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.[1] 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

[1]The typical technology for enabling WPCNs is RF-energy transfer [13].

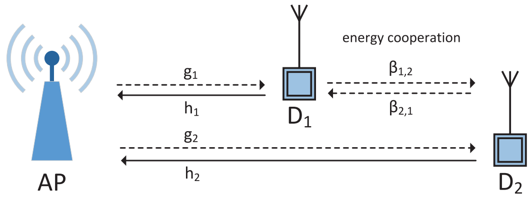

Figure 1: Block diagram of the system.

"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 [**?**], [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.

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

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 advantaged 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 $\Delta_q \triangleq B_{i,\max}/b_{i,\max}$ J and for a consistent formulation we choose $\Delta_q = B_{1,\max}/b_{1,\max} = B_{2,\max}/b_{2,\max}$.

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,[2] 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], 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., $s = (\mathbf{b}, \mathbf{h}, \mathbf{g})$, where $\mathbf{b} = (b_1, b_2)$ describes the current state of charge of the two batteries, with $b_i \in \{0, 1, \ldots, b_{i,\max}\}$, $\mathbf{h} = (h_1, h_2)$ represents the uplink channel gains of the two devices, and $\mathbf{g} = (g_1, g_2)$ 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 $h_i = \tilde{h}_i \Theta_i$, where $\tilde{h}_i$ is the path loss component and depends upon the distance and Θ_i represents the fading component and is i.i.d. over time. We assume $\tilde{h}_i = h_{0,i}\,(d_i)^{-\gamma_i}$, 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 γ_i is the path loss exponent. The random variables Θ_1, Θ_2 are assumed independent because of the different positions of D_1 and D_2, and have joint pdf $f_{\boldsymbol{\Theta}}(\boldsymbol{\theta}) = f_{\Theta_1}(\theta_1) f_{\Theta_2}(\theta_2)$.

The transmission powers ρ_1, ρ_2 and durations τ_1, τ_2 change dynamically in every slot according to the decision process. The energy spent by device D_i is $E_i \triangleq \rho_i \tau_i$ and must satisfy the energy constraint $E_i \leq B_i$, 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 $\tau_{\mathrm{AP}} \equiv T - \tau_1 - \tau_2$ s, 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 $\tilde{g}_i = g_{0,i}\,(d_i)^{-\delta_i}$ and $\Psi_i \sim f_{\Psi_i}(\psi_i)$. $g_{0,i}$, d_i, δ_i and Ψ_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

$$C_i = \tau_{\mathrm{AP}} Q_i \eta g_i \tag{1}$$

where η 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 $\beta_{i,\tilde{i}} \in (0, 1)$ ($\tilde{i} = 1$ 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. $\beta_{i,\tilde{i}}$ 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 $\mathrm{D}_{\tilde{i}}$ 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

[2] With 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

$$B_i \leftarrow \min\{B_{i,\max}, \underbrace{B_i}_{\substack{\text{initial batt.}\\ \text{level}}} - \overbrace{E_i}^{\substack{\text{uplink tx}\\ \text{energy}}} + \underbrace{C_i}_{\substack{\text{downlink transf.}\\ \text{energy}}} + \overbrace{\beta Z_{\bar{i}}}^{\substack{\text{energy coop.}\\ \text{from } \mathrm{D}_{\bar{i}}\to\mathrm{D}_i}} - \underbrace{Z_i}_{\substack{\text{energy coop.}\\ \text{from } \mathrm{D}_i\to\mathrm{D}_{\bar{i}}}}\} \quad (2)$$

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* μ as an action probability measure over the state set $\mathcal{S}$. The central controller which computes the policy knows all the system parameters (battery levels and channel states). For every state $s = (\mathbf{b}, \mathbf{h}, \mathbf{g}) \in \mathcal{S}$, μ defines with which probability an action a is performed. a summarizes the transmission durations τ_1, τ_2, τ_{AP}, the transmission powers ρ_1, ρ_2, the downlink transmission powers Q_1, Q_2 and the energy transfer energies Z_1, Z_2. Formally, μ defines $\mathbb{P}_\mu(a|s)$, with $\sum_{a\in\mathcal{A}(s)} \mathbb{P}_\mu(a|s) = 1$, where $\mathcal{A}(s)$ 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 μ, i.e., $\mathbb{P}_\mu(a|s)$ is equal to 1 for $a = \bar{a}$ and to 0 for $a \neq \bar{a}$, where $\bar{a}$ is an action in $\mathcal{A}(s)$. 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

$$\mu^\star = \arg\max_{\mu} \{\alpha G_{1,\mu} + (1-\alpha) G_{2,\mu}\} \quad (3)$$

$$G_{i,\mu} \triangleq \liminf_{K\to\infty} \frac{1}{K} \sum_{k=0}^{K-1} \mathbb{E}\left[\tau_i R(\rho_i, h_i)\right], \qquad i \in \{1,2\}, \quad (4)$$

where $\mu^\star$ 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 α with a bisection search [1]. $R(\rho_i, h_i)$ is the normalized transmission rate of device D_i in a single slot, approximated with Shannon's capacity formula

$$R(\rho_i, h_i) = \ln\left(1 + \frac{h_i \rho_i}{\sigma_0^2}\right), \quad (5)$$

where ρ_i is the transmission power, h_i is the uplink channel gain and σ_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, Vol. 2, Sec. 4]. The basic step of VIA is the policy improvement step

$$J_s^{(k+1)} = \max_{a\in\mathcal{A}(s)} \left\{ r_\alpha(\boldsymbol{\tau}, \boldsymbol{\rho}|\boldsymbol{h}) + \sum_{s'} \mathbb{P}(s'|s,a) J_{s'}^{(k)} \right\}, \quad (6)$$

$$r_\alpha(\boldsymbol{\tau}, \boldsymbol{\rho}|\boldsymbol{h}) \triangleq \alpha \tau_1 R(\rho_1, h_1) + (1-\alpha)\tau_2 R(\rho_2, h_2) \quad (7)$$

where $J_s^{(k)}$ is the value function at iteration k and the bold notation indicates a pair of values. The $\max$ operation requires to compute the tuple $(\boldsymbol{\tau}, \tau_{\text{AP}}, \boldsymbol{\rho}, \boldsymbol{Q}, \boldsymbol{T})$ for every state of the system which, in general, is a computationally demanding task. Therefore, even if (6) provides the optimal solution to the problem, the focus of this paper is on approximate strategies which achieve close to optimal performance while keeping the numerical computation simple.

IV. Approximate Scheme

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_s has to be computed (policy evaluation step [17, 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 $s = (\mathbf{b}, \mathbf{g}, \mathbf{h})$. However, since $\mathbf{g}$ and $\mathbf{h}$ evolve independently over time, the state space can be reduced to $s = (\mathbf{b})$ only, as follows. Define a new cost-to-go function

$$K_{\mathbf{b}} \triangleq \sum_{\mathbf{g},\mathbf{h}} J_{(\mathbf{b},\mathbf{g},\mathbf{h})}. \quad (8)$$

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

$$K_{\mathbf{b}} \leftarrow \sum_{\mathbf{g},\mathbf{h}} f(\mathbf{g},\mathbf{h}) \max_{a\in\mathcal{A}_{(\mathbf{b},\mathbf{g},\mathbf{h})}} \left\{ r_\alpha(\boldsymbol{\tau}, \boldsymbol{\rho}|\mathbf{h}) + \sum_{s'} \mathbb{P}(s'|s,a) J_{s'} \right\} \quad (9a)$$

$$= \sum_{\mathbf{g},\mathbf{h}} f(\mathbf{g},\mathbf{h}) \max_{a\in\mathcal{A}_{(\mathbf{b},\mathbf{g},\mathbf{h})}} \{ r_\alpha(\boldsymbol{\tau}, \boldsymbol{\rho}|\mathbf{h}) + K_{\mathbf{b}'} \}, \quad (9b)$$

where $\mathbf{b}'$ is defined using (2) as

$$b_i' = \min\{b_{i,\max}, b_i - e_i + c_i + \beta z_{\bar{i}} - z_i\}. \quad (10)$$

$e_i \triangleq \lfloor E_i/\Delta_q \rfloor$, $c_i \triangleq \lfloor C_i/\Delta_q \rfloor$ and $z_i \triangleq \lfloor Z_i/\Delta_q \rfloor$ are the discrete versions of E_i, C_i and Z_i. Using the $\lfloor\cdot\rfloor$ 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 $K_{\mathbf{b}}$ with another one simpler to compute. We follow this approach in the remainder of this section, and derive an

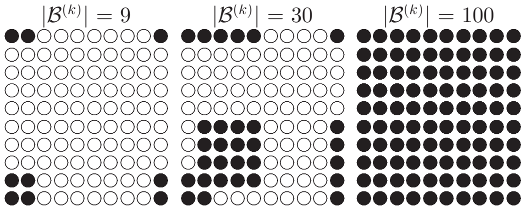

Figure 2: Different sets $\bar{\mathcal{B}}^{(k)}$ when $b_{1,\max} = b_{2,\max} = 9$. Rows and columns correspond to b_1 and b_2, respectively.

Approximate Value Iteration Algorithm (App-VIA) (see [18, Sec. 6.5]).

B. Approximate Value Iteration

In the classic VIA, the optimal policy is derived by iteratively solving (9) until the cost-to-go function converges. In the approximate approach, we modify every iteration of VIA according to the following two steps:

1) compute $K_{\mathbf{b}}^{(k)}$ for every $\mathbf{b} \in \bar{\mathcal{B}}^{(k)}$ performing the policy improvement step (Eq. (9)), with $\bar{\mathcal{B}}^{(k)} \subseteq \mathcal{B} = \{0, \ldots, b_{1,\max}\} \times \{0, \ldots, b_{2,\max}\}$. The superscript (k) denotes the k-th iteration of VIA and $\mathcal{B}$ is the set of all battery levels;
2) interpolate $K_{\mathbf{b}}^{(k)}$ for every $\mathbf{b} \in \mathcal{B} \backslash \bar{\mathcal{B}}^{(k)}$ using the values of $K_{\mathbf{b}}^{(k)}$ computed in the previous step.

The advantage is that the policy improvement is performed only for a subset $\bar{\mathcal{B}}^{(k)}$ rather that for every battery level in $\mathcal{B}$. See Figure 2 for a graphical interpretation. A black circle means that $\mathbf{b} \in \bar{\mathcal{B}}^{(k)}$. In the last case, all the battery levels are in $\bar{\mathcal{B}}^{(k)}$, i.e., $\bar{\mathcal{B}}^{(k)} = \mathcal{B}$. In general, $\bar{\mathcal{B}}^{(k)}$ can dynamically change in every step of the algorithm in a deterministic, stochastic or simulation based way. We further discuss our approach in the numerical evaluation section.

We now discuss in more detail the two previous points. The policy improvement step becomes, for every $\mathbf{b} \in \bar{\mathcal{B}}^{(k+1)}$,

$$\widehat{K}_{\mathbf{b}}^{(k+1)} = \sum_{\mathbf{g},\mathbf{h}} f(\mathbf{g},\mathbf{h}) \max_{a \in \mathcal{A}_{(\mathbf{b},\mathbf{g},\mathbf{h})}} \left\{ r_\alpha(\boldsymbol{\tau}, \boldsymbol{\rho} | \mathbf{h}) + \widetilde{K}_{\mathbf{b}'}^{(k)} \right\} \quad (11)$$

where $\mathbf{b}'$ is defined according to (10). $\widehat{K}_{\mathbf{b}}^{(k+1)}$ represents the approximate value function at step $k+1$ and is defined only in the subset $\bar{\mathcal{B}}^{(k+1)}$, whereas $\widetilde{K}_{\mathbf{b}}^{(k)}$ is such that

$$\widetilde{K}_{\mathbf{b}}^{(k)} = \widehat{K}_{\mathbf{b}}^{(k)}, \quad \text{if } \mathbf{b} \in \bar{\mathcal{B}}^{(k)}. \quad (12)$$

In the second phase of the algorithm, for all $\mathbf{b} \notin \bar{\mathcal{B}}^{(k)}$, $\widetilde{K}_{\mathbf{b}}^{(k)}$ is derived exploiting (12) with an interpolation process or using a mean square error approximation. In practice, $\widetilde{K}_{\mathbf{b}}(r_k)$ is designed in order to approximate the true function $K_{\mathbf{b}}^{(k)}$. We remark that $\widehat{K}_{\mathbf{b}}^{(k+1)}$ is defined only in $\bar{\mathcal{B}}^{(k)}$, whereas $\widetilde{K}_{\mathbf{b}}^{(k)}$ is defined for every $\mathbf{b} \in \mathcal{B}$.

C. Complexity

For every given $(\mathbf{b}, \mathbf{g}, \mathbf{h})$, the policy update step requires to iterate over all the possible actions $a \in \mathcal{A}_{(\mathbf{b},\mathbf{g},\mathbf{h})}$ (see Equation (9)), and thus to perform $|\mathcal{A}_{(\mathbf{b},\mathbf{g},\mathbf{h})}|$ operations. Therefore, the number of operations of a generic step of the traditional VIA is

$$\sum_{\mathbf{b} \in \mathcal{B}} \sum_{\mathbf{g},\mathbf{h}} |\mathcal{A}_{(\mathbf{b},\mathbf{g},\mathbf{h})}|, \quad (13)$$

whereas for iteration k of App-VIA, we have

$$\sum_{\mathbf{b} \in \bar{\mathcal{B}}^{(k)}} \sum_{\mathbf{g},\mathbf{h}} |\mathcal{A}_{(\mathbf{b},\mathbf{g},\mathbf{h})}|. \quad (14)$$

The key point is that the complexity of the traditional value iteration depends upon the size of $\mathcal{B}$, whereas App-VIA depends only upon $|\bar{\mathcal{B}}^{(0)}|, |\bar{\mathcal{B}}^{(1)}|, \ldots$. Exploiting the structure of the state space, $\bar{\mathcal{B}}^{(k)}$ can be chosen as in [2, Sec. VI] to obtain $|\bar{\mathcal{B}}^{(k)}| \ll |\mathcal{B}|$ and significantly reduce the complexity of the problem. For example, in our numerical results of Section V, we have $|\mathcal{B}| \approx 25000$ (two batteries with $b_{1,\max} = b_{2,\max} = 160$) and $|\bar{\mathcal{B}}^{(k)}| = 41, \ \forall k$.

D. Convergence Properties

In the following we show that, provided that the approximation $\widetilde{K}_{\mathbf{b}}^{(k)}$ is sufficiently good, the long-term reward of App-VIA is a good approximation of VIA.

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

$$K_{\mathbf{b}}^{(k+1)} = T\left(\mathcal{K}^{(k)}, \mathbf{b}\right), \quad \forall \mathbf{b} \in \mathcal{B}, \quad (15)$$

$$\widehat{K}_{\mathbf{b}}^{(k+1)} = T\left(\widetilde{\mathcal{K}}^{(k)}, \mathbf{b}\right), \quad \forall \mathbf{b} \in \bar{\mathcal{B}}^{(k+1)}, \quad (16)$$

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

Proposition 1. *After N iterations, the cost-to-go functions of App-VIA and VIA differ by at most $N\epsilon$, i.e.,*[3]

$$\| \underbrace{\mathcal{K}^{(N)}}_{\text{VIA}} - \underbrace{\widetilde{\mathcal{K}}^{(N)}}_{\text{App}-\text{VIA}} \|_\infty \leq N\epsilon \quad (17)$$

with

$$\epsilon \triangleq \max_{k=0,\ldots,N-1} \max_{\mathbf{b} \in \mathcal{B}} \left\{ \widetilde{K}_{\mathbf{b}}^{(k+1)} - T\left(\widetilde{\mathcal{K}}^{(k)}, \mathbf{b}\right) \right\} \quad (18)$$

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 $K_{\mathbf{b}}^{(k+1)}$ 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), $Q_{\max} = 3$ W (maximum transfer power), $P_{1,\min} = P_{2,\min} = 1$ mW and $P_{1,\max} = P_{2,\max} = 10$ mW, $B_{\max} \triangleq B_{1,\max} = B_{2,\max} = 0.125$ 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

[3]We adopt the notation $\|\mathcal{K}^{(N)} - \widetilde{\mathcal{K}}^{(N)}\|_\infty \triangleq \max_{\mathbf{b} \in \mathcal{B}} |K_{\mathbf{b}}^{(N)} - \widetilde{K}_{\mathbf{b}}^{(N)}|$.

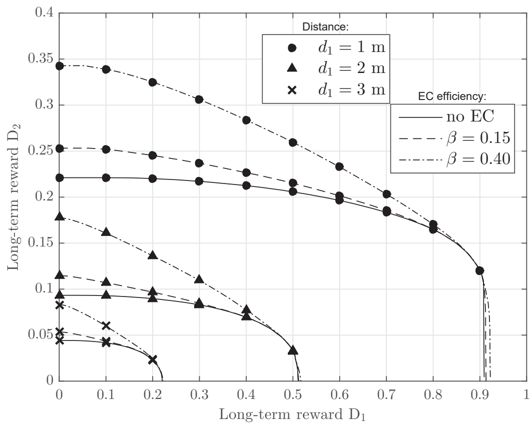

Figure 3: Throughput region of D_1 and D_2 for several distances d_1, $d_2 = d_1 + 2$ with and without energy cooperation with Rayleigh fading.

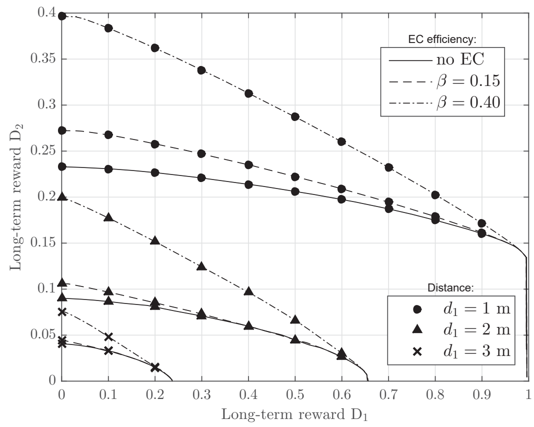

Figure 4: Throughput region of D_1 and D_2 for several distances d_1, $d_2 = d_1 + 2$ with and without energy cooperation with Nakagami fading with parameter 5.

meters apart, the energy cooperation efficiency $\beta \triangleq \beta_{1,2} = \beta_{2,1}$ 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 β and d_1. The curves are generated by changing the weight value α in Equation (3). When $\alpha = 0$, (3) degenerates to $\arg\max_\mu\{G_{2,\mu}\}$ 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 α 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

[1] A. Biason and M. Zorzi, “Long-term throughput optimization in WPCN with battery-powered devices,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2016.

[2] ——, “Battery-powered devices in WPCNs,” in *arXiv:1601.06847*. submitted to IEEE Trans. Commun., Jan. 2016.

[3] H. Ju and R. Zhang, “Throughput maximization in wireless powered communication networks,” *IEEE Trans. Wireless Commun.*, vol. 13, no. 1, pp. 418–428, Jan. 2014.

[4] ——, “User cooperation in wireless powered communication networks,” in *IEEE Global Communications Conference (GLOBECOM)*, Dec. 2014, pp. 1430–1435.

[5] ——, “Optimal resource allocation in full-duplex wireless-powered communication network,” *IEEE Trans. Commun.*, vol. 62, no. 10, pp. 3528–3540, Oct. 2014.

[6] L. Liu, R. Zhang, and K.-C. Chua, “Multi-antenna wireless powered communication with energy beamforming,” *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4349–4361, Dec. 2014.

[7] D. T. Hoang, D. Niyato, P. Wang, and D. I. Kim, “Optimal decentralized control policy for wireless communication systems with wireless energy transfer capability,” in *Proc. IEEE Int. Conf. Commun. (ICC)*, June 2014, pp. 2835–2840.

[8] A. Biason and M. Zorzi, “Joint transmission and energy transfer policies for energy harvesting devices with finite batteries,” *IEEE J. Sel. Areas in Commun.*, vol. 33, no. 12, pp. 2626–2640, Dec. 2015.

[9] B. Gurakan, O. Ozel, J. Yang, and S. Ulukus, “Energy cooperation in energy harvesting communications,” *IEEE Trans. Commun.*, vol. 61, no. 12, pp. 4884–4898, Dec. 2013.

[10] B. Gurakan, O. Ozel, and S. Ulukus, “Optimal energy and data routing in networks with energy cooperation,” *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 857–870, Feb. 2016.

[11] W. Ni and X. Dong, “Energy harvesting wireless communications with energy cooperation between transmitter and receiver,” *IEEE Trans. Commun.*, vol. 63, no. 4, pp. 1457–1469, Apr. 2015.

[12] K. Tutuncuoglu and A. Yener, “Energy harvesting networks with energy cooperation: Procrastinating policies,” *IEEE Trans. Comm*, vol. 63, no. 11, pp. 4525–4538, Nov. 2015.

[13] “Powercast corporation, TX91501 users manual & P2110s datasheet.”

[14] G.-M. Yang, C.-C. Ho, R. Zhang, and Y. Guan, “Throughput optimization for massive MIMO systems powered by wireless energy transfer,” *IEEE J. Sel. Areas in Commun.*, vol. 33, no. 8, pp. 1640–1650, Aug. 2015.

[15] X. Kang, C. K. Ho, and S. Sun, “Full-duplex wireless-powered communication network with energy causality,” *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5539–5551, Oct. 2015.

[16] A. Kurs, A. Karalis, R. Moffatt, J. D. Joannopoulos, P. Fisher, and M. Soljačić, “Wireless power transfer via strongly coupled magnetic resonances,” *Science*, vol. 317, no. 5834, pp. 83–86, July 2007.

[17] D. Bertsekas, *Dynamic programming and optimal control*. Athena Scientific, Belmont, Massachusetts, 2005.

[18] D. Bertsekas and J. N. Tsitsiklis, *Neuro-Dynamic Programming*. Athena Scientific, Belmont, Massachusetts, 1996.