Empirical Phase Transitions in Sparsity-Regularized Computed Tomography

Jakob Sauer Jørgensen

Postdoc, DTU

Sparse Tomo Days, DTU, March 28, 2014

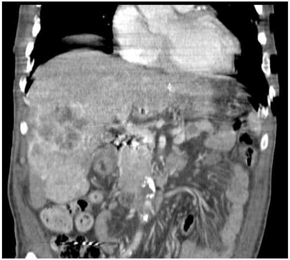
Joint work with

Per Christian Hansen, DTU

Emil Sidky and **Xiaochuan Pan**, U. Chicago

Christian Kruschel and **Dirk Lorenz**, TU Braunschweig

Exploiting prior knowledge in CT



Discrete imaging model:

$$Ax = b$$

Typical CT images:

- Regions of homogeneous tissue.
- Separated by sharp boundaries.

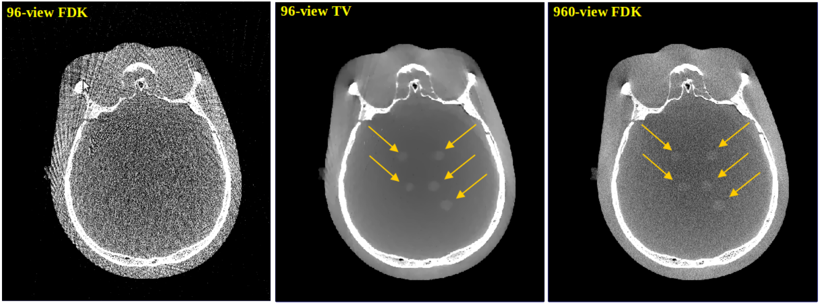
Reconstruction by regularization:

$$x^\star \quad = \quad \underset{x}{\operatorname{argmin}} \quad \mathcal{D}(Ax, b) \quad + \quad \lambda \cdot \mathcal{R}(x)$$

Sparsity-promoting choices:

- $\mathcal{R}(x) = \|x\|_1$ (ℓ_1/basis pursuit)
- $\mathcal{R}(x) = \|x\|_{\mathrm{TV}}$ (total variation)
- $\mathcal{R}(x) = \|D^T x\|_1$ (analysis-ℓ_1)

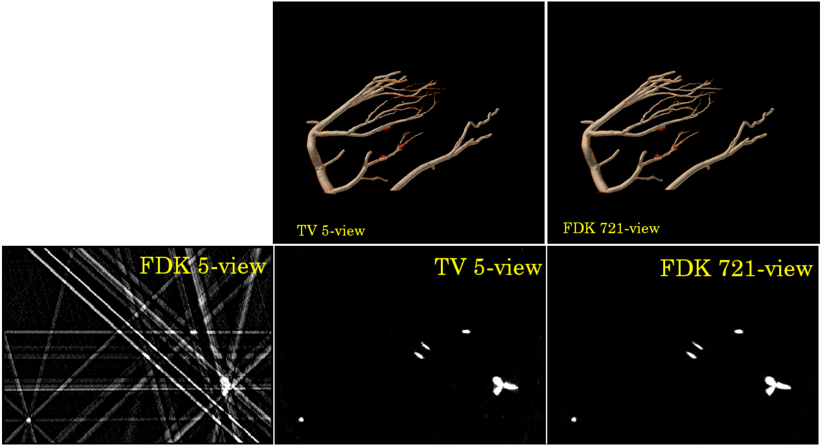
TV example: Physical head phantom, CB-CT



(Bian 2010).

Courtesy of X. Pan, U. Chicago.

TV example: Human coronary artery, CB-CT



Courtesy of X. Pan, U. Chicago.

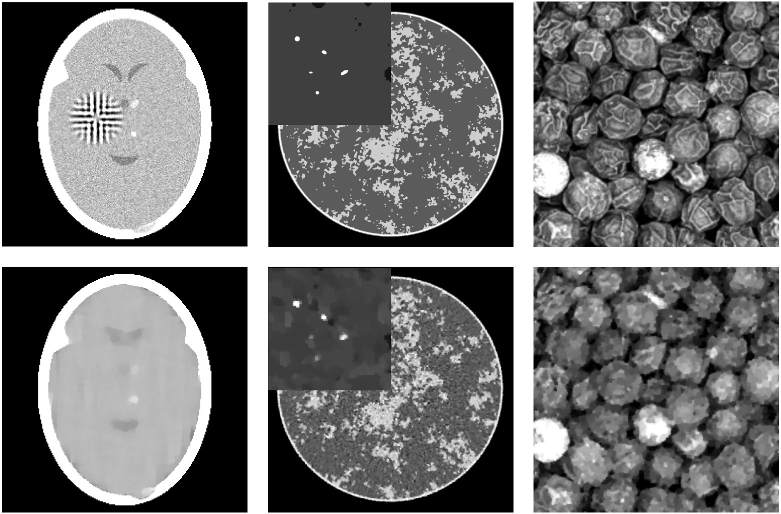
Data collected with a bench-top CB-CT of Dr. E. Ritman at Mayo

Less successful CT cases for TV:

(Herman & Davidi, 2008) (J. et al, 2011) (Courtesy of S. Soltani, DTU)

Original

TV



Lack of quantitative understanding

Some fundamental questions remain unanswered:

- Under what conditions will reconstruction work?
- Robustness to noise?
- Which types of images?
- *What is sufficient sampling?*

Application-specific vs. general

- Focus on the imaging model.

Classical CT sampling results

Continuous image and data:

- Based on invertibility and stability of Radon transform etc.
- Fan-beam: 180° plus fan-angle
- Cone-beam: Tuy's condition

Discrete data:

- Nyquist sampling
- Assumption of bandlimited signal
- (Huesmann 1977, Natterer)

Reconstruction with sparse/compressible signal assumption?

Compressed Sensing

Guarantees of accurate reconstruction

- Under suitable assumptions, a *sufficiently sparse signal* can be *recovered* from few measurements by ℓ_1-minimization.
- RIP, incoherence, spark, ...

For tomography?

- So far no practically useful guarantees.
- Results for certain discrete tomography cases (Petra et al.)

This study:

- Empirical study of sampling conditions for tomographic reconstruction of sparse signals
- Recoverability of single images
- Worst-case vs. average case

Reconstruction problems

Inequality-constrained regularization:

$$x^\star = \operatorname*{argmin}_x \mathcal{R}(x) \quad \text{s.t.} \quad \|Ax - b\|_2 \le \epsilon$$

Simplified reconstruction problems:

$$\begin{array}{lllll} \textbf{BP} & x_{\textbf{BP}} = \operatorname*{argmin}_x & \|x\|_1 & \text{s.t.} & Ax = b \\ \textbf{ATV} & x_{\textbf{ATV}} = \operatorname*{argmin}_x & \|D^T x\|_1 & \text{s.t.} & Ax = b \end{array}$$

finite-difference approximation of gradient

Algorithms:

- Our interest: Reliably obtaining accurate solution, not speed.
- Recast as linear programs (LPs) and solve by MOSEK.

Non-uniqueness of solutions

Both BP and ATV can have multiple solutions for same data:

- 1-norm convex, but not strictly convex.
- Even if x_{orig} is a minimizer, others may exist.

Consequences:

- Different algorithms may produce different solutions.
- Decision of recoverability of x_{orig} is algorithm-dependent.

Alternative idea:

- Can we test for uniqueness of solution?

Uniqueness test for BP

Given:

- $b = Ax_{\text{orig}}$
- $I = \text{support}(x_{\text{orig}})$
- A_I is A with columns I

Characterization of solution uniqueness:

- x_{orig} uniquely minimizes $\min_x \|x\|_1$ s.t. $Ax = b$

if and only if

- A_I is injective, and
- $\exists w : \quad A_I^T w = \text{sign}(x_{\text{orig}})_I$ and $\|A_{I^c}^T w\|_\infty < 1$

(Plumbley 2007, Fuchs 2004, Grasmair et al. 2011)

Uniqueness test for ATV

Given:

- $b = Ax_{\text{orig}}$
- $I = \text{support}(D^T x_{\text{orig}})$
- D_I is D with columns I

Characterization of solution uniqueness:

- x_{orig} uniquely minimizes $\min_x \|D^T x\|_1$ s.t. $Ax = b$

if and only if

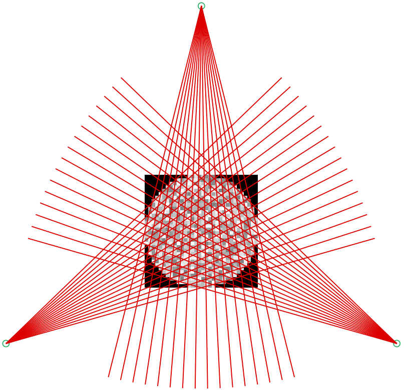
- $\binom{A}{D_{I^c}^T}$ is injective, and
- $\exists w, v : \quad Dv = A^T w, \quad v_I = \text{sign}(D_I^T x_{\text{orig}}), \quad \|v_{I^c}\|_\infty < 1$

Application of (Haltmeier 2013)

Uniqueness testing procedure using LP

	BP	**ATV**
1. Check injectivity:	A_I	$\begin{pmatrix} A \\ D_{I^c}^T \end{pmatrix}$
2. Solve LP:	$\begin{aligned} t^\star &= \operatorname{argmin}\ t \\ -te \leq A_{I^c}^T &w \leq te \\ A_I^T w &= \operatorname{sign}(x_{\mathsf{orig}})_I \end{aligned}$	$\begin{aligned} t^\star &= \operatorname{argmin}\ t \\ -te \leq v_{I^c} &\leq te \\ Dv &= A^T w \\ v_I &= \operatorname{sign}(D_I^T x_{\mathsf{orig}}) \end{aligned}$
3. Unique iff:	$t^\star < 1$	$t^\star < 1$

The geometry and system matrix



- Disk-shaped image inscribed in $N_{\text{side}} \times N_{\text{side}}$ square.
- Number of pixels:

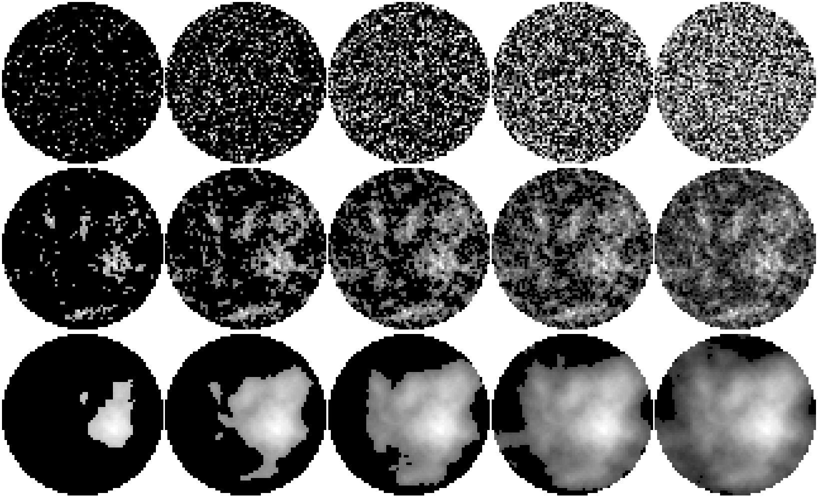
$$N \approx \frac{\pi}{4} N_{\text{side}}^2$$

- Fan-beam, equi-angular views ($N_{\text{views}} = 3$ shown)
- Number of rays per view: $2N_{\text{side}}$
- System matrix A size:

$$\begin{array}{cc} & N \\ M = N_{\text{views}} \cdot 2N_{\text{side}} & \boxed{\quad A \quad} \end{array}$$

Elements A_{ij} computed by the line intersection method (implementation: www.imm.dtu.dk/~pch/AIRtools/)

BP image class examples images



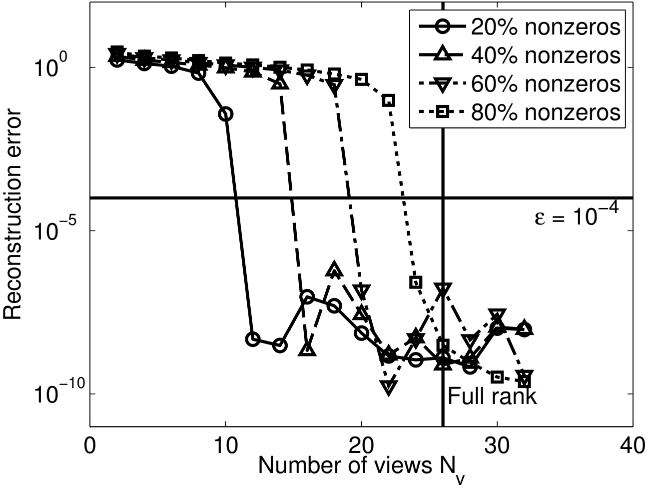
spikes

1-power

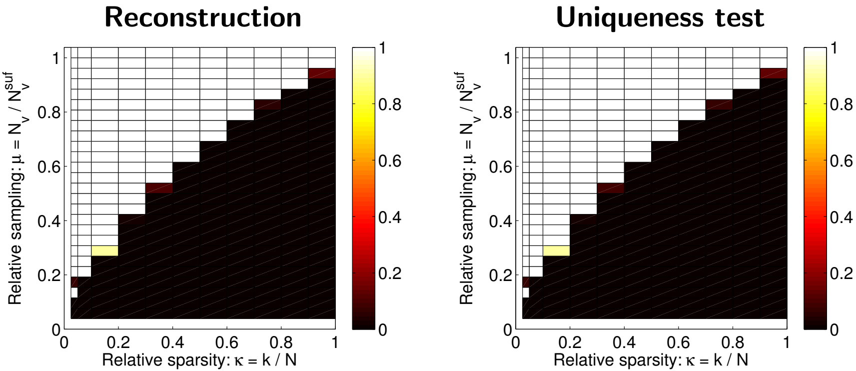
2-power

Relative sparsity: 0.1, 0.3, 0.5, 0.7, 0.9

Reconstruction error vs. sampling and sparsity



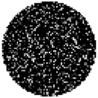
Phase diagrams: spikes with BP

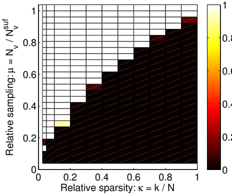


- Fraction recovered/unique of 100 instances at each point (κ, μ) of relative sparsity and sampling.
- Excellent agreement of reconstruction and uniqueness test.
- Well-separated "no-recovery" and "full-recovery" regions.
- Phase transition as in compressed sensing (Donoho-Tanner).

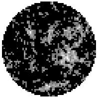
Comparing image classes, BP

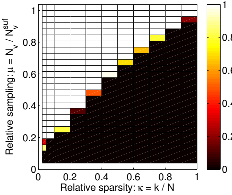
spikes



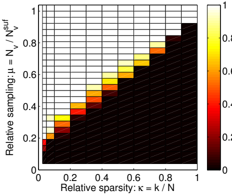


1-power

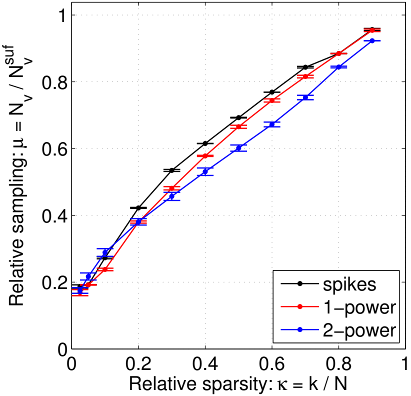




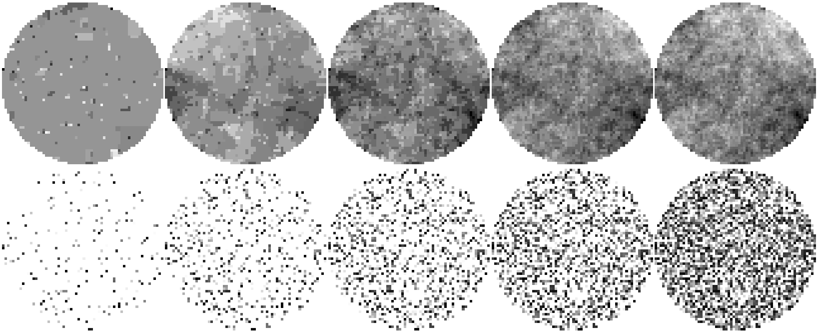
2-power



Average sufficient sampling



Example images: altproj, trununif

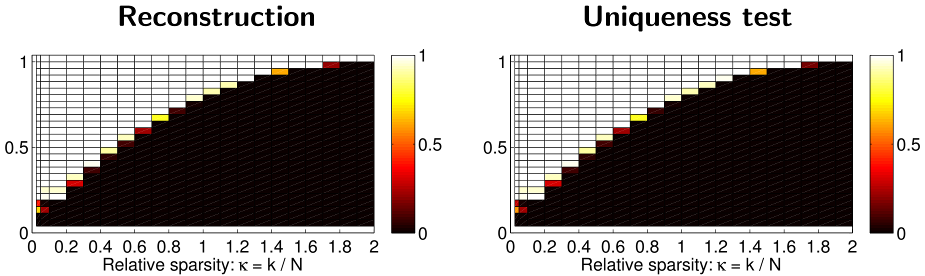


altproj

trununif

Relative sparsity 0.1 0.3 0.5 0.7 0.9

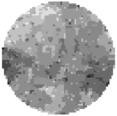
Phase diagrams: altproj with ATV

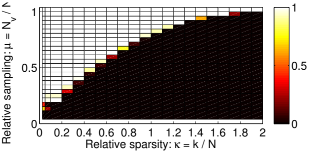


- Fraction recovered/unique of 100 instances at each point (κ, μ) of relative sparsity and sampling.
- Excellent agreement of reconstruction and uniqueness test.
- Well-separated "no-recovery" and "full-recovery" regions.
- Phase transition as in compressed sensing (Donoho-Tanner).

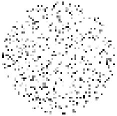
Comparing image classes, ATV

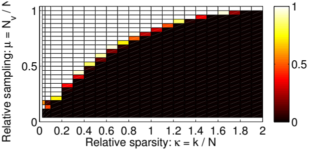
altproj



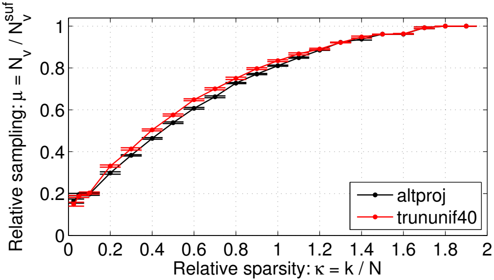


trununif

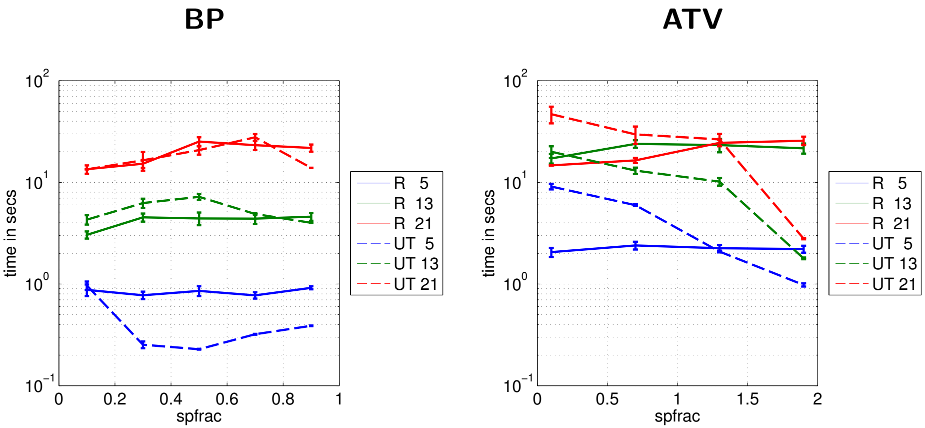




Average sufficient sampling



Time: Reconstruction vs. uniqueness test



- 10 repetitions at each relative sparsity and 5, 13, 21 views.
- Comparable time of reconstruction (R) and uniqueness test (UT).

A more well-known image: Shepp-Logan on disk

BP

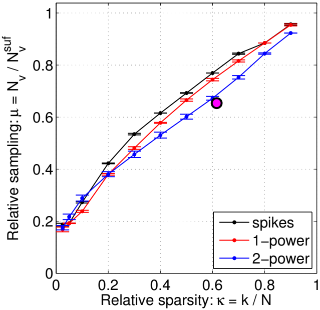
$\|x_{\text{orig}}\|_0 = \mathbf{1988}$

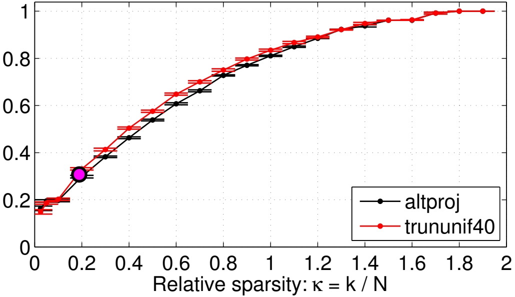
$N_{\text{v}}^{\textbf{BP}} = \mathbf{17}$

ATV

$\|Dx_{\text{orig}}\|_0 = \mathbf{610}$

$N_{\text{v}}^{\textbf{ATV}} = \mathbf{8}$





Conclusions and future work

Conclusions

- Empirical evidence of relation between sparsity and sampling
- Reconstruction and uniqueness test
- Phase transition from no to full recovery
- Small dependence on image class, mostly sparsity
- Additional results (not shown): limited angle, robustness to noise, scaling with image size.

Future work and open questions

- Extensions: Isotropic TV, more realistic image classes, ...
- Theoretical/compressed sensing explanation?
- Connection to classical CT sampling results?