IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 21 June 2024, accepted 3 July 2024, date of publication 5 July 2024, date of current version 17 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3424315

RESEARCH ARTICLE

FBMC/OQAM-Based Secure Voice Communications Over Voice Channels

SARA ČUBRILOVIĆ[1,2], (Member, IEEE), ZVEZDANA KUZMANOVIĆ[1,2], (Member, IEEE), MARIJA PUNT[1], DESIMIR VUČIĆ[2,3], AND BRANKO KOVAČEVIĆ[1], (Senior Member, IEEE)

[1]School of Electrical Engineering, University of Belgrade, 11000 Belgrade, Serbia
[2]Vlatacom Institute, 11070 Belgrade, Serbia
[3]School of Computing, Union University, 11000 Belgrade, Serbia

Corresponding author: Sara Čubrilović (sara.cubrilovic@vlatacom.com)

ABSTRACT The core objective is to securely transmit digital data over widely accessible voice channels like VoIP or GSM networks. Encryption at the highest security level is applied digitally, but a challenge lies in selecting a modulation technique that effectively converts encrypted digital data into analogue signals capable of navigating the obstacles inherent in voice channels. These channels are lossy and optimized for speech transmission, employing sophisticated signal processing techniques such as Voice Activity Detection (VAD) to maintain speech quality and suppress noise. This paper introduces a novel system design and approach for transmitting Data over Voice (DoV) channels using Offset Quadrature Amplitude Modulation (OQAM) combined with Filter Bank Multicarrier (FBMC) modulation. FBMC/OQAM generates speech-like waveforms that can evade detection by non-speech signal detection mechanisms inherent in voice channels, marking its debut in this domain. In order to demonstrate the feasibility of the proposed system, we have simulated the distortion of FBMC/OQAM signals when processed by an AMR-NB codec using MATLAB®. Subsequently, the performance of the proposed modulation technique is evaluated across various real voice channels, considering actual time and phase deformations, real-time voice transmission bit-rate constraints and possible bit errors, which all demand specific waveform structures for synchronization and algorithms for correct demodulation and decoding at the receiver. Evaluation is done for waveforms that were sent through GSM and VoIP channels. The obtained results demonstrate a favorable bit error rate (BER) on the 3G channel compared to existing methods, such as conventional Orthogonal Frequency Division Modulation (OFDM), thereby proving that FBMC/OQAM is suitable for secure DoV transmission. Finally, the decoded speech from all channels is assessed using Perceptual Evaluation of Speech Quality (PESQ) metric.

INDEX TERMS AMR-based compression, data over voice channels, FBMC/OQAM, secure voice communications, speech-like waveform.

I. INTRODUCTION

In wireless mobile networks, Data over Voice (DoV) techniques are indispensable for ensuring secure voice communication, particularly in specialized devices such as crypto phones. When implementing digital voice channels, signal parameters are carefully chosen to align with the speech model of the system, aiming to maintain intelligibility and speech quality while minimizing perceptually redundant information. However, this approach imposes notable constraints on the application of DoV, as voice channels are primarily optimized for voice transmission. In contrast to traditional data channels, voice channels tend to significantly alter the original signal waveform and shape due to transcoding and audio signal processing. To address this challenge, DoV techniques centered on encoding the data signal into speech-like parameters, designing codebooks, or employing modulation techniques are being actively researched and proposed [1]. These approaches aim to mitigate the distortion introduced by voice channel processing and enhance the reliability of data transmission over voice channels.

The associate editor coordinating the review of this manuscript and approving it for publication was Jun Wang.

Most DoV techniques primarily involve codebook-based methods utilizing parameter mapping or codebook optimization approaches to generate speech-like symbols or waveforms tailored for transmission over voice channels in wireless mobile networks. Parameter mapping entails converting the data to be transmitted into speech parameters, such as line spectral frequencies, pitch frequency and speech frame energy [2], [3], based on a speech production model. A similar approach is presented in [4] and [5], where [4] employs the LPC (Linear Predictive Coding) method to generate speech-like symbols, while [5] explores speech-like waveform generation using an autoregressive (AR) speech model. Additionally, DoV techniques leveraging classical modulation techniques like PSK, QAM, and M-FSK have been proposed [6], [7], [8]. For instance, [9] examines a solution employing orthogonal frequency division multiplexing (OFDM) transmission technique with QPSK and BPSK modulation on low-frequency and high-frequency subcarriers, respectively. For a comprehensive understanding of these and other DoV techniques for secure voice communication, refer to [10]. In [11], a novel distortion-tolerant perceptually-oriented speech encryption scheme is introduced for secure voice communication over voice channels. This scheme demonstrates robustness against typical moderate distortions encountered in such applications.

This paper presents a novel system and analysis of the performance sensitivity of a multicarrier filter bank employing the offset quadrature amplitude modulation (FBMC/OQAM) transmission technique within the context of secure voice communication over a voice channel utilizing an AMR-NB speech LPC codec. Notably, this study represents the first exploration of FBMC/OQAM in this specific application domain. Leveraging MATLAB® software, the study investigates the distortion introduced by the AMR-NB codec on the FBMC/OQAM speech-like signal. Additionally, it examines the impact of various FBMC/OQAM pulse shaping filters and amplitude pulsing of speech-like signals on the received signal quality of this DoV technique when applied to a voice channel employing the AMR-NB speech codec. Subsequently, the proposed modulation technique is successfully implemented for real-time encrypted transmission and evaluated over widely-used GSM (Global System for Mobile communication) and VoIP (Voice over Internet Protocol) channels. A key advantage of FBMC/OQAM over many other DoV techniques for secure voice communication lies in its simplicity and computational efficiency, simultaneously yielding acceptably low BER in a realistic environment and enabling secure DoV communication. Specifically, a set of orthogonal subcarriers can be generated using an inverse fast Fourier transform (IFFT) and a polyphase network (PPN). The FBMC/OQAM speech-like waveform, designed to withstand moderate voice channel distortions, is structured with pulse shaping, silent pauses and amplitude pulsing, thereby enhancing its robustness.

Since the data under examination stems from analogue speech signals generated by MELPe 1200 bps vocoder [12], the results are given in terms of the perceptual evaluation of speech quality (PESQ), as defined by ITU-T recommendation P.862 [13]. PESQ serves as an objective, standardized metric for assessing speech quality. Furthermore, the performance is evaluated also in terms of bit error rate (BER). Even though BER might not be the best metric for speech quality evaluation, given the fact that the bit sequence with the same BER value can result in different speech signals, it is a good error assessment method for data transmission. The subjective speech study was out of scope of this work, due to its extensive time and resource demand.

Chapter II explains the transmission and modulation technique of the proposed FBMC/OQPSK modulation. Chapter III presents simulation results, while Chapter IV analyzes and presents experimental results over real voice channels with real-time bit rate considerations. Chapter V provides the conclusion of the paper.

II. FBMC/OQAM TRANSMISSION TECHNIQUE

An interesting and practical variant of multi-carrier modulation (MCM) is FBMC modulation. FBMC represents a more intricate implementation technique compared to the classic OFDM technique, primarily due to the substitution of the IFFT/FFT module with a filter bank (FB). However, the utilization of a FB makes this technique spectrally more efficient. FBMC/OQAM employs time-frequency well-localized transmitter pulse shaping, resulting in superior spectral efficiency and a significant reduction in both intersymbol interference (ISI) and intercarrier interference (ICI). Moreover, OQAM introduces alternating $\cos(2\pi kt/T)$ and $\sin(2\pi kt/T)$ subcarriers (using phase factors $e^{jk\pi/2}$) and includes an offset of $T/2$ (where k denotes the subcarrier index and T represents the symbol duration) between the in-phase and quadrature components. In this way, orthogonality between subchannels in the FBMC/OQAM transmission system is achieved.

The fundamental concept behind FBMC/OQAM transmission involves the transmission of OQAM symbols. These complex QAM symbols $s_k(n) = s_k^{\Re}(n) + js_k^{\Im}(n)$ are split into their real $s_k^{\Re}(n)$ and imaginary $s_k^{\Im}(n)$ components, where k represents the subcarrier index and n denotes the time index. Subsequently, they are transmitted with a shift of half the symbol interval and a phase shift ensuring that adjacent symbols in both time (n) and frequency (k) positions are in quadrature with each other. Therefore, the FBMC/OQAM baseband signal can be expressed as [14], [15], and [16]:

$$\begin{aligned} x(m) &= x^{\Re}(m) + x^{\Im}(m - M/2) \\ &= \sum_n \sum_{k=0}^{M-1} s_k^{\Re}(n)\, h(m - nM) e^{j\frac{2}{M}k\left(m - \frac{KM-1}{2}\right)}{}_k(n) \\ &\quad + \sum_n \sum_{k=0}^{M-1} s_k^{\Re}(n)\, h(m - nM - \frac{M}{2}) \\ &\quad e^{j\frac{2}{M}k\left(m - \frac{(K+1)M-1}{2}\right)}{}_{k+1}(n) \end{aligned} \tag{1}$$

where the phase rotation factor takes the form $\theta_k(n) = j^k$ [17], ensuring orthogonality between adjacent subchannels and successive data symbols. Here, $h(m)$ represents the impulse response of the prototype filter with a length of $L_h = KM$,where K denotes the overlapping factor and M indicates the total number of subcarriers.

The real-valued symmetric impulse response of the popular PHYDYAS *half*-Nyquist prototype filter is expressed as:

$$h(m) = 1 + 2\sum_{k=1}^{K-1} (-1)^k H_k \cos\left(\frac{2\pi k}{KM} m\right), h(0) = 0 \tag{2}$$

where $m = 1, \ldots, KM - 1$ and H_k are frequency coefficients of the prototype filter. The $2K - 1$ non-zero frequency coefficients for $K = 4$ are given by $H_0 = 1, H_1 = H_{-1} = 0.97196, H_2 = H_{-2} = \sqrt{2}/2, H_3 = H_{-3} = 0.235147$ [16].

In the realm of FBMC/OQAM implementations, two main approaches exist: frequency spreading FBMC/OQAM (FS-FBMC/OQAM) and poly-phase network FBMC/OQAM (PPN-FBMC/OQAM). Both implementations yield the same FBMC transmit signal. However, in this paper, PPN-FBMC/OQAM is adopted due to its computational efficiency. Based on [15] and [16], this approach significantly reduces complexity by utilizing two M-point IFFTs and two PPNs.

A. PPN-FBMC/OQAM

The PPN-FBMC-OQAM primarily comprises synthesis filter bank (SFB) at the transmitter side and analysis filter bank (AFB) at the receiver side. The SFB is composed of OQAM preprocessing, IFFT, PPN filtering and parallel/serial (P/S) conversion. Accordingly, the AFB consists of OQAM postprocessing, FFT, PPN filtering and serial/parallel (S/P) conversion. This system achieves a notable reduction in complexity by employing two M-point IFFTs and two PPNs. The real $\Re\{\cdot\}$ and imaginary $\Im\{\cdot\}$ parts of the data are extracted from the input sequence of complex data symbols $s(n)$ and each part undergoes separate processing via IFFT and PPN. PPN-FBMC/OQAM transmitter is illustrated in Figure 1 [15], [16].

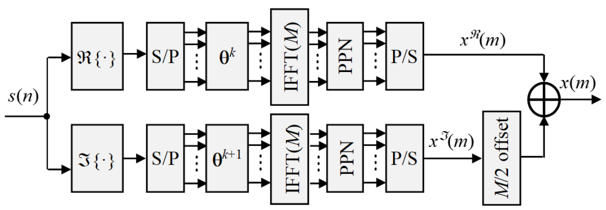

FIGURE 1. PPN-FBMC/OQAM transmitter.

In the OQAM preprocessing and IFFT stages, both the real and imaginary components of the complex QAM symbols are processed separately. Each component undergoes modulation with M real (in-phase branch) and M imaginary (quadrature branch) components of the complex QAM symbols using the same implementation. However, in the quadrature processing, an additional multiplication by $j(\theta k + 1(n) = j\theta k(n))$ is applied in the OQAM preprocessing. As a result, signals from both the in-phase and quadrature branches are obtained. These signals are then filtered (PPN) before undergoing P/S conversion. Finally, the resulting signals from both branches, with $M/2$ samples delay applied in the quadrature branch, are summed to produce a PPN-FBMC/OQAM signal. This approach ensures that both the in-phase and quadrature components are properly processed and combined to form the final PPN-FBMC/OQAM signal.

The PPN-FBMC/OQAM receiver executes matched signal demodulation, which reverses the modulation process performed by the modulator. Upon receiving the signal $r(n)$, it is split into two branches. The real part of the modulated symbols is demodulated in one branch, while the imaginary part is demodulated in the other branch, which includes a delay of $M/2$ samples. Initially, the received signal $r(n)$ is filtered using a PPN, identical to the one utilized at the transmitter. Subsequently, the filtered signal is transformed into the frequency domain using an M-point FFT. Finally, the phase shift introduced during transmission is compensated and the original symbols are reconstructed. This receiver architecture ensures that the received signal is properly processed to extract the transmitted symbols, facilitating accurate recovery of the original data.

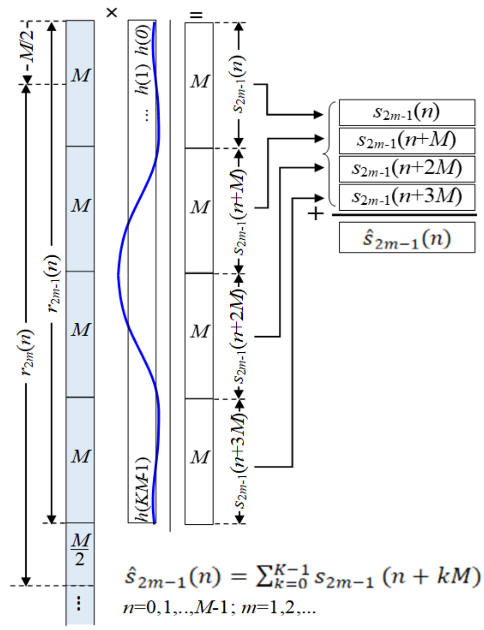

FIGURE 2. Illustration of PPN in-phase branch processing at reception (K=4).

The illustration of the PPN in-phase branch processing of the PPN-FBMC/OQAM signal at reception ($K = 4$) is depicted in Figure 2. Following the reception, the blocks of the received signal $r_l(n) = r[n + (l - 1)M/2]$, where $n = 0, 1, \ldots, M - 1$ and $l = 1, 2, \ldots,$ are shaped by multiplying them with the impulse response of the prototype filter $h(n)$ with KN samples. This yields sequences $s_l(n) = h(n) \cdot r_l(n)$. For sequences $s_l(n)$ with odd l (i.e., $s_{2m-1}(n)$ for $m = 1, 2, \ldots$), the in-phase branch processing is performed. Equally, for sequences $s_l(n)$ with even l (i.e.,

$s_{2m}(n)$ for $m = 1, 2, \ldots$), the quadrature branch processing is performed. By overlapping and summing K of these sequences $s_l(n+kM)$, for $k = 0, 1, \ldots, K-1$, a sequence $\hat{s}_l(n) = \sum_{k=0}^{K-1} s_l(n+kM)$ is obtained. Finally, the FFT operation is applied to these sequences. For the quadrature branch processing, it is necessary to delay the received signal by $M/2$ samples. This ensures proper synchronization and alignment with the in-phase processing.

B. SECURE VOICE COMMUNICATIONS OVER VOICE CHANNELS

Secure voice communications over voice channel entail several crucial steps: digitalizing the speech, compressing it to a low bit rate suitable for efficient transmission and encrypting the digital data. However, the resulting bit stream may encounter challenges passing through the speech transcoding process seamlessly. Therefore, the low bit rate data must be modulated onto speech-like waveforms capable of successfully passing through the speech codec (e.g., AMR), undergoing speech processing and meeting specific network requirements. It has been shown that OFDM can be perceived as a speech-like waveform [9] and the same assumption can be applied to FBMC/OQAM signal in this context. A special system or device for secure voice communication can be designed either as a standalone device or as an addon connected to a mobile phone through audio interface, via Bluetooth connection or audio cable.

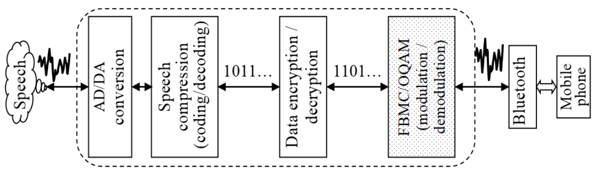

FIGURE 3. Simplified block diagram of the proposed FBMC/OQAM-based secure voice communication system.

The block diagram of the proposed FBMC/OQPSK system for secure voice communication is shown in Figure 3. This system facilitates the transmission of low bit rate data in a manner that is compatible with speech processing and network requirements, generating waveforms suitable for transmission over voice channels while ensuring security and efficiency.

In order to examine the characteristics of the proposed speech encryption system, we analyze the voice channel effects on the FBMC/OQPSK waveform. As an initial test, the waveform is processed by the AMR-NB codec and successfully decoded in MATLAB®. After these positive results, the information carrying waveform is enhanced with different structures for phase and time synchronization, with the goal of achieving even more robustness needed for the real-time communication over available voice channels. Simulations with such complete waveform are performed on the constructed signal waveform both in MATLAB® with the AMR codec and over real voice channels (VoIP and GSM).

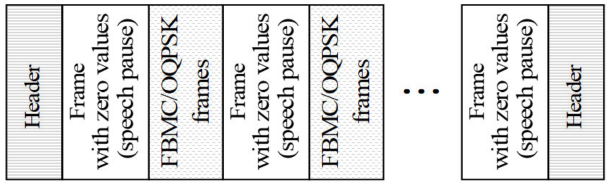

FIGURE 4. Structure of the speech-like FBMC/OQPSK signal.

III. SIMULATION RESULTS

The performance of the proposed DoV technique is assessed through MATLAB® simulations. Figure 4 illustrates the simulated structure of the speech-like FBMC/OQPSK signal. In this implementation of PPN-FBMC/OQPSK, FBMC/OQPSK with 7 subcarriers, an IFFT/FFT of size $M = 64$ and a sampling frequency $fs = 8$kHz are chosen. Various overlap factors K and prototype pulse shaping filters are utilized and analyzed to evaluate their impact on the system performance.

Analyzing and predicting the performance of Data over Voice (DoV) techniques can indeed be challenging, as it relies heavily on the ability of the FBMC/OQPSK signal to withstand distortions introduced by the voice channel. In order to achieve a speech-like structure of the FBMC/OQSK signal and to comprehensively examine the effect of these factors on DoV performance, two parameters are varied:

- the frame duration of the pauses with respect to the frame duration of the speech data
- the amplitude of the FBMC/OQPSK signal

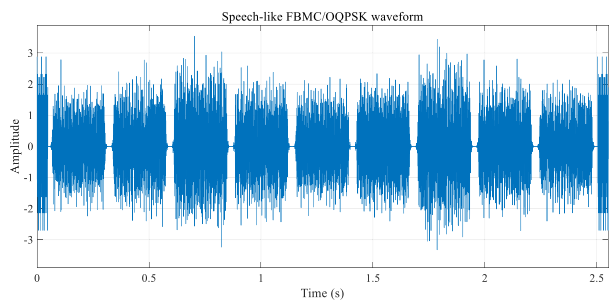

FIGURE 5. Speech-like FBMC/OQPSK waveform with the PHYDYAS half-Nyquist pulse shaping.

Figure 5 depicts an example of a speech-like FBMC/OQPSK waveform featuring PHYDYAS *half*-Nyquist ($K = 4$) pulse shaping, amplitude pulsing, a FBMC/OQPSK frame length of 29 FBMC symbols and a pause (silent period) with a length of 1 FBMC symbol. In this illustration, amplitude pulsing is implemented using amplification factors of 1, 0.8, and 0.7, which periodically adjust the amplitude of the FBMC/OQPSK signal from frame to frame. The introduction of amplitude pulsing results in a 10% improvement in BER. Thus, the periodically varying amplification factors in combination with silent periods contribute to enhancing the BER and robustness of the proposed FBMC/OQAM-based transmission technique when applied to the voice channel with the AMR-NB codec.

Using the speech-like FBMC/OQPSK signal with *half*-Nyquist pulse shaping and an overlap factor of $K = 4$, along with appropriate pauses and amplitude pulsing, a BER of $3.91 \cdot 10^{-2}$ and $1.28 \cdot 10^{-3}$ is achieved at a data rate of 1750 bps for transmission through the theoretical voice AMR-NBcodec with speeds of 7.4 kbps and 12.2 kbps, respectively. These BER values are obtained without any additional distortion analysis or phase corrections. Figure 6 illustrates a segment of the speech-like FBMC/OQPSK waveform distorted by theoretical AMR-NB codec with speeds of 7.4 kbps and 12.2 kbps. As expected, significantly higher distortion is introduced by the AMR-NB 7.4 kbps compared to the AMR-NB 12.2 kbps.

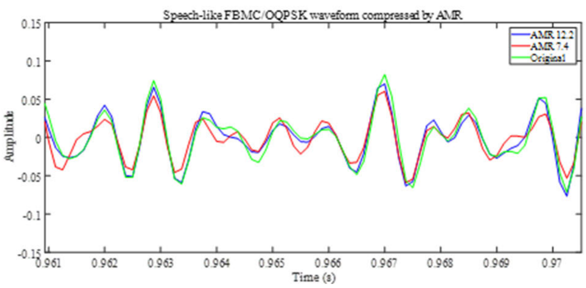

FIGURE 6. Speech-like FBMC/OQPSK waveform compressed and distorted by AMR codec.

Improvements in BER performance can also be attained through the utilization of suitable pulse shaping filters, which significantly impact the spectral properties of FBMC/OQPSK signal. Among these filters, the widely used orthogonal PHYDYAS half-Nyquist pulse shaping filters [17] with various overlapping factors ($K = 2, 3$ and 4) are considered. Additionally, other prototype filters like rectangular, Hamming, Hann, Chebyshev and Gaussian are also evaluated. By employing non-orthogonal prototype filters, such as the Chebyshev filter, BER performance improvement of up to 18% can be achieved compared to the half-Nyquist filter with an overlapping factor of $K = 4$, although orthogonality preservation among subcarriers is not maintained.

Indeed, the phase dispersion introduced by theoretical AMR-NB codec which uses LPC reconstruction filters with non-uniform phase response, stands out as the predominant source of errors in DoV techniques based on speech-like FBMC/OQPSK signal. In Figure 7 (AMR-NB 12.2 kbps), the scatter plot illustrates the received FBMC/OQPSK symbols (blue dots), while the red dots represent the sample means of four constellation points of distorted QPSK symbols. The resultant mean phase shift, attributed to AMR-induced distortion, measures approximately +2.25 degrees. Notably, this mean shift remains consistent across both AMR-NB 12.2 kbps and AMR-NB 7.4 kbps scenarios.

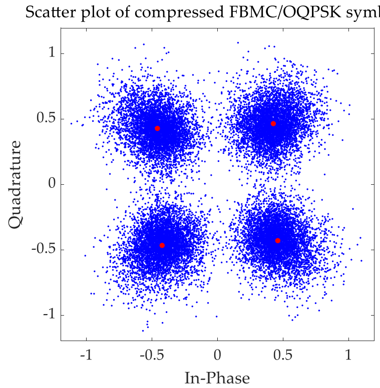

FIGURE 7. Scatter plot of AMR-based distorted FBMC/OQPSK symbols.

However, variations in the probability density function (PDF) and variance of phase shift distortion occur between AMR-NB 12.2 kbps and AMR-NB 7.4 kbps due to AMR-based compression. This discrepancy in PDF characteristics contributes to differences in the BER between the two codecs. Furthermore, it is important to note that this correlation between PDF width (variance) and BER may vary across different subcarrier frequencies.

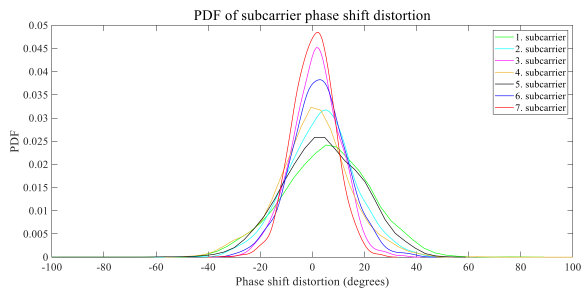

FIGURE 8. Averaged sample probability density functions of FBMC/OQPSK subcarrier phase shift distortion.

Figure 8 provides a visual representation of the averaged sample probability density functions of the FBMC/OQPSK subcarrier phase shift distortion introduced by AMR-NB 12.2 kbps -based compression. Likewise, Figure 8 illustrates that the variances of the mean phase shift distortion vary across different subcarriers. Specifically, the variances of phase shift distortion are notably higher when compression is introduced by AMR 7.4 kbps compared to AMR 12.2 kbps.

In comparison with similar works based on OFDM modulation [9], the proposed FBMC/OQPSK modulation has shown comparable results, but with larger spectral efficiency. Furthermore, it opens new opportunities for further FBMC/OQPSK modulation method implementations in this field.

IV. REAL-TIME CASE STUDY

In order to evaluate the objective performance of the proposed transmission technique, the modulated signal that carries information had to be adapted to overcome all possible distortions and noise conditions that are usually met on the voice channel. Firstly, the signal must exhibit speech characteristics, as it is intended to mimic human speech and effectively bypass all voice detection checks throughout the whole path. Real voice channels are susceptible to nonlinear effects such as signal inversion, insertion and loss of samples, jitter, data loss, non-stationary transmission and many more. Considering all these phenomena, precise time synchronization at the beginning of information segment is crucial for correct decoding of the modulated signal. To address this problem, signaling and synchronization sequences with good autocorrelation properties are added in the predefined points of the signal.

Another pivotal aspect is phase synchronization, which significantly influences the decoding of the FBMC/OQPSK signal. This synchronization is achieved by using predefined pilot symbols capable of determining possible phase shifts. Additionally, 15% of the encoded bits are subjected to Golay (12, 24) Forward Error Correction (FEC) coding, aiding in bit-error detection and correction. The block scheme of the proposed communication system is shown in Figure 9.

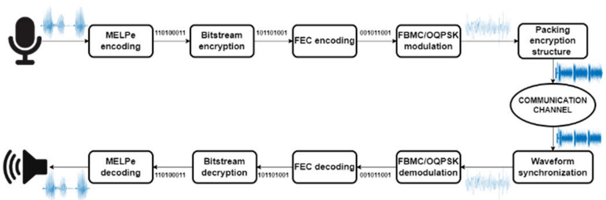

FIGURE 9. System block scheme.

To achieve the required bit-rate of 1200 bps for real-time voice transmission, the effective bit-rate is set significantly higher, specifically at 1407 bps. This augmented rate encompasses several components: the information bits generated by the MELPe 1200 bps vocoder, additional Golay Forward Error Correction (FEC) bits, and overheads for time and phase synchronization. The cumulative signal structure, that is generated in described manner and ready to be sent over various voice channels, is visually represented in Figure 10.

To meet these bit-rate requirements, the following parameters for FBMC/OQPSK and for real-time testing are configured as follows:

- IFFT/FFT size: 128 points
- Subcarrier count: 16
- Sampling frequency: 8000 Hz
- Number of pilot bits: 32
- Filter overlapping factor K: 4
- Filter length: 512 samples
- Transmitted symbols per test: 80000
- Transmitted bits per test: 160000
- Duration of the test signal: 2 minutes

V. EXPERIMENTAL RESULTS

In order to evaluate the performance of the FBMC/OQPSK transmission technique in a realistic scenario, the waveform generated in the described manner underwent initial testing

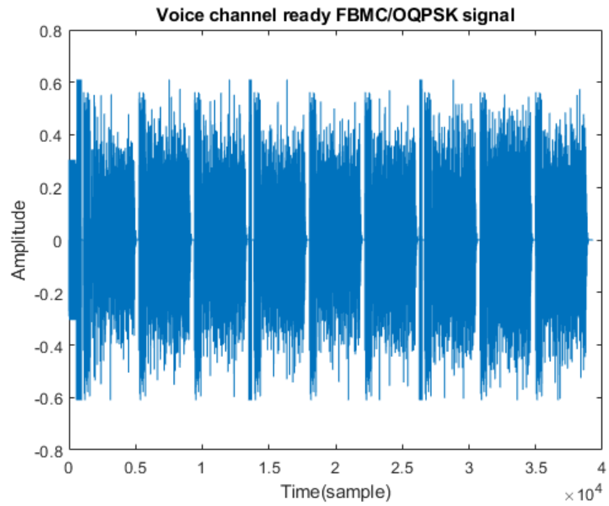

FIGURE 10. Waveform of FBMC/OQPSK signal prepared for transmitting over the real voice channel.

TABLE 1. Results using theoretical voice codec.

AMR-NB rate (kbps)	BER without FEC	BER with FEC
4.75	$1.91\cdot10^{-2}$	$1.89\cdot10^{-2}$
5.15	$1.86\cdot10^{-2}$	$1.84\cdot10^{-2}$
5.9	$1.47\cdot10^{-2}$	$1.43\cdot10^{-2}$
7.4	$7.29\cdot10^{-3}$	$6.56\cdot10^{-3}$
7.95	$8.31\cdot10^{-3}$	$7.65\cdot10^{-3}$
10.2	$0.85\cdot10^{-3}$	$0.66\cdot10^{-3}$
12.2	$0.63\cdot10^{-3}$	$0.49\cdot10^{-3}$

using theoretical AMR-NB simulation codecs. This approach aimed to simulate various compression rates typical of real-world mobile voice channels. The simulations were executed using MATLAB® software. As a real-world performance assessment, the signal was transmitted over three commonly used and publicly available voice channels: standard GSM 3G channel, Telegram VoIP channel and a Signal VoIP channel, serving as representatives of open-source applications.

The evaluation included measuring the BER across all available AMR-NB bit rates, considering both GSM and VoIP channels, with and without FEC coding. Furthermore, the quality of the transmitted voice signal was assessed using the PESQ metric across these mentioned voice channels. This comprehensive evaluation provides insights into how the FBMC/OQAM technique performs under varying conditions, offering a clear picture of its suitability for secure and reliable DoV transmission applications.

A. BER EVALUATION

The BER evaluation for the AMR-NB simulations, including scenarios with and without the implementation of Golay FEC code, is summarized in Table 1. It is evident that the BER experiences a notable increase at lower rates, aligning with expectations attributed to higher compression rates applied.

TABLE 2. Results using real voice channels.

Voice channel	BER without FEC		BER with FEC	
	median	mean	median	mean
3G-3G	$1.87 \cdot 10^{-3}$	$1.97 \cdot 10^{-3}$	$1.58 \cdot 10^{-3}$	$1.69 \cdot 10^{-3}$
Telegram	$5.41 \cdot 10^{-3}$	$5.39 \cdot 10^{-3}$	$4.59 \cdot 10^{-3}$	$4.59 \cdot 10^{-3}$
Signal	$6.41 \cdot 10^{-3}$	$6.49 \cdot 10^{-3}$	$5.52 \cdot 10^{-3}$	$5.62 \cdot 10^{-3}$

The results on the realistic voice channels are provided in Table 2, where the BERs are presented as median and mean statistics derived from 50 independent calls for each voice channel type. As a remark, the mean statistics closely resemble the median statistics, which implies the consistency of the signal quality across the respective channels.

The GSM 3G-3G voice channel has demonstrated highly satisfying BER performances, requiring no further extensive analysis. However, the noticeably higher BER observed with VoIP channels necessitated a deeper investigation into the received signal. By employing spectral and phase analysis techniques on the received symbols, it was determined that the majority of errors did not originate from mean phase shifts. This phenomenon, attributed largely to the utilization of OQPSK and phase synchronization pilot bits, underscored the robustness of the transmission in terms of mean phase stability. However, further analysis revealed a prominent issue of symbol dispersion, wherein the symbols experienced significant spreading or distortion. This dispersion was identified as the primary contributor to the observed BER increase. Figure 11 visually depicts this phenomenon by highlighting the erroneous symbols, providing a clear illustration of the dispersion effect on the received signal. Additionally, it is apparent that symbol amplitude was not preserved and exhibited notably fluctuations. Despite the potential importance in symbol mapping at the recipient's end, these fluctuation in symbol amplitude did not significantly impact the mapping process.

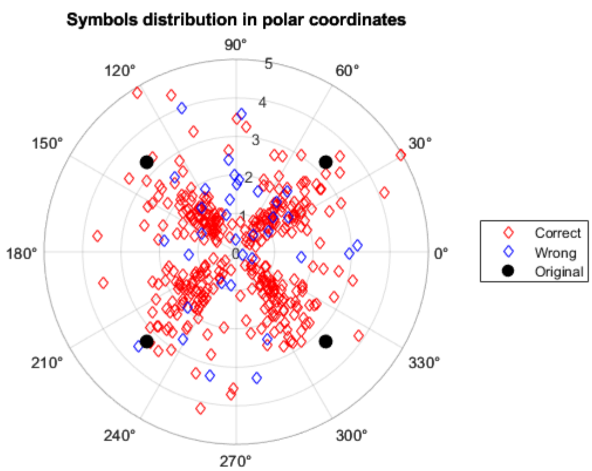

FIGURE 11. Symbols distribution in polar coordinates for Telegram signal as an example.

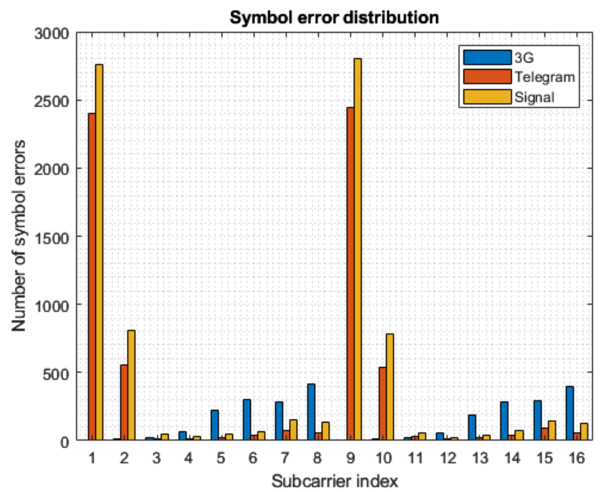

FIGURE 12. Symbol error distribution per subcarrier: comparison for 3G, Telegram and Signal voice channels.

Upon further expanded study of symbol errors, a significant correlation emerged between subcarrier frequency and the number of incorrectly mapped OQPSK symbols, indicating a potential underlying cause for the observed rise in BER. Figure 12 illustrates symbol error distribution per subcarrier for all three tested voice channels: 3G, Telegram and Signal. Analysis of the data revealed distinct patterns in error occurrence, particularly concerning specific subcarrier indices. In VoIP channels, namely Telegram and Signal, subcarriers with indices 1, 2, 9, and 10 were found to bear the brunt of symbol errors. This observation suggests a localized susceptibility to errors within these frequency ranges, potentially influenced by unique channel characteristics or environmental factors inherent to VoIP transmission. Similarly, on the 3G channel, a comparable effect was observed, albeit with different subcarrier indices. In this case, subcarriers 5, 6, 7, 8, 13, 14, 15, and 16 exhibited a higher prevalence of symbol errors. Furthermore, a discernible repetition pattern in error occurrence was noted, indicating specific and repetitive bandwidth limitations inherent to the voice channel.

The observed discrepancy in error distribution underscores the channel-specific nature of symbol error propagation and emphasizes the need for tailored optimization strategies for each communication channel. This finding suggests the presence of frequency-dependent non-linear behavior within the transmission environment, wherein certain subcarrier frequencies exhibit heightened susceptibility to symbol errors. To validate this hypothesis, adjustments were made to the frequencies of the subcarriers, deliberately avoiding those known to carry errors. Upon identifying a repetitive frequency pattern of increased symbol errors, certain frequencies were altered accordingly. The results of this optimization effort are presented in Table 3, detailing the BER for the adjusted subcarriers specifically tailored to VoIP channels, since those are more prone to mentioned effect.

TABLE 3. Results using real voice channels with changed subcarriers.

Voice channel	BER without FEC		BER with FEC	
	median	mean	median	mean
Telegram	$2.24 \cdot 10^{-3}$	$2.22 \cdot 10^{-3}$	$1.87 \cdot 10^{-3}$	$1.88 \cdot 10^{-3}$
Signal	$2.72 \cdot 10^{-3}$	$2.71 \cdot 10^{-3}$	$2.32 \cdot 10^{-3}$	$2.35 \cdot 10^{-3}$

By strategically manipulating subcarrier frequencies based on observed error patterns, significant improvements in BER were achieved, underscoring the effectiveness of targeted optimization strategies in mitigating frequency-dependent non-linear behavior and enhancing real-time secure voice communication reliability. These results expand the potential applications of the proposed technique for secure DoV transmission, offering promising avenues for further advancements in this field.

B. PESQ EVALUATION

Even though BER is a relevant evaluation metric for data transmission, in the context of speech, the information bits are generated by MELPe 1200 bps and synthesized back to speech after going through the voice channel. It is reasonable that the focus of the analysis should shift to the decoded speech intelligibility, rather than the number of incorrectly decoded bits solely. Moreover, there is a drawback of multiple possible speech signals originating from bit sequences with the same BER value, but different error patterns. Therefore, the quality of the decoded speech signals is evaluated using the PESQ metric. Although PESQ is considered obsolete and has been succeeded by POLQA (P.863), in our tested scenario where continuous time-varying delay is not a concern, PESQ remains sufficient for indicating the quality of the transmitted speech. PESQ algorithm compares two speech signals, taking the original, undistorted transmission signal as a reference and the decoded speech signal for comparing. PESQ algorithm outputs a score ranging from -0.5 for extremely degraded speech signal up to 4.5 for speech signal that perfectly resembles the original.

Another possible evaluation method would be a subjective speech quality assessment study, which would include volunteer listeners rating decoded speech signals on a given scale. This kind of analysis is time consuming and requires extensive planning and organizational resources and as such was out of scope for the purpose of this research.

The PESQ results for both theoretical AMR-NB channel simulation and real-world voice channels are presented in Table 4. The received speech signals are evaluated after MELPe 1200 bps synthesis and compared with the original voice signal recorded on the microphone of the transmitting device. For comparison purposes, the PESQ metric is also evaluated for the original undistorted signal that underwent MELPe 1200 bps compression only. The evaluated speech signals are assessed after using Golay correction code. PESQ code is implemented in MATLAB® as a C code wrapper.

TABLE 4. PESQ score results.

Voice channel	PESQ score
MELPe 1200 bps	2.96
Signal	2.59
Telegram	2.48
3G-3G	2.35
AMR-NB 12.2 kbps	2.68
AMR-NB 10.2 kbps	2.59
AMR-NB 7.95 kbps	1.35
AMR-NB 7.4 kbps	1.46
AMR-NB 5.9 kbps	0.85
AMR-NB 5.15 kbps	0.73
AMR-NB 4.75 kbps	0.70

The calculated PESQ scores correspond with the BER results and it is noticeable that even the MELPE 1200 bps signal without channel errors is highly compressed, resulting in a low PESQ score of 2.96. This is because MELPe 1200 bps is a low bit-rate voice compression vocoder designed primarily to generate as few bits as possible for use in secure communication systems, rather than prioritizing the preservation of speech characteristics not crucial for intelligibility, such as pitch, tone or intonation.

VI. CONCLUSION

In this paper, we proposed a new design for an encryption system utilizing a speech-like FBMC/OQPSK based waveform for transmitting encrypted digitized speech over the voice channel. This information carrying waveform was first tested in a MATLAB® simulations using AMR-NB codec of different bitrates and showed promising results, comparable to previously available similar work [9], but with larger spectral efficiency.

For real-time communication over publicly available voice channels, there was a need for additional measures in order to overcome numerous effects that distort the waveform, while respecting the required bit-rate constraint. These measures include time and phase synchronization, as well as the FEC Golay (12, 24) algorithm. The resulting robust signal was tested over the 3G, Signal and Telegram voice channels. The latter two showed frequency dependent behavior in the sense of repeatable erroneous symbol frequency distribution, which was successfully mitigated by adjusting the positions of information carriers in the frequency domain. This new approach was validated through a new series of tests and showed much better performance.

The results are given in terms of BER, which is valid in the context of pure DoV communication. In terms of its primary purpose of transmitting encrypted digitized speech, the evaluation focus shifted to decoded speech intelligibility metric – PESQ. FBMC/OQPSK signal showed consistent and promising results in both assessment criteria.

Future research endeavors will focus on expanding the testing scope to include a wider array of real-world voice channels. Additionally, exploration and implementation of a new modulation technique based on adapted subcarrier selection will be undertaken to further enhance performance and adaptability in diverse communication environments.

REFERENCES

[1] P. Krasnowski, J. Lebrun, and B. Martin, "Introducing a novel data over voice technique for secure voice communication," *Wireless Pers. Commun.*, vol. 124, no. 4, pp. 3077–3103, Jun. 2022.

[2] N. Katugampala, "Secure voice over GSM and other low bit rate systems," in *Proc. IEE Seminar Secur. GSM Beyond, End Secur. Mobile Commun.*, 2003, pp. 1–12.

[3] N. N. Katugampala, K. T. Al-Naimi, S. Villette, and A. M. Kondoz, "Real-time end-to-end secure voice communications over GSM voice channel," in *Proc. 13th Eur. Signal Process. Conf.*, Sep. 2005, pp. 1–4.

[4] M. A. Özkan and S. Berna Örs, "Data transmission via GSM voice channel for end to end security," in *Proc. IEEE 5th Int. Conf. Consum. Electron.*, Sep. 2015, pp. 378–382.

[5] M. Rashidi, A. Sayadiyan, and P. Mowlaee, "A harmonic approach to data transmission over GSM voice channel," in *Proc. 3rd Int. Conf. Inf. Commun. Technol.*, Apr. 2008, pp. 1–4.

[6] Z. Xu, "Data transmission method based on single carrier over GSM voice channel," *Revista De La Facultad De Ingeniera*, vol. 32, no. 9, pp. 23–29, Jan. 2017.

[7] T. Chmayssani and G. Baudoin, "Data transmission over voice dedicated channels using digital modulations," in *Proc. 18th Int. Conf. Radioelektronika*, 2008, pp. 1–10.

[8] B. Taleb Ali, G. Baudoin, and O. Venard, "Data transmission over mobile voice channel based on M-FSK modulation," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2013, pp. 4416–4421.

[9] L. Chen and Q. Guo, "An OFDM-based secure data communicating scheme in GSM voice channel," in *Proc. Int. Conf. Electron., Commun. Control (ICECC)*, Sep. 2011, pp. 723–726.

[10] P. Krasnowski, "Joint source-cryptographic-channel coding for real-time secure voice communications on voice channels," Ph.D. dissertation, Universite Cote d'Azur, Nice, France, 2021.

[11] P. Krasnowski, J. Lebrun, and B. Martin, "A novel distortion-tolerant speech encryption scheme for secure voice communication," *Speech Commun.*, vol. 143, pp. 57–72, Sep. 2022.

[12] *The 1200 and 2400 Bit/s NATO Interoperable Narrow Band Voice Coder*, Standard STANAG-4591.

[13] *Perceptual Evaluation of Speech Quality (PESQ): An Objective Method for End-to-End Speech Quality Assessment of Narrow-Band Telephone Networks and Speech Codecs*, ITU-T Standard P.862, 2001, p. 862.

[14] P. Siohan, C. Siclet, and N. Lacaille, "Analysis and design of OFDM/OQAM systems based on filterbank theory," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1170–1183, May 2002.

[15] A.-A. Husam and Z. Kollár, "Complexity comparison of filter bank multicarrier transmitter schemes," in *Proc. 11th Int. Symp. Commun. Syst., Netw. Digit. Signal Process. (CSNDSP)*, Jul. 2018, pp. 1–4.

[16] M. Bellanger et al., "FBMC physical layer: A primer," *PHYDYAS*, vol. 25, no. 4, pp. 7–10, Jan. 2010.

[17] S. Cubrilovic, D. Mandic, and A. Krstic, "Evaluation of improved classification of speech-like waveforms used for secure voice transmission," in *Proc. 21st Int. Symp. Infoteh-Jahorina*, Mar. 2022.

SARA ČUBRILOVIĆ (Member, IEEE) received the B.Sc. and M.Sc. degrees from the School of Electrical Engineering, University of Belgrade, Serbia, in 2017 and 2019, respectively, where she is currently pursuing the Ph.D. degree with the Signals and Systems Department. After graduating, she started her professional career as a Research and Development Engineer with the Vlatacom Institute of High Technologies. Her current research interests include speech and data encryption, artificial intelligence, and digital signal and speech processing.

ZVEZDANA KUZMANOVIĆ (Member, IEEE) received the B.Sc. degree from TU Wien, in 2017, and the M.Sc. degree from the Technical University of Munich, in 2020. She is currently pursuing the Ph.D. degree with the Signals and Systems Department, School of Electrical Engineering, University of Belgrade. After graduating, she started her professional career as a Research and Development Engineer with the Vlatacom Institute of High Technologies. Her current research interests include applied mathematics, data security, as well as speech and digital signal processing.

MARIJA PUNT received the B.Sc., M.Sc., and Ph.D. degrees in electrical and computer engineering from the School of Electrical Engineering, University of Belgrade, Serbia, in 2004, 2009, and 2015, respectively. She is currently an Associate Professor with the University of Belgrade. She teaches several courses on computer architecture and organization, web design, human–computer interaction, and multimedia. Her research interests include computer architecture, digital systems simulation, consumer electronics, data analysis, human–computer interaction, and multimedia.

DESIMIR VUČIĆ received the B.Sc. and M.Sc. degrees from the School of Electrical Engineering, University of Belgrade, Serbia, in 1975 and 1986, respectively, and the Ph.D. degree from the School of Technical Sciences, University of Novi Sad, Serbia, in 1998. From 1975 to 2006, he was with the Military Technical Institute, Belgrade. Since 2006, he has been an Associate Professor with the School of Computing, Union University, Belgrade. He is the author or co-author of two monographs, four scientific-technical publications, more than 120 articles published in international and national journals and conferences, and three technical solutions. His current research interests include digital signal processing, signal identification, and digital communications.

BRANKO KOVAČEVIĆ (Senior Member, IEEE) received the Ph.D. degree from the University of Belgrade, Belgrade, Serbia, in 1984. In 1981, he joined the Faculty of Electrical Engineering, University of Belgrade, where he is currently a Professor Emeritus. He is the author of eight books and more than 80 articles in scientific journals. His current research interests include robust estimation, system identification, adaptive and nonlinear filtering, optimal and adaptive control, and digital signal processing. He is a member of the EURASIP, the WSAES, and the National Association ETRAN, and a Corresponding Member of the Academy of Engineering Sciences of Serbia. He was a recipient of the Outstanding Research Prize of the Institute of Applied Mathematics and Electronics, the Prize of the Serbian Association for Informatics, and the Prize of the Association of Radio Systems Engineers. He is a Reviewer of IEEE Transactions, *Automatica* (IFAC), and *Signal Processing*.