QoE-aware Cross-Layer Adaptation for D2D Video Communication in Cooperative Cognitive Radio Networks

Subhankar Chatterjee and Swades De

***Abstract*—In this work, a cross-layer parameter optimization is considered for H.264 scalable video transmission in cognitive radio-enabled device-to-device networks. Adaptive modulation and coding is considered at the physical layer along with layer aware automatic repeat request (VLA-ARQ) at the link layer. The VLA-ARQ scheme is aware of video characteristics. Scalable video coding (SVC) based video streaming is carried out at the application layer. A novel cognitive radio network frame structure is proposed accommodating cooperative spectrum sensing (CSS), relaying primary user's (PU) transmission, and secondary users' (SU) opportunistic transmission, to maximize quality of experience (QoE) and to minimize total power consumption. In SVC based video streaming, successful reception of a particular video layer is dependent upon the successful reception of its higher layers, which influences the eventual QoE at the user end. A set of optimal values of CSS time duration, sensing decision threshold, cooperation, and transmission power control are obtained under the constraints of detection reliability, PU spectrum efficiency, and SU power budget. Extensive simulation results with respect to the system parameters show the efficacy of the proposed system. The numerical results demonstrate that the proposed technique offers over 14% gain in overall utility with proposed CSS and cross-layer parameter optimization strategy.**

***Index Terms*—Cognitive radio, cooperative communication, cross-layer optimization, device-to-device (D2D) communication, quality of experience (QoE), scalable video coding (SVC), video layer aware automatic repeat request (VLA-ARQ).**

I. Introduction

Wide proliferation of wireless devices has triggered an increasing demand for mobile data traffic [1], [2]. However, the erroneous and bandwidth constrained nature of wireless links causes packet loss and delay, leading to quality degradation for real-time video transmission. Various techniques at different layers, such as, forward error correction (FEC), adaptive modulation and coding (AMC), and automatic repeat request (ARQ) are utilized to mitigate the effects of wireless channel.

Device-to-device (D2D) communication in licensed cellular bands also provides an alternate way to deal with this increasing mobile traffic scenario. D2D users operate within close proximity and exchange information directly with each other without the help of base stations. Thus, bandwidth intensive and delay sensitive applications, such as, video streaming in social networking [3] can be performed reliably between D2D pairs. In spite of the latest advancements in wireless networks, the existing licensed cellular communication suffers from spectrum under-utilization, leading to scarcity of radio frequency spectrum. Cognitive radio (CR) enabled D2D communication can combat this spectrum scarcity by reusing the licensed band of the primary users (PUs).

In order to provide reliable and high quality service to both cellular and D2D users in a heterogeneous environment, high quality of experience (QoE) must be ensured at the user end. On the other hand, the quality of service (QoS) metrics, namely delay and packet error rate focus only on the network side. Better QoS parameters can lead to better QoE, i.e., better user perception in video streaming. However, this is not true always. In case of multimedia applications, maximizing data rate does not always guarantee high QoE, as QoE can be influenced by several other parameters, such as video length, type of service, etc. [4]. Therefore, it is necessary to study the relationship between QoE and QoS, and focus on QoE-driven cross-layer design for cognitive radio network (CRN).

A. Related Works

We classify the related works section into three categories:

1) Key operations in CRN: The main challenge in CRN is to efficiently integrate reliable cooperative spectrum sensing (CSS) with opportunistic secondary user's (SU) transmission to highlight various objectives, primarily the impact of PU traffic on SU throughput [5] and SU throughput improvement considering reliable CSS [6]. Energy-efficient system design is explored in recent studies, namely, optimal resource (sensing time and transmission power) allocation in [7], [8], soft combining CSS in [9], dynamic spectrum sensing in [10], and considering the traffic behaviour of PU and SU nodes in [11].

To enhance spectrum efficiency, cooperative spectrum sharing [12]–[14] is considered where SU nodes act as relay nodes to meet the QoS requirements of PU. As a reward, the SU nodes are allowed to access the licensed frequency band. In [12], the transmission of PU was assisted by SU in a full-duplex cooperative non-orthogonal multiple access scheme. In [13], an optimal power allocation and spectrum leasing scheme was proposed for cognitive multi-hop relay network. An ARQ scheme was proposed in [14], where the SU node acts as relay for PU's transmission during the retransmission phase. A common feature in all these works is that, they did not consider the QoE aspect as performance optimization measure. Rather the focus has been at the physical layer in terms of rate guarantee and allowable interference margin, which do not necessarily translate to application layer QoE optimization.

This work was supported by the Science and Engineering Research Board, DST, under Grant CRG/2019/002293.

S. Chatterjee and S. De are with the Department of Electrical Engineering and Bharti School of Telecommunication, IIT Delhi, New Delhi 110016, India.

2) Video communication over cellular network and CRN: Recent works in the literature addressed video transmission in licensed cellular network. The work in [15] focused on QoE improvement of video transmission in wireless sensor networks. A relay selection technique for QoE enhancement was proposed in [16] for H.264 scalable video communication. Recent works also highlighted the aspect of video communication over CRN. The authors in [17] considered both rate adaptation and scalable video coding (SVC), for video streaming over CRN. A QoE-aware resource allocation technique was proposed in [18] for CRN. Instead of considering cellular network and CRN separately, cross-layer parameter adaptation for video streaming application was proposed in [19], for heterogeneous wireless network considering subcarrier assignment, power control, and user association. The authors in [20] addressed cross-layer QoE improvement problem for D2D users in CRN-enabled Heterogeneous network.

3) Cross-layer parameter adaptation for video communication: In order to meet fairness in throughput under time-varying wireless channel, AMC has been extensively studied at the physical layer [21]. An alternate way to mitigate the variations in wireless channel is to consider ARQ protocol at the link layer [22]. To further improve the network performance, recent works have focused on cross-layer parameter adaptation, considering both ARQ at the link layer and AMC at the physical layer [23]. This approach was further extended in [24] considering the QoS constraints. However, the parameters related to video quality were not discussed in [23], [24]. An adaptive truncated Hybrid ARQ scheme was discussed in [25] that makes use of the inter-layer encoding. Literature have also focused on application layer video coding and physical layer characteristics to improve the quality of video transmission over wireless media. The work in [26] discussed optimal power allocation for distortion minimization considering SVC video transmission. SVC based video streaming over multiple input multiple output (MIMO) system was studied in [27] using equal power allocation and space-time block coding. An adaptive channel selection was discussed in [28] where video layers with high priority are sent through spatial channels with high signal-to-noise ratio (SNR).

B. Motivation and Contributions

Recently, SVC based video communication has gained more popularity as it is capable of meeting the different preferences of end users and adapt to the changes in network conditions [29]. However, one of the main challenges of SVC based video streaming is that the usefulness of the received lower priority (i.e. higher enhancement) layer i depends on successful decoding of the lower layers j, $0 < j < i$. The decoding errors in higher priority layers will result in propagation of errors for lower priority enhancement video layers. Therefore, simple bit error rate (BER) minimization may not guarantee successful reception of video frames below the higher priority enhancement layers. A cross-layer parameter adaptation for SVC video streaming is thus necessary, considering physical layer characteristics (AMC), link layer protocol (ARQ), and application layer video coding schemes [30]. However, the work in [30] did not focus on optimal resource allocation for QoE maximization. Further, physical layer and application layer source encoding specific details were not addressed. The authors in [31] considered physical layer and application layer attributes for SVC based video streaming over MIMO networks. Although they considered QoE maximization using near-optimal power allocation technique, video layer aware link layer transmission was not considered.

Cross-layer resource allocation for CRN-enabled cellular network is necessary to combat the issue of spectrum scarcity and to meet QoE requirements in video transmission. QoE improvement in CRN-enabled heterogeneous network was investigated in [20], considering user association and resource block assignment. The authors did not take note of the impact of AMC and ARQ on the overall system reliability. Furthermore, they did not consider realistic SVC based video streaming application, where the video coding characteristics affect the overall QoE of the users. From the existing literature, we identify that QoE-driven cross-layer parameter optimization based on AMC in physical layer, VLA-ARQ in link layer, and SVC in application layer for D2D communication in cooperative CRN has not been investigated. To this end, the contributions of the paper are as follows:

1. The present work focuses on cross-layer parameter adaptation considering AMC at the physical layer, video layer aware automatic repeat request (VLA-ARQ) at the link layer, and SVC based video streaming at the application layer. Unlike the previous works in the literature, VLA-ARQ protocol considered here is unique. It is shown via simulation results that the proposed VLA-ARQ scheme offers almost same video quality compared to the conventional ARQ protocol. However, it is more energy efficient approach compared to conventional one.
2. The advantages of the proposed approach are explored for D2D communication in cooperative CRN. A unique framework is also proposed for CRN where CSS, application aware opportunistic secondary transmission, and cooperation in PU's transmission are addressed together.
3. The current work focuses on optimal resource allocation, i.e. layer-wise power allocation for SVC video and CSS parameter allocation, to maximize QoE and simultaneously minimize power consumption of the SUs, hence ensuring energy-efficient CRN design. Reliable CSS operation is ensured here through high probability of detection. Furthermore, spectrum efficiency of the PUs is maintained above a predefined threshold.
4. Extensive simulation studies are conducted to determine the significance of the proposed scheme. The results demonstrate improved QoE along with reduced power consumption compared to the other competitive schemes.

C. Paper Organization

The proposed system model is described in Section II. The analytical framework of the proposed system performance is presented in Section III. Section IV presents the problem formulation and the proposed solution. The performance of the proposed technique is evaluated via simulations in Section V, before concluding the paper in Section VI.

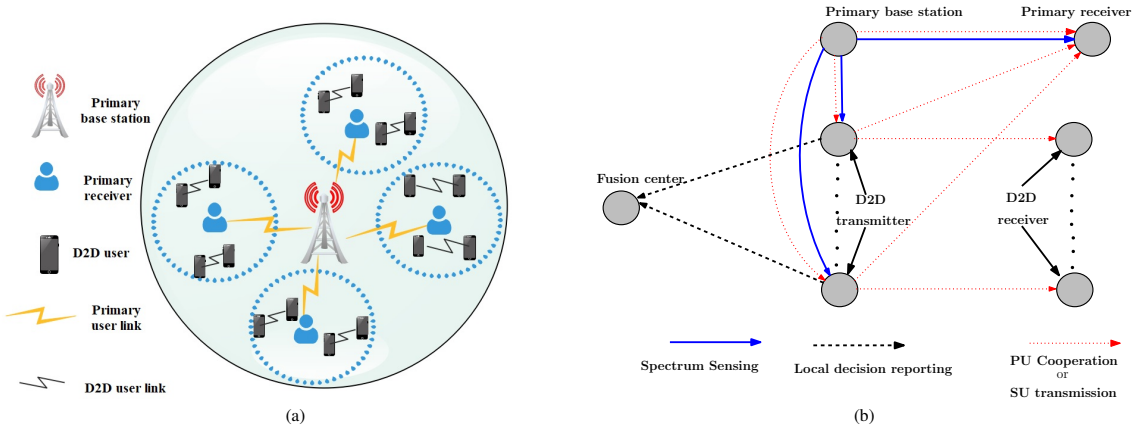


Fig. 1: (a) A heterogeneous cellular and D2D CRN scenario, and (b) cooperative cellular and D2D CRN system model.

II. System Model

We consider the downlink transmission in a D2D enabled cellular network. A heterogeneous cellular and D2D CRN scenario is shown in Fig. 1(a). The cellular network users (CNs) are considered as PUs and the D2D users are considered as SUs. PU network consists of one PU base station (PBS) and K number of primary receivers. SU network consists of $\mathcal{S}$ number of D2D transmitter-receiver pairs. Each node in both PU and SU network has single antenna. Let the total time frame of operation be T ms. During T time duration, PBS transmits data to K number of primary receivers using K dedicated orthogonal subchannels. Now, the D2D pairs are clustered into K groups (using K-means clustering) based on their path loss characteristics, in the close proximity of respective primary receiver nodes. Thus, the k^{th} group (at the vicinity of k^{th} primary receiver) consists of $\mathcal{S}^{\text{k}}$ number of D2D pairs and $\mathcal{S} = \{\mathcal{S}^1 \cup \mathcal{S}^2 \cup \cdots \mathcal{S}^{\text{K}}\}$.

The system model is further illustrated in Fig. 1(b), focusing on the operations of k^{th} group of D2D transmitter-receiver pairs. The time frame T is elaborated in Fig. 2, consisting of CSS time τ_s, reporting time for each D2D transmitter τ_r, and cooperation or D2D transmitters' transmission time $T - \tau_s - S\tau_r$. During τ_s, the $\mathcal{S}^{\text{k}}$ number of D2D transmitter nodes perform CSS to detect the activity of PBS in the k^{th} subchannel. If PBS is found to be in transmit mode, D2D transmitters act as amplify-and-forward (AF) relays and cooperate in the downlink transmission from PBS to k^{th} primary receiver (using k^{th} subchannel). When PBS is in active mode, the D2D transmitter nodes do not have access to the licensed spectrum, and therefore cannot perform own transmissions. When the PBS is in idle mode, as a reward for their cooperation the D2D nodes can use the channel and communicate their data to their respective receivers using time division multiple access (TDMA). *This integrated framework of CSS, D2D transmission, and cooperation in PU's transmission is novel, which has not been addressed in literature.*

Since only one primary receiver node is receiving data using each orthogonal subchannel, interference does not come into play here. Furthermore, D2D network analysis is identical for any primary receiver node. As a result, we focus on the

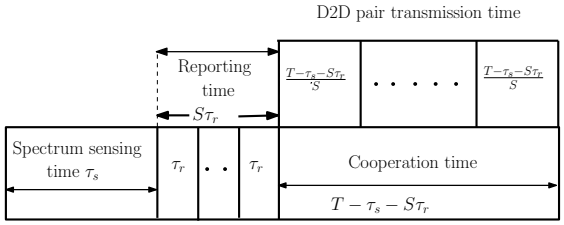


Fig. 2: Frame structure.

mathematical calculations considering a specific k^{th} primary receiver. For brevity, in the rest of this paper, $\mathcal{S}^{\text{k}}$ is denoted by S. Furthermore we assume primary receiver to be either in idle mode or receiving information for the entire time frame T. All the wireless channels are modelled as Rayleigh block fading channels. The channel fading coefficient of the wireless link $i - j$ is denoted by h_{ij} where $h_{ij} \sim \mathcal{CN}(0, d_{ij}^{-\beta})$. The path loss exponent is denoted by β and the distance between transmitter i and receiver j is denoted by d_{ij}. A list of symbols and their descriptions is provided in Table I.

III. Analytical Framework of the Proposed System Performance

In this section, we describe the relevant operations related to the proposed framework. The detailed analysis on CSS and cooperation in PBS transmission are first conducted, followed by discussion on SVC based video quality model. QoE parameter quantification considering AMC and VLA-ARQ is provided next. This section ends with the power consumption model for D2D pair.

A. Cooperative Spectrum Sensing Analysis

During τ_s, the D2D transmitter nodes perform spectrum sensing to identify the active/ idle state of PBS. The received signal at any m^{th} D2D transmitter node is

$$y_{pm} = \phi\sqrt{P_T}h_{pm}x_p + w_m. \quad (1)$$

Here ϕ acts as a binary indicator, $\phi = 1$ or 0 indicates PBS to be in transmit mode or in idle mode, respectively. x_p is

TABLE I: List of symbols used

S	Number of D2D transmitter-receiver pairs
K	Number of PU receivers
T	Time frame duration
τ_s	Spectrum sensing time
τ_r	Reporting time
β	Path loss exponent
h_{ij}	Channel fading coefficient of wireless link $i-j$
d_{pr}	Distance between PBS and primary receiver
d_{mm}	Distance between m^{th} D2D transmitter-receiver pair
d_{pm}	Distance between PBS and m^{th} D2D transmitter
d_{mr}	Distance between m^{th} D2D transmitter and primary receiver
ϕ	Binary indicator
P_T	PBS transmission power
P_n	Noise variance
λ_m	Local decision threshold
f_s	Sampling frequency
$\hat{\gamma}$	Average SNR at D2D transmitter node during spectrum sensing
P_{cm}	Transmission power of m^{th} D2D transmitter during cooperation
L	Number of layers of SVC video stream
q	Quantization step size
q^{min}	Minimum quantization step size
QP	Quantization parameter
t	Video frame rate
t^{max}	Maximum frame rate
$N_{k,m}$	Number of retransmissions for k^{th} layer of m^{th} D2D transmitter
z_k	Average size of layer k
F	Frame rate of video stream
$\hat{E}_k$	Differential rate of layer k
P_{ac}	The probability that PBS is in active mode
P_d	Cooperative probability of detection
$\hat{P}_d$	Target probability of detection
P_f	Cooperative probability of false alarm
$p_b^{k,m}$	Bit error rate probability for k^{th} layer of m^{th} D2D pair
$p_{fr}^{l,m}$	Probability that layer l is decoded correctly at m^{th} D2D receiver
P_{pp}^m	Processing power of m^{th} D2D transmitter
P_{DT}^m	Transmission power consumption of m^{th} D2D transmitter
P_S	Power consumption in sensing phase
P_R	Power consumption in reporting phase

PBS transmitted signal with zero mean and unit variance, and P_T denotes PBS transmission power. The noise w_m at m^{th} D2D transmitter is considered to be circularly symmetric complex Gaussian (CSCG) random variable with zero mean and variance $E\left[|w_m|^2\right]$=P_n.

Each D2D transmitter node performs energy detection using the received PBS signal. The local probability of detection p_d^m and the local probability of false alarm p_f^m at m^{th} D2D transmitter are respectively given by [6]:

$$p_d^m = Q\left(\left(\frac{\lambda_m}{P_n} - \hat{\gamma}_m - 1\right)\frac{\sqrt{f_s.\tau_s}}{\hat{\gamma}_m + 1}\right) \quad (2)$$

$$p_f^m = Q\left(\left(\frac{\lambda_m}{P_n} - 1\right)\sqrt{f_s.\tau_s}\right). \quad (3)$$

Here $Q(x) = \frac{1}{\sqrt{2\Pi}}\int_x^\infty \exp(-\frac{y^2}{2})dy$. The average signal-to-noise ratio (SNR) at m^{th} D2D transmitter is denoted by $\hat{\gamma}_m = \frac{d_{pm}^{-\beta}P_T}{P_n}$, and λ_m is the local decision threshold. The sampling frequency is denoted by f_s. Since the D2D transmitter nodes are located in a cluster, the distances between D2D transmitters to primary receiver and PBS are much larger than the distances among D2D nodes. Therefore, we consider $d_{mr} = d_r$, $d_{pm} = d_p$, $\hat{\gamma}_m = \hat{\gamma}$, $\lambda_m = \lambda$, $p_d^m = p_d$, and $p_f^m = p_f$; $\forall m = 1$ to S.

The local sensing results are forwarded sequentially to fusion center during the reporting phase using a common control channel. In this work, the control channel is considered to be error free. The reporting time for any particular D2D transmitter node is τ_r. These local binary decisions are decoded at fusion center and then combined using counting rule. Fusion center determines the global decision about the activity of PBS and broadcasts this decision to the D2D transmitter nodes. The probability of false alarm P_f and the probability of detection P_d at fusion center can be calculated as [6]:

$$P_d = \sum_{j=S_1}^{S}\binom{S}{j}p_d^j(1-p_d)^{S-j} = I_{p_d}(S_1, S-S_1+1) \quad (4)$$

$$P_f = \sum_{j=S_1}^{S}\binom{S}{j}p_f^j(1-p_f)^{S-j} = I_{p_f}(S_1, S-S_1+1). \quad (5)$$

Here, $I_y(a,b)$ is the regularized incomplete beta function, and S_1 denotes the final decision threshold. Let $P_d = \hat{P}_d$. Accordingly, local p_d can be written as follows:

$$p_d = I^{-1}(\hat{P}_d, S_1, S-S_1+1). \quad (6)$$

If $x = I_y(a,b)$, then $y = I^{-1}(x,a,b)$. Here $I^{-1}(x,a,b)$ is the inverse regularized incomplete beta function. The local p_f is determined using (2) and (3):

$$p_f = Q\left(Q^{-1}\left(p_d\right)\left(\hat{\gamma}+1\right) + \hat{\gamma}\sqrt{f_s\tau_s}\right). \quad (7)$$

The overall probability of false alarm P_f is then determined using (5).

B. Cooperation in PBS Transmission

If fusion center found PBS to be active, D2D transmitter nodes act as AF relays and cooperate in PU transmission in two parts. Primary receiver directly receives signal from PBS during the first part, which can be expressed as:

$$y_{pr} = \sqrt{P_T}h_{pr}x_p + w_{r1}. \quad (8)$$

At the same time, D2D transmitter nodes also receive signal from PBS (please see (1)). D2D transmitter nodes amplify the received signal from PBS and then forward it to primary receiver. Therefore, the signal received by primary receiver is:

$$y_{mr} = \sum_{m=1}^{S}\sqrt{\alpha_{cm}}h_{mr}y_{pm} + w_{r2}. \quad (9)$$

The noise signals w_{r1} (for direct link) and w_{r2} (for relay link) are CSCG random variables with zero mean and variance P_n. α_{cm} denotes the amplifying power gain of m^{th} D2D transmitter during cooperation.

Now, average spectral efficiency of PU is made of two terms, R_{pu}^1 and R_{pu}^2. In case of correct detection, spectrum efficiency is denoted by R_{pu}^1, and R_{pu}^2 indicates spectrum efficiency in case of missed detection. In case of missed detection, D2D transmitter nodes transmit data to its respective D2D receiver nodes, resulting in collision with PBS transmission and subsequently reduction in spectrum efficiency, i.e. $R_{pu}^1 \gg R_{pu}^2$. Therefore, the overall spectrum efficiency is $R_p \sim R_{pu}^1$.

$$R_{pu}^1 = P_{ac}P_d\frac{T-\tau_s-S\tau_r}{2T}\left(\log_2\left(1+\frac{|h_{pr}|^2P_T}{P_n}\right) + \log_2\left(1+\frac{\sum_{m=1}^{S}\alpha_{cm}|h_{mr}|^2|h_{pm}|^2P_T}{\sum_{m=1}^{S}\alpha_{cm}|h_{mr}|^2P_n+P_n}\right)\right). \quad (10)$$

Here P_{ac} represents the probability that PBS is in active mode. The factor $\frac{1}{2}$ accounts for the usage of two equal slots for cooperation.

C. Scalable Coding Based Video Model for SU Transmission

D2D transmitter nodes sequentially send information to the respective D2D receiver nodes using TDMA when PBS is detected idle. The transmission duration for m^{th} D2D transmitter is $\frac{T-\tau_s-S\tau_r}{S}$. In this work, video transmission is considered between D2D pairs. Video application is more delay sensitive and bandwidth intensive compared to any other applications, e.g. audio and web browsing.

In this paper, H.264 based SVC is considered[1]. Here, the users can dynamically adjust the received video quality based on their individual requirements. The original video stream is encoded into L layers of sub-streams, i.e. one base layer (layer 1) and $(L-1)$ enhancement layers (layer 2 to layer L). The basic quality of a video sequence is provided by base layer. The base layer is decoded independently of other enhancement layers. Now, the enhancement layer l_i has higher priority compared to layer l_j if $i < j$. The layer l_j is correctly decoded at the receiver only when all layers l_i, $i \leq j$ are correctly decoded. H.264 SVC encoder supports three types of scalabilities- temporal, spatial, and quality scalabilities. The temporal scalability is based on variable frame transmission rate, spatial scalability is linked with the display resolution of D2D receiver, and quality scalability is governed by encoding rate of SVC. Quality scalable video transmission is addressed in this work.

The quality parametric model described in [32] is utilized here. The quality Q of scalable video stream for a specific spatial resolution is given by:

$$Q = Q^{max}\ Q(t)\ Q(q) \quad \text{where}$$
$$Q(t) = \frac{1-e^{-a\frac{t}{t^{max}}}}{1-e^{-a}},\ Q(q) = \frac{e^{-b\frac{q}{q^{min}}}}{e^{-b}},\ \text{and}\ q = 2^{\frac{QP-4}{6}}. \tag{11}$$

The quantization parameter is specified by QP. q and t denote quantization step size and frame rate, respectively. The symbols a and b are video-specific parameters. The maximum quality of video sequence is Q^{max} when $t = t^{max}$, i.e., the maximum frame rate, and $q = q^{min}$, i.e., the minimum quantization step size. Q_{max} is set at 100% for normalization purpose. Since, only quality scalability is considered here, the frame rate of l^{th} layer $t_l = t = t^{max}$, and $Q(t) = 1, \forall l = 1$ to L. Accordingly, the quality for l^{th} layer is represented here as:

$$Q_l = \frac{e^{-b\frac{q_l}{q^{min}}}}{e^{-b}}. \tag{12}$$

Here, q_l is the quantization step size for l^{th} layer.

[1]The overall analysis is applicable to any layered or scalable video coding formats. For other video coding techniques, such as multi-description coded video, the video reception performance analysis would have to be modified.

TABLE II: Coefficients for different AMC [31]

Type (k)	Modulation	Code Rate	c_k	d_k
1	BPSK	1/2	1.1369	7.5556
2	QPSK	1/2	0.3351	3.2543
3	QPSK	3/4	0.2197	1.5244
4	16 QAM	9/16	0.2081	0.6250
5	16 QAM	3/4	0.1936	0.3484
6	64 QAM	3/4	0.1887	0.0871

D. QoE Parameter Quantification Considering AMC and VLA-ARQ

In this paper, we consider the following BER expression for AMC [31]:

$$p_b^{k,m} = c_{k,m}e^{-d_{k,m}\gamma_{k,m}}. \tag{13}$$

Here, $p_b^{k,m}$ is the BER (probability) for k^{th} layer of m^{th} D2D pair (considering AMC only). The values of the coefficients $c_{k,m}$ and $d_{k,m}$ depend upon the AMC of layer k. Table II shows the list of different coefficients and their corresponding AMCs. $\gamma_{k,m}$ is the instantaneous SNR for transmission of k^{th} layer between m^{th} D2D pair, i.e. $\gamma_{k,m} = \frac{|h_{mm}^k|^2 P_{km}}{P_n}$. The L layers of scalable video are sent over L orthogonal subcarriers. h_{mm}^k is the channel fading coefficient of the k^{th} subcarrier. The transmission power consumption for layer k is P_{km}.

We assume that perfect statistical and instantaneous channel state information (CSI) are available at the receiver. Since the base layer, i.e. layer 1 is of utmost importance, it is sent using the subcarrier with the best instantaneous SNR value. The remaining layers are matched to other subcarriers based on their priorities. For example, if $i < j$, priority of layer $i >$ priority of layer j, and instantaneous SNR of the subcarrier matched to layer $i >$ instantaneous SNR of the subcarrier matched to layer j.

In this work, VLA-ARQ is considered along with AMC. Unlike previous works in literature that deals with conventional ARQ, here we propose ARQ scheme only for the base layer, since it is of utmost importance. The same modulation mode is utilized for all the retransmissions. m^{th} D2D receiver estimates the modulation mode for transmission and the required number of retransmissions for base layer. These information are fed back to m^{th} D2D transmitter.

The time duration between different transmissions of base layer are assumed to be larger than the channel coherence time. Therefore, the fading coefficients corresponding to the multiple transmissions are independent and identically distributed (i.i.d.). The initial and the retransmitted packets of base layer ($k = 1$) are combined at m^{th} D2D receiver using maximum ratio combining (MRC). Let $\gamma_{k,m}^r$ be the instantaneous SNR in case of r^{th} retransmission and $N_{k,m}$ be the total number of retransmissions for base layer. The BER expression (for base layer) considering both AMC and ARQ is expressed as:

$$p_b^{k,m} = c_{k,m}e^{-d_{k,m}\left(\gamma_{k,m}+\sum_{r=1}^{N_{k,m}}\gamma_{k,m}^r\right)}, \quad \text{when } k = 1. \tag{14}$$

The average BER is determined at m^{th} D2D receiver with respect to $N_{k,m}$ different retransmissions. For brevity, we

consider $N_{k,m} = N$ for the subsequent analysis.

$$\begin{aligned}\hat{p}_b^{k,m} &= \int_0^\infty \cdots \int_0^\infty p_b^{k,m} \quad p_{\gamma_{k,m}^1,\cdots,\gamma_{k,m}^N}(\gamma_{k,m}^1,\cdots,\gamma_{k,m}^N) \\ &\quad \times d\gamma_{k,m}^1,\cdots,d\gamma_{k,m}^N \\ &= c_{k,m} e^{-d_{k,m}\gamma_{k,m}} \prod_{r=1}^{N} \int_0^\infty e^{-d_{k,m}\gamma_{k,m}^r} \quad p_{\gamma_{k,m}^r}(\gamma_{k,m}^r) \\ &\quad \times d\gamma_{k,m}^r, \quad \text{when} \quad k=1. \end{aligned} \tag{15}$$

Since Rayleigh fading is considered, $p_{\gamma_{k,m}^r}(\gamma_{k,m}^r) = \frac{1}{\hat{\gamma}_{k,m}} e^{-\frac{\gamma_{k,m}^r}{\hat{\gamma}_{k,m}}}$, the average SNR is denoted by $\hat{\gamma}_{k,m}$. Accordingly,

$$\begin{aligned}\hat{p}_b^{k,m} &= c_{k,m} e^{-d_{k,m}\gamma_{k,m}} \prod_{r=1}^{N} \int_0^\infty e^{-d_{k,m}\gamma_{k,m}^r} \\ &\quad \times \frac{1}{\hat{\gamma}_{k,m}} e^{-\frac{\gamma_{k,m}^r}{\hat{\gamma}_{k,m}}} \quad d\gamma_{k,m}^r \\ &= c_{k,m} e^{-d_{k,m}\gamma_{k,m}} (1 + d_{k,m}\hat{\gamma}_{k,m})^{-N}, \quad \text{when} \quad k=1. \end{aligned} \tag{16}$$

The average BER $\hat{p}_b^{k,m}$ should always be below a predefined threshold $\hat{p}_b^{th}$, i.e. $\hat{p}_b^{k,m} \leq \hat{p}_b^{th}$. That is,

$$\begin{aligned} & c_{k,m} e^{-d_{k,m}\gamma_{k,m}} (1 + d_{k,m}\hat{\gamma}_{k,m})^{-N} \leq \hat{p}_b^{th} \\ & \text{or, } N \geq \frac{d_{k,m}}{\ln(1 + d_{k,m}\hat{\gamma}_{k,m})} \left[\frac{1}{d_{k,m}} \ln\left(\frac{c_{k,m}}{\hat{p}_b^{th}} \right) - \gamma_{k,m} \right] \triangleq N^*. \end{aligned} \tag{17}$$

When $\gamma_{k,m} < \frac{1}{d_{k,m}} ln\left(\frac{c_{k,m}}{\hat{p}_b^{th}}\right)$, $N > 0$. In this case retransmission will be required for base layer. In this work, we set $\hat{p}_b^{k,m} = \hat{p}_b^{th}$. Therefore (17) is set to equality. So, $N \triangleq \lceil N^* \rceil$, where $\lceil a \rceil$ denotes the nearest integer higher than a.

As mentioned in (12), the quality of layer l is denoted by Q_l, without taking into account the impairments of wireless channel. Therefore, the received video quality in presence of fading channel is $Q_l \cdot p_{fr}^{l,m}$. Here $p_{fr}^{l,m}$ denotes the probability that layer l is decoded correctly at m^{th} D2D receiver. As mentioned in Section III-C, the layer l is correctly received only when all the layers $k \leq l$ are correctly received. Hence,

$$p_{fr}^{l,m} = \prod_{k=1}^{l} (1 - p_b^{k,m})^{z_k}. \tag{18}$$

$$\text{Here,} \quad p_b^{k,m} = c_{k,m} e^{-d_{k,m}\gamma_{k,m}} = c_{k,m} e^{-d_{k,m} P_{km} \rho_{k,m}}$$
$$\text{when} \quad k > 1 \text{ and } \rho_{k,m} = \frac{|h_{mm}^k|^2}{P_n}.$$

$$\begin{aligned} \text{and} \quad p_b^{k,m} &= c_{k,m} e^{-d_{k,m}\left(\gamma_{k,m} + \sum_{r=1}^{N} \gamma_{k,m}^r\right)} \\ &= c_{k,m} e^{-d_{k,m} P_{km} \rho_{k,m}}, \end{aligned}$$
$$\text{when} \quad k = 1 \text{ and } \rho_{k,m} = \frac{|h_{mm}^k|^2}{P_n} + \sum_{r=1}^{N} \frac{|h_{mm}^{k,r}|^2}{P_n}.$$

Here, z_k denotes the average layer size, i.e. the average number of bits in layer k. Let the frame rate of a particular video stream be F fps. Now, z_k can be represented as:

$$z_k = \frac{\hat{E}_k}{F}, \quad \text{where } k = \{1, 2, \cdots L\} \tag{19}$$

where $\hat{E}_k$ is the differential rate of layer k [33], which is determined as follows:

$$\hat{E}_k = \begin{cases} \mathrm{E}_k, & k = 1 \\ \mathrm{E}_k - \mathrm{E}_{k-1}, & k > 1 \end{cases} \tag{20}$$

Now, E_k denotes the cumulative rate upto layer k. Then, the total quality at m^{th} D2D receiver is

$$Q^m = \sum_{l=1}^{L} (1 - P_{ac})(1 - P_f) \frac{T - \tau_s - S\tau_r}{ST} Q_l \cdot p_{fr}^{l,m}. \tag{21}$$

The term $(1 - P_{ac})(1 - P_f)$ denotes the spectrum access probability of m^{th} D2D transmitter.

E. Power Consumption at D2D Pair

The average power consumption of m^{th} D2D transmitter consists of two parts, processing power P_{PP}^m and transmission power consumption P_{DT}^m. Following [34], processing power is calculated as:

$$P_{PP}^m \approx P_{filter} + P_{DAC} + P_{mix} + P_{syn}. \tag{22}$$

The active filter power consumption P_{filter}, digital-to-analog converter power consumption P_{DAC}, mixer power consumption P_{mix}, and synthesizer power consumption P_{syn} are all considered here to be constants. Now, the transmission power consumption:

$$\begin{aligned} P_{DT}^m &= \frac{\tau_s P_S + \tau_r P_R}{T} + P_{cm} \frac{T - \tau_s - S\tau_r}{2T} + ((1 - P_{ac}) \\ &\quad \times (1 - P_f) + P_{ac}(1 - P_d)) \frac{T - \tau_s - S\tau_r}{ST} \sum_{l=1}^{L} P_{lm}. \end{aligned} \tag{23}$$

Here P_R and P_S denote the power consumption in reporting phase and sensing phase, respectively. The power consumption during cooperation is $P_{cm} = ((1 - P_{ac})P_f P_n + P_{ac}P_d(|h_{pm}|^2 P_T + P_n))\alpha_{cm}$. The power consumption at m^{th} D2D transmitter for transmission of layer l is P_{lm}. We rewrite (23) as follows:

$$\begin{aligned} P_{DT}^m &= \sum_{l=1}^{L} \Bigg(\frac{\tau_s P_S + \tau_r P_R}{LT} + P_{cm} \frac{T - \tau_s - S\tau_r}{2LT} \\ &\quad + \Big((1 - P_{ac})(1 - P_f) + P_{ac}(1 - P_d) \Big) \frac{T - \tau_s - S\tau_r}{ST} P_{lm} \Bigg) \\ &= \sum_{l=1}^{L} P_{DT}^{lm}. \end{aligned} \tag{24}$$

The power consumption at m^{th} D2D receiver is [34]:

$$P_{DR}^m \approx P_{filter} + P_{ADC} + P_{mix} + P_{syn} + P_{IFA} + P_{LNA}. \tag{25}$$

Here, P_{ADC}, P_{IFA}, and P_{LNA} are analog-to-digital converter power consumption, inverted-F-antenna power consumption,

and low noise amplifier power consumption, respectively. Hence, total power consumption at m^{th} D2D pair is:

$$P_{DD}^{m} = \sum_{l=1}^{L} \eta^{l} \left(\frac{P_{DR}^{m}}{L} + P_{DT}^{lm} + \frac{P_{PP}^{m}}{L} \right) = \sum_{l=1}^{L} P_{DD}^{lm}. \quad (26)$$

The symbol $\eta^l = N+1$ when $l=1$ (i.e. base layer), where N indicates the number of retransmissions for base layer and $\eta^l = 1$ when $l > 1$ (i.e. enhancement layers).

Now, $p_b^{k,m}$ is a function of P_{km}; as P_{km} increases $p_b^{k,m}$ decreases. On the other hand, P_{km}, i.e., P_{DD}^{m} needs to be minimized from energy efficient system design purpose. Hence, there lies a trade-off. To highlight this trade-off, we define an utility function for the m^{th} D2D pair:

$$U_m = \sum_{l=1}^{L} (1-P_{ac})(1-P_f) \frac{T-\tau_s - S\tau_r}{ST} \frac{Q_l \cdot p_{fr}^{l,m}}{P_{DD}^{lm}}. \quad (27)$$

IV. Problem Formulation and Proposed Solution

This section addresses the multiuser resource allocation for sum utility maximization with the problem formulation, followed by the proposed solution.

A. Problem Formulation

The objective is to jointly determine the optimal sensing time (τ_s), the final decision threshold S_1, the cooperation power (P_{cm}), and the transmission power consumption for layer l (P_{lm}), $\forall l = \{1, \cdots, L\}$ and $m = \{1, \cdots, S\}$, to maximize the sum utility of all D2D pairs. The sum utility of S number of D2D pairs is denoted by $U = \sum_{m=1}^{S} U_m$. Here, the value of U_m depends on the values of Q_l, $p_{fr}^{l,m}$, and P_{DD}^{lm}. Q_l is a function of the quantization step size q_l. As shown in [35], the video quality is a concave function with respect to the quantization parameter QP. Determination of optimal value of Q_l is not the focus of this work. Here, the values of Q_l are calculated for specific q_l values, $\forall l = \{1, \cdots, L\}$. In this work, we mainly focus on optimizing $p_{fr}^{l,m}$ and P_{DD}^{lm} for L layers of scalable video transmission. Thus, the optimization problem is written as follows:

$$\max_{\tau_s, S_1, P_{cm}, P_{lm}} \sum_{m=1}^{S} \sum_{l=1}^{L} (1-P_{ac})(1-P_f) \frac{T-\tau_s - S\tau_r}{ST} \frac{Q_l \cdot p_{fr}^{l,m}}{P_{DD}^{lm}} \quad (28)$$

where

$$Q_l = \frac{e^{-b\frac{q_l}{q^{min}}}}{e^{-b}}, \; p_{fr}^{l,m} = \prod_{k=1}^{l} (1-p_b^{k,m})^{z_k},$$

$$P_{DD}^{lm} = \eta^{l} \left(\frac{P_{DR}^{m}}{L} + P_{DT}^{lm} + \frac{P_{PP}^{m}}{L} \right),$$

$$\text{and } P_{DT}^{lm} = \Big(\frac{\tau_s P_S + \tau_r P_R}{LT} + P_{cm} \frac{T-\tau_s - S\tau_r}{2LT} + \Big((1-P_{ac})(1-P_f) + P_{ac}(1-P_d) \Big) \frac{T-\tau_s - S\tau_r}{ST} P_{lm} \Big)$$

subject to

$$\text{C1: } 0 < \tau_s < T - S\tau_r, \quad (29)$$

$$\text{C2: } P_d \geq \hat{P}_d \quad (30)$$

$$\text{C3: } 1 \leq S_1 \leq S \quad (31)$$

$$\text{C4: } P_{lm} \geq 0, \quad \forall l = \{1, \cdots, L\}, \quad \forall m = \{1, \cdots, S\} \quad (32)$$

$$\text{C5: } \sum_{l=1}^{L} \eta^{l} \left(\frac{P_{DR}^{m}}{L} + P_{DT}^{lm} + \frac{P_{PP}^{m}}{L} \right) \leq P_{th}, \quad \forall m = \{1, \cdots, S\} \quad (33)$$

$$\text{C6: } P_{ac} P_d \frac{T-\tau_s - S\tau_r}{2T} \left(\log_2 \left(1 + \frac{|h_{pr}|^2 P_T}{P_n} \right) + \log_2 \left(1 + \frac{\sum_{m=1}^{S} \alpha_{cm} |h_{mr}|^2 |h_{pm}|^2 P_T}{\sum_{m=1}^{S} \alpha_{cm} |h_{mr}|^2 P_n + P_n} \right) \right) \geq R_p^{th}. \quad (34)$$

The optimization problem jointly maximizes the QoE of D2D users and minimizes the power consumption while ensuring the CSS duration and detection reliability constraints (see C1 and C2, respectively). C3 highlights the final decision threshold constraint for CSS operation. To ensure fairness, the average power consumption of m^{th} D2D pair should be within a predefined domain $[0 \;\; P_{th}]$ (see C4 and C5). Finally, a minimum spectrum efficiency requirement (R_p^{th}) is set for each primary receiver node (see C6).

B. Proposed Solution

The solution to the proposed optimization problem is elaborated here. First we derive the optimal cooperation power P_{cm} for individual D2D transmitter node. Then primal decomposition method is utilized to decompose the original optimization problem into two sub-problems in order to determine the optimal values of τ_s, S_1 and P_{lm}.

The objective function (28) reaches maximum value when P_D is set at the lower limit $\hat{P}_d$. Therefore, $P_d = \hat{P}_d$ is considered in the modified optimization problem.

Next, the following expression is obtained from (34):

$$\frac{\sum_{m=1}^{S} \gamma_{pm} \gamma_{mr} / \delta_{1m}}{\sum_{m=1}^{S} \gamma_{mr} / \delta_{1m} + \frac{1}{P_n}} \geq \frac{2^{\frac{2TR_p^{th}}{P_{ac} P_d (T-\tau_s - S\tau_r)}}}{1 + \frac{|h_{pr}|^2 P_T}{P_n}} - 1 \quad (35)$$

Here, $\gamma_{pm} = \frac{|h_{pm}|^2 P_T}{P_n}$, $\gamma_{mr} = \frac{|h_{mr}|^2 P_{cm}}{P_n}$, and $\delta_{1m} = ((1-P_{ac}) P_f P_n + P_{ac} P_d (|h_{pm}|^2 P_T + P_n))$. Let $\delta_2 = \frac{2^{\frac{2TR_p^{th}}{P_{ac} P_d (T-\tau_s - S\tau_r)}}}{1 + \frac{|h_{pr}|^2 P_T}{P_n}} - 1$.

In this work, each D2D transmitter is given equal transmission opportunity. Therefore, D2D transmitter nodes utilize an 'equal paying with SNR' method [36] for cooperation purpose. All D2D transmitter nodes must have the average cooperation

power to satisfy $\gamma_{ir} = \gamma_{jr},\ \forall i,j = \{1,\cdots,S\}$ and $i \neq j$. From (35), we get

$$\alpha_{cm} \geq \frac{1}{\delta_{1m}|h_{mr}|^2} \frac{\delta_2}{\sum_{m=1}^{S} \frac{\gamma_{pm}}{\delta_{1m}} - \delta_2 \sum_{m=1}^{S} \frac{1}{\delta_{1m}}} \tag{36}$$

The denominator of the objective function in (28) consists of P_{DD}^{lm}, which is a function of α_{cm}. The objective function is maximized when P_{DD}^{lm} is minimized, i.e. α_{cm} in (36) is set at equality. Thus,

$$\alpha_{cm} = \frac{1}{\delta_{1m}|h_{mr}|^2} \frac{\delta_2}{\sum_{m=1}^{S} \frac{\gamma_{pm}}{\delta_{1m}} - \delta_2 \sum_{m=1}^{S} \frac{1}{\delta_{1m}}} \tag{37}$$

We can represent (33) as follows:

$$\text{C7: } \sum_{l=1}^{L} \left(\varrho_1^l P_{lm} + \varrho_2^l \right) \leq P_{th}, \forall m = \{1,\cdots,S\} \tag{38}$$

where

$$\varrho_1^l = \eta^l \left((1-P_{ac})(1-P_f) + P_{ac}(1-P_d) \right) \frac{T - \tau_s - S\tau_r}{ST}$$

$$\varrho_2^l = \eta^l \left(\frac{P_{DR}^m}{L} + \frac{P_{PP}^m}{L} + \frac{\tau_s P_S + \tau_r P_R}{LT} + P_{cm} \frac{T - \tau_s - S\tau_r}{2LT} \right).$$

Following primal decomposition method [37], the modified optimization problem is broken into two sub-problems, P^1 and P^2.

$$\text{P}^1: \quad \max_{\{P_{lm}\}_{l=1}^{L}} \sum_{l=1}^{L} (1-P_{ac})(1-P_f) \frac{T - \tau_s - S\tau_r}{ST} \frac{Q_l \cdot p_{fr}^{l,m}}{P_{DD}^{lm}} \tag{39}$$

subject to

$$\text{C4: } P_{lm} \geq 0, \quad \forall l = \{1,\cdots,L\}, \quad \forall m = \{1,\cdots,S\} \tag{40}$$

$$\text{C7: } \sum_{l=1}^{L} \left(\varrho_1^l P_{lm} + \varrho_2^l \right) \leq P_{th}, \forall m = \{1,\cdots,S\} \tag{41}$$

and

$$\text{P}^2: \quad \max_{\tau_s, S_1} \sum_{m=1}^{S} \sum_{l=1}^{L} (1-P_{ac})(1-P_f) \frac{T - \tau_s - S\tau_r}{ST} \frac{Q_l \cdot p_{fr}^{l,m}}{P_{DD}^{lm}} \tag{42}$$

subject to

$$\text{C1: } 0 < \tau_s < T - S\tau_r \tag{43}$$

$$\text{C3: } 1 \leq S_1 \leq S. \tag{44}$$

The optimal value of τ_s is determined from P^2. Let

$$O = \sum_{m=1}^{S} \sum_{l=1}^{L} (1-P_{ac})(1-P_f) \frac{T - \tau_s - S\tau_r}{ST} \frac{Q_l \cdot p_{fr}^{l,m}}{P_{DD}^{lm}}.$$

The partial derivative of O with respect to τ_s is $\frac{\partial O}{\partial \tau_s}$. The equation $\frac{\partial O}{\partial \tau_s} = 0$ is non-linear, thus mathematically intractable. Now, $\lim_{\tau_s \to 0} \frac{\partial O}{\partial \tau_s} = \infty$ and $\lim_{\tau_s \to T - S\tau_r} \frac{\partial O}{\partial \tau_s} < 0$. So, there exists a unique stationary point which maximizes O. Hence, the optimal value of τ_s is determined following bisection search. Now, S is an integer; therefore the optimal value of S_1 can be determined using exhaustive search within $1 \leq S_1 \leq S$.

The optimal value of $P_{lm},\ \forall l = \{1,\cdots,L\}$ is determined using P^1. The problem in P^1 is further divided into L sub-problems $\text{P}^{1,l},\ \forall l = \{1,\cdots,L\}$. The sub-problem $\text{P}^{1,l}$ focuses on optimizing $\frac{p_{fr}^{l,m}}{P_{DD}^{lm}}$, as shown below:

$$\text{P}^{1,l}: \max_{\{P_{km}\}_{k=1}^{l}} \log \left(\frac{p_{fr}^{l,m}}{P_{DD}^{lm}} \right) = \sum_{k=1}^{l} z_k \log(1 - p_b^{k,m}) - \log \left(P_{DD}^{lm} \right) \tag{45}$$

subject to

$$\text{C4: } P_{km} \geq 0, \quad \forall k = \{1,\cdots,l\} \quad \forall m = \{1,\cdots,S\} \tag{46}$$

$$\text{C8: } \sum_{k=1}^{l} \left(\varrho_1^k P_{km} + \varrho_2^k \right) \leq P_{th}. \tag{47}$$

In case of $\text{P}^{1,l}$, $P_{km} = 0\ \forall\ k = \{l+1,\cdots,L\}$. Let the solution of $\text{P}^{1,l}$ be $\mathbf{P}_{\mathbf{lm}}^*$. Then the optimal solution of P^1 is:

$$\mathbf{P}_{\mathbf{m}}^* = \arg \max_{\{\mathbf{P}_{\mathbf{1m}}^*, \cdots, \mathbf{P}_{\mathbf{Lm}}^*\}} \sum_{l=1}^{L} \left((1-P_{ac})(1-P_f) \frac{T - \tau_s - S\tau_r}{ST} \times \frac{Q_l \cdot p_{fr}^{l,m}}{P_{DD}^{lm}} \right). \tag{48}$$

P^1 is solved here by choosing the best solution among all the solutions from corresponding L sub-problems. Next we describe how to solve the sub-problem $\text{P}^{1,l}$.

Lemma 1. $\log(1 - p_b^{k,m})$ *is concave in nature.*

Proof. Since, the coefficients $c_{k,m}$ and $d_{k,m}$ are positive, $p_b^{k,m}$ is a convex function of P_{km}. The logarithm function $\log(\cdot)$ is concave and increasing for any positive input value. Since $p_b^{k,m}$ is convex, $-p_b^{k,m}$ is concave. The composition rule [38] states that if $h(x)$ is concave and the function $g(\cdot)$ is concave and non-decreasing, then $f(x) = g(h(x))$ is concave in nature. Following composition rule, we can state that $\log(1 - p_b^{k,m})$ is concave function. □

Lemma 2. $\sum_{k=1}^{l} z_k \log(1 - p_b^{k,m})$ *is concave in nature.*

Proof. $z_k > 0$, and $\log(1 - p_b^{k,m})$ is concave function $\forall k, \forall m$. Convexity is preserved by the positive weighted summations [38]. Therefore, $\sum_{k=1}^{l} z_k \log(1 - p_b^{k,m})$ is concave in nature. □

Now, $-\log \left(P_{DD}^{lm} \right)$ is convex in nature. Therefore, the objective function in (45) is a summation of a convex and a concave function.

The derivative of P_{DD}^{lm} with respect to P_{lm} is ϱ_1^l. Let $O_1 = \sum_{k=1}^{l} z_k \log(1-p_b^{k,m}) - \log\left(P_{DD}^{lm}\right)$. The derivative of O_1 with respect to P_{km} is

$$\frac{\partial O_1}{\partial P_{km}} = \begin{cases} \frac{z_k c_{k,m} d_{k,m} \rho_{k,m} e^{-d_{k,m} P_{km} \rho_{k,m}}}{1-p_b^{k,m}}, & k \neq l \\ \frac{z_k c_{k,m} d_{k,m} \rho_{k,m} e^{-d_{k,m} P_{km} \rho_{k,m}}}{1-p_b^{k,m}} - \frac{\varrho_1^l}{\varrho_1^l P_{lm} + \varrho_2^l}, & k = l \end{cases} \tag{49}$$

Since $P_{km} \geq 0$, $\frac{\partial O_1}{\partial P_{km}} > 0$, $\forall k \neq l$. But when $k = l$, $\frac{z_k c_{k,m} d_{k,m} \rho_{k,m} e^{-d_{k,m} P_{km} \rho_{k,m}}}{1-p_b^{k,m}} > 0$ and $-\frac{\varrho_1^l}{\varrho_1^l P_{lm} + \varrho_2^l} < 0$. So, it is difficult to analytically determine the variation of O_1 with respect to P_{km}. Furthermore $\frac{\partial O_1}{\partial P_{km}} = 0$ is non-linear. So, the optimal value of P_{km} is determined using numerical method.

C. Computation Complexity

The optimization problem in (28) provides the optimal values of the system parameters for L-layer SVC video streaming via D2D network. These computations are done at the fusion center. The fusion center has high processing capability and the computation process is parallelized here to meet the deadlines. The computation complexity of P^2 is $S \cdot \left\lceil \log_2\left(\frac{T - S\tau_r}{\upsilon}\right) \right\rceil$. Here υ denotes the tolerance level for bisection search. In P^1, the optimal value of P_{lm} is determined following L number of sub-problems. These sub-problems correspond to L layers of SVC video and are solved in parallel. The computation complexity of P^1 is $\left\lceil \log_2\left(\frac{P_{th}}{\upsilon}\right) \right\rceil$ considering bisection search method.

V. Numerical Results

In this section, numerical results are used to demonstrate the performance of the proposed system. 1000 Monte Carlo simulations are carried out to highlight the variations in wireless channel. For simulation purpose, the normalized average utility U_{norm}^{av} of CRN is considered, $U_{norm}^{av} = \frac{U^{av}}{\max(U^{av})}$. Here $U^{av} = \frac{1}{S}\sum_{m=1}^{S} U_m$ represents the average utility. The variations of U_{norm}^{av} are shown with respect to spectrum sensing duration, sensing decision threshold, number of D2D nodes involved etc. 'Crew' video clip is considered here with parameter $b = 0.17$. In practical systems the number of enhancement layers are limited, mostly about 2 to 6 enhancement layers are considered. Previous work in the literature [39] focused on video layer selection for SVC based video streaming. However, this is not within the scope of the present work. Here, the number of enhancement layers are considered to be 2. So, the total number of layers $L = 3$.

TABLE III: AMC schemes and bit rates of different video layers

	Base layer	Enhancement layer 1	Enhancement layer 2
Bit rate E_k	79.2 kbps	165.80 kbps	315.80 kbps
AMC scheme	BPSK 1/2	QPSK 3/4	64 QAM 3/4

TABLE IV: List of simulation parameters

Name	Value
PBS transmission power P_T	46 dBm
Probability that PBS in active mode P_{ac}	0.3
Time slot duration T	100 ms
Distance between PBS and D2D transmitter d_p	500 m
Distance between D2D transmitter and primary receiver d_r	560 m
Distance between PBS and primary receiver d_{pr}	750 m
Average distance between D2D pairs	100 m
Sampling frequency f_s	100 KHz
Reporting time slot τ_r	0.1 ms
Reporting power consumption P_R	10 dBm
CSS power consumption P_S	15 dBm
Noise Variance P_n	-30 dBm
Target probability of detection $\hat{P}_d$	0.90
Average BER threshold $\hat{P}_b^{th}$	10^{-5}
Number of D2D pairs S	10
Target spectrum efficiency of PU R_p^{th}	0.3 bps/Hz
Power budget P_{th}	-10 dBW
Path loss exponent β	3.5

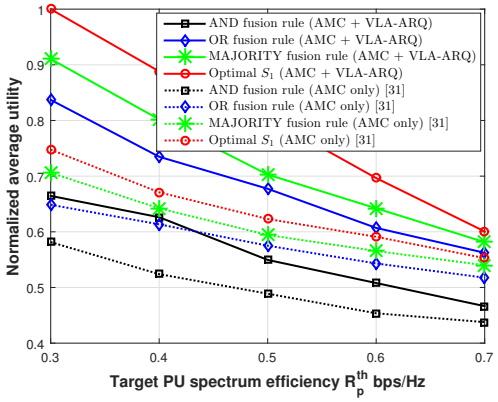


Fig. 3: Normalized average utility versus target PU spectrum efficiency R_p^{th}.

The frame rate F is set at 30 fps. The quantization parameters of the three layers (i.e. base layer and two enhancement layers) are set as QP $= [40, 36, 32]$. Considering the energy consumption circuits for D2D pair, the following parameters are set as $P_{filter} = 4$ dBm, $P_{mix} = 14$ dBm, $P_{syn} = 16$ dBm, $P_{LNA} = 13$ dBm, $P_{IFA} = 4$ dBm, $P_{DAC} = 11$ dBm, and $P_{ADC} = 10$ dBm [34]. The bit rates and the modulation schemes of video layers are listed in Table III. The list of other simulation parameters are shown in Table IV.

Four different transmission strategies are considered here. The first strategy considers fixed modulation scheme similar to the one in [20], ARQ is not considered here. In the second case, only AMC is considered (no ARQ). Here, different modulation and coding schemes are used for the transmission of different video layers [31]. The third strategy is AMC + conventional ARQ similar to the one in [23]. In this case ARQ is considered for all video layers. The final one is our proposed scheme (AMC + VLA-ARQ). Here along with AMC a video layer specific ARQ scheme is considered. For fair comparison, the other competitive schemes are implemented in our proposed optimization framework.

A. Performance Analysis for Different Fusion Rules

Fig. 3 shows normalized average utility of D2D network versus target spectrum efficiency of PU R_p^{th}. In this case, the

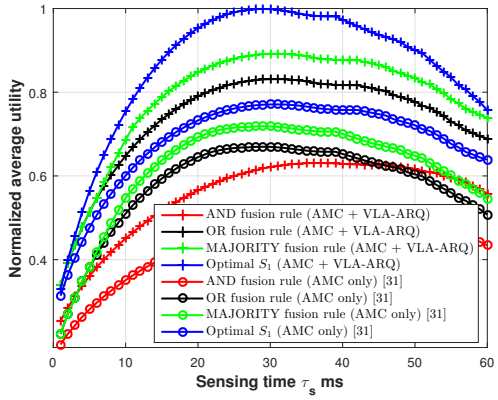
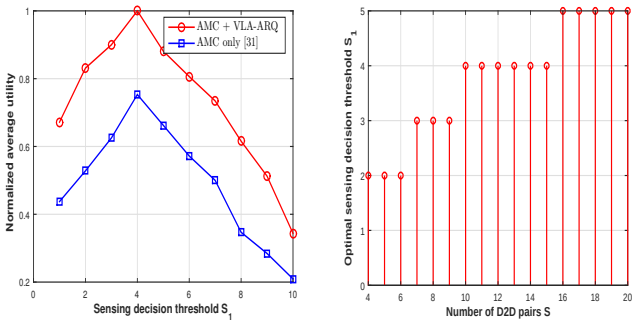


Fig. 4: Normalized average utility versus sensing time τ_s.



(a) (b)

Fig. 5: (a) Normalized average utility versus sensing decision threshold S_1, and (b) Optimal sensing decision threshold S_1 versus number of D2D pairs S.

optimal values of τ_s, S_1, P_{cm}, and P_{lm} are determined. The average utility value reduces with the increase in R_p^{th} value as more power is now needed for cooperation purpose. For the 'AMC + VLA-ARQ' approach (optimal fusion rule), the normalized utility value reduces about 39% as R_p^{th} is increased from 0.3 bps/Hz to 0.7 bps/Hz. Similar observations are also noted for 'MAJORITY','OR' and 'AND' fusion rules , characterized by $S_1 = \lfloor \frac{S}{2} \rfloor$, $S_1 = 1$, and $S_1 = S$, respectively. Here $\lfloor a \rfloor$ indicates the nearest integer lower than a. The performance of the proposed technique is also compared with respect to 'AMC only' approach considering optimal S_1 rule. About 11% gain is observed when $R_p^{th} = 0.7$ bps/Hz.

Fig. 4 illustrates the variation of normalized utility of CRN with respect to τ_s. The results are compared with 'MAJORITY', 'OR', and 'AND' fusion rules. For a given value of τ_s, the optimal values of S_1, P_{cm}, and P_{lm} are determined. It can be clearly seen that the proposed optimal fusion rule gives higher utility values compared to 'MAJORITY', 'OR', and 'AND' fusion rules. The worst performance can be seen with 'AND' rule. For $\tau_s = 30$ ms, the optimal fusion rule offers 58% higher utility value compared to 'AND' rule. From Fig. 4, we can also see that the utility value initially increases with τ_s till it reaches the optimal value. The optimal τ_s value for the proposed 'AMC + VLA-ARQ' is at 29 ms. When $\tau_s > 29$ ms, additional power is necessary to cooperate in PU's transmission (since cooperation time decreases). At this point, the transmission opportunity for D2D transmitters also reduces. As a result, the normalized average utility value gradually decreases. The performance of 'AMC + VLA-ARQ' is compared with the 'AMC only' approach considering optimal fusion rule. About 29% gain can be seen for the proposed approach when $\tau_s = 30$ ms. At the same operating point (i.e. $\tau_s = 30$ ms), the gain values are 24% and 28% for 'OR' and 'AND' fusion rules, respectively.

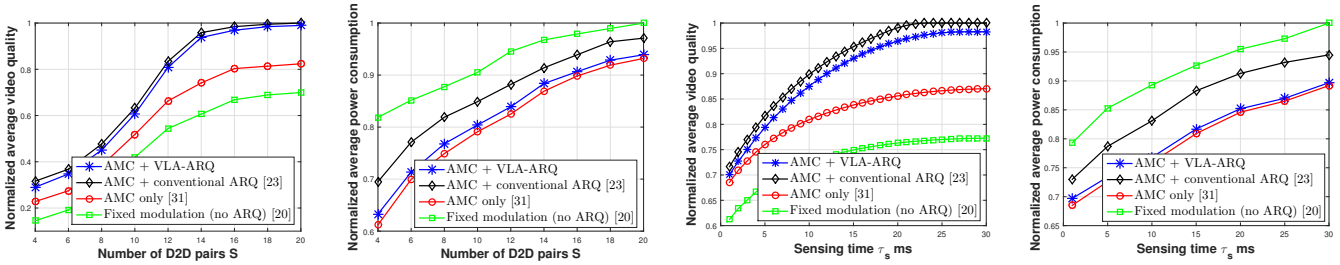
Fig. 5(a) shows the variation of average utility with the CSS decision threshold S_1. In this case, the optimal values of τ_s, P_{cm}, and P_{lm} are determined. It is observed that the utility is maximized when $S_1 = 4$. Furthermore, the proposed approach gives better performance compared to 'MAJORITY' ($S_1 = \lfloor \frac{S}{2} \rfloor$), 'OR' ($S_1 = 1$), and 'AND' ($S_1 = S$) fusion rules. The method in [31] also offers optimal utility value when $S_1 = 4$. Fig. 5(b) shows the variations of optimal S_1 with the number of D2D pairs S. The optimal value of S_1 changes from 2 to 5 when the value of S increases from 4 to 20.

The variations of U_{norm}^{av} are observed here with respect to τ_s, S_1, and R_P^{th}. The optimal fusion rule offers best performance of all the fusion rules. Therefore, the remaining analysis focuses only on the optimal fusion rule.

B. Performance Analysis for Different Transmission Strategies

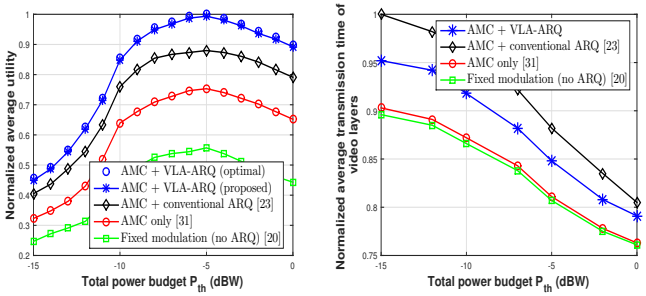
Variation of normalized video quality with the number of D2D pairs S is shown in Fig. 6(a). The number of D2D pairs are increased here from 4 to 20. The average video quality increases with the increase in number of D2D pairs S. The quality tends to saturate at higher values of S. The proposed approach offers a higher video quality compared to 'AMC only' and 'fixed modulation scheme'. Although the proposed approach offers reduced video quality compared to 'AMC + conventional ARQ' scheme, the performance loss is negligible. Maximum 4% performance loss can be observed here. Fig. 6(b) illustrates the variation of normalized average power consumption of D2D nodes with the number of D2D pairs S. Similar to Fig. 6(a), the average power consumption increases with the increase in S. It can be clearly observed that the proposed scheme is more energy efficient compared to 'fixed modulation' and 'AMC + conventional ARQ scheme'. Up to 11% reduction in power consumption is noted compared to the 'AMC + conventional ARQ scheme'. Although 'AMC only' approach has the least power consumption, video quality performance is quite poor compared to the proposed approach.

The variations of normalized video quality and normalized average power consumption versus sensing time τ_s are shown in Figs. 6(c) and 6(d), respectively. It can be clearly seen that the video quality gradually increases with the increase in τ_s value and then tends to saturate. The performance of the proposed approach is also compared here with the three competitive schemes. From Fig. 6(c), it can be noted that the proposed technique performs better than 'AMC only', and the 'fixed modulation' schemes. About 13% gain in video quality can be seen compared to 'AMC only' approach when $\tau_s = 20$ ms. 'AMC along with conventional ARQ' offers a marginally



(a) (b) (c) (d)

Fig. 6: (a) Normalized average video quality versus number of D2D pairs S, (b) Normalized average power consumption versus number of D2D pairs S, (c) Normalized average video quality versus sensing time τ_s (optimal S_1), and (d) Normalized average power consumption versus sensing time τ_s (optimal S_1).



(a) (b)

Fig. 7: (a) Normalized average utility versus power budget P_{th}, and (b) Normalized average transmission time versus power budget P_{th}.

higher video quality by layer oblivious re-transmissions, but at the cost of high power consumption – as seen in Fig. 6(d). From Figs. 6(c) and 6(d), it can be clearly inferred that the proposed scheme offers a better balance between high video quality and reduced average power consumption values compared to all three competitive schemes. This can be better understood from the analysis of utility function, since it takes into account the effects of both video quality and power consumption at D2D users. The high values of the utility function indicate improved video quality while maintaining reduced power consumption.

Keeping this in mind, Fig. 7(a) illustrates the variations of average utility with total power budget P_{th} for the proposed approach, 'AMC + conventional ARQ', 'AMC only', and 'fixed modulation scheme'. At lower values of P_{th}, the utility increases with the increase in P_{th}, as the received video quality improves in presence of additional power. At higher P_{th} (i.e., $P_{th} > -5$ dBW), the utility gradually decreases as the received video quality tends to saturate. As shown in Fig. 6(c), 'AMC + conventional ARQ' has a higher performance in terms of video quality. However, 'AMC + VLA-ARQ' scheme considers re-transmission only for the base layer, therefore it is more energy efficient compared to 'AMC + conventional ARQ' (see Fig. 6(d)). Overall, from Fig. 7(a) it can be seen that the proposed approach offers improved performance in terms average utility when compared with the other three competitive techniques. About 14% gain in overall utility is noted compared to 'AMC + conventional ARQ' when $P_{th} = -5$ dBW. The performance of the proposed approach is also compared with the optimal solution, obtained by exhaustive searching technique. It can be seen that our proposed technique provides almost similar performance to the optimal one. Finally, Fig. 7(b) demonstrates that the signaling overhead in terms of normalized transmission time of video stream (consisting of 3 layers) for the proposed approach is lower than closest competitive 'AMC + conventional ARQ' strategy [23]. The proposed approach requires higher transmission time compared to the other schemes in [20], [31]; however, these schemes hold no competitive advantage as significant improvement in average utility is also noted with the proposed scheme (see Fig. 7(a)).

VI. Conclusions

In this paper, a cross-layer parameter adaptation scheme has been proposed for scalable video transmission over wireless channel. This approach combines AMC at the physical layer, video layer aware ARQ at the link layer, and scalable coding at the application layer. A novel frame structure for CSS-cooperation-transmission has been proposed for cognitive radio-enabled D2D network to maximize received video quality and simultaneously minimize total power consumption. Performance of the proposed approach has been compared with three other competitive schemes. The proposed approach provides improved video quality as well as reduced energy consumption at the D2D users. About 14% gain in overall utility has been observed compared to the conventional ARQ based technique.

References

[1] "Cisco visual networking index: Global mobile data traffic forecast update, 2015–2020," San Jose, CA, USA, Cisco, White Paper, Jan. 2016. [Online]. Available: ''http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html"

[2] D. Song and C. W. Chen, "Maximum-throughput delivery of SVC-based video over MIMO systems with time-varying channel capacity," *J. Vis. Commun. Image Rep.*, vol. 19, no. 8, pp. 520 – 528, Dec. 2008.

[3] T. Q. Duong, N. Vo, T. Nguyen, M. Guizani, and L. Shu, "Energy-aware rate and description allocation optimized video streaming for mobile D2D communications," in *Proc. IEEE Int. Conf. Commun. (ICC)*, June 2015, pp. 6791–6796.

[4] S. Khan, S. Duhovnikov, E. Steinbach, M. Sgroi, and W. Kellerer, "Application-driven cross-layer optimization for mobile multimedia communication using a common application layer quality metric," in *Proc. IEEE ACM IWCMC*, July 2006, pp. 213–218.

[5] L. Tang, Y. Chen, E. L. Hines, and M. Alouini, "Effect of primary user traffic on sensing-throughput tradeoff for cognitive radios," *IEEE Trans. Wireless Commun.*, vol. 10, no. 4, pp. 1063–1068, Apr. 2011.

[6] Y. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[7] A. Ahmed and K. Baishnab, "Joint optimal design of sensing time and transmission power for maximizing energy efficiency in cognitive radio system," *Wireless Pers. Commun.*, vol. 110, p. 1839–1857, Feb. 2020.

[8] A. Ostovar, Y. B. Zikria, H. S. Kim, and R. Ali, "Optimization of Resource Allocation Model With Energy-Efficient Cooperative Sensing in Green Cognitive Radio Networks," *IEEE Access*, vol. 8, pp. 141 594–141 610, 2020.

[9] C. Wang, T. Song, J. Wu, W. Jiang, and J. Hu, "Energy-efficient optimal sensing and resource allocation of soft cooperative spectrum sensing in CRNs," in *Proc. 11th Int. Conf. Wireless Commun. Sig. Process. (WCSP)*, 2019, pp. 1–6.

[10] J. Adu Ansere, G. Han, H. Wang, C. Choi, and C. Wu, "A reliable energy efficient dynamic spectrum sensing for cognitive radio iot networks," *IEEE Int. Things J.*, vol. 6, no. 4, pp. 6748–6759, 2019.

[11] M. R. Amini, M. Mahdavi, and M. J. Omidi, "Maximizing dynamic access energy efficiency in multiuser crns with primary user return," *IEEE Syst. J.*, vol. 13, no. 2, pp. 1702–1713, 2019.

[12] X. Wang, M. Jia, Q. Guo, I. W.-H. Ho, and F. C.-M. Lau, "Full-duplex relaying cognitive radio network with cooperative nonorthogonal multiple access," *IEEE Syst. J.*, vol. 13, no. 4, pp. 3897–3908, 2019.

[13] A. H. Bastami and P. Kazemi, "Cognitive multi-hop multi-branch relaying: Spectrum leasing and optimal power allocation," *IEEE Trans. Wireless Commun.*, vol. 18, no. 8, pp. 4075–4088, Aug. 2019.

[14] N. Li, L. K. Rasmussen, and M. Xiao, "Performance analysis of cognitive user cooperation using binary network coding," *IEEE Trans. Veh. Technol.*, vol. 67, no. 8, pp. 7355–7369, Aug. 2018.

[15] M. Usman, N. Yang, M. A. Jan, X. He, M. Xu, and K.-M. Lam, "A joint framework for QoS and QoE for video transmission over wireless multimedia sensor networks," *IEEE Trans. Mobile Comput.*, vol. 17, no. 4, pp. 746–759, 2018.

[16] Z. Xu, Y. Cao, W. Wang, T. Jiang, and Q. Zhang, "Incentive mechanism for cooperative scalable video coding (svc) multicast based on contract theory," *IEEE Trans. Multimedia*, vol. 22, no. 2, pp. 445–458, 2020.

[17] Q. Jiang, V. C. M. Leung, M. T. Pourazad, H. Tang, and H. Xi, "Energy-efficient adaptive transmission of scalable video streaming in cognitive radio communications," *IEEE Syst. J.*, vol. 10, no. 2, pp. 761–772, June 2016.

[18] H. Zarini, A. Khalili, H. Tabassum, and M. Rasti, "Joint Transmission in QoE-Driven Backhaul-Aware MC-NOMA Cognitive Radio Network," in *Proc. IEEE Global Commun. Conf.*, 2020, pp. 1–6.

[19] J. Jia, Y. Deng, J. Chen, A. Aghvami, and A. Nallanathan, "Availability analysis and optimization in CoMP and CA-enabled hetnets," *IEEE Trans. Commun.*, vol. 65, no. 6, pp. 2438–2450, June 2017.

[20] J. Chen, Y. Deng, J. Jia, M. Dohler, and A. Nallanathan, "Cross-layer QoE optimization for D2D communication in CR-enabled heterogeneous cellular networks," *IEEE Trans. Cogn. Commun. Netw.*, vol. 4, no. 4, pp. 719–734, Dec. 2018.

[21] A. J. Goldsmith and S. Chua, "Adaptive coded modulation for fading channels," *IEEE Trans. Commun.*, vol. 46, no. 5, pp. 595–602, May 1998.

[22] E. Malkamaki and H. Leib, "Performance of truncated type-ii hybrid ARQ schemes with noisy feedback over block fading channels," *IEEE Trans. Commun.*, vol. 48, no. 9, pp. 1477–1487, Sep. 2000.

[23] Qingwen Liu, Shengli Zhou, and G. B. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1746–1755, Sep. 2004.

[24] X. Wang, Q. Liu, and G. B. Giannakis, "Analyzing and optimizing adaptive modulation coding jointly with ARQ for QoS-guaranteed traffic," *IEEE Trans. Veh. Technol.*, vol. 56, no. 2, pp. 710–720, Mar. 2007.

[25] C. Zhu, Y. Huo, B. Zhang, R. Zhang, M. El-Hajjar, and L. Hanzo, "Adaptive-truncated-HARQ-aided layered video streaming relying on interlayer FEC coding," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1506–1521, Mar. 2016.

[26] H. Lu, X. Jiang, and C. W. Chen, "Distortion-Aware Cross-Layer Power Allocation for Video Transmission Over Multi-User NOMA Systems," *IEEE Trans. Wireless Commun.*, vol. 20, no. 2, pp. 1076–1092, 2021.

[27] M. K. Jubran, M. Bansal, L. P. Kondi, and R. Grover, "Accurate distortion estimation and optimal bandwidth allocation for scalable H.264 video transmission over MIMO systems," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 106–116, Jan. 2009.

[28] D. Song and C. W. Chen, "Scalable H.264/AVC video transmission over MIMO wireless systems with adaptive channel selection based on partial channel information," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 17, no. 9, pp. 1218–1226, Sep. 2007.

[29] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[30] C. Taneja, S. Chatterjee, and S. De, "QoE-aware cross-layer adaptation for delay-constrained video transmission over wireless channels," in *Proc. IEEE Wireless Commun. Net. Conf.*, 2019, pp. 1–6.

[31] X. Chen, J. Hwang, C. Lee, and S. Chen, "A near optimal QoE-driven power allocation scheme for scalable video transmissions over MIMO systems," *IEEE J. Sel. Topics Sig. Process.*, vol. 9, no. 1, pp. 76–88, Feb. 2015.

[32] Yao Wang, Zhan Ma, and Yen-Fu Ou, "Modeling rate and perceptual quality of scalable video as functions of quantization and frame rate and its application in scalable video adaptation," in *Proc. 17th Int. Packet Video Workshop*, May 2009, pp. 1–9.

[33] V. R. Reddyvari and A. K. Jagannatham, "Quality optimal policy for H.264 scalable video scheduling in broadband multimedia wireless networks," in *Proc. Int. Conf. Signal Process. Commun. (SPCOM)*, July 2012, pp. 1–5.

[34] Y. Shen, C. Jiang, T. Q. S. Quek, and Y. Ren, "Device-to-device-assisted communications in cellular networks: An energy efficient approach in downlink video sharing scenario," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1575–1587, Feb. 2016.

[35] C. Singhal, S. De, R. Trestian, and G. Muntean, "Joint optimization of user-experience and energy-efficiency in wireless multimedia broadcast," *IEEE Trans. Mobile Comput.*, vol. 13, no. 7, pp. 1522–1535, July 2014.

[36] H.-J. Lim, M.-G. Song, and G.-H. Im, "Cooperation-based dynamic spectrum leasing via multi-winner auction of multiple bands," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1254–1263, Apr. 2013.

[37] D. P. Palomar and M. Chiang, "A tutorial on decomposition methods for network utility maximization," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1439–1451, Aug. 2006.

[38] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[39] J. Yang, P. Si, Z. Wang, X. Jiang, and L. Hanzo, "Dynamic resource allocation and layer selection for scalable video streaming in femtocell networks: A twin-time-scale approach," *IEEE Trans. Commun.*, vol. 66, no. 8, pp. 3455–3470, Aug. 2018.

Subhankar Chatterjee (S'16-M'18) is a postdoctoral fellow in the Department of Electrical Engineering at IIT Delhi. He completed his M.Sc. in Mobile and Satellite Communications from University of Surrey, United Kingdom in 2010 with distinction, and Ph.D. in Engineering from IIEST Shibpur, India in 2017. His research interests are broadly in cognitive radio, and cross-layer optimization for broadband communications.

Swades De (S'02-M'04-SM'14) is a Professor in the Department of Electrical Engineering at IIT Delhi. Before moving to IIT Delhi in 2007, he was a Tenure-Track Assistant Professor of Electrical and Computer Engineering at the New Jersey Institute of Technology (2004-2007). He worked as an ERCIM post-doctoral researcher at ISTI-CNR, Pisa, Italy (2004), and has nearly five years of industry experience in India on telecom hardware and software development (1993-1997, 1999). His research interests are broadly in communication networks, with emphasis on performance modeling and analysis. Current directions include energy harvesting sensor networks, broadband wireless access and routing, cognitive/white-space access networks, smart grid networks, and IoT communications. Dr. De currently serves as an Area Editor for IEEE COMMUNICATIONS LETTERS and Elsevier Computer Communications, and an Associate Editor for IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY and IEEE WIRELESS COMMUNICATIONS LETTERS.