This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2009 proceedings.

# Relay selection schemes for uniformly distributed wireless sensor networks

Keyvan Zarifi*†, Mohammed Abuthinien†, Ali Ghrayeb†, and Sofiene Affes*
*INRS-EMT, 800, de la Gauchetière west, Montreal, QC, H5A 1K6, Canada
{zarifi,affes}@emt.inrs.ca
†Concordia University, 1455 de Maisonneuve west, Montreal, QC, H3G 1M8, Canada
{m_abuthi,aghrayeb}@ece.concordia.ca

***Abstract*—The relay selection problem in a large wireless sensor network (WSN) with uniformly distributed identical nodes is investigated for a two-phase cooperative protocol where the signal transmitted from a single source is overheard by the network and is then relayed by multiple selected nodes subject to a total average transmission power.**

**First, a relay selection technique is considered that maximizes the average signal-to-noise ratio (SNR) at the access point (AC) while is applicable to a distributed WSN where only a limited information is available to the nodes. As this technique is shown to be energy inefficient, two other alternative relay selection techniques are also presented that substantially reduce the energy consumption of the network. One of these alternative techniques takes advantage of the static nature of the network topology and, as verified by simulations, achieves near-optimal SNR performance, while the other one randomly selects the relays from a neighborhood around the source and further guarantees a fair power consumption among the nodes at the cost of possible SNR performance drop at AC.**

**Accounting for the randomness of the nodes locations as well as the inter-terminal fading channel coefficients, the SNR performances of the energy efficient relay selection schemes are analyzed and a condition is derived under which the average SNR at AC is independent from the technique used to select the relaying nodes. This condition clarifies when the energy efficient and fair random relay selection scheme may be used without compromising the quality of the received signal at AC.**

## I. Introduction

Recent advances in mass production of low-cost low-power micro sensors have inspired growing interests in the applications of wireless sensor networks (WSNs) in various fields of research such as industrial automation, remote monitoring of biological signals, military surveillance, as well as space communications [1]. Sensor networks suitable for such applications are typically comprised of a group of light-weight battery-powered nodes that are randomly spread over a large region and should be able to communicate with a distant access point (AC) while having a very low power dissipation to avoid battery depletion over a long period of time.

Similar to other wireless communication systems, a major challenge in WSNs is the fluctuations in the received signal power at AC due to the inherent wireless channel fading effect. An efficient technique to counter this effect is providing spatial diversity by means of transmitting independently faded replicas of the same signal from multiple locations. As the nodes in WSNs are typically equipped with a single transmit antenna, spatial diversity can be achieved by means of cooperation among nodes [2]. Nodes cooperation can additionally serve as a means to increase the durability of WSNs: Each node can reduce its transmission power in view of the fact that a copy of its signal is also retransmitted from other relaying nodes. This results in a more even power dissipation among nodes, and, consequently, increasing the network lifetime [3].

As a major factor influencing the quality of reception at AC, the choice of relaying nodes from the set of potential candidates is a growing subject of research [4], [5]. In this paper, we investigate the relay selection problem in a large WSN with uniformly distributed identical nodes [1], [6]. We consider a two-phase transmission protocol where the signal transmitted from a single source is overheard by other nodes in the first phase and is then relayed by multiple selected nodes through orthogonal channels in the second phase.

Aiming to maximize the average signal-to-noise ratio (SNR) at AC, we show that if the forward channels from the nodes to AC are unknown, the optimal relay set is the set of nodes with the maximum received power from the source. Building on the protocol introduced in [4], we then develop a simple relay selection technique that selects the optimal relay set while complying with the distributed nature of WSN and the limited information available to the nodes. Although the above relay selection scheme delivers the highest possible SNR at AC, we explain why this technique can be excessively energy wasteful as the number of nodes grow large. This motivates us to take advantage of the static nature of the network topology and develop a substantially more energy efficient relay selection alternative that selects the relays from the closest nodes to the source. Simulation results verify that the latter technique provides AC a SNR performance close to that of the optimal relay selection counterpart. However, in spite of its clear advantages, the topology-based relay selection approach may result in an over-exploitation of the nodes that are close to the source. This drives us to develop a third relay selection technique that randomly chooses the relays from a neighborhood around the source and is both energy efficient and fair towards all nodes. The price that may have to be paid, is a noticeable SNR drop at AC due to the fact the relays are selected randomly.

As all of the above schemes have their own advantages, a sensible choice of the relay selection scheme requires a clear knowledge of how the performance of either of the above schemes depends on the system parameters such as the node density, the source transmission power, and the path-loss exponent. Leaving the analytical results regarding the optimal relay selection scheme to our later reports, we contribute to this knowledge by obtaining the average SNRs at every selected relay as well as AC for both the topology-based and the random relay selection schemes while taking into account the facts that all node-source distances and all inter-terminal channel links are random quantities. Based on the so-obtained results, we derive a condition under which the average SNR at AC becomes independent from the scheme used to select the relays. This condition is of particular practical importance, as it clarifies when the power-efficient and fair random relay selection scheme can be used without deteriorating the signal reception quality at AC.

This paper is organized as follows. Section II provides the system description and presents the signal model and the relay selection schemes. The SNR performances of both the topology-based and the random relay selection schemes are analyzed in Section III. Simulation results are given in Section IV and concluding remarks are presented in Section V.

## II. System Description

### A. Transmission protocol

Consider a large WSN wherein $N$ identical sensor nodes are uniformly distributed with density $\rho$ on a plane [1], [6] and assume that a node $s$ aims to transmit the signal-of-interest to a distant AC located outside the WSN through the following two-phase amplify-and-forward (AF) cooperative transmission protocol [2], as shown in Fig. 1: In the first phase, $s$ transmits its signal while all other nodes within its transmission range receive a noisy version of the transmitted signal. In the second phase, $K$ receiver nodes switch to the transmission mode, amplify the signal they received in the first phase and, subsequently, relay it to AC through orthogonal channels and subject to a total average transmission power constraint.

The policy under which the nodes choose to act as relays and the technique to enforce orthogonality of the relays transmitted signals are the subjects of our later discussion in Subsections II-D and II-E, respectively. Note that, the rationale behind the above multiple-relay cooperative scheme is its increased diversity and network durability in comparison to the simpler single-relay cooperative communication counterpart: The $K$ relays in the above scheme provide $K$ independently-faded communication links, and, hence, further decrease the sensitivity of the quality of signal reception at AC to a severe channel fading at each individual link. Moreover, as the transmission in the second phase is subject to a total average transmission power constraint, an increased number of relays directly results into a decreased power dissipation from each relay, and, therefore, more network lifetime.

### B. System assumptions

We apply the above cooperative communication protocol in a practical unsupervised scenario in which there is no channel state information at the source and no information exchange or synchronization among the relays while there is some low rate feedback from AC to the nodes. It is assumed that relay $k$ is only aware of $D_{s,k}$, its distance to $s$, along with $h_{s,k}$, the fading coefficient of the channel between $s$ and $k$, and

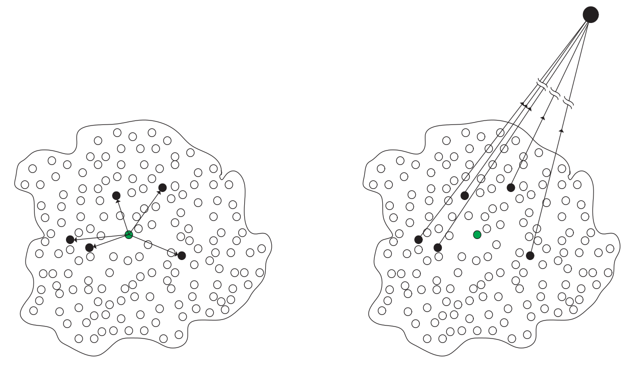


Fig. 1. The two-phases cooperation system description.

periodically sends this information to AC. In addition to the information sent from the relays, AC is also aware of $D$, its distance from the network, as well as $h_{k,d},\ k = 1, \ldots, K$, the fading coefficients of the channels from the relays to AC. Note that AC is located far enough such that its distance from every node on the network can be considered to be equal to $D$.

It is assumed that all inter-terminal channel links experience zero-mean unit-variance Rayleigh fading and noise at all terminals is zero-mean white Gaussian with variance $\sigma^2$.

### C. Signal Model

The received signal during the first phase at the relaying nodes is given by

$$y_k = \sqrt{p_s} h_{s,k} D_{s,k}^{-\nu/2} x + n_k \qquad k = 1, \ldots, K \tag{1}$$

where $p_s$ is the power transmitted from $s$, $\nu$ is the path-loss exponent with a value in the interval [2,6), $x$ is the zero-mean unit-variance transmitted data from $s$, and $n_k$ is noise. We use a common assumption that there is a relay-free zone of unit distance around the source [7], [8] to avoid the problem of received power divergence at small distances. Moreover, we should stress that the signal transmitted in the first phase is solely intended to reach the potential relays in a neighborhood around the source. Therefore, $p_s$ is selected small enough such that $s$ does not induce a considerable interfering effect on further parts of the network.

From (1) it follows that SNR at node $k$ is given by

$$\gamma_k = \frac{p_s}{\sigma^2} |h_{s,k}|^2 D_{s,k}^{-\nu} \tag{2}$$

and $\bar{\gamma}_k$, the average SNR at this node, can be represented as

$$\bar{\gamma}_k = \frac{p_s}{\sigma^2} \cdot \mathrm{E}\left\{ |h_{s,k}|^2 D_{s,k}^{-\nu} \right\} = \frac{p_s}{\sigma^2} \cdot \mathrm{E}\left\{ D_{s,k}^{-\nu} \right\}. \tag{3}$$

Equation (3) holds due to the fact that $h_{s,k}$ is a unit-variance random variable (r.v.) independent from $D_{s,k}$. Note that, the randomness of $D_{s,k}$ is due to the fact that the nodes are uniformly distributed on a plane. We hasten to clarify that throughout the paper all channel coefficients, distances between the source and relays, the transmitted signal from the source, and terminals' noises are treated as r.v.s and the statistical expectations are always with respect to the joint distribution of all r.v.s inside the argument.

The received signal at AC due to the $k$-th relaying node can be represented as

$$y_d^{[k]} = \sqrt{\beta_k} h_{k,d} D^{-\nu/2} y_k + n_d^{[k]} \qquad k = 1, \ldots, K \tag{4}$$

where $n_d^{[k]}$ is the noise at AC corrupting the relayed signal from $k$ and $\beta_k$ is the normalization factor at the relaying node $k$ and is given by [9]

$$\beta_k = \frac{P_T}{K \cdot \mathrm{E}\left\{|y_k|^2\right\}} \qquad k = 1, \ldots, K \tag{5}$$

where, according to (1) and (3), we have

$$\mathrm{E}\left\{|y_k|^2\right\} = p_s \mathrm{E}\left\{D_{s,k}^{-\nu}\right\} + \sigma^2 = \sigma^2\left(\bar{\gamma}_k + 1\right). \tag{6}$$

Equation (5) implies that the instantaneous transmitted power at node $k$ is equal to

$$\xi_k = \beta_k |y_k|^2 = \frac{P_T |y_k|^2}{K \cdot \mathrm{E}\left\{|y_k|^2\right\}}. \tag{7}$$

As such, *regardless* of which $K$ nodes decide to act as relays, the average transmitted power from all relays are equal and is given by $\mathrm{E}\{\xi_k\} = P_T/K, k = 1, \ldots, K$, and, therefore, the average of the total transmitted power from the whole network during the relaying phase is equal to $P_T$. This property facilitates a fair comparison between the performances of different relay selection schemes.

Assume that AC uses the optimal maximum ratio combining (MRC) scheme to estimate $x$ from the set of signals received from the relaying nodes, it can be readily shown from (1) and (4) that the soft symbol estimate at AC output is given by

$$r_d = \sum_{k=1}^{K} \frac{\sqrt{p_s \alpha_k} h_{s,k}^* h_{k,d}^* D_{s,k}^{-\nu/2}}{\sigma^2\left(1 + \alpha_k |h_{k,d}|^2\right)} \cdot y_d^{[k]} \tag{8}$$

where $^*$ is the conjugate operation and $\alpha_k \triangleq \beta_k D^{-\nu}$. It is direct to show from (4) and (8) that the resulting SNR at AC is

$$\gamma_d = \sum_{k=1}^{K} \theta_k \tag{9}$$

where

$$\theta_k \triangleq \frac{\alpha_k \left|h_{k,d}\right|^2}{1 + \alpha_k \left|h_{k,d}\right|^2} \cdot \gamma_k = \frac{\alpha_k \left|h_{k,d}\right|^2}{1 + \alpha_k \left|h_{k,d}\right|^2} \cdot \frac{p_s |h_{s,k}|^2}{\sigma^2 D_{s,k}^{\nu}}. \tag{10}$$

### D. Relay Selection Rule

It remains to set policies based on which $K$ nodes switch to the transmission mode and relay a scaled version of their received signal from $s$ to AC. Note that, these policies should be applicable to a distributed WSN in which there is no information exchange or synchronization among the nodes and no knowledge about $h_{k,d},\ k = 1, \ldots, K$ at the nodes.

Assuming that the aim is to maximize $\gamma_d$, it immediately follows from (9) that the best relaying set is the set of $K$ nodes having the $K$ largest $\theta_k$. Unfortunately, as $\theta_k$ depends on $h_{k,d}$, it is unknown at relay $k$ and, hence, cannot be directly used as a measure to select the relays. However, note from (10) that $\theta_k$ is an increasing function of $\gamma_k$, which, itself, is proportional to $|h_{s,k}|^2 D_{s,k}^{-\nu}$ that is known at the node $k$. It can be inferred from the above observation that, being unaware $h_{k,d}$, the optimal relaying set is the set of $K$ nodes with the largest $\gamma_k$. In what follows, we generalize the approach presented in [4] to show that how the $K$ nodes with the largest $\gamma_k$ can be identified and used as relays. First, a clear-to-send (CTS) signal from AC triggers each node $k$ to start its down-timer from the initial value of

$$T_k^{(\gamma)} = \frac{\lambda^{(\gamma)} D_{s,k}^{\nu}}{|h_{s,k}|^2} \tag{11}$$

where $\lambda^{(\gamma)}$ is a scalar. How to choose a proper $\lambda^{(\gamma)}$ is outlined in Subsection II-E. As the distances of AC from all nodes can be considered to be equal to $D$, all nodes simultaneously receive the CTS signal and, hence, the $k$-th node that expires its timer is the one having the $k$-th smallest $D_{s,k}^{\nu}/|h_{s,k}|^2$, or, equivalently, the $k$-th largest $\gamma_k$. As soon as a node's timer expires to zero, the node switches to the transmission mode and relays its signal to AC. This procedure continues until the total number of the received signals at AC reaches $K$. Then, AC broadcasts a hold flag and, subsequently, all nodes switch from the transmission mode back to the listening mode.

Note that as the channel fading coefficients $h_{s,k}$ may change frequently, the set of nodes with the $K$ largest SNRs may vary substantially from one source transmission frame to the next one. This implies that, in practice, all nodes may be viable relay candidates and, hence, should always be in either listening or transmission mode. This can be very energy wasteful specially in the scenarios in which $K$ is a small fraction of the total number of nodes within the transmission range of $s$. Note that the power consumption of a typical transceiver in the listening mode is three order of magnitude higher than its power consumption in the sleeping mode (when both transmit and receive circuitries are switched off) [10].

Note that the network topology is fixed or changes with a very slow rate in many WSN applications. This implies that the rate of changes in $D_{s,k}$ is typically much slower than the rate of changes in $h_{s,k}$. The latter observation along with the fact that $\gamma_k$ is inversely proportional to $D_{s,k}^{\nu}$ can be used to properly modify the above relay selection scheme such that a substantial amount of network energy is preserved at the cost of small degradation in the quality of reception at AC. Let us assume that the initial values of the nodes' counters are set to

$$T_k^{(t)} = \lambda^{(t)} D_{s,k}, \tag{12}$$

where $\lambda^{(t)}$ is a scalar. When the network receives the CTS signal, all timers start to tick down and the first $K$ nodes with the expired timers switch to the transmission mode and relay their signal. Once AC receives the $K$-th relayed signal, it issues a hold flag. Then, the $K$ nodes that have succeeded to relay their signals switch back to the listening mode while all other nodes (except the source) switch to the sleeping mode for $T^{(t)}$ seconds. $T^{(t)}$ depends on the prior knowledge about the rate of changes in the network topology and can be in the order of the time required for the transmission of several hundred data frame. In fact, IEEE 802.15.4 standard allows nodes to stay in the sleeping mode for more than 99% of the time [11]. After $T^{(t)}$ seconds, all the sleeping nodes switch back to the listening mode, set their timers to (12), and, again participate in another round of competition for acquiring the relay status.

Using the above simple approach, the $K$ closest nodes around $s$ act as relays while all other nodes preserve energy by switching to the sleeping mode. Meanwhile, every $T^{(t)}$ seconds the list of the $K$ closest nodes is updated to adapt to the possible changes in the network topology. Note that, the $k$-th closest node to $s$ is not necessarily the one with the $k$-th largest SNR. Therefore, some SNR performance degradation is expected at AC. However, simulation results in Section IV show that this performance loss is in fact negligible in our multiple-relay cooperative communication scheme.

Although the above topology-based relay selection scheme is substantially more energy efficient when compared to its optimal SNR-based counterpart, it may overexploit certain nodes around the source. In fact, if the WSN topology changes too slowly, the set of $K$ closest nodes around the source rarely changes. As such, these nodes switch back and forth between the listening and the transmission modes leading to their early battery depletion while all the other nodes keep staying in the sleeping mode. To avoid this problem, we propose to randomly select the relays from a neighborhood of radius $R$ around $s$ by letting the nodes on $O(s, R)$, the disc of radius $R$ centering at $s$, set their initial counter to a randomly generated value

$$T_k^{(r)} = \lambda_k^{(r)}. \tag{13}$$

When CTS is received, all the nodes on $O(s, R)$ start down-counting and the $K$ nodes whose timers expire first, relay their received signals. After receiving a hold flag from AC, the relaying nodes switch back to the listening mode for another round of relaying while the rest switch to the sleeping mode for a predetermined $T^{(r)}$ seconds. After $T^{(r)}$ seconds, all the sleeping nodes switch back to the listening mode and set the initial value of their counters to a *newly generated* random value. Using the above scheme, while a significant amount of energy is saved, all nodes on $O(s, R)$ have a similar chance of acquiring the relay status over a long period of time. Therefore, the risk of over-exploiting certain nodes diminishes. The price that may have to be paid is the possibility of considerable SNR reduction at AC due to the fact that the relaying nodes are chosen randomly. Note that, while $R$ should be large enough to include more than $K$ nodes around the source, it should not be excessively large to reduce the chance of selecting the nodes that are very far away from $s$ and receive a low quality version of the transmitted signal. In fact, as shown in Subsection III-B, if the randomly selected nodes are close enough to the source, they enjoy an average SNR level that is enough to provide AC an average SNR close to that of the topology-based and the optimal relay selection schemes.

### E. Collision-free transmission

The probability that two relays start their transmissions exactly at the same time is zero in all of the above relay selection schemes. However, it is possible that while a relay is transmitting, some other nodes expire their counters and start their transmission. A simple approach to decrease the probability of such a collision and maintain orthogonal transmissions among relays is to increase $\lambda^{(\bullet)}$ where, depending on the applied relay selection scheme, $\bullet$ is "$\gamma$", "$t$", or "$c$" (see, [4] for a similar treatment). While the collision probability can be arbitrarily decreased by increasing $\lambda^{(\bullet)}$, the price that has to paid is increasing the length of the relaying phase. To prevent large gaps between two consecutive relay transmissions due to excessively large $\lambda^{(\bullet)}$, we briefly outline the following alternative collision avoidance approach: Assume that nodes $k+1, \ldots, k+l$ expire their counters while relay $k$ is transmitting. Then, each node $w \in \{k+1, \ldots, k+l\}$ waits for $\varsigma_w$ seconds after the relay $k$ ends its transmission, and, then, if any other relay does not transmit, starts its transmission, otherwise, it waits for another $\varsigma_w$ seconds. Note that $\varsigma_w$ is a small node-specific random variable and waiting for $\varsigma_w$ seconds is necessary to avoid collision among nodes $k+1, \ldots, k+l$. Note also that, the above approach requires each node to be able to overhear the signal transmitted from the others. It should also be mentioned that, using this approach, a node $w$ may lose its place in the relaying queue to other nodes. However, the nodes that have the chance to replace node $w$ in the queue are those whose original places closely follow that of $w$ in the queue. Therefore, the signal reception performance at AC should not be significantly affected by such a queue jumping. Other collision avoidance schemes that require some coordination among nodes can be found in, e.g., [10], [12].

In the next section, we analyze the SNR performances of the above topology-based and random relay selection schemes and derive a sufficient condition under which the above two schemes result in approximately equal SNR performances at AC. We leave the analytical results concerning the optimal SNR-based scheme to our reports elsewhere.

## III. Analysis of the Selection Schemes

In this section, we obtain the average SNR at the $k$-th selected relay $\bar{\gamma}_k$ as well as the average SNR at AC $\bar{\gamma}_d = \mathrm{E}\{\gamma_d\}$ for the two latter relay selection schemes presented in Subsection II-D. The superscripts "$(t)$" and "$(r)$" are used to denote the average SNR values corresponding to the topology-based and the random relay selection schemes, respectively.

### A. Average SNR at the $k$-th selected relay

It can be observed from (3) that $\bar{\gamma}_k$ is proportional to $\mathrm{E}\left\{D_{s,k}^{-\nu}\right\}$, which, itself, depends on the probability density function (PDF) of $D_{s,k}$. Note that the PDF of $D_{s,k}$ depends on how the $k$-th relay is selected. It is direct to show that [6]

$$f_{D_{s,k}}^{(t)}(x) = \frac{2(\rho\pi)^k}{(k-1)!} x^{2k-1} e^{-\rho\pi x^2} \tag{14}$$

where $f_{D_{s,k}}^{(t)}(x)$ is the PDF of $D_{s,k}$ at the distance $x$ in the topology-based relay selection scheme, while [1]

$$f_{D_{s,k}}^{(r)}(x) = \frac{2x}{R^2} \tag{15}$$

where $f_{D_{s,k}}^{(r)}(x)$ is the PDF of $D_{s,k}$ at the distance $x$ in the random relay selection scheme. It follows from (3) and (14) that for the topology-based relay selection scheme we have

$$\bar{\gamma}_k^{(t)} = \frac{2p_s(\rho\pi)^k}{\sigma^2(k-1)!} \int_1^\infty x^{2k-1-\nu} e^{-\rho\pi x^2} dx. \tag{16}$$

Depending on the sign of $k-\frac{\nu}{2}$, the solution to the integral in (16) takes two different forms. When $k-\frac{\nu}{2}\geq 0$, it can be shown that

$$\bar{\gamma}_k^{(t)}=\frac{p_s(\rho\pi)^{\frac{\nu}{2}}}{\sigma^2(k-1)!}\Gamma\left(k-\frac{\nu}{2},\rho\pi\right) \quad (17)$$

where $\Gamma(a,z)=\int_z^\infty t^{a-1}e^{-t}dt$ is the upper incomplete Gamma function. For the case when $k-\frac{\nu}{2}<0$, the integration by parts can be used to show that

$$\bar{\gamma}_k^{(t)}=\frac{p_s(\rho\pi)^{\frac{\nu}{2}}}{\sigma^2(k-1)!}\left(\Gamma(\zeta,\rho\pi)-\sum_{j=0}^{\lceil\frac{\nu}{2}-k\rceil-1}\frac{e^{-\rho\pi}(\rho\pi)^{j+k-\frac{\nu}{2}}}{a_j}\right) \quad (18)$$

where $\zeta=\lceil\frac{\nu}{2}-k\rceil-(\frac{\nu}{2}-k)$ with $\lceil z\rceil$ being the smallest integer greater than $z$, $a_{-1}=1$, and $a_j=a_{j-1}(k-\frac{\nu}{2}+j)$.

Equations (3) and (15) can be used to show that for the random relay selection scheme we have

$$\bar{\gamma}_k^{(r)}=\frac{p_s}{\sigma^2}\int_1^R\frac{2x^{1-\nu}}{R^2}dx=\begin{cases}\frac{p_s}{\sigma^2}\cdot\frac{2\ln(R)}{R^2} & \nu=2\\ \frac{p_s}{\sigma^2}\cdot\frac{2(R^{2-\nu}-1)}{R^2(2-\nu)} & \nu\neq 2\end{cases}. \quad (19)$$

Note from (17) and (18) that $\bar{\gamma}_k^{(t)}$ is an explicit function of $k$. This, along with (5) and (6), implies that $\beta_k^{(t)}$, the normalization factor at the $k$-th relaying node in the topology-based relay selection scheme, depends on $k$. Therefore, when the latter relay selection scheme is used, each relaying node should be aware of its place in the relaying queue. This may be done by counting the total number of overheard signals transmitted from the preceding $k-1$ relays. Then, using, for instance, a look-up table, $\beta_k^{(t)}$ can be determined and used as the $k$-th relay normalization factor.

Note from (19) that $\bar{\gamma}_k^{(r)}$, and, hence, $\beta_k^{(r)}$, are independent from $k$. Therefore, when the random relay selection scheme is used, the relays may be unaware of their position in the relaying queue.

### B. Average SNR at AC

From (9) and (10) it follows that the average SNR at AC is given for both relay selection schemes by

$$\bar{\gamma}_d^{(\bullet)}=\sum_{k=1}^K\underbrace{\mathrm{E}\left\{\frac{\alpha_k^{(\bullet)}|h_{k,d}|^2}{1+\alpha_k^{(\bullet)}|h_{k,d}|^2}\right\}}_{\phi_k^{(\bullet)}}\bar{\gamma}_k^{(\bullet)} \quad (20)$$

where $\bullet$ is either "$t$" or "$r$" depending on the applied relay selection scheme. As the channel coefficients $h_{k,d}, k=1,\ldots,K$ are unit-variance Rayleigh distributed, we have

$$\phi_k^{(\bullet)}=\int_0^\infty\frac{\alpha_k^{(\bullet)}xe^{-x}}{1+\alpha_k^{(\bullet)}x}dx=1-\alpha_k^{(\bullet)^{-1}}e^{\alpha_k^{(\bullet)^{-1}}}\mathrm{E}_1\left(\alpha_k^{(\bullet)^{-1}}\right) \quad (21)$$

where

$$\mathrm{E}_1(z)=\int_z^\infty\frac{e^{-t}}{t}dt\approx\frac{e^{-z}}{z}\left(1-\frac{1}{z}\right) \quad (22)$$

and the approximation is valid for $z\gg 1$. Substituting (21) into (20) yields

$$\bar{\gamma}_d^{(\bullet)}=\sum_{k=1}^K\bar{\gamma}_k^{(\bullet)}\left(1-\alpha_k^{(\bullet)^{-1}}e^{\alpha_k^{(\bullet)^{-1}}}\mathrm{E}_1\left(\alpha_k^{(\bullet)^{-1}}\right)\right). \quad (23)$$

Note that due to the assumption that AC is in the far field, we have $\alpha_k^{(\bullet)^{-1}}\gg 1$. Therefore, the approximation in (22) can be used in (23) to show that

$$\bar{\gamma}_d^{(\bullet)}\approx\sum_{k=1}^K\bar{\gamma}_k^{(\bullet)}\alpha_k^{(\bullet)}=\sum_{k=1}^K\frac{\bar{\gamma}_k^{(\bullet)}\beta_k^{(\bullet)}}{D^\nu}=\frac{1}{K}\sum_{k=1}^K\frac{P_T}{\sigma^2D^\nu}\cdot\frac{\bar{\gamma}_k^{(\bullet)}}{\bar{\gamma}_k^{(\bullet)}+1} \quad (24)$$

where the expression of $\beta_k^{(\bullet)}$ from (5) and (6) is used in the second equation of (24). It is important to notice from (24) that when the average SNRs at all selected relays are high enough, that is,

$$\bar{\gamma}_k^{(\bullet)}\gg 1,\qquad k=1,\ldots,K \quad (25)$$

then, the average SNR at AC can be further simplified to

$$\bar{\gamma}_d^{(\bullet)}\approx\frac{P_T}{\sigma^2D^\nu} \quad (26)$$

Interestingly, (26) is independent of the relay selection scheme. The above discussion reveals the simple and very useful fact that as long as the average SNRs at all selected relays are high enough, the SNR performance at AC is insensitive to the approach used to select the relays. In such a case, the random relay selection scheme that is not only energy efficient but also fair towards all nodes can be applied without compromising the performance.

### C. SNR variance at AC for large $K$

It can be observed from (9) and (10) that $\gamma_d^{(\bullet)}$ depends on $D_{s,k}$, $h_{s,k}$, and $h_{k,d}$ for $k=1,\ldots,K$. This implies that the joint PDFs of all possible pairs from the above r.v.s are required to compute $\mathrm{var}(\gamma_d^{(\bullet)})$. When the random relay selection scheme is used, all the above r.v.s are independent from one another and it is direct to show that, as $K$ grows, $\mathrm{var}(\gamma_d^{(r)})$ converges to zero with rate $\mathcal{O}(1/K)$. On the other hand, when the topology-based relay selection scheme is used, $D_{s,l}$ and $D_{s,m}$ are not independent anymore. In such a case, it can be proven that the joint PDF of $D_{s,l}$ and $D_{s,m}$ for $m>l$ is given by

$$f_{D_{s,l},D_{s,m}}^{(t)}(r,t)=\frac{4(\rho\pi)^m t^{2(m-l)-1}r^{2l-1}}{(l-1)!(m-l-1)!e^{\rho\pi t^2}}\left(1-\frac{r^2}{t^2}\right)^{m-l-1} \quad (27)$$

when $t\geq r$, while $f_{D_{s,l},D_{s,m}}^{(t)}(r,t)=0$ when $t<r$. Using the above joint PDF, it can also be shown that $\mathrm{var}(\gamma_d^{(t)})$ converges to zero as $K$ grows.

Convergence of $\mathrm{var}(\gamma_d^{(\bullet)})$ to zero implies that, if $K$ is large enough, then, for any arbitrary set of realizations of $D_{s,k}$, $h_{s,k}$, and $h_{k,d}$, $\gamma_d^{(\bullet)}$ should be close to $\bar{\gamma}_d^{(\bullet)}$. This further confirms that $\bar{\gamma}_d^{(\bullet)}$ is a sensible performance measure for the considered cooperative WSN. Moreover, the above result verifies that the proposed relay selection schemes effectively decrease the signal power fluctuations at AC, and, hence, provide system diversity.

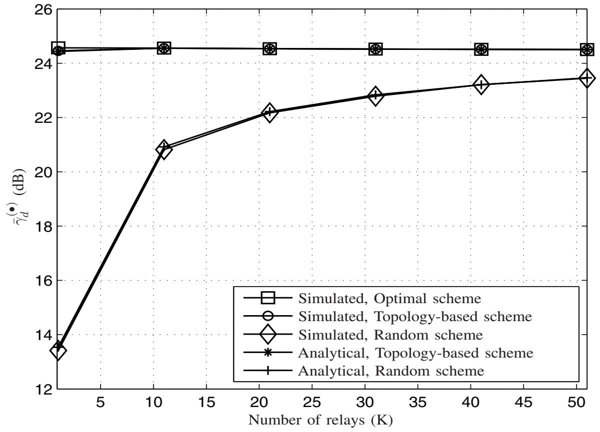


Fig. 2. The $\bar{\gamma}_d^{(\bullet)}$ curves versus $K$ for the different selection schemes.

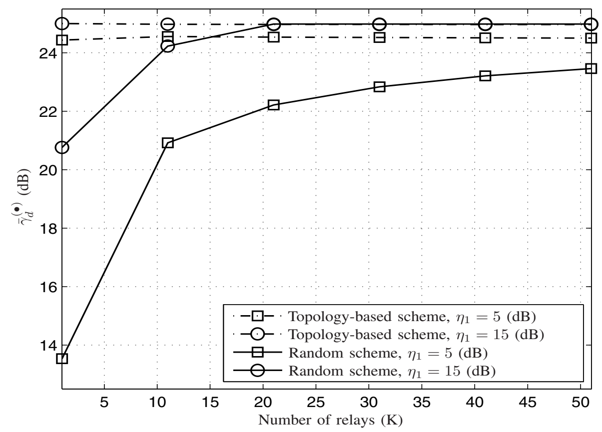


Fig. 3. The $\bar{\gamma}_d^{(\bullet)}$ curves for the topology-based and random selection schemes for different values of $\eta_1$.

## IV. Simulations

In this section, we use computer simulations to verify the analytical results presented in Section III. Fig. 2 shows the analytical and the simulated $\bar{\gamma}_d^{(\bullet)}$ curves for the topology-based and random relay selection schemes. The simulated SNR curve of the optimal relay selection scheme is also plotted in the figure. In Fig. 2, $R = 50$, $D = 1000$, $\rho = 0.1$, and $\nu = 2$ are chosen. Moreover, $P_T$ and $p_s$ are selected such that $\eta_1 \triangleq p_s R^{-\nu}/\sigma^2 = 5$ (dB) while $\eta_2 \triangleq P_T D^{-\nu}/\sigma^2 = 25$ (dB). Note that such a value of $p_s$ implies that (25) does not necessarily hold, and, therefore, there may be a significant difference between $\bar{\gamma}_d^{(t)}$ and $\bar{\gamma}_d^{(r)}$. As can be observed from the figure, this difference is in fact noticeable. Fig. 2 also shows that the analytical curves obtained from (23) closely follow their simulated counterparts. The figure also verifies that there is no significant loss in the SNR performance of the topology-based relay selection scheme comparing to the optimal relay selection counterpart.

Fig. 3 shows the curves of $\bar{\gamma}_d^{(\bullet)}$ for the topology-based and random selection schemes and for two different values of $\eta_1 = 5$ (dB) and $\eta_1 = 15$ (dB). All other parameters remain unchanged from Fig. 2. Note that, as $\eta_1$ increases from 5 (dB) to 15 (dB), $\bar{\gamma}_k^{(\bullet)}$ increase to the level that (25) holds for most of the selected relays. As can be observed from Fig. 3, in such a scenario, $\bar{\gamma}_d^{(r)}$ is very close to $\bar{\gamma}_d^{(t)}$ specially if the number of relays is large enough. This corroborates the discussion in Subsection III-B where it is shown that if the average SNRs at all selected relays are high enough, then the average SNR performance at AC is independent from the scheme used to select the relays.

## V. Conclusions

We have considered a multiple-relay selection problem in a large wireless sensor network (WSN) and presented an optimal relay selection scheme that maximizes the SNR at the access point (AC) while complying with the constraints posed by the unsupervised nature of WSN. Two other more energy efficient relay selection alternative techniques have also been developed: The first technique takes advantage of the slow changes in the network topology and let those nodes that deemed inappropriate to relay preserve energy by staying in the sleeping mode. The second technique randomly selects the relays from a neighborhood around the source while let all unselected nodes to stay in the sleeping mode. It is shown through simulations that the topology-based relay selection technique achieves a SNR close to that of its optimal counterpart while the random relay selection alternative that fairly treats all nodes may result in a SNR reduction at AC.

The SNR performances of the two energy efficient relay selection schemes have also been analytically studied and it has been proven that as long as the received signal power from the source is high enough at all selected relays, the SNR at AC is insensitive to the approach used to select the relays.

## References

[1] H. Ochiai, P. Mitran, H. V. Poor, and V. Tarokh, “Collaborative beamforming for distributed wireless ad hoc sensor networks,” *IEEE Trans. Signal Process.,* vol. 53, pp. 4110-4124, Nov. 2005.

[2] J. N. Laneman and G. W. Wornell, “Distributed space-time coded protocols for exploiting cooperative diversity in wireless networks,” *IEEE Trans. Inf. Theory,* vol. 49, pp. 2415-2425, Oct. 2003.

[3] I. Maric and R. D. Yates, “Cooperative multicast for maximum network lifetime,” *IEEE J. Sel. Areas Commun.,* vol. 23, pp. 127-135, Jan. 2005.

[4] A. Blestas, A. Khisti, D. P. Reed, and A. Lippman, “Asimple cooperative diversity method based on network path selection,” *IEEE J. Sel. Areas Commun.,* vol. 24, pp. 659-672, Mar. 2006.

[5] A. Bletsas, H. Shin, and M. Z. Win, “Cooperative communications with outage-optimal opportunistic relaying,” *IEEE Trans. Wireless Commun.,* vol. 6, pp. 3450-3460, Sep. 2007.

[6] M. Haenggi, “On distances in uniformly random networks,” *IEEE Trans. Information Theory,* vol. 51, pp. 3584-3586, Oct. 2005.

[7] Y. Fan and J. Thompson, “MIMO configurations for relay channels: Theory and practice,” *IEEE Trans. Wireless Communications,* vol. 6, pp. 1774-1786, May 2007.

[8] V. I. Morgenshtern and H. Bölcskei, “Crystallizations in large wireless networks,” *IEEE Trans. Inf. Theory,* vol. 53, pp. 3319-3349, Oct. 2007.

[9] X. Deng and A. M. Haimovich, “Power allocation for cooperative relaying in wireless networks,” *IEEE Commun. Lett.,* vol.9, no.11, pp. 994-996, Nov. 2005.

[10] W. Ye, J. Heidemann, and D. Estrin, “Medium access control with coordinated adaptive sleeping for wireless sensor networks,” *IEEE/ACM Trans. Networking,* vol. 12, pp. 493-506, June 2004.

[11] *Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LRW-PANs),* IEEE Std. 802.15.4, 2003.

[12] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *IEEE Trans. Inf. Theory,* vol. 46, pp. 388-404, Mar. 2000.