

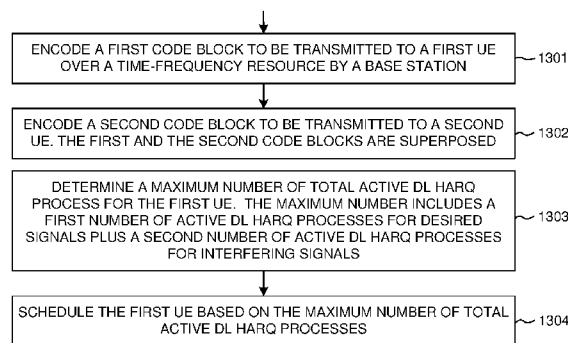


- (51) **International Patent Classification:**  
H04L 1/18 (2006.01) H04W 72/12 (2009.01)
- (21) **International Application Number:**  
PCT/CN2015/099664
- (22) **International Filing Date:**  
30 December 2015 (30.12.2015)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**  
62/097,813 30 December 2014 (30.12.2014) US  
14/980,867 28 December 2015 (28.12.2015) US
- (71) **Applicant:** **MEDIATEK INC.** [CN/CN]; No.1, Dusing Road 1st, Science-Based Industrial Park, Hsin-Chu, Taiwan 300 (CN).
- (72) **Inventors:** **HWANG, Chien-Hwa**; No.255, Guangming 9th Rd., Zhubei City, Hsinchu County, Taiwan 302 (CN). **LIAO, Yi-Ju**; 1F., No.40, Ln. 73, Wuling Rd., North Dist., Hsinchu City, Taiwan 300 (CN). **LOW, Tze-Ping**; 159 Bedford St, Lexington, MA Massachusetts 02420 (US).  **TSAI, Lung-Sheng**; No.176, Sec. 2, Zhongyi Rd., West Central Dist., Tainan City, Taiwan 700 (CN).
- (74) **Agent:** **BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.**; 16th Fl., Block A, Corporate Square, No.35 Jinrong Street, Beijing 100033 (CN).

- (81) **Designated States** (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) **Designated States** (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**  
— with international search report (Art. 21(3))

(54) **Title:** SOFT BUFFER PARTITION FOR SUPERPOSITION CODING



**FIG. 13**

(57) **Abstract:** When the codeword level interference cancellation (CW-IC) is used at the receiver in conjunction with the superposition coding scheme at the transmitter, it is helpful if the soft buffer at the receiver is reserved not only for the desired transport blocks (TBs) but also for the interfering TBs handled by the CW-IC. In so doing, the soft channel bits of interfering TBs at multiple (re) transmissions can be combined to enhance the success rate of data decoding. A soft buffer partition method for the soft channel bits of the desired and interfering TBs in the superposition coding scheme is proposed. The proposed method has a full flexibility in adjusting the soft buffer sizes for the desired and interfering TBs.



# **SOFT BUFFER PARTITION FOR SUPERPOSITION** **CODING**

## **CROSS REFERENCE TO RELATED APPLICATIONS**

5 This application claims priority under 35 U.S.C. §119 from U.S. Provisional Application Number 62/097,809, entitled “Rate Matching and Soft Channel Bits Storage for Superposition Coding,” filed on December 30, 2014; U.S. Provisional Application Number 62/097,813, entitled “Soft Buffer Partition for Superposition Coding,” filed on December 30, 2014, the subject matter of which is incorporated herein by reference.

## 10 **TECHNICAL FIELD**

The disclosed embodiments relate generally to mobile communication networks, and, more particularly, to rate matching, soft channel bits storage, and soft buffer partition for superposition coding.

## **BACKGROUND**

15 In a wireless cellular communications system, multiuser multiple-input multiple-output (MU-MIMO) is a promising technique to significantly increase the cell capacity. In MU-MIMO, the signals intended to different users are simultaneously transmitted with orthogonal (or quasi-orthogonal) precoders. On top of that, the concept of a joint optimization of multiuser operation from both transmitter and  
20 receiver’s perspective has the potential to further improve multiuser system capacity even if the transmission/precoding is non-orthogonal which could come from, for example but not limited to, the simultaneous transmission of a large number of non-orthogonal beams/layers with the possibility of more than one layer of data transmission in a beam. Such non-orthogonal transmission could allow multiple users  
25 to share the same resource elements without spatial separation, and allow improving the multiuser system capacity for networks with a small number of transmit antennas (i.e. 2 or 4, or even 1), where MU-MIMO based on spatial multiplexing is typically limited by wide beamwidth. An example of such joint Tx/Rx optimization associated with adaptive Tx power allocation and codeword level interference cancellation (CW-  
30 IC) receiver is recently a remarkable technical trend, including non-orthogonal multiple access (NOMA) and other schemes based on superposition coding.

In LTE, the rate-matching algorithm repeats or punctures the bits of a mother codeword to generate a requested number of bits according to the size of the time-frequency resource and a desired code rate that may be different from the mother code  
35 rate of the channel encoder. Besides, rate matching also needs to take into account the

soft buffer size of a code block at the receiver if soft packet combining is to be used to enhance the decoding performance. When superposition coding is used, a transport block intended to one user may also need to be decoded at another user's receiver. Nevertheless, according to the specifications of LTE, the soft buffer size per code block depends on the UE category. Therefore, the soft buffer sizes of the two receivers may not be the same. To enable the superposition coding in LTE, some issues related to rate matching need to be investigated. Specifically, two issues related to the soft buffer size setting at the transmitter and soft channel bits storage at the receiver in the superposition coding are discussed and resolved.

Furthermore, the soft buffers in the receiver are reserved for the desired signal as well as the interfering signals to be handled by CW-IC. As a result, the UE and the network should have the same understanding on the partition of the soft buffer for the two types of signals. Otherwise, the soft channel bits of some transport blocks (TBs) may not be accumulated, and soft packet combining of multiple (re)transmissions does not work efficiently. Up to LTE Rel-12, the soft buffer is reserved only for desired TBs, and the partition schemes for TBs is related to the number of DL HARQ processes. Thus, a method of soft buffer partition and restricting a total number of active DL HARQ processes for both desired and interfering TBs in superposition coding is proposed.

## SUMMARY

Methods for rate matching with soft buffer size setting at the transmitter and soft channel bits storage at the receiver for superposition coding are proposed. In the superposition coding scheme, a transport block intended to one UE needs to be decoded by another UE's receiver. However, the soft buffer sizes per code block of the two receivers may not be the same since the size depends on the UE category. The base station can signal the soft buffer size used at the transmitter for rate matching to the UEs for superposition decoding. A UE stores soft channel bits of encoded information bits associated with an interfering signal in its soft buffers in accordance with the soft buffer size used at the transmitter for rate matching. As a result, the UE can decode and subtract the interfering signal from the desired signal for superposition coding.

When the codeword level interference cancellation (CW-IC) is used at the receiver in conjunction with the superposition coding scheme at the transmitter, it is helpful if the soft buffer at the receiver is reserved not only for the desired transport blocks (TBs) but also for the interfering TBs handled by the CW-IC. In so doing, the soft channel bits of interfering TBs at multiple (re)transmissions can be combined to enhance the success rate of data decoding. A soft buffer partition method for the soft channel bits of the desired and interfering TBs in the superposition coding scheme is proposed. The proposed method has a full flexibility in adjusting the soft buffer sizes

for the desired and interfering TBs.

Other embodiments and advantages are described in the detailed description below. This summary does not purport to define the invention. The invention is defined by the claims.

5

## BRIEF DESCRIPTION OF DRAWINGS

Figure 1 illustrates a mobile communication network with rate matching and soft channel bits storage for superposition coding and interference cancellation in accordance with one novel aspect.

10 Figure 2 is a simplified block diagram of a base station and a user equipment that carry out certain embodiments of the present invention.

Figure 3 illustrates functional blocks in a communication system that maps information bits of a transport block to codewords and then maps to baseband signals for transmission.

15 Figure 4 illustrates an LTE rate matching procedure at the eNodeB and HARQ soft packet combining at the UE.

Figure 5 illustrates a first embodiment of LTE rate matching for superposition coding, where eNodeB determines soft buffer size for rate matching based on the code block soft buffer size reserved at each UE.

20 Figure 6 illustrates a second embodiment of LTE rate matching for superposition coding, where eNodeB determines soft buffer size for rate matching based on the minimum code block soft buffer size reserved at both UEs involved in superposition coding.

Figure 7 is a flow chart of a method of rate matching and soft channel bits storage from eNB perspective in accordance with one novel aspect.

25 Figure 8 is a flow chart of a method of rate matching and soft channel bits storage from UE perspective in accordance with one novel aspect.

Figure 9 illustrates the soft buffer partition based on the formula of  $N_{IR}$  in 3GPP LTE.

30 Figure 10 illustrates downlink HARQ processes associated with desired and interfering signals.

Figure 11 illustrates soft buffer partition based on a maximum total number of active DL HARQ processes. The maximum total number of active DL HARQ processes includes a first number of active DL HARQ processes for desired signals plus a second number of active DL HARQ processes for interfering signals.

35 Figure 12 illustrates a procedure of superposition coding with soft channel bits storage and soft buffer partition for both desired and interfering signals.

Figure 13 is a flow chart of a method of soft buffer partitioning from eNB perspective in accordance with one novel aspect.

Figure 14 is a flow chart of a method of soft buffer partitioning from UE

perspective in accordance with one novel aspect.

## DETAILED DESCRIPTION

Reference will now be made in detail to some embodiments of the invention, examples of which are illustrated in the accompanying drawings.

5 Figure 1 illustrates a mobile communication network 100 with superposition coding in accordance with one novel aspect. Mobile communication network 100 is an OFDM network comprising a plurality of user equipments UE 101, UE 102, and a serving base station eNB 104. In 3GPP LTE system based on OFDMA downlink, the radio resource is partitioned into subframes in time domain, each subframe is  
10 comprised of two slots and each slot has seven OFDMA symbols in the case of normal Cyclic Prefix (CP), or six OFDMA symbols in the case of extended CP. Each OFDMA symbol further consists of a number of OFDMA subcarriers in frequency domain depending on the system bandwidth. The basic unit of the resource grid is called Resource Element (RE), which spans an OFDMA subcarrier over one OFDMA  
15 symbol. Resource elements are grouped into resource blocks, where each resource block (RB) consists of 12 consecutive subcarriers in one slot.

Several physical downlink channels and reference signals are defined to use a set of resource elements carrying information originating from higher layers. For downlink channels, the Physical Downlink Shared Channel (PDSCH) is the main  
20 data-bearing downlink channel in LTE, while the Physical Downlink Control Channel (PDCCH) is used to carry downlink control information (DCI) in LTE. The control information may include scheduling decision, information related to reference signal information, rules forming the corresponding transport block (TB) to be carried by PDSCH, and power control command. For reference signals, Cell-specific reference  
25 signals (CRS) are utilized by UEs for the demodulation of control/data channels in non-precoded or codebook-based precoded transmission modes, radio link monitoring and measurements of channel state information (CSI) feedback. UE-specific reference signals (DM-RS) are utilized by UEs for the demodulation of control/data channels in non-codebook-based precoded transmission modes.

30 In the example of Figure 1, UE 101 (UE#1) is served by its serving base station eNB 104. UE#1 receives desired radio signal 111 transmitted from eNB 104. However, UE#1 also receives interfering radio signals. In one example, UE#1 receives interfering radio signal 112 transmitted from the same serving eNB 104 due to non-orthogonal multiple access (NOMA) operation intended for multiple UEs (e.g.,  
35 UE 102/UE#2) in the same serving cell. UE#1 may be equipped with an interference cancellation (IC) receiver that is capable of cancelling the contribution of the interfering signals from the desired signals.

Assume superposition coding scheme is used at the transmitter for the NOMA operation. Given the scenario of NOMA where UE#1 and UE#2 are scheduled in the

same time-frequency resource and the transport blocks (TBs) intended to UEs #1 and #2 are superposed and multi-casted to the two users with different transmission power levels ( $P_1 > P_2$ ). Suppose that UE#1 is closer to the base station (eNB 104) than UE#2, and the former and latter UEs are referred to as high- and low-geometry UEs, respectively. According to the rule of signal reception of NOMA, the receiver of UE#1 should perform codeword level interference cancellation (CW-IC) for the TB intended to UE#2. Specifically, UE#1 decodes the TB intended to UE#2, reconstructs the contribution of the UE#2's signal in the received signal, and then subtracts the reconstructed signal from the received signal to form a clean received signal. UE#1 can therefore decode its own signal via the clean received signal. When UE#1 has a decoding failure for UE#2's TB, it is helpful for the next decoding of the TB if UE#1 stores the soft channel bits of that TB. Therefore, when NOMA is used, it is beneficial for the high-geometry UE to be equipped with a soft buffer not only for its own signal but also for those interfering signals to be processed with the CW-IC. In so doing, multiple (re)transmissions of the interfering TB can be combined to enhance the success rate of decoding.

In LTE, the rate-matching algorithm repeats or punctures the bits of a mother codeword to generate a requested number of bits according to the size of the time-frequency resource and a desired code rate that may be different from the mother code rate of the channel encoder. Besides, rate matching also needs to take into account the soft buffer size of a code block at the receiver if soft packet combining is to be used to enhance the decoding performance. In the superposition coding scheme, the TB intended to UE#2 is to be decoded by both UE#1 and UE#2. However, the soft buffer sizes per code block of the two receivers may not be the same since the size depends on the UE category. In accordance with one novel aspect, two issues of rate matching, i.e., the soft buffer size setting at the transmitter (eNB 104) and soft channel bits storage at the receiver (UE#1), in the superposition coding are discussed and resolved.

Furthermore, the soft buffers in the receiver are reserved for the desired signal as well as the interfering signals to be handled by CW-IC. As a result, the UE and the network should have the same understanding on the partition of the soft buffer for the two types of signals. Otherwise, the soft channel bits of some TBs may not be accumulated, and soft packet combining of multiple (re)transmissions does not work efficiently. Up to LTE Rel-12, the soft buffer is reserved only for desired TBs, and the partition schemes for TBs are related to the number of DL HARQ processes. Thus, a method of soft buffer partition and restricting a total number of active DL HARQ processes for both desired and interfering TBs in superposition coding is proposed.

Figure 2 is a simplified block diagram of a base station 201 and a user equipment 211 that carry out certain embodiments of the present invention in a mobile communication network 200. For base station 201, antenna 221 transmits and

receives radio signals. RF transceiver module 208, coupled with the antenna, receives RF signals from the antenna, converts them to baseband signals and sends them to processor 203. RF transceiver 208 also converts received baseband signals from the processor, converts them to RF signals, and sends out to antenna 221. Processor 203 processes the received baseband signals and invokes different functional modules to perform features in base station 201. Memory 202 stores program instructions and data 209 to control the operations of the base station. Similar configuration exists in UE 211 where antenna 231 transmits and receives RF signals. RF transceiver module 218, coupled with the antenna, receives RF signals from the antenna, converts them to baseband signals and sends them to processor 213. The RF transceiver 218 also converts received baseband signals from the processor, converts them to RF signals, and sends out to antenna 231. Processor 213 processes the received baseband signals and invokes different functional modules to perform features in UE 211. Memory 212 stores program instructions and data 219 to control the operations of the UE. Memory 212 also contains a plurality of soft buffers 220 for storing soft channel bits of encoded code blocks.

Base station 201 and UE 211 also include several functional modules to carry out some embodiments of the present invention. The different functional modules are circuits that can be configured and implemented by software, firmware, hardware, or any combination thereof. The function modules, when executed by the processors 203 and 213 (e.g., via executing program codes 209 and 219), for example, allow base station 201 to schedule (via scheduler 204), encode (via encoder 205), mapping (via mapping circuit 206), and transmit control information and data (via control circuit 207) to UE 211, and allow UE 211 to receive, de-mapping (via de-mapper 216), and decode (via decoder 215) the control information and data (via control circuit 217) accordingly with interference cancellation capability. In one example, base station 201 provides code block soft buffer size of other UEs to UE 211. Under NOMA operation, upon receiving superposed code blocks intended for other UEs and corresponding soft buffer size, UE 211 is then able to perform Hybrid Automatic Repeat Request (HARQ) via HARQ handler 232, store soft channel bits into partitioned soft buffers for both desired and interfering TBs, and perform codeword level interference cancellation (CW-IC) via IC circuit 233 to decode the superposed code blocks and cancel the contribution of the interfering signals accordingly.

Figure 3 illustrates functional blocks of a transmitting device in a communication system that map information bits of a transport block (TB) to codewords and then map to baseband signals for transmission. In step 301, the information bits are arranged into transport blocks (TBs) and attached with CRC. In addition, the TBs are segmented into code blocks and attached with CRC. In step 302, channel coding (forward error correction such as Turbo coding) is performed with certain code rate. In step 303, rate matching is performed, which creates an output with a desired code

rate, and where the TBs are mapped into codewords. In step 304, the codewords are scrambled based on predefined scrambling rule (e.g., scramble with a corresponding Radio Network Temporary Identifier (RNTI) of the UE). In step 305, modulation mapping is performed, where the codewords are modulated based on various modulation orders (e.g., PSK, QAM) to create complex-valued modulation symbols. In step 306, layer mapping is performed, where the complex-valued symbols are mapped onto different MIMO layers depending on the number of transmit antenna used. In step 307, precoding is performed with certain precoding matrix index (PMI) for each antenna port. In step 308, the complex-valued symbols for each antenna are mapped onto corresponding resource elements (REs) of physical resource blocks (PRBs). Finally, in step 309, OFDM signals are generated for baseband signal transmission via antenna ports.

At the receiving side, a UE reserves a soft buffer to store soft channel bits for HARQ soft packet combining and decoding of the received information bits. The soft buffer size of a TB is denoted by  $N_{IR}$  bits (also referred as sub-buffer size), and the soft buffer size for a code block is denoted by  $N_{cb}$  bits (also referred as code block soft buffer size). According to 3GPP TS 36,212, the sizes  $N_{cb}$  can be determined as follows:

$$N_{cb} = \min\left(\left\lfloor \frac{N_{IR}}{C} \right\rfloor, K_w\right) \quad (1)$$

Where

- $N_{IR}$  is the soft buffer size for a TB;
- $C$  is the number of code blocks contained in a code word; and
- $K_w$  is the size of the output of the turbo encoder.

### **Rate Matching and Soft Channel Bits Storage**

Figure 4 illustrates an LTE rate matching procedure at the eNodeB and HARQ soft packet combining at the UE. At the eNodeB transmitter, the information bits are turbo encoded with the code rate of  $R=1/3$  to generate  $K_w$  coded bits. The number of transmitted coded bits is determined based on the size of the allocated time-frequency resource and the modulation coding scheme (MCS) assigned to the UE. Two-step rate matching is applied. The first step is applied only if  $N_{cb} < K_w$ . The purpose is to truncate the coded bits so that the truncated coded bits do not exceed the soft buffer size  $N_{cb}$ . In the second step,  $E$  consecutive coded bits are selected from the truncated coded bits (output of the first step), where  $E$  is the number of bits determined according to the size of allocated resource and the MCS level. The starting point of  $E$  coded bits is decided by the value of the redundancy version  $RV_i$ ,  $i = 0, 1, 2, 3$ , as shown in Figure 4. In the event of retransmission, a different  $RV_i$  is used to obtain a higher coding gain for an incremental redundancy soft packet-combining scheme.

At the UE receiver, the log likelihood ratio (LLRs),  $\{b_j(k); k=0,1,\dots,E-1\}$ , for

the  $j$ -th (re)-transmission, called soft channel bits, are computed. If the soft buffer for the code block is empty, the soft channel bits  $\{b_j(k)\}$  are stored in the  $N_{cb}$ -sized soft buffer; otherwise, the soft channel bits stored at the soft buffer are updated based on the newly computed  $\{b_j(k)\}$ . Finally, turbo decoding is performed to recover the information bits.

Figure 5 illustrates a first embodiment of LTE rate matching for superposition coding, where eNodeB determines soft buffer size for rate matching based on the code block soft buffer size  $N_{cb}$  reserved at each UE. In the example of Figure 5, UE#1 and UE#2 are two UEs scheduled in the same time-frequency resource, and the TBs intended to UE#1 and UE#2 are superposed. According to the rule of signal reception of NOMA, if UE#1 is the high-geometry UE, then the TB intended to UE#2 is to be decoded at both receivers of UE#1 and UE#2. Denote the soft buffer sizes per code block of UE#1 and UE#2 by  $N_{cb,1}$  and  $N_{cb,2}$ , respectively. It is assumed that  $N_{cb,1} < N_{cb,2}$ . In this embodiment, the first step of rate matching is performed at the eNodeB based on the soft buffer size  $N_{cb,2}$  of UE#2 (the target UE of the TB). That is, the first  $N_{cb,2}$  bits are chosen from the length- $K_w$  turbo encoder output. In the second step of the rate matching,  $E$  consecutive bits are selected from the output bits sequence of the first step.

In accordance with one novel aspect, the soft buffer size  $N_{cb,2}$  used at the eNodeB transmitter for rate matching should be signaled to UE#1 for the purpose of code block decoding. The redundancy version of the code block should also be signaled. As depicted by Figure 5, at the UE#1's receiver, all the soft channel bits  $\{b_j(k); k = 0, 1, \dots, E-1\}$  are input into the turbo decoder. If the decoding fails and the soft channel bits are to be stored at the soft buffer, in the case of  $N_{cb,1} < N_{cb,2}$ , the soft buffer size reserved at UE#1's receiver is insufficient to store all  $N_{cb,2}$  soft channel bits. In this case, only those soft channel bits corresponding to the first  $N_{cb,1}$  bits of the length- $K_w$  turbo encoder output are stored; those corresponding to the remaining  $N_{cb,2} - N_{cb,1}$  coded bits are discarded. At the UE#2's receiver, since the receiver's soft buffer size per code block  $N_{cb,2}$  is the same as the soft buffer size used at the first step of rate matching, the behavior of the soft channel bits storage is the same as in the conventional LTE systems.

Figure 6 illustrates a second embodiment of LTE rate matching for superposition coding, where eNodeB determines soft buffer size for rate matching based on the minimum code block soft buffer size  $N_{cb}$  reserved at both UEs involved in superposition coding. In the example of Figure 6, UE#1 and UE#2 are two UEs scheduled in the same time-frequency resource, and the TBs intended to UE#1 and UE#2 are superposed. According to the rule of signal reception of NOMA, if UE#1 is the high-geometry UE, then the TB intended to UE#2 is to be decoded at both receivers of UE#1 and UE#2. Denote the soft buffer sizes per code block of UE#1 and UE#2 by  $N_{cb,1}$  and  $N_{cb,2}$ , respectively. It is assumed that  $N_{cb,1} < N_{cb,2}$ . In this

embodiment, the first step of the rate matching is based on the code block soft buffer size  $N_{cb} = \min(N_{cb,1}, N_{cb,2})$ . The soft buffer size  $N_{cb}$  used at eNodeB rate matching should be signaled to UE#1 and UE#2 for code block decoding. The redundancy version of the code block should also be signaled. Since  $N_{cb} \leq N_{cb,1}$  and  $N_{cb} \leq N_{cb,2}$ , the behavior of the soft channel bits storage is the same as in the conventional LTE systems.

Figure 7 is a flow chart of a method of rate matching and soft channel bits storage from eNB perspective in accordance with one novel aspect. In step 701, a base station encodes a first code block to be transmitted to a first UE over a time-frequency resource. In step 702, the base station encodes a second code block to be transmitted to a second UE. The first code block and the second code block are superposed. In step 703, the base station performs rate matching for the first UE based on a first soft buffer size of the first code block. In step 704, the base station performs rate matching for the second UE based on a second soft buffer size of the second code block. In step 705, the base station transmits encoded information blocks of the code blocks. The base station also transmits information of the second soft buffer size to the first UE for interference cancellation.

Figure 8 is a flow chart of a method of rate matching and soft channel bits storage from UE perspective in accordance with one novel aspect. In step 801, a first UE receives a first plurality of encoded information bits associated with a first code block from a base station intended to the first UE. In step 802, the first UE receives a second plurality of encoded information bits associated with a second code block from the base station intended to a second UE. The first code block and the second code block are superposed. In step 803, the first UE decodes the first code block and stores soft channel bits of the first plurality of information bits in accordance with a first code block soft buffer size if the decoding of the first code block fails. In step 804, the first UE decodes the second code block and stores soft channel bits of the second plurality of information bits in accordance with a second code block soft buffer size if the decoding of the second code block fails.

### 30 **Soft Buffer Partition**

Up to Rel-12 LTE, the soft buffer is reserved only for the desired TBs. The partition scheme for TBs carried in different component carriers, HARQ processes, and spatial layers is briefly described below. The soft buffer size for a TB is denoted by  $N_{IR}$  bits (also referred as sub-buffer size). According to the latest version of 3GPP TS 36.212, the size  $N_{IR}$  can be determined as:

$$N_{IR} = \left\lfloor \frac{N_{soft}}{K_C \cdot K_{MIMO} \cdot \min(M_{DL\_HARQ}, M_{limit})} \right\rfloor \quad (2)$$

Where

- $N_{soft}$ : total number of soft channel bits;

- $K_C \in \{1,2,5\}$ : a number depending on  $N_{soft}$  and UE capability in the number of supportable spatial layers;
- $K_{MIMO}$ : equal to 2 if the UE is configured to receive the physical downlink shared channel (PDSCH) transmissions based on transmission modes 3, 4, 8, 9 or 10, and is equal to 1 otherwise;
- $M_{DL\_HARQ}$ : the maximum number of downlink (DL) HARQ processes; and
- $M_{limit}$ : a constant equal to 8.

For a frequency division duplex (FDD) LTE system, the maximum number of downlink (DL) HARQ processes  $M_{DL\_HARQ}$  is equal to 8. For a time division duplex (TDD) LTE system,  $M_{DL\_HARQ}$  varies from 4 to 15 depending on the TDD uplink-downlink configuration. The implication of  $\min(M_{DL\_HARQ}, M_{limit})$  is that a BS would not schedule more than  $\min(M_{DL\_HARQ}, M_{limit})$  “active” DL HARQ processes per component carrier for a UE. Therefore, there is no need for the UE to reserve soft buffers for more than  $K_C \times K_{MIMO} \times \min(M_{DL\_HARQ}, M_{limit})$  TBs. An active DL HARQ process means one that the TB associated with the DL HARQ process had been already received but not successfully decoded by the UE, and the soft channel bits are stored at the soft buffer.

In a TDD system with the maximum number of DL HARQ processes less than  $M_{limit}$  or an FDD system, the equality  $\min(M_{DL\_HARQ}, M_{limit})=M_{DL\_HARQ}$  holds. No issue regarding insufficiency of soft buffer for TBs would occur as long as no more than  $M_{DL\_HARQ}$  DL HARQ processes are used. However, for TDD system with the maximum number of DL HARQ processes greater than  $M_{limit}$  (i.e., TDD uplink-downlink configurations 2, 3, 4, and 5), the BS would not schedule a user with more than  $M_{limit}$  “active” DL HARQ processes concurrently per component carrier to avoid insufficient space of the soft buffer for TBs.

Figure 9 illustrates the soft buffer partition based on the formula of  $N_{IR}$  given in equation (2). The soft buffer can accommodate  $N_{soft}$  soft channel bits. The soft buffer is divided into  $K_C \times K_{MIMO} \times \min(M_{DL\_HARQ}, M_{limit})$  equal-sized sub-buffers. A shaded sub-buffer 901 is depicted at the left-bottom corner of Figure 9. The soft channel bits of one TB, with a length of  $N_{IR}$ , are stored in the shaded sub-buffer 901. In one novel aspect, the idea of restricting the number of active DL HARQ processes to save the soft buffer size for TDD uplink-downlink configurations 2, 3, 4, and 5 can be used for the soft buffer partition in superposition coding.

Figure 10 illustrates downlink HARQ processes associated with desired and interfering signals. Figure 10 depicts the meaning of “DL HARQ process of an interfering signal”. In the example of Figure 10, TB1 and TB2 are superposed in the same time-frequency resource, and TB1 and TB2 are desired TBs (intended for a high-geometry UE#1) and interference TBs (intended for other UEs), respectively, from UE#1’s perspective. The transmission of TB1 and TB2 both take place at subframe 3. Assume both TBs suffer from decoding failure at subframe 3. As a

result, the number of active DL HARQ processes for the desired signal is one from subframes 3 to 10, and the number of active DL HARQ processes for the interfering signal is one during the same interval. Retransmissions of TB1 and TB2 occur at subframe 11. Assume the decoding of TB1 fails, and the decoding for TB2 succeeds.

5 As a result, the soft channel bits for TB2 can be cleared, and the number of active DL HARQ processes for the interfering signal becomes zero. The number of active DL HARQ processes for the desired signal remains to be one.

The proposed method for the soft buffer partition in the superposition coding is described below. The formula of  $N_{IR}$  is kept unchanged except that the term  
10  $\min(M_{DL\_HARQ}, M_{limit})$  is replaced with a value  $P_{MAX}$ . That is,

$$N_{IR} = \left\lfloor \frac{N_{soft}}{K_C \cdot K_{MIMO} \cdot P_{MAX}} \right\rfloor \quad (3)$$

Where

- $N_{soft}$ : total number of soft channel bits;
- $K_C \in \{1,2,5\}$ : a number depending on  $N_{soft}$  and UE capability in the number  
15 of supportable spatial layers;
- $K_{MIMO}$ : equal to 2 if the UE is configured to receive the physical downlink shared channel (PDSCH) transmissions based on transmission modes 3, 4, 8, 9 or 10, and is equal to 1 otherwise;
- $P_{MAX}$ : a maximum total number of active DL HARQ processes.

20 Figure 11 illustrates the proposed soft buffer partition method based on a maximum total number of active DL HARQ processes  $P_{MAX}$ . The maximum total number of active DL HARQ processes  $P_{MAX}$  includes a first number of active DL HARQ processes for desired signals plus a second number of active DL HARQ processes for interfering signals. Given a UE for which codeword level interference  
25 cancellation in the superposition coding is needed, the BS scheduler in general would not schedule more than  $P_{MAX}$  active DL HARQ processes per component carrier for the UE, including the HARQ processes associated with desired TBs and the HARQ processes associated with interfering TBs. For example, if a high-geometry UE has  $T$  active DL HARQ processes in a component carrier for desired TBs, then the number  
30 of active DL HARQ processes in the component carrier for interfering TBs is no more than  $(P_{MAX} - T)$ .

The value of  $P_{MAX}$  can be determined by a predefined function of any of the parameters such as the maximum number of DL HARQ processes, and the UE and network can compute  $P_{MAX}$  on its own. The value of  $P_{MAX}$  can also be configurable  
35 by the BS and sent to the UE. The sizes of the soft buffer for desired and interfering TBs can be dynamically adjusted by the BS scheduler in each component carrier and at any time instant. As long as the total number of active DL HARQ processes does not exceed  $P_{MAX}$ , the soft buffer size is enough for active TBs. An extreme case is

that the soft buffer size for interfering TBs is equal to zero; in this case, no soft packet combining is used for interfering TBs.

Using the proposed method, the soft buffer size is the same for a DL HARQ process corresponding to a desired TB and to an interfering TB. A slight adjustment on the soft buffer sizes can be made for the two types of DL HARQ processes if it is desirable to have different soft buffer sizes for them. In one example, the ratio between the size of a sub-buffer corresponding to a desired signal and the size of a sub-buffer corresponding to an interfering signal is a predefined value. In another example, the ratio between the size of a sub-buffer corresponding to a desired signal and the size of a sub-buffer corresponding to an interfering signal is configurable.

Figure 12 illustrates a procedure of superposition coding with soft channel bits storage and soft buffer partition for both desired and interfering signals. In step 1211, a serving base station BS 1201 schedules a first UE#1 and a second UE#2 for NOMA operation. UE#1 is a high-geometry UE and UE#2 is a low-geometry UE. In step 1212, the BS sends control information to UE#1 and UE#2 via PDCCH. The control information may include information of the code block soft buffer size for UE#2, and may include the maximum number of total active DL HARQ processes for UE#1. In step 1213, the BS transmits superposed transport blocks TB1 and TB2 to UE#1 and UE#2 under NOMA. The encoding and rate matching of TB1 is based on a first soft buffer size for UE#1 and the encoding and rate matching of TB2 is based on a second soft buffer size for UE#2.

In step 1221, UE#1 tries to decode TB1 based on received encoded information bits associated with desired signal. In step 1222, UE#1 tries to decode TB2 based on received encoded information bits associated with interfering signal. Note that UE#1 can try to decode TB2 properly only if UE#1 knows the second code block soft buffer size from the BS. If the decoding of TB1/TB2 is not successful, UE#1 stores the soft channels bits of the encoded information bits in a soft buffer. TB1/TB2 will be retransmitted with another redundancy version and UE#1 will try to perform HARQ soft packet combining and decoding again. In step 1223, UE#1 reconstructs the contribution of the interfering signal TB2 in the received signal and performs CW-IC by subtracting the reconstructed interfering signal TB2 from the received signal. In step 1231, the BS schedules more downlink transmission for UE#1 and UE#2. In step 1232, the BS transmits more TBs to UE#1 and UE#2. Although in general the BS would not schedule more than  $P_{MAX}$  DL HARQ processes to a high-geometry UE in the NOMA scheme, exceptions may occur in some occasions. If the number of active DL HARQ processes is larger than  $P_{MAX}$ , then the UE can rearrange soft buffers allocation based on the priority setting for desired and interfering TBs.

In a first option (step 1241), the TB corresponding to a desired signal has a higher priority than the TB of an interfering signal. Assume all partitioned soft buffers (sub-buffers) have been occupied. In a first scenario, a new DL HARQ

process corresponds to a desired signal is activated. If a sub-buffer has been occupied by the DL HARQ process of an interfering signal, then the soft channel bits of the interfering signal stored at the sub-buffer are cleared and the soft channel bits of the new desired TB are stored at the cleared sub-buffer. Otherwise, the soft channel bits of the new DL HARQ process are discarded. In a second scenario, a new DL HARQ process corresponds to an interfering signal is activated. The soft channel bits of the new DL HARQ process are discarded. In a second option (step 1242), TBs corresponding to a desired signal and an interfering signal have the same priority. The soft channel bits of a newly active DL HARQ process are always discarded if all sub-buffers have been occupied.

Figure 13 is a flow chart of a method of soft buffer partitioning from eNB perspective in accordance with one novel aspect. In step 1301, a base station encodes a first code block to be transmitted to a first UE over a time-frequency resource. In step 1302, the base station encodes a second code block to be transmitted to a second UE. The first and the second code blocks are superposed. In step 1303, the base station determines a maximum number of total active downlink HARQ process for the first UE. The maximum number includes a first number of active DL HARQ processes for desired signals plus a second number of active DL HARQ processes for interfering signals. In step 1304, the base station schedules the first UE for DL transmission based on (e.g., not exceeding) the maximum number of total active DL HARQ processes.

Figure 14 is a flow chart of a method of soft buffer partitioning from UE perspective in accordance with one novel aspect. In step 1401, a UE obtains a maximum number of total active DL HARQ processes. The maximum number includes a first number of active DL HARQ processes for desired signals plus a second number of active DL HARQ processes for interfering signals. In step 1402, the UE partitions a buffer into a first number of sub-buffers having a first sub-buffer size and a second number of sub-buffers having a second sub-buffer size. In step 1403, the UE stores a first plurality of soft channel bits associated with a desired signal in one of the first number of sub-buffers. In step 1404, the UE stores a second plurality of soft channel bits associated with an interfering signal in one of the second number of sub-buffers.

Although the present invention has been described in connection with certain specific embodiments for instructional purposes, the present invention is not limited thereto. Accordingly, various modifications, adaptations, and combinations of various features of the described embodiments can be practiced without departing from the scope of the invention as set forth in the claims.

## CLAIMS

1. A method, comprising:

encoding a first code block to be transmitted to a first user equipment (UE) over a time-frequency resource by a base station;

5 encoding a second code block to be transmitted to a second UE, wherein the second code block is superposed with the first code block;

determining a maximum number of total active downlink (DL) hybrid automatic request (HARQ) processes for the first UE, wherein the maximum number includes a first number of active DL HARQ processes for desired signals for the first UE plus a  
10 second number of active DL HARQ processes for interfering signals; and

scheduling the first UE for DL transmission based on the maximum number of total active DL HARQ processes.

2. The method of Claim 1, wherein the base station schedules the first UE and the second UE for non-orthogonal multiple access NOMA operation, and wherein the  
15 first UE is a high-geometric UE and the second UE is a low-geometric UE respectively.

3. The method of Claim 2, wherein the first code block and the second code block are superposed and transmitted over the same time-frequency resource but with different transmission power levels.

20 4. The method of Claim 1, wherein the base station schedules the first UE to decode and cancel the superposed second code block via codeword level interference cancellation (CW-IC).

5. The method of Claim 1, wherein a sub-buffer size reserved at the first UE for a transport block is inversely proportional to the maximum number of total active DL  
25 HARQ processes.

6. The method of Claim 5, wherein the maximum number of total active DL HARQ processes is predefined or signaled to the first UE.

7. A method comprising:

obtaining a maximum number of total active downlink (DL) HARQ processes  
30 for a user equipment (UE), wherein the maximum number includes a first number of active DL HARQ processes for desired signals plus a second number of active DL HARQ processes for interfering signals;

partitioning a soft buffer into a first number of sub-buffers having a first sub-

buffer size and a second number of sub-buffers having a second sub-buffer size;

storing a first plurality of soft channel bits associated with a desired signal in one of the first number of sub-buffers; and

5 storing a second plurality of soft channel bits associated with an interfering signal in one of the second number of sub-buffers.

8. The method of Claim 7, wherein the UE is scheduled for non-orthogonal multiple access (NOMA) operation, and wherein the UE is a high-geometric UE.

9. The method of Claim 8, wherein the desired signal and the interfering signal are superposed and transmitted over the same time-frequency resource but with  
10 different transmission power levels.

10. The method of Claim 7, wherein the UE is scheduled by a base station to decode and cancel the interfering signal via codeword level interference cancellation (CW-IC).

11. The method of Claim 7, wherein the first and second sub-buffer sizes are  
15 inversely proportional to the first and second numbers of active DL HARQ processes respectively.

12. The method of Claim 7, wherein a ratio between the first sub-buffer size and the second sub-buffer size is configurable.

13. The method of Claim 7, wherein the UE clears one of the second sub-buffers  
20 for storing desired signals if all of the first number of sub-buffers are occupied.

14. The method of Claim 7, wherein soft channel bits of a newly active DL HARQ process are discarded if all sub-buffers are occupied.

15. A user equipment (UE), comprising:

a processor that is configured to obtain a maximum number of total active  
25 downlink (DL) HARQ processes, wherein the maximum number includes a first number of active DL HARQ processes for desired signals plus a second number of active DL HARQ processes for interfering signals;

a first number of sub-buffers having a first sub-buffer size, wherein the UE stores a first plurality of soft channel bits associated with a desired signal in one of the first  
30 number of sub-buffers; and

a second number of sub-buffers having a second sub-buffer size, wherein the UE stores a second plurality of soft channel bits associated with an interfering signal in one of the second number of sub-buffers.

16. The UE of Claim 15, wherein the UE is scheduled for non-orthogonal

multiple access (NOMA) operation, and wherein the UE is a high-geometric UE.

17. The UE of Claim 16, wherein the desired signal and the interfering signal are superposed and transmitted over the same time-frequency resource but with different transmission power levels.

5 18. The UE of Claim 15, wherein the UE is scheduled by a base station to decode and cancel the interfering signal via codeword level interference cancellation (CW-IC).

10 19. The UE of Claim 15, wherein the first and second sub-buffer sizes are inversely proportional to the first and second numbers of active DL HARQ processes respectively.

20. The UE of Claim 15, wherein a ratio between the first sub-buffer size and the second sub-buffer size is configurable.

21. The UE of Claim 15, wherein the UE clears one of the second sub-buffers for storing desired signals if all of the first number of sub-buffers are occupied.

15 22. The UE of Claim 15, wherein soft channel bits of a newly active DL HARQ process are discarded if all sub-buffers are occupied.

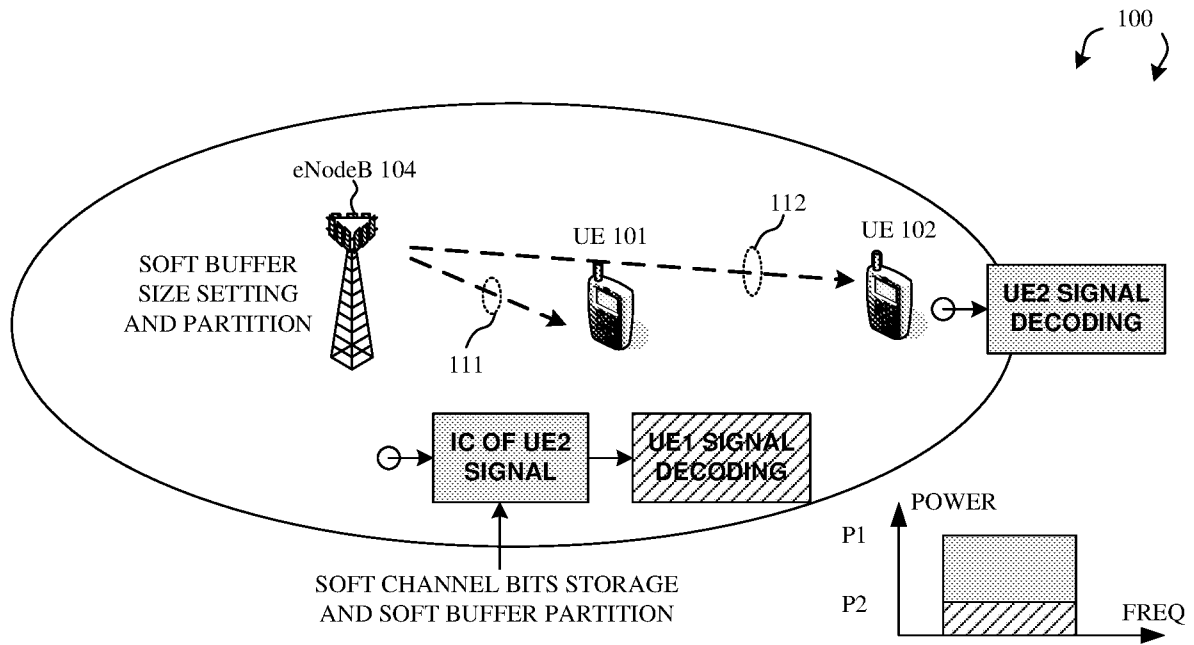


FIG. 1

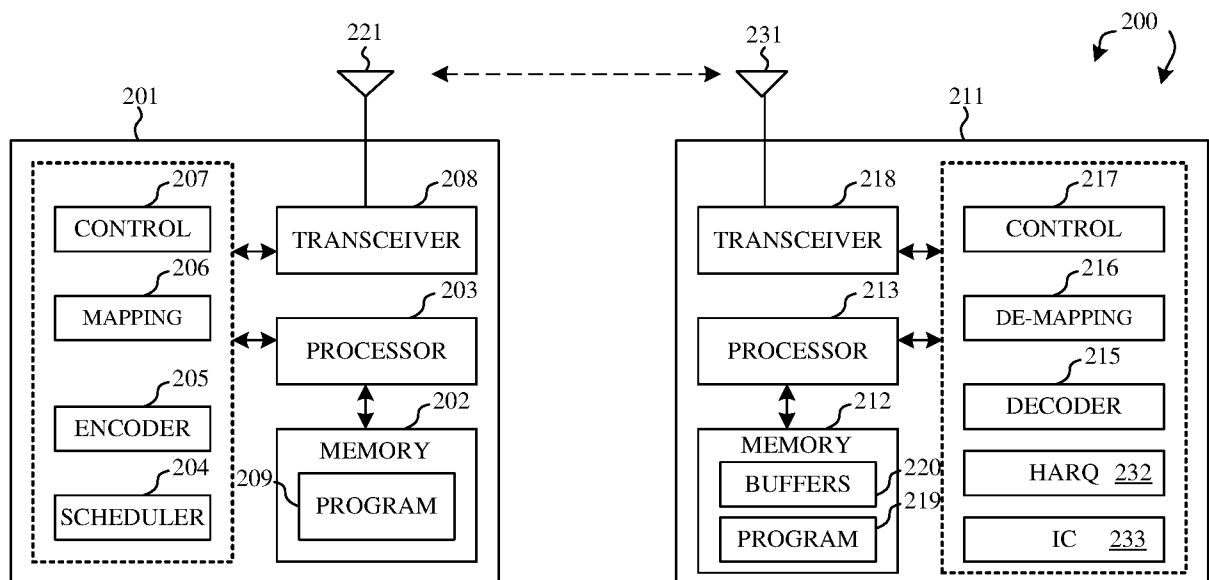


FIG. 2

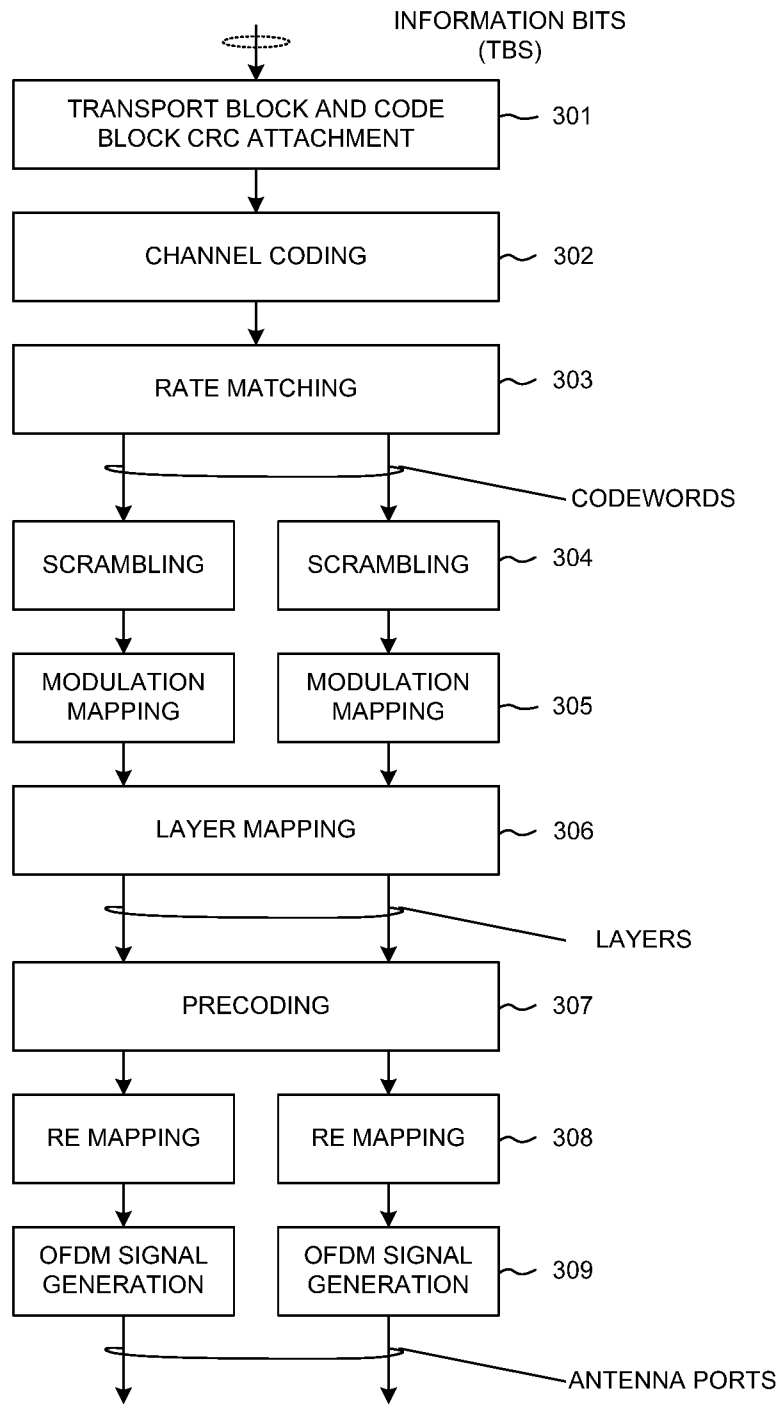


FIG. 3

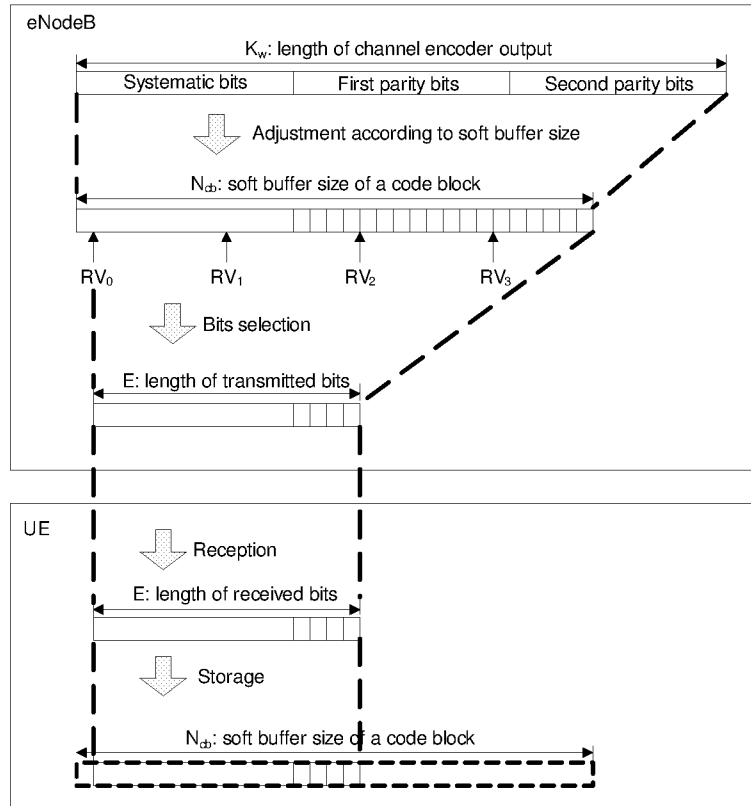


FIG. 4

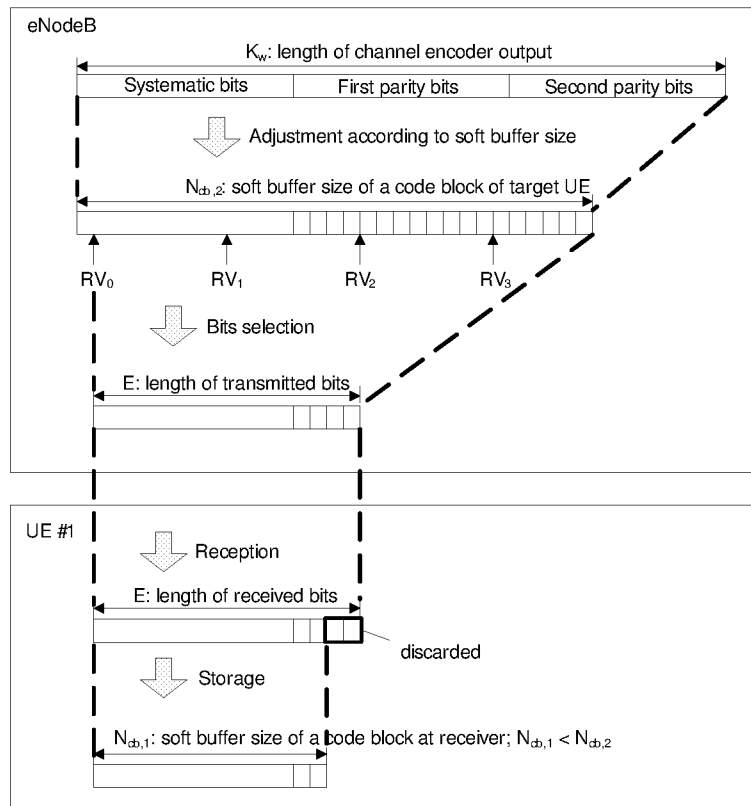


FIG. 5

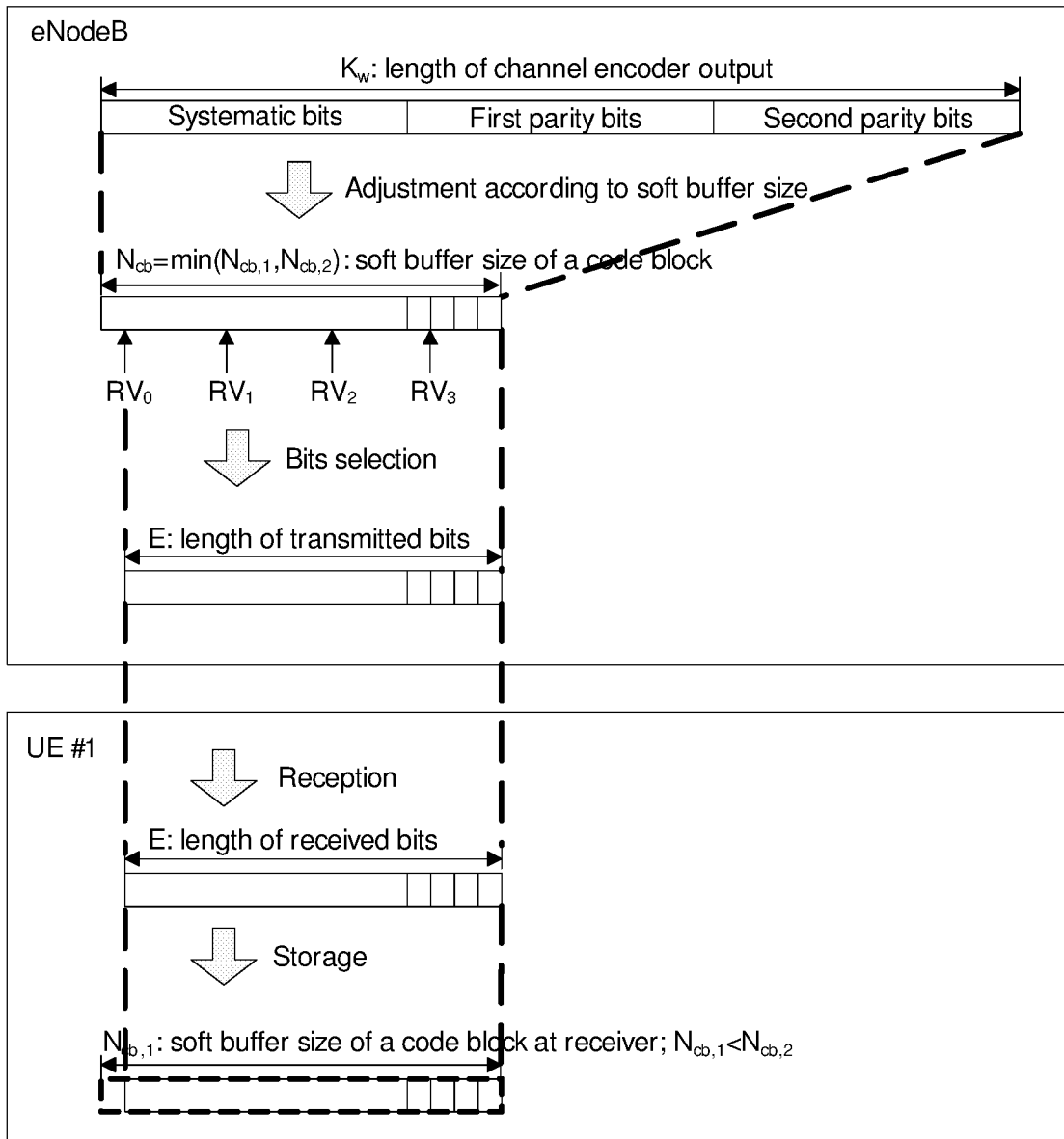


FIG. 6

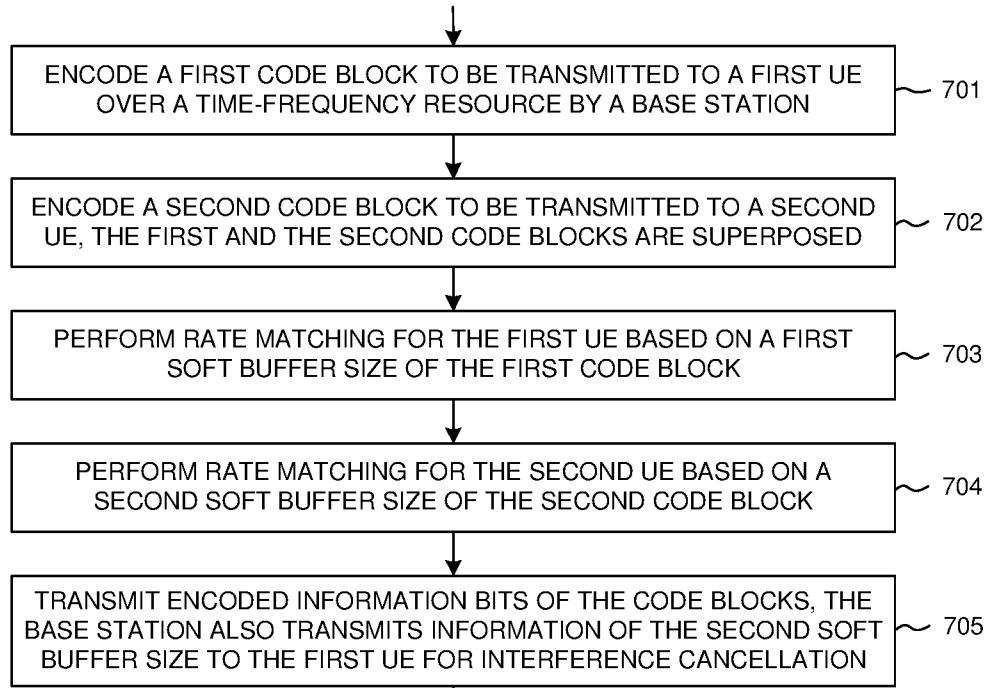


FIG. 7

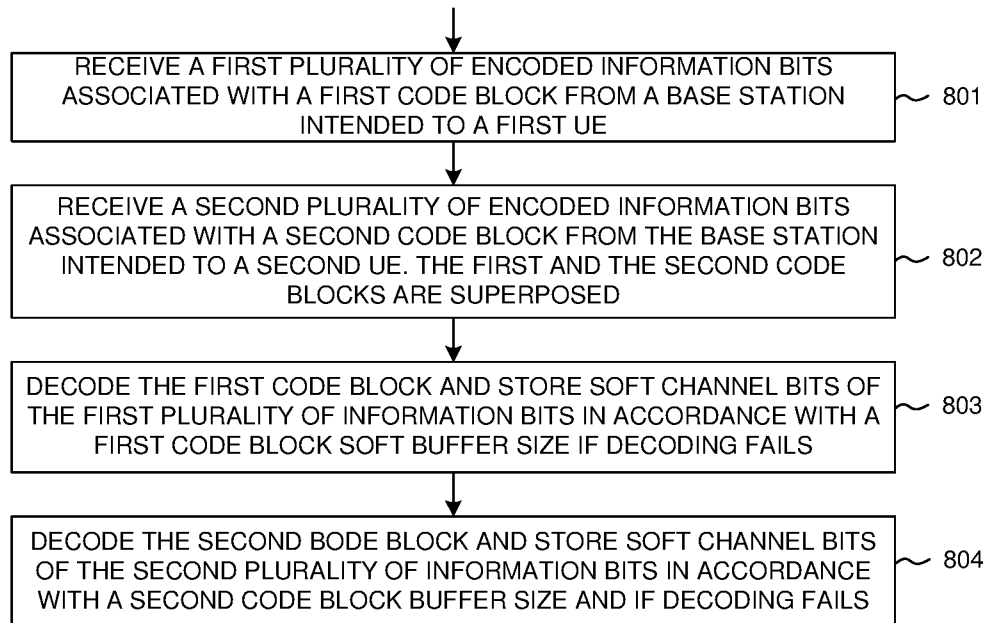


FIG. 8

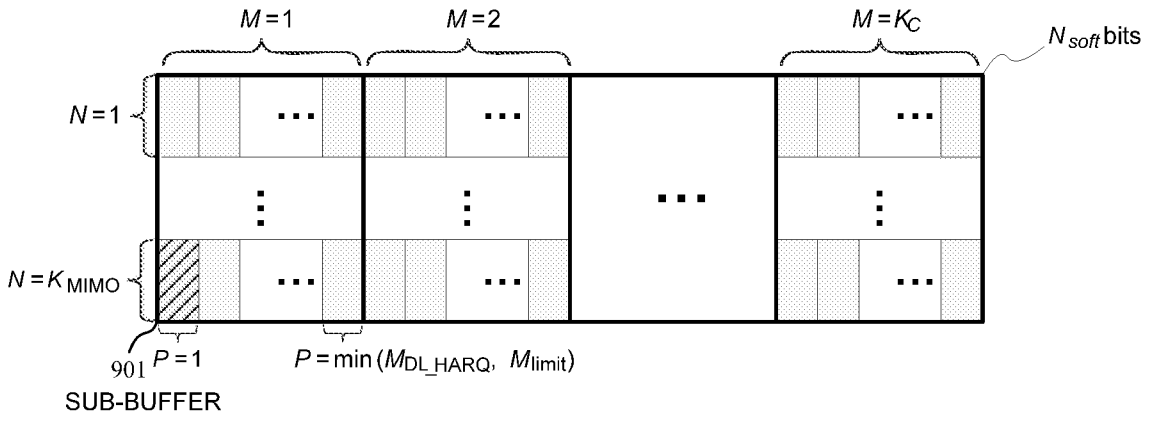


FIG. 9

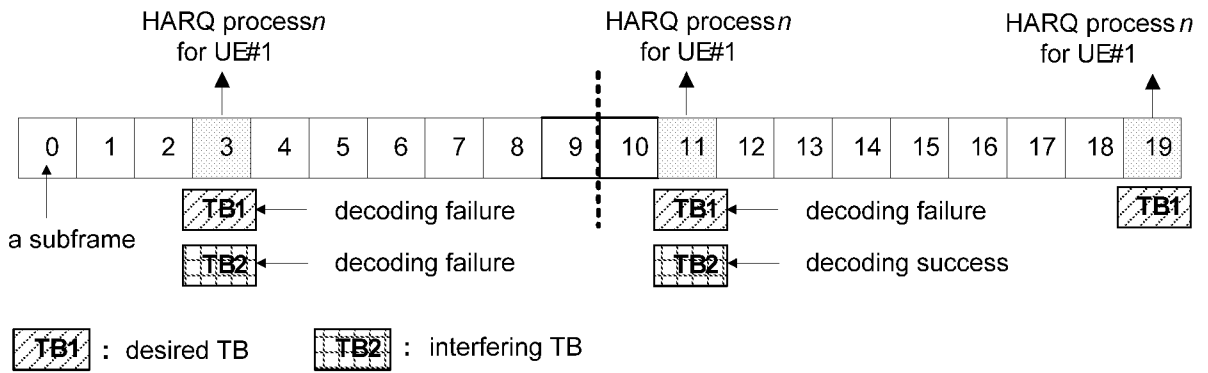


FIG. 10

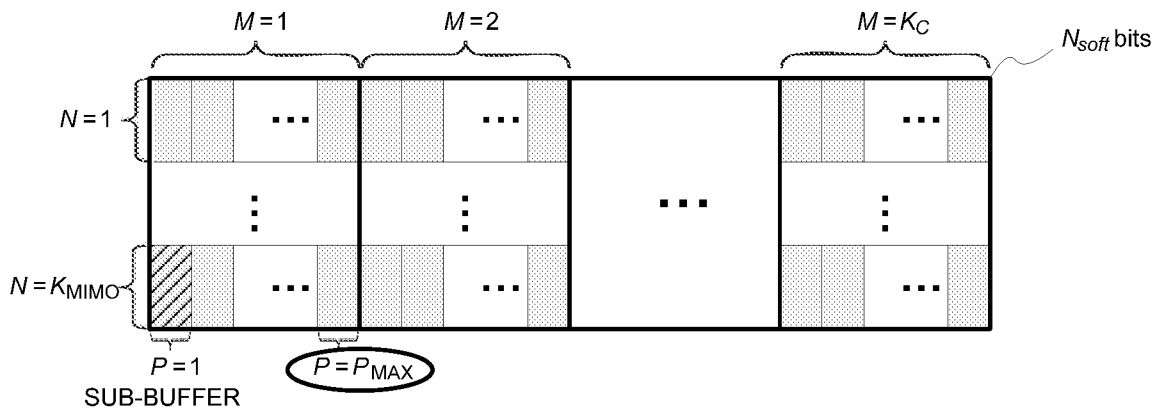


FIG. 11

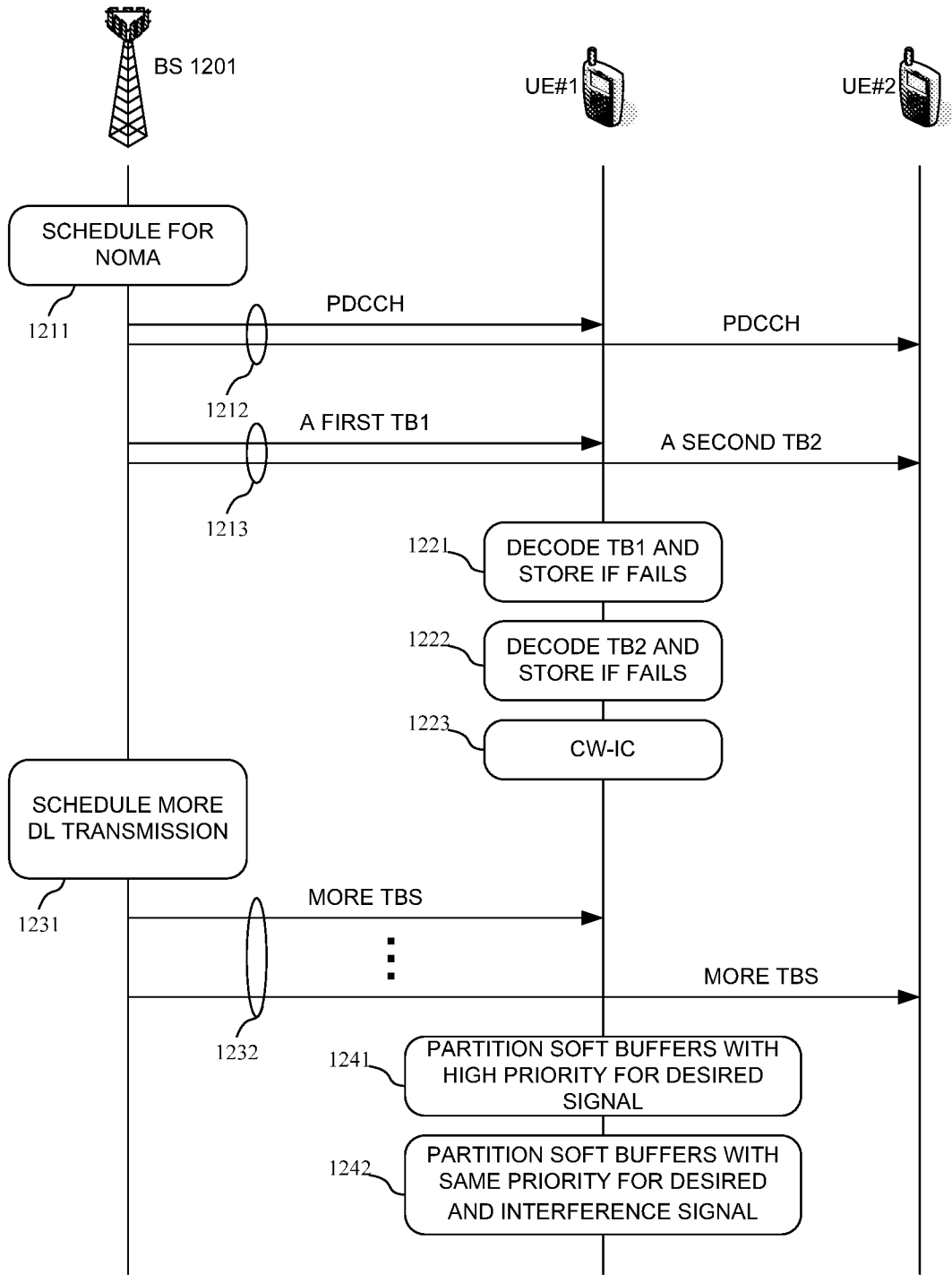
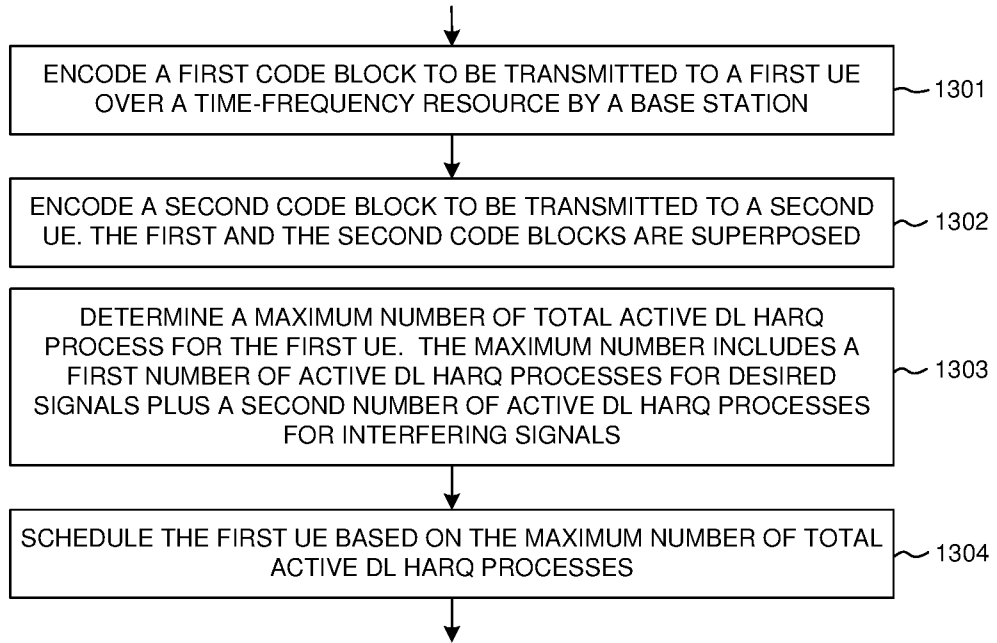
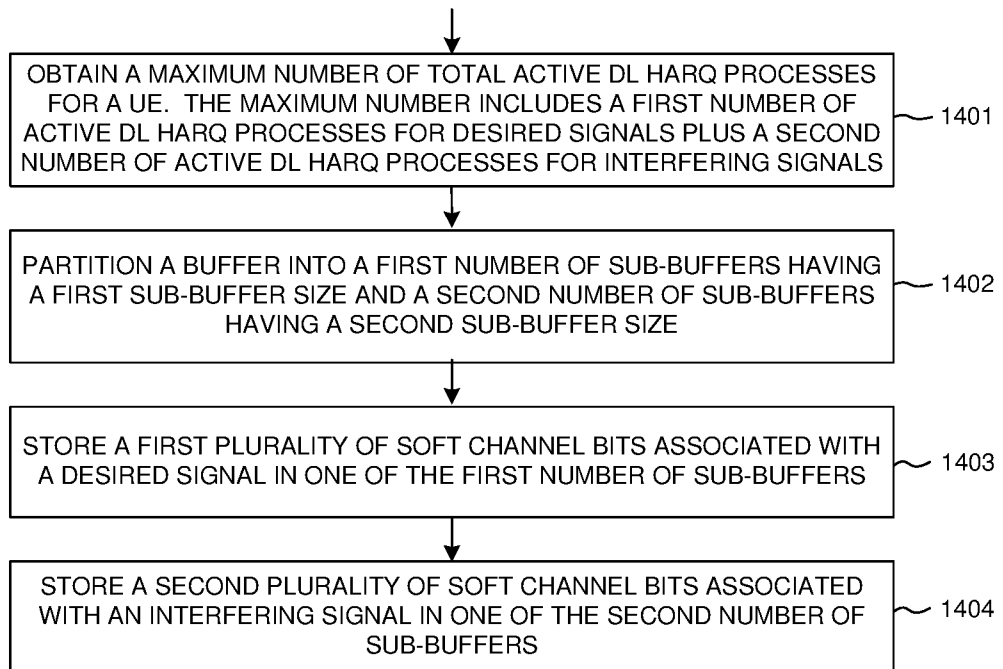


FIG. 12



**FIG. 13**



**FIG. 14**

## INTERNATIONAL SEARCH REPORT

International application No.

**PCT/CN2015/099664**

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
H04L 1/18(2006.01)i; H04W 72/12(2009.01)n		
According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols) H04L, H04Q, H04W		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) CPRSABS, CNABS, CNTXT, CNKI, VEN, WOTXT, EPTXT, USTXT, GOOGLE: cod+, block?, superpos+, downlink, HARQ, interfer????, soft, buffer, partition, size, channel bit		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 103155470 A (MOTOROLA MOBILITY INC.) 12 June 2013 (2013-06-12) see the whole document	1-22
A	CN 103051435 A (IND. TECHNOLOGY RES. INST.) 17 April 2013 (2013-04-17) see the whole document	1-22
A	US 2012057529 A1 (LG ELECTRONICS INC.) 08 March 2012 (2012-03-08) see the whole document	1-22
A	CN 102761510 A (ACER INC.) 31 October 2012 (2012-10-31) see the whole document	1-22
A	CN 101965707 A (NOKIA CORP.) 02 February 2011 (2011-02-02) see the whole document	1-22
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents:		
“A”	document defining the general state of the art which is not considered to be of particular relevance	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“E”	earlier application or patent but published on or after the international filing date	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“L”	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“O”	document referring to an oral disclosure, use, exhibition or other means	“&” document member of the same patent family
“P”	document published prior to the international filing date but later than the priority date claimed	
Date of the actual completion of the international search <b>25 February 2016</b>		Date of mailing of the international search report <b>24 March 2016</b>
Name and mailing address of the ISA/CN <b>STATE INTELLECTUAL PROPERTY OFFICE OF THE P.R.CHINA 6, Xitucheng Rd., Jimen Bridge, Haidian District, Beijing 100088, China</b>		Authorized officer <b>WANG,Hongli</b>
Facsimile No. <b>(86-10)62019451</b>		Telephone No. <b>(86-10)62088425</b>

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2015/099664**

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)				
CN	103155470	A	12 June 2013	US	8724742	B2	13 May 2014				
				KR	20130098389	A	04 September 2013				
				WO	2012047457	A1	12 April 2012				
				EP	2713541	A2	02 April 2014				
				EP	2625810	A1	14 August 2013				
				KR	101524662	B1	01 June 2015				
				US	2012087396	A1	12 April 2012				
				MX	2013003865	A	28 June 2013				
				US	2014198758	A1	17 July 2014				
				US	9078278	B2	07 July 2015				
CN	103051435	A	17 April 2013	TW	I483579	B	01 May 2015				
				TW	201310940	A	01 March 2013				
				US	2013051289	A1	28 February 2013				
				US	8885525	B2	11 November 2014				
US	2012057529	A1	08 March 2012	US	8588142	B2	19 November 2013				
				EP	2395692	A2	14 December 2011				
				EP	2395692	A4	03 December 2014				
				WO	2010110607	A3	23 December 2010				
				WO	2010110607	A2	30 September 2010				
				KR	20100107418	A	05 October 2010				
CN	102761510	A	31 October 2012	AU	2012202420	A1	15 November 2012				
				TW	I466557	B	21 December 2014				
				US	2012275397	A1	01 November 2012				
				BR	102012010079	A2	04 June 2013				
				RU	2012115861	A	27 October 2013				
				CN	102761510	B	20 May 2015				
				EP	2518927	A1	31 October 2012				
				JP	2012235454	A	29 November 2012				
				US	8891466	B2	18 November 2014				
				TW	201244507	A	01 November 2012				
				KR	20120122903	A	07 November 2012				
				IN	201201492	I4	21 March 2014				
				CN	101965707	A	02 February 2011	US	8286047	B2	09 October 2012
								EP	2266238	A2	29 December 2010
ES	2537341	T3	05 June 2015								
KR	101236310	B1	22 February 2013								
CN	101965707	B	12 March 2014								
US	2010050034	A1	25 February 2010								
WO	2009118595	A3	25 February 2010								
KR	20100126455	A	01 December 2010								
EP	2266238	B1	22 April 2015								
WO	2009118595	A2	01 October 2009								
TW	201010356	A	01 March 2010								
TW	I437861	B	11 May 2014								
IN	201006780	P4	08 July 2011								