Kalman Combining Based Iterative Detection and Decoding for MIMO Systems With Hybrid ARQ

Sangjoon Park, *Member, IEEE*

***Abstract*—This study proposes Kalman combining-based iterative detection and decoding (KC-IDD) schemes for multiple-input multiple-output (MIMO) systems with hybrid automatic repeat request (HARQ). The conventional Kalman filtering (KF) operation for Kalman combining performs the linear minimum mean-square error (LMMSE) detection with symbol-level combining (SLC), but it is unable to utilize *a priori* information of retransmitted symbols. Therefore, a new KF operation is derived, wherein an observation adjustment is employed to adjust the observation for a given state of the state-space model with the *a priori* information instead of directly utilizing it into the KF operation. Based on the modified KF operation, two KC-IDD schemes are developed: i) KC-IDD with the single-state observation adjustment, i.e., the observation adjustment for the current HARQ round, and ii) KC-IDD with multi-state observation adjustment (KC-IDD-MS), i.e., the observation adjustments throughout the HARQ rounds of a packet. Therefore, the proposed schemes can perform SLC-based LMMSE-IDD, where the complexity for a given number of turbo iterations is similar or smaller to that of the conventional LMMSE-IDD scheme with bit-level combining (BLC). Furthermore, the simulation results show that regardless of the retransmission strategy, the proposed schemes, especially the KC-IDD-MS scheme, outperformed the conventional LMMSE-IDD scheme in terms of error performance and decoding convergence speed for retransmissions.**

***Index Terms*—MIMO systems, hybrid ARQ, Kalman filtering, Kaman combining, iterative detection and decoding, LMMSE.**

I. Introduction

HYBRID automatic repeat request (HARQ) is the combined scheme of channel coding and retransmission to mitigate the link adaption errors in wireless communication systems [1], [2], [3]. The characteristics of the HARQ-employed systems vary based on the combining process of the retransmitted signals for detection. In particular, in multiple-input multiple-output (MIMO) systems with HARQ, the system performance, such as error performance and throughput, is considerably altered by the utilized combining and detection approach [4], [5]. Primarily, there are two combining and detection methods for MIMO systems with HARQ: bit-level combining (BLC) and symbol-level combining (SLC). In context, BLC can be applied in various system configurations with relatively low complexity, whereas SLC can improve the error performance and throughput. Therefore, several SLC schemes with an efficient reception procedure have been investigated for MIMO systems with HARQ [6], [7], [8], [9], [10], [11], [12].

In channel coded systems such as HARQ employed systems, the *a priori* information obtained from the channel decoding procedure can be used in the detection procedure for an improved estimation accuracy [13]. Upon iteratively conducting this process, the iterative detection and decoding (IDD) schemes can achieve a significantly better performance than the simple linear detection schemes such as linear minimum mean-square error (LMMSE) [14], [15], [16], [17]. Therefore, the receiver in HARQ-based MIMO systems can be designed while considering this IDD procedure with packet combining and detection method to achieve a suboptimal performance similar to that of the maximum-likelihood (ML) based combining and detection methods [7], [8] with less complexity.

The BLC can be considered as a simple extension of the conventional detection for MIMO non-ARQ systems using additional log-likelihood ratio (LLR) combining [4], [7]. Thus, the implementation of IDD to BLC becomes straightforward, e.g., performing the IDD procedure for the received signal in the current HARQ round and combining the output LLRs from the IDD detector with the LLRs computed at the previous transmission time slots. On the other hand, the direct implementation of IDD to SLC requires a significantly high complexity as the HARQ round of a packet increases. Further, the development of computationally efficient SLC-based IDD schemes is still limited owing to the complicated modeling for each system environment.

In [12], the Kalman combining scheme based on the Kalman filtering (KF) operation is developed for MIMO systems with HARQ. In this scheme, following the proper definition of the state-space model, the transition of a state vector can reflect the changes in a transmit signal vector by retransmission in HARQ-based MIMO systems, thereby enabling the Kalman combining scheme to perform the SLC-based LMMSE detection using KF, which requires a computational complexity similar to that of the BLC-based LMMSE detection method. However, in the state-space model of Kalman combining, it is unable to apply the *a priori* information of retransmitted symbols during the KF operation for detection, making it unsuitable for IDD procedures in HARQ-based MIMO systems.

Therefore, this paper proposes efficient SLC-based IDD schemes based on the Kalman combining scheme for MIMO systems with HARQ. A new KF operation is developed for the proposed IDD schemes, which enables the utilization of the *a priori* information for retransmitted symbols. Instead of

Manuscript received 31 March 2022; revised 26 July 2022; accepted 28 September 2022. Date of publication 3 October 2022; date of current version 13 February 2023. This work was supported by the Ministry of Science and ICT under Grant NRF-2022R1F1A1064163. The review of this article was coordinated by Prof. Yue Gao.

The author is with the Department of Electronic Engineering, Kyonggi University, Suwon 16227, Korea (e-mail: sj.park@kgu.ac.kr).

Digital Object Identifier 10.1109/TVT.2022.3211307

directly utilizing the *a priori* information at the KF operation, such as that in the conventional KF-based iterative detection approaches for frequency selective fading channels [18], [19], [20], the proposed KF operation adjusts the observations for the KF operation by using the *a priori* information. Thus, for both newly transmitted and retransmitted symbols, the proposed KF operation with the observation adjustment enables the utilization of the *a priori* information for iterative processing.

Based on the proposed KF operation, the basic Kalman combining-based IDD (KC-IDD) scheme performs the observation adjustment for the ongoing HARQ round (current state). Compared to the conventional LMMSE-IDD scheme with BLC, the basic KC-IDD scheme improves the error performance with similar or lower complexity. However, as the HARQ round of a packet increases, the ratio of the state with the observation adjustment over all the states decreases and the performance improvement can be limited. Thus, in addition to the basic KC-IDD scheme, the KC-IDD scheme with multi-state observation adjustment (denoted as the KC-IDD-MS scheme in the sequel) is developed. While the actual state transition occurs by the retransmission or transmission of a new packet, the proposed KC-IDD-MS scheme considers the virtual state transition for every new turbo iteration, which enables the observation adjustment throughout multiple states (HARQ rounds). Therefore, the KC-IDD-MS scheme can significantly improve the error performance compared to the LMMSE-IDD scheme with BLC, without the need for additional computational complexity from the basic KC-IDD scheme.

The remainder of this paper is organized as follows. Section II describes the MIMO system model with HARQ. Section III presents the KF operation for the conventional Kalman combining and the new KF operation with the observation adjustment. Section IV details the reception procedures of the proposed KC-IDD schemes. Section V presents the simulation results, and Section VI concludes the paper.

Throughout the paper, vectors and matrices are denoted by lowercase and uppercase boldface letters, respectively. Superscripts T, H, and -1 denote the transpose, conjugate-and-transpose, and inverse operators, respectively. $\mathbf{0}_{n\times k}$ and $\mathbf{1}_{n\times k}$ represent the $n\times k$ all-zero and all-one matrices, respectively, and $\mathbf{I}_n$ denotes the $n\times n$ identity matrix. $\mathbf{A}=\text{diag}(\mathbf{a})$ is a diagonal matrix with the main diagonal $\mathbf{a}$, and $\mathbb{E}[\cdot]$ denotes the mathematical expectation.

II. System Model

In this study, a spatially multiplexed MIMO system with N_i inputs and N_o outputs is considered, as illustrated in Fig. 1. The maximum number of HARQ rounds, i.e., the maximum number of transmissions of each packet, is set to R. For simple notations, we assume a single ARQ process (single packet transmission) within each transmission time slot. A data bit sequence for a packet is encoded at the transmitter, and a coded bit sequence is generated according to the employed retransmission strategy, e.g., Chase combining (CC) or incremental redundancy (IR), for the transmission at the $r(1\leq r\leq R)$th HARQ round. The coded bit sequence is modulated using a 2^Q-ary constellation $\mathcal{X}$, where the elements in $\mathcal{X}$ are assumed to satisfy $\sum_{x\in\mathcal{X}} x=0$ and $\sum_{x\in\mathcal{X}} |x|^2/2^Q=1$.

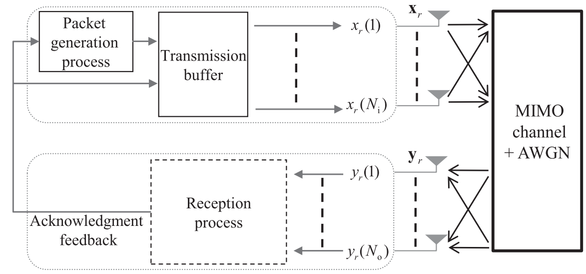


Fig. 1. System model block diagram.

Let J denote the number of transmit signal vectors for the transmission at each HARQ round. Then, $\mathbf{x}_{j,r}=[x_{j,r}(1),\ldots,x_{j,r}(N_\text{i})]^T$ denotes the $j(1\leq j\leq J)$th transmit signal vector for the rth HARQ round of a packet, where each $x_{j,r}(n)$ for $1\leq n\leq N_\text{i}$ is generated from Q coded bits. Without loss of generality, the index j is omitted throughout the remainder of this paper. Therefore, $\mathbf{x}_r=[x_r(1),\ldots,x_r(N_\text{i})]^T$ denotes the $N_\text{i}\times 1$ transmit signal vector for the rth HARQ round of a packet. Among $\mathbf{x}_r$, there are C symbols repeatedly transmitted throughout the HARQ rounds of a packet, i.e., $\{x_1(1),\ldots,x_1(C)\}=\{x_r(1),\ldots,x_r(C)\}$ for $1\leq r\leq R$. Further, the remaining last $U(=N_\text{i}-C)$ symbols are newly transmitted in every transmission time slot. Thus, the cases of $C=N_\text{i}(U=0)$ and $C<N_\text{i}(U>0)$ correspond to CC and IR retransmission strategies, respectively.

Based on above modeling, the received signal vector at the rth HARQ round can be expressed as

$$\mathbf{y}_r=\mathbf{H}_r\mathbf{x}_r+\mathbf{n}_r, \tag{1}$$

where $\mathbf{y}_r=[y_r(1),\ldots,y_r(N_\text{o})]^T$ is the $N_\text{o}\times 1$ receive signal vector at the rth HARQ round, $\mathbf{n}_r=[n_r(1),\ldots,n_r(N_\text{o})]^T$ is the $N_\text{o}\times 1$ zero-mean additive white complex Gaussian noise (AWGN) vector with the covariance matrix $\mathbb{E}[\mathbf{n}_r\mathbf{n}_r^H]=\sigma^2\mathbf{I}_{N_\text{o}}$, and $\mathbf{H}_r$ is the $N_\text{o}\times N_\text{i}$ MIMO channel matrix wherein $\mathbf{h}_{r,n}$ is the nth row of $\mathbf{H}_r$ for $1\leq n\leq N_\text{o}$. We assume that the full channel state information (CSI) at the receiver, i.e., all $\mathbf{H}_r$ and σ^2 are perfectly known to the receiver, and no CSI at the transmitter.

Following the completion of the iterative detection and decoding procedure with packet combining at the receiver, the transmission for the next $(r+1)$th HARQ round of the packet is performed during the upcoming transmission time slot if the decoded packet is detected in errors when $r<R$. If no errors are detected in the decoded packet or $r=R$, then the packet is terminated, and the initial transmission of a new packet is performed during the upcoming transmission time slot.

III. Kalman Filtering Operation With *A Priori* Information

A. State-Space Model and KF Procedure for Kalman Combining

First, we briefly review the state-space model and conventional Kalman combining procedure in [12] based on the system model described in Section II, which interprets a packet retransmission by the HARQ procedures as the state transition. The state-space model can be described using a state equation and an observation equation [21]. Let the transmit signal vector at the rth HARQ round, $\mathbf{x}_r$, be the rth state vector of size N_i for the state-space model. Then, the state equation is expressed as

$$\mathbf{x}_r = \mathbf{F}_r\mathbf{x}_{r-1} + \mathbf{w}_r, \tag{2}$$

where $\mathbf{F}_r$ is the $N_\mathrm{i} \times N_\mathrm{i}$ state transition matrix from $\mathbf{x}_{r-1}$ to $\mathbf{x}_r$ and $\mathbf{w}_r$ is the corresponding $N_\mathrm{i} \times 1$ process noise vector. Based on the definition of $\mathbf{x}_r$ in Section II, the state transition matrix $\mathbf{F}_r$ can be defined as

$$\mathbf{F}_r = \begin{cases} \mathrm{diag}([\mathbf{1}_{1\times\mathrm{C}}, \mathbf{0}_{1\times\mathrm{U}}]) & r > 1 \\ \mathbf{0}_{N_\mathrm{i}\times N_\mathrm{i}} & r = 1 \end{cases}, \tag{3}$$

which represents the state transition property of the system according to the retransmission strategy and r. When $r = 1$, $\mathbf{F}_r = \mathbf{0}_{N_\mathrm{i}\times N_\mathrm{i}}$ owing to the transmission of a new packet. Similarly, $\mathbf{w}_r$ can be defined as

$$\mathbf{w}_r = \begin{cases} [\mathbf{0}_{1\times C}, x_r(C+1), \ldots, x_r(N_\mathrm{i})]^T & r > 1 \\ \mathbf{x}_r & r = 1 \end{cases}, \tag{4}$$

which includes the symbols in $\mathbf{x}_r$ newly transmitted at the rth HARQ round of the packet.

Meanwhile, the observation equation for the given system model is equivalent to (1), i.e.,

$$\mathbf{y}_r = \mathbf{H}_r\mathbf{x}_r + \mathbf{n}_r, \tag{5}$$

where $\mathbf{y}_r$ and $\mathbf{n}_r$ become the rth observation vector and observation noise vector of size N_o for the state-space model, respectively.

Accordingly, the KF operation can be performed for Kalman combining [12] based on this state-space model. Let $\hat{\mathbf{x}}_{r|t}$ and $\mathbf{P}_{r|t} = \mathbb{E}[\{\mathbf{x}_r - \hat{\mathbf{x}}_{r|t}\}\{\mathbf{x}_r - \hat{\mathbf{x}}_{r|t}\}^H]$ denote the estimate of the state vector $\mathbf{x}_r$ given all the observations up to t and the corresponding error covariance matrix, respectively. Subsequently, $\hat{\mathbf{x}}_{r|r-1}$, the one-step predicted estimate of $\mathbf{x}_r$, and the corresponding error covariance matrix $\mathbf{P}_{r|r-1}$ are calculated from $\hat{\mathbf{x}}_{r-1|r-1}$ and $\mathbf{P}_{r-1|r-1}$ as [12], [21]

$$\hat{\mathbf{x}}_{r|r-1} = \mathbf{F}_r\hat{\mathbf{x}}_{r-1|r-1} + \overline{\mathbf{w}}_r \tag{6}$$

and

$$\mathbf{P}_{r|r-1} = \mathbf{F}_r\mathbf{P}_{r-1|r-1}\mathbf{F}_r^T + \mathbf{Z}_r. \tag{7}$$

As no *a priori* information is utilized in Kalman combining, $\overline{\mathbf{w}}_r = \mathbf{0}_{N_\mathrm{i}\times 1}$ and $\mathbf{Z}_r = \mathbf{I}_{N_\mathrm{i}} - \mathbf{F}_r$ regardless of r [12].

To avoid a matrix inversion operation after the prediction step in (6) and (7), the sequential correction steps can be performed using each element of the observation vector $\mathbf{y}_r$, that is, with $\hat{\mathbf{x}}_{r,0} = \hat{\mathbf{x}}_{r|r-1}$ and $\mathbf{P}_{r,0} = \mathbf{P}_{r|r-1}$, $\hat{\mathbf{x}}_{r,n}$ and $\mathbf{P}_{r,n}$ for $1 \le n \le N_\mathrm{o}$ can be obtained as

$$\hat{\mathbf{x}}_{r,n} = \hat{\mathbf{x}}_{r,n-1} + \mathbf{k}_{r,n}(y_r(n) - \mathbf{h}_{r,n}\hat{\mathbf{x}}_{r,n-1}) \tag{8}$$

and

$$\mathbf{P}_{r,n} = \mathbf{P}_{r,n-1} - \mathbf{k}_{r,n}\mathbf{v}_{r,n}^H, \tag{9}$$

where $\mathbf{h}_{r,n}$ is the nth row of $\mathbf{H}_r$, $\mathbf{v}_{r,n} = \mathbf{P}_{r,n-1}\mathbf{h}_{r,n}^H$, and the $N_\mathrm{i} \times 1$ Kalman gain vector $\mathbf{k}_{r,n}$ is given by

$$\mathbf{k}_{r,n} = \frac{\mathbf{v}_{r,n}}{\mathbf{h}_{r,n}\mathbf{v}_{r,n} + \sigma^2}. \tag{10}$$

The last estimate $\{\hat{\mathbf{x}}_{r,N_\mathrm{o}}, \mathbf{P}_{r,N_\mathrm{o}}\}$ becomes the value $\{\hat{\mathbf{x}}_{r|r}, \mathbf{P}_{r|r}\}$. Thus, the Kalman combining process at the rth HARQ round can obtain the LMMSE estimate of $\mathbf{x}_r$ given all $\{\mathbf{y}_1, \ldots, \mathbf{y}_r\}$ and $\{\mathbf{H}_1, \ldots, \mathbf{H}_r\}$ until the rth HARQ round, i.e., the SLC-based LMMSE estimate of $\mathbf{x}_r$. For example, in case of CC, $\hat{\mathbf{x}}_{r|r}$ can be rewritten as [12]

$$\hat{\mathbf{x}}_{r|r} = \left(\breve{\mathbf{H}}_r^H\breve{\mathbf{H}}_r + \sigma^2\mathbf{I}_{N_\mathrm{i}}\right)^{-1}\breve{\mathbf{H}}_r^H\breve{\mathbf{y}}_r, \tag{11}$$

where $\breve{\mathbf{y}}_r = [\mathbf{y}_1^T \cdots \mathbf{y}_r^T]^T$ is the $rN_\mathrm{o} \times 1$ aggregated received signal vector at the rth HARQ round and $\breve{\mathbf{H}}_r = [\mathbf{H}_1^T \cdots \mathbf{H}_r^T]^T$ is the $rN_\mathrm{o} \times N_\mathrm{i}$ aggregated channel matrix.

B. Proposed KF Procedure With Observation Adjustment

The KF operation presented in Section III-A is designed to utilize no *a priori* information by $\overline{\mathbf{w}}_r = \mathbf{0}_{N_\mathrm{i}\times 1}$ and $\mathbf{Z}_r = \mathbf{I}_{N_\mathrm{i}} - \mathbf{F}_r$ in (6) and (7), respectively. Considering the existence of available *a priori* information for the KF operation, e.g., $\overline{\mathbf{w}}_r \neq \mathbf{0}_{N_\mathrm{i}\times 1}$ in conjunction with the definition of the state noise vector $\mathbf{w}_r$ in (4), the KF operation to obtain $\hat{\mathbf{x}}_{r|r}$ from $\hat{\mathbf{x}}_{r-1|r-1}$ in (6)–(10) allows for the utilization of the *a priori* information of the following symbols: i) $\mathbf{x}_1$ (initially transmitted symbols) and ii) $\{x_r(C+1), \ldots, x_r(N_\mathrm{i})\}$ for $r > 1$ (new symbols transmitted for IR), which correspond to a non-zero state noise owing to a new symbol transmission. In other words, unlike the cases of newly transmitted symbols, the *a priori* information of the retransmitted symbols, i.e., every retransmitted symbols in CC and repeatedly transmitted symbols in IR, cannot be utilized. Thus, the KF operation in (6)–(10) for Kalman combining is unsuitable for the iterative detection procedure to fully utilize the *a priori* information of all the transmitted symbols, regardless of r and retransmission strategy.

An alternative approach for utilizing *a priori* information for retransmitted symbols as well as newly transmitted symbols is reiterating the Kalman combining procedure from the first to the rth state. That is, after the nonzero *a priori* information has been obtained, the receiver at the rth HARQ round can sequentially perform the KF operations from $\hat{\mathbf{x}}_{1|1}$ to $\hat{\mathbf{x}}_{r|r}$ with the given *a priori* information. Nevertheless, this approach requires r KF operations for one turbo iteration, which entails a high computational complexity as r increases.

To this end, we propose a modified KF operation based on observation adjustment, which can utilize the *a priori* information for both newly transmitted and retransmitted symbols by a single KF operation. First, we consider the relation $\mathbf{x}_r = \overline{\mathbf{x}}_r + \tilde{\mathbf{x}}_r$, where $\overline{\mathbf{x}}_r$ is the deterministic sequence provided by the *a priori* information of $\mathbf{x}_r$ and $\tilde{\mathbf{x}}_r$ is the residual part of $\mathbf{x}_r$ approximated as a zero-mean uncorrelated stochastic process [18], [19], [20]. $\overline{\mathbf{x}}_r$ can be expressed as

$$\overline{\mathbf{x}}_r = \mathbb{E}[\mathbf{x}_r|\mathbf{x}_r^\mathrm{P}] = [\overline{x}_r(1), \ldots, \overline{x}_r(N_\mathrm{i})]^T, \tag{12}$$

where $\mathbf{x}_r^\mathrm{P}$ is the *a priori* information of $\mathbf{x}_r$, e.g., the information obtained from the channel decoder in the IDD scheme. In addition, the real diagonal covariance matrix of $\tilde{\mathbf{x}}_r$, $\mathbf{V}_r$ can be expressed as

$$\mathbf{V}_r = \mathbb{E}[\tilde{\mathbf{x}}_r\tilde{\mathbf{x}}_r^H] = \mathrm{diag}(\mathbf{v}_r) = \mathrm{diag}([v_r(1), \ldots, v_r(N_\mathrm{i})]^T), \tag{13}$$

where the nth element of $\mathbf{v}_r$, $v_r(n) = \mathbb{E}[\tilde{x}_r(n)\tilde{x}_r^H(n)]$.

In accordance with the above-mentioned definitions, the proposed KF operation based on the observation adjustment can be explained as follows. First, $\mathbf{y}_r^*$, the new observation vector at the rth state for the proposed KF operation, is defined as

$$\mathbf{y}_r^* = \mathbf{y}_r - \mathbf{H}_r\overline{\mathbf{x}}_r = \mathbf{H}_r\tilde{\mathbf{x}}_r + \mathbf{n}_r. \tag{14}$$

Let $\tilde{\mathbf{x}}_r = \mathbf{\Lambda}_r\mathbf{x}_r$ with a real diagonal matrix $\mathbf{\Lambda}_r$. Owing to the relations $\mathbb{E}[\tilde{\mathbf{x}}_r\tilde{\mathbf{x}}_r^H] = \mathbf{V}_r$ and $\mathbb{E}[\mathbf{x}_r\mathbf{x}_r^H] = \mathbf{I}_{N_\text{i}}$, the following equation holds true:

$$\mathbf{V}_r = \mathbb{E}[\tilde{\mathbf{x}}_r\tilde{\mathbf{x}}_r^H] = \mathbf{\Lambda}_r\mathbb{E}[\mathbf{x}_r\mathbf{x}_r^H]\mathbf{\Lambda}_r = \mathbf{\Lambda}_r\mathbf{\Lambda}_r. \tag{15}$$

Consequently, $\mathbf{V}_r^{1/2} = \mathbf{\Lambda}_r = \text{diag}([v_r^{1/2}(1), \ldots, v_r^{1/2}(N_\text{i})]^T)$, and the new observation equation for the proposed KF operation can be derived as

$$\mathbf{y}_r^* = \mathbf{H}_r\mathbf{\Lambda}_r\mathbf{x} + \mathbf{n}_r = \mathbf{H}_r^*\mathbf{x}_r + \mathbf{n}_r, \tag{16}$$

where $\mathbf{H}_r^* = \mathbf{H}_r\mathbf{\Lambda}_r$. In this manner, the new observation equation (16) can be used instead of (5) to utilize the *a priori* information in the KF operation for MIMO systems with HARQ, regardless of r and retransmission strategy. (16) can also be interpreted as a reduced channel gain modeling of (5) by the soft interference cancellation (IC) on the receive signal vector.

As shown in (14)–(16), all the *a priori* information of $\mathbf{x}_r$ is utilized to obtain the new observation equation for the given r. Therefore, any *a priori* information of $\mathbf{x}_r$ should not be reutilized during the one-step prediction step. Consequently, $\hat{\mathbf{x}}_{r|r-1}^*$ and $\mathbf{P}_{r|r-1}^*$, the predicted estimates in the proposed KF operation, can be obtained as follows:

$$\hat{\mathbf{x}}_{r|r-1}^* = \mathbf{F}_r\hat{\mathbf{x}}_{r-1|r-1}^* \tag{17}$$

and

$$\mathbf{P}_{r|r-1}^* = \mathbf{F}_r\mathbf{P}_{r-1|r-1}^*\mathbf{F}_r^T + (\mathbf{I}_{N_\text{i}} - \mathbf{F}_r), \tag{18}$$

where $\hat{\mathbf{x}}_{r-1|r-1}^*$ and $\mathbf{P}_{r-1|r-1}^*$ are the outputs of the proposed KF operation at the previous $(r-1)$th state, and $(\mathbf{I}_{N_\text{i}} - \mathbf{F}_r)$ is used to initialize the error covariances of the newly transmitted symbols in $\mathbf{x}_r$ to 1, i.e., i) the symbols in $\mathbf{x}_1$ or ii) the last U symbols in $\mathbf{x}_r$ with $r > 1$ in case of IR. Thereafter, the correction step can be performed from $n = 1$ to $n = N_\text{o}$ by setting $\hat{\mathbf{x}}_{r,0}^* = \hat{\mathbf{x}}_{r|r-1}^*$ and $\mathbf{P}_{r,0}^* = \mathbf{P}_{r|r-1}^*$, which can be performed as

$$\hat{\mathbf{x}}_{r,n}^* = \hat{\mathbf{x}}_{r,n-1}^* + \mathbf{k}_{r,n}^*(y_r^*(n) - \mathbf{h}_{r,n}^*\hat{\mathbf{x}}_{r,n-1}^*) \tag{19}$$

and

$$\mathbf{P}_{r,n}^* = \mathbf{P}_{r,n-1}^* - \mathbf{k}_{r,n}^*(\mathbf{v}_{r,n}^*)^H. \tag{20}$$

In (19) and (20), $y_r^*(n)$ is the nth element of $\mathbf{y}_r^*$, $\mathbf{h}_{r,n}^*$ is the nth row of $\mathbf{H}_r^*$, $\mathbf{v}_{r,n}^* = \mathbf{P}_{r,n-1}^*(\mathbf{h}_{r,n}^*)^H$, and the $N_\text{i} \times 1$ Kalman gain vector $\mathbf{k}_{r,n}^*$ is expressed as

$$\mathbf{k}_{r,n}^* = \frac{\mathbf{v}_{r,n}^*}{\mathbf{h}_{r,n}^*\mathbf{v}_{r,n}^* + \sigma^2}. \tag{21}$$

Therefore, the final estimate $\{\hat{\mathbf{x}}_{r,N_\text{o}}^*, \mathbf{P}_{r,N_\text{o}}^*\}$ becomes the value $\{\hat{\mathbf{x}}_{r|r}^*, \mathbf{P}_{r|r}^*\}$.

As established in (12)–(21), the proposed KF operation utilizes the *a priori* information by adjusting the observation $\mathbf{y}_r$ and observation matrix $\mathbf{H}_r$ rather than directly utilizing the information in the prediction and correction steps. Thus, the proposed KF operation utilizes all the *a priori* information of the symbols regardless of r and retransmission strategy, i.e., the *a priori* information of both newly transmitted and retransmitted symbols. Meanwhile, the *a priori* information can be utilized in the conventional KF operation only for the newly transmitted symbols. Thus, the proposed KF operation can provide more accurate estimation results than the conventional KF operation for retransmitted symbols using the available *a priori* information, which leads to an overall error performance improvement. Additionally, considering the calculation of $\mathbf{y}_r^*$ and $\mathbf{H}_r^*$ results in slight computational complexity, the modified KF operation in (14)–(21) involves computational complexity that is approximately equal to that of the KF operation in (6)–(10) owing to the identical dimensions of the utilized matrices and vectors.

IV. Kalman Combining Based Iterative Detection and Decoding

The iterative reception procedures of the proposed KC-IDD schemes with the modified KF operation described in Section III-B are explained herein. For simple notations, the prediction step in (17) and (18) and the correction step in (19)–(21) including the observation adjustment in (14)–(16) for the modified KF operation are respectively expressed as

$$\{\hat{\mathbf{x}}_{r|r-1}, \mathbf{P}_{r|r-1}\} = \text{KF}^\text{P}\left(\{\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{P}_{r-1|r-1}\}\right) \tag{22}$$

and

$$\{\hat{\mathbf{x}}_{r|r}, \mathbf{P}_{r|r}\} = \text{KF}^\text{C}\left(\{\hat{\mathbf{x}}_{r|r-1}, \mathbf{P}_{r|r-1}\}, \{\overline{\mathbf{x}}_r, \mathbf{V}_r\}\right). \tag{23}$$

Further, K denotes the maximum number of the turbo iterations at the receiver.

A. Basic Reception Procedures

1) Initialization: The reception procedure of the proposed KC-IDD scheme is illustrated in Fig. 2. Prior to the first iteration at the rth HARQ round, the conditional symbol mean and variance for the first turbo iteration ($k = 1$), $\overline{\mathbf{x}}_r^k$ and $\mathbf{v}_r^k$ are initialized to $\mathbf{0}_{N_\text{i}\times 1}$ and $\mathbf{1}_{N_\text{i}\times 1}$, respectively.[1] Further, the proposed KC-IDD scheme conducts the prediction step as

$$\{\hat{\mathbf{x}}_{r|r-1}, \mathbf{P}_{r|r-1}\} = \text{KF}^\text{P}\left(\{\hat{\mathbf{x}}_{r-1|r-1}^\text{S}, \mathbf{P}_{r-1|r-1}^\text{S}\}\right), \tag{24}$$

where $\hat{\mathbf{x}}_{r-1|r-1}^\text{S}$ and $\mathbf{P}_{r-1|r-1}^\text{S}$ respectively represent the estimates of the state vector $\mathbf{x}_{r-1}$ and the corresponding error covariance matrix, which are stored in the receiver buffer after the reception procedure at the previous $(r-1)$th HARQ round. For $r = 1$, $\hat{\mathbf{x}}_{0|0}$ and $\mathbf{P}_{0|0}$ are $\mathbf{0}_{N_\text{i}\times 1}$ and $\mathbf{I}_{N_\text{i}}$, respectively.

2) KF and LLR Calculation Stage: After the initialization including the prediction step, the proposed KC-IDD scheme performs the remaining KF operation, i.e., the correction step. At the starting of the kth iteration, the correction step is initially performed with $\overline{\mathbf{x}}_r^k$ and $\mathbf{v}_r^k$, where $\overline{\mathbf{x}}_r^k$ and $\mathbf{v}_r^k$ include the conditional symbol mean and variance generated from the output LLRs of the decoder at the previous $(k-1)$th iteration,

[1]When $r > 1$, the soft information for retransmitted symbols calculated at previous transmission time slots can be reutilized. Such an application could be implemented in a straightforward manner and thereby it is omitted here.

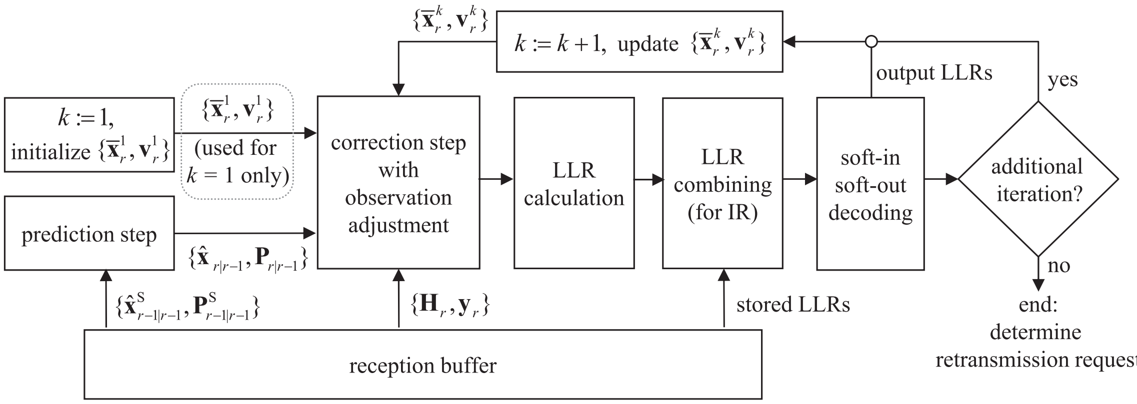


Fig. 2. The reception procedure of the basic KC-IDD scheme.

respectively. During the correction step, separated detection procedures are performed for each $x_r(n)$ for $1 \leq n \leq N_{\mathrm{i}}$, whereas the conditional symbol mean and variance of $x_r(n)$, $\overline{x}_r^k(n)$ and $v_r^k(n)$, should not be applied during the detection procedures for $x_r(n)$ according to the turbo principle [20]. Therefore, we define $\overline{\mathbf{x}}_{r,n}^k$ and $\mathbf{v}_{r,n}^k$, which are equivalent to $\overline{\mathbf{x}}_r^k$ and $\mathbf{v}_r^k$, except for the nth entries replaced by 0 and 1, respectively. Thereafter, the correction step for $x_r(n)$ at the kth turbo iteration can be performed as

$$\{\hat{\mathbf{x}}_{r|r,n}^k, \mathbf{P}_{r|r,n}^k\} = \mathrm{KF}^{\mathrm{C}}\left(\{\hat{\mathbf{x}}_{r|r-1}, \mathbf{P}_{r|r-1}\}, \{\overline{\mathbf{x}}_{r,n}^k, \mathbf{v}_{r,n}^k\}\right). \tag{25}$$

The output $\{\hat{\mathbf{x}}_{r|r,n}^k, \mathbf{P}_{r|r,n}^k\}$ obtained from (25) is used to compute the output LLRs of all the bits for $x_r(n)$. Let $\hat{x}_{r|r,n}^k(n)$ and $P_{r|r,n}^k(n)$ denote the nth element and diagonal element of $\hat{\mathbf{x}}_{r|r,n}^k$ and $\mathbf{P}_{r|r,n}^k$, respectively, where $P_{r|r,n}^k(n) = \mathbb{E}[\{x_r(n) - \hat{x}_{r|r,n}^k(n)\}\{x_r - \hat{x}_{r|r,n}^k(n)\}^H]$ denotes the variance of the residual interference. Subsequently, by a Gaussian approximation on the output $\hat{x}_{r|r,n}^k(n)$, the output LLR $l_{r,n}^{\mathrm{K},k}(q)$ of the $q(1 \leq q \leq Q)$th bit $b_{r,n}(q)$ for $x_r(n)$ after the KF operation at the kth turbo iteration can be obtained as [12]

$$l_{r,n}^{\mathrm{K},k}(q) = \log \frac{\sum\limits_{\forall x \in \mathcal{X}_q^1} \exp\left(\frac{-\left|\hat{x}_{r|r,n}^k(n) - x\right|^2}{P_{r|r,n}^k(n)}\right)}{\sum\limits_{\forall x \in \mathcal{X}_q^0} \exp\left(\frac{-\left|\hat{x}_{r|r,n}^k(n) - x\right|^2}{P_{r|r,n}^k(n)}\right)}, \tag{26}$$

where $\mathcal{X}_q^b$ with $b \in \{0,1\}$ includes the symbols in $\mathcal{X}$ with the qth bit $b_q = b$. (26) can be simplified using the max-log approximation as

$$l_{r,n}^{\mathrm{K},k}(q) \cong \max_{\forall x \in \mathcal{X}_q^0} \frac{\left|\hat{x}_{r|r,n}^k(n) - x\right|^2}{P_{r|r,n}^k(n)} - \max_{\forall x \in \mathcal{X}_q^1} \frac{\left|\hat{x}_{r|r,n}^k(n) - x\right|^2}{P_{r|r,n}^k(n)}. \tag{27}$$

(25) and (26) (or (25) and (27)) are repeated for $1 \leq n \leq N_{\mathrm{i}}$ to obtain all the LLRs of the bits for the symbols in $\mathbf{x}_r$.

3) LLR Combining and Decoding Stage: In case of CC, because $\mathbf{x}$ for a given packet is repeatedly transmitted throughout the HARQ rounds, no additional LLR combining procedure is required. Therefore, all $l_{r,n}^{\mathrm{K},k}(q)$ are utilized for the soft-in soft-output decoding operation, without any additional procedures. Meanwhile, in case of IR with $r > 1$, the LLRs calculated from the KF and LLR calculation stage need to be combined with the LLRs of the bits for $x_{r'}(n)$ when $r' < r$ and $C + 1 \leq n \leq N_{\mathrm{i}}$, i.e., non-repeatedly transmitted symbols, in a code-combining manner [1]. In this case, the combined LLRs are utilized for decoding. Thereafter, the output LLRs from the decoder are obtained after the decoding operation.

4) Iteration Stage: After the decoding stage, if an iteration stopping criterion, such as the maximum number of turbo iterations K, is not satisfied, then the output LLRs from the decoder are utilized to update $\overline{\mathbf{x}}_r$ and $\mathbf{v}_r$ for the KF stage at the subsequent $(k+1)$th turbo iteration. Let $\mathbf{l}_{r,n}^{\mathrm{D},k} = [l_{r,n}^{\mathrm{D},k}(1), \ldots, l_{r,n}^{\mathrm{D},k}(Q)]$ for $1 \leq n \leq N_{\mathrm{i}}$ denote the decoder output LLRs of the bits for $x_r(n)$, where $l_{r,n}^{\mathrm{D},k}(q)$ for $1 \leq q \leq Q$ represents the decoder output LLR of $b_{r,n}(q)$. Then, $\overline{x}_r^{k+1}(n)$ and $v_r^{k+1}(n)$ can be updated as [17], [19]

$$\overline{x}_r^{k+1}(n) = \sum_{x \in \mathcal{X}} x p(x_r(n) = x | \mathbf{l}_{r,n}^{\mathrm{D},k}) \tag{28}$$

and

$$v_r^{k+1}(n) = \sum_{x \in \mathcal{X}} |x - \overline{x}_r^{k+1}(n)|^2 p(x_r(n) = x | \mathbf{l}_{r,n}^{\mathrm{D},k}). \tag{29}$$

As expressed in (28) and (29), the conditional probability $p(x_r(n) = x | \mathbf{l}_{r,n}^{\mathrm{D},k})$ for all $x \in \mathcal{X}$ must be obtained to evaluate $\overline{x}_r^{k+1}(n)$ and $v_r^{k+1}(n)$. Based on the definition of LLR, this can be obtained as

$$p(x_r(n) = x | \mathbf{l}_{r,n}^{\mathrm{D},k}) = \prod_{q=1}^{Q} \frac{1}{1 + \exp\left((1 - 2b_q) l_{r,n}^{\mathrm{D},k}(q)\right)}, \tag{30}$$

where $b_q \in \{0,1\}$ denotes the qth bit of a given $x \in \mathcal{X}$.

On the contrary, if an iteration stopping criterion is satisfied after the current iteration, the receiver is required to determine whether the received packet is successfully decoded. If the decoded packet is determined to be successfully decoded or $r = R$ (maximum HARQ round of a packet), no information is required to be stored in the buffer as a new packet containing a new data bit sequence will be transmitted from the upcoming transmission time slot. In contrast, if the decoded packet is determined to be erroneous for $r < R$, the information utilized

TABLE I
SUMMARIZED IDD PROCEDURE OF KC-IDD

0) [Initialization] Set $k := 1$, and $\overline{\mathbf{x}}_r^1 := \mathbf{0}_{N_\mathrm{i}\times 1}$ and $\mathbf{v}_r^1 := \mathbf{1}_{N_\mathrm{i}\times 1}$. Also, do $\{\hat{\mathbf{x}}_{r\|r-1}, \mathbf{P}_{r\|r-1}\} = \mathrm{KF}^\mathrm{P}\left(\{\hat{\mathbf{x}}^\mathrm{S}_{r-1\|r-1}, \mathbf{P}^\mathrm{S}_{r-1\|r-1}\}\right)$.
1) [KF and LLR Calculation Stage] From $n = 1$ to $n = N_\mathrm{i}$, do
1-1) $\{\hat{\mathbf{x}}^k_{r\|r,n}, \mathbf{P}^k_{r\|r,n}\} = \mathrm{KF}^\mathrm{C}\left(\{\hat{\mathbf{x}}_{r\|r-1}, \mathbf{P}_{r\|r-1}\}, \{\overline{\mathbf{x}}^k_{r,n}, \mathbf{v}^k_{r,n}\}\right)$.
1-2) From $q = 1$ to $q = Q$, calculate $l^{\mathrm{K},k}_{r,n}(q)$ as in (26) or (27).
2) [LLR Combining and Decoding Stage] Perform LLR combining if necessary, and do decoding operations to generate the decoder output LLRs $l^{\mathrm{D},k}_{r,n}(q)$ for $1 \le q \le Q$ and $1 \le n \le N_\mathrm{i}$.
3) [Iteration Stage]
3-1) If an additional turbo iteration is required, compute $\overline{x}^{k+1}_r(n)$ and $v^{k+1}_r(n)$ using $l^{\mathrm{D},k}_{r,n}(q)$ as in (28)–(30) and go back to **1**). Otherwise, go to **3-2**).
3-2) If the decoding is failed when $r < R$, save $\{\hat{\mathbf{x}}^\mathrm{S}_{r\|r}, \mathbf{P}^\mathrm{S}_{r\|r}\}$ and LLRs (for IR) to the receiver buffer. Otherwise, clear the buffer.

during the reception procedures at the subsequent $(r+1)$th HARQ round of the packet needs to be stored.

In the proposed scheme, $\{\hat{\mathbf{x}}^\mathrm{S}_{r|r}, \mathbf{P}^\mathrm{S}_{r|r}\}$ should be stored for the prediction step in the future HARQ rounds. During the KF stage at the kth turbo iteration, the soft information $\{\overline{x}^k_r(n), v^k_r(n)\}$ is utilized to obtain the estimates $\{\hat{\mathbf{x}}^k_{r|r,n}, \mathbf{P}^k_{r|r,n}\}$. Because the modified KF operation of the proposed scheme can be interpreted as an LMMSE-based soft IC operation, a slight incorrectness in the $\overline{x}^k_r(n)$ and $v^k_r(n)$ can yield a significant degree of errors in the output estimates $\{\hat{\mathbf{x}}^k_{r|r,n}, \mathbf{P}^k_{r|r,n}\}$. Therefore, considering that a retransmission will be requested by the decoding errors during the IDD procedure at the current HARQ round, no soft information should be utilized for the KF estimate for the application in future HARQ rounds. Therefore, the error propagation to future HARQ rounds can be prevented by determining $\{\hat{\mathbf{x}}^\mathrm{S}_{r|r}, \mathbf{P}^\mathrm{S}_{r|r}\}$ as

$$\left\{\hat{\mathbf{x}}^\mathrm{S}_{r|r}, \mathbf{P}^\mathrm{S}_{r|r}\right\} = \left\{\hat{\mathbf{x}}^1_{r|r,n}, \mathbf{P}^1_{r|r,n}\right\} \text{ for any } 1 \le n \le N_\mathrm{i}. \quad (31)$$

Thus, any $\{\hat{\mathbf{x}}^1_{r|r,n}, \mathbf{P}^1_{r|r,n}\}$ with arbitrary n can be stored as $\{\hat{\mathbf{x}}^\mathrm{S}_{r|r}, \mathbf{P}^\mathrm{S}_{r|r}\}$ as $\overline{\mathbf{x}}^1_{r,n} = \mathbf{0}_{N_\mathrm{i}\times 1}$ and $\mathbf{v}^1_{r,n} = \mathbf{1}_{N_\mathrm{i}\times 1}$ for all n.

Meanwhile, when IR is employed and the current packet is not terminated, the LLRs of the bits for the symbols which will not be sent from the upcoming transmission time slot need to be stored for LLR combining in future HARQ rounds; in other cases, the LLR storing is not necessary.

5) Summary: The entire IDD procedure of the proposed KC-IDD scheme is summarized in Table I. As explained earlier, the KF operation in (25) can be interpreted as the LMMSE detection with soft IC operation using $\overline{\mathbf{x}}^k_r$ and $\mathbf{v}^k_r$. Meanwhile, $\{\hat{\mathbf{x}}^\mathrm{S}_{r-1|r-1}, \mathbf{P}^\mathrm{S}_{r-1|r-1}\}$ from the previous HARQ round experiences no IC operation, as every $\{\hat{\mathbf{x}}^\mathrm{S}_{r'|r'}, \mathbf{P}^\mathrm{S}_{r'|r'}\}$ for $1 \le r' \le r-1$ is obtained from the first turbo iteration at the r'th HARQ round with $\overline{\mathbf{x}}^1_{r,'n} = \mathbf{0}_{N_\mathrm{i}\times 1}$ and $\mathbf{v}^1_{r,'n} = \mathbf{1}_{N_\mathrm{i}\times 1}$. Thus, during the KF operations at the rth HARQ round, the soft IC operation is performed for only $\mathbf{y}_r$ and not for $\{\mathbf{y}_1, \ldots, \mathbf{y}_{r-1}\}$ via the single-state observation adjustment. Therefore, the proposed KC-IDD scheme in Table I can be interpreted as a KF-based LMMSE-IDD with partial soft IC operation for only the current HARQ round.

B. Extension for Multi-State Observation Adjustment

The basic KC-IDD scheme proposed in Section IV-A performs the observation adjustment for a single state corresponding to the current HARQ round. Therefore, the *a priori* information is utilized for the receiving signal of the current HARQ round only during the IDD procedure. Thus, the performance improvement of the proposed KC-IDD scheme for a large K can be restricted. Therefore, in this subsection, we develop the extended version of the basic KC-IDD scheme, KC-IDD-MS, which can perform the observation adjustment for multiple states corresponding to the previous and current HARQ rounds. The overall reception procedure of the proposed KC-IDD-MS scheme is illustrated in Fig. 3.

For the proposed KC-IDD-MS scheme, we define the virtual state transition concept, where the state (HARQ round) considered at the subsequent iteration can vary from the state (HARQ round) considered at the current iteration. That is, the state transition can be considered for the two consecutive iterations, thereby this virtual state transition varies from the actual state transition of the state-space model in Section III-A. Although the actual state transition is defined for the two consecutive HARQ rounds by additional transmission, the virtual state transition can be defined for any two HARQ rounds in the proposed KC-IDD-MS scheme.

Let o_k denote the virtual state (HARQ round) estimated at the kth iteration of the proposed KC-IDD-MS scheme. Further, let $\{\hat{\mathbf{x}}^k_{o_k,n}, \mathbf{P}^k_{o_k,n}\}$ denote the output of the KF operation for $x_{o_k}(n)$ at the kth iteration. Also, $\{\hat{\mathbf{x}}^0_{o_0,n}, \mathbf{P}^0_{o_0,n}\}$ for $k = 0$ is set to $\{\hat{\mathbf{x}}^\mathrm{S}_{r-1|r-1}, \mathbf{P}^\mathrm{S}_{r-1|r-1}\}$ regardless of n. Then, in the kth iteration, the prediction and correction steps are performed as

$$\left\{\hat{\mathbf{x}}^{k,(-)}_{o_k,n}, \mathbf{P}^{k,(-)}_{o_k,n}\right\} = \mathrm{KF}^\mathrm{P}\left(\{\hat{\mathbf{x}}^{k-1}_{o_{k-1},n}, \mathbf{P}^{k-1}_{o_{k-1},n}\}\right) \quad (32)$$

and

$$\{\hat{\mathbf{x}}^k_{o_k,n}, \mathbf{P}^k_{o_k,n}\} = \mathrm{KF}^\mathrm{C}\left(\{\hat{\mathbf{x}}^{k,(-)}_{o_k,n}, \mathbf{P}^{k,(-)}_{o_k,n}\}, \{\overline{\mathbf{x}}^k_{o_k,n}, \mathbf{v}^k_{o_k,n}\}\right), \quad (33)$$

where $\{\overline{\mathbf{x}}^k_{o_k,n}, \mathbf{v}^k_{o_k,n}\}$ is generated from the decoder output LLRs at the prior $(k-1)$th iteration. After the KF operation, the LLR calculation is performed as expressed in (26) or (27), and this is repeated for $1 \le n \le N_\mathrm{i}$ to obtain all the LLRs of the bits for the symbols in $\mathbf{x}_{o_k}$. After the LLR calculation, the KC-IDD and KC-IDD-MS schemes involve the identical detection procedures; therefore, the remainder of the detection procedure is herein omitted.

For the prediction step in (32), the state transition matrix $\mathbf{F}_{o_{k-1},o_k}$ is used instead of $\mathbf{F}_r$ in (24), where $\mathbf{F}_{o_{k-1},o_k}$ represents the state transition from $\mathbf{x}_{o_{k-1}}$ to $\mathbf{x}_{o_k}$. According to the relationship between $\mathbf{x}_{o_{k-1}}$ and $\mathbf{x}_{o_k}$ defined in Section II, $\mathbf{F}_{o_{k-1},o_k}$ can be defined as

$$\mathbf{F}_{o_{k-1},o_k} = \begin{cases} \mathbf{F}_r & k = 1 \\ \mathrm{diag}([\mathbf{1}_{1\times C}, \mathbf{0}_{1\times U}]) & o_{k-1} \ne o_k \\ \mathbf{I}_{N_\mathrm{i}} & o_{k-1} = o_k \end{cases}. \quad (34)$$

In case of CC with $U = 0$, because $\mathbf{F}_{o_{k-1},o_k} = \mathbf{I}_{N_\mathrm{i}}$ for $k > 1$, the prediction step (32) for the virtual state transition can be

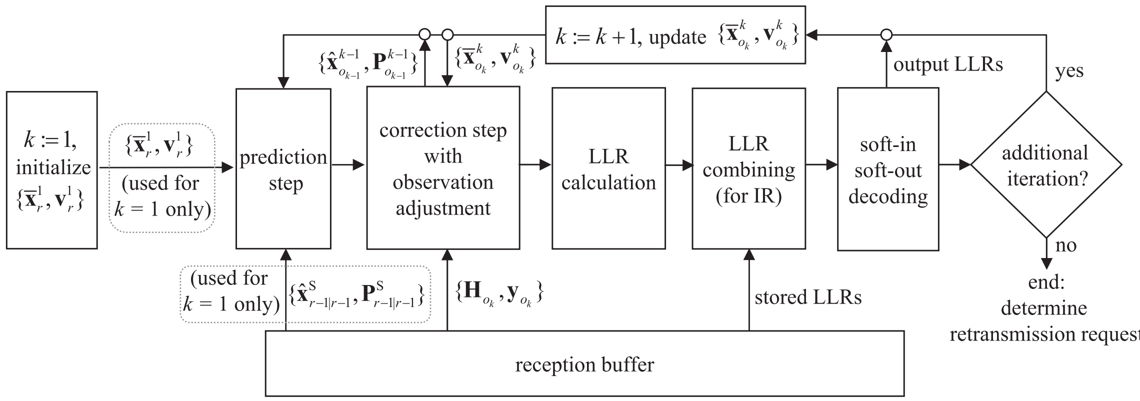


Fig. 3. The reception procedure of the KC-IDD-MS scheme. The index n indicating the transmit symbol is omitted for simplicity.

TABLE II
COMPUTATIONAL COMPLEXITY OF IDD SCHEMES AT THE rTH HARQ ROUND PER TURBO ITERATION

IDD Scheme	Detection (Complex Multiplications)	Detection (Complex Additions)	LLR Calculation
LMMSE-IDD w/ BLC	$N_\mathrm{i}(4N_\mathrm{o}^3/3+N_\mathrm{o}^2(N_\mathrm{i}+1)+N_\mathrm{o}(2N_\mathrm{i}+5/3)+1)$	$N_\mathrm{i}(4N_\mathrm{o}^3/3+N_\mathrm{o}^2(N_\mathrm{i}-3/2)+N_\mathrm{o}(N_\mathrm{i}+13/6)-1)$	$\mathcal{O}(N_\mathrm{i}Q2^Q)$
Proposed KC-IDD	$2N_\mathrm{o}N_\mathrm{i}^2(N_\mathrm{i}+3)+N_\mathrm{i}^2$	$N_\mathrm{o}N_\mathrm{i}^2(2N_\mathrm{i}+3)$	$\mathcal{O}(N_\mathrm{i}Q2^Q)$
Proposed KC-IDD-MS	$2N_\mathrm{o}N_\mathrm{i}^2(N_\mathrm{i}+3)+N_\mathrm{i}^2$	$N_\mathrm{o}N_\mathrm{i}^2(2N_\mathrm{i}+3)$	$\mathcal{O}(N_\mathrm{i}Q2^Q)$
Multiple LMMSE-IDD[a]	$r\cdot c_{\mathrm{CM,BLC}}(N_\mathrm{i},N_\mathrm{o})$	$r\cdot c_{\mathrm{CA,BLC}}(N_\mathrm{i},N_\mathrm{o})$	$\mathcal{O}(r\cdot N_\mathrm{i}Q2^Q)$
LMMSE-IDD w/ DSLC[a]	$c_{\mathrm{CM,BLC}}(C+rU,rN_\mathrm{o})$	$c_{\mathrm{CA,BLC}}(C+rU,rN_\mathrm{o})$	$\mathcal{O}((C+rU)Q2^Q)$

[a] $c_{\mathrm{CM,BLC}}(N_\mathrm{i},N_\mathrm{o})$ and $c_{\mathrm{CA,BLC}}(N_\mathrm{i},N_\mathrm{o})$ are the numbers of CMs and CAs for LMMSE-IDD with BLC in $N_\mathrm{o}\times N_\mathrm{i}$ MIMO systems, respectively.

omitted. Meanwhile, in case of IR with $U>0$, the prediction step should be performed for $k>1$.[2]

Although the updated *a priori* information is utilized for the current state in the basic KC-IDD scheme, the KC-IDD-MS scheme can utilize the updated *a priori* information for a distinct state at the following iteration. Therefore, if o_k for $1\leq k\leq K$ includes every possible state from the first to rth state, then the proposed KC-IDD-MS scheme can perform the observation adjustment for every available state for the current packet. Thus, to perform similar numbers of observation adjustments for utilizing the *a priori* information for each state, o_k is required to be set in a circular manner for $\{1,\ldots,r\}$, e.g., $r\to\cdots\to 1\to r\to\cdots$ to initially perform the KF operation for the rth state, which includes the new information on a packet from the channel. Notably, the KD-IDD-MS scheme does not require to be implemented for $r=1$ because the basic KC-IDD scheme can perform the observation adjustment for the current state, which is the only state available for the current packet when $r=1$.

C. Complexity and Memory Requirement

In this subsection, the computational complexity, space complexity, and memory requirement of the proposed schemes required in the detection stage with the *a priori* information (i.e., the KF and LLR calculation stage in the proposed scheme) are derived and compared with those of other IDD schemes. In addition to the proposed schemes and conventional LMMSE-IDD scheme with BLC, the following two IDD approaches are considered: i) the multiple LMMSE-IDD and combining scheme considering each HARQ round of a packet, e.g., performing an IDD procedure for all $\{\mathbf{x}_1,\ldots,\mathbf{x}_r\}$ separately and combining the results of r IDD procedures, and ii) the LMMSE-IDD scheme with direct SLC (DSLC), which performs the LMMSE-IDD procedure based on the aggregated system model including all the HARQ rounds of a packet, e.g., the aggregated channel matrix $[\mathbf{H}_1^T,\ldots,\mathbf{H}_r^T]^T$ for CC.

Table II shows the computational complexities of the IDD schemes at the rth HARQ round per turbo iteration. For the proposed schemes, the observation adjustment in (14)–(16) for each transmit symbol requires $(2N_\mathrm{o}N_\mathrm{i}+N_\mathrm{i})$ complex multiplications (CMs) and $N_\mathrm{o}N_\mathrm{i}$ complex additions (CAs), which should be performed N_i times. In addition, the correction step in (17)–(21) for each transmit symbol requires $N_\mathrm{o}(2N_\mathrm{i}^2+4N_\mathrm{i})$ CMs and $N_\mathrm{o}(2N_\mathrm{i}^2+N_\mathrm{i})$ CAs, which should also be performed N_i times. Moreover, considering the state transition matrix is a binary diagonal matrix, the prediction step does not require CMs and CAs for both schemes. Finally, the LLR calculation in (26) or (27) requires a computational complexity order of $\mathcal{O}(2^Q)$ for each bit, where the number of bits in a transmit signal vector is QN_i.

As shown in Table II, considering the highest-order terms, although the conventional LMMSE-IDD scheme with BLC requires the numbers of CMs and CAs similar to the proposed schemes for a small antenna configuration, it requires a larger

[2]By the diagonal structure of the state transition matrix with 0 and 1 only, the prediction steps in (24) and (32) become the selection of elements from the previous estimates. Thus, for both CC and IR, the prediction steps in (24) and (32) incur no complex multiplications or additions.

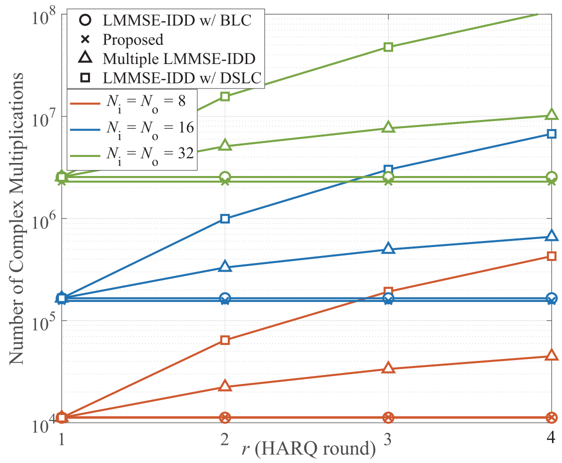


Fig. 4. The numbers of complex multiplications for IDD schemes according to r and antenna configuration. $N_\mathrm{i} = N_\mathrm{o}$ and CC ($C = N_\mathrm{i}$ and $U = 0$) is considered.

number of complex operations as N_i or N_o increases, especially when $N_\mathrm{o} >> N_\mathrm{i}$. Further, because r detection stages including the LLR calculation need to be simultaneously performed in each turbo iteration, the computational complexity of the multiple LMMSE-IDD scheme is r times higher than that of the conventional LMMSE-IDD scheme with BLC. In addition, because the size of the aggregated channel matrix for the IDD procedure is $rN_\mathrm{o} \times (C + rU)$, the LMMSE-IDD scheme with DSLC yields a significantly higher computational complexity than the other schemes, including the additional overhead for LLR calculation, as r increases.

Fig. 4 illustrates the number of CMs at the detection stage for the IDD schemes in Table II according to r and antenna configuration, where $N_\mathrm{i} = N_\mathrm{o}$ and CC ($C = N_\mathrm{i}$ and $U = 0$) is considered. The cases of CAs are omitted because the results for CAs are similar to those for CMs in Fig. 4. It is observed that the proposed schemes and conventional LMMSE-IDD scheme with BLC require similar numbers of CMs, fixed regardless of r, although the proposed schemes require a relatively smaller number of CMs for a larger antenna configuration. Meanwhile, unlike the proposed schemes and conventional LMMSE-IDD scheme with BLC, the number of CMs for the multiple LMMSE-IDD scheme increases with r, and the number of CMs for the LMMSE-IDD scheme with DSLC increases more rapidly with r. If IR is considered, the LMMSE-IDD scheme with DSLC requires a further larger number of CMs owing to the increased number of transmit antennas in the aggregated system model, while the numbers of CMs for the other schemes remain unchanged.

In Table III, the space complexity of the IDD schemes at the rth HARQ round per turbo iteration is derived, assuming parallel processing of the detection stage.[3] The proposed KC-IDD and KC-IDD-MS schemes exhibit identical space complexity order, although the KC-IDD-MS scheme can require a larger memory space than the KC-IDD scheme through multi-state observation adjustment. Meanwhile, considering $N_\mathrm{o} \geq N_\mathrm{i}$ for spatial

[3]The space complexity can be significantly altered according to the algorithm implementation, e.g., serial or parallel processing of the algorithm.

TABLE III
SPACE COMPLEXITY OF IDD SCHEMES AT THE rTH HARQ ROUND PER TURBO ITERATION

IDD Scheme	Space Complexity Order
LMMSE-IDD w/ BLC	$\mathcal{O}(N_\mathrm{i}(N_\mathrm{o}^2 + N_\mathrm{o}N_\mathrm{i} + 2^Q))$
Proposed KC-IDD	$\mathcal{O}(N_\mathrm{i}(N_\mathrm{i}^2 + N_\mathrm{o}N_\mathrm{i} + 2^Q))$
Proposed KC-IDD-MS	$\mathcal{O}(N_\mathrm{i}(N_\mathrm{i}^2 + N_\mathrm{o}N_\mathrm{i} + 2^Q))$
Multiple LMMSE-IDD	$\mathcal{O}(rN_\mathrm{i}(N_\mathrm{o}^2 + N_\mathrm{o}N_\mathrm{i} + 2^Q))$
LMMSE-IDD w/ DSLC	$\mathcal{O}((C+rU)(r^2N_\mathrm{o}^2+rN_\mathrm{o}(C+rU)+2^Q))$

TABLE IV
MEMORY REQUIREMENT OF IDD SCHEMES AFTER THE rTH HARQ ROUND

IDD Scheme	Matrix & Vector (complex units)	LLR (real units)
LMMSE-IDD w/ BLC	0	$(C + rU)Q$
Proposed KC-IDD	$N_\mathrm{i}(N_\mathrm{i} + 1)$	rUQ
Proposed KC-IDD-MS	$N_\mathrm{i}(N_\mathrm{i} + 1) + rN_\mathrm{o}(N_\mathrm{i} + 1)$	rUQ
Multiple LMMSE-IDD	$rN_\mathrm{o}(N_\mathrm{i} + 1)$	0
LMMSE-IDD w/ DSLC	$rN_\mathrm{o}(N_\mathrm{i} + 1)$	0

multiplexing, the conventional LMMSE-IDD scheme with BLC requires a larger space complexity compared to the proposed schemes. In addition, similar to the cases of the computational complexity in Table II, the space complexity order of the multiple LMMSE-IDD scheme is r times larger than that of the conventional LMMSE-IDD scheme with BLC. Furthermore, by using the aggregated system model, the LMMSE-IDD scheme with DSLC yields a significantly larger space complexity compared to the other schemes.

In Table IV, the memory requirement (receiver buffer size) of the IDD schemes after the reception procedure at the rth HARQ round is derived for when the next HARQ round of a packet is required (i.e., a decoding failure with $r < R$). In the proposed KC-IDD and KC-IDD-MS schemes, $N_\mathrm{i} \times 1$ vector ($\hat{\mathbf{x}}^\mathrm{S}_{r|r}$) and $N_\mathrm{i} \times N_\mathrm{i}$ Hermitian matrix ($\mathbf{P}^\mathrm{S}_{r|r}$) need to be stored. Further, in the proposed KC-IDD-MS scheme, multiple LMMSE-IDD scheme, and LMMSE-IDD scheme with DSLC, the $N_\mathrm{o} \times 1$ received signal vectors and $N_\mathrm{o} \times N_\mathrm{i}$ channel matrices from the first to rth HARQ rounds are required to be stored. Meanwhile, in the conventional LMMSE-IDD scheme with BLC, only the combined LLRs for the packet need to be stored, where the number of coded bits sent until the rth HARQ round is $(C + rU)Q$. In addition, the proposed schemes need to store the LLRs for rUQ bits, which will not be sent from the next transmission time slot, while the multiple LMMSE-IDD scheme and the LMMSE-IDD scheme with DSLC need no previous LLR values by the LLR recalculation for every transmitted bit. Consequently, in terms of memory requirement, the proposed KC-IDD-MS scheme requires a larger number of memory units than the other schemes.

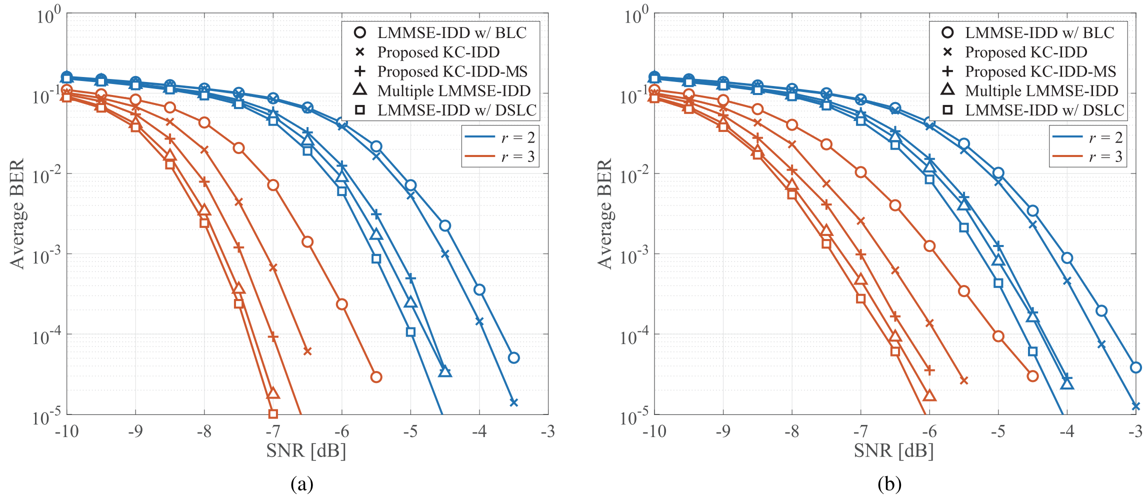


Fig. 5. Average BER performance with $K = 6$ for Chase combining: (a) quasi-static and (b) static channels.

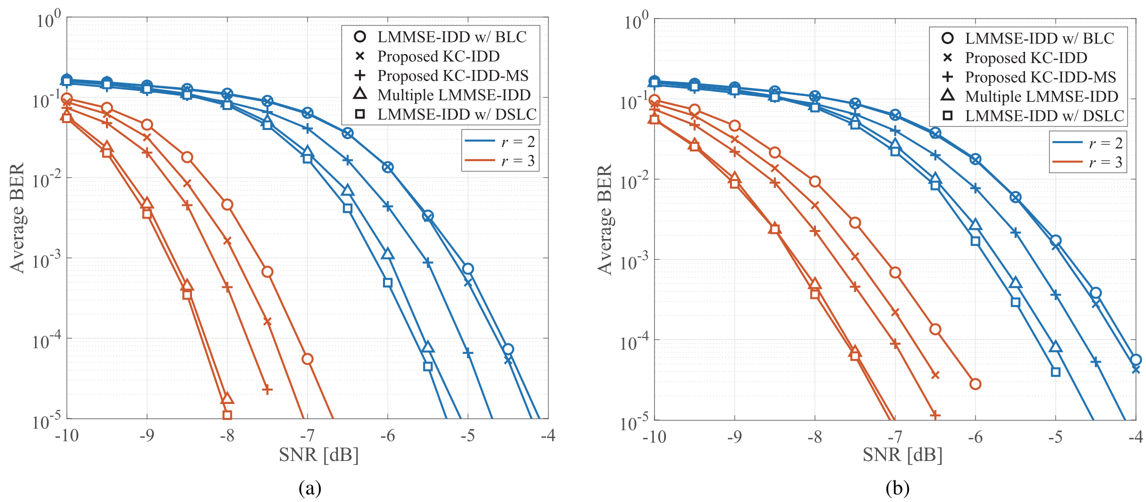


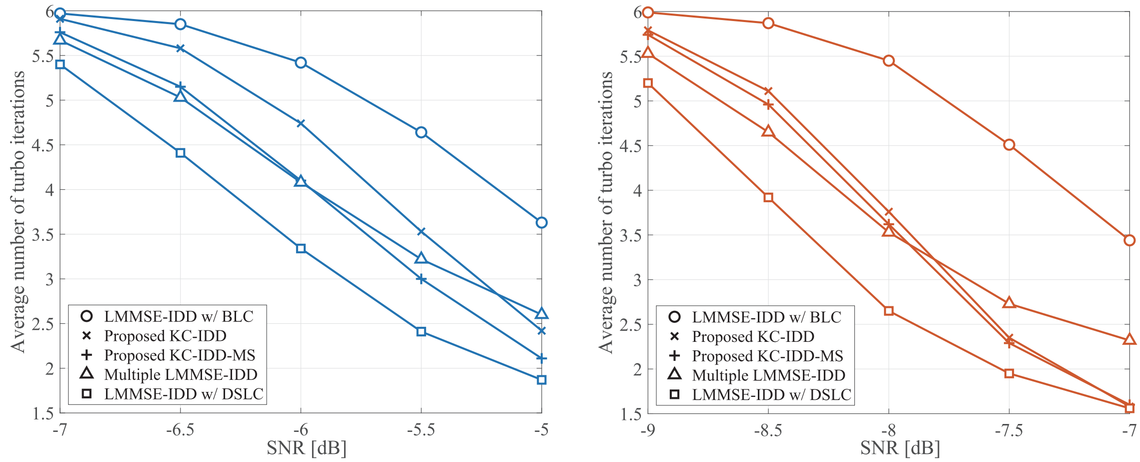
Fig. 6. Average BER performance with $K = 6$ for incremental redundancy: (a) quasi-static and (b) static channels.

V. Simulation Results

In this section, the error performance of IDD schemes is evaluated from numerical simulation results. For the proposed KC-IDD-MS scheme at the rth HARQ round, o_k is set to $\mathrm{mod}\ (r-k, r)+1$ to utilize the state (HARQ round) for each turbo iteration in a circular manner, as in Section IV-B. Because the error performance of the IDD schemes is identical for initial transmission (i.e., $r = 1$), the average bit-error rate (BER) results for retransmission (i.e., $r > 1$) are compared.

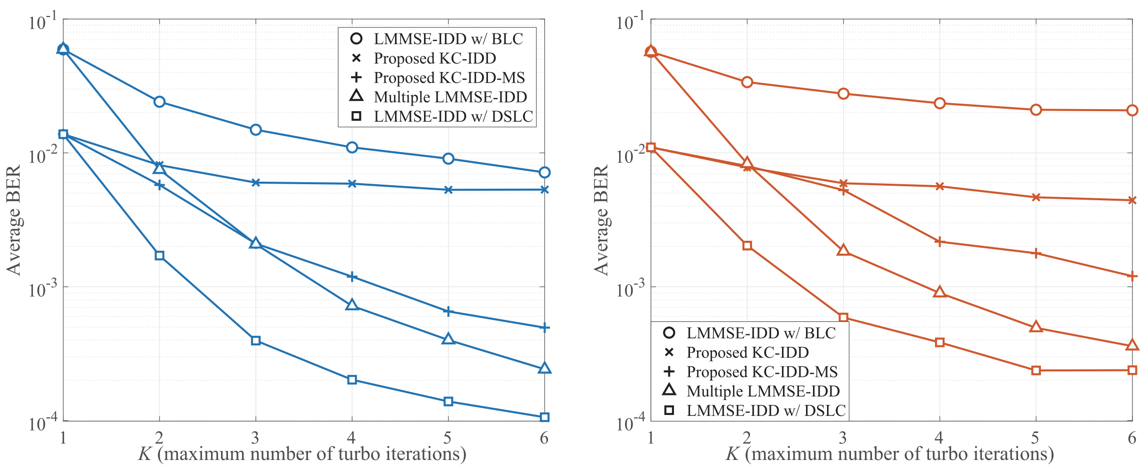
In the simulation environment, we considered $N_{\mathrm{i}} = N_{\mathrm{o}} = 8$, $R = 3$, and $U = N_{\mathrm{i}}/4$ for IR. A low-density parity-check (LDPC) code with a rate of 0.5 and a codeword length of 1152 was considered for the mother codeword, and the code rate of each packet was set to 3/4, i.e., 768 coded bits for each packet, by puncturing. A cyclic redundancy check (CRC) of length 32 was considered for error detection and retransmission request, and quadrature phase shift keying (QPSK) modulation (i.e., $Q = 2$) was employed. We considered i) the quasi-static Rayleigh fading channel, wherein the channel response is static in a given transmission time slot, and ii) the static Rayleigh fading channel, wherein the channel response is static until the termination of a given packet. A diagonal space interleaver was utilized for the symbol interleaving [14]. The min-sum decoding algorithm was applied at the receiver, and the number of decoding iterations for each turbo iteration was set to 10. Finally, the average signal-to-noise ratio (SNR) was set to $1/\sigma^2$.

Figs. 5 and 6 illustrate the average BER performance of IDD schemes with $K = 6$ for CC and IR, respectively. It is observed that the average BER characteristics of IDD schemes are similar for both quasi-static and static channels. Specifically, the proposed schemes achieve a better error performance than the conventional LMMSE-IDD scheme with BLC as r increases for both CC and IR. Because the performance improvement of Kalman combining over BLC is relatively limited for IR compared to the cases for CC [12], the performance improvement of the basic KC-IDD scheme over the conventional LMMSE-IDD scheme with BLC for IR also becomes limited compared to the cases for CC. On the other hand, by employing the multi-state observation adjustment, the proposed KC-IDD-MS scheme obtains a huge SNR gain from the conventional LMMSE-IDD scheme with BLC regardless of retransmission strategy. Meanwhile, the



(a) (b)

Fig. 7. Average number of turbo iterations until decoding convergence with $K = 6$ for Chase combining under quasi-static channel: (a) $r = 2$ and (b) $r = 3$.



(a) (b)

Fig. 8. Average BER performance according to K for Chase combining under quasi-static channel: (a) $r = 2$ with SNR $= 5$ dB and (b) $r = 3$ with SNR $=$ 7.5 dB.

multiple LMMSE-IDD scheme and LMMSE-IDD scheme with DSLC achieve a better average BER compared to the proposed schemes and conventional LMMSE-IDD scheme with BLC. However, they also require significantly larger computational and space complexities for the given K and r, as depicted in Section IV-C.

Fig. 7 compares the average number of turbo iterations of the IDD schemes until the decoding convergence when $K = 6$, where Chase combining and quasi-static channel are considered. The decoding convergence is defined by the satisfaction of all the parity check equations during the decoding operation. It is shown that the proposed schemes, especially, the KC-IDD-MS scheme, yield a faster convergence speed than the conventional LMMSE-IDD scheme with BLC for a given SNR. Considering the complexity per turbo iteration of the proposed schemes is similar or less than that of the conventional LMMSE-IDD scheme with BLC, the effective complexity of the proposed scheme can be significantly reduced by employing an iteration stopping criterion based on the decoding convergence. Meanwhile, the LMMSE-IDD scheme with DSLC requires the smallest number of turbo iterations for decoding convergence; however, its complexity per turbo iteration is significantly larger than that of the other schemes. Further, the multiple LMMSE-IDD scheme requires a larger number of turbo iterations than the proposed schemes in the high SNR region. This is because the first turbo iteration of the multiple LMMSE-IDD scheme has outputs identical to that of the conventional LMMSE-IDD scheme with BLC, which are also identical to that of the simple LMMSE detection with BLC having no iterative procedure.

Finally, Fig. 8 illustrates the average BER performance of the IDD schemes according to K for Chase combining under a quasi-static channel. When $K = 1$ (i.e., no iterative procedure), by fundamental limitation of the BLC compared to SLC in terms of error performance, the proposed schemes and LMMSE-IDD scheme with DSLC (which are identical to the LMMSE detection with SLC) outperform the conventional LMMSE-IDD scheme with BLC and multiple LMMSE-IDD scheme (which are identical to the LMMSE detection with BLC). Therefore, for a small K, the proposed schemes can achieve a fine error performance at a relatively low complexity overhead compared to the other IDD schemes. Meanwhile, as the number of maximum turbo iterations, K, increases, the performance improvement by the proposed schemes becomes smaller than those of the multiple LMMSE-IDD scheme and LMMSE-IDD scheme with DSLC. This is because they utilize the information from all the available HARQ rounds of a packet for each turbo iteration, while the proposed schemes utilize the information from a given HARQ round of the packet. Nevertheless, by employing the multi-state observation adjustment, the proposed KC-IDD-MS scheme can constantly improve the error performance as K increases, without having a high complexity for each turbo iteration as in the multiple LMMSE-IDD scheme and LMMSE-IDD scheme with DSLC.

VI. Conclusion

In this paper, we proposed and investigated the Kalman combining based IDD schemes for MIMO systems with HARQ. The proposed KC-IDD schemes based on the new KF procedure with observation adjustment outperform the conventional LMMSE-IDD scheme with BLC at a similar or lower complexity, regardless of the HARQ round and retransmission strategy. In particular, the proposed KC-IDD-MS scheme can achieve a significantly better error performance than the conventional LMMSE-IDD scheme with BLC owing to the observation adjustment performed for multiple states (HARQ rounds) without extra complexity, and the improvement was verified through numerical simulations. Therefore, the proposed schemes can be considered effective IDD schemes for MIMO systems with HARQ.

Throughout the paper, the HARQ model of the single ARQ process is considered for notational simplicity, which can be extended to multiple ARQ processes. In addition, the order of the states utilized for KC-IDD-MS in each turbo iteration for a given HARQ round can be additionally optimized to further improve the performance. Furthermore, it is considered that the KF operation in the proposed schemes is separately performed for each transmit symbol, which can be modified to simultaneously perform the KF operations of transmit symbols. In addition, when the nonlinear MIMO system model is considered for more practical applications, the proposed schemes can be extended using the extended Kalman filter to address such nonlinearities. These topics require further investigation in future works.

References

[1] Jung-Fu Cheng, "Coding performance of hybrid ARQ schemes," *IEEE Trans. Commun.*, vol. 54, no. 6, pp. 1017–1029, Jun. 2006.

[2] B. Makki and T. Eriksson, "On the performance of MIMO-ARQ systems with channel state information at the receiver," *IEEE Trans. Commun.*, vol. 62, no. 5, pp. 1588–1603, May 2014.

[3] A. Ahmed, A. Al-Dweik, Y. Iraqi, H. Mukhtar, M. Naeem, and E. Hossain, "Hybrid automatic repeat request (HARQ) in wireless communications systems and standards: A contemporary survey," *IEEE Commun. Surv. Tut.*, vol. 23, no. 4, pp. 2711–2752, Oct.–Dec. 2021.

[4] J. Lee, H. Lou, D. Toumpakaris, E. W. Jang, and J. M. Cioffi, "Transceiver design for MIMO wireless systems incorporating hybrid ARQ," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 32–40, Jan. 2009.

[5] S. Park and S. Choi, "Performance of symbol-level combining and bit-level combining in MIMO multiple ARQ systems," *IEEE Trans. Commun.*, vol. 64, no. 4, pp. 1517–1528, Apr. 2016.

[6] E. N. Onggosanusi, A. G. Dabak, Y. Hui, and G. Jeong, "Hybrid ARQ transmission and combining for MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, Anchorage, AK, USA, 2003, pp. 3205–3209.

[7] E. W. Jang, J. Lee, H. Lou, and J. M. Cioffi, "On the combining schemes for MIMO systems with hybrid ARQ," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 836–842, Feb. 2009.

[8] E. W. Jang, J. Lee, L. Song, and J. M. Cioffi, "An efficient symbol-level combining scheme for MIMO systems with hybrid ARQ," *IEEE Trans. Wireless Commun.*, vol. 8, no. 5, pp. 2443–2451, May 2009.

[9] C. Bai, W. A. Krzymien, and I. J. Fair, "Hybrid-ARQ for layered space time MIMO systems with channel state information only at the receiver," *IET Commun.*, vol. 4, no. 14, pp. 1765–1773, Sep. 2010.

[10] M. Zia and Z. Ding, "Bandwidth efficient variable rate HARQ under orthogonal space-time block codes," *IEEE Trans. Signal Process.*, vol. 62, no. 13, pp. 3360–3370, Jul. 2014.

[11] S. Park, "Aggregation-assisted combining for MIMO multiple ARQ systems," *IEEE Commun. Lett.*, vol. 22, no. 2, pp. 348–351, Feb. 2018.

[12] S. Park, "Kalman filtering based combining for MIMO systems with hybrid ARQ," *IEEE Trans. Signal Process.*, vol. 69, pp. 5250–5258, Sep. 2021.

[13] M. Tüchler, A. C. Singer, and R. Koetter, "Minimum mean squared error equalization using a priori information," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 673–683, Mar. 2002.

[14] H. Jafarkhani, *Space-Time Coding: Theory and Practice*, Cambridge, U.K.: Cambridge Univ. Press, 2005.

[15] J. Choi, "Iterative receivers with bit-level cancellation and detection for MIMO-BICM systems," *IEEE Trans. Signal Process.*, vol. 53, no. 12, pp. 4568–4577, Dec. 2005.

[16] C. Studer, S. Fateh, and D. Seethaler, "ASIC implementation of soft-input soft-output MIMO detection using MMSE parallel interference cancellation," *IEEE J. Solid-State Circuits*, vol. 46, no. 7, pp. 1754–1765, Jul. 2011.

[17] A. G. D. Uchoa, C. T. Healy, and R. C. de Lamare, "Iterative detection and decoding algorithms for MIMO systems in block-fading channels using LDPC codes," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2735–2741, Apr. 2016.

[18] S. Roy and T. M. Duman, "Soft input soft output Kalman equalizer for MIMO frequency selective fading channels," *IEEE Tran. Wireless Commun.*, vol. 6, no. 2, pp. 506–514, Feb. 2007.

[19] H. Kim and J. K. Tugnait, "Turbo equalization for doubly-selective fading channels using nonlinear Kalman filtering and basis expansion models," *IEEE Tran. Wireless Commun.*, vol. 9, no. 6, pp. 2076–2087, Jun. 2010.

[20] S. Park and S. Choi, "Iterative equalizer based on Kalman filtering and smoothing for MIMO-ISI channels," *IEEE Trans. Signal Process.*, vol. 63, no. 19, pp. 5111–5120, Oct. 2015.

[21] D. Simon, *Optimal State Estimation: Kalman, H_∞, and Nonlinear Approaches*, New York, USA: Wiley, 2006.

Sangjoon Park (Member, IEEE) was born in Seoul, Korea, in 1981. He received the B.S., M.S., and Ph.D. degrees in electrical and electronic engineering from Yonsei University, Seoul, Korea, in 2004, 2006, and 2011, respectively. From August 2011 to February 2014, he was with LG Electronics, Seoul, South Korea, as a Senior Research Engineer participating in research about telematics and vehicular communication systems. From April 2014 to August 2016, he was a Research Professor with the School of Electrical and Electronic Engineering, Yonsei University, Seoul, South Korea. From September 2016 to February 2019, he was an Assistant Professor with the Department of Information and Communication Engineering, Wonkwang University, Iksan, South Korea. From March 2019, he has been an Assistant Professor with the Department of Electronic Engineering, Kyonggi University, Suwon, Korea. His research interests mainly include communication signal processing techniques, such as detection and equalization algorithms, hybrid ARQ and MIMO applications for wireless communications, error correcting codes, algorithms for physical layer security, and other signal processing techniques.