PAPER

# Polling Schedule Algorithms for Data Aggregation with Sensor Phase Control in In-Vehicle UWB Networks*

**Hajime MIGITA**[†a)], ***Student Member***, **Yuki NAKAGOSHI**[††b)], **Patrick FINNERTY**[†††c)], ***Nonmembers***, **Chikara OHTA**[†††d)], ***Senior Member***, ***and*** **Makoto OKUHARA**[†e)], ***Nonmember***

**SUMMARY** To enhance fuel efficiency and lower manufacturing and maintenance costs, in-vehicle wireless networks can facilitate the weight reduction of vehicle wire harnesses. In this paper, we utilize the Impulse Radio-Ultra Wideband (IR-UWB) of IEEE 802.15.4a/z for in-vehicle wireless networks because of its excellent signal penetration and robustness in multipath environments. Since clear channel assessment is optional in this standard, we employ polling control as a multiple access control to prevent interference within the system. Therein, the preamble overhead is large in IR-UWB of IEEE 802.15.4a/z. Hence, aggregating as much sensor data as possible within each frame is more efficient. In this paper, we assume that reading out data from sensors and sending data to actuators is periodical and that their respective phases can be adjusted. Therefore, this paper proposes an integer linear programming-based scheduling algorithm that minimizes the number of transmitted frames by adjusting the read and write phases. Furthermore, we provide a heuristic algorithm that computes a suboptimal but acceptable solution in a shorter time. Experimental validation shows that the data aggregation of the proposed algorithms is robust against interference.

***key words:*** *UWB, in-vehicle network, wireless network control system, scheduling optimization, heuristic scheduling, IEEE 802.14.5*

## 1. Introduction

In-vehicle wire harnesses used for data transmission and power delivery are gaining weight due to the increase in electrical components, reducing fuel efficiency and causing manufacturing and maintenance costs. For example, the weight of a harness can reach 10 kg, 20 kg, and in excess of 40 kg in light, standard, and large cars, respectively. In the era of automated driving, we anticipate an increase in the number of electric devices in motor vehicles, further adding to the weight of wire harnesses. Replacing part of a wire harness with a "wireless harness" is expected to relieve these problems [1].

Manuscript received September 19, 2023.
Manuscript revised December 8, 2023.
Manuscript publicized March 5, 2024.

[†]Graduate School of Science, Technology and Innovation, Kobe University, Kobe-shi, 657-8501 Japan.
[††]Faculty of Engineering, Kobe University, Kobe-shi, 657-8501 Japan.
[†††]Graduate School of System Informatics, Kobe University, Kobe-shi, 657-8501 Japan.

*This paper is an extended version of [4], including a new heuristic scheduling algorithm and experimental results.

a) E-mail: 224p212p@gsuite.kobe-u.ac.jp
b) E-mail: 2065019t@gsuite.kobe-u.ac.jp
c) E-mail: finnerty.patrick@fine.cs.kobe-u.ac.jp
d) E-mail: ohta@port.kobe-u.ac.jp
e) E-mail: okuhara@fine.cs.kobe-u.ac.jp

DOI: 10.23919/transcom.2023EBP3150

In collaboration with DENSO TEN Limited and the Advanced Telecommunications Research Institute International (ATR), we have been trying to reduce the weight of wire harnesses by substituting them for a combination of Impulse Radio Ultra-Wideband (UWB) and Power Line Communication (PLC) [2]–[4]. IR-UWB is suitable for in-vehicle applications due to its high tolerance to multipath propagation and penetration, and it has already shown high reliability in the in-vehicle environment [5]–[11]. The power lines of current harnesses need to remain to power these electronic devices. Thus, there is an opportunity to use these lines as a medium for PLC [1]. In this paper, we focus on the wireless part and leave the integration with PLC as future work. The power consumption of the devices and their impact on the performance of the motor vehicle remains outside the scope of this study.

In contrast to using wireless harnesses to reduce the weight of in-vehicle networks, adopting Ethernet represents a different approach. This trend aims to provide the high-speed connectivity and ample bandwidth necessary to support advanced driver assistance systems (ADAS). However, most in-vehicle electronic devices do not require such high-speed communications. In practice, a reference-model integrated board [12] engineered for vehicle computers and service-oriented gateways (SoG) supports Ethernet and controller area network (CAN). Therefore, we aim to substitute CAN by UWB.

An in-vehicle ECU collects sensing data from sensors and sends control data to actuators, usually cyclically, via the harness. Since vehicles are controlled in real-time, scheduling is essential to satisfy the strict latency and data loss constraints. Further, a lightweight scheduling algorithm is desired so that even a weak ECU can compute it quickly. Increasing the channel efficiency is also essential because, as explained in Sect. 2.2, the preamble overhead is not negligible. Thus, this paper proposes polling scheduling with data aggregation for the UWB wireless harness. In this context, "aggregation" means aggregating and conveying as much data as possible within a MAC frame.

In-vehicle systems can be regarded as networked control systems. Some wireless schedule algorithms are found in Refs. [14]–[17]. These papers propose scheduling algorithms for Time Division Multiple Access (TDMA) using linear programming and heuristic. These algorithms aim to transmit data as evenly as possible in each subframe to satisfy latency constraints under periodic data generation

conditions. They purposely leave unused space in every subframe so that event-triggered sensors can transmit with minimal latency, and lost data can be retransmitted before their deadline. These scheduling algorithms, however, limit the data transmission cycle for each sensor or actuator to a multiple of the base cycle. For this reason, these scheduling algorithms cannot be adapted to our scenario, which will be explained in Sect. 2. In particular, Ref. [14] primarily targets in-vehicle UWB networks, but it does not consider UWB preamble overhead.

In terms of aggregation, Ref. [18] proposes a frame aggregation method using the aggregated-MAC protocol data unit (A-MPDU) that works with the IEEE 802.15.4 MAC layer to implement a patient monitoring system. In this context, many sensing data generated cyclically in the system must be collected with strict latency constraints and quality-of-service requirements satisfied. This paper shows that analyzing the traffic pattern and aggregating and transmitting frames is more effective than sending frames frequently. Reference [19] shows that aggregation can reduce physical layer overhead and improve communication efficiency in wireless time-sensitive networking. However, it also points out that a trade-off between aggregation and packet error rate (PER) is needed, indicating that aggregation has a side effect when the bit error rate (BER) is large.

Regarding computationally lightweight real-time scheduling algorithms, Early Deadline First (EDF) [20] and Round-robin (RR) [21] are well-known. EDF prioritizes data with the earliest deadline and selects them for transmission. This contributes to satisfying latency constraints, but considering only the current state makes aggregation difficult. Further, its work-conserving feature can cause transmissions to be temporally concentrated, leaving insufficient room for retransmission. As its name suggests, RR is a fair scheduling scheme in which a parent terminal (PT) gives rights to send data to each child terminal (CT) equally in a round-robin fashion. RR, however, cannot take data aggregation into account either.

Considering the above, this paper proposes scheduling algorithms based on integer linear programming (ILP) and heuristic, minimizing the number of transmitted frames to increase channel efficiency by adjusting the phases of the sensor data. This paper is an extended version of [2]–[4]. In Refs. [2], [3], the scheduling algorithm is formulated as an ILP problem to reduce the number of polling slots. However, the freedom in the phase of periodic data generation was not fully exploited and presents opportunities to further reduce the number of response frames. Consequently, in [4], we endeavored to directly minimize the response frames by incorporating the data generation phase as a variable. While effective, this approach led to an increase in the number of variables, causing optimization algorithms computation time to increase. Moreover, periodically recalculating the schedule as sensors are added or removed is necessary. To mitigate this increase in computation time, we fix the polling target in the ILP, as will be elaborated in Sect. 2.3. Furthermore, given the limited computational resources of ECUs, we developed a heuristic to further reduce the computation time.

The contributions of this paper are as follows:

1. The modified scheduling algorithm is described in detail.
2. A new heuristic scheduling algorithm that reduces the number of transmitted frames is proposed.
3. Experimental results demonstrating the effectiveness of our heuristic scheduling algorithm are shown.
4. Additional experiments with different data emission patterns for both the ILP-based and heuristic-based algorithms are also presented.

The rest of this paper is organized as follows. First, we describe an overview of the in-vehicle UWB network, UWB features, and polling control methods in Sect. 2. In Sect. 3, we formulate the schedule as an Integer Linear Programming (ILP) optimization problem. Then, we propose a heuristic scheduling algorithm to compute within reasonable time a sub-optimal schedule in Sect. 4. In Sect. 5, we present our experimental results and demonstrate the robustness of the schedules against interference. Finally, we conclude in Sect. 6.

## 2. System Model

This section presents an overview of in-vehicle networks, communication requirements, and control schemes.

### 2.1 UWB/PLC Integrated In-Vehicle Network

An overview of the UWB/PLC integrated in-vehicle network we envision is presented in Fig. 1. The network comprises one PT and five CTs. The PT serves as the electronic control unit and reads out data from sensors connected to CTs. When needed, the PT also sends control data to the actuators connected to CTs. In this paper, we only consider the uplink communication of the UWB part. More precisely, we consider that the PT periodically collects sensor data from the CTs via UWB.

Each sensor has a specific generation cycle, depending on its type. The length of sensor data is 6 octets, consisting of 2 octets for the sensor ID and 4 octets for sensor value. In this study, we consider two scenarios with different sensor data cycles: short cycles and long cycles. Tables 1 and 2 show the number of sensors connected to each CT and their

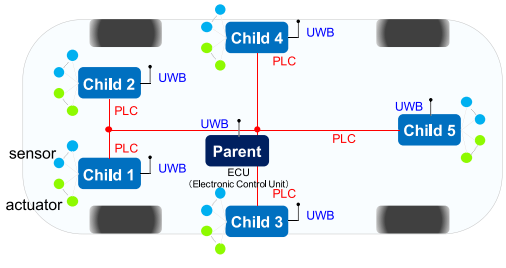



**Fig. 1** UWB/PLC integrated in-vehicle network.

**Table 1** Number of sensors and generate cycles for the short cycles case.

| CT | Generation Cycle (ms) | | | |
|---|---|---|---|---|
| | 24 | 56 | 72 | 108 |
| 1 | 16 | 12 | 10 | 2 |
| 2 | 14 | 14 | 10 | 8 |
| 3 | 16 | 14 | 6 | 4 |
| 4 | 14 | 18 | 6 | 6 |
| 5 | 14 | 16 | 10 | 4 |

**Table 2** Number of sensors and generate cycles for the long cycles case.

| CT | Generation Cycle (ms) | | | |
|---|---|---|---|---|
| | 24 | 56 | 72 | 108 |
| 1 | 6 | 12 | 18 | 30 |
| 2 | 8 | 16 | 16 | 38 |
| 3 | 6 | 16 | 21 | 26 |
| 4 | 6 | 14 | 18 | 30 |
| 5 | 5 | 16 | 20 | 27 |

**Table 3** Communication requirements for the in-vehicle UWB/PLC integrated network.

| Data type | Communication system | Load | Allowable latency | Allowable data loss rate | Allowable jitter |
|---|---|---|---|---|---|
| Non-low latency | UWB | 250 kbit/s | $\leq$ 25 ms | $\leq 10^{-3}$ | $\leq$ 10 ms |
| | PLC | 250 kbit/s | $\leq$ 25 ms | $\leq 10^{-3}$ | $\leq$ 10 ms |
| Low latency | PLC | 250 kbit/s | $\leq$ 15 ms | $\leq 10^{-4}$ | $\leq$ 5 ms |

generation cycles for the short and the long cycles cases respectively. The data transfer load is about 250 kbit/s in both cases.

Table 3 summarizes our performance target. The target data loss rate for UWB networks is $10^{-3}$. The time between when the data is created by the sensor and when it is received by the PT is considered latency. Regarding latency, communications are categorized into *low latency* for safety-critical data related to driving operation and *non-low latency*. Non-low latency corresponds to Class A or B of the Society of Automotive Engineers (SAE) classification. These classes comprise body-equipment data such as wiper operation, communication about windows, mirrors, air conditioning and meter displays [22]. Therefore, the maximum latency requirement is 25 ms or less and the jitter requirement is 10 ms or less. On the other hand, the target schedule computation time is 0.5 s. This corresponds to the approximate time it takes for the in-vehicle ECU to start up.

## 2.2 Preamble Overhead and Aggregation

Figure 2 shows the format of the polling and response frames. The polling frame from the PT to CTs includes the IDs of the sensors to be read and the data for the actuators. The response frame from CTs includes the readout data of the sensors specified in the polling frame. Standard IEEE 802.15.4 UWB frames can carry up to 127-byte Physical Service Data Unit (PDSU). As shown in Fig. 2, the MAC payload is 116 octets or less, and the sensor data in the response frame consists of 2 octets of ID and 4 octets of data; one frame can store the data from 19 sensors.

Figure 3 shows the ratio of PSDU transmission time to PLCP Protocol Data Unit (PPDU) frame transmission

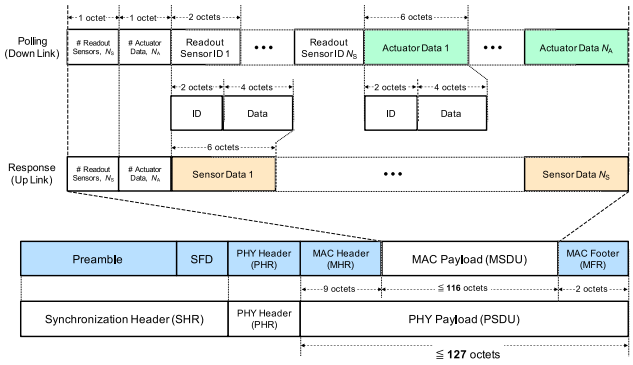



**Fig. 2** MAC SDU format/PPDU format.

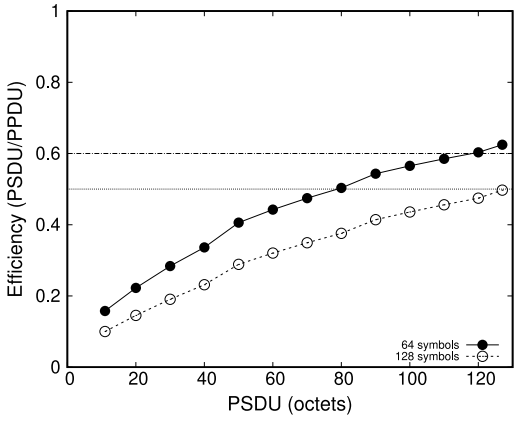



**Fig. 3** Ratio of PSDU transmission time to PPDU frame transmission time.

time for each PSDU size, estimated according to [23]. The communication efficiency in the case of 64 symbols is better than that of 128 symbols, but it remains at most 60%. In the case of 128 symbols, efficiency is further reduced to 50%, to the point that the preamble and payload transmission times are almost the same. Therefore, it is more efficient to use a 64-symbol preamble.

The UWB module we used, Qorvo DW3000, can extend the PSDU length up to 1023 octets using the extended format. This, however, makes the PHY Header (PHR) non-compliant with IEEE 802.15.4 because this format is an option in the IEEE 802.15.8 standard [24]. Therefore, when this long frame mode is chosen, the DW3000 cannot communicate with any device operating under the IEEE 802.15.4 standard frame encoding. Considering the implementation difficulty and the trade-off between frame length and PER, we choose to keep the IEEE 802.15.4 standard and to aggregate as much data as possible into the PSDU of up to 127 octets.

## 2.3 Polling Control and Retransmission

In IEEE 802.15.4, Clear Channel Assessment (CCA) is enabled by periodically inserting a preamble symbol into the data symbol [25]. This is, however, optional and is not implemented in UWB modules from major manufacturers such as NXP Semiconductors or Qorvo. The UWB module Qorvo

DW3000 we used is equipped with a simple pseudo-CCA, but it can detect only the preamble and not the PHR or PSDU. Therefore, we employ a polling control where the PT allows CTs to communicate one at a time. The DW3000 is compliant with 802.15.4 and already implements Reed-Solomon coding as Forward Error Correction (FEC) [24], [25].

When and which sensor data is read is determined through offline scheduling. The schedule is generated slot-by-slot, with a single CT communicating per slot. The generation cycle of each sensor is an integer multiple of the slot time. The slot length is determined by taking into account both the common divisor of the data generation cycle and the time required to adequately receive a response frame following the transmission of a polling frame.

In the short and long cycles, we introduced in Tables 1 and 2, the common divisors of the generation cycles are 1, 2, and 4 ms. Preliminary experiments showed that the time from sending a polling frame to receiving a single response frame is about 0.86 ms; until two response frames are received, it is about 1.35 ms; and until three response frames are received, it is about 2.1 ms. Considering these factors, we opted for a slot length of 4 ms in both of our scenarios. Furthermore, the number of sensor data specified to be read out in one polling frame should be approximately 38 data that can be transmitted in two response frames, leaving room for retransmission.

Polling is conducted sequentially from CT1 to CT5. Furthermore, as indicated in Table 3, the allowable latency is set at 25 ms, within which each CT should be polled at least once. Consequently, in our case, each CT must be polled at least every sixth slot, equating to 24 ms. Frequent polling, however, can lead to a waste of channel capacity. Therefore, we keep the unallocated slot after CT5 to reduce interference with other UWB systems and provide room for retransmission in case of communication failures [3]. This results in a total of 6 slots of 4 ms each that repeat every 24 ms as depicted in Fig. 4.

The polling control flowchart is shown in Fig. 5. The PT keeps the readouts to perform in a priority queue in which they are sorted according to their deadline. At the beginning of a slot, the planned readouts for that slot are added to the queue, prompting the PT to start polling the data from the corresponding CT in the earliest-deadline-first order. If the PT fails to receive a response for some sensor data, the readout request specifying the CT and sensor IDs is re-added to the priority queue, causing the other planned readouts to be performed later in the slot. As a result, failed transmissions are attempted again as soon as possible until they succeed, minimizing the latency even in case of failure. To ensure that latency constraints are met, data that would not satisfy latency requirements are discarded from the transmission queue. The empty slot serves as a buffer in case retransmissions cause the planned schedule to shift beyond the originally planned slots. Forward error correction is not further considered, given that the PSDU is encoded using Reed-Solomon code.

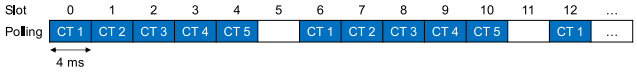



**Fig. 4** Pre-allocated slot to each chile device and unallocated slot.

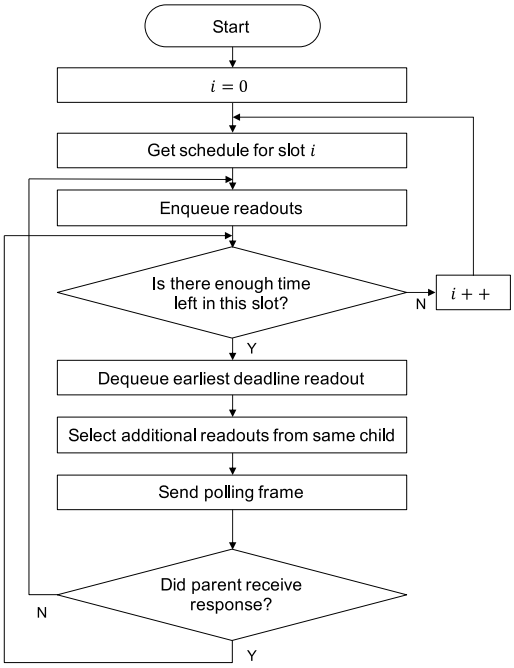



**Fig. 5** Polling control flowchart.

## 3. Optimization of Polling Schedule

This section explains the basic ideas for aggregating readout data into a frame and how to generate an optimal schedule using Integer Linear Programming (ILP).

### 3.1 Adjustment of Data Generation Phase

We explain the idea of phase control to reduce the number of frames transmitted. In this paper, each sensor is read out every corresponding cycle, and the polling order for each CT is also predetermined, as shown in Fig. 4. On the other hand, the generation cycle phase is still variable. Therefore, we adjust it to convey the readout data over fewer frames.

Let us explain the phase control using a simple example. Figures 6 and 7 depict scenarios without and with phase control, respectively. In this instance, we consider three types of periodic data that need to be read out before their respective deadlines. Sensor 1's data is expected to be read out every sixth slot, while Sensors 2 and 3 are read out every 12th slot. Polling is performed every 6th slot. For simplicity, we assume that two pieces of datum can be transmitted in a response frame. In the absence of phase adjustment and given these constraints, six frames are required, with the headers and trailers of frames transmitting only one datum, as shown in Fig. 6. In contrast, in Fig. 7, the generation cycle phase of Sensor 3 is shifted to avoid having to transmit the

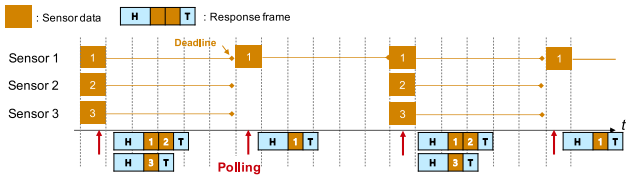



**Fig. 6** Polling without control of cyclic generation phase.

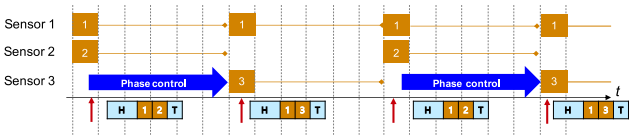



**Fig. 7** Polling with control of cyclic generation phase.

data from all 3 sensors in the same slot, reducing the required number of frames to four.

## 3.2 Scheduling Formulation

Based on this idea of phase control, we optimize when and which sensor data is read out using the following ILP formulation.

### 3.2.1 Variable Definitions

The variables used in our formulation are the following:

- $T$: schedule cycle (least common multiple of polling cycle and generation cycles),
- $\delta$: slot duration in milliseconds,
- $P$: polling cycle for each CT in milliseconds,
- $L$: allowable readout latency in milliseconds,
- $M$: the maximum number of data that can be read within a slot, where $M$ is a multiple of $N$,
- $N$: the maximum number of data that can be stored in a single frame,
- $k$: index of the allocated slots for an ST,
- $\mathcal{S}$: set of sensors,
- $\mathcal{J}_s$: set of readout data for sensor $s$,
- $\mathcal{J}$: set of readout data of all sensors,
- $c_s$: data generation cycle of sensor $s$,
- $g_{s,j}$: generation time of data $j$ from sensor $s$ before time shifting,
- $r_{s,j}$: allocated readout time in milliseconds for data $j$ of sensor $s$,
- $\phi_s$: integer variable, indicating the number of phase shift slots for the sensor $s$,
- $l_{s,j}$: latency from the generation time of data $j$ of sensor $s$ to the polling occasion for allocated readout time,
- $d_k$: the number of data transmitted in the $k$th allocated slot,
- $f_k$: the number of frames transmitted in the $k$th slot,

and

- $x_{s,j,k}$: binary variable, 1 if data $j$ of sensor $s$ is readout in the $k$th allocated slot, 0 otherwise.

It is assumed that time starts from 0 ms and slots are numbered in ascending order from zero. Note that the selection and timing of sensor data for readout repeat in cycles, with each cycle $T$ being the least common multiple of the sensor-generated cycles, as follows:

$$T = \mathrm{LCM}(\{P\} \cup \{c_s | s \in \mathcal{S}\}). \tag{1}$$

### 3.2.2 Objective Function and Constraints

We formulate a multi-objective function to reduce the total number of frames and avoid unnecessary polling latency. Recall that each CT is polled every $P/\delta$ slots, i.e., every six slot in the system model explained in Sect. 2. Therefore, the $k$th allocatable slot for CT $n$ is the $\{(P/\delta)k + (n-1)\}$st slot. Then, each CT is allocated $T/P$ polling slots during $T$ ms. For the latency of data occurring towards the end of the schedule cycle, $\{\lfloor (L-\delta)/P \rfloor + 1\}$ slots are added. The additional polling corresponds to the $\{\lfloor (L-\delta)/P \rfloor + 1\}$st allocatable slot from 0 in the early part of the cycle. Based on these, the range of $k$ is from 0 to $\{T/P + \lfloor (L-\delta)/P \rfloor + 1\}$. Taking this into consideration, we employ the following formulation, which is then solved for each CT to generate a schedule:

Minimize

$$\alpha \sum_{k=0}^{T/P-1} f_k + (1-\alpha) \sum_{s\in\mathcal{S}} \sum_{j\in\mathcal{J}_s} l_{s,j}, \tag{2}$$

subject to

$$l_{s,j} \le L - \delta, \quad \forall j \in \mathcal{J}_s, \forall s \in \mathcal{S}, \tag{3}$$

$$l_{s,j} \ge 0, \quad \forall j \in \mathcal{J}_s, \forall s \in \mathcal{S}, \tag{4}$$

$$l_{s,j} = r_{s,j} - (g_{s,j} + \phi_s \delta), \quad \forall j \in \mathcal{J}_s, \forall s \in \mathcal{S}, \tag{5}$$

$$r_{s,j} = \sum_{k=0}^{T/P+\lfloor (L-\delta)/P \rfloor} k\, P\, x_{s,j,k}, \quad \forall j \in \mathcal{J}_s, \forall s \in \mathcal{S}, \tag{6}$$

$$\phi_s \le c_s/\delta - 1, \quad \forall j \in \mathcal{J}_s, \forall s \in \mathcal{S}, \tag{7}$$

$$\phi_s \ge 0, \quad \forall j \in \mathcal{J}_s, \forall s \in \mathcal{S}, \tag{8}$$

$$\sum_{k=0}^{T/P+\lfloor (L-\delta)/P \rfloor} x_{s,j,k} = 1, \quad \forall j \in \mathcal{J}_s, \forall s \in \mathcal{S}, \tag{9}$$

$$d_k \le M, \quad 0 \le k \le T/P - 1, \tag{10}$$

$$d_k = \begin{cases} \sum_{s\in\mathcal{S}} \sum_{j\in\mathcal{J}_s} (x_{s,j,k} + x_{s,j,k+T/P}), \\ \qquad 0 \le k \le \lfloor (L-\delta)/P \rfloor, \\ \sum_{s\in\mathcal{S}} \sum_{j\in\mathcal{J}_s} x_{s,j,k}, \\ \qquad \lfloor (L-\delta)/P \rfloor < k \le T/P - 1, \end{cases} \tag{11}$$

$$f_k = \min\{1, d_k\} - \sum_{p=1}^{M/N} \min\{pN, d_k\} + \sum_{p=1}^{M/N} \min\{pN+1, d_k\}, \quad 0 \le k \le T/P - 1. \tag{12}$$

The weight parameter $\alpha$ is set to a significantly large value to primarily minimize the number of response frames while concurrently reducing readout latency. To achieve this, we choose the value of $\alpha$ to satisfy the following condition:

$$\frac{NL}{1+NL} \ll \alpha. \tag{13}$$

The derivation of the above is explained in Appendix.

The inequalities of Eqs. (3) and (4) are constraints on the time of data transmission allocation. The inequality of Eq. (3) guarantees the transmission of data within the allowable latency. Since the response frame will follow its polling frame in a slot, the latter should be sent $\delta$ before the allowable latency. The inequality of Eq. (4) forces the data transmission to occur after its creation, i.e., the latency should be non-negative. Equation (5) defines the latency, where $g_{s,j} + \phi_s\delta$ represents the generation time after phase shifting. Equation (6) converts the data readout time in the $k$th allocated slot to the time in slots. Equations (7) and (8) limit the range of phase shift to between 0 to the data generation cycle. Equation (9) ensures that all data are allocated for transmission only once. The inequality of Eq. (10) restricts the number of data transmitted in a slot to $M$ or less. Equation (11) represents the number of data transmitted in $k$th allocated slot. Data allocated for transmission after $T$ ms is concurrently transmitted within time slots 0 to $\lfloor L/P \rfloor - 1$. Equation (12) converts the number of readout data to the number of response frames to convey them.

Note that Eq. (12) is not a linear form. Nevertheless, it can be rewritten with Eq. (11) as a linear form of the column vector $(d_k, f_k)^\top$ by introducing the SOS2 (Special Ordered Set of Type 2) [26] auxiliary variables $t_1, \ldots, t_{2(M/N+1)}$ and binary variables $z_1, \ldots, z_{2(M/N)+1}$ as shown below:

$$\begin{pmatrix} d_k \\ f_k \end{pmatrix} = \sum_{p=0}^{M/N} t_{2p+1} \begin{pmatrix} pN \\ p \end{pmatrix} + \sum_{p=0}^{M/N} t_{2(p+1)} \begin{pmatrix} pN+1 \\ p+1 \end{pmatrix}, \tag{14}$$

$$t_i \geq 0, \quad i = 0, 1, \ldots, 2(M/N+1), \tag{15}$$

$$\sum_{i=1}^{2(M/N+1)} t_i = 1, \tag{16}$$

$$\sum_{i=1}^{2(M/N)+1} z_i = 1, \tag{17}$$

$$t_1 \leq z_1, \tag{18}$$

$$t_i \leq z_{i-1} + z_i, \quad i = 2, 3, \ldots, 2(M/N)+1, \tag{19}$$

$$t_{2(M/N+1)} \leq z_{2(M/N)+1}. \tag{20}$$

## 4. Heuristic Algorithm

The computational complexity of ILP grows exponentially with the number of inputs, i.e., sensors, posing challenges in finding a solution within a reasonable time. In our case, we want the low computing power of a motor vehicle's ECU to be capable of recomputing the schedule as the sensors polled and their cycles may evolve depending on the operation of the vehicle. Consequently, we introduce a heuristic algorithm designed to swiftly yield an acceptable solution to the schedule optimization problem.

**Algorithm 1** Heuristic algorithm

**Input:** Set of sensors $\mathcal{S}$
**Output:** Sensor data generation time and transmission schedule

Schedule cycle: $T$
Time slot length: $\delta$
Data generation cycle of sensor $s$: $c_s$
Polling cycle: $c_p$
Sort the set of sensors $\mathcal{S}$ in order of shortest generation cycle $c_s$
**for** $s \in \mathcal{S}$ **do**
  **for** $\phi = 0$ to $c_s/\delta - 1$ **do**
    **for** $j = 0$ to $T/c_s - 1$ **do**
      Calculate $j$th data generation time of sensor s: $g_{s,j,\phi}$
        $g_{s,j,\phi} = j\, c_c + \phi\, \delta$
      Calculate $j$th data readout time of sensor s: $r_{s,j,\phi}$
      **if** $g_{s,j,\phi} < T - c_p$ **then**
        $r_{s,j,\phi} = \lceil (g_{s,j,\phi} - \delta)/c_p \rceil\, c_p$
      **else**
        $r_{s,j,\phi} = \lceil (g_{s,j,\phi} - \delta - T)/c_p \rceil\, c_p$
      **end if**
    **end for**
    Calculate the total number of frames, the maximum number of frames per slot, the maximum room in payload among the response frames, and total latency
  **end for**
  Select $\phi$ and determine the phase of $s$ (priority is given to above criteria)
  Add readout time $r_{s,j,\phi}$ into schedule
**end for**

Our heuristic scheduling algorithm is shown in Algorithm 1. The algorithm performs the phase optimization for each CT independently, taking the set of sensors and their corresponding generation cycles as input. The algorithm gradually builds the schedule by adding the sensors to be read out one by one, starting with those exhibiting the shortest generation cycle.

The phase adjustment consists of selecting the optimal slot for each sensor based on the following evaluation criteria from most (top) to least important (bottom):

1. Minimizing the total number of response frames,
2. Minimizing the maximum number of response frames per slot,
3. Minimizing the maximum room in payload among the response frames,
4. Minimizing the total latency.

If a tie among multiple phase candidates occurs, one is selected at random.

In the above, Criterion 1 aims to suppress the total number of response frames, corresponding to the first term in Eq. (2). Criterion 2 intends to leave residual time in a slot for retransmissions. Criterion 3 aims to evenly distribute the room of payload, which we call "payload margin," among response frames, allowing flexibility for sensors whose phases have not yet been adjusted. Finally, Criterion 4 focuses on minimizing latency, corresponding to the second term in Eq. (2).

### 4.1 Applying the Heuristic Algorithm to Simple Examples

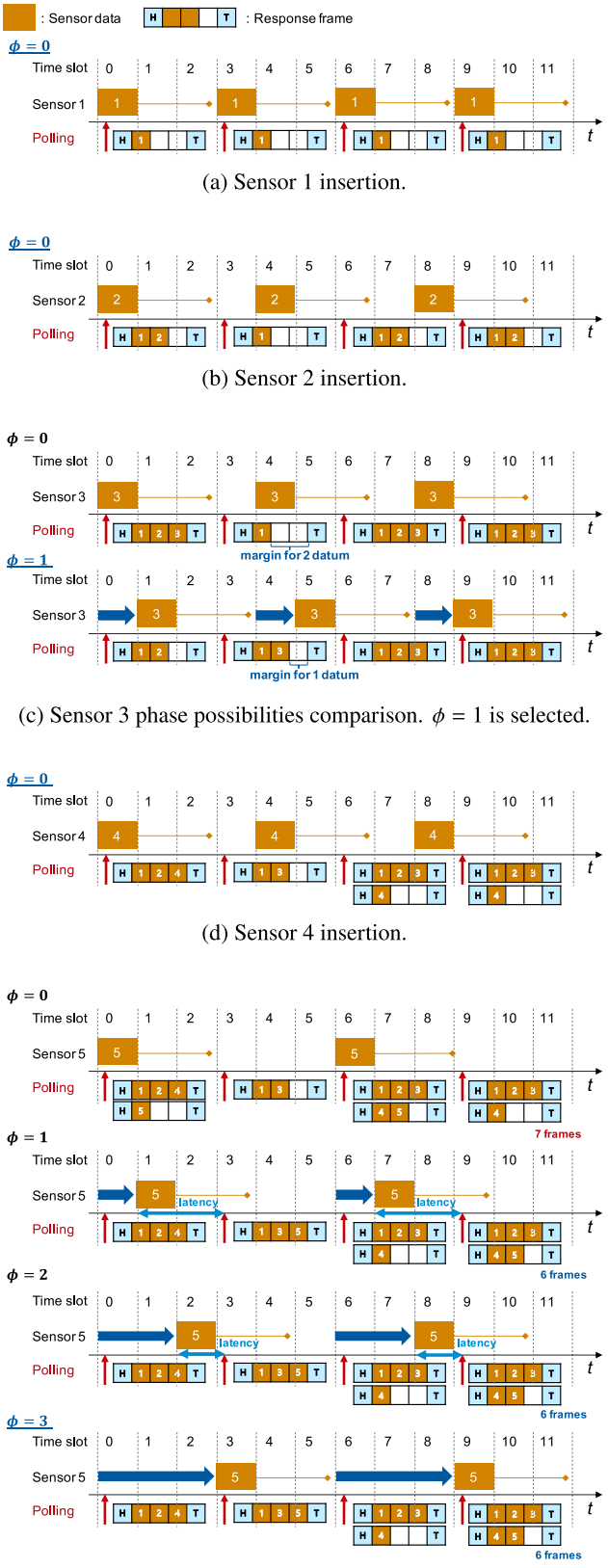
We illustrate the heuristic algorithm using a simple example presented in Fig. 8. In this example, we consider a CT equipped with five sensors: Sensor 1 through 5, each having generation cycles of $c_1 = 12$ ms, $c_2 = c_3 = c_4 = 16$ ms, and $c_5 = 24$ ms, respectively. The CT is polled every three slots, which amounts to every 12 ms. Unlike the example from Sect. 3.1, three pieces of data can be conveyed in a response frame.

The algorithm starts by determining the phase for Sensor 1, which has the shortest cycle. Since the generation cycle of Sensor 1 is equal to the polling cycle, shifting the generation phase only increases the latency, so it is scheduled as shown in Fig. 8(a). Next, Sensor 2 is equivalent in the above criteria, even if its phase is shifted over 16 ms, i.e., four slots. In such a case, its phase is randomly chosen; Figure 8(b) shows that a shift of 0 was picked. In the scheduling of Sensor 3, there is a difference in terms of Criterion 3 over its cycle, i.e., four slots. As shown in Fig. 8(c), the maximum payload margin is two, while it is one when the offset of the phase is one. Therefore, the latter is chosen. Sensor 4 is scheduled as shown in Fig. 8(d) for the same reason as Sensor 2. The last sensor has four potential phase offset candidates, as illustrated in Fig. 8(e). In the scenario without any offset from the 0th slot, the total number of response frames is seven, whereas it is six for the other offset options. The former case is eliminated based on Criterion 1. Ultimately, the choice is determined by Criterion 4 in favor of the offset by three slots, as it incurs no latency compared to the one- and two-slot offsets.

For reference, the schedule obtained without the heuristic algorithm is depicted in Fig. 9. With the heuristic algorithm, one less frame is necessary to transmit the data, reducing the total number of frames from seven down to six. Overall, the heuristic algorithm is effective in reducing overhead.

### 4.2 Comparison of ILP and Heuristic Scheduling Results

We compare the schedules generated by the heuristic algorithm and those by the ILP method discussed in Sect. 3. With the condition of the weight $\alpha$ of the objective function (2), the left-hand side of Eq. (13) is 0.997. Therefore, we set $\alpha = 0.999$. The number, $M$, of sensor data specified to be read out in one polling frame is 38, corresponding to two frames. The evaluation metrics include the total number of response frames and the average latency. Note that Latency is defined as the time from when the data is generated until the readout is allocated and polling begins. All calculations were performed on a MacBook Air equipped with an Apple M1 chip and 16 GB of memory, with user CPU time measured using the '`time`' command. The ILP was implemented in Python, using the Gurobi Optimizer as the optimization solver. In contrast, the heuristic algorithm, which does not rely on an optimization solver, was implemented in C to




(a) Sensor 1 insertion.

(b) Sensor 2 insertion.

(c) Sensor 3 phase possibilities comparison. $\phi = 1$ is selected.

(d) Sensor 4 insertion.

(e) Sensor 5 phase possibilities comparison. $\phi = 3$ is selected.

**Fig. 8** Heuristic algorithm execution example.

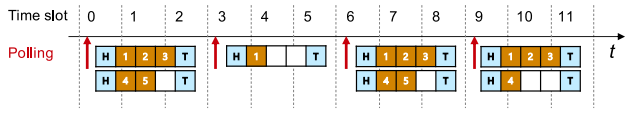



**Fig. 9** Schedule without phase adjustment by the heuristic algorithm.

**Table 4** Schedule characteristics for the short cycles case.

| Scheduling | Phase adjustment | Number of response frames | Average latency (ms) | Latency std deviation (ms) |
|---|---|---|---|---|
| RR w/ empty | ILP | 466 | 2.28 | 4.99 |
| RR w/ empty | Heuristic | 538 | 2.42 | 5.20 |
| RR w/ empty | w/o | 543 | 2.29 | 4.99 |
| RR w/o empty | w/o | 554 | 7.87 | 5.73 |
| PO w/o empty | w/o | 546 | 2.93 | 5.14 |

**Table 5** Schedule characteristics for the long cycles case.

| Scheduling | Phase adjustment | Number of response frames | Average latency (ms) | Latency std deviation (ms) |
|---|---|---|---|---|
| RR w/ empty | ILP | 439 | 3.53 | 5.71 |
| RR w/ empty | Heuristic | 525 | 3.53 | 5.71 |
| RR w/ empty | w/o | 598 | 3.53 | 5.71 |
| RR w/o empty | w/o | 605 | 7.80 | 5.77 |
| PO w/o empty | w/o | 606 | 7.04 | 7.50 |

maximize processing speed.

For comparison, we include three polling methods without phase adjustment: "RR without empty slots," which is the simplest polling without empty slots, and "RR with empty slots," which is described in Sect. 2.3. Note that both ILP and heuristic also keep an empty slot. "Polling-order optimized (PO)" approach optimizes the sequence of polling without incorporating empty slots, utilizing the ILP model described in [2], [3]. This method is designed to minimize polling frequency as much as possible. The two scenarios were used using short and long generation cycles previously introduced in Tables 1 and 2.

The results of the calculations are summarized in Tables 4 and 5. These tables show that the ILP and the heuristic algorithm successfully reduce the number of response frames, indicating effective aggregation of readout data. It is, however, important to note that this reduction in response frames comes at a cost. The average latency for the ILP and the heuristic algorithm is higher than that for no phase adjustment in the RR with empty slots. This is primarily due to readouts being postponed (within latency constraints) for aggregation. The reason for the extremely high latency of the round-robin without empty slots is that the polling cycle is 20 ms, while most sensors have a generation cycle of 24 ms. This forces polling timing to shift gradually and cyclically from the expected generation timing, resulting in a large latency. In contrast, this does not happen thanks to the matching polling cycle of 24 ms in the round-robin schedules with empty slots. In scenarios involving a long cycle, there is a significant disparity in the schedules when comparing cases with and without phase adjustment. Although the data rates for both patterns remain identical, these differences arise due to the increased flexibility in sensor generation timing afforded by the long cycle. Consequently, in "Polling-order optimized," which lacks phase adjustment, there is an observed increase in both the number of frames and latency.

**Table 6** Schedule computation time.

| Method | Short cycles | Long cycles |
|---|---|---|
| ILP | 987 s | 1676 s |
| Heuristic | 70 ms | 190 ms |

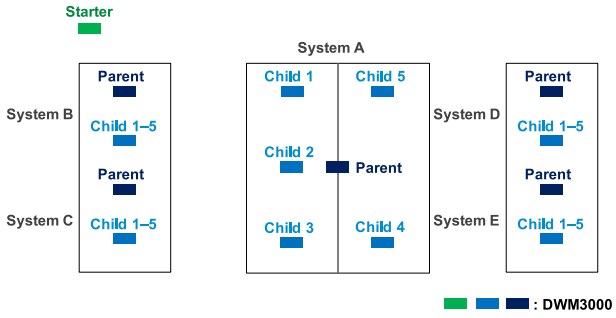



**Fig. 10** Physical layout of the UWB modules inside of the EMC tent.

Next, as depicted in Fig. 6, the heuristic algorithm achieves a remarkable reduction in computation time, approximately 99.9% less than the ILP, for both scenarios. The heuristic is below the target value of 0.5 s in both scenarios, while the ILP requires about 15 minutes or more. This suggests that the heuristic algorithm may enable on-board ECUs to implement phase adjustment in response to changing situations.

## 5. Experimental Validation

To verify the effectiveness of our various methods, we conducted an experiment in which we ran the obtained schedules and measured the data loss ratio and the latency. The experiment was conducted in an Electromagnetic Compatibility (EMC) tent manufactured by Medical Aid Co. LTD, and Qorvo DWM3000 UWB modules were used as the UWB equipment. In this experiment, a set of UWB modules is used to simulate the wireless harness (System A), and four sets of interfering UWB modules (Systems B through E) are placed, as shown in Fig. 10. System A uses six modules to emulate the in-vehicle network, while for Systems B through E, one module serves as the PT, and one module serves the role of the five CTs. The configuration used was channel 9 (center frequency 7987.2 MHz, bandwidth 499.2 MHz), data rate of 6.81 Mbit/s, and a preamble length of 64 symbols (as opposed to the default 128 symbols). This change in the preamble length was made to increase the time available for retransmission in case of reception failure, as explained in Sect. 2.

### 5.1 Experimental Methods

Figure 11 shows our experimental method. Five types of transmission schedules for each system were employed, as shown in Tables 4 and 5. The duration of a single schedule is 1512 ms, corresponding to the schedule cycle $T$, and it repeats three times in an experimental trial. To start the

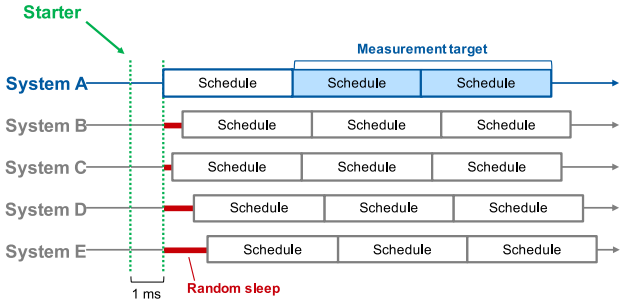



**Fig. 11** Synchronization method between Systems and measurement target.

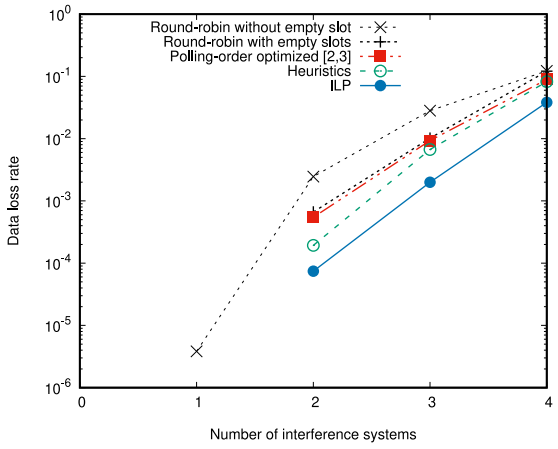
experiment, the starter UWB module sends a synchronization signal to the PT of Systems A through E. As soon as the PT in System A receives the synchronization signal, it polls the CT according to its schedule. On the other hand, Systems B through E sleep for a random period of time upon receiving the start signal and then communicate with their CT according to the same schedule. The random sleep time is selected in the range 0 through 1511.999 ms using the Mersenne-Twister method. We measured the communication performance of System A's second and third schedules, focusing on aspects such as data loss rate and latency while it is receiving interference from the other systems. To increase the reliability of the experiment, we repeated the experimental trial 100 times. As a result, we obtained data loss rate, average latency, and latency standard deviation.

## 5.2 Experimental Results

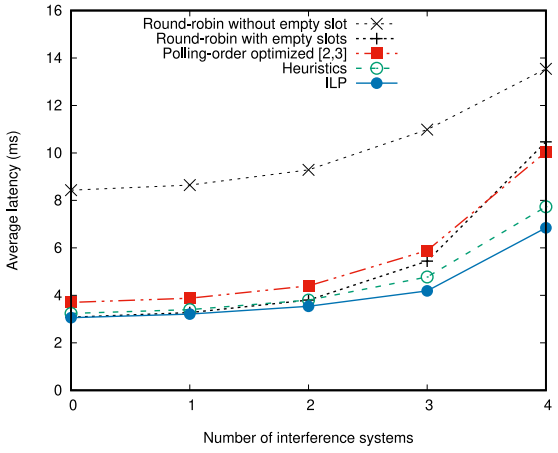
The data loss rate, average latency, and latency standard deviation for the short and long data cycles are shown in Figs. 12 and 13, respectively. Note that Figs. 12(a) and 13(a) use a logarithmic scale; cases with zero loss rate are therefore not plotted.

As a general trend, the data loss rate is lower when the number of frames is smaller, indicating that aggregation is effective. Since "Round-robin without empty slots" has a larger data loss rate even when there is only one interfering system, we conclude that it is worthwhile to prepare empty slots for retransmission in case of data loss. The schedules obtained through the "ILP" and the "heuristic" algorithm achieve the target data loss rate, $\leq 10^{-3}$, in the presence of up to two interfering systems in both patterns. It can be inferred that the degradation of radio link quality was not a factor in our experiments, as they were conducted within the EMC tent, thus avoiding external interference. In scenarios devoid of interfering systems, data loss was not observed. However, in the presence of one or more interfering systems, data loss occurred due to collisions between the communications of different UWB systems. "PO" demonstrated lower values than "RR without empty slot," yet it showed values nearly identical to those of "RR with empty slots." Moreover, it was confirmed that both the "ILP" and "heuristic" methods yield better outcomes than the existing method for both patterns.

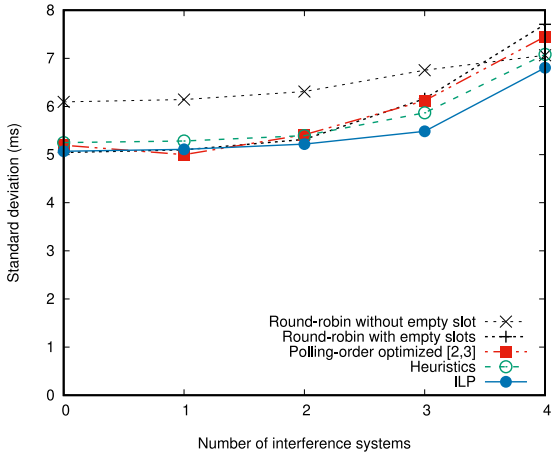
For average latency and its standard deviation, both the




(a) Data loss rate depending on the number of interference systems. Loss rates of 0 are not plotted.



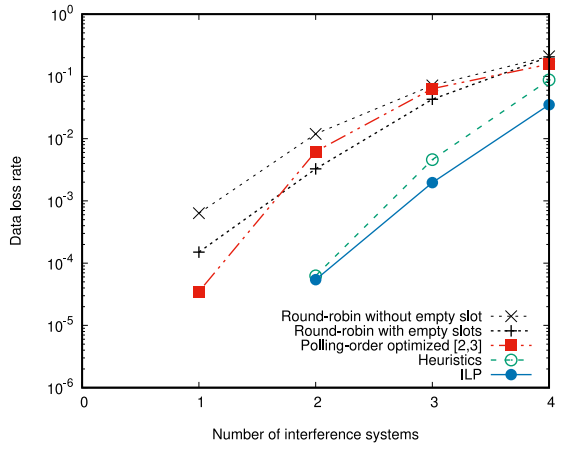

(b) Average latency depending on the number of interference systems.



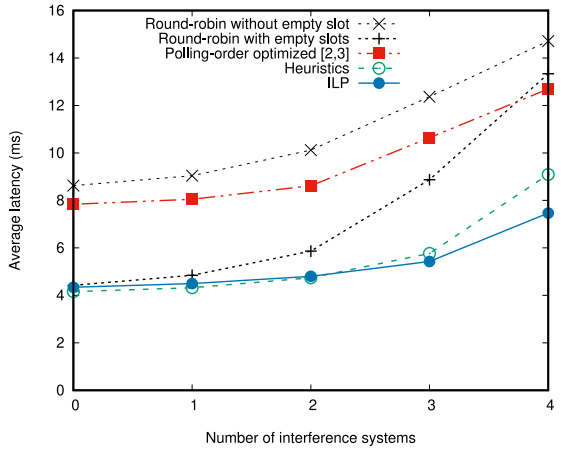

(c) Standard deviation of latency depending on the number of interference systems.

**Fig. 12** Results of the short cycle experiments.

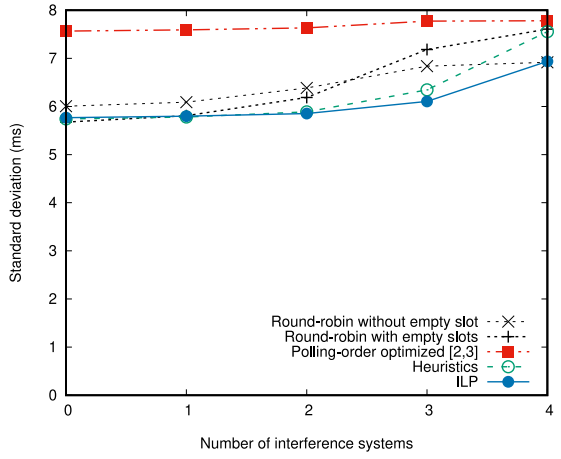
ILP and heuristic algorithms exhibit values that are comparable to those without interfering systems. However, in the case of the round-robin without empty slots, data loss




(a) Data loss rate depending on the number of interference systems. Loss rates of 0 are not plotted.




(b) Average latency depending on the number of interference systems.




(c) Standard deviation of latency depending on the number of interference systems.

**Fig. 13** Results of the long cycle experiments.

increases as interference grows, leading to more frequent retransmissions and higher latency. Conversely, the ILP and heuristic algorithms show a smaller increase in latency compared to the other methods. Moreover, schedules in the existing methods tended to tolerate a higher latency if it could reduce the number of polling, resulting in more significant latency compared to ILP and heuristics. In particular, the long cycle exhibited a significant concentration of data generation at certain points, resulting in increased latency. Data pollings that would fail to meet their latency requirements are systematically discarded from the polling queue and are therefore treated as transmission failures. Therefore, the observed maximum latency remains within the specified requirements. The standard deviations of the latency representing the jitter were all below the target value of 10 ms.

As a result, the ILP produces the most favorable schedule regarding data loss rate and latency. While not as good as the schedule obtained through the ILP, the heuristic algorithm still provides superior scheduling compared to standard polling methods without any adjustments.

We estimated the average power consumption of the UWB modules in the experiment using the DW3000 Datasheet [27]. Power consumption is mostly influenced by the number of frames transmitted. With the "PO" scheduler, where the most frames were transmitted, the PT consumes 41.07 mW and each CT consume 10.61 mW. These are of the same order of magnitude as the average power consumption of the CAN-FD (CAN Flexible Data Rate) transceiver "TCAN1042" which consumes 58.75 mW [28]. The power consumption of the UWB devices is not a concern.

## 6. Conclusion

This paper introduced a polling control method and scheduling algorithm designed for in-vehicle UWB networks. We developed an ILP formulation to generate schedules with minimum response frames. Also, we proposed a heuristic algorithm capable of producing schedules within a few hundred milliseconds. By comparing the quality of schedules obtained from both methods, we demonstrated that our proposed heuristic approach yields nearly optimal solutions. Additionally, we conducted experimental validation using a UWB module to assess the robustness of our schedules against interference. The ILP optimization and the heuristic schedules showed low data loss rates and minimized readout latency.

In this paper, we have primarily addressed scheduling for sensor readouts. Future directions include developing scheduling schemes incorporating data transmission to actuators and improving system reliability, especially in scenarios with three or more interference systems.

## Acknowledgments

We thank Mr. T. Tanaka, Mr. S. Yamaguchi, and Mr. M. Takenaka from Kobe University, Mr. N. Kurioka from DENSO TEN Limited, and Dr. S. Shimizu from ATR for their cooperation and valuable advice. This work was supported by MIC SCOPE Grant Number JP215007006.

**References**

[1] S.C. Ergen and A. Sangiovanni-Vincentelli, "Intravehicular energy-harvesting wireless networks: Reducing costs and emissions," IEEE Veh. Technol. Mag., vol.12, no.4, pp.77–85, Dec. 2017.
[2] C. Ohta, T. Tanaka, H. Migita, S. Yamaguchi, M. Takenaka, P. Finnerty, and T. Kamada, "A study on optimization of polling scheduling for in-vehicle UWB wireless networks," IEICE Commun. Express, vol.11, no.7, pp.429–434, July 2022.
[3] H. Migita, T. Tanaka, S. Yamaguchi, M. Takenaka, P. Finnerty, T. Kamada, and C. Ohta "Optimization of polling-based MAC schedule considering data aggregation for in-vehicle UWB wireless networks," Proc. 2022 IEEE 8th World Forum on Internet of Things (WF-IoT), 6 pages, Oct.–Nov. 2022.
[4] H. Migita, M. Takenaka, P. Finnerty, M. Okuhara, and C. Ohta, "A study on optimization of polling schedule to minimize the number of frames for in-vehicle UWB wireless network," IEICE Commun. Express, vol.12, no.6, pp.333–338, June 2023.
[5] M. Koca, G. Gurbilek, B. Soner, and S. Coleri, "Empirical feasibility analysis for energy harvesting intravehicular wireless sensor networks," IEEE Internet Things J., vol.8, no.1, pp.179–186, Jan. 2021.
[6] U. Demir, C.U. Bas, and S. Coleri Ergen, "Engine compartment UWB channel model for intravehicular wireless sensor networks," IEEE Trans. Veh, Technol., vol.63, no.6, pp.2497–2505, July 2014.
[7] C.U. Bas and S.C. Ergen, "Ultra-wideband channel model for intravehicular wireless sensor networks beneath the chassis: from statistical model to simulations," IEEE Trans. Veh. Technol., vol.62, no.1, pp.14–25, Jan. 2013.
[8] A. Chandra, A. Prokeš, T. Mikulášek, J. Blumenstein, P. Kukolev, T. Zemen, and C.F. Mecklenbräuker, "Frequency-domain in-Vehicle UWB channel modeling," IEEE Trans. Veh. Technol., vol.65, no.6, pp.3929–3940, June 2016.
[9] J. Li and T. Talty, "Channel characterization for ultra-wideband intravehicle sensor networks," Proc. IEEE Mil. Commun. Conf. (MILCOM), pp.1–5, 2006.
[10] T. Tsuboi, J. Yamada, N. Yamauchi, M. Nakagawa, and T. Maruyama, "UWB radio propagation inside vehicle environments," Proc. 2007 7th International Conference on ITS Telecommunications, pp.1–5, 2007.
[11] P.C. Richardson, W. Xiang, and W. Stark, "Modeling of ultra-wideband channels within vehicles," IEEE J. Sel. Areas Commun., vol.24, no.4, pp.906–912, April 2006.
[12] NXP Semiconductors, "S32G3 Vehicle Networking Reference Design," accessed on 5 Dec., 2023. https://www.nxp.com/design/designs/s32g3-vehicle-networking-reference-design:S32G-VNP-RDB3
[13] L.A. Perişoară, D.I. Săcăleanu, and C. Dănişor, "Automotive ethernet architecture for the connected dacia logan electric vehicle," Proc. 2022 14th International Conference on Communications (COMM), pp.1–6, 2022.
[14] Y. Sadi and S.C. Ergen, "Optimal power control, rate adaptation, and scheduling for UWB-based intravehicular wireless sensor networks," IEEE Trans. Veh. Technol., vol.62, no.1, pp.219–234, Sept. 2013.
[15] Y. Sadi and S.C. Ergen, "Energy and delay constrained maximum adaptive schedule for wireless networked control systems," IEEE Trans. Wireless Commun., vol.14, no.7, pp.3738–3751, July 2015.
[16] G. Karadag, M.S. Iqbal, and S. Coleri, "Optimal power control, scheduling, and energy harvesting for wireless networked control systems," IEEE Trans. Commun., vol.69, no.3, pp.1789–1801, March 2021.
[17] B. Farayev and S.C. Ergen, "Towards ultra-reliable M2M communication: Scheduling policies in fading channels," Proc. 2016 23rd International Conference on Telecommunications (ICT), pp.1–6, May 2016.
[18] M.S. Akbar, H. Yu, and H. Cang, "IEEE 802.15.4 frame aggregation enhancement to provide high performance in life-critical patient monitoring systems," Sensors, vol.17, no.2, 241, Jan. 2017.
[19] L. Zhang, Y. Gu, R. Wang, K. Yu, Z. Pang, Y. Li, and B. Vucetic "Enabling real-time quality-of-service and fine-grained aggregation for wireless TSN," Sensors, vol.22, no.10, 3901, May 2022.
[20] C.L. Liu and J.W. Layland, "Scheduling algorithms for multiprogramming in a hard-real-time environment," J. ACM (JACM), vol.20, no.1, pp.46–61, Jan. 1973.
[21] J. Nagle, "On packet switches with infinite storage," IEEE Trans. Commun., vol.COM-35, no.4, pp.435–438, April 1987.
[22] C. Lupini, "Vehicle multiplex communication," Vehicle Multiplex Communication: Serial Data Networking Applied to Vehicular Engineering, pp.i–viii, SAE, 2004.
[23] M. Charlier, B. Quoitin, S. Bette, and J. Eliasson, "Support for IEEE 802.15.4 ultra wideband communications in the Contiki operating system," Proc. 2016 Symposium on Communications and Vehicular Technologies, 6 pages, Nov. 2016.
[24] Qorvo Inc., "DW3000 Family User Manual," 2019. https://www.qorvo.com/products/d/da008154
[25] "IEEE Standard for Low-Rate Wireless Networks," IEEE Std 802.15.4-2020 (Revision of IEEE Std 802.15.4-2015), pp.1–800, July 2020.
[26] E.M.L. Beale and J.J.H. Forrest, "Global optimization using special ordered sets," Mathematical Programming, vol.10, no.1, pp.52–69, Dec. 1976.
[27] Qorvo Inc., "DW3000 Datasheet," 2020. https://www.qorvo.com/products/d/da008142
[28] Texas Instruments, "TCAN1042-Q1Automotive Fault Protected CAN Transceiver with CAN FD datasheet (Rev. D)," 2021. https://www.ti.com/lit/gpn/tcan1042-q1

## Appendix: Derivation of the Inequality for the Weight Parameter $\alpha$

Let us consider the value of $\alpha$ that mainly minimizes the total response frames and also reduces the total latency when minimizing the objective function of Eq. (2). For this sake, the first term should be much larger than the second term of the objective function, which means

$$(1-\alpha)\sum_{s\in\mathcal{S}}\sum_{j\in\mathcal{J}_s} l_{s,j} \ll \alpha \sum_{k=0}^{T/C} f_k. \tag{A·1}$$

First, let us consider an ideal situation such that all readout data during $T$ are conveyed via the least number, $F_{\min}$, of the response frames. The value of $F_{\min}$ can be obtained by

$$F_{\min} = \frac{1}{N}\sum_{s\in\mathcal{S}}\frac{C}{c_s}, \tag{A·2}$$

where recall that $N$ is the maximum number of readout data stored in a response frame, and $C/c_s$ is the number of readout data transmitted during the schedule cycle $T$. The value of $F_{\min}$ can be regarded as the lower bound of the total number of response frames, $\sum_{k=0}^{T/C} f_k$, which appears in the first term of the objective function. Thus, we have

$$F_{\min} \le \sum_{k=0}^{T/C} f_k. \tag{A·3}$$

Second, let us consider the worst case where all data are read out with the maximum latency $L$. In this case, the total maximum latency $L_{\max}$ is obtained by

$$L_{\max} = L \sum_{s \in \mathcal{S}} \frac{C}{c_s}. \quad (\mathrm{A}\cdot 4)$$

This can be regarded as the upper bound of the total latency appearing in the second term of the objective function. Thus, we have

$$\sum_{s \in \mathcal{S}} \sum_{j \in \mathcal{J}_s} l_{s,j} \leq L_{\max}. \quad (\mathrm{A}\cdot 5)$$

From Eqs. (A· 3) and (A· 5), a sufficient condition satisfying the inequality of Eq. (A· 1) is given as

$$(1-\alpha) L_{\max} \ll \alpha F_{\min}. \quad (\mathrm{A}\cdot 6)$$

Finally, by substituting Eqs. (A· 2) and (A· 4) to Eq. (A· 6) and solving it for $\alpha$, we have the inequality of Eq. (13).

**Hajime Migita** was born in Hyogo, Japan, in 2000. He received the B.E. degree in engineering in 2022, from Kobe University, Hyogo, Japan. He is currently enrolled in the master's program at the Graduate School of Science, Technology and Innovation, Kobe University. He is mainly engaged in research on in-vehicle communications. He is a student member of IEICE.

**Yuki Nakagoshi** is an undergraduate student of the Department of Computer Science and System Engineering at Kobe University. He is mainly engaged in research on in-vehicle communications.

**Patrick Finnerty** was born in Cholet, France, in 1995. He received his French Engineering degree in computer science and information technology from INSA Lyon in 2018 and his Ph.D. (Eng.) from Kobe University, Japan, in 2022. Since February 2022, he has been working as an assistant professor at the Graduate School of System Informatics, Kobe University. His research interests include parallel and distributed computing techniques and distributed systems. He is a member of IPSJ, IEEE Computer Society, and ACM.

**Chikara Ohta** was born in Osaka, Japan 1967. He received the B.E., M.E., and Ph.D. (Eng.) degrees in communication engineering from Osaka University, Osaka, Japan, in 1990, 1992, and 1995, respectively. From April 1995, he was an assistant professor with the Department of Computer Science, Faculty of Engineering, Gunma University. From October 1996, he was a lecturer at the Department of Information Science and Intelligent Systems, Faculty of Engineering, University of Tokushima, and an associate professor there from March 2001. From November 2002, he was an associate professor of the Department of Computer and Systems Engineering, Faculty of Engineering, Kobe University. From April 2010, he was an associate professor at the Graduate School of System Informatics, Kobe University, and a professor there from January 2015. From April 2016, he was a professor at the Graduate School of Science, Technology and Innovation, Kobe University. Since April 2022, he has been a professor at the Graduate School of System Informatics, Kobe University. From March 2003 to February 2004, he was a visiting scholar at the University of Massachusetts at Amherst, USA. His current research interests include the performance evaluation of communication networks. He is a member of IPSJ, IEEE, and ACM SIGCOMM.

**Makoto Okuhara** was born in Shimane, Japan, in 1977. He received the B.E. in intelligent information engineering from Tottori University, Japan, in 2000. In 2006, he Joined DENSO TEN Co., Ltd, Japan. He serves as Chief Inspector of the Mobility Solution Development Department of the Automotive Electronics Business Headquarters and Project Leader of the Innovation Development Center. He is mainly engaged in in-vehicle ECU development and wireless communication technology research. He is enrolled in the doctor's program at the Graduate School of Science, Technology and Innovation, Kobe University.

PAPER

# Waveguide Slot Array with Code-Division Multiplexing Function for Single RF Chain Digital Beamforming

**Narihiro NAKAMOTO**[†a)], ***Member***, **Kazunari KIHIRA**[†], **Toru FUKASAWA**[†], ***Senior Members***, **Yoshio INASAWA**[†], ***Member***, ***and*** **Naoki SHINOHARA**[††], ***Senior Member***

**SUMMARY** This study presents a novel waveguide slot array with a code-division multiplexing function for single RF chain digital beamforming. The proposed antenna is comprised of a rectangular metallic waveguide's bottom part and a multilayer printed circuit board (PCB) with the rectangular waveguide's top wall and slot apertures. Multiple pairs of two symmetric longitudinal slots are etched on the metal surface of the PCB, and a PIN diode is mounted across each slot. The received signals of each slot pair are multiplexed in a code-division multiplexing fashion by switching the diodes' bias according to the Walsh Hadamard code, and the original signals are then recovered through a despreading process in the digital domain for digital beamforming. A prototype antenna with eight slot pairs has been fabricated and tested for proof of concept. The measured results show the feasibility of the proposed antenna.

***key words:*** *digital beamforming, waveguide slot array, single RF chain*

## 1. Introduction

Digital beamforming (DBF) array antennas have been actively developed in recent years to meet the demands of modern wireless communication and radar systems [1]–[3]. In the conventional element-level DBF array antenna, each antenna element is connected to a separate radio frequency (RF) chain (amplifier, down-converter, and filter) and analog-to-digital converter (ADC) to obtain an individual complex received signal from each antenna element. Therefore, high hardware costs and increased power consumption make it difficult to realize a large-scale element-level DBF array antenna.

Many works have been reported to realize an element-level DBF array antenna that reduces the number of RF chains and ADCs [4]–[16]. To share a single RF chain among multiple antennas, received signals from multiple antennas are multiplexed into a single RF chain and ADC. Subsequently, each signal is recovered in the digital domain. In analogy with multiple access methods in wireless communication systems, three major multiplexing techniques, such as time-division multiplexing (TDM), frequency-division multiplexing (FDM), and code-division multiplexing (CDM), have been applied to realize a DBF array antenna with a single RF chain. In [4]–[7], received signals from each antenna element are transferred to a single RF chain sequentially using switches in a TDM fashion. In [8] and [9], received signals from each antenna element are converted to signals with a different intermediate frequency (IF) from each other and then multiplexed in an FDM fashion. However, TDM and FDM schemes have a major drawback: when an adjacent channel interferer (ACI) is present, recovered signals in TDM and FDM schemes suffer from corruption by the ACI [10], [11].

CDM is the most promising method because, unlike TDM and FDM, the original signals can be recovered even in the presence of ACI [11]. In the CDM scheme, received signals from each antenna element are multiplied by spreading codes, and the spread signals are transferred to a single ADC. The original signals are recovered by a despreading process in the digital domain. In [11], received signals are code-modulated in the RF domain with a code-modulating low noise amplifier (LNA). In [12] and [13], code-modulation is performed in the IF domain with multipliers after amplification and frequency conversion of received signals. In these systems, each antenna element still requires a separate amplifier, which increases cost and power consumption in a large-scale array.

In [14], received signals are code-modulated with reflection-type 0°/180° phase shifters, and then all spread signals are combined by a microstrip power combiner and fed into a single RF chain. In [15], [16], a metasurface with a CDM function is presented. The responses of each unit cell are code-modulated by switching their integrated diodes, and then spread signals are spatially combined and fed into a feeding antenna with a single RF chain. In these antennas, signal amplification is implemented after combining the spread signals, and thus they can reduce the number of amplifiers.

In this study, we present a novel waveguide slot array antenna with a CDM function. The proposed antenna uses slot pairs with RF switches (e.g., PIN diodes) on the broad wall of a rectangular waveguide. Waveguide slot arrays with slots having integrated PIN diodes are presented in several studies [7], [17]–[19]. In [17]–[19], slots with appropriate radiation phases are activated, and the others are inactivated by switching the diodes' bias to achieve desired radiation patterns in the RF domain. In [7], only one slot is activated at a time, and an activated slot is sequentially switched by switching the diodes' bias. The original signal of each slot is then recovered in the digital domain in a TDM fashion.

Manuscript received July 27, 2023.
Manuscript revised October 19, 2023.
Manuscript publicized February 16, 2024.

[†]Mitsubishi Electric Corporatin, Kamakura-shi, 247-8501 Japan.
[††]Research Institute for Sustainable Humanosphere, Kyoto University, Uji-shi, 611-0011 Japan.
a) E-mail: Nakamoto.Narihiro@ap.MitsubishiElectric.co.jp

DOI: 10.23919/transcom.2023EBP3123

However, in the proposed antenna, the received signal of each slot-pair can be code-modulated by switching the RF switches' bias. Subsequently, the spread signals are combined in the hollow rectangular waveguide, and fed into a single RF chain. Following digitization, the despreading process recovers the original complex signals received by each slot pair, and the DBF is performed to achieve scanned beams, multiple beams, adaptive beams, or direction of arrival (DOA) estimation. To the best of our knowledge, this is the first waveguide slot array with a CDM function. The proposed antenna reduces the number of RF chains, amplifiers, and ADCs and has the potential to offer higher efficiency than previous works using microstrip lines or transmittarray/reflectarray architecture [14], [15].

The rest of this paper is organized as follows: Section 2 explains the detailed configuration and operating principle of the proposed antenna; Sect. 3 describes the antenna design; Sect. 4 presents the experimental results of the prototype antenna; and Sect. 5 presents the conclusion.

## 2. Antenna Configuration and Operating Principle

Figure 1 shows the configuration of the proposed antenna with four slot pairs. The antenna comprises the bottom part of a metallic waveguide, a multilayer printed circuit board (PCB), RF switches, a single RF chain, an ADC, and a digital signal processing (DSP) unit. Both surfaces of the PCB are covered with copper foil, and slot apertures are formed by etching. Moreover, the back surface of the PCB is the top wall of the rectangular waveguide. A rectangular waveguide is formed by covering the bottom part of the metallic waveguide with the PCB. The waveguide is connected to the single RF chain through its feeding port. A slot pair comprises two longitudinal slots, which are offset symmetrically about the centerline of the waveguide, with the same dimensions. An RF switch is mounted across each slot aperture on the top surface of the PCB, and a DC bias line is printed on the inner layer of the PCB. The detailed configuration of the slot pair will be presented in Sect. 3.

The waveguide is here assumed to operate in only a single dominant mode ($TE_{10}$ mode). In that case, the transverse currents along the top inner waveguide wall are symmetrical about the centerline of the waveguide [20]. Because the longitudinal slot radiates by interrupting the transverse current, the field radiated by each slot of the slot pair has the same amplitude in anti-phase to each other. In the proposed antenna, one RF switch of each slot pair is turned off (OFF state), and the other is turned on (ON state), allowing one slot to receive an incident wave, whereas the other cannot. Therefore, the phase of the received signal of each slot pair can be reversed by switching the RF switches' state. The states of each switch are sequentially controlled according to high-speed spreading codes, and the received signal of each slot pair is code-modulated and spectrum-spread. Subsequently, spread signals are combined in the rectangular waveguide and fed into a single RF chain and an ADC. Following digitization, the original complex signals received by

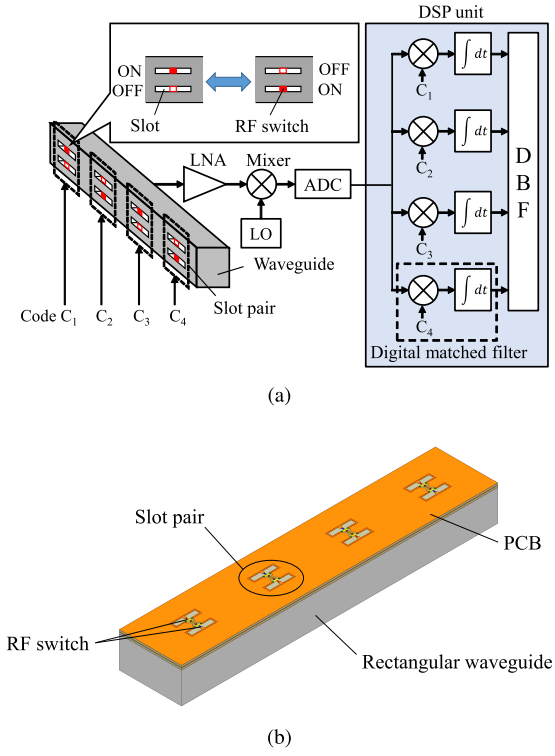



**Fig. 1** Configuration of the proposed antenna. (a) Block diagram and (b) perspective view of the waveguide slot array.

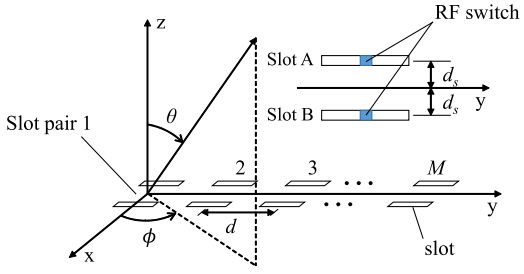



**Fig. 2** Radiating aperture of the proposed antenna with $M$ slot pairs.

each slot pair are demultiplexed by the despreading process in the DSP unit and then used for further beamforming or DOA estimation with DSP.

To understand the operating principle and signal-to-noise ratio (SNR) performance of the proposed antenna, we now present a mathematical model of the proposed antenna. Figure 2 illustrates a coordinate system of the radiating aperture of the proposed antenna. $M$ slot pairs are arranged parallel to the $y$-axis, and each slot is offset by $\pm d_s$ from it. Each slot is now assumed to have a radiation pattern given by

$$\boldsymbol{f}^{(\pm)}(\theta,\phi) = \boldsymbol{f}_m(\theta,\phi)e^{\pm jk_0 d_s \sin\theta\cos\phi} \tag{1}$$

where the superscripts (+) and (−) indicate the radiation pattern of the slot A and B, respectively, $\boldsymbol{f}_m(\theta,\phi)$ is the far-field vector of the $m$th slots, and $k_0$ is the free space wavenumber. We consider that a plane wave arrives at the antenna with an incident angle of $\theta$ in the $y$-$z$ plane ($\phi$ =

$\pi/2$). In this case, the received signal of the $m$th slot pair is given by

$$r_m(t) = c_m(t) f_{m\phi} s_m(t) e^{j2\pi f_c t} \tag{2}$$

where $f_{m\phi}$ is the $\phi$-component of $\boldsymbol{f}_m(\theta, \phi)$ in the $y$-$z$ plane (we ignore the cross-polarization component), $f_c$ is the carrier frequency, $s_m(t)$ is an incident signal to the slot pair, and $c_m(t)$ is a designated spreading code and indicates an operating state of the slot pair: $c_m(t) = 1$ when the slot A is activated and $c_m(t) = -1$ when the slot B is activated. Here we assume that the slot with an RF switch in the ON state is perfectly inactivated and does not receive any signals.

All the received signals $r_m(t)$ (for $1 \le m \le M$) are combined by the waveguide, and the output signal is then amplified by an LNA. The resultant signal is expressed as

$$R(t) = g\left[\sum_{m=1}^{M} \alpha_m r_m(t) + n(t)\right] \tag{3}$$

where $\alpha_m$ is the transmission characteristics of each slot pair through the waveguide, $g$ is the gain of the LNA, and $n(t)$ is the white Gaussian noise referred to the input of the LNA. $R(t)$ is then downconverted to the baseband, and the baseband signal is given by

$$R_b(t) = g\left[\sum_{m=1}^{M} \alpha_m f_{m\phi} c_m(t) s_m(t) + n(t)\right]. \tag{4}$$

Following digitization, the signal is separated into I and Q components by a digital IQ demodulator, but we ignore the IQ demodulator and show only a single path for simplicity in Fig. 1(a). The signal is then input to $M$ digital matched filters for the despreading process, where the $m$th matched filter is matched to the $m$th spreading code. When we assume that the incident signal $s_m(t)$ varies more slowly than the spreading code $c_m(t)$, the output signal $S_m(t)$ of the $m$th matched filter is expressed as

$$S_m(t) = \frac{1}{T_b}\int_0^{T_b} R_b(t) c_m(t)\, dt = S_{m,d}(t) + S_{m,i}(t) \tag{5}$$

$$S_{m,d}(t) = g\alpha_m f_{m\phi} C_{m,m} s_m(t) \tag{6}$$

$$S_{m,i}(t) = g{\sum}' \alpha_{m'} f_{m'\phi} C_{m,m'} s_{m'}(t) + \frac{g}{T_b}\int_0^{T_b} n(t) c_m(t)\, dt \tag{7}$$

where $T_b$ is the symbol duration, and $\sum'$ indicates a summation from $m' = 1$ to $M$ except for $m' = m$. $S_{m,d}(t)$ represents the desired signal for the $m$th slot-pair, whereas $S_{m,i}(t)$ is the interference from other signals and noise. $C_{m,m'}$ is the cross-correlation of the spreading codes and is defined as

$$C_{m,m'} = \frac{1}{T_b}\int_0^{T_b} c_m(t) c_{m'}(t)\, dt. \tag{8}$$

In (7), the first term of the right side is the interference from other signals, and a spreading code with $C_{m,m'} = 0$ must be used to minimize $S_{m,i}(t)$. Therefore, although there are several options for the spreading codes [21], the perfectly orthogonal codes, namely the Walsh-Hadamard (WH) codes (see Appendix A), with a length of $L_c \ge M$ must be used as the spreading codes. When we assume that synchronization between spreading and despreading is achieved, the orthogonality is satisfied for the WH codes as follows:

$$C_{m,m'} = \begin{cases} 1 & \text{for } m = m' \\ 0 & \text{for } m \ne m', \end{cases} \tag{9}$$

and $S_{m,d}(t)$ and $S_{m,i}(t)$ are reduced to

$$S_{m,d}(t) = g\alpha_m f_{m\phi} s_m(t) \tag{10}$$

$$S_{m,i}(t) = \frac{g}{T_b}\int_0^{T_b} n(t) c_m(t)\, dt. \tag{11}$$

As expected, the original signal can be perfectly recovered for the WH codes. The SNR at each element output becomes

$$SNR_m = \frac{|\alpha_m|^2 G_m P_s}{\sigma_n^2} \tag{12}$$

where $P_s = E[|s_m(t)|^2]$, $\sigma_n^2 = E[|n(t)|^2]$ ($E[\cdot]$ is an expectation operator), and $G_m = |f_{m\phi}|^2$ is the gain of each slot pair (see Appendix B). Similarly, SNR at the array output is given as

$$SNR_a = \frac{M\alpha_0^2 G_e P_s}{\sigma_n^2} \tag{13}$$

where we assume that $|\alpha_m| = \alpha_0$ and $G_m = G_e$ for all slot pairs. For the hollow waveguide, $|\alpha_m|^2 = \alpha_0^2 \simeq 1/M$ in (12) and (13), and thus the SNRs are degraded to $1/M$ of those for the element-level DBF array antenna or the previous works in [11] and [12]. This is because the signal amplification is performed after signal combination, and the signal power of each slot pair is degraded due to the insertion loss of the waveguide combiner. This SNR degradation is a drawback of the proposed antenna, and there is another restriction. When we use the spreading code with length $L_c$, the chip duration becomes $T_b/L_c$. Therefore, the ADC sampling rate and PIN diode's switching rate must be faster than $M/T_b$. Thus, the available ADC sampling rate and PIN diode's switching speed dictate the maximum number of multiplexing signals and maximum symbol rate in the proposed antenna.

Although we have here considered the receiving antenna, the same idea is applicable to a transmitting antenna. In that case, the transmitting signals from each slot-pair are code-modulated in the digital domain and recovered by the despreading process by switching the RF switches' bias. However, there are some difficult problems for the transmitting antenna such as spurious emissions and power handling capability of the RF switch, which will hinder practical application. Therefore, we focus only on the receiving antenna in this study. In addition, we here assume that the incident wave lies in the $y$-$z$ plane. For other incident angles out of this plane, the factor $e^{\pm jk_0 d_s \sin\theta\cos\phi}$ remains in (2), which introduces errors in the spreading process. The effect of these errors will be discussed in Sect. 4.

## 3. Antenna Design

We designed a prototype waveguide slot array with eight slot pairs as a proof of concept of the proposed antenna. Figure 3 demonstrates a detailed configuration of a slot pair for the prototype antenna. Each slot pair and its DC bias line are printed on a 4-layers PCB. The conductor layers are denoted as L1–L4, and two substrates with a thickness of 0.75 mm are bound with a prepreg. The total thickness of the PCB is 1.67 mm. Megtron6 ($\epsilon_r = 3.6$ and $\tan\delta = 0.003$) by Panasonic Industry is selected as a substrate material. Slot apertures with a length of $L_s$ and a width of $w_s$ are formed by etching all the layers. Additionally, multiple through-hole vias with a diameter of 0.4 mm are placed around the slot apertures at 0.7 mm intervals to prevent RF leakage into the parallel plate mode between layers.

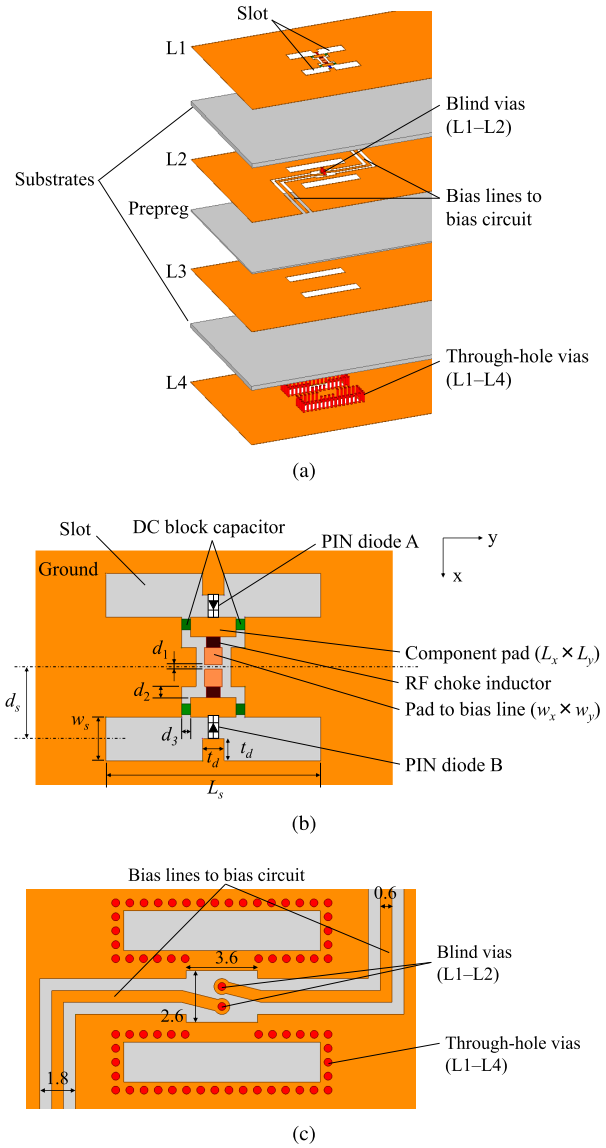



**Fig. 3** Configuration of a slot pair for the prototype antenna. (a) Expanded view of the PCB, (b) detailed configuration of slot pair with PIN diodes, RF choke inductors, and DC block capacitors, and (c) layout of the L2 layer.

PIN diode (SMP1345-079LF by Skyworks) is used as an RF switch. A single PIN diode is placed across each slot at its center on the L1 layer. DC bias lines of each PIN diode are printed on the L2 layer and connected to the bias circuit (not shown in Fig. 3). An RF choke inductor of 15 nH (LQW15AN15NG00 by Murata Manufacturing) is placed between the PIN diode and the bias line to prevent the RF leakage into the bias line. Two DC block capacitors of 1.2 pF (GJM1555C1H1R2WB01 by Murata Manufacturing) are placed between the component pad and the ground to isolate the bias line from the ground. In the proposed antenna, one diode of each slot pair is reverse biased (OFF state), and the other is forward biased (ON state). When the PIN diode is in the OFF state, it approximates an open circuit and has minimal effect on the radiation characteristics of the slot. In contrast, when the PIN diode is in the ON state, it approximately short-circuits the slot at its center, effectively deactivating it.

The inter-element spacing $d$, distance between the centers of adjacent slot pairs, is set to 33.1 mm for the prototype antenna to have a field of view of ±15° without any grating lobes at the operating frequency of 7 GHz. A standing-wave waveguide slot array is adopted in the prototype antenna, and the dimensions of the bottom part of metallic waveguide is designed such that a guided wavelength becomes twice the inter-element spacing. We select a single-ridged waveguide, and the designed dimensions are listed in Table 1.

A slot pair with the integrated components is designed using the electromagnetic simulator Ansys HFSS. In the prototype antenna, we use slot pairs with identical dimensions to simplify the antenna design, and we determine them such that the slots resonate at 7 GHz with an appropriate radiation conductance for impedance matching. The aperture distribution would degrade due to this simplified design. However, we accept it here because the signal variations among the slot pairs caused by it can be calibrated in the digital domain after demultiplexing. In the simulations, the PIN diode is modeled as an equivalent circuit of lumped elements to consider its effect on the slot performance. Specifically, the PIN diode in the ON state is modeled as a series circuit of a 2.0 Ω resistor and a 0.7 nH inductor, whereas one in the OFF state is modeled as a 0.13 pF capacitor [22], [23]. The designed dimensions of a slot pair are shown in Table 1.

Figure 4 depicts an exploded view of the prototype

**Table 1** Dimensions of the prototype antenna.

| Parameter | Value [mm] | Parameter | Value [mm] |
|---|---|---|---|
| Waveguide | | Pad and clearance | |
| inner width | 23.0 | $L_x$ | 0.9 |
| inner height | 11.5 | $L_y$ | 2.1 |
| ridge height | 5.0 | $w_x$ | 0.8 |
| ridge width | 1.0 | $w_y$ | 0.8 |
| Slot-pair | | $d_1$ | 0.2 |
| $L_s$ | 9.9 | $d_2$ | 0.5 |
| $w_s$ | 2.0 | $d_3$ | 0.4 |
| $d_s$ | 3.3 | $t_d$ | 1.0 |

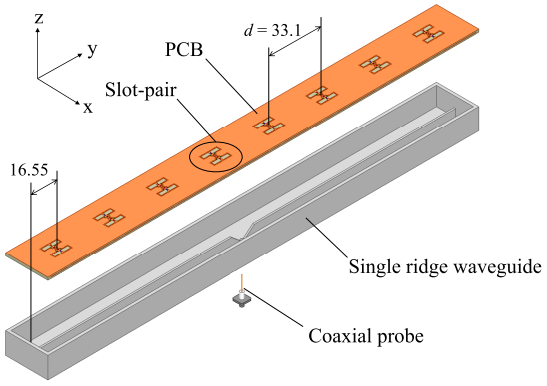



**Fig. 4** Exploded view of the prototype waveguide slot array (dimensions in millimeter).

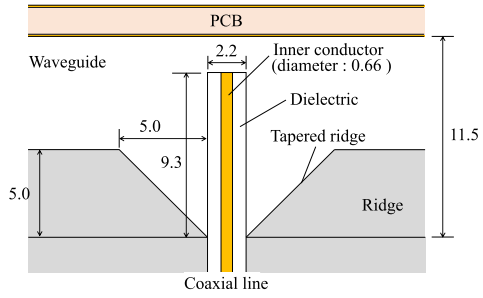



**Fig. 5** Configuration of the designed transition from the coaxial line to the single-ridged waveguide (dimensions in millimeter).

antenna. Eight designed slot pairs are arranged linearly, and both ends of the ridge waveguide are short-circuited at 16.55 mm away from the center of the end slot pairs. The ridge waveguide is fed by an SMA coaxial probe inserted at its center from its bottom side. Figure 5 shows the detailed structure of the transition from the coaxial line to the ridge waveguide. The coaxial inner conductor and the dielectric are extended into the waveguide, and the ridges are linearly tapered at the transition to achieve impedance matching.

## 4. Experiments

The prototype antenna was fabricated and tested to verify the feasibility of the proposed antenna. Figure 6 shows photos of the fabricated antenna. The prototype antenna was installed in a ground plane of dimension 400 mm×400 mm. Bias lines for PIN diodes were connected to bias circuits (DC power suppliers) through biasing wires. The PIN diodes were forward biased with the forward current of 10 mA and were reverse biased with the reverse voltage of 22 V.

Here we consider that a continuous wave (CW) with an amplitude $s_0$ impinges on the proposed antenna with the incident angle $\theta$ in the $y$-$z$ plane. The incident signal to the $m$th slot pair is given by $s_m(t) = s_0 e^{jk_0 md\sin\theta}$, and thus the recovered desired signal (10) becomes $S_{m,d}(t) = g\alpha_m s_0 f_{m\phi} e^{jk_0 md\sin\theta}$ and is proportional to the element pattern of the $m$th slot pair. In the experiment, therefore, CWs at 7 GHz were transmitted and received while changing $\theta$,

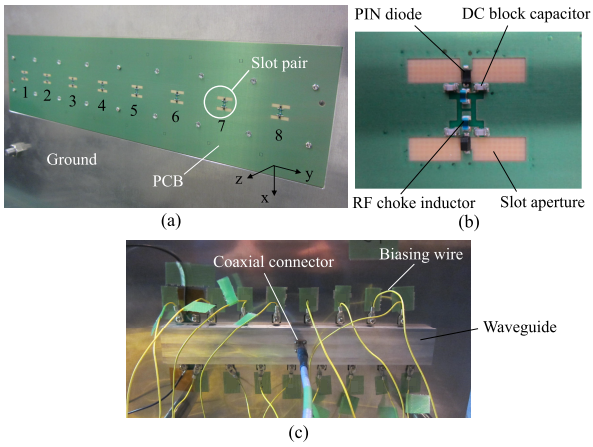
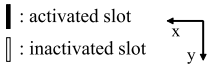



**Fig. 6** Fabricated prototype antenna.

and we then compared the recovered signals by the prototype antenna with simulated element patterns. The measurement was performed using a compact antenna test range setup in an anechoic chamber. CW transmission and reception, transmission coefficient ($S_{21}$) measurement, between a transmitting antenna and the prototype one was conducted with a vector network analyzer. To change the incident angle $\theta$, the prototype antenna was rotated around the $x$-axis, which passed through the waveguide center, with a rotary positioner. For a proof of concept, we manually switched the biasing state of each PIN diode to the specified states described later and measured the $S_{21}$ at various $\theta$ (from $-90°$ to $90°$ in $0.1°$ increments) in each state. Subsequently, we recovered the received signal $e_m$ of each slot-pair from the measured signals at $\theta$ by the following expression:

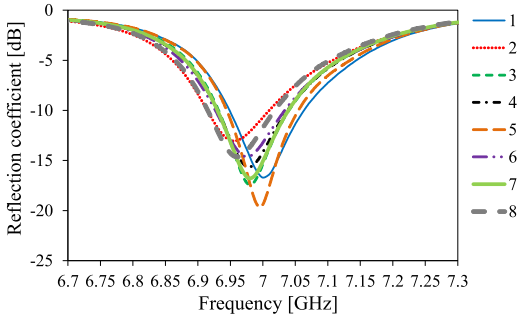
$$e_m(\theta) = \frac{1}{L_c}\sum_{l=1}^{L_c} E_l(\theta)c_m(l) \tag{14}$$

where $E_l(\theta)$ and $c_m(l)$ are the measured $S_{21}$ at $\theta$ and the designated spreading code for the $m$th slot pair at the $l$th state, respectively. This process was performed in software. We assumed that synchronization between spreading and despreading was perfectly achieved and ignored the transient characteristics of the PIN diodes. We used the WH codes with $L_c = 8$ or 16 as the spreading codes in the experiment. Figure 7 demonstrates the relationship between the WH codes with $L_c = 8$ and the positions of activated slots (with the PIN diode in the OFF state). In the 1st state, the activated slots were located to be excited in phase at 7 GHz. We assigned this arrangement to +1 of each code and determined the positions of activated slots for the other states. For the WH codes with $L_c = 16$, we avoided using the WH code whose elements were all +1 for a reason discussed later and used the codes corresponding to the 2nd to 10th rows of the Hadamard matrix (see Fig. A· 1). We determined the positions of the activated slots for the WH codes with $L_c = 16$ in a similar manner to $L_c = 8$, and we here omit the detailed arrangement.

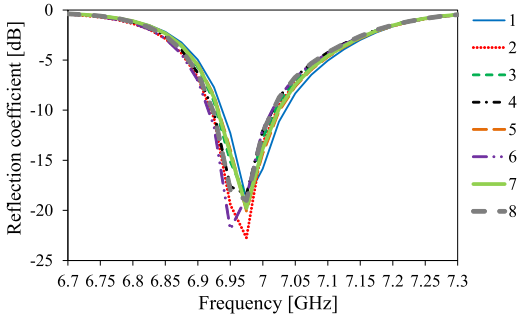



| Slot-pair number | WH code | State number | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | $c_1 = (+,+,+,+,+,+,+,+)$ | | | | | | | | |
| 2 | $c_2 = (+,-,+,-,+,-,+,-)$ | | | | | | | | |
| 3 | $c_3 = (+,+,-,-,+,+,-,-)$ | | | | | | | | |
| 4 | $c_4 = (+,-,-,+,+,-,-,+)$ | | | | | | | | |
| 5 | $c_5 = (+,+,+,+,-,-,-,-)$ | | | | | | | | |
| 6 | $c_6 = (+,-,+,-,-,+,-,+)$ | | | | | | | | |
| 7 | $c_7 = (+,+,-,-,-,-,+,+)$ | | | | | | | | |
| 8 | $c_8 = (+,-,-,+,-,+,+,-)$ | | | | | | | | |

**Fig. 7** Positional relationship between the WH codes with $L_c = 8$ and the activated slots (+ and − indicate +1 and −1, respectively).




(a)




(b)

**Fig. 8** Reflection coefficients in each state for the WH codes with $L_c = 8$: (a) Measurement and (b) simulation.

### 4.1 Experimental Results

Figures 8(a) and (b) show the measured and simulated reflection coefficients at each state for the WH codes with $L_c = 8$, respectively. Although there are some discrepancies between the measured and simulated reflection coefficients, the prototype antenna has a reflection coefficient of less than −10 dB at the operating frequency in all states. The causes for the discrepancies seem to be fabrication errors and errors in electrical properties of the integrated components. The reflection coefficients vary depending on the states due to the variation of the active admittance of each slot pair. The prototype antenna has a bandwidth of only 1.5% with a reflection coefficient of less than −10 dB, and thus a bandwidth enhancement is one of the future works. The bandwidth of the slot can be improved by replacing the PCB material with one with a low dielectric constant, and the slot array's bandwidth can be improved by replacing its series-feed structure with a corporate-feed type [24]. Since the measured reflection coefficients for the WH codes with $L_c = 16$ were almost the same as those with $L_c = 8$, we here omit the detailed results. The measured gain in the broadside direction at 7 GHz for the 1st state, corresponding to an ordinary probe-fed waveguide slot array, is 11.1 dBi, and the estimated radiation efficiency is approximately 47%. The loss due to the reverse parallel resistance $R_p$ of the PIN diode is a significant factor in the reduction in radiation efficiency. $R_p$ of the PIN diode for the prototype antenna is estimated to be approximately 10 kΩ by simulations. We can improve the radiation efficiency by replacing the PIN diode with one with a larger $R_p$, which is also a future challenge.

Figure 9(a) shows the recovered signals of the 1st to 4th slot pairs at each sample angle from the measured $S_{21}$ at 7 GHz for the WH codes with $L_c = 8$. In addition, recovered signals from the simulated $S_{21}$ and simulated individual element patterns of each slot pair are shown in Fig. 9(a). Due to the symmetry of the antenna structure, the recovered signals of the 5th to 8th slot pairs have similar patterns, which we omit here due to space constraints. The amplitudes of each signal are normalized by their respective peak amplitudes, and the phases are calibrated to be 0 deg. in the broadside direction. The recovered signals of each slot pair agree well with their respective element pattern in simulation, whereas there are some discrepancies in measurement. Particularly, there are larger discrepancies between measured and simulated recovered signals for the 1st slot pair, and we find that demultiplexing is not working well. We will later discuss the cause of this deterioration in recovered signals. Similarly, Fig. 9(b) shows the recovered signals with the WH codes with $L_c = 16$. In this case, the recovered signals from the measured $S_{21}$ agree with simulated element patterns for all slot pairs. Although there are some discrepancies between measured and simulated recovered signals, the prototype antenna is successful in demultiplexing.

### 4.2 Discussions

We first discuss the effect of the differences in properties of each PIN diode on the recovered signals to reveal the cause of the deterioration in the recovered signals of the 1st slot-pair with the WH codes with $L_c = 8$.

We now consider the same model as that in Sect. 2. When the PIN diodes have errors in their electrical properties, the received signal of each slot-pair deviates from its ideal. Now consider $\delta$ be the fractional error of the received signal due to the errors in the PIN diode's properties, and the received signal $r'_m(t)$ of the mth slot pair under these errors

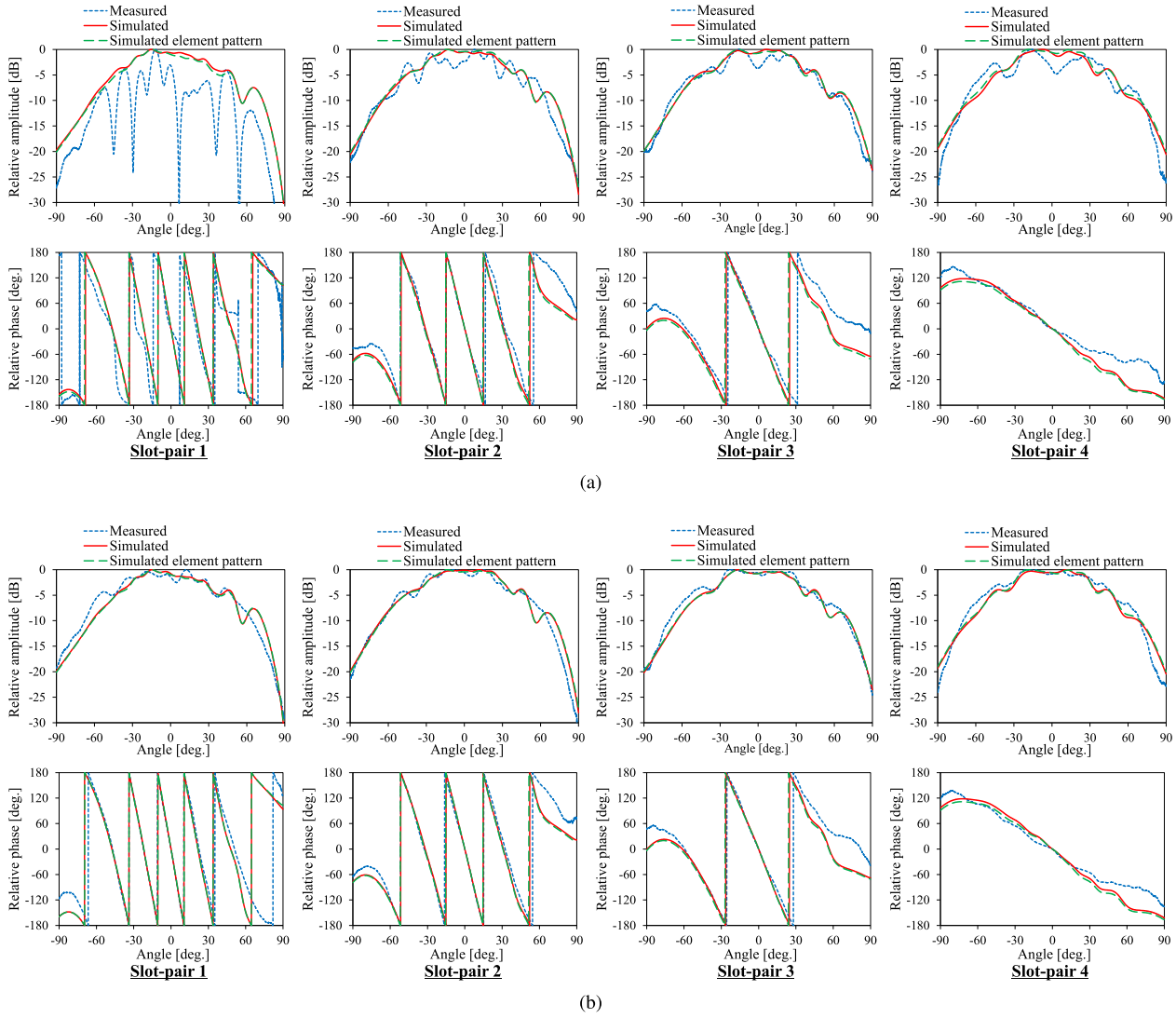



**Fig. 9** Recovered signals of the 1st to 4th slot-pairs with the WH codes with (a) $L_c = 8$ and (b) $L_c = 16$ at 7 GHz. Blue dashed and red solid lines indicate the recovered signals from the measured and simulated $S_{21}$, respectively, and long-dashed green lines indicate the simulated element patterns of each slot pair. The amplitude and phase patterns are shown in the upper and lower rows, respectively.

is then expressed by

$$r'_m(t) = \{1 + \delta_m(t)\}\, r_m(t). \tag{15}$$

The errors depend on the operating states of the slot pair, and we denote that $\delta_m(t) = \delta_m^{(\pm)}$ when the slot corresponding to $\pm 1$ is activated. Replacing $r_m(t)$ in (3) with $r'_m(t)$ and using (4) and (5), we obtain the following recovered signals:

$$S'_{m,d}(t) = \frac{g\alpha_m f_{m\phi} s_m(t)}{T_b} \times \int_0^{T_b} c_m^2(t)\,\{1 + \delta_m(t)\}\, dt \tag{16}$$

$$S'_{m,i}(t) = \frac{g}{T_b} {\sum}' \alpha_{m'} f_{m'\phi} s_{m'}(t) \times \int_0^{T_b} c_{m'}(t)\,\{1 + \delta_{m'}(t)\}\, c_m(t)\, dt \tag{17}$$

where we ignore the noise for simplicity. When we consider a case where the WH code whose elements are all +1 is used for the $m$th slot-pair, (16) and (17) are reduced to

$$S'_{m,d}(t) = g\alpha_m f_{m\phi} s_m(t)(1 + \delta_m^{(+)}) \tag{18}$$

$$S'_{m,i}(t) = \frac{g}{2} {\sum}' \alpha_{m'} f_{m'\phi} s_{m'}(t)(\delta_{m'}^{(+)} - \delta_{m'}^{(-)}). \tag{19}$$

According to (19), $S'_{m,i}(t) \neq 0$ when $\delta_{m'}^{(+)} \neq \delta_{m'}^{(-)}$, and thus the recovered signal suffers from the remaining interference from other slot-pairs' signals. However, when the WH codes, including both +1 and −1, are used for the $m$th slot pair, (16) and (17) are given by

$$S'_{m,d}(t) = g\alpha_m f_{m\phi} s_m(t) \left(1 + \frac{\delta_m^{(+)} + \delta_m^{(-)}}{2}\right) \tag{20}$$

$$S'_{m,i}(t) = 0. \tag{21}$$

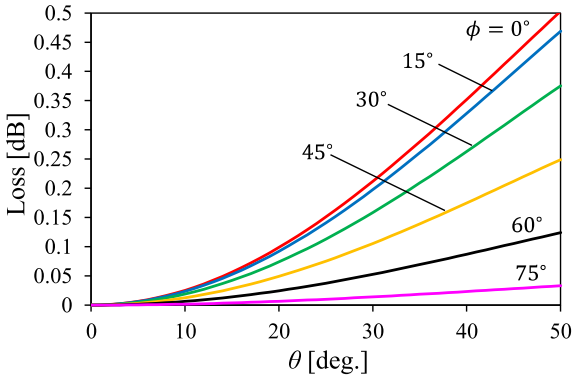



**Fig. 10** Calculated losses due to the factor $\cos(k_0 d_s \sin\theta\cos\phi)$.

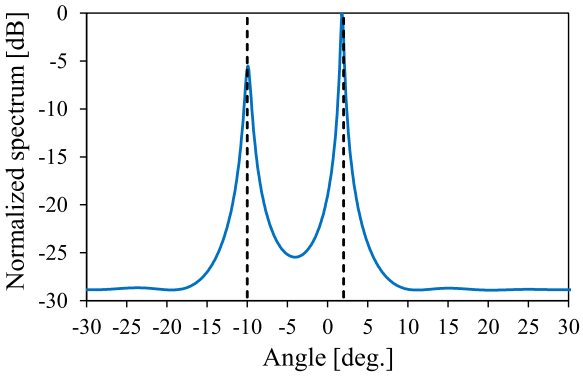



**Fig. 11** Simulated MUSIC spatial spectrum for two uncorrelated incident signals with $(\theta_1, \theta_2) = (-10.0°, 2.0°)$ and SNR = 20 dB.

Because $c_m(t)$ and $c_{m'}(t)\{1+\delta_m(t)\}$ are also orthogonal when $m \neq 1$ (see Appendix A), the interference from other slot-pairs' signals disappear in (21). Therefore, in an actual situation where the PIN diodes have errors in their electrical properties, we must use the WH codes including both +1 and −1, with a length of $L_c > M$, to prevent deterioration in recovered signals.

Next, we consider that a plane wave arrives at the antenna with an incident angle $(\theta, \phi)$ out of the $y$-$z$ plane in Fig. 2. In this case, the received signal $r''_m(t)$ is given by

$$r''_m(t) = c_m(t) f_m e^{\pm jk_0 d_s \sin\theta\cos\phi} s_m(t) e^{j2\pi f_c t} \tag{22}$$

where $f_m = |\boldsymbol{f}_m(\theta,\phi)|$, and we assume that the polarization of the antenna is matched to that of the incident plane wave. In (22), the factor $e^{\pm jk_0 d_s \sin\theta\cos\phi}$ causes an error depending on the operating state of the slot pair for the received signals. The effect of this error can be evaluated with (20) by replacing $\delta_m^{(\pm)}$ and $f_{m\phi}$ by $(e^{\pm jk_0 d_s \sin\theta\cos\phi} - 1)$ and $f_m$, respectively, and the recovered signal becomes

$$S''_{m,d}(t) = g\alpha_m f_m s_m(t)\cos(k_0 d_s \sin\theta\cos\phi). \tag{23}$$

The factor $\cos(k_0 d_s \sin\theta\cos\phi)$ introduces an additional loss into the recovered signal depending on the incident angle. Figure 10 shows calculated losses with various incident angles for the prototype antenna. The prototype antenna suffers from larger losses when $\phi$ is around 0° (for example, the loss of 0.43 dB for $(\theta,\phi) = (45°, 0°)$). Slot pairs with a smaller slot offset should be used to mitigate this loss, and it requires miniaturizing slot pairs. Slot offset is closely related to slot admittance and mounting space for the inductors and capacitors. Therefore, miniaturization of the slot pair is also one of the future challenges.

### 4.3 Application Example: DOA Estimation

Simulation tests of DOA estimation are performed to demonstrate the feasibility of DBF function by the prototype antenna. In the simulation, the one-dimensional MUSIC algorithm [25] is employed, and the received signal of the antenna is calculated using the measured $S_{21}$. Thus, received signals at the $l$th state of the prototype antenna are calculated by

$$R_b(t) = \sum_{k=1}^{K} s^{(k)}(t,\theta_k) E_l(\theta_k) + n(t) \tag{24}$$

where $s^{(k)}(t,\theta_k)$ represents the $k$th incident signal to the antenna, $\theta_k$ is its incident angle, and $K$ is the number of the incident signals. Recovered signals $S_m(t)$ are calculated by (5), and the MUSIC spectrum is then calculated by [25]

$$P_{MU}(\theta) = \frac{1}{\boldsymbol{a}^{\mathrm{H}}(\theta)\boldsymbol{E}_n\boldsymbol{E}_n^{\mathrm{H}}\boldsymbol{a}(\theta)} \tag{25}$$

where the superscript H represents the Hermitian transpose operation, $\boldsymbol{a}(\theta) = [C_1 e^{-jk_0 d\sin\theta}, \cdots, C_M e^{-jk_0 M d\sin\theta}]$, and $C_m$ are calibration weights and are determined using a reference signal from $\theta = 0°$, ignoring the mutual coupling effect for simplicity. $\boldsymbol{E}_n$ is a matrix whose columns are the noise eigenvectors of the covariance matrix $\boldsymbol{R} = E[\boldsymbol{S}(t)\boldsymbol{S}^{\mathrm{H}}(t)]$ where $\boldsymbol{S}(t) = [S_1(t), \cdots, S_M(t)]^{\mathsf{T}}$, and the superscript T represents the transpose operation.

Here, we consider two uncorrelated QPSK modulated incident signals with the same power. The incident angles are $(\theta_1, \theta_2) = (-10.0°, 2.0°)$. The operating state of the antenna is switched 16 times during each symbol according to the WH codes with $L_c = 16$. We assume that synchronization between spreading and despreading is perfectly achieved. The covariance matrix $\boldsymbol{R}$ is calculated using recovered signals of 100 symbols. Figure 11 shows the calculated MUSIC spatial spectrum for SNR = 20 dB. We can find two obvious peaks in the spectrum, and estimated incident angles are −10.0° and 1.8°, which agree well with the actual ones. The estimation error would be due to an effect of mutual couplings between slot pairs and imperfection in demultiplexing.

## 5. Conclusion

In this study, we presented a novel waveguide slot array with slot pairs for the CDM function. The received signals of each slot pair are multiplexed in the CDM fashion by switching the diodes' bias according to the WH codes, and they are recovered through the despreading process in the digital domain for further DSP. We also presented the

operating principle of the proposed antenna using the mathematical model, and then designed, fabricated, and tested a proof of concept prototype antenna with eight slot pairs. Although we manually switched the PIN diodes' bias, ignored the transient characteristics of the PIN diodes, and assumed perfect synchronization between spreading and despreading, the experimental results demonstrated the feasibility of the proposed antenna. We also showed that the WH codes including both +1 and −1, with a length of $L_c > M$, must be used to prevent recovered signals' deterioration due to errors in the PIN diodes' electrical properties.

Future works of the proposed antenna include further tests considering the transient characteristics of the PIN diodes and actual synchronization issues, bandwidth enhancement, radiation efficiency improvement, and miniaturization of the slot pair for mitigating the losses depending on the incident angle of signals. The proposed antenna has the potential to provide a single RF Chain DBF antenna with lower cost and higher radiation efficiency.

**References**

[1] C. Fulton, M. Yeary, D. Thompson, J. Lake, and A. Mitchell, "Digital phased arrays: Challenges and opportunities," Proc. IEEE, vol.104, no.3, pp.487–503, March 2016.
[2] P.K. Bailleul, "A new era in elemental digital beamforming for space-borne communications phased arrays," Proc. IEEE, vol.104, no.3, pp.623–632, March 2016.
[3] B. Yang, Z. Yu, J. Lan, R. Zhang, J. Zhou, and W. Hong, "Digital beamforming-based massive MIMO transceiver for 5G millimeter-wave communications," IEEE Trans. Microw. Theory Tech., vol.66, no.7, pp.3403–3418, July 2018.
[4] J.D. Fredrick, Y. Wang, and T. Itoh, "Smart antennas based on spatial multiplexing of local elements (SMILE) for mutual coupling reduction," IEEE Trans. Antennas Propag., vol.52, no.1, pp.106–114, Jan. 2004.
[5] K. Koga, N. Kikuma, H. Hirayama, K. Sakakibara, T. Koike, H. Iwasaki, and Y. Mizuno, "A study of switching methods for an adaptive array with a single receiver using time-division multiplexing," Trans. IEICE Commun. (Japanese Edition), vol.J96-B, no.2, pp.149–162, Feb. 2013.
[6] S. Henault, B.R. Jackson, and Y.M.M. Antar, "Compensation of time-division multiplexing distortion in switched antenna arrays with a single RF front-end and digitizer," IEEE Trans. Antennas Propag., vol.61, no.8, pp.4383–4388, Aug. 2013.
[7] Y. Yashchyshyn, K. Derzakowski, P.R. Bajurko, J. Marczewski, and S. Kozłowski, "Time-modulated reconfigurable antenna based on integrated S-PIN diodes for mm-Wave communication systems," IEEE Trans. Antennas Propag., vol.63, no.9, pp.4121–4131, Sept. 2015.
[8] M.A. Johnson, "Phased-array beam steering by multiplex sampling," Proc. IEEE, vol.56, no.11, pp.1801–1811, Nov. 1968.
[9] N. Akram, A. Madanayake, S. Pulipati, V. Ariyarathna, S.B. Venkatakrishnan, D. Psychogiou, J. Volakis, T.S. Rappaport, and T.L. Marzetta, "Frequency-multiplexed array digitization for MIMO receivers: 4-antennas/ADC at 28 GHz on Xilinx ZCU-1285 RF SoC," IEEE Access, vol.9, pp.142743–142753, 2021.
[10] A. Jahanian, F. Tzeng, and P. Heydari, "Code-modulated path-sharing multi-antenna receivers: Theory and analysis," IEEE Trans. Wireless Commun., vol.8, no.5, pp.2193–2201, May 2009.
[11] F. Tzeng, A. Jahanian, D. Pi, and P. Heydari, "A CMOS code-modulated path-sharing multi-antenna receiver front-end," IEEE J. Solid-State Circuits, vol.44, no.5, pp.1321–1335, May 2009.
[12] E.A. Alwan, A. Akhiyat, M. LaRue, W. Khalil, and J.L. Volakis, "Low cost, power efficient, on-site coding receiver (OSCR) for ultra-wideband digital beamforming," 2013 IEEE Int. Symp. Phased Array Systems Tech., pp.202–206, 2013.
[13] S.B. Venkatakrishnan, D.K. Papantonis, A.A. Akhiyat, E.A. Alwan, and J.L. Volakis, "Experimental validation of on-site coding digital beamformer with ultra-wideband antenna arrays," IEEE Trans. Microw. Theory Techn., vol.65, no.11, pp.4408–4417, Nov. 2017.
[14] J. Zhang, W. Wu, and D.-G. Fang, "Single RF channel digital beamforming multibeam antenna array based on time sequence phase weighting," IEEE Antennas Wireless Propag. Lett., vol.10, pp.514–516, 2011.
[15] X. Wang and C. Caloz, "Direction-of-arrival (DOA) estimation based on spacetime-modulated metasurface," 2019 IEEE Int. Symp. Antennas Propag. and USNC-URSI Radio Science Meeting, pp.1613–1614, 2013.
[16] X. Wang and C. Caloz, "Pseudorandom sequence (space–) time-modulated metasurfaces: Principles, operations, and applications," IEEE Antennas Propag. Mag., vol.64, no.4, pp.135–144, Aug. 2022.
[17] Y. Yashchyshyn, J. Marczewski, K. Derzakowski, J.W. Modelski, and P.B. Grabiec, "Development and investigation of an antenna system with reconfigurable aperture," IEEE Trans. Antennas Propag., vol.57, no.1, pp.2–8, Jan. 2009.
[18] Z. Zhou, H. Su, B. Yang, S. Wu, and X.Y. Zhang, "Digital encoding holographic antenna based on SIW waveguide," 2021 IEEE Int. Workshop on Electromagnetics: Applications and Student Innovation Competition (iWEM), 2021.
[19] H. Li, S. Li, B. Hou, X. Zhang, W. Wen, and C. Hu. "A digital SIW-slot antenna array with FPGA implementation of beamforming," Sci. Rep., vol.12, 8927, May 2022.
[20] R.E. Collin, "Waveguides and cavities," Field Theory of Guided Waves, pp.329–410, IEEE, 1991.
[21] E.H. Dinan and B. Jabbari, "Spreading codes for direct sequence CDMA and wideband CDMA cellular networks," IEEE Commun. Mag., vol.36, no.9, pp.48–54, Sept. 1998.
[22] Skyworks, "Design with PIN diodes," Skyworks Solutions, Inc., https://www.skyworksinc.com/-/media/skyworks/documents/products/1-100/Design_With_PIN_Diodes_200312E.pdf, July 2023.
[23] Skyworks, "SMP1345 series: very low capacitance, plastic packaged silicon PIN diodes," Skyworks Solutions, Inc., https://www.skyworksinc.com/-/media/skyworks/documents/products/101-200/smp1345_series_200046u.pdf, July 2023.
[24] Y. Miura, J. Hirokawa, M. Ando, Y. Shibuya, and G. Yoshida, "Double-layer full-corporate-feed hollow-waveguide slot array antenna in the 60-GHz band," IEEE Trans. Antennas Propag., vol.59, no.8, pp.2844–2851, Aug. 2011.
[25] R. Schmidt, "Multiple emitter location and signal parameter estimation," IEEE Trans. Antennas Propag., vol.34, no.3, pp.276–280, March 1986.
[26] W.K. Pratt, J. Kane, and H.C. Andrews, "Hadamard transform image coding," Proc. IEEE, vol.57, no.1, pp.58–68, Jan. 1969.
[27] M.P. Deisenroth, A.A. Faisal, and C. S Ong, Mathematics for Machine Learning, Cambridge University Press, Cambridge, 2020.
[28] G.B. Arfken and H.J. Weber, Mathematical Methods for Physicists, 5th ed., Academic Press, San Diego, 2000.

## Appendix A: Walsh Hadamard Code

The Walsh Hadamard (WH) codes are based on the Hadamard matrices, which are square arrays of +1 and −1 and whose rows or columns are completely orthogonal to each other. The Hadamard matrices are generated as follows [26]:

$$\boldsymbol{H}_2 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \boldsymbol{H}_{2N} = \begin{bmatrix} \boldsymbol{H}_N & \boldsymbol{H}_N \\ \boldsymbol{H}_N & -\boldsymbol{H}_N \end{bmatrix} \quad (\text{A}\cdot 1)$$

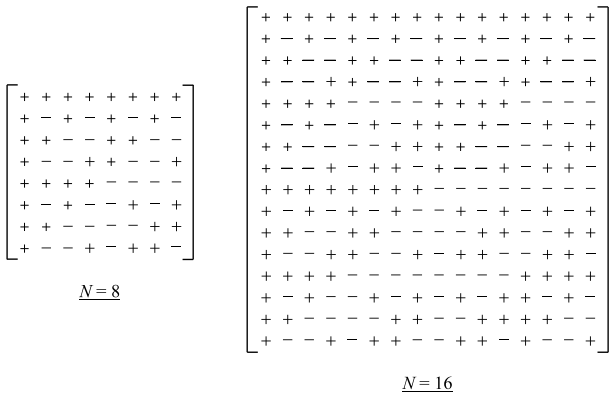



**Fig. A · 1** The Hadamard matrices of orders 8 and 16. + and − indicate +1 and −1, respectively.

where the subscripts indicate the orders of the Hadamard matrices. We show the Hadamard matrices of orders 8 and 16 in Fig. A · 1 (the signs of ±1 are only shown). The WH code corresponds to one of the row(column) vectors of the Hadamard matrices. The orthogonality of the WH codes can also be expressed as follows:

$$\frac{1}{N}\boldsymbol{h}_p\boldsymbol{h}_q^{\mathsf{T}} = \begin{cases} 1 & \text{for } p = q \\ 0 & \text{for } p \neq q \end{cases} \tag{A·2}$$

where $\boldsymbol{h}_p$ and $\boldsymbol{h}_q$ are the $p$th and $q$th row vectors of the Hadamard matrix of order $N$, respectively.

We now consider a vector $\boldsymbol{\Delta}_q$ obtained by replacing +1 of $\boldsymbol{h}_q$ by $(1+\delta_q^{(+)})$ and −1 by $-(1+\delta_q^{(-)})$. Using a property of an orthogonal matrix [27], the Euclidean norm of $\boldsymbol{H}_N\boldsymbol{\Delta}_q^{\mathsf{T}}$ can be expressed as follows:

$$\frac{1}{N}\|\boldsymbol{H}_N\boldsymbol{\Delta}_q^{\mathsf{T}}\|^2 = \|\boldsymbol{\Delta}_q^{\mathsf{T}}\|^2. \tag{A·3}$$

When $q \neq 1$, the right side of (A · 3) is $N\{(1+\delta_q^{(+)})^2 + (1+\delta_q^{(-)})^2\}/2$. The first and $q$th elements of $\boldsymbol{H}_N\boldsymbol{\Delta}_q^{\mathsf{T}}$ are $N(\delta_q^{(+)} - \delta_q^{(-)})/2$ and $N(2+\delta_q^{(+)}+\delta_q^{(-)})/2$, respectively, and a sum of squares of them becomes $N^2\{(1+\delta_q^{(+)})^2 + (1+\delta_q^{(-)})^2\}/2$. To satisfy (A · 3), therefore, the other elements of $\boldsymbol{H}_N\boldsymbol{\Delta}_q^{\mathsf{T}}$ must be all 0. Similarly, when $q = 1$, the right side of (A · 4) is $N(1+\delta_q^{(+)})^2$, and the first element of $\boldsymbol{H}_N\boldsymbol{\Delta}_q^{\mathsf{T}}$ is $N(1+\delta_q^{(+)})$. Therefore, the other elements of $\boldsymbol{H}_N\boldsymbol{\Delta}_q^{\mathsf{T}}$ must be all 0. Since each element of $\boldsymbol{H}_N\boldsymbol{\Delta}_q^{\mathsf{T}}$ is a dot product of $\boldsymbol{h}_p$ and $\boldsymbol{\Delta}_q^{\mathsf{T}}$, the following orthogonality is then obtained:

$$\boldsymbol{h}_p\boldsymbol{\Delta}_q^{\mathsf{T}} = 0 \quad \text{for } p \neq q,\ p \neq 1. \tag{A·4}$$

## Appendix B: SNR Caluculation

The average power of the interference signal for the $m$th slot pair is given by

$$E[|S_{m,i}(t)|^2] = E\left[\left|\frac{g}{T_b}\int_0^{T_b} n(t)c_m(t)\,dt\right|^2\right]. \tag{A·5}$$

Using the Schwarz's inequality [28], the right side of (A · 5) becomes

$$\begin{aligned} &E\left[\left|\frac{g}{T_b}\int n(t)c_m(t)\,dt\right|^2\right] \\ &\le \frac{|g|^2}{T_b^2}E\left[\int_0^{T_b}|n(t)|^2\,dt\cdot\int_0^{T_b}|c_m(t)|^2\,dt\right] \\ &= \frac{|g|^2}{T_b}\int_0^{T_b}E[|n(t)|^2]\,dt = |g|^2\sigma_n^2. \end{aligned} \tag{A·6}$$

Therefore, the worst-case SNR at the element output becomes (12).

Next, we consider an SNR at the array output. When we assume that all incident signals to the slot pairs have an equal power, the gains of each slot pair are identical ($|f_{m\phi}|^2 = G_e$), and the amplitude of transmission characteristics of each slot pair through the waveguide are identical ($|\alpha_m| = \alpha_0$), the maximum average power of the combined desired signal is then calculated as

$$\begin{aligned} E\left[\left|\sum_{m=1}^{M} w_m S_{m,d}(t)\right|^2\right] &= E\left[\left|g\sum_{m=1}^{M} w_m\alpha_m f_{m\phi}s_m(t)\right|^2\right] \\ &= M^2|g|^2\alpha_0^2 G_e P_s. \end{aligned} \tag{A·7}$$

where $w_m$ are excitation weights and have an unity amplitude. Similarly, a combined interference signal at the array output is given by

$$\begin{aligned} \sum_{m=1}^{M} w_m S_{m,i}(t) &= \frac{g}{T_b}\sum_{m=1}^{M} w_m\int_0^{T_b} n(t)c_m(t)\,dt \\ &= \frac{g}{T_b}\int_0^{T_b} n(t)\sum_{m=1}^{M} w_m c_m(t)\,dt. \end{aligned} \tag{A·8}$$

Using the Schwarz's inequality and the orthogonality of the WH codes, an average power of the combined interference signal is then calculated as

$$\begin{aligned} &E\left[\left|\frac{g}{T_b}\int_0^{T_b} n(t)\sum_{m=1}^{M} w_m c_m(t)\,dt\right|^2\right] \\ &\le \frac{|g|^2}{T_b^2}\int_0^{T_b}E[|n(t)|^2]\,dt\cdot E\left[\int_0^{T_b}\left|\sum_{m=1}^{M} w_m c_m(t)\right|^2 dt\right] \\ &= \frac{|g|^2\sigma_n^2}{T_b}E\left[\int_0^{T_b}\sum_{m=1}^{M}|w_m|^2 c_m^2(t)\,dt\right] \\ &= \frac{|g|^2\sigma_n^2}{T_b}\cdot MT_b = M|g|^2\sigma_n^2. \end{aligned} \tag{A·9}$$

Using (A · 8) and (A · 9), the SNR at the array output becomes (13).

**Narihiro Nakamoto** received the B.E. and M.E. degrees in electrical engineering from Kyoto University, Kyoto, Japan, in 2005 and 2007, respectively. In 2007, he joined Mitsubishi Electric Corporation, Kanagawa, Japan. From 2010 to 2012, he was with Advanced Telecommunications Research Institute International (ATR) Wave Engineering Laboratories, Kyoto, Japan. In 2012, he returned to Mitsubishi Electric Corporation and has been engaged in research and development on array antennas for satellite communication and radar systems. He received the IEEE Antennas and Propagation Society Tokyo Chapter Young Engineer Award in 2013 and the Young Engineer Award of IEICE in 2012. He is a member of IEEE.

**Kazunari Kihira** received the B.S., M.S., and Ph.D. degrees in electrical and computer engineering from the Nagoya Institute of Technology, Japan, in 1996, 1998, and 2007, respectively. In 1998, he joined Mitsubishi Electric Corporation, Kanagawa, Japan. His current research interests include phased array ad adaptive antenna. He was a recipient of the Young Engineer Award of the IEICE in 2001 and the IEEE Antennas and Propagation Society Japan Chapter Young Engineer Award in 2005. He is a member of the IEEE.

**Toru Fukasawa** was born in Tokyo, Japan, in 1969. He received the B.E., M.E., and Ph.D. degrees in electronic engineering from Hokkaido University, Sapporo, Japan, in 1992, 1994, and 2004, respectively. In 1994, he joined the Mitsubishi Electric Corporation, Tokyo, Japan. His current research interests include small antenna analysis and measurement methods. He received the Young Engineer Award in 2000 and the Best Paper Award in 2011, 2012, 2016, and 2018 from IEICE. He is a member of IEEE.

**Yoshio Inasawa** received B.E., M.E., and Ph.D. degrees in Electrical and Electronics Engineering from the Tokyo Institute of Technology in 1991, 1993, and 2008, respectively. In 1993, he joined Mitsubishi Electric Corporation, Tokyo. Since then, he has been engaged in research on electromagnetic field analysis and reflector antennas. He is currently a manager in the Antenna Technology Department of the Information Technology Research and Development Center, Mitsubishi Electric Corporation. Dr. Inasawa received the Young Researcher's Award of IEICE Japan in 1998 and the R&D 100 Awards in 2005. He is a senior member of the IEEE.

**Naoki Shinohara** received the B.E. degree in electronic engineering, the M.E. and Ph.D (Eng.) degrees in electrical engineering from Kyoto University, Japan, in 1991, 1993 and 1996, respectively. He was a research associate in Kyoto University from 1996. From 2010, he has been a professor in Kyoto University. He has been engaged in research on Solar Power Station/Satellite and Microwave Power Transmission system. He was IEEE MTT-S Distinguish Microwave Lecturer (2016–18), and is IEEE AdCom member, IEEE MTT-S Technical Committee 25 (Wireless Power Transfer and Conversion) former chair and member, IEEE MTT-S MGA (Member Geographic Activities) Region 10 regional coordinator, IEEE WPT Initiative Member, IEEE MTT-S Kansai Chapter TPC member, IEEE Wireless Power Transfer Conference founder and Steering committee member, URSI commission D chair, international journal of Wireless Power Transfer (Hindawi) executive editor, the first chair and technical committee member on IEICE Wireless Power Transfer, Japan Society of Electromagnetic Wave Energy Applications adviser, Space Solar Power Systems Society vice chair, Wireless Power Transfer Consortium for Practical Applications (WiPoT) chair, and Wireless Power Management Consortium (WPMc) chair. His books are "Wireless Power Transfer via Radiowaves" (ISTE Ltd. and John Wiley & Sons, Inc.), "Recent Wireless Power Transfer Technologies Via Radio Waves (ed.)" (River Publishers), and "Wireless Power Transfer: Theory, Technology, and Applications (ed.)" (IET), and some Japanese textbooks of WPT.

IEICE TRANS. COMMUN., VOL.E107–B, NO.8 AUGUST 2024

PAPER

# Differential Active Self-Interference Cancellation for Asynchronous In-Band Full-Duplex GFSK

**Shinsuke IBI** [†a)], ***Senior Member***, **Takumi TAKAHASHI** [††], ***Member***, ***and*** **Hisato IWAI**[†], ***Fellow***

**SUMMARY** This paper proposes a novel differential active self-interference canceller (DASIC) algorithm for asynchronous in-band full-duplex (IBFD) Gaussian filtered frequency shift keying (GFSK), which is designed for wireless Internet of Things (IoT). In IBFD communications, where two terminals simultaneously transmit and receive signals in the same frequency band, there is an extremely strong self-interference (SI). The SI can be mitigated by an active SI canceller (ASIC), which subtracts an interference replica based on channel state information (CSI) from the received signal. The challenging problem is the realization of asynchronous IBFD for wireless IoT in indoor environments. In the asynchronous mode, pilot contamination is induced by the non-orthogonality between asynchronous pilot sequences. In addition, the transceiver suffers from analog front-end (AFE) impairments, such as phase noise. Due to these impairments, the SI cannot be canceled entirely at the receiver, resulting in residual interference. To address the above issue, the DASIC incorporates the principle of the differential codec, which enables to suppress SI without the CSI estimation of SI owing to the differential structure. Also, on the premise of using an error correction technique, iterative detection and decoding (IDD) is applied to improve the detection capability while exchanging the extrinsic log-likelihood ratio (LLR) between the maximum *a-posteriori* probability (MAP) detector and the channel decoder. Finally, the validity of using the DASIC algorithm is evaluated by computer simulations in terms of the packet error rate (PER). The results clearly demonstrate the possibility of realizing asynchronous IBFD.

***key words:*** *asynchronous in-band full-duplex, pilot contamination, GFSK, active self-interference canceller, differential encoder, iterative detection and decoding, analog front-end impairment*

## 1. Introduction

Wireless Internet of Things (IoT) is now spreading rapidly due to the demand for big databases for machine learning-aided decision-making and control in sensor networks. Classical Bluetooth, as well as Bluetooth low energy (BLE), have been widely installed in commercial products that require power-saving functionality, such as sensors, remote control units, personal computer (PC) peripherals, smartwatches, and so on [1], [2]. The Bluetooth family utilizes Gaussian frequency-shift keying (GFSK) [3], [4], which achieves high power efficiency without nonlinear distortion in the power amplifier (PA) at the transmitter. Although GFSK is a classical modulation technique, it might be vital in constructing IoT information infrastructure. This paper focuses mainly on in-band full-duplex (IBFD) mutual communications for wireless IoT.

IBFD is a technique in which two nodes (terminals) simultaneously transmit and receive signals in the same frequency band. In recent years, the study of IBFD has attracted much attention from both theoretical and practical standpoints [5]–[7]. Even in spectrum utilization for 5G mobile communication systems, a reasonable implementation method is one of the topics of current interest [8]–[10]. Compared with half-duplex (HD), such as time division duplex (TDD) and frequency division duplex (FDD), the advantage of IBFD is that a 50% reduction in wireless resource occupancy can be achieved.

When nodes A and B cooperatively perform IBFD communication, node A transmits its own signal to node B and simultaneously receives the signal transmitted from node B. In addition to the signal arriving from node B, the signal of node A itself is included in the received signal at node A as self-interference (SI). SI is usually mitigated by a combination of passive and active interference cancellers [11]–[13]. In the passive canceller, an antenna structure is designed to enhance the path loss of the SI signals [13]. On the other hand, in the active canceller, SI is eliminated as much as possible by subtracting a replica of SI from the received signal. Ideally, the replica is regenerated because its own signal is perfectly known at the receiver of node A [11], [14]. Thus, in a broad sense, IBFD is a type of physical layer network-coded communication [15], [16].

However, the receiver of IBFD suffers from severe residual SI even after the passive interference cancellation, resulting in a low signal-to-interference power ratio (SIR). An analog-to-digital converter (ADC) is an essential interface with analog baseband signals for sophisticated digital signal processing in the physical layer. A lower SIR requires a higher resolution of ADC to avoid the disappearance of the desired signal component or fatal information loss, and it is not undesirable in terms of hardware implementation [17]–[19]. To alleviate the inconvenience caused by the low SIR, it is necessary to actively cancel SI in the analog domain at the front stage of the ADC [11], [20]. Typically, the active SI cancellation capability relies on the accuracy of the self-channel state information (CSI) of SI. In other words, the channel estimation should be conducted with the aid of a pilot sequence before invoking IBFD transmissions. Therefore, the synchronous frame format of the signals and medium access control (MAC) should be appropriately designed [21], [22]. Furthermore, in [12], [23], it is shown

Manuscript received July 10, 2023.
Manuscript revised November 11, 2023.
Manuscript publicized February 16, 2024.

[†]Faculty of Science and Engineering, Doshisha University, Kyotanabe-shi, 610-0394 Japan.
[††]Graduate School of Engineering, Osaka University, Suita-shi, 565-0871 Japan.
a) E-mail: sibi@mail.doshisha.ac.jp

DOI: 10.23919/transcom.2023EBP3119

that analog front-end (AFE) impairments, such as the IQ imbalance of the IQ mixers, nonlinear distortion of the PA, and phase noise at the local oscillators (LOs), are crucial problems for the active canceller. As a result, it is subject to residual SI in the analog domain, and the residuals are suppressed by the additional active canceller in the digital domain after the ADC.

In many wireless IoT scenarios, the communication link establishment is event-driven. Nodes A and B are not always in a full-buffer state, i.e., there is a possibility not to have a data message to send, resulting in the HD mode. Nevertheless, to invoke the IBFD mode anytime, the CSI should be periodically updated at the expense of pilot overhead signals. Due to the increased overheads, IBFD is not suitable for event-driven communications. Because the amount of data is considered small, the data should be sent in HD mode instead of frequent transmission of pilot signals. Alternatively, it is also possible to estimate the CSI for every IBFD request. However, orthogonal pilot sequences should be synchronously sent from both nodes A and B. If the strength of the node B signal is low, the pilot contamination is small, even in the case of non-orthogonal pilot usage. In this paper, the target SIR ranges between $-20$ and $-10$ [dB] in indoor environments. It means relatively high strength of the node B signal, resulting in severe pilot contamination. To solve this problem, it is practical to use pilot overheads that are orthogonal to each other at nodes A and B.

In the event-driven scenario, asynchronous IBFD is preferable. Due to the asynchronous mode, it is impossible to guarantee the orthogonality of pilot sequences, resulting in "pilot contamination," which induces residual SI after the active canceller. Asynchronous IBFD has been studied in a few papers so far [24]–[27]. However, most ideas rely on interference avoidance networking in MAC protocol aspects. Therefore, these papers motivate us to consider how to realize *the effective asynchronous SI canceller in pilot contamination environments* from the viewpoint of physical layer aspects.

Motivated by the above, this paper presents a differential active SI canceller (DASIC) that incorporates the principle of the differential codec. Without AFE impairments, the DASIC is capable of eliminating SI without self-CSI estimation. As a result, the communication links are asynchronously established, making them suitable for event-driven communications. However, the DASIC yields superposed signals of plural transmitted symbols, resulting in low reception sensitivity. This issue is relaxed by maximum *a-posteriori* probability (MAP) detection [28], [29]. Under the differential structure, on the premise of using an error correction code (ECC) technique of channel coding, iterative signal detection and decoding (IDD), which improves signal detection capability while exchanging the extrinsic log-likelihood ratio (LLR) between the MAP detector and the channel decoder [30]–[33], is also designed.

As a notable approach to reduce the computational complexity involved in the detection process, neural network (NN)-aided detectors attract attention for detecting signals or channel states under unknown statistical model behavior [34], [35]. In principle, wireless environments in the training phase of network structure should be the same as the test phase. Otherwise, the mismatch of the structure deteriorates detection capability. To overcome the penalty, the deep-unfolding technique [36], [37] is useful to improve the robustness against the model error. The proposed DASIC algorithm fully captures the statistical behavior by MAP algorithm with a few states. Therefore, there are no model mismatch errors, and the NN technique is structurally unnecessary.

The contributions of this paper are summarized as follows:

1. A DASIC for blind SI cancellation without self-CSI estimations, which enables asynchronous IBFD, is proposed.
2. To improve the detection capability, IDD for the DASIC is investigated from the standpoint of the packet error rate (PER) performance.
3. The negative effects of the AFE impairments are evaluated by computer simulations.

To the best of the authors' knowledge, this is the first paper that proposes asynchronous IBFD with the aid of differential signal structure for blind SI cancellation in terms of physical layer aspects.

The rest of this paper is organized as follows. Section 2 presents a schematic of the IBFD transceiver for GFSK signaling in the context of the signal model. In Sect. 3, the DASIC algorithm for blind SI cancellation is proposed. Section 4 describes signal detection algorithms designed on the basis of observations of the DASIC output from standpoints of computational complexity and detection capability. The proposed method with and without AFE impairments is validated by computer simulations in Sect. 5. Finally, Sect. 6 concludes the paper with a summary.

The mathematical notation employed in this paper is as follows. The continuous time and discrete time indexes are denoted by $(t)$ and $[k]$, respectively. $\mathbb{R}$ and $\mathbb{C}$ represent the fields of real numbers and complex numbers, respectively. $\mathfrak{R}\{\cdot\}$ indicates the real part of complex numbers. The imaginary unit $\sqrt{-1}$ is denoted by j. $\Pr[a|b]$ and $p(a|b)$, respectively, represent the conditional probability mass function (PMF) and the probability density function (PDF) of an event $a$ given the occurrence of an event $b$. $\mathbb{E}\{\cdot\}$ indicates an expected value.

## 2. Signal Model

This section formulates a signal model for the IBFD GFSK transceiver illustrated in Fig. 1. Nodes A and B simultaneously transmit and receive signals in the same frequency channel to communicate with each other. The transmitted RF signal $s_i(t) \in \mathbb{R}$ of each A and B is expressed as

$$s_i(t) = \sqrt{2}\alpha \mathfrak{R}\left\{x_i(t)\xi_i(t)\right\}, \tag{1}$$

$$\xi_i(t) = \exp\left[\mathrm{j}(\omega_i t + \theta_i(t))\right], \tag{2}$$

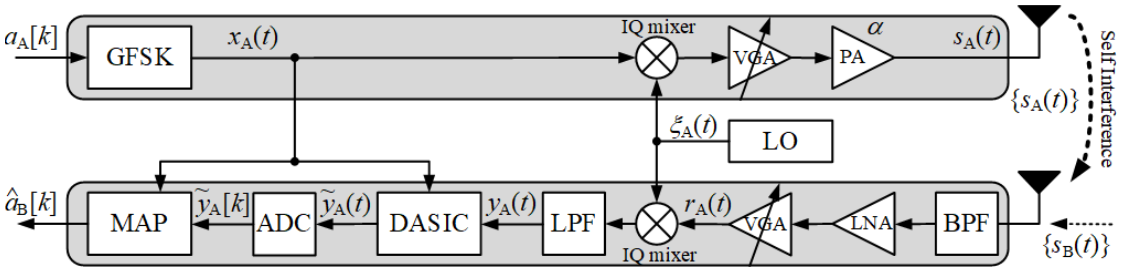



**Fig. 1** Block diagram of IBFD GFSK transceiver.

where $i \in \{\mathrm{A},\mathrm{B}\}$. $\alpha$ is the amplitude of transmitted RF signal, and $\omega_i = \omega_\mathrm{c} + \omega_i^\mathrm{o}$ is angle frequency with offset $\omega_i^\mathrm{o}$, where $\omega_\mathrm{c}$ is the center angle of frequency. More specifically, $\omega_\mathrm{A} = \omega_\mathrm{c} + \omega_\mathrm{A}^\mathrm{o}$ and $\omega_\mathrm{B} = \omega_\mathrm{c} + \omega_\mathrm{B}^\mathrm{o}$. It is noteworthy that $\omega_\mathrm{A}$ is not equal to $\omega_\mathrm{B}$. Furthermore, $\theta_i(t) = \theta_i'(t) + \theta_i^\mathrm{o}$ is the phase noise with offset $\theta_i^\mathrm{o}$, where $\theta_i'(t)$ obeys random-walk. The complex-valued baseband (CBB) signal $x_i(t) \in \mathbb{C}$ is represented by

$$x_i(t) = \exp\left[\mathrm{j}\phi_i(t)\right], \tag{3}$$

where $\phi_i(t)$ is an information-bearing variable. Denoting the energy and time duration of one symbol by $E_\mathrm{s}$ and $T_\mathrm{s}$, respectively, the amplitude is expressed as $\alpha = \sqrt{E_\mathrm{s}/T_\mathrm{s}}$. The amplitude is controlled by the variable gain amplifier (VGA) and PA with ideal characteristics.

Supposing that nodes A and B obey the same modulation rule, the information-bearing variable $\phi_i(t)$ during the interval $kT_\mathrm{s} \le t(= t' + kT_\mathrm{s}) < (k+1)T_\mathrm{s}$ is represented as

$$\phi_i(t) = 2\pi\eta \sum_{l=k-L+1}^{k} a_i[l]q(t' + (k-l)T_\mathrm{s}) + \pi\eta \sum_{l'=0}^{k-L} a_i[l'], \tag{4}$$

where $a_i[k] \in \{-1, 1\}$ is a message bit in a bipolar expression, $\eta$ is a modulation index, and $L$ is the memory size that is determined by $B_\mathrm{G}T_\mathrm{s}$. Here, $B_\mathrm{G}$ is the 3–dB sideband of the Gaussian filter. In this paper, minimum shift keying (MSK) with ($\eta = 0.5, B_\mathrm{G}T_\mathrm{s} = \infty$) and Gaussian filtered MSK (GMSK) with ($\eta = 0.5, B_\mathrm{G}T_\mathrm{s} = 0.5$) are used.

The function $q(t)$ represents a phase pulse, which is defined by [38]

$$q(t) = \int_{-\infty}^{t} g(\tau)d\tau, \tag{5}$$

under a constraint of

$$q(t) = \begin{cases} 0 & t \le 0 \\ 1/2 & t > LT_\mathrm{s} \end{cases}. \tag{6}$$

The function $g(\tau)$ is a frequency pulse given by

$$g(\tau) = \frac{1}{2T_\mathrm{s}}\left[Q\left(2\pi B_\mathrm{G}\frac{\tau - \frac{T_\mathrm{s}}{2}}{\sqrt{\ln 2}}\right) - Q\left(2\pi B_\mathrm{G}\frac{\tau + \frac{T_\mathrm{s}}{2}}{\sqrt{\ln 2}}\right)\right], \tag{7}$$

where $Q(\cdot)$ is the Q function.

In this paper, let us focus on the signal detection of node B's message at node A[†]. The received RF signal $r_\mathrm{A}(t)$ after the band-pass filter (BPF) is given by

$$\begin{aligned} r_\mathrm{A}(t) = \; & \Re\left\{h_\mathrm{AA}\sqrt{2}\alpha x_\mathrm{A}(t+\tau_\mathrm{A})\xi_\mathrm{A}(t+\tau_\mathrm{A})\right\} \\ & + \Re\left\{h_\mathrm{AB}\sqrt{2}\alpha x_\mathrm{B}(t+\tau_\mathrm{B})\xi_\mathrm{B}(t+\tau_\mathrm{B})\right\} + n_\mathrm{A}(t), \end{aligned} \tag{8}$$

where $h_\mathrm{AA} \in \mathbb{C}$ and $h_\mathrm{AB} \in \mathbb{C}$ are the complex-valued channel coefficients of the A − A and A − B links, respectively, which reflect the impacts of path loss, shadowing, and fading phenomena. $h_\mathrm{AA}$ and $h_\mathrm{AB}$ obey flat Rayleigh fading without frequency selectivity. The propagation delay time is denoted by $\tau_i$. Furthermore, $n_\mathrm{A}(t) \in \mathbb{R}$ is a noise term.

The CBB signal $y_\mathrm{A}(t)$ after low-noise amplifier (LNA), VGA, IQ mixer, and ideal low-pass filter (LPF) processes is expressed as

$$\begin{aligned} y_\mathrm{A}(t) &= \sqrt{2}\left\{r_\mathrm{A}(t)\xi_\mathrm{A}^*(t)\right\}_\mathrm{LPF} \\ &= h_\mathrm{AA}\kappa_\mathrm{AA}(t)\alpha x_\mathrm{A}(t+\tau_\mathrm{A}) \\ &\quad + h_\mathrm{AB}\kappa_\mathrm{AB}(t)\alpha x_\mathrm{B}(t+\tau_\mathrm{B}) + z_\mathrm{A}(t), \end{aligned} \tag{9}$$

where the noise term is $z_\mathrm{A}(t) = \sqrt{2}\left\{n_\mathrm{A}(t)\xi_\mathrm{A}^*(t)\right\}_\mathrm{LPF}$, which is modeled by complex-valued Gaussian noise with zero mean and $\beta \triangleq N_0/T_\mathrm{s}$ variance. Note that $N_0$ is the noise energy for one symbol's time duration, whereas $\beta$ is the noise density. $\kappa_{\mathrm{A}i}(t)$, ($i \in \{\mathrm{A},\mathrm{B}\}$), represents LO impairments of both nodes, which is given by

$$\kappa_{\mathrm{A}i}(t) = \xi_i(t+\tau_i)\xi_\mathrm{A}^*(t) = \kappa_{\mathrm{A}i}'(t)\kappa_{\mathrm{A}i}^\mathrm{o}, \tag{10}$$

$$\kappa_{\mathrm{A}i}'(t) = \exp\left[\mathrm{j}(\theta_i'(t+\tau_i) - \theta_\mathrm{A}'(t))\right], \tag{11}$$

$$\kappa_\mathrm{AA}^\mathrm{o} = 1, \tag{12}$$

$$\kappa_\mathrm{AB}^\mathrm{o} = \exp\left[\mathrm{j}(\omega_\mathrm{B}^\mathrm{o} - \omega_\mathrm{A}^\mathrm{o})\right]\exp\left[\mathrm{j}(\theta_\mathrm{B}^\mathrm{o} - \theta_\mathrm{A}^\mathrm{o})\right]. \tag{13}$$

In IBFD systems, the longer distance between nodes A and B induces a stronger SI because $|h_\mathrm{AA}|^2$ is much higher than $|h_\mathrm{AB}|^2$. Now, let us denote the SIR by

$$\zeta = \frac{|h_\mathrm{AB}|^2}{|h_\mathrm{AA}|^2}. \tag{14}$$

In principle, the signal components of $h_\mathrm{AB}x_\mathrm{B}(t)$ disappear or deteriorate after the ADC while it is subject to the huge difference between $|h_\mathrm{AA}|^2$ and $|h_\mathrm{AB}|^2$, i.e., extremely low $\zeta$. Thus, the SI of the first term of the right-hand side in

[†]The signal detection of node A's message at node B is readily explained by replacing the variables.

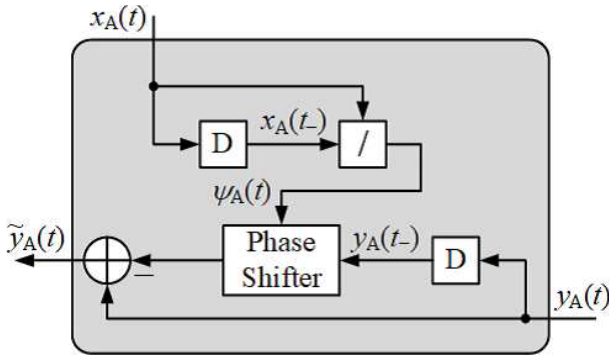



**Fig. 2** Block diagram of the DASIC process.

Eq. (8) should be canceled before the ADC process, i.e., in the analog domain. Assuming that $h'_{AA} \triangleq h_{AA}\alpha\xi_A(t+\tau_A)$ is accurately estimated at the receiver, a typical active SI canceller (ASIC) ideally subtracts the replica of the SI from the received signal as

$$r_A(t) - \Re\left\{h'_{AA}x_A(t+\tau_A)\right\}. \tag{15}$$

Instead of using an ASIC, this study employs a DASIC for blindly canceling the SI in a regime of differential detection without knowledge of $h_{AA}\alpha\xi_A(t+\tau_A)$. The CBB signal $\tilde{y}_A(t)$ is extracted from the output of the DASIC. Finally, the MAP detector yields the estimates $\hat{a}_B[k]$ of node B's message.

## 3. Principle of DASIC

A schematic of the DASIC is illustrated in Fig. 2, which is intended as an analog circuit. It is assumed that an analog multiplier is used for division and a mixer is used for phase adjustment; hence, a significant increase in power consumption is not expected owing to the use of passive elements. The problem arising here is the realization of a delay of one symbol duration of an analog signal. This paper assumes that this delay can be achieved ideally. There are still many issues to be solved in the implementation of analog circuits, which we regard as a future issue[†].

More specifically, by observing the CBB signal $x_A(t)$, a phase shift is generated as

$$\psi_A(t) = \frac{x_A(t)}{x_A(t-T_s)}. \tag{16}$$

Then, assuming an ideal delay circuit, the phase of a delayed CBB signal $y_A(t-T_s)$ is shifted by $\psi_A(t)$ at the phase shifter. Denoting $t_- = t - T_s$ to simplify the mathematical notations, we have the modified differential signal, which is represented as

$$\begin{aligned} y_A(t_-)\psi_A(t) &= h_{AA}\alpha x_A(t) + \varepsilon_A(t) \\ &\quad + h_{AB}\kappa_{AB}(t)\alpha x_B(t_- + \tau_B)\psi_A(t) \\ &\quad + z_A(t_-)\psi_A(t), \end{aligned} \tag{17}$$

where

$$\varepsilon_A(t) = h_{AA}\alpha\left[\kappa_{AA}(t)x_A(t+\tau_A) - x_A(t)\right]. \tag{18}$$

Equations (10)–(12) indicate that $\varepsilon_A(t)$ is zero if the propagation delay $\tau_A$ of SI is zero. In other words, the additional noise component $\varepsilon_A(t)$ is induced by the propagation delay time $\tau_A$.

Eventually, the analog interference canceller is operated as

$$\begin{aligned} \tilde{y}_A(t) &= y_A(t) - y_A(t_-)\psi_A(t) \\ &= \tilde{h}_{AB}\alpha\tilde{x}_B(t+\tau_B) + \tilde{\varepsilon}_A(t) + \tilde{z}_A(t), \end{aligned} \tag{19}$$

where we have

$$\tilde{h}_{AB} = h_{AB}\kappa^{o}_{AB}, \tag{20}$$

$$\tilde{x}_B(t+\tau_B) = x_B(t+\tau_B) - \psi_A(t)x_B(t_- + \tau_B), \tag{21}$$

$$\begin{aligned} \tilde{\varepsilon}_A(t) &= \varepsilon_A(t) + \tilde{h}_{AB}[\kappa'_{AB}(t) - 1]\alpha x_B(t+\tau_B) \\ &\quad - \tilde{h}_{AB}[\kappa'_{AB}(t_-) - 1]\psi_A(t)\alpha x_B(t_- + \tau_B), \end{aligned} \tag{22}$$

$$\tilde{z}_A(t) = z_A(t) - \psi_A(t)z_A(t_-). \tag{23}$$

$\tilde{\varepsilon}_A(t)$ denotes the residual SI after the DASIC process that is induced by the AFE impairments. Eq. (22) implies that the residual SI $\tilde{\varepsilon}_A(t)$ is zero only if $\theta'_A(t) = \theta'_B(t+\tau_B)$ and $\tau_A = 0$. More specifically, if $\theta'_A(t) = \theta'_B(t+\tau_B)$, $\kappa'_{AB}(t)$ of Eq. (11) is one for any $t$, resulting in $\tilde{\varepsilon}_A(t) = \varepsilon_A(t)$ from Eq. (22). On the other hand, if $\tau_A = 0$, we derive $\kappa_{AA}(t) = 1$ from Eqs. (10)–(12), resulting in $\varepsilon_A(t) = 0$ from Eq. (18). As the result, we have $\tilde{\varepsilon}_A(t) = 0$.

Figure 3 shows eye diagrams of the DASIC output for MSK and GMSK ($B_G T_s = 0.5$) signaling in an ideal case of $\tau_A = \tau_B = 0$ without AFE impairments. For the visualization, the noise term $z_A(t)$ is zero and $\tilde{h}_{AB}\alpha = 1$. In both cases, the eye apertures are maximized every $2(k+1)T_s$ and $2kT_s$ in the real and imaginary amplitudes, respectively. However, in GMSK, due to the presence of inter-symbol interference (ISI) induced by Gaussian filtering, the aperture is reduced.

Assuming the ideal sampling time at $t = kT_s + \delta$ that maximizes the eye aperture, the ADC of Eq. (19) yields

$$\tilde{y}_A[k] = \tilde{h}_{AB}\alpha\tilde{x}_B[k] + \tilde{\varepsilon}_A[k] + \tilde{z}_A[k], \tag{24}$$

where we have

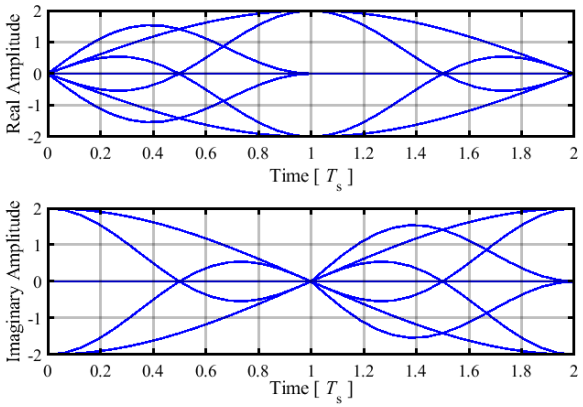
$$\tilde{x}_B[k] = x_B[k] - \psi_A[k]x_B[k-1], \tag{25}$$

$$\tilde{z}_A[k] = z_A[k] - \psi_A[k]z_A[k-1]. \tag{26}$$

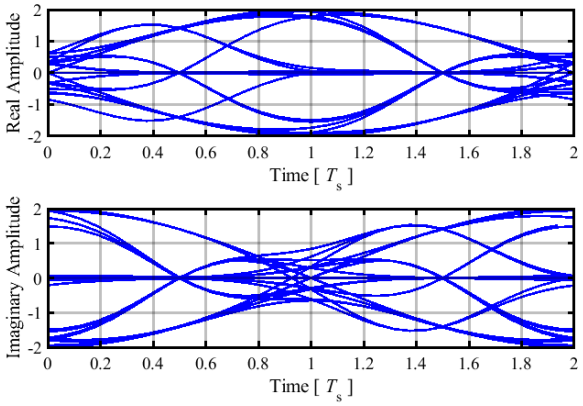
Equation (24) implies that $\tilde{y}_A[k]$ is represented by a simple linear model where $\tilde{x}_B[k]$ is given by the superposed CBB of $x_B[k]$ and $x_B[k-1]$.

The second moments of $\tilde{x}_B[k]$ and $\tilde{z}_A[k]$ are given by $\mathbb{E}\left\{|\tilde{x}_B[k]|^2\right\} = 2\mu$, $(0 \le \mu \le 1)$ and $\mathbb{E}\left\{|\tilde{z}_A[k]|^2\right\} = 2\beta = 2\frac{N_0}{T_s}$, respectively. Figure 4 characterizes equivalent signal density $\mathbb{E}\left\{|\tilde{x}_B[k]|^2\right\} = 2\mu$ at the continuous-time $t$ according to the delay time of the node B $\tau_B$. The value of $\mu$ is

[†]If the analog circuitry of the DASIC is difficult to implement from the cost perspective, Fig. 2 can be implemented in the digital domain as an alternative. In this case, the ADC shown in Fig. 1 is placed in front of the DASIC and can operate at the same sampling rate as the symbol rate of the MAP detector.

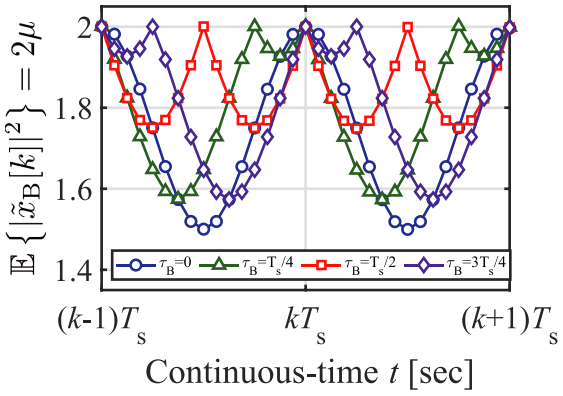



(a) MSK: $B_G T_s = \infty$.

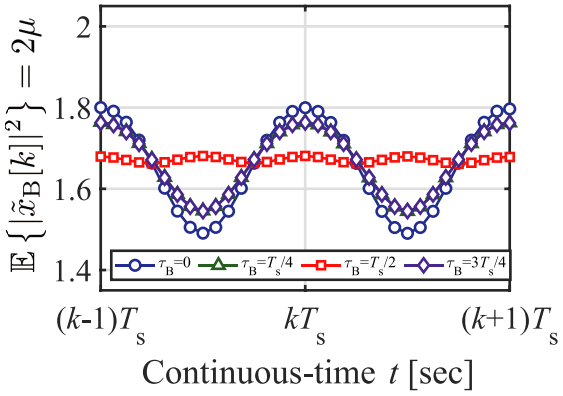



(b) GMSK: $B_G T_s = 0.5$.

**Fig. 3** Eye diagrams of the DASIC output $\tilde{y}_A(t)$ for MSK and GMSK: $\eta = 0.5$ where $\tilde{h}_{AB}\alpha = 1$ without AFE impairments.




(a) MSK: $B_G T_s = \infty$.




(b) GMSK: $B_G T_s = 0.5$.

**Fig. 4** Equivalent signal density $\mathbb{E}\{|\tilde{x}_B[k]|^2\} = 2\mu$ at the continuous time $t$ according to the delay time of the node B $\tau_B$.

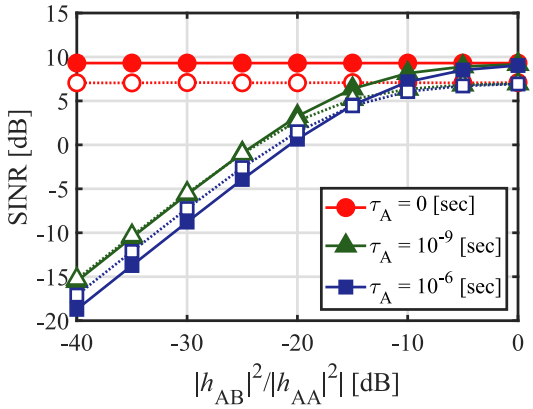
maximized at the timing $kT_s$ by adjusting the sample timing $\delta$ of ADC. As shown in (a), MSK can keep $\mu = 1$ if $\delta$ is appropriate. On the other hand, GMSK with $B_G T_s = 0.5$ in (b) is subject to the energy loss caused by the delay time of the node B $\tau_B$. The maximum loss from the ideal delay time $\tau_B = 0$ is $10\log_{10}\frac{1.80}{1.68} = 0.3$ [dB] at $\tau_B = T_s/2$. Note that the curves of $\tau_B = T_s/4$ and $3T_s/4$ are almost the same because of the phase-circularity. In the case of asynchronous IBFD, structurally, $\tau_B$ cannot maintain at zero. However, the maximum loss is only 0.3 [dB]; thus, the negative effect is not significant. From this point forward, we assume $\tau_B = 0$ for ease of discussions in this paper.

The signal-to-interference plus noise power ratio (SINR) of the DASIC output is given by

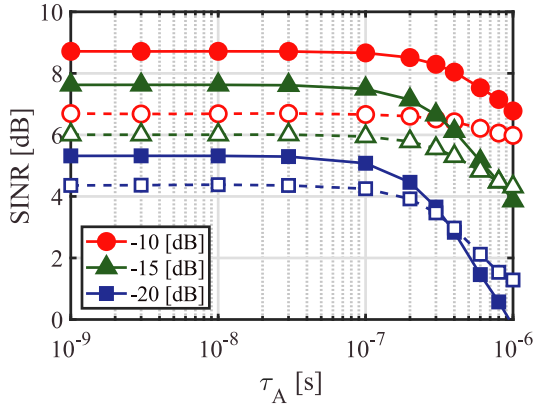
$$\rho = \frac{\mathbb{E}\left\{\left|\tilde{h}_{AB}\alpha\tilde{x}_B[k]\right|^2\right\}}{\mathbb{E}\left\{|\tilde{\varepsilon}_A[k] + \tilde{z}_A[k]|^2\right\}} = \frac{2\left|\tilde{h}_{AB}\right|^2\alpha^2\mu}{\mathbb{E}\left\{|\tilde{\varepsilon}_A[k]|^2\right\} + 2\beta} = \frac{1}{\frac{\mathbb{E}\{|\tilde{\varepsilon}_A[k]|^2\}}{2|\tilde{h}_{AB}|^2\alpha^2\mu} + \left(\mu|h_{AB}|^2\frac{E_s}{N_0}\right)^{-1}}. \quad (27)$$

To demonstrate the validity of the DASIC, in terms of the interference suppression capability, Fig. 5 exhibits the SINR $\rho$ property of the DASIC output against $\zeta = |h_{AB}|^2/|h_{AA}|^2$ and $\tau_A$. As an example, the instantaneous $|h_{AB}|^2\frac{E_s}{N_0}$ is 10 [dB]. The phase noise at the LOs of the transmitter and receiver is emulated by a model of the phased lock loop (PLL) based CMOS frequency synthesizer with the power spectrum density (PSD). In accordance with the requirements for Bluetooth applications, the PSD is defined as $-89$ [dBc/Hz] at 500 [kHz] and $-121$ [dBc/Hz] at 2 [MHz]. The symbol rate $1/T_s$ is 1 [MHz]. In Fig. 5(a), the delay time is $\tau_A \in \{0, 10^{-9}, 10^{-6}\}$ [sec]. In Fig. 5(b), the value of $\zeta = |h_{AB}|^2/|h_{AA}|^2$ is $\{-20, -15, -10\}$ [dB]. As a comparison with the traditional IBFD, the performance of the ASIC is depicted in the figure. In the ASIC, the channel state of $h_{AA}\alpha\xi_A(t+\tau_A)$ is estimated by the least square (LS) criterion with the aid of a 16-bit pilot sequence. Note that the 16-bit pilot sequences of nodes A and B are not orthogonal due to asynchronous IBFD.

Let us focus on Fig. 5(a). In the case of $\tau_A = 0$ [sec], the phase noise behavior can be completely eliminated by the IQ mixer in the receiver of node A. The ASIC suffers from pilot contamination due to asynchronous IBFD, resulting in lower SINR due to residual SI. In the case of $\tau_A = 10^{-9}$ [sec] = 1 [nsec], the SINR is severely degraded if $\zeta$ is low. This behavior makes it difficult to implement IBFD. If $\tau_A = 10^{-6}$ [sec] = 1 [$\mu$sec], the SINR of the DASIC is worse than that of the ASIC. This is because the differential




(a) SINR vs. $\zeta = |h_{AB}|^2/|h_{AA}|^2$.




(b) SINR vs. $\tau_A$: The values in the legend indicate $\zeta = |h_{AB}|^2/|h_{AA}|^2$.

**Fig. 5** SINR with phase noise impairments: Dashed and solid lines are of the ASIC and DASIC, respectively.

structure is broken. However, because the communication distance is short, the probability that $\tau_A = 1$ [$\mu$sec] will occur is low. Next, let us move our focus to Fig. 5(b). The SINR of DASIC is higher than ASIC when $\tau_A$ is shorter than 300 [nsec] when $\zeta = |h_{AB}|^2/|h_{AA}|^2 = -20$ [dB] of the lowest target in this paper. The figure emphasizes the importance of phase noise behavior in IBFD considerations.

## 4. MAP Detector for DASIC Output

As an example, let us consider GFSK signaling with $B_G T_s$, which can be fully captured by a memory size of $L = 3$. Denoting a weight factor as $w_l \triangleq 2\pi\eta q(\Delta t + lT_s)$, $(l \in \{0,1,2\})$ for Gaussian spectrum shaping, the components $x_B[k]$ and $x_B[k-1]$ in the superposed CBB signal $\tilde{x}_B[k]$ of Eq. (25) are expressed as

$$\begin{aligned} x_B[k] = &\exp\{\mathrm{j}(w_0 a_B[k] + w_1 a_B[k-1] \\ &+ w_2 a_B[k-2]) + \mathrm{j}\sigma[k]\}, \end{aligned} \tag{28}$$

$$\begin{aligned} x_B[k-1] = &\exp\{\mathrm{j}(w_0 a_B[k-1] + w_1 a_B[k-2]) \\ &+ \mathrm{j}\sigma[k]\}, \end{aligned} \tag{29}$$

where $\sigma[k]$ denotes a state that is defined by

$$\sigma[k] = \pi\eta \sum_{l'=0}^{k-3} a_B[l']. \tag{30}$$

On the other hand, the phase shift $\psi_A[k]$ in Eq. (25) is given by

$$\begin{aligned} \psi_A[k] = &\exp\{\mathrm{j}w_0(a_A[k] - a_A[k-1]) \\ &+ \mathrm{j}w_1(a_A[k-1] - a_A[k-2]) \\ &+ \mathrm{j}\pi\eta a_A[k-2]\}. \end{aligned} \tag{31}$$

### 4.1 MSK Signaling

For an intuitive understanding of the structure of the superposed CBB signal, at first, let us consider MSK signaling, i.e., $\eta = 0.5$ and $B_G T_s = \infty$. In this case, the principal values, in terms of the exponential function, of $\sigma[k]$ are one of four states: $\mathcal{U} = \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$. When the ADC samples at an appropriate timing $\Delta t$ that maximizes the eye aperture, the weight factors are $w_0 = 0$ and $w_1 = w_2 = \pi/2$, resulting in

$$x_B[k] = \exp\left\{\mathrm{j}\frac{\pi}{2}(a_B[k-1] + a_B[k-2]) + \mathrm{j}\sigma[k]\right\}, \tag{32}$$

$$x_B[k-1] = \exp\left\{\mathrm{j}\frac{\pi}{2}a_B[k-2] + \mathrm{j}\sigma[k]\right\}, \tag{33}$$

$$\psi_A[k] = \exp\left\{\mathrm{j}\frac{\pi}{2}a_A[k-1]\right\}. \tag{34}$$

Now, denoting $\delta[k] = a_B[k] - a_A[k] \in \{-2, 0, 2\}$, Eq. (25) is rewritten as

$$\begin{aligned} \tilde{x}_B[k] &= \exp\left\{\mathrm{j}\frac{\pi}{2}(a_B[k-1] + a_B[k-2]) + \mathrm{j}\sigma[k]\right\} \\ &\quad - \exp\left\{\mathrm{j}\frac{\pi}{2}(a_B[k-1] + a_B[k-2]) + \mathrm{j}\sigma[k]\right\} \\ &= \mathrm{j}\delta[k-1]\exp\left\{\mathrm{j}\frac{\pi}{2}a_B[k-2] + \mathrm{j}\sigma[k]\right\} \\ &= \begin{cases} 2\exp[\mathrm{j}u_A[k]], & (\delta[k-1] \neq 0) \\ 0, & (\delta[k-1] = 0) \end{cases}, \end{aligned} \tag{35}$$

where we have

$$u_A[k] = \frac{\pi}{2}\left(a_B[k-2] + \frac{\delta[k-1]}{2}\right) + \sigma[k] \in \mathcal{U}. \tag{36}$$

Equation (35) implies that the superposed CBB $\tilde{x}_B[k+1]$ is always zero if node B's information $a_B[k]$ is identical to node A's information $a_A[k]$. Otherwise, $\tilde{x}_B[k+1]$ belongs to one of the states $\mathcal{U}$.

An example of a scatter diagram of the DASIC outputs $\tilde{y}_A[k]$ is depicted in Fig. 6. The Euclidean distance from the DASIC output $\tilde{y}_A[k]$ to each of the five states $\mathcal{X} = \{0, 2\exp[\mathrm{j}\mathcal{U}]\} = \{0, 2, 2\mathrm{j}, -2, -2\mathrm{j}\}$ obeys the complex-valued Gaussian distribution with zero mean and $2\beta$ variance. When the occurrences of $a_A[k] = \pm 1$ and $a_B[k] = \pm 1$ are equiprobable, the occurrence probability of $\tilde{x}_B[k]$ is given by

$$\Pr[\tilde{x}_B[k] = 0] = 1/2, \tag{37}$$

$$\Pr[\tilde{x}_B[k] = \{2, 2\mathrm{j}, -2, -2\mathrm{j}\}] = 1/8. \tag{38}$$

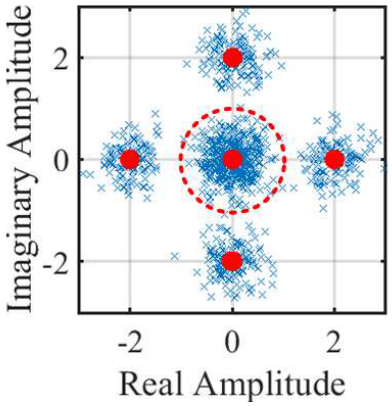



**Fig. 6** Scatter plot of the DASIC output $\tilde{y}_A[k]$ where $\tilde{h}_{AB}\alpha = 1$: $\times$ and $\bullet$ indicate the scatter of $\tilde{y}_A[k]$ and $\tilde{x}_B[k] \in \{0, 2, 2j, -2, -2j\}$, respectively.

Therefore, the threshold value for simple hard decision is set to $\pm 1$, which is the stochastic midpoint. According to (35), a criterion for the simple hard decision is formulated as

$$\hat{a}_B[k-1] = \begin{cases} a_A[k-1], & |\tilde{x}_B[k]| < |\tilde{h}_{AB}|\alpha \\ -a_A[k-1], & \text{others} \end{cases}. \quad (39)$$

The bits of nodes A and B are the same ($a_A[k-1] = a_B[k-1]$), $|\tilde{x}_B[k]|$ is inside the circle with radius $|\tilde{h}_{AB}|\alpha$. Otherwise, $|\tilde{x}_B[k]|$ is outside the circle. This method does not request phase information of $h_{BA}$ and the computational complexity is extremely low. However, probabilistic behavior cannot be captured.

For optimal detection in Gaussian channels, the information bit $a_B[k]$ coming from node B to node A is derived by finding the maximum *a-posteriori* probability (MAP) as

$$\hat{a}_B[k] = \arg \max_{a \in \{\pm 1\}} \Pr[a_B[k] = a | \tilde{y}_A[k+1]]. \quad (40)$$

Recalling the relationship discussed in Eq. (35), with the aid of perfect knowledge on the own message $a_A[k]$, the constituent *a-posteriori* probability is obtained by

$$\begin{aligned} &\Pr[a_B[k] = a_A[k] | \tilde{y}_A[k+1]] \\ &\quad = \Pr[\tilde{x}_B[k+1] = 0 | \tilde{y}_A[k+1]], \end{aligned} \quad (41)$$

$$\begin{aligned} &\Pr[a_B[k] = -a_A[k] | \tilde{y}_A[k+1]] \\ &\quad = \Pr[\tilde{x}_B[k+1] \neq 0 | \tilde{y}_A[k+1]], \end{aligned} \quad (42)$$

where

$$\begin{aligned} &\Pr[\tilde{x}_B[k+1] \neq 0 | \tilde{y}_A[k+1]] \\ &= \sum_{x \in \{2,2j,-2,-2j\}} \Pr[\tilde{x}_B[k+1] = x | \tilde{y}_A[k+1]]. \end{aligned} \quad (43)$$

On the basis of Bayes' theorem, the *a-posteriori* probability is rewritten as

$$\begin{aligned} &\Pr[\tilde{x}_B[k] = x | \tilde{y}_A[k]] \\ &\quad = \frac{p(\tilde{y}_A[k] | \tilde{x}_B[k] = x) \Pr[\tilde{x}_B[k] = x]}{p(\tilde{y}_A[k])}, \end{aligned} \quad (44)$$

where $p(\tilde{y}_A[k] | \tilde{x}_B[k] = x)$ is the PDF given by

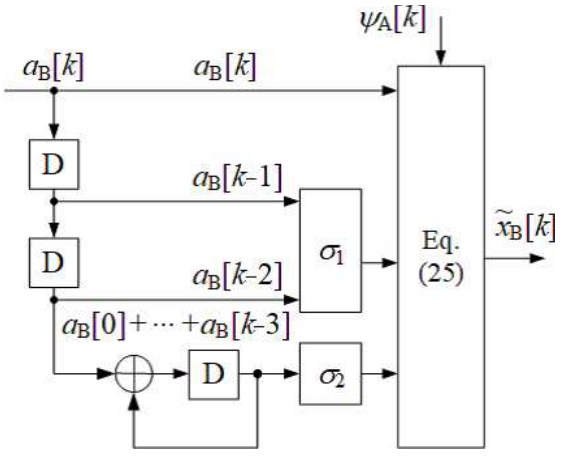



**Fig. 7** Structure of IIR filter of signal component $\tilde{x}_B[k]$ included in DASIC output $\tilde{y}_A[k]$.

$$\begin{aligned} &p(\tilde{y}_A[k] | \tilde{x}_B[k] = x) \\ &\quad = \frac{1}{2\pi\beta} \exp\left[ -\frac{\left| \tilde{y}_A[k] - \tilde{h}_{AB}\alpha x \right|^2}{2\beta} \right]. \end{aligned} \quad (45)$$

Note that $p(\tilde{y}_A[k])$ is a constant value in terms of the search problem of the maximum value in Eq. (40). Thus, we may replace $p(\tilde{y}_A[k])$ with unity.

### 4.2 GMSK Signaling

In GMSK signaling, we must pay attention to the fact that we should take into account the infinite impulse response (IIR) filter structure of Eq. (4) to fully capture the amount of mutual information between $\tilde{y}_A[k]$ and $a_B[k]$. Here, let us investigate the MAP detector for the IIR-filtered signal, which can achieve optimal performance in terms of probabilistic theory. According to the MAP framework, the *a-posteriori* LLR is defined by

$$\lambda_B[k] = \ln \frac{\Pr[a_B[k] = +1 | \{\tilde{y}_A[k]\}]}{\Pr[a_B[k] = -1 | \{\tilde{y}_A[k]\}]}, \quad (46)$$

where $\{\tilde{y}_A[k]\}$ indicates a sequence of $\tilde{y}_A[k]$. To compute the LLR $\lambda_B[k]$ with reasonable computational efforts, we focus on the IIR filter structure of Eq. (4).

Figure 7 illustrates the filter structure using memory size $L = 3$ as an example. On the basis of the trellis diagram depicted from Fig. 7, the LLR can be obtained by the Bahl–Cocke–Jelinek–Raviv (BCJR) algorithm [28] without requiring a large computational burden†. Assuming BCJR is applied for GFSK demodulation in both ASIC and DASIC, the complexity relies on the number of states in the BCJR

†Although this paper assumes the identical GFSK modulation in nodes A and B, the same DASIC technique can be used for SI cancellation when changing to other modulation schemes. However, note that changes in the combination of modulation schemes complicate the trellis structure, which causes changes in the error correction effect in the MAP signal detector, as well as in the iterative gain for the IDD regime.

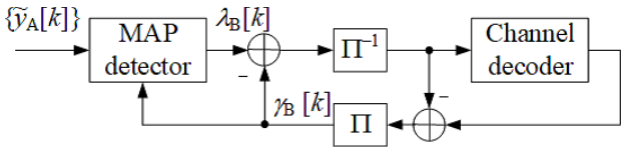



**Fig. 8** Schematic of the IDD receiver.

algorithm. In this paper, $L = 3$ is used for $B_G T_s = 0.5$ pulse shaping of both ASIC and DASIC. If $B_G T_s$ is smaller than 0.5, the $L$ should be a larger value than 3, resulting in a large number of states. In this paper, the number of states is negligibly small, i.e., $4^{L-1} = 16$ states. It is equivalent to binary convolutional code with constraint length 4. Thus, the complexity is not so high.

Here, one element of the Cartesian product set (four states) of $a_B[k-1] = \{\pm 1\}$ and $a_B[k-2] = \{\pm 1\}$ is represented by $\sigma_1$. In addition, one element of the four-state set $\mathcal{U}$ is represented by $\sigma_2$. From these states $\sigma_1$ and $\sigma_2$ as well as the DASIC input $a_B[k]$, the superposed CBB signal $\tilde{x}_B[k]$ can be reproduced, and the input/output relationship is denoted by $\tilde{x}_B[k] = f(a_B[k], \sigma_1, \sigma_2)$. The transition (joint) probability required in the BCJR process is given by

$$\begin{aligned}&\Pr\left[a_B[k] = \{\pm 1\}, \sigma_1, \sigma_2 | \tilde{y}_A[k]\right] \\ &\propto \exp\left\{-\frac{\left|\tilde{y}_A[k] - \tilde{h}_{AB}\alpha f(a_B[k], \sigma_1, \sigma_2)\right|^2}{2\beta}\right\} \\ &\quad \cdot \exp\left\{a_B[k]\frac{\gamma_B[k]}{2}\right\}, \end{aligned} \tag{47}$$

where $\gamma_B[k]$ is the *a-priori* LLR, which represents the confidence in the prior knowledge of $a_B[k]$, and it is defined by

$$\gamma_B[k] = \ln \frac{\Pr[a_B[k] = +1]}{\Pr[a_B[k] = -1]}. \tag{48}$$

Note that $\gamma_B[k]$ is zero if there is no prior information.

### 4.3 IDD Receiver

If ECC and an interleaver ($\Pi$) are used at the transmitter side, it is possible to perform IDD, in which signal detection is achieved by exchanging extrinsic LLRs between the MAP detector and the error correction decoder [31].

Figure 8 shows the schematic of the IDD receiver. With the aid of the *a-priori* LLR $\gamma_B[k]$, Eq. (46) calculates the *a-posteriori* LLR $\lambda_B[k]$. Subsequently, the extrinsic LLR is obtained by subtracting the *a-priori* LLR $\gamma_B[k]$ from the *a-posteriori* LLR $\lambda_B[k]$. The resulting extrinsic LLR is deinterleaved ($\Pi^{-1}$) and applied to the channel decoder. The extrinsic LLR is updated at the channel decoder after the error correction process. The updated LLR is fed back to the MAP detector via the interleaver as the *a-priori* LLR $\gamma_B[k]$. After repeating these LLR exchanges, the decision value of the information bit is obtained from the channel decoder.

Note that the computational complexity of the iteration process linearly increases according to the allowed number of iterations. Thus, the number of states in BCJR is a dominating factor in terms of the complexity of ASIC and DASIC.

**Table 1** Simulation parameters.

| | |
|---|---|
| Modulation | GMSK ($\eta = 0.5$, $B_G T_s = 0.5$, $L = 3$) |
| Detector | Log-MAP (Jacobian-log) |
| Packet length | 1024 symbols |
| Channel encoder | None, Convolutional code |
| Coding rate | 1, 1/2 (Constraint length = 4) |
| Decoder | None, Log-MAP (Jacobian-log) |
| Num. of iterations | 1, 16 |
| Interleaver | Random |
| Channel model | AWGN + Rayleigh fading |

### 4.4 Channel Estimations

To carry out the MAP detection of Eqs. (40) and (46), the values $h_{AB}$ and $\beta$ are estimated from the preamble. Now, we assume that an $M$-bit pilot sequence is appended in front of the data sequence. Based on the LS criterion, the estimated values can be obtained by

$$\hat{h}_{AB} = \frac{\sum_{k=1}^{M} \tilde{y}_A[k]\tilde{x}_B[k]^*}{\alpha \sum_{k=1}^{M} |\tilde{x}_B[k]|^2}, \tag{49}$$

$$\hat{\beta} = \frac{1}{M}\sum_{k=1}^{M} \frac{|\tilde{y}_A[k]|^2}{\alpha^2} - \left|\hat{h}_{AB}\right|^2. \tag{50}$$

As shown in Eq. (24), the signal model after SI cancellation is characterized by the single-input single-output (SISO) aditive white Gaussian noise (AWGN) model, assuming that random components $\tilde{\varepsilon}_A[k] + \tilde{z}_A[k]$ obey Gaussian distribution. In this simple signal model, the typical channel estimation is LS estimation. If long-term statistics of the channel state $\mathbb{E}\{|h_{AB}|^2\}$ are available, MAP estimation with high accuracy is feasible even in the channel estimation, but such long-term statistics are not available because we assume random access in the IoT scenario. Iterative channel estimation is a promising technique for IDD structures. However, owing to space limitations in this paper, we do not evaluate the effect of iterative channel estimation in simulations.

## 5. Numerical Results

Computer simulations were conducted to verify the effectiveness of the proposed DASIC. The simulation conditions are summarized in Table 1. The modulation scheme is GMSK ($\eta = 0.5$, $B_G T_s = 0.5$, $L = 3$), and the MAP algorithm based on the Jacobian logarithm is used for its demodulation. At this time, to cope with the tail bit problem of the MAP demodulator, dummy bits of $L - 1$ length are appended after the information data block. In the non-coded scenario without ECC, any channel code is not applied, resulting in 1024 data bits in one packet with 1024 symbols. On the other hand, in the coded scenario with the aid of ECC, half-rate non-systematic convolutional (NSC) code with Generator polynomials $G_0(x) = 1 + x + x^2 + x^3$ and $G_1(x) = 1 + x^2 + x^3$ is utilized. The code length is 1024 coded bits, i.e., the

information length is 509 bits and the tail length for state termination of the convolutional code is 3 bits. Its decoding employs the MAP algorithm. In the case of IDD, the number of iterations for LLR exchange between the MAP detector and the channel decoder is 16. At this time, we used a random permutation in the interleaver. The channel model assumes the AWGN channel is affected by flat Rayleigh fading. IQ imbalance is assumed to be negligibly small in this paper.

As a comparison with the traditional IBFD, i.e., a baseline scheme, the performance of the ASIC with the aid of CSI is evaluated. $L = 3$ is used for both ASIC and DASIC. Thus, the dominating complexity order of BCJR is the same. When IDD is applied in DASIC, the complexity linearly increases according to the allowed number of iterations.

### 5.1 Without Phase Noise Impairments

Figure 9 shows the PER characteristics of the DASIC without phase noise impairments for $|h_{\mathrm{AB}}|^2 E_{\mathrm{s}}/N_0$. Although the channel obeys Rayleigh fading, the horizontal axis is the instantaneous $|h_{\mathrm{AB}}|^2 E_{\mathrm{s}}/N_0$, not the average. As a result, it is equivalent to the channel affected by phase rotation without amplitude fluctuations.

Figure 9(a) presents the detection capability when the channel states and noise variance are ideally estimated. Thus, $\tilde{h}_{\mathrm{AB}}$ and $\beta$ are perfectly known at node A in the DASIC. On the other hand, $h_{\mathrm{AA}}$, $h_{\mathrm{AB}}$, and $N_0$ are perfectly captured in the ASIC. Three types of PER are evaluated for each DASIC and ASIC. The first and second types are without and with the channel encoder ("w/o ECC," "w/ ECC"). The last type of PER is IDD with the aid of LLR exchange between the MAP detector and the decoder. As can be seen in the figure, the ASIC achieves a gain of approximately 2 [dB] compared to the DASIC for all types of PER. However, this comparison is unfair because the ASIC can completely eliminate SI in this case. It is obvious that the DASIC is not useful in the ideal scenario.

Now, let us shift the focus to Figs. 9(b) and (c), which show the PER when the channel and noise variance are estimated by 32- and 16-bit pilot sequences, respectively. Due to the pilot contamination induced by asynchronous IBFD between nodes A and B, a longer pilot is necessary for the ASIC to suppress channel estimation errors enhanced by the non-orthogonal pilot sequences. Figure 9(b) shows the impairments caused by the pilot contamination. Even if the 32-bit pilot is available, the slope of the PER curves of the ASIC is low. As a result, DASIC outperforms ASIC at the

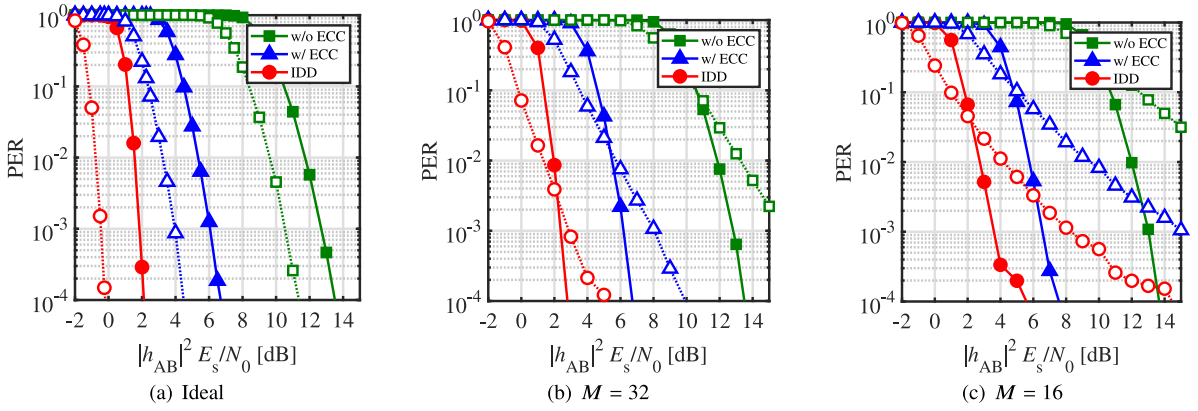



**Fig. 9** PER performance without phase noise impairments ($\zeta = -20$ [dB]): dashed and solid lines are of the ASIC and DASIC, respectively.

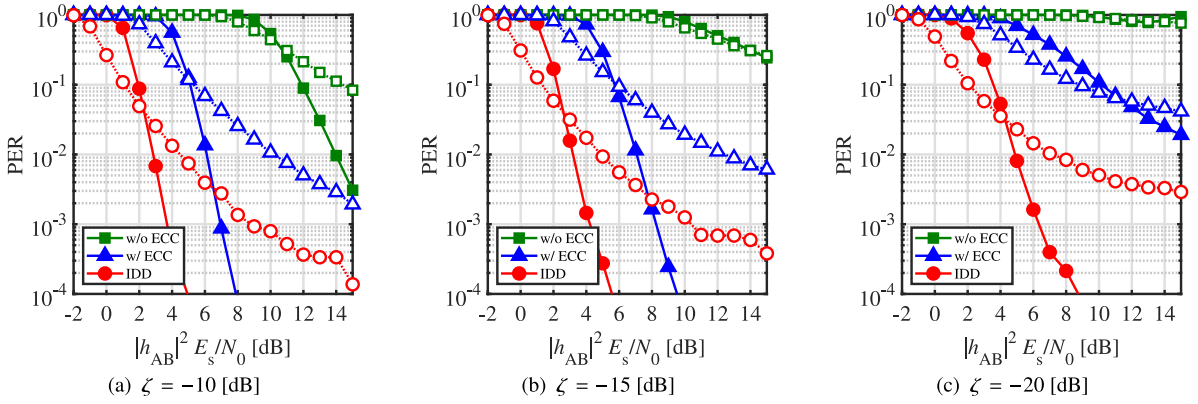



**Fig. 10** PER performance with phase noise impairments ($\tau_{\mathrm{A}} = 1$ [nsec], $M = 16$): dashed and solid lines are of the ASIC and DASIC, respectively.

lower PER target. Furthermore, when the pilot length is shorter, such as $M = 16$, as shown in Fig. 9(c), we can see that the DASIC has large advantages. The results explicitly demonstrate that the proposed DASIC algorithm is effective for the pilot contamination problem in asynchronous IBFD.

### 5.2 With Phase Noise Impairments

As discussed in Fig. 5, IBFD seriously suffers from the negative effects of phase noise. Figure 10 visualizes the negative effects of phase noise impairments for the delay $\tau_A = 1$ [nsec] and 16-bit pilot sequence ($M = 16$). In the case of $\zeta = -10$ [dB] of Fig. 10(a), the performance of the ASIC and DASIC-"w/o ECC" is degraded compared to that in Fig. 9(c). However, "w/ ECC" and "IDD" of the DASIC achieve almost the same performance. In the case of $\zeta = -15$ [dB], it is subject to 1 [dB] penalty at PER = $10^{-3}$ in the DASIC "w/ ECC." Finally, in the case of $\zeta = -20$ [dB], it is subject to 2 [dB] penalty at PER = $10^{-3}$ even in "IDD." However, there is no PER error floor in the curve. Furthermore, in the other simulation, we have confirmed the PER = 0.1 is achieved by the DASIC-"IDD" at $|h_{AB}|^2 E_s/N_0 = 8$ [dB] in the case of $\zeta = -25$ [dB] in the other simulations. These results explicitly validate the proposed DASIC-IDD algorithm even in the presence of phase noise when SIR is lager than $\zeta = -20$ [dB], despite the fact that the ASIC is more sensitive to phase noise behavior. The most impressive feature of the DASIC is the fact that asynchronous IBFD is realized.

When the SIR is below $\zeta = -20$ [dB], the proposed scheme is not so effective in the presence of phase noise effects because it leads to a significant increase in the transmission power to meet the required PER. Unfortunately, it is difficult to realize IBFD using the proposed method alone, which needs to support typical SIRs of −90 to −70 [dB] in short-range communications for wireless IoT. It is assumed that the proposed method will be used in combination with conventional self-interference cancellation techniques with a total of approximately 60 [dB] interference reduction: approximately 20 [dB] attenuation with antenna isolation, and approximately 40 [dB] with the other passive analog canceller. However, the scale of analog circuits required to achieve significant interference suppression with a passive analog canceller is large, making them unsuitable for wireless IoT applications. To address this problem, it is necessary to consider implementing the proposed method with a technique to suppress the effects of phase noise.

## 6. Conclusion

In this paper, we proposed a novel DASIC algorithm for asynchronous IBFD GFSK, which is designed for wireless IoT. The DASIC incorporates the principle of the differential codec. Because of the differential structure, the DASIC can suppress SI without the CSI estimation of SI. To improve the detection capability, IDD for the DASIC was appropriately investigated.

In ideal channel estimation conditions, the DASIC has the disadvantage of decreasing the reception sensitivity compared to the ASIC. However, when the CSI contains estimation errors induced by pilot contamination, the detection capability of the DASIC is better than that of the ASIC. Furthermore, with the aid of IDD, the SIR tolerance is improved to approximately 20 [dB] for maintaining PER = $10^{-3}$. Although the utilization field is limited to indoor environments, the results explicitly demonstrate the possibility of the realization of asynchronous IBFD in combination with conventional antenna isolation and analog SI cancellation techniques. The asynchronous blind SI cancellation will open new vistas in the utilization of IBFD in wireless IoT. As a future work, practical implementation of analog DASIC circuit is an attractive issue.

## Acknowledgments

A part of this work was financially supported by JSPS KAKENHI Grant Number JP21H01332 and grants-in-aid from the Haris Science Research Institute of Doshisha University.

### References

[1] K. Chang, "Bluetooth: A viable solution for IoT?," IEEE Wireless Commun., vol.21, no.6, pp.6–7, Dec. 2014.
[2] M.R. Palattella, M. Dohler, A. Grieco, G. Rizzo, J. Torsner, T. Engel, and L. Ladid, "Internet of Things in the 5G era: Enablers, architecture, and business models," IEEE J. Sel. Areas Commun., vol.34, no.3, pp.510–527, March 2016.
[3] K. Hirade, M. Ishizuka, F. Adachi, and K. Ohtani, "Error-rate performance of digital FM with differential detection in land mobile radio channels," IEEE Trans. Veh. Technol., vol.28, no.3, pp.204–212, Aug. 1979.
[4] B.E. Rimoldi, "A decomposition approach to CPM," IEEE Trans. Inf. Theory, vol.34, no.2, pp.260–270, March 1988.
[5] D. Bharadia, E. McMilin, and S. Katti, "Full duplex radios," Proc. ACM SIGCOMM 2013, pp.375–386, Aug. 2013.
[6] Z. Zhang, K. Long, A.V. Vasilakos, and L. Hanzo, "Full-duplex wireless communications: Challenges, solutions, and future research directions," Proc. IEEE, vol.104, no.7, pp.1369–1409, July 2016.
[7] K.E. Kolodziej, B.T. Perry, and J.S. Herd, "In-band full-duplex technology: Techniques and systems survey," IEEE Trans. Microw. Theory Techn., vol.67, no.7, pp.3025–3041, July 2019.
[8] A. Gupta and R.K. Jha, "A survey of 5G network: Architecture and emerging technologies," IEEE Access, vol.3, pp.1206–1232, 2015.
[9] Z. Zhang, X. Chai, K. Long, A.V. Vasilakos, and L. Hanzo, "Full duplex techniques for 5G networks: Self-interference cancellation, protocol design, and relay selection," IEEE Commun. Mag., vol.53, no.5, pp.128–137, May 2015.
[10] V. Wong, Key Technologies for 5G Wireless Systems, Cambridge University Press, 2017.
[11] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," IEEE Trans. Wireless Commun., vol.11, no.12, pp.4296–4307, Dec. 2012.
[12] A. Sahai, G. Patel, C. Dick, and A. Sabharwal, "On the impact of phase noise on active cancelation in wireless full-duplex," IEEE Trans. Veh. Technol., vol.62, no.9, pp.4494–4510, Nov. 2013.
[13] E. Everett, A. Sahai, and A. Sabharwal, "Passive self-interference suppression for full-duplex infrastructure nodes," IEEE Trans. Wireless Commun., vol.13, no.2, pp.680–694, Feb. 2014.
[14] B.P. Day, A.R. Margetts, D.W. Bliss, and P. Schniter, "Full-duplex MIMO relaying: Achievable rates under limited dynamic range," IEEE J. Sel. Areas Commun., vol.30, no.8, pp.1541–1553, Sept.

2012.
[15] P. Popovski and H. Yomo, "Physical network coding in two-way wireless relay channels," Proc. IEEE ICC 2007, pp.707–712, June 2007.
[16] Z. Zhang, Z. Ma, Z. Ding, M. Xiao, and G.K. Karagiannidis, "Full-duplex two-way and one-way relaying: Average rate, outage probability, and tradeoffs," IEEE Trans. Wireless Commun., vol.15, no.6, pp.3920–3933, June 2016.
[17] Y. Choi and H. Shirani-Mehr, "Simultaneous transmission and reception: Algorithm, design and system level performance," IEEE Trans. Wireless Commun., vol.12, no.12, pp.5992–6010, Dec. 2013.
[18] D. Korpi, T. Riihonen, V. Syrjälä, L. Anttila, M. Valkama, and R. Wichman, "Full-duplex transceiver system calculations: Analysis of ADC and linearity challenges," IEEE Trans. Wireless Commun., vol.13, no.7, pp.3821–3836, July 2014.
[19] A. Tang and X. Wang, "A-duplex: Medium access control for efficient coexistence between full-duplex and half-duplex communications," IEEE Trans. Wireless Commun., vol.14, no.10, pp.5871–5885, Oct. 2015.
[20] A. Sabharwal, P. Schniter, D. Guo, D.W. Bliss, S. Rangarajan, and R. Wichman, "In-band full-duplex wireless: Challenges and opportunities," IEEE J. Sel. Areas Commun., vol.32, no.9, pp.1637–1652, Sept. 2014.
[21] D. Kim, H. Lee, and D. Hong, "A survey of in-band full-duplex transmission: From the perspective of PHY and MAC layers," IEEE Commun. Surveys Tuts., vol.17, no.4, pp.2017–2046, Fourth quarter 2015.
[22] M. Kobayashi, R. Murakami, K. Kizaki, S. Saruwatari, and T. Watanabe, "Wireless full-duplex medium access control for enhancing energy efficiency," IEEE Trans. Green Commun. Networking, vol.2, no.1, pp.205–221, March 2018.
[23] K. Komatsu, Y. Miyaji, and H. Uehara, "Basis function selection of frequency-domain Hammerstein self-interference canceller for in-band full-duplex wireless communications," IEEE Trans. Wireless Commun., vol.17, no.6, pp.3768–3780, June 2018.
[24] G. Qiao, S. Gan, S. Liu, L. Ma, and Z. Sun, "Digital self-interference cancellation for asynchronous in-band full-duplex underwater acoustic communication," Sensors, vol.18, no.6, 1700, May 2018.
[25] J. Choi, D. Kim, S. Lee, H. Lee, J. Bang, and D. Hong, "A new frame structure for asynchronous in-band full-duplex systems," Proc. IEEE PIMRC 2015, pp.487–491, 2015.
[26] A. Munari, P. Mähönen, and M. Petrova, "A stochastic geometry approach to asynchronous aloha full-duplex networks," IEEE/ACM Trans. Netw., vol.25, no.6, pp.3695–3708, 2017.
[27] V. Towhidlou and M. Shikh-Bahaei, "Adaptive full-duplex communications in cognitive radio networks," IEEE Trans. Veh. Technol., vol.67, no.9, pp.8386–8395, 2018.
[28] L. Hanzo, T. Liew, and B. Yeap, Turbo Coding, Turbo Equalisation and Space-Time Coding, Wiley, New York, 2002.
[29] D. MacKay, Information Theory, Inference and Learning Algorithms, Cambridge University Press, 2003.
[30] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," IEEE Trans. Inf. Theory, vol.42, no.2, pp.429–445, March 1996.
[31] M. Tuchler and A.C. Singer, "Turbo equalization: An overview," IEEE Trans. Inf. Theory, vol.57, no.2, pp.920–952, Feb. 2011.
[32] S. Ibi, T. Matsumoto, R. Thoma, S. Sampei, and N. Morinaga, "EXIT chart-aided adaptive coding for multilevel BICM with turbo equalization in frequency-selective MIMO channels," IEEE Trans. Veh. Technol., vol.56, no.6, pp.3757–3769, Nov. 2007.
[33] T. Takahashi, S. Ibi, and S. Sampei, "Design of adaptively scaled belief in multi-dimensional signal detection for higher-order modulation," IEEE Trans. Commun., vol.67, no.3, pp.1986–2001, March 2019.
[34] M. Elsayed, A.A.A. El-Banna, O.A. Dobre, W. Shiu, and P. Wang, "Low complexity neural network structures for self-interference cancellation in full-duplex radio," IEEE Commun. Lett., vol.25, no.1, pp.181–185, 2021.
[35] Z. Wang, M. Ma, and F. Qin, "Neural-network-based nonlinear self-interference cancelation scheme for mobile stations with dual-connectivity," IEEE Access, vol.9, pp.53566–53575, 2021.
[36] A. Balatsoukas-Stimming and C. Studer, "Deep unfolding for communications systems: A survey and some new directions," Proc. IEEE SiPS 2019, pp.266–271, 2019.
[37] D. Shirase, T. Takahashi, S. Ibi, K. Muraoka, N. Ishii, and S. Sampei, "Deep unfolding-aided gaussian belief propagation for correlated large mimo detection," Proc. GLOBECOM 2020, pp.1–6, 2020.
[38] J.B. Anderson, T. Aulin, and C.E. Sundberg, Digital Phase Modulation, 1st ed., Springer Publishing Company, 2013.

**Shinsuke Ibi** received the B.E. degree in advanced engineering from Suzuka College of Technology, Japan, in 2002, and the ME and Ph.D. degrees in communication engineering from Osaka University, Japan, in 2004 and 2006, respectively. From 2005 to 2006, he was a visiting researcher at the Centre for Wireless Communications, University of Oulu, Finland. In 2006, he joined the Graduate School of Engineering, Osaka University. From 2010 to 2011, he was a visiting researcher at the University of Southampton, United Kingdom. In 2019, he moved to Doshisha University, and he is currently a professor in the Faculty of Science and Engineering. His research interests include EXIT-based coding theory, iterative detection, digital signal processing, cognitive radio, and communication theory. He received the 64th, 71th, and 80th Best Paper Awards from IEICE, 2017, 2018, 2019, and 2023 Best Paper Awards from IEICE Communication Society, and the 24th Telecom System Technology Award from the Telecommunication Advancement Foundation. He is a senior member of IEEE.

**Takumi Takahashi** received his B.E., M.E., and Ph.D. degrees in communication engineering from Osaka University, Osaka, Japan, in 2016, 2017, and 2019, respectively. From 2018 to 2019, he was a visiting researcher at the Centre for Wireless Communications, University of Oulu, Finland. In 2019, he joined the Graduate School of Engineering, Osaka University as an assistant professor. His research interests include belief propagation, compressed sensing, signal processing, and wireless communications. He is a member of IEEE.

**Hisato Iwai** received the B.E., M.E., and Ph.D. degrees from Kyoto University, Kyoto, Japan in 1987, 1989, and 2001, respectively. In 1989, he joined KDD R&D Laboratories, where he was engaged in research on antennas and propagation in radio communication systems. From 1996 to 1997, he was a visiting scholar at University of California, San Diego (UCSD). Also, he was with Telcordia Technologies NJ as a visiting researcher from 1999 to 2001. From 2004 to 2005, he was a research head at ATR Adaptive Communication Research Laboratories. He is currently a professor at Doshisha University, Kyoto, Japan. His current research interests include propagation models in various radio communication systems, radio transmission systems, applied radio systems, and so on. He is a member of IEEE.

IEICE TRANS. COMMUN., VOL.E107–B, NO.8 AUGUST 2024

PAPER

# Sum Rate Maximization for Multiuser Full-Duplex Wireless Powered Communication Networks

**Keigo HIRASHIMA**[†], *Student Member and* **Teruyuki MIYAJIMA** [†a)], *Member*

**SUMMARY** In this paper, we consider an orthogonal frequency division multiple access (OFDMA)-based multiuser full-duplex wireless powered communication network (FD WPCN) system with beamforming (BF) at an energy transmitter (ET). The ET performs BF to efficiently transmit energy to multiple users while suppressing interference to an information receiver (IR). Multiple users operating in full-duplex mode harvest energy from the signals sent by the ET while simultaneously transmitting information to the IR using the harvested energy. We analytically demonstrate that the FD WPCN is superior to its half-duplex (HD) WPCN counterpart in the high-SNR regime. We propose a transmitter design method that maximizes the sum rate by determining the BF at the ET, power allocation at both the ET and users, and sub-band allocation. Simulation results show the effectiveness of the proposed method.

***key words:*** *wireless energy transfer, energy beamforming, interference suppression, power allocation, sub-band allocation*

## 1. Introduction

With the emergence of Internet-of-Everything systems, large-scale Internet-of-Things (IoT) networks are expected to be developed. Because significant energy resources are required to operate such a large network stably, an energy-efficient communication system is required for 6G systems [1]. For this requirement, wireless energy transmission (WET) via radio frequency (RF) signals is an effective technology to provide energy to energy-constrained communication terminals distributed in a wide area [2]. Simultaneous wireless information and power transfer (SWIPT) systems and wireless powered communication network (WPCN) systems have been actively studied as candidates for WET-based communication systems. In SWIPT, user terminals harvest energy simultaneously while receiving information from RF signals sent by a transmitter. On the other hand, in a WPCN, user terminals harvest energy from RF signals sent by an energy transmitter (ET) and then the user terminals use the harvested energy for wireless information transmission (WIT) to an information receiver (IR) [3].

Most previous studies on WPCN systems assumed half-duplex (HD) transmission based on a harvest-then-transmit protocol, which temporally separates WET and WIT to prevent interference from the ET to the IR [4]–[6]. Several studies have been conducted on transceiver design for HD WPCN systems. In [4], the time allocation for WET and WIT in a time division multiple access (TDMA)-based multiuser WPCN was optimized by maximizing the sum rate. A joint time and power allocation method that maximizes the energy efficiency of a TDMA-based multiuser WPCN was proposed in [5]. In [6], the beamformer in a space division multiple access (SDMA)-based multiuser WPCN was designed by maximizing the minimum throughput. The major disadvantage of HD WPCN systems is the loss of spectral efficiency because information transmission cannot be performed during the WET phase.

Full-duplex (FD) transmission, which enables simultaneous WET and WIT, is an effective method for overcoming this problem. An inherent problem with FD WPCN systems is that the signal sent by the ET causes interference at the IR, which has a significant negative impact on system performance. The authors of [7] proposed a time allocation method that maximizes the sum rate in a TDMA-based multiuser FD WPCN and demonstrated its superiority over an HD WPCN if the interference can be effectively suppressed. However, specific methods for suppressing the interference were not provided in [7]. In [8], a beamforming (BF) technique at an ET equipped with multiple antennas was employed to suppress the interference in an FD WPCN by directing the null to the IR, demonstrating its effectiveness over an HD WPCN. However, the authors in [8] assumed a single-user system. For large-scale IoT WPCNs, the system should be extended to multiuser systems. In [9], [10], the extensions of the single-user system in [8] to multiuser one were proposed. However, they operate in the HD mode.

Orthogonal frequency division multiple access (OFDMA) has been widely used in narrowband IoT networks to accommodate multiple users because of its low complexity and efficient spectrum utilization [11]. In [12], [13], resource allocation, such as sub-bands and power, for an OFDMA-based WPCN was considered. However, these systems are assumed to operate in the HD mode. The authors in [14] proposed an allocation method for power and sub-bands for an OFDMA-based multiuser FD WPCN with a hybrid access point (HAP), where the ET and IR are collocated and communicate with multiple users. The authors in [15] extends the system in [14] to include relay nodes to support WIT. A HAP equipped with a single antenna transmits the energy signal to the users and simultaneously receives the information signals sent by the users using the harvested energy. Although the IR in the HAP suffers from interference form the ET, concrete methods to suppress it were not provided in [14], [15]. BF at the ET as in [8] is expected to

Manuscript received October 9, 2023.
Manuscript revised January 5, 2024.
Manuscript publicized March 5, 2024.

[†]Graduate School of Science and Engineering, Ibaraki University, Hitachi-shi, 316-8511 Japan.
a) E-mail: teruyuki.miyajima.spc@vc.ibaraki.ac.jp

DOI: 10.23919/transcom.2023EBP3165

effectively mitigate interference even in OFDMA-based FD WPCN systems as in [14], but no such studies have been conducted. This motivates us to study the interference suppression by incorporating BF at the ET into OFDMA-based FD WPCN systems.

In this paper, we consider an OFDMA-based multiuser FD WPCN system with a BF in the ET. In the system, the ET performs BF to efficiently transmit energy to multiple users while suppressing interference to the IR, and the users operating in the FD mode harvest energy from the signals sent by the ET while transmitting information on the allocated sub-bands using the harvested energy to the IR. We propose a transmitter design method that maximizes the sum rate by determining the BF at the ET, power allocated to each sub-band at the ET, sub-bands allocated to each user, and power allocated to each sub-band at each user. Analytical and simulation results are provided to demonstrate the effectiveness of the proposed method.

Throughout this paper, the following notations are used: $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^*$ indicate the transpose, Hermitian transpose, and complex conjugate of a vector or matrix, respectively. $\otimes$ denotes the Kronecker product, $\mathbf{0}_{N\times M}$ denotes a $N \times M$ matrix whose elements are all zero, $\mathbf{I}_N$ denotes a $N \times N$ identity matrix, $\mathbf{J}_n \in \mathbb{R}^{N\times N}$denotes a matrix with $(n,n)$ elements of 1 and 0 otherwise, $[x]^+ = \max\{0, x\}$, $|S|$ denotes the number of elements in the set $S$.

## 2. System Model

We consider an OFDMA-based FD WPCN system that contains an ET, $K$ users, and an IR (Fig. 1). The ET transmits energy signals toward the users. Each user operating in FD mode harvests energy from its received signal and simultaneously transmits its information data to the IR using the harvested energy. The IR detects the data from its received signal. ET uses $M_\mathrm{T}$ transmit antennas to suppress interference to the IR as in the system in [8]. The IR is equipped with a single receiving antenna, and each user is equipped with a single transmitting antenna and a single receiving antenna. We assume that the system has perfect knowledge of all channels, which are frequency-selective. OFDMA with $N$ sub-bands is applied to the energy transmission by the ET and information transmission by the users.

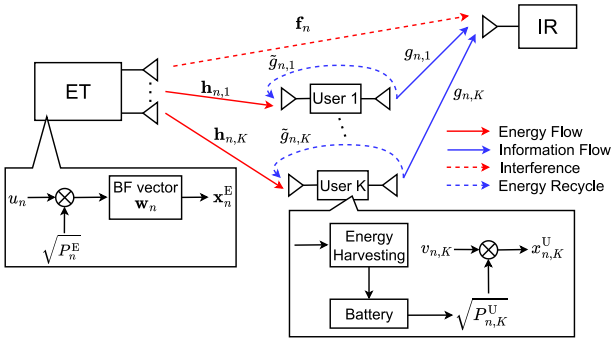



**Fig. 1** Proposed OFDMA-based multiuser FD WPCN system.

The transmitted signal of the ET in the $n$th sub-band is expressed as

$$\mathbf{x}_n^\mathrm{E} = \sqrt{P_n^\mathrm{E}}\mathbf{w}_n u_n, \tag{1}$$

where $P_n^\mathrm{E}$ is the transmit power of the ET in the $n$th sub-band, $\mathbf{w}_n \in \mathbb{C}^{M_\mathrm{T}\times 1}$ is the BF vector in the $n$th sub-band, and $u_n$ is a random signal for energy transmission in the $n$th sub-band. We assume that $\{u_n\}$ are normalized and independent of each other, i.e., $\mathbb{E}[u_n u_{n'}^*] = 1$ for $n = n'$, and otherwise 0, and the BF vectors $\mathbf{w}_n$ are normalized, i.e., $\|\mathbf{w}_n\| = 1, \forall n$. Thus, the total transmit power of the ET is given by

$$P^\mathrm{E} = \sum_{n=0}^{N-1} \mathbb{E}\left[\|\mathbf{x}_n^\mathrm{E}\|^2\right] = \sum_{n=0}^{N-1} P_n^\mathrm{E}. \tag{2}$$

The transmitted signal from the $k$th user is expressed as

$$x_{n,k}^\mathrm{U} = \sqrt{P_{n,k}^\mathrm{U}}v_{n,k}, \tag{3}$$

where $P_{n,k}^\mathrm{U}$ is the transmit power of the $k$th user in the $n$th sub-band, and $v_{n,k}$ is the data symbol of the $k$th user in the $n$th sub-band. We also assume that the transmit symbols $\{v_{n,k}\}$ are normalized and independent of each other, i.e., $\mathbb{E}[v_{n,k}v_{n'}^*] = 1$ for $n = n'$, and otherwise 0. Each user receives the signals sent from the ET and its own transmitted signal owing to the FD operation and harvests energy from the received signals. The harvested power of the $k$th user is expressed as [12]

$$P_k^\mathrm{EH} = \eta\left(\sum_{n=0}^{N-1} P_n^\mathrm{E}|\mathbf{h}_{n,k}^T\mathbf{w}_n|^2 + \sum_{n\in\mathcal{N}_k} P_{n,k}^\mathrm{U}|\tilde{g}_{n,k}|^2\right), \tag{4}$$

where $\mathbf{h}_{n,k} \in \mathbb{C}^{M_\mathrm{T}\times 1}$ is the channel from the ET to the $k$th user in the $n$th sub-band, $\tilde{g}_{n,k} \in \mathbb{C}$ is the loopback channel of the $k$th user in the $n$th sub-band, $\mathcal{N}_k$ is the set of the sub-bands allocated to the $k$th user, and $\eta$ is the energy conversion efficiency. Note that the users do not suffer from so-called self-interference caused by the FD operation since they do not aim for information reception.

Assuming that each user uses all the harvested energy to transmit its information data, the following relationship holds for all $k$[†]:

$$\sum_{n\in\mathcal{N}_k} P_{n,k}^\mathrm{U} = P_k^\mathrm{EH}. \tag{5}$$

Furthermore, from Eqs. (4) and (5), we obtain a closed-form expression for $P_{n,k}^U$ as

$$P_{n,k}^\mathrm{U} = \alpha_{n,k}\frac{\eta\sum_{n'=0}^{N-1} P_{n'}^\mathrm{E}|\mathbf{h}_{n',k}^T\mathbf{w}_{n'}|^2}{1-\eta|\tilde{g}_{n,k}|^2}, \tag{6}$$

where a power allocation parameter $\alpha_{n,k} \in \mathbb{R}$ is introduced

[†]Past literature on OFDMA-based WPCNs [12]–[15] imposed an inequality constraint for WIT energy consumption. Instead, we impose the equality constraint, which makes the transmitter design problem tractable.

in place of $P^{\mathrm{U}}_{n,k}$, subject to $\alpha_{n,k} \geq 0$, and $\sum_{n\in\mathcal{N}_k} \alpha_{n,k} = 1$, $\forall k$. We assume that the signals from the ET and all users arriving at the IR are time-aligned within the cyclic-prefix interval. Subsequently, the received signals of the IR in the $n$th sub-band can be expressed as

$$y^{\mathrm{I}}_n = \sqrt{P^{\mathrm{U}}_{n,k}} g_{n,k}\, v_{n,k} + \sqrt{P^{\mathrm{E}}_n} \mathbf{f}^T_n \mathbf{w}_n u_n + z_n, \tag{7}$$

where $g_{n,k} \in \mathbb{C}$ is the channel of the $k$th user in the $n$th sub-band, $\mathbf{f}_n \in \mathbb{C}^{M_{\mathrm{T}}\times 1}$ is the channel vector from the ET to the IR in the $n$th sub-band, the first term is the received signal from the $k$th user in the $n$th sub-band, the second term is the interference from the ET to the IR in the $n$th sub-band, and $z_n$ is the additive white Gaussian noise in the $n$th sub-band at the IR. From Eqs. (6) and (7), the received signal-to-interference-plus-noise ratio (SINR) for the $k$th user in the $n$th sub-band at the IR is obtained by

$$\begin{aligned}\gamma_{n,k} &= \frac{P^{\mathrm{U}}_{n,k}|g_{n,k}|^2}{P^{\mathrm{E}}_n|\mathbf{f}^T_n\mathbf{w}_n|^2 + \sigma^2_z} \\ &= \frac{\alpha_{n,k}\sum_{n'=0}^{N-1} P^{\mathrm{E}}_{n'}|\mathbf{h}^T_{n',k}\mathbf{w}_{n'}|^2}{P^{\mathrm{E}}_n|\mathbf{f}^T_n\mathbf{w}_n|^2 + \sigma^2_z} G_{n,k},\end{aligned} \tag{8}$$

where $G_{n,k} = \frac{\eta|g_{n,k}|^2}{1-\eta|\bar{g}_{n,k}|^2}$. Note that, according to (5), the SINR includes $\{\mathbf{w}_n\}$ and $\{P^{\mathrm{E}}_n\}$ in both the numerator and denominator. This makes the WET BF design and WET power allocation tractable, as explained in Sect. 4. Subsequently, the achievable sum rate of the system is expressed as

$$R_{\mathrm{sum}} = \sum_{k=1}^{K}\sum_{n\in\mathcal{N}_k} \log_2(1+\gamma_{n,k}). \tag{9}$$

The aim of the transmitter design explained in Sect. 4 is to determine the variables $\{\mathbf{w}_n\}$, $\{P^{\mathrm{E}}_n\}$, $\{\alpha_{n,k}\}$, and $\{\mathcal{N}_k\}$ to maximize $R_{\mathrm{sum}}$[†].

## 3. Performance Analysis

We compare the transmission rate of the FD WPCN with that of the HD WPCN. For simplicity, we assume that the same power $\alpha_{n,k}$ is allocated to all sub-bands in both FD and HD WPCNs and that the sub-bands allocated to a user $\mathcal{N}_k$ are the same for the FD and HD WPCNs. It is reasonable to assume that the harvested power in the HD WPCN is larger than that in the FD WPCN by appropriately setting $\mathbf{w}_n$ and $P^{\mathrm{E}}_n$ because no interference from the ET to the IR exists; thus, the BF at the ET can focus only on transmitting energy to users in the HD WPCN.

In the HD WPCN, the users harvest energy from the signals sent by the ET in the first phase of duration $\tau T$ and send the information data using the harvested energy in the second phase of duration $(1-\tau)T$, where $0 < \tau < 1$, and $T$ is the block transmission time. Thus, the following relationship between the harvested and transmitted energies of the $k$th user holds:

$$\tau T P^{\mathrm{HD}}_{\mathrm{EH},k} = (1-\tau)T \sum_{n\in\mathcal{N}_k} P^{\mathrm{U}}_{n,k}, \tag{10}$$

where $P^{\mathrm{HD}}_{\mathrm{EH},k}$ is the harvested power of the $k$th HD user. From Eq. (10), we obtain $P^{\mathrm{U}}_{n,k} = \tau P^{\mathrm{HD}}_{\mathrm{EH},k}/(1-\tau)|\mathcal{N}_k|$ under the assumption of equal power allocation. Therefore, the transmission rate of the $k$th user in the $n$th sub-band in the HD WPCN is expressed as

$$R^{\mathrm{HD}}_{n,k} = \log_2\left(1 + \frac{\tau P^{\mathrm{HD}}_{\mathrm{EH},k}|g_{n,k}|^2}{(1-\tau)|\mathcal{N}_k|\sigma^2_z}\right), \tag{11}$$

for $n \in \mathcal{N}_k$.

For the FD WPCN, we obtain $P^{\mathrm{U}}_{n,k} = P^{\mathrm{FD}}_{\mathrm{EH},k}/|\mathcal{N}_k|$ under the assumption of equal power allocation, where $P^{\mathrm{FD}}_{\mathrm{EH},k}$ is the power harvested by the $k$th FD user. Subsequently, the transmission rate of the $k$th user in the $n$th sub-band in the FD WPCN is expressed as

$$R^{\mathrm{FD}}_{n,k} = \log_2\left(1 + \frac{P^{\mathrm{FD}}_{\mathrm{EH},k}|g_{n,k}|^2}{|\mathcal{N}_k|(I_n + \sigma^2_z)}\right), \tag{12}$$

for $n \in \mathcal{N}_k$, where $I_n > 0$ represents the power of the interference from the ET to the IR in the $n$th sub-band. From the assumption on the harvested power mentioned above, we obtain $P^{\mathrm{HD}}_{\mathrm{EH},k} \geq P^{\mathrm{FD}}_{\mathrm{EH},k} = \mu_k P^{\mathrm{HD}}_{\mathrm{EH},k}$ for $0 \leq \mu_k \leq 1$. In addition, $\sigma^2_z \leq I_n + \sigma^2_z = \nu_n\sigma^2_z$ holds for $\nu_n > 1$. Therefore, we can rewrite (12) as $R^{\mathrm{FD}}_{n,k} = \log_2\left(1 + \mu_k P^{\mathrm{HD}}_{\mathrm{EH},k}|g_{n,k}|^2/\nu_n|\mathcal{N}_k|\sigma^2_z\right)$.

By defining the signal-to-noise ratio (SNR) as $\gamma_{n,k} = P^{\mathrm{HD}}_{\mathrm{EH},k}\|g_{n,k}\|^2/|\mathcal{N}_k|\sigma^2_z$, the transmission rates can be rewritten as

$$R^{\mathrm{HD}}_{n,k} = (1-\tau)\log_2\left(1 + \frac{\tau}{1-\tau}\gamma_{n,k}\right) \tag{13}$$

$$R^{\mathrm{FD}}_{n,k} = \log_2\left(1 + \lambda_{n,k}\gamma_{n,k}\right), \tag{14}$$

where $\lambda_{n,k} = \mu_k/\nu_n$, which satisfies $0 < \lambda_{n,k} < 1$ and indicates the SNR degradation of the FD WPCN. Taking the limit of $\gamma_{n,k} \to \infty$ for $R^{\mathrm{HD}}_{n,k}/R^{\mathrm{FD}}_{n,k}$, we obtain

$$\begin{aligned}\lim_{\gamma_{n,k}\to\infty} \frac{R^{\mathrm{HD}}_{n,k}}{R^{\mathrm{FD}}_{n,k}} &= (1-\tau)\lim_{\gamma_{n,k}\to\infty} \frac{\log_2(1+\frac{\tau\gamma_{n,k}}{1-\tau})}{\log_2(1+\lambda_{n,k}\gamma_{n,k})} \\ &= 1-\tau < 1.\end{aligned} \tag{15}$$

The result $R^{\mathrm{FD}}_{n,k} > R^{\mathrm{HD}}_{n,k}$ holds for any $k$ and $n$. Thus, we can conclude that the sum rate of the FD WPCN is higher

[†]A few sub-bands may be sufficient for the energy transmission and can avoid interference from the ET to the IR. This is valid if there are no restrictions on the power spectral density of the transmitted signals. Even in this case, it is crucial to determine the sub-bands used and appropriate power allocation to the sub-bands for efficient energy transmission. Moreover, it was reported in [16] that it is necessary to distribute the transmit power over multiple sub-bands to comply with regulations such as the power spectral density constraints in the ISM band. These facts justify the power allocation $\{P^{\mathrm{E}}_n\}$ to the sub-bands for the energy transmission.

than that of the HD WPCN. Note that, however, the result holds only for a sufficiently high SNR. To evaluate their performance at moderate SNRs, we conducted simulations, which are described in Sect. 5.

## 4. Transmitter Design

Now, we design the transmitters of the ET and users. Specifically, the problem of determining $\{\mathbf{w}_n\}, \{P_n^{\mathrm{E}}\}, \{\alpha_{n,k}\}, \{\mathcal{N}_k\}$ that maximizes $R_{\mathrm{sum}}$ is formulated as follows:

$$\text{(P1):} \quad \max_{\{\mathbf{w}_n\},\{P_n^{\mathrm{E}}\},\{\alpha_{n,k}\},\{\mathcal{N}_k\}} \quad R_{\mathrm{sum}} \tag{16}$$

$$\text{s.t.} \quad \|\mathbf{w}_n\| = 1, \forall n, \tag{17}$$

$$\sum_{n=0}^{N-1} P_n^{\mathrm{E}} \le P_{\max}, \tag{18}$$

$$0 \le P_n^{\mathrm{E}} \le P_{\mathrm{peak}}, \forall n, \tag{19}$$

$$\sum_{n\in\mathcal{N}_k} \alpha_{n,k} = 1, \alpha_{n,k} \ge 0, \forall n, k. \tag{20}$$

Equation (18) implies that the total power at the ET is limited to the maximum power $P_{\max}$. Equation (19) implies that the power allocated to each sub-band is limited to the peak power $P_{\mathrm{peak}}$. This is introduced to address power spectral density constraints of transmitted signals [16]. Jointly optimizing problem (P1) for all variables is difficult. Therefore, we consider solving problem (P1) for each variable sequentially.

### 4.1 Beamforming for Energy Transmission

We consider determining the transmission BF vector $\{\mathbf{w}_n\}$ at the ET when $\{P_n^{\mathrm{E}}\}$, $\{\mathcal{N}_k\}$, and $\{\alpha_{n,k}\}$ are given. We rewrite SINR in (8) as

$$\gamma_{n,k} = \frac{\mathbf{w}^H \tilde{\mathbf{H}}_{n,k} \mathbf{w}}{\mathbf{w}^H \tilde{\mathbf{F}}_n \mathbf{w}}, \tag{21}$$

where $\mathbf{w} = [\mathbf{w}_0^T \ \dots \ \mathbf{w}_{N-1}^T]^T \in \mathbb{C}^{M_{\mathrm{T}}N\times 1}$, and $\tilde{\mathbf{H}}_{n,k}$ and $\tilde{\mathbf{F}}_n$ are $M_{\mathrm{T}}N \times M_{\mathrm{T}}N$ block diagonal matrices defined by

$$\tilde{\mathbf{H}}_{n,k} = \alpha_{n,k} G_{n,k} \cdot \begin{bmatrix} P_0^{\mathrm{E}} \mathbf{h}_{0,k}^* \mathbf{h}_{0,k}^T & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & P_{N-1}^{\mathrm{E}} \mathbf{h}_{N-1,k}^* \mathbf{h}_{N-1,k}^T \end{bmatrix}, \tag{22}$$

$$\tilde{\mathbf{F}}_n = \mathbf{J}_n \otimes P_n^{\mathrm{E}} \mathbf{f}_n^* \mathbf{f}_n^T + \frac{\sigma_z^2}{N} \mathbf{I}_{M_{\mathrm{T}}N}. \tag{23}$$

Assuming a high SINR, the sum rate can be approximated as

$$R_{\mathrm{sum}} \simeq \sum_{k=1}^{K} \sum_{n\in\mathcal{N}_k} \log_2 \left( \frac{\mathbf{w}^H \tilde{\mathbf{H}}_{n,k} \mathbf{w}}{\mathbf{w}^H \tilde{\mathbf{F}}_n \mathbf{w}} \right) = \log_2 \left( \prod_{k=1}^{K} \prod_{n\in\mathcal{N}_k} \frac{\mathbf{w}^H \tilde{\mathbf{H}}_{n,k} \mathbf{w}}{\mathbf{w}^H \tilde{\mathbf{F}}_n \mathbf{w}} \right). \tag{24}$$

**Algorithm 1** GPI Algorithm [17] for Energy BF Design

1: $\kappa = 0$, $\mathbf{w}^{(0)} = \frac{1}{\sqrt{NM_{\mathrm{T}}}}[1, \dots, 1]^T$
2: **repeat**
3: Compute $\mathbf{H}(\mathbf{w}^{(\kappa)})$ in (27) and $\mathbf{F}(\mathbf{w}^{(\kappa)})$ in (28)
4: $\mathbf{y} = \mathbf{F}(\mathbf{w}^{(\kappa)})^{-1}\mathbf{H}(\mathbf{w}^{(\kappa)})\mathbf{w}^{(\kappa)}$
5: $\mathbf{w}^{(\kappa+1)} = \mathbf{y}/\|\mathbf{y}\|$
6: $\kappa = \kappa + 1$
7: **until** $\|\mathbf{w}^{(\kappa)} - \mathbf{w}^{(\kappa-1)}\| < \epsilon$
8: **for** $n = 0$ to $N-1$ **do**
9: $\mathbf{w}_n = \frac{\mathbf{J}_n \otimes \mathbf{I}_{M_{\mathrm{T}}} \mathbf{w}^{(\kappa)}}{\|\mathbf{J}_n \otimes \mathbf{I}_{M_{\mathrm{T}}} \mathbf{w}^{(\kappa)}\|}$
10: **end for**

Therefore, the optimization problem for determining the BF vector $\mathbf{w}$ that maximizes the sum rate is formulated as

$$\text{(P2):} \quad \max_{\mathbf{w}} \prod_{k=1}^{K} \prod_{n\in\mathcal{N}_k} \frac{\mathbf{w}^H \tilde{\mathbf{H}}_{n,k} \mathbf{w}}{\mathbf{w}^H \tilde{\mathbf{F}}_n \mathbf{w}} \quad \text{s.t.} \ \|\mathbf{w}\| = N. \tag{25}$$

Problem (P2) can be solved by using a general power iterative (GPI) algorithm [17]. Defining the objective function as $\lambda(\mathbf{w}) = \prod_{k=1}^{K} \prod_{n\in\mathcal{N}_k} \frac{\mathbf{w}^H \tilde{\mathbf{H}}_{n,k} \mathbf{w}}{\mathbf{w}^H \tilde{\mathbf{F}}_n \mathbf{w}}$, the solutions to problem (P2) should satisfy the condition $\frac{\partial \lambda(\mathbf{w})}{\partial \mathbf{w}^H} = \mathbf{0}$. From the condition, we obtain

$$\mathbf{H}(\mathbf{w})\mathbf{w} = \lambda(\mathbf{w})\mathbf{F}(\mathbf{w})\mathbf{w} \tag{26}$$

where

$$\mathbf{H}(\mathbf{w}) = \sum_{k=1}^{K} \sum_{n\in\mathcal{N}_k} \left( \prod_{k=1}^{K} \prod_{\substack{i\in\mathcal{N}_k \\ i\ne n}} \mathbf{w}^H \tilde{\mathbf{H}}_{i,k} \mathbf{w} \right) \tilde{\mathbf{H}}_{n,k}, \tag{27}$$

$$\mathbf{F}(\mathbf{w}) = \sum_{n=0}^{N-1} \left( \prod_{\substack{i=0 \\ i\ne n}}^{N-1} \mathbf{w}^H \tilde{\mathbf{F}}_i \mathbf{w} \right) \tilde{\mathbf{F}}_n. \tag{28}$$

Because Eq. (26) has a similar form to a generalized eigenvalue problem, we employ the power method, which is an iterative algorithm for solving eigenvalue problems. However, unlike ordinary generalized eigenvalue problems, the matrices in Eq. (26) depend on $\mathbf{w}$. Therefore, the GPI algorithm updates the matrices during the iterations. The GPI algorithm is summarized in Algorithm 1. In Steps 8 to 10, the BF vector on the $n$th sub-band is extracted from the resulting vector $\mathbf{w}^{(\kappa)}$ and normalized.

### 4.2 Power Allocation for Energy Transmission

We consider determining the transmission power $\{P_n^{\mathrm{E}}\}$ at the ET when $\{\mathbf{w}_n\}$, $\{\mathcal{N}_k\}$ and $\{\alpha_{n,k}\}$ are given. The problem of determining $\{P_n^{\mathrm{E}}\}$ that maximizes the sum rate is formulated as

$$\max_{\mathbf{P}_{\mathrm{E}}} \sum_{k=1}^{K} \sum_{n\in\mathcal{N}_k} \log_2 \left( 1 + \frac{\sum_{n'=0}^{N-1} P_{n'}^{\mathrm{E}} |\mathbf{h}_{n',k}^T \mathbf{w}_{n'}|^2}{P_n^{\mathrm{E}} |\mathbf{f}_n^T \mathbf{w}_n|^2 + \sigma_z^2} \alpha_{n,k} G_{n,k} \right)$$

$$\text{s.t.} \ \sum_{n=0}^{N-1} P_n^{\mathrm{E}} \le P_{\max}, \ 0 \le P_n^{\mathrm{E}} \le P_{\mathrm{peak}} \ \forall n, \tag{29}$$

where $\mathbf{P}_E = [P_0^{\mathrm{E}} \ \cdots \ P_{N-1}^{\mathrm{E}}]^T$. This problem is difficult to solve because the objective function $R_{\mathrm{sum}}(\mathbf{P}_{\mathrm{E}})$ is generally nonconvex with respect to $\mathbf{P}_{\mathrm{E}}$. Hence, to solve this problem as a convex optimization problem, we employ the approach described in [18]. First, the objective function is rewritten as the difference of concave functions as $R_{\mathrm{sum}}(\mathbf{P}_{\mathrm{E}}) = f(\mathbf{P}_{\mathrm{E}}) - g(\mathbf{P}_{\mathrm{E}})$, where

$$f(\mathbf{P}_{\mathrm{E}}) = \sum_{k=1}^{K} \sum_{n\in\mathcal{N}_k} \log_2 \left( P_n^{\mathrm{E}} |\mathbf{f}_n^T \mathbf{w}_n|^2 + \sigma_z^2 + \alpha_{n,k} G_{n,k} \sum_{n'=0}^{N-1} P_{n'}^{\mathrm{E}} |\mathbf{h}_{n',k}^T \mathbf{w}_{n'}|^2 \right), \tag{30}$$

$$g(\mathbf{P}_{\mathrm{E}}) = \sum_{n=0}^{N-1} \log_2 \left( P_n^{\mathrm{E}} |\mathbf{f}_n^T \mathbf{w}_n|^2 + \sigma_z^2 \right). \tag{31}$$

Next, the objective function $R_{\mathrm{sum}}(\mathbf{P}_{\mathrm{E}}) = f(\mathbf{P}_{\mathrm{E}}) - g(\mathbf{P}_{\mathrm{E}})$ is approximated by using a first-order approximation of $g(\mathbf{P}_{\mathrm{E}})$. Subsequently, the optimization problem is expressed as

$$\text{(P3):} \quad \max_{\mathbf{P}^{\mathrm{E}}} \ f(\mathbf{P}_{\mathrm{E}}) - g(\mathbf{P}_{\mathrm{E}}^0) - \nabla g^T(\mathbf{P}_{\mathrm{E}}^0) \cdot (\mathbf{P}_{\mathrm{E}} - \mathbf{P}_{\mathrm{E}}^0)$$
$$\text{s.t.} \quad \sum_{n=0}^{N-1} P_n^{\mathrm{E}} \le P_{\max}, \ 0 \le P_n^{\mathrm{E}} \le P_{\mathrm{peak}} \ \forall n, \tag{32}$$

where $\mathbf{P}_{\mathrm{E}}^0 \in \mathbb{C}^{N\times 1}$ is any vector satisfying the constraints and

$$\nabla g^T(\mathbf{P}_{\mathrm{E}}) = \frac{1}{\ln 2} \left[ \frac{|\mathbf{f}_0^T \mathbf{w}_0|^2}{P_0^{\mathrm{E}} |\mathbf{f}_0^T \mathbf{w}_0|^2} \ \cdots \ \frac{|\mathbf{f}_{N-1}^T \mathbf{w}_{N-1}|^2}{P_{N-1}^{\mathrm{E}} |\mathbf{f}_{N-1}^T \mathbf{w}_{N-1}|^2} \right]. \tag{33}$$

Problem (P3) is concave and solvable via the solver CVXPY [19]. Although an iterative approach that updates $\mathbf{P}_{\mathrm{E}}$ in (32) may provide a better solution, a one-shot solution can provide satisfactory performance according to our preliminary simulation results.

### 4.3 Power Allocation for Information Transmission

We determine the transmission power allocation parameters $\{\alpha_{n,k}\}$ for users when $\{\mathbf{w}_n\}$, $\{P_n^{\mathrm{E}}\}$, and $\{\mathcal{N}_k\}$ are provided. The problem of determining the allocation parameters $\{\alpha_{n,k}\}$ can be expressed as

$$\text{(P4):} \quad \max_{\alpha_{n,k}, \forall n,k} \sum_{k=1}^{K} \sum_{n\in\mathcal{N}_k} \log_2 \left( 1 + \alpha_{n,k} G'_{n,k} \right)$$
$$\text{s.t.} \quad \sum_{n\in\mathcal{N}_k} \alpha_{n,k} = 1, \ \alpha_{n,k} \ge 0, \ \forall n,k, \tag{34}$$

where

$$G'_{n,k} = \frac{\sum_{n'=0}^{N-1} P_{n'}^{\mathrm{E}} |\mathbf{h}_{n',k}^T \mathbf{w}_{n'}|^2}{P_n^{\mathrm{E}} |\mathbf{f}_n^T \mathbf{w}_n|^2 + \sigma_z^2} G_{n,k}. \tag{35}$$

This problem can be solved using the water-filling algorithm. Therefore, the solution of (P4) is given by

$$\alpha_{n,k} = \left[ \nu_k - \frac{1}{G'_{n,k}} \right]^+, \ n \in \mathcal{N}_k, \tag{36}$$

where the water level $\nu_k$ is calculated numerically to satisfy the constraints in (P4).

### 4.4 Sub-Band Allocation for Information Transmission

We consider determining the sub-band allocation $\{\mathcal{N}_k\}$. Because information transmission can be considered an uplink OFDMA transmission, we can exploit the concept of a sub-band allocation algorithm for uplink OFDMA in [20]. The key concept of the algorithm is to allocate the $n$th sub-band to the $k$th user with the highest transmission rate as

$$\text{(P5):} \quad k = \arg\max_{k'} \{\log_2(1 + \alpha_{n,k'} G'_{n,k'})\}, \tag{37}$$

where the transmission rate $\log_2(1 + \alpha_{n,k'} G'_{n,k'})$ is obtained by allocating all unallocated sub-bands to the $k'$th user. To adopt criterion (37) in our system, we require further assumptions. To compute the transmission rates, we require $\{\alpha_{n,k}\}$, $\{G'_{n,k}\}$, $\{\mathbf{w}_n\}$ and $\{P_n^{\mathrm{E}}\}$, as shown in (35). Here, we assume that $P_n^{\mathrm{E}}$ are identical for all $n$. To obtain $\{\mathbf{w}_n\}$, we use Algorithm 1 under the assumption of equal power allocation for $\alpha_{n,k}$. Subsequently, we refine $\{\alpha_{n,k}\}$ using (36) and allocate sub-bands to users based on (37) until all sub-bands are allocated. The resulting algorithm for sub-band allocation is summarized in Algorithm 2.

### 4.5 Summary of Transmitter Design

The BF design and resource allocation described above are performed according to the following steps.

Step1) Determine $\{\mathcal{N}_k\}$ using Algorithm 2.
Step2) Initialize $P_n^{\mathrm{E}} = P_{\max}/N, \alpha_{n,k} = 1/|\mathcal{N}_k| \ \forall n,k$.

**Algorithm 2** Sub-band Allocation Algorithm

```
N = {0, 1, ..., N − 1}, Ñ = ∅, Ñ_k = ∅, ∀k
P_n^E = P_max/N, ∀n
for k = 1 to K do
    α_{n,k} = 1/N, ∀n
    Determine {w_{n,k}} using Algorithm1
    Determine {G'_{n,k}} using (35)
end for
while Σ_{k=1}^{K} |Ñ_k| = N do
    for k = 1 to K do
        N_k = N, w_n = w_{n,k}, ∀n
        Determine {α_{n,k}} using (36)
    end for
    for n ∈ N do
        if r_{n,k} = log2(1 + α_{n,k} G'_{n,k}) = 0, ∀k then
            Ñ = Ñ + {n}
        else
            k = arg max_{k'} {r_{n,k'}}, Ñ_k = Ñ_k + {n}
        end if
    end for
    N = Ñ, Ñ = ∅
end while
N_k = Ñ_k
```

Step3) Using $\{P_n^{\mathrm{E}}\}, \{\alpha_{n,k}\}, \{\mathcal{N}_k\}$, determine $\{\mathbf{w}_n\}$ using Algorithm 1.
Step4) Using $\{\mathbf{w}_n\}, \{\alpha_{n,k}\}, \{\mathcal{N}_k\}$, determine $\{P_n^{\mathrm{E}}\}$ by solving (P3).
Step5) Using $\{\mathbf{w}_n\}, \{P_n^{\mathrm{E}}\}, \{\mathcal{N}_k\}$, determine $\{\alpha_{n,k}\}$ by solving (36).

The computational complexity of Steps 1, 3, 4, and 5 are $O(K(N^2M_{\mathrm{T}}^2 + NM_{\mathrm{T}}^3))$, $O(N^2M_{\mathrm{T}}^2 + NM_{\mathrm{T}}^3)$, $O(N^3)$ [18], and $O(N^2M_{\mathrm{T}})$, respectively. Note that the proposed design method can only provide suboptimal solutions. However, simulations in the next section validate that the suboptimal solutions can provide satisfactory performances.

## 5. Simulation Results

We evaluated the performance of the proposed system using computer simulations. Unless otherwise stated, the simulation settings were as follows: The number of sub-bands was $N = 16$, the number of transmit antennas of the ET was $M_{\mathrm{T}} = 5$, the number of users was $K = 2$, the energy-conversion efficiency was $\eta = 0.6$ [21], the maximum transmit power of the ET was $P_{\max} = 40\,\mathrm{dBm}$, the maximum transmit power per sub-band of the ET was $P_{\mathrm{peak}} = 8P_{\max}/N$, the loopback gain of the user was $-10\,\mathrm{dB}$, the noise power was $\sigma_z^2 = -90\,\mathrm{dBm}$, the distance attenuation was $(c/4\pi f_c)^2 d^{2.4}$, where $c$ is the speed of light, $d$ is the distance, and the carrier frequency was $f_c = 900\,\mathrm{MHz}$. The Rayleigh fading channels were modeled as linear FIR filters with a 6 tap exponentially decaying power profile. The locations of the ET, users, and IR are shown in Fig. 2. The ET and IR were located at $(-6\,\mathrm{m}, 0\,\mathrm{m})$ and $(6\,\mathrm{m}, 0\,\mathrm{m})$, respectively, and the users were uniformly distributed in a circle with the center $(x_{\mathrm{U}}, 0)$ and radius 3 m. Except for Fig. 8, we set $x_{\mathrm{U}} = 0$. The average sum rates were obtained from $10^3$ simulation trials in which each trial had an independent channel realization.

First, we confirmed the convergence of the GPI algorithm because convergence cannot be ensured analytically. Figure 3 shows convergence characteristics for $M_{\mathrm{T}} = 2, 5$, and 10. The GPI algorithm converged after about 10 iterations for any number of antennas.

Next, we evaluated the performance of the proposed BF method. Figure 4 shows the sum rate characteristics for various BF methods as a function of the number of antennas $M_{\mathrm{T}}$ at the ET. "EH maximization BF" maximized the harvested power in Eq. (4), which was obtained using a normalized eigenvector corresponding to the largest eigenvalue of the matrix derived from (4). "Interference Nulling BF" completely canceled the interference at the IR, i.e., $\mathbf{f}_n^T\mathbf{w}_n = 0$ for $M_{\mathrm{T}} \geq 2$. When no BF was performed (labeled as "w/o BF"), we set $\mathbf{w}_n = 1/\sqrt{M_{\mathrm{T}}}[1, \ldots, 1]$. For "w/o BF" and "Interference Nulling BF," the sum rate did not increase as the number of antennas increased because they did not consider the sum rate improvement. For "EH maximization BF," the sum rate increased as the number of antennas increased. However, owing to the significant influence of interference,

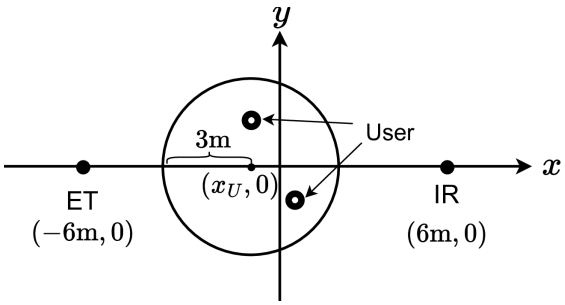



**Fig. 2** Locations of ET, IR, and users.

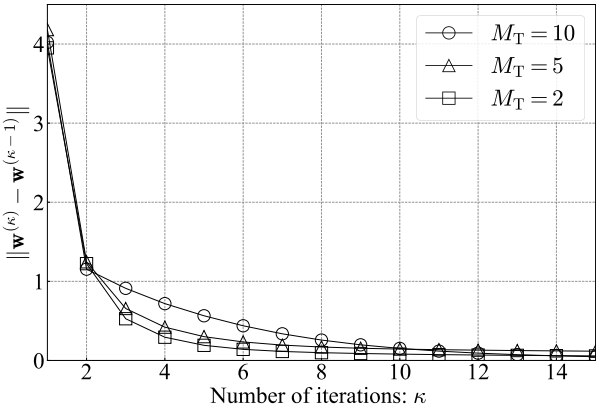



**Fig. 3** Convergence characteristics of GPI algorithm.

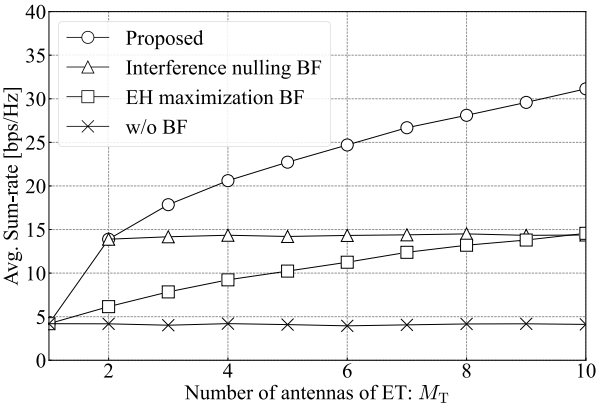



**Fig. 4** Performance comparison of the various BF methods.

its performance was not even comparable to that of "Interference Nulling BF." The proposed method also increased the sum rate as the number of antennas increases and achieved a higher sum rate than the other methods. This is because the proposed method determines BF by considering both interference suppression and efficient energy transmission.

Next, we considered power allocation at both the ET and users. Figure 5 shows the sum rates of the power allocation methods as a function of the maximum transmission power of the ET, where "Equal Power Allocation" was the case in which $P_n^{\mathrm{E}} = P_{\max}/N$ and $\alpha_{n,k} = 1/|\mathcal{N}_k|, \forall n, k$. The sum rates of both methods increased with increasing $P_{\max}$ because the BF at the ET considerably suppressed the interference at

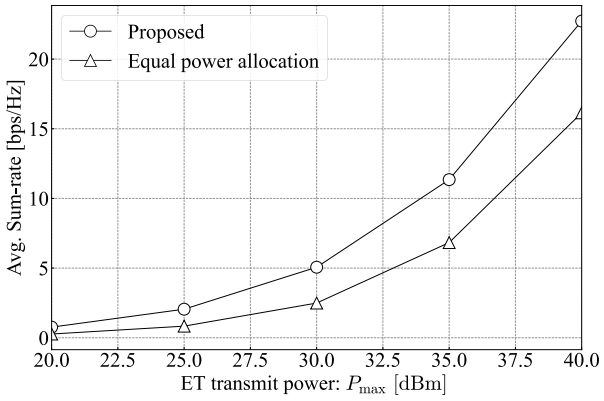



**Fig. 5** Performance of power allocation methods.

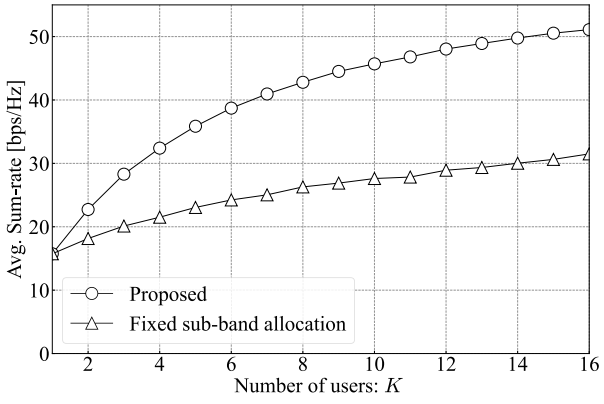



**Fig. 6** Performance of sub-band allocation methods.

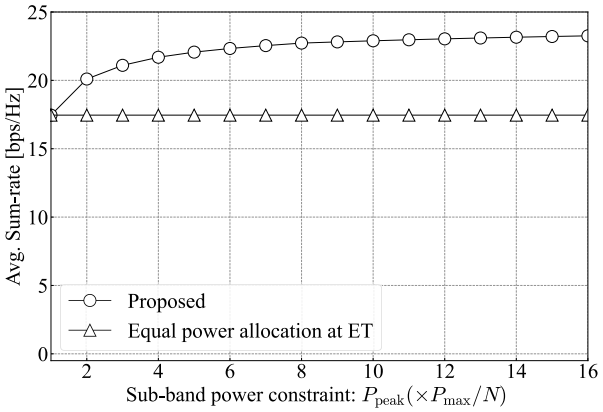



**Fig. 7** Effect of transmission power per sub-band at the ET.

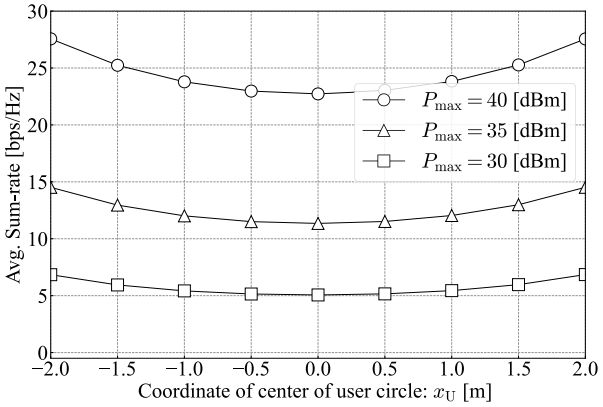



**Fig. 8** Effect of user location.

the IR. The proposed method had a higher rate at any $P_{\max}$ than "Equal Power Allocation" owing to appropriate power allocation.

Next, we considered sub-band allocation methods. Figure 6 shows the sum rate characteristics with respect to the number of the users, where "Fixed Sub-band Allocation" was the case in which the sub-bands allocated to the $k$th user were $n = (k-1) + iK, i = 0, 1, \ldots, \lfloor N/K - 1 \rfloor$. When $K = 1$, the proposed method was identical to a fixed allocation because no sub-band allocation was performed. The performance difference between the proposed method and fixed allocation increased as the number of users increased. This is because an increase in the number of users provides multi-user diversity, which utilizes combinations of users and sub-bands with favorable channel conditions.

Next, we considered the effect of the transmission power constraint per sub-band at the ET. Figure 7 shows the sum rate characteristics with respect to the transmission power constraint per sub-band $P_{\text{peak}}$, where "Equal Power Allocation at ET" was the case in which $P_n^{\text{E}} = P_{\max}/N$. By increasing $P_{\text{peak}}$, the power could be allocated intensively to a few specific sub-bands as long as the maximum transmit power was not exceeded. The figure shows that the sum rate improvement using the proposed method increased with an increase $P_{\text{peak}}$. This result indicates that to improve the sum rate, a few specific sub-bands, rather than all the sub-bands in energy transmission, should be used. However, such a large power allocation to a few sub-bands is not always possible in practical scenarios because it might violate the legal constraints of the transmit power [22].

Next, we considered the effect of the locations of users. Figure 8 shows the sum rates when the x-coordinate $x_{\text{U}}$ of the center of the circle in which the users existed changed. The sum rate had a minimum at $x_{\text{U}} = 0$ and was symmetric around $x_{\text{U}} = 0$ for any $P_{\max}$. Therefore, it is preferable for users to move toward either ET or IR. This is not the case when using an HAP in which both ET and IR are collocated.

Finally, we compared the proposed FD WPCN with the HD WPCN designed using the proposed method with $\mathbf{f}_n = \mathbf{0}$ and $\tilde{g}_{n,k} = 0$. Figure 9 shows the sum rates as a function of the time allocated for energy transmission $\tau$ where the sum rate of the HD WPCN was obtained using $\sum_{k=1}^{K} \sum_{n \in \mathcal{N}_k} (1-\tau) \log_2(1 + \frac{\tau}{1-\tau} \gamma_{n,k})$ and $M_{\text{T}} = 2$. Unlike the analytical results for the high-SNR regime in Sect. 3, the transmission rate of the HD WPCN depended on $\tau$ because the SNR was not sufficiently high in the simulation setup. The results show that the proposed FD WPCN exhibited a higher sum rate than the HD WPCN, regardless of $\tau$ even in the moderate-SNR regime.

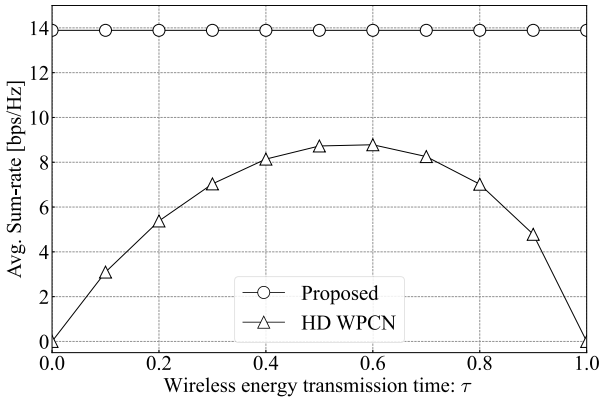



**Fig. 9** Comparison with the half-duplex WPCN system.

## 6. Conclusion

This paper has investigated an OFDMA-based multiuser FD WPCN system with BF. We proposed a sum rate maximization design method that determines the BF and transmit power allocation at the ET, as well as the transmit power and sub-band allocation of each user. The analytical results demonstrated the superiority of the FD WPCN over the HD WPCN in the high-SNR regime. The simulation results showed that the proposed transmitter design is effective in improving the sum rate even in a moderate SNR regime.

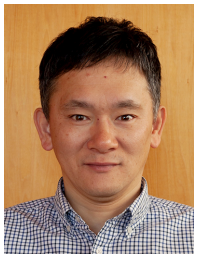
## Acknowledgments

This study was supported by JSPS KAKENHI (Grant number: JP20K04479, JP24K07472).

### References

[1] D.C. Nguyen, M. Ding, P.N. Pathirana, A. Seneviratne, J. Li, D. Niyato, O. Dobre, and H.V. Poor, "6G Internet of Things: A comprehensive survey," IEEE Internet Thing J., vol.9, no.1, pp.359–386, Jan. 2022.
[2] X. Lu, P. Wang, D. Niyato, D.I. Kim, and Z. Han, "Wireless charging technologies: Fundamentals, standards, and network applications," IEEE Commun. Surveys Tuts., vol.18, no.2, pp.1413–1452, 2nd Quart. 2016.
[3] S. Bi, Y. Zeng, and R. Zhang, "Wireless powered communication networks: An overview," IEEE Wireless Commun., vol.23, no.2, pp.10–18, April 2016.
[4] H. Ju and R. Zhang, "Throughput maximization in wireless powered communication networks," IEEE Trans. Wireless Commun., vol.13, no.1, pp.418–428, Jan. 2014.
[5] Q. Wu, M. Tao, D.W. K. Ng, W. Chen, and R. Schober, "Energy-efficient transmission for wireless powered multiuser communication networks," Proc. IEEE Int. Conf. Commun. (ICC), London, UK, pp.154–159, June 2015.
[6] L. Liu, R. Zhang, and K. Chua, "Multi-antenna wireless powered communication with energy beamforming," IEEE Trans. Commun., vol.62, no.12, pp.4349–4361, Dec. 2014.
[7] H. Ju, K. Chang, and M.-S. Lee, "In-band full-duplex wireless powered communication networks," Proc. 2015 17th Int. Conf. Adv. Commun. Technol. (ICACT), Pyeong Chang, pp.23–27, July 2015.
[8] Y. Che, J. Xu, L. Duan, and R. Zhang, "Multiantenna wireless powered communication with cochannel energy and information transfer," IEEE Commun. Lett., vol.19, no.12, pp.2266–2269, Dec. 2015.
[9] M.-L. Ku and C.-Y. Lin, "Beamforming and resource allocation for charging power minimization in multiuser wireless-powered networks," IEEE Access, vol.9, pp.136231–136242, 2021.
[10] J. Choi, C. Song, and J. Joung, "Wireless powered information transfer based on zero-forcing for multiuser MIMO systems," IEEE Trans. Veh. Technol., vol.67, no.9, pp.8561–8570, Sept. 2018.
[11] A.A. Zaidi,R. Baldemair, V. Moles-Cases, N. He, K. Werner, and A. Cedergren, "OFDM numerology design for 5G new radio to support IoT, eMBB, and MBSFN," IEEE Commun. Standards Mag., vol.2, no.2, pp.78–83, June 2018.
[12] T. Nguyen, Q. Pham, V. Nguyen, J. Lee, and Y. Kim, "Resource allocation for energy efficiency in OFDMA-enabled WPCN," IEEE Wireless Commun. Lett., vol.9, no.12, pp.2049–2053, Dec. 2020.
[13] X. Zhou, C.K. Ho, and R. Zhang, "Wireless power meets energy harvesting: A joint energy allocation approach in OFDM-based system," IEEE Trans. Wireless Commun., vol.15, no.5, pp.3481–3491, May 2016.
[14] F. Wang and X. Zhang, "Resource allocation for wireless power transmission enabled full-duplex OFDMA mobile wireless networks," Proc. IEEE Global Commun. Conf. (GLOBECOM), Abu Dhabi, UAE, pp.1–6, Dec. 2018.
[15] X. Zhang and F. Wang, "Resource allocation for wireless power transmission over full-duplex OFDMA/NOMA mobile wireless networks," IEEE J. Sel. Areas Commun., vol.37, no.2, pp.327–344, Feb. 2019.
[16] Y. Zeng and R. Zhang, "Optimized training for net energy maximization in multi-antenna wireless energy transfer over frequency-selective channel," IEEE Trans. Commun., vol.63, no.6, pp.2360–2373, June 2015.
[17] N. Lee, H.J. Yang, and J. Chun, "Achievable sum-rate maximizing AF relay beamforming scheme in two-way relay channels," Proc. IEEE Int. Conf. Commun. Workshops, Beijing, May 2008.
[18] H.H. Kha, H.D. Tuan, and H.H. Nguyen, "Fast global optimal power allocation in wireless networks by local D.C. programming," IEEE Trans. Wireless Commun., vol.11, no.2, pp.510–515, Feb. 2012.
[19] S. Diamond and S. Boyd, "CVXPY: A python-embedded modeling language for convex optimization," J. Machine Learning Research, vol.17, pp.2909–2913, 2016.
[20] C.Y. Ng and C.W. Sung, "Low complexity subcarrier and power allocation for utility maximization in uplink OFDMA systems," IEEE Trans. Wireless Commun., vol.7, no.5, pp.1667–1675, May 2008.
[21] X. Lu, P. Wang, D. Niyato, D.I. Kim, and Z. Han, "Wireless networks with RF energy harvesting: A contemporary survey," IEEE Commun. Surveys Tuts., vol.17, no.2, pp.757–789, 2nd Quart., 2015.
[22] FCC Rules and Regulations Part 15 Section 247, Operation within the bands 902–928 MHz, 2400–2483.5 MHz, and 5725–5850 MHz.

**Keigo Hirashima** received the B.Eng. degree in electrical and electronic engineering from Ibaraki University, Hitachi, Japan in 2022. He is currently a M.Eng. candidate at Ibaraki University. His research interests include signal processing for wireless communications.



**Teruyuki Miyajima** received B.Eng., M.Eng., and Ph.D. in electrical engineering from Saitama University, Saitama, Japan, in 1989, 1991, and 1994, respectively. In 1994, he joined Ibaraki University, Hitachi, Japan, where he is currently a professor in the domain of electrical and electronic systems engineering. His current interests include signal processing for wireless communications. Dr. Miyajima is an IEEE member.