

Unary Coding Design for Simultaneous Wireless Information and Power Transfer with Practical M-QAM

Yizhe Zhao, *Student Member, IEEE*, Jie Hu, *Member, IEEE*, Kun Yang, *Senior Member, IEEE*, and Kai-Kit Wong, *Fellow, IEEE*

Abstract

Relying on the propagation of modulated radio-frequency (RF) signals, we can achieve simultaneous wireless information and power transfer (SWIPT) to support low-power communication devices. In this paper, we studied a unary coding controlled SWIPT system by considering a practical M-QAM. Markov chains are exploited for characterising coherent binary information source and for modelling the generation process of modulated symbols. Therefore, both mutual information between the unary coded information source and destination and average energy harvesting performance of the SWIPT receiver are analysed in semi-closed-form. With the aid of the genetic algorithm, the sub-optimal codeword distribution of coded information source is obtained by maximising the average energy harvesting performance, while satisfying the requirement of the mutual information. Simulation results show that a higher-level unary code and a lower-order M-QAM results in higher WPT performance, when the maximum transmit power of the modulated symbol is fixed.

Index Terms

Simultaneous wireless information and power transfer (SWIPT), unary coding, modulation, Markov chain, genetic algorithm

I. INTRODUCTION

A. Background

In 5G and B5G, a huge number of communication devices result in a more crowded and more energy-sensitive network [1]. Moreover, many miniature-sized, battery-powered and batteryless devices play important roles for environmental monitoring [2]. In order to prolong the lifetime of these devices and reduce their maintenance cost, they are enabled to harvest energy from ambient environment [3]. However, uncontrollable and unpredictable energy harvesting processes limit communication performance of battery-powered and batteryless devices.

By contrast, radio frequency (RF) signal based wireless power transfer (WPT) [4] is capable of providing controllable and on-demand energy supply services to battery-powered and batteryless devices without extra hardware implementation at RF signal emitters, such as base stations (BSs). However, coordinating both wireless information transfer (WIT) and WPT in the same spectral bands imposes a lot of challenges on transceiver design, interference management and resource allocation, which yields researches on simultaneous wireless information and power transfer (SWIPT) [5]. Distinctive WIT and WPT services compete for various resources in time-domain [6], power-domain [7], frequency domain [8] and spatial domain [9] to deliver their quality of services (QoS).

However, current research [6]–[11] on SWIPT mainly focus on design of front-ends, while largely ignoring the impact of coding and modulation on SWIPT performance. The tradeoff between WIT and WPT also resides in modulation and coding design of a SWIPT system [12]. For instance, the WPT performance can be improved by designing appropriate codebooks and by increasing the transmit probabilities of the modulated symbols carrying higher transmit power. However, the WIT performance would be degraded.

B. Related Works

Plenty of existing works in SWIPT investigated transceiver design, resource allocation and networking. For an instance, Qi *et. al* [13] studied non-orthogonal-multi-access (NOMA) based SWIPT in IoT, where a robust beamformer was designed in order to maximize the weighted

sum rate and to minimize the total power consumption. Yu *et. al* [14] considered a distributed antenna aided SWIPT system by considering composite fading channels, where the energy efficiency performance is maximised. In [15], Chen *et. al* investigated a SWIPT aided mobile edge computing (MEC) system in full-duplex cellular networks, where users offload their partial computation tasks to a base station (BS). By exploiting SWIPT, these users may also be charged when they download computation results from the BS. When we lack terrestrial infrastructure, unmanned aerial vehicles (UAVs) are capable of transmitting wireless power to batteryless devices, while user associations [16] and workload scheduling [4] are optimised for improving both WPT and WIT performance. Moreover, in the network layer, Zhao *et. al* studied optimal hybrid access point (H-AP) deployment schemes in a smart city to enable SWIPT for mobile users. Shiming *et. al* [17] considered a SWIPT aided routing problem in a multi-hop energy-constrained wireless network, where energy harvested by a node can be an energy compensation for data forwarding.

However, analyses of WIT performance are based on the classic Shannon-Hartley channel capacity [14]–[16], which assumes Gaussian distributed transmit signals. This is far from practice, since modulated signals are practically transmitted with finite alphabet. The overestimated WIT performance may occupy more resources, which hence reduces attainable WPT performance in practice. Recently, very few works discussed the modulation and coding design for SWIPT. In [18], Zhao *et. al* studied a joint interleaving and modulation design in a NOMA-SWIPT system, where the energy interleaver and the constellation rotator were optimally designed for ensuring constructive combination of signals targeting different users. Chaehun *et. al* [19] proposed a multi-tone amplitude modulation scheme and the receiver architecture for SWIPT, where the energy harvested amount and the symbol error rate (SER) at the receiver were both improved. Other existing works about modulation design were studied in amplitude [20], frequency [21] or the tone index of multisine waveforms [22].

Coding design for SWIPT has also been studied in recent years. For instance, the SWIPT performance of the nonlinearity low density parity check (LDPC) code [23], run-length-limited (RLL) code [24] and subblock-constrained code [25] were analysed and optimised, respectively.

Kim *et. al* [26] analysed the rate-energy tradeoff by considering the case of the finite codelength. Hu *et. al* [27] adopted finite blocklength codes for achieving ultra-reliable low latency in a SWIPT aided relaying system. The reliability is maximized by considering both PS and TS protocols at a receiver. Moreover, the end-to-end WIT rate is maximised in a SWIPT aided cooperative communication system by adopting a rateless code [28]. In [29], Hu *et. al* studied an unary coding controlled SWIPT system, where the input distribution of unary codewords was optimised, in order to maximize WPT performance by considering infinite and finite battery capacity, while satisfying WIT requirement.

However, the impact of modulation [18]–[22] and that of coding on SWIPT performance [23]–[29] were separately studied. In fact, binary codewords generated by a SWIPT encoder cannot solely determine the SWIPT performance. The binary signs have to be mapped onto modulated symbols carrying different amount of energy, which results in various WPT performance. Different modulators also have various bit-error-rate (BER) performance, which results in distinctive WIT performance. For instance, in [29], the unary code controlled SWIPT is only feasible for on-off-keying (OOK) based modulator, where a binary sign ‘1’ is represented by the presence of a RF signal, but ‘0’ is represented by the absence of RF signals.

C. Contributions

Unary coding [29]–[31] has been widely adopted in modern wireless communication systems due to its low encoding and decoding complexity. Its flexible codeword structure is suitable to achieve SWIPT to low-power communication devices. Therefore, we aim for optimising the codeword distribution of an encoded binary information source for maximising the SWIPT performance, when the binary codewords are modulated by a practical M-QAM based modulator. Our novel contributions are then summarised as follows:

- A complete transceiver is studied, which includes a unary coding based SWIPT encoder, a M-QAM based modulator, a multiple-antenna aided transmitter and a power-splitting based SWIPT receiver.

- Discrete Markov chains are exploited for characterising the correlation among the binary signs generated by the SWIPT encoder and for modelling the transmit pattern of the modulated symbols.
- Both mutual information between the encoded information source and destination as well as energy harvested by the SWIPT receiver are analysed in semi-closed-form.
- The codeword distribution of the SWIPT encoder is obtained by exploiting the genetic algorithm, which maximises the average energy harvested by the SWIPT receiver, while guaranteeing a specific mutual information requirement.
- Our theoretical analysis and sub-optimal design are validated by numerical results by considering various unary coding and modulation schemes. The tradeoff between WIT and WPT performance is also demonstrated.

The rest of the paper is organised as follows: The SWIPT transceiver architecture is introduced in Section II, which is followed by our Markov chain based modelling in Section III. Both the WIT and WPT performance is analysed in Section IV, while the unary coding design is obtained in Section V. After presenting the numerical results in Section VI, we finally conclude our paper in Section VII. Some important mathematical notations are summarised in TABLE. I.

Notations: \mathbb{C} denotes the set of complex numbers. The scalar variable is denoted in the form of x , while the vector and the matrix are denoted in the form of \mathbf{x} and \mathbf{X} , respectively. \mathbf{X}^* (or \mathbf{x}^*) denotes the conjugate transpose of the matrix \mathbf{X} (or the vector \mathbf{x}). Moreover, $|x|$ represents the absolute value of a complex variable x , while $\|\mathbf{x}\|$ denotes the 2-norm of the vector \mathbf{x} .

II. SYSTEM MODEL AND TRANSCIEVER ARCHITECTURE

A SWIPT transceiver is illustrated in Fig. 1, which includes a SWIPT transmitter, a wireless channel and a SWIPT receiver. The SWIPT transmitter should be equipped with these functional modules, which are described as follows:

- *SWIPT encoder.* The original messages generated by an information source \mathbb{X} is encoded into binary codewords by a SWIPT encoder, which constitutes a binary coded information

TABLE I
THE DESCRIPTION OF THE NOTATIONS

Notations	The meaning of notations
\mathbb{X}	information source
$\overline{\mathbb{X}}$	coded information source
\mathbb{Y}	information destination
$\overline{\mathbb{Y}}$	coded information destination
\mathbf{X}_L	binary sequence generated by $\overline{\mathbb{X}}$
χ_j	j -th message of $\overline{\mathbb{X}}$
$\overline{\chi}_k$	k -th codeword of $\overline{\mathbb{X}}$
\mathbf{Pr}_x	codeword distribution
\mathbf{Pr}_c	state transition probabilities of coded information source
\mathbf{Pr}_s	state transition probabilities of modulators
s_m	m -th modulated symbol in M-QAM
ζ_m	binary sequence corresponding to s_m
$\zeta_{m,n}$	n -th binary sign in ζ_m
$\lambda_m^{(b)}$	backward run-length of '1' of s_m
$\lambda_m^{(f)}$	forward run-length of '1' of s_m
ξ	minimum Euclidean distance in a modulation constellation
$\kappa_{t,m}$	transmit power of s_m
$\kappa_{r,m}$	receive power of s_m
P_t	average transmit power
P_r	average receive power
ϵ_s	symbol error rate
ϵ_b	bit error rate

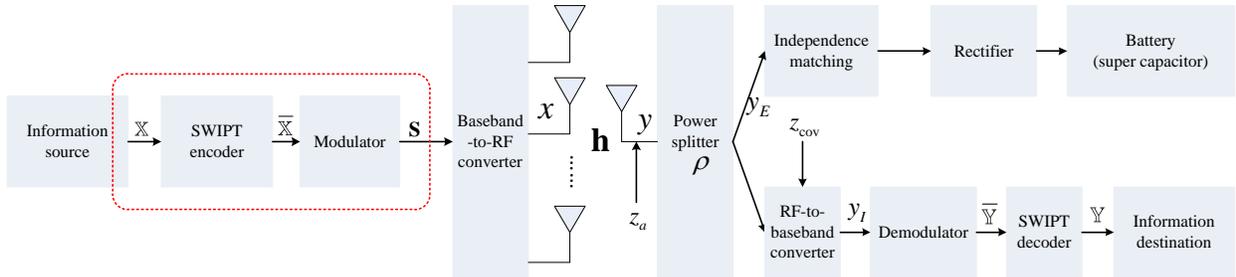


Fig. 1. Transceiver architecture

source $\overline{\mathbb{X}}$. The unary coding technique is used for the SWIPT encoder, since it's more friendly for low-power devices.

- *Modulator*. A digital modulator maps the binary sequences onto a range of modulated

symbols \mathbf{S} . Traditional M-QAM modulator is conceived at the SWIPT encoder.¹

- *Baseband-to-RF converter.* The baseband-to-RF converter is responsible for converting the baseband signals to the RF band. An appropriate beamformer is adopted by the multi-antennas transmitter, in order to maximize the received power at the receiver.

After passing the baseband-to-RF converter, the wireless signals are transmitted by N_t antennas. A Rayleigh block fading based multiple-input-single-output (MISO) channel is assumed between the SWIPT transmitter and receiver. The channel coefficient is expressed as

$$\mathbf{h} = \Omega \mathbf{g}, \quad (1)$$

where Ω is the large scale fading caused by path loss, $\mathbf{g} = [g_1, \dots, g_{N_t}] \in \mathbb{C}^{1 \times N_t}$ represents the multipath fading of the MISO channel.

The SWIPT receiver has the following main functional modules:

- *Power splitter.* The power splitter enables the SWIPT receiver to operate energy harvesting and information decoding simultaneously by splitting the received RF signals into two portions according to a power splitting ratio ρ .
- *WIT modules.* One portion of the received signal is converted to the baseband signal y_I by an RF-to-baseband converter. The signal y_I is then demodulated by a digital demodulator into binary sequences, which constitutes a coded information destination $\overline{\mathbb{Y}}$. The resultant binary sequences are then decoded into messages by an information decoder, which constitutes the final information destination \mathbb{Y} .
- *WPT modules.* The other portion of the received signals flow into an energy harvesting circuit, including an independence matching and a rectifier. The RF signal is then transformed to direct current (DC). The DC energy is finally stored in a battery or a super-capacitor.

The signal received by the single antenna of the SWIPT receiver is expressed as

$$y = \mathbf{h}\mathbf{w}x + z_a, \quad (2)$$

¹Note that any modulator is applicable in the following analysis. Without loss of generality, we select M-QAM, since it is widely adopted in modern digital communication systems.

where x is the transmit RF signal corresponds to the baseband modulated symbol s having a maximum transmit power P_{\max} , \mathbf{w} is the transmitter beamforming vector, and z_a represents the complex additive-white-Gaussian-noise (AWGN) imposed on the receive antenna, which has a zero mean and a variance of $\sigma_a/2$ per dimension. The maximum ratio based beamformer [32] has been proved to be optimal in a $N_t \times 1$ MISO system, where we have $\mathbf{w} = \mathbf{h}^*/\|\mathbf{h}\|$. A portion of $\sqrt{\rho}$ of the received RF signal is exploited for energy harvesting, which is expressed as

$$y_E = \sqrt{\rho}(\Omega\|\mathbf{g}\|x + z_a), \quad (3)$$

Moreover, the other portion $\sqrt{1-\rho}$ is relied upon for information demodulation and decoding, which is expressed as

$$y_I = \sqrt{1-\rho}(\Omega\|\mathbf{g}\|x + z_a) + z_{\text{cov}}, \quad (4)$$

where z_{cov} represents the AWGN imposed on the RF-to-baseband module, which has a zero mean and a variance of $\sigma_{\text{cov}}/2$ per dimension.

III. MARKOV CHAINS FOR CODING AND MODULATION

A. Markov Modelling for Unary Coded Information Source

Assuming that the original information source \mathbb{X} randomly generates a message chosen from a message book $\{\chi_1, \dots, \chi_J\}$, a message is then encoded by one of the unary codewords chosen from a codebook $\{\bar{\chi}_1, \dots, \bar{\chi}_K\}$. The k -th unary codeword $\bar{\chi}_k$ has $k-1$ binary signs ‘1’ followed by a binary sign ‘0’. If we denote the probability distribution of the original messages as $\{p(\chi_1), \dots, p(\chi_J)\}$, the codeword distribution is $\mathbf{Pr}_{\mathbf{x}} = \{p(\bar{\chi}_1), \dots, p(\bar{\chi}_K)\}$, where $p(\bar{\chi}_k) = \sum_{j=1}^J \Pr(\bar{\chi}_k|\chi_j)p(\chi_j)$ and $\Pr(\bar{\chi}_k|\chi_j)$ represents the probability that the original message χ_j is mapped onto the codeword $\bar{\chi}_k$.

The resultant coded information source $\bar{\mathbb{X}}$ may generate a binary sequence $\mathbf{X}_L = \{X_1, \dots, X_L\}$ having a length of L . We define a range of states $\mathbf{C}_L = \{C_1, \dots, C_L\}$ for the coded information source $\bar{\mathbb{X}}$, where C_l represents the run-length of ‘1’ after the transmission of the l -th binary sign in \mathbf{X}_L . Since the run-length of ‘1’ has a maximum of $(K-1)$ when the K -level unary

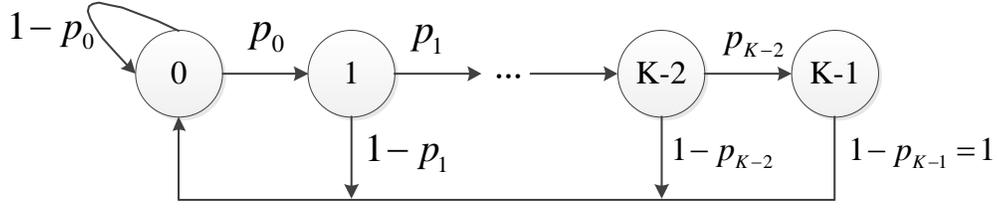


Fig. 2. Markov chain for a K -level unary coded information source $\bar{\mathbf{X}}$

code is conceived, we have $C_l \in \{0, \dots, k, \dots, K-1\}$ for $\forall l = 1, \dots, L$. Generally, there is always a definite mapping relationship between \mathbf{X}_L and the corresponding state sequence \mathbf{C}_L . For example, if we have $L = 8$ and $\mathbf{X}_8 = \{0, 1, 1, 0, 1, 1, 1, 1\}$, the corresponding state sequence should be $\mathbf{C}_8 = \{0, 1, 2, 0, 1, 2, 3, 4\}$. The state transitions follow the Markov property, as portrayed in Fig. 2. We define $p_k = \Pr\{C_{l+1} = k+1 | C_l = k\}$ ($k = 0, \dots, K-1$) as the transition probability from the state $C_l = k$ to the state $C_{l+1} = k+1$, which is expressed as

$$p_k = \frac{p(\bar{X}_{k+2}) + \dots + p(\bar{X}_K)}{p(\bar{X}_{k+1}) + p(\bar{X}_{k+2}) + \dots + p(\bar{X}_K)}, k = 0, \dots, K-2. \quad (5)$$

If the current state is $C_l = k$ ($k = 0, \dots, K-2$), the coded information source $\bar{\mathbf{X}}$ outputs a binary sign '1' with a probability of p_k , and it then transits to a new state of $C_{l+1} = k+1$. $\bar{\mathbf{X}}$ outputs a binary sign '0' with a probability of $1-p_k$, and it then transits to the state $C_{l+1} = 0$. Specifically, when the current state is $C_l = K-1$, $\bar{\mathbf{X}}$ must output a binary sign '0', since the run-length of '1' reaches its maximum. The state transition probabilities $\Pr_{\text{c}}(k'|k) = \Pr\{C_{l+1} = k' | C_l = k\}$ of the coded information source $\bar{\mathbf{X}}$ are then formulated as:

$$\Pr_{\text{c}}(k'|k) = \begin{cases} p_k, & k \leq K-2, k' = k+1, \\ 1-p_k, & k \leq K-2, k' = 0, \\ 1, & k = K-1, k' = 0, \\ 0, & \text{otherwise,} \end{cases} \quad (6)$$

which is also illustrated in Fig. 2.

Let us denote the state transition probabilities as a $K \times K$ matrix \mathbf{Pr}_{c} , whose element in the k -th row and in the k' -th column represents the probability of $\Pr_{\text{c}}(k'-1|k-1)$. Therefore,

TABLE II
MAPPING RELATIONSHIP BETWEEN BINARY SEQUENCES AND SYMBOLS AND THEIR CHARACTERISTICS FOR 16-QAM

m	ζ_m	$\lambda_m^{(b)}(\leftarrow)$	$\lambda_m^{(f)}(\rightarrow)$	m	ζ_m	$\lambda_m^{(b)}(\leftarrow)$	$\lambda_m^{(f)}(\rightarrow)$
1	0000 _← →	0	0	9	1000 _← →	0	1
2	0001 _← →	1	0	10	1001 _← →	1	1
3	0010 _← →	0	0	11	1010 _← →	0	1
4	0011 _← →	2	0	12	1011 _← →	2	1
5	0100 _← →	0	0	13	1100 _← →	0	2
6	0101 _← →	1	0	14	1101 _← →	1	2
7	0110 _← →	0	0	15	1110 _← →	0	3
8	0111 _← →	3	0	16	1111 _← →	4,5,...	4

stationary distribution π_c is then derived by solving the following equations

$$\begin{cases} \pi_c = \pi_c \cdot \mathbf{Pr}_c, \\ \sum_{k=1}^K \pi_{c,k} = 1, \end{cases} \quad (7)$$

where $\pi_{c,k}$ is the k -th element in π_c representing the stationary probability of $C_l = k - 1$.

B. Markov Analysis of Transmitted Symbols

By conceiving M-QAM as the digital modulator at the transmitter, the binary sequences generated by the SWIPT encoder are modulated as different symbols by obeying the Gray mapping rule. Given the m -th symbol s_m ($m = 1, \dots, M$) in the M-QAM constellation, its corresponding binary sequence is denoted as $\zeta_m = \{\zeta_{m,1}, \dots, \zeta_{m,n}, \dots, \zeta_{m,N}\}$ ($n = 1, \dots, N$), where ζ_m is the binary formation of the decimal number $(m - 1)$ and $N = \log_2(M)$.

Given a specific symbol s_m and its corresponding binary sequence ζ_m , we define $\lambda_m^{(b)}$ as the number of successive binary sign '1' counting from $\zeta_{m,N}$ in a backward direction. For $\forall m = 1, \dots, M - 1$, we have $\lambda_m^{(b)} < N$, since there is at least a single sign '0' in the sequence ζ_m . For the symbol s_M , we have $\lambda_M^{(b)} \in \{N, N + 1, \dots, K - 1\}$ if $K - 1 \geq N$, which depends on the

binary sequences of previously transmitted symbols. Furthermore, we define $\lambda_m^{(f)}$ as the number of successive binary sign ‘1’ counting from $\zeta_{m,1}$ to $\zeta_{m,N}$ in a forward direction. The mapping relationship of the symbols in 16-QAM and their corresponding binary sequences ζ_m , $\lambda_m^{(b)}$, $\lambda_m^{(f)}$ are exemplified in TABLE. II.

Two cases of (a) $K = 4, N = 4$ and (b) $K = 6, N = 4$ are exemplified in Fig. 3 in order to help readers understand all these concepts. For instance, in Fig. 3(a), the binary sequence of the modulated symbol s_{12} is $\zeta_{12} = 1011$. Its backward run of ‘1’ of its sub-sequence $\{\zeta_{12,3}, \zeta_{12,4}\} = 11$ has a length of $\lambda_{12}^{(b)}$, while the forward run of ‘1’ of the sub-sequence $\{\zeta_{12,1}\} = 1$ has a length of $\lambda_{12}^{(f)}$. Observe from Fig. 3(a) that the maximum run-length of ‘1’ is $K - 1 = 3 < N = 4$, which indicates that the sequence ζ_{16} and the corresponding symbol s_{16} are never output. Transmissions of a pair of consecutive modulated symbols also follow the Markov property, due to the correlation among binary signs output by the coded information source. By contrast, in the case of Fig. 3(b), we have $K - 1 = 5 > N = 4$. Therefore, the binary sequence ζ_{16} can be output by the coded information source $\bar{\mathbb{X}}$. Observe from Fig. 3(b) that the backward run-length $\lambda_{16}^{(b)}$ of ‘1’ of the binary sign ζ_{16} also depends on the previous sequence ζ_{14} in this example and its corresponding value is $\lambda_{16}^{(b)} = 5$. In this case, the probability that ζ_4 is output after ζ_{16} also depends on ζ_{14} . As a result, when we have $K - 1 \geq N$, state transitions no longer obey the Markov property, since a state of the modulator transmitting a symbol does not only depend on its adjacent peer, but also depend on more previous ones. Therefore, in the following discussions on deriving the transmit probabilities of modulated symbols, we consider two cases of $K - 1 < N$ and $K - 1 \geq N$, respectively.

1) *Case 1* : $K - 1 < N$. We define $\delta_{m,n}$ as the backward run-length of ‘1’ starting from the binary sign $\zeta_{m,n}$, which can be expressed as

$$\delta_{m,n} = \begin{cases} \zeta_{m,n} * (\delta_{m',N} + 1) = \zeta_{m,n} * (\lambda_{m'}^{(b)} + 1), n = 1, \\ \zeta_{m,n} * (\delta_{m,n-1} + 1), 1 < n \leq N, \end{cases} \quad (8)$$

where $\zeta_{m'}$ is the previous transmitted binary sequence before ζ_m .

Further, a Markov chain having M states, each of which corresponds to a specific symbol

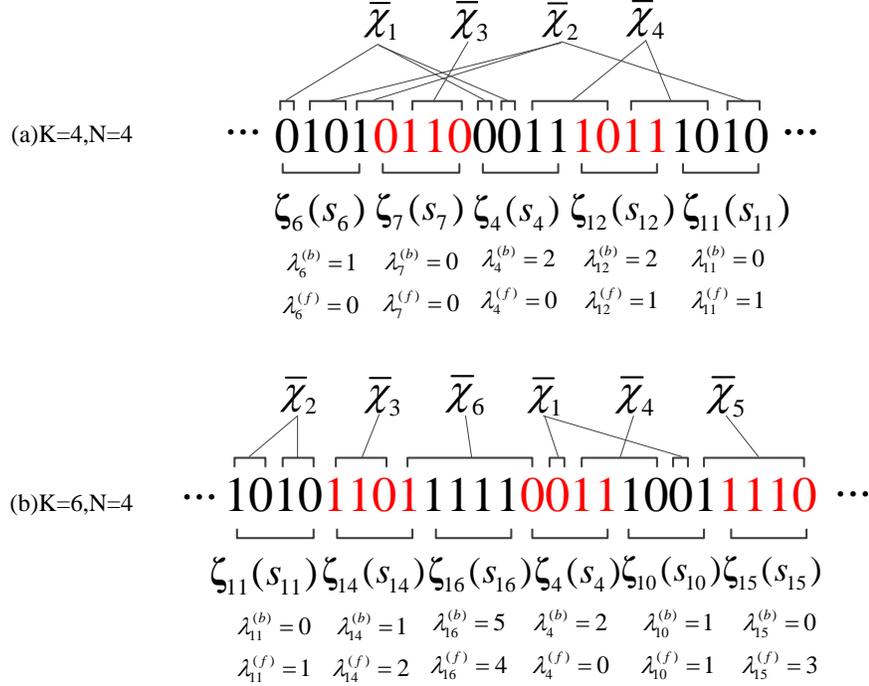


Fig. 3. Example of transmit binary sequences with two cases of $K = 4, N = 4$ and $K = 6, N = 4$.

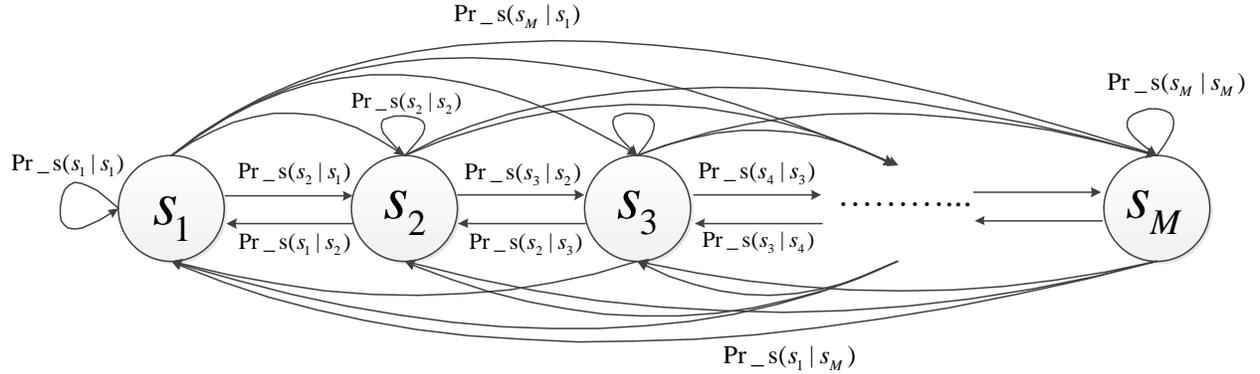


Fig. 4. Markov chain for the case of $K - 1 < N$.

s_m ($m = 1, \dots, M$), is output by the digital modulator, which is exemplified in Fig. 4. Given a specific time-slot, the Markov chain stays at the state $s_{m'}$. In the next time-slot, the probability of state $s_{m'}$ transiting to state s_m is expressed as

$$\text{Pr}_s(s_m | s_{m'}) = \begin{cases} 0, & \lambda_{m'}^{(b)} + \lambda_m^{(f)} > K - 1, \\ \text{Pr}_c(\delta_{m,1} | \lambda_{m'}^{(b)}) \prod_{n=1}^{N-1} \text{Pr}_c(\delta_{m,n+1} | \delta_{m,n}), & \lambda_{m'}^{(b)} + \lambda_m^{(f)} \leq K - 1. \end{cases} \quad (9)$$

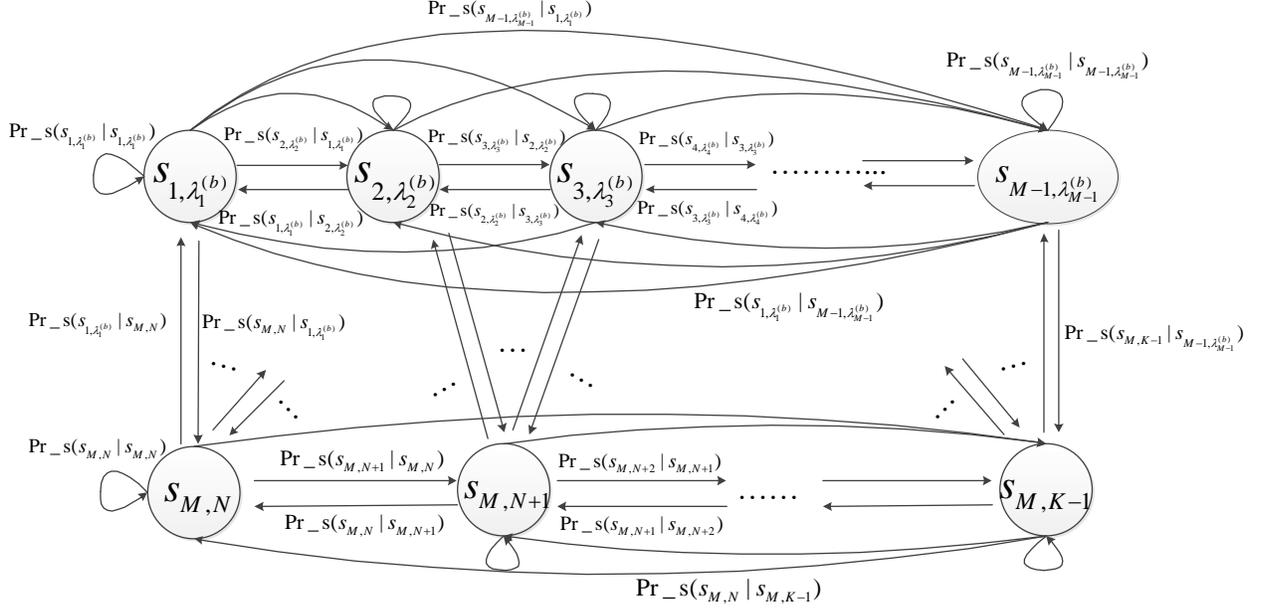


Fig. 5. Markov chain for the case of $K - 1 \geq N$.

The transition probability matrix of this Markov chain is denoted as \mathbf{Pr}_s having the size of $M \times M$, whose element on the m' -th row and the m -th column represents the state transition probability $\Pr_s(s_{m'} | s_m)$. Then, the stationary probability vector π_s of all possible states is obtained by solving the following equations

$$\begin{cases} \pi_s = \pi_s \cdot \mathbf{Pr}_s, \\ \sum_{m=1}^M \pi_{s,m} = 1. \end{cases} \quad (10)$$

The stationary probability τ_m of the symbol s_m being transmitted is expressed as $\tau_m = \pi_{s,m}$, where $\pi_{s,m}$ represents the m -th element in π_s .

2) *Case 2: $K - 1 \geq N$.* When we have $K - 1 \geq N$, we cannot use a similar Markov chain as we did in Case 1, where the Markov property cannot be satisfied, as exemplified in Fig. 3. We have to redefine the states for retrieving the Markov property.

Different from the Markov chain of the case $K - 1 < N$, we denote each state as $s_{m,q}$, where $q = \lambda_m^{(b)}$. As discussed before, when we have $m < M$, the binary sequence ζ_m has a definite value of $\lambda_m^{(b)}$. Therefore, a modulated symbol s_m may correspond to a specific state $s_{m,q}$. However, if we have $m = M$, the backward run-length $\lambda_M^{(b)}$ of '1' can be $\{N, N + 1, \dots, K - 1\}$, which

depends on the previous transmitted binary sequence. Therefore, the modulated symbol s_M may correspond to $K - N$ different states of $s_{M,q}$ ($q = N, N + 1, \dots, K - 1$). Given a state $s_{m',q'}$, the backward run-length of ‘1’ of the binary sequence $\zeta_{m'}$ is already known as q' . Then, the transition probability from the state $s_{m',q'}$ to $s_{m,q}$ is readily obtained according to the Markov analysis of the unary code. Therefore, the transition between two adjacent states $s_{m',q'}$ and $s_{m,q}$ obeys the Markov property. The Markov chain for the case of $K - 1 \geq N$ is illustrated in Fig. 5, while the state transition probability $\Pr_s(s_{m,q}|s_{m',q'})$ is derived as follows:

- Case 2-1: $m < M$. Since the backward run-length of sign ‘1’ is determined for every state, the state transition probability $\Pr_s(s_{m,q}|s_{m',q'})$ can be obtained similar to Case 1, which is formulated as

$$\Pr_s(s_{m,q}|s_{m',q'}) = \begin{cases} 0, & q' + \lambda_m^{(f)} > K - 1, \\ \Pr_c(\delta_{m,1}|q') \prod_{n=1}^{N-1} \Pr_c(\delta_{m,n+1}|\delta_{m,n}), & q' + \lambda_m^{(f)} \leq K - 1. \end{cases} \quad (11)$$

- Case 2-2: $m = M$. In this case, since there is no sign ‘0’ in the binary sequence ζ_M , the backward run-length $q = \lambda_M^{(b)}$ depends on the symbol $s_{m'}$ transmitted in the previous time-slot, which is expressed as $q = \lambda_M^{(b)} = \lambda_{m'}^{(b)} + N = q' + N$. Therefore, the probability of the state $s_{m',q'}$ transiting to $s_{M,q}$ is expressed as

$$\Pr_s(s_{M,q}|s_{m',q'}) = \begin{cases} 0, & q \neq q' + N, \\ \prod_{k=q'}^q \Pr_c(k+1|k), & \text{otherwise.} \end{cases} \quad (12)$$

The resultant state transition probability matrix of this Markov chain is denoted as $\mathbf{Pr_s}$ having the size of $(M + K - N - 1) \times (M + K - N - 1)$. Then, the stationary probabilities π_s of all $M + K - N - 1$ states can be obtained by solving similar equations as Eq. (10). For the cases of $m = 1 \dots, M - 1$, the probability that s_m being transmitted is just the stationary probability of $s_{m,q}$, where $m = 1 \dots, M - 1$ and $q = \lambda_m^{(b)}$. For the case of $m = M$, the probability that s_M being transmitted is the sum of the stationary probabilities $s_{M,q}$, where $q = N, N + 1, \dots, K - 1$.

Finally, the resultant probability τ_m of the symbol s_m being transmitted is then expressed as

$$\tau_m = \begin{cases} \pi_{s,m}, & m < M, \\ \sum_{i=M}^{M+K-N-1} \pi_{s,i}, & m = M. \end{cases} \quad (13)$$

IV. SWIPT PERFORMANCE ANALYSIS

A. WPT Performance Analysis

When M-QAM modulator is conceived, the minimum Euclidean distance in its constellation is expressed as $\xi = \sqrt{\frac{4P_{\max}}{(\sqrt{M}-1)^2}}$, where P_{\max} is the maximum transmit power of the modulated signal. For a complex modulated symbol $s_m = A_{m,I} + jA_{m,Q}$, we have $A_{m,I}, A_{m,Q} \in \{\frac{1-\sqrt{M}}{2}\xi, \frac{3-\sqrt{M}}{2}\xi, \dots, \frac{\sqrt{M}-3}{2}\xi, \frac{\sqrt{M}-1}{2}\xi\}$. After obtaining the stationary transmission probabilities $\{\tau_m | m = 1, \dots, M\}$ of all possible symbols, the average transmit power P_t of the SWIPT transmitter is then formulated as

$$P_t = \sum_{m=1}^M \tau_m \kappa_{t,m}, \quad (14)$$

where $\kappa_{t,m} = ((A_{m,I})^2 + (A_{m,Q})^2)/2$ represents the transmit power of the symbol s_m . When s_m is transmitted, the received power for WPT is obtained as

$$\kappa_{r,m} = \rho \left((\Omega \|\mathbf{g}\| A_{m,I} + z_{a,I})^2 + (\Omega \|\mathbf{g}\| A_{m,Q} + z_{a,Q})^2 \right) / 2, \quad (15)$$

by exploiting Eq. (4) and by reformulating the complex antenna noise as $z_a = z_{a,I} + jz_{a,Q}$.

The envelope $\{|g_i| | i = 1, \dots, N_t\}$ of every element in the multipath fading vector \mathbf{g} follows a Rayleigh distribution. Therefore, the corresponding power gains $\{|g_i|^2 | i = 1, \dots, N_t\}$ all obeys an exponential distribution. As a result, $\|\mathbf{g}\|^2 = \sum_{i=1}^{N_t} |g_i|^2$ follows a chi-square distribution having $2N_t$ degrees of freedom. The probability density function (PDF) of $\|\mathbf{g}\|$ is then formulated as

$$f(\|\mathbf{g}\|) = \frac{1}{2^{N_t-1} \Gamma(N_t)} \|\mathbf{g}\|^{2N_t-1} e^{-\frac{\|\mathbf{g}\|^2}{2}}, \quad (16)$$

where we have

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt. \quad (17)$$

When the symbol s_m is transmitted, the amount of energy harvested is expressed as

$$P_{r,m} = \begin{cases} k_{r,m}, & \text{if } k_{r,m} \geq P_{\text{th}}, \\ 0, & \text{otherwise,} \end{cases} \quad (18)$$

where P_{th} is the sensitivity of an energy harvester. Therefore, the average energy harvesting performance is formulated as

$$P_r = \sum_{m=1}^M \tau_m \int_0^{+\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\pi \sigma_a^2} e^{-\frac{z_{a,I}^2 + z_{a,Q}^2}{\sigma_a^2}} \frac{\|\mathbf{g}\|^{2N_r-1} e^{-\frac{\|\mathbf{g}\|^2}{2}}}{2^{N_r-1} \Gamma(N_r)} P_{r,m} dz_{a,I} dz_{a,Q} d\|\mathbf{g}\|. \quad (19)$$

Since we cannot obtain the exact expressions for τ_m , P_r of (19) is in a semi-closed form.

B. WIT Performance Analysis

1) *BER Analysis*: By substituting a specific modulated symbol s_m into Eq. (4), when the symbol s_m is transmitted, the received symbol \hat{s}_m for WIT is formulated as

$$\begin{aligned} \hat{s}_m &= \sqrt{(1-\rho)}(\Omega\|\mathbf{g}\|s_m + z_a) + z_{\text{cov}} \\ &= \sqrt{(1-\rho)}\Omega\|\mathbf{g}\|(A_{m,I} + jA_{m,Q}) + (z_{n,I} + jz_{n,Q}), \end{aligned} \quad (20)$$

where $z_n = \sqrt{(1-\rho)}z_a + z_{\text{cov}} \sim \mathbb{C}\mathcal{N}(0, (1-\rho)\sigma_a^2 + \sigma_{\text{cov}}^2)$ is total AWGN having two orthogonal dimensions $z_{n,I}$ and $z_{n,Q}$. Since the receiver has no knowledge of the input distribution, maximum likelihood (ML) detector is conceived for demodulating the received signals. Each modulated symbol has a successful demodulation region at the receiver, which is expressed as a square having the length of $\sqrt{(1-\rho)}\Omega\|\mathbf{g}\|\xi$. Therefore, in order to successfully demodulate the original symbol s_m , the offset $z_{n,I} + jz_{n,Q}$ should be constrained within the range of $[-\sqrt{(1-\rho)}\Omega\|\mathbf{g}\|\xi/2, \sqrt{(1-\rho)}\Omega\|\mathbf{g}\|\xi/2]$ for both the in-phase and the quadrature dimension.

Generally, the symbol demodulation error always occurs between two adjacent constellation points, since the probability that one symbol is mis-demodulated as any other non-adjacent

symbols can be neglected. According to the Gray mapping rule, there is only one binary difference between the sequences of two adjacent constellation points. Therefore, the BER at the SWIPT receiver could be approximated as $\epsilon_b \approx \frac{\epsilon_s}{N}$, where the SER ϵ_s is calculated as

$$\begin{aligned} \epsilon_s &= 1 - \int_0^{+\infty} \int_{-\frac{\sqrt{(1-\rho)\Omega}\|\mathbf{g}\|\xi}{2}}^{\frac{\sqrt{(1-\rho)\Omega}\|\mathbf{g}\|\xi}{2}} \int_{-\frac{\sqrt{(1-\rho)\Omega}\|\mathbf{g}\|\xi}{2}}^{\frac{\sqrt{(1-\rho)\Omega}\|\mathbf{g}\|\xi}{2}} \frac{e^{-\frac{z_{n,I}^2+z_{n,Q}^2}{(1-\rho)\sigma_a^2+\sigma_{\text{cov}}^2}}}{\pi((1-\rho)\sigma_a^2+\sigma_{\text{cov}}^2)} f(\|\mathbf{g}\|) dz_{n,I} dz_{n,Q} d\|\mathbf{g}\| \\ &= 1 - \int_0^{+\infty} \left[1 - 2Q\left(\frac{\sqrt{(1-\rho)\Omega}\|\mathbf{g}\|\xi/2}{\sqrt{(1-\rho)\sigma_a^2/2+\sigma_{\text{cov}}^2/2}}\right) \right]^2 f(\|\mathbf{g}\|) d\|\mathbf{g}\|. \end{aligned} \quad (21)$$

Since the successful demodulation regions of all the symbols have the same size, the probabilities that any received symbols \hat{s}_m and $\hat{s}_{m'}$ are mis-demodulated to each other are also the same. The BER from ‘0’ to ‘1’ is the same as that from ‘1’ to ‘0’.

2) *Mutual Information*: We consider the mutual information between the encoded information source $\overline{\mathbf{X}}$ and destination $\overline{\mathbf{Y}}$ as the WIT performance of the SWIPT system. The coded source $\overline{\mathbf{X}}$ outputs a binary sequence $\mathbf{X}_L = \{X_1, \dots, X_L\}$ having a length of L . Then, the encoded destination $\overline{\mathbf{Y}}$ receive a binary sequence $\mathbf{Y}_L = \{Y_1, \dots, Y_L\}$. Information theoretical channel between the encoded information source and destination can be modelled as a binary symmetric channel having the average BER ϵ_b as the crossover probabilities. The mutual information between $\overline{\mathbf{X}}$ and $\overline{\mathbf{Y}}$ can be formulated as

$$I(\overline{\mathbf{X}}; \overline{\mathbf{Y}}) = H(\overline{\mathbf{X}}) - H(\overline{\mathbf{X}}|\overline{\mathbf{Y}}). \quad (22)$$

According to Eq. (6) of [29], the entropy $H(\overline{\mathbf{X}})$ can be expressed as

$$\begin{aligned} H(\overline{\mathbf{X}}) &\triangleq \lim_{L \rightarrow \infty} \frac{H(\mathbf{X}_L)}{L} = \lim_{L \rightarrow \infty} \frac{H(\mathbf{C}_L)}{L} \\ &= \lim_{L \rightarrow \infty} H(C_L|C_1, \dots, C_{L-1}) \\ &= \lim_{L \rightarrow \infty} H(C_L|C_{L-1}) = \lim_{L \rightarrow \infty} H(C_2|C_1). \end{aligned} \quad (23)$$

The conditional entropy $H(\bar{\mathbf{X}}|\bar{\mathbf{Y}})$ is then formulated as

$$\begin{aligned} H(\bar{\mathbf{X}}|\bar{\mathbf{Y}}) &= \lim_{L \rightarrow \infty} \frac{H(\mathbf{X}_L|\mathbf{Y}_L)}{L} = \lim_{L \rightarrow \infty} \frac{H(\mathbf{C}_L|\mathbf{Y}_L)}{L} \\ &= \lim_{L \rightarrow \infty} \frac{1}{L} \left[H(C_1|Y_1, \dots, Y_L) + \sum_{l=2}^L H(C_l|C_{l-1}, \dots, C_1, Y_1, \dots, Y_L) \right]. \end{aligned} \quad (24)$$

Given the state C_l at the l -th instant, the state C_{l+1} is independent of $\{Y_1, \dots, Y_{l-1}, Y_l\}$ at the previous instants, while the state C_l is also correlated with $\{Y_{l+1}, Y_{l+2}, \dots\}$, according to Lemma 1 and Lemma 2 in [29]. Therefore, we have

$$\begin{aligned} H(C_l|C_{l-1}, \dots, C_1, Y_1, \dots, Y_L) &= H(C_l|C_{l-1}, Y_1, \dots, Y_l, \dots, Y_L) \\ &= H(C_l|C_{l-1}, Y_l, \dots, Y_L) \\ &\leq H(C_l|C_{l-1}, Y_l). \end{aligned} \quad (25)$$

As a result, the upper-bound of the conditional entropy $H(\bar{\mathbf{X}}|\bar{\mathbf{Y}})$ is obtained as

$$\begin{aligned} H(\bar{\mathbf{X}}|\bar{\mathbf{Y}}) &\leq \lim_{L \rightarrow \infty} \frac{1}{L} \left[H(C_1|Y_1, \dots, Y_L) + \sum_{l=2}^L H(C_l|C_{l-1}, Y_l) \right] \\ &\stackrel{(a)}{=} \lim_{L \rightarrow \infty} \frac{H(C_1|Y_1, \dots, Y_L) + (L-1)H(C_2|C_1, Y_2)}{L} \\ &= H(C_2|C_1, Y_2), \end{aligned} \quad (26)$$

where the equality (a) is derived by considering the stationary property of $\bar{\mathbf{X}}$. Then, the lower-bound of mutual information $I(\bar{\mathbf{X}}; \bar{\mathbf{Y}})$ is finally derived as

$$\begin{aligned} I(\bar{\mathbf{X}}; \bar{\mathbf{Y}}) &\geq H(C_2|C_1) - H(C_2|C_1, Y_2) = H(Y_2|C_1) - H(Y_2|C_2, C_1) \\ &= \sum_{k=0}^{K-1} \pi_{c,k} [H(p_k \epsilon_b + (1-p_k)(1-\epsilon_b)) - H(\epsilon_b)] \\ &\triangleq I_{\text{inf}}(\bar{\mathbf{X}}; \bar{\mathbf{Y}}), \end{aligned} \quad (27)$$

where $H(x) = -x \log_2(x) - (1-x) \log_2(1-x)$. Since we cannot obtain the exact expression for π_c , $I_{\text{inf}}(\bar{\mathbf{X}}; \bar{\mathbf{Y}})$ of Eq. (27) is in a semi-closed form.

V. JOINT CODING AND MODULATION DESIGN

The unary codeword distribution $\mathbf{Pr}_{\mathbf{x}}$ as well as the power splitting ratio ρ are jointly optimised, in order to maximize the average power P_r for energy harvesting of the receiver, while guaranteeing the lower-bound of mutual information $I_{\text{inf}}(\overline{\mathbf{X}}; \overline{\mathbf{Y}})$ is higher than a threshold. This optimization problem is formulated as

$$\begin{aligned} \text{(P1)} \quad & \max_{\mathbf{Pr}_{\mathbf{x}}, \rho} P_r \\ \text{s.t.} \quad & I_{\text{inf}}(\overline{\mathbf{X}}; \overline{\mathbf{Y}}) \geq I_{\text{th}}. \end{aligned} \quad (28)$$

Given a threshold I_{th} , it's obvious that the actual mutual information $I(\overline{\mathbf{X}}; \overline{\mathbf{Y}})$ also satisfies the constraint, since we have $I(\overline{\mathbf{X}}; \overline{\mathbf{Y}}) \geq I_{\text{inf}}(\overline{\mathbf{X}}; \overline{\mathbf{Y}}) \geq I_{\text{th}}$. Since both P_r and $I_{\text{inf}}(\overline{\mathbf{X}}; \overline{\mathbf{Y}})$ are expressed in semi-closed form, (P1) can only be solved by a heuristic algorithm.

A. Power Splitting Ratio

Given a specific codeword distribution $\mathbf{Pr}_{\mathbf{x}}$, we firstly optimise the power splitting ratio ρ .

Theorem 1: The average energy harvesting power P_r monotonously increases with respect to (w.r.t.) the power splitting ratio ρ , while the lower-bound mutual information $I_{\text{inf}}(\overline{\mathbf{X}}; \overline{\mathbf{Y}})$ monotonously decreases w.r.t. ρ , when we have $\epsilon_b < 0.5$.

Proof: Please refer to Appendix A for detailed proof. ■

Since the BER ϵ_b is normally lower than 0.5, the optimal power splitting ratio ρ^* could be obtained by solving $I_{\text{inf}}(\overline{\mathbf{X}}; \overline{\mathbf{Y}}) = I_{\text{th}}$ with the aid of a bisection method, which is summarized in Algorithm. 1.

B. GA Aided Coding Design

Given the optimal power splitting ratio ρ^* , a genetic algorithm is proposed for optimising the codeword distribution $\mathbf{Pr}_{\mathbf{x}}$, which is detailed in the pseudo code of Algorithm. 2. The main steps are summarised as follows:

Algorithm 1 Bisection method based algorithm for obtaining ρ^* .

Input: Codeword distribution $\mathbf{Pr}_{\mathbf{x}}$; Mutual information threshold I_{th} ; Mutual information threshold I_{th} ; Convergence threshold ϕ and all the other physical layer parameters.

Output: Optimal power splitting ratio ρ^* ;

```

1: Initialise  $\rho_{\text{upper}} \leftarrow 1, \rho_{\text{lower}} \leftarrow 0, I \leftarrow 0$ ;
2: while  $|I - I_{\text{th}}| > \phi$  do
3:   Update  $\rho = (\rho_{\text{upper}} + \rho_{\text{lower}})/2$ ;
4:   Obtain lower-bound mutual information  $I = I_{\text{inf}}(\bar{\mathbf{X}}; \bar{\mathbf{Y}})$  according to Eq. (27);
5:   if  $I > I_{\text{th}}$  then
6:      $\rho_{\text{lower}} = \rho$ ;
7:   else
8:      $\rho_{\text{upper}} = \rho$ ;
9:   end if
10: end while
11: Return  $\rho^* = (\rho_{\text{upper}} + \rho_{\text{lower}})/2$ .

```

- *Step 1: Initialization.* Initialize a generation \mathbb{G} having ind_num individuals. Each individual is a legal codeword distribution $\mathbf{Pr}_{\mathbf{x}_j} = \{p_j(\bar{\mathcal{X}}_1), \dots, p_j(\bar{\mathcal{X}}_K)\}$ satisfying $\sum_{k=1}^K p_j(\bar{\mathcal{X}}_k) = 1$ for $\forall j = 1, \dots, ind_num$. This step corresponds to line 1 in Algorithm. 2.
- *Step 2: Natural selection.* Obtain the fitness of each individual according to the mutual information constraint of (P1). The fitness of the j -th individual is expressed as

$$fitness_j = \begin{cases} P_r, & I_{\text{inf}}(\bar{\mathbf{X}}; \bar{\mathbf{Y}}) \geq I_{\text{th}}, \\ 0, & \text{otherwise.} \end{cases} \quad (29)$$

The selection probability of each individual is calculated as

$$prob_j = \frac{fitness_j}{\sum_{j'=1}^{ind_num} fitness_{j'}}. \quad (30)$$

Then, we update the generation \mathbb{G} by selecting new individuals according to the probabilities $\{prob_j | \forall j\}$. This step corresponds to line 6 and line 10 in Algorithm. 2.

- *Step 3: Crossover.* Crossover is operated after a natural selection. The generation \mathbb{G} is updated once more by generating offsprings of two randomly selected parental individuals one after another. Given a specific pair of parental individuals $\mathbf{Pr}_{\mathbf{x}_{(father)}}$ and $\mathbf{Pr}_{\mathbf{x}_{(mother)}}$,

their offsprings $\mathbf{Pr_x}_{(offspring)}$ is expressed as

$$p_{(offspring)}(\bar{\chi}_k) = a \cdot p_{(father)}(\bar{\chi}_k) + (1 - a) \cdot p_{(mother)}(\bar{\chi}_k), \forall j = 1, \dots, ind_num, k = 1, \dots, K. \quad (31)$$

Note that the offspring satisfies $\sum_{k=1}^K p_{(offspring)}(\bar{\chi}_k) = 1$, since we have $\sum_{k=1}^K p_{(father)}(\bar{\chi}_k) = 1$ and $\sum_{k=1}^K p_{(mother)}(\bar{\chi}_k) = 1$. The crossover factor a determines the fraction of characteristics that the offspring inherits from the parental individuals. In our algorithm, we choose $a = 0.5$. This step corresponds to line 11 in Algorithm. 2.

- *Step 4: Mutation.* The last step of genetic algorithm is mutation, in order to prevent the solution from local-optimum. For each individual $\mathbf{Pr_x}_j$, the mutation is operated by adding a random Δp following the normal distribution $\mathcal{N}(0, \nu)$, which is expressed as

$$p_j(\bar{\chi}_k) = p_j(\bar{\chi}_k) + \Delta p, k = 1, \dots, K. \quad (32)$$

Then, the individual is updated as

$$p_j(\bar{\chi}_k) = \frac{p_j(\bar{\chi}_k) - \min(0, \min(\mathbf{Pr_x}_j))}{\sum \mathbf{Pr_x}_j - K \min(0, \min(\mathbf{Pr_x}_j))}, \quad (33)$$

for $\forall k = 1, \dots, K$. A larger variance ν results in a faster convergence of the genetic algorithm, but a lower optimality. This step corresponds to line 12 in Algorithm. 2.

VI. SIMULATION RESULTS

Our optimal coding design and impact of M-QAM on SWIPT performance are validated by both Monte-Carlo simulation and theoretical analysis. We consider $N_t = 4$ antennas at the transmitter and one antenna at the receiver. The square root of path-loss is set to $\Omega = 10^{-1.5}$, while the noise power is $\sigma_a^2 = \sigma_{cov}^2 = -50\text{dBm}$. The transmitter has a maximum transmit power constraint of $P_{\max} = 0.1W$, while the energy harvesting threshold is set to $P_{\text{th}} = -13\text{dBm}$. The Gray mapping from binary sequence to modulated symbols is conceived, where the constellation of 16-QAM is exemplified in Fig. 6.

Algorithm 2 Genetic algorithm for obtaining ρ^* and $\mathbf{Pr}_{\mathbf{x}^*}$.

Input: Mutual information threshold I_{th} ; Mutation probability mutation_prob ; Maximum generation number max_gen and all the other physical layer parameters.

Output: Optimal power splitting ratio ρ^* ; Optimal input distribution $\mathbf{Pr}_{\mathbf{x}^*}$;

- 1: Initialise a generation \mathbb{G} according to Step 1; Generation label $i = 1$; Optimal fitness fitness^* ;
 - 2: **while** $i < \text{max_gen}$ **do**
 - 3: $i \leftarrow i + 1$;
 - 4: Obtain the optimal power splitting ratio ρ_j for each individual $\mathbf{Pr}_{\mathbf{x}_j}$ in \mathbb{G} according to Algorithm. 1;
 - 5: Obtain the average energy harvesting power P_r and lower-bound mutual information $I_{\text{inf}}(\bar{\mathbf{X}}; \bar{\mathbf{Y}})$ for each individual;
 - 6: Obtain the fitness of \mathbb{G} according to (29);
 - 7: **if** $\max\{\text{fitness}_j\} > \text{fitness}^*$ **then**
 - 8: Update $\text{fitness}^* \leftarrow \max\{\text{fitness}_j\}$. Update ρ^* and $\mathbf{Pr}_{\mathbf{x}^*}$ according to the individual having the maximum fitness;
 - 9: **end if**
 - 10: Update the generation \mathbb{G} by selecting individuals according to the probabilities (30);
 - 11: Obtain offsprings after crossover according to Step 3 and update the generation \mathbb{G} ;
 - 12: As for each individual, generate a value φ following uniformly distribution $\mathcal{U}(0, 1)$. If $\varphi > \text{mutation_prob}$, operate mutation according to Step 4. Then, update the generation G after traversing all the individuals;
 - 13: **end while**
-

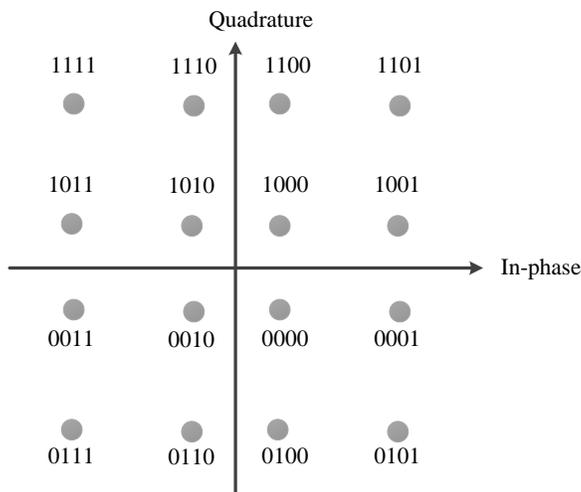


Fig. 6. Gray mapping for 16-QAM

A. Validation of Theoretical Analysis

Firstly, we validate our theoretical analysis, when 16-QAM is adopted. In Fig. 7, we compare the theoretical transmit probabilities of all the modulated symbols to the simulation counterparts, given an arbitrary codeword distribution $\mathbf{Pr}_{\mathbf{x}} = \{0.42, 0.25, 0.07, 0.26\}$ for 4-level unary code and $\mathbf{Pr}_{\mathbf{x}} = \{0.16, 0.25, 0.13, 0.04, 0.05, 0.05, 0.23, 0.09\}$ for 8-level unary code, respectively. The Monte Carlo simulation is adopted by generating 10^6 codewords for obtaining those probabilities.

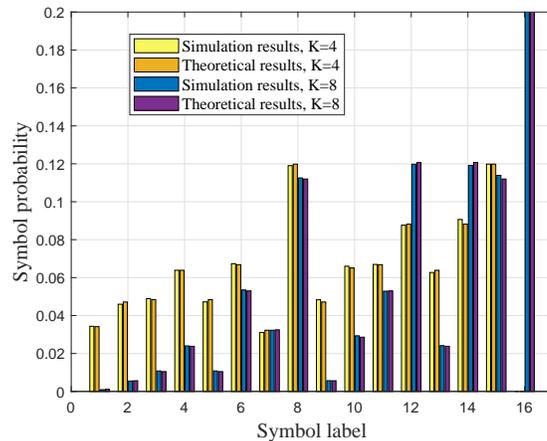


Fig. 7. Validation of symbol probability Markov analysis.

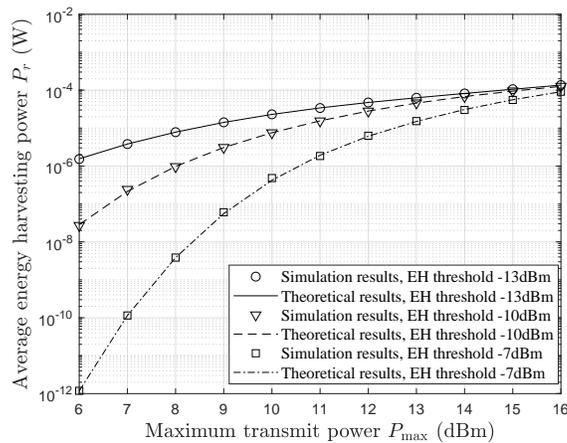


Fig. 8. Validation of WPT analysis.

Observe from Fig. 7 that the simulation and theoretical results perfectly match with each other, which confirms the accuracy of our Markov modelling. Note that when 4-level unary code is conceived, the probability of the 16-th symbol is zero, since the maximum run-length of binary sign ‘1’ is $K-1 = 3$. In Fig. 8 and 9, we depict the average energy harvesting power as well as the receive BER versus maximum transmit power P_{\max} . Observe from Fig. 8 that a lower threshold P_{th} results in a higher energy harvesting power. When we increase the maximum transmit power P_{\max} , P_{th} has less effect on energy harvesting, since the received power is sufficiently high to activate the energy harvester. Moreover, observe from Fig. 9 that a higher power splitting ratio ρ results in a higher BER, since the signal strength for the demodulation is weakened.

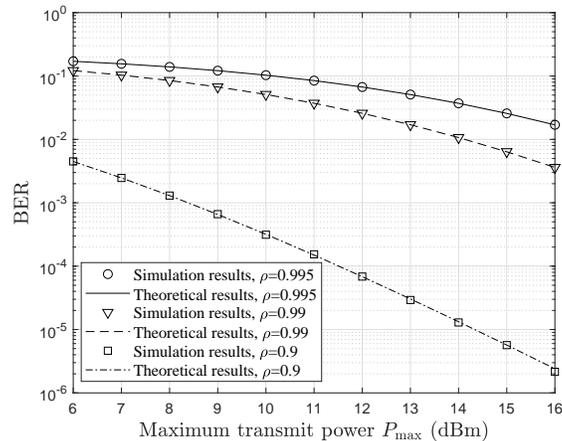


Fig. 9. Validation of WIT analysis.

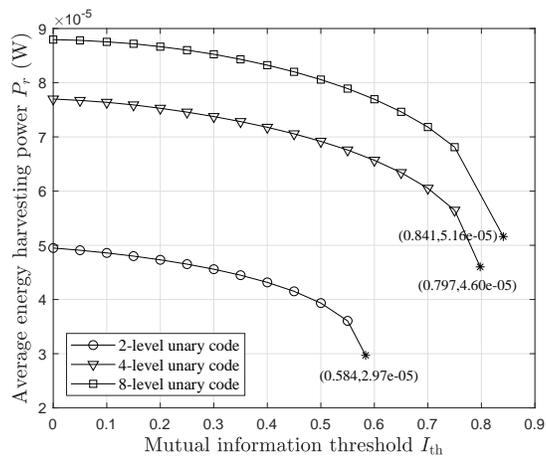


Fig. 10. Average energy harvesting power versus mutual information threshold, when 16-QAM modulator is conceived.

B. Genetic Algorithm Aided Coding Design

We then evaluate the SWIPT performance of our genetic algorithm aided coding design in Figs. 10-12. Observe from Fig. 10 that when we increase the minimum mutual information constraint I_{th} , the average energy harvesting power P_r is reduced. Given a K-level unary code and a fixed modulator, any coding design cannot achieve a mutual information higher than the corresponding channel capacity. Therefore, observe from Fig. 10 that, every coding and modulation scheme can only satisfy the mutual information constraint lower than a certain value. For example, when 16-QAM modulator is adopted, the upper-bound mutual information of 2-level unary code is 0.584, while that of 8-level unary code is 0.841. Observe from Fig. 10 that

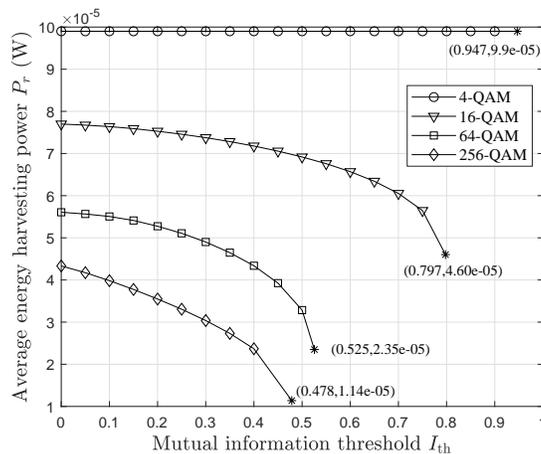


Fig. 11. Average energy harvesting power versus modulation order, when 4-level unary code is conceived.

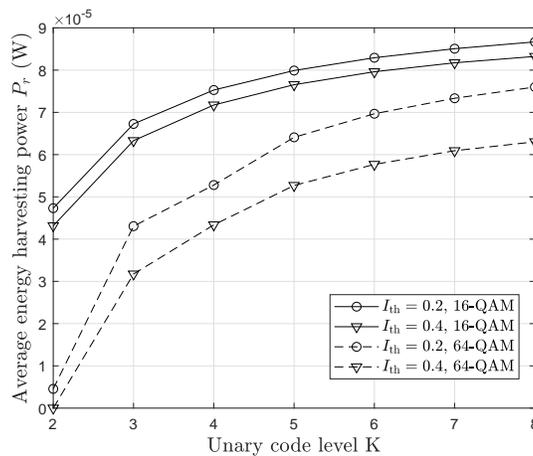


Fig. 12. Average energy harvesting power versus unary code level K .

8-level unary code has the best WPT performance. Moreover, we may observe from Fig. 11 that when 4-QAM modulator is adopted, the average energy harvesting power P_r does not change as we increase I_{th} . This is because that all the symbols in 4-QAM's constellation have the same transmit power. Therefore, the codeword distribution does not have any impact on the WPT performance. Moreover, a higher order modulator results in a lower WPT performance. This is because when the maximum transmit power P_{max} is fixed, a higher order modulator achieves a lower average transmit power. Furthermore, a higher order modulator also results in a higher BER, which indicates that WPT performance has to be sacrificed in order to compensate for the degradation of WIT performance.

Observe from Fig. 12 that the average energy harvesting power P_r increases as we increase the level of the unary code. This is because when K becomes higher, the SWIPT encoder becomes more flexible to control the attainable SWIPT performance. Furthermore, a higher order modulator results in a more rapidly increasing P_r . This is because a modulated symbol may represent more binary signs, which requires a higher K to achieve the optimal WPT performance.

VII. CONCLUSION

The unary coding design is proposed for a SWIPT system with practical digital modulators. By exploiting the Markov chain, the stationary probabilities of modulated symbols being transmitted are obtained, which results in semi-closed-form theoretical analysis of both WPT and WIT performance. By exploiting the genetic algorithm aided coding design, the amount of energy harvested by the SWIPT receiver is maximised, while satisfying the mutual information constraint. At last, Monte-Carlo simulation validates our theoretical analysis. The tradeoff between the WIT and WPT is also revealed in the coding and modulation level. Both the simulation and theoretical results demonstrate that a higher level of the unary code and a lower order of M-QAM results in a higher WPT performance, when the maximum transmit power of the modulated symbol is fixed.

APPENDIX A

PROOF OF THEOREM. 1

We firstly discuss the monotonicity of P_r w.r.t. ρ . If we increase ρ , $\kappa_{r,m}$ also increases. According to Eq. (19), P_r monotonously increases w.r.t. ρ .

Since function $Q\left(\frac{\sqrt{(1-\rho)\Omega}\|\mathbf{g}\|\xi/2}{\sqrt{(1-\rho)\sigma_a^2/2+\sigma_{\text{cov}}^2/2}}\right)$ monotonically increase w.r.t. ρ , observe from Eq. (21) that ϵ_s also increases when we increase ρ . So does ϵ_b . We then discuss the monotonicity of $I_{\text{inf}}(\bar{\mathbf{X}}; \bar{\mathbf{Y}})$ w.r.t. ϵ_b . The lower-bound mutual information of Eq. (27) can be reformulated as

$$I_{\text{inf}}(\bar{\mathbf{X}}; \bar{\mathbf{Y}}) = \sum_{k=0}^{K-1} \pi_{c,k} \Delta_k, \quad (34)$$

where $\Delta_k = H(p_k \epsilon_b + (1 - p_k)(1 - \epsilon_b)) - H(\epsilon_b)$. The first-order derivative of Δ_k w.r.t. ϵ_b is derived as

$$\frac{d\Delta_k}{d\epsilon_b} = (1 - 2p_k) \log_2 \frac{(2p_k - 1)\epsilon_b + 1 - p_k}{p_k - (2p_k - 1)\epsilon_b} + \log_2 \frac{\epsilon_b}{1 - \epsilon_b}. \quad (35)$$

We have $\frac{d\Delta_k}{d\epsilon_b} = 0$ when $\epsilon_b = 0.5$. The second-order derivative of Δ_k w.r.t. ϵ_b is expressed as

$$\begin{aligned} \frac{d^2\Delta_k}{d\epsilon_b^2} &= \frac{1}{\ln 2} \left(\frac{1}{\epsilon_b(1 - \epsilon_b)} - \frac{(1 - 2p_k)^2}{((2p_k - 1)\epsilon_b + 1 - p_k)(p_k - (2p_k - 1)\epsilon_b)} \right) \\ &= \frac{1}{\ln 2} \frac{p_k(1 - p_k)}{\epsilon_b(1 - \epsilon_b)((2p_k - 1)\epsilon_b + 1 - p_k)(p_k - (2p_k - 1)\epsilon_b)}. \end{aligned} \quad (36)$$

When the BER ϵ_b satisfies $0 < \epsilon_b < 0.5$, we have $\min(0.5, 1 - p_k) < (2p_k - 1)\epsilon_b + 1 - p_k < \max(0.5, 1 - p_k)$ and $\min(0.5, 1 - p_k) < p_k - (2p_k - 1)\epsilon_b < \max(0.5, 1 - p_k)$. Therefore, the second-order derivative satisfies $\frac{d^2\Delta_k}{d\epsilon_b^2} > 0$ subject to constraint of $0 < \epsilon_b < 0.5$. As a result, $\frac{d\Delta_k}{d\epsilon_b}$ monotonically increase w.r.t. ϵ_b . Since we have $\frac{d\Delta_k}{d\epsilon_b}|_{\epsilon_b=0.5} = 0$, it's easy to obtain $\frac{d\Delta_k}{d\epsilon_b}|_{\epsilon_b < 0.5} < 0$, which indicates that Δ_k monotonically decreases w.r.t. ϵ_b . Furthermore, since ϵ_b is a monotonously increasing function of ρ , the lower-bound mutual information $I_{\text{inf}}(\bar{\mathbf{X}}; \bar{\mathbf{Y}})$ monotonically decreases w.r.t. ρ , when we have $\epsilon_b < 0.5$.

REFERENCES

- [1] L. Chettri and R. Bera, "A comprehensive survey on internet of things (IoT) toward 5G wireless systems," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 16–32, Jan 2020.
- [2] B. Buurman, J. Kamruzzaman, G. Karmakar, and S. Islam, "Low-power wide-area networks: Design goals, architecture, suitability to use cases and research challenges," *IEEE Access*, vol. 8, pp. 17 179–17 220, 2020.
- [3] Y. Zhang, S. Shen, C. Y. Chiu, and R. Murch, "Hybrid RF-Solar energy harvesting systems utilizing transparent multiport micromeshed antennas," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 11, pp. 4534–4546, Nov 2019.
- [4] Y. Du, K. Yang, K. Wang, G. Zhang, Y. Zhao, and D. Chen, "Joint resources and workflow scheduling in UAV-enabled wirelessly-powered MEC for IoT systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 10, pp. 10 187–10 200, Oct 2019.
- [5] J. Tang, Y. Yu, M. Liu, D. K. C. So, X. Zhang, Z. Li, and K. Wong, "Joint power allocation and splitting control for SWIPT-Enabled NOMA systems," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 120–133, Jan 2020.
- [6] D. Lim, J. Kang, C. Chun, and H. Kim, "Joint transmit power and time-switching control for device-to-device communications in SWIPT cellular networks," *IEEE Communications Letters*, vol. 23, no. 2, pp. 322–325, Feb 2019.

- [7] Y. Xu, G. Li, Y. Yang, M. Liu, and G. Gui, "Robust resource allocation and power splitting in SWIPT enabled heterogeneous networks: A robust minimax approach," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 799–10 811, Dec 2019.
- [8] D. Xu and H. Zhu, "Outage minimized resource allocation for multiuser OFDM systems with SWIPT," *IEEE Access*, vol. 7, pp. 79 714–79 725, 2019.
- [9] D. Lee, Y. Jang, M. Jung, and S. Choi, "SCLNR-based precoding scheme for multi-user MIMO SWIPT systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 12, pp. 12 392–12 395, Dec 2019.
- [10] H. Lee, K. Lee, H. Kim, and I. Lee, "Joint transceiver optimization for MISO SWIPT systems with time switching," *IEEE Transactions on Wireless Communications*, vol. 17, no. 5, pp. 3298–3312, May 2018.
- [11] Y. Xu, G. Li, Y. Yang, M. Liu, and G. Gui, "Robust resource allocation and power splitting in SWIPT enabled heterogeneous networks: A robust minimax approach," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10 799–10 811, Dec 2019.
- [12] J. Hu, Y. Zhao, and K. Yang, "Modulation and coding design for simultaneous wireless information and power transfer," *IEEE Communications Magazine*, vol. 57, no. 5, pp. 124–130, May 2019.
- [13] Q. Qi, X. Chen, and D. W. K. Ng, "Robust beamforming for NOMA-based cellular massive iot with SWIPT," *IEEE Transactions on Signal Processing*, vol. 68, pp. 211–224, 2020.
- [14] X. Yu, J. Chu, K. Yu, T. Teng, and N. Li, "Energy-efficiency optimization for IoT-distributed antenna systems with SWIPT over composite fading channels," *IEEE Internet of Things Journal*, vol. 7, no. 1, pp. 197–207, Jan 2020.
- [15] F. Chen, J. Fu, Z. Wang, Y. Zhou, and W. Qiu, "Joint communication and computation resource optimization in FD-MEC cellular networks," *IEEE Access*, vol. 7, pp. 168 444–168 454, 2019.
- [16] F. Huang, J. Chen, H. Wang, G. Ding, Y. Gong, and Y. Yang, "Multiple-UAV-assisted SWIPT in internet of things: User association and power allocation," *IEEE Access*, vol. 7, pp. 124 244–124 255, 2019.
- [17] S. He, K. Xie, W. Chen, D. Zhang, and J. Wen, "Energy-aware routing for SWIPT in multi-hop energy-constrained wireless network," *IEEE Access*, vol. 6, pp. 17 996–18 008, 2018.
- [18] Y. Zhao, J. Hu, Z. Ding, and K. Yang, "Joint interleaver and modulation design for multi-user SWIPT-NOMA," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7288–7301, Oct 2019.
- [19] C. Im, J. Lee, and C. Lee, "A multi-tone amplitude modulation scheme for wireless information and power transfer," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 1, pp. 1147–1151, Jan 2020.
- [20] A. Rajaram, D. N. K. Jayakody, B. Chen, R. Dinis, and S. Affes, "Modulation-based simultaneous wireless information and power transfer," *IEEE Communications Letters*, vol. 24, no. 1, pp. 136–140, Jan 2020.
- [21] S. Claessens, N. Pan, D. Schreurs, and S. Pollin, "Multitone FSK Modulation for SWIPT," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 5, pp. 1665–1674, May 2019.
- [22] I. Krikidis and C. Psomas, "Tone-index multisine modulation for SWIPT," *IEEE Signal Processing Letters*, vol. 26, no. 8, pp. 1252–1256, Aug 2019.
- [23] M. Dabirnia and T. M. Duman, "Nonlinear code design for joint energy and information transfer," in *2015 IEEE International Conference on Communications (ICC)*, June 2015, pp. 4247–4252.
- [24] A. M. Fouladgar, O. Simeone, and E. Erkip, "Constrained codes for joint energy and information transfer," *IEEE Transactions on Communications*, vol. 62, no. 6, pp. 2121–2131, June 2014.

- [25] A. Tandon, M. Motani, and L. R. Varshney, "Subblock-constrained codes for real-time simultaneous energy and information transfer," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 4212–4227, July 2016.
- [26] I. Kim, D. I. Kim, and J. Kang, "Rate-energy tradeoff and decoding error probability-energy tradeoff for SWIPT in finite code length," *IEEE Transactions on Wireless Communications*, vol. 16, no. 12, pp. 8220–8234, Dec 2017.
- [27] Y. Hu, Y. Zhu, M. C. Gursoy, and A. Schmeink, "SWIPT-enabled relaying in IoT networks operating with finite blocklength codes," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 1, pp. 74–88, Jan 2019.
- [28] X. Di, K. Xiong, P. Fan, and H. Yang, "Simultaneous wireless information and power transfer in cooperative relay networks with rateless codes," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 4, pp. 2981–2996, April 2017.
- [29] J. Hu, M. Li, K. Yang, S. X. Ng, and K. Wong, "Unary coding controlled simultaneous wireless information and power transfer," *IEEE Transactions on Wireless Communications*, pp. 1–1, 2019.
- [30] Z. Babar, M. A. Mohd Izhar, H. V. Nguyen, P. Botsinis, D. Alanis, D. Chandra, S. X. Ng, R. G. Maunder, and L. Hanzo, "Unary-coded dimming control improves ON-OFF keying visible light communication," *IEEE Transactions on Communications*, vol. 66, no. 1, pp. 255–264, Jan 2018.
- [31] W. Zhang, Z. Song, M. F. Brejza, T. Wang, R. G. Maunder, and L. Hanzo, "Learning-aided unary error correction codes for non-stationary and unknown sources," *IEEE Access*, vol. 4, pp. 2408–2428, 2016.
- [32] F. Yuan, S. Jin, K. Wong, J. Zhao, and H. Zhu, "Wireless information and power transfer design for energy cooperation distributed antenna systems," *IEEE Access*, vol. 5, pp. 8094–8105, 2017.