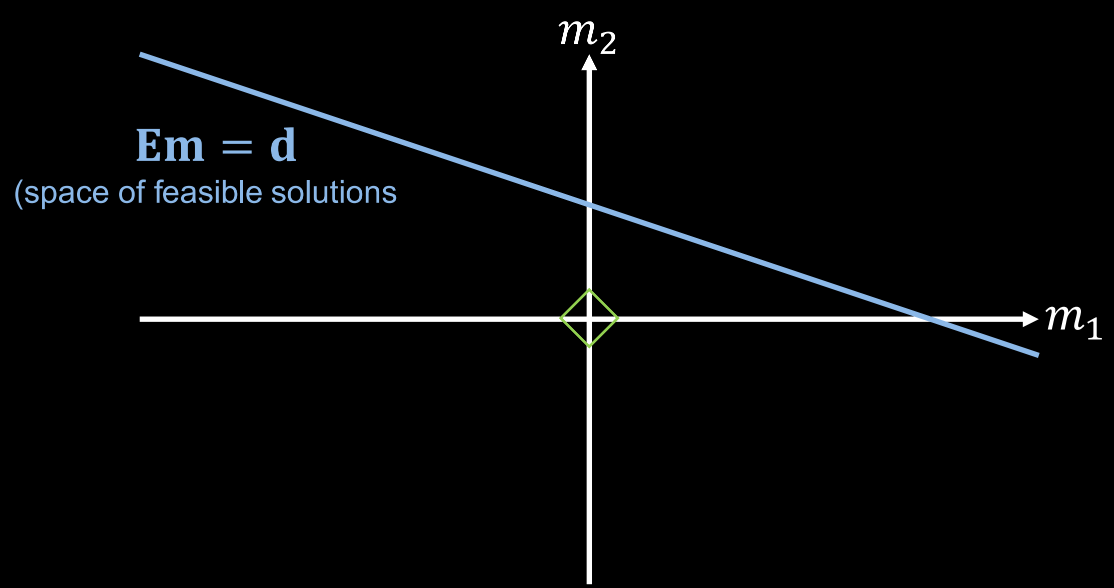
(Laplace distribution)



Heavy concentration near 0

Geometric interpretation

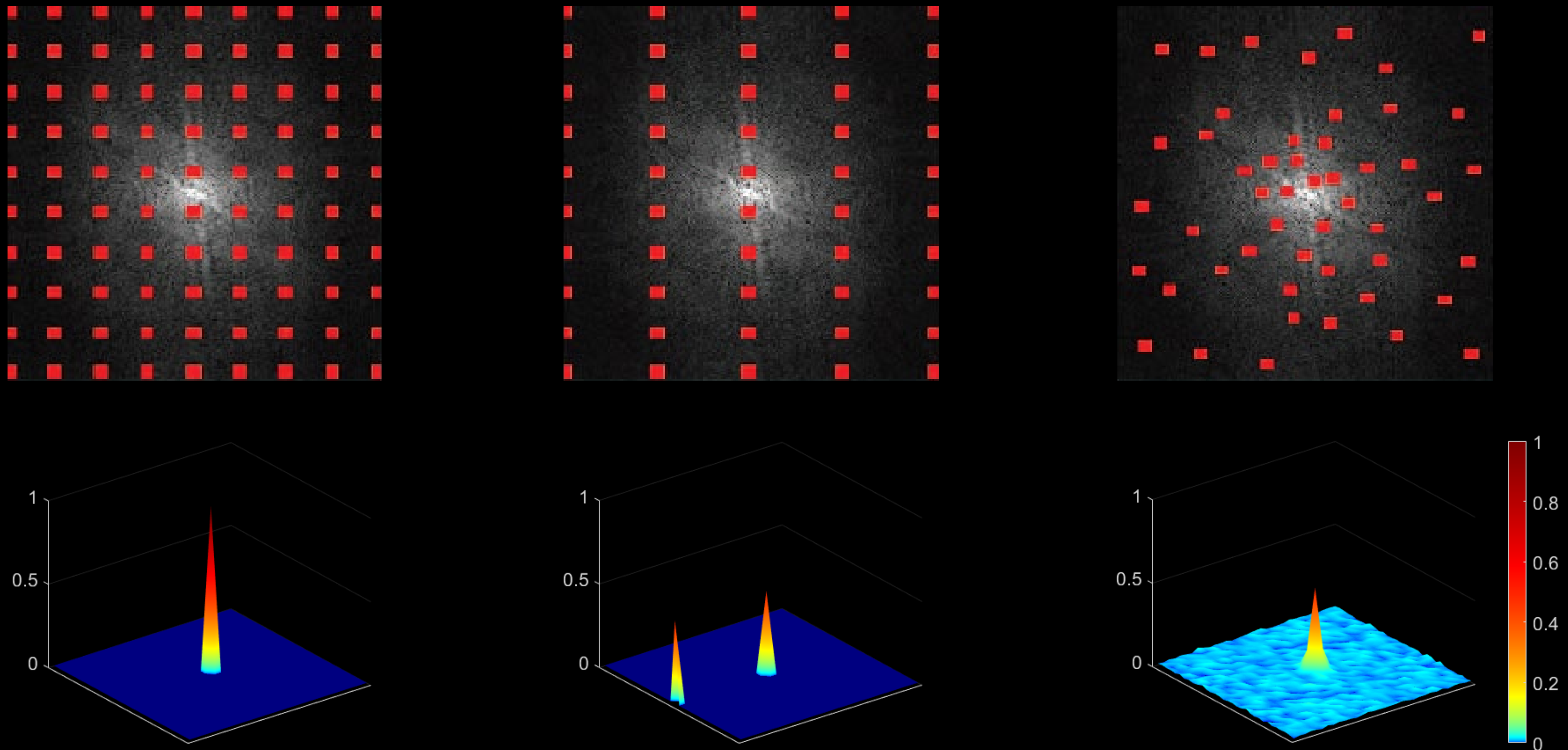
Find smallest $\|\boldsymbol{\Psi}\mathbf{m}\|_1$ that produces feasible solution



$\|\boldsymbol{\Psi}\mathbf{m}\|_1 = \mathit{constant}$ has points along the axes and can often find the sparse solution

Compressed sensing: Sampling

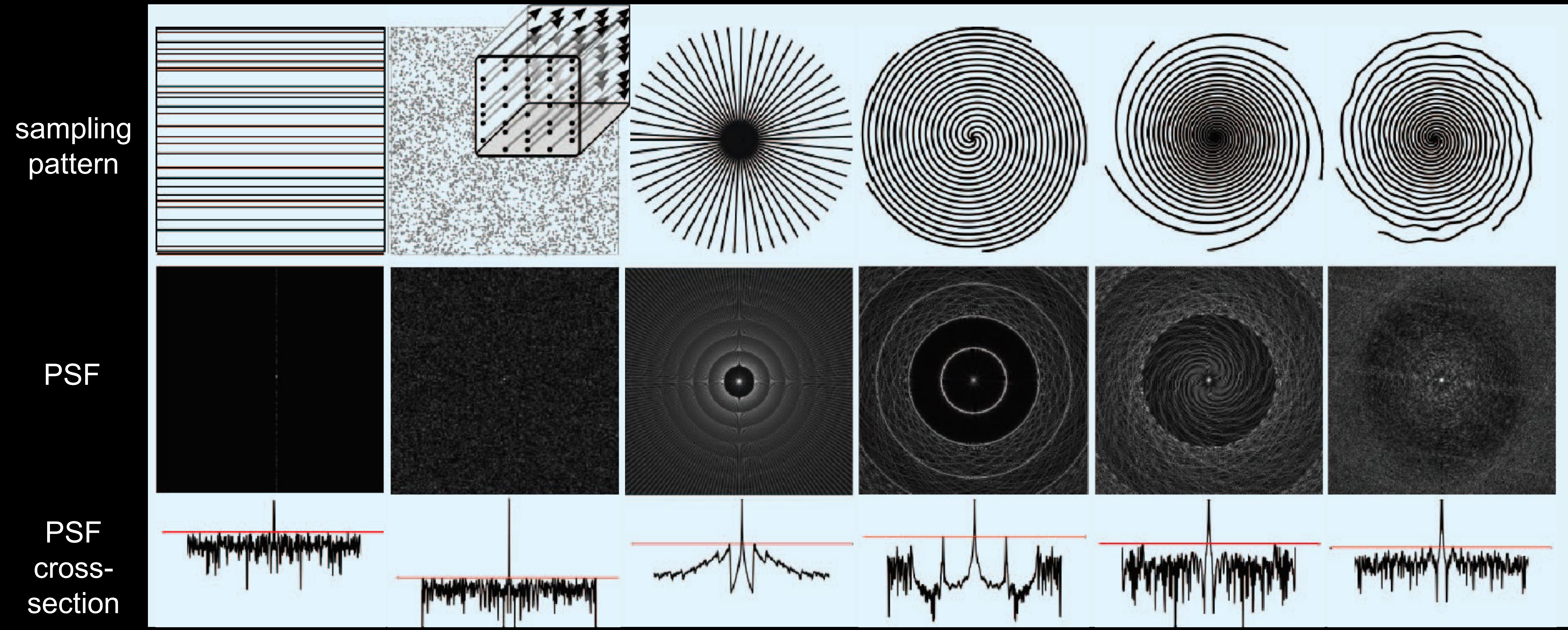
Nyquist sampling | Uniform undersampling | Incoherent undersampling



point spread function (PSF) | point spread function (PSF)

Figures adapted from
Lustig M et al., *IEEE Signal Process Mag* 2008

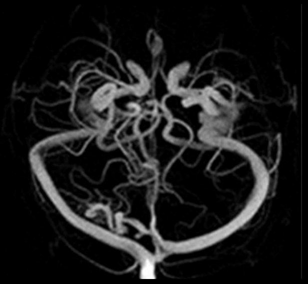
Compressed sensing: Sampling



Figures adapted from
Lustig M et al., *IEEE Signal Process Mag* 2008

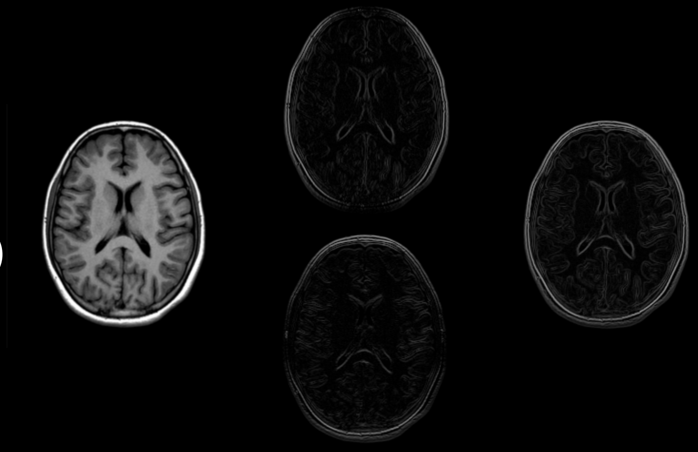
Compressed sensing: Example transforms

- No transformation
  - Suitable when image itself is sparse
    - e.g., angiograms (no background contrast)



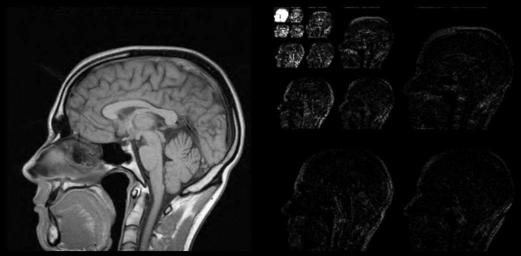
$\mathbf{m}$ is sparse

- Finite difference transformation (total variation)
  - Suitable when edge map is sparse
    - e.g., brain images (discrete tissue compartments)



$\nabla\mathbf{m}$ is sparse

- Wavelet transformation (~multiscale edge information)
  - Suitable for wide range of medical and natural images
    - e.g., MR images in general



$\boldsymbol{\Psi}\mathbf{m}$ is sparse

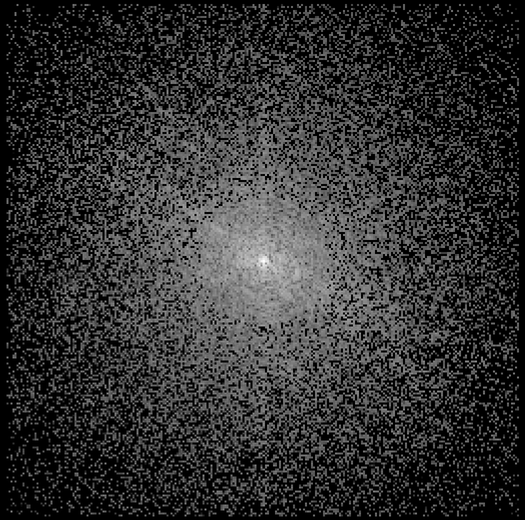
Chang W et al., *AJNR* 2015
Lustig M et al., *Magn Reson Med* 2007

Compressed sensing: alternating minimization

Two goals:

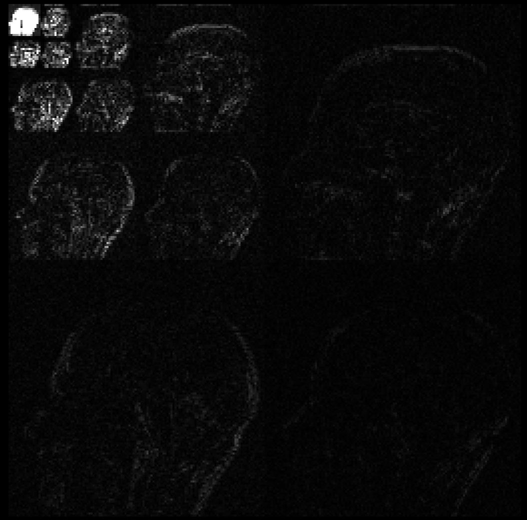
1. Impose sparsity
2. Maintain data consistency

Find the image with sparsest representation that also fits the data



k-space domain

image domain



wavelet domain

Compressed sensing: alternating minimization

Two goals:

1. Impose sparsity
2. Maintain data consistency

Find the image with sparsest representation that also fits the data

Impose sparsity
(e.g., threshold)

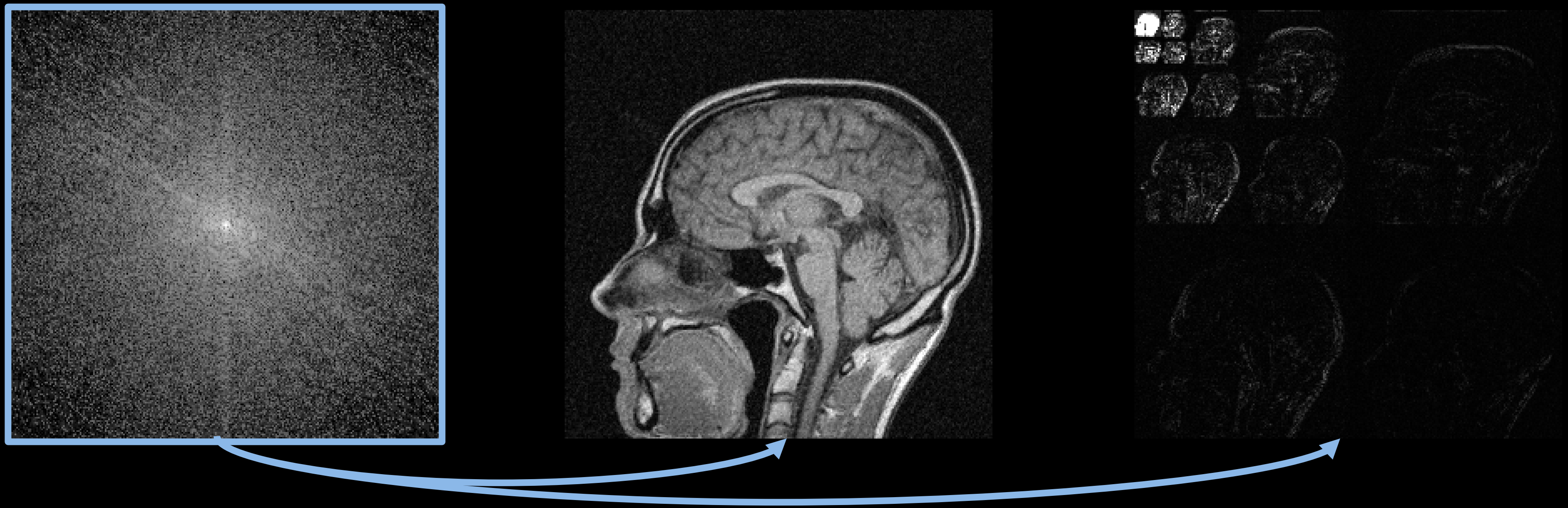
Compressed sensing: alternating minimization

Two goals:

1. Impose sparsity
2. Maintain data consistency

Find the image with sparsest representation that also fits the data

Maintain data consistency (e.g., replace)



Compressed sensing: alternating minimization

Two goals:

1. Impose sparsity
2. Maintain data consistency

Find the image with sparsest representation that also fits the data

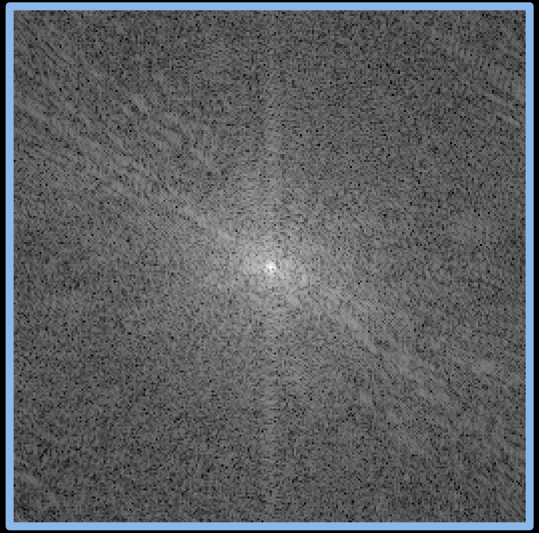
Impose sparsity
(e.g., threshold)

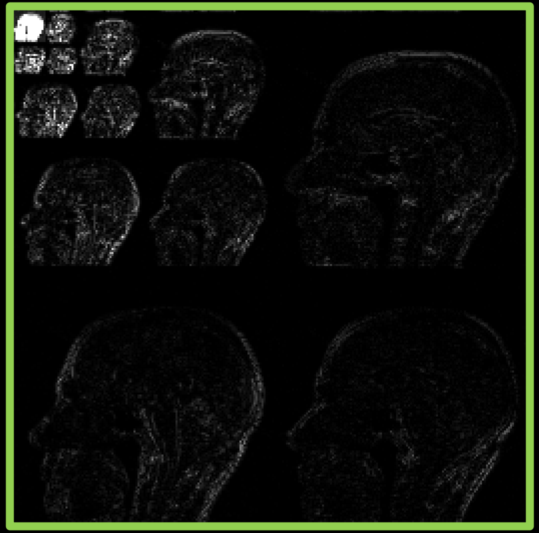
Compressed sensing: alternating minimization

Two goals:

1. Impose sparsity
2. Maintain data consistency

Find the image with sparsest representation that also fits the data





After several iterations, a balance between both goals is achieved

Questions?

Please fill out the evaluation form!

(see QR code)