



**HAL**  
open science

## Joint source channel coding optimization for heterogeneous access networks in multiuser scenario

Laura Toni, Lorenzo Rossi, Nazim Agoulmine, Jean-Guy Fontaine

### ► To cite this version:

Laura Toni, Lorenzo Rossi, Nazim Agoulmine, Jean-Guy Fontaine. Joint source channel coding optimization for heterogeneous access networks in multiuser scenario. 12th IFIP/IEEE International Symposium on Integrated Network Management (IM 2011), IEEE Communications Society, May 2011, Dublin, Ireland. pp.1005-1012, 10.1109/INM.2011.5990514 . hal-00745105

**HAL Id: hal-00745105**

**<https://hal.science/hal-00745105v1>**

Submitted on 20 Jan 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

# Joint Source Channel Coding Optimization for Heterogeneous Access Networks in Multiuser Scenario

Laura Toni<sup>1</sup>, Lorenzo Rossi<sup>1</sup>, Nazim Agoulmine<sup>2</sup>, and Jean-Guy Fontaine<sup>1</sup>

<sup>1</sup>IIT, Italian Institute of Technology, Genova, Italy.

<sup>2</sup>IBISC/LRSM, University of Evry, Val d'Essonne, France

E-mail: {laura.toni, lorenzo.rossi}@iit.it, nazim.agoulmine@iup.univ-evry.fr, jean-guy.fontaine@iit.it

**Abstract**—In a multiple access technologies scenario, aimed at providing the “always best connected” (ABC) service, the future fourth-generation wireless networks will be characterized by an increasing heterogeneity. When a multimedia transmission is considered, the ABC can be improved by the optimization of the source coding. In this work, we propose a joint source channel coding (JSCC) optimization technique, able to adapt the system parameters to channel variations, providing at the same time a high level of scalability of the encoded video bitstream. With the proposed JSCC method, we address the ABC issue in multiuser scenario affected by multiple interference. The results show that, despite the simplicity of the algorithm, the adaptive system experiences an improvement of the performance compared to fixed coding schemes.

## I. INTRODUCTION

In recent years, with the rapid convergence of multimedia, Internet and wireless communications, there is a growing trend in the heterogeneity of emerging integrated wired/wireless networks. Future wireless systems are aimed at providing always best connected (ABC) service, which means that in a multitude of access technologies such as long term evolution networks (LTE), universal mobile telecommunications systems (UMTS), wireless local area network (WLAN), and worldwide interoperability for microwave access (WiMAX), the user is not only connected but he experiences the best possible quality of service (QoS) anytime and anywhere. This goal leads to heterogeneity of the network, in terms of channel bandwidths, mobility levels of terminals, end-user quality-of-service (QoS) requirements. To face this need an intelligent access network selection method, able to select the network that better satisfies the user requirements [1]–[4], has to be considered. In the literature, several works have investigated the best network selection algorithm and the related handoff techniques to achieve the ABC taking into account user requirements and mobility.

In multimedia transmission application, several works have investigated source coding techniques able to guarantee a sufficient quality level in ABC heterogeneous networks. In a multiuser scenario, in which each user has its own constraints and different scheduling priorities, for example, the bandwidth dedicated to the multimedia transmission could be variable. Thus, the decoder should be able to reconstruct the video from

several ranges of received bit rate. For these reasons, the aim of video coding studies is changed into the optimization of the video quality over a given bit rate range instead of a given bit rate [5], [6], meeting the requirements of heterogeneous networks. Progressive source coding techniques are able to meet these requirements, allowing partial decoding at various resolution and quality levels from a single compressed bitstream. The fine granularity scalable (FGS) video coding proposed by the MPEG-4 [7] committee, for example, is characterized by a base layer (BL), which guarantees the basic quality of the encoded video, and a single enhanced layer (EL), encoded in a progressive manner. Although the BL is encoded at a certain rate such that it will not overflow the channel capacity, the EL, which provides improved quality with respect to the BL, can be compressed over any desired bit rate range. Thanks to the progressive nature of the EL, the decoder is able to reconstruct the EL at any received bit rate. While in previous scalable video coding the EL is entirely received or lost, with FGS any portion of the EL correctly received improves the quality of the decoded video bitstream. However, FGS maintain the main drawback of progressive source coders: they are usually extremely sensitive to channel impairments which can be severe in mobile wireless links due to multipath signal propagation, delay and Doppler spreads, and other effects. Sometimes a single error can cause an unrecoverable loss in synchronization between encoder and decoder, and produce substantial quality degradation.

To overcome this limit, network-adaptive scalable video coding and cross-layer optimization techniques have always been under intense research [8]–[10]. Due to the robustness achieved by multiple description (MD) coding<sup>1</sup>, early studies were focused on the optimization of the forward error correction (FEC), protecting the information symbols against channel impairments by means of a MD based channel coding [12]–[14]. In [15], the authors studied the packet loss behavior of an Orthogonal Frequency-Division Multiplexing (OFDM) system and constructed multiple independent descriptions using a FEC-based strategy for progressive images transmission.

<sup>1</sup>In MD coding, the source bitstream is fragmented in several descriptions, leading to a more reliable image transmission over wireless bursty channels [11].

Coding in the time domain have been included in [16], where the authors proposed a bidimensional channel coding. In our previous work [17], we assumed that the progressive source stream exploits the MD based coding not at the physical layer, but at the application layer. We proposed a source coder able to encode a video into layered packet-streams, robust to channel impairments. The scalability of the proposed method was reached by fragmenting the layered packet-streams in independent descriptions, and it could be tuned by varying the setup parameters in the source coder.

In this work, we present an optimization technique for the transmission of video images encoded with the source coder recently presented in [17]. In a multi-user scenario in which each single user has to cope with the channel impairments and the interference due to other users, we propose an adaptive joint source channel coding (JSCC) technique for the transmission of progressive video sequences. Based on both the bandwidth dedicated to the considered terminal, and the level of interference due to the other terminals, the proposed method adapts the channel code rate in the time domain to the small-scale fading variations of the channel, allowing the maximization of the transmitted throughput, without degradation in the system performance. The different amount of redundancy in each transmission leads to a variable source information rate which will be met by opportunistically varying the source coding setup parameters.

This paper is organized as follows. In Section II, the system model is presented. Section III describes the source coding mechanism considered in this paper, and discusses some of the associated tradeoff issues. In Section IV, the description of the algorithm for practical packetized channel is provided, together with the problem formulation for the JSCC optimization. Section V illustrates the chosen metrics for performance benchmarking. In Section VI, simulation results and discussion are provided. Finally, Section VII gives summary and conclusions.

## II. SYSTEM MODEL

We consider a multiuser heterogeneous scenario for multimedia transmission over a packetized network, in which the transmission of the useful terminal might be interfered by the other users, as depicted in Fig. 1. The source information is protected by Reed-Solomon (RS) channel coding in the time domain and modulated with  $M$ -ary quadrature amplitude modulation ( $M$ -QAM) constellation signaling over Rayleigh fading, with  $N$ -branches multiple channels reception. Independent, identically distributed (i.i.d.) fading and same shadowing level over all branches, (i.e., microdiversity) are considered. The optimum combining technique is employed to process signals over branches [18]. In particular, the minimum mean square error (MMSE) and maximum ratio combining (MRC) techniques are considered for interfered and uninterfered systems, respectively.

We denote by  $E_s$  the mean (averaged over the fading) transmitted symbol energy, by  $\mathbf{h} = [h_1, h_2, \dots, h_N]^T$  the channel vector, whose elements  $h_k$  represent fading gain for

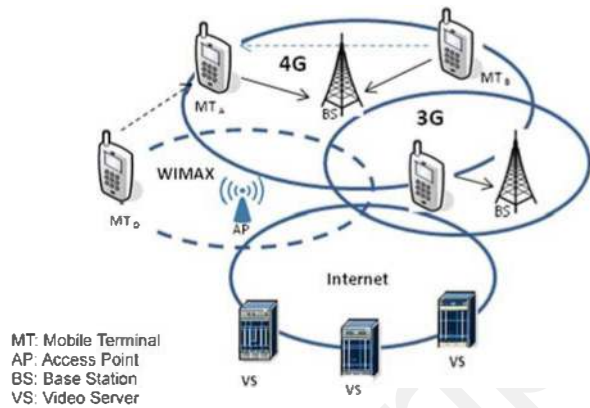


Fig. 1. Considered scenario.

the  $k$ -th channel and are complex Gaussian random variables (RVs) with  $\mathbb{E}\{h\} = 0$  and  $\mathbb{E}\{|h|^2\} = 1$  (i.e.,  $h \sim \mathcal{CN}(0, 1)$ ), and by  $\mathbf{n}(k)$  the additive white Gaussian noise (AWGN) with mean zero and  $\mathbb{E}\{\mathbf{n}(k)\mathbf{n}(k)^H\} = N_0\mathbf{I}_N$ , where  $N_0$  is the one-sided thermal noise power spectral density. The mean signal-to-noise ratio (SNR) per branch can be expressed as  $\bar{\gamma} = \mathbb{E}\{|h|^2\}E_s/N_0 = E_s/N_0$ .

Adaptive channel coding techniques allow to achieve the best throughput accordingly with channel conditions, minimizing the outage probability [19]–[21]. In particular, a set of  $(J+1)$  channel code rates can be adopted,  $\{R_0, R_1, \dots, R_J\}$ . For uninterfered adaptive systems, the optimal channel code rate is chosen depending on the instantaneous SNR,  $\gamma$ . When, even with the lowest channel code rate  $R_0$ ,  $\gamma$  is not sufficient to guarantee the target packet error rate (PER)  $P_e^*$ , the system is in outage. The forward error correction (FEC) level is chosen comparing the  $\gamma$  value with thresholds, that are set to guarantee the target PER when the  $j$ -th code rate is adopted, that is  $P_e(\gamma^*, R_j) = P_e^*$ . In particular, when the SNR value falls within the  $j$ -th region  $(\gamma_j^*, \gamma_{j+1}^*]$ , the  $j$ -th channel code rate  $R_j$  is adopted.

In the presence of  $N_1$  interfering signals, the received signal on the generic  $k$ -th branch after matched filtering and sampling at the symbol rate is given by  $z_k(l) = z_k^{(P)}(l) + \mathbb{I}z_k^{(Q)}(l)$ , with  $l$  representing the time index<sup>2</sup>, and  $\mathbb{I}$  the imaginary unit. The signal at the output of the receiving antennas  $\mathbf{z}(l) = [z_1(l), \dots, z_N(l)]^H$  can be written as [22]

$$\mathbf{z}(l) = \sqrt{E_s}\mathbf{h}_s b_s(l) + \sum_{i=1}^{N_1} \sqrt{E_i} \mathbf{h}_i b_i(l) + \mathbf{n}(l) \quad (1)$$

where  $E_i$  is the mean symbol energy transmitted by the  $i$ -th interfering signals,  $b_s(l) = b_s^{(P)}(l) + \mathbb{I}b_s^{(Q)}(l)$  is the useful signal sample, and  $b_i(l) = b_i^{(P)}(l) + \mathbb{I}b_i^{(Q)}(l)$  is the interfering signal

<sup>2</sup>Note that the superscripts (P) and (Q) stand for the in-phase and quadrature components, respectively.

sample from the  $i$ -th interferer, with  $i = 1, \dots, N_I$ . Without loss of generality, we consider ordered interferers such that  $E_i \geq E_{i+1}$ . We assume a slow frequency flat Rayleigh distributed fading. The vectors  $\mathbf{h}$  and  $\mathbf{h}_i = [h_{i,1}, h_{i,2}, \dots, h_{i,N}]^T$ , with  $i = 1, 2, \dots, N_I$  have distribution  $\mathcal{CN}(0, \mathbf{I}_N)$ , and we denote by  $\mathbf{H}_I = \{\mathbf{h}_i\}_{i=1, \dots, N_I}$  the interference channel matrix, where  $N_I$  is the number of interfering signals. As analyzed in [23], when the useful signal is affected by fading, then the Gaussian approximation for the interferers is justified. Thus,  $b_1(l), \dots, b_{N_I}(l)$  are independent zero-mean complex Gaussian RVs, each with unitary variance.

As far as the optimum combining technique is concerned, we recall that the optimum weight vector to combine signals over branches and maximize the SNIR is given by

$$\mathbf{w} = \alpha \mathbf{R}^{-1} \mathbf{h}$$

where  $\alpha$  is an arbitrary constant and  $\mathbf{R}$  the covariance matrix. The components of the received vector  $\mathbf{z}(k)$  are combined through the vector of the optimum weight  $\mathbf{w}^H$  giving

$$r(k) = \mathbf{w}^H \mathbf{z}(k).$$

In the absence of interference,  $\mathbf{w} = \mathbf{h}$  and the optimum combiner (OC) reduces to the well known MRC [24], [25]. We consider an overloaded system<sup>3</sup> (i.e.,  $N_I \geq N$ ) for which the signal-to-noise-plus-interference ratio (SINR) at the combiner output can be written as  $\gamma_T = E_s \mathbf{h}^H \mathbf{R}^{-1} \mathbf{h}$ .

We consider that the information symbols are protected by systematic  $(n, k)$  RS codes<sup>4</sup> which can correct up to  $t = \lfloor (n - k + 1)/2 \rfloor$  erroneous bits, where  $\lfloor x \rfloor$  denotes the largest integer less than or equal to  $x$ . Starting from the exact bit error probability (BEP) [26], [27]

$$P_b(\gamma_T) = \frac{2}{\sqrt{M} \log_2(\sqrt{M})} \sum_{h=1}^{\log_2(\sqrt{M})} \sum_{n=0}^{(1-2^{-h})\sqrt{M}-1} (-1)^{\lfloor \frac{in2^{h-1}}{\sqrt{M}} \rfloor} \times \left( 2^{h-1} - \left\lfloor \frac{n2^{h-1}}{\sqrt{M}} + \frac{1}{2} \right\rfloor \right) Q \left( (2n+1) \sqrt{\frac{3\gamma_T}{(M-1)}} \right) \quad (2)$$

where  $Q(\cdot)$  is the Gaussian- $Q$  function [24], the PER is derived as

$$P_e(\gamma_T, R) = 1 - \sum_{j=0}^{t(R)} \binom{N_t}{j} P_b(\gamma_T)^j (1 - P_b(\gamma_T))^{N_t-j} \quad (3)$$

where  $N_t$  is the total number of bits in which  $k$  information bits are encoded.

### III. PROPOSED SOURCE CODING

We remind the reader to [17] for a detailed description of the proposed source coding. In the following, we just provide a brief overview. Starting from the Joint Photographic Experts

<sup>3</sup>In this paper, for sake of brevity, we consider the overloaded case only, but the analysis can be extended to the underloaded case.

<sup>4</sup>In each  $(n, k)$  RS codeword,  $k$  information symbols are encoded into  $n$  total symbols. While  $k$  is variable,  $n$  is constant in each codeword and equals  $L$ .

Group (JPEG) encoder, we customize the algorithm for a sequence of frame (or video), keeping in mind both the goals of scalability and robustness to packet loss. We assume that one I-frame occurs every  $(N_f - 1)$  P-frames.<sup>5</sup> The I is an intra frame, encoded without any references, while the P is an inter frame encoded with the previous I-frame as reference. For both the I- and P-frames, first the image is layered encoded, then each layer is compressed in a scalable manner.

#### A. Layered Coding

The proposed algorithm allows to encode each YUV<sup>6</sup> frame into  $L$  layers, each one providing a decoded image at a better quality than the previous layer. When an I-frame is considered, the layer 0 (BL) is obtained encoding the original frame after a lowpass filtering and downsampling, leading to a low quality stream with high coding efficiency. Denoting by  $F_I^0$  the filtered and downsampled I-frame<sup>7</sup>, the JPEG encoded BL is  $\mathcal{I}^0 = \text{JPG}(F_I^0)$ . To obtain the successive ELs, the original I-frame is subtracted to the frame reconstructed from the lower layers. The difference is encoded at a given resolution and compression coefficient. Note that higher layers will achieve a better quality of the reconstructed image, at the price of a lower compression efficiency. Each  $l$ -th layer of each I-frame is recursively encoded as follows:

$$\begin{cases} \mathcal{I}^0 &= \text{JPG}(F_I^0), \\ \mathcal{I}^1 &= \text{JPG}[F_I^0 - \text{IJPG}(\mathcal{I}^0)], \\ \mathcal{I}^l &= \text{JPG}[F_I^0 - (\text{IJPG}(\mathcal{I}^{l-1}) + \text{IJPG}(\mathcal{I}^{l-2}))], \quad l \geq 2 \end{cases} \quad (4)$$

where  $\text{IJPG}(\mathcal{I}^{l-1})$  is the  $(l-1)$  JPEG decoded layer, and  $F_I^0$  is the original frame.

For the P-frames, the encoding process uses the  $\text{IJPG}(\mathcal{I}^0)$  as reference frame. Since the BL of the I-frame is a reference for the successive frames, it has to be correctly received to avoid drift issues at the decoder. To encode the BL, after a lowpass filtering and a downsampling, the P-frame  $F_P$  is compared to the BL of the I-frame, and it is encoded as follows

$$\mathcal{P}^0 = \text{JPG}[\text{IJPG}(\text{JPG}(F_P^0)) - \text{IJPG}(\mathcal{I}^0)]. \quad (5)$$

The other layers are similarly encoded using the lower layer as reference: the image reconstructed from the lower layer is compared to the original P frame, and the current layer is recursively encoded as follows

$$\begin{cases} \mathcal{P}^0 &= \text{JPG}[\text{IJPG}(\text{JPG}(F_P^0)) - \text{IJPG}(\mathcal{I}^0)], \\ \mathcal{P}^1 &= \text{JPG}[F_P^0 - (\text{IJPG}(\mathcal{P}^0) + \text{IJPG}(\mathcal{I}^0))], \\ \mathcal{P}^l &= \text{JPG}[F_P^0 - (\text{IJPG}(\mathcal{P}^{l-1}) + \text{IJPG}(\mathcal{P}^{l-2}))], \quad l \geq 2 \end{cases} \quad (6)$$

where  $F_P^0$  and  $F_P$  are the downsampled and filtered P-frame and the original one, respectively.

<sup>5</sup> $N_f$  is an arbitrary parameter; in this paper  $N_f = 12$ .

<sup>6</sup>YUV 4:2:0 is the color space of each frame of the sequence.

<sup>7</sup>Note that  $k$ -th I-frame should be denoted by  $F_{I,k}$ . For sake of notation we omit the temporal index  $k$ , and we explicit it only if necessary.

## B. Scalable Coding

In order to reach a scalability with a resolution higher than the  $L$  levels of the layered stream, we adopt a MD-based coding technique. Every layer is divided into a matrix of blocks, each one independent from the others. Based on the user requirements (in terms of perceived quality and available bandwidth), a detector will allow the transmission of only a sub-set of the encoded blocks. The selection of each block, based on the mean square error (mse) parameter, is fully described in [17]. In short, we assume that each block has dimension  $M \times M$  pixels, and  $B_j^k$  denotes the  $j$ -th block of the  $k$  frame<sup>8</sup>. Once the  $j$ -th block is encoded, the detector compares it to the  $j$ -th block of the previous frame, by evaluating the mse as follows

$$\text{mse} = \frac{1}{M^2} \sum_{n=1}^M \sum_{m=1}^M |X_{n,m}(B_j^k) - X_{n,m}(B_j^{k-1})|^2 \quad (7)$$

where  $X_{n,m}(B_j^k)$  is the pixel at the position  $(n, m)$  for the  $j$ -th block of the  $k$ -th frame. If the mse value is higher than a given threshold ( $\text{TH}_{\text{mse}}$ ), the block is transmitted. The mse threshold is tuned on the user requirements and the available bandwidth. The lower the threshold, the higher the quality and the transmitted rate.

At the decoder side, each  $i$ -th block that has been filtered by the selector (i.e., that has not been transmitted) is replaced with the  $i$ -th block of the previous frame. Unfortunately, replacement of blocks with mse below the threshold can propagate permanent distortion through the video sequence. To reduce this effect, we introduce a transmission *aging counter parameter* and an *aging threshold* (denoted by  $A_{\text{th,tx}}$ ). The aging counter parameter of each block is incremented every time the block is filtered by the mse comparator, and reset when transmitted. If the counter reaches  $A_{\text{th,tx}}$ , the block is sent in any case, disregarding the mse threshold. In summary, the block  $B_j^k$  is transmitted if one of the following conditions is satisfied

$$\begin{cases} \text{mse} > \text{TH}_{\text{mse}}, \\ a_j \geq A_{\text{th,tx}}. \end{cases} \quad (8)$$

where  $a_j$  is the aging counter parameter of the  $j$ -th block.

## IV. PROBLEM FORMULATION

In the previous section we described the source encoder in details, here, for practical packetized channels, we state the problem formulation and analyze the algorithm, describing in details the process of packet loss concealment.

In an ideal transmission (with no packet loss), the decoding process is strictly correlated to the encoding algorithm: the inverse of the operations computed during the encoder occurs at the decoder. For example,  $\hat{I}^l$  is the  $l$ -th layer of an I-frame reconstructed from the decoder. It is evaluated as the JPEG decoding of the difference between the  $l$ -th layer and

<sup>8</sup>Each layer  $l$  has its own  $j$ -th block, that should be denoted by  $B_j^{k,l}$ . Since each block  $B_j^{k,l}$  is always compared to blocks of the same layer ( $B_j^{m,l}$ ), we omit the layer superscript.

the previous ones, plus the JPEG decoding of the previous layers. It means that  $\hat{I}^l$  is evaluated as follows

$$\hat{I}^l = \text{IJPEG}(I^l) + \text{IJPEG}(I^{(l-1)}) + \text{IJPEG}(I^{(l-2)}), \quad l \geq 2. \quad (9)$$

This decoding process can be extended to any block of the layers. Even at the receiver side, we introduce an aging counter parameter and an aging threshold  $A_{\text{th,rx}}$ . Every time that  $B_j^{k,l}$  is lost, the aging counter is incremented and the block is replaced by  $B_j^{k-1,l}$ . When the counter reaches  $A_{\text{th,rx}}$ , the missing block is no more replaced, producing a ‘‘hole’’ in the layer  $l$ . The aging counter is reset at each correct reception of the block.

The problem of the mse threshold and aging parameters (i.e., set-up parameters) optimization can be formulated as follows. In a multi-user scenario, based on both the access network constraints and the user requirements, a given bit-rate  $R_c$  is available for the transmission of the video sequence over wireless channel. Since the main focus of this work is to propose a JSCC algorithm for the transmission of video sequences, the dynamic allocation of bandwidth in a multiuser scenario is not considered this paper. Although here we assume a fixed bit-rate available for the useful user, the algorithm can be extended to any dynamic bandwidth allocation algorithm. Each frame is transmitted in  $N_p$  packets, each of those consisting of  $N_t$  modulated symbols. Denoting by  $M$  the adopted modulation,  $n = N_t \times \log_2(M)$  is the number of bits per packet, and  $S = n \times N_p = N_t \times \log_2(M) \times N_p$  is the bunch of bits per frame. Note that, we impose that  $S \leq \bar{b}_{\text{pf}}$ , where  $\bar{b}_{\text{pf}}$  is the channel capacity in terms of average bits per frame, given by

$$\bar{b}_{\text{pf}} = R_c[\text{bps}] \times (R_f[\text{fps}])^{-1} \quad (10)$$

where  $R_f$  is the frame-rate usually set to 25 – 30 fps.

Since a slow fading scenario is considered, the opportunist amount of FEC per packet is chosen once per frame. As already mentioned, the channel code rate is chosen depending on the instantaneous SINR. In particular, when  $\gamma_T \in (\gamma_i^*, \gamma_{i+1}^*]$ , the  $j$ -th channel code rate  $R_{\text{rs},j}(k)$  is chosen for transmission of the  $k$ -th frame. When  $\gamma_T < \gamma_0^*$  the system is in outage and no packets are transmitted. Given the total amount of FEC imposed for the  $k$ -th frame, the total source information available for the transmission of the current frame is

$$L_F(k) = N_t \times \log_2(M) \times R_{\text{rs},j}(k) \times N_p.$$

With a minimum pre-processing at the transmitter, the  $\text{TH}_{\text{mse}}$  and the aging parameter  $A_{\text{th,tx}}$  are evaluated as that values that optimize the expected distortion, meeting the constraint of the total bit-budget available for the current frame. This means that the best setup parameters  $\text{TH}_{\text{mse}}^*$ ,  $A_{\text{th,tx}}^*$  are chosen as follows

$$\begin{aligned} & \text{TH}_{\text{mse}}^*, A_{\text{th,tx}}^* \text{ such that} \\ & \min_{\{\text{TH}_{\text{mse}}, A_{\text{th,tx}}\}} \{E[D]\} B(\text{TH}_{\text{mse}}, A_{\text{th,tx}}) \leq L_F(k) \leq \bar{b}_{\text{pf}} \times R_{\text{rs},j}(k) \} \end{aligned} \quad (11)$$

where  $E[D]$  is the expected distortion of the current frame, and  $B(\text{TH}_{\text{mse}}, A_{\text{th,tx}})$  is the number of bits in which the current

frame is encoded when  $(TH_{mse}, A_{th,tx})$  are considered as setup parameters. Then, the amount of FEC selected per frame is able to guarantee a minimum PER at the receiver side, and at the same time, the source coding parameters are evaluated taking into account the video content. For a slow motion video, for example, in order to compress the video bitstream to the available source information, it might be worthy to increase the aging parameter rather than the threshold one.

It is worth noting that the mse threshold and the aging threshold enable to maximize the system performance and to optimize the quality of service even when the users handoff from a high speed technology, such as WiMAX, to a lower bit-rate technology, such as WiFi.

## V. BENCHMARK

Unfortunately, for video-streams, there is not an objective technique for the assessment of the user's perceived quality. The long-standing benchmark used for evaluating video quality is the ITU-R Recommendation BT.500 [28]. More recently the ITU-T developed the Recommendation P.910 [29] for multimedia quality assessment and the ITU-T Study Group 9 is working in cooperation with the Video Quality Expert Group (VQEG) on the maintenance and enhancement of Recommendations P.910, P.911, P.920, P.930, P.931, J.146, J.148, J.149. The VQEG was formed in 1997 to address video quality issues and to develop new objective methods using the characteristic of the human visual system. The reference objective model proposed by the VQEG is the Peak Signal-to-Noise Ratio<sup>9</sup> (PSNR). In the following section we will use the peak signal-to-noise ratio (PSNR) as benchmarking metric between the proposed codec and the state of the art MPEG-4 FGS<sup>10</sup>. FGS is adopted by the MPEG-4 standard for efficient and flexible distribution of multimedia streams over heterogeneous wired and wireless networks [31]–[33]. FGS provide also a down-scaling mechanism which adapts MPEG-4 media to network traffic and/or consumer device resources.

## VI. RESULTS

We carried out simulations on the  $352 \times 288$  CIF News, and Foreman video sequences. Similar results were obtained for both of them. Hence, in this paper, we only present the results coming from the News sequence. The physical layer simulator (channel coding, modulator, wireless channel, etc.) was implemented in MATLAB, while the source coding/decoding in C. To better understand the effect of the MD-based encoding, we consider only 2 layers, the BL and one EL, and we assume  $A_{th,tx} = A_{th,rx} = A_{th}$ . Note that in our system the BL is always correctly received, while the EL is subject to the channel impairments. We consider transmission over packetized networks. We assume 4- and 64-QAM modulation and Reed-Solomon (RS) channel code in the time domain with channel code rate  $R_{rs} \in \{0.5, 0.6, 0.7, 0.8, 1\}$ . We assume the codeword length after the RS coder fixed equal to the

<sup>9</sup>PSNR =  $10 \cdot \log_{10} \left( \frac{255^2}{mse} \right)$ .

<sup>10</sup>We refer to [6], [30] for further details on this standard.

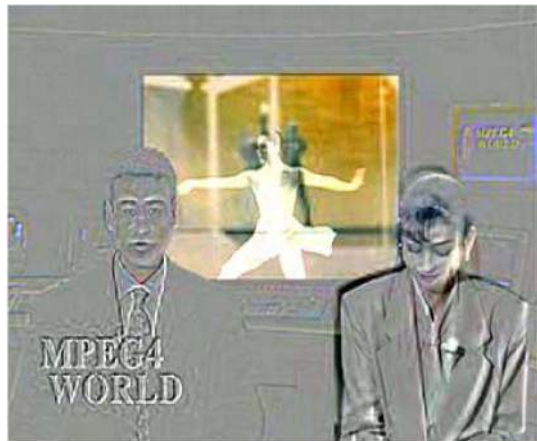


Fig. 2. Grabbing of the EL  $\mathcal{P}^1$ .

coherence time, which means  $N_t = 500$  modulated symbols. Note that for a mobile terminal, the coherence time of the fast fading is inversely proportional to the maximum Doppler frequency: with a carrier frequency of 900 MHz, the coherence time is about 72 ms and 4 ms for a mobile speed of 3 km/h and 50 km/h, respectively. Assuming a symbol period of  $66\mu s$  [34],  $N_t = 60 \div 1000$  symbols in a coherence time.

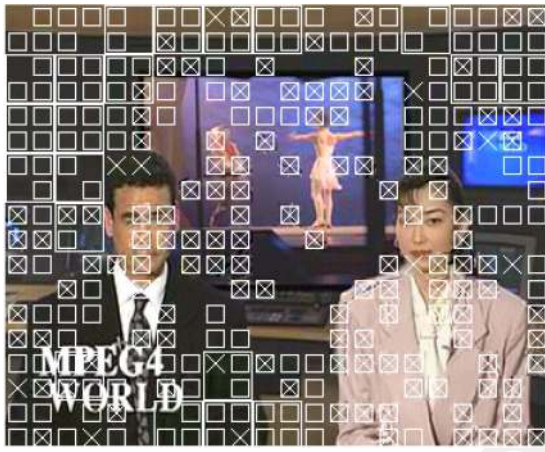
In the following, we first provide some grabbing of the encoded or decoded frames, then we provide results representing the systems performance in terms of PSNR or bit-rate. For each  $i$ -th interference, we evaluate the performance conditioned to a given level of signal-to-interference ratio (SIR), that is defined as  $SIR_i = \bar{\gamma}/\bar{\gamma}_{1,i}$ , where  $\bar{\gamma}_{1,i}$  is the mean SNR for the  $i$ -th interfering signal. This means we assume that the large-scale fading variations of the interfering level follow the mean SNR variations of the useful signal [35]. Moreover, we assume the number of interfering signals  $N_I = 2$  and  $5$ , and we define  $\Gamma_i$  as the ratio between the mean symbol energy transmitted by the  $i$ -th interfering signals  $E_i$  (from other terminals) and the energy transmitted by the considered terminal (i.e.,  $\Gamma_i = E_i/E_1$ ). We consider an ordered interfering scenario, then  $\Gamma_i < 1$ , and in particular, we define the ratio of the transmitted energies as  $\mathbf{\Gamma} = [\Gamma_1, \dots, \Gamma_{M_I}] = [1, 1/2, 1/2, 1/3, 1/3]$ .

In Fig. 2, a middle step of the codec process is displayed.  $\mathcal{P}^1$  is the JPEG encoding of the difference between the original P-frame and the reconstruction of the reference frame and the BL of the current frame. Note that this difference for the background or the static area is mostly represented by edges (contours) of the image. The area with higher information is the one with higher motion, in this case the central one.

We now consider the block decomposition proposed in our algorithm. In Fig. 3, a screen capture is shown with superimposed information about MD-based coding and packet loss concealment, with  $A_{th} = 20$ , and  $TH_{mse} = 2$ . In Fig. 3(a), to better understand the MD-based encoder, we assume to transmit the video sequence over a channel with no packet



(a)  $Plr = 0$ .



(b)  $Plr = 50\%$ .

Fig. 3. Decoded images shown with a graphical representation of the block decomposition. The sequence is encoded with  $A_{th} = 2$ , and  $TH_{mse} = 20$ , and two values of  $Plr$ . The blocks that are not transmitted are denoted by the symbol  $\square$  (big box for the BL and smaller boxes for the EL), the blocks lost during the transmission but replaced at the receiver are denoted by the symbol  $\boxtimes$ . Finally, the blocks lost during the transmission but not replaced due to the expired aging are denoted by the symbol  $\times$ .

loss<sup>11</sup> (i.e.,  $Plr = 0$ ). The big and small empty white boxes represent respectively the BL and the EL blocks filtered by the detector, that is, the BL and the EL blocks not transmitted. As expected, the central area of the frame, that is the one with highest motion in this sequence, is the most transmitted area (i.e., the one with less white boxes). In Fig. 3(b), we consider a worse case, characterized by a packet loss probability of 50%. Here, we denote by the symbol  $\boxtimes$  the packets that are lost and replaced (packet loss concealment), while the ones that are lost and not replaced (aging counter  $\geq A_{th,rx}$ ) are denoted by the symbol  $\times$ . For a given channel (i.e., for a given  $Plr$

<sup>11</sup>In this condition the parameter  $A_{th,rx}$  can be neglected.

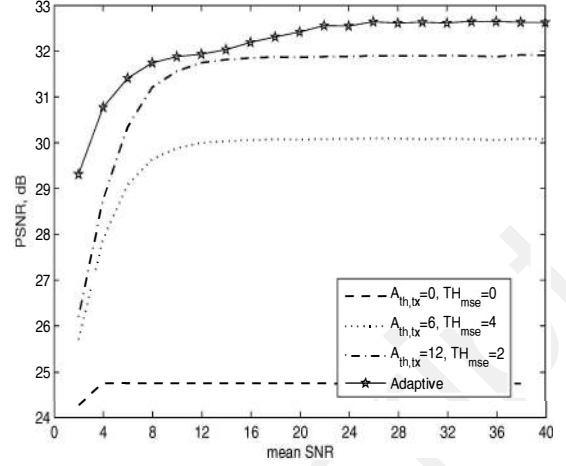


Fig. 4. Comparison between fixed and adaptive coding scheme for systems with  $M = 4$ ,  $P_c^* = 10^{-2}$ ,  $N_1 = 2$ ,  $SIR=15\text{dB}$ , and  $\Gamma_2 = 1/2$ .

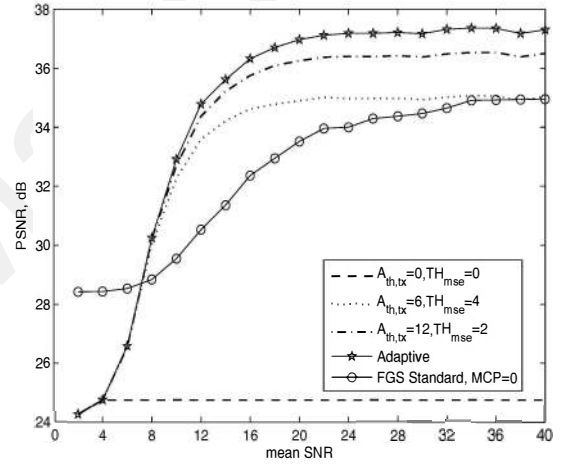


Fig. 5. Comparison between fixed and adaptive coding scheme for systems with  $M = 64$ ,  $P_c^* = 10^{-2}$ ,  $N_1 = 2$ ,  $SIR=15\text{dB}$ , and  $\Gamma_2 = 1/2$ .

parameter), the occurrence of the  $\times$  symbols depends on the receiver side aging parameter: the higher  $A_{th,rx}$ , the higher the packet loss concealment.

In the following figures, we evaluate the adaptive JSCC algorithm for the proposed source coding. We first compare the adaptive method to the fixed one, then we show the system performance for several considered scenarios. In Fig. 4 and Fig. 5, the fixed coding scheme is compared to the adaptive one for systems with constellation size 4- and 64-QAM, respectively. Moreover, in Fig. 4, the proposed source coder is compared to the standard MPEG-4 FGS. For the standard codec, we consider the same bit-rate as the one our approach achieves with  $TH_{mse} = 0$  (i.e., the highest R value). The

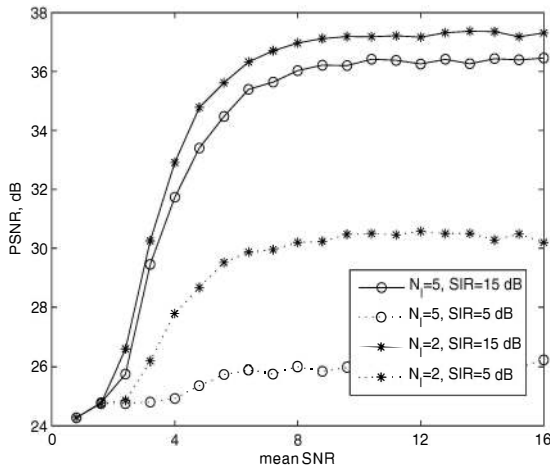


Fig. 6. PSNR vs. mean SNR for adaptive coding scheme systems with  $M = 64$ ,  $P_c^* = 10^{-2}$ ,  $N_I = 2$ , and 5,  $SIR = 15$  and 5 dB, and  $\Gamma_2 = 1/2$ ,  $\Gamma_3 = 1/2$ ,  $\Gamma_4 = 1/3$ ,  $\Gamma_5 = 1/3$ .

scenario considered is the one with  $P_c^* = 10^{-2}$ ,  $N_I = 2$ ,  $SIR = 15$  dB, and  $\Gamma_2 = 1/2$ . Both the figures depict the PSNR (mean PSNR per video sequence) as a function of the mean SNR. In both the figures, it can be observed that a fixed code rate and fixed set-up parameters are not always the best choice. For very low set-up parameters (e.g.,  $A_{th,tx} = 0$ ,  $TH_{mse} = 0$ ), the compressed bitstream might be greater than the bit-budget available for the transmission. This would lead the system to be in outage and no transmission takes place. Then, only the BL is received and a low PSNR curve might be observed. Even if some combinations of setup parameters ( $A_{th,tx}$ ,  $TH_{mse}$ ) provide high PSNR, adaptive coding scheme systematically achieves better performances. Note also that, even for high SNR values, the adaptive coding scheme outperforms the standard.

Fig. 6 depicts the PSNR as a function of the mean SNR for systems with  $M = 64$ ,  $P_c^* = 10^{-2}$ ,  $N_I = 2$ , and 5,  $SIR_1 = 15$  and 5 dB. In all the considered scenarios, for low  $\bar{\gamma}$  the system is in outage and no bits are considered for the transmission. By increasing the mean SNR, the lower the interference the greater the PSNR of the decoded video. In particular, for both the  $SIR_1$  values considered in the figures, the system with 2 interfering signals outperforms the one with 5 interfering signals, which have a total amount of interfering power greater than the case with  $N_I = 2$ .

Finally, the PSNR as a function of the number of interfering terminals  $N_I$  is provided in Fig. 7 for systems with  $M = 64$ ,  $P_c^* = 10^{-2}$ ,  $\bar{\gamma} = 20$  dB, and  $SIR = 15$  for both fixed and adaptive coding schemes. Even for high  $N_I$  values, the adaptive system achieves very good performance (PSNR > 30 dB) and outperforms the fixed coding scheme.

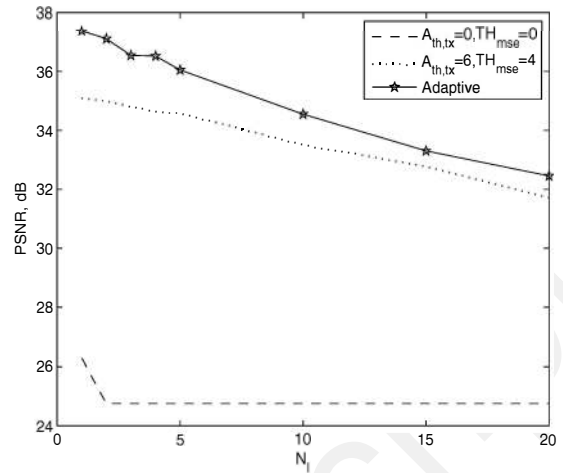


Fig. 7. PSNR vs.  $N_I$  for adaptive coding scheme systems with  $M = 64$ ,  $P_c^* = 10^{-2}$ ,  $\bar{\gamma} = 20$  dB,  $SIR = 15$  and  $\Gamma_2 = 1/2$ ,  $\Gamma_3 = 1/2$ ,  $\Gamma_4 = 1/3$ ,  $\Gamma_5 = 1/3$ ,  $\{\Gamma_i\}_{i=5...20} = 1/5$ .

## VII. CONCLUSION

In fourth-generation communications, the access network selection is a fundamental issue to achieve the always best connected (ABC) service. In multiuser heterogeneous scenarios, each terminal has to cope with channel impairments, together with the interference due to the other terminals and with the limited bandwidth imposed by any dynamic bandwidth allocation algorithm for multiuser systems. When multimedia streaming is considered, the ABC can be improved by optimizing the source coding. Starting from a multiple description based source coding method, we propose a joint source channel coding (JSCC) optimization technique, able to maximize the throughput without sacrificing the received quality of the transmitted video bitstream. In particular, in a multiuser scenario, by adapting the channel code rate to the small-scale fading variations and to the interfering level due to the other terminals, the source coder is able to impose the best setup parameters which maximize the received video quality, meeting the bitbudget constraint. The results showed that despite the interfering terminals and the limited bandwidth dedicated to the current terminal, the adaptive JSCC technique allows the user to transmit a video sequence with an acceptable video quality.

## REFERENCES

- [1] J. Lv, Y. Ma, and S. Yoshizawa, "Intelligent seamless vertical handoff algorithm for the next generation wireless networks," in *Proceedings of the 1st international conference on MOBILE Wireless MiddleWARE, Operating Systems, and Applications*, 2007.
- [2] Q.-T. Nguyen-Vuong, N. Agoulmine, and Y. Ghamri-Doudane, "Terminal-controlled mobility management in heterogeneous wireless networks," *IEEE Communications Magazine*, vol. 45, no. 4, pp. 122–129, April 2007.



- [3] L. Eastwood, S. Migaldi, Q. Xie, and V. Gupta, "Mobility using IEEE 802.21 in a heterogeneous IEEE 802.16/802.11-based, IMT-advanced (4G) network," *IEEE Wireless Commun.*, vol. 15, no. 2, pp. 26–34, April 2008.
- [4] R. Ferrus, O. Sallent, and R. Agustí, "Interworking in heterogeneous wireless networks: Comprehensive framework and future trends," vol. 17, no. 2, pp. 22–31, 2010.
- [5] M. van der Schaar and H. Radha, "Adaptive motion-compensation fine-granular-scalability (AMC-FGS) for wireless video," *IEEE Trans. on Circuits and Syst. for Video Tech.*, vol. 12, no. 6, pp. 360–371, Jun 2002.
- [6] F. Wu, S. Li, and Y.-Q. Zhang, "A framework for efficient progressive fine granularity scalable video coding," *IEEE Trans. on Circuits and Syst. for Video Tech.*, vol. 11, no. 3, pp. 332–344, Mar 2001.
- [7] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," *IEEE Trans. on Circuits and Syst. for Video Tech.*, vol. 11, no. 3, pp. 301–317, Mar 2001.
- [8] P. G. Sherwood and K. Zeger, "Progressive image coding for noisy channels," *IEEE Signal Processing Lett.*, vol. 4, no. 7, pp. 191–198, July 1997.
- [9] Y. Liu, P. Salama, Z. Li, and E. J. Delp, "An enhancement of leaky prediction layered video coding," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 15, no. 11, pp. 1317–1331, Nov. 2005.
- [10] S. Chang, P. Cosman, and L. Milstein, "Performance analysis of n-channel symmetric fec-based multiple description coding for ofdm networks," *IEEE Trans. on Image Processing*, 2010.
- [11] V. K. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Processing Mag.*, vol. 18, no. 5, pp. 74–93, Sept. 2001.
- [12] A. E. Mohr, E. A. Riskin, and R. E. Ladner, "Graceful degradation over packet erasure channels through forward error correction," in *Proc. IEEE Data Compression Conference*, Snowbird, UT, March 1999, pp. 92–101.
- [13] R. Puri, K.-W. Lee, K. Ramchandran, and V. Bharghavan, "An integrated source transcoding and congestion control paradigm for video streaming in the Internet," *IEEE Trans. on Multimedia*, vol. 3, no. 1, pp. 18–32, Mar. 2001.
- [14] D. G. Sachs, R. Anand, and K. Ramchandran, "Wireless image transmission using multiple-description based concatenated codes," in *Proc. SPIE*, vol. 3974, San Jose, CA, Jan. 2000, pp. 300–311.
- [15] Y. S. Chan, P. C. Cosman, and L. B. Milstein, "A cross-layer diversity technique for multi-carrier OFDM multimedia networks," *IEEE Trans. on Image Processing*, vol. 15, no. 4, pp. 833–847, Apr. 2006.
- [16] L. Toni, Y. S. Chan, P. C. Cosman, and L. B. Milstein, "Channel coding for progressive images in a 2-D time-frequency OFDM block with channel estimation errors," *To appear on IEEE Trans. on Image Processing*, vol. 18, no. 11, pp. 2476–2490, 2009.
- [17] L. Rossi, L. Toni, N. Agoulmine, and J.-G. Fontaine, "Source coding optimization for heterogeneous access networks," in *2010 IEEE/IFIP Network Operations and Management Symposium Workshops (NOMS Wksp)*, April 2010, pp. 8–15.
- [18] J. Winters, "Optimum combining in digital mobile radio with cochannel interference," *IEEE J. Select. Areas Commun.*, vol. 2, no. 4, pp. 528–539, Jul 1984.
- [19] P. Örmeci, X. Liu, D. L. Goeckel, and R. D. Wesel, "Adaptive bit-interleaved coded modulation," *IEEE Trans. on Commun.*, vol. 49, no. 9, pp. 1572–1581, Sep. 2001.
- [20] D. L. Goeckel, "Adaptive coding for time-varying channels using outdated fading estimates," *IEEE Trans. on Commun.*, vol. 47, no. 6, pp. 844–855, Jun. 1999.
- [21] S. T. Chung and A. J. Goldsmith, "Degree of Freedom in Adaptive Modulation: A Unified View," *IEEE Trans. on Commun.*, vol. 49, no. 9, pp. 1561–1571, Sep. 2001.
- [22] M. Chiani, M. Z. Win, and A. Zanella, "On optimum combining of  $M$ -ary PSK signals with unequal-power interferers and noise," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 44–47, Jan. 2005.
- [23] A. Giorgetti and M. Chiani, "Influence of fading on the Gaussian approximation for BPSK and QPSK with asynchronous cochannel interference," *IEEE Trans. on Wireless Commun.*, vol. 4, no. 2, pp. 384–389, March 2005.
- [24] J. G. Proakis, *Digital Communications*, 4th ed. New York, NY, 10020: McGraw-Hill, Inc., 2001.
- [25] M. K. Simon and M.-S. Alouini, *Digital Communication over Fading Channels: A Unified Approach to Performance Analysis*, 1st ed. New York, NY, 10158: John Wiley & Sons, Inc., 2000.
- [26] K. Cho and D. Yoon, "On the general BER expression of one- and two-dimensional amplitude modulations," *IEEE Trans. Commun.*, vol. 50, no. 7, pp. 1074–1080, Jul. 2002.
- [27] A. Conti, M. Z. Win, and M. Chiani, "Slow adaptive M-QAM with diversity in fast fading and shadowing," *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 895–905, May 2007.
- [28] International Telecommunications Union, "Methodology for the subjective assessment of the quality of television pictures," Recommendation ITU-R BT.500-7, 1997.
- [29] —, "Subjective video quality assessment methods for multimedia applications," Recommendation ITU-T P.910, 1996.
- [30] "Editor proposed draft text modifications for joint video specification (ITU-T Rec. H.264 — ISO/IEC 14496-10 AVC), Geneva Modification draft 37, Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-E146d37, Geneva, Oct. 2002."
- [31] C. H. Chia and M. S. Beg, "MPEG-4 video transmission over bluetooth links," in *IEEE International Conference on Personal Wireless Communications*, 2002, pp. 280–284.
- [32] P. Ikkurthy and M. A. Labrador, "Characterization of MPEG-4 traffic over IEEE 802.11b wireless lans," in *27th Annual IEEE Conference on Local Computer Networks*, 2002, pp. 421–427.
- [33] G. Kuhne and C. Kuhmich, "Transmitting MPEG-4 video streams over the internet: Problems and solutions," in *ACM Multimedia*, 1999.
- [34] *Universal Mobile Telecommunications System (UMTS) standard, 3GPP TS 25.211; Physical channels and mapping of transport channels onto physical channels (FDD)*. Version 8.5.0 Release 8, Sept. 2009.
- [35] L. Toni, A. Conti, and et al., "Slow adaptive modulation with diversity in the presence of interference," *Submitted to IEEE Transaction on Comm.*