Average Throughput Maximization for Energy Harvesting Transmitters With Causal Energy Arrival Information

Qing Bai, Rana Ali Amjad and Josef A. Nossek
Institute for Circuit Theory and Signal Processing
Technische Universität München, Munich, Germany
Email: bai.qing@nws.ei.tum.de

***Abstract*—We consider an energy harvesting node which transmits data using the energy it harvests from the environment. In the simple scenario of point-to-point communication, the performance of the node in terms of throughput is greatly influenced by the transmission strategy it employs and its knowledge about the energy arriving process. We assume in this work that the transmitting node does not have non-causal information about the energy to be harvested in future, but only has available the statistics of the energy arriving process which is stationary. The practical considerations that the energy storage capacity of the node is limited and there is additional energy consumption within the circuitry of the node are taken into account by our system model. Viewing the system as a finite-state Markov decision process, we optimize the transmission policy the node employs as a function of the energy storage state after an energy arrival, by using the policy-iteration algorithm. The asymptotic performance of the system in terms of average throughput is studied within the established theoretical and algorithmic framework, under the several transmission strategies we propose. Simulation results indicate the advantage of each strategy with respect to a certain range of values that the system parameters may take.**

I. Introduction

As a promising alternative to traditional communication devices powered by batteries or fixed utilities, energy harvesting nodes are able to recharge their energy storage with the energy they harvest from the ambience, and therefore serve as an attractive solution for *green communications* and admit longer lifetime of the system or network they are employed in. Depending on the external energy source, the energy arrival process at the node exhibits different properties [1][2]. In general, the energy that the node is able to harvest has a random nature and can not be reliably predicted. For point-to-point communications with an energy harvesting transmitter, the performance limit of the system measured in short-term throughput is evaluated in [3][4], where perfect energy arrival information during the time slot of interest has been assumed available. The optimal control theory has been applied in [4][5] for finding the optimal transmit power allocation in time. In the more practical scenario that the transmitter only has statistical information about the energy arrivals, different methodologies are required for analysing the control and the subsequent behaviour of the system. In [6], transmission strategies that achieve the channel capacity with an energy harvesting transmitter and random amounts of energy arrivals are discussed, where infinite energy storage capacity is assumed. The dynamic programming technique is applied in [7] for maximizing the expected throughput on a finite time interval, where quantization is performed in time for the value-iteration method. In the more recent work [8], the optimal transmit power as a function of stored energy to maximize long-term average throughput is determined by solving a first order non-linear ordinary differential equation.

With the assumption of *compound Poisson energy arrivals* where the exact probability distributions are known to the transmitter, the focus and novelty of our work here include the formulation and solution of a long-term average throughput maximization problem, where the system is treated as a *finite-state Markov decision process* and the *policy-iteration* algorithm is applied for optimizing transmission policies. The optimization framework is not restricted to the specific energy arrival distributions or proposed transmission strategies. Moreover, a generalized circuit power model is incorporated, the impact of which is studied from the perspectives of theory, algorithm, and also the system performance.

The rest of the paper is organized as follows: in Section II we introduce the system model from the aspects of data transmission, energy consumption and energy harvesting. The maximization of average throughput per stage is formulated by an optimization in Section III. We first consider the scenario with no circuit power consumption in Section IV, where we start with tackling the single-stage problem and then explain the mechanism and implementation of the policy-iteration algorithm. We turn to the more general case with non-zero circuit power in Section V, where the emphasis is put on the necessary modifications that should be made to existing conclusions and algorithms. Simulation results are shown at the end of Section IV and V corresponding to the two test scenarios, and finally we summarize the paper in Section VI.

II. System Model

The scenario that an energy harvesting node transmits data over a single, invariant link for a boundless time period using the energy it collects in the meanwhile is considered. We adopt a continuous-time model and assume that the transmit power of the node, denoted with $p_{\mathrm{tx}}(t)$, can be adapted continuously.

A. Data Transmission Model

Let $f(p_{\mathrm{tx}}(t))$ be the instantaneous data rate as dependent on the transmit power, where the function f is assumed to be nonnegative, strictly concave as well as monotonically increasing. The dependency given by f is time-invariant since we assume that the channel stays constant. The *throughput* of the system over a certain time interval is defined as the integral of f on it. For numerical studies and simulations we employ the rate function $f(p_{\mathrm{tx}}) = \log(1 + p_{\mathrm{tx}})$, which is in analogy to the Shannon capacity formula with normalized bandwidth and channel-to-noise power ratio. Yet the analysis and the algorithms we propose in this work can be applied with all functions f that satisfy the aforementioned conditions.

B. Energy Consumption Model

Besides the radiated power p_{tx}, there is additional power consumption within the circuit of the transmitter incurred by the D/A converter, the mixer, the filters and the power amplifier, *etc*. For the typical communication range in a wireless sensor network where energy harvesting nodes are of particular interest, the transmit power required to achieve sufficiently good receive signal-to-noise ratio (SNR) is on the same order of magnitude as the analog/digital processing power of the transmitting node [9]. Therefore, it is necessary and also important to include circuit power into the energy consumption model of the system. Let $p_{\mathrm{c}}(t)$ be the power consumption of the transmitter circuitry at time t. We formulate p_{c} as a function of the transmit power, *i.e.*, $p_{\mathrm{c}}(t) = g_{\mathrm{c}}(p_{\mathrm{tx}}(t))$, and obtain the total power dissipation formula as $p(t) = g(p_{\mathrm{tx}}(t)) = p_{\mathrm{tx}}(t) + p_{\mathrm{c}}(t)$. The function g_{c} is assumed nondecreasing on $[0, +\infty)$, continuous and continuously differentiable on $(0, +\infty)$. Note that a single discontinuous point of $g_{\mathrm{c}}(p_{\mathrm{tx}})$ at $p_{\mathrm{tx}} = 0$ is allowed due to the different modes that the transmitter could be operating on. When the transmitter is not sending any signal, it can be turned into *sleep* mode for which the circuit power consumption is low enough to be neglected, *i.e.*, we assume $g_{\mathrm{c}}(0) = 0$. Otherwise, the transmitter is considered in *active* mode and its circuit consumes ineligible energy, which means $g_{\mathrm{c}}(p_{\mathrm{tx}}) > 0$ for $p_{\mathrm{tx}} > 0$. Furthermore, as the convexity of g_{c} plays a crucial role in preserving the convex structure of the throughput maximization problem [4], we assume here that g_{c} is convex on $(0, +\infty)$. This condition is indeed met by more practical system models, *e.g.*, the MQAM model presented in [9].

On investigating the optimal transmission policy of the energy harvesting node, we start with the simple case that $g(p_{\mathrm{tx}}) = p_{\mathrm{tx}}$, *i.e.*, circuit power is completely neglected, and then extend the study to non-zero circuit power. Since the power dissipation of the power amplifier is usually assumed to be linearly dependent on transmit power, and power consumption of other components of the transmitter can be assumed constant [9], we employ the following linear power consumption model for numerical analysis as well as simulations

$$g(p_{\mathrm{tx}}) = \begin{cases} (1+b) \cdot p_{\mathrm{tx}} + c, & p_{\mathrm{tx}} > 0, \\ 0, & p_{\mathrm{tx}} = 0, \end{cases} \tag{1}$$

where b and c are both nonnegative constants. Without loss of generality, let the system begin operation at the time instance $t_0 = 0$. The total energy consumption of the node by time t, represented by the non-decreasing function $W(t)$, is given as

$$W(t) = \int_0^t g(p_{\mathrm{tx}}(\tau))\,\mathrm{d}\tau.$$

Naturally, there is the initial condition $W(0) = 0$, and $W(t)$ is subject to causality constraint which is explained next.

C. Energy Harvesting Model

An energy harvesting node gathers energy from the environment and stores them in its storage medium. We consider here that the harvested energy becomes available to the transmitter in the form of *energy packets*, *i.e.*, the energy arrives at discrete time instances with various amounts. Let U_n and t_n denote the size of the nth packet and the time instance at which it arrives, respectively. The interval between two consecutive energy arrivals is called a *stage*, and the inter-arrival time $t_n - t_{n-1}$ gives the duration of the nth stage. Let the function $N(t)$ indicate the number of arrivals until time t. The cumulative harvested energy by t, denoted with $U(t)$, can be written as

$$U(t) = A_0 + \sum_{n=1}^{N(t)} U_n, \quad \forall t \geq 0,$$

where A_0 stands for the amount of energy in storage at t_0.

The values of U_n, $n = 1, 2, \ldots$ are determined by the ambience of the node and the energy harvesting technology it is using. Due to the limited capacity of the energy storage, the actual amount of energy input at the beginning of each stage might be smaller than what the environment has to offer. In other words, the energy input to the system at the nth energy arrival, denoted with A_n, is upper bounded by U_n. Let $E_{\max}$ be the maximum amount of energy that the node can store and assume it a finite constant. The state of energy storage at time t, represented by the function $Z(t)$, is assumed perfectly known by the node $\forall t$. The sequence of energy inputs, the cumulative energy input function $A(t)$, and the relation between energy storage $Z(t)$, energy input $A(t)$ and energy consumption $W(t)$, are respectively given by

$$\begin{aligned} A_n &= \min\left(U_n, E_{\max} - Z(t_n)\right), \quad n = 1, 2, \ldots, \\ A(t) &= A_0 + \sum_{n=1}^{N(t)} A_n, \quad \forall t \geq 0, \\ Z(t) &= A(t) - W(t), \quad \forall t \geq 0. \end{aligned}$$

Note that Z and A are influenced by the transmit power p_{tx} the node chooses to use as well as the storage capacity $E_{\max}$. The causality restriction requires $W(t) \leq A(t)$ to hold $\forall t \geq 0$.

In practice, when and how much energy can be harvested is generally random and not exactly predictable at the transmitting node. We assume here that the energy harvesting node has perfect causal information about the energy arrivals as well as their statistics, and this statistical information does not change over time. To be more specific, the energy arrivals follow a stationary Poisson process with known intensity λ_0, *i.e.*, the inter-arrival times, denoted with λ_n, are exponentially distributed with $\frac{1}{\lambda_0}$ as their mean value. The quantities of energy arrivals U_n, $n = 1, 2, \ldots$ are i.i.d. random variables taking positive values, and the availability of the probability density function is assumed. The random process indicated by $U(t)$ is compound Poisson under these conditions.

III. Average Throughput Maximization

The maximization of throughput on a given finite time slot $[0, T]$ for an energy harvesting transmitter has been investigated in [3][4], where non-causal energy arrival information during $[0, T]$ is assumed available. When the system is to operate for a long time or there is no specific termination time, it is appropriate to maximize the *average throughput* instead, for the total throughput is not well defined or approaches infinity as time evolves. To this end, the energy harvesting node aims at maximizing the *average throughput per stage* by properly adapting its transmit power based on the knowledge of energy arrival statistics, which is formulated as the optimization

$$\begin{aligned}\max_{p_{\rm tx}\in\mathcal{P}} \quad & \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} f\left(p_{\rm tx}(t)\right) \mathrm{d}t \\ \text{s.t.} \quad & W(t) = \int_0^t g\left(p_{\rm tx}(\tau)\right) \mathrm{d}\tau \le A(t), \quad \forall t \ge 0, \\ & W(0) = 0,\end{aligned} \tag{2}$$

where $\mathcal{P}$ is the set of finite, nonnegative, piecewise continuous functions defined for $t \ge 0$. As the number of stages N goes to infinity, the termination time t_N converges to its expected value $\frac{N}{\lambda_0}$. From this observation and by replacing t_N with T, we see that the optimization objective in (2) is equivalent to

$$\frac{1}{\lambda_0} \cdot \lim_{T\to\infty} \frac{1}{T} \int_0^T f\left(p_{\rm tx}(t)\right) \mathrm{d}t,$$

which is the average throughput scaled by the mean inter-arrival time. As we will analyse and solve the optimization within the theoretical framework of Markov decision processes, it is convenient to employ the formulation (2). Let the objective function of (2) be denoted with $\rho(p_{\rm tx})$.

IV. Transmission Strategies Without Circuit Power Consideration

The energy harvesting node is influenced by the energy arrival events in a discrete manner. We therefore consider a *transmission policy*, as the decision on how much resources to take for the current stage which is made at the beginning of each stage, and a *transmission strategy*, as how these resources are to be allocated on the single stage. Transmission policies are optimized with the policy-iteration algorithm as a function of storage state, whereas transmission strategies are proposed based the analysis on the single-stage throughput maximization problem, and are assumed unchanged during operation. For the content of this section we assume $g(p_{\rm tx}) = p_{\rm tx}$.

A. Single-stage Solution

Consider a single-stage of duration λ with initial energy state A_0. As the arrivals of energy follow a Poisson process, λ is an exponentially distributed random variable. The expected throughput $I(p_{\rm tx})$ on the stage can be calculated as

$$I(p_{\rm tx}) = \mathrm{E}\left[\int_0^{\lambda} f\left(p_{\rm tx}\right) \mathrm{d}t\right] = \int_0^{+\infty} \mathrm{e}^{-\lambda_0 t} f\left(p_{\rm tx}\right) \mathrm{d}t.$$

The maximization of the expected throughput can then be written in the standard form of a control problem as

$$\begin{aligned}\max_{p_{\rm tx}\in\mathcal{P}} \quad & \int_0^{+\infty} \mathrm{e}^{-\lambda_0 t} f\left(p_{\rm tx}\right) \mathrm{d}t \\ \text{s.t.} \quad & \dot{W} = g(p_{\rm tx}), \\ & W(0) = 0, \quad W(t) \le A_0, \ \forall t > 0,\end{aligned} \tag{3}$$

where we have omitted the t-argument in the time-dependent functions, and $\dot{W}$ stands for the derivative of W with respect to t. Note that the objective depends explicitly on t and it converges as $\mathrm{e}^{-\lambda_0 t}$ serves as a discounting factor for the data rate. Let the optimal control to (3) be denoted with $p_{\rm tx}^*$. The Hamiltonian of the problem is given by

$$H(t, p_{\rm tx}, \mu) = -\mathrm{e}^{-\lambda_0 t} f\left(p_{\rm tx}\right) + \mu g\left(p_{\rm tx}\right).$$

Since H does not explicitly depend on W, the co-state equation suggests that $\dot{\mu}^* = 0$, *i.e.*, μ^* is constant. According to *Pontryagin Maximum Principle* (PMP) [10][11], we have

$$H_{p_{\rm tx}}(t, p_{\rm tx}^*, \mu^*) = -\mathrm{e}^{-\lambda_0 t} f_{p_{\rm tx}}\left(p_{\rm tx}^*\right) + \mu^* g_{p_{\rm tx}}\left(p_{\rm tx}^*\right) = 0, \tag{4}$$

where a function with a variable as its subscript refers to the partial derivative of the function with respect to this variable. Plugging $f\left(p_{\rm tx}\right) = \log(1 + p_{\rm tx})$ and $g\left(p_{\rm tx}\right) = p_{\rm tx}$ into (4) and considering the non-negativity of $p_{\rm tx}$, we arrive at

$$p_{\rm tx}^*(t) = \max\left(\frac{\mathrm{e}^{-\lambda_0 t}}{\mu^*} - 1, 0\right), \quad t \ge 0. \tag{5}$$

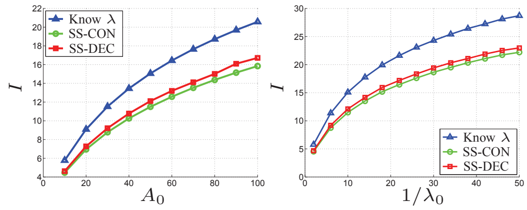
It is clear from (5) that $\mu^* \in (0, 1)$ and that the optimal transmit power $p_{\rm tx}^*$ decays exponentially in time. Starting with $\mu = 1$, the search for μ^* can be performed with decreasing μ over the iterations until the total energy consumption reaches A_0. We term this transmission strategy as SS-DEC strategy.

As a simpler alternative to the optimal transmit power given by (5), we also consider a constant power strategy where the transmit power ξ is used for the whole stage. The expected throughput $I(\xi)$ achieved under this strategy is calculated as

$$I(\xi) = \mathrm{E}\left[f(\xi) \min\left(\lambda, \frac{A_0}{g(\xi)}\right)\right] = \frac{f(\xi)}{\lambda_0}\left(1 - \mathrm{e}^{-\frac{\lambda_0 A_0}{g(\xi)}}\right), \tag{6}$$

which can be maximized by searching for the optimal constant transmit power ξ^*, which does not have a close-form expression though. This strategy is termed as SS-CON.

We compare the two transmit strategies with the simulation results shown in Figure 1. For each initial energy state A_0 and for each mean inter-arrival time $1/\lambda_0$, μ^* and ξ^* are computed offline for the SS-DEC and the SS-CON strategies,



(a) $\lambda_0 = 0.1$ (b) $A_0 = 100$

Figure 1. Throughput achieved on a single-stage

respectively. The values are used for 2×10^4 simulations and we plot the average achieved throughput in both figures. The optimal throughput with λ known in advance is depicted for reference. It can be seen that the gap between SS-CON and the optimal SS-DEC strategy is not significant. Also notice in Figure 1(b) that, although the available energy is the same, with the stage lasting longer, larger throughput can be achieved due to the better energy efficiency endowed by time.

B. Policy-Iteration Algorithm

We are aiming at the problem of maximizing the average throughput per stage (2) which involves an infinite number of stages on which the stationary system operates. With such a scenario and optimization goal, the *policy-iteration* technique is more favourable than the common dynamic programming technique based on *value-iteration* [12][13]. Moreover, it is convenient to consider the system as a finite-state Markov decision process in the context of applying the policy-iteration method for infinite horizon problems. To this end, we assume that the energy harvesting node can only change its transmission policy at the beginning of each stage. The decision made for stage n is based on the energy state $Z(t_{n-1})$ of the storage. Let the sequence of state variables of the system, namely, the initial energy level and the amount of stored energy right after each energy arrival, be denoted with Z_n, $n = 0, 1, \ldots$. The continuous state space $[0, E_{\max}]$ is discretized with granularity δ to a discrete set of M states, which means the true values of Z_n are approximated by

$$s_i = (i-1)\delta, \quad i = 1, \ldots, M \quad \text{with} \quad M - 1 = \left\lfloor \frac{E_{\max}}{\delta} \right\rfloor.$$

Under a good transmission strategy, the energy harvesting node tries to reduce, at least to some extent, the chance that energy miss events defined by $A_n < U_n$ happen, which are resulted from insufficient storage capacity. This can be accomplished by employing a reasonably large transmit power. Since p_{tx} is not bounded from above as assumed in this work, the underlying Markov process of a good transmitter has only one recurrent chain and is completely ergodic. This means, the initial state of the system, *i.e*, $Z_0 = A_0$, does not influence the average throughput that can be achieved with the strategy.

The policy-iteration algorithm consists of the *Value-Determination Operation* and the *Policy-Improvement Routine* in each of its iteration cycles [14]. In the value-determination phase, the system is operated under a given policy so that for each state of the system, the throughput to be expected on the current stage can be computed. In addition, the transition probability of the system from the current state to any possible next state can be specified. Let i be the state index of the system at the current stage and j be the state index for the next stage. A set of M linear equations can be established as

$$\rho + v_i = q_i + \sum_{j=1}^{M} P_{ij} v_j, \quad i = 1, \ldots, M, \qquad (7)$$

where P_{ij} stands for the transition probability from state i to state j, and q_i stands for the expected throughput to be achieved on the current stage with state i. It is obvious from (7) that shifting the unknowns $v_1, \ldots, v_M$ by the same constant amount does not change the equations since $\sum_{j=1}^{M} P_{ij} = 1$. Therefore, we can set the unknown v_M to 0 which results in exactly M unknowns to be solved by the equation system, namely, $\rho, v_1, v_2, \ldots, v_{M-1}$. The resulting $v_1, v_2, \ldots, v_{M-1}$ are called the *relative values* of the given policy and serve as inputs to the policy-improvement routine that follows.

The policy-improvement routine, as contrary to the value-determination operation which yields relative values based on a given policy, produces new policies as a function of input relative values. Mathematically, it solves the following optimizations with respect to all feasible policies in each state

$$\max_{p_{\text{tx}}} \quad q_i(p_{\text{tx}}) + \sum_{j=1}^{M} P_{ij}(p_{\text{tx}}) v_j, \quad i = 1, \ldots, M. \qquad (8)$$

The optimal policy for each state is then recorded and used for computing new relative values. The computation of transition probabilities can be rather tedious depending on the distribution of the quantity of each energy arrival, and might turn out more demanding than solving (7) and (8). The policy-iteration algorithm can be started with the policy-improvement routine with all relative values initialized to 0. For each iteration cycle, the obtained average throughput per stage is improved. We find via numerical experiments that the algorithm converges already with very few iterations, usually less than 5. Details on the algorithm as well as the proof for the monotonically increasing property of ρ over iterations can be found in [14].

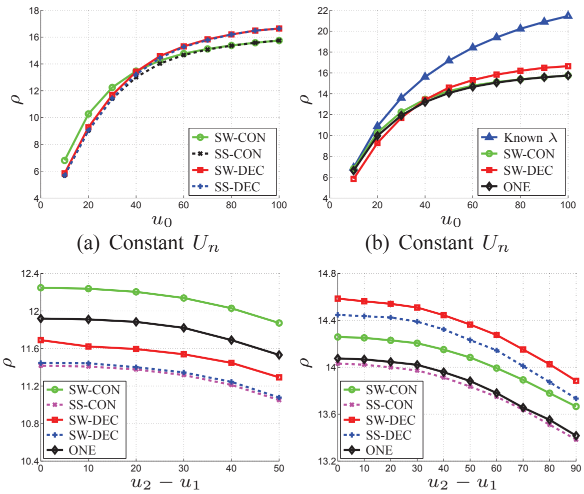
What is obtained in the end with the policy-iteration algorithm is a look-up table with M entries. With each entry which corresponds to one system state, the transmission policy that the node should take in that state is specified. The remaining task now is to figure out the general strategy the energy harvesting node employs, which determines how the system operates in between two states. Based on our study on the single-stage problem in the previous section, we propose 3 transmission strategies, the *constant* (ONE) strategy, the *stage-wise constant* (SW-CON) strategy and the *stage-wise decreasing* (SW-DEC) strategy. With strategy ONE, a constant transmit power is always used regardless of which state the system is in. With the SW-CON strategy, the transmitter employs constant transmit power on each stage, whereas with the SW-DEC strategy, the transmitter uses decreasing transmit power indicated by (5). In the last case, the key to the transmission policy which we optimize iteratively is in fact the constant μ^*. It should be noted that the policy-iteration algorithm provides the optimal transmission policy with respect to the given transmission strategy, but not necessarily the global optimal solution to (2). We choose ONE and SW-CON for their simplicity and SW-DEC for its optimality for the single-stage expected throughput maximization problem.

C. Simulation Results

Simulations are performed to verify the implementation of the policy-iteration algorithm, and more importantly, to enable comparisons between the proposed transmission strategies. Fixed simulation parameters are listed in Table I, and for each optimized policy with the varying system parameters, average results for N operation stages are shown in Figure 2.

Table I
SIMULATION PARAMETERS

$E_{\max}$	λ_0	N	δ
100	0.1	2×10^4	0.25



(a) Constant U_n (b) Constant U_n

(c) Uniform U_n, $u_0 = 30$ (d) Uniform U_n, $u_0 = 50$

Figure 2. Average throughput per stage achieved with different strategies

In Figure 2(a) and 2(b), the amount of energy in each arrival is identically u_0. The SW-CON and SW-DEC strategies are compared with their single-stage counterparts SS-CON and SS-DEC in Figure 2(a), for which the policies at each stage are not optimized with the policy-iteration algorithm but are only optimal for the current stage. With u_0 approaching $E_{\max}$, the myopic policies perform as good as the corresponding optimized policies, whereas for small u_0 their performance is relatively worse. The performance gap between the optimized policy and the myopic policy for the DEC strategy is in general quite small. Note that in terms of memory and online computational requirements, the optimized policy and the myopic policy are equivalent. For the SW-CON and SW-DEC strategies, the maximal average throughput found with the policy-iteration algorithm coincides very well with the simulation results, and therefore we do not need to distinguish them on figures. This applies to all simulation results we show in the paper. We compare the performance of SW-CON, SW-DEC and ONE strategies in Figure 2(b), where the optimal average throughput with non-causal energy arrival information is plotted as a reference. The behaviour of strategy ONE is found to be similar to SW-CON. It is interesting to note the crossing point of the curves representing the SW-CON and SW-DEC strategies: for $u_0 \le 40$ approximately, SW-CON performs better while for $u_0 > 40$, SW-DEC yields more throughput. Such a result can be expected as with larger energy arrivals, using decreasing power effectively reduces the chance of energy miss, which might yet sacrifice energy efficiency too much when the energy arrivals are with a low quantity.

Energy arrivals with uniformly distributed quantities on the interval $[u_1, u_2]$ are assumed for the simulations shown in Figure 2(c) and 2(d), where the mean values are $u_0 = 30$ and $u_0 = 50$ respectively. The performance of the system deteriorates as the energy arrivals become more diverse, and the gaps between different strategies can be clearly observed.

V. TRANSMISSION STRATEGIES WITH CIRCUIT POWER CONSIDERATION

The effect of circuit power to the transmission strategies and algorithms we propose as well as to system performance is discussed in this section. We have shown in [4] that, when g_c is convex and continuous on $[0, +\infty)$, the same algorithm to find the optimal transmission strategy without circuit power can be applied, and we only need to recover p^*_{tx} from the obtained optimal total power dissipation p^* with the inverse function g^{-1}. When g_c is discontinuous at $p_{\text{tx}} = 0$, the so-called *energy-efficient transmit power*, denoted with p_{tx_0} and defined by the solution to the equation

$$(f_{p_{\text{tx}}} g - f g_{p_{\text{tx}}})(p_{\text{tx}}) = 0,$$

plays a critical role in the optimal transmission strategy in that $p^*_{\text{tx}} \ge p_{\text{tx}_0}$ has to be satisfied $\forall t$. We mainly focus on the latter case here for the special treatment it requires.

A. Single-stage and Infinite-horizon Solutions

Our derivation of the solution to the single-stage problem (3) still holds until the relation (4) obtained with the PMP. Due to the discontinuity of g_c at the zero point, we confine the end-point of the control problem to t_1 with $p_{\text{tx}} > 0$ for $t \le t_1$ and $p_{\text{tx}} = 0$ for $t > t_1$, such that an equivalent control problem with free termination time and a continuous state equation is attained. The transversality condition at end-point t_1 requires

$$H(t_1, p^*_{\text{tx}}(t_1), \mu^*) = -\mathrm{e}^{-\lambda_0 t_1} f(p^*_{\text{tx}}(t_1)) + \mu^* g(p^*_{\text{tx}}(t_1)) = 0.$$

Evaluating (4) at t_1 and plugging the result above, we have

$$-f\left(p^*_{\text{tx}}(t_1)\right) g_{p_{\text{tx}}}\left(p^*_{\text{tx}}(t_1)\right) + f_{p_{\text{tx}}}\left(p^*_{\text{tx}}(t_1)\right) g\left(p^*_{\text{tx}}(t_1)\right) = 0,$$

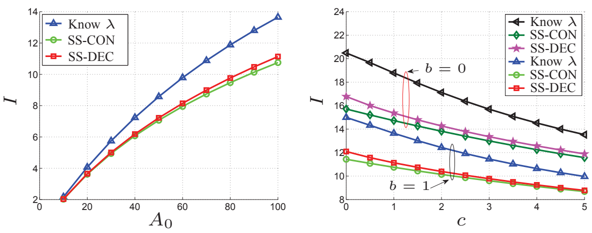
which means the optimal transmit power at the end-point t_1 equals the energy-efficient transmit power p_{tx_0}. When $f = \log(1 + p_{\text{tx}})$ and g is given by the linear model (1), we have

$$p^*_{\text{tx}}(t) = \begin{cases} \frac{\mathrm{e}^{-\lambda_0 t}}{(1+b)\mu^*} - 1, & \frac{\mathrm{e}^{-\lambda_0 t}}{(1+b)\mu^*} - 1 \ge p_{\text{tx}_0}, \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

Similar to the case without circuit power, a search for μ^* is performed which terminates when the energy consumption until end-point t_1 equals A_0. The SS-CON strategy can also be considered, where the optimal solution ξ^* to (6) can be shown to satisfy $\xi^* > p_{\mathrm{tx}_0}$ for general concave rate functions f and convex circuit power functions g. In the infinite horizon scenario, the policy-iteration algorithm works with transmission strategies suggested by solutions to the single-stage problem, which does not change in principle as previously.

B. Simulation Results

We use the same set of parameters as given in Table I and the linear circuit power model (1) for simulations. The optimal and suboptimal solutions to the single-stage problem are first studied in Figure 3, where we vary the initial energy state A_0 as well as the circuit power parameters b and c. The near-optimal performance of strategy SS-CON can be observed.



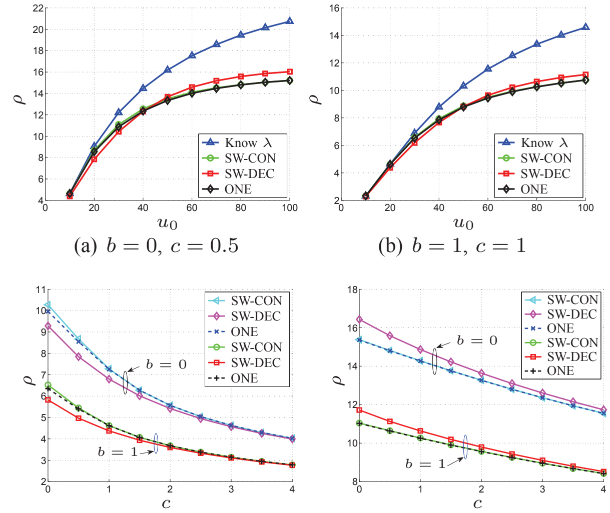
(a) $b = 1, c = 1$ (b) $A_0 = 100$

Figure 3. Throughput achieved on a single-stage

For the average throughput maximization over an infinite number of stages, we assume all energy arrivals are of the same quantity u_0. From Figure 4 we immediately observe the deterioration of system performance due to the non-zero circuit power. A shift of the crossing point of strategies SW-CON and SW-DEC to the right can be found on the first two figures where circuit power parameters are increased. The conclusion that strategy SW-CON is more favourable in the low-energy regime while SW-DEC is preferred in the high-energy regime is further emphasized with the last two figures. Notice that in a less dynamic system with constant energy arrivals, strategy ONE performs almost as good as SW-CON, and the myopic strategies also come close to their optimized counterparts for which reason they are not shown on the figures.

VI. Conclusions

We investigate in this paper transmission strategies that an energy harvesting node with only causal and statistical energy arrival information should employ in order to maximize the average throughput. The system is modelled as a finite-state Markov decision process for which we apply the policy-iteration algorithm to find the optimal transmission policies. The theoretical framework and the algorithm we use are not restricted to one specific energy arrival distribution or one particular circuit power model. Stage-wise constant transmit power, Stage-wise decreasing transmit power and constant transmit power are the three strategies we propose for the energy harvesting transmitter. Depending on the amount and diversity of energy arrivals and the energy storage capacity



(a) $b = 0,\ c = 0.5$ (b) $b = 1,\ c = 1$

(c) $u_0 = 20$ (d) $u_0 = 80$

Figure 4. Average throughput per stage achieved with different strategies

of the system, one strategy might outperform or be beaten by another strategy, which are demonstrated with simulation results, yet the common advantage of the three strategies is their low computational requirement both offline and online.

References

[1] S. Sudevalayam, P. Kulkarni, *Energy Harvesting Sensor Noses: Survey and Implications*, IEEE Communications Surveys & Tutorials, vol. 13, no. 3, 2011.
[2] M. Gorlatova, A. Wallwater, G. Zussman, *Networking Low-Power Energy Harvesting Devices: Measurements and Algorithms*, IEEE INFOCOM 2011, April 2011.
[3] K. Tutuncuoglu, A. Yener, *Optimum Transmission Policies for Battery Limited Energy Harvesting Nodes*, IEEE Transactions on Wireless Communications, vol. 11, no. 3, March 2012.
[4] Q. Bai and J. A. Nossek, *Throughput Maximization for Energy Harvesting Nodes with Generalized Circuit Power Modelling*, the 13th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2012), June 2012.
[5] M. Zafer, E. Modiano, *A Calculus Approach to Energy Efficient Data Transmission with Quality-of-Service Constraints*, IEEE/ACM Transactions on Networking, vol. 17, pp. 898-911, March 2010.
[6] O. Ozel, S. Ulukus, *Information-theoretic analysis of an energy harvesting communication system*, IEEE International Symposium on Personal, Indoor, Mobile Radio Communications, pp. 330-335, September 2010.
[7] O. Ozel, K. Tutuncuoglu, J. Yang, S. Ulukus, A. Yener, *Transmission with Energy Harvesting Nodes in Fading Wireless Channels: Optimal Policies*, IEEE Journal on Selected Areas in Communications, Sept. 2011, vol. 29, issue 8, pp. 1732-1743.
[8] P. Mitran, *On Optimal Online Policies in Energy Harvesting Systems For Compound Poisson Energy Arrivals*, IEEE International Symposium on Information Theory, Cambridge, MA, USA, July 2012.
[9] S. Cui, A. J. Goldsmith, *Energy-constrained modulation optimization*, IEEE Transactions on Wireless Communications, Sept. 2005, vol. 4, issue 5, pp. 2349-2360.
[10] L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze, E.F. Mishchenko, *The Mathematical Theory of Optimal Processes*, Pergamon Press, 1964.
[11] D. Liberzon, *Calculus of Variations and Optimal Control Theory: a Concise Introduction*, Princeton University Press, December 2011.
[12] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley-Interscience, March 2005.
[13] D. P. Bertsekas, *Dynamic Programming and Optimal Control*, vol. I, Athena Scientific, 2005, ISBN 1-886529-26-4.
[14] R. A. Howard, *Dynamic Programming and Markov Processes*, The Technology Press of The Massachusetts Institute of Technology and John Wiley & Sons, Inc., New York · London.