Journal Pre-proof

Video coding for OFDM systems with imperfect CSI: A hybrid digital-analog approach

Pradeepa Yahampath

PII: S0923-5965(20)30106-5
DOI: https://doi.org/10.1016/j.image.2020.115903
Reference: IMAGE 115903

To appear in: *Signal Processing: Image Communication*

Received date : 29 March 2019
Revised date : 15 April 2020
Accepted date : 7 June 2020

Please cite this article as: P. Yahampath, Video coding for OFDM systems with imperfect CSI: A hybrid digital-analog approach, *Signal Processing: Image Communication* (2020), doi: https://doi.org/10.1016/j.image.2020.115903.

Video Coding for OFDM Systems with Imperfect CSI: A Hybrid Digital-Analog Approach

Pradeepa Yahampath

University of Manitoba, Department of Electrical and Computer Engineering, Winnipeg, R3T 5V6, Canada

Abstract

Performance of an orthogonal frequency division multiplexing (OFDM) system is greatest when the exact channel state information (CSI) is used for transmitter rate control and power allocation. However, in real systems CSI can only be approximately known. Moreover, in video communication, it can be difficult to use any CSI for rate control of a video codec if the channel changes significantly during a group of pictures coded jointly, such as when the receiver is moving. We address this issue through a hybrid digital-analog (HDA) coding system where a standard video codec is used to generate a fixed-rate base layer upon which the analog quantization error is superimposed as a refinement layer. The system adapts to channel variations by proper transmit-power allocation between digital and analog components and across OFDM subcarriers, based on CSI. We present a power allocation scheme for this system which explicitly takes into account the imprecise nature of the available CSI. Experimental results obtained with simulated OFDM channel traces show that proposed scheme is able to achieve a much better quality-vs-reliability trade-off in video transmission, compared to the best known digital-only and analog-only alternatives.

Keywords:

Video coding, hybrid-digital analog coding, scalable video coding, OFDM, power allocation

1. Introduction

Discrete-time analog transmission of video signals in wireless systems, either as a part of an HDA coding scheme, or on its own as a purely analog scheme (commonly referred to as *uncoded* transmission) have been widely investigated [1–29]. The attractiveness of analog transmission is due to its inherent ability to adapt to the time-varying channel signal-to-noise ratio (CSNR) of wireless channels and reproduce at the receiver a video signal whose fidelity varies smoothly with the CSNR. In order to achieve the same result with a digital transmission, one must essentially "mimic" an analog transmission by using layered video compression, as well as complex channel adaptive coding and modulation schemes [30]. However, it is needless to emphasize that under fixed channel conditions, a well designed digital scheme is far superior to any practically realizable analog scheme, due to the possibility of better source and channel coding. High bandwidth-compression required in most real-time wireless video applications is difficult to achieve in analog form. Digital video coding techniques which achieve very high compression ratios with only a minor loss of quality are now common place [31–33]. The same is also true for channel coding. It is therefore not surprising that various hybrid methods have been investigated for effectively combining the natural channel-adaptation property of analog transmission with the coding advantage of digital transmission. In general, HDA transmission can be effective whenever fixed encoding has to be used on a time-varying channel [34].

This paper considers the application of HDA coding to improve the efficiency and the reliability of video unicast through orthogonal frequency division multiplexing (OFDM) [35] to mobile receivers when the available channel state information

(CSI) is only approximate. OFDM is widely used to tackle frequency-selective fading typical in mobile wireless systems. It is well known that maximizing the data rate over an OFDM system requires adaptive rate and power allocation to multiple subcarriers at the transmitter based on any available CSI. However, in reality video transmission to mobile receivers in CSI-optimized OFDM systems involves the following challenges.

1. In real systems, the transmitter adaptation as well as decoding has to be carried out based on noisy CSI rather than the exact values. First, CSI is usually estimated from pilot symbols and the channel noise in these symbols results in noisy estimates. Second, channel estimation may only be carried out at relatively longer intervals due to system limitations, even though the channel itself may be varying more rapidly. In such cases, the channel prediction is used to generate CSI for shorter intervals [36] and consequently the available CSI will include prediction errors. Third, the available CSI usually include quantization noise. The use of such CSI estimates for transmitter optimization, disregarding their inherent inaccuracies, will render a digital decoder susceptible to outages.

2. Adaptive transmission at the physical layer also requires adaptive video coding at the application layer. This may be difficult to implement. In particular, it is not possible to do so when video is being transmitted to a mobile receiver undergoing channel coherence times much shorter than the duration of a block of inter-frame coded video frames, referred to a group-of-pictures (GOP). For example, at a carrier frequency of 2.5 GHz, a mobile-receiver traveling at 100 km/h experiences a channel coherence time of $T_{\text{COH}} \approx 4$ ms [37]. The most efficient transmission therefore calls for video rate adaptation at comparable intervals. However, at a video frame rate of 30 Hz and a typical GOP size of 8 frames, the duration of a GOP is $T_{\text{GOP}} = 267$ ms, and therefore the rate adaptation is only feasible at such intervals. In these circumstances the quantization parameters (QPs) of the video codec for a given GOP will have to be chosen independent of actual CSI, for example based on predicted worst-case channel conditions. This results in inefficient use of transmission resources such as power and bandwidth.

The purely digital solution to the aforementioned CSI uncertainty problem is layered video coding combined with suitable power, coding, and modulation adaptation [38, 39]. Layered video coding allows encoding a GOP into a bit stream which contains a base layer and one or more refinement layers, with the property that while the base layer alone is sufficient for an acceptable reconstruction of the original video, the refinement layers, if available, can be used to progressively improve the base layer quality. However, multi-layer coding is not as efficient as single layer coding at the same bit rate. Furthermore, to obtain smooth video quality variations under varying channel conditions, many refinement layers may be required. More importantly, when encoding and decoding are based on imperfect CSI estimates, the refinement layers can become undecodable.

In contrast to previous work, this paper presents an new HDA coding solution which is shown to substantially improve the reliability and efficiency of video transmission over multi-carrier communication systems with low channel coherence times and noisy CSI. The motivation is to achieve the best "(video) quality-reliability" trade-off by using an efficient video codec to generate a fixed-rate base layer independent of CSI, and transmitting the quantization error along side in analog form as a single, infinite resolution refinement layer to efficiently utilize in a CSI-adaptive manner, the power and bandwidth made redundant by fixed-rate coding. The proposed scheme is based on power-splitting [34] wherein digital and analog quadrature amplitude modulated (QAM) symbols are superposed so that the same bandwidth is shared, but the total power is allocated, between the analog and digital transmissions. One the one hand, by spending more power on the digital transmission, the bit rate ate of the video codec and hence the video quality can be increased, but this accompanies an increase in the probability of complete outage that can result from using bit rates incompati-

ble with the actual channel state. On the other hand, spending more power on the analog quantization error reduces the allowable codec bit rate (hence the coding efficiency), but it also reduces the vulnerability to complete outages due to CSI errors. The best possible quality-reliability trade-off is thus achieved by optimal power allocation.

A key aspect of the proposed HDA scheme is that it decouples the video encoding problem and the CSI-based multicarrier transmission problem. In the video coding-phase, the encoder uses a GOP-size suitable for efficient inter-frame coding to generate the base layer at a bit rate determined by the average CSNR rather than any CSI that cannot be known exactly in advance. In the transmission-phase, the CSI-based optimal power allocation is performed at much shorter intervals (determined by the receiver's channel coherence time) compared to the GOP duration.

The main features of the proposed approach and the specific contribution this work are summarized below.

- Our ideal goal is to minimize the probability of base layer outage subject to a lower-bound on decoded video quality. Given the complex nature of the relationship between the QPs, the bit rate, and the distortion of state-of-the-art video codecs, this goal seems impossible to achieve. We therefore implicitly approach this goal by first choosing the video codec (base layer) bit rate to ensure a minimum acceptable decoded video quality, and allocating the minimum amount of power required for reliable transmission of the complete base layer within a GOP duration. Given that the CSI is unreliable, rather than using the available CSI directly, we use it through a noise model, making the power allocation solution robust against CSI errors. To be able to find a solution without being specific to any particular channel estimation or prediction algorithm, for which there exists many, we adopt the widely used additive Gaussian noise model for estimated CSI [40–42].

- Upon digital power allocation, the remaining power is then used to transmit superimposed analog QAM symbols (quantization error samples from video codec) on each subchannel. The performance improvement achievable by sending the analog quantization error depends crucially on how we assign quantization errors from different video macro-blocks to subchannels as well as how we allocate the given power budget among the subchannels. First, there can be large variation in CSNR across the subchannels. For example, in OFDM systems, the CSNR can vary as much as 10 dB across the subcarriers [43]. Second, the quantization errors across different video macro blocks do not have the same importance. If the exact subchannel CSNRs are known to the transmitter and receiver, it can be shown that (under the MSE criterion) the close-form optimal solution to the aforementioned channel- assignment/power-allocation problem is a combination of waterfilling and inverse waterfilling [44]. However, when the estimates of the subchannel CSNRs are noisy, the optimal solution is not obvious and depends on the CSI error-model. To this end, we establish under the aforementioned additive noise model for CSI, a waterfilling solution for combined channel-assignment and power-allocation. This solution is optimal for each OFDM channel use.

- We present a complete implementation of an HDA video coding system incorporating the proposed power allocation scheme for color video transmission over an OFDM system wherein the base layer is generated using a standard H264/AVC video codec [45]. Our experimental results, obtained using numerous standard test video sequences, show that the proposed HDA scheme achieves better video quality [up to 3-6 dB in peak signal-to-noise ratio (PSNR)] during motion intensive video segments, compared to a typical layered video transmission scheme implemented using a standard H264/SVC video codec [39] as well as a purely analog transmission scheme which is a modification (to account for CSI errors) of the recently reported Parcast+ [6].

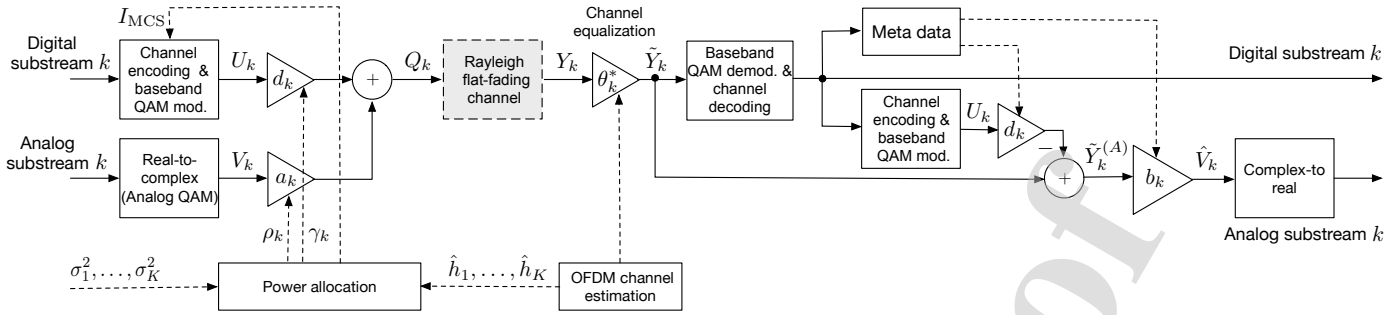


Figure 1: A single subchannel of the multichannel communication system model considered in this paper.

Related previous work- HDA video transmission schemes for time-varying channels reported in the literature fall into two distinct categories: (i) broadcast/multicast where the transmitter cannot be adapted to the channel state [2, 4, 5, 7, 8, 11, 14, 16, 17, 21, 22, 46] and (ii) unicast where the transmitter is adapted to the channel state, but a better power/bandwidth efficiency is achieved by using HDA encoding rather than digital layered coding. The problem considered in this paper belongs to the latter category. Some of the related previous work in this category only applies to single-carrier transmission [12, 19, 21]. The work closely related to this paper includes [6, 9] which consider CSI-based power allocation for video transmission in OFDM systems. However, in both [6] and [9] the availability of ideal CSI has been assumed. Ref. [6] presents a purely analog video transmission method for MIMO-OFDM channels where CSI is used for MIMO-channel precoding and power allocation. Unlike SoftCast [4], which relies on 3D-DCT for spatio-temporal video compression, [6] combines 2D-DCT with motion- compensated temporal filtering (MCTF) which is shown to improve the performance. The main shortcoming of purely analog video transmission is the difficulty of achieving coding gains, and hence the bandwidth efficiencies comparable to those possible with the state-of-the-art digital video codecs. Ref. [9] partially addresses this issue by using a hybrid scheme where an H264/AVC codec is used to first compress the video signal, and the analog quantization error is transmitted with the approach in [6]. The bandwidth efficiency of this approach, though better than purely analog transmission, is still limited by the use of BPSK and the orthogonal transmission of the analog and digital parts.

To the author's knowledge, none of the previous work on HDA video coding, either in relation to single-carrier or multicarrier transmission, has not considered the effect of noise in available CSI estimates on the system performance. Systems optimized for noisy CSI can be highly vulnerable to channel dependent video quality degradations, including complete decoder outages. This issue has been addressed recently in [47], in the context of communicating a memoryless Gaussian source over a fading channel using a simple HDA scheme based on entropy constrained scalar quantization (ECSQ). That analysis, which relies on analytical expressions for rate-distortion characteristics of an ECSQ as well as an assumption of orthogonal transmission of the analog and digital parts, however cannot be applied to a system such as ours, which uses a state-of-the art video codec for quantization and digital-analog superposition for HDA transmission.

The rest of the paper is organized as follows. Section 2 describes the mathematical model of the HDA system considered in this paper. Section 3 formulates and solves the power allocation problem. Section 4 describes the OFDM video transmission system implemented in this paper. Sec. 5 presents and discusses the experimental results. Conclusions appear in Sec. 6.

2. System Description

2.1. Channel model

The basic model is a set of K parallel subchannels with equal bandwidths and different gains. We assume the follow-

ing: (1) total average transmitted power on all K subchannel has a limit P_T, (2) each subchannel is an AWGN channel which undergoes Rayleigh flat fading independent of the others, and (3) channel gains are made available to the encoder and decoder at T_{CSU} s intervals, which we refer to as the channel state update (CSU) period.

Let the gains of the subchannels during a given CSU period be $h_1, \ldots, h_K$. Under the Rayleigh fading model, the real and imaginary parts of each h_k are iid Gaussian variables with mean zero [48]. Let the variances of these variables be a known value σ_h^2, so that $E\{|h_k|^2\} = 2\sigma_h^2$.

2.2. Modeling CSI Error

The statistics of CSI are highly dependent on the channel estimation and prediction methods used as well as various other system details. The use of a system specific model can not only make the analytical derivations overly complicated, but also results in solutions that are not broadly applicable. For this reason, we adopt the the additive Gaussian noise model which has been widely used in the literature for studying the effects of noisy CSI, see for example [40–42] and the references therein. Under this model, the estimate $\hat{h}_k$ is related to its true value h_k by $\hat{h}_k = h_k + \epsilon_k$ where ϵ_k is a mean zero complex Gaussian variable with iid real and imaginary parts, independent of h_k. For jointly Gaussian variables, we can also write

$$h_k = \hat{h}'_k + \epsilon'_k, \tag{1}$$

where $\hat{h}'_k = \alpha \hat{h}_k$, $\alpha = \frac{\sigma_h^2}{\sigma_h^2 + \sigma_\epsilon^2}$, and σ_ϵ^2 is the variance of real and imaginary parts of ϵ_k, and ϵ'_k is a complex Gaussian variable independent of $\hat{h}_k$, with mean zero and the variance $\sigma_{\epsilon'}^2 = \alpha\sigma_\epsilon^2$ per real/imaginary part. For obvious reasons, we define $S_h = \sigma_h^2/\sigma_\epsilon^2$ as the *CSI estimation signal-to-noise ratio* (CSI-SNR) which measures the accuracy of the channel estimates.

2.3. Video encoding

Let $\boldsymbol{S}$ be a GOP consisting N_{GOP} consecutive video frames. The basic problem at hand is to compress and communicate $\boldsymbol{S}$ to a single receiver over the channel described above, by using a total power of P_T per single use of the bank of K subchannels (equivalent to one OFDM symbol). If the GOP reconstructed after decoding is $\hat{\boldsymbol{S}}$, our goal is to minimize the average distortion between $\boldsymbol{S}$ and $\hat{\boldsymbol{S}}$. Given that the random subchannel gains change every T_{CSU} s, minimizing the average distortion requires selecting appropriately the source and channel coding rates, as well as the modulation order according to the estimated channel state. Since the gains across the K subchannels can be different, the transmitter power must be allocated among the K channels accordingly. Recall that the duration of $\boldsymbol{S}$, T_{GOP} is longer than T_{CSU} so that power allocation has to be repeated many times during the transmission of a GOP. For maximum coding efficiency, all video frames in a GOP have to be encoded jointly, using a given target bit rate. However, it is not possible to determine the optimal target bit rate to be used by the video encoder (determined by the choice of QP) as (1) the channel state for the entire duration of the GOP is not fixed, and (b) any available CSI is not exact. Therefore, we employ a two-layer HDA coding scheme with a fixed rate base layer whose bit rate is chosen based on a prior estimate of the worst-case CSI.

2.3.1. Transmission scheme

The first step in our HDA transmission scheme is to digitally encode the input sequence $\boldsymbol{S}$ to generate the base layer bit sequence, followed by the computation of the quantization error (QE) sequence $\boldsymbol{S} - \hat{\boldsymbol{S}}_b$, where $\hat{\boldsymbol{S}}_b$ is the reconstructed base layer. The base layer bit sequence is demultiplexed into K substreams each of which is channel encoded and mapped to QAM symbols to be transmitted over one of the subchannels. The QE of the source encoder, which acts as the refinement layer, is transmitted in analog form by superposition on the digital QAM symbols. During each CSU period, the total power is allocated between the digital and analog QAM symbols as well as across the K subchannels based on the estimated channel gains and the variances of QEs transmitted on the subchannels, so as to minimize the MSE of video reconstructed at the receiver. A block diagram of a single subchannel in our system is shown in Fig. 1.

In [49], we considered an HDA scheme where QE samples

were used to form analog QAM symbols without any other pre-processing. However, it was found that with most video sequence, the encoder output typically has a non-negligible residual intra- and inter-frame redundancies. Therefore it is inefficient in terms of channel bandwidth to send the QE samples without eliminating the correlation. In order to improve the efficiency, we first apply the 3D discrete cosine transform (3D-DCT) [50] to decorrelated the blocks of pixels in QE frames, and then transmit the resulting *quantization-error transform coefficients* (QETCs) using analog QAM. We partition the sequence of QETCs into K sub-streams such that each substream can be assigned to one subchannel. As described in Sec. 4, these substreams are formed by grouping QETCs with the same importance (this is possible due to the energy compaction property of the DCT). Two consecutive values in a QETC substream are used as the in-phase and quadrature components of an analog QAM symbol. A sequence of HDA channel symbols for each of the K subchannels is formed by the superposition of the QAM symbols from the respective digital and the analog substreams.

Let $\boldsymbol{X} = (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_K)^T$ be the sequence of n QETCs transmitted during a given CSU period where $\boldsymbol{Z}_k$ denotes the substream of n/K real-valued QETCs transmitted on the k-th subcarrier, $k = 1, \ldots, K$. Assume that all QETCs in $\boldsymbol{Z}_k$ are independent with zero mean and variance ν_k^2. The following description applies to equivalent complex baseband symbols transmitted during one CSU period. Denote an arbitrary pair of digital and analog QAM symbols transmitted on the k-th subchannel by two complex random variables U_k and V_k respectively. Without a loss of generality, let $E\{|U_k|^2\} = 1$. Since the complex symbol V_k is formed by combining a pair of QETCs in the substream $\boldsymbol{Z}_k$, $E\{|V_k|^2\} = 2\nu_k^2$. The HDA QAM symbol is then formed by the superposition of the analog QAM symbol V_k and the digital QAM symbol U_k. The transmitted HDA QAM symbol is thus given by $Q_k = d_k U_k + a_k V_k$, where d_k and a_k are real constants chosen such that $\sum_{k=1}^{K} E\{|Q_k|^2\} = P_T$. Suppose the fractions of total transmitter power allocated to the digital and analog QAM symbols on the k-th subchannel are $0 \le \gamma_k \le 1$ and $0 \le \rho_k < 1$ respectively. Then, we have $d_k = \sqrt{\gamma_k P_T}$ and $a_k = \sqrt{\rho_k P_T/(2\nu_k^2)}$. For convenience, we define $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)$ and $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_K)$.

2.4. Video decoding

The output of the k-th subchannel is given by $Y_k = h_k Q_k + W_k$, where W_k is the complex white Gaussian channel noise with variance σ_w^2 per real/imaginary component. At the receiver, the output signal on each subchannel is first equalized by using a single-tap linear equalizer θ_k which has to be computed from the estimated channel gain $\hat{h}_k$. The equalized channel output $\tilde{Y}_k = \theta_k Y_k$ is the input to the channel-decoder used to recover the bit sub-stream transmitted on the subchannel k. From (1), it follows that

$$\tilde{Y}_k = Q_k \hat{h}'_k \theta_k + W'_k, \tag{2}$$

where we have defined $W'_k = (Q_k \epsilon'_k + W_k)\theta_k$. Note that the data term $Q_k \hat{h}'_k \theta_k$ and the noise term W'_k in (2) are uncorrelated. As the digital QAM symbols are uniformly distributed and the superimposed analog transform coefficients are approximately Gaussian distributed [51], it can be shown that the real/imaginary parts of W'_k have zero-mean, symmetrical and unimodal pdfs which can be well approximated by a Gaussian pdf, a fact which has also been confirmed by experimentally observed histograms. In this case, the zero-forcing equalizer for the channel output is simply $\theta_k = 1/\hat{h}'_k$, and we have

$$\tilde{Y}_k = d_k U_k + a_k V_k + W'_k. \tag{3}$$

The total noise component at the input to the channel decoder of the subchannel k is the sum of interference from the analog signal $a_k V_k$ and the noise W'_k, which are independent random variables. We assume that the total noise $\tilde{W}_k = a_k V_k + W'_k$ is iid with a zero-mean Gaussian distribution. In order to keep the decoding error probability negligible, the power allocation to the digital transmission must be chosen to ensure that the CSNR at the input to the channel decoder for each subchannel is above a certain threshold which depends on the MCS used on the subchannel. All modern wireless transmission standards

Table 1: MCS set defined in the IEEE802.11a OFDM standard [35, Table 78]. The last column shows the estimated minimum CSNR required to maintain the bit-error below 10^{-7}.

MCS no.	Modulation	Channel Coderate	Bits/symbol	Threshold SNR (dB)
1	BPSK	1/2	0.5	7
2	BPSK	3/4	0.75	9
3	QPSK	1/2	1	10
4	QPSK	3/4	1.5	13
5	16-QAM	1/2	2	14
6	16-QAM	3/4	3	20
7	64-QAM	2/3	4	24
8	64-QAM	3/4	4.5	26

include a set of allowed MCSs. The set of MCSs used in this paper (taken from [35]) is shown in Table 1, along with the experimentally determined threshold CSNRs. In this paper we will assume that all subchannels use the same MCS for a given CSU period. This assumption can be relaxed at the expense of an increase in the complexity of the digital power-allocation problem considered below. When restricted to the same MCS on all channels, the CSNR threshold, or equivalently the index $I_{\text{MCS}} \in \{1, \ldots, N_{\text{MCS}}\}$ identifying the MCS, is an optimization variable in the power allocation problem, where N_{MCS} is the number of MCSs supported by the transmitter. We next determine an expression for the CSNR at the input of the channel decoder k.

First note that CSNR expression for binary phase shift keying (BPSK) is different to that of QAM as for the former, only the inphase component [real part of (3)] of the equalized channel output matters. Note also that, in the case of BPSK, $d_k U_k$ in (3) is real. Therefore, in general we can express the noise variance at the input to channel decoder k as

$$E\{|\tilde{W}_k|^2\} = \frac{2\sigma_w^2}{|\hat{h}'_k|^2}\left[S_\epsilon \gamma_k + (\delta \hat{S}_k + S_\epsilon)\rho_k + 1\right],$$

where

$$\delta = \begin{cases} 0.5 & \text{for BPSK} \\ 1 & \text{for QAM} \end{cases}$$

and, we have defined $S_\epsilon = \alpha P_T \sigma_\epsilon^2/\sigma_w^2$ and $\hat{S}_k = P_T|\hat{h}'_k|^2/(2\sigma_w^2)$. Note that as $\sigma_\epsilon^2 \to 0$ (when CSI becomes exact), we have $S_\epsilon = 0$ and $\hat{S}_k = S_k = P_T|h_k|^2/(2\sigma_w^2)$. Here, S_k is the true but unknown CSNR when total power is allocated to channel k and $\hat{S}_k$ is its estimate. The predicted CSNR at the input of the digital channel decoder k is thus

$$\hat{\Gamma}_k(\hat{S}_k) = \frac{\gamma_k \hat{S}_k}{\left[1 + S_\epsilon \gamma_k + (\delta \hat{S}_k + S_\epsilon)\rho_k\right]}, \tag{4}$$

$k = 1, \ldots, K$. In Sec. 3, we will use this expression to solve the power allocation problem.

Now suppose that the base layer bit stream has been decoded error-free at the receiver, i.e., $d_k U_k$ is known. In order to estimate the analog component, the transmitted digital signal is first canceled from the channel output to obtain $\tilde{Y}_k^{(A)} = \tilde{Y}_k - d_k U_k$ (see Fig. 1). The analog signal (a pair of QETC) V_k is then estimated from $\tilde{Y}_k^{(A)} = a_k V_k + W'_k$, where W'_k is uncorrelated with V_k and is Gaussian by assumption. Therefore, the optimal (MMSE) estimator for V_k, given $\tilde{Y}_k^{(A)}$ is linear. It can be shown that (see [52]) the optimal linear estimate for V_k is $\hat{V}_k = b_k \tilde{Y}_k^{(A)}$, where

$$b_k = \frac{1}{a_k} \frac{\rho_k P_T}{(\rho_k P_T + 2\sigma_{w'k}^2)}, \tag{5}$$

and

$$\sigma_{w'k}^2 = \frac{1}{2}E\{|W'_k|^2\} = \frac{\sigma_w^2}{|\hat{h}'_k|^2}\left[(\gamma_k + \rho_k)S_\epsilon + 1\right].$$

The resulting MMSE is $E\{|V_k - \hat{V}_k|^2\} = 4\nu_k^2 \sigma_{w'k}^2/(2\sigma_{w'k}^2 + \rho_k P_T)$. The estimates $\hat{V}_1, \ldots, \hat{V}_k$ are then multiplexed to form the MMSE estimate $\hat{\boldsymbol{Z}}$ of the transmitted QETC vector $\boldsymbol{Z}$. It now follows that the total MSE of QETCs transmitted during the CSU period is

$$\begin{aligned} D(\boldsymbol{\gamma}, \boldsymbol{\rho}, I_{\text{MCS}}) &= \frac{1}{n}E\{\|\boldsymbol{Z} - \hat{\boldsymbol{Z}}\|^2\} \\ &= \frac{1}{K}\sum_{k=1}^{K} \frac{2\nu_k^2(1 + S_\epsilon \gamma_k + S_\epsilon \rho_k)}{1 + S_\epsilon \gamma_k + (S_\epsilon + \hat{S}_k)\rho_k}, \end{aligned} \tag{6}$$

where we have made explicit the fact this MSE depends on the power allocation as well as the choice of MCS. Since the transform used to map QEs to QETCs is orthogonal, (6) is also the total MSE of the reconstructed video frames in the transmitted GOP.

3. Power Allocation Problem and Solution

The goal of optimal power allocation is to determine $\boldsymbol{\gamma}$, $\boldsymbol{\rho}$, and I_{MCS} which minimize (6) subject to a constraint on the total transmit power. This problem has to be solved at the beginning of each CSU period using the estimated CSI, $\hat{\boldsymbol{h}} = (\hat{h}_1, \ldots, \hat{h}_K)$. First we recall that, in addition to the power constraint, we also have the constraint that all bits of the base layer must be transmitted in a time less than or equal to a GOP duration, or equivalently, using at most $N_{\text{CSU}} = T_{\text{GOP}}/T_{\text{CSU}}$ CSU periods. A failure to complete the transmission of all base layer bits within N_{CSU} CSU periods will result in no decoded video output for the given GOP, a condition we will refer to as a *decoder outage*. Since the estimated CSI is only available for the current CSU period, the total number bits that can be supported during the remaining CSU periods cannot be known exactly. Given that the base layer bit rate is fixed, there is always a non-zero probability that an outage will occur. It follows that, to minimize the outage probability, we have to maximize the number of base layer bits that can be transmitted during each CSU period $n = 1, \ldots, N_{\text{CSU}}$. The rate maximizing power allocation can be found by solving the following problem.

3.1. Digital power allocation

$$\begin{aligned} \max_{i,\boldsymbol{\gamma}} \ & \sum_k R_{\text{MCS}}(i)\mathcal{I}(\gamma_k > 0) \\ \text{subject to} \ & \sum_k \gamma_k \le 1 \\ & i \in \{1, \ldots, N_{\text{MCS}}\} \\ & \rho_k = 0 \text{ and } 0 < \gamma_k \le 1, k = 1, \ldots, K, \end{aligned}$$

where $R_{\text{MCS}}(i)$ is bit rate of the MCS i (column 4, Table 1) and $\mathcal{I}(\cdot)$ is the indicator function. First consider digital power allocation assuming a single MCS with bit rate r bits/symbol and CSNR threshold η. It is required that $\hat{\Gamma}_k(\hat{S_k}) \ge \eta$ for $k = 1, \ldots, K$. Using (4), we obtain the solution

$$\gamma_k \ge \max\left\{0, \frac{\eta}{\hat{S}_k - \eta S_\epsilon}\right\}, \tag{7}$$

and $\sum \gamma_k \le 1$. The total supported bit rate is $r \sum \mathcal{I}(\gamma_k > 0)$. The total power allocated is $\sum \gamma_k P_T$ which can be minimized by ordering CSI such that that $\hat{S}_k \ge \hat{S}_{k+1}$, $k = 1, \ldots, K-1$ and solving (7) in that order. Now if we have a set of MCSs as in Table 1, we can solve the optimal power allocation problem for each MCS and choose the one that maximizes the total bit rate. Let the solution to this problem be $\bar{\boldsymbol{\gamma}}$. If $\sum \bar{\gamma}_k < 1$, then the excess power $\tilde{P} = (1 - \sum \bar{\gamma}_k)P_T$ can be allocated to the analog transmission. It is worth pointing out that, the digital power allocation using (7) is equivalent to adaptive selection of the optimal CSNR threshold for each MCS based on noisy CSI (rather than directly using the ideal thresholds in Table 1). Without doing so, the base layer cannot be transmitted reliably.

3.2. Analog power allocation

In this section we solve the analog power allocation problem in close-form and interpret it as an instance of the familiar waterfilling [44]. We first note that, whenever power is allocated to an analog QAM symbol, i.e., $\rho_k > 0$, the CSNR of the digital QAM symbol on that subchannel degrades due to interference, see (4). We therefore have to simultaneously allocate power to both analog and digital symbols such that the CSNR at the input of the digital decoder is unchanged. Let $\gamma'_k P_T$ be the additional power allocated to digital QAM symbols on subchannel k to compensate for analog interference. The total power allocated to digital QAM symbols on subchannel k is then $\gamma_k P_T = (\gamma'_k + \bar{\gamma}_k)P_T$. Now, to maintain the same digital CSNR as before, we need

$$\frac{\bar{\gamma}_k \hat{S}_k}{1 + \bar{\gamma}_k S_\epsilon} = \frac{\gamma_k \hat{S}_k}{1 + \gamma_k S_\epsilon + \rho_k(\delta \hat{S}_k + S_\epsilon)}.$$

It follows that $\gamma'_k = \rho_k \bar{\gamma}_k(\delta \hat{S}_k + S_\epsilon)$. Defining $\hat{S}'_k = \delta \hat{S}_k + S_\epsilon$, we have $\gamma_k = \bar{\gamma}_k(1 + \rho_k \hat{S}'_k)$, which is the requirement for preventing analog interference to the digital symbols. The power constraint $\sum(\gamma_k + \rho_k) = 1$ can now be restated as $\sum \beta_k \rho_k = 1$, where

$$\beta_k = \frac{1 + \bar{\gamma}_k \hat{S}'_k}{1 - \sum \bar{\gamma}_k}.$$

We can also re-write the MSE in (6) as (ignoring the normalization constant K)

$$D(\boldsymbol{\rho}|\bar{\boldsymbol{\gamma}}) = \sum_{k=1}^{K} \frac{2\nu_k^2(F_k + G_k\rho_k)}{F_k + (G_k + \hat{S}_k)\rho_k},$$

where $F_k = 1 + S_\epsilon \bar{\gamma}_k$ and $G_k = S_\epsilon(1 + \bar{\gamma}_k \hat{S}'_k)$.

We can now state the analog power allocation problem as follows. Given P_T, $\hat{\boldsymbol{S}} \triangleq (\hat{S}_1, \ldots, \hat{S}_K)$, I_{MCS} and $\bar{\boldsymbol{\gamma}}$, solve

$$\min_{\boldsymbol{\rho}} D(\boldsymbol{\rho}|\bar{\boldsymbol{\gamma}}) \tag{8}$$

subject to $\sum \beta_k \rho_k = 1$ and $0 \le \rho_k \le 1$ for $k = 1, \ldots, K$. This minimization problem is convex, and can be solved by the Lagrange multiplier method. The Lagrangian is

$$J(\boldsymbol{\rho}) = \sum_k D(\boldsymbol{\rho}|\bar{\boldsymbol{\gamma}}) + \lambda \sum_k (\beta_k \rho_k - 1),$$

where $\lambda \ge 0$ is the Lagrange multiplier. By letting $\partial J / \partial \rho_k = 0$ we obtain the solution $\lambda = \Delta_k(\rho_k)$ for $k = 1, \ldots, K$, where

$$\Delta_k(\rho_k) = \frac{2\nu_k^2 F_k \hat{S}_k}{\beta_k [F_k + \rho_k (G_k + \hat{S}_k)]^2}.$$

Now using the Karush-Kuhn-Tucker (KKT) conditions we find the optimal analog power allocation as

$$\rho_k = \begin{cases} \frac{F_k}{G_k + \hat{S}_k} \left(\sqrt{\frac{2\nu_k^2 \hat{S}_k}{\lambda \beta_k F_k}} - 1 \right) & \text{if } \lambda < \frac{2\nu_k^2 \hat{S}_k}{\beta_k F_k} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

This solution can be found as follows. For convenience let $T_k = \beta_k F_k / (G_k + \hat{S}_k)$ and $\phi_k = \hat{S}_k / (\beta_k F_k)$. Without a loss of generality assume $\nu_1^2 \phi_1 \ge \nu_2^2 \phi_2 \ldots \ge \nu_K^2 \phi_K^2$. Also define the function

$$\Lambda(K') \triangleq \frac{\left(\sum_{k=1}^{K'} T_k \sqrt{2\nu_k^2 \phi_k} \right)^2}{\left(1 + \sum_{k=1}^{K'} T_k \right)^2}, \; 1 \le K' \le K. \tag{10}$$

Now determine $K^* \in \{1, \ldots, K\}$ such that $2\nu_j^2 \phi_j \ge \Lambda(K^*)$ for $j = 1, \ldots, K^*$, and $2\nu_j^2 \phi_j < \Lambda(K^*)$ for $j = K^* + 1, \ldots, K$. The optimal Lagrange multiplier is then given by $\lambda^* = \Lambda(K^*)$. Let the optimal power allocation be (γ_k^*, ρ_k^*), The resulting MMSE is

$$D^* = \frac{1}{K} \sum_{k=1}^{K^*} \frac{2\nu_k^2 [1 + (\gamma_k^* + \rho_k^*) S_\epsilon]}{1 + (\gamma_k^* + \rho_k^*) S_\epsilon + \rho_k^* \hat{S}_k} + \frac{1}{K} \sum_{k=K^*+1}^{K} 2\nu_k^2. \tag{11}$$

This solution can be interpreted as waterfilling on $\Delta_k(\rho_k)$ with a water level λ^* which satisfies (10). This solution also shows how QETCs $Z_1, \ldots, Z_K$ should be assigned to subchannels. Suppose coefficients $Z_1, \ldots, Z_K$ are labeled such that $\nu_1^2 \ge \nu_2^2 \ldots \ge \nu_K^2$ and the subchannels are labeled such that $\phi_1 \ge \phi_2 \ldots \ge \phi_K$. Then, assigning k-th QETC to the k-th subchannel minimizes (11). Note that those subchannels for which $\nu_k^2 \phi_k$ is too small will not be used and hence Z_k will not be transmitted, $k = K^* + 1, \ldots, K$. This can be seen as optimal bandwidth compression of the transmitted analog signal to meet the power and channel-capacity constraints. Note also that the optimal solution does not necessarily assign QETCs with the larger variances to subchannels with the larger power-gains, as in the case of purely analog transmission [4, 6].

4. Application to Video over OFDM Systems

In this section we describe an implementation of an HDA system for real-time color video transmission over a 20 MHz, 64 subcarrier OFDM system ($N_{\text{FFT}} = 64$) whose symbol rate and MCS set were adopted from the IEEE 802.11a standard [35]. While the format of source video and the codec used to generate the base layer are not particularly important, the way we divide a QE frame sequence into substreams with different variances will impact the quality of the HDA transmitted video. In our implementation we have used color video in YCbCr format [32] as it already stores a video sequence as 3 separate subsequences, the more important luminance (Y) component, and the relatively less important chrominance components Cr and Cb. The QETCs for analog transmission are formed by applying the 3D-DCT to QE frames. Due to the energy compaction property of the DCT, the QET coefficients can be used to form substreams with unequal variances as required for power allocation. In previous work on analog video transmission, 3D-DCT has been shown to be effective method of decorrelating successive video frames [4]. An alternative is using the 2D-DCT with motion-compensated temporal filtering (MCTF) [6], but motion-compensation will be much less useful with QE frames, and it will also introduce an unnecessary overhead as the motion vectors must also be sent to the decoder.

The values of various parameters used for video encoding and the packetization of analog QAM symbols are shown in Table 2. The coded bits of the base layer and the corresponding QETCs of a single GOP consisting of a total of N_s pixels

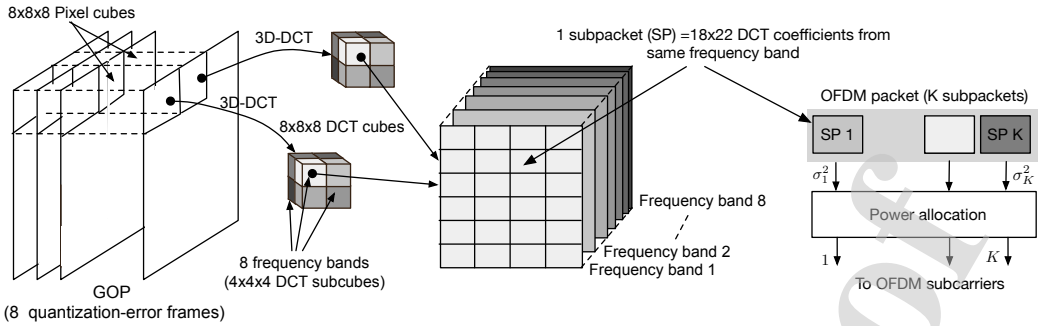


Figure 2: Mapping quantization errors to OFDM Packets. Subpackets are generated for Y-GOP, U-GOP, and V-GOP separately, pooled together, and sorted according to their variances for power-allocation.

Table 2: Summary of important HDA transmitter parameters.

Parameter	Symbol	Value
Video frame rate		30 fps
Frame size (YCbCr 4:2:0), CIF		352×288
4CIF		704×576
Video resol. (bits/YCbCr-pixel)		24
GOP size (frames)	N_{GOP}	8
OFDM symbol duration	T_{SYM}	4 μs
Preamble duration	T_{PA}	20 μs
OFDM video data symbols/packet	L_{P}	792
OFDM meta data symbols/packet (CIF/4CIF)	N_{H}	64/68

are transmitted in a sequence of $N_{\text{P}} = T_{\text{GOP}}/T_{\text{P}}$ OFDM packets of fixed-size (see Table 2.) We let an OFDM packet be a sequence of L_{P} OFDM data symbols, where an OFDM symbol is a set of N_{FFT} QAM symbols. An OFDM system typically transmits a preamble prior to transmitting data symbols in a packet. Additionally, each OFDM packets should carry in its header the meta-data required for HDA decoding (power allocation factors, and mean and variances of analog data, see Sec. 5). Let the preamble and header durations respectively be T_{PA} and T_{H} ($= N_{\text{H}} T_{\text{SYM}}$). The total duration of an OFDM packet is therefore $T_{\text{P}} = T_{\text{PA}} + T_{\text{H}} + T_{\text{SYM}} L_{\text{P}}$. In our implementation, we assume that CSI estimates are updated for each OFDM packet (i.e., $T_{\text{CSU}} = T_{\text{P}}$), and therefore the power allocation problem is solved for each transmitted packet.

The digital power allocation is based on the estimates of minimum CSNR required by each MCS to achieve a biterror rate less than 10^{-7}, as shown in Table 1. The coderates shown are those of the rate-1/2 punctured convolutional code specified for OFDM-PHY in IEEE802.11a [35, Fig. 114]. Despite error correction by OFDM-PHY, occasional biterrors do occur which can affect the correct decoding of the base layer. In order to prevent this, the base layer data is further protected by an outercode, before being passed to the OFDM-PHY. We have used (255,239) Read-Solomon code specified in [53] for this purpose.

4.1. Packetization of Analog Data

For the sake of reproducibility of the experimental results, we provide in the following a reasonably complete description of the procedure used to map QETCs to OFDM subchannels according to the variances of the QETCs and the parameters ϕ_k, $k = 1, \ldots, K$ of subchannels. The process is illustrated in Fig. 2. The total number QETCs that can be transmitted during a GOP is $2KL_{\text{P}}N_{\text{P}}$. It should be noted that unless the number of subchannels used for HDA transmission, $K \geq \frac{N_{\text{S}}}{2L_{\text{P}}N_{\text{P}}}$, some of the QETCs will be dropped, where $K < N_{\text{FFT}}$. This happens when the channel bandwidth is smaller than the video bandwidth. Recall also that (see Sec. 3.2) the optimal power allocation may result in zero power being allocated for analog transmission on one or more subchannels.

The first-step in mapping QETCs to analog OFDM packets is to apply the 3D-DCT to Y, Cr, and Cb components of

the QE frames (referred to as Y-GOP, Cr-GOP, and Cb-GOP) separately. In this work a GOP size of 8 frames has been used. Therefore, 3D-DCT is applied to 8×8×8 non-overlapping cubes of QE pixels in each component GOP. We then divide each coefficient cube (a DCT *cube*) into eight 4×4×4 sub-cubes which can be considered as different frequency sub-bands. We use the coefficients in the same frequency sub-band of different DCT cubes in a single component GOP to form one *frequency band*. Each of the 3 component GOPs is therefore transformed into 8 frequency bands, as shown in Fig. 2. The quantization errors of the base layer still contains some details from the original video frames and as a result the variances of the DCT coefficients in different frequency bands can be significantly different. Therefore, we use the frequency bands to generate substreams with different variances. Towards this end, each frequency band is divided into non-overlapping 18×22 rectangular sub-blocks which we refer to as *subpackets*. In order to appear as white noise to digital symbols, each subpacket is whitened by applying the non-ordered Haddamard transform. We form an OFDM packet in such a way that it consists of K subpackets which get transmitted in parallel over K subchannels. The assignment of subpackets to subchannels is dictated by the power-allocation solution in Sec. 3.2. In other words, $\boldsymbol{Z}_k$, $k = 1, \ldots, K$ in Sec. 2 represent subpackets in a single OFDM packet. In assembling OFDM packets, we consider only the variances of the subpackets and not the nature of the data (i.e, frequency band or color component).

5. Experimental Results

5.1. Test video sequences

The HDA system has been tested by using 10 standard CIF color video sequences, *Football* (360 frames), *Coastguard*, *Mobile*, *Stefan*, *Tabletennis*, *Foreman*, *Soccer*, *Crew* all having (300 frames), *Bus* (150 frames), and *Students* (1007 frames), as well as 4 standard 4CIF color video sequences *Soccer*, *Harbor*, *Crew*, and *City* (300 frames each). Experimental results for select sequences are presented in this Section.

5.2. Video coding and meta-data

The base layer of the HDA system has been generated using the H264/AVC reference implementation JM 19.0 in baseline profile [45]. The HDA system requires the transmission of certain amount meta-data for decoding each OFDM packet. These include the quantized values of the power allocation factors γ_k and ρ_k, $k = 1, \ldots, K$ (5 bits each), the means (6 bits each) and variances (6 bits each) of the K subpackets, as well as the K sequence numbers (10 bits each for CIF and 12 bits each for 4CIF) indicating the locations of subpacket data in video frames. This meta-data, which is critical to decoding, is sent at the start of every OFDM packet by using the lowest MCS (1/2 rate channel coding and BPSK) to ensure the maximum reliability (analog symbols are not transmitted in the packet-header). The number of header symbols N_{H} required to accommodate the meta-data bits is shown in Table 2.

5.3. OFDM Channel simulation

In order to model a wireless channel with frequency selective fading, the SISO ITU extended urban channel model with the classic Doppler spectrum from [54, Table 14] has been used. This model has 9 filter taps with relative delays (in ns) $\{0, 50, 120, 200, 230, 500, 1600, 2300, 5000\}$ and normalized received powers (in dB) $\{-1, -1, -1, 0, 0, 0, -3, -5, -7\}$. The Doppler spectrum has been generated using a maximum Doppler shift of 85 Hz ($T_{\text{COH}} \approx 5$ ms.) The channel traces were generated using the `comm.RayleighChannel` system object in the MATLAB Communication Toolbox®. The sampling rate of CIF video at 30 fps is 4.56×10^6 pixels per second, and hence the bandwidth is about 2.2 MHz. We therefore only considered using $K = 4$ (1.25 MHz total bandwidth) and $K = 6$ (1.87 MHz total bandwidth) subcarriers for CIF video transmission, out of the total of 48 data subcarrirers available in our OFDM system. For CIF sequences, the experimental results are presented here for two OFDM systems which use the same channel trace: (1) 10 dB mean CSNR and $K = 6$ sub-channels [*(6,10) channel*], and (2) 20 dB mean CSNR and $K = 4$ sub-channels [*(4,20) channel*]. In the case for 4CIF sequences, which have a band-

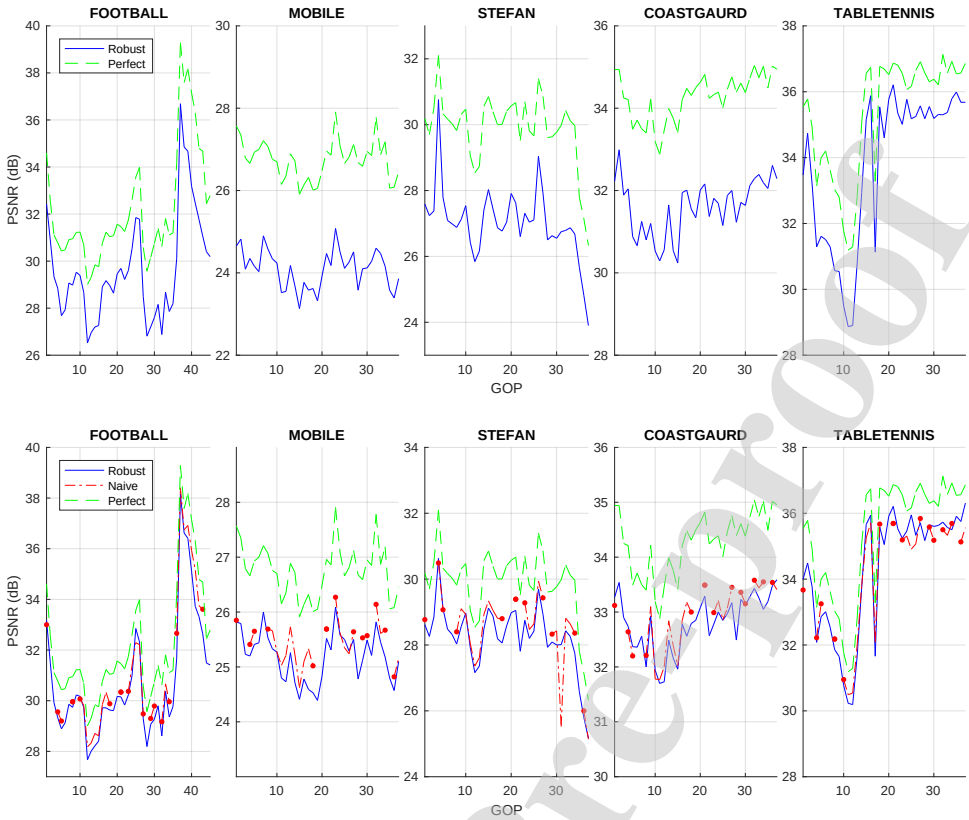


Figure 3: Performance comparison of HDA systems which perform power allocation as proposed in this paper (Robust), by ignoring CSI-errors (Naive), and using perfect CSI (Perfect.) on the (6,10) channel : (top row) CSI-SNR = 8 dB and (bottom row) CSI-SNR = 12 dB. At 10 dB CSI-SNR, the naive system lost all GOPs due to decoder outages. GOPs between filled circles (on naive system curve for 12 dB-CSNR) have been lost due to decoder outage. Base layer bit rate is 300 kbps.

width of 8.4 MHz, $K = 16$ subchannels (5 MHz total bandwidth) were allocated.

5.4. Advantage of Optimal Power Allocation

We first demonstrate the advantage of the proposed power allocation procedure when the CSI estimates are noisy. To this end, we compare the performance of a system which uses the power allocation method presented in Sec. 3, which we will refer to as a *robust* system, with that of (a) a system with perfect CSI, that is, a system with $S_h = \infty$, and (b) a *naive* system which simply ignores the errors in CSI, that is, assumes $S_h = \infty$ in power allocation and decoding.

Figures 3 and 4 show the PSNRs of these three systems on (6,10) and (4,20) channels respectively. In the case of (6,10) channel, a base layer bit rate of 300 kbps was used. Increasing the bit rate any higher resulted in some outages due to the supported channel bit-payload becoming too small during the periods of deep-fading. Similarly, the base layer bit rate for the (4,20) channel was set to 1 Mbps. In each case, the robust and naive systems were tested under two different CSI-SNR values as indicated in the figures [8 and 12 dB on (6,10) channel and, 17 and 23 dB on (4,20) channel]. Note that, in the case of 10 dB CSNR and 8 dB CSI-SNR, no results are shown for the naive system. This is because, in this case, most of the OFDM packets were undecodable, and consequently no video output was produced by the H264/AVC decoder. The gaps between the filled circles on the PSNRs curves for the naive system indicate video segments that were subject to outages, mostly due to excessive channel errors caused by power-allocations that did not match the actual channel conditions. The situation is not much better with 17 dB CSI-SNR on (4,20) channel. Even at 23 dB CSI-SNR for which the naive systems performs very close to the robust system, a loss of a GOP occurs in the *Coastguard* sequence. It was observed that, provided the base layer bit rate is

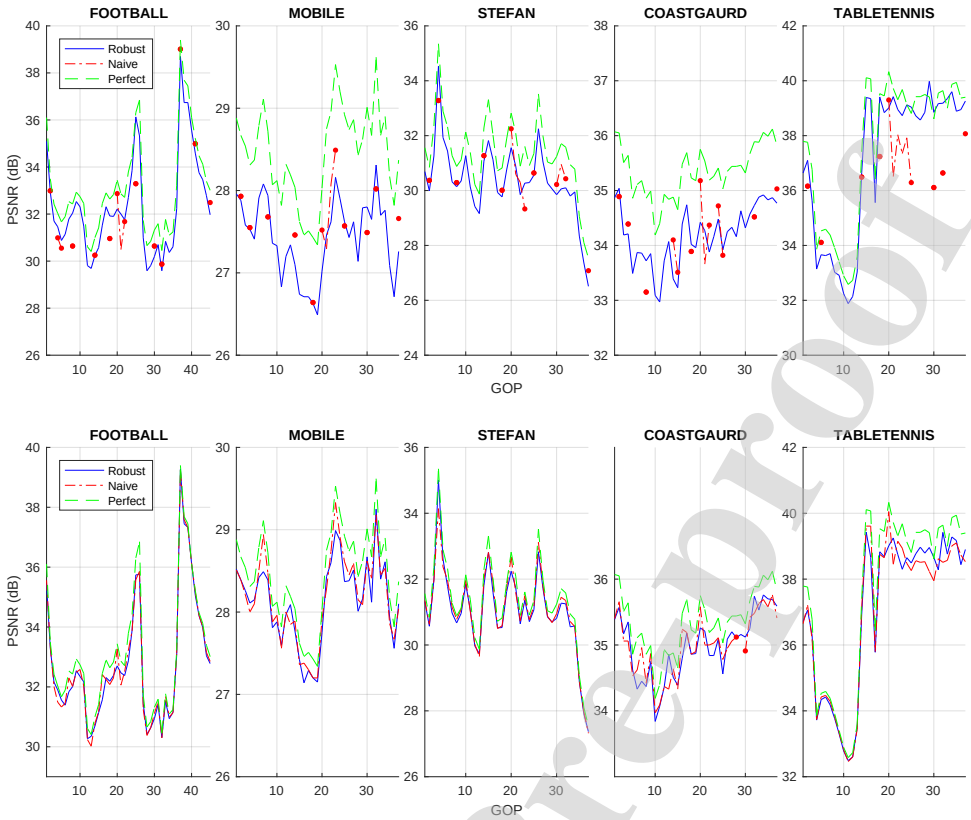


Figure 4: Performance comparison of HDA systems which perform power allocation as proposed in this paper (Robust), by ignoring CSI-errors (Naive), and using perfect CSI (Perfect.) on the (4,20) channel : (top row) CSI-SNR = 17 dB and (bottom row) CSI-SNR = 23 dB. GOPs between filled circles (on naive system curves) have been lost due to decoder outage. Base layer bit rate is 1 Mbps.

chosen appropriately (not too high), the robust power allocation rarely resulted in outages.

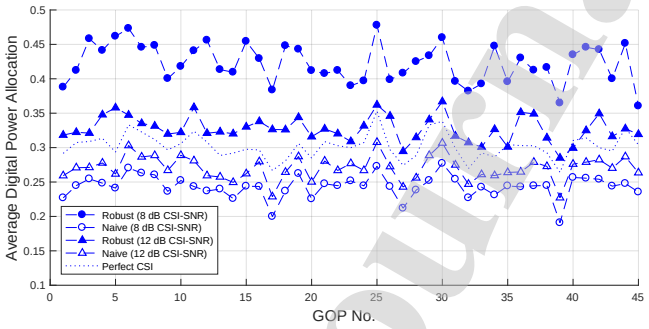


Figure 7: Average digital power allocation factor of the robust and naive systems for the Footbal sequence on (6, 10) channel (see Fig. 3.)

Further insight into the effect of the proposed power allocation procedure can be gained from Fig. 7 which compares the average digital power allocation factors of the robust, naive, and perfect-CSI systems for the Football sequence on the (6,10) channel (Fig. 3). The naive system blindly uses the noisy CSI

Table 3: Percentages of dropped analog packet in HDA systems considered in Figs. 3 and 4.

Channel	Perfect	CSI-SNR	Naive	Robust
(6,10)	54%	8 dB	56%	76%
		12 dB	57%	64%
(4,20)	68%	17 dB	68%	73%
		23 dB	68%	68%

to determine the digital power allocation based on the ideal CSNR thresholds in the Table 1 and therefore tends to allocate a smaller amount of power to the digital part than that is warranted by the unobserved true channel state. This "over estimation" of the MCS leads to mismatches between the chosen MCS and the actual channel state during the transmission, resulting in outages as shown in Figs 3. Note that, ignoring noise in CSI makes the digital power allocation relatively insensitive to the

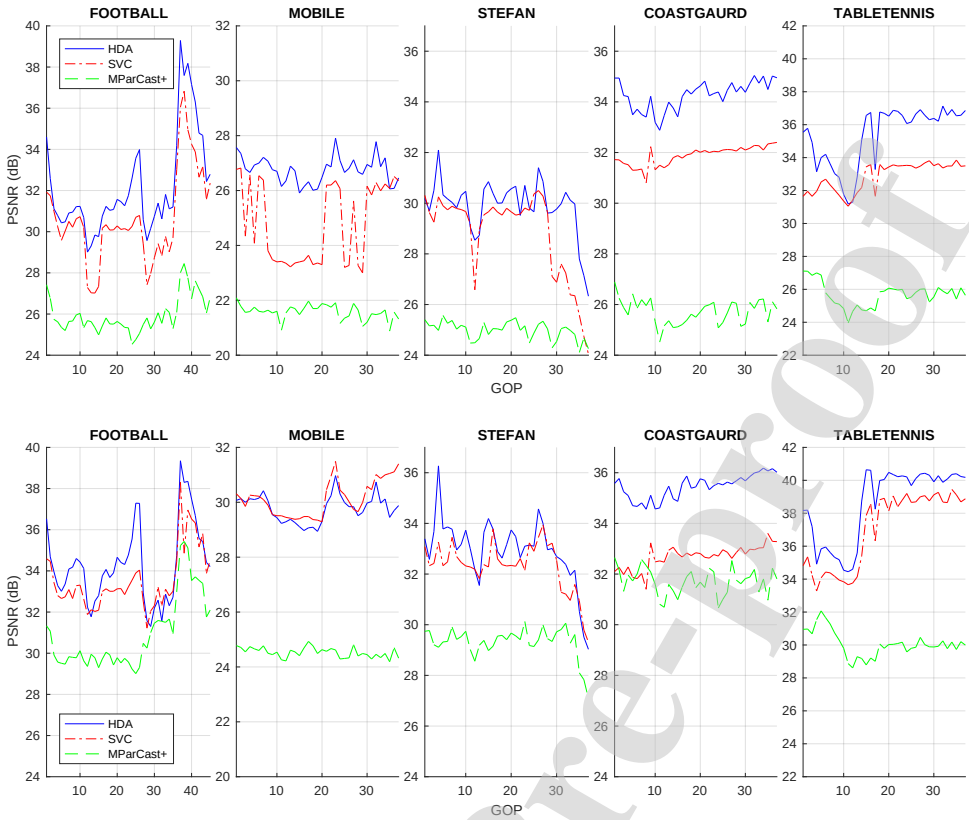


Figure 5: Performance of HDA, SVC, and MParCast+ systems with perfect CSI used for power-allocation and decoding on (6,10) channel (top row) and (4,20) channel (bottom row).

Original MParCast+ SVC HDA

Figure 6: . The original and reconstructed versions of the frame 241 of the Football sequence, when HDA, SVC, and MParCast+ systems use perfect CSI for power allocation and decoding [300 kbps base layer bit rate and (6,10) channel.]

CSI-SNR (note also that the naive system always allocates less power to the digital part compared to the system with perfect CSI, making it more susceptible to outages.) In contrast, the robust system computes the CSNR thresholds using (7), effectively making the power allocation a function of the CSI-SNR. This results in choosing an MCS for each packet that is much better matched to the likely actual channel state and thus greatly reduces the outage probability with a minimal sacrifice of the end-to-end video PSNR. The effect of optimal power allocation is also reflected in the Table 3 which shows the percentages of analog packets dropped by each system due to the lack of power and channel bandwidth. As the naive system ignores noise in CSI, it has percentages of dropped analog packets comparable to those of the perfect-CSI system. In contrast, with optimal power allocation in the robust system, a lower CSI-SNR not only results in more power being allocated to the digital part,

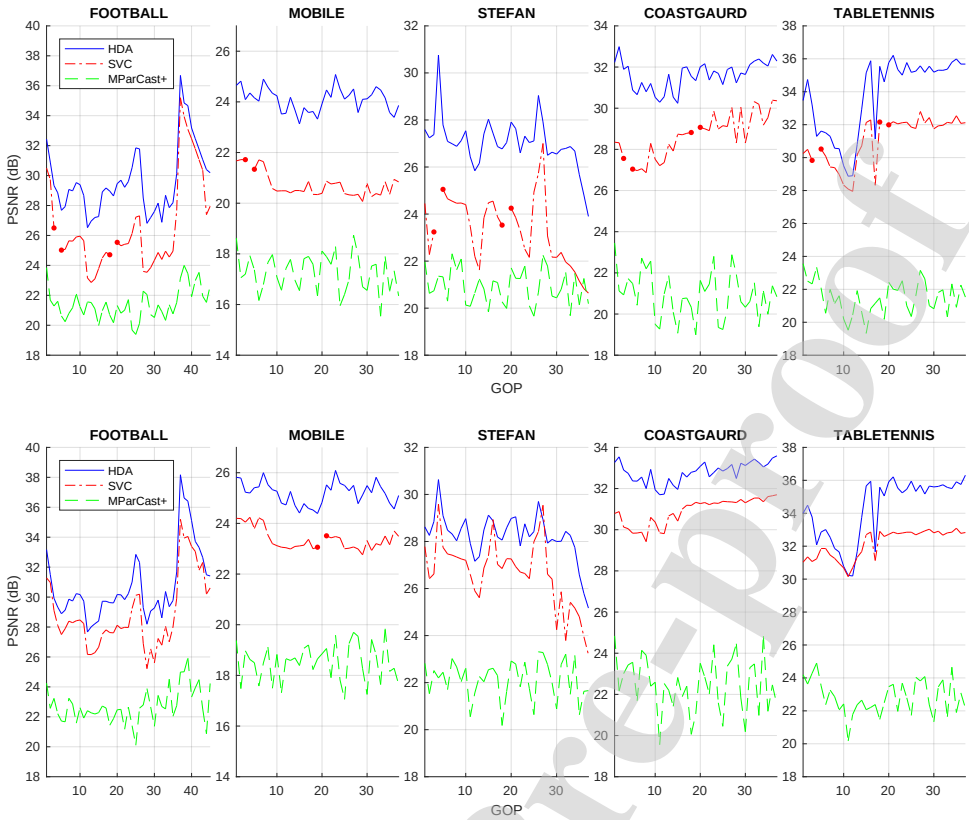


Figure 8: Performance of HDA, SVC, and MParCast+ systems under CSI errors on (6,10) channel: (top row) 8 dB CSI-SNR, and (bottom row) 12 dB CSI-SNR. Base layer bitrate is 300 kbps.

but also results in more of the remaining power being allocated to fewer analog packets prioritized based on their variances to minimize the end-to-end video MSE, as implied by (9).

5.5. *Comparisons*

We next compare the HDA system with two alternative schemes which are derived from the best digital-only and analog-only alternatives reported in recent literature.

5.5.1. *Layered digital coding*

An alternative to the proposed HDA scheme is a fully-digital video transmission system based on layered video coding [55]. For comparison purposes a 3-layer (a base layer and 2 refinement layers) system was implemented using the H264/SVC JSVM reference software [39]. The latter is the scalable video coding (SVC) extension of the H264/AVC encoder used in the HDA system. The two refinements layers were generated using the SNR scaling option (with fixed spatial and temporal resolutions) and the bit rate of the base layer being set equal to that of the HDA system. Bit rate control of a layered video encoder is complicated. In order to avoid the dependence of our results on any particular rate control algorithm, we used an off-line search algorithm to determine the QPs required to control the bit rates of the refinement layers based on the following rule: the bit rate of the first refinement layer is about 50%-80% of that of the base layer, and the basic QP of the second refinement layer is equal to that of the first refinement layer. Each layer-bitstream is formed by extracting the independently coded network access layer units (NALUs) corresponding to the same quality index, from the H264SVC encoder output. As in the HDA system, the power allocation to subcarriers was carried out on a per packet basis (using the same OFDM packet-size) so as to maximize the total channel bit rate subject to an average power constraint. The same procedure used with the HDA base layer was used for each layer, starting with the base layer, so that the base layer is

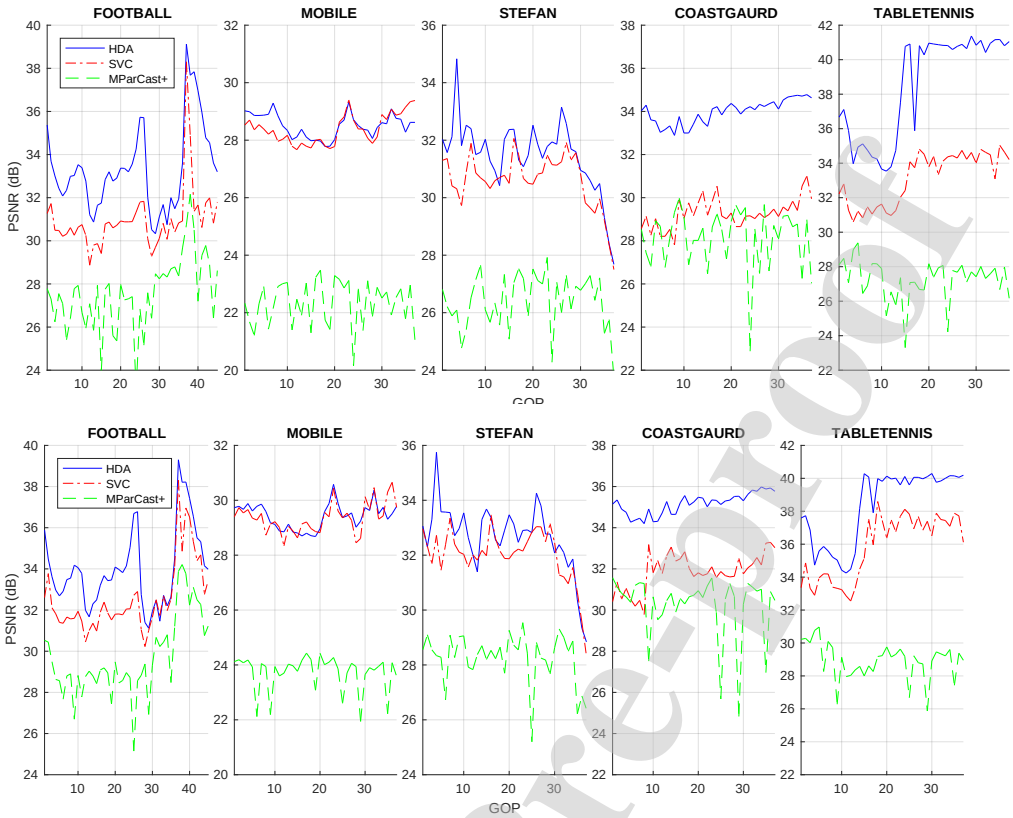


Figure 9: Performance of HDA, SVC, and MParCast+ systems under CSI errors on (4,20) channel: (top row) 17 dB CSI-SNR, and (bottom row) 23 dB CSI-SNR. Bae layer bit rate is 1 Mbps.

transmitted with the highest priority, followed by the two refinement layers.

5.5.2. *Uncoded (analog) transmission*

A fully analog system widely used in the literature as a baseline for comparisons with HDA coding is SoftCast [4]. SoftCast uses 3D-DCT to decorrelate a GOP and transmit-power is allocated to cubes of DCT coefficients, referred to as chunks, based on the variances of the chunks. SoftCast has been designed for broadcast applications and does not use CSI at the transmitter. Obviously such a system would fare poorly on a fading channel compared to our HDA system which adapts the transmitter power based on CSI estimates. ParCast+ described in [6] is an adaptation of SoftCast for OFDM channels (former also uses MCTF instead of 3D-DCT for improved decorrelation). However, as ParCast+ simply assumes perfect CSI for power allocation, we modified it to accommodate CSI errors by replacing the power-allocation scheme in [6] by ours, but with digital power allocation factors set to zero. In the following, we will refer to this scheme as modified ParCast+ (MParCast+). In a related work [17], a pure analog transmission for an OFDM channel, has been considered, where periodic pilot-symbols are used for channel estimation. Different to our setup, [17] considers assigning multiple subchannels to each analog symbol based on its variance so that one can benefit from diversity combining to mitigate the effect of CSI estimation errors. This approach cannot be easily adopted for comparisons here as the modification required to fit it to our setup appears non-trivial.

5.5.3. *Test results*

We first compare the performance of HDA, digital-layered (simply referred to as SVC), and purely analog MParCast+ systems when perfect CSI is available to both encoder and decoder. Figure 5 shows the PSNRs of these three systems on (6,10) and (4,20) channels. Both HDA and SVC systems use the same base layer bit rates (300 kbps at 10 dB CSNR and 1 Mbps at 20 dB

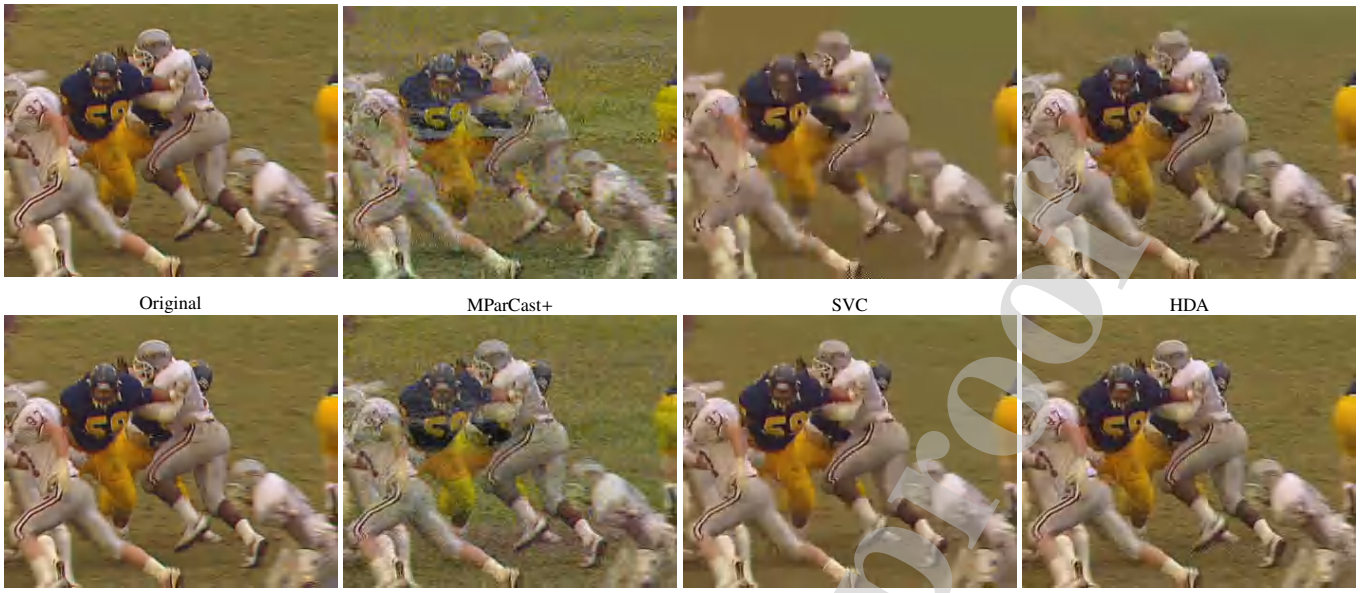


Figure 10: The original and reconstructed versions of the frame 241 of the Football sequence, when HDA, SVC, and Analog systems on (6,10) channel at 8 dB CSI-SNR (top row) and 12 dB CSI-SNR (bottom row).

CSNR), which have been chosen to ensure that no outages occurred. Large PSNR drops in the SVC system at 10 dB CSNR occur when the 1st refinement layer is sent only partially due to the lack of channel capacity during certain intervals. The issue is much less severe at 20 dB CSNR. The reconstruction quality of a digital refinement layer, unlike an analog refinement layer, does not scale smoothly with a varying channel capacity. The relatively poor performance of the MParCast+ system is due to lower analog coding gain and a relatively high amount of channel noise "seeping" into reconstructed video frames through the linear analog decoder. In [6], authors observe that their analog system also works well only at high CSNRs. As an example, we show in Fig. 6, the 241-st frame of the *Football* sequence transmitted on the (6,10) channel. Channel noise does not affect the HDA system to the same degree, as in this case the noise only appears in the quantization error and not the base layer. Note the less detailed "washed-out" background in the SVC-frame compared to that of the HDA-frame, caused by missing refinement layers.

Finally, we compare the performance the three systems under realistic conditions where there are CSI errors. In practice CSI-SNR will depend on the mean CSNR, the details of the CSI acquistion method, and the bit rate of the feedback link from the receiver to the transmitter. For the sake of our system comparisons, we have assumed that the total CSI error is Gaussian with a CSI-SNR close to the mean CSNR. Figs. 8 and 9 present comparisons of the three systems on the (6,10) system (8 dB or 12 dB CSI-SNRs) and the (4,20) system (17 dB or 23 dB CSI-SNRs) respectively. In addition, Fig. 11 presents a similar comparison for high-resolution 4CIF video sequences which require a much higher channel bandwidth than the CIF video considered in Figs. 8 and 9. Accordingly, the base layer bit rate of both HDA and SVC systems in this case have been set to 4 Mb/s. The (16,20) *channel* in Fig. 11 uses the same 20 dB channel trace as before, but with $K = 16$ subchannels.

Even though all three systems suffer performance losses as a result of using wrong CSI, the HDA system performs noticebly better than the SVC and MParCast+ systems. These plots show that the HDA system can outperform the 3-layer SVC system by about 3-6 dB in PSNR. In general, it was observed that the advantage of the HDA system is most significant during the video segments with high frame-to-frame variations. (e.g. rapid movements of objects or the camera.) Fig. 10 shows the frame 241 in the Football sequence, transmitted by the MParCast+, SVC, and HDA systems on the (6,10) channel. This particular frame comes from a segment of video with rapid movements

Figure 11: Performance of HDA, SVC, and MParCast+ systems for 4CIF resolution video on a (16,20) channel: (top row) 17 dB CSI-SNR, and (bottom row) 23 dB CSI-SNR. Base layer bit rate is 4 Mbps.

of a group of football players occupying much of the screen. When rapid inter-frame variations occur, the quantization error variance of the base layer increases. In this case, the HDA system transmits the most significant quantization errors all which is received by the decoder. The effect on the SVC system is the increased size of the refinement layers which can result in the transmission of only a part of a refinement layer, at times resulting in undesirable artifacts in reconstructed video. This is the main reason behind the inconsistency of the performance gap between the SVC and HDA systems across different video sequences. Figure 12 shows a close-up of this problem at the beginning of the 4CIF Crew sequence where a large PSNR gap exists between HDA and SVC systems.

6. Conclusion

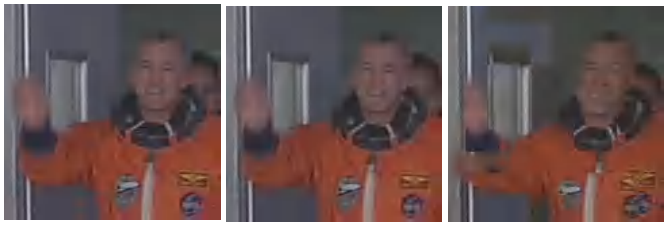
We have investigated the application of HDA coding to minimize the performance loss that can result from the mismatch between the CSI available at the transmitter/receiver and the actual channel state, in video transmission over OFDM channels. To this end we considered an approach to incorporating a CSI error-model into the power allocation problem and presented a close-form solution. Experimental results obtained with many test video sequences have consistently shown that, under realistic conditions where exact CSI cannot be known, HDA video coding with proper power allocation can outperform well known pure digital and pure analog alternatives.

References

[1] W. F. Schreiber, Advanced television systems for terrestrial broadcasting: Some problems and some proposed solutions, Proc. IEEE 83 (6) (1993) 958–981 (Jun. 1993).

[2] S. J. Wee, M. O. Polley, W. F. Schreiber, A scalable source coder for a hybrid HDTV terrestrial broadcasting system, in: IEEE Int. Conf. Image Processing (ICIP), 1994 (1994).

[3] I. Kozintsev, K. Ramachandran, Hybrid compressed-uncompressed



Original HDA SVC

Figure 12: A part of the original frame 6 of the 4CIF Crew sequence and the versions reconstructed using the HDA and SVC schemes. (16,20) channel at 17 dB CSI-SNR.

framework for wireless image transmission, in: IEEE ICIP, 1997, pp. 77–80 (1997).

[4] S. Jakubczak, D. Katabi, Softcast: Clean-slate scalable wireless video, Tech. Report MIT-CSAIL-TR-2011-008, M.I.T (2011).

[5] H. Cui, Z. Song, Z. Yang, C. Luo, R. Xiong, F. Wu, Cactus: A hybrid digital-analog wireless video communication, in: 16th ACM Int. Conf. Modelling, analysis & simulation of wireless and mobile systems, 2013, pp. 273–28 (2013).

[6] X. L. Liu, W. Hu, C. Luo, Q. Pu, F. Wu, Y. Zhang, Parcast+: Parallel video unicast in MIMO-OFDM WLANs, IEEE Trans. Multimedia 16 (7) (2014) 2038–2051 (Nov. 2014).

[7] H. Cui, C. Luo, C. W. Chen, F. Wu, Robust uncoded vodeo transmission over wireless fast fading channels, in: IEEE INFOCOM 2014, 2014, pp. 73–81 (2014).

[8] L. Yu, H. Li, W. Li, Wireless scalable video coding using hybrid digital-analog scheme, IEEE Trans. Circuits Syst. Video Technol. 24 (2) (2014) 331–345 (Feb. 2014).

[9] N. Fan, Y. Liu, L. Zhang, Hybrid digital-analog video transmission based on H.264/AVC and Parcast in MIMO-OFDM WLANs, in: 21st Int. Conf. Telecom. (ICT), 2014 (2014).

[10] Z. Song, R. Xiong, S. Ma, W. Gao, Hybridcast: A wireless image/video softcast scheme using layered representation and hybrid digital-analog modulation, in: IEEE ICIP, 2014, pp. 6001–6005 (2014).

[11] N. Fan, Y. Liu, Q. Wang, L. Zhang, Hybrid digital-analog video multicast scheme based on H.264/AVC and softcast, in: 17th Int. Symp Wireless Personal Multimedia Com. (WMPC2014), 2014 (2014).

[12] X. Zhao, H. Lu, C. W. Chen, J. Wu, Adaptive hybrid digital-analog video transmission in wireless fading channel, IEEE Trans. Circuits Syst. Video Technol. 26 (6) (2015) 1117 – 1130 (Jun. 2015).

[13] B. Tan, X. Huang, J. Wu, P. Xia, An adaptive hybrid digital-analog modulation scheme, in: IEEE China SIP, 2015, pp. 766–771 (2015).

[14] T. Fujihashi, T. Koike-Akino, T. Watanabe, P. V. Orlik, Compressive sensing for loss-resilient hybrid wireless video transmission, in: IEEE GLOBECOM, 2015 (2015).

[15] F. Zhang, A. Wang, H. Wang, S. Li, X. Ma, Channel-aware video softcast scheme, in: IEEE China SIP, 2015, pp. 578–581 (2015).

[16] C. L. D. He, C. Lan, F. Wu, W. Zeng, Structure-preserving hybrid digital-analog video delivery in wireless networks, IEEE Trans. Multimedia 17 (9) (2015) 945–956 (Sep. 2015).

[17] H. Cui, D. Liu, Y. Han, J. Wu, Robust uncoded video transmission under practical channel estimation, in: IEEE GLOBECOM, 2016 (2016).

[18] Y. Liu, X. Lin, N. Fan, L. Zhang, Hybrid digital-analog video transmission in wireless multicast and multiple-input multiple-output system, Journal of Electronic Imaging 25 (2016) 1–12 (2016).

[19] B. Tan, H. Cui, J. Wu, C. W. Chen, An optimal resource allocation for superposition coding-based hybrid digital-analog system, IEEE Internet of Things Journal 4 (4) (2017) 945–956 (Aug. 2017).

[20] D. He, C. Lan, C. Luo, E. Chen, F. Wu, W. Zeng, Progressive pseudo-analog transmission for mobile video streaming, IEEE Trans. Multimedia 19 (8) (2017) 1894–1907 (Aug. 2017).

[21] Z. Zhang, D. Liu, X. Ma, X. Wang, Ecast: An enhanced video transmission design for wireless multicast systems over fading channels, IEEE Systems Journal 11 (4) (2017) 2566 – 2577 (Dec. 2017).

[22] B. Tan, J. Wu, Y. Li, H. Cui, W. Yu, C. W. Chen, Analog coded softcast: A network slice design for multimedia broadcast/multicast, IEEE Trans. Multimedia 19 (10) (2017) 2293–2306 (Oct. 2017).

[23] D. Liu, J. Wu, H. Cui, D. Zhang, C. Luo, F. Wu, Cost-distortion optimization and resource control in pseudo-analog visual communications, IEEE Trans. Multimedia 20 (11) (2018) 3097– 3110 (Nov. 2018).

[24] J. Shen, F. Liang, C. Luo, H. Li, W. Zeng, Cooperative hybrid digital-analog video transmission in D2D networks, in: IEEE ICIP, 2018, pp. 3274– 3278 (2018).

[25] C. Lan, C. Luo, W. Zeng, F. Wu, A practical hybrid digital-analog scheme for wireless video transmission, IEEE Trans. Circuits Syst. Video Technol. 28 (7) (2018) 1634 – 1647 (Jul. 2018).

[26] T. Fujihashi, T. Koike-Akino, T. Watanabe, P. V. Orlik, High-quality soft video delivery with GMRF-based overhead reduction, IEEE Trans. Multimedia 20 (2) (2018) 472–483 (Feb. 2018).

[27] F. Liang, C. Luo, R. Xiong, W. Zeng, F. Wu, Hybrid digital-analog video delivery with Shannon-Kotelnikov mapping, IEEE Trans. Multimedia 20 (8) (2018) 2138– 2152 (Aug. 2018).

[28] F. Liang, C. Luo, R. Xiong, W. Zeng, F. Wu, Superimposed modulation for soft video delivery with hidden resources, IEEE Trans. Circuits Syst. Video Technol. 28 (9) (2018) 2345 – 2358 (Sep. 2018).

[29] J. Zhang, A. Wang, J. Liang, H. Wang, S. Li, X. Zhang, Distortion estimation-based adaptive power allocation for hybrid digital-analog video transmission, IEEE Trans. Circuits Syst. Video Technol. 29 (6) (2019) 1806–1818 (Jun. 2019).

[30] L. Ramachandran, A. Ortega, M. Vetterli, Bit allication for dependent quantization with application to MPEG video coders, IEEE Trans. Image Process. 3 (5) (1994) 533–545 (Sep. 1994).

[31] D. Marpe, T. Wiegand, G. J. Sullivan, The H.264/MPEG4 advanced video coding standard and its applications, IEEE Trans. Commun. (2006) 134–143 (Aug. 2006).

[32] I. E. Richardson, The H.264 Advanced Video Compression Standard, 2nd Edition, John Wiley, 2010 (2010).

[33] L. Yu, S. Chen, J. Wang, Overview of AVS-video coding standards, Signal Processing: Image Communication 24 (4) (2009) 247–262 (Apr. 2009).

[34] U. Mittal, N. Phamdo, Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communication, IEEE Trans. Inf. Theory 48 (5) (2002) 1082–1102 (May 2002).

[35] Supplement to IEEE Standard for Information technology, Part 11: Wireless LAN Medium Access control (MAC) and Physical Layer (PHY) specifications: High-speed Physical Layer in the 5 GHZ Band, IEEE Std 802.11-1999.

[36] D. Schafhuber, G. Matz, MMSE and adaptive prediction of time-varying channels for OFDM systems, IEEE Trans. Wireless Commun. 4 (2) (2005) 593 – 602 (Mar. 2005).

[37] J. G. Andrews, A. Ghosh, R. Muhamed, Fundamentals of WiMAX: Understanding Broadband Wireless Networking, Prentice-Hall, 2007 (2007).

[38] Y. Pei, D. W. Modestino, Multi-layered video transmission over wireless channels using an adaptive modulation and coding scheme, in: Proc. Int. Conf. Image Processing (ICIP) 2001, 2001, pp. 1009–1012 (2001).

[39] H.264/SVC JSVM Reference Software, Available at `https://www.hhi.fraunhofer.de/en/departments/vca.html`.

[40] R. Narasimhan, Performance of diversity schemes for OFDM systems with frequency offset, phase noise, and channel estimation errors, IEEE Trans. Commun. 50 (11) (2002) 1561–1565 (Oct. 2002).

[41] S. Ye, R. S. Blum, L. J. C. Jr., Adaptive OFDM systems with imperfect channel state information, IEEE Trans. Wireless Commun. 5 (11) (2006) 3255–3265 (Nov. 2006).

[42] G. Caire, N. Jindal, M. Kobayashi, N. Ravindran, Multiuser MIMO achievable rates with downlink training and channel state feedback, IEEE Trans. Inf. Theory 57 (6) (2010) 2845–2866 (Jun. 2010).

[43] D. Halperin, W. Hu, A. Sheth, D. Wetherall, Predictable 802.11 packet delivery from wireless channel measurements, in: Proc. ACM SIGCOMM, 2010, pp. 159–170 (2010).

[44] T. M. Cover, J. A. Thomas, Elements of Information Theory, 2nd Edition, John Wiley, 2006 (2006).

[45] H.264/AVC JM Reference Software, Available at `https://www.hhi.fraunhofer.de/en/departments/vca.html`.

[46] M. O. Polley, S. J. Wee, W. F. Schreiber, Hybrid channel coding for multiresolution for HDTV terrestrial broadcasting, in: IEEE Int. Conf. Image Processing (ICIP), 1994 (1994).

[47] X. Jiang, H. Lu, Joint rate and resource allocation in hybrid digital-analog transmission over fading channels, IEEE Trans. Veh. Technol. 67 (10) (2018) 9528 – 9541 (Oct. 2018).

[48] G. L. Stuber, Principles of Mobile Communications, 2nd Edition, Kluwer Academic Publishers, 2002 (2002).

[49] P. Yahampath, Digital-analog superposition coding for OFDM channels with application to video transmission, in: IEEE Int. Conf. Acostics, Speech and Sig. Proc. (ICASSP), 2018, pp. 1802–1806 (April 2018).

[50] J. A. Roese, W. K. Pratt, G. S. Robinson, Interframe cosine transform image coding, IEEE Trans. Commun. 25 (11) (1977) 1329–1339 (Nov. 1977).

[51] G. S. Yovanof, S. Liu, Statistical analysis of the DCT coefficients and their quantization error, in: The 30th Asilomar Conf. on Signals, Systems and Computers, 1996, pp. 601–605 (1996).

[52] S. M. Kay, Fundamentals of Statistical Signal Processing: Estimation Theory, Prentice Hall, 1993 (1993).

[53] E. T. S. Institute, Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television, ETSI EN 300 744 V1.6.1 (2009-01), 2009 (2009).

[54] M. Hernandez, H.-B. Li, I. Dotlic, R. Muira, Channel models for TG8, Tech. report, IEEE P802.15 Working Group for Wireless Personal Area Networks (2012).

[55] Z. Yang, X. Wang, Scalable video broadcast over downlink MIMO-OFDM systems, IEEE Trans. Circuits Syst. Video Technol. 23 (2) (2013) 212–223 (Feb. 2013).

Highlights:

- A new hybrid digital-analog video coding approach for OFDM systems with noisy channel state information is presented.
- A water-filling solution to optimal power allocation problem is obtained by considering a noise model for channel state information.
- Experimental results show that proposed scheme achieves a much better video quality-vs-reliability trade-off, compared to the best known digital-only and analog-only alternatives.

We wish to draw the attention of the Editor to the following facts which may be considered as potential conflicts of interest and to significant financial contributions to this work. [OR]
We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us.

We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

We further confirm that any aspect of the work covered in this manuscript that has involved either experimental animals or human patients has been conducted with the ethical approval of all relevant bodies and that such approvals are acknowledged within the manuscript.

We understand that the Corresponding Author is the sole contact for the Editorial process (including Editorial Manager and direct communications with the office). He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs. We confirm that we have provided a current, correct email address which is accessible by the Corresponding Author and which has been configured to accept email from Pradeepa.Yahampath@umanitoba.ca

Signed by all authors as follows:

Pradeepa Yahampath