COMPRESSIVE SIGNAL PROCESSING WITH CIRCULANT SENSING MATRICES

Diego Valsesia   *Enrico Magli*

Politecnico di Torino (Italy) – Dipartimento di Elettronica e Telecomunicazioni

ABSTRACT

Compressive sensing achieves effective dimensionality reduction of signals, under a sparsity constraint, by means of a small number of random measurements acquired through a sensing matrix. In a signal processing system, the problem arises of processing the random projections directly, without first reconstructing the signal. In this paper, we show that circulant sensing matrices allow to perform a variety of classical signal processing tasks such as filtering, interpolation, registration, transforms, and so forth, directly in the compressed domain and in an exact fashion, *i.e.*, without relying on estimators as proposed in the existing literature. The advantage of the techniques presented in this paper is to enable direct measurement-to-measurement transformations, without the need of costly recovery procedures.

Index Terms— Compressed sensing, circulant matrix, compressive filtering

1. INTRODUCTION

Compressive sensing (CS) [1] has successfully shown that a small number of measurements, obtained through a suitable sensing matrix, is able to acquire a sparse signal and enable its exact recovery from its measurements. Remarkably, knowledge of the sparsity basis is not explicitly needed at the moment of acquisition, but only for recovery. Since the beginnings of CS, it was clear that working in the compressed domain presented advantages in terms of complexity because many tasks do not really require recovery but rather to solve inference problems. The seminal paper by Davenport *et al.* [2] on signal processing with compressive measurements provided techniques to perform detection, classification, estimation and interference cancellation tasks. In this paper, we focus on a different, although similar, problem. The goal is to perform classical signal processing operations such as filtering, interpolation and others directly on the compressive measurements, thereby avoiding to first apply the computationally expensive reconstruction algorithms in order to later apply the operator on the reconstructed samples. The resulting operation is a measurement-to-measurement transformation, which takes as input the acquired measurements and outputs

the measurements of the processed signal, without ever performing reconstruction. This is achieved by using a circulant sensing matrix, or a distributed circulation property in single node and multi-node scenarios respectively. Conversely, [2] is concerned with parameter estimation from the measurements. This can also be applied to the estimation of transform or filtered coefficients, but the estimate is always affected by some error, whereas our proposed method yields an exact solution.

2. BACKGROUND AND NOTATION

We use the subscript $A_{[a,b]}$ to denote a submatrix of A composed by the rows indexed by interval $[a, b]$. The subscript $\mathbf{x}_{\to a}$ denotes a circular shift to the right (downwards for column vectors) by a positions.

In the standard CS framework, introduced in [3], a signal $\mathbf{x} \in \mathbb{R}^{n \times 1}$ which has a sparse representation in some basis $\Psi \in \mathbb{R}^{n \times n}$, *i.e.*, $\mathbf{x} = \Psi \boldsymbol{\theta}$, $\|\boldsymbol{\theta}\|_0 = k$, $k \ll n$, can be recovered by a smaller vector $\mathbf{y} \in \mathbb{R}^{m \times 1}$, $k < m < n$, of linear measurements $\mathbf{y} = \Phi \mathbf{x}$, where $\Phi \in \mathbb{R}^{m \times n}$ is the *sensing matrix*. The optimum solution, seeking the sparsest vector compliant with $\mathbf{y}$, is an NP-hard problem, but one can resort to a convex optimization reconstruction problem by minimizing the l_1 norm, provided enough measurements $(m \sim k \log(n/k))$ are available. This algorithm is robust when it is used to reconstruct signals which are not exactly sparse, but rather compressible, meaning that the magnitude of their sorted coefficients (in some basis Ψ) decays exponentially.

The most used sensing matrices in the literature are random matrices whose elements are i.i.d. Gaussian random variables. However, circulant matrices have been recently proposed [4][5] because they can be implemented easily or arise naturally in various problems. Circulant matrices have been shown to be as effective as Gaussian sensing matrices, when the signal is acquired in its sparsity domain and have a slightly reduced performance otherwise [6]. The form of such matrices is

$$\Phi = \begin{bmatrix} \phi_1 & \phi_2 & \phi_3 & \cdots & \phi_n \\ \phi_n & \phi_1 & \phi_2 & \cdots & \phi_{n-1} \\ & & \vdots & & \\ \phi_{n-m+2} & \phi_{n-m+3} & \phi_{n-m+4} & \cdots & \phi_{n-m+1} \end{bmatrix}$$

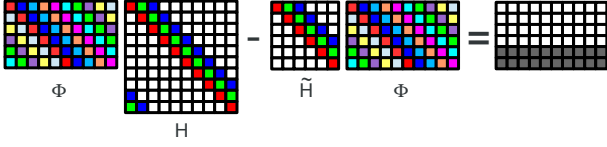where the first row, called seed in the following, is drawn at

Fig. 1. Partial commutation condition

random (*e.g.*, i.i.d. Gaussian). Other constructions (*e.g.*, randomly selecting m rows from the full matrix) are reported in [6], but we will not consider them in this paper. The n-point circular convolution of two sequences $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{h} \in \mathbb{R}^{N_f}$ is $\mathbf{x}_f = \mathbf{h} \circledast \mathbf{x} = H\mathbf{x}$, where H is an $n \times n$ circulant matrix.

The commutator of two linear operators $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times n}$ is zero if and only if the two operators commute and is defined as

$$[A, B] = AB - BA. \tag{1}$$

3. SINGLE-NODE COMPRESSIVE FILTERING

The main goal of this paper is to show that it is possible to compute a measurement-to-measurement transformation that allows us to find the measurements of a *filtered* version of the acquired signal, directly from its linear measurements. The classic filtering operation consists in convolving the filter impulse response with the signal, and circular convolution can be used if periodic boundary conditions are considered. This amounts to taking the product between a square circulant matrix H and the signal of interest. It can be observed that square circulant matrices form a commutative group with respect to matrix product, hence if we used a square circulant sensing matrix Φ we would have $\mathbf{y}_f = \Phi H\mathbf{x} = H\Phi\mathbf{x} = H\mathbf{y}$. CS uses rectangular sensing matrices in order to achieve compression, hence Φ is a "partial" circulant matrix. Nevertheless, we can still exploit a partial commutation property. Let us define an extension of the commutator introduced in (1) to handle rectangular matrices. The *m-partial commutator* is defined as:

$$[\Phi, H]_m = \Phi H - \tilde{H}\Phi \tag{2}$$

where $\tilde{H}$ is the submatrix of H restricted to the first m rows and m columns. The following theorem shows that obtaining the measurements of a filtered version of the signal is as straightforward as filtering the measurements. The price to pay for this operation is the "corruption" of $N_f - 1$ measurements in the positions corresponding to the non-zero rows of $[\Phi, H]_m$, thus the impulse response should not be too long to avoid corruption of too many measurements, and acquisition should take an extra $N_f - 1$ measurements to account for this. A graphical depiction of Theorem 1 is shown in Fig. 1.

Theorem 1. *Let H be an $n \times n$ circulant matrix obtained from an impulse response of length N_f, Φ be an $m \times n$ partial circulant sensing matrix, measurements $\mathbf{y} = \Phi\mathbf{x}$, and measurements of the filtered signal $\mathbf{y}_f = \Phi H\mathbf{x}$. Then, $\left(\tilde{H}\mathbf{y}\right)_i = (\mathbf{y_f})_i$ if and only if $i \in [1, m - N_f + 1]$.*
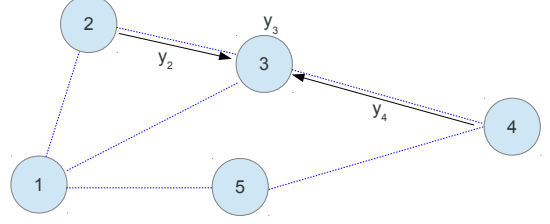

Fig. 2. Second derivative computation in multi-node scenario

Proof. First, we notice that it is enough to prove that only that the first $m - N_f + 1 = m'$ rows of the m-partial commutator are zero. Let us call $C \in \mathbb{R}^{n \times n}$ the square circulant matrix having the same seed as Φ, $H_T = H_{[1,m']}$ and call $\tilde{H}_T = \tilde{H}_{[1,m']}$. Then, $([C, H])_{[1,m']} = (CH)_{[1,m']} - (HC)_{[1,m']} = ([\Phi, H]_m)_{[1,m']} + \left(\tilde{H}\Phi\right)_{[1,m']} - (HC)_{[1,m']} = ([\Phi, H]_m)_{[1,m']} + \tilde{H}_T\Phi - H_TC = 0$. Since, by assumption, the first row of matrix H has at most N_f non-zero entries (all in the first N_f positions), we have $\left(H_T^T\right)_{[m,n]} = 0$ and $\tilde{H}_T = \left(\left(H_T^T\right)_{[1,m]}\right)^T$, so that $H_TC = \left(\left(H_T^T\right)_{[1,m]}\right)^T \Phi = \tilde{H}_T\Phi$. We thus find $([\Phi, H]_m)_{[1,m']} = 0$. The last $N_f - 1$ rows of $[\Phi, H]_m$ are surely non-zero because $\left(H_T^T\right)_{[m,n]} \neq 0$. $\square$

4. MULTI-NODE COMPRESSIVE FILTERING

In this section we deal with an extension of the previous framework, which can encompass two scenarios. The first one considers a multi-node distributed system in which agents cooperate by sharing some measurements in order to perform the filtering operation. The second one is a single-node system having as many measurements as those from all the nodes in the previous interpretation, but with a structured sensing matrix. Let us describe the principles of operation referring to the former scenario for ease of explanation, but always keeping in mind the latter.

In the multi-node scenario, each node is equipped with a random sensing matrix (*e.g.*, Gaussian), *not* circulant. However, nodes are assigned matrices that have a circulant property over the network. Numbering the nodes from 1 to J, this means that $\Phi^{(1)}$ is drawn at random, then row i in matrix $\Phi^{(j)}$ is a circularly right-shifted version of the same row in $\Phi^{(j-1)}$. If all the nodes observe the same signal $\mathbf{x}$, then they can collaborate to perform the filtering operation in the compressed domain. Given an impulse response $\mathbf{h}$, node j can obtain the measurements of the filtered signal $\mathbf{y}_f^{(j)} = \Phi^{(j)}\mathbf{x}_f$ as:

$$\mathbf{y}_f^{(j)} = \sum_{i=0}^{N_f - 1} h_i \mathbf{y}^{(j-i)} \qquad \text{for } j \in [N_f, J] \tag{3}$$

We can derive a commutation theorem similar to Thm. 1, by defining the (J, m)-*distributed partial commutator*:

$$\left[\tilde{\Phi}, H\right]_{J,m} = \tilde{\Phi}H - (H_J \otimes I_m)\tilde{\Phi} \tag{4}$$

where $\tilde{\Phi} = \left[\Phi^{(1)^T} \Phi^{(2)^T} \cdots \Phi^{(J)^T} \right]^T$, H_J is a submatrix of H restricted to the first J rows and columns, I_m is the $m \times m$ identity matrix, and $\otimes$ denotes Kronecker product.

Theorem 2. *Given J nodes, the l-th measurement of the filtered signal at node j is $\left(\mathbf{y}_f^{(j)} \right)_l = \sum_{i=0}^{N_f-1} h_i \left(\mathbf{y}^{(j-i)} \right)_l$ if and only if $j > N_f - 1$.*

We skip the proof for brevity and because it follows that of Thm. 1, now using the (J, m)-distributed partial commutator.

The main advantage of this multi-node framework is that each node has a fully random matrix on its own. Therefore, if its number of measurements is sufficiently high, it can perform reconstruction without suffering the slight performance degradation shown by circulant matrices when the sparsity domain is not the identity. Moreover, no measurement is corrupted in this process, except for $N_f - 1$ nodes, which cannot compute the filtered measurements with respect to their own sensing matrix. Moreover, multiple nodes can stack their measurements, corresponding to the second scenario discussed at the beginning of the section. In this case, the resulting sensing matrix would be obtained by stacking the individual sensing matrices, thus having a block-circulant structure (a total of m independent rows, used as seed of J circulant blocks).

5. MAIN APPLICATIONS OF THE PROPOSED FRAMEWORK

The following sections describe some applications of the presented framework. For reasons of brevity, we shall focus on the single-node case, but all the results can be readily extended to the multi-node scenario. The general case of filtering using circular convolution with a sequence $\mathbf{h}$ has already been treated in the previous section. The following examples consider problems that can be cast in a similar manner.

5.1. Finite differences

It is of interest to apply a discretized differential operator to the signal and compute the measurements of the resulting sequence directly from the original measurements. Let us consider a discretization of the second derivative, so that the n-th sample in the output is computed, in vector form, as:

$$\mathbf{x}_f = \mathbf{x}_{\to 1} - 2\mathbf{x} + \mathbf{x}_{\leftarrow 1}$$

where the subscript $\to k$ denotes a circular right-shift by k positions. It is easy to check that the measurements $\mathbf{y}_f = \Phi \mathbf{x}_f$ can be obtained directly in the compressed domain as:

$$\mathbf{y}_f = \mathbf{y}_{\to 1} - 2\mathbf{y} + \mathbf{y}_{\leftarrow 1}$$

with the corruption of the first measurement. In the multi-node setting it is immediate to derive that node j can request

the measurements of nodes $j-1$ and $j+1$ in order to compute $\Phi^{(j)} \mathbf{x}_f$ directly in the compressed domain:

$$\mathbf{y}_f^{(j)} = \mathbf{y}^{(j-1)} - 2\mathbf{y}^{(j)} + \mathbf{y}^{(j+1)}$$

This example is shown in Fig. 2.

5.2. Compressive interpolation

The objective is to compute the measurements of an interpolated version of the signal of interest from its compressive measurements. Interpolation is modelled as the cascade of an upsampler and a low-pass filter. Therefore it is possible to perform the operation in the compressed domain if we assume that the original signal was acquired by a circulant matrix with a few zeroed columns, whose number depends on the upsampling factor. The low-pass filter is implemented by the standard compressive filtering technique explained in the previous sections, bearing in mind that $N_f - 1$ measurements will be corrupted. As an example, let us consider a piecewise linear interpolator with an interpolation factor of 2. It is easy to verify that the interpolated sequence and the corresponding measurement-domain transformation are:

$$\mathbf{x}_{\text{INT}} = \frac{1}{2}\mathbf{x}_{\to 1} + \frac{1}{2}\mathbf{x}_{\leftarrow 1} + \mathbf{x}$$

$$\mathbf{y}_{\text{INT}} = \frac{1}{2}\mathbf{y}_{\to 1} + \frac{1}{2}\mathbf{y}_{\leftarrow 1} + \mathbf{y}$$

In this example the original measurements have been acquired as $\mathbf{y} = \hat{\Phi}\mathbf{x}$, where the i-th column of $\hat{\Phi}$ is $\hat{\Phi}_i = \Phi_{2i-1}$ for a given circulant matrix Φ. The measurements of the interpolated signal with respect to matrix Φ are now available.

5.3. Shift retrieval and registration

The problem of integer shift retrieval in the compressed domain was considered in [7]. The authors showed that the test $\max_s \Re \{ \langle \mathbf{z}, \Phi D^s \Phi^\star \mathbf{v} \rangle \}$ exactly recovers the integer shift s from as low as one measurement ($\mathbf{z}$ and $\mathbf{v}$ are the measurements of the signal and of its shifted version using cyclic shift operator D^s), given some conditions on Φ. The partial Fourier matrix was considered as an example of sensing matrix satisfying the prescribed properties. We notice that circulant matrices do not satisfy the conditions listed in Theorem 1 in [7]; in particular, $\alpha\Phi^\star\Phi \neq I$ for any scaling factor α. However, we can exploit Theorem 1 to perform shift retrieval in a different way. One of the implications of Theorem 1 is that a shift in the signal corresponds to a shift in the measurements by the same amount, being the cyclic shift operator D^s a circulant matrix, with the corruption of s measurements.

Theorem 3. *Let $\Phi \in \mathbb{R}^{m \times n}$ be a partial circulant matrix. Let $\mathbf{z} = \Phi \mathbf{x}$ and $\mathbf{v} = \Phi \mathbf{x}_{\to s^\star}$. If $|s^\star| < m$, then (5) retrieves the correct shift $\hat{s} = s^\star$.*

$$\hat{s} = \arg\min_s \|\tilde{\mathbf{z}} - \tilde{\mathbf{v}}\|_2 \tag{5}$$

with $\tilde{\mathbf{z}} = \mathbf{z}_{[1,m-s]}$ and $\tilde{\mathbf{v}} = \mathbf{v}_{[s+1,m]}$ for $s \geq 0$, or $\tilde{\mathbf{z}} = \mathbf{z}_{[|s|+1,m]}$ and $\tilde{\mathbf{v}} = \mathbf{v}_{[1,m-|s|]}$ for $s < 0$.

Proof. Let us call $C \in \mathbb{R}^{n \times n}$ the square circulant matrix having the same seed as Φ. Suppose that $0 < s^\star < m$. If $s = s^\star$, we have $\tilde{\mathbf{v}} = C_{[s^\star+1,m]}\mathbf{x}_{\to s^\star} = C_{[1,m-s^\star]}\mathbf{x}$, and $\tilde{\mathbf{z}} = C_{[1,m-s^\star]}\mathbf{x}$, hence $\tilde{\mathbf{v}} = \tilde{\mathbf{z}}$, so $\|\tilde{\mathbf{z}} - \tilde{\mathbf{v}}\|_2 = 0$. If $s > s^\star$ or $s < s^\star$ we have $\tilde{\mathbf{v}} = C_{[s+1,m]}\mathbf{x}_{\to s^\star} = C_{[s-s^\star+1,m]}\mathbf{x}$ and $\tilde{\mathbf{v}} = \left[C_{[m-s^\star,m]}^T C_{[1,m-s^\star-s]}^T \right]^T \mathbf{x}$, respectively and $\tilde{\mathbf{z}} = C_{[1,m-s]}\mathbf{x}$ in both cases. Hence, by construction of C, we have $\tilde{\mathbf{v}} \neq \tilde{\mathbf{z}}$. The same reasoning applies for $-m < s^\star < 0$. Thus, we proved retrieval of the correct shift. When $|s^\star| \geq m$, $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{z}}$ cannot be constructed because $\tilde{\mathbf{v}}$ and $\tilde{\mathbf{z}}$ contain measurements coming from two disjoint submatrices of C. $\square$

Notice that thanks to compressive interpolation (section 5.2), it is also possible to retrieve sub-integer shifts; we leave this as future work.

A different problem is registration. If we know the shift s, we can register a signal, meaning that we can compute the measurements of the shifted version simply by shifting the measurement vector. If want to compute the measurements with respect to the original matrix Φ, then s measurements will be corrupted by the registration. Otherwise, we can suppose that the registered measurements are measurements acquired with a different sensing matrix Φ', which is circulant, having as seed the $(m - s + 1)$-th row of the full circulant matrix, that the partial Φ was extracted from.

5.4. Compressive wavelet transform

We show that it is possible to obtain the measurements of the wavelet coefficients from the measurements of the signal directly in the compressed domain. Changing the signal domain after the sensing process can be useful in several ways. In [8], the authors propose a technique to improve the quality of the reconstruction of signals compressible in the wavelet domain by exploiting the tree-based structure within the wavelet coefficients. The method requires CS measurements of the wavelet coefficients in order to be applied. However, in many scenarios it is not practical to sense the wavelet coefficients, either because the sensing is implemented in a low-complexity, low-power hardware so that calculating the transform would be computationally and energetically costly or because specialized hardware directly acquires random projections of the data (*e.g.*, [9]). Moreover, being able to compute the measurements of the wavelet coefficients, which typically hold a sparse or compressible representation of the signal, allows to recover them directly, thus avoiding any issues regarding the coherence [10] of sensing matrix and spasifying wavelet dictionary.

In order to implement the wavelet transform in the compressed domain we exploit the lifting scheme [11]. The lifting method to compute the wavelet transform consists in a sequence of lifting steps, each composed by two filters operating on two subsequences. Initially, the two subsequences are the even and odd samples of the signal; the predict filter is used to predict the odd sequence from the even one, and the
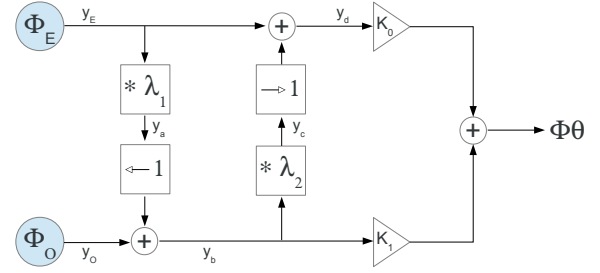

Fig. 3. Compressive wavelet transform

difference is then passed through the update filter and the result summed back to the even sequence. It can be noticed that this framework can be mapped in the compressed domain because the filtering steps can be implemented as explained in the previous sections under the choice of periodic boundary conditions. We thus only need to impose a structured sensing that acquires measurements of the even and odd subsequences separately. Let us call $\Phi^{(e)}$ and $\Phi^{(o)}$ the sensing matrices used to acquire the even and odd subsequences, respectively. Both have size $m \times n$, but the even (odd) columns are zeroed in $\Phi^{(o)}$ ($\Phi^{(e)}$). Moreover, the sum $\Phi^{(e)} + \Phi^{(o)}$ is a circulant matrix. The measurements of the even and odd sequences form the two input sequences in the compressive scheme, depicted in Fig. 3. The prediction step applies the prediction filter to the even sequence and the output is left-shifted by 1. The update step applies the update filter and right-shifts the output by 1. This is the structure of a single lifting step, which may be repeated depending on the particular transform to be implemented. Finally, gains are applied and the two sequences are summed. It can be shown that the final output is $\Phi\boldsymbol{\theta}$ being $\boldsymbol{\theta}$ the vector of wavelet coefficients. We remark that, as explained in section 3, $2(N_f - 1) + 2(N_f - 1)$ measurements per lifting step are corrupted. In more detail, let us consider a practical example taken from [12]. The spline 5/3 biorthogonal transform has the following prediction and update filters $\lambda_1(z) = -\frac{1}{2}(1+z)$ and $\lambda_2(z) = \frac{1}{4}(1+z^{-1})$ and gain factors $K_0 = 1$, $K_1 = \frac{1}{2}$. Applying the compressive transform procedure we can obtain the measurements of the wavelet coefficients with a corruption of the first 2 and last 2 measurements, which will be discarded.

6. CONCLUSIONS

In this paper we have shown how endowing the sensing matrix with circulant properties (*e.g.*, fully circulant as in the single-node case or block-circulant as in the multi-node case) allows to translate a variety of classical signal processing tasks to the reduced dimensionality domain in an exact form or corrupting few measurements. We have discussed some applications of the presented paradigm, including various forms of filtering, shift retrieval and registration and a technique to transform the measurements of a signal into the measurements of its wavelet coefficients. Future directions for the work presented in this paper include extending the theory to 2D signals.

7. REFERENCES

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

[2] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal Processing With Compressive Measurements," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 445–460, 2010.

[3] E. J. Candes and T. Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.

[4] W. U. Bajwa, J. D. Haupt, G. M. Raz, S. J. Wright, and R. D. Nowak, "Toeplitz-Structured Compressed Sensing Matrices," in *Statistical Signal Processing, 2007. SSP '07. IEEE/SP 14th Workshop on*, 2007, pp. 294–298.

[5] H. Rauhut, "Circulant and Toeplitz matrices in compressed sensing," *arXiv preprint arXiv:0902.4394*, 2009.

[6] W. Yin, S. Morgan, J. Yang, and Y. Zhang, "Practical compressive sensing with Toeplitz and circulant matrices," in *Proc. SPIE*, vol. 7744, 2010, p. 77440K.

[7] H. Ohlsson, Y. C. Eldar, A. Y. Yang, and S. S. Sastry, "Compressive shift retrieval," in *Acoustics, Speech and Signal Processing, 2013. ICASSP 2013. IEEE International Conference on*, vol. 1, 2013, pp. 6034–6038.

[8] L. He and L. Carin, "Exploiting structure in wavelet-based bayesian compressive sensing," *Signal Processing, IEEE Transactions on*, vol. 57, no. 9, pp. 3488–3497, 2009.

[9] J. Yoo, S. Becker, M. Monge, M. Loh, E. Candes, and A. Emami-Neyestanak, "Design and implementation of a fully integrated compressed-sensing signal acquisition system," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, 2012, pp. 5325–5328.

[10] E. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse problems*, vol. 23, no. 3, p. 969, 2007.

[11] W. Sweldens, "The lifting scheme: a construction of second generation wavelets," *SIAM J. Math. Anal.*, vol. 29, no. 2, pp. 511–546, Mar. 1998. [Online]. Available: http://dx.doi.org/10.1137/S0036141095289051

[12] D. S. Taubman, M. W. Marcellin, and M. Rabbani, "JPEG2000: Image compression fundamentals, standards and practice," *Journal of Electronic Imaging*, vol. 11, no. 2, pp. 286–287, 2002.