Loss Visibility Optimized Real-Time Video Transmission Over MIMO Systems

Amin Abdel Khalek, *Member, IEEE*, Constantine Caramanis, *Member, IEEE*, and Robert W. Heath, *Fellow, IEEE*

***Abstract*—The structured nature of video data motivates introducing video-aware decisions that make use of this structure for improved video transmission over wireless networks. In this paper, we introduce an architecture for real-time video transmission over multiple-input multiple-output (MIMO) wireless communication systems using loss visibility side information. We quantify the perceptual importance of a packet through the packet loss visibility and use the loss visibility distribution to provide a notion of relative packet importance. To jointly achieve high video quality and low latency, we define the optimization objective function as the throughput weighted by the loss visibility of each packet, a proxy for the total perceptual value of successful packets per unit of time. We solve the problem of mapping video packets to MIMO subchannels and adapting per-stream rates to maximize the proposed objective. We show that the solution enables jointly reaping gains in terms of improved video quality and lower latency. Optimized packet-stream mapping enables transmission of more relevant packets over more reliable streams while unequal modulation opportunistically increases the transmission rate on the stronger streams to enable low latency delivery of high priority packets. Tested on H.264-encoded video sequences, for a 4×4 MIMO system with three spatial streams, the proposed architecture achieves 8 dB power reduction for the same video quality and supports $2.4\times$ higher throughput due to unequal modulation. Furthermore, the gains are achieved at the expense of few bits of cross-layer overhead rather than a complex cross-layer design.**

***Index Terms*—Limited feedback, loss visibility, multiple-input multiple-output (MIMO), packet loss, unequal error protection, video signal processing.**

I. Introduction

THE delay-sensitive nature of real-time video transmission motivates the use of unreliable transport protocols, such as user datagram protocol (UDP) for video delivery. This causes the wireless channel impairments, such as losses and delays, to be visible at the APP layer. Consequently, achieving good overall video quality for real-time video requires mitigating channel-induced distortions. Since video quality is the metric of interest from the user perspective, transmission policies should be designed to minimize the impact of losses on video quality. Generally, incorporating video quality-based optimization into lower layer protocols requires a complex, and practically prohibitive, cross-layer design that jointly adapts the video server and the base station. In this paper, we incorporate video quality based optimization into the network while maintaining strict inter-layer separation and avoiding a complex feedback mechanism between the network and the server. Instead, we propose an easy-to-implement feedforward mechanism that communicates packet loss visibility to lower layers to enable video quality optimization. At the cost of few additional bits that can be embedded in the packet header, video quality-based optimization is enabled by prioritizing video packets at the PHY layer based on perceptual relevance.

The response to video packet losses and distortions is inherently unequal due to the features of state-of-the-art video codecs (e.g. [1], [2]) such as inter-frame coding, motion compensation, and error concealment. For example, inter-frame coding introduces packet dependencies in the temporal domain, thus causing different error propagation patterns, and increasing the loss visibility variability. Furthermore, the non-uniform motion across different spatial locations causes loss visibility to be unequal across slices and dependent on the error concealment method. Video packet loss visibility captures this unequal response by training a statistical model that maps a set of features per packet to a measure of visibility of that packet loss. More formally, video packet loss visibility is defined as the probability that the artifact due to the loss of a given packet is visible to the average user. The objective of loss visibility modeling and estimation (e.g. [3], [4]) is to find the model that best correlates the loss visibility estimate with the results reported through subjective tests, thus naturally capturing the viewer perception. Quantizing the loss visibility information and embedding it into the packet headers provides an inexpensive tool for perceptual quality optimization.

Advanced PHY layer designs, such as multiple-input multiple-output (MIMO) processing, have become an integral part of state-of-the-art wireless standards such as 3GPP Long Term Evolution (LTE) and IEEE 802.11n, which deliver the bulk of stored and real-time video traffic. In this paper, we leverage the spatial degrees of freedom of the MIMO channel to map video packets to MIMO subchannels based on channel quality and packet loss visibility. In short, the proposed technique makes use of the unequal gains of MIMO substreams to provide unequal protection of video packets resulting in a video quality gain. Jointly, unequal modulation is leveraged on the better streams, resulting in a throughput gain and timely delivery of

Manuscript received February 08, 2015; revised May 31, 2015; accepted August 02, 2015. Date of publication August 12, 2015; date of current version September 15, 2015. This work was supported by the Intel-Cisco Video Aware Wireless Networks (VAWN) Program. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Yap-Peng Tan.

The authors are with the Wireless Networking & Communications Group, Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712-1687 USA (e-mail: akhalek@utexas.edu; constantine@utexas.edu; rheath@utexas.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/TMM.2015.2468196

perceptually relevant packets. Consequently, packet prioritization is achieved both in terms of reliability and rate. The major contributions in this paper are summarized as follows.

A. Paper Contributions

1) Low Overhead Video-Aware PHY Optimization: We propose a new low overhead architecture for real-time video transmission to mitigate channel-induced video distortions. Our proposed architecture uses quantized loss visibility scores embedded in the packet header at the expense of only few extra bits per packet while avoiding a complex cross-layer design. We argue that the loss visibility scores of buffered video packets are not sufficient to fully capture the loss visibility variability since real-time video only supports small buffers. Thus, we also estimate the loss visibility distribution inexpensively to capture this variability and provide a notion of relative packet importance that is used in optimizing transmission decisions.

2) Packet Prioritization for High Quality and Low Latency: At the PHY layer, we propose to use the loss visibility values to classify video packets into different priority classes. To optimize the loss visibility-based transmission policy for high video quality and low latency, we define an optimization metric that generalizes the conventional notion of throughput by weighting each packet in the optimization objective by its loss visibility. Since loss visibility reflects the visual perception of a corresponding packet loss, our optimization metric is a proxy for *the total perceptual value of packets successfully delivered per unit time*. Given the proposed objective function that enables joint optimization of video quality and latency, we derive optimized PHY layer packet prioritization schemes. We emphasize that the proposed metric is used for optimization rather than evaluation of the algorithm. For assessment of video quality gains, we use objective video quality metrics.

3) Loss Visibility Optimized MIMO Precoding: For a MIMO system, each class of packets is transmitted through a different spatial stream corresponding to a decomposed subchannel of the MIMO channel. We derive the optimal packet-stream mapping that maximizes the loss visibility weighted throughput objective. The solution can be summarized as follows: (1) The MIMO channel is decomposed into parallel streams, (2) the per-stream transmission rate, i.e. modulation order, is chosen to maximize the corresponding throughput per stream, (3) the spatial streams are ordered by their probability of packet error, a function of both the per-stream SNRs and (potentially unequal) modulation orders, (4) the packets are classified according to a thresholding policy whereby higher priority packets are mapped to high order streams as defined by the ordering in (3). The optimal thresholding policy is such that the load is balanced across streams based on the fraction of packets per priority class, the modulation order per stream, and the retransmission overhead. We show that the solution enables jointly reaping gains in terms of improved video quality and lower latency: A packet prioritization gain results from transmission of more relevant packets over more reliable streams and an unequal modulation gain results from opportunistically increasing the transmission rate on the stronger streams to enable low latency delivery of high priority packets.

B. Related Work

We review related work on loss visibility estimation and modeling [3]–[5], loss visibility based optimization [6], and adaptive MIMO transmission for video content [7]–[12]. In [3], a generalized linear model is proposed for video packet loss visibility modeling considering factors within a packet and its temporal and spatial vicinity to capture the temporal and spatial distortions. The set of features used to estimate loss visibility is versatile by being applicable over a range of encoding standards, GoP structures, and error concealment methods. Some features such as motion magnitude, motion variance, distance from scene cut, and camera motion capture the video source properties. Other features such as initial structural similarity index (SSIM), maximum per-macroblock (MB) mean square error (MSE), and spatial extent capture the distortions caused by the loss in the spatial domain. Temporal error propagation is also captured through features related to the number of frames affected by the loss, distance to reference frame, error concealment method, and other scene loss concealment. The generalized linear model using these features is fit based on subjective tests. Other related loss visibility modeling approaches can be found in [4] and [5]. Besides generalized linear models, [4] proposes a classification-based approach using a statistical tool called classification and regression trees (CART) to classify each packet loss as visible or invisible. The loss visibility model developed in [4] is applied in [6] for selecting unequal coding rates for different slices and for resource allocation in an OFDM system.

In this paper, we propose a generic framework that allows using loss visibility models to optimize transmission policies at the PHY and MAC protocol layers. Specifically, we apply the generalized linear modeling approach in [3] for loss visibility estimation of encoded sequences due to its versatility and high classification accuracy. We further argue that the loss visibility distribution provides a notion of relative packet importance for real-time video where only a small number of packets are buffered, and thus, we propose to inexpensively learn and update the distribution, and subsequently use it in loss visibility based adaptation.

While loss visibility-based adaptation approaches are not heavily investigated in the literature, other adaptive video transmission techniques such as joint source-channel coding (JSCC) [13]–[16], unequal error protection (UEP) [17]–[19], and prioritized scheduling [20], and distortion-aware resource allocation [21], [22] have been proposed to increase video quality and error resilience. Previous work, however, does not present a generic framework for incorporating loss visibility-based decisions into wireless networks. To the best of our knowledge, this is the first comprehensive work that defines a generic cross-layer design for using loss visibility in wireless networks, develops MIMO transmission strategies for prioritized delay-sensitive video delivery, and derives corresponding closed-form gain expressions.

Adaptive MIMO transmission for video content has been investigated in [7]–[10]. In [7], a cross-layer framework for MIMO video broadcast is proposed by allocating scalable video layers to the end-users jointly with precoder computation to ensure that delay and buffer constraints are met. In [8], a layered video transmission scheme over MIMO is proposed. It periodically switches each bit stream among multiple antennas to match the ordering

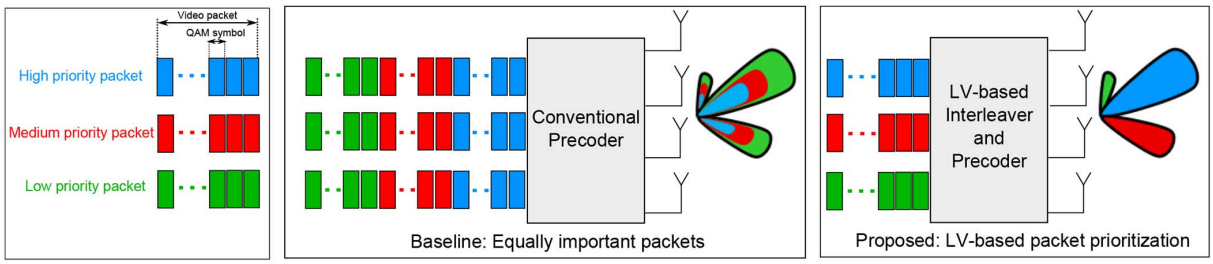

Fig. 1. Illustration of the proposed precoder and interleaver design for packet prioritization over MIMO channels.

of subchannel SNRs, thus providing prioritized delivery. In [9], a method is proposed to adaptively control the diversity and multiplexing gain of a MIMO system to minimize the cumulative video distortion and satisfy delay constraints. Finally, in [10], distortion-aware MIMO link adaptation techniques are proposed for MCS and MIMO mode selection. Since [7], [8] are only applicable to scalable video coded bitstreams, the application scope of the proposed techniques is limited as the majority of current video content is non-scalable. Furthermore, [9], [10] relies on rate-distortion information which is typically not available for real-time encoded or transcoded video.

II. System Model

This section introduces the proposed MIMO system model that enables loss visibility-based packet prioritization as well as the model for the APP, MAC, and PHY layers.

A. Prioritized MIMO Transmission

Consider P packets buffered for transmission where packet p is represented as $\mathbf{s}_p = [s_p[1], \ldots, s_p[b(\mathbf{s}_p)]]$ where $b(\mathbf{s}_p)$ is the number of QAM symbols. The vector of symbols corresponding to all buffered packets is denoted $\mathbf{s} = [\mathbf{s}_1, \ldots, \mathbf{s}_P]^T$.

Consider a narrowband MIMO wireless system with N_t transmit antennas and N_r receive antennas. The system uses S spatial streams where $S \leq \min(N_\mathrm{t}, N_\mathrm{r})$ and each stream corresponds to a stream of constellation symbols. Our general framework enables the size of the constellation to vary per substream, as well as the number of substreams, known as mode adaptation. Thus, we have $1 \leq S \leq \min(N_\mathrm{t}, N_\mathrm{r})$. Linear precoding enables mapping a symbol vector from each spatial stream to an N_t-dimensional spatial signal using an $N_\mathrm{t} \times S$ linear precoding matrix $\mathbf{F}_S$. The spatial signal encounters a channel matrix $\mathbf{H}$ and an additive noise vector $\mathbf{n}$ with elements each distributed according to $\mathcal{CN}(0, N_0)$. The corresponding input-output relationship is

$$\mathbf{y}[i] = \sqrt{\frac{E_\mathrm{s}}{N_\mathrm{t}}}\mathbf{H}\mathbf{F}_S\mathbf{T}[i]\mathbf{s} + \mathbf{n}[i] \tag{1}$$

where $\mathbf{y}[i]$ is the received signal and $\mathbf{T}[i]$ is an interleaver matrix that determines the mapping between symbols and spatial streams in the ith channel use and is proposed to enable loss visibility-based prioritized transmission. Note that $\mathbf{T}[i]$ has dimensions $S \times \sum_p b(\mathbf{s}_p)$. Conventionally, in the absence of loss visibility information, the symbols are transmitted sequentially. Thus, the interleaver for the ith channel use can be represented mathematically as

$$\mathbf{T}[i] = \left[\mathbf{0}_{S,(i-1)S} | \mathbf{I}_S | \mathbf{0}_{S,\sum_p b(\mathbf{s}_p) - iS}\right] \tag{2}$$

where $\mathbf{0}_{m,n}$ is an all zeros $m \times n$ matrix and $\mathbf{I}_m$ is an $m \times m$ identity matrix. In this paper, we propose designing an interleaver matrix that provides packet prioritization based on loss visibility. Consider a classification policy whereby a set of packets $\mathcal{V}_m$ is classified into priority level m corresponding to packets transmitted through spatial stream m. The following interleaver design ensures that packets $p \in \mathcal{V}_m$ are transmitted through stream m

$$\mathbf{T}[1]_{m,n} = \begin{cases} 1 & \text{if } n = 1 + \sum_{j=1}^{m-1} \sum_{p \in \mathcal{V}_m} b(\mathbf{s}_p) \\ 0 & \text{otherwise.} \end{cases}$$
$$\mathbf{T}[i+1] = \left[\mathbf{0}_{S,1} | \mathbf{T}[i]_{1:S,1:\sum_p b(\mathbf{s}_p)-1}\right]. \tag{3}$$

The matrix $\mathbf{T}[1]$ selects the first S QAM symbols from S packets mapped to each of the spatial streams. In the next channel use, $\mathbf{T}[i+1]$ is obtained from $\mathbf{T}[i]$ by rotating the rows one position to the right. This ensures that the next transmission captures the next S QAM symbols from each of the S packets. The process is repeated for each subset of packets in the buffer mapped to different spatial streams. This implementation is useful for practical signal processing purposes, whereby the interleaver matrix in (3) is updated inexpensively by "sliding" the interleaver from the previous channel use. The resulting mapping is illustrated in Fig. 1 and the physical interpretation of the process is that high priority packets are sent over the more reliable MIMO subchannels.

We note that if the set of packets buffered P is larger than the channel coherence time multiplied by the packet throughput, then the prioritization policy is applied on a subset of the packets that can be transmitted within a single coherence time. The policy is then updated with every new channel estimate and new subset of the buffered packets are transmitted.

Given the simple interleaving procedure in (3) that enables packet prioritization, the main question we address in subsequent sections is how to determine the classification policy, i.e., given a set of P packets with loss visibility values $\{v(\mathbf{s}_i)\}_{i=1}^P$, how to determine the priority sets $\mathcal{V}_i$ to maximize a video quality-based utility function.

B. Precoder Design

The matrix $\mathbf{H}\mathbf{F}_S$ can be thought of as an effective channel. The receiver decodes $\mathbf{y}$ using this effective channel and a zero

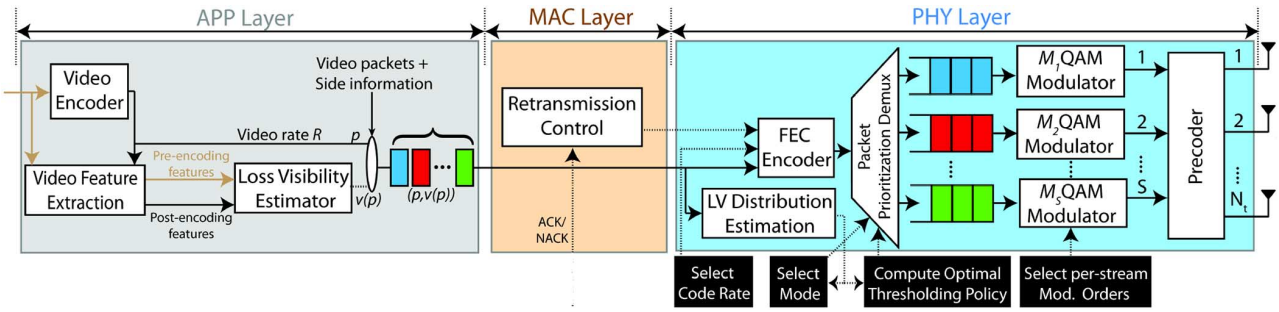

Fig. 2. System block diagram for loss visibility-based prioritized MIMO transmission.

forcing receiver. Each entry of the channel matrix $\mathbf{H}$ corresponds to a flat fading channel. Further, we assume a block-fading model whereby the channel coherence time is fixed over the duration of at least a single packet transmission and then independently takes a new realization. All the transmission decisions are adapted every channel coherence time which could be as small as one packet duration, thus being applicable over a range of mobility scenarios. For a zero forcing receiver, it is shown in [23] that the SNR on the ith stream is

$$\gamma_i(\mathbf{H}) = \frac{E_s}{N_0} \frac{1}{[\mathbf{F}_S^* \mathbf{H}^* \mathbf{H} \mathbf{F}_S]^{-1}_{i,i}}. \tag{4}$$

We consider both cases of perfect and imperfect, i.e., quantized transmitter channel state information (CSIT). In both scenarios, we assume that the feedback delay is negligible and the transmitter and receiver are fully synchronized. With perfect CSIT, the MIMO channel can be converted to parallel, noninterfering single-input single-output (SISO) channels through a singular value decomposition (SVD) of the channel matrix [24]. We consider unitary precoding whereby the columns of $\mathbf{F}_S$ are restricted to be orthogonal. While this could be further generalized to a non-unitary power constraint, we note that using the unitary constraint along with multimode precoding results in performance near the capacity achieved by waterfilling. [25]. Thus, we create $\mathbf{F}_S$ from a normalized version of the right singular vectors of $\mathbf{H}$ as follows:

$$\mathbf{F}_S = \frac{1}{\sqrt{S}} [\mathbf{V}]_{:,1:S} \tag{5}$$

where $\mathbf{H} = \mathbf{U}\Sigma\mathbf{V}^*$ is the singular value decomposition of $\mathbf{H}$ and the notation $[\mathbf{A}]_{:,i:j}$ refers to the submatrix formed by columns i to j of matrix $\mathbf{A}$. Under the precoding structure in (3), the SNR for the ith stream simplifies to

$$\gamma_i(\mathbf{H}) = \frac{E_s}{N_0} \frac{\sigma_i^2}{S} \tag{6}$$

where σ_i is the ith singular value of $\mathbf{H}$. For quantized CSIT, the receiver chooses a precoding matrix $\mathbf{F}_S$ from a codebook $\mathcal{F}_S$ consisting of a finite set of precoding matrices. There are $\log_2(|\mathcal{F}_S|) = B_S$ bits of feedback used to convey the index of the chosen precoding matrix back to the transmitter if S spatial streams are used. For simulations, the codebook $\mathcal{F}_S$ is designed using Grassmannian subspace packing with the chordal subspace distance measure as described in [26]. The criterion for selecting the precoder at the receiver is to maximize the minimum singular value, that is, $\mathbf{F}_S = \arg\max_{\mathbf{F}\in\mathcal{F}} \lambda_{\min}(\mathbf{HF})$.

C. Modulation, Coding, and Retransmission

We apply unequal modulation per stream. The data through stream i are modulated with a QAM constellation of size $M_i \in \mathcal{M}$ resulting in a data rate $R_i = B\log_2 M_i$. Each constellation is normalized such that the average symbol energy is unity. Given a channel realization, the vector of modulation schemes is denoted $\mathbf{M} = \{M_i\}_{i=1}^{i=S}$. The set of coding rates is $\mathcal{C}$ and the data through all streams are coded with coding rate $C \in \mathcal{C}$.

The probability of packet error through stream i conditioning on the modulation scheme M_i, the coding rate C, and the ith post-processing SNR $\gamma_i(\mathbf{H})$ is denoted $\alpha_i = \mathrm{PER}(M_i, C, \gamma_i(\mathbf{H}))$. While we use the notation α_i for brevity, the dependence on the modulation order, coding rate, and SNR is implied. The uncoded M-QAM error probability expressions $\mathrm{PER}_{\mathrm{uncoded}}(M, \gamma)$ are provided in the literature (e.g. [27]). Given a set of channel codes $\mathcal{C}$, we estimate the coding gain $g(C)$ of each particular code $C \in \mathcal{C}$ using Monte-Carlo simulations to deduce the coded PER expression $\mathrm{PER}(M_i, C, \gamma_i(\mathbf{H})) = \mathrm{PER}_{\mathrm{uncoded}}(M_i, \gamma_i(\mathbf{H}) + g(C))$.

Retransmission is modeled with a limit of L retransmissions, determined by the MAC protocol. The number of retransmissions follows a truncated geometric distribution assuming the channel is fixed during retransmission. Thus, the mean number of transmissions through stream i is

$$r_i = \sum_{k=1}^{L+1} k(1-\alpha_i)\alpha_i^{k-1} + (L+1)\alpha_i^{L+1} = \frac{1-\alpha_i^{L+1}}{1-\alpha_i} \tag{7}$$

since $(1-\alpha_i)\alpha_i^{k-1}$ is the probability of success in k transmissions and α_i^{L+1} is the post-retransmission failure probability. We define the post-retransmission probability of successful packet delivery through stream i as

$$p_i^{\mathrm{success}} = 1 - \alpha_i^{L+1}. \tag{8}$$

The complete system block diagram including APP layer loss visibility estimation, MAC layer retransmissions and PHY layer packet prioritization is shown in Fig. 2. Commonly used notation is summarized in Table I.

III. Loss Visibility Estimation and Loss Visibility-Based Optimization

In this section, we first present background on loss visibility estimation and present a framework for using loss visibility side

TABLE I
COMMONLY USED NOTATION

N_t	Number of transmit antennas
N_r	Number of receive antennas
S	Number of spatial streams
$f_v(v)$	Packet loss visibility distribution
V_i	Cumulative loss visibility values of class i packets
$\hat{\mathbf{v}} = \{\hat{v}_i\}_{i=2}^{S}$	Vector of loss visibility thresholds
$\gamma_i(\mathbf{H})$	Post-processing SNR on i^{th} stream
t_i	Mean time to transmit a class i packet
$\mathbf{M} = \{M_i\}_{i=1}^{i=S}$	Vector of modulation schemes per stream
$R_i = B\log_2(M_i)$	Data rate on stream i
$C \in \mathcal{C}$	Coding rate
α_i	PER for packets transmitted through stream i
α_{baseline}	PER for packets multiplexed through all streams
p_i^{success}	Prob. of successful packet delivery through stream i
$p_{\text{baseline}}^{\text{success}}$	Prob. of success by multiplexing through all streams
r_i	Avg. retransmissions for packets through stream i

information to characterize the video content. We further propose an optimization metric that uses loss visibility to jointly maximize video quality and network throughput.

A. Background: Loss Visibility Estimation

The objective of loss visibility estimation is to associate a packet p with a value $v(\mathbf{s}_p)$ ranging from 0 to 1 and indicating the *loss visibility* of the packet. A value $v(\mathbf{s}_p) = 0$ indicates that losing packet p does not have a visible impact on the end video quality whereas a value $v(\mathbf{s}_p) = 1$ indicates that the loss of packet p will be visible with probability 1. A PHY packet is composed of one or more slices. If the PHY packet is composed of multiple slices, the packet loss visibility is the mean of the individual slice visibility.

To estimate the loss visibility of APP layer slices, we use the generalized linear model (GLM) approach proposed in [3]. We extract video features both from the raw video reference as well as the encoded bitstream. We note that, for real-time video transmission, the raw video is available at the server since encoding is done in real-time. A video frame is divided into a set of slices, each corresponding to a horizontal group of MBs. We apply forward motion estimation to each MB to estimate the motion magnitude for each MB and compute the slice motion magnitude as the average per-MB motion magnitude. The residual energy for each MB is computed from the corresponding motion-compensated residual signal. By thresholding the average motion in the entire video frame, we detect if the scene consists of a still background or if there is camera motion. In addition to these features, we extract features from the encoded bitstream. Specifically, based on the frame type and the inter-frame prediction settings, we flag each packet as affecting one or multiple frames. To capture spatial-domain distortions, we further compute the initial SSIM feature corresponding to the SSIM in the frame affected by the loss, and max initial mean square error (IMSE) representing maximum per-MB MSE in the same frame. For videos sequences with multiple scenes, we detect scene cuts and use that to flag packets concealed using a reference corresponding to a previous scene for which losses are more visible. We also flag packets before scene cuts for which losses will be barely visible. Scene cuts are detected simply by comparing the residual energy between each two consecutive frames to a preset threshold. While other features are defined in [3], subjective tests show that only the ones mentioned above have high (positive or negative) correlation with loss visibility as reported by viewers. Using all these features, we use the following logistic regression model:

$$\log\left(\frac{v(\mathbf{s}_p)}{1-v(\mathbf{s}_p)}\right) = \beta_0 + \sum_{i=1}^{F}\beta_i x_{pi} \tag{9}$$

where $\beta = \{\beta_0, \beta_1, \ldots, \beta_F\}$ are the intercept and the coefficients associated with the different features. We use the coefficients as reported in Table IV in [3]. We assume the loss visibility $v(\mathbf{s}_p)$ of packet p is communicated to the physical layer through the packet header and deep packet inspection can be performed at the network edge to extract the loss visibility.

We note that the feature set used for estimating loss visibility is scalable across major coding standards, namely H.265, also known as High efficiency video coding (HEVC), and H.264. HEVC [28], [29] addresses the need for superior coding efficiency with the popularity of high definition (HD) and beyond-HD formats (e.g., 4 k × 2 k resolution). It uses a hierarchical block structure with coding blocks, prediction blocks, and transform blocks to provide more flexibility in coding decisions and improve coding efficiency. However, the underlying block-based coding approach remains the same as H.264. Thus, block-based features that capture spatial domain distortions such as initial SSIM and IMSE can be similarly computed. Furthermore, both represent P and B frame blocks by a motion vector specifying the reference block location and a residual signal, and using quarter-sample motion compensation prediction. Thus, features that capture temporal domain distortions relating to frame type, motion magnitude, and error concealment can be similarly extracted.

Our system model allows for unequal packet sizes and the packet value $v(\mathbf{s}_p)$ and the packet size $b(\mathbf{s}_p)$ may in general be correlated, as is the case in practice. We assume, however, that if $v(\mathbf{s}_1) > v(\mathbf{s}_2)$, then $b(\mathbf{s}_1) > b(\mathbf{s}_2)$. This is typically the case since low visibility packets (e.g. B frame packets) are predictively encoded, and thus compressed more efficiently.

B. Loss Visibility Distribution Estimation

Over a sufficiently long timescale, the distribution of the loss visibility values characterizes the video source and the codec. For instance, a GoP structure $IBPBP$ … results in a larger concentration of low visibility packets than $IPPPP$ …. Thus, we estimate the loss visibility distribution to be used in packet classification. We propose to estimate the *loss visibility distribution* using kernel density estimation (KDE) [30], update it with incoming packets, and use it to derive the optimal packet prioritization policy. With KDE, the density estimate at $0 \le x \le 1$, denoted by $\hat{f}_v(x)$, is

$$\begin{aligned}\hat{f}_v(x) &= \sum_{i=1}^{W}\frac{K_h(x - v(\mathbf{s}_{p-i}))}{W} \\ &= \frac{1}{Wh}\sum_{i=1}^{W}K\left(\frac{x - v(\mathbf{s}_{p-i})}{h}\right)\end{aligned} \tag{10}$$

where W is the window corresponding to the number of packets over which the estimate is obtained and $K_h(\cdot)$ is a kernel with smoothing parameter $h > 0$. Adjusting the kernel density estimation window W and smoothing parameter h provides a

bias/variance tradeoff between factual estimation of the loss visibility and fine adaptation to changing video characteristics. The distribution is inexpensive to compute and update as it only consists of linear operations.

The main advantage of using the loss visibility distribution is that for real-time video, where large buffers are not available, the buffered packet values are not fully representative of the loss visibility variability. Thus, the loss visibility distribution is used instead to capture this variability and provide a notion of relative packet importance.

C. LV-Weighted Throughput: An Optimization Metric

To jointly capture the two desirable objectives of high video quality and low latency video delivery, we propose optimizing throughput weighted by per-packet loss visibility. This generalizes the conventional notion of throughput to unequally important packets. Maximizing loss visibility-weighted throughput is equivalent to *maximizing the total perceptual value of packets successfully delivered per unit time*. This enables composite gains in video quality and throughput. The expression is

$$WT = \frac{\sum_v q^{\text{success}}(v) v}{t(\mathbf{H}, \mathbf{M}, C, \{\mathcal{V}_i\}_{i=1}^S)} \tag{11}$$

where $q^{\text{success}}(v)$ is the probability that a packet with loss visibility v is successfully delivered (after potential retransmission), and $t(\mathbf{H}, \mathbf{M}, \{\mathcal{V}_i\}_{i=1}^S)$ is the time to transmit the packets given the packet-stream mapping $\{\mathcal{V}_i\}_{i=1}^S$, the channel matrix $\mathbf{H}$, modulation orders $\mathbf{M}$, and the coding rate C. The dependence of the success probability on the packet values is intended to capture general unequal error protection policies. In the proposed packet prioritization policy presented in § II-A, the expression reduces to

$$WT_{\text{prioritized}} = \frac{\sum_{i=1}^S p_i^{\text{success}}(\gamma_i(\mathbf{H}), \mathbf{M}, C) \sum_{v \in \mathcal{V}_i} v}{\max_i t_i(\gamma_i(\mathbf{H}), M_i, C, \mathcal{V}_i)} \tag{12}$$

since $q^{\text{success}}(v) = p_i^{\text{success}}(\gamma_i(\mathbf{H}), \mathbf{M}, C) = 1 - \alpha_i^{L+1}$ is the probability of post-retransmission successful packet delivery defined in (8) if $v \in \mathcal{V}_i$. Alternatively, for the baseline where no loss visibility side information is used, the loss visibility-weighted throughput expression is

$$WT_{\text{baseline}} = \frac{p^{\text{success}}(\mathbf{H}, \mathbf{M}, C) \sum_{v \in \cup_i \mathcal{V}_i} v}{t(\mathbf{H}, \mathbf{M}, C)} \tag{13}$$

where $p_{\text{baseline}}^{\text{success}}(\mathbf{H}, \mathbf{M}, C) = 1 - \alpha_{\text{baseline}}^{L+1}$ for the baseline case whereby each packet is multiplexed over all streams. We note that packet error rate in the baseline case α_{baseline} and the prioritized transmission case α_i can be related as follows. Consider a packet of b QAM symbols with a symbol error rate SER_i through stream i, in the prioritized transmission scenario, the packet error rate corresponding to transmission through stream i is $\alpha_i = 1 - (1 - \text{SER}_i)^b$ assuming symbol errors are uncorrelated in time. Alternatively, without packet prioritization, the packet is transmitted over b/S channel uses through all streams and the corresponding packet error rate is $\alpha_{\text{baseline}} = 1 - \prod_i (1 - \text{SER}_i)^{b/S}$ assuming symbol errors are uncorrelated in time and across spatial streams. Substituting for α_i, we obtain

$$\alpha_{\text{baseline}} = 1 - \prod_{i=1}^S (1 - \alpha_i)^{1/S}. \tag{14}$$

IV. Loss Visibility-Based Packet Prioritization

In this section, we formulate the prioritized video transmission problem over MIMO channels and we derive the optimal packet prioritization policy that maximizes the loss visibility-weighted throughput.

A. Problem Formulation

We propose to solve the problem

$$\max_{\{\mathcal{V}_i\}, \mathbf{M}, C, S} \quad WT_{\text{prioritized}}(\{\mathcal{V}_i\}, \mathbf{M}, C, S) \tag{15}$$

$$\text{s.t.} \quad \cup_{i=1}^S \mathcal{V}_i = [0, 1] \tag{16}$$

$$M_i \in \mathcal{M} \forall i = 1, \ldots, S; C \in \mathcal{C}. \tag{17}$$

The objective is to select the number of packet classes S and the mapping of the set of packets $\mathcal{V}_i$ to spatial stream i, as well as the modulation orders $\mathbf{M}$ and the coding rate C such that the weighted throughput objective is maximized.

B. Stream Ordering

First, we show that the set $\mathcal{V}_i$ that maximizes the proposed objective is obtained by ordering the spatial streams by the corresponding probability of error and mapping the packets onto the ordered streams according to a set of thresholds.

Lemma 1: The optimal packet-stream mapping is such that $\mathcal{V}_i$ has the form $\mathcal{V}_i = [\hat{v}_i, \hat{v}_{i+1}]$ where $\cup_{i=1}^S \mathcal{V}_i = [0, 1]$. Furthermore, for any two packets $\mathbf{s}_1$ and $\mathbf{s}_2$ s.t. $v(\mathbf{s}_1) < v(\mathbf{s}_2)$, $\mathbf{s}_1 \in \mathcal{V}_i$ and $\mathbf{s}_2 \in \mathcal{V}_k$ where $p_i^{\text{success}} \leq p_k^{\text{success}}$. It follows that the streams should be ordered by the probability of success $p_i^{\text{success}} \leq p_{i+1}^{\text{success}}$.

Proof: See Appendix A. ∎

Note that the ordering in Lemma 1 captures the effect of modulation, coding, retransmission, and channel state because p_i^{success} is a function of M_i, C, r, and $\boldsymbol{\gamma}(\mathbf{H})$. In fact, the result represents a generalization of SNR ordering to the case of unequal modulation per stream.

The classification policy reduces into a thresholding policy completely determined by the vector of thresholds $\hat{\mathbf{v}} = \{\hat{v}_i\}_{i=1}^{i=S+1}$. Furthermore, the constraint in (16) can be rewritten as $0 \leq \hat{v}_i \leq \hat{v}_{i+1} \leq 1$ where $\hat{v}_1 = 0$ and $\hat{v}_{S+1} = 1$ by definition. Thus, we have

$$WT_{\text{prioritized}} = \frac{\sum_{i=1}^S p_i^{\text{success}}(\gamma_i(\mathbf{H}), \mathbf{M}, C) \sum_{p \in \mathcal{V}_i} v(\mathbf{s}_p)}{\max_i t_i(\gamma_i(\mathbf{H}), M_i, C, \mathcal{V}_i)}. \tag{18}$$

Now, we expand (18) by writing $t_i(\gamma_i(\mathbf{H}), M_i, C, \mathcal{V}_i)$ in terms of the respective parameters. The time to transmit a packet through stream i is $b(\mathbf{s}_p)(1 - \alpha_i^{L+1})/(CR_i(1 - \alpha_i))$ where

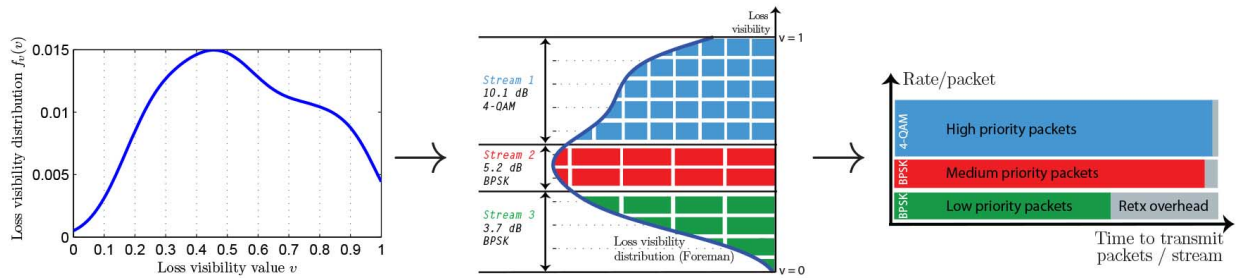

Fig. 3. Graphical illustration of loss visibility optimized transmission policy for $\boldsymbol{\gamma}(\mathbf{H}) = [10.1; 5.2; 3.7]$ dB and the Foreman video sequence: (a) obtain loss visibility distribution (shown for the Foreman video sequence); (b) decompose MIMO channel; (c) determine throughput-maximizing modulation order per stream; and (d) find the optimal thresholding policy. Note that high priority packets achieve both higher rate and reliability.

$b(\mathbf{s}_p)$ is the size of packet p. Taking the expectation over class i packets, we obtain

$$t_i(\gamma_i(\mathbf{H}), M_i, C, \mathcal{V}_i) = \mathbb{E}\left[\frac{b(\mathbf{s}_p)(1-\alpha_i^{L+1})}{CR_i(1-\alpha_i)}\right] \times (F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i)) = \frac{\mathbb{E}[b(\mathbf{s}_p)](1-\alpha_i^{L+1})}{CR_i(1-\alpha_i)} \times (F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i))$$

where $\mathbb{E}[b(\mathbf{s}_p)]$ is the mean packet size. Thus, the weighted throughput expression is

$$WT_{\text{prioritized}}(\hat{\mathbf{v}}, \mathbf{M}, C, S) = \frac{\left[\sum_{i=1}^{S}(1-\alpha_i^{L+1})\int_{\hat{v}_i}^{\hat{v}_{i+1}} v f_v(v)\mathrm{d}v\right]/\mathbb{E}[b(\mathbf{s}_p)]}{\max_i\{(F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i))(1-\alpha_i^{L+1})/CR_i(1-\alpha_i)\}} \quad (19)$$

$$= \underbrace{\frac{CR_{\tilde{i}}(1-\alpha_{\tilde{i}})}{\mathbb{E}[b(\mathbf{s}_p)](1-\alpha_{\tilde{i}}^{L+1})}}_{\text{Throughput component}} \underbrace{\frac{\sum_{i=1}^{S}(1-\alpha_i^{L+1})\int_{\hat{v}_i}^{\hat{v}_{i+1}} v f_v(v)\mathrm{d}v}{(F_v(\hat{v}_{\tilde{i}+1}) - F_v(\hat{v}_{\tilde{i}}))}}_{\text{Video quality component}} \quad (20)$$

where $\tilde{i} = \arg\max_i\{(F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i))(1-\alpha_i^{L+1})/CR_i(1-\alpha_i)\}$ denotes the stream with the longest average transmission time.

C. Optimal Thresholding Policy: A Load Balancing Solution

In this section, we derive the optimal thresholding policy $\hat{\mathbf{v}}^*$ for any continuous loss visibility distribution given the optimal ordering in § IV-B.

In Lemma 2 and Lemma 3, we derive properties of the gradient $\partial WT_{\text{prioritized}}/\partial\hat{v}_i$ that will be used to find the thresholds $\hat{v}_i$ that maximize the weighted throughput expression in Theorem 1.

Lemma 2: If the streams are ordered by the post-retransmission success probability, i.e., $p_i^{\text{success}} \leq p_{i+1}^{\text{success}} \forall i = 1, \cdots, N_s - 1$, then the gradient $\partial WT_{\text{prioritized}}/\partial\hat{v}_i$ satisfies the following properties:

1) $\partial WT_{\text{prioritized}}/\partial\hat{v}_{\tilde{i}} \geq 0$ where $\tilde{i} = \arg\max t_i$
2) $\partial WT_{\text{prioritized}}/\partial\hat{v}_i \leq 0 \forall i \neq \tilde{i}$

Proof: See Appendix B. ∎

We use Lemma 2 to derive a more general condition on the behavior of the gradient for the case where $\exists\tilde{j} \neq \tilde{i}$ s.t. $\tilde{i} = \tilde{j} = \arg\max t_i$, i.e., more than one stream have the same average transmission time. This extension will be key to proving the result in Theorem 1.

Lemma 3: Define $\mathcal{I} = \{\arg\max t_i\}$. If $\{\hat{v}_i; i \in \mathcal{I} \text{ or } i-1 \in \mathcal{I}\}$ are jointly scaled to keep $\mathcal{I}$ fixed, then

1) $\partial WT_{\text{prioritized}}/\partial\hat{v}_i \geq 0$ if $i \in \mathcal{I}$ and $i-1 \notin \mathcal{I}$
2) $\partial WT_{\text{prioritized}}/\partial\hat{v}_i \leq 0$ if $i \notin \mathcal{I}$ and $i-1 \in \mathcal{I}$

Proof: See Appendix C. ∎

Theorem 1 provides the optimal thresholding policy among streams and applies for any continuous loss visibility distribution obtained using kernel density estimation based on (10).

Theorem 1: Thresholding Policy: The optimal loss visibility thresholds $\hat{\mathbf{v}}^* = \{\hat{v}_i^*\}_{i=2}^{S}$ satisfy

$$F_v(\hat{v}_{i+1}^*) - F_v(\hat{v}_i^*) = \frac{R_i/r_i}{\sum_{j=1}^{S} R_j/r_j} \forall i = 1, \cdots, S \quad (21)$$

where $r_i = (1-\alpha_i^{L+1})/(1-\alpha_i)$.

Proof: See Appendix D. ∎

The solution is such that the post-retransmission throughput is equal among streams. Thus, the thresholds are selected to balance the load among spatial streams in proportion to the achievable throughput on each stream and the corresponding fraction of packets in each of the S classes. Correspondingly, the solution is referred to as the load balancing solution.

Fig. 3 illustrates the result for a specific channel realization and the Foreman video sequence. First, we show the loss visibility distribution obtained using kernel density estimation. Next, the MIMO channel is decomposed to obtain $\boldsymbol{\gamma}(\mathbf{H})$. Based on the SNR per stream, the throughput-maximizing constellation is chosen per stream. Given the loss visibility distribution, the constellation order, and the corresponding packet error rate, the set of thresholds are determined. The most prominent result in Fig. 3 is that the *high priority packets are sent with higher reliability (lower packet error rate / retransmission overhead) and lower latency (higher order constellation)*. Thus, utilizing the MIMO channel structure in the manner proposed enables both fewer errors and lower latency for the video packets that matter most making it particularly suitable for real-time video.

We further emphasize the cross-layer nature of the solution based on the components of (21) in the following three aspects:

1) *Non-uniform loss visibility distribution* (*APP*): The loss visibility thresholds are selected to balance the fraction of packets through each stream based on the loss visibility

distribution. In Fig. 3, this can be seen on the second stream where $\hat{v}_3 - \hat{v}_2$ is made small enough to compensate for the larger concentration of medium priority packets.

2) *Unequal modulation per stream* (*PHY*): If the SNR on stream i supports a higher modulation order M_i, the fraction of stream i packets is increased accordingly. In Fig. 3, this can be seen on the uppermost stream.
3) *Retransmission overhead* (*MAC*): If a particularly low SNR on spatial stream i incurs a large retransmission overhead r_i, the fraction of packets through stream i is reduced accordingly. In Fig. 3, this can be seen on the lowermost stream.

Under the load balancing solution in Theorem 1, we have

$$WT_{\text{prioritized}}(\hat{\mathbf{v}}^*, \mathbf{M}, C, S) = \underbrace{\frac{C}{\mathbb{E}[b(\mathbf{s}_p)]} \sum_{i=1}^{S} \frac{1-\alpha_i}{1-\alpha_i^{L+1}} R_i}_{\text{Post-retx sum throughput}} \underbrace{\left[\sum_{i=1}^{S} (1-\alpha_i^{L+1}) \int_{\hat{v}_i}^{\hat{v}_{i+1}} v f_v(v) \mathrm{d}v\right]}_{\text{Loss-penalized quality measure}}. \quad (22)$$

We note that for the special case of full retransmission, i.e., $L = \infty \forall i$, (22) reduces to the sum throughput as follows:

$$WT_{\text{prioritized}}(\hat{\mathbf{v}}^*, \mathbf{M}, C, S) = \frac{C}{\mathbb{E}[b(\mathbf{s}_p)]} \sum_{i=1}^{S} (1-\alpha_i) R_i. \quad (23)$$

In this limiting case of infinite retransmissions, all packets are eventually delivered reliably and providing packet prioritization on the basis of video quality becomes obsolete. Thus, the objective function reduces to throughput optimization.

We further note that although the packet prioritization policy is considered across spatial streams within a single coherence bandwidth, i.e., a single resource block (RB) in an OFDM system, the same solution can be readily extended to the frequency domain by sorting the per-RB effective SNR across RBs and allocating resources to different packets according to loss visibility as suggested by Theorem 1. In a similar manner, resources can also be allocated to video packets in both frequency and spatial domain in a MIMO-OFDM system.

D. MIMO Mode Selection and Link Adaptation

Next, we discuss the selection of the modulation order per stream, the coding rate, and the MIMO mode to optimize the target objective. Link adaptation enables adapting the modulation and coding to the channel conditions. We optimize the modulation and coding order to maximize the throughput component of (22). Thus, we have the equation shown at the bottom of the page. Thus, the optimal modulation and coding combination can be found in the following two steps. First, for any given code rate, a corresponding set of modulation orders per stream are selected as follows:

$$\tilde{M}_i(C) = \arg \max_{M_i \in \mathcal{M}} \left\{ \frac{1-\alpha_i(\gamma_i, M_i, C)}{1-\alpha_i(\gamma_i, M_i, C)^{L+1}} R_i \right\}.$$

Next, given $\{\tilde{M}_i(C)\}_{i=1}^{S}$ for every $C \in \mathcal{C}$, we select the optimal code rate and the corresponding optimal modulation order per stream as follows:

$$C^* = \arg\max_{C \in \mathcal{C}} \left\{ C \sum_i \log_2 \tilde{M}_i(C) \frac{1-\alpha_i(\gamma_i, M_i, C)}{1-\alpha_i(\gamma_i, M_i, C)^{L+1}} \right\},$$

$M_i^* = \tilde{M}_i(C^*)$ and $R_i^* = B \log_2(M_i^*)$.

Substituting $\{\mathbf{M}^*, C^*\}$ into (22), we obtain

$$WT_{\text{prioritized}}(\hat{\mathbf{v}}^*, \mathbf{M}^*, C^*, S) = \frac{C^*}{\mathbb{E}[b(\mathbf{s}_p)]} \sum_{i=1}^{S} \frac{1-\alpha_i}{1-\alpha_i^{L+1}} R_i^* \left[\sum_{i=1}^{S} (1-\alpha_i^{L+1}) \int_{\hat{v}_i}^{\hat{v}_{i+1}} v f_v(v) \mathrm{d}v\right]. \quad (24)$$

Practical MIMO link adaptation should include a mechanism for switching the mode, i.e., the number of spatial streams based on channel state matrix $\mathbf{H}$ to optimize system performance and provide a suitable diversity-multiplexing tradeoff. In this work, the MIMO mode selection criterion is intended to capture video quality as well as throughput. On one hand, if the loss visibility distribution experiences higher variability, it may be preferable to use more streams to provide prioritized delivery by adding more packets classes if the channel quality is good. On the other hand, if the variability in packet importance is low, then the contribution of packet prioritization is minimal and reliable delivery with a smaller number of spatial streams may be preferred. Thus, mode selection can adapt according to the video source in a content-aware manner. Consequently, the mode selection criterion is to maximize the weighted throughput expression:

$$S^* = \arg\max WT_{\text{prioritized}} \text{s.t.} C^* \sum_{i=1}^{S} \frac{1-\alpha_i}{1-\alpha_i^{L+1}} R_i^* > R \quad (25)$$

where R is the video source rate. The constraint $C^* \sum_{i=1}^{S} (1-\alpha_i) R_i^* / (1-\alpha_i^{L+1}) > R$ ensures the throughput with the selected mode at least matches the rate of the video to ensure that the wireless link can serve the requirements of the video source.

Algorithm 1 describes the complete algorithm for loss visibility-optimized video transmission including selecting the optimal thresholding policy, the MCS per stream given the post-processing SNRs corresponding to the MIMO channel decomposition, as well as MIMO mode adaptation.

$$\{\mathbf{M}^*, C^*\} = \arg\max_{M_i \in \mathcal{M}, C \in \mathcal{C}} \left\{ C \sum_i \frac{1-\alpha_i(\gamma_i, M_i, C)}{1-\alpha_i(\gamma_i, M_i, C)^{L+1}} R_i \right\} = \arg\max_{C \in \mathcal{C}} \left\{ C \sum_i \arg\max_{M_i \in \mathcal{M}, C \in \mathcal{C}} \left\{ \frac{1-\alpha_i(\gamma_i, M_i, C)}{1-\alpha_i(\gamma_i, M_i, C)^{L+1}} R_i \right\} \right\}$$

Algorithm 1 Loss Visibility Optimized Video Transmission over MIMO.

Given channel state $\mathbf{H}$
for $i = 1 \rightarrow S$ **do**
Step 1. Precoder Computation
Compute precoder $\mathbf{F}_S$ and post-processing SNRs $\boldsymbol{\gamma}(\mathbf{H}) = \{\gamma_i(\mathbf{H})\}_{i=1}^{S}$
Step 2. MCS Selection
for $C \in \mathcal{C}$ **do**
$\tilde{M}_i(C) = \arg\max_{M_i \in \mathcal{M}} \{\frac{1-\alpha_i}{1-\alpha_i^{L+1}} R_i\}$
end for
$C^* = \arg\max_{C \in \mathcal{C}} \{C \sum_i \log_2 \tilde{M}_i(C) \frac{1-\alpha_i}{1-\alpha_i^{L+1}}\}$
$M_i^* = \tilde{M}_i(C^*)$
Order streams according to post-retransmission success probability, i.e., $p_i^{\text{success}} \leq p_{i+1}^{\text{success}} \forall i = 1, \cdots, S-1$.
Step 3. Loss Visibility Distribution Update
Use kernel density estimation to update the loss visibility distribution $\hat{f}_v(x) = \frac{1}{Wh} \sum_{i=1}^{W} K(\frac{x - v(\mathbf{s}_{p-i})}{h})$
Step 4. Thresholding Policy Selection
Compute $\hat{\mathbf{v}}^* = \{\hat{v}_i^*\}_{i=2}^{S}$ to satisfy $F_v(\hat{v}_{i+1}^*) - F_v(\hat{v}_i^*) = \frac{R_i/r_i}{\sum_{j=1}^{S} R_j/r_j} \forall i = 1, \cdots, S$
end for
Step 5. Mode Selection

Select the optimal mode S^* according to (25).

V. Video Quality and Throughput Gains

To quantify the gains from using the loss visibility side information as proposed in Algorithm 1, we compare with conventional MIMO transmission whereby no side information is used for packet prioritization. Instead, the symbols corresponding to each packet are multiplexed on all spatial streams.

A. Gain Analysis

In the absence of packet prioritization, each packet is multiplexed over all streams, thus, the packet error rate expression should be modified to capture the new packet error rate. As shown in (14), the corresponding PER relates to the packet prioritization case as follows $\alpha_{\text{baseline}} = 1 - \prod_{i=1}^{S} (1 - \alpha_i)^{1/S}$. Further, the probability of success for the baseline case is expressed as $p_{\text{baseline}}^{\text{success}}(\mathbf{H}, \mathbf{M}, C) = (1 - \alpha_{\text{baseline}}^{L+1})$. Thus, for a representative set of P packets, the cumulative value of packets received successfully is $P(1 - \alpha_{\text{baseline}}^{L+1})\mathbb{E}[v(\mathbf{s}_p)]$ where $\mathbb{E}[v(\mathbf{s}_p)] = \int_0^1 v f_v(v) \mathrm{d}v$ is the average packet loss visibility. Furthermore, the transmission time is $\max_i \{\mathbb{E}[b(\mathbf{s}_p)](1 - \alpha_{\text{baseline}}^{L+1})/CR_i(1 - \alpha_{\text{baseline}})\} P/S = \mathbb{E}[b(\mathbf{s}_p)](1 - \alpha_{\text{baseline}}^{L+1})P/S(1 - \alpha_{\text{baseline}})C \min_i\{R_i\}$. Thus, the weighted throughput objective for the baseline follows from (26) as shown at the bottom of the page. We make the following two key observations regarding the result in (26):

1) In the absence of packet prioritization, unequal modulation is not beneficial. This is because the throughput is limited by the worst spatial stream as evidenced by the term $\min_i\{R_i\}$.
2) In the absence of packet prioritization, the objective does not depend on the retransmission limit r. This is due to the fact that the loss in throughput due to retransmission is compensated by a gain in video quality and vice versa.

Therefore, for the baseline case, we consider the same modulation order M for all streams. Further, we select the modulation order M and coding rate C to maximize the post retransmission throughput, that is

$$\{M^*, C^*\} = \arg \max_{M \in \mathcal{M}, C \in \mathcal{C}} \left\{ CR \frac{1 - \alpha_{\text{baseline}}(\boldsymbol{\gamma}, M, C)}{1 - \alpha_{\text{baseline}}(\boldsymbol{\gamma}, M, C)^{L+1}} \right\}$$

where $R = B \log_2(M)$. Now, we write the gain $G = \mathbb{E}_{\mathbf{H}}[WT_{\text{prioritized}}]/\mathbb{E}_{\mathbf{H}}[WT_{\text{baseline}}]$ as follows:

$$G = G_{\text{PP}} \times G_{\text{UM}} = \underbrace{\frac{\mathbb{E}_{\mathbf{H}}[\sum_{i=1}^{S}(1 - \alpha_i^{L+1}) \int_{\hat{v}_i}^{\hat{v}_{i+1}} v f_v(v) \mathrm{d}v]}{\mathbb{E}_{\mathbf{H}}[(1 - \alpha_{\text{baseline}}^{L+1}) \int_0^1 v f_v(v) \mathrm{d}v]}}_{\text{Packet Prioritization Gain } G_{\text{PP}}} \times \underbrace{\frac{\mathbb{E}_{\mathbf{H}}[\max_C \left\{ C \sum_i \max_{M_i} \{R_i/r_i\} \right\}]}{\mathbb{E}_{\mathbf{H}}[S \max_{M,C} \{CR/r_{\text{baseline}}\}]}}_{\text{Unequal Modulation Gain } G_{\text{UM}}} \tag{27}$$

where $r_i = (1 - \alpha_i^{L+1})/(1 - \alpha_i)$ and $r_{\text{baseline}} = (1 - \alpha_{\text{baseline}}^{L+1})/(1 - \alpha_{\text{baseline}})$ are the average number of retransmissions for the proposed and baseline scenarios respectively.

$$\begin{aligned} WT_{\text{baseline}} &= \frac{P(1 - \alpha_{\text{baseline}}^{L+1})\mathbb{E}[v(\mathbf{s}_p)]}{\mathbb{E}[b(\mathbf{s}_p)](1 - \alpha_{\text{baseline}}^{L+1})P/S(1 - \alpha_{\text{baseline}})C \min_i\{R_i\}} \\ &= \underbrace{\frac{C}{\mathbb{E}[b(\mathbf{s}_p)]} \frac{(1 - \alpha_{\text{baseline}})}{(1 - \alpha_{\text{baseline}}^{L+1})} S \min_i\{R_i\}}_{\text{Throughput component}} \underbrace{(1 - \alpha_{\text{baseline}}^{L+1})\mathbb{E}[v(\mathbf{s}_p)]}_{\text{Quality component}} \\ &= \frac{\mathbb{E}[v(\mathbf{s}_p)]CS(1 - \alpha_{\text{baseline}}) \min_i\{R_i\}}{\mathbb{E}[b(\mathbf{s}_p)]}. \end{aligned} \tag{26}$$

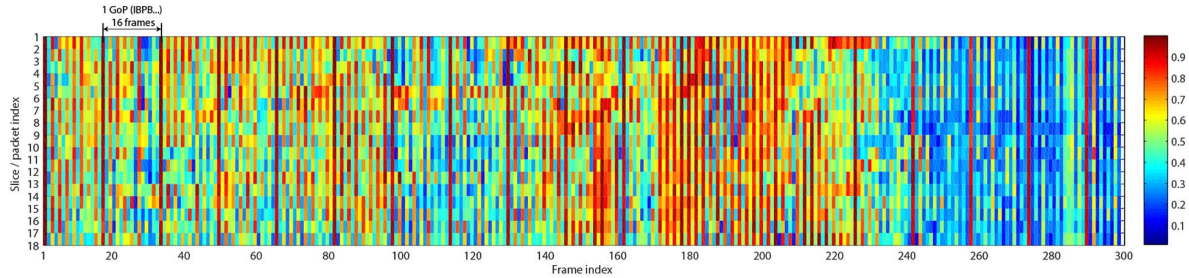

Fig. 4. Loss visibility map of the Foreman video sequence encoded with H.264/AVC using an IBPBP GoP structure with 18 horizontal slices per frame and a GoP duration of 16.

B. Packet Prioritization Gain

The first component of (27) is referred to as *packet prioritization gain* and is expressed as follows:

$$G_{\mathrm{PP}} = \frac{\mathbb{E}_{\mathbf{H}}\left[\sum_{i=1}^{S}(1-\alpha_i^{L+1})\mathbb{E}\left[v(\mathbf{s}_p)|\hat{v}_i \le v(\mathbf{s}_p) \le \hat{v}_{i+1}\right]\right]}{\left(1-\left(1-\prod_{i=1}^{S}\left(1-\mathbb{E}_{\mathbf{H}}\left[\alpha_i\right]\right)^{1/S}\right)^{L+1}\right)\mathbb{E}[v(\mathbf{s}_p)]}. \tag{28}$$

It results from the fact that the more relevant packets are transmitted through the more reliable streams. Because streams are ordered by the post-retransmission success probability $1-\alpha_i^{L+1}$, the packet prioritization gain is always greater than 1. We note that this gain is highest when both the packet loss visibility and the per-stream SNRs exhibit high variability. Furthermore, if infinite retransmissions are allowed, this gain converges to one since all packets are eventually received successfully. The dependence on $\mathbf{H}$ in (28) is through both the loss visibility thresholds $\{\hat{v}_i\}_{i=1}^{S}$ and the PERs $\{\alpha_i\}_{i=1}^{S}$. *The packet prioritization gain represents a reduction in loss visibility, i.e., a video quality gain.*

C. Unequal Modulation Gain

The second component of (27) is referred to as the *unequal modulation gain* and is expressed as follows:

$$G_{\mathrm{UM}} = \frac{\mathbb{E}_{\mathbf{H}}[\max_{C}\left\{C\sum_i \max_{M_i}\{R_i/r_i\}\right\}]}{\mathbb{E}_{\mathbf{H}}[S\max_{M,C}\{CR/r_{\mathrm{baseline}}\}]}. \tag{29}$$

It corresponds to the throughput averaged over spatial streams divided by the throughput on the worst spatial stream. It results from the fact that the optimal transmission policy can opportunistically increase the rate on the stronger streams to enable low latency delivery of high priority packets. Conversely, in conventional MIMO transmission with a fixed modulation order, the performance achieved is limited by the performance on the worst stream. This justifies why the unequal modulation gain is the achievable throughput averaged over all streams divided by the achievable throughput on the worst stream. The dependence on $\mathbf{H}$ in (29) is through the PERs $\{\alpha_i\}_{i=1}^{S}$ which impact the per-stream throughputs $\{R_i\}_{i=1}^{S}$ and the retransmission overhead $\{r_i\}_{i=1}^{S}$. *The unequal modulation gain results in an increase in throughput.*

D. Impact of Limited Feedback

The expressions for $WT_{\mathrm{prioritized}}$ and WT_{baseline} are in terms of the error probability $p_i^{\mathrm{success}_s}$, which in turn depends on the post processing SNR vector $\boldsymbol{\gamma} = \{\gamma_i(\mathbf{H})\}_{i=1}^{S}$. Thus, they apply equivalently under limited feedback given that $\gamma_i(\mathbf{H})$ is computed using (4) according to the selected precoder. We compute the gains under limited feedback by taking the expectation of the individual gains for each channel state given its mapping to the corresponding codeword. This corresponds to $G = \mathbb{E}_{\mathcal{F}}[WT_{\mathrm{prioritized}}]/\mathbb{E}_{\mathcal{F}}[WT_{\mathrm{baseline}}]$.

VI. Results and Analysis

In this section, we first evaluate the proposed loss visibility based MIMO transmission policies using H.264 encoded bit streams under different antenna configurations. Next, we present numerical results to quantify the packet prioritization and unequal modulation gains. Each entry of the channel matrix corresponds to a flat Rayleigh fading channel. The system bandwidth is 1 MHz. The set of possible M-QAM constellations is $\mathcal{M} = \{2, 4, 16, 64\}$ corresponding to BPSK, 4-QAM, 16-QAM, and 64-QAM. The set of possible coding rates is $\mathcal{C} = \{1/2, 2/3, 3/4, 5/6\}$.

A. Video Quality Gains on H.264 Sequences

To evaluate the video quality gain from the loss visibility based prioritization policy, we test the proposed algorithm on the Foreman video sequence [31] encoded with H.264/AVC. The GoP structure is $IBPBP$ … and the GoP duration is 16 frames. The MB size is 16×16 and we use the CIF resolution of 352×288. The video frame is divided into horizontal slices where each slice is 22 MBs wide and 1 MB high. Thus, each frame corresponds to 18 slices. The decoder uses motion copy error concealment. Loss visibility estimation is applied based on [3] as described in § III-A. Fig. 4 shows the resulting loss visibility scores for each frame/slice for the Foreman video sequence. Several observations are in order.

1) *Frame type:* The variability of the visibility across frames is clear. For instance, the I frames can be noticed as dark red every GoP interval. Furthermore, the odd-numbered frames corresponding to P have higher loss visibility than the even-numbered B frames.
2) *Subject/background motion:* Face motion between Frame 1 and Frame 170 cause high loss visibility for some slices depending on the spatial location of motion. Background motion between Frame 170 and Frame 220 contributes an overall increase in loss visibility. Beyond that, the lack of object and background motion causes an overall drop in loss visibility.

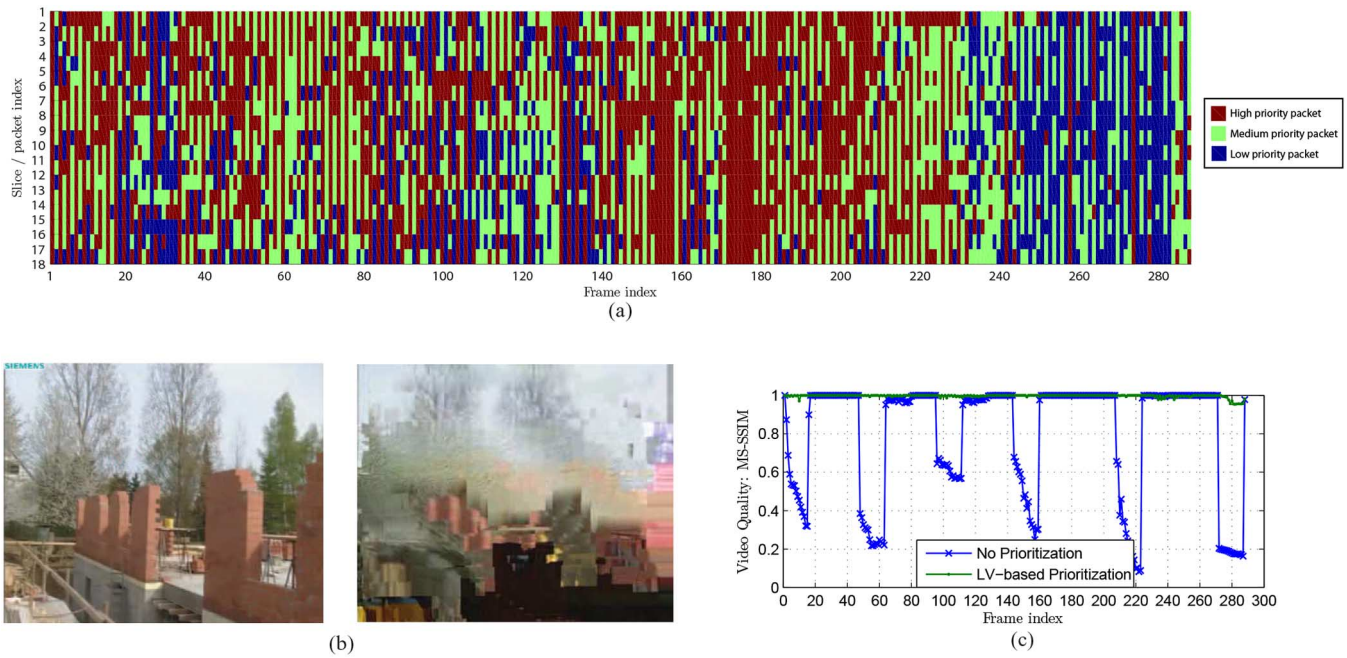

Fig. 5. Case study of the loss visibility-based prioritization policy for the Foreman video sequence with 4×4 MIMO system, $S = 3$ streams, and $E_s/N_0 = 5$ dB. The retransmission limit is $r = 4$. (a) Packet-stream mapping. (b) Received frame 223 with and without prioritization. (c) Comparison of video quality of the received videos.

3) *Error propagation:* For odd-numbered P frames, it can be noticed that the packet loss visibility captures the severity of potential error propagation by decaying for P frames towards the end of the GoP, i.e., close to the next reference frame.

Fig. 5 applies the loss visibility based prioritization policy to the Foreman video sequence [31] for a 4×4 MIMO system, $S = 3$ streams/classes, and $E_s/N_0 = 5$ dB. The retransmission limit is $r = 4$ and the channel coherence time is equal to 1 GoP corresponding to a low mobility environment. Fig. 5(a) shows the mapping of each video packet to the corresponding spatial stream. Packets mapped to the best spatial stream are referred to as high priority packets and vice versa. The corresponding video quality is shown in Fig. 5(c) in comparison with the baseline, whereby the symbols corresponding to each packet are mapped to all spatial streams, for the same channel realization. *Despite having 460 packet losses post-retransmission, the mean video quality with prioritization is 0.997 on the MS-SSIM scale whereas the mean video quality without prioritization is 0.802.* With packet prioritization, losses affect only packets where error concealment can conceal the loss from being visible to the average viewer. In contrast, the error propagation effect is very severe in the baseline case. The received and concealed frames with index 223 of the Foreman sequence are shown in Fig. 5(b) to further demonstrate the quality difference.

Fig. 6 demonstrates the video quality gains for a range of antenna configurations for the Foreman video sequence encoded with the same properties as previously described. The video quality at each data point is the frame-averaged quality further averaged over 10 different channel realizations. The same channel realizations are used for the two cases. The first observed trend is that for a fixed antenna configuration, the gains are maximized when $S = \min(N_t, N_r)$. This is because

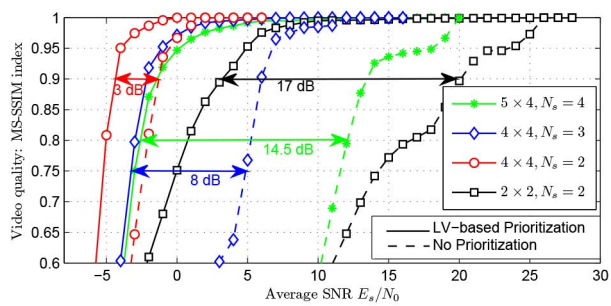

Fig. 6. Comparison of the loss visibility-based packet prioritization versus non-prioritized MIMO precoding for H.264-encoded Foreman sequence for different antenna configurations over a range of SNRs. The retransmission limit is $r = 4$ and the channel coherence time is 1 GoP.

the large variability in the post-processing SNRs across streams enables more effective packet prioritization. Furthermore, increasing the number of antennas for a fixed number of streams improves video quality but reduces the video quality gain. The maximum gain is reported for a 2×2 setting where a video quality of 0.9 requires $E_s/N_0 = 3$ dB with prioritization versus $E_s/N_0 = 20$ dB without prioritization. Furthermore, gains in excess of 10 dB are achieved over a range of antenna configurations.

Next, we consider our proposed loss visibility based packet prioritization in comparison to slice type-based traffic prioritization introduced in [32]. The method in [32] exploits the information found in SVC or AVC bitstream to order packets in decreasing order of importance, considering both the layer to which data belong and the dependencies in the decoding process. Specifically, three groups of slices are defined: I-type, P-type and B-type. Within each category, the slices are sorted based on picture decoding order resulting in an ordering of

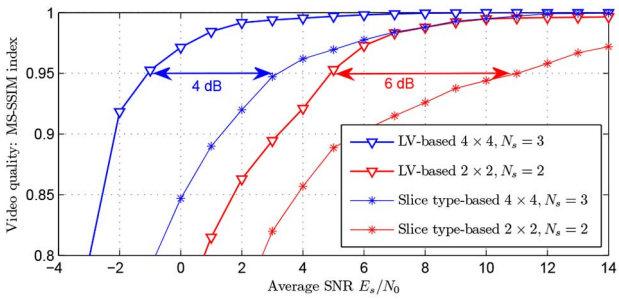

Fig. 7. Comparison of the loss visibility-based packet prioritization versus frame-based packet prioritization for different antenna configurations over a range of SNRs.

priorities all slices in one GoP. This ordering captures the inter-frame dependency that causes error propagation and video quality degradation in the case of loss of high priority packets. In our comparison, we use these priorities to map the video packets to the different MIMO spatial streams and compare to our proposed loss visibility based mapping. Fig. 7 shows the video quality achieved with loss visibility based prioritization in comparison to slice type based prioritization for two MIMO configurations, namely 4×4 with 3 streams and 2×2 with 2 streams. Results show that slice type based prioritization as proposed in the literature alone cannot fully capture the packet relevance. This is because our model implicitly captures slice type, but also several other highly relevant properties relating to motion magnitude and energy, error concealment, number of dependent frames, slice distortion visibility, etc. Furthermore, our model optimizes the weights of these features to provide the best representation of packet relevance to video quality, which prior work is unable to capture. This explains the gains observed, which are most significant in the relevant operating region, that is, for a target MS-SSIM greater than 0.9 where an order of 4 to 6 dB gain is obtained.

B. Throughput Gains

Having shown that significant video quality gains are achieved by the loss visibility-based video transmission policies, we then examine the throughput gains by plotting the closed-form unequal modulation gain expression derived in § V. Recall that the throughput gain is achieved due to leveraging unequal modulation on the stronger spatial streams.

In Fig. 8, we examine the unequal modulation gain G_{UM}, defined in (27). Fig. 8(a) shows the gain for $S = 2$ spatial streams with different antenna configurations. Recall from the unequal modulation gain expression that the gain is maximized when the per-stream throughputs exhibit the highest variability among streams. In a two stream setup, this corresponds to the case where the difference between the throughputs on the two streams is maximal. Thus, for $S = 2$, a 2×2 system gains more than a 4×4 system. In a 4×4 system with $S = 2$, the diversity and channel hardening reduce the gains from the proposed prioritization policy because the supported modulation orders per stream are equivalent for most channel realizations and the achievable throughput on the two streams is comparable. In Fig. 8(b), we plot the unequal modulation gain for a 4×4 system for different numbers of spatial streams S. In the medium to high SNR regime, for the same $N_{\mathrm{t}} \times N_{\mathrm{r}}$ configuration, more streams provide higher gains versus non-video aware approaches since the condition number of the effective channel $\mathbf{HF}_S$ is likely to be higher making it possible for video-aware techniques to make use of the diverse channel statistics among streams. For $S = 2$ and $S = 4$, we show the fractional use of each modulation scheme at the peak operating points. For $S = 2$ at $E_{\mathrm{s}}/N_0 = -1$ dB, the best stream can support 4-QAM for most realizations while the worst stream can only support BPSK. Similarly, at 8 dB and 15 dB for 16-QAM and 64-QAM. Conversely, at 4 dB (resp. 12 dB), both streams support 4-QAM (resp. 16-QAM) for most channel realizations. Thus, the gain G_{UM} is close to 1.

C. Practical Considerations: Quantized LV Information and Quantized Channel Feedback

The two key pieces of side information available to the physical layer to optimize video transmission for high perceptual quality are (1) MIMO channel feedback from the client, and (2) loss visibility communicated in the video packet headers. In practice, neither is available in a non-quantized form. Thus, we assess the performance impact of using quantized loss visibility as well as quantized channel feedback.

First, considering loss visibility quantization, we show that the same video quality gain can be achieved even if the loss visibility is quantized down to $n_b = \log_2(S)$ bits where S is the maximum number of spatial streams. Quantization of loss visibility is performed based on the loss visibility distribution such that each quantization level consists of a similar number of video packets. Fig. 9 shows the performance of quantized vs. unquantized loss visibility. In the case of 2 spatial streams, there is no performance penalty incurred by sending only a single bit of loss visibility information. Similarly, in the case of 4 spatial streams, 2 bit of loss visibility information is sufficient. This is very promising performance-wise because for all practical purposes, 2 bits of side information are enough to approach optimal physical layer video quality optimizations.

Next, we consider the impact of limited channel feedback. Fig. 10 shows the unequal modulation gain with limited feedback for different codebook sizes and antenna configurations. The codebooks are obtained using Grassmannian subspace packing with the chordal subspace distance measure.[1] As expected, the gains increase as the codebook size increases as well as for larger number of spatial streams. With only 2 spatial streams in a 4×2 antenna configuration and a 3 bit codebook, 27% throughput increase is achieved. With 4 spatial streams in a 7×4 antenna configuration and a 4 bit codebook, 56% throughput increase is achieved. The trends of the gains closely follows those in Fig. 8 corresponding to perfect CSI feedback. In terms of the nominal gain values, we observe that with codebook-based limited feedback, the gain drops because the unequal stream quality cannot be fully utilized due to channel quantization errors. Such errors cause the gap between the post processing SNRs on the best and worst stream to tighten, thus reducing the achievable gain.

[1]"Grassmannian subspace packing codebooks," [Online]. Available: https://engineering.purdue.edu/~djlove/grass.html

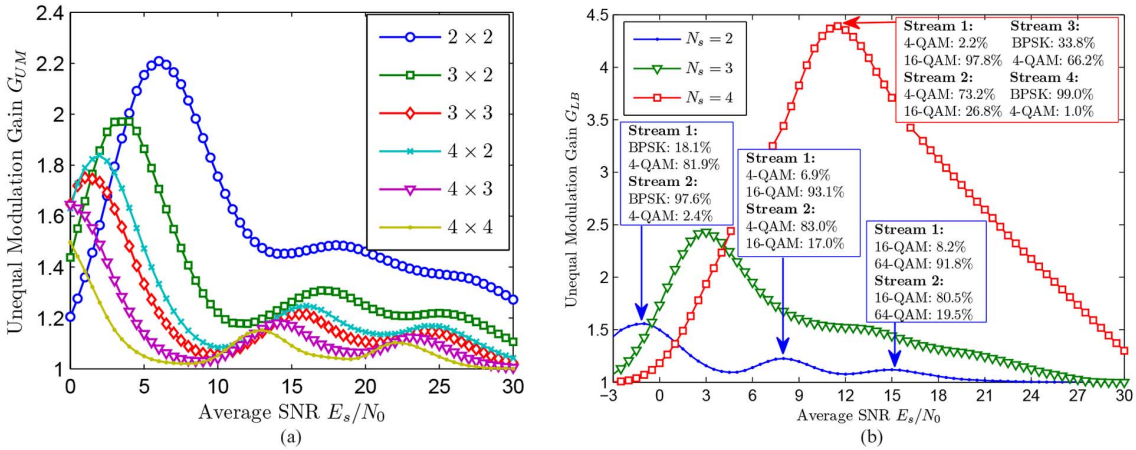

Fig. 8. Analysis of the unequal modulation gain G_{UM}. It corresponds to the throughput averaged over spatial streams divided by the throughput on the worst spatial stream. The peaks correspond to operating points where the modulation orders across streams are very likely to be "unequal". (a) Different antenna configurations with $S = 2$ streams (b) Different number of streams for a 4×4 system.

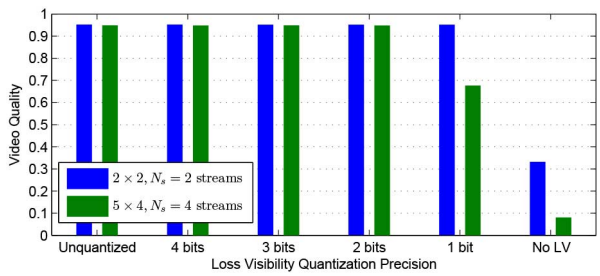

Fig. 9. Quantized versus unquantized loss visibility performance. Note that similar performance is achieved as long as the number of bits is $n_b \geq \log_2(S)$ where S is the maximum number of spatial streams.

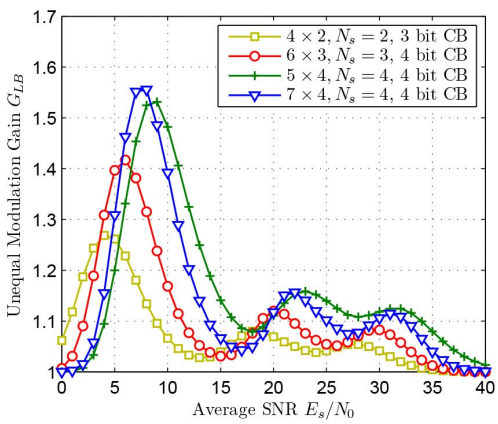

Fig. 10. Unequal modulation gain achieved with limited feedback for different codebook sizes and antenna configurations.

VII. Conclusion

We proposed a cross-layer architecture for prioritized packet delivery over a MIMO PHY layer based on loss visibility taking advantage of the large variability in loss visibility due to the video source and encoder features. We presented a loss visibility-based thresholding policy that maps different packets to different spatial streams and derived the optimal thresholding policy for any loss visibility distribution. The proposed architecture requires minimal additional cross-layer overhead while achieving quality and capacity gains. We demonstrated gains in excess of 10 dB with different antenna configurations on H.264 encoded video sequences.

Appendix A
Proof of Lemma 1

Consider two video packets $\mathbf{s}_1$ and $\mathbf{s}_2$ such that $v(\mathbf{s}_1) < v(\mathbf{s}_2)$. Assume the packet-stream mapping is such that $\mathbf{s}_2 \in \mathcal{V}'_i$ and $\mathbf{s}_1 \in \mathcal{V}'_k$ where $p_i^{\text{success}} = 1 - \alpha_i^{L+1} \leq 1 - \alpha_k^{L+1} = p_k^{\text{success}}$. We switch the mapping of packets $\mathbf{s}_1$ and $\mathbf{s}_2$, that is, $\mathcal{V}_i = \mathcal{V}'_i + \{\mathbf{s}_2\} - \{\mathbf{s}_1\}$ and $\mathcal{V}_k = \mathcal{V}'_k + \{\mathbf{s}_1\} - \{\mathbf{s}_2\}$. We show that the corresponding objective function $WT' \leq WT$. The expression for $WT' = V'/T'$ is as shown in (30), at the bottom of the page. We can simplify the numerator V' as follows:

$$V' = \left(\sum_{l \notin \{i,k\}} p_l^{\text{success}} \sum_{\mathbf{s} \in \mathcal{V}_l} v(\mathbf{s}) \right) + p_i^{\text{success}} \sum_{\mathbf{s} \in \mathcal{V}'_i} v(\mathbf{s}) + p_k^{\text{success}} \sum_{\mathbf{s} \in \mathcal{V}'_k} v(\mathbf{s}) \tag{31}$$

$$= \left(\sum_{l \notin \{i,k\}} p_l^{\text{success}} \sum_{\mathbf{s} \in \mathcal{V}_l} v(\mathbf{s}) \right) + p_i^{\text{success}} \sum_{\mathbf{s} \in \mathcal{V}_i} v(\mathbf{s}) + p_k^{\text{success}} \sum_{\mathbf{s} \in \mathcal{V}_k} v(\mathbf{s}) + (p_i^{\text{success}} - p_k^{\text{success}})(v(\mathbf{s}_2) - v(\mathbf{s}_1)) \tag{32}$$

$$< \sum_l p_l^{\text{success}} \sum_{\mathbf{s} \in \mathcal{V}_l} v(\mathbf{s}) \tag{33}$$

$$WT' = \frac{(\sum_{l \notin \{i,k\}} p_l^{\text{success}} \sum_{\mathbf{s} \in \mathcal{V}_l} v(\mathbf{s})) + p_i^{\text{success}} \sum_{\mathbf{s} \in \mathcal{V}'_i} v(\mathbf{s}) + p_k^{\text{success}} \sum_{\mathbf{s} \in \mathcal{V}'_k} v(\mathbf{s})}{\max\{\max_{l \notin \{i,k\}} t_l(\gamma_l, M_l, C, \mathcal{V}_l), t_i(\gamma_i, M_i, C, \mathcal{V}'_i), t_k(\gamma_k, M_k, C, \mathcal{V}'_k)\}} = \frac{V'}{T'} \tag{30}$$

where (33) follows because $v(\mathbf{s}_1) < v(\mathbf{s}_2)$ and $p_i^{\text{success}} \leq p_k^{\text{success}}$ by definition and $p_i^{\text{success}} \leq p_k^{\text{success}}$ by the proposed ordering. Next, the denominator T' can be reduced as follows:

$$\begin{aligned} T' &= \max\{\max_{l\notin\{i,k\}} t_l(\gamma_l, M_l, C, \mathcal{V}_l), t_i(\gamma_i, M_i, C, \mathcal{V}_i'), \\ &\quad t_k(\gamma_k, M_k, C, \mathcal{V}_k')\} \quad (34) \\ &\geq \max\{\max_{l\notin\{i,k\}} t_l(\gamma_l, M_l, C, \mathcal{V}_l), t_i(\gamma_i, M_i, C, \mathcal{V}_i), \\ &\quad t_k(\gamma_k, M_k, C, \mathcal{V}_k)\} \quad (35) \\ &= \max_l t_l. \quad (36) \end{aligned}$$

We show the key transition to (35) separately in all the following four possible cases. For brevity, we denote by $\text{Thr}(i)$ the throughput on the ith stream in the derivation below.

1) $\{\arg\max(t_l, t_i(\mathcal{V}_i'), t_k(\mathcal{V}_k'))) = i,\ \arg\max(t_l, t_i(\mathcal{V}_i), t_k(\mathcal{V}_k)) = i\}$: In this case, switching the ordering improves the objective since $v(\mathbf{s}_1) < v(\mathbf{s}_2)$ and $b(\mathbf{s}_1) < b(\mathbf{s}_2)$, thus $t_i(\mathcal{V}_i) < t_i(\mathcal{V}_i')$.
2) $\{\arg\max(t_l, t_i(\mathcal{V}_i'), t_k(\mathcal{V}_k'))) = k,\ \arg\max(t_l, t_i(\mathcal{V}_i), t_k(\mathcal{V}_k)) = k\}$: While this reduces the objective since $t_k(\mathcal{V}_k) > t_k(\mathcal{V}_k)$, we show by contradiction that it never occurs. We have $\arg\max(t_l, t_i(\mathcal{V}_i'), t_k(\mathcal{V}_k'))) = k \Rightarrow b(\mathbf{s}_2)/\text{Thr}(i) < b(\mathbf{s}_1)/\text{Thr}(k)$ and $\arg\max(t_l, t_i(\mathcal{V}_i), t_k(\mathcal{V}_k))) = k \Rightarrow b(\mathbf{s}_2)/\text{Thr}(k) > b(\mathbf{s}_1)/\text{Thr}(i)$. Thus, $\text{Thr}(i)/\text{Thr}(k) < b(\mathbf{s}_2)/b(\mathbf{s}_1) < \text{Thr}(k)/\text{Thr}(i)$. Since $b(\mathbf{s}_2)/b(\mathbf{s}_1) > 1$, we have $1 < \text{Thr}(k)/\text{Thr}(i) \Rightarrow \text{Thr}(i) < \text{Thr}(k) \Rightarrow b(\mathbf{s}_1)/\text{Thr}(i) > b(\mathbf{s}_1)/\text{Thr}(k)$. Combining with $b(\mathbf{s}_2)/\text{Thr}(i) < b(\mathbf{s}_1)/\text{Thr}(k)$, we obtain $b(\mathbf{s}_2)/\text{Thr}(i) < b(\mathbf{s}_1)/\text{Thr}(i) \Rightarrow b(\mathbf{s}_2) < b(\mathbf{s}_1)$. Thus, we have a contradiction.
3) $\{\arg\max(t_l, t_i(\mathcal{V}_i'), t_k(\mathcal{V}_k'))) = i,\ \arg\max(t_l, t_i(\mathcal{V}_i), t_k(\mathcal{V}_k)) = k\}$: We show by contradiction that this case never occurs. $\arg\max(t_l, t_i(\mathcal{V}_i'), t_k(\mathcal{V}_k'))) = i \Rightarrow b(\mathbf{s}_2)/\text{Thr}(i) > b(\mathbf{s}_1)/\text{Thr}(k)$ and $\arg\max(t_l, t_i(\mathcal{V}_i), t_k(\mathcal{V}_k)) = k \Rightarrow b(\mathbf{s}_2)/\text{Thr}(k) > b(\mathbf{s}_1)/\text{Thr}(i)$. Thus, $b(\mathbf{s}_2)/b(\mathbf{s}_1) > \text{Thr}(k)/\text{Thr}(i)$ and $b(\mathbf{s}_2)/b(\mathbf{s}_1) > \text{Thr}(i)/\text{Thr}(k)$ which is a contradiction.
4) $\{\arg\max(t_l, t_i(\mathcal{V}_i'), t_k(\mathcal{V}_k'))) = k,\ \arg\max(t_l, t_i(\mathcal{V}_i), t_k(\mathcal{V}_k)) = i\}$: We show by contradiction that this case never occurs. $\arg\max(t_l, t_i(\mathcal{V}_i'), t_k(\mathcal{V}_k'))) = k \Rightarrow b(\mathbf{s}_2)/\text{Thr}(i) < b(\mathbf{s}_1)/\text{Thr}(k)$ and $\arg\max(t_l, t_i(\mathcal{V}_i), t_k(\mathcal{V}_k)) = i \Rightarrow b(\mathbf{s}_2)/\text{Thr}(k) < b(\mathbf{s}_1)/\text{Thr}(i)$. Thus, $b(\mathbf{s}_2)/b(\mathbf{s}_1) < \text{Thr}(k)/\text{Thr}(i)$ and $b(\mathbf{s}_2)/b(\mathbf{s}_1) < \text{Thr}(i)/\text{Thr}(k)$ which is a contradiction.

Thus, it follows that $T' \geq \max_l t_l$. Consequently, we have

$$WT' < \frac{\sum_l p_l^{\text{success}} \sum_{\mathbf{s}\in\mathcal{V}_l} v(\mathbf{s})}{\max_l t_l} = WT. \quad (37)$$

Thus, the proposed ordering maximizes the objective function and $\hat{v}_i$ is the threshold between ordered spatial stream $i-1$ and i.

Appendix B
Proof of Lemma 2

First, the gradient of $WT(\hat{\mathbf{v}}, \mathbf{M}, C, S)$ with respect to $\hat{v}_i$ is $\partial WT/\partial\hat{v}_i = (h\partial g/\partial\hat{v}_i - g\partial h/\partial\hat{v}_i)/h^2$ where $g = [\sum_{i=1}^{S}(1-\alpha_i^{L+1})\int_{\hat{v}_i}^{\hat{v}_{i+1}} v f_v(v)\text{d}v]$ and $h = \mathbb{E}[b(\mathbf{s}_p)](F_v(\hat{v}_{\tilde{i}+1}) - F_v(\hat{v}_{\tilde{i}}))(1-\alpha_{\tilde{i}}^{L+1})/CR_{\tilde{i}}(1-\alpha_{\tilde{i}})$ are the numerator and denominator of (19). The components of the gradient are

$$\frac{\partial g}{\partial\hat{v}_i} = (\alpha_i^{L+1} - \alpha_{i-1}^{L+1})\hat{v}_i f_v(\hat{v}_i) \quad (38)$$

where we used the fact that $\partial(\int_0^{\hat{v}_1} v f_v(v)\text{d}v)/\partial\hat{v}_1 = \lim_{\epsilon\to 0}(\int_{\hat{v}_1}^{\hat{v}_1+\epsilon} v f_v(v)\text{d}v/\epsilon) = \hat{v}_1 f_v(\hat{v}_1)$. Furthermore, the gradient corresponding to h is

$$\frac{\partial h}{\partial\hat{v}_i} = \frac{\mathbb{E}[b(\mathbf{s}_p)]f_v(\hat{v}_i)}{C}\begin{cases} \frac{1-\alpha_{i-1}^{L+1}}{R_{i-1}(1-\alpha_{i-1})} & \text{if } i = \tilde{i}+1 \\ -\frac{1-\alpha_i^{L+1}}{R_i(1-\alpha_i)} & \text{if } i = \tilde{i} \\ 0 & \text{otherwise.} \end{cases}$$

Next, we prove part 1 of the Lemma. From the expressions for $\partial g/\partial\hat{v}_i$ and $\partial h/\partial\hat{v}_i$, it follows that

$$\begin{aligned} &\frac{\partial WT}{\partial\hat{v}_{\tilde{i}}} h^2 \frac{CR_{\tilde{i}}(1-\alpha_{\tilde{i}})}{\mathbb{E}[b(\mathbf{s}_p)]} \quad (40) \\ &= (\alpha_{\tilde{i}}^{L+1} - \alpha_{\tilde{i}-1}^{L+1})\hat{v}_i f_v(\hat{v}_i)(F_v(\hat{v}_{\tilde{i}+1}) - F_v(\hat{v}_{\tilde{i}}))(1-\alpha_{\tilde{i}}^{L+1}) \\ &\quad + \left(\sum_{i=1}^{S}(1-\alpha_i^{L+1})\int_{\hat{v}_i}^{\hat{v}_{i+1}} v f_v(v)\text{d}v\right) f_v(\hat{v}_{\tilde{i}})(1-\alpha_{\tilde{i}}^{L+1}) \quad (41) \\ &= f_v(\hat{v}_{\tilde{i}})(1-\alpha_{\tilde{i}}^{L+1})\left[\left(\sum_{i=1}^{S}(1-\alpha_i^{L+1})\int_{\hat{v}_i}^{\hat{v}_{i+1}} v f_v(v)\text{d}v\right)\right. \\ &\quad \left. - (\alpha_{\tilde{i}-1}^{L+1} - \alpha_{\tilde{i}}^{L+1})\hat{v}_i(F_v(\hat{v}_{\tilde{i}+1}) - F_v(\hat{v}_{\tilde{i}}))\right] \quad (42) \\ &\geq f_v(\hat{v}_{\tilde{i}})(1-\alpha_{\tilde{i}}^{L+1})\left[\left((1-\alpha_{\tilde{i}}^{L+1})\int_{\hat{v}_{\tilde{i}}}^{\hat{v}_{\tilde{i}+1}} v f_v(v)\text{d}v\right)\right. \\ &\quad \left. -(\alpha_{\tilde{i}-1}^{L+1} - \alpha_{\tilde{i}}^{L+1})\hat{v}_i(F_v(\hat{v}_{\tilde{i}+1}) - F_v(\hat{v}_{\tilde{i}}))\right] \quad (43) \end{aligned}$$

where (43) follows because $\sum_{i=1}^{S}(1-\alpha_i^{L+1})V_i \geq (1-\alpha_{\tilde{i}}^{L+1})V_{\tilde{i}}$. Next, using the fact that $\alpha_{\tilde{i}-1}^{L+1} \leq 1$, we further reduce the expression to

$$\begin{aligned} &\frac{\partial WT}{\partial\hat{v}_{\tilde{i}}} h^2 \frac{CR_{\tilde{i}}(1-\alpha_{\tilde{i}})}{\mathbb{E}[b(\mathbf{s}_p)]} \\ &\quad \geq f_v(\hat{v}_{\tilde{i}})(1-\alpha_{\tilde{i}}^{L+1})^2 \\ &\quad\quad \times\left[\int_{\hat{v}_{\tilde{i}}}^{\hat{v}_{\tilde{i}+1}} v f_v(v)\text{d}v - \hat{v}_i\left(F_v(\hat{v}_{\tilde{i}+1}) - F_v(\hat{v}_{\tilde{i}})\right)\right]. \quad (44) \end{aligned}$$

Finally, applying the fact that $\int_a^b x f(x)dx \geq \int_a^b a f(x)dx$ if $a \geq 0$ where $a = \hat{v}_{\tilde{i}}$ and $x = v$, it follows that $\partial WT/\partial\hat{v}_{\tilde{i}} \geq 0$.

We prove part 2 of Lemma 1 by investigating the terms of the gradient $\partial WT/\partial\hat{v}_i = (h\partial g/\partial\hat{v}_i - g\partial h/\partial\hat{v}_i)/h^2$. We have $h \geq 0$ and $\partial g/\partial\hat{v}_i \leq 0$ unconditionally. Furthermore, $\partial h/\partial\hat{v}_i \geq 0 \forall i \neq \tilde{i}$ and $g \geq 0$. Thus, $\partial WT/\partial\hat{v}_i \leq 0 \forall i \neq \tilde{i}$.

Appendix C
Proof of Lemma 3

The special case of $|\mathcal{I}| = 1$ is proved in Lemma 1. The case of $|\mathcal{I}| > 1$ where the elements of $|\mathcal{I}|$ are non-consecutive also directly follows from Lemma 1 as one could jointly decrease $\{\hat{v}_i\}\forall i \in \mathcal{I}$ and increase $\{\hat{v}_{i+1}\}\forall i \in \mathcal{I}$ such that the set $\mathcal{I}$ is fixed. For the general case where some elements of $\mathcal{I}$ are

consecutive, the set $\mathcal{I}$ can be divided into subsets of consecutive streams. For example, if $\mathcal{I} = \{1, 3, 4\}$, the first subset is $\{1\}$ and the second subset is $\{3,4\}$. Within each subset, $\partial WT/\partial \hat{v}_i \geq 0$ for the lower-most stream satisfying $i \in \mathcal{I}$ and $i-1 \notin \mathcal{I}$ by part 1 of Lemma 1 and $\partial WT/\partial \hat{v}_i \leq 0$ for the upper-most stream satisfying $i \notin \mathcal{I}$ and $i-1 \in \mathcal{I}$ by part 2 of Lemma 1. Thus, there exist an infinitesimal step $\boldsymbol{\epsilon} = \{\epsilon_1, \cdots, \epsilon_S\}$ such that $\epsilon_i \geq 0$ if $i \in \mathcal{I}$ and $i - 1\mathcal{I}$ and $\epsilon_i \leq 0$ if $i \notin \mathcal{I}$ and $i - 1 \in \mathcal{I}$ keeping $\mathcal{I}$ fixed and improving the objective and the result follows.

Appendix D
Proof of Theorem 1

We present a convergent method that takes as input any feasible solution and obtains a solution with an improved objective satisfying the condition stated above. Start with any feasible solution and define the initial set of streams with the longest average transmission time $\mathcal{I} = \{i \text{s.t.} t_i = \max_j t_j\}$. Construct an infinitesimal step $\boldsymbol{\epsilon} = \{\epsilon_1, \cdots, \epsilon_S\}$ such that $\epsilon_i \geq 0$ if $i \in \mathcal{I}$ and $i - 1 \notin \mathcal{I}$ and $\epsilon_i \leq 0$ if $i \notin \mathcal{I}$ and $i - 1 \in \mathcal{I}$. By Lemma 3, there exists such an $\boldsymbol{\epsilon}$ such that $\mathcal{I}$ is unchanged and $WT(\hat{\mathbf{v}} + \boldsymbol{\epsilon}) > WT(\hat{\mathbf{v}})$. Repeat until $\min_{i \in \mathcal{I}, j \notin \mathcal{I}}\{t_i - t_j\} < \delta$ where δ is an arbitrarily small positive number. This necessarily increases $\mathcal{I}$. Repopulate $\mathcal{I}$ according to the new $\{\hat{v}_i\}$. Repeat until $\mathcal{I} = \{2, \cdots, S\}$. Thus, the optimal policy necessarily satisfies $t_1 = t_2 = \cdots = t_S$, equivalently, $(F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i))/(R_i(1-\alpha_i)) = (F_v(\hat{v}_2) - F_v(\hat{v}_1))/(R_1(1-\alpha_1)) \forall i$. By taking $1 = \sum_i F_v(\hat{v}_{i+1}) - F_v(\hat{v}_i)$, the Theorem follows.

References

[1] T. Wiegand, G. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H. 264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[2] 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 Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

[3] T. Lin, S. Kanumuri, Y. Zhi, D. Poole, P. Cosman, and A. Reibman, "A versatile model for packet loss visibility and its application to packet prioritization," *IEEE Trans. Image Process.*, vol. 19, no. 3, pp. 722–735, Mar. 2010.

[4] S. Kanumuri, P. Cosman, A. Reibman, and V. Vaishampayan, "Modeling packet-loss visibility in MPEG-2 video," *IEEE Trans. Multimedia*, vol. 8, no. 2, pp. 341–355, Apr. 2006.

[5] S. Kanumuri, S. Subramanian, P. Cosman, and A. Reibman, "Predicting H.264 packet loss visibility using a generalized linear model," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2006, pp. 2245–2248.

[6] L. Toni, P. Cosman, and L. Milstein, "Unequal error protection based on slice visibility for transmission of compressed video over OFDM channels," in *Proc. IEEE Int. Conf. Multimedia Expo*, Jul. 2011, pp. 1–6.

[7] J. Xu, R. Hormis, and X. Wang, "MIMO video broadcast via transmit-precoding and SNR-scalable video coding," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 3, pp. 456–466, Apr. 2010.

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

[9] R. Hormis, E. Linzer, and X. Wang, "Adaptive mode and diversity-control for video transmission on MIMO wireless channels," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3624–3637, Sep. 2009.

[10] O. Oyman and J. Foerster, "Distortion-aware MIMO link adaptation for enhanced multimedia communications," in *Proc. Int. Symp. Personal, Indoor Mobile Radio Commun. Workshops*, Sep. 2010, pp. 387–392.

[11] A. A. Khalek, C. Caramanis, and R. W. Heath, Jr., "Video-aware MIMO precoding with packet prioritization and unequal modulation," in *Proc. Eur. Signal Process. Conf.*, Aug. 2012, pp. 1905–1909.

[12] A. A. Khalek, C. Caramanis, and R. W. Heath, Jr., "Prioritized multimode precoding for joint minimization of source-channel video distortions," in *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, Nov. 2012, pp. 925–929.

[13] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proc. IEEE*, vol. 93, no. 1, pp. 71–83, Jan. 2005.

[14] L. Kondi, F. Ishtiaq, and A. Katsaggelos, "Joint source-channel coding for motion-compensated DCT-based SNR scalable video," *IEEE Trans. Image Process.*, vol. 11, no. 9, pp. 1043–1052, Sep. 2002.

[15] Y. Zhang, W. Gao, Y. Lu, Q. Huang, and D. Zhao, "Joint source-channel rate-distortion optimization for H.264 video coding over error-prone networks," *IEEE Trans. Multimedia*, vol. 9, no. 3, pp. 445–454, Apr. 2007.

[16] A. Abdel Khalek, C. Caramanis, and R. W. Heath, Jr., "Joint source-channel adaptation for perceptually optimized scalable video transmission," in *Proc. IEEE GLOBECOM*, Dec. 2011, pp. 1–5.

[17] J. Kim, R. Mersereau, and Y. Altunbasak, "Error-resilient image and video transmission over the Internet using unequal error protection," *IEEE Trans. Image Process.*, vol. 12, no. 2, pp. 121–131, Feb. 2003.

[18] M. Gallant and F. Kossentini, "Rate-distortion optimized layered coding with unequal error protection for robust Internet video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 357–372, Mar. 2001.

[19] A. Abdel Khalek, C. Caramanis, and R. W. Heath, Jr, "A cross-layer design for perceptual optimization of H.264/SVC with unequal error protection," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 7, pp. 1157–1171, Aug. 2012.

[20] M. van der Schaar and D. S. 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.

[21] Q. Zhang, W. Zhu, and Y.-Q. Zhang, "Channel-adaptive resource allocation for scalable video transmission over 3G wireless network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 8, pp. 1049–1063, Aug. 2004.

[22] H. Luo, S. Ci, D. Wu, J. Wu, and H. Tang, "Quality-driven cross-layer optimized video delivery over LTE," *IEEE Commun. Mag.*, vol. 48, no. 2, pp. 102–109, Feb. 2010.

[23] R. Heath, Jr., S. Sandhu, and A. Paulraj, "Antenna selection for spatial multiplexing systems with linear receivers," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 142–144, Apr. 2001.

[24] E. Telatar, "Capacity of multi-antenna gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, Nov. 1999.

[25] D. Love and R. Heath, Jr, "Multimode precoding for MIMO wireless systems," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3674–3687, Oct. 2005.

[26] D. Love and R. Heath, "Limited feedback unitary precoding for spatial multiplexing systems," *IEEE Trans. Inf. Theory*, vol. 51, no. 8, pp. 2967–2976, Aug. 2005.

[27] A. Goldsmith, *Wireless communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[28] G. J. Sullivan, J. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[29] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards—including high efficiency video coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669–1684, Dec. 2012.

[30] D. Scott, *Multivariate Density Estimation*. Hoboken, NJ, USA: Wiley, 1992, vol. 139.

[31] "YUV video sequences," Arizona State Univ.. Tempe, AZ, USA, Jan. 2010 [Online]. Available: http://trace.eas.asu.edu/yuv/

[32] A. Fiandrotti, D. Gallucci, E. Masala, and E. Magli, "Traffic prioritization of h. 264/svc video over 802.11e ad hoc wireless networks," in *Proc. 17th Int. Conf. Comput. Commun. Netw.*, 2008, pp. 1–5.

Amin Abdel Khalek (S'11–M'13) received the B.E. degree in computer and communication engineering (with highest distinction) from Notre Dame University, Notre Dame, IN, USA, in 2008, the M.E. degree in electrical and computer engineering from the American University of Beirut, Beirut, Lebanon, in 2010, and the Ph.D. degree in electrical and computer engineering from the University of Texas at Austin, Austin, TX, USA, in 2013.

He was a member of the Wireless Networking and Communications Group, The University of Texas at Austin, Austin, TX, USA, from 2010 to 2013. In 2011 and 2012, he was an Intern with the Intel Corporation, Hillsboro, OR, USA. In 2013, he was an Intern

with Samsung Telecommunications America, Richardson, TX, USA. In 2014, he became a Senior Design Engineer with Freescale Semiconductor, Austin, TX, USA. His current interests include wireless communications and multimedia signal processing, 3GPP LTE and IEEE 802.11 L1 and L2 algorithms, and emerging communication standards such as 5G.

Constantine Caramanis (S'05–M'06) received the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 2006.

Since 2006, he has been a Faculty Member with the Department of Electrical and Computer Engineering, The University of Texas at Austin, where he is currently the Associate Director of the Wireless Networking and Communications Group. His current research interests include robust and large scale optimization and control, machine learning and high-dimensional statistics with applications to large scale networks, and computer aided design.

Dr. Caramanis received the NSF CAREER award in 2011.

Robert W. Heath Jr. (S'96–M'01–SM'06–F'11) received the B.S. and M.S. degrees from the University of Virginia, Charlottesville, VA, USA, in 1996 and 1997 respectively, and the Ph.D. degree from Stanford University, Stanford, CA, USA, in 2002, all in electrical engineering.

From 1998 to 2001, he was a Senior Member of the Technical Staff then a Senior Consultant with Iospan Wireless Inc, San Jose, CA, USA. Since January 2002, he has been with the Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX, USA, where he is a Cullen Trust for Higher Education Endowed Professor, as well as a Member of the Wireless Networking and Communications Group. He is also President and CEO of MIMO Wireless Inc., Austin, TX, USA, and Chief Innovation Officer at Kuma Signals LLC, Austin, TX, USA. He co-authored the book *Millimeter Wave Wireless Communications* (Prentice Hall, 2014). His research interests include several aspects of wireless communication and signal processing: limited feedback techniques, multihop networking, multiuser and multicell MIMO, interference alignment, adaptive video transmission, manifold signal processing, and millimeter wave communication techniques.

Dr. Heath was a 2003 Frontiers in Education New Faculty Fellow. He is also a licensed amateur radio operator and is a Registered Professional Engineer in the state of Texas. He has been an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, an Associate Editor for the IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, and Lead Guest Editor for the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS special issue on limited feedback communication, as well as the IEEE JOURNAL ON SELECTED TOPICS IN SIGNAL PROCESSING special issue on heterogenous networks. He was on the Steering Committee for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS from 2011 to 2014. He was a Member of the Signal Processing for Communications Technical Committee in the IEEE Signal Processing Society, and is a former Chair of the IEEE COMSOC Communications Technical Theory Committee. He was a Technical Co-Chair for the 2007 Fall Vehicular Technology Conference; General Chair of the 2008 Communication Theory Workshop; General Co-Chair, Technical Co-Chair, and Co-Organizer of the 2009 IEEE Signal Processing for Wireless Communications Workshop; Local Co-Organizer for the 2009 IEEE CAMSAP Conference; Technical Co-Chair for the 2010 IEEE International Symposium on Information Theory; Technical Chair for the 2011 Asilomar Conference on Signals, Systems, and Computers; General Chair for the 2013 Asilomar Conference on Signals, Systems, and Computers; Founding General Co-Chair for the 2013 IEEE GlobalSIP conference; and Technical Co-Chair for the 2014 IEEE GLOBECOM conference. He was a co-author of papers that received Best Student Paper Awards at IEEE VTC 2006 Spring, WPMC 2006, IEEE GLOBECOM 2006, IEEE VTC 2007 Spring, and IEEE RWS 2009, as well as co-recipient of the Grand Prize in the 2008 WinTech WinCool Demo Contest. He was co-recipient of the 2010 and 2013 EURASIP *Journal on Wireless Communications and Networking* Best Paper Awards, the 2012 Signal Processing Magazine Best Paper Award, a 2013 Signal Processing Society Best Paper Award, the 2014 EURASIP Journal on Advances in Signal Processing Best Paper Award, and the 2014 *Journal of Communications and Networks* Best Paper Award.