Optimal Sampling Rates in Infinite-Dimensional Compressed Sensing

Mitra Fatemi, Loic Baboulaz and Martin Vetterli
Laboratory of Audiovisual Communications
École Polytechnique Fédérale de Lausanne (EPFL)
Email: {mitra.fatemi, loic.baboulaz, martin.vetterli}@epfl.ch

***Abstract*—The theory of compressed sensing studies the problem of recovering a high dimensional sparse vector from its projections onto lower dimensional subspaces. The recently introduced framework of infinite-dimensional compressed sensing [1], to some extent generalizes these results to infinite-dimensional scenarios. In particular, it is shown that the continuous-time signals that have sparse representations in a known domain can be recovered from random samples in a different domain. The range M and the minimum number m of samples for perfect recovery are limited by a balancing property of the two bases. In this paper, by considering Fourier and Haar wavelet bases, we experimentally show that M can be optimally tuned to minimize the number of samples m that guarantee perfect recovery. This study does not have any parallel in the finite-dimensional CS.**

I. Introduction

Real-world signals are inherently analog or continuous-time and we often observe them through digital measuring devices. Imaging devices such as digital cameras and magnetic resonance imaging (MRI) machines are well known examples that measure light fields and brain signals, respectively. A linear measuring process consists of sampling the signal using certain sampling kernels. The samples of a continuous-time signal f can be regarded as its coefficients in an infinite-dimensional sampling domain $\mathcal{S}$ with a basis made of the sampling kernels. In general, infinite number of samples is required to precisely represent f. By adapting the sampling kernels to a specific type of signal, it is possible to reduce the infinite dimensional representation to a finite one. However, in most of the acquisition devices, the sampling kernels are limited by the physics of the device, and are rarely controllable. Therefore, it is very likely that a finite collection of samples captured by a measuring device result in a poor approximation of the signal.

An approach to reconstructing a satisfactory approximation of the signal is to calculate its coefficients in another domain $\mathcal{R}$ that efficiently represents the class of signals subject to the measurement. This means that any signal f in this class has sparse or fast decaying coefficients in $\mathcal{R}$ and N-term approximations of f in $\mathcal{R}$ rapidly converge to the signal. Wavelets are examples of the representation domains that provide fast converging approximations for piecewise continuous signals with pointwise singularities. Also, piecewise smooth images have compressible coefficients in the curvelet [2] and contourlet [3] domains.

First introduced in [4] and further improved in [5], consistent reconstruction is concerned with the problem of calculating the coefficients of a signal in a domain from its samples in a different domain. The consistent reconstruction method uses N samples in the sampling domain to calculate N coefficients in the reconstruction domain. Adcock and Hansen revisited this problem in [6], [7] and they argued that in general, N samples may not be enough to stably find N coefficients in $\mathcal{R}$. Also, they introduced a new *generalized sampling* (GS) approach to stably recover N coefficients in $\mathcal{R}$ from M samples in $\mathcal{S}$, where usually the stable sampling rate M is larger than N.

With the GS framework, we can perfectly reconstruct the signals that have sparse coefficients in a known domain $\mathcal{R}$ from a finite number of samples. However, similar to the finite-dimensional compressed sensing (CS) [8], [9], we are interested to take advantage of the sparsity of coefficients to reduce the number of samples. This problem can be considered as an infinite-dimensional variant of the CS problem where the goal is to recover a sparse vector $\mathbf{x}$ from linear measurements $\mathbf{y} = U\mathbf{x}$. It is shown that if the sensing matrix U has the so-called *restricted isometry property* (RIP) of order $2k$, any k-sparse vector $\mathbf{x}$ can be uniquely recovered from the measurements $\mathbf{y} = U_{m\times n}\mathbf{x}$ [10]. However, verifying the RIP condition for a matrix is computationally hard. In [11], Candès and Romberg considered orthonormal matrices $U \in \mathbb{R}^{n\times n}$ and they showed that in this case the coherence $\mu(U) = \max_{i,j} u_{i,j}$ can be used to determine the subsampling rate m.

Adcock and Hansen recently extended this idea to GS to address infinite-dimensional compressed sensing [1]. In this theory, a set of k-sparse coefficients in $\mathcal{R}$ with the support of nonzero coefficients in $\{1, ..., N\}$ are recovered with high probability from m samples in $\mathcal{S}$ chosen uniformly at random from the range $\{1, ..., M\}$ by solving the *basis pursuit* problem. The subsampling rate m depends on the coherence of the underlying sensing matrix. In addition, the parameters (N, k, M, m) should satisfy a *balancing condition* (refer to II-B).

The infinite-dimensional CS developed in [1] is a promising framework that allows us to obtain far better approximations of signals and images. However, it is not clear in this theory how the parameters (M, m) change with respect to (N, k) and what are the optimum values of the sampling rate m and the support range M. In this paper, we study this problem. Specifically, we study the change of m as a function of M for some specific choices of sampling and reconstruction domains and find the optimum values of (M, m) for given values N and k, through the experiments.

The paper is organized as follows. In Section II, we define the problem and briefly review GS and infinite-dimensional CS theories. In Section III, we study the balancing condition in infinite-dimensional CS and discuss the optimum choices of sampling rate and support. Also, we present some experiment results to calculate the optimum values of (M, m) for some given pairs (N, k) when the sampling and reconstruction kernels are Fourier exponentials and Haar wavelets. We use the optimum values calculated in this section to recover the sparse coefficients of different signals in Section IV. Finally, we conclude the paper in Section V.

II. Problem Description

Let $\mathcal{H}$ be a Hilbert space and $\mathcal{S}, \mathcal{R} \subseteq \mathcal{H}$ represent the sampling and reconstruction spaces with the orthonormal bases $\{\psi_j\}_{j=1}^{\infty}$ and $\{\phi_i\}_{i=1}^{\infty}$, respectively. Let $f = \sum_{i=1}^{\infty} \alpha_i \phi_i$ be the signal we wish to recover and suppose that we have access to the collection of samples

$$\beta_1, \beta_2, ... \quad \text{with} \quad \beta_j = \langle f, \psi_j \rangle. \tag{1}$$

The problem throughout this paper is to recover the best approximation of f in terms of $\{\phi_j\}_{j=1}^{\infty}$ from the samples in (1). Equivalently, we seek the best approximation of the coefficients $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ...]^T$ from measurements $\boldsymbol{\beta} = [\beta_1, \beta_2, ...]^T = U\boldsymbol{\alpha}$, with

$$U = \begin{pmatrix} \langle \phi_1, \psi_1 \rangle & \langle \phi_2, \psi_1 \rangle & \dots \\ \langle \phi_1, \psi_2 \rangle & \langle \phi_2, \psi_2 \rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{2}$$

A. *Consistent reconstruction and generalized sampling*

The consistent reconstruction of f is a point $\hat{f} \in \mathcal{R}$ that generates the same samples $\langle \hat{f}, \psi_j \rangle = \beta_j, \quad j = 1, 2, ...$. If we represent the orthogonal projection of f onto $\mathcal{S}$ by $P_{\mathcal{S}} f = \sum_{j=1}^{\infty} \beta_j \psi_j$, this is equivalent to

$$\hat{f} \in \mathcal{R} : \quad P_{\mathcal{S}} f = P_{\mathcal{S}} \hat{f}. \tag{3}$$

When the two subspaces satisfy $\mathcal{R} \oplus \mathcal{S}^{\perp} = \mathcal{H}$, equation (3) has a unique solution that can be found by solving the infinite-dimensional system of linear equations $U\boldsymbol{\alpha} = \boldsymbol{\beta}$ [4]. Clearly in practice, we have access to a finite number of samples. Therefore, we must consider truncations of this linear system and seek the first N coefficients $\boldsymbol{\alpha}^N$ of $\boldsymbol{\alpha}$. This is equivalent to looking for the N-term approximation of f in $\mathcal{R}$, i.e. $P_{\mathcal{R}_N} f = \sum_{i=1}^{N} \alpha_i \phi_i$.

We may think of solving this problem by taking N samples in $\mathcal{S}$ and considering the consistency condition in the N-dimensional subspace $\mathcal{S}_N$:

$$\hat{f} \in \mathcal{R}_N \quad \text{s.t.} \quad P_{\mathcal{S}_N} \hat{f} = P_{\mathcal{S}_N} f.$$

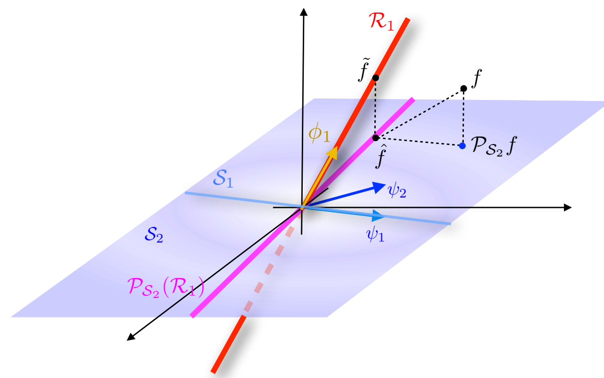


Fig. 1. Generalized sampling reconstruction $\tilde{f}$ of f in $\mathcal{R}_1$ from samples in $\mathcal{S}_2$.

The above equation has a stable solution only if

$$\mathcal{R}_N \oplus \mathcal{S}_N^{\perp} = \mathcal{H}. \tag{4}$$

If we define the angle between two subspaces $\mathcal{R}, \mathcal{S}$ as

$$\cos(\theta_{\mathcal{RS}}) = \inf_{\substack{\mathbf{r} \in \mathcal{R} \\ \|\mathbf{r}\|=1}} \|P_{\mathcal{S}} \mathbf{r}\|,$$

then the condition in (4) is equivalent to $\cos(\theta_{\mathcal{R}_N \mathcal{S}_N}) \neq 0$. In general, this condition may not hold for an arbitrary N, even if the infinite-dimensional spaces satisfy $\mathcal{R} \oplus \mathcal{S}^{\perp} = \mathcal{H}$ [6]. The generalized sampling approach to this problem is to increase the number of samples $M > N$ such that the condition $\cos(\theta_{\mathcal{R}_N \mathcal{S}_M}) \neq 0$ is met. In this case, the projection of $\mathcal{R}_N$ onto $\mathcal{S}_M$ is an N dimensional subspace $P_{\mathcal{S}_M}[\mathcal{R}_N] = \overline{span\{P_{\mathcal{S}_M} \phi_i\}_{i=1}^{N}}$. Now, we find an approximation of $P_{\mathcal{R}_N} f$ by verifying the consistency condition in this subspace [7]

$$\hat{f} \in \mathcal{R}_N \quad \text{s.t.} \quad P_{P_{\mathcal{S}_M}[\mathcal{R}_N]} \hat{f} = P_{P_{\mathcal{S}_M}[\mathcal{R}_N]} f. \tag{5}$$

Note that $P_{P_{\mathcal{S}_M}[\mathcal{R}_N]} f = P_{P_{\mathcal{S}_M}[\mathcal{R}_N]} P_{\mathcal{S}_M} f$ can be derived from the samples.

In Figure 1, we try to explain the GS reconstruction through an example in $\mathbb{R}^3$. In this example, we find the approximation of f in $\mathcal{R}_1$ from two samples in $\mathcal{S}_2$. Note, that since $\mathcal{R}_1$ is orthogonal to $\mathcal{S}_1 = \overline{span\{\psi_1\}}$, one sample of f in $\mathcal{S}_1$ is not sufficient for the stable approximation of f in $\mathcal{R}_1$.

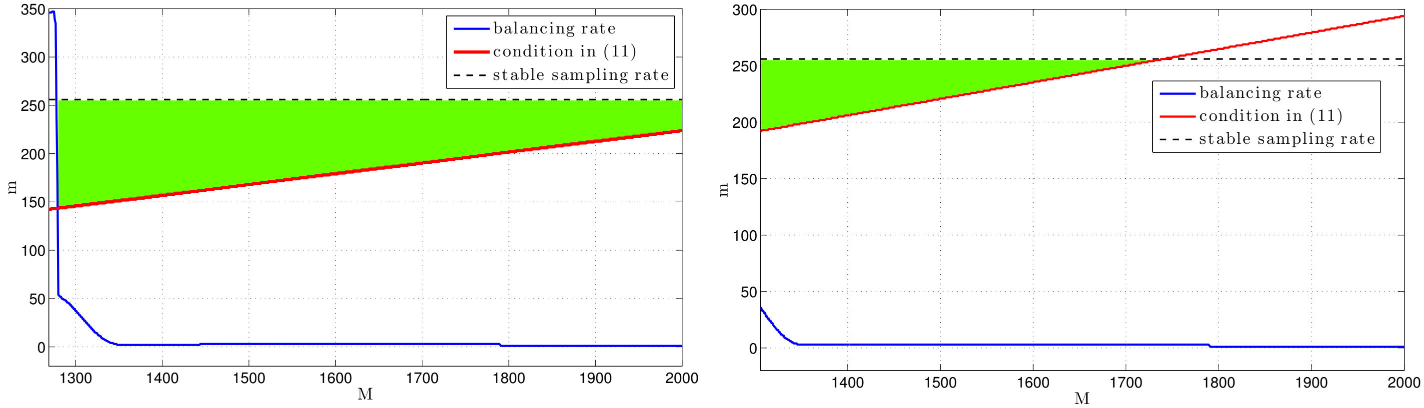
The solution of the GS equation in (5) is a stable approximation of f in $\mathcal{R}_N$ and it satisfies

$$\|f - \hat{f}\| \leq \frac{1}{\cos(\theta_{\mathcal{R}_N \mathcal{S}_M})} \|f - P_{\mathcal{R}_N} f\|.$$

Also, the coefficients of $\hat{f}$ can be calculated as $\boldsymbol{\alpha}^N = ((U^{M,N})^* U^{M,N})^{-1} (U^{M,N})^* \boldsymbol{\beta}^M$, where $U^{M,N}$ is the $M \times N$ subsection of U.

B. *Infinite-dimensional compressed sensing*

Now, assume that the coefficients $\boldsymbol{\alpha}$ are k-sparse with a support $\Delta \in \{1, ..., N\}$. In this case, we can perfectly recover f from equation (5). The infinite-dimensional CS approach in [1] exploits the sparsity to reduce the number of samples.



(a) (b)

Fig. 2. The acceptable range of sampling rate m and sampling support M for samples in the Fourier domain and sparse coefficients in the Haar domain, $N = 200$ and (a) $k = 30$, (b) $k = 40$. The blue and red plots display the minimum values of m as a function of M that are dictated by the balancing property and the equation (7) with $\epsilon = 0.05$, respectively. The black lines show the stable sampling rate in GS. The green regions display the acceptable ranges of (M, m).

The price of the subsampling, however, is to trade the stable recovery in GS with a probabilistic recovery.

Before we recall the main results in [1] for recovery of sparse or compressible signals in $\mathcal{R}$, we need to define the balancing property.

Definition 1. *Let U be the isometry matrix in (2). Then M and m satisfy the balancing property with respect to U, N and k if*

$$\|(U^{M\times N})^* U^{M\times N} - I_{N\times N}\| \leq \left(4\sqrt{\log_2(4M\sqrt{k}/m)}\right)^{-1},$$

$$\|(U^{M\times N})^* U^{M\times N} - \mathit{diag}((U^{M\times N})^* U^{M\times N})\|_{mr} \leq \frac{1}{8\sqrt{k}},$$

where $\|U\|_{mr}$ denotes the maximum ℓ^2 norm of different rows of U.

Theorem 1. *Let U be an isometry matrix with the coherence $\mu(U) = \max_{i,j} |u_{i,j}|$. Let the coefficients $\boldsymbol{\alpha} \in \ell^1(\mathbb{N})$ in $\mathcal{R}$ can be written as $\boldsymbol{\alpha} = \boldsymbol{\alpha}_0 + \boldsymbol{\alpha}_1$ with $\boldsymbol{\alpha}_0, \boldsymbol{\alpha}_1 \in \ell^1(\mathbb{N})$ and $supp(\boldsymbol{\alpha}_0) = \Delta \subset \{1, ..., N\}$ and $supp(\boldsymbol{\alpha}_1) = \{1, ..., N\}$. Also let $\epsilon > 0$ and $\Omega \subset \{1, ..., M\}$ be chosen uniformly at random with $|\Omega| = m$. If $\boldsymbol{\beta} = U\boldsymbol{\alpha}$ and $\hat{\boldsymbol{\alpha}}$ is a minimizer of*

$$\inf_{\boldsymbol{\eta}\in\ell^1(\mathbb{N})} \|\boldsymbol{\eta}\|_{\ell^1} \quad s.t. \quad U_{\Omega}^{M\times N}\boldsymbol{\eta}^N = \boldsymbol{\beta}_{\Omega}, \tag{6}$$

then with probability exceeding $1 - \epsilon$ we have

$$\|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}\| \leq \left(\frac{20M}{m} + 11 + \frac{m}{2M}\right)\|\boldsymbol{\alpha}_1\|_{\ell^1},$$

given that $(N, |\Delta|, M, m)$ satisfy the balancing property and m satisfies

$$m \geq CM\mu^2(U)|\Delta|(\log(\epsilon^{-1}) + 1)\log\left(\frac{MN\sqrt{|\Delta|}}{m}\right), \tag{7}$$

for a universal constant C.

In case that $\boldsymbol{\alpha}_1 = 0$ and $\boldsymbol{\alpha}$ is a k-sparse vector with $k = |\Delta|$, the equation (6) has a unique solution that coincides with $\boldsymbol{\alpha}$ with probability greater than $1 - \epsilon$.

III. Optimal Sampling Rate

Theorem 1 indicates that a signal with a k-sparse representation in $\mathcal{R}_N$ can be recovered with high probability from m random samples in $\mathcal{S}_M$, if m fulfills the condition in (7) and (N, k, M, m) satisfy the balancing property with respect to U. The condition (7) has a simple structure and we can easily track the change in m based on changes in M, N and k. On the contrary, it is not clear which values of (N, k, M, m) satisfy the balancing property with respect to a given U and how changes in (N, k) affect the sampling rate m and sampling support M. In other words, it is not clear what the subsampling gain of this theory is with respect to the stable sampling rate of GS, for a given sparsity.

In this section, we investigate the balancing property when the underlying sampling and reconstruction domains are formed by Fourier exponentials and Haar wavelet functions in $L^2[0, 1]$. This special choice of basis functions has applications in the MRI problem.

We use the following setup to find efficient sampling rates for fixed pairs of N and k. First, we find all values of M in the range $\{k, k+1, ..., M_{\max}\}$ such that the submatrix $U^{M\times N}$ satisfies the constraint

$$\|(U^{M\times N})^* U^{M\times N} - \mathrm{diag}((U^{M\times N})^* U^{M\times N})\|_{mr} \leq \frac{1}{8\sqrt{k}}.$$

The upper bound $M_{\max}$ on the range of samples is usually determined by the sampling device. We point out that in general, the maximum row norm in the above equation does not change monotonically with M. Thus, we should find the acceptable values of M by checking all numbers in $\{k, k+1, ..., M_{\max}\}$.

In the next step, for each verified M, we find the minimum m that satisfies (7) and the first constraint in Definition 1. Finally, we accept the pair (M, m) if $m < \min(M, M_1)$ where M_1 denotes the stable sampling rate in GS corresponding to N.

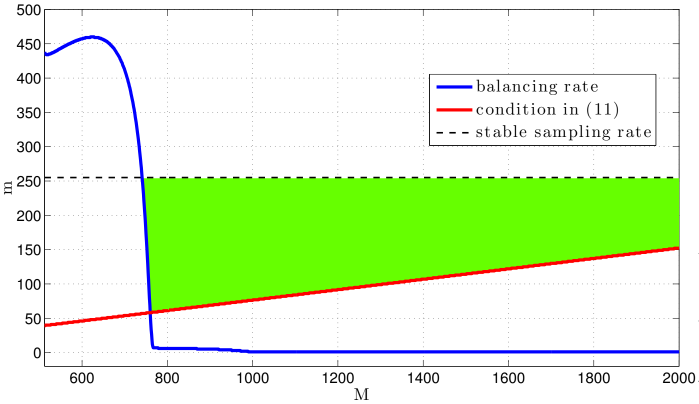


Fig. 3. The acceptable range of sampling rate m and sampling support M for samples in the Haar domain and sparse Fourier coefficients with $N = 200$ and $k = 20$.

Figures 2(a) and 2(b) display the acceptable pairs (M, m) for $N = 200$, $M_{\max} = 2000$ and two different sparsity values $k = 30, 40$, for sampling in Fourier and reconstruction in Haar domains. Figure 3 depicts the same variables for $k = 20$, when the sampling and sparsity domains are reversed. In these figures, the minimum values of m as a function of M satisfying the balancing property and the equation (7) are indicated in blue and red, respectively. The error probability is $\epsilon = 0.05$. Also, the black lines display the stable sampling rate corresponding to $N = 200$.

The green region in each figure shows the acceptable range of (M, m). The optimal sampling rate is determined by the point in this region that corresponds to the smallest m. For instance, Figure 3 shows that a signal with 20-sparse Fourier coefficients in the range $\{1, ..., 200\}$ can be recovered with probability greater than 0.95 from 58 samples that are chosen uniformly at random from the first 760 coefficients in the Haar domain. This means that we get a large subsampling gain by solving the basis pursuit problem in equation (6). On the contrary, Figure 2(b) illustrates that we do not get too much subsampling gain by replacing the basis pursuit problem in (6) with the stable reconstruction in GS for the specific values of the parameters in this plot.

IV. Numerical Experiments

In this section, we use the optimal values of (M, m) in Figure 2(a) to recover signals having sparse representations in the wavelet domain from randomly chosen Fourier coefficients.

In the first experiment, we consider signals of the form

$$f(t) = \sum_{i=1}^{200} \alpha_i \phi_i(t),$$

with only 20 nonzero coefficients, where $\{\phi_i(t)\}_{i\in\mathbb{N}}$ are Haar wavelets on $[0, 1]$. In the second experiment we consider signals of the form

$$f(t) = \sum_{i=1}^{200} \alpha_{0,i} \phi_i(t) + \sum_{i=1}^{200} \alpha_{1,i} \phi_i(t),$$

TABLE I
The approximation errors for the wavelet coefficients (avg. 100 trials)

	$\|\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}}\|_{\ell_\infty} / \|\boldsymbol{\alpha}\|_{\ell_\infty}$	SNR
Noiseless coefficients	0.1024×10^{-6}	104 dB
Noisy coefficients	0.7921×10^{-3}	64.1 dB

where the coefficient vector $[\alpha_{0,1}, ..., \alpha_{0,200}]^T$ is 20-sparse and $[\alpha_{1,1}, ..., \alpha_{1,200}]^T$ has a small ℓ_1 norm. For each case, we take $m = 144$ Fourier samples chosen uniformly from the first 1280 Fourier coefficients and we recover the signal by finding the solution to (6). Table I summarizes the approximation errors in the wavelet coefficients. The results in this table are averages over 100 trials.

V. Conclusion

We studied the sampling problem of infinite-dimensional signals that have sparse representations in a known domain. We adopted the random sampling approach of compressed sensing. Unlike the finite-dimensional case, the sampling scheme involves a pair (M, m), where m samples are randomly chosen among a size M subset of possible sampling kernels. For a given setup, there are various pairs which provide high probability of reconstruction. A counter intuitive result is that the required number of samples m does not necessarily decrease as M increases. We experimentally showed that one can find the optimum M that results in the minimum number of samples. We also observed that by swapping the sampling and sparsity domains, the optimal sampling schemes drastically change.

References

[1] B. Adcock and A. Hansen, "Generalized sampling and infinite-dimensional compressed sensing," 2011, preprint.
[2] E. Candès and D. Donoho, "New tight frames of curvelets and optimal representations of objects with piecewise c^2 singularities," *Communications on Pure and Applied Mathematics*, vol. 57, no. 2, pp. 219–266, 2003.
[3] M. Do and M.Vetterli, "The contourlet transform: an efficient directional multiresolution image representation," *IEEE Trans. Image Proc.*, vol. 14, no. 12, pp. 2091–2106, 2005.
[4] M. Unser and A. Aldroubi, "A general sampling theory for nonideal acquisition devices," *IEEE Trans. Signal Proc.*, vol. 42, no. 11, pp. 2915–2925, 1994.
[5] M. Unser and J. Zerubia, "A generalized sampling theory without band-limiting constraints," *IEEE Trans. Circuits Syst. II.*, vol. 45, no. 8, pp. 959–969, 1998.
[6] B. Adcock and A. Hansen, "A generalized sampling theorem for stable reconstruction in arbitrary bases," *Journal of Fourier Analysis and Applications*, pp. 1–32, 2010.
[7] ——, "Sharp bounds, optimality and a geometric interpretation for generalized sampling in hilbert spaces," 2011, preprint.
[8] E. Candes and T. Tao, "Near optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Info Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
[9] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
[10] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Info Theory*, vol. 52, no. 2, pp. 489–509, 2006.
[11] E. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Problems*, vol. 23, no. 3, pp. 969–985, 2007.