

Binary Energy Harvesting Channel with Finite Energy Storage

Kaya Tutuncuoglu¹, Omur Ozel², Aylin Yener¹, and Sennur Ulukus²

¹Department of Electrical Engineering, Pennsylvania State University, University Park, PA 16802

²Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742

Abstract—We consider the capacity of an energy harvesting communication channel with a finite-sized battery. As an abstraction of this problem, we consider a system where energy arrives at the encoder in multiples of a fixed quantity, and the physical layer is modeled accordingly as a finite discrete alphabet channel based on this fixed quantity. Further, for tractability, we consider the case of binary energy arrivals into a unit-capacity battery over a noiseless binary channel. Viewing the available energy as state, this is a state-dependent channel with causal state information available only at the transmitter. Further, the state is correlated over time and the channel inputs modify the future states. We show that this channel is equivalent to an additive geometric-noise timing channel with causal information of the noise available at the transmitter. We provide a single-letter capacity expression involving an auxiliary random variable, and evaluate this expression with certain auxiliary random variable selection, which resembles noise concentration and lattice-type coding in the timing channel. We evaluate the achievable rates by the proposed auxiliary selection and extend our results to noiseless ternary channels.

I. INTRODUCTION

We consider an energy harvesting communication system, where energy needed for communication is harvested by the transmitter during the course of communication; see Fig. 1. We study the capacity of such communication channels. Capacity of such channels has been identified in previous work for two extreme cases: When the battery-size is unlimited, [1] showed that the capacity is equal to the capacity of the same system with an average power constraint equal to the average recharge rate; an unlimited-sized battery averages out the fluctuations in energy arrivals, and reduces energy causality constraints at all channel uses into a single average power constraint. At the other extreme, when the battery-size is zero, the system becomes a stochastic amplitude-constrained channel. Reference [2] has found the capacity of this system for an AWGN channel with independent and identically distributed (i.i.d.) energy arrivals, by modeling it as a state-dependent channel with causal state information available at the transmitter, and showed that capacity is achieved by using a Shannon strategy [3]; in particular, the capacity achieving input distribution is discrete as in the case of [4]. The capacity for the case of a finite-sized battery is unknown, and is the problem considered in this paper.

This work was supported by NSF Grants CNS 09-64364 and CNS 09-64632.

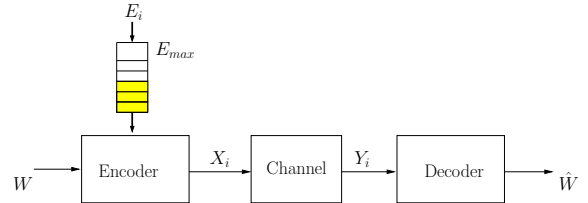


Fig. 1. Energy harvesting communication system with a finite-sized battery.

This problem has the following characteristics: The battery can be viewed as an energy queue where energy arrives as a stochastic process over time; for tractability, we assume an i.i.d. energy arrival process. The codebook used to transmit messages acts as a server to this energy queue, and determines its dynamics. The energy available at the energy queue is the state, which determines the set of feasible symbols that can be transmitted. In particular, at each channel use, the transmitted symbols are instantaneously amplitude-constrained to the (square root of the) available energy in the battery. This state is naturally causally known to the transmitter, but unknown to the receiver. This state is correlated over time, even when energy arrivals are i.i.d. In addition, transmitter's own actions further affect the future of the state.

From [4] and [2], we know that, when the channel inputs are constant amplitude-constrained, or i.i.d. stochastic amplitude-constrained, over a Gaussian channel, the optimum input distributions are discrete. However, these discrete mass points are arbitrary real numbers, and it is hard to track the dynamics of the energy queue, if it is served with codebooks generated by arbitrary real mass points. For a tractable abstraction of the system, we model energy arrivals as multiples of a fixed quantity, and correspondingly, consider a physical layer which has a discrete alphabet based on this fixed quantity. For further analytical tractability, we assume that the physical layer is a noiseless binary channel, energy arrivals are binary, and the battery is unit-sized. This abstraction is reminiscent of the one in [5]. We will see that, even in this simple model, unavailability of the battery state to the receiver, memory of the state in time, and the fact that the state evolves based on the previous channel inputs, render the problem challenging. The fact that channel inputs affect future states is reminiscent of action dependent channels in [6]. While Shannon strategy, which is optimal in the zero-battery case in [2], yields achievable rates for the finite-battery case, the transmitter may utilize the

memory in the battery state to achieve higher rates.

We first present the system model which is a state-dependent channel, and derive achievable rates based on certain Shannon strategies. We next show that our system may equivalently be modeled as a timing channel [7], where information is transmitted by timings between 1s, as opposed to the actual places of 1s and 0s. This converts our problem into a timing channel with additive geometric noise (service time), where the service time is causally known to the transmitter. We combine [7] and [3] to find a single-letter capacity expression for the capacity of this equivalent channel. Shannon strategy [3] is optimal here, because the state is i.i.d. in time. The capacity expression involves an auxiliary random variable, and its optimization is difficult. For this reason, we determine an achievable rate based on a certain selection of this auxiliary random variable. This selection resembles the concentration idea in [8], and may be interpreted as a lattice-type coding for the timing channel. We evaluate the achievable rate of the proposed selection and compare it with rates achieved by Shannon strategies in the original channel. Finally, we extend our results to the case of a noiseless ternary channel by means of its equivalence to a timing channel with extra information sent inside the symbols in addition to the timings of the symbols.

II. THE CHANNEL MODEL

Consider a noiseless binary channel with an energy harvesting transmitter. The battery in the transmitter can store at most one unit of energy. The channel input symbols, namely 0 and 1, have zero and one unit energy cost, respectively. When channel input X_i is transmitted in the i th channel use, the receiver gets $Y_i = X_i$. At each channel use, the transmitter both harvests energy and also transmits a symbol. The order of harvesting and transmission in a channel use is as follows: At each channel use, S_i denotes the energy available in the battery. The transmitter observes the available battery energy S_i and transmits a symbol X_i . This symbol is constrained by the battery energy: $X_i \in \{0, 1\}$ when $S_i = 1$ and $X_i = 0$ when $S_i = 0$. After sending the symbol, the transmitter harvests energy. Energy arrivals (harvesting) is modeled as an i.i.d. Bernoulli process with $E_i \in \{0, 1\}$ and $\Pr[E_i = 1] = q$. Incoming energy E_i is first stored in the battery, if there is space, before it is used for transmission. Since the battery has one-unit energy storage capacity, energies may overflow and get wasted. The battery state is updated as

$$S_{i+1} = \min\{S_i - X_i + E_i, 1\}. \quad (1)$$

We note that S_i is a state for this channel that is causally available to the transmitter but not to the receiver. Moreover, perfect channel output feedback is available to the transmitter as the channel is noiseless. Note that even though the channel is noiseless, uncertainty of the battery energy at the transmitter side makes it challenging for the receiver to decode the messages of the transmitter. A crucial property of this channel model is the fact that the state of this channel S_i has memory and the evolution of the state process is affected by the channel

input X_i . Even though the channel state is not i.i.d. in time, Shannon strategy [3] provides some achievable rates, which we next specify for our channel model.

A. Achievable Rates with Shannon Strategies

Consider the Shannon strategy $(0, 0)$ and $(0, 1)$ as in [3] where the first and second entries denote the actual channel input when battery state S is 0 and 1, respectively. We represent $(0, 0), (0, 1)$ as $U = 0, 1$ for simplicity, and let U be i.i.d. with $\Pr[U = 1] = p$. The state transition probabilities are $\Pr[S_{i+1} = 1|S_i = 0] = q$ and $\Pr[S_{i+1} = 0|S_i = 1] = p(1 - q)$, yielding the stationary probability $\Pr[S = 1] = \frac{q}{p+q-pq}$. Note that the battery state is ergodic and the receiver can use the stationary probability of the battery state for joint typicality decoding. In this case, the channel takes the form

$$p(y|u) = \Pr[S = 1]\delta(y - u) + (1 - \Pr[S = 1])\delta(y) \quad (2)$$

where $\delta(\cdot)$ is a unit impulse at 0 and $u \in \{0, 1\}$. Here, $I(U; Y)$ is an achievable rate due to [3] and the fact that the battery energy state is an ergodic process, see also [5]. We refer to this as the naïve i.i.d. Shannon strategy (NIID). The best achievable rate by this strategy is

$$R_{NIID} = \max_{p \in [0, 1]} H\left(\frac{pq}{p+q-pq}\right) - pH\left(\frac{q}{p+q-pq}\right) \quad (3)$$

where $H(\cdot)$ is the binary entropy function.

The naïve approach yields a sub-optimal rate since the memory of the system is not exploited by the decoder. Indeed, the decoder can better exploit the memory of the system by using the n -letter joint probability $p(u^n, y^n)$. The maximum rate that can be achieved by using this scheme is given by

$$R_{OIID} = \max_{p \in [0, 1]} \lim_{n \rightarrow \infty} \frac{1}{n} I(U^n; Y^n). \quad (4)$$

We will call the scheme that achieves R_{OIID} the optimal i.i.d. Shannon strategy (OIID). Here, the limit $\lim_{n \rightarrow \infty} \frac{1}{n} I(U^n; Y^n)$ can be calculated numerically using the method in [9]. Notice that when the input u_i is i.i.d., the joint probability $p(y_i, u_i, s_{i+1}|s_i)$ can be expressed in the form

$$p(y_i, u_i, s_{i+1}|s_i) = p(y_i, s_{i+1}|u_i, s_i)p(u_i) \quad (5)$$

where $p(y_i, s_{i+1}|u_i, s_i)$ is independent of i . This allows using the iterative method in [9] to calculate the achievable rate for a given input distribution, which is then optimized with respect to the i.i.d. probability distribution to yield (4).

III. EQUIVALENT TIMING CHANNEL

Since the channel input is binary, the encoding and decoding can be performed over the number of channel uses between two 1s. Let us define $T_n \in \{1, 2, \dots\}$ as the number of channel uses between the $n - 1$ st transmitted 1 and the n th transmitted 1. After a 1 is transmitted in the i th channel use, in view of (1), available energy in the battery drops to zero in the $i + 1$ st channel use unless $E_i = 1$ as the battery can store at most one unit of energy. The node cannot transmit another 1 until the

next energy arrival. Define the idle time $Z_n \in \{0, 1, 2, \dots\}$ as the number of channel uses between the $n - 1$ st transmission of a 1 and the next energy arrival after that. This representation yields the following timing channel:

$$T_n = V_n + Z_n \quad (6)$$

where $V_n \in \{1, 2, \dots\}$ is the number of channel uses the transmitter chooses to wait to transmit a 1 after the first energy arrival proceeding the $n - 1$ st transmission of a 1. T_n , V_n and Z_n are depicted in Fig. 2, where circles represent energy arrivals and triangles represent transmissions of a 1. Note that Z_n is independent of any action of the transmitter; it is the number of i.i.d. Bernoulli trials until the first energy arrives after the $n - 1$ st transmission of a 1, and is a geometric random variable with parameter q . As the original channel is noiseless, when the transmitter puts a 1 to the channel, the receiver observes it and calculates T_n perfectly. Therefore, the transmitter observes Z_n and decides on V_n , but all that the receiver observes is T_n . Even though the physical channel is noiseless, the uncertainty in Z_n creates difficulty for the receiver. In fact, in terms of a timing channel, the overall channel is an additive noisy timing channel with a geometric noise, which is known causally to the transmitter.

It is worthwhile to mention that when the channel symbols are considered as the time difference between two 1s, the resulting model becomes similar to a noiseless channel with symbols of varying durations as in Shannon's original work [10]. However, in our problem, the symbol durations are affected by a random energy arrival process, while in [10] the symbol durations are fixed. Therefore, the problem in [10] is to *pack* as many symbols as possible within a given block length, while our problem is also concerned with the randomness introduced by energy harvesting.

In the timing channel, the transmitter has the feedback of previous channel outputs T^{n-1} and causal knowledge of the idle time Z_n before deciding V_n . In fact, the transmitter observes energy arrivals E_i even when the battery is full. We note that the timing channel in this section and the classical state-dependent channel in Section II are equivalent in the sense that they have the same capacity, using the definition of capacity of a timing channel in [7]. This is due to the fact that the encoders and decoders of these channels have different representations of the same object, and the rate and capacity of these channels are properly defined taking into account the time cost of the codewords, i.e., the average number of channel uses needed, following [7]. We state this fact as a lemma.

Lemma 1 *The timing channel capacity with additive causally known state C_T and the classical state-dependent channel capacity C are equal, i.e., $C = C_T$.*

IV. THE CAPACITY

The capacity of a discrete memoryless channel with causal side information is [3], [11, Thm. 7.2]

$$C = \max_{p(u), v(u, z)} I(U; T), \quad (7)$$

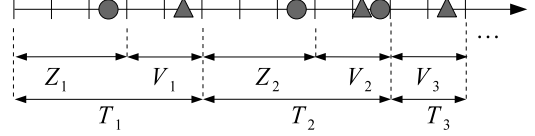


Fig. 2. Graphical representation of T_n , V_n and Z_n . An energy arrival of $E_i = 1$ is represented with a circle and a channel input of $X_i = 1$ is represented with a triangle. Note that $Z_3 = 0$ as a transmission of a 1 and an energy harvesting occur in the same channel use.

where T is the channel output. The maximization requires finding the optimal distribution for the auxiliary variable U and the mapping $v(\cdot, \cdot)$ which can be taken as deterministic without losing optimality [11, Thm. 7.2]. Cardinality bound on U is $|\mathcal{U}| \leq \min\{(|\mathcal{V}|-1)|\mathcal{Z}|+1, |\mathcal{T}|\}$. We combine techniques from [3] and [7] to prove the following theorem.

Theorem 1 *The capacity of the timing channel with additive causally known state, C_T , is:*

$$C_T = \max_{p(u), v(u, z)} \frac{I(U; T)}{\mathbb{E}[T]} \quad (8)$$

Proof: Let $W \in \{1, \dots, M\}$ denote the message, which is uniform and let n denote the maximum number of channel uses on average (over the message W and the energy arrivals E_i) to send any message $W = w$. We note that $\sum_{i=1}^m \mathbb{E}[T_i] \leq n$ where the expectation is over the energy arrival sequence E_i and the message W .

Define $U_i = (W, T^{i-1})$. Note that even though T^i depends on W , since E_i is an i.i.d. random process, Z_i is independent of W and T^{i-1} and hence of U_i . We have the following sequence of inequalities:

$$\log(M) - H(W|T^m) = H(W) - H(W|T^m) \quad (9)$$

$$= I(W; T^m) \quad (10)$$

$$= \sum_{i=1}^m I(W; T_i | T^{i-1}) \quad (11)$$

$$\leq \sum_{i=1}^m I(W, T^{i-1}; T_i) \quad (12)$$

$$= \sum_{i=1}^m I(U_i; T_i) \quad (13)$$

$$\leq \frac{n}{\sum_{i=1}^m \mathbb{E}[T_i]} \sum_{i=1}^m I(U_i; T_i) \quad (14)$$

$$\leq n \sup_U \frac{I(U; T)}{\mathbb{E}[T]} = nC_T \quad (15)$$

where (14) follows from the fact that $n \geq \sum_{i=1}^m \mathbb{E}[T_i]$, and the inequality in (15) holds due to the fact that U_i are independent of Z_i and $\frac{\sum_{i=1}^m a_i}{\sum_{i=1}^m b_i} \leq \max_i \frac{a_i}{b_i}$ for $a_i, b_i > 0$. Whenever the probability of error goes to zero as $m \rightarrow \infty$, $H(W|T^m) \rightarrow 0$ by Fano's inequality and therefore, $\frac{\log(M)}{n} = R \leq C_T$, which completes the converse proof.

For the direct part, we simply use an encoding scheme that does not use the available feedback information T^{i-1} and the

achievability of $R = \frac{I(U;T)}{\mathbb{E}[Z]}$ follows from [3], [7] and [11]. ■

We note that the capacity expression in (8) is similar to the capacity of a telephone signaling channel [7], augmented with Shannon strategy to utilize the causal state information at the transmitter. It is worth remarking that the transmitter, in fact, has more freedom in that, it is free to update its decision after observing Z^i ; however, as the converse proof of Theorem 1 shows, this does not yield a higher achievable rate.

For the timing channel, the input, output and state alphabets have infinite cardinalities. Therefore, the cardinality bound on U is infinite and the solution to the maximization problem in (8) is difficult. While the capacity achieving auxiliary selection is still an open problem, we continue our capacity analysis by resorting to a family of auxiliary random variables with finite cardinality, in the next section. We will also devise an upper bound in Section VI in closed form which will help us assess the performance of the proposed finite cardinality auxiliary random variables.

V. FINITE CARDINALITY AUXILIARY VARIABLES

In this section, we propose to use a finite cardinality auxiliary random variable selection parameterized by a variable N to be optimized later. Let U have a probability mass function (pmf) $p(u)$ over a support set $U \in \{0, 1, \dots, N-1\}$, where N is the cardinality of U . Moreover, we fix the mapping from U and Z to V as

$$V = (U - Z \bmod N) + 1 \quad (16)$$

The output of the timing channel is $T = V + Z$. Let $T' = T - 1 \bmod N$. Note that $T' = U$. We have

$$\max_{p(u)} \frac{I(U;T)}{\mathbb{E}[V+Z]} \leq \max_{p(u)} \frac{H(U)}{\mathbb{E}[V+Z]} \quad (17)$$

$$= \max_{p(u)} \frac{I(U;T')}{\mathbb{E}[V+Z]} \quad (18)$$

where (17) is due to the non-negativity of entropy. Therefore, for fixed N , this scheme achieves

$$R_A^{(N)} = \max_{p(u)} \frac{H(U)}{\mathbb{E}[V+Z]} \quad (19)$$

We further maximize $R_A^{(N)}$ over N and obtain the best achievable rate by this scheme: $R_A = \max_N R_A^{(N)}$.

In the original binary energy harvesting channel, this strategy corresponds to channel uses being divided into frames of length N and indexed in base N . The transmitter constructs a codebook with symbols $U \in \{0, 1, \dots, N-1\}$ and conveys each symbol by transmitting a 1 at the earliest channel use indexed with this symbol as illustrated in Fig. 3 for $N = 5$. This coding strategy can be interpreted as concentrating the effective state Z to multiples of a *frame length* N as in [8]. The receiver decodes intended codeword U with $U = T - 1 \bmod N$, which treats Z as if it is eliminated with the modulo operation. Therefore, this scheme is a variation of the concentration scheme of Willems in [8] for the timing channel. It can also be interpreted as lattice-coding the timing via modulo operation.

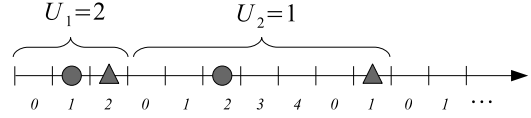


Fig. 3. Coding scheme: each message symbol U_i is conveyed by transmitting a 1 at the earliest channel use possible with index equal to U_i . Here, $N = 5$.

VI. AN UPPER BOUND

In order to get an upper bound, we provide the channel state of the timing channel Z to the receiver.

$$C_{UB} = \max_{p(v)} \frac{H(V)}{\mathbb{E}[V] + \mathbb{E}[Z]} \quad (20)$$

$$= \max_{\mu \geq 0} \frac{1}{\mu + \mathbb{E}[Z]} \max_{\mathbb{E}[V] \leq \mu} H(V) \quad (21)$$

Here, the inner problem requires finding the entropy maximizing probability distribution with an expectation constraint of μ over a discrete positive support, which is solved by a geometrically distributed V with parameter $\frac{1}{\mu}$. Thus, $H(V) = \frac{H(p)}{p}$ where $H(p)$ is the binary entropy function. Substituting in (21) and noting that $Z \in \{0, 1, \dots\}$ is geometrically distributed with parameter q , the upper bound is found as

$$C_{UB} = \max_{p \geq 0} \frac{H(p)/p}{\frac{1}{p} + \frac{1-q}{q}} = \max_{p \geq 0} \frac{qH(p)}{q + p(1-q)} \quad (22)$$

VII. NUMERICAL RESULTS

In this section, we provide simple numerical results for the achievable scheme proposed in Section V, how it compares to the i.i.d. Shannon strategies in the classical state-dependent channel given in Section II-A, and the upper bound in Section VI. We also provide comparisons with the capacities in the extreme cases of batteries with zero and unlimited energy storage (still operating over a noiseless binary channel). Next, we briefly mention the capacities under these extreme cases.

A. Capacity with Zero Energy Storage

Consider the model with no energy storage, that is, $X_i = 1$ is allowed only if $E_i = 1$, which is known causally by the transmitter. In this case, the state of the system, E_i , is i.i.d. The capacity of this channel is achieved by a Shannon strategy as in [2], which chooses among $X_i \in \{0, 1\}$ with $\Pr[X_i = 1 | E_i = 1] = p$ whenever $E_i = 1$, and is forced to transmit $X_i = 0$ whenever $E_i = 0$. The capacity is:

$$C_{ZS} = \max_p H(pq) - pH(q) \quad (23)$$

where $H(p)$ is the binary entropy function.

B. Capacity with Infinite Energy Storage

When there is an unlimited battery, as the average energy arrival rate is q , the capacity is the maximum of $H(p)$ over $p \leq q$, which is achieved by using the save-and-transmit scheme

in [1]. When evaluated, the capacity becomes:

$$C_{IS} = \begin{cases} H(q), & q \leq \frac{1}{2} \\ 1, & q > \frac{1}{2} \end{cases} \quad (24)$$

We are now ready to compare the capacities and achievable rates, which are shown in Fig. 4 as a function of the energy arrival probability q . We first remark that the achievable rate by the specified family of finite cardinality auxiliary variables in Section V is higher than the best achievable rate by the naïve i.i.d. Shannon strategy in the classical state-dependent channel where the receiver uses the stationary distribution of the battery energy. This achievable rate also outperforms the optimum i.i.d. Shannon strategy for low arrival probabilities. We also remark that in comparison to zero and infinite energy storage capacities, a battery of unit size provides a significant rate improvement.

For the achievable rate proposed in Section V, a large N provides a larger cardinality for U , sending more information with a single unit of energy. However, as N increases, each symbol takes more time and more harvested energy is potentially wasted. We observe numerically that for low harvest rates the optimal N decreases with increasing q so that more of the scarcely available energy is utilized. This trend is reversed for high harvest rates, e.g., $q > 0.7$, as wasting energy becomes less of a concern then.

VIII. EXTENSION TO NOISELESS TERNARY CHANNEL

In this section, we consider an extension with a third channel input -1 , which also requires one unit of energy. Specifically, we represent the ternary output vector Y^n with $T^\ell \in \{1, 2, \dots\}^\ell$, defined as the number of zeros between the $\ell - 1$ st and the ℓ th nonzero output, and $S^\ell \in \{-1, 1\}^\ell$, defined as the sign of the ℓ th nonzero output. Once again, note that Y^n can be perfectly reconstructed from (T^ℓ, S^ℓ) and vice versa. With this representation, the equivalent model consists of two parallel channels, $T_n = V_n + Z_n$ and $S_n = R_n$, where R_n is the sign of the ℓ th nonzero input, conveyed noise-free to the receiver. The equivalence to the original channel can be established by an extension of Lemma 1. Noting that the binary side channel and the timing channel are used equally many times, the information conveyed with each of their use is $I(U; T) + 1$ where U is the Shannon strategy. This is similar to queues with information bearing packets, since each nonzero symbol can be considered as a packet bearing 1 bit of information. Hence, as in [7, Section IV], coding over symbol energy and over symbols of the same energy are done independently and this yields the capacity.

Theorem 2 *The capacity with ternary symbols is:*

$$C_{ter} = \max_{p(u), v(u, s)} \frac{I(U; T) + 1}{\mathbb{E}[T]} \quad (25)$$

IX. CONCLUSION

In this paper, we considered an energy harvesting transmitter with a unit-capacity battery communicating over a noiseless

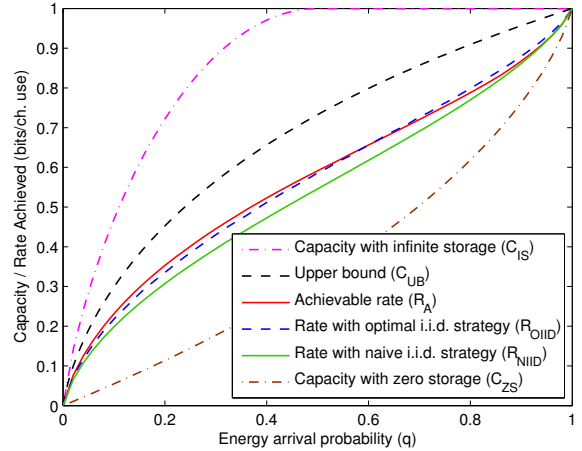


Fig. 4. Upper bound, achievable rate, no battery capacity, infinite battery capacity, and rates by naïve and optimum i.i.d. Shannon strategies.

binary channel. Classical view of this channel is a state-dependent channel with causal state information at the transmitter where the evolution of the state is affected by the channel input. This coupling of the state and the channel input creates a challenge. We show that this channel is equivalent to a timing channel with feedback and causal state information at the transmitter, and find a single-letter capacity expression for its capacity. As evaluation of this capacity expression is difficult, we resort to finite cardinality auxiliary random variables and numerically study the achievable rates. We compare these achievable rates with rates achievable by naïve and optimum i.i.d. Shannon strategies in the original channel. Finally, we extend our results to a noiseless ternary channel where information is transmitted via timing and contents of symbols.

REFERENCES

- [1] O. Ozel and S. Ulukus, "Achieving AWGN capacity under stochastic energy harvesting," *IEEE Trans. on Inform. Theory*, vol. 58, no. 10, pp. 6471–6483, Oct. 2012.
- [2] —, "AWGN channel under time-varying amplitude constraints with causal information at the transmitter," in *Asilomar Conf.*, Nov. 2011.
- [3] C. E. Shannon, "Channels with side information at the transmitter," *IBM journal of Research and Development*, vol. 2, no. 4, pp. 289–293, 1958.
- [4] J. G. Smith, "The information capacity of amplitude and variance-constrained scalar Gaussian channels," *Information and Control*, vol. 18, pp. 203–219, April 1971.
- [5] P. Popovski, A. Fouladgar, and O. Simeone, "Interactive joint transfer of energy and information," *IEEE Trans. on Commun.*, submitted, Sep. 2012. Available at arXiv:1209.6367.
- [6] T. Weissman, "Capacity of channels with action-dependent states," *IEEE Trans. on Inform. Theory*, vol. 56, no. 11, pp. 5396–5411, Nov. 2010.
- [7] V. Anantharam and S. Verdú, "Bits through queues," *IEEE Trans. on Inform. Theory*, vol. 42, no. 1, pp. 4–18, Jan. 1996.
- [8] F. M. J. Willems, "Signaling for the Gaussian channel with side information at the transmitter," in *IEEE ISIT*, Jun. 2000.
- [9] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. on Inform. Theory*, vol. 52, no. 8, pp. 3498–3508, August 2006.
- [10] C. E. Shannon, "A mathematical theory of communication," *Bell System Tech. Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [11] A. El Gamal and Y. H. Kim, *Network information theory*. Cambridge University Press, 2011.