# Single-pixel Imaging and ADMM

EE367/CS448I: Computational Imaging
stanford.edu/class/ee367
Lecture 11

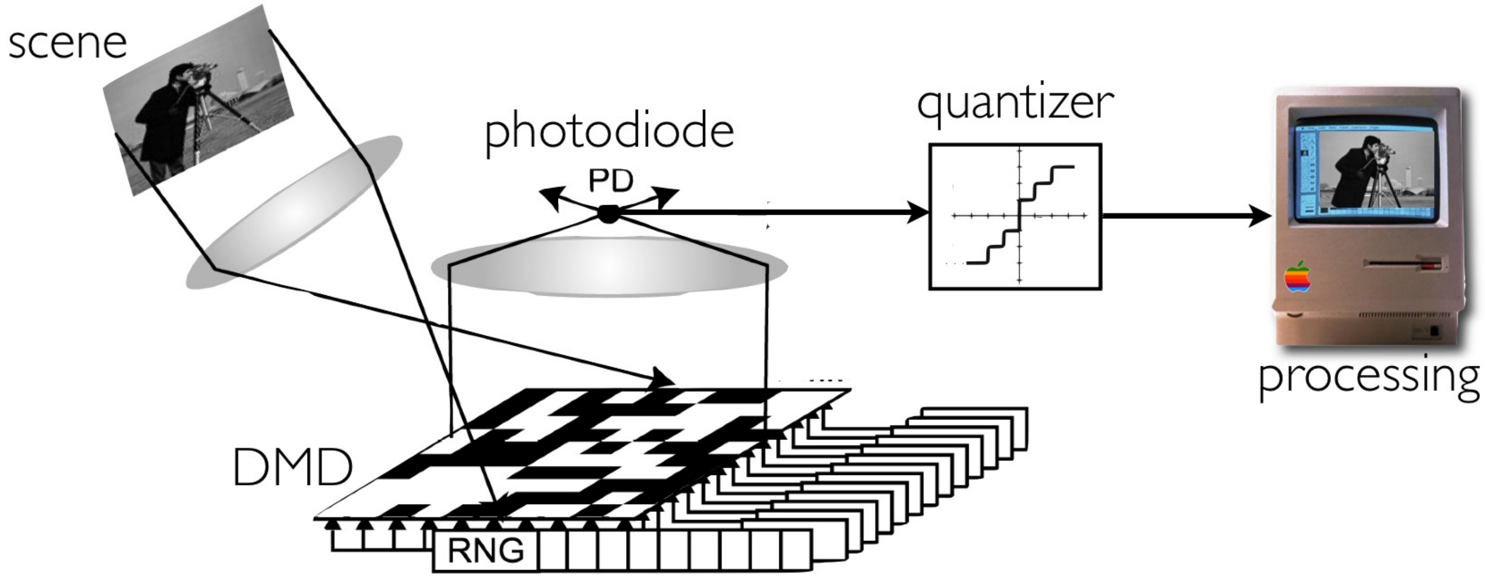
Gordon Wetzstein
Stanford University

# Overview

- Single-pixel imaging and other compressive imaging problems
- HQS for general inverse problems & compressive imaging
- The Alternating Direction Methods of Multipliers (ADMM)
- ADMM for general inverse problems & compressive imaging
- Outlook on using ADMM with Poisson noise and multiple regularizers

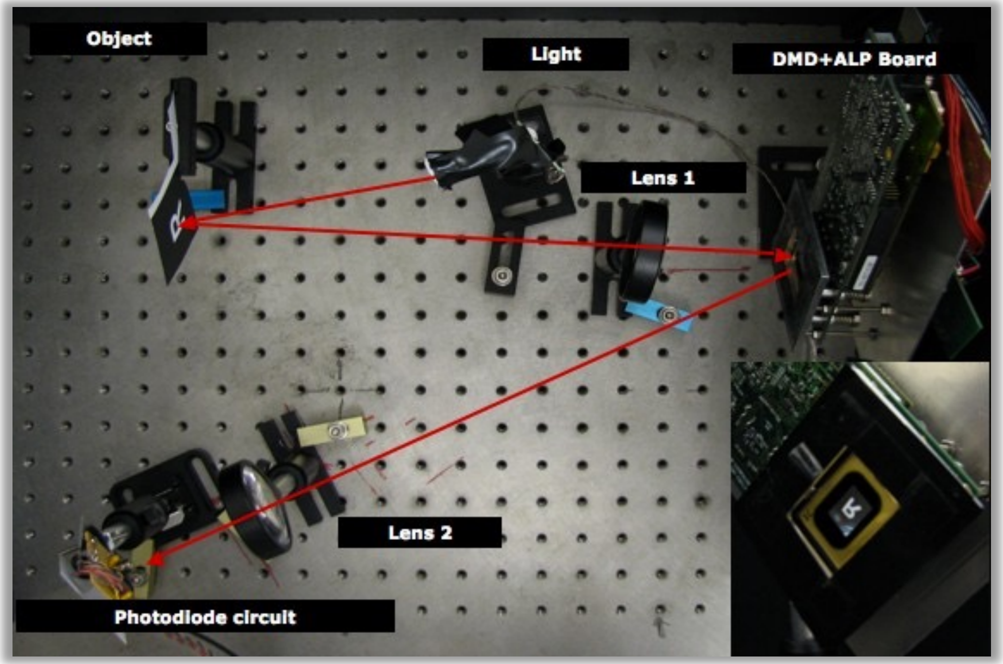
Must read: course notes on Solving Regularized Inverse Problems with ADMM!

# Single-pixel Imaging




Duarte et al. 2008

# Single-pixel Imaging




Duarte et al. 2008

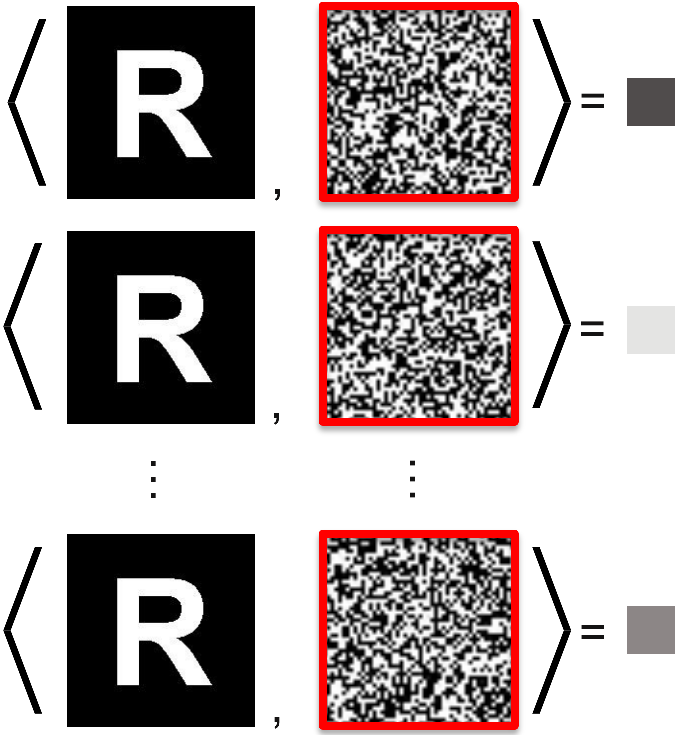
original

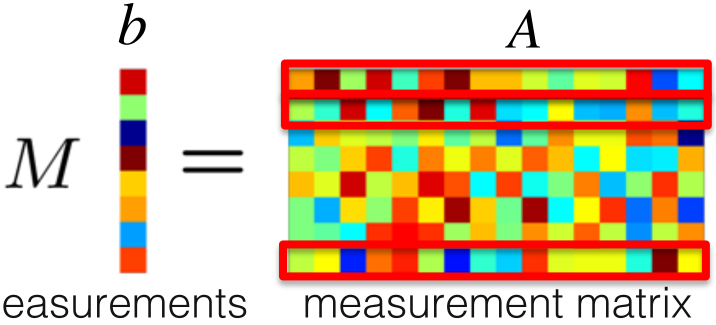
10%

5%

2%

# Single-pixel Imaging





# Under-determined Inverse Problems

- Image formation model: $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\eta}, \quad \boldsymbol{b} \in \mathbb{R}^M, \boldsymbol{x} \in \mathbb{R}^N, \boldsymbol{A} \in \mathbb{R}^{M \times N}$

- What makes it under-determined (or a compressive imaging problem): $M < N$

- Problem: infinitely many solutions satisfy the observations!

  Same problem as ill-posed problems! → need **image priors**

# Under-determined Inverse Problems

- Image formation model: $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\eta}$, $\quad \boldsymbol{b} \in \mathbb{R}^{M}, \boldsymbol{x} \in \mathbb{R}^{N}, \boldsymbol{A} \in \mathbb{R}^{M \times N}$

- Standard approach – the least-norm solution: $\tilde{\boldsymbol{x}}_{\mathrm{ln}} = \boldsymbol{A}^{\boldsymbol{T}}\left(\boldsymbol{A}\boldsymbol{A}^{\boldsymbol{T}}\right)\boldsymbol{b}$

- This is the solution of optimization problem $$\begin{aligned} &\text{minimize}_{x} \ \|\boldsymbol{x}\|_2 \\ &\text{subject to} \quad \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{aligned}$$

Note: among the infinitely many solutions satisfying the observations, the least-norm solution is the one with the smallest L2 norm, thus equivalent to $\|\cdot\|_2$ regularizer

# Under-determined Inverse Problems

- Image formation model: $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\eta}, \quad \boldsymbol{b} \in \mathbb{R}^M, \boldsymbol{x} \in \mathbb{R}^N, \boldsymbol{A} \in \mathbb{R}^{M \times N}$

- Standard approach – the least-norm solution: $\tilde{\boldsymbol{x}}_{\mathrm{ln}} = \boldsymbol{A}^T\left(\boldsymbol{A}\boldsymbol{A}^T\right)\boldsymbol{b}$

- Results (not great):

Compression Factor $^N/_M$

2x

PSNR 12.3

4x

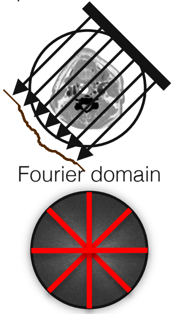
PSNR 10.4

8x

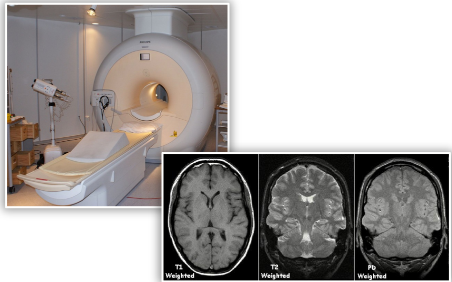
PSNR 9.7

# Other Inverse Problems in Imaging

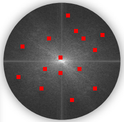
Images: Wikipedia



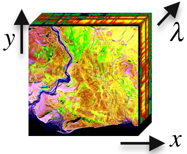

Computed Tomography (CT)

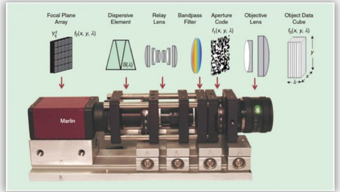



Fourier domain



Magnetic Resonance Imaging (MRI)




Hyperspectral Imaging

- Computational photography
- Light field imaging
- Thermal imaging
- …

# Other Inverse Problems in Imaging

- All these inverse problems have important applications and are very different
- Yet, they all boil down to the same inverse problem, each with a different matrix $\boldsymbol{A}$: $\quad \text{minimize}_{\boldsymbol{x}} \ \frac{1}{2}||\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}||_2^2 + \lambda\Psi(\boldsymbol{x})$
- The methods derived here also apply to all those problems and applications; single-pixel imaging is a great example problem → "if you can solve this, you can solve anything"

# Review of HQS for General Inverse Problems

- Objective or “loss” function of general inverse problem:

$$\text{minimize}_{\boldsymbol{x}} \ \frac{1}{2}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2^2 + \underbrace{\lambda}_{\text{weight of regularizer}}\Psi(\boldsymbol{D}\boldsymbol{x})$$

- Reformulate as:

$$\text{minimize}_{\{x,z\}} \ \underbrace{\frac{1}{2}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2^2}_{f(\boldsymbol{x})} + \underbrace{\lambda\Psi(\boldsymbol{z})}_{g(\boldsymbol{z})}$$

$$\text{subject to} \ \ \boldsymbol{D}\boldsymbol{x}-\boldsymbol{z}=0$$

- Remove constraints using penalty term (equivalent for large $\rho$):

$$L_\rho(\boldsymbol{x},\boldsymbol{z}) = f(\boldsymbol{x}) + g(\boldsymbol{z}) + \underbrace{\frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x}-\boldsymbol{z}\|_2^2}_{\text{penalty term}}$$

# Review of HQS for General Inverse Problems

$$L_\rho(\boldsymbol{x},\boldsymbol{z}) = f(\boldsymbol{x}) + g(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x} - \boldsymbol{z}\|_2^2$$

- Alternating gradient descent approach to solving penalty formulation leads to following iterative algorithm:

```
while not converged:
```

$$\boldsymbol{x} \leftarrow \mathrm{prox}_{f,\rho}(\boldsymbol{z}) = \arg\min_{\boldsymbol{x}} L_\rho(\boldsymbol{x},\boldsymbol{z}) = \arg\min_{\boldsymbol{x}} f(\boldsymbol{x}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x} - \boldsymbol{z}\|_2^2$$

$$\boldsymbol{z} \leftarrow \mathrm{prox}_{g,\rho}(\boldsymbol{D}\boldsymbol{x}) = \arg\min_{\boldsymbol{z}} L_\rho(\boldsymbol{x},\boldsymbol{z}) = \arg\min_{\boldsymbol{z}} g(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x} - \boldsymbol{z}\|_2^2$$

# Review of HQS for General Inverse Problems

$$L_\rho(\boldsymbol{x},\boldsymbol{z}) = f(\boldsymbol{x}) + g(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x}-\boldsymbol{z}\|_2^2$$

$$L_\rho(\boldsymbol{x},\boldsymbol{z}) = \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|_2^2 + \lambda\Psi(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x}-\boldsymbol{z}\|_2^2$$

$\boldsymbol{x} \in \mathbb{R}^N$ unknown image

$\boldsymbol{A} \in \mathbb{R}^{M\times N}$ matrix describing image formation model

$\boldsymbol{z} \in \mathbb{R}^{2N}, \boldsymbol{D} = \begin{bmatrix}\boldsymbol{D}_x \\ \boldsymbol{D}_y\end{bmatrix} \in \mathbb{R}^{2N\times N}$ for TV regularizer

$\boldsymbol{z} \in \mathbb{R}^{N}, \boldsymbol{D} = \boldsymbol{I} \in \mathbb{R}^{N\times N}$ for denoising or other regularizers

# Review of HQS for General Inverse Problems

<u>$\boldsymbol{x}$ – update:</u>

$$\boldsymbol{x} \leftarrow \mathrm{prox}_{\|\cdot\|_2,\rho}(\boldsymbol{z}) = \arg\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{Ax}-\boldsymbol{b}\|_2^2 + \frac{\rho}{2}\|\boldsymbol{Dx}-\boldsymbol{z}\|_2^2$$

$$\boldsymbol{x} \leftarrow \underbrace{\left(\boldsymbol{A}^T\boldsymbol{A} + \rho\boldsymbol{D}^T\boldsymbol{D}\right)}_{\widetilde{\boldsymbol{A}}}{}^{-1}\underbrace{\left(\boldsymbol{A}^T\boldsymbol{b} + \rho\boldsymbol{D}^T\boldsymbol{z}\right)}_{\widetilde{\boldsymbol{b}}}$$

- For general inverse problems, we don't necessarily have an efficient closed-form solution for this problem, like we did for the deconvolution problem in lecture 10

- Use matrix-free iterative solver, such as the conjugate gradient method, to solve $\widetilde{\boldsymbol{A}}\boldsymbol{x} = \widetilde{\boldsymbol{b}}$ (e.g., `scipy.sparse.linalg.cg`)

# Review of HQS for General Inverse Problems

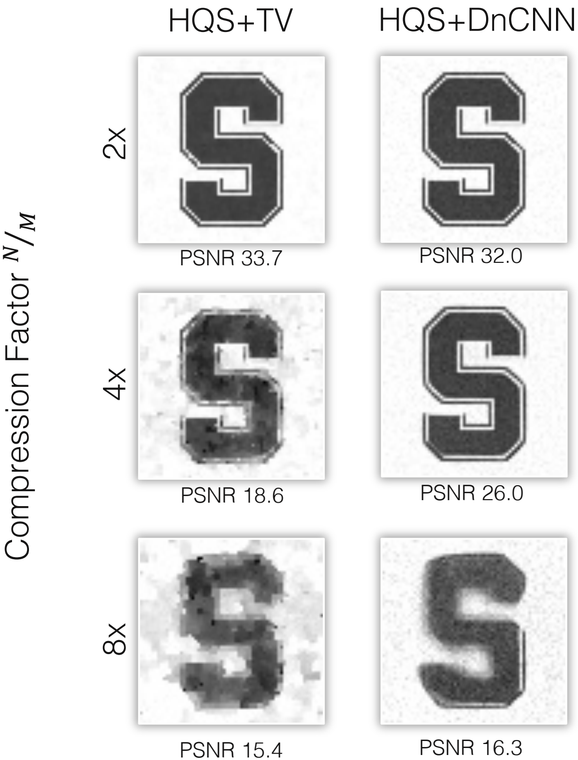
$\boldsymbol{z}$ – update for TV regularizer in closed form:

$$\boldsymbol{z} \leftarrow \operatorname{prox}_{\|\cdot\|_1,\rho}(\boldsymbol{D}\boldsymbol{x}) = \arg\min_{\boldsymbol{z}} \lambda \|\boldsymbol{z}\|_1 + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x} - \boldsymbol{z}\|_2^2 = \mathcal{S}_\kappa(\boldsymbol{v})$$

$\boldsymbol{z}$ – update for denoising-based regularizer in closed form:

$$\boldsymbol{z} \leftarrow \operatorname{prox}_{\mathcal{D},\rho}(\boldsymbol{x}) = \arg\min_{\boldsymbol{z}} \lambda \Psi(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{x} - \boldsymbol{z}\|_2^2 = \mathcal{D}\left(\boldsymbol{x}, \sigma^2 = \frac{\lambda}{\rho}\right)$$

# HQS for Single-pixel Imaging




- Works okay for low compression factor, i.e., when $M$ is close to $N$
- Not very robust for larger compression factors
- Formulation using penalty term is not adequate → need something more robust

# HQS vs. ADMM

- Objective function: $$\text{minimize}_{\boldsymbol{x}} \ \frac{1}{2}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2^2 + \lambda\Psi(\boldsymbol{D}\boldsymbol{x})$$

- Reformulate as: $$\text{minimize}_{\{x,z\}} \ \underbrace{\frac{1}{2}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2^2}_{f(\boldsymbol{x})} + \underbrace{\lambda\Psi(\boldsymbol{z})}_{g(\boldsymbol{z})}$$
$$\text{subject to } \ \boldsymbol{D}\boldsymbol{x}-\boldsymbol{z}=0$$

- Penalty Method of HQS: $$L_\rho^{(\mathrm{HQS})}(\boldsymbol{x},\boldsymbol{z}) = f(\boldsymbol{x}) + g(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x}-\boldsymbol{z}\|_2^2$$

- Augmented Lagrangian: $$L_\rho^{(\mathrm{ADMM})}(\boldsymbol{x},\boldsymbol{z},\boldsymbol{y}) = f(\boldsymbol{x}) + g(\boldsymbol{z}) + \boldsymbol{y}^T(\boldsymbol{D}\boldsymbol{x}-\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x}-\boldsymbol{z}\|_2^2$$
$$\overset{\boldsymbol{u}=(1/\rho)\boldsymbol{y}}{=} f(\boldsymbol{x}) + g(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x}-\boldsymbol{z}+\boldsymbol{u}\|_2^2 - \frac{\rho}{2}\|\boldsymbol{u}\|_2^2$$

# ADMM

$$L_\rho^{(\mathrm{ADMM})}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{u}) = f(\boldsymbol{x}) + g(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x} - \boldsymbol{z} + \boldsymbol{u}\|_2^2 - \frac{\rho}{2}\|\boldsymbol{u}\|_2^2$$

$\boldsymbol{x} \in \mathbb{R}^N$ unknown image

$\boldsymbol{A} \in \mathbb{R}^{M \times N}$ matrix describing image formation model

$\boldsymbol{z}, \boldsymbol{u} \in \mathbb{R}^{2N}, \boldsymbol{D} = \begin{bmatrix} \boldsymbol{D}_x \\ \boldsymbol{D}_y \end{bmatrix} \in \mathbb{R}^{2N \times N}$ for TV regularizer

$\boldsymbol{z}, \boldsymbol{u} \in \mathbb{R}^N, \boldsymbol{D} = \boldsymbol{I} \in \mathbb{R}^{N \times N}$ for denoising or other regularizers

# ADMM

$$L_\rho^{(\mathrm{ADMM})}(\boldsymbol{x},\boldsymbol{z},\boldsymbol{u}) = f(\boldsymbol{x}) + g(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x} - \boldsymbol{z} + \boldsymbol{u}\|_2^2 - \frac{\rho}{2}\|\boldsymbol{u}\|_2^2$$

- Alternating gradient descent approach to solving Augmented Lagrangian:

```
while not converged:
```

$$\boldsymbol{x} \leftarrow \mathrm{prox}_{f,\rho}(\boldsymbol{z}) = \arg\min_x L_\rho^{(\mathrm{ADMM})}(\boldsymbol{x},\boldsymbol{z},\boldsymbol{u}) = \arg\min_{\boldsymbol{x}} f(\boldsymbol{x}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x} - \boldsymbol{z} + \boldsymbol{u}\|_2^2$$

$$\boldsymbol{z} \leftarrow \mathrm{prox}_{g,\rho}(\boldsymbol{D}\boldsymbol{x}) = \arg\min_z L_\rho^{(\mathrm{ADMM})}(\boldsymbol{x},\boldsymbol{z},\boldsymbol{u}) = \arg\min_{\boldsymbol{z}} g(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x} - \boldsymbol{z} + \boldsymbol{u}\|_2^2$$

$$\boldsymbol{u} \leftarrow \boldsymbol{u} + \boldsymbol{D}\boldsymbol{x} - \boldsymbol{z}$$

# ADMM

$\boldsymbol{x}$ – update:

$$\boldsymbol{x} \leftarrow \mathrm{prox}_{\|\cdot\|_2,\rho}(\boldsymbol{z}) = \arg\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 + \frac{\rho}{2}\|\boldsymbol{D}\boldsymbol{x} - \boldsymbol{z} + \boldsymbol{u}\|_2^2,$$

$$\boldsymbol{x} \leftarrow \underbrace{\left(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{A} + \rho\boldsymbol{D}^{\boldsymbol{T}}\boldsymbol{D}\right)}_{\widetilde{\boldsymbol{A}}}{}^{-1}\underbrace{\left(\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{b} + \rho\boldsymbol{D}^{\boldsymbol{T}}(\boldsymbol{z} - \boldsymbol{u})\right)}_{\widetilde{\boldsymbol{b}}}$$

- Same general x-update as HQS, use matrix-free iterative solver, such as the conjugate gradient method, to solve $\widetilde{\boldsymbol{A}}\boldsymbol{x} = \widetilde{\boldsymbol{b}}$ (e.g., `scipy.sparse.linalg.cg`)

# ADMM

**$\boldsymbol{z}$ – update for TV regularizer in closed form:**

$$\boldsymbol{z} \leftarrow \mathrm{prox}_{\|\cdot\|_1,\rho}(\boldsymbol{v}) = \arg\min_z \lambda\|\boldsymbol{z}\|_1 + \frac{\rho}{2}\|\boldsymbol{v}-\boldsymbol{z}\|_2^2 = \mathcal{S}_\kappa(\boldsymbol{v}),\ \boldsymbol{v} = \boldsymbol{D}\boldsymbol{x} + \boldsymbol{u}$$

**$\boldsymbol{z}$ – update for denoising-based regularizer in closed form:**

$$\boldsymbol{z} \leftarrow \mathrm{prox}_{\mathcal{D},\rho}(\boldsymbol{x}+\boldsymbol{u}) = \arg\min_z \lambda\Psi(\boldsymbol{z}) + \frac{\rho}{2}\|\boldsymbol{x}-\boldsymbol{z}+\boldsymbol{u}\|_2^2 = \mathcal{D}\left(\boldsymbol{x}+\boldsymbol{u}, \sigma^2 = \frac{\lambda}{\rho}\right)$$

→ Same z-update rules as HQS!

# ADMM

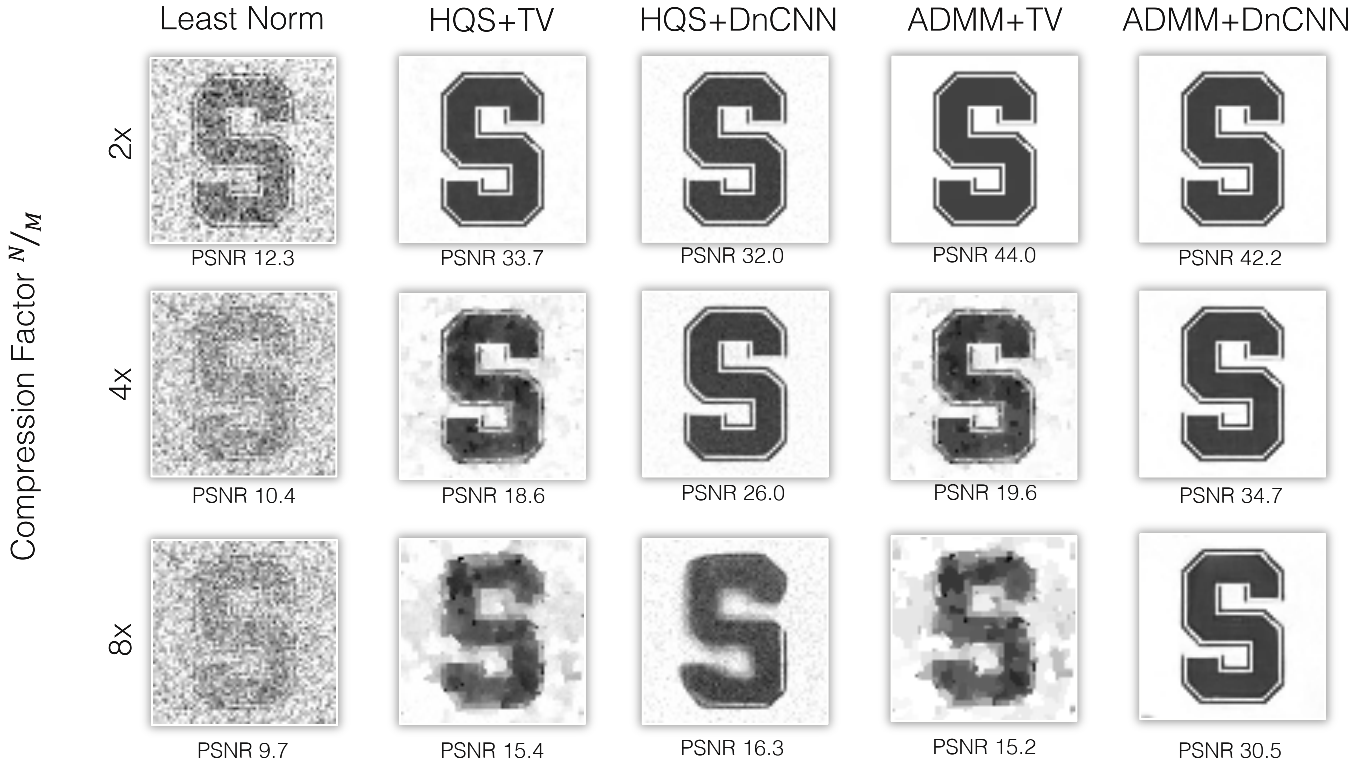
## ADMM for inverse problem with denoiser

initialize $\rho$ and $\lambda$
$\mathbf{x} = zeros\,(W, H)\,;$
$\mathbf{z} = zeros\,(W, H)\,;$
$\mathbf{u} = zeros\,(W, H)\,;$
**for** $k = 1$ **to** $max\_iters$ **do**
  $\mathbf{x} = \mathbf{prox}_{\|\cdot\|_2,\rho}\,(\mathbf{v}) = \text{cg\_solve}\left(\mathbf{A}^T\mathbf{A} + \rho\mathbf{I}, \mathbf{A}^T\mathbf{b} + \rho\,(\mathbf{z} - \mathbf{u})\right)$
  $\mathbf{prox}_{\mathcal{D},\rho}\,(\mathbf{x} + \mathbf{u}) = \mathcal{D}\left(\mathbf{x} + \mathbf{u}, \sigma^2 = \frac{\lambda}{\rho}\right)$
  $\mathbf{u} = \mathbf{u} + \mathbf{x} - \mathbf{z}$
**end for**

## ADMM for inverse problem with TV

initialize $\rho$ and $\lambda$
$\mathbf{x} = zeros\,(W, H)\,;$
$\mathbf{z} = zeros\,(W, H, 2)\,;$
$\mathbf{u} = zeros\,(W, H, 2)\,;$
**for** $k = 1$ **to** $max\_iters$ **do**
  $\mathbf{x} = \mathbf{prox}_{\|\cdot\|_2,\rho}\,(\mathbf{z} - \mathbf{u}) = \text{cg\_solve}\left(\mathbf{A}^T\mathbf{A} + \rho\mathbf{D}^T\mathbf{D}, \mathbf{A}^T\mathbf{b} + \rho\mathbf{D}^T\,(\mathbf{z} - \mathbf{u})\right)$
  $\mathbf{z} = \mathbf{prox}_{\|\cdot\|_1,\rho}\,(\mathbf{Dx} + \mathbf{u}) = \mathcal{S}_{\lambda/\rho}\,(\mathbf{Dx} + \mathbf{u})$
  $\mathbf{u} = \mathbf{u} + \mathbf{Dx} - \mathbf{z}$
**end for**

# ADMM – Results



# Back to the Bayesian Perspective of Inverse Problems

Note: the following material is optional and not part of any homework or the midterm!

# Bayesian Perspective of Gaussian Noise

- Image formation model: $\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\eta}, \quad \boldsymbol{b} \in \mathbb{R}^M, \boldsymbol{x} \in \mathbb{R}^N, \boldsymbol{A} \in \mathbb{R}^{M \times N}$

- Joint probability of all observations: $p(\boldsymbol{b}|\boldsymbol{x},\sigma) = \prod_{i=1}^{M} p(\boldsymbol{b}_i|\boldsymbol{x}_i,\sigma) \propto e^{-\frac{\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2^2}{2\sigma^2}}$

- Bayes' rule: $p(\boldsymbol{x}|\boldsymbol{b},\sigma) = \frac{p(\boldsymbol{b}|\boldsymbol{x},\sigma)p(\boldsymbol{x})}{p(\boldsymbol{b})} \propto p(\boldsymbol{b}|\boldsymbol{x},\sigma)\boldsymbol{p}(\boldsymbol{x})$

- Maximum-a-posterior (MAP) solution:

$$\begin{aligned}\boldsymbol{x}_{MAP} &= \arg\min_{\boldsymbol{x}} \ -\log\big(p(\boldsymbol{x}|\boldsymbol{b},\sigma)\big) \\ &= \arg\min_{\boldsymbol{x}} \ \frac{1}{2\sigma^2}\|\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}\|_2^2 + \Psi(\boldsymbol{x})\end{aligned}$$

# Bayesian Perspective of Poisson Noise

- Image formation model: $\boldsymbol{b} = \mathcal{P}(\boldsymbol{Ax})$, $\quad \boldsymbol{b} \in \mathbb{R}^M, \boldsymbol{x} \in \mathbb{R}^N, \boldsymbol{A} \in \mathbb{R}^{M \times N}$

- Probability of observation $i$:
$$p(\boldsymbol{b}_i|\boldsymbol{x}) = \frac{(\boldsymbol{Ax})_i^{\boldsymbol{b}_i}\, e^{-(\boldsymbol{Ax})_i}}{\boldsymbol{b}_i!}$$

- Joint probability of all observations:
$$p(\boldsymbol{b}|\boldsymbol{x}) = \prod_{i=1}^{M} p(\boldsymbol{b}_i|\boldsymbol{x})$$
$$= \prod_{i=1}^{M} e^{\log((\boldsymbol{Ax})_i)\boldsymbol{b}_i} \cdot e^{-(\boldsymbol{Ax})_i} \cdot \frac{1}{\boldsymbol{b}_i!}$$

# Bayesian Perspective of Poisson Noise

- Image formation model: $\boldsymbol{b} = \mathcal{P}(\boldsymbol{Ax})$, $\quad \boldsymbol{b} \in \mathbb{R}^M, \boldsymbol{x} \in \mathbb{R}^N, \boldsymbol{A} \in \mathbb{R}^{M \times N}$

- Bayes' rule: $$p(\boldsymbol{x}|\boldsymbol{b},\sigma) = \frac{p(\boldsymbol{b}|\boldsymbol{x},\sigma)p(\boldsymbol{x})}{p(\boldsymbol{b})} \propto p(\boldsymbol{b}|\boldsymbol{x},\sigma)p(\boldsymbol{x})$$

- Maximum-a-posterior (MAP) solution:

$$\boldsymbol{x}_{MAP} = \arg\min_x \; -\log\big(p(\boldsymbol{x}|\boldsymbol{b},\sigma)\big) = -\log\big(p(\boldsymbol{b}|\boldsymbol{x})\big) - \log\big(p(\boldsymbol{x})\big)$$

$$= \arg\min_x \; -\log\big(p(\boldsymbol{b}|\boldsymbol{x})\big) + \lambda\Psi(\boldsymbol{x})$$

# ADMM+TV for Poisson Noise & Nonnegativity

- Objective function: $$\text{minimize}_x \; -\log\big(\underset{\uparrow \text{ includes } \boldsymbol{A}}{p(\boldsymbol{b}|\boldsymbol{x})}\big) + \lambda\Psi(\boldsymbol{D}\boldsymbol{x})$$

- Reformulate as: $$\text{minimize}_{\{x,z\}} \; \underbrace{-\log\big(\overset{\text{does not include } \boldsymbol{A} \downarrow}{p(\boldsymbol{b}|\boldsymbol{z}_1)}\big)}_{g_1(\boldsymbol{z}_1)} + \underbrace{\lambda_1\|\boldsymbol{z}_2\|_1}_{g_2(\boldsymbol{z}_2)} + \underbrace{\mathcal{I}_{\mathbb{R}_+}(\boldsymbol{z}_3)}_{g_3(\boldsymbol{z}_3)}$$

- Indicator function:

$$\mathcal{I}_{\mathbb{R}_+}(v) = \begin{cases} 0 & v > 0 \\ \infty & \text{otherwise} \end{cases}$$

$$\text{subject to} \quad \underbrace{\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{D} \\ \boldsymbol{I} \end{bmatrix}}_{\boldsymbol{K}} \boldsymbol{x} - \underbrace{\begin{bmatrix} \boldsymbol{z}_1 \\ \boldsymbol{z}_2 \\ \boldsymbol{z}_3 \end{bmatrix}}_{\boldsymbol{z}} = 0$$

- Scaled Augmented Lagrangian:

$$L_\rho^{(\text{ADMM})}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{u}) = \sum_i g_i(\boldsymbol{z}_i) + \frac{\rho}{2}\|\boldsymbol{K}\boldsymbol{x} - \boldsymbol{z} + \boldsymbol{u}\|_2^2 - \frac{\rho}{2}\|\boldsymbol{u}\|_2^2$$

# ADMM+TV for Poisson Noise & Nonnegativity

$$L_\rho^{(\mathrm{ADMM})}(\boldsymbol{x},\boldsymbol{z},\boldsymbol{u}) = \sum_i g_i(\boldsymbol{z}_i) + \frac{\rho}{2}\|\boldsymbol{K}\boldsymbol{x}-\boldsymbol{z}+\boldsymbol{u}\|_2^2 - \frac{\rho}{2}\|\boldsymbol{u}\|_2^2$$

- Alternating gradient descent approach to solving Augmented Lagrangian:

```
while not converged:
```

$$\boldsymbol{x} \leftarrow \mathrm{prox}_{\|\cdot\|_2,\rho}(\boldsymbol{z}) = \arg\min_x L_\rho^{(\mathrm{ADMM})}(\boldsymbol{x},\boldsymbol{z},\boldsymbol{u}) = \arg\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{K}\boldsymbol{x}-\boldsymbol{z}+\boldsymbol{u}\|_2^2$$

for all i:

$$\boldsymbol{z_i} \leftarrow \mathrm{prox}_{g_i,\rho}(\boldsymbol{x}) = \arg\min_{z_i} L_\rho^{(\mathrm{ADMM})}(\boldsymbol{x},\boldsymbol{z},\boldsymbol{u}) = \arg\min_{z_i} g_i(\boldsymbol{z}_i) + \frac{\rho}{2}\|\boldsymbol{K}\boldsymbol{x}-\boldsymbol{z}+\boldsymbol{u}\|_2^2$$

$$\boldsymbol{u} \leftarrow \boldsymbol{u}+\boldsymbol{K}\boldsymbol{x}-\boldsymbol{z}$$

# ADMM+TV for Poisson Noise & Nonnegativity

- Derivation of all these proximal operators in the course notes on Noise, Denoising, and Image Reconstruction with Noise!

```
while not converged:
```

$$\boldsymbol{x} \leftarrow \mathrm{prox}_{\|\cdot\|_2,\rho}(\boldsymbol{z}) = \arg\min_x L_\rho^{(\mathrm{ADMM})}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{u}) = \arg\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{K}\boldsymbol{x} - \boldsymbol{z} + \boldsymbol{u}\|_2^2$$

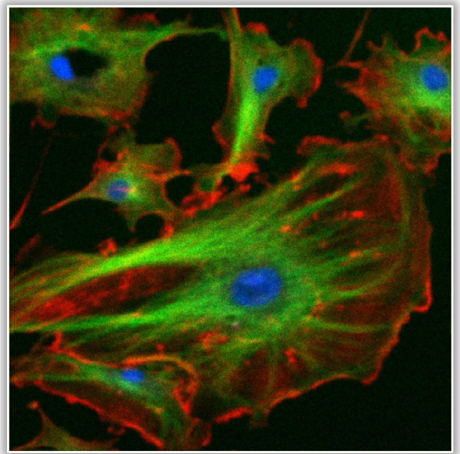
for all i:

$$\boldsymbol{z_i} \leftarrow \mathrm{prox}_{g_i,\rho}(\boldsymbol{x}) = \arg\min_{z_i} L_\rho^{(\mathrm{ADMM})}(\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{u}) = \arg\min_{z_i} g_i(\boldsymbol{z}_i) + \frac{\rho}{2}\|\boldsymbol{K}\boldsymbol{x} - \boldsymbol{z} + \boldsymbol{u}\|_2^2$$

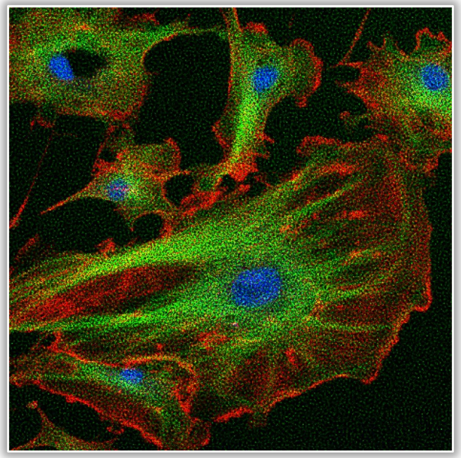
$$\boldsymbol{u} \leftarrow \boldsymbol{u} + \boldsymbol{K}\boldsymbol{x} - \boldsymbol{z}$$

# ADMM+TV for Poisson Noise & Nonnegativity

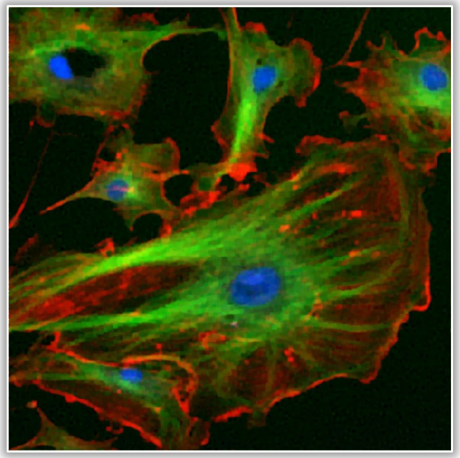
Blurry & Noisy Measurements



Richardson-Lucy Method
(maximum likelihood solution)



ADMM+TV+Nonnegativity
(maximum-a-posteriori solution)



# References and Further Reading

**Must read: EE367 course notes on Solving Regularized Inverse Problems with ADMM!**

**Optional read: EE367 course notes on Noise, Denoising, and Image Reconstruction with Noise**

ADMM

- S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein "Distributed optimization and statistical learning via the alternating direction method of multipliers", Foundation and Trends in Machine Learning, 2001

Single-pixel Imaging

- M. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, R. Baraniuk "Single-pixel imaging via compressive sampling", IEEE Signal Processing Magazine 2008