

Resource Allocation and Performance Analysis for Multiuser Video Transmission Over Doubly Selective Channels

Dawei Wang, *Student Member, IEEE*, Laura Toni, *Member, IEEE*,
Pamela C. Cosman, *Fellow, IEEE*, and Laurence B. Milstein, *Fellow, IEEE*

Abstract—We consider an uplink multicarrier system with multiple video users who want to send compressed video data to the base station. In the time domain, we model the time-varying channel using Jakes' model, and in the frequency domain, each subcarrier is assumed to be independently fading. The video is scalably coded in units of a group of pictures (GOP), and users have different video rate distortion (RD) functions. At the beginning of the GOP, the base station collects both the RD information and the instantaneous channel state information (CSI) for subcarrier allocation purposes. We design a cross-layer resource allocation algorithm to assign subcarriers to users based on both the demand of the video and the quality of the channel. Once the resource allocation decision is made, the users then periodically adapt the modulation format of the subcarriers allocated according to the evolution of the CSI for the duration of the GOP. We show that our cross-layer resource allocation robustly outperforms two baseline algorithms, each of which uses only one layer of information for resource allocation.

Index Terms—Time-varying channel, radio spectrum management, multimedia communication, multicarrier system, wireless power allocation.

I. INTRODUCTION AND RELATED WORK

OVER the past decade, the high demand for data rate for multimedia transmission and the limitation of communication bandwidth have become the bottleneck to multimedia communications. For cellular and wireless local area network systems, various PHY/APP cross layer techniques have been studied to both improve video quality and increase cell capacity. Among them, point-to-point PHY/APP cross layer optimization

seeks to exploit the unique characteristics of video, and applies schemes like multiple description coding to provide different protection levels and achieve higher end-to-end user quality of experience (QoE). Most research has focused on slow-varying channels [1]–[3].

To combat the uncertainty of a time-varying channel, video communication with automatic repeat request (ARQ) was proposed in [4] and [5]. Although ARQ is easy to implement, the time delay and uncertainty for exchanging the ARQ signals might not be suitable for delay sensitive video applications. More importantly, for a system with different Doppler spreads, the packet loss rate (PLR) varies dramatically with respect to channel estimation accuracy, which is determined jointly by the pilot spacing, the pilot power and the number of pilots used for interpolation [6]. Most papers on ARQ-based video communication oversimplify the PLR model. In [4], PLR is treated as a constant for all Doppler spreads. In [5], the authors study the performance of adaptive modulation with ARQ in a data communication system, and perfect channel estimation is assumed for choosing the modulation format and demodulation at the receiver. Channel estimation accuracy could be improved by reducing the interval between the pilots. However, the throughput loss due to pilot insertion might significantly reduce the number of video source bits delivered to the channel. Under the perfect channel state information (CSI) assumption, the critical tradeoff between the channel estimation accuracy and source encoding rate is missed in [5]. Channel variation can sometimes be accommodated by using buffers at both the transmitter and receiver sides. Recently, video oriented protocols such as HTTP adaptive streaming (HAS) and dynamic adaptive streaming over HTTP (DASH) have been introduced for delivering enhanced QoE video over the Internet [7], [8]. In this paper, we are interested in a scenario of tight delay constraints for video delivery, and do not consider the presence of buffers.

Forward Error Correction (FEC)-based video communication with no retransmission is often used for delay-sensitive video data. To achieve higher average image quality in systems with high mobility, [9] utilizes the coding diversity across both time and frequency, and analyzes the performance of progressive image transmission in the presence of inter-carrier interference and channel estimation error in a multicarrier setting. In [10], the authors study a joint link and source adaptation system, where the modulation and coding scheme at the PHY layer is chosen according to instantaneous CSI and different

Manuscript received November 10, 2013; revised March 19, 2014 and August 2, 2014; accepted November 16, 2014. Date of publication December 4, 2014; date of current version April 7, 2015. This work was supported by the Intel-Cisco Video Aware Wireless Networks program and in part by the National Science Foundation under Grant CCF-0915727. The associate editor coordinating the review of this paper and approving it for publication was A. A. Abouzeid.

D. Wang was with the University of California at San Diego, La Jolla, CA 92093 USA. He is now with Adaptive Spectrum and Signal Alignment (ASSIA) Inc., Redwood City, CA 94065 USA (e-mail: dwangnb@gmail.com; daw017@ucsd.edu).

L. Toni was with the University of California at San Diego, La Jolla, CA 92093 USA. She is now with École Polytechnique Fédérale de Lausanne, Lausanne 1015, Switzerland (e-mail: laura.toni@epfl.ch).

P. C. Cosman and L. B. Milstein are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92037 USA (e-mail: pcosman@ucsd.edu; milstein@ece.ucsd.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TWC.2014.2377172

importance levels of the packet, while the source rate at the APP layer is chosen based on the visibility of the packet. Without any assumption of knowledge of the channel in the future, the adaptation scheme in [10] has the potential to be applicable for systems with arbitrary mobility. Both [9] and [10] focus on point-to-point multimedia scenarios where the user is assumed to have a fixed resource allocation regardless of the demand.

Another technique is to exploit the relative diversity (both at the PHY layer and the APP layer) in a multiple access environment. To utilize multiple user channel diversity, multicarrier systems are widely used in cellular systems, as the resource allocation for multicarrier systems can be done flexibly. Multiple user subcarrier assignment with different user rate demand is investigated in [11], where the authors try to maximize the weighted sum of throughputs for different user priority levels. In [12] and [13], user priority is further abstracted as a utility function, and the goal of the resource allocation is to maximize the sum of utilities. In [14], a multiple subcarrier assignment problem is presented, where both user rate distortion (RD) functions and CSI of multiple subcarriers are used for resource allocation to minimize the sum of distortions. Inspired by the equal-slope condition for video multiplexing [15], the authors derived a necessary condition for optimal spectrum sharing and designed an iterative algorithm for multiple subcarrier assignment. In references such as [12] on multiuser multicarrier resource allocation, the objective function (e.g., the utility function in [12]) is measured over some interval of time, and the channel is assumed to be stable for that duration.

In this paper, we study a multiuser scalable video uplink system in a doubly selective environment. Unlike [14], we do not assume any relation between the channel Doppler spread and video group of pictures (GOP) duration. For the physical layer, we use pilot symbol assisted modulation (PSAM) and allow users to adaptively change modulation format between resource allocation decisions. With the goal of minimizing average distortion, we design a PHY/APP cross layer resource allocation algorithm which takes into account both throughput loss and channel estimation error. As discussed above, for systems with high mobility, the end-to-end video performance highly depends on the tradeoff between the channel estimation accuracy and video source encoding rate. Unlike [14], we take the throughput loss due to pilot insertion into consideration in our resource allocation design in this paper. The assumption of low mobility in [14] and arbitrary mobility in this paper make the system design different. By adjusting the spacing between the pilot symbols, we show via simulation the optimized values for channel adaptation. We further quantify the tradeoff between the loss of throughput due to pilot insertion and loss of performance due to channel outdating. We show that our previously designed resource allocation algorithm in [14] can be made robust to Doppler spread and outperform the baseline resource allocation algorithms that only use a single layer of information. To the best of our knowledge, this paper is the first work concerning video communication resource allocation with arbitrary user mobility.

The rest of the paper is organized as follows: In Section II, we describe the system model for a multiple user multicarrier video communication system. We also provide details on the

application layer scalable video codec and the physical layer doubly selective fading channel model. We then discuss the resource allocation framework and the use of PSAM scheme for arbitrary user mobility in Section III. We present three resource allocation algorithms, each with a different degree of knowledge of the system information in Section IV, and we show their performance results in Section V. We conclude the paper in Section VI.

II. SYSTEM MODEL

We consider a cellular multicarrier video communication system with a set of video users indexed by k , $k = \{1, 2, 3 \dots K\}$. The system occupies a total frequency band of W (Hz) equally divided into M_c ($M_c \geq K$) subcarriers indexed by m , $m = \{1, 2, 3 \dots M_c\}$. We assume users experience the same Doppler spread. We focus on an uplink system, and the task of the resource allocation is to assign subcarriers to users based on both RD functions and CSI to minimize the sum of video distortions. The same principle of the system design presented in this paper can also be applied in a downlink system.

A. Doubly Selective Channel Model

The system operates in a slotted manner, and the length of one time slot is T_s (sec), equal to both the video display time and the transmission duration of one GOP. As depicted in Fig. 1, we assume a block fading model in the frequency domain with coherence bandwidth $B_c = \Psi W / M_c$. Here, W / M_c is the bandwidth of each subcarrier, and Ψ is the coherence bandwidth in units of subcarriers. Let $H_{k,m}[l]$ be the complex channel gain of user k for subcarrier m at the l -th symbol. The subcarrier assignment as well as the power allocation decision will be made on a slot-to-slot basis. Each subcarrier can only be used by one user, but it is possible for one user to get more than one subcarrier. Further, $H_{k,m}[l] = \gamma_k \alpha_{k,m}[l]$, where γ_k depends on the path-loss coefficient, the distance between the mobile user and the base station, and the shadowing caused by obstacles. For user k , we assume that γ_k is fixed for the duration of a GOP. The variable $\alpha_{k,m}[l]$ captures the multipath fading and is modeled as a zero-mean complex stationary Gaussian random process [16]. The magnitude of $\alpha_{k,m}[l]$ is Rayleigh distributed with a variance of unity for a non-line-of-sight system. The band-limited spectrum of $\alpha_{k,m}(t)$ is given by

$$S(f) = S(0) \left[1 - \frac{f}{f_d} \right]^{-1/2}, \quad |f| < f_d \quad (1)$$

where f_d is the Doppler spread, and $S(0) = 2/\pi$. Define $f_{nd} = f_d T_0$ as the normalized Doppler spread, where T_0 is the symbol duration. The autocorrelation between two symbols l and $l + \Delta l$ can be written as [16]

$$\mathbb{E} [\alpha_{k,m}[l] \alpha_{k,m}^*[l + \Delta l]] = J_0 [2\pi \cdot f_{nd} \cdot (\Delta l)] \quad (2)$$

where $J_0(\cdot)$ is the zeroth order Bessel function of the first kind. The first zero crossing of the correlation function occurs when the product $f_{nd} \cdot (\Delta l)$ is about 0.4.

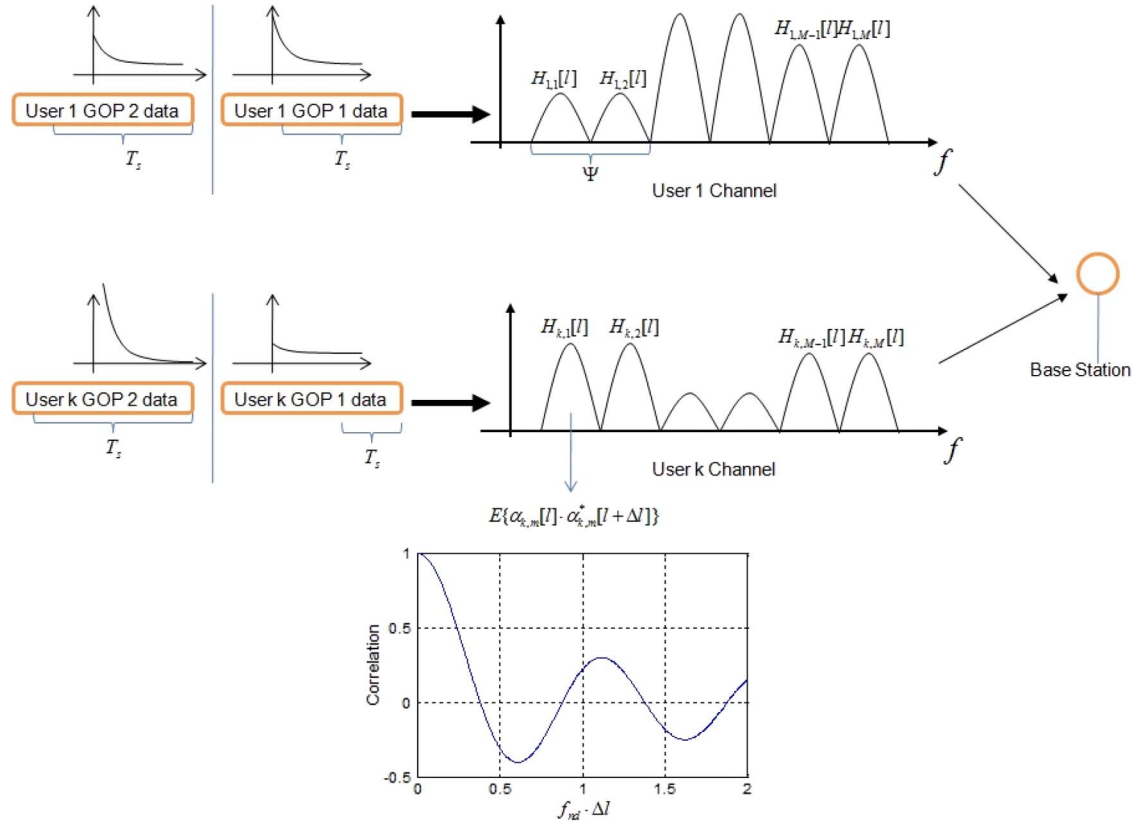


Fig. 1. Doubly selective channel model and resource allocation for multiple user video transmission system.

For a multicarrier system, the complex envelope of the transmitted signal for user k can be written as

$$x_k(t) = \sum_l \sum_{m=1}^{M_c} \sqrt{P_{k,m}} X_{k,m}[l] \exp\left(\frac{j2\pi m t}{T_0}\right) g(t - lT_0) \quad (3)$$

where $P_{k,m}$ and $X_{k,m}[l]$ are the transmission power and coded symbol with unit variance, respectively, of user k on subcarrier m . The power $P_{k,m}$ is assumed to be fixed for the duration of a time slot, and $P_{k,m} = 0$ if subcarrier m is not allocated to user k . Also, $g(t)$ is a zero-excess bandwidth Nyquist pulse, with $G(f) = 1, \forall f \in [-1/2T_0, 1/2T_0]$, and $G(f) = 0$ otherwise.

Since we assume flat fading for each subcarrier, the lowpass equivalent received signal of user k on subcarrier m is given by

$$y_{k,m}(t) = \sqrt{P_{k,m}} H_{k,m}[l] X_{k,m}[l] \exp\left(\frac{j2\pi m t}{T_0}\right) g(t - lT_0) + n_{k,m}(t) \quad (4)$$

where $n_{k,m}(t)$ is complex additive white Gaussian noise (AWGN) with two-sided power spectral density N_0 . Research in [17]–[19] shows that timing and frequency synchronization problems can be properly solved by using joint estimation schemes with a reasonable training sequence. Thus, for simplicity, we ignore the effect of intercarrier interference and intersymbol interference. To detect the signal on subcarrier m , a correlation operation is performed:

$$Y_{k,m} \triangleq \int_{-\infty}^{\infty} y_{k,m}(t) \exp\left(\frac{-j2\pi m t}{T_0}\right) g(t) dt \quad (5)$$

The noise power is given by $P_N = E[|N_{k,m}|^2]/2 = N_0/T_0$, and the power for the desired signal is $P_{k,m}|H_{k,m}[l]|^2$. If the modulation

format is adaptive QAM, from [20] and [21], the symbol error rate (SER) for an AWGN channel can be approximated as

$$SER \approx 4Q \left(\sqrt{\frac{3}{M-1} \frac{P_{k,m}|H_{k,m}[l]|^2}{P_N}} \right) \quad (6)$$

Here, M is the alphabet size of a QAM waveform, and for a given fixed SER_t , the information rate (number of bits each symbol can carry) $R_{k,m}(P_{k,m}, H_{k,m}[l])$ (in bits/symbol) can be written as a function of transmission power and channel response gain:

$$R_{k,m}(P_{k,m}, H_{k,m}[l]) = \min \left\{ \left\lceil \log_2 \left[1 + \eta P_{k,m} |H_{k,m}[l]|^2 \right] \right\rceil, R_{max} \right\} \quad (7)$$

where $\eta = \frac{3}{P_N} [Q^{-1}(SER_t/4)]^{-2}$ and R_{max} is the largest alphabet size the system allows. The term $\log_2[1 + \eta P_{k,m}|H_{k,m}[l]|^2]$ is obtained by solving (6) for M , and then taking the log of the result. We take the flooring operation since the number of bits per symbol has to be an integer. The bit rate (in bits/sec) then can be written as $R_{k,m}(P_{k,m}, H_{k,m}[l])/T_0$. In the following sections, we will replace $H_{k,m}[l]$ by the estimate $\tilde{H}_{k,m}[l]$ in (7) and use $R_{k,m}(P_{k,m}, \tilde{H}_{k,m}[l])$ to determine the modulation alphabet size. Note that, because of the channel estimation error and the floor operation in (7), the actual SER of the symbols might be different from the parameter SER_t . The effect of the channel estimation accuracy and the choice of SER_t will be investigated in the simulation section.

B. Scalable Video Codec

A scalable video codec is designed such that the decoder only needs a portion of the encoded bitstream (a substream) to display the video. The decoded fidelity of the video depends on the length of the substream, as well as the rate distortion characteristics of the video content. Scalable coding allows flexible adaptation to time-varying wireless channels and throughput variations in multiple-hop communication systems.

For each bitstream, the most important video information (e.g., coding modes) is contained in a substream called the base layer (BL). One or more enhancement layers (EL) are added such that the mean square error (MSE) will decrease when more enhancement bits are received by the decoder. Previously, a fine granular-scalability (FGS) codec was proposed [22], [23] for accurate source-to-channel adaptation. FGS allows every successfully delivered video bit to improve the video quality, but the granularity of the scalability will sacrifice the video compression efficiency. The scalable extension of H.264/AVC (known as H.264/SVC) [24] with medium granular scalability (MGS) has emerged as a balanced solution for the tradeoff between compression efficiency and scalable granularity. An H.264/SVC MGS codec features temporal, spatial and quality scalability, and allows the flexibility of dropping a combination of substreams according to the communication channel. In this paper, we are interested in the quality scalability function defined in [24], which packetizes the encoded bitstream according to the zonal location of the DCT coefficients and ranks the packets based on their importance in the GOP. The encoder assigns the highest priority for transmission to the packets which can most effectively reduce the compression distortion. If an error occurs in the transmission, as one of the properties of the H.264/SVC decoder [24], the entire packet and all successive packets within the GOP will be dropped, but previous packets (all of which have higher priority) will be used for decoding the GOP.

To characterize the tradeoff between the compression fidelity and the number of bits used to describe the source, we model the rate distortion (RD) curve of the video using a parameterized function. Since the video is compressed in units of GOPs, this RD function is also measured on a GOP-by-GOP basis. Let $D_k(B)$ be the RD function of user k , where B is the number of bits in the substream (the length of the truncated bit stream). For each GOP, the MSE distortion can be approximated as [25]

$$D_k(B) = a_k + \frac{w_k}{B + v_k} \quad (8)$$

where a_k , v_k , and w_k are constants which depend on the video content. When the picture is relatively spatially uniform, and the motion of the video is slow, the time and spatial redundancy can be easily compressed, and one would expect a relatively flat RD function. For a video with high complexity and fast motion, the RD function is normally steep and w_k is relatively large. The difference of the RD tradeoff between different users constitutes application layer diversity. Since the path-loss and shadowing are assumed to be constant for the duration of the GOP, we drop the index of the GOP for simplicity in the following sections.

III. CROSS LAYER VIDEO UPLINK SYSTEM WITH ARBITRARY USER MOBILITY

In a multiple user system, to minimize the sum of the MSEs across all users, the base station collects the RD information (coefficients of a_k , w_k , and v_k in (8)) as well as the CSI of the subcarriers, and allocates the subcarriers jointly according to application layer and physical layer information. In [14], we studied video resource allocation for a system in which the channel varies slowly over the duration of a GOP. Under the condition of constant CSI for the entire GOP, we assumed that the modulation format remains unchanged for one GOP, and the throughput of each subcarrier could be perfectly estimated at the beginning of each GOP. The resource allocation problem then becomes a mixed integer programming problem, and allocation decisions are made on a GOP-by-GOP basis.

A. System Operation Overview

For mobile users operating in an environment with a high enough Doppler spread, the CSI estimated at the beginning of the GOP will be outdated prior to the end of the GOP. For subcarrier m , if the modulation format determined by the CSI at the beginning of the GOP is held constant for the entire GOP duration, it is likely that the video data will either be over-protected or under-protected. On the other hand, the resource allocation decision based on both RD information and CSI is normally of high complexity [26], [27]. If we update the resource allocation decision every coherence time, the base station will need to collect the instantaneous CSI of all the subcarriers for all the users, as well as the amount of the GOP that has already been transmitted. If the allocator makes a resource allocation decision at each coherence time, it is not only computationally difficult, but also requires a large amount of information exchange. We thus propose a scheme having two phases which balances the computational complexity with the adaptation accuracy.

Phase I Cross Layer Resource Allocation: At the beginning of each GOP, user k submits the RD function $D_k(B)$ of the current GOP to the base station. A duration of K symbols is dedicated for channel estimation for every subcarrier for all users. Each user is assigned one symbol and users sequentially send pilot symbols to the base station for channel estimation. The instantaneous CSI of user k , $H_{k,m}^c$, ($k \in \{1, 2, \dots, K\}$, $m \in \{1, 2, \dots, M_c\}$) for every subcarrier for each user is jointly used with the RD information by the base station to make an allocation decision. The allocation decision is fed back to the users and each user is allowed to access the subcarriers assigned to him for the entire GOP duration. The resource allocation algorithm and the information exchange for the RD function is conducted once per GOP.

Phase II Pilot Assisted Adaptive Modulation: After resource allocation, when each individual user knows the subset of assigned subcarriers, each user periodically sends a pilot symbol to the base station for channel estimation purposes. Based on the instantaneous CSI at each period, the base station updates the modulation format of each subcarrier and feeds the new modulation formats back to the corresponding users. The

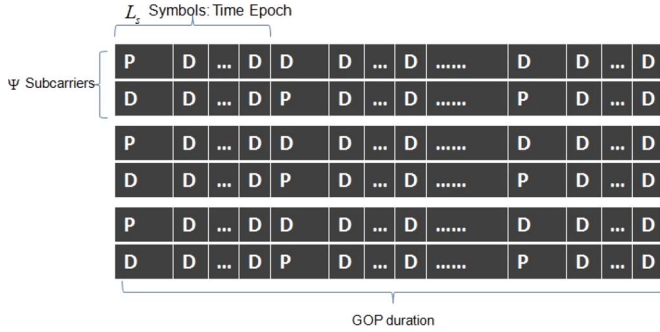


Fig. 2. Pilot assisted modulation. One pilot symbol is added for each subcarrier to estimate the channel for every time epoch. If the number of correlated subcarriers is larger than 1, only one pilot symbol is needed for every correlated band.

estimated CSI is also used for demodulation purposes. Since the modulation format is updated periodically, the number of information bits transmitted cannot be estimated accurately at the beginning of the GOP. The important bits will be transmitted first, and the actual number of bits transmitted is determined by the overall channel conditions over the duration of a GOP (T_s seconds). Note that in this phase, no information about the RD is exchanged, and the resource allocation decision is not updated. The details of the adaptive modulation scheme are discussed below.

B. Pilot Symbol Assisted Modulation (PSAM)

As depicted in Fig. 2, at most one pilot symbol will be sent from the user to the base station in each subcarrier every L_s symbols for channel estimation. Define a time epoch as a group of L_s symbols, which is also the distance between two pilot symbols. The modulation format for every time epoch will be kept the same, and is determined by the CSI of the pilot symbol. We define λ as the ratio between the number of information symbols and the total number of symbols. In most of this paper, we will focus on a system with independently faded subcarriers, i.e., $\Psi = 1$, so that, $\lambda = (L_s - 1)/L_s$. In a system with coherence bandwidth $\Psi > 1$, since all the subcarriers within the coherence bandwidth will have the same fade realization, fewer pilot symbols need to be sent. For example, in Fig. 2, two subcarriers share the same fade ($\Psi = 2$), and one pilot symbol can be saved in each time epoch for every group of two subcarriers spanned by the same coherence bandwidth.

Let $P_{k,m}$ be the power for the k -th user on the m -th subcarrier, where $P_{k,m} = P_k > 0$ for the duration of GOP if the m -th subcarrier is allocated to the k -th user, and $P_{k,m} = 0$ otherwise. We assume open loop power control, and P_k is chosen such that the average received power per subcarrier at the base station is the same for all users. Let $\mu_{k,m}$ be the ratio between the power of the pilot symbols and average data power for the k -th user on the m -th subcarrier. The power of the pilot symbol, $P_{k,m}^p$, and the data power, $P_{k,m}^d$, are then given by

$$P_{k,m}^d = \frac{P_{k,m} L_s}{L_s - 1 + \mu_{k,m}} \quad (9)$$

$$P_{k,m}^p = \frac{P_{k,m} L_s \mu_{k,m}}{L_s - 1 + \mu_{k,m}} \quad (10)$$

We assume that the value of $\mu_{k,m}$ is the same for all subcarriers of all users, and hence we drop the indices k and m . The average power is given by

$$P_{k,m} = \mu \frac{1}{L_s} P_{k,m}^p + \frac{L_s - 1}{L_s} P_{k,m}^d \quad (11)$$

For a given SER_t , the modulation format for subcarrier m of user k is updated every L_s symbols based on $\tilde{H}_{k,m}[iL_s + 1]$, which is the estimate of the channel response of the i -th pilot, $H_{k,m}[iL_s + 1]$. The information rate (number of bits each symbol can carry in bits/symbol) for the i -th group of L_s symbols can be written as $R_{k,m}(P_{k,m}^d, \tilde{H}_{k,m}[iL_s + 1])$ using (7). To estimate the channel response $H_{k,m}[iL_s + 1]$, a Wiener filter with K_e pilots is used for interpolation. Since the decision of the modulation format needs to be fed back to the users immediately after the pilot symbol is sent, we can only use the pilots prior to the current one for channel estimation. In other words, to estimate $H_{k,m}[iL_s + 1]$, the pilots at indices $l = jL_s + 1$, $\{j = (i, i - 1 \dots (i - K_e + 1))\}$ are used, where K_e is chosen to be even. To estimate the channel gain of the data symbols $H_{k,m}[iL_s + u]$, $u = (2, 3 \dots L_s)$ for demodulation purposes, pilots from both sides can be used. That is, the pilots at time indices $l = jL_s + 1$, $\{j = ((i - K_e/2 + 1), (i - K_e/2 + 2) \dots (i + K_e/2))\}$ are jointly used to interpolate the channel gain.

From [6], [28], the channel estimation error $e = \tilde{H}_{k,m}[l] - H_{k,m}[l]$ can be modeled as a Gaussian random variable with zero mean and variance equal to

$$\sigma_e^2 = \sigma_u^2 - \mathbf{w}(\mathbf{l})^+ \mathbf{R}^{-1} \mathbf{w}(\mathbf{l}) / \left(P_{k,m}^p \overline{|H_{k,m}|^2} \right) \quad (12)$$

where $\sigma_u^2 = P_{k,m}^x \overline{|H_{k,m}|^2}$ is the average received power of the data/pilot symbols, and $+$ represents conjugate transpose. If we use K_e pilots for interpolation, \mathbf{R} is a $K_e \times K_e$ matrix with the entry in the i -th row and j -th column given by

$$R_{ij} = P_N \delta_{ij} + P_{k,m}^p \overline{|H_{k,m}|^2} J_0(2\pi f_{nd}(i - j)L_s) \quad (13)$$

where δ is the Kronecker delta, P_N is the noise power, and $\mathbf{w}(\mathbf{l})$ is a K_e column vector for the l -th channel sample. The v -th row of $\mathbf{w}(\mathbf{l})$ is given by

$$w_v(l) = P_{k,m}^p \overline{|H_{k,m}|^2} \rho_v(l) \quad (14)$$

where $\rho_v(l)$ is the correlation coefficient of the l -th channel sample and the channel estimate obtained from the v -th pilot. For example, as discussed above, to estimate the channel sample $H_{k,m}[l]$, if the symbol at time l is a data symbol and belongs to the i -th pilot time epoch, i.e., $iL_s + 2 \leq l \leq (i + 1)L_s$, the v -th pilot used for interpolation is at $(i - K_e/2 + v)L_s + 1$, so that

$$\rho_v(l) = J_0(2\pi f_{nd}[l - ((i - K_e/2 + v)L_s + 1)]) \quad (15)$$

From (12), (13), and (14), we see that one of the crucial parameters for deciding the channel estimation error is the distance between pilots, L_s . For a given Doppler spread, a smaller L_s results in a larger overhead, and a larger L_s results in larger channel mismatch due to CSI outdated, but less throughput loss. Besides L_s , the variance of e is also dependent on the number of pilot symbols for interpolation K_e , the average SNR, and the ratio μ between the pilot symbol power and data symbol power.

C. Resource Allocation Problem Formulation

At the beginning of a GOP, the base station estimates the throughput of each subcarrier using (7), assuming that the adaptive QAM format will last for T_s seconds. The number of bits transmitted over subcarrier m of user k can then be written as $R_{k,m}(P_{k,m}^d, \tilde{H}_{k,m}^e) \cdot T_s/T_0$, where T_s/T_0 is the number of QAM symbols for a GOP. We denote by λ ($0 < \lambda \leq 1$) the fraction of data symbols, and $\lfloor (1-\lambda) \cdot T_s/T_0 \rfloor$ symbols will be pilots. To protect the data, a channel code of fixed rate u is added. If the channel stays constant, the number of information bits that the physical layer can support for user k across all M_c subcarriers is given by

$$B_k = \left\lfloor \sum_{m=1}^{M_c} u \cdot \lambda \cdot R_{k,m} \left(P_{k,m}^d, \tilde{H}_{k,m}^e \right) \cdot T_s/T_0 \right\rfloor \quad (16)$$

with $P_{k,m}^d = P_k^d$, if subcarrier m is assigned to user k , and $P_{k,m}^d = 0$ otherwise.

Although to obtain the performance of the resource allocation algorithm we will include the effects of channel errors and time varying modulation choices, for the allocation algorithm design, we ignore the effect of channel errors and assume that the modulation format is constant for the GOP duration. We use (16) as the channel throughput for our algorithm design. If we plug (16) into (8), then the MSE distortion for user k can be written as

$$D_k = a_k + \frac{b_k}{\sum_{m=1}^{M_c} \lambda R_{k,m} \left(P_{k,m}^d, \tilde{H}_{k,m}^e \right) + c_k} \quad (17)$$

Here, we have divided both the numerator and denominator by $u \cdot T_s/T_0$ for simplicity. So

$$b_k = \frac{w_k}{(u \cdot T_s/T_0)} \quad c_k = \frac{v_k}{(u \cdot T_s/T_0)} \quad (18)$$

The base station needs to assign M_c subcarriers to K users at the beginning of each GOP, and users can access the subcarriers for the duration of the GOP. The allocation decision will be updated at the beginning of the next GOP as both CSI and RD are updated. Mathematically, our resource allocation goal is to minimize the sum of distortions among K users at each time slot. The optimization objective is

$$\min_{\underline{P}} \sum_{k=1}^K \frac{b_k}{\sum_{m=1}^{M_c} \lambda R_{k,m} \left(P_{k,m}^d, \tilde{H}_{k,m}^e \right) + c_k} \quad (19)$$

where

$$\underline{P} = \begin{bmatrix} P_{1,1}^d & P_{1,2}^d & \cdots & P_{1,M}^d \\ P_{2,1}^d & P_{2,2}^d & \cdots & P_{2,M}^d \\ \cdots & \cdots & P_{k,m}^d & \cdots \\ P_{K,1}^d & P_{K,2}^d & \cdots & P_{K,M}^d \end{bmatrix} \quad (20)$$

and the entry in the k -th row and m -th column, $P_{k,m}^d$, is the power allocation of the m -th subcarrier for user k . The base station sends the allocation decision to the users. Note that we dropped the a_k term, as it is constant with respect to \underline{P} . We assume that

any subcarrier is used by one user exclusively, so the feasibility constraint for the optimization problem is as follows:

For $m \in \{1, 2, 3 \dots M_c\}$, $P_{k',m}^d P_{k,m}^d = 0$, $\forall k \neq k'$ and $P_{k,m}^d = \{0, P_k^d\}$.

Mathematically, (19) is an NP-hard integer programming problem [29], and an exhaustive search approach would need K^{M_c} calculations. In the next section, we will propose a sub-optimal algorithm which gives priority to users with steep RD curvatures to access the subcarriers. We compare the performance of our algorithm with two baseline algorithms, each of which has limited information about the state of the channel and the state of the videos. Note that, compared to [14], the problem formulation in (19) includes the ratio λ between the numbers of data symbols and total symbols. Given a fixed power $P_{k,m}$, the throughput of each subcarrier is also controlled by the power ratio between the pilot symbols and data symbols, $\mu_{k,m}$. For a system with relatively high Doppler spread, the end-to-end video performance is closely related to the tradeoff between the channel estimation accuracy and data rate. A more frequent pilot insertion and larger pilot power will improve the channel estimation accuracy at the expense of sacrificing data rate. This tradeoff was not a key issue in [14], but will be discussed in depth in Section V after we introduce the resource allocation algorithms in Section IV.

IV. RESOURCE ALLOCATION ALGORITHMS

Before we state the cross layer algorithm that solves the optimization problem (19), we introduce two baseline algorithms, which each use only one layer of information when allocating resources in Phase I. The three algorithms are similar to those in [14], so we omit most of the details. After resource allocation, the channel adaptation mechanism is the same for all three resource allocation algorithms in Phase II.

Application Layer Resource Allocation Algorithm: We assume the base station knows only the application layer information (RD function) when allocating. Define L_k , ($k = \{1, 2 \dots K\}$) as the number of subcarriers allocated to user k , which is determined by the relative complexity of the RD functions $D_k(B)$. To measure GOP relative complexity, we set a common video MSE distortion target D_{target} . Let B_k be the number of bits for user k to achieve D_{target} , where B_k can be found from $D_k(B)$. Given the constraint of $\sum_{k=1}^N L_k = M_c$, the subcarriers will be split such that $L_k \approx M_c \frac{B_k}{\sum_{k=1}^K B_k}$. Since CSI is

not used, the base station randomly chooses L_k subcarriers for user k . After allocation, the user applies Phase II to update the modulation format using (7). For the application layer system, since the mechanism of Phase II, which includes the choice of the system parameters (L_s , μ , etc.), is the same as for the cross layer system, the performance difference between the two systems depends solely on the resource allocation.

Physical Layer Resource Allocation Algorithm: Here only the instantaneous CSI $\tilde{H}_{k,m}^e$ is available at the base station. Past work has attempted to balance allocation efficiency and fairness [30]. In this paper, similar to [14], we are interested in the

metric of channel gain ratio $|\tilde{H}_{k,m}^e|^2/|\tilde{H}_k^e|^2$, where $|\tilde{H}_{k,m}^e|^2$ is the instantaneous channel gain of subcarrier m for user k and $|\tilde{H}_k^e|^2 \triangleq \frac{1}{M_c} \sum_{m=1}^{M_c} |\tilde{H}_{k,m}^e|^2$ is the empirical average channel gain of user k at the beginning of the GOP. To control the degree of imbalance in the number of subcarriers that users receive, we impose a set of thresholds ψ_n , $n = 1, 2, \dots, K-1$, such that the sum of the number of subcarriers for any group of n users will not exceed ψ_n . We set $\psi_1 = \lceil \epsilon \frac{M_c}{K} \rceil$ and

$$\psi_n = \psi_{n-1} + \left\lceil \epsilon \left(\frac{M_c - \psi_{n-1}}{K - (n-1)} \right) \right\rceil \quad (21)$$

for $2 \leq n \leq K-1$. The parameter $\epsilon \geq 1$ controls allocation imbalance; larger values of ϵ allow more unbalanced resource allocation, biased towards users with better channels.

To assign subcarriers, we find the subcarrier m^* which has the best channel gain ratio among all users and subcarriers, and assign subcarrier m^* to user k^* for whom channel gain ratio is largest. That subcarrier is removed from consideration, and the next highest channel gain ratio is found. During this sequential process, we check if the total number of subcarriers assigned to any group of n users equals ψ_n , and exclude all users in the group from receiving more resources if it does. Details about the parameter ϵ and algorithm steps are in [14].

Cross Layer Resource Allocation Algorithm: When both RD information and CSI are used, a cross layer algorithm attempts to satisfy the two goals of giving more subcarriers to users with demanding RD curves and to those with high channel gains. With the consideration of λ and $\mu_{k,m}$ in the problem formulation in (19), we use the same resource allocation algorithm as [14]. The algorithm first assigns subcarriers using a MUD algorithm and estimates the number of bits that the channel can support using (16). The algorithm then tests possible reassignments of subcarriers to the user with the steepest slope of his/her RD curve, and uses the CSI to find the subcarrier which maximally reduces the sum of distortions through reassignment. After reassignment, we update the user with the steepest slope and continue the iteration until we exhaust all possibilities for subcarrier switching.

V. SYSTEM PARAMETER OPTIMIZATION

We use a video with a resolution of 352×240 . The video consists of 150 frames at 30 frames/second, and is organized into GOPs of 15 frames (IPPP). The content of the video includes both high motion segments and low motion segments. Each user is assigned the same video, but with random starting points. The simulation runs for one cycle of the entire video sequence, and users are then assigned another random set of starting points for the next cycle. By assigning random starting points of the same cyclic video to different users, we create instantaneous application layer diversity among users and yet have the same average complexity over time for the different users. Between cycles, different realizations of starting points will generate different levels of application diversity. We encode the video using H.264/SVC reference software JSVM version 9.19.12. For the MGS layer, the 4×4 DCT coefficients

of each macroblock are split using MGS vector [1, 1, 2, 2, 2, 8] [24]. To specify the a_k , b_k , and c_k values of RD curves in (17), we extract bitstreams at 20 different encoding rates from 60 to 1200 kbps offline and use these operational points to find the RD function by non-linear regression. At the decoder side, the bit stream after the first channel error is discarded. In a very crowded system, it is possible that some users will not be allocated any subcarriers and the transmission rate for the GOP is zero. If that happens, the last frame of the previous GOP is held over for the duration of the current GOP. The video performance is reported using Peak Signal to Noise Ratio (PSNR) defined as

$$PSNR = 10 \log_{10} \frac{255^2}{\mathbb{E}[MSE]} \quad (22)$$

We consider a single cell of radius equal to 500 meters. The bandwidth for each subcarrier is 100 kHz. The channel response consists of both path loss and multipath fading, and the amplitude squared of the multiplicative channel coefficient $H_{k,m}[l]$ is given by

$$|H_{k,m}[l]|^2 = |\alpha_{k,m}[l]|^2 \cdot K_0 \cdot \left(\frac{d_0}{d_k} \right)^\beta \quad (23)$$

where $\alpha_{k,m}[l]$ follows the Jakes' model and is generated using the statistical model proposed by [31]. Also, d_k is the distance of user k to the base station, $d_0 = 10$ m is the reference distance, and the path-loss model is accurate when $d_k > d_0$ [32]. The users are perfectly power controlled and the average received power is 17.8 dB per subcarrier. We set the path-loss exponent $\beta = 3$, and $K_0 = -30$ dB is a constant. For all three optimization schemes, we apply a rate $u = 1/2$ convolutional code with code generator polynomial [23, 35] in octal. The codeword length is equal to the length of the entire GOP bitstream, and the coded bits are interleaved across different subcarriers. To give the system an appropriate level of protection at low Doppler spreads, the value of SER_t was set to 0.15, after having tested values both lower and higher than 0.15. The PSNR performance will degrade as the Doppler spread increases and channel estimation accuracy becomes worse. At the receiver, we use soft-decision decoding with eight reliability ranges. For all three resource allocation algorithms, the users are allowed to choose $M = \{4, 8, 16, 32, 64, 128, 256\}$ as their MQAM modulation alphabet size every L_s symbols. For the physical layer optimization algorithm, we set $\epsilon = 1.5$, so that each individual user cannot be assigned more than 150% of the average number of subcarriers ($\psi_1 = M_c/K$) defined in (21) for ψ_1 .

A. System With Different L_s

In Fig. 3, we show the performance of the three resource allocation algorithms with respect to different Doppler spreads. The pilot insertion spacing L_s equals 100, the number of pilots used for channel estimation K_e is 24, and the ratio μ between pilot and data power equals unity. The PSNR performance here consists of both the degradation caused by source compression and that caused by channel errors. The effects of packet loss, errors in RD curve fitting, and imperfection of encoder rate control, are included in the simulation.

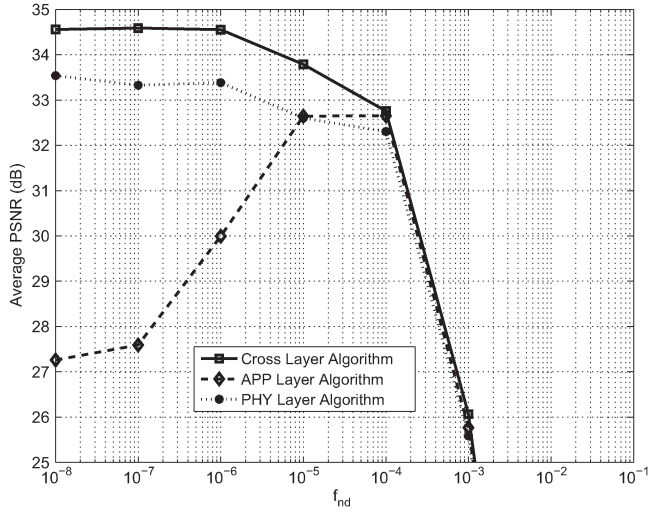


Fig. 3. $L_s = 100$, $M_c = 16$ subcarriers, $K = 3$ users, $SER_t = 0.15$, $\mu = 1$. The PSNR performance for a system with $f_{nd} > 10^{-2}$ is less than 15 dB.

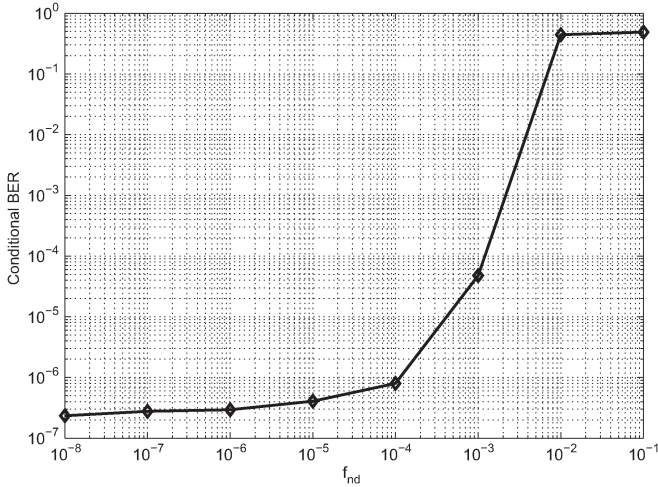


Fig. 4. Decoded BER vs f_{nd} , $L_s = 100$, $SER_t = 0.15$.

In Fig. 4, we show the decoded BER, conditioned on the event that the user transmits, i.e., $R_{k,m}(p_{k,m}^d, \tilde{H}_{k,m}[iL_s + 1]) \geq 2$, where the number “2” corresponds to the number of bits per symbol of 4QAM. When $1/f_{nd}$ is significantly larger than $L_s = 100$, the channel is relatively constant within each duration of L_s symbols. The decoded bit error rate for this scenario is in the range of 10^{-8} to 10^{-6} . Since we update the modulation format every L_s symbols, the correlation between the data symbols and pilot symbols is relatively high, and the channel estimation error given in (12) is small. The actual raw SER for every group of 100 symbols is similar to the SER of an AWGN channel conditioned on the instantaneous channel fade, and is close to the SER estimated using (6). In other words, for $L_s = 100$ and $f_{nd} < 10^{-4}$ in Fig. 3, the proposed mechanism of PSAM (Phase II) can accurately adapt to the variation of the channel and properly control the channel error rate.

In Fig. 3, when $f_{nd} = 10^{-3}$, we see a large performance drop for all three algorithms. Since the channel estimates become increasingly outdated, the modulation format chosen based upon the CSI becomes increasingly meaningless. We see a significant increase of the decoder BER when f_{nd} reaches 10^{-3} ,

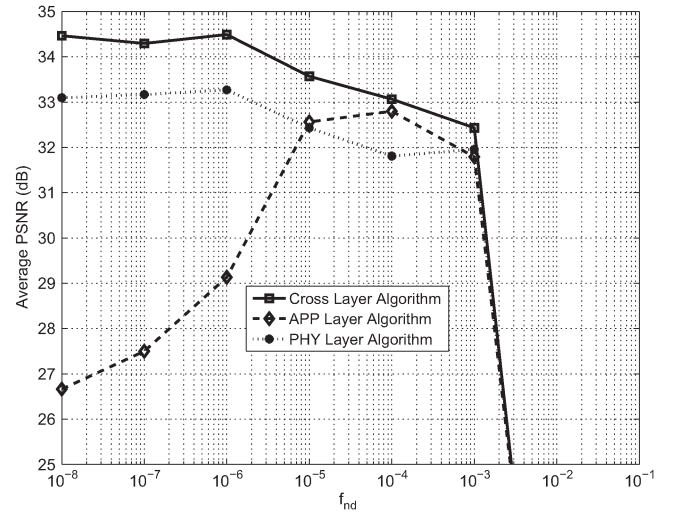


Fig. 5. $L_s = 25$, $M_c = 16$ subcarriers, $K = 3$ users, $SER_t = 0.15$, $\mu = 1$.

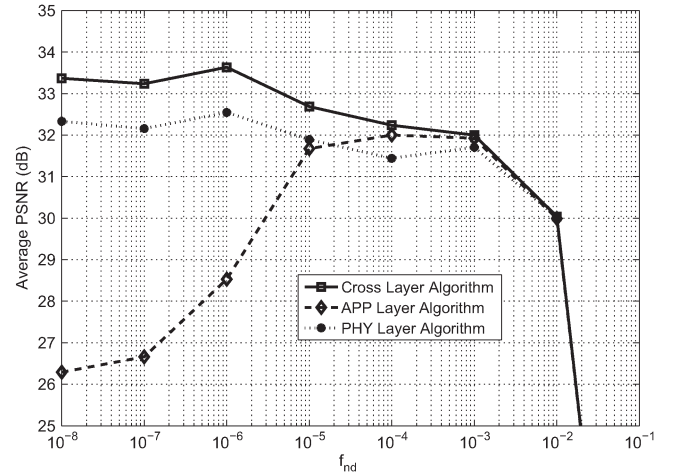


Fig. 6. $L_s = 5$, $M_c = 16$ subcarriers, $K = 3$ users, $SER_t = 0.15$, $\mu = 1$.

and the decoded PSNR is much worse than that at $f_{nd} = 10^{-4}$. When the normalized Doppler spread is 10^{-2} , the decoded BER is too large for the system to function.

In Fig. 5, we decrease the value of L_s to 25. Recall that the fraction of data is $\lambda = (L_s - 1)/L_s$. Decreasing the value of L_s will increase the amount of overhead (pilot symbols) used for channel estimation. However, decreasing the value of L_s will help the system to achieve better estimation accuracy at high Doppler. Compared to a system with $L_s = 100$ in Fig. 3, we see that a pilot spacing $L_s = 25$ has better performance when $f_{nd} = 10^{-3}$, despite the drop of the source rate. If we further decrease the value of L_s to 5 as in Fig. 6, the loss of the source data further increases to 20%, but the benefits of accurate channel estimation and modulation adaptation allow the system to operate at a reasonable PSNR value at $f_{nd} = 10^{-2}$.

B. Comparison of the Three Algorithms

Comparing the systems with different L_s , in Figs. 2, 4, and 5, we see that the performance follows a very similar trend. For both the physical layer and cross layer algorithms, we see that PSNR decreases when f_{nd} increases from 10^{-6} to 10^{-3} .

Since we use instantaneous CSI $\tilde{H}_{k,m}^e$ for resource allocation, the multiuser diversity will allow both physical layer and cross layer optimization algorithms to assign subcarriers that are, with high probability, experiencing a strong channel compared to the average channel gain. In other words, if a subcarrier m^* is assigned to a user k^* , the instantaneous CSI \tilde{H}_{k^*,m^*}^e at the beginning of the GOP has a high probability of being better than the average CSI. If the coherence time of the channel is long, the CSI will be relatively constant over the duration of the GOP. In this particular system, if f_{nd} is in the range of 10^{-8} to 10^{-6} , the channel is not varying much and the modulation format changes little over a GOP. As f_{nd} increases, at some point, we expect to see different fades of the channel over a GOP. The average of these different fades with high probability will be worse than the CSI at the beginning of the GOP. In particular, the initial CSI will be outdated very soon for systems with f_{nd} in the vicinity of 10^{-3} to 10^{-1} .

In Fig. 3, for the application layer algorithm, the performance at high Doppler spreads is similar to that of the other two algorithms. However, the performance at low Doppler spreads is markedly different. Because the application layer algorithm does not use physical layer information, it does not benefit from the multiuser diversity as the other two algorithms do. Therefore, at low Doppler spreads, the application layer algorithm has virtually no diversity, and hence exhibits poor performance. That performance improves as the Doppler spread increases up to the point where all three algorithms fail due to channel outdating.

Comparing the application layer and physical layer optimization algorithms, we see that the physical layer algorithm has better performance when f_{nd} is small, and the application layer algorithm wins when f_{nd} increases. That is, it is more important to exploit the multiuser channel diversity and allocate the resources based on the channel realization in a slowly varying environment. As the channel will stay relatively constant for a GOP, the resource allocation made based on the CSI at the beginning of the GOP will be meaningful over the entire GOP. On the other hand, when the coherence time is sufficiently small, the CSI will become outdated and the throughput of each subcarrier will be very similar regardless of the initial state of the channel. Thus, when f_{nd} is high, it is more important to allocate resources based on RD information, and the application layer algorithm has better performance. Note that the cross layer optimization algorithm is robust with respect to different normalized Doppler spreads and always outperforms the two baselines, until the Doppler spread is too large for the system to track the fades, at which point all three algorithms fail.

C. Systems With Different Resources

We now study the performance for different parameter values. In Fig. 7, we increase the number of users to 4. In Fig. 8(a) and (b), we show the performance of a system of 24 subcarriers with 3 and 4 users, respectively. Comparing Figs. 3, 7, 8(a) and (b), we see that the system with 24 subcarriers and 3 users not only has the best performance, but also has the largest gap between the cross layer and physical layer optimization algorithms. This is because when the system has 24 subcarriers

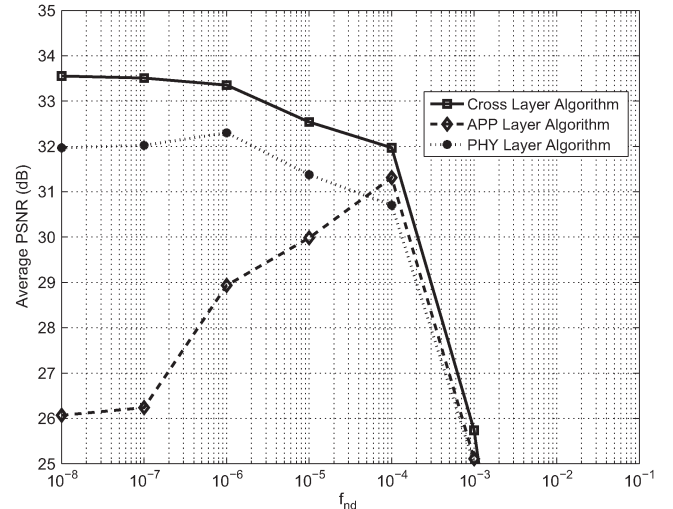


Fig. 7. $L_s = 100$, $M_c = 16$ subcarriers, $K = 4$ users, $SER_t = 0.15$, $\mu = 1$.

and 3 users, the cross layer algorithm has more degrees of freedom to allocate the resources among the users.

In Table I, we show a typical example of the performance evolution of a three-user system with 16 subcarriers and $f_{nd} = 10^{-6}$. The resource allocation is done based on the estimate of the first sample of the GOP, and we show the change of the estimated MSE for each individual user at each step of the iteration. Among the three users, the first and second users have demanding RD curves and the corresponding b_k of the RD information is much larger than that of the third user. For the cross layer algorithm initialization, the average PSNR is equal to 32.48 dB, and subcarriers are assigned evenly to all the users. The algorithm converges in 3 steps (3 iterations) and the final subcarrier allocation is 6 and 7 for the two demanding users and 3 for the third user. The average PSNR value of the cross layer algorithm is 33.16 dB. In this case, the cross layer optimization improves the initial performance of the system by only 0.7 dB.

In Table II, we study a system with 24 subcarriers and 3 users. To compare with the previous example fairly, the channel realizations for 16 out of the 24 subcarriers are the same as the realizations of the previous example. Realizations for the remaining 8 subcarriers are generated using the same mechanism as for the first 16 subcarriers. For the 24-subcarrier system, the initialization of the allocation is (7, 8, 9) subcarriers for the three users, and the corresponding average PSNR is 34.82 dB. It takes the cross layer algorithm 5 switches of subcarriers, and the cross layer algorithm assigns (9, 10, 5) subcarriers to the three users. The resulting average PSNR performance is 36.17 dB, and so the cross layer algorithm improves the performance by 1.35 dB. Going from 16 to 24 subcarriers, we see that it takes the cross layer algorithm more steps to converge and the PSNR improvement is larger.

D. Systems With Different μ

Besides lowering the value of L_s , another way of achieving accurate channel estimation is to allocate more power to pilot symbols. In Fig. 9, we show the system performance versus the pilot-to-data power ratio μ for the cross layer algorithm

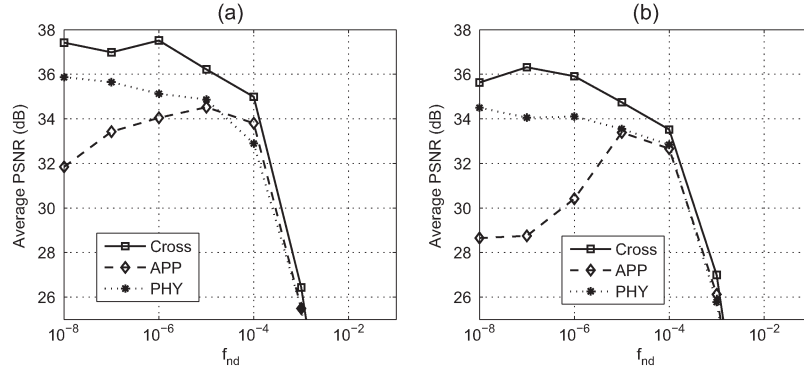


Fig. 8. $L_s = 100$, $M_c = 24$ subcarriers, $SER_t = 0.15$, $\mu = 1$; (a) $K = 3$ users, (b) $K = 4$ users.

TABLE I
EVOLUTION OF PERFORMANCE: $M_c = 16$ SUBCARRIERS, $U1 = \text{User 1}$, $U2 = \text{User 2}$, $U3 = \text{User 3}$. $N_{\text{Sub}} = \# \text{ Subcarriers Allocated}$

	U1 MSE	U1 NSub	U2 MSE	U2 NSub	U3 MSE	U3 NSub	Avg. MSE	PSNR
Initialization	53.68	5	56.52	5	~ 0	6	36.73	32.48
Step 1	53.68	5	45.89	6	~ 0	5	33.19	32.97
Step 2	44.82	6	45.89	6	4.66	4	31.79	33.10
Step 3	44.82	6	40.35	7	9.08	3	31.41	33.16

TABLE II
EVOLUTION OF PERFORMANCE: $M_c = 24$ SUBCARRIERS, $U1 = \text{User 1}$, $U2 = \text{User 2}$, $U3 = \text{User 3}$. $N_{\text{Sub}} = \# \text{ Subcarriers Allocated}$

	U1 MSE	U1 NSub	U2 MSE	U2 NSub	U3 MSE	U3 NSub	Avg. MSE	PSNR
Initialization	34.38	7	29.94	8	~ 0	9	21.43	34.82
Step 1	29.41	8	29.94	8	~ 0	8	19.78	35.17
Step 2	25.27	9	29.94	8	~ 0	7	18.40	35.48
Step 3	25.27	9	25.31	9	~ 0	6	16.86	35.86
Step 4	23.11	10	25.31	9	~ 0	5	16.14	36.05
Step 5	23.80	9	23.31	10	~ 0	5	15.71	36.17

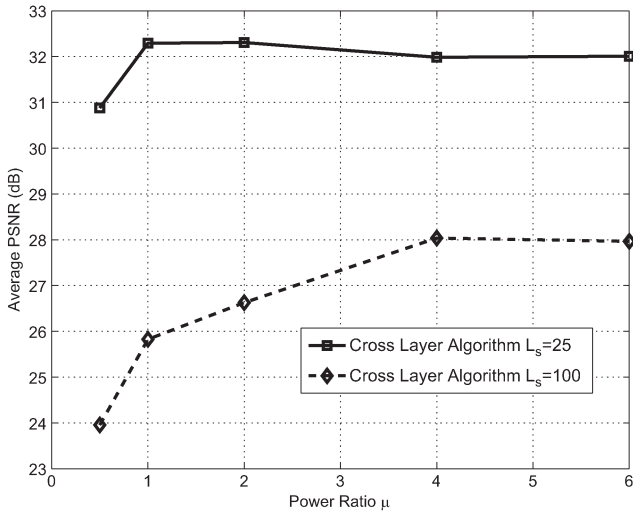


Fig. 9. $M_c = 16$ subcarriers, $K = 3$ users, $SER_t = 0.15$, $f_{nd} = 10^{-3}$.

when $f_{nd} = 10^{-3}$. As expected, for small values of μ , the pilot power is too small, and the PSNR is correspondingly small. As μ increases, the PSNR increases up to a point where it levels off, and eventually (as μ continues to increase) it will decrease, because the data symbols have too little power. Note that the $L_s = 25$ curve performs better than the $L_s = 100$ curve for all values of μ , because the normalized Doppler spread of 10^{-3} results in too much outdating when pilots are spaced every 100 symbols apart.

E. Systems With Different Numbers of Users

We now study the system performance when the number of users increases. We fix the number of subcarriers to 16 and change the number of users in the system from 3 to 9. In Fig. 10(a), we show the average performance of all users when $f_{nd} = 10^{-6}$ and $L_s = 100$. We see that in this case, when f_{nd} is small, the PSNR gain of using the cross layer scheme is very large compared to the application layer scheme. With 9 users, the users will normally get a small number of subcarriers. For the application layer algorithm, the subcarriers are randomly assigned, and because the number of subcarriers for each user is small or zero, it is very likely that at least one user will be in a very bad situation. The user will thus have to function with, at best, a low rate, and this has a large impact on the average distortion of the group.

Compared to the physical layer algorithm, the PSNR gain of the cross layer scheme is always more than 1 dB. Alternatively, rather than fixing the number of users, if we fix the average PSNR at, say, 30 dB, the capacity gain of the cross layer algorithm is a factor of three times that of the application layer algorithm, and about a factor of 1.5 times that of the physical layer algorithm.

In Fig. 10(b), we let $f_{nd} = 10^{-3}$ and $L_s = 25$. Because of the benefit of time diversity, we see that the gap between the cross layer and application layer schemes is less than 0.5 dB for a system of 3 users. However, the performance of the application layer algorithm decreases sharply when the number of users

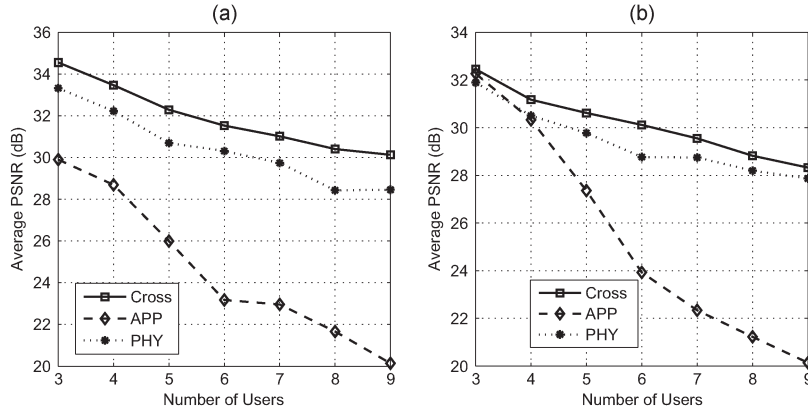


Fig. 10. $M_c = 16$ subcarriers, $SER_t = 0.15$, $\mu = 1$: (a) $L_s = 100$, $f_{nd} = 10^{-6}$, (b) $L_s = 25$, $f_{nd} = 10^{-3}$.

increases, because some of the less demanding users in the system will not be allocated any subcarriers. The performance of these users will be determined by holding the last frame of the previous GOP, resulting in a low average PSNR value. The gap between the cross layer and the physical layer algorithms ranges from 0.6 to 1.3 dB, which is smaller than the corresponding gap in Fig. 10(a). This is because when f_{nd} increases, the CSI used for resource allocation is outdated more quickly, and this affects the ability of the cross layer algorithm to balance the resources among the users. In other words, at high f_{nd} , the effect of the CSI of the first symbol of the GOP has less impact than that for a slow fading system.

For a fixed average PSNR performance of about 29 dB, the physical layer algorithm can support 6 users. The same average performance can be achieved for an 8-user system if the cross layer algorithm is adopted, while the application layer algorithm can only support less than 5 users. In this sense, the capacity gain of the cross layer algorithm for a relatively fast fading system is smaller than that for a slow fading system.

F. Comparison With the Exhaustive Search Algorithm

We now compare the performance of the iterative algorithm with an exhaustive search algorithm. In Fig. 11, we plot the average MSE distortion against the number of iterations for a system with $f_{nd} = 10^{-6}$. We see that the algorithm can very effectively reduce the sum of the distortions among the users in a couple of iterations, and the distortion diminishes as the number of iterations increases. For the example shown in Fig. 11, the algorithm converges when the number of iterations reaches 6.

In Table III, we show a performance comparison between the iterative algorithm and an exhaustive search for the case of 16 subcarriers and 3 users for the normalized Doppler spread varying from $f_{nd} = 10^{-8}$ to $f_{nd} = 10^{-4}$. The exhaustive search will calculate 3^{16} different subcarrier assignment combinations, while the iterative algorithm normally will converge in less than 8 iterations. For each iteration, the iterative algorithm will search at most 16 possible subcarrier switches, so the total number of subcarrier assignment combinations the iterative algorithm will consider is normally less than $8 \times 16 = 128$. As

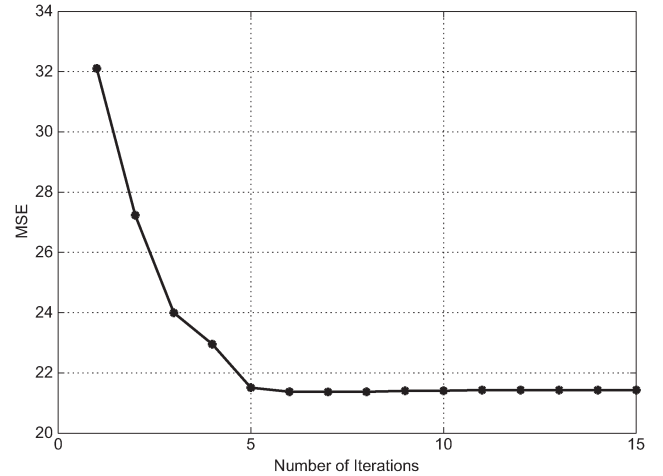


Fig. 11. MSE vs number of iterations: $L_s = 100$, $M_c = 16$ subcarriers, $K = 3$ users, $f_{nd} = 10^{-6}$.

TABLE III
PSNR PERFORMANCE COMPARISON, ITERATIVE
ALGORITHM VS. EXHAUSTIVE SEARCH

	10^{-8}	10^{-7}	10^{-6}	10^{-5}	10^{-4}
Exhaustive Search	34.6507	34.6945	34.7468	33.9645	33.2068
Iterative Algorithm	34.6185	34.6199	34.6686	33.9621	33.1483

illustrated in the table, the performance gap between the two algorithms is always less than 0.08 dB, while the complexity of the algorithms differs by orders of magnitude.

VI. CONCLUSION

In this paper, we study a multicarrier uplink video communication system over a channel with arbitrary normalized Doppler spread. We use both the application layer RD information and the physical layer CSI to allocate subcarriers to video users. After the resource allocator assigns the subcarriers to the users, each user continues to send pilot symbols and update the modulation format of each subcarrier based on feedback from the base station. The critical parameter L_s controls the tradeoff between the channel outdated and source rate. Our key results can be summarized as follows: for a system that has to

function well over a large range of Doppler spread, robustness in performance is arguably the key characteristic that it should exhibit. The cross layer design we presented in this paper satisfies this criterion. In particular, it was seen that if only a single layer design is employed, for some values of the Doppler spread the application layer algorithm was superior to the physical layer algorithm, with the opposite result holding for the other ranges of the Doppler spread. However, the cross layer algorithm outperformed both single layer algorithms over the range of Doppler spreads that allowed meaningful performance.

REFERENCES

- [1] A. Mohr, E. Riskin, and R. Ladner, "Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction," *IEEE J. Select. Areas Commun.*, vol. 18, no. 6, pp. 819–828, Jun. 2000.
- [2] Y. Chan, P. Cosman, and L. Milstein, "A multiple description coding and delivery scheme for motion-compensated fine granularity scalable video," *IEEE Trans. Image Process.*, vol. 17, no. 8, pp. 1353–1367, Aug. 2008.
- [3] G.-M. Su, Z. Han, M. Wu, and K. Liu, "Joint uplink and downlink optimization for real-time multiuser video streaming over WLANs," *IEEE J. Select. Areas Commun.*, vol. 1, no. 2, pp. 280–294, Aug. 2007.
- [4] M. Van der Schaar and D. Turaga, "Cross-layer packetization and retransmission strategies for delay-sensitive wireless multimedia transmission," *IEEE Trans. Multimedia*, vol. 9, no. 1, pp. 185–197, Jan. 2007.
- [5] Q. Liu, S. Zhou, and G. Giannakis, "Cross-layer combining of adaptive modulation and coding with truncated ARQ over wireless links," *IEEE Trans. Commun.*, vol. 3, no. 5, pp. 1746–1755, Sep. 2004.
- [6] X. Cai and G. Giannakis, "Adaptive PSAM accounting for channel estimation and prediction errors," *IEEE Trans. Wireless Commun.*, vol. 4, no. 1, pp. 246–256, Jan. 2005.
- [7] O. Oyman and S. Singh, "Quality of experience for HTTP adaptive streaming services," *IEEE Commun. Mag.*, vol. 50, no. 4, pp. 20–27, Apr. 2012.
- [8] Y. Liu, "Dynamic adaptive streaming over CCN: A caching and overhead analysis," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2013, pp. 1–5.
- [9] L. Toni, Y. Chan, P. Cosman, and L. Milstein, "Channel coding for progressive images in a 2-D time–frequency OFDM block with channel estimation errors," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2476–2490, Nov. 2009.
- [10] A. Khalek, C. Caramanis, and R. Heath, "A cross-layer design for perceptual optimization of H.264/SVC with unequal error protection," *IEEE J. Select. Areas Commun.*, vol. 30, no. 7, pp. 1157–1171, Jul. 2012.
- [11] Z. Shen, J. Andrews, and B. Evans, "Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints," *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, pp. 2726–2737, Nov. 2005.
- [12] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks-Part I: Theoretical framework," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 614–624, Mar. 2005.
- [13] G. Song and Y. Li, "Cross-layer optimization for OFDM wireless networks-Part II: Algorithm development," *IEEE Trans. Wireless Commun.*, vol. 4, no. 2, pp. 625–634, Mar. 2005.
- [14] D. Wang, L. Toni, P. Cosman, and L. Milstein, "Uplink resource management for multiuser OFDM video transmission systems: Analysis and algorithm design," *IEEE Trans. Commun.*, vol. 61, no. 5, pp. 2060–2073, May 2013.
- [15] A. Ortega, K. Ramchandran, and M. Vetterli, "Optimal trellis-based buffered compression and fast approximations," *IEEE Trans. Image Process.*, vol. 3, no. 1, pp. 26–40, Jan. 1994.
- [16] W. Jakes, *Microwave Mobile Communications*. New York, NY, USA: Wiley, 1974.
- [17] J.-J. van de Beek *et al.*, "A time and frequency synchronization scheme for multiuser OFDM," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 11, pp. 1900–1914, Nov. 1999.
- [18] S. Barbarossa, M. Pompili, and G. Giannakis, "Channel-independent synchronization of orthogonal frequency division multiple access systems," *IEEE J. Sel. Areas Commun.*, vol. 20, no. 2, pp. 474–486, Feb. 2002.
- [19] M.-O. Pun, M. Morelli, and C.-C. Kuo, "Maximum-likelihood synchronization and channel estimation for OFDMA uplink transmissions," *IEEE Trans. Commun.*, vol. 54, no. 4, pp. 726–736, Apr. 2006.
- [20] J. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2000.
- [21] Q. Qu, L. Milstein, and D. Vaman, "Cognitive radio based multi-user resource allocation in mobile ad hoc networks using multi-carrier CDMA modulation," *IEEE J. Select. Areas Commun.*, vol. 26, no. 1, pp. 70–82, Jan. 2008.
- [22] F. Wu, S. Li, and Y. Zhang, "A framework for efficient progressive fine granularity scalable video coding," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 11, no. 3, pp. 332–344, Mar. 2001.
- [23] M. van der Schaar and H. Radha, "Adaptive motion-compensation fine granular-scalability (AMC-FGS) for wireless video," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 12, no. 6, pp. 360–371, Jun. 2002.
- [24] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.
- [25] K. Stuhlmüller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Select. Areas Commun.*, vol. 18, no. 6, pp. 1012–1032, Jun. 2000.
- [26] G.-M. Su, Z. Han, M. Wu, and K. Liu, "A scalable multiuser framework for video over OFDM networks: Fairness and efficiency," *IEEE Trans. Circuits Syst. Video Tech.*, vol. 16, no. 10, pp. 1217–1231, Oct. 2006.
- [27] H. Ha, C. Yim, and Y. Y. Kim, "Cross-layer multiuser resource allocation for video communication over OFDM networks," *Comput. Commun.*, vol. 31, no. 15, pp. 3553–3563, Sep. 2008.
- [28] J. Cavers, "An analysis of pilot symbol assisted modulation for Rayleigh fading channels [mobile radio]," *IEEE Trans. Veh. Technol.*, vol. 40, no. 4, pp. 686–693, Nov. 1991.
- [29] R. Agarwal, V. R. Majjigi, Z. Han, R. Vannithamby, and J. M. Cioffi, "Low complexity resource allocation with opportunistic feedback over downlink OFDMA networks," *IEEE J. Select. Areas Commun.*, vol. 26, no. 8, pp. 1462–1472, Oct. 2008.
- [30] P. Viswanath, D. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1277–1294, Jun. 2002.
- [31] Y. R. Zheng and C. Xiao, "Simulation models with correct statistical properties for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 51, no. 6, pp. 920–928, Jun. 2003.
- [32] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge Univ. Press, 2005.



processing.



Dawei Wang (S'11) received the B.Eng. (with first-class honors) degree in electronic engineering from the Hong Kong University of Science and Technology, Kowloon, Hong Kong, in 2008 and the M.S. and Ph.D. degrees from the University of California at San Diego, La Jolla, CA, USA, in 2011 and 2013, respectively.

He is currently a Senior System Engineer with Adaptive Spectrum and Signal Alignment (ASSIA) Inc., Redwood City, CA. His research interests are in the areas of communication theory and video

Laura Toni (S'06–M'09) received the M.S. and Ph.D. degrees in electrical engineering from the University of Bologna, Bologna, Italy, in 2005 and 2009, respectively.

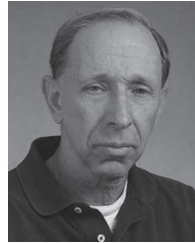
During 2007, she was a Visiting Scholar with the University of California at San Diego (UCSD), La Jolla, CA, USA, and since 2009, she has been a frequent visitor at UCSD, working on joint source and channel coding for wireless communication systems. From June 2009 to November 2011, she was with the Telerobotics and Application Department, Italian Institute of Technology, investigating wireless sensor networks for robotics applications. From November 2011 to November 2012, she was a Postdoctoral Fellow with the Department of Electrical and Computer Engineering, UCSD. Since December 2012, she has been a Postdoctoral Fellow with the Signal Processing Laboratory (LTS4), Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland. Her research mainly involves wireless communication systems, interactive streaming, multicamera networks, joint source and channel coding, underwater communications, and learning frameworks for online optimizations.



Pamela C. Cosman (S'88–M'93–SM'00–F'08) received the B.S.(Hons.) degree in electrical engineering from California Institute of Technology, Pasadena, CA, USA, in 1987 and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1993.

During 1993–1995, she was an NSF Postdoctoral Fellow at Stanford University and a Visiting Professor at the University of Minnesota. In 1995, she joined the faculty of the Department of Electrical and Computer Engineering, University of California, San Diego, CA, where she was the Director of the Center for Wireless Communications from 2006 to 2008 and is currently a Professor and an Associate Dean for Students of the Jacobs School of Engineering. Her research interests are in the areas of image and video compression and processing and wireless communications.

Dr. Cosman was the Technical Program Chair of the 1998 Information Theory Workshop in San Diego. She has been a member of the Technical Program Committee or the Organizing Committee for numerous conferences, including ICIP 2008–2011, QOMEX 2010–2012, ICME 2011–2013, VCIP 2010, PacketVideo 2007–2013, WPMC 2006, ICISP 2003, ACIVS 2002–2012, and ICC 2012. She was an Associate Editor of the *IEEE Communications Letters* (1998–2001), a Guest Editor of the June 2000 special issue of the *IEEE Journal on Selected Areas in Communications* on "Error-resilient image and video coding," an Associate Editor of the *IEEE Signal Processing Letters* (2001–2005), and the Editor-in-Chief (2006–2009) and a Senior Editor (2003–2005, 2010–2013) of the *IEEE Journal on Selected Areas in Communications*. Her awards include the ECE Departmental Graduate Teaching Award, a Career Award from the National Science Foundation, a Powell Faculty Fellowship, the Globecom 2008 Best Paper Award, and the HISB 2012 Best Poster Award. She is a member of Tau Beta Pi and Sigma Xi.



Laurence B. Milstein (S'66–M'68–SM'77–F'85) received the B.E.E. degree from The City College of New York, NY, USA, in 1964 and the M.S. and Ph.D. degrees in electrical engineering from the Polytechnic Institute of Brooklyn, Brooklyn, NY, in 1966 and 1968, respectively.

From 1968 to 1974, he was with the Space and Communications Group, Hughes Aircraft Company, and from 1974 to 1976, he was a member of the Department of Electrical and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY. Since 1976, he has been with the Department of Electrical and Computer Engineering, University of California at San Diego (UCSD), La Jolla, CA, USA, where he is a former Department Chair and currently the Ericsson Professor of Wireless Communications Access Techniques, working in the area of digital communication theory with special emphasis on spread-spectrum communication systems. He has been also a Consultant to both government and industry in the areas of radar and communications.

Dr. Milstein was the Vice President for Technical Affairs of the IEEE Communications Society in 1990 and 1991 and a former Chair of the IEEE Fellows Selection Committee. He has been an Associate Editor for communication theory of the *IEEE TRANSACTIONS ON COMMUNICATIONS*, an Associate Editor for book reviews of the *IEEE TRANSACTIONS ON INFORMATION THEORY*, an Associate Technical Editor of the *IEEE Communications Magazine*, and the Editor-in-Chief of the *IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS*. He was a recipient of the 1998 Military Communications Conference Long Term Technical Achievement Award, an Academic Senate 1999 UCSD Distinguished Teaching Award, an IEEE Third Millennium Medal in 2000, the 2000 IEEE Communications Society Armstrong Technical Achievement Award, and various prize paper awards, including the 2002 Fred Ellersick MILCOM Award. He was also the recipient of the IEEE Communication Theory Technical Committee (CTTC) Service Award in 2009 and the CTTC Achievement Award in 2012.