Two cooperative multicast schemes of scalable video in relay-based cellular networks

Duckdong Hwang, Phuc Chau, Jitae Shin, Tae-Jin Lee

College of Information and Communication Engineering, Sungkyunkwan University, Suwon, Republic of Korea

Abstract: This study investigates the cooperative multicast of scalable video over dedicated relay-based cellular networks. Different layers in a scalable video tend to have different transmission quality requirements, and legacy cellular network has limitations in providing homogeneous service quality within its coverage area. The authors propose two realistic cooperative transmission schemes with a capacity performance analysis for scalable video over dedicated relay-based cellular networks. The proposed transmission schemes with multiple antennas at an evolved node B and the dedicated relay utilise the combinations of beamforming and space-time coding in order to enhance the multicast performance of scalable video. The proposed schemes ensure that the received base layer (BL) of scalable video guarantees the minimum video quality to all user terminals throughout the cell coverage with given outage probability, while the terminals inside the relay coverage can achieve better video quality by receiving both the BL and additional enhancement layer data according to the received channel condition. The antennas configuration and the video type determine which multicast scheme to choose.

1 Introduction

For the joint transmission of data composed of unequal quality requirements, such as scalable or layered video, hierarchical modulation (HM) [1] can be considered. In HM, a set of different constellations are superposed with different minimum distances among the constellation points and one class of the dataset can be mapped to one constellation. The differentiation of data at the modulation level ensures unequal error rate performances among the classes of datasets for the additive white Gaussian noise channel. However, in fading wireless channels, where random fading dominates over additive noise, other means are required to provide unequal error rate performances. The error rate performance or transmission reliability of wireless channels is often captured by the diversity gain.

To enhance the reliability in fading wireless channels, a multi-transmitting antenna-based diversity technique called as space-time block coding (STBC) was introduced [2, 3]. The advanced multiple-input–multiple-output (MIMO) wireless physical layer technique has enabled the significant increase in spatial diversity and spatial multiplexing (SM) [4]. Hence, layered scalable video transmission schemes over MIMO systems have been proposed in several research works, such as STBC-MIMO [5] and singular value decomposition (SVD)-based MIMO [6]. Song and Liu [5] demonstrated that the STBC-MIMO was able to achieve the desired video quality in terms of peak signal-to-noise ratio (PSNR) compared to a single-input–single-output (SISO) system. The SVD-based MIMO approach [6] could be adopted for maximising the throughput delivery of scalable video. Song and Chen [7] proposed an automatic unequal error protection (UEP) using adaptive selection among the MIMO channels, because the different layers in a scalable video have unequal importance. Based on the channel state information, the scalable layers are allocated in such a way that the base layer (BL) is transmitted over the best channel, the first enhancement layer (EL) over the second best channel and so on in order to obtain better video quality. However, the aforementioned studies considered the scalable video transmission for unicasting. In multicast over the wireless medium, the major challenge arises from the heterogeneous channel condition of different mobile stations (MSs). In conventional multicast schemes, the lowest transmission rate is used for BL of the scalable video to ensure all receivers decoding successfully with a minimum quality of video contents [8]. Zhou et al. [9] considered the multicast of scalable video contents. The branch-and-bound-based adaptive channel scheduling algorithm was proposed to map scalable video layers to appropriate MIMO antennas to provide flexible UEP according to the feedback of signal-to-noise ratio (SNR) powers strength. Yang and Wang [10] proposed a cross-layer design framework to enhance the scalable broadcasting over MIMO systems. Chen and Chung [11] proposed a novel dual-diversity space-time coding to exploit the intrinsic UEP capability by adopting a dual constellation pair of two consecutive transmission periods. The MS experiencing good channel is able to decode both layers and MSs in bad channel condition enjoys BL. However, the above-mentioned works result in reduction of allowable minimum throughput for basic video quality, for example, BL, when the number of users increases.

For cooperative transmission, relay operation can provide diversity gain in the data reception and system performance [12, 13]. Two-stage cooperative multicast transmission was proposed to achieve optimised power consumption and guaranteed coverage [14]. The evolved node B (eNB) (base station) broadcasts data to all MSs in the first stage. Then the MS decoding successful at the first stage will become a mobile relay to forward data to remain MSs in the second stage. The system throughput can be improved significantly with the assistance of a relay. The integration of network coding (NC), MIMO and HM for scalable video transmission over cellular relay networks could be found in [15], in which the proposed scheme provided better video quality gains than a SISO single-relay network system. The superposition of STBC and SM was introduced for the transmission of a scalable extension of H.264/AVC (SVC) transmission over wireless networks [16], where an MS opportunistically helps the destination MS by the amplify-and-forward relay operation. Unlike HM, this approach is effective in wireless channels because of the diversity from STBC and the relay in cellular relay networks. The two tiers of reliability grades for the base and the ELs of SVC are well matched with the multicast scenario where MSs in a multicast group experience different channel SNR and are required to obtain opportunistic video quality according to the SNR. The BL in SVC...
provides a minimum video quality targeting the MS in the worst channel condition, while the MS with a good channel gets better video quality with additionally received ELs. It is important that the cooperative multicast scheme in cellular relay networks is properly designed to match well with scalable video transmission.

In this paper, we cast a similar approach to that of [16] to the multicast of SVC in the dedicated decode-and-forward (DF)-assisted cellular networks. We superpose STBC and beamforming for the eNB transmission and allow the relay station (RS) to have multiple antennas. The beamforming at the eNB is controlled by the feedback from the RS. The number of antennas at the base station and the dedicated relay are limited by the size of the available STBC design. Compared with those in [16], where the unicast transmission of SVC with an opportunistic MS relaying with two eNB antennas is considered, our proposed schemes are more practical and match well with the real cellular multicast environment for scalable video. Moreover, the method in [16] lacks a capacity analysis, which assesses the performance limitations of the proposed transmission scheme with modern channel codes. In [8, 17–19], relay-based SVC multicast in the cellular network is considered without multi-antenna techniques at the base station. Therefore this paper provides an analysis of two MIMO-based scalable video multicast schemes that are better fits to the real cellular multicast environment, and ensures that the BL of scalable video is served throughout the cell coverage area within a given outage probability while the MSs near the relay coverage enjoy both layers. Depending on the video type and antenna configuration, two transmission modes are proposed for the eNB to choose between the two modes for better visual quality at MSs. The video type can be categorised according to the amount of movements in the scene, where the videos with lots of movements result in high BL data rates while the videos with mild movements generate moderate BL rates. We provide the capacity performances of the proposed schemes so that it is clear how the schemes work in real world cellular systems simply parameterised by the eNB and RS signal powers.

**Notations:** A bold lowercase letter represents a vector and a bold uppercase letter represents a matrix. The notations $A^T$, $A^H$ and $\text{Tr}$ of $A$ are the transpose, the Hermitian transpose and the trace of matrix $A$, respectively. The notation $\|A\|_F = \text{Tr}(A^H A)$ denotes the Frobenius norm of $A$. The identity matrix with dimension $a$ is denoted $I_a$. $\|a\|$ denotes the norm of a vector $a$. $\mathcal{CN}(0, 1)$ denotes the circular symmetric complex Gaussian random variable with a zero mean and a variance of one.

## 2 Proposed cooperative transmission schemes

Fig. 1 depicts a cellular network with an RS, in which the SVC bitstreams are transmitted from the eNB towards MSs within the cell coverage area. And the RS helps the mobile users within the coverage area of the RS to enhance the video quality of the received data, which sacrifices some diversity order in order to support the wide range covered by the multiple RSs. Cooperative transmissions from the DF at RS enhance the reception quality of the MSs within its coverage area. The eNB is assumed to be equipped with multiple antennas, and thus it employs multiple-antenna transmission techniques. The SVC bitstreams are composed of a base and ELs so that most of the MSs within the cell coverage can decode the BL and meet the minimum required video quality, while some of the MSs (such as those that are helped by the relay transmission) can decode full layers of video signal and enjoy higher quality video. Let the number of antennas at the eNB, RS and the MS be $M_b$, $M_r$ and $M_m$, respectively. Most likely, however, $M_m = 1$ in practice. The matrix channels from the eNB to the RS and MSs are denoted $H_r$ and $H_m$, respectively, and the channel matrix from RS to MSs is denoted $G$. The elements of those channel matrices are assumed to be independent identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$. The eNB power is denoted $P_b$ and the relay power is denoted $P_r$. Note that the values of $P_b$ and $P_r$ are determined by actual cell design which is largely dependent on non-technical factors such as geography and economic consideration. In this work, these two values parameterise the relay-based cellular systems. The transmission is composed of two phases, where the eNB broadcasts the two layers of SVC in the first phase and the RS transmits the BL of the SVC in the second phase towards the MSs in its coverage. Depending on BL and EL data rates, we have two transmission modes described in Sections 2.1 and 2.2, respectively.

### 2.1 Transmission mode I

Here, we propose a transmission scheme for videos, the BL data rate of which impacts much on the video quality such as videos with lots of movement. The transmission of SVC from the eNB is made by a superposition transmission of the BL and the EL signals. In the transmission mode I, the signal matrix transmitted by the $M_b$ antennas at the eNB for a $T$ symbol duration of the first phase is composed of two $M_b \times T$ matrices as

$$f_i x_i + \sqrt{\rho} / M_b y_e$$

where the first matrix $f_i x_i$ is the pre-coded (by the $M_b \times 1$ vector $f_i$) BL signal $x_i = [x_1, ..., x_T]^T$ and the second matrix $y_e$ is the $M_b \times T$ orthogonal space-time block-coded (OSTBC) signal of the EL signal $y_1, ..., y_T$. Depending on $M_b$, there are available OSTBC designs and $T$ is determined by the OSTBC design selected. The code rate $R$ of OSTBC is defined as $K/T$ and the famous Alamouti code for the $M_b = 2, T = 2$ case is given as

$$Y_e = \begin{bmatrix} y_1 & -y_2 \\ y_2 & y_1 \end{bmatrix}$$

The precoder $f_i$ is a unit norm $M_b \times 1$ vector whose direction is controlled by feedback from the RS. Most likely, RS commands the eNB so that $f_i$ directs towards the right singular vector of $H_r$, with the largest singular value (maximum ratio transmission, MRT). The beamforming is controlled by the feedback information from the RS. If the eNB–RS link is stable, then the eNB and the RS can steer the eNB beamformer very closely towards the right singular vector direction. By taking MRT at the eNB and maximum ratio combining (MRC) at the RS, a large diversity order (up to $M_b M_r$) and, thus, strongly reliable reception of the BL at the RS is guaranteed. If there are multiple RSs within a cell, we can set $f_i$ to be a linear combination of directions towards the RS channels. The symbol powers of the BL and the EL are defined as $P_b = E[|x|^2] = P_r = E[|y|^2]$, and they should meet $P_b = P_r (1 + \rho)$, where $\rho$ is the power dividing factor between the BL and ELs. Note that the symbols $(x_i, y)$ are channel-coded by strong codes, such as low-density parity check (LDPC) codes or turbo codes.
The signal matrices received at the RS and MSs in the first phase are given, respectively, as

\[ R_i = H_i (f_i x_i + \sqrt{\rho/M_i} y_1 + n_i) \]

\[ R_1 = H_1 (f_1 x_1 + \sqrt{\rho/M_1} y_1 + n_1) \]

(2)

Here \( N_i \) and \( N_1 \) are \( CN(0, 1) \) distributed white noise matrices at RS and MSs with dimensions \( M_i \times T \) and \( M_1 \times T \) respectively. RS applies unit norm receive beamformer vector \( g_i \) of dimensions \( M_i \times 1 \) to \( R_i \), whose direction is matched to the left singular vector of \( H_i \) with the largest singular value of MRC. Let the largest singular value of \( H_i \) be \( \lambda_i \), then the signal vector after the receive beamformer is

\[ r_i = \lambda_i^{-1} R_i = \lambda_i^{-1} + \sqrt{\rho/M_i} h_i y_1 + n_i \]

(3)

Here \( h_i = g_i^H H_i \) and \( n_i = g_i^H N_i \). Using (3), the RS decodes the BL symbol vector \( x_0 \) as

\[ \hat{x}_0 = \min_{x_0} \| (r_i - \lambda_i^{-1}) x_0 - (r_i - \lambda_i^{-1}) \| \]

(4)

Here \( X \) is the set composed of all the vectors that \( x_0 \) can take and the noise correlation matrix \( K = E[h_i^H Y_1 h_i] + I_T \). Upon decoding \( x_0 \), RS encodes the decoded symbol vector \( \hat{x}_0 \) for \( T_2 \) symbol duration of the second phase into another OSTBC matrix \( X_0 \) of size \( M_1 \times T_2 \) and sends it to the MSs inside its coverage area. At MSs, the received signal in the second phase is given as

\[ R_2 = \frac{1}{\sqrt{M_2}} G X_0 + N_2 \]

(5)

Here \( N_i \) is \( CN(0, 1) \) distributed white noise matrix at MSs with dimensions \( M_2 \times T_2 \) and \( E[Tr(X_0 X_0^H)] = M_2 P_t \). Note that the base and the relay powers, \( P_r \) and \( P_i \) respectively, denote the products of the actual powers at the sources (eNB and RS) and the path loss towards the terminals.

MSs can be inside the coverage area of an RS, or some are far away from the RS. In the latter case, the RS transmissions will be weakly heard by those MSs. For both cases, \( R_2 \) in (5) is used to decode the BL. Since \( X_0 \) is OSTBC-coded, the standard decoding principles are available to recover the vector \( x_0 \) [2, 3, 16]. Let the estimate vectors of received signal from eNB and RS be \( \hat{x}_0 \) and \( \hat{x}_1 \), respectively. From the first phase observation, MSs can make another estimate of the vector as \( \hat{x}_0 = g_1^H R_1 \). We make the MRC of these two estimate as \( \hat{x}_0 = \sqrt{\rho} \hat{x}_0 + \sqrt{1/\rho} \hat{x}_1 \) to reach the final decisions with this combined estimate of \( \hat{x}_0 \). Based on these new decisions, we subtract out the term with \( x_0 \) in \( R_2 \) of (2), which results in a new signal

\[ \hat{R}_1 = \sqrt{\rho/M_1} H_1 Y_0 + N_1 \]

Again, the standard decoding principles of OSTBC provide an estimate of \( Y_0 \) as well. At this point, the MSs have acquired both the base and the ELs of SVC.

2.2 Transmission mode II

In the transmission mode II, both the BL and EL data rates are important such as videos with mild movement. To balance the both layers, we switch the loading of the BL and the EL as

\[ \sqrt{\rho} x_1 + \sqrt{1/\rho} y_1 \]

(6)

where the first matrix \( f_1 x_1 \) is the pre-coded (by the \( M_1 \times 1 \) vector \( f_1 \)) EL signal \( x_1 = [y_1, \ldots, y_T] \) and the second matrix \( Y_1 \) is the \( M_2 \times T \) OSTBC signal of the BL signal \( y_1, \ldots, y_T \). The same definitions used in the previous subsection are applied. The signal models in (2) and (3) are the same except for the signalling matrices \( x_1 \) and \( Y_1 \) in the corresponding positions.

Similarly, the RS decodes the EL symbol vector \( x_1 \) this time as

\[ \hat{x}_1 = \min_{x_1} |(r_1 - \lambda x_1) k^{-1} (r_1 - \lambda x_1) | \]

(7)

Here \( X \) is the set composed of all the vectors that \( x_1 \) can take and the noise correlation \( k = E[h_1^H Y_1^H Y_1] + I_T \). Upon decoding \( x_1 \), RS encodes the decoded symbol vector \( \hat{x}_1 \) into another OSTBC matrix \( X_1 \) of size \( M_1 \times T_2 \) and sends it to the MSs inside its coverage area in the second phase. The received signal at MSs is the same as (5) with \( X_1 \) in the corresponding positions. The signal at MSs is the same as in (5) except for \( x_1 \) in the position.

From \( R_1 \) in (2), MSs decode the OSTBC-coded \( Y_1 \) first using the standard decoding principles. Let this estimate matrix be \( \tilde{X}_1 \). On the basis of this new estimate, we subtract out the term with \( x_1 \) in \( R_1 \) of (2), which results in a new signal

\[ \tilde{R}_1 = H_1 f_1 x_1 + N_1 \]

From this, MSs can make an estimate of the vector as \( \tilde{x}_1 = f_1^H H_1 \). From (5), MSs makes another estimate of \( \tilde{x}_1 \) using the standard OSTBC decoding. Note that \( \tilde{x}_1 \) and \( \hat{x}_1 \) are the estimates of received signal from eNB and RS in both phases, respectively. Again, we make the MRC of these two estimate as \( \tilde{x}_1 = \sqrt{\rho} \tilde{x}_1 + \sqrt{1/\rho} \hat{x}_1 \) to reach the final decisions with this combined estimate of \( \tilde{x}_1 \). At this point, the MSs have acquired both the base and the ELs of SVC.

3 Performance evaluation

In this section, we provide the capacity performances of the proposed schemes, while Section 4 concentrates on its impact on scalable video. Modern channel codes like LDPC or turbo codes provide huge coding gains in Rayleigh fading channels [20] and easily approach the vicinity of the channel capacity with negligible error rate performances. Therefore the capacity analysis is sufficient to concisely capture the transmission performance of modern channel-coded wireless systems.

3.1 Capacities of scalable video layers

In the transmission mode I, the relay attempts to decode the BL signal vector \( x_0 \) with the received signal (3). It is clear that the received signal power is \( \lambda^2 P_t \) and the noise plus interference (the EL signal \( x_1 \) is the interference, in this case) is \( \rho |h|^2 P_i + 1 \). Therefore the relevant channel capacity of the BL at the relay is given as

\[ C_{1,1} = \frac{1}{2} \log \left( 1 + \frac{\lambda^2 P_t}{\rho |h|^2 P_i + 1} \right) \]

(8)

where (1/2) comes from the fact that relay operates in the half-duplex mode.

On the other hand, the 4th MS decodes the BL signal with \( R_1 \) in (2) and \( R_2 \) in (5). From \( R_1 \) the signal power is \( ||H_1 f_1 s_1||^2 P_s \) and the noise plus interference power is \( ||H_0^H H_1 f_1 s_1||^2 P_s + 1 \). which results in the signal-to-interference and -noise power ratio (SINR) of

\[ \frac{||H_1 f_1 s_1||^2 P_s}{||H_0^H H_1 f_1 s_1||^2 P_s + 1} \]

The approximation holds in reasonably high \( P_s \). From \( R_2 \) the signal power is \( ||G_2||^2 P_s \) and the noise power is 1, which results in the signal-to-interference and -noise power ratio (SINR) of \( ||G_2||^2 P_s \). Therefore the capacity of the BL at the MS2 from the combined
Once the BL is decoded at MS, the MS tries to decode the EL with $R_e$, the SNR of which is calculated as $\rho ||H||_2 ||P||_2^2$ and results in the capacity expression for the EL at MS as

$$C_{k,b,1} = \frac{1}{2} \log \left( 1 + \frac{\rho ||H||_2^2 ||P||_2^2}{\rho ||H||_2 ||P||_2} \right)$$

In transmission mode II, the relay tries to decode the EL signal vector $x$, with the received signal. It is not easy to see the EL channel capacity $C_{k,b,2}$ is the same as the BL channel capacity $C_{k,b,1}$ given in (8). At MS, the BL signal is decoded first and thus the BL capacity at MS is given as

$$C_{k,b,2} = \frac{1}{2} \log \left( 1 + \frac{\rho ||H||_2^2 ||P||_2^2}{\rho ||H||_2 ||P||_2} \right)$$

In mode II, the relay decodes EL data, the capacity of which is given as

$$C_{k,e,2} = \frac{1}{2} \log \left( 1 + \frac{\rho P_{e}}{\rho ||H||_2 ||P||_2} \right)$$

Finally, the EL capacity at MS is given as

$$C_{k,e,1} = \frac{1}{2} \log \left( 1 + ||G||_2^2 ||P||_2 + \rho ||H||_2 ||P||_2^2 \right)$$

From (9) and (10), it is noted that mode I emphasises the BL capacity $C_{k,b,1}$, while, from (11) and (13), we can see that mode II underscores the EL capacity $C_{k,e,2}$. We will show in Section 4 that mode I is advantageous when the number of antennas in the system increases. The reason behind this fact comes from comparing (10) and (11), where the former one is more susceptible to the number of antennas than the latter one is.

$$C_{k,b,1} = \frac{1}{2} \log \left( 1 + ||G||_2^2 ||P||_2 + \rho ||H||_2 ||P||_2^2 \right)$$

### 3.2 Outage events and the data rates of the layers

Suppose the transmission of the original video signal requires the total capacity of $R_b$ bps/Hz, and we divide the original video into a scalable video with a base rate $R_b$ bps/Hz and an enhancement rate of $R_e$ bps/Hz, where $R_b + R_e$ in general. In mode I, we can define three different types of outage events:

1. $O_{b,1}$ is the event where $C_{k,b,1} < R_b$ and $O_{b,1}$ is the complement event of $O_{b,1}$.
2. $O_{b,2,1}$ is the event where $C_{k,b,1} = R_b$ with $O_{b,1}$.
3. $O_{k,e,1}$ is the event where $C_{k,e,1} = R_e$ with $O_{b,2}$.

We must allocate $R_b$ and $R_e$ such that the MS near the outskirt of the macro cell area is in $O_{b,1}$ outage with less than a certain probability, while the MS near the edge of the relay coverage area is in $O_{b,2}$ outage with less than a certain probability. For mode II, we can define $O_{e,2}$, $O_{b,2,2}$, and $O_{b,2,3}$ similarly.

From Figs. 2–4, the cumulative density functions (CDF) of $C_{k,b,1}$ and $C_{k,e,2}$, respectively, with different network configurations are plotted. These curves allow us to assess the probabilities $P(O_{b,1})$ and $P(O_{b,2})$ of the BL and the EL outage events, respectively. From Fig. 2, we can see that the outage probability of $O_{b,1}$ will be negligible by keeping the eNB to RS link channel strong enough and the base rate $R_b$ small enough. For example, we can keep $P(O_{b,1}) < 0.02$ with $R_b \le 0.65$ bps/Hz when $P_b = 6 \text{ dB}, \rho = 0.2, M_b = 2, M_e = 1, \text{ and } M_m = 1$ (the dotted curve with plus markers). Once, $P(O_{b,1})$ is made to be negligible, we can use Fig. 3a ($M_b = 2, M_e = 1, M_m = 1$ case) or Fig. 4a ($M_b = 4, M_e = 2, M_m = 2$ case) to determine the $R_b$ and $R_e$ that satisfy the outage probability constraints. Consider the same case as above ($P_b = 6 \text{ dB}, \rho = 0.2, M_b = 2, M_e = 1, \text{ and } M_m = 1$) when $P(O_{b,1})$ is on the outskirt of the cell coverage area (let us say $P_f = 3 \text{ dB}$ here) and is $< 10^{-1}$ and $P(O_{b,2})$ is on the outskirt of the RS coverage area (also, $P_h = 6 \text{ dB}$ here) and is $< 10^{-1}$ (the dash-dot curve with plus markers for BL and the dashed curve with cross markers for EL, respectively). We then, can use Fig. 3a to find that $R_b \le 1.1$ bps/Hz and $R_e \le 0.1$ bps/Hz satisfy the given constraints.

Increasing the number of antennas in the network increases the rates $R_b$ and $R_e$. For $M_b = 4, M_e = 2, M_m = 2$ with $\rho = 0.2$ and $P_b = 6 \text{ dB}$ (the solid curve with plus markers), we can keep $P(O_{b,1})$
negligible with $R_b \leq 1.9$ bps/Hz, as shown in Fig. 2. Then, $R_b \leq 2.45$ bps/Hz and $R_e \leq 2$ bps/Hz are found to be sufficient from Fig. 4a. On the other hand, $\rho$ provides a trade-off between $R_b$ and $R_e$ since increasing $\rho$ stresses $R_e$ at the expense of $R_b$. Similarly, we can find $R_b$ and $R_e$ values for mode II transmission from Figs. 3b and 4b. For instance, $M_b = 4$, $M_r = 2$, $M_m = 2$ with $\rho = 0.05$ and $P_b = 6$ dB, we can keep $P(O_{b,2})$ and $P(O_{e,2})$ negligible with $R_b \leq 2.1$ bps/Hz and $R_e \leq 1.85$ bps/Hz from Fig. 4b. Finally, the total

Fig. 3  CDF curves of the BL and the EL capacities of

a Mode I
b Mode II in case of two eNB antennas and a single antenna at each RS and MS ($M_b = 2$, $M_r = 1$, $M_m = 1$)
BL (3, 0.5) in the legend stands for a BL with $P_b = 3$ dB and $\rho = 0.5$
EL (6, 0.2) in the legend stands for an EL with $P_b = (1 + \rho)P_x = 6$ dB and $\rho = 0.2$

Fig. 4  CDF curves of the BL and the EL capacities of

a Mode I
b Mode II in case of two eNB antennas and a single antenna at each RS and MS ($M_b = 2$, $M_r = 1$, $M_m = 1$)
BL (3, 0.5) in the legend stands for a BL with $P_b = 3$ dB and $\rho = 0.5$
EL (6, 0.2) in the legend stands for an EL with $P_b = (1 + \rho)P_x = 6$ dB and $\rho = 0.2$
throughputs of the BL and the EL are given as $R_b$ Wbps and $R_e$ Wbps when WHz is allowed for the scalable video transmission.

4 Achievable bit-rate and video quality performance

On the basis of the analysis of Sections 3.1 and 3.2, we provide the performance of the corresponding quality of scalable video. For video quality analysis of the proposed transmission schemes, we use scalable video bitstreams encoded as two layers with the bit-rates of BL and EL by JSVM 9.19 reference software [21] with SNR scalability, and the decoded video quality is measured by a widely used objective metric, PSNR. The test sequences, Soccer and Mobile, with 150 frames, each, are used with CIF resolution ($352 \times 288$ pixels) and are encoded with a group of picture size of 16 frames with a hierarchical prediction structure. Note that Soccer sequence has lots of movement and is expected to require high BL.

![Fig. 5] Total bit-rates of BL and EL as a function of SNR

Elements inside the parenthesis in the legend are ordered as $(M_b, M_r, \rho)$ with $M_1, M_2$ for mode I and mode II, respectively.

![Fig. 6] Scalable video PSNR performance of the worst-case MSs for a Soccer video

Elements inside the parenthesis in the legend are ordered as $(M_b, M_r, \rho)$ with $M_1, M_2$ for mode I and mode II, respectively.

a Soccer video

b Mobile video
data rate while Mobile sequence has mild movement and both layers have similar importance.

From the results of Figs. 3 and 4, we can derive the corresponding capacities of BL and EL for a given outage probability, which is matched directly with the CDF of capacity. The bandwidth of this performance is assumed to be 100 kHz in consideration of the LTE system and other overheads. First, Fig. 5 shows the achievable total bit-rates of both BL and EL of Soccer video in MSs under the constraints of the target outage probability (e.g. 10%). When channel SNR for an MS is not strong enough to decode the BL, then, for a given ρ, the power dividing factor between BL and EL in eNB, the MS has the BL data only. However, the MS with good channel SNR condition, that is, near RS or eNB, can decode both the BL and additional EL. Note that there exist only a single BL rate and a single EL rate for the multicast of SVC throughout the cellular area. Therefore, the SNR in the horizontal axis is designated as the target SNR, at which the system sets the value for the edge of relay coverage. The total bit-rates in Fig. 5 can be interpreted as the rates that can be cast over the network with the given parameters (M_b=M_r,M_m, ρ) set and the target MSs SNR level, which we set to 10% outage of EL guaranteed. For example, a point at the target SNR = 6 dB denotes the sum of BL rate and EL rate, where the BL can be decoded throughout the cell area with 10% outage and EL can be decoded with 10% outage for MSs with SNR ≥ 6 dB. As we increase the target SNR, the total bit-rate increases while the number of MSs that can decode EL should decrease (the area where EL can be decoded is shrinking).

Fig. 5 shows that the total bit-rates of two modes are largely affected by the antenna settings: 4–2–2 case is for the antennas M_b=4, M_r=2, M_m=2 and 2–1–1 case is for M_b=2, M_r=1, M_m=1. On the other hand, the impact of ρ, the power dividing factor, is marginal. A large ρ uses more power for EL at the expense of BL power. Therefore a trade-off between BL and EL performance exists depending on ρ value. A smaller ρ (emphasising BL transmission) gives better data rates with an exception of 4–2–2 case of mode I, where larger ρ results in better data rates. Note that this trend diminishes as the target SNR level increases. Note also that mode I shows a superior performance in 4–2–2 case while mode II performs better in 2–1–1 case. Therefore mode I benefits more from additional antennas than mode II in data rate sense as explained in Section 3.1.

Secondly, the video qualities related to the total bit-rates in Fig. 5 are shown in Fig. 6. The PSNRs, as a typical metric for video quality, are obtained based on the achievable bit-rates of both BL and EL using a SVC decoder. Fig. 6a plots the PSNRs of Soccer sequence while Fig. 6b plots those of Mobile sequence. In Soccer, PSNR trends similar to that of the bit-rate results are observed. In Fig. 6, the interpretation of the horizontal axis should be the same as in Fig. 5, that is, the target SNR, at which the system is setting the value for the edge of relay coverage. As we increase the target SNR, the PSNR increases while the number of MSs that can decode EL should decrease (the area where EL can be decoded is shrinking).

The wide PSNR variation in the 2–1–1 case indicates that the BL data contribute much bigger influence on the resulting PSNR than the EL data, even though the total bit-rate gap is narrow between ρ values in 2–1–1 case of Fig. 5. The results of Mobile show that the PSNR performance of mode II is improved much when comparing M(2–1–1, 0.05) to M(2–1–1, 0.2). Since mode II is designed for videos with relatively mild movements, this result shows that the design objective is somehow achieved. If we choose ρ the values that give the best total data-rates for both modes (ρ = 0.2 for mode I and ρ = 0.05 for mode II), an eNB, depending on the video type and the antenna configuration, can adjust its transmission mode for better visual quality. From the observations of Fig. 6, we can summarise the mode selection given in Table 1.

5 Conclusion

We propose practical cooperative transmission schemes with performance analysis of both BER and capacity performance for scalable video over dedicated relay-based cellular networks. The cooperative transmissions with multiple antennas at the eNB and the dedicated relay utilise the combination of beamforming and space-time coding in order to enhance the multicast performance of scalable video. The proposed schemes ensure that the received BL of scalable video guarantees the minimum video quality to all MSs throughout the cell coverage area with a given outage probability. The MSs near the relay coverage area can achieve better video quality because they received both the BL and additional EL data, according to the received channel SNR configuration. Depending on the video types and antenna configuration, an eNB can choose between two transmission modes for better video quality at MSs.

6 Acknowledgment

This research was funded by the MSIP (Ministry of Science, ICT & Future Planning), Korea in the ICT R&D Program 2014.

7 References


