Simultaneous Information and Energy Transmission with Short Packets and Finite Constellations

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

Abstract

This paper characterizes the trade-offs between information and energy transmission over an additive white Gaussian noise channel in the finite block-length regime with finite channel input symbols. These trade-offs are characterized in the form of inequalities involving the information transmission rate, energy transmission rate, decoding error probability (DEP) and energy outage probability (EOP) for a given finite block-length code. The first set of results identify the tuples of information rate, energy rate, DEP and EOP that cannot be achieved by a given code. Following this, a novel method for constructing a family of codes that can satisfy a target information rate, energy rate, DEP and EOP is proposed. Finally, the achievability results identify the set of tuples of information rate, energy rate, DEP and EOP that can be simultaneously achieved by the constructed family of codes.

Index Terms

Simultaneous information and energy transmission, finite block-length, and finite channel inputs.

1. Introduction

Simultaneous information and energy transmission (SIET), also known as simultaneous wireless information and power transfer (SWIPT), is a technology that employs radio frequency (RF) signals to simultaneously transmit both information and energy to (possibly different) remote devices [1]–[3]. A central question in modern SIET is to characterize the information rate, energy rate, decoding error probability (DEP) and energy outage probability (EOP) that can be simultaneously achieved by any given transmission technique. Such question has been traditionally studied under at least one of two simplifying assumptions. The first assumption is that the transmission duration is infinitely long, as in [2], [4]–[8]. Under this assumption, the DEP and the EOP can be made arbitrarily close to zero, and thus, the focus is exclusively on the information and energy transmission rates. If this assumption is dropped, as in the remainder of this work, the DEP and EOP are bounded away from zero, and thus, the performance analysis of SIET shall take into account four metrics: information rate, energy rate, DEP, and EOP. The second common assumption is that channel inputs symbols can be drawn from probability distributions whose support is uncountable, e.g., Gaussian distributions as in [9]. This presumes the ability to transmit with infinitely many channel input symbols, which is not the case in practice. In a departure from this assumption, this work considers the more practical case of finite channel input symbols, which provides insights for practical systems operating with modulation schemes built upon commercial constellations.

A. State of the art

The existing body of work in SIET can be divided into three main categories. The first category is the study of the trade-offs between the information and energy rates that can be simultaneously transmitted by an RF signal. In the asymptotic block-length regime, the focus is on the notion of information-energy regions to characterize the set of all information and energy rate pairs that can be simultaneously achieved [5]. To capture the trade-off between the information and energy rates, a capacity-energy function is defined in [4] for various channels including the discrete memoryless channel, binary symmetric channel, and the additive white Gaussian noise (AWGN) channel. The information-energy capacity region of the Gaussian multiple access channel is characterized in [7], [10], whereas the information-energy capacity region of the Gaussian interference channel is approximated in [8]. In the finite block-length regime, [11] and [12] provide a characterization of the information-energy

TABLE I: Summary of the state-of-art

Reference	Channel inputs	Block-length	EH non-linearities	Channel	DEP	EOP
[4]	Infinite	Asymptotic	-	DMC, AWGN	-	-
[2]	Infinite	Asymptotic	-	AWGN + Frequency-selective fading	-	-
[6]	Infinite	Asymptotic	✓	Rayleigh fading	-	-
[32]	Infinite	Asymptotic	-	Flat fading	-	-
[19]	Infinite	Asymptotic	-	AWGN	-	-
[33]	Finite	Asymptotic	-	AWGN + Fading	-	-
[26]	Finite	Asymptotic	-	AWGN + Flat-fading	-	-
[34]	Finite	-	-	AWGN + Rayleigh fading	✓	-
[27]	Finite	-	-	AWGN	-	-
[16]	Infinite	Asymptotic	✓	AWGN	-	-
[9]	Infinite	Asymptotic	✓	Multi-path fading	-	-
[35]	Uncountable	Asymptotic	-	AWGN + Rayleigh fading	✓	-
[36]	Uncountable	Asymptotic	-	AWGN + Rayleigh and Rician fading	-	-
[37]	Uncountable	Asymptotic	-	AWGN	-	-
[38]	Uncountable	Asymptotic	✓	Multi-path fading	-	-
[39]	Uncountable	Asymptotic	-	Multi-path fading	-	-
[40]	Uncountable	Asymptotic	-	AWGN + Flat-fading	-	-
[41]	Finite	Finite	✓	AWGN	-	-
[11]	Finite	Finite	-	BSC	✓	✓
[12]	Finite	Finite	-	BSC	-	-
This work	Finite	Finite	✓	AWGN	✓	✓

capacity region with binary antipodal channel inputs and fixed DEP and EOP. Converse and achievability bounds on these metrics for finite channel input symbols are presented in [13] and [14], respectively. The fundamental limits of SIET for the case of semantic communications are characterized in [15].

The second category of research studies the modeling of energy harvesting (EH) circuits. One of the key questions in this regard tackles the estimation of the amount of energy that can be harvested 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 second and fourth powers of the signal magnitude [16]. Recent research on EH non-linearities [9], [17] has shown that energy models that do not account for these non-linearities result in inaccurate estimates of the harvested energy. The impact of the energy harvester non-linearities on the fundamental limits of SIET in the finite block-length regime has been studied in [18]. In [19], a learning based approach is proposed for modeling the EH circuits where the memory of non-linear elements in the EH circuit is modeled as a Markov decision process.

The third line of research covers the system design, implementation, and applications of SIET networks. Optimal waveform design for SIET from a multi-antenna transmitter to multiple single antenna receivers is studied in [20]. Signal and system design exclusively for wireless energy transmission has been studied in [17], [21]–[24] and [25]. In [26], the authors optimize resource allocation and beamforming for intelligent reflecting surfaces aided SIET. An algorithm for designing a circular quadrature amplitude modulation scheme for SIET that maximizes the peak-to-average power ratio has been proposed in [27].

More comprehensive overviews of the work on SIET in the second and third categories detailed above, can be found in [28], [29], [30], and [31]. 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 while a dash indicates the opposite.

B. Contributions

This work provides a comprehensive characterization of the relationships between the parameters of short packet codes with finite channel input symbols (finite constellations, e.g., QPSK, 16-QAM, 64-QAM, among others) for simultaneous transmission of information and energy. The main contributions can be summarized as follows.

- The first set of results, presented in Section 3, provide bounds on the information transmission rate, energy transmission rate, DEP and EOP of any given code over a point-to-point AWGN channel. All these bounds emerge to be functions of common code parameters such as the block-length and the type. Together these bounds form the necessary conditions that any code must satisfy for accomplishing the tasks of information and energy transmission simultaneously. Through the dependence on common parameters of the code, these bounds also characterize the trade-offs between information rate, energy rate, DEP and EOP of a given code.
- A method for constructing codes for SIET over short packet communication is proposed. Using this method, a family of codes can be designed to achieve given rates of information and energy transmission while also ensuring that the DEP

and EOP are kept within the required thresholds. The feasible tuples of information rate, energy rate, DEP and EOP that can be achieved by the codes thus constructed are also characterized.
- The performance of the constructed codes is evaluated by comparing the achievability results with the identified necessary conditions. Comprehensive numerical analyses are also conducted to illustrate the trade-offs between the competing objectives of information and energy transmission identified by the results in this work.

C. Notation

The sets of natural, real and complex numbers are denoted by $\mathbb{N}$, $\mathbb{R}$ and $\mathbb{C}$, respectively. In particular, $0 \notin \mathbb{N}$. 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 empty set is denoted by $\emptyset$. The sinc function is defined as follows

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

The tail distribution function of the standard normal distribution is referred to as the Q-function where

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

2. 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 while also delivering B Joules of energy to the EH over an AWGN channel. The transmission takes place over a finite duration of $n \in \mathbb{N}$ channel uses. Thus, the rate of energy transmission is B/n Joules per channel use. The transmitter uses L symbols from the set of channel input symbols

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

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 transmitted over n channel uses by

$$\boldsymbol{\nu} = (\nu_1, \nu_2, \ldots, \nu_n)^{\mathsf{T}}. \tag{4}$$

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

$$x(t) = \sum_{m=1}^n \nu_m \, \mathrm{sinc}\left(f_w\left(t - (m-1)T\right)\right), \tag{5}$$

where the sinc function is defined in (1). The signal $x(t)$ in (5) 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 (5) as follows:

$$\tilde{x}(t) = \Re\left(x(t)\sqrt{2}\exp(\mathrm{i}2\pi f_c t)\right), \tag{6}$$

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

$$Y(t) = \tilde{x}(t) + N_1(t), \text{ and} \tag{7a}$$

$$Z(t) = \tilde{x}(t) + N_2(t), \tag{7b}$$

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; and the random variables $Y(t)$ and $Z(t)$ are the inputs to the IR and the EH, respectively.

At the IR, the received signal $Y(t)$ in (7a) 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 $f_w \, \mathrm{sinc}\left(f_w t\right)$ 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 [42, Section 2.2.4]:

$$\boldsymbol{Y} = \boldsymbol{\nu} + \boldsymbol{N}, \tag{8}$$

where the vector $\boldsymbol{Y} = (Y_1, Y_2, \ldots, Y_n)^\mathsf{T} \in \mathbb{C}^n$ is the input to the IR; the vector of channel input symbols $\boldsymbol{\nu}$ is in (4); and the noise vector $\boldsymbol{N} = (N_1, N_2, \ldots, N_n)^\mathsf{T} \in \mathbb{C}^n$ is 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 [42, Section 2.2.4]). Therefore, for all $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)^\mathsf{T} \in \mathbb{C}^n$, and for 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 (8) is given by

$$f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{\nu}) = \prod_{m=1}^{n} f_{Y|X}(y_m|\nu_m), \tag{9a}$$

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

$$f_{Y|X}(y_m|\nu_m) = \frac{1}{\pi\sigma^2} \exp\left(-\frac{|y_m - \nu_m|^2}{\sigma^2}\right). \tag{9b}$$

At the EH, the RF signal $Z(t)$ in (7b) is not down-converted or filtered. See for instance [9], [17].

A. Information and Energy Transmission

Let M be the cardinality of the set of symbols at the output of the information source. To accomplish the tasks of information and energy transmission, the transmitter makes use of an $(n, M, \mathcal{X})$-code defined as follows.

Definition 2.1 $((n, M, \mathcal{X})$-code$)$. *An $(n, M, \mathcal{X})$-code, with $\mathcal{X}$ in (3), is a set of pairs*

$$\{(\boldsymbol{u}(1), \mathcal{D}_1), (\boldsymbol{u}(2), \mathcal{D}_2), \ldots, (\boldsymbol{u}(M), \mathcal{D}_M)\}, \tag{10}$$

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

$$\boldsymbol{u}(i) = (u_1(i), u_2(i), \ldots, u_n(i)) \in \mathcal{X}^n, \tag{11a}$$

$$\mathcal{D}_i \subseteq \mathbb{C}^n, \text{ and} \tag{11b}$$

$$\mathcal{D}_i \cap \mathcal{D}_j = \emptyset. \tag{11c}$$

Note that Definition 2.1 specifies the exact set of channel input symbols via the set $\mathcal{X}$. Thus, constraints on the average power or peak-amplitude power are vacuous as those are fixed by the definition of $\mathcal{X}$.

The associated encoding, decoding and energy harvesting operations for a given $(n, M, \mathcal{X})$-code are defined as follows.

1) Encoding and decoding: Assume that the transmitter uses the $(n, M, \mathcal{X})$-code

$$\mathscr{C} \triangleq \{(\boldsymbol{u}(1), \mathcal{D}_1), (\boldsymbol{u}(2), \mathcal{D}_2), \ldots, (\boldsymbol{u}(M), \mathcal{D}_M)\}, \tag{12}$$

that satisfies (11).

To transmit 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 (8). Let W be a random variable that represents the output of the information source. For all $i \in \{1, 2, \ldots, M\}$, the probability of transmitting codeword $\boldsymbol{u}(i)$ is

$$P_W(i) = \frac{1}{M}, \tag{13}$$

which represents a maximum entropy information source. The IR decides that message index i was transmitted, if the following event takes place:

$$\boldsymbol{Y} \in \mathcal{D}_i, \tag{14}$$

with $\mathcal{D}_i$ in (12). Therefore, the DEP associated with the transmission of message index i is

$$\gamma_i(\mathscr{C}) \triangleq 1 - \int_{\mathcal{D}_i} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{u}(i)) \mathrm{d}\boldsymbol{y}, \tag{15}$$

with the conditional density $f_{\boldsymbol{Y}|\boldsymbol{X}}$ defined in (9). Alternatively, the average DEP for code $\mathscr{C}$ is

$$\gamma(\mathscr{C}) \triangleq \frac{1}{M} \sum_{i=1}^{M} \gamma_i(\mathscr{C}). \tag{16}$$

The information transmission rate of any $(n, M, \mathcal{X})$-code $\mathscr{C}$, denoted by $R(\mathscr{C})$, satisfies

$$R(\mathscr{C}) \triangleq \frac{\log_2 M}{n}, \tag{17}$$

in bits per channel use.

2) Energy Harvesting: The channel output observed at the EH while transmitting codeword $\boldsymbol{u}(i)$, with $i \in \{1, 2, \ldots, M\}$, is denoted by $Z_i(t)$, with $t \in [0, nT]$. From (6) and (7b), such a signal $Z_i(t)$ is

$$Z_i(t) = \Re\Big(\sqrt{2} \sum_{m=1}^{n} u_m(i) \tag{18}$$

$$\operatorname{sinc}\left(f_w(t - (m-1)T)\right) \exp\left(\mathrm{i}2\pi f_c t\right)\Big) + N_2(t),$$

where, for all $m \in \{1, 2, \ldots, n\}$, the complex $u_m(i)$ is the mth symbol of the codeword $\boldsymbol{u}(i)$ in (11a) and the random variable $N_2(t)$ is a real Gaussian random variable with zero mean and variance σ^2. For all $i \in \{1, 2, \ldots, M\}$ and $t \in [0, nT]$, denote by $x_i(t)$ the signal

$$x_i(t) = \tag{19}$$

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

From (18) and (19) it follows that,

$$Z_i(t) = x_i(t) + N_2(t). \tag{20}$$

Thus, for all $t \in [0, nT]$, the channel output $Z_i(t)$ in (20) is a real Gaussian random variable with mean $x_i(t)$ and variance σ^2.

The energy harvesting model used in this work is the non-linear model introduced in [9] and [17]. This model states that the energy harvested from an RF signal is proportional to the DC component of the second and fourth powers of the signal. Hence, for all $i \in \{1, 2, \ldots, M\}$, the expected energy harvested from the channel output $Z_i(t)$ in (18) during the whole duration of the source symbol i is

$$e_i \triangleq k_1 \sum_{m=1}^{n} |u_m(i)|^2 + k_2 \sum_{m=1}^{n} |u_m(i)|^4, \tag{21}$$

where the constants k_1 and k_2 are positive, with $k_1 = 0.0034$ and $k_2 = 0.3829$ [9]; and for all $m \in \{1, 2, \ldots, n\}$, the complex number $u_m(i)$ is in (11a).

The energy harvested during the transmission of a codeword, i.e., during n channel uses, is a random variable denoted by E. The probability of harvesting energy e given that codeword $\boldsymbol{u}(i)$ was transmitted is given by

$$P_{E|W}(e|i) = \mathbb{1}_{\{e=e_i\}}, \tag{22}$$

where the random variable W possesses the probability mass function defined in (13); and the energy level e_i is defined in (21). The equality in (22) follows from the fact that once the transmitted codeword is known to be $\boldsymbol{u}(i)$, the energy harvested is a deterministic quantity e_i, as defined in (21). This is essentially due to the fact that the noise $N_2(t)$ in (20) does not contribute any significant amount of energy. The signal from which energy is harvested is $x_i(t)$. Therefore, the probability mass function of the harvested energy, given that codeword $\boldsymbol{u}(i)$ was transmitted is a point mass concentrated at e_i.

The probability mass function of the random variable E, denoted by P_E is given by the following:

$$P_E(e) = \sum_{i=1}^{M} P_{E|W}(e|i) P_W(i) \tag{23}$$

$$= \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\{e=e_i\}}. \tag{24}$$

Let $B \in \mathbb{R}$ be the energy required at the EH. Then, using (23), the EOP associated with code $\mathscr{C}$ is the probability of transmitting a message $i \in \{1, 2, \ldots, M\}$, such that e_i in (21) satisfies $e_i < B$. That is,

$$\theta(\mathscr{C}, B) \triangleq \sum_{i \in \{j \in \{1,2,\ldots,M\}: e_j < B\}} P_E(e_i) \tag{25}$$

$$= \sum_{i \in \{j \in \{1,2,\ldots,M\}: e_j < B\}} \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\{e=e_i\}} \tag{26}$$

$$= \frac{1}{M} \left| \{i \in \{1, 2, \ldots, M\} : e_i < B\} \right| \tag{27}$$

$$= \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\{e_i < B\}}, \tag{28}$$

where the equality in (26) follows from (24).

Definition 2.1 can now be refined to include the DEP and the EOP as follows.

Definition 2.2 $((n, M, \mathcal{X}, \epsilon, B, \delta)$-code). *An $(n, M, \mathcal{X})$-code $\mathscr{C}$ is said to be an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code for the random transformation in (9) if the following hold:*

$$\gamma(\mathscr{C}) \leq \epsilon, and \tag{29}$$

$$\theta(\mathscr{C}, B) \leq \delta, \tag{30}$$

where $\gamma(\mathscr{C})$ and $\theta(\mathscr{C}, B)$ are defined in (16) and (25), respectively.

The $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes thus defined can be employed for transmitting at most B energy units in n channel uses and at most $\frac{\log_2 M}{n}$ bits per channel use, while ensuring that the resulting DEP is smaller than ϵ and the EOP is smaller than δ.

The primary objective of what follows in Section 3 is a characterization of the trade-offs between the various information and energy parameters of $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes. These trade-offs are presented in the form of inequalities involving the code parameters $n, M, \mathcal{X}, \epsilon, B$ and δ. This characterization serves two main purposes. First, these inequalities define what tuples $(n, M, \mathcal{X}, \epsilon, B, \delta)$ are "*possible*". In other words, these inequalities define the necessary conditions that an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code must satisfy. Second, the dependence of M, ϵ, B, δ on common parameters precisely quantifies the trade-offs between the objectives of information and energy transmission.

3. NECESSARY CONDITIONS

This section introduces some inequalities involving the information and energy rates, the DEP and the EOP that any $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code satisfies. From this perspective, these inequalities are recognized to be necessary conditions for SIET. Furthermore, these inequalities quantify the mutual dependence between the information rate, the energy rate, the DEP and the EOP of $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes, which in turn, provides insights into the feasibility of SIET for specific reliability guarantees. For a specific $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code, e.g., the code $\mathscr{C}$ in (12), this mutual dependence is presented in terms of the *type* induced by each of the codewords of such a code. The type induced by the codeword $\boldsymbol{u}(i)$, with $i \in \{1, 2, \ldots, M\}$, is a probability mass function, denoted by $P_{\boldsymbol{u}(i)}$, such that for all $x \in \mathcal{X}$,

$$P_{\boldsymbol{u}(i)}(x) \triangleq \frac{1}{n} \sum_{m=1}^{n} \mathbb{1}_{\{u_m(i)=x\}}. \tag{31}$$

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

$$P_{\mathscr{C}}(x) \triangleq \frac{1}{M} \sum_{i=1}^{M} P_{\boldsymbol{u}(i)}(x). \tag{32}$$

For the ease of presentation, the main results are presented for a specific class of $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes, referred to as *constant composition* codes. Nonetheless, the results presented in this section can be easily extended to the whole class of $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes.

Definition 3.1 (Constant Composition Codes). *An $(n, M, \mathcal{X})$-code $\mathscr{C}$ is said to be a constant composition code if for all $i \in \{1, 2, \ldots, M\}$ and for all $x \in \mathcal{X}$, it holds that*

$$P_{\boldsymbol{u}(i)}(x) = P_{\mathscr{C}}(x), \tag{33}$$

where, $P_{\boldsymbol{u}(i)}$ and $P_{\mathscr{C}}$ are the types defined in (31) and (32), respectively.

Constant composition codes are $(n, M, \mathcal{X})$-codes in which a given channel input symbol is used the same number of times in all the codewords. As a consequence, for a constant composition code $\mathscr{C}$, the amount of energy e_i in (21) harvested while transmitting the codeword $i \in \{1, 2, \ldots, M\}$, is invariant with respect to i. That is, for all $i \in \{1, 2, \ldots, M\}$,

$$e_i = e_{\mathscr{C}}, \tag{34}$$

and the positive real $e_{\mathscr{C}}$ satisfies

$$e_{\mathscr{C}} \triangleq k_1 \sum_{x \in \mathcal{X}} n P_{\mathscr{C}}(x) |x|^2 + k_2 \sum_{x \in \mathcal{X}} n P_{\mathscr{C}}(x) |x|^4, \tag{35}$$

where the constants k_1 and k_2 are the same as in (21); $\mathcal{X}$ in (3); and $P_{\mathscr{C}}$ in (32).

The following theorem introduces a first caracterization of the dependences among the performance metrics of an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code.

Theorem 3.1. *If a given constant composition $(n, M, \mathcal{X})$-code, $\mathscr{C}$, with $\mathcal{X}$ in (3), is an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code for the random transformation in (9), then the following conditions simultaneously hold:*

$$\epsilon \geq (M-1)Q\left(\sqrt{\frac{\sum_{\ell=1}^{L} n P_{\mathscr{C}}(x^{(\ell)}) |x^{(\ell)} - \bar{x}^{(\ell)}|^2}{2\sigma^2}} \right), \tag{36a}$$

$$R(\mathscr{C}) \leq \frac{1}{n} \log_2\left(\frac{n!}{\prod_{\ell=1}^{L} (n P_{\mathscr{C}}(x^{(\ell)}))!} \right), \tag{36b}$$

$$\delta \geq \mathbb{1}_{\{e_{\mathscr{C}} < B\}}, \tag{36c}$$

$$B \leq e_{\mathscr{C}}, \tag{36d}$$

where, the information rate $R(\mathscr{C})$ is defined in (87); the function Q is defined in (2); for all $\ell \in \{1, 2, \ldots, L\}$, the complex $\bar{x}^{(\ell)} \in \mathcal{X}$ is

$$\bar{x}^{(\ell)} \in \arg \max_{x \in \mathcal{X} \setminus \{x^{(\ell)}\}} |x^{(\ell)} - x|; \tag{37}$$

$P_{\mathscr{C}}$ is the type defined in (32); and $e_{\mathscr{C}}$ is defined in (35).

Proof. The proofs are presented in Appendix A. $\qquad \square$

Theorem 3.1 introduces the fundamental limits on the main performance metrics, namely the information transmission rate, energy transmission rate, DEP, and EOP, of all constant composition $(n, M, \mathcal{X})$-codes. Essentially, if a given constant composition $(n, M, \mathcal{X})$-code $\mathscr{C}$ is an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code, then the DEP ϵ cannot be smaller that the right-hand side of (36a); the information transmission rate $R(\mathscr{C})$ in (17) cannot be bigger than the right-hand side of (87); the EOP δ cannot be smaller than the right-hand side of (36c); and the energy transmitted cannot be bigger than the right-hand side of (36d). These fundamental upper and lower bounds on the performance metrics depend on the type induced by the code $\mathscr{C}$. This observation reveals a profound conclusion: Two different $(n, M, \mathcal{X})$-codes $\mathscr{C}_1$ and $\mathscr{C}_2$ of the form in (12), might exhibit different codewords and different decoding regions, nonetheless, if they induce the same types, i.e., $P_{\mathscr{C}_1} = P_{\mathscr{C}_2}$, then their performance metrics to simultaneously transmit information and energy are subject to identical constraints. This reveals the relevance of the type $P_{\mathscr{C}}$ in (36) as the key parameter in code design for SIET.

In the remainder of this section, the necessary conditions on the information and energy transmission parameters of $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes are discussed in further detail.

A. DEP

To minimize the DEP for the code $\mathscr{C}$, the decoding regions $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_M$ are defined using the maximum a posteriori (MAP) decision rule [43, Chapter 21], *i.e.*, for all $i \in \{1, 2, \ldots, M\}$,

$$\mathcal{D}_i = \left\{ \boldsymbol{y} \in \mathbb{C} : \forall j \in \{1, 2, \ldots, M\}, \frac{f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{u}(i))}{f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{u}(j))} \geq 1 \right\}. \tag{38}$$

The lower bound on the DEP, ϵ, in (36a), is independent of the decoding regions $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_M$ of the code $\mathscr{C}$. This lower bound is the exact DEP of a code in which all the channel input symbols in $\mathcal{X}$ are equidistant from one another, such as in the case of binary phase-shift keying. Hence, the lower bound on the DEP, ϵ, in (36a) is tight for such a code.

B. Energy Transmission Rate and EOP

The lower and upper bounds in (36c) and (36d), which might appear redundant with respect to each other in the case of constant composition codes, are more informative in the general case as discussed below. The following lemma provides a lower bound on δ for the general case of $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes.

Lemma 3.2. *If a given $(n, M, \mathcal{X})$-code, $\mathscr{C}$, with $\mathcal{X}$ in (3), is an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code for the random transformation in (9), then the following holds for δ in (30):*

$$\delta \geq \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\{e_i < B\}}, \tag{39}$$

where, for all $i \in \{1, 2, \ldots, M\}$, the real $e_i \in [0, \infty)$ is in (21).

Proof. The result follows from (28) and (30). □

In Lemma 3.2, the lower bound on the EOP, δ, is the probability of the set of symbols that deliver energy that is smaller than the required value of B. That is, $\delta \geqslant P_W(\{i \in \{1, 2, \ldots, M\} : e_i < B\})$, with P_W in (13). Hence, if there exists at most one codeword $\boldsymbol{u}(i)$, for some $i \in \{1, 2, \ldots, M\}$, such that $e_i < B$, then, the EOP is bounded away from zero. More specifically, $\delta \geqslant \frac{1}{M}$. If there are ρ codewords $\boldsymbol{u}(\ell_1)$, $\boldsymbol{u}(\ell_2)$, $\cdots$, $\boldsymbol{u}(\ell_\rho)$, such that for all $j \in \{1, 2, \ldots, \rho\}$, $e_{\ell_j} < B$, then, $\delta \geqslant \frac{\rho}{M}$. If all the codewords $\boldsymbol{u}(1)$, $\boldsymbol{u}(2)$, ..., $\boldsymbol{u}(M)$ are such that $\max_i e_i < B$, then the EOP is equal to one. That is, it is impossible to guarantee that B units of energy will be harvested at the EH. This reveals the entanglement between B and δ.

To quantify the relationship between B and δ, denote by $M' \leq M$, the number of unique values in the vector $(e_1, e_2, \ldots, e_M)^\mathsf{T}$ with e_i in (21). These M' unique energy levels are represented by $\{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_{M'}\} \subset \mathbb{R}$. More specifically, for all $i \in \{1, 2, \ldots, M\}$, there exists $j \in \{1, 2, \ldots, M'\}$ such that $e_i = \bar{e}_j$.

Assume without loss of generality that the following holds:

$$0 < \bar{e}_1 < \bar{e}_2 < \ldots < \bar{e}_{M'}. \tag{40}$$

Using this notation, the following lemma provides an upper bound on the energy requirement B for any $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code as a function of the EOP δ.

Lemma 3.3. *Given an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ for the random transformation in (9), with $\mathcal{X}$ in (3), let $j^+ \in \mathbb{N}$ be defined as follows.*

$$j^+ \triangleq \min \left\{ j \in \{1, \ldots M'\} : \delta \leqslant \frac{\sum_{k=1}^{j} \sum_{i=1}^{M} \mathbb{1}_{\{e_i = \bar{e}_k\}}}{M} \right\}, \tag{41}$$

where, the positive integer M' and the reals $\bar{e}_1$, $\bar{e}_2$, ..., $\bar{e}_{M'}$ are in (40). Then, the following holds for the energy requirement B:

$$B \leq \bar{e}_{j^+} \tag{42}$$

Proof. The proof is given in Appendix C. □

C. Information Transmission Rate

Though the bound on $R(\mathscr{C})$ in (87) is tight, it can quickly become computationally infeasible as the block-length n increases. The following theorem provides a tractable approximation of this bound.

Theorem 3.4. *If a given constant composition $(n, M, \mathcal{X})$-code, $\mathscr{C}$, with $\mathcal{X}$ in (3), is an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code for the random transformation in (9), then the following holds for $R(\mathscr{C})$ in (17):*

$$R(\mathscr{C}) \leq H(P_\mathscr{C}) + \frac{1}{n^2} \left(\frac{1}{12} - \sum_{\ell=1}^{L} \frac{1}{12 P_\mathscr{C}(x^{(\ell)}) + 1} \right) \tag{43}$$

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

where, $P_\mathscr{C}$ is the type defined in (32); and L is the number of channel input symbols in (3).

Proof. The proof is given in Appendix B. $\qquad\square$

Note that all terms in (43), except the entropy $H(P_{\mathscr{C}})$, vanish with the block-length n. This implies the well known information theoretic result that information rate is essentially constrained by the entropy of the channel input symbols. In particular, note that $H(P_{\mathscr{C}}) \leqslant \log_2 L$. Furthermore, this bound holds with equality for a code $\mathscr{C}$ when the type $P_{\mathscr{C}}$ is uniform, *i.e.*, all the symbols from the set of channel inputs $\mathcal{X}$ are used with the same frequency in $\mathscr{C}$. It is important to note that, even though it may not be explicitly clear in Theorem 3.1 and Theorem 3.4, the inequalities on the information rate and the DEP are in fact dependent on each other. Each of the inequalities are a function of the type $P_{\mathscr{C}}$, which defines the dependence between these parameters.

Remark 3.1. *The value $\frac{n!}{\prod_{\ell=1}^{L}(n P_{\mathscr{C}}(x^{(\ell)}))!}$ in (87) is the exact number of different codewords of length n, while exhibiting an average type $P_{\mathscr{C}}$. Hence, codes such that every channel input symbol is used the same number of times among all the codewords 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 reduce the energy transmission rate significantly. 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 $|x_1| < |x_2|$. Then, from (35), it is clear that using the symbol x_1 equally often as x_2 constraints the value of $e_{\mathscr{C}}$ and hence, the energy B. Alternatively, codes that exhibit the largest energy rates are those in which the symbols that have the largest magnitude are used most often. This clearly deviates from the uniform distribution and thus, constrains the information rate R.*

4. CODE CONSTRUCTION FOR SIET

This section introduces a method for constructing codes for SIET that ensure certain reliability guarantees. The tuples of information rate, energy rate, DEP and EOP that can be achieved by the constructed codes are also characterized. This is accomplished in two steps. The first step is the construction of the constellation $\mathcal{X}$ defined in (3) and the construction of a particular $(n, M, \mathcal{X})$-code, whose codewords and decoding sets are subject to certain conditions such that such the code is an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code.

A. Constellation Design

Consider a constellation represented by the set $\mathcal{X}$ in (3) and assume that $\mathcal{X}$ is the union of $C \in \mathbb{N}$ sets of channel input symbols, which will be referred to as *layers*. A layer is a set of symbols that have the same magnitude. For all $c \in \{1, 2, \ldots, C\}$, denote by $L_c \in \mathbb{N}$ the number of symbols and by $A_c \in \mathbb{R}^+$ the amplitude of the symbols in layer c. The underlying assumption is that symbols in layer c are equally spaced along the circle of radius A_c. This is typically the case in phase-shift keying (PSK) modulations [43, Chapter 16]. This is because, given the number of symbols in a layer, the symbols being equally spaced is the most favorable in terms of the DEP, as formally shown later in the proof of the main results of this section.

Let $\alpha_c \in [0, 2\pi]$ denote the phase shift of the symbols in layer c. The layer c is denoted as

$$\mathcal{U}(A_c, L_c, \alpha_c) \triangleq \{x_c^{(1)}, x_c^{(2)}, \ldots, x_c^{(L_c)}\} \tag{44}$$

$$= \left\{ x \in \mathbb{C} : x = A_c \exp\left(\mathrm{i}\left(\frac{2\pi}{L_c}\ell + \alpha_c\right)\right), \right.$$

$$\left. \ell \in \{0, 1, 2, \ldots, (L_c - 1)\} \right\}, \tag{45}$$

where i is the complex unit. Hence, the set $\mathcal{X}$ is given by

$$\mathcal{X} = \bigcup_{c=1}^{C} \mathcal{U}(A_c, L_c, \alpha_c). \tag{46}$$

The vector of the amplitudes A_c in (46) is denoted by

$$\boldsymbol{A} = (A_1, A_2, \ldots, A_C)^{\mathsf{T}}, \tag{47}$$

and without any loss of generality, assume that

$$A_1 > A_2 > \ldots > A_C. \tag{48}$$

The vector of the number of symbols in each layer in (46) is denoted by

$$\boldsymbol{L} = (L_1, L_2, \ldots, L_C)^{\mathsf{T}}; \tag{49}$$

and the vector of the phase shifts of the symbols in each layer in (46) is denoted by

$$\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_C)^{\mathsf{T}}. \tag{50}$$

The total number of symbols L in (3) satisfies

$$L = \sum_{c=1}^{C} L_c. \tag{51}$$

Using the design in (46) for the set of channel input symbols, $\mathcal{X}$ in (3), the $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ is constructed in two steps. The first step consists of designing the codewords, while the second step designs the corresponding decoding regions.

B. Codeword Design

Consider an $(n, M, \mathcal{X})$-code of the form in (12), with $\mathcal{X}$ in (46), whose codewords are all different and are constructed satisfying a constraint on the number of times the symbols in the different layers are used. For all $c \in \{1, 2, \ldots, C\}$, let p_c be the empirical probability of symbols of the cth layer. That is,

$$p_c = \frac{1}{Mn} \sum_{\ell=1}^{L_c} \sum_{i=1}^{M} \sum_{m=1}^{n} \mathbb{1}_{\{u_m(i)=x_c^{(\ell)}\}} \tag{52}$$

$$= \frac{1}{M} \sum_{\ell=1}^{L_c} \sum_{i=1}^{M} P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right) \tag{53}$$

$$= \sum_{\ell=1}^{L_c} P_{\mathscr{C}}\left(x_c^{(\ell)}\right), \tag{54}$$

where, for all $i \in \{1, 2, \ldots, M\}$ and for all $m \in \{1, 2, \ldots, n\}$, the complex $u_m(i)$ is the mth channel input symbol of the codeword $\boldsymbol{u}(i)$ in (11a); the types $P_{\boldsymbol{u}(i)}$ and $P_{\mathscr{C}}$ are defined in (31) and in (32), respectively; and $x_c^{(\ell)}$ is in (44). The resulting probability vector is denoted by

$$\boldsymbol{p} = (p_1, p_2, \ldots, p_C)^{\mathsf{T}}, \tag{55}$$

which is a free parameter in the construction of code $\mathscr{C}$. The constraint on the codewords of code $\mathscr{C}$ is that 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 $\mathscr{C}$ satisfies

$$P_{\mathscr{C}}(x_c^{(\ell)}) = \frac{p_c}{L_c}. \tag{56}$$

That is, symbols within the same layer are used with the same frequency in $\mathscr{C}$. The motivation for this assumption is that intuitively, while using two symbols within the same layer with different frequencies does not impact the energy transmission rate B or EOP δ, it does have an impact on the information transmission rate and DEP of $\mathscr{C}$.

C. Decoding Regions

The focus is now on the design of the decoding regions $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_M$ of the code $\mathscr{C}$, for the codewords $\boldsymbol{u}(1), \boldsymbol{u}(2), \ldots, \boldsymbol{u}(M)$ constructed above. For all $c \in \{1, 2, \ldots, C\}$ and $\ell \in \{1, 2, \ldots, L_c\}$, let the set $\mathcal{G}_c^{(\ell)} \subseteq \mathbb{C}$ be a circle of radius $r_c \in \mathbb{R}^+$ centered at $x_c^{(\ell)}$, with $x_c^{(\ell)}$ in (44). That is,

$$\mathcal{G}_c^{(\ell)} = \left\{ y \in \mathbb{C} : \left| y - x_c^{(\ell)} \right|^2 \leq r_c^2 \right\}. \tag{57}$$

For all $i \in \{1, 2, \ldots, M\}$, the decoding region $\mathcal{D}_i$ for codeword $\boldsymbol{u}(i)$ is built as follows:

$$\mathcal{D}_i = \mathcal{D}_{i,1} \times \mathcal{D}_{i,2} \times \ldots \times \mathcal{D}_{i,n}, \tag{58}$$

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

$$\mathcal{D}_{i,m} = \mathcal{G}_c^{(\ell)}, \tag{59}$$

with c and ℓ satisfying $u_m(i) = x_c^{(\ell)}$.

To ensure that the decoding regions $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_M$ are mutually disjoint, for all $c \in \{1, 2, \ldots, C-1\}$, the amplitudes A_c in (44) and the radii r_c in (57) satisfy the following:

$$A_{c+1} - A_c \geq r_{c+1} + r_c. \tag{60}$$

The vector of the radii in (57) is denoted by

$$\boldsymbol{r} = (r_1, r_2, \ldots, r_C)^{\mathsf{T}}. \tag{61}$$

Note that there are multiple ways of choosing the parameters $\boldsymbol{A}, \boldsymbol{\alpha}, \boldsymbol{p}$ and $\boldsymbol{r}$ in the proposed construction to satisfy the same information rate, energy rate, DEP and EOP requirements. Thus, the construction defines a family of codes

$$\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right), \tag{62}$$

with the number of layers C in (46); $\boldsymbol{A}$ in (47); $\boldsymbol{L}$ in (49); $\boldsymbol{\alpha}$ in (50); $\boldsymbol{p}$ in (55); and $\boldsymbol{r}$ in (61). The codewords $\boldsymbol{u}(1)$, $\boldsymbol{u}(2)$, ..., $\boldsymbol{u}(M)$ and their respective decoding regions $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_M$ completely define an $(n, M, \mathcal{X})$-code within the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$.

D. Achievable Tuples

The following theorem characterizes the tuples of information rate, energy rate, DEP and EOP that can be achieved by codes within the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in (62).

Theorem 4.1. *A code $\mathscr{C}$ of the form in* (12) *from the family* $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ *in* (62) *is an* $(n, M, \mathcal{X}, \epsilon, B, \delta)$-*code for the random transformation in* (9)*, if the following inequalities are simultaneously satisfied:*

$$\epsilon \geq 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \left(1 - \exp\left(-\frac{r_c^2}{\sigma^2}\right)\right)^{n \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)}, \tag{63a}$$

$$R(\mathscr{C}) \leq \log_2 \left(\sum_{c=1}^{C} \left\lfloor \frac{\pi}{2 \arcsin \frac{r_c}{2 A_c}} \right\rfloor \right), \tag{63b}$$

$$\delta \geq \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\left\{\sum_{c=1}^{C} \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)\left(k_1 A_c^2 + k_2 A_c^4\right) < \frac{B}{n}\right\}}, \tag{63c}$$

$$B \leq \bar{e}_{j+}, \tag{63d}$$

where, the information rate $R(\mathscr{C})$ is defined in (87)*; the energy level $\bar{e}_{j+}$ is defined in* (42)*; and the constants k_1 and k_2 are the same as in* (21)*; and the type $P_{\boldsymbol{u}(i)}$ is defined in* (31)*.*

Proof. The proofs are presented in Appendix D. □

Theorem 4.1 characterizes the conditions on the parameters $C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}$ and $\boldsymbol{r}$ such that all codes within the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in (62) are $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes. Hence, given some target information transmission rate R; energy requirement B; DEP ϵ; and EOP δ, Theorem 4.1 provides the conditions on the parameters $(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ for the existence of an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code. Such a code can be constructed by designing the codewords and the decoding regions as suggested in Section 4-B and Section 4-C, respectively.

Remark 4.1. *It is essential to note the following about the choice of decoding sets in* (57) *and the derived bound on the achievable DEP in* (63a)*. Firstly, the circular construction of decoding sets in* (57) *is a choice made in order to obtain closed form tractable expressions for the lower bound on the DEP in* (63a)*. This is enabled by the inherent circular symmetry of the complex AWGN noise. The bound in* (63a) *reveals the dependence of the DEP on the various parameters of the code such as the type. This is instructive for studying the trade-offs between the parameters of the constructed family of codes. Secondly, this choice of decoding sets provides guidelines for constructing the set of channel input symbols $\mathcal{X}$ while ensuring that the DEP does not exceed a required value which is controlled using the choice of the radii of the decoding sets in* (61)*.*

5. Performance Analysis

The performance of the constructed family of codes can be evaluated by comparing the achievability bounds identified for the constructed family of codes $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in Theorem 4.1 with the necessary conditions in Theorem 3.1 that any $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code must satisfy. This comparison is presented in the form of information-energy regions in Fig. 1 where the information transmission rates are plotted as a function of the energy transmission rates of the codes. For this comparison, consider constant composition $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes in the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in (62) that employ the set of channel input symbols $\mathcal{X}$ of the form in (46) with number of layers $C = 3$. The duration of the transmission in channel uses is $n = 80$. 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 (57). 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 to satisfy (60). The points on the curves in Fig. 1 are obtained by varying ϵ and the probability vector $\boldsymbol{p}$ in (55).

The first important observation from Fig. 1 is that the achievability results for the codes constructed in this work match the necessary conditions identified in Theorem 3.1 except for the DEP ϵ. This is illustrated by the overlap of the necessary conditions and the achievable information-energy curves. However, for the same information rate and energy pair, the DEP

for the achievable curves is higher than the DEP bound identified by the necessary conditions. The sub-optimality in DEP is a result of the choice of circular decoding regions in (57) (See Remark 4.1).

Fig. 1 also shows a clear trade-off between the information and energy transmission rates that can be simultaneously supported by a given code. The maximum achievable information transmission rate is $R = 4.93$ 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 that can be delivered at $R = 4.93$ bits/channel use is $B = 3.2 \times 10^5$ energy units. This corresponds to the point D_1 in Fig. 1. 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 (55) 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 $R = 3.8$ bits/channel use. This corresponds to the point D_2 in Fig. 1. The curves between the points D_1 and D_2 in Fig. 1 illustrate the trade-off between the information and energy transmission. As B is increased from 3.2×10^5 energy units at point D_1, R begins to decreases. Similarly, as R is increased from 3.8 bits/channel use at point D_2, B begins to decrease.
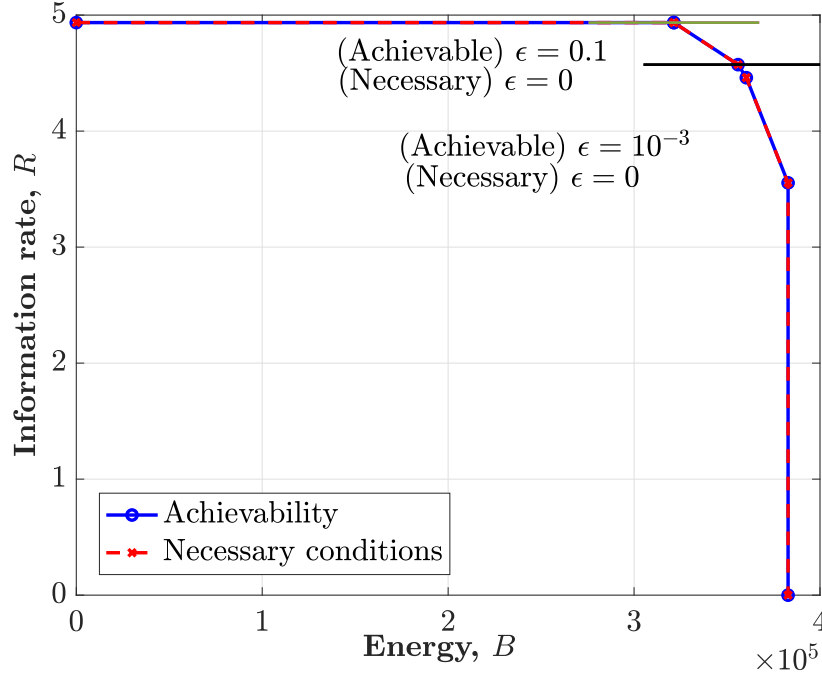

Fig. 1: A comparison of the information-energy region defined by the necessary conditions (Theorem 3.1) versus the achievable information-energy region (Theorem 4.1) for constant composition codes in the family $\mathsf{C}\,(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$.

6. Discussion

1) Trade-offs: The results presented in this work reveal several interesting insights into the trade-offs between the information transmission rate R, the energy requirement B, the DEP ϵ and the EOP δ. For instance, as shown by (36c) and (63c), δ increases as B increases and vice versa. The consequence of this relationship is that a smaller δ can be achieved at the cost of lower B. Similarly, for increasing B in (36d) and (63d), a larger value of δ has to be tolerated. These relationships also reveal that both B and δ can be improved by a code that has higher values of e_i in (21). This is achieved by using the symbols with greater energy more frequently in the code.

From the necessary condition in (36a), it follows that ϵ decreases as the distance between the symbols in the set of channel inputs $\mathcal{X}$ in (3) increases. This implies that, a lower DEP can be achieved by increasing the distances between the channel input symbols. However, with the peak-amplitude constraint in place, increasing the distance between symbols implies that the number of symbols L decreases. This in turn, decreases the upper bound on the information rate R in (87). In (63c), ϵ decreases as a function of the radii r_c of the decoding regions in (57). On the other hand, from (63b) it follows that the information rate R increases as the radii r_c decrease. These trade-offs between R, B, ϵ, and δ are more comprehensively illustrated in the following example.

Consider a family of constant composition $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes $\mathscr{C}$ in the family $\mathsf{C}\,(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62). The set of channel input symbols $\mathcal{X}$ in (46) is composed of two layers with 5 symbols in each layer, *i.e.*, $C = 2$ and $L_1 = L_2 = 5$. The
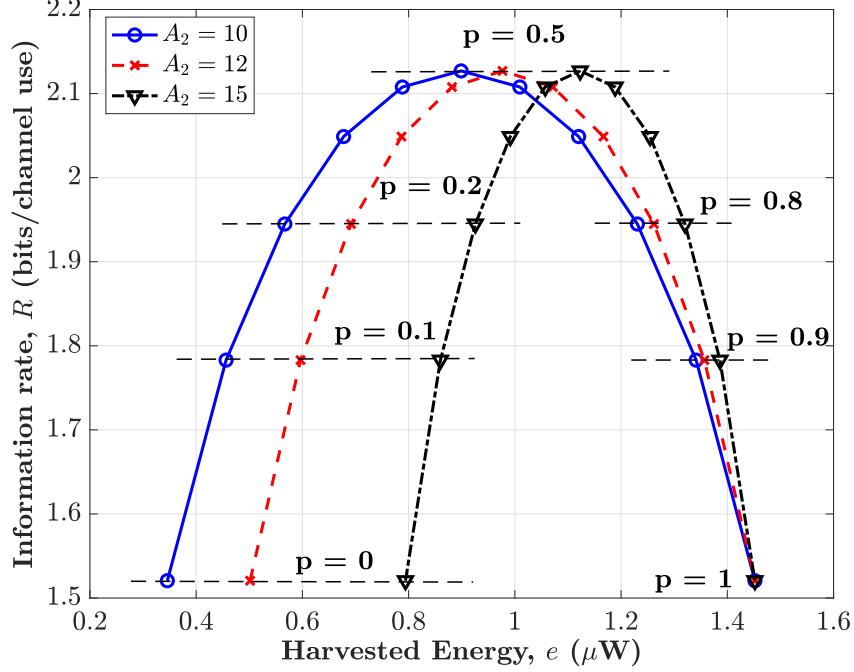
Fig. 2: Bounds on the information transmission rate R in (87) as a function of the harvested energy e in (65).

radius of the first layer is $A_1 = 20$ millivolts and the radius of the second layer A_2 is varied to illustrate the trade-offs between the various parameters. 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 (55) is given by

$$\boldsymbol{p} = (p, (1-p))^{\mathsf{T}}. \tag{64}$$

The duration of the transmission in channel uses is $n = 100$. Since $\mathscr{C}$ is a constant composition code, from (34), it holds that, for all $i \in \{1, 2, \ldots, M\}$,

$$e_i = e, \tag{65}$$

where $e \in [0, \infty)$ is calculated as in (35).

Fig. 2 shows the trade-offs between the the information transmission rate R in (87) and the harvested energy e in microwatts (μW) in (65) as a function of p in (64). 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 information rate R first increases and then decreases as a function of e in (65). For each of these curves, the maximum $R = 2.13$ bits/channel use corresponds to the 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 in (87) is only a function of the type. 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 in turn increases e.

2) Comparison with state of the art: Fig. 3 compares the bound on the information rate in (87) (Necessary 1) and (43) (Necessary 2) with the finite block-length converse bound on the information rate in [44, Theorem 54] for a QPSK (*i.e.*, $L = 4$) constellation and signal to noise ratio (SNR) equal to 18 dB. It should be noted that the converse bounds in [44] are concerned with information transmission alone. The following observations can be made from these plots. Firstly, the proposed bounds are very close to the finite block-length converse for small block-lengths. As the block-length increases, the difference between the two bounds increases. Secondly, the bounds in (87) and (43) are very close. This shows that the approximation in (43) is tight, especially as the block-length increases. Finally, it should be noted that the bounds in (87) and (43) are for the case of constant composition codes whereas the bound from [44] has no such restriction. In fact, the information rate R can be at most $\log_2 L$ in this case. As can be observed from Fig. 3, the proposed bounds actually come very close to the maximum value of $\log_2 L$ as the block-length n increases.
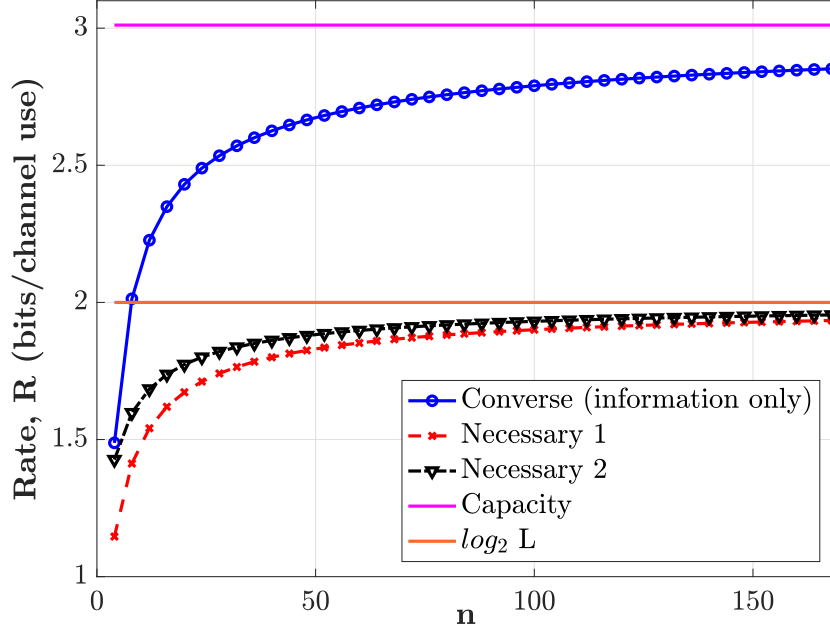
Fig. 3: A comparison of the bounds on the information rate from the necessary conditions in (87) (Necessary 1) and (43) (Necessary 2) for finite block-length SIET with the finite block-length converse in [44] as a function of the block-length n.

REFERENCES

[1] N. Tesla, *Apparatus for Transmitting Electrical Energy*. New York, NY: United States Patent Office, Dec. 1914, vol. US1119732 A.

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

[3] A. Iqbal, A. Smida, M. Al-Hasan, I. B. Mabrouk, and T. A. Denidni, "Highly Isolated Wireless Power Transfer and Information Co-Delivery Using a Pacemaker Duplex Antenna," *IEEE Transactions on Microwave Theory and Techniques*, p. 1–13, 2024.

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

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

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

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

[8] N. Khalfet and S. M. Perlaza, "Simultaneous Information and Energy Transmission in the Two-User Gaussian Interference Channel," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 1, pp. 156 –170, Jan. 2019.

[9] 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.

[10] S. B. Amor, S. M. Perlaza, I. Krikidis, and H. V. Poor, "Feedback Enhances Simultaneous Wireless Information and Energy Transmission in Multiple Access Channels," *IEEE Transactions on Information Theory*, vol. 63, no. 8, pp. 5244–5265, 2017.

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

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

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

[14] 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 *Proceedings of the IEEE International Symposium on Information Theory (ISIT)*, Espoo, Finland, Jun. 2022, pp. 2106–2111.

[15] N. Khalfet, C. Psomas, S. Chatzinotas, and I. Krikidis, "Semantic Communications for Simultaneous Wireless Information and Power Transfer," *IEEE Transactions on Communications*, pp. 1–1, 2024.

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

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

[18] 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 *Proceedings of the IEEE Information Theory Workshop (ITW)*, Mumbai, India, Nov. 2022.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

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

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

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

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

[35] 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.

[36] 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.

[37] 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.

[38] 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.

[39] 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.

[40] C. Song, C. Ling, J. Park, and B. Clerckx, "MIMO Broadcasting for Simultaneous Wireless Information and Power Transfer: Weighted MMSE Approaches," in *Proceedings of the IEEE Globecom Workshops (GC Wkshps)*, 2014, pp. 1151–1156.

[41] 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.

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

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

[44] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Channel Coding Rate in the Finite Blocklength Regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, 2010.

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

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

APPENDIX

A. Proofs for Theorem 3.1

Theorem A.1. *Given a constant composition* $(n, M, \mathcal{X}, \epsilon, B, \delta)$*-code* $\mathscr{C}$ *for the random transformation in* (9) *of the form in* (12)*, for all* $\ell \in \{1, 2, \ldots, L\}$*, let the complex* $\bar{x}^{(\ell)} \in \mathcal{X}$ *in* (3) *be such that*

$$\bar{x}^{(\ell)} \in \arg \max_{x \in \mathcal{X} \backslash \{x^{(\ell)}\}} \left| x^{(\ell)} - x \right|. \tag{66}$$

Then, ϵ *in* (29) *must satisfy the following:*

$$\epsilon \geq (M-1)\mathrm{Q}\left(\sqrt{\frac{\sum_{\ell=1}^{L} n P_{\mathscr{C}}(x^{(\ell)}) \left| x^{(\ell)} - \bar{x}^{(\ell)} \right|^2}{2\sigma^2}} \right), \tag{67}$$

where, $P_{\mathscr{C}}$ *is the type defined in* (32)*;* L *is the number of symbols in* (3)*, and the real* σ^2 *is the noise variance in* (9b)*.*

Proof. To determine the lower bound on the DEP, the decoding regions $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_M$ are defined using the MAP decision rule [43, Chapter 21] as in (38). Using the definition of $f_{\boldsymbol{Y}|\boldsymbol{X}}$ in (9b) followed simple mathematical manipulations, the decision rule in (38) simplifies to the minimum distance decoder [45, Chapter 4] which decides that message i was transmitted if

$$i \in \arg \min_{i' \in \{1, 2, \ldots, M\}} \left| \boldsymbol{y} - \boldsymbol{u}(i') \right|. \tag{68}$$

The decoding error probability $\gamma_i(\mathscr{C})$ in (15) for the code $\mathscr{C}$, given that the message index $i \in \{1, 2, \ldots, M\}$ is transmitted is

$$\gamma_i(\mathscr{C}) = \Pr\left(\boldsymbol{Y} \notin \mathcal{D}_i \big| \boldsymbol{u}(i) \right) \tag{69}$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{M} \Pr\left(\left| \boldsymbol{Y} - \boldsymbol{u}(j) \right|^2 < \left| \boldsymbol{Y} - \boldsymbol{u}(i) \right|^2 \big| \boldsymbol{u}(i) \right) \tag{70}$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{M} \Pr\left(\left| \boldsymbol{u}(i) + \boldsymbol{N} - \boldsymbol{u}(j) \right|^2 < \left| \boldsymbol{u}(i) + \boldsymbol{N} - \boldsymbol{u}(i) \right|^2 \right) \tag{71}$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{M} \Pr\left(\left| \boldsymbol{u}(i) + \boldsymbol{N} - \boldsymbol{u}(j) \right|^2 < \left| \boldsymbol{N} \right|^2 \right) \tag{72}$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{M} \Pr\left(\sum_{m=1}^{n} \left| u_m(i) + N_m - u_m(j) \right|^2 < \sum_{m=1}^{n} \left| N_m \right|^2 \right) \tag{73}$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{M} \Pr\Big(\sum_{m=1}^{n} (\Re(u_m(i) + N_m - u_m(j)))^2 +$$

$$(\Im(u_m(i) + N_m - u_m(j)))^2 < \sum_{m=1}^{n} \Re(N_m)^2 + \Im(N_m)^2 \Big) \tag{74}$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{M} \Pr\Big(\sum_{m=1}^{n} (\Re(u_m(i)) - \Re(u_m(j)))^2 + (\Im(u_m(i)) - \Im(u_m(j)))^2$$

$$+ 2\Re(N_m)(\Re(u_m(i)) - \Re(u_m(j))) +$$

$$2\Im(N_m)(\Im(u_m(i)) - \Im(u_m(j))) < 0 \Big) \tag{75}$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{M} \Pr\Big(\sum_{m=1}^{n} \Re(N_m)(\Re(u_m(i)) - \Re(u_m(j))) +$$

$$\Im(N_m)(\Im(u_m(i)) - \Im(u_m(j))) < -\frac{1}{2} \sum_{m=1}^{n} (\Re(u_m(i)) - \Re(u_m(j)))^2$$

$$+ (\Im(u_m(i)) - \Im(u_m(j)))^2 \Big), \tag{76}$$

where, the equality in (70) follows from (68), and $\boldsymbol{N} = (N_1, N_2, \ldots, N_n)^\mathsf{T}$ is the AWGN noise vector in (8) 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 $\frac{1}{2}\sigma^2$. Therefore, the random variable

$$\overline{N}_{i,j} = \sum_{m=1}^{n} \Re(N_m)(\Re(u_m(i)) - \Re(u_m(j)))$$
$$+ \Im(N_m)(\Im(u_m(i)) - \Im(u_m(j))) \tag{77}$$

in (76) is a linear combination of $2n$ Gaussian random variables, each with mean zero and variance $\frac{1}{2}\sigma^2$. Thus, $\overline{N}_{i,j}$ is also a zero mean Gaussian random variable with variance given by

$$\sigma_{\overline{N}_{i,j}}^2 = \frac{\sigma^2}{2} \sum_{m=1}^{n} (\Re(u_m(i)) - \Re(u_m(j)))^2 +$$
$$(\Im(u_m(i)) - \Im(u_m(j)))^2. \tag{78}$$

From (76),(77) and (78), it follows that,

$$\gamma_i(\mathscr{C}) = \sum_{\substack{j=1 \\ j \neq i}}^{M} \Pr\left(\overline{N}_{i,j} < -\frac{\sigma_{\overline{N}_{i,j}}^2}{\sigma^2}\right). \tag{79}$$

Since $\overline{N}_{i,j} \sim \mathcal{N}(0, \sigma_{\overline{N}_{i,j}}^2)$, from (79) it follows that

$$\gamma_i(\mathscr{C}) = \sum_{\substack{j=1 \\ j \neq i}}^{M} Q\left(\frac{\sigma_{\overline{N}_{i,j}}}{\sigma^2}\right) \tag{80}$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{M} Q\left(\sqrt{\frac{1}{2\sigma^2} \sum_{m=1}^{n} (\Re(u_m(i)) - \Re(u_m(j)))^2 + (\Im(u_m(i)) - \Im(u_m(j)))^2}\right) \tag{81}$$

$$= \sum_{\substack{j=1 \\ j \neq i}}^{M} Q\left(\sqrt{\frac{\sum_{m=1}^{n} |u_m(i) - u_m(j)|^2}{2\sigma^2}}\right), \tag{82}$$

where, the Q function is in (2). For all $i \in \{1, 2, \ldots, M\}$ and all $m \in \{1, 2, \ldots, n\}$, $u_m(i) \in \mathcal{X} = \{x^{(1)}, x^{(2)}, \ldots, x^{(L)}\}$ in (3). For $u_m(i) = x^{(\ell)}$, from (66) it follows that

$$|u_m(i) - u_m(j)| \leq \left|x^{(\ell)} - \bar{x}^{(\ell)}\right| \tag{83}$$

Since $\mathscr{C}$ is a constant composition code, from (31) and Definition 3.1, it follows that, for all $i \in \{1, 2, \ldots, M\}$, the symbol $x^{(\ell)} \in \mathcal{X}$ appears in the codeword $\boldsymbol{u}(i)$, $nP_{\mathscr{C}}(x^{(\ell)})$ number of times. Therefore, from (82) and (83), it follows that,

$$\gamma_i(\mathscr{C}) \geq \sum_{\substack{j=1 \\ j \neq i}}^{M} Q\left(\sqrt{\frac{\sum_{\ell=1}^{L} nP_{\mathscr{C}}(x^{(\ell)}) \left|x^{(\ell)} - \bar{x}^{(\ell)}\right|^2}{2\sigma^2}}\right)$$
$$= (M-1)Q\left(\sqrt{\frac{\sum_{\ell=1}^{L} nP_{\mathscr{C}}(x^{(\ell)}) \left|x^{(\ell)} - \bar{x}^{(\ell)}\right|^2}{2\sigma^2}}\right). \tag{84}$$

The average DEP $\gamma(\mathscr{C})$ in (16) is given by

$$\gamma(\mathscr{C}) \geq \frac{1}{M} \sum_{i=1}^{M} (M-1)Q\left(\sqrt{\frac{\sum_{\ell=1}^{L} nP_{\mathscr{C}}(x^{(\ell)}) \left|x^{(\ell)} - \bar{x}^{(\ell)}\right|^2}{2\sigma^2}}\right)$$
$$= (M-1)Q\left(\sqrt{\frac{\sum_{\ell=1}^{L} nP_{\mathscr{C}}(x^{(\ell)}) \left|x^{(\ell)} - \bar{x}^{(\ell)}\right|^2}{2\sigma^2}}\right). \tag{85}$$

Finally, from (29) and (85), it follows that

$$\epsilon \geq (M-1)Q\left(\sqrt{\frac{\sum_{\ell=1}^{L} nP_{\mathscr{C}}(x^{(\ell)})\left|x^{(\ell)} - \bar{x}^{(\ell)}\right|^2}{2\sigma^2}}\right), \tag{86}$$

which completes the proof. $\qquad\square$

The following lemma proves the bound on the information rate $R(\mathscr{C})$ in (87).

Lemma A.2. *Given a constant composition* $(n, M, \mathcal{X}, \epsilon, B, \delta)$*-code* $\mathscr{C}$ *for the random transformation in* (8) *of the form in* (12)*, the information transmission rate* $R(\mathscr{C})$ *in* (17) *is such that*

$$R(\mathscr{C}) \leq \frac{1}{n}\log_2\left(\frac{n!}{\prod_{\ell=1}^{L}(nP_{\mathscr{C}}(x^{(\ell)}))!}\right), \tag{87}$$

where, $P_{\mathscr{C}}$ *is the type defined in* (32).

Proof. Given a code type $P_{\mathscr{C}}$ that satisfies (33), the number of codewords that can be constructed is given by

$$M = \binom{n}{nP_{\mathscr{C}}(x^{(1)})}\binom{n - nP_{\mathscr{C}}(x^{(1)})}{nP_{\mathscr{C}}(x^{(2)})}\cdots\binom{n - \sum_{\ell=1}^{L-1} nP_{\mathscr{C}}(x^{(\ell)})}{nP_{\mathscr{C}}(x^{(L)})}$$
$$= \frac{n!}{\prod_{\ell=1}^{L}(nP_{\mathscr{C}}(x^{(\ell)}))!}. \tag{88}$$

Therefore, the information rate $R(\mathscr{C})$ in (17) satisfies

$$R(\mathscr{C}) \leq \frac{1}{n}\log_2\left(\frac{n!}{\prod_{\ell=1}^{L}(nP_{\mathscr{C}}(x^{(\ell)}))!}\right), \tag{89}$$

which completes the proof. $\qquad\square$

The next result provides proof of the bound on δ in (36c).

Theorem A.3. *Given a constant composition* $(n, M, \mathcal{X}, \epsilon, B, \delta)$*-code* $\mathscr{C}$ *for the random transformation in* (9) *of the form in* (12)*, the following holds for* δ *in* (30)*:*

$$\delta \geq \mathbb{1}_{\{e_{\mathscr{C}} < B\}}, \tag{90}$$

where $e_{\mathscr{C}} \in [0, \infty)$ *is in* (34).

Proof. From (28), the average EOP for the $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ is given by

$$\theta(\mathscr{C}, B) = \frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\{e_i < B\}}. \tag{91}$$

Since $\mathscr{C}$ is a constant composition code, from (91) and (34) it follows that

$$\theta(\mathscr{C}, B) = \frac{1}{M}\sum_{i=1}^{M}\mathbb{1}_{\{e_{\mathscr{C}} < B\}} \tag{92}$$
$$= \mathbb{1}_{\{e_{\mathscr{C}} < B\}}. \tag{93}$$

From (30) and (93), it follows that

$$\delta \geq \mathbb{1}_{\{e_{\mathscr{C}} < B\}}. \tag{94}$$

This completes the proof. $\qquad\square$

B. Proof of Theorem 3.4

Proof. From (87), for the information transmission rate $R(\mathscr{C})$ of code $\mathscr{C}$, it holds that

$$R(\mathscr{C}) \leq \frac{1}{n} \log(n!) - \frac{1}{n} \sum_{\ell=1}^{L} \log\left(\left(nP_{\mathscr{C}}(x^{(\ell)})\right)!\right). \tag{95}$$

Using the Stirling's approximation [46] on the factorial terms yields

$$(nP_{\mathscr{C}}(x^{(\ell)}))! \geqslant \sqrt{2\pi} \left(nP_{\mathscr{C}}(x^{(\ell)})\right)^{nP_{\mathscr{C}}(x^{(\ell)})+\frac{1}{2}}$$

$$\exp\left(-nP_{\mathscr{C}}(x^{(\ell)}) + \frac{1}{12nP_{\mathscr{C}}(x^{(\ell)})+1}\right), \text{ and} \tag{96}$$

$$n! \leqslant \sqrt{2\pi} n^{n+\frac{1}{2}} \exp\left(-n + \frac{1}{12n}\right). \tag{97}$$

From (96) and (97), it follows that,

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

$$\log(nP_{\mathscr{C}}(x^{(\ell)})) - nP_{\mathscr{C}}(x^{(\ell)}) + \frac{1}{12nP_{\mathscr{C}}(x^{(\ell)})+1} \tag{98}$$

$$= \log\left(\sqrt{2\pi}\right) + nP_{\mathscr{C}}(x^{(\ell)}) \log(P_{\mathscr{C}}(x^{(\ell)})) + \frac{1}{2} \log(P_{\mathscr{C}}(x^{(\ell)}))$$

$$+ \left(nP_{\mathscr{C}}(x^{(\ell)}) + \frac{1}{2}\right) \log(n) - nP_{\mathscr{C}}(x^{(\ell)}) + \frac{1}{12nP_{\mathscr{C}}(x^{(\ell)})+1}, \tag{99}$$

and ,

$$\log(n!) \leq \log\left(\sqrt{2\pi}\right) + \left(n + \frac{1}{2}\right) \log(n) - n + \frac{1}{12n}$$

$$= n\log(n) - n + \frac{1}{12n} + \frac{1}{2} \log(2\pi n). \tag{100}$$

The sum in (95) satisfies,

$$\sum_{\ell=1}^{L} \log\left(\left(nP_{\mathscr{C}}(x^{(\ell)})\right)!\right) \geq L \log\left(\sqrt{2\pi}\right) - nH\left(P_{\mathscr{C}}\right)$$

$$+ \frac{1}{2} \sum_{\ell=1}^{L} \log(P_{\mathscr{C}}(x^{(\ell)})) + n\log(n) + \frac{L}{2} \log(n) - n$$

$$+ \sum_{\ell=1}^{L} \frac{1}{12nP_{\mathscr{C}}(x^{(\ell)})+1}. \tag{101}$$

Using (100) and (101) in (95) yields,

$$R(\mathscr{C}) \leq \log(n) - 1 + \frac{1}{12n^2} + \frac{1}{2n} \log(2\pi n)$$

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

$$- \frac{L}{2n} \log(n) + 1 - \frac{1}{n} \sum_{\ell=1}^{L} \frac{1}{12nP_{\mathscr{C}}(x^{(\ell)})+1} \tag{102}$$

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

$$\frac{1}{2n} \left(\log(2\pi n) - \sum_{\ell=1}^{L} \log(2\pi nP_{\mathscr{C}}(x^{(\ell)}))\right) \tag{103}$$

$$= H\left(P_{\mathscr{C}}\right) + \frac{1}{n^2} \left(\frac{1}{12} - \sum_{\ell=1}^{L} \frac{1}{12P_{\mathscr{C}}(x^{(\ell)})+1}\right) + \tag{104}$$

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

which completes the proof. $\qquad\square$

C. Proof of Lemma 3.3

Proof. Given that channel input $i \in \{1, 2, \ldots, M\}$ is transmitted, the harvested energy is given by the constant e_i in (21). Furthermore, given the set of unique energy levels $\{\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_{M'}\}$ in (40), there are a finite number of values that the EOP δ can take. More precisely,

$$\delta \in \left\{ \frac{\sum_{k=1}^{j} \sum_{i=1}^{M} \mathbb{1}_{\{e_i = \bar{e}_k\}}}{M}, j \in \{1, 2, \ldots, M'\} \right\}. \tag{105}$$

This implies that, for all $j \in \{1, 2, \ldots, M'\}$, for $\delta \leqslant \dfrac{\sum_{k=1}^{j} \sum_{i=1}^{M} \mathbb{1}_{\{e_i = \bar{e}_k\}}}{M}$, at most $\lfloor M\delta \rfloor = \sum_{k=1}^{j} \sum_{i=1}^{M} \mathbb{1}_{\{e_i = \bar{e}_k\}}$ codewords can have energy less than B. This is possible if and only if

$$B \leq \bar{e}_j. \tag{106}$$

Define $j^+ \in \mathbb{N}$ as

$$j^+ \triangleq \min \left\{ j \in \{1, \ldots M'\} : \delta \leqslant \frac{\sum_{k=1}^{j} \sum_{i=1}^{M} \mathbb{1}_{\{e_i = \bar{e}_k\}}}{M} \right\}, \tag{107}$$

where, the positive integer M' and the reals $\bar{e}_1$, $\bar{e}_2$, $\ldots$, $\bar{e}_{M'}$ are in (40).

Then, from (106) and (107), it follows that

$$B \leq \bar{e}_{j^+} \tag{108}$$

This completes the proof. $\qquad\square$

D. Proofs for Theorem 4.1

The following lemma provides a lower bound on the DEP of codes from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62).

Lemma A.4. *Consider an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62). The parameters $r_1, r_2, \ldots, r_C$ in (61) satisfy the following:*

$$\epsilon \geq 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \left(1 - \exp\left(-\frac{r_c^2}{\sigma^2} \right) \right)^{n \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{109}$$

where, the type $P_{\boldsymbol{u}(i)}$ is defined in (31), the real σ^2 is defined in (9b), and $x_c^{(\ell)} \in \mathcal{U}(A_c, L_c, \alpha_c)$, with $\mathcal{U}(A_c, L_c, \alpha_c)$ in (44).

Proof. From (15) and (16), the average DEP of the $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ is given by

$$\gamma(\mathscr{C}) = 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{\mathcal{D}_i} f_{\boldsymbol{Y}|\boldsymbol{X}}(\boldsymbol{y}|\boldsymbol{u}(i)) \mathrm{d}\boldsymbol{y} \tag{110}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \int_{\mathcal{D}_{i,1} \times \mathcal{D}_{i,2} \times \ldots \times \mathcal{D}_{i,n}} \prod_{m=1}^{n} f_{Y|X}(y|u_m(i)) \, \mathrm{d}y \tag{111}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{m=1}^{n} \int_{\mathcal{D}_{i,m}} f_{Y|X}(y|u_m(i)) \, \mathrm{d}y \tag{112}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \prod_{\ell=1}^{L_c} \left(\int_{\mathcal{G}_c^{(\ell)}} f_{Y|X}(y|x_c^{(\ell)}) \mathrm{d}y \right)^{n P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{113}$$

where, the equality in (111) follows due to (9) and (58), and (112) follows due to Fubini's theorem. Using (9b) in (113) yields,

$$\gamma\left(\mathscr{C}\right) = 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \prod_{\ell=1}^{L_c} \left(\int_{\mathcal{G}_c^{(\ell)}} \frac{1}{\pi\sigma^2} \right. \tag{114}$$

$$\left. \exp\left(-\frac{(\Re(y) - \Re(x_c^{(\ell)}))^2 + (\Im(y) - \Im(x_c^{(\ell)}))^2}{\sigma^2} \right) dy \right)^{nP_{\boldsymbol{u}(i)}(x_c^{(\ell)})}.$$

Evaluating the integral term in (114) for all $\ell \in \{1, 2, \ldots, L_c\}$ yields,

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

$$= \int_{\Im(x_c^{(\ell)})-r_c}^{\Im(x_c^{(\ell)})+r_c} \int_{\Re(x_c^{(\ell)})-\sqrt{r_c^2-(v-\Im(x_c^{(\ell)}))^2}}^{\Re(x_c^{(\ell)})+\sqrt{r_c^2-(v-\Im(x_c^{(\ell)}))^2}}$$

$$\frac{1}{\pi\sigma^2} \exp\left(-\frac{(u - \Re(x_c^{(\ell)}))^2 + (v - \Im(x_c^{(\ell)}))^2}{\sigma^2} \right) du dv, \tag{115}$$

$$= \int_{-r_c}^{r_c} \int_{-\sqrt{r_c^2-v^2}}^{\sqrt{r_c^2-v^2}} \frac{1}{\pi\sigma^2} \exp\left(-\frac{u^2 + v^2}{\sigma^2} \right) du dv, \tag{116}$$

$$= \int_{0}^{\pi} \int_{0}^{\frac{r_c}{\sigma}} \frac{1}{2\pi} \exp\left(-\zeta^2 \right) \zeta d\zeta d\eta, \tag{117}$$

$$= \left(1 - \exp\left(-\frac{r_c^2}{\sigma^2} \right) \right). \tag{118}$$

The equality in (117) is obtained from the change of variables $u = \sigma\zeta\cos\eta, v = \sigma\zeta\sin\eta$. Plugging (118) in (114) yields,

$$\gamma\left(\mathscr{C}\right) = 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \prod_{\ell=1}^{L_c} \left(1 - \exp\left(-\frac{r_c^2}{\sigma^2} \right) \right)^{nP_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{119}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \left(1 - \exp\left(-\frac{r_c^2}{\sigma^2} \right) \right)^{n \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}. \tag{120}$$

From (29), for $\mathscr{C}$ to be an $(n, M, \mathcal{X}, \epsilon)$-code, the following must hold:

$$\gamma\left(\mathscr{C}\right) \leq \epsilon. \tag{121}$$

This implies that,

$$\epsilon \geq 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \left(1 - \exp\left(-\frac{r_c^2}{\sigma^2} \right) \right)^{n \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{122}$$

which completes the proof. $\qquad\square$

The bound on ϵ given by Lemma A.4 provides the minimum value of the DEP that can be achieved by the constructed family of codes. This implies that any DEP requirement greater than this value can be satisfied by the constructed family of codes.

From Definition 3.1 and (56), it follows that, for a constant composition code $\mathscr{C}$ of the form in (??) from the family $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ in (62), it holds that,

$$P_{\boldsymbol{u}(i)}(x_c^{(\ell)}) = P_{\mathscr{C}}(x_c^{(\ell)}) = \frac{p_c}{L_c}. \tag{123}$$

The following result for constant composition codes follows from (123) and Lemma A.4.

Corollary A.5. *Consider a constant composition* $(n, M, \mathcal{X}, \epsilon, B, \delta)$-*code* $\mathscr{C}$ *from the family* $\mathsf{C}\left(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r}\right)$ *in (62). The parameters* $r_1, r_2, \ldots, r_C$ *in (61) satisfy the following:*

$$\epsilon \geq 1 - \prod_{c=1}^{C} \left(1 - \exp\left(-\frac{r_c^2}{\sigma^2} \right) \right)^{np_c}, \tag{124}$$

where, the real σ^2 is defined in (9b), and $x_c^{(\ell)} \in \mathcal{U}(A_c, L_c, \alpha_c)$, with $\mathcal{U}(A_c, L_c, \alpha_c)$ in (44).

In the following lemma, we provide an upper bound on the achievable information transmission rate $R(\mathscr{C})$ for a code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62). Proving Lemma A.6 requires some additional results that provide bounds on the radius of decoding regions r_c in (61) as well as the number of symbols in each layer of the constructed codes L_c in (49). In the interest of better readability, these results are presented as Lemmas A.11 and A.12 in Appendix E and we directly state the bound on the information transmission rate $R(\mathscr{C})$ here.

Lemma A.6. *Given an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62), the information transmission rate $R(\mathscr{C})$ satisfies the following:*

$$R(\mathscr{C}) \leq \log_2 \sum_{c=1}^{C} \left\lfloor \frac{\pi}{2 \arcsin \frac{r_c}{2A_c}} \right\rfloor, \tag{125}$$

where, for all $c \in \{1, 2, \ldots, C\}$, the radius r_c is in (61), and the amplitude A_c is in (44).

Proof. The largest number of codewords of length n that can be formed with L different channel input symbols is L^n. Hence, from (17), it follows that

$$R(\mathscr{C}) \leq \frac{\log_2 M}{n} \tag{126}$$

$$\leq \frac{\log_2 L^n}{n} \tag{127}$$

$$= \log_2 L \tag{128}$$

$$\leq \log_2 \sum_{c=1}^{C} \left\lfloor \frac{\pi}{2 \arcsin \frac{r_c}{2A_c}} \right\rfloor, \tag{129}$$

where, the inequality in (129) follows from Lemma A.12 and (51). This completes the proof. $\square$

The following lemma provides a bound on the information transmission rate for a constant composition $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code.

Lemma A.7. *For a constant composition $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62), the information transmission rate $R(\mathscr{C})$ is given by:*

$$R(\mathscr{C}) \leq \frac{1}{n} \log_2 \left(\frac{n!}{\prod_{c=1}^{C} \left(\left(n \frac{p_c}{L_c} \right)! \right)^{L_c}} \right). \tag{130}$$

Proof. For a constant composition $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ for which the type $P_{\mathscr{C}}$ satisfies (33), the number of codewords that can be constructed is given by

$$\begin{aligned}
M = & \binom{n}{nP_{\mathscr{C}}(x_1^{(1)})} \binom{n - nP_{\mathscr{C}}(x_1^{(1)})}{nP_{\mathscr{C}}(x_1^{(2)})} \cdots \binom{n - \sum_{\ell=1}^{L_1-1} nP_{\mathscr{C}}(x_1^{(\ell)})}{nP_{\mathscr{C}}(x_1^{(L_1)})} \\
& \times \binom{n - \sum_{\ell=1}^{L_1} nP_{\mathscr{C}}(x_1^{(\ell)})}{nP_{\mathscr{C}}(x_2^{(1)})} \binom{n - \sum_{\ell=1}^{L_1} nP_{\mathscr{C}}(x_1^{(\ell)}) - nP_{\mathscr{C}}(x_2^{(1)})}{nP_{\mathscr{C}}(x_2^{(2)})} \\
& \cdots \binom{n - \sum_{\ell=1}^{L_1} nP_{\mathscr{C}}(x_1^{(\ell)}) - \sum_{\ell=1}^{L_2-1} nP_{\mathscr{C}}(x_2^{(\ell)})}{nP_{\mathscr{C}}(x_2^{(L_2)})} \times \cdots \\
& \times \binom{n - \sum_{c=1}^{C-1} \sum_{\ell=1}^{L_c} nP_{\mathscr{C}}(x_c^{(\ell)})}{nP_{\mathscr{C}}(x_C^{(1)})} \binom{n - \sum_{c=1}^{C-1} \sum_{\ell=1}^{L_c} nP_{\mathscr{C}}(x_c^{(\ell)}) - nP_{\mathscr{C}}(x_C^{(1)})}{nP_{\mathscr{C}}(x_C^{(2)})} \\
& \cdots \binom{n - \sum_{c=1}^{C-1} \sum_{\ell=1}^{L_c} nP_{\mathscr{C}}(x_c^{(\ell)}) - \sum_{\ell=1}^{L_C-1} nP_{\mathscr{C}}(x_C^{(\ell)})}{nP_{\mathscr{C}}(x_C^{(L_C)})} \\
= & \binom{n}{n\frac{p_1}{L_1}} \binom{n - n\frac{p_1}{L_1}}{n\frac{p_1}{L_1}} \cdots \binom{n - \sum_{\ell=1}^{L_1-1} n\frac{p_1}{L_1}}{n\frac{p_1}{L_1}} \\
& \times \binom{n - np_1}{n\frac{p_2}{L_2}} \binom{n - np_1 - n\frac{p_2}{L_2}}{n\frac{p_2}{L_2}} \cdots \binom{n - np_1 - \sum_{\ell=1}^{L_2-1} n\frac{p_2}{L_2}}{n\frac{p_2}{L_2}} \times
\end{aligned} \tag{131}$$

$$\ldots \times \binom{n - \sum_{c=1}^{C-1} np_c}{n\frac{p_C}{L_C}} \binom{n - \sum_{c=1}^{C-1} np_c - n\frac{p_C}{L_C}}{n\frac{p_C}{L_C}} \ldots$$

$$\binom{n - \sum_{c=1}^{C-1} np_c - \sum_{\ell=1}^{L_C-1} n\frac{p_C}{L_C}}{n\frac{p_C}{L_C}} \tag{132}$$

$$= \frac{n!}{\prod_{c=1}^{C} \left(\left(n\frac{p_c}{L_c}\right)!\right)^{L_c}} \tag{133}$$

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

$$R(\mathscr{C}) \leq \frac{\log_2 M}{n} = \frac{1}{n} \log_2 \left(\frac{n!}{\prod_{c=1}^{C} \left(\left(n\frac{p_c}{L_c}\right)!\right)^{L_c}} \right). \tag{134}$$

This completes the proof. $\qquad\square$

The following lemma provides the achievable bound on the EOP δ in (63c) for codes from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$.

Lemma A.8. *For an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62), the following holds:*

$$\delta \geq \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\left\{ \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)(k_1 A_c^2 + k_2 A_c^4) < \frac{B}{n} \right\}}, \tag{135}$$

where, k_1 and k_2 are positive real constants defined in (21) and $P_{\boldsymbol{u}(i)}$ is the type in (31).

Proof. From (44), 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 (44) are given by

$$x_c^{(\ell)} = A_c \exp\left(\mathrm{i} \left(\frac{2\pi}{L_c} \ell + \alpha_c \right) \right). \tag{136}$$

This implies that

$$\left| x_c^{(\ell)} \right| = A_c. \tag{137}$$

Using (28) and (21), the EOP for the $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ is given by:

$$\theta(\mathscr{C}, B) = \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\left\{ \left(k_1 \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} n P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right) \left| x_c^{(\ell)} \right|^2 + k_2 \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} n P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right) \left| x_c^{(\ell)} \right|^4 \right) < B \right\}} \tag{138}$$

$$= \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\left\{ \left(k_1 \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} n P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right) A_c^2 + k_2 \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} n P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right) A_c^4 \right) < B \right\}}. \tag{139}$$

From (30) and (139), the code $\mathscr{C}$ is an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code if the following holds:

$$\frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\left\{ \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}\left(x_c^{(\ell)}\right)(k_1 A_c^2 + k_2 A_c^4) < \frac{B}{n} \right\}} \leq \delta. \tag{140}$$

This completes the proof. $\qquad\square$

The achievable bound on the EOP δ in (30) for a constant composition code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62) is given by the following lemma:

Lemma A.9. *For a constant composition $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62), for all $c \in \{1, 2, \ldots, C\}$, the parameters p_c in (52) satisfy the following:*

$$\delta = \mathbb{1}_{\left\{ \left(k_1 \sum_{c=1}^{C} np_c A_c^2 + k_2 \sum_{c=1}^{C} np_c A_c^4 \right) < B \right\}}, \tag{141}$$

where, k_1 and k_2 are positive real constants defined in (21).

Proof. From (28) and (34), the EOP for the constant composition code $\mathscr{C}$ is given by:

$$\theta(\mathscr{C}, B) = \frac{1}{M} \sum_{i=1}^{M} \mathbb{1}_{\left\{ \begin{array}{l} (k_1 \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} n P_{\mathscr{C}}(x_c^{(\ell)}) |x_c^{(\ell)}|^2 + \\ k_2 \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} n P_{\mathscr{C}}(x_c^{(\ell)}) |x_c^{(\ell)}|^4) < B \end{array} \right\}} \tag{142}$$

$$= \mathbb{1}_{\left\{ \begin{array}{l} (k_1 \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} n P_{\mathscr{C}}(x_c^{(\ell)}) A_c^2 + \\ k_2 \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} n P_{\mathscr{C}}(x_c^{(\ell)}) A_c^4) < B \end{array} \right\}} \tag{143}$$

$$= \mathbb{1}_{\left\{ (k_1 \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} n \frac{p_c}{L_c} A_c^2 + k_2 \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} n \frac{p_c}{L_c} A_c^4) < B \right\}} \tag{144}$$

$$= \mathbb{1}_{\left\{ (k_1 \sum_{c=1}^{C} n p_c A_c^2 + k_2 \sum_{c=1}^{C} n p_c A_c^4) < B \right\}}. \tag{145}$$

The equality in (143) follows from (137) and (144) follows from (56). From (30) and (145), it follows that the code $\mathscr{C}$ is an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code if the following holds:

$$\delta = \mathbb{1}_{\left\{ (k_1 \sum_{c=1}^{C} n p_c A_c^2 + k_2 \sum_{c=1}^{C} n p_c A_c^4) < B \right\}}. \tag{146}$$

This completes the proof. $\qquad\square$

The following lemma provides an upper bound on the energy transmission rate B for $(n, M, \mathcal{X}, \epsilon, B, \delta)$-codes from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$.

Lemma A.10. *For an $(n, M, \mathcal{X}, \epsilon, B, \delta)$-code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62), let $j^+ \in \mathbb{N}$ be*

$$j^+ \triangleq \min \left\{ j \in \{1, \ldots M'\} : \delta \leqslant \frac{\sum_{k=1}^{j} \sum_{i=1}^{M} \mathbb{1}_{\{e_i = \bar{e}_k\}}}{M} \right\}, \tag{147}$$

where, the positive integer M' and the reals $\bar{e}_1$, $\bar{e}_2$, ..., $\bar{e}_{M'}$ are in (40). Then, the following holds for the energy transmission rate B:

$$B \leq \bar{e}_{j^+} \tag{148}$$

Proof. The proof follows on the same lines as that for Lemma 3.3 where, for all $i \in \{1, 2, \ldots, M\}$, the energy e_i in (21) is given by

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

$\qquad\square$

E. Miscellaneous Results

For the special case where for all $c \in \{1, 2, \ldots, C\}$, the radius of the decoding regions $r_c = r$ in (57), the following lemma provides a lower bound on the value of r.

Lemma A.11. *Consider an $(n, M, \mathcal{X}, P, \epsilon, B, \delta)$-code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62) such that, for all $c \in \{1, 2, \ldots, C\}$, $r_c = r$ in (61). Then, the parameter r satisfies:*

$$r \geq \sqrt{\sigma^2 \log\left(\frac{1}{1 - (1 - \epsilon)^{\frac{1}{n}}}\right)}, \tag{149}$$

where, the real σ^2 is defined in (9b).

The lower bound on r in Lemma A.11 characterizes the relationship between the DEP requirement ϵ and the spacing between channel input symbols r. This is a key element in the code construction since it helps determine the precise structure of the constellation that can guarantee a given DEP. Another crucial insight from Lemma A.11 is the relationship between the noise variance σ^2 and r. It is known from classical information theoretic literature that the channel noise variance dictates how clearly two neighboring symbols can be distinguished at the decoder. While this is a well known result, the precise relationship between these quantities cannot be determined due to the complexity of calculating the DEP. The unique choice of decoding regions made in (57) allows us to arrive at (149), which clearly shows the trade-off between these two quantities and provides the required minimum distance between channel input symbols as a function of the noise variance σ^2.
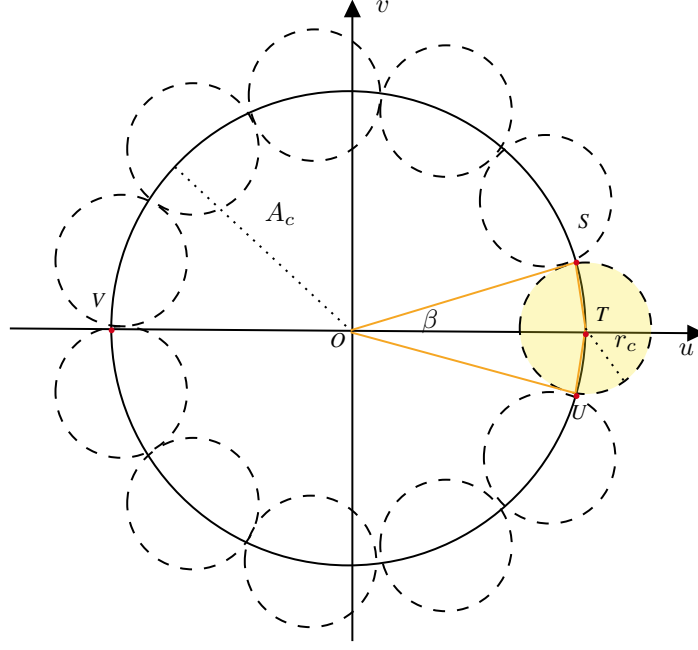
Fig. 4: Graphical representation of the symbols in layer c defined in (44)

Proof. If the parameters r_c in (61) are such that, for all $c \in \{1, 2, \ldots, C\}$, $r_c = r$, the average DEP in (120) is given by:

$$\gamma(\mathscr{C}) = 1 - \frac{1}{M} \sum_{i=1}^{M} \prod_{c=1}^{C} \left(1 - e^{-\frac{r^2}{\sigma^2}}\right)^{n \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{150}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \left(1 - e^{-\frac{r^2}{\sigma^2}}\right)^{n \sum_{c=1}^{C} \sum_{\ell=1}^{L_c} P_{\boldsymbol{u}(i)}(x_c^{(\ell)})}, \tag{151}$$

$$= 1 - \frac{1}{M} \sum_{i=1}^{M} \left(1 - e^{-\frac{r^2}{\sigma^2}}\right)^{n}, \tag{152}$$

$$= 1 - \left(1 - e^{-\frac{r^2}{\sigma^2}}\right)^{n}. \tag{153}$$

From (29), for $\mathscr{C}$ to be an $(n, M, \mathcal{X}, P, \epsilon)$-code, the following must hold:

$$\epsilon \geq 1 - \left(1 - e^{-\frac{r^2}{\sigma^2}}\right)^{n}. \tag{154}$$

This implies that

$$r \geq \sqrt{\sigma^2 \log\left(\frac{1}{1 - (1-\epsilon)^{\frac{1}{n}}}\right)}. \tag{155}$$

This completes the proof. $\qquad\square$

The information rate achievable by a code is a function of the number of channel input symbols L in (51) which in turn is a function of the number of symbols in each layer of the set of channel inputs $\mathcal{X}$ in (46). 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.

Lemma A.12. *Consider an $(n, M, \mathcal{X}, P, \epsilon, B, \delta)$-code $\mathscr{C}$ from the family $\mathsf{C}(C, \boldsymbol{A}, \boldsymbol{L}, \boldsymbol{\alpha}, \boldsymbol{p}, \boldsymbol{r})$ in (62). Then, for all $c \in \{1, 2, \ldots, C\}$, the number of symbols in layer c of $\mathcal{X}$ must satisfy the following:*

$$L_c \leq \left\lfloor \frac{\pi}{2 \arcsin \frac{r_c}{2 A_c}} \right\rfloor, \tag{156}$$

where, r_c is the radius of the decoding regions in (57) and A_c is the amplitude in (44).

Lemma A.12 provides an upper bound on the number of symbols that can be packed into a given layer of the proposed code construction while respecting the DEP requirement ϵ in (29). Using (156), the maximum number of symbols L_c that should

be placed in layer c of the code can be precisely determined. For all $c \in \{1, 2, \ldots, C\}$, the number of symbols in a layer L_c is a function of the radius of decoding sets r_c and the amplitude of the layer A_c. Note that the radii r_c are in turn functions of the noise variance σ^2 and the DEP requirement ϵ as shown in Lemma A.11.

Proof. For an $(n, M, \mathcal{X}, P, \epsilon, B, \delta)$-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 (61) and the amplitude A_c in (47) determine the number of symbols L_c that can be accommodated in the layer c.

Layer c of the form in (44) is illustrated in Fig. 4. The symbols in layer c are distributed uniformly along the circle of radius A_c centered at the origin O. The maximum number of symbols in layer c is equal to the number of non-overlapping circles of radius r_c corresponding to the decoding regions defined in (57) that can be placed along the circumference of the circle of radius A_c. From Fig. 4, 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

$$L_c \leq \left\lfloor \frac{2\pi}{\beta} \right\rfloor. \tag{157}$$

To determine the value of the angle β in Fig. 4, 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

$$2\pi - \beta = 2\angle STU, \tag{158}$$

which implies that

$$\angle STU = \frac{2\pi - \beta}{2}. \tag{159}$$

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

$$\angle STO = \frac{2\pi - \beta}{4}, \tag{160}$$

$$\angle SOT = \frac{\beta}{2}. \tag{161}$$

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

$$\frac{\sin(\angle SOT)}{ST} = \frac{\sin(\angle STO)}{SO}. \tag{162}$$

This implies that,

$$\frac{\sin\left(\frac{\beta}{2}\right)}{r_c} = \frac{\sin\left(\frac{2\pi - \beta}{4}\right)}{A_c}, \tag{163}$$

$$= \frac{1}{A_c} \sin\left(\frac{\pi}{2} - \frac{\beta}{4}\right), \tag{164}$$

$$= \frac{1}{A_c} \cos\left(\frac{\beta}{4}\right). \tag{165}$$

From (165), it follows that,

$$\frac{2}{r_c} \sin\left(\frac{\beta}{4}\right) \cos\left(\frac{\beta}{4}\right) = \frac{1}{A_c} \cos\left(\frac{\beta}{4}\right), \tag{166}$$

which implies that,

$$\sin\left(\frac{\beta}{4}\right) = \frac{r_c}{2A_c}, \quad \text{and} \tag{167}$$

$$\beta = 4 \arcsin \frac{r_c}{2A_c}. \tag{168}$$

Substituting the value of β from (168) in (157), it follows that the number of symbols in layer c of $\mathcal{X}$ is at most

$$L_c \leq \left\lfloor \frac{\pi}{2 \arcsin \frac{r_c}{2A_c}} \right\rfloor. \tag{169}$$

This completes the proof. □