IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 12, 2019, accepted November 5, 2019, date of publication November 19, 2019, date of current version December 4, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2954163

Design of Coded ALOHA With ZigZag Decoder

MASARU OINAGA, (Student Member, IEEE), SHUN OGATA, (Student Member, IEEE), AND KOJI ISHIBASHI, (Member, IEEE)

Advanced Wireless and Communication Research Center, The University of Electro-Communications, Chofu 182-8585, Japan

Corresponding author: Masaru Oinaga (oinaga@awcc.uec.ac.jp)

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant JP16K06338 and Grant JP18H03765.

ABSTRACT In this study, we investigated the effects of combining a zigzag decoder (ZD) with a coded ALOHA using the successive interference cancellation (SIC) technique to retrieve packets from collisions. We proposed *zigzag decodable coded slotted* ALOHA (ZDCSA) and *enhanced*-ZDCSA (E-ZDCSA) as a scheme which applies ZD before and after SIC, respectively. Furthermore, we derived the asymptotic analysis for the throughput and packet loss rate (PLR) performances of E-ZDCSA and validated its accuracy with the Monte Carlo simulations. Through numerical and asymptotic analyses, we showed that E-ZDCSA outperforms ZDCSA and the conventional coded ALOHA schemes in terms of throughput and PLR performances in most of the offered load regime. Moreover, we demonstrated that our proposed schemes outperform the conventional protocol in a practical scenario in which ZD cannot resolve all the collisions.

INDEX TERMS Slotted ALOHA, successive interference cancellation, ZigZag decoding.

I. INTRODUCTION

Due to the recent advances in the field of Internet of Things (IoT), a system composed of a massive number of devices connected via a wireless network is gathering attention. In particular, massive Machine Type Communications (mMTC) require extensive connectivity with the number of users in the system being more than 10^6 [2]. Further, the system is typically uplink-dominant, and the transmission of users are sporadic. Therefore, the design of a multiple access scheme which satisfies the requirements mentioned above is demanding.

The multiple access scheme can be categorized into two major types: grant access and random access (RA). The grant access scheme such as the time division multiple access (TDMA) is known to achieve high efficiency in terms of communication; however, an overhead of the resource allocation becomes critical when the network size is large. On the other hand, RA schemes such as pure ALOHA [3] does not allocate any resources and allows users to transmit at their own decision. Therefore, its overhead becomes much smaller compared to that of the resource allocation scheme. Nevertheless, when multiple packets arrive simultaneously at a receiver, this causes a collision of packets, thus, resulting in discarding of packets, which significantly limits the throughput performance of pure ALOHA. In [4], *Slotted* ALOHA (SA) has been proposed where users are time-synchronous, and each transmission is done within a time structure referred to as time slots. While SA improves the performance from the classical ALOHA, its peak throughput is still severely limited to $1/e \approx 0.37$.

To conquer the problem of the packet collision, many of the proposed ALOHA protocols exploit the successive interference cancellation (SIC) as a technique for retrieving packets from the collision [5]–[7]. An ALOHA scheme with SIC is referred to as *coded* ALOHA, and among them, the contention resolution diversity slotted ALOHA (CRDSA) [5] is regarded as a milestone, which was originally proposed for satellite communications and later realized as a reasonable option for mMTC. In CRDSA, every user transmits its packet for a given number of times within a time frame composed of several time slots. In [6], the irregular repetition slotted ALOHA (IRSA) was proposed as a generalization of CRDSA, where users individually determine the number of retransmissions based on a degree distribution. Moreover, an asymptotic analysis of the performance of IRSA based on a density evolution [8] was provided, which was inspired from graph-based codes such as low-density parity-check (LDPC) codes [9] and the peeling decoder [10]. It was shown that the IRSA achieved a high-throughput performance comparable with TDMA by optimizing the degree distribution via differential evolution [11]. However, the IRSA suffered from an extreme degradation of the throughput performance after its peak performance. This degradation is due to the limitation

The associate editor coordinating the review of this manuscript and approving it for publication was Oussama Habachi.

of SIC, where its decoding process only starts from packets without a collision, which the occurrence becomes rare in a high-load regime.

As another approach for resolving the collision of packets, a zigzag decoder (ZD) was proposed [12]. ZD was first proposed as a solution to the hidden terminal problem in IEEE 802.11 networks and was later applied to the ALOHA schemes [13]–[15]. In ZD, every packet is assumed to be equipped with a unique word or a *pilot* at both ends. This pilot enables the receiver to detect a time slot with the collision caused by two packets, if the packets are not completely over-wrapped. The receiver then immediately appends an additional time slot and requests the users who sent the packets to retransmit and other users to interrupt its transmission. This operation forms a set of collisions composed of packets from two different users. If the delay of the arrived packets is different between two patterns, ZD is able to retrieve the packets from both users. In [13], it was shown that ZD improved the throughput performance when it was applied to slotted ALOHA. As another example, the performance of the frameless ALOHA [16], a variant of the coded ALOHA inspired by the rateless codes [17], can also be enhanced with the application of ZD [15]. In both cases, it was pointed out that ZD is capable of suppressing the degradation of throughput after the peak. Hence, it is natural to think that ZD is expected to compensate for the problem of the coded ALOHA in the high-load regime. To the best of our knowledge, the performance analysis of a protocol which applies ZD into a framed coded ALOHA scheme has not yet been discussed.

In this study, a zigzag decodable coded slotted ALOHA (ZDCSA) is proposed, where ZD is introduced to the coded ALOHA [6]. First, ZD is introduced before conducting SIC in order to reduce the number of packets that are not yet retrieved. Then, ZD is introduced after SIC in order to retrieve the remaining packets, which is referred to as *enhanced* ZDCSA (E-ZDCSA) in this study. The contributions of this study are summarized as follows:

- We discussed an approach referred to as ZDCSA, where ZD is applied before SIC. We also provided the optimization method of the degree distribution for ZDCSA.
- To improve the throughput performance, we proposed a novel scheme referred to as E-ZDCSA, in which ZD is applied after SIC. We also derived the theoretical performance of E-ZDCSA along with the optimization method of the degree distribution for E-ZDCSA.
- Based on the theoretical and numerical analyses, we showed that E-ZDCSA outperforms IRSA, especially in the high-load regime.

The remainder of this paper is organized as follows. In Section II, a system model is provided. In Section III, SIC is explained in detail, along with a bipartite graph representation. Section IV describes ZD, followed by our proposed protocols and its optimization method in Section V. The numerical results on the throughput performance are presented in Section VI. The conclusions follow in Section VII.

II. SYSTEM MODEL

In this study, a network model as shown in Fig. 1 is considered, where N users transmit their packets to a common receiver. Additionally, all the transmissions are conducted within a time frame that is comprised of M time slots. Therefore, an offered load of the system is given by $G \triangleq N/M$. Every user and the receiver are assumed to be time-synchronous, and the packets are transmitted in a time slot domain. Each user also possesses a packet at the beginning and does not generate new packets during the time frame. If multiple packets simultaneously arrive at the receiver, all of them cannot be immediately retrieved due to packet collision. The collision at the receiver is considered the sole reason for packet loss, *i.e.,* the effects from a physical layer (*e.g.,* fading) is not considered. Moreover, multiple packet reception (MPR) techniques such as exploiting capture effects [18] are also not considered, following a channel model proposed in [19].

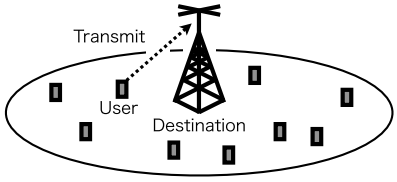

FIGURE 1. Proposed system model. *N* users transmit packets to a common destination (receiver). Each user individually decides its transmission.

Each packet is assumed to arrive at the receiver with a propagation delay within a guard interval installed in the time slot. Therefore, a relationship among the duration of the packet d_{packet}, the propagation delay d_{prop}, and the duration of the time slot d_{slot} is represented as $d_{\text{packet}} + d_{\text{prop}} \leq d_{\text{slot}}$. Besides that, the packet is assumed to be equipped with a unique word at both ends for user identification. This assumption is practical because the information of the transmitter should be included in each transmitted packet.

III. CODED ALOHA

In this section, we briefly describe the conventional coded ALOHA scheme proposed in [6]. We first provide the assumption for the receiver, and then describe the graph representation of the relationship between the transmission of each packet and the observation at each time slot. Finally, we explain the asymptotic analysis and the optimization method for the conventional ALOHA scheme.

First, the receiver is assumed to be able to distinguish the following states of each time slot:

1) *Idle*, where no users have transmitted.
2) *Singleton*, where only one user has transmitted.
3) Multiple users have transmitted, and packets collide.

The packet can be immediately retrieved if and only if it arrived at a singleton time slot. If a packet arrived at a time slot with a collision, it could not be retrieved unless:

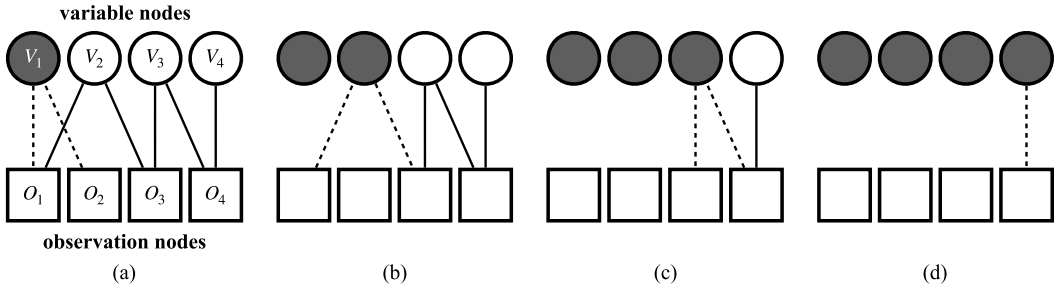

FIGURE 2. Toy example of a bipartite graph where a set of variable nodes V_1, V_2, V_3 and V_4 are connected to observation nodes O_1, O_2, O_3 and O_4 with edges. Each variable node, observation node, and the edge is depicted as a circle, a square, and a line segment, respectively. The variable node is colored grey when the edges connected to that node is removed by SIC, *i.e.,* when the corresponding packet is retrieved.

1) the interference of other packets involved in the collision is removed and the time slot becomes singleton
2) the same packet which arrived at a different time slot is retrieved.

A. BIPARTITE GRAPH REPRESENTATION

The relationship between the transmission of a user and the observation at a time slot can be represented by a bipartite graph composed of *variable nodes*, *observation nodes*, and *edges*. The variable node corresponds to the packet of each user, whereas the observation node corresponds to the packet reception at each time slot. In this study, a set consists of N variable nodes and M observation nodes which are denoted by $\mathcal{V}$ and $\mathcal{O}$, respectively. On the other hand, the edge corresponds to the relationship between each node, *i.e.,* if the i-th user has transmitted a packet at the j-th time slot, $V_i \in \mathcal{V}$ will be connected to $O_j \in \mathcal{O}$ with an edge. The number of edges connected to each node is referred to as a *degree*, and it represents the number of packet transmissions for variable nodes and the number of packet arrivals for observation nodes.

If a packet arrives at the time slot without a collision, it is assumed to be immediately retrieved. On the contrary, if multiple packets arrive simultaneously at the receiver, all of them cannot be retrieved unless an interference cancellation is performed. The coded ALOHA schemes proposed in [5], [6] utilized SIC after receiving M time slots. The decoding procedure of SIC can be described in a bipartite graph as such:

1) Search for an observation node with degree 1, which means a packet arrived at the corresponding time slot without a collision and thus can be immediately retrieved.
2) Track a variable node with the edge connected to the degree 1 observation node found in Step 1 and remove all the edges connected to that variable node.
3) Iterate Steps 1) and 2) until the number of observation nodes with degree 1 becomes 0.

Note that Step 2) leads to additional observation nodes with degree 1 and corresponds to removing the interference of a retrieved signal of the packet from other received signals at the time slot. Additionally, Step 2) is realized by assuming each packet is equipped with a pointer which indicates the time slot in which its replica was transmitted [5]. A dominant factor of the throughput degradation in a high-load regime is the rare occurrence of singleton observation nodes.

The packet recovery via SIC can be explained using a toy example. The bipartite graph shown in Fig. 2 consists of a set of variable nodes $\mathcal{V} = \{V_1, V_2, V_3, V_4\}$, a set of observation nodes $\mathcal{O} = \{O_1, O_2, O_3, O_4\}$ and a number of edges that connect the nodes. (a) First, the edges connected to V_1 are removed as one of them is connected to O_2, which is a singleton node. (b) Next, the edges connected to V_2 are removed after O_1 becomes a singleton node. (c) Then, O_3 becomes a singleton node and thus enables the removal of the edges connected to V_3. (d) Finally, the edge connected to V_4 is removed, and SIC is terminated as the number of nodes with degree 1 became 0. Therefore, SIC can retrieve all the packets in the case of the toy example shown in Fig. 2.

B. PACKET LOSS RATE ANALYSIS BASED ON DENSITY EVOLUTION

1) DEGREE DISTRIBUTIONS

In a system which adopts the IRSA approach, each user determines its packet transmission rate based on a given degree distribution function. Specifically, a user with degree-k selects k random time slots within the time frame. Let us denote the probability that the user transmits k times with L_k. Therefore, the degree distribution of the variable node is given by

$$L(x) \triangleq \sum_{k=1}^{k_{\max}} L_k x^k, \tag{1}$$

where $k_{\max}$ represents the maximum number of the packet transmissions and x is a dummy variable. The distribution, $L(x)$, is configurable by a system designer and is a target of optimization in order to achieve the highest peak throughput for a given $k_{\max}$. Note that the IRSA is regarded as a generalization of CRDSA; CRDSA can be realized by setting

$$L_k = \begin{cases} 1 & (k = k_{\max}) \\ 0 & (\text{otherwise}). \end{cases} \tag{2}$$

Similarly, the degree distribution of the observation node is given by

$$R(x) \triangleq \sum_{k=0}^{N} R_k x^k, \tag{3}$$

where R_k is the probability that the observation node has degree-k. Unlike $L(x)$, $R(x)$ cannot be configured by the system designer and depends on the realization of the system. However, each R_k can be calculated from a binomial distribution as follows:

$$R_k = \binom{N}{k} p^k (1-p)^{N-k}, \tag{4}$$

where p represents the transmission probability of a user per time slot, which is calculated as follows:

$$p = \frac{\sum_\ell \ell R_\ell}{N}. \tag{5}$$

When N, M is sufficiently large and p is very small, R_k can be approximated by Poisson distribution as

$$R_k \approx \frac{(\sum_\ell \ell R_\ell)^k \mathrm{e}^{-(\sum_\ell \ell R_\ell)}}{k!}. \tag{6}$$

From the definition, the average number of packet transmissions per user and packet arrivals per time slot is given by $\sum_k kL_k = L'(1)$ and $\sum_k kR_k = R'(1)$, respectively, where $(\cdot)'$ represents the first-order derivative. Therefore, the offered load of the system can be written as $G = N/M = R'(1)/L'(1)$.

Moreover, the *edge-perspective* degree distribution can be defined by considering the probability of the edge connected to a node with degree-k. Let the probability of an edge connected to a variable node with degree-k be denoted by λ_k. Similarly, the probability of an edge connected to an observation node with degree-k can be denoted by ρ_k. If L_k and R_k are given, then λ_k and ρ_k can be defined as

$$\lambda_k \triangleq \frac{L_k k}{\sum_\ell \ell L_\ell} \tag{7}$$

and

$$\rho_k \triangleq \frac{R_k k}{\sum_\ell \ell R_\ell}, \tag{8}$$

respectively. Then, the edge-perspective degree distribution functions for the variable and observatin nodes can be given by

$$\lambda(x) \triangleq \sum_{k=1}^{k_{\max}} \lambda_k x^{k-1} \tag{9}$$

and

$$\rho(x) \triangleq \sum_{k=1}^{N} \rho_k x^{k-1}, \tag{10}$$

respectively.

2) DENSITY EVOLUTION

The theoretical analysis of the packet loss rate (PLR) of the coded ALOHA can be performed with density evolution [20] which involves iterative calculation over the aforementioned degree distributions. Let q_i denote the probability that the edge is connected to an observation node and not yet removed in the i-th iteration. With the distribution functions given by Eqs. (9) and (10), q_i is given by

$$q_i = \sum_{k=1}^{N} \rho_k \left(1 - (1 - \lambda(q_{i-1}))^{k-1}\right) \tag{11}$$

$$= 1 - \rho\left(1 - \lambda(q_{i-1})\right), \tag{12}$$

where $q_0 = 1$. For the sake of notation simplicity, the index of q_i will be dropped and be denoted by q for the rest of this paper. After q is sufficiently updated with Eq. (12), the resulting PLR is calculated by substituting q for $L(x)$ in Eq. (1). Although this theoretical analysis considers an asymptotic setting where N and M are both infinite, it is still useful for an analysis of the practical settings [6].

C. OPTIMIZATION OF DEGREE DISTRIBUTION

For each value of the offered load G, $R(x)$ and $\rho(x)$ can be derived from $L(x)$ and $\lambda(x)$. Therefore, if G and $L(x)$ are given, the probability q can be derived via the density evolution given by Eq. (12). Then, the offered load G^* in which the throughput is maximized can be derived from

$$\max_G \; G \times \left(1 - L(q)\right) \tag{13}$$

$$\text{s.t. } G > 0, \tag{14}$$

where q is the result of density evolution. Then, the peak throughput, when $L(x)$ is given, can be defined by the following function

$$T\left(L(x)\right) \triangleq G^* \times \left(1 - L(q)\right). \tag{15}$$

The optimum $L(x)$ for each $k_{\max}$ can be obtained from

$$\max_{L(x)} \; T(L(x)) \tag{16}$$

$$\text{s.t. } \sum_{k=1}^{k_{\max}} L_k = 1. \tag{17}$$

In Table 1, the optimized degree distributions derived in [6] are provided. Note that the optimization problem given by Eqs. (16) and (17) can be solved by differential evolution [11].

TABLE 1. Optimized degree distributions for $k_{\max} = 4, 5, 6, 8$ provided in [6].

$k_{\max}$	$L(x)$	$T(L(x))$
4	$0.5102x^2 + 0.4898x^4$	0.868
5	$0.5631x^2 + 0.0436x^3 + 0.3933x^5$	0.898
6	$0.5465x^2 + 0.1623x^3 + 0.2912x^6$	0.915
8	$0.5x^2 + 0.28x^2 + 0.22x^8$	0.938

IV. ZigZag DECODER

As mentioned in the previous section, SIC requires a singleton node in order to start its decoding process. Therefore, if a transmission is done in a manner shown in the bipartite graph in Fig. 3(a), SIC cannot retrieve the packets anymore. The realization of the bipartite graph without the singleton observation nodes is referred to as a *stopping set*, which is regarded as a significant reason behind the degradation of throughput in a high-load regime. In this paper, we refer to the bipartite graph in Fig. 3(a) as the stopping set unless otherwise specified.

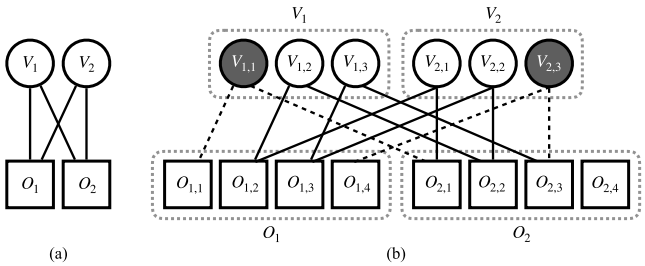

FIGURE 3. Toy example of a bipartite graph of a stopping set composed of two variable nodes, V_1, V_2, and two observation nodes O_1, O_2. The packet-wise representation (a) can be decomposed into segment-wise representation (b).

On the other hand, the zigzag decoder (ZD) proposed in [12] has a chance of resolving the stopping sets such as Fig. 3(a). When considering ZD, each time slot will be further divided into *segments*, and supposes that the packet arrives at the time slot with a *segment-wise* propagation delay. Then, the receiver is capable of distinguishing the following states of each time slot:

1) *Idle*, where no users have transmitted.
2) *Singleton*, where only one user has transmitted.
3) Two users have transmitted, and packets have collided.
4) More than two users have transmitted, and packets have collided.

Note that State 3) is only distinguishable if two packets arrive with a different propagation delay and are not completely overlapped [12]. If State 3) occurs at the time slot, the receiver immediately appends an additional time slot. Then, the receiver broadcasts a signal to request the user to retransmit its packet at the additional time slot if it transmitted at the most recent time slot. If only the requested users transmit their packet at the additional time slot, a stopping set as shown in Fig. 3(a) can be manually formed. Then, ZD will perform a segment-wise interference cancellation over the stopping set. In this study, the two packets involved in the stopping set are retrieved with a probability of ω. Note that $(1-\omega)$ includes the probability that the arrived two packets were overlapped entirely, and therefore, the segment-wise SIC could not be applied.

The decoding process of ZD can be regarded as a segment-wise SIC, which can be explained using the bipartite graph. Suppose that the packet and the time slot is composed of N_p and N_t segments respectively, then, the variable node corresponding to the i-th user can be decomposed into $V_i = \{v_{i,j} \mid 0 \leq j \leq N_p\}$. Similarly, the observation node corresponding to the i-th time slot can be decomposed into $O_i = \{o_{i,j} \mid 0 \leq j \leq N_t\}$. Fig. 3(b) shows an example with $N_p = 3$ and $N_t = 4$, where each node is decomposed into $V_1 = \{V_{1,1}, V_{1,2}, V_{1,3}\}$, $V_2 = \{V_{2,1}, V_{2,2}, V_{2,3}\}$, $O_1 = \{O_{1,1}, O_{1,2}, O_{1,3}, O_{1,4}\}$, and $O_2 = \{O_{2,1}, O_{2,2}, O_{2,3}, O_{2,4}\}$. The decoding process is initiated by retrieving segments $V_{1,1}$ and $V_{2,3}$ which arrive respectively at the singleton segments $O_{1,1}$ and $O_{1,4}$. After removing the interference of $V_{1,1}$ and $V_{2,3}$ from all the other segments, segments $O_{2,1}$ and $O_{2,3}$ become singleton. Then, segments $V_{1,3}$ and $V_{2,1}$ can be retrieved respectively from the singleton segments $O_{2,1}$ and $O_{2,3}$. Finally, segments $V_{1,2}$ and eventually $V_{2,2}$ can be retrieved after the interference cancellation. Therefore, all the segments of both V_1 and V_2 can be retrieved via ZD in this example.

V. PROPOSED PROTOCOLS

In this section, a combination of the zigzag decoder and the conventional coded ALOHA scheme is considered.

A. ZDCSA

In ZDCSA, the receiver changes its behavior only if two packets arrived at the same time slot, and otherwise operate similarly with IRSA. Moreover, based on the state of the time slot, the receiver successively operates while receiving packets. When State 3) shown in the previous section occurs, the receiver appends a time slot immediately afterward and requests users who sent the packet to retransmit their packet. The requested packet retransmission is done in the appended time slot, where the left of the users halt their transmission until the retransmission is completed. Thus, the stopping set is manually created, where involved packets can be retrieved via ZD with the probability of ω. If packets are successfully retrieved, the transmitters of that packet halt its packet transmission for the left of the time frame. The time frame is terminated if the receiver received M time slots without counting the appended time slots.

In order to consider the overhead of ZD, we defined the ratio of the number of time slots added via the ZD to M as α. Therefore, we redefined the throughput calculation as

$$T_{\mathrm{ZD}} \triangleq \frac{G}{1+\alpha} \times (1-\beta) \tag{18}$$

where β represents PLR. Furthermore, the peak throughput when $L(x)$ is applied is denoted by $T_{\mathrm{ZD}}(L(x))$. Note that for schemes without ZD, such as the IRSA and CRDSA, α is set to 0.

Unlike the conventional coded ALOHA, ZDCSA is incapable of strict analysis based on density evolution. The problem of analytical tractability is caused by whether the ZD event that occurs at a particular time slot depends on the previous slots to that time slot. Therefore, the degree distribution for ZDCSA is optimized to maximize the peak throughput in a finite setting. The optimization problem can be written as

$$\max_{L(x)} \; T^{\mathrm{ZD}}(L(x)) \tag{19}$$

$$\text{s.t.} \sum_{k=0}^{k_{\max}} L_k = 1, \tag{20}$$

where $T^{\mathrm{ZD}}(L(x))$ in this study was obtained from the Monte Carlo simulation with $N = 10^4$. In Table 2, the results of the optimization via differential evolution for $k_{\max} = 4, 5, 6, 8$ are presented.

The peak throughput comparison between the IRSA and ZDCSA with $k_{\max} = 4, 5, 6, 8$ is presented in Fig. 4. As seen

TABLE 2. Optimized distribution function for ZDCSA in case of $k_{\max} = 4, 5, 6, 8$.

$k_{\max}$	$L(x)$
4	$0.3761x^2 + 0.0683x^3 + 0.5556x^4$
5	$0.4162x^2 + 0.0852x^3 + 0.2170x^4 + 0.2816x^5$
6	$0.4677x^2 + 0.0296x^4 + 0.1646x^5 + 0.2524x^6$
8	$0.4280x^2 + 0.1999x^3 + 0.0523x^4$ $+0.0071x^5 + 0.0223x^6 + 0.2016x^7 + 0.0888x^8$

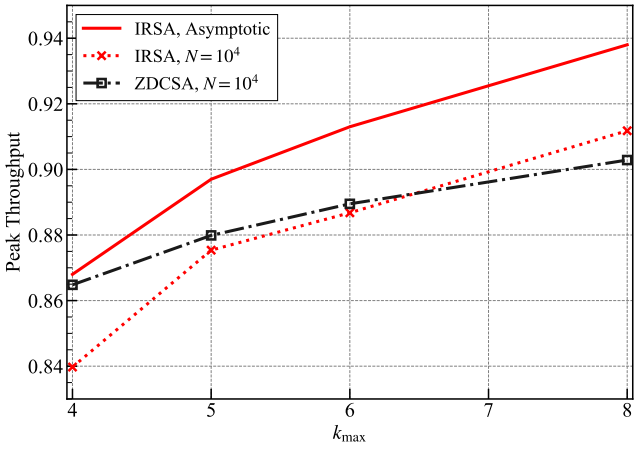

FIGURE 4. Peak throughput comparison between IRSA and ZDCSA with $k_{\max} = 4, 5, 6, 8$.

in the figure, the peak throughput performance of the IRSA and ZCDSA improves as $k_{\max}$ increases. The figure also reveals that the IRSA outperforms ZDCSA when $k_{\max}$ becomes 8 or in the asymptotic setting. This outcome is due to two reasons:

1) The degree distribution of the IRSA is optimized via an asymptotic setting. The optimization $L(x)$ for ZDCSA is done in a non-asymptotic setting *i.e.,* when N becomes large, the IRSA will eventually outperform ZDCSA in terms of peak throughput.
2) Time slots are added excessively by the ZD, even in a load region where SIC is capable of retrieving packets of all the users.

Therefore, a scheme which achieves the performance comparable to the IRSA with less addition of time slots is required.

B. E-ZDCSA

Next, another novel approach referred to as the *enhanced*-ZDCSA (E-ZDCSA) is proposed. In E-ZDCSA, the packet transmission and packet recovery is first conducted in an IRSA manner. Within the received time frame, the receiver first appends time slot if it detects the time slot with a collision caused by two packets. Then, it addresses the transmitter of the packets involved in the collided time slot to retransmit its packet to the appended time slot in order to form the stopping set. Finally, the stopping set is resolved via ZD with a probability of ω. As ZD is operated if SIC could not resolve a collision of packets, E-ZDCSA enables a decrease in the number of appended time slots compared to ZDCSA. Furthermore, E-ZDCSA is capable of optimizing its performance via asymptotic analysis based on density evolution.

1) ASYMPTOTIC ANALYSIS BASED ON DENSITY EVOLUTION

First, the node-perspective degree distributions after SIC are derived in order to track the asymptotic performance of E-ZDCSA. Let L_k^{SIC} denote the probability that a variable node has degree-k after SIC. L_k^{SIC} can be derived by $L_k^{\mathrm{SIC}} = l_k / \sum_\ell l_\ell$, where l_k $(k \geq 0)$ is calculated by

$$l_k = \begin{cases} 1 - \sum_{\ell=1} L_\ell q^\ell & (k = 0) \\ L_k q^k & (k > 0), \end{cases} \tag{21}$$

and q is the result of density evolution given by Eq. (12).

In the same manner as L_k^{SIC}, let R_k^{SIC} denote the probability that an observation node has degree-k after SIC. R_k^{SIC} can be derived by $R_k^{\mathrm{SIC}} = r_k / \sum_\ell r_\ell$, where r_k $(k \geq 0)$ is calculated by

$$r_k = \begin{cases} \displaystyle\sum_{m=0}^{1} \sum_{\ell=m}^{N} R_\ell \binom{\ell}{m} \lambda(q)^m (1 - \lambda(q))^{\ell - m} & (k = 0) \\ 0 & (k = 1) \\ \displaystyle\sum_{\ell=k}^{N} R_\ell \binom{\ell}{k} \lambda(q)^k (1 - \lambda(q))^{\ell - k} & (k > 1). \end{cases} \tag{22}$$

From the above definitions, the degree distributions of a variable and an observation node after SIC can be defined as

$$L^{\mathrm{SIC}}(x) \triangleq \sum_{k=0}^{k_{\max}} L_k^{\mathrm{SIC}} x^k \tag{23}$$

and

$$R^{\mathrm{SIC}}(x) \triangleq \sum_{k=0}^{N} R_k^{\mathrm{SIC}} x^k, \tag{24}$$

respectively. Next, let λ_k^{SIC}, ρ_k^{SIC} denote the probability that an edge is connected to a degree-k variable and observation node, respectively. From L_k^{SIC} and R_k^{SIC}, λ_k^{SIC} and ρ_k^{SIC} can be derived by

$$\lambda_k^{\mathrm{SIC}} \triangleq \frac{L_k^{\mathrm{SIC}} k}{\sum_\ell \ell L_\ell^{\mathrm{SIC}}} \tag{25}$$

and

$$\rho_k^{\mathrm{SIC}} \triangleq \frac{R_k^{\mathrm{SIC}} k}{\sum_\ell \ell R_\ell^{\mathrm{SIC}}}, \tag{26}$$

respectively. Then, the edge-perspective degree distributions after SIC are given by

$$\lambda^{\mathrm{SIC}}(x) \triangleq \sum_{k=1}^{k_{\max}} \lambda_k^{\mathrm{SIC}} x^{k-1} \tag{27}$$

and

$$\rho^{\mathrm{SIC}}(x) \triangleq \sum_{k=1}^{N} \rho_k^{\mathrm{SIC}} x^{k-1}, \tag{28}$$

respectively. When ZD is utilized, the edges connected to an observation node with degree-2 can be removed with a probability of ω. Hence, the probability that an edge is not removed after ZD is derived from

$$q^{\mathrm{ZD}} \triangleq q(1-\omega\rho_2^{\mathrm{SIC}}). \tag{29}$$

Therefore, the PLR after ZD can be calculated by $L(q^{\mathrm{ZD}})$.

2) OPTIMIZATION OF $L(x)$ IN CASE OF E-ZDCSA

In the same manner as IRSA, the offered load $G^\star$ in which the throughput is maximized can be derived by

$$\max_{G} \; G \times \left(1 - L(q^{\mathrm{ZD}})\right) \tag{30}$$

$$\text{s.t. } G > 0. \tag{31}$$

Then, the peak throughput of E-ZDCSA when $L(x)$ is given can be defined by the following function:

$$T^{\mathrm{E-ZD}}\big(L(x)\big) \triangleq \frac{G^\star}{1+\alpha} \times \left(1 - L(q^{\mathrm{ZD}})\right). \tag{32}$$

Therefore, the optimum $L(x)$ for each $k_{\max}$ can be obtained from

$$\max_{L(x)} \; T^{\mathrm{E-ZD}}(L(x)) \tag{33}$$

$$\text{s.t. } \sum_{k=1}^{k_{\max}} L_k = 1. \tag{34}$$

In Table 3, the results of optimization via differential evolution for $k_{\max} = 4, 5, 6, 8$ are presented. Note that the IRSA and E-ZDCSA share the same peak throughput performance when the utilized node distributions are the same.

TABLE 3. Optimized degree distributions for E-ZDCSA in case of $k_{\max} = 4, 5, 6, 8$.

$k_{\max}$	$L(x)$	$T(L(x))$
4	$0.5130x^2 + 0.0013x^3 + 0.4857x^4$	0.8682
5	$0.5630x^2 + 0.0413x^3 + 0.3956x^5$	0.8977
6	$0.5485x^2 + 0.1389x^3 + 0.0276x^4$ $+0.0168x^5 + 0.2682x^6$	0.9130
8	$0.5116x^2 + 0.2633x^3 + 0.0003x^4 + 0.0019x^5$ $+0.0048x^6 + 0.0405x^7 + 0.1776x^8$	0.9381

VI. NUMERICAL ANALYSIS

In this section, the performance of our proposed protocols are compared with the conventional coded ALOHA schemes. First, the accuracy of the proposed node distribution analysis provided in Section IV is validated. Then, the throughput and PLR performances of ZDCSA and E-ZDCSA are compared to that of the conventional IRSA. Finally, the analysis when ω is altered is provided.

A. DISTRIBUTIONS AFTER SIC

In this subsection, the theoretical and simulation results of node distribution after SIC are compared. In order to measure the difference between the distributions, the Kullback-Leiber (KL) divergence [21] denoted by $D_{\mathrm{KL}}(P||Q)$ is introduced, where P and Q are both a discrete probability mass function (PMF). In this study, P and Q are derived via the asymptotic and numerical analysis, respectively.

Fig. 5 shows the result of the KL divergence when comparing the theoretical and numerical results of the distribution of the variable and observation node after SIC. For both the asymptotic and simulation settings, the node distribution $L(x)$ is set as

$$L(x) = 0.5x^2 + 0.28x^3 + 0.22x^8, \tag{35}$$

and $N = 10^4$ for the Monte Carlo simulation. From the result shown in Fig. 5, the theoretical analysis proposed in this study enables the node distributions after SIC to be tracked in most cases except in the moderate-to-high load regime, or the so-called *waterfall region*. Although a finite analysis of IRSA in the waterfall region is provided in [22], it is not introduced in this study for the sake of simplicity.

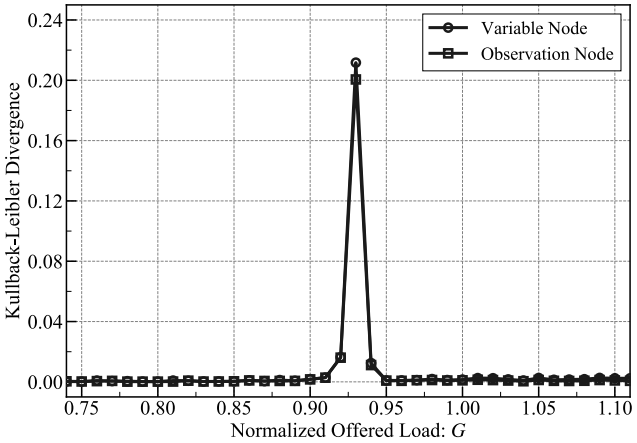

FIGURE 5. The KL divergence when comparing the theoretical and numerical results of the variable and observation node distribution after SIC.

B. THROUGHPUT AND PLR PERFORMANCES

Next, the throughput and PLR performances of the proposed schemes (ZDCSA and E-ZDCSA) are compared with the conventional scheme (IRSA). The node distribution $L(x)$ is set as

$$L(x) = 0.5116x^2 + 0.2633x^3 + 0.0003x^4 + 0.0019x^5 + 0.0048x^6 + 0.0405x^7 + 0.1776x^8, \tag{36}$$

for IRSA and E-ZDCSA, and

$$L(x) = 0.4280x^2 + 0.1999x^3 + 0.0523x^4 + 0.0071x^5 + 0.0223x^6 + 0.2016x^7 + 0.0888x^8 \tag{37}$$

for ZDCSA. Further, N is set to 10^4 for a non-asymptotic setting. Fig. 6 and Fig. 7 show the comparison between each scheme in terms of the throughput and PLR performances, respectively. For both analyses, the results of the Monte Carlo simulation validated the asymptotic analysis presented in the previous section.

As seen from Fig. 6, both the proposed schemes with ZD clearly suppressed the degradation of the throughput

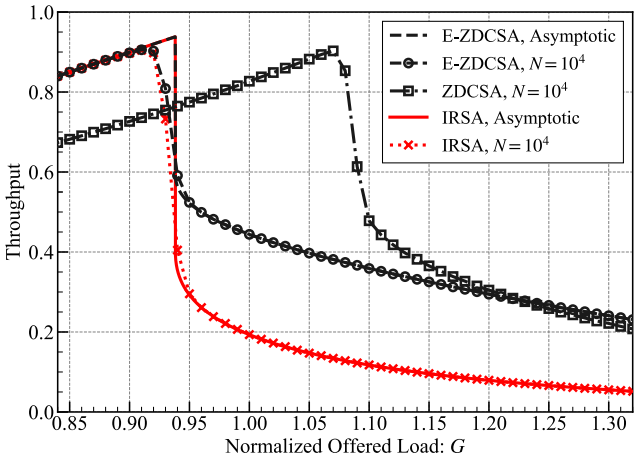

FIGURE 6. Throughput performances of IRSA, ZDCSA, and E-ZDCSA.

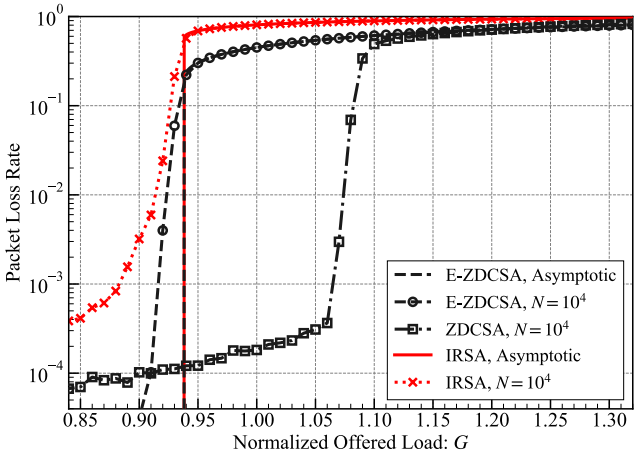

FIGURE 7. PLR performances of IRSA, ZDCSA, and E-ZDCSA.

performance in a high-load regime compared to the conventional IRSA. Besides that, the effect of the overhead caused by appending time slots via ZD is significantly depicted in the result of ZDCSA; although the peak throughput was relatively high, its throughput performance was degraded in a moderate-load regime. On the other hand, E-ZDCSA showed a throughput performance similar to that of IRSA, where the throughput rises linearly to N/M until its peak. After its peak, E-ZDCSA outperformed the IRSA thanks to ZD, and eventually, ZDCSA because of less overhead. In the peak throughput performance comparison, both E-ZDCSA and IRSA achieved 0.9381 in the asymptotic setting. For non-asymptotic setting with $N = 10^4$, the peak throughputs of IRSA, ZDCSA, and E-ZDCSA were 0.9045, 0.9029 and 0.9056, respectively. Therefore, it can be deduced that E-ZDCSA is the most effective scheme in terms of throughput performance.

From Fig. 7, it can be seen that E-ZDCSA and IRSA did not have an error floor in the asymptotic setting. Conversely, in the non-asymptotic setting, the proposed schemes showed a lower error floor compared to IRSA. Notably, E-ZDCSA drew its error floor lower than the other two schemes and achieved the lowest PLR in the region of $G \leq 0.91$. While the offered load G was in the range of [0.91, 1.23], ZDCSA showed the lowest PLR owing to the structure of applying ZD before SIC. After G became higher than 1.23, E-ZDCSA outperformed ZDCSA by achieving the lowest PLR among all the schemes in comparison. Therefore, it can be concluded that E-ZDCSA is the most effective scheme in almost all the load regimes in terms of PLR performance.

C. THROUGHPUT PERFORMANCE IN A PRACTICAL SCENARIO

Finally, this subsection observes the peak throughput performance of the proposed schemes when ZD success probability ω is changed. As the transmission of users is typically sporadic [2] in the mMTC scenario, the number of active users is assumed to be relatively small compared to the number of all users. Therefore, this subsection considered a smaller number of users compared to the previous analyses. As described in Section IV, ZD can be applied if two packets arrived with a different propagation delay. For simplicity, the delay of the packet of each user can be regarded as a random selection of back-off patterns. Then, it can be said that the ZD fails only if two users picked the same back-off pattern. Therefore, the ZD success probability can be calculated as

$$\omega = 1 - N_{\mathrm{B}}^{-1}, \tag{38}$$

where N_{B} denotes the number of back-off patterns.

Fig. 8 shows the peak throughput performance comparison of IRSA, ZDCSA and E-ZDCSA in $\omega \in [0, 1]$ when N is set to 10^3. Note that the node distribution is configured according to Eq. (36) for IRSA and E-ZDCSA, and Eq. (37) for E-ZDCSA. From the result of $\omega = 0$, it can be seen that E-ZDCSA significantly decreased the number of appending time slots compared to ZDCSA. For E-ZDCSA, it outperformed IRSA when ω was greater than 0.5. Therefore, only $(1 - \omega)^{-1} = 2$ back-off patterns were required in order for E-ZDCSA to achieve better performance than the conventional IRSA. In the case of ZDCSA, it outperformed

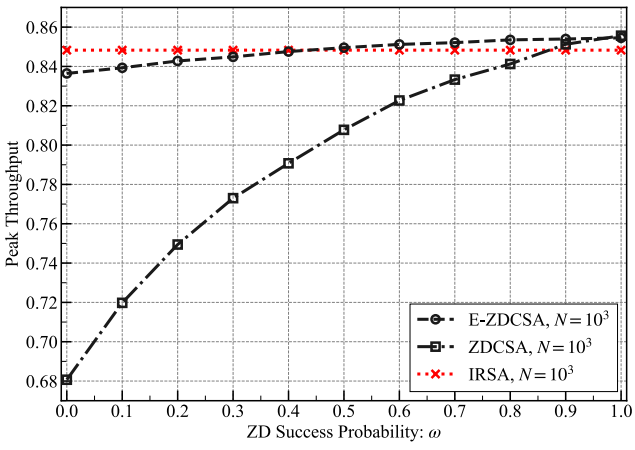

FIGURE 8. Peak throughput performance with $\omega \in [0, 1]$ and $N = 10^3$.

IRSA with $\omega \geq 0.9$, meaning that the number of back-off patterns in order to outperform IRSA was approximately 10. In particular, N_B was initialized to 32 in IEEE 802.11 [12], which yielded $\omega = (1 - 32)^{-1} \approx 0.969$. As shown in the figure, both ZDCSA and E-ZDCSA outperformed the IRSA when ω was set to 0.969. Therefore, the proposed protocols are effective in terms of the practical scenario.

VII. CONCLUSION

In this study, the effects of applying ZD into a coded ALOHA scheme in terms of throughput and PLR performances were investigated. Namely, ZDCSA and E-ZDCSA were proposed as a scheme which applies ZD before and after SIC is conducted, respectively. Furthermore, the asymptotic analysis for E-ZDCSA was derived, and its accuracy was validated with the Monte Carlo simulations. Through the numerical and asymptotic analyses, it was revealed that E-ZDCSA outperformed ZDCSA and the conventional IRSA in terms of throughput performance and PLR performance in most of the offered load regime. Moreover, the proposed schemes have also outperformed the conventional scheme in a practical scenario with $\omega \leq 1$. Note that it remains as a future work to take the effects of the physical layer such as capture effects into consideration.

ACKNOWLEDGMENT

This article was presented in part at the 15th Workshop on Positioning, Navigation and Communications (WPNC), Bremen, Germany, October 2018 [1].

REFERENCES

[1] M. Oinaga, S. Ogata, and K. Ishibashi, "ZigZag decodable coded slotted ALOHA," in *Proc. IEEE 15th Workshop Positioning, Navigat. Commun. (WPNC)*, Bremen, Germany, Oct. 2018, pp. 1–6.

[2] C. Bockelmann, "Towards massive connectivity support for scalable mMTC communications in 5G networks," *IEEE Access*, vol. 6, pp. 28969–28992, 2018.

[3] N. Abramson, "THE ALOHA SYSTEM: Another alternative for computer communications," in *Proc. Fall Joint Comput. Conf.*, New York, NY, USA, vol. 37, Nov. 1970, pp. 281–285.

[4] L. G. Roberts, "ALOHA packet system with and without slots and capture," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 5, no. 2, pp. 28–42, Apr. 1975.

[5] E. Casini, R. De Gaudenzi, and O. del Rio Herrero, "Contention resolution diversity slotted ALOHA (CRDSA): An enhanced random access scheme for satellite access packet networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 4, pp. 1408–1419, Apr. 2007.

[6] G. Liva, "Graph-based analysis and optimization of contention resolution diversity slotted ALOHA," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 477–487, Feb. 2011.

[7] E. Paolini, G. Liva, and M. Chiani, "Coded slotted ALOHA: A graph-based method for uncoordinated multiple access," *IEEE Trans. Inf. Theory*, vol. 61, no. 12, pp. 6815–6832, Dec. 2015.

[8] T. Richardson and R. Urbanke, *Modern Coding Theory*. New York, NY, USA: Cambridge Univ. Press, 2008.

[9] R. G. Gallagers, *Low-Density Parity-Check Codes*. Cambridge, MA, USA: MIT Press, 1963.

[10] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Improved low-density parity-check codes using irregular graphs," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 585–598, Feb. 2001.

[11] R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, Dec. 1997.

[12] S. Gollakota and D. Katabi, "ZigZag decoding: Combating hidden terminals in wireless networks," in *Proc. ACM SIGCOMM Conf. Data Commun.*, Seattle, WA, USA, Aug. 2008, pp. 159–170.

[13] J. Paek and M. J. Neely, "Mathematical analysis of throughput bounds in random access with ZIGZAG decoding," in *Proc. 7th Int. Symp. Modeling Optim. Mobile, Ad Hoc, Wireless Netw. (WiOpt)*, Jun. 2009, pp. 1–7.

[14] S. Ogata and K. Ishibashi, "ZigZag decodable frameless ALOHA," in *Proc. 52nd Asilomar Conf. Signals, Syst., Comput.*, Oct. 2018, pp. 504–507.

[15] S. Ogata and K. Ishibashi, "Application of ZigZag decoding in frameless ALOHA," *IEEE Access*, vol. 7, pp. 39528–39538, 2019.

[16] C. Stefanovic, P. Popovski, and D. Vukobratovic, "Frameless ALOHA protocol for wireless networks," *IEEE Commun. Lett.*, vol. 16, no. 12, pp. 2087–2090, Dec. 2012.

[17] M. Luby, "LT codes," in *Proc. 43rd Annu. IEEE Symp. Found. Comput. Sci.*, Vancouver, BC, Canada, Nov. 2002, pp. 271–280.

[18] A. Zanella and M. Zorzi, "Theoretical analysis of the capture probability in wireless systems with multiple packet reception capabilities," *IEEE Trans. Commun.*, vol. 60, no. 4, pp. 1058–1071, Apr. 2012.

[19] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[20] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 619–637, Feb. 2001.

[21] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Statist.*, vol. 22, no. 1, pp. 79–86, 1951.

[22] A. G. I. Amat and G. Liva, "Finite-length analysis of irregular repetition slotted ALOHA in the waterfall region," *IEEE Commun. Lett.*, vol. 22, no. 5, pp. 886–889, May 2018.

MASARU OINAGA (S'18) received the B.E. degree in engineering from The University of Electro-Communications, Tokyo, Japan, in 2018, where he is currently pursuing the M.E. degree. His current research interests are random access protocols and information theory.

SHUN OGATA (S'14) received the B.E., M.E., and Ph.D. degrees in engineering from The University of Electro-Communications, Tokyo, Japan, in 2014, 2016, and 2019, respectively. His current research interests include communication theory, channel coding, and information theory.

KOJI ISHIBASHI (S'01–M'07) received the B.E. and M.E. degrees in engineering from The University of Electro-Communications, Tokyo, Japan, in 2002 and 2004, respectively, and the Ph.D. degree in engineering from Yokohama National University, Yokohama, Japan, in 2007. From 2007 to 2012, he was an Assistant Professor with the Department of Electrical and Electronic Engineering, Shizuoka University, Hamamatsu, Japan. From 2010 to 2012, he was a Visiting Scholar with the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA. Since 2012, he has been with the Advanced Wireless and Communication Research Center, The University of Electro-Communications, where he is currently an Associate Professor. His current research interests are energy harvesting communications, wireless power transfer, channel codes, signal processing, and information theory.