Supplementary document of paper:
Joint source, channel and space-time coding of progressive bitstream in MIMO channels

PROOF OF LEMMA 2

From (13), it follows that

$$\begin{aligned} &D^{\mathrm{p}}_{1,2,\ldots,L-i+1}(r_i^*, r_{i+1}^*, \ldots, r_L^*;\ c_i^*, c_{i+1}^*, \ldots, c_L^*;\ \alpha) \\ &\le D^{\mathrm{p}}_{1,2,\ldots,L-i+1}(r_i^*, r_{i+1}, \ldots, r_L;\ c_i^*, c_{i+1}, \ldots, c_L;\ \alpha) \\ &\text{for any } r_{i+1}, r_{i+2}, \ldots, r_L \in \mathcal{R} \text{ and } c_{i+1}, c_{i+2}, \ldots, c_L \in \mathcal{C}. \end{aligned} \tag{49}$$

From (31), the inequality given by (49) can be rewritten as

$$\begin{aligned} &\sigma^2 p(r_i^*, c_i^*) + g\big(b(r_i^*, c_i^*)\big)\big(1 - p(r_i^*, c_i^*)\big) D^{\mathrm{p}}_{1,2,\ldots,L-i}(r_{i+1}^*, r_{i+2}^*, \ldots, r_L^*;\ c_{i+1}^*, c_{i+2}^*, \ldots, c_L^*;\ \alpha) \\ &\le \sigma^2 p(r_i^*, c_i^*) + g\big(b(r_i^*, c_i^*)\big)\big(1 - p(r_i^*, c_i^*)\big) D^{\mathrm{p}}_{1,2,\ldots,L-i}(r_{i+1}, r_{i+2}, \ldots, r_L;\ c_{i+1}, c_{i+2}, \ldots, c_L;\ \alpha) \\ &\text{for any } r_{i+1}, r_{i+2}, \ldots, r_L \in \mathcal{R} \text{ and } c_{i+1}, c_{i+2}, \ldots, c_L \in \mathcal{C}. \end{aligned} \tag{50}$$

Since $p(r_i^*, c_i^*) < 1$, $g\big(b(r_i^*, c_i^*)\big) = 2^{-\alpha r_i^* c_i^* T_{\mathrm{pkt}} W_{\mathrm{pkt}}} > 0$, and from (50), we have

$$\begin{aligned} &D^{\mathrm{p}}_{1,2,\ldots,L-i}(r_{i+1}^*, r_{i+2}^*, \ldots, r_L^*;\ c_{i+1}^*, c_{i+2}^*, \ldots, c_L^*;\ \alpha) \\ &\le D^{\mathrm{p}}_{1,2,\ldots,L-i}(r_{i+1}, r_{i+2}, \ldots, r_L;\ c_{i+1}, c_{i+2}, \ldots, c_L;\ \alpha) \\ &\text{for any } r_{i+1}, r_{i+2}, \ldots, r_L \in \mathcal{R} \text{ and } c_{i+1}, c_{i+2}, \ldots, c_L \in \mathcal{C}. \end{aligned} \tag{51}$$

Based on (51), we will prove (14) by induction on the number of packets: We first consider $L-i$ packets. Eq. (51) is identical to (14) when we let $j = i+1$ in (14) (i.e., $L-i$ packets). We next suppose that (14) holds for $j = n$ $(\ge i+1)$. In other words, for $L-n+1$ $(\le L-i)$ packets, we have

$$\begin{aligned} &D^{\mathrm{p}}_{1,2,\ldots,L-n+1}(r_n^*, r_{n+1}^*, \ldots, r_L^*;\ c_n^*, c_{n+1}^*, \ldots, c_L^*;\ \alpha) \\ &\le D^{\mathrm{p}}_{1,2,\ldots,L-n+1}(r_n, r_{n+1}, \ldots, r_L;\ c_n, c_{n+1}, \ldots, c_L;\ \alpha) \\ &\text{for any } r_n, r_{n+1}, \ldots, r_L \in \mathcal{R} \text{ and } c_n, c_{n+1}, \ldots, c_L \in \mathcal{C}. \end{aligned} \tag{52}$$

Eq. (52), which is the induction hypothesis, is identical to (13) when we let $i = n$ in (13). Since (13) implies (51), (51) holds for $i = n$, i.e.,

$$
\begin{aligned}
& D^{\mathrm{p}}_{1,2,\ldots,L-n}(r^*_{n+1}, r^*_{n+2}, \ldots, r^*_L;\; c^*_{n+1}, c^*_{n+2}, \ldots, c^*_L;\; \alpha) \\
& \leq D^{\mathrm{p}}_{1,2,\ldots,L-n}(r_{n+1}, r_{n+2}, \ldots, r_L;\; c_{n+1}, c_{n+2}, \ldots, c_L;\; \alpha) \\
& \text{for any } r_{n+1}, r_{n+2}, \ldots, r_L \in \mathcal{R} \text{ and } c_{n+1}, c_{n+2}, \ldots, c_L \in \mathcal{C}.
\end{aligned} \tag{53}
$$

Letting $j = n+1$ in (14) (i.e., $L-n$ packets), we obtain a result that is identical to (53). Hence, (14) holds for $j = n+1$. We have thus shown that (14) is valid for $j \geq i+1$. $\square$

PROOF OF LEMMA 3

From (9), it can be shown that $D^{\mathrm{p}}_{1,2,\ldots,L}(r_1, r_2, \ldots, r_L;\; c_1, c_2, \ldots, c_L;\; \alpha)$ is rewritten as

$$
\begin{aligned}
& D^{\mathrm{p}}_{1,2,\ldots,L}(r_1, r_2, \ldots, r_L;\; c_1, c_2, \ldots, c_L;\; \alpha) \\
& = \sum_{n=0}^{L-2} \sigma^2 \left(\prod_{i=1}^{n} g\big(b(r_i, c_i)\big) \right) p(r_{n+1}, c_{n+1}) \prod_{i=1}^{n} \big(1 - p(r_i, c_i)\big) \\
& + \sigma^2 \prod_{i=1}^{L-1} g\big(b(r_i, c_i)\big) \prod_{i=1}^{L-1} \big(1 - p(r_i, c_i)\big) \Big\{ p(r_L, c_L) + g\big(b(r_L, c_L)\big)\big(1 - p(r_L, c_L)\big) \Big\}.
\end{aligned} \tag{54}
$$

In addition, from (9), $D^{\mathrm{p}}_{1,2,\ldots,L-1}(r_1, r_2, \ldots, r_{L-1};\; c_1, c_2, \ldots, c_{L-1};\; \alpha)$ is given by

$$
\begin{aligned}
& D^{\mathrm{p}}_{1,2,\ldots,L-1}(r_1, r_2, \ldots, r_{L-1};\; c_1, c_2, \ldots, c_{L-1};\; \alpha) \\
& = \sum_{n=0}^{L-2} \sigma^2 \left(\prod_{i=1}^{n} g\big(b(r_i, c_i)\big) \right) p(r_{n+1}, c_{n+1}) \prod_{i=1}^{n} \big(1 - p(r_i, c_i)\big) \\
& + \sigma^2 \prod_{i=1}^{L-1} g\big(b(r_i, c_i)\big) \prod_{i=1}^{L-1} \big(1 - p(r_i, c_i)\big).
\end{aligned} \tag{55}
$$

From (54) and (55), we obtain

$$
\begin{aligned}
& D^{\mathrm{p}}_{1,2,\ldots,L}(r_1, r_2, \ldots, r_L;\; c_1, c_2, \ldots, c_L;\; \alpha) \\
& = D^{\mathrm{p}}_{1,2,\ldots,L-1}(r_1, r_2, \ldots, r_{L-1};\; c_1, c_2, \ldots, c_{L-1};\; \alpha) \\
& \quad + \sigma^2 \prod_{i=1}^{L-1} g\big(b(r_i, c_i)\big) \prod_{i=1}^{L-1} \big(1 - p(r_i, c_i)\big) \Big\{ p(r_L, c_L) + g\big(b(r_L, c_L)\big)\big(1 - p(r_L, c_L)\big) - 1 \Big\} \\
& < D^{\mathrm{p}}_{1,2,\ldots,L-1}(r_1, r_2, \ldots, r_{L-1};\; c_1, c_2, \ldots, c_{L-1};\; \alpha),
\end{aligned} \tag{56}
$$

where the inequality follows from $p(r_L, c_L) < 1$, and $0 < g\big(b(r_L, c_L)\big) = 2^{-\alpha r_L c_L T_{\text{pkt}} W_{\text{pkt}}} < 1$. $\square$

PROOF OF LEMMA 4

From (15) of Lemma 3, it is clear that for $1 \leq i \leq L-1$, we have

$$\begin{aligned} & D^{\text{p}}_{1,2,\ldots,L-i+1}(r^*_{i+1}, r^*_{i+2}, \ldots, r^*_L, r_k;\ c^*_{i+1}, c^*_{i+2}, \ldots, c^*_L, c_k;\ \alpha) \\ & < D^{\text{p}}_{1,2,\ldots,L-i}(r^*_{i+1}, r^*_{i+2}, \ldots, r^*_L;\ c^*_{i+1}, c^*_{i+2}, \ldots, c^*_L;\ \alpha) \\ & \qquad \text{for any } r_k \in \mathcal{R} \text{ and } c_k \in \mathcal{C}. \end{aligned} \tag{57}$$

From the condition of this lemma given by (13) and (57), we can derive

$$\begin{aligned} & D^{\text{p}}_{1,2,\ldots,L-i+1}(r^*_i, r^*_{i+1}, \ldots, r^*_L;\ c^*_i, c^*_{i+1}, \ldots, c^*_L;\ \alpha) \\ & \leq D^{\text{p}}_{1,2,\ldots,L-i+1}(r^*_{i+1}, r^*_{i+2}, \ldots, r^*_L, r_k;\ c^*_{i+1}, c^*_{i+2}, \ldots, c^*_L, c_k;\ \alpha) \\ & < D^{\text{p}}_{1,2,\ldots,L-i}(r^*_{i+1}, r^*_{i+2}, \ldots, r^*_L;\ c^*_{i+1}, c^*_{i+2}, \ldots, c^*_L;\ \alpha) \\ & \qquad \text{for any } r_k \in \mathcal{R} \text{ and } c_k \in \mathcal{C}, \end{aligned} \tag{58}$$

where the first inequality follows from (13), and the second inequality follows from (57).

Based on (58), we will prove (16) by induction on the number of packets: We first consider $L-i$ packets. Eq. (58) is identical to (16) when we let $j = i+1$ in (16) (i.e., $L-i$ packets). We next suppose that (16) holds for $j = n$ $(\geq i+1)$. In other words, for $L-n+1$ $(\leq L-i)$ packets, we have the following induction hypothesis.

$$\begin{aligned} & D^{\text{p}}_{1,2,\ldots,L-i+1}(r^*_i, r^*_{i+1}, \ldots, r^*_L;\ c^*_i, c^*_{i+1}, \ldots, c^*_L;\ \alpha) \\ & < D^{\text{p}}_{1,2,\ldots,L-n+1}(r^*_n, r^*_{n+1}, \ldots, r^*_L;\ c^*_n, c^*_{n+1}, \ldots, c^*_L;\ \alpha). \end{aligned} \tag{59}$$

Note that the right hand side of (59) is also a parametric distortion-based optimum for $L-n+1$ progressive packets, because Lemma 2 indicates that the condition of this lemma, which is given by (13), implies (14) for some integer $j \geq i+1$. From the fact that a parametric distortion-based optimal solution satisfies (58), and that the right hand side of (59) equals the first line of (58) when setting $i = n$ in (58), it follows that

$$\begin{aligned} & D^{\text{p}}_{1,2,\ldots,L-n+1}(r^*_n, r^*_{n+1}, \ldots, r^*_L;\ c^*_n, c^*_{n+1}, \ldots, c^*_L;\ \alpha) \\ & < D^{\text{p}}_{1,2,\ldots,L-n}(r^*_{n+1}, r^*_{n+2}, \ldots, r^*_L;\ c^*_{n+1}, c^*_{n+2}, \ldots, c^*_L;\ \alpha). \end{aligned} \tag{60}$$

From the induction hypothesis given by (59) and (60), we have

$$\begin{aligned} &D^{\mathrm{p}}_{1,2,\ldots,L-i+1}(r^*_i, r^*_{i+1}, \ldots, r^*_L;\ c^*_i, c^*_{i+1}, \ldots, c^*_L;\ \alpha) \\ &< D^{\mathrm{p}}_{1,2,\ldots,L-n}(r^*_{n+1}, r^*_{n+2}, \ldots, r^*_L;\ c^*_{n+1}, c^*_{n+2}, \ldots, c^*_L;\ \alpha). \end{aligned} \tag{61}$$

Letting $j = n+1$ in (16) (i.e., $L-n$ packets), we obtain a result identical to (61). Hence, (16) holds for $j = n+1$. We have thus shown that (16) holds for $j \geq i+1$. $\square$

PROOF OF COROLLARY 6

The condition of this corollary, given by (18), is identical to (13) of Lemma 2 when we let $i = 1$ in (13). Thus, (14) of Lemma 2 holds for some integer k in the range of $2 \leq k \leq L$ as follows:

$$\begin{aligned} &D^{\mathrm{p}}_{1,2,\ldots,L-k+1}(r^*_k, r^*_{k+1}, \ldots, r^*_L;\ c^*_k, c^*_{k+1}, \ldots, c^*_L;\ \alpha) \\ &\leq D^{\mathrm{p}}_{1,2,\ldots,L-k+1}(r_k, r_{k+1}, \ldots, r_L;\ c_k, c_{k+1}, \ldots, c_L;\ \alpha) \\ &\text{for any } r_k, r_{k+1}, \ldots, r_L \in \mathcal{R} \text{ and } c_k, c_{k+1}, \ldots, c_L \in \mathcal{C}. \end{aligned} \tag{62}$$

If we let $i = k$ in the condition of Theorem 5, given by (13), then it equals (62) in the range of $2 \leq k \leq L-1$. Note that this range of k is a subset of $1 \leq k \leq L-1$ and $2 \leq k \leq L$ given by (13) and (62), respectively. As a result, it follows from Theorem 5 that, for $2 \leq k \leq L-1$ and $k+1 \leq j \leq L$, (19) holds with k being substituted into i. In addition, from Theorem 5 and (18), it follows immediately that for $2 \leq j \leq L$, (19) holds with $i = 1$. We have thus shown that (19) holds for $1 \leq i \leq L-1$ and $i+1 \leq j \leq L$. Letting $i = 1, 2, \ldots, L-1$ and $j = i+1, i+2, \ldots, L$ in (19), we obtain at least $(L^2 - L)/2$ $(= \sum_{i=1}^{L-1} L - i)$ constraints on $r^*_1, r^*_2, \ldots, r^*_{L-1}$ or $c^*_1, c^*_2, \ldots, c^*_{L-1}$ of all the L packets except the last one. $\square$

COMPUTATION APPROACH FOR LOCAL SEARCH SOLUTION IN A MIMO SYSTEM

In Step 3 of the pseudocode in Section IV of [7], the expected distortion, denoted by $E_L[d]$, is given by [Eq. (1), 7]:

$$E_L[d] = \sum_{i=0}^{L} P_i(R_1, \ldots, R_L) f(V_i(R_1, \ldots, R_L)), \tag{63}$$

where $P_i(R_1, \ldots, R_L)$ is the probability that no decoding errors occur in the first i packets with an error in the next one, when a set of spectral efficiencies, denoted by $R_1, \ldots, R_L$, is assigned to a series of L packets in a SISO system; $f(x)$ is the operational distortion-rate function, and $V_i(R_1, \ldots, R_L)$ is the number of total source bits in the first i packets. For a MIMO system, when computing $E_L[d]$, we have replaced $P_i(R_1, \ldots, R_L)$ in (63) by $P_i(R_1, \ldots, R_L; C_1, \ldots, C_L)$; this is the probability that no decoding errors occur in the first i packets with an error in the next one, when a set of spatial multiplexing rates, $C_1, \ldots, C_L$, and a set of spectral efficiencies, $R_1, \ldots, R_L$, are assigned to L packets in a MIMO system. We have also replaced $V_i(R_1, \ldots, R_L)$ in (63) by $V_i(R_1, \ldots, R_L; C_1, \ldots, C_L)$, in which spatial multiplexing rates as well as spectral efficiencies are used to compute the number of source bits (note that just the spectral efficiencies are considered to compute $V_i(R_1, \ldots, R_L)$). Regarding a set of spatial multiplexing rates, all the packets are encoded by the same space-time code (e.g., OSTBC). Accordingly, only a set of spectral efficiencies is optimally chosen, and is assigned to L packets following the approach specified in Section IV of [7], whereas the same space-time code (e.g., OSTBC) is assigned to L packets.