



Research | October 02, 2017

From Global to Local: Getting More from Compressed Sensing

By Alexander Bastounis, Ben Adcock, and Anders C. Hansen

Over the last decade, compressed sensing and sparse recovery techniques have had a significant impact on applied mathematics and its uses in science and engineering. Compressed sensing applications have moved beyond experimentation and are beginning to be seen in new implementations. An area of particular note is imaging, where compressed sensing can be used in magnetic resonance imaging (MRI), electron tomography, and radio interferometry, among other applications. With this in mind, it is timely to revisit the mathematics of compressed sensing as it pertains to imaging. While the standard theory of compressed sensing is justifiably celebrated, it falls somewhat short of explaining phenomena that result from the application of these techniques to imaging. In this article, we describe recent work that seeks to bridge this gap. As we demonstrate, our approach yields significant practical benefits in imaging, allowing researchers to further push the limits of performance.

Standard Compressed Sensing

Compressed sensing [3, 4, 6] concerns the recovery of an object from an incomplete set of linear measurements. In a discrete setting, one can formulate this as the linear system

$$y = Ax,$$

where $y \in \mathbb{C}^m$ is the vector of measurements, $x \in \mathbb{C}^N$ is the object to recover, and $A \in \mathbb{C}^{m \times N}$ is the so-called measurement matrix. In practice, the number of measurements m is often substantially smaller than the dimension N , making recovery generally impossible. To overcome this, compressed sensing leverages two key properties: *sparsity* of the vector x and *incoherence* of the measurement vectors (rows of A). The first property asserts that x should have at most $s \leq m$ significant components, with the remaining components relatively small, while the second states that the measurement vectors should be (in a formally-defined sense) spread out, rather than concentrated around a small number of entries.

A popular tool in compressed sensing theory is the *Restricted Isometry Property* (RIP). A matrix has the RIP or order s if there exists a $\delta \in (0, 1)$, such that

$$(1 - \delta)\|x\|_{\ell^2}^2 \leq \|Ax\|_{\ell^2}^2 \leq (1 + \delta)\|x\|_{\ell^2}^2, \text{ for all } s\text{-sparse vectors } x.$$

For instance, if recovery is performed by solving the convex *basis pursuit* problem

$$\min_{z \in \mathbb{C}^N} \|z\|_{\ell^1}, \text{ subject to } Az = y, \quad (1)$$

then the RIP (with sufficiently small δ) implies exact recovery of any s -sparse x and robustness with respect to perturbations in x (i.e., inexact sparsity) or y (i.e., noise).

Typically, the rows of A are drawn independently according to some random distribution. An elegant demonstration of compressed sensing mathematics considers Gaussian random vectors. These are incoherent, and a signature result asserts that A has the RIP with an optimal number of measurements $m \approx Cs \log(N/s)$.

The Flip Test

Imaging is an ideal fit for compressed sensing, and one of its original areas of application [5, 8, 11]. Though not typically sparse themselves, images can be represented sparsely in certain dictionaries, such as wavelets. Furthermore, acquisition devices found in many imaging applications are modelled with the Fourier transform, which tends to be fairly incoherent.

In their seminal work [3], Emmanuel Candès, Justin Romberg, and Terence Tao demonstrated the benefits of compressed sensing by recovering the classical Shepp-Logan phantom from incomplete Fourier measurements. This experiment is repeated in Figures 1a-1c. The theoretical

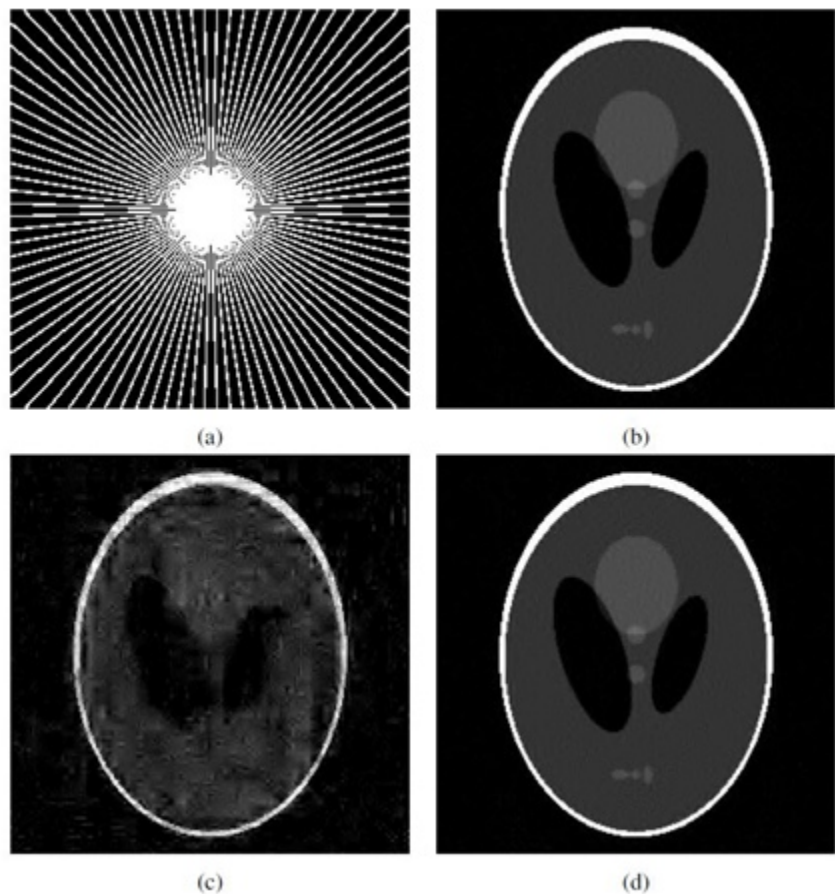


Figure 1. The flip test. 1a. Radial sampling map in k -space. White pixels denote the frequencies sampled. 1b. Image \hat{x} , recovered using Haar wavelets (PSNR = 28.7dB). 1c. Flipped recovery \check{x} (PSNR = 15dB). 1d. Flipped recovery where the flipping is done in levels (PSNR = 29.0dB). Image credit: Alexander Bastounis, Ben Adcock, and Anders C. Hansen.

basis for this result is twofold. First, the image $x \in \mathbb{C}^N$ has a sparse representation in a wavelet basis. Specifically, if $\Phi \in \mathbb{C}^{N \times N}$ is the matrix of the wavelet transform, then

$$x = \Phi c$$

for some s -sparse vector of coefficients $c \in \mathbb{C}^N$. Second, the matrix

$$A = PF\Phi \quad (2)$$

has the RIP, where $F \in \mathbb{C}^{N \times N}$ is the discrete Fourier matrix and $P \in \mathbb{C}^{m \times N}$ is the matrix of the *sampling map*, i.e., P selects rows of F according to frequencies shown in Figure 1b. As per the theory, x should recover to high accuracy as $\hat{x} = \Phi \hat{c}$, where \hat{c} is a solution of (1).

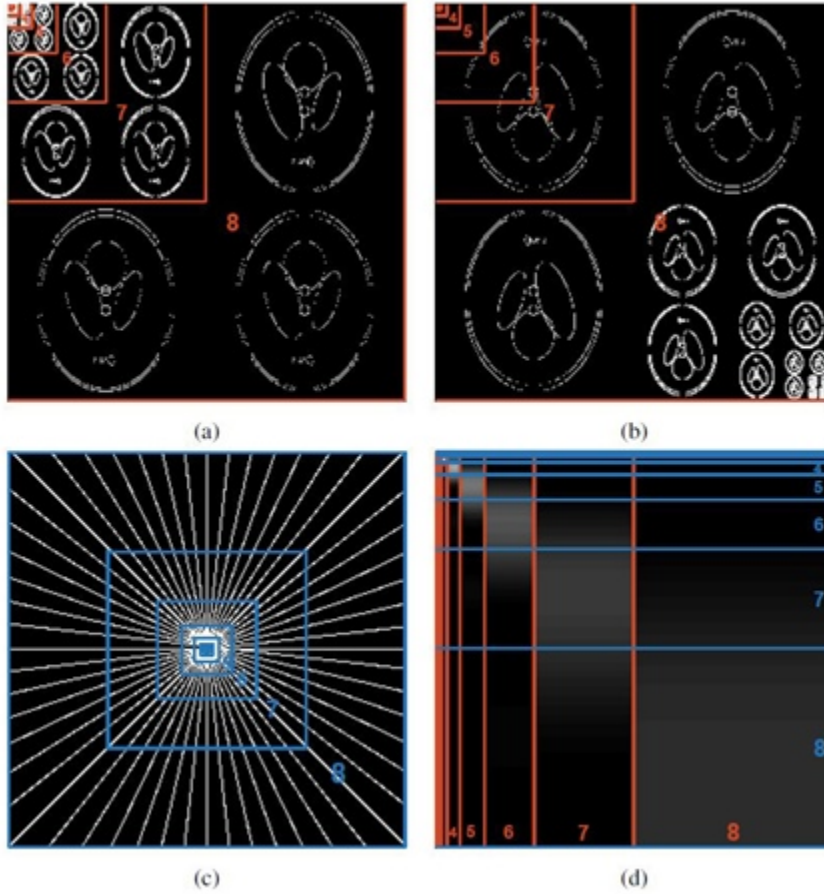


Figure 2. Asymptotic sparsity and asymptotic incoherence. 2a. Wavelet coefficients of the Shepp-Logan phantom, arranged according to increasing scale. 2b. Coefficients of the flipped phantom. 2c. Radial sampling map in k -space. The square annulus regions denote the essential frequency concentration of wavelets at a given scale. 2d. Absolute values of the matrix $U = \Phi F$, with larger and smaller values corresponding to lighter or darker colours respectively. Vertical lines indicate the wavelet scales and horizontal lines indicate the annular frequency bands. Image credit: Alexander Bastounis, Ben Adcock, and Anders C. Hansen.

To examine the extent to which this theory explains the results observed in Figure 1, we perform the following experiment, known as the *flip test* [1]. Let $Q \in \mathbb{C}^{N \times N}$ be a permutation matrix and define the permuted coefficients $c_p = Qc$ and corresponding image $x_p = \Phi c_p$. Now let \hat{c}_p be the coefficients recovered by solving (1) with $y = PFx_p$, and define $\check{c} = Q^{-1}\hat{c}_p$ and $\tilde{x} = \Phi \check{c}$. Since permutations do not affect sparsity, coefficients c_p are s -sparse and image x_p has an s -sparse representation in the wavelet basis. Hence, if matrix A has the RIP, one would expect both the *unflipped* reconstruction \hat{x} and *flipped* reconstruction \tilde{x} to yield similar recoveries of the original image x .

Figure 1d demonstrates that this is not the case. The flipped reconstruction—in this example, the permutation simply reverses the ordering of the coefficients—is significantly worse than the unflipped reconstruction. We therefore conclude that the RIP cannot hold. Additionally, since

certain sparse vectors are recovered better than others, *distribution* of the wavelet coefficients is crucial to recovery.

Classical wavelet theory can intuitively explain these results. As illustrated in Figure 2a, wavelet coefficients, when arranged according to dyadic scales, are sparser at finer scales than coarser scales. Moreover, wavelets at a given scale are essentially concentrated in square annular regions of k -space (see Figure 2c). The radial sampling pattern samples less densely in regions corresponding to the fine scales, where the image is more sparse, and more densely at coarse scales, where the image is less sparse. However, if the coefficients are permuted (see Figure 2b), too many coefficients exist at fine scales (compared to the number of high-frequency samples) to ensure good recovery.

A Levels-based Framework

To provide a more comprehensive compressed sensing framework, the approach in [1, 2] first replaces the global concepts of sparsity and incoherence with suitable local quantities. Specifically, let r be a number of levels and $\mathbf{M} = (M_1, \dots, M_r)$, where

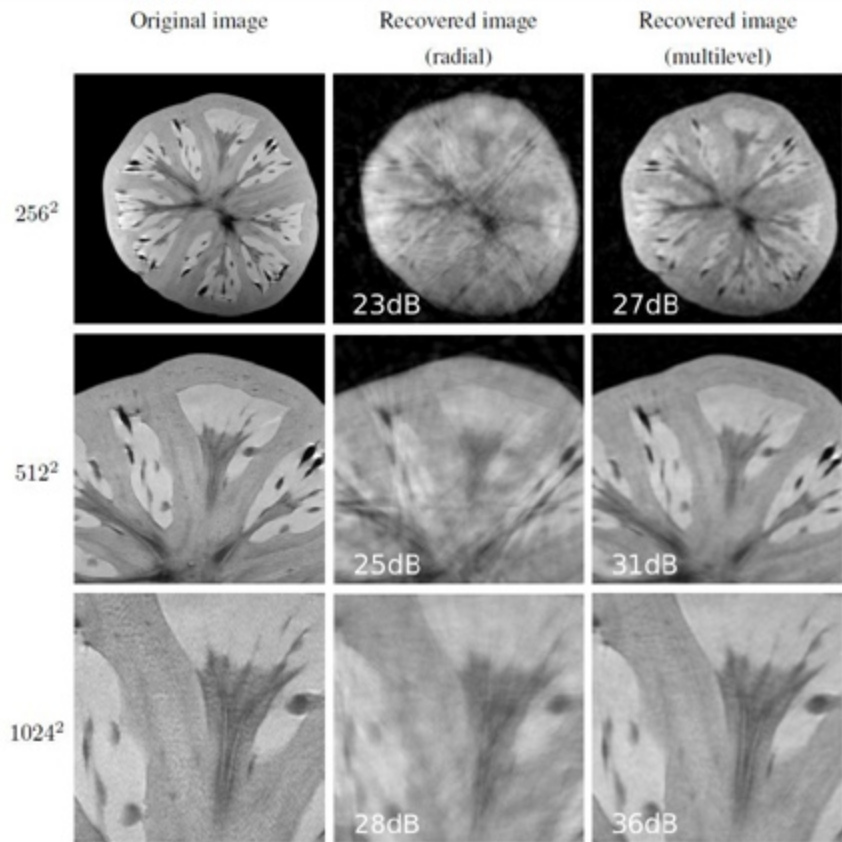


Figure 3. Compressed sensing using 6.25% Fourier measurements at various resolutions. Original image courtesy of Andy Ellison, recovered images by Alexander Bastounis, Ben Adcock, and Anders C. Hansen.

$1 \leq M_1 < \dots < M_r = N$, a vector of *sparsity levels*. These may typically correspond to wavelet scales. Rather than a single sparsity index s , the new model considers a vector $\mathbf{s} = (s_1, \dots, s_r)$ of local sparsities, with s_k as the sparsity at the k th level. We refer to vector $\mathbf{x} \in \mathbb{C}^N$ with this sparsity pattern as (\mathbf{s}, \mathbf{M}) -sparse in levels.

Note that permutations performed in Figure 2c destroy sparsity in levels but not global sparsity. Conversely, Figure 2d demonstrates that permutations within scales do not unduly alter the reconstruction quality, thus demonstrating the appropriateness of the (\mathbf{s}, \mathbf{M}) -sparsity model.

A modified version of the RIP helps analyze recovery with this model [2]. Matrix A has the *RIP in levels* (RIPL) of order (\mathbf{s}, \mathbf{M}) if there exists a $\delta \in (0, 1)$, such that

$$(1 - \delta)\|x\|_{\ell^2}^2 \leq \|Ax\|_{\ell^2}^2 \leq (1 + \delta)\|x\|_{\ell^2}^2, \text{ for all } (\mathbf{s}, \mathbf{M})\text{-sparse vectors } x.$$

Much like the standard RIP, if A has the RIPL (for small $\delta_{\mathbf{s}, \mathbf{M}}$), then all (\mathbf{s}, \mathbf{M}) -sparse vectors can be robustly recovered by solving (1) .

Returning to Fourier sampling with wavelet sparsity, this novel sparsity model calls for a new type of sampling, known as *multilevel random subsampling*. The idea is to break up the rows of the matrix U into levels [1], following the block-diagonal structure illustrated in Figure 2d. Specifically, we introduce *sampling levels* $\mathbf{N} = (N_1, \dots, N_r)$, where $1 \leq N_1 < \dots < N_r = N$, and a vector $\mathbf{m} = (m_1, \dots, m_r)$ of local numbers of measurements. Within each sampling level, m_k samples are chosen uniformly at random. Using this approach, one can show that the matrix (2) satisfies the RIPL (in the one-dimensional setting) [7], provided

$$m_k \approx C \left(s_k + \sum_{\substack{l=1 \\ j \neq k}}^r 2^{-|l-k|} s_l \right) \cdot \log^3(N) \cdot \log^2(s), \quad k = 1, \dots, r. \quad (3)$$

That is, the number of measurements m_k required to capture each wavelet scale should be roughly proportional to the corresponding sparsity s_k .

Applications and Benefits

By refining the sparsity model and sampling procedure, this framework not only explains the observations of the flip test but also significantly enhances compressed sensing performance in various imaging settings. By following sparsity patterns of the wavelet coefficients, one can exploit (3) to develop sampling patterns that target the sparsity in levels structure and thereby enhance reconstruction quality.

We conclude by demonstrating these benefits in several practical settings (see [10] for further experiments). First, Figure 3 compares the recovery of a magnetic resonance image at various resolutions from Fourier measurements, taken according to radial and multilevel sampling patterns. Multilevel sampling is consistently superior to radial sampling because it better targets the image's sparsity structure. This benefit also increases with the resolution, since the multilevel sampling pattern aligns increasingly well with the wavelet coefficients' *asymptotic* sparsity.



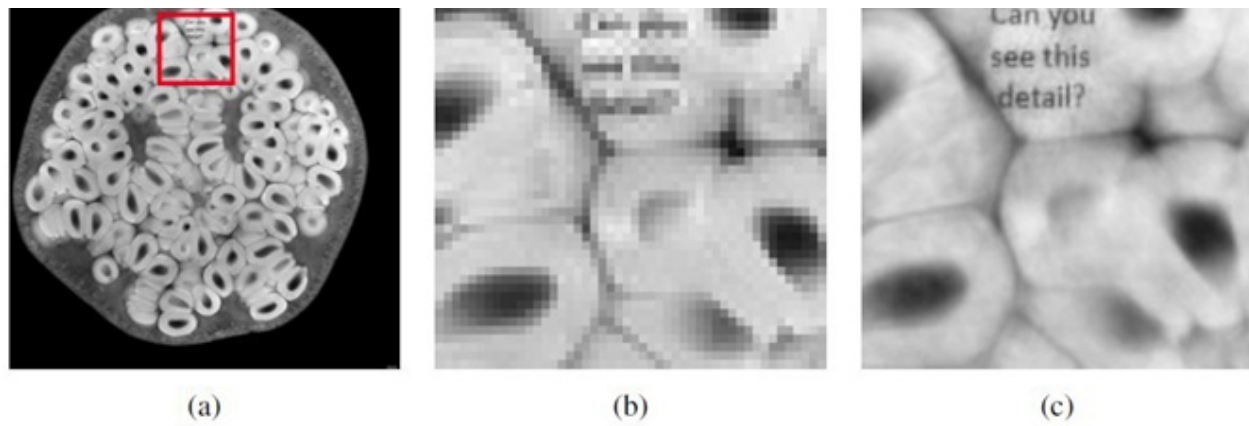


Figure 4. Resolution enhancing in MRI. 4a. Original image with small synthetic detail added. 4b. Linear recovery from the lowest $m=256^2$ Fourier measurements. 4c. Compressed sensing with multilevel subsampling using $m=256^2$ measurements. Original image courtesy of Andy Ellison, recovered images by Alexander Bastounis, Ben Adcock, and Anders C. Hansen.

From this latter observation we conclude the following: instead of using compressed sensing at lower resolutions to reduce acquisition time, one can best realize the full benefits by subsampling at higher resolutions and seeking to improve image quality. In other words, compressed sensing is most beneficial as a *resolution enhancer*. Figure 4 demonstrates this effect. For a fixed budget of measurements, subsampling from higher resolutions yields a vastly superior reconstruction when compared to full sampling at low frequencies. Both Siemens—a leading manufacturer of MRI scanners [13]—and [10] further verify this phenomenon in a practical MRI setting.

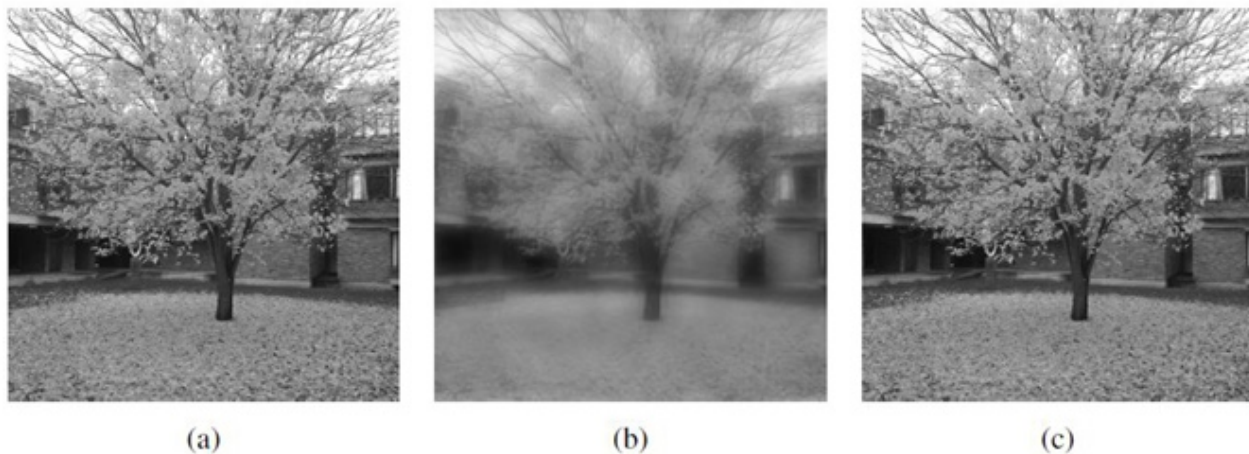


Figure 5. Compressive imaging. 5a. Original image. 5b. Recovery from $m=16.5\%$ scrambled Hadamard measurements. 5c. Recovery from $m=16.5\%$ multilevel subsampled Hadamard measurements. Image credit: Alexander Bastounis, Ben Adcock, and Anders C. Hansen.

Finally, Figure 5 considers a class of problems informally known as *compressive imaging* [11]. In these problems—the applications of which include single-pixel [5] and lensless imaging, infrared imaging, and fluorescence microscopy [12]—one can choose the measurement matrix A , provide \mathbf{y} that its entries are binary. In this case, \mathbf{A} is a randomly-sampled Hadamard transform with

scrambled columns is a standard choice for A . This is a computationally-efficient procedure whose performance mimics that of random Gaussian sampling; it is near-optimal for recovering sparse vectors. It may therefore seem surprising that the reconstruction quality can be improved. However, as Figure 5 shows, a multilevel subsampled Hadamard transform (without column scrambling) does precisely this. Even though wavelet coefficients are sparse, the procedure targets the image's fine details (captured by the fine-scale wavelet coefficients) to achieve a significant performance gain.

Acknowledgments: The authors thank Andy Ellison for the MR images used in Figures 3 and 4.

References

- [1] Adcock, B., Hansen, A.C., Poon, C., & Roman, B. (2017). Breaking the coherence barrier: A new theory for compressed sensing. *Forum of Mathematics, Sigma*, 5.
- [2] Bastounis, A., & Hansen, A.C. (2017). On the absence of uniform recovery in many real-world applications of compressed sensing and the restricted isometry property and nullspace property in levels. *SIAM J. Imaging Sci.*, 10(1), 335-371.
- [3] Candès, E.J., Romberg, J., & Tao, T. (2006). Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Trans. Inform. Theory*, 52(2), 489-509.
- [4] Donoho, D.L. (2006). Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4), 1289-1306.
- [5] Duarte, M.F., Davenport, M.A., Takhar, D., Laska, J., Kelly, K., & Baraniuk, R.G. (2008). Single-pixel imaging via compressive sampling. *IEEE Signal Process. Mag.*, 25(2), 83-91.
- [6] Foucart, S., & Rauhut, H. (2013). *A Mathematical Introduction to Compressive Sensing: Applied and Numerical Harmonic Analysis*. New York, NY: Birkhauser.
- [7] Li, C., & Adcock, B. (2017). Compressed sensing with local structure: uniform recovery guarantees for the sparsity in levels class. *Appl. Comput. Harmon. Anal.* (to appear). *arXiv:1601.01988*.
- [8] Lustig, M., Donoho, D.L., Santos, J.M., & Pauly, J.M. (2008). Compressed Sensing MRI. *IEEE Signal Process. Mag.*, 25(2), 72-82.
- [9] Poon, C. (2015). On the role of total variation in compressed sensing. *SIAM J. Imaging Sci.*, 8(1), 682-720.
- [10] Roman, B., Calderbank, R., Adcock, B., Nietlispach, D., Bostock, M., Calvo-Almazán, I., Graves, M., & Hansen, A.C. (2017). Undersampling improves fidelity of physical imaging and the benefits grow with resolution. *Preprint*.
- [11] Romberg, J. (2008). Imaging via compressive sampling. *IEEE Signal Process. Mag.*, 25(2), 14-20.
- [12] Studer, V., Bobin, J., Chahid, M., Moussavi, H., Candès, E., & Dahan, M. (2011). Compressive fluorescence microscopy for biological and hyperspectral imaging. *Proc. Natl Acad. Sci.*, 109(26), 1679-1687.
- [13] Wang, Q., Zenge, M., Cetingul, H.E., Mueller, E., & Nadar, M.S. (2014). Novel sampling strategies for sparse MR image reconstruction. *Proc. Int. Soc. Mag. Res. in Med.*, 22.

Alexander Bastounis (a_bastounis@cam.ac.uk) is a final-year Ph.D. student in the Applied

Functional and Harmonic Analysis group in the Department of Applied Mathematics and Theoretical Physics at the University of Cambridge. His research interests include compressed sensing, optimisation theory, computability theory, and applied functional analysis. Ben Adcock (ben_adcock@sfu.ca) is an assistant professor at Simon Fraser University. He received the Leslie Fox Prize for Numerical Analysis in 2011, an Alfred P. Sloan Research Fellowship in 2015, and the CAIMS-PIMS Early Career Award in 2017. His research interests include applied and computational harmonic analysis, numerical analysis, and approximation theory. Anders C. Hansen (ach70@cam.ac.uk) is the head of the Applied Functional and Harmonic Analysis group at the University of Cambridge, where he is a reader (associate professor) in mathematics. He is also a Royal Society University Research Fellow. His research interests include computational harmonic analysis, foundations of computational mathematics, mathematical signal processing, functional analysis, numerical analysis, and mathematical physics.