IEICE TRANS. FUNDAMENTALS, VOL.E102–A, NO.8 AUGUST 2019

PAPER

Adaptive FIR Filtering for PAPR Reduction in OFDM Systems

Hikaru MORITA[†*], ***Student Member***, **Teruyuki MIYAJIMA**[†a)], ***and*** **Yoshiki SUGITANI**[†], ***Members***

SUMMARY This study proposes a Peak-to-Average Power Ratio (PAPR) reduction method using an adaptive Finite Impulse Response (FIR) filter in Orthogonal Frequency Division Multiplexing systems. At the transmitter, an iterative algorithm that minimizes the p-norm of a transmitted signal vector is used to update the weight coefficients of the FIR filter to reduce PAPR. At the receiver, the FIR filter used at the transmitter is estimated using pilot symbols, and its effect can be compensated for by using an equalizer for proper demodulation. Simulation results show that the proposed method is superior to conventional methods in terms of the PAPR reduction and computational complexity. It also shows that the proposed method has a trade-off between PAPR reduction and bit error rate performance.

key words: *OFDM, PAPR, p-norm, adaptive algorithm, channel estimation*

1. Introduction

Orthogonal Frequency Division Multiplexing (OFDM) has been widely adopted in various wireless standards, such as IEEE 802.11 [1] and LTE-Advanced [2], because it can achieve a reliable high-rate transmission over multipath fading channels [3]. However, a major disadvantage of OFDM is its high Peak-to-Average Power Ratio (PAPR), which causes a nonlinear distortion and low power efficiency at a nonlinear power amplifier. It can be concluded that PAPR reduction is an important issue in OFDM systems [4].

The simplest and most popular PAPR reduction method is clipping and filtering, which involves clipping the high peaks of an OFDM signal. A major drawback of this method is the generation of both in-band and out-of-band distortion, leading to an increase in Bit Error Rate (BER) and degradation in the spectral efficiency. So far, many researchers have developed various distortionless methods [5]–[8]. These distortionless methods comprise complex problems such as combinatorial problems or Quadratic Programming (QP) problems. Therefore, their common drawback is a high computational complexity to achieve a significant PAPR reduction.

An effective approach to reduce the complexity is to employ a gradient descent algorithm, which is a standard tool to solve optimization problems [8]–[12]. In general, they can provide a suboptimal solution by using a simple iterative procedure. In [8], [9], algorithms to obtain a peak reduction signal for Tone Reservation (TR) [8] were proposed. Since the reduction signal conveys no information, spectral efficiency is reduced and more transmit signal power is required. In [10], PAPR is reduced by iteratively finding a signal indicating the antenna position in Space Shift Keying (SSK). It requires a number of antennas, which makes it unsuitable for small terminals. In [11], the phase factors of Partial Transmit Sequences (PTS) [5] are determined by the Constant Modulus Algorithm (CMA). However, the minimum of the CMA cost function is not necessarily coincident with that of PAPR [12]. Furthermore, PTS is required to send phase factors as Side Information (SI), which causes a loss in transmission data rate and BER degradation due to an erroneous SI detection. In [12], the phase factors are determined by iteratively minimizing the p-norm of a transmitted signal vector. This method avoids sending SI by blindly estimating the phase factors at the receiver. However, it is difficult to apply this method to the higher-order QAM constellations because of the estimation difficulty of the phase factors.

Another interesting approach is to employ adaptive filtering. In [13], an adaptive FIR filter is placed before Inverse Discrete Fourier Transform (IDFT) and its filter coefficients are adjusted by a sequential QP algorithm. This method requires sending the filter coefficients as SI, which consumes transmission power. In addition, it is impractical to send continuous complex coefficients precisely. In [14], a single adaptive All-Pass Filter (APF) is placed after IDFT and its parameter is adjusted using the Newton–Raphson algorithm. Interestingly, this method does not require SI because the composite channel comprising APF and a physical channel is estimated using the pilot symbols at the receiver. The drawbacks of the method are, first, this method requires extra processing to stabilize the filter; second, the resulting impulse response of APF can be longer than a Cyclic Prefix (CP). As shown later, since this method has only one adjustable parameter, its PARP reduction capability is limited.

In this study, we propose a PAPR reduction method using an adaptive FIR filter, which is always stable and has an impulse response shorter than CP. The weight coefficients of the FIR filter are updated by iteratively minimizing the p-norm of a transmitted signal vector. It does not require sending any additional reduction signal or SI. At the receiver, either the FIR filter used at the transmitter or a composite channel is estimated using the pilot symbols, and the effect of the FIR filter can be compensated for by an equalizer. This method can be applied to the higher-order QAM constellations without any modification.

Throughout this paper, we use the following notations:

Manuscript received November 22, 2018.
Manuscript revised March 14, 2019.

[†]The authors are with the Graduate School of Science and Engineering, Ibaraki University, Hitachi-shi, 316-8511 Japan.
[*]Presently, with Denso Wave Inc.
a) E-mail: teruyuki.miyajima.spc@vc.ibaraki.ac.jp

DOI: 10.1587/transfun.E102.A.938

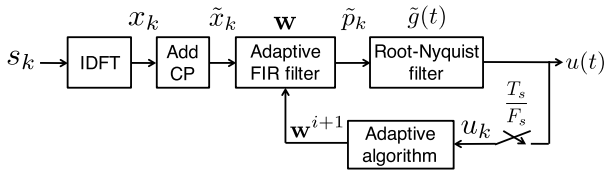


Fig. 1 Transmitter structure.

$(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^*$ represent the transpose, Hermitian transpose, and complex conjugate of a vector or matrix, respectively. $E[\cdot]$ denotes the expectation operator. diag($\mathbf{a}$) represents a diagonal matrix whose diagonal entries consist of the entries of vector $\mathbf{a}$.

2. System Model

2.1 Transmitter

Figure 1 shows the transmitter structure of the proposed system equipped with a single transmit antenna. Data symbols s_k are generated every T_s seconds, and N data symbols are grouped into a data block $\mathbf{s}_n = [s_{nN}\ s_{nN+1}\ \cdots\ s_{(n+1)N-1}]^T$. A time domain signal of the nth block is obtained by taking IDFT of $\mathbf{s}_n$ as

$$\mathbf{x}_n = \begin{bmatrix} x_{nN} & x_{nN+1} & \cdots & x_{(n+1)N-1} \end{bmatrix}^T = \mathbf{F}^H \mathbf{s}_n \tag{1}$$

where $\mathbf{F}$ denotes the N-point DFT matrix whose (k,l)th element is $\frac{\exp(-j2\pi(k-1)(l-1)/N)}{\sqrt{N}}$. A CP of length P is added to $\mathbf{x}_n$ to obtain a discrete time OFDM signal $\tilde{\mathbf{x}}_n = \begin{bmatrix}\tilde{x}_{nQ}\ \tilde{x}_{nQ+1}\ \cdots\ \tilde{x}_{(n+1)Q-1}\end{bmatrix}^T = [x_{(n+1)N-P}\ \cdots\ x_{(n+1)N-1}\ x_{nN}\ \cdots\ x_{(n+1)N-1}]^T$ for the nth block of length $Q = N + P$.

In the proposed system, the PAPR of the transmit signal is reduced by an adaptive FIR filter with input $\tilde{x}_k$. The output of the FIR filter for the nth block is given by

$$\tilde{p}_k = \sum_{l=0}^{L_w} w_l^{(n)} \tilde{x}_{k-l}, \qquad nQ \le k < (n+1)Q \tag{2}$$

where L_w is the FIR filter order and $w_l^{(n)}$ is a weight coefficient of the nth block, which is adjusted by an adaptive algorithm described later. According to [15], a continuous-time signal $u(t)$ is obtained by the root Nyquist filter with input $\tilde{p}_k$ as

$$u(t) = \sum_k \tilde{p}_k \tilde{g}(t - kT_s) \tag{3}$$

where $\tilde{g}(t)$ is a delayed version of a truncated root Nyquist pulse, defined as

$$\tilde{g}(t) = \begin{cases} \bar{g}(t - T_o) & 0 \le t \le 2T_o \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $T_o = MT_s$ is a delay, M is a pulse truncation length, and $\bar{g}(t)$ is the root Nyquist pulse. The continuous signal $u(t)$ is sent to the receiver.

To approximate the PAPR of a continuous-time transmitted signal $u(t)$, we consider an oversampled signal given by $u_k = u(kT_s/F_s)$ where F_s, which is an integer, is the oversampling factor. The PAPR of the transmitted signal is expressed as

$$\text{PAPR} = \frac{\max_{nF_sQ \le k < (n+1)F_sQ} |u_k|^2}{E\left[|u_k|^2\right]}. \tag{5}$$

Because a high PAPR causes a nonlinear distortion and low power efficiency, our purpose is to adjust the FIR filter coefficients using an adaptive algorithm such that the PAPR is reduced.

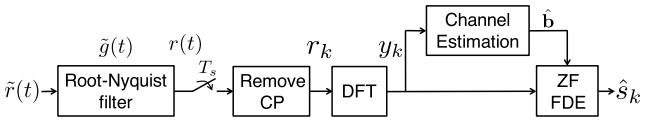


Fig. 2 Receiver structure.

2.2 Receiver

Figure 2 shows the receiver structure of the proposed system. The transmitted signal $u(t)$ is passed through a physical channel whose impulse response is $h(t) = \sum_{l=0}^{L_h} h_l \delta(t - lT_s)$, and is corrupted by additive white Gaussian noise, where L_h is the order of the channel. The received signal is filtered by the root-Nyquist filter $\tilde{g}(t)$ in (4) and its output $r(t)$ is sampled every T_s seconds to obtain a discrete-time signal as

$$r_k = r(kT_s) = \sum_{l=0}^{L_c} c_l \tilde{p}_{k-l} + z_k \tag{6}$$

where c_l is the impulse response of a composite channel of order $L_c = L_h + 4M$, which consists of the physical channel and Nyquist filter, and is given by

$$c_k = \sum_{l=0}^{L_h} h_l g_{k-l} \tag{7}$$

where $g_k = g(kT_s), k = 0, 1, \cdots, L_g$, $L_g = 4M$, $g(t) = \tilde{g}(t) * \tilde{g}(t)$ is the delayed version of the Nyquist pulse truncated within $[0, L_gT_s]$ and z_k is channel noise. Furthermore, by substituting (2) into (6), we have

$$r_k = \sum_{l=0}^{L_b} b_l^{(n)} \tilde{x}_{k-l} + z_k, \qquad nQ \le k < (n+1)Q \tag{8}$$

where

$$b_k^{(n)} = \sum_{l=0}^{L_c} c_l w_{k-l}^{(n)} \tag{9}$$

is the impulse response of the total channel for the nth block of order $L_b = L_w + L_g + L_h$, which consists of the FIR filter,

Nyquist filter, and physical channel. To prevent inter-block interference, the CP length P is larger than or equal to L_b.

The received samples after removing CP are written as

$$\mathbf{r}_n = \left[r_{nP}\ r_{nP+1}\ \cdots\ r_{(n+1)Q-1}\right]^T = \mathbf{B}_n\mathbf{x}_n + \mathbf{z}_n \quad (10)$$

where $\mathbf{B}_n$ is an $N \times N$ circulant matrix of the total channel $b_k^{(n)}$ and $\mathbf{z}_n$ is a noise vector. By applying DFT to $\mathbf{r}_n$, we obtain the frequency domain signal of the nth block as

$$\mathbf{y}_n = \mathbf{F}\mathbf{r}_n = \mathbf{F}\mathbf{B}_n\mathbf{F}^H\mathbf{s}_n + \mathbf{F}\mathbf{z}_n = \mathbf{D}_n\mathbf{s}_n + \mathbf{F}\mathbf{z}_n \quad (11)$$

where $\mathbf{D}_n = \mathbf{F}\mathbf{B}_n\mathbf{F}^H$ is an $N \times N$ diagonal matrix. The diagonal elements of $\mathbf{D}_n$ are identical to $\mathbf{F}_L\mathbf{b}_n$ where $\mathbf{F}_L$ is the $N\times(L_b+1)$ matrix consisting of the first L_b+1 columns of $\sqrt{N}\mathbf{F}$ and $\mathbf{b}_n = [b_0^{(n)}\ b_1^{(n)}\ \cdots\ b_{L_b}^{(n)}]^T$ is the total channel vector of the nth block. If the total channel is known at the receiver, the effect of the channel is compensated by the Zero-Forcing Frequency-Domain Equalizer (ZF-FDE) as

$$\hat{\mathbf{s}}_n = \mathbf{D}_n^{-1}\mathbf{y}_n = \mathbf{s}_n + \mathbf{D}_n^{-1}\mathbf{F}\mathbf{z}_n. \quad (12)$$

The total channel $b_k^{(n)}$ can be estimated by using pilot symbols. Thus, it should be noted that we need not transmit the weight coefficients $w_k^{(n)}$ of the FIR filter as SI because the FIR filter $w_k^{(n)}$ is a part of the total channel $b_k^{(n)}$ as shown in (9). In practice, instead of the total channel $b_k^{(n)}$, it is sufficient to estimate a part of the channel since we have the knowledge of the Nyquist pulse g_k as discussed in Sect. 4. Note that the estimation should be done for each block.

3. PAPR Reduction Adaptive Algorithm

3.1 Transmitted Signal Model

For notational convenience, we represent the transmitted signal samples u_k in a vector form. From (3), we have

$$u_k[m] = u_{kF_s+m} = \sum_{l=0}^{2M} \tilde{g}_l[m]\tilde{p}_{k-l} \quad (13)$$

for $k = nQ, nQ+1, \cdots, (n+1)Q-1$ and $m = 0, 1, \cdots, F_s-1$, where

$$\tilde{g}_k[m] = \tilde{g}\left(kT_s + \frac{m}{F_s}T_s\right). \quad (14)$$

Then, we collect F_s samples in a vector as

$$\tilde{\mathbf{u}}_k = [u_k[0]\, u_k[1]\ \cdots\ u_k[F_s-1]]^T = \sum_{l=0}^{2M} \tilde{\mathbf{g}}_l\tilde{p}_{k-l} \quad (15)$$

where $\tilde{\mathbf{g}}_k = [\tilde{g}_k[0]\, \tilde{g}_k[1]\ \cdots\ \tilde{g}_k[F_s-1]]^T$. All samples corresponding to the nth OFDM block are collected as

$$\begin{aligned}\mathbf{u}_n &= \left[\begin{array}{cccc} u_{nF_sQ} & u_{nF_sQ+1} & \cdots & u_{(n+1)F_sQ-1}\end{array}\right]^T \\ &= \left[\begin{array}{cccc} \tilde{\mathbf{u}}_{nQ}^T & \tilde{\mathbf{u}}_{nQ+1}^T & \cdots & \tilde{\mathbf{u}}_{(n+1)Q-1}^T\end{array}\right]^T \\ &= \tilde{\mathbf{G}}_{\mathrm{o}}\left[\begin{array}{c}\bar{\mathbf{p}}_{n-1}\\ \breve{\mathbf{p}}_n\end{array}\right] = \left[\begin{array}{cc}\bar{\mathbf{G}}_{\mathrm{o}} & \breve{\mathbf{G}}_{\mathrm{o}}\end{array}\right]\left[\begin{array}{c}\bar{\mathbf{p}}_{n-1}\\ \breve{\mathbf{p}}_n\end{array}\right] \quad (16)\end{aligned}$$

where

$$\begin{aligned}\tilde{\mathbf{G}}_{\mathrm{o}} &= \left[\begin{array}{ccc|ccc} \tilde{\mathbf{g}}_{2M} & \cdots & \tilde{\mathbf{g}}_1 & \tilde{\mathbf{g}}_0 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \vdots & \tilde{\mathbf{g}}_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \tilde{\mathbf{g}}_{2M} & \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \tilde{\mathbf{g}}_{2M} & \cdots & \tilde{\mathbf{g}}_1 & \tilde{\mathbf{g}}_0 \end{array}\right] \\ &= \left[\begin{array}{c|c}\bar{\mathbf{G}}_{\mathrm{o}} & \breve{\mathbf{G}}_{\mathrm{o}}\end{array}\right], \quad (17)\end{aligned}$$

$$\bar{\mathbf{p}}_{n-1} = \left[\begin{array}{cccc} \tilde{p}_{nQ-2M} & \cdots & \tilde{p}_{nQ-2} & \tilde{p}_{nQ-1}\end{array}\right]^T, \quad (18)$$

$$\breve{\mathbf{p}}_n = \left[\begin{array}{cccc} \tilde{p}_{nQ} & \tilde{p}_{nQ+1} & \cdots & \tilde{p}_{(n+1)Q-1}\end{array}\right]^T. \quad (19)$$

From (2), we have

$$\breve{\mathbf{p}}_n = \tilde{\mathbf{X}}_n\mathbf{w}_n \quad (20)$$

where $\mathbf{w}_n = [w_0^{(n)}\ w_1^{(n)}\ \cdots\ w_{L_w}^{(n)}]$ and

$$\tilde{\mathbf{X}}_n = \left[\begin{array}{cccc} \tilde{x}_{nQ} & \tilde{x}_{nQ-1} & \cdots & \tilde{x}_{nQ-L_w} \\ \tilde{x}_{nQ+1} & \tilde{x}_{nQ} & \cdots & \tilde{x}_{nQ+1-L_w} \\ \vdots & \ddots & \ddots & \vdots \\ \tilde{x}_{(n+1)Q-1} & \tilde{x}_{(n+1)Q-2} & \cdots & \tilde{x}_{(n+1)Q-(1+L_w)}\end{array}\right].$$

Then, $\mathbf{u}_n$ can be expressed as

$$\mathbf{u}_n = \bar{\mathbf{G}}_{\mathrm{o}}\bar{\mathbf{p}}_{n-1} + \breve{\mathbf{G}}_{\mathrm{o}}\tilde{\mathbf{X}}_n\mathbf{w}_n = \bar{\mathbf{G}}_{\mathrm{o}}\bar{\mathbf{p}}_{n-1} + \mathbf{V}_n\mathbf{w}_n \quad (21)$$

where $\mathbf{V}_n = \breve{\mathbf{G}}_{\mathrm{o}}\tilde{\mathbf{X}}_n$. Note that $\bar{\mathbf{p}}_{n-1}$ has been determined in the $n-1$th block and is not influenced by the determination of the FIR filter $\mathbf{w}_n$ in the nth block. Thus, when we determine $\mathbf{w}_n$, $\bar{\mathbf{p}}_{n-1}$ is fixed.

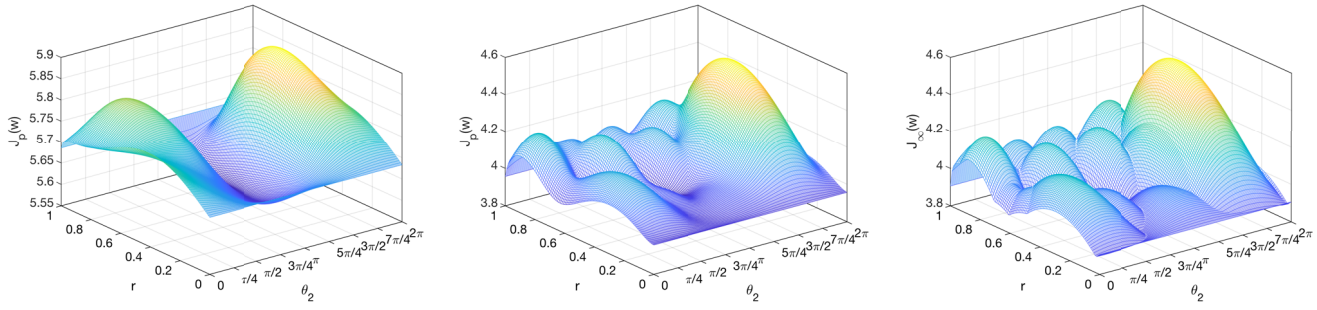
3.2 Cost Function and Adaptive Algorithm

If the average power $E[|u_k|^2]$ can be kept constant, the PAPR of u_k is identical to the infinity-norm of the transmitted signal vector $\mathbf{u}_n$. Thus, we can employ the infinity-norm of $\mathbf{u}_n$ as the cost function:

$$J_\infty(\mathbf{w}_n) = \|\mathbf{u}_n\|_\infty = \max_{nF_sQ \le k < (n+1)F_sQ} |u_k|. \quad (22)$$

We use a gradient descent algorithm to minimize $J_\infty(\mathbf{w}_n)$ because of its simplicity. We impose the norm constraint $\|\mathbf{w}_n\|_2 = 1$, which keeps the average power constant and prevents $\mathbf{w}_n$ from converging to zero. When a gradient descent algorithm is employed, a convex cost function is desirable. However, the infinity norm cost function generally has local minima.

Let us visualize the cost function $J_\infty(\mathbf{w}_n)$. We consider an FIR filter with two weight coefficients $\mathbf{w}_n = [w_0^{(n)}\ \ w_1^{(n)}]^T$. Since $J_\infty(\mathbf{w}_n)$ is invariant to a phase rotation of $\mathbf{w}_n$ and $\|\mathbf{w}_n\|_2 = 1$ holds, we can parameterize $\mathbf{w}_n$ as $\mathbf{w}_n = \left[\begin{array}{cc} r & \sqrt{1-r^2}e^{j\theta_2}\end{array}\right]^T$ where $0 \le r \le 1$, $0 \le \theta_2 \le 2\pi$. Figure 3(c) plots an example of $J_\infty(\mathbf{w}_n)$ as a function of r



(a) 8 - norm (b) 64 - norm (c) ∞ - norm

Fig. 3 Examples of cost function.

and θ_2. As can be seen from this figure, the infinity-norm cost function $J_\infty(\mathbf{w}_n)$ is non-smooth and has multiple local minima. Then, a gradient descent algorithm may be trapped in an undesired minimum.

An approach to overcome this issue is to approximate the infinity-norm cost function by the p-norm cost function given by

$$J_p(\mathbf{w}_n) = \|\mathbf{u}_n\|_p = \left(\sum_{k=nF_sQ}^{(n+1)F_sQ-1} |u_k|^p \right)^{\frac{1}{p}}. \tag{23}$$

This approximation has been successfully applied to the TR method [9], the SLM method [10], and the PTS method [12]. Although the p-norm cost function is convex, local minima can exist due to the norm constraint $\|\mathbf{w}_n\|_2 = 1$. Figures 3(a) and (b) show the examples of $J_p(\mathbf{w}_n)$ for p =8 and 64, respectively. It can be observed that the smaller p is, the smoother the cost function is. As p decreases, the shape of $J_p(\mathbf{w}_n)$ deviates from that of $J_\infty(\mathbf{w}_n)$. This observation suggests that $J_p(\mathbf{w}_n)$ with a properly chosen p well approximates $J_\infty(\mathbf{w}_n)$. The weight coefficients obtained by minimizing $J_p(\mathbf{w}_n)$ might be suboptimal but can be expected to reduce PAPR significantly.

The updation of the weight coefficients of the nth block is performed by the gradient descent method given by

$$\begin{aligned} \bar{\mathbf{w}}_n^{i+1} &= \mathbf{w}_n^i - \mu \frac{\partial J_p(\mathbf{w}_n)}{\partial \mathbf{w}_n^*} \\ &= \mathbf{w}_n^i - \frac{\mu}{2} (\|\mathbf{u}_n\|_p)^{1-p} \sum_{k=nF_sQ}^{(n+1)F_sQ-1} \mathbf{v}_k^H u_k |u_k|^{p-2} \end{aligned} \tag{24}$$

$$\mathbf{w}_n^{i+1} = \bar{\mathbf{w}}_n^{i+1} / \|\bar{\mathbf{w}}_n^{i+1}\|_2 \tag{25}$$

where μ is a step size and $\mathbf{v}_k$ is the $(k - nF_sQ + 1)$th row of $\mathbf{V}_n$. Given an initial value $\mathbf{w}_n^0$, the weight coefficients are updated according to (24) and (25) until they converge or the predetermined number of iterations is reached.

4. Channel Estimation

As mentioned in Sect. 2, knowledge of the total channel $b_k^{(n)}$ for each block is required for data demodulation. We consider time-domain channel estimation by Least Squares (LS) using pilot symbols known to the receiver. Let $\bar{\mathbf{s}}$ be an $N_\mathrm{p} \times 1$ vector composed of N_p pilot symbols. The DFT outputs corresponding to the pilot symbols are extracted from (11) and are expressed as

$$\bar{\mathbf{y}}_n = \bar{\mathbf{S}} \bar{\mathbf{F}}_L \mathbf{b}_n + \bar{\mathbf{z}}_n \tag{26}$$

where $\bar{\mathbf{S}} = \mathrm{diag}(\bar{\mathbf{s}})$, $\bar{\mathbf{F}}_L$ is the $N_\mathrm{p} \times (L_b + 1)$ matrix obtained by extracting N_p rows corresponding to pilot symbols from $\mathbf{F}_L$, and $\bar{\mathbf{z}}_n$ is a noise vector. The direct LS estimation of the total channel $\mathbf{b}_n$ needs at least $L_b + 1$ pilot symbols. However, L_b can be large, which can result in wastage of the bandwidth. To reduce the number of required pilot symbols, we consider the two channel estimation methods.

4.1 Estimation of Composite Channel

The total channel $b_k^{(n)}$ is composed of several subsystems as shown in Fig. 4. Since the Nyquist pulse g_k is known to the receiver, it is sufficient to estimate a composite channel $a_k^{(n)}$ composed of the FIR filter $w_k^{(n)}$ and the physical channel h_k. First, we consider estimating the composite channel whose impulse response is given by

$$\mathbf{a}_n = \begin{bmatrix} a_0^{(n)} & a_1^{(n)} & \cdots & a_{L_a}^{(n)} \end{bmatrix}^T = \mathbf{H}\mathbf{w}_n \tag{27}$$

where $L_a = L_h + L_w$ is the order of the channel $a_k^{(n)}$ and $\mathbf{H}$ is a Toeplitz matrix

$$\mathbf{H} = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ h_{L_h} & & & h_0 \\ 0 & h_{L_h} & & \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & h_{L_h} \end{bmatrix} \in \mathbb{C}^{(L_a+1)\times(L_w+1)}. \tag{28}$$

Then, the total channel becomes $\mathbf{b}_n = \mathbf{G}\mathbf{a}_n$ where $\mathbf{G}$ is also a Toeplitz matrix of size $(L_b+1)\times(L_a+1)$, whose first column is $[g_0\ g_1\ \cdots\ g_{L_g}\ 0\ \cdots\ 0]^T$ and first row is $[g_0\ 0\ \cdots\ 0]$. The LS estimation is given by

$$\hat{\mathbf{a}}_n = \arg\min_{\tilde{\mathbf{a}}_n} \|\bar{\mathbf{y}}_n - \mathbf{A}_a \tilde{\mathbf{a}}_n\|_2^2 \tag{29}$$

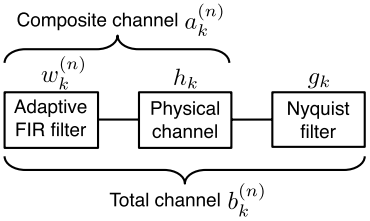


Fig. 4 Total channel composed of subsystems.

where $\mathbf{A}_a = \bar{\mathbf{S}}\bar{\mathbf{F}}_L\mathbf{G} \in \mathbb{C}^{N_\mathrm{p}\times(L_a+1)}$. By solving (29), the impulse response is obtained by

$$\hat{\mathbf{a}}_n = (\mathbf{A}_a^H\mathbf{A}_a)^{-1}\mathbf{A}_a^H\bar{\mathbf{y}}_n. \tag{30}$$

Since $\mathbf{A}_a$ must have the full column rank, the required number of pilot symbols in each OFDM block is

$$N_\mathrm{p} \geq L_a + 1 = L_h + L_w + 1. \tag{31}$$

4.2 Estimation of FIR Filter

When the physical channel h_k changes very slowly, i.e., it can be considered to remain constant for a long period, it is sufficient to estimate only the FIR filter coefficients $w_k^{(n)}$ instead of $a_k^{(n)}$. The DFT outputs in (26) can be rewritten as $\bar{\mathbf{y}}_n = \bar{\mathbf{S}}\bar{\mathbf{F}}_L\mathbf{G}\mathbf{H}\mathbf{w}_n + \bar{\mathbf{z}}_n$. Similar to (29), we have the LS estimation of $\mathbf{w}_n$ as

$$\hat{\mathbf{w}}_n = (\mathbf{A}_w^H\mathbf{A}_w)^{-1}\mathbf{A}_w^H\bar{\mathbf{y}}_n \tag{32}$$

where $\mathbf{A}_w = \bar{\mathbf{S}}\bar{\mathbf{F}}_L\mathbf{G}\mathbf{H} \in \mathbb{C}^{N_\mathrm{p}\times(L_w+1)}$. The required number of the pilot symbols in each OFDM block for $\mathbf{A}_w$ to have the full column rank is

$$N_\mathrm{p} \geq L_w + 1. \tag{33}$$

The estimation (32) requires fewer pilot symbols than the estimation (30) and thus, improves the bandwidth efficiency reduction owing to the use of pilot symbols.

When we use the estimation method (32), we need to estimate the physical channel h_k in advance. During the estimation period for h_k, we should fix the value of $w_k^{(0)}$ to a constant value. The number of pilot symbols for the estimation of h_k is $L_h + 1$. Note that since it is sufficient to perform the estimation only when the channel changes, the loss due to the pilot symbols is not significant in a quasi-static environment.

5. Simulation Results

Unless otherwise stated, the simulation parameters in Table 1 were used. The channel taps h_k and the noise z_k were generated as complex Gaussian random variables. We set the initial values as $\mathbf{w}_n^0 = [1\ 0\ \cdots\ 0]^T$ and $\bar{\mathbf{p}}_0 = \mathbf{0}$.

In Fig. 5, the convergence performances for various step sizes are shown. The vertical axis represents the PAPR that achieves Complementary Cumulative Distribution Functions (CCDF) of 10^{-3}, i.e., PAPR_0 of $\Pr(\mathrm{PAPR} > \mathrm{PAPR}_0) = 10^{-3}$. In the case of $\mu = 10^{-1}$, PAPR is not reduced at all. PAPR converges fastest when $\mu = 10^{-2}$ and reaches 4.9 dB at 10^3 iterations.

Figure 6 shows the effect of the FIR filter order L_w for various values of p. It is seen that the PAPR reduces as the FIR filter order increases. When $p = 8$, the performance is the worst. As mentioned in Sect. 3.2, this is because the cost function with $p = 8$ differs greatly from the infinity norm. The performance of $p = \infty$ also degrades because of

Table 1 Simulation parameters.

Modulation scheme	16QAM
Number of subcarriers N	256
CP length P	64
Oversampling factor F_s	4
p-norm	64-norm
Number of iterations I	10^3
Step size μ	10^{-2}
FIR filter order L_w	20
Data symbol interval T_s	10^{-7}
Pulse truncation length M	8
Roll-off factor β	0.5
Channel order L_h	10
SSPA parameter ρ	2

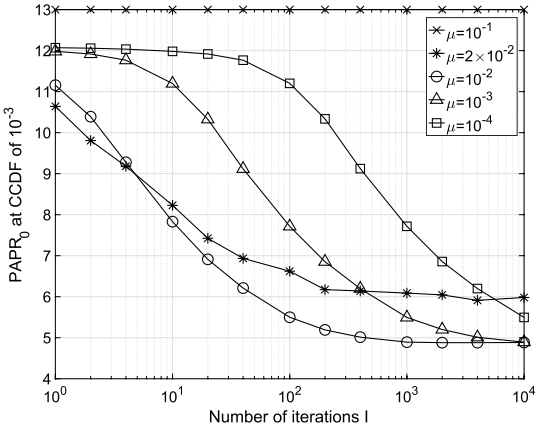


Fig. 5 Convergence performance.

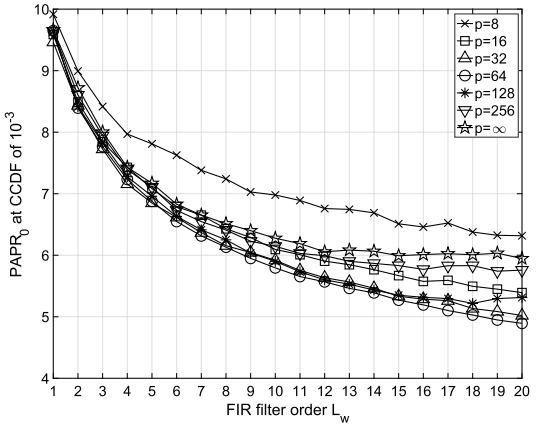


Fig. 6 Effect of FIR filter order L_w and p.

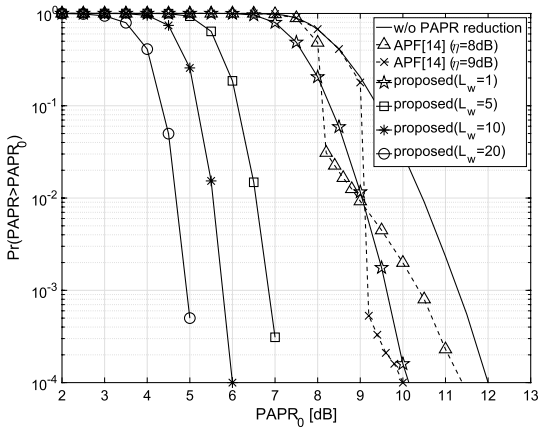


Fig. 7 CCDFs of PAPR.

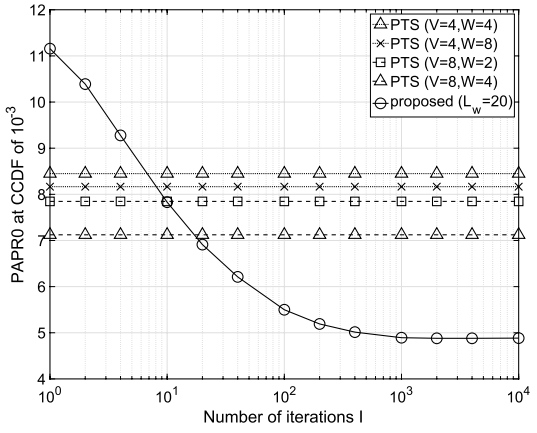


Fig. 8 Comparison with PTS.

the convergence to an undesired local minimum. The best performance is obtained when $p = 64$. The best p depends on conditions such as modulation scheme and the number of subcarriers. The determination of an appropriate p is our future research topic.

In Fig. 7, we show the CCDF of PAPR of the proposed method for various values of L_w. It is seen that the PAPR at CCDF of 10^{-3} achieved by the proposed method with $L_w = 20$ is approximately 6.5 dB lower than that without PAPR reduction. For the sake of comparison, results of the conventional APF method with η=8 dB, 9 dB, and $\lambda = 0.9$ [14] are also shown. The conventional APF method has only one adjustable parameter and its complexity found in [14] is slightly lower than that of the proposed method. However, as shown in this figure, its PAPR performance is very limited compared to the proposed method.

We compare the proposed method with PTS [5]. In Fig. 8, the PAPR for CCDF of 10^{-3} of PTS is shown, where the number of disjoint sub-blocks is $V = 4, 8$ and the number of allowed phase angles is $W = 2, 4, 8$, and the PAPR curve of the proposed method is same as that in Fig. 5. The larger V or W is, the lower PAPR of PTS is. The proposed method can achieve significantly lower PAPR than PTS.

The computational complexity (the number of multiplications) of the PTS C_{PTS} and the proposed method C_{FIR} are given by

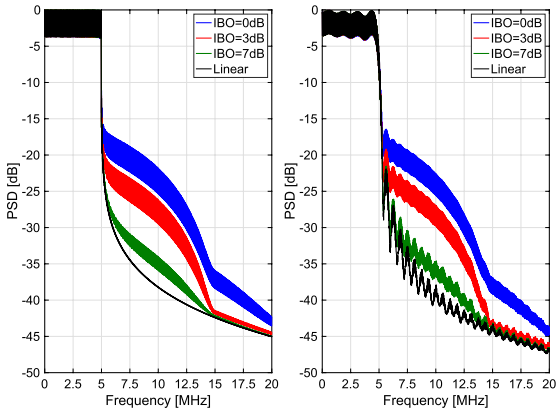
$$C_{\text{PTS}} = VN \log_2 N + W^{V-1}(VF_sN + F_sN), \tag{34}$$

$$C_{\text{FIR}} = N \log_2 N + I\{F_sQ(2q + L_w + 5) + 3(L_w + 1) + 2\}, \tag{35}$$

where we set $p = 2^q$. Table 2 shows C_{PTS} and C_{FIR} where the number of iterations I in (35) is identified from Fig. 8 such that the achieved PAPR of the proposed method at I is almost equal to that of PTS. The complexity of the proposed method is remarkably lower than that of PTS. We could not obtain the results for $V > 8$ or $W > 8$ owing to the huge complexity of PTS.

Table 2 Complexity comparison.

(V, W)	C_{PTS}	I	C_{FIR}
(4 , 4)	3.4×10^5	7	3.3×10^5
(4 , 8)	2.6×10^6	8	3.8×10^5
(8 , 2)	1.2×10^6	10	4.8×10^5
(8 , 4)	1.5×10^8	20	9.5×10^5



(a) PTS($V = 8, W = 2$) (b) proposed($I = 10$)

Fig. 9 PSD performance.

Figure 9 shows the Power Spectrum Density (PSD) performances of PTS and the proposed method. The PSD of the proposed method fluctuates and is slightly worse than that of PTS around 5 to 10 MHz when a linear amplifier is used. This is because the proposed method uses a truncated root Nyquist filter. If the pulse truncation length M can be lengthened, the fluctuation can be reduced. In Fig. 9, PSDs are also shown when we use a Solid State Power Amplifier (SSPA) whose AM/AM characteristic is $g(x) = x\left(1 + (x/A)^{2\rho}\right)^{-1/2\rho}$ where A is the saturation level and ρ controls the sharpness of the AM/AM curve, and input backoff (IBO) is 0,3, and 7 dB. Except for the fluctuation of the proposed method, the PSDs of the proposed method are almost the same as those of PTS.

In Fig. 10, we show the relationship between PAPR and BER where the perfect channel estimation is assumed and $E_b/N_0 = 40$ dB. Evidently, the proposed method has a trade-off between PAPR and BER. If the iteration is stopped when a required BER is achieved, the resulting PAPR of the FIR

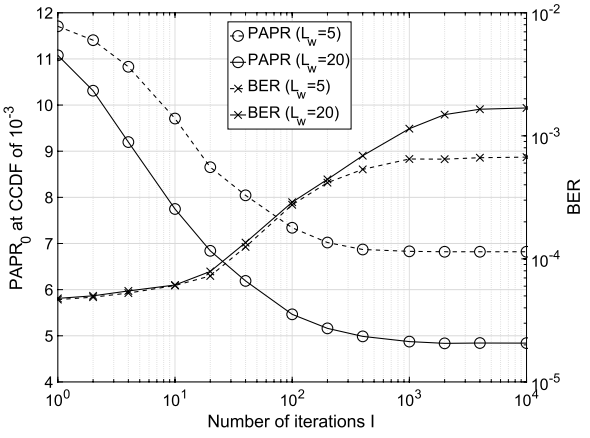


Fig. 10 Relationship between PAPR and BER.

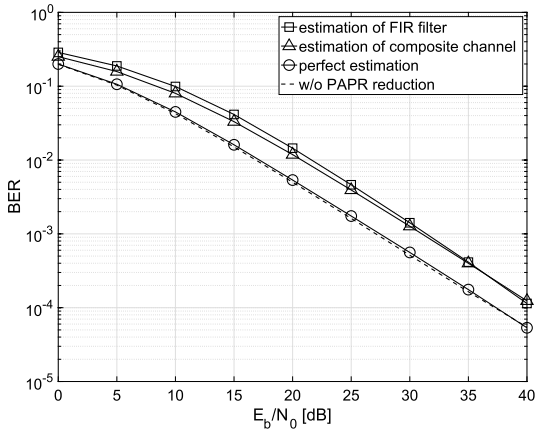


Fig. 12 Influence of channel estimation on BER.

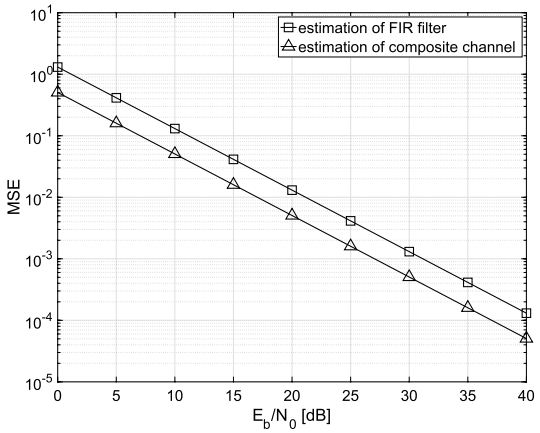


Fig. 11 MSE performance.

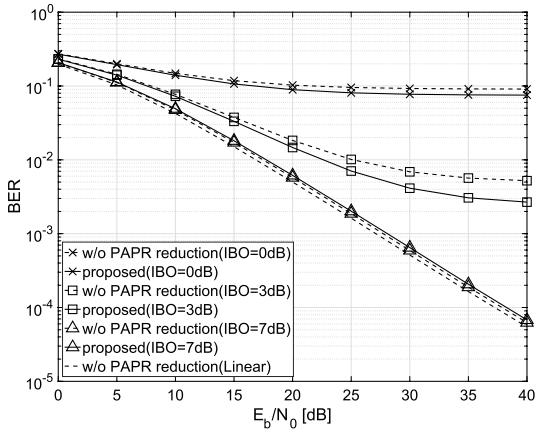


Fig. 13 BER performance when SSPA is used.

filter with a larger order L_w is lower than that with a smaller order. For example, when BER=10^{-4}, PAPR_0 of $L_w = 5$ and 20 are 8.3 dB and 6.4 dB, respectively. However, as shown in (33), the number of pilot symbols required for estimation is increased for a large FIR filter order.

Figure 11 shows the Mean Squared Error (MSE) performance of the channel estimation methods discussed in Sect. 4, where the number of iterations $I = 10$. MSE is defined by

$$\mathrm{MSE} = \|\hat{\mathbf{b}}_n - \mathbf{b}_n\|_2^2 / \|\mathbf{b}_n\|_2^2, \tag{36}$$

where $\hat{\mathbf{b}}_n = \mathbf{G}\hat{\mathbf{a}}_n$ for (30) and $\hat{\mathbf{b}}_n = \mathbf{GH}\hat{\mathbf{w}}_n$ for (32) with the perfect knowledge of h_k. The number of pilot symbols is $N_\mathrm{p} = L_a + 1 = 31$ for (30) and $N_\mathrm{p} = L_w + 1 = 21$ for (32). In both cases, MSE decreases as E_b/N_0 increases. MSE of the composite channel estimation is lower than that of the FIR filter estimation because the former uses more pilot symbols.

Figure 12 shows the influence of the channel estimation on BER. It is observed that if the channel is perfectly estimated, the degradation of the proposed method is insignificant when compared to the case without PAPR reduction. When the estimated channels are used, the performance degrades by approximately 4 dB. Although it can be improved by using more pilot symbols, this results in a spectral efficiency loss. We also observe that although the number of pilot symbols required for the FIR filter estimation (32) is fewer than that for the composite channel estimation (30), their BER performances are almost identical to each other.

In Fig. 13, the BER performances are shown when SSPA is used. When IBO is 0 and 3 dB, the proposed method shows superior performance compared to the original OFDM without PAPR reduction. When IBO is 7 dB, the BER of the proposed method is almost the same as that without PAPR reduction.

Finally, let us explain the limitations of the proposed method. First, as shown in Fig. 6, PAPR performance improves as the FIR filter order L_w increases. However, the increase of L_w leads to the increase of both the number of pilot symbols N_p in (33) and the computational complexity C_FIR in (35). Second, as shown in Fig. 10, the BER performance deteriorates as the PAPR performance improves. It is desirable that BER is independent of PAPR. It is worth investigating how to overcome these limitations.

6. Conclusion

In this study, we proposed an adaptive FIR filter based PAPR

reduction method in the OFDM systems. The weight coefficients of the FIR filter are determined by minimizing the p-norm of the transmitted signal vector by a gradient descent method. It was shown that PAPR significantly reduces as the FIR filter order increases if p is chosen properly. Moreover, the proposed method achieves a lower PAPR compared to the PTS method, while the complexity is significantly reduced. A simulation result suggested that we should take into account not only PAPR, but also BER. We presented two channel estimation methods using pilot symbols. It was shown that the BER performance using the estimated channel is close to the ideal case.

Improving the channel estimation performance with fewer pilot symbols is a challenging prospect. One approach is to employ a blind channel estimation method [16]. However, the blind method is too complex computationally to be implemented in small terminals. It is worth developing a simple blind method.

References

[1] B. Bellalta, L. Bononi, R. Bruno, and A. Kassler, "Next generation IEEE 802.11 wireless local area networks: Current status, future directions and open challenges," Comput. Commun., vol.75, pp.1–25, Feb. 2016.
[2] S. Parkvall, A. Furuskar, and E. Dahlman, "Evolution of LTE toward IMT-advanced," IEEE Commun. Mag., vol.49, no.2, pp.84–91, Feb. 2011.
[3] T. Hwang, C. Yang, G. Wu, S. Li, and G.Y. Li, "OFDM and its wireless applications: A survey," IEEE Trans. Veh. Technol., vol.58, no.4, pp.1673–1694, May 2009.
[4] S.H. Han and J.H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," IEEE Trans. Wireless Commun., vol.12, no.2, pp.56–65, April 2005.
[5] S.H. Müller and J.B. Huber, "A novel peak power reduction scheme for OFDM," Proc. 1997 IEEE 8th Int. Symp. on Personal, Indoor and Mobile Radio Commun. (PIMRC), pp.1090–1094, 1997.
[6] R.W. Bauml, R.F.H. Fisher, and J.B. Huber, "Reducing the peak-to-average power ratio of multicarrier modulation by selected mapping," Electron. Lett., vol.32, no.22, pp.2056–2057, Oct. 1996.
[7] B.S. Krongold and D.L. Jones, "PAR reduction in OFDM via active constellation extension," IEEE Trans. Broadcast., vol.49, no.3, pp.258–268, Sept. 2003.
[8] J. Tellado and J.M. Cioffi, "Peak power reduction for multicarrier transmission," Proc. IEEE CTMC GLOBECOM'98, pp.219–224, Sydney, Australia, Nov. 1998.
[9] S. Janaaththanan, C. Kasparis, and B.G. Evans, "A gradient based algorithm for PAPR reduction of OFDM using tone reservation technique," Proc. IEEE VTC Spring 2008, pp.2977–2980, May 2008.
[10] P. Yang, Y. Xiao, and S. Li, "An improved gradient-based PAPR reduction method for space shift keying (SSK)-OFDM systems," IEICE Trans. Commun., vol.E94-B, no.12, pp.3532–3539, Dec. 2011.
[11] S. Khademi and A.-J. van der Veen, "Constant modulus algorithm for peak-to-average power ratio (PAPR) reduction in MIMO OFDM/A," IEEE Signal Process. Lett., vol.20, no.5, pp.531–534, May 2013.
[12] M. Yoshida, H. Nashimoto, and T. Miyajima, "PTS-based PAPR reduction by iterative p-norm minimization without side information in OFDM systems," IEICE Trans. Commun., vol.E101-B, no.3, pp.856–864, March 2018.
[13] A. Yao and Y. Zheng, "Peak-to-average power reduction of OFDM signaling using adaptive digital filter," Proc. IEEE ICASSP, pp.IV-305–308, Toulouse, France, May 2006.
[14] E. Hong, H. Kim, and D. Har, "PAPR reduction by a single adaptive all-pass filter for OFDM systems," IEICE Electron. Express, vol.8, no.19, pp.1633–1639, Oct. 2011.
[15] Y.-P. Lin and S.-M. Phoong, "OFDM transmitters: Analog representation and DFT-based implementation," IEEE Trans. Signal Process., vol.51, no.9, pp.2450–2453, Sept. 2003.
[16] T. Cui and C. Tellambura, "Joint date detection and channel estimation for OFDM systems," IEEE Trans. Commun., vol.54, no.4, pp.670–679, April 2006.

Hikaru Morita received the B.Eng. and M.Eng. degrees in electrical and electronic engineering from Ibaraki University, Hitachi, Japan, in 2017 and 2019, respectively. In 2019, he joined Denso Wave Inc., Aichi, Japan. His research interests include signal processing for wireless communications.

Teruyuki Miyajima received the B.Eng., M.Eng., and Ph.D. degrees 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 are in signal processing for wireless communications. Dr. Miyajima is a member of IEEE.

Yoshiki Sugitani received the M.Eng., and Ph.D. degrees in electrical engineering from Osaka Prefecture University, Ibaraki, Japan, in 2013 and 2016, respectively. In 2016, he joined Ibaraki University, Hitachi, Japan, where he is currently an assistant professor in the Domain of Electrical and Electronic Systems Engineering. His current interests are in control for synchronization of complex networks. Dr. Sugitani is a member of IEICE.