ENERGY-CONSTRAINED THROUGHPUT MAXIMIZATION FOR POINT-TO-POINT COMMUNICATIONS

Qing Bai, Jingrui Li and Josef A. Nossek

Technische Universität München
Institute for Circuit Theory and Signal Processing
Theresienstr. 90, 80333 Munich, Germany

ABSTRACT

We consider a very basic communication scenario: a transmitter sends data to a receiver over a single invariant link, where each side has a certain energy budget which is a common restriction for sensor nodes powered by batteries or energy harvesting components. Given the channel coefficient, the instantaneous data rate that can be achieved depends on the transmit power employed by the transmitter, and the resolution of the analog-to-digital converter (ADC) employed by the receiver. The power consumptions of the transmitter and the receiver also depend, respectively, on these two parameters. In this paper we seek to answer the question: what is the transmit and receive strategy that leads to the maximal throughput achieved on a finite time interval $[0, T]$, given the individual energy constraints at the transmitter and the receiver? The underlying control problem, which is found non-convex in its original form, can be made convex by a reconstruction of the power-rate surface based on which the optimal solution can be conveniently determined.

Index Terms— Throughput maximization, optimal control, energy-efficiency

1. INTRODUCTION

Many communication devices, such as wireless sensor nodes and mobile terminals, are limited by the energy that is containable to the batteries they have. Although their sustainability can be improved by employing energy harvesting techniques, it is of fundamental importance to make efficient utilization of the available energy. In recent years, there have been various works discussing the trade-offs between energy/energy-efficiency and the traditional performance metrics such as delay and spectrum-efficiency, *e.g.*, [1][2], leading to constrained parameter optimizations in different system contexts. When it comes to the joint optimization of the communication between a transmitter and a receiver, the problem is multidimensional which usually has a rather complex structure. On the other hand, when the usage of energy is under discussion, the impact of the variation of power consumption over time needs to be explored, which is one of the central focuses of energy harvesting systems [3][4]. This is to say, instead of optimizing a scalar or vector parameter, we optimize one or more functions of time. In this work, we investigate the joint optimization of transmit power and the ADC resolution both as functions defined on the time interval $[0, T]$, with the goal of maximizing the throughput achieved by the point-to-point communication. Although the considered scenario is simple and the statement of the problem is straightforward, this problem has not been touched upon in the literature to our best knowledge, and requires careful treatment due to its non-convexity.

2. SYSTEM MODEL

We consider the data transmission from a transmitter to a receiver in a given time interval $[0, T]$, where both sides are not supported by constant power supplies but have a certain energy budget. It is assumed here that the transmitter and the receiver both have one single antenna, and the communication channel between them stays constant during the time interval of interest. Furthermore, we assume that there is a central control unit in the system which has perfect knowledge about all relevant system parameters as well as the channel state, so that it is able to jointly optimize the transmit and receive strategies to achieve the maximal possible throughput. Transmit power employed by the transmitter, denoted with p_{tx}, and the ADC resolution employed at the receiver, denoted with b, are taken as the two control variables which we can adapt continuously in time and magnitude. In this section, the relations between the control variables and the achievable data rate as well as the energy consumptions are to be explained.

Let the constant channel coefficient during $[0, T]$ be $h \in \mathcal{C}$, and the receive noise power be σ^2. With transmit power p_{tx}, the receive signal-to-noise ratio (SNR) γ can be written as $\gamma = |h|^2 \cdot p_{\text{tx}}/\sigma^2$. The channel capacity as dependent on γ and the ADC resolution b is lower bounded by [5]

$$f(\gamma, b) = \log_2 \left(\frac{1+\gamma}{1+\gamma \cdot 2^{-2b}} \right), \tag{1}$$

where the transmission bandwidth has been normalized without loss of generality. In the aforementioned paper it was shown that this lower bound is quite tight especially in the low SNR region. Therefore, we employ (1) as our rate function $R = f(\gamma, b)$, and define *throughput* ρ as the integral of the instantaneous data rate R from time 0 to T as

$$\rho = \int_0^T f(\gamma, b)\,\mathrm{d}t.$$

Note that ρ is a functional of $b(t)$ and $\gamma(t)$, or equivalently, $p_{\mathrm{tx}}(t)$, which are functions of time defined on $[0, T]$. Yet for notational simplicity we usually neglect the time index t.

The power consumption of the transmitter and the receiver, denoted with P_1 and P_2 respectively, are given by

$$P_1 = \xi \cdot p_{\mathrm{tx}} = \frac{\xi\sigma^2}{|h|^2} \cdot \gamma \triangleq a_1\gamma,$$
$$P_2 = \begin{cases} 0, & b = 0, \\ c\sigma^2 \cdot 2^{2b} \triangleq a_2 \cdot 2^{2b}, & b > 0. \end{cases} \quad (2)$$

For the transmitter, we assume that the power consumption is proportional to the transmit power, where the scaling factor ξ is related to the efficiency of the power amplifier. It is often considered that the power consumption consists of another constant part when $p_{\mathrm{tx}} > 0$, resulting from other circuit components [6], which we do not take into our model for simplicity. We distinguish the operation modes of the receiver, on the other hand, by whether a positive ADC resolution b is employed: the receiver is in *active* mode when $b > 0$ and is in *sleeping* mode when $b = 0$, for which we assume that the power consumption is negligible. For the active mode, the power consumption is proportional to 2^{2b} and the receive noise power, scaled by some constant parameter c determined by the specific design of the ADC [5]. In (2) the constant scaling factors in P_1 and P_2 are put into two parameters a_1 and a_2, leading to a more concise description of the system.

3. THROUGHPUT MAXIMIZATION

Let the available energy at the transmitter and the receiver be A_1 and A_2, respectively. Throughput of the system is to be maximized with efficient utilization of these energy, by adapting γ (p_{tx}) and b properly during $[0, T]$. Mathematically, this can be formulated as a standard control problem as

$$\begin{aligned} \max_{\gamma\geq 0, b\geq 0} \quad & \int_0^T f(\gamma, b)\,\mathrm{d}t \\ \text{s.t.} \quad & \dot{W}_1 = P_1(\gamma), \quad \dot{W}_2 = P_2(b), \\ & W_1(0) = 0, \quad W_2(0) = 0, \\ & W_1(T) \leq A_1, \quad W_2(T) \leq A_2, \end{aligned} \quad (3)$$

where W_1 and W_2 stand for the cumulative energy consumption of the transmitter and the receiver, respectively, which are referred to as *state* variables. The derivatives of W_1 and W_2 with respect to time correspond to the power consumption functions P_1 and P_2, which serve as state equations of the system. The energy consumptions at time 0 are naturally 0, and they should not exceed the energy budgets at time T.

For better tractability of the optimization, we assume that both p_{tx} and b can take any non-negative values although in practice, the ADC resolution b is an integer. The obtained fractional valued b can nevertheless be achieved via time-sharing of integer valued resolutions. However, the discontinuity of function P_2 at $b = 0$, and the fact that function f is not jointly concave in γ and b, render Problem (3) non-convex and prohibit the direct application of the optimal control theory, which requires convexity of the problem to guarantee the sufficiency of the first-order optimality conditions.

In our previous work [4], we resolve the throughput maximization problem of an energy-constrained transmitter, which also has a discontinuous state equation due to the circuit power associated only with the active operation mode. The optimal solution we derived there can be interpreted from a geometric point of view: a tangent line is made from the point $(P, R) = (0, 0)$, which corresponds to the sleeping mode of the transmitter, to the power-rate curve, and the power values between 0 and that of the tangent point should be achieved via time-sharing. This is to say, after transforming the control variable to the power consumption P and convexify the P-R curve, we obtain a convex optimization problem which can be handled quite easily. Observing this, we seek to convexify the power-rate surface in our problem, where both the discontinuous point $(P_1, P_2, R) = (0, 0, 0)$ and the non-concave region of the surface need to be taken into account.

4. OPTIMAL TRANSMIT AND RECEIVE STRATEGY

Following the approach we conceive and describe in the last section, we first reformulate (3) into an optimization on P_1 and P_2, where the rate function is now given by

$$R = g(P_1, P_2) = \begin{cases} 0, & P_2 = 0, \\ \log_2\left(\frac{a_1 + P_1}{a_1 + a_2 \cdot \frac{P_1}{P_2}}\right), & P_2 > a_2. \end{cases}$$

It can be shown that g is concave both in P_1 and P_2 for $P_2 > a_2$, yet the determinant of its Hessian matrix

$$|\boldsymbol{H}(g)| = -\frac{a_1 a_2}{(\ln 2)^2} \cdot \frac{a_1 a_2 + 2P_1(a_2 - P_2)}{(a_1 + P_1)^2 P_2^2 (a_2 P_1 + a_1 P_2)^2}$$

is not always positive. Let the ratio between P_1 and P_2 be α. With fixed $\alpha > 0$, g is concave in P_1 for $P_1 \in (a_2\alpha, +\infty)$. As $g(0, 0) = 0$, function g with respect to the fixed α can be convexified by making a tangent line from point $(0, 0, 0)$ to its graph, and replacing the part between $(0, 0, 0)$ and the tangent point with the straight tangent line. Let u denote the P_1-coordinate of the tangent point. The condition $\frac{g(u, u/\alpha)}{u} =$

$\frac{\partial g}{\partial P_1}\Big|_{P_1=u}$ should be satisfied as the tangent line passes point $(0,0,0)$, leading to the equation

$$\frac{u}{a_1+u} = \ln\left(\frac{a_1+u}{a_1+a_2\alpha}\right) \tag{4}$$

which implicitly defines u as a function of α. Consequently, for any $\alpha \in (0,+\infty)$, we can compute the coordinates of the corresponding tangent point, and convexify the function g with respect to the direction specified by α. In Fig. 1, we illustrate the power-rate surface, which is the graph of $g(P_1,P_2)$, with $a_1 = a_2 = 1$, and the line of tangent points. The tangent lines connecting the origin and the tangent points yield a *tangent region*. Keeping the part of the original surface beyond the tangent region unchanged, we construct a new power-rate surface, which we will show in the following to be concave.

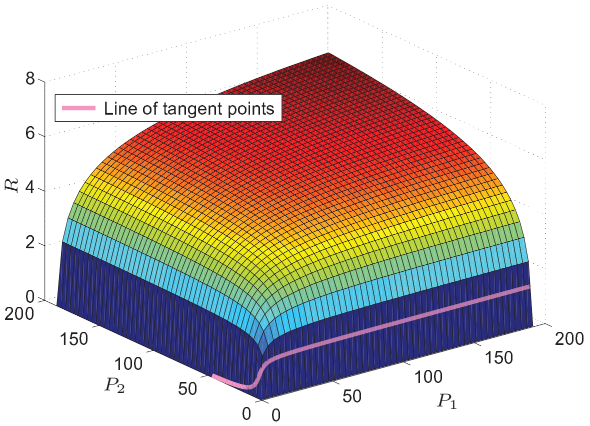


Fig. 1. Power-rate surface and the line of tangent points

We consider two points $X(\alpha_1 x, x)$ and $Y(\alpha_2 y, y)$. The point resulting from the time-sharing of X and Y is given by

$$Z\left(\beta\alpha_1 x + (1-\beta)\alpha_2 y, \beta x + (1-\beta)y\right), \quad 0 \le \beta \le 1,$$

$$\text{with} \quad \alpha_Z = \frac{\beta\alpha_1 x + (1-\beta)\alpha_2 y}{\beta x + (1-\beta)y}.$$

The data rate achieved at point Z, as suggested by the reconstructed power-rate surface, can be computed as

$$R_Z = \frac{1}{\ln 2}\cdot\frac{\beta\alpha_1 x + (1-\beta)\alpha_2 y}{a_1 + u(\alpha_Z)} \tag{5}$$

when Z is in the tangent region; otherwise R_Z can be directly obtained from evaluating the function g at Z. In the former case, the concavity of the surface can be proved by showing that $\frac{\partial^2 R_Z}{\partial \beta^2} < 0$, whereas in the latter, we can show that at any point beyond the tangent region, the rate function g is jointly concave in P_1 and P_2. Note that by taking this approach, we do not need to discuss separately, the cases that the points X and Y are inside or out of the tangent region.

From the definition of α_Z and (4), it can be computed that

$$\frac{\mathrm{d}\alpha_Z}{\mathrm{d}\beta} = \frac{(\alpha_1-\alpha_2)xy}{(\beta x+(1-\beta)y)^2}, \quad \frac{\mathrm{d}u}{\mathrm{d}\alpha} = a_2\cdot\frac{(a_1+u)^2}{u(a_1+a_2\alpha_Z)}.$$

By using the chain rule we can obtain $\frac{\mathrm{d}u(\alpha_Z)}{\mathrm{d}\beta}$. It then follows

$$\frac{\mathrm{d}R_Z}{\mathrm{d}\beta} = \frac{1}{\ln 2}\left[\frac{\alpha_1 x - \alpha_2 y}{a_1+u(\alpha_Z)} - \frac{a_2(\alpha_1-\alpha_2)xy\cdot\alpha_Z}{(\beta x+(1-\beta)y)\,u(a_1+a_2\alpha_Z)}\right]$$

$$\frac{\mathrm{d}^2 R_Z}{\mathrm{d}\beta^2} = -\frac{1}{\ln 2}\cdot\frac{a_2(\alpha_1-\alpha_2)^2(xy)^2}{(\beta x+(1-\beta)y)^3\,u(\alpha_Z)\,(a_1+a_2\alpha_Z)} \times\left[2 - \frac{a_2\alpha_Z}{a_1+a_2\alpha_Z}\cdot\frac{(a_1+u(\alpha_Z))^2+u^2(\alpha_Z)}{u^2(\alpha_Z)}\right]$$

From (4) which characterizes the tangent points, we have

$$\frac{u}{a_1+u} = \ln\left(1+\frac{u-a_2\alpha}{a_1+a_2\alpha}\right) < \frac{u-a_2\alpha}{a_1+a_2\alpha},$$

leading to the inequality $a_2\alpha < \frac{u^2}{2u+a_1}$. Consequently, we see that R_Z is strictly concave in β since

$$\frac{a_2\alpha_Z}{a_1+a_2\alpha_Z}\cdot\frac{(a_1+u(\alpha_Z))^2+u^2(\alpha_Z)}{u^2(\alpha_Z)} < \frac{u^2(\alpha_Z)\left(a_1^2+2a_1u(\alpha_Z)+2u^2(\alpha_Z)\right)}{\left(a_1^2+2a_1u(\alpha_Z)+u^2(\alpha_Z)\right)u^2(\alpha_Z)} < 2.$$

Suppose point Q with P_1-coordinate p and ratio α is outside the tangent region, *i.e.*, $p > u(\alpha)$. Since g is strictly concave in P_1 along the fixed direction specified by α, the R-intercept of the tangent line at Q must be positive, yielding

$$g(p,p/\alpha) - p\cdot\frac{\partial g}{\partial P_1}\Big|_{P_1=p} > 0.$$

Similarly as before, we apply the relation $\ln(1+x) < x$ for $x > 0$ and obtain that $\alpha < \frac{p^2}{a_2(2p+a_1)}$, which gives

$$a_1a_2 + 2p(a_2 - \frac{p}{\alpha}) < a_1a_2 + 2a_2p - 2p^2\cdot\frac{a_2(2p+a_1)}{p^2} = -a_1a_2 - 2a_2p < 0.$$

This means, the Hessian matrix $\boldsymbol{H}(g)$ is negative definite at any point outside the tangent region, suggesting the strict concavity of the surface over this region.

With the above proof, we assure that the reconstructed power-rate surface, which is obtained by making tangent lines from the origin to the surface determined by g for all directions $\alpha \in (0,+\infty)$, is overall concave. Based on this, the optimal solution to (3) can be achieved via Algorithm 1. Optimality of the algorithm can be understood as the straight line connecting any two points on the power-rate surface would lie below, or coincide with, the surface due to its concavity. If the two points have different α values, then the connecting

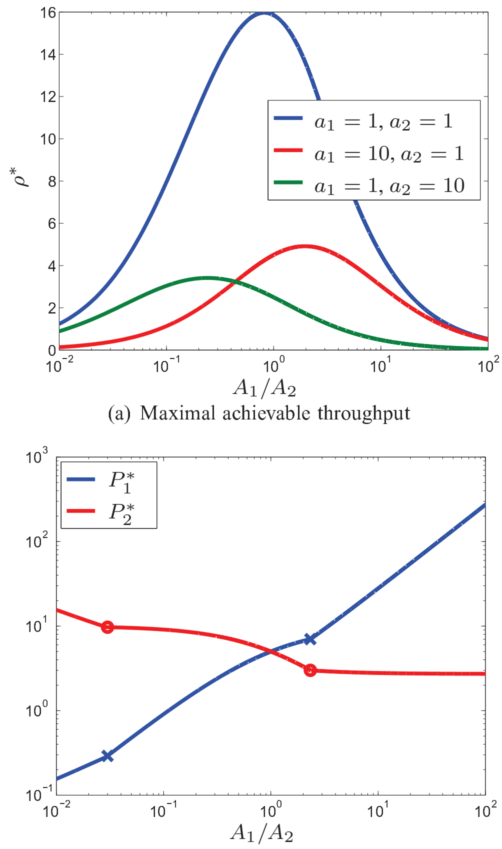
Algorithm 1 Obtaining the optimal operation power

$\alpha \leftarrow A_1/A_2$, compute $u(\alpha)$ by solving (4)
if $A_1/T > u(\alpha)$ **then**
  Tx and Rx operate actively for the whole time interval $[0, T]$ with $P_1^* = A_1/T$, $P_2^* = A_2/T$
else
  Tx and Rx operate actively for a time period of $A_1/u(\alpha)$ with power $P_1^* = u(\alpha)$, $P_2^* = u(\alpha)/\alpha$, and then turn into sleeping mode
end if

straight line lies below the surface, meaning that any way of time-sharing of the two operation points results in a smaller average data rate compared to using the average power constantly. The argument also applies when the two points have the same α, but at least one of them is outside the tangent region. When both points have the same α and are in the tangent region, then any point on the connecting line is also on the surface, and it can be achieved via time-sharing of the sleeping mode and the operation mode specified by the tangent point in the direction given by α. The above reasoning leads to the conclusion that both transmitter and receiver should consume their power constantly, and the sleeping mode comes into play when the desired power is smaller than that of the tangent point in the corresponding direction. Since the rate function g increases monotonically in P_1 and P_2, all available energy should be consumed by time T in order to maximize the throughput, and we can therefore determine the optimal transmit and receive strategy as described in Algorithm 1. The optimal transmit power p_{tx}^* and the optimal ADC resolution b^* can then be directly obtained from P_1^* and P_2^*.

5. NUMERICAL RESULTS

We set up some simple test scenarios to see how the maximal achievable throughput ρ^* and the optimal solution change with respect to the system parameters a_1, a_2, A_1 and A_2. We take $T = 10$, fix the total available energy at the transmitter and the receiver to 100, *i.e.*, $A_1 + A_2 = 100$, and vary the ratio between A_1 and A_2 from 0.01 to 100. The maximal throughput achieved with different a_1 and a_2 are shown in Fig. 2(a), and the optimal power values during the active mode for the case $a_1 = a_2 = 1$ are shown in Fig. 2(b). The peaks in the ρ^* curves indicate the optimal energy allocation between the transmitter and the receiver. It turns out that the positions of the peaks change not only with the parameters a_1 and a_2, but also with the sum energy $A_1 + A_2$. The crosses and circles as in Fig. 2(b) indicate the turning points at which the optimal transmit and receive strategies change. When A_1/A_2 is very small, the transmitter does not have enough energy to support the power $u(A_1/A_2)$ for the whole time interval, causing a sleeping period necessary. Similarly, a sleeping period is required when A_1/A_2 is very large. In the medium region, the transmitter and the receiver are active for the whole time interval, and therefore P_1^* increases linearly with A_1. Since A_1 is concave in A_1/A_2 when $A_1 + A_2$ is fixed, so should P_1^* be, which is exactly what can be seen in the figure.



(a) Maximal achievable throughput

(b) Optimal power consumption, $a_1 = a_2 = 1$

Fig. 2. Optimization results for a system with fixed total energy $A_1 + A_2 = 100$, $T = 10$

6. CONCLUSION

We investigate the throughput maximization problem over a given time interval for the point-to-point communication where the transmitter and the receiver have individual energy constraints, and assume perfect knowledge of all relevant system parameters. An information theoretic model is established, and the problem is formulated as a standard control problem with transmit power and ADC resolution as its two control variables. The essential proposal in finding the jointly optimal transmit and receive strategy is to transform the optimization variables to the power consumptions at the transmitter and the receiver, and then convexify the power-rate surface. The obtained result lays the basis for optimizations in more complicated scenarios, *e.g.*, with energy harvesting transceivers, and can be applied for improving the performance of energy-constrained communication systems.

7. REFERENCES

[1] Y. Chen, S. Zhang, S. Xu, and G. Y. Li, "Fundamental trade-offs on green wireless networks," *IEEE Communications Magazine*, vol. 49, pp. 30–37, June 2011.

[2] Q. Bai and J. A. Nossek, "On energy efficient cross-layer assisted resource allocation in multiuser multicarrier systems," in *Proceeding of IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, September 2009.

[3] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Transactions on Wireless Communications*, vol. 11, pp. 1180–1189, March 2012.

[4] Q. Bai and J. A. Nossek, "Throughput maximization for energy harvesting nodes with generalized circuit power modelling," in *Proceeding of 13th IEEE International Workshop on Signal Processing Advances in Wireless Communications*, June 2012.

[5] A. Mezghani and J. A. Nossek, "Modeling and minimization of transceiver power consumption in wireless networks," in *Proceeding of 2011 International ITG Workshop on Smart Antennas (WSA)*, February 2011.

[6] S. Cui and A. J. Goldsmith, "Energy-constrained modulation optimization," *IEEE Transactions on Wireless Communications*, vol. 4, pp. 2349–2360, September 2005.