Reconstruction methods for under-sampled MR data
aka: Constrained reconstruction methods

Jeffrey A. Fessler

EECS Department, BME Department, Dept. of Radiology
University of Michigan

ISMRM Imaging Strategies Course
May 8, 2011

Acknowledgments: Doug Noll, Sathish Ramani

Disclosure

Declaration of Relevant Financial Interests or Relationships
Speaker Name: Jeff Fessler

I have the following relevant financial interest or relationship to disclose with regard to the subject matter of this presentation:

Company name: GE Healthcare and GE Global Research
Type of relationship: X-ray CT image reconstruction collaborations

I have no conflicts of interest with regards to MR topics.

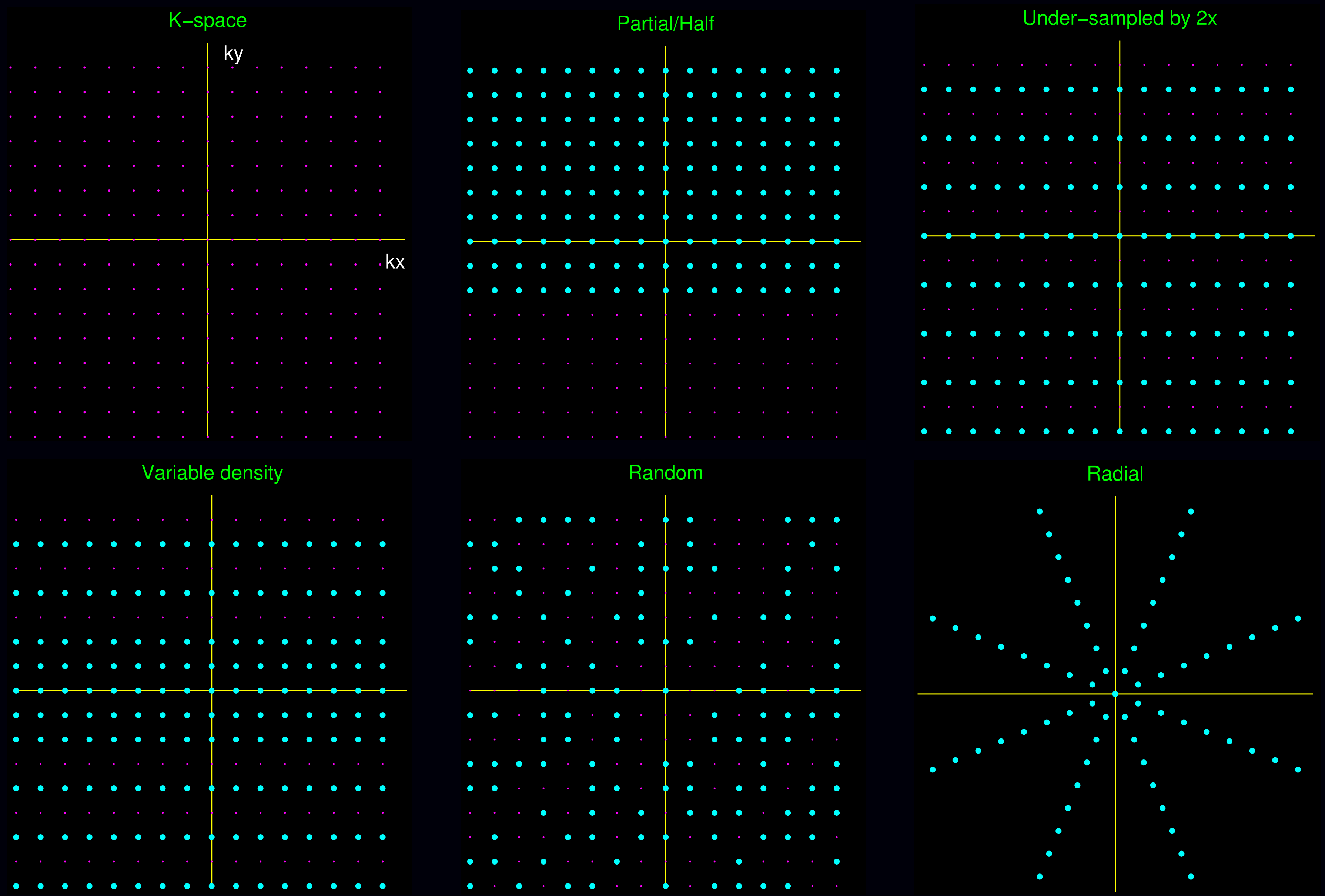
Introduction

Reasons for under-sampling:

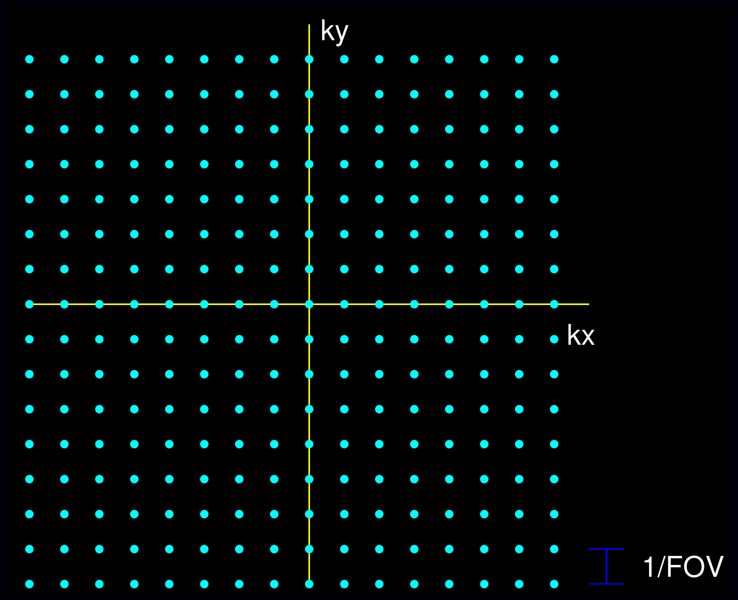
- Static imaging: reduce scan time
- Dynamic imaging: inherent
  - dynamic contrast studies (microscopic motion?)
  - bulk motion

All such situations require assumptions / constraints / models.

Under-Sampled K-space: Examples



Is This Under-Sampled K-space?

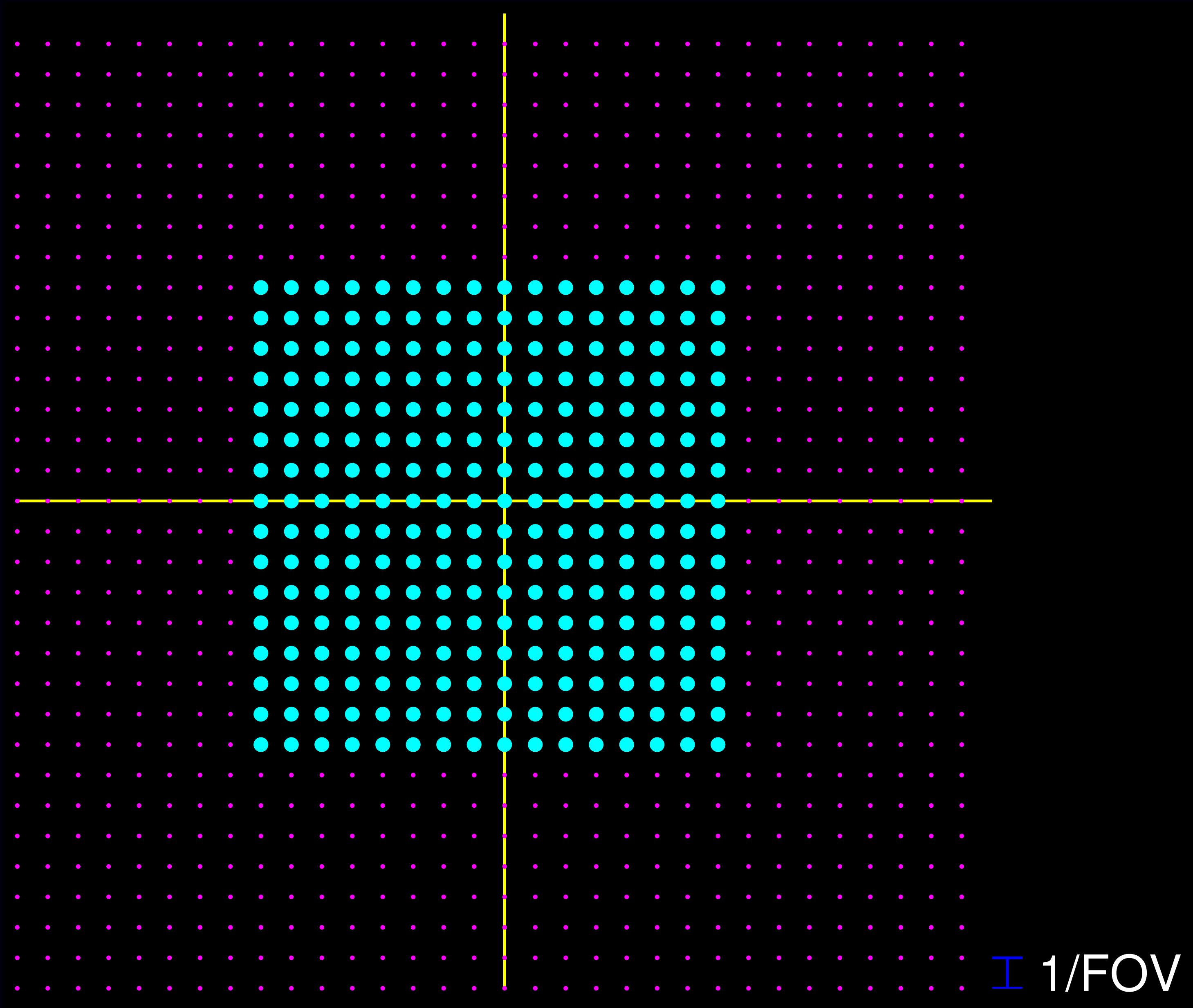


Note: k-space sample spacing is 1/FOV (Nyquist sample spacing).

Answers (audience response system):

1. No
2. Yes
3. Unsure
4. Will this be on the final exam?

Is This Under-Sampled K-space?



Basic MRI Signal Model

Ignoring many physical effects, the baseband signal in lth receive coil is approximately:

$$s_l(t) = \int f(\vec{r})\, c_l(\vec{r}) \exp\left(-\imath 2\pi \vec{k}(t) \cdot \vec{r}\right) \mathrm{d}\vec{r}. \tag{1}$$

- $\vec{r}$: spatial position
- $c_l(\vec{r})$: receive sensitivity of the lth coil, $l = 1, \ldots, L$
- $\vec{k}(t)$: k-space trajectory
- $f(\vec{r})$: (unknown) transverse magnetization of the object

MR scan data is noisy samples thereof:

$$y_{li} = s_l(t_i) + \varepsilon_{li}, \qquad i = 1, \ldots, M, \quad l = 1, \ldots, L \tag{2}$$

- y_{li}: ith sample of lth coil's signal
- $\boldsymbol{\varepsilon}_{li}$: additive complex white gaussian noise,
- M: number of k-space samples.

Goal: reconstruct object $f(\vec{r})$ from measurement vector $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_L)$, where $\boldsymbol{y}_l = (y_{l1}, \ldots, y_{l,M})$ is data from lth coil.

MR Image Reconstruction is Ill-Posed

$$y_{li} = \int f(\vec{r})\, c_l(\vec{r}) \exp\left(-\imath 2\pi \vec{k}(t_i) \cdot \vec{r}\right) \mathrm{d}\vec{r} + \varepsilon_{li}$$

- Unknown object $f(\vec{r})$ is a continuous space function
- Measurement vector $\boldsymbol{y}$ is finite dimensional

$\therefore$ All MRI data is under-sampled

Uncountably infinitely many objects $f(\vec{r})$ fit the data $\boldsymbol{y}$ exactly,
even for "fully sampled" data, even if there were no noise.

For "fully sampled" Cartesian k-space data,
how shall we choose one reconstructed image $\hat{f}(\vec{r})$ from among those?

1. Impose some assumptions / constraints / models
2. Just take an inverse FFT of the data
3. Both of the above
4. None of the above

Inverse FFT for MR Image Reconstruction

Using an inverse FFT for reconstruction from "fully sampled" single-coil data is equivalent to assuming the object lies in a finite-dimensional subspace:

$$f(\vec{r}) = f(x,y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f[n,m]\, b(x - n\triangle_{\mathrm{X}})\, b(y - n\triangle_{\mathrm{Y}})\,.$$

What choice of basis function $b(\cdot)$ is implicit in IFFT reconstruction?

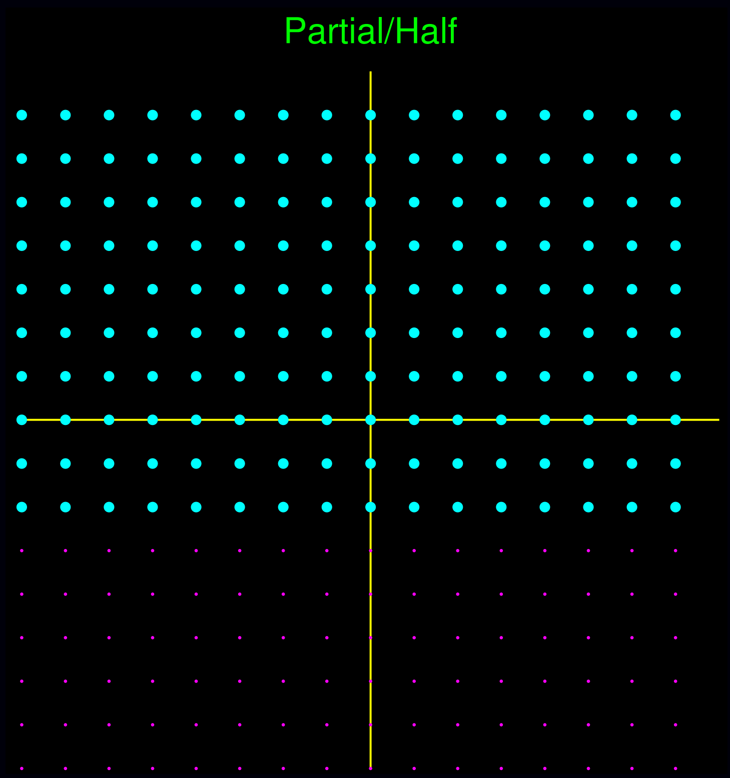
1. Dirac impulse
2. Rectangle (pixel)
3. Sinc
4. Dirichlet (periodic sinc)

$\therefore$ The use of assumptions / constraints / models is ubiquitous in MR.

In particular, constraining the estimate to lie in a finite-dimensional subspace is nearly ubiquitous.

(All models are wrong but some models are useful...)

Conventional Approach: Partial K-space



Conventional solution: Homodyning

Noll *et al.*, IEEE T-MI, June 1991

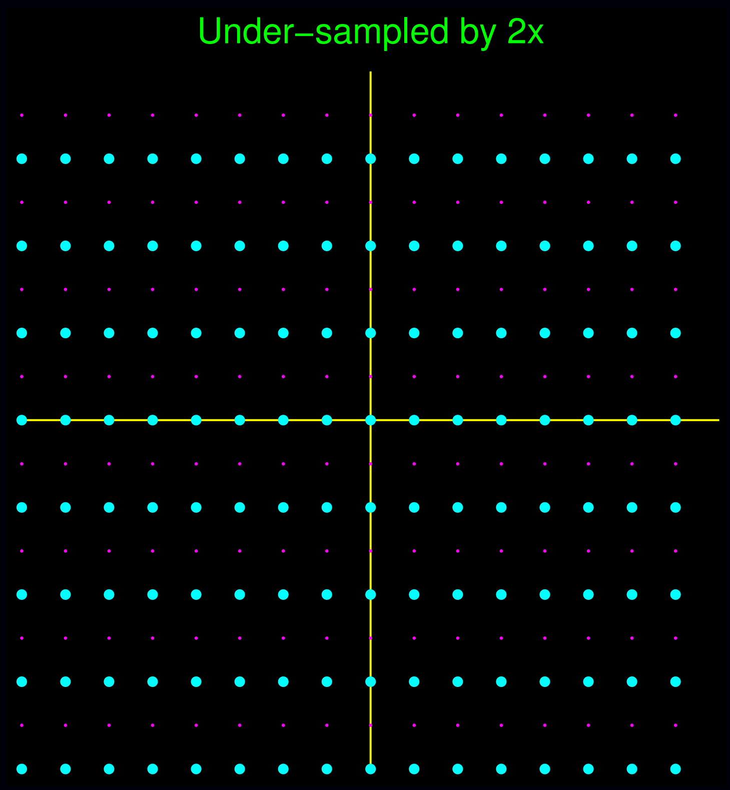
Constraint: object phase is smooth

Related iterative methods

Fessler & Noll, ISBI 2004

Bydder & Robson, MRM, June 2005

Conventional Approach: Decimation

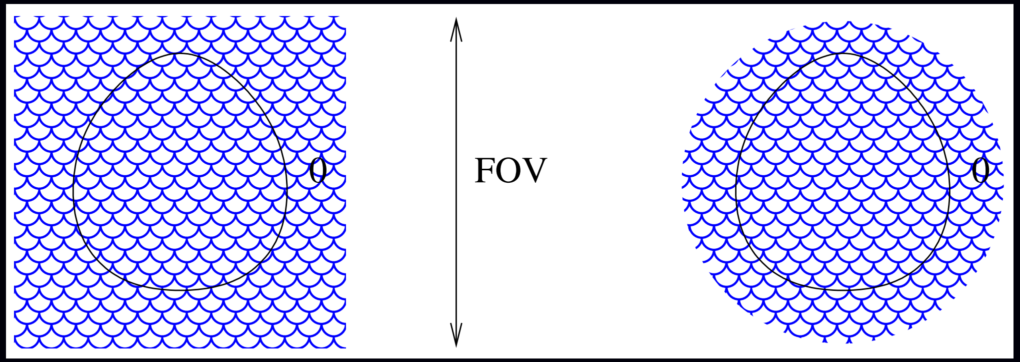


Conventional solutions: SENSE/GRAPPA (parallel imaging)

Pruessmann *et al.*, MRM, Nov. 1999

Griswold *et al.*, MRM, June 2002

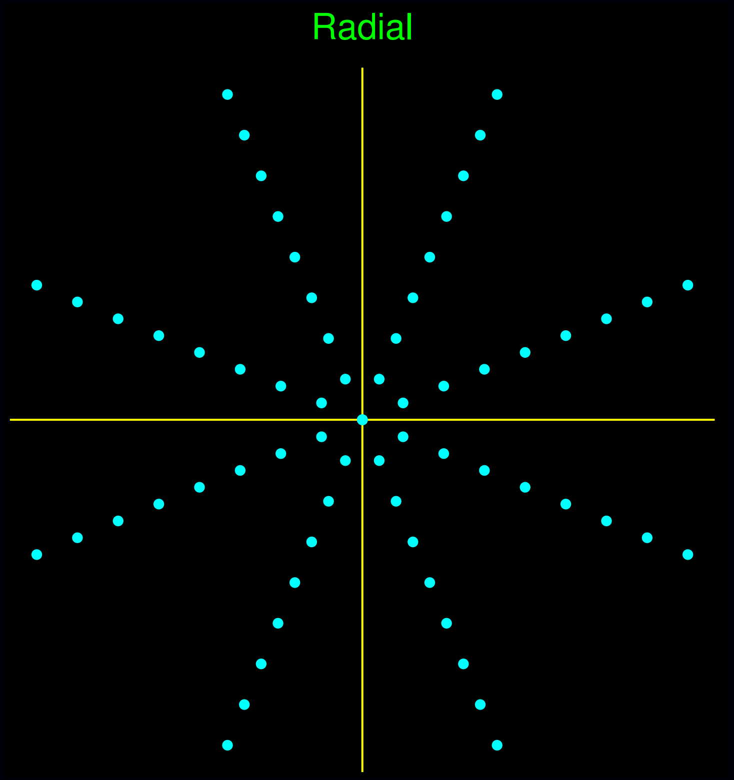
Constraint: object has finite support



Note: one can combine under-sampling strategies, *e.g.*, decimation and partial k-space

King & Angelos, ISMRM, 153. 2000

Conventional Approach: Non-Cartesian (Under) Sampling



Conventional solution: gridding

Jackson *et al.*, IEEE T-MI, Sep. 1991

Constraint??

1. object has finite support?
2. object has smooth phase?
3. object is band-limited?

Gridding alone is insufficient for “severely” under-sampled data.
For moderate amounts of under-sampling, consider non-Cartesian SENSE.

Pruessmann *et al.*, MRM, 2001

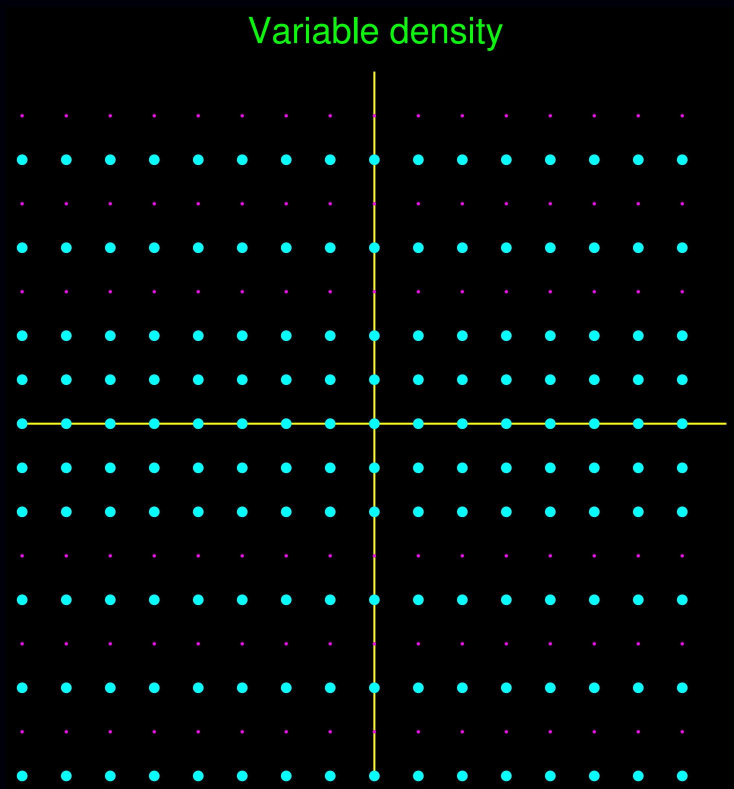
Or non-Cartesian GRAPPA.

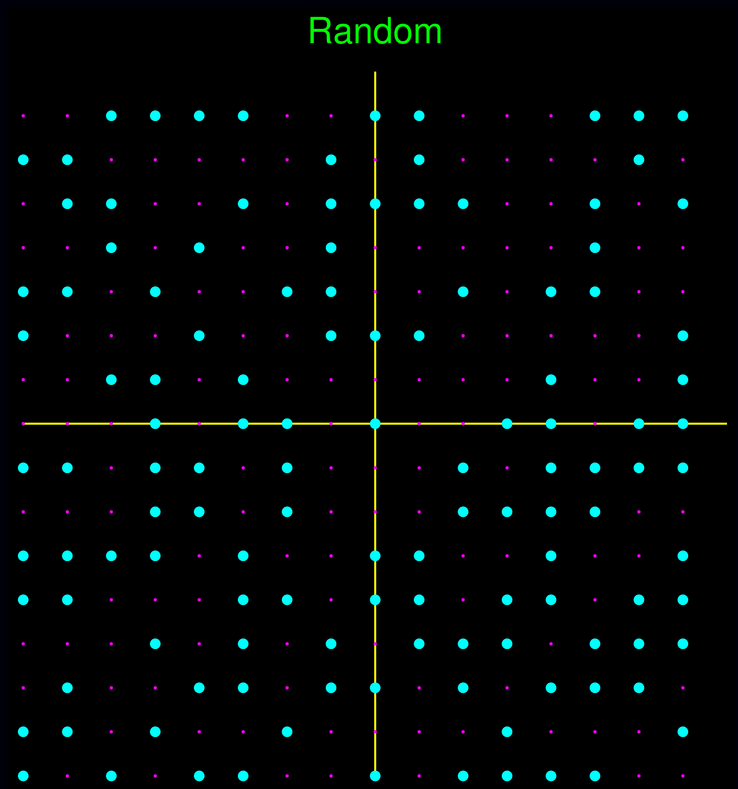
Seiberlich *et al.*, MRM, 2007

For “severely” under-sampled data, stronger constraints are needed.

Conventional Approach: Non-Cartesian Sampling

What about these sampling patterns?





Again, for “severely” under-sampled data, stronger constraints are needed.

Finite-dimensional subspace constraint

$$f(\vec{r}) = \sum_{j=1}^{N} x_j b_j(\vec{r}) = \sum_{j=1}^{N} x_j b(\vec{r} - \vec{r}_j) \tag{3}$$

- $b(\cdot)$: user-selected object basis function(s) (*e.g.*, rect function)
- $\vec{r}_j$: center of jth basis function translate
- N: number of parameters (*e.g.*, pixels)
- $\boldsymbol{x} = (x_1, \ldots, x_N)$: vector of unknown parameters (*e.g.*, pixel values).

Substituting this basis expansion into the signal model (1) yields:

$$\boldsymbol{y}_l = \boldsymbol{A}_l \boldsymbol{x} + \boldsymbol{\varepsilon}_l.$$

The elements $\{a_{lij}\}$ of the system matrix $\boldsymbol{A}_l$ for with the lth coil are:

$$a_{lij} = \int b(\vec{r} - \vec{r}_j)\, c_l(\vec{r})\, \mathrm{e}^{-\imath 2\pi \vec{k}(t_i) \cdot \vec{r}}\, \mathrm{d}\vec{r}, \tag{4}$$

Stacking up all L vectors and defining the $ML \times N$ matrix $\boldsymbol{A} = (\boldsymbol{A}_1, \ldots, \boldsymbol{A}_L)$ yields the “usual” linear model

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon}.$$

Parameterization alone is not enough

The finite-dimensional subspace constraint leads to the linear model

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{\varepsilon}.$$

- $\boldsymbol{y}$: measured data
- $\boldsymbol{A}$: known system model (k-space sampling and coil sensitivities)
- $\boldsymbol{x}$: unknown object
- $\boldsymbol{\varepsilon}$: additive noise

For severe under-sampling, $\boldsymbol{A}$ usually has fewer rows than columns.
In such under-determined situations, least-squares estimation is not unique:

$$\arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 = \{\boldsymbol{x} : \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}\}.$$

There may still be (uncountably) infinitely many solutions to $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$.

Need stronger assumptions / constraints to select a single estimate $\hat{\boldsymbol{x}}$.

Application-specific basis images

Instead of generic basis functions like rect, sinc, Dirac,
choose specialized basis functions $b_j(\vec{r})$, $j = 1, \ldots, N$, with N "small."

Constraint: $f(\vec{r})$ lies in a parsimonious subspace:

$$f(\vec{r}) = \sum_{j=1}^{N} x_j b_j(\vec{r}) \,.$$

If N is less than the number of k-space samples,
then the problem is over-determined and LS estimation is feasible:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\| = \left[\boldsymbol{A}'\boldsymbol{A}\right]^{-1} \boldsymbol{A}'\boldsymbol{y}.$$

Challenges

- choice of basis functions – preserving pathology?
- unstructured data-driven basis functions $\Longrightarrow$ lack of fast transforms

PCA of training data to design basis functions (and k-space samples)

Cao and Levin, MRM, Sep. 1993

Cao and Levin, IEEE T-MI, June 1995

Reference image?

In some applications, a "related" image $\boldsymbol{x}_0$ may be available.
Assume: unknown $\boldsymbol{x}$ is somehow "similar" to the reference image $\boldsymbol{x}_0$.

Regularization approach:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \underbrace{\|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2}_{\text{data fit}} + \beta \underbrace{\|\boldsymbol{L}(\boldsymbol{x} - \boldsymbol{x}_0)\|_p^p}_{\text{prior}},$$

where $\boldsymbol{L}$ is an optional weighting matrix.
For $p = 2$ and $\boldsymbol{L} = \boldsymbol{I}$:

$$\hat{\boldsymbol{x}} = \left[\boldsymbol{A}'\boldsymbol{A} + \beta\boldsymbol{I}\right]^{-1}\left(\boldsymbol{A}'\boldsymbol{y} + \beta\boldsymbol{x}_0\right) = \boldsymbol{x}_0 + \left[\boldsymbol{A}'\boldsymbol{A} + \beta\boldsymbol{I}\right]^{-1}\boldsymbol{A}'(\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}_0)$$

Constrained approach:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{x} - \boldsymbol{x}_0\|_p \text{ sub. to } \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}.$$

Sometimes equivalent to simply replacing the missing k-space samples with the spectrum of $\boldsymbol{x}_0$.

cf. "prior image constrained compressed sensing (PICCS)" approach
G H Chen *et al.*, Med. Phys., Feb. 2008. François *et al.*, ISMRM 3808, 2009.

Bayesian approach

Assume: $\boldsymbol{x}$ is a gaussian random field with mean $\boldsymbol{\mu}$, covariance matrix $\boldsymbol{K}$.

MAP / MMSE estimate:

$$\begin{aligned}\hat{\boldsymbol{x}} &= \arg\max_{\boldsymbol{x}} \mathrm{p}(\boldsymbol{x}\,|\,\boldsymbol{y}) = \arg\min_{\boldsymbol{x}} \frac{1}{2\sigma^2}\|\boldsymbol{y}-\boldsymbol{A}\boldsymbol{x}\|^2 + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{K}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) \\ &= \boldsymbol{\mu} + \left[\boldsymbol{A}'\boldsymbol{A} + \sigma^2\boldsymbol{K}^{-1}\right]^{-1}\boldsymbol{A}'(\boldsymbol{y}-\boldsymbol{A}\boldsymbol{\mu})\end{aligned}$$

Challenges

- Requires training data for $\boldsymbol{\mu}$ and $\boldsymbol{K}$.
- gaussian prior distribution is questionable
- Computation of $\hat{\boldsymbol{x}}$ if $\boldsymbol{K}$ is unstructured

Abandon training data and seek more “generic” constraints.

Sparsity / Compressibility

Start with the usual finite-dimensional subspace model:

$$f(\vec{r}) = \sum_{j=1}^{N} x_j b_j(\vec{r}) .$$

Usually generic image basis functions like rect or sinc are used here,
so x_j is just the jth pixel value.

Again, often N exceeds the number of measurements (under-determined).

Now constrain the coefficient vector $\boldsymbol{x} = (x_1, \ldots, x_N)$ somehow, as follows.

- 1. Synthesis approach
- 2. Analysis approach

Preliminaries:

- $\|\boldsymbol{x}\|_0 = \sum_k \mathbb{1}_{\{x_k \neq 0\}}$.
- $\|\boldsymbol{x}\|_1 = \sum_k |x_k|$.

We say $\boldsymbol{x}$ is "sparse" if $\|\boldsymbol{x}\|_0$ is "small."

We say $\boldsymbol{x}$ is "compressible" if $\|\boldsymbol{x}\|_1 \approx \|\tilde{\boldsymbol{x}}\|_1$,
where $\tilde{\boldsymbol{x}}$ retains only the "large" elements of $\boldsymbol{x}$.

1. Synthesis approach

For some $N \times K$ matrix $\boldsymbol{B}$ and K-dimensional coefficient vector $\boldsymbol{\theta}$:

$$\boldsymbol{x} = \boldsymbol{B}\boldsymbol{\theta}.$$

Usual choices:

- $K = N$ and $\boldsymbol{B}$ is an orthonormal basis (*e.g.*, wavelet synthesis)
- $K \gg N$ (!!) and $\boldsymbol{B}$ is an over-complete "dictionary"

Assume that the coefficient vector $\boldsymbol{\theta}$ is *sparse* or *compressible*.

Estimation strategies for synthesis approach use $\hat{\boldsymbol{x}} = \boldsymbol{B}\hat{\boldsymbol{\theta}}$ where:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_0 \text{ sub. to } \boldsymbol{y} = \boldsymbol{A}\boldsymbol{B}\boldsymbol{\theta} \quad \text{(sparsest possible)}$$

or

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \|\boldsymbol{\theta}\|_0 \text{ sub. to } \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{B}\boldsymbol{\theta}\| < \delta$$

or

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{B}\boldsymbol{\theta}\| \text{ sub. to } \|\boldsymbol{\theta}\|_0 \leq L$$

or

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{B}\boldsymbol{\theta}\|^2 + \beta \|\boldsymbol{\theta}\|_0 .$$

In all formulations, $\|\boldsymbol{\theta}\|_0$ often replaced by $\|\boldsymbol{\theta}\|_1$ to make problem convex.

2. Analysis approach

For some $K \times N$ analysis matrix $\boldsymbol{T}$, assume that the matrix-vector product

$$\boldsymbol{T}\boldsymbol{x}$$

is *sparse* or *compressible*.
Usual choices:

- $K = N$ and $\boldsymbol{T}$ is an orthonormal basis (*e.g.*, wavelet transform)
- $K \gg N$ and $\boldsymbol{T}$ includes wavelet transforms and finite-differences, aka total variation (TV), along various directions.

Estimation strategies for analysis approach:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{T}\boldsymbol{x}\|_0 \text{ sub. to } \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}$$

or

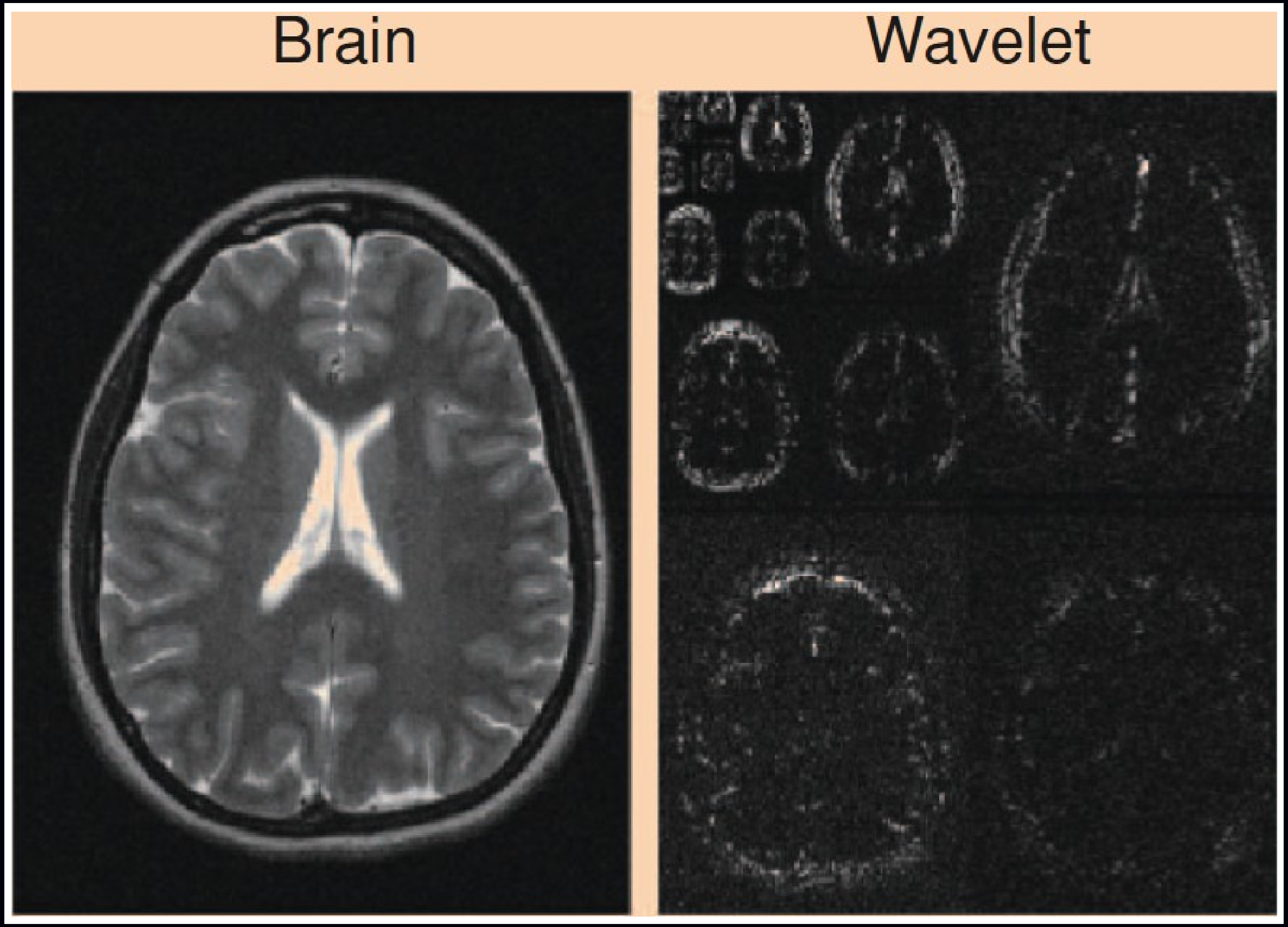
$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{T}\boldsymbol{x}\|_0 \text{ sub. to } \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\| < \delta$$

or

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|^2 + \beta \|\boldsymbol{T}\boldsymbol{x}\|_0 .$$

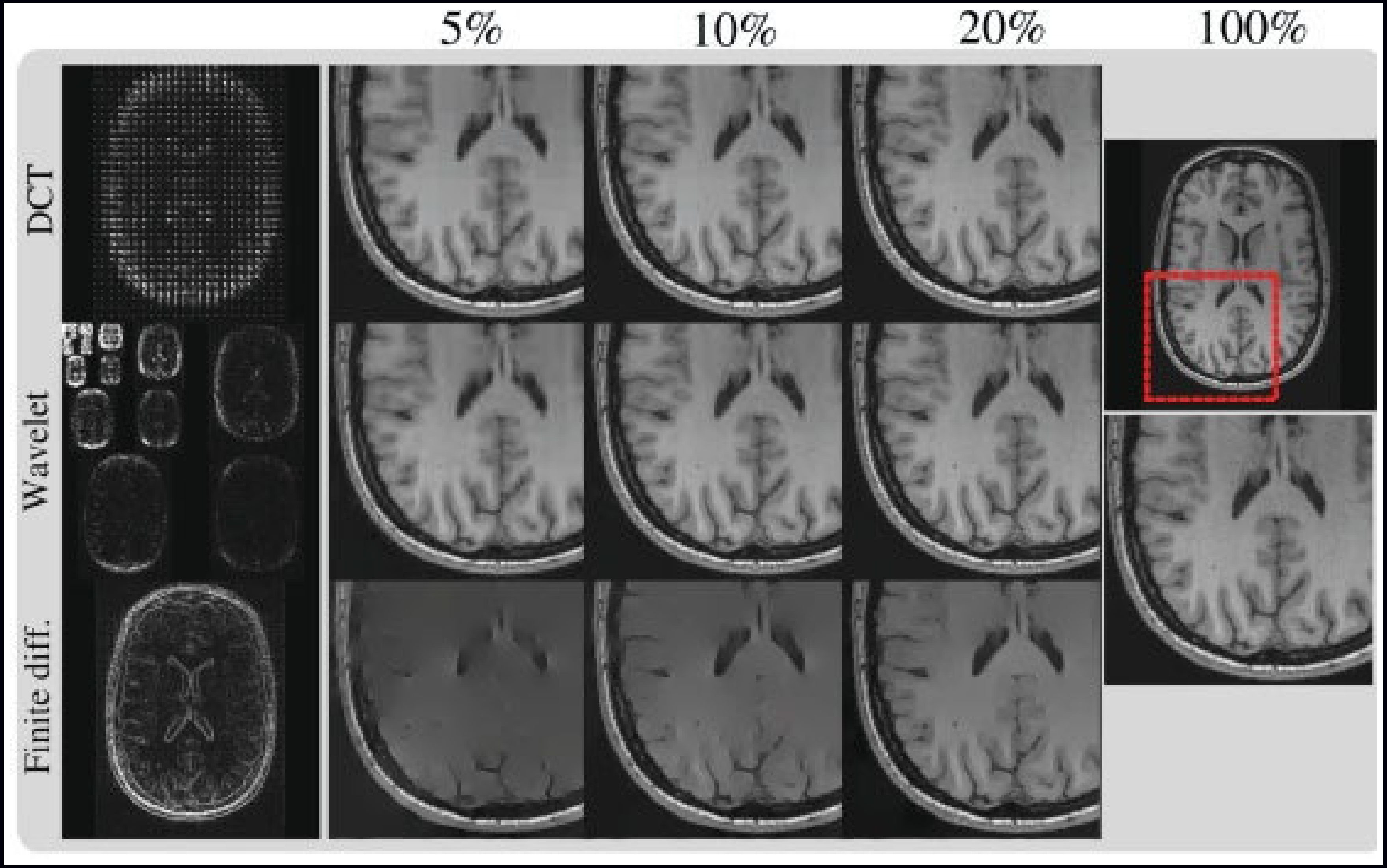
In all formulations, $\|\boldsymbol{\theta}\|_0$ often replaced by $\|\boldsymbol{\theta}\|_1$ to make problem convex.

Illustration of compressibility



Lustig *et al.*, IEEE Sig. Proc. Mag., Mar. 2008

Illustration of compressibility



Lustig *et al.*, MRM, Dec. 2007

Challenges with using sparsity constraints

- Choice of transform $\boldsymbol{T}$ or basis $\boldsymbol{B}$
  (These may affect image quality significantly.)
- Choice of β or δ
- Optimization algorithms / computation time

  ISMRM 2011 Poster 2873, Wed. 1:30. S. Ramani & J. Fessler: "An augmented Lagrangian method for regularized MRI reconstruction using SENSE"
  See also IEEE T-MI Mar. 2011: "Parallel MR image reconstruction using augmented Lagrangian methods"
- Cost function complexity
  - $\|x\|_0$ is non-convex ($\therefore$ local minimizers)
  - $\|x\|_1$ is non-differentiable, not strictly convex
    ($\therefore$ global minimizer may not be unique)
- Object phase variations

  ISMRM 2011 Poster 2841, Thu. 1:30. Zhao *et al.*, "Separate magnitude and phase regularization via compressed sensing"
- Efficacy of results depends on sampling pattern

Patch-wise sparsity

"Traditional" synthesis approach to sparsity uses bases for entire image:

$$\boldsymbol{x} = \boldsymbol{B}\boldsymbol{\theta}.$$

An alternative approach is to consider image a numerous small patches:

$$\boldsymbol{x} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_P\},$$

and represent each patch using a common basis or "dictionary" $\boldsymbol{D}$:

$$\boldsymbol{x}_p = \boldsymbol{D}\boldsymbol{\theta}_p, \; p = 1, \ldots, P$$

and assume that the coefficients $\boldsymbol{\theta}_p$ for each patch are sparse.

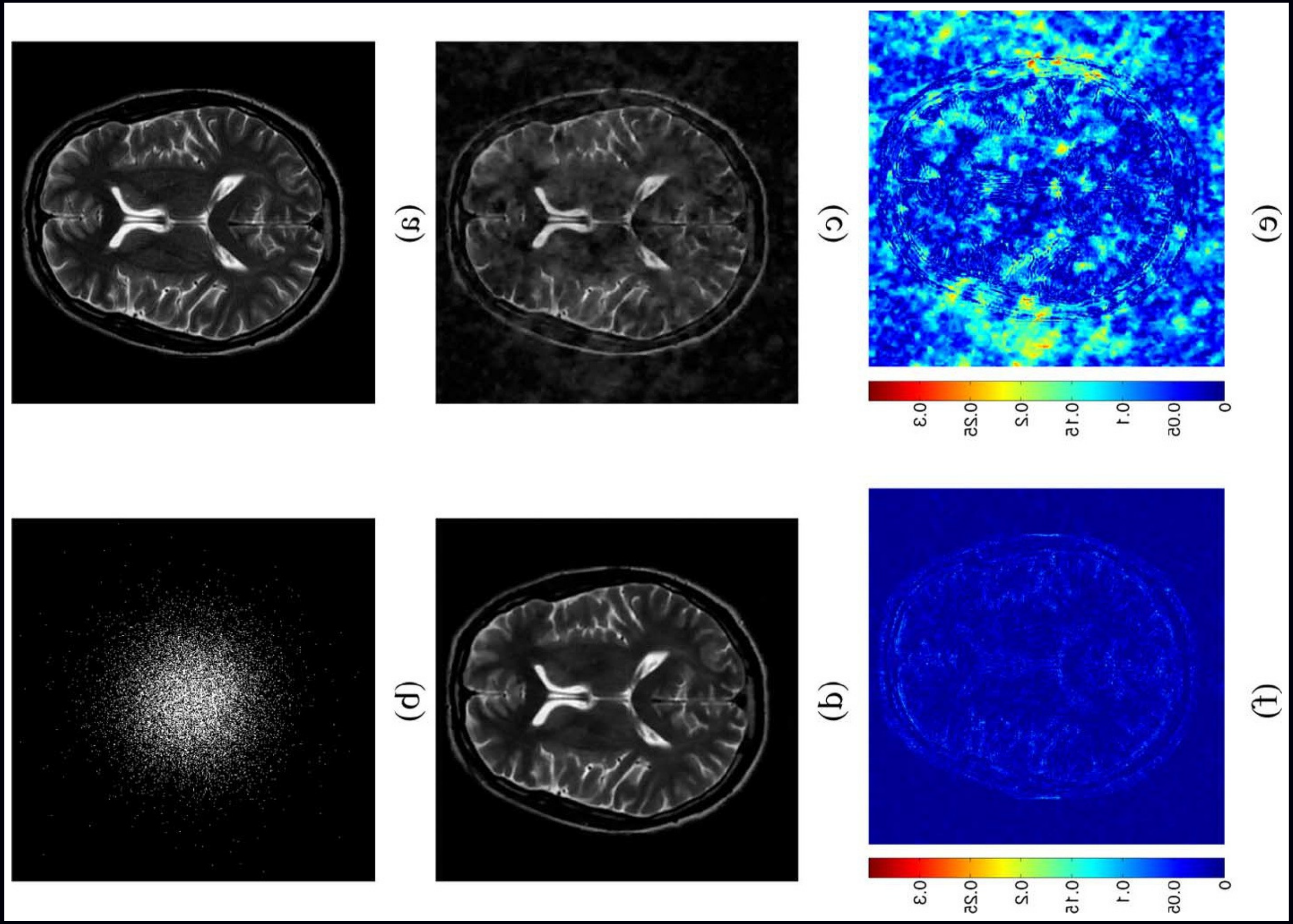
The patch basis or dictionary $\boldsymbol{D}$ can be

- predefined (*e.g.*, DCT)
- learned from training data
- estimated jointly during reconstruction ("adaptive" or "dictionary learning").

$$\arg\min_{\boldsymbol{x},\boldsymbol{\theta},\boldsymbol{D}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \beta_1 \sum_p \|\boldsymbol{x}_p - \boldsymbol{D}\boldsymbol{\theta}_p\|_2^2 + \beta_2 \sum_p \|\boldsymbol{\theta}_p\|_1$$

Ravishankar & Bresler, IEEE T-MI, May 2011 (and many references therein)

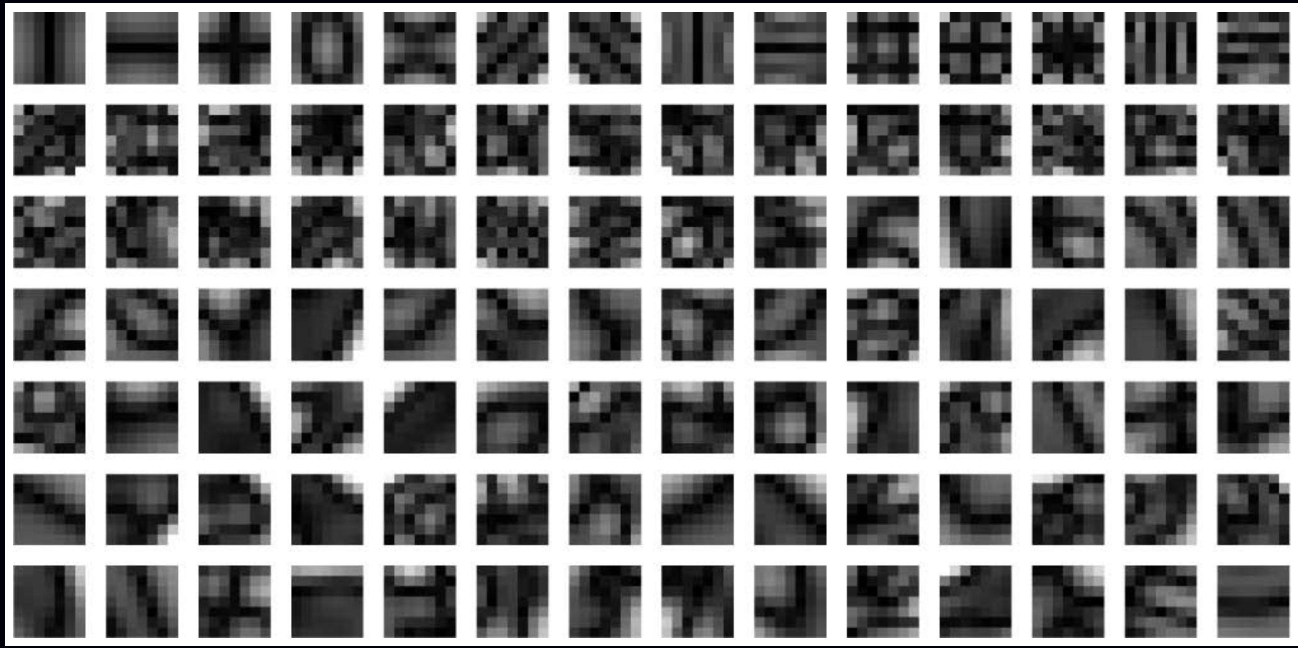
Adaptive Patch-Based Sparsity Example



(a) true. (b) sampling. (c) CS with wavelets/TV. (d) patch-based sparsity. (e),(f) errors.

20-fold under-sampling. Ravishankar & Bresler, IEEE T-MI, May 2011

Example of Adaptive Patch Dictionary



7×7 patches, *cf.* K-SVD
cf. K-SVD

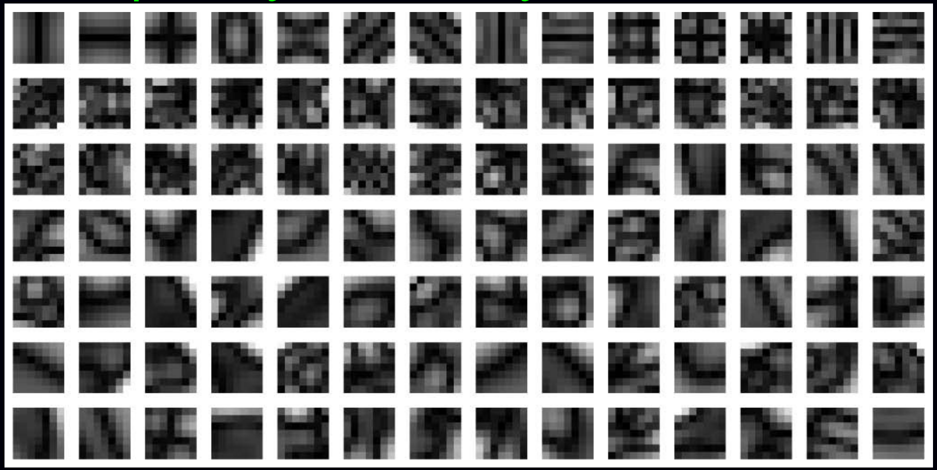
Ravishankar & Bresler, IEEE T-MI, May 2011
Aharon *et al.*, IEEE T-SP, Nov. 2006

How many basis patches would conventional SVD produce?

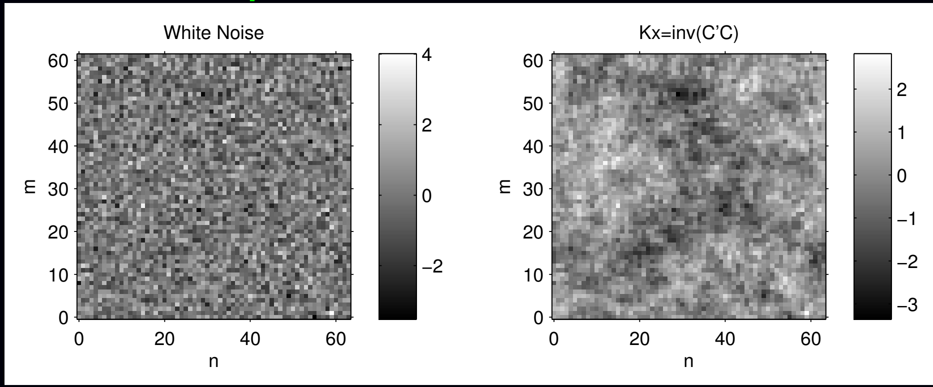
1. 7
2. 14
3. 49
4. 98

Comparison of Priors

Adaptive patch-based sparsity dictionary:



Gauss-Markov random field prior:



Challenges: non-convexity, choosing β_1 and β_2, ...

Nonlocal (patch-based) regularization

Based on success of nonlocal means (NLM) image denoising algorithm
Buades *et al.*, SIAM MMS 2005

Adapted to image reconstructions problems:

$$\arg\min_{\boldsymbol{x}} \|\boldsymbol{y} - \boldsymbol{A}\boldsymbol{x}\|_2^2 + \sum_{n,m} \psi(P_n(\boldsymbol{x}), P_m(\boldsymbol{x}))$$

- $P_n(\boldsymbol{x})$: nth patch of image
- $\psi(P_1, P_2)$: measures dissimilarity of two patches
- generalization of traditional edge-preserving regularization

Quite active research area:
Adluru *et al.*, JMRI, Nov. 2010
Manjón *et al.*, Med. Im. An., Dec. 2010
Yang & Jacob, ISBI 2011
Wang & Qi, ISBI 2011

Summary

Numerous possible assumptions / constraints / models
for image reconstruction from under-sampled k-space data:

- Using fewer basis functions
- Using reference image(s)
- Statistical priors
- Sparsity / compressibility
- Patch-based sparsity
- Patch-based regularization
- ...

Dynamic imaging

- same general principles apply
- even more variations / combinations possible
- space and/or time and/or Fourier transforms thereof
  - *e.g.*, dynamic cardiac imaging is pseudo-periodic
  - DCE MRI is amenable to small number of temporal basis components

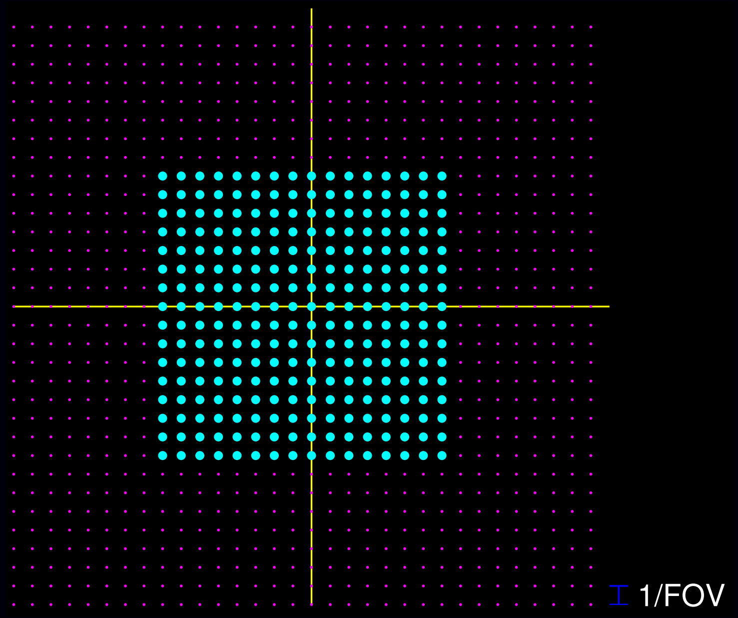
(innumerable references)

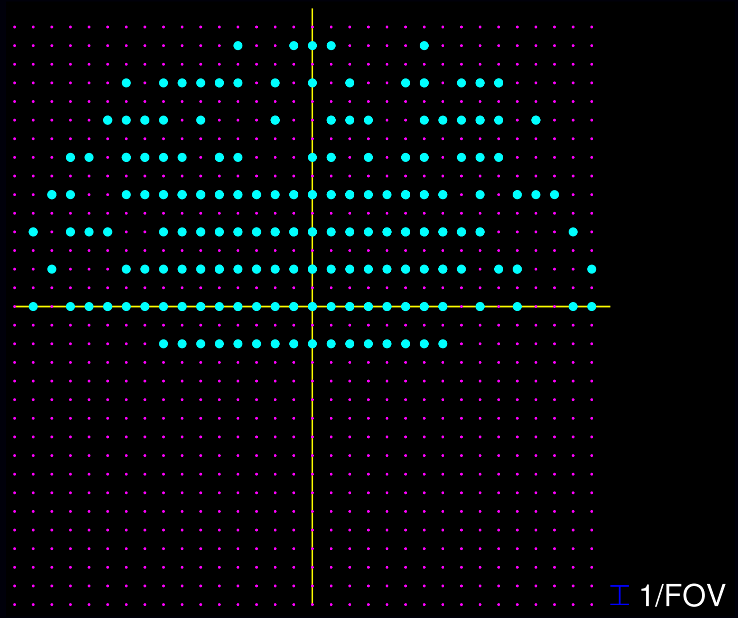
Conclusion

Because all MRI is under-sampled, the key question is not:

should we under-sample?

but rather *how* shall we under sample?





Caution: both produce "wrong" reconstructed images, but in different ways.

All MRI reconstruction methods involve constraints.

So the key question is:

which constraints are most appropriate for a given application?

Resources

Talk and code available online at

`http://www.eecs.umich.edu/~fessler`

References

[1] D. C. Noll, D. G. Nishimura, and A. Macovski. Homodyne detection in magnetic resonance imaging. *IEEE Trans. Med. Imag.*, 10(2):154–63, June 1991.

[2] J. A. Fessler and D. C. Noll. Iterative image reconstruction in MRI with separate magnitude and phase regularization. In *Proc. IEEE Intl. Symp. Biomed. Imag.*, pages 209–12, 2004.

[3] M. Bydder and M. D. Robson. Partial Fourier partially parallel imaging. *Mag. Res. Med.*, 53(6):1393–401, June 2005.

[4] K. P. Pruessmann, M. Weiger, M. B. Scheidegger, and P. Boesiger. SENSE: sensitivity encoding for fast MRI. *Mag. Res. Med.*, 42(5):952–62, November 1999.

[5] M. A. Griswold, P. M. Jakob, R. M. Heidemann, M. Nittka, V. Jellus, J. Wang, B. Kiefer, and A. Haase. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Mag. Res. Med.*, 47(6):1202–10, June 2002.

[6] K. F. King and L. Angelos. SENSE with partial Fourier homodyne reconstruction. In *Proc. Intl. Soc. Mag. Res. Med.*, page 153, 2000.

[7] J. I. Jackson, C. H. Meyer, D. G. Nishimura, and A. Macovski. Selection of a convolution function for Fourier inversion using gridding. *IEEE Trans. Med. Imag.*, 10(3):473–8, September 1991.

[8] K. P. Pruessmann, M. Weiger, P. Börnert, and P. Boesiger. Advances in sensitivity encoding with arbitrary k-space trajectories. *Mag. Res. Med.*, 46(4):638–51, October 2001.

[9] N. Seiberlich, F. A. Breuer, M. Blaimer, K. Barkauskas, P. M. Jakob, and M. A. Griswold. Non-Cartesian data reconstruction using GRAPPA operator gridding (GROG). *Mag. Res. Med.*, 58(6):1257–65, December 2007.

[10] Y. Cao and D. N. Levin. Feature-recognizing MRI. *Mag. Res. Med.*, 30(3):305–17, September 1993.

[11] Y. Cao and D. N. Levin. Using an image database to constrain the acquisition and reconstruction of MR images of the human head. *IEEE Trans. Med. Imag.*, 14(2):350–61, June 1995.

[12] G-H. Chen, J. Tang, and S. Leng. Prior image constrained compressed sensing (PICCS): A method to accurately reconstruct dynamic CT images from highly undersampled projection data sets. *Med. Phys.*, 35(2):660–3, February 2008.

[13] C. François, J. Tang, and G. H. Chen. Retrospective enhancement of radially undersampled cardiac cine MR images using prior image constrained compressed sensing (PICCS). In *Proc. Intl. Soc. Mag. Res. Med.*, page 3808, 2009.

[14] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed sensing MRI. *IEEE Sig. Proc. Mag.*, 25(2):72–82, March 2008.

[15] M. Lustig, D. Donoho, and J. M. Pauly. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Mag. Res. Med.*, 58(6):1182–95, December 2007.

[16] S. Ramani and J. A. Fessler. An augmented Lagrangian method for regularized MRI reconstruction using SENSE. In *Proc. Intl. Soc. Mag. Res. Med.*, page 2873, 2011. To appear as poster.

[17] S. Ramani and J. A. Fessler. Parallel MR image reconstruction using augmented Lagrangian methods. *IEEE Trans. Med. Imag.*, 30(3):694–706, March 2011. nihms 283021.

[18] F. Zhao, J. A. Fessler, J-F. Nielsen, and D. C. Noll. Separate magnitude and phase regularization via compressed sensing. In *Proc. Intl. Soc. Mag. Res. Med.*, page 2841, 2011. To appear as poster.

[19] S. Ravishankar and Y. Bresler. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans. Med. Imag.*, 30(5):1028–41, May 2011.

[20] M. Aharon, M. Elad, and A. Bruckstein. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Sig. Proc.*, 54(11):4311–22, November 2006.

[21] A. Buades, B. Coll, and J. M. Morel. A review of image denoising methods, with a new one. *SIAM Multiscale Modeling and Simulation*, 4(2):490–530, 2005.

[22] G. Adluru, T. Tasdizen, M. C. Schabel, and E. V. R. DiBella. Reconstruction of 3D dynamic contrast-enhanced magnetic resonance imaging using nonlocal means. *J. Mag. Res. Im.*, 32(5):1217–27, November 2010.

[23] J. V. Manjón, P. Coup, A. Buades, V. Fonov, D. L. Collins, and M. Robles. Non-local MRI upsampling. *Med. Im. Anal.*, 14(6):784–92, December 2010.

[24] Z. Yang and M. Jacob. A unified energy minimization framework for nonlocal regularization. In *Proc. IEEE Intl. Symp. Biomed. Imag.*, pages 1150–, 2011.

[25] G. Wang and J. Qi. Patch-based regularization for iterative PET image reconstruction. In *Proc. IEEE Intl. Symp. Biomed. Imag.*, pages 1508–, 2011.