On the Capacity of a Communication System with Energy Harvesting and a Limited Battery

Wei Mao and Babak Hassibi
Department of Electrical Engineering
California Institute of Technology, Pasadena, CA 91125
Email: {wmao, hassibi}@caltech.edu

***Abstract*—We consider the problem of determining the capacity of an energy-harvesting transmitter with finite battery communicating over a discrete memoryless channel. When the battery is unlimited, or zero, the capacity has been determined, but it remains unknown for a finite non-zero battery. In this paper we assume that the harvested energy at each time, the total battery storage, and the transmitter signal energy at each time can be quantized to the same unit (i.e., the same energy interval). Under this assumption, we show that the capacity can be described using the Verdú-Han general framework. If we further assume that the transmitted symbol at each time depends only on the energy currently available, and not on the entire past history of energy harvests and symbols transmitted, then we show that the system reduces to a finite state channel (FSC) with the required ergodic and Markov properties so that lower bounds on the capacity can be readily numerically computed. We conjecture that our numerical bounds are tight. Our numerical results indicate that even the minimal possible battery storage can reap a significant fraction of the infinite battery capacity.**

I. Introduction

In many future wireless systems, such as low-power wireless sensor networks, one may encounter transmitters that harvest and store energy for transmission. Such communication systems have recently been introduced and studied by Ulukus and coworkers [1], [2]. In particular [3] shows that with unlimited battery the entire capacity of an AWGN channel can be achieved. Using Shannon's method [4], [5] tries to analyze the AWGN channel capacity in the zero battery case, however, the proof is incomplete.[1] Nevertheless, treating such a case with a discrete channel is elementary (see Section III). The intermediate case, i.e., the case with a finite nonzero battery, is first considered in [6], where the optimum offline transmission policy for an energy harvesting node is obtained. However, determining the channel capacity in such a case remains open.

In this paper, we assume that the energy harvested at each time, the total battery storage, and the transmitter signal energy at each time can be quantized to the same unit (i.e., the same energy interval). Under this assumption, we show that the capacity can be described using the Verdú-Han general framework [7]. We then make a further assumption that the transmitted symbol at each time instant depends only on the energy currently available, and not on the entire history of harvested energy and transmitted symbols. We conjecture that such a strategy does not lose optimality. With this assumption, and assuming that the energy harvesting process is stationary and ergodic (e.g. i.i.d.), we show that the resulting channel reduces to a finite state channel (FSC). We demonstrate the required ergodicity and asymptotic mean stationarity (AMS) of the underlying processes so that the Shannon-McMillan-Breiman (SMB) theorem can be used to compute bounds on the capacity. The generalized Blahut-Arimoto algorithm [8] is then used to optimized these bounds. Our numerical results suggest that even the minimal battery storage of one unit can reap a significant fraction (around 70%) of the capacity obtained using an unlimited battery.

Further extensions of the work will also be discussed.

[1]Two major problems of the proof in [5] are: (15) cannot be analytically extended to $\mathbb{C}^2$, while (18) cannot be implied by the identity theorem on $\mathbb{C}^2$.

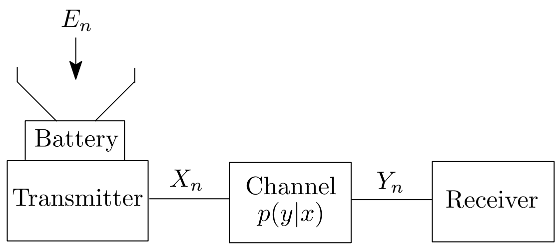

Fig. 1. System model.

II. System Model

Consider a communication system powered by some energy harvesting mechanism with a limited battery, as depicted in Fig 1. At time n the system harvests E_n units of energy, which together with the energy stored in the battery are used for the n-th transmission. The remainder, as much as the battery can hold, will be saved in the battery for future transmissions. At each time the transmitter sends a symbol $X_n \in \mathcal{X}$ over the channel $p(y|x)$, which outputs a symbol $Y_n \in \mathcal{Y}$ to the receiver. The alphabets $\mathcal{X}$ and $\mathcal{Y}$ are assumed to be finite with $\mathcal{X} \subset \mathbb{R}$ or $\mathbb{C}$, and the channel is a discrete memoryless channel (DMC). The difference between our system and an ordinary DMC is, the transmitter is not free to choose every letter in $\mathcal{X}$ during each transmission; instead, at each time n it can only send a symbol X_n that does not demand more than the current available energy. To be precise, let B_n be the energy remaining in the battery after the $(n-1)$-th transmission and let $\overline{B}$ denote the capacity limit of the battery, which is a finite number. The energy constraints on the system are

$$\begin{cases} |X_n|^2 \le B_n + E_n, \\ B_{n+1} = \min\left\{ B_n + E_n - |X_n|^2,\ \overline{B} \right\}. \end{cases} \tag{1}$$

Remark 1: We can also consider an alternative energy storage model as in [6], where the harvested energy is not immediately available to the transmitter: first it charges the battery, then the stored energy is used for transmission. In other words, $\min\left\{B_n + E_n, \overline{B}\right\}$ should be used in place of $B_n + E_n$ for the energy constraints (1). Such a model can also be analyzed using the same framework in this paper, however, the capacity will differ from the current model.

Assume the initial energy B_1 stored in the battery is a random variable and the sequence of arriving energies $\{E_n\}_{n=1}^{\infty}$ is a random process which is independent of B_1. We will assume that the transmitter has causal knowledge of B_1 and the process[2] $\{E_n\}$ but the receiver does not. In this scenario (c.f. [4]) a code of length N is defined as a function

$$f^{(N)} : (m, b_1, e^N) \mapsto x^N$$

such that the encoded vector x^N satisfies the energy constraints (1). Here m is the message to be encoded, b_1 and $e^N = (e_1, \cdots, e_N)$ are the realizations of B_1 and E^N respectively. Let f_n denote the n-th coordinate of the mapping $f^{(N)}$, then f_n should be causal in $\{E_n\}$, i.e., it is only a function of (m, b_1, e^n). The decoder receives the output y^N of the DMC and estimates m.

To simplify the problem, we assume that the energies involved are all quantized with the same interval size, i.e. all B_n, E_n, $|X_n|^2$ and $\overline{B}$ are integral multiples of some common unit of energy Δ_E. Hence without loss of generality we can assume all these quantities are integers. Furthermore, assume the alphabet of E_n is a finite set $\mathcal{E}_H$ of non-negative integers, so B_n can only take values in a finite set $\mathcal{E}_B$.

III. Equivalent Channel without Constraints

In a sense, the constraints (1) introduce memory to the system. The memory does not affect the channel transition probabilities, but limits the input alphabet instead. To capture such (infinite) memory of the channel model, we consider the energy arrival process $\{E_n\}_{n=1}^{\infty}$, together with the initial battery status B_1, as the state sequence of the channel. The states are causally known at the transmitter, but not at the receiver. Then using the approaches in [4], [9], the channel can be converted to an equivalent channel without states but with enlarged input alphabets, as defined below. Note that if $\{E_n\}$ is i.i.d. and $\overline{B} = 0$, then the channel states are i.i.d. and the capacity can be easily obtained as that of a DMC with a larger input alphabet, by the method in [4] with only a slight modification on the definition of the new alphabet (c.f. [5] and Example 1 in Section VII).

Let $\boldsymbol{W} = \{\mathcal{U}^{(N)}, W^N = p(y^N|u^{(N)}), \mathcal{Y}^N\}_{N=1}^{\infty}$ denote the new channel. It is defined by describing the input/output alphabets and the transition probabilities for each block length N, which corresponds to N operations of the original channel from the beginning of transmission. The output alphabet is the same as the original channel, whereas the input is different. We will see that the input alphabet for a block of length N is not the Cartesian product of N copies of the alphabet of a single symbol, so it is denoted by $\mathcal{U}^{(N)}$.

A. Input Symbols

The input symbol for the equivalent channel is denoted by $u^{(N)} = (u_1, \cdots, u_N)$, whose n-th entry is a function

$$u_n : \mathcal{E}_B \times \mathcal{E}_H^n \to \mathcal{X},$$

which can also be viewed as a vector in $\mathcal{X}^{|\mathcal{E}_B|\cdot|\mathcal{E}_H|^n}$. Such functions are sometimes called strategy letters. However, not all such functions are valid candidates for the input: for each state sequence (b_1, e^n), the encoded symbol $x_n = u_n(b_1, e^n)$ needs be compatible with all the previous symbols x^{n-1}, that is, together they should fulfill the energy constraints (1). Specifically, for every $B_1 = b_1$ and $E^N = e^N$, a feasible input vector $u^{(N)}$ of strategy letters needs to satisfy the following requirements:

$$|u_n(b_1, e^n)|^2 \le B_n + e_n, \quad \forall 1 \le n \le N,$$

where $B_n = B_n(u^{n-1}, b_1, e^{n-1})$ is determined recursively by

$$B_{n+1} = \min\left\{ B_n + e_n - |u_n(b_1, e^n)|^2,\ \overline{B} \right\}.$$

Thus the permitted choices of u_n depends not only on the states (b_1, e^n), but also on all previous strategy letters $u^{n-1} = (u_1, \cdots, u_{n-1})$. In other words, let

$$\mathcal{X}(a) := \{x \in \mathcal{X} : |x|^2 \le a\}$$

for $a \ge 0$, then the alphabet of u_n is

$$\mathcal{U}_n(u^{n-1}) = \prod_{b_1, e^n} \mathcal{X}(B_n(u^{n-1}, b_1, e^{n-1}) + e_n).$$

So the alphabet for the input vector $u^{(N)}$ is

$$\mathcal{U}^{(N)} = \{(u_1, \cdots, u_N) : u_n \in \mathcal{U}_n(u^{n-1}),\ \forall 1 \le n \le N\}.$$

Namely, $\mathcal{U}^{(N)}$ is the collection of all vectors of N causal mappings that are consistent with the energy constraint.

B. Transition Probabilities

With the transducer model in [4], the N-symbol transition probabilities $p(y^N|u^{(N)})$ for the new channel $\boldsymbol{W}$ is

$$\begin{aligned} p(y^N|u^{(N)}) &= \sum_{b_1, e^N} p(y^N, b_1, e^N|u^{(N)}) \\ &= \sum_{b_1, e^N} p(b_1)p(e^N)p(y^N|b_1, e^N, u^{(N)}) \\ &= \sum_{b_1, e^N} p(b_1)p(e^N) \prod_{n=1}^{N} p(y_n|u_n(b_1, e^n)). \end{aligned} \tag{2}$$

The new channel has the same capacity as the original one. This is because the new input symbols $u^{(N)}$ are just the

[2] To be concise we sometimes drop the sub-/super- scripts for a random process.

implementation of the codes $f^{(N)}$ for the original channel with the same constraints, and the output vector y^N has the same distribution in both cases. Thus codes in both models can be translated into each other with the same probability of error.

From this point on we only consider the equivalent channel $\boldsymbol{W}$ and call $U^{(N)} \in \mathcal{U}^{(N)}$ the input symbol. To avoid confusion we call X_n the immediate input symbol.

IV. Channel Capacity

To compute the capacity of the new channel we need to invoke the general capacity formula for arbitrary channels without feedback in [7]. Define the input distribution process $\boldsymbol{U}$ to be a collection of random vectors $\{U^{(N)}\}_{N=1}^{\infty}$, where each $U^{(N)}$ is a random vector of feasible strategy letters that corresponds to some probability distribution on $\mathcal{U}^{(N)}$, and such distributions for different N could be irrelevant to each other. The corresponding output distribution process $\boldsymbol{Y} = \{Y^{(N)}\}_{N=1}^{\infty}$ is the collection of random vectors $Y^{(N)} = (Y_1^{(N)}, \cdots, Y_N^{(N)})$ in $\mathcal{Y}^N$, each of which is induced by the input random vector $U^{(N)}$ and the conditional probability (2).

For each N define the normalized information density between $U^{(N)}$ and $Y^{(N)}$ to be

$$i_N(U^{(N)}; Y^{(N)}) := \frac{1}{N} \log \frac{p\left(Y^{(N)} | U^{(N)}\right)}{p\left(Y^{(N)}\right)}.$$

According to [7], the capacity of the channel is given by

$$C = \sup_{\boldsymbol{U}: U^{(N)} \in \mathcal{U}^{(N)}} \underline{\boldsymbol{I}}(\boldsymbol{U}; \boldsymbol{Y}) \tag{3}$$

where the inf-information rate $\underline{\boldsymbol{I}}(\boldsymbol{U}; \boldsymbol{Y})$ is defined as the liminf in probability of the sequence of random variables $i_N(U^{(N)}; Y^{(N)})$, and the supremum is taken over all possible input distribution processes $\boldsymbol{U}$.

V. Finite State Channel Reduction

As the alphabet size of each coordinate u_n for the input process is growing with n (possibly exponentially), the capacity formula (3) is difficult to evaluate. In the following we consider a special category of input symbols, whose alphabet size for each coordinate is essentially constant. With such input processes, together with some conditions on the energy arrival process $\{E_n\}$, the new channel reduces to a finite state channel defined in [10], whose capacity provides a lower bound for the original channel capacity C.

First let us define the channel state[3] S_n at time n to be the total energy available for the n-th transmission, i.e.,

$$S_n = B_n + E_n.$$

Then the energy constraint becomes

$$|X_n|^2 \leq S_n.$$

Note that the states depend on the input and there are only finitely many states, as with the assumptions at the end of Section II, $S_n \in \mathcal{S} = \mathcal{E}_B + \mathcal{E}_H$.

[3]Note that the channel states in this context are indeed "energy states" or "partial states", which only capture partial information of the full channel states $(B_1, \{E_n\})$ in Section III.

Consider an input symbol $u^{(N)} = (u_1, \cdots, u_N)$ whose n-th strategy letter u_n is only a function of S_n, then its alphabet size is a constant that does not depend on n. To be precise, each u_n is associated with an auxiliary function $v_n : \mathcal{S} \to \mathcal{X}$ that satisfies the energy constraint $|v_n(s)|^2 \leq s$ for all $s \in \mathcal{S}$. The strategy letter u_n is defined through v_n in the following way: for each (b_1, e^n), u_n first compute $s_n = B_n(u^{n-1}, b_1, e^{n-1}) + e_n$, then assign $u_n(b_1, e^n) = v_n(s_n)$. Hence the vector $v^N = (v_1, \cdots, v_N)$ uniquely determines the input symbol $u^{(N)}$, and for each N there is a one-to-one correspondence between the collection $\mathcal{U}'^{(N)}$ of all such special input symbols $u^{(N)}$ and $\mathcal{V}^N$, where $\mathcal{V} = \prod_{s \in \mathcal{S}} \mathcal{X}(s)$ is the alphabet for v_n.

Now restricting the input distribution process of (3) to only consisting of those random vectors $U^{(N)}$ that take values in $\mathcal{U}'^{(N)}$, we obtain a lower bound of the capacity:

$$C' = \sup_{\boldsymbol{U}: U^{(N)} \in \mathcal{U}'^{(N)}} \underline{\boldsymbol{I}}(\boldsymbol{U}; \boldsymbol{Y}).$$

Observe that by the one-to-one correspondence between $\mathcal{U}'^{(N)}$ and $\mathcal{V}^N$, for each $v^N \in \mathcal{V}^N$ and for all $y^N \in \mathcal{Y}^N$ we can define the conditional probability $p\left(y^N | v^N\right)$ through the corresponding $u^{(N)}$ and $p(y^N | u^{(N)})$, thus defining a surrogate channel $\boldsymbol{W}' = \{\mathcal{V}^N, W'^N = p(y^N | v^N), \mathcal{Y}^N\}_{N=1}^{\infty}$. Since the operation of the channel $\boldsymbol{W}$ restricted on the input alphabets $\mathcal{U}'^{(N)}$ is exactly the same as the surrogate channel $\boldsymbol{W}'$, they have the same capacity

$$C' = \sup_{\boldsymbol{V}: V^{(N)} \in \mathcal{V}^N} \underline{\boldsymbol{I}}(\boldsymbol{V}; \boldsymbol{Y}), \tag{4}$$

where $\boldsymbol{V} = \{V^{(N)}\}_{N=1}^{\infty}$ is the input distribution process for the surrogate channel, with each $V^{(N)} = (V_1^{(N)}, \cdots, V_N^{(N)})$ being a random vector in $\mathcal{V}^N$. The output distribution process $\boldsymbol{Y}$, the normalized information density $i_N(V^{(N)}; Y^{(N)})$ and the inf-information rate $\underline{\boldsymbol{I}}(\boldsymbol{V}; \boldsymbol{Y})$ are defined similarly as before for the channel $\boldsymbol{W}'$.

Remark 2: The advantage of using the process $\boldsymbol{V}$ and the surrogate channel $\boldsymbol{W}'$ is that each coordinate v_n now has a fixed finite alphabet $\mathcal{V}$. Furthermore, S_n contains all the information about the energy constraint on the current immediate input symbol X_n, which is the only influence the (full) channel states have on the transmission. Thus we conjecture that $C' = C$, but are not able to prove it yet.

Next to connect $\boldsymbol{W}'$ to the finite state channels (FSC) defined in [10], we assume $\{E_n\}$ to be i.i.d..[4] Then for the channel $\boldsymbol{W}'$ we have

$$\Pr(Y_n = y_n, S_{n+1} = s_{n+1} | V_n = v_n, S_n = s_n) = p(y_n | v_n(s_n)) p(e_{n+1}),$$

where $e_{n+1} = s_{n+1} - b_{n+1}$ with

$$b_{n+1} = \min\left\{ s_n - |v_n(s_n)|^2,\ \overline{B} \right\}$$

and $p(e_{n+1}) = 0$ for $e_{n+1} \notin \mathcal{E}_H$. Since Y_n is the output of a DMC with input $X_n = V_n(S_n)$ and $\{E_n\}$ is i.i.d., it is clear that conditioned on $V_n = v_n$ and $S_n = s_n$, Y_n and S_{n+1} are

[4]This condition can be relaxed to (finite-order) Markov, see Section VIII.

independent of all other random variables. Furthermore, the conditional probability above does not change with time, so we can denote it by

$$P(y_n s_{n+1}|v_n s_n) = P(s_{n+1}|v_n s_n)p(y_n|v_n(s_n)), \quad (5)$$

where $P(s_{n+1}|v_n s_n) = p(e_{n+1})$ as described above. By the definition of an FSC in [10], $\boldsymbol{W}'$ is an FSC.

As the input symbols V_n are independent of B_1, by Theorems 4.6.1, 4.6.2 and 5.9.2 in [10] the capacity of $\boldsymbol{W}'$ can also be represented as

$$C' = \lim_{N\to\infty} \underline{C}_N = \sup_N \left(\underline{C}_N - \frac{\log|\mathcal{S}|}{N} \right), \quad (6)$$

where

$$\underline{C}_N = \max_{V^N} \min_{s_1} \frac{1}{N} I(V^N; Y^N | S_1 = s_1).$$

Define an FSC (in our setting) to be indecomposable iff ([10, Theorem 4.6.3]) for some fixed N and each input sequence $v = \{v_n\}_{n=1}^{\infty}$, there exists an s_N (that may depend on v) such that

$$Q(s_N|v, s_1) := \Pr(S_N = s_N|\{V_n\} = v, S_1 = s_1) > 0 \quad (7)$$

for any s_1. If $\boldsymbol{W}'$ is indecomposable, then [10, Theorem 4.6.4]

$$C' = \lim_{N\to\infty} \max_{V^N} \frac{1}{N} I(V^N; Y^N). \quad (8)$$

In the next section we will show that under some mild conditions (see Theorem 1), $\boldsymbol{W}'$ is indeed indecomposable.

VI. Ergodicity Results

It is still difficult to evaluate any of the capacity formulas (4), (6) or (8) of C'. However, if we have a random process $\{V_n\}_{n=1}^{\infty}$ which induces a joint process $\{V_n, Y_n\}$ that satisfies SMB, then we have (see e.g. [11])

$$\underline{\boldsymbol{I}}(\boldsymbol{V}, \boldsymbol{Y}) = \lim_{N\to\infty} \frac{1}{N} I(V^N; Y^N) = H(\mathcal{V}) + H(\mathcal{Y}) - H(\mathcal{V}, \mathcal{Y}),$$

where $\boldsymbol{V}$ is the input distribution process corresponding to $\{V_n\}$ and $H(\mathcal{V})$, $H(\mathcal{Y})$ and $H(\mathcal{V}, \mathcal{Y})$ are the entropy rates of $\{V_n\}$, $\{Y_n\}$ and $\{V_n, Y_n\}$ respectively. Define the information rate $I(\mathcal{V}, \mathcal{Y})$ to be the right hand side of the equation above, then for each process $\{V_n\}$ with the aforementioned property, $I(\mathcal{V}, \mathcal{Y})$ serves as a lower bound both for (4) and (8). Furthermore, $I(\mathcal{V}, \mathcal{Y})$ is amenable to numerical computation. By SMB we can use a long sample sequence to estimate $I(\mathcal{V}, \mathcal{Y})$ (see e.g. [12]), which has linear complexity in many cases, as opposed to the exponential complexity (in N) of (6) and (8). See Section VII for an example on the numerical computation and optimization of such lower bounds of C'.

Having all these said, we still need to find conditions for such "SMB-amenable" processes $\{V_n\}$. That is why we need the ergodicity results. Since FSC's are Markov channels (see [11], [13] for definition), they are AMS [13]. Hence an AMS process $\{V_n\}_{n=1}^{\infty}$ induces an AMS joint process $\{V_n, S_n, Y_n\}$. Moreover, under some mild conditions on the channel model (see Theorem 1 below), the Markov channel is weakly ergodic a.e. (see [14] for definition) for a stationary and ergodic process $\{V_n\}_{n=1}^{\infty}$, and so the joint process $\{V_n, S_n, Y_n\}$ is AMS and ergodic [14]. Hence SMB applies [11] for such channels with stationary and ergodic input processes.

Remark 3: An indecomposable channel is ergodic [13], and is weakly ergodic a.e. [14] for any stationary and ergodic input process, so SMB applies in this case.

Theorem 1: For the FSC $\boldsymbol{W}'$, assume $\{E_n\}$ is i.i.d. and E_n takes positive probabilities on all of $\mathcal{E}_H$.

i) If there exists N such that for each input sequence $v = \{v_n\}_{n=1}^{\infty}$ and any $S_1 = s_1$, $B_N = \overline{B}$ with a positive probability, then $\boldsymbol{W}'$ is indecomposable.
ii) If $\mathcal{E}_H$ is a continuous interval of non-negative integers and $\max \mathcal{E}_H - \min \mathcal{E}_H \geq \overline{B}$, then $\boldsymbol{W}'$ is indecomposable.
iii) If for a stationary and ergodic process $\{V_n\}_{n=1}^{\infty}$, there is a finite vector v^N with $\Pr(V^N = v^N) > 0$ such that for any $S_1 = s_1$, $B_N = \overline{B}$ with positive probability, then the channel is weakly ergodic a.e..
iv) Both i) and iii) hold if $\max \mathcal{E}_H > \max\{|x|^2 : x \in \mathcal{X}\}$.

Proof: We only prove that the FSC is indecomposable.

i): Whenever such N exists, let $e_N \in \mathcal{E}_H$. Then for $s_N = B_N + e_N$ we have

$$Q(s_N|v, s_1) \geq \Pr(B_N = \overline{B}|v, s_1)p(e_N) > 0$$

for all v and arbitrary s_1. Hence (7) holds.

ii): With a positive probability S_2 can always be boosted up to $s_2 = \overline{B} + \min \mathcal{E}_H$, hence (7) holds.

iv): If $\max \mathcal{E}_H > \max\{|x|^2 : x \in \mathcal{X}\}$, then for any v and s_1, at most after $n = \overline{B}$ transmissions, $S_n \geq \overline{B}$ with a probability no smaller than $[\Pr(E_n = \max \mathcal{E}_H)]^n > 0$, in which case $B_{n+1} = \overline{B}$. ∎

Remark 4: The conditions in ii) and iv) are satisfied if E_n can reach a relatively high energy level (compared to $\mathcal{X}$ or $\overline{B}$) with even a very small positive probability, which is not a harsh requirement for many natural energy sources.

The following is another (elementary) ergodicity result.

Lemma 1: If $\{V_n\}$ and $\{E_n\}$ are i.i.d. and the corresponding $\{S_n\}$ form an irreducible and aperiodic Markov chain, then $\{V_n, S_n\}$ is an ergodic Markov process. The whole joint distribution is also ergodic and AMS [15], so SMB holds.

Remark 5: We have more results on weak ergodicity. However, due to the space constraint and extra notation and background knowledge we have to introduce, these results are omitted in this paper.

VII. Numerical Computation

As discussed in the last section, for a joint process $\{V_n, Y_n\}$ that satisfies SMB, we can use the sample entropies of a long sequence to estimate the information rate. When the joint process has a hidden Markov structure, the sample entropies can be efficiently computed [12]. Hence for example, for a stationary ergodic Markov input process $\{V_n\}$ with an indecomposable FSC $\boldsymbol{W}'$, the information rate can be efficiently computed. Furthermore, [8] developed a generalized Blahut-Arimoto algorithm that optimize the information rate for a

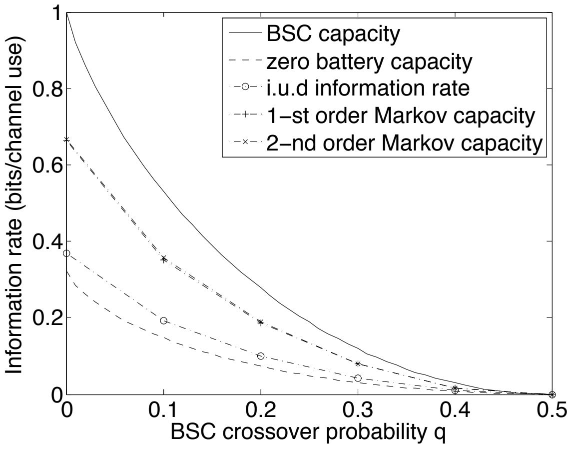

Fig. 2. The information rates.

given Markov order of the input process. In the following we will show an example of our model that fits into this regime.

Example 1: Assume $\{E_n\}$ is an i.i.d. Bernoulli(0.5) process with $\mathcal{E}_H = \mathcal{X} = \{0, 1\}$. Assume $\overline{B} = 1$, then $\mathcal{S} = \{0, 1, 2\}$, $\mathcal{V} = \{(0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1)\}$. Let the DMC be a binary symmetric (BSC) with crossover probability q. Then the condition in case ii) of Theorem 1 is satisfied and $\boldsymbol{W}'$ is indecomposable.

The capacity for the same BSC without energy constraint is $1 - H(q)$, which is an upper bound for the case of infinite battery. On the other extreme, when there is no battery, as commented in Section III Shannon's method [4] can be modified to obtain an equivalent DMC with input alphabet $\mathcal{U} = \{u_a, u_b\}$ where $u_a = (0, 0)$ and $u_b = (0, 1)$, both of which are functions of E_n. The transition probability of this DMC is

$$p(y|u) = \sum_{e \in \mathcal{E}_H} p(e)p(y|u(e)), \quad \forall y \in \mathcal{Y},\ u \in \mathcal{U}.$$

Its capacity can be easily calculated analytically.

For the channel $\boldsymbol{W}'$ we compute the i.u.d. rate, which is the information rate for the i.i.d. uniform input process, and optimize the information rate over Markov input processes of order 1 and 2. The numerical results are shown in Fig 2. For comparison, the capacities for the same BSC without energy constraint and with zero battery are also shown.

We have the following remarks on the numerical results:

1) The minimal non-zero battery storage can give us a great boost on the capacity, it even achieves a significant fraction (around 70%) of the capacity without energy constraints.
2) The Markov input processes achieves higher rate than the i.u.d. input, and with higher order the information rate is higher. So memory in the input helps, but increasing the Markov order by 1 only slightly increases the information rate.

VIII. Discussions

Our future work can extend the results in the following directions:

1) Continuous channel. It is possible to treat the continuous channels with finite input, e.g. AWGN with binary input. Although the FSC and Markov channel results are both for finite alphabets, we can consider only the state process itself as the output of an Markov channel, then a continuous memoryless channel is connected to the output. For such a case we can still derive ergodicity results and apply SMB [11].
2) The energy arrival process can be possibly extended from i.i.d. to k-step Markov, if we consider these Markov states as part of the channel states.
3) We can compute the numerical results for the source strategy letters that depends on more energy states, i.e., on some history of $\{S_n\}$.

References

[1] O. Ozel, J. Yang, and S. Ulukus, "Optimal broadcast scheduling for an energy harvesting rechargeable transmitter with a finite capacity battery," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2193–2203, Jun 2012.

[2] J. Yang and S. Ulukus, "Optimal packet scheduling in an energy harvesting communication system," *IEEE Trans. Commun.*, vol. 60, no. 1, pp. 220–230, January 2012.

[3] O. Ozel and S. Ulukus, "Achieving AWGN capacity under stochastic energy harvesting," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6471–6483, October 2012.

[4] C. Shannon, "Channels with side information at the transmitter," *IBM Journal of Research and Development*, vol. 2, no. 4, pp. 289–293, Oct. 1958.

[5] O. Ozel and S. Ulukus, "AWGN channel under time-varying amplitude constraints with causal information at the transmitter," in *Proc. 45th Asilomar Conf. on Signals, Systems and Computers*, Nov. 2011.

[6] K. Tutuncuoglu and A. Yener, "Optimum transmission policies for battery limited energy harvesting nodes," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1180–1189, Mar. 2012.

[7] S. Verdú and T. S. Han, "A general formula for channel capacity," *IEEE Trans. Inf. Theory*, vol. 40, no. 4, pp. 1147–1157, Jul. 1994.

[8] P. O. Vontobel, A. Kavčić, D. M. Arnold, and H.-A. Loeliger, "A generalization of the Blahut–Arimoto algorithm to finite-state channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1887–1918, May 2008.

[9] G. Caire and S. Shamai (Shitz), "On the capacity of some channels with channel state information," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2007–2019, Sep. 1999.

[10] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley & Sons, 1968.

[11] R. M. Gray, *Entropy and Information Theory*. New York: Springer-Verlag, 1990.

[12] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavčić, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inf. Theory*, vol. 52, no. 8, pp. 3498–3508, August 2006.

[13] J. C. Kieffer and M. Rahe, "Markov channels are asymptotically mean stationary," *Siam Journal of Mathematical Analysis*, vol. 12, no. 3, pp. 293–305, 1981.

[14] R. M. Gray, M. O. Dunham, and R. L. Gobbi, "Ergodicity of Markov channels," *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 656–664, Sep. 1987.

[15] Y. Ephraim and N. Merhav, "Hidden Markov processes," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1518–1569, Jun. 2002.