An Efficient Spectrum Sensing Scheme for Cognitive Radio

Samuel Cheng, Vladimir Stanković, and Lina Stanković

Abstract - The paper combines distributed source coding and compressive sampling for efficient spectrum estimation. Two or more cognitive radios sample the spectrum compressively and independently compress their observations using multiterminal source coding. A central hub collects compressed streams from these radios before performing joint multiterminal source decoding followed by iterative signal reconstruction. Simulation results are provided for two radios performing practical multiterminal source coding with uniform scalar quantization and systematic turbo codes for Slepian-Wolf coding and error protection.

S. Cheng is with the Dept. Electrical and Computer Engineering, University of Oklahoma, Tulsa, OK. Tel: +1 918 660-3234. Fax: +1 918 660-3238. Email: samuel.cheng@ou.edu. V. Stanković and L. Stanković are with the Dept. Electronic and Electrical Engineering, University of Strathclyde, Glasgow, UK. Tel: +44 141 548-2679. Fax: +44 141 552 4968. Email: vladimir.stankovic@eee.strath.ac.uk, lina.stankovic@eee.strath.ac.uk. EDICS:COM-CODE.

I. INTRODUCTION

Since the concept of cognitive radio (CR) was introduced, the topic has received much attention. The challenge is how CR devices, networks and services can co-exist with one another in the same spectrum space. This is achieved by CRs dynamically sensing the spectrum for 'unused' or white bands, and 'low energy use' or grey bands. In an environment where multiple CRs co-exist, possibly with other wireless devices, spectrum sharing could be managed by a central hub, based on the measurements received from all CRs. This will avoid spectrum misuse, competition among CRs, and will enable prioritization among CRs, based, for example, on their subscription.

In this paper, we consider the scenario whereby CRs sense the spectrum environment to create a central spectrum database and exploit spectrum holes. Two key issues are high complexity and power consumption of the sensing process, and large bandwidth needed for delivering the observations to a central hub. Since acquired samples at different CRs in a localized area are correlated, distributed source coding (DSC) [1], or more specifically, multiterminal (MT) source coding [2], is the most effective compression method. Moreover, since the spectrum is expected to be sparse, compressive sampling (CS) [3] techniques can be exploited.

In [4], joint sparsity between measurements of different terminals was discussed and several models proposed. In this paper, we assume that measured signals at each CR are sparse, but we do not address joint sparsity of the signals at different CRs to further reduce the number of needed measurements. Instead, we collect enough samples at each CR to independently recover its observation, and employ MT source coding to exploit correlation among the samples at different CRs. When employing MT source coding on compressively sampled data one problem lies in trading off acquisition complexity and compression. Indeed, if we acquire less samples, we will have to spend more bits to represent them. Based on the result of [5] for CS of noisy measurements, we derive a lower bound on the number of measurements needed for successful reconstruction as a function of MT coding distortion and rate. We focus on the case with only two CRs because practical MT code designs for more than two terminals are very complex. However, the system can be generalized to multiple CRs by grouping the CRs into pairs and processing them independently, though with reduced performance [1]. We assume perfect synchronization among the sensors; different sampling times at the two CRs will only result in the total rate increase.

Our main contribution is improving bandwidth efficiency and power usage of CRs with DSC in a typically noisy CR channel, and using CS to reduce acquisition time compared to the case where only DSC is used. Indeed, recently, the potential of using only CS for wideband CR spectrum sensing was shown [6]. However, exploiting MT source coding on compressively sensed data, to the best of our knowledge, has not been reported yet.

II. BACKGROUND

Compressive sampling: CS is a novel framework that enables sampling below the Nyquist rate, without (or with small) reduction in reconstruction quality. It is based on exploiting sparsity of the signal in some domain. Let $\mathbf{x} = \{x[1], \ldots, x[N]\}$ be a set of N samples of a real-valued, discrete-time random process X. Let $\mathbf{x} = \mathbf{\Psi s} = \sum_{i=1}^{N} s_i \psi_i$, where $\mathbf{s} = [s_1, \ldots, s_N]$ is an N-vector of weighted coefficients $s_i = \langle \mathbf{x}, \psi_i \rangle$, and $\mathbf{\Psi} = [\psi_1 | \psi_2 | \cdots | \psi_N]$ is an $N \times N$ orthonormal basis matrix with ψ_i being the i-th basis column vector. Vector $\mathbf{x}$ is considered K-sparse in the domain $\mathbf{\Psi}$, for $K \ll N$, if only K out of N coefficients of $\mathbf{s}$ are non-zero. CS removes "sampling redundancy" when sensing sparse signals by requiring only M samples of the signal, where $K < M \ll N$. Let $\mathbf{y}$ be an M-length measurement vector given by: $\mathbf{y} = \mathbf{\Phi x}$, where $\mathbf{\Phi}$ is an $M \times N$ measurement matrix. Signal $\mathbf{x}$ can be recovered losslessly from $M \approx K$ or slightly more measurements if the measurement matrix $\mathbf{\Phi}$ is properly designed [3].

Unfortunately, reconstruction of $\mathbf{x}$ from vector $\mathbf{y}$ of M samples is not trivial. The exact solution is NP-hard and consists of finding the minimum l_0 norm, i.e., the number of non-zero elements. However, excellent approximation can be obtained via the l_1 norm minimization. The l_1 norm minimization problem, basis pursuit, can be solved using linear programming with $O(N^3)$ complexity, and it requires in general more measurements than l_0 minimization; faster solutions were proposed at the expense of slightly more measurements. In our experiments, we use the modified LASSO reconstruction [7], which minimizes: $\hat{\mathbf{s}} = \arg\min ||s'||_1, \quad \text{subject to} \quad ||\mathbf{\Phi \Psi s'} - \mathbf{y}||_2 \leq \epsilon$, where ϵ is a threshold which trades-off precision and execution time. This algorithm is robust even when the measurements are corrupted by measurement noise.

MT source coding: Slepian and Wolf [8] first considered a simple DSC problem of separate lossless compression of two correlated sources, and showed that separate encoding and joint decoding suffer no rate loss compared to the case when the sources are compressed jointly. Berger [2] introduced the MT source coding problem by considering a more general case of separate *lossy* source coding.

In indirect MT source coding a single source is observed by two terminals, which compress their noisy observations independently before sending the result to a central point for joint decoding. Let U be the source to be reconstructed at the central point. It is observed by two distinct terminals, which are provided only with a noisy version, $V_i = U + Z_i$, $i = 1, 2$. Each terminal compresses its observation without knowing the observation present at the other terminal. The compressed bitstreams are sent to the central point over noiseless channels (assuming appropriate use of channel codes). The central point performs joint decoding and reconstructs source $\hat{U}$. Providing practical code designs capable of approaching these limits is a challenge [1]. In [9], a design was proposed that exploits quantization of the source observations at the terminal followed by Slepian-Wolf (SW) coding of the quantized indices via turbo or LDPC codes.

III. SYSTEM DESCRIPTION

Two cognitive radios, CR1 and CR2, sense the bandwidth and send their measurements to a central point, which in turn reconstructs the spectrum, and decides which bands are occupied. Besides the power spectral density occupied by other radios, each CR also senses i.i.d. additive white Gaussian noise (AWGN). Thus, the task of each CR is to sample the spectrum, compress the noisy values and communicate them to the central point.

Let $x(t)$ be the signal that occupies our chosen frequency band, whose frequency response is to be estimated. Then, CRi, $i = 1, 2$, receives: $x_i(t) = x(t) + n_i(t)$, where $n_i(t)$ is zero-mean AWGN independent of $x(t)$ and $n_j(t)$, $j \neq i$. Let x_s and x_{s_i} be vectors of N equidistant samples of $x(t)$ and $x_i(t)$, respectively, sampled at or above the Nyquist sampling rate. Then by the Nyquist Theorem, $x(t)$ can be reconstructed losslessly if all N samples are received. However, using the fact that x_s is sparse in the frequency domain, the sampling rate can be reduced significantly without sacrificing reconstruction quality exploiting the CS paradigm. Indeed, CRi can sample x_{s_i} in M points, $K < M \ll N$, as: $y_i = \Phi_i x_{s_i} = \Phi_i \Psi X_{s_i}$, where Φ_i is an $M \times N$ measurement matrix, Ψ is the inverse Fourier transform, and X_{s_i}, the Fourier representation of x_{s_i}, has only $K \ll N$ non-zero elements when noise-free.

Note that sampling is carried out in the time domain. $x_i(t)$ is down-sampled (non-uniformly) by setting measurement matrix to contain all zeros except one 1 in each of M rows, where the position of 1 is random in that row. Thus, y_i contains M random (not equidistant) samples of $x_i(t)$. Let us assume that CR1 and CR2 are perfectly synchronized via a centralized clock and use the same measurement matrix $\Phi_1 = \Phi_2 = \Phi$ which can be achieved by the same randomization seed. Then, y_1 and y_2 can be seen as two noisy replicas of the same source, hence they are correlated. This correlation can be exploited in compressing them via indirect MT source coding before transmission.

After acquisition, y_i is quantized sample-by-sample. To reduce complexity we employ uniform scalar quantization. Since the same scalar quantizers are used by both CRs, there will be remaining correlation among output quantizer indices, which can be exploited by SW coding. However, since least significant bits (LSB) are least correlated, we remove them before SW coding. LSB will be sent by one CR only. The remaining bitplanes are fed to the linear channel code for SW coding. In our simulations, we use the systematic punctured turbo code used in [9] to achieve symmetric compression rates among the two CRs. Due to the systematic nature of the employed SW code, the generated syndrome will consist of a portion of systematic bits and punctured "parity" turbo bits. Additional entropy coding can be employed on the systematic part to further reduce the rate. The resulting compressed bitstreams are sent to the central point after error protection by an appropriate channel code.

The central point collects signals from both CRs before attempting to estimate the spectrum. First, it performs indirect MT source decoding, that is, turbo decoding (possible entropy decoding) and dequantization to recover M-vectors $\hat{y}_1$ and $\hat{y}_2$ after padding the LSBs. Then, it performs minimum mean square error (MSE) estimation on $\hat{y}_1$ and $\hat{y}_2$ obtaining $\hat{y}$, estimation of x_s in the M sampling points. Note that when $n_1(t)$ and $n_2(t)$ are of the same power, this estimation boils down to computing the mean of the reconstructed $\hat{y}_1$ and $\hat{y}_2$. Finally, the decoder performs iterative reconstruction on $\hat{y}$ to estimate the spectrum. Note that this step may introduce significant distortion if the number of acquired measurements is not sufficient.

To reduce the acquisition complexity, it is desirable to minimize the number of required measurements. In [5], the influence of Gaussian measurement noise on the number of required measurements was studied, and the number of samples needed for successful reconstruction was approximatively lower bounded as: $M \geq \frac{2K \log(N/K)}{\log(1+SNR)}$, where K is the sparsity of the signal and SNR is measurement signal-to-noise ratio.

In our case, "measurement noise" at the decoder comes from noises $n_1(t)$ and $n_2(t)$ and from quantization noises. We can compute the expected distortion before the CS reconstruction as:

$D = \frac{1}{M}\sum_{j=1}^{M} E[(y(j) - \hat{y}(j))^2]$, where $y(j)$ and $\hat{y}(j)$, $j = 1, \ldots, M$, denote, respectively, M samples taken by the terminals assuming noise-free sensing and M reconstructed values after MT source decoding and minimum MSE estimation.

The distortion D can be seen as the power of the "measurement noise", thus, the above equation for the minimum number of measurements needed to be acquired by each CR becomes:

$$M \geq \frac{2K\log(N/K)}{\log(1 + P_y/D)}, \tag{1}$$

where P_y is the power of y.

The achievable compression sum-rate for the above distortion constraint D, assuming Gaussian i.i.d. sources, can be calculated as (see [9] and references therein):

$$R = R_1 + R_2 \geq \frac{1}{2}\log^+ [\frac{4P_y^2}{P_{n_1}^2 P_{n_2}^2 D(\frac{1}{P_y^2} - \frac{1}{D} + \frac{1}{P_{n_1}^2} + \frac{1}{P_{n_2}^2})^2}],$$

where P_{n_1} and P_{n_2} are power of the noises $n_1(t)$ and $n_2(t)$ on the measurements, respectively, and $\log^+ w = \log w$ for $w > 0$ and $\log^+ w = 0$ otherwise.

From the above equation and (1) we can find the lower bound on the minimum number of measurements M as a function of the total transmission rate $R_t = R/C$ over a channel with capacity C assuming the quadratic Gaussian case:

$$R_t = \frac{1}{2C}\log^+ \frac{4P_y(10^{\sigma/M} - 1)}{(A - P_{n_1}P_{n_2}\frac{10^{\sigma/M}-1}{P_y})^2}, \tag{2}$$

where $\sigma = 2K\log(N/K)$, $A = P_{n_1}/P_{n_2} + P_{n_2}/P_{n_1} + P_{n_1}P_{n_2}/P_y^2$. Note that there exists a trade-off between the number of measurements and the rate. That is, if we want to acquire less number of measurements, we would need more bits to represent them. We point out that, as the result of [5], the lower bound (1) is only an information-theoretical approximation, and thus is expected to be loose.

In general, $\Phi_1 \neq \Phi_2$, which can also come from asynchronous sampling. In this case, each CR would acquire less number of measurements, which means lower acquisition complexity. However, the correlation between the measurements of the two radios has decreased, thus the compression gain goes down. If Φ_1 and Φ_2 differ in all rows, then the number of necessary measurements will be minimized, but the MT compression rate will be maximized. Indeed, the increase in the source coding rate is needed to compensate for the difference between measured values at different time.

If the level of noise in the channel is very low, then each CR would need to acquire only slightly more than $M/2$ measurements, but these measurements would not be compressible (unless $x(t)$ is a slow-varying signal, such that samples taken at different time are mutually correlated). On the other hand, with $\Phi_1 = \Phi_2$, each CR needs M measurements, which can be efficiently MT compressed.

It is possible to trade-off the rate and M by allowing Φ_1 and Φ_2 to differ in some rows only. Then only the measurements taken at the same time instances can be MT compressed and sent at the rate given by (2); the rest would be sent uncompressed in the case of i.i.d. signals. Note that the proposed scheme of CS followed by MT source coding can be applied to any measurement matrices (e.g., Gaussian [3]) as long as the RIP condition is met.

IV. SIMULATION RESULTS AND DISCUSSION

We generate frequency response of x_s, such that K occupied portions of its spectrum are randomly allocated. We set $K = 25$, $N = 1000$, and the signal power to 100 and change noise power $P_{n_1} = P_{n_2}$ to obtain different SNRs. We set $\Phi_1 = \Phi_2$ and use 4 bits to represent each sample, and quantization step size of our scalar quantizer is 0.17. The compression sum-rate is set to 5.2 bit/sample (b/s), or in average 2.6 b/s at each CR, which is 1.27 b/s away from the theoretical limit calculated in Section III. This loss mainly comes from the use of a simple uniform scalar quantizer and short turbo codeword sizes. Puncturing of the turbo code output bits can be used to generate different SW rates. The same turbo code of rate 1/2 is used as the error-protection code against transmission noise. We note that the error-protection is not optimized as this is not the focus of the paper. We used the CVX package [10] for implementation of the LASSO algorithm with a threshold of $\epsilon = 10^{-5}$, which was found empirically as the best value for all SNRs. The MT compressed streams are sent over an AWGN with SNR=10 dB.

As performance indicator measure, we use MSE between the original spectrum and Fourier transform of the reconstructed $\hat{x}_s$: $MSE = 1/N \sum_{j=1}^{N} E[(\hat{X}_s(j) - X_s(j))^2]$. Fig. 1 shows MSE, averaged over a number of simulations, as a function of rate reduction factor per CR for three different values of SNRs in the channel. Rate reduction factor was calculated as the ratio of the number of bits sent by one CR after CS and MT source coding (including error protection), to the number of needed bits needed to represent measurements if CS, MT source coding, and error protection were not applied. Since the ratio is much less than 1, huge bandwidth gain is realized. The obtained points correspond to the number of collected measurements $M = [80, 100, \ldots, 200, 300]$. Note that we fixed the coding rate, and change the total transmission rate by varying the number of collected CS samples. Curve D represents the limit of the system found in Section III, assuming ideal SW coding and full sampling (no CS and residual errors) for SNR=20 dB (the difference for SNR of 24 and 28 dB is negligible). Note that D is shown as a straight line in Fig. 1 independent of the x-axis. For the practical curves, the remaining distortion comes from SW coding errors and non-ideal CS reconstruction.

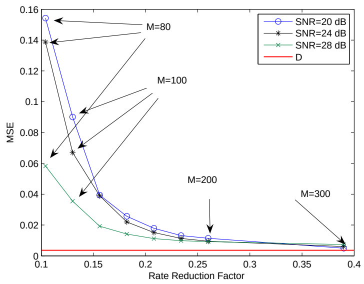

Fig. 1. MSE distortion of the reconstructed frequency response vs. rate reduction factor defined as the ratio of the number of bits sent by one CR after CS and MT source coding, to the number of bits that the CR would send if CS and MT source coding were not applied. D denotes the limit of the system assuming ideal SW coding.

It can be observed that with $M = 80$ significant distortion is present. But, after collecting $M = 200$ measurements, further measurements provide only small distortion reduction. Thus, with the increase of M the number of errors due to CS decreases, but the loss due to practical SW coding remains. The lower bound on the number of measurements calculated by (1), for this setup was between 52 and 55 depending on the SNR. Thus, this bound is very loose. For low M (80 or 100), residual SW coding errors are often present resulting in high distortion especially for low SNRs (20 and 24 dB). This is because the turbo code, employed for SW coding, does not perform well for small channel codeword sizes. As the channel codeword size increases, the performance improves. Note that the curves approach the ideal curve D but will never meet it due to SW coding errors.

Fig. 2 shows MSE as a function of SNR in the channel for three different values of M It can be seen from the figure that, as expected, MSE decreases with the increase of SNR in the channel for all values of M. However, there is a large jump in distortion when M is decreased from 160 to 120. On the other hand, decreasing M from 200 to 160 leads to a much smaller distortion increase. Note that the limit of the system D slowly decreases with the increase of SNR.

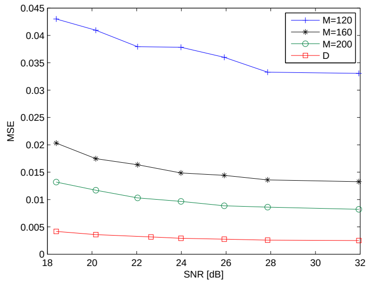

Fig. 2. MSE distortion of the reconstructed frequency response vs. SNR in the channel for three different numbers of acquired samples M. D denotes the limit of the system assuming ideal SW coding.

Fig. 3 shows the compression sum rate in b/s as a function of M under an MSE distortion constraint of 0.1. It can be seen that the required SW coding rate decreases with M as expected, because of small turbo codeword size. Note that for SNR=20 dB and $M = 80$ it was not possible to obtain distortion below 0.1.

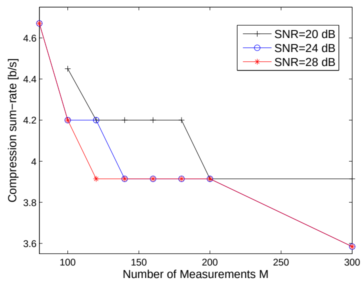

Fig. 3. The MT compression sum-rate vs. the number of collected measurements M for three different SNR's. Note that each measurement is represented by four bits before SW coding.

References

[1] Z. Xiong, A. Liveris, and S. Cheng, "Distributed source coding for sensor networks," *IEEE Signal Proc. Magazine,* vol. 21, pp. 80–94, Sept. 2004.

[2] T. Berger, "Multiterminal source coding," *The Information Theory Approach to Communications,* G. Longo, Ed., New York: Springer-Verlag, 1977.

[3] R.G. Baraniuk, "Compressive Sensing [Lecture Notes]," *IEEE Signal Proc. Magazine,* vol. 24, pp. 118–121, July 2007.

[4] M.F. Duarte, S. Sarvotham, D. Baron, M.B. Wakin, and R.G. Baraniuk, "Distributed compressed sensing of jointly sparse signals," in *Proc. 2005 Asilomar Conf.,* Pacific Grove, CA, Oct.-Nov. 2005.

[5] S. Sarvotham, D. Baron, and R.G. Baraniuk, "Measurements vs. bits: Compressed sensing meets information theory," in *Proc. Allerton Conf.,* Monticello, IL, Sept. 2006.

[6] Z. Tian and G.B. Giannakis, "Compressed sensing for wideband cognitive radios," in *Proc. IEEE ICASSP*, Honolulu, Hawaii, Apr. 2007.

[7] R. Tibshirani, "Regression shrinkage and selection via the Lasso," *J. of the Royal Statistical Society,* 1996. 58(1): pp. 267–288.

[8] D. Slepian and J.K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory,* vol. IT-19, pp. 471–480, July 1973.

[9] Y. Yang, V. Stanković, Z. Xiong, and W. Zhao, "On multiterminal source code design," *IEEE Trans. Inform. Theory,* vol. 54, pp. 2278–2302, May 2008.

[10] M. Grant, S. Boyd, and Y. Ye, "CVX: Matlab software for disciplined convex programming," available at http://www.stanford.edu/boyd/cvx.