Novel Two-Bit Adaptive Delta Modulation Algorithms

Zoran PERIC[1], Bojan DENIC[1*], Vladimir DESPOTOVIC[2],

[1]*University of Niš, Faculty of Electronic Engineering,*
 Aleksandra Medvedeva 14, 18000 Niš, Serbia
[2]*University of Belgrade, Technical Faculty in Bor, Vojske Jugoslavije 12, 19210, Bor, Serbia*
e-mail: zoran.peric@elfak.ni.ac.rs, bojan.denic@elfak.rs, vdespotovic@tfbor.bg.ac.rs

Abstract. This paper introduces two novel algorithms for the 2-bit adaptive delta modulation, namely 2-bit hybrid adaptive delta modulation and 2-bit optimal adaptive delta modulation. In 2-bit hybrid adaptive delta modulation, the adaptation is performed both at the frame level and the sample level, where the estimated variance is used to determine the initial quantization step size. In the latter algorithm, the estimated variance is used to scale the quantizer codebook optimally designed assuming Laplace distribution of the input signal. The algorithms are tested using speech signal and compared to constant factor delta modulation, continuously variable slope delta modulation and instantaneously adaptive 2-bit delta modulation, showing that the proposed algorithms offer higher performance and significantly wider dynamic range.

Key words: delta modulation, predictive coding, speech coding, signal to noise ratio, Laplacian source.

1. Introduction

Delta modulation (DM) is a simple analog-to-digital conversion technique widely used in coding (and compression) of correlated signals, including speech, audio, image, etc. It can be observed as low-complexity Differential Pulse Code Modulation (DPCM) (Jayant and Noll, 1984; Hanzo *et al.*, 2007; Uddin *et al.*, 2016; Gibson, 2017; Sarade, 2017), where the basic configuration includes one-bit quantization and the first-order prediction (Zrilic, 2005). Given its low-complexity and solid performance, DM is a good candidate for real time implementations. Both DPCM and DM belong to a group of predictive coding algorithms (Gibson, 2016, 2017), which are often used in adaptive signal processing, cognitive signal processing, speech enhancement (Hucha Arce *et al.*, 2017) and artificial intelligence (Hastie *et al.*, 2008).

DM makes a comparison between the current signal sample and its previous value, and outputs a single bit indicating the sign of the difference between these two samples. If the difference is positive, the approximated signal is increased by step $+\Delta$ (bit 1), otherwise, if it is negative, the approximated signal is decreased by step $-\Delta$ (bit -1). Since

---
[*]Corresponding author.

the step size is always constant, the maximum or minimum slopes of the approximated signal tend to occur along the straight lines. Therefore, DM with the fixed step size is also known as Linear Delta Modulation (LDM). The main advantages of LDM are the simple implementation of encoder and decoder and the low bit-rate. However, LDM also suffers from several limitations, such as slope overload and granular noise. Slope overload occurs when the step size Δ is not large enough; hence the approximated signal cannot follow the steep changes in the input signal. On the other hand, the granular noise occurs when the step size Δ is too large for small variations in the input signal.

To overcome the drawbacks of LDM and improve its performance, different modifications are proposed, such as e.g. Adaptive Delta Modulation (ADM). In ADM the step size is not constant, but updated according to the specific rule related to the changes in the input signal. Practical applications of ADM algorithms require appropriate restriction on minimum and maximum step size, which respectively controls the amount of idle channel noise and slope overload distortion (Jayant and Noll, 1984). The examples of ADM are Constant Factor Delta Modulation (CFDM) and Continuously Variable Slope Delta Modulation (CVSDM). CFDM uses one or two bit memory to determine the appropriate step size at each sampling instant, whereas in CVSDM the step size of the approximated signal is progressively increased or decreased, in case the same state has been observed three or four times in a row. Tombras (1990) has considered the 2-digit ADM that uses memory and looks ahead estimation of step size, generating at its output binary and ternary digits. The 2-bit ADM (Prosalentis and Tombras, 2007) is actually the modification of 2-digit ADM (Tombras, 1990), that eliminates the need for a ternary digit, which is in turn reflected in slightly reduced performance. The forward adaptive algorithm in Denic *et al.* (2017) is based on three-level delta modulation where the quantizer codebook is adapted framewise.

Another coding technique based on DM is sigma-delta modulation (SDM), where an integrator is added in front of the ordinary DM modulator, followed by a differentiator in front of the DM demodulator (Aldajani and Sayed, 2001; Prosalentis and Tombras, 2008, 2009; Bashir *et al.*, 2016; Gray, 1987). SDM is commonly used in analog-to-digital (A/D) signal conversion, where the sampling rates are considerably higher, leading to significantly increased transmission rate defined as the product of the sampling rate and the number of bits per sample used to represent the input amplitude. This can be a limiting factor for applications, such as e.g. speech coding, where smaller transmission rates are required. Hence, in such scenarios, ADM is a better solution.

ADM algorithms are widely used in speech coding, but also in other areas of signal processing, such as networked controlled systems (Gomez-Estern *et al.*, 2011), fiber optic based data transmission of signals from sensors (Visan *et al.*, 2016) or transmission over the noiseless binary channels (Dokuchaev, 2015).

In this paper, we propose two solutions for 2-bit ADM, with a goal to improve the overall performance of the one presented in Prosalentis and Tombras (2007). Both developed algorithms perform frame-by-frame processing of the input signal, and estimate the frame variance to adapt the systems to the input signal variations. Whereas the algorithm in Prosalentis and Tombras (2007) performs the step size adaptation sample-by-sample,

the first proposed algorithm, namely 2-bit hybrid ADM, performs adaptation both at the frame level and at the sample level. In this case, the estimated variance is used to determine the initial quantization step-size for each frame, and eventually, instantaneously adaptive logic for step size given in Prosalentis and Tombras (2007) is applied within the particular frame. Note that the ability to track well the time-varying signals (e.g. speech) and consequently achieve high performance is related to the good choice of the initial step size value, which in Prosalentis and Tombras (2007) is determined using an external LDM configuration. In our algorithm, the step size initialization is embedded and avoids excessive preprocessing. Information about the frame variance is required at the receiving end; hence, it needs to be quantized and transmitted once per each frame using the finite number of bits.

The second presented 2-bit ADM algorithm is introduced to improve the one in Prosalentis and Tombras (2007) in terms of the employed quantizer. In particular, we have upgraded the standard DM scheme with the optimal 2-bit scalar quantizer designed for Laplacian probability density function (pdf), that is applied in forward adaptive scheme (Denic *et al.*, 2017; Nikolic and Peric, 2008; Peric *et al.*, 2013). The algorithm employs the adaptive first-order prediction, i.e. the prediction coefficient is adapted to the signal statistics. The estimated frame variance is used to adapt the codebook of the quantizer to the input signal variations, once per each frame. In this scenario, the information about the predictor coefficient needs to be sent to the receiver in addition to the frame variance.

The performance of the proposed 2-bit ADM algorithms is tested for speech signal, as the long-term statistics of speech is well modelled by Laplacian pdf (Jayant and Noll, 1984; Chu, 2003; Gazor and Zhang, 2003) and compared to three baselines, i.e. CFDM, CVSDM and instantaneously adaptive 2-bit ADM (Prosalentis and Tombras, 2007).

The remaining of this paper is organized as follows: in Section 2 we present the proposed hybrid and optimal 2-bit ADM algorithms. In Section 3 the experimental results obtained using the real speech signal are presented and discussed. Finally, concluding remarks are given in Section 4.

2. Two-Bit Hybrid Adaptive Delta Modulation

The proposed 2-bit hybrid ADM algorithm is actually the improvement of the one discussed in Prosalentis and Tombras (2007). The 2-bit ADM (Prosalentis and Tombras, 2007) is characterized by an exponentially variable rate in step-size changes, where the employed quantizers generate output codewords $L_1(n)$ and $L_2(n)$ to represent the sign and the relative magnitude of step size, respectively. In particular, the implemented adaptive logic tries to fit the quantizer step-size to the variations of input signal with unknown variance, starting from some initial step-size value. To avoid an arbitrary step size initialization which might not be optimal, 2-bit ADM (Prosalentis and Tombras, 2007) requires employing an external LDM configuration in the preprocessing stage, for choosing the appropriate initial step size.

In this paper, we developed an algorithm where the step size initialization is embedded and hence, the excessive preprocessing procedures are avoided. Hence, the modification

of the algorithm described in Prosalentis and Tombras (2007) consists of dividing the input signal into frames, estimating frame variance and using this value for determining the initial step size for each frame. The detailed description of the proposed algorithm is illustrated in Fig. 1.

The available input signal is divided into frames using a buffer. Each frame contains certain number of input samples $x_j(n)$, $n = 1, \ldots, M$, where j is the index of the frame and M is the frame size. The frame variance is calculated in the variance estimation block as:

$$\sigma_j^2 = \frac{1}{M} \sum_{n=1}^{M} x_j^2(n), \quad j = 1, \ldots, F. \tag{1}$$

The next step is the variance quantization using L-levels log-uniform quantizer (Q_{LU}) (Denic *et al.*, 2017; Nikolic and Peric, 2008). It has the decision thresholds and the representative levels respectively given by (2) and (3):

$$t_i^{\mathrm{LU}} \, [\mathrm{dB}] = \alpha_{\min} + \Delta_L i, \quad i = 0, \ldots, L, \tag{2}$$

$$y_i^{\mathrm{LU}} \, [\mathrm{dB}] = \alpha_{\min} + \Delta_L (i - 0.5), \quad i = 1, \ldots, L, \tag{3}$$

where $\Delta_L \, [\mathrm{dB}] = \frac{\alpha_{\max} - \alpha_{\min}}{L}$ is the step size, and the dynamic range of the input signal variance is defined as $[\alpha_{\min} \, [\mathrm{dB}], \alpha_{\max} \, [\mathrm{dB}]]$.

For quantizing the logarithmic variance defined as $\alpha \, [\mathrm{dB}] = 10 \log 10(\sigma_j^2 / \sigma_{\mathrm{ref}}^2)$, it uses the mapping function given as: $Q_{\mathrm{LU}}(\alpha) = y_i^{\mathrm{LU}}$ if $\alpha \in (t_{i-1}^{\mathrm{LU}}, t_i^{\mathrm{LU}})$. In the linear domain, the outputs are given as:

$$V_i^2 = \sigma_{\mathrm{ref}}^2 10^{\frac{y_i^{\mathrm{LU}}}{10}}, \quad i = 1, \ldots, L. \tag{4}$$

Information about the employed level of Q_{LU} is transferred to the receiver once per frame as side information (index J in Fig. 1) using $R_{\mathrm{LU}} = \log_2 L$ bits.

Based on the Q_{LU} output determined for each frame j, $V_{i,j}^2$, we define the initial step size $\Delta_j(0)$ once per each frame as:

$$\Delta_j(0) = K V_{i,j}, \tag{5}$$

where K is the real constant determined such that it maximizes SNR. Upon selection of the initial step size for the current frame, it is used for quantization of the first prediction error sample, whereas for all other samples in the frame, the step size $\Delta(n)$, $n = 2, \ldots, M$ is updated according to the specific rule (Prosalentis and Tombras, 2007). Hence, the proposed 2-bit hybrid ADM algorithm includes both adaptation at the frame level and at the sample level, to provide a combination of variance availability and signal-tracking possibilities.

The prediction error signal $e_j(n) = x_j(n) - y_j(n)$ is formed, where $x_j(n)$ is the frame sample value and $y_j(n)$ is its predicted value, and the sign of $e_j(n)$ represented by $L_1(n)$
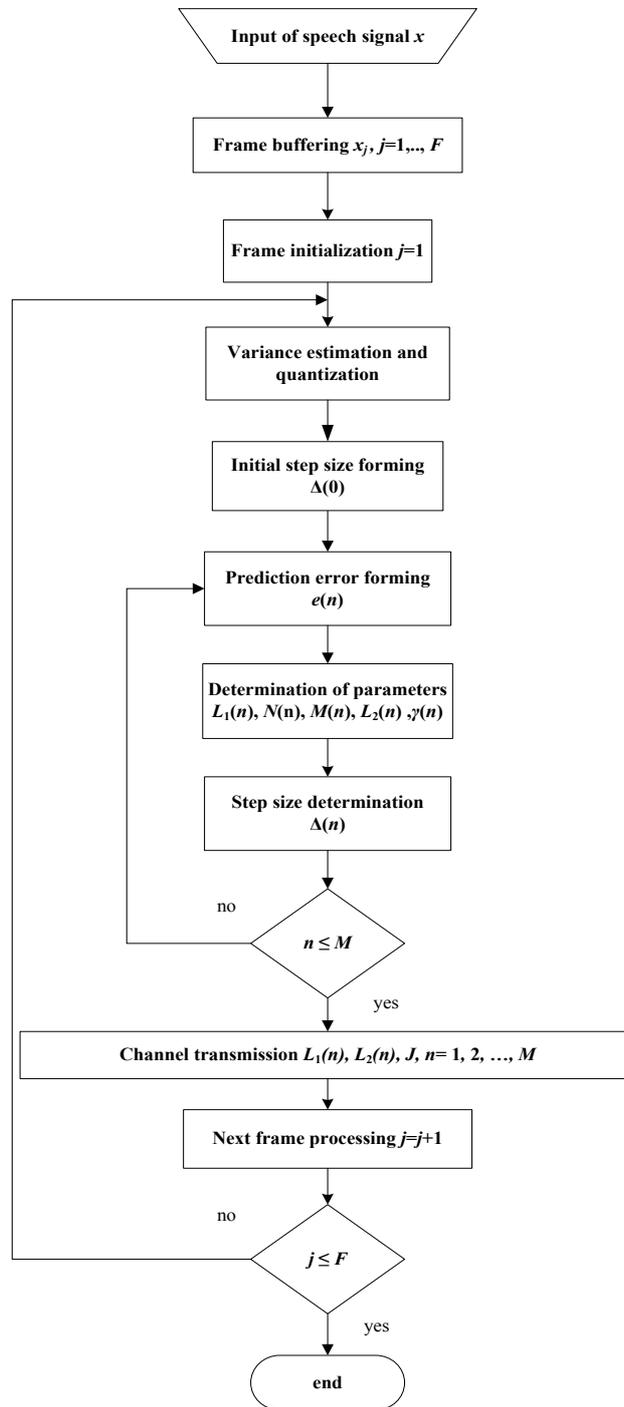
Input of speech signal *x*

Frame buffering *x_j* , *j*=1,.., *F*

Frame initialization *j*=1

Variance estimation and quantization

Initial step size forming Δ(0)

Prediction error forming *e*(*n*)

Determination of parameters *L_1*(*n*), *N*(n), *M*(*n*), *L_2*(*n*) ,*γ*(*n*)

Step size determination Δ(*n*)

no

n* ≤ *M

yes

Channel transmission *L_1(n)*, *L_2(n)*, *J*, *n*= 1, 2, …, *M*

Next frame processing *j*=*j*+1

no

j* ≤ *F

yes

end

Fig. 1. Block diagram of the proposed 2-bit hybrid ADM algorithm.

bit (positive $+1$ or negative -1) is determined as:

$$L_{1,j}(n) = \text{sgn}(e_j(n)) = \text{sgn}(x_j(n) - y_j(n)). \tag{6}$$

This is actually the output codeword of the two-level quantizer. The sign of the prediction error, i.e. $L_{1,j}(n)$ bit, is then compared to the previous $L_{1,j}(n)$ bit and as the comparison result the step parameter $N_j(n)$ is determined:

$$N_j(n) = \begin{cases} \alpha & \text{if } L_{1,j}(n) = L_{1,j}(n-1), \\ \frac{1}{\alpha} & \text{if } L_{1,j}(n) \neq L_{1,j}(n-1), \end{cases} \tag{7}$$

where $\alpha > 1$. Further, the magnitude of the prediction error $|e_j(n)|$ is compared with the appropriate threshold set to be in the middle of the distance between two possible step size values $N_j(n)\Delta_j(n-1)\beta$ and $N_j(n)\Delta_j(n-1)/\beta$, resulting in the step size multiplier $M_j(n)$ determination:

$$M_j(n) = \begin{cases} N_j(n)\beta & \text{if } |e_j(n)| \geqslant \frac{1}{2}(\beta + \frac{1}{\beta})N_j(n)\Delta_j(n-1), \\ \frac{N_j(n)}{\beta} & \text{otherwise}, \end{cases} \tag{8}$$

where $\beta > 1$.

The information about the selected multiplier is transferred to the receiver with the second bit $L_{2,j}(n)$ having two possible values $+1$ or -1, which actually present the output of the second employed quantizer. Then, $L_{2,j}(n)$ is compared to the previous $L_{2,j}(n-1)$ bit to define the parameter $\gamma_j(n)$:

$$\gamma_j(n) = \begin{cases} \gamma & \text{if } L_{2,j}(n) = L_{2,j}(n-1) = -1, \\ 1 & \text{otherwise}. \end{cases} \tag{9}$$

It represents an additional memory function, beside $N_j(n)$. Finally, the step size adaptation rule that applies to a particular frame is given by Prosalentis and Tombras (2007):

$$\Delta_j(n) = M_j(n)\gamma_j(n)\Delta_j(n-1), \quad n = 2, \dots, M, \tag{10}$$

or equivalently

$$\Delta_j(n) = \alpha^{L_{1,j}(n)L_{1,j}(n-1)}\beta^{L_{2,j}(n)}\Delta_j(n-1). \tag{11}$$

The samples of the reconstructed frame (provided at the local decoder as well as in the receiver) have the form:

$$y_j(n) = y_j(n-1) + L_{1,j}(n)\Delta_j(n). \tag{12}$$

The prediction error signal is formed with frame overlapping of one sample, i.e. the first sample in each frame $x_j(1)$ is predicted using the last sample from the previous reconstructed frame $y_j(M)$, except for the first frame where $y_j(1) = 0$, as there is no previous

frame in that case. In addition, $L_{2,j}(M)$ bit from the previous frame is taken into account in processing the next frame.

Regarding the parameters α, β and γ, their selection is explained in detail in Prosalentis and Tombras (2007) and adopted in this paper.

The bits $L_{1,j}(n)$ and $L_{2,j}(n)$ together with the side information that defines the number of bits per frame needed to represent the quantized variance $\frac{R_{\mathrm{LU}}}{M}$ are transferred to the receiver, leading to the overall bit rate:

$$R = R^{P,T} + \frac{R_{\mathrm{LU}}}{M}, \tag{13}$$

where $R^{P,T} = 2$ bit/sample is the rate of the algorithm described in Prosalentis and Tombras (2007).

3. Optimal Two-Bit Adaptive Delta Modulation

In this section, we introduce another variant of 2-bit ADM, designed to improve the one in Prosalentis and Tombras (2007). In a baseline algorithm, one bit codeword is used to represent the sign of the prediction error (positive or negative) and one bit is used to represent the relative magnitude of the prediction error. The parameters of such DM quantizer, i.e. the representative levels and decision thresholds, are determined with respect to the parameters α and β. In particular, depending on the case whether the current and the previous sample of the prediction error have the same or different sign, two four-level quantizers denoted as Q_1 and Q_2 are employed. If they have the same sign, Q_1 is used with the quantization levels (in the positive part) $\{y_3^{Q_1} = \alpha/\beta, y_4^{Q_1} = \alpha \cdot \beta\}$ and the decision threshold being exactly at the half distance between the corresponding levels $t_2^{Q_1} = \alpha \cdot (\beta + 1/\beta)$. Otherwise, Q_2 is used with the quantization levels (in the positive part) $\{y_3^{Q_2} = 1/(\alpha \cdot \beta), y_4^{Q_2} = \beta/\alpha\}$, and the decision threshold $t_2^{Q_2} = (\beta + 1/\beta)/\alpha$. However, the quantizer designed in this way is not the optimal solution.

In this paper, we develop the algorithm with optimal (in the minimum distortion sense) fixed rate ($R = 2$ bit/sample) scalar quantizer. In the following subchapter we will explain the optimal quantizer design followed by its implementation in the proposed 2-bit ADM.

3.1. *Optimal Quantizer Design*

An N-level scalar quantizer Q can be regarded as the functional mapping $Q : R \to Y$, where R is the set of real numbers and $Y = \{y_1, y_2, \ldots, y_N\} \subset R$ is the set of representative levels that forms the code book of size N (Na, 2004; Lee and Na, 2017). In particular, Q partitions the real line into N cells $S_i = (t_{i-1}, t_i]$, $i = 1, 2, \ldots, N$, where t_i, $i = 0, 1, \ldots, N$ are the decision thresholds ($t_0 = -\infty$ and $t_N = \infty$) and each cell is represented by the level $y_i \in S_i$. For the input value $x \in S_i$, the quantizer output is y_i, i.e. it holds $Q(x) = y_i$, if $x \in S_i$.
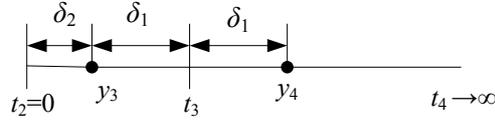
Fig. 2. Illustration of the proposed quantizer.

If we assume that the information source is memoryless and Laplacian with zero-mean and variance σ^2, $p(x, \sigma) = \frac{1}{\sqrt{2}\sigma}e^{-\frac{\sqrt{2}|x|}{\sigma}}$, which is commonly used model for speech (Gazor and Zhang, 2003), then, for a given source, the mean-squared distortion D is evaluated as Jayant and Noll (1984), Chu (2003):

$$D = \sum_i \int_{S_i} (x - y_i)^2 p(x)\, dx. \tag{14}$$

The optimized quantization parameters, i.e. the decision thresholds and the representative levels, that minimize (14), can be obtained by differentiating D over t_i and y_i, and equating with zero, resulting in:

$$t_i = \frac{y_i + y_{i+1}}{2}, \quad i = 1, \ldots, N-1, \tag{15}$$

$$y_i = \frac{\int_{t_{i-1}}^{t_i} x p(x)dx}{\int_{t_{i-1}}^{t_i} p(x)dx}, \quad i = 1, \ldots, N. \tag{16}$$

The equations (15) and (16) are known as the nearest neighbour and the centroid rule, respectively (Jayant and Noll, 1984; Hanzo *et al.*, 2007).

The symmetrical $N = 4$ levels and fixed rate ($R = \log_2 N = 2$ bit/sample) scalar quantizer is designed for zero mean and unit variance, with positive part shown in Fig. 2. Due to the symmetry one can write: $t_1 = -t_2$, $y_1 = -y_4$ and $y_2 = -y_3$. Parameters $\delta_1 = y_4 - t_3$ and $\delta_2 = y_3 - t_2$ in Fig. 2 are the offsets, representing the distance between the corresponding representative level and the lower decision threshold. Note that δ_i, $i = 1, 2$, completely defines the proposed quantizer, as its decision thresholds and representative levels can be specified as:

$$t_3 = \delta_1 + \delta_2, \qquad y_3 = \delta_2, \qquad y_4 = 2\delta_1 + \delta_2. \tag{17}$$

Theorem 1. *An optimal* 2*-bit scalar quantizer can be designed using the following iterative rule*:

$$\delta_2^{(i+1)} = \frac{1}{\sqrt{2}} - \sqrt{2}\exp\left(-\left(1 + \sqrt{2}\delta_2^{(i)}\right)\right). \tag{18}$$

Proof. Substituting Laplacian pdf (for $\sigma = 1$) into (16) we arrive at:

$$y_3 = \frac{1}{\sqrt{2}} - \frac{t_3 \exp(-\sqrt{2}t_3)}{1 - \exp(-\sqrt{2}t_3)}, \tag{19}$$

$$y_4 = t_3 + \frac{1}{\sqrt{2}}. \tag{20}$$

According to the basic definition of offset δ_1 and (20), it is obvious that $\delta_1 = 1/\sqrt{2}$. According to (17), we have $t_3 = 1/\sqrt{2} + \delta_2$, and substituting in (19), after some mathematical manipulations, we obtain:

$$\delta_2 = \frac{1}{\sqrt{2}} - \sqrt{2}\exp\left(-\left(1 + \sqrt{2}\delta_2\right)\right), \tag{21}$$

which can be solved iteratively; thus, completing the proof. $\qquad\square$

Corollary 1. *Total distortion of the optimal four-level (2-bit) quantizer is specified as:*

$$D = \delta_2^2. \tag{22}$$

Proof. Total distortion given by (14) can be rewritten as:

$$D = 2\int_{t_2=0}^{t_3} x^2 p(x)dx - 4y_3 \int_{t_2=0}^{t_3} xp(x)dx + 2y_3^2 \int_{t_2=0}^{t_3} p(x)dx$$
$$+ 2\int_{t_3}^{\infty} x^2 p(x)dx - 4y_4 \int_{t_3}^{\infty} xp(x)dx + 2y_4^2 \int_{t_3}^{\infty} p(x)dx. \tag{23}$$

Knowing that $\sigma_{ref}^2 = 2\int_0^{\infty} x^2 p(x)dx = 1$ and using (16) after some mathematical manipulations we arrive at:

$$D = 1 - y_3^2 P(y_3) - y_4^2 P(y_4), \tag{24}$$

where $P(y_3)$ and $P(y_4)$ are the probabilities of occurrence of the levels y_3 and y_4, respectively:

$$P(y_3) = \int_0^{t_3} p(x)dx = \frac{1}{2}\left(1 - \exp(-\sqrt{2}t_3)\right), \tag{25}$$

$$P(y_4) = \int_{t_3}^{\infty} p(x)dx = \frac{1}{2}\exp(-\sqrt{2}t_3). \tag{26}$$

Substituting (17) in (25) and (26) and further applying in (24) results in:

$$D = 1 - \delta_2^2 - 2(\sqrt{2}\delta_2 + 1)\exp\left(-\left(1 + \sqrt{2}\delta_2\right)\right). \tag{27}$$

Table 1
Performance of the proposed 2-bit optimal quantizer
and baselines for the Laplacian source with zero
mean and unit variance.

	Q	Q_1	Q_2
D	0.18	0.20	0.19
SQNR [dB]	7.54	7.00	7.24
R [b/s]	2	2	2

Finally, using (21), after some basic mathematical manipulations, (27) becomes (Na, 2004; Lee and Na, 2017):

$$D = \delta_2^2. \tag{28}$$

Furthermore, the performance of the developed 2-bit optimal quantizer (denoted as Q) is compared to the baselines Q_1 and Q_2 for $\alpha = 1.1$ and $\beta = 1.8$, used in Prosalentis and Tombras (2007), by assuming the memoryless Laplacian source with zero-mean and unit variance, which is the standard approach in scalar quantizer design (Jayant and Noll, 1984). The results in terms of distortion D and signal-to-quantization-noise ratio SQNR $= 10 \log_{10}(1/D)$ are provided in Table 1. It is evident that the 2-bit optimal quantizer outperforms baselines Q_1 and Q_2 by nearly 0.5 dB and 0.3 dB respectively, in terms of SQNR. □

3.2. *Implementation of the Optimal Quantizer in Two-Bit Delta Modulation*

The diagram of the proposed adaptive two-bit delta modulation is shown in Fig. 3, where the optimal two-bit ($N = 4$ levels) scalar quantizer, with framewise codebook adaptation (Denic *et al.*, 2017; Dincic *et al.*, 2016; Nikolic and Peric, 2008; Peric *et al.*, 2013) is applied. In addition to the buffering, variance estimation and log-uniform quantization steps in the previous algorithm, the additional step for correlation coefficient estimation is introduced. Particularly, for the current frame, the prediction error signal $e[n] = x[n] - \hat{x}[n]$ is fed to the quantizer input, where $x[n]$ denotes the original sample value, $\hat{x}[n] = a \cdot x[n-1]$ denotes the predicted sample value and a is the optimal predictor coefficient determined as in Jayant and Noll (1984):

$$a = \frac{E\{x[n]x[n-1]\}}{E[x^2[n-1]]} = \frac{R_x(1)}{R_x(0)} = \rho, \tag{29}$$

where $E\{\cdot\}$ is the mathematical expectation, $R_x(0)$ and $R_x(1)$ represent the autocorrelation function at lags 0 and 1, respectively, and ρ is the correlation coefficient. Since the information about the predictor coefficient is required at the receiving end (as well as in the local decoder), ρ is quantized using the N_g-levels uniform quantizer (Q_ρ):

$$\rho \in \rho_l \mid \rho_l = \rho_{\min} + (2l-1)\frac{\Delta^u}{2}, \quad l = 1, \ldots, N_g, \ \Delta^u = \frac{\rho_{\max} - \rho_{\min}}{N_g}, \tag{30}$$
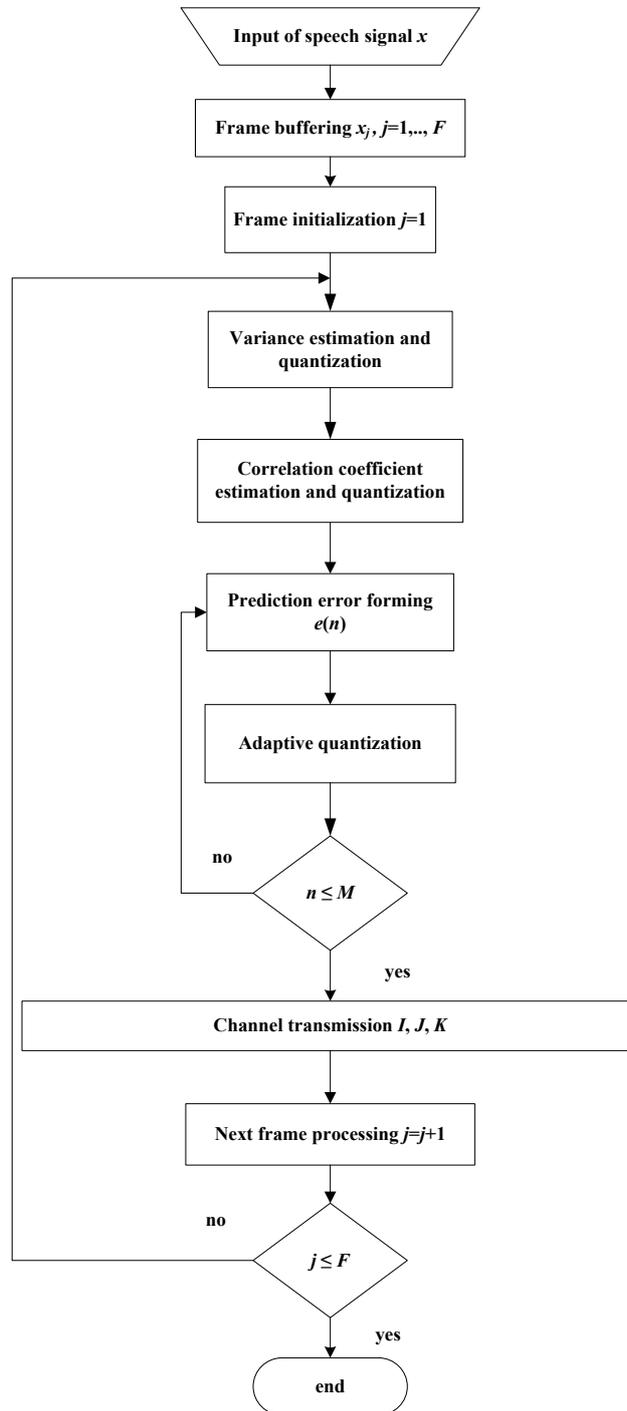
Fig. 3. Block diagram of the 2-bit optimal ADM algorithm.

and information about this is transferred once per each frame (i.e. the predictor coefficient is adjusted once per frame) with $R_\rho = \log_2 N_g$ bits. The adaptation to the variance of prediction error is performed for each frame, and the codebook of the employed two-bit quantizer is updated once per frame according to:

$$t_3^a = g \cdot t_3(\sigma = 1), \tag{31}$$

$$y_3^a = g \cdot y_3(\sigma = 1), \qquad y_4^a = g \cdot y_4(\sigma = 1), \tag{32}$$

where 'a' in the superscript indicates the adapted decision thresholds and representative levels, and gain g is defined as in Jayant and Noll (1984):

$$g = V_{k,j} \cdot \sqrt{1 - \rho_l^2}, \tag{33}$$

where $V_{k,j}$ is given by (4) and ρ_l is given by (30).

Reconstructed signal value $y(n)$ within the current frame is determined as:

$$y(n) = \rho_l \cdot y(n-1) + y_i^a \mathrm{sgn}(e(n)), \tag{34}$$

where y_i^a is defined using (32).

The bit rate of the proposed 2-bit optimal ADM is given by:

$$R^{\mathrm{opt}} = 2 + \frac{R_{\mathrm{LU}} + R_\rho}{M}, \tag{35}$$

where, compared to (13), the side information is increased by R_ρ/M bits, transmitting the information about the predictor coefficient.

4. Experimental Results and Discussion

This section presents and discusses the experimental results obtained in speech coding, since Laplacian pdf can be considered to be a good model for long-term statistics of speech (Gazor and Zhang, 2003). Experiments are performed using four different speech signals recorded in wav format (two male and two female American English speakers), with basic properties presented in Table 2. The amplitude range of the considered speech signals is normalized within the range $[-1, 1]$. All speech signals used in experiments contain both voiced and unvoiced speech.

As an objective measure of quality the segmental SNR (SNR$_{\mathrm{seg}}$) is used, which is calculated separately over all speech frames and then averaged. SNR$_{\mathrm{seg}}$ can be defined as (Hanzo *et al.*, 2007):

$$\mathrm{SNR}_{\mathrm{seg}} = \frac{1}{F} \sum_{j=1}^{F} 10 \log_{10} \left(\frac{\sigma_j^2}{D_j} \right), \tag{36}$$

Table 2
Basic information of the employed speech signals.

Speaker	Sampling frequency [Hz]	Duration [s]	No. of uttered sentences
Male 1	22050	9	2
Male 2	22050	6	1
Female 1	22050	9	2
Female 2	22050	4	1

where F is total number of frames, σ_j^2 is the variance of the j-th speech frame given by (1), and D_j is the distortion of the j-th frame:

$$D_j = \frac{1}{M} \sum_{n=1}^{M} (x(n) - y(n))^2, \quad j = 1, \ldots, F, \tag{37}$$

where M is the frame length.

The performance of the proposed 2-bit ADM algorithms is investigated for frame lengths of 10 ms, 20 ms and 30 ms. Hence, the total number of frames, denoted as F, depends on the duration of the employed speech signal and the frame size.

As the baselines we employ CFDM, CVSDM and 2-bit ADM algorithm (Prosalentis and Tombras, 2007). To have comparable results, all algorithms should generate the same bit rate at their output. Hence, different sampling rates have to be employed for different algorithms. For CFDM and CVSDM signal 22050 Hz sampling rate is used, while the baseline 2-bit ADM operates at half the sampling rate of CFDM, i.e. 22050/2 = 11025 Hz, to produce the same output baud. The sampling rates of the proposed solutions depend on the frame lengths and they are given by 22050/R kHz for 2-bit hybrid ADM and 22050/R^{opt} kHz for 2-bit optimal ADM, where R and R^{opt} are defined in (13) and (35), respectively.

For the proposed 2-bit hybrid ADM we choose parameters $\alpha = 1.1$, $\beta = 1.8$, $\gamma = 1.2$, same as in 2-bit ADM baseline (Prosalentis and Tombras, 2007). In addition, we use the log-uniform quantizer with $L = 32$ levels ($R_{\mathrm{LU}} = 5$ bits) for variance quantization, that is used to adapt the initial step size for each frame (2-bit hybrid ADM) or to adapt the quantizer codebook (2-bit optimal ADM), and $N_g = 32$ levels ($R_\rho = 5$ bits) for quantization of the predictor coefficient. For CFDM we adopt $\alpha = 1.1$ and for CVSDM we use $\beta = 0.9$.

In case of baselines the same initial step-size value, denoted as δ_0, is used, i.e. the one that maximizes SNR of LDM, while the variable step size is limited into upper $\Delta_{\max}$ and lower $\Delta_{\min}$ value, providing $\Delta_{\max}/\Delta_{\min} = 1000$ (i.e. 60 dB dynamic range).

For the proposed 2-bit hybrid ADM algorithm, the initial step size $\Delta_j(0)$ is, according to (5), determined once per each frame and depends on constant K, which should be chosen such that it maximizes SNR. Fig. 4 shows the selection of optimal K for a given speech signal (male 1 in Table 2) and frame length of 20 ms, indicating that $K_{\mathrm{opt}} = 0.29$ fulfils the criterion of maximal SNR. The optimal values of K are chosen in a similar way for the frame sizes of 10 and 30 ms. Table 3 lists the optimal values of K for all four speech signals included in the experiment and different frame lengths.
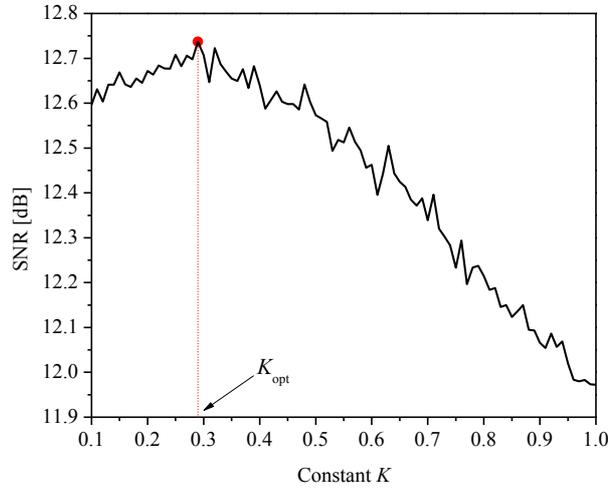
Fig. 4. Selection of the optimal value of constant K for speech frames of size 20 ms and $L = 32$-levels Q_{LU} (2-bit hybrid ADM; $\alpha = 1.1$, $\beta = 1.8$, $\gamma = 1.2$).

Table 3
The optimal values of K for different speech signals and different frame length.

Speaker	10 ms	20 ms	30 ms
Male 1	0.33	0.29	0.28
Male 2	0.26	0.24	0.29
Female 1	0.26	0.18	0.19
Female 2	0.16	0.13	0.22

Figure 5 illustrates SNR as a function of the input signal level (in dB) for two proposed 2-bit ADM algorithms for the frame size of 20 ms, as well as three baselines. The results for male speakers are presented in Fig. 5(a) and Fig. 5(b), whereas the remaining two subplots refer to the female speakers. It can be seen in Fig. 5 that CFDM offers stable SNR in a relatively wide dynamic range, while CVSDM provides slightly higher (Fig. 5(a), (c), (d)) or substantially higher (Fig. 5(b)) maximum SNR at the expense of significantly smaller dynamic range. Two-bit ADM (Prosalentis and Tombras, 2007) achieves higher maximum SNR values than CFDM with only slightly smaller dynamic range. On the other hand, in all scenarios, the proposed 2-bit ADM algorithms offer constant SNR in the entire dynamic range. It is evident that both proposed algorithms outperform the baselines. For example, in case of male 1 speaker (Fig. 5(a)), the proposed 2-bit hybrid ADM has nearly 1.4 dB higher SNR than 2-bit ADM baseline and over 3 dB higher than CVSDM and CFDM. In case of 2-bit optimal ADM, we report the gain in maximal SNR of 3 dB over 2-bit ADM baseline and over 5 dB in case of CVSDM and CFDM.

Table 4 lists the achieved SNR_{seg} values averaged over all frames (dynamic range [-40 dB, 40 dB]) for all considered speech signals and different frame lengths, for two proposed algorithms (2-bit hybrid ADM and 2-bit optimal ADM). According to the re-
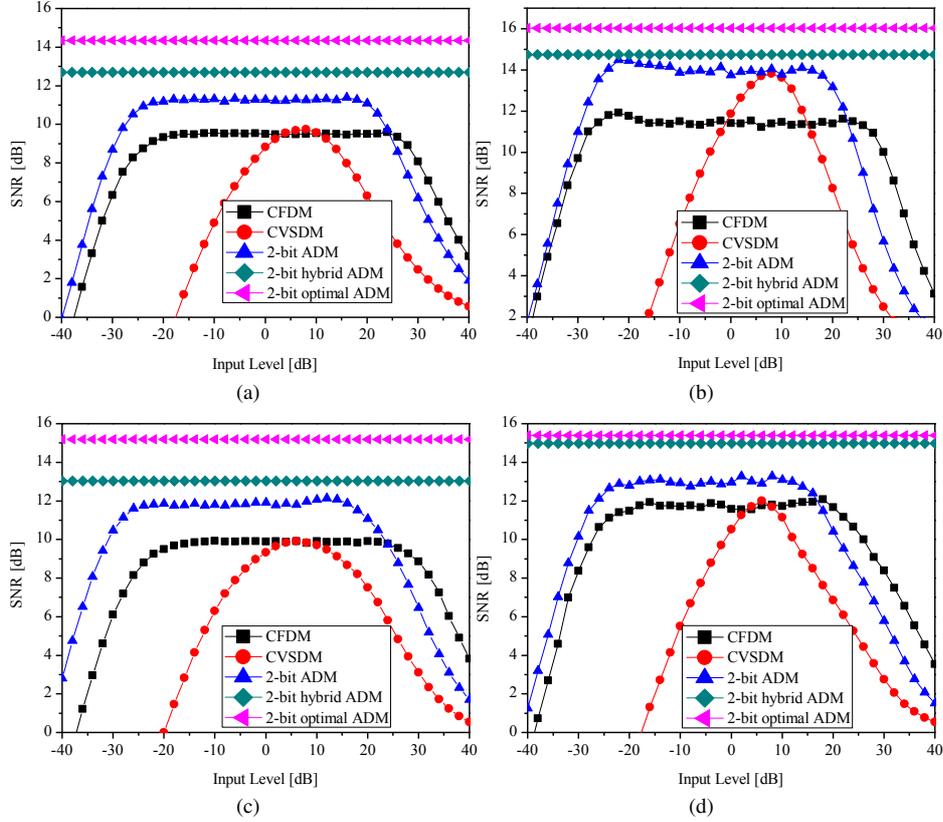
Fig. 5. SNR versus different variances of speech signal for CFDM, CVSDM, 2-bit ADM and the proposed 2-bit hybrid ADM ($\alpha = 1.1$, $\beta = 1.8$, $\gamma = 1.2$ and $L = 32$-levels Q_{LU}) and 2-bit optimal ADM (Q_{LU} with $L = 32$ levels and Q_ρ with $N_g = 32$ levels) (frame length 20 ms) operating at the same output bit rate for: (a) male 1, (b) male 2, (c) female 1 and (d) female 2 speaker.

Table 4

The average SNR_{seg} of the proposed 2-bit hybrid ADM ($\alpha = 1.1$, $\beta = 1.8$, $\gamma = 1.2$ and $L = 32$-levels Q_{LU}) and 2-bit optimal ADM (Q_{LU} with $L = 32$ levels and Q_ρ with $N_g = 32$ levels), obtained in the dynamic range (-40 dB–40 dB) for various frame lengths at the output rate of 22050 bps.

Speaker	10 ms		20 ms		30 ms	
	SNR^h	SNR^o	SNR^h	SNR^o	SNR^h	SNR^o
Male 1	12.74	14.60	12.73	14.34	12.57	13.81
Male 2	14.88	16.55	14.76	16.03	14.67	15.34
Female 1	13.06	15.35	13.05	15.19	13.02	15.05
Female 2	15.22	16.20	14.99	15.40	14.50	14.94

sults in a given table and the ones in Fig. 5, the attained gain in SNR over the 2-bit ADM (Prosalentis and Tombras, 2007) is in the range from 0.37 dB to 1.5 dB in case of 2-bit hybrid ADM and in the range from 1.04 dB to 2.95 dB in case of 2-bit optimal ADM, considering the frame length of 30 ms. Similarly, we report the gain in SNR with respect
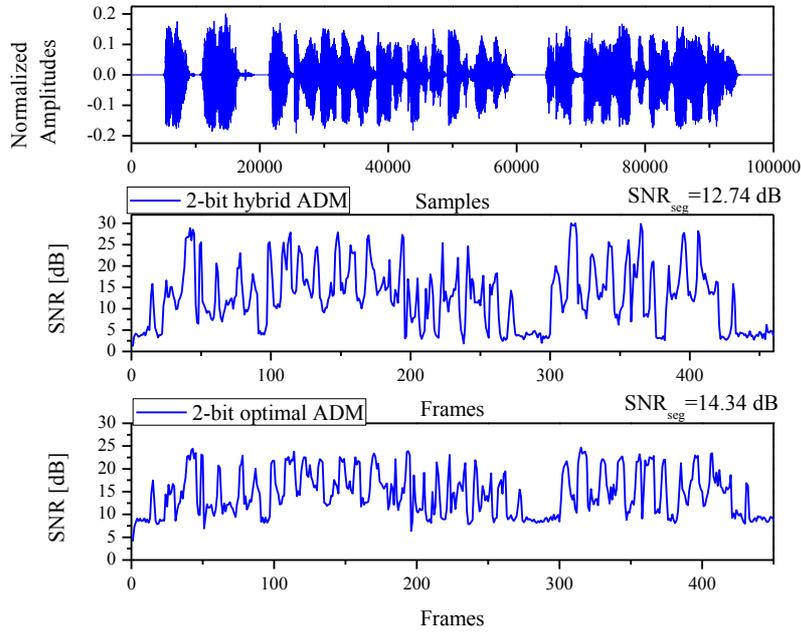
Fig. 6. The original speech signal and SNR over speech frames of size 20 ms ($F = 450$) for 2-bit hybrid ADM ($K_{\mathrm{opt}} = 0.29$, $\alpha = 1.1$, $\beta = 1.8$, $\gamma = 1.2$ and $L = 32$-levels Q_{LU}) and 2-bit optimal ADM (Q_{LU} with $L = 32$ levels and Q_ρ with $N_g = 32$).

to the 2-bit baseline within the range from 0.58 to 2.22 dB in case of 2-bit hybrid ADM and within the range from 2.25 to 3.3 dB in case of 2-bit optimal ADM, when the frames of 10 ms length are employed. Furthermore, it can be observed that, for both algorithms, better performance is obtained for shorter frames (10 ms), which is expected since the initial step size, as well as the quantizer codebook are updated more often. However, this improvement is obtained at the cost of increased bit rate, as for shorter frames the side information is transferred more often. Therefore, as the rate-quality compromise solution we recommend the implementation of the proposed algorithms with the frame size of 20 ms.

SNR across different frames with the length of 20 ms of the original speech signal (male 1 speaker) for both proposed algorithms is depicted in Fig. 6. Observe that smaller variations in SNR for both voiced and unvoiced frames are obtained in case of 2-bit optimal ADM, leading to the higher SNR$_{\mathrm{seg}}$ value.

5. Conclusion

This paper considers two solutions of the 2-bit adaptive delta modulation, namely 2-bit hybrid and 2-bit optimal ADM. In 2-bit hybrid ADM, the estimated variance is used to initialize the step size for each frame, followed by the same step size adaptation procedure as in the instantaneously 2-bit ADM baseline algorithm. Hence, the step size initialization is embedded in the algorithm and avoids using external algorithms for determining

the initial step size. In 2-bit optimal ADM the quantizer is optimally designed assuming Laplacian distribution. Both the quantizer codebook and the predictor coefficient are adapted framewise. The proposed algorithms have shown to be superior in speech coding, when compared to baselines, i.e. 2-bit ADM, CFDM and CVSDM, having wider dynamic range and offering higher performance, measured by SNR. According to the obtained results, there is a great possibility of implementation of the developed algorithms in practical processing of signals, which, as speech signal, have statistics modelled by the Laplacian pdf.

References

Aldajani, M.A., Sayed, A.H. (2001). Stability and performance analysis of an adaptive sigma-delta modulator. *IEEE Transactions on Circuits and Systems II*, 48(3), 233–244.

Bashir, S., Ahmed, S., Kakkar, V., (2016). Design and performance trends of low power sigma-delta A/D converters. *Journal of VLSI Design Tools & Technology*, 6(2), 5–12.

Chu, W.C. (2003). *Speech Coding Algorithms: Foundation and Evolution of Standardized Coders*. John Wiley & Sons, New Jersey, NJ.

Denic, B., Peric, Z., Despotovic, V. (2017). Three-level delta modulation for Laplacian source coding. *Advances in Electrical and Computer Engineering*, 17(1), 95–102.

Dincic, M., Peric, Z., Jovanovic, A. (2016). New coding algorithm based on variable-length codewords for piecewise uniform quantizers. *Informatica*, 27(3), 527–548.

Dokuchaev, N. (2015). On transmission of a continuous signal via a noiseless binary channel. *IEEE Signal Processing Letters*, 22(8), 1171–1174.

Gazor, S., Zhang, W. (2003). Speech probability distribution. *IEEE Signal Processing Letters*, 10(7), 204–207.

Gibson, J.D. (2016). Speech compression. *Information*, 7(32), 1–22.

Gibson, J.D. (2017). On the high rate, independence, and optimal prediction assumptions in predictive coding. In: *Proceedings of the IEEE Information Theory and Applications Workshop (ITA)*, San Diego, CA, USA.

Gomez-Estern, F., Canudas-de-Wit, C., Rubio, F.R. (2011). Adaptive delta modulation in networked controlled systems with bounded disturbances. *IEEE Transactions on Automatic Control*, 56(1), 129–134.

Gray, M.R. (1987). Oversampled sigma-delta modulation. *IEEE Transactions on Communications*, 35(5), 481–489.

Hanzo, L., Somerville, C., Woodard, J. (2007). *Voice and Audio Compression for Wireless Communications*. John Wiley & Sons, Chichester.

Hastie, T., Tibshirani, R., Friedman, R. (2008). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York.

Hucha Arce, F., Moonen, M., Verhelst, M., Bertrand, A. (2017). Adaptive quantization for multichannel Wiener filter-based speech enhancement in wireless acoustic sensor networks. *Wireless Communications and Mobile Computing*, 1–15, Article ID 3173196.

Jayant, N.S., Noll, P. (1984). *Digital Coding of Wavefors*. Prentice Hall, New Jersey, NJ.

Lee, J., Na, S. (2017). A rigorous revisit to the partial distortion theorem in the case of a Laplacian source. *IEEE Communications Letters*, 21(12), 2554–2557.

Na, S. (2004). On the support of fixed-rate minimum mean-squared error scalar quantizers for a Laplacian source. *IEEE Transactions on Information Theory*, 50(5), 937–944.

Nikolic, J., Peric, Z. (2008). Lloyd–Max's algorithm implementation in speech coding algorithm based on forward adaptive technique. *Informatica*, 19(2), 255–270.

Peric, Z., Nikolic, J., Mosic, A., Petkovic, M. (2013). Design of fixed and adaptive companding quantizer with variable-length codeword for memoryless Gaussian source. *Informatica*, 24(1), 71–86.

Prosalentis, E.A., Tombras, G.S. (2007). 2-bit adaptive delta modulation system with improved performance. *EURASIP Journal on Advances in Signal Processing*, Article ID 16286.

Prosalentis, E.A., Tombras, G.S. (2008). Description of a 2-bit adaptive sigma-delta modulation system with minimized idle tones. *EURASIP Journal on Advances in Signal Processing*, Article ID 426502.

Prosalentis, E.A., Tombras, G.S. (2009). Elimination of idle tones by a second order 2-bit adaptive sigma delta modulation system. *ETRI Journal*, 31(4), 393–398.

Sarade, S.S., (2017). Speech compression by using adaptive differential pulse code modulation (ADPCM) technique with microcontroller. *Journal of Electronics and Communication Systems*, 2(3), 1–9.

Tombras, G.S. (1990). New adaptation algorithm for a two-digit adaptive delta modulation system. *International Journal of Electronics*, 68(3), 343–349.

Uddin, S., Ansari, I.R., Naaz, S., (2016). Low bit rate speech coding using differential pulse code modulation. *Advances in Research*, 8(3), 1–6.

Visan, D.A., Jurian, M., Jurian, M., Cioc, I.B., Ionescu, L.M., Lita, A.I. (2016). Delta encoder for fiber optic based data transmission of signals from sensors. In: *Proceedings of the 8th IEEE International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, Ploiesti, Romania.

Zrilic, D.G. (2005). *Circuits and Systems Based on Delta Modulation*. Springer, Berlin.

Z. Peric was born in Niš, Serbia, in 1964. He received the BS, MS and PhD degrees from the Faculty of Electronic Engineering, University of Niš, Serbia, in 1989, 1994 and 1999, respectively. He is a full-time professor at Department of Telecommunications, Faculty of Electronic Engineering, University of Niš. His current research interests include the information theory and signal processing. He is an author and co-author of over 240 papers. Dr. Peric has been a reviewer of a number of journals, including *IEEE Transactions on Information Theory*, *IEEE Transactions on Signal Processing*, *IEEE Transactions on Communications*, *Compel*, *Informatica*, *Information Technology and Control*, *Expert Systems with Applications and Digital Signal Processing*.

B.D. Denic was born in Vrbeštica, township Uroševac, Serbia, in 1986. He received BS and MS degrees in electronics and telecommunications from the Faculty of Technical Sciences, University of Priština, Serbia. He is currently a research assistant and PhD student at the Faculty of Electronic Engineering, University of Niš, Serbia. His current research interests include scalar quantization and signal processing. He has published 5 papers in peer-reviewed international journals on the above subject.

V. Despotovic received his PhD degree in electrical engineering from the University of Niš, Serbia, in 2012. Currently he is working as assistant professor at the University of Belgrade, Serbia. Previously he was engaged as postdoctoral researcher at the University of Paderborn, Germany. His main research interests include statistical signal processing, natural language processing, speech coding, fractional calculus and machine learning.