Interactive Joint Transfer of Energy and Information

Popovski, Petar; Fouladgar, A. M. ; Simeone, Osvaldo

Published in:
I E E E Transactions on Communications

DOI (link to publication from Publisher):
10.1109/TCOMM.2013.031213.120723

Publication date:
2013

Document Version
Early version, also known as pre-print

Interactive Joint Transfer of Energy and Information

P. Popovski, A. M. Fouladgar, and O. Simeone

***Abstract*—In some communication networks, such as passive RFID systems, the energy used to transfer information between a sender and a recipient can be reused for successive communication tasks. In fact, from known results in physics, any system that exchanges information via the transfer of given physical resources, such as radio waves, particles and qubits, can conceivably reuse, at least part, of the received resources.**

This paper aims at illustrating some of the new challenges that arise in the design of communication networks in which the signals exchanged by the nodes carry both information and energy. To this end, a baseline two-way communication system is considered in which two nodes communicate in an interactive fashion. In the system, a node can either send an "on" symbol (or "1"), which costs one unit of energy, or an "off" signal (or "0"), which does not require any energy expenditure. Upon reception of a "1" signal, the recipient node "harvests", with some probability, the energy contained in the signal and stores it for future communication tasks. Inner and outer bounds on the achievable rates are derived. Numerical results demonstrate the effectiveness of the proposed strategies and illustrate some key design insights.

***Index Terms*—Two-way channel, interactive communication, energy transfer, energy harvesting.**

I. Introduction

THE conventional assumption made in the design of communication systems is that the energy used to transfer information between a sender and a recipient cannot be reused for future communication tasks. There are, however, notable exceptions. An example is given by communication based on wireless energy transfer, such as passive RFID systems [1] or some body area networks [2], in which a terminal can transfer both information and energy via the transmitted radio signal, and the delivered energy can be used for communication by the recipients. For instance, a passive RFID tag modulates information by backscattering the radio energy received from the reader (see, e.g., [1]). Another, less conventional, example is that of a biological system in which information is communicated via the transmission of particles (see, e.g., [3]), which can be later reused for other communication tasks. A further potential instance of this type of networks is one in which communication takes place via the exchange of quantum systems, such as photons, which may measured and then reused [4].

Manuscript received September 25, 2012; revised December 20, 2012. The associate editor coordinating the review of this paper and approving it for publication was M. Tao.

P. Popovski is with the Department of Electronic Systems Aalborg University, Denmark (e-mail: petarp@es.aau.dk).

A. M. Fouladgar and O. Simeone are with the CWCSPR, New Jersey Institute of Technology, Newark, NJ 07102 USA (e-mail: {af8, osvaldo.simeone}@njit.edu).

Part of this paper was presented at the IEEE Information Theory Workshop (ITW 2012), Lausanne, Switzerland, Sept. 2012.

Digital Object Identifier 10.1109/TCOMM.2013.031213.120723

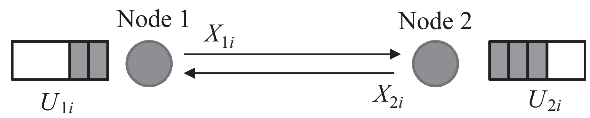

Fig. 1. Two-way noiseless binary communication with energy exchange. The total number of energy units is fixed (to five in the figure) and transmission of a "1" symbol transfers energy from the sender node to the recipient. See Fig. 3 for a generalized model.

To summarize, any system that exchanges information via the transfer of given physical resources, such as radio waves, particles or qubits, can conceivably reuse, at least part, of the received resources for later communication tasks. This conclusion is supported by physical considerations [5] and practically demonstrated by the existing systems based on this principle [1], [2]. It is emphasized that the possibility to deliver jointly energy and information promises not only to ease the energy requirements of various communication systems, but also, more importantly, to enable novel applications, such as the body area networks studied in [1], [2]. Moreover, an understanding of the interplay between energy and information flows could lead to insights on the workings of some communication systems in nature [3].

A. State of the Art

While the interaction between energy and information continues to be subject of research in the physics community (see, e.g., [6]), the topic has been tackled from a communication and information theoretic level only in a handful of pioneering works, as reviewed in the following. The references [7], [8], [9] have focused on the problem of maximizing the information rate of a point-to-point system subject to minimum received energy constraints. Specifically, in [7] a single point-to-point channel was studied, while [8], [9] investigated a set of parallel point-to-point channels. To illustrate the trade-offs between the transfer of energy and information in a point-to-point channel consider the noiseless transmission of a 4-PAM signal in the alphabet $\{-2, 1, 1, 2\}$. If one requires the received energy to be the maximum possible, that is, to be equal to 4, the maximum transferable information rate is 1 bit per symbol, since one is forced to communicate only with the larger energy symbols $\{-2, 2\}$. However, with no receive energy constraint, one can clearly convey 2 bits per symbol by choosing all available symbols with equal probability. This example also explains the substantial difference between the problems studied in [7], [8], [9] and that with maximum receive energy constraints studied in [10][1]. The optimization of beamforming

[1]With a maximal receive energy constraint of 4, one can still clearly transmit 2 bits per symbol.

strategies under a receive energy constraint was tackled in [11], [12] for multiantenna broadcast channels. Considerations on the design of the receiver under the constraint that, when harvesting energy from the antenna, the receiver is not able to use the same signal for information decoding, can be found in [13].

B. Contributions

In all of the previous work summarized above, the requirement on the energy harvested from the received signal is considered to be an additional constraint imposed to the system design. This work is instead motivated by the observation that, in more complex network scenarios, as mentioned above, the energy harvested from the received signal may be reused for future communication tasks. In this case, the energy and information content of the exchanged signals should be engineered so as to best suit the requirements of the communication network. To study this aspect, we consider a baseline two-way communication system, as illustrated in Fig. 1. This is incidentally the same topology selected by Shannon to initiate the study of networks from an information theoretic perspective [18]. In the considered model, the two nodes interact for the exchange of information and can harvest the received energy.

To enable analysis and insights, we assume that the two parties involved have a common clock and that, at each time, a node can either send an "on" symbol (or "1"), which costs one unit of energy, or an "off" signal (or "0"), which does not require any energy expenditure. Upon reception of a "1" signal, the recipient node can harvest, possibly with some loss, the energy contained in the signal and stores it for future communication tasks. Each node communicates in a full-duplex manner, that is, at a given instant, it can simultaneously send and receive an energy unit. The channel from Node 1 to Node 2 is orthogonal to the channel from Node 2 to Node 1, and hence the full-duplex channel is an ideal composition of two independent unidirectional channels. In order to introduce the main concepts with the minimum of the notation and technical complications, we first consider the case in which the two nodes start with a given number of energy units in their batteries, which can neither be lost or replenished from outside, and the binary channel in either direction is noiseless.

To see that even this simple scenario offers relevant research challenges, we observe the following. If there were no limitation on the number of energy units, the nodes could communicate 1 bit per channel use in either direction given that the channels are ideal. However, if there is, say, one energy unit available in the system, only the node that currently possesses the energy unit can transmit a "1", whereas the other node is forced to transmit a "0"[2]. Therefore, the design of the communication strategy at the nodes should aim not only at transferring the most information to the counterpart, but also to facilitate energy transfer to enable communication in the reverse direction. We study this problem, described in Sec. II by deriving inner and outer bounds on the achievable rate region as a function of the available energy units in Sec. III. The main results are then extended to a model that accounts for energy replenishments and losses, along with noisy channels. The generalized model is presented in Sec. IV and the generalized results are presented in Sec. V.

[2]In the case when there is a single energy unit, the system naturally operates in a half-duplex manner

It is finally observed that the class of problems at hand, in which terminals can harvest energy from the received signals, is related to the increasing body of work on energy harvesting (see, e.g., [14]-[17] and references therein). However, in this line of work, the energy is assumed to be harvested from the environment in a way that is not affected by the communication process, unlike the scenario under study.

Notation: $[m, n] = \{m, m+1, ..., n\}$ for integers $m \leq n$. $\mathbb{N}$ is the set of integer numbers. We use the standard notation in [19] for information theoretic quantities such as entropy and mutual information. If the distribution is $\texttt{Bern}(p)$ we will also write $H(p)$ for the entropy. Capital letters denote random variables and the corresponding lowercase quantities denote specific values of the random variables. X^i for an integer i denotes the vector $X^i = (X_1, ..., X_i)$.

II. System Model

We consider the binary and noiseless two-way system illustrated in Fig. 1, in which the total number of *energy units* in the system is equal to a finite integer number $\texttt{U} \geq 1$ at all times and the channels between the two nodes are noiseless. Each node has an energy buffer that can store at least $\texttt{U}$ units. In Sec. IV, the model will be extended to include stochastic energy losses and replenishments along with noisy channels. At any given time instant k, with $k \in [1, n]$, the *state* of the system $(U_{1,k}, U_{2,k}) \in \mathbb{N}^2$ is given by the current energy allocation between the two nodes. Specifically, a state $(U_{1,k}, U_{2,k})$ indicates that at the kth channel use there are $U_{j,k}$ energy units at Node j, with $j = 1, 2$. Since we assume here that $U_{1,k} + U_{2,k} = \texttt{U}$ for each channel use $k \in [1, n]$ (i.e., no energy losses occur), then, in this section and in the next, we will refer to $U_{1,k}$ as the state of the system, which always imply the equality $U_{2,k} = \texttt{U} - U_{1,k}$.

At any channel use $k \in [1, n]$, each Node j can transmit either symbol $X_{j,k} = 0$ or symbol $X_{j,k} = 1$, and transmission of a "1" costs one energy unit, while symbol "0" does not require any energy expenditure. Therefore, the available transmission alphabet for Node j, $j = 1, 2$ during the kth channel use is

$$\mathcal{X}_u = \{0, 1\} \quad \text{if} \quad U_{j,k} = u \geq 1 \tag{1a}$$

$$\text{and } \mathcal{X}_0 = \{0\} \quad \text{if} \quad U_{j,k} = 0, \tag{1b}$$

so that $X_{j,k} \in \mathcal{X}_u$ if $U_{j,k} = u$ energy units are available at Node j. The channel is noiseless so that the received signals at channel use k are given by

$$Y_{1,k} = X_{2,k} \text{ and } Y_{2,k} = X_{1,k} \tag{2}$$

for Node 1 and Node 2, respectively.

Transmission of a "1" transfers one energy unit from the sender node to the recipient node. Therefore, the state of Node 1 for $k \in [1, n]$ evolves as follows

$$U_{1,k} = (U_{1,k-1} - X_{1,k-1}) + X_{2,k-1}, \tag{3}$$

where we set $U_{1,1} = u_{1,1} \leq \mathtt{U}$ as some initial state and $U_{2,k} = \mathtt{U} - U_{1,k}$. We observe that the current state $U_{1,k}$ is a deterministic function of the number $\mathtt{U}$ of total energy units, of the initial state $U_{1,1}$ and of the previously transmitted signals X_1^{k-1} and X_2^{k-1}. We also note that both nodes are clearly aware of the state of the system at each time since $U_{1,k} + U_{2,k} = \mathtt{U}$ is satisfied for each channel use k.

Node 1 has message M_1, uniformly distributed in the set $[1, 2^{nR_1}]$, to communicate to Node 2, and similarly for the message $M_2 \in [1, 2^{nR_2}]$ to be communicated between Node 2 and Node 1. Parameters R_1 and R_2 are the transmission rates in bits per channel use (c.u.) for Node 1 and for Node 2, respectively. We use the following definitions for an $(n, R_1, R_2, \mathtt{U})$ code. Specifically, the code is defined by: the overall number of energy units $\mathtt{U}$; two sequences of encoding functions, namely, for Node 1, we have functions $\mathrm{f}_{1,k}$ for $k \in [1, n]$, which map the message M_1 and the past received symbols X_2^{k-1} (along with the initial state) into the currently transmitted signal $X_{1,k} \in \mathcal{X}_{U_{1,k}}$; similarly, for Node 2, we have functions $\mathrm{f}_{2,k}$ for $k \in [1, n]$, which map the message M_2 and the past received symbols X_1^{k-1} (along with the initial state) into the currently transmitted signal $X_{2,k} \in \mathcal{X}_{U_{2,k}}$; and two decoding functions, namely, for Node 1, we have a function g_1, which maps all received signals X_2^n and the local message M_1 into an estimate $\hat{M}_1$ of message M_2; and similarly, for Node 2, we have a function g_2, which maps all received signals X_1^n and the local message M_2 into an estimate $\hat{M}_1$ of message M_1.

We say that rates (R_1, R_2) are achievable with $\mathtt{U}$ energy units if there exists an $(n, R_1, R_2, \mathtt{U})$ code for all sufficiently large n that guarantees reliable communication. We are interested in studying the closure of the set of all the rate pairs (R_1, R_2) that are achievable with $\mathtt{U}$ energy units, which we refer to as capacity region $\mathcal{C}(\mathtt{U})$. Given the noiseless nature of the channels, we note that the initial state $U_{1,1} = u_{1,1} \leq \mathtt{U}$ does not affect the rate region since in a finite number of steps it is always possible to redistribute the energy according to any desired state.

III. Inner and Outer Bounds

In this section, we derive inner and outer bounds to the capacity region.

A. Inner Bounds

In order to gain insights into the nature of the problem under study, we consider here various communication strategies. We start by the simplest, but intuitively important, case with $\mathtt{U}=1$, and we then generalize to $\mathtt{U}>1$.

1) $\mathtt{U}= 1$ Energy Unit: We start with the special case of one energy unit ($\mathtt{U} = 1$) and assume the initial state $u_{1,1} = 1$, so that the energy unit is initially available at Node 1. The other case, namely $u_{1,1} = 0$, can be treated in a symmetric way. In this setting, during each channel use, "information" can be transferred only from the node where the energy unit resides towards the other node, and not vice versa, since the other node is forced to transmits the "0" symbol. This suggests that, when $\mathtt{U} = 1$, the channel is necessarily used in a time-sharing manner, and thus the sum-rate is at most one bit per channel use. The first question is whether the sum-rate of 1 bit/c.u. is achievable, and, if so, which strategy accomplishes this task.

A Naïve Strategy: We start with a rather naïve encoding strategy that turns out to be insufficient to achieve the upper bound of 1 bit/c.u.. The nodes agree on a frame size $F = 2^b > 1$ channel uses for some integer b and partition the n channel uses in n/F frames (assumed to be an integer for simplicity). The node that has the energy unit at the beginning of the frame communicates $b = \log_2 F$ bits to the other node by placing the energy unit in one specific channel use among the $F = 2^b$ of the frame. This process also transfers the energy unit to the other node, and the procedure is repeated. The sum-rate achieved by this scheme is

$$R_1 + R_2 = \frac{\log_2 F}{F} \text{ [bits/c.u.]}, \tag{4}$$

which is rather inefficient: the maximum is achieved with $F = 2$, leading to a sum-rate of $R_1 + R_2 = 1/2$ bits/ c.u..

The previous strategy can be easily improved by noting that the frame can be interrupted after the channel use in which the energy unit is used, since the receiving node can still decode the transmitted b bits. This strategy corresponds to using a variable-length channel code. Specifically, we can assign, without loss of optimality within this class of strategies, the codeword "01" to information bit "0" and the codeword "1" to bit "1". The average number of channel uses per bit is thus $1/2 + 1/2{\cdot}2 = 3/2$. Therefore, the overall number of channel uses necessary for the transmission of m bits is upper bounded by $\frac{3m}{2} + m\epsilon$ with arbitrarily small probability for large m by the weak law of large numbers (see, e.g., [19]). It follows that an achievable sum-rate is given by

$$R_1 + R_2 = \frac{1}{3/2} = \frac{2}{3}, \tag{5}$$

which is still lower than the upper bound of 1 bit/c.u..

An Optimal Strategy: We now discuss a strategy that achieves the upper bound of 1 bit/c.u.. The procedure is based on time-sharing, as driven by the transfer of the energy unit from one to the other node. Each node j encodes its data in codewords of size m, where each codeword is selected uniformly from the set of $\binom{m}{\frac{m}{2}}$ binary sequences $b_{j,1}, ..., b_{j,m}$ with an equal fraction of zeros and ones. The number of information bits carried by each codeword is $\log_2 \binom{m}{\frac{m}{2}} = mH(1/2) - O(\log m) = m - O(\log m)$, which can be shown by using Stirling's formula[3]. Since the initial state is $u_{1,1} = 1$, Node 1 is the first to transmit: it sends its information bits, starting with $b_{1,1}$ up until the first bit that equals "1". Specifically, assume that we have $b_{1,1} = b_{1,2} = \cdots b_{1,i_1-1} = 0$ and $b_{1,i_1} = 1$. Thus, in the i_1th channel use the energy unit is transferred to Node 2. From the $(i_1+1)-$th channel use, Node 2 then starts sending its first bit $b_{2,1}$ and the following bits until the first bit equal to "1". The process is then repeated. Since the codewords of the two nodes have the same number of ones, the process completes without stalling in $2m$ channel uses. Hence, the achieved sum-rate is equal to $R_1 + R_2 = \frac{\log_2 \binom{m}{\frac{m}{2}}}{m}$ and tends to 1 bit/c.u. as m tends to infinity.

[3]Stirling's formula leads to the bounds $\frac{1}{m+1} 2^{mH(k/m)} \leq \binom{m}{k} \leq 2^{mH(k/m)}$ see, e.g., [20].

2) $\mathtt{U} > 1$ *Energy Units:* In the sum-capacity strategy discussed above with $\mathtt{U} = 1$ energy unit, both nodes transmit equiprobable symbols "0" and "1". When there are $\mathtt{U} > 1$ energy units in the system, maximizing the sum-capacity generally requires a different approach. Consider the scenario with $\mathtt{U} = 2$ energy units: now it can happen that both energy units are available at one node, say Node 1. While Node 1 would prefer to transmit equiprobable symbols "0" and "1" in order to maximize the *information* flow to the recipient, one must now also consider the *energy* flow: privileging transmission of a "1" over that of a "0" makes it possible to transfer energy to Node 2, leading to a state in which both nodes have energy for the next channel use. This might be beneficial in terms of achievable sum-rate.

Based on this insight, in the following, we propose a coding strategy that employs rate splitting and codebook multiplexing. The strategy is a natural extension of the baseline approach discussed above for the case $\mathtt{U} = 1$. Each Node j constructs $\mathtt{U}$ codebooks, namely $\mathcal{C}_{j|u}$, with $u \in [1, \mathtt{U}]$, where codebook $\mathcal{C}_{j|u}$ is to be used when the Node j has u energy units. Each codebook $\mathcal{C}_{j|u}$ is composed of codewords having approximately a fraction $p_{1|u}$ of "1" symbols[4]. The main idea is that, when the number u of available energy units is large, one might prefer to use a codebook with a larger fraction $p_{1|u}$ of "1" symbols in order to facilitate energy transfer.

Proposition 1. *The rate pair* (R_1, R_2) *satisfying*

$$\begin{aligned} R_1 &\leq \sum_{u=1}^{\mathtt{U}} \pi_u H(X_{1|u}) \\ \textit{and } R_2 &\leq \sum_{u=1}^{\mathtt{U}} \pi_u H(X_{2|u}), \end{aligned} \tag{6}$$

where $X_{j|u} \sim \mathtt{Bern}(p_{j|u})$, $j = 1, 2$, *for some probabilities* $0 < p_{1|u}, p_{2|u} < 1$, $u = 1 \ldots \mathtt{U}$, *with* $p_{1|0} = p_{2|\mathtt{U}} = 0$, *is included in the capacity region* $\mathcal{C}(\mathtt{U})$. *The probabilities* $\pi_u \geq 0$, $u = 0 \ldots \mathtt{U}$, *in (6) satisfy the fixed-point equations*

$$\pi_u = \pi_u(\phi_{0,0|u} + \phi_{1,1|u}) + \pi_{u-1}\phi_{0,1|u} + \pi_{u+1}\phi_{1,0|u} \tag{7}$$

with $\pi_{-1} = \pi_{\mathtt{U}+1} = 0$, $\sum_{u=1}^{\mathtt{U}} \pi_u = 1$, *and we have defined*

$$\begin{aligned} \phi_{0,0|u} &= (1 - p_{1|u})(1 - p_{2|\mathtt{U}-u}) \\ \phi_{0,1|u} &= (1 - p_{1|u})p_{2|\mathtt{U}-u} \\ \phi_{1,0|u} &= p_{1|u}(1 - p_{2|\mathtt{U}-u}) \\ \textit{and } \phi_{1,1|u} &= p_{1|u}p_{2|\mathtt{U}-u}. \end{aligned} \tag{8}$$

This proposition is proved by resorting to random coding arguments, whereby codebook $\mathcal{C}_{j|u}$ is generated with independent and identically distributed (i.i.d.) entries $X_{j|u}$ distributed as $\mathtt{Bern}(p_{j|u})$, $j = 1, 2$. As introduced above, the idea is that, when the state is $U_{1,i} = u$, Node j transmits a symbol from the codebook associated with that state, namely codebook $\mathcal{C}_{1|u}$ for Node 1 and codebook $\mathcal{C}_{2|\mathtt{U}-u}$ for Node 2 (which has $\mathtt{U} - u$ energy units). Both nodes know the current state $U_{1,i}$ and thus can demultiplex the codebooks at the receiver side. According to the random coding argument, the state $U_{1,i}$ evolves according to a Markov chain: the system stays in the same state u with probability $\phi_{0,0|u} + \phi_{1,1|u}$ (both nodes transmit "0" or "1"), changes to the state $u + 1$ with probability $\phi_{1,0|u}$ (Node 1 transmits a "1" and Node 2 a "0") or changes to the state $u - 1$ with probability $\phi_{0,1|u}$ (Node 1 transmits a "0" and Node 2 a "1"). The definition of the conditional probabilities (8) reflects the fact that the codebooks are generated independently by the two nodes. A full proof is given in Appendix A.

[4]Since the proof is based on random coding, the fraction of "1" symbols is close to $p_{1|u}$ as guaranteed by the law of large numbers (see Appendix A for details).

B. Outer Bounds

In this section, we derive an outer bound to the capacity region $\mathcal{C}(\mathtt{U})$. Similar to the standard cut-set bound [19, Ch. 17], the outer bound differs from the inner bound of Proposition 1 in that it allows for a joint distribution $\phi_{x_1,x_2|u}$ of the variables $X_{1|u}$ and $X_{2|u}$.

Proposition 2. *If the rate pair* (R_1, R_2) *is included in the capacity region* $\mathcal{C}(\mathtt{U})$, *then there exist probabilities* $\pi_u \geq 0$ *with* $\sum_{u=1}^{\mathtt{U}} \pi_u = 1$, *and* $\phi_{x_1,x_2|u} \geq 0$ with $\sum_{x_1,x_2 \in \{0,1\}} \phi_{x_1,x_2|u} = 1$ *for all* $u \in \{0, 1, ..., \mathtt{U}\}$, *such that* $\phi_{1,x_2|0} = 0$ *for* $x_2 \in \{0, 1\}$, $\phi_{x_1,1|\mathtt{U}} = 0$ *for* $x_1 \in \{0, 1\}$, *condition (7) is satisfied, and the following inequalities hold*

$$R_1 \leq \sum_{u=0}^{U} \pi_u H\left(X_{1|u} \mid X_{2|u}\right) \tag{9}$$

$$R_2 \leq \sum_{u=0}^{U} \pi_u H\left(X_{2|u} \mid X_{1|u}\right) \tag{10}$$

$$\textit{and } R_1 + R_2 \leq \sum_{u=0}^{U} \pi_u H\left(X_{1|u}, X_{2|u}\right), \tag{11}$$

where variables $X_{1|u}$ *and* $X_{2|u}$ *are jointly distributed with distribution* $\phi_{x_1,x_2|u}$.

The outer bound is proved in Appendix B using information-theoretic inequalities. We remark here that, unlike the achievable strategy described in the previous section, the outer bound is evaluated using a *joint* distribution $\phi_{x_1,x_2|u}$ of the random variables $X_{1|u}$ and $X_{2|u}$ representing the transmitted symbols when the state is $U_{1,k} = u$. In fact, we recall that, in the achievable strategy described in the previous section, the codebooks are generated independently. Intuitively, allowing for a joint distribution $\phi_{x_1,x_2|u}$ leads to an enhanced performance and hence to an outer bound on the achievable rate region. Analyzing the tightness of the inner and outer bounds for arbitrary $\mathtt{U}$ is highly nontrivial, due to the fact that the distribution $\phi_{x_1,x_2|u}$ affects the bounds through the stationary probabilities of the Markov chain. We therefore resort to numerical analysis in the next section.

C. Numerical Results

Fig. 2 compares the achievable sum-rate obtained from Proposition 1 and the upper bound (11) on the sum-rate obtained from Proposition 2 versus the total number of energy units $\mathtt{U}$. As for the achievable sum-rate, we consider both a conventional codebook design in which the same probability $p_{j|u} = 0.5$ is used irrespective of the state $U_{1.i} = u$, and

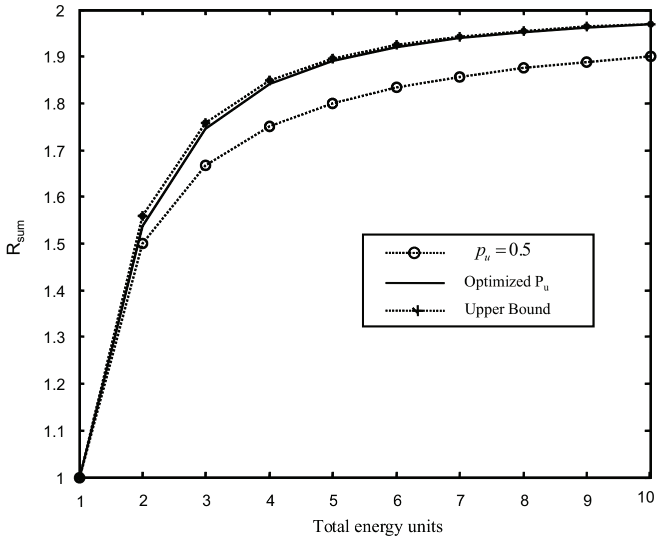

Fig. 2. Achievable sum-rate obtained from Proposition 1 and upper bound (11) versus the total number of energy units U.

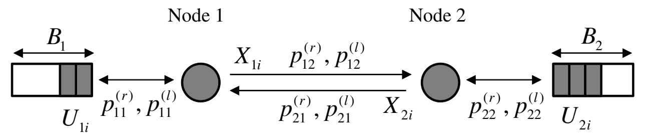

Fig. 3. Two-way noisy binary communication with energy exchange. The probabilities of replenishments through the channel or locally at the nodes are referred to as $p_{ij}^{(r)}$, with $i \neq j$ or $p_{ij}^{(r)}$, with $i = j$, respectively, and similarly for the probabilities of losses $p_{ij}^{(l)}$. See Fig. 4 for an illustration of the channel and Fig. 5 for an illustration of the harvesting process.

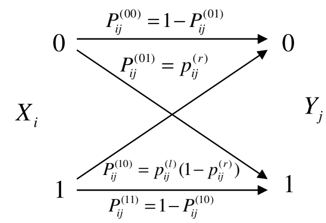

Fig. 4. Channel from Node i to Node j.

one in which the probabilities $p_{j|u}$ are optimized. It can be seen that using conventional codebooks, which only aim at maximizing information flow on a single link, leads to substantial performance loss. Instead, the proposed strategy with optimized probabilities $p_{j|u}$, which account also for the need to manage the energy flow in the two-way communication system, performs close to the upper bound. The latter is indeed achieved when U is large enough.

A remark on the optimal probabilities $p_{j|u}$ is in order. Due to symmetry, it can be seen that we have $p_{1|u} = p_{2|\mathtt{U}-u}$. Moreover, numerical results show that $p_{1|u}$ increases monotonically as u goes from 0 to U, such that $p_{1,\mathtt{U}} > 0.5$. In particular, when the number of states $\mathtt{U}+1$ is odd, it holds that $p_{1,\mathtt{U}/2} = p_{2,\mathtt{U}/2} = 0.5$. It is finally noted that the energy neutral transitions (both nodes emitting "0" or both emitting "1") occur with equal probability (i.e., $(1-p_{1,u})(1-p_{2,u}) = p_{1,u}p_{2,u}$).

IV. System Model with Stochastic Replenishments and Losses

In this section we extend the two-way communication system with energy exchange studied above to include energy losses and replenishments, which may occur in different parts of the system, as illustrated in Fig. 3. Specifically, the energy units can be lost either while in transit through a lossy channel or locally at either node during processing. The first event is due to, e.g., path loss or fading, while the second is due to the inefficiencies of the energy storage system, see, e.g., [21]. Similarly, energy units can be replenished either by harvesting energy from the channel, e.g., from an interfering signal or a source of RF energy [22], or through a source of power locally connected to the node, e.g., a solar panel. All loss and replenishment events are assumed to be independent. As above, we assume that the two parties involved have a common clock, and that, at each time, a node can either send a "1", which requires one unit of energy cost, or a "0", which does not require any energy expenditure. We also assume that Node 1 and Node 2 have energy buffers of capacities B_1 and B_2 energy units, respectively, to store the available energy.

Unlike in the previous sections, we assume that the binary channel from Node i to Node j with $i \neq j$, is noisy as shown in Fig. 4, with the probability of $P_{ij}^{(01)}$ of flipping a "0" symbol to a "1" symbol and the probability $P_{ij}^{(10)}$ of flipping symbol "1" to symbol "0". These probabilities can be interpreted in terms of replenishments and losses across the channel. To elaborate, let us define as $p_{ij}^{(r)}$ the probability of replenishment via harvesting from the channel (for $i \neq j$), e.g., thanks to an RF source that operates on the same bandwidth as the ij link. Moreover, define as $p_{ij}^{(l)}$ the probability that an energy unit is lost while in transmit through the channel for the ij link. With these definitions, assuming that losses and replenishments are independent, we can write the transition probabilities as in Fig. 4.

Losses and replenishments can also take place locally at the nodes with the probability $P_{ii}^{(01)}$ of flipping a "0" symbol to a "1" symbol and the probability $P_{ii}^{(10)}$ of flipping symbol "1" to symbol "0" at Node i upon reception. Specifically, let us define as $p_{ii}^{(r)}$ the probability of replenishment at Node i, whereby an energy unit is received by Node i from an external source of energy directly connected to the node, such as a solar panel. Note that this energy unit is not received through the channel but is directly stored in the buffer and therefore does not affect the decoder, unlike replenishment events over the channel. We emphasize that the model limits the peak harvested energy to one energy unit. Moreover, define as $p_{ii}^{(l)}$ the probability that an energy unit, while correctly received by the decoder at Node i, is lost during processing before reaching the energy buffer. Note that in this case the decoder at Node i correctly records a "1", but this energy unit cannot be reused for future channel uses. This event is thus different from a loss over the channel in which the decoder at Node i observes a "0" symbol. With these definitions, assuming that losses and replenishments are independent and that no more

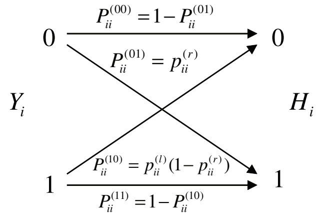

Fig. 5. Statistical relationship between the received signal Y_i and the energy H_i harvested by Node i.

than one energy unit can be harvested in each time instant, we can write the transition probabilities between the received signal Y_i and the harvested energy H_i at Node i as in Fig. 5.

Based on the discussion above, at any given time instant k, with $k \in [1, n]$, the *state* of the system $(U_{1,k}, U_{2,k}) \in \mathbb{N}^2$ is given by the current energy levels $U_{1,k}$ and $U_{2,k}$ in the buffers of Node 1 and Node 2, respectively. By the capacity limitations of the buffers, we have $u_1 \in [0, B_1]$ and $u_2 \in [0, B_2]$ for each channel use $k \in [1, n]$. The transmitted symbols are limited as per (1).

The channel is noisy with transition probabilities as in Fig. 3. Moreover, the relationship between received signal and harvested energy is as in Fig. 5. Therefore, the state of battery at Node 1 for $k \in [1, n]$ evolves as follows

$$U_{i,k} = (U_{i,k-1} - X_{i,k-1}) + H_{i,k-1}. \tag{12}$$

Similar to Sec. II, we use the following definitions for an (n, R_1, R_2, B_1, B_2). Specifically, the code is defined by: the buffer capacities B_1 and B_2; two sequences of encoding functions, $f_{i,k}$ for $k \in [1, n]$ and $i = 1, 2$, which map the message M_i, the past received symbols Y_i^{k-1} along with the past and current states (U_1^k, U_2^k) into the currently transmitted signal $X_{i,k} \in \mathcal{X}_{U_{i,k}}$; two decoding functions g_i, for $i = 1, 2$, which map the received signal Y_i^n, the local message M_i and the sequence of states U_1^n, U_2^n into an estimate $\hat{M}_j$ of message M_j of the other node $j \neq i$. Achievability is defined as in Sec. II. Finally, the closure of the set of all the rate pairs (R_1, R_2) is defined as the capacity region $\mathcal{C}_0(B_1, B_2)$, where the subscript "0" denotes the fact that the capacity region generally depends on the initial state and we have made explicit the dependence on the battery sizes (B_1, B_2). Note that the capacity region $\mathcal{C}_0(B_1, B_2)$ is non-decreasing with respect to both (B_1, B_2), since a node can always discard any additional battery capacity and thus achieve the same performance as with a smaller battery. The performance advantage of using large buffers arises from the possibility to better manage the energy received through replenishment by avoiding battery overflow. This, in turn, allows the nodes to focus the optimization of their transmission strategies towards information transfer rather than energy transfer. Related numerical results can be found in Sec. V-D.

Remark 3. In the definition of code given above, we have assumed that the nodes can track the state of the batteries $(U_{1,k}, U_{2,k})$ at both nodes. We refer to this scenario as having Global Energy Information (GEI). We remark that in the presence of losses and replenishment, the nodes generally cannot track the amount of energy available at the other node based only on the knowledge of the received signal. Instead, information about the state of the other node needs to be acquired through additional resources such as control channels or appropriate sensors. In general, the assumed model with GEI can thus be thought of providing a best-case bound on system performance. In Sec. V-C, we will study the scenario, referred to as having Local Energy Information (LEI), in which each node is only aware of the energy available in its own local battery.

V. Generalizing the Inner and Outer Bounds

In this section, we first propose a communication strategy that leads to an achievable rate region by generalizing the approach discussed in Sec. III. The outer bound of Sec. III is similarly extended. While the strategy at hand is based on GEI (see Remark 3), we then discuss an achievable strategy with LEI in Sec. V-C, and present some numerical results in Sec. V-D.

A. Transmission Strategy

The proposed strategy is an extension of the approach put forth in Sec. III, and operates as follows. Node i, with $i = 1, 2$, constructs one independent codebook $\mathcal{C}_{i|(u_1,u_2)}$ for each possible state $(u_1, u_2) \in [0, B_1] \times [0, B_2]$. As in Sec. III, at each time k, if the state is $(U_{1,k}, U_{2,k}) = (u_1, u_2)$, then Node i transmits the next symbol from the codebook $\mathcal{C}_{i|(u_1,u_2)}$. At the end of the last channel use, each node, being aware of the sequences of states, can demultiplex the transmission of the other node and decode the messages encoded in all the $(B_1 + 1)(B_2 + 1)$ codebooks.

The codebook of Node i corresponding to state (u_1, u_2) is generated by drawing each bit independently with a given probability $p_{i|(u_1,u_2)}$ for $i = 1, 2$ and all states $(u_1, u_2) \in [0, B_1] \times [0, B_2]$. Note that, due to (1), we have $p_{1|(0,u_2)} = 0$ for all $u_2 \in [0, B_2]$ since, when $U_{1,k} = 0$, Node 1 has no energy available and thus must transmit a "0" symbol; and similarly we have $p_{2|(u_1,0)} = 0$ for all $u_1 \in [0, B_1]$. Given the probabilities $p_{i|(u_1,u_2)}$ for $i = 1, 2$ and all states $(u_1, u_2) \in [0, B_1] \times [0, B_2]$, the $(B_1 + 1)(B_2 + 1) \times (B_1 + 1)(B_2 + 1)$ transition probability matrix $\boldsymbol{P}$ can be obtained that contains the transition probabilities from any state $(u_1, u_2) \in [0, B_1] \times [0, B_2]$ to any state $(u_1', u_2') \in [0, B_1] \times [0, B_2]$. These transition probabilities depend on the parameters $(p_{ij}^{(r)}, p_{ij}^{(l)})$, $(p_{ii}^{(r)}, p_{ii}^{(l)})$, and $p_{i|(u_1,u_2)}$ for $i, j = 1, 2$ and $(u_1, u_2) \in [0, B_1] \times [0, B_2]$, as detailed in Appendix C.

B. Inner and Outer Bounds

In order to derive the rates achievable with this strategy, denote as $\pi_{(u_1,u_2)}$ the average fraction of channel uses k such that we have $(U_{1,k}, U_{2,k}) = (u_1, u_2)$ for all states $(u_1, u_2) \in [0, B_1] \times [0, B_2]$, as done in Sec. III. Note that $\sum_{(u_1,u_2)} \pi_{(u_1,u_2)} = 1$. This function is also referred to as the steady-state probability and can be calculated as the limit

$$\boldsymbol{\pi}_{(u_1,u_2)} = \lim_{k \to \infty} \boldsymbol{P}^k \boldsymbol{\pi}(1), \tag{13}$$

where $\boldsymbol{\pi}_{(u_1,u_2)}$ is the $(B_1 + 1)(B_2 + 1) \times 1$ vector containing the steady-state probabilities $\pi_{(u_1,u_2)}$ for all states

$(u_1, u_2) \in [0, B_1] \times [0, B_2]$ and we recall that $\boldsymbol{P}$ is the transition probability matrix. Vector $\boldsymbol{\pi}(1)$ accounts for the initial state and is thus a vector of all zeros except for a one in the entry corresponding to the initial state. We note that the limit in (13) always exists for the model studied in Sec. III (for all non-trivial transmission probabilities), and is given by (7)-(8). The same is generally true here apart from degenerate cases. However, the transition matrix (13) is possibly reducible, and thus the calculation of the limit generally requires the factorization of the matrix according to the canonical form for reducible matrices. We refer to [26, ch. 8] for a detailed discussion on the existence and calculation of the limit (13).

Proposition 4. *Assuming that the limit (13) exists, the rate pair (R_1, R_2) satisfying the inequalities*

$$\begin{aligned} R_1 &\leq \sum_{\substack{(u_1,u_2)\in \\ [0,B_1]\times[0,B_2]}} \pi_{(u_1,u_2)} I(X_{1|(u_1,u_2)}; Y_2) \\ \text{and } R_2 &\leq \sum_{\substack{(u_1,u_2)\in \\ [0,B_1]\times[0,B_2]}} \pi_{(u_1,u_2)} I(X_{2|(u_1,u_2)}; Y_1) \end{aligned} \tag{14}$$

for some transmission probabilities $p_{i|(u_1,u_2)}$, for $i = 1, 2$ and $(u_1, u_2) \in [0, B_1] \times [0, B_2]$ is achievable, where we have denoted as $X_{i|(u_1,u_2)}$ as the Bernoulli variable $\mathtt{Bern}(p_{i|(u_1,u_2)})$. We also have

$$\begin{aligned} I(X_{1|(u_1,u_2)};Y_2) = &H\big((1-p_{1|(u_1,u_2)})P_{01} + p_{1|(u_1,u_2)}P_{11}\big) \\ &-\big[p_{1|(u_1,u_2)}H(P_{11}) + (1-p_{1|(u_1,u_2)})H(P_{01})\big] \end{aligned} \tag{15}$$

and similarly for $I(X_{2|(u_1,u_2)}; Y_1)$.

Remark 5. The achievability of the rates in (14) can be proved by adopting the multiplexing strategy described above and following the same main steps as in Appendix A. Here, we also point out that the achievability of (14) under the assumption that the limit (13) exists is a direct consequence of [27, Lemma 12.3.1].

An outer bound can be also derived by generalizing Proposition 2. In particular, following similar steps as in Appendix B, one can prove that an outer bound is obtained by allowing for joint probabilities, rather than product distributions as in Proposition 1. Moreover, one can add the sum-rate constraint that generalizes (11) as

$$R_1 + R_2 \leq \sum_{\substack{(u_1,u_2)\in \\ [0,B_1]\times[0,B_2]}} \pi_{(u_1,u_2)} I(X_{1|(u_1,u_2)}, X_{2|(u_1,u_2)}; Y_1, Y_2), \tag{16}$$

where $X_{1|(u_1,u_2)}, X_{2|(u_1,u_2)}$ are jointly distributed.

C. Local Energy Information

In the discussion above, we have assumed GEI, that is, each node knows the full current energy state $(U_{1,k}, U_{2,k})$ (see Remark 3). In this section, we consider instead the scenario with LEI, in which Node 1 only knows its local energy level U_1 and Node 2 only knows U_2.

We first observe that the energy $U_{1,k}$ can be considered to be the state of the link 12 at channel use k, since it affects the available input symbols via (1) (and similarly for $U_{2,k}$ and link 21). Therefore, the model at hand falls in the category of channels with states in which the state is known only at the transmitter. For these channels, under the assumption that the state sequence is i.i.d. and independent of the transmitted signal, it is known that so called Shannon strategies are optimal [19, Ch. 7]. In the model under study, unlike the conventional setting, the state sequence U_1^n (and U_2^n) is neither i.i.d. nor independent of the transmitted signal X_1^n (and X_2^n). Therefore, Shannon strategies are generally not optimal. We will see below that they can be nevertheless used to lead to non-trivial achievable rates. A related approach was proposed in [23] in the context of energy-harvesting systems with no batteries.

Following Shannon strategies, we draw auxiliary codebooks made of independent and i.i.d. codewords V_1^n and V_2^n using pmfs $p(v_1)$ and $p(v_2)$, respectively. Each symbol $V_{j,k}$ for Node j and time instant k is a vector consisting of B_j bits. The main idea is that, at each time k, Node j transmits the bit in $V_{j,k}$ corresponding to the current state $U_{j,k}$. Note that the latter can take B_j possible values at which the transmitted signal is non-trivial (for $U_{j,k} = 0$, we necessarily have $X_{j,k} = 0$).

At the receiver side, the decoder at Node 2 uses joint typicality decoding with respect to the distribution $p(v_1, y_2)$, which is given as

$$p(v_1, y_2) = p(v_1)\sum_{u_1} \pi(u_1) p(y_2 | f_1(v_1, u_1)) \tag{17}$$

where $\pi(u_1)$ is the marginal distribution of the steady-state probability of the Markov chain induced by the random coding strategy and the evolution of the system, as discussed above (see also Appendix C). Following standard information theoretic considerations, we obtain that the rate pair (R_1, R_2) satisfying

$$R_1 \leq I(V_1; Y_2), \tag{18a}$$

$$\text{and } R_2 \leq I(V_2; Y_1), \tag{18b}$$

for some pmfs $p(v_1)$, $p(v_2)$ is achievable, where $p(v_1, y_2)$ is as in (17) and similarly for $p(v_2, y_1)$. Regarding the details of the proof, being based on conventional tools (see [19, Ch. 3]), here we simply point out that it is based on the ergodicity of the Markov chain, which allows to conclude that the error event in which the correct codeword is not jointly typical takes place with negligible probability; and the packing lemma in [19, Lemma 3.1], which entails that the error events due to mistaking other codewords for the correct one have also negligible probability[5].

Remark 6. In the strategy proposed above, each node adapts the choice of the current transmitted symbol only to the current local energy state. A potentially better approach would be to perform adaptation based on a local state that includes also a number of past energy states of the node, along with the current one, and/or current and past received signals. This aspect is not further explored in this paper.

[5]The packing lemma does not assume that the received signal be i.i.d. and thus applies to our scenario (see [19, Lemma 3.1]).

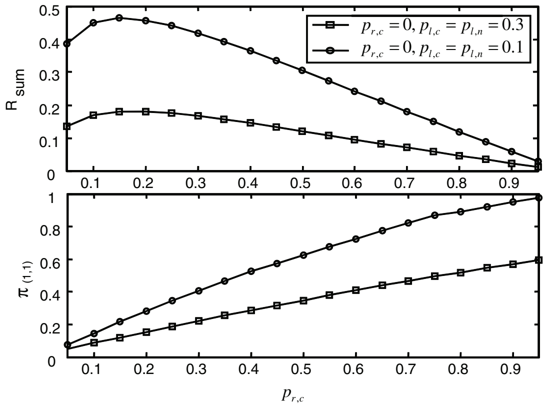

Fig. 6. Sum-rate R_{sum} and steady-state probability $\pi_{(1,1)}$ versus the probability $p_{r,c}$ of replenishment on the channel (see Fig. 4).

D. Numerical Results

In this section, we present some numerical examples in order to assess the impact of replenishment and loss processes. Unless stated otherwise, we will assume that each node has the ability to store only one unit of energy i.e., $B_1 = B_2 = 1$. We consider a symmetric system with $p_{12}^{(r)} = p_{21}^{(r)} = p_{r,c}$, $p_{11}^{(r)} = p_{22}^{(r)} = p_{r,n}$, $p_{12}^{(l)} = p_{21}^{(l)} = p_{l,c}$ and $p_{11}^{(l)} = p_{21}^{(l)} = p_{l,n}$, where the subscripts "c" and "n" stand for "channel" and "node", so that, e.g., $p_{r,n}$ is the probability of replenishment locally at a node. We first assume GEI.

Fig. 6 shows the sum-rate obtained by summing the right-hand sides of (14), optimized over the probabilities $p_{1|(u_1,u_2)}$ and $p_{2|(u_1,u_2)}$ for all states $(u_1, u_2) \in [0, B_1] \times [0, B_2]$ versus the replenishment probability on the channel $p_{r,c}$ (see Fig. 4) for two cases, namely $p_{r,n} = 0, p_{l,n}, p_{l,c} = 0.1$ and $p_{r,n} = 0, p_{l,n}, p_{l,c} = 0.3$. We also show in the same figure the steady-state probability $\pi_{(1,1)}$ of state $(u_1, u_2) = (1, 1)$ corresponding to the optimal values of $p_{1|(u_1,u_2)}$ and $p_{2|(u_1,u_2)}$. It is seen that increasing the probability $p_{r,c}$ increases the chance of being in state $(u_1, u_2) = (1, 1)$, due to the increased availability of energy. However, increasing $p_{r,c}$ has also the deleterious effect of flipping bits on the channel from "0"s to "1"s with larger probability. It is seen that, in the regime in which $p_{r,c}$ is sufficiently small, and the system is energy-limited, increasing $p_{r,c}$ is beneficial, while for $p_{r,c}$ large enough the second effect dominates and the achievable sum-rate decreases.

We now turn to assessing the effect of local replenishment at the node and the effect of the size of the energy buffers. Fig. 7 shows the optimized sum-rate versus $p_{r,n}$. Both nodes have identical buffer of size B, where $B = 1$ or $B = 2$. Increasing $p_{r,n}$ improves the sum-rate, since it enhances the probability of being in the state of full batteries, without any side effect on the channel quality. Moreover, having a larger buffer size increases the rates especially for small-to-intermediate values of $p_{r,n}$, since, as discussed in Sec. IV , a larger buffer enables a better management of the replenished energy. In particular, when the replenishment probability $p_{r,n}$ is large enough, the batteries tend to be full all the time, as seen in the lower part of Fig. 7, and thus there is no need for a more complex energy management, leading to reducing gains from having large batteries.

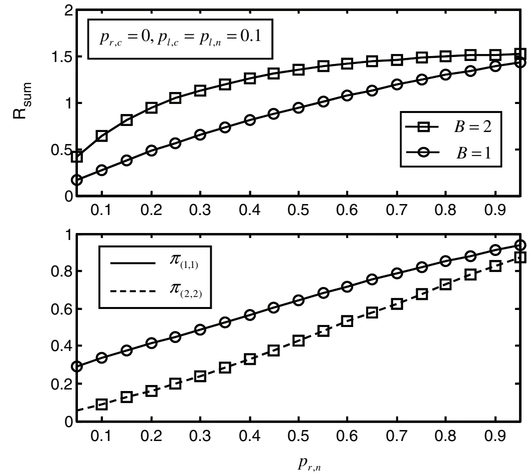

Fig. 7. Sum-rate R_{sum} and steady-state probability $\pi_{(1,1)}$ versus the probability $p_{r,n}$ of replenishment at the node (see Fig. 5). Both nodes have buffer of size B.

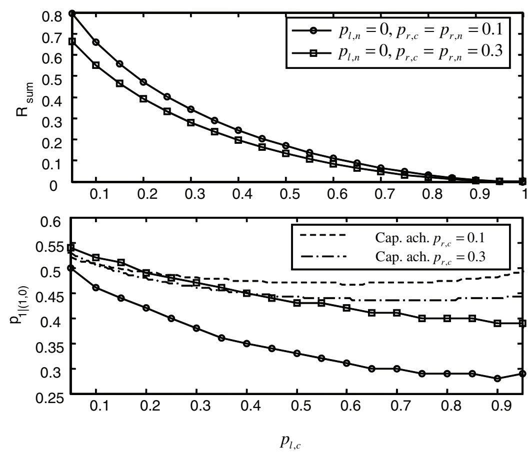

Fig. 8. Sum-rate R_{sum} and optimum transmission probability $p_{1|(1,0)}$ versus the probability $p_{l,c}$ of loss on the channel. Dotted lines show the capacity achieving probability.

Fig. 8 and Fig. 9 show the effect of loss events on the channel and at the nodes, respectively. We show both the sum-rate and the optimal transmission probability $p_{1|(1,0)}$, which equals the optimal probability $p_{2|(0,1)}$ by symmetry. The latter is also compared with the transmission probability that maximizes the mutual information $I(X_{1|(0,1)}; Y_2)$ in (14) and that is thus capacity achieving. It is noted that this is the probability that maximizes the information rate when there are no energy limitations. As it can be seen, by comparing Fig. 8 and Fig. 9, increasing the loss probability both on the channel and at the node decreases the sum-rate, although the rate of this decrease is larger for the latter, since, similar to the discussion above, a loss at the node does not affect the channel. Moreover, for small $p_{l,c}$ and $p_{l,n}$, the transmission probability $p_{1|(1,0)}$ is close to the capacity-achieving probability, while for larger loss probabilities $p_{l,c}$ and $p_{l,n}$, it becomes smaller than the capacity-achieving probability.

We now consider the effect of LEI. Fig. 10 compares the sum-rate achieved with GEI and LEI versus the replenishment probability $p_{r,c}$ on the channel. As it can be seen, LEI entails a

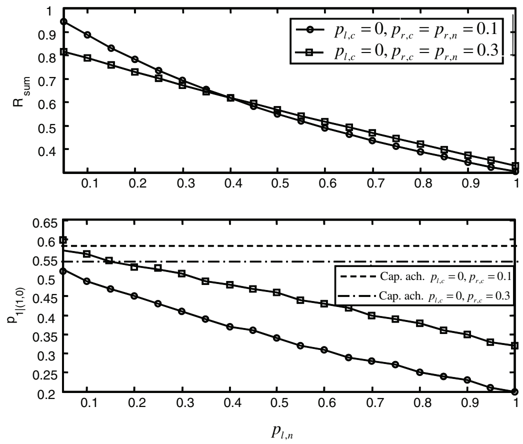

Fig. 9. Sum-rate R_{sum} and optimum transmission probability $p_{1|(1,0)}$ versus the probability $p_{l,n}$ of loss at the node. Dotted lines show the capacity achieving probability.

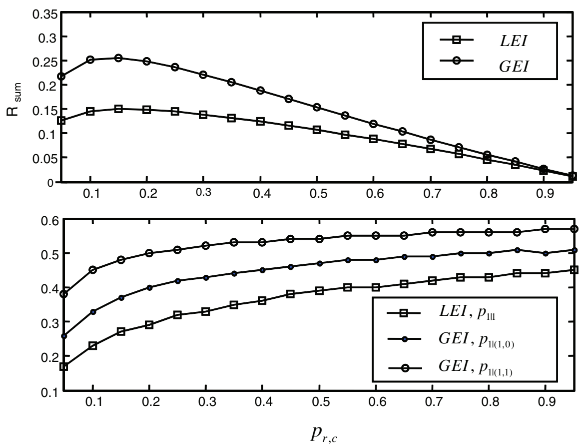

Fig. 10. Sum-rate R_{sum} and optimum transmission probabilities for Global Energy Information (GEI) and Local Energy Information (LEI) for $p_{l,c} = 0.3$ and $p_{r,n} = p_{l,n} = 0$.

significant performance loss with respect to GEI. To gain some insight as to the reasons of this loss, the figure also shows the optimal transmission probabilities $p_{1|(1,0)}$, $p_{1|(1,1)}$ with GEI and the probability $p(v_1) = p_{1|1}$, that is the probability of transmitting "1" if the local battery contains energy, for LEI (V_1 is a Bernoulli variable since $B_1 = B_2 = 1$). With GEI, the nodes can adapt the transmission strategy to the energy state of both nodes and thus choose different probabilities $p_{1|(1,0)}$ and $p_{1|(1,1)}$, while with LEI the nodes are forced to choose a single probability $p_{1|1}$ irrespective of the state of the battery at the other node.

VI. Conclusions

Energy and information content are two contrasting criteria in the design of a communication signal. In a number of emerging and envisaged communication networks, the participating nodes are able to reuse part of the energy in the received signal for future communication tasks. Therefore, it becomes critical to develop models and theoretical insights into the involved trade-offs between energy and information exchange at a system level. In this work, we have taken a first step in this direction, by considering a two-way channel under a simple binary “on-off” signaling model. The derived inner and outer bounds shed light into promising transmission strategies that adapt to the current energy state. It is emphasized that conventional strategies based solely on the maximization of the information flow entail substantial losses.

The results presented in this paper call for further studies on different fronts. One is the development of better models which strike a good balance between adherence to reality and analytical tractability. As an example, more complex models could account for non-orthogonal communication links in which the resoures used in the two directions can interact, as for particles possibly colliding. Also, larger input and output alphabets, along with a larger state space for the state of the batteries, could be considered in order to better model aspects such as path loss. Finally, as pointed out in [24], constraints on the receiver design favor structures in which the receiver operates either as an information decoder or as an energy harvester – this aspect could also be included in a more refined analysis. A second front of investigation is the development of better communication strategies for the practical scenario in which the energy state of the network is not fully known at the nodes.

Appendix A
Proof of Proposition 1

1) Code construction: We generate $\mathtt{U}$ codebooks for each Node $j = 1, 2$, namely $\mathcal{C}_{j|u}$, with $u \in [1, \mathtt{U}]$. The codebook $\mathcal{C}_{j|u}$ for $u > 0$ has $K_{j,u}$ codewords, each consisting of $n_{j,u}$ symbols $\tilde{x}_{j,u,l} \in \{0, 1\}$, which are randomly and independently generated as $\mathtt{Bern}(p_{j|u})$ variables, with $l = 1, 2, ..., n_{j,u}$ and $n_{j,u} = n\delta_{j,u}$, for some $0 \leq \delta_{j,u} < 1$. We denote the codewords as $\tilde{x}_{j,u}^{n_{j,u}}(m_{j,u})$ with $m_{j,u} \in [1, K_{j,u}]$. Note that the parameter $\delta_{j,u}$ does not depend on n, and hence, if $n \to \infty$, then we have $n_{j,u} \to \infty$ for all j, u. We set $2^{nR_j} = \prod_{u=1}^{\mathtt{U}} K_{j,u}$, while the relations among the remaining parameters ($K_{j,u}$,$\delta_{j,u}$,$p_{j|u}$) will be specified below.

2) Encoding: Each node performs rate splitting. Namely, given a message $M_j \in [1, 2^{nR_j}]$, Node j finds a $\mathtt{U}-$tuple $[m_{j,1}, ..., m_{j,\mathtt{U}}]$ with $m_{j,u} \in [1, K_{j,u}]$ that uniquely represents M_j. This is always possible since we have $2^{nR_j} = \prod_{u=1}^{\mathtt{U}} K_{j,u}$. Then, the selected codewords $\tilde{x}_{j,u}^{n_{j,u}}(m_{j,u})$ for $u \in [1, \mathtt{U}]$ are transmitted via multiplexing based on the current available energy. Specifically, each Node j initializes $\mathtt{U}$ pointers $l_{j,1} = l_{j,2} = \cdots = l_{j,\mathtt{U}} = 1$ that keep track of the number of symbols already sent from codewords $\tilde{x}_{j,1}^{n_{j,1}}(m_{j,1})$, $\tilde{x}_{j,2}^{n_{j,2}}(m_{j,2})$, ..., $\tilde{x}_{j,\mathtt{U}}^{n_{j,\mathtt{U}}}(m_{j,\mathtt{U}})$, respectively. At channel use i, if the state is $U_{1,i} = u$, then the nodes operate as follows.

- Node 1: If $u = 0$, then $x_{1,i} = 0$. Else, if $l_{1,u} \leq n_{1,u}$, Node 1 transmits $x_{1,i} = \tilde{x}_{1,u,l_{1,u}}(m_{1,u})$ and increments the pointer $l_{1,u}$ by 1. Finally, if $l_{1,u} = n_{1,u}+1$ the pointer $v_{1,u}$ is not incremented, and the transmitter uses *random padding*, i.e., it sends $x_{1,i} = 1$ with probability $p_{1,u}$ and $x_{1,i} = 0$ otherwise.
- Node 2: If $u = \mathtt{U}$ (i.e., no energy is available at Node 2), then $x_{2,i} = 0$. Else, if $l_{2,\mathtt{U}-u} \leq n_{2,\mathtt{U}-u}$, Node 2 transmits $x_{2,i} = \tilde{x}_{2,\mathtt{U}-u,2,l_{2,\mathtt{U}-u}}(m_{2,\mathtt{U}-u})$ and increments the pointer $l_{2,\mathtt{U}-u}$ by 1. Finally, if $l_{2,\mathtt{U}-u} = n_{2,\mathtt{U}-u} + 1$,

the pointer $l_{2,\mathtt{U}-u}$ is not incremented, and Node 2 sends $x_{2,i}=1$ with probability $p_{2,\mathtt{U}-u}$ and $x_{2,i}=0$ otherwise. The random padding method used above is done for technical reasons that will be clarified below.

3) Decoding: We first describe the decoding strategy for Node 2. By construction, the nodes are aware of the state sequence U_1^n, and thus can determine the ordered set

$$\mathcal{N}_u=\{i|U_{1,i}=u\}, \tag{19}$$

of channel use indices in which the state is u with $u\in[0,\mathtt{U}]$. For all $u\in[1,\mathtt{U}]$, if $|\mathcal{N}_u|\geq n_{1,u}$, then Node 2 takes the first $n_{1,u}$ indices $i_{u,1}<i_{u,2}<\cdots<i_{u,n_{1,u}}$ from the set $\mathcal{N}_u$ and obtains the list of messages $m_{1,u}\in[1,K_{1,u}]$ that satisfy $\tilde{x}_{1,u,k}(m_{1,u})=x_{1,i_{u,k}}$ for all $k\in[1,n_{1,u}]$. Note that the list cannot be empty due to the fact that the channel is noiseless. However, it contains more than one message, or if $|\mathcal{N}_u|<n_{1,u}$, then Node 2 puts out the estimate $\hat{m}_{1,u}=1$. Instead, if the list contains only one message $m_{1,u}$, then Node 2 sets $\hat{m}_{1,u}=m_{1,u}$. Finally, the message estimate is obtained as $\hat{m}_1=[\hat{m}_{1,1},...,\hat{m}_{1,\mathtt{U}}]$.

Node 1 operates in the same way, with the only caveat that the uth codebook $\mathcal{C}_{2|u}$ of Node 2 is observed at channel uses in the set $\mathcal{N}_{U-u}$ for $u\in[1,\mathtt{U}]$.

4) Analysis: We evaluate the probability of error on average over the messages and the generation of the codebooks, following the random coding principle. From the definition of the decoders given above, the event that any of the decoders is in error is included in the set $\mathcal{E}=\bigcup_{j=1,2}\bigcup_{u=1}^{\mathtt{U}}(\mathcal{E}_{j,u}^{(1)}\cup\mathcal{E}_{j,u}^{(2)})$, where: (*i*) $\mathcal{E}_{j,u}^{(1)}$ is the event that $|\mathcal{N}_u|<n_{1,u}$ for $j=1$ and that $|\mathcal{N}_{\mathtt{U}-u}|<n_{2,u}$ for $j=2$, that is, that the number of channel uses in which the system resides in the state in which the codeword $\tilde{x}_{j,u}^{n_{j,u}}(m_{j,u})$ from the codebook $\mathcal{C}_{j,u}$ is sent is not sufficient to transmit the codeword in full; (*ii*) $\mathcal{E}_{j,u}^{(2)}$ is the event that two different messages $m'_{j,u},m''_{j,u}\in[1,K_{j,u}]$ are represented by the same codewords, i.e., $\tilde{x}_{j,u}^{n_{1,u}}(m'_{j,u})=\tilde{x}_{1,u}^{n,1u}(m''_{1,u})$.

The probability of error can thus be upper bounded as

$$\Pr[\mathcal{E}]\leq\sum_{j=1}^{2}\sum_{u=1}^{\mathtt{U}}\left(\Pr[\mathcal{E}_{j,u}^{(1)}]+\Pr[\mathcal{E}_{j,u}^{(2)}]\right). \tag{20}$$

In the following, we evaluate upper bounds on this terms.

It immediately follows from the packing lemma of [19] that $\Pr[\mathcal{E}_{j,u}^{(2)}]\to 0$ as $n_{j,u}\to\infty$ as long as

$$\frac{\log_2 K_{j,u}}{n_{j,u}}<H(p_{j|u})-\delta(\epsilon) \tag{21}$$

with $\delta(\epsilon)\to 0$ as $\epsilon\to 0$. For analysis of the probabilities $\Pr[\mathcal{E}_{j,u}^{(1)}]$, we observe that, under the probability measure induced by the described random codes, the evolution of the state $U_{1,i}$ across the channel uses $i\in[1,n]$ is a Markov chain with $\mathtt{U}+1$ states. Specifically, the chain is a birth-death process, since, if the state is $U_{1,i}=u$ in channel use i, the next state $U_{1,i+1}$ can only be either $u-1$ or $u+1$. More precisely, let $q_{u|w}=\Pr(U_{1,i+1}=u|U_{1,i}=w)$ be the transition probability. Note that, due to the use of random padding, the transition probability $q_{u|w}$ remains constant during all n channel uses, so that the Markov chain is time-invariant.

We now elaborate on the Markov chain $U_{1,i}$. To this end, we first define as $\phi_{x_1,x_2|u}$, where $x_1,x_2\in\{0,1\}$ be the joint probability that Node 1 transmits $X_{1,i}=x_1$ and Node 2 transmits $X_{2,i}=x_2$ during the ith channel use in which the state is $U_{1,i}=u$. We can now write the non-zero values of the transition probability $q_{u|w}$ as follows:

$$\begin{aligned} q_{u,u-1}=\phi_{1,0|u}\quad q_{u,u+1}&=\phi_{0,1|u}\\ q_{u,u}=1-q_{u,u-1}&-q_{u,u+1} \end{aligned} \tag{22}$$

With a slight abuse of the notation and noting that $\phi_{1,0|0}=\phi_{1,1|0}=0$ and $\phi_{0,1|\mathtt{U}}=\phi_{1,1|\mathtt{U}}=0$ the expressions above also represent the transitions for the two extremal states $u=0$ and $u=\mathtt{U}$, as they imply $q_{0|-1}=0$ and $q_{\mathtt{U}|\mathtt{U}+1}=0$.

If $p_{1,0}=p_{2,0}=0$ and $0<p_{1,u},p_{2,u}<1$ for all $u>0$, then it can be seen that the Markov chain is aperiodic and irreducible, and thus there exist a unique set of stationary probabilities $\pi_0,\pi_1,\cdots,\pi_{\mathtt{U}}$, which are given by solving the linear system, defined by taking $\mathtt{U}$ equations of type (7) for $u=0\ldots\mathtt{U}-1$ and adding the condition $\sum_{u=0}^{\mathtt{U}}\pi_u=1$.

We are now interested in the statistical properties of the set $|\mathcal{N}_u|$ of channel uses in which the state satisfies $U_1=u$. Using the ergodic theorem and the strong law of large numbers [25, Theorem 1.10.2], it can be shown that $\lim_{n\to\infty}\frac{V_u(n)}{n}=\pi_u$ with probability 1. Therefore, if we choose:

$$l_{1,u}=l_{2,\mathtt{U}-u}=n(\pi_u-\epsilon) \tag{23}$$

then $\Pr[\mathcal{E}_{1,u}^{(2)}]=\Pr[\mathcal{E}_{2,\mathtt{U}-u}^{(2)}]$ can be made arbitrarily close to 0 as $n\to\infty$. This concludes the proof.

Appendix B
Proof of Proposition 2

Consider any $(n,R_1,R_2,\mathtt{U})$ code with zero probability of error, as per our definition of achievability in Sec. II. We have the following inequalities:

$$\begin{aligned} nR_1&=H(M_1)=H(M_1|M_2,U_{1,1}=u_{1,1})\\ &\overset{(a)}{=}H(M_1,X_1^n,U_1^n|M_2,U_{1,1}=u_{1,1})\\ &\overset{(b)}{=}H(X_1^n,U_1^n|M_2,U_{1,1}=u_{1,1})\\ &=\sum_{i=1}^{n}H(X_{1,i},U_{1,i}|X_1^{i-1},U_1^{i-1},M_2,U_{1,1}=u_{1,1})\\ &=\sum_{i=1}^{n}H(U_{1,i}|X_1^{i-1},U_1^{i-1},M_2,U_{1,1}=u_{1,1})\\ &\quad+H(X_{1,i}|X_1^{i-1},U_1^{i},M_2,U_{1,1}=u_{1,1})\\ &\overset{(c)}{=}\sum_{i=1}^{n}H(X_{1,i}|X_1^{i-1},U_1^{i},M_2,U_{1,1}=u_{1,1})\\ &\overset{(d)}{\leq}\sum_{i=1}^{n}H(X_{1,i}|U_{1,i},X_{2,i})\\ &\overset{(e)}{=}H(X_1|U_1,X_2,Q)\\ &\leq H(X_1|U_1,X_2), \end{aligned} \tag{24}$$

where (*a*) follows since X_1^n,U_1^n are functions of M_1,M_2 and $u_{1,1}$; (*b*) follows since $H(M_1|X_1^n,U_1^n,\,M_2,U_{1,1}=u_{1,1})=0$ holds due to the constraint of zero probability of error; (*c*) follows since $U_{1,i}$ is a function of X_1^{i-1},M_2 and $u_{1,1}$;

(d) follows by conditioning reduces entropy; (e) follows by defining a variable Q uniformly distributed in the set $[1, n]$ and independent of all other variables, along with $X_1 = X_{1Q}$, $X_2 = X_{2Q}$ and $U_1 = U_{1Q}$.

Similar for nR_2 we obtain the bound $nR_1 \leq H(X_1|U_1, X_2)$. Moreover, for the sum-rate, similar steps lead to

$$\begin{aligned} n(R_1+R_2) &= H(M_1, M_2) = H(M_1, M_2|U_{1,1} = u_{1,1}) \\ &= H(M_1 M_2, X_1^n, X_2^n, U_1^n|U_{1,1} = u_{1,1}) \\ &= H(X_1^n, X_2^n, U_1^n|U_{1,1} = u_{1,1}) \\ &= \sum_{i=1}^{n} H(U_{1,i}|X_1^{i-1}, X_2^{i-1}, U_1^{i-1}, M_2, U_{1,1}=u_{1,1}) \\ &+ H(X_{1,i}, X_{2,i}|X_1^{i-1}, X_2^{i-1}, U_1^{i}, M_2, U_{1,1}=u_{1,1}) \\ &\geq \sum_{i=1}^{n} H(X_{1,i}, X_{2,i}|U_{1,i}) \\ &= H(X_1, X_2|U_1). \end{aligned} \tag{25}$$

Let us now define $\pi_u = \Pr[U_1 = u]$ and $\phi_{x_1,x_2|u} = \Pr[X_1 = x_1, X_2 = x_2|U_1 = u]$ for $i, j \in \{0, 1\}$ and for all $u_1 \in \{0, 1, ..., \mho\}$. Probability conservation implies that the relationship (7) be satisfied. This concludes the proof.

Appendix C
Transition Probabilities for the Model in Sec. V

Here we discuss the transition probability matrix $\boldsymbol{P}$ used in Sec. V. To this end, define as $Q_{ij}^{(ab)}$ $(i, j = 1, 2)$ the probability that $H_j = a \in \{0, 1\}$ energy units are added to the battery at Node j conditioned on Node i sending symbol $X_j = b \in \{0, 1\}$, for $i \neq j$, namely

$$Q_{ij}^{(00)} = P_{ij}^{(00)} P_{jj}^{(00)} + P_{ij}^{(01)} P_{jj}^{(10)}, \tag{26a}$$
$$Q_{ij}^{(01)} = P_{ij}^{(00)} P_{jj}^{(01)} + P_{ij}^{(01)} P_{jj}^{(11)}, \tag{26b}$$
$$Q_{ij}^{(10)} = P_{ij}^{(10)} P_{jj}^{(00)} + P_{ij}^{(11)} P_{jj}^{(10)}, \tag{26c}$$
$$\text{and } Q_{ij}^{(11)} = P_{ij}^{(10)} P_{jj}^{(01)} + P_{ij}^{(11)} P_{jj}^{(11)}. \tag{26d}$$

Note that these transition probabilities correspond to the cascade of the channels in Fig 4 and Fig 5. Based on these probabilities, we can now evaluate all the possible transition probabilities from state (u_1, u_2) to any other state (u_1', u_2'). We start with $u_1 \in [1, B_1-1]$ and $u_2 \in [1, B_2-1]$ for $B_1, B_2 > 1$ whose outgoing transition probabilities are illustrated in Fig. 11. The "boundary" states with $u_j = 0$ or $u_j = B_j$ for some $j = 1, 2$ are discussed later.

By the stated assumptions, the state (u_1, u_2) can only transit to state $(u_1 + i_1, u_2 + i_2)$ with $i_1, i_2 \in \{-1, 0, 1\}$, so that the energy in the battery is increased or decreased by at most one energy unit. Therefore, for the "non-boundary" states (u_1, u_2) with $u_1 \in [1, B_1 - 1]$ and $u_2 \in [1, B_2 - 1]$, the probabilities in Fig. 11 can be easily obtained as

$$\begin{aligned} P_{0,0}(u_1, u_2) =& p_{1|(u_1,u_2)} p_{2|(u_1,u_2)} (Q_{21}^{(11)} Q_{12}^{(11)}) \\ &+ p_{1|(u_1,u_2)} \bar{p}_{2|(u_1,u_2)} (Q_{21}^{(01)} Q_{12}^{(10)}) \\ &+ \bar{p}_{1|(u_1,u_2)} p_{2|(u_1,u_2)} (Q_{21}^{(10)} Q_{12}^{(01)}) \\ &+ \bar{p}_{1|(u_1,u_2)} \bar{p}_{2|(u_1,u_2)} (Q_{21}^{(00)} Q_{12}^{(00)}), \end{aligned} \tag{27}$$

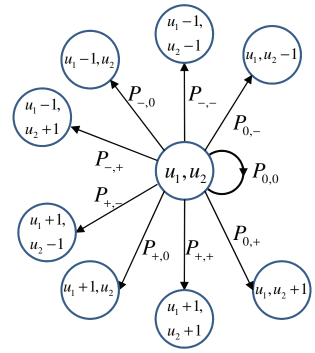

Fig. 11. The outgoing transition probabilities from a state (u_1, u_2).

$$\begin{aligned} P_{+,0}(u_1, u_2) =& \bar{p}_{1|(u_1,u_2)} p_{2|(u_1,u_2)} (Q_{21}^{(11)} Q_{12}^{(01)}) \\ &+ \bar{p}_{1|(u_1,u_2)} \bar{p}_{2|(u_1,u_2)} (Q_{21}^{(01)} Q_{12}^{(00)}), \end{aligned} \tag{28}$$

$$\begin{aligned} P_{0,+}(u_1, u_2) =& p_{1|(u_1,u_2)} \bar{p}_{2|(u_1,u_2)} (Q_{21}^{(01)} Q_{12}^{(11)}) \\ &+ \bar{p}_{1|(u_1,u_2)} \bar{p}_{2|(u_1,u_2)} (Q_{21}^{(00)} Q_{12}^{(01)}), \end{aligned} \tag{29}$$

$$P_{+,+}(u_1, u_2) = \bar{p}_{1|(u_1,u_2)} \bar{p}_{2|(u_1,u_2)} (Q_{21}^{(01)} Q_{12}^{(01)}), \tag{30}$$

$$P_{+,-}(u_1, u_2) = \bar{p}_{1|(u_1,u_2)} p_{2|(u_1,u_2)} (Q_{21}^{(11)} Q_{12}^{(00)}), \tag{31}$$

$$P_{-,+}(u_1, u_2) = p_{1|(u_1,u_2)} \bar{p}_{2|(u_1,u_2)} (Q_{21}^{(00)} Q_{12}^{(11)}), \tag{32}$$

$$\begin{aligned} P_{0,-}(u_1, u_2) =& p_{1|(u_1,u_2)} p_{2|(u_1,u_2)} (Q_{21}^{(11)} Q_{12}^{(10)}) \\ &+ \bar{p}_{1|(u_1,u_2)} p_{2|(u_1,u_2)} (Q_{21}^{(10)} Q_{12}^{(00)}), \end{aligned} \tag{33}$$

$$\begin{aligned} P_{-,0}(u_1, u_2) =& p_{1|(u_1,u_2)} p_{2|(u_1,u_2)} (Q_{21}^{(10)} Q_{12}^{(11)}) \\ &+ p_{1|(u_1,u_2)} \bar{p}_{2|(u_1,u_2)} (Q_{21}^{(00)} Q_{12}^{(10)}), \end{aligned} \tag{34}$$

and

$$P_{-,-}(u_1, u_2) = p_{1|(u_1,u_2)} p_{2|(u_1,u_2)} (Q_{21}^{(10)} Q_{12}^{(10)}), \tag{35}$$

where we recall that $p_{i|(u_1,u_2)}$ is the probability of sending a "1" symbol by Node i given the state (u_1, u_2) and $\bar{p}_{i|(u_1,u_2)} = 1 - p_{i|(u_1,u_2)}$.

For the "boundary" states (u_1, u_2) with u_1 and/or u_2 equal to 0 the outgoing transitions in Fig. 11 and probabilities in (27)-(35) still hold since the transitions to states with energy less than zero are disabled by the conditions $p_{i|(u_1,u_2)} = 0$ for $u_j = 0$, $j = 1, 2$. Instead, if $u_1 = B_1$ and $u_2 \in [0, B_2 - 1]$ then the probabilities in (33)-(35) remain the same, but we have $P_{+,-} = P_{+,0} = P_{+,+} = 0$ and $P_{0,0}$ equals the sum of the right-hand sides of (27) and (28), while $P_{0,+}$ equals the sum of the right-hand sides of (29) and (30), and $P_{0,-}$ equals the sum of the right-hand sides of (33) and (31) . The transition probabilities from the states with $u_1 \in [0, B_1-1]$ and $u_2 = B_2$ follow in a symmetric fashion. Finally if $u_1 = B_1$ and $u_2 = B_2$, then we have $P_{-,+} = P_{+,-} = P_{+,0} = P_{0,+} = P_{+,+} = 0$ and $P_{0,0}$ is the sum of (27), (30), (28), and (29), while $P_{0,-}$ is the sum of (33) and (31) and $P_{-,0}$ is the sum of (34) and (32).

By using the transition probabilities defined above, one can easily construct the transition matrix $\boldsymbol{P}$ of the corresponding Markov chain. For instance, for the case $B = 1$, we can write the transition matrix as

$$\boldsymbol{P} = \begin{bmatrix} P_{0,0}(0,0) & P_{0,-}(0,1) & P_{-,0}(1,0) & P_{-,-}(1,1) \\ P_{0,+}(0,0) & P_{0,0}(0,1) & P_{-,+}(1,0) & P_{-,0}(1,1) \\ P_{+,0}(0,0) & P_{+,-}(0,1) & P_{0,0}(1,0) & P_{0,-}(1,1) \\ P_{+,+}(0,0) & P_{+,0}(0,1) & P_{0,+}(1,0) & P_{0,0}(1,1) \end{bmatrix}, \quad (36)$$

where the column index represents the the initial state and the row index the final state.

REFERENCES

[1] V. Chawla and S. H. Dong, "An overview of passive RFID," *IEEE Commun. Mag.*, vol. 45, no. 9, pp. 11–17, Sept. 2007.

[2] F. Zhang, S. A. Hackworth, X. Liu, H. Chen, R. J. Sclabassi, and M. Sun, "Wireless energy transfer platform for medical sensors and implantable devices," in *Proc. 2009 Annual Int. Conf. IEEE Engineering in Medicine and Biology Society*, pp. 1045–1048.

[3] A. W. Eckford, "Achievable information rates for molecular communication with distinct molecules," in *Proc. 2007 Bio-Inspired Models of Network, Information and Computing Systems*, pp. 313–315.

[4] B. Schumacher and M. Westmoreland, *Quantum Processes Systems, and Information.* Cambridge University Press, 2010.

[5] R. Landauer, "Minimal energy requirements in communication," *Science,* vol. 272, no. 5270, pp. 1914–1918, Jun. 1996.

[6] H. Leff and A. F. Rex, *Maxwell's Demon 2: Entropy, Classical and Quantum Information*. Taylor & Francis, 2002.

[7] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. 2008 IEEE Int. Symposium on Inform. Theory*, pp. 1612–1616.

[8] P. Grover and A. Sahai, "Shannon meets Tesla: wireless information and power transfer," in *Proc. 2010 IEEE Int. Symposium on Inform. Theory*, pp. 2363–2367.

[9] L. R. Varshney, "On energy/information cross-layer architectures," in *Proc. 2012 IEEE Int. Symposium on Inform. Theory*.

[10] M. Gastpar, "On capacity under receive and spatial spectrum-sharing constraints," *IEEE Trans. Inf. Theory*, vol. 53, no. 2, pp. 471–487, Feb. 2007.

[11] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," in *Proc. 2011 IEEE Global Telecommunications Conference*, pp. 1–5.

[12] Z. Xiang and M. Tao, "Robust beamforming for wireless information and power transmission," *IEEE Wireless Commun. Lett.*, vol. 1, no. 4, pp. 372–375, Aug. 2012.

[13] L. Liu, R. Zhang and K. C. Chua, "Wireless information transfer with opportunistic energy harvesting," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 288–300, Jan. 2013.

[14] O. Ozel and S. Ulukus, "Information-theoretic analysis of an energy harvesting communication system," in *Proc. 2010 IEEE PIMRC,* pp. 330–335.

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

[16] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: optimal policies," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 8, pp. 1732–1743, Sept. 2011.

[17] C. K. Ho and R. Zhang, "Optimal energy allocation for wireless communications with energy harvesting constraints," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4808–4818, Sep. 2012.

[18] C. E. Shannon, "Two-way communication channels," in *Proc. 1961 Berkeley Symp. Probability and Statistics*, pp. 61l–644.

[19] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[20] T. Cover and J. A. Thomas, *Elements of Information Theory*. Wiley-Interscience, 2006.

[21] A. Kailas, M. A. Ingram, and Y. Zhang, "A novel routing metric for environmentally-powered sensors with hybrid energy storage systems," in *Proc. 2009 Int. Conf. on Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology*, pp. 42–46.

[22] W. C. Brown, "The history of power transmission by radio waves," *IEEE Trans. Microwave Theory and Techniques*, vol. 32, no. 9, pp. 1230–1242, Sept. 1984.

[23] O. Ozel and S. Ulukus, "AWGN channel under time-varying amplitude constraints with causal information at the transmitter," in *Proc. 2011 IEEE Asilomar Conf. on Signals, Systems and Computers*, pp. 373–377.

[24] X. Zhou, R. Zhang, and C. K. Ho, "Wireless information and power transfer: architecture design and rate-energy tradeoff," arXiv:1205.0618.

[25] J. Norris, *Markov Chains*. Cambridge University Press, 1997.

[26] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra.* Society for Industrial and Applied Mathematics, 2000.

[27] R. M. Gray, *Entropy and Information Theory*. Springer, 2011.

Petar Popovski (S'97-A'98-M'04-AM'10) received Dipl.-Ing. In electrical engineering (1997) and Magister Ing. in communication engineering (2000) from Sts. Cyril and Methodius University, Skopje, Macedonia, and Ph.D. from Aalborg University, Denmark, in 2004. He is currently a Professor at Aalborg University. From 2008 to 2009 he held part-time position as a wireless architect at Oticon A/S. He has more than 150 publications in journals, conference proceedings and books and has more than 25 patents and patent applications. He has received the Young Elite Researcher award and the SAPERE AUDE career grant from the Danish Council for Independent Research. He has received six best paper awards, including three from IEEE. Dr. Popovski serves on the editorial board of several journals, including IEEE COMMUNICATIONS LETTERS (Senior Editor), IEEE JSAC Cognitive Radio Series, and IEEE TRANSACTIONS ON COMMUNICATIONS. His research interests are in the broad area of wireless communication and networking, information theory and protocol design.

Ali Mohammad Fouladgar received the B.Sc. degree and the M.Sc. degree in Electrical engineering from Isfahan University of Technology, Isfahan, Iran, in 2006 and 2008, respectively. He is currently with the Center for Wireless Communications and Signal Processing Research (CWCSPR), New Jersey Institute of Technology (NJIT), Newark, where he is a Ph.D. student. His current research interests concern the analysis and design of communication networks with emphasis on information-theoretic and signal processing aspects.

Osvaldo Simeone received the M.Sc. degree (with honors) and the Ph.D. degree in information engineering from Politecnico di Milano, Milan, Italy, in 2001 and 2005, respectively. He is currently with the Center for Wireless Communications and Signal Processing Research (CWCSPR), New Jersey Institute of Technology (NJIT), Newark, where he is an Associate Professor. His current research interests concern the analysis and design of communication networks with emphasis on information-theoretic and signal processing aspects. Dr. Simeone is a co-recipient of Best Paper Awards of the IEEE SPAWC 2007 and IEEE WRECOM 2007. He currently serves as an Editor for IEEE TRANSACTIONS ON COMMUNICATIONS.