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

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

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

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

ISSN 0249-6399   ISRN INRIA/RR--9261--FR+ENG

RESEARCH REPORT
N° 9261
June 2019
Project-Teams Maracas

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

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

Project-Teams Maracas

Research Report n° 9261 — version 1 — initial version June 2019 — revised version Juin 2019 — 30 pages

Abstract: In this INRIA Research Report, 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 an energy harvester through the binary symmetric channel. 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 (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.

Key-words: Binary symmetric channel, simultaneous information and energy transmission, information-energy capacity region

---

Nizar Khalfet and Samir M. Perlaza are with the CITI Laboratory, a joint laboratory between INRIA, INSA de Lyon and the Université de Lyon. 6 Av. des Arts 69621 Villeurbanne, France. ({samir.perlaza, nizar.khalfet}@inria.fr). Ali Tajer is with the Department of Electrical, Computer, and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180. H. Vincent Poor (poor@princeton.edu) and Samir M. Perlaza are with the Department of Electrical Engineering at Princeton University, Princeton, NJ 08544 USA. This research was supported in part by the Ambassade de France aux Etats Unis under the Thomas Jefferson Funds, in part by the Agence Nationale de la Recherche under Grant ANR-15-NMED-0009-03, and in part by the U.S. National Science Foundation under Grants CNS-1702808 and ECCS-1647198.

RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

Inovallée
655 avenue de l'Europe Montbonnot
38334 Saint Ismier Cedex

Résumé : Dans ce rapport, les limites fondamentales de la transmission simultanée d'information et d'énergie dans le canal binaire symétrique sont déterminées. L'ensemble des débits atteignables de transmission d'information et d'énergie (en bits par utilisation canal et en unités d'énergie par utilisation canal respectivement) est identifié.

Mots-clés : Canal binaire symétrique , régime non-asymptotique, transmission simultanée d'information et d'énergie

Contents

1 Notation

Throughout this research report, 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 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. Given a binary vector $\boldsymbol{y}$, the Hamming distance between $\boldsymbol{x}$ and $\boldsymbol{y}$ is denoted by $d(\boldsymbol{x}, \boldsymbol{y})$ and

$$d(\boldsymbol{x}, \boldsymbol{y}) = \sum_{t=1}^{n} \mathbb{1}_{\{x_t \neq y_t\}}. \tag{1}$$

The expected value and the variance with respect to a random variable X is denoted by $\mathbb{E}_X[\cdot]$ and $\mathbb{V}_X[\cdot]$, respectively. The notation $\mathbb{S}_X[\cdot]$ denotes the third absolute moment with respect to the random variable X. The binary logarithm and natural logarithm functions are denoted by log and ln, respectively. The complementary cumulative distribution function $Q : \mathbb{R} \to [0,1]$ of the standard Gaussian distribution is

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

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 $\binom{n}{t} = 0$.

2 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 reseqrc 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).

The report is organized as follows. Section 3 formulates the problem and introduces the notion of the information-energy capacity region in the non-asymptotic block-length regime. Section 4 presents the main results. Section 5 concludes this work.

3 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

$$(\{0,1\}^n, \{0,1\}^n \times \{0,1\}^n, P_{\boldsymbol{YZ}|\boldsymbol{X}}), \tag{3}$$

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

$$P_{\boldsymbol{YZ}|\boldsymbol{X}}(\boldsymbol{y}, \boldsymbol{z}|\boldsymbol{x}) = \prod_{t=1}^{n} P_{Y|X}(y_t|x_t) P_{Z|X}(z_t|x_t), \tag{4}$$

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

$$P_{Y|X}(y|x) = \alpha_1 \mathbb{1}_{\{x \neq y\}} + (1-\alpha_1)\mathbb{1}_{\{x=y\}}, \tag{5}$$

$$P_{Z|X}(z|x) = \alpha_2 \mathbb{1}_{\{x \neq z\}} + (1-\alpha_2)\mathbb{1}_{\{x=z\}}, \tag{6}$$

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.

3.1 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 (3) is a system

$$\{(\boldsymbol{u}(1), \mathcal{D}_1), (\boldsymbol{u}(2), \mathcal{D}_2), \ldots, (\boldsymbol{u}(M), \mathcal{D}_M)\}, \tag{7}$$

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

$$\boldsymbol{u}(i) \triangleq (u_1(i), u_2(i), \ldots, u_n(i)) \in \{0, 1\}^n, \tag{8a}$$

$$\mathcal{D}_i \cap \mathcal{D}_j = \emptyset, \text{ and} \tag{8b}$$

$$\bigcup_{i=1}^{M} \mathcal{D}_i \subseteq \{0, 1\}^n. \tag{8c}$$

Given the system in (7), 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

$$(y_1, y_2, \ldots, y_n) \in \mathcal{D}_i. \tag{9}$$

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

$$\lambda_i \triangleq \Pr\left[\boldsymbol{Y} \in \mathcal{D}_i^{\mathsf{c}} \mid \boldsymbol{X} = \boldsymbol{u}(i)\right], \tag{10}$$

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

$$\lambda \triangleq \frac{1}{M} \sum_{m=1}^{M} \lambda_m. \tag{11}$$

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.

3.2 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

$$b_0 \triangleq g(0), \quad \text{and} \tag{12a}$$

$$b_1 \triangleq g(1) \tag{12b}$$

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

$$B_n(\boldsymbol{z}) \triangleq \frac{1}{n} \sum_{t=1}^{n} g(z_t) = (b_0 - b_1) \frac{N(0|\boldsymbol{z})}{n} + b_1. \tag{13}$$

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:

$$b_1 < b_0. \tag{14}$$

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

$$\theta_i \triangleq \Pr\left[B_n(\boldsymbol{Z}) < b \,|\, \boldsymbol{X} = \boldsymbol{u}(i)\right], \tag{15}$$

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

$$\theta \triangleq \frac{1}{M}\sum_{i=1}^{M}\theta_i. \tag{16}$$

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

$$b_1 \leqslant B_n(\boldsymbol{z}) \leqslant b_0. \tag{17}$$

The inequalities in (17) 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.

4 Fundamental Limits

The non-asymptotic fundamental limits of the system described in Section 3 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 an average or maximum DEP and an 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* (3) *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.*

4.1 Bounds on the Energy Transmission Rate

This section provides some upper-bounds on the energy transmission rate for any given $(n, M, \epsilon, \delta, b)$-code with either maximum or average ESP constraints. These bounds are expressed in terms of the parameters of the code (Definition 1), the parameters of the random transformation in (3) and the empirical input distributions induced by the code.

Definition 7 (Empirical Distributions) *Consider an (n, M)-code described by the system in* (7)*. For all $i \in \{1, 2, \ldots, M\}$, the empirical probability distribution of the channel input symbols induced by the codeword $\boldsymbol{u}(i)$ is*

$$\bar{P}_X^{(i)}(0) \triangleq \frac{1}{n} N(0|\boldsymbol{u}(i)) = 1 - \bar{P}_X^{(i)}(1), \tag{18}$$

The empirical distribution of the channel input symbols jointly induced by all codewords is denoted by

$$\bar{P}_X(0) \triangleq \frac{1}{nM} \sum_{i=1}^{M} N(0|\boldsymbol{u}(i)) = 1 - \bar{P}_X(1). \tag{19}$$

Often, the vector $\left(\bar{P}_X^{(i)}(0), \bar{P}_X^{(i)}(1)\right)$ is referred to as the *type* of the codeword $\boldsymbol{u}(i)$ [12].

Using the empirical distributions in Definition 7, some upper bounds on the energy transmission rate can be described. These upper bounds are obtained from the analysis of the ESP in (15). The following proposition provides the exact value of the ESP for any given $(n, M, \epsilon, \delta, b)$-code.

Proposition 1 (Ground-Truth ESP) *Consider an $(n, M, \epsilon, \delta, b)$-code described by the system in* (7) *for the random transformation in* (3) *satisfying* (14)*. Then, for all $i \in \{1, 2, \ldots, M\}$, the ESP in* (15) *satisfies*

$$\theta_i = \sum_{k=0}^{\left\lfloor \frac{n(b-b_1)}{b_0-b_1} \right\rfloor} \sum_{s=0}^{k} \binom{N(0|\boldsymbol{u}(i))}{s} \binom{N(1|\boldsymbol{u}(i))}{k-s} (1-\alpha_2)^{N(1|\boldsymbol{u}(i))-k+2s} \alpha_2^{N(0|\boldsymbol{u}(i))+k-2s}. \tag{20}$$

Proof: The proof of Proposition 1 is presented in Appendix B. ■

The equality in (20) together with Definition 4 and Definition 5 provide the first bounds on the energy transmission rate of any given code. The following corollary describes these bounds.

Corollary 1 *Consider an $(n, M, \epsilon, \delta, b)$-code described by the system in* (7) *for the random transformation in* (3) *satisfying* (14)*. Then, subject to a maximal ESP constraint, it holds that,*

$$b < \tilde{B} \tag{21}$$

where $\tilde{B}$ is the largest real that satisfies for all $i \in \{1, 2, \ldots, M\}$,

$$\sum_{k=0}^{\left\lfloor \frac{n(\tilde{B}-b_1)}{b_0-b_1} \right\rfloor} \sum_{s=0}^{k} \binom{N(0|\boldsymbol{u}(i))}{s} \binom{N(1|\boldsymbol{u}(i))}{k-s} (1-\alpha_2)^{N(1|\boldsymbol{u}(i))-k+2s} \alpha_2^{N(0|\boldsymbol{u}(i))+k-2s} < \delta, \tag{22}$$

and subject to an average ESP constraint, the energy rate b satisfies

$$b < \breve{B} \tag{23}$$

where $\breve{B}$ is the biggest positive real that satisfies

$$\frac{1}{M}\sum_{i=1}^{M}\sum_{k=0}^{\left\lfloor \frac{n(\breve{B}-b_1)}{b_0-b_1}\right\rfloor}\sum_{s=0}^{k}\binom{N\left(0|\boldsymbol{u}(i)\right)}{s}\binom{N\left(1|\boldsymbol{u}(i)\right)}{k-s}(1-\alpha_2)^{N(1|\boldsymbol{u}(i))-k+2s}\alpha_2^{N(0|\boldsymbol{u}(i))+k-2s} < \delta. \tag{24}$$

The bounds in Corollary 1 are not in closed-form and thus, are difficult to calculate. Moreover, they bring very little insight to obtain a bound on the energy rate b at which an $(n, M, \epsilon, \delta, b)$ code can transmit energy. Therefore, it would be desirable to approximate the individual ESP in order to obtain an upper bound on the energy transmission rate in a closed form expression, probably, at the expense of some precision. The following proposition provides some approximations on the ESP using tools from large deviations theory [13].

Proposition 2 *Consider an $(n, M, \epsilon, \delta, b)$-code described by the system in (7) for the random transformation in (3) satisfying (14). Then, for all $i \in \{1, 2, \dots, M\}$, the ESP in (15) satisfies*

$$\theta_i > 1-\exp\left(-n\frac{\left(\frac{b-b_1}{b_0-b_1}-\left(\left(1-2\alpha_2\right)\bar{P}_X^{(i)}(0)+\alpha_2\right)\right)^2}{\frac{b-b_1}{b_0-b_1}+\left(\left(1-2\alpha_2\right)\bar{P}_X^{(i)}(0)+\alpha_2\right)}\right), \tag{25}$$

and

$$\theta_i < \exp\left(-n\frac{\left(\frac{b-b_1}{b_0-b_1}-\left(\left(1-2\alpha_2\right)\bar{P}_X^{(i)}(0)+\alpha_2\right)\right)^2}{2\left(\left(1-2\alpha_2\right)\bar{P}_X^{(i)}(0)+\alpha_2\right)}\right). \tag{26}$$

Proof: The proof of Proposition 2 is presented in Appendix C. ■

Using Proposition 2, the energy transmission rate of an $(n, M, \epsilon, \delta, b)$-code can be upper-bounded.

Proposition 3 (Large Deviation Bound) *Consider an $(n, M, \epsilon, \delta, b)$-code described by the system in (7) for the random transformation in (3) satisfying (14). Then, subject to a maximal ESP constraint, it holds that for all $i \in \{1, 2, \dots, M\}$,*

$$b < (b_0-b_1)\left(\left(1-2\alpha_2\right)\bar{P}_X^{(i)}(0)+\alpha_2\right)+b_1+\frac{b_0-b_1}{\sqrt{n}}\sqrt{-2\left(\left(1-2\alpha_2\right)\bar{P}_X^{(i)}(0)+\alpha_2\right)\log(1-\delta)} - \frac{b_0-b_1}{n}\log(1-\delta). \tag{27}$$

and subject to an average ESP constraint, the energy rate b satisfies

$$b < \hat{B}, \tag{28}$$

where $\hat{B}$ is the biggest positive real that satisfies

$$1-\delta < \frac{1}{M}\sum_{i=1}^{M}\exp\left(-n\frac{\left(\frac{b-b_1}{b_0-b_1}-\left(\left(1-2\alpha_2\right)\bar{P}_X^{(i)}(0)+\alpha_2\right)\right)^2}{\frac{b-b_1}{b_0-b_1}+\left(\left(1-2\alpha_2\right)\bar{P}_X^{(i)}(0)+\alpha_2\right)}\right). \tag{29}$$

Proof: The proof of Proposition 3 is presented in Appendix D. ■
Another approximation to the ESP in (15) is obtained from the Berry-Esseen theorem (Theorem 3 in Appendix A). The following lemma presents this approximation.

Lemma 1 *Consider an $(n, M, \epsilon, \delta, b)$-code described by the system in (7) for the random transformation in (3) satisfying (14). Then, for all $i \in \{1, 2, \dots, M\}$, the ESP in (15) satisfies*

$$\theta_i \geq Q\left(\frac{n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0)+\alpha_2-\frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right)-\frac{n\left[\alpha_2(1-\alpha_2)^3+(1-\alpha_2)\alpha_2^3\right]}{2\left(n\alpha_2(1-\alpha_2)\right)^{3/2}} \tag{30}$$

and

$$\theta_i \leq Q\left(\frac{n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0)+\alpha_2-\frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right)+\frac{n\left[\alpha_2(1-\alpha_2)^3+(1-\alpha_2)\alpha_2^3\right]}{2\left(n\alpha_2(1-\alpha_2)\right)^{3/2}}. \tag{31}$$

Proof: The proof of Lemma 1 is presented in Appendix E. ■
Using Lemma 1, the energy transmission rate of an $(n, M, \epsilon, \delta, b)$-code can be upper-bounded.

Proposition 4 (Gaussian Approximation Bound) *Consider an $(n, M, \epsilon, \delta, b)$-code described by the system in (7) for the random transformation in (3) satisfying (14). Then, subject to a maximal ESP constraint, it holds that for all $i \in \{1, 2, \dots, M\}$,*

$$b \leqslant (b_0-b_1)\Big((1-2\alpha_2)\,\bar{P}_X^{(i)}(0)+\alpha_2\Big)+b_1-\sqrt{\frac{(b_0-b_1)^2\alpha_2(1-\alpha_2)}{n}}Q^{-1}\left(\delta+\frac{(1-\alpha_2)^2+\alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}\right), \tag{32}$$

and subject to an average ESP constraint, the energy rate b satisfies

$$b \leqslant \hat{B}, \tag{33}$$

where $\hat{B}$ is the biggest positive real that satisfies

$$\frac{1}{M}\sum_{i=1}^{M} Q\left(\frac{n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0)+\alpha_2-\frac{\hat{B}-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right)-\frac{(1-\alpha_2)^2+\alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}} < \delta. \tag{34}$$

Proof: The proof of Proposition 4 is presented in Appendix F. ■
Note that the upper bound in (32), is valid when the following condition is satisfied

$$0 < \delta + \frac{(1-\alpha_2)^2+\alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)^2}} < 1, \tag{35}$$

given that the domain of the function Q^{-1} is $(0, 1)$. In general, the bounds presented in Proposition 4 are tighter than those presented in Proposition 3 for small values of the block length n. Nonetheless, the bounds in Proposition 3 are easier to calculate and perform equally well for large n.

Note that the input distribution that achieves the largest information transmission rate, without any energy constraint, is the uniform distribution [10]. That is, $\bar{P}_X^{(i)}(0) = 1-\bar{P}_X^{(i)}(1) = \frac{1}{2}$. Hence, Proposition 4 provides an outer bound on the energy rate that can be transmitted by an (n, M, ϵ)-code that possesses an empirical input distribution that is uniform. The following corollary describes this observation.

Corollary 2 *Consider an $(n, M, \epsilon, \delta, b)$-code for the random transformation in* (4) *satisfying* (14)*. Assume that such a code exhibits an information-rate optimal empirical distribution. Then, it follows that $b < \underline{b}(n, \delta)$, with $\underline{b} : \mathbb{N} \times [0, 1] \to \mathbb{R}$, such that*

$$\underline{b}(n, \delta) \triangleq \frac{b_0 + b_1}{2} - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1 - \alpha_2)}{n}} Q^{-1}\left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}}\right). \tag{36}$$

Essentially, Corollary 2 determines a threshold on the energy rate b beyond which an $(n, M, \epsilon, \delta, b)$-code, if it exists, exhibits a conflict between the energy transmission task and the information transmission task. More specifically, if there exists an $(n, M, \epsilon, \delta, b)$-code whose energy transmission rate b is beyond the threshold $\underline{b}$ in Corollary 2, it exhibits an empirical input distribution for which $\bar{P}_X(0) > \bar{P}_X(1)$. This implies than a zero is transmitted more often than a one, which is not information-rate optimal.

Proposition 4 also provides upper bounds on the largest energy rate that can be transmitted by any $(n, M, \epsilon, \delta, b)$-code. Note that the largest energy-transmission rate is achieved by a zero information-rate code whose codewords contain only zeros, i.e., $\bar{P}_X(0) = 1 - \bar{P}_X(1) = 1$. The following corollary describes this observation.

Corollary 3 *Consider an $(n, M, \epsilon, \delta, b)$-code for the random transformation in* (3) *satisfying* (14)*. Then, it follows that: $b < \bar{b}(n, \delta)$, with $\bar{b} : \mathbb{N} \times [0, 1] \to \mathbb{R}$, such that*

$$\bar{b}(n, \delta) \triangleq (1 - \alpha_2)b_0 + \alpha_2 b_1 - \sqrt{\frac{(b_0 - b_1)^2 \alpha_2 (1 - \alpha_2)}{n}} Q^{-1}\left(\delta + \frac{(1 - \alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1 - \alpha_2)}}\right). \tag{37}$$

From Corollary 2 and Corollary 3, it might be expected that the energy transmission task and the information transmission task exhibit a conflicting interaction. The following section explores this particular interaction.

Let $\rho^* : [b_1, b_0] \to [0, 1]$ be defined as

$$\rho^*(b) \triangleq \min\left(1, \rho^+(b)\right), \tag{38}$$

with $\rho^+ : [b_1, b_0] \to \mathbb{R}_+$ such that

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

Note that for a fixed $b > 0$, $\rho^+(b)$ describes the empirical input distribution that saturates the inequality in (32). The following corollary from Proposition 4 highlights this observation.

Corollary 4 *Consider an $(n, M, \epsilon, \delta, b)$-code described by the system in* (7) *for the random transformation in* (3) *satisfying* (14)*. Then, subject to a maximal energy-shortage probability constraint, for all $i \in \{1, 2, \ldots, M\}$, the empirical input distribution $\bar{P}_X^{(i)}$ satisfies*

$$\bar{P}_X^{(i)}(0) \geqslant \rho^*(b), \tag{39}$$

where $\rho^*(b)$ is defined in (38)*.*

Corollary 4 leads to interesting conclusions by noticing that $\rho^*(b)$ is a lower bound on the fraction of zeros in each codeword (maximal energy-shortage probability) when energy is transmitted at an average energy rate b. This is natural from the perspective of the assumption in (14), which implies that the symbol zero carries more energy that the symbol one.

An interesting class of codes is that of homogeneous codes. A formal definition of these codes is hereunder.

Definition 8 (Homogeneous Codes) *A code $\mathcal{C}$ described by the system in* (7) *is said to be homogeneous if the following conditions hold:*

$$N(0|\boldsymbol{u}(1)) = N(0|\boldsymbol{u}(2)) = \ldots = N(0|\boldsymbol{u}(M)) \text{ and} \quad (40)$$
$$N(1|\boldsymbol{u}(1)) = N(1|\boldsymbol{u}(2)) = \ldots = N(1|\boldsymbol{u}(M)). \quad (41)$$

The interest in this class of codes stems from the fact that an average ESP constraint or a maximum ESP constraint leads to the same fundamental limits on the energy rate.

Corollary 5 *Consider an $(n, M, \epsilon, \delta, b)$-code described by the system in* (7) *for the random transformation in* (3) *satisfying* (14)*, and assume it is a homogeneous code. Then, the bounds on the energy rate b subject to a maximum ESP and average ESP are identical. That is, the bound in* (21) *is identical to* (23)*; the one in* (27) *is identical to* (28)*; and the one in* (32) *is identical to* (33)*.*

4.2 Bounds on the Information Transmission Rate

Given an $(n, M, \epsilon, \delta, b)$-code, the following lemma describes a bound on M, which does not take into account the decoding error probability ϵ and thus, it might be loose. However, it plays an important role when $\underline{b}(n, \delta) < b < \overline{b}(n, \delta)$.

Lemma 2 Consider an $(n, M, \epsilon, \delta, b)$-code for the random transformation in (3) satisfying (14). Then, subject to a maximal energy-shortage probability constraint, it holds that

$$M \leqslant \binom{n}{\lceil n\rho^*(b)\rceil} 2^{(n-\lceil n\rho^*(b)\rceil)}, \quad (42)$$

where $\rho^*(b)$ is defined by (38).

Proof of Proposition 2: Corollary 4 provides an approximation to the minimum number of zeros in each codeword in any given $(n, M, \epsilon, \delta, b)$-code with maximal energy-shortage probability. That is, for all $i \in \{1, 2, \ldots, M\}$ if follows that

$$N(0|\boldsymbol{u}(i)) \geqslant \lceil n\rho^*(b)\rceil. \quad (43)$$

This immediately provides an upper-bound on M given that all codewords must contain at least $\lceil n\rho^*(b)\rceil$ zeros. Hence, the right-hand side of (42) is the maximum number of codewords of length n for which at least $\lceil n\rho^*(b)\rceil$ symbols are zeros. This completes the proof. ■

Note that $\rho^*(b)$ is monotonically increasing with the energy rate b. Interestingly, when $\rho^*(b) \in (\frac{1}{2}, 1]$, the right-hand sides of (42) is monotonically decreasing with b. This highlights the existing trade-off between the information transmission task and the energy transmission task. That is, in the regime in which $\rho^*(b) \in (\frac{1}{2}, 1]$, increasing the energy rate would necessarily imply decreasing the information rate.

4.3 Information-Energy Capacity Region

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 (3) subject to (14) 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,

$$\underline{\mathcal{C}}(n, \epsilon, \delta) \subseteq \mathcal{C}(n, \epsilon, \delta) \subseteq \overline{\mathcal{C}}(n, \epsilon, \delta). \quad (44)$$

This approximation is obtained by considering an average DEP constraint and a maximum ESP constraint. The following notation is used to describe the set $\underline{\mathcal{C}}(n,\epsilon,\delta)$. Let functions $\phi : \mathbb{N} \times [0,1] \to [0,1]$, and $\chi : \mathbb{R}_+ \times [0,1] \to [0,1]$ be defined as,

$$\phi(m,\rho) \triangleq \min\Bigg\{1, (m-1) \sum_{\ell_0=0}^{n} \sum_{\ell_1=0}^{n} \sum_{\ell_2=0}^{\min\{\ell_0,\ell_1\}} \sum_{\ell_3=0}^{\ell_0+\ell_1-2\ell_2} \sum_{\ell_4=0}^{\ell_3} \binom{n}{\ell_0}\binom{\ell_0}{\ell_2}\binom{n-\ell_0}{\ell_1-\ell_2}\binom{\ell_0}{\ell_4}\binom{n-\ell_0}{\ell_3-\ell_4}$$
$$\cdot \alpha_1^{\ell_0+\ell_1-2\ell_2}(1-\alpha_1)^{n-\ell_0-\ell_1+2\ell_2}\rho^{\ell_3+\ell_0-2\ell_4+\ell_1}(1-\rho)^{2n-(\ell_3+\ell_0-2\ell_4+\ell_1)}\Bigg\}, \tag{45}$$

and

$$\chi(s,\rho) \triangleq \sum_{t=0}^{n}\binom{n}{t}\rho^t(1-\rho)^{n-t}Q\left(\frac{(1-2\alpha_2)t + n\left(\alpha_2 - \frac{s-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right) + \frac{(1-\alpha_2)^2+\alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}. \tag{46}$$

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

$$M_1^\star(\rho) \triangleq \begin{cases} \underset{m\in\mathbb{N}}{\operatorname{argmax}}\, \phi(m,\rho) \\ \text{s.t.} \quad \phi(m,\rho) < \epsilon \end{cases}, \tag{47}$$

and

$$B^\star(\rho) \triangleq \begin{cases} \underset{s\in\mathbb{R}_+}{\operatorname{argmax}}\, \chi(s,\rho) \\ \text{s.t.} \quad \chi(s,\rho) < \delta \end{cases}. \tag{48}$$

Note that functions ϕ in (45), χ in (46), $M_1^\star$ in (47), and $B^\star$ in (48) depend on the block length n, the parameters of the random transformation in (3), i.e., α_1 and α_2, and the energy harvested from symbols 0 and 1, i.e., b_0 and b_1 in (12). Nonetheless, none of these parameters is put as an argument 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* (3) *subject to* (14)*, contains the set*

$$\underline{\mathcal{C}}(n,\epsilon,\delta) \triangleq \Big\{(M,b) \in \mathbb{N}\times\mathbb{R}_+ : \exists\rho\in[0,1], M < M_1^\star(\rho) \,\text{and}\, b < B^*(\rho)\Big\}, \tag{49}$$

where $M_1^\star : [0,1] \to \mathbb{N}$ is defined in (47) *and $B^\star : [0,1] \to \mathbb{R}_+$ is defined in* (48).

Proof of Theorem 1: The proof of Theorem 1 is presented in Appendix G. ■
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

$$\gamma(\rho,q) = \sum_{\ell_0=0}^{n}\sum_{\ell_1=0}^{n}\sum_{\ell_2=0}^{\min\{\ell_0,\ell_1\}}\binom{n}{\ell_0}\binom{\ell_0}{\ell_2}\binom{n-\ell_0}{\ell_1-\ell_2}\rho^{\ell_1}(1-\rho)^{n-\ell_1}q^{\ell_0}(1-q)^{n-\ell_0} \tag{50}$$
$$\left(\mathbb{1}_{\left\{(\ell_0+\ell_1-2\ell_2)\log\frac{\alpha_1}{1-\alpha_1}+\ell_0\log\frac{1-q}{q}>L\right\}} + \lambda\mathbb{1}_{\left\{(\ell_0+\ell_1-2\ell_2)\log\frac{\alpha_1}{1-\alpha_1}+\ell_0\log\frac{1-q}{q}=L\right\}}\right)$$

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

$$1-\epsilon = \sum_{\ell_0=0}^{n}\sum_{\ell_1=0}^{n}\sum_{\ell_2=0}^{\min\{\ell_0,\ell_1\}} \binom{n}{\ell_0}\binom{\ell_0}{\ell_2}\binom{n-\ell_0}{\ell_1-\ell_2}(1-\alpha_1)^{n-(\ell_0+\ell_1-2\ell_2)}\alpha_1^{\ell_0+\ell_1-2\ell_2}\rho^{\ell_1}(1-\rho)^{n-\ell_1}$$
$$\cdot\left(\mathbb{1}_{\left\{(\ell_0+\ell_1-2\ell_2)\log\frac{\alpha_1}{1-\alpha_1}+\ell_0\log\frac{1-q}{q}>L\right\}} + \lambda\mathbb{1}_{\left\{(\ell_0+\ell_1-2\ell_2)\log\frac{\alpha_1}{1-\alpha_1}+\ell_0\log\frac{1-q}{q}=L\right\}}\right).$$

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

$$\Gamma(b) \triangleq \sup_{\rho>\rho^*(b)} \inf_{q\in[0,\frac{1}{2}]} \frac{1}{\gamma(\rho,q)}. \tag{51}$$

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

$$B^+ \triangleq (1-\alpha_2)b_0 + \alpha_2 b_1 - \sqrt{\frac{(b_0-b_1)^2\alpha_2(1-\alpha_2)}{n}} \cdot Q^{-1}\left(\delta + \frac{(1-\alpha_2)^2+\alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}\right). \tag{52}$$

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* (3) *subject to* (14)*, is contained in the set*

$$\bar{\mathcal{C}}(n,\epsilon,\delta) \triangleq \left\{(M,b) \in \mathbb{N}\times\mathbb{R}_+ : M < \Gamma(b) \text{ and } b < B^+\right\}, \tag{53}$$

where $\Gamma : \mathbb{R}_+ \to [0,1]$ is defined in (51) *and B^+ is defined in* (52).

Proof: The proof of Theorem 2 is presented in Appendix H. ■

5 Conclusions

In this research report, the fundamental limits of SIET have been studied under the assumption that the transmission occurs during a finite number of channel uses at the expense of strictly positive DEP and ESP. From this perspective, a non-asymptotic fundamental limit has been introduced: the information-energy capacity region, that is, the largest set of jointly achievable energy and information rates. The focus has been on the case of one transmitter, one IR and one EH communicating via binary symmetric memoryless channels. In this case, given a finite block length, a DEP, and an ESP, four scenarios have been observed depending on whether an average or maximal constraint is imposed on the DEP and the ESP. These results have revealed the competitive interaction between the information transmission task and energy transmission task. In particular, a certain regime in which increasing the information rate necessarily implies decreasing the energy rate and *vice versa* has been identified.

Appendices

A Preliminary Results

This section introduces some auxiliary results that play a key role in the following appendices.

Definition 9 (Moment Generating Function) *Given a random variable X, its moment generating function is denoted by $\phi_X : \mathbb{R} \to \mathbb{R}$ and*

$$\phi_X(\lambda) = \mathbb{E}_X \left[\mathrm{e}^{\lambda X} \right]. \tag{54}$$

The following lemmas highlight some properties of the moment generating function.

Lemma 3 *Let $Z = \sum_{i=1}^{n} X_i$ be a random variable formed by the sum of n independent random variables $X_1, X_2, \ldots, X_n$. Then , for all $\lambda \in \mathbb{R}$,*

$$\phi_Z(\lambda) = \prod_{t=1}^{n} \phi_{X_t}(\lambda). \tag{55}$$

Proof:

$$\begin{aligned}
\phi_Z(\lambda) &= \mathbb{E}_Z \left[\mathrm{e}^{\lambda Z} \right] && (56)\\
&= \mathbb{E}_{X_1, X_2, \ldots, X_N} \left[\mathrm{e}^{\lambda \sum_{t=1}^{n} X_t} \right] && (57)\\
&= \mathbb{E}_{X_1, X_2, \ldots, X_N} \prod_{t=1}^{n} \mathrm{e}^{\lambda X_t} && (58)\\
&= \prod_{t=1}^{n} \mathbb{E}_{X_t} \left[\mathrm{e}^{\lambda X_t} \right] && (59)\\
&= \prod_{t=1}^{n} \phi_{X_t}(\lambda). && (60)
\end{aligned}$$

■

Lemma 4 *Let X be a Bernoulli random variable with $P_X(1) = 1 - P_X(0) = \rho$. Then, for all $\lambda \in \mathbb{R}$*

$$\phi_X(\lambda) = 1 + \rho \left(\mathrm{e}^{\lambda} - 1 \right). \tag{61}$$

Proof:

$$\begin{aligned}
\mathbb{E}_X \left[\mathrm{e}^{\lambda X} \right] &= P_X(1)\mathrm{e}^{\lambda} + P_X(0) && (62)\\
&= \rho \mathrm{e}^{\lambda} + (1 - \rho) && (63)\\
&= 1 + \rho \left(\mathrm{e}^{\lambda} - 1 \right). && (64)
\end{aligned}$$

■

Theorem 3 (Berry-Esseen Theorem, [14]) *Let X_1, X_2, $\ldots$, X_n be independent random variables such that for all $t \in \{1, 2, \ldots, n\}$,*

$$\begin{aligned}
\mu_t &= \mathbb{E}_{X_t} \left[X_t \right], && (65)\\
\sigma_t^2 &= \mathbb{E}_{X_t} \left[X_t^2 \right] - \mu_t^2, && (66)\\
\phi_t &= \mathbb{E}_{X_t} \left[|X_t - \mu_t|^3 \right]. && (67)
\end{aligned}$$

Then, it holds for all $\lambda \in \mathbb{R}$ that

$$\left| \Pr \left[\sum_{t=1}^{n} X_t - \mu_t \geqslant \sigma \lambda \right] - Q(\lambda) \right| \leqslant \frac{c_0 \phi}{\sigma^3}, \tag{68}$$

where

$$\mu = \sum_{t=1}^{n} \mu_t, \quad \sigma^2 = \sum_{t=1}^{n} \sigma_t^2, \quad \textit{and} \quad \phi = \sum_{t=1}^{n} \phi_t. \tag{69}$$

The best value of the constant c_0 is $c_0 = 0.4748$ [15].

B Proof of Proposition 1

Consider the definition of the ESP in (15). Hence, for all $i \in \{1, 2, \ldots, M\}$,

$$\theta_i = \Pr\left[\sum_{t=1}^{n} \mathbb{1}_{\{Z_t=0\}} < \left(\frac{n(b-b_1)}{b_0 - b_1} \right) \middle| \boldsymbol{X} = \boldsymbol{u}(i) \right]. \tag{70}$$

Assume that the Transmitter uses an $(n, M, \epsilon, \delta, b)$-code and it aims at sending the message index $i \in \{1, 2, \ldots, M\}$. Then, for all $t \in \{1, 2, \ldots, n\}$, the random variable $\mathbb{1}_{\{Z_t=0\}}$ in (70) follows a Bernoulli distribution and the probability of a "one" is

$$P_{Z|X}(0|u_t(i)) = \begin{cases} \alpha_2 & \text{if} \quad u_t(i) = 1 \\ 1 - \alpha_2 & \text{if} \quad u_t(i) = 0 \end{cases}. \tag{71}$$

This implies that the random variable $\sum_{t=1}^{n} \mathbb{1}_{\{Z_t=0\}}$ can be expressed as the sum of two random variables with binomial distributions $\mathcal{B}(N(0|\boldsymbol{u}(i)), 1 - \alpha_2)$ and $\mathcal{B}(N(1|\boldsymbol{u}(i)), \alpha_2)$, respectively. That is,

$$\sum_{t=1}^{n} \mathbb{1}_{\{Z_t=0\}} = \sum_{t \in \{m : u_m(i)=0\}} \mathbb{1}_{\{Z_t=0\}} + \sum_{t \in \{m : u_m(i)=1\}} \mathbb{1}_{\{Z_t=0\}}. \tag{72}$$

Hence,

$$\begin{aligned}
\theta_i =& \sum_{m=0}^{\left\lfloor \frac{n(b-b_1)}{b_0-b_1} \right\rfloor} \Pr\left[\sum_{t \in \{m : u_m(i)=0\}} \mathbb{1}_{\{Z_t=0\}} + \sum_{t \in \{m : u_m(i)=1\}} \mathbb{1}_{\{Z_t=0\}} = m \middle| \boldsymbol{X} = \boldsymbol{u}(i) \right] \\
=& \sum_{k=0}^{\left\lfloor \frac{n(b-b_1)}{b_0-b_1} \right\rfloor} \sum_{s=0}^{k} \Pr\left[\sum_{t \in \{m : u_m(i)=0\}} \mathbb{1}_{\{Z_t=0\}} = s \middle| \boldsymbol{X} = \boldsymbol{u}(i) \right] \\
& \Pr\left[\sum_{t \in \{m : u_m(i)=1\}} \mathbb{1}_{\{Z_t=0\}} = k - s \middle| \boldsymbol{X} = \boldsymbol{u}(i) \right] \\
=& \sum_{k=0}^{\left\lfloor \frac{n(b-b_1)}{b_0-b_1} \right\rfloor} \sum_{s=0}^{k} \binom{N(0|\boldsymbol{u}(i))}{s} (1-\alpha_2)^s \alpha_2^{N(0|\boldsymbol{u}(i))-s} \binom{N(1|\boldsymbol{u}(i))}{k-s} (1-\alpha_2)^{N(1|\boldsymbol{u}(i))-k+s} \alpha_2^{k-s} \\
=& \sum_{k=0}^{\left\lfloor \frac{n(b-b_1)}{b_0-b_1} \right\rfloor} \sum_{s=0}^{k} \binom{N(0|\boldsymbol{u}(i))}{s} \binom{N(1|\boldsymbol{u}(i))}{k-s} (1-\alpha_2)^{N(1|\boldsymbol{u}(i))-k+2s} \alpha_2^{N(0|\boldsymbol{u}(i))+k-2s},
\end{aligned} \tag{73}$$

which completes the proof.

C Proof of Proposition 2

For all $i \in \{1, 2, \ldots, M\}$, let V_i be the following random variable:

$$V_i = \sum_{t \in \{m : u_m(i) = 0\}} \mathbb{1}_{\{Z_t = 0\}} + \sum_{t \in \{m : u_m(i) = 1\}} \mathbb{1}_{\{Z_t = 0\}}. \tag{74}$$

From Lemma 3 and Lemma 4, the following holds,

$$\begin{aligned}
\phi_{V_i}(\lambda) &= \prod_{t \in \{m:u_m(i)=0\}}^{n} 1 + (1-\alpha_2)(\mathrm{e}^{\lambda} - 1) \prod_{t \in \{m:u_m(i)=1\}}^{n} 1 + \alpha_2(\mathrm{e}^{\lambda} - 1) &(75)\\
&\leq \prod_{t \in \{m:u_m(i)=0\}}^{n} \exp\left((1-\alpha_2)(\mathrm{e}^{\lambda} - 1)\right) \prod_{t \in \{m:u_m(i)=1\}}^{n} \exp\left(\alpha_2(\mathrm{e}^{\lambda} - 1)\right) &(76)\\
&\leq \exp\left(N(0|\boldsymbol{u}(i))(1-\alpha_2)(\mathrm{e}^{\lambda} - 1) + N(1|\boldsymbol{u}(i))\alpha_2(\mathrm{e}^{\lambda} - 1)\right) &(77)\\
&= \exp\left((\mathrm{e}^{\lambda} - 1)\left(N(0|\boldsymbol{u}(i))(1-\alpha_2) + N(1|\boldsymbol{u}(i))\alpha_2\right)\right) &(78)\\
&= \exp\left(n(\mathrm{e}^{\lambda} - 1)\left(\bar{P}_X^{(i)}(0)(1-\alpha_2) + (1 - \bar{P}_X^{(i)}(0))\alpha_2\right)\right) &(79)\\
&= \exp\left(n(\mathrm{e}^{\lambda} - 1)(\bar{P}_X^{(i)}(0)(1-2\alpha_2) + \alpha_2)\right), &(80)
\end{aligned}$$

where (76) follows from the fact that for all $x \in \mathbb{R}, 1 + x \leq \mathrm{e}^x$. In order to ease the notation, let μ be

$$\mu \triangleq n\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2) + \alpha_2\right). \tag{81}$$

From Markov's inequality, it holds that for all $(\lambda, \gamma) \in \mathbb{R}^2$,

$$\begin{aligned}
\Pr\left[V_i > (1+\gamma)\mu\right] &= \Pr\left[\mathrm{e}^{\lambda V_i} > \mathrm{e}^{\lambda(1+\gamma)\mu}\right] &(82)\\
&\leq \frac{\mathbb{E}_{V_i}\left[\mathrm{e}^{\lambda V_i}\right]}{\mathrm{e}^{\lambda(1+\gamma)\mu}} &(83)\\
&= \frac{\phi_{V_i}(\lambda)}{\mathrm{e}^{\lambda(1+\gamma)\mu}} &(84)\\
&\leq \frac{\exp\left((\mathrm{e}^{\lambda} - 1)\mu\right)}{\mathrm{e}^{\lambda(1+\gamma)\mu}} &(85)\\
&= \exp\left((\mathrm{e}^{\lambda} - 1)\mu - \lambda(1+\gamma)\mu\right) &(86)\\
&= \exp\left(\mu(\mathrm{e}^{\lambda} - 1 - \lambda(1-\gamma))\right). &(87)
\end{aligned}$$

Note that the choice of λ can be improved to tight the bound in (87). Note that,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathrm{e}^{\lambda} - (1+\gamma)\lambda = \mathrm{e}^{\lambda} - (1+\gamma). \tag{88}$$

Then, the optimal λ is the solution to $\mathrm{e}^{\lambda} - (1+\gamma) = 0$. That is, $\lambda = \log(1+\gamma)$. This implies:

$$\begin{aligned}
\Pr\left[V_i > (1+\gamma)\mu\right] &\leq \exp\left(\mu\left(1 + \gamma - 1 - (1+\gamma)\log(1+\gamma)\right)\right) &(89)\\
&= \exp\left(\mu\left(\gamma - (1+\gamma)\log(1+\gamma)\right)\right) &(90)\\
&= \left(\frac{\mathrm{e}^{\lambda}}{(1+\gamma)^{(1+\gamma)}}\right)^{\mu} &(91)\\
&< \exp\left(-\frac{\gamma^2}{2+\gamma}\mu\right), &(92)
\end{aligned}$$

where (92) follows from the fact that

$$\log\left(\left(\frac{\mathrm{e}^{\lambda}}{(1+\gamma)^{(1+\gamma)}}\right)^{\mu}\right) = \mu\log(\exp\gamma) - \mu(1+\gamma)\log(1+\gamma) \tag{93}$$

$$\leq \mu\left(\gamma - \frac{2(1+\gamma)\gamma}{2+\gamma}\right) \tag{94}$$

$$= \mu\left(-\frac{\gamma^2}{2+\gamma}\right), \tag{95}$$

and the inequality in (94) is due to the fact that for all $x > 0$, $\log(1+x) \geq \frac{x}{1+\frac{x}{2}}$. Hence, it follows that

$$\Pr\left[\sum_{t=1}^{n}\mathbb{1}_{\{Z_t=0\}} > \frac{n(b-b_1)}{b_0-b_1}\right]$$

$$= \Pr\left[\sum_{t=1}^{n}\mathbb{1}_{\{Z_t=0\}} > \left(1+\left(\frac{b-b_1}{\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)}\right)-1\right)\right.$$

$$\left. n\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)\right] \tag{96}$$

$$\leq \exp\left(\frac{-n\left(\frac{b-b_1}{\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)}-1\right)^2\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)}{\frac{b-b_1}{\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)}}\right) \tag{97}$$

$$= \exp\left(-n\frac{\left(\frac{b-b_1}{b_0-b_1}-\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\right)^2}{\left(\frac{b-b_1}{b_0-b_1}+\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\right)}\right). \tag{98}$$

This completes the proof of (25).

On the other hand, from Markov's inequality, it holds that for all $(\lambda,\gamma) \in \mathbb{R}^2$,

$$\Pr\left[V_i < (1-\gamma)\mu\right] = \Pr\left[\mathrm{e}^{-\lambda V_i} > \mathrm{e}^{-\lambda(1-\gamma)\mu}\right] \tag{99}$$

$$\leq \frac{\mathbb{E}_{V_i}\left[\mathrm{e}^{-\lambda V_i}\right]}{\mathrm{e}^{\lambda(1-\gamma)\mu}} \tag{100}$$

$$= \frac{\phi_{V_i}(-\lambda)}{\mathrm{e}^{-\lambda(1-\gamma)\mu}} \tag{101}$$

$$\leq \frac{\exp\left((\mathrm{e}^{-\lambda}-1)\mu\right)}{\mathrm{e}^{-\lambda(1-\gamma)\mu}} \tag{102}$$

$$= \exp\left(\mu(\mathrm{e}^{-\lambda}-1+\lambda(1-\gamma))\right). \tag{103}$$

Note that the choice of λ can be improved to tight the bound in (103). Note that,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(\mathrm{e}^{-\lambda}+(1-\gamma)\lambda\right) = -\mathrm{e}^{-\lambda}+(1-\gamma). \tag{104}$$

Then, the optimal λ is the solution to $-\mathrm{e}^{-\lambda}+(1-\gamma) = 0$. That is, $\lambda = -\log(1-\gamma)$. This implies:

$$\begin{aligned}
\Pr\left[V_i < (1-\gamma)\mu\right] &\leq \exp\left(\mu(\mathrm{e}^{-\lambda} - 1 + \lambda(1-\gamma))\right) && (105)\\
&\leq \exp\left(-\mu\left(\gamma + (1-\gamma)\left(-\gamma + \frac{\gamma^2}{2}\right)\right)\right) && (106)\\
&= \exp\left(-\mu(\frac{\gamma^2}{2} + \gamma^2(1 - \frac{\gamma^3}{2}))\right) && (107)\\
&\leq \exp\left(-\frac{\mu\gamma^2}{2}\right) && (108)
\end{aligned}$$

Then,

$$\begin{aligned}
&\Pr\left[\sum_{t=1}^{n} \mathbb{1}_{\{Z_t=0\}} < \frac{n(b-b_1)}{b_0-b_1}\right]\\
&= \Pr\left[\sum_{t=1}^{n} \mathbb{1}_{\{Z_t=0\}} < \left(1 - \left(1 - \frac{b-b_1}{\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)}\right)\right)\right.\\
&\left. n\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)\right]\\
&\leq \exp\left(-n\frac{\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)}{2}\left(1 - \frac{b-b_1}{\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)}\right)^2\right)\\
&= \exp\left(-n\frac{\left(\frac{b-b_1}{b_0-b_1} - \left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\right)^2}{2\left(\bar{P}_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)}\right). && (109)
\end{aligned}$$

This completes the proof of (26).

D Proof of Proposition 3

From Proposition 2, it follows that

$$\Pr\left[\sum_{t=1}^{n} \mathbb{1}_{\{Z_t=0\}} > \frac{n(b-b_1)}{b_0-b_1}\right] < \exp\left(-n\frac{\left(\frac{b-b_1}{b_0-b_1} - \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\right)^2}{\left(\frac{b-b_1}{b_0-b_1} + \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\right)}\right), \quad (110)$$

and thus subject to a maximal ESP constraint, it holds that for all $i \in \{1, 2, \ldots, M\}$,

$$\delta > \theta_i > 1 - \exp\left(-n\frac{\left(\frac{b-b_1}{b_0-b_1} - \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\right)^2}{\left(\frac{b-b_1}{b_0-b_1} + \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\right)}\right). \quad (111)$$

Hence, it follows that

$$\exp\left(-n\frac{\left(\frac{b-b_1}{b_0-b_1} - \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\right)^2}{\left(\frac{b-b_1}{b_0-b_1} + \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\right)}\right) > 1-\delta, \quad (112)$$

which implies

$$\left(\frac{b-b_1}{b_0-b_1}\right)^2 - \left(2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) - \frac{1}{n}\log(1-\delta)\right)\left(\frac{b-b_1}{b_0-b_1}\right) + \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) \cdot \left(\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) + \frac{1}{n}\log(1-\delta)\right) < 0. \tag{113}$$

Denote Γ_1 and Γ_2 the roots of the quadratic function in the left-hand side of (113). Then,

$$\Gamma_1 = \frac{1}{2}\left(2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) - \frac{1}{n}\log(1-\delta)\right)$$

$$+\frac{1}{2}\sqrt{-\frac{8}{n}\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\log(1-\delta) + \frac{1}{n^2}\log(1-\delta)^2} \tag{114}$$

$$= \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) - \frac{1}{2n}\log(1-\delta) \tag{115}$$

$$+\sqrt{\frac{-\log(1-\delta)}{n}}\sqrt{2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) - \frac{1}{4n}\log(1-\delta)} \tag{116}$$

$$\Gamma_2 = \frac{1}{2}\left(2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) - \frac{1}{n}\log(1-\delta)\right) \tag{117}$$

$$-\frac{1}{2}\sqrt{-\frac{8}{n}\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\log(1-\delta) + \frac{1}{n^2}\log(1-\delta)^2} \tag{118}$$

$$= \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) - \frac{1}{2n}\log(1-\delta) \tag{119}$$

$$-\sqrt{\frac{-\log(1-\delta)}{n}}\sqrt{2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) - \frac{1}{4n}\log(1-\delta)} \tag{120}$$

From (113), it follows that $\left(\frac{b-b_1}{b_0-b_1}\right)$ must satisfy:

$$\Gamma_2 < \left(\frac{b-b_1}{b_0-b_1}\right) < \Gamma_1. \tag{121}$$

This implies that:

$$\left(\frac{b-b_1}{b_0-b_1}\right) \leq \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) - \frac{1}{2n}\log(1-\delta) \tag{122}$$

$$+\sqrt{\frac{-\log(1-\delta)}{n}}\sqrt{2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right) - \frac{1}{4n}\log(1-\delta)} \tag{123}$$

Thus,

$$\begin{aligned}
b \leq & \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)+b_1-\frac{(b_0-b_1)}{2n}\log(1-\delta) \\
& +\sqrt{\frac{-(b_0-b_1)\log(1-\delta)}{n}}\sqrt{2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)-\frac{(b_0-b_1)}{4n}\log(1-\delta)} \\
\leq & \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)+b_1-\frac{(b_0-b_1)}{2n}\log(1-\delta) \\
& +\sqrt{\frac{2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)^2\log(\frac{1}{1-\delta})}{n}}+\sqrt{\frac{(b_0-b_1)^2\left(\log\left(\frac{1}{1-\delta}\right)\right)^2}{4n^2}} \qquad (124) \\
= & \left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)(b_0-b_1)+b_1-\frac{(b_0-b_1)}{n}\log(1-\delta) \\
& +\frac{(b_0-b_1)}{\sqrt{n}}\sqrt{-2\left(P_X^{(i)}(0)(1-2\alpha_2)+\alpha_2\right)\log(1-\delta)}, \qquad (125)
\end{aligned}$$

and this completes the proof of (27). The proof of (28) follows immediately from (16) and (125).

E Proof of Lemma 1

For all $t \in \{1,2,\ldots,n\}$, consider the first moment, the second moment, and the third absolute moment of the random variable $\mathbb{1}_{\{Z_t=0\}}$ given that the channel input is $\boldsymbol{u}(i)$, with $i \in \{1,2,\ldots,M\}$:

$$\mathbb{E}_{Z|X=u_t(i)}\left[\mathbb{1}_{\{Z_t=0\}}\right]=P_{Z|X}(0|u_t(i)) \qquad (126a)$$

$$\begin{aligned}
\mathbb{V}_{Z|X=u_t(i)}\left[\mathbb{1}_{\{Z_t=0\}}\right]=&\mathbb{E}_{Z|X=u_t(i)}\left[\mathbb{1}^2_{\{Z_t=0\}}\right]-\mathbb{E}_{Z|X=u_t(i)}\left[\mathbb{1}_{\{Z_t=0\}}\right]^2 \\
=&\mathbb{E}_{Z|X=u_t(i)}\left[\mathbb{1}_{\{Z_t=0\}}\right]-\mathbb{E}_{Z|X=u_t(i)}\left[\mathbb{1}_{\{Z_t=0\}}\right]^2 \\
=&P_{Z|X}(0|u_t(i))-P_{Z|X}(0|u_t(i))^2 \\
=&P_{Z|X}(0|u_t(i))\left(1-P_{Z|X}(0|u_t(i))\right) \\
=&P_{Z|X}(0|u_t(i))P_{Z|X}(1|u_t(i)), \text{and} \qquad (126b)
\end{aligned}$$

$$\begin{aligned}
\mathbb{S}_{Z|X=u_t(i)}\left[\mathbb{1}_{\{Z_t=0\}}\right]=&\mathbb{E}_{Z|X=u_t(i)}\left[\left|\left(\mathbb{1}_{\{Z_t=0\}}-P_{Z|X}(0|u_t(i))\right)\right|^3\right] \\
=&P_{Z|X}(0|u_t(i))\left|1-P_{Z|X}(0|u_t(i))\right|^3+P_{Z|X}(1|u_t(i))P_{Z|X}(0|u_t(i))^3 \\
=&P_{Z|X}(0|u_t(i))P_{Z|X}(1|u_t(i))^3+P_{Z|X}(1|u_t(i))P_{Z|X}(0|u_t(i))^3. \qquad (126c)
\end{aligned}$$

Using the equalities above, the following holds,

$$\sum_{t=1}^{n} \mathbb{E}_{Z|X}\left[\mathbb{1}_{\{Z_t=0\}}\right] = nP_{Z|X}(0|0)\bar{P}_X^{(i)}(0) + nP_{Z|X}(0|1)\bar{P}_X^{(i)}(1)$$
$$= n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2\right), \tag{127a}$$

$$\sum_{t=1}^{n} \mathbb{V}_{Z|X=u_t(i)}\left[\mathbb{1}_{\{Z_t=0\}}\right] = \sum_{t=1}^{n} P_{Z|X}(0|u_t(i))P_{Z|X}(1|u_t(i))$$
$$= N(0|\boldsymbol{u}(i))P_{Z|X}(0|0)P_{Z|X}(1|0) + N(1|\boldsymbol{u}(i))P_{Z|X}(0|1)P_{Z|X}(1|1)$$
$$= N(0|\boldsymbol{u}(i))(1-\alpha_2)\alpha_2 + N(1|\boldsymbol{u}(i))\alpha_2(1-\alpha_2)$$
$$= n\left(\alpha_2(1-\alpha_2)\right), \text{and} \tag{127b}$$

$$\sum_{k=1}^{n} \mathbb{S}_{Z|X=u_t(i)}\left[\mathbb{1}_{\{Z_t=0\}}\right] = \sum_{t=1}^{n} P_{Z|X}(0|u_t(i))P_{Z|X}(1|u_t(i))^3 + P_{Z|X}(1|u_t(i))P_{Z|X}(0|u_t(i))^3$$
$$= N(0|\boldsymbol{u}(i))\left(P_{Z|X}(0|0)P_{Z|X}(1|0)^3 + P_{Z|X}(1|0)P_{Z|X}(0|0)^3\right)$$
$$+ N(1|\boldsymbol{u}(i))\left(P_{Z|X}(0|1)P_{Z|X}(1|1)^3 + P_{Z|X}(1|1)P_{Z|X}(0|1)^3\right)$$
$$= N(0|\boldsymbol{u}(i))\left(\alpha_2^3(1-\alpha_2) + \alpha_2(1-\alpha_2)^3\right)$$
$$+ N(1|\boldsymbol{u}(i))\left(\alpha_2^3(1-\alpha_2) + \alpha_2(1-\alpha_2)^3\right)$$
$$= n\left(\alpha_2^3(1-\alpha_2) + \alpha_2(1-\alpha_2)^3\right). \tag{127c}$$

Using (127), it follows that for all $i \in \{1, 2, \ldots, M\}$,

$$\theta_i = \Pr\left[B_n(\boldsymbol{Z}) < b | \boldsymbol{X} = \boldsymbol{u}(i)\right]$$
$$= \Pr\left[\sum_{k=1}^{n} \mathbb{1}\{Z_k = 0\} < \frac{n(b-b_1)}{b_0-b_1} \middle| \boldsymbol{X} = \boldsymbol{u}(i)\right]$$
$$= \Pr\left[\sum_{t=1}^{n} \left(\mathbb{1}\{Z_t = 0\} - P_{Z|X}(0|u_t(i))\right)\right.$$
$$\left. \leq \sqrt{n\alpha_2(1-\alpha_2)}\left(\frac{\frac{n(b-b_1)}{b_0-b_1} - n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right) \middle| \boldsymbol{X} = \boldsymbol{u}(i)\right].$$

From the Berry-Esseen theorem (Theorem 3 in Appendix A), it follows that

$$\theta_i \geq Q\left(\frac{n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right) - \frac{n\left[\alpha_2(1-\alpha_2)^3 + (1-\alpha_2)\alpha_2^3\right]}{2\left(n\alpha_2(1-\alpha_2)\right)^{3/2}}$$
$$= Q\left(\frac{n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right) - \frac{\left(\alpha_2(1-\alpha_2)\right)\left((1-\alpha_2)^2 + \alpha_2^2\right)}{2\sqrt{n}\left(\alpha_2(1-\alpha_2)\right)^{3/2}} \tag{128}$$
$$= Q\left(\frac{n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right) - \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}, \tag{129}$$

and

$$\theta_i \leq Q\left(\frac{n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0) + \alpha_2 - \frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right) + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}.$$

This completes the proof.

F Proof of Proposition 4

Using the fact that $Q^{-1}(\cdot)$ is a decreasing function, it follows from Lemma 1 that

$$\frac{n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0)+\alpha_2)-\frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}} \geqslant Q^{-1}\left(\delta+\frac{(1-\alpha_2)^2+\alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}\right). \tag{130}$$

From (130), it holds that

$$\frac{n(b-b_1)}{b_0-b_1} \leqslant n\left((1-2\alpha_2)\bar{P}_X^{(i)}(0)+\alpha_2\right)-\sqrt{n\alpha_2(1-\alpha_2)}Q^{-1}\left(\delta+\frac{(1-\alpha_2)^2+\alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}\right). \tag{131}$$

Finally, the energy rate b is upper bounded as the following

$$b \leq (b_0-b_1)\left((1-2\alpha_2)\bar{P}_X^{(i)}(0)+\alpha_2\right)+b_1-\sqrt{\frac{(b_0-b_1)^2\alpha_2(1-\alpha_2)}{n}}Q^{-1}\left(\delta+\frac{(1-\alpha_2)^2+\alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}\right),$$

which ends the proof of the energy bound (27). The proof of (28) follows immediately by following the definition of θ in (16) and using (129). This completes the proof. ■

G Proof of Theorem 1

The proof of Theorem 1 is based on the random coding arguments.
Codebook Generation: Let $\rho \in [0,1]$ be a fixed parameter. Consider a probability distribution P_X that satisfies

$$P_X(0) = 1 - P_X(1) = \rho. \tag{132}$$

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 $\boldsymbol{x} \in \mathcal{X}^n$,

$$P_{\boldsymbol{X}}(\boldsymbol{x}) = \rho^{N(0|\boldsymbol{x})}(1-\rho)^{n-N(0|\boldsymbol{x})}. \tag{133}$$

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

$$\imath(\boldsymbol{x}, \boldsymbol{y}) = \log\left(\frac{P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})}{P_{\boldsymbol{Y}}(\boldsymbol{y})}\right). \tag{134}$$

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

$$\mathcal{D}_\ell \triangleq \left\{\boldsymbol{y} \in \mathcal{Y}^n : \ell \in \arg\max_{k\in\{1,2,\ldots,M\}} \imath\left(\boldsymbol{u}(k), \boldsymbol{y}\right)\right\} \setminus \bigcup_{j=1}^{\ell-1} \mathcal{D}_j, \text{ and} \tag{135a}$$

$$\mathcal{D}_M \triangleq \mathcal{Y}^n \setminus \bigcup_{j=1}^{M-1} \mathcal{D}_j. \tag{135b}$$

Given a generated (n, M)-code, the transmitter inputs the symbol $u_t(i)$ at channel use t, with $t \in \{1, 2, \ldots, n\}$ to transmit the message index $i \in \{1, 2, \ldots, M\}$. After n channel uses, the IR observes an n-dimensional channel output vector $\boldsymbol{y}$. The IR decides upon the index i following the rule in (9).

Decoding Error Probability Analysis: Let $\bar{\lambda}$ be the average over all possible codebooks of the DEP λ in (11). An immediate consequence of the random coding union bound [10] is that the average DEP $\bar{\lambda}$ is upper bounded by:

$$\begin{aligned}\bar{\lambda} &< \mathbb{E}_{\boldsymbol{XY}}\left[\min\left\{1, (M-1)\mathrm{Pr}\left[\iota(\bar{\boldsymbol{X}};\boldsymbol{Y}) \geqslant \iota(\boldsymbol{X};\boldsymbol{Y})\right]\right\}\right] \\ &= \sum_{\boldsymbol{x}\in\mathcal{X}^n}\sum_{\boldsymbol{y}\in\mathcal{Y}^n} P_{\boldsymbol{X}}(\boldsymbol{x})P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})\min\left\{1, (M-1)\mathrm{Pr}\left[\iota(\bar{\boldsymbol{X}};\boldsymbol{y}) \geqslant \iota(\boldsymbol{x};\boldsymbol{y})\right]\right\}, \end{aligned} \tag{136}$$

where the probability in (136) is with respect to the random variable $\bar{\boldsymbol{X}}$, whose probability mass function is $P_{\boldsymbol{X}}$ in (133). Note that for all $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$, the following holds

$$\begin{aligned}\mathbb{1}_{\{\iota(\bar{\boldsymbol{x}};\boldsymbol{y})\geq\iota(\boldsymbol{x};\boldsymbol{y})\}} &= \mathbb{1}_{\left\{\log\frac{P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\bar{\boldsymbol{x}})}{P_{\boldsymbol{Y}}(\boldsymbol{y})}\geq\log\frac{P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})}{P_{\boldsymbol{Y}}(\boldsymbol{y})}\right\}} \\ &= \mathbb{1}_{\{d(\bar{\boldsymbol{x}},\boldsymbol{y})\log\alpha_1+(n-d(\bar{\boldsymbol{x}},\boldsymbol{y}))\log(1-\alpha_1)\geq d(\boldsymbol{x},\boldsymbol{y})\log\alpha_1+(n-d(\boldsymbol{x},\boldsymbol{y}))\log(1-\alpha_1)\}} \\ &= \mathbb{1}_{\{(d(\bar{\boldsymbol{x}},\boldsymbol{y})-d(\boldsymbol{x},\boldsymbol{y}))\log\alpha_1-(d(\bar{\boldsymbol{x}},\boldsymbol{y})-d(\boldsymbol{x},\boldsymbol{y}))\log(1-\alpha_1)\}} \\ &= \mathbb{1}_{\left\{(d(\bar{\boldsymbol{x}},\boldsymbol{y})-d(\boldsymbol{x},\boldsymbol{y}))\log\left(\frac{\alpha_1}{1-\alpha_1}\right)\geq 0\right\}} \qquad (137) \\ &= \mathbb{1}_{\{d(\bar{\boldsymbol{x}},\boldsymbol{y})\leq d(\boldsymbol{x},\boldsymbol{y})\}}, \qquad (138)\end{aligned}$$

where (137) follows from the fact that $\log\left(\frac{\alpha_1}{1-\alpha_1}\right) < 0$ for all $\alpha_1 \in \left[0, \frac{1}{2}\right)$. Now, from (138) it holds that for all $(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$:

$$\begin{aligned}\mathrm{Pr}\left[\iota(\bar{\boldsymbol{X}};\boldsymbol{y}) \geq \iota(\boldsymbol{x};\boldsymbol{y})\right] &= \sum_{\bar{\boldsymbol{x}}\in\mathcal{X}^n}\mathbb{1}_{\{d(\bar{\boldsymbol{x}},\boldsymbol{y})\leq d(\boldsymbol{x},\boldsymbol{y})\}}P_{\boldsymbol{X}}(\bar{\boldsymbol{x}}) \qquad (139) \\ &= \mathrm{Pr}\left[d\left(\bar{\boldsymbol{X}},\boldsymbol{y}\right) \leq d(\boldsymbol{x},\boldsymbol{y})\right] \\ &= \sum_{\ell=0}^{d(\boldsymbol{x},\boldsymbol{y})}\mathrm{Pr}\left[d\left(\bar{\boldsymbol{X}},\boldsymbol{y}\right) = \ell\right], \qquad (140)\end{aligned}$$

where,

$$\begin{aligned}\mathrm{Pr}\left[d\left(\bar{\boldsymbol{X}},\boldsymbol{y}\right) = \ell)\right] &= \mathrm{Pr}\left[\sum_{t=0}^{n}\mathbb{1}_{\{\bar{X}_t\neq y_t\}} = \ell\right] \\ &= \mathrm{Pr}\left[\sum_{t\in\{m:y_m=0\}}\mathbb{1}_{\{\bar{X}_t=1\}} + \sum_{t\in\{m:y_m=1\}}\mathbb{1}_{\{\bar{X}_t=0\}} = \ell\right], \end{aligned} \tag{141}$$

with $\ell \in \{0,1,\ldots,n\}$. Note that for all $\boldsymbol{y} \in \{0,1\}^n$,

$$\sum_{t\in\{m:y_m=0\}}\mathbb{1}_{\{\bar{X}_t=1\}} \sim \mathrm{Binomial}(N(0|\boldsymbol{y}), 1-\rho), \quad \text{and} \tag{142}$$

$$\sum_{t\in\{m:y_m=1\}}\mathbb{1}_{\{\bar{X}_t=0\}} \sim \mathrm{Binomial}(n-N(0|\boldsymbol{y}), \rho). \tag{143}$$

Therefore, it holds that

$$\begin{aligned}\Pr\left[d\left(\bar{\boldsymbol{X}},\boldsymbol{y}\right)=\ell\right]=&\sum_{s=0}^{\ell}\Pr\left[\sum_{t\in\{m:y_m=0\}}\mathbb{1}_{\left\{\bar{X}_t=1\right\}}=s\right]\Pr\left[\sum_{t\in\{m:y_m=1\}}\mathbb{1}_{\left\{\bar{X}_t=0\right\}}=\ell-s\right]\\ =&\sum_{s=0}^{\ell}\binom{N(0|\boldsymbol{y})}{s}(1-\rho)^{s}\rho^{N(0|\boldsymbol{y})-s}\binom{n-N(0|\boldsymbol{y})}{\ell-s}\rho^{\ell-s}(1-\rho)^{n-N(0|\boldsymbol{y})-\ell+s}\\ =&\sum_{s=0}^{\ell}\binom{N(0|\boldsymbol{y})}{s}\binom{n-N(0|\boldsymbol{y})}{\ell-s}\rho^{\ell+N(0|\boldsymbol{y})-2s}(1-\rho)^{n-(\ell+N(0|\boldsymbol{y})-2s)}.\end{aligned}\tag{144}$$

Plugging (144) in (140) yields

$$\begin{aligned}\Pr\left[\iota(\bar{\boldsymbol{X}};\boldsymbol{y})\geq\iota(\boldsymbol{x};\boldsymbol{y})\right]=&\sum_{\ell_1=0}^{d(\boldsymbol{x},\boldsymbol{y})}\Pr\left[d\left(\bar{\boldsymbol{X}},\boldsymbol{y}\right)=\ell_1\right]\\ =&\sum_{\ell_1=0}^{d(\boldsymbol{x},\boldsymbol{y})}\sum_{\ell_2=0}^{\ell_1}\binom{N(0|\boldsymbol{y})}{\ell_2}\binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2}\rho^{\ell_1+N(0|\boldsymbol{y})-2\ell_2}(1-\rho)^{n-(\ell_1+N(0|\boldsymbol{y})-2\ell_2)}.\end{aligned}\tag{145}$$

The term $\mathbb{E}_{\boldsymbol{X}\boldsymbol{Y}}\left[\Pr\left[\iota(\bar{\boldsymbol{X}};\boldsymbol{Y})\geq\iota(\boldsymbol{X};\boldsymbol{Y})\right]\right]$ in (136) can be calculated as follows:

$$\mathbb{E}_{\boldsymbol{X}\boldsymbol{Y}}\left[\Pr\left[\iota(\bar{\boldsymbol{X}};\boldsymbol{Y})\geq\iota(\boldsymbol{X};\boldsymbol{Y})\right]\right]\tag{146}$$

$$\begin{aligned}&=\sum_{\boldsymbol{y}\in\mathcal{Y}^n}\sum_{\boldsymbol{x}\in\mathcal{X}^n}P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})P_{\boldsymbol{X}}(\boldsymbol{x})\Pr\left[\iota(\bar{\boldsymbol{X}};\boldsymbol{y})\geq\iota(\boldsymbol{x};\boldsymbol{y})\right]\\ &=\sum_{\boldsymbol{y}\in\mathcal{Y}^n}\sum_{\boldsymbol{x}\in\mathcal{X}^n}(1-\alpha_1)^{d(\boldsymbol{x},\boldsymbol{y})}\alpha_1^{n-d(\boldsymbol{x},\boldsymbol{y})}\rho^{N(0|\boldsymbol{x})}(1-\rho)^{n-N(0|\boldsymbol{x})}\Pr\left[d\left(\bar{\boldsymbol{X}},\boldsymbol{y}\right)\leq d(\boldsymbol{x},\boldsymbol{y})\right].\end{aligned}$$

There are 2^n vectors in $\mathcal{X}^n$ and such a number can be written as follows for all $\boldsymbol{y}\in\mathcal{Y}^n$:

$$2^n=\sum_{\ell=0}^{n}\binom{n}{\ell}=\sum_{\ell_1=0}^{n}\sum_{\ell_2=0}^{\ell_1}\binom{N(0|\boldsymbol{y})}{\ell_2}\binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2},\tag{147}$$

where the second equality follows from the Vandermonde's inequality. Hence, given an $\ell\leqslant n$, $\binom{n}{\ell}$ can be interpreted as the number of vectors in $\mathcal{X}^n$ with ℓ zeros. Then, $\sum_{\ell=0}^{n}\binom{n}{\ell}$ is the number of all vectors in $\mathcal{X}^n$. To continue with this analysis, given a vector $\boldsymbol{y}\in\mathcal{Y}^n$ and a vector $\boldsymbol{x}\in\mathcal{X}^n$, let ℓ_1 be the number of zeros in $\boldsymbol{x}$; and let also ℓ_2 be the number of zeros in $\boldsymbol{x}$ that are in the same components in which the vector $\boldsymbol{y}$ also has zeros. Hence, the number $\binom{N(0|\boldsymbol{y})}{\ell_2}\binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2}$ can be interpreted as the number of vectors $\boldsymbol{x}$ that contain exactly ℓ_1 zeros among which ℓ_2 zeros are placed on components in which $\boldsymbol{y}$ also contains zeros. Therefore, $\sum_{\ell_2=0}^{\ell_1}\binom{N(0|\boldsymbol{y})}{\ell_2}\binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2}=\binom{n}{\ell_1}$ is the number of vectors in $\mathcal{X}^n$ with exactly ℓ_1 zeros. Using these interpretations, it holds that

$$N(0|\boldsymbol{x})=\ell_1\text{ and}\tag{148}$$

$$d(\boldsymbol{x},\boldsymbol{y})=N(0|\boldsymbol{y})-\ell_2+\ell_1-\ell_2\tag{149}$$

$$=N(0|\boldsymbol{y})+\ell_1-2\ell_2,\tag{150}$$

and

$$\mathbb{E}_{\boldsymbol{X}\boldsymbol{Y}}\left[\Pr\left[\iota(\bar{\boldsymbol{X}};\boldsymbol{Y}) \geq \iota(\boldsymbol{X};\boldsymbol{Y})\right]\right] \tag{151}$$

$$= \sum_{\boldsymbol{y}\in\mathcal{Y}^n}\sum_{\ell_1=0}^{n}\sum_{\ell_2=0}^{\min\{N(0|\boldsymbol{y}),\ell_1\}}\binom{N(0|\boldsymbol{y})}{\ell_2}\binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2}\alpha_1^{N(0|\boldsymbol{y})+\ell_1-2\ell_2}(1-\alpha_1)^{n-N(0|\boldsymbol{y})-\ell_1+2\ell_2}$$
$$\cdot\rho^{\ell_1}(1-\rho)^{n-\ell_1}\Pr\left[d(\bar{\boldsymbol{X}},\boldsymbol{y}) \leq N(0|\boldsymbol{y})+\ell_1-2\ell_2\right]$$
$$= \sum_{\boldsymbol{y}\in\mathcal{Y}^n}\sum_{\ell_1=0}^{n}\sum_{\ell_2=0}^{\min\{N(0|\boldsymbol{y}),\ell_1\}}\binom{N(0|\boldsymbol{y})}{\ell_2}\binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2}\alpha_1^{N(0|\boldsymbol{y})+\ell_1-2\ell_2}(1-\alpha_1)^{n-N(0|\boldsymbol{y})-\ell_1+2\ell_2}$$
$$\cdot\rho^{\ell_1}(1-\rho)^{n-\ell_1}\sum_{\ell_3=0}^{N(0|\boldsymbol{y})+\ell_1-2\ell_2}\sum_{\ell_4=0}^{\ell_3}\binom{N(0|\boldsymbol{y})}{\ell_4}\binom{n-N(0|\boldsymbol{y})}{\ell_3-\ell_4}\rho^{\ell_3+N(0|\boldsymbol{y})-2\ell_4}(1-\rho)^{n-(\ell_3+N(0|\boldsymbol{y})-2\ell_4)}$$
$$= \sum_{\ell_0=0}^{n}\sum_{\ell_1=0}^{n}\sum_{\ell_2=0}^{\min\{\ell_0,\ell_1\}}\binom{n}{\ell_0}\binom{\ell_0}{\ell_2}\binom{n-\ell_0}{\ell_1-\ell_2}\alpha_1^{\ell_0+\ell_1-2\ell_2}(1-\alpha_1)^{n-\ell_0-\ell_1+2\ell_2}\rho^{\ell_1}(1-\rho)^{n-\ell_1}$$
$$\cdot\sum_{\ell_3=0}^{\ell_0+\ell_1-2\ell_2}\sum_{\ell_4=0}^{\ell_3}\binom{\ell_0}{\ell_4}\binom{n-\ell_0}{\ell_3-\ell_4}\rho^{\ell_3+\ell_0-2\ell_4}(1-\rho)^{n-(\ell_3+\ell_0-2\ell_4)} \tag{152}$$
$$= \sum_{\ell_0=0}^{n}\sum_{\ell_1=0}^{n}\sum_{\ell_2=0}^{\min\{\ell_0,\ell_1\}}\sum_{\ell_3=0}^{\ell_0+\ell_1-2\ell_2}\sum_{\ell_4=0}^{\ell_3}\binom{n}{\ell_0}\binom{\ell_0}{\ell_2}\binom{n-\ell_0}{\ell_1-\ell_2}\binom{\ell_0}{\ell_4}\binom{n-\ell_0}{\ell_3-\ell_4}$$
$$\cdot\alpha_1^{\ell_0+\ell_1-2\ell_2}(1-\alpha_1)^{n-\ell_0-\ell_1+2\ell_2}\rho^{\ell_3+\ell_0-2\ell_4+\ell_1}(1-\rho)^{2n-(\ell_3+\ell_0-2\ell_4+\ell_1)}. \tag{153}$$

Replacing (153) into (136), leads to $\bar{\lambda} < \phi(m,\rho)$. This 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 (7) generated using the probability mass function in (132). Hence, at an energy transmission rate b, it follows from Lemma 1 that for all $i \in \{1, 2, \ldots, M\}$,

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

Let $\bar{\theta}_i$ be the average over all possible codebooks of the energy-shortage probability θ_i in (15) while transmitting at an energy rate b. Hence, the following holds:

$$\bar{\theta}_i = \sum_{\boldsymbol{x}\in\mathcal{X}^n} P_{\boldsymbol{X}}(\boldsymbol{x})\Pr\left[B_n(\boldsymbol{Z}) < b \mid \boldsymbol{X} = \boldsymbol{x}\right]$$
$$\leqslant \sum_{\boldsymbol{x}\in\mathcal{X}^n} P_{\boldsymbol{X}}(\boldsymbol{x}) Q\left(\frac{(1-2\alpha_2)N(0|\boldsymbol{x}) + n\left(\alpha_2 - \frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right) + \frac{(1-\alpha_2)^2+\alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}$$
$$= \sum_{t=0}^{n}\binom{n}{t}\rho^t(1-\rho)^{n-t} Q\left(\frac{(1-2\alpha_2)t + n\left(\alpha_2 - \frac{b-b_1}{b_0-b_1}\right)}{\sqrt{n\alpha_2(1-\alpha_2)}}\right) + \frac{(1-\alpha_2)^2+\alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}}$$
$$= \chi(b,\rho). \tag{154}$$

Hence from (154), it follows that for all $i \in \{1, 2, \ldots, M\}$, $\bar{\theta} < \chi(b, \rho)$. This completes the proof. ∎

H Proof of Theorem 2

Consider an $(n, M, \epsilon, \delta, b)$-code with maximum ESP described by the system in (7) and empirical input distribution

$$\bar{P}_X(0) = 1 - \bar{P}_X(1) \stackrel{\triangle}{=} \rho. \tag{155}$$

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

$$\mathcal{H}_0 : (\boldsymbol{X}, \boldsymbol{Y}) \sim \bar{P}_{\boldsymbol{X}} Q_{\boldsymbol{Y}} \text{ and} \tag{156a}$$
$$\mathcal{H}_1 : (\boldsymbol{X}, \boldsymbol{Y}) \sim \bar{P}_{\boldsymbol{X}} P_{\boldsymbol{Y}|\boldsymbol{X}}, \tag{156b}$$

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

$$Q_{\boldsymbol{Y}}(\boldsymbol{y}) = q^{N(0,\boldsymbol{y})} (1-q)^{n-N(0,\boldsymbol{y})}, \tag{157}$$
$$\bar{P}_{\boldsymbol{X}}(\boldsymbol{x}) = \rho^{N(0|\boldsymbol{x})}(1-\rho)^{n-N(0|\boldsymbol{x})}, \text{ and} \tag{158}$$
$$P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) = \alpha_1^{d(\boldsymbol{x},\boldsymbol{y})}(1-\alpha_1)^{n-d(\boldsymbol{x},\boldsymbol{y})}. \tag{159}$$

The goal of the binary hypothesis test in (156) 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 $P_{T|\boldsymbol{XY}}$ 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 (156). 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

$$\beta_{1-\epsilon}(P_{\boldsymbol{XY}}, \bar{P}_{\boldsymbol{X}} Q_{\boldsymbol{Y}}) = \inf_{P_{T|\boldsymbol{XY}}: \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \sum_{\boldsymbol{x}\in\mathcal{X}^n} P_{T|\boldsymbol{XY}}(1|\boldsymbol{x},\boldsymbol{y}) P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) \bar{P}_{\boldsymbol{X}}(\boldsymbol{x}) \geq 1-\epsilon} \left[\sum_{\boldsymbol{y}\in\mathcal{Y}^n} \sum_{\boldsymbol{x}\in\mathcal{X}^n} P_{T|\boldsymbol{XY}}(1|\boldsymbol{x},\boldsymbol{y}) \bar{P}_{\boldsymbol{X}}(\boldsymbol{x}) Q_{\boldsymbol{Y}}(\boldsymbol{y}) \right], \tag{160}$$

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 $1 - \epsilon$. Note that the corresponding log-likelihood ratio for these hypotheses is for all $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{X}^n \times \mathcal{Y}^n$,

$$\begin{aligned}
\log \frac{P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x})}{Q_{\boldsymbol{Y}}(\boldsymbol{y})} =& \log \frac{(1-\alpha_1)^{n-d(\boldsymbol{x},\boldsymbol{y})} \alpha_1^{d(\boldsymbol{x},\boldsymbol{y})}}{q^{N(0,\boldsymbol{y})} \left(1-q\right)^{n-N(0,\boldsymbol{y})}} \\
=& (n - d(\boldsymbol{x},\boldsymbol{y})) \log(1-\alpha_1) + d(\boldsymbol{x},\boldsymbol{y}) \log \alpha_1 - N(0|\boldsymbol{y}) \log\left(q\right) - (n - N(0|\boldsymbol{y})) \log\left(1-q\right) \\
=& n \log\left(\frac{1-\alpha_1}{1-q}\right) + d(\boldsymbol{x},\boldsymbol{y}) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q}.
\end{aligned}$$

Hence, from the Neyman-Pearson lemma [16], it follows that the optimal test is of the form

$$P^*_{T|\boldsymbol{XY}}(1|\boldsymbol{x},\boldsymbol{y}) = \begin{cases} 0 & \text{if} \quad d(\boldsymbol{x},\boldsymbol{y}) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} < L \\ \lambda & \text{if} \quad d(\boldsymbol{x},\boldsymbol{y}) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} = L \\ 1 & \text{if} \quad d(\boldsymbol{x},\boldsymbol{y}) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} > L, \end{cases} \tag{161}$$

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

$$\begin{aligned}
1-\epsilon =& \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \sum_{\boldsymbol{x}\ \in\mathcal{X}^n} P^*_{T|\boldsymbol{X}\boldsymbol{Y}}(1|\boldsymbol{x},\boldsymbol{y}) P_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) \bar{P}_{\boldsymbol{X}}(\boldsymbol{x}) \\
=& \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \sum_{\boldsymbol{x}\ \in\mathcal{X}^n} (1-\alpha_1)^{n-d(\boldsymbol{x},\boldsymbol{y})} \alpha^{d(\boldsymbol{x},\boldsymbol{y})} \rho^{N(0|\boldsymbol{x})} (1-\rho)^{N(0|\boldsymbol{x})} \\
&\cdot \left(\mathbb{1}_{\left\{ d(\boldsymbol{x},\boldsymbol{y}) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} > L \right\}} + \lambda \mathbb{1}_{\left\{ d(\boldsymbol{x},\boldsymbol{y}) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} = L \right\}} \right). \quad (162)
\end{aligned}$$

The equality in (162) can be rewritten by noting that there are 2^n vectors in $\mathcal{X}^n$ and for all $\boldsymbol{y} \in \mathcal{Y}^n$:

$$2^n = \sum_{\ell=0}^{n} \binom{n}{\ell} = \sum_{\ell_1=0}^{n} \sum_{\ell_2=0}^{\ell_1} \binom{N(0|\boldsymbol{y})}{\ell_2} \binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2}, \quad (163)$$

where the second equality follows from the Vandermonde's inequality. Given a vector $\boldsymbol{y} \in \mathcal{Y}^n$ and a vector $\boldsymbol{x} \in \mathcal{X}^n$, let ℓ_1 be the number of zeros in $\boldsymbol{x}$; and let also ℓ_2 be the number of zeros in $\boldsymbol{x}$ that are in the same components in which the vector $\boldsymbol{y}$ also has zeros. Hence, the number $\binom{N(0|\boldsymbol{y})}{\ell_2}\binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2}$ can be interpreted as the number of vectors $\boldsymbol{x}$ that contain exactly ℓ_1 zeros among which ℓ_2 zeros are placed on components in which $\boldsymbol{y}$ also contains zeros. Therefore, $\sum_{\ell_2=0}^{\ell_1} \binom{N(0|\boldsymbol{y})}{\ell_2}\binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2} = \binom{n}{\ell_1}$ is the number of vectors in $\mathcal{X}^n$ with exactly ℓ_1 zeros. Using these interpretations, it holds that

$$\begin{aligned}
N(0|\boldsymbol{x}) &= \ell_1 \text{ and} & (164)\\
d(\boldsymbol{x},\boldsymbol{y}) &= N(0|\boldsymbol{y}) - \ell_2 + \ell_1 - \ell_2 & (165)\\
&= N(0|\boldsymbol{y}) + \ell_1 - 2\ell_2, & (166)
\end{aligned}$$

and thus, the equality in (162) can be rewritten as follows:

$$\begin{aligned}
1-\epsilon =& \sum_{\boldsymbol{y}\in\mathcal{Y}^n} \sum_{\ell_1=0}^{n} \sum_{\ell_2=0}^{\ell_1} \binom{N(0|\boldsymbol{y})}{\ell_2} \binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2} (1-\alpha_1)^{n-(N(0|\boldsymbol{y})+\ell_1-2\ell_2)} \alpha_1^{N(0|\boldsymbol{y})+\ell_1-2\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1} \\
&\cdot \left(\mathbb{1}_{\left\{ (N(0|\boldsymbol{y})+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} > L \right\}} + \lambda \mathbb{1}_{\left\{ (N(0|\boldsymbol{y})+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} = L \right\}} \right) \\
=& \sum_{\ell_0=0}^{n} \sum_{\ell_1=0}^{n} \sum_{\ell_2=0}^{\ell_1} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} (1-\alpha_1)^{n-(\ell_0+\ell_1-2\ell_2)} \alpha_1^{\ell_0+\ell_1-2\ell_2} \rho^{\ell_1} (1-\rho)^{n-\ell_1} \\
&\cdot \left(\mathbb{1}_{\left\{ (\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} > L \right\}} + \lambda \mathbb{1}_{\left\{ (\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} = L \right\}} \right).
\end{aligned}$$

Plugging (161) into (160) yields

$$
\begin{aligned}
\beta_{1-\epsilon}(P_{\boldsymbol{XY}}, \bar{P}_{\boldsymbol{X}} Q_{\boldsymbol{Y}}) =& \sum_{\boldsymbol{y} \in \mathcal{Y}^n} \sum_{\boldsymbol{x} \in \mathcal{X}^n} P^*_{T|\boldsymbol{XY}}(1|\boldsymbol{x},\boldsymbol{y}) Q_{\boldsymbol{Y}}(\boldsymbol{y}) \bar{P}_{\boldsymbol{X}}(\boldsymbol{x}) \\
=& \sum_{\boldsymbol{y} \in \mathcal{Y}^n} \sum_{\boldsymbol{x} \in \mathcal{X}^n} \rho^{N(0|\boldsymbol{x})}(1-\rho)^{n-N(0|\boldsymbol{x})} q^{N(0|\boldsymbol{y})} (1-q)^{n-N(0|\boldsymbol{y})} \\
& \left(\mathbb{1}_{\left\{ d(\boldsymbol{x},\boldsymbol{y}) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} > L \right\}} + \lambda \mathbb{1}_{\left\{ d(\boldsymbol{x},\boldsymbol{y}) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} = L \right\}} \right) \\
=& \sum_{\boldsymbol{y} \in \mathcal{Y}^n} \sum_{\ell_1=0}^{n} \sum_{\ell_2=0}^{\ell_1} \binom{N(0|\boldsymbol{y})}{\ell_2} \binom{n-N(0|\boldsymbol{y})}{\ell_1-\ell_2} \rho^{\ell_1}(1-\rho)^{n-\ell_1} q^{N(0|\boldsymbol{y})} (1-q)^{n-N(0|\boldsymbol{y})} \\
& \Big(\mathbb{1}_{\left\{ (N(0|\boldsymbol{y})+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} > L \right\}} \\
& + \lambda \mathbb{1}_{\left\{ (N(0|\boldsymbol{y})+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + N(0|\boldsymbol{y}) \log \frac{1-q}{q} = L \right\}} \Big) \\
=& \sum_{\ell_0=0}^{n} \sum_{\ell_1=0}^{n} \sum_{\ell_2=0}^{\ell_1} \binom{n}{\ell_0} \binom{\ell_0}{\ell_2} \binom{n-\ell_0}{\ell_1-\ell_2} \rho^{\ell_1}(1-\rho)^{n-\ell_1} q^{\ell_0} (1-q)^{n-\ell_0} \\
& \left(\mathbb{1}_{\left\{ (\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} > L \right\}} + \lambda \mathbb{1}_{\left\{ (\ell_0+\ell_1-2\ell_2) \log \frac{\alpha_1}{1-\alpha_1} + \ell_0 \log \frac{1-q}{q} = L \right\}} \right) \\
\triangleq& \gamma(\rho, q).
\end{aligned}
$$

Finally, from Theorem 29 in [10], it follows that

$$
M < \Gamma(b), \tag{167}
$$

where the function Γ is defined in (51) and the optimization domain over ρ is due to Corollary 4, which requires that $\rho > \rho^*(b)$. This completes the proof of the information bound.

The proof continues with the proof of the bound on the energy rate. From Proposition 4, subject to a maximum ESP constraint, it follows that for any $(n, M, \epsilon, \delta, b)$-code described by the system in (7) for the random transformation in (3) satisfying (14), it holds that for all $i \in \{1, 2, \ldots, M\}$:

$$
b \leqslant (b_0 - b_1)\Big((1-2\alpha_2)\, \bar{P}_X^{(i)}(0) + \alpha_2 \Big) + b_1 - \sqrt{\frac{(b_0-b_1)^2 \alpha_2 (1-\alpha_2)}{n}} Q^{-1}\left(\delta + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}} \right), \tag{168}
$$

$$
= (1-\alpha_2) b_0 + \alpha_2 b_1 - \sqrt{\frac{(b_0-b_1)^2 \alpha_2 (1-\alpha_2)}{n}} \cdot Q^{-1}\left(\delta + \frac{(1-\alpha_2)^2 + \alpha_2^2}{2\sqrt{n\alpha_2(1-\alpha_2)}} \right) \tag{169}
$$

$$
= B^+. \tag{170}
$$

From (167) and (170), it follows that the information rate $\frac{\log M}{n}$ and the energy rate b of any $(n, M, \epsilon, \delta, b)$-code for the random transformation in (3) satisfying (14) subject to a maximum ESP constraint are such that, $M < \Gamma(b)$ and $b < B^+$. 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] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, 2nd ed. New York: Academic Press, 1981.

[13] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. New York: Springer, 1998.

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

[15] I. Shevtsova, "On the absolute constants in the Berry-Esseen type inequalities for identically distributed summands," *ArXiv e-prints*, Nov. 2011.

[16] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. New York: Springer, 1994.

RESEARCH CENTRE
GRENOBLE – RHÔNE-ALPES

Inovallée
655 avenue de l'Europe Montbonnot
38334 Saint Ismier Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399