Advanced Image Reconstruction

Incomplete Measurements and Compressed Sensing

Mark Chiew (mark.chiew@ndcn.ox.ac.uk)
Wellcome Centre for Integrative Neuroimaging (WIN)
31 Jan 2019

Outline

- Review: MR Imaging as a Linear System
- Solving under-determined problems
- Basics of Compressed Sensing
- Some algorithmic suggestions for the practical

MR Imaging as a Linear System

MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx$$

Physical Perspective

MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx$$

What the scanner "measures"

MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx$$

Magnetisation magnitude at every point in space

MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx$$

Magnetisation phase at every point in space

Dictated by Bloch equation and linear magnetic fields

Unfortunately we can't make these like "delta" functions

MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx$$

Coil integrates vector magnetisation over all space

MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx$$

Fourier Transform Perspective

MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx$$

Our signal in the "standard" pixel basis:

MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx$$

New basis of complex exponentials

MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx$$

Inner product to compute Fourier coefficients

i.e. change of basis from standard to Fourier

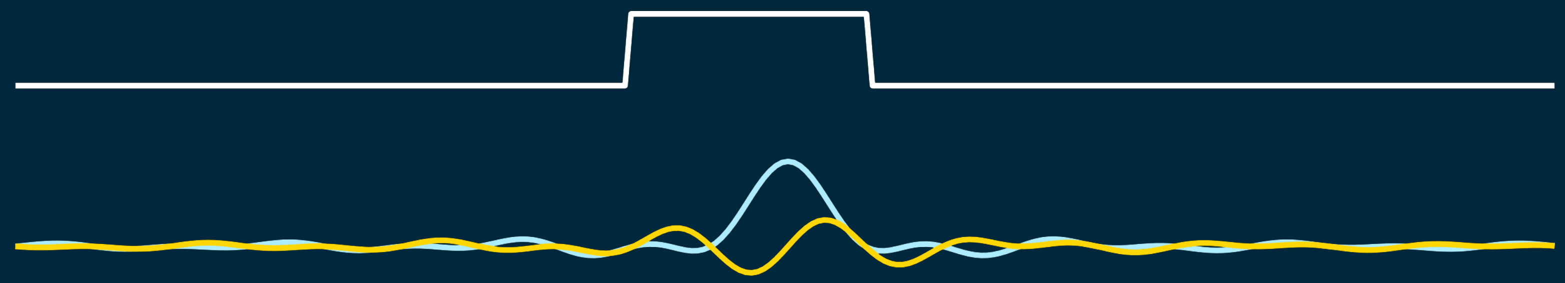
MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx \qquad [10]$$

Same signal in new (Fourier) basis

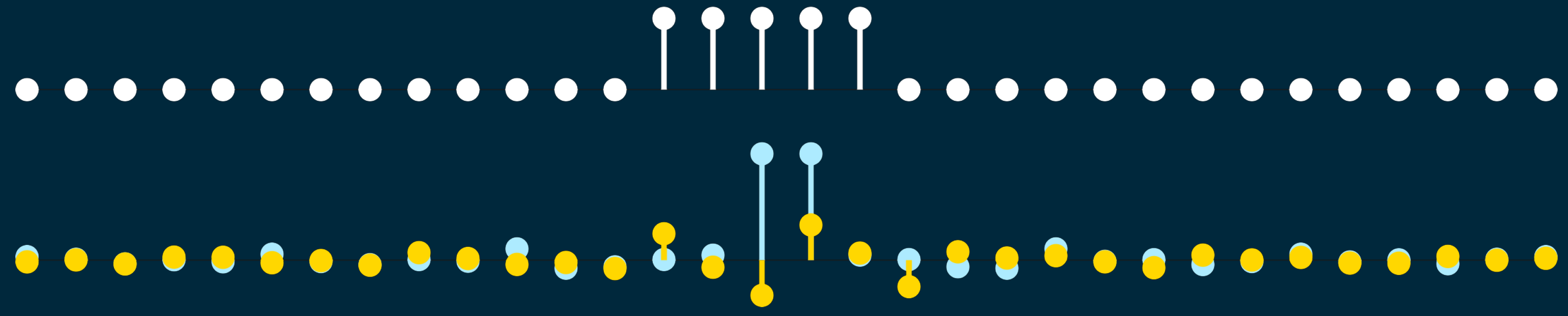
MR Signal Equation

$$\hat{s}(k) = \int \rho(x) e^{-i(2\pi k)x} dx \qquad [10]$$



Discretised Signal Equation

$$\hat{s}(k) = \sum_{n=0}^{N-1} \rho(x_n) e^{-i2\pi kn/N} \qquad [11]$$



Discretised Signal Equation

$$\hat{s}(k) = \sum_{n=0}^{N-1} \rho(x_n) e^{-i2\pi kn/N} \qquad [11]$$

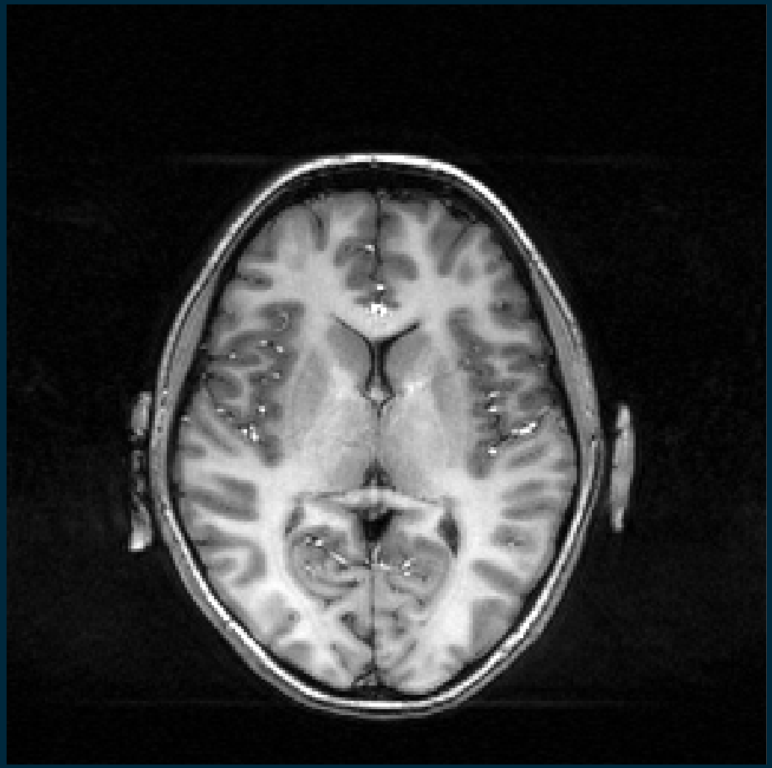
$$\begin{bmatrix} \hat{s}(k_0) \\ \hat{s}(k_1) \\ \vdots \\ \hat{s}(k_{N-1}) \end{bmatrix} = \begin{bmatrix} e^{-i2\pi k_0(0)/N} & e^{-i2\pi k_0(1)/N} & \dots & e^{-i2\pi k_0(N-1)/N} \\ e^{-i2\pi k_1(0)/N} & e^{-i2\pi k_1(1)/N} & \dots & e^{-i2\pi k_1(N-1)/N} \\ \vdots & \vdots & \vdots & \vdots \\ e^{-i2\pi k_{N-1}(0)/N} & e^{-i2\pi k_{N-1}(1)/N} & \dots & e^{-i2\pi k_{N-1}(N-1)/N} \end{bmatrix} \begin{bmatrix} \rho(x_0) \\ \rho(x_1) \\ \vdots \\ \rho(x_{N-1}) \end{bmatrix}$$

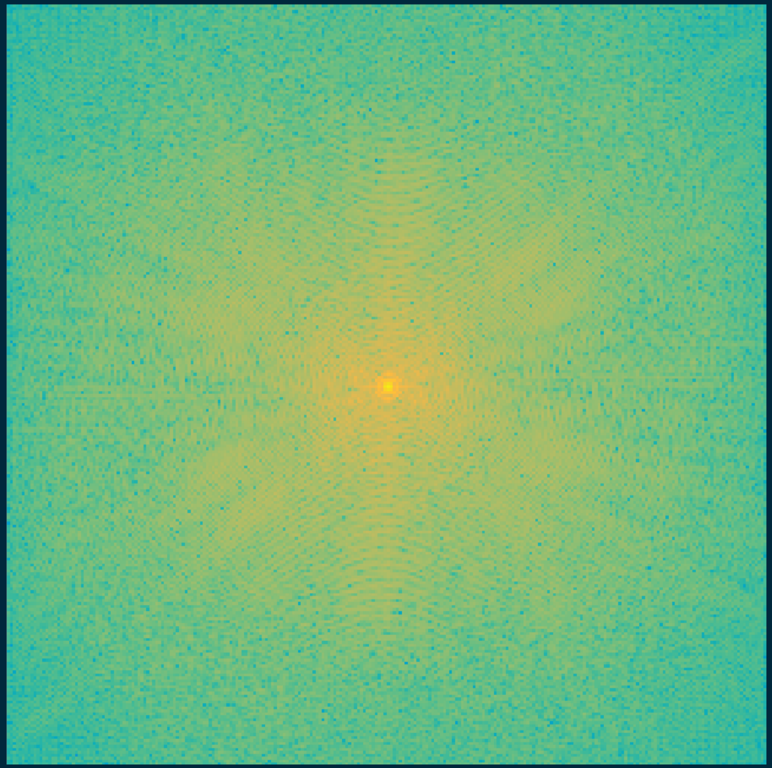
Matrix Signal Equation

$$\hat{s} = F\rho \qquad [12]$$

Continuous vs. Discrete

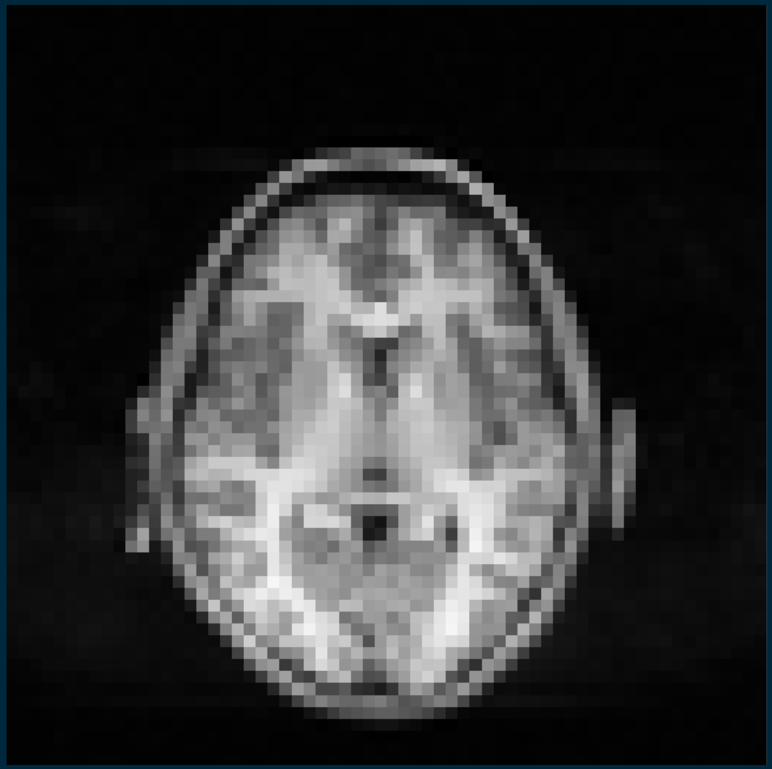
Continuous source space

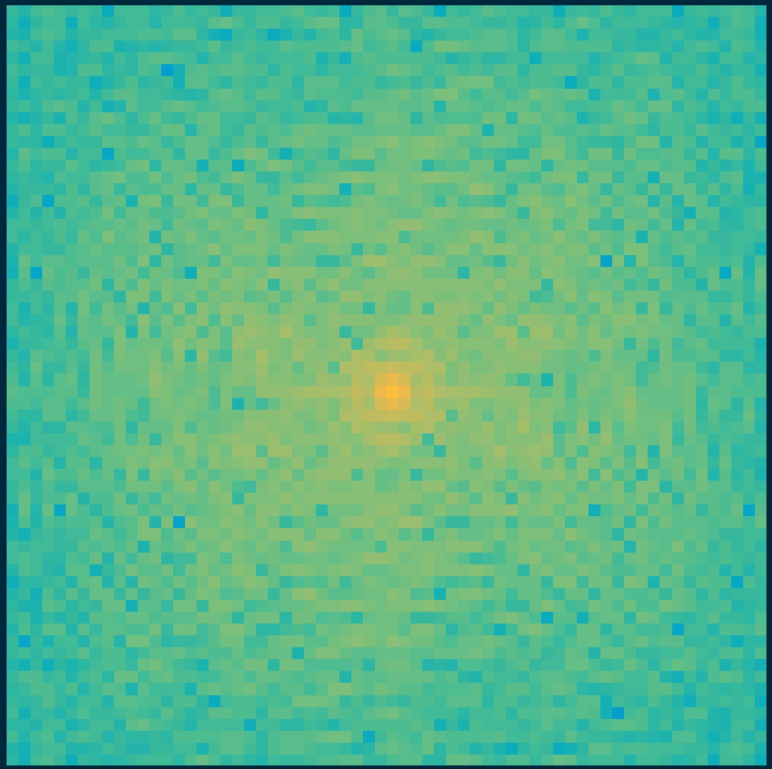




Continuous k-space

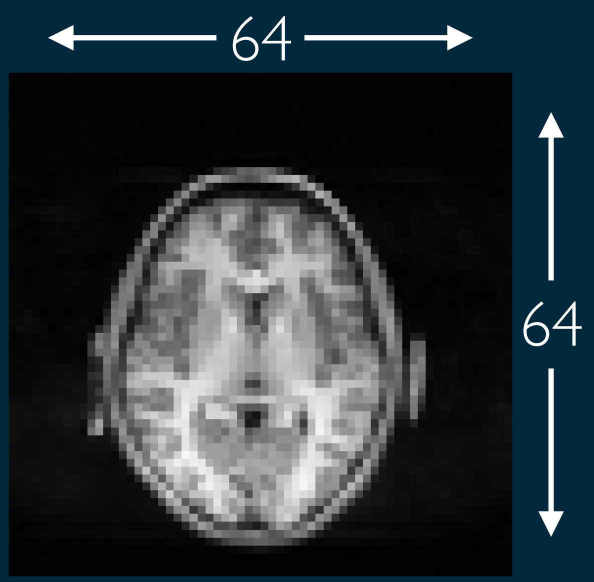
Discrete image space





Discrete, sampled k-space

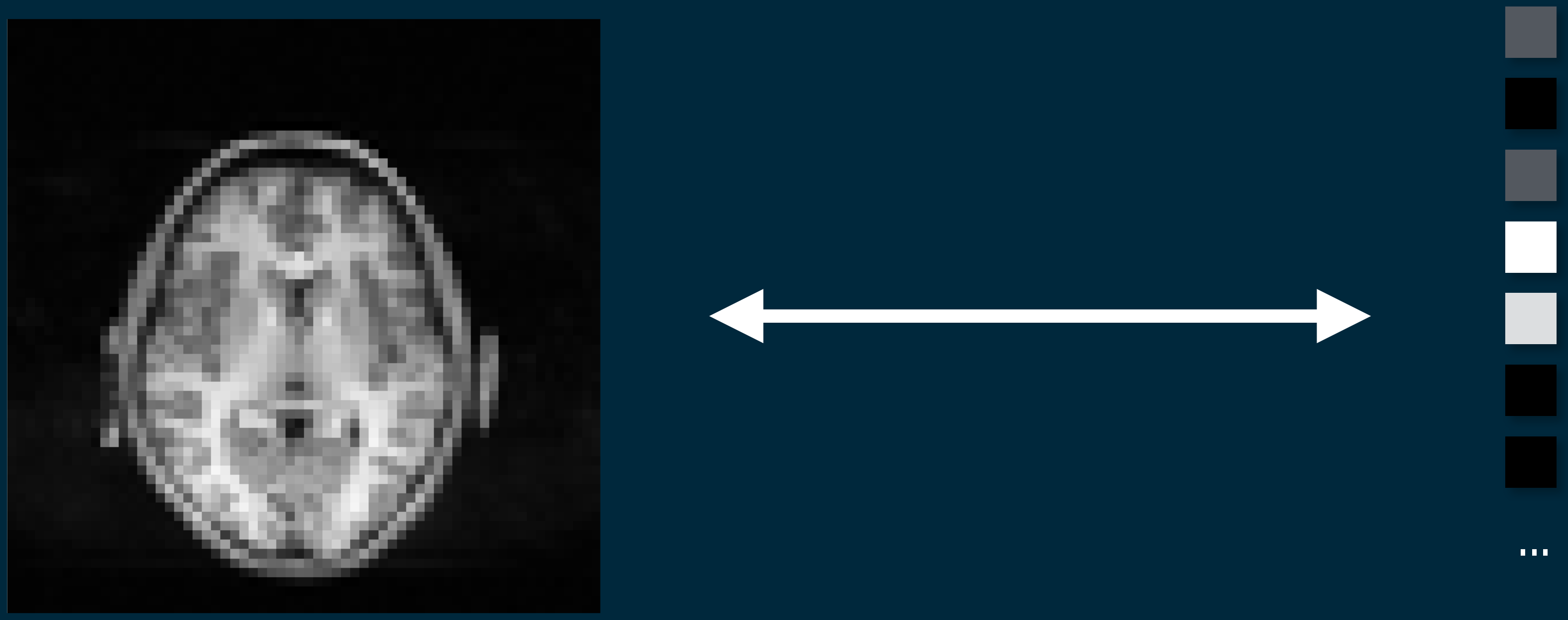
The Image as an Array of Voxels



- An ***m***x***n*** voxel image contains ***mn*** voxels
- With no knowledge or assumptions, each voxel's value is independent of the others
- Therefore the image *data* itself lies in $\mathbb{C}^{mn}$ (e.g. 4096 dimensional complex vector space)

The Image as a Vector

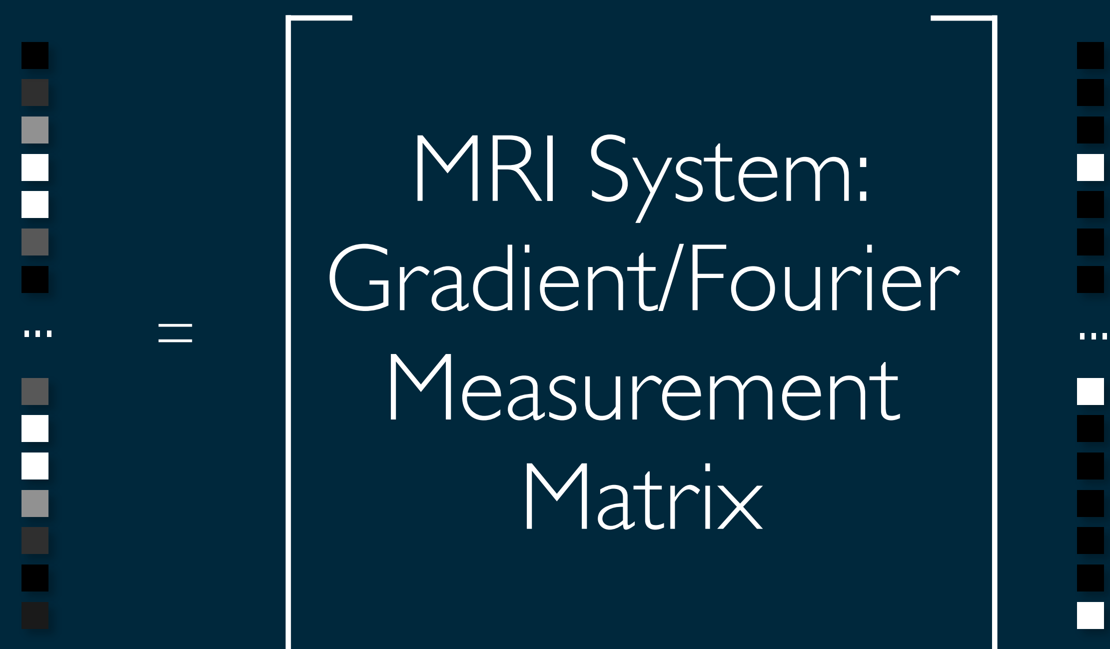
- Image data are vectors in an ***mn***-dimensional vector space
- Consider the data in conventional [***mn*** × *1*] vector form
  - We know exactly how to map the data vector to an "image"



- To find the discrete image, we need to **solve for this vector**

Image Encoding (Measurement)

1. MRI measurements are encoded via magnetic field gradients
2. Gradients impose sinusoidal phase variation prior to integration
3. This process is linear, and physically encodes a Fourier transform



$$k = E \cdot x$$

Our measurement can be described as a linear system

Imaging as a Linear System

- Imaging becomes a simple matter of solving:

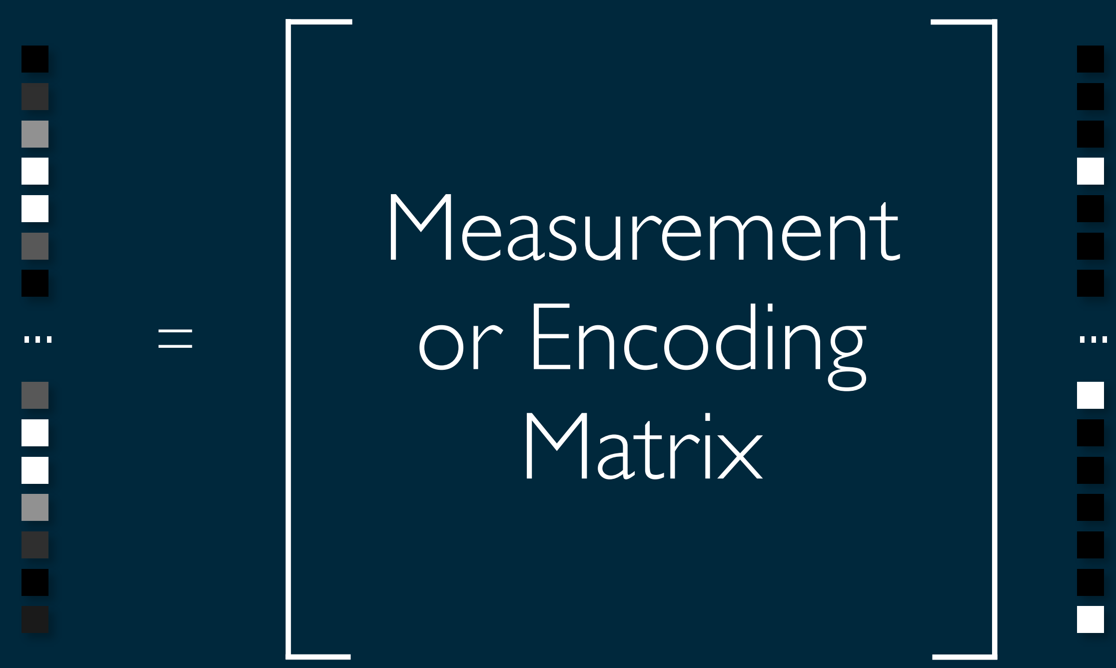
$$k = Ex$$

  - where ***k*** is some vector of measured k-space values, ***E*** is a Fourier *encoding matrix* (or *measurement transform*) modelling the action of the MRI hardware, and ***x*** is the unknown image

- Abstraction of the imaging problem offers flexibility and utility afforded by existing mathematics

The Inverse Problem

- Forward problem: “Given ***E*** and ***x***, generate the samples ***k***”
- Inverse problem: “Given ***E*** and ***k***, what should the image ***x*** be?”



$$k = E \cdot x$$

Solving for the Image

- "Apply the inverse Fourier transform to the k-space data"

$$k = Fx \longrightarrow \hat{x} = F^{-1}k$$

- We know that the *inverse* Fourier transform means the *inverse* or *conjugate transpose* of the (unitary) Fourier transform matrix
- A square matrix is necessary (but not sufficient) for invertibility
  - Square means equal numbers of measurements and unknowns
  - Need as many k-space samples as there are image points

Solving under-determined problems

Under-sampling

- Measurements we make in k-space are sequential, and take time
- To speed up data acquisition, we can reduce the # of measurements
  - With fewer measurements, total measurement time is reduced,, but our linear system is now underdetermined (#unknowns > #measurements)

- How do we solve for our unknown image given a reduced number of measurements?

Ill-Posed Problems

$$7 = 3x - y$$

Ill-Posed Problems

$$7 = 3x - y$$

You can ***never*** find a unique solution to an under-determined problem

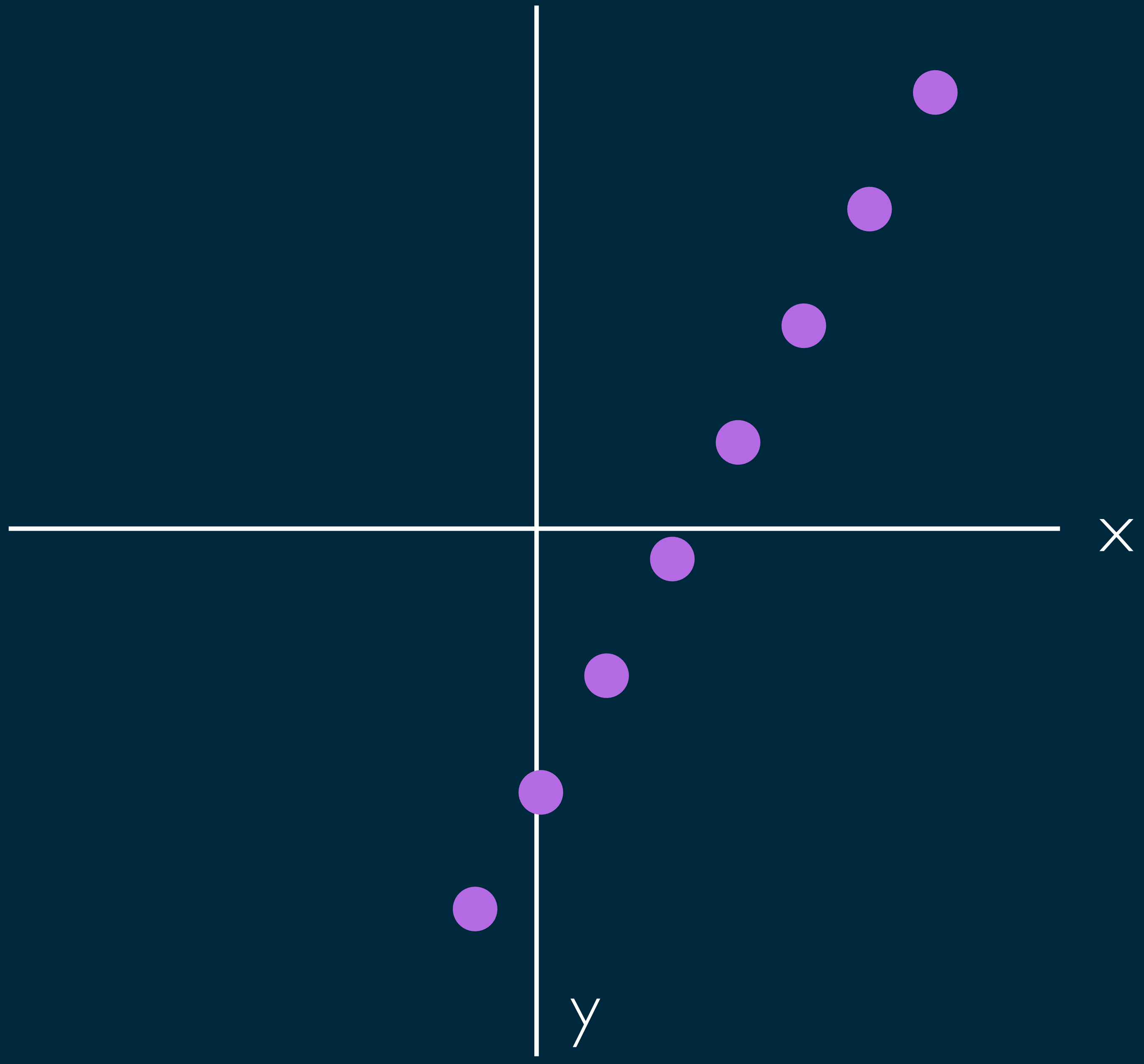
Ill-Posed Problems

$$7 = 3x - y$$

- Some additional information:
  - y is a positive integer (insufficient)
  - y is even (still insufficient)
  - y ≤ 5 (now sufficient)

Ill-Posed Problems

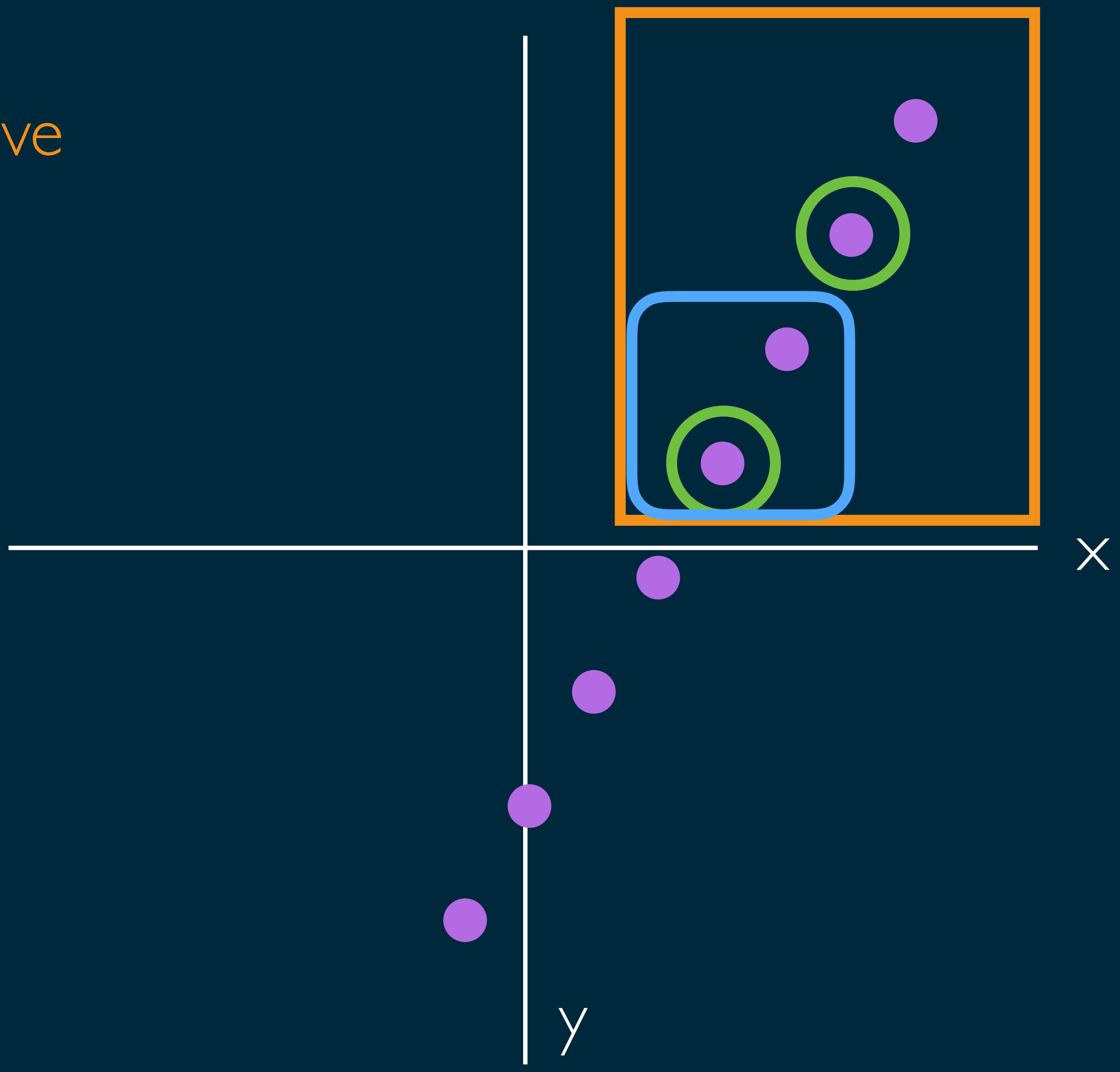
$$7 = 3x - y$$



Ill-Posed Problems

$$7 = 3x - y$$

- y is positive
- y is even
- $y \leq 5$



Regularisation

$$7 = 3x - y$$

- Regularisation is the act of providing extra information to help solve an ill-posed or under-determined problem
- Enforcing particular *constraints* on a solution is a means of regularisation

Regularisation

- Choosing the right regulariser depends on your *a priori* information

- e.g., From the set of infinite possible solutions, give me the one that
  - has minimum energy
  - has minimum entropy
  - has only positive coefficients
  - has the minimum L_1 norm
  - is the sparsest

Compressed Sensing

- **L_0 "norm"** (# non-zero values)
  - → $||x||_0 = \boldsymbol{\Sigma}\ (x_i \mathrel{!=} 0)$
- **L_1 norm** (sum of absolute values of coefficients)
  - → $||x||_1 = \boldsymbol{\Sigma}\ |x_i|$
- **L_2 norm** (root sum of squared coefficients)
  - → $||x||_2 = (x^T x)^{0.5} = (\ \boldsymbol{\Sigma}\ x_i^2\ )^{0.5}$
  - → Euclidean distance norm, easy to work with

A signal is **sparse** if it has few non-zero coefficients

Sparse

Non-sparse

Introduction to CS

- Compressed Sensing (CS) aims to recover images from heavily under-sampled measurements

- With fewer measurements than unknowns, we need to *know* or *assume* something about our data to find our image
  - Minimizing the L2 norm chooses the signal/image with the lowest energy
  - This is great if you *know* your signal should have low energy

- What other *features* or *structure* in our images can we use to our advantage?

MRI Data is Sparse

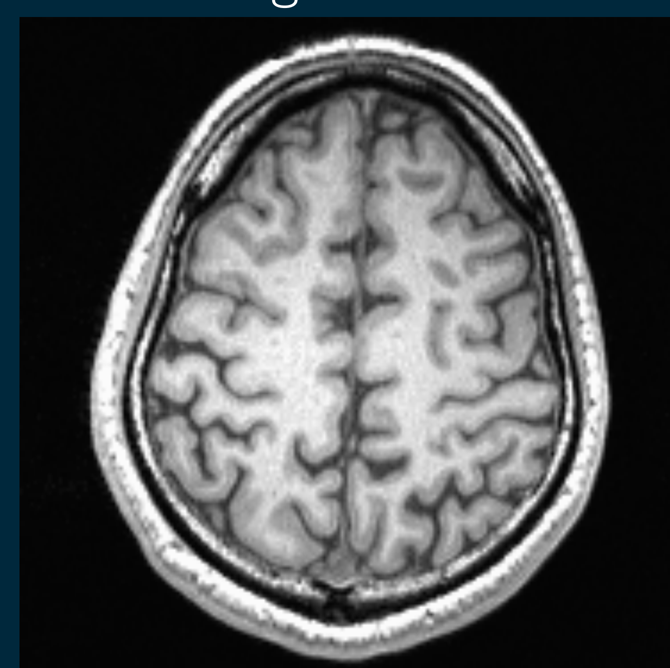
- Sparsity is a type of structure we can exploit

Image courtesy of T. Okell

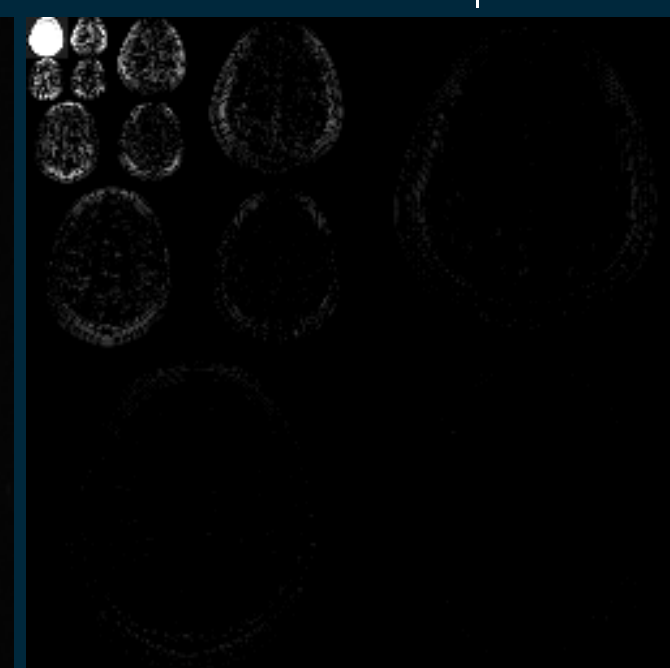
- MR Angiograms are sparse in native image space

- Structural brain images can be sparse using a wavelet representation

Image Domain



Wavelet Decomposition



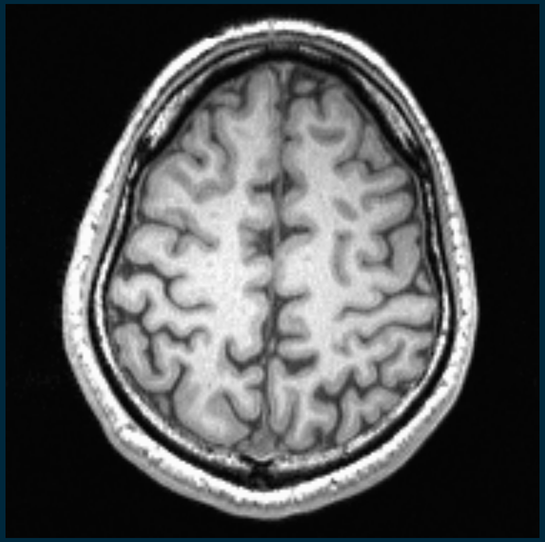
Compressed Sensing

- We can now define Compressed Sensing (CS):
  - Given data we know is sparse (in some representation)
  - Solve the under-determined system by finding the sparsest solution that is consistent with the data

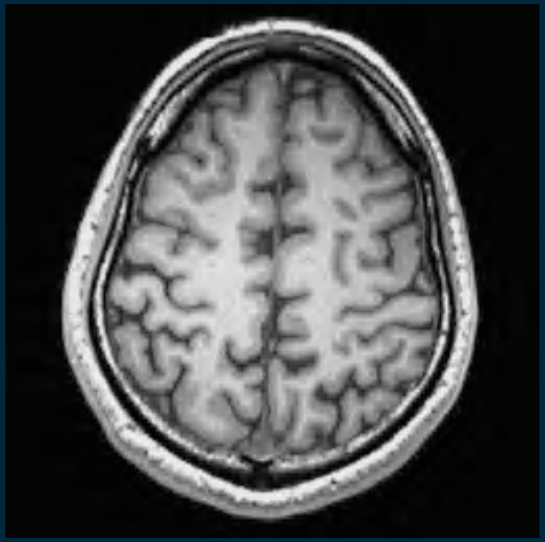
- Assuming some conditions and details are met, we are highly likely to reconstruct our full image vector, with far fewer measurements than unknowns

Intuition

- Sparsity is related to the idea of compressibility
  - Sparse data are highly compressible



100%
{16384 wavelet coefficients}



10%
{1638 wavelet coefficients}

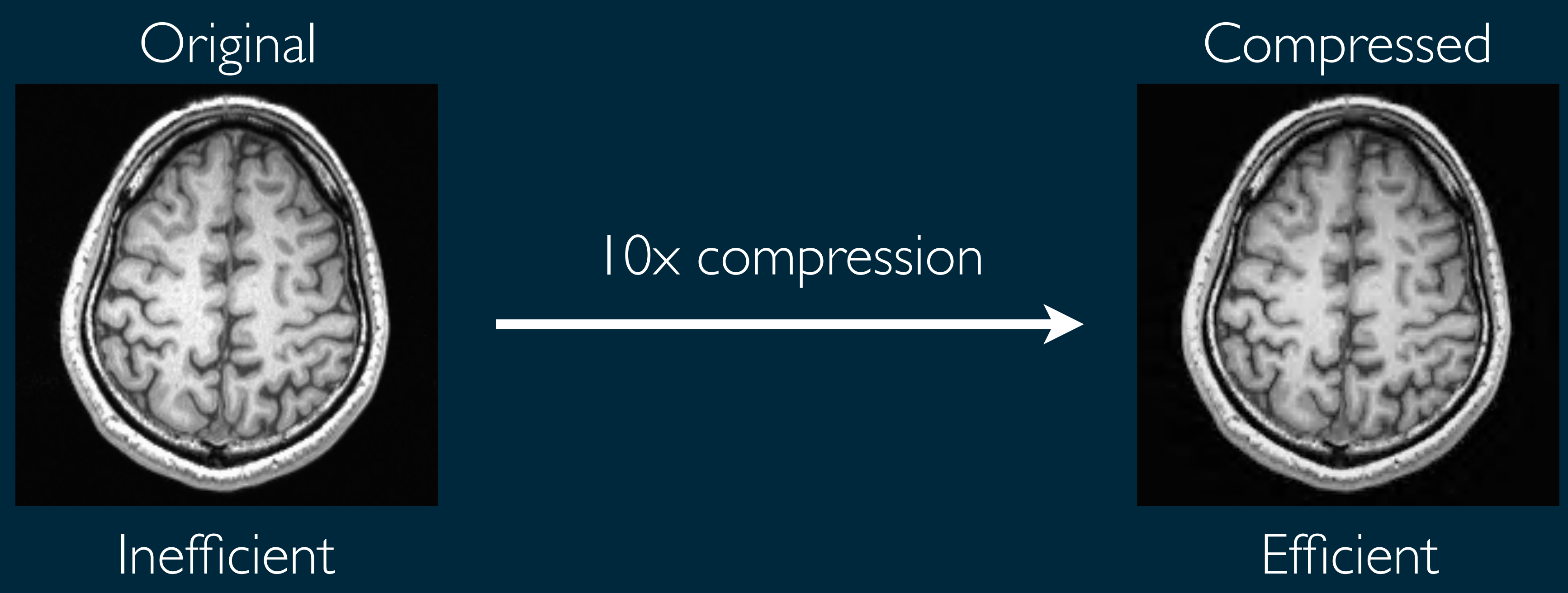
5%
{819 wavelet coefficients}

1%
{164 wavelet coefficients}

- Compressible signals contain less information

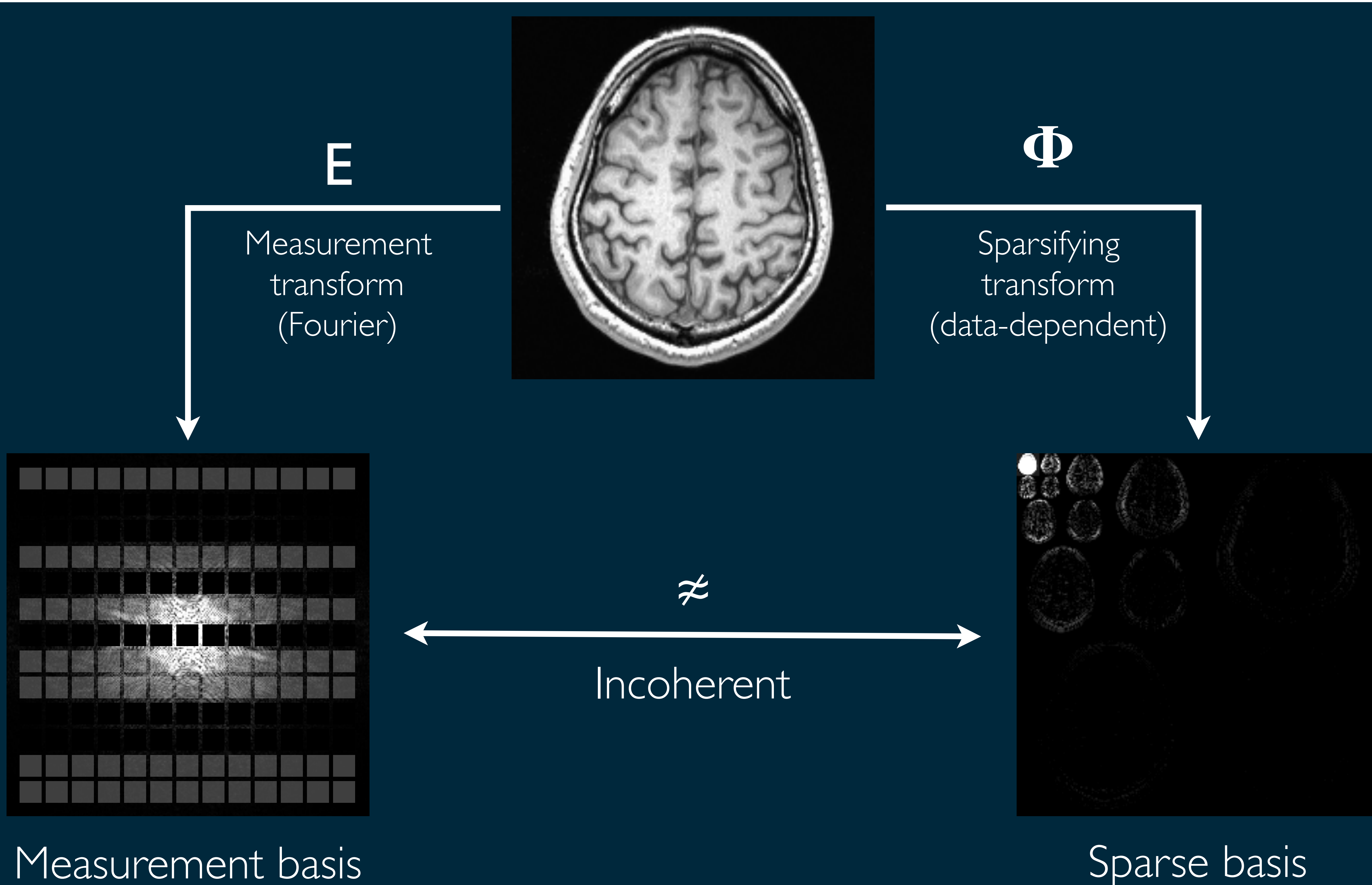
Sparsity and Compressibility

- Compression reduces the inefficiency in data representation



- CS exploits compressibility not for storage, but for speed
  - Fewer effective unknowns ➞ fewer measurements
  - Amount of acceleration is directly linked to the amount of sparsity (or compressibility)

Sampling Requirements



CS Reconstruction

- How to find sparse solutions?
  - The L1 norm is a good approximation of sparsity
- "Find the sparsest solution consistent with the data" becomes:

Sparsifying Transform

L1 norm

$$\min ||\Phi x||_1$$
$$\text{such that } k = Ex$$

Data consistency

CS Reconstruction

$$\begin{aligned} &\min ||\Phi x||_1 \\ &\text{s.t. } k = Ex \end{aligned} \xrightarrow[\text{to be noisy}]{\text{Allow measurements}} \begin{aligned} &\min ||\Phi x||_1 \\ &\text{s.t. } ||k - Ex||_2 < \epsilon \end{aligned}$$

- You will often see it in an unconstrained (Lagrangian) form:

$$\min_x \frac{1}{2}||k - Ex||_2^2 + \lambda||\Phi x||_1$$

Recon as an Optimisation

This is called the objective or cost function

$$\min_x \frac{1}{2}||k - Ex||_2^2 + \lambda||\Phi x||_1$$

Sparsity promoting term

Data consistency promoting term

Recon as an Optimisation

This formulation applies in conventional (non Compressed Sensing) imaging as well

$$\min_x \frac{1}{2}\underbrace{||k - Ex||_2^2}$$

This is just the least squares problem

Recon as an Optimisation

$$\min_x \frac{1}{2}||k - Ex||_2^2$$

Differentiating and setting the result equal to zero yields the normal equation:

$$E^* E x = E^* k$$

Which leads to the closed form solution:

$$x = (E^* E)^{-1} E^* k$$

Iterative Reconstruction

- We often want to avoid explicitly forming inverses
- In Compressed Sensing, the optimisation problem is non-linear
- We often turn to iterative methods
  - Different approaches can be taken depending on whether the objective function is differentiable at all or not
  - 1st order gradient methods are common in image reconstruction

Iterative Reconstruction

- Essentially construct a sequence of estimates $\boldsymbol{x^n}$ such that

$$x^{n+1} = f(x^n)$$

- where $\boldsymbol{f}$ consists of some "cheap" operations on $\boldsymbol{x^n}$
- The exact form of $\boldsymbol{f}$ depends on your objective function and your iterative algorithm

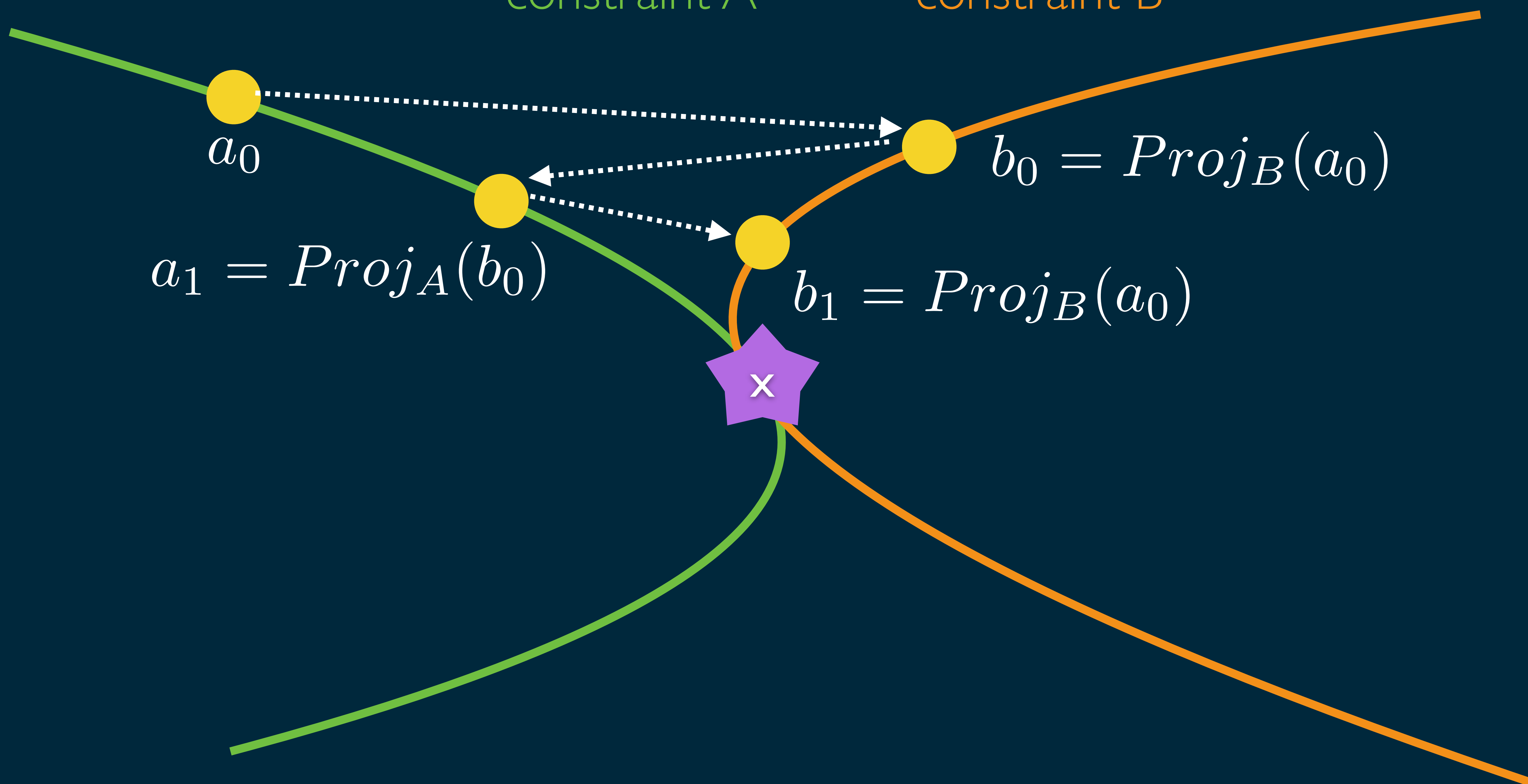
- After some criteria is met, for some n, $\boldsymbol{x^n}$ is your final answer

Some algorithmic suggestions…

Projection onto Convex Sets

$$\min_x \frac{1}{2} \boxed{||k - Ex||}_2^2 + \lambda \boxed{|\Phi x|}_1$$

constraint A   constraint B



Projection onto Convex Sets

$$u^{n+1} = f(x^n)$$

$$x^{n+1} = g(u^{n+1})$$

$$x^0, x^1, \ldots, x^n$$

should converge to a solution that is in the intersection of both constraint sets, if the sets are convex

$f(x)$ should be a projection onto the set of solutions where $\|Ex - k\|_2 < \epsilon$

$g(x)$ should be a projection onto the set of solutions where $\|\Psi x\|_1 < \alpha$

Some suggestions

Geometric constraint approach

Gradient-based approach

Guess and check approach

Good luck with the practical!

mark.chiew@ndcn.ox.ac.uk