Sparse approximation methods for the recovery of signals from undersampled data when the signal is sparse in an overcomplete dictionary have received much attention recently due to their practical importance. A common assumption is the $D$-restricted isometry property ($D$-RIP), which asks that the sampling matrix approximately preserve the norm of all signals sparse in $D$. While many classes of random matrices satisfy this condition, those with a fast-multiply stemming from subsampled bases require an additional randomization of the column signs, which is not feasible in many practical applications. In this work, we demonstrate that one can subsample certain bases in such a way that the $D$-RIP will hold without the need for random column signs.

I. INTRODUCTION

The compressed sensing paradigm, as first introduced by Candès and Tao [10] and Donoho [12], is based on the idea of using available degrees of freedom in a sensing mechanism to tune the measurement system so that it allows for efficient recovery of the underlying signal or image from undersampled measurement data. A model assumption that allows for such signal recovery is sparsity; the signal can be represented by just a few elements of a given representation system.

Often it is near-optimal to choose the measurements completely at random, for example following a Gaussian distribution [31]. Typically, however, some additional structure is imposed by the application at hand, and randomness can only be infused in the remaining degrees of freedom. For example, magnetic resonance imaging (MRI) is known to be well modeled by inner products with Fourier basis vectors. This structure cannot be changed, and the only aspect that can be decided at random is which of them are selected.

An important difference between completely random measurement systems and many structured random measurement systems is in the aspect of universality. Gaussian measurement systems and many other systems without much imposed structure yield similar reconstruction quality for all different orthonormal basis representations. For structured measurement systems, this is, in general, no longer the case. When the measurements are uniformly subsampled from an orthonormal basis such as the Fourier basis, one requires, for example that the measurement basis and the sparsity basis are incoherent; the inner products between vectors from the two bases are small. If the two bases are not incoherent, more refined concepts of incoherence are required. An important example is that of the Fourier measurement basis and a wavelet sparsity basis. Both contain the constant vector and are hence maximally coherent.

All of the above mentioned works, however, exclusively cover the case of sparsity on orthonormal basis representations. On the other hand, there are a vast number of applications in which sparsity is expressed not in terms of a basis but in terms of a redundant, often highly overcomplete, dictionary. Specifically, if $f$ is the signal of interest, then one expresses $f = Dx$, where $D \in \mathbb{C}^{n \times N}$ is an arbitrary overcomplete dictionary and $x \in \mathbb{C}^N$ is a sparse (or nearly sparse) coefficient vector. Redundancy is used widespread in practice either because no sparsifying orthonormal basis exists for the signals of interest, or because the redundancy itself is useful and allows for a significantly larger class of signals with significantly more sparsity. For example, it has been well-known that overcompleteness is the key to a drastic reduction in artifacts and recovery error in the denoising framework [32], [33].

Results using overcomplete dictionaries in the compressed sensing framework are motivated by the broad array of tight frames appearing in practical applications. For example, if one utilizes sparsity with respect to the Discrete Fourier Transform (DFT), then one implicitly assumes that the signal is well represented by frequencies along the lattice. To allow for more flexibility in this rigid assumption, one instead may employ the oversampled DFT, containing frequencies on a much finer grid, or even over intervals of varying widths. Gabor frames are used in imaging as well as radar and sonar applications, which are often highly redundant [25]. There are also many frames used in imaging applications such as undecimated wavelet frames [13], [32], curvelets [6], shearlets [23], [14], framelets [5], and many others, all of which are overcomplete with highly correlated columns.

II. STRUCTURED COMPRESSED SENSING WITH REDUNDANT DICTIONARIES

Recall that a dictionary $D \in \mathbb{C}^{n \times N}$ is a Parseval frame if $DD^* = I_n$, and that a vector $x$ is $s$-sparse if $\|x\|_0 := |\text{supp}(x)| \leq s$. Then the restricted isometry property with respect to a dictionary $D$ ($D$-RIP) is defined as follows.
Definition 1 ([7]). Fix a dictionary $D \in \mathbb{C}^{n \times N}$ and matrix $A \in \mathbb{C}^{m \times n}$. The matrix $A$ satisfies the $D$-RIP with parameters $\delta$ and $s$ if
\[
(1 - \delta)\|Ax\|_2^2 \leq \|ADx\|_2^2 \leq (1 + \delta)\|Dx\|_2^2
\]
for all $s$-sparse vectors $x \in \mathbb{C}^N$.

Note that when $D$ is the identity, this definition reduces to the standard definition of the restricted isometry property [9]. Under such an assumption on the measurement matrix, the following results bound the reconstruction error for the $\ell_1$-analysis method. This method consists of estimating a signal $f$ from noisy measurements $y = Af + e$ by solving the convex minimization problem
\[
\hat{f} = \arg\min_{\tilde{f} \in \mathbb{C}^N} \|D^* \tilde{f}\|_1 \text{ such that } \|A\tilde{f} - y\|_2 \leq \varepsilon, \quad (P_1)
\]
where $\varepsilon$ is the noise level, that is, $\|e\|_2 \leq \varepsilon$.

The $\ell_1$-analysis method (like alternative approaches) is based on the model assumption that for a signal $f = Dx$ not only the underlying (synthesis) coefficient sequence $x$ (which is typically unknown and hard to obtain), but also the analysis coefficients $D^* f$ are compressible, i.e., dominated by a few large entries. The assumption has been observed empirically for many dictionaries and also forms the basis for a number of thresholding approaches to signal denoising [6], [22], [14].

More precisely, Theorem 2 gives the precise formulation of the resulting recovery guarantees (as derived in [7]).

Theorem 2 ([7]). Let $D$ be a tight frame, $\varepsilon > 0$, and consider a matrix $A$ that has the $D$-RIP with parameters $2s$ and $\delta < 0.08$. Then for every signal $f \in \mathbb{C}^N$, the reconstruction $\hat{f}$ obtained from noisy measurements $y = Af + e$, $\|e\|_2 \leq \varepsilon$, via the $\ell_1$-analysis problem $(P_1)$ satisfies
\[
\|f - \hat{f}\|_2 \leq \varepsilon + \frac{\|D^* f - (D^* f)_s\|_1}{\sqrt{s}}.
\]

where $x_s$ denotes the best $s$-sparse approximation to $x$, $x_s = \inf \|z\|_0 : \|x - z\|_2 \leq \varepsilon$.

Thus one obtains recovery guarantees for signals $f$ whose analysis coefficients $D^* f$ have a suitable decay. This is the case for many natural dictionaries used in practice such as the Gabor frame, undecimated wavelets, curvelets, etc. (see e.g. [7] for a detailed description of such frames).

The results in this paper need a similar, but somewhat weaker assumption to hold for all signals corresponding to sparse synthesis coefficients $x$, namely one needs to control the localization factor as introduced in the following definition.

Definition 3. For a dictionary $D \in \mathbb{C}^{n \times N}$ and a sparsity level $s$, we define the localization factor as
\[
\eta_{s, D} = \eta = \sup_{\|Dz\|_2 = 1, \|z\|_0 \leq s} \frac{\|D^* Dz\|_1}{\sqrt{s}}.
\]

We use the term localization factor here because the quantity can be viewed as a measure of how compressible the objective in $(P_1)$ is. Note that if $D$ is orthonormal, then $\eta = 1$, and $\eta$ increases with the redundancy in $D$. Moreover, guarantees like (2) give a recovery guarantee of $Dz$ up to the scaled $\ell_1$ norm of the tail of $D^* Dz$ for any sparse vector $z$. Precisely, the right most term of (2) is analogous to a single instance of (3) up to an additive factor of 1. We believe that assuming that the localization factor is bounded or at least growing not too fast is a reasonable assumption and pose it as a problem for subsequent work to obtain estimates for this quantity for dictionaries of practical interest, discussing also a few simple examples below. To illustrate that our results go strictly beyond existing theory, we will show that harmonic frames with small redundancy indeed have a bounded localization factor. To our knowledge, this case is not covered by existing theories. In addition, we also bound the localization factor for certain redundant Haar wavelet frames.

For an extension of our result to weighted sparse expansions and for a more detailed discussion on variable density sampling, see [21].

A. The D-RIP revisited

When $D$ is the identity basis, the $D$-RIP of Definition 1 reduces to the classical restricted isometry property (RIP) first introduced in [9]. It is now well-known that any $m \times n$ matrix whose entries are i.i.d. subgaussian random variables will satisfy the RIP with a small constant $0 < \delta < c < 1$ with high probability as long as $m$ is on the order of $s \log(n/s)$ (see e.g. [26], [31]). In order to use a fast-multiply to obtain the measurements, one desires to form $A$ by subsampling a structured matrix.

Both the RIP and the $D$-RIP are closely related to the Johnson-Lindenstrauss lemma [16], [1], [2]. Indeed, any matrix $A \in \mathbb{C}^{m \times n}$ which for a fixed vector $z \in \mathbb{C}^n$ satisfies
\[
P(\|Az\|^2 - \|z\|^2 \geq \delta\|z\|^2) \leq Ce^{-Cm}
\]
will satisfy the $D$-RIP with high probability as long as $m$ is at least on the order of $s \log(N/s)$ [29]. From this, any matrix satisfying the Johnson-Lindenstrauss lemma will also satisfy the $D$-RIP. Random matrices known to have this property include matrices with independent subgaussian entries (such as Gaussian or Bernoulli matrices), see for example [11]. Moreover, it is shown in [19] that any matrix that satisfies the classical RIP will satisfy the Johnson-Lindenstrauss lemma and thus the $D$-RIP with high probability after randomizing the signs of the columns. The latter construction allows for structured random matrices such as randomly subsampled Fourier matrices (in combination with the results from [31]) and matrices representing subsampled random convolutions (in combination with the results from [28], [17]); in both cases, however, again with randomized column signs. While this gives an abundance of such matrices, as mentioned above, it is not always practical or possible to apply random column signs in the sampling procedure.

An important general set-up of structured random sensing matrices known to satisfy the regular RIP is the framework of bounded orthonormal systems, which includes as a special case subsampled discrete Fourier transform measurements (without column signs randomized). Such measurements are the only type of measurements possible in many physical systems.
where compressive sensing is of interest, such as in MRI, radar, and astronomy [24], [3], [15], [4]. We recall this set-up (in the discrete setting) below.

**Definition 4 (Bounded orthonormal system).** Consider a probability measure \( \nu \) on the discrete set \([n]\) and a system \( \{a_j \in \mathbb{C}^n, j \in [n]\} \) that is orthonormal with respect to \( \nu \) in the sense that
\[
\sum_i a_k(i) a_j(i) \nu_i = \delta_{j,k}
\]
where \( \delta_{j,k} = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{else} \end{cases} \) is the Kronecker delta function.

Suppose further that the system is uniformly bounded: there exists a constant \( K \geq 1 \) such that
\[
\sup_{i \in [n]} \sup_{j \in [n]} |a_j(i)| \leq K.
\]

Then the matrix \( \Psi \in \mathbb{C}^{n \times n} \) whose rows are indexed by \( a_j \) is called a bounded orthonormal system.

Drawing \( m \) indices \( i_1, i_2, \ldots, i_m \) independently from the orthogonalization measure \( \nu \), the sampling matrix \( A \in \mathbb{C}^{m \times n} \) whose rows are indexed by the (re-normalized) sampled vectors \( \sqrt{\nu} a_{i_k}(k) \in \mathbb{C}^n \) will have the restricted isometry property with high probability (precisely, with probability exceeding \( 1 - n^{-C \log^3(s)} \)) provided the number of measurements satisfies
\[
m \geq CK^2 s \log^3(s) \log(n).
\]

This result was first shown in the case where \( \nu \) is the uniform measure by Rudelson and Vershynin [31], for a slightly worse dependence of \( m \) on the order of \( s \log^2 s \log(s \log n) \log(n) \). These results were subsequently extended to the general bounded orthonormal system set-up by Rauhut [27], and the dependence of \( m \) was slightly improved to \( s \log^3 s \log n \) in [30].

An important special case where these results can be applied is that of incoherence between the measurement and sampling bases. Here the coherence of two sets of vectors \( A = \{a_i\} \) and \( B = \{b_j\} \) is given by \( \mu = \sup_{i,j} |\langle a_i, b_j \rangle| \). Two orthonormal bases \( A \) and \( B \) of \( \mathbb{C}^n \) are called incoherent if \( \mu \leq Kn^{-1/2} \). In this case, the renormalized system \( \tilde{A} = \{\sqrt{\nu_a} B^*\} \) is an orthonormal system with respect to the uniform measure, which is bounded by \( K \). Then the above results imply that signals which are sparse in basis \( B \) can be reconstructed from inner products with a uniformly subsampled subset of basis \( A \). These incoherence-based guarantees are a standard criterion to ensure signal recovery, as first observed in [8].

**B. New results**

Our main result shows that one can subsample certain bounded orthonormal matrices to obtain a matrix that has the \( D \)-RIP and hence yields recovery guarantees for sparse recovery in redundant dictionaries, as we will summarize in this subsection. As indicated above, our technical estimates below will imply such guarantees for various algorithms. As an example we focus on the method of \( \ell_1 \)-analysis, for which the first \( D \)-RIP based guarantees are available [7], while noting that our technical estimates below imply similar guarantees also for the other algorithms referenced above. The proof of the following theorem can be found in [18].

**Theorem 5.** Fix a probability measure \( \nu \) on \([n]\), sparsity level \( s < N \), constant \( 0 < \delta < 1 \), and let \( D \in \mathbb{C}^{n \times N} \) be a Parseval frame. Let \( A \) be an orthonormal systems matrix with respect to \( \nu \) as in Definition 4, and assume the incoherence satisfies,
\[
\max_j |\langle a_j, d_j \rangle| \leq K.
\]

Define the localization factor \( \eta = \eta_{D,s} \) as in Definition 3. Construct an \( m \times n \) submatrix \( \tilde{A} \) of \( A \) by sampling rows of \( A \) according to the measure \( \nu \). Then as long as
\[
m \geq C\delta^{-2} s \eta^2 \log^3(\eta^2) \log(N), \quad \text{and}
\]
\[
m \geq C\delta^{-2} s \eta^2 \log(1/\gamma)
\]
then with probability \( 1 - \gamma \), the normalized submatrix \( \sqrt{\frac{m}{N}} \tilde{A} \) satisfies the \( D \)-RIP with parameters \( [s/K^2] \) and \( \delta \).

Utilizing Theorems 5 and 2 we obtain the following corollary for signal recovery with redundant dictionaries using subsampled structured measurements.

**Corollary 6.** Fix a sparsity level \( s < N \). Let \( D \in \mathbb{C}^{n \times N} \) be a Parseval frame – i.e., \( DD^* \) is the identity – with columns \( \{d_1, \ldots, d_N\} \), and let \( A = \{a_1, \ldots, a_n\} \) be an orthonormal basis of \( \mathbb{C}^n \) that is incoherent to \( D \) in the sense that
\[
\sup_{i \in [n]} \sup_{j \in [n]} |\langle a_i, d_j \rangle| \leq Kn^{-1/2}
\]
for some \( K \geq 1 \).

Consider the localization factor \( \eta \) as in Definition 3. Construct \( \tilde{A} \in \mathbb{C}^{m \times n} \) by sampling vectors from \( A \) i.i.d. uniformly at random. Then as long as
\[
m \geq CsK^2 \eta^2 \log^3(\eta^2) \log(N),
\]
then with probability \( 1 - N^{-\log^3 s} \), \( \sqrt{\frac{m}{N}} \tilde{A} \) exhibits uniform recovery guarantees for \( \ell_1 \)-analysis. That is, for every signal \( f \), the solution \( \tilde{f} \) of the minimization problem (\( P_f \)) with \( y = \sqrt{\frac{m}{N}} \tilde{A} f + e \) for noise \( e \) with \( ||e||_2 \leq \varepsilon \) satisfies
\[
\|f - \tilde{f}\|_2 \leq C_1 \varepsilon + C_2 \frac{\|D^* f - (D^* f)_{\delta}\|_S}{\sqrt{S}}.
\]

Here \( C_1, C_2 \) and \( s \) are absolute constants independent of the dimensions and the signal \( f \).

**Proof.** Let \( A, D \) and \( s \) be given as in Corollary 6. We will apply Theorem 5 with \( K = 1/\sqrt{2} \), \( \nu \) the uniform measure on \([n]\), \( \gamma = N^{-\log^3 s} \), \( \delta = 0.08 \), and matrix \( \sqrt{n}A \). We first note that since \( A \) is an orthonormal basis, that \( \sqrt{n}A \) is an orthonormal systems matrix with respect to the uniform measure \( \nu \) as in Definition 4. In addition, (8) implies (6) for matrix \( \sqrt{n}A \) with \( K = 1/\sqrt{2} \). Furthermore, setting \( \gamma = N^{-\log^3 s} \), (9) implies both inequalities of (7) (adjusting constants appropriately). Thus the assumptions of Theorem 5 are in force. Theorem 5 then guarantees that with probability \( 1 - \gamma = 1 - N^{-\log^3 s} \), the uniformly subsampled matrix \( \sqrt{\frac{m}{N}}(\sqrt{n}A) \) satisfies the \( D \)-RIP with parameters \( 2s \) and \( \delta \). This shows that this implies the \( D \)-RIP with parameters \( 2s \) and \( \delta \). By Theorem 2, (10) holds and this completes the proof. \( \square \)
III. SOME EXAMPLE APPLICATIONS

A. Harmonic frames

It remains to find examples of a dictionary with bounded localization factor and an associated measurement system for which incoherence condition (8) holds. Our main example is that of sampling a signal that is sparse in an oversampled Fourier system, a so-called harmonic frame [34]; the measurement system is the standard basis. Indeed, one can see by direct calculation that the standard basis is incoherent in the sense of (8) to any set of Fourier vectors, even of non-integer frequencies. We will now show that if the number of frame vectors exceeds the dimension only by a constant, such a dictionary will also have bounded localization factor. This setup is a simple example, but our results apply, and it is not covered by previous theory.

More precisely, we fix $L \in \mathbb{N}$ and consider $N = n + L$ vectors in dimension $n$. We assume that $L$ is such that $L s \leq \frac{n}{4}$. Then the harmonic frame is defined via its frame matrix $D = (d_{jk})$, which results from the $N \times N$ discrete Fourier transform matrix $F = (f_{jk})$ by deleting the last $L$ rows. That is, we have $d_{jk} = \frac{1}{\sqrt{n+L}} \exp\left(\frac{2\pi i j k}{n+L}\right)$ for $j = 1 \ldots n$ and $k = 1 \ldots N$. The corresponding Gram matrix satisfies $(D^* D)_{ii} = \frac{n}{n+L}$ and, for $i \neq j$, by orthogonality of $F$, $(D^* D)_{jk} = \left|\frac{1}{\sqrt{n+L}} (F^* F)_{jk} - \sum_{\ell \in \{n+1\}^L} \hat{f}_j \hat{f}_{\ell k}\right| \leq \frac{L}{n+L}$.

As a consequence, we have for $z$ $s$-sparse with $S = \text{supp} z$, $(D^* D z)_j = \sum_{k \in S} (D^* D)_{jk} z_k \leq \frac{L}{n+L} \|z\|_1$, $j \notin S$.

So we obtain, using that $D^*$ is an isometry, $\|Dz\|_2^2 = \|D^* Dz\|_2^2 \geq \sum_{k \in S} (D^* Dz)_k^2 \geq \sum_{k \in S} (|z_k| - |(D^* D)_{jk} z_k|)^2 \geq \sum_{k \in S} |z_k|^2 - 2|z_k| \frac{L}{n+L} \|z\|_1 = \|z\|_2^2 - 2\|z\|_2 \frac{L}{n+L} \|z\|_1 \geq (1 - \frac{2L}{n+L}) \|z\|_2^2 \geq \frac{1}{2} \|z\|_2^2$.

That is, for $z$ with $\|Dz\|_2 = 1$, one has $\|z\|_2 \leq \sqrt{2}$. Consequently, $\eta = \sup_{\|Dz\|_2 = 1, \|z\|_0 \leq s} \frac{\|D^* Dz\|_1}{\sqrt{s}} = \sup_{\|Dz\|_2 = 1, \|z\|_0 \leq s} \frac{\|D^* Dz\|_1 + \|D^* Dz\|_{S^c}}{\sqrt{s}} \leq \sup_{\|Dz\|_2 = 1, \|z\|_0 \leq s} \frac{\|D^* Dz\|_2 + \|D^* Dz\|_{S^c}}{\sqrt{s}} = 1 + \frac{1}{\sqrt{s}} \|z\|_{S^c} \leq 1 + \frac{1}{\sqrt{s}} \frac{\|z\|_{S^c}}{\|z\|_2} \leq 1 + L \sqrt{2}$.

B. Fourier measurements and redundant Haar frames

We next present a second example of a sampling setup that satisfies the assumptions of incoherence and localization factor of Theorem 5. In contrast to the previous example, one needs to precondition and adjust the sampling density to satisfy these assumptions according to Theorem 5, which allows us to understand the setup of the Fourier measurement basis and a 1D Haar wavelet frame with redundancy 2, as introduced in the following. Let $n = 2^p$. Recall that the univariate discrete Haar wavelet basis of $\mathbb{C}^{2^n}$ consists of $h_0 = 2^{-p/2} (1, 1, \ldots, 1)$, $h = h_{00} = 2^{-p/2} (1, 1, \ldots, 1, -1, -1, \ldots, -1)$ and the frame basis elements $h_{\ell,k}$ given by $h_{\ell,k}(j) = h(2^\ell j - k)$

\[
\begin{cases}
2^{-\frac{\ell p}{2}} & \text{for } k 2^{p-\ell} \leq j < k 2^{p-\ell} + 2^{p-\ell-1} \\
-2^{-\frac{\ell p}{2}} & \text{for } k 2^{p-\ell} + 2^{p-\ell-1} \leq j < k 2^{p-\ell} + 2^{p-\ell} \\
0 & \text{else,}
\end{cases}
\]

for any $(\ell, k) \in \mathbb{Z}^2$ satisfying $0 \leq \ell < p$ and $0 \leq k < 2^\ell$. The corresponding basis transformation matrix is denoted by $H$.

One can now create a wavelet frame of redundancy 2 by considering the union of this basis and a circular shift of it by one index. That is, one adds the vector $\hat{h}^0 = h^0 = 2^{-p/2} (1, 1, \ldots, 1)$ and vectors of the form $h_{\ell,k}(j) = h_{\ell,k}(j + 1)$ for all $(\ell, k)$. Here we identify $2^p + 1 = 1$. This is also an orthonormal basis – its basis transformation matrix will be denoted by $H$ in the following, and the matrix $D \in \mathbb{C}^{2^p \times 2^{p+2}}$ with columns

\[
D(:, (\ell, k, 1)) = \frac{1}{\sqrt{2}} h_{\ell,k}, \\
D(:, (\ell, k, 2)) = -\frac{1}{\sqrt{2}} \hat{h}_{\ell,k}
\]

forms a Parseval frame with redundancy 2.

Theorem 5 applies to the example where sparsity is with respect to the redundant Haar frame and where sampling measurements are rows $\{a_k\}$ from the $n \times n$ orthonormal DFT matrix. We will utilize Theorem 5 with measure $\nu(k) = \kappa_k^2 / \|a_k\|_2^2$ where $\sup_{1 \leq k \leq n} \|a_k\| \leq \kappa_k$. Then note that if the diagonal
matrix \( W = \text{diag}(w) \in \mathbb{C}^{n \times n} \) with \( w_k = \|y_k\|_2 / \kappa_k \), that the renormalized system \( \Phi = \frac{1}{m} W A \) is an orthonormal system with respect to \( \nu \), bounded by \( \|x\|^2 \). Following Corollary 6.4 of [20], we have the following coherence estimates:

\[
\max_{i,j} \left| \left\langle a_i, \hat{h}_{j,i} \right\rangle \right| \leq \kappa_k = \frac{3\sqrt{2\pi}}{\sqrt{k}}.
\]

Since \( \|x\|^2 = 18\pi \sum_{k=1}^n k^{-1} \leq 18\pi \log_2(n) \) (the bound holding for large enough \( n \)) grows only mildly with \( n \), the Fourier / wavelet frame example is a good fit for Theorem 5, provided the localization factor of the Haar frame is also small. One can show that the localization factor of the Haar frame is bounded by \( \eta \leq 2\sqrt{\log_2(n)} \), leading to the following corollary, whose derivation details can be found in [18].

**Corollary 7.** Fix a sparsity level \( s < N \), and constant \( 0 < \delta < 1 \). Let \( D \in \mathbb{C}^{n \times N} \) be the redundant Haar frame as defined in (11) and let \( A \in \mathbb{C}^{n \times n} \) with rows \( \{a_1, \ldots, a_n\} \) be the orthonormal DFT matrix. Construct \( \tilde{A} \in \mathbb{C}^{n \times n} \) by sampling rows from \( A \) i.i.d. from the probability measure \( \nu \) on \( [n] \) with power-law decay \( \nu(k) = Ck^{-\delta} \).

As long as the number of measurements satisfies

\[
m \geq \delta^{-2}8\log^3(s \log(n) \log^3(n)) ,
\]

then with probability \( 1 - n^{-\log^3 s} \), the following holds for every signal \( f \): the solution \( \hat{f} \) of the \( l_1 \)-analysis problem (P1) with \( y = \tilde{A} f + e \) for noise \( e \) with error \( \|e\|_2 \leq \varepsilon \) satisfies

\[
\|f - \hat{f}\|_2 \leq C_1 \varepsilon + C_2 \frac{\|D^* f - (D^* f)_s\|_1}{\sqrt{s}}.
\]

Here, \( C_1 \) and \( C_2 \) are absolute constants independent of the dimensions and the signal \( f \).

**ACKNOWLEDGMENT**

The authors would like to thank the Institute for Computational and Experimental Research in Mathematics (ICERM) for its hospitality during a stay where this work was initiated.

In addition, Krahmer was supported by the German Science Foundation (DFG) in the context of the Emmy Noether Junior Research Group KR 4512/1-1 (RaSenQuaSI). Needell was partially supported by an NSF Career award, DOD-Navy grant N00014-12-1-0743, and an AFOSR Young Investigator Program award.

**REFERENCES**


