Administering quality-energy trade-off in IoT sensing applications by means of adapted compressed sensing

Administering Quality-Energy Trade-Off in IoT Sensing Applications by Means of Adapted Compressed Sensing

Mauro Mangia, Alex Marchioni, Fabio Pareschi, Riccardo Rovatti, Gianluca Setti

***Abstract*—A common scheme to let a very large number of low-resources sensing units communicate their readings to a remote concentrator is to deploy intermediate hubs that collect subsets of readings by means of local communication and perform the needed long-range transmission of a compressed version of the data. We here propose to exploit Compressed Sensing as an extremely lightweight lossy compression stage for which it is easy to address the trade-off between the quality of the reconstructed signal and the energy needed to complete acquisition. Over the huge set of parameters characterizing the design space (such as the number of intermediate hubs, the sensors transmission range, etc.), we analyze such a trade-off when the placements of the hubs is not completely random but aims at promoting diversity between the subsets of readings considered by each hub. With respect to the case of no intermediate data aggregation, numerical evidence suggests that, when an appropriate design strategy for the compressed sensing stage is adopted and diversity is promoted, an energy savings higher than 60% with high quality signal reconstruction can be obtained. This operative point corresponds to 20 intermediate hubs deployed to collect reading from 128 sensors.**

***Index Terms*—Internet of Things, Wireless Sensor Networks, Compressed Sensing, Signals on graphs, Smart Dust**

I. Introduction

NEXT steps in the development of the Information and Communication Technology, being known as Internet of Things [1], [2], Industry 4.0 [3], or Big Data Analytics [4], are all based on an increasing interaction between information processing and the physical world. In this sense, the very same concept of sensing is rapidly changing [5], [6], and the word *sensor* is no longer regarded as a simple device converting physical quantities into electrical signals or digital words, but a complex and smart system.

M. Mangia is with the Advanced Research Center on Electronic Systems (ARCES), University of Bologna, 40125 Bologna, Italy (e-mail: mauro.mangia2@unibo.it)

A. Marchioni is with the Department of Electrical, Electronic, and Information Engineering, University of Bologna, 40136 Bologna, Italy (e-mail: alex.marchioni@unibo.it).

F. Pareschi is with the Department of Engineering, University of Ferrara, 44122 Ferrara, Italy, and also with the Advanced Research Center on Electronic Systems (ARCES), University of Bologna, 40125 Bologna, Italy (e-mail: fabio.pareschi@unife.it).

R. Rovatti is with the Department of Electrical, Electronic, and Information Engineering, University of Bologna, 40136 Bologna, Italy, and also with the Advanced Research Center on Electronic Systems (ARCES), University of Bologna, 40125 Bologna, Italy (e-mail: riccardo.rovatti@unibo.it).

G. Setti is with the Department of Electronics and Telecommunications, Politecnico di Torino, 10129 Torino, Italy, and also with the Advanced Research Center on Electronic Systems (ARCES), University of Bologna, 40125 Bologna, Italy (e-mail: gianluca.setti@polito.it).

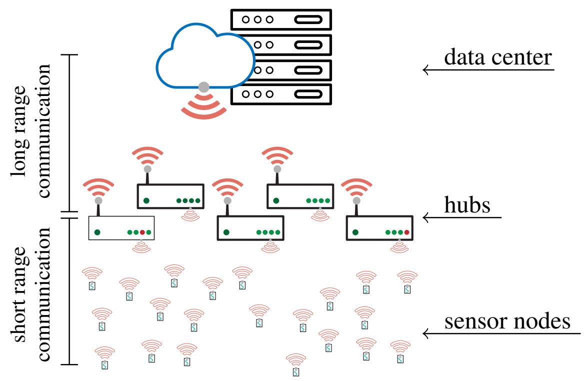


Fig. 1. Scenario considered in this paper, where many sensor nodes establish a short range communication (WSN) with a local hub, capable of long range communication (WAN) with a data collector.

The scenario considered in this paper belongs to this general framework. In particular, we focus on a system where a large amount of data is generated by ultra low-power, miniaturized autonomous sensor nodes, dispatching their readings to some remote data collector for processing purposes. Such a system is often referred to with the terms "Smart Dust" [7], [8].

Such a paradigm finds application in a wide range of new applications ranging from security/safety surveillance to structural health monitoring [9]–[11], smart building [12], and even to miniaturized biomedical implants [13]–[15].

The intuitive representation of the considered scenario is depicted in Figure 1. A number of sensor nodes is deployed according to the smart dust paradigm, i.e., with non-controllable geographical locations. They feature local communication capability and are joined into a Wireless Sensor Network (WSN). Sensor readings are eventually collected by a central data collector, assumed to be far away from nodes. To this aim, local hubs are introduced in the WNS collecting readings from the sensor nodes, pre-processing them, and delivering intermediate results to the data collector by means of long range transmissions in a Wide Area Network (WAN). Sensor nodes and local hubs are considered different devices, with a different hardware architecture and complexity. Nevertheless, both classes of devices are battery powered so that a quality-energy trade-off must consider the entire set of local/wide communications.

In order to introduce data compression so to reduce the overall energetic costs, we consider the Compressed Sensing (CS) paradigm [16], [17], as in other approaches [18]–[23].

The main novelties in this paper are: i- local hubs must be deployed according to certain geographical constraints, and their positions cannot be set a-priori (i.e., common coverage techniques cannot be applied); ii- a new figure of merit (i.e., *hub diversity*) is adopted to account for the quantity of additional signal information introduced by each hub; iii- the *rakeness*-based CS paradigm is applied to this framework to further reduce communication costs [24], [25] with respect to standard CS approach.

The investigated novelties, along with many framework features (e.g., number of local hubs, sensor nodes transmission range, communication failure probability, sensors node locations) are such that some quality-energy trade-off can be addressed.

The paper is organized as follows. In Section II we introduce the motivation of the work. We also present a survey on related CS-based works. Section III introduces the details of the considered system and recall some prior results used in this paper. Section IV is used to highlight what are the degrees of freedom and how we can tune them in order to optimize the system. In Section V we describe the simulation setting and the proposed energy-quality trade-off as well as the obtained results. Finally, we draw the conclusion.

II. Motivation and Related Works

The introduction of the two-layer structure of Figure 1 is justified by two considerations.

First, the energy required by transmission increases more than linearly with the distance. In other words, there is a very large difference between the energy involved in the local communication process (WSN) and in the wide-area (WAN) communication process. Indicating with α the ratio between the local and the wide-area transmission ranges, the ratio between the energies involved in the two communication processes, assuming that both have antennas with no particular directivity, is α^2. Assuming α at least in the order of 10^2, we expect a ratio between entailed communication powers in the order of 10^4.

Then, it is worth noticing that the number of sensor nodes may be relevant. Due to this, it is fundamental to administer data processing and communication at different levels, with the aim of reducing the use of system resources. In such a situation, introducing in the WSN either a certain amount of processing or a certain level of data aggregation, or both, may improve the energy-quality trade-off, with beneficial effects in terms of reduction of the energy to complete acquisition.

This is more than enough to introduce local readings collection and pre-processing before a long-range transmission is attempted. Yet, the problem of data aggregation in WSN has been widely investigated in the literature [26]. An important result is given in [27], where the authors consider a sensor network where the data collector is located far from the sensor nodes and propose a dynamic and adaptive low-energy clustering approach known as LEACH. However, in [27], any sensor is capable of promoting itself to the role of local hub, and this is not allowed according to the smart dust paradigm where hubs and sensors are different devices.

In order to improve performance in terms of energy required for transferring data to the collector, many works introduce CS in this framework. Since hubs are battery power devices, energy saving is fundamental. This is why data aggregation based on CS has been proposed: its computational cost is smaller with respect to standard algorithms and the achieved compression yields a comparable number of long range transmissions. Hence, hubs collect readings from sensor nodes, and provide to the collector only a small number of linear combinations of them. According to existing literature [25], [28]–[31], the additional energy-cost due to this kind of processing is in general negligible, and dominated by communication cost.

Among the works introducing CS in this scenario, [20] proposes CS-based compression and considers the number of clusters to reduce the total amount of local communications needed to first collect readings and then transmit intermediate results to a main collector. However, the main collector is assumed to be at the center or immediately outside the sensing area, but not far from it thus avoiding any trade-off between local and long-range communications. The authors of [21] propose, on the basis of the LEACH scheme, an adaptive and energy-balanced data gathering and aggregating approach. Similarly to previously considered cases, the collector is assumed to be inside the sensing area. The authors of [19] propose an aggregation technique, and investigate how the number of clusters is subjected to the overall power consumption. Yet, similarly to previous works, the collector is located at the edge of the sensing area. In [23] data aggregation before attempting long-range communication is investigated; however, hubs are selected among the sensor nodes in the WSN as in [27], while in the scenario considered here sensor nodes and hubs are different devices.

With respect to the aforementioned literature, the main innovative aspects of this paper can be summarized as follows.

- We assume that the deployment of the hubs follows rules that are similar to that used in the deployment of sensors, i.e. hubs cannot be placed at will, since their positions cannot be set. To cope with this, we model the hubs positions as random variables. Note that the immediate consequence is that common coverage techniques cannot be applied [32], [33].
- The *hub diversity* is introduced as a figure of merit to indicate how different is the coverage of two randomly deployed hubs. A low diversity value indicates a situation where two hubs communicates with almost the same set of sensor nodes, and so they provide a very similar information to the data collector. Based on this, it is possible to discard hub networks that do not ensure a minimum hub diversity.
- Each local hub compresses the sensors readings by CS paradigm. Compression is increased by adapting the rakeness-based CS to this framework [24], [25].

III. Scenario description

A. Input signal and node network configuration

According to what is commonly assumed by the smart dust paradigm, the input signal is a physical quantity that is

Fig. 2. Example of sensor node network with $n = 50$ and $d_{th} = 0.2$. A connection between nodes indicates that they are related to the each other, i.e., their distance is smaller than d_{th}.

TABLE I
FEATURES OF THE CONSIDERED SCENARIO.

Input signal and node network configuration

	description	**values**
N	dimensionality of the signal	128
κ	number of non-zero components of x along D	$\{6, 12\}$
D	sparsity basis of x, Fourier of the connectivity graph with $d_{th} = 0.15$	$\mathcal{F}[C]$
$\mathcal{X}$	correlation matrix of x	$\mathbf{E}[xx^\top]$
ISNR	intrinsic SNR of x	60 dB

Hubs network configuration

	description	**values**
M	number of hubs	$\{14, 16, 18, 20, 22, 24\}$
r_{TX}	maximum connection distance between a hub and nodes in its neighborhood	$\{0.2, 0.25, 0.3\}$
$\mathcal{N}_k$	set of sensors sending their reading to the k-th hub	

Reading processing strategy

	description	**values**
E_{TX}	energy required by a sensor for a reading transmission	
E_{RX}	energy required by a hub for a sensor reading reception	
E_{WAN}	energy required by a hub to transmit a digital word	
m	number of compressed value transmitted to the main data collector	$\{4, 5, \ldots, 128\}$
p_{f}	probability that a hub fails in receiving a single value from a sensor	$\{0, 0.05, \ldots, 0.8\}$

acquired through N sensor nodes that are distributed, without any direct control, in the area (or over/inside the structure) to be monitored. For the sake of simplicity, we propose the following simplified model.

Let us number all sensor nodes from 0 to $N-1$, and assume that we are monitoring the input signal inside the 2-D unit square. Indicating with $\nu_k \in [0,1]\times[0,1]$ the coordinates of the k-th sensor node, we model the ν_k as random variables with an uniform distribution inside the unit square. We define the input signal x as the column vector $x = (x_0, \ldots, x_{N-1})^\top$, where x_k is the reading from the k-th sensor, and where $\cdot^\top$ stands for vector transposition. We also consider a disturbance vector $\eta = (\eta_0, \ldots, \eta_{N-1})^\top$ modeled with a Gaussian distribution so that the Intrinsic Signal-to-Noise Ratio (ISNR) is $\|x\|_2/\|\eta\|_2$.

In order to simplify the mathematical notation, in the following we assume that the expected value of the input signal is $\mathbf{E}[x] = 0$. We also make the two additional and realistic assumptions that x is *sparse* and *localized*.

The first assumption is the formalization of the fact that x exhibits redundancy and thus is compressible. Given a proper orthonormal basis $D \in \mathbb{R}^{N\times N}$ such that the input signal is expressed as $x = D\xi$, x is sparse if the coefficient vector $\xi \in \mathbb{R}^N$ has only a few non-negligible components, that are indeed the only ones required to reconstruct x. We also say that, if one knows that not more than κ components of ξ are non-null, then x is a κ-sparse signal.

The intuition behind sparsity is that the number of degrees of freedom of x is smaller than N, i.e., readings are not independent of, but related to the each other. We consider that the readings from two nodes j and k are related if the distance $\|\nu_j - \nu_k\|$ is no larger than a threshold value d_{th}. This definition allows us to create an *undirected graph* associated to the network: each sensor node is a graph vertex, and an edge between vertices j and k exists if $\|\nu_j - \nu_k\| \leq d_{th}$. We also limit ourselves to consider connected graphs. An example of sensor node network for $N = 50$ along with its associated graph obtained with $d_{th} = 0.2$ is depicted in Figure 2.

The advantage of creating a graph representation of the sensor network is that some authors [18], [34]–[37] have recently suggested a relation between sparsity and connected graph, in particular with the adjacency matrix C associated to the graph and defined as $C_{j,k} = 1$ if an edge between vertices j and k exists, and 0 otherwise. Assuming that C can be diagonalized as $C = D\Lambda D^{-1}$, with D the non-singular eigenvectors matrix and Λ the diagonal matrix containing the eigenvalues, [34] and [35] observe that D is the generalization of the Fourier basis for discrete time periodic signals. As it happens for time domain signals, graph-supported signals are often sparse in their Fourier representation. In the following, we will refer to this by saying that $D = \mathcal{F}[C]$.

The other prior on x, i.e., localization, is the hypothesis that its energy is not uniformly distributed along its components [24]. This property is well described by using the second order statistics of x, i.e., by its $N \times N$ correlation matrix $\mathcal{X} = \mathbf{E}[xx^\top]$. Signals with independent components feature a trivial diagonal correlation matrix made of individual variances $\mathcal{X}_{k,k} = \mathbf{E}[x_k^2]$ and products of the means $\mathcal{X}_{j,k} = \mathbf{E}[x_j]\mathbf{E}[x_k] = 0$ for $j \neq k$.

It is worth noticing that real-world quantities usually feature some form of both redundancy and correlation between components. Even if the computation of both D and $\mathcal{X}$ may result non-trivial, these assumptions can be used as a prior to optimize the acquisition of x.

B. Hub network configuration

To allow long-range communication to the data collector, a number M of hubs are spread into the unit square, at coordinates θ_k, $k = 0, \ldots, M-1$. The k-th hub is able to communicate with its neighborhood $\mathcal{N}_k$, defined as the set of the node indexes j such that $\|\theta_k - \nu_j\| \leq r_{\mathrm{TX}}$, being r_{TX} the transmission range of sensor nodes. Each hub collects all readings from its neighborhood, pre-processes them, and sends intermediate results to the collector thanks to a long range communication capability.

In many situations it is useful to deploy hubs with a regular pattern [32], [33]. Instead, we here assume that the physical constraints imposing to sensor nodes to be deployed in non-controllable geographical positions regulate also the deployment of the hubs. In other words, and assuming to have no information on the underlying sensor network and on possible geographical constraints, we model also θ_k as random variables in the unit square. Note that, depending on the position of nodes and hubs, a node could be in one or more than one neighborhood, but it is also possible that a node is not included in any neighborhood, as more clearly detailed in Subsection IV-B.

This is actually not an issue. The sparsity property implies that input signal has redundancy, so that it may be correctly reconstructed even if some readings are not available. On the contrary, reducing the hub network coverage may represent a way to reduce system power consumption. As an example, [25] introduces *puncturing* as a technique based on the intentional skipping of some samples producing an effective reduction in the energy required to acquire a sparse signal.

C. Readings processing strategy

Each of the N sensor nodes broadcasts its readings, and all hubs in range (i.e., with distance smaller than r_{TX}) can read the transmitted value. We assume that all N transmissions are not superimposed in some domain and do not interfere with each other.

We indicate with E_{TX} the energy required by a sensor node to broadcast its reading. We also indicate with E_{RX} the energy required by a hub to receive a sensor reading and assume that hubs are smart enough to spend energy only on sensors that are within their neighborhood.

Furthermore, in order to take into account external interference, we model the communication process in a stochastic way: there is a non-null probability p_{f} that a hub fails in correctly receiving a measurement. Both E_{TX} and E_{RX} are spent by the node and by the hub, respectively, independently of the fact that the transmission is successful or not.

Each hub can send readings to a central collector with an energy E_{WAN} for each single piece of data, with $E_{\mathrm{WAN}} \gg E_{\mathrm{TX}}$. In order to reduce long range transmission costs, reading from the sensors are pre-processed, and only a limited number of linear combinations of them is sent.

In more detail, we apply Compressed Sensing (CS), a technique known to reduce the number of samples required to correctly reconstruct a sparse signal at a negligible cost in terms of energy requirements. Each hub computes and sends the same number m/M of linear combinations, so that the whole long range system rely on m transmissions. The amount of energy required for processing is considered negligible with respect to communication-related energy

If m is not an integer multiple of M, some hubs compute and send $\lfloor m/M \rfloor$ combinations, while some others $\lfloor m/M \rfloor + 1$, with $\lfloor \cdot \rfloor$ the largest integer less than or equal to its argument.

D. Compressed sensing and signal reconstruction

Compressed sensing is a technique [16], [17] leveraging the sparsity prior to reduce the amount of quantities required to reconstruct a signal with respect to a Nyquist rate sampling.

Given a signal $x \in \mathbb{R}^N$, instead of considering the N values x_k for $k = 0, \ldots, N-1$, the fundamental idea is to compute a certain number of their linear combinations $y_j = \sum_{k=0}^{N-1} A_{j,k} \left(x_k + \eta_k\right)$ for $j = 0, \ldots, m-1$ and with $m < N$, called *measurements*. By arranging measurement and the disturbance terms in the vectors y and η, respectively, and the linear combination coefficients in the matrix A, then CS is described by the relationship $y = A\left(x + \eta\right)$.

The pre-processing mechanism introduced in the hubs belongs to the CS framework, and we can base signal reconstruction on the many techniques proposed to get $\hat{x}$ as a correct estimation of the actual input signal x. The main issue in signal reconstruction is that, since $m < N$, $\hat{x}$ cannot be in principle computed starting from the knowledge of y only. Yet, a number of theoretical developments guarantees that, if x is sparse with respect to D, $\hat{x} = D\hat{\xi}$ can be obtained by the optimization problem [16]

$$\hat{\xi} = \arg\min_{\xi} \|\xi\|_1 \quad \text{s.t. } \|AD\xi - y\|_2 < \sigma_\eta \tag{1}$$

i.e., by looking at the sparsest $\hat{\xi}$ among all ξ for which $AD\xi \approx y$. In (1), $\|\cdot\|_1$ and $\|\cdot\|_2$ are the standard ℓ_1 and ℓ_2 norms[1], and σ_η bounds the effects of η. Such an approach is called basis pursuit with denoising (BPDN).

Most of the practical interest in CS comes from the fact that actual estimation algorithms largely outperform the theoretical bounds allowing an effective recovery of x from a small number of measurements, i.e., usually $m \ll N$ [38]. This makes CS a quite good and computationally light compression algorithm.

Furthermore, the conditions on A that allow correct reconstruction are simply achieved by using random matrices and, even if the formal results depend on specific matrix distributions [39], in practical cases a wide class of random matrices allows for effective signal recovery [40]–[43].

[1]It is a common practice to promote sparsity by means of the ℓ_1 norm instead of the computationally intractable count of non-zero components given by ℓ_0 pseudo-norm.

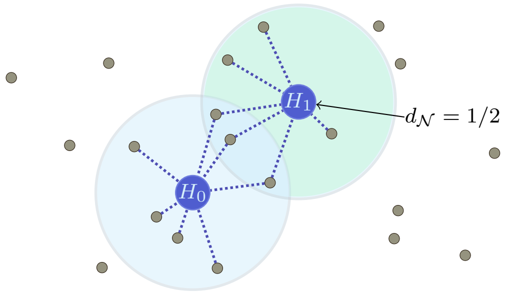


Fig. 3. Example of neighborhood diversity for a case with only two hubs: the neighborhood $\mathcal{N}_1$ of hub H_1, with respect to the neighborhood $\mathcal{N}_0$ of hub H_0 is such that $|\mathcal{N}_1 \setminus \mathcal{N}_0| \,/\, |\mathcal{N}_1| = 1/2$, so $d_{\mathcal{N}} = 1/2$.

Recently, some optimization techniques have been proposed capable of substantially improving CS performance with a proper adaptation of the sensing matrix A to some input signal features [24], [44], [45]. In the following, we will focus on the approach proposed in [24] that bases the design of A on the second-order statistics $\mathcal{X}$ of the input signal.

IV. System Design

The system described in previous section is defined by a large number of parameters. Some of them are constraints, while others are degrees of freedom that can be tuned to optimize some system performance. In this section we highlight which are the constraints and which are the degrees of freedom, and how the latter can be properly tuned.

A. System constraints

All input signal related parameters are to be considered constraints. The number of sensor nodes N, the sparsity property identified by the sparsity level κ and by the adjacency matrix C (or equivalently by the sparsity basis $D = \mathcal{F}[C]$), and the correlation matrix $\mathcal{X}$. Also the probability of a failure in a reading reception p_{f} is considered a constraint, since it is due to external interference.

B. Hub network design

The design of the hub network is a degree of freedom, and one can decide both the number of hubs M, and their deployment strategy. We suggest here two different strategies, both based on a random hubs deployment with some consistency checking. The first one checks for non-zero neighborhood cardinality only, and the second one for a minimum amount of neighborhood diversity. Both strategies will be used in the next sections and are detailed as follows.

- Coordinates of the hubs are randomly drawn in the unit square according to a uniform distribution. Every new hub must have at least one sensor node in its neighborhood, otherwise it is discarded and a new one is drawn. In this way, the cardinality of all neighborhoods is $|\mathcal{N}_k| \geq 1$, $\forall k$. We indicate this strategy as `rnd-H`.
- Coordinates of the hubs are randomly drawn in the unit square according to a uniform distribution. Every new hub must have a neighborhood that is non-negligibly different with respect to that of the already present ones, otherwise it is discarded and a new one is drawn. In mathematical terms, being j an already placed hub and k the new one, and being $\mathcal{N}_j$ and $\mathcal{N}_k$ their neighborhood, the number of nodes that are in $\mathcal{N}_k$ but not in $\mathcal{N}_j$ normalized by the number of nodes in $\mathcal{N}_k$ has a lower bound given by $d_{\mathcal{N}}$, i.e., $|\mathcal{N}_k \setminus \mathcal{N}_j| \,/\, |\mathcal{N}_k| \geq d_{\mathcal{N}}$, $\forall j < k$. A simple example illustrating the neighborhood diversity concept is illustrated in Figure 3. We refer to this strategy, illustrated in Algorithm 1, as `div-H`.

Algorithm 1 Adaptive positioning of M hubs

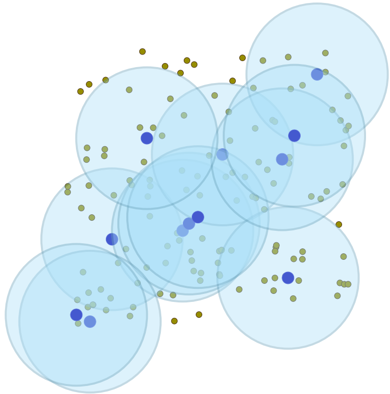
procedure DIVH(M, r_{TX}, $d_{\mathcal{N}}$)
  $\Theta \leftarrow \emptyset$
  $\mathcal{N} \leftarrow \emptyset$
  do M **times**
    cond $\leftarrow$ True
    while *cond* **do**
      $\theta_k \leftarrow$ random position in $[0,\, 1]^2$
      $\mathcal{N}_k \leftarrow$ sensors subset with distance $\leq r_{TX}$
      cond $\leftarrow$ False
      for all $\mathcal{N}_j \in \mathcal{N}$ **do**
        if $|\mathcal{N}_k \setminus \mathcal{N}_j| < d_{\mathcal{N}} |\mathcal{N}_k|$ **then**
          cond $\leftarrow$ True
        end if
      end for
    end while
    $\Theta \leftarrow \Theta \cup \theta_k$
    $\mathcal{N} \leftarrow \mathcal{N} \cup \mathcal{N}_k$
  end do
  return Θ, $\mathcal{N}$
end procedure
Θ: Set of hubs coordinates
$\mathcal{N}$: Set of hubs neighborhoods

Note that `div-H` includes `rnd-H` as a prerequisite to be able to compute $|\mathcal{N}_k \setminus \mathcal{N}_j| \,/\, |\mathcal{N}_k|$, while `rnd-H` can be considered a special case of `div-H` with $d_{\mathcal{N}} = 0$. Note also that none of the two strategies ensures a coverage of the whole unit square. As a consequence, sensor nodes may exist that are not covered by any hubs, in particular if M is small. This can be clearly observed in Figure 4, illustrating two examples of coverage for the `rnd-H` and the `div-H` strategies, respectively, with $N = 128$, $M = 12$, $r_{TX} = 0.25$ and $d_{\mathcal{N}} = 0.25$ (for to the `div-H` case only).

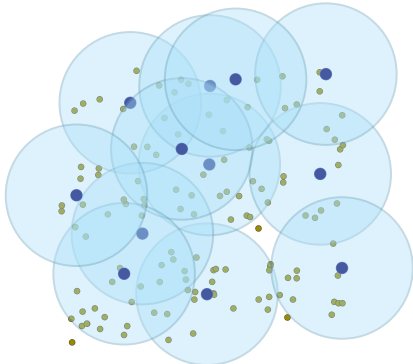
C. Compressed sensing design

The design of the readings pre-processing stage in the hubs follows the usual CS guidelines. Two parameters can be tuned to improve parameters: the number of measurements m and the sensing matrix A. Since the former has an impact on performance that is trivial, we focus here on the latter.

Temporarily neglecting the additive noise terms η_k considered in previous section, let us assume that the generic j-th measurement $y_j = \sum_{k=0}^{N-1} A_{j,k} x_k$, with $j = 0, \ldots, m-1$, is generated by the hub u, with $u = 0, \ldots, M-1$, by collecting



(a)



(b)

Fig. 4. Example of coverage for the proposed strategies, for $N = 128$, $M = 12$ and $r_{TX} = 0.25$. (a): rnd-H; (b): div-H, with $d_{\mathcal{N}} = 0.25$.

and linearly combining all the readings from the $n_u = |\mathcal{N}_u|$ nodes in its neighborhood $\mathcal{N}_u$. This implies that $A_{j,k} = 0$ if $k \notin \mathcal{N}_u$.

To allow a simpler notation, let us introduce the vector $a \in \mathbb{R}^n$ as the generic j-th row of A, i.e., $a = A_{j,\cdot}{}^\top$. With this, the generic j-th measurement is given by the scalar product

$$y_j = a^\top x = a_{|u}{}^\top x_{|u} \tag{2}$$

where the $\cdot_{|u}$ is an operator that, given an input indexed quantity, returns only the elements whose indexes are in $\mathcal{N}_u$. In other words, $a_{|u} \in \mathbb{R}^{n_u}$ and $x_{|u} \in \mathbb{R}^{n_u}$ are two vectors containing only the non-zero elements of a and the readings of the neighborhood $\mathcal{N}_u$, respectively.

According to well-known CS guarantees, the non-zeros elements of a (i.e., the elements of $a_{|u}$) can be taken as instances of zero-mean and unit-variance Gaussian random variables, independently of the each other. We refer to this strategy as rnd-CS.

Yet, if the second-order prior $\mathcal{X}$ for the considered input signal is known, it can be exploited to improve CS performance. This is what the rakeness concept, developed in [24], [25], [38], [46], [47], does by adapting the second-order statistic of a to that of $\mathcal{X}$ thus increasing the ability of the linear combination in (2) to *rake* energy from x. In other words, the rakeness approach introduces correlation among the elements of a single row of A, while all rows of A are still independent of each other as in standard CS.

More formally, a slight modification with respect to the rakeness mathematical framework in [25] is required, to cope with the constraint that only the terms in $a_{|u}$ need to be effectively designed, as described in the following.

Let $\mathcal{X}_{|u} \in \mathbb{R}^{n_u \times n_u}$ be the second order statistic characterization of $x_{|u}$. The average energy of the generic measurement as in (2) is

$$\mathbf{E}\left[a_{|u}{}^\top x_{|u} x_{|u}{}^\top a_{|u}\right] = \operatorname{tr}\left(\mathbf{E}\left[a_{|u} a_{|u}{}^\top\right] \mathbf{E}\left[x_{|u} x_{|u}{}^\top\right]\right) = \operatorname{tr}\left(\mathcal{A}_{|u} \mathcal{X}_{|u}\right)$$

where $\mathcal{A}_{|u} = \mathbf{E}[a_{|u} a_{|u}{}^\top]$ is the correlation matrix of $a_{|u}$ (i.e., the non-zeros elements of a), and where $\operatorname{tr}(\cdot)$ is the trace of its matrix argument.

Raked energy can be increased by generating vectors $a_{|u}$ whose correlation matrix $\mathcal{A}_{|u}$ is the solution of the optimization problem

$$\max_{\mathcal{A}_{|u}} \quad \operatorname{tr}\left(\mathcal{A}_{|u} \mathcal{X}_{|u}\right) \tag{3a}$$
$$\text{s.t.} \quad \mathcal{A}_{|u} = \mathcal{A}_{|u}{}^\top \tag{3b}$$
$$\text{s.t.} \quad \mathcal{A}_{|u} \succeq 0 \tag{3c}$$
$$\text{s.t.} \quad \operatorname{tr}\left(\mathcal{A}_{|u}\right) = n_u \tag{3d}$$
$$\text{s.t.} \quad \operatorname{tr}\left(\mathcal{A}_{|u}{}^2\right) \leq \frac{1}{2} n_u{}^2 \tag{3e}$$

where (3b) and (3c) ask for a symmetric and positive semidefinite $\mathcal{A}_{|u}$, respectively (i.e., $\mathcal{A}_{|u}$ is a feasible correlation matrix), and (3d) sets the energy of $a_{|u}$ according to the number of nodes in $\mathcal{N}_u$. Conversely, as discussed in detail in [24], [25], the aim of (3e) is to guarantee a minimum randomness level for $a_{|u}$ in order to span the whole signal space and allow a correct reconstruction even for the instances of x that are observed with a smaller frequency.

Following [25], the analytic solution of the above optimization problem is given by

$$\mathcal{A}_{|u} = \frac{1}{2}\left(\frac{n_u \mathcal{X}_{|u}}{\operatorname{tr}\left(\mathcal{X}_{|u}\right)} + I_{n_u}\right) \tag{4}$$

where I_{n_u} is the $n_u \times n_u$ identity matrix.

We indicate with rak-CS a second option for generating A, where every row is randomly generated using jointly-Gaussian variables such that the corresponding $a_{|u}$ is characterized by a correlation matrix (4).

D. *Energy costs*

Finally, also energy quantities are degrees of freedom. In our simplified model we have considered the energy E_{WAN} required to long range transmit the generic linear combination y_j, and the energy E_{TX} and E_{RX} required to short range transmit and receive a reading x_k, respectively. According to our model, the total energy required by all local transmissions is given by $n_{\mathrm{TX}} E_{\mathrm{TX}}$, where n_{TX} is the number of "readable"

TABLE II
SYSTEM CONFIGURATIONS USED AS REFERENCE CASES WITH CORRESPONDING AVERAGES FOR NUMBER OF TX AND NUMBERS OF RX.

Range	Technology	$E^b_{\rm TX}$ [nJ/bit]	$E^b_{\rm RX}/E^b_{\rm TX}$	Ref.
SR	IEEE 802.15.4	109	0.65	[48]
	BLE	27	0.65	[49]
	WiFi	18	0.23	[50]
LR	LoRa	39600	–	[51]
	GSM	17757	–	[52]

TABLE III
TECHNOLOGICAL PARAMETERS IN THE SYSTEM DESIGN.

	description	values
$r^{(\rm nom)}_{\rm TX}$	maximum value of $r_{\rm TX}$ for a fixed comm. technology	0.3
$E^{(\rm nom)}_{\rm TX}$	nominal value of $E_{\rm TX}$ for $r^{(\rm nom)}_{\rm TX}$	
γ	ratio between energy for TX and RX in short range communication	0.65
ϵ	ratio between energy for short and long range transmission	$[5\times 10^{-6}, 5\times 10^{-2}]$

TABLE IV
DEGREES OF FREEDOM IN THE SYSTEM DESIGN.

	description	values	
	hub positioning	rnd−H	random positioning, $d_{\mathcal{N}}=0$
		div−H	heuristic in Subsection IV-B with $d_{\mathcal{N}}>0$
$d_{\mathcal{N}}$	neighborhood diversity	$\{0, 0.05, 0.1, 0.15, 0.2, 0.25\}$	
	sensing matrix design paradigm	rnd−CS	classical random coefficients
		rak−CS	rakeness-based coefficients

transmission, with $n_{\rm TX} \leq N$ due to the possible presence of nodes not covered by any hub. The energy required for receiving these readings depends on the hubs positions, and it is given by $\sum_k |\mathcal{N}_k| E_{\rm RX} = n_{\rm RX}E_{\rm RX}$. The long range link, relying on m transmissions, has an energy cost given by $mE_{\rm WAN}$. As already anticipated, we consider the amount of energy required for pre-processing in the hub negligible.

The energy values depend on the adopted communication protocol, that also sets the transmission range. As an example, we have indicated in Table II the value of the energy required for the transmission of one bit $E^b_{\rm TX}$ and the ratio between reception and transmission energy $E^b_{\rm RX}/E^b_{\rm TX}$ for some reference solutions implementing either a short range (SR) or a long range (LR) communication protocol. The $E^b_{\rm TX}$ can be used as a starting value for computing either $E_{\rm TX}$ (SR protocols) or $E_{\rm WAN}$ (LR protocols); the ratio $E^b_{\rm RX}/E^b_{\rm TX}$ is a good estimator for $E_{\rm RX}/E_{\rm TX}$. The transmission range is not indicated in the table as it depends on many factors, but for all solutions is in the tens of meters range for SR, and in the kilometer range for the LR.

Interestingly, all reference solutions also allow a reduction of $E^b_{\rm TX}$ with a consequent reduction of the transmission range for energy saving purposes. We will exploit this in the following for the local communication introducing a quadratic dependence of $E_{\rm TX}$ from $r_{\rm TX}$. In detail, indicating with $E^{(\rm nom)}_{\rm TX}$ and $r^{(\rm nom)}_{\rm TX}$ the nominal values of energy required to transmit a reading and the transmission range of the selected communication protocol, respectively, the actual value of $E_{\rm TX}$ is given by

$$E_{\rm TX} = E^{(\rm nom)}_{\rm TX}\left(\frac{r_{\rm TX}}{r^{(\rm nom)}_{\rm TX}}\right)^2 \quad (5)$$

Furthermore, instead of directly using energy values, we focus on the two dimensionless quantities $\gamma = E_{\rm RX}/E^{(\rm nom)}_{\rm TX}$ and $\epsilon = E^{(\rm nom)}_{\rm TX}/E_{\rm WAN}$. The former, according to Table II, is set to $\gamma = 0.65$. Instead, the latter is considered in a wide range, with particular attention to the two corner cases of the table, given by $\epsilon = 5\cdot 10^{-4}$ and $\epsilon = 5\cdot 10^{-3}$.

V. SETTING AND NUMERICAL EVIDENCES

The effectiveness of the proposed design has been investigated by Montecarlo simulations for a huge number of different system configurations. The values considered for the parameters discussed in the Section III are listed in Table I, while Table III an Table IV list technological parameters that characterize communication protocols and degrees of freedom in the system design as described in Section IV. With the set of parameters in Table III it is possible to identify a pair of communication technologies, one for short and one for long range transmission, while the hub position and data compression mechanism depend on the degree of freedom in Table IV.

For the considered class of input signals, we focus here on the signal sparsity with $\kappa = \{6, 12\}$ in order to account different effectivenesses for the entire CS framework. The corresponding sparsity basis is the Fourier of adjacency matrix $D = \mathcal{F}[C]$ (with $d_{th} = 0.15$) which refers, for each trial, to a different set of sensor nodes randomly positioned in the unit square. Input signal characterization is completed by the empirical evaluation of $\mathcal{X}$. This is computed over a training set composed by 10000 signal instances.

For the network configuration, we focus here on neighborhood characterization that depends on the number of hubs M, on the adopted hub positioning policy and on the sensor transmission range $r_{\rm TX}$ with: $M \in \{14, 16, 18, 20, 22, 24\}$. Hubs are drawn according to both rnd−H and div−H and $r_{\rm TX} \in \{0.2, 0.25, 0.3\}$. For div−H we have $d_{\mathcal{N}} \in \{0.05, 0.1, 0.15, 0.2, 0.25\}$.

As described in Section III-C, the M hubs are cyclical used in the computation of the m-dimensional measurements vector y, i.e., each hub computes at least $\lfloor m/M \rfloor$ measurements.

The computation of y refers to the adopted policy for data compression. Two different sensing matrices A, obtained by drawing non-null elements either as instances of independent and identically distributed Gaussian random variable, i.e, by applying rnd−CS, or as instances of a Gaussian process with a correlation profile as in (4), i.e., by following the rak−CS approach.

Accordingly, all the possible combinations described above give rise to a total amount of 864 combinations. For each one of these cases, performance has been evaluated for a number of measurements m given by all possible integers in the range $[4, N]$, and by averaging 1500 different trials. Signals recovery problem (1) was solved according to [53]. Quality of the reconstructed instances $\hat{x}$ is evaluated by means of the Reconstruction Signal-to-Noise Ratio (RSNR), defined as

$$\mathrm{RSNR} = 20\log_{10}\left(\frac{\|x\|_2}{\|x-\hat{x}\|_2}\right)$$

The main figure of merit considered in this paper is the Probability of Correct Reconstruction for a given quality of service, PCR_Q, that is defined as the probability that the RSNR exceeds a threshold Q

$$\mathrm{PCR}_Q = \Pr\{\mathrm{RSNR} \geq Q\}$$

We prefer this figure of merit with respect to the average value of the RSNR because it also gives indications on the variance of the reconstruction quality. The PCR_Q is, in fact, capable of revealing undesired situations that the simple observation of the average RSNR value may mask. In particular we will focus on $\mathrm{PCR}_Q = 0.95$, implying that the required minimum RSNR value is obtained at least 95% of the times.

For a certain system configuration, i.e., a certain set of values assigned to the features described below, it is possible to identify an energy cost needed to send all the information to the data collector. The overall cost, as detailed in Section IV-D, is made by three different contributions: energy required by sensors to transmit readings to the hubs, energy required by the hubs to receive readings, and energy required by the hubs to transmit the computed measurement. Mathematically:

$$E_{\mathrm{CS}} = mE_{\mathrm{WAN}} + n_{\mathrm{TX}}\left(\frac{r_{TX}}{r_{\mathrm{TX}}^{(\mathrm{nom})}}\right)^2 E_{\mathrm{TX}}^{(\mathrm{nom})} + n_{\mathrm{RX}}E_{\mathrm{RX}}$$

where we have taken into account the possibility to reduce E_{TX} by reducing the r_{TX} according to (5).

We normalize this energy to that required by a straightforward acquisition scheme where N readings are long range transmitted to the data collector, i.e., $E_0 = NE_{\mathrm{WAN}}$. The obtained figure of merit is

$$\frac{E_{\mathrm{CS}}}{E_0} = \frac{m}{N} + \frac{n_{\mathrm{TX}}\left(\frac{r_{TX}}{r_{\mathrm{TX}}^{(\mathrm{nom})}}\right)^2 + n_{\mathrm{RX}}\gamma}{N}\epsilon \tag{6}$$

that depends on the two dimensionless quantities γ and ϵ defined in Subsection IV-D. Clearly, values of E_{CS}/E_0 lower than 1 indicate energy saving with respect to the straightforward approach.

A trade-off between two defined figures of merit (the quality of service Q and the energies ratio E_{CS}/E_0) is expected where a higher values of Q implies a lower energy saving.

A. Numerical Evidences

According to Table I, a very large number of scenarios can be identified by $\{\kappa, r_{\mathrm{TX}}, M, p_{\mathrm{f}}\}$. Additionally, the communication technologies are modeled by ϵ, the hub positioning depends on $d_{\mathcal{N}}$ (the rnd−H is achieved for $d_{\mathcal{N}} = 0$) and the adopted compression scheme is one among rnd−CS and rak−CS. The considered value of m is the smallest one that guarantees the desired PCR_Q.

TABLE V
SYSTEM CONFIGURATIONS USED AS REFERENCE CASES WITH CORRESPONDING AVERAGES FOR NUMBER OF TX AND NUMBERS OF RX.

	κ	M	$d_{\mathcal{N}}$	r_{TX}	$\mathbf{E}[n_{\mathrm{TX}}]$	$\mathbf{E}[n_{\mathrm{RX}}]$	rak−CS div−H	rak−CS rnd−H	rnd−CS div−H	rnd−CS rnd−H
							$\min_m\{\mathrm{PCR}_{55\,\mathrm{dB}} \geq 0.95\}$			
SYS1	6	16	0.2	0.3	126.7	461.0	45	56	67	82
SYS2	6	22	0.1	0.25	125.2	454.2	51	60	74	84
SYS3	12	16	0.1	0.3	125.4	455.6	78	92	120	-
SYS4	12	24	0.25	0.2	123.9	331.4	90	-	123	-

By now, we neglect the impact of both ϵ and p_{f} and we limit ourselves to consider only four configurations as case studies. We indicate them with labels from SYS1 to SYS4 that correspond to parameter values in Table V. To give an idea of the connectivity generated by those configurations, we also propose in the table the average number of links between nodes and hubs, expressed in terms of average number of achieved transmissions $E[n_{\mathrm{TX}}]$ and of total number of readings received by hubs $E[n_{\mathrm{RX}}]$. $E[n_{\mathrm{TX}}]$ is determined by the number of nodes whose reading cannot be received by any hub, while $E[n_{\mathrm{RX}}]$ refers to hub diversity along with M and r_{TX}.

For each of these reference cases we propose reconstruction performance as a function of m for any possible combination of rnd−CS or rak−CS, and of rnd−H or div−H. Results can be seen in Figure 5 in terms of probability of correct reconstruction given a 55 dB quality of service (PCR_{55dB}).

The rak−CS approach largely outperforms the rnd−CS one. Conversely, there is a non-negligible difference between rnd−H and div−H, with slightly advantage for the div−H approach, excluding SYS4 (characterized by a large M with small r_{TX}) for which also an improvement is more evident.

The figure has to be read as follows. By considering SYS1 with div−H and rak−CS, the desired quality of service $Q = 55$ dB is achieved with a probability higher than 0.95 for any value of m larger than 45. The computation of m is fundamental to assess long-range communication costs. The values of m for the other options are shown in Table V.

Once that the m value satisfying the desired quality of service has been computed, it is possible to evaluate the communication costs in terms of the normalized energy E_{CS}/E_0. Results for this second figure of merit are depicted in Figure 6 for both sparsity levels and as a function of ϵ, for the constant value of $\gamma = 0.65$. Results do not refer to a single system configuration, but each plot in the the obtained profiles refers to the configuration ensuring the lowest normalized energy among all configurations given by all possible values of m, M, r_{TX} and $d_{\mathcal{N}}$ that guarantee $\mathrm{PCR}_{55\,\mathrm{dB}} \geq 0.95$. Each plot reports performance for rnd−CS and rak−CS, in combination with rnd−H and div−H. Note that rnd−H implies $d_{\mathcal{N}} = 0$, i.e., the reported lowest energy ratios span combinations of m, M and r_{TX} only.

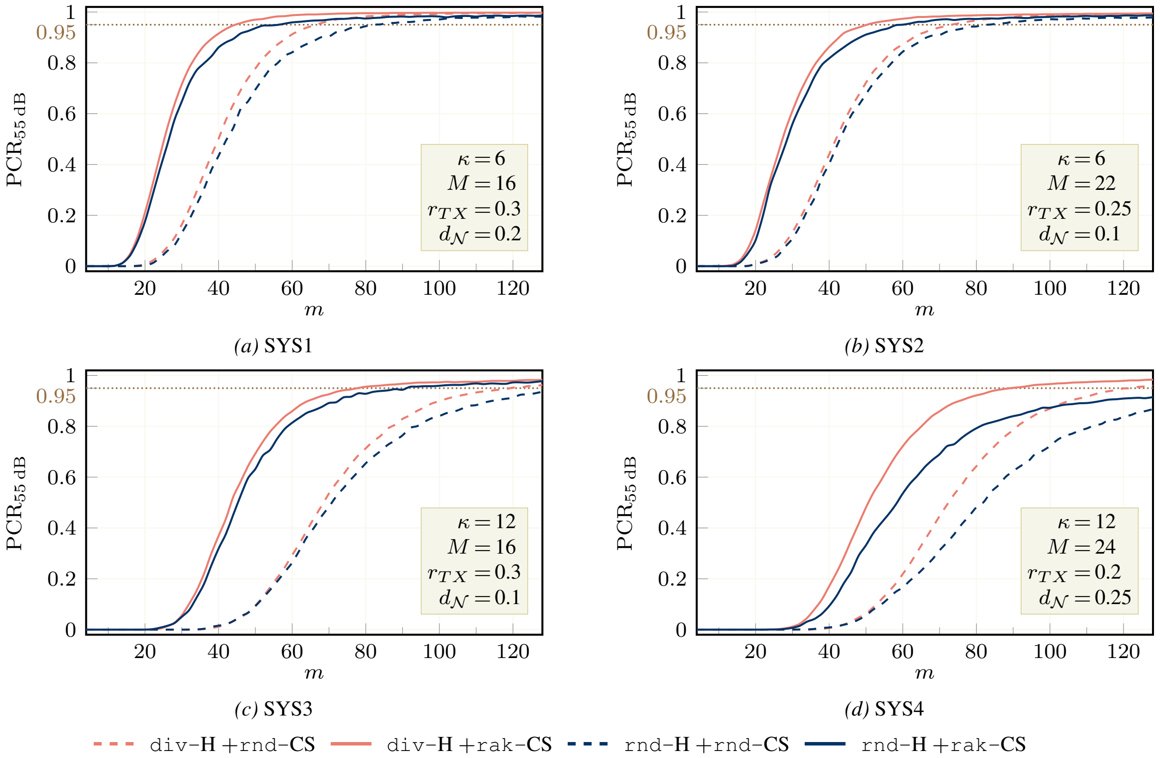


Fig. 5. PCR against m for the different system configurations of Table V. `rnd-H` lines refer to $d_{\mathcal{N}} = 0$.

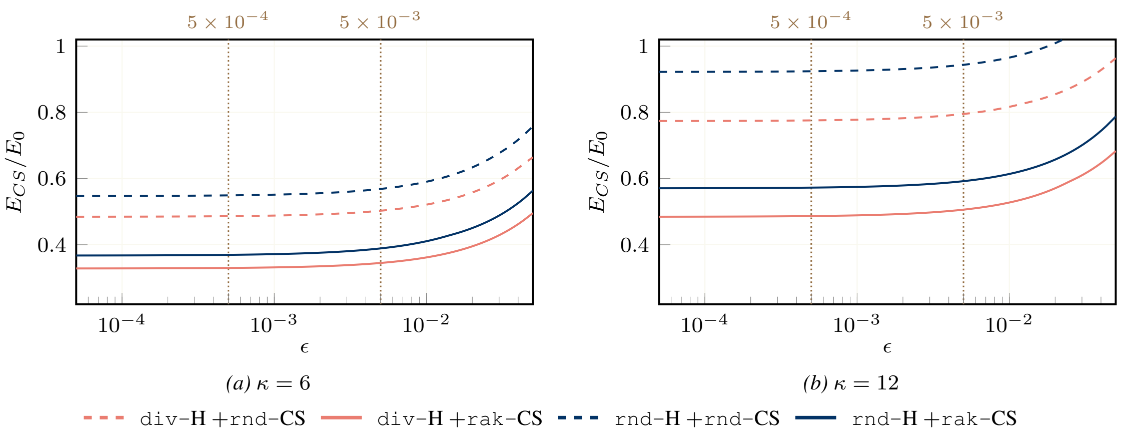


Fig. 6. Energy ratio against ϵ, for $\kappa = \{6, 12\}$ and with $\gamma = 0.65$. For each value of ϵ the corresponding E_{CS}/E_0 is the lowest along all possible combinations of m, M, r_{TX} and $d_{\mathcal{N}}$ that guarantees $\mathrm{PCR}_{55\,\mathrm{dB}} \geq 0.95$. `rnd-H` lines refer to $d_{\mathcal{N}} = 0$.

With $5 \times 10^{-4} \leq \epsilon \leq 5 \times 10^{-3}$, i.e., for values of ϵ intermediate between the two corner cases of Table II, performances are almost flat since energy cost is dominated by the long range transmissions that is proportional to m. This means that the lowest energy is always given by the same configuration. Optimal system configuration changes only for extreme values $\epsilon > 10^{-2}$.

Focusing on values of ϵ associated to technologies considered in Table II, the simultaneous adoption of `div-H` and `rak-CS` outperforms others possible combination of CS approach and hub positioning with $E_{\mathrm{CS}}/E_0 \approx 0.33$ for $\kappa = 6$ and $E_{\mathrm{CS}}/E_0 \approx 0.48$ for $\kappa = 12$. Here we have $(m, M, r_{\mathrm{TX}}, d_{\mathcal{N}})$ equal to $(42, 16, 0.3, 0.25)$ for $\kappa = 6$ and $(62, 22, 0.3, 0.25)$ for $\kappa = 12$. For this reason, in the rest of the paper we will focus only on the `div-H` and `rak-CS` case.

The proposed results show that the optimal configuration (including transmission cost) mainly depends on the desired quality of service so that a quality-energy trade-off can be investigated. Different levels of Q as a function of E_{CS}/E_0, always considering $\mathrm{PCR}_Q \geq 0.95$, are depicted in Figure 7 for $\kappa = \{6, 12\}$. Figure 7 shows also profiles for few values of ϵ that exploit current communication technologies along with the extreme case $\epsilon = 10^{-2}$ in which energy ratio between

This is the author's version of an article that has been published in this journal. Changes were made to this version by the publisher prior to publication.
The final version of record is available at http://dx.doi.org/10.1109/JETCAS.2018.2846884

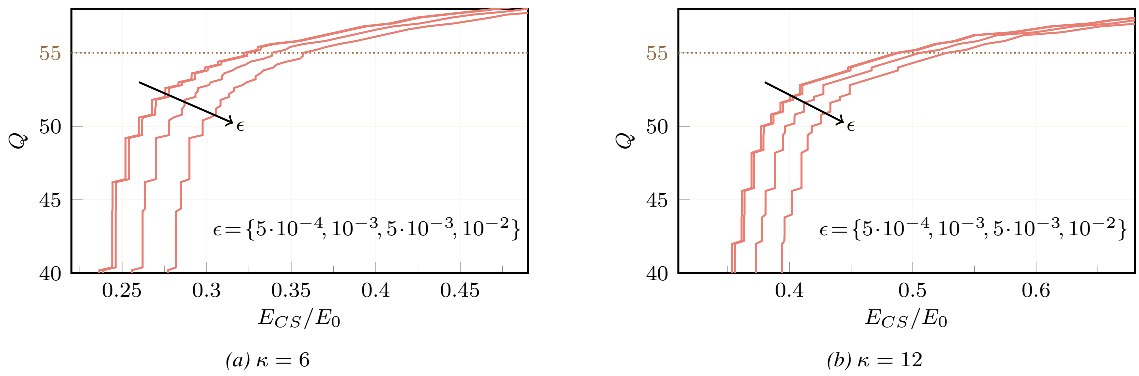


Fig. 7. Trade-offs between quality of service Q and energy ratio E_{CS}/E_0 for $\mathtt{div}{-}H + \mathtt{rak}{-}CS$ and for $\kappa = \{6, 12\}$ with $\gamma = 0.65$. Each line reports the lowest E_{CS}/E_0 along all possible combinations of m, M, r_{TX} and $d_{\mathcal{N}}$ that guarantee $\mathrm{PCR}_Q \geq 0.95$. rnd-H lines refer to $d_{\mathcal{N}} = 0$.

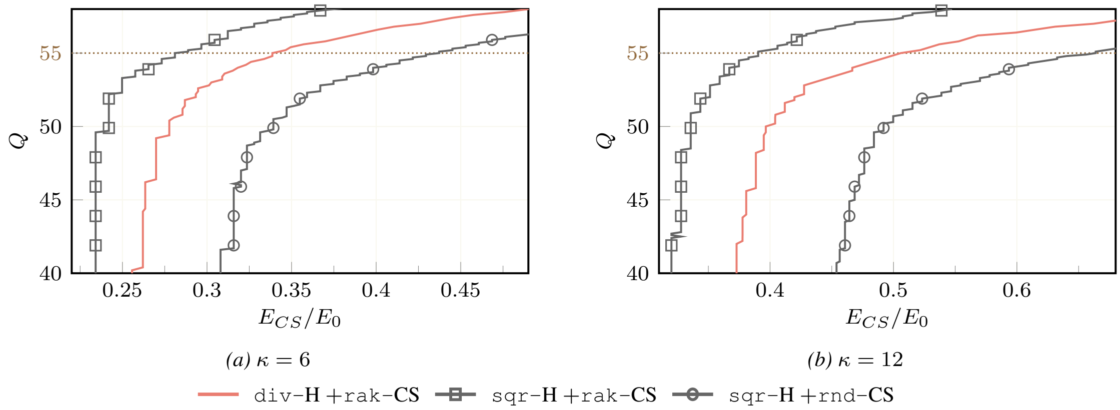


Fig. 8. Trade-offs between quality of service Q and energy ratio E_{CS}/E_0 for $\mathtt{div}{-}H + \mathtt{rak}{-}CS$ (with $d_{\mathcal{N}} = 0.25$) along with cases where hubs are on a regular grid (sqr-H) and either rnd-CS or rak-CS are considered for data compression. Results are for $\epsilon = 5 \times 10^{-3}$, $\gamma = 0.65$ and for $\kappa = \{6, 12\}$. Each line reports the lowest E_{CS}/E_0 along all possible combinations of m, M and r_{TX} that guarantee $\mathrm{PCR}_Q \geq 0.95$.

short and long range communication is drastically increased. Even for possible unfavorable future trends, represented by $\epsilon = 10^{-2}$, the energy saving obtained with $Q \leq 55\,\mathrm{dB}$ is greater than 65% for $\kappa = 6$ and it is greater than 45% for $\kappa = 12$.

Furthermore, for $Q \leq 50\,\mathrm{dB}$ an increasing in Q has a negligible cost in terms of overall energy consumption while for $Q > 50\,\mathrm{dB}$ the additional energy cost needed to increase Q rapidly grows. This interesting phenomenon is due to an intrinsic property of the CS paradigm: when strong signal degradation is allowed, i.e., $Q \ll \mathrm{ISNR}$, the slope of the dependency of the average performance on m is high. Conversely, average performance slowly increase with m when Q is close to ISNR.

As discussed in Section III-B, random hub positioning has been introduced to cope with possible physical constraints that do not permit to completely control the hubs deployment. Nevertheless, a comparison of the proposed (div-H +rak-CS) approach with hub positioning based on a regular pattern is depicted in Figure 8. These results refer to hubs deployed on a square grid, named sqr-H, that is a trivial solution for the coverage of a square area. As in the previous cases, trade-offs for sqr-H refer to the best option along different system configurations identified by $M = \{1, 4, 9, 16, 25\}$ and $r_{\mathrm{TX}} = \{^{0.75}/\sqrt{2M}, {}^{1}/\sqrt{2M}, {}^{1.25}/\sqrt{2M}, {}^{1.5}/\sqrt{2M}\}$.

The proposed div-H +rak-CS outperforms sqr-H combined with rnd-CS. Moreover the adoption of rak-CS with hubs on a regular grid guarantees the highest energy saving for all considered values of Q. Results suggest two important remarks: i) rak-CS outperforms rnd-CS also in case of hubs on a regular grid; ii) the limited deviation between profiles for sqr-H +rak-CS and div-H +rak-CS accounts for the possible physical constraints imposed to the hub positioning.

Finally, it is possible to evaluate the role of the probability of failure in a short-range transmission p_{f} in Figure 9. Also here we refer to the div-H +rak-CS cases only and to the four system configurations in Table V. As figure of merit, we define the ratio between E_{CS} and $E_{CS,p_{\mathrm{f}}=0}$ where the latter corresponds to the already presented results while, here, the E_{CS} values account for the impact of p_{f}. As before, both E_{CS} and $E_{CS,p_{\mathrm{f}}=0}$ correspond to the minimum value of m that satisfies $\mathrm{PCR}_{55\,\mathrm{dB}} \geq 0.95$.

The shown profiles highlight the robustness of this framework to this hard to avoid phenomenon. In particular, such results are to be considered as a trade-off energy *vs* tolerance in missing single communications. As an example, for SYS2 $m = 51$ is enough to ensure $\mathrm{PCR}_{55\,\mathrm{dB}} \geq 0.95$ with $p_{\mathrm{f}} = 0$ (see Table V), while to tolerate a 20% failure rate in data

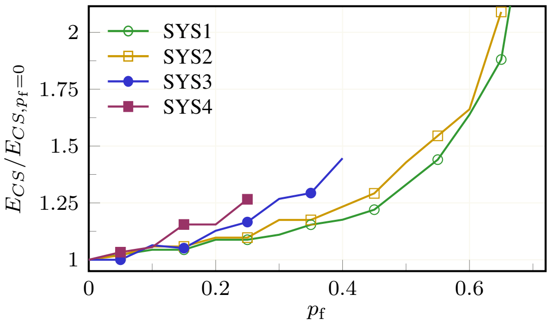


Fig. 9. Energy ratio against p_f, probability that a hub fails in receiving a single sample from a sensor, with the adoption of $\texttt{div}-H+\texttt{rak}-CS$ in the four system configurations of Table V. $E_{CS,p_\mathrm{f}=0}$ profiles are for $\epsilon = 10^{-3}$, $\gamma = 0.65$ and for the minimum m that guarantees $\mathrm{PCR}_{55\,\mathrm{dB}} \geq 0.95$.

reception ($p_\mathrm{f} = 0.2$) the minimum m is 56. In this example, 5 additional measurements are enough to compensate missing readings in the ratio of 1 out of 5.

This robustness to transmission failures is related to one of the CS properties discussed in [25], [54], i.e., a CS based acquisition system is able to tolerate some missing data that, in our system, correspond to sensors that do not communicate with any hub. This is the case of a sensor, staying in a single neighborhood, that is temporally unable to transmit its readings. In case of sensors that belong to more than one neighborhood, transmission failures towards a single hub is less restrictive than scenarios discussed in [25], [54].

The robustness to missing data from sensor to hubs is also related to the sparsity of the input signal, i.e., to the value of κ. For both SYS3 and SYS4 (where $\kappa = 12$) the system is less robust to this phenomenon with respect to the cases of SYS1 and SYS2 (where $\kappa = 6$).

VI. Conclusion

Compressed Sensing, especially its rakeness-based variant, is able to yield non-negligible lossy compression though entailing an extremely limited computational burden. Hence, it is the ideal candidate for the compression stage that may be implemented at the intermediate level in a sensor network architecture in which local hubs collect sensor readings by means of short-range communications and relay their compressed version to a remote concentrator by long-range transmission.

In the paper we were able to show that such an approach, paired with an empirical strategy aiming at promoting diversity between the set of readings collected by different hubs, is able to substantially reduce the energy requirements with respect to the no-compression and, though it clearly strips part of the redundancy in the sensed data, it is still quite robust with respect to the possible failure of local communications.

References

[1] J. Zheng, D. Simplot-Ryl, C. Bisdikian, and H. T. Mouftah, "The internet of things," *IEEE Communications Magazine*, vol. 49, no. 11, pp. 30–31, November 2011.
[2] L. D. Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 4, pp. 2233–2243, Nov 2014.
[3] R. Drath and A. Horch, "Industrie 4.0: Hit or hype? [industry forum]," *IEEE Industrial Electronics Magazine*, vol. 8, no. 2, pp. 56–58, June 2014.
[4] H. Hu, Y. Wen, T. S. Chua, and X. Li, "Toward scalable systems for big data analytics: A technology tutorial," *IEEE Access*, vol. 2, pp. 652–687, 2014.
[5] X. Liu *et al.*, "A fully integrated wireless compressed sensing neural signal acquisition system for chronic recording and brain machine interface," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 4, pp. 874–883, Aug 2016.
[6] S. Savazzi, V. Rampa, and U. Spagnolini, "Wireless cloud networks for the factory of things: Connectivity modeling and layout design," *IEEE Internet of Things Journal*, vol. 1, no. 2, pp. 180–195, April 2014.
[7] B. Warneke, M. Last, B. Liebowitz, and K. S. J. Pister, "Smart dust: communicating with a cubic-millimeter computer," *Computer*, vol. 34, no. 1, pp. 44–51, Jan 2001.
[8] T. Fischer, A. Agarwal, and H. Hess, "A smart dust biosensor powered by kinesin motors," *Nature Nanotechnology*, vol. 4, pp. 162–166, 2009.
[9] N. G. Elvin, N. Lajnef, and A. A. Elvin, "Feasibility of structural monitoring with vibration powered sensors," *Smart Materials and Structures*, vol. 15, no. 4, p. 977, 2006.
[10] N. Mohamed and I. Jawhar, "A fault tolerant wired/wireless sensor network architecture for monitoring pipeline infrastructures," in *2008 Second International Conference on Sensor Technologies and Applications (sensorcomm 2008)*, Aug 2008, pp. 179–184.
[11] J. P. Amezquita-Sanchez and H. Adeli, "Signal processing techniques for vibration-based health monitoring of smart structures," *Archives of Computational Methods in Engineering*, vol. 23, no. 1, pp. 1–15, Mar 2016.
[12] Y. Tachwali, H. Refai, and J. E. Fagan, "Minimizing hvac energy consumption using a wireless sensor network," in *IECON 2007 - 33rd Annual Conference of the IEEE Industrial Electronics Society*, Nov 2007, pp. 439–444.
[13] L. Schwiebert, S. K. Gupta, and J. Weinmann, "Research challenges in wireless networks of biomedical sensors," in *Proceedings of the 7th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '01. New York, NY, USA: ACM, 2001, pp. 151–165.
[14] E. Y. Chow, S. Chakraborty, W. J. Chappell, and P. P. Irazoqui, "Mixed-signal integrated circuits for self-contained sub-cubic millimeter biomedical implants," in *2010 IEEE International Solid-State Circuits Conference - (ISSCC)*, Feb 2010, pp. 236–237.
[15] J. Snoeijs, P. Georgiou, and S. Carrara, "CMOS Body Dust - Towards Drinkable Diagnostics," in *2017 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Oct 2016.
[16] D. L. Donoho, "Compressed sensing," *Information Theory, IEEE Transactions on*, vol. 52, no. 4, pp. 1289–1306, 2006.
[17] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, no. 2, pp. 21–30, 2008.
[18] J. Haupt, W. U. Bajwa, M. Rabbat, and R. Nowak, "Compressed sensing for networked data," *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 92–101, March 2008.
[19] M. T. Nguyen, K. A. Teague, and N. Rahnavard, "Ccs: Energy-efficient data collection in clustered wireless sensor networks utilizing block-wise compressive sensing," *Computer Networks*, vol. 106, pp. 171 – 185, 2016.
[20] R. Xie and X. Jia, "Transmission-efficient clustering method for wireless sensor networks using compressive sensing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 3, pp. 806–815, March 2014.
[21] X. Xing, D. Xie, and G. Wang, "Energy-balanced data gathering and aggregating in wsns: A compressed sensing scheme," *International Journal of Distributed Sensor Networks*, vol. 11, no. 10, p. 585191, 2015.
[22] M. Mangia, F. Pareschi, R. Varma, R. Rovatti, J. Kovaevi, and G. Setti, "Rakeness-based compressed sensing of multiple-graph signals for iot applications," *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2018. Doi: 10.1109/TCSII.2018.2821241.
[23] M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "Rakeness-based compressed sensing and hub spreading to administer short/long range communication tradeoff in iot settings," *IEEE Internet of Things Journal*, 2018. Doi: 10.1109/JIOT.2018.2828647.
[24] M. Mangia, R. Rovatti, and G. Setti, "Rakeness in the design of analog-to-information conversion of sparse and localized signals," *IEEE*

Transactions on Circuits and Systems I: Regular Papers, vol. 59, no. 5, pp. 1001–1014, May 2012.

[25] M. Mangia, F. Pareschi, V. Cambareri, R. Rovatti, and G. Setti, "Rakeness-based design of low-complexity compressed sensing," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 5, pp. 1201–1213, May 2017.

[26] R. Rajagopalan and P. K. Varshney, "Data-aggregation techniques in sensor networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 8, no. 4, pp. 48–63, Fourth 2006.

[27] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, Oct 2002.

[28] F. Chen, A. P. Chandrakasan, and V. M. Stojanović, "Design and Analysis of a Hardware-Efficient Compressed Sensing Architecture for Data Compression in Wireless Sensors," *IEEE Journal of Solid-State Circuits*, vol. 47, no. 3, pp. 744–756, Mar. 2012.

[29] G. Yang, V. Y. F. Tan, C. K. Ho, S. H. Ting, and Y. L. Guan, "Wireless compressive sensing for energy harvesting sensor nodes," *IEEE Transactions on Signal Processing*, vol. 61, no. 18, pp. 4491–4505, Sept 2013.

[30] J. Y. Park, M. B. Wakin, and A. C. Gilbert, "Modal analysis with compressive measurements," *IEEE Transactions on Signal Processing*, vol. 62, no. 7, pp. 1655–1670, April 2014.

[31] M. Mangia, D. Bortolotti, F. Pareschi, A. Bartolini, L. Benini, R. Rovatti, and G. Setti, "Zeroing for hw-efficient compressed sensing architectures targeting data compression in wireless sensor networks," *Microprocessors and Microsystems*, vol. 48, pp. 69–79, Feb. 2017, extended papers from the 2015 Nordic Circuits and Systems Conference.

[32] D. Cox, "Cochannel interference considerations in frequency reuse small-coverage-area radio systems," *IEEE Transactions on Communications*, vol. 30, no. 1, pp. 135–142, January 1982.

[33] J. Thornton, D. Grace, M. H. Capstick, and T. C. Tozer, "Optimizing an array of antennas for cellular coverage from a high altitude platform," *IEEE Transactions on Wireless Communications*, vol. 2, no. 3, pp. 484–492, May 2003.

[34] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Graph fourier transform," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2013, pp. 6167–6170.

[35] A. Sandryhaila and J. M. F. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Transactions on Signal Processing*, vol. 62, no. 12, pp. 3042–3054, June 2014.

[36] S. Chen, R. Varma, A. Singh, and J. Kovacevic, "Signal recovery on graphs: Random versus experimentally designed sampling," in *2015 International Conference on Sampling Theory and Applications (SampTA)*, May 2015, pp. 337–341.

[37] S. Chen, R. Varma, A. Singh, and J. Kovacevic, "Signal recovery on graphs: Fundamental limits of sampling strategies," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 4, pp. 539–554, Dec 2016.

[38] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A rakeness-based design flow for analog-to-information conversion by compressive sensing," in *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*. IEEE, May 2013, pp. 1360–1363.

[39] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[40] E. J. Candes and T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.

[41] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, no. 12, pp. 4655–4666, 2007.

[42] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18 914–18 919, Nov. 2009.

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

[44] F. Krahmer and R. Ward, "Stable and robust sampling strategies for compressive imaging," *IEEE Transactions on Image Processing*, vol. 23, no. 2, pp. 612–622, Feb 2014.

[45] M. A. Davenport, A. K. Massimino, D. Needell, and T. Woolf, "Constrained adaptive sensing," *IEEE Transactions on Signal Processing*, vol. 64, no. 20, pp. 5437–5449, Oct 2016.

[46] V. Cambareri, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A case study in low-complexity ecg signal encoding: How compressing is compressed sensing?" *IEEE Signal Processing Letters*, vol. 22, no. 10, pp. 1743–1747, Oct 2015.

[47] F. Pareschi, P. Albertini, G. Frattini, M. Mangia, R. Rovatti, and G. Setti, "Hardware-Algorithms Co-Design and Implementation of an Analog-to-Information Converter for Biosignals Based on Compressed Sensing," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 1, pp. 149–162, Feb. 2016.

[48] "A True System-on-Chip Solution for 2.4-GHz IEEE 802.15.4 and ZigBee Applications," Texas Instruments Datasheet, Feb. 2011, http://www.ti.com/lit/ds/symlink/cc2530.pdf.

[49] "CC2640R2F SimpleLink® Bluetooth™ low energy Wireless MCU," Texas Instruments Datasheet, Jan. 2017, http://www.ti.com/lit/ds/symlink/cc2640r2f.pdf.

[50] "ATWINC15x0-MR210xB - IEEE 802.11 b/g/n SmartConnect IoT Module," Microchip Technology Inc. Datasheet, 2017, http://ww1.microchip.com/downloads/en/DeviceDoc/70005304A.pdf.

[51] "SX1276/77/78/79 - 137 MHz to 1020 MHz Low Power Long Range Transceiver," Semtech Corporation Datasheet, Aug. 2017, http://www.semtech.com/images/datasheet/sx1276.pdf.

[52] "SARA-G3 series - Dual and quad-band GSM/GPRS module," u-blox Datasheet, Aug. 2017, https://www.u-blox.com/sites/default/files/SARA-G3_DataSheet_(UBX-13000993).pdf.

[53] E. van den Berg and M. P. Friedlander, "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2009.

[54] D. E. Bellasi, R. Rovatti, L. Benini, and G. Setti, "A low-power architecture for punctured compressed sensing and estimation in wireless sensor-nodes," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, no. 5, pp. 1296–1305, May 2015.

Mauro Mangia (S'09-M'13) received the B.Sc. and M.Sc. in Electronic Engineering and the Ph.D. degree in Information Technology from the University of Bologna (Bologna, Italy), respectively in 2005, 2009 and 2013. He is currently a Postdoctoral Researcher in the statistical signal processing group of ARCES - University of Bologna. In 2009 and 2012, he was a visiting Ph.D. student at the Ecole Polytechnique Federale de Lausanne (EPFL). His research interests are in nonlinear systems, Compressed Sensing, Internet of Things , Ultra-Wideband Systems, and Systems Biology. He was the recipient of the 2013 IEEE CAS Society Guillemin-Cauer Award and best student paper award at ISCAS2011. He is also the Web and Social Media Chair for ISCAS2018.

Alex Marchioni received the B.S. and M.S. degree (with honors) in electronic engineering from the University of Bologna, respectively in 2011 and 2015. In 2018, he joined the Department of Electrical, Electronic, and Information Engineering "Guglielmo Marconi" (DEI) of the University of Bologna, where he is currently working as research fellow. His research interests include compressed sensing, biomedical applications and signal processing for the Internet of Things and Big Data analytics.

Fabio Pareschi (S'05-M'08) received the Dr. Eng. degree (with honours) in Electronic Engineering from University of Ferrara, Italy, in 2001, and the Ph.D. in Information Technology under the European Doctorate Project (EDITH) from University of Bologna, Italy, in 2007. He is currently an Assistant Professor in the Department of Engineering, University of Ferrara. He is also a faculty member of ARCES - University of Bologna, Italy. He served as Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS - PART II (2010-2013). His research activity focuses on analog and mixed-mode electronic circuit design, statistical signal processing, compressed sensing, random number generation and testing, and electromagnetic compatibility. He was recipient of the best paper award at ECCTD 2005 and the best student paper award at EMC Zurich 2005.

Riccardo Rovatti (M'99-SM'02-F'12) received the M.S. degree in Electronic Engineering and the Ph.D. degree in Electronics, Computer Science, and Telecommunications both from the University of Bologna, Italy in 1992 and 1996, respectively. He is now a Full Professor of Electronics at the University of Bologna. He is the author of approximately 300 technical contributions to international conferences and journals, and of two volumes. His research focuses on mathematical and applicative aspects of statistical signal processing and on the application of statistics to nonlinear dynamical systems. He received the 2004 IEEE CAS Society Darlington Award, the 2013 IEEE CAS Society Guillemin-Cauer Award, as well as the best paper award at ECCTD 2005, and the best student paper award at EMC Zurich 2005 and ISCAS 2011. He was elected IEEE Fellow in 2012 for contributions to nonlinear and statistical signal processing applied to electronic systems.

Gianluca Setti (S'89-M'91-SM'02-F'06) received a Ph.D. degree in Electronic Engineering and Computer Science from the University of Bologna in 1997. Since 1997 he has been with the School of Engineering at the University of Ferrara, Italy, where he is currently a Professor of Circuit Theory and Analog Electronics and is also a permanent faculty member of ARCES, University of Bologna. His research interests include nonlinear circuits, implementation and application of chaotic circuits and systems, electromagnetic compatibility, statistical signal processing and biomedical circuits and systems. Dr. Setti received the 2013 IEEE CAS Society Meritorious Service Award and co-recipient of the 2004 IEEE CAS Society Darlington Award, of the 2013 IEEE CAS Society Guillemin-Cauer Award, as well as of the best paper award at ECCTD2005, and the best student paper award at EMCZurich2005 and at ISCAS2011. He held several editorial positions and served, in particular, as the Editor-in-Chief for the IEEE Transactions on Circuits and Systems - Part II (2006-2007) and of the IEEE Transactions on Circuits and Systems - Part I (2008-2009). Dr. Setti was the Technical Program Co-Chair ISCAS2007, ISCAS2008, ICECS2012, BioCAS2013 as well as the General Co-Chair of NOLTA2006 and ISCAS2018. He was Distinguished Lecturer of the IEEE CAS Society (2004-2005 and 2014-2015), a member of its Board of Governors (2005-2008), and he served as the 2010 President of CASS. He held several other volunteer positions for the IEEE and in 2013-2014 he was the first non North-American Vice President of the IEEE for Publication Services and Products.