A rakeness-based design flow for Analog-to-Information conversion by Compressive Sensing

This is the author's version of the article that has been presented at IEEE ISCAS2013
The editorial version of the paper is available at http://dx.doi.org/10.1109/ISCAS.2013.6572107

A Rakeness-based Design Flow for Analog-to-Information Conversion by Compressive Sensing

Valerio Cambareri*, Mauro Mangia*, Fabio Pareschi†, Riccardo Rovatti*, Gianluca Setti†
* ARCES – University of Bologna, via Toffano 2/2, Bologna, Italy. {vcambareri,mmangia}@arces.unibo.it, riccardo.rovatti@unibo.it
† ENDIF – University of Ferrara, via Saragat 1, Ferrara, Italy. {fabio.pareschi, gianluca.setti}@unife.it

***Abstract*—Classical design of Analog-to-Information converters based on Compressive Sensing uses random projection matrices made of independent and identically distributed entries. Leveraging on previous work, we define a complete and extremely simple design flow that quantifies the statistical dependencies in projection matrices allowing the exploitation of non-uniformities in the distribution of the energy of the input signal. The energy-driven reconstruction concept and the effect of this design technique are justified and demonstrated by simulations reporting conspicuous savings in the number of measurements needed for signal reconstruction that approach 50%.**

I. Introduction

Analog-to-Information (A2I) converters aim at working on an amount of resources (hardware, time, energy, etc.) that depends on the rate of information contained in the signals to acquire, rather than their physical features (e.g. their overall bandwidth). Their design and implementation has attracted increasing interest as a way, for example, of deploying extremely economical interfaces between the real world and digital processing devices that, under suitable assumptions, are able to work below classical resource limits such as the Nyquist rate [1].

The core idea derives from Compressive Sensing (CS) [2, 3] in which the information content of a signal is identified with the number K of non-null coefficients needed to express it in a suitably defined signal basis, and the acquisition amounts to projecting the signal itself on a number M of testing waveforms.

The architecture of the most straightforward Analog-to-Information converter is equivalent to Figure 1, in which the signal is sampled N times in a given time window, the samples are multiplied by suitably generated coefficients, and the products are accumulated to yield the projections $y_0, \dots, y_{M-1}$. Projection y_0 is obtained by using coefficients $\Phi_{0,0}, \dots, \Phi_{0,N-1}$, projection y_1 by using coefficients $\Phi_{1,0}, \dots, \Phi_{1,N-1}$, and so on. The measurements are then ready to be possibly converted into digital words at the end of the sampling window.

In a given time window, the signal $x(t)$ is naturally represented by a set of N samples contained in the vector x, thus can be expressed as a linear combination of vectors in an N-dimensional basis. Yet, if it is known that the number of non-null coefficients in that linear combination is $K \ll N$ for any interesting instance, one may prove that the whole x can

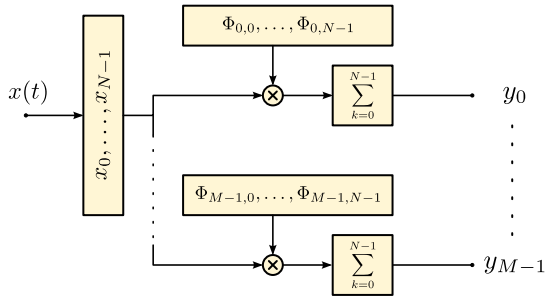

Fig. 1. Block scheme of an Analog-to-Information converter based on Compressive Sensing.

be recovered from a number of projections M that is much smaller than N and depends on K.

To formalize this concept align the basis vectors as columns of an $N \times N$ matrix Ψ, so that $x = \Psi\alpha$ for a certain N-dimensional vector of coefficients α, of which only K are known to be non-null for every realization of x.

Projection is a linear operation, thus we define an $M \times N$ matrix Φ such that the vector y of M measurements is $y = \Phi x = \Phi\Psi\alpha = \Theta\alpha$, where Θ is implicitly defined. Since Θ is an $M \times N$ matrix and $M < N$, recovering α from y is only possible by leveraging on the *a priori* assumption that only K of N components α are non-null, i.e. that α is K-sparse.

There is a flourishing literature on algorithms that perform sparse signal recovery [4, 5]; for the sake of simplicity we will adopt here the most classical solution, i.e. recovering α by minimizing $\sum_{j=0}^{N-1} |\alpha_j|$ subject to the constraint $y = \Theta\alpha$, which is a linear programming (LP) problem [6, 7].

Such a method is formally guaranteed to work if the matrix Θ obeys some conditions, the most widely employed being the Restricted Isometry Property (RIP) [8]. Since it turns out that many classes of random matrices are such that almost all the corresponding instances satisfy the RIP, the architecture in Figure 1 is commonly implemented to deliver M random projections of the signal x.

This common design method can be improved when, in addition to the sparsity assumption, we know that (as it is almost always the case) the energy of the signal to acquire is not evenly distributed along all directions, i.e., x is not the realization of a white process. We call such signals *localized*.

This has been shown in [9–11] where a criterion was introduced to statistically design the matrix Φ so that the $\Phi_{j,k}$

This is the author's version of the article that has been presented at IEEE ISCAS2013
The editorial version of the paper is available at http://dx.doi.org/10.1109/ISCAS.2013.6572107

are not independent and identically distributed (i.i.d.) random variables but are tuned to rake as much average energy as possible from the process to sense.

Brought to its extreme consequences, this would imply projecting only along a single direction, thus preventing the detection of signals in all other orthogonal directions. To prevent this overtuning, the second-order statistics of the projection waveforms appearing as the rows of Φ are chosen by solving an optimization problem maximizing the average energy of the measurements, while simultaneously guaranteeing that the projections are still random enough to capture energy even from less probable instances.

With this, the trade-off between the specialization needed to increase the measurement energy and the randomness needed to statistically span the whole signal space is administered by empirically setting a single real parameter r whose value is established by numerical and application-specific optimization.

The method described in [11] also lacked a formal definition of the localization concept on which rakeness-based design hinges, of its consequences on the measurement energy, as well as a clear connection between measurement energy and quality of the reconstruction.

This paper aims at taking a few steps forward in the complete definition of a rakeness-based design flow based on a clearer understanding of the link between localization, measurement energy and reconstruction quality, so that no trial-and-error tuning on r is needed to outperform blind i.i.d. design.

II. Energy-driven reconstruction

One of the most useful features of CS by linear projections is the so called "phase transition" [12, 13], i.e., the fact that for a fixed K and N, there is a critical $\bar{M}$ for the number of measurements in y such that $M > \bar{M}$ measurements are almost always enough to successfully reconstruct the input signal x from y, while $M < \bar{M}$ measurements are almost always insufficient.

Clearly, the exact $\bar{M}$ depends on the definition of reconstruction quality and on the threshold deciding when it is satisfactory, and its quantification is a key step in CS theory.

RIP-based guarantees on $\bar{M}$ [8] intrinsically consider isotropic projections since the main requirement on which they rely is that the matrix Θ is approximately an isometry when it is applied to *any* K-sparse vector.

On the contrary, rakeness-based design exploits non-uniformities in the statistics of the K-sparse vectors to design Θ and thus cannot exploit RIP-related results but at the price of a worst-case analysis that loosens performance guarantees.

Regrettably, RIP-based guarantees themselves may substantially underestimate phase transitions [14] and this makes the RIP analysis of the matrices implied by rakeness-based design useless for performance optimization.

For this reason, what we explore here is energy-driven reconstruction and, in preparation for a further theoretical step, we do so by numerically analyzing a toy system.

In particular, we consider unit-energy, K-sparse vectors α whose nonzero entries are i.i.d. Gaussian random variables with zero mean and variance $\frac{1}{K}$, and express $x = \Psi\alpha$ along the discrete Fourier basis Ψ, whose entries

$$\Psi_{j,k} = \frac{1}{\sqrt{N}} \begin{cases} 1 & \text{if } j = 0 \\ 2\cos\left(2\pi\frac{jk}{N}\right) & \text{if } j = 1, \dots, N/2-1 \\ (-1)^k & \text{if } j = N/2 \\ 2\sin\left(2\pi\frac{jk}{N}\right) & \text{if } j = N/2+1, \dots, N-1 \end{cases}$$

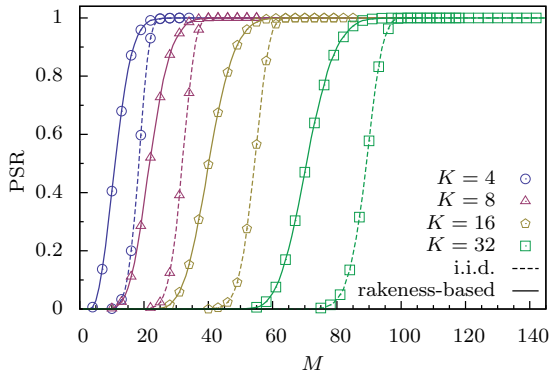

Fig. 2. Probability of successful reconstruction (PSR) for maximum energy projections (solid lines) versus i.i.d. projections (dotted lines) at different sparsity levels

A $10^4 \times N$ matrix Φ' of i.i.d. Gaussian entries with zero-mean and variance $\frac{1}{N}$ is used as a linear random projection operator to obtain a vector of candidate measurements $y' = \Phi' x$.

Two reconstructions of α are performed by solving the standard LP problem: one uses the first M components of y' (thus reproducing the simplest, classical CS setting); the other uses the M largest-modulus components of y'.

The results are reported in Figure 2 for $N = 256$ and $K = 4, 8, 16, 32$ by reporting the fraction of instances of x that are successfully reconstructed against the number M of measurements used. Since we do not consider additional noise in the system, assessing a successful reconstruction $\hat{\alpha}$ of the sparse vector α boils down to computing the relative error $\|\hat{\alpha} - \alpha\|^2 / \|\alpha\|^2$ and matching it against a $-120\,\mathrm{dB}$ threshold, which discriminates between successes up to machine precision (at about $-320\,\mathrm{dB}$) and failures (at about $0\,\mathrm{dB}$).

It is clear that, even in a noiseless configuration, choosing the M measurements with the largest energy improves sparse signal reconstruction w.r.t. the classical setting, since it moves the phase transition to lower values of $\bar{M}$.

Though this is not a proof, though we are dealing with a toy system that has no practical advantage since it relies on choosing the best M of 10^4 projections, and though numerical evidence is useless for what concerns system optimization, we still derive the clear indication that increasing the average energy of the projections is beneficial.

III. Localization

Assuming that the overall signal energy is normalized to 1, the most natural way to collect more energy is to identify the directions with which x preferentially aligns and concentrate the projections along them.

This is the author's version of the article that has been presented at IEEE ISCAS2013
The editorial version of the paper is available at http://dx.doi.org/10.1109/ISCAS.2013.6572107

Signal	Original Sampling Rate	N	$\mathcal{L}(\mu)$
ECG [15]	720 Hz	360	0.187
Speech [16]	20 KHz	200	0.069
EMG [15]	400 Hz	200	0.021
B&W printed letters [11]	24×24 pixels	576	0.016

TABLE I
VALUES OF $\mathcal{L}(\mu)$ FOR REAL-WORLD SIGNAL CLASSES

To explore how the energy of x distributes in the signal space, consider the correlation matrix $C_x = \mathbf{E}[xx^\top]$ together with its eigenvalues $\mu_0 \geq \mu_1 \geq \cdots \geq \mu_{N-1} \geq 0$ and corresponding orthonormal eigenvectors $u_0, \ldots, u_{N-1}$.

To formally define localization, normalize the average energy of x as $\mathbf{E}[x^2] = \operatorname{tr}(C_x) = 1$ so that $\sum_{l=0}^{N-1} \mu_l = 1$, and quantify the localization of x as

$$\mathcal{L}(\mu) = \sum_{l=0}^{N-1} \left(\mu_l - \frac{1}{N}\right)^2$$

that is null when all the eigenvalues are equal and increases up to $1 - 1/N$ when $\mu_0 = 1$ and $\mu_l = 0$ for $l = 1, \ldots, N-1$, i.e., when there is only one direction (u_0) along which all the energy of x is concentrated.

Almost all signals are localized since only perfectly uniform energy distributions yield $\mathcal{L}(\mu) = 0$. Table I reports localization values for some real-world signal classes that may undergo A2I conversion, namely, ElectroCardioGrams, ElectroMyoGrams, images of black-and-white isolated printed letters, 10 ms-long speech segments.

IV. RAKENESS-BASED DESIGN

Rakeness-based design (for a thorough discussion and derivation see [11]) implies a matrix Φ whose rows are i.i.d. vectors and their correlation matrix $C_\Phi = \mathbf{E}[\Phi_{j,\cdot}\Phi_{j,\cdot}^\top] = \sum_{l=0}^{N-1} \lambda_l u_l u_l^\top$ is decided by the eigenvalues $\lambda_0 \geq \lambda_1 \geq \cdots \geq \lambda_{N-1} \geq 0$. The expression of these new eigenvalues is derived in [11] and can be rearranged into $\lambda_l = \frac{1}{N} + \left(\mu_l - \frac{1}{N}\right)\sqrt{\frac{r - \frac{1}{N}}{\mathcal{L}(\mu)}}$, to highlight the newly defined localization $\mathcal{L}(\mu)$ and a positive real parameter r that must be empirically set to administer the trade-off between focusing on the most energetic directions and providing a random span of the whole signal space.

To avoid the hand-tuning of r, let us begin by defining $r = \frac{1}{N} + \tau^2 \mathcal{L}(\mu)/(1 - N\mu_{N-1})^2$ for some $\tau \geq 0$ so that $\lambda_l = \frac{1}{N} + \tau\left(\mu_l - \frac{1}{N}\right)/(1 - N\mu_{N-1})$. With this, straightforward calculations give the average energy raked by projecting along the rows of Φ as $\mathbf{E}[y_j^2] = \sum_{l=0}^{N-1} \lambda_l \mu_l = \frac{1}{N} + \tau\mathcal{L}(\mu)/(1 - N\mu_{N-1})$ that increases with $\mathcal{L}(\mu)$. The localization of the process generating the rows of Φ is

$$\mathcal{L}(\lambda) = \sum_{l=0}^{N-1} \left(\lambda_l - \frac{1}{N}\right)^2 = \left(\frac{\tau}{1 - N\mu_{N-1}}\right)^2 \mathcal{L}(\mu)$$

The trade-off between the exploitation of preferred directions and random spanning of the whole signal space can be recast into a localization inequality. In fact, $\mathcal{L}(\lambda) \leq \mathcal{L}(\mu)$ is the mathematical equivalent of the natural requirement that sensing waveforms are not more "specialized" than the waveforms they have to sense. In terms of the design parameter τ, this translates into $0 \leq \tau \leq 1 - N\mu_{N-1}$. It is observed that when τ is chosen in this interval, performances do not vary significantly; it is therefore sensible to choose τ in mid-range, i.e., $\tau = (1 - N\mu_{N-1})/2$ that implies $\lambda_l = \left(\frac{1}{N} + \mu_l\right)/2$.

Note that the smallest eigenvalue of C_Φ is surely positive thus ensuring that C_Φ is nonsingular and that no direction is neglected by random projections.

The overall design flow is therefore completely defined, requires no fine-tuning of r, and consists of these few steps:

1) Obtain the correlation matrix C_x of the signal to acquire and compute its eigenvalues μ_l end eigenvectors u_l;
2) Set $C_\Phi = \frac{1}{2}\sum_{l=0}^{N-1}\left(\frac{1}{N} + \mu_l\right) u_l u_l^\top$;
3) Generate every row of Φ as an independent random vector with correlation C_Φ.

Step 3) is straightforward when the rows of Φ are jointly Gaussian vectors, since we may start from a classical projection matrix Φ' containing i.i.d. Gaussian entries with zero mean and unit variance, and multiply it by the matrix $\sqrt{C_\Phi} = \sum_{l=0}^{N-1} \sqrt{\frac{1}{2}\left(\frac{1}{N} + \mu_l\right)} u_l u_l^\top$ to obtain $\Phi = \Phi'\sqrt{C_\Phi}$.

V. PERFORMANCE EVALUATION

To demonstrate the effectiveness of this design flow we test it on synthetic signals that are K-sparse in the same Fourier basis Ψ used in Section II.

We modulate the localization of the input signal x by drawing the K indexes at which its sparse coefficients α are non-null at random, with the same non-flat probability assignment $p : \{0, \ldots, N-1\} \to [0, 1]$ and ensuring that they are all different. If we indicate with $k_0 < k_1 < \cdots < k_{K-1}$ the indexes identifying the support of α, the signal is $x = \sum_{l=0}^{K-1} \alpha_{k_l} \Psi_{\cdot,k_l}$. In our cases, the α_{k_l} are i.i.d. Gaussian random variables with zero mean and variance $1/K$.

With these assumptions, the joint probability of the support of α is $p(k_0, \ldots, k_{K-1}) = \prod_{s=0}^{K-1} p(k_s) \left[\sum_{t=k_s+1}^{N-K+s} p(t)\right]^{-1}$ when $k_0 < k_1 < \cdots < k_{N-1}$ and zero otherwise. From this we may derive the marginal probabilities $q_l(k) = \Pr\{k_l = k\}$ and, exploiting the independence between the α_{k_l} and the fact that they have unit variance, the correlation matrix of x, $C_x = \sum_{l=0}^{K-1}\sum_{k=0}^{N-1} q_l(k)\Psi_{\cdot,k}\Psi_{\cdot,k}^\top$. By exchanging the order of summation we have $C_x = \sum_{k=0}^{N-1}\left[\sum_{l=0}^{K-1} q_l(k)\right]\Psi_{\cdot,k}\Psi_{\cdot,k}^\top$, from which we get that the very same columns of Ψ are the eigenvectors corresponding to eigenvalues $\lambda_k = \sum_{l=0}^{K-1} q_l(k)$ that are needed to trigger the rakeness-based design flow, yielding the random projection operator Φ.

In our case, the initial probability assignment p is taken with a doubly triangular profile with maxima in $N/4 - 1$ and $3N/4 - 1$ and varying widths for the corresponding triangles, thus producing different values for $\mathcal{L}(\mu)$. This choice allows the generation of K-sparse test signals with a given amount of localization. For each value of K and $\mathcal{L}(\mu)$, we generate 2000 realizations of α and evaluate the probability of successful reconstruction from $y = \Phi x$.

Figure 3 plots the minimum number of measurements needed to ensure that at least 90% of the K-sparse instances are correctly reconstructed as a function of $\mathcal{L}(\mu)$ when

This is the author's version of the article that has been presented at IEEE ISCAS2013
The editorial version of the paper is available at http://dx.doi.org/10.1109/ISCAS.2013.6572107

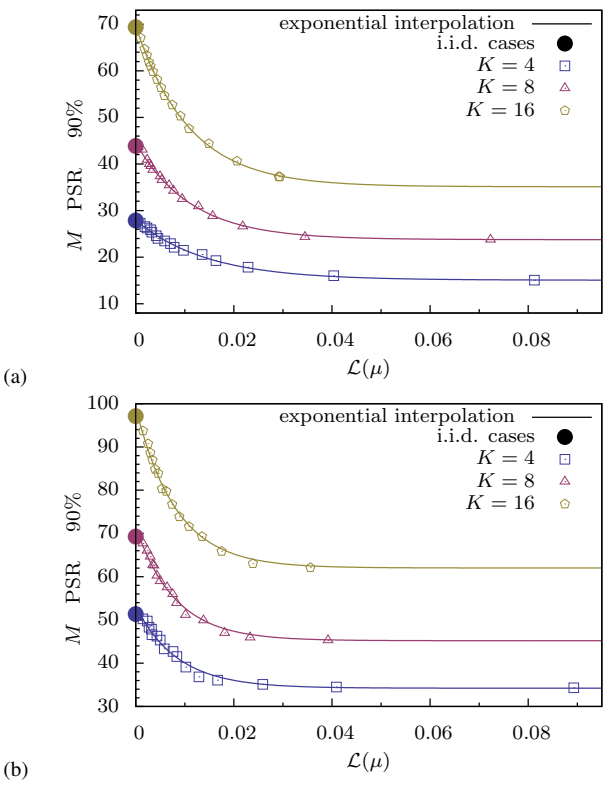

Fig. 3. The minimum number of measurements needed by rakeness-based design to have 90% probability of successful reconstruction as a function of localization when (a) no noise is superimposed to the signal and (b) additive Gaussian noise is superimposed to the signal with an SNR of 30dB.

rakeness-based design of Φ is used. Note that i.i.d. design disregards localization, thus behaving as if $\mathcal{L}(\mu) = 0$ and needing a number of measurements corresponding to the value of the curves at that abscissa.

Results are reported in the noiseless case with a quality requirement of 120 dB and when the input signal is perturbed by an additive white Gaussian noise with an SNR of 30 dB and reconstruction is considered successful when the relative error is below -20 dB. Note how, in both cases, even relatively small localization values (in agreement with those reported in Table I) can be exploited to substantially decrease the number of measurements needed for correct reconstruction of the signal and that the benefit increases as K increases.

As a final measure of performance, Figure 4 reports a comparison of the phase transition boundaries resulting from rakeness-based design and from i.i.d. design. Boundaries are reported in the (δ, ρ) plane [12] with $\delta = M/N$ and $\rho = K/M$ and correspond to system configurations in which 90% of the reconstructions of a noiseless signal (under the same conditions of Figure 3(a)) are successful. Note how the rakeness-based design largely dominates the i.i.d. design since the same number of measurements M allows the recovery of more informative signals with a significantly larger K.

VI. Conclusion

If the energy of the signal being acquired is not uniformly distributed in the signal space, we suggest the adoption of a rakeness-based design of the sampling waveforms entailed by an RMPI-A2I system. The corresponding, very simple design flow yields a substantial reduction in the number of measurements needed to reach a prescribed performance level.

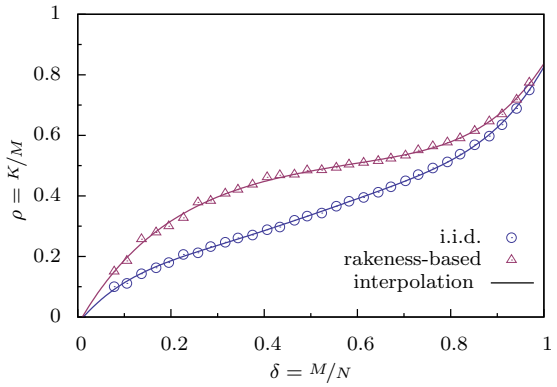

Fig. 4. Phase transition boundaries for rakeness-based (triangles) and i.i.d. (circles) random matrix designs.

References

[1] J. Haboba, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A pragmatic look at some compressive sensing architectures with saturation and quantization," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, Sep. 2012.

[2] E. J. Candes and M. B. Wakin, "An Introduction to Compressive Sampling," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[3] D. L. Donoho, "Compressed Sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[4] A. Maleki and D. Donoho, "Optimally tuned iterative reconstruction algorithms for compressed sensing," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 330–341, 2010.

[5] D. Needell and J. Tropp, "Cosamp: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, no. 3, pp. 301–321, 2009.

[6] E. J. Candes and T. Tao, "Decoding by linear programming," *IEEE Transactions on Information Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.

[7] E. J. Candes, J. K. Romberg, K. Justin, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, Aug. 2006.

[8] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, no. 9, pp. 589–592, 2008.

[9] J. Ranieri, R. Rovatti, and G. Setti, "Compressive sensing of localized signals: Application to analog-to-information conversion," in *Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on*, June 2010, pp. 3513–3516.

[10] M. Mangia, R. Rovatti, and G. Setti, "Analog-to-information conversion of sparse and non-white signals: Statistical design of sensing waveforms," in *Proceedings of 2011 IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2011, pp. 2129–2132.

[11] ——, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, no. 5, pp. 1001 –1014, May 2012.

[12] D. Donoho and J. Tanner, "Precise undersampling theorems," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 913 –924, June 2010.

[13] ——, "Counting faces of randomly-projected polytopes when the projection radically lowers dimension," *J. Amer. Math. Soc*, vol. 22, no. 1, pp. 1–53, 2009.

[14] J. Blanchard, C. Cartis, and J. Tanner, "Compressed sensing: How sharp is the restricted isometry property?" *SIAM Review*, vol. 53, no. 1, pp. 105–125, 2011.

[15] A.L. Goldberger *et al.*, "Physiobank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. 215–220, 2000.

[16] J. Harrington *et al.*, "The emu speech database system," available at http://emu.sourceforge.net/.