Analytic Conditions For Energy Neutrality In Uniformly-formed Wireless Sensor Networks

Hana Besbes, George Smart *Student Member, IEEE*, Dujdow Buranapanichkit, Christos Kloukinas and Yiannis Andreopoulos* *Member, IEEE*

***Abstract*—Future deployments of wireless sensor network (WSN) infrastructures for environmental or event monitoring are expected to be equipped with energy harvesters (e.g. piezoelectric, thermal, photovoltaic) in order to substantially increase their autonomy. In this paper we derive conditions for *energy neutrality*, i.e. perpetual energy autonomy per sensor node, by balancing the node's expected energy consumption with its expected energy harvesting capability. Our analysis assumes a uniformly-formed WSN, i.e. a network comprising identical transmitter sensor nodes and identical receiver/relay sensor nodes with a balanced cluster-tree topology. The proposed framework is parametric to: *(i)* the duty cycle for the network activation; *(ii)* the number of nodes in the same tier of the cluster-tree topology; *(iii)* the consumption rate of the receiver node(s) that collect (and possibly relay) data along with their own; *(iv)* the marginal probability density function (PDF) characterizing the data transmission rate per node; *(v)* the expected amount of energy harvested by each node. Based on our analysis, we obtain the number of nodes leading to the *minimum* energy harvesting requirement for each tier of the WSN cluster-tree topology. We also derive closed-form expressions for the difference in the minimum energy harvesting requirements between four transmission rate PDFs in function of the WSN parameters.**

Our analytic results are validated via experiments using TelosB sensor nodes and an energy measurement testbed. Our framework is useful for feasibility studies on energy harvesting technologies in WSNs and for optimizing the operational settings of hierarchical WSN-based monitoring infrastructures prior to time-consuming testing and deployment within the application environment.

***Keywords*—wireless sensor networks, energy harvesting, energy neutrality, analytic modeling**

I. Introduction

ENERGY AUTONOMY is widely recognized as one of the key challenges for monitoring infrastructures based on wireless sensor networks (WSNs) [1]. Several works approach the problem of energy efficiency focusing on a particular aspect of the WSN-based monitoring [2]. Technology-oriented approaches design new circuits and systems for more efficient harvesting [3], [4], [5], or strive for more efficient scheduling and transmission protocols [6], [7], [8], [9], [10]. These try to bridge the gap between data sensing and transmission requirements and the corresponding energy harvesting and energy storage capability of the underlying hardware. Other approaches propose optimal energy management policies under given energy harvesting, sensing and transmission capabilities [11], [12], [13], [14], [15]. Such policies optimize the manner each sensor node performs its data gathering and buffer management in order to minimize the required energy consumption. For a node that has an energy storage unit (battery) that can store hundreds of hours worth of operating energy, if the expected energy dissipation over a time interval, e.g. 24 hours, is matched with the amount of energy expected to be harvested from the environment within the same interval, it can be said that the sensor node achieves *energy neutrality* [2]. That is, the node is expected to be able to operate in perpetuity without the requirement for human intervention. Due to the physical limitations of harvesting technology, practical energy neutrality is achievable today under the notion of *duty cycling* [16], where sensor nodes are suspended during long periods of inactivity in order to preserve (and replenish) their battery resources.

In this paper, we approach the problem of energy neutrality in a more holistic, system-oriented, manner. Specifically, we focus on the common application scenario of a monitoring infrastructure where sensor nodes follow a periodic duty cycle in order to capture and transmit measurements to a base station, or to another node that relays the information to a base station. We derive a parametric model for energy neutrality in function of the system settings under the assumption of a *uniformly-formed* WSN, i.e. a network of identical sensor nodes that are: *(i)* producing data traffic with the same statistical characterization and *(ii)* connected to the base station via a *cluster-tree topology* [8] represented by a symmetric and acyclic graph with balanced bandwidth allocation per link. Within this framework, the key advance of our work in comparison to previous work on optimal energy management policies [2] [11], [12], [13], [14] is that we provide closed-form expressions for the minimum-required harvested energy in order for each node to remain energy neutral. Our specific contributions are:

- For each tier of a WSN cluster-tree topology, analytic derivation of the number of nodes that leads to the minimum requirement for harvested energy under four commonly encountered marginal PDFs for the data transmission rate per sensor.

*Corresponding author. HB, GS, DB and YA are with the Electronic and Electrical Engineering Department, University College London, Roberts Building, Torrington Place, London, WC1E 7JE; tel. +442076797303; fax. +442073889325; email: {hana.besbes.11, george.smart, d.buranapanichkit, i.andreopoulos}@ucl.ac.uk. CK is with the School of Informatics, City University London, Northampton Square, London, EC1V 0HB; tel. +442070408848; fax. +442070400244; email: c.kloukinas@city.ac.uk. This work was supported in part by the UK EPSRC (grant EP/K033166/1) and the EU (project IoT@Work FP7-257367). D. Buranapanichkit was supported by a Royal Thai Government Scholarship (Science and Technology).

- Analytic comparison of the minimum requirements for energy harvesting under different application parameters and different data transmission rates.
- Validation of the theoretical results via: *(i)* an energy measurement testbed and TelosB nodes employing a recently-proposed collision-free protocol (TFDMA [6]) *(ii)* establishment of optimal operational parameters within two application scenarios for WSN-based monitoring.

In Section II, we present the system model corresponding to the application scenarios under consideration. The analytic derivations characterizing energy-neutral operation under different data transmission rates are presented in Section III, where we also derive the minimum requirement for harvested energy under various widely-used statistical characterizations for the data transmission rate. Section IV presents the experimental validation of the proposed analytic formulations for energy-neutral operation, Section V presents results within two applications and Section VI concludes the paper.

II. System Model

We consider a set of wireless sensor nodes connected to a "sink" node, which represents the collecting unit, i.e. a base station with power supply. This connection could be direct; alternatively, under a symmetric and balanced cluster-tree topology [1], [17], each node could be linked to a "relay" node that conveys its measurements (along with its own) to another relay node or, eventually, to the base station. Interference between neighboring nodes can be avoided by using simple heuristics or graph coloring approaches in conjunction with transmission and reception in different channels. For example, a node can listen to Channel X and transmit in Channel Y, with $X \neq Y$ [6], [18], [19], [20], [21], [22], [23], [24]. We illustrate such examples in Figure 1. The ellipses indicate the coverage of each receiver, with their channel allocated such that, under appropriate scheduling of transmissions within each tier, no interference is caused. The links indicate the bandwidth available to each transmitting sensor node. The figure shows that the essentials of the problem boil down to the analysis of the interaction between each sensor node and its corresponding base station or relay node at the same tier of the cluster-tree topology.

A. System Description

For our analysis, we assume that, for harvesting interval of T seconds, the sensor nodes are continuously active for T_{act} seconds. This defines the *duty cycle*

$$c = \frac{T_{\text{act}}}{T}. \tag{1}$$

The network activation can be triggered by external events or by scheduled data gathering with rate c over the duration of the application, $0 < c < 1$. Examples are: data acquisition and transmission in environmental monitoring [1], event-driven activation for surveillance [2], and adaptive control of duty cycling for energy management [12] [16]. Thus, the value of c can be adjusted statically or dynamically, depending on the application environment.

When the sensor nodes are activated, they first converge into a *balanced time-frequency steady-state mode*, where each node joins the base station (or a relay node) on a particular channel such that: *(i)* the number of nodes coupled to the base station or each relay node is balanced; *(ii)* each cluster-tree tier accommodates transmissions from n nodes without collisions. Several low-energy (centralized or distributed) WSN protocols, such as EM-MAC [7], wirelessHART [19], IEEE 802.15.4 GTS [8] and TFDMA [6] can achieve this goal. For example, TFDMA achieves this for 16 nodes and 4 channels within 3-5 seconds [6], while the centralized IEEE 802.15.4 GTS can establish collision-free single-channel time division multiple access (TDMA) within 1-2 seconds [8]. While energy is consumed for the protocol setup and the establishment of the cluster-tree configuration, the payoff for the WSN is the achievement of balanced, collision-free, steady-state operation with predictable characteristics during the active period. Examples of several uniformly-formed topologies that can operate in collision-free steady-state mode are given in Figure 1.

Each sensor captures, processes and transmits (and potentially relays) data. We assume that the transmission data rate varies; it will thus be modeled as a random variable. The rate variability may stem from: adaptive sensing strategies [25], packet retransmissions or protocol adaptivity to mitigate interference effects [7], and variable-rate data encoding [26] to reduce the transmission bitrate and ensure robustness to packet erasures [27]. Thus, due to these factors, the number of bits sent within each transmission slot of the utilized protocol varies, despite the fact that the physical layer rate is fixed for most WSN systems using the IEEE 802.15.4 PHY.

Within each tier of the cluster-tree topology, depending on the amount of data to be transmitted, a node may need to: *(i)* stay awake (beaconing and radio on) if less bits have to be sent than what is possible within its transmission slot; *(ii)* buffer the residual data if more bits must be sent than what its slot permits. Once the active period of T_{act} seconds lapses, each node suspends its activity (i.e. goes into "sleep" mode) in order to conserve energy. Figure 2 shows two examples of TDMA transmission slots during the active period. During both the active and sleep modes, each sensor harvests energy based on its on-board harvesting unit (e.g. piezoelectric harvester or solar cell).

Concerning the on-board battery of each sensor, given that modern IEEE 802.15.4 compliant sensors (e.g. TelosB, micaZ, STM32W motes, etc.) can be powered by their on-board batteries for very long time intervals (e.g. hundreds of hours of continuous operation), the battery capacity can be assumed to be infinite compared to the energy budget spent and harvested within each interval of T seconds [2], [10], [15]. In addition, due to the assumption of infinite battery capacity, issues such as leakage current and battery aging do not need to be considered.

We remark that practical WSN transceiver hardware reacts in intervals proportional to one packet transmission (or to the utilized time-frequency slotting mechanism). Thus, the

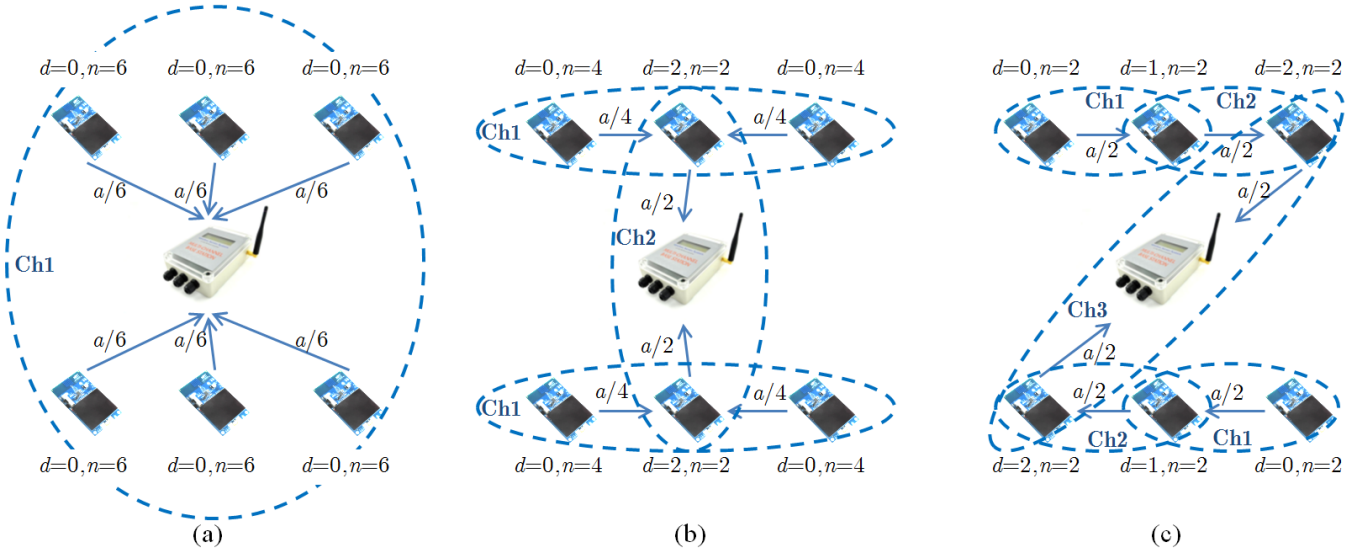

Figure 1: Three interference-free uniformly-formed topologies within a WSN comprising six identical sensor nodes and one base station, with a indicating the consumption rate of each receiver/relay node (in bits-per-second); (a): direct (one-tier) connection to the base station using a single channel. (b): two-tier cluster-tree topology using two channels. (c): three-tier cluster-tree topology using three channels. For illustration purposes the channel number coincides with the tier number, with the highest tier containing the base station. We indicate (via d) the additional nodes whose traffic is relayed by each node, as well as the number of nodes in the same tier of the cluster-tree topology (via n).

transmission and reception of data is not strictly a continuous process. However, energy consumption within each sensor node *is* strictly continuous as, regardless of the transceiver, each sensor node is active for the entire duration of T_{act} seconds by sensing, processing data (e.g. to remove noise or to perform data encoding) and other runtime operations related to data gathering, processing and transmission (such as buffer management at the application, medium access and physical layers and service interrupts of the runtime environment).

B. Definitions

When the WSN goes into the active state, we assume that k Joule is consumed by each sensor node in order to reach the balanced, collision-free, steady-state operation via one of the well-known centralized or distributed mechanisms suitable for this purpose [6], [18], [19], [17], [7]. During the steady-state operation of each node, the average energy rate consumed to process and transmit data is g Joule-per-bit.

1) Data Production and Energy Harvesting : Because the data production and transmission by each sensor node is a non-deterministic process, the data transmission rate (in bits-per-second) is modeled by random variable (RV) Ψ with PDF $P(\psi)$. The statistical modeling of this rate can be gained by observing the occurred physical phenomena and analyzing the behavior of each node when it captures, processes and transmits bits, in conjunction with the data relayed by other nodes of the same tier (if the node is also a relay in the WSN). Alternatively, the data production and transmission rate can be controlled (or "shaped") by the system designer in order to achieve a certain goal, such as limiting the occurring latency or, in our case, to minimize the harvested energy required in order to operate each node in perpetuity. Examples of systems with variable data transmission rates include visual sensor networks transmitting compressed video frames or image features [28], [29], [30], [31], as well as activity monitoring or localization networks where the data acquisition is irregular and depends on the events occurring in the monitored area [32], [33], [34].

The energy harvesting process is also a non-deterministic process. This means that the harvested energy profile changes over time, depending on the surrounding environmental conditions and availability of the energy source [3]. For example, solar panel or piezoelectric energy scavenging mechanisms produce different levels of power at different times of the day, depending on the environmental conditions and on whether they are placed indoors or outdoors [3], [35]. Therefore, the power (Watt) produced by the harvesting mechanism is modeled by RV $X \sim P(\chi)$.

Since both the data rate and the power produced by the harvester may be non-stationary, we assume their marginal statistics for $P(\psi)$ and $P(\chi)$, which are derived starting from a doubly stochastic model for these processes. Specifically, such marginal statistics can be obtained by [36], [37]: *(i)* fitting PDFs to sets of past measurements of data rates and power, with the statistical moments (parameters) of such distributions characterized by another PDF; *(ii)* integrating over the parameter space to derive the final form of $P(\psi)$ and $P(\chi)$. For example, if the data transmission rate is modeled as a Half-Gaussian distribution with variance parameter that is itself exponentially distributed, by integrating over the parameter

space, the marginal statistics of the data rate become Laplacian [36], [37]. The disadvantage of using marginal statistics for the data transmission rate and the power produced by the harvester is the removal of the stochastic dependencies to transient physical properties of these quantities. However, in this work we are interested in the *expected* requirements for energy harvesting to maintain energy neutrality over a lengthy time interval (e.g. several hours) and not in the *variations* of energy harvesting over short time intervals. Such variations are irrelevant since the on-board batteries of each node can support its stand-alone operation for hundreds of hours if needed. Thus, a mean-based analysis using the marginal statistics is suitable for this purpose.

2) Data Consumption and Energy Penalties : The data consumption rate of the application layer of each receiver under the employed collision-free steady-state operation is a bits-per-second (bps). For example, under the IEEE 802.15.4 physical layer and the CC2420 transceiver, $a \cong 144$ kbps at the application layer under the NullMAC and NullRDC options of Contiki operating system[1]. Since n identical sensor nodes transmit data at the same tier of the cluster-tree topology (Figure 1), we define the ratio $\frac{a}{n}$ as the *coupling point* of each receiver within each tier. This means that, in the ideal case, each sensor node should transmit its captured data at the rate of $\frac{a}{n}$ bps. However, given the time-varying nature of the data transmission rate per node, beyond the energy for data processing (e.g. encoding) and transmission we encounter the following two cases: *(i)* receiver underloading, where $\Psi < \frac{a}{n}$ and "idle" energy is consumed by the node with rate b Joule-per-bit (J/b) by staying active during transmission opportunities for synchronization and other runtime purposes (e.g. transmitting beacon messages [17], [6]); *(ii)* receiver overloading, where $\Psi > \frac{a}{n}$ and "penalty" energy is consumed with rate p J/b by the sensor to buffer (and retrieve) the data prior to transmission. Examples of both are illustrated in Figure 2 for TDMA-based collision-free transmission [6], [38]. The nomenclature summary of our system model is given in Table I.

III. Characterization of Energy Neutrality

We derive the analytic conditions that correspond to the *minimum energy harvesting required* in order to maintain energy neutrality in the system model described previously. There are two modes of operation with complementary energy profiles: the *active mode*, where energy is (primarily) consumed, and the *sleep* (or suspend) *mode*, where each node is suspended and energy is harvested in order to replenish the node's battery resources. During both the sleep and the active modes, each sensor node is expected to harvest $T\int_0^\infty \chi P(\chi)\, d\chi = TE[X]$ Joule from the surrounding environment.

During the active mode period of cT seconds we define five components for the energy consumption for each sensor node, most of which are pictorially illustrated in Figure 2:

[1] https://github.com/kb2ma/contiki/wiki/Change-mac-or-radio-duty-cycling-protocols contains more details; the NullMAC mechanism does not do any MAC-level processing and leads to the maximum energy efficiency, assuming that the application layer handles the transmission opportunities and data buffering.

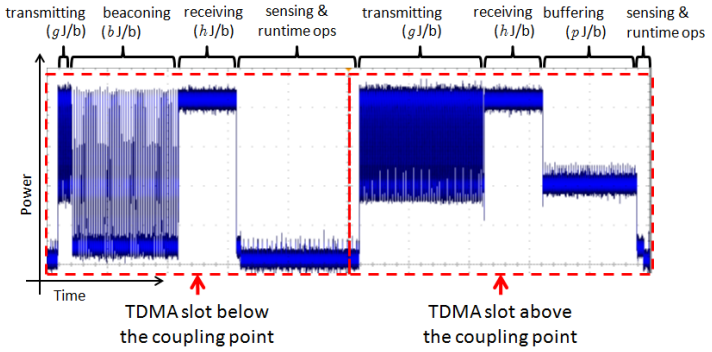

Figure 2: Energy profile of a TelosB sensor node within an undercoupled and an overcoupled TDMA slot during the active period. The indicated metrics (in Joule-per-bit) are defined in Table I.

Table I: Nomenclature table.

Symbol	Unit	Definition
c	–	Duty cycle
T, T_{act}	s	Harvesting time interval, active time interval
n	–	Number of transmitting sensor nodes *at the same tier* of the cluster-tree topology
d	–	Number of *additional* sensor nodes whose traffic is relayed by each node at a given tier of the cluster-tree topology
k	J	Energy consumed for wake-up, set-up and convergence
g	J/b	Energy for processing and transmitting one bit
p	J/b	Penalty energy for storing one bit during receiver overloading
b	J/b	Energy during idle periods for the time interval corresponding to one bit transmission
h	J/b	Energy for receiving and temporary buffering one bit under the relay case
a	bps	Data consumption rate of a relay node (or base station)
r	bps	Average data transmission rate per node
$\Psi \sim P_{d+1}(\psi)$	bps	RV modeling the data production and transmission rate per node that is also relaying data from d other nodes
$E_{d+1}[\Psi]$	bps	Expected data production and transmission rate per node that is also relaying data from d other nodes
$X \sim P(\chi)$	W	RV modeling the power harvested by each node
$E[X]$	W	Expected power harvested by each node
E_n	J	Node residual energy (harvested minus consumed) over the harvesting time interval T

1) *Setup and convergence energy*. Each node is activated once during the harvesting time interval. Thus the energy to converge to steady state is k J. We remark that the convergence time is at least two orders of magnitude smaller than T_{act} (e.g. 1−5 s vs. $T_{\text{act}} = 400$ s) and can be considered negligible in comparison to T_{act}.

2) *Energy for processing and transmitting* the node's own data and the data relayed to it from d other nodes, given by $cTg\int_0^\infty \psi P_{d+1}(\psi)\,d\psi = cTgE_{d+1}[\Psi]$ J, with $E_{d+1}[\Psi] \equiv (d+1)\,E[\Psi]$ and $E[\Psi]$ the expected transmission rate of each node that is not a relay. If $E_{d+1}[\Psi] > \frac{a}{n}$ (i.e. the mean transmission rate is higher than the coupling point), then T_{act} includes the time each node has to remain active without producing new data, in order to complete the transmission of the data buffered in its flash memory.
3) *Energy for receiving and buffering data* (in low-power on-chip memory) from d nodes prior to relaying it, given by $cTh\int_0^\infty \psi P_d(\psi)\,d\psi = cThE_d[\Psi]$ J. This energy is dominated by the receiver power requirements. Moreover, for a node that has an energy storage unit (battery) that can store hundreds of hours worth of operating energy, if the expected energy dissipation over a time interval, e.g. 24 hours, is matched with the amount of energy expected to be harvested from the environment within the same interval, it can be said that the sensor node achieves *energy neutrality* [2]. In practical IEEE 802.15.4 hardware, the average transceiver power under receive mode is virtually the same whether the node is actually receiving data or not. It is thus irrelevant to the receiver power whether the transmitting node used its entire transmission slot or not.
4) *Idle energy*, consumed when the data rate Ψ is smaller than the receiver coupling point $\frac{a}{n}$: $cTb\int_0^{\frac{a}{n}}\left(\frac{a}{n}-\psi\right)P_{d+1}(\psi)\,d\psi$ J. This energy corresponds to beaconing for synchronization and other run-time operations carried out during the transmit mode.
5) *Penalty energy*, consumed when the data rate Ψ is larger than the receiver coupling point $\frac{a}{n}$ and the data is buffered in high-power, typically off-chip, memory prior to transmission at the next available opportunity: $cTp\int_{\frac{a}{n}}^\infty\left(\psi-\frac{a}{n}\right)P_{d+1}(\psi)\,d\psi$ J.

Notice that, apart from the setup and convergence energy, the energy consumption for all the remaining components is affected by the total number of additional nodes (d) relaying their traffic via the current node. Example cluster-tree topologies providing instantiations for d and n in WSNs are given in Figure 1.

The *residual energy* of each node in a tier of the cluster-tree topology is defined as the difference between the produced (harvested) energy and the consumed energy over the harvesting time interval. It can be calculated for each sensor node by:

$$\begin{aligned} E_n &= TE[X] - k - cT \times \left[E_{d+1}[\Psi]\left(g+\frac{hd}{d+1}\right)\right. \\ &+ b\int_0^{\frac{a}{n}}\left(\frac{a}{n}-\psi\right)P_{d+1}(\psi)\,d\psi \\ &+ \left. p\int_{\frac{a}{n}}^\infty\left(\psi-\frac{a}{n}\right)P_{d+1}(\psi)\,d\psi\right]. \end{aligned} \tag{2}$$

Clearly, $E_n < 0$ corresponds to energy deficit (the expected energy produced by the harvesting process is lower than the expected consumption during the harvesting time interval), $E_n > 0$ corresponds to energy surplus, and $E_n = 0$ corresponds to energy neutrality. Notice that we used the relationship $\forall d > 0:\ E_{d+1}[\Psi] = \frac{d+1}{d}E_d[\Psi]$ in (2), since the expected transmission rate of each node increases linearly with respect to d in a uniformly-formed WSN. Adding and subtracting $cTp\int_0^{\frac{a}{n}}\left(\psi-\frac{a}{n}\right)P_{d+1}(\psi)\,d\psi$ in E_n, we get:

$$\begin{aligned} E_n &= TE[X] - k - cT \\ &\times \left[E_{d+1}[\Psi]\left(g+\frac{hd}{d+1}+p\right) - \frac{ap}{n}\right. \\ &+ \left.(b+p)\int_0^{\frac{a}{n}}\left(\frac{a}{n}-\psi\right)P_{d+1}(\psi)\,d\psi\right]. \end{aligned} \tag{3}$$

Evidently, the residual energy depends on the coupling point, $\frac{a}{n}$, as well as on the PDF of the data transmission rate per sensor node, $P_{d+1}(\psi)$. In the remainder of this section, we consider different cases for $P_{d+1}(\psi)$ to derive the residual energy under different statistical characterizations for the data transmission rate of each node and examine the conditions under which $E_n = 0$, i.e. energy neutrality is achieved.

A. Illustrative Case: Uniform Distribution

When no knowledge of the underlying statistics of the data generation process exists, one can assume that $P_{d+1}(\psi)$ is uniform over the interval $[0, 2(d+1)r]$:

$$P_{d+1,\text{U}}(\psi) = \begin{cases} \frac{1}{2(d+1)r}, & 0 \le \psi \le 2(d+1)r \\ 0, & \text{otherwise} \end{cases}. \tag{4}$$

The expected value of Ψ is $E_{d+1,\text{U}}[\Psi] = (d+1)r$ bps. If $\frac{a}{n} > 2(d+1)r$, then the coupling point is always overprovisioned; thus, each node will remain in idle state consuming energy for beaconing and radio on, which cannot lead to optimal energy efficiency. Thus, this case is not detailed here. For $\frac{a}{n} \le 2(d+1)r$, by using (4) in (3), we obtain:

$$\begin{aligned} E_{n,\text{U}} &= TE[X] - k - cT \\ &\times \left[(d+1)\,r\left(g+\frac{hd}{d+1}+p\right)\right. \\ &- \left.\frac{ap}{n} + \frac{a^2(b+p)}{4(d+1)rn^2}\right] \end{aligned} \tag{5}$$

If $p = 0$ then $E_{n,\text{U}}$ is monotonically increasing with n as there is no energy penalty for buffering data and the optimal number of nodes is (trivially) infinity. Moreover, if $b = p = 0$, then (3) is independent of n as this assumes no energy penalties. Given that these cases lead to trivial solutions, we do not investigate them further. For $b, p \neq 0$, the first derivative of $E_{n,\text{U}}$ to n is

$$\frac{dE_{n,\text{U}}}{dn} = cT\left[-\frac{ap}{n^2} + \frac{a^2(b+p)}{2(d+1)rn^3}\right]. \tag{6}$$

For $n \in (0,\infty)$, the number of nodes for which $\frac{dE_{n,\text{U}}}{dn} = 0$ is

$$n_{0,\text{U}} = \frac{a(b+p)}{2p(d+1)r} \tag{7}$$

As (7) is the only admissible solution of $\frac{dE_{n,\mathrm{U}}}{dn}=0$ and $E_{n,\mathrm{U}}$ is differentiable for $n\in(0,\infty)$, $n_{0,\mathrm{U}}$ is the global extremum or inflection point of $E_{n,\mathrm{U}}$. The second derivative of $E_{n,\mathrm{U}}$ is

$$\frac{d^2E_{n,\mathrm{U}}}{dn^2}=cT\left[\frac{2ap}{n^3}-\frac{3a^2\left(b+p\right)}{2(d+1)rn^4}\right]. \tag{8}$$

By evaluating $\frac{d^2E_{n,\mathrm{U}}}{dn^2}$ for $n_{0,\mathrm{U}}$ nodes, we obtain

$$\frac{d^2E_{n,\mathrm{U}}}{dn^2}\left(n_{0,\mathrm{U}}\right)=-\frac{8cT\left(d+1\right)^3p^4r^3}{a^2\left(b+p\right)^3}, \tag{9}$$

which is negative (since all the variables are positive). Thus, the maximum-possible residual energy for $n\in(0,\infty)$ is achieved under $n=n_{0,\mathrm{U}}$, and it is:

$$\begin{aligned}\max\left\{E_{n,\mathrm{U}}\right\} &= TE\left[X\right]-k-cT\left(d+1\right)r\\ &\times \left[g+\frac{hd}{d+1}+\frac{pb}{b+p}\right].\end{aligned} \tag{10}$$

The last equation demonstrates that the maximum residual energy obtained is zero, i.e. we achieve balanced consumption and production over the harvesting interval, under energy harvesting with rate given by:

$$\begin{aligned}\min\left\{E\left[X\right]\right\}_{\mathrm{U}} &= \frac{k}{T}+c\left(d+1\right)r\\ &\times \left(g+\frac{hd}{d+1}+\frac{pb}{b+p}\right).\end{aligned} \tag{11}$$

Hence, if the energy harvester of the node achieves *at least* $\min\left\{E\left[X\right]\right\}_{\mathrm{U}}$ W (averaged over the interval of T seconds), this suffices for perpetual (energy-neutral) operation of a WSN comprising $n_{0,\mathrm{U}}$ nodes at the same tier of the cluster-tree topology, with each node transmitting data with uniform rate between $[0,2(d+1)r]$ bps. The minimum power shown in (11) is obtained under the operational parameters: c, T, d, k, g, h, b, p (see Table I), $n_{0,\mathrm{U}}$ nodes and $E\left[\Psi\right]=r$. These parameters can be derived based on the utilized technology and the application specifics, as we shall show in Section IV and Section V.

The value derived for $n_{0,\mathrm{U}}$ by (7) is a real number. Within a practical setting, we have to select $\lfloor n_{0,\mathrm{U}}\rfloor$ (if greater than zero) or $\lceil n_{0,\mathrm{U}}\rceil$, depending on which one derives the highest residual energy value in (5). Since the minimum harvested power required for energy neutral operation and the number of nodes achieving it have a critical dependence on the data transmission rate and its characteristics, in the next subsection we derive this result under various characterizations for Ψ that are encountered often in practical data gathering applications based on WSNs. Similarly as for this subsection, once the result for the continuous case is derived, we can immediately derive the discrete-case equivalent by converting the optimal value of n to the nearest integer that provides for the highest residual energy.

B. Minimum Harvested Power Required for Data Transmission Rate Modeled by the Pareto, Exponential and Half-Gaussian Distributions

We can now generalize the previous calculation to other distributions expressing commonly observed data transmission rates in practical applications. We consider three additional PDFs for Ψ that have been used to model the marginal statistics of many real-world data transmission applications. We provide the obtained analytic results in this subsection. Since the proofs follow the same process as for the uniform distribution, they are given in Appendix I in summary form. For each distribution, we couple its parameters to the average transmission rate of the uniform distribution, $(d+1)\,r$, such that it is possible to achieve the same average data transmission rate over any uniformly-formed WSN cluster-tree topology where each node relays data from d additional nodes. This facilitates comparisons of the minimum power-harvesting capability required under different characterizations for the data rate.

1) Pareto distribution and fixed data rate: This distribution has been used, amongst others, to model the marginal data size distribution of TCP sessions that contain substantial number of small files and a few very large ones [39], [40]. Consider $P_{d+1,\mathrm{P}}\left(\psi\right)$ as the Pareto distribution with scale v and shape $\alpha\geq 2$ $(\alpha\in\mathbb{N})$,

$$P_{d+1,\mathrm{P}}\left(\psi\right)=\begin{cases}\alpha\frac{v^{\alpha}}{\psi^{\alpha+1}}, & \psi\geq v\\ 0, & \text{otherwise}\end{cases}. \tag{12}$$

The expected value of Ψ is $E_{d+1,\mathrm{P}}\left[\Psi\right]=\frac{\alpha v}{\alpha-1}$ bps. Thus, if we set

$$v=\frac{\alpha-1}{\alpha}(d+1)r \tag{13}$$

we obtain $E_{d+1,\mathrm{P}}\left[\Psi\right]=(d+1)r$ bps, i.e. we match the expected data transmission rate to that of the Uniform distribution. For the case of the Pareto distribution, if $\frac{a}{n}<v$, this corresponds to each node always attempting to transmit more data than what is allowed by the coupling point. This case will always incur energy penalty for buffering the residual bits beyond the coupling point and it is thus not investigated further as it will not lead to an optimal solution. For $\frac{a}{n}\geq v$, we obtain via (3):

$$\begin{aligned}E_{n,\mathrm{P}} &= TE\left[X\right]-k-cT\left[\alpha v\frac{g+\frac{hd}{d+1}+p}{\alpha-1}\right.\\ &+ \left.\frac{ab}{n}+(b+p)\left(\frac{v^{\alpha}n^{\alpha-1}}{a^{\alpha-1}\left(\alpha-1\right)}-\frac{\alpha v}{\alpha-1}\right)\right].\end{aligned} \tag{14}$$

Since $b+p\neq 0$, the number of nodes that derives the *minimum power from the harvester* to allow for energy neutrality under data transmission rate following the Pareto distribution of (12) is

$$n_{0,\mathrm{P}}=\frac{a}{v}\left(\frac{b}{b+p}\right)^{\frac{1}{\alpha}} \tag{15}$$

The minimum harvested power required under (15) is:

$$\min\left\{E\left[X\right]\right\}_{\mathrm{P}} = \frac{k}{T}+c\left(d+1\right)r\left[g+\frac{hd}{d+1} - b+b^{\frac{\alpha-1}{\alpha}}\left(b+p\right)^{\frac{1}{\alpha}}\right] \quad (16)$$

A special case for this distribution is when $\alpha = r$, which leads to $v = (d+1)(r-1)$ from (13). Then, the expected value of Ψ is $E_{d+1,\mathrm{F}}[\Psi] = (d+1)r$ bps and its standard deviation is $\sigma_{d+1,\mathrm{F}}[\psi] = (d+1)\sqrt{\frac{r}{r-2}}$. For $r > 150$ bps, the standard deviation is less than 0.7% of the mean value. Thus, in practice this case corresponds to transmission with fixed rate of $(d+1)r$ bps. This scenario occurs in WSNs capturing and transmitting data with fixed rate during their active time, e.g. in periodic temperature or humidity measurements gathered by WSNs [33], [32]. For this case, the number of nodes leading to the minimum harvested power is:

$$n_{0,\mathrm{F}} = \frac{a}{(d+1)(r-1)}\left(\frac{b}{b+p}\right)^{\frac{1}{r}} \quad (17)$$

For the vast majority of values for a, d and r used in practical WSN applications, $n_{0,\mathrm{F}}$ is equal to either $\left\lfloor\frac{a}{(d+1)r}\right\rfloor$ (if greater than zero) or $\left\lceil\frac{a}{(d+1)r}\right\rceil$ when converted into an integer. This agrees with the intuitive answer for balancing fixed-rate transmission of $(d+1)r$ bps to consumption rate of a bps. The minimum harvested power required under (17) is:

$$\min\left\{E\left[X\right]\right\}_{\mathrm{F}} = \frac{k}{T}+c\left(d+1\right)r\left[g+\frac{hd}{d+1} - b+b^{\frac{r-1}{r}}\left(b+p\right)^{\frac{1}{r}}\right] \quad (18)$$

2) Exponential distribution : The marginal statistics of MPEG video traffic have often been modeled as exponentially decaying [41]. Consider $P_{d+1,\mathrm{E}}(\psi)$ as the Exponential distribution with rate parameter $\frac{1}{(d+1)r}$

$$P_{d+1,\mathrm{E}}(\psi) = \frac{1}{(d+1)r}\exp\left(-\frac{1}{(d+1)r}\psi\right) \quad (19)$$

for $\psi \geq 0$. In this case, the expected value of Ψ is $E_{d+1,\mathrm{E}}[\Psi] = (d+1)r$ bps. Via (3), we obtain

$$E_{n,\mathrm{E}} = TE\left[X\right]-k-cT\left[\left(d+1\right)r\left(g+\frac{hd}{d+1}+p\right) + \frac{ab}{n}+\left(d+1\right)r\left(b+p\right) \times \left[\exp\left(-\frac{a}{n\left(d+1\right)r}\right)-1\right]\right]. \quad (20)$$

Assuming $b \neq 0$, the value of

$$n_{0,\mathrm{E}} = \frac{a}{(d+1)r\ln\left(\frac{b+p}{b}\right)} \quad (21)$$

is the number of nodes that requires the *minimum power from the harvester* to allow for the system to maintain energy neutrality under data transmission following the exponential distribution of (19). The minimum harvested power required under this number of nodes is:

$$\min\left\{E\left[X\right]\right\}_{\mathrm{E}} = \frac{k}{T}+c\left(d+1\right)r \times \left[b\ln\left(\frac{b+p}{b}\right)+g+\frac{hd}{d+1}\right]. \quad (22)$$

3) Half-Gaussian distribution : We conclude this part by considering $P_{d+1,\mathrm{H}}(\psi)$ as the Half-Gaussian distribution with mean $E_{d+1,\mathrm{H}}[\Psi] = (d+1)r$

$$P_{d+1,\mathrm{H}}(\psi) = \begin{cases} 0, & \psi < 0 \\ \frac{2}{\pi(d+1)r}\exp\left(-\frac{\psi^2}{\pi(d+1)^2r^2}\right), & \psi w \geq 0 \end{cases} \quad (23)$$

This distribution has been widely used in data gathering problems in science and engineering when the modeled data has non-negativity constraints. Some recent examples include the statistical characterization of motion vector data rates in Wyner-Ziv video coding algorithms suitable for WSNs [30], or the statistical characterization of sample amplitudes captured by an image sensor [36], [42]. Via (3), we obtain

$$E_{n,\mathrm{H}} = TE\left[X\right]-k-cT\left[(d+1)r\left(g+\frac{hd}{d+1}+p\right) - \frac{ap}{n}+(b+p)\left[(d+1)r\left[\exp\left(-\frac{a^2}{\pi(d+1)^2r^2n^2}\right) - 1\right]+\frac{a}{n}\mathrm{erf}\left(\frac{a}{\sqrt{\pi}(d+1)rn}\right)\right]\right], \quad (24)$$

with $\mathrm{erf}(\cdot)$ the error function that can be approximated by its Taylor series expansion. Under $b \neq 0$ and $p \neq 0$, the number of nodes that leads to the *minimum power required from the harvester* in order for the system to maintain energy neutrality under data transmission rate (per node) characterized by $P_{d+1,\mathrm{H}}(\psi)$ is

$$n_{0,\mathrm{H}} = \frac{a}{\sqrt{\pi}(d+1)r\mathrm{erf}^{-1}\left(\frac{p}{b+p}\right)}, \quad (25)$$

with $\mathrm{erf}^{-1}(\cdot)$ the inverse error function, which can be approximated by its series expansion. The minimum harvested power required under (25) is:

$$\min\left\{E\left[X\right]\right\}_{\mathrm{H}} = \frac{k}{T}+c\left(d+1\right)r\left[g+\frac{hd}{d+1}-b + (b+p)\exp\left(-\left[\mathrm{erf}^{-1}\left(\frac{p}{b+p}\right)\right]^2\right)\right]. \quad (26)$$

C. Considering the Relay Case under a Multi-hop Topology

When expanding this analysis to multi-layer topologies, one can consider a variety of settings as illustrated in Figure 1. Here we distinguish three cases, which are discussed in the following.

Firstly, when each node shapes its overall data transmission rate (which includes their own data and the data received from other nodes) according to one of the distributions considered in the previous subsection, the results will follow what was discussed before.

Secondly, when each node simply aggregates the received data with its own data within each TFDMA transmission opportunity, thereby leading to a new data production rate PDF, one must consider this new distribution in the proposed analytic framework. Such distributions will be the convolutions of identical Uniform, Pareto, Exponential and Half-Gaussian distributions. For small values of d, e.g. $1 \leq d \leq 3$, the results can be derived following the steps given in Subsection III-A and III-B if functions

$$P_{d+1,\mathrm{Z}}(\psi) = \underbrace{P_{\mathrm{Z}}(\psi) \star \ldots \star P_{\mathrm{Z}}(\psi)}_{d \text{ times}},\ \mathrm{Z} \in \{\mathrm{U},\mathrm{P},\mathrm{E},\mathrm{H}\}$$

are derived. Given that $P_{d+1,\mathrm{Z}}(\psi)$ and the $\int_0^{\frac{a}{n}} \left(\frac{a}{n} - \psi\right) P_{d+1,\mathrm{Z}}(\psi)\, d\psi$ term of (3) can be computed with the help of a numerical package (e.g. Mathematica or Matlab Symbolic) and that these will vary for each value of d, we do not expand on these cases further.

Finally, when $d \geq 4$, according to the central limit theorem [43], all data rate PDFs will begin to converge to a Gaussian distribution. By considering $P_{d+1,\mathrm{N}}(\psi)$ as the Gaussian distribution with mean $E_{d+1,\mathrm{N}}[\Psi] = (d+1)\, r$ and standard deviation σ

$$P_{d+1,\mathrm{N}}(\psi) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\psi - (d+1)\, r)^2}{2\sigma^2}\right), \quad (27)$$

via (3), we obtain (see Appendix I for details on this derivation):

$$\begin{aligned} E_{n,\mathrm{N}} &= TE[X] - k - cT\Bigg[(d+1)\, r\left(g + \frac{hd}{d+1} + p\right) \\ &- \frac{ap}{n} + (b+p)\left[\frac{n(d+1)\, r - a}{2n}\right. \\ &\times \left.\left[\mathrm{erf}\left(\frac{(d+1)\, r - \frac{a}{n}}{\sqrt{2}\sigma}\right) - \mathrm{erf}\left(\frac{(d+1)\, r}{\sqrt{2}\sigma}\right)\right]\right. \quad (28) \\ &+ \frac{\sigma}{\sqrt{2\pi}}\left[\exp\left(\frac{-\left((d+1)\, r - \frac{a}{n}\right)^2}{2\sigma^2}\right)\right. \\ &- \left.\left.\left.\exp\left(\frac{-\left((d+1)\, r\right)^2}{2\sigma^2}\right)\right]\right]\Bigg]. \end{aligned}$$

The residual energy of (28) has a global maximum for $n \in (0, \infty)$, i.e. a global minimum in the required harvesting power $E[\chi]$, if: *(i)* $b \neq 0$ or $p \neq 0$ and *(ii)* the following condition is satisfied:

$$\left|\mathrm{erf}\left(\frac{(d+1)\, r}{\sqrt{2}\sigma}\right) - \frac{2p}{b+p}\right| < 1. \quad (29)$$

Then, the number of nodes that leads to the *minimum power* required in order for the system to maintain energy neutrality under data transmission (per node) following $P_{d+1,\mathrm{N}}(\psi)$ is

$$n_{0,\mathrm{N}} = \frac{a}{(d+1)\, r - \sqrt{2}\sigma c_{\mathrm{N}}}, \quad (30)$$

with $(d+1)\, r \neq \sqrt{2}\sigma c_{\mathrm{N}}$,

$$c_{\mathrm{N}} = \mathrm{erf}^{-1}\left(\mathrm{erf}\left(\frac{(d+1)\, r}{\sqrt{2}\sigma}\right) - \frac{2p}{b+p}\right). \quad (31)$$

The minimum harvested power required under (30) is:

$$\begin{aligned} \min\{E[X]\}_{\mathrm{N}} &= \frac{k}{T} + c(d+1)\, r\left[g + \frac{hd}{d+1}\right. \\ &+ \left.\frac{\sigma(b+p)}{\sqrt{2\pi}(d+1)\, r}\left[\exp\left(-c_{\mathrm{N}}^2\right)\right.\right. \quad (32) \\ &- \left.\left.\exp\left(-\frac{((d+1)\, r)^2}{2\sigma^2}\right)\right]\right]. \end{aligned}$$

D. Comparison of Minimum Required Harvested Power under the Same Expected Data Transmission Rate per Node

We can now address two interesting questions for a uniformly-formed WSN with application parameters given in Table I: *What is difference of the minimum-required harvested power under the various data transmission rate distributions studied in Section III? Can we rank these distributions with respect to their incurred energy efficiency?* We use the Uniform distribution as the basis of our comparisons and include all distributions except of the Gaussian PDF that emerges as the convergence of any data rate PDF when $d \geq 4$. To match all cases, we can set:

$$\begin{aligned} n_{0,\mathrm{P}} &\equiv s_{\mathrm{PU}} \times n_{0,\mathrm{U}} \\ n_{0,\mathrm{F}} &\equiv s_{\mathrm{FU}} \times n_{0,\mathrm{U}} \\ n_{0,\mathrm{E}} &\equiv s_{\mathrm{EU}} \times n_{0,\mathrm{U}} \\ n_{0,\mathrm{H}} &\equiv s_{\mathrm{HU}} \times n_{0,\mathrm{U}} \end{aligned} \quad (33)$$

and

$$\begin{aligned} E_{d+1,\mathrm{U}}[\Psi] &\equiv E_{d+1,\mathrm{P}}[\Psi] \equiv E_{d+1,\mathrm{F}}[\Psi] \quad (34) \\ &\equiv E_{d+1,\mathrm{E}}[\Psi] \equiv E_{d+1,\mathrm{H}}[\Psi] \equiv (d+1)r, \end{aligned}$$

where factors s_{PU}, s_{FU}, s_{EU}, s_{HU} in (33) express the ratio of the number of nodes[2] corresponding to each transmission rate PDF and the same expected transmission rate is assumed per node. Moreover, given that our energy expressions depend on both the beaconing power (during idle transmission intervals) and the penalty power (when writing residual bits to the flash memory), without loss of generality we define their ratio as

$$c_{bp} \equiv \frac{b}{p}. \quad (35)$$

[2]which can be calculated via the definition of $n_{0,\mathrm{P}}$, $n_{0,\mathrm{E}}$, $n_{0,\mathrm{H}}$ and $n_{0,\mathrm{U}}$ as shown in Appendix II.

We can then derive the conditions that make a particular PDF preferable in terms of minimizing the required power from the harvesting mechanism, parametrically to the operational settings of the application. The remainder of this section addresses this problem.

Proposition 1. (Pareto vs. Uniform PDF). *The difference between the minimum harvested power required under a pareto-distributed and a uniformly-distributed transmission rate per node when their average WSN data rates are matched via* (33) *and* (34) *is:*

$$\begin{aligned} & \min\left\{E\left[X\right]\right\}_{\mathrm{P}} - \min\left\{E\left[X\right]\right\}_{\mathrm{U}} \\ = \;& c\left(d+1\right) rb\left[\left(\frac{2(d+1)r}{s_{\mathrm{PU}} v} - 1\right)\frac{p}{b+p} - 1\right]. \end{aligned} \quad (36)$$

Proof: See Appendix II. ∎

Corollary 1. *For* $\alpha \geq 2$ *and* $\alpha \in \mathsf{N}$:

$$\min\left\{E\left[X\right]\right\}_{\mathrm{P}} < \min\left\{E\left[X\right]\right\}_{\mathrm{U}}$$

iff

$$2 - \sqrt[\alpha]{\frac{c_{bp}+1}{c_{bp}}} - \frac{c_{bp}}{c_{bp}+1} > 0. \quad (37)$$

Proof: If

$$s_{\mathrm{PU}} > \frac{2p\alpha}{\left(\alpha-1\right)\left(b+2p\right)},$$

then Proposition 1 shows that the Pareto distribution requires less harvested power in comparison to the Uniform distribution (under the settings leading to the minimum requirement for both). Based on (60) of Appendix II and (35), the last inequality becomes

$$\frac{c_{bp}^{\frac{1}{\alpha}}}{\left(c_{bp}+1\right)^{\frac{\alpha+1}{\alpha}}} - \frac{1}{\left(c_{bp}+2\right)} > 0. \quad (38)$$

Knowing that $c_{bp} > 0$, the last condition can be simplified to (37). ∎

When enforcing the equality condition in (37), then data transmission rates characterized by $\alpha-$Pareto and Uniform PDF (under the same mean) will lead to the same harvested energy requirements. Since there is no closed-form solution for (37) under the equality condition, the numerical evaluation of c_{bp} under this condition is provided in Figure 3 for $\alpha \in \{2, \ldots, 26\}$. The results demonstrate that, for the vast majority of α values, the two distributions attain the same minimum harvested energy only under very small values for c_{bp}, i.e. when the beaconing power becomes negligible in comparison to the penalty power to write residual data to the flash memory. Thus, for the vast majority of cases, data transmission rates characterized by the Pareto distribution are expected to require less harvested energy than those characterized by the Uniform distribution (under the same mean and under the number of nodes leading to the minimum harvesting requirements). Finally, by comparing (16) with $\alpha < r$ and (18), it is straightforward to show that $\min\left\{E\left[X\right]\right\}_{\mathrm{F}} < \min\left\{E\left[X\right]\right\}_{\mathrm{P}}$.

Proposition 2. (Exponential vs. Uniform PDF). *The difference between the minimum harvested power required under an exponentially-distributed and a uniformly-distributed transmission rate per node when their average WSN data transmission rates are matched via* (33) *and* (34) *is:*

$$\min\left\{E\left[X\right]\right\}_{E} - \min\left\{E\left[X\right]\right\}_{U} = c\left(d+1\right) r \frac{pb}{b+p}\left(\frac{2}{s_{EU}} - 1\right). \quad (39)$$

Proof: See Appendix II. ∎

Corollary 2. $\min\left\{E\left[X\right]\right\}_{\mathrm{E}} > \min\left\{E\left[X\right]\right\}_{\mathrm{U}}$.

Proof: Proposition 2 shows that, under the settings leading to the minimum power requirement, the Exponential distribution requires more harvesting power in comparison to the Uniform distribution if $s_{\mathrm{EU}} < 2$. Based on (62) of Appendix II, this inequality becomes:

$$\ln\left(\frac{c_{bp}+1}{c_{bp}}\right) - \frac{1}{c_{bp}+1} > 0. \quad (40)$$

By investigating the extrema and monotonicity via the first derivative of (40), it is straightforward to verify that (40) holds for all $c_{bp} > 0$. ∎

Proposition 3. (Half-Gaussian vs. Uniform PDF). *The difference between the minimum harvested power required under a half-Gaussian-distributed and a uniformly-distributed transmission rate per node when their average WSN data rates are matched via* (33) *and* (34) *is:*

$$\begin{aligned} & \min\left\{E\left[X\right]\right\}_{\mathrm{H}} - \min\left\{E\left[X\right]\right\}_{\mathrm{U}} \\ = \;& c\left(d+1\right) r\left[\frac{p^2}{b+p} - \left(b+p\right)\right. \\ \times \;& \left.\left[1 - \exp\left(-\frac{4p^2}{s_{\mathrm{HU}}^2 \pi (b+p)^2}\right)\right]\right]. \end{aligned} \quad (41)$$

Proof: See Appendix II. ∎

Corollary 3. $\min\left\{E\left[X\right]\right\}_{\mathrm{H}} > \min\left\{E\left[X\right]\right\}_{\mathrm{U}}$.

Proof: Proposition 3 shows that the Half-Gaussian distribution requires more harvesting power in comparison to the Uniform distribution when

$$s_{\mathrm{HU}} > \frac{2p}{\sqrt{\pi}\left(b+p\right)\sqrt{\ln\left(\frac{(b+p)^2}{b^2+2bp}\right)}}, \quad (42)$$

which, based on (64) of Appendix II, becomes:

$$\sqrt{\ln\left(\frac{\left(c_{bp}+1\right)^2}{c_{bp}^2+2c_{bp}}\right)} - \mathrm{erf}^{-1}\left(\frac{1}{c_{bp}+1}\right) > 0. \quad (43)$$

By differentiating the last expression we can verify that (43) holds for all $c_{bp} > 0$. ∎

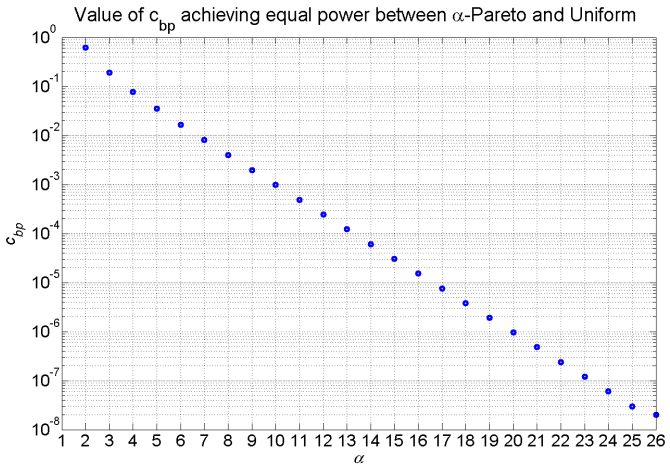

Figure 3: Values obtained for c_{bp} [ratio between beaconing and penalty power per bit defined in (35)] by setting (37) under the equality condition.

Given that both the Exponential and the Half-Gaussian distribution require more power than the Uniform distribution at the number of nodes providing for the minimum harvesting power per PDF, it is thus beneficial to establish the difference between these two PDFs.

Proposition 4. (Half-Gaussian vs, Exponential PDF). *The difference between the minimum harvested power required under a half-Gaussian−distributed and an exponentially-distributed transmission rate per node when their average WSN data rates are matched via* (33) *and* (34) *is:*

$$\begin{aligned} & \min\left\{E\left[X\right]\right\}_{\mathrm{H}} - \min\left\{E\left[X\right]\right\}_{\mathrm{E}} \\ = \;& c\left(d+1\right)r\left[-b - b\ln\left(\frac{b+p}{b}\right) + \left(b+p\right)\right. \\ \times \;& \left.\exp\left(-\frac{1}{s_{\mathrm{HE}}^2\pi}\left[\ln\left(\frac{b+p}{b}\right)\right]^2\right)\right]. \end{aligned} \quad (44)$$

with $s_{\mathrm{HE}} \equiv \frac{n_{0,\mathrm{H}}}{n_{0,\mathrm{E}}}$ the ratio between the number of nodes providing for the minimum power (per PDF).

Proof: See Appendix II. ■

Corollary 4. $\min\left\{E\left[X\right]\right\}_{\mathrm{H}} < \min\left\{E\left[X\right]\right\}_{\mathrm{E}}$.

Proof: From (44) we establish that if

$$s_{\mathrm{HE}} < \frac{\ln\left(\frac{b+p}{b}\right)}{\sqrt{\pi\ln\left(\frac{b+p}{b+b\ln\left(\frac{b+p}{b}\right)}\right)}}$$

then the Half-Gaussian distribution will require less harvested power than the Exponential distribution (at the settings leading to the minimum power). Via (66) of Appendix II, the last inequality becomes:

$$\mathrm{erf}^{-1}\left(\frac{1}{c_{bp}+1}\right) - \sqrt{\ln\left(\frac{1+\frac{1}{c_{bp}}}{\ln\left(1+\frac{1}{c_{bp}}\right)+1}\right)} > 0, \quad (45)$$

By differentiating the last expression we can verify that (45) holds for all $c_{bp} > 0$. ■

Proposition 5. *The ranking of the different transmission rate PDFs for the minimum-required harvested power to maintain energy neutrality is (from lowest to highest requirement):*

$$\textbf{Fixed Rate} \prec \textbf{Pareto} \overset{\textit{iff}\ (37)}{\prec} \textbf{Uniform} \prec \textbf{Half-Gaussian} \prec \textbf{Exponential}$$

Proof: The proof follows from the combination of Corollaries 1−4. ■

Beyond the ranking of the different distributions, Propositions 1−4 show that the decrease in the required harvested power offered by each case is proportional to the average data transmission rate, $(d+1)\,r$, with a proportionality factor that depends on the energy penalties of the system, b and p but not on the energy rates for transmitting and receiving each bit, g and h.

E. Discussion

The results of this section can be used in practical applications to assess the impact in the required harvesting power and harvesting time when the statistics of the transmission data rate follow a certain PDF and the network parameters are fixed. Conversely, if a particular technology, such as an array of photovoltaic cells or a piezoelectric circuit, has been shown to provide for certain power generation capability per sensor, under the knowledge of the system and data gathering parameters and the duty cycle of the network, one can establish the appropriate network parameters per tier. Finally, for given network and system parameters, one can assess the achievable data transmission rates such that the WSN infrastructure remains energy neutral.

Thus, as shown in Figure 4, our analytic results allow for the linkage of network, data gathering and energy and system parameters within uniformly-formed cluster-tree WSN topologies. Hence, our analysis can be used for early-stage exploration of the capabilities of a particular WSN infrastructure in conjunction with the data gathering requirements of a particular application, prior to embarking in cumbersome development and testing in the field. Finally, Propositions 1-5 provide the ranking of various data transmission PDFs under the same mean rate and the number of nodes leading to the minimum energy requirements per PDF. This can be used to control the way data processing and compression algorithms produce data in WSN applications aiming for energy-neutral operation with minimum harvesting requirements.

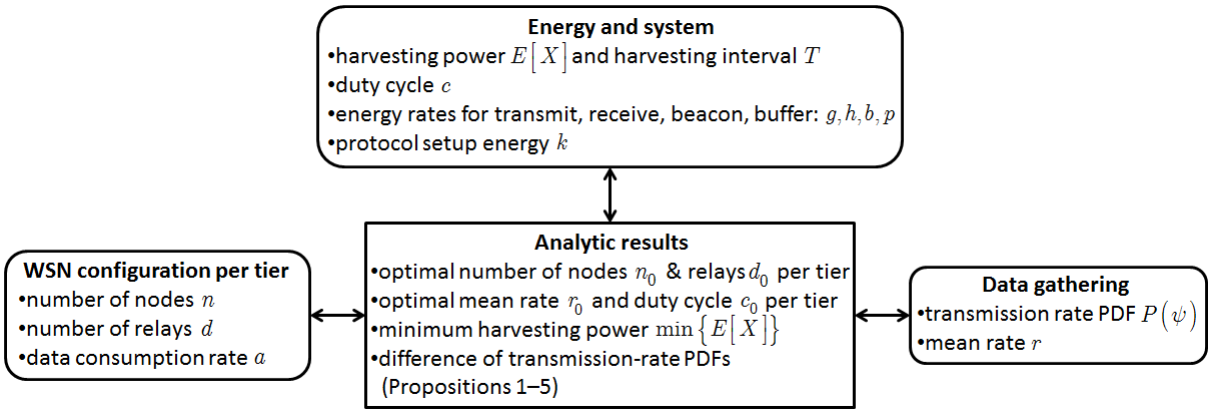

Figure 4: Conceptual illustration of the linkage between: network, system and data gathering via the proposed analysis. When parameters from two out of three domains are provided, our analytic framework can tune the parameters of the third. The symbol definitions are provided in Table I.

IV. Evaluation of the Analytic Results

A. WSN and System Settings

We consider a typical WSN setup comprising of several TelosB nodes (using the IEEE 802.15.4 standard with the CC2420 transceiver) running the low-power Contiki 2.6 operating system. All nodes use our recently-proposed TFDMA protocol (which is available open source [6]) to communicate with the receiver (or base station) existing on the same channel, following topologies such as the ones shown in Figure 1. The TFDMA protocol uses biology-inspired self-maintaining algorithms in wireless sensor nodes and achieves near optimum TDMA characteristics in a decentralized manner and over multiple channels (frequencies). This is achieved by extending the concept of collaborative reactive listening in order to balance the number of nodes in all available channels of IEEE 802.15.4 [6]. Consequently, TFDMA can be deployed at the application layer with very low complexity and provides for balanced multichannel coordination of multiple nodes. We opted for its usage as it allows for quick convergence to the steady state and permits collision-free multichannel communications once steady state has been established. It also provides for comparable or superior bandwidth utilization to channel-hopping approaches like TSMP and EM-MAC [7]. However, similar results can be obtained with any other protocol offering collision-free single- or multi-channel communications under a cluster-tree topology, such as TSMP [18], IEEE 802.15.4 GTS [17], [9], etc.

For the utilized TFDMA and active time $T_{\text{act}} = 400$ s, convergence has been shown to occur in less than 1.3% of T_{act} (3−5s) and, on average, the energy dissipation for convergence has been found to be $k = 165.6$ mJ in our setup. Concerning the communications side, following the default TFDMA setup, for all our measurements we set the packet size to 114 bytes, the DESYNC interval to 1s and the DESYNC constant to 0.95 [38]. Each node transmits 1-byte beacon packets every 8 ms when it is not transmitting data packets during its transmission slot to maintain connectivity and synchronization. Finally, since the TFDMA protocol ensures no collisions occur during the steady-state active mode, we are utilizing the very-low complexity NullMAC and NullRDC options of Contiki OS (see footnote 1), which lead to maximum data consumption rate at the application layer of $a = 144$ kbps.

Concerning the data gathering itself, we created artificial data via a custom Matlab function that, starting from the `rand()` function, generates data with Uniform, Pareto, Exponential and Half-Gaussian distributions (considered in Section III) via rejection sampling [44], with mean transmission rate equal to $r = 24$ kbps. The data is copied onto each node and it is read from its external flash memory during the steady-state active mode. This ensures that: *(i)* we match the different PDFs under consideration and *(ii)* the energy to retrieve this data from the flash memory replaces the sensing and processing energy that would have been dissipated if the data had come from an actual sensing process.

Under these operational settings, our energy measurement setup comprises a high-tolerance 1 Ohm resistor placed in series with each TelosB node. By measuring the current consumption at the resistor and knowing that each node operates at 3 Volt, we derive the real-time energy consumption (see Figure 2 for examples). The utilized time resolution for the power measurements was 12.5 KHz using a Tektronix MDO4104-6 oscilloscope. Under this setup, we also measured the different energy rates of Table I by enabling transmission, listening, writing to flash memory and beaconing during the idle state to maintain synchronization. They were found to be: $g = 2.29262 \times 10^{-7}$ J/b, $h = 2.92309 \times 10^{-6}$ J/b, $p = 3.89392 \times 10^{-7}$ J/b and $b = 2.17324 \times 10^{-7}$ J/b.

B. Model Validation

We consider the fully-connected (single-hop) topology shown in Figure 1(a). Each node sends only its own data, which corresponds to $d = 0$ (no relay) and we test with various values for n (total number of nodes within a single channel). We present the results in Figure 5. For each data production

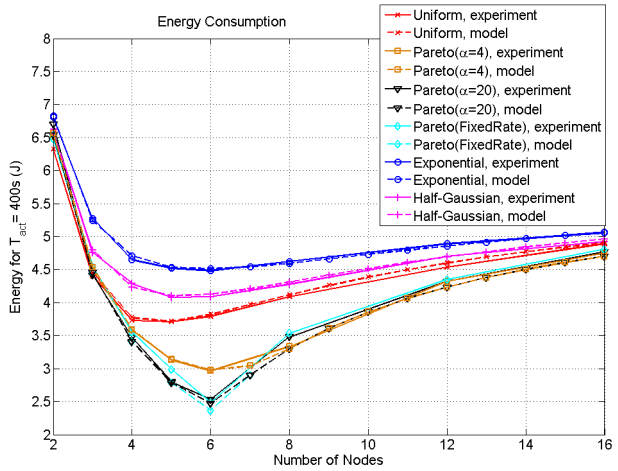

Figure 5: Energy consumption per node under different data transmission PDFs. The experiments correspond to $T_{\text{act}} = 400$ s, $k = 165.6$ mJ and $d = 0$.

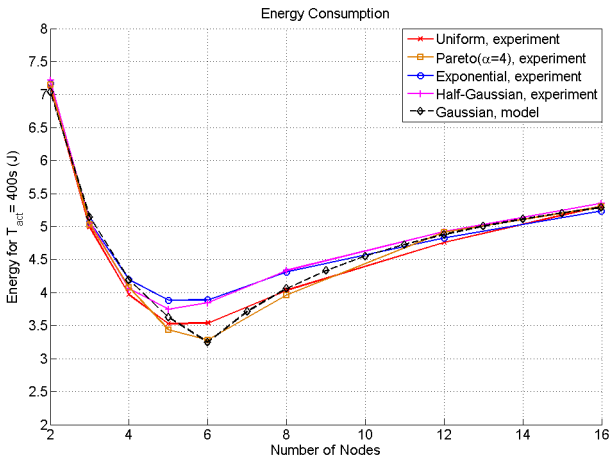

Figure 6: Energy consumption per node with different data production PDFs, $d = 4$ and $r = 4.8$ kbps; each node aggregates the received data with its own data within each transmission opportunity.

PDF, the "theoretical" results have been produced via (5), (14), (20), (24) by considering only the energy dissipation part; the energy harvesting part, $TE[X]$, is discussed separately in Section V. Evidently, for the vast majority of cases, the theoretical and experimental results are in agreement, with the maximum difference between them limited to within 0.237 J, i.e. maximum error of 7.2%. We have observed the same level of accuracy under a variety of different rates r and for more than one tier (under multi-channel TFDMA operation) but omit these repetitive experiments for brevity of exposition.

The results of Figure 5 demonstrate that each transmission rate distribution incurs different energy consumption and the ranking of the PDFs is precisely as predicted by Proposition 5. Thus, even under the same mean transmission rate, the manner the data traffic is shaped in a WSN plays an important role in the system's requirements for energy neutrality. Moreover, the results show that, depending on the transmission rate PDF, the number of nodes where the minimum energy consumption occurs, i.e. $n_{0,\text{U}}$, $n_{0,\text{P}}$, $n_{0,\text{F}}$, $n_{0,\text{E}}$ and $n_{0,\text{H}}$, may differ. This is reflected by Propositions 1−4. The accuracy of these analytic estimations is quantified in Table II in comparison to the experimentally-obtained values for the difference in the minimum energy consumption. Following the balancing conditions of (33) and (34), we also report the values for s_{PU}, s_{FU}, s_{EU}, s_{HU} and s_{HE} parameters corresponding to this experiment. The table demonstrates that the theoretically-calculated difference via Propositions 1-4 is very close to the experimentally-obtained values, as the average percentile error is only 4.20%.

Finally, with respect to a multi-hop scenario, Figure 6 presents the results under $d = 4$ and $r = 4.8$ kbps (with all other settings being the same). As expected, all experimental curves converge towards the model results of the Gaussian distribution. This convergence improves further when higher values of d are considered.

V. Applications

A. Maximizing Active Time Interval under Given Energy Harvesting Capability

The first application concerns WSNs where every sensor is equipped with certain energy harvesting technology, e.g. a piezoelectric unit or solar panel. Under given data transmission rate PDF with mean r (fitted to the experimentally-observed data rate histogram), the aim is to derive the optimal number of sensors (n_0) and the maximum duty cycle (c_0) so that the network performs data gathering and transmission for the maximum amount of active time under energy neutrality. Such a scenario occurs in energy management systems for WSNs and indoor or outdoor monitoring systems that are expected to be active for the maximum amount of time possible. [14], [2], [33].

We utilize the expected power produced by harvesting, $E[\chi]$, for photovoltaic and piezoelectric technologies under stable indoor conditions (as reported in the relevant literature) and consider a single-tier network topology [Figure 1(a)]. The goal is to match the expected energy harvested within T s with the expected energy consumption within T_{act} s and report what are the highest-possible values for the duty cycle c and T_{act}.

For the data rate PDFs considered in this paper and the system settings of Section IV, we present the obtained results in Table III under harvesting time $T = 21600$ s (6 hr) and $r = 3000$ bps. Under the given value for $TE[X]$, the values for n_0 were derived using: (7), (15), (17), (21) and (25); c_0 (and T_{act}) were derived solving: (11), (16), (18), (22) and (26) for c. Evidently, depending on the technology used and the utilized PDF, the results can vary, i.e. from energy neutrality achieved with $c_0 = 0.112$ and $T_{\text{act}} = 2424$ s for exponentially-distributed data gathering and transmission rate, to $c_0 = 0.279$ and $T_{\text{act}} = 6038$ s for Pareto-distributed (or fixed) data gathering and transmission. In an application that

Table II: Differences in the minimum harvested energy required amongst the considered PDFs under the settings of Figure 5.

Proposition	Theoretical difference (J)	Experimental difference (J)	Percentile error (%)
1, (Pareto $\alpha = 4$ vs. Uniform), $s_{\text{PU}} = 1.3$	-0.729	-0.742	1.81
1, (Pareto $\alpha = 20$ vs. Uniform), $s_{\text{PU}} = 1.3$	-1.229	-1.176	4.32
1, [Fixed-rate (Pareto $\alpha = r$) vs. Uniform], $s_{\text{FU}} = 1.3$	-1.339	-1.223	8.63
2 (Exponential vs. Uniform), $s_{\text{EU}} = 1.2$	+0.803	+0.775	3.45
3 (Half-Gaussian vs. Uniform), $s_{\text{HU}} = 1.1$	+0.394	+0.372	5.54
4 (Half-Gaussian vs. Exponential), $s_{\text{HE}} = 0.9$	-0.409	-0.403	1.44

Table III: Maximum active time T_{act} and duty cycle c (in parentheses) required by two different harvesting technologies under harvesting time $T = 21600$ s (6 hr) and with mean data rate $r = 3000$ bps.

Production Rate PDF	16 cm^2 Solar Panel $E[X] = 160\mu$W [35]	Piezoelectric Unit $E[X] = 200\mu$W [3]
Uniform, $n_0 = 37$	2975 s (0.138)	3755 s (0.174)
Pareto ($\alpha = 4$), $n_0 = 50$	3754 s (0.173)	4729 s (0.219)
Pareto ($\alpha = 20$), $n_0 = 48$	4556 s (0.211)	5753 s (0.266)
Fixed rate, $n_0 = 48$	4782 s (0.221)	6038 s (0.279)
Exponential, $n_0 = 47$	2424 s (0.112)	3061 s (0.142)
Half-Gaussian, $n_0 = 42$	2677 s (0.124)	3380 s (0.156)

requires $c_0 = 1$ (continuous monitoring), based on the results of Table III we can calculate how many independent sets of n_0 nodes to install so that continuous monitoring is achieved under energy neutrality. For example, since $c_0 > 0.25$ under the Pareto PDF (with $\alpha \geq 20$) and piezoelectric harvesting, we can predict that, by installing four independently-operating sets of 48 nodes and imposing that only one set is active at any given time, constant monitoring and transmission is ensured under energy neutrality.

Our framework allows for such studies to be done at early design stages and can incorporate all the relevant parameters of the WSN (protocol-related parameters, system settings, active time, etc.) in order to meet the requirements imposed by each application.

B. Minimizing Power Harvesting Requirements under a Fixed Network Setup

In the second application example, we consider a typical structural monitoring system, such as the one proposed by Notay and Safdar [45]. In such WSN-based systems, several sensors are embedded into a structure (e.g. sensors embedded within an airplane's wings or within the steel structure of a suspension bridge) in order to gather and transmit measurements to collection points. These collection points relay these measurements (along with their own) to WiFi-equipped access points [45] that have power supply. The sensors can harvest energy via the vibrations of the structure (e.g. airplane wing vibrations during flight) but energy neutrality must be ensured with the minimum possible power harvesting as the sensors are placed in difficult-to-service areas and must be able to operate in perpetuity. In such applications there is no strict real-time constraint for the data collection, as a volume of V_{fixed} bytes of measurements is collected from each sensor for batch off-line analysis of structural properties and the reaction times are in the order of hours or even days. Finally, the harvesting time interval is imposed by the application context, e.g. the average duration of a flight or the structural vibrations occurring in a suspension bridge during peak usage hours each day. Thus, the mean data transmission rate can be adjusted so as to minimize the required power harvesting under the pre-established network setup and harvesting time interval.

Under a given two-tier cluster-tree topology, such as the one shown in Figure 1(b), with [45]:

- fixed number of sensors and fixed relay configuration per tier ($n \equiv n_{\text{fixed}}$, $d \equiv d_{\text{fixed}}$),
- fixed requirements for the harvesting time and the volume of data to be collected by each node, i.e.
$$T \equiv T_{\text{fixed}} \text{ and } r \times c \times T_{\text{fixed}} \equiv V_{\text{fixed}}, \tag{46}$$
- the assumption of Pareto-distributed or fixed data transmission rate (i.e. $\alpha \equiv \alpha_{\text{fixed}}$),

we derive the mean rate and duty cycle setting per tier that minimize the harvested power requirements. This is achieved by: *(i)* deriving v_0 by solving (15) for v under $n \equiv n_{\text{fixed}}$; *(ii)* deriving r_0 by solving (13) for r under $v \equiv v_0$ and $d \equiv d_{\text{fixed}}$; *(iii)* deriving $c_0 \equiv \frac{V_{\text{fixed}}}{r_0 T_{\text{fixed}}}$. This effectively "tunes" the duty cycle and the mean data rate so that the fixed network and data transmission settings listed above become optimal, i.e. they lead to the minimum power harvesting that ensures energy neutrality. Under: the system settings of Section IV, $V_{\text{fixed}} = 25$ Mbit per sensor and $T_{\text{fixed}} = 86400$ s, Table IV shows the derived minimum power harvesting requirements in comparison to the results obtained under an ad-hoc allocation of data rates and duty cycles per tier.

Both the ad-hoc and the proposed settings satisfy the conditions imposed by (46) and lead to energy neutrality. However, deriving the mean rate and duty cycle per tier under the proposed framework meets these constraints with substantial savings in the required power harvesting, which were experimentally found to range between 19% and 37% in comparison to the ad-hoc settings. Hence, under piezoelectric harvesting providing $E[\chi] = 200\ \frac{\mu\text{W}}{\text{cm}^2}$ [3] the proposed approach requires an active harvesting area of only 2.7 - 3.0 cm^2 per node of Tier 2 while the ad-hoc approach requires 3.6 - 3.7 cm^2.

VI. CONCLUSIONS

We proposed an analytic framework for characterizing practical energy neutrality in uniformly-formed wireless sensor networks (WSNs). Our framework recognizes the importance of the application data transmission rate in the WSN's energy

Table IV: Minimum power harvesting requirement ($\min\{E[X]\}$) under ad-hoc settings and under optimized mean data rate and duty cycle adjustment with the proposed framework. The energy saving shows the percentile difference between the minimum harvesting requirement for the ad-hoc and proposed cases. The network parameters of this example correspond to Figure 1(b).

Network Parameters	Ad-hoc		Proposed		Energy Saving (%)	
	Pareto ($\alpha = 4$)	Fixed rate	Pareto ($\alpha = 4$)	Fixed rate	Pareto	Fixed rate
Tier 1	$r_{\text{adhoc}} = 22000$ bps	$r_{\text{adhoc}} = 22000$ bps	$r_0 = 37135$ bps	$r_0 = 36000$ bps	22.32	37.04
$d_{\text{fixed}} = 0$	$c_{\text{adhoc}} = 0.0132$	$c_{\text{adhoc}} = 0.0132$	$c_0 = 0.0078$	$c_0 = 0.0080$		
$n_{\text{fixed}} = 4$	$\min\{E[X]\} = 112$ μW	$\min\{E[X]\} = 108$ μW	$\min\{E[X]\} = 87$ μW	$\min\{E[X]\} = 68$ μW		
Tier 2	$r_{\text{adhoc}} = 12000$ bps	$r_{\text{adhoc}} = 12000$ bps	$r_0 = 24757$ bps	$r_0 = 24000$ bps	19.18	25.96
$d_{\text{fixed}} = 2$	$c_{\text{adhoc}} = 0.0241$	$c_{\text{adhoc}} = 0.0241$	$c_0 = 0.0117$	$c_0 = 0.0121$		
$n_{\text{fixed}} = 2$	$\min\{E[X]\} = 735$ μW	$\min\{E[X]\} = 728$ μW	$\min\{E[X]\} = 594$ μW	$\min\{E[X]\} = 539$ μW		

dissipation. Specifically, it provides for an analytic assessment of the expected energy dissipation in function of the system parameters, under a variety of statistical characterizations for the data transmission rate of each sensor node. The experimental assessment via: *(i)* low-power TelosB nodes, *(ii)* a recently-proposed, collision-free, communication protocol with rapid, low-energy, convergence and *(iii)* an energy measurement testbed, validates that our analytic framework matches experiments with TelosB nodes with accuracy that is within 7% of the measured energy consumption.

Our framework can be used in conjunction with particular harvesting technologies to predict the smallest possible energy harvesting interval leading to an energy-neutral deployment before costly and cumbersome testing in the field. Finally, our analysis could be used in conjunction with future energy-harvesting WSN systems and technologies in order to predict the best possible data transmission rate that can be accommodated in function of the system's operational settings.

VII. Appendix I

For all distributions, we provide the first derivative and show that, when set to zero and under $n \in (0, \infty)$, this leads to a single admissible extremum value. We then demonstrate that, under this value, the second derivative is guaranteed to be negative. Thus, the extremum value maximizes the residual energy under each data transmission rate PDF.

Pareto distribution: The first derivative of $E_{n,\text{P}}$ to n, $n \in (0, \infty)$, is

$$\frac{dE_{n,\text{P}}}{dn} = cT\left[\frac{ab}{n^2} - \frac{a}{n^2}(b+p)\left(\frac{vn}{a}\right)^{\alpha}\right]. \quad (47)$$

Assuming that $b \neq 0$, the only admissible solution of $\frac{dE_{n,\text{P}}}{dn} = 0$ is given in (15), as all other solutions are complex numbers. In conjunction with the fact that $E_{n,\text{P}}$ is differentiable for $n \in (0, \infty)$, this demonstrates that $n_{0,\text{P}}$ is the global extremum or inflection point of $E_{n,\text{P}}$. The second derivative of $E_{n,\text{P}}$ is:

$$\frac{d^2E_{n,\text{P}}}{dn^2} = cT\left[-\frac{2ab}{n^3} - \frac{a}{n^3}(\alpha-2)(b+p)\left(\frac{vn}{a}\right)^{\alpha}\right]. \quad (48)$$

By evaluating $\frac{d^2E_{n,\text{P}}}{dn^2}$ for $n_{0,\text{P}}$ nodes, we obtain

$$\frac{d^2E_{n,\text{P}}}{dn^2}(n_{0,\text{P}}) = -\frac{cT[2b+(\alpha-2)](b+p)^{\frac{3}{\alpha}}v^3}{a^2 b^{\frac{3}{\alpha}}}, \quad (49)$$

which is negative since $\alpha \geq 2$ and all variables are positive. This means that the maximum-possible residual energy for $n \in (0, \infty)$ is achieved under $n = n_{0,\text{P}}$. This derivation also covers the case of fixed-rate data production and transmission if we set $v = (d+1)(r-1)$ and $\alpha = r$.

Exponential distribution: The first derivative of $E_{n,\text{E}}$ to n, $n \in (0, \infty)$, is

$$\frac{dE_{n,\text{E}}}{dn} = cT\left[\frac{ab}{n^2} - \frac{a}{n^2}(b+p)\exp\left(-\frac{a}{n(d+1)r}\right)\right]. \quad (50)$$

Assuming that $b \neq 0$, the only admissible solution for $\frac{dE_{n,\text{E}}}{dn} = 0$ is given in (21), as all other solutions are complex numbers. In conjunction with the fact that $E_{n,\text{E}}$ is differentiable for $n \in (0, \infty)$, this demonstrates that $n_{0,\text{E}}$ is the global extremum or inflection point of $E_{n,\text{E}}$. The second derivative of $E_{n,\text{E}}$ is:

$$\begin{aligned}\frac{d^2E_{n,\text{E}}}{dn^2} &= cT\left[-\frac{2ab}{n^3} - \frac{a}{n^4(d+1)r}(b+p)\right. \\ &\times \left.\exp\left(-\frac{a}{n(d+1)r}\right)[a - 2n(d+1)r]\right] \quad (51)\end{aligned}$$

By evaluating $\frac{d^2E_{n,\text{E}}}{dn^2}$ for $n_{0,\text{E}}$ nodes, we obtain

$$\frac{d^2E_{n,\text{E}}}{dn^2}(n_{0,\text{E}}) = -\frac{cT(d+1)^3 br^3}{a^2}\left[\ln\left(\frac{b+p}{b}\right)\right]^4, \quad (52)$$

which is negative since all variables are positive and the natural logarithm is raised to an even power. This means that the maximum-possible residual energy for $n \in (0, \infty)$ is achieved under $n = n_{0,\text{E}}$.

Half-Gaussian distribution: The first derivative of $E_{n,\text{H}}$ to n, $n \in (0, \infty)$, is

$$\frac{dE_{n,\text{H}}}{dn} = cT\left[-\frac{ap}{n^2} + \frac{a}{n^2}(b+p)\,\text{erf}\left(\frac{a}{\sqrt{\pi}(d+1)rn}\right)\right]. \quad (53)$$

Assuming that $p \neq 0$, the only admissible solution for $\frac{dE_{n,\text{H}}}{dn} = 0$ is given in (25). In conjunction with the fact that $E_{n,\text{H}}$ is differentiable for $n \in (0, \infty)$, this demonstrates that $n_{0,\text{H}}$ is the global extremum or inflection point of $E_{n,\text{H}}$. The second derivative of $E_{n,\text{H}}$ is:

$$\begin{aligned}\frac{d^2E_{n,\mathrm{H}}}{dn^2} &= cT\left[\frac{2ap}{n^3}-\frac{2a}{n^3}(b+p)\right.\\ &\times \operatorname{erf}\left(\frac{a}{\sqrt{\pi}(d+1)rn}\right)-\frac{2a^2}{\pi(d+1)rn^4} \quad (54)\\ &\times \left.(b+p)\exp\left(-\frac{a^2}{\pi((d+1)r)^2n^2}\right)\right].\end{aligned}$$

By evaluating $\frac{d^2E_{n,\mathrm{H}}}{dn^2}$ for $n_{0,\mathrm{H}}$ nodes, we obtain

$$\begin{aligned}\frac{d^2E_{n,\mathrm{H}}}{dn^2}(n_{0,\mathrm{H}}) &= -\frac{2\pi cT(d+1)^3r^3(b+p)}{a^2}\\ &\times \left[\operatorname{erf}^{-1}\left(\frac{p}{b+p}\right)\right]^4 \quad (55)\\ &\times \exp\left(-\left[\operatorname{erf}^{-1}\left(\frac{p}{b+p}\right)\right]^2\right),\end{aligned}$$

which is negative since the inverse error function is raised to an even power and all variables are positive. This means that the maximum-possible residual energy for $n\in(0,\infty)$ is achieved under $n=n_{0,\mathrm{H}}$.

Gaussian distribution: The first derivative of $E_{n,\mathrm{N}}$ to n, $n\in(0,\infty)$, is

$$\begin{aligned}\frac{dE_{n,\mathrm{N}}}{dn} &= cT\left[-\frac{ap}{n^2}+\frac{a}{2n^2}(b+p)\left[\operatorname{erf}\left(\frac{(d+1)r}{\sqrt{2}\sigma}\right)\right.\right.\\ &- \left.\left.\operatorname{erf}\left(\frac{(d+1)r-\frac{a}{n}}{\sqrt{2}\sigma}\right)\right]\right]. \quad (56)\end{aligned}$$

Assuming that $p\neq 0$, the only admissible solution for $\frac{dE_{n,\mathrm{N}}}{dn}=0$ is given in (30). In conjunction with the fact that $E_{n,\mathrm{N}}$ is differentiable for $n\in(0,\infty)$, $n_{0,\mathrm{N}}$ is the global extremum or inflection point of $E_{n,\mathrm{N}}$. The second derivative of $E_{n,\mathrm{N}}$ is:

$$\begin{aligned}\frac{d^2E_{n,\mathrm{N}}}{dn^2} &= cT\left[\frac{2ap}{n^3}+\frac{a}{n^3}(b+p)\right.\\ &\times \left[\operatorname{erf}\left(\frac{(d+1)r-\frac{a}{n}}{\sqrt{2}\sigma}\right)-\operatorname{erf}\left(\frac{(d+1)r}{\sqrt{2}\sigma}\right)\right] \quad (57)\\ &- \left.\frac{a^2}{\sqrt{2\pi}\sigma n^4}(b+p)\exp\left(\frac{-\left[(d+1)r-\frac{a}{n}\right]^2}{2\sigma^2}\right)\right].\end{aligned}$$

By evaluating $\frac{d^2E_{n,\mathrm{N}}}{dn^2}$ for $n_{0,\mathrm{N}}$ nodes, we obtain

$$\frac{d^2E_{n,\mathrm{N}}}{dn^2}(n_{0,\mathrm{N}}) = -\frac{cT(b+p)}{a^2\sqrt{2\pi}\sigma}\left[(d+1)r-\sqrt{2}\sigma c_{\mathrm{N}}\right]^4\exp\left(-c_{\mathrm{N}}^2\right), \quad (58)$$

which is negative since $\left[(d+1)r-\sqrt{2}\sigma c_{\mathrm{N}}\right]$ is raised to an even power and all variables are positive. This means that the maximum-possible residual energy for $n\in(0,\infty)$ is achieved under $n=n_{0,\mathrm{N}}$.

VIII. Appendix II

Proof of Proposition 1: After a few straightforward derivations, we reach:

$$\begin{aligned}&\min\{E[X]\}_{\mathrm{P}}-\min\{E[X]\}_{\mathrm{U}}\\ &= c(d+1)r\left[-b+b^{\frac{\alpha-1}{\alpha}}(b+p)^{\frac{1}{\alpha}}-\frac{pb}{b+p}\right] \quad (59)\end{aligned}$$

Replacing (15) and (7) in the first condition of (33) we reach:

$$(b+p)^{\frac{1}{\alpha}} = \frac{2p(d+1)r}{v}\times\frac{b^{\frac{1}{\alpha}}}{s_{\mathrm{PU}}(b+p)} \quad (60)$$

that leads to the final result when used within (59).

Proof of Proposition 2: After a few straightforward derivations, we reach:

$$\begin{aligned}&\min\{E[X]\}_{\mathrm{E}}-\min\{E[X]\}_{\mathrm{U}}\\ &= c(d+1)r\left[b\ln\left(\frac{b+p}{b}\right)-\frac{pb}{b+p}\right] \quad (61)\end{aligned}$$

Replacing (21) and (7) in the second condition of (33), we reach the condition:

$$\ln\left(\frac{b+p}{b}\right) = \frac{2p}{s_{\mathrm{EU}}(b+p)} \quad (62)$$

that leads to the final result when used within (61).

Proof of Proposition 3: In this case, we reach the expression:

$$\begin{aligned}&\min\{E[X]\}_{\mathrm{H}}-\min\{E[X]\}_{\mathrm{U}}\\ &= c(d+1)r[-b+(b+p)\\ &\times \exp\left(-\operatorname{erf}^{-1}\left(\frac{p}{b+p}\right)^2\right)-\frac{pb}{b+p}\Bigg] \quad (63)\end{aligned}$$

Replacing (25) and (7) in the third condition of (33), we reach the condition:

$$\operatorname{erf}^{-1}\left(\frac{p}{b+p}\right) = \frac{2p}{\sqrt{\pi}s_{\mathrm{HU}}(b+p)}. \quad (64)$$

Taking the square of the last expression and replacing in (63) leads to the final result.

Proof of Proposition 4: We reach the expression:

$$\begin{aligned}&\min\{E[X]\}_{\mathrm{H}}-\min\{E[X]\}_{\mathrm{E}}\\ &= c(d+1)r[-b+(b+p)\\ &\times \exp\left(-\operatorname{erf}^{-1}\left(\frac{p}{b+p}\right)^2\right)-b\ln\left(\frac{b+p}{b}\right)\Bigg] \quad (65)\end{aligned}$$

Replacing (25) and (21) in (33), we reach the condition:

$$\operatorname{erf}^{-1}\left(\frac{p}{b+p}\right) = \frac{1}{\sqrt{\pi}s_{\mathrm{HE}}}\ln\left(\frac{b+p}{b}\right) \quad (66)$$

Taking the square of the last expression and replacing in (65) leads to the final result.

REFERENCES

[1] A. Bachir, M. Dohler, T. Watteyne, and K.K. Leung, "MAC essentials for wireless sensor networks," *IEEE Comm. Surv. Tut.*, vol. 12, no. 2, pp. 222 –248, quarter 2010.

[2] A. Kansal, J. Hsu, S. Zahedi, and M. B. Srivastava, "Power management in energy harvesting sensor networks," *ACM Trans. Embed. Comput. Syst.*, vol. 6, no. 4, Sep. 2007.

[3] S. Sudevalayam and P. Kulkarni, "Energy harvesting sensor nodes: Survey and implications," *IEEE Comm. Surv. Tut.*, vol. 13, no. 3, pp. 443 –461, quarter 2011.

[4] S. Chalasani and J.M. Conrad, "A survey of energy harvesting sources for embedded systems," in *IEEE Southeastcon, 2008*, Apr. 2008, pp. 442 –447.

[5] He C., M.E. Kiziroglou, D.C. Yates, and E.M. Yeatman, "A MEMS self-powered sensor and RF transmission platform for WSN nodes," *IEEE Sensors J.*, vol. 11, no. 12, pp. 3437 –3445, Dec. 2011.

[6] D. Buranapanichkit and Y. Andreopoulos, "Distributed time-frequency division multiple access protocol for wireless sensor networks," *IEEE Wirel. Comm. Lett.*, vol. 1, no. 5, pp. 440 –443, Oct. 2012.

[7] L. Tang, Y. Sun, O. Gurewitz, and D. B. Johnson, "EM-MAC: a dynamic multichannel energy-efficient MAC protocol for wireless sensor networks," in *Proc. of the 12th ACM Int. Symp. on Mob. Ad Hoc Netw. and Comput.*, New York, NY, USA, 2011, MobiHoc '11, pp. 23:1–23:11, ACM.

[8] A. Koubaa, A. Cunha, and M. Alves, "A time division beacon scheduling mechanism for IEEE 802.15.4/zigbee cluster-tree wireless sensor networks," in *19th Euromicro Conf. on Real-Time Syst., ECRTS '07*, july 2007, pp. 125 –135.

[9] G. Lu, B. Krishnamachari, and C.S. Raghavendra, "Performance evaluation of the IEEE 802.15. 4 MAC for low-rate low-power wireless networks," in *IEEE Internat. Conf. on Perf., Comput., and Comm.* IEEE, 2004, pp. 701–706.

[10] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Trans. on Communications*, vol. 60, no. 1, pp. 220–230, 2012.

[11] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Trans. on Wireless Comm.*, vol. 9, no. 4, pp. 1326 –1336, Apr. 2010.

[12] L. Benini, A. Bogliolo, and G. De Micheli, "A survey of design techniques for system-level dynamic power management," *IEEE Trans. on Very Large Scale Integr. (VLSI) Syst.*, vol. 8, no. 3, pp. 299–316, 2000.

[13] B. Zhang, R. Simon, and H. Aydin, "Energy management for time-critical energy harvesting wireless sensor networks," in *Lecture Notes in Comp. Sc.: Stabil., Safety, and Secur. of Distr. Syst.*, 2010, vol. 6366, pp. 236–251.

[14] C. Alippi, G. Anastasi, M. Di Francesco, and M. Roveri, "Energy management in wireless sensor networks with energy-hungry sensors," *IEEE Instr., & Meas. Mag.*, vol. 12, no. 2, pp. 16–23, 2009.

[15] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Trans. on Wireless Communications*, vol. 11, no. 3, pp. 1180–1189, 2012.

[16] C.M. Vigorito, D. Ganesan, and A.G. Barto, "Adaptive control of duty cycling in energy-harvesting wireless sensor networks," in *IEEE 4th Annual Comm. Soc. Conf. Sensor, Mesh and Ad Hoc Comm. and Netw.*, Jun. 2007, pp. 21 –30.

[17] A. Koubaa, M. Alves, M. Attia, and A. Van Nieuwenhuyse, "Collision-free beacon scheduling mechanisms for IEEE 802.15.4/zigbee cluster-tree wireless sensor networks," in *Proc. of the 7th Int. Worksh. on Appl. and Serv. in Wireless Netw. (ASWN)*, 2007.

[18] K. Pister and L. Doherty, "TSMP: Time synchronized mesh protocol," *IASTED Distr. Sensor Netw.*, pp. 391–398, 2008.

[19] J. Song, S. Han, A.K. Mok, D. Chen, M. Lucas, and M. Nixon, "WirelessHART: Applying wireless technology in real-time industrial process control," in *IEEE Real-Time and Embed. Tech. and Appl. Symp., 2008. RTAS'08*. IEEE, 2008, pp. 377–386.

[20] J. Degesys and R. Nagpal, "Towards desynchronization of multi-hop topologies," in *IEEE International Conference on Self-Adaptive and Self-Organizing Systems, SASO'08*, 2008, pp. 129–138.

[21] A. Motskin, T. Roughgarden, P. Skraba, and L. Guibas, "Lightweight coloring and desynchronization for networks," in *IEEE INFOCOM 2009*. IEEE, 2009, pp. 2383–2391.

[22] T. Wu and S. Biswas, "Minimizing inter-cluster interference by self-reorganizing mac allocation in sensor networks," *Wireless Networks*, vol. 13, no. 5, pp. 691–703, Oct. 2007.

[23] G. Haigang, L. Ming, W. Xiaomin, C. Lijun, and X Li, "An interference free cluster-based tdma protocol for wireless sensor networks," *Wireless Algorithms, Systems, and Applications*, vol. 4138, pp. 217–227, 2006.

[24] X. Zhuang, Y. Yang, and W. Ding, "A tdma-based protocol and implementation for avoiding inter-cluster interference of wireless sensor network," in *IEEE Industrial Tech. Conf. ICIT 2009.*, 2009, pp. 1–6.

[25] J. Kho, A. Rogers, and N.R. Jennings, "Decentralized control of adaptive sampling in wireless sensor networks," *ACM Trans. on Sensor Netw.*, vol. 5, no. 3, pp. 19, 2009.

[26] A. McCree, K. Brady, and T.F. Quatieri, "Multisensor very lowbit rate speech coding using segment quantization," in *IEEE Internat. Conf. on Acoust., Speech and Signal Proc.* IEEE, 2008, pp. 3997–4000.

[27] M. Pursley and L. Davisson, "Variable rate coding for nonergodic sources and classes of ergodic sources subject to a fidelity constraint," *IEEE Trans. on Inf. Theory*, vol. 22, no. 3, pp. 324–337, 1976.

[28] P. Kulkarni, D. Ganesan, P. Shenoy, and Q. Lu, "Senseye: a multi-tier camera sensor network," in *13th annual ACM international conference on Multimedia*. ACM, 2005, pp. 229–238.

[29] A. Rowe, D. Goel, and R. Rajkumar, "Firefly mosaic: A vision-enabled wireless sensor networking system," in *Real-Time Systems Symposium, 2007. RTSS 2007. 28th IEEE International*. IEEE, 2007, pp. 459–468.

[30] M. Tagliasacchi, S. Tubaro, and A. Sarti, "On the modeling of motion in Wyner-Ziv video coding," in *IEEE Int. Conf. on Image Process.* IEEE, 2006, pp. 593–596.

[31] A. Redondi, M. Cesana, and M. Tagliasacchi, "Low bitrate coding schemes for local image descriptors," in *MMSP*, 2012, pp. 124–129.

[32] G. WernerAllen, K. Lorincz, J. Johnson, J. Lees, and M. Welsh, "Fidelity and yield in a volcano monitoring sensor network," in *7th symposium on Operating systems design and implementation*. ACM, 2006, pp. 381–396.

[33] D. Palma, L. Bencini, G. Collodi, G. Manes, F. Chiti, R. Fantacci, and A. Manes, "Distributed monitoring systems for agriculture based on wireless sensor network technology," *International Journal on Advances in Networks and Services,*, vol. 3, 2010.

[34] A. Redondi, M. Chirico, L. Borsani, M. Cesana, and M. Tagliasacchi, "An integrated system based on wireless sensor networks for patient monitoring, localization and tracking," *Ad Hoc Networks*, vol. 11, no. 1, pp. 39–53, 2013.

[35] S. Roundy, D. Steingart, L. Frechette, P. Wright, and J. Rabaey, "Power sources for wireless sensor networks," *J. of Wireless Sensor Netw.*, pp. 1–17, 2004.

[36] E.Y. Lam and J.W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Trans. on Image Process.*, vol. 9, no. 10, pp. 1661–1666, Oct. 2000.

[37] B. Foo, Y. Andreopoulos, and M. van der Schaar, "Analytical rate-distortion-complexity modeling of wavelet-based video coders," *IEEE Trans. on Signal Process.*, vol. 56, no. 2, pp. 797–815, Feb. 2008.

[38] J. Degesys, I. Rose, A. Patel, and R. Nagpal, "DESYNC: self-organizing desynchronization and TDMA on wireless sensor networks," in *Proc. 6th Int. Conf. on Inf. Process. in Sensor Netw.* ACM, 2007, pp. 11–20.

[39] V. Paxson and S. Floyd, "Wide area traffic: the failure of Poisson modeling," *IEEE/ACM Trans. on Networking*, vol. 3, no. 3, pp. 226–244, Mar. 1995.

[40] K. Park, G. Kim, and M. Crovella, "On the relationship between file sizes, transport protocols, and self-similar network traffic," in *Proc. 1996 Int. Conf. on Network Prot.* IEEE, 1996, pp. 171–180.

[41] M. Dai, Y. Zhang, and D. Loguinov, "A unified traffic model for MPEG-4 and H.264 video traces," *IEEE Trans. on Multimedia*, vol. 11, no. 5, pp. 1010–1023, May 2009.

[42] Y. Andreopoulos and I. Patras, "Incremental refinement of image salient-point detection," *IEEE Trans. on Image Process.*, vol. 17, no. 9, pp. 1685–1699, Sept. 2008.

[43] A. Papoulis, *Probability & statistics (3rd edition)*, Prentice-Hall, 1990.

[44] W.R. Gilks and P. Wild, "Adaptive rejection sampling for Gibbs sampling," *Applied Statistics*, pp. 337–348, 1992.

[45] J.K. Notay and G.A. Safdar, "A wireless sensor network based structural health monitoring system for an airplane," in *Automation and Computing (ICAC), 2011 17th International Conference on*. IEEE, 2011, pp. 240–245.

[46] K.W. Fan, Z. Zheng, and P. Sinha, "Steady and fair rate allocation for rechargeable sensors in perpetual sensor networks," in *Proc. of the 6th ACM Conf. on Embed. Network Sensor Syst.* ACM, 2008, pp. 239–252.

Hana Besbes was MSc student in the MSc Telecommunications program of the Department of Electronic and Electrical Engineering of University College London (UK) when this work was performed. She is now studying at the Technical University of Munich, Germany. Her research interests are in wireless sensor networks analysis and design aspects.

George Smart is PhD student in the Department of Electronic and Electrical Engineering of University College London (UK). His research interests are in wireless sensor networks and embedded systems software and hardware design.

Dujdow Buranapanichkit is PhD student in the Department of Electronic and Electrical Engineering of University College London (UK). Her research interests are in wireless sensor networks and distributed synchronization mechanisms and protocol design.

Christos Kloukinas is Senior Lecturer in the Department of Computer Science, City University London (UK). His research interests are in the area of software engineering and embedded systems, focusing in the development of methods and tools for the analysis, fine-grain control, optimization and implementation of mission-critical and safety-critical, hard real-time embedded systems.

Yiannis Andreopoulos (M'00) is Senior Lecturer in the Department of Electronic and Electrical Engineering of University College London (UK). His research interests are in wireless sensor networks, error-tolerant computing and multimedia systems.