Received February 5, 2019, accepted February 22, 2019, date of publication March 4, 2019, date of current version April 2, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2902798

Speech Quality Assessment in Wireless Communications With MIMO Systems Using a Parametric Model

DEMÓSTENES Z. RODRÍGUEZ[1], (Senior Member, IEEE), RENATA L. ROSA[1], FRANCISCONE L. ALMEIDA, Jr.[1], GABRIEL MITTAG[2,3], AND SEBASTIAN MÖLLER[2,3], (Senior Member, IEEE)

[1]Department of Computer Science, Federal University of Lavras, 37200-000 Lavras, Brazil
[2]Quality and Usability Lab, Technische Universität Berlin, 10623 Berlin, Germany
[3]Deutsches Forschungszentrum für Künstliche Intelligenz, Berlin, Germany

Corresponding author: Renata L. Rosa (renata.rosa@dcc.ufla.br)

This work was supported in part by the CAPES -Von Humboldt Foundations Process under Grant 88881.145495/2017-01, and in part by the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) under Grant 2015/25512-0 and Grant 2015/24496-0.

ABSTRACT Communication service providers use specialized solutions to evaluate the quality of their services. Also, different mechanisms that increase network robustness are incorporated in current communication systems. One of the most accepted techniques to improve transmission performance is the MIMO system. In communication services, voice quality is important to determine the user's quality of experience. Nowadays, different speech quality assessment methods are used, one of them is the parametric method that is used for network planning purposes. The ITU-T Rec. G107.1 is the most accepted model for wide-band communication systems. However, it does not consider the degradations occurring in a wireless network nor the quality improvement, caused by the MIMO systems. Thus, we propose a speech quality model, based on wireless parameters, such as signal-to-noise ratio, Doppler shift, MIMO configurations, and different modulation schemes. Also, real speech signals encoded by 3 operation modes of the AMR-WB codec are used in test scenarios. The resulting speech samples were evaluated by the algorithm described in ITU-T Rec. P.862.2, which scores are used as a reference. With these results, a wireless function, named I_{W-M} that relates the wireless network parameters with speech quality is proposed and inserted into the wide-band E-model algorithm. It is worth noting that the main novelty of the proposed I_{W-M} is the consideration of the quality improvement incorporated by MIMO systems with different antenna array configurations. The performance validation results demonstrated that the inclusion of I_{W-M} values into the R global score determined a reliable model, reaching a Pearson correlation coefficient and a normalized RMSE of 0.976 and 0.144, respectively.

INDEX TERMS Speech quality, wireless communications, MIMO systems, antenna arrays, fading channels, parametric models, E-model.

I. INTRODUCTION

Cellular network operators perform operation and maintenance tasks. However, these actions are not enough to discover all network coverage problems in areas, where the cellular operator provides the service and consequently some users may have unsatisfactory experiences using the phone call service. To avoid these problems, it is important that network engineers have specialized tools to help them in planning tasks. The number of mobile devices is continuously increasing and according to [1] smart-phones will be responsible for 85.8% of the Internet traffic across the world at the end of 2021; then networks need to be prepared to properly manage this volume of traffic. In communication systems there are many factors that can impact the users' quality of experience (QoE), especially in wireless environments [2].

Communication systems are evolving and new technologies are being incorporated to improve the global network performance. These advances make it possible to integrate different services and technologies; for example,

The associate editor coordinating the review of this manuscript and approving it for publication was Hayder Al-Hraishawi.

IP Multimedia Subsystem (IMS) solution is implemented, which permits to establish phone calls based on IP networks and Long Term Evolution (LTE) networks. Another evolution from Public Switched Telephone Network (PSTN) is the bandwidth used, PSTN systems operate in a narrow-band (NB) network that transmits signals considering frequencies between 300 and 3400 Hz. Current networks allow the transmission of wide-band (WB) signals (50 - 7000 Hz) supported by speech codecs for WB signals, such as Adaptive Multi-Rate codec (AMR-WB) [3], that improves the speech quality perception.

Additionally, to decrease the impact of wireless degradations on the transmitted signals, the latest network generations have incorporated solutions that use the principle of spatial diversity in the radio channel, such as the Multiple-Input Multiple-Output (MIMO) system [4]–[6]. As the available spectrum assigned to an operator is physically limited, new radio resource management techniques are required [7]. The use of multiple antennas on transmitter and receptor equipment has the main objective to minimize the bit error rate (BER) in the wireless transmission channel and thereby improving the signal quality at the reception point. Also, modulation schemes have an influence on the transmission performance [8]—to work properly each one needs different channel conditions.

In telephone services, one of the most important factor to determine the user's QoE is the perceived speech quality [9]. The International Telecommunication Union (ITU) describes several methods for speech quality evaluation. These methods are classified into subjective and objective methods. Subjective methods are performed in a controlled environment and following a test procedure [10] that is generally well explained before starting the tests. Thus, a speech sample is scored by several assessors and the average score is computed. This average value is called Mean Opinion Score (MOS).

On the other hand, objective methods use an algorithm to predict the MOS value. Thus, their results can be obtained almost in real-time; then, they may be used in commercial network solutions [11]. According to [12], objective methods are classified into three models, based on the input type used by the algorithm. Algorithms that use speech signals as input are subdivided into intrusive and non-intrusive methods. The most representative non-intrusive method for NB signals is described in ITU-T Recommendation P.563, but it does not present reliable results [13]. The ITU-T Recommendations P.862 [14] that is used in NB signals, P.862.2 used in WB signals [15], and P.863 [16] used in super-WB (SWB) signals are intrusive methods.

The E-Model algorithm is the most known parametric method that is described in ITU-T Recommendation G.107 [17]. This algorithm predicts the overall quality of a phone call conversation. It considers different degradation factors, such as acoustic echoes, electrical coupling on the user's device, transmission delays, packet loss, and specific characteristics of speech codecs. In recent years, different improvements were proposed [18]–[20] to adapt the E-model algorithm for WB and SWB networks.

Most of the current research related to speech quality assessment focuses on wired network impairments. However, in wireless network scenarios, different impairments occur in the transmission channel [21], such as signal power variations, path delays, several paths, channel noise insertion into speech signal [6], [22], among others. Therefore, the effect of those degradations on the speech quality needs to be properly evaluated. Additionally, it is important to evaluate the impact of MIMO systems on the transmission system quality improvement, specifically for network planning purposes.

The main objective of this research is to introduce a wireless function, named I_{W-M}, which is incorporated into the E-model algorithm. In short, the contributions of this work are stated as follows: (i) Different wireless degradations and a MIMO system considering different antenna array configurations are studied, using a test scenario, in which a speech signal is transmitted. (ii) The proposed I_{W-M} considers the effect of MIMO systems under different wireless impairments. (iii) Different AMR-WB operation modes are used in the test scenarios. (iv) A recommended list of MIMO antenna configurations is introduced for a specific wireless network configuration to reach an acceptable speech quality. Thus, this paper shows how to improve the E-model to be used for wireless communication systems, based on wireless parameter values and also considering current technologies, such as MIMO.

The remainder of this paper is organized as follows: Section II presents a theoretical revision that considers the E-model algorithm, the AMR-WB codec, and MIMO systems. A description of wireless communication channel characteristics is presented in Section III. In Section IV, the test methodology and the proposed wireless factor I_{W-M} are described. Section V presents the results. Finally, the conclusions and future proposal works are presented in Section VI.

II. THEORETICAL REVISION

In this section, a general overview of the E-model algorithm for WB signals, the AMR-WB codec and MIMO systems are presented.

A. THE E-MODEL ALGORITHM

As stated before, the E-model algorithm uses several parameters as input, which are called impairment factors, and each of them affects the speech quality score. In short, these impairment factors are related to ambient noise at the end-points, transmission channel noise, network delays and packet loss, the impairment inserted by codecs, among others. The E-model uses the transmission rating scale (R-scale) and it is defined by:

$$R = R_0 - I_s - I_d - I_e + A \tag{1}$$

where, R is the quality score and it ranges from 0 to 100 for NB networks. R_0 represents an ideal scenario without impairments, for NB networks its default value is 93.2.

I_s represents all the impairments which occur simultaneously with the voice. I_d is related to network delays. I_e is known as the equipment impairment factor and represents codec impairments [23]. And A is the advantage factor; it takes, for example, a value of 0 for a wired network and 20 for a satellite connection.

Also, ITU-T Rec. G.107 introduces a function to convert R scores to scores of the 5-point MOS scale for a conversational situation, which is presented in (2).

$$MOS_{CQE} = \begin{cases} 1, & R < 0 \\ 1 + 0.035R + R(R-60)(100-R)(7.10^{-6}), & \\ \quad 0 < R < 100 & \\ 4.5, & R > 100 \end{cases} \tag{2}$$

A function that converts MOS scores to R scores is also presented. In [24], a third order polynomial function to convert MOS scores to R scores is proposed, which is presented in (3):

$$R = \Big\{3.026(MOS)^3 - 25.314(MOS)^2 + 87.060(MOS) - 57.336\Big\} \tag{3}$$

As previously stated, the first E-model implementation focused on NB scenarios. Nowadays, the ITU-T Rec. G.107.1 [25] introduces a new R_0 value for WB signals equal to 129—the impairment factors, however, are the same. Scores from mixed NB/WB scenario ($R_{NB/WB}$) can be obtained based on NB scenarios (R_{NB}) according to:

$$R_{NB/WB} = 1.29(R_{NB}) \tag{4}$$

Also, a methodology to determine I_e using an instrumental model for WB speech codecs is described in ITU-T Rec. P.834.1 [26].

B. AMR-WB CODEC

Traditionally, telecommunication services only provide a limited bandwidth (NB), resulting in the typically muffled audio that is a characteristic of telephone conversations. In mobile communication, one common codec that is used in these networks is the AMR coding standard. To increase the perceived speech quality, a WB transmission was introduced in communication services. The AMR-WB codec uses a sample rate of 16 kHz and extends the frequency range of the AMR-NB codec by around one octave, while it is able to run with the same bitrates. Nowadays, AMR-WB is widely used by network operators (e.g. in UMTS networks) to allow high quality telephone conversations. Similar to AMR-NB, the AMR-WB is based on algebraic code excited linear prediction (ACELP), which uses a vector quantization technique.

In total the AMR-WB codec provides nine different operation modes that are presented in Table 1. A higher bitrate can be useful for background noise conditions or music and lower bitrates provide better quality than NB codecs.

TABLE 1. AMR-WB operation modes and their bit-rates.

AMR-WB Operation Modes	*Bit Rates* (kbps)
Mode 0	6.60
Mode 1	8.85
Mode 2	12.65
Mode 3	14.25
Mode 4	15.85
Mode 5	18.25
Mode 6	19.85
Mode 7	23.05
Mode 8	23.85

3GPP specified the AMR-WB codec in $TS26.190$ [3] and it is also described by the ITU-T as G.722.2 [27].

C. MIMO SYSTEM

In order to reduce the BER at the receiver [28], diversity techniques repeat the information via different transmission channels. One example of diversity is the repeat code technique. These codes repeat the bits of information in the transmitter before being sent by the channel. Thus, in lossy channels, the information can be successfully retrieved as long as at least one of the repeated bits is received correctly. However, the repeat code technique is not efficient because it uses the available bandwidth to send redundant information, reducing the spectral efficiency of the communication system. Therefore, a trade-off must be considered between the robustness of the system and the spectral efficiency.

In general, three kinds of diversity are used: temporal, frequency and spatial. Repetition code is a type of temporal diversity. The frequency diversity works similarly, sending information across different frequency ranges that exhibit independent fading. Spatial diversity is the technique that uses different antennas to transmit and receive the same information. Thus, an important characteristic of MIMO systems is the ability to turn multipath propagation into a benefit for the user. MIMO takes advantage of fading channels and multipath delay spread [29], [30] for increasing transfer rates [31]. When the channel has a favorable condition, the system uses different antennas to transmit different information, increasing the spectral efficiency and consequently the capacity of the system. On the other hand, when the channel has an unfavorable condition for the signal transmission, the transmitter and the receiver use various antennas to transmit and receive redundant bits, increasing the system robustness in wireless networks [28]. MIMO systems require complex algorithms and a greater computational complexity to optimize the management of transmitted information. The significant performance improvement in wireless communication without extra spectrum is largely responsible for the success of MIMO. It is one of the most significant technical breakthroughs in modern communications, because this technology presents itself as an important option for more efficient use of the scarce electromagnetic spectrum [31].

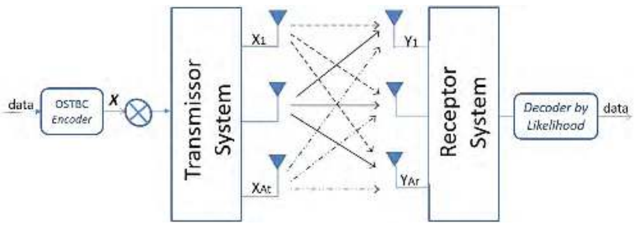

FIGURE 1. Basic diagram of a MIMO system.

A basic block diagram of a MIMO system is presented in Fig. 1 with A_t and A_r being the transmitting and receiving antennas, respectively. Also, an Orthogonal Space-Time Block Code (OSTBC) encoder is considered. The vectors of transmitted and received symbols in discrete time are $x[m] = \{x_1[m], \ldots, x_{A_t}[m]\}^T$ and $y[m] = \{y_1[m], \ldots, y_{A_r}[m]\}^T$, respectively, where x_k represents the symbol sent by k-esim transmitting antenna and y_j represents the symbol received on the j-esim receiving antenna.

In the MIMO system, the transmission channels between each pair of transmitting and receiving antennas is represented by h_{ji}, which builds a matrix $\mathbf{H} = [h_{ji}]$ of dimension $A_r \times A_t$. Considering that the variable $\mathbf{w}$ represents the noise at the receiver, the following relation between the transmitted signal and the received signal is obtained at the discrete time instant m [28]:

$$\mathbf{y}[m] = \sqrt{\frac{\mathrm{SNR}}{A_t}}\mathbf{Hx}[m] + \mathbf{w}[m], \tag{5}$$

where $\mathrm{SNR} = P/N_0$, P is the power of the transmission signal, and N_0 is the power of additive white Gaussian noise (AWGN).

In order for the system to achieve a good performance, wireless channels that compose the MIMO system must have a low correlation. That means that the signals transmitted in each channel suffer different impairment types, such as fading and noise intensity. In communications carried out in urban centers, in which the transmitted signal is affected by several propagation phenomena, the implementation of MIMO systems becomes feasible due to the de-correlation of this environment.

In this work, we use a simulator, in which the signal can be transmitted and received by $A_t(A_r) = [1, 2, 3, 4]$ antennas. When $A_t(A_r) > 1$, the system uses the OSTBC encoder. The OSTBC maps the symbols coming from previous phases in the transmitter in order to generate temporal vectors containing the symbols that will be sent by each of the antennas. To perform this mapping, the OSTBC uses a precoding matrix. For the case $A_t = 2$, this matrix is given by

$$C_2 = \begin{bmatrix} x_1[m] & x_2[m] \\ -x_2^*[m+1] & -x_1^*[m+1] \end{bmatrix}, \tag{6}$$

where * denotes the conjugate transposed from the complex symbol x. Note that at instant m, the antenna 1 transmits $x_1[m]$ and the antenna 2 transmits $x_2[m]$. In the next instant, $m+1$, the antennas transmit $-x_2^*[m+1]$ and $-x_1^*[m+1]$, respectively. Thus, in two instants of time, the receiver will have received the symbols x_1 and x_2. Therefore, there was an insertion of redundancy and no loss of capacity, since two instants were used to receive two different symbols.

At the receiver, the OSTBC combines the received signals by all the antennas to extract the information from the symbols encoded in the transmitter. In the sequence, the decoding scheme named maximum likelihood (ML) is used to obtain the transmitted information.

In this work, a MIMO configuration was used, based on the OSTBC system with a low spatial correlation, in which the signal to be transmitted is repeated in the transmission antennas allowing the system to be more robust against fading degradations. In this work, the P.862.2 algorithm is used to analyze its impact on the speech signal quality.

III. WIRELESS COMMUNICATION CHANNEL CHARACTERISTICS

In wireless networks, there are several degradation factors. In a communication scenario, the transmitter and the receiver are usually in different places with various obstacles between them. Thus, it is probable that there is no line of sight between the sender and the receiver; thus, the signal at the end-point is composed by different signals that arrive from different paths [32]. In the Rayleigh fading channel model there is not a direct path and in the Rician model there is at least one direct path between sender and receiver, but also different phases and amplitudes of the signal arriving at the receiver.

Some physical phenomena that occur in wireless communication are multipath, reflection, diffraction, mirroring and absorption. These phenomena lead to fading in the signal and it can be classified into two types: the large and the small scale [32]–[34]. The first one occurs in longer distances from the transmitter. The power of the received signal decreases as the distance increases due to the path loss or obstacles in the path, such as buildings, vegetation and mountains. Atmospheric phenomena, such as rain, snow, and hail also contribute to the fading; the shadowing factor is also another well-known problem. The fading [35], [36] is an impairment factor in wireless channels.

In a transmission that presents fading, the signal suffers variation in its amplitude and frequency [37]. The Doppler effect and the multipath are characteristics of a fading channel model. The multipath condition generates different amplitudes and phases of the signal at the receptor. The Doppler results in signal frequency variation, and it is related to the relative movement between transmitter and receptor. In [38], a method to calculate the Doppler shift in high speed scenario is presented.

IV. METHODOLOGY

In this section, the methodology that is used to implement the test scenario is explained. Furthermore, the criteria which is used to determine the proposed I_{W-M} factor related to wireless networks is treated.

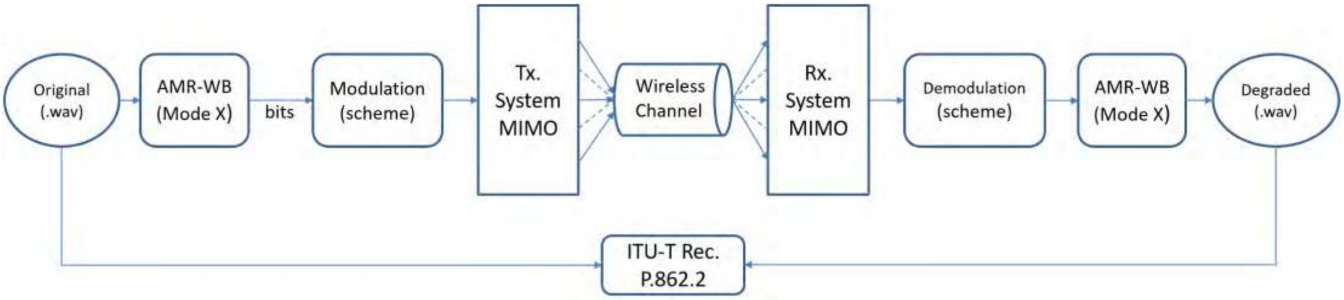

FIGURE 2. Block diagram of the test scenario that considers AMR-WB codec and a MIMO system.

A. TEST SCENARIO

Fig. 2 presents a block diagram of the test scenario implemented. Different impairment conditions were created, using several test scenario configurations. Thus, at the receiver, speech samples with different impairments are obtained.

As a first step, the speech samples were encoded using the AMR-WB speech codec [3], which supports WB signals. As previously stated, the AMR-WB codec has nine operation modes, each one with a different bitrate as indicated in Table 1. It is important to note that the AMR-WB codec is currently used in commercial telecommunication services.

In our test scenario, the following AMR-WB operation modes were used: Mode 0, Mode 4 and Mode 8. The AMR-WB header includes the frame quality indicator (FQI) field, which length is 1 bit. The FQI indicates if the data in the frame contains errors and this information can be used by error concealment algorithms implemented in the decoder [39]. In the test scenarios, the FQI status is determined using the information obtained by the cyclic redundancy check (CRC) code implemented in the AMR-WB using the CRC field composed by 8 bits. The speech samples were taken from ITU-T Rec. P.501 [40] that correspond to full-band signals with a sampling frequency of 48 kHz. They were converted to WB signals using a band-pass filter [41] and down-sampling process to obtain 16 kHz.

The transmission channel corresponds to the Rayleigh fading channel model with additive white Gaussian noise (AWGN) [42], [43]. The parameters configured in this channel model were Signal-to-Noise-Ratio (SNR) and the maximum Doppler frequency shift. Also, a MIMO system was implemented and different intensities of SNR (dB) and maximum Doppler frequency shifts (Hz) were performed in the transmission channel. The implemented MIMO system considers different antenna array configurations, considering the number of antennas in the transmitter and receiver. It uses the Orthogonal Space-Time Block Code (OSTBC) encoder available in Matlab.

In a wireless communication, the modulation scheme is an important factor that is related to the data transmission rate [44]. Also, the channel background noise affects each modulation scheme in different manners, and a different bit-error-rate is obtained. The modulation schemes used in the test scenarios were: BPSK, QPSK and QAM-16. It is important to note that PSK and QAM modulation schemes are used in the sub carriers of OFDM that is used in current communication networks.

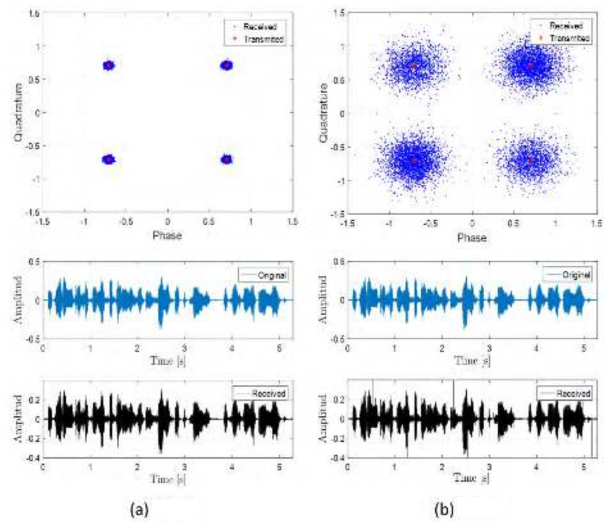

FIGURE 3. Impact of wireless channel degradation on speech signal considering QPSK modulation scheme. (a) NSc-1: transmission without wireless channel degradation and (b) NSc-2 Transmission with a SNR = 8 dB and a maximum doppler shift = 100 Hz.

The commands used for the implementation of the wireless transmission channel used in test scenarios are presented in the Appendix of this document.

In the last step of the test, the algorithm described in ITU-T Rec. P.862.2 was used; thus, each impaired file received a MOS score as shown in Fig. 2. It is important to note that P.862.2 algorithm is able to evaluate WB speech signals and additionally, this algorithm returns as output two MOS scores, the MOS index that corresponds to the scale 0 - 4.5, and the MOS-Listening Quality Objective (MOS-LQO) index that considers the scale in the range 1 - 5. This last output value was used in this work.

Fig. 3 presents two network scenarios, NSc-1 and NSc-2. The first one represents no wireless channel degradation, specifically, SNR = 30 dB and Doppler shift = 0 Hz. Conversely, NSC-2 represents a severe channel degradation, with a SNR = 8 dB and maximum Doppler shift = 100 Hz. In both scenarios, QPSK modulation scheme and AMR-WB operation mode 8 were used. It can be observed how the SNR affect the symbol transmission and the speech signal at the end-point. The MOS score of the speech signal in NSc-1 and NSc-2 were 4.25 and 3.05, respectively.

TABLE 2. Parameter used in the configuration of test scenarios.

Parameter	Values
SNR (dB)	0, 4, 8, 12, 16, 20, 24
Doppler shift (Hz)	0, 50, 100, 150, 200
Modulation	BPSK, QPSK, QAM-16
Number of Tx. Antennas (nTx)	1, 2, 4
Number of Rx. Antennas (mRx)	1, 2, 4

The network parameter values used in the configuration of different test scenarios are presented in Table 2. Note that the maximum SNR value was 24, this is because with this value, the maximum MOS score is reached in all modulation schemes. Also, a maximum Doppler shift value higher than 200 Hz does not add a considerable negative effect. It is worth noting that these configuration parameters were established when the MIMO system was not used, in order to obtain more scenarios with wireless channel impairments and hence be able to evaluate the impact of different MIMO configurations.

In total, 315 different network scenarios were evaluated. The MIMO antenna configuration MIMO(nTx,mRx): $MIMO2x2$ and $MIMO4x4$ were used in the tests, in addition to the Single-Input-Single-Output (SISO) configuration. We restricted the antenna array configuration in order to limit the number of network scenario configurations. It is worth noting that the simulation of each scenario was performed 50 times, and 8 original WB speech files (4 male and 4 female speakers) with a sampling frequency of 16 kHz were used. Thus, 126, 000 impaired speech files with different degradations were evaluated using P.862.2 algorithm.

B. PROPOSED FACTOR FOR WIRELESS CHANNELS

As stated before, the main goal of this work is to propose a factor related to wireless channels, named I_{W-M}, that considers the wireless impairments and the impact of MIMO systems on the speech quality. Thus, I_{W-M} is included in the WB E-model algorithm as presented in (7):

$$R_w = R_0 - I_s - I_d - I_e - I_{W-M} + A \tag{7}$$

The default values of I_s, I_d impairment factors and the A enhanced factor are used, which are introduced in the ITU-T Rec G.107. In this work, the global R-score is denominated as R_w to highlight that it considers the wireless impairment.

In order to determine the impairment factor related to degradations inserted by codec, I_e, it is necessary to find an ideal scenario for each AMR-WB operation modes 0, 4 and 8. These ideal network conditions are found when the maximum speech quality determined by a MOS score is reached considering each modulation scheme. Also, in these initial simulations no MIMO system was used. The parameter values that reached the best speech quality were: an SNR value of 24, a maximum Doppler shift equal to 0 Hz. This procedure was performed for each AMR-WB mode and the SNR and Doppler shift values to reach the maxim speech quality were the same. Due to the MOS scores (P.862.2 results) obtained in the test scenarios correspond to a 5-point MOS quality scale, the relations (3) and (4) were used to convert MOS scores to the R-scale values.

With the quality scores in the R-scale, the I_e is calculated using the following relation:

$$I_e = R_0 - R_w - I_W \tag{8}$$

where, I_e is the codec impairment parameter, R_0 is equal to 129 that is the highest quality score for WB scenarios, R_w is experimentally determined by simulation test results that for this case it is the maximum quality score reached in an ideal test network scenario for each AMR-WB mode operation; and I_W is the wireless impairment factor that is equal to zero in an ideal scenario. Note that I_W does not consider MIMO system.

Then, the proposed wireless factor that includes MIMO system, I_{W-M}, is obtained by:

$$I_{W-M} = R_0 - R_w - I_e - G_{(nTx,mRx)} \tag{9}$$

where R_0, R_w and I_e were previously defined, and the proposed function $G_{(nTx,mRx)}$ corresponds to the speech quality gain when a MIMO system is used. The MIMO system considers different antenna array configurations, with n antennas in the transmitter (nTx) and m antennas in the receiver (mRx). This quality improvement is determined in relation to a transmission system that uses only one antenna in the transmitter and receiver. It is important to note that to calculate I_e, the $G_{(nTx,mRx)}$ output was set to zero because the MIMO system was not considered.

In general, the $G_{(nTx,mRx)}$ function depends on the modulation scheme, the AMR-WB operation mode and the SNR; and its outputs are numerical values presented in tables. The quality improvement needs to be quantified using the R-scale. In this research, different tables are presented that contain a specific quality gain for each modulation and AMR-WB mode operation. It is worth noting that because these quality scores were obtained using the ITU-T Rec. P.862.2, they are expressed in the 5-point MOS scale; and thus relations (4) and (3) need to be applied before.

Preliminary test results, with no MIMO system implemented, showed a high correlation between MOS scores and SNR or Doppler shift values, each of them with a different effect. Also, the modulation techniques showed to have a considerable impact on the speech quality. For this reason, a different function for each modulation scheme and AMR-WB mode was considered. To determine I_{W-M}, different functions were tested. The best accuracy was obtained by a second-order polynomial function to model the SNR impact and a logarithmic function to model the Doppler shift effect.

Note that the $G_{(nTx,mRx)}$ values represent a quality enhancement and therefore it reduces the wireless impairment score contained in I_{W-M}, which is defined by:

$$I_{W-M} = \begin{cases} a.S^2 + b.S + c + d.Ln(D) - G_{(nTx,mRx)}, & S < Th \\ 0, & S \geq Th \end{cases} \tag{10}$$

where, S represents the SNR (dB) and D represents the maximum Doppler shift (Hz). The coefficients of the 2^{nd} order polynomial function are a, b and c; and d is the weigh degradation related to the Doppler shift (Hz); Th is a threshold value that depends on the modulation scheme used; and $G_{(nTx,mRx)}$ is the quality gain corresponding to a specific antenna configuration of the MIMO system. It is important to note that a, b, c, d values and the $G_{(nTx,mRx)}$ outputs depend on the modulation scheme and the AMR-WB mode operation. Firstly, a, b, c, d values are determined without MIMO; later, $G_{(nTx,mRx)}$ outputs are obtained.

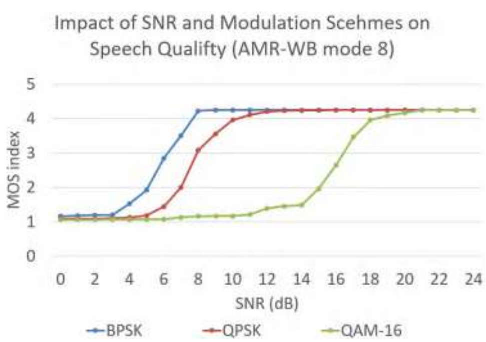

FIGURE 4. Impact of SNR (dB) on speech signal quality (MOS) for BPSK, QPSK and QAM-16 modulation schemes.

V. RESULTS

Firstly, the effect of different SNR (dB) values on the speech quality, considering BPSK, QPSK and QAM-16 modulation schemes are evaluated. Fig. 4 shows the results obtained for the AMR-WB with operation mode 8. In this scenario, a SISO configuration was evaluated and for a better visualization, 25 SNR values were used. As can be observed in Fig. 4, SNR values (dB) have a high negative impact on speech quality. Also, it is important to note that the speech quality starts to improve at different SNR values, which depend on the modulation scheme used. Because of this behavior, each scheme needs to have a different model. As expected, the BPSK modulation scheme is more robust in lower SNR values.

The impact of different maximum Doppler shift (Hz) values is presented in Fig. 5, in which the speech quality degradation is represented by the absolute value of MOS index. The BPSK, QPSK and QAM-16 modulation schemes and AMR-WB with operation mode 8 were used in the test scenario. From Fig. 4 and Fig. 5, it is observed that Doppler

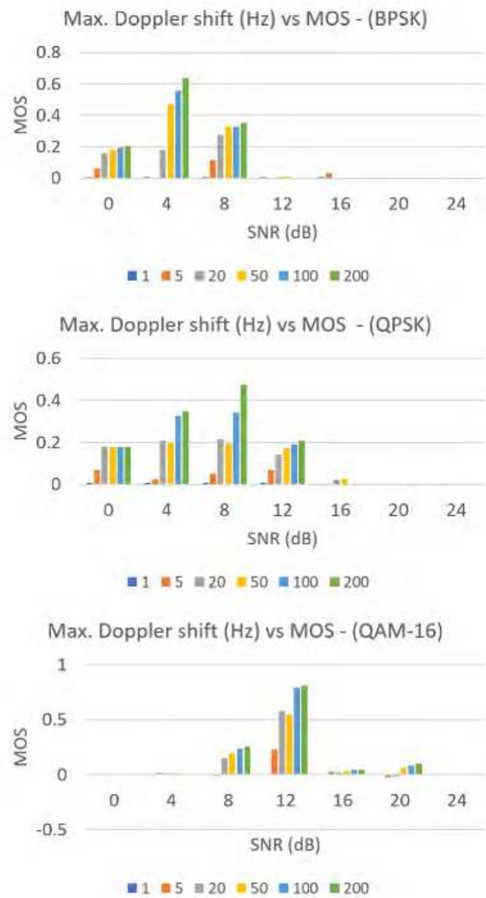

FIGURE 5. Impact of maximum doppler shift (Hz) on speech signal quality (absolute value of MOS) for BPSK, QPSK and QAM-16 modulation schemes.

TABLE 3. Impairment factors (I_e) of AMR-WB codec (Modes 0, 4 and 8) found in ideal network scenarios.

AMR-WB - Mode 0		AMR-WB - Mode 4		AMR-WB - Mode 8	
Max. MOS	I_e	*Max. MOS*	I_e	*Max. MOS*	I_e
3.43	44.3	4.16	19.5	4.25	15.6

Shift effect depends on SNR values, and it has a lower impact on speech quality than SNR. In scenarios in which the speech quality reached the maximum value, the Doppler shift effect is almost negligible.

As previously stated, in order to determine the I_e impairment factor for each AMR-WB mode, a network scenario without the MIMO system that reaches a maximum MOS index value was found. For this test scenario, the network parameters values were: SNR = 24 dB, maximum Doppler shift = 0 Hz; and the three modulation schemes were considered. The results are presented in Table 3.

Once I_e was calculated for each AMR-WB mode, the coefficients introduced in (10) need to be determined. The experimental results obtained to calculate I_{W-M} for each modulation scheme and AMR-WB mode were used to determine the variables, a, b, c and d. It should be noted that the proposed functions reached the best correlation performance

TABLE 4. Coefficient values used to determine I_{W-M} for AMR-WB mode 0.

Modulation Scheme	a (0)	b (0)	c (0)	d (0)	Th
BPSK	0.4601	-10.2090	49.5184	1.6097	12
QPSK	0.0559	-6.3104	61.3105	1.5987	12
QAM-16	-0.2809	0.5003	65.2103	1.9019	16

TABLE 5. Coefficient values used to determine I_{W-M} for AMR-WB mode 4.

Modulation Scheme	a (4)	b (4)	c (4)	d (4)	Th
BPSK	0.6209	-14.6439	74.2874	1.7389	12
QPSK	-0.0312	-6.0019	76.4458	1.1672	12
QAM-16	-0.2539	0.8362	88.2678	1.9016	16

TABLE 6. Coefficient values used to determine I_{W-M} for AMR-WB mode 8.

Modulation Scheme	a (8)	b (8)	c (8)	d (8)	Th
BPSK	0.8237	-18.5675	100.8899	1.8513	12
QPSK	-0.1517	-5.7482	92.4085	0.8507	12
QAM-16	-0.3832	1.1621	101.2768	1.9025	16

in relation to other functions. Table 4, Table 5 and Table 6 present these coefficient values for modes 0, 4 and 8, respectively.

Fig. 6 shows the relation between I_{W-M} and the Doppler shift parameters, considering a test scenario with BPSK modulation scheme and SNR values of 0dB, 4dB, 8dB, and 12dB. It can be observed that there is a relation between channel condition and speech quality, for instance, for an SNR equal to 12dB the Doppler effect is minimum and its behavior cannot be defined, the I_{W-M} values are lower than 0.8. Conversely, when SNR is equal to 0, 4, and 8 there is a direct relation between Doppler shift and I_{W-M} values; also the I_{W-M} are higher, for example, for $SNR = 0$, I_{W-M} reached 87.

TABLE 7. Performance validation of I_{W-M} model in wireless scenarios without MIMO system and considering P.862.2 scores as reference values.

Codec Mode	BPSK PCC / NRMSE	QPSK PCC / NRMSE	QAM-16 PCC / NRMSE
8	0.995 / 0.108	0.994 / 0.079	0.977 / 0.117
4	0.980 / 0.115	0.984 / 0.095	0.981 / 0.124
0	0.988 / 0.086	0.986 / 0.082	0.990 / 0.057

For performance assessment, a comparison between R_w scores using (7) that includes I_{W-M} defined in (10), and the P.862.2 results converted to the R-scale, is performed. In this case, $G_{(nTx,mRx)}$ is equal to zero, that means only wireless degradations (SNR and Doppler shift) are considered. The Pearson correlation coefficient (PCC) and normalized root-mean-squared error (NRMSE) values are presented in Table 7.

After having evaluated the impairments inserted by wireless channel, a MIMO system was configured, and the antenna array configurations $2x2$ and $4x4$ were considered. Additionally, the SISO configuration is implemented and it is used as reference to determine the quality gain of each MIMO configuration. The impact of the MIMO system on the speech quality, considering BPSK, QPSK, QAM-16 modulation schemes, and AMR-WB modes 0, 4, and 8 is presented in Fig. 7.

Based on the results presented in Fig. 7, the values of $G_{(nTx,mRx)}$ are calculated. Thus, the difference between MOS scores reached in each MIMO configuration (2x2 or 4x4) and SISO configuration, for each SNR value, is computed. Table 8, Table 9 and Table 10 present the $G_{(nTx,mRx)}$ values, for AMR-WB operation modes 0, 4, and 8, respectively.

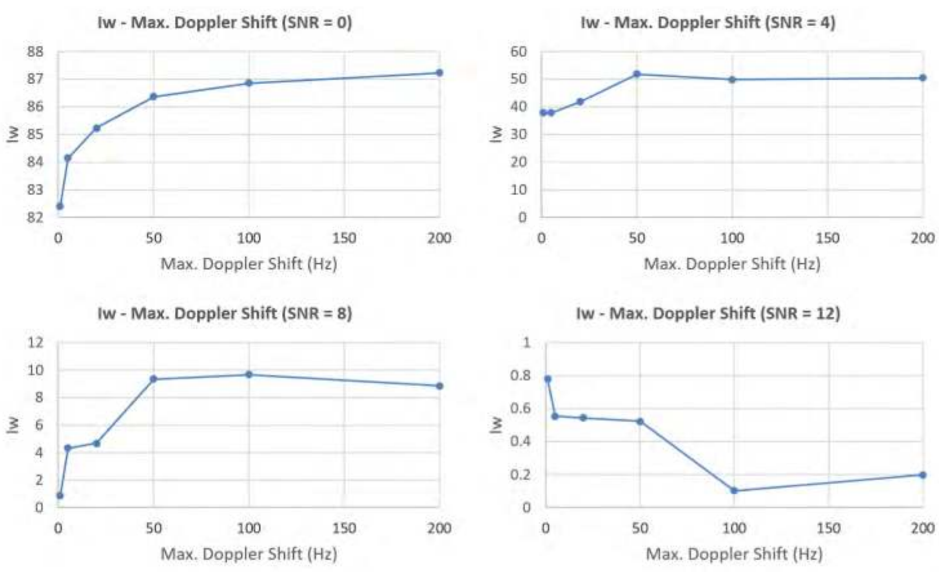

FIGURE 6. Impact of doppler Shift on the I_{W-M} parameter considering a test scenario with BPSK and SNR values of 0dB, 4dB, 8dB and 12dB.

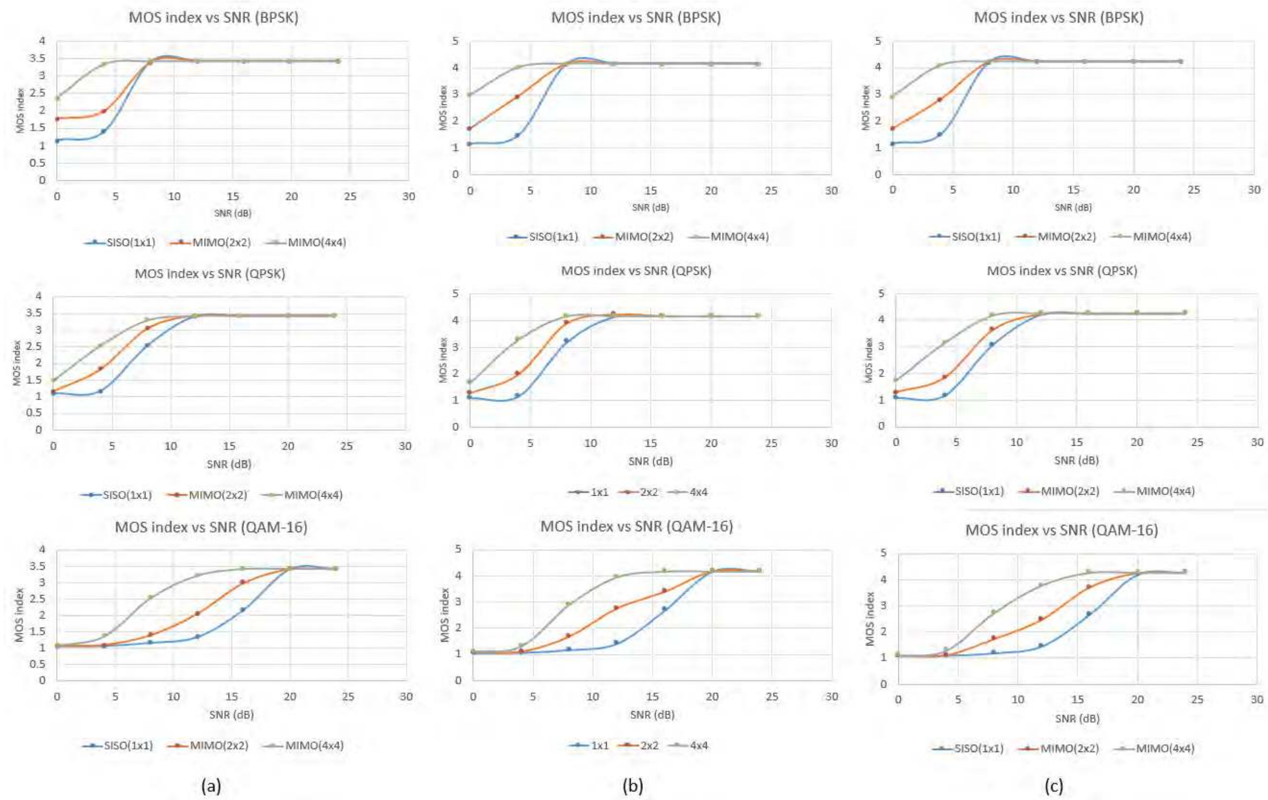

FIGURE 7. Relation between MOS index and SNR (dB) in MIMO transmission, considering (a) AMR-WB mode 0, (b) AMR-WB mode 4, and (c) AMR-WB mode 8.

TABLE 8. Values of MIMO configurations gains that define $G_{(nTx,mRx)}$ considering AMR-WB mode 0.

SNR	BPSK		QPSK		QAM-16	
(dB)	G(2x2)	G(4x4)	G(2x2)	G(4x4)	G(2x2)	G(4x4)
0	0.63	1.21	0.07	0.36	0.02	0.04
4	0.57	**1.97**	0.65	1.37	0.04	0.20
8	0.00	0.01	0.52	0.79	0.22	1.37
12	0.00	0.00	0.03	0.02	0.72	**1.91**
16	0.00	0.00	0.00	0.01	0.88	1.26
20	0.00	0.00	0.00	0.00	0.00	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00

TABLE 9. Values of MIMO configuration gains that define $G_{(nTx,mRx)}$ considering AMR-WB mode 4.

SNR	BPSK		QPSK		QAM-16	
(dB)	G(2x2)	G(4x4)	G(2x2)	G(4x4)	G(2x2)	G(4x4)
0	0.53	**1.87**	0.18	0.55	0.02	0.03
4	1.42	**2.55**	0.79	**2.13**	0.02	0.21
8	0.01	0.01	0.70	0.98	0.49	1.73
12	0.00	0.00	0.08	0.02	1.34	**2.54**
16	0.00	0.00	0.00	0.00	0.70	1.48
20	0.00	0.00	0.00	0.00	0.00	0.00
24	0.00	0.00	0.00	0.00	0.00	0.00

As previously mentioned, the values of Table 8, Table 9, and Table 10 use the 5-point MOS scale; these values are converted to R-scale using (3). From theses tables, it can be observed that in some cases the speech quality improvement reached values closer to 2 that is a great improvement

TABLE 10. Values of MIMO configuration gains that define $G_{(nTx,mRx)}$ considering AMR-WB mode 8.

SNR	BPSK		QPSK		QAM-16	
(dB)	G(2x2)	G(4x4)	G(2x2)	G(4x4)	G(2x2)	G(4x4)
0	0.56	1.72	0.11	0.46	0.02	0.05
4	1.28	**2.61**	0.57	**1.91**	0.02	0.17
8	0.01	0.01	0.57	1.08	0.59	1.53
12	0.00	0.00	0.03	0.05	1.12	**2.38**
16	0.00	0.00	0.00	0.00	1.06	1.64
20	0.00	0.00	0.00	0.00	0.09	0.12
24	0.00	0.00	0.00	0.00	0.00	0.00

considering a 5-point MOS scale. Also, it is observed that in some network conditions, the MIMO implementation has not improved the speech quality compared to the SISO scenario. We used the 5-point MOS scale because it is the quality scale mostly used in both subjective and objective speech assessment methods, and thus, the impact of the MIMO system can be better interpreted.

A comparison between R_w scores, including wireless degradations and $G_{(nTx,mRx)}$, and P.862.2 results is performed. The PCC and NRMSE values are presented in Table 11.

As can be observed, the performance results presented in Table 11 and Table 7 are similar. These results demonstrate that the inclusion of $G_{(nTx,mRx)}$ as part of I_{W-M} helps to evaluate the impact of MIMO systems in different wireless network conditions with high accuracy.

TABLE 11. Performance validation of I_{W-M} model in wireless scenarios considering a MIMO system and P.862.2 scores as reference values.

Codec Mode	BPSK PCC / NRMSE	QPSK PCC / NRMSE	QAM-16 PCC / NRMSE
8	0.993 / 0.135	0.984 / 0.111	0.975 / 0.139
4	0.987 / 0.146	0.988 / 0.107	0.976 / 0.135
0	0.980 / 0.149	0.978 / 0.104	0.983 / 0.096

Finally, the performance of R_w, based on the proposed I_{W-M} was evaluated using 4 additional speech files (2 male and 2 female speakers). To this end, the P.862.2 algorithm results were also considered as reference. In total, 21, 600 speech samples were tested; considering the following parameter: SNR = [1, 6, 13] dB, maximum Doppler shift = [15, 80] Hz for each one of the 3 modulation schemes (BPSK, QPSK and QAM-16) and the 3 AMR-WB operation modes (0, 4 and 8). Each test scenario was repeated 50 times and the same MIMO antenna array configuration (MIMO2x2 and MIMO4x4) and SISO configuration were used. Because the SNR values are different to those used to model $G_{nTx,mRx}$, a linear regression is applied to the corresponding scenarios presented in Tables 8, 9 and 10 to obtain the speech quality gain values. The experimental results show that proposed model reached a global PCC and NRMSE of 0.976 and 0.144, respectively.

It can be concluded that the performance assessment results—reached by the proposed method—are reliable. It is worth noting that to the best of our knowledge, in the current literature there is not a parametric model that evaluates the impact of MIMO systems on the speech quality. For this reason, the performance assessment was done only using a standardized speech-based model, namely the P.862.2 algorithm. Another relevant characteristic of the proposed I_{W-M} factor is that other impairment factors of the WB E-model algorithm are not affected.

VI. CONCLUSION

Experimental test results demonstrated that there is a high correlation between both wireless channel degradations, improvements introduced by the MIMO system, and the speech quality scores. To obtain those results, an useful test scenario was implemented, in which a real speech signal is coded and transmitted at bit-level, passed through different stages of a wireless communication system and finally the quality of the decoded speech signal is evaluated. The proposed wireless factor I_{W-M} takes into account a fading channel model with different SNR and maximum Doppler shift values, different modulation schemes, and also two MIMO system configurations. The results obtained by the modified WB E-model algorithm, which global score R_w includes the I_{W-M} factor are reliable because they are highly correlated with the P.862.2 results, which are used as ground-truth. In the performance validation results, which are carried out with different test materials and different impairment conditions, the proposed R_w reached a PCC and NRMSE of 0.976 and 0.144, respectively.

Currently, the most advanced communication systems adopted MIMO techniques for their intrinsic advantages. Also, with the growth of mobile devices, most of the voice communication is performed in wireless networks; therefore, the speech signal can suffer different degradations to those originated in wired networks. In this arena, communication service providers need to have specialized tools for network planning purposes available. Thus, it is convenient to incorporate parameters related to wireless communication systems to the WB E-model algorithm. In this sense, this work presents a solution to improve the WB E-model, because it considers common wireless degradations and the MIMO technique to improve the transmission quality. In a future work, more parameters will be tested and incorporated into the E-model algorithm, such as different techniques adopted in current wireless communication systems; also, different approaches that perform minimum changes in the E-model algorithm will be studied.

TABLE 12. Commands used to implement the wireless transmission channel of network simulation scenarios.

Line	Commands
1	**% Audio file**
2	info = audioinfo('audio.wav');
3	[y,Fs] = audioread('audio.wav');
4	t = 0:(1 / Fs):info.Duration;
5	**% Modulation scheme**
6	if modtype == BPSK
7	hModulator = comm.BPSKModulator;
8	y_{mod} = step(hModulator, y_{coded});
9	**% Channel (MIMO system)**
10	hEnc = comm.OSTBCEncoder('NumTransmitAntennas', ...
11	NumTxAnt,'SymbolRate');
12	txSig = step(hEnc,y_{mod});
13	hChan = comm.MIMOChannel('SampleRate',S-Rate, ...
14	'MaximumDopplerShift',MaxDopplerShift, ...
15	'SpatialCorrel',low,'NumTransmitAnt',NumTxAnt, ...
16	'NumReceiveAntennas',NumRxAnt);
17	hAWGN = comm.AWGNChannel('NoiseMethod',...
18	'Signal to noise ratio (SNR)',...
19	'SNR',SNR,'SignalPower',1);
20	[chanOut,pathGains] = step(hChan,txSig);
21	rxSig = step(hAWGN,chanOut);
22	**% Demodulation scheme**
23	hDemod = comm.BPSKDemodulator;
24	$sig_{demod-bin}$ = step(hDemod, rxSig);
25	hError = comm.ErrorRate;
26	errorStats = step(hError, y_{coded}, $sig_{demod-bin}$);

APPENDIX
PROGRAMMING COMMANDS FOR THE TRANSMISSION CHANNEL IMPLEMENTATION

Table 12 presents the Matlab commands used to implement the main blocks of the wireless transmission channel used in different simulation test scenarios.

REFERENCES

[1] *Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016 2021*, Cisco, San Jose, CA, USA, Jun. 2017.

[2] D. Z. Rodríguez, R. L. Rosa, R. D. Nunes, and E. T. Affonso, "Assessment of quality-of-experience in telecommunication services," *Int. J. Digit. Inf. Wireless Commun.*, vol. 6, no. 4, pp. 241–259, Oct. 2016.

[3] *Speech Codec Speech Processing Functions; Adaptive Multi-Rate—Wideband (AMR-WB) Speech Codec; Transcoding Functions*, document TS 26.190, 3GPP, Jun. 2018.

[4] G. R. Patil and V. K. Kokate, "Reduced complexity tree search algorithms for MIMO decoding," *J. Commun. Softw. Syst.*, vol. 10, no. 4, pp. 230–236, Dec. 2014.

[5] F. Benykhlef and N. Boukli-Hacene, "Ebg structures for reduction of mutual coupling in patch antennas arrays," *J. Commun. Softw. Syst.*, vol. 13, no. 1, pp. 9–14, Mar. 2017.

[6] D. Z. Rodríguez, G. Pívaro, R. L. Rosa, G. Mittag, and S. Möller, "Quantifying the quality improvement of MIMO transmission systems in VoIP communication," in *Proc. 26th Int. Conf. Softw., Telecommun. Comput. Netw. (SoftCOM)*, Sep. 2018, pp. 1–5.

[7] H. Fall, O. Zytoune, and M. Yahyai, "Theory of algorithm suitability on managing radio resources in next generation mobile networks," *J. Commun. Softw. Syst.*, vol. 14, no. 2, pp. 180–188, Jun. 2018.

[8] M. L. Ammari and F. Gagnon, "Impact of uncertain channel estimation and outdated feedback on the adaptive M-PSK modulation," *J. Commun. Softw. Syst.*, vol. 10, no. 1, pp. 41–47, Mar. 2014.

[9] D. Rodríguez, G. Pivaro, and J. Sousa, "Apparatus and method for evaluating voice quality in a mobile network," U.S. Patent 9 078 143 B2, Jul. 10, 2015, pp. 1–15.

[10] *Methods for Subjective Determination of Transmission Quality*, document ITU-T Rec. P.800, Jun. 1996. [Online]. Available: http://www.itu.int/rec/T-REC-P.800/en

[11] C. Sloan, N. Harte, D. Kelly, A. C. Kokaram, and A. Hines, "Objective assessment of perceptual audio quality using visqolaudio," *IEEE Trans. Broadcast.*, vol. 63, no. 4, pp. 693–705, Dec. 2017.

[12] S. Möller, W. Y. Chan, N. Côté, T. H. Falk, A. Raake, and M. Wältermann, "Speech quality estimation: Models and trends," *IEEE Signal Process. Mag.*, vol. 28, no. 6, pp. 18–28, Nov. 2011.

[13] J. Polacký and P. Po ta, "An analysis of the impact of packet loss, codecs and type of voice on internal parameters of P.563 model," in *Proc. 10th Int. Conf. Digit. Technol.*, Zilina, Slovakia, Jul. 2014, pp. 281–284.

[14] document ITU-T Rec.-P.862, 2007, Accessed: Feb. 1, 2017. [Online]. Available: http://www.itu.int/rec/T-REC-P.862/en

[15] *Wideband Extension to Recommendation P.862 for the Assessment of Wideband Telephone Networks and Speech Codecs*, document ITU-T Rec. P.862.2, Apr. 2007. [Online]. Available: https://www.itu.int/rec/T-REC-P.862.2-200711-I/en

[16] *Perceptual Objective Listening Quality Assessment (POLQA)*, document ITU-T Rec. P.863, Sep. 2014. [Online]. Available: http://www.itu.int/rec/T-REC-P.863-201409-I

[17] *The E-Model: A Computational Model for Use in Transmission Planning*, document ITU-T Rec. G.107, Jun. 2014. [Online]. Available: https://www.itu.int/rec/T-REC-G.107

[18] S. Möller, N. Côté, V. Gautier-Turbin, N. Kitawaki, and A. Takahashi, "Instrumental estimation of E-model parameters for wideband speech codecs," *EURASIP J. Audio, Speech, Music Process.*, vol. 2010, Nov. 2010, Art. no. 782731.

[19] J. Abichandani *et al.*, "A comparative study of voice quality and coverage for voice over long term evolution calls using different codec mode-sets," *IEEE Access*, vol. 5, pp. 10315–10322, 2017.

[20] G. Mittag, S. Möller, V. Barriac, and S. Ragor, "Quantifying quality degradation of the EVS super-wideband speech codec," in *Proc. 10th Int. Conf. Qual. Multimedia Exper. (QoMEX)*, Sardinia, Italy, May/Jun. 2018, pp. 1–6.

[21] E. T. Affonso, D. Z. Rodríguez, R. L. Rosa, T. Andrade, and G. Bressan, "Voice quality assessment in mobile devices considering different fading models," in *Proc. IEEE Int. Symp. Consumer Electron. (ISCE)*, Sep. 2016, pp. 21–22.

[22] T. Al-Sharabati, "Effect of pseudo random noise (PRN) spreading sequence generation of 3GPP usersâ Ź codes on GPS operation in mobile handset," *J. Commun. Softw. Syst.*, vol. 12, no. 4, pp. 182–189, Dec. 2018.

[23] E. T. Affonso, R. L. Rosa, and D. Z. Rodríguez, "Speech quality assessment over lossy transmission channels using deep belief networks," *IEEE Signal Process. Lett.*, vol. 25, no. 1, pp. 70–74, Jan. 2018.

[24] H. Assem, D. Malone, J. Dunne, and P. O'Sullivan, "Monitoring VoIP call quality using improved simplified E-model," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Jan. 2013, pp. 927–931.

[25] *Wideband E-Model*, document ITU-T Rec. G.107.1, Jun. 2015. [Online]. Available: https://www.itu.int/rec/T-REC-G.107.1/en

[26] *Extension of the Methodology for the Derivation of Equipment Impairment Factors from Instrumental Models for Wideband Speech Codecs*, document ITU-T Rec. P.834.1, Jun. 2015. [Online]. Available: https://www.itu.int/rec/T-REC-P.834.1/en

[27] *Wideband Coding of Speech at Around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB)*, document ITU-T Rec. G.722.2, Jul. 2003. [Online]. Available: https://www.itu.int/rec/T-REC-G.722.2-200307-I/en

[28] D. Tse and P. Viswanath, *Fundamentals Wireless Communication*. New York, NY, USA: Cambridge Univ. Press, 2005.

[29] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, no. 3, pp. 311–335, Mar. 1998.

[30] G. C. Raleigh and J. M. Cioffi, "Spatio-temporal coding for wireless communication," *IEEE Trans. Commun.*, vol. 46, no. 3, pp. 357–366, Mar. 1998.

[31] D. Gesbert, M. Shafi, D.-S. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.

[32] S. Haykin and M. Moher, *Modern Wireless Communications*. Upper Saddle River, NJ, USA: Pearson Prentice-Hall, 2005.

[33] V. J. Arokiamary, *Cellular and Mobile Communications*, 2nd ed. Chennai, Tamil Nadu: Technical, 2009.

[34] J. D. Gibson, *The Mobile Communications Handbook*, 2nd ed. Boca Raton, FL, USA: CRC Press, 1999.

[35] H. Lee, S. Byeon, B. Kim, K. B. Lee, and S. Choi, "Enhancing Voice over WLAN via Rate Adaptation and Retry Scheduling," *IEEE Trans. Mobile Comput.*, vol. 13, no. 12, pp. 2791–2805, Dec. 2014.

[36] K. S. Shanmugan, "Simulation-based estimate of QoS for voice traffic over WCDMA radio links," in *Proc. 5th Int. Conf. Wireless Commun., Netw. Mobile Comput.*, Beijing, China, Sep. 2009, pp. 1–4.

[37] J. D. Gibson, *The Communications Handbook*. Boca Raton, FL, USA: CRC Press, 2002.

[38] T. Liu, X. Ma, R. Zhao, H. Dong, and L. Jia, "Doppler shift estimation for high-speed railway scenario," in *Proc. IEEE 83rd Veh. Technol. Conf. (VTC Spring)*, May 2016, pp. 1–5.

[39] *Mandatory Speech Codec Speech Processing Functions; Adaptive Multi-Rate (AMR) Speech Codec Frame Structure*, document TS 26.101, 3GPP, Jun. 2018.

[40] *Test Signals for Use in Telephonometry*, document ITU-T Rec. P.501, Mar. 2017. [Online]. Available: https://www.itu.int/rec/T-REC-P.501-201703-I/en

[41] *Transmission Characteristics for Wideband Digital Loudspeaking and Hands-Free Telephony Terminals*, document ITU-T Rec. P.341, Mar. 2012. [Online]. Available: https://www.itu.int/rec/T-REC-P.341-201103-I/en

[42] M. Abo-Zahhad, M. Farrag, and A. Ali, "A fast accurate method for calculating symbol error probabilities for AWGN and Rayleigh fading channels," in *Proc. 33rd Nat. Radio Sci. Conf. (NRSC)*, Feb. 2016, pp. 241–248.

[43] F. Alberge, "Deep learning constellation design for the AWGN channel with additive radar interference," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1413–1423, Feb. 2019.

[44] M. N. Rajesh, B. K. Shrisha, N. Rao, and H. V. Kumaraswamy, "An analysis of BER comparison of various digital modulation schemes used for adaptive modulation," in *Proc. IEEE Int. Conf. Recent Trends Electron., Inf. Commun. Technol. (RTEICT)*, May 2016, pp. 241–245.

DEMÓSTENES Z. RODRÍGUEZ (M'12–SM'15) received the B.S. degree in electronic engineering from the Pontifical Catholic University of Peru, and the M.S. and Ph.D. degrees from the University of São Paulo, in 2009 and 2013, respectively. He has a solid knowledge in telecommunication systems and computer science based on 15 years of Professional experience in major companies. He is currently an Adjunct Professor with the Department of Computer Science, Federal University of Lavras, Brazil. His research interests include QoS and QoE in multimedia services, and architect solutions in telecommunication systems.

RENATA L. ROSA received the M.S. degree from the University of São Paulo, in 2009, and the Ph.D. degree from the Polytechnic School, University of São Paulo, in 2015. She is currently an Adjunct Professor with the Department of Computer Science, Federal University of Lavras, Brazil. Her current research interests include computer networks, quality of experience of multimedia service, social networks, and recommendation systems.

FRANCISCONE L. ALMEIDA, Jr. received the B.S. degree in computer science from the Federal University of Lavras. He was a Technician in computer systems with the Federal Center of Technological Education, Brazil, where started learning about computing and technology. His current research interests include open source systems, computer architecture, multimedia services, telecommunication systems, and signal processing.

GABRIEL MITTAG received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from the Technische Universität Berlin (TU Berlin), Germany, where he is currently pursuing the Ph.D. degree in quality of experience of speech communication services. During his master's degree, he spend two semesters at RMIT University, Melbourne, Australia, where he focused primarily on biomedical and speech signal processing. Since 2016, he has been a Research Assistant with the Quality and Usability Labs, TU Berlin. His current research interests include psychoacoustics, quality evaluation, signal processing, and diagnostic speech quality prediction.

SEBASTIAN MÖLLER received the degree in electrical engineering from the University of Bochum, Bochum, Germany; the University of Orleans, Orléans cedex, France; and the University of Bologna, Bologna, Italy, and the Doctor of Engineering degree in 1999. He received the Venia Legendi with a book on the quality of telephone-based spoken dialog systems in 2004. In 2005, he joined Deutsche Telekom Laboratories, Technische Universität Berlin (TU Berlin), Berlin, Germany, where he was appointed as a Professor for Quality and Usability in 2007. His current interests include speech signal processing, speech technology, and quality and usability evaluation.