This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2846657, IEEE Transactions on Communications

Resource Allocation under Sequential Resource Access

Ali Tajer, *Senior Member*, Maha Zohdy *Student Member*, Khawla Alnajjar *Member*

***Abstract*—This paper treats the problem of optimal resource allocation over time in a finite-horizon setting, in which the resource become available only sequentially and in incremental values and the utility function is concave and can freely vary over time. Such resource allocation problems have direct applications in data communication networks (e.g., energy harvesting systems). This problem is studied extensively for special choices of the concave utility function (time-invariant and logarithmic) in which case the optimal resource allocation policies are well-understood. This paper treats this problem in its general form and analytically characterizes the structure of the optimal resource allocation policy, and devises an algorithm for computing the exact solutions analytically. An observation instrumental to devising the provided algorithm is that there exist time instances at which the available resources are exhausted, with no carry-over to future. This algorithm identifies all such instances, which in turn facilitates breaking the original problem into multiple problems with significantly reduced dimensions. Furthermore, some widely-used special cases in which the algorithm takes simpler structures are characterized, and the application to the energy harvesting systems is discussed. Numerical evaluations are provided to assess the key properties of the optimal resource allocation structure and to compare the performance with the generic convex optimization algorithms.**

***Index Terms*—Sequential access, energy harvesting, delay-limited, optimization.**

I. Introduction

A. Overview

Consider a resource allocation problem over a finite time horizon $T \in \mathbb{N}$. The resource is made available for utilization *sequentially* over time and in *increments*. Such resource allocation models manifest in a wide range of power allocation and scheduling objectives in communication systems. For instance, in energy harvesting networks the transmitters rely partly or entirely on ambient sources in their surrounding environments. In such systems, the energy resources are available only sequentially and incrementally over time as they are harvested. Similarly, the packet transmission systems under stringent quality-of-service (QoS) constraints constitute another class of resource allocation problems in which the data packets to be transmitted arrive sequentially over time at the transmitter, while all the arriving information packets are required to be delivered to their destination by a given deadline or by using a given amount of energy. In order to set the context for further discussions, and by relegating closer review of the motivations and related literature to Section I-B, we next provide the statement of the problem.

A. Tajer and M. Zohdy are with the department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180.

K. Alnajjar is with the Department of Electrical Engineering, University of Sharjah, Sharjah, UAE.

This research was supported in part by the U. S. National Science Foundation under Grants ECCS-1455228 and DMS-1737976, and the CAREER Award ECCS-1554482.

In a time-slotted setting, we denote the incremental amount of resource made available during time slot $t \in \{1, \dots, T\}$ by $s_t \in \mathbb{R}^+$, and denote the actual amount of resource utilized during time slot $t \in \{1, \dots, T\}$ by $x_t \in \mathbb{R}^+$. The resource is assumed to be used only causally, leading to the following set of T resource utilization constraints:

$$\sum_{i=1}^{t} x_i \leq \sum_{i=1}^{t} s_i \ , \qquad \forall t \in \{1, \dots, T\} \ . \tag{1}$$

Accordingly, we denote the resource vector by $\boldsymbol{s} \triangleq [s_1, \dots, s_T]$ and denote the vector of utilized resource over time by $\boldsymbol{x} \triangleq [x_1, \dots, x_T]$. Also, we define the utility function $f_t : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ as the measure of the contribution of the amount of resource utilized during time slot $t \in \{1, \dots, T\}$, i.e., x_t. We assume that all functions $\{f_t : t \in \{1, \dots, T\}\}$ are *differentiable*, *non-decreasing*, and *strictly concave*, and denote the aggregate utility gleaned over the entire time horizon by $F(\boldsymbol{x}) \triangleq \sum_{t=1}^{T} f_t(x_t)$. Based on these definitions, the resource allocation problem under the sequential access to the resource over a finite time-horizon can be formalized as

$$\mathcal{P}(\boldsymbol{s}) \triangleq \begin{cases} \max\limits_{\boldsymbol{x}} & F(\boldsymbol{x}) \\ \text{s.t.} & \sum\limits_{i=1}^{t} x_i \leq \sum\limits_{i=1}^{t} s_i \ , \ \ \forall t \in \{1, \dots, T\} \\ & \boldsymbol{x} \succeq 0 \end{cases} . \tag{2}$$

The problem in (2), in its special cases with some constraints relaxed, subsumes an extensive body of well-understood problems, e.g., power allocation in parallel channels [1] and power allocation in single-user multi-antenna channels [2] when $\{s_i = 0 : i \in \{1, \dots, T-1\}\}$. In this paper, we leverage the structure of the convex optimization problem formalized in (2) and provide the optimal solution for the general form analytically. First, we study the properties of the optimal solution, based on which we devise an algorithm to provide the optimal solution as will be shown in Section I-C.

B. Motivation and Related Work

In this subsection we provide a more detailed overview of two classes of communication systems and their existing relevant literature in which resource allocation objectives can be formalized as problem $\mathcal{P}(\boldsymbol{s})$ defined in (2).

Energy Harvesting Communication Systems: Among the varying conditions communication systems experience, causal and incremental availability of the energy resource introduces a new dimension in resource allocation that does not exist in systems facing *average* or *aggregate* constraints on the resource. Energy harvesting networks, in which the transmitters rely partly or entirely on ambient sources in their surrounding environments, represent one class of such communication systems in which the resource is available only sequentially.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2846657, IEEE Transactions on Communications

Energy harvesting networks empowered by perpetual sources of power, are especially promising alternatives to systems with lifetime-limited batteries. In such systems, nevertheless, the availability of energy becomes sporadic and temporally volatile, in which case devising optimal policies for efficient utilization of the harvested energy directly translates into how continually the communication link can be sustained by relying on the harvested energy. In such systems, optimally balancing energy consumption over time leads to solving problems of the form in (2).

Optimal resource allocation policies under different settings and objectives are studied extensively. In particular, and most relevant to the scope of this paper, in the single-user energy harvesting channels, optimal power allocation policies are studied under a number of assumptions on the battery size for storing the harvested energy (finite versus infinite), and information available regarding the causality of energy harvesting, and wireless channel fading process (slow versus fast). Specifically, the studies in [3] and [4] consider infinite-capacity batteries, establish certain properties of the optimal policies, and devise the *directional water-filling* approach to power allocation in static as well as fading wireless channels. Extensions to random channel conditions and finite-capacity batteries for static channels are studied in [5]–[7].

Enforcing a finite battery capacity induces constraints on the policies, which are driven by the possibility of battery overflow at the instances of harvesting energy. Extensions to such finite-battery settings when facing inefficiencies in battery storage is investigated in [8]. The studies in [9] and [10] address causal and non-causal availability of the channel state information. The settings in which the channels undergo random fading processes as well as the associated policies are studied in [11]–[14]. A closely related problem is studied in [15], which investigates an optimal power allocation scheme that minimizes the delay by which a given amount of data is successfully transmitted. A greedy power allocation policy and the conditions under which the policy is optimal is studied in [16]. Properties of online optimal policies in fading channels are analyzed in [17]. The impacts of finite-horizon on optimal power control is studied in [18] and [19]. The undesired effect of non-ideal power circuit at the transmitters of the additive white Gaussian noise channel (AWGN) is analyzed in [20]. A universal approach that incorporates the effects of the energy arrival process is presented in [21], and energy harvesting sensor networks are studied in [22] and [23].

In addition to the single-user systems, multiuser energy harvesting systems are also studied extensively. Specifically, the multiple access channels under different settings are studied in [24]–[28], broadcast channels are studied in [29], and interference channels are studied in [30]–[32]. Other studies that do not primarily focus on resource (power) optimization, but indirectly involve that include analyzing the capacity of the AWGN channel [33]–[39], and that of the Gaussian multiple access channel [40], [41].

QoS-constrained Systems: Optimizing the efficiency of packet transmission systems under stringent quality-of-service constraints is another class of resource allocation problems solving which is equivalent to the problem in (2). In such systems, the data packets to be transmitted arrive sequentially over time at the source and all the arriving information packets requires to be delivered to their destination by a given deadline or by using a given amount of energy. For instance, the studies in [42], [43] consider minimizing the energy-cost used to transmit data packets through wireless channels subject to given delay or other quality of service constraints. Maximizing the transmission throughput of an energy- or time-constrained transmitter over fading channels is studied in [44]. Under a fixed delay constraint, a transmission schedule that maximizes the battery life-time is derived in [45], while the study in [46] considers minimum-energy scheduling problems over fading multiple-access and broadcast channels. Also, the recent study in [47] analyzes proactive content caching from an energy efficiency perspective. Moreover, a scheduling algorithm with real-time constraints was presented in [48].

C. Contributions

In this paper we treat the problem in (2) in its general from. This problem in the special form that the utility functions are homogeneous (identical) over time, i.e., $f_t = f$ for all $t \in \{1, \dots, T\}$, is considered in [43] to address problems arising in QoS-constrained energy optimization. This study develops a calculus approach and offers an optimal algorithm for determining the optimal solution. Furthermore, in the context of energy harvesting systems, while the optimal structure is not fully characterized, some of the properties of the optimal solution are delineated in [4], [49], and the special case of homogeneous utility functions, i.e., static channels, is treated in [4], [7].

In this paper, we analytically characterize the properties of the optimal solution, based on which we provide an algorithm that determines the exact optimal solution. The key component of the structure of the optimal solution is that there exist time instances at which all the available resources are exhausted and the amount of available resources is set to zero. This is in contrast to the other time instances, at which always a fraction of the resources are reserved to be consumed in the future. The proposed algorithm progressively determines all these time instances, and we call the associated constraints the *dominant* constraints. Once these instances are known, the finite-horizon optimization problem reduces to a collection of smaller problems that do *not* include any inequality constraint, and each involves only one equality constraint. Furthermore, we also comment on a stochastic account of the problem in (2) and show that when the amount of available resources and/or the utility functions bear stochastic uncertainties, optimizing the *expected* aggregated utility subject to chance constraints on the available resources can be translated into and solved by a problem of the form in (2).

The remainder of the paper is organized as follows. In Section II we provide the optimal structure of the optimal solution to (2) and algorithms that identify the optimal solution in its general form, and for some common special cases. In Section III we discuss the direct application to the single-user energy harvesting system. The numerical evaluations are provided in Section IV to assess the structure of the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2846657, IEEE Transactions on Communications

solution and also compare the performance with that of the generic convex optimization algorithms. Section V concludes the paper, and all the proofs are relegated to the appendices.

II. Optimal Solution: Properties and Algorithm

The objective in this section is to analytically characterize $\boldsymbol{x}^*$, which we define as the solution to $\mathcal{P}(\boldsymbol{s})$. The solution $\boldsymbol{x}^*$ is unique since all the constraints are linear and the utility function is strictly concave. We start by considering the *offline* resource allocation problem, in which the resource vector $\boldsymbol{s}$ and the utility functions $\{f_t : t \in \{1,\dots,T\}\}$ are known *deterministically*. We characterize the optimal solution *analytically*, and discuss the generalization to the settings in which these terms bear stochastic uncertainties in Section II-H.

A. Algorithm for Finding the Optimal Solution

We start by providing an algorithm that identifies the exact solution to $\mathcal{P}(\boldsymbol{s})$, discuss its complexity in Section II-B, present an overview of the scheme of the proofs in Section II-C, and present the detailed steps of the analysis for establishing its optimality properties in sections II-D and II-E. In these latter two subsections, specifically, we show that the optimal solution $\boldsymbol{x}^*$ has two key properties, which constitute the main structure of Algorithm 1 for analytically solving $\mathcal{P}(\boldsymbol{s})$. The first property is that the set of optimal values $\{x_1^*,\dots,x_T^*\}$ can be partitioned into d mutually exclusive subsets separated at time instants $t \in \{u_0,u_1,\dots,u_d\}$, which we can find analytically. We denote these subsets by

$$\{x_1^*,\dots,x_{u_1}^*\}, \quad \dots \quad \{x_{u_{d-1}+1}^*,\dots,x_T^*\}\ , \tag{3}$$

where we have set $u_0 = 0$ and $U_d = T$. We show that each subset can be characterized analytically and independently of the rest. Built on this observation, secondly, we show that among all the constraints of $\mathcal{P}(\boldsymbol{s})$, i.e.,

$$\sum_{i=1}^{t} x_i \ \leq\ \sum_{i=1}^{t} s_i\ , \qquad \forall t \in \{1,\dots,T\}\ . \tag{4}$$

the constraints corresponding to $t \in \{u_1,\dots,u_d\}$ hold with *equality*, and all others hold with *strict inequality*. Finally, based on these two properties we show that finding $\boldsymbol{x}^*$ via solving $\mathcal{P}(\boldsymbol{s})$ reduces to solving a number of problems with a similar structure, but with reduced dimension.

The detail steps of solving $\mathcal{P}(\boldsymbol{s})$ are provided in Algorithm 1. This algorithm receives the resource vector $\boldsymbol{s}$ as its input and produces the optimal resource allocation solution $\boldsymbol{x}^*$. It consists of one outer loop (lines 3-13) the purpose of which is progressively determining the indices of the time instants $\{u_i : i \in \{1,\dots,d\}\}$. Each of the d outer loops involves an inner loop (lines 6-9), which finds a part of the optimal solution, and specifically in the iteration i of the outer loop, the inner loop finds the optimal values $\{x_i^* : i \in \{u_{i-1}+1,\dots,u_i\}\}$. This inner loops within the $i^{\rm th}$ iteration solve optimization problems $\mathcal{Q}_{u_{i-1}\to t}(\boldsymbol{s})$ for all values of $t \in \{u_{i-1}+1,\dots,T\}$, where corresponding to each pair $m<n$ we have defined the auxiliary problem

$$\mathcal{Q}_{m\to n}(\boldsymbol{s}) \triangleq \begin{cases} \max\limits_{\boldsymbol{x}} & \sum\limits_{i=m+1}^{n} f_i(x_i) \\ \text{s.t.} & \sum\limits_{i=m+1}^{n} x_i = \sum\limits_{i=m+1}^{n} s_i\ , \quad \forall t \in \{1,\dots,T\} \\ & \boldsymbol{x} \succeq 0 \end{cases}, \tag{5}$$

It is noteworthy that $\mathcal{Q}_{m\to n}(\boldsymbol{s})$ has a unique globally optimal solution, since its utility function is strictly concave.

Algorithm 1 - Solving $\mathcal{P}(\boldsymbol{s})$ for any given resource vector $\boldsymbol{s}$

1: **input** $\boldsymbol{s}$
2: **initialize** $t=1$, $d=0$ and $u_0=0$,
3: **while** $u_d \leq T-1$
4: $\quad d \leftarrow d+1$
5: $\quad$**set** $\mathcal{A}_d \triangleq \{u_{d-1}+1,\dots,T\}$
6: $\quad$**for** $t \in \mathcal{A}_d$
7: $\quad\quad$set $\boldsymbol{w}^{d,t}$ as the solution to $\mathcal{Q}_{u_{d-1}\to t}(\boldsymbol{s})$
8: $\quad\quad$set $q^{d,t} \triangleq \min\left\{\frac{{\rm d}f_i}{{\rm d}x}(w_i^{d,t}) : i \in \{u_{d-1}+1,\dots,t\}\right\}$
9: $\quad$**end for**
10: $\quad u_d \triangleq \arg\max\limits_{t\in\mathcal{A}_d} q^{d,t}$ (if not unique, select the smallest [a])
11: $\quad v_d \triangleq \max\limits_{t\in\mathcal{A}_d} q^{d,t}$
12: $\quad \boldsymbol{z}^d \triangleq \boldsymbol{w}^{d,u_d}$
13: **end while**
14: **for** $i \in \{1,\dots,d\}$
15: $\quad$**for** $t \in \mathcal{D}_i = \{u_{i-1}+1,\dots,u_i\}$
16: $\quad\quad x_t \triangleq z_t^i$
17 : $\quad$**end for**
18 : **end for**
19: **output** $\boldsymbol{x}$ and d and $\{u_1,\dots,u_d\}$ and $\{v_1,\dots,v_d\}$

[a] For the convenience in the analyses, throughout the rest of the paper we assume that u_d is unique. In case that it is not unique, by selecting the smallest choice all the analyses remain valid.

B. Computational Complexity

The significance of obtaining the optimal solution $\boldsymbol{x}^*$ *analytically* is the substantial reduction in the computational complexity. To furnish the relevant context, we remark that since the utility functions are strictly concave, the generic approaches in convex optimization can be readily applied to the problem at hand. In particular, the primal-dual interior-point (IP) methods are known to be extremely efficient and capable of handling large-scale nonlinear problems. From a computational perspective, the complexity of IP methods is shaped primarily by two factors, namely the desired level of accuracy in the solution they provide (i.e., closeness to the optimal solution) and the nature of the utility functions (e.g., linear or quadratic). In the IP methods, it is well-investigated that for *linear* utility functions, the computational complexity scales at the rate $O(\sqrt{T}\ln\frac{1}{\epsilon})$, where T is the dimension of the problem and ϵ accounts for the error of the solution provided by the IP method, i.e., the difference between the optimal solution and the solution provided by the IP method. For *non-linear* utility functions, which is the case in this paper, the complexity is higher, and except for special cases (e.g., quadratic) the general complexity is unknown. On the other hand, Algorithm 1 provides the *exact* optimal solutions, which corresponds to guaranteeing that $\epsilon = 0$ for the output of

Algorithm 1, achieving which by the IP method results in theoretically unbounded computational complexity. The same trend is true for other numerical approaches as well, and in Section IV we provide numerical comparisons between the computational complexities. Finally we remark that the complexity of Algorithm 1 is $O(T)$, since in the worst case it has T iterations. Each iteration involves solving a problem of the form $\mathcal{Q}_{m\to n}(\boldsymbol{s})$. The solution to $\mathcal{Q}_{m\to n}(\boldsymbol{s})$ often has a closed-form when the utility functions are specified, and as a result as it is customary, the computational complexity is considered negligible.

C. Scheme of the Proofs

Before proceeding to the details the proofs, we provide a scheme of the steps involved. The objective is to characterize the key properties of $\boldsymbol{x}^*$ as the optimal solution of $\mathcal{P}(\boldsymbol{s})$. For the analytical purposes, we construct another resource allocation vector $\tilde{\boldsymbol{x}}$ as the output of Algorithm 1 when its input $\boldsymbol{s}$ is replaced by $\boldsymbol{x}^*$. It is noteworthy that this serves merely an an auxiliary solution which are not interesting in computing, but rather we investigate its properties. Specifically, we show the following properties for $\tilde{\boldsymbol{x}}$:

1) From the construction of $\tilde{\boldsymbol{x}}$, it can be readily verified that $\tilde{\boldsymbol{x}}$ satisfies all the constraints of $\mathcal{P}(\boldsymbol{s})$. As a result due to the optimality of $\boldsymbol{x}^*$, the utility corresponding to $\tilde{\boldsymbol{x}}$ cannot exceed the utility corresponding to $\boldsymbol{x}^*$, i.e., $F(\boldsymbol{x}^*)\geq F(\tilde{\boldsymbol{x}})$. This is established in Lemma 2.
2) Also, from the construction of $\tilde{\boldsymbol{x}}$, we prove that $F(\boldsymbol{x}^*)\leq F(\tilde{\boldsymbol{x}})$. This is established in Lemma 3.
3) By leveraging the results of lemmas 2 and 3 we subsequently have $\tilde{\boldsymbol{x}}=\boldsymbol{x}^*$. This implies that if we initiate Algorithm 1 with $\boldsymbol{x}^*$, it will produce the same vector $\boldsymbol{x}^*$ as its output. This is established in Theorem 1.
4) Finally, we show that initiating Algorithm 1 with inputs $\boldsymbol{s}$ and $\boldsymbol{x}^*$ results in the same resource allocation vectors. This is formalized in Theorem 2, which in conjunction with Theorem 1 establishes that the output of the Algorithm 1 is the unique desired vector $\boldsymbol{x}^*$.

Besides these main items, we also show that $\tilde{\boldsymbol{x}}$ and the value of the utility functions corresponding to this resource allocation vector have a number of algebraic properties established in lemmas 1, 4, and 5, which $\boldsymbol{x}^*$ also inherits due to the observation that $\boldsymbol{x}^*=\tilde{\boldsymbol{x}}$.

D. Grouping the Constraints

We start the analysis by showing that the set of the optimal values $\{x_1^*,\dots,x_T^*\}$ has the key property that this set can be partitioned into smaller subsets, such that the elements within one subset are closely related. These properties are established via lemmas 1-5. For this purpose, we first establish a number of properties for $\tilde{\boldsymbol{x}}$, which is the output of Algorithm 1 when its input $\boldsymbol{s}$ is replaced with the optimal solution $\boldsymbol{x}^*$. It is noteworthy that it is not our objective to actually compute $\tilde{\boldsymbol{x}}$, but rather we aim to show that when such an auxiliary term is constructed according to the rules specified in Algorithm 1, it satisfies certain desired properties. Hence, the purpose of this algorithm is only proving the properties, as a result of which for proving the properties and does *not* involve knowing the optimal solution $\boldsymbol{x}^*$, or actually computing $\tilde{\boldsymbol{x}}$.

In order to construct $\tilde{\boldsymbol{x}}$, Algorithm 1 admits $\boldsymbol{x}^*$ as its input, and based on that *successively* partitions the set of constraints $\{\sum_{i=1}^t x_i \leq \sum_{i=1}^t s_i : t\in\{1,\dots,T\}\}$ into d disjoint subsets of constraints. Specifically, it returns time indices $0=u_0<u_1<\cdots<u_d=T$, and partitions the set $\{1,\dots,T\}$ into d disjoint sets:

$$\mathcal{D}_i \triangleq \{u_{i-1}+1,\dots,u_i\}\ , \qquad \text{for} \quad i\in\{1,\dots,d\}\ . \quad (6)$$

Furthermore, this algorithm computes the metrics $\{v_i : i\in\{1,\dots,d\}\}$ and assigns v_i to the set $\mathcal{D}_i$. Once the dominant constraints are known, solving $\mathcal{P}(\boldsymbol{x}^*)$ reduces to solving a collection of smaller problems in the form of $\mathcal{Q}_{u_{u-1}\to u_i}(\boldsymbol{x}^*)$ defined in (5). The properties of $\tilde{\boldsymbol{x}}$ are formalized in the following lemmas.

Lemma 1. *When Algorithm 1 is initiated with $\boldsymbol{x}^*$, for given $m\in\{1,\dots,d\}$ and $t\in\mathcal{A}_m\triangleq\{u_{m-1}+1,\dots,T\}$, we have*

$$\begin{array}{rcll}\frac{\mathrm{d}f_i}{\mathrm{d}x}(w_i^{m,t}) &=& \lambda_{m,t}\ , & \forall i\in\{u_{m-1}+1,\dots,t : w_i^{m,t}>0\}\\ \frac{\mathrm{d}f_i}{\mathrm{d}x}(w_i^{m,t}) &>& \lambda_{m,t}\ , & \forall i\in\{u_{m-1}+1,\dots,t : w_i^{m,t}=0\}\end{array},$$

where we have defined $\boldsymbol{w}^{m,t}\triangleq[w_1^{m,t},\dots,w_T^{m,t}]$, and $\lambda_{m,t}\in\mathbb{R}_+$ is a strictly positive real constant. Furthermore we have $q^{m,t}=\lambda_{m,t}$.

Proof: See Appendix A. ■

Lemma 2. *Vector $\tilde{\boldsymbol{x}}$ generated by Algorithm 1 satisfies all the constraints of $\mathcal{P}(\boldsymbol{s})$.*

Proof: See Appendix B. ■

Lemma 3. *The vector $\tilde{\boldsymbol{x}}$ satisfies $F(\tilde{\boldsymbol{x}})\geq F(\boldsymbol{x}^*)$, and the equality holds if and only if $\boldsymbol{x}^*=\tilde{\boldsymbol{x}}$.*

Proof: See Appendix C. ■

The results of lemmas 1-3, collectively, establish the optimality of $\tilde{\boldsymbol{x}}$ generated by Algorithm 1, which is formalized by the following theorem.

Theorem 1. *By initiating Algorithm 1 with $\boldsymbol{x}^*$ as the optimal solution to $\mathcal{P}(\boldsymbol{s})$, then $\tilde{\boldsymbol{x}}$ generated by Algorithm 1 is equal to the optimal solution to $\mathcal{P}(\boldsymbol{s})$, i.e., $\tilde{\boldsymbol{x}}=\boldsymbol{x}^*$.*

E. Dominant Constraints

By leveraging the results in the previous subsection, which essentially partition the set of all constraints into a collection of d disjoint constraint sets, next we provide additional properties for these sets of constraints. Specifically, we show that in each of the given d sets, at least one constraint holds with equality, which we refer to as the *dominant* constraint. These d dominant constraints are the only constraints needed to characterize the optimal solution to $\boldsymbol{x}^*$. The following lemma represents an intermediate and instrumental step towards characterizing the set of dominant constraints of $\mathcal{P}(\boldsymbol{s})$. In particular, it establishes a connection among the derivative measures $q^{d,t}$ and v^d defined in Algorithm 1.

Lemma 4. *The sequence $\{v_1,\dots,v_d\}$ is strictly decreasing.*

Proof: See Appendix D. ■

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2846657, IEEE Transactions on Communications

We remark that the indices $\{u_i : i \in \{1,\dots,d\}\}$ and their associated constraint indices $\{v_i : i \in \{1,\dots,d\}\}$ have significant physical meanings in resource allocation. Specifically, the elements of $\{u_i : i \in \{1,\dots,d\}\}$ specify the time instances at which all the resources arrived by that time instance are consumed in their entirely. At other time instances, a fraction of the available resources is reserved for being consumed in the future time instances. This observation is formally demonstrated in the following lemma. Also, the measures $\{v_i : i \in \{1,\dots,d\}\}$ are the derivatives of the utility functions at the optimal solution $\boldsymbol{x}^*$ over time. Specifically, for all the indices in the range $t \in \mathcal{D}_{i+1}$, the derivatives of all the utility terms f_t at the non-zero optimal values of $\boldsymbol{x}^*$ are all the same, and equal to v_i, i.e., for to the set $\mathcal{D}_i$ is defined in (6) we have

$$v_i = \frac{\mathrm{d}f_t(x_t)}{\mathrm{d}x_t} \ , \quad \forall t \in \mathcal{D}_{i+1}, \quad \text{and} \quad \forall x_t \neq 0 \ .$$

Lemma 5. *Under the optimal solution $\boldsymbol{x}^*$, all the inequality constraints with indices included in $\{u_m : m \in \{1,\dots,d\}\}$ hold with equality, i.e.,*

$$\forall m \in \{1,\dots,d\} : \qquad \sum_{i=1}^{u_m} x_i^* = \sum_{i=1}^{u_m} s_i \ . \tag{7}$$

Proof: See Appendix E. ■

F. Initiating $\boldsymbol{x}^*$ via Algorithm

By leveraging the results of Lemma 4 and Lemma 5 in this subsection, we poof the optimality of Algorithm 1 for obtaining $\boldsymbol{x}^*$. So far we have shown that if we modify Algorithm 1 such that instead of inputting $\boldsymbol{s}$ we input the resource vector $\boldsymbol{x}^*$, then the output will be in fact the optimal solution $\boldsymbol{x}^*$. Next we show that initiating Algorithm 1 with either $\boldsymbol{x}^*$ or $\boldsymbol{s}$ yields the same output. The underlying insight is that this algorithm depends on $\boldsymbol{x}^*$ primarily for determining the metrics $\{v_i : i \in \{1,\dots,d\}\}$ and their associated constraint indices $\{u_i : i \in \{1,\dots,d\}\}$. By invoking the result of Lemma 5, we next show that for determining the sets $\{v_i : i \in \{1,\dots,d\}\}$ and $\{u_i : i \in \{1,\dots,d\}\}$, alternatively, we can also use the resource vector $\boldsymbol{s}$, based on which subsequently we can show that the outcome of Algorithm 1 based on the input $\boldsymbol{s}$ will be in fact the optimal solution $\boldsymbol{x}^*$. Insensitivity of Algorithm 1 to the choice of $\boldsymbol{x}^*$ by $\boldsymbol{s}$ as the input is formalized in the next lemma.

Lemma 6. *Denote the set of constraint indices yielded by Algorithm 1 by $\{u_i : i \in \{1,\dots,d\}\}$, and denote the counterpart set when in Algorithm 1 $\boldsymbol{x}^*$ is replaced with $\boldsymbol{s}$ by $\{\bar{u}_i : i \in \{1,\dots,\bar{d}\}\}$. We have $\{u_i : i \in \{1,\dots,d\}\} = \{\bar{u}_i : i \in \{1,\dots,\bar{d}\}\}$.*

Proof: See Appendix F. ■

Based on the result of Lemma 6, in the following theorem we establish the optimality of Algorithm 1, that is it produces $\boldsymbol{x}^*$ with it is initiated with input $\boldsymbol{s}$.

Theorem 2. *By admitting $\boldsymbol{s}$ as its input, Algorithm 1 generates the optimal solution of $\mathcal{P}(\boldsymbol{s})$.*

Proof: See Appendix G. ■

G. Homogeneous Utility Functions

In this subsection we consider the settings in which the utility functions are all identical, i.e., $f_t = f$ for all $t \in \{1,\dots,T\}$. While in such settings we can solve $\mathcal{P}(\boldsymbol{s})$ directly via Algorithm 1, nevertheless, by leveraging the homogeneity structure, this algorithm can be significantly simplified. Specifically, we show that the inner loop that solves an optimization problem for all the future time instances (lines 6-9) can be avoided, and the indices of the dominant constraints and the associated resource allocation scheme can be found directly based on $\boldsymbol{s}$. Specifically, in the following lemma, we show that the set of dominant constraints $\{u_m : m \in \{1,\dots,d\}\}$ can be found without solving the optimization problems of the form $\mathcal{Q}_{u_{d-1}\to t}(\boldsymbol{s})$, unlike in the general form.

Lemma 7. *For problem $\mathcal{P}(\boldsymbol{s})$ with identical utility function $f_t = f$, the indices of the dominant constraints $\{u_m : m \in \{1,\dots,d\}\}$ are given by*

$$u_m = \arg\min_{t \in \mathcal{A}_m} \frac{1}{t - u_{m-1}} \sum_{i=u_{m-1}+1}^{t} s_i \ . \tag{8}$$

Proof: See Appendix H. ■

Based on the result of Lemma 7, we provide Algorithm 2 as a simpler algorithm for obtaining the optimal solution $\boldsymbol{x}^*$ to the problem in (2) by admitting the resource vector $\boldsymbol{s}$ as the input. The optimality of the outcome of the algorithm $\boldsymbol{x}^*$ is stated in Theorem 3.

Algorithm 2 - Computing $\boldsymbol{x}^*$ under homogeneous utility functions

```
1: input s
2: initialize d = 0 and u_d = 0
3: while u_d ≤ T − 1
4:     d ← d + 1
5:     set A_d ≜ {u_{d−1} + 1, ..., T}
6:     u_d = arg min_{t ∈ A_d} 1/(t − u_{d−1}) Σ_{i=u_{d−1}+1}^{t} s_i
7:     β_d ≜ 1/(u_d − u_{d−1}) Σ_{i=u_{d−1}+1}^{u_d} s_i
8: end while
9: x* = Σ_{m=1}^{d} β_m · [0_{u_{m−1}} , 1_{u_m−u_{m−1}} , 0_{T−u_m}]
```

Theorem 3. *The optimal solution to the problem $\mathcal{P}(\boldsymbol{s})$ under homogeneous utility functions is yielded by Algorithm 2, and takes the closed form $\boldsymbol{x}^* = \sum_{m=1}^{d} \beta_m \cdot \left[\mathbb{0}_{u_{m-1}}\ ,\ \mathbb{1}_{u_m-u_{m-1}}\ ,\ \mathbb{0}_{T-u_m}\right]$. where $\mathbb{0}_\ell$ and $\mathbb{1}_\ell$ are ℓ-dimensional vectors of all zeros and all ones, respectively.*

Proof: See Appendix I. ■

We comment that a similar solution structure is provided in [15] for treating the problem of optimal power allocation over a point-to-point *static* channel in an energy harvesting system.

H. Stochastic Uncertainties

In this subsection we consider a class of utility functions and resource vectors the true values of which are known only causally, and otherwise bear stochastic uncertainties. We show that solving this class of stochastic problems can be reduced to solving problems of the form in (2). To formalize such settings, we assume $\{s_t : t \in \{1,\dots,T\}\}$ are independent and identically distributed (i.i.d.) random variables unknown non-causally. Furthermore, to capture the uncertainties in $f_t(x)$ we assume that the function depends on an unknown random variable α_t, and denoted it by $f_t(x,\alpha_t)$. We also assume that f_t is concave in its both arguments. Given these notations, a stochastic account of (2) can be formalized by optimizing the *expected* value of the aggregate utility subject to *chance* constraints on the availability of the resource, i.e.,

$$\mathcal{Q}(\gamma) \triangleq \left\{ \begin{array}{ll} \max\limits_{\boldsymbol{x}} & \sum\limits_{t=1}^{T} \mathbb{E}_{\alpha_t}[f_t(x_t,\alpha_t)] \\ \text{s.t.} & \mathbb{P}\left(\sum\limits_{i=1}^{t} x_i \leq \sum\limits_{i=1}^{t} s_i\right) \geq \gamma \ , \quad \forall t \\ & \boldsymbol{x} \succeq 0 \end{array} \right. . \quad (9)$$

It can be readily verified that the function $\bar{f}_t(x) = \mathbb{E}_{\alpha_t}[f_t(x,\alpha_t)]$ is concave in x. Also, by denoting the cumulative distribution function of $\sum_{i=1}^{t} s_i$ by G_t, the stochastic constraints can be rewritten as $\sum_{i=1}^{t} x_i \leq G_t^{-1}(1-\gamma)$ By setting $\gamma_1 \triangleq G_1^{-1}(\gamma)$, defining

$$\gamma_t \ \triangleq\ G_t^{-1}(1-\gamma) - G_{t-1}^{-1}(1-\gamma) \ , \qquad \forall t \in \{2,\dots,T\} \ , \quad (10)$$

and noting that $G_t(x) \geq G_{t-1}(x)$ for all $t \in \{2,\dots,T\}$, it can be readily verified that the solution of $\mathcal{Q}(\gamma)$ can be found by solving the problem $\mathcal{P}(\boldsymbol{s})$ since $\mathcal{Q}(\gamma) = \mathcal{P}([\gamma_1,\dots,\gamma_T])$.

III. Application: Energy Harvesting Systems

In this section we discuss the application of the general approach developed in Section II to the problem of power allocation in a single-user point-to-point communication channel in which the transmitter's battery is equipped with an energy harvesting unit, gathering its power entirely from ambient sources in its surrounding environment. Hence, power, as the resource, is made available for transmission only sequentially and incrementally over time. For this purpose, consider a time-slotted transmission over a single-antenna channel in which the channel input at time $t \in \{1,\dots,T\}$ is denoted by X_t, and the output is given by

$$Y_t \ = \ h_t \cdot X_t \ + \ N_t \ , \qquad \text{for} \quad t \in \{1,\dots,T\} \ , \quad (11)$$

where h_t denotes the channel coefficient at time $t \in \{1,\dots,T\}$, and N_t account for additive white Gaussian noise distributed according to $\mathcal{N}_{\mathbb{C}}(0,1)$. In this model, x_t denotes the transmission power at time $t \in \{1,\dots,T\}$ and s_t denotes energy increments harvested at time t. Throughout the analysis we assume that the battery has infinite capacity.

A. Sum-rate Maximization in Fading Channels

By setting the utility function as $f_t(x_t) \triangleq \log(1+\alpha_t \cdot x_t)$ where $\alpha_t \triangleq |h_t|^2$, the optimal power consumption scheme over time for the purpose of maximizing the sum-rate capacity in this energy harvesting system can be obtained via solving

$$\mathcal{P}(\boldsymbol{s}) = \left\{ \begin{array}{ll} \max\limits_{\boldsymbol{x}} & \sum\limits_{t=1}^{T} \log(1+\alpha_t \cdot x_t) \\ \text{s.t.} & \sum\limits_{i=1}^{t} x_i \leq \sum\limits_{i=1}^{t} s_i \ , \quad t \in \{1,\dots,T\} \\ & \boldsymbol{x} \succeq 0 \end{array} \right. . \quad (12)$$

For solving $\mathcal{P}(\boldsymbol{s})$ we can directly apply Algorithm 1, which can identify the set of the dominant constraints recursively. In each recursion cycle, the algorithm solves a power allocation problem that is equivalent to optimizing power allocation across independent parallel channels and can be solved via the well-known water-filling algorithm. Nevertheless, when there is more structure to be leveraged, solving such power allocation problems can be avoided, and the indices of the dominant constraints, and the associated power allocation schemes can be found directly. Hence, for a general fading model, the optimal solution of $\mathcal{P}(\boldsymbol{s})$ consists in identifying the dominant constraints index by $\{u_i : i \in \{1,\dots,d\}\}$ in conjunction with applying the water-filling algorithm d times for solving $\mathcal{Q}_{u_{i-1}\to u_i}(\boldsymbol{e})$. We remark that the indices $\{u_i : i \in \{1,\dots,d\}\}$ mark the instances at which the entire energy available at those instances is exhausted, and there is no energy carry-over to the following instances. The fact that the optimal solution involves elements similar to water-filling is pointed out and discussed in details in [4], and the result in this paper complements this observation by determining the exact time intervals $\{u_{i-1}+1,\dots,u_i\}$ over which the optimal power solution is the water-filling solution of $\mathcal{Q}_{u_{i-1}\to u_i}(\boldsymbol{s})$. In the following corollary, we also address a special cases of interest, in which the fading process can be time-varying, but the rate of variations is small enough to be bounded by a measure specified by the variations of the harvested energy over time. Specifically, if the deviations of $\frac{1}{\alpha_t}$ from their average $\frac{1}{T}\sum_{t=1}^{T}\frac{1}{\alpha_t}$ are smaller than the average harvested energy $\frac{1}{T}\sum_{t=1}^{T} s_t$, i.e., when

$$\frac{1}{\min_t \alpha_t} \ \leq \ \frac{1}{T}\sum_{t=1}^{T}(s_t + \frac{1}{\alpha_t}) \ , \quad (13)$$

then the structure of Algorithm 1 simplifies significantly, as specified on the following corollary and Algorithm 3. It is noteworthy that a static channel (i.e., α_t constant) satisfies (13) and power allocation in static channels can be also determined by Algorithm 3.

Corollary 1 (Slowly Fading Channels)**.** *For a fading model that satisfies* (13)*, for the optimal solution of* $\mathcal{P}(\boldsymbol{s})$ *the time instants at which the available resources are exhausted), for* $m \in \{1,\dots,d\}$*, are given by*

$$u_m \ = \ \arg\min_{t\,\in\,\mathcal{A}_m} \ \frac{1}{t-u_{m-1}} \cdot \sum_{i=u_{m-1}+1}^{t} \left(s_i + \frac{1}{\alpha_i}\right) \ . \quad (14)$$

Proof: See Appendix J. ∎

Based on the result of Corollary 1, we provide Algorithm 3 as a simple approach to obtain the optimal power allocation $\boldsymbol{p}^*$ to the problem in (12) by admitting the vector of harvested energy $\boldsymbol{e}$ as an input. For the convenience in notation, we define the channel power gain vector as $\boldsymbol{\alpha} \triangleq [\alpha_1, \dots, \alpha_T]$. The optimality of the provided power allocation $\boldsymbol{p}^*$ is stated in Theorem 4.

Theorem 4. *For a quasi-static fading model that satisfies* (13)*, power allocation* $\boldsymbol{p}^*$ *yielded by Algorithm 3 is the optimal solution to the problem* $\mathcal{P}(\boldsymbol{e})$*.*

Proof: Follows the same line of argument as in the proof of Theorem 3. ■

Algorithm 3 - Optimal power allocation $\boldsymbol{p}^*$ over quasi-static fading channels

input $\boldsymbol{s}$
initialize $d = 0$ and $u_d = v_d = 0$
while $u_d \leq T - 1$
  $d \leftarrow d + 1$
  set $\mathcal{A}_d \triangleq \{u_{d-1} + 1, \dots, T\}$
  $u_d = \arg\min_{t \in \mathcal{A}_d} \frac{1}{t - u_{d-1}} \cdot \sum_{i=u_{d-1}+1}^{t} \left(s_i + \frac{1}{\alpha_i}\right)$
  $\beta_d \triangleq \frac{1}{u_d - u_{d-1}} \sum_{i=\bar{u}_{d-1}+1}^{u_d} (s_i + \frac{1}{\alpha_i})$
end while
for $t \in \{1, \dots, T\}$
  set $p_t^* \triangleq \beta_j - \frac{1}{\alpha_t}$, where $j = \inf\{u_i : u_i \geq t\}$
end for

B. Special Cases

In this subsection we present two special cases that specialize the sum-rate optimization of interest to the two special cases studies in Section II, namely homogeneous utility functions and utility functions with stochastic uncertainties.

Example 1 (Homogeneous Utility Functions)**.** In the context of energy harvesting, the utility functions turn out to be homogeneous when the fading process is static and the fading coefficients do not vary over time, i.e., $\alpha_1 = \cdots = \alpha_t = \alpha$, as a result of which the utility functions remain unchanged over time, i.e., $f_t(x) = \log(1 + \alpha x)$.

Example 2 (Stochastic Uncertainty)**.** When the fading coefficient α_t is random and unknown to the transmitter, the utility function $f_t(x, \alpha_t) = \log(1 + \alpha_t x)$, which is concave in both α_t and x, becomes also random and unknown. Based on the discussion in Section II-H, the expected utility function $\bar{f}_t(x) = \mathbb{E}_{\alpha_t}[f_t(\alpha_t, x)] = \mathbb{E}_{\alpha_t}[\log(1 + \alpha_t x)]$ is concave in x, and as result the stochastically-constrained power allocation problem can be solved via (9).

IV. Numerical Evaluations

In this section, we present numerical evaluations to highlight the structure and the properties of Algorithm 1 provided in Section II and compare its performance with the generic numerical algorithms for solving convex problems. Throughout the simulations we pursue two objectives. First, we aim to numerically assess the structure of the optimal solution given in lemmas 4 and 5, and assess the number of the variations of the dominant constraints, as well as the utility value with respect to different resource arrival processes. Secondly, we compare the structure of the optimal solution and the performance yielded by the optimal solution characterized with those yielded by two generic convex optimization approaches, namely the interior point (IP) method and the Matlab CVX solver.

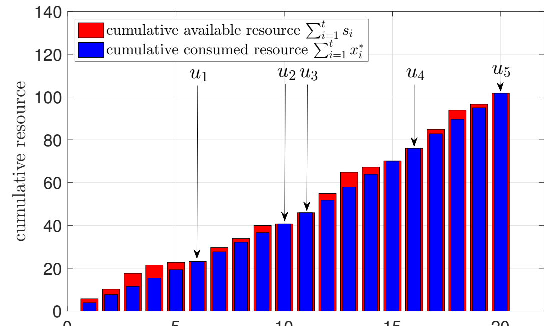

Fig. 1: Resource arrival and allocation.

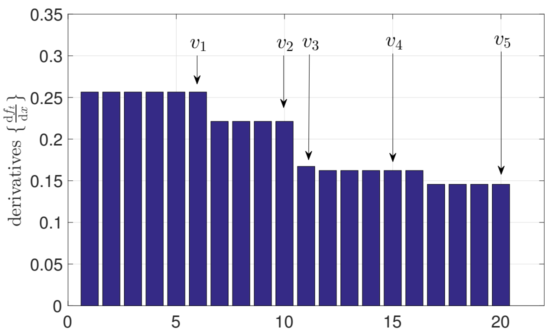

Fig. 2: Derivative measures $\{\mathrm{d}f_t/\mathrm{d}t\}$.

Throughout the simulations we focus on the slowly-fading energy harvesting application specified in Section III. In this model the utility function at time t is $f_t(x_t) = \log(1+|h_t|^2 x_t)$, where the channel coefficients h_t follow a Rayleigh fading and are distributed according to $\mathcal{N}_{\mathbb{C}}(0, 1)$. The amount of energy harvested at different time slots, i.e., $\{s_t : t \in \{1, \dots, T\}\}$, randomly varies over time, and for the purpose of implementation we consider three different models for the energy arrival process, namely $\text{Unif}(0, 2\eta)$, $\text{Exp}(\frac{1}{\eta})$, and $\text{Poisson}(\eta)$, where η denotes the average resource arrival rate.

A. Constraint Groups

The key structure of the solution to $\mathcal{P}(\boldsymbol{s})$ is that it can be reduced by partitioning $\{x^*, \dots, x_T^*\}$ into d disjoint sets, where the values in each set or related (their respective functions have same derivatives) and can be computed independently of each other. The set $\{u_i :\in \{1, \dots, d\}\}$ specifies the time instances at which all the available resources are exhausted. In order to demonstrate this numerically, we set the time horizon to $T = 10$, and generate one realization of the harvested energy vector $\boldsymbol{s}$. We solve the problem $\mathcal{P}(\boldsymbol{s})$ for this realization, and in Fig. 1 plot the variations of $\sum_{i=1}^{t} x_i^*$ over time to asses the optimal properties stated in lemmas 4 and 5. For this evaluation we consider $\text{Unif}(0, 2\eta)$ as the energy arrival process, with $\eta = 5$. The light (red) bar at time t shows the level of available resources at time

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2846657, IEEE Transactions on Communications

t, and the dark (blue) bar depicts the amount of available resources to be consumed at time t. It is observed that at certain time instants the two bars have exactly same heights indicating the available resources are exhausted. These time instants occur at $\{u_1, u_2, u_3, u_4, u_5\} = \{6, 10, 11, 16, 20\}$. We have also evaluated the variations of $\sum_{i=1}^{T} x_i$ for the solution $\boldsymbol{x}$ provided by the CVX solver as well as the IP method, where we have observed that the solutions match with the optimal solution with high accuracy (albeit with higher complexity analyzed in Section IV-B). Furthermore, for the same system realization used for the evaluations in Fig. 1, the variations of the derivatives of the utility functions, i.e., $\{\mathrm{d}f_t/\mathrm{d}t : t \in \{1, \ldots, T\}\}$ are depicted in Fig. 2. It shows two main properties associated with the derivative measures $\{v_i : i \in \{1, \ldots, d\}\}$. First, the solutions in the range $\{u_i + 1, \ldots, u_{i+1}\}$ have the same derivatives, and secondly, the metrics $\{v_i : i \in \{1, \ldots, d\}\}$ are strictly decreasing over time. These values are marked in Fig. 2.

B. Computational Complexity

An important practical advantage of $\{x^*, \ldots, x_T^*\}$ is that the elements in each partition are computed independently of each other. This leads to significant reduction in the computational complexity since instead of solving a T-dimensional problem we face solving a number of problems with dimensions much smaller than T. To compare the complexity of Algorithm 1 with those of CVX solver and IP method, we consider the setting of Section IV-A, and provide Table I, which demonstrates the processing times of the algorithm for different values of T and three energy arrival processes (uniform, exponential, and Poisson). This table shows that the algorithm, in designing which the structure of the problem is taken into account, is considerably faster than the CVX solver and the IP method.

C. Number of Partitions

Figure 3 depict the variations of the number of partitions d with respect to different rates of energy arrival η under three different processes, and for different problem dimensions $T = 10, 50, 100$. It is observed that for a given T, d remains rather insensitive to energy arrival process, and the the average arrival rate η. The underlying reason is that the expected values of $\{u_i : i \in \{1, \ldots, d\}\}$ do not depend on the exact distribution of the resource arrival process, and rather they depend on the relative changes of these distributions over time. When the distributions are identical over time, as is the case in this setting, their exact choices do not have a significant impact. As a result, varying the energy arrival rate η does not affect the average values of $\{u_i : i \in \{1, \ldots, d\}\}$, and subsequently, the expected value of d.

Increasing T has two opposing effects on d. On the one hand it increases the spacing between the consecutive time indices in $\{u_i : i \in \{1, \ldots, d\}\}$. The reason underlying this is that according to Algorithm 1, these time indices are determined by selecting the maximum derivative measure, $q^{d,t}$, for every $t \in \mathcal{A}_i$ and $i \in \{1, \ldots, d\}$. Thus, by increasing the time horizon T, the maximum values of the derivative measures, $q^{d,t}$, appear, on average, at later time instances. This effect tends to decrease d. On the other hand, a larger T is expected to lead to a larger number of constraint groups. Figure 3 shows the combined effect of these two opposing trends is in favor of an increasing trend for d.

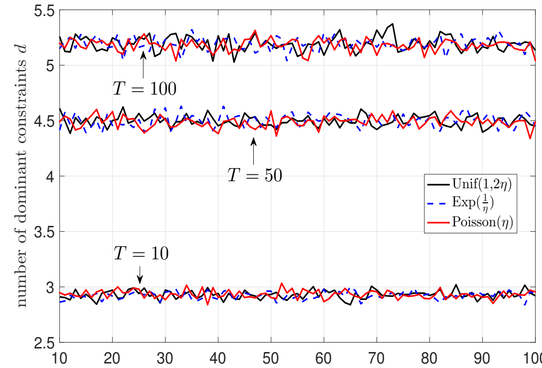

Fig. 3: Average d versus η.

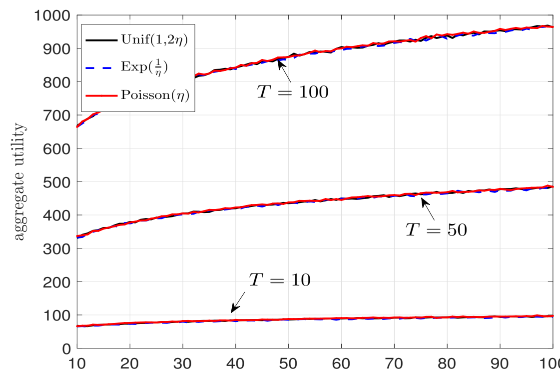

Fig. 4: Aggregate utility versus η.

D. Aggregate Utility

Finally, we evaluate the maximum sum-rate as the solution to (12) and in Fig. 4 depict the variations of the sum-rate versus the resource arrival rate η for different values of T. We use a setting similar to those described in the previous subsection. As expected, increasing the transmission horizon leads to increased throughput. Specifically, it can be readily shown that the sum-rate scales linearly with the time horizon T. Also it is observed that increasing η leads to increase in the sum-rate, which is also expected as higher η, on average, indicates higher amount of resources, and leads to a larger utility. Finally, we remark that the CVX method and the IP method achieve the similar solutions with a range of bounded error, albeit at the expense of higher computation complexities.

V. Conclusion

In this paper, we have analysed and solved the problem of optimal resource allocation over time, with the resource becomes available sequentially and incrementally over time. Such problems in their general forms subsume a wide range of conventional resource allocation problems in communication systems (e.g., resource allocation over parallel channels), and have direct application in certain applications in which

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2846657, IEEE Transactions on Communications

Table I: Computational time in seconds

T \ s_t	Unif (0,10)			Exp $(\frac{1}{30})$			Poisson (50)		
	Alg. 1	IP	CVX	Alg. 1	IP	CVX	Alg. 1	IP	CVX
10	29×10^{-6}	0.11	1.84	51×10^{-6}	0.15	1.75	0.58×10^{-3}	0.12	1.95
100	31×10^{-6}	0.87	11.25	55×10^{-6}	0.87	10.55	1.15×10^{-3}	0.87	11.51
1000	42×10^{-6}	2.09	443.70	77×10^{-6}	2.08	376.23	1.91×10^{-3}	2.09	369.72

resources are accessible sequentially (e.g., energy harvesting and quality-constrained systems). Such problems are well-investigated in their special forms for the utility functions, and in this paper, we have treated the problem in its general form. First, we have established certain key properties of the optimal solution, based on which we have proposed an algorithm for obtaining the solution. A key observation has been that there exist time instants at which the available resource is entirely utilized, and characterizing the optimal solution depends on identifying those instants. The proposed algorithm provides closed-form characterization of these instants. Furthermore, we have shown that the proposed algorithm can be applied to a stochastic version of the resource allocation problem, in which only the statistical properties of the resource arrivals are known. Moreover, we have applied the obtained optimal solution to identify a closed-form optimal power allocation policy under energy harvesting constraints over the single-user fading channels. Finally, we have provided numerical evaluations to depict the key properties of the optimal resource allocation policy and to also compare the performance with those of generic convex optimization algorithms.

Appendix A
Proof of Lemma 1

By recalling the definitions of $\boldsymbol{w}^{d,t}$ (line 6 of Algorithm 1) and $\mathcal{Q}_{m\rightarrow n}(\boldsymbol{x}^*)$ in (5), it can be readily verified that vector $\boldsymbol{w}^{m,t}$, where $m \in \{1,\dots,d\}$ and $t \in \mathcal{A}_m$, is a solution to $\mathcal{Q}_{u_{m-1}\rightarrow t}(\boldsymbol{x}^*)$. Corresponding to each pair of m and t where $m \in \{1,\dots,d\}$ and $t \in \mathcal{A}_m \triangleq \{u_{m-1}+1,\dots,T\}$, we define the $\lambda_{m,t}$ and $\eta^{\ell}_{m,t}$ as positive real numbers Lagrange multipliers corresponding to the equality and inequality constraints in (5). Enforcing the Karush-Kuhn-Tucker (KKT) condition yields that for all $i \in \{u_{m-1}+1,\dots,t\}$ we have

$$\frac{\mathrm{d}f_i}{\mathrm{d}x}(w_i^{m,t}) \;=\; \lambda_{m,t} + \eta^i_{m,t}\ , \tag{15}$$

where $\eta^i_{m,t} > 0$ if $w_i^{m,t} = 0$, and otherwise $\eta^i_{m,t} = 0$.

Appendix B
Proof of Lemma 2

Recall that in Algorithm 1 we have defined $\tilde{x}_t \triangleq z^i_t$ where $i = \inf\{j \ : \ u_j \geq t\}$. Also, define

$$h_t(\lambda) \;\triangleq\; \left[\left(\frac{\mathrm{d}f_t}{\mathrm{d}\lambda}\right)^{-1}(\lambda)\right]^+ , \qquad \forall t \in \{1,\dots,T\}\ , \tag{16}$$

where $[x]^+$ denotes $\max\{x,0\}$. Since functions $\{f_t : t \in \{i,\dots,T\}\}$ are strictly concave, the inverse of their derivatives, i.e., $\{h_t : t \in \{1,\dots,T\}\}$ are non-increasing functions. Hence, for a given combination of m, t, and ℓ, such that $m \in \{0,\dots,d-1\}$, $t \in \mathcal{D}_m = \{u_{m-1}+1,\dots,u_m\}$, and $\ell \in \{u_{m-1}+1,\dots,t\}$, we have

$$\begin{aligned}
\tilde{x}_\ell &= z^m_\ell && \text{(definition of } \tilde{x}_t \text{ in Algorithm 1)} && (17)\\
&= w^{m,u_m}_\ell && \text{(definition of } \boldsymbol{z}^d \text{ in Algorithm 1)} && (18)\\
&= h_\ell(\lambda_{m,u_m}) && \text{(according to 16)} && (19)\\
&= h_\ell(q^{m,u_m}) && \text{(definition of } q^{d,t} \text{ in Algorithm 1)} && (20)\\
&\leq h_\ell(q^{m,t}) && (h_t \text{ is non-decreasing}) && (21)\\
&= h_\ell(\lambda_{m,t}) && \text{(definition of } q^{d,t} \text{ in Algorithm 1)} && (22)\\
&= w^{m,t}_\ell\ . && \text{(according to (16) and Lemma 1)} && (23)
\end{aligned}$$

Consequently, corresponding to each $m \in \{1,\dots,d\}$, for all $t \in \mathcal{D}_m$ we have

$$\sum_{\ell=u_{m-1}+1}^{t} \tilde{x}_\ell \overset{(17)-(23)}{\leq} \sum_{\ell=u_{m-1}+1}^{t} w^{m,t}_\ell = \sum_{t=u_{m-1}+1}^{t} x^*_\ell\ , \tag{24}$$

where the last equality in (24) holds since for $t \in \{u_{m-1}+1,\dots,u_m\}$, $\boldsymbol{w}^{m,t}$ is the solution of $\mathcal{Q}_{u_{m-1}\rightarrow t}(\boldsymbol{x}^*)$, as specified in Algorithm 1. By invoking the inequalities in (24) we find that for each $m \in \{1,\dots,d\}$, and for all $t \in \mathcal{D}_m$ we have

$$\begin{aligned}
\sum_{\ell=1}^{t} \tilde{x}_\ell &= \sum_{i=1}^{m-1}\sum_{\ell\in\mathcal{D}_i} \tilde{x}_\ell + \sum_{\ell=u_{m-1}+1}^{t} \tilde{x}_\ell && (25)\\
&= \sum_{i=1}^{m-1}\sum_{\ell\in\mathcal{D}_i} z^i_\ell + \sum_{\ell=u_{m-1}+1}^{t} \tilde{x}_\ell && (26)\\
&= \sum_{i=1}^{m-1}\sum_{\ell\in\mathcal{D}_i} w^{i,u_i}_\ell + \sum_{\ell=u_{m-1}+1}^{t} x^*_\ell && (27)\\
&= \sum_{i=1}^{m-1}\sum_{\ell\in\mathcal{D}_i} \tilde{x}_\ell + \sum_{\ell=u_{m-1}+1}^{t} x^*_\ell && (28)\\
&\leq \sum_{i=1}^{m-1}\sum_{\ell\in\mathcal{D}_i} \tilde{x}_\ell + \sum_{\ell=u_{m-1}+1}^{t} \tilde{x}_\ell && (29)\\
&= \sum_{\ell=1}^{t} \tilde{x}_\ell \;\leq\; \sum_{\ell=1}^{t} s_\ell\ , && (30)
\end{aligned}$$

where (27) holds by noting the definition of $\boldsymbol{w}^{d,t}$ in Algorithm 1; the transition from (27) to (28) holds since for each $i \in \{1,\dots,d\}$, it can be readily verified that $\boldsymbol{w}^{i,u_i}$ is the solution of $\mathcal{Q}_{u_{i-1}\rightarrow u_i}(\boldsymbol{x}^*)$, and subsequently, $\sum_{\ell\in\mathcal{D}_i} w^{i,u_i}_\ell = \sum_{\ell\in\mathcal{D}_i} x^*_\ell$; the transition from (28) to (29) holds by invoking the inequalities in (24); and (30) follows by recalling that $\tilde{\boldsymbol{x}}$ is an optimal solution of $\mathcal{P}(\boldsymbol{s})$, which as a result, should satisfy all constraints of $\mathcal{P}(\boldsymbol{s})$.

Appendix C
Proof of Lemma 3

Based on the definition of $\boldsymbol{z}^d$, for each $m \in \{1,\dots,d\}$, it can be readily verified that $\boldsymbol{z}^m$ is the solution of $\mathcal{Q}_{u_{m-1}\to u_m}(\boldsymbol{x}^*)$. Based on the definition of $\mathcal{Q}_{m\to n}(\boldsymbol{s})$ in (5), this observation indicates that for any vector $\boldsymbol{x} \succeq 0$ that satisfies $\sum_{i=u_{m-1}+1}^{u_m} x_i = \sum_{i=u_{m-1}+1}^{u_m} x_i^*$ we have

$$\sum_{t=u_{m-1}+1}^{u_m} f(z_t^m) \geq \sum_{t=u_{m-1}+1}^{u_m} f(x_t) \ . \tag{31}$$

Clearly $\boldsymbol{x}^*$ satisfies the above constraint, based on which we obtain

$$\sum_{t=u_{m-1}+1}^{u_m} f(z_t^m) \geq \sum_{t=u_{m-1}+1}^{u_m} f(x_t^*) \ . \tag{32}$$

By recalling the definition of $\tilde{\boldsymbol{x}}$ in Appendix B, and by invoking the inequality in (32) we obtain

$$F(\tilde{\boldsymbol{x}}) = \sum_{t=1}^{T} f_t(\tilde{x}_t) \overset{(6)}{=} \sum_{m=1}^{d} \sum_{t=u_{m-1}+1}^{u_m} f_t(\tilde{x}_t) \tag{33}$$

$$= \sum_{m=1}^{d} \sum_{t=u_{m-1}+1}^{u_m} f_t(z_t^m) \tag{34}$$

$$\overset{(32)}{\geq} \sum_{m=1}^{d} \sum_{t=u_{m-1}+1}^{u_m} f_t(x_t^*) = F(\boldsymbol{x}^*) \ . \tag{35}$$

Appendix D
Proof of Lemma 4

We start by establishing $\lambda_{m,u_m} \geq \lambda_{m,u_{m+1}}$ and $\lambda_{m,u_{m+1}} \geq \lambda_{m+1,u_{m+1}}$.

1) $\underline{\lambda_{m,u_m} \geq \lambda_{m,u_{m+1}}}$: By noting the definitions of u_d and $q^{d,t}$ in Algorithm 1, and using Lemma 1

$$\lambda_{m,u_m} = q^{m,u_m} = \max_{t\in\mathcal{A}_m} q^{m,t} \geq q^{m,u_{m+1}} = \lambda_{m,u_{m+1}} \ . \tag{36}$$

2) $\underline{\lambda_{m,u_{m+1}} \geq \lambda_{m+1,u_{m+1}}}$: By recalling the definition of h_t in (16) and that $\boldsymbol{w}^{d,t}$ is the solution to $\mathcal{Q}_{u_{d-1}\to t}(\boldsymbol{x}^*)$ we have

$$\sum_{t=u_{m-1}+1}^{u_m} h_t(\lambda_{m,u_m}) = \sum_{t=u_{m-1}+1}^{u_m} x_t^* \ , \tag{37}$$

$$\sum_{t=u_m+1}^{u_{m+1}} h_t(\lambda_{m+1,u_{m+1}}) = \sum_{t=u_m+1}^{u_{m+1}} x_t^* \ , \tag{38}$$

$$\text{and} \quad \sum_{t=u_{m-1}+1}^{u_{m+1}} h_t(\lambda_{m,u_{m+1}}) = \sum_{t=u_{m-1}+1}^{u_{m+1}} x_t^* \ . \tag{39}$$

Now, by contradiction assume that $\lambda_{m,u_{m+1}} < \lambda_{m+1,u_{m+1}}$. Based on (39) we find

$$\sum_{t=u_{m-1}+1}^{u_{m+1}} h_t(\lambda_{m,u_{m+1}}) = \sum_{t=u_{m-1}+1}^{u_m} \underbrace{h_t(\lambda_{m,u_{m+1}})}_{\geq h_t(\lambda_{m,u_m})} \tag{40}$$

$$+ \sum_{t=u_m+1}^{u_{m+1}} \underbrace{h_t(\lambda_{m,u_{m+1}})}_{> h_t(\lambda_{m+1,u_{m+1}})} \tag{41}$$

$$> \sum_{t=u_{m-1}+1}^{u_m} h_t(\lambda_{m,u_m}) + \sum_{t=u_m+1}^{u_{m+1}} h_t(\lambda_{m+1,u_{m+1}}) \tag{42}$$

$$\overset{(37)-(38)}{=} \sum_{t=u_{m-1}+1}^{u_m} x_t^* + \sum_{t=u_m+1}^{u_{m+1}} x_t^* = \sum_{t=u_{m-1}+1}^{u_{m+1}} x_t^* \ , \tag{43}$$

which contradicts (39). Hence, given the inequalities in (36) we have

$$\begin{aligned} v_m &= \max_{t\in\mathcal{A}_m} q^{m,t} && (\text{where} \quad \mathcal{A}_m = \{u_{m-1}+1,\dots,T\}) \\ &= \lambda_{m,u_m} && (\text{definition of } u_d \text{ and Lemma 1}) \\ &\geq \lambda_{m+1,u_{m+1}} \\ &= \max_{t\in\mathcal{A}_{m+1}} q^{m+1,t} && (\text{where} \quad \mathcal{A}_{m+1} = \{u_m+1,\dots,T\}) \\ &= v_{m+1} \ . \end{aligned} \tag{44}$$

Appendix E
Proof of Lemma 5

We provide the proof via backward induction.

Basis: The statement is true for $m = d$. We start by proving that the statement holds for $m = d$. To this end note that by construction, $u_d = T$. Hence, we need to show that $\sum_{t=1}^T x_t^* = \sum_{t=1}^T s_t$. Assume that, by contradiction, we have $\sum_{t=1}^T x_t^* < \sum_{t=1}^T s_t$., in which case we generate another vector $\bar{\boldsymbol{x}}$ according to

$$\bar{x}_t = \begin{cases} x_t^* & \text{for} \quad t \in \{1,\dots,T-1\} \\ x_t^* + [\sum_{t=1}^T s_t - \sum_{t=1}^T x_t^*] & \text{for} \quad t = T \end{cases} . \tag{45}$$

It can be readily verified that $bar\boldsymbol{x}$ satisfies the constraints of $\mathcal{P}(\boldsymbol{s})$, and since f_T is an monotonic $F(\bar{\boldsymbol{x}}) > F(\boldsymbol{x}^*)$, which contradicts the optimality of $\boldsymbol{x}^*$. Hence, for $m = d$ the statement is valid.

Induction hypothesis: The statement is true for some $(m+1) \in \{2,\dots,d\}$.

Induction step: The statement is true for m. By contradiction, assume that $\sum_{t=1}^{u_m} x_t^* < \sum_{t=1}^{u_m} s_t$. Next, select indices $i \in \{u_{m-1}+1,\dots,u_m\}$ and $j \in \{u_m+1,\dots,u_{m+1}\}$ such that x_i and x_j are strictly positive, and select a positive constant δ in the range $[0, \sum_{t=1}^i s_t - \sum_{t=1}^i x_t^*]$, and generate $\bar{\boldsymbol{x}}$ according to

$$\bar{x}_t = \begin{cases} x_t^* & \text{for} \quad t \in \{1,\dots,T\} \setminus \{i,j\} \\ x_t^* + \delta & \text{for} \quad t = i \\ x_t^* - \delta & \text{for} \quad t = j \end{cases} \tag{46}$$

It can be easily verified that $\bar{\boldsymbol{x}}$ satisfies all the constraints of $\mathcal{P}(\boldsymbol{s})$. Next, we show that for structure of $\bar{\boldsymbol{x}}$ characterized in (46) there exist $\delta \in (0, \sum_{t=1}^i s_t - \sum_{t=1}^i x_t^*]$ corresponding to which $F(\bar{\boldsymbol{x}}) > F(\boldsymbol{x}^*)$. This conclusion violates the optimality of $\boldsymbol{x}^*$, and proves the induction statement by contradiction. For this purpose, we define the function $g : \mathbb{R} \to \mathbb{R}$ as $g(\delta) \triangleq F(\bar{\boldsymbol{x}})$. Clearly, $g(0) = F(\bar{\boldsymbol{x}})|_{\delta=0} = F(\boldsymbol{x}^*)$. Hence, from the expansion of F we have

$$\frac{\mathrm{d}g(\delta)}{\mathrm{d}\delta} = \sum_{t=1}^{T} \frac{\mathrm{d}f_t(\bar{x}_t)}{\mathrm{d}\delta} = \frac{\mathrm{d}f_i(x_i^*+\delta)}{\mathrm{d}\delta} + \frac{\mathrm{d}f_j(x_j^*-\delta)}{\mathrm{d}\delta} \ . \tag{47}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCOMM.2018.2846657, IEEE Transactions on Communications

By recalling the definitions of $q^{t,d}$ and v_d, noting the range of i and j, and following the same line of derivations as in the proof of Lemma 2 (specifically (20)) we can readily show that

$$\frac{\mathrm{d}f_i(x_i^*+\delta)}{\mathrm{d}\delta}\Big|_{\delta=0} \ = \ v_m \text{ and } \frac{\mathrm{d}f_j(x_j^*-\delta)}{\mathrm{d}\delta}\Big|_{\delta=0} \ = \ v_{m+1} \ . \tag{48}$$

From (47)-(48) we find that $\mathrm{d}g(\delta)/\mathrm{d}\delta|_{\delta=0} = v_m - v_{m+1}$, which in conjunction with Lemma 4 establishes that $\mathrm{d}g(\delta)/\mathrm{d}\delta|_{\delta=0} > 0$, Hence, for a sufficiently small $\delta > 0$ we have $F(\bar{\boldsymbol{x}}) = g(\delta) > g(0) = F(\boldsymbol{x}^*)$ which contradicts the optimality of $\boldsymbol{x}^*$, and as a result $\sum_{t=1}^{u_m} x_t^* = \sum_{t=1}^{u_m} s_t$

Appendix F
Proof of Lemma 6

We demonstrate this statement by induction.

Basis: $u_1 = \bar{u}_1$. For $m = 1$, Algorithm 1 compute T vectors $\{\boldsymbol{w}^{1,1},\dots,\boldsymbol{w}^{1,T}\}$ and the associated gradient terms $\{q^{1,1},\dots,q^{1,T}\}$ by admitting an input $\boldsymbol{x}^*$. Similarly, it generates the corresponding sets by admitting the vector $\boldsymbol{s}$, which we denote by $\{\bar{\boldsymbol{w}}^{1,1},\dots,\bar{\boldsymbol{w}}^{1,T}\}$ and $\{\bar{q}^{1,1},\dots,\bar{q}^{1,T}\}$, to highlight their potential discrepancies. By noting that

$$u_1 \ = \ \arg\max\left\{q^{1,t} : t\in\{1,\dots,T\}\right\} \tag{49}$$

$$\text{and} \quad \bar{u}_1 \ = \ \arg\max\left\{\bar{q}^{1,t} : t\in\{1,\dots,T\}\right\} \ , \tag{50}$$

in order to show that $u_1 = \bar{u}_1$, we provide the following two properties.

Property 1: $q^{1,t} = \bar{q}^{1,t}$ for $t = u_1$:

As shown in Lemma 5, constraint u_1 for the optimal solution $\boldsymbol{x}^*$ holds with equality, i.e.,

$$\sum_{i=1}^{u_1} x_i^* \ = \ \sum_{i=1}^{u_1} s_i \ . \tag{51}$$

This implies that $\boldsymbol{w}^{1,u_1}$ is the solution to $\mathcal{Q}_{0\to u_1}(\boldsymbol{x}^*) = \mathcal{Q}_{0\to u_1}(\boldsymbol{s})$. Similarly, from the construction of Algorithm 1 by admitting $\boldsymbol{s}$ we have that $\bar{\boldsymbol{w}}^{1,u_1}$ is the solution to $\mathcal{Q}_{0\to u_1}(\boldsymbol{s})$. As a result, $\boldsymbol{w}^{1,u_1} = \bar{\boldsymbol{w}}^{1,u_1}$, which in turn indicates that

$$q^{1,u_1} = \min\left\{\frac{\mathrm{d}f_i}{\mathrm{d}x}(w_i^{1,u_1}) : i\in\{1,\dots,u_1\}\right\} \tag{52}$$

$$= \ \min\left\{\frac{\mathrm{d}f_i}{\mathrm{d}x}(\bar{w}_i^{1,u_1}) : i\in\{1,\dots,u_1\}\right\} \tag{53}$$

$$= \bar{q}^{1,u_1} \ . \tag{54}$$

Property 2: $q^{1,t} \geq \bar{q}^{1,t}$ for $t\in\{1,\dots,T\}\setminus\{u_1\}$:

For all $t\in\{1,\dots,T\}\setminus\{u_1\}$, the constraints do not hold necessarily with equality, and are valid only in their general form

$$\sum_{i=1}^{t} x_i^* \ \leq \ \sum_{i=1}^{t} s_i \ . \tag{55}$$

Since $\boldsymbol{w}^{1,t}$ is the solution to $\mathcal{Q}_{0\to t}(\boldsymbol{x}^*)$ and $\bar{\boldsymbol{w}}^{1,t}$ is the solution to $\mathcal{Q}_{0\to t}(\boldsymbol{s})$, by recalling the definition of h_t in (16) we find

$$\sum_{i=1}^{t} h_i(\lambda_{1,t}) = \sum_{i=1}^{t} x_i^* \text{ and } \sum_{i=1}^{t} h_i(\bar{\lambda}_{1,t}) = \sum_{i=1}^{t} s_i \ , \tag{56}$$

where we have defined the Lagrangian multipliers $\{\bar{\lambda}_{1,t}\}$ as the counterparts of $\{\lambda_{1,t}\}$ corresponding to $\bar{\boldsymbol{w}}^{1,\bar{u}_1}$. By noting that functions h_t are decreasing, and invoking (55) we consequently have $\lambda_{1,t} \geq \bar{\lambda}_{1,t}$, and subsequently $q^{1,t} \geq \bar{q}^{1,t}$. Property 2 in conjunction with (49) demonstrates that

$$q^{1,u_1} \ = \ \max\left\{q^{1,t} : t\in\{1,\dots,T\}\right\} \tag{57}$$

$$\geq \ \max\left\{\bar{q}^{1,t} : t\in\{1,\dots,T\}\right\} \ . \tag{58}$$

Additionally, from Property 1 we have $q^{1,u_1} = \bar{q}^{1,u_1} \in \{\bar{q}^{1,t} : t\in\{1,\dots,T\}\}$, which combined with (57) establishes the desired property that $q^{1,u_1} = \max\{\bar{q}^{1,t} : t\in\{1,\dots,T\}\}$. and subsequently, $\bar{u}_1 = \arg\max\{\bar{q}^{1,t} : t\in\{1,\dots,T\}\} = u_1$.

Induction hypothesis: $u_{m-1} = \bar{u}_{m-1}$ for $(m-1) \in \{1,\dots,d-1\}$.

Induction step: $u_m = \bar{u}_m$ for $m\in\{2,\dots,d\}$. We assume that $u_{m-1} = \bar{u}_{m-1}$, and the proof follows the same line of arguments as in the proof for the case $m = 1$ with the necessary modifications, i.e., replacing indices u_0, $\bar{u}_0$, u_1, and u_1 by u_{m-1}, $\bar{u}_{m-1}$, u_m, and u_m, respectively. Besides these changes, all other steps are exactly the same, which are omitted for brevity.

Appendix G
Proof of Theorem 2

Based on the construction of Algorithm 1 by admitting the input $\boldsymbol{x}^*$, $\tilde{\boldsymbol{x}}$ can be found based on the vectors $\{\boldsymbol{z}^i : i\in\{1,\dots,d\}$ (described in lines 14-18), which in turn can be computed from the vectors $\{\boldsymbol{w}^{i,\bar{u}_i} : i\in\{1,\dots,d\}$ (described in line 11). Let the vectors $\{\bar{\boldsymbol{z}}^i : i\in\{1,\dots,\bar{d}\}$ and $\{\boldsymbol{w}^{i,\bar{u}_i} : i\in\{1,\dots,d\}$ denote the corresponding vectors generated by admitting the vector $\boldsymbol{s}$ as an input. Using the result of Lemma 6 we show that the sets $\{\bar{\boldsymbol{z}}^i : i\in\{1,\dots,\bar{d}\}$ and $\{\bar{\boldsymbol{w}}^{i,\bar{u}_i} : i\in\{1,\dots,\bar{d}\}$ generated by Algorithm 1, admitting $\boldsymbol{x}^*$ and $\boldsymbol{s}$ are identical. Note that from Lemma 6 we have $u_m = \bar{u}_m$ for all $m\in\{1,\dots,d\}$. As a result $\bar{\boldsymbol{w}}^{m,\bar{u}_m}$ is the solution to $\mathcal{Q}_{\bar{u}_{m-1}\to\bar{u}_m}(\boldsymbol{s}) = \mathcal{Q}_{u_{m-1}\to u_m}(\boldsymbol{s}) = \mathcal{Q}_{u_{m-1}\to u_m}(\boldsymbol{x}^*)$ where the solution of the last problem is $\boldsymbol{w}^{m,u_m}$. Hence, $\bar{\boldsymbol{w}}^{m,\bar{u}_m} = \boldsymbol{w}^{m,u_m}$ for all $m\in\{1,\dots,d\}$, which also implies that $\bar{\boldsymbol{z}}^m = \boldsymbol{z}^m$ for all $m\in\{1,\dots,d\}$. Hence, the vector $\tilde{\boldsymbol{x}}$ computed based on $\{\bar{\boldsymbol{z}}^i : i\in\{1,\dots,d\}$ is the same $\boldsymbol{x}^*$ computed based on $\{\boldsymbol{z}^i : i\in\{1,\dots,d\}$. Finally, the solution $\tilde{\boldsymbol{x}}$ computed by using $\boldsymbol{x}^*$ lends its optimality to that computed based on $\boldsymbol{s}$.

Appendix H
Proof of Lemma 7

By proving Lemma 6 and Theorem 2 stated in Subsection II-F, we showed that the solution $\boldsymbol{x}^*$ to problem (2) yielded computed by using $\boldsymbol{s}$ through Algorithm 1 is optimal. In this lemma, we characterize the set of constraint indices $\{u_i : i\in\{1,\dots,T\}\}$ (defined in line 9 in Algorithm 1) as well as the set of derivative measures $\{v_i : i\in\{1,\dots,T\}\}$ (defined in line 10 in Algorithm 1) that are specialized to the setting of interest, i.e., homogeneous utility functions. The result of this lemma constitutes an intermediate step that will be used in proving the optimality $\boldsymbol{x}^*$ determined by Algorithm 3 in Subsection II-G. For this purpose, note that according to Lemma 1, and based on the symmetry involved, all the terms $\{w_i^{m,t} : i \in \mathcal{A}_m\}$ must be non-zero and satisfy

$$\frac{\mathrm{d}f}{\mathrm{d}w}(w_i^{m,t}) = \lambda_{m,t} \ , \quad \forall i \in \{u_{m-1}+1, \ldots, t\} \ , \tag{59}$$

which in turn indicates that the terms $\{w_i^{m,t} : i \in \mathcal{A}_m\}$ are equal. Hence,

$$w_i^{m,t} = \left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)^{-1} (\lambda_{m,t}) \ . \tag{60}$$

Since $\boldsymbol{w}^{m,t}$ is the solution to $\mathcal{Q}_{u_{m-1}\to t}(\boldsymbol{e})$, we obtain

$$\sum_{i=u_{m-1}+1}^{t} s_i = \sum_{i=u_{m-1}+1}^{t} w_i^{m,t} \overset{(60)}{=} [t - \bar{u}_{m-1}] \left(\frac{\mathrm{d}f}{\mathrm{d}x}\right)^{-1} (\lambda_{m,t}) \ , \tag{61}$$

or equivalently,

$$\lambda_{m,t} = \frac{\mathrm{d}f}{\mathrm{d}x}\left(\frac{1}{t-u_{m-1}} \sum_{i=u_{m-1}+1}^{t} s_i\right) \ . \tag{62}$$

Based on the definition of the sets $\mathcal{A}_m$ for u_m defined in line 4 in Algorithm 1, we obtain

$$u_m = \arg\max_{t\in\mathcal{A}_m} \ q^{m,t} \tag{63}$$

$$= \arg\max_{t\in\mathcal{A}_m} \ \min\left\{\frac{\mathrm{d}f_i}{\mathrm{d}x}(\bar{w}_i^{d,t}) : i \in \{u_{d-1}+1, \ldots, t\}\right\} \tag{64}$$

$$= \arg\max_{t\in\mathcal{A}_m} \ \lambda_{m,t} \tag{65}$$

$$\overset{(62)}{=} \arg\max_{t\in\mathcal{A}_m} \ \frac{\mathrm{d}f}{\mathrm{d}x}\left(\frac{1}{t-u_{m-1}} \sum_{i=u_{m-1}+1}^{t} s_i\right) \tag{66}$$

$$= \arg\min_{t\in\mathcal{A}_m} \ \frac{1}{t-u_{m-1}} \sum_{i=u_{m-1}+1}^{t} s_i \ . \tag{67}$$

APPENDIX I
PROOF OF THEOREM 3

According to Algorithm 2, for every $m \in \{1, \ldots, T\}$, and for all $t \in \{u_{m-1}+1, \ldots, u_m\}$, the optimal solution is $x_t^* = z_t^m = w_t^{m,u_m}$ By recalling (60) and (62), for all $t \in \mathcal{D}_m$ we have

$$w_t^{m,u_m} = \beta_m = \frac{1}{u_m - u_{m-1}} \sum_{i=u_{m-1}+1}^{u_m} s_i \ . \tag{68}$$

APPENDIX J
PROOF OF COROLLARY 1

First note that under the condition specified in (13), by using the results of Lemma 1, it can be readily verified that all optimal solutions $\{w_i^{t,m} : i \in \{u_{m-1}+1, \ldots, t\}\}$ for all $m \in \{1, \ldots, d\}$ and all $t \in \mathcal{A}_m$ are strictly positive, and satisfy

$$\frac{\mathrm{d}f_i}{\mathrm{d}w}(w_i^{m,t}) = \frac{\alpha_i}{1 + \alpha_i \cdot w_i^{m,t}} = \lambda_{m,t} \ . \tag{69}$$

Also, since $\boldsymbol{w}^{m,t}$ is a solution to $\mathcal{Q}_{u_{m-1}\to t}(\boldsymbol{s})$, it satisfies

$$\sum_{i=u_{m-1}+1}^{t} w_i^{m,t} = \sum_{i=u_{m-1}+1}^{t} s_i \, . \tag{70}$$

Based on (60) and (70) it can be readily verified that

$$\lambda_{m,t} = \left[\frac{1}{t-u_{m-1}} \sum_{i=u_{m-1}+1}^{t} (s_i + \frac{1}{\alpha_i})\right]^{-1} \ . \tag{71}$$

Hence,

$$u_m = \arg\max_{t\in\mathcal{A}_m} \ q^{m,t} = \arg\max_{t\in\mathcal{A}_m} \ \lambda_{m,t} \tag{72}$$

$$\overset{(71)}{=} \arg\min_{t\in\mathcal{A}_m} \ \frac{1}{t-u_{m-1}} \cdot \sum_{i=u_{m-1}+1}^{t} \left(s_i + \frac{1}{\alpha_i}\right) \ . \tag{73}$$

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley, 2006.

[2] E. Telatar, “Capacity of multi-antenna gaussian channels,” *Transactions on Emerging Telecommunications Technologies*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[3] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, “Resource management for fading wireless channels with energy harvesting nodes,” in *Proc. IEEE International Conference on Computer Communications*, Shanghai, China, Apr. 2011, pp. 456–460.

[4] ——, “Transmission with energy harvesting nodes in fading wireless channels: Optimal policies,” *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1732–1743, Aug. 2011.

[5] A. Sinha, “Optimal power allocation for a renewable energy source,” in *National Conference on Communications*, Kharagpur, India, Feb. 2012.

[6] K. Tutuncuoglu and A. Yener, “Short-term throughput maximization for battery limited energy harvesting nodes,” in *Proc. IEEE International Conference on Communications*, Kyoto, Japan, Jun. 2011.

[7] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, “Optimum transmission policies for battery limited energy harvesting nodes,” *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 1180–1189, Mar. 2012.

[8] K. Tutuncuoglu and A. Yener, “Optimal power policy for energy harvesting transmitters with inefficient energy storage,” in *Proc. Annual Conference on Information Sciences and Systems*, Princeton, NJ, Mar. 2012.

[9] C. K. Ho and R. Zhang, “Optimal energy allocation for wireless communications powered by energy harvesters,” in *Proc. IEEE International Symposium on Information Theory*, Austin, TX, Jun. 2010, pp. 2368–2372.

[10] ——, “Optimal energy allocation for wireless communications with energy harvesting constraints,” *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4808–4818, Sep. 2012.

[11] C. Huang, R. Zhang, and S. Cui, “Optimal power allocation for outage probability minimization in fading channels with energy harvesting constraints,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 2, pp. 1074–1087, Feb. 2014.

[12] O. Ozel, K. Tutuncuoglu, S. Ulukus, and A. Yener, “The binary energy harvesting channel with on-off fading,” in *Proc. IEEE International Symposium on Information Theory*, Hong Kong, China, Jun. 2015, pp. 2722–2726.

[13] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, “Adaptive transmission policies for energy harvesting wireless nodes in fading channels,” in *Proc. Annual Conference on Information Sciences and Systems*, Baltimore, MD, Mar. 2011.

[14] B. Varan, K. Tutuncuoglu, and A. Yener, “Energy harvesting communications with continuous energy arrivals,” in *Information Theory and Applications Workshop*, San Diego, CA, Feb. 2014.

[15] J. Yang and S. Ulukus, “Optimal packet scheduling in an energy harvesting communication system,” *IEEE Transactions on Communications*, vol. 60, no. 1, pp. 220–230, Jan. 2012.

[16] X. Wang, J. Gong, C. Hu, S. Zhou, and Z. Niu, "Optimal power allocation on discrete energy harvesting model," *EURASIP Journal on Wireless Communications and Networking*, vol. 2015, no. 1, p. 48, Mar. 2015.

[17] A. Kazerouni and A. Özgür, "Optimal online strategies for an energy harvesting system with Bernoulli energy recharges," in *Proc. International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, Mumbai, India, May 2015, pp. 235–242.

[18] Q. Wang and M. Liu, "When simplicity meets optimality: Efficient transmission power control with stochastic energy harvesting," in *Proc. International Conference on Computer Communications*, Turin, Italy, Apr. 2013, pp. 580–584.

[19] F. Amirnavaei and M. Dong, "Online power control strategy for wireless transmission with energy harvesting," in *IEEE International Workshop on Signal Processing Advances in Wireless Communications*, Turin, Italy, Jul. 2015.

[20] J. Xu and R. Zhang, "Throughput optimal policies for energy harvesting wireless transmitters with non-ideal circuit power," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 2, pp. 322–332, Feb. 2014.

[21] D. Shaviv and A. Özgür, "Universally near optimal online power control for energy harvesting nodes," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 12, pp. 3620–3631, Dec. 2016.

[22] V. Sharma, U. Mukherji, V. Joseph, and S. Gupta, "Optimal energy management policies for energy harvesting sensor nodes," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 1536–1276, Apr. 2010.

[23] S. Mao, M. H. Cheung, and V. W. Wong, "An optimal energy allocation algorithm for energy harvesting wireless sensor networks," in *Proc. IEEE International Conference on Communications*, Ontario, Canada, Jun. 2012, pp. 265–270.

[24] Z. Wang, V. Aggarwal, and X. Wang, "Iterative dynamic water-filling for fading multiple-access channels with energy harvesting," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 3, pp. 382–395, Mar. 2015.

[25] H. A. Inan and A. Özgür, "Online power control for the energy harvesting multiple access channel," in *International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*, Tempe, AZ, May 2016.

[26] M. B. Khuzani and P. Mitran, "On online energy harvesting in multiple access communication systems," *IEEE Transactions on Information Theory*, vol. 60, no. 3, pp. 1883–1898, Mar. 2014.

[27] K. Tutuncuoglu and A. Yener, "The energy harvesting multiple access channel with energy storage losses," in *Proc. IEEE Information Theory Workshop*, Lausanne, Switzerland, Sept. 2012, pp. 94–98.

[28] J. Yang and S. Ulukus, "Optimal packet scheduling in a multiple access channel with rechargeable nodes," in *Proc. IEEE International Conference on Communications*, Kyoto, Japan, Jun. 2011.

[29] O. Ozel, J. Yang, and S. Ulukus, "Optimal transmission schemes for parallel and fading Gaussian broadcast channels with an energy harvesting rechargeable transmitter," *Elsevier Computer Communications*, vol. 36, no. 12, pp. 1360–1372, Jul. 2013.

[30] K. Tutuncuoglu and A. Yener, "Optimal power control for energy harvesting transmitters in an interference channel," in *Proc. Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2011, pp. 378–382.

[31] ——, "Sum-rate optimal power policies for energy harvesting transmitters in an interference channel," *Journal of Communications and Networks*, vol. 14, no. 2, pp. 151–161, Apr. 2012.

[32] ——, "Transmission policies for asymmetric interference channels with energy harvesting nodes," in *Proc. IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, San Juan, Puerto Rico, Dec. 2011, pp. 197–200.

[33] O. Ozel and S. Ulukus, "Achieving awgn capacity under stochastic energy harvesting," *IEEE Transactions on Information Theory*, vol. 58, no. 10, pp. 6471–6483, Oct. 2012.

[34] Y. Dong, F. Farnia, and A. Özgür, "Near optimal energy control and approximate capacity of energy harvesting communication," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 3, pp. 540–557, Mar. 2015.

[35] Y. Dong and A. Ozgur, "Approximate capacity of energy harvesting communication with finite battery," in *Proc. IEEE International Symposium Information Theory*, Honolulu, HI, Jul. 2014, pp. 801–805.

[36] D. Shaviv and A. Ozgur, "Capacity of the AWGN channel with random battery recharges," in *Proc. IEEE International Symposium on Information Theory*, Hong Kong, China, Jun. 2015, pp. 136–140.

[37] D. Shaviv, P.-M. Nguyen, and A. Özgür, "Capacity of the energy-harvesting channel with a finite battery," *IEEE Transactions on Information Theory*, vol. 62, no. 11, pp. 6436–6458, Nov. 2016.

[38] K. Tutuncuoglu, O. Ozel, A. Yener, and S. Ulukus, "Binary energy harvesting channel with finite energy storage," in *Proc. IEEE International Symposium on Information Theory*, Istanbul, Turkey, Jul. 2013, pp. 1591–1595.

[39] V. Jog and V. Anantharam, "An energy harvesting awgn channel with a finite battery," in *Proc. IEEE International Symposium on Information Theory*, Honolulu, HI, Jul. 2014, pp. 806–810.

[40] O. Ozel and S. Ulukus, "On the capacity region of the Gaussian MAC with batteryless energy harvesting transmitters," in *Proc. IEEE Global Communications Conference*, Anaheim, CA, Dec. 2012, pp. 2385–2390.

[41] H. A. Inan, D. Shaviv, and A. Ozgur, "Capacity of the energy harvesting Gaussian MAC," in *Proc. IEEE International Symposium on Information Theory*, Barcelona, Spain, Jul. 2016, pp. 2744–2748.

[42] E. Uysal-Biyikoglu, B. Prabhakar, and A. El Gamal, "Energy-efficient packet transmission over a wireless link," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 487–499, Aug. 2002.

[43] M. A. Zafer and E. Modiano, "A calculus approach to energy-efficient data transmission with quality-of-service constraints," *IEEE/ACM Transactions on Networking*, vol. 17, no. 3, pp. 898–911, Jun. 2009.

[44] A. Fu, E. Modiano, and J. Tsitsiklis, "Optimal energy allocation for delay-constrained data transmission over a time-varying channel," in *Proc. Annual Joint Conference of the IEEE Computer and Communications Societies*, vol. 2, San Francisco, CA, Apr. 2003, pp. 1095–1105.

[45] P. Nuggehalli, V. Srinivasan, and R. R. Rao, "Delay constrained energy efficient transmission strategies for wireless devices," in *Proc. IEEE Annual Joint Conference of the Computer and Communications Societies*, vol. 3, New York, NY, Jun. 2002, pp. 1765–1772.

[46] E. Uysal-Biyikoglu and A. El Gamal, "On adaptive transmission for energy efficiency in wireless data networks," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3081–3094, Nov. 2004.

[47] A. C. Güngör and D. Gündüz, "Proactive wireless caching at mobile user devices for energy efficiency," in *International Symposium on Wireless Communication Systems*, Brussels, Belgium, Aug. 2015, pp. 186–190.

[48] B. Gaujal, N. Navet, and C. Walsh, "Shortest-path algorithms for real-time scheduling of FIFO tasks with minimal energy use," *ACM Transactions on Embedded Computing Systems*, vol. 4, no. 4, pp. 907–933, Nov. 2005.

[49] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, and A. Yener, "Transmission with energy harvesting nodes in fading wireless channels: Optimal policies," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 8, pp. 1732–1743, Sep. 2011.

Ali Tajer (S05, M10, SM15) is an Assistant Professor of Electrical, Computer, and Systems Engineering at Rensselaer Polytechnic Institute. He received the Ph.D. degree in Electrical Engineering from Columbia University. His research interests include mathematical statistics and network information theory, with applications in wireless communications and power grids. Dr. Tajer serves as an Editor for the *IEEE Transactions on Communications*, and an Editor for *the IEEE Transactions on Smart Grid*. He is a senior member of the IEEE and received an NSF CAREER award in 2016.

Maha Zohdy received the B.Sc. degree in electrical engineering from Cairo University, Egypt and the M.Sc. degree from the School of Communication and Information Technology at Nile University. She is a PhD graduate student in the Department of Electrical, Computer, and Systems Engineering at Rensselaer Polytechnic Institute. Her research interests include wireless communication theory and networks Information theory, optimization in wireless networks, and caching in wireless networks.

K. Alnajjar received the PhD degree in Electrical and Electronics Engineering from the University of Canterbury, New Zealand in 2015. Currently she is an assistant professor with the department of Electrical and Computer Engineering, University of Sharjah, UAE. Her research interests include multiple-input multiple-output communications and analysis of communication systems