Transmission Policies in Wireless Powered Communication Networks with Energy Cooperation

Alessandro Biason and Michele Zorzi
{biasonal,zorzi}@dei.unipd.it

Department of Information Engineering, University of Padova - via Gradenigo 6b, 35131 Padova, Italy

Abstract—Energy Harvesting (EH) has been recognized as one of the most appealing solutions for extending the devices lifetime in wireless sensor networks. Despite the vast literature available about ambient EH, in the last few years Energy Transfer (ET) has been introduced as a new and promising paradigm. With ET, it becomes possible to actively control the energy source and thus improve the network performance. We focus on two particular applications of ET which have been studied separately in the literature so far: Energy Cooperation (EC) and Wireless Powered Communication Networks (WPCNs). In the first case, energy is wirelessly shared among terminal devices according to their requirements and energy availability, whereas, in a WPCN, energy can be purposely transferred from an energy-rich network node (e.g., an access point) to terminal devices. We solve a weighted throughput optimization problem for the two-node case using optimal as well as sub-optimal schemes. Numerically, we explain the role of EC in improving the system performance.

I. INTRODUCTION

Traditionally, in wireless sensor networks or cellular networks the devices are battery powered and therefore, unless the batteries are periodically replaced, the network lifetime is limited. However, battery replacement is not always a viable option due to the excessive costs (e.g., if the number of devices is high) and may even be infeasible in some cases (e.g., devices located in toxic environments or hard-to-reach areas, or implanted inside the human body). In these scenarios, Energy Harvesting (EH) can be adopted to prolong the devices lifetime for, ideally, an unlimited amount of time. The challenge is to understand how to optimally exploit the external renewable energy source in order to achieve high performance. In addition to the traditional ambient energy sources, e.g., sunlight, wind, electromagnetic field, vibrations, etc., man-made energy sources have been recently considered. In particular, thanks to recent technology developments, it is possible to transfer energy wirelessly to/among different devices and recharge their batteries in a controlled fashion.

Two of the most studied applications of the Energy Transfer (ET) mechanism are Wireless Powered Communication Networks (WPCNs) [1]–[7] and Energy Cooperation (EC) [8]–[12]. In the first case, an Access Point (AP) with unlimited available energy but with an average power constraint supplies different mobile or terminal devices simultaneously. The devices use the harvested energy to perform computational tasks, e.g., sensing, and uploading data packets. The challenge is to correctly deliver the energy to the terminals which need it most. A common approach to solving this problem is the “harvest-then-transmit” scheme [3], in which energy transfer and data uplink phases are temporally interleaved. In [4] the previous paradigm was extended to consider data cooperation among terminal devices, so that the closer user can be used as a hybrid relay. Other works investigated the benefits of multi-antenna transmission to improve the harvesting performance [6], or the use of massive MIMO [14]. In general, energy transfer and uplink data transmission are performed in the same frequency band, thus full-duplex schemes with interference cancellation were studied in [5], [15].

Differently, the EC paradigm is more challenging to analyze and realize. In this case, terminal devices share their energy in order to achieve a higher common global reward. Energy cooperation can be realized using Strongly Coupled Magnetic Resonances (SCMR) [16], which provides high energy efficiency (e.g., 15–40% at a distance of 2 m). This new paradigm was introduced in [9], where the authors set up an offline throughput optimization problem. Other works [7], [11] studied the case of a transmitter-receiver pair and introduced performance upper bounds with and without cooperation.

In this work, we combine the two previous approaches. Even though our model is similar to [4], here we consider cooperation in terms of energy and not information. As in [1], we set up an online optimization problem in order to focus on the long-term performance. However, differently from [1], to reduce the number of system variables (duration of uplink and downlink phases, transmission and transfer powers, etc.), we introduce an approximate technique which is easier to compute and useful to understand the performance of the system [2]. Using this new scheme, we compare the system throughput with and without energy cooperation, and show that EC can significantly increase the system performance when a doubly near-far phenomenon is present.

The paper is organized as follows. Section II defines the system model. In Section III we present the long-term optimization problem and its optimal solution. The approximate approach is described in Section IV. Section V presents our numerical results. Finally, Section VI concludes the paper.
II. SYSTEM MODEL

Consider a Wireless Powered Communication Network (WPCN) in which an Access Point (AP) transfers energy to two terminal devices $D_1$ and $D_2$ located at different distances. $D_1$ and $D_2$ use the harvested energy to uplink data packets to AP.

Due to the different locations, one device may be more advanced than the other in terms of average harvested energy and average energy required for the uplink phase (doubly near-far effect [3]). With the goal of improving the system performance in terms of throughput, we allow energy cooperation between $D_1$ and $D_2$, i.e., the two devices are able to exchange energy between them according to their current energy levels and requirements.

$D_i$ $(i \in \{1,2\})$ is equipped with a finite battery of size $B_{i,max}$ which is discretized in $b_{i,max} + 1$ energy quanta and can be considered as a buffer. One energy quantum corresponds to $\Delta_q = B_{i,max}/b_{i,max}$ and for a consistent formulation we choose $\Delta_q = B_{1,max}/b_{1,max}$.

The link between AP and the two devices is half-duplex, so $D_i$ can either receive energy or transmit data in uplink at a given time. Moreover, thanks to the different nature of the technologies used for energy cooperation, the two devices can exchange energy independently of the current uplink or downlink states. Uplink and downlink phases are interleaved over time in a TDMA fashion. In every time slot $k$, with $k = 0,1,\ldots$, four operations are performed (decision, uplink, downlink and energy cooperation), which will be discussed in more detail in the next subsections.

A. Decision

At the beginning of a slot, the choice of the transmission parameters (transmission powers and durations, transferred energy via RF and via SCMR) is performed. We consider a centralized approach in which a coordinator (e.g., the AP) decides and disseminates these parameters. Future work includes the study of distributed schemes as in [7]. The duration of this first phase is assumed to be negligible.

As in [2], [2], we formulate the problem as a Markov Decision Process (MDP) and apply stochastic optimization tools to solve it. The decision on the transmission parameters is based on the current state of the system $s$, which is given by the channel status and the battery levels, i.e., $s = (b, h, g)$, where $b = (b_1, b_2)$ describes the current state of charge of the two batteries, with $b_i \in [0, 1, \ldots, b_{i,max}]$, $h = (h_1, h_2)$ represents the uplink channel gains of the two devices, and $g = (g_1, g_2)$ is the downlink channel gain pair.

B. Uplink

The first $\tau_1 + \tau_2$ seconds of a slot are dedicated to the uplink phase. In order to avoid collision, $D_1$ and $D_2$ transmit their data to the common access point AP in a TDMA fashion. Due to the different distances from AP, the two devices have to use different powers to achieve the same transmission rate. In particular, the uplink channel gain of $D_i$ is modeled as $h_i = h_i \Theta_i$, where $h_i$ is the path loss component and depends upon the distance and $\Theta_i$ represents the fading component and is i.i.d. over time. We assume $h_i = h_{0,i} \left( d_i \right)^{-\gamma_i}$, where $d_i$ is the distance between $D_i$ and AP, $h_{0,i}$ is the reference attenuation at the distance of 1 m and $\gamma_i$ is the path loss exponent. The random variables $\Theta_1, \Theta_2$ are assumed independent because of the different positions of $D_1$ and $D_2$, and have joint pdf $f_{\Theta}(\theta) = f_{\Theta_1}(\theta_1) f_{\Theta_2}(\theta_2)$.

The transmission powers $\rho_1, \rho_2$ and durations $\tau_1, \tau_2$ change dynamically in every slot according to the decision process. The energy spent by device $D_i$ is $E_i \leq B_i$, where $B_i = b_i/\Delta_q$ is the energy stored in the battery at the beginning of the slot.

C. Downlink

During the second phase, which lasts for $\tau_{AP} = T - \tau_1 - \tau_2$, AP transfers energy to both devices. AP is equipped with more than one antenna in order to perform energy beamforming and to direct the energy toward $D_1$ or $D_2$. We assume that also the downlink channel is affected by flat fading and thus model the channel attenuation as $g_i = \tilde{g}_i \Psi_i$, with $\tilde{g}_i = g_{0,i} \left( d_i \right)^{-\beta_i}$ and $\Psi_i \sim f_{\Psi_i}(\psi_i)$, $g_{0,i}$, $\beta_i$ and $\Psi_i$ are defined as their counterparts in the previous section.

When AP transfers a power $Q_i \leq Q_{max}$ (e.g., $Q_{max} = 3$ W) to $D_i$, the corresponding harvested energy is

$$ C_i = \tau_{AP} Q_i \eta_i $$

where $\eta$ is a conversion efficiency factor in $[0, 1]$. When beamforming is used, AP is subject to the constraint $Q_1 + Q_2 \leq Q_{max}$.

D. Energy Cooperation

Energy Cooperation (EC) is useful to boost the system performance when one device receives and/or consumes much less energy than the other. Assume, for example, that $D_2$ is farther from AP and thus receives only a small amount of energy due to the term $g_{0.2} \ll g_{0.1}$ (on average). In this case, $D_1$ can operate as energy relay and share part of its energy with $D_2$, if necessary.

Since EC and downlink ET are performed with different technologies, we assume that for the duration of the entire slot, energy cooperation can be performed, if necessary. We model the EC efficiency with a constant $\beta_{i,1} \in (0, 1)$ ($i = 1$ if $i = 2$ and vice-versa): if $x$ joules of energy are sent from $D_1$ to $D_2$, then only $\beta_{1,2} x$ J are effectively received by $D_2$. $\beta_{i,1}$ is strongly distance dependent (some energy efficiency curves can be found in [16]). In this work, we use, as a baseline, $\beta_{1,2} = \beta_{2,1} \in [0.15, 0.4]$ which represent the losses at a distance of two meters (considering also the conversion inefficiencies).

EC is performed for all the slot length in parallel to the uplink or downlink phases. The energy extracted from $D_i$ and delivered to $D_1$ is $Z_i$ and must satisfy $E_i + Z_i \leq B_i$, i.e., it is not possible to use more energy than the stored amount at the beginning of the slot.

E. Batteries Evolution

Differently from [3], [4], and similar to [1], in this work we want to fully exploit the potential of the batteries. According to the resonant coils, an EC is performed at frequencies around 10 MHz [16], whereas an energy transfer from AP to the devices uses much higher frequencies (e.g., 915 MHz [13]).
to the previous subsections, the battery level of D_i changes dynamically in every slot according to the parameters set in the decision phase. The update formula becomes

$$B_i \leftarrow \min \{B_i, B_{i, \text{max}}, B_{i, \text{min}} - E_i + C_i + \beta Z_i - Z_i \}$$  \hspace{1cm} (2)$$

The min is used to consider overflow situations. Note that the right term is always greater than zero because \(E_i + Z_i \leq B_i\).

### III. Optimization Problem

#### A. Problem Statement

We define a policy \(\mu\) as an action probability measure over the state set \(\Delta\). The central controller which computes the policy knows all the system parameters (battery levels and channel states). For every state \(s = (b, h, g) \in \Delta\), \(\mu\) defines with which probability an action \(a\) is performed. \(a\) summarizes the transmission durations \(\tau_1, \tau_2, \tau_{\text{AP}}\), the transmission powers \(p_1, p_2\), the downlink transmission powers \(Q_1, Q_2\) and the energy transfer energies \(Z_1, Z_2\). Formally, \(\mu\) defines \(P_{\mu}(a|s)\), with \(\sum_{a \in A(s)} P_{\mu}(a|s) = 1\), where \(A(s)\) is the set of the possible actions in state \(s\). For the sake of presentation simplicity, in the next sections we use a deterministic policy \(\mu\), i.e., \(P_{\mu}(a|s)\) is equal to 1 for \(a = \hat{a}\) and to 0 for \(a \neq \hat{a}\), where \(\hat{a}\) is an action in \(A(s)\). However, in our numerical evaluation we consider a more general random policy.

The goal is to maximize the long-term average weighted-sum throughput reached by the two devices. Formally, we study the following problem

$$\mu^* = \arg \max_{\mu} \{\alpha G_1, \mu + (1 - \alpha) G_2, \mu\}$$ \hspace{1cm} (3)$$

where \(\mu^*\) is the optimal policy of our problem, and the expectation is taken with respect to the channel conditions. If the weight parameter \(\alpha = 0\) \([\alpha = 1]\) then we are focusing only on \(D_2\) \([D_1]\). To maximize the sum-throughput we can use \(\alpha = 1/2\), whereas if we want to achieve fairness we can choose \(\alpha\) with a bisection search \([1]\). \(R(p_i, h_i)\) is the normalized transmission rate of device \(D_i\) in a single slot, approximated with Shannon’s capacity formula

$$R(p_i, h_i) = \log \left(1 + \frac{h_i p_i}{\sigma_0^2}\right),$$

where \(p_i\) is the transmission power, \(h_i\) is the uplink channel gain and \(\sigma_0^2\) is the noise power.

#### B. Optimal Solution

The problem previously defined is a Markov Decision Process (MDP) which can be solved optimally with standard optimization techniques, e.g., the Value Iteration Algorithm \([17, \text{Vol. 2, Sec. 4}]\). The basic step of VIA is the policy improvement step

$$J^k_{\alpha}(s) = \max_{a \in A(s)} \left\{ r_\alpha(\tau, \rho|s, a) + \sum_{s'} \mathbb{P}(s'|s, a) J^k_{\alpha'}(s') \right\}, \hspace{1cm} (6)$$

where \(J^k_{\alpha}(s)\) is the value function at iteration \(k\) and the bold notation indicates a pair of values.

Finding the optimal policy is practically feasible only for a relatively small number of discrete values which however corresponds to a rough quantization. Therefore, in this section we propose a method which is based on the characteristics of the original solution but is faster to compute and achieves approximately the same performance as the optimal scheme. This is particularly useful to characterize the system performance and identify the system trade-offs.

We first reduce the state space complexity by exploiting the channel i.i.d. properties.

#### A. Reducing State Space Complexity

In a general step of VIA, given the current policy, the corresponding cost-to-go function \(J_{\alpha}\) has to be computed (policy evaluation step \([17, \text{Vol. 1, Sec. 7.4}]\). This process is challenging when the state space is large.

So far, the state of the system is the tuple \(s = (b, g, h)\). However, since \(g\) and \(h\) evolve independently over time, the state space can be reduced to \(s = (b)\) only, as follows. Define a new cost-to-go function

$$K_{\hat{b}} \triangleq \sum_{g, h} J_{(b, g, h)}. \hspace{1cm} (8)$$

\(K_{\hat{b}}\) substitutes \(J_{(b, g, h)}\) in the original problem. Indeed, we can rewrite the policy improvement step as

$$K_{\hat{b}} \leftarrow \sum_{g, h} f(g, h) \max_{a \in A(s, b, g, h)} \left\{ r_\alpha(\tau, \rho|s, a) + \sum_{s'} \mathbb{P}(s'|s, a) J_{\alpha'} \right\}$$

$$= \sum_{g, h} f(g, h) \max_{a \in A(s, b, g, h)} \left\{ r_\alpha(\tau, \rho|s, a) + K_{\hat{b}} \right\}, \hspace{1cm} (9a)$$

where \(b'\) is defined using (2) as

$$b' = \min \{b_{\text{max}}, b - e_i + c_i + \beta z_i - z_i\}. \hspace{1cm} (10)$$

\(e_i \triangleq \lceil E_i / \Delta q \rceil\), \(c_i \triangleq \lceil C_i / \Delta q \rceil\) and \(z_i \triangleq \lceil Z_i / \Delta q \rceil\) are the discrete versions of \(E_i, C_i\) and \(Z_i\). Using the \([\cdot]\) operation leads to a lower bound to the real performance.

This procedure simplifies the numerical computation because it reduces the complexity of the policy evaluation step (there is a lower number of states) and the number of elementary operations inside the \(\max\) operation in the policy improvement step, and will be used in the next subsection to derive the approximate scheme.

Even with the simplification introduced in this subsection, performing the policy improvement step, i.e., solving (9) for all system states, remains challenging. To manage this problem, several different approximated techniques have been proposed in the literature so far. An interesting idea is to approximate the function \(K_{\hat{b}}\) with another one simpler to compute. We follow this approach in the remainder of this section, and derive an
Therefore, the number of operations of a generic step of the algorithm in a deterministic, complexity of the problem. For example, in our numerical results of Section V, we have $|B| \approx 25000$ (two batteries with $b_{1,\text{max}} = b_{2,\text{max}} = 160$) and $|\mathcal{B}(k)| = 41$, $\forall k$.}

**D. Convergence Properties**

In the following we show that, provided that the approximation $\tilde{K}_b(k)$ is sufficiently good, the long-term reward of $\text{App-VIA}$ is a good approximation of $\text{VIA}$.

First, we introduce the notation $T(\cdot)$ as follows. Define the two sets $\mathcal{R}_b \triangleq \{K_b(k), \forall b \in \mathcal{B}\}$ and $\mathcal{\tilde{R}}_b \triangleq \{\tilde{K}_b(k), \forall b \in \mathcal{B}\}$. Then, Equations (9) and (11) can be written as

$$K_b^{(k+1)} = T(\mathcal{R}_b(k), b), \quad \forall b \in \mathcal{B},$$

$$\tilde{K}_b^{(k+1)} = T(\mathcal{\tilde{R}}_b(k), b), \quad \forall b \in \mathcal{B},$$

respectively. Also, assume that the initial configurations are equal, i.e., $\mathcal{R}(0) = \mathcal{\tilde{R}}(0)$. Note that $K_b^{(k+1)}$ is evaluated for every $b$, whereas we compute $\tilde{K}_b^{(k+1)}$ only in the subset $\mathcal{B}(k+1)$.

**Proposition 1.** After $N$ iterations, the cost-to-go functions of $\text{App-VIA}$ and $\text{VIA}$ differ by at most $Ne$, i.e.,

$$\|\mathcal{R}_b^{(N)} - \mathcal{\tilde{R}}_b^{(N)}\|_{\infty} \leq Ne$$

with

$$\epsilon \triangleq \max_{k=0,...,N-1} \max_{b \in \mathcal{B}} \left\{ \tilde{K}_b^{(k+1)} - T(\mathcal{\tilde{R}}_b(k), b) \right\}$$

**Proof.** See Appendix A in [2].

We first remark that, because of (18), Proposition 1 describes a worst case analysis. $N$ corresponds to the number of iterations of VIA and, in our problem, it can be numerically verified that $N$ is typically small, e.g., $N \approx 10$. The previous proposition provides a bound to the algorithm performance and guarantees convergence, provided that the approximation of $K_b^{(k+1)}$ is sufficiently good.

**V. Numerical Results**

If not otherwise specified, in our numerical evaluation we used the following parameters: Nakagami fading with parameter 1 (Rayleigh fading with no Line-of-Sight (LoS)) or 5 (strong LoS component), energy conversion efficiency $\eta = 0.8$, $h_{0,1} = h_{0,2} = 1.25 \times 10^{-3}$, $\gamma_1 = \gamma_2 = 3$ (path loss exponents), $\sigma_0^2 = -155$ dBm/Hz (noise power), a bandwidth of 1 MHz, $T = 500$ ms (slot duration), $Q_{\text{max}} = 3$ W (maximum transfer power), $P_{1,\text{min}} = P_{2,\text{min}} = 1$ mW and $P_{1,\text{max}} = P_{2,\text{max}} = 10$ mW, $B_{1,\text{max}} \triangleq B_{2,\text{max}} = 0.125$ mJ. The distances between AP and the two devices are $(d_1, d_2) \in \{(1,3),(2,4),(3,5)\}$ m. Since the devices are 2

\[\text{3We adopt the notation } \|\mathcal{R}(N) - \mathcal{\tilde{R}}(N)\|_{\infty} \preceq \max_{b \in \mathcal{B}} |K_b^{(N)} - \tilde{K}_b^{(N)}|.\]
When \( d_1 = 0 \), \( d_2 = \beta_{1,2} = \beta_{1,2} \) is set to 0.15 or 0.4 [16]. However, even if we consider bi-directional ET, for most of the time the term \( Z_2 \) (energy transferred from \( D_2 \) to \( D_1 \)) is zero. Also, we present our results for the approximate scheme of Section IV and refer the reader to [2] for a more detailed comparison between optimal and sub-optimal approaches.

Figure 3 shows the normalized throughput region of the two devices for different values of \( \beta \) and \( d_1 \). The curves are generated by changing the weight value \( \alpha \) in Equation (3). When \( \alpha = 0 \), (3) degenerates to \( \arg \max \{ G_{2,\mu} \} \) and we obtain the points on the y-axis, i.e., the optimization focuses on \( D_2 \) only (similarly for \( \alpha = 1 \) with \( D_1 \)). It can be observed that the distance strongly influences the performance of the system. This is mainly due to the path loss effects, which limit the operating range of the energy transfer technology to a few meters. As can be seen, energy cooperation among the two devices can greatly improve the system performance, especially when \( \alpha \) is small. Indeed, in this case more importance is given to \( D_2 \), which can benefit from part of the energy of \( D_1 \) for uploading more data. Thus, employing \( D_1 \) as a relay to solve the doubly near-far effect may be a suitable solution.

Finally, Figure 4 is similar to the previous one but was obtained using Nakagami fading (stronger LoS component for the same average channel quality). Since the LoS is stronger, the scenario is closer to the deterministic energy arrival case, which can be shown to be an upper bound for the energy harvesting systems. The stronger LoS directly improves the performance of \( D_1 \), whereas the throughput of \( D_2 \) increases only when EC is considered. In this case, the benefits of EC are even larger than in the Rayleigh case.

VI. CONCLUSIONS

A wireless powered communication network consisting of an access point and two terminal devices with energy transfer capabilities was analyzed. We solved the long-term throughput optimization problem and showed the role of energy cooperation in improving the performance of the system when a doubly near-far effect is present. An approximation to the optimal solution was introduced and its quality was analytically discussed.

REFERENCES