On Ultra-Reliable and Low Latency Simultaneous Information and Energy Transmission Systems

Nizar Khalfet, Samir Perlaza, Ali Tajer, Harold Vincent Poor

HAL Id: hal-02136606
https://hal.archives-ouvertes.fr/hal-02136606

Submitted on 22 May 2019

On Ultra-Reliable and Low Latency Simultaneous Information and Energy Transmission Systems

Nizar Khalfet, Samir M. Perlaza, Ali Tajer, and H. Vincent Poor

Abstract—In this paper, the fundamental limits of simultaneous information and energy transmission (SIET) are studied in the non-asymptotic block-length regime. The focus is on the case of a transmitter simultaneously sending information to a receiver and energy to a harvester through binary symmetric channels. Given a finite number of channel uses (latency constraint) as well as tolerable average decoding error probability and energy shortage probability (reliability constraints), two sets of information and energy transmission rates are presented. One consists in rate pairs for which the existence of at least one code achieving such rates under the latency and reliability constraints is proved (achievable region). The second one consists in a set whose complement contains the rate pairs for which there does not exist a code capable of achieving such rates and satisfying both latency and reliability constraints (converse region). These two sets approximate the information-energy capacity region, which allows analyzing the trade-offs among performance, latency, and reliability in SIET systems.

Index Terms—Information and Energy Transmission, Information-Energy Capacity Region, Finite Block-Length Regime.

I. Introduction

Simultaneous information and energy transmission (SIET) refers to systems in which at least one transmitter aims to simultaneously send information to a set of information receivers (IRs) and energy to a set of energy harvesters (EHs). This idea traces back to Nikola Tesla, who proposed SIET in 1914 [1]. In modern communications systems, SIET is one of the central ideas for wirelessly powering up devices with low-energy consumption [2].

The fundamental limits of SIET are characterized by the *information-energy capacity region* [3]. This region consists in the set of all information and energy transmission rates that can be simultaneously achieved. In general, it can be characterized in two different regimes: (i) the asymptotic block-length regime; and (ii) the non-asymptotic block-length regime. The former refers to a case in which the block length is assumed to be infinitely long, while the decoding error probability (DEP) and the energy-shortage probability (ESP)

are assumed to be arbitrarily close to zero. From this perspective, the asymptotic block-length regime does not capture the constraints on latency. Essentially, these limits apply only to the scenarios in which the duration of the transmission is arbitrarily long. The non-asymptotic regime, on the other hand, refers to the case in which the block length is assumed to be finite and both the DEP and the ESP are bounded away from zero. In this case, the information-energy capacity region is parametrized by a finite block length, an upper bound on the DEP, and an upper bound on the ESP. This allows taking into account the constraints on latency in terms of channel uses, and reliability in terms of DEP and ESP.

The information energy capacity region in the asymptotic regime was characterized for point-to-point memoryless channels in [3], [4], and [5]. Alternatively, in multi-user channels, characterizations of the information-energy capacity region of multiple access channels were presented in [6] and [7]. A characterization of this region in the context of the interference channel was presented in [8]. In the non-asymptotic regime, however, the information-energy capacity region in point-to-point channels is not well-investigated. A first attempt to characterize it was made in [9], building upon the existing results on the fundamental limits on information transmission in the non-asymptotic block-length regime in [10] and [11]. In multi-user channels, a characterization of the information-energy capacity region is unknown.

The focus of this paper is on a system in which a transmitter simultaneously sends information to an information receiver and energy to an energy harvester through binary symmetric channels. The main contribution is characterizing the information-energy capacity region. This characterization is achieved by providing a set that is confined by the information-energy capacity region and another set that contains it. The inner set contains the information and energy transmission rates for which there always exists at least one code achieving such rates (achievable region). The outer set is a set whose complement contains the information and energy transmission rates that cannot be achieved by any code (converse region).

This paper is organized as follows. Section II provides the notation. Section III introduces the system model. Section IV presents an approximation to the information-energy capacity region in terms of the number of channel uses, the DEP, and the ESP. Section V presents the main ideas in the proofs of the main results. Complete proofs are presented in [12].

II. Notation

Throughout this paper, sets are denoted with uppercase calligraphic letters, e.g., $\mathcal{X}$. Random variables are denoted by

uppercase letters, e.g., X, and their realizations are denoted by lower case letters, e.g., x. The probability distribution of X is denoted by P_X. Whenever a second random variable Y is involved, P_{XY} and $P_{Y|X}$ denote, respectively, the joint probability distribution of (X, Y) and the conditional probability distribution of Y given X. Let n be a fixed natural number. An n-dimensional vector of random variables is denoted by bold upper case letters, e.g., $\boldsymbol{X} \triangleq (X_1, X_2, \ldots, X_n)^\mathsf{T}$, and its corresponding realization by bold lower case letters, e.g., $\boldsymbol{x} \triangleq (x_1, x_2, \ldots, x_n)^\mathsf{T}$. Let $\boldsymbol{x}$ be a binary vector. Then, the number of zeros and ones in $\boldsymbol{x}$ are denoted by $N(0|\boldsymbol{x})$ and $N(1|\boldsymbol{x})$, respectively. The notation $\mathbb{E}_X[\cdot]$ is for the expected value of the random variable X. The complementary cumulative distribution function $Q : \mathbb{R} \to [0, 1]$ of the standard Gaussian distribution is

$$Q(t) = \frac{1}{\sqrt{2\pi}} \int_t^\infty \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x, \qquad (1)$$

and the functional inverse of Q is $Q^{-1} : [0, 1] \to \mathbb{R}$. Given two integers n and t, the coefficient of the term x^t in the expansion of the binomial power $(1 + x)^n$ is denoted by $\binom{n}{t}$. Therefore, for all $t < 0$ or $t > n$, it is assumed that $\binom{n}{t} = 0$.

III. SYSTEM MODEL

Consider a three-party communication system in which a transmitter aims at simultaneously sending information to an IR and energy to an EH through a binary symmetric channel. Such a system can be modeled by a random transformation

$$(\{0, 1\}^n, \{0, 1\}^n \times \{0, 1\}^n, P_{\boldsymbol{YZ}|\boldsymbol{X}}), \qquad (2)$$

where $n \in \mathbb{N}$ is the block length. Given an input $\boldsymbol{x} \triangleq (x_1, x_2, \ldots, x_n) \in \{0, 1\}^n$, the outputs $\boldsymbol{y} \triangleq (y_1, y_2, \ldots, y_n) \in \{0, 1\}^n$ and $\boldsymbol{z} \triangleq (z_1, z_2, \ldots, z_n) \in \{0, 1\}^n$ are observed at the IR and at the EH, respectively, with probability

$$P_{\boldsymbol{YZ}|\boldsymbol{X}}(\boldsymbol{y}, \boldsymbol{z}|\boldsymbol{x}) = \prod_{t=1}^n P_{Y|X}(y_t|x_t) P_{Z|X}(z_t|x_t), \qquad (3)$$

where for all $(x, y, z) \in \{0, 1\}^3$,

$$P_{Y|X}(y|x) = \alpha_1 \mathbb{1}_{\{x \neq y\}} + (1 - \alpha_1) \mathbb{1}_{\{x = y\}}, \qquad (4)$$
$$P_{Z|X}(z|x) = \alpha_2 \mathbb{1}_{\{x \neq z\}} + (1 - \alpha_2) \mathbb{1}_{\{x = z\}}, \qquad (5)$$

and $\alpha_1 \in [0, \frac{1}{2})$ and $\alpha_2 \in (0, \frac{1}{2})$. In this context, two tasks are carried out by the transmitter: (a) the information transmission task; and (b) the energy transmission task.

A. Information Transmission Task

The purpose of this task is to send a message from the transmitter to the IR. The message index is a realization of a random variable uniformly distributed in $\{1, 2, \ldots, M\}$, with $M \in \mathbb{N}$. To carry out this task within n channel uses, the transmitter uses an (n, M)-code.

Definition 1 ((n, M)-code): An (n, M)-code for the random transformation in (2) *is a system*

$$\{(\boldsymbol{u}(1), \mathcal{D}_1), (\boldsymbol{u}(2), \mathcal{D}_2), \ldots, (\boldsymbol{u}(M), \mathcal{D}_M)\}, \qquad (6)$$

where for all $(i, j) \in \{1, 2, \ldots, M\}^2$, with $i \neq j$,

$$\boldsymbol{u}(i) \triangleq (u_1(i), u_2(i), \ldots, u_n(i)) \in \{0, 1\}^n, \qquad (7a)$$
$$\mathcal{D}_i \cap \mathcal{D}_j = \emptyset, \; and \qquad (7b)$$
$$\bigcup_{i=1}^M \mathcal{D}_i \subseteq \{0, 1\}^n. \qquad (7c)$$

Given the system in (6), for all $i \in \{1, 2, \ldots, M\}$, to transmit the message with index i, the transmitter inputs the symbol $u_t(i)$ to the channel at time $t \in \{1, 2, \ldots, n\}$. The IR observes the output y_t at the end of channel use t. At the end of n channel uses, the IR decides that the symbol i was transmitted if it satisfies the rule

$$(y_1, y_2, \ldots, y_n) \in \mathcal{D}_i. \qquad (8)$$

The decoding error probability associated with the transmission of message index i, denoted by $\lambda_i \in [0, 1]$, is

$$\lambda_i \triangleq \Pr\left[\boldsymbol{Y} \in \mathcal{D}_i^\mathsf{c} \,|\, \boldsymbol{X} = \boldsymbol{u}(i)\right], \qquad (9)$$

where the probability is with respect to the marginal $P_{\boldsymbol{Y}|\boldsymbol{X}}$, and $\mathcal{D}_i^\mathsf{c}$ represents the complement of $\mathcal{D}_i$ with respect to $\{0, 1\}^n$. The average probability of error, denoted by λ, is

$$\lambda \triangleq \frac{1}{M} \sum_{m=1}^M \lambda_m. \qquad (10)$$

Information transmission is said to be reliable if the average or maximum DEP is controlled. This leads to the following refinements of Definition 1.

Definition 2 ((n, M, ϵ)-code with maximum DEP): Let $\epsilon \in [0, 1]$ be fixed. An (n, M)-code that satisfies $\lambda_i < \epsilon$, for all $i \in \{1, 2, \ldots, M\}$, is said to be an (n, M, ϵ)-code with maximum DEP.

Definition 3 ((n, M, ϵ)-code with average DEP): Let $\epsilon \in [0, 1]$ be fixed. An (n, M)-code that satisfies $\lambda < \epsilon$ is said to be an (n, M, ϵ)-code with average DEP.

Note that any (n, M, ϵ)-code with maximum DEP is also a (n, M, ϵ)-code with average DEP. Nonetheless, the converse is not necessarily true.

B. Energy Transmission Task

Let $g : \{0, 1\} \to \mathbb{R}_+$ be a positive real-valued function that determines the energy harvested from the channel output symbols. Let

$$b_0 \triangleq g(0), \quad and \qquad (11a)$$
$$b_1 \triangleq g(1) \qquad (11b)$$

be the energy harvested when the channel outputs at the EH are 0 and 1, respectively. At the end of n channel uses, the average energy delivered to the EH by the channel outputs $\boldsymbol{z} = (z_1, z_2, \ldots, z_n)$ is given by the function $B_n : \{0, 1\}^n \to \mathbb{R}_+$, with

$$B_n(\boldsymbol{z}) \triangleq \frac{1}{n} \sum_{t=1}^n g(z_t) = (b_0 - b_1) \frac{N(0|\boldsymbol{z})}{n} + b_1. \qquad (12)$$

The objective of the transmitter is to ensure that energy is harvested at the EH at a rate not smaller than b energy units per channel use, with $b \geqslant 0$. An energy-shortage event occurs when the energy harvested at the EH is less than b at the end of the transmission. The case in which $b_0 = b_1$ is trivial, since for all channel outputs $\boldsymbol{z} \in \{0,1\}^n$, it holds that $B_n(\boldsymbol{z}) = b_0 = b_1$. That is, the average energy rate at the input of the EH is independent of the codebook, and either an energy shortage is never observed if $b \geqslant b_0 = b_1$; or the the system is always under energy shortage if $b < b_0 = b_1$. Hence, to avoid these trivial cases, the following assumption is adopted without loss of generality:

$$b_1 < b_0. \tag{13}$$

The probability of energy-shortage when transmitting the message with index $i \in \{1, 2, \ldots, M\}$ is

$$\theta_i \triangleq \Pr\left[B_n(\boldsymbol{Z}) < b \,\middle|\, \boldsymbol{X} = \boldsymbol{u}(i)\right], \tag{14}$$

where the probability is with respect to the marginal $P_{\boldsymbol{Z}|\boldsymbol{X}}$. The average probability of energy-shortage, denoted by θ, is

$$\theta \triangleq \frac{1}{M} \sum_{i=1}^{M} \theta_i. \tag{15}$$

Note that for all $\boldsymbol{z} \in \mathcal{Z}^n$, $B_n(\boldsymbol{z})$ is bounded according to

$$b_1 \leqslant B_n(\boldsymbol{z}) \leqslant b_0. \tag{16}$$

The inequalities in (16) imply that there exists a case in which energy transmission might occur with zero (maximal or average) ESP for all energy transmission rates $b \leqslant b_1$. This is because the event $B_n(\boldsymbol{Z}) < b_1$ is observed with zero probability. Alternatively, any energy transmission rate $b > b_0$ cannot be achieved with an average or maximal energy-shortage probability strictly smaller than one.

Energy transmission is said to be reliable if the average or maximum ESP is controlled. This leads to the following refinements of Definition 1.

Definition 4 ($(n, M, \epsilon, \delta, b)$-code with maximum ESP): Let $\delta \in [0,1]$ and $b \geq 0$ be fixed. An (n, M, ϵ)-code that satisfies $\theta_i < \delta$, for all $i \in \{1, 2, \ldots, M\}$, is said to be an $(n, M, \epsilon, \delta, b)$-code with maximum ESP.

Definition 5 ($(n, M, \epsilon, \delta, b)$-code with average ESP): Let $\delta \in [0,1]$ and $b \geq 0$ be fixed. An (n, M, ϵ)-code that satisfies $\theta < \delta$ is said to be an $(n, M, \epsilon, \delta, b)$-code with average ESP.
Note that any $(n, M, \epsilon, \delta, b)$-code with maximum ESP is also a $(n, M, \epsilon, \delta, b)$-code with average ESP. Nonetheless, the converse is not necessarily true.

C. Fundamental Limits

The non-asymptotic fundamental limits of the system described in Section III are described by the notion of information-energy capacity region. That is, the set of all information and energy transmission rates that are achievable within a given block length subject to constraints on the average or maximum DEP and average or maximum ESP. Note that an average or maximum DEP constraint leads to different definitions of the information-energy capacity region, and so does an average or maximum ESP constraint.

Definition 6 (Information-Energy Capacity Region): The information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$ with average or maximum DEP and average or maximum ESP of the random transformation in (2) is the set of all pairs $\left(\frac{\log_2(M)}{n}, b\right)$ for which there exists an $(n, M, \epsilon, \delta, b)$-code with average or maximum DEP and average or maximum ESP, respectively.

IV. MAIN RESULTS

Given a fixed block length n and a pair $(\epsilon, \delta) \in [0,1]^2$, the information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$ (Definition 6) of the random transformation in (2) subject to (13) is approximated by a set $\underline{\mathcal{C}}(n, \epsilon, \delta)$ that is contained in $\mathcal{C}(n, \epsilon, \delta)$ (Theorem 1) and another set $\overline{\mathcal{C}}(n, \epsilon, \delta)$ that contains $\mathcal{C}(n, \epsilon, \delta)$ (Theorem 2). That is,

$$\underline{\mathcal{C}}(n, \epsilon, \delta) \subseteq \mathcal{C}(n, \epsilon, \delta) \subseteq \overline{\mathcal{C}}(n, \epsilon, \delta). \tag{17}$$

This approximation is obtained by considering an average DEP constraint and a maximum ESP constraint. To describe the set $\underline{\mathcal{C}}(n, \epsilon, \delta)$ the following notation is used.

Let the functions $\phi : \mathbb{N} \times [0,1] \to [0,1]$, and $\chi : \mathbb{R}_+ \times [0,1] \to [0,1]$ be defined as,

$$
\begin{aligned}
\phi(m,\rho) \triangleq \min\Bigg\{ & 1, (m-1) \sum_{\ell_0=0}^{n} \sum_{\ell_1=0}^{n} \sum_{\ell_2=0}^{\min\{\ell_0,\ell_1\}} \sum_{\ell_3=0}^{\ell_0+\ell_1-2\ell_2} \sum_{\ell_4=0}^{\ell_3} \\
& \cdot \binom{n}{\ell_0}\binom{\ell_0}{\ell_2}\binom{n-\ell_0}{\ell_1-\ell_2}\binom{\ell_0}{\ell_4}\binom{n-\ell_0}{\ell_3-\ell_4} \\
& \cdot \alpha_1^{\ell_0+\ell_1-2\ell_2}(1-\alpha_1)^{n-\ell_0-\ell_1+2\ell_2} \\
& \rho^{\ell_3+\ell_0-2\ell_4+\ell_1}(1-\rho)^{2n-(\ell_3+\ell_0-2\ell_4+\ell_1)} \Bigg\},
\end{aligned} \tag{18}
$$

and

$$
\begin{aligned}
\chi(s,\rho) \triangleq & \sum_{t=0}^{n} \binom{n}{t} \rho^t (1-\rho)^{n-t} \\
& \cdot Q\left(\frac{(1-2\alpha_2)t + n\left(\alpha_2 - \frac{s-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \right) \\
& + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}.
\end{aligned} \tag{19}
$$

Let also the functions $M_1^\star : [0,1] \to \mathbb{N}$, and $B^\star : [0,1] \to \mathbb{R}_+$ be defined as

$$
M_1^\star(\rho) \triangleq \left\{
\begin{array}{l}
\operatorname*{argmax}_{m \in \mathbb{N}} \phi(m,\rho) \\
\text{s.t.} \quad \phi(m,\rho) < \epsilon
\end{array}
\right. , \tag{20}
$$

and

$$
B^\star(\rho) \triangleq \left\{
\begin{array}{l}
\operatorname*{argmax}_{s \in \mathbb{R}_+} \chi(s,\rho) \\
\text{s.t.} \quad \chi(s,\rho) < \delta
\end{array}
\right. . \tag{21}
$$

Note that the functions ϕ in (18), χ in (19), $M_1^\star$ in (20), and $B^\star$ in (21) depend on the block length n; the parameters of the random transformation in (2), i.e., α_1 and α_2; and the energy harvested from the symbols 0 and 1, i.e., b_0 and b_1 in (11).

Nonetheless, none of these parameters is put as an arguement of these functions given that they remain constant during this analysis. Using this notation, given a fixed block length n and a pair $(\epsilon, \delta) \in [0,1]^2$, the following theorem introduces the set $\underline{\mathcal{C}}(n, \epsilon, \delta)$, that is contained in the information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$.

Theorem 1: The information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$ of the random transformation in (2) subject to (13), contains the set

$$\underline{\mathcal{C}}(n, \epsilon, \delta) \triangleq \Big\{ (M, b) \in \mathbb{N} \times \mathbb{R}_+ : \exists \rho \in [0,1], \qquad (22)$$

$$M < M_1^\star(\rho) \text{ and } b < B^*(\rho) \Big\},$$

where $M_1^\star : [0,1] \to \mathbb{N}$ is defined in (20) and $B^* : [0,1] \to \mathbb{R}_+$ is defined in (21).

The description of the set $\overline{\mathcal{C}}(n, \epsilon, \delta)$ uses the following notation. Consider the function $\gamma : [0,1]^2 \to [0,1]$, such that

$$\gamma(\rho, q)$$

$$\triangleq \sum_{\ell_0=0}^{n} \sum_{\ell_1=0}^{n} \sum_{\ell_2=0}^{\min\{\ell_0,\ell_1\}} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1 - \ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1}$$

$$\cdot q^{\ell_0} (1-q)^{n-\ell_0} \Big(\mathbb{1}_{\left\{ (\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} > L \right\}}$$

$$+ \lambda \mathbb{1}_{\left\{ (\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} = L \right\}} \Big), \qquad (23)$$

where $\lambda \in [0,1]$ and $L \in \mathbb{R}$ are chosen to satisfy

$$1-\epsilon = \sum_{\ell_0=0}^{n} \sum_{\ell_1=0}^{n} \sum_{\ell_2=0}^{\min\{\ell_0,\ell_1\}} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1 - \ell_2}$$

$$\cdot (1-\alpha_1)^{n-(\ell_0+\ell_1-2\ell_2)} \alpha_1^{\ell_0+\ell_1-2\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1}$$

$$\cdot \Big(\mathbb{1}_{\left\{ (\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} > L \right\}}$$

$$+ \lambda \mathbb{1}_{\left\{ (\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} = L \right\}} \Big). \qquad (24)$$

Consider also the function $\Gamma : [b_1, b_0] \to \mathbb{R}_+$ defined as

$$\Gamma(b) \triangleq \sup_{\rho > \rho^*(b)} \inf_{q \in [0, \frac{1}{2}]} \frac{1}{\gamma(\rho, q)}. \qquad (25)$$

Finally, let also $B^+ \in \mathbb{R}_+$ defined as,

$$B^+ \triangleq (1-\alpha_2) b_0 + \alpha_2 b_1 \qquad (26)$$

$$- \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1-\alpha_2)}{n}} Q^{-1} \left(\delta + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}} \right).$$

Using this notation, given a fixed block length n and a pair $(\epsilon, \delta) \in [0,1]^2$, the following theorem introduces a set, denoted by $\bar{\mathcal{C}}(n, \epsilon, \delta)$, that contains the information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$.

Theorem 2: The information-energy capacity region $\mathcal{C}(n, \epsilon, \delta)$ of the random transformation in (2) subject to (13), is contained in the set

$$\bar{\mathcal{C}}(n, \epsilon, \delta) \triangleq \Big\{ (M, b) \in \mathbb{N} \times \mathbb{R}_+ : M < \Gamma(b) \text{ and } b < B^+ \Big\}, \qquad (27)$$

where $\Gamma : \mathbb{R}_+ \to [0,1]$ is defined in (25) and B^+ is defined in (26).

V. PROOFS

A. Proof of Theorem 1

The proof of Theorem 1 is based on random coding arguments.

Codebook Generation: Let $\rho \in [0,1]$ be a fixed parameter. Consider a probability distribution P_X that satisfies

$$P_X(0) = 1 - P_X(1) = \rho. \qquad (28)$$

Let also $M \in \mathbb{N}$ and $b \in \mathbb{R}_+$ be fixed parameters. An (n, M)-code is randomly generated as follows: first, the codewords $\boldsymbol{u}(1), \boldsymbol{u}(2), \ldots, \boldsymbol{u}(M)$ are realizations of a random variable $\boldsymbol{X}$ following a distribution $P_{\boldsymbol{X}}$ such that for all $\boldsymbol{x} \in \mathcal{X}^n$,

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = \rho^{N(0|\boldsymbol{x})} (1-\rho)^{n-N(0|\boldsymbol{x})}. \qquad (29)$$

Second, the decoding sets $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_M$ are defined using the information density function $\imath : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$, such that

$$\imath(\boldsymbol{x}, \boldsymbol{y}) = \log \left(\frac{P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})}{P_{\boldsymbol{Y}}(\boldsymbol{y})} \right). \qquad (30)$$

Using this notation, for all $\ell \in \{1, 2, \ldots, M-1\}$:

$$\mathcal{D}_\ell \triangleq \Big\{ \boldsymbol{y} \in \mathcal{Y}^n : \ell \in \arg \max_{k \in \{1,2,\ldots,M\}} \imath(\boldsymbol{u}(k), \boldsymbol{y}) \Big\} \setminus \bigcup_{j=1}^{\ell-1} \mathcal{D}_j,$$

and $\qquad \mathcal{D}_M \triangleq \mathcal{Y}^n \setminus \bigcup_{j=1}^{M-1} \mathcal{D}_j. \qquad (31a)$

Decoding Error Probability Analysis: Let $\bar{\lambda}$ be the average over all possible codebooks of the DEP λ in (10). An immediate consequence of the random coding union bound [10] is that the average DEP $\bar{\lambda}$ is upper bounded by:

$$\bar{\lambda} < \mathbb{E}_{\boldsymbol{X}\boldsymbol{Y}} \left[\min \Big\{ 1, (M-1) \Pr \left[\imath(\bar{\boldsymbol{X}}; \boldsymbol{Y}) \geqslant \imath(\boldsymbol{X}; \boldsymbol{Y}) \right] \Big\} \right]$$

$$= \phi(m, \rho), \qquad (32)$$

where the probability in (32) is with respect to the random variable $\bar{\boldsymbol{X}}$, which is independent of $\boldsymbol{X}$ and $\boldsymbol{Y}$, and whose probability mass function is $P_{\boldsymbol{X}}$ in (29). The function $\phi : \mathbb{N} \times [0,1] \to [0,1]$ is defined in (18). The inequality in (32) leads to $\bar{\lambda} < \phi(m, \rho)$, which completes the proof of the bound on the information rate. The proof continues with the proof of the bound on the energy rate.

Energy-Shortage Probability Analysis: Consider an (n, M)-code described by the system in (6) generated using the probability mass function in (28). Hence, at an energy transmission rate b, it follows from the Berry-Esseen Theorem [13] that for all $i \in \{1, 2, \ldots, M\}$,

$$\theta_i \leq Q \left(\frac{(1-2\alpha_2) N(0|\boldsymbol{u}(i)) + n \left(\alpha_2 - \frac{b-b_1}{b_0-b_1} \right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \right)$$

$$+ \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}. \qquad (33)$$

Let $\bar{\theta}_i$ be the average over all possible codebooks of the energy-shortage probability θ_i in (14) while transmitting at an energy rate b. Hence, the following holds:

$$
\begin{aligned}
\bar{\theta}_i =& \sum_{\boldsymbol{x} \in \mathcal{X}^n} P_{\boldsymbol{X}}(\boldsymbol{x}) \Pr\left[B_n(\boldsymbol{Z}) < b \,|\, \boldsymbol{X} = \boldsymbol{x}\right] \\
\leqslant& \sum_{t=0}^{n} \binom{n}{t} \rho^t (1-\rho)^{n-t} Q\left(\frac{(1-2\alpha_2)t + n\left(\alpha_2 - \frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \right) \\
&+ \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}} = \chi(b,\rho).
\end{aligned}
\tag{34}
$$

Hence from (34), it follows that for all $i \in \{1, 2, \ldots, M\}$, $\bar{\theta} < \chi(b,\rho)$. This completes the proof. ∎

B. Proof of Theorem 2

Consider an $(n, M, \epsilon, \delta, b)$-code with maximum ESP described by the system in (6) and empirical input distribution

$$
\bar{P}_X(0) = 1 - \bar{P}_X(1) \triangleq \rho.
\tag{35}
$$

The proof is based on the notion of the *meta converse* introduced in [10]. Consider the following hypotheses:

$$
\mathcal{H}_0 : (\boldsymbol{X}, \boldsymbol{Y}) \sim \bar{P}_{\boldsymbol{X}} Q_{\boldsymbol{Y}} \text{ and} \tag{36a}
$$
$$
\mathcal{H}_1 : (\boldsymbol{X}, \boldsymbol{Y}) \sim \bar{P}_{\boldsymbol{X}} P_{\boldsymbol{Y}|\boldsymbol{X}}, \tag{36b}
$$

where for all $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$

$$
Q_{\boldsymbol{Y}}(\boldsymbol{y}) = q^{N(0,\boldsymbol{y})} (1-q)^{n-N(0,\boldsymbol{y})}, \tag{37}
$$
$$
\bar{P}_{\boldsymbol{X}}(\boldsymbol{x}) = \rho^{N(0|\boldsymbol{x})} (1-\rho)^{n-N(0|\boldsymbol{x})}, \text{ and} \tag{38}
$$
$$
P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) = \alpha_1^{d(\boldsymbol{x},\boldsymbol{y})} (1-\alpha_1)^{n-d(\boldsymbol{x},\boldsymbol{y})}. \tag{39}
$$

The goal of the binary hypothesis test in (36) is to determine, based on the observation of $\boldsymbol{x} \in \mathcal{X}^n$ and $\boldsymbol{y} \in \mathcal{X}^n$, whether these vectors are realizations of the random variables in hypothesis $\mathcal{H}_0$ or $\mathcal{H}_1$. Consider a random transformation $P_{T|\boldsymbol{X}\boldsymbol{Y}}$ from $\mathcal{X}^n \times \mathcal{Y}^n \to \{0,1\}$. Note that this transformation can be a randomized test for the hypothesis test in (36). More specifically, $P_{T|\boldsymbol{X},\boldsymbol{Y}}(1|\boldsymbol{x},\boldsymbol{y}) = 1 - P_{T|\boldsymbol{X},\boldsymbol{Y}}(0|\boldsymbol{x},\boldsymbol{y})$ is the probability with which $\mathcal{H}_1$ is accepted given $\boldsymbol{x}$ and $\boldsymbol{y}$. Define the function $\beta_{1-\epsilon} : \Delta(\mathcal{X}^n \times \mathcal{Y}^n)^2 \to [0,1]$ by

$$
\beta_{1-\epsilon}(P_{\boldsymbol{X}\boldsymbol{Y}}, \bar{P}_{\boldsymbol{X}} Q_{\boldsymbol{Y}}) =
$$
$$
\inf_{P_{T|\boldsymbol{X}\boldsymbol{Y}}: \sum_{\boldsymbol{y} \in \mathcal{Y}^n} \sum_{\boldsymbol{x} \in \mathcal{X}^n} P_{T|\boldsymbol{X}\boldsymbol{Y}}(1|\boldsymbol{x},\boldsymbol{y}) P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) \bar{P}_{\boldsymbol{X}}(\boldsymbol{x}) \geq 1-\epsilon}
$$
$$
\left[\sum_{\boldsymbol{y} \in \mathcal{Y}^n} \sum_{\boldsymbol{x} \in \mathcal{X}^n} P_{T|\boldsymbol{X}\boldsymbol{Y}}(1|\boldsymbol{x},\boldsymbol{y}) \bar{P}_{\boldsymbol{X}}(\boldsymbol{x}) Q_{\boldsymbol{Y}}(\boldsymbol{y}) \right], \tag{40}
$$

that is, the minimum probability of falsely rejecting $\mathcal{H}_0$ given that the probability of successfully accepting $\mathcal{H}_1$ is lower bounded by $1-\epsilon$. Therefore, the following holds:

$$
\beta_{1-\epsilon}(P_{\boldsymbol{X}\boldsymbol{Y}}, \bar{P}_{\boldsymbol{X}} Q_{\boldsymbol{Y}}) = \gamma(\rho, q),
$$

where the function $\gamma : [0,1]^2 \to [0,1]$ is defined in (23).

Finally, from Theorem 29 in [10], it follows that $M < \Gamma(b)$, where the function Γ is defined in (25) and the optimization domain over ρ is due to Corollary 4 in [12]. This completes the proof of the information bound.

Finally, subject to a maximum ESP constraint, it follows that for any $(n, M, \epsilon, \delta, b)$-code described by the system in (6) for the random transformation in (2) satisfying (13), it follows from the Berry-Esseen Theorem [13] that for all $i \in \{1, 2, \ldots, M\}$:

$$
\begin{aligned}
\theta_i \geq& Q\left(\frac{(1-2\alpha_2)N(0|\boldsymbol{u}(i)) + n\left(\alpha_2 - \frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \right) \\
&- \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}.
\end{aligned}
\tag{41}
$$

Using the fact that for all $i \in \{1, 2, \ldots, M\}$, $\delta > \theta_i$, it follows that $b \leqslant B^+$, where the constant B^+ is defined in (26). This completes the proof of Theorem 2. ∎

REFERENCES

[1] N. Tesla, "Apparatus for transmitting electrical energy." Dec. 1 1914, US Patent 1,119,732.

[2] F. Tariq, M. R. A. Khandaker, K.-K. Wong, M. Imran, M. Bennis, and M. Debbah, "A speculative study on 6G," *online: https://arxiv.org/pdf/1902.06700.pdf*, Feb. 2019.

[3] S. Belhadj Amor and S. M. Perlaza, "Fundamental limits of simultaneous energy and information transmission," in *Proc. International Symposium on Telecommunications*, Thessaloniki, Greece, May 2016.

[4] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. IEEE International Symposium on Information Theory*, Toronto, ON, Canada, Jul. 2008, pp. 1612–1616.

[5] P. Grover and A. Sahai, "Shannon meets Tesla: Wireless information and power transfer," in *Proc. IEEE International Symposium on Information Theory*, Jun. 2010, pp. 2363–2367.

[6] A. M. Fouladgar and O. Simeone, "On the transfer of information and energy in multi-user systems," *IEEE Communications Letters*, vol. 16, no. 11, pp. 1733–1736, Nov. 2012.

[7] S. Belhadj Amor, S. M. Perlaza, I. Krikidis, and H. V. Poor, "Feedback enhances simultaneous wireless information and energy transmission in multiple access channels," *IEEE Transactions on Information Theory*, vol. 63, no. 8, pp. 5244–5265, Aug. 2017.

[8] N. Khalfet and S. M. Perlaza, "Simultaneous information and energy transmission in the two-user Gaussian interference channel," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 1, pp. 156–170, Jan. 2019.

[9] S. M. Perlaza, A. Tajer, and H. V. Poor, "Simultaneous energy and information transmission: A finite block-length analysis," in *Proc. of the 19th IEEE International Workshop on Signal Processing Advances in Wireless Communication*, Kalamata, Greece, Jun. 2018.

[10] Y. Polyanskiy, "Channel coding: Non-asymptotic fundamental limits," Ph.D. dissertation, Electrical Engineering, Princeton University, Princeton, NJ, USA, Nov. 2010.

[11] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[12] N. Khalfet, S. M. Perlaza, A. Tajer, and H. V. Poor, "On ultra-reliable and low latency simultaneous information and energy transmission systems," INRIA Grenoble - Rhône-Alpes, Tech. Rep. 9261, Mar. 2019.

[13] W. Feller, *An Introduction to Probability Theory and Its Application*, 2nd ed. New York, NY: John Wiley and Sons, 1971, vol. 2.