

Amplify-and-Forward Compressed Sensing as an Energy-Efficient Solution in Wireless Sensor Networks

Joan Enric Barceló-Lladó, Antoni Morell, and Gonzalo Seco-Granados

Abstract—In this paper, we propose a novel distributed compressed sensing transmission scheme, which is referred to as amplify-and-forward compressed sensing (AF-CS), in order to improve the existing tradeoff among reconstruction error, energy consumption, and resource utilization. The goal is twofold. First, to take advantage of the time correlation in order to produce sparse versions of the signal vector, which collects the transmitted signals of all the sensors. Second, to benefit from the nature of the multiple access channel in order to perform random measurements of the signal vector. Additionally, a simple model that accurately approximates the distortion introduced by the proposed scheme is presented. This model is then used to select the number of active nodes and relays based on a cost function that controls the tradeoff between reconstruction error and energy consumption. Simulation results show that the AF-CS outperforms other techniques in terms of distortion and number of transmissions, providing simultaneously, energy savings and a significant reduction in the number of channel uses.

Index Terms—Compressed sensing, decentralized communications, energy efficiency, wireless sensor networks.

I. INTRODUCTION

WIRELESS Sensor Networks (WSNs) design is currently one of the most challenging topics in the field of wireless communications. In particular, WSNs are severely energy-constrained because they consist of many small, cheap and power limited nodes whose batteries cannot be recharged in most cases. Hence, the application of energy-efficient algorithms turns out to be crucial.

Interestingly, sensor readings tend to be highly time-correlated in many scenarios because of oversampling. Note that even sampling at the widely accepted Nyquist frequency, which is useful to capture the maximum possible variations in the signal of interest, there may be time windows where the signal fluctuates less and, in fact, this consideration applies to most detection and many monitoring scenarios. In some occasions the sensor readings are also space-correlated (as it happens in most of the environmental monitoring applications)

Manuscript received October 28, 2013; accepted January 14, 2014. Date of publication January 28, 2014; date of current version March 24, 2014. This work was supported by the Spanish Research and Development under Project TEC2011-28219. The associate editor coordinating the review of this paper and approving it for publication was Dr. Thilo Sauter.

The authors are with the Telecommunications and Systems Engineering Department, Universitat Autònoma de Barcelona, Cerdanyola del Valles 08193, Spain (e-mail: joanenic.barcelo@uab.es; antoni.morell@uab.es; gonzalo.seco@uab.es).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSEN.2014.2303080

and this can also be considered in extensions to the ideas presented in this paper.

In this context, *Compressed Sensing* (CS) appears to be a good candidate in order to exploit signal correlations in the time domain. In a nutshell, CS permits the recovery of a given signal $\mathbf{x} \in \mathbb{R}^S$ from a small number of measurements when some conditions are satisfied. Specifically, this is possible when \mathbf{x} can be accurately (or exactly) represented as a linear combination of K vectors taken from a proper basis Ψ [1]. If so, the vector \mathbf{x} is said to be compressible or that it has a K -sparse representation in Ψ . Roughly speaking, the CS theory studies how to recover the signal \mathbf{x} from a reduced set of measurements. In particular, a central result in CS theory states that the original signal can be estimated with arbitrarily small error when the number of measurements is proportional (up to a logarithmic factor [2]) to the information contained in the signal \mathbf{x} , i.e. K , instead of proportional to the number of samples S .

A. Compressed Sensing in WSNs

Compressed Sensing (CS) has been successfully applied in WSNs for localization [3], detection [4] and in general for data collection [5]–[8] purposes in the recent years. In this paper we focus on data collection, which embraces monitoring but also some event detection applications, and our aim is to reduce the overall power consumption in the network. Motivated by the fact that radio transmission and reception is the main sink of energy (or at least a very important one) in the sensor nodes [9], we explore how to benefit from the compressed sensing theory in order to transmit as little as possible. To the best of the authors' knowledge, the most representative techniques in the literature that use compressed sensing for data collection purposes in WSNs have already been cited and suffer from two important drawbacks: i) either they assume impractical hypotheses about the sparsity of the signal of interest and/or ii) the resulting number of transmissions is high.

Probably one of the most relevant works (and also one of the most referenced) in the field of CS applied to WSNs that circumvents in part these drawbacks is [6], where the authors first introduced what they termed *Compressive Wireless Sensing* (CWS). In brief, all S nodes that constitute the WSN transmit their processed data at the same time and coherently towards one fusion center in order to compute one projection. In this

case the projection process is also in charge of representing the sensor readings in the required sparse basis and therefore, once R projections are available at the fusion center, it is able to recover an approximation of the original signal from its sparse representation with some bounded distortion. Consequently, this scheme uses $S \times R$ transmissions and R channel uses (one channel use per iteration).

B. Our Contribution

In this paper, we build upon the CWS scheme in the sense that we also assume analog phase-coherent transmissions and, as in CWS, the multiple access channel plays an important role. The proposed solution is termed Amplify-and-Forward Compressed Sensing (AF-CS) and has two main differences with respect to CWS, namely: i) the sparse representation of the sensor readings is obtained individually at the different nodes by means of a Conditional Downsampling Encoder (CDE) - Predictive Decoder (PD) pair [10] that exploits the inner time-correlations in order to reduce the number of necessary transmissions and ii) we consider a two-hop transmission from the subset of K active nodes that aim to send their readings to the fusion center through a subset of R relying nodes. Thanks to this configuration, we obtain R projections (in the terminology of CWS) with a single channel use. In other words, the proposed AF-CS scheme naturally and elegantly copes with the multi-user interference in its first hop transmission. Afterwards, R additional transmissions are required to send the projections to the fusion center, so the total number of channel uses is nearly the same as in the CWS solution. However, the total number of transmissions is $(K + 1) \times R$ and thus, there is a reduction with respect to the CWS case as far as $S > K$.

Being our main contribution the design of a complete energy-efficient solution for WSNs based on CS, the partial contributions in this paper are: i) performance analysis and design of the CDE-PD scheme (based on the previous authors' work in [10]), ii) performance analysis and design of the CS part, and iii) selection of the system parameters in order to operate in the desired trade-off between energy consumption and accuracy in the reconstructed signal.

C. Notation

Boldface upper-case letters denote matrices, boldface lower-case letters denote column vectors, and italics denote scalars. $(\cdot)^T$, $(\cdot)^*$, $(\cdot)^H$ denote transpose, complex conjugate, and conjugate transpose (Hermitian) respectively. $[\mathbf{X}]_{i,j}$, $[\mathbf{x}]_i$ is the $(i$ th, j th) element of matrix \mathbf{X} and i th position of vector \mathbf{x} respectively. $[\mathbf{X}]_i$ is a row vector that contains the i th row of \mathbf{X} . $(\cdot)^*$ denotes the optimal value. Let \mathbf{a}_K be a K -sparse approximation of \mathbf{a} . $|\cdot|$ is the absolute value. $\|\mathbf{a}\|_1$ and $\|\mathbf{a}\|$ mean the l^1 -norm and the Euclidean norm of \mathbf{a} respectively. The notation $\|\mathbf{A}\|$ indicates the Frobenius norm of a matrix \mathbf{A} . \mathbf{I}_S is the identity matrix of dimension S . $\mathbb{E}[\cdot]$ is the statistical expectation. Function $\text{erf}(\cdot)$ represents the error function. $\mathcal{N}(\boldsymbol{\mu}, \mathbf{R})$ is a Gaussian vector distribution with mean $\boldsymbol{\mu}$ and covariance matrix \mathbf{R} , σ_x^2 is the variance of \mathbf{x} . The notation \hat{x} denotes the estimation of the scalar x .

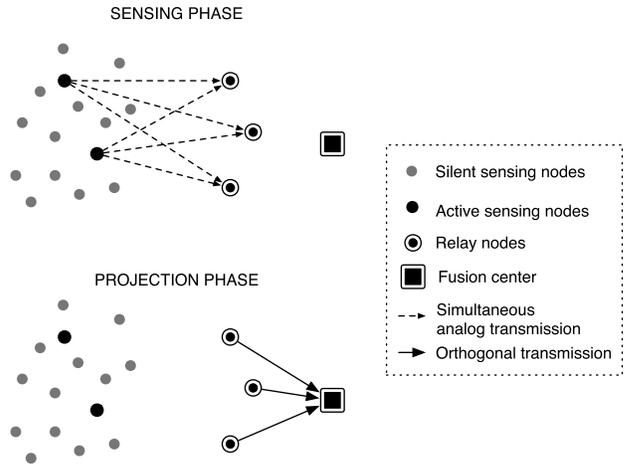


Fig. 1. WSN scenario composed by K active sensing nodes, Q sleep sensing nodes, R relay sensors, and one fusion center. The signal reconstruction phase is carried out by the fusion center after the projection phase.

D. Organization of the Paper

The rest of the paper is organized as follows. In Section II we present the system model and the assumptions considered throughout the paper. Section III explains the proposed AF-CS algorithm and describes each of its phases. The expressions of the distortion introduced by the AF-CS decoder are detailed in Section IV. The design of the network following a given cost function and its subsequent optimization are presented in Section V. Simulation results are shown in Section VI, and conclusions are drawn in Section VII.

II. SYSTEM MODEL

The proposed scheme implements a two-hop WSN (see Fig. 1) that monitors a given physical scalar magnitude (e.g., temperature or humidity) and is composed of:

- A set \mathcal{R} (of cardinality R) of *Amplify-and-Forward* (AF) relay sensors, connected wirelessly to one fusion center.
- A set $\mathcal{K}(n)$ (of average cardinality K) of *active sensors* transmitting at sample time n and connected (wirelessly) to \mathcal{R} . The remaining sensors $\mathcal{Q}(n)$ (of average cardinality Q) stay in *sleep mode*.
- The set \mathcal{S} of S sensing nodes contains all the nodes in the network, i.e., $\mathcal{S} = \mathcal{K}(n) \cup \mathcal{Q}(n) \cup \mathcal{R}$.

The two-hop transmission scheme works as explained next: First, the sensors in $\mathcal{K}(n)$ send their measurements to the relay nodes in \mathcal{R} by means of a simultaneous coherent analog transmission where signal overlapping is allowed. Second, the relay nodes retransmit the received signal to the fusion center using orthogonal transmissions. This transmission scheme is summarized in Fig. 1 and the details are explained in Section III.

We assume that the sensor readings are time correlated and we model them as an S -dimensional stochastic process,

$$\mathbf{X} = [\mathbf{x}(1) \mathbf{x}(2) \dots \mathbf{x}(N)], \quad (1)$$

where $\mathbf{x}(n) = [x_1(n), x_2(n) \dots x_S(n)]^T$ and $x_s(n)$ denotes the reading of the s th sensor at the sample time n and N denotes the number of time samples in the observation window.

Also, we assume that the channel matrix, here relabeled as *sensing matrix* $\Phi \in \mathbb{R}^{R \times S}$, follows the Gaussian measurement ensemble, where

$$[\Phi]_{i,j} \sim \mathcal{N}(0, R^{-1}). \quad (2)$$

Note that this is not an idealized channel assumption. For example, the well-known MIMO i.i.d. Rayleigh fading model has Gaussian real and imaginary parts. This fading model is valid whenever the antenna arrays are sparsely spaced or the transmissions take place in the so-called richly scattered environment [11, Sec. 7.3.8]. Since sensors will be placed in general more than half a wavelength apart, a low to moderate scattering environment suffices to validate the MIMO i.i.d. fading model. In this case, transmissions using only the in-phase or the quadrature component have Gaussian channel coefficients.

Note also that the variance of every element in the sensing matrix is arbitrarily set to R^{-1} since it follows the convention in the literature in order to maintain the relation $\mathbb{E}[\|\Phi \mathbf{x}\|^2] = \mathbb{E}[\|\mathbf{x}\|^2]$ for an arbitrary vector \mathbf{x} and does not affect to the generality of the model since the channel gain can be adjusted at the receiver if necessary.

III. AMPLIFY-AND-FORWARD COMPRESSED SENSING

This section details our proposed CS scheme, i.e., AF-CS strategy. It has three phases: *i)* the *sensing phase*, *ii)* the *projection phase*, and *iii)* the *signal reconstruction phase*.

A. Sensing Phase

Compressed sensing exploits sparsity to acquire high-dimensional signals that can be represented in a low-dimensional subspace. A priori, a vector with readings from a certain physical phenomena is not sparse in general but expected to be correlated in time and space. Thus, correlated signals may have a (pseudo)sparse representation if they are expressed in a proper basis $\Psi \in \mathbb{C}^{S \times S}$. In order to carry out this task, some central entity is needed to gather all the measurements (i.e., collect $\mathbf{x}(n)$) and compute all the transformed coefficients $\Psi \mathbf{x}(n)$. This is not a problem in centralized applications, e.g. image processing, but the extension to decentralized applications is not straightforward.

In a WSN scenario, one possibility in order to compute the transformed coefficients is that all the S *sensing nodes* transmit their readings along several iterations towards a fusion center in order to get a sparse version of $\mathbf{x}(n)$ (as in [6]). However, note that this approach is signaling intensive and highly energy consuming in terms of number of transmissions. In order to overcome this problem, one possible solution is to artificially create a sparse representation of $\mathbf{x}(n)$ with only K loaded entries and to set the rest to zero. Specifically, we propose to replace the linear transformation $\Psi \mathbf{x}(n)$ by a non-linear *Conditional Downsampling Encoder* (CDE), which exploits the time correlation properties of the signal. As we show next in Section IV, the distortion introduced for high-correlated signals is quite low, even for a high downsampling rate γ .

Algorithm 1 Sensing nodes

```

for  $n = 1$  to end do
  during the sensing phase
  for each  $s \in S$  do
    get the  $s$ th measurement  $x_s(n)$ .
    compute linear prediction  $\hat{x}_s(n)$  as  $\hat{x}_s(n) = \mathbf{w}^H \tilde{\mathbf{x}}(n)$ .
    compute  $x_s(n) - \hat{x}_s(n)$ .
    if  $|x_s(n) - \hat{x}_s(n)| > \Delta$  then
      active sensor mode (i.e., belongs to  $\mathcal{K}(n)$ )
      broadcast  $[\mathbf{x}_K(n)]_s = x_s(n)$ .
      store  $x_s(n)$ .
    else
      stay in sleep mode (i.e., belongs to  $\mathcal{Q}(n)$ )
      store  $\hat{x}_s(n)$ .
    end if
  end for
end for.

```

Shortly, the CDE uses the time correlation of the signal in order to decide whether the current sample should be transmitted or not. In particular, the CDE uses a linear prediction of $x_s(n)$ denoted as $\hat{x}_s(n)$. The value $\hat{x}_s(n)$ is compared then to the signal of interest $x_s(n)$. If the absolute value of the difference is higher than a given threshold Δ , the encoder transmits the sample. Otherwise, the transmission is blocked. It is important to remark that the value of Δ can be chosen to have K active sensors in mean, as it is detailed in Section IV.

Therefore, a prediction $\hat{x}_s(n)$ of each element of $\mathbf{x}(n)$ based on the N past readings of each sensor is required. Note that thanks to this mechanism a sparse version of $\mathbf{x}(n)$, i.e., $\mathbf{x}_K(n)$, containing (on average) only the K most relevant readings with low distortion is obtained, i.e.,

$$[\mathbf{x}_K(n)]_j = \begin{cases} x_j(n) & \text{if } j \in \mathcal{K}(n) \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

In this work, a linear predictor $\hat{x}_s(n)$ is considered. In particular, the *Linear Wiener Filter* (LWF) is selected due to its optimality in terms of Mean Squared Error (MSE). Mathematically, $\hat{x}_s(n) = \mathbf{w}^H \tilde{\mathbf{x}}(n)$, where $\mathbf{w} = \mathbf{R}^{-1} \mathbf{r}$ is the N -dimensional LWF solution, \mathbf{r} is known and denotes the $N \times 1$ *cross-correlation* vector between the past stored samples and the desired measurement $x_s(n)$, and $\tilde{\mathbf{x}}(n) \in \mathbb{R}^N$ is the *observation vector* that collects the N last outputs of the CDE. It is defined as:

$$[\tilde{\mathbf{x}}(n)]_j = \begin{cases} x(n-j) & \text{if } s \in \mathcal{K}(n-j) \\ \hat{x}(n-j) & \text{otherwise} \end{cases} \quad (4)$$

Note that the straightforward option is to build the observation vector with past readings only. However, the predictions computed at the s th sensor and at the fusion center (necessary in the decoding process) would not be the same as far as the latter does not receive all the readings. Thanks to the observation vector defined in (4), we guarantee that both the s th sensor and the fusion center operate accordingly. Thus if the s th sensor transmits, $x_s(n)$ is stored in the observation vector at the sensing node. Otherwise, if the s th sensor is silent, it stores $\hat{x}(n)$ as detailed in Algorithm 1. Following this approach, only

Algorithm 2 Relay nodes

```

for  $n = 1$  to end do
  for each  $r \in \mathcal{R}$  do
    measure projection  $y_r(n) = [\Phi]_r \mathbf{x}_K(n) + w(n)$ .
    transmit  $y_r(n)$  to the fusion center.
  end for
end for.
    
```

the subset of sensors $\mathcal{K}(n)$ transmit their readings while the rest $\mathcal{Q}(n)$ are in sleep mode.

B. Projection Phase

First, let $y_r(n)$ define the received signal at the r th relay as

$$\begin{aligned}
 y_r(n) &= \sum_{s \in \mathcal{K}} [\Phi]_{r,s} x_s(n) + w(n), \\
 &= [\Phi]_r \mathbf{x}_K(n) + w(n), \quad \text{for } r = 1, \dots, R \quad (5)
 \end{aligned}$$

where $[\Phi]_{r,s}$ models the flat-fading channel from the active sensor s towards the relay r defined in (2) and $w(n)$ models the AWGN with zero mean and variance σ_w^2 .

Note that $y_r(n)$ is in fact the projection of $\mathbf{x}_K(n)$ into $[\Phi]_r$ and thus, from a medium access control point of view, there is no need to make orthogonal transmissions during the *sensing phase*, so this phase requires one channel use according to (5). Differently, during the *projection phase*, it is assumed that the *relay nodes* transmit through R orthogonal channels to send a coded version of all the projected values $y_r(n)$. Therefore, this phase requires R channel uses as detailed in Algorithm 2. The sensing phase and the projection phase are graphically summarised in Fig. 1.

C. Signal Reconstruction Phase

First, note that perfect Channel State Information (CSI) is assumed at the fusion center. Although this is a classical assumption in the communication literature, it may be hard to satisfy in some specific scenarios that are power limited. Hence, some energy efficient techniques to estimate the channel matrix are needed [12]. However, in case that only an imperfect channel estimation of the type

$$\hat{\Phi} = \Phi + \Sigma \quad (6)$$

is available, where $\Sigma \in \mathbb{R}^{R \times S}$ is a random matrix with i.i.d. Gaussian entries with zero mean and variance σ_Σ^2 , the authors in [13] show that an upper bound for the quadratic error at the decoder grows linearly with the power of the distortion σ_Σ^2 .

The fusion center gathers all the received measurements in the R dimensional projection vector, denoted as $\mathbf{y}(n) = [y_1(n) \ y_2(n) \ \dots \ y_R(n)]^T$. Hence, the goal of the reconstruction phase is to recover an approximation of $\mathbf{x}_K(n)$, i.e. $\hat{\mathbf{x}}_K(n)$, given $\mathbf{y}(n)$ and Φ . The most prevalent decoder for the noiseless case in the CS literature is the l^1 -norm minimization program which is a convex relaxation of the original NP-hard problem (with the l^0 -norm) as in e.g., [1] and [2]:

$$\begin{aligned}
 \text{CSD : } & \underset{\hat{\mathbf{x}}_K(n) \in \mathbb{R}^S}{\text{minimize}} \quad \|\hat{\mathbf{x}}_K(n)\|_1 \\
 & \text{subject to} \quad \|\mathbf{y}(n) - \Phi \hat{\mathbf{x}}_K(n)\|_2 < \varepsilon. \quad (7)
 \end{aligned}$$

Algorithm 3 fusion center

```

for  $n = 1$  to end do
  while during the sensing phase do
    compute observation vector  $\tilde{\mathbf{x}}_s(n)$  for each sensor.
  end while
  while during the projection phase do
    collect  $y_1(n) \dots y_R(n)$  projection from relays.
  end while
  solve CSD in (7) to recover  $\hat{\mathbf{x}}_K(n)$ .
  obtain  $\hat{\mathbf{x}}(n)$  replacing the zeros by their estimations  $\hat{x}_s(n)$ .
end for.
    
```

where ε is an upper bound on the magnitude of the noise, i.e. $\varepsilon \geq \|w(n)\|_2$. Note that for the noiseless case, the constraint in the CSD turns into $\mathbf{y}(n) = \Phi \hat{\mathbf{x}}(n)$ and that other approaches in the literature such as the unconstrained LASSO [14], [15] can be applied too.

Let us remark that l^1 -norm based detection has been considered in this paper mainly because it is a well-established solution in the literature. However, other approaches such as the Bayesian Compressive Sensing [16] might be applied and interesting too, but this is out of the scope of this work.

Hence, the proposed AF-CS decoder has two main building blocks: *i*) a *CS decoder* (CSD) block that recovers $\hat{\mathbf{x}}_K(n)$, and *ii*) a *Predictive Decoder* (PD) that outputs $\hat{x}(n)$ (see Fig. 2).

Afterwards, the $S-K$ zeros in $x_K(n)$ are replaced by the corresponding predicted entries as detailed in Algorithm 3. Therefore, we propose to use an instance of the following PD,

$$\text{PD : } \begin{cases} [\hat{\mathbf{x}}(n)]_s = [\hat{\mathbf{x}}_K(n)]_s & \text{if } [\hat{\mathbf{x}}_K(n)]_s \neq 0 \\ [\hat{\mathbf{x}}(n)]_s = \hat{x}_s(n) & \text{otherwise.} \end{cases} \quad (8)$$

IV. DISTORTION ANALYSIS

In this section, we assess the error introduced by the AF-CS decoder, $\mathcal{D}_{\text{AF-CS}}$ seen as a combination of two partial decoders: *i*) the CS decoder (named CSD) and *ii*) the predictive decoder (PD). In the following, we analyze the distortion introduced by both steps separately.

A. Distortion Due to the CSD Step

In this subsection we focus our analysis on the distortion introduced by the CSD, denoted as \mathcal{D}_{CSD} , and defined as

$$\mathcal{D}_{\text{CSD}} = \mathbb{E} \left[\|\mathbf{x}_K(n) - \hat{\mathbf{x}}_K(n)\|_2^2 \right], \quad (9)$$

where our interest is to adjust the values of K , R , and S in order to satisfy an upper-bound of the distortion.

According to [14], and being σ_w^2 the noise variance in (5), the distortion can be upper-bounded by

$$\mathcal{D}_{\text{CSD}} \leq \left(\frac{K}{R} \sigma_w^2 \log S \right), \quad (10)$$

when the sensing matrix satisfies the *Restricted Isometry Property* (RIP). The RIP was first introduced by Candès and Tao in [17, Definition 1.1].

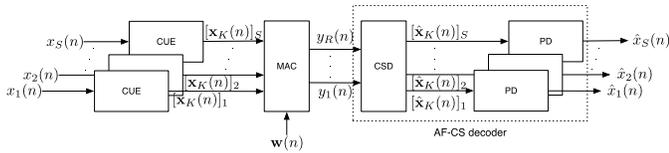


Fig. 2. Block diagram of the AF-CS transmission scheme.

If a given sensing matrix Φ verifies the RIP, it means that it behaves like a nearly orthogonal system but only for sparse linear combinations. In addition, it is shown in [2] and [18] that this condition suffices for the exact reconstruction of sparse linear combination of these vectors in the noiseless case.

The current state-of-the-art in CS theory [14] reveals that if the number of random measurements R is on the order of $K \log S$ (with $R \ll S$), it is possible to recover $\mathbf{x}_K(n)$ with an error bounded by (10). This condition was first introduced for the Fourier basis case in [2] and for the Gaussian ensemble in [19]. However, little more than $C_0 > 0$ is known. Because of this limitation, we can also find some practical results, e.g., that R should be between $3K$ and $5K$ [20], [18], and maybe others that we have unintentionally omitted. Existing bounds in the literature are quite heterogeneous and, in fact, it is still unknown whether or not perfect recovery (in the noiseless case) can be guaranteed when the number of measurements R is on the order of $K \log S$ [14], i.e., it is still an open problem in the CS literature.

Since no practical results to relate K , R and S are available, an empirical model to describe the distortion precisely has been developed. Let us motivate our results with the following experiment.

Experiment 1: Let us consider $S = 200$, $K = [0, S]$, and $R = [0, S]$. The CSD is applied to the noiseless case. In particular, we have used CVX, a package for solving convex programs [21], [22]. In our experiment, $\mathbf{x}_K(n)$ has been generated as an S dimensional all-zero vector except for K loaded entries with independent and Gaussian values with zero mean and unit variance, and randomly located along the dimension of $\mathbf{x}_K(n)$. The $R \times S$ matrix Φ has been generated following the Gaussian ensemble with entries distributed according to $\mathcal{N}(0, R^{-1})$. The probability of recovery has been averaged 5 times for each combination of K and R .

In Fig. 3 we represent the results of the Experiment 1. We plot the probability of recovery as a function of both K and R . The grey zone indicates that the probability of recovery is zero. The white zone means that perfect recovery is achieved. From the results, the reader can appreciate that the division of the recovery and non-recovery zones is well defined.

In the same plot, we also represent the theoretical bound $R > K \log S$ and the practical $R \in (3K, 5K)$. Note that the proposed bound accurately describes the result of the experiment for the entire range of values, while the other conditions only approximate the experimental results for low values of R and K .

Moreover, note from the results that two boundary points must be necessarily included. The first one is the trivial point $(K, R) = (0, 0)$, and all the proposed models already contain this point. The second point is $(K, R) = (S, S)$ since in

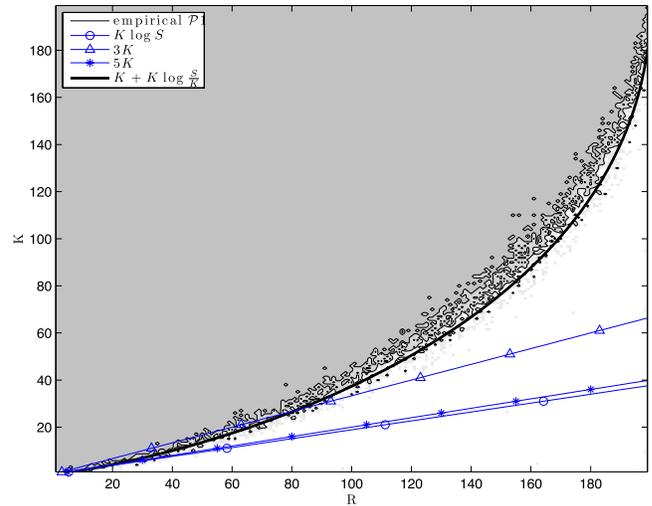


Fig. 3. Empirical results for Experiment 1. The empirical probability of recovery map for each value of $K, R \in 1, \dots, S$ is compared with different models in the literature. The white area means perfect recovery with high probability while the grey area means that the probability of recovery is zero. The area below the curves indicates which are the values (K, R) that the model relates to high probability of recovery. We omitted the parameter C_0 for simplicity.

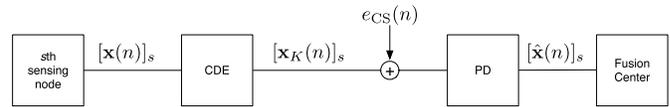


Fig. 4. Simplified CDE-PD transmission block diagram for a given sensor s with the presence of a perturbation error e_{CS} in the communication link.

this case Φ is a square full rank matrix (with almost sure probability [23]), and perfect recovery is achievable with $\hat{\mathbf{x}}_K(n) = \Phi^{-1}\mathbf{y}(n)$ for any value of K (even for $K = S$). However, none of the proposed methods contains this second point. In order to overcome this problem, we propose an empirical model called \mathcal{C}_{CS} ,

$$\mathcal{C}_{CS} : R > K + K \log \left(\frac{S}{K} \right) \quad (11)$$

which is plotted with double-width solid line in Fig. 3.

Note that it can be interpreted as an extension of the known condition $R > K \log \left(\frac{S}{K} \right)$ in [2] to take into account the entire set of values of K and R .

B. Distortion Due to the PD Step

In this section, we focus on the distortion introduced by the Predictive Decoder (PD) step, denoted as \mathcal{D}_{PD} and defined as,

$$\mathcal{D}_{PD} = \mathbb{E} \left[\|\mathbf{x}(n) - \hat{\mathbf{x}}(n)\|_2^2 \right]. \quad (12)$$

In the following, we focus without loss of generality on the MSE obtained in the reconstruction of the s th sensor. Hence, the studied scheme takes into account both the CDE (since it is a lossy encoder) and the PD and can be simplified as in Fig. 4, where $e_{CS}(n)$ is an external source of distortion, in our case, the error propagated from the CSD phase.

Albeit our proposed scheme can deal with any time correlated signals, and for analysis and design purposes, let $x_s(n)$

be a real and time-discrete *auto-regressive* model of order 1 ($AR-1$), which is commonly assumed in the signal processing literature in order to model real sources. It is defined as

$$x_s(n) = \rho x_s(n-1) + z(n), \quad \text{for } n = 1, 2, \dots \quad (13)$$

with covariance matrix \mathbf{R} , and $[\mathbf{R}]_{n,n-i} = \rho^i$, where the *auto-regression coefficient* is denoted by $\rho \in [0, 1]$ and assumed to be constant during the transmission. The random process $z(n)$ is a sequence of Gaussian distributed and independent random variables with zero mean and variance σ_z^2 with pdf denoted as $f(x)$. Also, and without loss of generality, let us assume unit variance of $x_s(n)$, i.e. $\sigma_x^2 = 1$; therefore, the variance of the noise is $\sigma_z^2 = 1 - \rho^2$.

This problem was studied in [10] together with other encoding-decoding schemes in the context of generic WSNs and assuming the ideal noiseless case, that is $\mathbb{E}[\|e_{CS}(n)\|^2] = \sigma_{CS}^2 = 0$. In this paper the analysis of the perturbed scenario is mandatory and [10] is a necessary tool to develop it. In the following, let us summarize first the necessary contribution from [10] and deal with the perturbed case afterwards.

In the unperturbed case, the transmission scheme is modeled according to a Markov Chain where each state t corresponds to the elapsed time since the last transmission. In other words, $t = 0$ means that the sensor s transmits in the current time slot, $t = 1$ means that the sensor s does not transmit but it transmitted in the previous time slot, and so on. An accurate approximation of the MSE related to each state can be computed as [10, Lemma 2]:

$$\underline{\text{MSE}}_t^{\text{CDE-PD}} \simeq h\left(1 - \rho^2 + \rho^2 \underline{\text{MSE}}_{t-1}^{\text{CDE-PD}} | \Delta_t\right) \quad (14)$$

where $i)$ Δ_t is the threshold at the CDE related to the state t and computed as

$$\Delta_t = \sqrt{2\left(1 - \rho^2 + \underline{\text{MSE}}_{t-1}^{\text{CDE-PD}}(\Delta_{t-1})\right)} \text{erf}^{-1}\left(1 - \frac{K}{S}\right), \quad \text{for } t = 1, 2, \dots \quad (15)$$

and $ii)$ the function $h(\sigma^2 | \Delta_t) : \mathbb{R} \rightarrow \mathbb{R}$ is defined as

$$h(\sigma^2 | \Delta_t) = \frac{2}{\eta(\Delta_t) \sqrt{2\pi\sigma^2}} \times \left(-\Delta_t \sigma^2 e^{-\frac{\Delta_t^2}{2\sigma^2}} + \frac{1}{2} \sqrt{2\pi} \sigma^6 \text{erf}\left(\frac{\Delta_t}{\sqrt{2\sigma^2}}\right)\right), \quad (16)$$

where $\eta(\Delta_t) = \int_{-\Delta_t}^{\Delta_t} f(x) dx$ and $f(x)$ denotes the pdf of $z(n)$.

The result in [10, Lemma 2] reveals the MSE for each state t . However, the mean distortion will be obtained by averaging over all the possible states as

$$\mathcal{D}_{\text{PD}} \simeq \sum_{t=0}^{\infty} P_t \underline{\text{MSE}}_t^{\text{CDE-PD}}, \quad (17)$$

where $P_t = K/S(1 - K/S)^t$ is the probability of sensor s being in state t .

For the perturbed case, let us consider the result in the following lemma.

Lemma 1: Let the $\underline{\text{MSE}}_t^{\text{CDE-PD}}$ define an approximation of the mean square error when the observation vector is $\tilde{\mathbf{x}}_t(n)$ and with the presence of a perturbation $e_{CD}(n)$ of mean 0

and power σ_{CS}^2 in the transmitted signal and $\underline{\text{MSE}}_t$ as defined in (14).

$$\text{MSE}_t^{\text{CDE-PD}} = \underline{\text{MSE}}_t^{\text{CDE-PD}} + \rho^{2t} \sigma_{CS}^2. \quad (18)$$

proof First, let $\tilde{\mathbf{x}}_t(n)$ define an observation vector composed by sensor readings and estimations as in (4), where the first reading is at position t . Also, $x'(n) = x(n) + e_{CS}(n)$.

For $t = 0$, the error is σ_{CS}^2 by definition since

$$\text{MSE}_0^{\text{CDE-PD}} = \mathbb{E}[\|x(n) - x'(n)\|^2] = \mathbb{E}[\|e_{CS}\|^2] = \sigma_{CS}^2. \quad (19)$$

For $t = 1$, the error $\text{MSE}_1^{\text{CDE-PD}}$ follows the conditional variance¹

$$\text{MSE}_1^{\text{CDE-PD}} = \mathbb{E}\left[\left(x(n) - \rho x'(n-1)\right)^2 \mid |x(n) - \rho x(n-1)| < \Delta_1\right]. \quad (20)$$

After some algebraic manipulations and taking into account that e_{CS} and $x(n)$ are statistically independent, the MSE is

$$\text{MSE}_1^{\text{CDE-PD}} = \mathbb{E}\left[\left(x(n) - \rho x(n-1)\right)^2 \mid |x(n) - \rho x(n-1)| < \Delta_1\right] + \rho^2 \sigma_{CS}^2. \quad (21)$$

According to the results in [10], the term $\mathbb{E}\left[\left(x(n) - \rho x(n-1)\right)^2 \mid |x(n) - \rho x(n-1)| < \Delta_1\right]$ can be expressed as $h(1 - \rho^2 | \Delta_1)$. Hence the previous MSE can be written as

$$\text{MSE}_1^{\text{CDE-PD}} = h(1 - \rho^2 | \Delta_1) + \rho^2 \sigma_{CS}^2. \quad (22)$$

For $t = 2$ the derivation is similar.

$$\begin{aligned} \text{MSE}_2^{\text{CDE-PD}} &= \mathbb{E}\left[\left(x(n) - \rho \hat{x}(n-1)\right)^2 \mid |x(n) - \rho \hat{x}(n-1)| < \Delta_2, |z(n)| < \Delta_1\right], \\ &= \mathbb{E}\left[\left(x(n) - \rho^2 x'(n-2)\right)^2 \mid |x(n) - \rho^2 x(n-2)| < \Delta_2, |z(n)| < \Delta_1\right], \\ &= \underline{\text{MSE}}_2^{\text{CDE-PD}} + \rho^4 \sigma_{CS}^2. \end{aligned} \quad (23)$$

Again, according to [10], the term $\mathbb{E}[\left(x(n) - \rho^2 x(n-2)\right)^2 \mid |x(n) - \rho^2 x(n-2)| < \Delta_2, |z(n)| < \Delta_1]$ can be accurately approximated by $h(1 - \rho^2 + \rho^2 \underline{\text{MSE}}_1^{\text{CDE-PD}} | \Delta_2)$, hence the MSE is

$$\text{MSE}_2^{\text{CDE-PD}} = \underline{\text{MSE}}_2^{\text{CDE-PD}} + \rho^4 \sigma_{CS}^2. \quad (24)$$

It is easy to conclude for $t > 0$ that

$$\text{MSE}_t^{\text{CDE-PD}} = \underline{\text{MSE}}_t^{\text{CDE-PD}} + \rho^{2t} \sigma_{CS}^2. \quad (25)$$

¹The conditional variance of a continuous random variable X given the condition $Y = y$ is defined as $\text{var}(X|Y = y) = \mathbb{E}[X^2|Y = y] - \left(\mathbb{E}[X|Y = y]\right)^2$, where $f(X|Y = y)$ is the conditional pdf of X given $Y = y$.

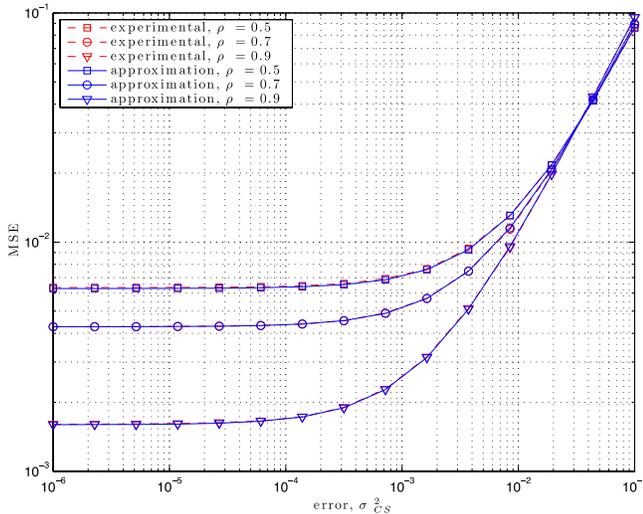


Fig. 5. Experimental and theoretical approximation of the distortion of the pair CDE-PD as a function of the perturbation power σ_{CS}^2 . The correlation coefficients are $\rho = \{0.5, 0.7, 0.9\}$ and the coding rate is $\gamma = 0.25$. Each point of the experimental curves corresponds to a transmission of 10^4 samples.

Again, the mean distortion will be obtained by averaging over all the possible states as

$$\begin{aligned} \mathcal{D}_{PD} &\simeq \sum_{t=0}^{\infty} P_t \text{MSE}_t^{\text{CDE-PD}} \\ &= \sum_{t=0}^{\infty} P_t \left(\text{MSE}_t^{\text{CDE-PD}} + \rho^{2t} \sigma_{CS}^2 \right) \\ &= \sum_{t=0}^{\infty} P_t \text{MSE}_t^{\text{CDE-PD}} + \frac{\gamma \sigma_{CS}^2 \rho^2}{1 - (1 - \gamma)\rho} \end{aligned} \quad (26)$$

where $\gamma = K/R$. Therefore, the second term in (26) is the impact of the noise in the communication link.

The distortion of the CDE-PD pair is evaluated in the following experiment.

Experiment 2: Let us simulate the CDE-PD scheme in Fig. 4 for the values of $\rho = \{0.5, 0.7, 0.9\}$ and $\gamma = K/S = 0.25$ as a function of the perturbation power σ_{CS}^2 during a transmission of 10^4 samples. The threshold values Δ_t have been computed following (15). The results have been compared with the theoretical approximation in (26).

The results of Experiment 2 have been plotted in Fig. 5. The reader can appreciate that the analytical approximations accurately describe the simulated results.

V. NETWORK DESIGN

The parameters K and R have a direct influence on the distortion level at the AF-CS decoding stage. The higher the value of K , the lower the distortion at the PD. At the same time, a sufficiently large value of R ensures zero distortion at the CSD. On the other hand, K and R also influence the number of transmissions as we have discussed in Section III. In order to adjust both K and R , we define the cost function $\nu(\beta)$ as

$$\nu(\beta) = \beta \mathcal{D}_{\text{AF-CS}} + (1 - \beta)\mathcal{E}, \quad (27)$$

where \mathcal{E} is the energy consumption modeled as the number of transmissions normalized by the total number of sensors,

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Number of fusion nodes:	$F = 1$
Number of sensing nodes:	$S = 200$
Number of active sensors:	$K = [0, 200]$
Number of relay nodes:	$R = [0, 200]$
Cost factor:	$\beta = 0.5$
Correlation model, $[\mathbf{R}]_{n,n+k} = \rho^{ k }$:	$\rho = 0.95$

i.e., $\mathcal{E} = \frac{K+R}{S}$. Therefore, the cost factor β controls the trade-off between energy consumption and error. Hence, for a given β , the design problem corresponds to find the values of K^* and R^* that minimize $\mathcal{P}2$, as

$$\begin{aligned} \mathcal{P}2 : \quad &\underset{K,R}{\text{minimize}} \quad \nu(\beta), \\ &\text{subject to} \quad \mathcal{C}_{CS}. \end{aligned} \quad (28)$$

Recall that the distortion of the AF-CS decoder can be modeled as $\mathcal{D}_{\text{AF-CS}} = \mathcal{D}_{\text{PD}}$ if \mathcal{C}_{CS} holds. Thus, this problem can be solved very fast because \mathcal{D}_{PD} depends only on K . Once K^* is obtained, the lowest value of R that satisfies \mathcal{C}_{CS} is computed.

VI. NUMERICAL RESULTS

In this section, the simulation results test the performance of our proposed energy-efficient scheme. Table I summarizes the parameters in our set-up.

A. Results About the Proposed CS Design Rules.

In order to study the accuracy of our analytical design strategy, we compare the theoretical to the real (simulated) results for the noiseless case. Fig. 6 shows the contour lines of the empirical cost function $\tilde{\nu}(\beta)$, defined as

$$\tilde{\nu}(\beta) = \beta \mathcal{D}_{\text{AF-CS}}^{\text{sim}} + (1 - \beta)\mathcal{E}. \quad (29)$$

where $\mathcal{D}_{\text{AF-CS}}^{\text{sim}} = \mathbb{E}[|\mathbf{x}(n) - \hat{\mathbf{x}}(n)|^2]$ is calculated empirically. In this simulation, we set $\beta = 0.5$ in order to equally prioritize the energy consumption and the reconstruction error. Clearly, we can observe a low-valued region in $\tilde{\nu}(0.5)$ for low values of K and R (i.e., the contour lines around the marker \times). This is a direct consequence of the CS principia, that is, the signal can be accurately recovered from a small amount of the total data. On the contrary, for high values of K and R , the proposed CS scheme performs inefficiently due to either a wrong signal recovery or a higher energy consumption or both.

We have also plotted in Fig. 6 the feasible regions and the optimal results for the minimization problem $\mathcal{P}2$ for two different constraints on the number of measurements R , i.e.: *i*) the proposed $\mathcal{C}_{CS} : R > K + K \log(\frac{S}{K})$ and *ii*) the conventional $\mathcal{C}'_{CS} : R > K \log S$.

In Fig. 7 we evaluate the existing gap between the design parameters obtained in (28) using both \mathcal{C}_{CS} and \mathcal{C}'_{CS} , i.e., K^* and R^* , and K'^* and R'^* , respectively, and the ones obtained *a posteriori* minimizing $\tilde{\nu}(\beta)$ by simulation, i.e., K_{sim} and R_{sim} . In Table II these values are given for $\beta = 0.5$.

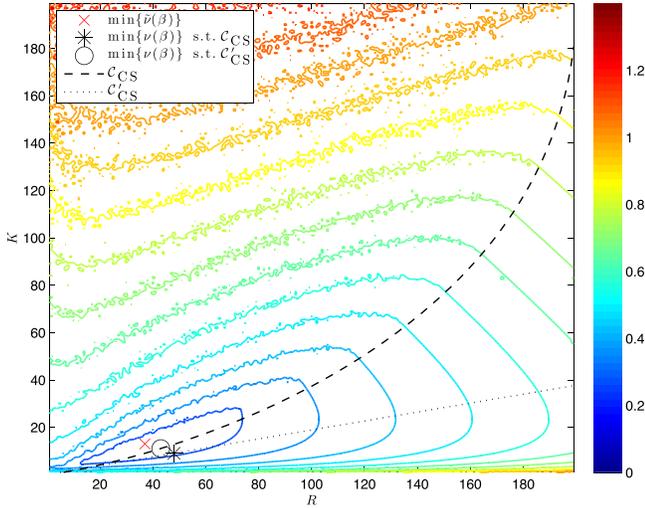


Fig. 6. Contour map of $\tilde{\nu}(\beta)$ for $\beta = 0.5$. Cold colors represent low values and high values are hot-colored (in color printed version). For black and white version, contours near the cross marker are showing the minimum values. This figure has been averaged over 100 realizations. The areas under the curves C_{CS} and C'_{CS} correspond to the feasible set of $\mathcal{P}2$ subject to C_{CS} and C'_{CS} respectively.

Although C'_{CS} guarantees an exact reconstruction of $\mathbf{x}_K(n)$, the results in Fig. 7 show that this constraint is too conservative. On the other hand, we can observe that our proposed model using C_{CS} is much closer to the real performance of the system. Indeed, it slightly upper-bounds the actual results, giving a slightly conservative solution for the network design parameters. However, the gap between R^* and R_{sim} is 9.8 sensors on average (it means only a relative error of 4.9% over the total number of sensors). The gap between K^* and K_{sim} is even smaller, 1.4 on average (0.7%). On the other hand, using C'_{CS} , the error increases up to 17.5 in mean (8.75%) for R'^* and 12.35 (6%) for K'^* .

For low values of β , our approach prioritizes the energy savings by decreasing the number of active sensors. Although it provides less accuracy, this situation may be interesting in some WSN monitoring applications. For example in temperature or humidity monitoring there may be long periods of time where changes are not expected, or even periods of less interest, e.g., during the nights. On the contrary, other situations would require higher accuracy in the measurements at the expense of more energy. Using our approach, the parameter β can be adjusted in order to accommodate all these situations.

B. Comparison to Standard CS Schemes Available in the WSN Literature.

We compare the performance of the AF-CS scheme to two reference systems in the WSN literature: *i*) the Classical Approach (CA) where the group of S sensing nodes transmit their measurements directly to the fusion node each time slot, and *ii*) the Compressive Wireless Sensing (CWS) [6].

For a fair comparison, we add a *space correlation* (i.e., a correlation among the measures of the sensors) equal to the applied *time correlation* in Table I, because algorithms of the type of the CWS scheme compress their readings according to some prior knowledge of the spatial structure of the signal

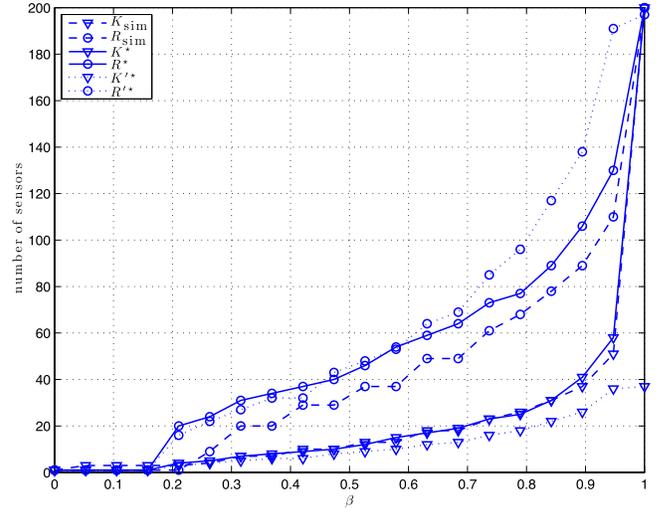


Fig. 7. Comparison between the optimal K and R parameters obtained analytically (solid lines) with the optimal simulated ones using C_{CS} (dotted lines) and C'_{CS} (dashed lines).

TABLE II
OPTIMAL VALUES FOR K AND R WITH $\beta = 0.5$

	K	R
$\min\{\tilde{\nu}(\beta)\}$	$K_{sim} = 13$	$R_{sim} = 38$
$\min\{\nu(\beta)\}$ s.t. C_{CS}	$K^* = 11$	$R^* = 43$
$\min\{\nu(\beta)\}$ s.t. C'_{CS}	$K'^* = 9$	$R'^* = 48$

(as e.g., spatial correlation). In addition, the matrix Ψ is constructed following a DCT basis and Φ is a gaussian random matrix (with i.i.d. entries). In principle, this type of transform is suitable for temperature-like readings, as it is showed in [8].

We compare the following figures of merit:

- 1) *Empirical distortion*, $\mathcal{D}_{AF-CS}^{sim}$, averaged over 100 realizations and for different instances of Φ .
- 2) The *relative energy consumption* \mathcal{E} measures the energy consumption in comparison to a standard star-topology WSN (in terms of the number of transmissions). For a fair comparison among all considered schemes, we assume that they all have the same duration and transmission power in mean. In the evaluated strategies:

- $\mathcal{E} = 1$ for the case of CA. All S sensors transmit their readings to the *fusion center*.
- $\mathcal{E} = R$ for CWS. All S sensors transmit using the same resource to the *fusion center* to perform one projection. This process is repeated R times.
- $\mathcal{E} = \frac{(R+K)}{5}$ for AF-CS. The nodes in the subset $\mathcal{K}(n)$ broadcast their messages towards the R relays, involving K transmissions. In addition, the *relay nodes* retransmit the computed R projections to the *fusion center*.

- 3) The number of *channel uses*:

- CA performs S channel uses assuming a given orthogonal medium access control policy.
- CWS requires R channel uses.
- AF-CS requires one channel use in the *sensing phase* and R channel uses in the *projection phase*, hence a total of $R + 1$ channel uses.

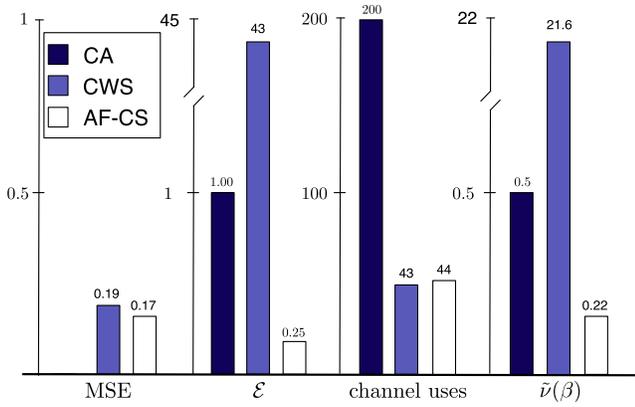


Fig. 8. Graphic bar comparison of the different figures of merit among the studied CS schemes. Lower values are better.

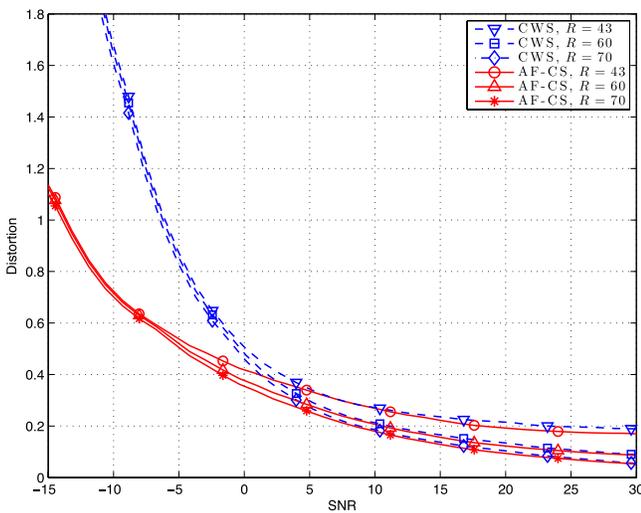


Fig. 9. Comparison of $\mathcal{D}_{\text{AF-CS}}^{\text{sim}}$ as a function of the SNR between CWS and AF-CS. The simulation setup is for $K = 11$ and $R = 43, 60, 70$. The figure has been averaged over 100 realizations.

- 4) The empirical *cost function* $\bar{v}(\beta)$ as an indicator of the existing trade-off between energy savings and accuracy. We have set $\beta = 0.5$.

Fig. 8 summarizes the results for the configuration $K^* = 11$, $R^* = 43$ obtained in Fig. 6 and Table II for the evaluated topologies.

In terms of MSE, it is obvious that the best performance is obtained with the CA, since there is no reconstruction error. However, our proposed method performs slightly better than CWS (0.17 against 0.19).

The major difference among the schemes can be observed in terms of the energy consumption. Our method performs the best compared to the other evaluated techniques. Note also that the CWS performs even worse than the CA because it is extremely inefficient in terms of number of transmissions. The same conclusions can be extracted from $\bar{v}(\beta)$.

Regarding the number of *channel uses*, CWS performs the best with 43 channel uses. However, our proposed method requires just an additional use. Besides, a big improvement in comparison to CA can be appreciated with 200 channel uses.

C. Robustness Against Additive White Gaussian Noise

In this section, we evaluate by simulation the performance reduction due to the AWGN comparing our method to the CWS approach. Simulation results in Fig. 9 show the performance of the methods analyzed as a function of the SNR. The SNR is defined as the ratio between the power of the useful signal $\Phi \mathbf{x}_K(n)$ and the power of the noise $\mathbf{w}(n)$:

$$\text{SNR} = 10 \log \left(\frac{\|\Phi \mathbf{x}_K(n)\|^2}{\|\mathbf{w}(n)\|^2} \right). \quad (30)$$

In Fig. 9, we evaluate the robustness of the techniques in terms of $\mathcal{D}_{\text{AF-CS}}^{\text{sim}}$ against thermal noise and for different values of R . We observe that AF-CS outperforms clearly CWS for low values of SNR (below 5dB). However, both techniques perform similarly for high values of SNR.

VII. CONCLUSION

This paper has introduced a *decentralized* solution for a *compressed sensing* implementation in a two-hop Wireless Sensor Network scenario, which is referred to as *Amplify-and-Forward Compressed Sensing* (AF-CS). First, the sensing nodes exploit inner time correlation in order to reduce the number of transmissions, only keeping as active nodes the K sensors with the most unpredictable readings. In such a way, we distributedly transform the vector of interest $\mathbf{x}(n)$ in a K -sparse approximation $\mathbf{x}_K(n)$. Second, the relay nodes receive *random projections* formed from the linear combination of the readings of the K active sensors, that transmit time-synchronized, each one multiplied by its channel gain. Thus, the number of channel uses is significantly reduced with respect to a classical approach. The relays retransmit the random projections to the fusion center, where finally, it reconstructs the original vector using an l^1 -norm minimization.

Next, we have described an error model that allows us to configure the number of active sensors K and the minimum number of relay nodes R that are required in order to guarantee the desired performance by means of an specifically designed optimization problem. It takes into account the existing trade-off between energy consumption and signal distortion.

Finally, the simulation results indicate that our *proposed* scheme drastically reduces the number of transmissions (by a factor of 4.54) and the number of *channel uses* (by a factor of 4) when compared to a *classical transmission scheme*. The AF-CS scheme also outperforms other distributed compressed-sensing-based techniques for wireless sensor networks not only in terms of energy-efficiency (reducing the number of transmissions by a factor of 172) but also in terms of robustness against noise.

REFERENCES

- [1] D. L. Donoho, "Compressed Sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [3] S. Nikitaki and P. Tsakalides, "Localization in wireless networks based on jointly compressed sensing," in *Proc. 19th Eur. Signal Process. Conf.*, Barcelona, Spain, Aug. 2011, pp. 1–5.

- [4] J. Meng, H. Li, and Z. Han, "Sparse event detection in wireless sensor networks using compressive sensing," in *Proc. 43rd Annu. Conf. Inf. Sci. Syst.*, Mar. 2009, pp. 181–185.
- [5] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Compressive data gathering for large-scale wireless sensor networks," in *Proc. ACM Mobicom*, Beijing, China, Sep. 2009, pp. 145–156.
- [6] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressive wireless sensing," in *Proc. 5th Int. Conf. Inf. Process. Sensor Netw.*, New York, NY, USA, Apr. 2006, pp. 134–142.
- [7] J. Haupt, W. U. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 92–101, Mar. 2008.
- [8] C. Chou, R. Rana, and W. Hu, "Energy efficient information collection in wireless sensor networks using adaptive compressive sensing," in *Proc. IEEE 34th Conf. Local Comput. Netw.*, Zurich, Switzerland, Oct. 2009, pp. 443–450.
- [9] R. Rugin, A. Conti, and G. Mazzini, "Experimental investigation of the energy consumption for wireless sensor network with centralized data collection scheme," in *Proc. 15th Int. Conf. Softw., Telecommun. Comput. Netw.*, Sep. 2007, pp. 1–5.
- [10] J. E. Barceló-Llado, A. Morell, and G. Seco-Granados, "Conditional downsampling for energy-efficient communications in wireless sensor networks," *J. Adv. Signal Process.*, vol. 101, pp. 1–16, May 2013.
- [11] P. Viswanath David Tse, *Fundamentals of Wireless Communication*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
- [12] P. Lioliou, M. Viberg, and M. Coldrey, "Efficient channel estimation techniques for amplify and forward relaying systems," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3150–3155, Nov. 2012.
- [13] M. A. Herman and T. Strohmer, "General deviants: An analysis of perturbations in compressed sensing," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 342–349, Apr. 2010.
- [14] E. J. Candès and Y. Plan, "A probabilistic and RIPless theory of compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 11, pp. 7235–7254, Nov. 2011.
- [15] J. L. Paredes and G. R. Arce, "Compressive sensing signal reconstruction by weighted median regression estimates," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2585–2601, Jun. 2011.
- [16] S. Ji, Y. Xue, and L. Carin, "Bayesian compressive sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, Jun. 2008.
- [17] E. J. Candès and T. Tao, "Decoding by linear programming," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, Dec. 2005.
- [18] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.
- [19] E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies," *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [20] E. Candès, "The restricted isometry property and its implications for compressed sensing," *Comput. Rendus Math.*, vol. 346, nos. 9–10, pp. 589–592, 2008.
- [21] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," Apr. 2011.
- [22] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Advances in Learning and Control* (Lecture Notes in Control and Information Sciences), V. Blondel, S. Boyd, and H. Kimura, Eds. New York, NY, USA: Springer-Verlag, 2008, pp. 95–110, 2008.
- [23] X. Feng and Z. Zhang, "The rank of a random matrix," *Appl. Math. Comput.*, vol. 185, pp. 689–694, Jul. 2007.

Joaquín Enric Barceló-Lladó received the M.Sc. degree in electrical engineering and the Ph.D. degree from Universitat Autònoma de Barcelona (UAB), Spain, in 2007 and 2012, respectively.

In 2007, he was a Research Assistant with Università degli Studi di Siena, Italy, involved in cross-layer techniques for satellite networks. From 2008 to 2009, he held a visiting position at the UNIK-University Graduate Center, Oslo, Norway, where he was involved in resource allocation for satellite communications based on game theory. He joined the Signal Processing for Communications and Navigation Group, Department of Telecommunication and Systems Engineering, UAB, in 2009. He has been involved in several national and international projects with public and private fundings. His current areas of interest are related to algorithms for energy-efficient communications with correlated sources in wireless sensor networks.

Antoni Morell received the M.Sc. degree in electrical engineering and the Ph.D. degree from Universitat Politècnica de Catalunya (UPC), Barcelona, in 2002 and 2008, respectively.

He was with the Signal Theory and Communications Department, UPC, from 2002 to 2005. He joined Universitat Autònoma de Barcelona, as a Research and Teaching Assistant in 2005 and an Assistant Professor in 2008, teaching courses on communications, signals, and systems. He has been involved over ten research and development projects, national and international, being the Principal Investigator in one of them. He has expertise in optimization techniques applied to communications, also in the field of wireless sensor networks. He has published over 30 papers in recognized international journals and conference proceedings and his current research interests include distributed multiple access protocols, distributed source coding, compressed sensing, routing strategies in WSNs, and inertial-aided indoor positioning.

Gonzalo Seco-Granados (S'97–M'02–SM'08) received the Ph.D. degree in telecommunication engineering from Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in 2000, and the M.B.A. degree from the IESE-University of Navarra, Barcelona, in 2002. From 2002 to 2005, he was a member of the technical staff within the RF Payload Division, European Space Research and Technology Center, European Space Agency, Noordwijk, The Netherlands, where he was involved in the Galileo project. He led the activities concerning navigation receivers and indoor positioning for GPS and Galileo. Since 2006, he has been an Associate Professor with the Department of Telecommunications and Systems Engineering, Universitat Autònoma de Barcelona, Barcelona. From 2007 to 2011, he was a coordinator of the Telecommunications Engineering degree; and since 2011, he has been the Vice-Director of the UAB Engineering School. In 2008, he was the Director of one of the six Chairs of Technology and Knowledge Transfer UAB Research Park Santander.

Dr. Seco-Granados has been principal investigator of over 20 national and international research projects, and acts often as an advisor of the European Commission in topics related to communications and navigation. He has had several visiting appointments at Brigham Young University and the University of California at Irvine. He has published over 30 journal papers and 110 conference contributions, and holds two patents under exploitation. His research interests include signal processing for wireless communications and navigation, estimation theory, synchronization, location-based communications, and optimization.