Information-Energy Regions in the Finite Block-Length Regime with Finite Channel Inputs

Sadaf Ul Zuhra, Samir M Perlaza, H. Vincent Poor, Eitan Altman, Mikael Skoglund

Information-Energy Regions in the Finite Block-Length Regime with Finite Channel Inputs

Sadaf ul Zuhra, *Member, IEEE,* Samir M. Perlaza, *Senior Member, IEEE,* H. Vincent Poor, *Fellow, IEEE,* Eitan Altman, *Fellow, IEEE,* Mikael Skoglund, *Fellow, IEEE*

Abstract

This paper provides a complete characterization of the information-energy region of simultaneous information and energy transmission over an additive white Gaussian noise channel in the finite block-length regime with finite sets of channel input symbols. Given a set of channel input symbols, the converse characterizes the tuples of information rate, energy rate, decoding error probability (DEP) and energy outage probability (EOP) that cannot be achieved by any code built using the given set of channel inputs. A novel method for constructing a family of codes that respects the given information rate, energy rate, DEP and EOP requirements is proposed. The achievable region identifies the set of tuples of information rate, energy rate, DEP and EOP that can be achieved by the constructed family of codes. The proposed construction proves to be information rate, energy rate, and EOP optimal. The achieved DEP is, however, sub-optimal, owing to the choice of the decoding regions made during the construction.

Index Terms

Simultaneous information and energy transmission, SIET, SWIPT, information-energy regions, achievability, converse, finite block-length, finite channel inputs

1. Introduction

Simultaneous information and energy transmission (SIET) (also known as simultaneous wireless information and power transfer (SWIPT)) employs radio frequency (RF) signals to simultaneously accomplish the tasks of conveying information and providing energy to (possibly different) devices. In the following, for the sake of correctness, the denomination "energy transmission" is preferred against "power transfer", and thus, the acronym SIET is adopted in the remainder of this paper. A fundamental question of interest in SIET is that, given a certain code, what can and cannot be accomplished by such a code in terms of the information and energy transmission rates. It is particularly interesting to answer this question in the finite block-length regime with finite sets of channel input symbols. The trade-offs between information and energy transmission rates in SIET have previously been studied in the asymptotic regime [1]–[6] where the assumption of infinitely long transmissions guarantees that the decoding error probability (DEP) and the energy outage probability (EOP) can be made arbitrarily close to zero. Thus, the focus of the asymptotic results is only on the information and energy transmission rates. In the finite block-length regime, however, the DEP and EOP are bounded away from zero and pose additional constraints on the fundamental limits of SIET.

A common assumption in the study of SIET over an additive white Gaussian noise (AWGN) channel is that the channel inputs are derived independently from a Gaussian distribution. In a departure from this norm, this work considers the more meaningful, discrete set of channel input symbols that define the channel. This work characterizes the fundamental limits of SIET for such channels in the finite block-length regime in the presence of AWGN noise. Unlike Gaussian inputs, the fundamental limits obtained for finite discrete sets of channel input symbols provide insights for practical systems that operate with finite block-lengths and finite sets of channel inputs.

Sadaf ul Zuhra, Samir M. Perlaza and Eitan Altman are with INRIA, Centre de Recherche de Sophia Antipolis - Méditerranée, 2004 Route des Lucioles, 06902 Sophia Antipolis, France. {sadaf-ul.zuhra, samir.perlaza, eitan.altman}@inria.fr

H. Vincent Poor, Sadaf ul Zuhra, and Samir M. Perlaza are with the Department of Electrical and Computer Engineering, Princeton University, 08540 Princeton, NJ, USA. (poor@princeton.edu)

Mikael Skoglund is with the School of Electrical Engineering and Computer Science, Malvinas Väg 10, KTH Royal Institute of Technology, 11428 Stockholm, Sweden. (skoglund@kth.se)

Samir M. Perlaza is also with the Laboratoire de Mathématiques GAATI, Université de la Polynésie Française, BP 6570, 98702 Faaa, French Polynesia.

Eitan Altman is also with the Laboratoire d'Informatique d'Avignon, Université d'Avignon, 84911 Avignon, France; and with the Laboratory of Information, Network and Communication Sciences, 75013 Paris, France.

This research was supported in part by the European Commission through the H2020-MSCA-RISE-2019 program under grant 872172; in part by the Agence Nationale de la Recherche (ANR) through the project MAESTRO-5G (ANR-18-CE25-0012); in part by the U.S. National Science Foundation under Grant CCF-1908308; and in part by the French Government through the "Plan de Relance" and "Programme d'investissements d'avenir".

This paper was presented in part at the IEEE Information Theory Workshop (ITW) 2021, IEEE International Symposium on Information Theory 2022 and IEEE ITW 2022.

The existing body of work in SIET can be studied under three main categories. The first is the study of the the trade-offs between the information and energy rates that can be simultaneously transmitted by an RF signal. Earlier research in this area is focused primarily on the asymptotic regime [1]–[4]. In this case, the notion of information-energy region generalizes to the set of all information and energy rate tuples that can be simultaneously achieved in the asymptotic block-length regime [2]. To capture the trade-off between the information and energy rates, [1] defines a capacity-energy function for various channels including the discrete memoryless channel, binary symmetric channel, and the AWGN channel. In [4], the information-energy trade-off is studied for a coupled-inductor circuit that models a slow frequency-selective fading channel. The information-energy capacity region of the Gaussian multiple access channel is characterized in [5], whereas the information-energy capacity region of the Gaussian interference channel is approximated in [6]. Nonetheless, the information-energy trade-off is not the only trade-off involved in SIET. In the finite block-length regime, several other trade-offs appear which are taken into consideration in this paper.

Within the finite block-length regime, [7] and [8] provide a characterization of the information-energy capacity region with binary antipodal channel inputs. A converse and an achievable information-energy region of SIET with arbitrary number of channel inputs is presented in [9] and [10], respectively. It has been shown in [9] and [10] that, in the non-asymptotic regime, these limits are a function of the type induced by the code used for the transmission. A type is understood in the sense of the empirical frequency with which each channel input symbol appears in the codewords [11]. The impact of energy harvester non-linearities on the fundamental limits of SIET in the finite block-length regime has been studied in [12].

The second area of research that has received considerable attention in SIET is the modelling of the energy harvester (EH) circuits with the aim of providing accurate estimates of the energy gathered from an RF signal. This line of inquiry has revealed that, due to the presence of non-linear elements such as diodes in the EH circuits, the expected energy harvested from a signal is a function of the fourth power of the signal magnitude [13] in addition to the squared magnitude as was conventionally assumed (see [2], [5] and [8]). Recent research on EH non-linearities [14], [15] has shown that energy models that do not account for these non-linearities result in inaccurate estimates of the harvested energy. This work accounts for the EH non-linearities by adopting the models proposed in the literature for determining the energy harvested from the transmitted RF signals.

The third category of research deals with aspects related to the design and implementation of SIET such as signal and system design, resource allocation, receiver architectures, energy harvester circuits, and decoding strategies. Optimal waveform design for SIET from a multi-antenna transmitter to multiple single antenna receivers is studied in [16]. Signal and system design exclusively for wireless energy transmission has been studied in [15], [17]–[20] and [21]. In [22], the authors optimize resource allocation and beamforming for intelligent reflecting surfaces aided SIET. The memory of non-linear elements in the EH circuit is modeled as a Markov decision process in [23] and a learning based model is proposed for the EH circuit. An algorithm for designing a circular quadrature amplitude modulation scheme for SIET that maximizes the peak-to-average power ratio has been proposed in [24]. This paper contributes to this line of research by providing a method of code construction for SIET that proves to be information rate, energy rate and EOP optimal.

More comprehensive overviews of the work on SIET in the second and third categories detailed above, can be found in [25], [26], [27], and [28]. A comparison of relevant aspects of the existing literature on SIET with this work is provided in Table I below. A tick mark indicates that the specific factor has been taken into consideration in the paper while a dash indicates that the concerned feature has not been considered in that reference.

The main contributions of this paper are summarized below.

- The converse for finite block-length SIET over an AWGN channel with a peak-amplitude constraint is characterized for a given finite set of channel input symbols. The converse defines the set of information rate, energy rate, DEP and EOP tuples that cannot be achieved by any code that employs the given set of channel input symbols.
  It is of interest to characterize the information-energy regions for a given set of channel input symbols because channel inputs are usually designed for very specific functions and are, therefore, fixed by design in a system. Ideally, the converse information-energy region would be characterized for the optimal set of channel inputs. However, determining the optimal set of channel inputs is known to be a very difficult problem [44]–[46], even without energy considerations. Interestingly, the process of characterizing the information-energy regions for given sets of channel input symbols results in guidelines for designing optimal channel input symbols for SIET. This characterization also reveals the dependence of the information rate, the energy rate, the DEP and the EOP on various parameters of the code which, in turn, provides insights into which parameters should be changed and how in order to achieve a certain performance objective in SIET.
- Based upon the insights provided by the converse information-energy region, a method of constructing sets of channel input symbols and codes for SIET over an AWGN channel in the finite block-length regime is proposed.
- The achievable information-energy region for the constructed family of codes is characterized. The achievable region defines the set of information rate, energy rate, DEP and EOP tuples that can be achieved by at least one code from the

TABLE I: Summary of the state-of-art

	Point-to-point					
Reference	**Channel inputs**	**Block-length**	**EH non-linearities**	**Channel**	**DEP**	**EOP**
[1]	Infinite	Asymptotic	-	DMC, AWGN	-	-
[4]	Infinite	Asymptotic	-	AWGN + Frequency-selective fading	-	-
[3]	Infinite	Asymptotic	✓	Rayleigh fading	-	-
[29]	Infinite	Asymptotic	-	Flat fading	-	-
[23]	Infinite	Asymptotic	-	AWGN	-	-
[30]	Finite	Asymptotic	-	AWGN + Fading	-	-
[22]	Finite	Asymptotic	-	AWGN + Flat-fading	-	-
[31]	Finite	-	-	AWGN + Rayleigh fading	✓	-
[24]	Finite	-	-	AWGN	-	-
[13]	Infinite	Asymptotic	✓	AWGN	-	-
[14]	Infinite	Asymptotic	✓	Multi-path fading	-	-
[32]	Uncountable	Asymptotic	-	AWGN + Rayleigh fading	✓	-
[33]	Uncountable	Asymptotic	-	AWGN + Rayleigh and Rician fading	-	-
[34]	Uncountable	Asymptotic	-	AWGN	-	-
[35]	Uncountable	Asymptotic	✓	Multi-path fading	-	-
[36]	Uncountable	Asymptotic	-	Multi-path fading	-	-
[37]	Uncountable	Asymptotic	-	AWGN + Flat-fading	-	-
[38]	Finite	Finite	✓	AWGN	-	-
[7]	Finite	Finite	-	BSC	✓	✓
[8]	Finite	Finite	-	BSC	-	-
This work	Finite	Finite	✓	AWGN	✓	✓
	MAC					
[5]	Finite	Asymptotic	-	AWGN	-	-
[6]	Finite	Asymptotic	-	Gaussian interference	-	-
[39]	Uncountable	Asymptotic	-	Path-loss, Rayleigh fading	✓	✓
[40]	Uncountable	Asymptotic	✓	AWGN	-	-
[41]	Uncountable	Asymptotic	-	AWGN + flat fading	✓	✓
[42]	Uncountable	Asymptotic	-	AWGN + flat fading	-	-
[43]	Uncountable	Asymptotic	-	AWGN + flat fading	-	-

constructed family of codes.

- The trade-offs between the information rate, energy rate, DEP and EOP are illustrated through various instructive examples. The information-energy region is also illustrated for a given set of channel input symbols. The study of the gaps between the converse and achievability regions shows that the constructed codes are information rate, energy rate and EOP optimal. The DEP achieved by the constructed codes is, however, sub-optimal.

The rest of this paper is organized as follows. The notations used in this paper and the system model are presented in Sections 2 and 3, respectively. The converse region for finite block-length SIET is characterized in Section 4. Section 5 provides the construction of codes for finite block-length SIET followed by the characterization of an achievable region. The trade-offs between various parameters of finite block-length SIET are elucidated using a simple set of channel input symbols in Section 6. Section 7, illustrates the characterized converse and achievable regions and provides insights into the optimality of the constructed class of codes for SIET.

2. Notation

The sets of natural, real and complex numbers are denoted by $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$, respectively. In particular, $0 \notin \mathbb{N}$. The Borel σ-algebra on $\mathbb{R}$ is denoted by $\mathscr{B}(\mathbb{R})$. Random variables and random vectors are denoted by uppercase letters and uppercase bold letters, respectively. Scalars are denoted by lowercase letters and vectors by lowercase bold letters. The real and imaginary parts of a complex number $c \in \mathbb{C}$ are denoted by $\Re(c)$ and $\Im(c)$, respectively. The complex conjugate of $c \in \mathbb{C}$ is denoted by $c^\star$ and the magnitude of c is denoted by $|c|$. The imaginary unit is denoted by i, *i.e.*, $\mathrm{i}^2 = -1$. The Kullback-Leibler (KL) divergence between two measures P and R, with P absolutely continuous with R, is denoted by $D(P||R) = \int \log \frac{\mathrm{d}P}{\mathrm{d}R} \mathrm{d}P$, where $\frac{\mathrm{d}P}{\mathrm{d}R}$ denotes the Radon-Nykodim derivative of P with respect to R. The Fourier transform of an integrable signal $x(t)$ is denoted by $\hat{x}(f)$. The sinc function is defined as follows

$$\mathrm{sinc}(t) \triangleq \frac{\sin(\pi t)}{\pi t}, \tag{1}$$

and the Q function is given by the following:

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

3. System Model

Consider a communication system formed by a transmitter, an information receiver (IR), and an energy harvester (EH). The objective of the transmitter is to simultaneously send information to the IR at a rate of R bits per second; and energy to the EH at a rate of B Joules per second over an additive white Gaussian noise (AWGN) channel. The transmission takes place over a finite duration of $n \in \mathbb{N}$ channel uses. The transmitter uses L symbols from the set

$$\mathcal{X} \triangleq \{x^{(1)}, x^{(2)}, \ldots, x^{(L)}\} \subset \mathbb{C} \tag{3}$$

that contains all possible channel input symbols. That is,

$$L \triangleq |\mathcal{X}|\,. \tag{4}$$

For all $m \in \{1, 2, \ldots, n\}$, denote by $\nu_m \in \mathcal{X}$, the symbol to be transmitted during channel use m. Denote the vector of channel input symbols over n channel uses by

$$\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_n)^{\mathsf{T}}. \tag{5}$$

The baseband frequency of the transmitter in Hertz (Hz) is denoted by f_w. Denote by $T = \frac{1}{f_w}$, the duration of a channel use in time units. Hence, the transmission takes place during nT time units. The complex baseband signal at time t, with $t \in [0, nT]$ is given by

$$x(t) = \sum_{m=1}^{n} \nu_m \operatorname{sinc}\left(f_w\left(t - (m-1)T\right)\right), \tag{6}$$

where the sinc function is in (1). The signal $x(t)$ in (6) has a bandwidth of $\frac{f_w}{2} > 0$ Hz. Let $f_c > \frac{f_w}{2}$ denote the center frequency of the transmitter. The RF signal input to the channel at time t, denoted by $\tilde{x}(t)$, is obtained by the frequency up-conversion of the baseband signal $x(t)$ in (6) as follows:

$$\tilde{x}(t) = \Re\left(x(t)\sqrt{2}\exp(\mathrm{i}2\pi f_c t)\right), \tag{7}$$

where i is the complex unit. The RF outputs of the AWGN channel at time $t \in [0, nT]$ are the random variables

$$Y(t) = \tilde{x}(t) + N_1(t), \text{ and} \tag{8a}$$

$$Z(t) = \tilde{x}(t) + N_2(t), \tag{8b}$$

where, for all $t \in [0, nT]$, the random variables $N_1(t)$ and $N_2(t)$ represent real white Gaussian noise with zero mean and variance σ^2. $Y(t)$ and $Z(t)$ are the inputs to the IR and the EH, respectively.

At the IR, the received signal $Y(t)$ in (8a) is first multiplied with $\sqrt{2}\exp(-\mathrm{i}2\pi f_c t)$ to obtain the down-converted output. The down-converted output is then passed through a unit gain low pass filter with impulse response $f_w \operatorname{sinc}(f_w t)$ that has a cut-off frequency of $\frac{f_w}{2}$ Hz to obtain the complex baseband equivalent of $Y(t)$. This is followed by ideally sampling the complex baseband output at intervals of $1/f_w$. The resulting discrete time baseband output at the end of n channel uses is given by the following random vector [47, Section 2.2.4]:

$$\boldsymbol{Y} = \boldsymbol{\nu} + \boldsymbol{N}, \tag{9}$$

where the vector $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)^{\mathsf{T}} \in \mathbb{C}^n$ is the input to the IR; $\boldsymbol{\nu}$ is the vector of channel input symbols in (5); and $\boldsymbol{N} = (N_1, N_2, \ldots, N_n)^{\mathsf{T}} \in \mathbb{C}^n$ is the noise vector such that, for all $m \in \{1, 2, \ldots, n\}$, the random variable N_m is a complex circularly symmetric Gaussian random variable whose real and imaginary parts have zero means and variances $\frac{1}{2}\sigma^2$. Moreover, the random variables $N_1, N_2, \ldots, N_n$ are mutually independent (see [47, Section 2.2.4]). That is, for all $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)^{\mathsf{T}} \in \mathbb{C}^n$, and all $\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_n)^{\mathsf{T}} \in \mathbb{C}^n$, the conditional probability density function of the channel output $\boldsymbol{Y}$ in (9) is given by

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{x}) = \prod_{m=1}^{n} f_{Y|X}(y_m|\nu_m) \tag{10}$$

and for all $m \in \{1, 2, \ldots, n\}$,

$$f_{Y|X}(y_m|\nu_m) = \frac{1}{\pi\sigma^2}\exp\left(-\frac{|y_m - \nu_m|^2}{\sigma^2}\right) \tag{11}$$

$$= \frac{1}{\pi\sigma^2}\exp\left(-\frac{(\Re(y_m) - \Re(\nu_m))^2 + (\Im(y_m) - \Im(\nu_m))^2}{\sigma^2}\right). \tag{12}$$

The EH does not down-convert or filter the received input $Z(t)$ (see [15] and [14]). The RF signal $Z(t)$ in (8b) is used as is for harvesting the energy contained in it.

Within this framework, two tasks must be accomplished: information transmission and energy transmission.

A. Information Transmission

Let M be the resolution of the set of indices from which a message is transmitted within n channel uses. That is,

$$M \leqslant 2^{n \log L}, \tag{13}$$

where L is the number of channel input symbols in (4). To reliably transmit a message index, the transmitter uses an (n, M)-code defined as follows.

Definition 3.1 ((n, M)-code)**.** *An (n, M)-code for the random transformation in* (9) *is a system:*

$$\{(\boldsymbol{u}(1), \mathcal{D}_1), (\boldsymbol{u}(2), \mathcal{D}_2), \ldots, (\boldsymbol{u}(M), \mathcal{D}_M)\}, \tag{14}$$

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

$$\boldsymbol{u}(i) = (u_1(i), u_2(i), \ldots, u_n(i)) \in \mathcal{X}^n, \tag{15a}$$

$$\mathcal{D}_i \cap \mathcal{D}_j = \phi, \tag{15b}$$

$$\bigcup_{i=1}^{M} \mathcal{D}_i \subseteq \mathbb{C}^n, \text{ and} \tag{15c}$$

$$|u_m(i)| \leqslant P, \tag{15d}$$

where the real $P > 0$ is the peak-amplitude constraint.

Assume that the transmitter uses the (n, M)-code

$$\mathscr{C} \triangleq \{(\boldsymbol{u}(1), \mathcal{D}_1), (\boldsymbol{u}(2), \mathcal{D}_2), \ldots, (\boldsymbol{u}(M), \mathcal{D}_M)\}, \tag{16}$$

that satisfies (15). The results in this paper are presented in terms of the types induced by the codewords of the given code $\mathscr{C}$ in (16). The type induced by the codeword $\boldsymbol{u}(i)$, with $i \in \{1, 2, \ldots, M\}$, is a probability mass function (pmf) whose support is equal to or a subset of $\mathcal{X}$ in (3). This pmf is denoted by $P_{\boldsymbol{u}(i)}$ and for all $x \in \mathcal{X}$,

$$P_{\boldsymbol{u}(i)}(x) \triangleq \frac{1}{n} \sum_{m=1}^{n} \mathbb{1}_{\{u_m(i)=x\}}. \tag{17}$$

The type induced by all the codewords in $\mathscr{C}$ is also a pmf on the set $\mathcal{X}$ in (3). This pmf is denoted by $P_{\mathscr{C}}$ and for all $x \in \mathcal{X}$,

$$P_{\mathscr{C}}(x) \triangleq \frac{1}{M} \sum_{i=1}^{M} P_{\boldsymbol{u}(i)}(x). \tag{18}$$

The information rate of any (n, M)-code $\mathscr{C}$ is given by

$$R(\mathscr{C}) = \frac{\log M}{n} \tag{19}$$

in bits per channel use. To transmit the message index i, with $i \in \{1, 2, \ldots, M\}$, the transmitter uses the codeword $\boldsymbol{u}(i) = (u_1(i), u_2(i), \ldots, u_n(i))$. That is, at channel use m, with $m \in \{1, 2, \ldots, n\}$, the transmitter inputs the RF signal corresponding to symbol $u_m(i)$ into the channel. At the end of n channel uses, the IR observes a realization of the random vector $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)^{\mathsf{T}}$ in (9). The IR decides that message index j, with $j \in \{1, 2, \ldots, M\}$, was transmitted, if the following event takes place:

$$\boldsymbol{Y} \in \mathcal{D}_j, \tag{20}$$

with $\mathcal{D}_j$ in (16). That is, the set $\mathcal{D}_j \subseteq \mathbb{C}^n$ is the region of correct detection for message index j. Therefore, the DEP associated with the transmission of message index i is given by

$$\gamma_i(\mathscr{C}) \triangleq 1 - \int_{\mathcal{D}_i} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{u}(i)) \mathrm{d}\boldsymbol{y}, \tag{21}$$

and the average DEP for code $\mathscr{C}$ is given by

$$\gamma(\mathscr{C}) \triangleq \frac{1}{M} \sum_{i=1}^{M} \gamma_i(\mathscr{C}). \tag{22}$$

Using this notation, Definition 3.1 can be refined as follows.

Definition 3.2 ((n, M, ϵ)-codes)**.** *An (n, M)-code $\mathscr{C}$ for the random transformation in (9), is said to be an (n, M, ϵ)-code if*

$$\gamma(\mathscr{C}) \leq \epsilon. \tag{23}$$

Using (17) and (18), a special class of codes of particular interest in this study, called homogeneous codes, is defined hereunder.

Definition 3.3 (Homogeneous Codes)**.** *An (n, M, ϵ)-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16) *is said to be homogeneous if for all $i \in \{1, 2, \ldots, M\}$ and for all $x \in \mathcal{X}$, with $\mathcal{X}$ in* (3) *it holds that*

$$P_{\boldsymbol{u}(i)}(x) = P_{\mathscr{C}}(x), \tag{24}$$

where, $P_{\boldsymbol{u}(i)}$ and $P_{\mathscr{C}}$ are the types defined in (17) *and* (18)*, respectively.*

Homogeneous codes are essentially (n, M, ϵ)-codes in which a given channel input symbol is used the same number of times in all codewords.

B. Energy Transmission

While transmitting message $i \in \{1, 2, \ldots, M\}$, the channel output observed at the EH is denoted by $Z_i(t)$, with $t \in [0, nT]$. From (7) and (8b), the channel output $Z_i(t)$ is given by

$$Z_i(t) = \Re\Big(\sqrt{2} \sum_{m=1}^{n} u_m(i) \operatorname{sinc}\left(f_w(t - (m-1)T)\right) \exp\left(\mathrm{i}2\pi f_c t\right) \Big) + N_2(t) \tag{25}$$

$$= x_i(t) + N_2(t), \tag{26}$$

where, for all $m \in \{1, 2, \ldots, n\}$, the complex $u_m(i)$ is the m^{th} symbol of the codeword $\boldsymbol{u}(i)$ in (15a); for all $t \in [0, nT]$, the signal $x_i(t)$ in (26) is

$$x_i(t) = \Re\Big(\sqrt{2} \sum_{m=1}^{n} u_m(i) \operatorname{sinc}\left(f_w(t - (m-1)T)\right) \exp\left(\mathrm{i}2\pi f_c t\right) \Big); \tag{27}$$

and the random variable $N_2(t)$ is a real Gaussian random variable with zero mean and variance σ^2 in (10). For all $t \in [0, nT]$, the channel output $Z_i(t)$ in (26) is a real Gaussian random variable with mean $x_i(t)$ and variance σ^2.

The non-linear energy model in [14] and [15] states that the energy harvested from a signal is proportional to the DC component of the second and fourth powers of the signal. Using this model, for all $i \in \{1, 2, \ldots, M\}$, the expected energy harvested from the channel output $Z_i(t)$ in (25) during the time $t \in [0, nT]$ is given by the following:

$$e_i \triangleq k_1 \sum_{m=1}^{n} |u_m(i)|^2 + k_2 \sum_{m=1}^{n} |u_m(i)|^4 \tag{28}$$

$$= k_1 \sum_{x \in \mathcal{X}} nP_{\boldsymbol{u}(i)}(x) |x|^2 + k_2 \sum_{x \in \mathcal{X}} nP_{\boldsymbol{u}(i)}(x) |x|^4, \tag{29}$$

where $x \in \mathcal{X}$, with $\mathcal{X}$ in (3); k_1 and k_2 are positive constants, with $k_1 = 0.0034$ and $k_2 = 0.3829$ [14]; for all $m \in \{1, 2, \ldots, n\}$, the complex $u_m(i)$ is in (15a) and $P_{\boldsymbol{u}(i)}$ is the type defined in (17). From (29), for a homogeneous code $\mathscr{C}$ (Definition 3.3), the expected energy harvested from the channel output $Z_i(t)$ in (26) during the time $t \in [0, nT]$ is equal for all $i \in \{1, 2, \ldots, M\}$ and is given by the following:

$$e_{\mathscr{C}} \triangleq k_1 \sum_{x \in \mathcal{X}} nP_{\mathscr{C}}(x) |x|^2 + k_2 \sum_{x \in \mathcal{X}} nP_{\mathscr{C}}(x) |x|^4, \tag{30}$$

where $x \in \mathcal{X}$, with $\mathcal{X}$ in (3) and $P_{\mathscr{C}}$ is the type defined in (18).

Denote by $M' \leq M$, the number of unique values in the vectior $(e_1, e_2, \ldots, e_M)^{\mathsf{T}}$ with e_i in (28). The M' unique energy levels are given by $\{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_{M'}\}$. Assume without loss of generality that the following holds:

$$0 < \bar{e}_1 < \bar{e}_2 < \ldots < \bar{e}_{M'}. \tag{31}$$

More specifically, for all $i \in \{1, 2, \ldots, M\}$, there exists $j \in \{1, 2, \ldots, M'\}$ such that $e_i = \bar{e}_j$ with e_i in (28).

For all $j \in \{1, 2, \ldots, M'\}$, define variables y_j to be the number of codewords that carry energy $\bar{e}_j$. More precisely, for all $j \in \{1, 2, \ldots, M'\}$, y_j is given by

$$y_j = \sum_{i=1}^{M} \mathbb{1}_{\{e_i = \bar{e}_j\}}. \tag{32}$$

It follows that $\sum_{j=1}^{M'} y_j = M$.

Let W be a random variable that denotes the message index sent during time $t \in [0, nT]$. For all $i \in \{1, 2, \ldots, M\}$, the probability of transmitting message index i is given by

$$P_W(i) = \frac{1}{M}. \tag{33}$$

The energy harvested during time $t \in [0, nT]$ is also a random variable denoted by E. The probability of harvesting energy e given that message index i was transmitted is given by

$$P_{E|W}(e|i) = \mathbb{1}_{\{e=e_i\}}, \tag{34}$$

where e_i is in (28). The pmf of the random variable E, denoted by P_E is given by the following:

$$P_E(e) = \sum_{i=1}^{M} P_{E|W}(e|i) P_W(i) \tag{35}$$

$$= \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\{e=e_i\}}. \tag{36}$$

Using (35), the EOP associated with code $\mathscr{C}$ is defined as follows:

$$\theta(\mathscr{C}, B) \triangleq \Pr(E < B) = \sum_{i \in \{j \in \{1,2,\ldots,M\}: e_j < B\}} P_E(e_i) \tag{37}$$

$$= \sum_{i \in \{j \in \{1,2,\ldots,M\}: e_j < B\}} \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\{e=e_i\}} \tag{38}$$

$$= \frac{1}{M} \Big| \{i \in \{1, 2, \ldots, M\} : e_i < B\} \Big| \tag{39}$$

$$= \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\{e_i < B\}}, \tag{40}$$

where, P_E is the pmf defined in (35), e_i is defined in (28) and, the equality in (38) follows from (36).

From (40), an interesting insight can be derived about $\theta(\mathscr{C}, B)$. The EOP $\theta(\mathscr{C}, B)$ can only take discrete values in the set $\{0, \frac{1}{M}, \frac{2}{M}, \ldots, 1\}$. Moreover, for homogeneous codes [9, Definition 4], it holds that $\theta(\mathscr{C}, B) \in \{0, 1\}$. The following refinement of Definition 3.2 follows from (37).

Definition 3.4 ($(n, M, \epsilon, B, \delta)$-code)**.** *An (n, M, ϵ)-code $\mathscr{C}$ for the random transformation in* (9) *is said to be an $(n, M, \epsilon, B, \delta)$-code if*

$$\theta(\mathscr{C}, B) \leq \delta. \tag{41}$$

4. Information-Energy Converse Region

The main results in this section provide upper bounds on the information rate R in (19), lower bounds on the average DEP ϵ in (23) and lower bounds on the EOP δ in (41) for all possible $(n, M, \epsilon, B, \delta)$-codes that employ a given set of channel input symbols $\mathcal{X}$ of the form in (3).

A. Energy Transmission Rate and EOP

The following corollary introduces a lower bound on the EOP that holds for all possible $(n, M, \epsilon, B, \delta)$-codes.

Corollary 4.1. *Given an $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16)*, the following holds:*

$$\delta \geq \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\{e_i < B\}}, \tag{42}$$

where, for all $i \in \{1, 2, \ldots, M\}$, the real $e_i \in [0, \infty)$ is in (28)*.*

Proof. The result follows from (40) and (41). □

The following lemma provides the bound on the EOP for the class of homogeneous codes.

Lemma 4.2. *Given a homogeneous $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16)*, the following holds:*

$$\delta = \mathbb{1}_{\{e_{\mathscr{C}} < B\}}, \tag{43}$$

where $e_{\mathscr{C}} \in [0, \infty)$ *as in* (30)*.*

Proof. From (40), the average EOP of code $\mathscr{C}$ is given by

$$\theta(\mathscr{C}, B) = \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\{e_i < B\}}. \tag{44}$$

Since $\mathscr{C}$ is homogeneous, from (44) and (30) it follows that

$$\theta(\mathscr{C}, B) = \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\{e_{\mathscr{C}} < B\}} \tag{45}$$

$$= \mathbb{1}_{\{e_{\mathscr{C}} < B\}}. \tag{46}$$

From (41) and (46), it follows that

$$\delta = \mathbb{1}_{\{e_{\mathscr{C}} < B\}}. \tag{47}$$

This completes the proof. □

Note that, while (41) would suggest that (47) should be an inequality, namely $\delta \geq \mathbb{1}_{\{e_{\mathscr{C}} < B\}}$. However, since $\delta \in \{0, 1\}$ for the homogeneous code $\mathscr{C}$ in Lemma 4.2, (47) holds with equality.

The derived lower bound on δ in Corollary 4.1 can now be used for determining the upper bound on the energy transmission rate B as follows.

Lemma 4.3. *Given an* $(n, M, \epsilon, B, \delta)$*-code* $\mathscr{C}$ *for the random transformation in* (9) *of the form in* (16)*, the energy transmission rate* B *satisfies,*

$$B \leq \begin{cases} \bar{e}_1 & \text{if } \ 0 \leq \delta < \frac{y_1}{M}, \\ \bar{e}_j & \text{if } \ \frac{\sum_{k=1}^{j-1} y_k}{M} \leq \delta < \frac{\sum_{k=1}^{j} y_k}{M}, j \in \{2, 3, \ldots, M'\}. \end{cases} \tag{48}$$

where, for all $i \in \{1, 2, \ldots, M\}$*, energy* e_i *is in* (28)*; the positive integer* M' *is in* (31)*; and, for all* $j \in \{1, 2, \ldots, M'\}$*,* $\bar{e}_j$ *is in* (31) *and* y_j *is in* (32)*.*

Proof. From (42), it follows that,

$$\sum_{i=1}^{M} \mathbb{1}_{\{e_i < B\}} \leq M\delta. \tag{49}$$

This implies that, in order to achieve an EOP less than or equal to δ, the number of codewords that have energy less than B (given by $\sum_{i=1}^{M} \mathbb{1}_{\{e_i < B\}}$) can be at most equal to $\lfloor M\delta \rfloor$. This allows calculating the upper bound on the energy transmission rate B as follows.

From (32) and the definition of the EOP in (40), it can be seen that the EOP δ can only take values from the following set:

$$\left\{0, \frac{y_1}{M}, \frac{y_1 + y_2}{M}, \ldots, \frac{\sum_{j=1}^{M'-1} y_j}{M}, 1\right\}. \tag{50}$$

For $0 \leq \delta < \frac{y_1}{M}$, at most $\lfloor M\delta \rfloor = y_1$ codewords can have energy less than B which is possible only if

$$B \leq \bar{e}_1. \tag{51}$$

For $\frac{y_1}{M} \leq \delta < \frac{y_1 + y_2}{M}$, at most $\lfloor M\delta \rfloor = y_1 + y_2$ codewords can have energy less than B which is possible only if

$$B \leq \bar{e}_2. \tag{52}$$

The upper bound on B can similarly be calculated for all possible values of δ in (50). Therefore, the upper bound for B is given by the following:

$$B \leq \begin{cases} \bar{e}_1 & \text{if } 0 \leq \delta < \frac{y_1}{M}, \\ \bar{e}_j & \text{if } \frac{\sum_{1}^{j-1} y_j}{M} \leq \delta < \frac{\sum_{1}^{j} y_j}{M}, j \in \{2, 3, \ldots, M'\}. \end{cases} \tag{53}$$

This completes the proof. □

Consider the special case in which all the M codewords carry a different amount of energy. That is, for all $i \in \{1, 2, \ldots, M\}$, $e_1 \neq e_2 \neq \ldots \neq e_M$. Furthermore, assume without loss of generality that

$$e_1 < e_2 < \ldots < e_M. \tag{54}$$

For this particular case, the upper bound on B is given by the following:

$$B \leq e_i, \quad \text{if } \frac{(i-1)}{M} \leq \delta < \frac{i}{M}, i \in \{1, 2, \ldots, M\}. \tag{55}$$

Note that (55) is a special case of (53) for $M' = M$.

B. Average Decoding Error Probability

The average DEP of a given code $\mathscr{C}$ of the form in (16) depends on the choice of the decoding sets $\mathcal{D}_1$, $\mathcal{D}_2$, ..., $\mathcal{D}_M$. For all $i \in \{1, 2, \ldots, M\}$, consider the smallest sets $\mathcal{D}_{i,1}$, $\mathcal{D}_{i,2}$, ..., $\mathcal{D}_{i,n}$ such that the decoding sets $\mathcal{D}_i$s are of the following form:

$$\mathcal{D}_i \subseteq \mathcal{D}_{i,1} \times \mathcal{D}_{i,2} \times \ldots \times \mathcal{D}_{i,n}. \tag{56}$$

Note that these sets exist by construction as follows. For all $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)^{\mathsf{T}} \in \mathcal{D}_i$, let for all $m \in \{1, 2, \ldots, n\}$, $y_m \in \mathcal{D}_{i,m}$. Hence, the inclusion in (56) holds.

The following lemma proves that it is sub-optimal to have the decoding set of a symbol $u_m(i)$ be dependent on the codeword $i \in \{1, 2, \ldots, M\}$ and/or the position of the symbol $m \in \{1, 2, \ldots, n\}$ in the codeword. In other words, for all $(i, i') \in \{1, 2, \ldots, M\}^2$, all $(m, m') \in \{1, 2, \ldots, n\}^2$, and all $\ell \in \{1, 2, \ldots, L\}$, if $u_m(i) = u_{m'}(i') = x^{(\ell)}$, with $x^{(\ell)} \in \mathcal{X}$, it is sub-optimal to have $\mathcal{D}_{i,m} \neq \mathcal{D}_{i',m'}$.

Lemma 4.4. *Consider an $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16) *with the set of channel input symbols $\mathcal{X}$ in* (3) *and decoding sets of the form in* (56)*. For all $i \in \{1, 2, \ldots, M\}$, all $m \in \{1, 2, \ldots, n\}$, and all $\ell \in \{1, 2, \ldots, L\}$, if $u_m(i) = x^{(\ell)}$, with $x^{(\ell)} \in \mathcal{X}$, the sets $\mathcal{D}_{i,m}$ in* (56) *are given by the following:*

$$\mathcal{D}_{i,m} = \bar{\mathcal{E}}_\ell, \tag{57}$$

where, for all $\ell \in \{1, 2, \ldots, L\}$,

$$\bar{\mathcal{E}}_\ell \triangleq \mathcal{D}_{j,\omega(j,\ell)}, \tag{58}$$

with,

$$j \in \arg \max_{k \in \{1,2,\ldots,M\}} \int_{\mathcal{D}_{k,\omega(k,\ell)}} f_{Y|X}\left(y|x^{(\ell)}\right) \mathrm{d}y, \tag{59}$$

where,

$$\omega(k, \ell) \in \arg \max_{m \in \{1,2,\ldots,n\}} \int_{\mathcal{D}_{k,m}} f_{Y|X}\left(y|u_m(k) = x^{(\ell)}\right) \mathrm{d}y. \tag{60}$$

Given any other $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}'$ for the random transformation in (9) *of the form in* (16)*, with the same set of channel input symbols $\mathcal{X}$ in* (3) *and decoding sets $\mathcal{D}_i$ of the form in* (56)*, the average DEP γ in* (22) *of $\mathscr{C}$ and $\mathscr{C}'$ satisfy the following:*

$$\gamma\left(\mathscr{C}'\right) \geq \gamma\left(\mathscr{C}\right). \tag{61}$$

Proof. From (21), for all $i \in \{1, 2, \ldots, M\}$, the DEP for codeword i of code $\mathscr{C}'$ is given by

$$\gamma_i(\mathscr{C}') = 1 - P_{\boldsymbol{Y}|\boldsymbol{X}=\boldsymbol{u}(i)}\left(\mathcal{D}_i\right) \tag{62}$$

$$\geq 1 - P_{\boldsymbol{Y}|\boldsymbol{X}=\boldsymbol{u}(i)}\left(\mathcal{D}_{i,1} \times \mathcal{D}_{i,2} \times \ldots \times \mathcal{D}_{i,n}\right) \tag{63}$$

$$= 1 - \int_{\mathcal{D}_{i,1} \times \mathcal{D}_{i,2} \times \ldots \times \mathcal{D}_{i,n}} \prod_{m=1}^{n} f_{Y|X}\left(y|u_m(i)\right) \mathrm{d}y \tag{64}$$

$$= 1 - \prod_{m=1}^{n} \int_{\mathcal{D}_{i,m}} f_{Y|X}\left(y|u_m(i)\right) \mathrm{d}y, \tag{65}$$

where $P_{\boldsymbol{Y}|\boldsymbol{X}}$ is the probability measure induced by the pdf $f_{\boldsymbol{Y}|\boldsymbol{X}}$ in (10); for all $i \in \{1, 2, \ldots, M\}$ and all $m \in \{1, 2, \ldots, n\}$, $\boldsymbol{u}(i)$ and $u_m(i)$ are in (15a); the inequality in (63) follows from (56); equality in (64) follows from (10); and (65) follows from (64) due to Fubini's theorem. Let $\omega : \{1, 2, \ldots, M\} \times \{1, 2, \ldots, L\} \to \{1, 2, \ldots, n\}$ be defined as follows

$$\omega(i, \ell) \in \arg \max_{m \in \{1,2,\ldots,n\}} \int_{\mathcal{D}_{i,m}} f_{Y|X}\left(y|u_m(i) = x^{(\ell)}\right) \mathrm{d}y, \tag{66}$$

with $x^{(\ell)} \in \mathcal{X}$, and $\mathcal{X}$ in (3). For all $i \in \{1, 2, \ldots, M\}$, let the sets $\mu_{i1}, \mu_{i2}, \ldots, \mu_{iL}$ be a partition of $\{1, 2, \ldots, n\}$ such that, for all $\ell \in \{1, 2, \ldots, L\}$, the set $\mu_{i\ell}$ is given by

$$\mu_{i\ell} = \left\{ m \in \{1, 2, \ldots, n\} : u_m(i) = x^{(\ell)} \right\}. \tag{67}$$

Hence, for all $i \in \{1, 2, \ldots, M\}$, from (65), it follows that,

$$\gamma_i(\mathscr{C}') \geq 1 - \prod_{\ell=1}^{L} \left(\prod_{m \in \mu_{i\ell}} \int_{\mathcal{D}_{i,m}} f_{Y|X}\left(y|x^{(\ell)}\right) \mathrm{d}y \right) \tag{68}$$

$$\geq 1 - \prod_{\ell=1}^{L} \left(\prod_{m \in \mu_{i\ell}} \int_{\mathcal{D}_{i,\omega(i,\ell)}} f_{Y|X}\left(y|x^{(\ell)}\right) \mathrm{d}y \right) \tag{69}$$

$$= 1 - \prod_{\ell=1}^{L} \left(\int_{\mathcal{D}_{i,\omega(i,\ell)}} f_{Y|X}\left(y|x^{(\ell)}\right) \mathrm{d}y \right)^{nP_{\boldsymbol{u}(i)}(x^{(\ell)})}, \tag{70}$$

where, the inequality in (69) follows from (66). For all $\ell \in \{1, 2, \ldots, L\}$, let the sets $\bar{\mathcal{E}}_1, \bar{\mathcal{E}}_2, \ldots, \bar{\mathcal{E}}_L$ be such that

$$\bar{\mathcal{E}}_\ell \triangleq \mathcal{D}_{j,\omega(j,\ell)}, \tag{71}$$

$$\text{with } j \in \arg \max_{k \in \{1,2,\ldots,M\}} \int_{\mathcal{D}_{k,\omega(k,\ell)}} f_{Y|X}\left(y|x^{(\ell)}\right) \mathrm{d}y. \tag{72}$$

Hence, from (70) and (71), for all $i \in \{1, 2, \ldots, M\}$, it follows that

$$\gamma_i(\mathscr{C}') \geqslant 1 - \prod_{\ell=1}^{L} \left(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \right)^{nP_{\boldsymbol{u}(i)}(x^{(\ell)})}. \tag{73}$$

From (22) and (73), it follows that the average DEP $\gamma(\mathscr{C}')$ is given by

$$\gamma\left(\mathscr{C}'\right) \geq \frac{1}{M} \sum_{i=1}^{M} \left(1 - \prod_{\ell=1}^{L} \left(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \right)^{nP_{\boldsymbol{u}(i)}(x^{(\ell)})} \right) \tag{74}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{\ell=1}^{L} \left(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \right)^{nP_{\boldsymbol{u}(i)}(x^{(\ell)})}, \tag{75}$$

$$= \gamma\left(\mathscr{C}\right). \tag{76}$$

This completes the proof. □

The following lemma introduces a lower bound on the DEP ϵ in (23) for an $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ of the form in (16).

Lemma 4.5. *Given an $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16) *and decoding sets of the form in* (56)*, the average DEP ϵ in* (23) *satisfies the following:*

$$\epsilon \geq 1 - \frac{1}{M} \sum_{i=1}^{M} \exp\left(-nH(P_{\boldsymbol{u}(i)}) - nD(P_{\boldsymbol{u}(i)}||F) + n \log L \right), \tag{77}$$

where, $P_{\boldsymbol{u}(i)}$ is the type defined in (17)*; L is in* (4)*; the function F is a pmf on $\mathcal{X}$ in* (3) *such that for all $\ell \in \{1, 2, \ldots, L\}$,*

$$F\left(x^{(\ell)}\right) = \frac{\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y}{\sum_{j=1}^{L} \int_{\bar{\mathcal{E}}_j} f_{Y|X}(y|x^{(j)}) \mathrm{d}y}; \tag{78}$$

and for all $\ell \in \{1, 2, \ldots, L\}$, the set $\bar{\mathcal{E}}_\ell$ is in (71)*.*

Proof. From (75), the average DEP of the $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ is given by

$$\gamma(\mathscr{C}) \geq 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{\ell=1}^{L} \left(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \right)^{nP_{\boldsymbol{u}(i)}(x^{(\ell)})}. \tag{79}$$

Evaluating the product of integrals in (79), the following holds:

$$\prod_{\ell=1}^{L}\Big(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)})\mathrm{d}y\Big)^{nP_{\boldsymbol{u}(i)}(x^{(\ell)})}$$
$$= \exp\left(\log\prod_{\ell=1}^{L}\Big(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)})\mathrm{d}y\Big)^{nP_{\boldsymbol{u}(i)}(x^{(\ell)})}\right), \tag{80}$$
$$= \exp\left(\sum_{\ell=1}^{L} nP_{\boldsymbol{u}(i)}(x^{(\ell)})\log\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)})\mathrm{d}y\right), \tag{81}$$
$$= \exp\left(\sum_{\ell=1}^{L} nP_{\boldsymbol{u}(i)}(x^{(\ell)})\left(\log\frac{\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)})\mathrm{d}y}{\sum_{j=1}^{L}\left(\int_{\bar{\mathcal{E}}_j} f_{Y|X}(y|x^{(j)})\mathrm{d}y\right)}\right.\right.$$
$$\left.\left.+\log\sum_{j=1}^{L}\left(\int_{\bar{\mathcal{E}}_j} f_{Y|X}(y|x^{(j)})\mathrm{d}y\right)+\log P_{\boldsymbol{u}(i)}(x^{(\ell)})-\log P_{\boldsymbol{u}(i)}(x^{(\ell)})\right)\right). \tag{82}$$

Plugging (78) in (82) yields,

$$\prod_{\ell=1}^{L}\Big(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)})\mathrm{d}y\Big)^{nP_{\boldsymbol{u}(i)}(x^{(\ell)})}$$
$$= \exp\left(\sum_{\ell=1}^{L}\left(nP_{\boldsymbol{u}(i)}(x^{(\ell)})\log\frac{F\left(x^{(\ell)}\right)}{P_{\boldsymbol{u}(i)}(x^{(\ell)})}\right.\right.$$
$$\left.\left.+nP_{\boldsymbol{u}(i)}(x^{(\ell)})\log\left(\sum_{j=1}^{L}\int_{\bar{\mathcal{E}}_j} f_{Y|X}(y|x^{(j)})\mathrm{d}y\right)+nP_{\boldsymbol{u}(i)}(x^{(\ell)})\log P_{\boldsymbol{u}(i)}(x^{(\ell)})\right)\right) \tag{83}$$
$$= \exp\left(-nH(P_{\boldsymbol{u}(i)})-nD(P_{\boldsymbol{u}(i)}||F)+n\log\sum_{j=1}^{L}\int_{\bar{\mathcal{E}}_j} f_{Y|X}(y|x^{(j)})\mathrm{d}y\right). \tag{84}$$

Using (84) in (74) yields,

$$\gamma(\mathscr{C}) \geq 1-\frac{1}{M}\sum_{i=1}^{M}\exp\left(-nH(P_{\boldsymbol{u}(i)})-nD(P_{\boldsymbol{u}(i)}||F)+n\log\sum_{j=1}^{L}\int_{\bar{\mathcal{E}}_j} f_{Y|X}(y|x^{(j)})\mathrm{d}y\right) \tag{85}$$
$$\geq 1-\frac{1}{M}\sum_{i=1}^{M}\exp\left(-nH(P_{\boldsymbol{u}(i)})-nD(P_{\boldsymbol{u}(i)}||F)+n\log L\right). \tag{86}$$

Finally, from (23) and (86), it follows that

$$\epsilon \geq 1-\frac{1}{M}\sum_{i=1}^{M}\exp\left(-nH(P_{\boldsymbol{u}(i)})-nD(P_{\boldsymbol{u}(i)}||F)+n\log L\right), \tag{87}$$

which completes the proof. □

The following lemma provides another lower bound on the average DEP ϵ that recovers the optimal decoding sets for the channel input symbols as defined by the maximum a posteriori (MAP) decision rule [48, Chapter 21]. For all $\ell \in \{1,2,\ldots,L\}$, define the sets $\mathcal{E}_\ell$ as follows:

$$\mathcal{E}_\ell \triangleq \begin{cases} \left\{y\in\mathbb{C}:\ \forall\ \ell'\in\{1,2,\ldots,L\}, P_{\mathscr{C}}(x^{(\ell)})f_{Y|X}(y|x^{(\ell)}) \geq P_{\mathscr{C}}(x^{(\ell')})f_{Y|X}(y|x^{(\ell')})\right\} & \text{if}\ \ \ell=1, \\ \left\{y\in\mathbb{C}:\ \forall\ \ell'\in\{1,2,\ldots,L\}, P_{\mathscr{C}}(x^{(\ell)})f_{Y|X}(y|x^{(\ell)}) \geq P_{\mathscr{C}}(x^{(\ell')})f_{Y|X}(y|x^{(\ell')})\right\} \setminus \bigcup_{k=1}^{\ell-1}\mathcal{E}_k & \text{otherwise.} \end{cases} \tag{88}$$

Note that the definition of the sets in (88) is not unique because the set of channel input symbols $\mathcal{X}$ in (3) is not an ordered set. Therefore, one could go through the L symbols in $\mathcal{X}$ in any order ($L!$ possibilities) to define the sets in (88) without impacting the resulting DEP.

Lemma 4.6. *Given an $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16) *such that* (56) *holds, the average DEP ϵ satisfies the following:*

$$\epsilon \geq 1 - \sum_{\ell=1}^{L} \int_{\mathcal{E}_\ell} P_{\mathscr{C}}(x^{(\ell)}) f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y, \tag{89}$$

where, $P_{\mathscr{C}}$ is the type defined in (18)*; L is in* (4)*; $x^{(\ell)} \in \mathcal{X}$ in* (3)*, and for all $\ell \in \{1, 2, \ldots, L\}$, the set $\mathcal{E}_\ell$ is defined in* (88)*.*

Proof. From (75), the average DEP of the $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ is given by

$$\gamma(\mathscr{C}) \geq 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{\ell=1}^{L} \Big(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \Big)^{n P_{\boldsymbol{u}(i)}(x^{(\ell)})}. \tag{90}$$

The second term on the right hand side of (90) can be written as follows.

$$\begin{aligned}
\frac{1}{M} \sum_{i=1}^{M} \prod_{\ell=1}^{L} \Big(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \Big)^{n P_{\boldsymbol{u}(i)}(x^{(\ell)})} &= \frac{1}{M} \sum_{i=1}^{M} \exp\left(\log \left(\prod_{\ell=1}^{L} \Big(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \Big)^{n P_{\boldsymbol{u}(i)}(x^{(\ell)})} \right) \right) && (91) \\
&= \frac{1}{M} \sum_{i=1}^{M} \exp\left(\sum_{\ell=1}^{L} n P_{\boldsymbol{u}(i)}(x^{(\ell)}) \log \Big(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \Big) \right) && (92) \\
&\leq \frac{1}{M} \sum_{i=1}^{M} \exp\left(n \log \left(\sum_{\ell=1}^{L} P_{\boldsymbol{u}(i)}(x^{(\ell)}) \int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \right) \right) && (93) \\
&= \frac{1}{M} \sum_{i=1}^{M} \exp\left(\log \left(\sum_{\ell=1}^{L} P_{\boldsymbol{u}(i)}(x^{(\ell)}) \int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \right)^{n} \right) && (94) \\
&= \frac{1}{M} \sum_{i=1}^{M} \left(\sum_{\ell=1}^{L} P_{\boldsymbol{u}(i)}(x^{(\ell)}) \int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \right)^{n} && (95) \\
&\leq \frac{1}{M} \sum_{i=1}^{M} \left(\sum_{\ell=1}^{L} P_{\boldsymbol{u}(i)}(x^{(\ell)}) \int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \right) && (96) \\
&= \sum_{\ell=1}^{L} \left(\frac{1}{M} \sum_{i=1}^{M} P_{\boldsymbol{u}(i)}(x^{(\ell)}) \right) \int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y && (97) \\
&= \sum_{\ell=1}^{L} P_{\mathscr{C}}(x^{(\ell)}) \int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y && (98) \\
&= \sum_{\ell=1}^{L} \int_{\bar{\mathcal{E}}_\ell} P_{\mathscr{C}}(x^{(\ell)}) f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y, && (99)
\end{aligned}$$

where, the inequality in (93) follows from the concavity of the log function and Jensen's inequality; the inequality in (96) follows because $\sum_{\ell=1}^{L} P_{\boldsymbol{u}(i)}(x^{(\ell)}) \int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \leq 1$; and the equality in (98) follows from (18). From (99) it follows that,

$$\begin{aligned}
\frac{1}{M} \sum_{i=1}^{M} \left(\prod_{\ell=1}^{L} \Big(\int_{\bar{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \Big)^{n P_{\boldsymbol{u}(i)}(x^{(\ell)})} \right) &\leq \sum_{\ell=1}^{L} \int_{\bar{\mathcal{E}}_\ell} P_{\mathscr{C}}(x^{(\ell)}) f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y && (100) \\
&\leq \sum_{\ell=1}^{L} \int_{\mathcal{E}_\ell} P_{\mathscr{C}}(x^{(\ell)}) f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y, && (101)
\end{aligned}$$

where, for all $\ell \in \{1, 2, \ldots, L\}$, the set $\mathcal{E}_\ell$ is in (88). From (90) and (101), it follows that

$$\gamma(\mathscr{C}) \geq 1 - \sum_{\ell=1}^{L} \int_{\mathcal{E}_\ell} P_{\mathscr{C}}(x^{(\ell)}) f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y. \tag{102}$$

Finally, from (23) and (102), it follows that

$$\epsilon \geq 1 - \sum_{\ell=1}^{L} \int_{\mathcal{E}_\ell} P_{\mathscr{C}}(x^{(\ell)}) f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y. \tag{103}$$

This completes the proof. □

The following result simplifies the lower bound on the DEP in Lemma 4.6 for the case of homogeneous codes (Defintion 3.3) as follows.

Corollary 4.7. *Given a homogeneous $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16) *such that* (56) *holds, the DEP ϵ satisfies that*

$$\epsilon \geq 1 - \exp\Big(- nH\left(P_{\mathscr{C}}\right) - nD\left(P_{\mathscr{C}}||F\right) + n\log L\Big), \tag{104}$$

where, $P_{\mathscr{C}}$ is the type defined in (18) *and the pmf F is defined in* (78)*.*

The lower bound on the DEP in (104) provides interesting insights into the trade-offs between the DEP and the energy transmission rate B for an $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$. As the divergence between the type $P_{\mathscr{C}}$ in (18) and the pmf F in (78) increases, the lower bound on the DEP increases. Furthermore, if the type $P_{\mathscr{C}}$ of a symbol $x^{(\ell)}$ is increased/decreased to increase the harvested energy B, the bound in (104) dictates that the corresponding decoding set $\mathcal{E}_\ell$ should be proportionately made bigger/smaller to respect the same DEP ϵ.

Though the previous lemmas provide interesting insights into the behavior of the lower bound on the DEP of an $(n, M, \epsilon, B, \delta)$-code, these are often difficult to calculate for large sets of channel inputs. The following lemma introduces a more tractable lower bound on the DEP ϵ that can be easily calculated for any form of the set of channel input symbols.

Lemma 4.8. *Consider an $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16)*. For all $\ell \in \{1, 2, \ldots, L\}$, the complex $\bar{x}^{(\ell)} \in \mathcal{X}$ in* (3)*, denotes the nearest neighbor of the symbol $x^{(\ell)} \in \mathcal{X}$, i.e.,*

$$\bar{x}^{(\ell)} \in \arg \min_{x \in \mathcal{X} \setminus \{x^{(\ell)}\}} \left| x^{(\ell)} - x \right|. \tag{105}$$

Then, the DEP ϵ satisfies,

$$\epsilon \geq \mathrm{Q}\left(\frac{|x^{(\ell^\star)} - \bar{x}^{(\ell^\star)}|}{\sqrt{2\sigma^2}} - \frac{\sigma}{\sqrt{2}|x^{(\ell^\star)} - \bar{x}^{(\ell^\star)}|} \log\left(\frac{P_{\mathscr{C}}(\bar{x}^{(\ell^\star)})}{P_{\mathscr{C}}(x^{(\ell^\star)})} \right) \right), \tag{106}$$

where, the function Q is defined in (2)*, the real σ^2 is defined in* (12) *and $\ell^\star \in \{1, 2, \ldots, L\}$ is such that*

$$\ell^\star \in \arg \max_{\ell \in \{1,2,\ldots,L\}} \left(1 - \mathrm{Q}\left(\frac{|x^{(\ell)} - \bar{x}^{(\ell)}|}{\sqrt{2\sigma^2}} - \frac{\sigma}{\sqrt{2}|x^{(\ell)} - \bar{x}^{(\ell)}|} \log\left(\frac{P_{\mathscr{C}}(\bar{x}^{(\ell)})}{P_{\mathscr{C}}(x^{(\ell)})} \right) \right) \right). \tag{107}$$

Proof. From Lemma 4.6, the average DEP ϵ for code $\mathscr{C}$ satisfies the following:

$$\epsilon \geq 1 - \sum_{\ell=1}^{L} P_{\mathscr{C}}(x^{(\ell)}) \int_{\mathcal{E}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y, \tag{108}$$

with the set $\mathcal{E}_\ell$ in (88). For all $\ell \in \{1, 2, \ldots, L\}$ denote by $\bar{x}^{(\ell)}$, a nearest neighbor of $x^{(\ell)}$ as in (105). Denote by $\hat{\mathcal{E}}_\ell$, the following region

$$\hat{\mathcal{E}}_\ell = \left\{ y \in \mathbb{C} : P_{\mathscr{C}}(x^{(\ell)}) f_{Y|X}(y|x^{(\ell)}) \geq P_{\mathscr{C}}(\bar{x}^{(\ell)}) f_{Y|X}(y|\bar{x}^{(\ell)}) \right\}. \tag{109}$$

From the definition of $\mathcal{E}_\ell$ in (88) and the definition of $\hat{\mathcal{E}}_\ell$ in (109), it follows that,

$$\mathcal{E}_\ell \subseteq \hat{\mathcal{E}}_\ell \tag{110}$$

Therefore, from (110) and (108) it follows that,

$$\epsilon \geq 1 - \sum_{\ell=1}^{L} P_{\mathscr{C}}(x^{(\ell)}) \int_{\hat{\mathcal{E}}_\ell} f_{Y|X}(y|x^{(\ell)}) \mathrm{d}y \tag{111}$$

$$= 1 - \sum_{\ell=1}^{L} P_{\mathscr{C}}(x^{(\ell)}) \left(1 - \mathrm{Q}\left(\frac{|x^{(\ell)} - \bar{x}^{(\ell)}|}{\sqrt{2\sigma^2}} - \frac{\sigma}{\sqrt{2}|x^{(\ell)} - \bar{x}^{(\ell)}|} \log\left(\frac{P_{\mathscr{C}}(\bar{x}^{(\ell)})}{P_{\mathscr{C}}(x^{(\ell)})} \right) \right) \right), \tag{112}$$

where, the equality in (112) follows from (111) due to [48, Lemma 20.14.1]. Let $\ell^\star \in \{1, 2, \ldots, L\}$ such that

$$\ell^\star \in \arg \max_{\ell \in \{1,2,\ldots,L\}} \left(1 - \mathrm{Q}\left(\frac{|x^{(\ell)} - \bar{x}^{(\ell)}|}{\sqrt{2\sigma^2}} - \frac{\sigma}{\sqrt{2}|x^{(\ell)} - \bar{x}^{(\ell)}|} \log\left(\frac{P_{\mathscr{C}}(\bar{x}^{(\ell)})}{P_{\mathscr{C}}(x^{(\ell)})} \right) \right) \right). \tag{113}$$

From (112) and (113) it follows that,

$$\epsilon \geq 1 - \sum_{\ell=1}^{L} P_{\mathscr{C}}(x^{(\ell)}) \left(1 - \mathrm{Q}\left(\frac{|x^{(\ell^\star)} - \bar{x}^{(\ell^\star)}|}{\sqrt{2\sigma^2}} - \frac{\sigma}{\sqrt{2}|x^{(\ell^\star)} - \bar{x}^{(\ell^\star)}|} \log\left(\frac{P_{\mathscr{C}}(\bar{x}^{(\ell^\star)})}{P_{\mathscr{C}}(x^{(\ell^\star)})}\right)\right)\right) \tag{114}$$

$$= 1 - \left(1 - \mathrm{Q}\left(\frac{|x^{(\ell^\star)} - \bar{x}^{(\ell^\star)}|}{\sqrt{2\sigma^2}} - \frac{\sigma}{\sqrt{2}|x^{(\ell^\star)} - \bar{x}^{(\ell^\star)}|} \log\left(\frac{P_{\mathscr{C}}(\bar{x}^{(\ell^\star)})}{P_{\mathscr{C}}(x^{(\ell^\star)})}\right)\right)\right) \tag{115}$$

$$= \mathrm{Q}\left(\frac{|x^{(\ell^\star)} - \bar{x}^{(\ell^\star)}|}{\sqrt{2\sigma^2}} - \frac{\sigma}{\sqrt{2}|x^{(\ell^\star)} - \bar{x}^{(\ell^\star)}|} \log\left(\frac{P_{\mathscr{C}}(\bar{x}^{(\ell^\star)})}{P_{\mathscr{C}}(x^{(\ell^\star)})}\right)\right). \tag{116}$$

This completes the proof □

C. Information Transmission Rate

A first upper bound on the information rate is obtained by upper bounding the number of codewords that a code might possess given the particular types $P_{\boldsymbol{u}(1)}$, $P_{\boldsymbol{u}(2)}$, …, $P_{\boldsymbol{u}(n)}$ in (17); or the average type $P_{\mathscr{C}}$ in (18). The following lemma introduces such an upper bound for the case of a homogeneous code.

Lemma 4.9. *Given a homogeneous $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16)*, the information transmission rate $R(\mathscr{C})$ in* (19) *is such that*

$$R(\mathscr{C}) = \frac{1}{n} \log\left(\frac{n!}{\prod_{\ell=1}^{L}(nP_{\mathscr{C}}(x^{(\ell)}))!}\right) \leq \log L, \tag{117}$$

where, $P_{\mathscr{C}}$ is the type defined in (18) *and L is the number of channel input symbols in* (4)*.*

Proof. The largest number of codewords of length n that can be formed using L channel input symbols is L^n. In this case, the codewords do not exhibit the type $P_{\mathscr{C}}$. Hence, from (19), in the absence of a constraint on the type $P_{\mathscr{C}}$, the largest information transmission rate is $\log L$ bits per channel use. This justifies the inequality on the right-hand side of (117).

Alternatively, given a code type $P_{\mathscr{C}}$ that satisfies (24), the number of codewords that can be constructed is given by

$$M = \binom{n}{nP_{\mathscr{C}}(x^{(1)})}\binom{n - nP_{\mathscr{C}}(x^{(1)})}{nP_{\mathscr{C}}(x^{(2)})} \cdots \binom{n - \sum_{\ell=1}^{L-1} nP_{\mathscr{C}}(x^{(\ell)})}{nP_{\mathscr{C}}(x^{(L)})} = \frac{n!}{\prod_{\ell=1}^{L}(nP_{\mathscr{C}}(x^{(\ell)}))!}. \tag{118}$$

Therefore, the information rate $R(\mathscr{C})$ in (19) satisfies

$$R(\mathscr{C}) = \frac{1}{n} \log\left(\frac{n!}{\prod_{\ell=1}^{L}(nP_{\mathscr{C}}(x^{(\ell)}))!}\right), \tag{119}$$

which completes the proof. □

Equality in (117) holds when the type induced by the codewords of the code $\mathscr{C}$ is uniform, i.e., for all $x \in \mathcal{X}$ it holds that

$$P_{\mathscr{C}}(x) = \frac{1}{L}. \tag{120}$$

Lemma 4.9 can be written in terms of the entropy of the type $P_{\mathscr{C}}$ as shown in the following result.

Lemma 4.10. *Given a homogeneous $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16)*, the information transmission rate $R(\mathscr{C})$ in* (19) *is such that*

$$R(\mathscr{C}) \leq H\left(P_{\mathscr{C}}\right) + \frac{1}{n^2}\left(\frac{1}{12} - \sum_{\ell=1}^{L} \frac{1}{12P_{\mathscr{C}}(x^{(\ell)}) + 1}\right) + \frac{1}{n}\left(\log\left(\sqrt{2\pi}\right) - \sum_{\ell=1}^{L} \log\sqrt{2\pi P_{\mathscr{C}}(x^{(\ell)})}\right) - \frac{\log n}{n}\left(\frac{L-1}{2}\right), \tag{121}$$

where, $P_{\mathscr{C}}$ is the type defined in (18) *and L is the number of channel input symbols in* (4)*.*

Proof. From (117), the following holds

$$R(\mathscr{C}) \leq \frac{1}{n}\log(n!) - \frac{1}{n}\sum_{\ell=1}^{L} \log\left(\left(nP_{\mathscr{C}}(x^{(\ell)})\right)!\right), \tag{122}$$

and using the Stirling's approximation [49] on the factorial terms yields

$$(nP_{\mathscr{C}}(x^{(\ell)}))! \geqslant \sqrt{2\pi}\left(nP_{\mathscr{C}}(x^{(\ell)})\right)^{nP_{\mathscr{C}}(x^{(\ell)})+\frac{1}{2}} \exp\left(-nP_{\mathscr{C}}(x^{(\ell)}) + \frac{1}{12nP_{\mathscr{C}}(x^{(\ell)})+1}\right), \text{ and} \tag{123}$$

$$n! \leqslant \sqrt{2\pi} n^{n+\frac{1}{2}} \exp\left(-n + \frac{1}{12n}\right). \tag{124}$$

From (123) and (124), it follows that,

$$\log\left((nP_{\mathscr{C}}(x^{(\ell)}))!\right) \geq \log\left(\sqrt{2\pi}\right) + \left(nP_{\mathscr{C}}(x^{(\ell)}) + \frac{1}{2}\right)\log(nP_{\mathscr{C}}(x^{(\ell)})) - nP_{\mathscr{C}}(x^{(\ell)}) + \frac{1}{12nP_{\mathscr{C}}(x^{(\ell)})+1} \tag{125}$$

$$\begin{aligned} &= \log\left(\sqrt{2\pi}\right) + nP_{\mathscr{C}}(x^{(\ell)})\log(P_{\mathscr{C}}(x^{(\ell)})) + \frac{1}{2}\log(P_{\mathscr{C}}(x^{(\ell)})) \\ &\quad + \left(nP_{\mathscr{C}}(x^{(\ell)}) + \frac{1}{2}\right)\log(n) - nP_{\mathscr{C}}(x^{(\ell)}) + \frac{1}{12nP_{\mathscr{C}}(x^{(\ell)})+1} \text{ and }, \end{aligned} \tag{126}$$

$$\begin{aligned} \log(n!) &\leq \log\left(\sqrt{2\pi}\right) + \left(n+\frac{1}{2}\right)\log(n) - n + \frac{1}{12n} \\ &= n\log(n) - n + \frac{1}{12n} + \frac{1}{2}\log(2\pi n). \end{aligned} \tag{127}$$

The sum in (122) satisfies,

$$\begin{aligned} \sum_{\ell=1}^{L}\log\left(\left(nP_{\mathscr{C}}(x^{(\ell)})\right)!\right) \geq L\log\left(\sqrt{2\pi}\right) - nH\left(P_{\mathscr{C}}\right) + \frac{1}{2}\sum_{\ell=1}^{L}\log(P_{\mathscr{C}}(x^{(\ell)})) + n\log(n) \\ + \frac{L}{2}\log(n) - n + \sum_{\ell=1}^{L}\frac{1}{12nP_{\mathscr{C}}(x^{(\ell)})+1}. \end{aligned} \tag{128}$$

Using (127) and (128) in (122) yields,

$$\begin{aligned} R(\mathscr{C}) &\leq \log(n) - 1 + \frac{1}{12n^2} + \frac{1}{2n}\log(2\pi n) - \frac{L}{n}\log\left(\sqrt{2\pi}\right) + H\left(P_{\mathscr{C}}\right) - \frac{1}{2n}\sum_{\ell=1}^{L}\log(P_{\mathscr{C}}(x^{(\ell)})) - \log(n) \\ &\quad - \frac{L}{2n}\log(n) + 1 - \frac{1}{n}\sum_{\ell=1}^{L}\frac{1}{12nP_{\mathscr{C}}(x^{(\ell)})+1} \end{aligned} \tag{129}$$

$$\leq H\left(P_{\mathscr{C}}\right) + \frac{1}{n^2}\left(\frac{1}{12} - \sum_{\ell=1}^{L}\frac{1}{12P_{\mathscr{C}}(x^{(\ell)})+1}\right) + \frac{1}{2n}\left(\log(2\pi n) - \sum_{\ell=1}^{L}\log(2\pi n P_{\mathscr{C}}(x^{(\ell)}))\right) \tag{130}$$

$$= H\left(P_{\mathscr{C}}\right) + \frac{1}{n^2}\left(\frac{1}{12} - \sum_{\ell=1}^{L}\frac{1}{12P_{\mathscr{C}}(x^{(\ell)})+1}\right) + \frac{1}{n}\left(\log\left(\sqrt{2\pi}\right) - \sum_{\ell=1}^{L}\log\sqrt{2\pi P_{\mathscr{C}}(x^{(\ell)})}\right) - \frac{\log n}{n}\left(\frac{L-1}{2}\right), \tag{131}$$

which completes the proof. □

Note that all terms in (131), except the entropy $H\left(P_{\mathscr{C}}\right)$, vanish with the block-length n. This implies that the information rate is essentially constrained by the entropy of the channel input symbols. In particular, note that $H\left(P_{\mathscr{C}}\right) \leqslant \log L$.

D. Information and Energy Trade-Off

The bounds provided in subsections 4-A, 4-B and 4-C together define the converse information-energy region for $(n, M, \epsilon, B, \delta)$-codes for the random transformation in (9) of the form in (16) for a given set of channel inputs. These bounds prove to be a function of the type $P_{\mathscr{C}}$ and quantify the trade-offs between the information rate R, the energy rate B, the DEP ϵ and the EOP δ for these codes. The following theorem illustrates the dependence between R, B, ϵ, and δ for the class of homogeneous codes.

Theorem 4.11. *Given a homogeneous $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in* (9) *of the form in* (16)*, the following hold:*

$$\delta = \mathbb{1}_{\{e_{\mathscr{C}} < B\}}; \tag{132}$$

$$R(\mathscr{C}) = \frac{1}{n} \log \left(\frac{n!}{\prod_{\ell=1}^{L} (n P_{\mathscr{C}}(x^{(\ell)}))!} \right); \text{ and} \tag{133}$$

$$\epsilon \geq \mathrm{Q}\left(\frac{|x^{(\ell^\star)} - \bar{x}^{(\ell^\star)}|}{\sqrt{2\sigma^2}} - \frac{\sigma}{\sqrt{2}|x^{(\ell^\star)} - \bar{x}^{(\ell^\star)}|} \log \left(\frac{P_{\mathscr{C}}(\bar{x}^{(\ell^\star)})}{P_{\mathscr{C}}(x^{(\ell^\star)})} \right) \right), \tag{134}$$

where, for all $\ell \in \{1, 2, \ldots, L\}$, the complex $\bar{x}^{(\ell)} \in \mathcal{X}$ is in (105) *with $\mathcal{X}$ in* (3)*; the function* Q *is defined in* (2)*; the real σ^2 is in* (12)*; $\ell^\star \in \{1, 2, \ldots, L\}$ is defined in* (107)*; the real $e_{\mathscr{C}} \in [0, \infty)$ is in* (30)*; and $P_{\mathscr{C}}$ is the type defined in* (18)*.*

Proof. The result follows from Lemma 4.2, Lemma 4.9, and Lemma 4.8. □

Theorem 4.11 illustrates the trade-offs between all the parameters of a homogeneous $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in (9). The upper bound on the information rate in (133) and the lower bound on the DEP in (134) depend on the type $P_{\mathscr{C}}$. Though it is not made obvious from (132), the lower bound on the EOP also depends on the type $P_{\mathscr{C}}$ via (30). The information rate is essentially upper bounded by the entropy of $P_{\mathscr{C}}$. Hence, codes whose codewords are such that every channel input symbol is used the same number of times are less constrained in terms of the information rate. This is the case in which $P_{\mathscr{C}}$ is a uniform distribution. Alternatively, using a uniform type $P_{\mathscr{C}}$ might dramatically constrain the energy transmission rate. For instance, assume that the set of channel input symbols is such that for at least one pair $(x^{(1)}, x^{(2)}) \in \mathcal{X}^2$, with $\mathcal{X}$ in (3), it holds that $\left|x^{(1)}\right| < \left|x^{(2)}\right|$. Then, from (30), using the symbol $x^{(1)}$ equally often as $x^{(2)}$ certainly constraints the energy $e_{\mathscr{C}}$ and hence, the energy rate B. Codes that might potentially exhibit the largest energy rates are those in which the symbols that have the largest magnitude are used more often. This clearly deviates from the uniform distribution and thus, constraints the information rate R.

5. Information-Energy Achievable Region

This section characterizes an achievable information-energy region for SIET. An achievable region defines tuples of information rate, energy rate, DEP and EOP that can be achieved by some code. As the name suggests, for a tuple to be achievable, a code needs to exist that can achieve the tuple. Therefore, the first step in the characterization of an achievable region is the construction of a code which is accomplished in Section 5-A. This is followed by characterizing the bounds on the various parameters that can be achieved by the constructed code in Section 5-B.

A. Code Construction

The process of characterizing an achievable information-energy capacity region for SIET in the finite block-length regime with finite set of channel input symbols begins with the construction of an (n, M)-code $\mathscr{C}$ of the form in (16). The construction of the code begins with the construction of the channel input symbols. The set of channel input symbols is a modulation constellation represented by a finite subset of $\mathbb{C}$. Consider a set of channel input symbols formed by C *layers*, with $C \in \mathbb{N}$. A layer is a subset of symbols in $\mathbb{C}$ that have the same magnitude. For all $c \in \{1, 2, \ldots, C\}$, denote by $L_c \in \mathbb{N}$ the number of symbols in the c^{th} layer and let $A_c \in \mathbb{R}^+$ and $\alpha_c \in [0, 2\pi]$ be the amplitude and phase shift of the symbols in layer c. Denote such a layer by $\mathcal{U}(A_c, L_c, \alpha_c)$. That is,

$$\mathcal{U}(A_c, L_c, \alpha_c) \triangleq \left\{ x_c^{(\ell)} = A_c \exp\left(\mathrm{i} \left(\frac{2\pi}{L_c} \ell + \alpha_c \right) \right) \subseteq \mathbb{C} : \ell \in \{0, 1, 2, \ldots, (L_c - 1)\} \right\}, \tag{135a}$$

where i is the complex unit. Using this notation, the set of channel input symbols can be described by the following set

$$\mathcal{X} = \bigcup_{c=1}^{C} \mathcal{U}(A_c, L_c, \alpha_c). \tag{135b}$$

The vector of the amplitudes in (135b) is denoted by

$$\boldsymbol{A} = (A_1, A_2, \ldots, A_C)^{\mathsf{T}}; \tag{135c}$$

the vector of the number of symbols in each layer in (135b) is denoted by

$$\boldsymbol{L} = (L_1, L_2, \ldots, L_C)^{\mathsf{T}}; \tag{135d}$$

and the vector of the phase shifts of the symbols in each layer in (135b) is denoted by

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_C)^{\mathsf{T}}. \tag{135e}$$

The total number of symbols L in (4) for $\mathcal{X}$ in (135b) is

$$L = \sum_{c=1}^{C} L_c. \tag{135f}$$

Without any loss of generality, assume that

$$A_1 > A_2 > \ldots > A_C. \tag{135g}$$

For all $c \in \{1, 2, \ldots, C\}$, the symbols in $\mathcal{U}(A_c, L_c, \alpha_c)$ in (135a), are equally spaced along a circle of radius A_c (see Fig. 1). The set of channel input symbols $\mathcal{X}$ in (135b) is thus made up of points uniformly distributed along C concentric circles.

The construction of the (n, M)-code $\mathscr{C}$ is accomplished as follows. For all $c \in \{1, 2, \ldots, C\}$, let p_c be the frequency with which symbols of the c^{th} layer appear in $\mathscr{C}$. The resulting probability vector is denoted by

$$\boldsymbol{p} = (p_1, p_2, \ldots, p_C)^{\mathsf{T}}. \tag{135h}$$

For transmitting message index $i \in \{1, 2, \ldots, M\}$, the transmitter uses the codeword

$$\boldsymbol{u}(i) = (u_1(i), u_2(i), \ldots, u_n(i)) \in \mathcal{X}^n, \tag{135i}$$

with $\mathcal{X}$ in (135b). Using the probability vector $\boldsymbol{p}$ in (135h) which determines the type $P_{\mathscr{C}}$ in (18), and the set of symbols $\mathcal{X}$ in (135b), the set of codewords $\{\boldsymbol{u}(1), \boldsymbol{u}(2), \ldots, \boldsymbol{u}(M)\}$ is completely defined.

For all $c \in \{1, 2, \ldots, C\}$, it holds that,

$$p_c = \frac{1}{Mn} \sum_{\ell=1}^{L_c} \sum_{i=1}^{M} \sum_{m=1}^{n} \mathbb{1}_{\{u_m(i) = x_c^{(\ell)}\}} \tag{135j}$$

$$= \frac{1}{M} \sum_{\ell=1}^{L_c} \sum_{i=1}^{M} P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right), \tag{135k}$$

where, for all $i \in \{1, 2, \ldots, M\}$ and all $m \in \{1, 2, \ldots, n\}$, the complex $u_m(i)$ in (135i) is the m^{th} channel input symbol of the codeword $\boldsymbol{u}(i)$ and $x_c^{(\ell)}$ is in (135a). The equality in (135k) follows from (17). The symbols within a layer are used with the same frequency in $\mathscr{C}$. Hence, for all $c \in \{1, 2, \ldots, C\}$ and for all $\ell \in \{1, 2, \ldots, L_c\}$, the frequency with which the symbol $x_c^{(\ell)}$ appears in $\mathscr{C}$ is given by

$$P_{\mathscr{C}}(x_c^{(\ell)}) \triangleq \frac{1}{Mn} \sum_{i=1}^{M} \sum_{m=1}^{n} \mathbb{1}_{\{u_m(i) = x_c^{(\ell)}\}}, \tag{135l}$$

$$= \frac{p_c}{L_c}. \tag{135m}$$

The decoding region for codeword $\boldsymbol{u}(i)$ is denoted by $\mathcal{D}_i$ and that for the symbol $u_m(i)$ is denoted by $\mathcal{D}_{i,m}$ such that the following constraint on $\mathcal{D}_i$ holds:

$$\mathcal{D}_i = \mathcal{D}_{i,1} \times \mathcal{D}_{i,2} \times \ldots \times \mathcal{D}_{i,n}, \tag{135n}$$

where, for all $i \in \{1, 2, \ldots, M\}$, $c \in \{1, 2, \ldots, C\}$, $t \in \{1, 2, \ldots, n\}$, and $\ell \in \{1, 2, \ldots, L_c\}$, when $u_m(i) = x_c^{(\ell)}$, then, $\mathcal{D}_{i,m} = \mathcal{G}_c^{(\ell)}$.

The decoding set $\mathcal{G}_c^{(\ell)} \subseteq \mathbb{C}$ associated with symbol $x_c^{(\ell)}$ is a circle of radius $r_c \in \mathbb{R}^+$ centered at $x_c^{(\ell)}$. That is,

$$\mathcal{G}_c^{(\ell)} = \left\{ y \in \mathbb{C} : \left| y - x_c^{(\ell)} \right|^2 \leq r_c^2 \right\}, \tag{135o}$$

$$= \left\{ y \in \mathbb{C} : \left(\Re(y) - \Re(x_c^{(\ell)}) \right)^2 + \left(\Im(y) - \Im(x_c^{(\ell)}) \right)^2 \leq r_c^2 \right\}. \tag{135p}$$

The radii r_1, r_2, …, r_C are chosen such that the decoding sets $\mathcal{G}_c^{(\ell)}$ are mutually disjoint. To ensure that the decoding regions are mutually disjoint, for all $c \in \{1, 2, \ldots, C\}$, the amplitudes A_c in (135a) are chosen to satisfy the following:

$$A_{c+1} - A_c \geq r_{c+1} + r_c. \tag{135q}$$

The vector of these radii is denoted by

$$\boldsymbol{r} = (r_1, r_2, \ldots, r_C)^{\mathsf{T}}. \tag{135r}$$

This defines a family of (n, M)-codes denoted by

$$\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right), \tag{135s}$$

with the number of layers C in (135b), $\boldsymbol{A}$ in (135c), $\boldsymbol{L}$ in (135d), $\boldsymbol{\alpha}$ in (135e) $\boldsymbol{p}$ in (135h) and $\boldsymbol{r}$ in (135r). The following section characterizes the achievable region for codes in the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$.

B. Achievability Bounds

The following lemma provides a lower bound on the DEP ϵ of codes from the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in (135s).

Lemma 5.1. *Consider an (n, M)-code $\mathscr{C}$ of the form in* (135) *from the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in* (135s)*. The code $\mathscr{C}$ is an (n, M, ϵ)-code if the parameters $r_1, r_2, \ldots, r_C$ in* (135p) *satisfy the following:*

$$\epsilon \geq 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \left(1 - e^{-\frac{r_c^2}{\sigma^2}}\right)^{n \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{136}$$

where, the type $P_{\boldsymbol{u}(i)}$ is defined in (17)*, the real σ^2 is defined in* (12)*, and $x_c^{(\ell)} \in \mathcal{U}(A_c, L_c, \alpha_c)$, with $\mathcal{U}(A_c, L_c, \alpha_c)$ in* (135a)*.*

Proof. From (21) and (22), the average DEP of code $\mathscr{C}$ is given by

$$\gamma(\mathscr{C}) = 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{\mathcal{D}_i} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{u}(i)) \mathrm{d}\boldsymbol{y} \tag{137}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{\mathcal{D}_{i,1} \times \mathcal{D}_{i,2} \times \ldots \times \mathcal{D}_{i,n}} \prod_{m=1}^{n} f_{Y|X}\left(y|u_m(i)\right) \mathrm{d}y \tag{138}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{m=1}^{n} \int_{\mathcal{D}_{i,m}} f_{Y|X}\left(y|u_m(i)\right) \mathrm{d}y \tag{139}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \prod_{\ell=1}^{L_c} \left(\int_{\mathcal{G}_c^{(\ell)}} f_{Y|X}(y|x_c^{(\ell)}) \mathrm{d}y\right)^{n P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{140}$$

where, the equality in (138) follows due to (10) and (135n), and (139) follows due to Fubini's theorem. Using (12) in (140) yields,

$$\gamma\left(\mathscr{C}\right) = 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \prod_{\ell=1}^{L_c} \left(\int_{\mathcal{G}_c^{(\ell)}} \frac{1}{\pi\sigma^2} \exp\left(-\frac{(\Re(y) - \Re(x_c^{(\ell)}))^2 + (\Im(y) - \Im(x_c^{(\ell)}))^2}{\sigma^2}\right) \mathrm{d}y\right)^{n P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}. \tag{141}$$

Evaluating the integral term in (141) for all $\ell \in \{1, 2, \ldots, L_c\}$ yields,

$$\int_{\mathcal{G}_c^{(\ell)}} \frac{1}{\pi\sigma^2} \exp\left(-\frac{(\Re(y) - \Re(x_c^{(\ell)}))^2 + (\Im(y) - \Im(x_c^{(\ell)}))^2}{\sigma^2}\right) \mathrm{d}y$$

$$= \int_{\Im(x_c^{(\ell)}) - r_c}^{\Im(x_c^{(\ell)}) + r_c} \int_{\Re(x_c^{(\ell)}) - \sqrt{r_c^2 - (v - \Im(x_c^{(\ell)}))^2}}^{\Re(x_c^{(\ell)}) + \sqrt{r_c^2 - (v - \Im(x_c^{(\ell)}))^2}} \frac{1}{\pi\sigma^2} e^{-\frac{(u - \Re(x_c^{(\ell)}))^2 + (v - \Im(x_c^{(\ell)}))^2}{\sigma^2}} \mathrm{d}u \mathrm{d}v, \tag{142}$$

$$= \int_{-r_c}^{r_c} \int_{-\sqrt{r_c^2 - v^2}}^{\sqrt{r_c^2 - v^2}} \frac{1}{\pi\sigma^2} e^{-\frac{u^2 + v^2}{\sigma^2}} \mathrm{d}u \mathrm{d}v, \tag{143}$$

$$= \int_0^{\pi} \int_0^{\frac{r_c}{\sigma}} \frac{1}{2\pi} e^{-\zeta^2} \zeta \mathrm{d}\zeta \mathrm{d}\eta, \tag{144}$$

$$= (1 - e^{-\frac{r_c^2}{\sigma^2}}). \tag{145}$$

The equality in (144) is obtained from the change of variables $u = \sigma\zeta\cos\eta, v = \sigma\zeta\sin\eta$. Plugging (145) in (141) yields,

$$\gamma\left(\mathscr{C}\right) = 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \prod_{\ell=1}^{L_c} \left(1 - e^{-\frac{r_c^2}{\sigma^2}}\right)^{n P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{146}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \left(1 - e^{-\frac{r_c^2}{\sigma^2}}\right)^{n \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}. \tag{147}$$

From (23), for $\mathscr{C}$ to be an (n, M, ϵ)-code, the following must hold:

$$\gamma\left(\mathscr{C}\right) \leq \epsilon. \tag{148}$$

This implies that,

$$\epsilon \geq 1 - \frac{1}{M}\sum_{i=1}^{M}\prod_{c=1}^{C}\left(1 - e^{-\frac{r_c^2}{\sigma^2}}\right)^{n\sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{149}$$

which is the desired result. □

From Definition 3.3 and (135m), it follows that, for a homogeneous code $\mathscr{C}$ of the form in (135) from the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in (135s), it holds that,

$$P_{\boldsymbol{u}(i)}(x_c^{(\ell)}) = \frac{p_c}{L_c}. \tag{150}$$

The following result for homogeneous codes follows from (150) and Lemma 5.1.

Corollary 5.2. *Consider a homogeneous (n, M)-code $\mathscr{C}$ of the form in* (135) *from the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in* (135s)*. The code $\mathscr{C}$ is an (n, M, ϵ)-code if the parameters $r_1, r_2, \ldots, r_C$ in* (135p) *satisfy the following:*

$$\epsilon \geq 1 - \prod_{c=1}^{C}\left(1 - e^{-\frac{r_c^2}{\sigma^2}}\right)^{np_c}, \tag{151}$$

where, the real σ^2 is defined in (12)*, and $x_c^{(\ell)} \in \mathcal{U}(A_c, L_c, \alpha_c)$, with $\mathcal{U}(A_c, L_c, \alpha_c)$ in* (135a)*.*

For the special case where for all $c \in \{1, 2, \ldots, C\}$, $r_c = r$ in (135r), the following lemma provides a lower bound on the radius r.

Lemma 5.3. *Consider an (n, M)-code $\mathscr{C}$ of the form in* (135) *from the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in* (135s) *such that, for all $c \in \{1, 2, \ldots, C\}$, $r_c = r$ in* (135p)*. The code $\mathscr{C}$ is an (n, M, ϵ)-code if the parameter r satisfies:*

$$r \geq \sqrt{\sigma^2 \log\left(\frac{1}{1 - (1-\epsilon)^{\frac{1}{n}}}\right)}, \tag{152}$$

where, the real σ^2 is defined in (12)*.*

Proof. If the parameters r_c in (135p) are such that, for all $c \in \{1, 2, \ldots, C\}$, $r_c = r$, the average DEP in (147) is given by:

$$\gamma\left(\mathscr{C}\right) = 1 - \frac{1}{M}\sum_{i=1}^{M}\prod_{c=1}^{C}\left(1 - e^{-\frac{r^2}{\sigma^2}}\right)^{n\sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{153}$$

$$= 1 - \frac{1}{M}\sum_{i=1}^{M}\left(1 - e^{-\frac{r^2}{\sigma^2}}\right)^{n\sum_{c=1}^{C}\sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{154}$$

$$= 1 - \frac{1}{M}\sum_{i=1}^{M}\left(1 - e^{-\frac{r^2}{\sigma^2}}\right)^{n}, \tag{155}$$

$$= 1 - \left(1 - e^{-\frac{r^2}{\sigma^2}}\right)^{n}. \tag{156}$$

From (23), for $\mathscr{C}$ to be an (n, M, ϵ)-code, the following must hold:

$$\epsilon \geq 1 - \left(1 - e^{-\frac{r^2}{\sigma^2}}\right)^{n}. \tag{157}$$

This implies that

$$r \geq \sqrt{\sigma^2 \log\left(\frac{1}{1 - (1-\epsilon)^{\frac{1}{n}}}\right)}. \tag{158}$$

This completes the proof. □

The information rate achievable by a code is a function of the number of channel input symbols L in (135f) which in turn is a function of the number of symbols in each layer of the constructed set of channel inputs $\mathcal{X}$ in (135b). The following lemma provides an upper bound on the number of symbols in a layer $c \in \{1, 2, \ldots, C\}$ denoted by L_c.

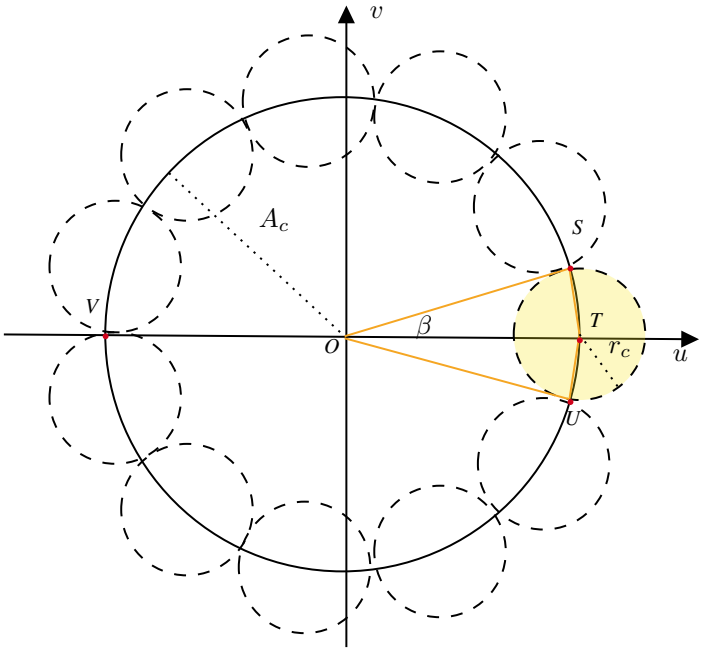

Fig. 1: Graphical representation of the symbols in layer c defined in (135a)

.

Lemma 5.4. *Consider an (n, M)-code $\mathscr{C}$ of the form in* (135) *from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in* (135s)*. Then, for all $c \in \{1, 2, \ldots, C\}$, the number of symbols in layer c of $\mathcal{X}$ is given by*

$$L_c \leq \left\lfloor \frac{\pi}{2 \arcsin \frac{r_c}{2A_c}} \right\rfloor, \tag{159}$$

where, r_c is the radius of the decoding regions $\mathcal{G}_c^{(1)}, \mathcal{G}_c^{(2)}, \ldots, \mathcal{G}_c^{(L_c)}$ in (135p) *and A_c is the amplitude in* (135a)*.*

Proof. For an (n, M)-code $\mathscr{C}$ in the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$, for all $c \in \{1, 2, \ldots, C\}$, the radius r_c of the decoding regions in (135r) and the amplitude A_c in (135c) determine the number of symbols L_c that can be accommodated in the layer c.

Layer c of the form in (135a) is illustrated in Fig. 1. Symbols in layer c are distributed uniformly along the circle of radius A_c centered at the origin O. The maximum number of symbols that can be accommodated in layer c is equal to the number of non-overlapping circles of radius r_c corresponding to the decoding regions defined in (135p) that can be placed along the circumference of the circle of radius A_c. From Fig. 1, a circle of radius r_c centered at a symbol in layer c subtends angle $\angle SOU = \beta$ at O. Therefore, the maximum number of symbols L_c that can be accommodated along the circle of radius A_c is given by

$$L_c \leq \left\lfloor \frac{2\pi}{\beta} \right\rfloor. \tag{160}$$

The following lemma determines the value of the angle β in Fig. 1.

Lemma 5.5. *Consider a circle of radius A_c centered at O and a circle of radius r_c centered on the circumference of the first circle as shown in Fig. 1. The two circles intersect at points S and U. Then, the angle $\angle SOU$ is given by the following:*

$$\angle SOU = \beta = 4 \arcsin \frac{r_c}{2A_c}. \tag{161}$$

Proof. In Fig. 1, consider the circle of radius r_c centered at T and the larger circle of radius A_c centered at the origin O. The circles intersect at points S and U. The angle subtended by the major arc $\overset{\frown}{SU}$ at O is the reflex angle $2\pi - \beta$. Since the angle subtended by an arc of a circle at its centre is two times the angle that it subtends anywhere on the circumference, it holds that

$$2\pi - \beta = 2\angle STU, \tag{162}$$

which implies that

$$\angle STU = \frac{2\pi - \beta}{2}. \tag{163}$$

The line segment TO bisects angles $\angle STU$ and $\angle SOU$. Therefore, the following hold:

$$\angle STO = \frac{2\pi - \beta}{4}, \tag{164}$$

$$\angle SOT = \frac{\beta}{2}. \tag{165}$$

From the triangle $\triangle SOT$, it holds that:

$$\frac{\sin\left(\angle SOT\right)}{ST} = \frac{\sin\left(\angle STO\right)}{SO}. \tag{166}$$

This implies that,

$$\frac{\sin\left(\frac{\beta}{2}\right)}{r_c} = \frac{\sin\left(\frac{2\pi-\beta}{4}\right)}{A_c}, \tag{167}$$

$$= \frac{1}{A_c}\sin\left(\frac{\pi}{2} - \frac{\beta}{4}\right), \tag{168}$$

$$= \frac{1}{A_c}\cos\left(\frac{\beta}{4}\right). \tag{169}$$

From (169), it follows that,

$$\frac{2}{r_c}\sin\left(\frac{\beta}{4}\right)\cos\left(\frac{\beta}{4}\right) = \frac{1}{A_c}\cos\left(\frac{\beta}{4}\right), \tag{170}$$

which implies that,

$$\sin\left(\frac{\beta}{4}\right) = \frac{r_c}{2A_c}, \quad \text{and} \tag{171}$$

$$\beta = 4\arcsin\frac{r_c}{2A_c}. \tag{172}$$

This completes the proof. □

Substituting the value of β from (161) in (160), the number of symbols in layer c of the set $\mathcal{X}$ is given by

$$L_c \leq \left\lfloor \frac{\pi}{2\arcsin\frac{r_c}{2A_c}} \right\rfloor. \tag{173}$$

This completes the proof. □

The information transmission rate $R\left(\mathscr{C}\right)$ is a function of the number of symbols L in (135f). Therefore, using Lemma 5.4, the achievable bound on the information transmission rate $R\left(\mathscr{C}\right)$ for code $\mathscr{C}$ from the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in (135s) can now be calculated. The following lemma provides this result.

Lemma 5.6. *For a homogeneous (n, M)-code $\mathscr{C}$ of the form in* (135) *from the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in* (135s)*, the information transmission rate $R\left(\mathscr{C}\right)$ satisfies the following:*

$$R\left(\mathscr{C}\right) = \frac{1}{n}\log\left(\frac{n!}{\prod_{c=1}^{C}\left(\left(n\frac{p_c}{L_c}\right)!\right)^{L_c}}\right) \leq \log\sum_{c=1}^{C}\left\lfloor\frac{\pi}{2\arcsin\frac{r_c}{2A_c}}\right\rfloor, \tag{174}$$

where, for all $c \in \{1, 2, \ldots, C\}$, the radius r_c is in (135p)*, and the amplitude A_c is in* (135a)*.*

Proof. The largest number of codewords of length n that can be formed with L different channel input symbols is L^n. Hence, from (19), it follows that

$$R\left(\mathscr{C}\right) = \frac{\log M}{n} \tag{175}$$

$$\leq \frac{\log L^n}{n} \tag{176}$$

$$= \log L \tag{177}$$

$$\leq \log\sum_{c=1}^{C}\left\lfloor\frac{\pi}{2\arcsin\frac{r_c}{2A_c}}\right\rfloor, \tag{178}$$

where, the inequality in (178) follows from Lemma 5.4 and (135f).

The bound in (178) can only be achieved by a code for which the type $P_{\mathscr{C}}$ is uniform, *i.e.*, for all $x \in \mathcal{X}$ with $\mathcal{X}$ in (135b), it holds that,

$$P_{\mathscr{C}}(x) = \frac{1}{L}. \tag{179}$$

Given the type $P_{\mathscr{C}}$ that satisfies (24), the number of codewords that can be constructed is given by

$$\begin{aligned}
M = &\binom{n}{nP_{\mathscr{C}}(x_1^{(1)})}\binom{n-nP_{\mathscr{C}}(x_1^{(1)})}{nP_{\mathscr{C}}(x_1^{(2)})}\cdots\binom{n-\sum_{\ell=1}^{L_1-1}nP_{\mathscr{C}}(x_1^{(\ell)})}{nP_{\mathscr{C}}(x_1^{(L_1)})}\times\binom{n-\sum_{\ell=1}^{L_1}nP_{\mathscr{C}}(x_1^{(\ell)})}{nP_{\mathscr{C}}(x_2^{(1)})}\\
&\binom{n-\sum_{\ell=1}^{L_1}nP_{\mathscr{C}}(x_1^{(\ell)})-nP_{\mathscr{C}}(x_2^{(1)})}{nP_{\mathscr{C}}(x_2^{(2)})}\cdots\binom{n-\sum_{\ell=1}^{L_1}nP_{\mathscr{C}}(x_1^{(\ell)})-\sum_{\ell=1}^{L_2-1}nP_{\mathscr{C}}(x_2^{(\ell)})}{nP_{\mathscr{C}}(x_2^{(L_2)})}\times\ldots\times\binom{n-\sum_{c=1}^{C-1}\sum_{\ell=1}^{L_c}nP_{\mathscr{C}}(x_c^{(\ell)})}{nP_{\mathscr{C}}(x_C^{(1)})}\\
&\binom{n-\sum_{c=1}^{C-1}\sum_{\ell=1}^{L_c}nP_{\mathscr{C}}(x_c^{(\ell)})-nP_{\mathscr{C}}(x_C^{(1)})}{nP_{\mathscr{C}}(x_C^{(2)})}\cdots\binom{n-\sum_{c=1}^{C-1}\sum_{\ell=1}^{L_c}nP_{\mathscr{C}}(x_c^{(\ell)})-\sum_{\ell=1}^{L_C-1}nP_{\mathscr{C}}(x_C^{(\ell)})}{nP_{\mathscr{C}}(x_C^{(L_C)})}
\end{aligned} \tag{180}$$

$$\begin{aligned}
= &\binom{n}{n\frac{p_1}{L_1}}\binom{n-n\frac{p_1}{L_1}}{n\frac{p_1}{L_1}}\cdots\binom{n-\sum_{\ell=1}^{L_1-1}n\frac{p_1}{L_1}}{n\frac{p_1}{L_1}}\times\binom{n-np_1}{n\frac{p_2}{L_2}}\binom{n-np_1-n\frac{p_2}{L_2}}{n\frac{p_2}{L_2}}\cdots\binom{n-np_1-\sum_{\ell=1}^{L_2-1}n\frac{p_2}{L_2}}{n\frac{p_2}{L_2}}\times\ldots\times\\
&\binom{n-\sum_{c=1}^{C-1}np_c}{n\frac{p_C}{L_C}}\binom{n-\sum_{c=1}^{C-1}np_c-n\frac{p_C}{L_C}}{n\frac{p_C}{L_C}}\cdots\binom{n-\sum_{c=1}^{C-1}np_c-\sum_{\ell=1}^{L_C-1}n\frac{p_C}{L_C}}{n\frac{p_C}{L_C}}
\end{aligned} \tag{181}$$

$$= \frac{n!}{\prod_{c=1}^{C}\left(\left(n\frac{p_c}{L_c}\right)!\right)^{L_c}} \tag{182}$$

Therefore, the information transmission rate $R(\mathscr{C})$ is given by

$$R(\mathscr{C}) = \frac{\log M}{n} = \frac{1}{n}\log\left(\frac{n!}{\prod_{c=1}^{C}\left(\left(n\frac{p_c}{L_c}\right)!\right)^{L_c}}\right). \tag{183}$$

The result follows from (178) and (183). □

Given the number of layers C in (135b), the vector of amplitudes $\boldsymbol{A}$ in (135c), the vector of the number of symbols in each layer $\boldsymbol{L}$ in (135d), the vector of phase shifts $\boldsymbol{\alpha}$ in (135e), the probability vector $\boldsymbol{p}$ in (135h) and the radius vector $\boldsymbol{r}$ in (135r), infinitely many sets of channel input symbols $\mathcal{X}$ in (135b) can be generated by rotating $\mathcal{X}$. Rotating $\mathcal{X}$ implies changing the phase shift of all the symbols in $\mathcal{X}$ by the same angle. It is easy to see that the information transmission rate $R(\mathscr{C})$ of the code $\mathscr{C}$ is independent of such rotations since it only depends on the number of symbols L in (135f) and the probability vector $\boldsymbol{p}$. However, the impact of such rotations on the DEP $\gamma(\mathscr{C})$ and the EOP $\theta(\mathscr{C}, B)$ of $\mathscr{C}$ is not immediately obvious. The following lemma proves that the DEP and the EOP also remain unchanged for any rotation of the set $\mathcal{X}$.

Lemma 5.7. *Consider two (n, M)-codes $\mathscr{C}$ and $\mathscr{C}'$, both of the form in* (135) *from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in* (135s)*. For all $c \in \{1, 2, \ldots, C\}$ and all $\ell \in \{1, 2, \ldots, L_c\}$, denote the symbols in layer c in* (135a) *of the code $\mathscr{C}$ by $x_c^{(\ell)}$ and those of $\mathscr{C}'$ by $\hat{x}_c^{(\ell)}$. For all $c \in \{1, 2, \ldots, C\}$, all $\ell \in \{1, 2, \ldots, L_c\}$, and $\omega \in [0, 2\pi]$, it holds that,*

$$\hat{x}_c^{(\ell)} = e^{\mathrm{i}\omega}x_c^{(\ell)}. \tag{184}$$

All other parameters of the codes $\mathscr{C}$ and $\mathscr{C}'$ are identical. Then, the average DEP γ in (22) *and the average EOP θ in* (37) *of $\mathscr{C}$ and $\mathscr{C}'$ satisfy the following:*

$$\gamma(\mathscr{C}) = \gamma(\mathscr{C}'), \text{and} \tag{185}$$

$$\theta(\mathscr{C}, B) = \theta(\mathscr{C}', B). \tag{186}$$

Proof. From (141), the average DEP for $\mathscr{C}$ is given by the following:

$$\gamma(\mathscr{C}) = 1 - \frac{1}{M}\sum_{i=1}^{M}\prod_{c=1}^{C}\prod_{\ell=1}^{L_c}\left(\int_{\mathcal{G}_c^{(\ell)}}\frac{1}{\pi\sigma^2}\exp\left(-\frac{(\Re(y)-\Re(x_c^{(\ell)}))^2+(\Im(y)-\Im(x_c^{(\ell)}))^2}{\sigma^2}\right)\mathrm{d}y\right)^{nP_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{187}$$

$$= 1 - \frac{1}{M}\sum_{i=1}^{M}\prod_{c=1}^{C}\prod_{\ell=1}^{L_c}\left(\int_{y:\left|y-x_c^{(\ell)}\right|^2\le r_c^2}\frac{1}{\pi\sigma^2}\exp\left(-\frac{\left|y-x_c^{(\ell)}\right|^2}{\sigma^2}\right)\mathrm{d}y\right)^{nP_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{188}$$

where, the expression in (188) follows from (187) due to (11) and (135o). The change of variable $y = z + x_c^{(\ell)} - \hat{x}_c^{(\ell)}$ in the integral in (188) yields

$$\int_{y:\left|y-x_c^{(\ell)}\right|^2 \le r_c^2} \frac{1}{\pi\sigma^2} \exp\left(-\frac{\left|y-x_c^{(\ell)}\right|^2}{\sigma^2}\right) \mathrm{d}y = \int_{z:\left|z-\hat{x}_c^{(\ell)}\right|^2 \le r_c^2} \frac{1}{\pi\sigma^2} \exp\left(-\frac{\left|z-\hat{x}_c^{(\ell)}\right|^2}{\sigma^2}\right) \mathrm{d}z. \tag{189}$$

Using (189) in (188) yields:

$$\gamma(\mathscr{C}) = 1 - \frac{1}{M}\sum_{i=1}^{M}\prod_{c=1}^{C}\prod_{\ell=1}^{L_c}\left(\int_{z:\left|z-\hat{x}_c^{(\ell)}\right|^2 \le r_c^2} \frac{1}{\pi\sigma^2} \exp\left(-\frac{\left|z-\hat{x}_c^{(\ell)}\right|^2}{\sigma^2}\right) \mathrm{d}z\right)^{nP_{\boldsymbol{u}(i)}(x_c^{(\ell)})}. \tag{190}$$

The empirical probability of usage of symbols is equal for the symbols $x_c^{(\ell)}$ and $\hat{x}_c^{(\ell)}$. That is, for all $I \in \{1, 2, \ldots, M\}$, all $c \in \{1, 2, \ldots, C\}$ and all $\ell \in \{1, 2, \ldots, L_c\}$, it holds that

$$P_{\boldsymbol{u}(i)}(x_c^{(\ell)}) = P_{\boldsymbol{u}(i)}(\hat{x}_c^{(\ell)}). \tag{191}$$

From (190) and (191), it follows that

$$\gamma(\mathscr{C}) = 1 - \frac{1}{M}\sum_{i=1}^{M}\prod_{c=1}^{C}\prod_{\ell=1}^{L_c}\left(\int_{z:\left|z-\hat{x}_c^{(\ell)}\right|^2 \le r_c^2} \frac{1}{\pi\sigma^2} \exp\left(-\frac{\left|z-\hat{x}_c^{(\ell)}\right|^2}{\sigma^2}\right) \mathrm{d}z\right)^{nP_{\boldsymbol{u}(i)}(\hat{x}_c^{(\ell)})} \tag{192}$$

$$= \gamma(\mathscr{C}'), \tag{193}$$

which is the desired result in (185).

From (184), for all $c \in \{1, 2, \ldots, C\}$, all $\ell \in \{1, 2, \ldots, L_c\}$, and $\omega \in [0, 2\pi]$, it holds that,

$$\hat{x}_c^{(\ell)} = e^{\mathrm{i}\omega} x_c^{(\ell)}. \tag{194}$$

This implies that

$$\left|\hat{x}_c^{(\ell)}\right| = \left|x_c^{(\ell)}\right|. \tag{195}$$

Using (40) and (29), the EOP for $\mathscr{C}$ is given by

$$\theta(\mathscr{C}, B) = \frac{1}{M}\sum_{i=1}^{M} \mathbb{1}_{\left\{k_1 \sum_{c=1}^{C}\sum_{\ell=1}^{L_c} nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)\left|x_c^{(\ell)}\right|^2 + k_2 \sum_{c=1}^{C}\sum_{\ell=1}^{L_c} nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)\left|x_c^{(\ell)}\right|^4 < B\right\}} \tag{196}$$

$$= \frac{1}{M}\sum_{i=1}^{M} \mathbb{1}_{\left\{k_1 \sum_{c=1}^{C}\sum_{\ell=1}^{L_c} nP_{\boldsymbol{u}(i)}\left(\hat{x}_c^{(\ell)}\right)\left|\hat{x}_c^{(\ell)}\right|^2 + k_2 \sum_{c=1}^{C}\sum_{\ell=1}^{L_c} nP_{\boldsymbol{u}(i)}\left(\hat{x}_c^{(\ell)}\right)\left|\hat{x}_c^{(\ell)}\right|^4 < B\right\}} \tag{197}$$

$$= \theta(\mathscr{C}', B), \tag{198}$$

where, the equality in (197) follows from (191) and (195). This completes the proof. □

Lemma 5.7 proves that the DEP and the EOP of a code $\mathscr{C}$ are independent of the rotations of the underlying set of channel input symbols $\mathcal{X}$. In fact, *for the given construction* of codes in this section, the DEP and the EOP are actually independent of the vector of phase shifts $\boldsymbol{\alpha}$ in (135e) altogether. The following lemma proves this result.

Lemma 5.8. *Consider (n, M)-codes $\mathscr{C}$ and $\mathscr{C}'$ of the form in* (135) *from the family* $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ *in* (135s). *The codes $\mathscr{C}$ and $\mathscr{C}'$ are identical except for the phase shift α_c in* (135a). *More specifically, the set of channel input symbols of $\mathscr{C}$ and $\mathscr{C}'$ are $\mathcal{X}$ and $\mathcal{X}'$, respectively such that*

$$\mathcal{X} = \bigcup_{c=1}^{C} \mathcal{U}(A_c, L_c, \alpha_c), \text{ and} \tag{199}$$

$$\mathcal{X}' = \bigcup_{c=1}^{C} \mathcal{U}(A_c, L_c, \alpha_c'). \tag{200}$$

Then, the average DEP γ in (22) *and the average EOP θ in* (37) *of $\mathscr{C}$ and $\mathscr{C}'$ satisfy the following:*

$$\theta(\mathscr{C}, B) = \theta(\mathscr{C}', B), \text{ and} \tag{201}$$

$$\gamma(\mathscr{C}) = \gamma(\mathscr{C}'). \tag{202}$$

Proof. Denote the ℓ^{th} symbol in layer c of $\mathscr{C}$ by $x_c^{(\ell)}$ and that of $\mathscr{C}'$ by $\bar{x}_c^{(\ell)}$. From (135a), it follows that

$$x_c^{(\ell)} = A_c \exp\left(\mathrm{i}\left(\frac{2\pi}{L_c}\ell + \alpha_c\right)\right) \tag{203}$$

$$= A_c \exp\left(\mathrm{i}\left(\frac{2\pi}{L_c}\ell + \alpha_c + \alpha_c' - \alpha_c'\right)\right) \tag{204}$$

$$= A_c \exp\left(\mathrm{i}\left(\frac{2\pi}{L_c}\ell + \alpha_c'\right)\right) \exp\left(\mathrm{i}\left(\alpha_c - \alpha_c'\right)\right) \tag{205}$$

$$= \exp\left(\mathrm{i}\left(\alpha_c - \alpha_c'\right)\right) \bar{x}_c^{(\ell)} \tag{206}$$

$$= \exp\left(\mathrm{i}\omega_c\right) \bar{x}_c^{(\ell)}, \tag{207}$$

where, $\omega_c = \alpha_c - \alpha_c' \in [0, 2\pi]$. This implies that for all $c \in \{1, 2, \ldots, C\}$ and all $\ell \in \{1, 2, \ldots, L_c\}$, it holds that

$$\left|x_c^{(\ell)}\right| = \left|\bar{x}_c^{(\ell)}\right|. \tag{208}$$

The result in (201) then follows from (208) and Lemma 5.7.

From (188), the average DEP for $\mathscr{C}$ is given by

$$\gamma\left(\mathscr{C}\right) = 1 - \frac{1}{M}\sum_{i=1}^{M}\prod_{c=1}^{C}\prod_{\ell=1}^{L_c}\left(\int_{y:\left|y-x_c^{(\ell)}\right|^2 \le r_c^2} \frac{1}{\pi\sigma^2}\exp\left(-\frac{\left|y - x_c^{(\ell)}\right|^2}{\sigma^2}\right)\mathrm{d}y\right)^{nP_{\boldsymbol{u}(i)}(x_c^{(\ell)})} \tag{209}$$

$$= 1 - \frac{1}{M}\sum_{i=1}^{M}\prod_{c=1}^{C}\prod_{\ell=1}^{L_c}\left(\int_{y:\left|y-x_c^{(\ell)}\right|^2 \le r_c^2} \frac{1}{\pi\sigma^2}\exp\left(-\frac{\left|y - x_c^{(\ell)}\right|^2}{\sigma^2}\right)\mathrm{d}y\right)^{nP_{\boldsymbol{u}(i)}(\bar{x}_c^{(\ell)})}, \tag{210}$$

where, the equality in (210) follows from (209) because $P_{\boldsymbol{u}(i)}(x_c^{(\ell)}) = P_{\boldsymbol{u}(i)}(\bar{x}_c^{(\ell)})$. Using the change of variable $y = z + x_c^{(\ell)} - \bar{x}_c^{(\ell)}$ in (210) yields

$$\gamma\left(\mathscr{C}\right) = 1 - \frac{1}{M}\sum_{i=1}^{M}\prod_{c=1}^{C}\prod_{\ell=1}^{L_c}\left(\int_{z:\left|z-\bar{x}_c^{(\ell)}\right|^2 \le r_c^2} \frac{1}{\pi\sigma^2}\exp\left(-\frac{\left|z - \bar{x}_c^{(\ell)}\right|^2}{\sigma^2}\right)\mathrm{d}z\right)^{nP_{\boldsymbol{u}(i)}(\bar{x}_c^{(\ell)})} \tag{211}$$

$$= \gamma\left(\mathscr{C}'\right), \tag{212}$$

which completes the proof of (202). □

The following lemma provides the achievable bound on the EOP δ in (41) for codes from the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in (135s).

Lemma 5.9. *Consider an (n, M, ϵ)-code $\mathscr{C}$ of the form in* (135) *from the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in* (135s)*. The code $\mathscr{C}$ is an $(n, M, \epsilon, B, \delta)$-code if the following holds:*

$$\delta \ge \frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c} nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_c^2 + k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c} nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_c^4\right) < B\right\}}, \tag{213}$$

where, k_1 and k_2 are positive real constants defined in (29)*.*

Proof. From (135a), for all $c \in \{1, 2, \ldots, C\}$ and all $\ell \in \{1, 2, \ldots, L_c\}$, the symbols $x_c^{(\ell)} \in \mathcal{U}(A_c, L_c, \alpha_c)$ in (135a) are given by

$$x_c^{(\ell)} = A_c \exp\left(\mathrm{i}\left(\frac{2\pi}{L_c}\ell + \alpha_c\right)\right). \tag{214}$$

This implies that

$$\left|x_c^{(\ell)}\right| = A_c. \tag{215}$$

From (40) and (29), the EOP for the code $\mathscr{C}$ is given by:

$$\theta(\mathscr{C},B)=\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)\left|x_c^{(\ell)}\right|^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)\left|x_c^{(\ell)}\right|^4\right)<B\right\}} \tag{216}$$

$$=\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_c^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_c^4\right)<B\right\}}. \tag{217}$$

From (41) and (217), the code $\mathscr{C}$ is an (n,M,ϵ,B,δ)-code if the following holds:

$$\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_c^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_c^4\right)<B\right\}}<\delta. \tag{218}$$

This completes the proof. □

The achievable bound on the EOP δ in (41) for a homogeneous code $\mathscr{C}$ from the family $\mathsf{C}\left(C,\boldsymbol{A},\boldsymbol{L},\boldsymbol{\alpha},\boldsymbol{p},\boldsymbol{r}\right)$ in (135s) is given by the following lemma:

Lemma 5.10. *Consider a homogeneous (n,M,ϵ)-code $\mathscr{C}$ of the form in* (135) *from the family $\mathsf{C}\left(C,\boldsymbol{A},\boldsymbol{L},\boldsymbol{\alpha},\boldsymbol{p},\boldsymbol{r}\right)$ in* (135s)*. The code $\mathscr{C}$ is an (n,M,ϵ,B,δ)-code if, for all $c\in\{1,2,\ldots,C\}$, the parameters p_c in* (135j) *satisfy the following:*

$$\delta=\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}np_cA_c^2+k_2\sum_{c=1}^{C}np_cA_c^4\right)<B\right\}}, \tag{219}$$

where, k_1 and k_2 are positive real constants defined in (29)*.*

Proof. From (40) and (30), the EOP for the homogeneous code $\mathscr{C}$ is given by:

$$\theta(\mathscr{C},B)=\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\mathscr{C}}\left(x_c^{(\ell)}\right)\left|x_c^{(\ell)}\right|^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\mathscr{C}}\left(x_c^{(\ell)}\right)\left|x_c^{(\ell)}\right|^4\right)<B\right\}} \tag{220}$$

$$=\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\mathscr{C}}\left(x_c^{(\ell)}\right)A_c^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\mathscr{C}}\left(x_c^{(\ell)}\right)A_c^4\right)<B\right\}} \tag{221}$$

$$=\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}n\frac{p_c}{L_c}A_c^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}n\frac{p_c}{L_c}A_c^4\right)<B\right\}} \tag{222}$$

$$=\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}np_cA_c^2+k_2\sum_{c=1}^{C}np_cA_c^4\right)<B\right\}}. \tag{223}$$

The equality in (221) follows from (215) and (222) follows from (135m). From (41) and (223), the code $\mathscr{C}$ is an (n,M,ϵ,B,δ)-code if the following holds:

$$\delta=\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}np_cA_c^2+k_2\sum_{c=1}^{C}np_cA_c^4\right)<B\right\}}. \tag{224}$$

This completes the proof. □

The following lemma provides an upper bound on the energy transmission rate B for (n,M,ϵ)-code $\mathscr{C}$ of the form in (135) from the family $\mathsf{C}\left(C,\boldsymbol{A},\boldsymbol{L},\boldsymbol{\alpha},\boldsymbol{p},\boldsymbol{r}\right)$.

Lemma 5.11. *Consider an (n,M,ϵ)-code $\mathscr{C}$ of the form in* (135) *from the family $\mathsf{C}\left(C,\boldsymbol{A},\boldsymbol{L},\boldsymbol{\alpha},\boldsymbol{p},\boldsymbol{r}\right)$ in* (135s)*. The code $\mathscr{C}$ is an (n,M,ϵ,B,δ)-code if the following holds:*

$$B\leq\begin{cases}\bar{e}_1 & \text{if } \ 0\leq\delta<\frac{y_1}{M},\\ \bar{e}_j & \text{if } \ \frac{\sum_{k=1}^{j-1}y_k}{M}<\delta\leq\frac{\sum_{k=1}^{j}y_k}{M},j\in\{2,3,\ldots,M'\}.\end{cases} \tag{225}$$

where, for all $i\in\{1,2,\ldots,M\}$, energy e_i is in (28)*; the positive integer M' is in* (31)*; and for all $j\in\{1,2,\ldots,M'\}$ $\bar{e}_j$ is in* (31) *and y_j is in* (32)*.*

Proof. The proof follows on the same lines as that for Lemma 4.3 where, for all $i\in\{1,2,\ldots,M\}$, the energy e_i in (28) is given by

$$e_i=\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_c^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_c^4. \tag{226}$$

□

The information transmission rate $R(\mathscr{C})$ of a code $\mathscr{C}$ is dictated by the number of channel input symbols L in (135f) and the probability vector $\boldsymbol{p}$ in (135h) that defines the empirical frequency with which each symbol appears in $\mathscr{C}$. The following lemma provides the value of $\boldsymbol{p}$ that achieves the upper bound on $R(\mathscr{C})$ for a given set of channel input symbols.

Lemma 5.12. *Consider an (n, M)-code $\mathscr{C}'$ of the form in* (135) *from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in* (135s) *with $\boldsymbol{p} = (p_1, p_2, \ldots, p_C)^{\mathsf{T}}$ in* (135h) *such that, for all $c \in \{1, 2, \ldots, C\}$,*

$$p_c = \frac{L_c}{L}, \tag{227}$$

where, L_c and L are as defined in (135a) *and* (135f) *respectively. Then, given any other (n, M)-code $\mathscr{C}$ in $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ that is identical to $\mathscr{C}'$ except for the probability distribution $\boldsymbol{p}$, it holds that,*

$$R(\mathscr{C}') \geq R(\mathscr{C}), \tag{228}$$

where, $R(\mathscr{C}')$ and $R(\mathscr{C})$ are the information transmission rates for $\mathscr{C}'$ and $\mathscr{C}$, respectively.

Proof. For the (n, M)-code $\mathscr{C}'$ in the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ with p_c of the form in (227), from (135m), for all $c \in \{1, 2, \ldots, C\}$ and all $\ell \in \{1, 2, \ldots, L_c\}$, the type $P_{\mathscr{C}'}$ is given by the following:

$$P_{\mathscr{C}'}(x_c^{(\ell)}) = \frac{p_c}{L_c} = \frac{1}{L}. \tag{229}$$

For the (n, M)-code $\mathscr{C}'$ with the type $P_{\mathscr{C}'}$ of the form in (229), the number of codewords that can be represented using L symbols is given by $M = L^n$. From (19), the information transmission rate for code $\mathscr{C}'$ is given by

$$R(\mathscr{C}') = \frac{\log M}{n} = \log L. \tag{230}$$

For the (n, M)-code $\mathscr{C}$, from (19) and Lemma 4.9, it follows that,

$$R(\mathscr{C}) \leq \log L = R(\mathscr{C}'). \tag{231}$$

This completes the proof. □

The following lemma provides the value of $\boldsymbol{p}$ that achieves the lower bound on the EOP $\theta(\mathscr{C}, B)$ for code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (135s).

Lemma 5.13. *Consider an (n, M)-code $\mathscr{C}$ of the form in* (135) *from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in* (135s) *with $\boldsymbol{p} = (p_1, p_2, \ldots, p_C)^{\mathsf{T}}$ in* (135h) *such that,*

$$p_c = \begin{cases} 1 & \text{if } c = 1, \\ 0 & \text{otherwise}. \end{cases} \tag{232}$$

Then, given any other (n, M)-code $\mathscr{C}'$ in the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ that is identical to $\mathscr{C}$ except for the probability distribution $\boldsymbol{p}$, it holds that

$$\theta(\mathscr{C}', B) \geq \theta(\mathscr{C}, B), \tag{233}$$

where θ is the average EOP in (37)*.*

Proof. Since $p_c = 0$ for all $c \in \{2, 3, \ldots, C\}$, from (135k) it follows that, for all $i \in \{1, 2, \ldots, M\}$, all $c \in \{2, 3, \ldots, C\}$ and all $\ell \in \{1, 2, \ldots, L_c\}$, the type

$$P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right) = 0. \tag{234}$$

Since $P_{\boldsymbol{u}(i)}$ is a pmf as defined in (17), it holds that, for all $i \in \{1, 2, \ldots, M\}$,

$$\sum_{c=1}^{C} \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right) = 1. \tag{235}$$

From (234) and (235), it follows that,

$$\sum_{\ell=1}^{L_1} P_{\boldsymbol{u}(i)}\left(x_1^{(\ell)}\right) = 1. \tag{236}$$

From (40) and (29), the EOP for the code $\mathscr{C}$ is given by

$$\theta(\mathscr{C},B)=\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)\left|x_c^{(\ell)}\right|^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)\left|x_c^{(\ell)}\right|^4<B\right\}} \tag{237}$$

$$=\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{k_1\sum_{\ell=1}^{L_1}nP_{\boldsymbol{u}(i)}\left(x_1^{(\ell)}\right)A_1^2+k_2\sum_{\ell=1}^{L_1}nP_{\boldsymbol{u}(i)}\left(x_1^{(\ell)}\right)A_1^4<B\right\}} \tag{238}$$

$$=\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{k_1nA_1^2+k_2nA_1^4<B\right\}}, \tag{239}$$

where, the equality in (238) follows from (234) and (239) follows from (236). Similarly, the EOP for the code $\mathscr{C}'$ is given by

$$\theta(\mathscr{C}',B)=\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)\left|x_c^{(\ell)}\right|^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)\left|x_c^{(\ell)}\right|^4<B\right\}} \tag{240}$$

$$=\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_c^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_c^4<B\right\}} \tag{241}$$

$$\geq\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{k_1\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_1^2+k_2\sum_{c=1}^{C}\sum_{\ell=1}^{L_c}nP_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)A_1^4<B\right\}} \tag{242}$$

$$=\frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\left\{k_1nA_1^2+k_2nA_1^4<B\right\}} \tag{243}$$

$$=\theta(\mathscr{C},B), \tag{244}$$

where, the equality in (241) follows from (215); the inequality in (242) follows from (135g); the equality in (243) follows from (235); and (244) follows from (239). This completes the proof. □

C. An Information-Energy Achievable Region

The achievable upper bound on the information rate R in (19), lower bound on the average DEP ϵ in (23) and lower bound on the average EOP δ in (41) determined in Section 5-B together define the achievable region of the codes of the form in (135) from the family $\mathsf{C}(C,\boldsymbol{A},\boldsymbol{L},\boldsymbol{\alpha},\boldsymbol{p},\boldsymbol{r})$ in (135s). The following theorem provides the achievable region thus defined, for the class of homogeneous codes in the family $\mathsf{C}(C,\boldsymbol{A},\boldsymbol{L},\boldsymbol{\alpha},\boldsymbol{p},\boldsymbol{r})$.

Theorem 5.14. *A homogeneous (n,M)-code $\mathscr{C}$ of the form in* (135) *from the family $\mathsf{C}(C,\boldsymbol{A},\boldsymbol{L},\boldsymbol{\alpha},\boldsymbol{p},\boldsymbol{r})$ in* (135s) *is an (n,M,ϵ,B,δ)-code if the number of layers C, and for all $c\in\{1,2,\ldots,C\}$, the parameters A_c, L_c, r_c, and p_c satisfy the following:*

$$\epsilon\geq 1-\prod_{c=1}^{C}\left(1-e^{-\frac{r_c^2}{\sigma^2}}\right)^{np_c}, \tag{245a}$$

$$L_c\leq\left\lfloor\frac{\pi}{2\arcsin\frac{r_c}{2A_c}}\right\rfloor, \tag{245b}$$

$$M\leq\frac{n!}{\prod_{c=1}^{C}\left((n\frac{p_c}{L_c})!\right)^{L_c}}, \tag{245c}$$

$$\delta=\mathbb{1}_{\left\{\left(k_1\sum_{c=1}^{C}np_cA_c^2+k_2\sum_{c=1}^{C}np_cA_c^4\right)<B\right\}}, \tag{245d}$$

where, the real σ^2 is defined in (12).

Proof. The result follows from Corollary 5.2; and Lemmas 5.4, 5.6, 5.10. □

6. Discussion

The converse and achievable bounds presented in this work reveal several interesting insights into the trade-offs between the information transmission rate R, the energy transmission rate B, the DEP ϵ and the EOP δ. For instance, the converse as well as the achievable lower bounds on δ in (42) and (213), respectively, increase as B increases and vice versa. The consequence of this relationship is that a lower δ can be achieved at the cost of lower B. Similarly, for increasing the converse or achievable upper bounds on B in (48) and (225), respectively, a higher value of δ has to be tolerated. These bounds also reveal that both B and δ can be improved by a code that has higher values of e_i in 28. This is achieved by using the symbols with greater

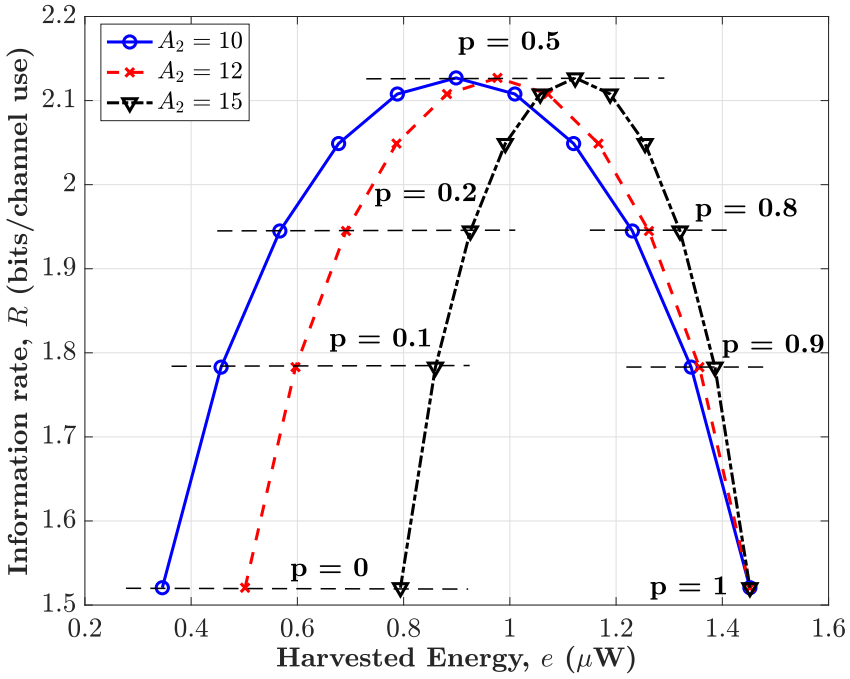

Fig. 2: Converse bounds on the information transmission rate R in (117) as a function of the harvested energy e in (247).

energy more frequently in the code. In fact, for the constructed codes in the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in (135s), the largest B and the least δ are achieved by a code with $p_1 = 1$ in (135h).

The converse lower bound on the DEP ϵ in (134) decreases as the distance between the symbols in the set of channel inputs $\mathcal{X}$ in (135b) increases. This implies that, a lower DEP can be achieved by increasing the distances between the channel input symbols. However, within the peak-amplitude constraint, increasing the distance between symbols implies that the number of symbols L decreases. This in turn, decreases the converse upper bound on the information rate R in (117). The achievable lower bound on ϵ in (245a) decreases as a function of the radii r_c of the decoding regions in (135p). On the other hand, the upper bound on the achievable information rate R in (174) increases as the radii r_c decrease. The trade-offs between R, B, ϵ, and δ are further illustrated using the following example.

Consider a homogeneous $(n, M, \epsilon, B, \delta)$-code $\mathscr{C}$ in the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in (135s) with the peak-amplitude constraint $P = 20$ millivolts in (15d). The set of channel input symbols $\mathcal{X}$ in (135b) is composed of two layers with 5 symbols in each layer, *i.e.*, $C = 2$ and $L_1 = L_2 = 5$. The radius of the first layer is $A_1 = P$ and the radius of the second layer A_2 is varied to illustrate the trade-offs between various parameters of $\mathscr{C}$. The frequency with which symbols from the first layer appear in the code is $p_1 = p = 1 - p_2$. That is, the vector $\boldsymbol{p}$ in (135h) is given by

$$\boldsymbol{p} = (p, (1-p))^{\mathsf{T}}. \tag{246}$$

The duration of the transmission in channel uses is $n = 100$. Since $\mathscr{C}$ is a homogeneous code, from (30), it holds that, for all $i \in \{1, 2, \ldots, M\}$,

$$e_i = e, \tag{247}$$

where $e \in [0, \infty)$ is calculated as in (30).

Fig. 2 shows the trade-offs between the converse bound on the information transmission rate R (117) and the harvested energy e in microwatts (μW) in (247) as a function of p in (246). Each curve in the figure is generated for some value of $A_2 < A_1$ by varying the value of $p \in [0, 1]$. The following trade-offs can be observed from this figure. The harvested energy e increases as p increases. This is because higher p corresponds to the symbols from the first layer $c = 1$ which have higher energy (since $A_1 > A_2$) being used more frequently in $\mathscr{C}$. For a fixed value of A_2 in Fig. 2, the bound on the information rate R first increases and then decreases as a function of e in (247). For each of these curves, the maximum $R = 2.13$ bits/channel use corresponds to uniform type, *i.e.*, $p = 0.5$. For p lesser or greater than 0.5, the bound on R decreases. Furthermore, the bounds on R are independent of the values of A_1 and A_2. This is due to the fact that the information rate R is only a function of the number of codewords M in (118).

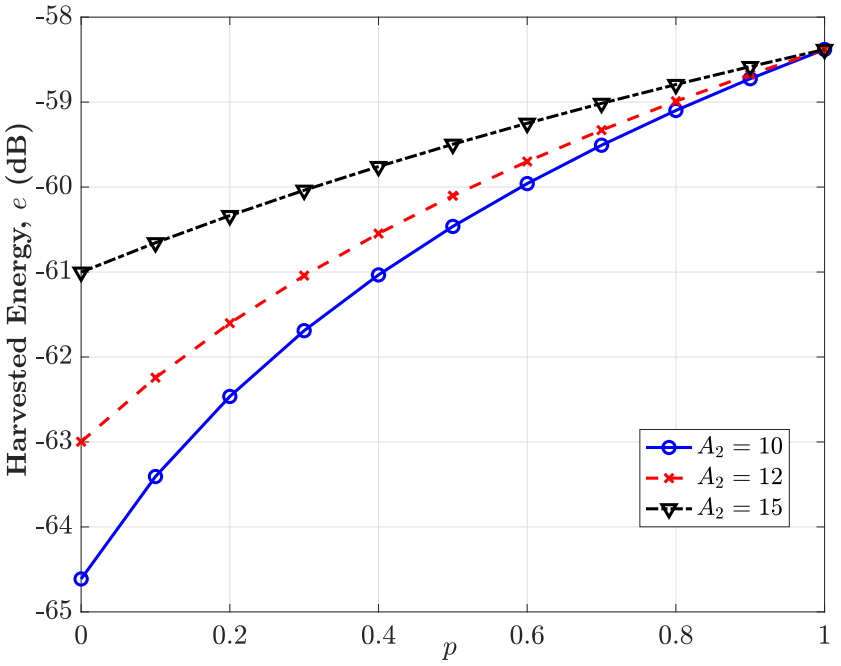

Fig. 3: The bounds on the harvested energy e in (247) as a function of p in (246).

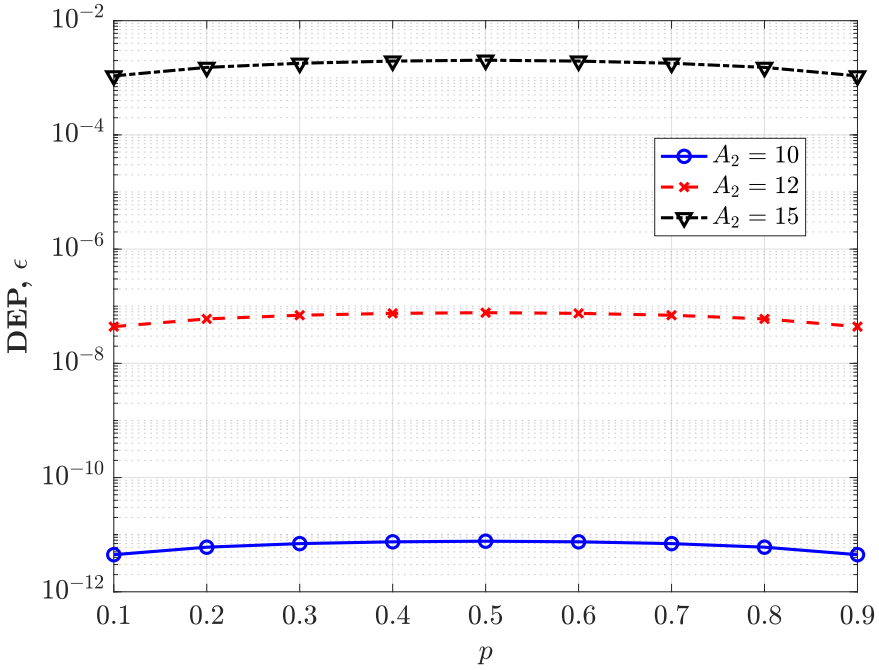

Fig. 4: The bounds on the DEP ϵ in (134) as a function of p in (246).

Fig. 3 shows the variation of the harvested energy e in (247) as a function of p in (246). The value of e increases as p increases. This is because increasing p means that the symbols from the first layer which have greater energy are used more frequently in $\mathscr{C}$ which increases the harvested energy. Furthermore, the harvested energy e also increases as A_2 increases. This is because, higher values of A_2 imply higher energy contained in the symbols in the second layer which inturn increases e. Fig. 4 shows the variation of the converse bound on the DEP ϵ in (134) as a function of the probability p in (246). It is observed that, as A_2 increases, the bound on ϵ increases. This is because, increasing A_2 results in the distances between the symbols in the two layers of $\mathcal{X}$ in (135b) to decrease which results in an increase in ϵ according to (134).

7. EXAMPLES

In this section, the characterized converse and achievable information-energy regions for SIET are illustrated using the following numerical example. Consider homogeneous $(n, M, \epsilon, B, \delta)$-codes in $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (135s) that employ the set of channel input symbols $\mathcal{X}$ of the form in (135b) with number of layers $C = 3$. The duration of the transmission in channel uses is $n = 80$.

In Fig. 5, the achievable bounds on the energy transmission rate B in (225) are plotted for homogeneous $(n, M, \epsilon, B, \delta)$-codes in the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ as a function of the achievable bounds on the DEP ϵ in (245a). The figure also shows the converse bounds on B in (48) as a function of the converse bounds on ϵ in (134). The amplitude of the first layer of the set of channel inputs $\mathcal{X}$ is fixed at $A_1 = 50$. Amplitudes of the second and third layers A_2 and A_3 are determined by the radii of the decoding regions according to (135q). For all $c \in \{1, 2, \ldots, C\}$, the number of symbols in layer c i.e., L_c, is determined by the radii r_c and the amplitudes A_c according to (245b). The probability vector in (135h) is $\boldsymbol{p} = (0.5, 0.3, 0.2)^{\mathsf{T}}$. The points on the curves are generated by varying r_c between 2 and 10.

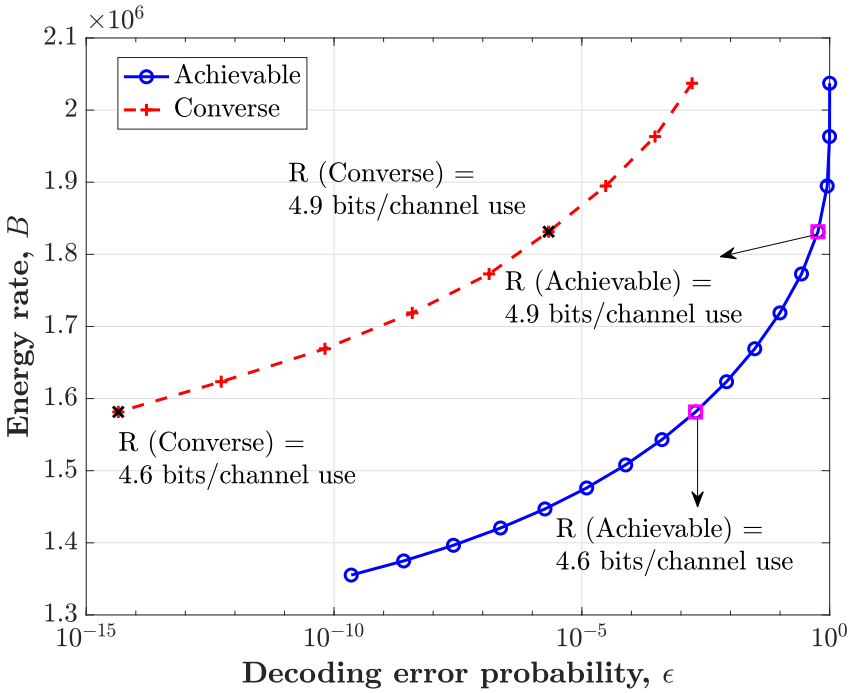

Fig. 5: Converse (48) and achievable (225) bounds on the energy transmission rate B for homogeneous codes in the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ as a function of the converse and achievable bounds on the DEP ϵ in (134) and (245a), respectively.

Fig. 5 shows several trade-offs between the energy transmission rate B, the DEP ϵ, and the information transmission rate R. Since the codes under consideration are homogeneous, the EOP $\delta \in \{0, 1\}$. Here, the δ is fixed to be 0 and the upper bounds on B in (225) and (48) are calculated accordingly. It is observed that the converse and achievable upper bounds on B increase as ϵ increases. This is due to the fact that increasing ϵ allows decreasing the radii of the decoding regions r_c in (135r) according to (245a). At the same time, decrease in r_c allows increasing the amplitudes A_2 and A_3 according to (135q). Increased amplitudes A_2 and A_3 result in higher values of the energy $e_{\mathscr{C}}$ in (30) which increases B. The information rate R also increases as a function of the DEP ϵ in Fig. 5. This is because, increasing ϵ allows decreasing the radii r_c according to (245a). At the same time, decrease in r_c increases the number of symbols in a layer L_c according to (245b) which increases the converse and achievable bounds on R in (117) and (174), respectively.

Fig. 6 illustrates the converse and achievable information-energy regions of the constructed homogeneous $(n, M, \epsilon, B, \delta)$-codes in the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (135s). In this case, the radii of the decoding regions r_c are assumed to be the same for all the layers i.e., for all $c \in \{1, 2, \ldots, C\}$, the radius $r_c = r$ in (135p). The value of r is obtained according to (245a). The amplitude of the first layer is $A_1 = 30$. Amplitudes of the second and third layers A_2 and A_3 are determined by r according to (135q). The points on the curves in Fig. 6 are obtained by varying ϵ and the probability vector $\boldsymbol{p}$ in (135h).

Fig. 6 shows the following trade-offs between the information and energy transmission rates in the converse and achievable curves. Firstly, the maximum achievable information transmission rate is $R = 4$ bits/channel use. This R is achieved by a code in which all the symbols in the set of channel inputs $\mathcal{X}$ are used with the same frequency. The maximum energy transmission

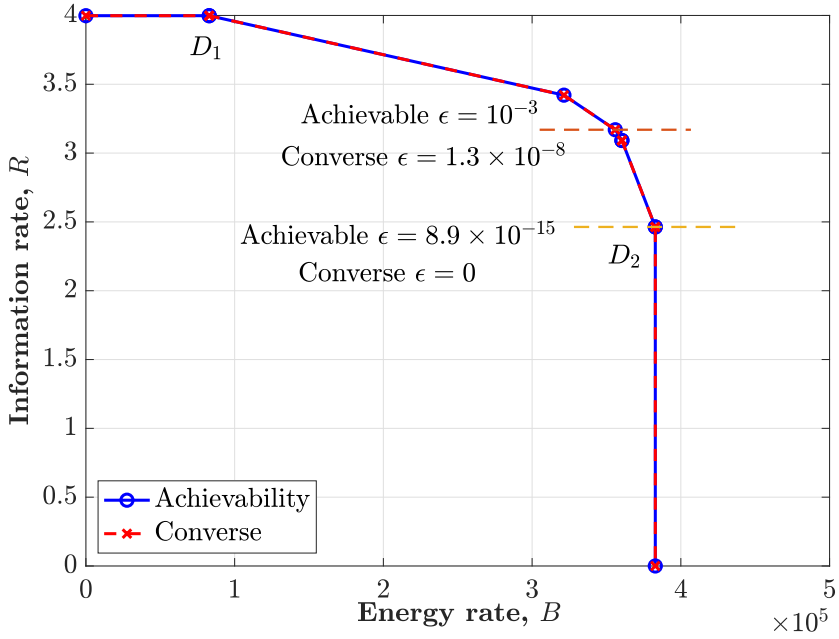

Fig. 6: Converse and achievable information-energy regions for homogeneous codes in the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$.

rate that can be achieved at $R = 4$ bits/channel use is $B = 8.3 \times 10^4$ energy units. This corresponds to the point D_1 in Fig. 6. Secondly, the maximum achievable B is 3.8×10^5 energy units. This is achieved by a code that exclusively uses the symbols in the first layer i.e., the probability vector $\boldsymbol{p}$ in (135h) is $\boldsymbol{p} = (1, 0, 0)^{\mathsf{T}}$. The maximum R that can be achieved at $B = 3.8 \times 10^5$ energy units is $R = 2.46$ bits/channel use. This corresponds to the point D_2 in Fig. 6. Thirdly, the curves between the points D_1 and D_2 in Fig. 6 show the trade-off between the information and energy transmission rates. As B is increased from 8×10^4 energy units at point D_1, R begins to decreases. Similarly, as R is increased from 2.46 bits/channel use at point D_2, B begins to decrease.

The codes constructed in this work are optimal in the sense of the converse results in Section 4 except for the DEP ϵ. Fig.5 shows that for the same values of the converse and achievable bounds on the energy transmission rate B, the lower bound on the DEP ϵ for the achievability curve is much higher than the converse. Similarly, for the same value of the information rate R in the converse and achievability curves, the corresponding ϵ for the achievability curve is higher than that of the converse. Fig. 6 shows that the converse and achievable information-energy rate curves for $\mathscr{C}$ overlap. However, for the same information and energy rate pair, the DEP for the achievable curves is higher than the converse. The sub-optimality in DEP arises due to the choice of circular decoding regions in (135p).

REFERENCES

[1] L. R. Varshney, "Transporting Information and Energy Simultaneously," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Toronto, ON, Canada, Jul. 2008, pp. 1612–1616.

[2] S. B. Amor and S. M. Perlaza, "Fundamental Limits of Simultaneous Energy and Information Transmission," in *Proc. International Conference on Telecommunications (ICT)*, Thessaloniki, Greece, May 2016, pp. 1–5.

[3] N. Khalfet and I. Krikidis, "The Capacity of SWIPT Systems over Rayleigh-Fading Channels with HPA," in *Proc. IEEE Information Theory Workshop (ITW)*, Kanazawa, Japan, 2021, pp. 1–6.

[4] P. Grover and A. Sahai, "Shannon Meets Tesla: Wireless Information and Power Transfer," in *Proc. IEEE International Symposium on Information Theory*, 2010, pp. 2363–2367.

[5] S. B. Amor, S. M. Perlaza, I. Krikidis, and H. V. Poor, "Feedback Enhances Simultaneous Energy and Information Transmission in Multiple Access Channels," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Barcelona, Spain, Jul. 2016, pp. 1974–1978.

[6] 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.

[7] S. M. Perlaza, A. Tajer, and H. V. Poor, "Simultaneous Information and Energy Transmission: A Finite Block-length Analysis," in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Kalamata, Greece, Jun. 2018, pp. 1–5.

[8] N. Khalfet, S. M. Perlaza, A. Tajer, and H. V. Poor, "On Ultra-reliable and Low Latency Simultaneous Information and Energy Transmission Systems," in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, Cannes, France, Jul. 2019, pp. 1–5.

[9] S. u. Zuhra, S. M. Perlaza, and E. Altman, "Simultaneous Information and Energy Transmission with Finite Constellations," in *Proc. IEEE Information Theory Workshop (ITW)*, Kanazawa, Japan, Oct. 2021, pp. 1–6.

[10] S. u. Zuhra, S. M. Perlaza, H. V. Poor, and E. Altman, "Achievable Information-Energy Region in the Finite Block-Length Regime with Finite Constellations," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, Espoo, Finland, Jun. 2022, pp. 2106–2111.

[11] I. Csiszár, "The Method of Types," *IEEE Transactions on Information Theory*, vol. 44, no. 5, pp. 2505 – 2523, Oct. 1998.

[12] S. u. Zuhra, S. M. Perlaza, H. V. Poor, and M. Skoglund, "Information-Energy Trade-offs with EH Non-linearities in the Finite Block-Length Regime with Finite Constellations ," in *Proc. IEEE Information Theory Workshop (ITW)*, Mumbai, India, Nov. 2022.

[13] M. Varasteh, B. Rassouli, and B. Clerckx, "Wireless Information and Power Transfer over an AWGN Channel: Nonlinearity and Asymmetric Gaussian Signaling," in *Proc. IEEE Information Theory Workshop (ITW)*, Kaohsiung, Taiwan, Nov. 2017, pp. 181–185.

[14] B. Clerckx, "Wireless Information and Power Transfer: Nonlinearity, Waveform Design, and Rate-Energy Tradeoff," *IEEE Transactions on Signal Processing*, vol. 66, no. 4, pp. 847–862, 2018.

[15] B. Clerckx and E. Bayguzina, "Waveform Design for Wireless Power Transfer," *IEEE Transactions on Signal Processing*, vol. 64, no. 23, pp. 6313–6328, 2016.

[16] S. Abeywickrama, R. Zhang, and C. Yuen, "Refined Nonlinear Rectenna Modeling and Optimal Waveform Design for Multi-User Multi-Antenna Wireless Power Transfer," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 5, pp. 1198–1210, 2021.

[17] S. Shen and B. Clerckx, "Joint Waveform and Beamforming Optimization for MIMO Wireless Power Transfer," *IEEE Transactions on Communications*, vol. 69, no. 8, pp. 5441–5455, 2021.

[18] O. L. A. López, F. A. Monteiro, H. Alves, R. Zhang, and M. Latva-Aho, "A Low-Complexity Beamforming Design for Multiuser Wireless Energy Transfer," *IEEE Wireless Communications Letters*, vol. 10, no. 1, pp. 58–62, 2021.

[19] Y. Zeng, B. Clerckx, and R. Zhang, "Communications and Signals Design for Wireless Power Transmission," *IEEE Transactions on Communications*, vol. 65, no. 5, pp. 2264–2290, 2017.

[20] P. Mukherjee, C. Psomas, and I. Krikidis, "Differential Chaos Shift Keying-Based Wireless Power Transfer with Nonlinearities," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 5, pp. 1185–1197, 2021.

[21] J. Kim, B. Clerckx, and P. D. Mitcheson, "Signal and System Design for Wireless Power Transfer: Prototype, Experiment and Validation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 11, pp. 7453–7469, 2020.

[22] A. Khalili, S. Zargari, Q. Wu, D. W. K. Ng, and R. Zhang, "Multi-Objective Resource Allocation for IRS-Aided SWIPT," *IEEE Wireless Communications Letters*, vol. 10, no. 6, pp. 1324–1328, 2021.

[23] N. Shanin, L. Cottatellucci, and R. Schober, "Markov Decision Process Based Design of SWIPT Systems: Non-Linear EH Circuits, Memory, and Impedance Mismatch," *IEEE Transactions on Communications*, vol. 69, no. 2, pp. 1259–1274, 2021.

[24] G. M. Kraidy, C. Psomas, and I. Krikidis, "Fundamentals of Circular QAM for Wireless Information and Power Transfer," in *Proc. IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2021, pp. 616–620.

[25] T. D. P. Perera, D. N. K. Jayakody, S. K. Sharma, S. Chatzinotas, and J. Li, "Simultaneous Wireless Information and Power Transfer (SWIPT): Recent Advances and Future Challenges," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 1, pp. 264–302, Dec. 2018.

[26] J. Huang, C.-C. Xing, and C. Wang, "Simultaneous Wireless Information and Power Transfer: Technologies, Applications, and Research Challenges," *IEEE Communications Magazine*, vol. 55, no. 11, pp. 26–32, Nov. 2017.

[27] B. Clerckx, R. Zhang, R. Schober, D. W. K. Ng, D. I. Kim, and H. V. Poor, "Fundamentals of Wireless Information and Power Transfer: From RF Energy Harvester Models to Signal and System Designs," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 1, pp. 4–33, 2019.

[28] B. Clerckx, J. Kim, K. W. Choi, and D. I. Kim, "Foundations of Wireless Information and Power Transfer: Theory, Prototypes, and Experiments," *Proceedings of the IEEE*, vol. 110, no. 1, pp. 8–30, Jan. 2022.

[29] L. Liu, R. Zhang, and K.-C. Chua, "Wireless Information Transfer with Opportunistic Energy Harvesting," *IEEE Transactions on Wireless Communications*, vol. 12, no. 1, pp. 288–300, 2013.

[30] N. Shanin, L. Cottatellucci, and R. Schober, "Rate-Power Region of SWIPT Systems Employing Nonlinear Energy Harvester Circuits with Memory," in *Proc. IEEE International Conference on Communications (ICC)*, Dublin, Ireland, Jun. 2020, pp. 1–7.

[31] E. Goudeli, C. Psomas, and I. Krikidis, "Sequential Decoding for Simultaneous Wireless Information and Power Transfer," in *Proc. International Conference on Telecommunications (ICT)*, 2017, pp. 1–5.

[32] G. Lin, Y. Zhou, W. Jiang, X. He, X. Zhou, G. He, and P. Yang, "LF-SWIPT: Outage Analysis for SWIPT Relaying Networks Using Lossy Forwarding With QoS Guaranteed," *IEEE Internet of Things Journal*, pp. 1–1, 2022.

[33] H. T. Thien, P.-V. Tuan, and I. Koo, "A Secure-Transmission Maximization Scheme for SWIPT Systems Assisted by an Intelligent Reflecting Surface and Deep Learning," *IEEE Access*, vol. 10, pp. 31 851–31 867, 2022.

[34] J. Liu, C.-H. R. Lin, Y.-C. Hu, and P. K. Donta, "Joint Beamforming, Power Allocation, and Splitting Control for SWIPT-Enabled IoT Networks with Deep Reinforcement Learning and Game Theory," *Sensors*, vol. 22, no. 6, p. 2328, 2022.

[35] B. Clerckx, K. Huang, L. R. Varshney, S. Ulukus, and M.-S. Alouini, "Wireless Power Transfer for Future Networks: Signal Processing, Machine Learning, Computing, and Sensing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 15, no. 5, pp. 1060–1094, 2021.

[36] R. Zhang and C. K. Ho, "MIMO Broadcasting for Simultaneous Wireless Information and Power Transfer," *IEEE Transactions on Wireless Communications*, vol. 12, no. 5, pp. 1989–2001, 2013.

[37] C. Song, C. Ling, J. Park, and B. Clerckx, "MIMO broadcasting for simultaneous wireless information and power transfer: Weighted MMSE approaches," in *2014 IEEE Globecom Workshops (GC Wkshps)*, 2014, pp. 1151–1156.

[38] M. Varasteh, J. Hoydis, and B. Clerckx, "Learning to Communicate and Energize: Modulation, Coding, and Multiple Access Designs for Wireless Information-Power Transmission," *IEEE Transactions on Communications*, vol. 68, no. 11, pp. 6822–6839, 2020.

[39] M. B. Goktas and Z. Ding, "A Wireless Power Transfer Assisted NOMA Transmission Scheme for 5G and Beyond mMTC," *IEEE Wireless Communications Letters*, pp. 1–1, 2022.

[40] C. E. García, M. R. Camana, and I. Koo, "Low-Complexity PSO-Based Resource Allocation Scheme for Cooperative Non-Linear SWIPT-Enabled NOMA," *IEEE Access*, vol. 10, pp. 34 207–34 220, 2022.

[41] D. Mishra and G. C. Alexandropoulos, "Jointly Optimal Spatial Channel Assignment and Power Allocation for MIMO SWIPT Systems," *IEEE Wireless Communications Letters*, vol. 7, no. 2, pp. 214–217, 2018.

[42] Q. Wu and R. Zhang, "Joint Active and Passive Beamforming Optimization for Intelligent Reflecting Surface Assisted SWIPT Under QoS Constraints," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 8, pp. 1735–1748, 2020.

[43] J. Park and B. Clerckx, "Joint Wireless Information and Energy Transfer in a Two-User MIMO Interference Channel," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 4210–4221, 2013.

[44] A. Dytso, M. Egan, S. M. Perlaza, H. V. Poor, and S. S. Shitz, "Optimal Inputs for Some Classes of Degraded Wiretap Channels," in *Proc. IEEE Information Theory Workshop (ITW)*, 2018, pp. 1–5.

[45] A. Dytso, S. Yagli, H. V. Poor, and S. Shamai Shitz, "The Capacity Achieving Distribution for the Amplitude Constrained Additive Gaussian Channel: An Upper Bound on the Number of Mass Points," *IEEE Transactions on Information Theory*, vol. 66, no. 4, pp. 2006–2022, 2020.

[46] S. Yagli, A. Dytso, H. V. Poor, and S. S. Shitz, "An Upper Bound on the Number of Mass Points in the Capacity Achieving Distribution for the Amplitude Constrained Additive Gaussian Channel," in *Proc. IEEE International Symposium on Information Theory (ISIT)*, 2019, pp. 1907–1911.

[47] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[48] A. Lapidoth, *A Foundation in Digital Communication*, 2nd ed. Cambridge University Press, 2017.

[49] H. Robbins, "A Remark on Stirling's Formula," *The American mathematical monthly*, vol. 62, no. 1, pp. 26–29, 1955.

[50] J. G. Proakis and M. Salehi, *Digital Communications*, 5th ed. McGraw-Hill Higher Education, 2008.