=============== **INFORMATION THEORY** ===============

Finite Blocklength Analysis of Energy Harvesting Channels

K. G. Shenoy[a,*] and V. Sharma[a,**]

[a] *Electrical Communications Engineering Dept., Indian Institute of Science, Bangalore, India*
e-mail: *konchady@iisc.ac.in, **vinod@iisc.ac.in

Abstract—We consider additive white Gaussian noise channels and discrete memoryless channels where the transmitter harvests energy from the environment. These can model wireless sensor networks as well as Internet of Things. By providing a unifying framework that works for any energy harvesting channel, we study these channels assuming an infinite energy buffer and provide the corresponding achievability and converse bounds on the channel capacity in the finite blocklength regime. We additionally provide moderate deviation asymptotic bounds.

Key words: achievable rates, converse, channel capacity, finite blocklength, EH-AWGN, EH-DMC, moderate deviations.

1. INTRODUCTION

In the information theoretic analysis of channels, channel capacity is the maximum rate at which a source can transmit messages to the receiver subject to an arbitrarily small probability of error. However, channel capacity can be achieved arbitrarily closely by using very large blocklength codes. In practice, we are restricted by blocklength, and as a result, we would like to study the backoff from capacity as well as the variation in maximal code size as a function of blocklength. For a fixed probability of error, the study of achievable rates in the finite blocklength regime, with a focus on the second-order coefficient, is known as second-order analysis in literature.

Like channel capacity, a finite blocklength characterization consists of two parts, namely the achievability and the converse bound on the maximal code size (number of messages) M. Given the probability of error, the achievability part usually deals with the existence of a code using, for instance, random coding arguments or manipulating general achievability bounds and showing that the bound can be achieved. The converse, on the other hand, is an upper bound on the maximal code size which is to be satisfied by every feasible code. This paper focuses on developing both for energy harvesting channels.

Energy harvesting (EH) channels and networks have gained considerable interest recently due to advances in wireless sensor networks and green communications (see [2–4]). Transmitting symbols requires energy at the encoder end. Thus, the study of the channel is done in tandem with the energy harvesting system. The energy harvesting section is modeled as a buffer or a rechargeable battery which stores incoming energy from some ambient source (e.g., solar energy from the sun). The energy buffer may be of finite or infinite length, and the energy arrival process may be discrete or continuous. A problem of interest is to compare the performance of a channel with and without an energy harvesting system (e.g., whether we can quantify the impact on the channel capacity, finite blocklength capacity, etc.).

Finite blocklength analysis for discrete memoryless channels (DMC) was first carried out in [5]. Nonasymptotic second-order results for additive white Gaussian noise (AWGN) channels in addition to other channel types were provided in [6, 7]. Further, in [7, 8], third-order terms were provided and a meta-converse was developed, a converse result that recovered and improved upon known converses. Later, tighter results for various DMCs were studied in [9]. Nonasymptotic analysis of channels with state was carried out in [10]. Under the energy harvesting setup, assuming infinite buffer, the channel capacity for EH-AWGN channels was obtained in [11, 12]. The study of finite blocklength achievability for energy harvesting noiseless binary channels was carried out in [13]. The achievability result with second-order $O(\sqrt{n})$ was shown in [1] for EH-AWGN channels. Both achievability and converse results for EH-AWGN channels were further refined recently in [14], which considered block i.i.d. energy arrivals.

In addition to finite blocklength analysis, we also give bounds on the moderate deviations coefficient for EH-AWGN channels and EH-DMC. In this analysis, we transmit at a rate less than capacity where the backoff goes to zero at a certain rate called the moderate deviation regime. In this regime, the probability of error will go to zero with increasing blocklength n. The goal is to characterize the moderate deviation error exponent. Moderate deviation analysis has been studied for memoryless channels in [15, 16]. In [16], the authors characterize the moderate deviation coefficient in terms of the channel dispersion. For DMCs with variable length feedback, the moderate deviation analysis was carried out in [17].

Main contributions. In this paper, we provide a scheme that can directly be used to compute achievable rates for a wide class of energy harvesting channels. We focus on analyzing EH-AWGN and EH-DMC with infinite buffer. In particular,

1. We provide finite blocklength achievable rates for EH-AWGN and EH-DMC channels assuming a fixed maximum probability of error. It is shown that a save and transmit scheme where the saving phase is $O(\sqrt{n})$ long is sufficient to allow for reliable communication in an energy harvesting setup. When compared with the non-energy harvesting case (but with an equivalent average power constraint), we observe that the second-order term is still $\Theta(\sqrt{n})$. Note that the coefficients of the second-order term would not necessarily be same.

2. Next, we provide a finite blocklength converse bound, i.e., an upper bound on achievable rates, for EH-AWGN channels. This is derived by modifying the meta converse from [7] and specifically applying it to EH-AWGN channels. In quasi-static fading channels, [18] also modified the meta converse suitably to obtain the desired bounds. Moreover we give an alternate, shorter proof of the converse bound for EH-AWGN channels, which was first shown in [14]. Additionally, we analyze DMCs with energy harvesting and provide the finite blocklength converse bounds for them. The analysis of both achievability and converse in this framework for EH-DMCs is novel. We are able to show that in both the achievability and converse, the second-order term for EH-AWGN and EH-DMC is $O(\sqrt{n})$. This also gives us the strong converse for these channels as a corollary, since the first-order term is unaffected by the probability of error term. Furthermore, we extend our results to the case where a sequence of messages is sent in the system where residual energy from previous transmissions can be beneficial.

3. Next, we provide moderate deviation lower and upper bounds for both types of channels. This is done by showing that the bounds on channel dispersion obtained while proving the finite blocklength bounds also bound the moderate deviations coefficient. These bounds are novel, i.e., unavailable elsewhere for energy harvesting channels. Finally, we plot our lower and upper rate bounds for certain parameters and provide suitable inferences.

4. Finally, we plot our finite blocklength bounds for the EH-AWGN channel under low, medium, and high SNR regimes and compare them with equivalent non-energy harvesting AWGN channel. For the EH-DMC, we choose, as examples, an energy harvesting binary symmetric channel (EH-BSC) and an energy harvesting binary erasure channel (BEC) and plot the corresponding

results. Moreover, the cases where the plots reveal nontrivial facts about the channel rates are discussed.

Later, in Section 10, we will compare and discuss our results and methods in detail with those available in [14, 18].

2. PRELIMINARIES

2.1. Basic Notation

We shall use boldface letters (e.g., $\boldsymbol{x}$) to denote vectors (belonging to $\mathbb{R}^n$ for a specified $n \in \mathbb{N}$). When the length of a vector needs to be specified, we shall mention it as $\boldsymbol{x}^k = (x_1, x_2, \ldots, x_k)$. Similarly, $\boldsymbol{x}_i^j = (x_i, x_{i+1}, \ldots, x_j)$. Lowercase letters denote deterministic scalars or vectors, whereas uppercase letters denote random variables or random vectors respectively. We shall use $[M]$ to denote the set $\{1, 2, \ldots, M\}$. We shall denote by $\mathcal{P}(\mathcal{X})$ the set of probability distributions on $\mathcal{X}$ (in cases where the alphabet is clear, we simply use $\mathcal{P}$). The expectation operator will be denoted by $\mathbf{E}$, and if the distribution (say P) needs to be specified, then it shall be denoted as $\mathbf{E}_P$. We will occasionally use the Bachmann–Landau notation $O(\cdot)$, $\Theta(\cdot)$, etc. to denote appropriate orders. All logarithms are by default to the base 2, and the base may be explicitly mentioned in some cases.

2.2. Channels, Probability of Error, and Capacity

Given an input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$, a *channel*, denoted by $W(y \,|\, x)$ or equivalently $P_{Y|X}$, is a conditional probability measure on $\mathcal{Y}$ given $x \in \mathcal{X}$. If the probability density function exists for the channel, we shall denote it by $f_{Y|X}$.

Given a probability distribution P on $\mathcal{X}$ and a channel W, we define the output measure PW as

$$PW(y) = \sum_{x \in \mathcal{X}} P(x) W(y \,|\, x).$$

There are two notions of probability of error which we will use. Given a code $\mathcal{C}$ with M messages, let $U \in [M]$ be the random variable, uniformly distributed on $[M]$, denoting the message to be transmitted and $\widehat{U} \in [M]$ the message that is decoded at the receiver. The *maximal probability of error* (max p.o.e.) of the code $\mathcal{C}$ is

$$P_{e,\mathrm{max}}(\mathcal{C}) := \max_{1 \le m \le M} \Pr[\widehat{U} \neq m \,|\, U = m]. \tag{1}$$

Similarly, the *average probability of error* (avg p.o.e.) is defined as

$$P_{e,\mathrm{avg}}(\mathcal{C}) := \frac{1}{M} \sum_{m=1}^{M} \Pr[\widehat{U} \neq m \,|\, U = m]. \tag{2}$$

The channel capacity is the same in both cases. However, in the finite blocklength regime, the differences are in higher-order terms, resulting in an $O(\log n)$ difference [7]. In this paper, we will stick to the maximal probability of error criterion, since it is advantageous while analyzing the energy harvesting DMC results. As a matter of technique, max p.o.e. bounds usually involve working with sequences as opposed to avg p.o.e., where we work with distributions.

An (n, M, ε) code is a code with M codewords of codeword length n and probability of error at most ε. We define

$$M^*(n, \varepsilon) := \max\{M : \text{there exists an } (n, M, \varepsilon) \text{ code}\}.$$

Given an (n, M, ε) code, we shall call $\dfrac{\log M}{n}$ the *rate* of the code. For $0 < \varepsilon < 1$, the ε-*capacity* C_ε is defined as

$$C_\varepsilon = \lim_{n \to \infty} \frac{\log M^*(n, \varepsilon)}{n}$$

and the *capacity* of the channel is defined as

$$C = \lim_{\varepsilon \to 0} C_\varepsilon.$$

Note that both limits exist. It is clear that $C_\varepsilon \geq C$. However, for certain classes of channels like DMCs and standard AWGN channels with average power constraints, we have $C_\varepsilon = C$ for every $0 < \varepsilon < 1$. Then we say that the channel satisfies the *strong converse*, which means that if we transmit at rates greater than capacity, the probability of error of the code tends to 1 as the blocklength n tends to infinity.

Given $0 < \varepsilon < 1$ and any channel, let $M_m(\varepsilon)$ be the maximal number of codewords of the code under max p.o.e. and $M_a(\varepsilon)$ the one for avg p.o.e. for that channel. Clearly, we have $M_m(\varepsilon) \leq M_a(\varepsilon)$. Using code expurgation (see, e.g., [19]), we also have

$$M_m(\varepsilon') \geq \frac{\varepsilon' - \varepsilon}{\varepsilon'} M_a(\varepsilon) \tag{3}$$

for $\varepsilon' > \varepsilon$. This means that any upper bound on rates for average p.o.e. also is an upper bound for the maximal case. However, for achievable rates, the lower bound for avg p.o.e. will be slightly penalized when extended to max p.o.e.

2.3. AWGN Channel

Given $\boldsymbol{a} \in \mathbb{R}^n$ and a covariance matrix $\boldsymbol{K} \in \mathbb{R}^{n \times n}$, denote

$$\mathcal{N}(\boldsymbol{a}; \boldsymbol{K}) := \frac{\exp\{-(\boldsymbol{x} - \boldsymbol{a})^T \boldsymbol{K}^{-1} (\boldsymbol{x} - \boldsymbol{a})\}}{(2\pi)^{n/2} (\det(\boldsymbol{K}))^{1/2}}$$

as the multivariate normal distribution with mean $\boldsymbol{a}$ and covariance matrix $\boldsymbol{K}$ whose determinant is nonzero. An additive white Gaussian noise (AWGN) channel with noise variance σ^2 is given by

$$Y = x + Z$$

where $x \in \mathbb{R}$ is the input to the channel and $Z \sim \mathcal{N}(0; \sigma^2)$. The n-dimensional version is obtained by applying the one-dimensional version ($n = 1$) case independently on each input x_i, $1 \leq i \leq n$. The AWGN channel with average power constraint S is an AWGN channel where the input $\boldsymbol{x}$ satisfies

$$\|\boldsymbol{x}\|_2^2 := \sum_{i=1}^n x_i^2 \leq nS. \tag{4}$$

The capacity of an AWGN channel (denoted by C_G) with average power constraint P is given by

$$C_G := \frac{1}{2} \log_2\left(1 + \frac{P}{\sigma^2}\right) \text{ bits per channel use.}$$

In [5,6], it was shown that for an AWGN channel with average power constraint P, the maximum code size $M^*(n, \varepsilon, P)$, for n sufficiently large, satisfies

$$\log M^*(n, \varepsilon, P) = nC_G + \sqrt{nV_G}\Phi^{-1}(\varepsilon) + O(\log(n))$$

where the probability of error is at most ε, $V_G = \dfrac{P(P+2)}{2(P+1)^2} \log_2^2(e)$, and $\Phi(x) = \int\limits_{-\infty}^{x} \dfrac{e^{-u^2/2}}{\sqrt{2\pi}}\, du$.

2.4. Discrete Memoryless Channels (DMC)

A DMC is characterized by a finite input alphabet $\mathcal{X}$, finite output alphabet $\mathcal{Y}$, and the transition probabilities given by $W = P_{Y|X}$, which satisfies for every $n \geq 1$

$$P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,\boldsymbol{x}) = \prod_{i=1}^{n} P_{Y|X}(y_i\,|\,x_i).$$

While the output is, in principle, allowed to depend on past outputs (which is known as a DMC with feedback), we only consider DMCs without feedback. The capacity C_D of a DMC W that is nonexotic[1] (see [7, Appendix H] and also [9]) is given by Shannon's formula as

$$C_D = \sup_{P \in \mathcal{P}(\mathcal{X})} I(P; W) \triangleq \sup_{P \in \mathcal{P}(\mathcal{X})} \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P(x) W(y\,|\,x) \log\left(\frac{W(y\,|\,x)}{PW(y)}\right),$$

where $I(P; W)$ is the mutual information (see [19]) between the input and output of the channel.

Now we define a few terms that will be required later.

Definition 1. Given a channel W and an output distribution Q, the information density [20] of the channel is given by

$$i(x, y; Q) = \log\left(\frac{W(y\,|\,x)}{Q(y)}\right). \tag{5}$$

Often $Q = PW$ for some $P \in \mathcal{P}(\mathcal{X})$, in which case we shall denote the information density by $i_P(x, y)$.

Observe that $I(P; W) = \sum_{x,y} P(x) W(y\,|\,x) i_P(x, y)$. Similarly, the variance of information density is given by

$$V(P; W) := \left[\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} P(x) W(y\,|\,x)(i_P(x, y))^2\right] - (I(P; W))^2.$$

The finite blocklength result for nonexotic DMCs with channel W, probability of error $0 < \varepsilon < 1$, and $V(P; W) > 0$ for a capacity achieving distribution P is given by (see [5–9])

$$\log M^*(n, \varepsilon) = nC_D + \sqrt{nV_D}\Phi^{-1}(\varepsilon) + O(\log(n)),$$

where

$$V_D = \begin{cases} V_{\min} := \min_{P \in \Pi} V(P; W), & \varepsilon \leq 1/2, \\ V_{\max} := \max_{P \in \Pi} V(P; W), & \varepsilon > 1/2, \end{cases}$$

and $\Pi = \{P \in \mathcal{P}(\mathcal{X}) : I(P; W) = C_D\}$ is the set of capacity achieving distributions.

2.5. DMC with Cost Constraints

Let $\Lambda \colon \mathcal{X} \to \mathbb{R}$ be a nonnegative function, which we will refer to as the energy function. The energy function simply returns the energy of the symbol x, which is a generalization of the standard

---

[1] A DMC is exotic if the maximum variance of its information density, i.e., $V_{\max}$, is 0 and for some input symbol x_0, $P(x_0) = 0$ for any capacity achieving distribution P but $D(W(\cdot\,|\,x_0)\,\|\,Q_Y^*) = C$ and $V(W(\cdot\,|\,x_0)\,\|\,Q_Y^*) > 0$. Here $D(P_1\,\|\,P_2)$ is the KL divergence between P_1 and P_2 and $V(P_1\,\|\,P_2) = \sum_x P_1(x) \log^2\left(\frac{P_1(x)}{P_2(x)}\right) - D^2(P_1\,\|\,P_2)$.
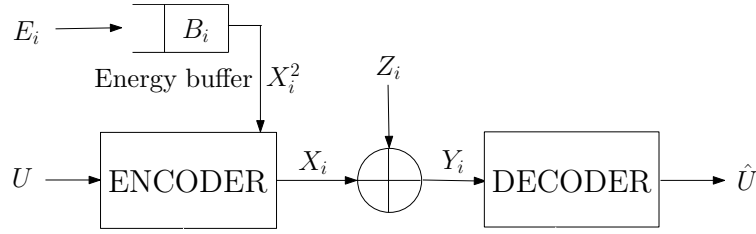
Fig. 1. Block diagram of an AWGN energy harvesting system.

energy function $\Lambda(x) = x^2$ for an AWGN channel. We further assume that the energy function is separable, i.e., given a vector $\boldsymbol{x} \in \mathcal{X}^n$,

$$\Lambda(\boldsymbol{x}) := \sum_{i=1}^{n} \Lambda(x_i). \tag{6}$$

Define the constrained sets $\mathbb{F}_a$ and $\mathcal{F}_a$ for $a \geq 0$ as follows:

$$\mathbb{F}_a = \{\boldsymbol{x} \in \mathcal{X}^n : \Lambda(\boldsymbol{x}) \leq na\}, \tag{7}$$

$$\mathcal{F}_a = \{P \in \mathcal{P} : \mathbf{E}_P[\Lambda(X)] \leq na\}. \tag{8}$$

In a DMC with cost constraints (see [21, 22]), where the codewords are drawn from $\mathbb{F}_a$, the capacity changes to

$$C_D(a) = \sup_{P \in \mathcal{F}_a} I(P; W). \tag{9}$$

Moreover, the maximum achievable code size, for any $a > 0$, denoted by $M^*(n, \varepsilon, a)$, under some regularity conditions (see [7] for the result and [22] for some refinements), is given by

$$\log M^*(n, \varepsilon, a) = nC_D(a) + \sqrt{nV_D(a)}\Phi^{-1}(\varepsilon) + O(\log n)$$

where $V_D(a)$ is the channel dispersion (see [6]).

An energy harvesting DMC (EH-DMC) may be viewed as a generalization of a DMC with cost constraints, and its finite blocklength analysis is reserved to Section 4.

2.6. Energy Harvesting AWGN Channel

An energy harvesting system consists of an energy buffer, which stores energy from various sources over a period of time. Energy is usually harvested from some ambient source, e.g., solar power. An EH-AWGN channel consists of an energy harvesting system at the transmitter end followed by an AWGN channel, as is shown in Fig. 1. To send symbol x on the channel, we would require x^2 units of energy from the buffer, and if sufficient energy is available, the transmission succeeds; otherwise, an outage occurs. An outage event may be handled as an error event, or a suitably truncated symbol may be transmitted. In this paper, as far as achievability is concerned, the outage will be treated as an error event. We assume that the energy buffer has infinite capacity and energy leakages do not occur. Additionally, the incoming energy process $\{E_i\}$ is assumed to be i.i.d., nonnegative, with finite mean $\mathbf{E}[E_1]$ and finite variance σ_E^2.

The system works as follows. We first harvest energy E_i in slot i, use it along with some energy in the buffer if needed to transmit the symbol, and then store the remaining energy. Let B_i be the

energy in the buffer at the ith transmission slot. Assume $B_0 = 0$. Then the energy in the buffer, for $1 \leq i \leq n$, evolves as

$$B_i = (B_{i-1} + E_i - X_i^2)^+$$

where $(x)^+ = \max\{x, 0\}$. In our case, we will be designing $\boldsymbol{X}$ such that the term inside $(\cdot)^+$ is nonnegative.

For the ordinary AWGN channel with power constraint S, the sequences were supposed to satisfy (4). The constraint for the energy harvesting AWGN channel with input $\boldsymbol{x}$ and energy vector $\boldsymbol{e}$ is

$$\left\|\boldsymbol{x}^k\right\|_2^2 \leq \left\|\boldsymbol{e}^k\right\|_1, \quad 1 \leq k \leq n,$$

which is another way of saying that there should, at every time instant, be enough energy to transmit the desired symbol. To ensure this, $\boldsymbol{x}$ is allowed to depend on $\boldsymbol{e}$.

The capacity of an EH-AWGN channel (see [11, 12]) is

$$C_{EG} = \frac{1}{2} \log\left(1 + \frac{\mathbf{E}[E_1]}{\sigma^2}\right). \tag{10}$$

Additionally, the strong converse was also shown to hold for this channel (see [11]). This would logically imply a converse of the form $\log M \leq nC_{EG} + o(n)$. However, since we seek for a refinement of this expression, we would need finer tools to extract a finite blocklength converse. In this regard, we will be using several results from [7]. For clarity, we use the notation of that paper.

We now state the following bounds on finite blocklength capacity for EH-AWGN channels.

Theorem 1. *Consider an EH-AWGN channel with AWGN variance σ^2, with energy harvesting process $\{E_i\}$, i.i.d. at the encoder, with mean $\mathbf{E}[E_1]$ and variance $\sigma_E^2 < \infty$. Given maximal probability of error $\varepsilon > 0$, we have the following.*

1 (Achievable bound). *For sufficiently large blocklength n, the maximum size of the code, $M^*(n, \varepsilon)$, satisfies*

$$\log M^*(n, \varepsilon) \geq nC_{EG} + \sqrt{n}\left[\sqrt{V_{EG}}\Phi^{-1}(\lambda\varepsilon) - K_{\varepsilon,\lambda}C_{EG}\right] - \log n + O(1), \tag{11}$$

where C_{EG} is defined in (10), $V_{EG} = \dfrac{\mathbf{E}[E_1]}{\mathbf{E}[E_1] + \sigma^2}\log_2^2(e)$, $K_{\varepsilon,\lambda} = \sqrt{\dfrac{4(2\mathbf{E}[E_1]^2 + \sigma_E^2)}{(1-\lambda)\varepsilon\mathbf{E}[E_1]^2}}$, and the above holds for any $0 < \lambda < 1$;

2 (Converse bound). *Also,*

$$\log M^*(n, \varepsilon) \leq nC_{EG} + \sqrt{nV_{EG2}}\Phi^{-1}(\varepsilon) + \frac{1}{2}\log n + O(1), \tag{12}$$

where $V_{EG2} = \dfrac{\mathbf{E}[E_1]^2 + \mathbf{E}[E_1^2] + 4\sigma^2\mathbf{E}[E_1]}{4(\mathbf{E}[E_1] + \sigma^2)^2}\log_2^2(e)$.

We defer the proof of the achievable bound to Section 3 and the converse bound to Section 6. While the second-order terms (coefficients of $\sqrt{n}$) do not match, we can conclude that $\log M^*(n, \varepsilon) = nC_{EG} + \Theta(\sqrt{n})$.

2.7. Energy Harvesting DMC

An energy harvesting DMC is a DMC with an energy harvesting setup at the encoder. Let $\Lambda(\cdot)$ be the energy function (see (6)) associated with this DMC. The model is the same as that of an EH-AWGN channel except for the following differences and assumptions:

1. The AWGN channel is replaced with a DMC;
2. Energy consumed by symbol x_i is $\Lambda(x_i)$. Also, there is a symbol x_0 with $\Lambda(x_0) = 0$;
3. We additionally assume that the DMCs are not exotic.

The analysis for energy harvesting DMCs is, by and large, analogous to the analysis of EH-AWGN channels. However, using method of types (refer to [19, 21] for more information on types), we are able to improve the converse bound to resemble that of the original non-energy harvesting DMC.

The capacity of an EH-DMC, where the energy harvesting process has mean $\mathbf{E}[E_1]$, is given by

$$C_{ED} := \sup_{P \in \mathcal{F}_{\mathbf{E}[E_1]}} I(P; W) \tag{13}$$

where $\mathcal{F}_a$ was defined in (8).

The following are the finite blocklength bounds on rate for EH-DMC proved in this paper.

Theorem 2. *Given $0 < \varepsilon < 1$, for maximal probability of error, consider an EH-DMC with HUS architecture, and the energy process $\{E_i\}$ i.i.d. with $\mathbf{E}[E_1^2] < \infty$.*

1 (Achievable bound). *Given the input distribution $P_X \in \mathcal{F}_{\mathbf{E}[E_1]}$, the maximal size of the code $M^*(n, \varepsilon)$ with blocklength n sufficiently large satisfies*

$$\log M^*(n, \varepsilon) \geq nI(P_X; W) - \sqrt{n}K_{\varepsilon,\lambda}I(P_X; W) + \sqrt{nV(P_X; W)}\Phi^{-1}(\lambda\varepsilon) - \log n + O(1), \tag{14}$$

for any $0 < \lambda < 1$. Here $K_{\varepsilon,\lambda} = \dfrac{2\sqrt{\mathrm{Var}(\Delta_1)}}{\mathbf{E}[E_1]\sqrt{(1-\lambda)\varepsilon}}$ and $\Delta_1 = E_1 - \Lambda(X_1)$;

2 (Converse bound). *Given $\eta > 0$, the maximal size of the code $M^*(n, \varepsilon)$ satisfies*

$$\log M^*(n, \varepsilon) \leq nC_{ED} + \sqrt{n}C'(\mathbf{E}[E_1])D_{\varepsilon} + \sqrt{nV_{\varepsilon}^*(\eta)}\left(\Phi^{-1}(\varepsilon) + \frac{K_{\varepsilon}\varepsilon}{4}\right) + O(\log n), \tag{15}$$

where $C'(\cdot)$ is the derivative of the capacity cost function given in (9) and where D_{ε}, K_{ε}, and $V_{\varepsilon}^*(\eta)$ are functions of ε independent of n.

2.8. Encoder and Decoder for Energy Harvesting Channels

For traditional channels (AWGN, DMC, etc.), the encoder and decoder have access to the codebook (random or otherwise) for the purposes of encoding and decoding respectively. In the energy harvesting setup, the encoder has access to the incoming energy values. Hence, any codeword $c \in \mathcal{C}$, where $\mathcal{C}$ is the codebook, is a length-n vector $c(m, \boldsymbol{e}^n)$, where n is the block length, for message m and energy vector $\boldsymbol{e}^n$. Due to causality requirements, the ith symbol of the codeword can only depend on $\boldsymbol{e}^i$. The decoder does not have access to these energy values and therefore does not have access to this energy dependent codebook. The decoder is, on the other hand, allowed to have access to an energy independent pre-codebook. Henceforth, in the context of energy harvesting channels, a codeword corresponding to message m shall mean the mapping $m \to c(m, \cdot)$. We note that the definitions of code size M, probability of error, etc. as defined in Section 2.2 remain unchanged. This is in the same spirit of the analysis of channels with state information available at encoder only.

We shall see in the achievability proofs that a codebook that is independent of energy values is created which is also available at the decoder. Then, at the encoder, the energy values are used to modify the codewords so as to meet the necessary constraints. This is one way of creating an encoder-decoder pair. This concept is similar to the one used in channels with state where the encoder has state information but not decoder [23].

3. FINITE BLOCKLENGTH ACHIEVABILITY BOUND FOR EH-AWGN

This section deals with the proof of Theorem 1, part 1. Assume that $0 < \varepsilon < 1$ is provided. We will first construct an avg p.o.e. $\varepsilon_n = \varepsilon - 1/\sqrt{n}$ code, where $n > \varepsilon^{-2}$. The code expurgation

argument yields the following bound:

$$\log M_m(n, \varepsilon) \geq \log M_a(n, \varepsilon_n) - \frac{1}{2}\log n - \log \varepsilon, \tag{16}$$

where M_m is the size of the codebook for max p.o.e. obtained via expurgation and M_a is for avg p.o.e. Assume that the buffer is empty at the beginning. This gives the worst case scenario, since if the buffer were nonempty at the start, it could only aid communications, and therefore our achievability bound would still hold. The coding scheme that we propose has two phases; namely a saving phase and a transmission phase. This is known in literature as the *save and transmit scheme* (see [12]).

3.1. Saving Phase

In this phase, the transmitter transmits 0, the symbol that uses zero energy, for a set number of slots. During this period, it allows the energy buffer to build up. The receiver is aware of the number of slots and chooses to ignore the output during those slots, since they are not information bearing. The caveat is that slots are wasted, as far as information transfer is concerned, in gathering energy. To ensure that this scheme does not affect the coefficient of the first-order term, it is required that the number of slots set for gathering energy scale at most as $o(n)$.

Fix $0 < \lambda < 1$ and let $K_{\varepsilon,\lambda} = \sqrt{\dfrac{4(2\,\mathbf{E}[E_1]^2 + \sigma_E^2)}{(1-\lambda)\varepsilon\,\mathbf{E}[E_1]^2}}$. Let N_n represent the number of slots reserved for the saving phase. During this phase, the buffer fills up with energy, and after N_n time slots we expect it to have crossed some threshold energy value, which we will denote by E_{0n}. Let $N_n = K_{\varepsilon,\lambda}\sqrt{n}$ (we take the ceiling if N_n is not an integer wherever applicable) and $E_{0n} = N_n\,\mathbf{E}[E_1]/2$. Let $\mathcal{E}_0$ denote the event that the system failed to gather E_{0n} energy. We have

$$\Pr(\mathcal{E}_0) = \Pr\left[\sum_{i=1}^{N_n} E_i \leq E_{0n}\right]$$

$$= \Pr\left[\sum_{i=1}^{N_n}(E_i - \mathbf{E}[E_1]) \leq -E_{0n}\right]$$

$$\leq \Pr\left[\left|\sum_{i=1}^{N_n}(E_i - \mathbf{E}[E_1])\right| \geq E_{0n}\right]$$

$$\leq \frac{4\sigma_E^2}{K_{\varepsilon_n,\lambda}\,\mathbf{E}[E_1]^2\sqrt{n}} \leq \frac{4\sigma_E^2}{K_{\varepsilon,\lambda}\,\mathbf{E}[E_1]^2\sqrt{n}}, \tag{17}$$

where in the last step we used Chebyshev's inequality and also that $K_{\varepsilon,\lambda}$ is monotone decreasing in ε. The above bound ensures a decay of $O(n^{-1/2})$ in the probability of error and hence can be made arbitrarily small by increasing n suitably.

3.2. Encoding and Decoding Scheme

We use a random coding argument where we generate a codebook with M_a codewords and codeword length n, with each entry i.i.d. Gaussian with zero mean and variance $\mathbf{E}[E_1]$. This codebook is also available at the decoder. Denote by $V_i(m)$ the ith symbol of the mth codeword.

At the decoder, we use a form of threshold decoding. Recall the definition of $i_P(x, y)$, where we take W to be the Gaussian channel under consideration and PW to be the output distribution (which is Gaussian with mean 0 and variance $\mathbf{E}[E_1] + \sigma^2$) for the corresponding Gaussian input. The decoding rule is to pick the unique message $\widehat{m}$ such that

$$i_P(\boldsymbol{V}^n(\widehat{m}), \boldsymbol{Y}) \geq \log \gamma_n, \tag{18}$$

where γ_n is a nonnegative number that will be chosen later. This is the same detector that was used for standard AWGN channels.

Note that as far as the transmitter end is concerned, this codebook is not an *energy dependent* codebook as discussed in Section 2.8. This is keeping in line with decoder not having access to the energy arrival process. However, in the next section, we will design the actual transmitted symbol $\boldsymbol{X}^n$ which will be a function of the energy arrivals. Thus, we refer to the above codebook as a *pre-codebook*.

3.3. Transmit Phase

Let n be the number of slots wherein we transmit symbols on the AWGN channel. We count channel uses from the $N_n + 1$ instant onwards. Once we gather at least E_{0n} energy, we must ensure, with high probability, that subsequent transmissions will not cause an outage. Let $\boldsymbol{v}^n$ be the input before checking the energy buffer. At transmission instant i, $1 \leq i \leq n$, there are two cases, i.e.,

1. There is sufficient energy, in which case the input to the channel $x_i = v_i$;
2. There is insufficient energy, in which case we transmit $x_i = 0$.

Assuming message m is to be transmitted, given $\boldsymbol{V}^n(m)$ as in Section 3.2, the corresponding symbol $\boldsymbol{X}^n(m)$ is obtained by the above rules.

Denote the set of sequences $(\boldsymbol{v}^n, \boldsymbol{e}^n)$ that satisfy the above requirements by $\mathcal{A}_n$, where

$$\mathcal{A}_n = \bigcap_{\ell=1}^{n} \left\{ (\boldsymbol{v}^n, \boldsymbol{e}^n) : s_\ell \geq -E_{0n} \right\}, \tag{19}$$

and $s_\ell = \sum_{k=1}^{\ell} e_k - v_k^2$. Note that the transmitted codeword satisfies the energy harvesting conditions, since E_{0n} energy has already been harvested before the transmission started. Denote by $\mathcal{E}_1$ the event that the energy constraints are violated. Let $\{V_i\}$, $1 \leq i \leq n$, be i.i.d. random variables (not necessarily Gaussian) with zero mean, variance $\mathbf{E}[E_1]$, and $\mathbf{E}[V_1^4] < \infty$. Formally,

$$\Pr(\mathcal{E}_1) = \Pr(\mathcal{A}_n^c)$$

$$= \Pr\left[\bigcup_{\ell=1}^{n} \{ S_\ell \leq -E_{0n} \} \right]$$

$$\leq \Pr\left[\bigcup_{\ell=1}^{n} \{ |S_\ell| \geq E_{0n} \} \right]$$

$$= \Pr\left[\max_{1 \leq \ell \leq n} |S_\ell| \geq E_{0n} \right] \tag{20}$$

and $S_\ell = \sum_{k=1}^{\ell} E_k - V_k^2$. Now S_ℓ is a sum of i.i.d. random variables with zero mean and finite variance. We now invoke Kolmogorov's inequality [24, ch. 3], which is stated as follows.

Lemma 1 (Kolmogorov's inequality). *Let Z_i be independent zero mean random variables and $S_n = \sum_{i=1}^{n} Z_i$. If $\mathbf{E}[Z_i] = 0$ and $\mathbf{E}[Z_i^2] < \infty$, then for any $0 < a < \infty$*

$$\Pr\left(\max_{1 \leq i \leq n} |S_i| \geq a \right) \leq \frac{\mathbf{E}[S_n^2]}{a^2}.$$

Hence, we have

$$\Pr(\mathcal{E}_1) \leq \frac{\mathbf{E}[S_n^2]}{E_{0n}^2} = \frac{4(2\mathbf{E}[E_1]^2 + \sigma_E^2)}{K_{\varepsilon_n,\lambda}^2 \, \mathbf{E}[E_1]^2} \leq \frac{4(2\mathbf{E}[E_1]^2 + \sigma_E^2)}{K_{\varepsilon,\lambda}^2 \, \mathbf{E}[E_1]^2}. \tag{21}$$

Unlike (17), the right-hand side above is independent of n. However, by a clever choice of $K_{\varepsilon,\lambda}$, it can be made small enough. Our choice of $K_{\varepsilon,\lambda}$ will ensure that $\Pr(\mathcal{E}_1) \leq (1-\lambda)\varepsilon$. Thus, a total of $N_n + n$ slots are used for both saving and transmission in this scheme.

We would like to remark that the aforementioned results do not assume that V_i is Gaussian and the channel part has no role here except for the input constraint. This means that the above bound holds for non-Gaussian energy harvesting channels with independent inputs satisfying the aforementioned moment constraints.

3.4. Lower Bound Derivation

Let $\mathcal{E}_H = \mathcal{E}_0 \cup \mathcal{E}_1$. Now under average probability of error (see (2)), we see that

$$
\begin{aligned}
P_{e,\text{avg}} &= \frac{1}{M_a} \sum_{i=1}^{M_a} \Pr[\widehat{U} \neq i \,|\, U = i] \\
&= \frac{1}{M_a} \sum_{i=1}^{M_a} \Pr[\widehat{U} \neq i,\, \mathcal{E}_H^c \,|\, U = i] + \Pr[\widehat{U} \neq i,\, \mathcal{E}_H \,|\, U = i] \\
&\leq \frac{1}{M_a} \sum_{i=1}^{M_a} \Pr[\widehat{U} \neq i,\, \mathcal{E}_H^c \,|\, U = i] + \Pr[\mathcal{E}_H].
\end{aligned}
\tag{22}
$$

Note that the above probabilities have an inherent expectation with respect to all codebooks. Now we make an important observation. The pre-codebook, Gaussian channel, and decoder all function independent of the energy harvesting system. The only way that energy harvesting variables come into picture is via $\boldsymbol{Y}^n$, since $Y_i = X_i + Z_i$. Hence, under event $\mathcal{E}_H^c$, it follows that $\boldsymbol{X}^n = \boldsymbol{V}^n$. This is the same as saying that when energy harvesting constraint is met, the symbol from the pre-codebook passes unimpeded, only corrupted by noise.

Consider $\Pr[\widehat{U} \neq i,\, \mathcal{E}_H^c \,|\, U = i]$. As was mentioned earlier, under event $\mathcal{E}_H^c$, the channel part behaves like a standard AWGN channel, and so the error events are the ones corresponding to that. Hence, using the union bound, we have

$$
\begin{aligned}
\Pr[\widehat{U} \neq i,\, \mathcal{E}_H^c \,|\, U = i] &\leq \Pr[\{i_P(\boldsymbol{V}^n(i); \boldsymbol{Y}^n) \leq \log \gamma_n\},\, \mathcal{E}_H^c \,|\, U = i] \\
&\quad + \sum_{\substack{1 \leq j \leq M_a \\ j \neq i}} \Pr[\{i_P(\boldsymbol{V}^n(j); \boldsymbol{Y}^n) \geq \log \gamma_n\},\, \mathcal{E}_H^c \,|\, U = i] \\
&\leq \Pr[i_P(\boldsymbol{V}^n(i); \boldsymbol{V}^n(i) + \boldsymbol{Z}^n) \leq \log \gamma_n \,|\, U = i] \\
&\quad + \sum_{\substack{1 \leq j \leq M_a \\ j \neq i}} \Pr[i_P(\boldsymbol{V}^n(j); \boldsymbol{V}^n(i) + \boldsymbol{Z}^n) \geq \log \gamma_n \,|\, U = i].
\end{aligned}
\tag{23}
$$

The rest of the proof is similar to the bounds derived for an AWGN channel (see [7]), and hence we get the following bound for any $\gamma_n > 0$:

$$
\Pr[\widehat{U} \neq i,\, \mathcal{E}_H^c \,|\, U = i] \leq \Pr\left[\log\left(\frac{W^n(\boldsymbol{V}^n + \boldsymbol{Z}^n \,|\, \boldsymbol{V}^n)}{P_{\boldsymbol{Y}^n}(\boldsymbol{V}^n + \boldsymbol{Z}^n)}\right) \leq \log \gamma_n\right] + \frac{M_a}{\gamma_n}.
\tag{24}
$$

Thus, the first term on right-hand side of (22) is upper bounded by (24). We have already derived an upper bound on $\Pr(\mathcal{E}_H)$ via (17), (21), and the union bound.

We have

$$
\Pr\left[\log\left(\frac{W^n(\boldsymbol{V}^n + \boldsymbol{Z}^n \,|\, \boldsymbol{V}^n)}{P_{\boldsymbol{Y}^n}(\boldsymbol{V}^n + \boldsymbol{Z}^n)}\right) \leq \log \gamma_n\right] = \Pr\left\{\sum_{i=1}^{n} G_i \leq \log \gamma_n\right\},
\tag{25}
$$

where $G_i = \log\left(\dfrac{\boldsymbol{W}(V_i + Z_i \mid V_i)}{\boldsymbol{P}_Y(V_i + Z_i)}\right)$. Note that the G_i are i.i.d. based on the remarks provided earlier. Moreover, we have

$$C_{EG} := \mathbf{E}[G_i] = \frac{1}{2}\log\left(1 + \frac{\mathbf{E}[E_1]}{\sigma^2}\right), \tag{26}$$

$$V_{EG} := \mathrm{Var}(G_i) = \frac{\mathbf{E}[E_1]}{\mathbf{E}[E_1] + \sigma^2}\log_2^2(e). \tag{27}$$

Also, the third moment, $\mathbf{E}[|G_i|^3]$, is finite. To proceed further, we state the Berry–Esseen theorem (see [24, Theorem 6.4.1]).

Lemma 2 (Berry–Esseen theorem). *Let X_i, $1 \le i \le n$, be an i.i.d. sequence of random variables with mean μ, variance $\sigma^2 < \infty$, and $\mathbf{E}[|X_1|^3] < \infty$. Let $S_n = \sum\limits_{i=1}^{n} X_i$. Then we have, for any $x \in \mathbb{R}$,*

$$\left|\mathrm{Pr}\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \le x\right) - \Phi(x)\right| \le C\frac{E|X_1 - \mu|^3}{\sigma^3\sqrt{n}},$$

where $C < 1/2$ (see [25]). Note that the bound is uniform in x.

Let $K = \dfrac{\mathbf{E}[|G_i - \mathbf{E}[G_i]|^3]}{2V_{EG}^{3/2}}$. Applying the Berry–Esseen theorem, we have, for any $u \in \mathbb{R}$,

$$\left|\mathrm{Pr}\left\{\frac{\left(\sum\limits_{i=1}^{n} G_i\right) - nC_{EG}}{\sqrt{nV_{EG}}} \le u\right\} - \Phi(u)\right| \le \frac{K}{\sqrt{n}}.$$

Substituting $u = \dfrac{\log\gamma_n - nC_{EG}}{\sqrt{nV_{EG}}}$, we get

$$\mathrm{Pr}\left\{\sum_{i=1}^{n} G_i \le \log\gamma_n\right\} \le \Phi\left(\frac{\log\gamma_n - nC_{EG}}{\sqrt{nV_{EG}}}\right) + \frac{K}{\sqrt{n}}. \tag{28}$$

Let $\alpha_n = \lambda\varepsilon_n - \dfrac{4\sigma_E^2}{K_{\varepsilon_n,\lambda}\,\mathbf{E}[E_1]^2\sqrt{n}} - \dfrac{2K}{\sqrt{n}}$. In terms of ε, we have

$$\alpha_n \ge \lambda\varepsilon - \frac{4\sigma_E^2}{K_{\varepsilon,\lambda}\,\mathbf{E}[E_1]^2\sqrt{n}} - \frac{2K + \lambda}{\sqrt{n}} := \alpha_n'. \tag{29}$$

Set $\log\gamma_n = nC_{EG} + \sqrt{nV_{EG}}\Phi^{-1}(\alpha_n)$. We pick n large enough to ensure $\alpha_n > 0$. From (22), (24), (25), and (28), we have

$$\log M_a(n, \varepsilon_n) \ge \log\gamma_n - \frac{1}{2}\log n + O(1)$$
$$\ge nC_{EG} + \sqrt{nV_{EG}}\Phi^{-1}(\alpha_n) - \frac{1}{2}\log n + O(1). \tag{30}$$

Using (16) and (29), we get

$$\log M_m(n, \varepsilon) \ge nC_{EG} + \sqrt{nV_{EG}}\Phi^{-1}(\alpha_n') - \log n + O(1), \tag{31}$$

noting that Φ^{-1} is a monotone increasing function.

We further simplify $\Phi^{-1}(\alpha_n')$ using Taylor's theorem. There exists $u \in (\alpha_n', \lambda\varepsilon)$ such that

$$f(\alpha_n') = f(\lambda\varepsilon) + (\alpha_n' - \lambda\varepsilon)f'(u),$$

where $f(x) = \Phi^{-1}(x)$. Note that $f(x)$ has a derivative that is positive, strictly decreasing up to $x = 1/2$, beyond which it increases. Thus, in $(\alpha'_n, \lambda\varepsilon)$, $f'(u) \leq \widehat{f} = \max\{f'(\alpha'_{n_0}), f'(\lambda\varepsilon)\}$, where n_0 is the smallest n for which $\alpha'_n > 0$. Hence we get, with our choice of α'_n, that

$$\log M^*(n, \varepsilon) \geq \log M_m(n, \varepsilon') \geq nC_{EG} + \sqrt{nV_{EG}}\Phi^{-1}(\lambda\varepsilon) - \log(n) + O(1). \tag{32}$$

Let $\widehat{n} = n + N_n$. We have used $\widehat{n}$ slots, out of which n were for data transmission. We will express the result as a function of $\widehat{n}$, the total number of slots used. Hence we have

$$\log M^*(\widehat{n}, \varepsilon) \geq (\widehat{n} - N_n)C_{EG} + \sqrt{nV_{EG}}\Phi^{-1}(\lambda\varepsilon) - \log(\widehat{n} - N_n) + O(1),$$
$$\geq \widehat{n}C_{EG} - K_{\varepsilon,\lambda}\sqrt{\widehat{n}}C_{EG} + \sqrt{nV_{EG}}\Phi^{-1}(\lambda\varepsilon) - \log\widehat{n} + O(1). \tag{33}$$

Note that $\sqrt{n} \leq \sqrt{\widehat{n}}$ and $\sqrt{n} \geq \sqrt{\widehat{n}} - \dfrac{K_{\varepsilon,\lambda}}{2}$; the latter follows from

$$\sqrt{\widehat{n}} = \sqrt{n + K_{\varepsilon,\lambda}\sqrt{n}} = \sqrt{n}\sqrt{1 + \frac{K_{\varepsilon,\lambda}}{\sqrt{n}}} \leq \sqrt{n}\left(1 + \frac{K_{\varepsilon,\lambda}}{2\sqrt{n}}\right) = \sqrt{n} + \frac{K_{\varepsilon,\lambda}}{2}, \tag{34}$$

where we have used $(1+x)^{1/2} \leq 1 + \dfrac{x}{2}$ for $x > 0$. From (33) and (34), we observe that regardless of the sign of $\Phi^{-1}(\lambda\varepsilon)$, the lower bounds obtained differ by a constant which does not depend on n. Putting it all together, we get for $\widehat{n}$ large enough

$$\log M^*(\widehat{n}, \varepsilon) \geq \widehat{n}C_{EG} + \sqrt{\widehat{n}}\left[\sqrt{V_{EG}}\Phi^{-1}(\lambda\varepsilon) - K_{\varepsilon,\lambda}C_{EG}\right] - \log\widehat{n} + O(1).$$

For completeness, the exact bound is

$$\log M^*(\widehat{n}, \varepsilon) \geq \widehat{n}C_{EG} + \sqrt{\widehat{n}}\left[\sqrt{V_{EG}}\Phi^{-1}(\lambda\varepsilon) - K_{\varepsilon,\lambda}C_{EG}\right] - \log\widehat{n} - \log\varepsilon K$$
$$- \frac{K_{\varepsilon,\lambda}}{2} - \sqrt{V_{EG}}\Phi^{-1}(\lambda\varepsilon)\widehat{f}\left[\frac{4\sigma_E^2}{K_{\varepsilon,\lambda}\mathbf{E}[E_1]^2} + 2K + \lambda\right]. \tag{35}$$

This concludes the proof of the achievable bound for Theorem 1.

4. FINITE BLOCKLENGTH ACHIEVABILITY BOUND FOR EH-DMC

We use the same random coding strategy as in the EH-AWGN channel case. Choose any input distribution $P_X \in \mathcal{F}_{\mathbf{E}[E_1]}$. Generate an $M \times n$ matrix with each element distributed i.i.d. with distribution P_X. Now follow the proof exactly as in the achievability of the EH-AWGN channel case, replacing the term X_i^2 with $\Lambda(X_i)$ wherever it is encountered.

In particular, we could substitute $P_X^* \in \Gamma$ (where Γ is the set of capacity achieving input distributions that are contained in $\mathcal{F}_{\mathbf{E}[E_1]}$) to obtain the best bound. If there are many capacity achieving distributions, then $V(P_X^*; W)$ may change with the choice of distribution P_X^*. Hence, consider

$$V_{ED} = \begin{cases} V_{\min} := \min_{P \in \Gamma} V(P; W) & \text{if } \varepsilon \leq \dfrac{1}{2\lambda}, \\ V_{\max} := \max_{P \in \Gamma} V(P; W) & \text{if } \varepsilon > \dfrac{1}{2\lambda}. \end{cases}$$

Putting it all together, we obtain the following achievability bound:

$$\log M^*(\widehat{n}, \varepsilon) \geq \widehat{n}C_{ED} - \sqrt{\widehat{n}}K_{\varepsilon,\lambda}C_{ED} + \sqrt{\widehat{n}V_{ED}}\Phi^{-1}(\lambda\varepsilon) - \log\widehat{n} + O(1), \tag{36}$$

for all $\widehat{n}$ sufficiently large. The exact bound is

$$\log M^*(\widehat{n}, \varepsilon) \geq \widehat{n}C_{ED} - \sqrt{\widehat{n}}K_{\varepsilon,\lambda}C_{ED} + \sqrt{\widehat{n}V_{ED}}\Phi^{-1}(\lambda\varepsilon) - \log\widehat{n} - \log\varepsilon K$$
$$- \frac{K_{\varepsilon,\lambda}}{2} - \sqrt{V_{ED}}\Phi^{-1}(\lambda\varepsilon)\widehat{f}\left[\frac{4\sigma_E^2}{K_{\varepsilon,\lambda}\mathbf{E}[E_1]^2} + 2K + \lambda\right]. \tag{37}$$

5. CONVERSE THEOREMS

In this section, we will provide a general upper bound on finite blocklength rates for energy harvesting channels. We resort to methods used in [7] to derive these new bounds. Then we apply these to the EH-AWGN and the EH-DMC.

We recall the following error probability function $\beta_\alpha(P, Q)$ (see [7]).

Definition 2. Given two distributions P and Q on $\mathcal{X}$, define, for $\alpha \in [0, 1]$,

$$\beta_\alpha(P, Q) := \min Q[T = 1] := \min \int_{\mathcal{X}} P_{T|X}(1 \mid x)\, dQ(x), \tag{38}$$

where the minimum is over all distributions $(P_{T|X})$ of test functions $T: \mathcal{X} \to \{0, 1\}$ such that $P[T = 1] \geq \alpha$.

This function is essentially the type 2 error probability (probability of deciding P when Q is true) when the type 1 error probability is less than $1 - \alpha$.

The meta converse, proved in [7], is one of the tightest known general converse bounds for any channel. There are two versions, one for average probability of error and the other for maximal probability of error. Note that these are single shot bounds and can be naturally extended for blocklength n.

Lemma 3 (meta converse (avg p.o.e.)). *Every* (M, ε) *average probability of error code satisfies*

$$M \leq \sup_{P_X} \frac{1}{\beta_{1-\varepsilon}(P_{XY}, P_X Q_Y)}$$

for any output distribution Q_Y.

Lemma 4 (meta converse (max p.o.e.)). *Every* (M, ε) *maximal probability of error code satisfies*

$$M \leq \frac{1}{\beta_{1-\varepsilon}(P_{Y|X=c(\overline{m})}, Q_Y)} \leq \sup_{x \in \mathbb{F}} \frac{1}{\beta_{1-\varepsilon}(P_{Y|X=x}, Q_Y)}$$

for any output distribution Q_Y *and codewords coming from* $\mathbb{F} \subset \mathcal{X}$, *where* $\mathcal{X}$ *is the input alphabet and* $c(\overline{m})$ *is the codeword of the message* $\overline{m}$ *satisfying*

$$\overline{m} = \arg\min_{m \in [M]} \Pr[\widehat{U} = m \mid U = m] \tag{39}$$

under channel Q_Y.

However, it is not immediately clear as to the technique of incorporating the effects of energy harvesting in the above expression. This is due to the fact that the set $\mathbb{F}$ above, which is the constrained set, changes with energy. Also unlike traditional channels, the codebook will change depending on available energy. Hence, any codeword is of the form $c(m, \boldsymbol{e})$ for message m and energy vector $\boldsymbol{e}$.

Energy Harvesting Converse (General Version). Under the energy harvesting setup described earlier, we obtain the following converse bounds.

Theorem 3. *Given an energy harvesting setup with channel* W *and incoming energy process* $E \sim P_E$ *i.i.d., every* (M, ε) *code (average p.o.e.) satisfies*

$$M \leq \sup_{P_{X^n|E^n}} \frac{1}{\beta_{1-\varepsilon}(P_{E^n X^n Y^n}, P_{E^n X^n} Q_{Y^n})}, \tag{40}$$

where $P_{E^n X^n Y^n}(e^n, x^n, y^n) = P_{E^n}(e^n) P_{X^n|E^n}(x^n|e^n) W(y^n|x^n)$, *for any output distribution* Q_{Y^n}. *The supremum is over all distributions that satisfy the energy harvesting constraints. Under the maximal probability of error case, we have*

$$M \leq \frac{1}{\beta_{1-\varepsilon}\left(W(\cdot\,|\,c(\overline{m},*))P_{E^n}(*), Q_{Y^n}P_{E^n}\right)} \tag{41}$$

for any output distribution Q_{Y^n} *and* $c(\overline{m},*)$ *is the codeword whose message* $\overline{m}$ *satisfies* (39). *Here* $\cdot$ *represents the output alphabet and* $*$ *represents the energy alphabet.*

Proof. The proof of (40) is available in [14]. For the proof of (41), refer to Appendix A. $\triangle$

The bound in (40) was used to develop a finite blocklength converse for EH-AWGN channels extended to the block i.i.d. energy arrivals regime [14]. We shall derive the same result for EH-AWGN channels under maximal probability of error criterion but using (41).

There is a weaker, but analytically convenient, converse bound under maximal p.o.e. stated as follows.

Theorem 4. *Consider an energy harvesting setup with channel* W, *incoming energy process* $E \sim P_E$ *i.i.d., and cost function* Λ *as defined in Section* 2.7. *Under the requirement that every codeword* $\boldsymbol{x}(m, \boldsymbol{e}^n)$ *satisfying the energy harvesting constraint, i.e.,*

$$\sum_{i=1}^{n} \Lambda(x_i(m, \boldsymbol{e}^n)) \leq \sum_{i=1}^{n} e_i, \tag{42}$$

for energy vector $\boldsymbol{e}^n$ *and maximal probability of error* ε, *we have*

$$M \leq \sup_{\boldsymbol{x}^n \in \mathbb{F}_{\overline{E}_n}} \frac{1}{\beta_{1-\varepsilon-\tau_n}\left(W(\cdot\,|\,\boldsymbol{x}^n), Q_{Y^n}\right)}, \tag{43}$$

where $\tau_n = \Pr\left(\sum\limits_{i=1}^{n} E_i \geq n\overline{E}_n\right)$,

$$\mathbb{F}_{\overline{E}_n} = \left\{\boldsymbol{x}^n : \sum_{i=1}^{n} \Lambda(x_i) \leq n\overline{E}_n\right\}, \tag{44}$$

and $\overline{E}_n$ *is a nonnegative sequence chosen such that* $\tau_n < 1 - \varepsilon$.

Proof. See Appendix B. $\triangle$

There is a nice structure for EH-AWGN channels that helps in getting sharper bounds when using (40) or (41). These details are clarified in the proof of the converse bound for EH-AWGN channels. However, that structure is absent when dealing with EH-DMCs. Theorem 4 will be used to get a useful upper bound in this case.

6. FINITE BLOCKLENGTH CONVERSE BOUND FOR EH-AWGN

We argue that it suffices to look at codewords $\boldsymbol{x}^n$ that satisfy

$$\sum_{k=1}^{n} x_k^2 = \sum_{k=1}^{n} e_k, \tag{45}$$

where $\boldsymbol{e}^n$ is the energy vector. In short, we are ignoring the outage events that can happen for $1 \leq k < n$ and we are using up all the energy in transmission at time n. The former is justified by noting that doing so merely relaxes the constraints and that can only increase capacity. Hence, any upper bound on the relaxed version is an upper bound on the original version. As for the

latter, it is a well known Yaglom-map trick where given the best code of codeword length n but satisfying (45) with a strict inequality ($<$), we can construct a new code with the same probability of error but with codeword length $n+1$. The extra symbol is picked so as to exhaust all remaining energy. This new code clearly satisfies (45), is an upper bound for the original length-n code, and is further upper bounded by the largest code of codeword length $n+1$ satisfying (45).

Let $0 < \varepsilon < 1$, the maximal probability of error, be fixed. Pick W as a Gaussian channel with variance σ^2 and $Q_{Y^n} = \prod_{i=1}^{n} Q_Y$, where Q_Y is Gaussian with mean 0 and variance $\mathbf{E}[E_1] + \sigma^2$. Now for two distributions P_1 and P_2 and any $\gamma > 0$, $\beta_\alpha(P_1, P_2)$ is lower bounded as (see [7, equation 106])

$$\beta_\alpha(P_1, P_2) \geq \frac{1}{\gamma}\left(\alpha - P_1\left[\frac{dP_1}{dP_2} \geq \gamma\right]\right). \tag{46}$$

From (41) and (46), we have, for any $\gamma_n > 0$,

$$M \leq \frac{\gamma_n}{1 - \varepsilon - \Pr\left[\log \dfrac{W(\boldsymbol{Y}^n \mid \boldsymbol{x}^n(\overline{m}, \boldsymbol{E}))}{Q_{Y^n}} \geq \log \gamma_n\right]}, \tag{47}$$

where the probability is under $W(\,\cdot\, \mid \boldsymbol{x}(\overline{m}, *))P_{E^n}(*)$. Since W is a Gaussian channel here, we can replace Y_i with $x_i(\overline{m}, \boldsymbol{e}) + Z_i$ where Z_i are i.i.d. $\mathcal{N}(0, \sigma^2)$. The probability term in the denominator then simplifies to

$$\Pr\left[\log \frac{W(\boldsymbol{Y}^n \mid \boldsymbol{x}^n(\overline{m}, \boldsymbol{E}))}{Q_{Y^n}} \geq \log \gamma_n\right]$$

$$= \Pr\left[\sum_{i=1}^{n} \frac{(x_i(\overline{m}, \boldsymbol{E}) + Z_i)^2}{2(\mathbf{E}[E_1] + \sigma^2)} \log_2(e) - \sum_{i=1}^{n} \frac{Z_i^2}{2\sigma^2} \log_2(e) \geq \log(\gamma_n) - nC_{EG}\right]$$

$$= \Pr\left[\sum_{i=1}^{n}\left(\frac{Z_i}{\sigma} - \frac{x_i(\overline{m}, \boldsymbol{E})\sigma}{\mathbf{E}[E_1]}\right)^2 \leq \frac{2(\mathbf{E}[E_1] + \sigma^2)}{\mathbf{E}[E_1]}(nC_{EG} - \log \gamma_n)\ln 2\right.$$

$$\left. + \sum_{i=1}^{n} x_i^2(\overline{m}, \boldsymbol{E})\left(\frac{\sigma^2}{\mathbf{E}[E_1]^2} + \frac{1}{\mathbf{E}[E_1]}\right)\right]$$

$$= \Pr\left[\sum_{i=1}^{n}\left(\frac{Z_i}{\sigma} - \frac{x_i(\overline{m}, \boldsymbol{E})\sigma}{\mathbf{E}[E_1]}\right)^2 \leq \frac{2(\mathbf{E}[E_1] + \sigma^2)}{\mathbf{E}[E_1]}(nC_{EG} - \log \gamma_n)\ln 2\right.$$

$$\left. + \sum_{i=1}^{n} E_i\left(\frac{\sigma^2}{\mathbf{E}[E_1]^2} + \frac{1}{\mathbf{E}[E_1]}\right)\right], \tag{48}$$

where (48) follows from (45). Now, we condition the above probability term on $\boldsymbol{E} = \boldsymbol{e}$, noting that $\boldsymbol{E}$ is independent of $\boldsymbol{Z}$. We observe then that the probability is the cumulative distribution function (CDF) of a noncentral χ^2 distribution with n degrees of freedom and noncentrality parameter

$$B = \sum_{i=1}^{n} \frac{x_i^2(\overline{m}, \boldsymbol{e})\sigma^2}{\mathbf{E}[E_1]^2} = \sum_{i=1}^{n} \frac{e_i \sigma^2}{\mathbf{E}[E_1]^2}. \tag{49}$$

The CDF of a noncentral χ^2 random variable $\widehat{Z}$ equals

$$\Pr(\widehat{Z} \leq u) = 1 - Q_{n/2}^M(\sqrt{B}, \sqrt{u}), \tag{50}$$

where $Q_d^M(a, b)$ is the Marcum Q function of order d (see [26]). Now we observe that the CDF does not depend on the individual x_i or e_i but rather on the sum of e_i. Replacing $x_i(\overline{m}, \boldsymbol{E})$ with $\sqrt{E_i}$

in (48) will not change the CDF. Hence, from (49) and (50), (48) equals

$$\Pr\left[\sum_{i=1}^{n}\left(\frac{Z_i}{\sigma}-\frac{\sqrt{E_i}\sigma}{\mathbf{E}[E_1]}\right)^2 \leq \frac{2(\mathbf{E}[E_1]+\sigma^2)}{\mathbf{E}[E_1]}(nC_{EG}-\log\gamma_n)\ln 2+\sum_{i=1}^{n}E_i\left(\frac{\sigma^2}{\mathbf{E}[E_1]^2}+\frac{1}{\mathbf{E}[E_1]}\right)\right]. \quad (51)$$

This is precisely the structure mentioned earlier that allows us to work with a simplified expression. As a result, the terms in the summation are i.i.d. (as opposed to just being independent). By suitably rearranging the terms, (51) equals

$$\Pr\left[\frac{\sum\limits_{i=1}^{n}\eta_i}{\sqrt{nV_{EG2}}}\leq\frac{nC_{EG}-\log\gamma_n}{\sqrt{nV_{EG2}}}\right], \quad (52)$$

where η_i are i.i.d. with zero mean and variance $V_{EG2}=\dfrac{\mathbf{E}[E_1]^2+\mathbf{E}[E_1^2]+4\sigma^2\,\mathbf{E}[E_1]}{4(\mathbf{E}[E_1]+\sigma^2)^2}\log_2^2(e)$. The third moment of η_i is finite. Applying the Berry–Esseen theorem (Lemma 2) and picking $\log\gamma_n=nC_{EG}-\sqrt{nV_{EG2}}\Phi^{-1}(\alpha_n)$, where α_n is picked such that $0<\alpha_n<1-\varepsilon$, gives us

$$\Pr\left[\frac{\sum\limits_{i=1}^{n}\eta_i}{\sqrt{nV_{EG2}}}\leq\frac{nC_{EG}-\log\gamma_n}{\sqrt{nV_{EG2}}}\right]\leq\alpha_n+\frac{\varkappa}{\sqrt{n}}, \quad (53)$$

where $\varkappa=\mathbf{E}[|\eta_i|^3]/V_{EG2}^{3/2}$.

Pick $\alpha_n=1-\varepsilon-\dfrac{2\varkappa}{\sqrt{n}}$. For n sufficiently large, $0<\alpha_n<1-\varepsilon$. From (47), (51), and (53), we get

$$\log M\leq nC_{EG}-\sqrt{nV_{EG2}}\Phi^{-1}(\alpha_n)-\log(\varkappa/\sqrt{n}).$$

Using Taylor series expansion on Φ^{-1} as well as bounding steps similar to the proof of achievability of Theorem 1, we obtain

$$\log M\leq nC_{EG}+\sqrt{nV_{EG2}}\Phi^{-1}(\varepsilon)+\frac{1}{2}\log n+O(1),$$

which gives us the desired converse bound.

For completeness, the exact bound is

$$\log M\leq nC_{EG}+\sqrt{nV_{EG2}}\Phi^{-1}(\varepsilon)+\frac{1}{2}\log n-\log\varkappa+2\varkappa\sqrt{V_{EG2}}\,\widehat{f_2}, \quad (54)$$

where $\widehat{f_2}=\max\left\{f(\varepsilon),f\left(\varepsilon+\dfrac{2\varkappa}{\sqrt{n_0}}\right)\right\}$, $f(x)=\dfrac{d}{dx}\Phi^{-1}(x)$, and n_0 is the smallest n for which $\varepsilon+\dfrac{2\varkappa}{\sqrt{n}}<1$.

7. FINITE BLOCKLENGTH CONVERSE BOUND FOR EH-DMC

Unfortunately, we cannot simply mirror the proof of the EH-AWGN channel converse in Section 6, since the AWGN channel structure that was exploited there is absent here. However, there is a different structure that can be exploited here, namely the method of types (see [21]). We will be using the framework of Theorem 4. Let $0<\varepsilon<1$ be given and the DMC of the EH-DMC be denoted by $W(y\,|\,x)$. The incoming energy random variables E_i are i.i.d. as before.

Recall the definitions given in (7) and (8). We have, from (43),

$$M\leq\sup_{\boldsymbol{x}^n\in\mathbb{F}_{\overline{E}_n}}\frac{1}{\beta_{1-\varepsilon-\tau_n}\left(W(\,\cdot\,|\,\boldsymbol{x}^n),Q_{Y^n}\right)}. \quad (55)$$

We pick $\overline{E}_n = \mathbf{E}[E_1] + \delta_n$, where $\delta_n > 0$. Then τ_n is given by

$$\tau_n = \Pr\left(\sum_{i=1}^{n} E_i \geq n(\mathbf{E}[E_1] + \delta_n)\right). \tag{56}$$

We will ensure that $\tau_n \leq \frac{\varepsilon}{4}$. To do this, pick $\delta_n = \frac{D_\varepsilon}{\sqrt{n}}$, where $D_\varepsilon = \sqrt{\frac{4\sigma_E^2}{\varepsilon}}$, and use Chebyshev's inequality.

We can rewrite (55) as

$$M \leq \sup_{P \in \mathcal{F}_{\overline{E}_n} \cap \mathcal{P}_n} \sup_{\boldsymbol{x}^n \in T_P} \frac{1}{\beta_{1-\varepsilon-\tau_n}\left(W(\cdot \,|\, \boldsymbol{x}^n), Q_{Y^n}\right)}, \tag{57}$$

where T_P denotes the type class of distribution P and $\mathcal{P}_n$ is the set of all types for sequences of length n. Consider the inner supremum term,

$$\sup_{\boldsymbol{x}^n \in \mathcal{T}_P} \frac{1}{\beta_{1-\varepsilon-\tau_n}(W(\cdot \,|\, \boldsymbol{x}^n), Q_{Y^n})}.$$

The beta error function above is independent of which sequence $\boldsymbol{x}$ is picked provided that the sequences have the same type [7] and $Q_{Y^n} = \prod_{k=1}^{n} Q_Y$ for some distribution Q_Y on $\mathcal{Y}$. Hence, pick any sequence $\boldsymbol{x}$ from $\mathcal{T}_{P_0}$, where $P_0 \in \mathcal{F}_{\overline{E}_n} \cap \mathcal{P}_n$.

Let $Q_Y = P_0 W$. We recall [7, Theorem 48] for standard, nonexotic DMCs. Although this bounded the maximal subcode of type P_0 of the maximal code, we note that the term actually being bounded is the beta error function as mentioned below.

Lemma 5. *For $0 < \varepsilon < 1$, for all $P_0 \in \mathcal{P}_n$, $\boldsymbol{x} \in \mathcal{T}_{P_0}$, and n sufficiently large, we have*

$$-\log \beta_{1-\varepsilon}(W^n(\cdot \,|\, \boldsymbol{x}), (P_0 W)^n) \leq nC_D + \sqrt{nV_D}\Phi^{-1}(\varepsilon) + \frac{1}{2}\log n + O(1),$$

where

$$V_D = \begin{cases} V_{\min} = \min\limits_{P \in \Gamma} V(P; W), & 0 < \varepsilon \leq 1/2, \\ V_{\max} = \max\limits_{P \in \Gamma} V(P; W), & 1/2 < \varepsilon < 1, \end{cases}$$

and Γ is the set of capacity achieving distributions.

Note that the bound on the right-hand side does not depend on the distribution of the type. Hence, if we make the following substitutions:

1. Replace Γ with

$$\Gamma_{\overline{E}_n} = \{P \in \mathcal{F}_{\overline{E}_n} : I(P; W) = C_{ED}\}. \tag{58}$$

   This is because the outer supremum in (57) is over $\mathcal{F}_{\overline{E}_n}$. Note that the original proof of Lemma 5 used the fact that Γ was compact and convex. These properties hold for $\Gamma_{\overline{E}_n}$, so we may substitute this wherever Γ was used;
2. The final supremum that gives the uniform (over input distributions) bound was over $\mathcal{P}$. Here we substitute $\mathcal{F}_{\overline{E}_n}$ in its place;
3. ε is replaced by $\varepsilon + \tau_n$;

then

$$\log M^*(n, \varepsilon) \leq nC_D(\overline{E}_n) + \sqrt{n\widehat{V}(\overline{E}_n)}\Phi^{-1}(\varepsilon + \tau_n) + O(\log(n)), \tag{59}$$

where $C_D(\cdot)$ is defined in (9) and

$$\widehat{V}(\overline{E}_n) = \begin{cases} V_{\min}^{(n)} = \min_{P \in \Gamma_{\overline{E}_n}} V(P; W), & 0 < \varepsilon + \tau_n \leq 1/2, \\ V_{\max}^{(n)} = \max_{P \in \Gamma_{\overline{E}_n}} V(P; W), & 1/2 < \varepsilon + \tau_n < 1. \end{cases} \tag{60}$$

We can further simplify (59) by expanding $C_D(\overline{E}_n)$, $\widehat{V}(\overline{E}_n)$, and $\Phi^{-1}(u)$.

Now $C_D(a)$ is a nondecreasing concave function (see [21]). Hence we have for any $a > 0$, $b > 0$,

$$C_D(a + b) \leq C_D(a) + bC'_D(a),$$

where $C'_D(\cdot)$ is the derivative of $C_D(a)$. Let $a = \mathbf{E}[E_1]$ and $b = \delta_n$. Note that $C'_D(a)$ in this case is a constant, since $\mathbf{E}[E_1]$ is a constant.

Using Taylor series expansion, we get that for some constant K_ε,

$$\Phi^{-1}(\varepsilon + \tau_n) \leq \Phi^{-1}(\varepsilon) + \tau_n K_\varepsilon.$$

Now let ε_R be the root of

$$\Phi^{-1}(\varepsilon) + \frac{K_\varepsilon \varepsilon}{4} = 0.$$

Pick any $\eta > 0$. Observe that for n sufficiently large, $\Gamma_{\overline{E}_n} \subset \Gamma_{\mathbf{E}[E_1]+\eta}$. Hence, we can replace $\widehat{V}(\overline{E}_n)$ with

$$V_\varepsilon^*(\eta) = \begin{cases} \min_{P \in \Gamma_{\mathbf{E}[E_1]+\eta}} V(P; W), & 0 < \varepsilon \leq \varepsilon_R, \\ \max_{P \in \Gamma_{\mathbf{E}[E_1]+\eta}} V(P; W), & \varepsilon_R < \varepsilon < 1. \end{cases}$$

Note that $C_D(\mathbf{E}[E_1]) \equiv C_{ED}$. Thus, we have for n sufficiently large

$$\log M^*(n, \varepsilon) \leq nC_{ED} + \sqrt{n}C'(\mathbf{E}[E_1])D_\varepsilon + \sqrt{nV_\varepsilon^*(\eta)}\left(\Phi^{-1}(\varepsilon) + \frac{K_\varepsilon \varepsilon}{4}\right) + O(\log n).$$

Unlike the earlier cases, we do not provide an exact bound here. This is due to the original bounds in [7] for DMC being described with unknown constants in the $O(1)$ term.

8. INITIAL AND RESIDUAL ENERGY

Our analysis concerns the transmission of a single message, which is the usual case in information theory [19] as well as information theoretic treatments of energy harvesting [11, 12]. In practice, multiple messages are transmitted and in the beginning of transmission of some messages, there will be some random residual energy. If the buffer length is finite, then with positive probability the residual energy will be zero. Thus, our bounds correspond to the worst case scenario where there is no initial energy in the buffer. Secondly, if the energy buffer had some residual energy e_0, then if $e_0 > E_{0n}$, we do not need to wait until we get E_{0n} energy. However, to take into account the general case of multiple transmissions, for any e_0, we continue to have a block of N_n slots before the transmission starts (otherwise, due to random e_0, we will need to change the slot structure and coding/decoding strategies for each message transmission). The following steps are to be noted while deriving the lower bound.

1. We gather energy for N_n slots as before in the saving phase. The only change is that the energy threshold is increased to $E_{0n}+e_0$. However, there is no change in (17), since we are still targeting to gather E_{0n} energy as before;

2. V_i is now chosen Gaussian with variance $\mathbf{E}[E_1] + e_0/n$. In (19), $\mathcal{A}_n$ is replaced with $\mathcal{A}'_n$ which is the same as $\mathcal{A}_n$ but with E_{0n} replaced with $E_{0n} + e_0$. Let

$$\mathcal{B}_n = \bigcap_{\ell=1}^{n} \left\{ (\boldsymbol{v}^n, \boldsymbol{e}^n) : s'_\ell \geq -E_{0n} \right\}, \tag{61}$$

where $s'_\ell = \sum_{k=1}^{\ell} (e_\ell - v_\ell^2 + e_0/n)$. Clearly, each term in the sum is i.i.d. zero mean as required for Kolmogorov's inequality and $\mathcal{B}_n \subset \mathcal{A}'_n$. There are no changes to the bound on $\Pr(\mathcal{E}_1)$.

Keeping in mind the aforementioned changes, it can be shown that

$$\log M^*(n, \varepsilon) \geq nC_{EG}^{(n)} + \sqrt{n}\left[\sqrt{V_{EG}^{(n)}} \Phi^{-1}(\lambda\varepsilon) - K_{\varepsilon,\lambda} C_{EG}^{(n)} \right] - \log n + O(1), \tag{62}$$

where $C_{EG}^{(n)} = \frac{1}{2}\log\left(1 + \frac{\mathbf{E}[E_1] + e_0/n}{\sigma^2}\right)$ and $V_{EG}^{(n)} = \frac{\mathbf{E}[E_1] + e_0/n}{\mathbf{E}[E_1] + \sigma^2 + e_0/n}\log_2^2(e)$. The exact bound is

$$\log M^*(n, \varepsilon) \geq nC_{EG} + \sqrt{n}\left[\sqrt{V_{EG}} \Phi^{-1}(\lambda\varepsilon) - K_{\varepsilon,\lambda} C_{EG} \right] - \log n + \frac{e_0 K_{\varepsilon,\lambda}}{2\sigma^2 \sqrt{n}}$$
$$- \log \varepsilon K - \frac{K_{\varepsilon,\lambda}}{2} - \sqrt{V_{EG}} \Phi^{-1}(\lambda\varepsilon) \widehat{f}\left[\frac{4\sigma_E^2}{K_{\varepsilon,\lambda} \mathbf{E}[E_1]^2} + 2K + \lambda \right]. \tag{63}$$

The bound for EH-DMC is similar to (36) with $\mathbf{E}[E_1]$ replaced with $\mathbf{E}[E_1] + e_0/n$. Note that we could further simplify to remove the e_0/n terms from the first- and second-order terms, thereby collecting them in $O(1)$. However, our emphasis here was on the effects of initial energy which is that the bounds improve, as does the achievable rate.

Considering converse bounds, assume that the energy buffer is pre-filled with some deterministic energy, say e_0, before communication starts. As an example, consider the derivation of EH-AWGN converse. We replace $\sum_{k=1}^{n} e_k$ in (45) with $e_0 + \sum_{k=1}^{n} e_k$. Let $\overline{E}_n = \mathbf{E}[E_1] + e_0/n$. We take Q_Y as $\mathcal{N}(0, \overline{E}_n)$. Similar steps until (48) follow except for the above substitutions. We replace E_i with $E_i + e_0/n$ in (51), which does not change the CDF in this case. This would yield

$$\log M^*(n, \varepsilon) \leq nC_{EG}^{(n)} + \sqrt{nV_{EG2}^{(n)}} \Phi^{-1}(\varepsilon) + \frac{1}{2}\log n - \log \varkappa + 2\varkappa\sqrt{V_{EG2}^{(n)}} \widehat{f}_2, \tag{64}$$

where

$$V_{EG2}^{(n)} = \frac{2\overline{E}_n^2 + \sigma_E^2 + 4\sigma^2 \overline{E}_n}{4(\overline{E}_n + \sigma^2)^2}\log_2^2(e).$$

Now, $C_{EG}^{(n)} \leq C_{EG} + \frac{1}{2}\log\left(1 + \frac{e_0}{n\sigma^2}\right)$ and $|V_{EG2}^{(n)} - V_{EG2}| \leq \frac{c_v e_0}{n}$ for some nonnegative constant c_v. Thus, the upper bound is only affected by an $O(1)$ term. This does not mean that initial energy has little to no impact but rather that with increasing n, its impact reduces significantly; e.g., if $e_0 = O(n)$, then it would affect the first-order term. For the EH-DMC converse, a similar change is made where we replace E_i with $E_i + e_0/n$. The e_0 term gets absorbed in the definition of δ_n as defined in (56). Once again, it only yields $O(1)$ changes, which still allow our bounds to hold. We therefore conclude that initial deterministic energy in the buffer does not affect the first- or second-order terms in the achievable and converse bounds.

Residual Energy. The above discussion assumed a fixed amount of initial energy. In practice, when multiple messages are transmitted, one after another, then the left-over energy after each successful transmission would vary randomly. Consider an EH-AWGN channel setup as before with n,

the slots for transmission, and maximal probability of error ε fixed. Each message undergoes a saving and transmission phase as before. In the following we assume that if during the transmission of a message we run out of energy at some point, then the transmission of that message stops and this is conveyed to the receiver. However, the harvested energy during the *rest* of that slot is stored, to be used for the transmission of the next message. In the following we denote the residual energy after the transmission of the ℓth slot by R_ℓ. We obtain lower bounds, r_ℓ, on it.

Define the random sequence r_ℓ as follows. Let $r_0 = e_0$ for initial energy e_0 in the beginning of slot $\ell = 1$. Then r_ℓ, for $\ell \geq 1$ evolves as

$$r_\ell = (r_{\ell-1} + \zeta_\ell)^+, \tag{65}$$

where $\zeta_\ell = \sum_{k=1}^{n+N_n} E_{(\ell-1)(n+N_n)+k} - \sum_{j=1}^{n} \Lambda(X_{(\ell-1)(n+N_n)+j})$, and $\Lambda(\cdot)$ is the energy function as discussed in EH-DMC and for EH-AWGN, $\Lambda(x) = x^2$. Also $\mathbf{E}[\zeta_\ell] = (n + N_n)\,\mathbf{E}[E_1] - nP$ for $1 \leq \ell \leq L$, where we assume that the channel inputs X_i are i.i.d. random variables with finite mean and $\mathbf{E}[\Lambda(X_i)] = P$, and P is to be decided. It can be shown that $r_\ell \leq R_\ell$ for $\ell \geq 1$. In the study of GI/GI/1 queues [27], (65) is the well-known Lindley equation. Thus, r_ℓ will have a stationary distribution if $\mathbf{E}[\zeta_1] < 0$. For our purposes, we do not gain anything by working in this regime. Instead, we will actually take $\mathbf{E}[\zeta_1] > 0$ but close to 0 and therefore $P \leq \left(1 + \dfrac{N_n}{n}\right)\mathbf{E}[E_1]$. This way we will stay within the framework used so far and use the above results. Moreover, the residual energy r_ℓ, will diverge to ∞ as ℓ increases. We have from [28] that r_ℓ obeys

$$\frac{r_\ell}{\ell} \xrightarrow{p} \mathbf{E}[\zeta_1], \qquad \frac{r_\ell - \ell\,\mathbf{E}[\zeta_1]}{\sqrt{\ell}\sigma_\zeta} \xrightarrow{d} \mathcal{N}(0,1) \tag{66}$$

as ℓ tends to infinity, where $\sigma_\zeta^2 = \mathrm{Var}(\zeta_1)$. Thus, for ℓ sufficiently large, one could resort to setting $e_0 = r_\ell \approx \ell\,\mathbf{E}[\zeta_1] - a\sqrt{\ell}\sigma_\zeta$, where a is chosen large enough so that $\dfrac{r_\ell - \ell\,\mathbf{E}[\zeta_1]}{\sqrt{\ell}\sigma_\zeta} \geq -a$ is a high probability event. Using the e_0 dependent bounds mentioned earlier, we could derive a lower bound that considers residual energy. Further refinements on this approximation are available, but for sake of brevity we skip those details.

9. MODERATE DEVIATION ASYMPTOTICS

In this section, we discuss the bounds on the moderate deviation asymptotics for the EH-AWGN channel and the EH-DMC. In this analysis, unlike in the second-order analysis in the previous sections, we allow probability of error to go to zero as a function of blocklength n. However, we do so in the moderate deviations regime which is defined formally as follows (see [16]).

Definition 3 (moderate deviation coefficient). Given a channel W, let ρ_n be a sequence of nonnegative real numbers such that $\rho_n \to 0$ and $n\rho_n^2 \to \infty$. Then for codes of size M_n satisfying $\log M_n = n(C - \rho_n)$, where C is the channel capacity, the moderate deviations coefficient (MDC) ξ, if it exists, is defined as

$$\xi = \lim_{n \to \infty} \frac{\log \varepsilon(n)}{n\rho_n^2},$$

where $\varepsilon(n)$ is the probability of error as a function of blocklength n.

For memoryless channels with channel dispersion $V > 0$, it was shown in [16] that $\xi = -\dfrac{1}{2V}$ is the moderate deviation coefficient. In the case of energy harvesting channels, it is more involved. This is due to not knowing the exact dispersion value as well as the fact that energy harvesting channels are not truly memoryless due to the energy vector. However, they have a part which is memoryless, and this is what we have been exploiting so far in our analysis.

9.1. MDC for EH-AWGN Channels

We now state the following theorem bounding the MDC for EH-AWGN channels.

Theorem 5. *For an EH-AWGN channel with energy process E_i i.i.d. with variance σ_E^2, the MDC is bounded as*

$$\liminf_{n \to \infty} \frac{\log \varepsilon(n)}{n\rho_n^2} \geq -\frac{1}{2V_{EG2}}, \tag{67}$$

$$\limsup_{n \to \infty} \frac{\log \varepsilon(n)}{n\rho_n^2} \leq -\frac{1}{2V_{EG}}, \tag{68}$$

where V_{EG} is defined in (11) and V_{EG2} is defined in (12).

Proof. To show (67), let us consider (47) whose terms are rearranged, replacing ε with $\varepsilon(n)$, as

$$\varepsilon(n) \geq \Pr\left[\log \frac{W(\boldsymbol{Y}^n \mid \boldsymbol{x}^n(\overline{m}, \boldsymbol{E}))}{Q_{Y^n}} \leq \log \gamma_n\right] - \frac{\gamma_n}{M}.$$

We also have from (52) that

$$\Pr\left[\log \frac{W(\boldsymbol{Y}^n \mid \boldsymbol{x}^n(\overline{m}, \boldsymbol{E}))}{Q_{Y^n}} \leq \log \gamma_n\right] = \Pr\left[\sum_{i=1}^{n} \eta_i \geq nC_{EG} - \log \gamma_n\right].$$

Now let $\log M = n(C_{EG} - \rho_n)$ and $\log \gamma_n = n(C_{EG} - \alpha\rho_n)$ for any $\alpha > 1$. From [29, Theorem 3.7.1], we get

$$\liminf_{n \to \infty} \frac{\log \Pr\left[\sum\limits_{i=1}^{n} \eta_i \geq nC_{EG} - \log \gamma_n\right]}{n\rho_n^2} \geq -\inf_{x \geq \alpha} \frac{x^2}{2V_{EG2}} = -\frac{\alpha^2}{2V_{EG2}},$$

where noting that V_{EG2} is the variance of η_i and letting $\alpha \to 1$, we get (67).

To show (68), we need to modify some of our arguments which we used while discussing the save and transmit scheme. This is because we need to show that codes of $\log M = n(C_{EG} - \rho_n)$ exist. The analysis so far was done so as to work with the optimum order of $\sqrt{n}$. This is not valid anymore, since $\rho_n > 1/\sqrt{n}$.

Recalling error events $\mathcal{E}_0$ from (17) and $\mathcal{E}_1$ from (20), we will show that with an appropriate choice for N_n and E_{0n}, we can set

$$\Pr(\mathcal{E}_0) + \Pr(\mathcal{E}_1) \leq \frac{\varepsilon(n)}{2}.$$

To ensure this, let us choose

$$N_n = \max\left\{\frac{16\sigma_E^2}{\varepsilon(n)\,\mathbf{E}[E_1]^2}, \frac{4\sqrt{n(2\,\mathbf{E}[E_1]^2 + \sigma_E^2)}}{\mathbf{E}[E_1]\sqrt{\varepsilon(n)}}\right\}.$$

Clearly, $N_n \to \infty$ as $n \to \infty$, and both $\Pr(\mathcal{E}_0)$ and $\Pr(\mathcal{E}_1)$ are each upper bounded by $\varepsilon(n)/4$.

Hence, the probability of error $\varepsilon(n)$ is bounded by

$$\varepsilon(n) \leq \frac{\varepsilon(n)}{2} + \Pr\left[\log\left(\frac{W^n(\boldsymbol{Y}^n \mid \boldsymbol{X}^n)}{P_{\boldsymbol{Y}^n}(\boldsymbol{Y}^n)}\right) \leq \log \gamma_n\right] + \frac{M}{\gamma_n},$$

$$\frac{\log(\varepsilon(n)/2)}{n\rho_n^2} \leq \frac{1}{n\rho_n^2} \log\left[\Pr\left\{\sum_{i=1}^{n} G_i \leq \log \gamma_n\right\} + 2^{-(1-\alpha)n\rho_n}\right].$$

Now let $\log \gamma_n = n(C_{EG} - \alpha \rho_n)$, where $\alpha < 1$ and $\log M = n(C_{EG} - \rho_n)$. Codes of the latter size are assured by Feinstein's lemma. Now from (25) and [29, Theorem 3.7.1], we have

$$\limsup_{n \to \infty} \frac{1}{n\rho_n^2} \log \Pr\left\{ \sum_{i=1}^{n} G_i \leq \log \gamma_n \right\} \leq - \inf_{x \leq -\alpha} \frac{x^2}{2V_{EG}} = -\frac{\alpha^2}{2V_{EG}}. \tag{69}$$

Letting $\alpha \to 1$, we get (68). $\triangle$

9.2. MDC for EH-DMC

The MDC bounds for EH-DMC should be analogous to that of the EH-AWGN channel. However, since V_{ED} varies with the choice of λ, we need to refine it slightly.

Theorem 6. *For the EH-DMC, the following bounds on MDC apply*:

$$\liminf_{n \to \infty} \frac{\log \varepsilon(n)}{n\rho_n^2} \geq - \inf_{\eta > 0} \frac{1}{2V_{\min,\eta}}, \tag{70}$$

$$\limsup_{n \to \infty} \frac{\log \varepsilon(n)}{n\rho_n^2} \leq -\frac{1}{2V_{\min}}, \tag{71}$$

where $V_{\min} = \min\limits_{P \in \Gamma_{\mathbf{E}[E_1]}} V(P;W)$ *and* $V_{\min,\eta} = \min\limits_{P \in \Gamma_{\mathbf{E}[E_1]+\eta}} V(P;W)$, *where* Γ *is the set of capacity achieving input distributions that are in* $\mathcal{F}_{\mathbf{E}[E_1]}$.

Proof. Bound (70) follows from [16, Theorem 6] with the following changes:

1. The distributions need to be admissible, i.e., from $\mathcal{F}_{\overline{E}_n}$;
2. $\varepsilon(n)$ is to be replaced with $\varepsilon(n) + \tau_n$. But as per our construction, $\tau_n < \varepsilon(n)/4$. Hence, it is the same as replacing $\varepsilon(n)$ with $\frac{5}{4}\varepsilon(n)$.

To prove (71), we note that the steps are very similar to the proof of (68). To begin with, pick a capacity achieving distribution P_X and follow all the steps exactly as before. We get

$$\limsup_{n \to \infty} \frac{\log \varepsilon(n)}{n\rho_n^2} \leq -\frac{1}{2V(P_X;W)}.$$

Since this is valid for any $P_X \in \Gamma$, the tightest bound is obtained when we replace $V(P_X;W)$ with $V_{\min}$. $\triangle$

10. DISCUSSION AND COMPARISON WITH RECENT WORK

10.1. Comparison with [14]

In [14], the authors considered a version of the EH-AWGN channel with block energy arrivals. In this model, each block is of length L and energy arrivals are i.i.d. across blocks. Within a block, they are equal. The authors then studied the effects on finite blocklength rates when L was constant and when L was sub-linearly increasing with n (i.e., $L = \omega(1)$). For average probability of error $0 \leq \varepsilon < 1/2$, they showed that for an EH-AWGN channel with block energy arrivals as described above, for constant L, large enough n, and unit noise variance,

$$C_{EG} + V_\varepsilon^- \sqrt{\frac{L}{n}} - o\left(\sqrt{\frac{L}{n}}\right) \leq \frac{1}{n} M^*(n, \varepsilon) \leq C_{EG} + V_\varepsilon^+ \sqrt{\frac{L}{n}} + o\left(\sqrt{\frac{L}{n}}\right), \tag{72}$$

where

$$V_\varepsilon^- = \sup_{0<\lambda<1} -C_{EG}\sqrt{2\left(\frac{\mathbf{E}[E_1^2]}{\mathbf{E}[E_1]^2}+1\right)\log\frac{1}{\lambda\varepsilon}} + \sqrt{\frac{\mathbf{E}[E_1](\log e)^2}{L(1+\mathbf{E}[E_1])}}\Phi^{-1}((1-\lambda)\varepsilon), \qquad (73)$$

$$V_\varepsilon^+ = \frac{\log e}{2(1+\mathbf{E}[E_1])}\sqrt{\sigma_E^2 + \frac{2\,\mathbf{E}[E_1](\mathbf{E}[E_1]+2)}{L}}\ \Phi^{-1}(\varepsilon). \qquad (74)$$

The EH-AWGN model that we study is essentially the same as that in [14] but with $L=1$, and they were analyzing it under the average probability of error criterion. Comparing with (11), we see that they are identical except for the first term, where they have $\log\frac{1}{\lambda\varepsilon}$ as opposed to our $\frac{2}{\lambda\varepsilon}$. Thus, their second-order coefficients are slightly tighter than ours. As for the V_ε^+ term, we recover it exactly, since in this case we provided an alternate proof. In terms of proof techniques, in our case we split the energy harvesting constraints into two parts, namely energy gathering phase and the transmission phase, and analyzed them individually (via Kolmogorov's inequality), whereas in their work it was tackled together using moment generating functions. Additionally, they took advantage of the Gaussianness of the channel to simplify expressions for both achievability and converse.

10.2. Alternate EH-AWGN Converse Using [7, 18]

The method of modifying the meta converse in [7], as we did for energy harvesting channels, has been carried out correspondingly for quasi-static fading AWGN channels in [18, Appendix III]. While we were unable to use the methods directly due to the models being very different, our framework is in a sense inspired by it. We provide a sketch of another derivation of converse bound, as suggested by an anonymous reviewer.

Assuming avg p.o.e. ε, let $\varepsilon(E^n)$ be the average p.o.e. given the energy realization E^n. Note that the avg p.o.e. bound will upper bound the max p.o.e. bound. Letting $\overline{E}_n = \sum_{i=1}^{n} E_i/n$ and noting that the energy harvesting condition to avoid outage is tantamount to $\sum_{i=1}^{n} X_i^2 \leq n\overline{E}_n$, fixing n large, we have, from [7],

$$\varepsilon(E^n) \geq \Phi\left(\frac{\log M - nC(\overline{E}_n) - 1/2\log n - K(\overline{E}_n)}{\sqrt{nV(\overline{E}_n)}}\right). \qquad (75)$$

Taking expected value with respect to the distribution of E^n in (75) should yield a desirable converse. We get

$$\varepsilon \geq \mathbf{E}\,\Phi\left(\frac{\log M - nC(\overline{E}_n) - 1/2\log n - K(\overline{E}_n)}{\sqrt{nV(\overline{E}_n)}}\right). \qquad (76)$$

In order to take the expectation inside the Φ function, we have, for any real-valued random variable U,

$$\mathbf{E}\,\Phi\left(\frac{U+a}{b}\right) = \Pr(U \geq bZ - a) = \int_{-\infty}^{\infty} f_Z(z)\Pr(U \geq bz - a)\,dz, \qquad (77)$$

where $Z \sim N(0,1)$. When U is independent zero mean Gaussian with variance σ_U^2, (77) equals $\Phi\left(\frac{a}{\sqrt{b^2+\sigma_U^2}}\right)$. Using Taylor's approximation on $V(\overline{E}_n)$, (77), and the Berry–Esseen theorem, (76) simplifies to

$$\varepsilon \geq \mathbf{E}\,\Phi\left(\frac{\log M - nC(\overline{E}_n) - 1/2\log n - K(\overline{E}_n)}{\sqrt{nV_{EG2}}}\right) - \frac{c_1(\mathcal{E}[E_1])}{n^{1/3}} \qquad (78)$$

for a constant $c_1(\mathcal{E}[E_1])$ independent of n. By using $C(\overline{E}_n) \le C_{EG} + \dfrac{\overline{E}_n - \mathbf{E}[E_1]}{2(\mathbf{E}[E_1] + \sigma^2)}$ and applying a uniform bound on $K(\overline{E}_n)$, we get the upper bound

$$\log M \le nC(P) + \sqrt{nV_{EG}}\Phi^{-1}\left(\varepsilon + \frac{c_1(\mathbf{E}[E_1])}{n^{1/3}} + \frac{c_2}{\sqrt{n}}\right) + 0.5\log n + O(1)$$

$$\le nC(P) + \sqrt{nV_{EG}}\Phi^{-1}(\varepsilon) + n^{1/6}c_1(\mathbf{E}[E_1]) + 0.5\log n + O(1), \tag{79}$$

which is weaker than (12). Moreover, this proof implicitly uses properties of the Gaussian channel structure which cannot directly be applied to EH-DMCs. However, it is a direct way to get an upper bound (not necessarily the best bound) on rates for special channels (like fading channels) perturbed by AWGN noise, and further details are available in [18].

11. NUMERICAL RESULTS

We now evaluate and plot the finite blocklength bounds on rate as well as the slots consumed in the saving part of save and transmit as a function of blocklength. We use the formulae derived in the earlier sections, for a specified set of parameters, to evaluate the aforementioned quantities. For the EH-DMC, we describe an energy harvesting binary symmetric channel (BSC) and a binary erasure channel (BEC) and plot the corresponding bounds for these. Note that in all plots, we are ignoring the constant terms in the bounds, i.e., coefficients of $O(1/n)$ in the rates. Additionally, we compare our results with the finite blocklength lower bounds of an equivalent non-energy harvesting channel. For example, in the EH-AWGN case, we consider an AWGN channel with average power constraint $\mathbf{E}[E_1]$, while in the EH-DMC cases, we consider corresponding DMCs with power constraint $\mathbf{E}[E_1]$. This will allow us, when the equivalent channel's lower bound is above the energy harvesting upper bound, to comment on the effects of energy harvesting on rate. The gap in rates mentioned henceforth will be the difference between the bounds divided by the upper bound, expressed as a percentage.

11.1. EH-AWGN Results

We take the maximal probability of error $\varepsilon = 0.1$, $\mathbf{E}[E_1] = 1$, and $\sigma_E^2 = 5$. We consider blocklengths n between 5000 to 10 000. We consider three different regimes, i.e.,

1. Low SNR (-20 dB). In this regime (Fig. 2), we observe that the lower bound is a poor approximation to the finite blocklength rate. Due to a larger number of errors, this regime also requires more slots to harvest energy to lower the error due to outage (about 20.5% to 27.6%);
2. Moderate SNR (0 dB). Compared to low SNR, this regime (see Fig. 3) gives a better approximation to finite blocklength. The gap in rates is significantly lowered to approximately 19% to 27%. Additionally, the number of slots required in the saving phase are also considerably reduced (16% to 22%);
3. High SNR (20 dB). In this regime (Fig. 4), the gap between rates is about 18.2% to 24.2% and the slots required in saving energy is between 15.8% to 21.6%. While this is an improvement from moderate SNR, it is not as significant as that between low SNR to moderate SNR.

To summarize, the finite blocklength bounds are decent approximations to the finite blocklength rate in the moderate to high SNR regime. Further improvements would require an improved lower bound, which would require changing the transmission scheme. Except for the low SNR case, we observe that the energy harvesting upper bound is below the lower bound of the equivalent AWGN channel. We can infer from this that the finite blocklength energy harvesting rates are lower than that of the non-energy harvesting case in the moderate and high SNR regime.
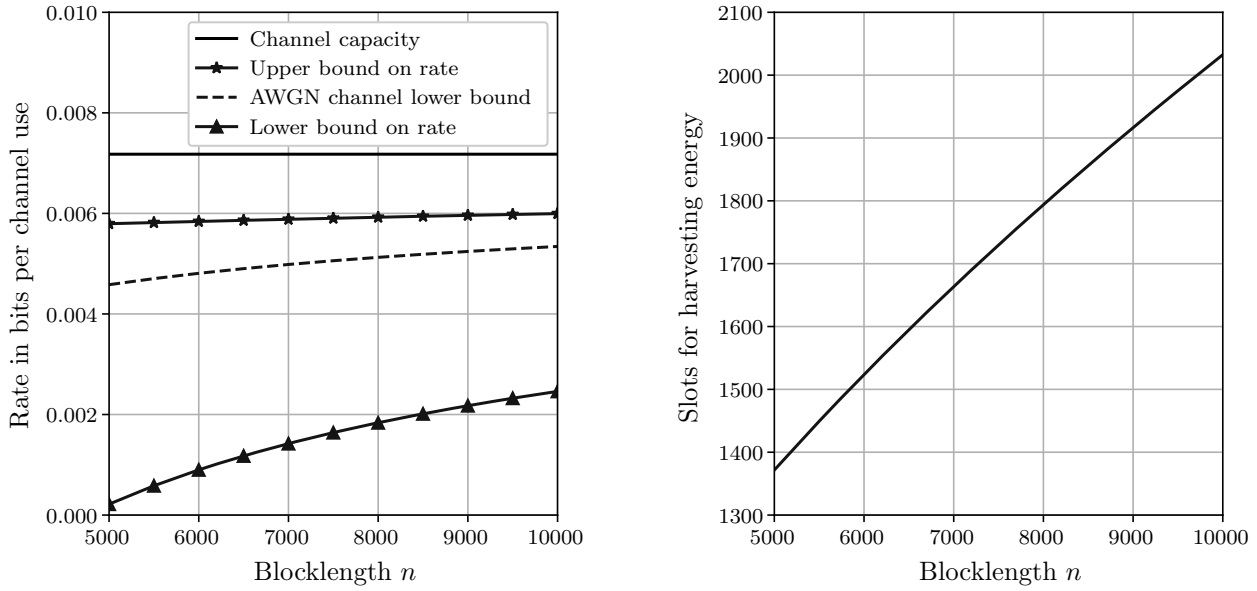
Fig. 2. Plot of FB rates for an EH-AWGN channel versus the total blocklength (harvesting plus transmission) in low SNR regime. The other plot shows the number of slots used for harvesting energy.
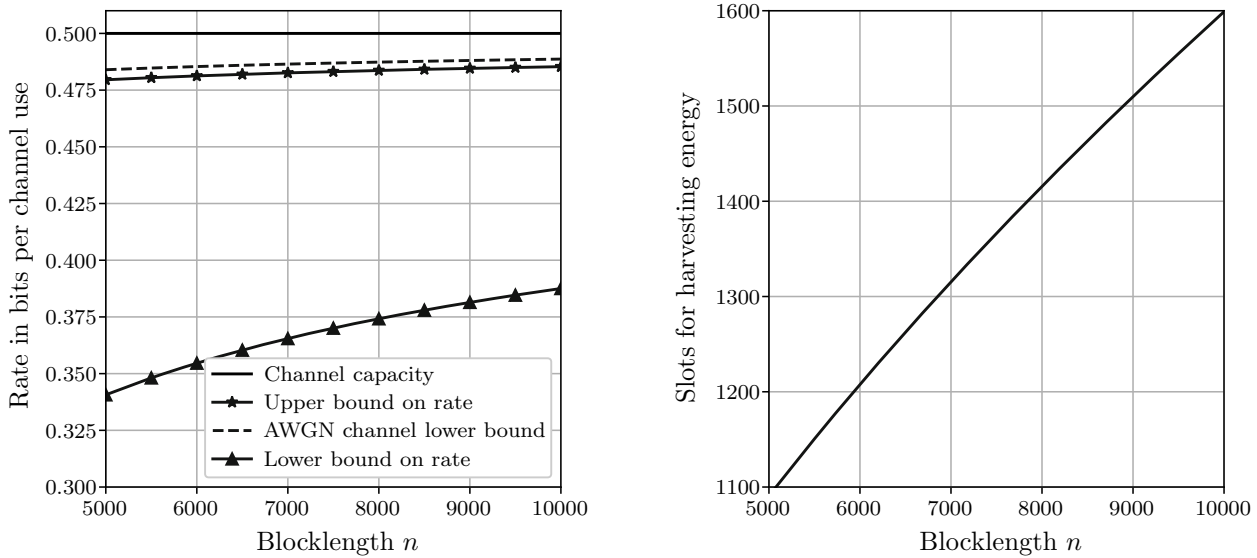

Fig. 3. Plot of FB rates for an EH-AWGN channel versus the total blocklength (harvesting plus transmission) in moderate SNR regime. The other plot shows the number of slots used for harvesting energy.

11.2. EH-BSC

Consider a binary symmetric channel W with crossover probability α. That is $\mathcal{X} = \mathcal{Y} = \{0, 1\}$ and $W(0|1) = W(1|0) = \alpha$. Let $p_0 := \Pr(X = 0)$. If the capacity achieving distribution, which satisfies the energy harvesting requirements, is unique with p_0 as before, then

$$C_{ED} = C_{BSC} = h(\alpha p_0 + \overline{\alpha}\,\overline{p_0}) - h(\alpha),$$

$$V(P; W) = V_{BSC} = \sum_{x \in \{\alpha, \overline{\alpha}\}} \sum_{y \in \{p_0, \overline{p_0}\}} xy \left[\log\left(\frac{x}{xy + \overline{x}\,\overline{y}} \right) \right]^2 - C_{BSC}^2,$$
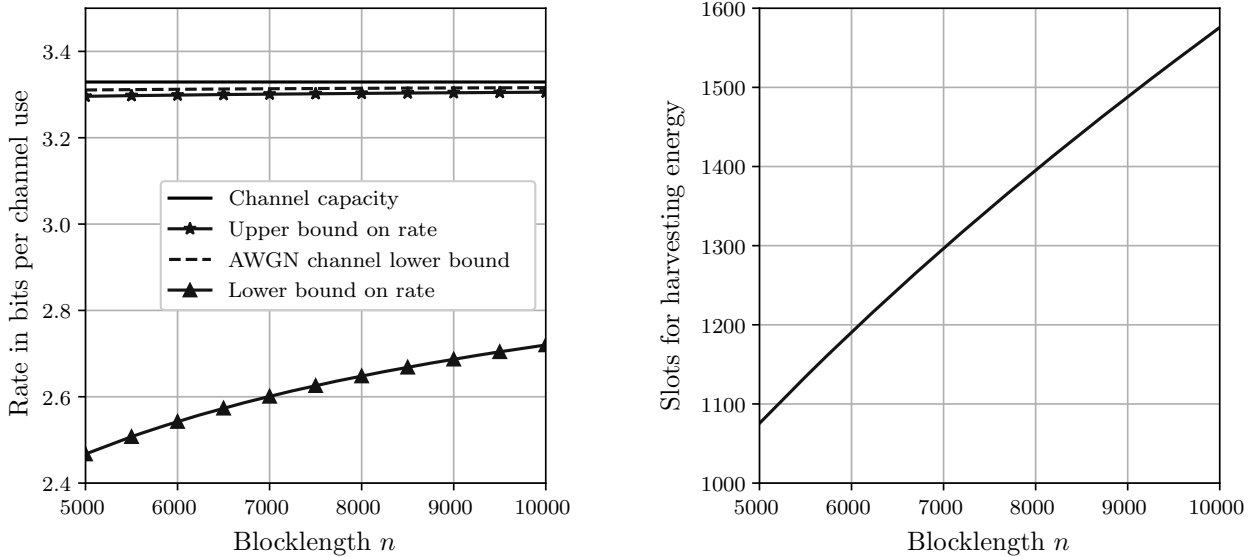
Fig. 4. Plot of FB rates for an EH-AWGN channel versus the total blocklength (harvesting plus transmission) in high SNR regime. The other plot shows the number of slots used for harvesting energy.
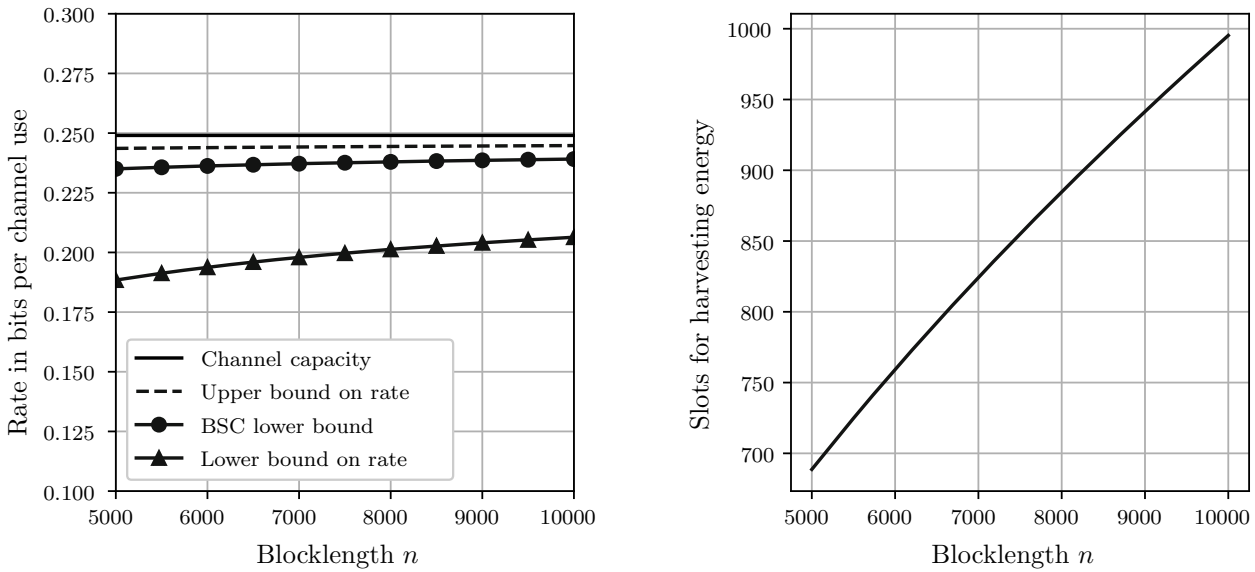

Fig. 5. Plot of FB rates for an EH-BSC channel versus the total blocklength (harvesting plus transmission). The plot on the right gives the number of slots used for harvesting energy.

where $\overline{u} := 1 - u$ and $h(x) = -x \log_2(x) - \overline{x} \log_2(\overline{x})$ is the binary entropy function. Note that the choice of p_0 is influenced by energy harvesting constraints. In this example, we pick $\alpha = 0.05$, the energy function $\Lambda(x) = 3x$, and $\mathbf{E}[E_1] = 1$. This ensures the uniqueness of the capacity achieving distribution with $p_0 = 2/3$. We take $\varepsilon = 0.1$ and $\sigma_E^2 = 0.2$ here. Figure 5 plots the lower and upper bounds for this example where n is between 5000 and 10 000.

We observe that the difference between upper and lower bounds for this example is between 13.7% to 23%. The blocklength required for saving energy varies from 9.8% to 13.8% in this range. In this case, the non-energy harvesting lower bound is below the energy harvesting upper bound. Hence, we cannot infer anything about the rates as a function of σ_E^2 here.
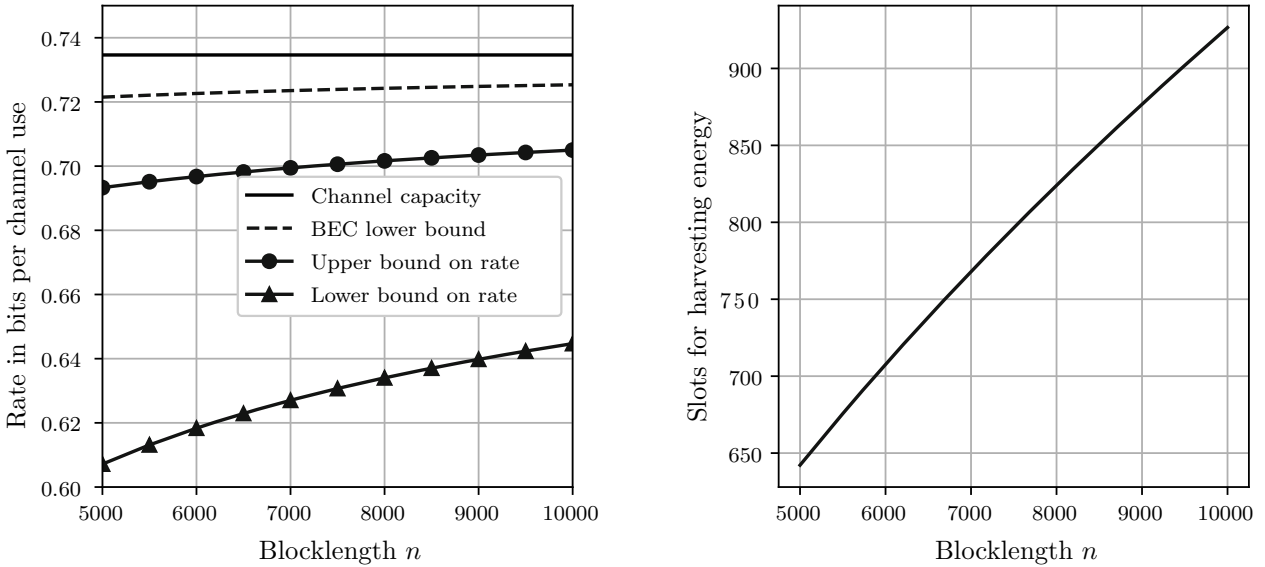
Fig. 6. Plot of FB rates for an EH-BEC channel versus the total blocklength (harvesting plus transmission). The plot on the right gives the number of slots used for harvesting energy.

11.3. EH-BEC

A binary erasure channel W is a channel with binary inputs $\mathcal{X} = \{0, 1\}$ and ternary outputs $\mathcal{Y} = \{0, E_R, 1\}$ with $W(0|0) = W(1|1) = 1 - \alpha$ and $W(E_R|0) = W(E_R|1) = \alpha$, where α is the erasure probability. Similarly to the BSC case, if we have a unique capacity achieving distribution, $p_0 = \Pr(X = 0)$, then

$$C_{ED} = C_{BEC} = (1 - \alpha)h(p_0),$$

$$V(P; W) = V_{BEC} = (1 - \alpha)p_0(\log(p_0))^2 + (1 - \alpha)(1 - p_0)(\log(1 - p_0))^2 - C_{BEC}^2.$$

Using the same parameters as in the BSC case, we plot the bounds in Fig. 6.

We observe a difference of 8.6% to 12.2% between the upper and lower bounds as well as saving energy slot utilization of 9.3% to 12.8% for the specified range of parameters. Here our bounds appear to better approximate the rates as opposed to BSC. Moreover, the non-energy harvesting lower bound is above the upper bound, meaning that in this case the effects of energy harvesting are detrimental to the rate.

11.4. Effects of Energy Harvesting Variance σ_E^2

Comparing the bounds (11) and (12) derived for EH-AWGN channels, we observe that both bounds are lowered with increasing σ_E^2. This is illustrated in Fig. 7. Interestingly, when compared to the AWGN lower bound, the EH-AWGN upper bound appears to only differ by $O(\log n/n)$ when $\sigma_E^2 = 0$. However, the lower bound is strongly affected by the variance.

12. DISCUSSION OF RESULTS AND CONCLUSION

In this paper, we have shown that for both EH-AWGN and EH-DMC channels, the finite blocklength code size varies as $nC - \Theta(\sqrt{n})$ under the maximal probability of error criterion. This was shown by deriving lower and upper bounds with second-order $\sqrt{n}$. We also bounded the moderate deviation asymptotics for both channel types.
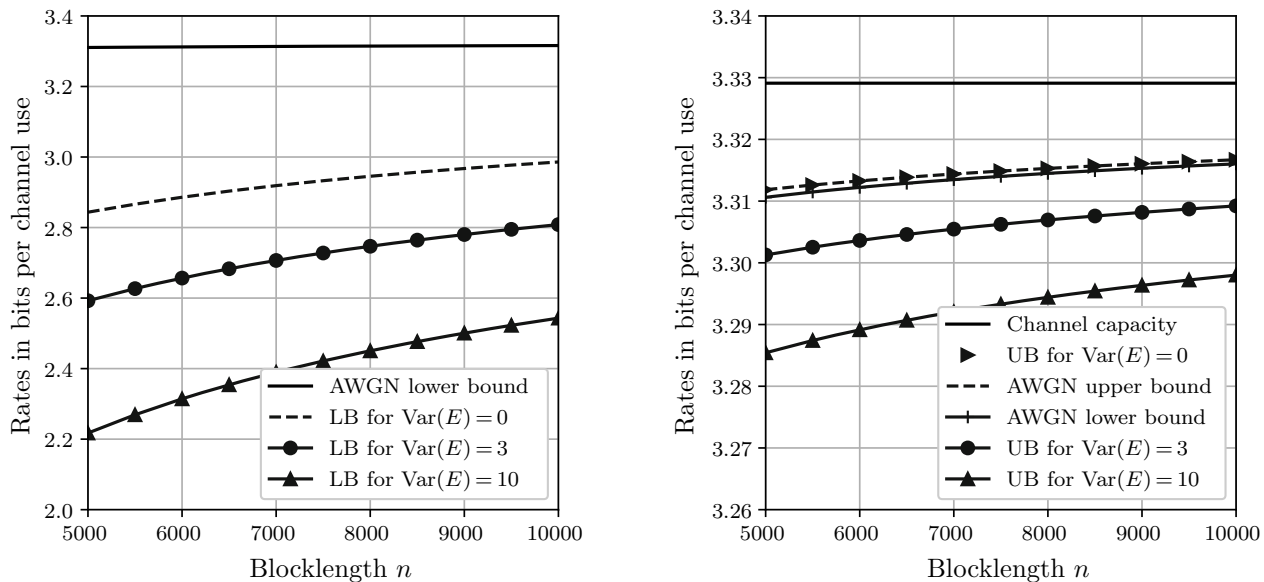
Fig. 7. Plot of finite blocklength rates for an EH-AWGN for different energy harvesting variances.

Additionally, the bounds were plotted for a few examples. In certain cases, such as the AWGN channel with moderate to high SNR as well as the BEC case, we observed that the rates are exacerbated with increased variance of the energy harvesting process. It is desirable to tighten the gap between lower and upper bounds so that this conjecture may be further verified. As future work, obtaining matching bounds in the finite blocklength as well as the moderate deviations regime will be useful.

ACKNOWLEDGMENTS

ADDITIONAL INFORMATION

A part of this paper has been published in ISIT'2016 proceedings [1].

APPENDIX A: *PROOF OF* (41)

Let $U \in [M]$ denote the message to be transmitted and similarly $\widehat{U}$ the decoded message. We have for channel W, if the maximal probability of error is ε, the following steps:

$$
\begin{aligned}
1 - \varepsilon &\leq \Pr[\widehat{U} = m \,|\, U = m] \\
&= \int_{\boldsymbol{y}, \boldsymbol{e}} \Pr[\widehat{U} = m \,|\, \boldsymbol{Y} = \boldsymbol{y}] W(\boldsymbol{y} \,|\, c(m, \boldsymbol{e})) \, dP_{\boldsymbol{E}}(\boldsymbol{e}),
\end{aligned}
\tag{80}
$$

where the above holds for any message m.

Now $\Pr[\widehat{U} = m \,|\, \boldsymbol{Y} = \boldsymbol{y}]$ is a test on the decoder end that achieves the probability of error requirement. Even though it does not depend on $\boldsymbol{e}$, since the decoder does not have access to the energy samples, it is still a valid test on $(\boldsymbol{y}, \boldsymbol{e})$.

Now assume that instead of channel W, the message is sent on channel Q_Y which is an auxiliary channel that ignores the input but has the same output alphabet. Using the above decoder,

let $\overline{m}$ be the message that achieves the maximal probability of error under Q_Y. Then, clearly, $P(\widehat{U} = \overline{m} \,|\, U = \overline{m}) \leq \dfrac{1}{M}$ under Q_Y. But then, from (80) and the definition of the beta error function, we have

$$\beta_{1-\varepsilon}\left(W(\,\cdot\,|\,c(\overline{m},*))P_{\boldsymbol{E}}(*), Q_{\boldsymbol{Y}}P_{\boldsymbol{E}}\right) \leq \int\limits_{\boldsymbol{y}} P(\widehat{U} = \overline{m} \,|\, \boldsymbol{Y} = \boldsymbol{y})\, dQ_{\boldsymbol{Y}}(\boldsymbol{y}) \leq \frac{1}{M}.$$

APPENDIX B: PROOF OF THEOREM 4

The proof follows the steps used in proving the original meta-converse (see [7]) up to a point. Given distribution $Q_{\boldsymbol{Y}}$, which is essentially a reference channel that does not depend on input, let the maximal probability of error for this "channel" be ε'. Let U be the random variable denoting the message to be sent and $\widehat{U}$ be the message that was decoded.

Consider the definition of maximal probability of error. We see that there is a message, call it $\overline{m}$, such that

$$1 - \varepsilon' = \Pr[\widehat{U} = \overline{m} \,|\, U = \overline{m}] = \int\limits_{\boldsymbol{y}} P_{\widehat{U}|\boldsymbol{Y}}(\overline{m}\,|\,\boldsymbol{y})\, dQ_{\boldsymbol{Y}}(\boldsymbol{y}). \tag{81}$$

But we also have

$$1 - \varepsilon' = \min_m \Pr[\widehat{U} = m \,|\, U = m]$$

$$\leq \frac{1}{M} \sum_{m=1}^{M} \Pr[\widehat{U} = m \,|\, U = m]$$

$$= \frac{1}{M} \sum_{m=1}^{M} \int\limits_{\boldsymbol{y}} \Pr[\widehat{U} = m \,|\, \boldsymbol{Y} = \boldsymbol{y}]\, dQ_{\boldsymbol{Y}}(\boldsymbol{y})$$

$$= \frac{1}{M} \int\limits_{\boldsymbol{y}} \left(\sum_{m=1}^{M} \Pr[\widehat{U} = m \,|\, \boldsymbol{Y} = \boldsymbol{y}] \right) dQ_{\boldsymbol{Y}}(\boldsymbol{y})$$

$$= \frac{1}{M}. \tag{82}$$

Combining equation (81) and (82), we get

$$M \leq \frac{1}{\int\limits_{\boldsymbol{y}} P_{\widehat{U}|\boldsymbol{Y}}(\overline{m}\,|\,\boldsymbol{y})\, dQ_{\boldsymbol{Y}}(\boldsymbol{y})}. \tag{83}$$

Now we have, for any $\mathcal{E}_1 \subset \mathbb{R}^n_+$,

$$1 - \varepsilon \leq \int\limits_{\boldsymbol{e}} \int\limits_{\boldsymbol{y}} P_{\widehat{U}|\boldsymbol{Y}}(\overline{m}\,|\,\boldsymbol{y})\, dP_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,c(\overline{m},\boldsymbol{e}))\, dP_{\boldsymbol{E}}(\boldsymbol{e})$$

$$\leq \int\limits_{\boldsymbol{e}\in\mathcal{E}_1} \int\limits_{\boldsymbol{y}} P_{\widehat{U}|\boldsymbol{Y}^n}(\overline{m}\,|\,\boldsymbol{y})\, dP_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,c(\overline{m},\boldsymbol{e}))\, dP_{\boldsymbol{E}}(\boldsymbol{e}) + P_{\boldsymbol{E}}(\mathcal{E}_1^c).$$

Rearranging and using the definitions given in the statement of the lemma, letting $\mathcal{E}_1 = \Big\{ \boldsymbol{e} : \sum\limits_{i=1}^{n} e_i \leq n\overline{E}_n \Big\}$ and $\tau_n = P_E(\mathcal{E}_1^c)$, we get

$$1 - \varepsilon - \tau_n \leq \int\limits_{\boldsymbol{e}\in\mathcal{E}_1} \int\limits_{\boldsymbol{y}} P_{\widehat{U}|\boldsymbol{Y}}(\overline{m}\,|\,\boldsymbol{y})\, dP_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,c(\overline{m},\boldsymbol{e}))\, dP_{\boldsymbol{E}}(\boldsymbol{e}) \tag{84}$$

$$\Rightarrow 1 - \varepsilon - \tau_n \leq \frac{1 - \varepsilon - \tau_n}{1 - \tau_n} \leq \int\limits_{\boldsymbol{y}} P_{\widehat{U}|\boldsymbol{Y}}(\overline{m}\,|\,\boldsymbol{y}) \Bigg\{ \int\limits_{\boldsymbol{e}\in\mathcal{E}_1} dP_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}\,|\,c(\overline{m},\boldsymbol{e})) \frac{dP_{\boldsymbol{E}}(\boldsymbol{e})}{1 - \tau_n} \Bigg\}. \tag{85}$$

Note that we divide by $1 - \tau_n$ to ensure that the term in braces is a probability distribution. From (83), (85), and the definition of β error function, we get

$$\frac{1}{M} \geq \beta_{1-\varepsilon-\tau_n}\left(\int_{\boldsymbol{e} \in \mathcal{E}_1} dP_{\boldsymbol{Y}|\boldsymbol{X}}(\,\cdot\,|\,c(\overline{m}, \boldsymbol{e}))\frac{dP_{\boldsymbol{E}}(\boldsymbol{e})}{1-\tau_n}, Q_{\boldsymbol{Y}}\right)$$

$$\geq \inf_{\boldsymbol{x} \in \mathbb{F}_{\overline{E}_n}} \beta_{1-\varepsilon-\tau_n}\left(\int_{\boldsymbol{e} \in \mathcal{E}_1} dP_{\boldsymbol{Y}|\boldsymbol{X}}(\,\cdot\,|\,\boldsymbol{x})\frac{dP_{\boldsymbol{E}}(\boldsymbol{e})}{1-\tau_n}, Q_{\boldsymbol{Y}}\right)$$

$$= \inf_{\boldsymbol{x} \in \mathbb{F}_{\overline{E}_n}} \beta_{1-\varepsilon-\tau_n}\left(P_{\boldsymbol{Y}|\boldsymbol{X}}(\,\cdot\,|\,\boldsymbol{x}), Q_{\boldsymbol{Y}}\right).$$

Note that we could take the infimum over $\mathbb{F}_{\overline{E}_n}$, a nonrandom set here, because when $e^n \in \mathcal{E}_1$, it implies that $c(\overline{m}, e^n) \in \mathbb{F}$. Hence, we have (43).

REFERENCES

1. Shenoy, K.G. and Sharma, V., Finite Blocklength Achievable Rates for Energy Harvesting AWGN Channels with Infinite Buffer, in *Proc. 2016 IEEE Int. Symp. on Information Theory (ISIT'2016), Barcelona, Spain, July 10–15, 2016*, pp. 465–469. https://doi.org/10.1109/ISIT.2016.7541342

2. Kamalinejad, P., Mahapatra, C., Sheng, Z., Mirabbasi, S., Leung, V.C., and Guan, Y.L., Wireless Energy Harvesting for the Internet of Things, *IEEE Commun. Mag.*, 2015, vol. 53, no. 6, pp. 102–108. https://doi.org/10.1109/MCOM.2015.7120024

3. Ku, M.-L., Li, W., Chen, Y., and Liu, K.R., Advances in Energy Harvesting Communications: Past, Present, and Future Challenges, *IEEE Commun. Surv. Tutor.*, 2016, vol. 18, no. 2, pp. 1384–1412. https://doi.org/10.1109/COMST.2015.2497324

4. Raza, M., Aslam, N., Le-Minh, H., Hussain, S., Cao, Y., and Khan, N.M., A Critical Analysis of Research Potential, Challenges, and Future Directives in Industrial Wireless Sensor Networks, *IEEE Commun. Surv. Tutor.*, 2017, vol. 20, no. 1, pp. 39–95. https://doi.org/10.1109/COMST.2017.2759725

5. Strassen, V., Asymptotische Abschätzungen in Shannons Informationstheorie, *Trans. 3rd Prague Conf. on Information Theory, Statistical Decision Functions, Random Processes held at Liblice near Prague, June 5–13, 1962*, Kozesnik, J., Ed, Prague: Czechoslovak Acad. Sci., 1964, pp. 689–723.

6. Hayashi, M., Information Spectrum Approach to Second-Order Coding Rate in Channel Coding, *IEEE Trans. Inform. Theory*, 2009, vol. 55, no. 11, pp. 4947–4966. https://doi.org/10.1109/TIT.2009.2030478

7. Polyanskiy, Y., Poor, H.V., and Verdú, S., Channel Coding Rate in the Finite Blocklength Regime, *IEEE Trans. Inform. Theory*, 2010, vol. 56, no. 5, pp. 2307–2359. https://doi.org/10.1109/TIT.2010.2043769

8. Polyanskiy, Y., Poor, H.V., and Verdú, S., New Channel Coding Achievability Bounds, in *Proc. 2008 IEEE Int. Symp. on Information Theory (ISIT'2008), Toronto, ON, Canada, July 6–11, 2008*, pp. 1763–1767. https://doi.org/10.1109/ISIT.2008.4595291

9. Tomamichel, M. and Tan, V.Y.F., A Tight Upper Bound for the Third-Order Asymptotics for Most Discrete Memoryless Channels, *IEEE Trans. Inform. Theory*, 2013, vol. 59, no. 11, pp. 7041–7051. https://doi.org/10.1109/TIT.2013.2276077

10. Tomamichel, M. and Tan, V.Y.F., Second-Order Coding Rates for Channels with State, *IEEE Trans. Inform. Theory*, 2014, vol. 60, no. 8, pp. 4427–4448. https://doi.org/10.1109/TIT.2014.2324555

11. Rajesh, R., Sharma, V., and Viswanath, P., Capacity of Gaussian Channels with Energy Harvesting and Processing Cost, *IEEE Trans. Inform. Theory*, 2014, vol. 60, no. 5, pp. 2563–2575. https://doi.org/10.1109/TIT.2014.2311822

12. Ozel, O. and Ulukus, S., Achieving AWGN Capacity under Stochastic Energy Harvesting, *IEEE Trans. Inform. Theory*, 2012, vol. 58, no. 10, pp. 6471–6483. `https://doi.org/10.1109/TIT.2012.2204389`

13. Yang, J., Achievable Rate for Energy Harvesting Channel with Finite Blocklength, in *Proc. 2014 IEEE Int. Symp. on Information Theory (ISIT'2014), Honolulu, HI, USA, June 29 – July 4, 2014*, pp. 811–815. `https://doi.org/10.1109/ISIT.2014.6874945`

14. Fong, S.L., Tan, V.Y.F., and Özgür, A., On Achievable Rates of AWGN Energy-Harvesting Channels with Block Energy Arrival and Non-vanishing Error Probabilities, *IEEE Trans. Inform. Theory*, 2018, vol. 64, no. 3, pp. 2038–2064. `https://doi.org/10.1109/TIT.2017.2765545`

15. Altuğ, Y. and Wagner, A.B., Moderate Deviations in Channel Coding, *IEEE Trans. Inform. Theory*, 2014, vol. 60, no. 8, pp. 4417–4426. `https://doi.org/10.1109/TIT.2014.2323418`

16. Polyanskiy, Y. and Verdú, S., Channel Dispersion and Moderate Deviations Limits for Memoryless Channels, in *Proc. 48th Annu. Allerton Conf. on Communication, Control, and Computation, Monticello, IL, USA, Sept. 29 – Oct. 1, 2010*, pp. 1334–1339. `https://doi.org/10.1109/ALLERTON.2010.5707068`

17. Truong, L.V. and Tan, V.Y.F., Moderate Deviation Asymptotics for Variable-Length Codes with Feedback, `https://arXiv.org/abs/1707.04850v2` [cs.IT], 2017.

18. Yang, W., Durisi, G., Koch, T., and Polyanskiy, Y., Quasi-static Multiple-Antenna Fading Channels at Finite Blocklength, *IEEE Trans. Inform. Theory*, 2014, vol. 60, no. 7, pp. 4232–4265. `https://doi.org/10.1109/TIT.2014.2318726`

19. Cover, T.M. and Thomas, J.A., *Elements of Information Theory*, Hoboken, NJ, USA: Wiley, 2006, 2nd ed.

20. Han, T.S., *Information-Spectrum Methods in Information Theory*, Berlin: Springer, 2003.

21. Csiszár, I. and Körner, J., *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Cambridge: Cambridge Univ. Press, 2011, 2nd ed.

22. Kostina, V. and Verdú, S., Channels with Cost Constraints: Strong Converse and Dispersion, *IEEE Trans. Inform. Theory*, 2015, vol. 61, no. 5, pp. 2415–2429. `https://doi.org/10.1109/TIT.2015.2409261`

23. El Gamal, A. and Kim, Y.-H., *Network Information Theory*, Cambridge: Cambridge Univ. Press, 2011.

24. Athreya, K.B. and Lahiri, S.N., *Probability Theory*, New Delhi, India: Hindustan Book Agency, 2006.

25. Tyurin, I.S., An Improvement of Upper Estimates of the Constants in the Lyapunov Theorem, *Uspekhi Mat. Nauk*, 2010, vol. 65, no. 3 (393), pp. 201–202 [*Russian Math. Surveys* (Engl. Transl.), 2010, vol. 65, no. 3, pp. 586–588]. `http://dx.doi.org/10.1070/RM2010v065n03ABEH004688`

26. Sun, Y., Baricz, Á., and Zhou, S., On the Monotonicity, log-Concavity, and Tight Bounds of the Generalized Marcum and Nuttall Q-Functions, *IEEE Trans. Inform. Theory*, 2010, vol. 56, no. 3, pp. 1166–1186. `https://doi.org/10.1109/TIT.2009.2039048`

27. Wolff, R.W., *Stochastic Modeling and the Theory of Queues*, Englewood Cliffs, NJ: Prentice Hall, 1989.

28. Gut, A., *Stopped Random Walks*, New York: Springer, 2009, 2nd ed.

29. Dembo, A. and Zeitouni, O., *Large Deviations Techniques and Applications*, Berlin: Springer, 2009. Corrected printing of the 1998 ed.

F