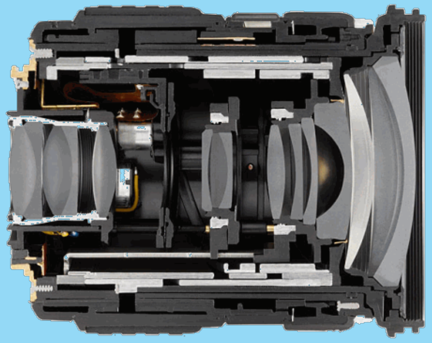
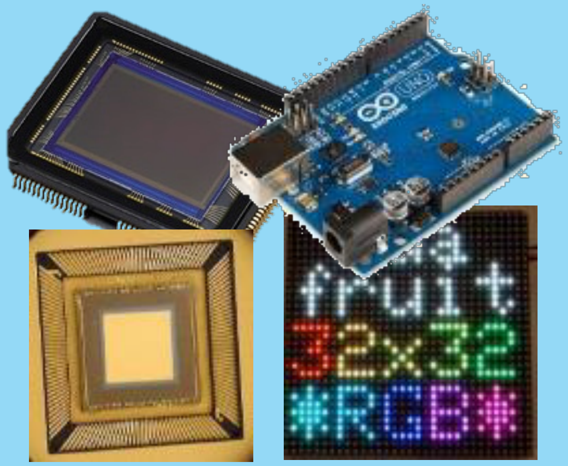


Optics

+



Sensors
&
devices

+

Signal
processing
&
algorithms

COMPUTATIONAL METHODS FOR IMAGING (AND VISION)

LECTURE 24: NONLINEAR INVERSION - BEYOND ℓ_2 REGULARIZATION

PROF. JOHN MURRAY-BRUCE

WHERE ARE WE

10	15-Mar-21	Forward Models and Inverse Problems	Linear Inversion - Inverse problems - Deconvolution and Denoising	IIP 4, Appendix E	**HW 4**	
	17-Mar-21		Intro to Regularized Inversion I - Tikhonov	IIP 5, Appendix E		
11	22-Mar-21	Regularization	Intro to Regularized Inversion II - Iterative methods - Steepest descent	IIP 6		
	24-Mar-21		Statistical methods I - ML estimation - Bayesian estimation	IIP 7.1 - 7.5		
12	29-Mar-21	Forward models and Inverse Problems II	LSV imaging systems: Forward problem - SVD - Inversion	IIP 8.1, 9, 10		
	31-Mar-21	Non-linear Regularization	Beyond L_2-regularization - Sparsity (l_0- and l_1-priors) - TV prior	SMIV 1.1 - 1.5 Papers & Handout		**HW 4**
13	5-Apr-21		~~Algorithm overview~~ ~~- ISTA/FISTA~~ ~~- ADMM~~	Papers & Handout		
	7-Apr-21		Beyond L_2 regularization			
14	12-Apr-21 14-Apr-21	**Spring Break (no classes)**				
15	19-Apr-21 21-Apr-21	Applications of Comp. Imaging	Intro to optics & Applications			
16	26-Apr-21 28-Apr-21	**Group Presentations (Teams)**				
17	3-May-21	***no class**				
	5-May-21	**Final Exam: 12:30 PM - 2:30 PM**				

WE ARE HERE!

NONLINEAR REGULARIZATION

OUTLINE

- **Beyond ℓ_2 regularization**
  - ℓ_1-norm regularization
  - TV regularization
- **Robustness to noise**
- **Matlab demonstrations**

NONLINEAR REGULARIZATION

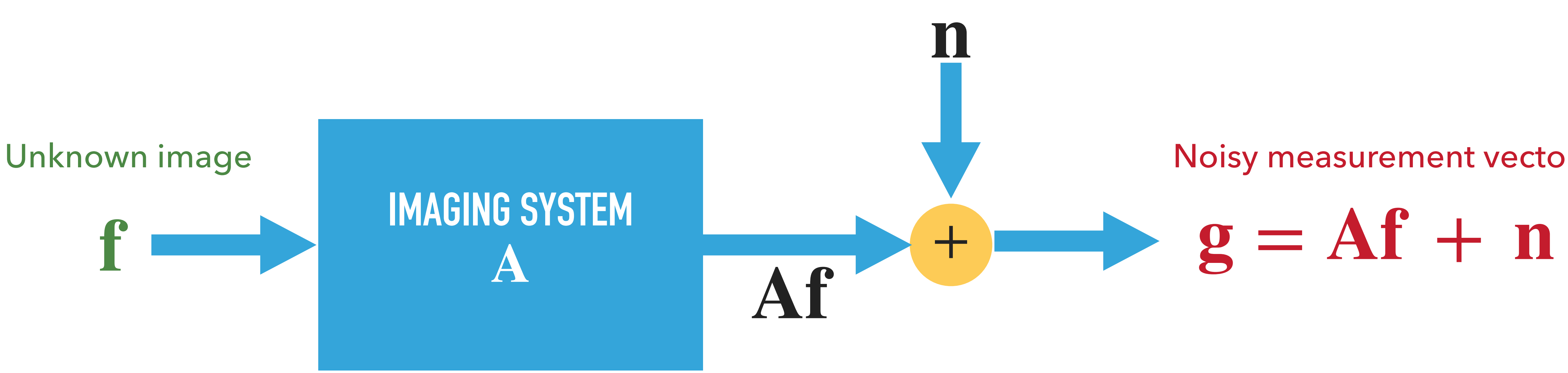
LEARNING OUTCOME

- Identify the key benefits of ℓ_1 regularization
- Identify the key benefits of TV regularization
- Solve linear inverse imaging problems with ℓ_1 and TV regularization

BRIEF REVIEW OF LINEAR INVERSION

ℓ_2 REGULARIZATION

DISCRETE MODEL FOR IMAGING SYSTEMS



- Given the measurements $\mathbf{g}$, we wish to recover an estimate of the input $\mathbf{f}$
  - This is a discrete **linear inverse problem**
- Discussed different linear inverses:
  - Standard inverse (square matrix), Pseudo-inverse, Tikhonov regularized inverse

DISCRETE MODEL FOR IMAGING SYSTEMS
LINEAR INVERSION

- **Discrete model:** linear imaging systems modeled as, $\mathbf{g} = \mathbf{A}\mathbf{f} + \mathbf{n}$
- **Standard Inverse:** works only if $\mathbf{A}$ can be inverted
- **Moore-Penrose pseudo-inverse:**

$$\min_{\mathbf{f}} \|\mathbf{g} - \mathbf{A}\mathbf{f}\|^2$$

  - Solution: $\mathbf{f}_{\text{est}} = \mathbf{A}^{\dagger}\mathbf{g}$, where $\mathbf{A}^{\dagger} = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$

- **Tikhonov regularized inverse**

$$\min_{\mathbf{f}} \|\mathbf{g} - \mathbf{A}\mathbf{f}\|_2^2 + \lambda\|\mathbf{f}\|_2^2$$

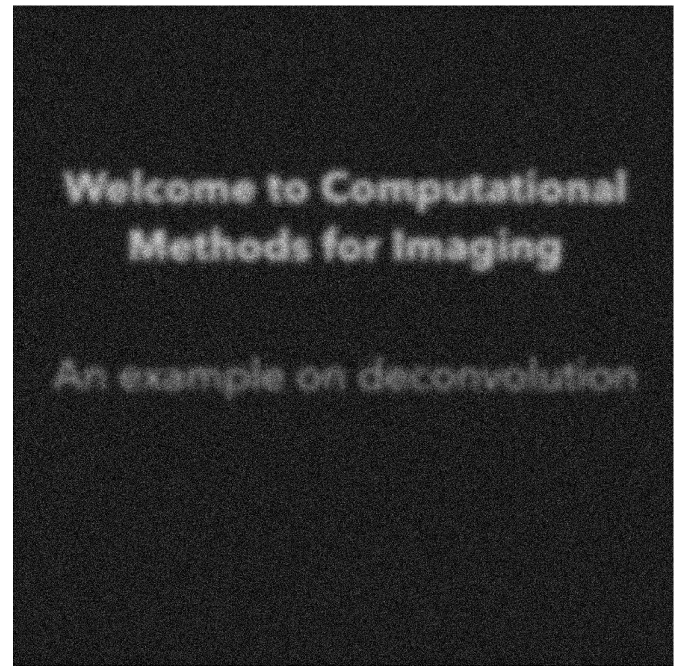
Represents our prior/initial beliefs about the structure of f, i.e. **Bounded 2-norm**

  - Solution: $\mathbf{f}_{\text{est}} = \mathbf{A}^{\dagger}_{\text{Tik}}\mathbf{g}$, where $\mathbf{A}^{\dagger}_{\text{Tik}} = (\mathbf{A}^H\mathbf{A} + \lambda\mathbf{I})^{-1}\mathbf{A}^H$

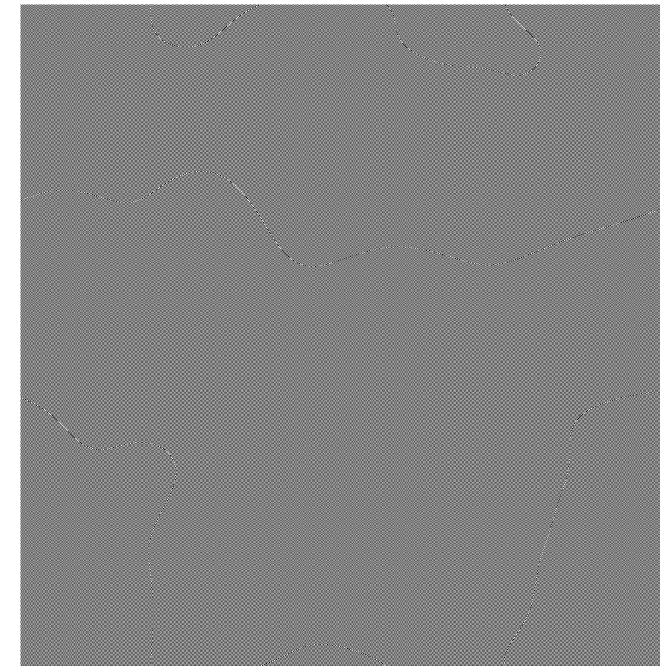
COMPARISON

- Standard inverse and pseudo-inverse produce garbage
- Tikhonov-regularized estimate provides some improvement

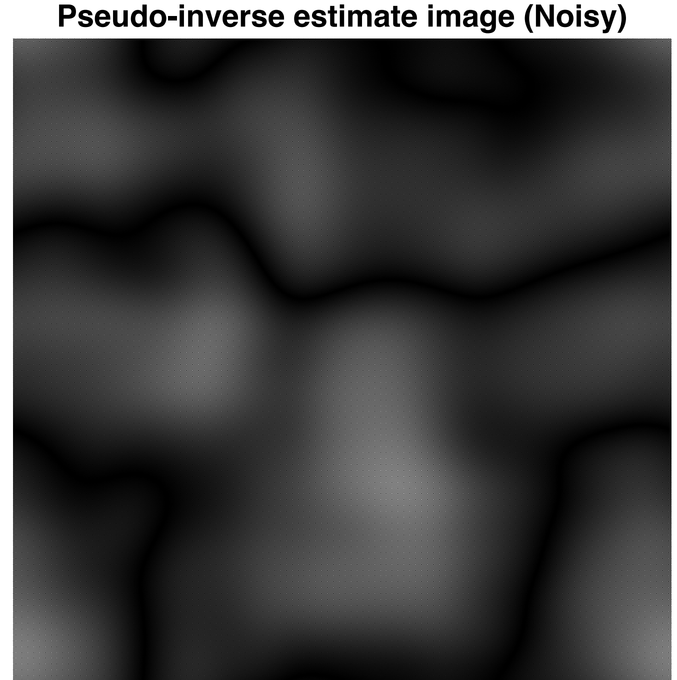
Measurement



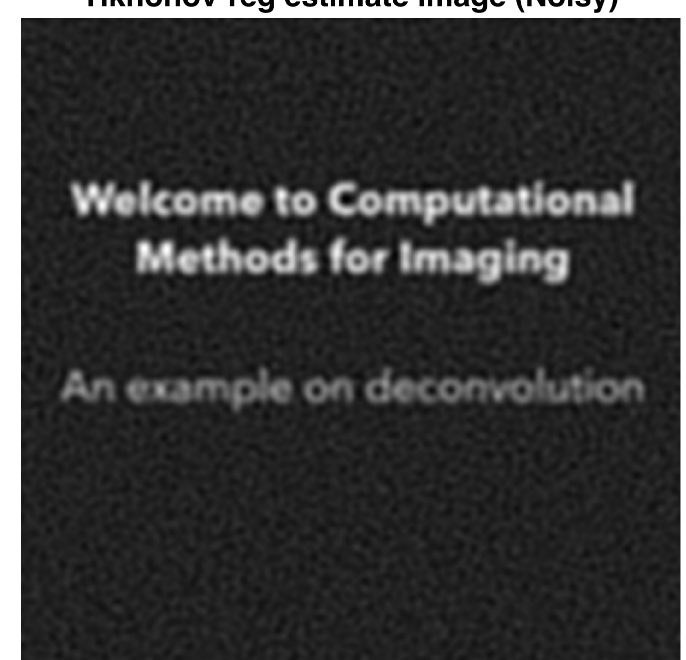
Standard inverse estimate image (Noisy)



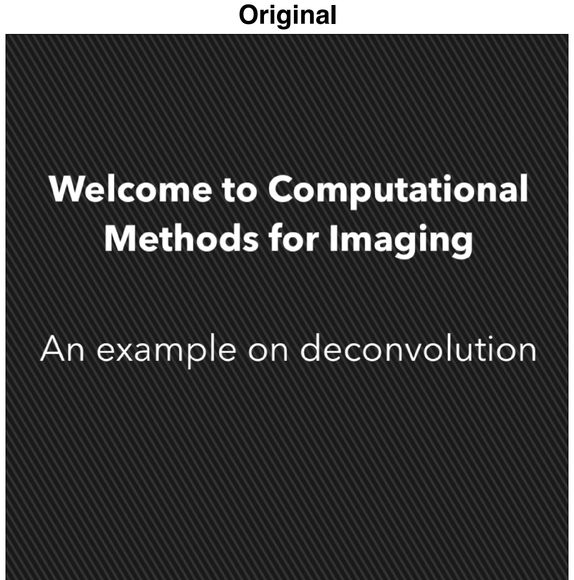
Pseudo-inverse estimate image (Noisy)



Tikhonov-reg estimate image (Noisy)



Original

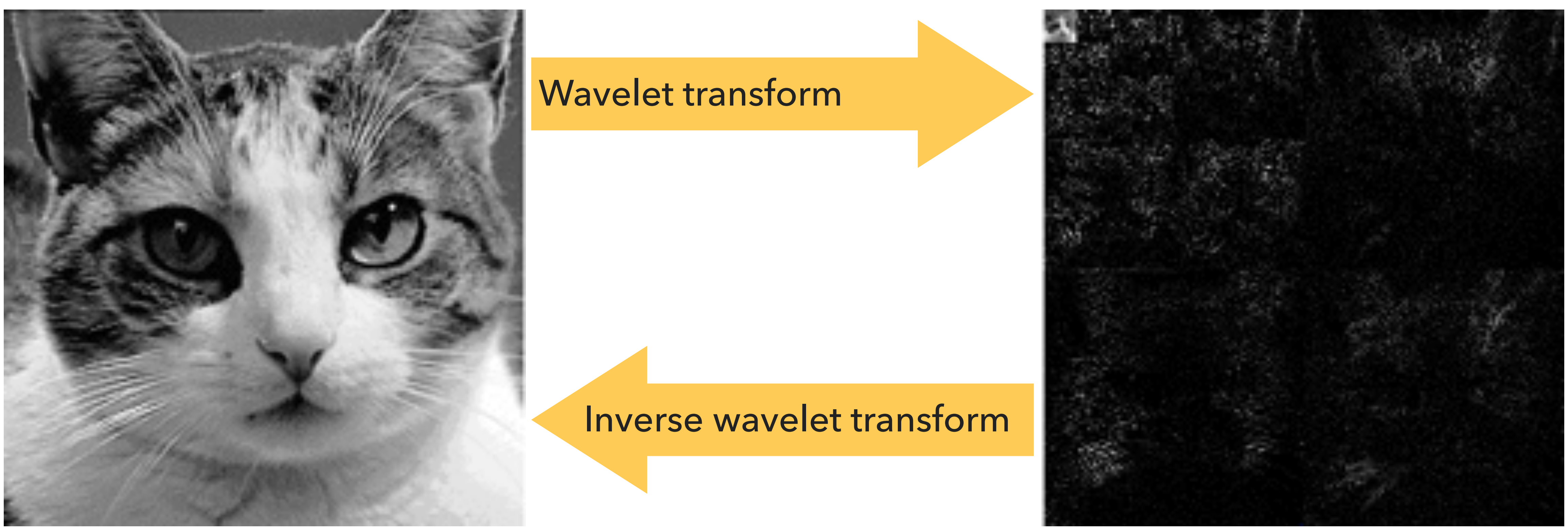


NONLINEAR INVERSION

BEYOND ℓ_2 REGULARIZATION

SPARSITY IS A NATURAL PRIOR FOR IMAGES

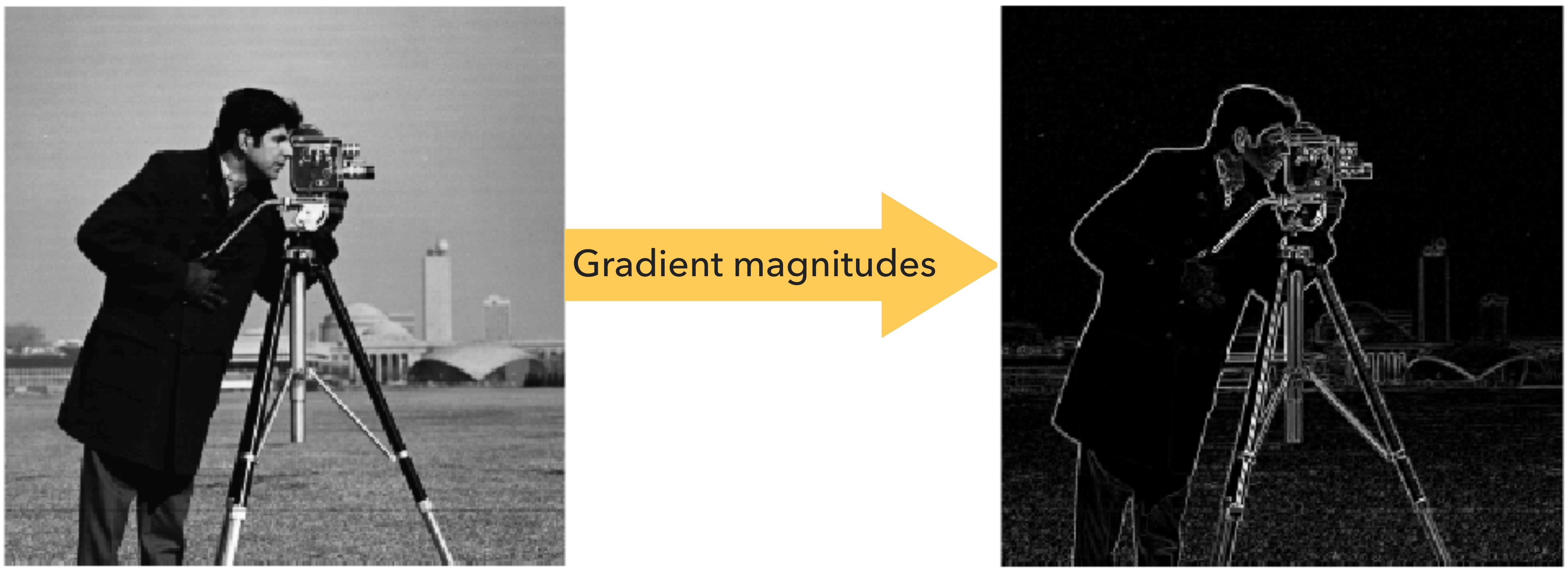
- A vector/matrix is said to be **sparse** if an overwhelming majority of its entries are equal to zero
- **Example:** sparsity in a *wavelet* basis (wavelet domain)



$$\mathbf{f} = \mathbf{W}^T\mathbf{c} \qquad \mathbf{c} = \mathbf{W}\mathbf{f}$$

SPARSITY IS A NATURAL PRIOR FOR IMAGES

- **Example 2:** sparsity in of the image *gradient*



$\mathbf{f}$ $\quad\quad$ $\mathbf{u} = \mathbf{D}\mathbf{f}$

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix}$$

A SPARSITY ENFORCING PRIORS

- **Idea:** Incorporate that information as a prior when estimating images
  - Seek estimates that are sparse in some transform domain
  - For natural images, sparsity in the wavelet domain or gradient is common
  - One could try to learn the transforms **(Dictionary Learning)**
    - But, in general, a basis where the object is sparse is a bit of an art form

THE ASSOCIATED MINIMIZATION PROBLEMS

ENFORCING SPARSITY

A SPARSITY ENFORCING PRIOR

THE ℓ_1-NORM

- **Sparse solutions:** the ℓ_1-norm constraint (under certain mathematical conditions) leads to sparse solutions

- Specifically,

This is the ℓ_1 regularization term (or sparsity promoting term).

$$\min_{\mathbf{f}} \|\mathbf{g} - \mathbf{A}\mathbf{f}\|_2^2 + \lambda\|\mathbf{W}\mathbf{f}\|_1$$

- $\mathbf{W}$ is some *a priori* chosen basis, the wavelet, Fourier, DCT

A SPARSITY ENFORCING PRIOR

THE ℓ_1-NORM

- **ℓ_1-regularized inversion:**

$$\min_{\mathbf{f}} \|\mathbf{g} - \mathbf{A}\mathbf{f}\|_2^2 + \lambda \|\mathbf{W}\mathbf{f}\|_1$$

  - $\mathbf{W}$ is some *a priori* chosen basis, the wavelet, Fourier, DCT

- **No closed form solution, unfortunately :-(**
- Must solve numerically using (iterative optimization methods). E.g.:
  - Conjugate gradient methods (like FISTA, ISTA)
  - Alternating Direction Method of Multipliers (ADMM)
  - Many implementations exist in pretty much any programming language.
  - We will use it as a black box! ;-)

A SPARSITY ENFORCING PRIOR

THE ℓ_1-NORM INTUITION

- **ℓ_1-regularized inversion:**

$$\min_{\mathbf{f}} \|\mathbf{g} - \mathbf{A}\mathbf{f}\|_2^2 + \lambda\|\mathbf{W}\mathbf{f}\|_1$$

  - $\mathbf{W}$ is some *a priori* chosen basis, the wavelet, Fourier, DCT
  - It could also be the identity matrix $\mathbf{I}$
- Recall the definition for ℓ_p-norm of a vector $\mathbf{f} \in \mathbb{R}^N$:

$$\|\mathbf{f}\|_p = \left(|f_0|^p + |f_1|^p + \cdots + |f_{N-1}|^p \right)^{1/p},$$

  where $|f_n|$ is the absolute value of f_n
- **Sparse solutions intuition:** Let's look at a simple linear system

WHY DOES THIS WORK?

ENFORCING SPARSITY

SOLUTIONS TO LINEAR SYSTEMS
USING DIFFERENT NORMS

- Consider a very simple linear system (i.e. $\mathbf{g} = \mathbf{A}\mathbf{f}$)
  - The solutions behave very differently under different **norm constraints** on $\mathbf{f}$
  - Let the linear system be:

$$g_0 = (1/5)f_0 + (2/5)f_1$$

  - Or, in matrix-vector form

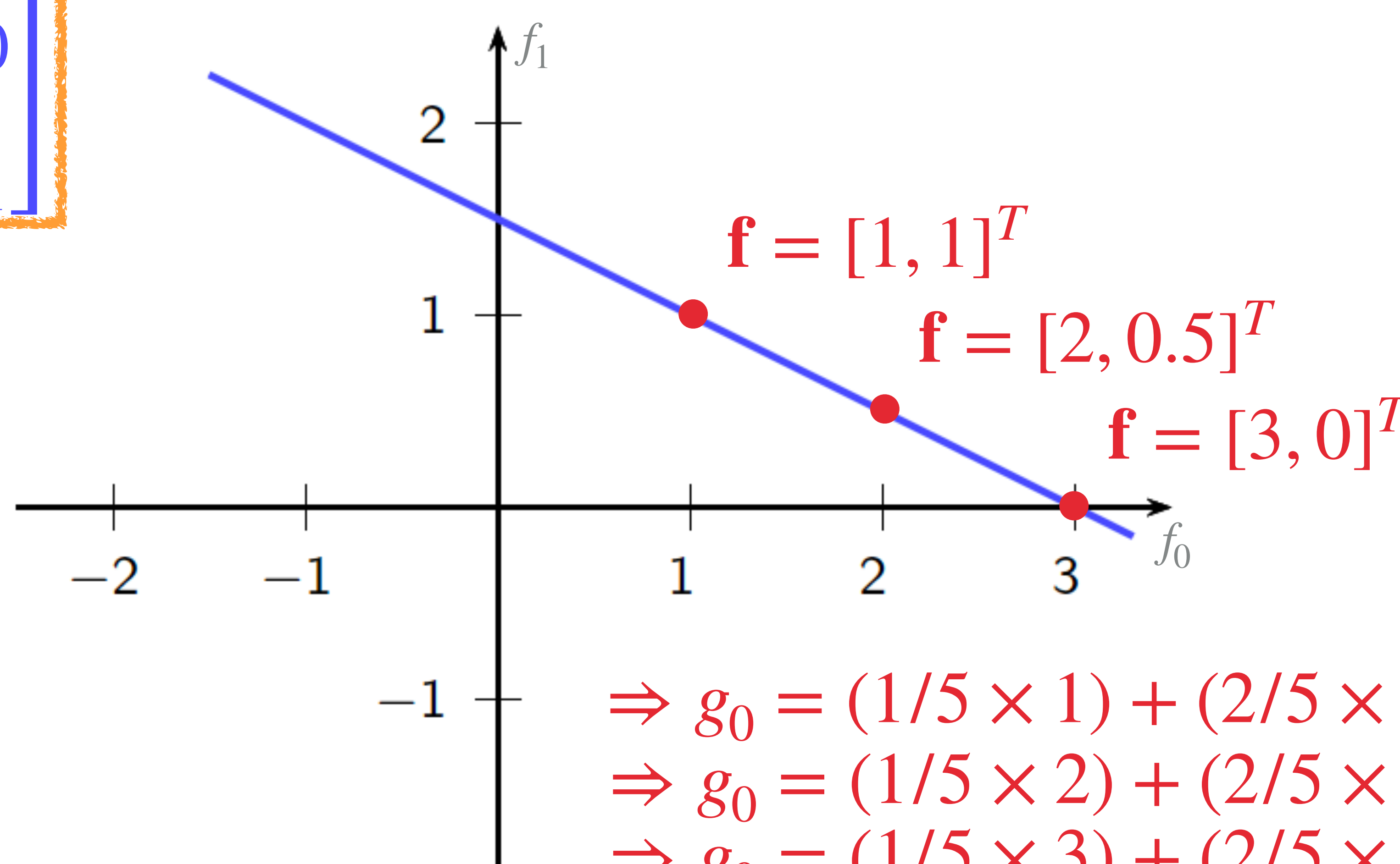
$$\underbrace{\boxed{g_0}}_{\mathbf{g}} = \underbrace{\boxed{\begin{bmatrix} 1/5 & 2/5 \end{bmatrix}}}_{\mathbf{A}} \underbrace{\boxed{\begin{bmatrix} f_0 \\ f_1 \end{bmatrix}}}_{\mathbf{f}}$$

SOLUTIONS TO LINEAR SYSTEMS
USING DIFFERENT NORMS

- Consider a very simple linear system (i.e. $\mathbf{g} = \mathbf{A}\mathbf{f}$)
  - The solutions behave very differently under different **norm constraints** on $\mathbf{f}$

$$\boxed{g_0 = [1/5 \quad 2/5] \begin{bmatrix} f_0 \\ f_1 \end{bmatrix}}$$

$$g_0 = 3/5$$



$$\Rightarrow g_0 = (1/5 \times 1) + (2/5 \times 1) = 3/5$$
$$\Rightarrow g_0 = (1/5 \times 2) + (2/5 \times 0.5) = 3/5$$
$$\Rightarrow g_0 = (1/5 \times 3) + (2/5 \times 0) = 3/5$$

INVESTIGATE SOLUTIONS UNDER DIFFERENT NORM CONSTRAINTS

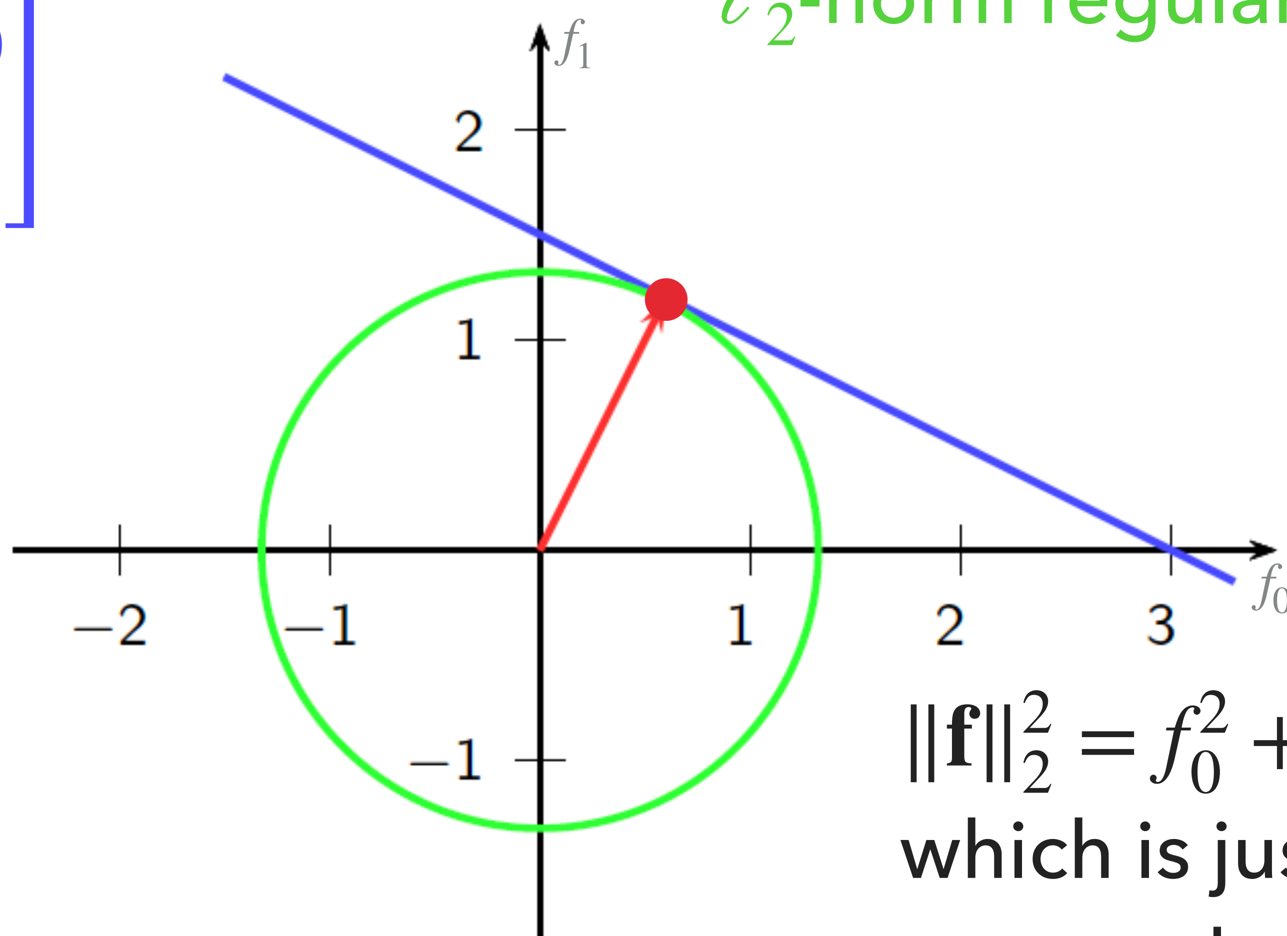
SOLUTIONS TO LINEAR SYSTEMS
USING DIFFERENT NORMS

- Consider a very simple linear system (i.e. $\mathbf{g} = \mathbf{A}\mathbf{f}$)
  - The solutions behave very differently under different **norm constraints** on $\mathbf{f}$

$$g_0 = [1/5 \quad 2/5] \begin{bmatrix} f_0 \\ f_1 \end{bmatrix}$$

$$g_0 = 3/5$$

ℓ_2-norm regularization, i.e. $||\mathbf{f}||_2^2$



$$||\mathbf{f}||_2^2 = f_0^2 + f_1^2$$

which is just a circle centered at the origin

SOLUTIONS TO LINEAR SYSTEMS
USING DIFFERENT NORMS

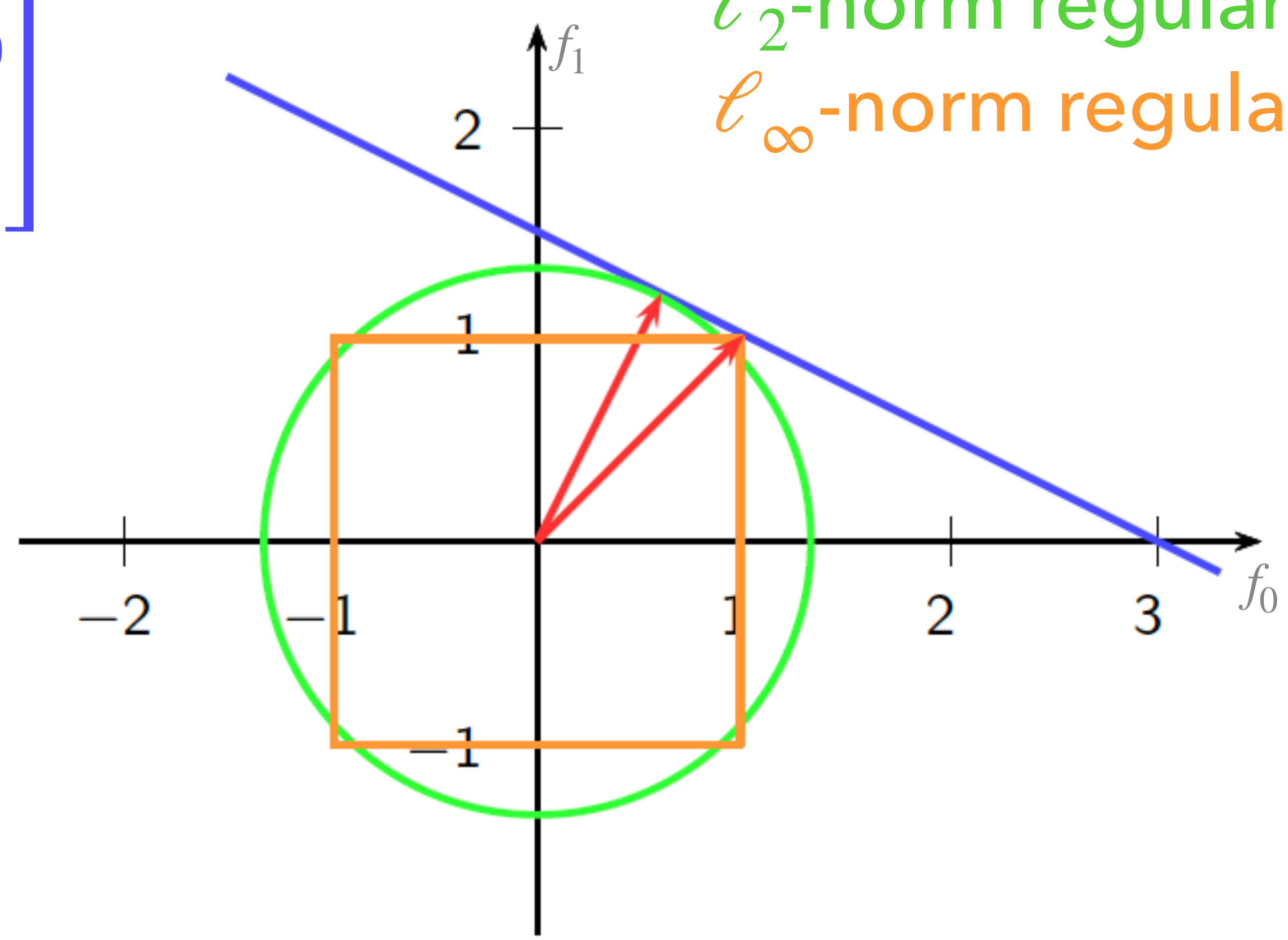
- Consider a very simple linear system (i.e. $\mathbf{g} = \mathbf{A}\mathbf{f}$)
  - The solutions behave very differently under different **norm constraints** on $\mathbf{f}$

$$g_0 = [1/5 \quad 2/5] \begin{bmatrix} f_0 \\ f_1 \end{bmatrix}$$

$$g_0 = 3/5$$

ℓ_2-norm regularization, i.e. $||\mathbf{f}||_2^2$

ℓ_∞-norm regularization



SOLUTIONS TO LINEAR SYSTEMS
USING DIFFERENT NORMS

- Consider a very simple linear system (i.e. $\mathbf{g} = \mathbf{A}\mathbf{f}$)
  - The solutions behave very differently under different **norm constraints** on $\mathbf{f}$

$$g_0 = [1/5 \quad 2/5] \begin{bmatrix} f_0 \\ f_1 \end{bmatrix}$$

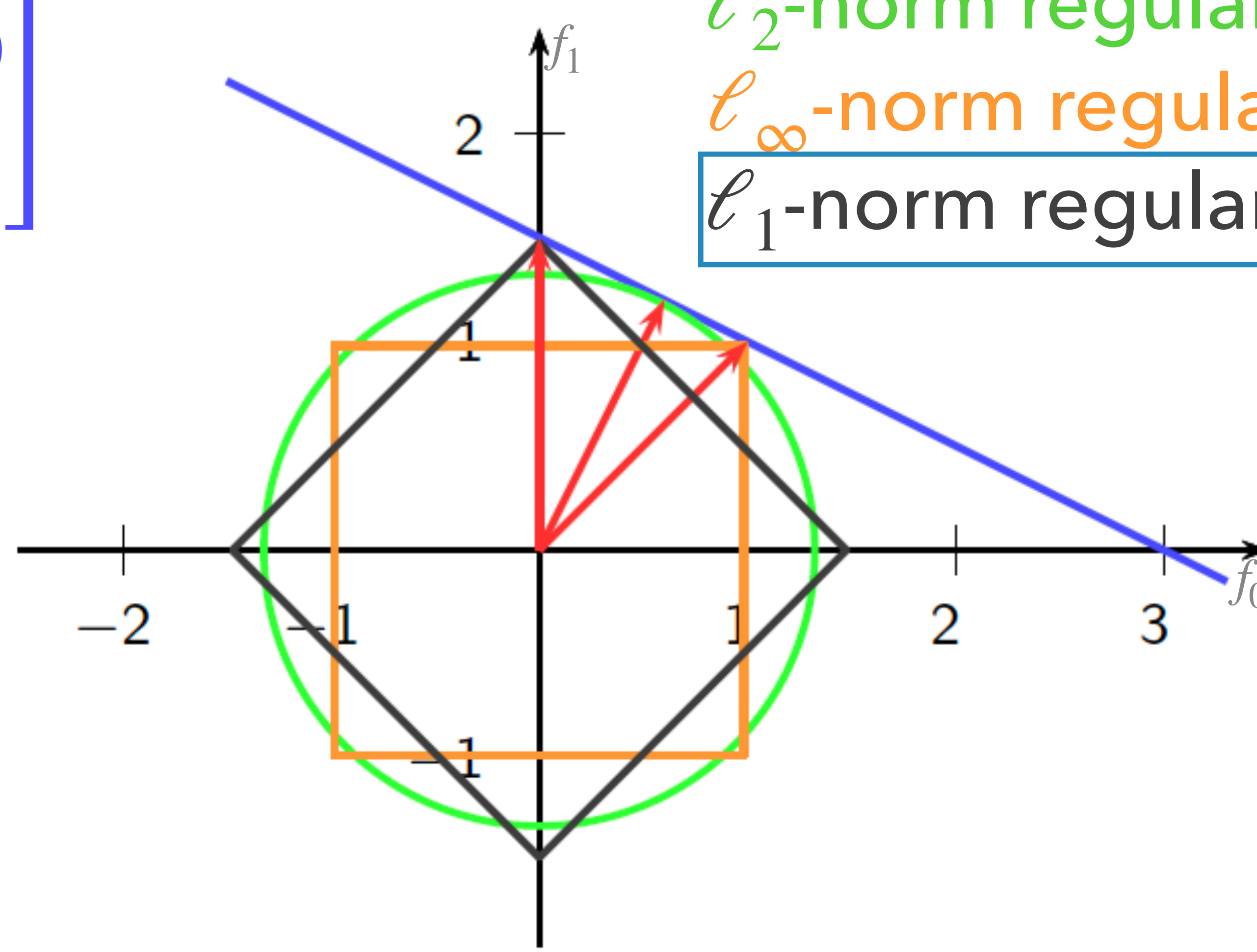
$$g_0 = 3/5$$

ONE OF THESE NORMS GIVES THE SPARSEST SOLUTION!

ℓ_2-norm regularization, i.e. $||\mathbf{f}||_2^2$

ℓ_∞-norm regularization

$\boxed{\ell_1\text{-norm regularization, i.e. } ||\mathbf{f}||_1}$

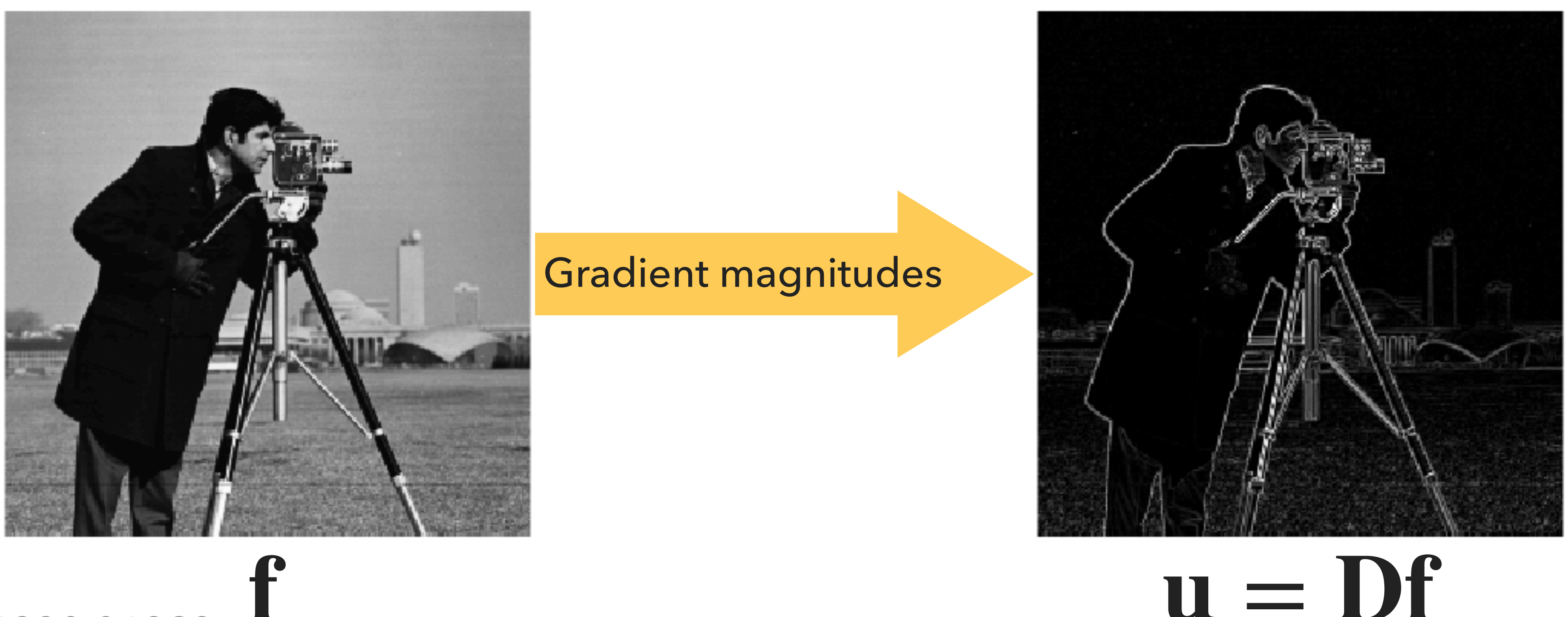


ENFORCING SPARSITY IN THE GRADIENT
- PIECEWISE SMOOTHNESS

TOTAL VARIATION

TOTAL VARIATION PRIOR

- **Total Variation:** Here, the gradient of the image is assumed to be sparse
  - This is called the total variation regularization



TOTAL VARIATION (TV) REGULARIZATION

THE TV SEMI-NORM

- **ℓ_1-regularized inversion:**

$$\min_{\mathbf{f}} \|\mathbf{g} - \mathbf{A}\mathbf{f}\|_2^2 + \lambda \, \mathrm{TV}(\mathbf{f})$$

  - The total variation (semi)-norm term is defined as:
  - $\mathrm{TV}(\mathbf{f}) = |f_1 - f_0| + |f_2 - f_1| + \cdots + |f_{N-1} - f_{N-2}|$
- This constraint promotes piecewise constant reconstructions
  - Does a better job at preserving edges (compared to Tikhonov)

MATLAB CODE

See '`Lecture24_TVReg_Inverse_Prob.m`' on canvas under *Beyond L2 Regularization* module.

An implementation of FISTA algorithm for solving the TV regularized inverse problem is the function '`fistatvrecon.m`' also on canvas (Lecture 24).

For mathematical details of the algorithm, can look at the papers:

[1] Amir Beck, and Marc Teboulle. "A fast iterative shrinkage-thresholding algorithm for linear inverse problems." *SIAM journal on imaging sciences,* 2(1), 183-202, 2009.
[2] Amir Beck, and Marc Teboulle. "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems." *IEEE transactions on image processing,* 18(11), 2419-2434, 2009.

ℓ_1 REGULARIZATION

TV REGULARIZATION

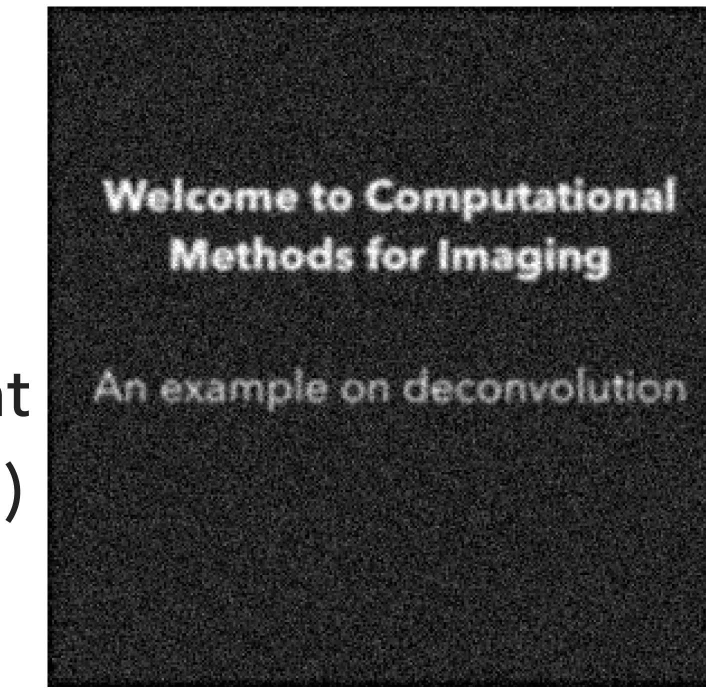
MATLAB PRACTICE 10: NONLINEAR INVERSION

NUMERICAL RECONSTRUCTIONS

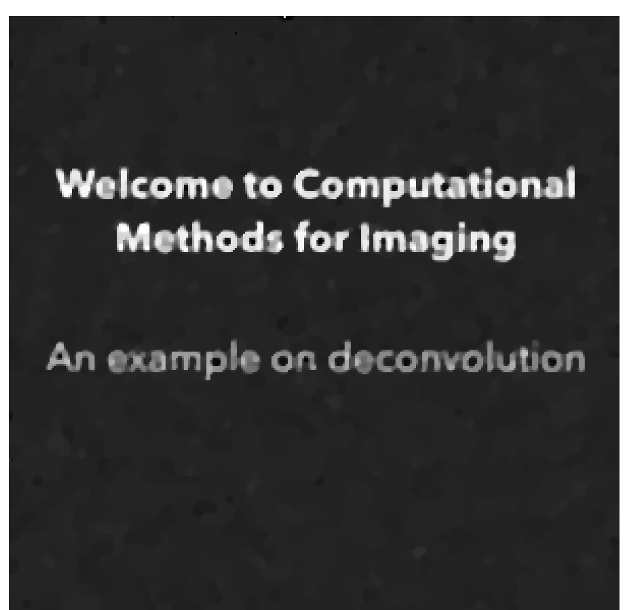
NUMERICAL RESULTS: TOTAL VARIATION REGULARIZED

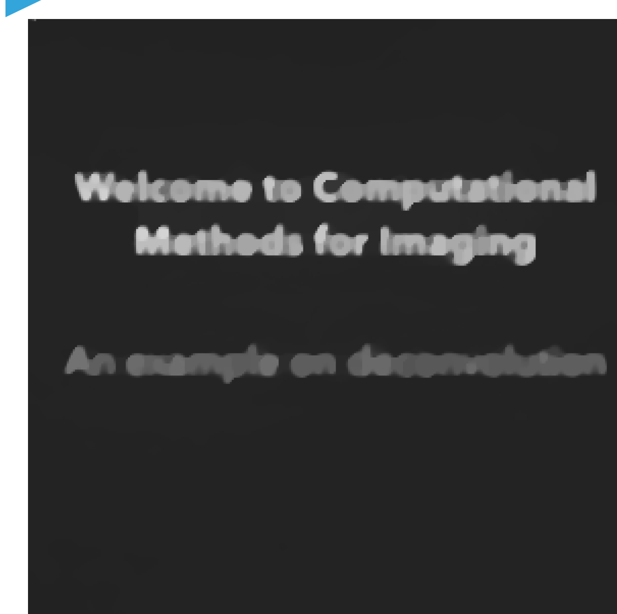
Ground truth

Blurred and Noisy Measurement (SNR = 10dB)



Increasing λ



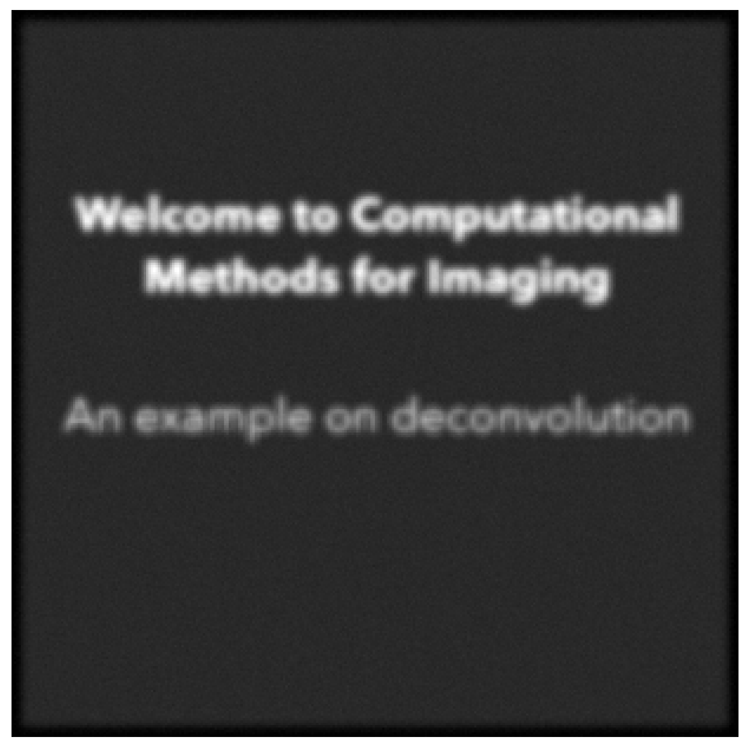


NUMERICAL RESULTS: TOTAL VARIATION REGULARIZED

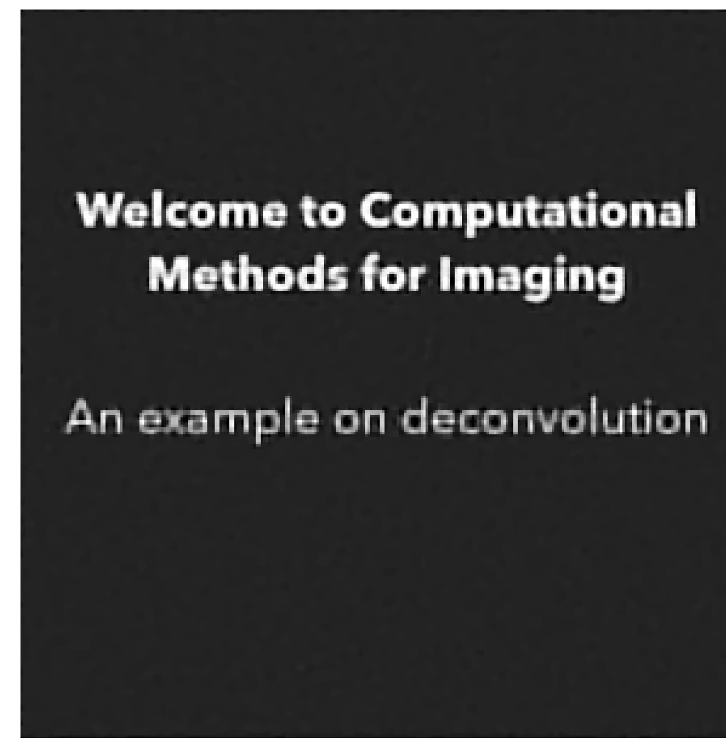
INCREASING LEVEL OF BLURRING

Ground truth

Measurement



TV reconstruction

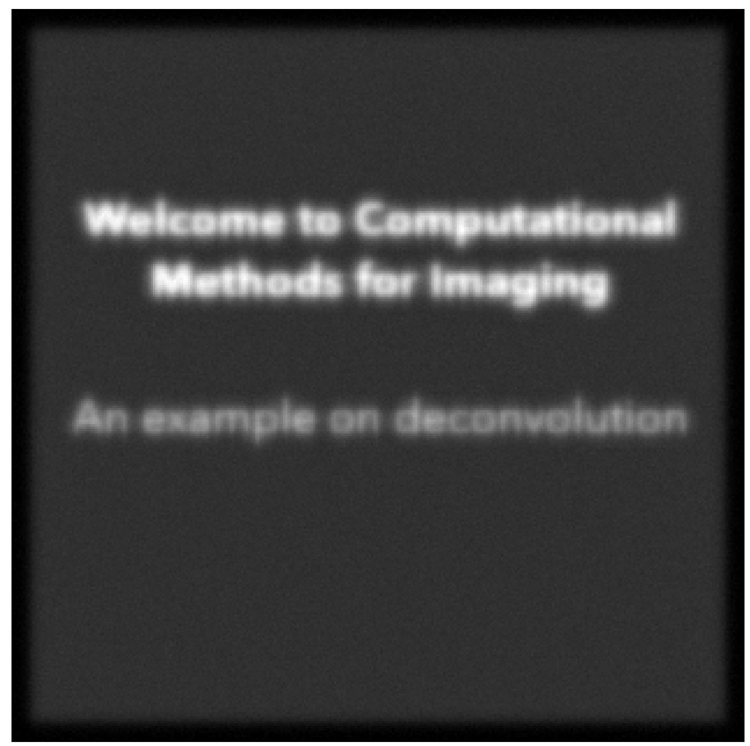


NUMERICAL RESULTS: TOTAL VARIATION REGULARIZED

BLURRING INCREASED FURTHER

Ground truth

Measurement



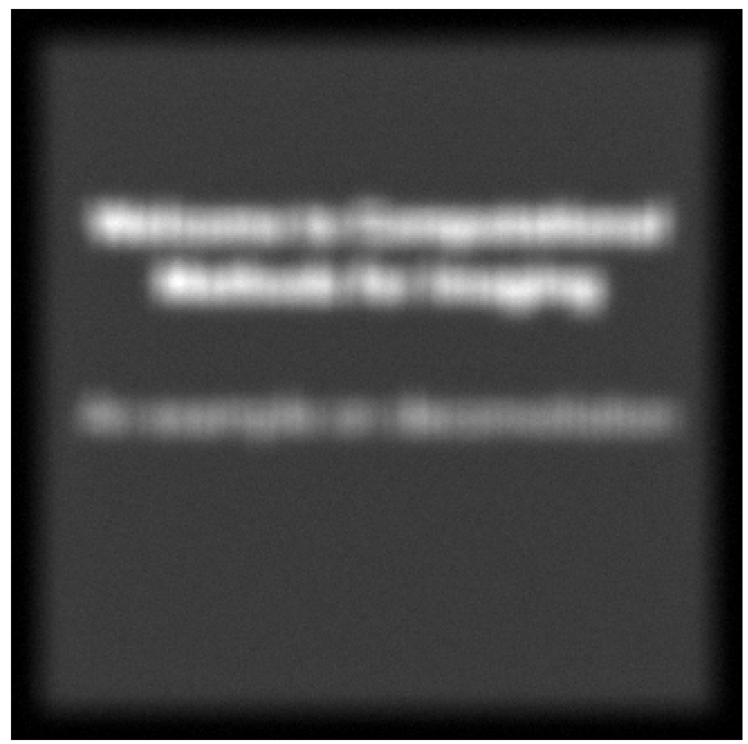
TV reconstruction

NUMERICAL RESULTS: TOTAL VARIATION REGULARIZED

BLURRING INCREASED (EVEN) FURTHER

Ground truth

Measurement



TV reconstruction

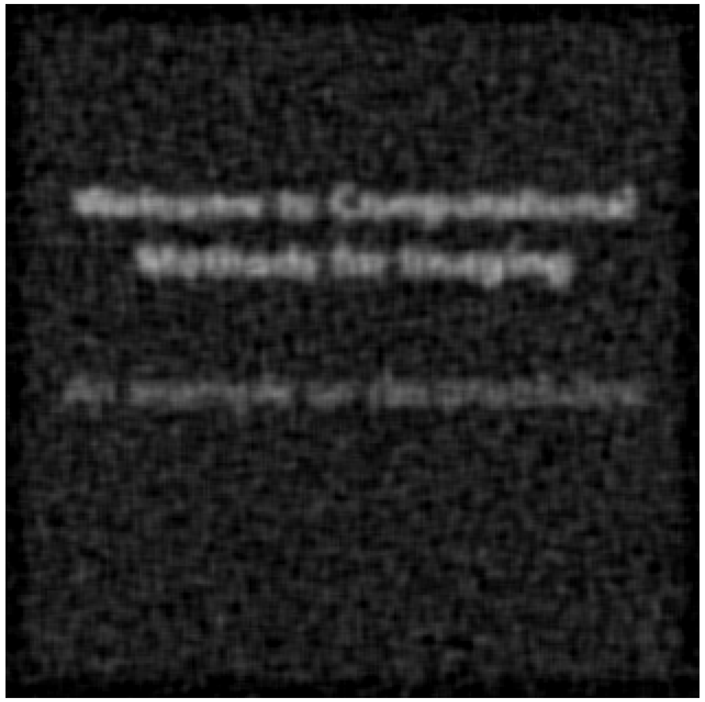
Severe blurring, TV struggles but text still somewhat legible

NUMERICAL RESULTS: TOTAL VARIATION REGULARIZED
COMPARING TV AND TIKHONOV

Measurement

TV reconstruction

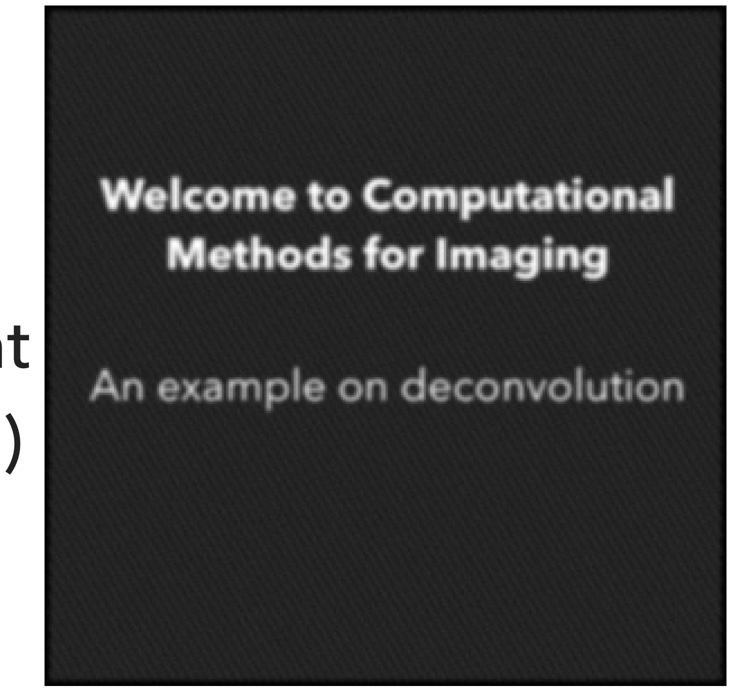
Tikhonov



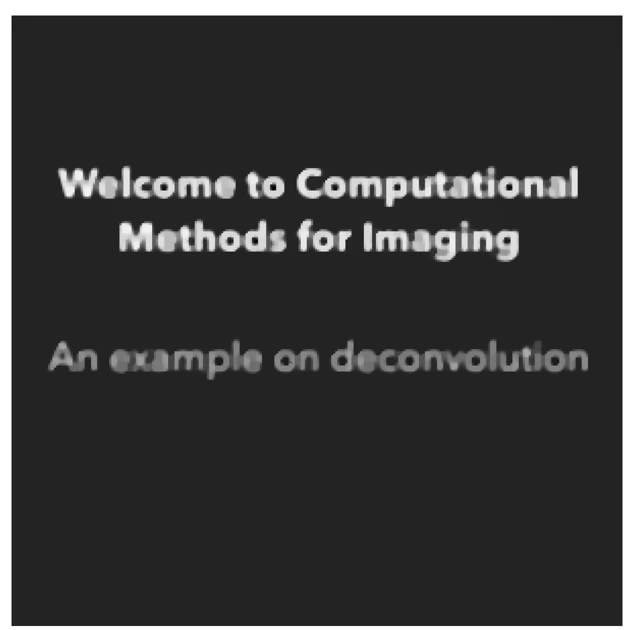
NUMERICAL RESULTS: TOTAL VARIATION REGULARIZED

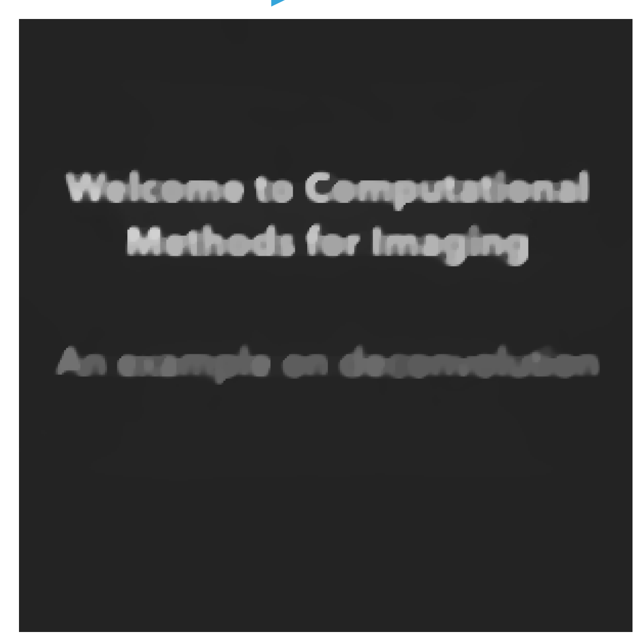
Ground truth

Noisy Measurement (SNR = 30dB)



Increasing λ





SUMMARY

- We learnt about LSI systems
  - Deriving continuous and discrete models
  - Discrete models enable computational/digital implementation
- Forward and inverse problems in imaging
- Solving inverse problems (examples on blurring)
  - Non-regularized inversion:
    - Standard inverse and Least squares (MP Pseudo-inverse)
  - Regularized inversion
    - **Linear:** Tikhonov regularization
    - **Non-linear:** ℓ_1 and total variation regularization

THIS LECTURE

REGULARIZATION IMPORTANT CONCEPT IN MACHINE LEARNING TOO!

NEXT TIME!

INTRO TO OPTICS