DIGITAL-ANALOG SUPERPOSITION CODING FOR OFDM CHANNELS WITH APPLICATION TO VIDEO TRANSMISSION

Pradeepa Yahampath

Department of Electrical and Computer Engineering, University of Manitoba
Winnipeg, Manitoba R3T 5V6, Canada
email: Pradeepa.Yahampath@umanitoba.ca

ABSTRACT

A new approach to hybrid digital-analog (HDA) video transmission over orthogonal frequency division multiplexing (OFDM) channels is presented. The goal is to achieve the best video quality by optimal power allocation, when the OFDM sub-channels have unequal and time-varying signal-to-noise ratios (SNR). In this method, the quantization error of a video encoder is superimposed on digital quadrature amplitude modulation (QAM) symbols. A solution to the power allocation problem is presented. Experimental comparisons with layered video coding and adaptive modulation shows that the proposed HDA approach is able to achieve a better video quality most of the time, particularly whenever there is a high motion content.

***Index Terms*—** HDA coding, video coding, OFDM channels, power-allocation, layered coding

1. INTRODUCTION

Recently, analog transmission has been considered for video communication over wireless channels due to its ability to achieve a graceful variation of video quality as the channel SNR varies [1–3]. However, high bandwidth compression required for real-time video transmission is difficult to achieve in analog form, whereas digital video coding techniques which achieve very high compression ratios with only a modest loss of quality are now common place [4]. Recent theoretical and experimental studies indicate that HDA transmission schemes can be designed to jointly achieve both compression efficiency and robustness to channel variations [5–8].

In purely digital transmission, the aforementioned problem is dealt with by using layered transmission [9, 10], where video is first encoded into multi-resolution layers, and then transmitted by applying adaptive channel-coding and modulation (ACM) to the layers to match the channel conditions. ACM typically involves selecting the channel coderate and the QAM-order from a pre-defined set of combinations, referred to as *modulation and coding schemes* (MCS) [11]. In practice, one has to make do with a limited number of coded layers and MCSs. However, an infinite number of refinement layers and infinite modulation resolution can be achieved in a very simple manner if the refinement information (e.g., quantization error) is transmitted in analog form.

In this paper, we present a new approach to HDA video transmission over frequency-selective fading channels using OFDM [12]. Typically, SNRs on different OFDM sub-channels are not identical and can vary as much as 20 dB [13]. Therefore, when the channel state information (CSI) is available at the transmitter, the optimal power loading across OFDM sub-channels can be used to significantly increase the bitrate. With HDA transmission though power allocation should account for both digital analog signals for best performance. Unlike the previous work on HDA video transmission, our approach uses superposition of analog and digital signals to more efficiently use the bandwidth and power. A digital baselayer is first generated by using a standard video encoder [4]. The analog quantization error of the baselayer is then superimposed on the in-phase and quadrature components of the digital QAM signal. The quantization error is reconstructed at the receiver by optimal linear estimation. We present a complete solution to the underlying digital-analog power allocation problem. Experimental results obtained with video transmission over an OFDM system loosely based on the IEEE 802.11a [11] are presented to compare the HDA scheme with a completely digital, layered transmission scheme. These results clearly demonstrate the advantage of the HDA approach due to the more efficient use of power across the OFDM sub-channels.

Relation to prior work- Previous work on analog and HDA video transmission falls into two categories- (i) broadcasting over AWGN channels (no transmitter-side CSI)) [1, 3, 5, 7, 8, 14, 15], and (ii) adaptive-transmission over fading channels (transmitter uses CSI) [2, 6, 8]. This paper falls into the second category. Work in [2] involves pure analog transmission of motion-alinged 3D transform coefficients of the video signal. As a result the power allocation problem is much different to ours. Furthermore, 3D transforms cannot achieve the level of bandwidth compression that can be reached with the state-of-the art video codecs. A direct comparison with ours is not possible at present, as the results in [2] are actually for MIMO channels. However, our HDA approach will likely outperform purely analog transmission as we leave the bandwidth compression to a highly optimized video encoder and only the quantization error is sent in analog. Ref. [6] transmits the quantization error of the baselayer generated by an H.264/AVC encoder in analog form. However, unlike ours, [6] transmits the analog and digital components orthogonally on the QAM carriers, and hence only binary phase shift keying (BPSK) can be used for digital transmission. Therefore, the power allocation problem is different (the analog decoder used in [6] is also not the optimal linear estimator). In contrast, in our approach the power allocation algorithm chooses the best MCS from a set. To the author's knowledge, the HDA video transmission using superposition coding and QAM has not been reported earlier.

2. SYSTEM DESCRIPTION AND PROBLEM SETUP

We will formulate the basic problem and its solution using a generic multi-channel communication system. Suppose $(x_1, \ldots, x_K)$ is a set of random variables (sources), all mean-zero, but having variances $(\sigma_1^2, \ldots, \sigma_K^2)$. The K sources are transmitted over a set of K independent Gaussian Rayleigh-fading channels, a single channel of which is shown in Fig. 1. All channels have iid Gaussian noise with power spectral density $\frac{N_0}{2}$. Let the gains of the channels be $g_1, \ldots, g_K$. We assume slow-fading and therefore the channel power gains remain constant during the transmission of a block of samples [e.g., a group-of-pictures (GOP) in video coding]. Each channel uses QAM and the signal constellations of different channels need not be the same. However, the transmitter is subject to a total power constraint P_T.

In our HDA system, the quantization errors of a digital source encoder are superimposed on each quadrature component of QAM symbols (i.e., 2 error samples per symbol). Our goal is to use the digital source encoder output as the base-layer and the analog quantization error as the refinement layer. As such the bitrate of the source encoder is fixed, to say R_s bits/s, regardless of the channel conditions. Let the output of the digital modulator on the k-th channel be the complex QAM symbol sequence $\{u_k(n)\}$, n being the discrete-time. It is assumed that the QAM constellation is power-normalized so that $E\{|u_k(n)|^2\} = 1$. Let the sequence of complex-valued quantization error samples on the k-th channel be $\{e_k(n)\}$. It is assumed that quantization errors are mean-zero and iid. The input-symbol of the k-th channel at time n is

$$v_k(n) = d_k u_k(n) + a_k e_k(n),\ k = 1, \ldots, K, \quad (1)$$

where d_k and a_k are constants chosen to satisfy the power constraint, $\sum_{k=1}^{K} E\{|v_k|^2\} \leq P_T$ (we will assume that $u_k(n)$ and $e_k(n)$ are independent). Suppose the fractions of total transmitter power allocated to the digital and analog components on the k-th channel be $0 \leq \gamma_k \leq 1$ and $0 \leq \rho_k < 1$ respectively. Then, we have $d_k = \sqrt{\gamma_k P_T}$ and $a_k = \sqrt{\rho_k P_T/(2\sigma_{e,k}^2)}$, where $\sigma_{e,k}^2$ is the variance of the quantization error samples transmitted on the k-th channel. The output of the k-th channel is

$$y_k(n) = g_k v_k(n) + z_k(n), \quad (2)$$

where $z_k(n)$ is the complex Gaussian channel noise. With OFDM in mind, we assume that channel gains $g_1, \ldots, g_K$ are known to both transmitter and receiver. Therefore, the output signal on each channel is first equalized to obtain $\tilde{y}_k(n) = y_k(n)/g_k$, which is then channel-decoded to recover the output bits of the source encoder. In order to ensure a sufficiently low bit error probability, we choose the power allocation factors γ_k such that the SNR at the digital receiver on every sub-channel is above a certain threshold which depends on the MCS of the channel. In state-of-the-art OFDM systems, a standard set of MCS is available for adaptive transmission (e.g., IEEE802.11a standard specifies 8 different coderate/QAM-order combinations). For such a set of MCSs, the minimum SNRs required for ensuring a specified receiver bit-error probability are a priory known. Given a set of MCSs and the corresponding minimum SNRs, we can solve a power allocation problem to determine $\gamma_1, \ldots, \gamma_K$ which guarantees a source bitrate of R_s while minimizing the average power consumption of the digital transmission.

Once the digital encoder output has been recovered, the transmitter's digital modulator output is duplicated by the receiver to cancel the transmitted digital signal from the channel output. The quantization error samples are then estimated from the residual signal $y'_k(n) = \tilde{y}_k(n) - d_k u_k(n)$. We will use minimum mean square error (MMSE) estimation, and to facilitate analytical developments as well as practical implementation, we confine our-selves to linear estimators. It can be shown that [16] the MMSE estimator for the quantization error is given by $\hat{e}_k(n) = b_k y'_k(n)$, where

$$b_k = \frac{1}{a_k} \frac{\rho_k S_k}{(1 + \rho_k S_k)}, \quad (3)$$

and $S_k \triangleq P_T |g_k|^2 / N_0$. Furthermore, the resulting MMSE is $E\{|e_k(n) - \hat{e}_k(n)|^2\} = \sigma_{e,k}^2/(1 + \rho_k S_k)$. The source signal on the k-th channel is then reconstructed as $\hat{x}_k(n) = \hat{x}'_k(n) + \hat{e}_k(n)$, where $\hat{x}'_k(n)$ is output of the digital source decoder. Since the reconstruction error of the sources $x_1, \ldots, x_K$ is solely due to the error in estimated quantization errors, the total MSE is

$$D = \sum_{k=1}^{K} \frac{\sigma_{e,k}^2}{1 + \rho_k S_k}. \quad (4)$$

Note that D is a function of both $\boldsymbol{\rho} = (\rho_1, \ldots, \rho_K)$ and $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)$. Given P_T, R_s, a set of permissible MCSs, and CSI $\boldsymbol{S} \triangleq (S_1, \ldots, S_K)$, our goal is to determine $\boldsymbol{\rho}$ and $\boldsymbol{\gamma}$ which minimize D. Since the bitrate R_s of the digital base-layer is fixed, the problem involves determining the bit loading profile and the analog power allocation across K channels which minimize D, subject to a total power constraint. This

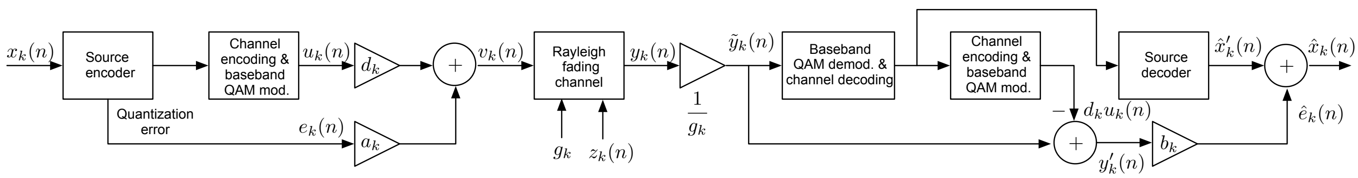

Fig. 1: A single channel of the multi-channel HDA transmission system considered in this paper.

is a mixed discrete-continuous optimization problem that appears difficult to solve. Hence, we solve the following two sub-problems in sequence.

Problem 1: Given P_T, $\boldsymbol{S}$, a set of MCSs (i.e., a set of permitted transmission bitrates) and their minimum SNR requirements, we determine the optimal allocation of R_s (bits/s) for K channels which minimizes the total average power required. This gives us an initial solution for $\boldsymbol{\gamma}$, say $\boldsymbol{\gamma}_0 = (\gamma_1^{(0}, \ldots, \gamma_K^{(0)})$.

Problem 2: Given P_T, $\boldsymbol{\gamma}_0$, $\boldsymbol{S}$, and $\sigma_{e,k}^2$, determine $\boldsymbol{\gamma}_1 \triangleq (\gamma_1^{(1)}, \ldots, \gamma_K^{(1)})$ and $\boldsymbol{\rho}$, which minimize D, where $\gamma_k^{(1)} = \gamma_k - \gamma_k^{(0)}$, such that $\sum_{k=1}^{K}(\gamma_k + \rho_k) \leq 1$ for $0 \leq \gamma_k^{(0)} \leq 1$, $0 \leq \gamma_k^{(1)} \leq 1$, and $0 \leq \rho_k < 1$.

The rationale behind this approach is as follows. In problem 1, we allocate the minimum power required to reliably transmit the digital base-layer. In problem 2, we then use the remaining power to transmit the analog quantization error so that the average distortion is minimized.

3. OPTIMAL POWER ALLOCATION

Methods for efficiently solving the Problem 1 are well known (see for example [17]), and the details of which we skip for brevity. We here present the solution to Problem 2.

Given a solution to Problem 1, say $\gamma_k^{(0)}, k = 1, \ldots, K$, we know that an SNR of at least $\eta_k^* = \gamma_k^{(0)} S_k$ is required at the input to the digital decoder on the k-th channel to maintain a negligibly low bit-error rate. Since the analog transmission on any channel essentially acts as noise to the digital transmission on that channel, whenever $\rho_k > 0$ we must choose $\gamma_k^{(1)}$ such that the SNR

$$\frac{(\gamma_k^{(0)} + \gamma_k^{(1)})S_k}{1 + \rho_k S_k} = \eta_k^*. \tag{5}$$

It follows that $\gamma_k^{(1)} = \rho_k \gamma_k^{(0)} S_k$. The total power constraint is thus given by $\sum_{k=1}^{K}\left[\rho_k(1 + \gamma_k^{(0)} S_k) + \gamma_k^{(0)}\right] \leq 1$. By defining $\beta_k \triangleq (1 + \gamma_k^{(0)} S_k)$ and $B \triangleq (1 - \sum \gamma_k^{(0)})$, the power constraint can be written as

$$\sum \beta_k \rho_k \leq B. \tag{6}$$

We can solve the Problem 2 using the Lagrange multiplier method. The Lagrangian is

$$J(\boldsymbol{\rho}) = D(\boldsymbol{\rho}) + \lambda \sum(\beta_k \rho_k - B), \tag{7}$$

where $\lambda \geq 0$ is the Lagrange multiplier. By letting $\partial J / \partial \rho_k = 0$ for $k = 1, \ldots, K$ and using the KKT conditions, we find the following optimal solution (details omitted for brevity).

Let $T_k \triangleq S_k/\beta_k$. Without a loss of generality, we assign the sources to channels such that $\sigma_{e,1}^2 T_1 \geq \sigma_{e,2}^2 T_2 \ldots \geq \sigma_{e,K}^2 T_K$. Let K' be such that, $\sigma_{e,j}^2 T_j \geq \lambda^*$ for $j = 1, \ldots, K'$, and $\sigma_{e,j}^2 T_j < \lambda^*$ for $j = K'+1, \ldots, K$, where

$$\lambda^* = \frac{\left(\sum_{k=1}^{K'} \sqrt{\frac{\sigma_{e,k}^2}{T_k}}\right)^2}{\left(B + \sum_{k=1}^{K'} \frac{1}{T_k}\right)^2}. \tag{8}$$

The optimal analog power allocation is then given by

$$\rho_k^* = \left(\sqrt{\frac{\sigma_{e,k}^2}{\lambda^* \beta_k S_k}} - 1\right)^+, \tag{9}$$

where $(x)^+$=x if $x > 0$ and $(x)^+ = 0$ otherwise. The corresponding digital power allocation is $\gamma_k^* = (1 + \rho_k^* S_k)\gamma_k^{(0)}$.

4. VIDEO TRANSMISSION

This section presents preliminary experimental results obtained with video transmission using the proposed HDA approach. An extended version, including a comprehensive set of results will be reported in a future publication.

OFDM channel model- OFDM system parameters were taken from the IEEE 802.11a WLAN standard [11] which divides a 20 kHz bandwidth into 64 sub-channels (312.5 kHz each). The frequency selective fading channel was simulated using the ITU pedestrian-B model [18].

Video source- We used 328×288, (CIF) resolution, 30 frames/sec, color video sequences in YUV 4:2:0 format, where each component is 8 bits/pixel each. However, the chrominance components U and V are sub-sampled by a factor of two. Since the bandwidth of these sequences in uncompressed form is about 2.3 MHz, only 4 OFDM sub-channels were used, resulting in a bandwidth compression of about 1.8.

HDA system- The digital base-layer was generated using the JM 19.0 reference software for the H.264/AVC video codec [19]. The baseline profile and a GOP size of 8 frames were used, and it was assumed that the channel state remains unchanged for the duration of a GOP. A packet length of 256 QAM symbols was used. For packetization, video data was

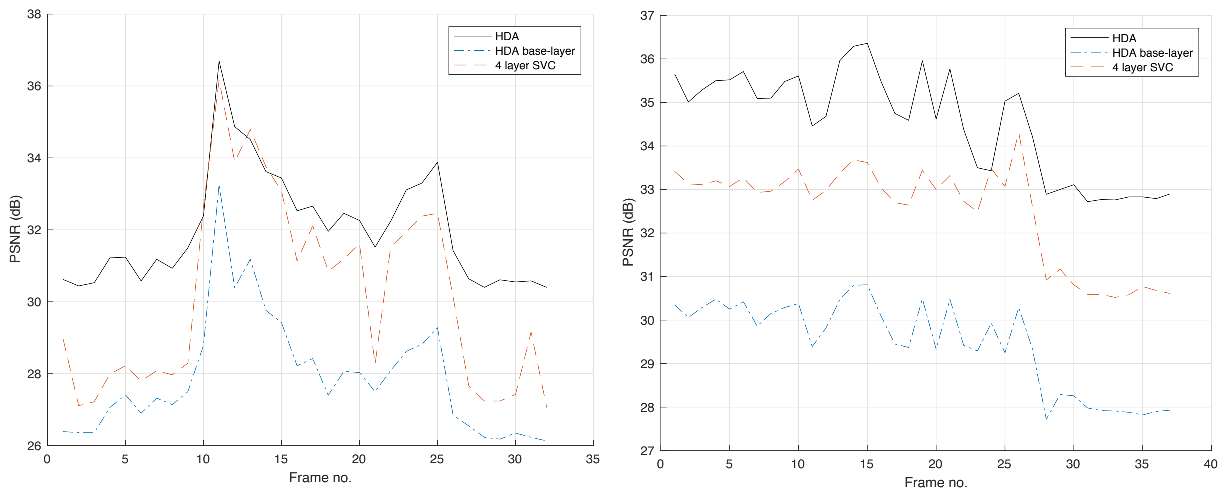

Fig. 2: *PSNRs of HDA coding and layered coding of two video sequences: Football (left) and Foreman (right).*

encapsulated in RTP packets, each of which used a 24-bit header to carry sequencing information of video data. In addition to the H.264 bitstream, some data packets were also used to carry the side-information required for decoding the analog packets. The bit-payload of a packet would depend on the particular MCS chosen by the digital power loading algorithm. The same set of MCSs as in IEEE802.11a which includes channel coderates 1/2, 2/3, and 3/4, and 2, 4, 16, and 64 QAM formats [11, Table 78] were used. For channel coding, the rate-1/2 convolutional encoder with puncturing as specified in [11] was used.

For analog transmission, each quantization error frame in a GOP was divided into 8×8 pixel blocks, and the corresponding blocks in all 8 frames of a GOP (512 samples), which we will henceforth refer to as a macro GOP (MGOP), were transmitted in a single packet (256 QAM symbols). Thus each GOP consists of 1584 (Y component)+ 2×396 (U and V components)= 2376 MGOPs. The power allocation was performed across all 4 sub-channels on a packet-by-packet basis. For this purpose, all MGOPs were sorted according to their variances [$\sigma^2_{e,k}$ in (8) and (9)], and the MGOPs with higher variances were assigned to channels with higher SNRs. Since we only used 4 sub-channels to send all 2376 analog packets and also due to the power constraint, some MGOPs with the lowest variances were entirely dropped by the power allocation algorithm. The variance of the samples in each MGOP [required by the linear decoder (3)] was quantized to 5 bits and transmitted to the decoder in data packets. Additionally, the sequencing information of each analog packet was sent using 12 bits, requiring an overhead of 17 bits/analog-packet.

4-layer SVC- For comparison purposes, we also implemented a digital layered video transmission scheme in which MCSs for OFDM sub-channels were adaptively selected based on CSI, This was achieved by a bit-loading algorithm which maximized the total bitrate subject to a total power constraint [17]. Thus, the total available bitrate and hence the number of video layers transmitted would vary depending on the channel state. For layered transmission, the source video was encoded into a base-layer and 3 refinement layers using the JSVM reference software for the H.264/SVC video codec [20]. As in the case of the HDA system, a GOP size of 8 frames was used. The video bit-stream was transmitted using RTP packets with a 24-bit header.

A number of standard CIF video sequences were used to test the HDA system. In order to compare the HDA system and the 4-layer SVC system, we computed the peak signal-to-noise (PSNR) of video-frames, defined by $10\log_{10}\left(\frac{255^2}{D}\right)$ dB, where D is the average MSE of R, G, and B color components of a frame. Fig. 2 shows the observed PSNRs of two common video sequences, when the average channel SNR is 18 dB. The figures also include the PSNR of only the base-layer of the HDA system. The intervals where there is a larger gap between the HDA and layered coding are segments containing high motion activity. The *Football* sequence has high motion activity than the *Foreman* sequence. One clear advantage of the HDA coding is due to the fact that power is more efficiently used when the refinement information has an infinite resolution. The use of a few discrete layers is less power efficient. Another issue is that any multi-layer video codec has to sacrifice the quality to achieve the scalablity property. Of course, the HDA baselayer can be generated using a single layer codec. This is also evident from Fig. 2 (left) which shows that, at times layered coding is not much better than the HDA baselayer.

5. FUTURE WORK

HDA system will be compared with a purely analog system similar to [1,2]. An extension to MIMO-OFDM channels will be considered.

6. REFERENCES

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

[2] X. L. Liu, W Hu, C. Luo, Q. Pu, F. Wu, and Y. Zhang, "Parcast+: Parallel video unicast in MIMO-OFDM WLANs," *IEEE Trans. Multimedia*, vol. 16, no. 7, pp. 2038–2051, Nov. 2014.

[3] H. Cui, C. Luo, C. W. Chen, and F. Wu, "Robust uncoded vodeo transmission over wireless fast fading channels," in *IEEE INFOCOM 2014*, 2014, pp. 73–81.

[4] K. R. Rao, D. N. Kim, and J. J. Hwang, *Video coding standards: AVS China, H.264/MPEG-4 PART 10, HEVC, VP6, DIRAC and VC-1*, Signals and Communication Technology. Springer Verlag, 2014.

[5] L. Yu, H. Li, and W. Li, "Wireless scalable video coding using hybrid digital-analog scheme," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 2, pp. 331–345, Feb. 2014.

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

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

[8] Y. Liu, X. Lin, N. Fan, and L. Zhang, "Hybrid digital-analog video transmission in wireless multicast and multiple-input multiple-output system," in *Journal of Electronic Imaging*, 2016, vol. 25 (1), pp. 013006:1–12.

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

[10] D. Dardari, M. G. Martini, M. Mazzotti, and M. Chiani, "Layered video transmission on adaptive ofdm wireless systems," *EURASIP Journal on Applied Signal Processing*, vol. 10, pp. 1557–1567, 2004.

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

[12] R. van Nee and R. Prasad, *OFDM for Wireless Multimedia Communications*, Artech House Inc., 2000.

[13] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Predictable 802.11 packet delivery from wireless channel measurements," in *Proc. ACM SIGCOMM*, 2010, pp. 159–170.

[14] S. J. Wee, M. O. Polley, and W. F. Schreiber, "A scalable source coder for a hybrid HDTV terrestrial broadcasting system," in *Int. Conf. Image Processing (ICIP)*, 1991.

[15] W. F. Schreiber, "Advanced television systems for terrestrial broadcasting: Some problems and some proposed solutions," *Proc. IEEE*, vol. 83, no. 6, pp. 958–981, Jun. 1993.

[16] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall, 1993.

[17] J. Campello and J. M. Cioffi, "Optimal discrete loading (98-341)," Tech. report, Information Systems Lab., Stanford University, 1998.

[18] "Universal Mobile Telecommunications system (UMTS); Selection procedures for the choice of radio transmission technologies of the UMTS (UMTS 30.03 version 3.2.0)," Tech. report, European Telecom. Standards Institute, 1998.

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

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