On random and deterministic compressed sensing and the Restricted Isometry Property in Levels

Alexander Bastounis
CCA, Centre for Mathematical Sciences
University of Cambridge
Wilberforce Rd, Cambridge CB3 0WA

Anders C. Hansen
DAMTP, Centre for Mathematical Sciences
University of Cambridge
Wilberforce Rd, Cambridge CB3 0WA

***Abstract*—Compressed sensing (CS) is one of the great successes of computational mathematics in the past decade. There are a collection of tools which aim to mathematically describe compressed sensing when the sampling pattern is taken in a random or deterministic way. Unfortunately, there are many practical applications where the well studied concepts of *uniform recovery* and the *Restricted Isometry Property* (RIP) can be shown to be insufficient explanations for the success of compressed sensing. This occurs both when the sampling pattern is taken using a deterministic or a non-deterministic method. We shall study this phenomenon and explain why the RIP is absent, and then propose an adaptation which we term 'the RIP in levels' which aims to solve the issues surrounding the RIP. The paper ends by conjecturing that the RIP in levels could provide a collection of results for deterministic sampling patterns.**

I. Introduction

In sampling theory, a typical problem is to reduce the number of measurements required for the effective operation of a scanning device. This can be expressed in formal mathematics as solving an underdetermined system of linear equations. More precisely, let $x \in \mathbb{C}^n$ be a vector and let $M \in \mathbb{C}^{n\times n}$ be a scanning device (e.g. a Magnetic Resonance Imaging (MRI) scanner, a Computerized Tomography (CT) scanner etc.). We are tasked with finding x from the values of $y := P_\Omega Mx$, where P_Ω is a projection map from $\{1, 2, \dots, n\}$ onto the sampling set $\Omega \subseteq \{1, 2, \dots, n\}$ so that $(P_\Omega y)_i = y_i$ if $i \in \Omega$ and $(P_\Omega y)_i = 0$ if $i \notin \Omega$.

It is clear that, in general, it will be impossible to recover x from such information unless $\Omega = \{1, 2, \dots, n\}$. However, if we know a priori that x can be represented in a sparse way in a different basis $\{W_1, W_2, \dots W_n\}$, recovery may still be possible. Indeed, we can then write $y = P_\Omega MBw$ where w is a representation of x using the basis $(W_i)_{i=1}^n$ and B is the change of basis matrix from $(W_i)_{i=1}^n$ to the standard basis in $\mathbb{C}^n$. The theory of compressed sensing [1]–[6] states that under certain conditions which shall be discussed throughout this paper, solutions to

$$\min_{w'\in\mathbb{C}^n} \|w'\|_1 \text{ such that } y = P_\Omega MBw' \tag{1}$$

will be very good or exact approximations to w. The focus of this paper will be on structured sampling, which is needed in most applications, and ℓ^1 minimization. This is, however,

978-1-4673-7353-1/15/$31.00 ©2015 IEEE

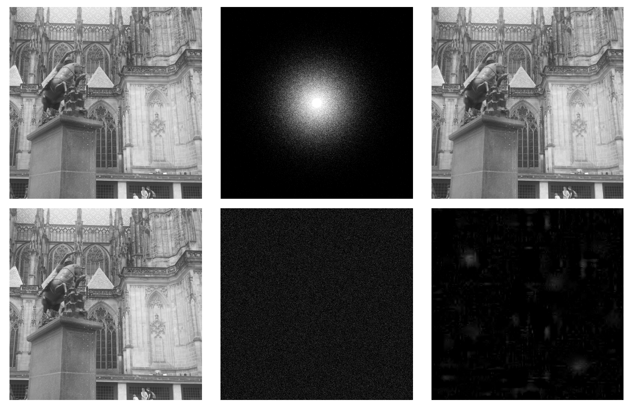

Fig. 1. The results of different subsampling patterns when compressed sensing with the matrix $\mathrm{DFT}\,\mathrm{DWT}^{-1}$ is performed. The same number of samples (7.9%) were taken in both cases

entirely distinct from theories of model based compressed sensing [7], which shall not be discussed further here.

A standard example that we shall use repeatedly is the case where M is a discrete Fourier transform (DFT), B is the inverse wavelet transform (henceforth denoted DWT^{-1}) and w is a collection of wavelet coefficients of an image x. This Fourier-Wavelet example can be used to simulate an MRI scanner, spectroscopy, electron microscopy and computerized tomography among many others. If we are permitted to choose Ω ourselves, then a strong choice of Ω with $|\Omega| \ll n$ is to exploit so-called multi-level sampling. This is detailed in [8] and an example is provided in Figure 1.

One of our main focuses is that of *recovery with deterministic sampling*. In this case, we are given a fixed Ω and are then tasked with mathematically describing the effectiveness of solving (1). Of particular importance is the case where Ω is a collection of radial lines (see Figure 2), and as before M, B and w represent DFT, DWT^{-1} and the wavelet coefficients of an image respectively. It will, however, be helpful to begin by discussing non-deterministic sampling, i.e. where Ω is permitted to be randomly chosen.

II. Uniform recovery in compressed sensing

With a random choice of Ω, there are a number of possible ways of mathematically quantifying recovery. Firstly, one can

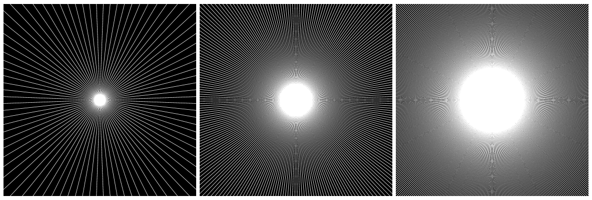

Fig. 2. Various radial line sampling patterns on a 2048×2048 pixel grid, with 44, 128 and 256 radial lines respectively.

randomly choose Ω and then ask whether it is possible to recover all s-sparse w (i.e. the collection of w which have at most s non-zero entries) exactly. This type of recovery is called *uniform recovery*. Asking whether or not uniform recovery is possible leads to the development of the *Restricted Isometry Property (RIP)* [9].

A. Uniform recovery and the RIP

More precisely, a matrix $U \in \mathbb{C}^{m\times n}$ is said to satisfy the RIP of order s with constant δ_s if

$$(1-\delta_s)\|w'\|_2^2 \le \|Uw'\|_2^2 \le (1+\delta_s)\|w'\|_2^2$$

for all s-sparse $w' \in \mathbb{C}^n$. If the matrix $P_\Omega MB$ satisfies the RIP of order s with a sufficiently small constant δ_s (e.g. $\delta_s < \frac{1}{3}$ as in [10]) then solutions w_1 to (1) satisfy

$$\|w_1 - w\|_1 \le C\sigma_s(w)_1 \tag{2}$$

$$\|w_1 - w\|_2 \le \frac{C}{\sqrt{s}}\sigma_s(w)_1 \tag{3}$$

where

$$\sigma_s(w)_1 := \min_{w' \text{ is } s\text{-sparse}} \|w - w'\|_1$$

and we say that *the RIP holds*.

All that remains is to prove that the examples of interest satisfy the RIP. Unfortunately, showing that a given matrix satisfies the RIP is an NP-hard problem [11]. In fact, it can be demonstrated easily that many practical matrices $P_\Omega MB$ do not satisfy the RIP. Indeed, one can perform the so-called *flip test* (taken from [8]) as follows:

1) Compute the wavelet coefficients w of a typical image.
2) Suppose that $P_\Omega MB$ satisfies the necessary conditions for (2) and (3). Use standard ℓ^1 minimization to find a solution to w_1 to (1).
3) Flip the coefficients of w to obtain $\tilde{w}$. Set $\tilde{w}_2$ to be a solution of (1) with $y = P_\Omega MB\tilde{w}$.
4) Flip the coefficients of $\tilde{w}_2$ to obtain w_2.
5) If (2) and (3) are even close to sharp, w_1 should approximate w in the same way that $\tilde{w}_2$ approximates $\tilde{w}$ (since $\sigma_s(w)_1 = \sigma_s(\tilde{w})_1$).
6) Therefore, $\|w_1 - w\| \approx \|\tilde{w}_2 - \tilde{w}\| = \|w_2 - w\|$.

Performing the flip test shows that uniform recovery does not hold in many cases (see Figure 3). In particular, for various sampling patterns using a Fourier-Wavelet matrix, the RIP does not hold. Therefore, the RIP is insufficient for describing successful compressed sensing.

Fig. 3. The flip test using various sampling patterns with $U = P_\Omega \mathrm{DFT}\,\mathrm{DWT}^{-1}$. The column CS reconstruction represents the standard reconstruction w_1 and CS flip represents the flipped reconstruction w_2.

B. Is uniform recovery necessary?

In the discussion that follows, we shall assume that $B = \mathrm{DWT}^{-1}$, although our arguments apply equally to other level based reconstruction basis. Let us now assume that we are trying to recover a vector w which represents the wavelet coefficients of an image. In this situation, the location of the non-zero values is extremely important. Indeed, typical images have larger and therefore more important entries in their coarser wavelet coefficients, whereas the finer wavelet coefficients are unlikely to contain important details.

Uniform recovery suggests that the location is unimportant, and that *all* s-sparse vectors can be recovered. The class of s-sparse wavelet coefficients is clearly much larger than the class of standard images and as such, it is unrealistic to expect recovery of all s-sparse vectors.

III. Mathematically Modelling CS

A. (s,M)-sparsity and the RIP in levels

In a level based reconstruction basis, it no longer makes sense to speak of s-sparsity. As discussed above, the location of the non-zero values is extremely important. In particular, we should expect relatively more sparsity in the finer levels. To model this situation, we can introduce (as in [8]) the concept of $(\mathbf{s},\mathbf{M})$-*sparsity.*

Definition 1. *For an* $l \in \mathbb{N}$, *which we call* the number of levels, *let* $\mathbf{s} = (s_1, s_2, \ldots, s_l) \in \mathbb{N}^l$ *and* $\mathbf{M} = (M_0, M_1, \ldots, M_l) \in \mathbb{N}^{l+1}$ *satisfy*

$$0 = M_0 < 1 \le M_1 < M_2 < \cdots < M_l$$

and $s_i \le M_i - M_{i-1}$. *A set* Λ *is said to be an* $(\mathbf{s},\mathbf{M})$-sparse set *if* $\Lambda \subset \{M_0+1, M_0+2, \ldots, M_l\}$ *and, for* $i = 1, 2, \ldots l,$

we have

$$|\Lambda \cap \{M_{i-1}+1, M_{i-1}+2, \ldots, M_i\}| \leq s_i.$$

We say that w is $(\mathbf{s}, \mathbf{M})$-sparse *if the support of w is an* $(\mathbf{s}, \mathbf{M})$-*sparse set.*

A careful choice of $\mathbf{M}$ allows us to say that w is $(\mathbf{s}, \mathbf{M})$-sparse if and only if it contains at most s_1 non-zero coefficients in the first wavelet level, s_2 in the second, s_3 in the third and so on. Even though the set of $(\mathbf{s}, \mathbf{M})$-sparse vectors is still larger than the set of wavelet coefficients that represent typical images, experimental evidence (see Section V) suggests that recovering all $(\mathbf{s}, \mathbf{M})$-sparse vectors is an attainable goal in the case where $M = \mathrm{DFT}^{-1}$, Ω is a collection of radial lines and $\mathbf{s}$ and $\mathbf{M}$ are chosen sensibly.

Now that we have a good concept of sparsity in a level based reconstruction basis, it is time to define an analogue to the RIP. The natural adaptation, which we term the *'RIP in levels'* (as described in [12]), is defined as follows:

Definition 2. *A matrix $U \in \mathbb{C}^{m\times n}$ is said to satisfy the RIP in levels (RIP_L) of order $(\mathbf{s}, \mathbf{M})$ with constant $\delta_{\mathbf{s},\mathbf{M}}$ if*

$$(1 - \delta_{\mathbf{s},\mathbf{M}})\|w'\|_2^2 \leq \|Uw'\|_2^2 \leq (1 + \delta_{\mathbf{s},\mathbf{M}})\|w'\|_2^2$$

for all $(\mathbf{s}, \mathbf{M})$-sparse vectors $w' \in \mathbb{C}^n$.

We shall see later that the RIP_L allows for recovery of all $(\mathbf{s}, \mathbf{M})$-sparse vectors.

B. The weighted RIP

In [13], an alternative to the RIP was described, termed the *'weighted RIP'*. Here, we do not consider s-sparse or $(\mathbf{s}, \mathbf{M})$-sparse vectors, but *weighted sparse* vectors. To be precise, given a vector $w \in \mathbb{C}^n$ and a vector of weights $\omega := (\omega_1, \omega_2, \ldots, \omega_n) \in \mathbb{R}^n$ with $\omega_j \geq 1$ for each j, we set $\|w\|_{\omega,0} := \sum_{j \in \mathrm{supp}(w)} \omega_j^2$. w is said to be (ω, s)-*sparse* if $\|w\|_{\omega,0} \leq s$. In the same way as with the RIP_L and the RIP, we can define a weighted RIP.

Definition 3. *A matrix $U \in \mathbb{C}^{m\times n}$ is said to satisfy the weighted RIP of order (ω, s) with constant $\delta_{\omega,s}$ if*

$$(1 - \delta_{\omega,s})\|w'\|_2^2 \leq \|Uw'\|_2^2 \leq (1 + \delta_{\omega,s})\|w'\|_2^2$$

for all (ω, s)-sparse vectors $w' \in \mathbb{C}^n$.

Weighted sparsity has been shown to be useful when analysing recovery of smooth functions from undersampled measurements [13]. It is thus natural to ask whether weighted sparsity also is a good model for recovery involving wavelets, curvelets, shearlets etc. As the following theorem and numerical example suggest, this may not be the case. We have demonstrated above that the class of sparse signals is too big to be able to explain the success of compressed sensing using X-lets. We now argue that the set of weighted sparse vectors is also too big.

Example 1 (Flip test with weighted sparsity). The flip test above is designed to reveal if the given sparsity model can explain the success of the sampling strategy Ω used. We can do exactly the same with weighted sparsity. Consider the successful recovery using the sampling pattern Ω in Figure 4, where we have recovered perfectly a sparse image (this image has only 1% non-zero wavelet coefficients). To reveal if this successful sampling pattern could also recover all weighted sparse vectors we do the following: Find the smallest s such that $\|w\|_{\omega,0} \leq s$ where w denotes the wavelet coefficients of the image. We then permute w to a vector $\tilde{w}$ such that $\|\tilde{w}\|_{\omega,0} \leq s$. Next, we try to recover $\tilde{w}$ via ℓ^1 optimisation, and finally we perform the inverse permutation on the recovered vector (similar to the flip test above). As shown in Figures 4 and 5 this gives a highly suboptimal result. The weights used were $1, 2^{d/2}, 2^{2d/2}, \ldots, 2^{ld/2}$ for all the indices corresponding to the levels $1, 2, \ldots, l$ respectively, where d is the dimension of the problem, however, other choices of weights give exactly the same phenomenon. Note that the result is the same if the ℓ^1 recovery is replaced by weighted ℓ^1 recovery. This is illustrated in Figure 5 where we have performed the same flip test for a one dimensional signal and a different sampling strategy.

Fig. 4. With the sampling pattern on the right, the leftmost image (with wavelet coefficients w) is recovered perfectly under a discrete fourier transform with DB2 wavelets. The image in the middle is the result of ℓ^1 recovery after applying a permutation to w preserving $\|\cdot\|_{\omega,0}$, and then permuting back. Similar issues are observed when ℓ^1 recovery is replaced by weighted ℓ^1 recovery.

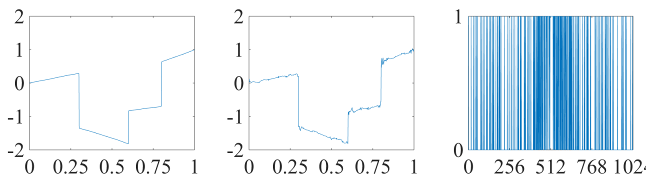

Fig. 5. With the sampling pattern on the right, the leftmost image (with wavelet coefficients w) is recovered perfectly under a discrete fourier transform with DB2 wavelets. The image in the middle is the result of weighted ℓ^1 recovery after applying a permutation to w preserving $\|\cdot\|_{\omega,0}$, and then permuting back. The rightmost image shows where samples were taken - a value of 1 indicates that a sample was taken at that fourier frequency.

Remark 1. If the previous flip test is changed to only permute within the levels, i.e. the $(\mathbf{s}, \mathbf{M})$-sparse structure is preserved rather than the weighted sparsity, then the recovery is perfect. See Section V.

The issue is that the class of weighted sparse vectors is too big and allows one to change the sparsity of the wavelet levels. The following theorem provides additional insight as to what goes wrong.

Theorem 1. *Suppose that* $\mathbf{s}, \mathbf{M}$ *is a sparsity pattern with l levels and that the set of* $(\mathbf{s}, \mathbf{M})$-*sparse vectors is contained in*

Fig. 6. A demonstration of the near block diagonality of typical CS matrices. From left to right, the images represent the absolute values of Fourier to DB2, Fourier to DB10 and Haar to Hadmard matrices.

the set of (ω, s)*-sparse vectors. Then there is an* l_0 *for which the set of* $(\mathbf{s}', \mathbf{M})$ *is contained in the* (ω, s)*-sparse vectors, where* $\mathbf{s}' := (\underbrace{0, \dots, 0}_{l_0 - 1}, ls_{l_0}, 0, \dots, 0)$.

Remark 2 (The consequences of Theorem 1). In many practical circumstances (such as Fourier or Hadamard to wavelets, see Figure 6), the matrix $P_\Omega MB \approx \bigoplus_{j=1}^{l} P_{\Omega_j} X_j$ for some complex matrices $X_j \in \mathbb{C}^{M_j - M_{j-1}}$ and projection matrices P_{Ω_j}. When this occurs, Theorem 1 suggests that Ω_j must be selected in such a way so that $P_{\Omega_{l_0}} X_{l_0}$ is capable of recovering all ls_{l_0}-sparse vectors. This requires a suboptimal number of measurements. On the other hand, to recover all $(\mathbf{s}, \mathbf{M})$-sparse vectors we would only require Ω_{l_0} to be chosen so that $P_{\Omega_{l_0}} X_{l_0}$ is capable of recovering all s_{l_0}-sparse vectors. Theorem 1 highlights another significant issue with the weighted RIP. Specifically, given w, it is possible to permute w in such a way that the number of non-zero coefficients in each level is changed. This is exactly what is going on in the above flip test.

C. Non-uniform recovery

The final mathematical description discussed here is to fix w and then randomly choose Ω. With a certain probability, good recovery of w will be possible. This is the theory of non-uniform recovery. A great deal of research has been done in this area. Notably, the *golfing scheme* (see [14] can be applied to generate probabilistic results on the number of measurements required for acceptable recovery. The results in this area have successfully been applied to schemes with a random Ω and the Fourier-Wavelet example above.

IV. Deterministic sampling in compressed sensing

It is clear that the RIP is an inadequate description of compressed sensing in the deterministic case, by applying the flip test as before. Moreover, the weighted RIP suffers from the same issues, albeit to a smaller degree. Additionally, the non-uniform recovery results rely on randomness. In our current situation, there is no randomness involved - we are simply fixing Ω and w. Consequently, we shall focus the remainder of our analysis on the RIP_L.

A. Recovery results using the RIP_L

If a matrix satisfies the RIP of order s with a sufficiently small constant δ_s, then uniform recovery of all s-sparse vectors is guaranteed. Similarly, one suspects that if the RIP_L constant of order $(\mathbf{s}, \mathbf{M})$ is sufficiently small (henceforth, if this condition is satisfied then a matrix is said to satisfy the RIP_L), then recovery of all $(\mathbf{s}, \mathbf{M})$-sparse vectors is guaranteed. Indeed, one can show the following result:

Theorem 2. *Let* $U \in \mathbb{C}^{m \times n}$*, and let* $(\mathbf{s}, \mathbf{M})$ *be a sparsity pattern satisfying the following:*

1) $M_l = n$ *where* l *is the number of levels.*
2) $s_i \neq 0$ *for each* $i = 1, 2, \dots l$.

Suppose that U *has* RIP_L *constant* $\delta_{2\mathbf{s},\mathbf{M}}$ *satisfying*

$$\delta_{2\mathbf{s},\mathbf{M}} < \frac{1}{\sqrt{l\left(\sqrt{\eta_{\mathbf{s},\mathbf{M}}} + \frac{1}{4}\right)^2 + 1}}$$

where $\eta_{\mathbf{s},\mathbf{M}} = \max_{1 \le i,j \le l} s_i / s_j$.

Let $w \in \mathbb{C}^n$ *and* $y \in \mathbb{C}^m$ *satisfy* $\|Uw - y\|_2 \le \epsilon$*. Then any* $\widetilde{w} \in \mathbb{C}^n$ *which satisfies both* $\|\widetilde{w}\|_1 \le \|w\|_1$ *and* $\|U\widetilde{w} - y\|_2 \le \epsilon$ *also satisfies*

$$\|w - \widetilde{w}\|_1 \le C_1 \sigma_{\mathbf{s},\mathbf{M}}(w)_1 + D_1 \sqrt{\tilde{s}}\,\epsilon \tag{4}$$

and

$$\|w - \widetilde{w}\|_2 \le \frac{\sigma_{\mathbf{s},\mathbf{M}}(w)_1}{\sqrt{\tilde{s}}} \left(C_2 + C_2' \sqrt[4]{l \eta_{\mathbf{s},\mathbf{M}}}\right) + \epsilon \left(D_2 + D_2' \sqrt[4]{l \eta_{\mathbf{s},\mathbf{M}}}\right) \tag{5}$$

where

$$\sigma_{\mathbf{s},\mathbf{M}}(w)_1 := \min \|w - \widehat{w_2}\|_1 \text{ such that } \widehat{w_2} \text{ is } (\mathbf{s}, \mathbf{M})\text{-sparse},$$

$\tilde{s} = s_1 + s_2 + s_3 + \dots + s_l$ *and* C_1, C_2, C_2', D_1, D_2 *and* D_2' *depend only on* $\delta_{2\mathbf{s},\mathbf{M}}$*. In particular,* (4) *and* (5) *also hold for solutions to*

$$\min_{w' \in \mathbb{C}^n} \|w'\|_1 \text{ subject to } \|Uw' - y\|_2 \le \epsilon.$$

The change from solving $Uw' = y$ to solving $\|Uw' - y\|_2 \le \epsilon$ allows us to cover the case where the measurements we obtain are no longer given by Uw but instead by $Uw + \gamma$ where γ is a noise term with $\|\gamma\| \le \epsilon$. Similar results that account for noise can be obtained for the standard RIP. It is worth noting that although Theorem 2 reduces to a standard result on the RIP in the case that $l = 1$, the requirement on $\delta_{\mathbf{s},\mathbf{M}}$ involves the number of levels l and a parameter $\eta_{\mathbf{s},\mathbf{M}}$. Further work in [12] has shown that this requirement cannot be avoided.

V. Numerical results

It is now pertinent to ask whether or not recovery of all $(\mathbf{s}, \mathbf{M})$-sparse vectors is a realistic expectation of compressed sensing on a level based reconstruction problem. To give experimental evidence suggesting that this is the case, one can adapt the flip test to a 'permutations in levels' test.

1) As before, we compute the wavelet coefficients w of a typical image.
2) Suppose that $P_\Omega MB$ satisfies the necessary conditions for (4) and (5). Use standard ℓ^1 minimization to find a solution to w_1 to (1).

TABLE I
THE FLIP IN LEVELS TEST PERFORMED USING RADIAL LINES.

Maximum relative error	Minimum relative error	Standard deviation	Number of lines
2.16%	2.14%	0.0025%	128
1.03%	1.02%	0.0011%	256

3) Let Q be a permutation satisfying

$$Q(\{M_{i-1}+1,\dots,M_i\}) = \{M_{i-1}+1,\dots,M_i\} \quad (6)$$

for $i=1,2,\dots,l$. Apply the permutation Q to w to obtain $\tilde{w}$. Set $\tilde{w}_2$ to be a solution to (1) with $y = P_\Omega MB\tilde{w}$.

4) Apply the inverse permutation Q^{-1} to $\tilde{w}_2$ to obtain w_2.

5) If (2) and (3) are even close to sharp, w_1 should approximate w in the same way that $\tilde{w}_2$ approximates $\tilde{w}$ (since $\sigma_{\mathbf{s},\mathbf{M}}(w)_1 = \sigma_{\mathbf{s},\mathbf{M}}(\tilde{w})_1$).

6) Therefore, $\|w_1 - w\| \approx \|w_2 - w\| = \|\tilde{w}_2 - \tilde{w}\|$.

Of course, it is far from practical to do this test for every single possible permutation. However, for any choice of Q satisfying (6) we certainly require $\|w_1 - w\|$ to be very close to $\|w_2 - w\|$ for the RIP_L to hold. Conversely, if the test succeeds using a large collection of randomly generated Q on a fixed image w with a fixed matrix $P_\Omega MB$ then one would suspect that equations (4) and (5) hold, in turn providing evidence that the RIP_L holds for $P_\Omega MB$. We performed this test with fifty randomly chosen permutations on the radial lines image from Figure 3. The results are displayed in I. The small standard deviations and differences between the maximum and minimum relative error strongly suggest that equations (4) and (5) hold when $U = P_\Omega \,\mathrm{DFT}\,\mathrm{DWT}^{-1}$, with Ω taken to be a sufficiently large collection of radial lines. This will certainly occur when U satisfies the RIP_L.

VI. PERTURBATIONS AND THE RIP IN LEVELS

It may however be the case that $P_\Omega \,\mathrm{DFT}\,\mathrm{DWT}^{-1}$ does not satisfy the RIP_L, but instead approximates a matrix that does. In this case, let us set $U = P_\Omega \,\mathrm{DFT}\,\mathrm{DWT}^{-1}$ and suppose that $U = R + E$ where $R \in \mathbb{C}^{m\times n}$ satisfies the RIP_L and $E \in \mathbb{C}^{m\times n}$ is a perturbation matrix such that $\|E\|_2$ is small. We can then prove the following result:

Theorem 3. *Suppose that the sparsity pattern* $(\mathbf{s},\mathbf{M})$ *satisfies the following two conditions:*

1) $M_l = n$ *where* l *is the number of levels.*
2) $s_i \neq 0$ *for each* $i = 1,2,\dots l$.

Furthermore, suppose that R *has* RIP_L *constant* $\delta_{2\mathbf{s},\mathbf{M}}$ *satisfying*

$$\delta_{2\mathbf{s},\mathbf{M}} < \frac{1}{\sqrt{l\left(\sqrt{\eta_{\mathbf{s},\mathbf{M}}}+\frac{1}{4}\right)^2+1}}.$$

Let $w \in \mathbb{C}^n$ *and* $y \in \mathbb{C}^m$ *satisfy* $\|Uw - y\|_2 \le \epsilon$. *Then any* $\widetilde{w} \in \mathbb{C}^n$ *which solves the* ℓ^1 *minimization problem*

$$\min_{w' \in \mathbb{C}^n} \|w'\|_1 \text{ subject to } \|Uw' - y\|_2 \le \epsilon$$

also satisfies

$$\|w - \widetilde{w}\|_1 \le C_1 \sigma_{\mathbf{s},\mathbf{M}}(x)_1 + K_1\sqrt{\tilde{s}}\,(\epsilon + \|E\|\|w\|_1) \quad (7)$$

and

$$\|w - \widetilde{w}\|_2 \le \frac{\sigma_{\mathbf{s},\mathbf{M}}(w)_1}{\sqrt{\tilde{s}}}\left(C_2 + C_2'\sqrt[4]{l\eta_{\mathbf{s},\mathbf{M}}}\right) + (\|E\|\|w\| + \epsilon)(K_2 + K_2'\sqrt[4]{l\eta_{\mathbf{s},\mathbf{M}}}) \quad (8)$$

where C_1, C_2, C_2' *are the same as in Theorem 2 and* K_1, K_2 *and* K_2' *depend only on* $\delta_{2\mathbf{s},\mathbf{M}}$.

It is straightforward to prove this Theorem using Theorem 2, but we shall not provide these details here.

Following this perturbation result, we claim that in the radial line case, $P_\Omega \,\mathrm{DFT}\,\mathrm{DWT}^{-1}$ satisfies the RIP_L, or $P_\Omega \,\mathrm{DFT}\,\mathrm{DWT}^{-1}$ is a slight perturbation of a matrix that satisfies the RIP_L whenever sufficiently many radial lines are taken. A proof of this conjecture along with an estimate on the required number of radial lines would provide a strong result in deterministic sampling, but as of yet it remains unsolved.

REFERENCES

[1] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[2] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

[3] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok, "Introduction to compressed sensing," in *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2011.

[4] Y. C. Eldar and G. Kutyniok, Eds., *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[5] M. Fornasier and H. Rauhut, "Compressive sensing," in *Handbook of Mathematical Methods in Imaging*. Springer, 2011, pp. 187–228.

[6] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing*, ser. Applied and Numerical Harmonic Analysis. Birkhäuser/Springer, New York, 2013. [Online]. Available: http://dx.doi.org/10.1007/978-0-8176-4948-7

[7] R. G. Baraniuk, V. Cevher, M. F. Duarte, and C. Hegde, "Model-based compressive sensing," *IEEE Trans. Inform. Theory*, vol. 56, no. 4, pp. 1982–2001, 2010. [Online]. Available: http://dx.doi.org/10.1109/TIT.2010.2040894

[8] B. Adcock, A. C. Hansen, C. Poon, and B. Roman, "Breaking the coherence barrier: A new theory for compressed sensing." *CoRR*, vol. abs/1302.0561, 2013.

[9] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inform. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005. [Online]. Available: http://dx.doi.org/10.1109/TIT.2005.858979

[10] T. T. Cai and A. Zhang, "Sharp RIP bound for sparse signal and low-rank matrix recovery," *Appl. Comput. Harmon. Anal.*, vol. 35, no. 1, pp. 74–93, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.acha.2012.07.010

[11] A. M. Tillmann and M. E. Pfetsch, "The computational complexity of the restricted isometry property, the nullspace property, and related concepts in compressed sensing," *IEEE Trans. Inform. Theory*, vol. 60, no. 2, pp. 1248–1259, 2014. [Online]. Available: http://dx.doi.org/10.1109/TIT.2013.2290112

[12] A. Bastounis and A. C. Hansen, "On the absence of the rip in real-world applications of compressed sensing and the rip in levels," *CoRR*, vol. abs/1411.4449, 2014.

[13] H. Rauhut and R. Ward, "Interpolation via weighted l1 minimization," *CoRR*, vol. abs/1308.0759, 2014.

[14] D. Gross, "Recovering low-rank matrices from few coefficients in any basis," *IEEE Trans. Inform. Theory*, vol. 57, no. 3, pp. 1548–1566, March 2011.