ENERGY EFFICIENT TELEMONITORING OF WHEEZES

Aris S. Lalos and Konstantinos Moustakas
Department of Electrical and Computer Engineering
University of Patras, Greece

Emails: aris.lalos@ece.upatras.gr, moustakas@ece.upatras.gr

ABSTRACT

Wheezes are abnormal continuous adventitious lung sounds that are strongly related to patients with obstructive airways diseases. Wireless telemonitoring of these sounds facilitate early diagnosis (short, long term) and management of chronic inflammatory disease of the airways (e.g., asthma) through the use of an accurate and energy efficient mhealth system. Therefore, low complexity breath compression schemes with high compression ratio are required. To this end, we propose a compressed sensing based compression/reconstruction solution that enables wheeze detection from a small number of linearly encoded samples, by exploiting the block sparsity of the breath eigenspectrum during reconstruction at the receiver. Simulation studies, carried out with publicly available breath sounds, show the energy efficiency benefits of the proposed CS scheme, compared to traditional CS recovery approaches.

Index Terms— Compressed Sensing, Wheezes, Time-Frequency Analysis, PCA

1. INTRODUCTION

Wheezes are continious adventitious sounds, similar to musical sounds, superimposed with the sound of normal breathing. The presence of wheezes in a breath sound is an insensitive sign for severity of obstructive airways diseases [1], such as asthma and chronic obstructive pulmonary disease (COPD). Several experts in the fields of information and communication technologies (sensing, signal processing, wireless communication, HW/SW design) , respiratory medicine, inhaler devices focus on building novel mhealth systems that facilitate the early diagnosis and management of such diseases through a personalised approach. Although medical devices that detect these sounds, are already commercially available [2], they operate on-demand in hand held form and cannot be used for continious tracking the intensity of symptoms. To that end, the authors in [3] proposed a low-cost wearable sensing system consisting of low power miniaturized wearable sensor, recording breath signals, and transmitting them to a body node coordinator (BNC) e.g., a smartphone, that serves as a gateway and facilitates the continuous monitoring of asthmatic wheezes.

The concept of an mhealth sensing system that facilitates the continious monitoring of asthmatic wheezes is shown in Fig.1. This systems requires new schemes and algorithms to be implemented in order to optimize the energy consumption and the total hardware cost at the transmitter. Low energy consumption significantly increases the battery lifetime of the breath sensor, while the hardware cost reduction makes the mhealth system economically viable and more easily accepted by the individual customers. Both requirements motivate the design of compression/reconstruction schemes, with high compression ratio capabilities and reduced computational requirements. The vast majority of audio compression schemes available in the literature [4] charge the transmitter with most of the processing, thus not coping effectively with the above requirements.

Compressed Sensing (CS) approaches for signal compression/reconstruction offers an affordable solution for audio compression in wireless sensor networks [5], by allowing the reconstructing of audio signals from a small number of random linear observations. To the best of our knowledge, this is the first work that demonstrates the benefits of CS based compression/reconstruction schemes for the efficient telemonitoring of breath sounds in wireless body ares networks (WBANs). More specifically, we enhance the benefits of the conventional CS schemes proposed in [5], by taking into account specific characteristics (e.g., block sparsity, sample correlation) of the breath sounds in the eigen spectrum domain. The proposed novel recovery algorithm, named PCA based Group LASSO, increases the mhealth system energy efficiency by a factor of 1.8 as compared to traditional CS recovery approaches.

The rest of the paper is outlined as follows: Section II, presents the system model. Section III describes the proposed compression/reconstruction schemes. In Section VI, the performance of the proposed scheme is evaluated and compared to state-of-the-art algorithms. Finally, Section VII concludes this paper.

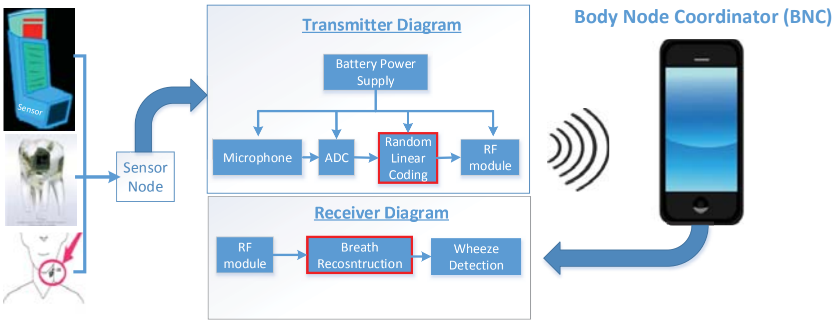


Fig. 1. *mHelath system for wheeze telemonitoring.*

2. BREATH TELEMONITORING MODEL

Figure 1 illustrates the telemonitoring system model under study. In particular, we consider a WBAN formed by an accoustic sensor (source node) that records a real time breath sound and transmits it to a BNC (e.g., smartphone). We assume that the sensor can be attached to the neck, to the tooth or even the inhaler of the patient.

The breath signal, is recorded by the microphne, digitized and divided into segments of N samples. Each segment is represented as a vector $\mathbf{x} = [x_1, \ldots, x_N]^T$, where $x_i \in \mathbb{R}$. We assume that the recorded signal to be transmitted contains noise itself and, as a result, it may be written as $\mathbf{u} = \mathbf{x}+ \mathbf{w}_s$, where $\mathbf{u} = [u_1, \ldots, u_N]^T$ are samples of the noisy signal and $\mathbf{w}_s = [w_1, \ldots, w_N]^T$ is the random noise. For each segment, the source generates M random linear combinations (see Fig. 1) by using a random matrix $\mathbf{A}$ of dimension $M \times N$ (Random Linear Coding - RLC) and performs quantization as follows:

$$\mathbf{y}_q = Q\left(\mathbf{A}\mathbf{u}\right) = \mathbf{A}\mathbf{x} + \mathbf{w}_q, \tag{1}$$

where $Q : \Re \rightarrow \mathbf{Y_i}$ is a scalar quantization function-Typical quantizers are usually optimized by selecting decision boundaries and output levels in order to minimize the distortion (e.g., mean square error) between the input real number and its quantized representation. that discretizes its input, by performing a mapping of each real element of $\mathbf{y}$ to a finite set of codewords $\mathbf{Y_i}$ and $\mathbf{w}_q$ represents the combination of the sensing and quantization error. The encoded samples are then quantized $\mathbf{y}_q$ transmitted to the BNC where the reconstruction of the original signal and breath reconstruction and analysis is taking place. At this point it should be mentioned that the overhead introduced by the transmission of the random encoding coefficients, can be significantly reduced by: i) adopting the policy that was used in [6], that is, instead of transmitting a full encoding matrix, the authors propose the transmission of the coefficient in the first row and then the generation of the $M-1$ rows at the destination by performing predefined shifts of the received row, and ii) considering that the same encoding coefficients remain fixed for a number of L segments, where $L >> N$. [1]

3. EFFICIENT RECONSTRUCTION OF BREATH SOUNDS

Several sparse representation methods, such as the discrete cosine/wavelet transfrom (DCT,DWT) [7], have been used in the past for sparsifying audio signals. Their sparse representation efficiency is based on the fact that they tend to decor-relate (e.g., DCT, DWT) a given signal and redistribute the energy contained in the signal, so that most of energy is contained in a small number of components. In this case, $\mathbf{x}$ can be expressed as $\mathbf{x} = \mathbf{\Psi}\mathbf{s}$, where $\mathbf{\Psi} \in \mathbb{R}^{N \times N}$ is an orthonormal basis matrix of an appropriate transform domain, and $\mathbf{s}$ is a sparse representation vector. The reconstruction of breath sound at the BNC, can be performed by exploiting the sparsity of the breath signal in some transform domain (e.g., DCT, DWT) through the use of the classical CS theory.

3.1. Conventional CS approach

According to the classical CS approach, vector $\mathbf{s}$ may be recovered from $\mathbf{y}$ by solving the problem:

$$min_s \left\{ \|\mathbf{s}\|_0 : \|\mathbf{y}_q - \mathbf{A}\mathbf{\Psi}\mathbf{s}\|_2^2 \leq \epsilon \right\}, \tag{2}$$

where the parameter ϵ corresponds to the predefined error tolerance and $\| \cdot \|_2$ is the ℓ_2 - norm of the input vector, respectively. The above optimization problem is computationally intractable and it cannot be used

[1] Experimental results have shown that the aforementioned strategies do not affect the performance of the decoding algorithms presented below in the BNC. Thus, it is reasonable to neglect the communication overhead that is introduced by the transmission of the encoding coefficients to the BNC and assume that matrix $\mathbf{A}$ is considered to be known at the receiver.

for practical applications. CS suggests replacing the ℓ_0 quasi-norm by the convex ℓ_1-norm and solving the following problem: $min_s \left\{\|\mathbf{s}\|_1 : \|\mathbf{y}_q - \mathbf{A\Psi s}\|_2^2 \leq \epsilon\right\}$, where $\|\mathbf{s}\|_1 = \sum_{i=1}^{N} |s_i|$. Langrange relaxation allow us to efficiently approximate the solution of the aforementioned problem by solving the problem:

$$\hat{\mathbf{s}}_L := \arg\min_{\mathbf{s}} \|\mathbf{y}_q - \mathbf{A\Psi s}\|_2^2 + \lambda\|\mathbf{s}\|_1, \qquad (3)$$

where λ is a penalty parameter that can be tuned, to trade off the value of the ordinary least square error $\|\mathbf{y} - \mathbf{A\Psi s}\|_2^2$ for the number of the nonzero entries (degree of sparsity) in $\mathbf{s}$. Algorithmically, the aforementioned convex optimization problem in eq. (3), known as LASSO problem, can be tackled by any generic second-order cone program (SOCP) solver (e.g., interrior point methods). The original signal in the time domain can be reconstructed by computing $\hat{\mathbf{x}}_L = \mathbf{\Psi}\hat{\mathbf{s}}_L$.

3.2. Exploitation of the Group Sparsity in a transform domain

Many wheezing sounds are represented in the frecuency domain, as a cluster of one main peak in parallel with several lower amplitude peaks. Thus, vector $\mathbf{s}$ can be viewed as a concatenation of R blocks of length d:

$$\mathbf{s} = [\ \mathbf{s}^T[1], \mathbf{s}^T[2], \ldots, \mathbf{s}^T[R]\ ]^T, \qquad (4)$$

where $\mathbf{s}[i] = \left[s_{(i-1)d+1}, \ldots, s_{id}\right]^T$ denotes the i^{th} block and $N = Rd$. Note that only few clusters, including those that correspond to wheezes, consist of large amplitudes and can be considered as non zero blocks. The aforementioned structure is known as block sparse structure and enables the signal recovery from a reduced number of samples [8], compared to sparse structures. By simply using the ℓ_1 relaxation for reconstructing $\mathbf{s}$, we ignore the fact that the signal is block-sparse, i.e., the non-zero entries occur in consecutive positions. To exploit block sparsity, we reconstruct vector $\mathbf{s}$ by solving:

$$\hat{\mathbf{s}}_{GL} := \arg\min_{\mathbf{s}} \|\mathbf{y}_q - \mathbf{A\Psi s}\|_2^2 + \lambda \sum_{i=1}^{R} \|\mathbf{s}[i]\|_2, \qquad (5)$$

which is also known as group LASSO problem [9]. Then, the breath signal in the native domain $\mathbf{x}$ can be reconstructed as in the LASSO case, by computing $\hat{\mathbf{x}}_{GL} = \mathbf{\Psi}\hat{\mathbf{s}}_{GL}$

3.3. PCA based group LASSO

Orthogonal transforms such as discrete cosine transform, and wavelet transform construct effective dictionaries for sparse modeling of several kind of signals such as audio, biosignals [5, 6]. Their ability to sparsify the signal, is based on the fact that they i) tend to decorrelate the components of a given signal ii) tend to redistribute the energy contained in the signal so that most of energy is contained in a small number of components [7]. The principal component analysis (PCA) transform, is a linear transform that completely decorrelates the signal components and concentrates the signal energy to a small number of coefficients. This is achieved by fitting a low-dimensional subspace to the data in a way that minimizes the l_2 approximation error [7]. More specifically, the PCA transform is defined by the eigenvalues of the covariance matrix. In mathematical terms, suppose that $\mathbf{R}$ corresponds to the covariance matrix of the signal:

$$\mathbf{R} = \mathbb{E}\left[\mathbf{x}\mathbf{x}^T\right]. \qquad (6)$$

By modeling breath sounds as an auto regressive - 1 process (AR-1) with parameter ρ [10] [2] the correlation matrix $\mathbf{R}$ can be written as a Toeplitz symmetric matrix with elements $\mathbf{R}_{i,j} = \rho^{|i-j|}$, $\forall i, j \in [1, \ldots, N]$.

Now, let $\mathbf{R} = \mathbf{E\Lambda E}^T$ be the eigenvalue decomposition of matrix $\mathbf{R}$. Matrix $\mathbf{E}$ is an orthogonal matrix with the eigenvectors of $\mathbf{R}$ and $\mathbf{\Lambda}$ is a diagonal matrix with the eigenvalues of $\mathbf{R}$ in decreasing order. The sparsity and the block sparsity in the PCA domain can be exploited by solving the convex optimization problems in eqs. (3) and (5) after setting $\mathbf{\Psi} = \mathbf{E}$. The breath signal in the native domain $\mathbf{x}$ can be then reconstructed as in the previous sections, by multiplying the resulted vectors $\hat{\mathbf{s}}_L$ and $\hat{\mathbf{s}}_{GL}$ with $\mathbf{E}$.

4. SIMULATION RESULTS

The focus of this study is to identify the benefits of applying different compressed sensing schemes, in the mhealth system energy efficiency. The proposed telemonitoring scheme are studied by using wheezing sounds from a database of prerecorded respiratory sounds. The database consisted of a total of 26 sounds that contained more than one uninterrupted interval of wheezing. These sounds were recorded using 11 kHZ sampling rate and 8-bit depth. Due to the lack of a single standard respiratory sound database, the recordings used in our study were drawn from multiple commonly referenced Internet sources [11–14].

4.1. Simulation Setup

We assume that each breath signal, is divided into segments of $N = 256$ samples. Each segment is encoded at the sensor side and the M encoded RLC measurements are transmitted to a BNC. The BNC can reconstruct the

[2]Higher order models can be also included but without having any significant impact on the performance of the PCA Group LASSO scheme

original signal by performing either i) the conventional LASSO in the DCT domain ii) the group LASSO in the DCT domain iii) the conventional LASSO in the DWT domain[3], iv) the group LASSO in the DWT domain, v) the proposed PCA LASSO vi) the proposed PCA group LASSO. The convex problems were solved by using the cyclic coordinate descent approach [9].

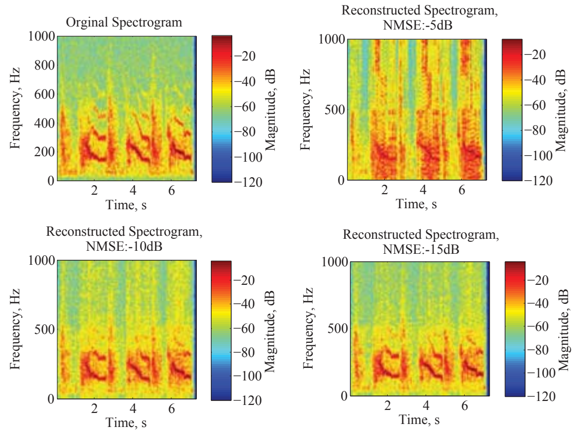
At this point it should be noted that many wheeze detection schemes (e.g., [3,15]) are based on the time frequency analysis of the breath sound signals. To be more specific, they make use of the short time Fourier Transform (STFT) and a set of rules related to the duration, the temporal and spectral continuity of peaks in expected frequency bands. To evaluate the diagnostic quality of the compressed breath recordings and the accuracy of the system, we make use of whether the wheezes can be identified in the spectrogram of the breath signal. An empirical study after using the aforementioned approaches for compression/reconstruction and wheeze detection, showed that the wheezes can be identified in the spectrogram of the reconstructed signal when the NMSE between the spectrogram of the original and reconstructed signals is less than -10 dB. In fig. 2 (a) we provide the spectrogram of a reconstructed breath sound that corresponds to a spectrogram NMSE of (a) -5 (b) -10 and (c) -15 dB. By inspecting this figure it is shown that the harmonic components originating from wheezing, clearly appear as continuous frequency peaks elevated against the noise of normal respiration in cases (b) and (c), while in case (a) they are not identifiable. The peaks of wheezing are localized along the frequency axis and spread in the time axis.

The energy efficiency (EE) of the considered mhalth system, after assuming that the transmit and receive power is equal to 3.8 mW and 4.6 mW respectively, the bit depth of the encoded measurements is $l_q = 8$ bits, while the duration of a packet transmission is 2.94 m s [16] [4] can be evaluated by:

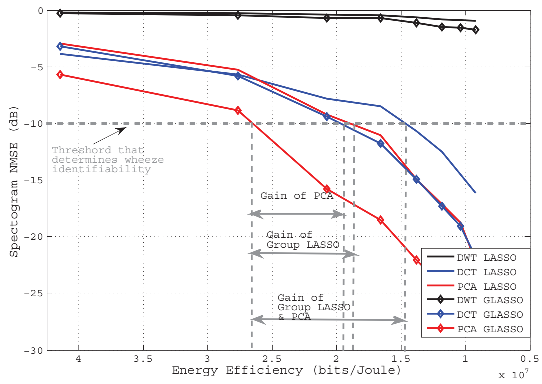
$$EE = \frac{Segment\ Bits}{Total\ Energy} = \frac{8N}{M(P_T + P_R)T_s/10}\ bits/Joule. \tag{7}$$

4.2. Performance Evaluation

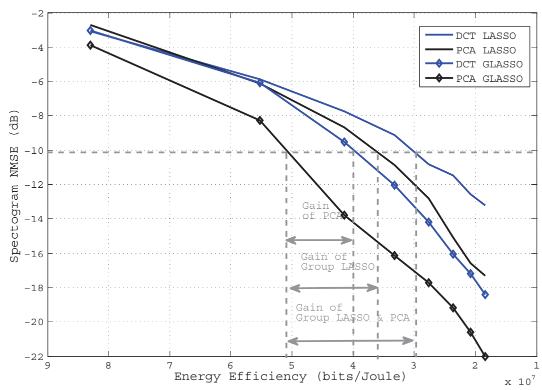
In Fig.2 (b), the obtained spectogram NMSE, is plotted against the energy efficiency values determined from the M transmitted RLC data by the sensor. A block length $d = 16$ was selected and the scaling rules for the parameter λ in the LASSO and group LASSO approaches



(a)



(b)



(c)

Fig. 2. *(a) Time frequency analysis of original and reconstructed signals (b) Evaluation in term of energy efficiency (c) Evaluation of the Quantization Effects*

[3]Note that for the study we utilized the orthogonal Daubechies wavelets (db 10).

[4]We have assumed packets with 14 bytes header and 80 bytes payload (10 audio samples/packet), and a data rate equal to 256 kbps [17].

follow the results of [18]. By inspecting this figure, it is clearly shown that the exploitation of group sparsity in the PCA domain increases the energy efficiency as compared to the DCT group LASSO and the conventional DCT LASSO cases by 38.5%, and 98% respectively. This is attributed to the fact that the PCA method results in much more sparse representations, while the exploitation of group sparsity reduces the degrees of freedom of the compressible signal, by allowing group of samples to have zero or non zero norm values. This property enable us to reduce, the number of measurements M required for a stable recovery of the breath signal at the BNC.

Finally, to evaluate the quantization effects in the recovery performance of the algorithms (i) - (iv), we conducted experiments assuming that the bit depth of the encoded measurements was $l_q = 4$ bits. The quantization was performed by applying the Lloyd max algorithm. Fig.2 (c) shows the obtained spectogram NMSE against the energy efficiency values determined from the M transmitted RLC data by the sensor. As compared to the previous case the EE values are doubled due to the reduced bit depth. The recovery in both the DCT and PCA domain remains almost unaffected as compared to the aforementioned case ($l_q = 8$). As a result, the exploitation of group sparsity in both the DCT and PCA domain is robust to quantization errors, increasing significantly the mhealth system energy efficiency. The robustness to quantization effects can be further increased by adopting the policies described in [19].

5. CONCLUSION

Wireless telemonitoring of breath sounds facilitate the early diagnosis and management of chronic inflammatory disease of the airways but introduces challenges related to the real time compression and transmission of this audio signals. To this end, we propose a novel Compressed Sensing framework that offers significant gains in terms of energy efficiency of the considered mhealth system, by exploiting the benefits of the group LASSO approaches in the eigen-spectrum domain. The presented schemes are going to be implemented on smartphones with Android operating system.

6. ACKNOWLEDGMENTS

This work has been funded by the H2020-PHC-2014-2015 Project MyAirCoach (Grant Agreement No. 643607).

REFERENCES

[1] F. Leahy R. Fenton H. Pasterkamp, A. Tal and V. Chernick, "The effect of anticholinergic treatment on postexertional wheezing in asthma studied by phonopneumography and spirometry," *Am. Rev. Respir. Dis.*, , no. 132, pp. 16–21, 1985.

[2] "iSonea: AirSonea Acoustic Respiratory Monitoring Device.A vailable online:," `http://isoneamed.com/products/devices/`, [Online; accessed 19-Nov-2013].

[3] Dinko Oletic, Bruno Arsenali, and Vedran Bilas, "Low-power wearable respiratory sound sensing," *Sensors*, vol. 14, no. 4, pp. 6535–6566, 2014.

[4] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 451–515, April 2000.

[5] A. Griffin, T. Hirvonen, C. Tzagkarakis, A. Mouchtaris, and P. Tsakalides, "Single-channel and multi-channel sinusoidal audio coding using compressed sensing," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 19, no. 5, pp. 1382–1395, July 2011.

[6] H. Mamaghanian, N. Khaled, D. Atienza, and P. Vandergheynst, "Compressed sensing for real-time energy-efficient ecg compression on wireless body sensor nodes," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 9, pp. 2456–2466, Sept 2011.

[7] R. Rubinstein, A.M. Bruckstein, and M. Elad, "Dictionaries for sparse representation modeling," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1045–1057, June 2010.

[8] Y.C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *IEEE Trans. Signal Process*, vol. 58, no. 6, pp. 3042–3054, 2010.

[9] Jerome Friedman, Trevor Hastie, Holger Hfling, and Robert Tibshirani, "Pathwise coordinate optimization," Tech. Rep., Ann. Appl. Stat., 2007.

[10] V. Iyer, P.A. Ramamoorthy, and Y. Ploysongsang, "Autoregressive modeling of lung sounds: characterization of source and transmission," *Biomedical Engineering, IEEE Transactions on*, vol. 36, no. 11, pp. 1133–1137, Nov 1989.

[11] "The R.A.L.E. Repository. Available online:," `http://www.rale.ca/`, [Online; accessed on 16-Nov-2013].

[12] "Stethographics Lung Sound Samples. Available online:," `http://www.stethographics.com/main/productsdownloads.html`, [Online; accessed on 16-Nov-2013].

[13] "Mediscuss Respiratory Sounds. Available online:," `http://www.mediscuss.org/downloads.php?do=file&id=2`, [Online; accessed on 12-Dec-2013].

[14] "East Tennessee State University Pulmonary Breath Sounds. Available online:," `http://faculty.etsu.edu/arnall/www/publichtml/heartlung/breathsounds/contents.html`, [Online; accessed on 12-Dec-2013].

[15] Styliani A. Taplidou and Leontios J. Hadjileontiadis, "Wheeze detection based on time-frequency analysis of breath sounds," *Computers in Biology and Medicine*, vol. 37, no. 8, pp. 1073 – 1083, 2007.

[16] LAN/MAN Standards Committee of the IEEE Computer Society, "IEEE Standard for Local and metropolitan area networks – Part 15.6: Wireless Body Area Networks," *IEEE Std 802.15.6-2012*, Feb. 2012.

[17] S. Movassaghi et al., "Wireless body area networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. PP, no. 99, pp. 1–29, 2014.

[18] D. Angelosante, J.A. Bazerque, and G.B. Giannakis, "Online adaptive estimation of sparse signals: Where rls meets the ℓ_1 -norm," *IEEE Trans. Signal Process*, vol. 58, no. 7, pp. 3436–3447, 2010.

[19] "Democracy in action: Quantization, saturation, and compressive sensing," *Applied and Computational Harmonic Analysis*, vol. 31, no. 3, pp. 429 – 443, 2011.