Rakeness-based compressed sensing and hub spreading to administer short/long-range communication tradeoff in IoT Settings

# Rakeness-based Compressed Sensing and Hub Spreading to Administer Short/Long Range Communication Tradeoff in IoT settings

Mauro Mangia, Fabio Pareschi, Riccardo Rovatti, Gianluca Setti

***Abstract*—In common distributed sensing scenarios, a number of local Wireless Sensor Networks perform sets of acquisitions that must be sent to a central collector which may be far from the measurement fields. Hence, readings from individual nodes may reach their destination by exploiting both local and long-range transmission capabilities. The Compressed Sensing (CS) paradigm may help finding a convenient mix of the two options, especially if it follows the rakeness-based design flow that has been recently introduced.**

**CS is exploited by identifying local hubs that aggregate many sensor readings in a smaller number of quantities that are then transmitted to the central collector.**

**We here show that, depending on the relative cost of local versus long-range transmission, carefully administering the choice of the hubs, the breadth of the neighborhood from which they collect readings, as well as the coefficients with which those readings a linearly aggregated, one may significantly reduce the energy needed to sample the field.**

**Simulations indicate that savings may be over 50% for values of the parameters modeling nowadays local and long-range transmission technologies.**

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

## I. Introduction

As the next steps in the technology development see an increasing interaction between information processing and the physical world, the deployment and the very same concept of sensing is changing rapidly [1]–[4]. The word *sensor* no longer indicates only the device translating some physical quantity into an electrical signal, neither it limits to the conditioning and digitalization of such a signal to make it compatible with digital processing.

In order to fit into the grand view of an information processing structure overlying and improving physical reality, *sensors* must be able to play their role in a system in which processing and communication have to be administered at many different levels in order to reach a globally optimized use of resources. This is surely true when, for example, many sensors are deployed over an area and must report their readings to a remote collector. A certain amount of processing within the local sensor network may benefit the long-range transmission needed to reach the remote collector.

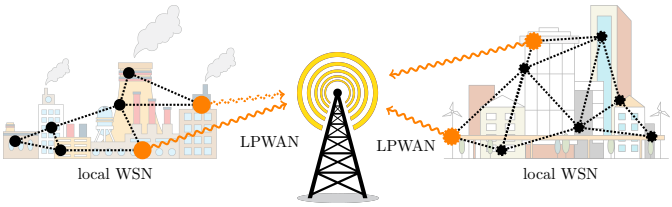



Fig. 1. A grand view of systems made of local Wireless Sensor Networks that communicate their readings to a geographically separated collector.

Though the paper deals with this issue from a methodological point of view, Figure 1 gives an intuitive representation of a structures in which sensors belong to a Wireless Sensor Network (WSN) whose nodes have local communication capabilities and finally deliver their acquisitions to the collector by means of long range transmissions in some Wide Area Network (WAN).

In this paper we investigate the trade-off between local and wide-area communication by arranging sensor nodes in neighborhoods, each of them controlled by a hub that aggregates data by means of a Compressed Sensing-based (CS) coding technique before communicating with the remote collector. While the concept of data aggregation is well studied [5], the application of CS-based techniques is relatively new in the literature [6]. With respect to previous proposals we here introduce three notable novelties: $i$) the application of the CS optimization technique known as *rakeness* [7], [8] to improve performance in terms of data compression; $ii$) a modified Dijkstra's algorithm to define hub neighborhoods under the assumption that not more than a predefined hops budget is used to transmit all sensor readings to the hub using a multi-hop protocol; $iii$) the algorithm according to which we select the hubs, that is specifically designed to cope with the proposed neighborhoods construction. Simulations show that the energy required to locally collect sensor readings and transmit compressed data to the remote collector is strongly reduced with respect to other CS-based approaches proposed in the literature.

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

TABLE I
ENERGY PER BIT USED IN SOME SHORT- AND LONG-RANGE COMMUNICATION PROTOCOLS. VALUES REFER TO THE MAXIMUM OUTPUT POWER.

| Range | Technology | $E_{\text{bit}}$ [nJ/bit] | Ref. |
|---|---|---|---|
| short | IEEE 802.15.4 | 109 | [10] |
| | BLE | 27 | [11] |
| | WiFi | 18 | [12] |
| long | LoRa | 39600 | [13] |
| | GSM | 17757 | [14] |

To describe and test the proposed scheme, the paper is organized as follows. In Section II we introduce the motivation of this work. We also present related works on CS-based data gathering techniques and briefly introduce some mathematical concepts used in the paper. Section III and IV introduce the proposed aggregation technique and the hub selection algorithms, respectively. Section V sketches the principles of classical and rakeness-based CS that are applied to locally collected data. Section VI describes a general framework that allows to simultaneously model the estimation accuracy of the scheme as well as its needs in terms of energy employed in transmissions. It also proposes a number of options for the connectivity graph and the sparsity basis that characterize a system and are worth testing to observe the effect of the degrees of freedom on the overall performance. Finally, Section VII performs a thorough numerical assessment of the configurations discussed in the previous section and shows how the overall method may be used to save resources from a global point of view.

## II. MOTIVATION AND RELATED WORK

The trade-off between local communication and wide-area transmission in the scenario of Figure 1 is investigated under the commonly adopted assumption [9] that a number of sensor nodes, all equal to each other, with limited processing power and highly constrained energy resources, are randomly deployed in a sensing area. The investigation is motivated by the fact that, assuming that the ratio between the distance covered by long-range and short-range communications is about $10^2$ (e.g., tens of meters vs. kilometers) and that no particular directivity can be provided by sensor nodes antennas, according to the free space propagation model one expects that the ratio between entailed powers is of the order of $10^4$.

This theoretical prediction is matched by actual consumption of nowadays implementations. Table I compares the transmission efficiency in terms of Joule per bit of some communication protocols commonly used in the most common machine-to-machine (M2M) communication scenarios [15]. The ratio $\epsilon$ between short- and long-range efficiencies varies in a wide range, depending on the considered protocols. Yet, we focus on the two corner cases identified by the table, that we approximate in $\epsilon = 5 \cdot 10^{-4}$ and $\epsilon = 5 \cdot 10^{-3}$. Values in this order of magnitude are more than enough to justify the introduction of substantial local data exchange to provide data aggregation and compression before long-range transmission is attempted.

The problem of data aggregation in WSN has been widely investigated in the Literature [5], and has been recently improved by the adoption of CS-based local processing techniques, mainly due to the low cost in terms of complexity. Differently from classic compression mechanism based on some domain transformation such as Fourier or Wavelet, CS is based on simple linear combinations of input samples. This approach lowers the additional cost in terms of required energy, thus making power budget typically dominated by local communication cost [8], [16]–[19]

Authors of [20] introduce RACS (Random Access Compressed Sensing) as a scheme that, thanks to a strong undersampling of nodes in underwater sensor networks, can improve energy and bandwidth efficiency. A similar approach has been proposed in [21] and addressed as STCDG (Spatio-Temporal Compressive Data Collection). The results of all these studies reveal that CS represents a good alternative to solutions based on standard compression schemes.

Yet, in many of the works that can be found in the Literature, CS is applied, either in centralized or in a distributed way, in sensor networks adopting a *multi-hop* scheme: sensor readings are propagated through the nodes before reaching the collector (typically indicated as sink). The advantage is that the energy required for propagating a reading increases linearly with the number of hops, that can be assumed proportional to distance, while in the open space model power should increase quadratically with distance. These works aim is to improve the lifetime of the WSN while ensuring a sufficient performance in terms of overall data reconstruction. This is the case of [6] that, given a multi-hop WSN, highlights the potential of applying CS to the data aggregation problem. In [22] authors propose CS as a mean to reduce communication cost by reducing the number of transmissions in large-scale WSNs. To this aim, they apply CS both in the spatial and also in the spatial-temporal domain. Authors of [23] focus on the trade-off between the energy spent for transmission and for compression using data sets gathered by a real-life deployment. In [24] the authors compare the results of a time-domain based CS scheme in terms of energetic costs with that achievable by conventional schemes. In [25] authors are able to write a joint optimization problem that allows to collect data at the sink, minimizing both the number of transmissions (i.e., the required energy) and the latency. Authors of [26] investigates performance, in terms of both energy and latency, of a proposed coding scheme based on CS where the linear combinations are computed in a distributed way.

Yet, all the aforementioned works assume that there is a single data collector (sink) located not far from the WSN. The scenario of Figure 1 requires a slightly different network modeling, similar to that envisioned in [15] for the future M2M growth in the IoT scenario with a heterogeneous wireless connectivity capability. To cope with this scenario, many works consider *clustering*, i.e. nodes are assigned to non-overlapping clusters, each of them controlled by a hub collecting their readings (possibly by means of a multi-hop protocol) and providing long-range communication capability.

A similar network model appeared in [9]. Authors consider microsensor networks where the data collector is located far

from the nodes, and proposed a dynamic and adaptive low-energy clustering approach known as LEACH. Yet, a very few works are considering a similar model within the CS framework. In [27] the clustering technique is considered, and the optimal number of clusters that provides the minimum power consumption is investigated. However, data 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. Authors of [28] propose an energy-balanced data gathering and aggregating scheme with an adaptive approach very similar to LEACH. Also here, the collector is assumed to be inside the sensing area. Finally, [29] investigates a clustering approach with a CS-based compression that finds the optimal number of clusters reducing the total amount of local hops needed to first collect data inter-cluster (by multi-hop protocol) and then transmit compressed data from each cluster to a main collector. Yet, similarly to previous works, the collector is located at the edge of the sensing area.

The scenario considered here is more similar to that in [9], where nodes communicate with a multi-hop protocol to some hubs. In our work hubs apply CS-base compression to readings before sending them to the collector. The main innovative aspects with respect to the aforementioned literature can be summarized as follows.

- We aim at reducing the joint energy required by both local WSN communication and WAN communication, while almost all of the cited works aim at reducing the WSN energy only. This is reflected into an improvement in the network life assuming a proper hub rotation policy [9].
- We introduce a modified Dijkstra's algorithm for defining hub neighborhoods, and the *hop budget* $H$ as a very simple design parameter to set the neighborhood size. This helps in the energy consumption design of the WSN, whose is directly related to $H$.
- We relax the definition of cluster introducing the *neighborhood*. It is possible for a node to not belong to any neighborhood, but also to belong to more than one. A CS system can easily cope with this, with beneficial effects on the performance [19], [30]. Such a situation has never been considered in the Literature [27]–[29].
- The compression achieved by CS is increased by applying the recently introduced rakeness-based design [7], [8].

In the following, we rely on some mathematical definitions and concepts that are briefly summarized here.

### A. Sensor nodes deployment

We assume that $N$ sensors numbered from 0 to $N-1$ are randomly deployed in the sensing field. At a given time, each of them acquires its reading $x_k$ with $k = 0, \ldots, N-1$. Mathematically, we introduce $x$ as the state vector of the sensing field by arranging readings as $x = (x_0, \ldots, x_{N-1})^\top$, where $\cdot^\top$ stands for vector transposition.

The final goal of the method we propose is to let the collector compute an estimate $\hat{x}$ of the original vector $x$. Given a desired estimation accuracy depending on $x - \hat{x}$, the trade-off between short- and long-range communication may be administered to reduce the overall energy need.

### B. Connectivity graph

We assume that all sensor nodes have the same fixed transmission power, and consequently a fixed transmission range [29]. Accordingly, communication capabilities of the nodes depend on their distance. We model this by introducing an undirected graph with a vertex at each node and the $N \times N$ adjacency matrix $C$ such that $C_{j,k} = 1$ if the $k$-th vertex/node can communicate with the $j$-th vertex/node and $C_{j,k} = 0$ if this cannot happen.

This very simple model of connectivity derives from the need of keeping the implementation of each node as simple as possible thus avoiding, for example, variable power transmission that would be better modeled by a complete weighted graph with communication costs attached to each edge.

In our setting, when feasible, communication happens by equal cost *hops* between nodes.

### C. Second-order signal statistics

One of the priors that is often available when designing acquisition systems is how the energy of the signal is distributed along its components.

This is accounted for by the second order statistics of the vector $x$, i.e., by its $N \times N$ correlation matrix $\mathcal{X} = \mathbf{E}[xx^\top]$. Signals with independent components feature a trivial correlation matrix made of individual variances $\mathcal{X}_{k,k} = \mathbf{E}[x_k^2]$ and products of the means $\mathcal{X}_{j,k} = \mathbf{E}[x_j]\mathbf{E}[x_k]$ for $j \neq k$.

Yet, real-world quantities usually feature some form of correlation between components and the resulting non-trivial $\mathcal{X}$ can be used as a prior to optimize acquisition of the vector $x$ just like power-spectrum can be used to properly design the acquisition of a time domain signal.

### D. Compressibility and sparsity

Real-world signals, especially high-dimensional ones, usually exhibit redundancy and thus are compressible. For vectors, the simplest form of compressibility is with respect to a suitable orthonormal basis and can be formalized arranging the vectors of such a basis in the matrix $D$, and saying that if the signal is expressed as $x = D\xi$, then the coefficient vector $\xi \in \mathbb{R}^N$ has few non-negligible components that are the only ones that would be needed to reconstruct the original signal.

An extreme form of compressibility is $\kappa$-*sparsity*, in which one knows that not more than $\kappa < N$ components of $\xi$ are non-null. The sparsity basis $D$ is clearly application-dependent and real world time-domain signals are usually found to be compressible or even sparse along some suitably defined Fourier-like or wavelet-like basis.

Our vector $x$ does not contain time-domain samples of a signal. Yet, relationships may exists between the nodes, different from belonging to a sequence ordered in time, that imply the existence of a sparsity basis.

Some authors have recently suggested [6], [31]–[33] that such relationships may be modeled by laying down a further graphs connecting the vertices/nodes with edges whose weight accounts for the link between the corresponding extrema. In our case, this would amount to a further $N \times N$ matrix $G$

such that $G_{j,k}$ is the weight of the edge connecting the $k$-th vertex with the $j$-th vertex. Assuming that such a matrix $G$ can be diagonalized as $G = D\Lambda D^{-1}$, with $D$ a non-singular matrix and $\Lambda$ the diagonal matrix containing the eigenvalues of $G$, evidence has been put forward that some signals (often indicated as *signals on graphs*) are compressible or sparse with respect to $D$.

As an intuitive justification for such a prior, note that the acquisition of time-domain signals can be modeled by associating each time instant to a vertex and setting $G_{k+1,k} = 1$ for each $k = 0, \dots, N-2$ to express their ordering. If we further set $G_{0,n-1} = 1$ and $G_{j,k} = 0$ in all other cases, the corresponding $x_k$ can be thought as samples of a periodic time-domain waveform.

In this case, it is easy to see that if $G = D\Lambda D^{-1}$ we have $D_{j,k} = e^{-2\pi \mathbf{i}\, jk/n}$ for $j, k = 0, \dots, n-1$, i.e., $D$ contains the Fourier basis. Being compressible or sparse with respect to $D$ is therefore equivalent of being compressible or sparse with respect to what can be taken as the generalization of the Fourier basis for graph-supported signals [34], [35].

### E. Compressed sensing

Compressed Sensing (CS) is a technique leveraging the sparsity prior to reduce the amount of scalars that needs to be used to identify a signal. It has been proven effective when dealing with structured signals like time-domain waveforms or images and has been recently proposed for the acquisition of graph-supported signals [6].

The fundamental idea is that, instead of acquiring all the values $x_k$ for $k = 0, \dots, N-1$, one may consider only a certain number of their linear combinations $y_j = \sum_{k=0}^{N-1} A_{j,k} x_k + \eta_j$ where $\eta_j$ is a noise component, for $j = 0, \dots, m-1$ and with $m < N$, called *measurements*.

Since $m < N$, if the measurements and the disturbances are arranged in the vectors $y$ and $\eta$, and the linear combination coefficients in the matrix $A$, then the relationship $y = Ax + \eta$ does not, in general, allow a straightforward estimation of $x$. Yet, if $x$ is known to be sparse with respect to the basis $D$, then $y = AD\xi + \eta$, where at most $\kappa$ components of $\xi$ are known to be non-zero. A number of recent theoretical and algorithmic developments guarantees that, exploiting this prior, $x$ can be effectively estimated with some known bounds on the estimation error depending on $\eta$ and on specific properties of the matrix $A$.

Beyond theoretical guarantees, it is fair to say that most of the practical interest in CS comes from few key facts:

- the actual performance of estimation algorithms largely outperforms the theoretical bounds allowing an effective recovery of $x$ from a small number of measurements, i.e., usually $m \ll N$ [36];
- the mathematical conditions on $A$ can be matched by using random matrices and, although the formal results depend on specific matrix distributions [37], in practice a wide class of random matrices allows for effective signal recovery [38]–[41];
- leveraging on this, it is possible to substantially improve acquisition performance by adapting the statistical distribution of $A$ to that of the signal $x$, in particular to its second-order statistics in $\mathcal{X}$. Once the distribution of $A$ is chosen, a single instance can be drawn and used for sensing. This is what rakeness-based design of $A$ does [7], [36];
- the amount of processing entailed by CS is negligible as it amounts only to linear combinations [18], [19], [42]. This perfectly suits the low resource budget of sensor nodes.

All the above can be combined in a data collection scheme that uses inter-node communications over the connectivity graph to gather subsets of the readings and computes the corresponding linear combinations $y_j$ by means of suitable coefficients, possibly designed with a rakeness-based approach to CS [8]. The measurements, that come in smaller number with respect to the sensor readings, are the quantities transmitted to the central collector for reconstructing $\hat{x}$.

This blending of local communication and CS is a distinctive and useful improvement of what is classically proposed for the acquisition of graph-supported signals where $A$ is made of $m$ rows of the $N \times N$ identity matrix [33] thus implementing a subsampling instead of a true, possibly optimized, linear combination.

## III. Collecting readings by local communication

Some sensing nodes are promoted to the role of hub, and collect readings from nearby nodes (their *neighborhood*) providing linear combinations of them to the common collector.

Neighborhood design should be accurately investigated, as their size is critical for system performance [27], [29]. Given a hub $u$, we indicate with $W$ the associated neighborhood and with $n$ its size, i.e., $n = |W|$.

In this work, we set $W$ as the largest set of nodes whose readings can be transferred to $u$ within a *hop budget* $H$, i.e., within a number of short-range transmissions not higher than $H$. There is a twofold reason that suggests that $H$ should be considered as an important design parameter and that its value should be limited by some threshold.

- The overall energy for transferring all readings to the hub is directly proportional to $H$. Setting $H$ is equivalent to set an upper bound for all short-range transmissions within each neighborhood.
- Also the time for collecting readings of all nodes in $W$ is regulated by $H$, as it will be detailed in the following.

More formally, the connectivity matrix $C$ allows to define a distance $h(j,k)$ between the $k$-th and the $j$-th node, that is the minimum number of hops needed to reach the latter from the former. Said $V = \{0, \dots, N-1\}$ the set of nodes/vertices, each hub $u$ selects a subset $W \subseteq V$ such that $\sum_{w \in W} h(w,u) \leqslant H$.

This can be done by modifying the classical Dijkstra algorithm for the shortest path to a given root $u$ so that it adds a new vertex to the tree only if there are enough hops left to go from that vertex to the root. Algorithm 1 formalizes this in a procedure that accepts the adjacency matrix $C$, the hub $u$, and the budget $H$ to give the set $W$ of vertices/nodes that communicate their readings to $u$. Note that at line 8, possible ties are resolved by randomly selecting a node $w$ among those that exhibit the same minimum distance $d_{\min}$.

**Algorithm 1** Dijkstra's algorithm modified to spend not more than $H$ hops

**procedure** BUDGETDIJKSTRA($C, u, H$)
  $V \leftarrow \{0, \ldots, N-1\}$
  $W \leftarrow \varnothing$
  $d \leftarrow (\infty, \ldots, \infty)^\top$
  $d(u) \leftarrow 0$
  **while** $V \neq \varnothing$ **do**
    $d_{\min} \leftarrow \min_{v \in V} d(v)$
    $w \leftarrow \text{randompick}(\{v \in V | d(v) = d_{\min}\})$
    $V \leftarrow V \backslash \{w\}$
    **foreach** $z | C_{z,w} = 1$ **do**
      $d(z) \leftarrow \min\{d(z), d_{\min} + 1\}$
    **end foreach**
    **if** $H \geqslant d_{\min}$ **then**
      $W \leftarrow W \cup \{w\}$
      $H \leftarrow H - d_{\min}$
    **else**
      **exit while**
    **end if**
  **end while**
  **return** $W$
**end procedure**

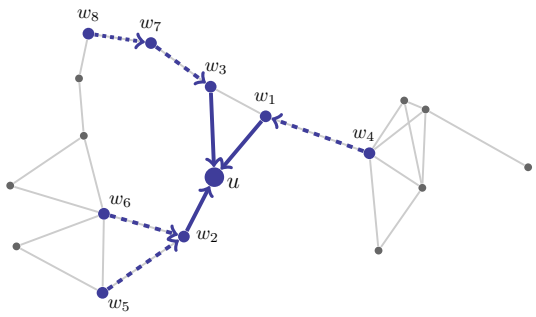



Fig. 2. The application of Algorithm 1 for $H = 16$.

Figure 2 exemplifies the procedure for $H = 16$. Black dots represent nodes, and edges are drawn between pairs of vertices $j, k$ such that $C_{j,k} = 1$, the largest blue disk represents the hub $u$ and is inserted in $W$ by the first execution of the **while** loop.

The nodes $w_1$, $w_2$, and $w_3$ can reach $u$ with just one hop along the solid blue edges and they are added to $W$ by the three subsequent iterations of the **while** loop. This uses 3 out of the $H = 16$ in the budget.

The nodes $w_4$, $w_5$, $w_6$, and $w_7$ can reach $u$ by transmitting their reading to one of the nodes just added to $W$ along the dashed blue edges, and thus use 2 hops each. They are added to $W$ and use a total of 8 hops out of the $H = 16 - 3 = 13$ in the budget.

This leaves us with only $H = 16 - 3 - 8 = 5$ hops in the budget. The nodes with the shortest path to $u$ that are not yet in $W$, need 3 hops to reach the hub. Hence, only one of them may be added to $W$. In this case $w_8$ is chosen, which can communicate its reading to $w_7$ along the dotted blue edge and thus to $u$ by means of a further hop through $w_3$.

In general, $H$ is less than what is needed to collect all the readings in the network and Algorithm 1 limits the nodes contributing to a measurement to those in a neighborhood of the hub.

To have an idea on how large such a neighborhood can be, let us refer to a typical node arrangement that we will use in all our simulations. The $N$ nodes are deployed at random in a square area with a normalized unit side. Then a transmission range $r$ is defined and any pair of nodes with an Euclidean distance not larger than $r$ gives rise to an edge in the connectivity graph. We will consider $N = 128$ and $r \in \{0.125, 0.15, 0.175\}$ to test scenarios with increasing connectivity.

It is intuitive that $H$ sets the number of nodes $n$ in a neighborhood. The experimentally obtained average value $\mathbf{E}[n]$ as a function of $H$ is plotted in Figure 3-(a) for three different values of $r$. Figures 3-(b), (c), and (d) show typical neighborhoods corresponding to $H = 32$ for graphs with increasing connectivity. Note how, as $r$ increases, the number of nodes that can be reached using the same hop budget increases.

The value of $H$ also sets the neighborhood *latency*, i.e. the time required by the hub to gather all $n$ associated readings, that is another crucial parameter when considering a multi-hop approach [25]. We estimate this time under the following assumption: $i$) the collection procedure must avoid data collisions; $ii$) at each time step a node can either transmit or receive a piece of data (e.g., a single antenna or RF amplifier is shared for transmission and reception). In this case latency is in the range $[n-1, 2n-3]$.

To see why, let us refer to Figure 2. Only nodes $w_1$, $w_2$ and $w_3$ can directly transmit to $u$. Readings at $w_4$, $w_5$, $w_6$, $w_7$ and $w_8$ need to be passed to a node one-step-closer to the hub, and can reach the hub only through $w_1$, $w_2$ or $w_3$. In other words, we can identify 3 subtrees, indicated with $S_1$, $S_2$ and $S_3$, whose cardinality is $s_1$, $s_2$ and $s_3$, respectively, and whose root node (i.e., the node that can communicate with $u$) is $w_1$, $w_2$ and $w_3$, respectively. A generic subtree $S_i$ can transfer all its $s_i$ readings to $u$ in $2s_i - 1$ time steps. The proof by construction is shown in Figure 4, and it is based on the observation that the bottleneck is given by the root node: it has to transmit $s_i$ pieces of data to $u$, and receive $s_i - 1$ pieces of data from other nodes, with an overall number $2s_i - 1$ of data exchanges. Hence, the total number of time steps for $u$ to collect all readings is $\sum_i (2s_i - 1)$ where $i$ spans the subtree covering on the same hub, and ranges from $n-1$ (when $s_i = 1$, $\forall i = 1, \ldots, n-1$) to $2n-3$ (when only the subtree $S_1$ exists, with $s_1 = n - 1$).

From Figure 3-(a) we get that as $H$ increases each hub has the possibility of accumulating a larger number of readings and thus is presumably able to compute measurements containing a larger amount of information on the original signal. Yet, as $H$ increases, both the energy and the time needed to collect readings increases so that too large values of $H$ are to be avoided.

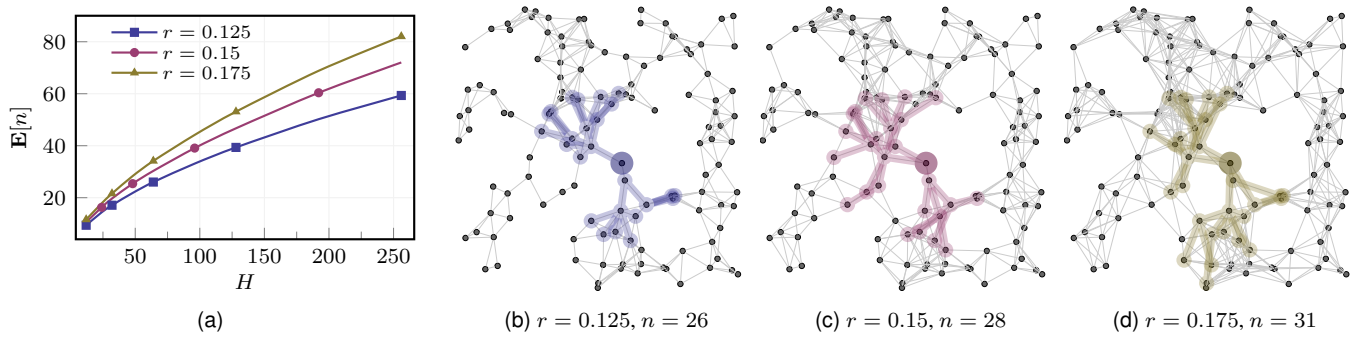



Fig. 3. The average number of nodes in a neighborhood as a function of the hop budget (a). The three graphs in (b), (c), and (d) show instances of connectivity and neighborhoods for three configurations corresponding to the points highlighted in (a).

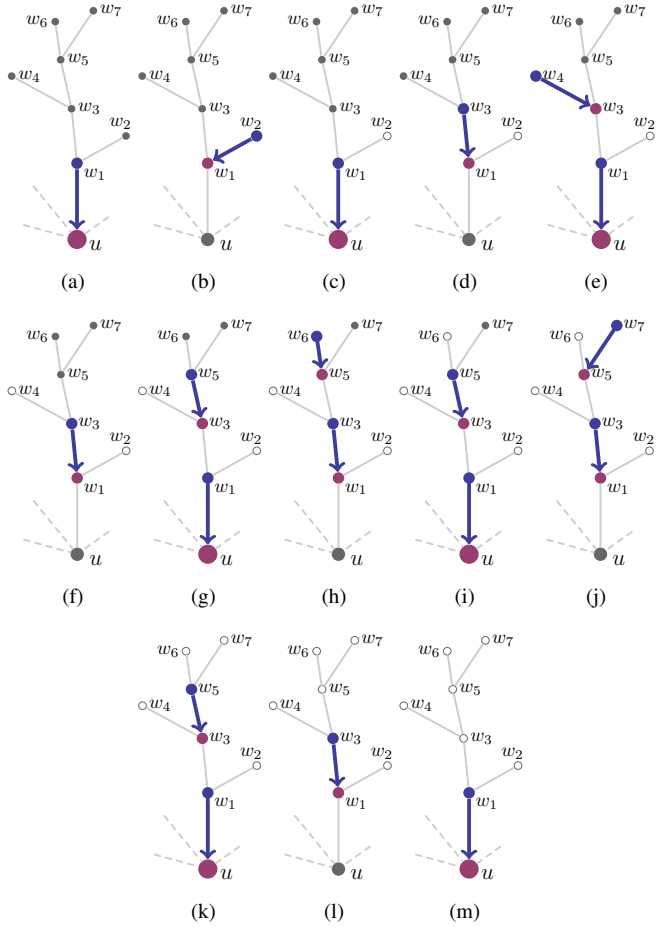



Fig. 4. A sketch of the method giving the number of time steps required to send all readings of a subtree with $s_i$ nodes to the hub. Bold blue edges stand for an active data transmission, nodes in bold are those involved in the transmission process; empty nodes have already transmitted their readings and are not involved anymore in the data propagation process. Starting from its own reading, the root node $w_1$ sends to the hub $u$ all readings of the subtree, alternating a step with a data transmission and a step with a data reception. Being $s_i$ the number of transmission steps, and $s_i - 1$ of reception steps, the total number of steps is $2s_i - 1$. In the figure, $s_i = 7$, so 13 steps are required, from (a) to (m).

## IV. Hub Spreading Algorithms

Among the $N$ sensing nodes, $M > 1$ of them are promoted to the role of hub.

A first, straightforward option for hub selection is the random one, that chooses any subset of $M$ out of $N$ nodes with equal probability to be used as hubs. This approach is common in the literature [9], [27] as it ensures a satisfactory hub rotation policy.

Yet, a pure random hub selection combined with the proposed modified Dijkstra's algorithm does not prevent overlapping neighborhoods, and does not even ensures full coverage of the sensing field.

Indeed, one of the main advantage of the CS is its capability to take advantage from substantially overlapping neighborhoods [30] and also to cope with uncovered sensors. However, since both the number of neighborhoods $M$ and the hop budget $H$ must be kept as low as possible to save communication resources, any avoidable overlap may actually be a waste if the same resources could have been spent to incorporate the reading of uncovered sensors into a measurement.

To treat the issue more formally, once that the neighborhoods $W_0, \ldots, W_{M-1}$ of the $M$ hubs are decided one may define

$$\gamma = \max_k \left| \{ W_j | k \in W_j, j = 0, \ldots, M-1 \} \right|,$$

i.e., the maximum number of times that any vertex appears in a neighborhood.

Minimizing $\gamma$ is a covering-type problem that may be extremely hard to solve. Yet, we may hope to reduce it by a simple heuristic. In particular we propose to select the hub one after the other randomly from a pool $U$ that is initialized with all the available nodes and from which we drop the nodes already participating in a neighborhood. If $U$ ends up containing no nodes then it is reset to include all the nodes with the exception of those already chosen as hubs. The procedure is formalized in Algorithm 2 to accept the connectivity matrix $C$, the number of hubs $M$ and the hop budget associated to each neighborhood $H$ to yield a set of hubs.

Though very simple, such an heuristic is able to reduce $\gamma$ as reported in Figure 5-(a) where we plot the average $\gamma$

**Algorithm 2** An heuristic procedure used to select the $M$ hubs

```
procedure HUBSELECT(C, H)
    V ← {0, ..., N − 1}
    U ← ∅
    do M times
        u ← randompick(V)
        U ← U ∪ {u}
        W ← BUDGETDIJKSTRA(C, u, H)
        V ← V\W
        if V = ∅ then
            V ← {0, ..., N − 1}\U
        end if
    end do
    return U
end procedure
```

over a large number of randomly drawn geometric connectivity graphs with $N = 128$ and $r = 0.15$ when $M = 32$ hubs are selected and as a function of the hop budget $H$ available to build each neighborhood. Figure 5-(b) and (c) show 6 of neighborhoods as they are placed by the two methods that in the following we will name as `rnd-`$\gamma$ and `low-`$\gamma$.

## V. Classical and Rakeness-based CS on collected readings

Let us now concentrate on the hub $u$ with its neighborhood $W^u$ that contains $n^u = |W^u|$ nodes. Once that readings are collected, the hub uses CS (see Section II-E) to processes them and produce a number $m^u$ of linear combinations that enter the measurements vector $y$

$$y_j^u = \sum_{k=0}^{n^u-1} A_{j,k}^u x_k \tag{1}$$

for $j = 0, \ldots, m^u - 1$ for some coefficients $A_{j,k}^u$. The coefficients are such that $A_{j,k}^u = 0$ if $k \notin W^u$ to model the fact the the readings combined in $y_j^u$ are only those collected by the hub $u$.

According to well-known guarantees on CS acquisition, the $n^u$ non-zeros in the $m^u$ rows of $A^u$ can be taken as instances of zero-mean and unit-variance Gaussian random variables, each row of $A^u$ being independent of the other rows. Classical CS assumes independence also between the different columns of $A^u$.

Yet, whenever the prior $\mathcal{X}$ on the second-order statistic of the signal is available, it has been proved that [7], [8], [36], [43], [44] this can be exploited by resorting to rakeness-based design. The main idea is to adapt the statistic of each of the rows $a$ of $A^u$ to increase the ability of the resulting projection of *raking* energy from the signal itself.

More formally, let us indicate with $\cdot^{|u}$ an indexed quantity when its indexes are limited to the integers contained in $W^u$, and with $a^{|u}$ a vector corresponding to a generic row of $A^u$. With this, the generic measurement is the scalar product between the generic row of $A^u$ and the signal $x$, a scalar product that can be limited to the non-zero components to yield $a^{|u\top} x^{|u}$. The average energy of such a generic measurement 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 the non-zeros in the row.

Hence, the energy that is raked from the signal can be increased by solving an optimization problem of the kind

$$\max_{\mathcal{A}^{|u}} \quad \operatorname{tr}\left(\mathcal{A}^{|u}\mathcal{X}^{|u}\right) \tag{2}$$
$$\text{s.t.} \quad \mathcal{A}^{|u} \succeq 0 \tag{3}$$
$$\text{s.t.} \quad \mathcal{A}^{|u} = \mathcal{A}^{|u\top} \tag{4}$$
$$\text{s.t.} \quad \operatorname{tr}\left(\mathcal{A}^{|u}\right) = n^u \tag{5}$$
$$\text{s.t.} \quad \operatorname{tr}\left(\mathcal{A}^{|u 2}\right) \leqslant \zeta \left(n^u\right)^2 \tag{6}$$

where (3) and (4) ensure that $A^{|u}$ is a symmetric and positive-semidefinite matrix, i.e., a proper correlation matrix, and (5) sets the energy of the row proportional to the number of coefficients $n^u$. As far as (6) is concerned, note that, due to the random nature of the signal, observing only its maximum-energy component (i.e., its principal component) is not enough to reconstruct it, and energy maximization should be tempered by the need to span the whole signal space. This is obtained by suitably bounding the sum of the squares of the eigenvalues of $\mathcal{A}^{|u}$ to prevent them to concentrate only on the principal components [7], [8]. In its simplest form, (2)-(6) has the analytical solution

$$\mathcal{A}^{|u} = \frac{1}{2}\left(\frac{n^u \mathcal{X}^{|u}}{\operatorname{tr}\left(\mathcal{X}^{|u}\right)} + I_{n^u}\right) \tag{7}$$

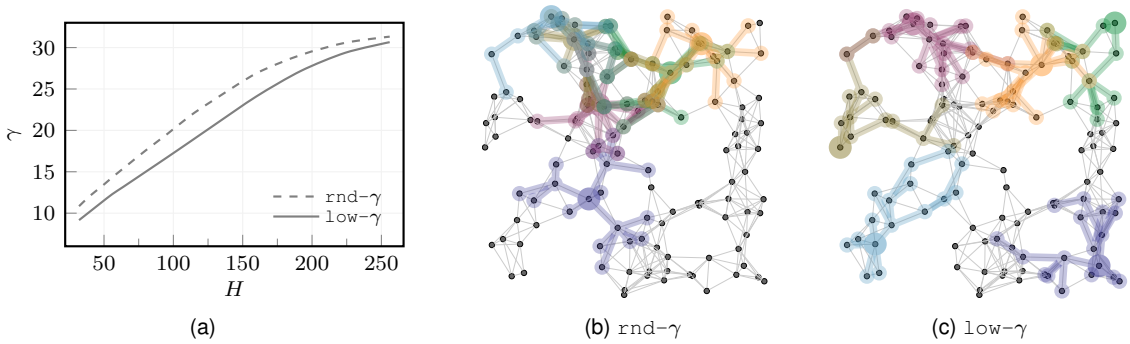
where $I_{n^u}$ is the $n^u \times n^u$ matrix and $\zeta$ was set as suggested in [8].

Hence, as a second option, instead of drawing the coefficients as random independent normals, for every measurement $y_j^u$ depending on the vertices in $W^u$ we generate random jointly-Gaussian coefficients with correlation (7).

## VI. Setting for performance assessment

The effectiveness of the techniques described above has been tested by Montecarlo simulations in a number of configurations. In all cases $N = 128$.

As anticipated, the connectivity graph is a geometric graph in which nodes correspond to points randomly chosen within a unit-length square according to a uniform distribution. A connection between two nodes is established whenever the two corresponding points have a distance not larger than $r$, with $r \in \{0.125, 0.15, 0.175\}$. This allows to test different levels of connectivity since the larger the $r$, the higher the number of potential connections between nodes that may be exploited. The adjacency matrix of the connectivity graph is $C$ and is




(a) (b) rnd-γ (c) low-γ

Fig. 5. The overlap $\gamma$ between $M = 32$ neighborhoods against the hop budget $H$ when the hubs are placed randomly and with the simple heuristic of 2 (a). Examples of hubs and neighborhoods obtained by the two medhods (b) and (c).

TABLE II
OPTIONS AND PARAMETERS CONSIDERED IN THE SIMULATIONS.

| parameter | meaning | values |
|---|---|---|
| $N$ | dimensionality of the signal | 128 |
| $\kappa$ | number of non-zero components of $x$ along $D$ | $\{6, 12\}$ |
| $D$ | sparsity basis of $x$ | $\mathcal{F}[C]$ Fourier of connectivity graph; $\mathcal{F}[\tilde{C}]$ Fourier of perturbed connectivity graph; $\mathcal{N}$ Orthonormalized random normal basis |
| ISNR | intrinsic signal-to-noise ratio of $x$ | 60 dB |
| $r$ | maximum connection distance between nodes uniformly distributed in $[0, 1]^2$ | $\{0.125, 0.15, 0.175\}$ |
| $M$ | number of hubs | $\{4, 5, 6, 7, 8, 10, 12, 14, 16, 20, 24, 28, 32, 40, 48, 56, 64\}$ |
| $H$ | maximum number of node-to-node communications to collect readings at each hub | $\{32, 40, 48, 56, 64, 80, 96, 112, 128, 160, 192, 224, 256\}$ |
| | hub selection | low-$\gamma$ selected with the heuristic in Section III; rnd-$\gamma$ selected randomly |
| | coefficients for the computation of the measurements | rnd-CS classical random coefficients; rak-CS rakeness-based coefficients |
| $\epsilon$ | ratio between energy/bit in local communication and energy/bit in long-range communication | $[5 \times 10^{-6}, 5 \times 10^{-2}]$ |

such that $C_{j,k} = 1$ whenever the $k$-th node can communicate with the $j$-th node and $C_{j,k} = 0$ otherwise.

The connectivity graph regulates the collection of sensor readings at the $M$ hubs with $M \in \{4, 5, 6, 7, 8, 10, 12, 14, 16, 20, 24, 28, 32, 40, 48, 56, 64\}$. Hubs may be selected either randomly among the available nodes (rnd-$\gamma$) or according to the heuristic of Section III that aims at reducing the overlap between neighborhood (low-$\gamma$).

The local communication budget for each collection is $H \in \{32, 40, 48, 56, 64, 80, 96, 112, 128, 160, 192, 224, 256\}$. Clearly, as $H$ increases the number of readings collected by each hub also increases so that measurements combines a larger number of entries of the vectors $x$, intuitively increasing their information content.

The $M$ hubs are used cyclically to provide the $m$ measurements, i.e., the first $M$ measurement come from different hubs that are reused to provide the subsequent $M$, and so on up to the total $m$ measurements. If $m$ is not a multiple of $M$ not all the hubs provide the same number of measurements.

Measurements are computed according to (1) by usign random normal coefficients. We evaluate both the classical CS case in which coefficients are independent (this option is labeled rnd–CS) and the rakeness-based option in which the array of coefficients used for each measurement has an auto-correlation matrix given by (7) (this option is labeled rak–CS) .

The signal $x$ is $\kappa$ sparse with $\kappa \in \{6, 12\}$ with respect to a sparsity basis $D$ for which we consider three options.

As a first option, $D$ is the eigenvector basis of the adjacency matrix of the connectivity graph $C$. As discussed above this amounts to say that $x$ is sparse with respect to the Fourier basis of the connectivity graph. We label this option with $\mathcal{F}[C]$.

To test a configuration in which this situation is only approximately verified, we also consider a second option that takes $D$ as the eigenvector basis of a perturbation of $C$ by means of a Watts-Strogatz-like adjustment that randomly provides non-local links [45]. More formally, starting from $C$ each of the initial edges has a probability $p$ of being deleted and substituted by an edge randomly chosen between those not already present in the graph. We consider $p = 0.2$ and label this option with as $\mathcal{F}[\tilde{C}]$.

As a third option we generate a random basis $D$ by orthonormalizing an $N \times N$ random matrix with independent normal entries. We label this option as $\mathcal{N}$.

In all cases the readings of the sensor that are processed correspond to the true signal pertubed by a white Gaussian noise with an Intrinsic Signal-to-Noise Ratio (ISRN) of 60 dB to emulate noise and system non-idealities.

The options described above give raise to a total of 15912 configurations. Each of such configurations has a cost associated to the energy needed to send all the information to the central collector. Such a cost is made of two contributions. The contribution due to the local communication is proportional to the number of hubs and to the hop budged available to collect the readings at each of them. The contribution due to long-range communication is proportional to the number of measurements. Overall the energy needed to transmit all the information is

$$E_{\text{tot}} = mE_{\text{long range}} + \min\{m, M\}HE_{\text{short range}}$$

where $E_{\text{long range}}$ is the energy needed to transmit a sensor reading from a node to the central collector and $E_{\text{short range}}$ is the energy needed to transmit a sensor reading between two nodes. As anticipated, we do not take into account the energy spent in computing the measurements since this is usually negligible with respect to transmission [18], [19], [42].

A straightforward acquisition scheme would need $E_0 = NE_{\text{long range}}$ that may be used to normalize the above cost into

$$e = \frac{E_{\text{tot}}}{E_0} = \frac{m}{N} + \frac{\min\{m, M\}H}{N}\epsilon \tag{8}$$

where $\epsilon$ is the ratio between the energy/bit entailed in a short-range communication and the energy/bit entailed in a long-range communication. Clearly, values of $e$ lower than 1 indicate energy saving with respect to the straightforward approach.

In return for such a cost, the information gathered at the collector allows to recover the original signal $x$ with a certain fidelity that varies from instance to instance.

We run extensive Montecarlo simulations to assess the statistics of such a fidelity, measured as the Reconstruction Signal-to-Noise Ratio, i.e., considering the original signal $x$ and the recovered one $\hat{x}$ to compute

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

where signals recovery was performed by [46].

We summarize the performance of a method by analyzing its Probability of Correct Reconstruction (PCR)

$$\text{PCR} = \Pr\{\text{RSNR} \geqslant \text{RSNR}_{\min}\}$$

that is estimated by simple frequency counts against a minimum quality level fixed at $\text{RSNR}_{\min} = \text{ISNR} - 3\,\text{dB} = 57\,\text{dB}$. PCR acts as a guarantee in the spirit of outage probabilities of communication systems since, for example, PCR$= 0.95$ implies that the required RSNR is obtained at least 95% of the times.

## VII. Numerical evidence and discussion

Equation (8) shows that the energy cost is made of a compression term $m/N$ that depends on the ability of CS to express an $N$-dimensional signal with only $m \ll N$ scalars and of an overhead term $\epsilon \min\{m, M\}H/N$ that accounts for the cost of the local transmissions needed to apply CS itself.

We may show the effectiveness of CS in lowering the first term by plotting the achievable PCR against the number $m$ of measurements. This is done in Figure 6 for few significant configurations that allow to highlight the main trends. Such trends are maintained in all the cases we tested.

In each plot of Figure 6 dashed tracks correspond to `rnd`−CS while solid tracks correspond to `rak`−CS. As far as colors are concerned, blue tracks correspond to `rnd`−$\gamma$ while red tracks correspond to `low`−$\gamma$.

Since $m$ on the horizontal axis is the number of measurements, the more to the left a curve, the higher the performance. This allows to appreciate that in all cases maximum performance is given by `low`−$\gamma$ and `rak`−CS systems with non-negligible advantages over more classical approaches.

Considering Figure 6-(a) as a reference case, the plots in Figure 6-(c) to (f) show how the performance profile change when one of the options (highlighted in red) changes.

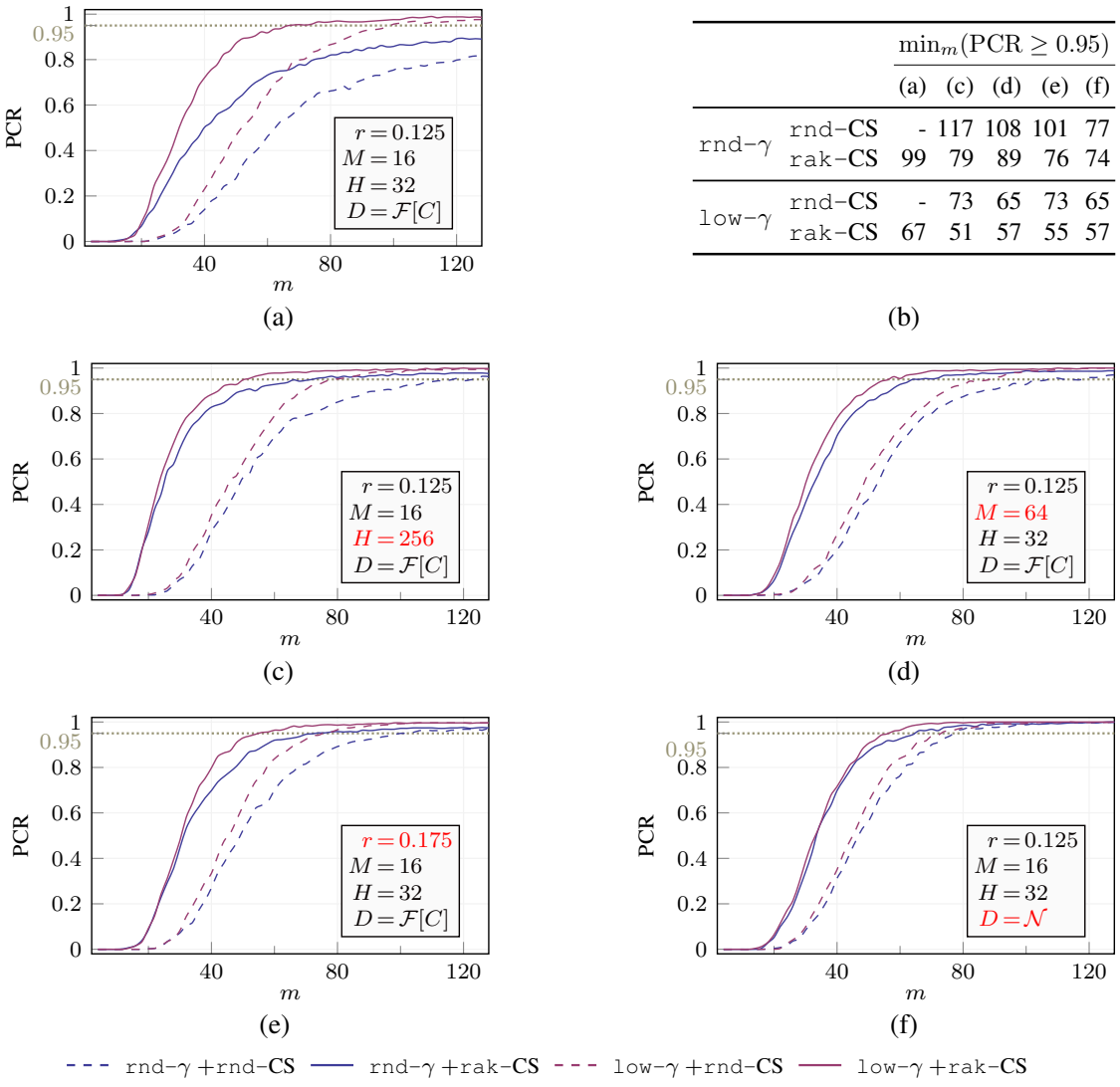
By comparing Figure 6-(a) with Figure 6-(c) and Figure 6-(d), one observes that the effectiveness of `low`−$\gamma$ is maximum when $M$ and $H$ are small. In that case, neighborhoods are small in size and few in number, and their correct administration brings noteworthy advantages. As either $H$ or $M$ increases, most of the nodes are able to contribute their readings to at least one hub almost independently of the hub selection strategy.

On the contrary, the effectiveness of `rak`−CS increases with $M$ and $H$, since this increases the diversity in the radings that are combined in the measurements, a diversity that is exploited by rakeness-based design.

Beyond qualitative considerations, better performance translates in the need of less resources to obtain the same quality in signal recovery. To quantify this, we intersect the plots with the PCR$= 0.95$ line to find the minimum number of measurements that is needed to ensure that an RSNR not smaller than 57 dB is obtained at least 95% of the times.

These values are reported in Figure 6-(b) and are less than $N = 128$, thus indicating that the first term in (8) would be less than 1. Yet, by adopting `low`−$\gamma$ and `rak`−CS one my hope to make $m/N$ to be not more than $67/128 < 0.53$ and thus, once payed for the overhead, to reach substantial savings in $e$.

To estimate the achievable savings we fix the required performance level to PCR$= 0.95$ and scan the design space $(H, M)$ for the point that guarantees such a performance while minimizing (8). This is done for $\epsilon \in [5 \times 10^{-5}, 5 \times 10^{-2}]$ and for signals that are characterized by different sparsity levels $\kappa \in \{6, 12\}$ with respect to different bases $D \in \{\mathcal{F}[C], \mathcal{F}[\tilde{C}], \mathcal{N}\}$. The result is reported in Figure 7, in which $e$ is plotted against $\epsilon$. All plots are clipped to $e \leqslant 1$ as larger values of $e$ imply no saving. The two values $\epsilon = 5 \times 10^{-4}$ and $\epsilon = 5 \times 10^{-3}$ corresponding to currently viable technologies are marked as vertical dotted lines.




| | | $\min_m(\text{PCR} \geq 0.95)$ | | | | |
|---|---|---|---|---|---|---|
| | | (a) | (c) | (d) | (e) | (f) |
| rnd-γ | rnd-CS | - | 117 | 108 | 101 | 77 |
| | rak-CS | 99 | 79 | 89 | 76 | 74 |
| low-γ | rnd-CS | - | 73 | 65 | 73 | 65 |
| | rak-CS | 67 | 51 | 57 | 55 | 57 |

Fig. 6. PCR against $m$ for different example configurations (a)(c)(d)(e)(f). The table reports the minimum number of measurements needed to ensure that PCR$\geqslant 0.95$ (b).

By comparing the plots in that Figure, one notes that $\kappa$, i.e., the intrinsic complexity of the signal, affects energy needs since $e$ plots are higher in the second row ($\kappa = 12$) of Figure 7 with respect to first row ($\kappa = 6$).

Rakeness-based CS is always a key player since the gap between solid (rak−CS) and dashed (rnd−CS) tracks is always substantial. In more complex cases ($\kappa = 12$) the adoption of rakeness-based CS is the key to make the approach profitable since standard CS would not yield significant energy savings.

Moreover, when the relative cost of local communication ($\epsilon$) increases, the effectiveness of the methods we propose decreases since trends tend to rise towards and beyond $e = 1$. Yet, as this happens, the need for a careful administration of local transmission makes hub selection more effective since the gap between red (low−$\gamma$) and blue (rnd−$\gamma$) curves widens.

Hub spreading by means of low−$\gamma$ substantially affects savings when time constraint come into play. From Section III we know that the time needed by the hubs to collect readings is proportional to the number of nodes in their neighborhoods and Figure 3-(a) shows how this depends on the hop budget $H$. As an example, we may think that time constraints force us to set $H \leqslant 32$ and look for minimum-energy design choices within such a limitation. The result is reported in Figure 8 that deals only with rak−CS cases. Comparing Figure 7 with Figure 8, one immediately gets that the $H \leqslant 32$ constraint results in a lower energy saving, though the gap between blue and red tracks shows that trying to lower the overlap between neighborhoods has non-negligible beneficial effefts.

Overall, both in the $H \leqslant 32$ and in the time-unconstrained cases, using values of $\epsilon$ corresponding to current technology, the energy needed to send all the information to the central collector can be cut significantly by adopting the methods we propose, with peak reductions of almost 70% in the most favorable cases.

This is obtained by observing that the triples $(m, M, H)$ matching the performance constraint PCR$=$ 0.95 for

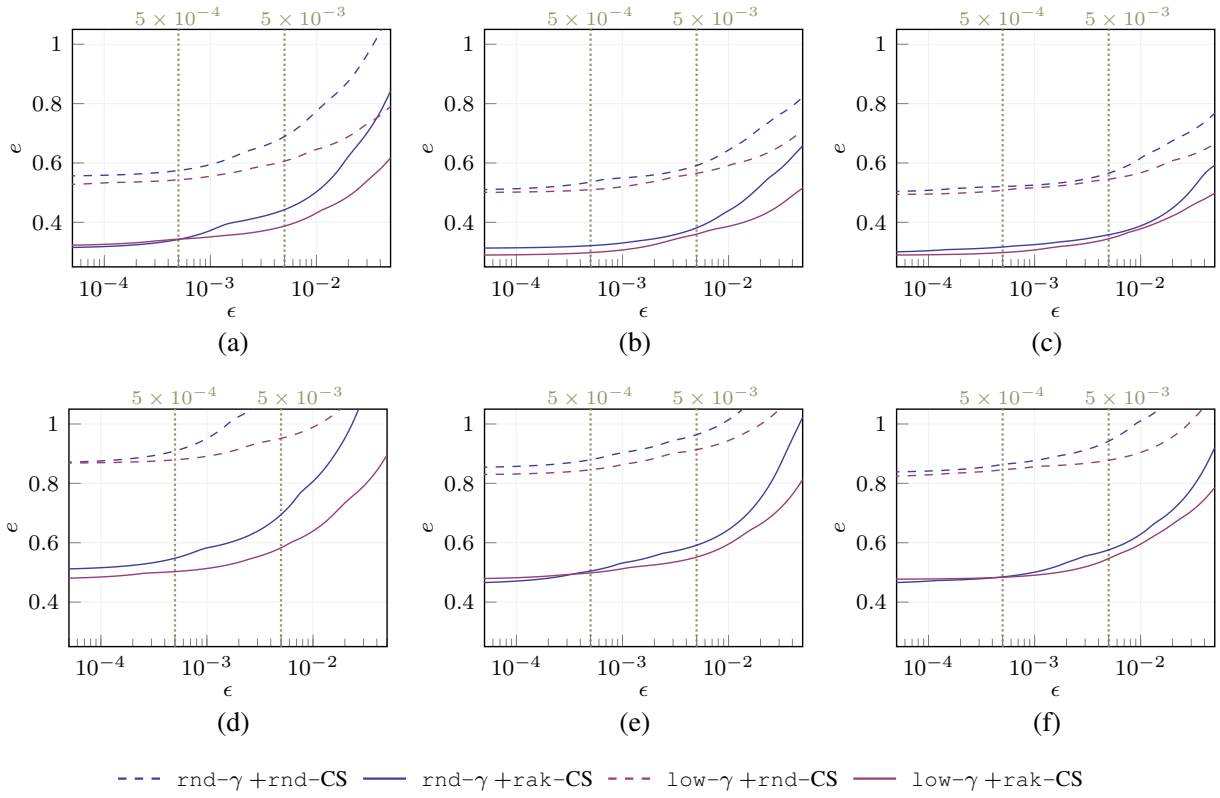



Fig. 7. Minimum energy needed to transmit to the collector all the information needed for signal recovery with PCR$\geqslant 0.95$. In (a), (b), and (c) $\kappa = 6$ while in (d), (e), and (f) $\kappa = 12$. In (a) and (d), $D = \mathcal{F}[C]$. In (b) and (e), $D = \mathcal{F}[\tilde{C}]$. In (c) and (f), $D = \mathcal{N}$.

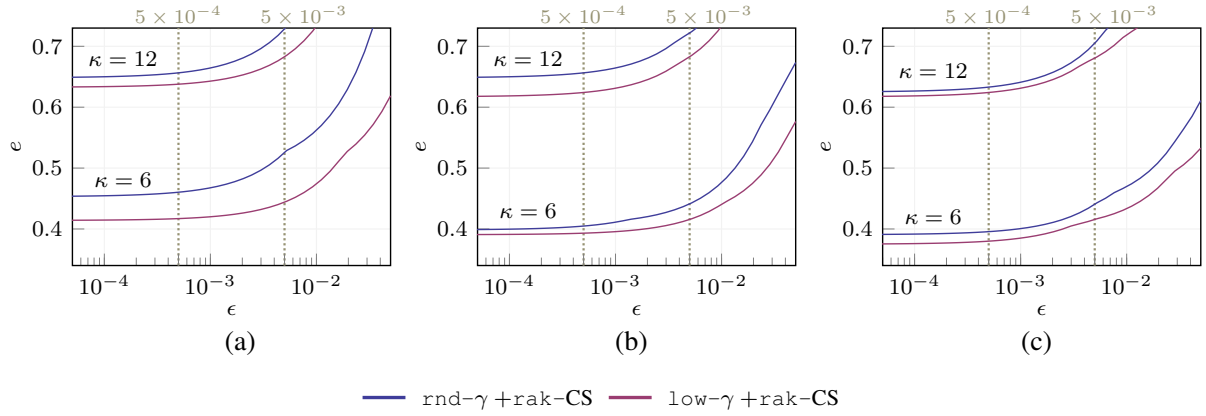



Fig. 8. Minimum energy needed to transmit to the collector all the information needed for signal recovery with PCR$\geqslant 0.95$ where $H = 32$ is set by a time constraint in the acquisition. Only rak–CS cases are reported. In (a) $D = \mathcal{F}[C]$, in (b) $D = \mathcal{F}[\tilde{C}]$, and in (c) $D = \mathcal{N}$.

$\mathrm{RSNR}_{\min} = 57\,\mathrm{dB}$ implies a different ratio between the energy spent in long-range communication (proportional to $m$) and the energy spent in local communication (proportional to $\min\{m, M\}H$). The ratio yielding the minimum total normalized energy (8) depends on the relative cost of the local versus long-range communication $\epsilon$. This is shown in Figure 9 that considers the case of Figures 7-(a) and (d). Even though with different trends, independently of the method one uses for hub selection and for measurement computation, as the cost of a single local transmission increases the number of local transmissions involved in the minimum-energy configuration decreases.

Finally, we propose a comparison between results obtained by the low–$\gamma$ spreading approach and that obtained by a state-of-the-art approach based on clustering and CS compression. In details, we consider the approach discussed in [29] that works in a scenario very similar to one we adopt. In [29] each sensor nodes is assigned to a single cluster and communicate with the corresponding hub using a multi-hop protocol, while data is locally compressed at each hub by rnd–CS. Hubs are identified according to the following iterative algorithm.

1) $M$ nodes are randomly elected as hubs.

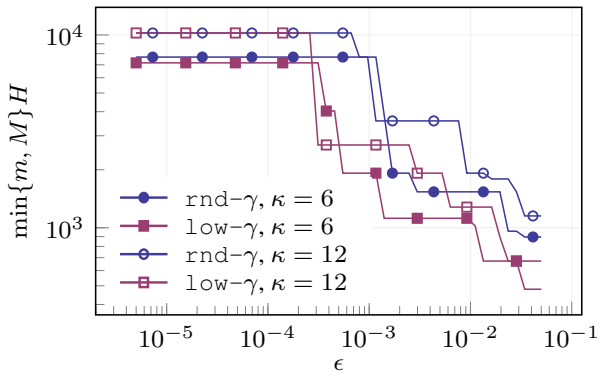



Fig. 9. The number of local transmissions involved in the minimum-energy configurations for the case in Figure 7-(a) and (d) as a function of $\epsilon$.

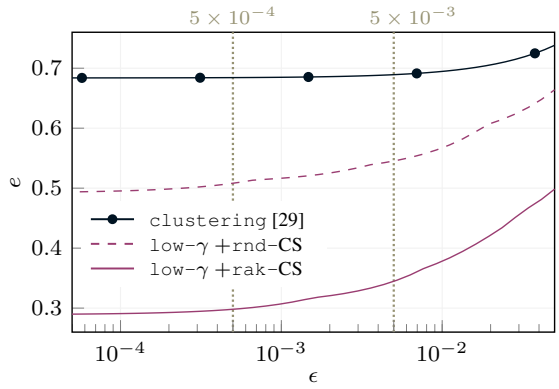



Fig. 10. Minimum energy needed to transmit to the collector all the information needed for signal recovery with PCR $\geqslant 0.95$ where $\kappa = 6$ and $D = \mathcal{N}$ for the clustering based approach in [29] and for the proposed low−$\gamma$ +rnd−CS and low−$\gamma$ +rak−CS.

2) Each node is assigned to the cluster identified by the hub that is the closest one in terms of euclidean distance.
3) For each cluster, select as a new hub the node for which the sum of the distances with all other nodes is minimized. In other words, new selected hubs are the cluster centroids.
4) Repeat steps 2) and 3) until there is no change in the hubs election.

According to [29], this algorithm always converges after very few iterations.

The degree of freedom that is investigated is the number of cluster $M$. Even though authors set $M$ in order to minimize the total transmission cost, in their model the data collector is located close to the sensing area and long-range communication costs are not considered in the proposed optimization.

To allow a comparison with the method proposed here, we adopt the clustering algorithm of [29] considering, for each value of $\epsilon$, all values of $M$ in Table II and, for each of them, the minimum value of $m$ that guarantees PCR $\geqslant 0.95$. For each value of $\epsilon$ the couple $(M, m)$ ensuring the lowest normalized energy $e$ is selected. Obtained values of $e$ against $\epsilon$ are shown in Figure 10 compared with those obtained using the proposed hub spreading technique both without sensing optimization (low−$\gamma$ +rnd−CS) and when the rakeness-based CS is adopted (low−$\gamma$ +rak−CS). Here, $\kappa = 6$ and $D = \mathcal{N}$ are considered as for Figure 7-(c).

One can observe that, for all considered $\epsilon$ values, the hub spreading technique we propose and the adoption of the rakeness-based CS guarantee a non negligible reduction in terms of energy needed to transmit to the collector all the information needed to correctly recover sensor nodes readings.

## VIII. Conclusion

CS can be used to locally aggregate the readings of a number of sensors into a smaller number of quantities to be transmitted to a local collector that is able to reconstruct the individual acquisition providing it can exploit a sparsity prior on their ensemble.

Exploiting this and carefully administering the choice of the local hubs, the breadth of the neighborhood from which they collect readings, as well as the coefficients with which those readings a linearly aggregated, one may significantly reduce the energy needed to sample the field.

Simulations indicate that savings may be over 50% for values of the parameters modeling nowadays local and long-range transmission technologies.

## References

[1] 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.
[2] 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.
[3] A. Attwood, D. J. Lamb, and O. Abuelmaatti, "Position-relative identities in the internet of things: An evolutionary ght approach," *IEEE Internet of Things Journal*, vol. 1, no. 5, pp. 497–507, Oct 2014.
[4] S. Li, L. D. Xu, and X. Wang, "A continuous biomedical signal acquisition system based on compressed sensing in body sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 9, no. 3, pp. 1764–1771, Aug 2013.
[5] 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.
[6] 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.
[7] 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.
[8] 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.
[9] 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.
[10] "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.
[11] "CC2640R2F SimpleLink® Bluetooth™ low energy Wireless MCU," Texas Instruments Datasheet, Jan. 2017, http://www.ti.com/lit/ds/symlink/cc2640r2f.pdf.
[12] "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.
[13] "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.

[14] "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.

[15] "LTE MTC: Optimizing LTE Advanced for Machine-Type Communications," Qualcomm Whitepaper, Jan. 2014.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] F. Fazel, M. Fazel, and M. Stojanovic, "Random access compressed sensing for energy-efficient underwater sensor networks," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1660–1670, September 2011.

[21] J. Cheng, Q. Ye, H. Jiang, D. Wang, and C. Wang, "Stcdg: An efficient data gathering algorithm based on matrix completion for wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 850–861, February 2013.

[22] C. Luo, F. Wu, J. Sun, and C. W. Chen, "Efficient measurement generation and pervasive sparsity for compressive data gathering," *IEEE Transactions on Wireless Communications*, vol. 9, no. 12, pp. 3728–3738, December 2010.

[23] C. Caione, D. Brunelli, and L. Benini, "Distributed compressive sampling for lifetime optimization in dense wireless sensor networks," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 1, pp. 30–40, Feb 2012.

[24] C. Karakus, A. C. Gurbuz, and B. Tavli, "Analysis of energy efficiency of compressive sensing in wireless sensor networks," *IEEE Sensors Journal*, vol. 13, no. 5, pp. 1999–2008, May 2013.

[25] D. Ebrahimi and C. Assi, "On the interaction between scheduling and compressive data gathering in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2845–2858, April 2016.

[26] X. Yu and S. J. Baek, "Joint routing and scheduling for data collection with compressive sensing to achieve order-optimal latency," *International Journal of Distributed Sensor Networks*, vol. 13, no. 10, p. 1550147717737968, 2017.

[27] 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.

[28] 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.

[29] 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.

[30] J. Haboba, M. Mangia, F. Pareschi, R. Rovatti, and G. Setti, "A pragmatic look at some compressive sensing architectures with saturation and quantization," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, pp. 443–459, Sept 2012.

[31] 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.

[32] 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.

[33] S. Chen, R. Varma, A. Sandryhaila, and J. Kovacevic, "Discrete signal processing on graphs: Sampling theory," *IEEE Transactions on Signal Processing*, vol. 63, no. 24, pp. 6510–6523, Dec 2015.

[34] 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.

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] 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.

[41] 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.

[42] D. Bortolotti, M. Mangia, A. Bartolini, R. Rovatti, G. Setti, and L. Benini, "Rakeness-based compressed sensing on ultra-low power multi-core biomedicai processors," in *Proceedings of the 2014 Conference on Design and Architectures for Signal and Image Processing*, Oct. 2014, pp. 1–8.

[43] 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.

[44] 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.

[45] M. Newman, *Networks: An Introduction*. London, U.K.: Oxford Univ, 2010.

[46] 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.

**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.

**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.