Received June 23, 2019, accepted August 7, 2019, date of publication September 10, 2019, date of current version September 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2940640

The Adaptive Multi-Level Phase Coding Method in Audio Steganography

AHMED ABDULJABBAR ALSABHANY[1], FARIDA RIDZUAN[1,2], AND A. H. AZNI[1,2]

[1]Faculty of Science and Technology (FST), Universiti Sains Islam Malaysia (USIM), Nilai 71800, Malaysia
[2]CyberSecurity and Systems Research Unit, Islamic Science Institute (ISI), Universiti Sains Islam Malaysia (USIM), Nilai 71800, Malaysia

Corresponding authors: Ahmed Abduljabbar Alsabhany (ahmad88sabhany@gmail.com) and Farida Ridzuan (farida@usim.edu.my)

This work was supported in part by the Ministry of Higher Education (MOHE) Malaysia through the research under Grant USIM/FRGS/FST/055002/50117, in part by the Universiti Sains Islam Malaysia (USIM) and MOHE through the Graduate Financial Assistance Scheme (PhD), and in part by the Universiti Sains Islam Malaysia under the Partial Graduate Research Assistantship (GRA) Scheme.

ABSTRACT Audio steganography allows and inspires many researchers to design methods for secure communication. Based on the evaluation on the existing methods, it was found that most methods focused on one or two requirements while disregarding others, causing imbalanced performance. Moreover, most methods lack adaptivity and dynamic allocation. Therefore, in this research, a method called Adaptive Multi-level Phase Coding (AMPC) was proposed to optimize the above issues. The reverse logic of the main tradeoffs was used to empirically design several embedding levels that that simultaneously attained good performance for all aspects as much as possible. Then, an adaptive component was added by selecting the embedding level that provided the best performance for each embedding process. Moreover, the error spreading factor was introduced to achieve a fair payload distribution. The performance balance objective requires a new formulation that will enable the accurate selection of the degree of modification, multiple-bit embedding per modification, and reduced retrieval errors. As a result, the interval centering quantization (ICQ) was formulated and implemented in the proposed method. The experimental results show that AMPC successfully fulfilled the research objectives. Also, AMPC surpassed other phase coding methods in all aspects while time-domain methods achieved the highest transparency and capacity with the lowest robustness. Moreover, experiments show that the implementation of adaptive multi-level concept is able to improve the existing method's performance significantly. In summary, AMPC was able to achieve a stable embedding rate of 33 Kbps at 35 dB of SNR, which is higher than the recorded embedding rate of other phase coding methods.

INDEX TERMS Audio steganography, phase coding, adaptive multi-level, AMPC, LSB.

I. INTRODUCTION

Audio steganography is the process of hiding secret data inside an audio file. Early audio steganography methods exploited the Human Auditory System (HAS) to convey secret messages. However, more advanced statistical steganalysis approaches have been introduced recently [1], such as the methods in [2]–[4]. The main challenge in audio steganography is that three main requirements (embedding capacity, transparency, and robustness) must be fulfilled simultaneously [5], [6]. The embedding capacity often referred to as embedding rate, is defined as the maximum message size per 1-time unit. The second requirement is transparency, which indicates the ability to avoid suspicions and is usually related to the degree of similarity and error between the original cover signal and the stego signal. Due to the fundamental trade-off between capacity and transparency, many methods have been crafted to improve the capacity and transparency such as those proposed in [1], [7]–[11], mainly in the time or wavelet domains. The third requirement is robustness, which is defined as the ability of the method to withstand intentional and accidental signal attacks. This area is dominated by audio watermarking methods operating in the frequency domain such as those in the Discrete Fourier Transform (DFT) and the Discrete Cosine Transform (DCT). DFT methods can be divided into two categories. The first category operates by modifying the magnitude or the main spectrum components of the discrete Fourier transform [12]–[15].

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

This research focuses on the second category, which operates by modifying the phase components of the discrete Fourier transform [16]–[20]. Phase coding and its robustness potential have been demonstrated originally in [21]. Phase coding methods have achieved high robustness and variable transparency levels, leading to a low embedding capacity. In phase coding, the maximum embedding rate when transparency and robustness are measured is 24 kbps at 32 dB, achieved by the method proposed in [22]. However, even such an embedding rate is considered low when compared to the wavelet or time-domain based methods that lack the robustness advantage.

Most of the methods show a clear tendency to fulfill one or two requirements over the others. Consequently, some requirements will be fulfilled at the cost of others, causing performance imbalance. Although phase coding methods have a positive reputation in many review articles such as [5], [23] and [24], not much work has been done to uncover the full potential of this method as compared to the time or wavelet domains.

Moreover, most of the proposed methods are based on static embedding behavior that lacks adaptivity. Adaptive methods involve features that enable the adjustment of embedding routines to capitalize on performance, which gives the methods an advantage over static methods. In adaptive methods such as that of [25] and [26], the degree of modification depends on special criteria such as amplitude, coefficient value and energy level, while the signal length is not considered. Meanwhile, in the Matrix Embedding Strategy (MES) in [27], a high adaptivity level was achieved when the ratio between the message and the signal length was considered.

Another limitation in most existing methods is the sequential embedding approach, which causes error condensing and noise-and-quality discrepancies within parts of the stego audio itself. This limitation is most evident in cover underloading scenarios, which occurs when the message size is considerably lesser than the maximum capacity of the method for a certain cover file [28]. Sequential embedding in this scenario indicates that the embedded message in the first parts of the audio had produced a noise-and-quality gap between the clean and the embedded parts of the audio. As a consequence, these discrepancies could be picked up as a feature via steganalysis methods for message detection. The main reason for this issue is the unequal and non-dynamic message allocation over the cover. In this research, this issue is referred to as the lack of dynamic security.

Based on the limitations discussed earlier and the robustness advantage in phase coding, this study revisits audio steganography by phase coding to improve the capacity of this method while maintaining similar transparency and robustness levels. Moreover, adaptivity and dynamic security are also targeted. The rest of the paper is structured as follows; the related works are presented in Section II; the proposed method is presented in Section III; the experimental results are presented in Section IV; and the conclusion is presented in Section V.

II. RELATED WORKS

In this section, several audio steganography methods are reviewed including the phase coding methods followed by the time-domain methods. Finally, a brief summary of the reviewed methods is presented to conclude the section.

The main trade-off in phase coding exists between robustness and transparency, as highlighted in [29]. It is shown that high robustness demands a higher error rate; therefore reducing the SNR and vice versa. The method operates by dividing the phase range into multiple degrees where 0 and 1 are represented by every two consecutive degrees, and the nearest degree to match the message bit is selected. With consecutive embedding, specific degrees will be repeated over the embedding segment, which might be used to construct a pattern and therefore compromise the embedding. Hence, one of the main characteristics recommended for phase values is randomness [30]. Moreover, the embedding rate in this method is 333 bps, which is considered low. Another phase coding method proposed by Rivas [18] was based on interpolation. The embedding is carried out by taking the average of two neighboring phases (before and after). The modified phase is equal to the average value plus or minus a constant shift phase based on the value of the message bit. However, the highest embedding rate of this method was 172 bps, concluding that it suffers from low embedding capacity.

Parab *et al.* [19] also proposed a phase coding method based on the difference between phases. The method sets the difference as even or odd to hide 1 message bit of 0 or 1. However, the frame length in this method was selected as 10 ms, which also resulted in a low embedding rate. A key observation of this method is the high Bit Error Rate (BER), which reached 21% when 25 bins were modified per frame. In contrast, the methods proposed in [16], [17], [22] and [31] achieved high embedding rates. These methods used thresholding criteria that avoided embedding at low frequencies, while in the selected frequencies, explicit layers of the Least Significant Bit (LSB) in each phase were exploited to hide the secret message.

The BER is a metric used to calculate the correctness of embedding; it shows the percentage of the message bits that were retrieved incorrectly. In [31], the author demonstrated that embedding at low LSBs, such as the last two LSBs maintained lower robustness against Additive White Gaussian Noise (AWGN) than embedding at high LSBs such as the 5th and 6th LSBs. In [16] and [17], the BER was not measured at high embedding rates. However, in [22], the BER was around 0% against 10 dB of SNR after white Gaussian noise addition. The contradictory BER results in [19] and [22] combined reveal a research-worthy issue. In this context, the method in [19] can be expressed as phase coding by LSB substitution at the last bit only.

Therefore, this research further investigates whether a retrieval error in phase coding exists when LSB substitution is employed and the reasons that cause this error, if any. Some randomly selected cover signals were tested using the methods in [19] and [22]. Originally, the last 4 LSBs were utilized for embedding in [22]. However, for clarity, the method is implemented in this experiment modified one phase bin at the fourth LSB only, in a 4 ms frame (176 samples in 44.1 kHz audio). Moreover, the method uses an embedding threshold to select the phases. To get a better view of this comparison, another method in which one phase bin was modified at the third LSB without the embedding condition was implemented. On the other hand, the method in [19] was implemented by modifying one phase only in a 256-sample frame. The BER results of these methods are illustrated in Fig. 1.

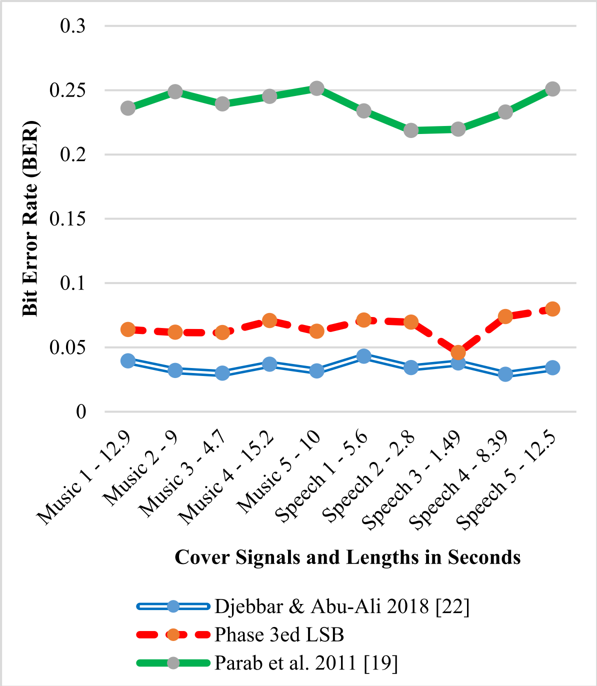


FIGURE 1. Phase coding using LSB substitution methods against BER.

Fig. 1 explicitly shows the presence of errors. Moreover, it shows that modifying deeper level of LSB yielded a lower error rate. The highest error rate occurred when the first LSB was used, fewer errors occurred at the third LSB, and the least was observed at the fourth LSB.

As for the time-domain methods, in the method proposed by Ahmed *et al.* [8], 8 LSBs of each selected audio sample of 16-bit length were replaced with message bits. The selection was carried out using an amplitude threshold. Meanwhile, Bazyar and Sudirman [7] proposed another time-domain method based on varying the payload per sample based on the first 2 Most Significant Bits (MSBs), and embedded 4, 5, 6, or 7 message bits per sample. Both methods achieved high embedding capacity. Other low bit encoding methods in the time-domain that aimed for high transparency were proposed in [32]–[35], but as a result of higher transparency, these methods achieved low embedding capacity.

In general, the robustness of the low bit encoding methods under the time-domain to withstand against signal processing attacks is poor [5], [6], [36]. Moreover, such methods follow a linear approach in embedding, indicating error condensing in the first parts of the signal.

In summary, three general issues could be observed in most existing audio steganography methods, namely:

- Imbalanced Performance
- Lack of adaptivity
- Lack of Dynamic Security

Furthermore, two main issues can be found in existing phase coding methods, namely:

- Low embedding capacity such as 172 bps and 333 bps in the methods of [18], [29], respectively.
- A high retrieval error rate in LSB-based methods [19], [22].

Thus, the objective of this research is to design a method to improve upon these limitations.

III. THE PROPOSED METHOD

The proposed method attempts to improve the general limitations in the existing methods by proposing an adaptive multilevel method. Meanwhile, the specific limitations in phase coding are addressed using a newly designed method of data injection called the Interval Centering Quantization (ICQ). The adaptive multilevel method uses the main trade-off logic in reverse order to design multiple embedding levels with various performances in capacity, transparency, and robustness. The trade-off itself states that capacity has an inverse relationship with both transparency and robustness, and those two have their own inverse relationship. Therefore, the inverse logic is that for low capacity, one of these cases should be obtained:

1. High transparency and low robustness,
2. High robustness and low transparency,
3. Moderate transparency and robustness.

Moreover, when the capacity is increased, the other aspects are affected negatively. In this method, the high and low rates of all three aspects were determined empirically. The adaptivity concept breaks down the maximum degree of modification when the capacity is low and spreads the error horizontally over the signal to achieve better performance in the other aspects. Hence, the three levels of capacity were initially defined. Then, other embedding parameters were tweaked to provide the best performance in terms of transparency and robustness. In the embedding operation, the method calculates the ratio between the message size and the signal length to select the embedding level that provides the best performance. The ratio between the message size and the signal length is denoted as Payload per Second (PPS) calculated as per Equation (1) while the level selection function was calculated as per Equation (2).

$$PPS = Ms/L \tag{1}$$

where Ms represents the message size and L represents the audio signal length in seconds.

$$F(PPS) = \begin{cases} Level 1 & Level\ 1_{Min} < PPS \leq Level\ 1_{Max} \\ \cdots & \cdots \\ Level\ K & Level\ K - 1_{max} < PPS \leq Level\ k_{Max} \\ No\ Selection & Level\ k_{Max} < PPS \end{cases} \quad (2)$$

where K is the number of embedding levels, Min is the minimum embedding rate and Max is the maximum embedding rate.

To solve the dynamic security issue, the message size was considered among the embedding keys and each message was distributed fairly on the cover signal to eliminate the formation of noise hotspots and quality difference. In the implementation, a dynamic parameter, namely the Error Spreading Factor (ESF), which represents the number of samples that are left as a gap between two modified frames, was calculated as in Equation (3).

$$ESF = \left\lfloor \frac{((L * Fs) - ini) - (\frac{Ms}{BPB} * FL)}{(\frac{Ms}{BPB})} \right\rfloor \quad (3)$$

where ini is the initial embedding point, BPB is the message bit number per block and FL is the frame length.

After the investigation, it was found that the retrieval errors in LSB-based phase coding were due to floating-point errors or rewriting errors. In most phase coding methods, the phase value is subjected to two transformations, which are the conversion from radians to degree format and the Inverse Fast Fourier Transform (IFFT), which includes rounding. On the other hand, rewriting errors occur when two or more DFT bins in the same frame are modified whereas the second modification corrupts or cancels the first one during the IFFT process.

To bypass floating-point errors, multiple bit embedding per quantization and an accurate selection of the degree of modification were enabled, as proposed in the ICQ. Meanwhile, only one phase was modified per frame to eliminate rewriting errors. The idea of this injection method is to encode a new value in the center of an interval that is more immune against floating-point errors and signal attacks, as it will be less likely to flip the value out of the interval. Moreover, in the retrieval process, any value within the interval reads the same value as the interval center. In this method, the first process involved calculating the modulus m as in Equation (4).

$$m = Value \bmod Q, \quad \forall\, Q \in \mathbb{N} \quad (4)$$

where $Value$ is the data unit value (which could be the phase value, coefficient value or sample value) and Q is the selected degree of modification. The result of this operation will always be in the range of 0 to Q-1. Then, this range was further divided into intervals, where the number of intervals, denoted as Noi, corresponds to the number of message bits to be embedded per modification, as calculated in Equation (5).

$$Noi = 2^n, \quad \forall\, n \in \mathbb{N} \quad (5)$$

where n is the number of message bits per modification. For example, to embed two message bits in one $Value$, the range of the modulus, m, is divided into 4 equal intervals, where each interval corresponds to one combination of the two bits (00, 01, 10, 11). The interval length is denoted as IL, and its condition is given in Equation (6).

$$IL = Q/Noi, \quad \forall\, IL > 2 \quad (6)$$

After acquiring the parameters m, Noi, and IL, the retrieval and embedding processes can begin. The general approach here is to change the value of the modulus to the value that represents the message combination. However, the key design feature here is that the modified value should be at the rough center of the interval. Equations (7) and (8) present the general retrieval and embedding rules, respectively.

$$R = \begin{cases} 0, & 0 \leq m < 1 * IL \\ 1, & 1 * IL \leq m < 2 * IL \\ \cdots & \cdots \\ Noi - 1 & (Noi - 1) * IL \leq m < Q \end{cases} \quad (7)$$

where R is the retrieved message segment of the data unit $Value$, in decimal representation.

$$NV = Value - m + Dec * IL + (IL - 1)/2 + c \quad (8)$$

where NV is the new value of the phase after quantization, Dec is the message segment in decimal to be embedded, and c is a security randomization factor ranging between $-1/10$ * IL to $1/10$ * IL to slightly move the new value off-center in order to deny the newly generated values the ability to form patterns.

An empirical study was conducted to set the embedding parameters of the three levels. The experiment started by setting three embedding rates to provide the best performance for each level within the boundaries of the main trade-off. The parameter setting of the levels is illustrated in Fig. 2.

The design of the embedding levels aims to find the best settings for all variable performances in terms of the aspects to achieve the performance balance. The relative performance of the levels is illustrated in Fig. 3.

After level selection, the ESF was calculated as per Equation (3) based on the unique parameter setting of each level and the embedding process starts.

In all levels, the default binary value of each frame was read using Equation (7), which is based on the Fast Fourier transform (FFT) process, radians to degree conversion and the modulus m calculated as per Equation (4). Then, the retrieved value was compared to the corresponding encrypted message bits. If a change is needed, the write step is initiated as in Equation (8), followed by converting the new value back to radians format and the IFFT process, respectively. Otherwise, the counter would be incremented by the ESF and FL, and the process moves to the next frame. The flow diagram of AMPC is illustrated in Fig. 4. In the following sub-sections, the embedding process in each level is explained.

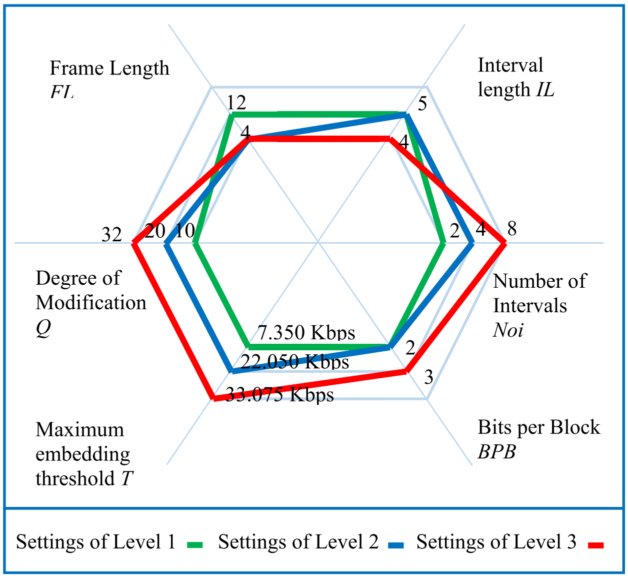


FIGURE 2. Embedding levels parameter settings.

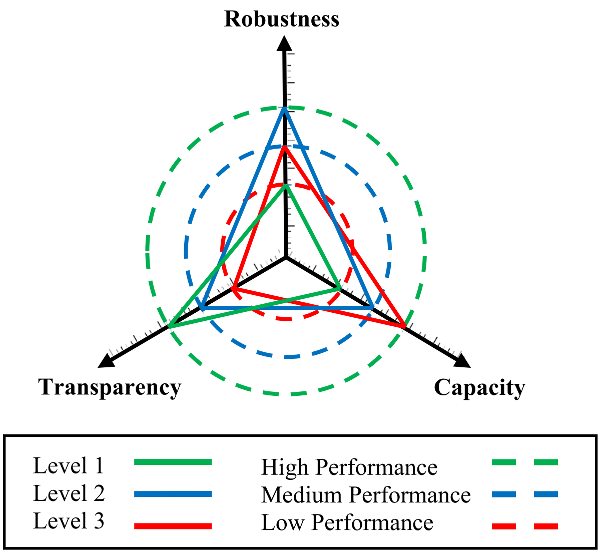


FIGURE 3. The intended performance of the embedding levels.

A. LEVEL 1

In Level 1, the first case of the MES is employed to boost transparency. In this level, the Frame Length used is 12 samples, which hides two message bits. However, each sample is further divided into three sub-blocks, where each sub-block is interpreted as one bit. After reading the value of each sub-block using Equation (7), three bits are achieved as output. Then, the MES is used to translate the three bits into two bits and compare it with the two message bits. If a change is needed, the MES selects the sub-block to be modified and the writing step is initiated as in Equation (8). In any case, only one sub-block will be changed to hide two message bits. The reading and writing processes in Level 1 are illustrated in Fig. 5(a).

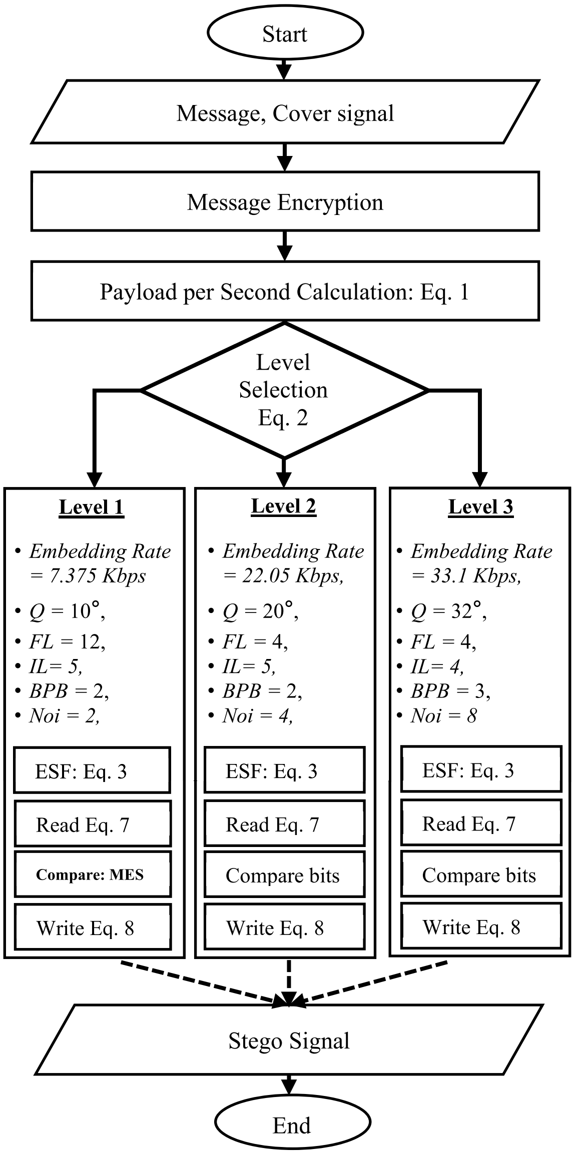


FIGURE 4. The flow diagram of the proposed method.

B. LEVEL 2

In level 2, a higher degree of modification of 20° is selected. In this level, the second phase bin of each frame of 4 samples is processed for modification to hide two bits. The range from 0 to Q-1 is divided into four intervals to represent the binary combination of the two bits stream, namely 00, 01, 10 and 11. After reading the value of the default phase using Equation (7), the result is then compared to the message bits. If a change is needed, the write step is initiated as in Equation (8). The reading and writing processes in Level 2 are illustrated in Fig. 5(b).

C. LEVEL 3

This level includes the highest degree of modification and the highest embedding rate. In this level, the second phase bin of each frame of four samples is processed for modification

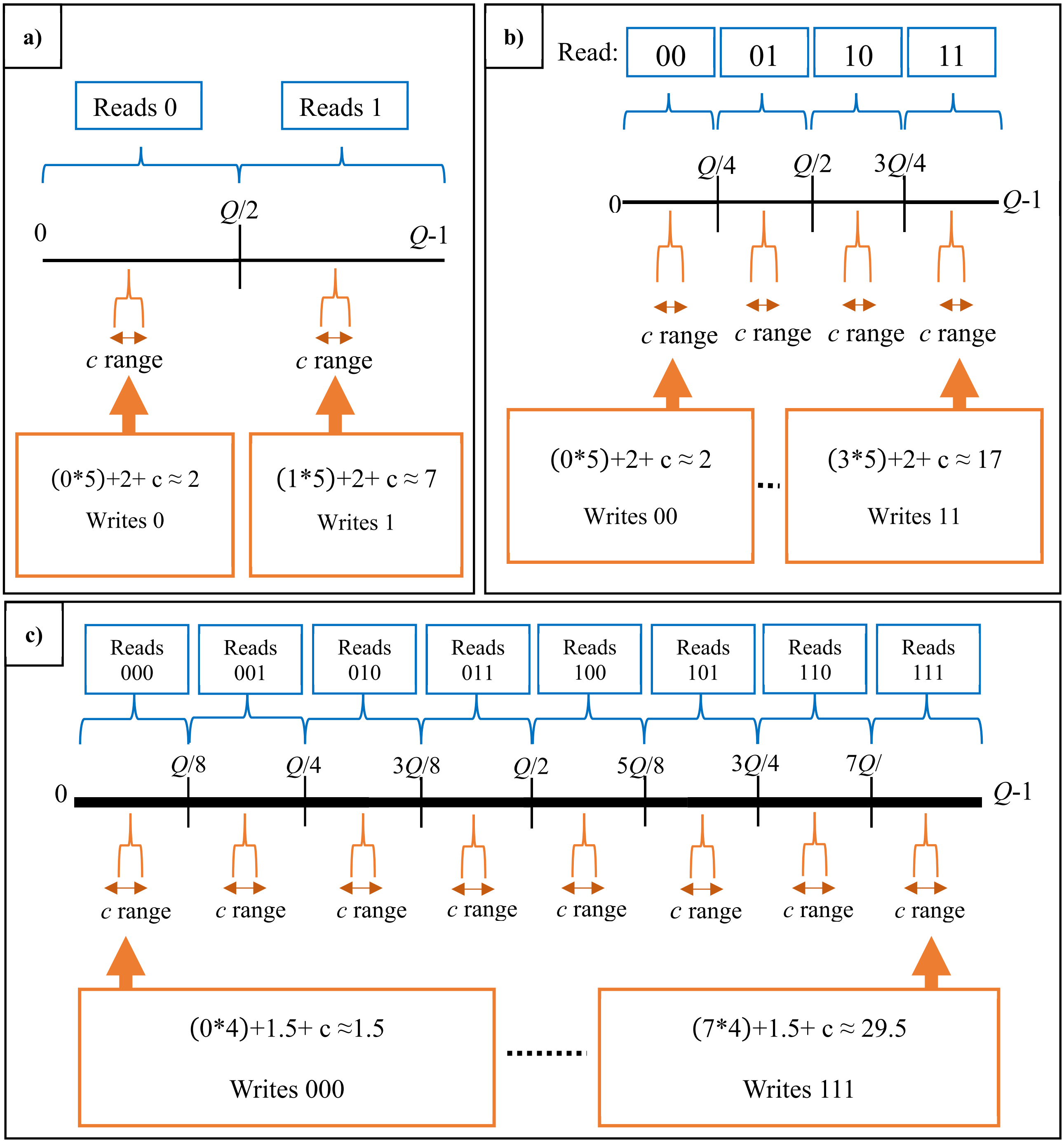


FIGURE 5. An illustration of the reading and writing processes using a) Level 1, b) Level 2, and c) Level 3.

to hide three bits. The range from 0 to Q-1 is divided into eight intervals to represent the binary combination of the 3-bit stream, namely 000, 001, 010, 011, 100, 101, 110 and 111. After reading the value of the default phase using Equation (7), the result is then compared to the message bits. If a change is needed, the writing step is initiated as in Equation (8). The reading and writing processes in Level 3 are illustrated in Fig. 5(c).

The retrieval process is done similarly without the writing steps where *BPB* is extracted from each block to assemble the message in binary form. Next, the message is decrypted and converted to its original form as in text, image or audio.

IV. EXPERIMENTAL RESULTS

The existing phase coding methods such as the ones proposed in [19] and [22] were included in the comparison study

TABLE 1. AMPC levels performance at maximum capacity.

Cover File		Level 1			Level 2			Level 3		
Name	**Length**	**Payload**	**SNR**	**BER**	**Payload**	**SNR**	**BER**	**Payload**	**SNR**	**BER**
Music 1	12.9	94919	49.95	1.93E-03	284759	39.68	3.00E-03	427138	35.73	4.63E-03
Music 2	9	66389	50.55	2.90E-03	199167	39.93	3.97E-03	298750	36.04	5.43E-03
Music 3	4.7	34561	46.18	3.86E-05	103685	36.19	1.90E-04	155527	32.17	1.63E-03
Music 4	15.2	111743	52.64	1.20E-03	335231	42.78	2.03E-03	502846	38.56	4.77E-03
Music 5	10	73499	51.07	5.00E-04	220499	40.67	6.67E-04	330748	36.96	1.83E-03
Speech 1	5.6	41385	44.98	8.67E-04	124159	35.43	5.00E-03	186238	31.44	9.13E-03
Speech 2	2.8	20821	45.83	6.33E-04	62463	35.79	8.77E-03	93694	31.76	1.26E-02
Speech 3	1.49	11007	49.49	1.36E-03	33023	40.12	8.50E-03	49534	35.53	1.31E-02
Speech 4	8.39	61727	48.01	4.93E-03	185181	37.74	9.70E-03	277771	33.99	1.06E-02
Speech 5	12.5	92501	47.81	9.33E-04	277503	37.95	9.40E-03	416254	33.69	1.57E-02
Min	1.49	11007	44.98	3.86E-05	33023	35.43	1.90E-04	49534	31.44	1.63E-03
Max	15.2	111743	52.64	4.93E-03	335231	42.78	9.70E-03	502846	38.56	1.57E-02
Mean	8.27	60941	48.68	1.53E-03	182827	38.71	5.09E-03	273850	34.59	7.93E-03

because of their performance. In [19], the frame length used was 10 ms and 25 bits were embedded per frame. Meanwhile, in [22], the frame length used was 4 ms and the embedding condition was used to select the phase bins for embedding. Then, embedding was carried out only in the 4 LSB of the phase for the best BER results. In addition to the phase coding methods, high capacity time-domain methods such as that of [7] and [8] were included to capture a better image of the proposed method's performance. In this experiment, the method proposed by [8] was implemented with 1024 embedding threshold. The Signal to Noise Ratio (SNR) was used to evaluate transparency, while BER against AWGN and Lossy Compression were used to evaluate robustness. Also, SNR was used to highlight the adaptivity concept against capacity ratios. The AMPC was compared in terms of visual error distribution and SegSNR spikes to capture the effect of the fair payload distribution against sequential embedding. Finally, the execution time was evaluated and compared. The setting of each experiment and the results are discussed in detail in the following sections.

A. AMPC LEVELS AT MAXIMUM CAPACITY

In the first experiment, the proposed method represented by the three embedding levels was evaluated. The SNR, maximum embedding capacity and BER of each level were compared. Table 1 shows the payload in bits, SNR in decibel (dB), and the BER in each level for the embedding of 10 randomly selected samples of both music and speech types.

In Table 1, the average payload for Level 3 was the highest, followed by Level 2 and Level 1 due to the different embedding rates in each level i.e. 7.35 Kbps for Level 1, 22.05 Kbps for Level 2, and 33.1 Kbps for Level 3. The shortest signal in this experiment was speech 3 of 1.5 seconds length, which achieved 11 kbps in Level 1 at 50 dB, 33 Kbps in Level 2 at 40 dB and almost 50 Kbps in Level 3 at 36 dB.

The results show that all the levels maintained a BER rate below 0.001%, indicating that the embedded messages were all retrieved correctly, which was not the case in many LSB-based phase coding methods as shown in Fig. 1. The difference between the proposed method and the methods in Fig. 1 is the ICQ and variable degrees of embedding. In the ICQ, new values are injected at the center of a virtual interval, which enables the new value to keep its hidden bit(s) even after a small change in its value. It is noticed that speech signals maintained slightly higher BER than music signals in all levels, due to the higher number of zero phases in speech signals that have silent intervals.

The SNR results for the levels show significant differences, mainly due to the variable payloads and the variable degree of the modification Q.

B. SNR COMPARISON EXPERIMENT

An SNR comparison experiment with equal payloads was conducted to capture the performance differences between the levels and other related methods. In this experiment, three randomly selected audio signals were used to hide equal payloads where the payloads were selected based on the capacity of the phase coding methods. If the payload was higher than the method's capacity, a value of 1 would be used to represent an embedding failure case as in Fig. 6.

Moreover, in this experiment, each embedding level was evaluated individually for comparison purposes. Fig. 6(a), Fig. 6(b) and Fig. 6(c) show the SNR results of the included

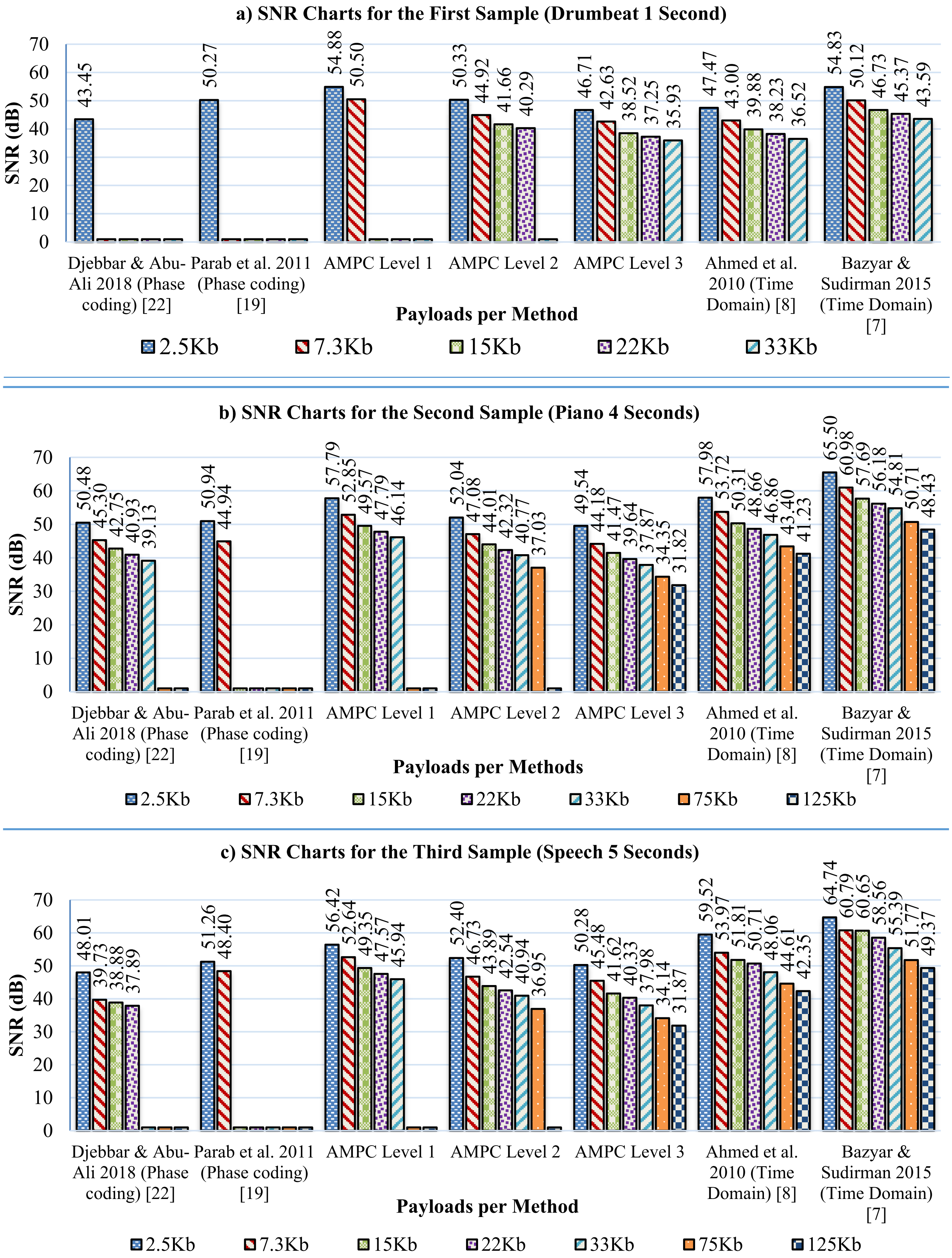


FIGURE 6. SNR comparison against equal payloads of the proposed method, existing phase coding methods, and existing time-domain methods using a) 1 second drumbeat, b) 4 seconds piano, and c) 5 seconds of human speech.

methods using the three cover samples, where a higher SNR indicates a better level of transparency.

In Fig. 6(a), the cover signal is the sound of a 1-second duration (44100 samples) drumbeat, which was selected to show the capacity potential in the one-time unit. For the first payloads, the first AMPC level achieved the highest SNR followed by the time-domain method in [7]. AMPC Level 1 maintained the lowest degree of modification where $Q = 10$. Moreover, in the first level only, the first MES case was used to take advantage of the unutilized capacity to boost transparency. For example, for the embedding of two bits, only one phase only was modified, while the ICQ hid 1 bit per sub-block (4 samples).

Based on Fig. 6, it is observed that for most of the payloads, the time-domain methods, which are [7] and [8] achieved the highest SNR, followed by the selected AMPC level and the existing phase coding methods, which are [19] and [22]. The selected AMPC level is the level that provides the highest SNR for each scenario. For example, the payload 33 Kb in the second cover signal was achieved using all AMPC levels. However, in the actual practice of AMPC, the first level will always be selected for cover signals of similar length, which is realized by the PPS calculation in Equation (1) and the level selection in Equation (2). Moreover, the highest recorded result in phase coding was 24 Kbps at 32 dB of SNR [22]. The results in Table 1 and Fig. 6 show that the third AMPC level was able to stably provide 33 Kbps at almost 35 dB of SNR. Therefore, it is concluded that the proposed method provides balanced performance in terms of capacity and transparency, as it achieved better embedding rates and better SNR for the same covers and payloads. Moreover, the proposed method provided a median performance between the phase coding and time-domain methods.

C. BER COMPARISON AGAINST AWGN AND LOSSY COMPRESSION

A BER comparison against two signal attacks was carried out to capture the robustness of the proposed method. In this experiment, each method embeds a random message in three audio covers, where the payload is dependent on the method itself. Then, in the AWGN test, the stego signals are subjected to various levels of noise. Next, the retrieval code of each method is used to recover as much as possible of the noisy signal. Meanwhile, in the Lossy compression attack, the stego signals are compressed to MPEG-4 format using variable bit rates and converted back to WAV format before attempting to retrieve the message. In both tests, a lower BER will indicate more robust embedding and a higher BER will indicate that the attack destroyed the embedded message.

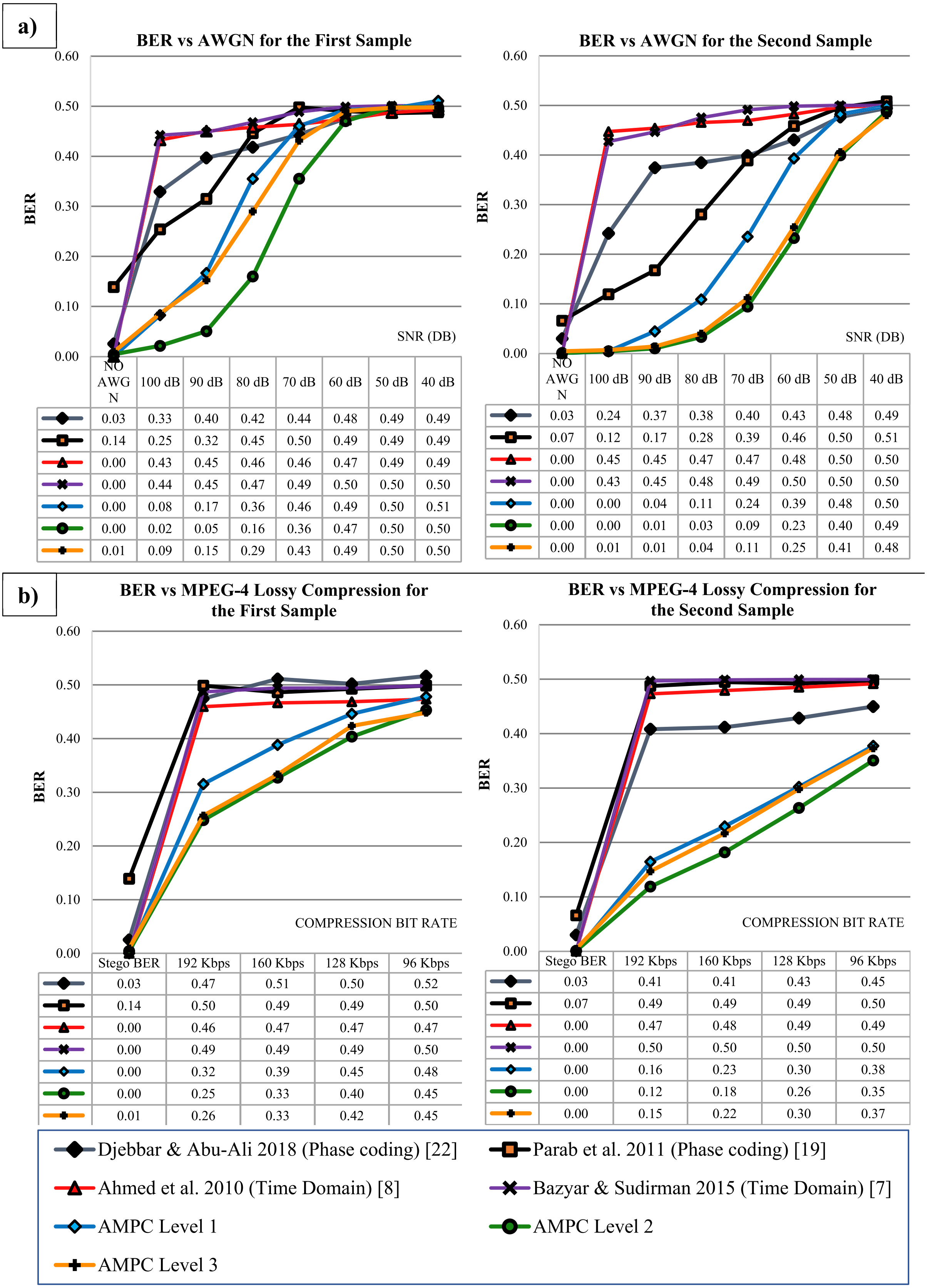
1) AWGN TEST

Fig. 7(a) shows the BER of the methods after adding variable levels of white Gaussian noise to the stego signals. The main highlight of the results in both signals is that AMPC showed better resistance than the other methods. In more detail, in the time-domain methods, the BER jumped from 0% to around 45% at the first noise level, indicating total destruction of the embedded message. In contrast, existing phase coding methods showed better resistance to noise addition, which is justified by the concept of phase coding itself, where each DFT point is extracted after including a number of time-domain samples (FFT frame length). Therefore, the stored value in such a DFT point is in a way repeated in variable rates over the frame elements. The frame length sustains a positive relationship with the noise resistance and robustness in general as shown in [29], where a low error rate was achieved even at 30 dB of noise. Another factor that affects the noise resistance is the degree of modification, which can be represented by the maximum potential error after each embedding. Hence, in [19], the frame length was more than the frame length of [22], while in [22], a higher degree of modification was attained than in [19], which caused an intersection in the study's BER charts. The advantage of AMPC in this test is its ICQ design, where a margin of error (interval half) is left to ensure correct retrieval even after the value is changed by errors, noise or other similar attacks. Specifically, Level 2 achieved the best BER, followed by Level 3 and Level 1. To justify these results, it is important to consider three factors: the degree of modification Q, the Interval Length IL and the MES in the first level. Although the degree of modification in the Level 3 was 32°, and 20° in Level 1 and 2, the second level achieved better BER, which is explained by the IL difference, which is equal to 5 in Level 1 and 2, and 4 in Level 3.

In all levels, the modified phase value was made at the center of the interval. Therefore, in Level 1 and 2, the modified phase value will maintain its embedded bits until it is changed by 2.5°. Due to the contiguous nature of the intervals, when the phase is modified by 2.5°, the modulus m will point at an adjacent interval and hence create a retrieval error. Meanwhile, in Level 3, the modified phase needs to be changed by only 2° to point to another interval. Therefore, Level 2 achieved better robustness. The MES in Level 1 required that all three sub-blocks carry the correct value to retrieve the two-bit stream correctly, which explains the lower robustness of this level.

2) MPEG-4 LOSSY COMPRESSION TEST

This test is more destructive than AWGN. The main trend of the results of this test is similar to that of the AWGN experiment where the proposed method dominated and yielded similar results for the order of the level. Similarly, Level 2 achieved the best BER, followed by Level 3 and Level 1. In both samples, most of the selected methods jumped to the total loss rate after a compression rate of 192. However, both [22] and the proposed method showed better robustness in the second sample, as the first sample had a one-second length. The AMPC achieved better results than other methods because of its ICQ injection method, which leaves a margin of error for the modified values. Although AMPC came first in this test, the Lossy Compression still had a very destructive effect, which is more evident in the first sample.



BER vs AWGN for the First Sample

	NO AWGN	100 dB	90 dB	80 dB	70 dB	60 dB	50 dB	40 dB
◆	0.03	0.33	0.40	0.42	0.44	0.48	0.49	0.49
■	0.14	0.25	0.32	0.45	0.50	0.49	0.49	0.49
▲	0.00	0.43	0.45	0.46	0.46	0.47	0.49	0.49
✖	0.00	0.44	0.45	0.47	0.49	0.50	0.50	0.50
◇	0.00	0.08	0.17	0.36	0.46	0.49	0.50	0.51
●	0.00	0.02	0.05	0.16	0.36	0.47	0.50	0.50
✚	0.01	0.09	0.15	0.29	0.43	0.49	0.50	0.50

BER vs AWGN for the Second Sample

	NO AWGN	100 dB	90 dB	80 dB	70 dB	60 dB	50 dB	40 dB
◆	0.03	0.24	0.37	0.38	0.40	0.43	0.48	0.49
■	0.07	0.12	0.17	0.28	0.39	0.46	0.50	0.51
▲	0.00	0.45	0.45	0.47	0.47	0.48	0.50	0.50
✖	0.00	0.43	0.45	0.48	0.49	0.50	0.50	0.50
◇	0.00	0.00	0.04	0.11	0.24	0.39	0.48	0.50
●	0.00	0.00	0.01	0.03	0.09	0.23	0.40	0.49
✚	0.00	0.01	0.01	0.04	0.11	0.25	0.41	0.48

BER vs MPEG-4 Lossy Compression for the First Sample

	Stego BER	192 Kbps	160 Kbps	128 Kbps	96 Kbps
◆	0.03	0.47	0.51	0.50	0.52
■	0.14	0.50	0.49	0.49	0.50
▲	0.00	0.46	0.47	0.47	0.47
✖	0.00	0.49	0.49	0.49	0.50
◇	0.00	0.32	0.39	0.45	0.48
●	0.00	0.25	0.33	0.40	0.45
✚	0.01	0.26	0.33	0.42	0.45

BER vs MPEG-4 Lossy Compression for the Second Sample

	Stego BER	192 Kbps	160 Kbps	128 Kbps	96 Kbps
◆	0.03	0.41	0.41	0.43	0.45
■	0.07	0.49	0.49	0.49	0.50
▲	0.00	0.47	0.48	0.49	0.49
✖	0.00	0.50	0.50	0.50	0.50
◇	0.00	0.16	0.23	0.30	0.38
●	0.00	0.12	0.18	0.26	0.35
✚	0.00	0.15	0.22	0.30	0.37

FIGURE 7. **BER against A) AWGN levels and B) MPEG-4 Lossy Compression for 2 samples.**

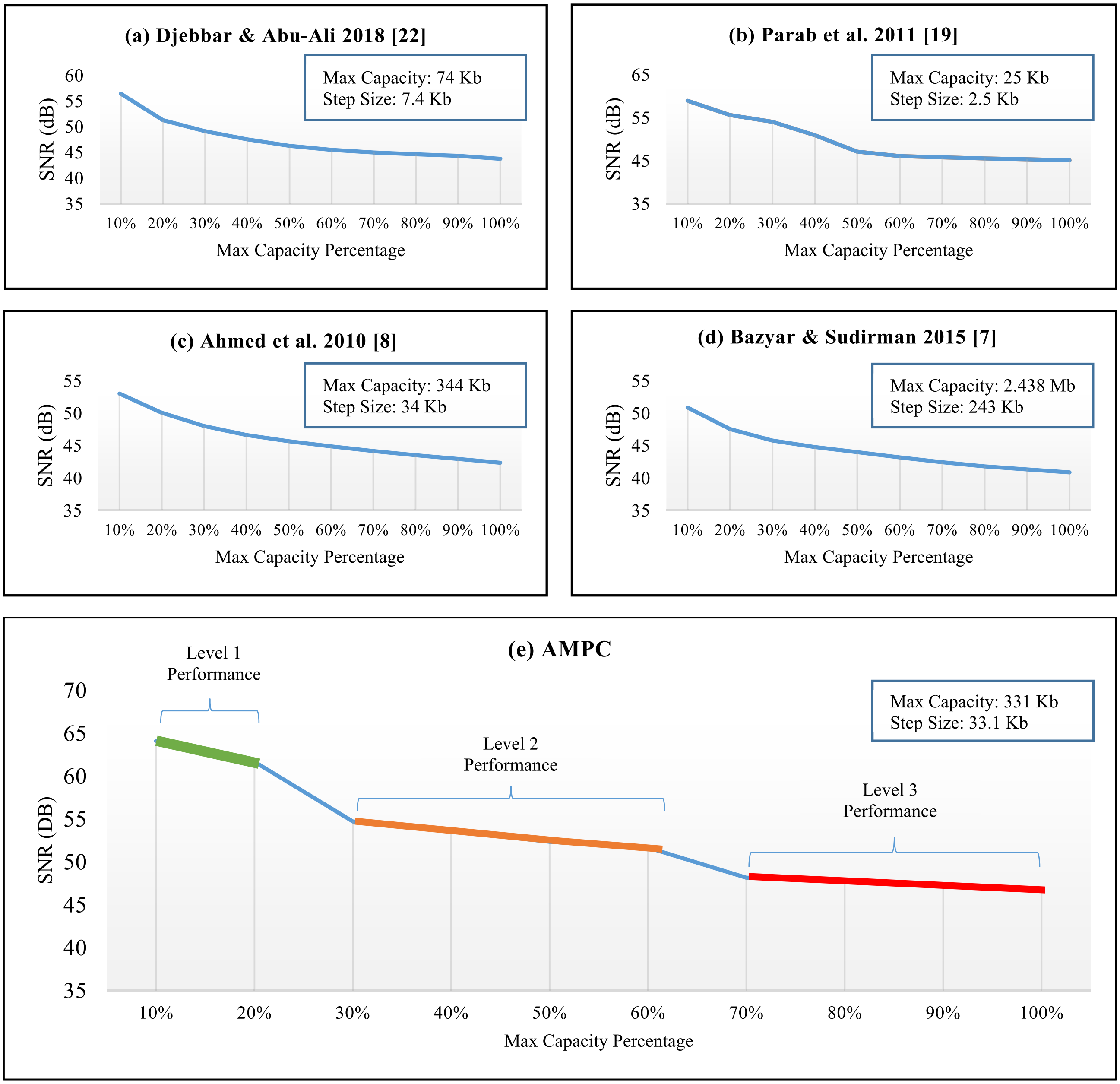


FIGURE 8. SNR performance of the methods against incremental payloads.

In summary, both tests showed an advantage of using ICQ as an injection method over other methods, namely yielding better robustness. Based on the results of the experiments, the proposed AMPC achieved a better balance over other existing methods, coming in second in terms of capacity and transparency after the time-domain methods and first in terms of robustness. Moreover, a new record in phase coding was achieved where AMPC recorded a stable embedding rate of 33 Kbps at approximately 35 dB of SNR.

D. THE EFFECT OF ADAPTIVE MULTI-LEVEL CONCEPT ON THE PERFORMANCE

This experiment highlights the effect of the adaptive concept on overall performance. In this experiment, SNR was calculated incrementally from 10% to 100% of the maximum capacity of each method. After determining the maximum capacity of each method for the selected cover signal, the payload scale was set up with a 10% increment. Fig. 8 shows the SNR rates for the selected methods against low to high capacity. The main highlight in Fig. 8 is the linear performance in both existing phase coding and time-domain methods. On the other hand, the AMPC chart in Fig 8 shows clear shifts in performance, which are caused by the change in levels. The multi-level concept is designed to capture variable capacity, transparency and robustness by adjusting the degree of modification and the other variables adaptively. As a result, when the message is smaller than the maximum capacity of the cover, a lower degree of modification is invoked to achieve optimal performance. Similarly, when the message is large or almost equal to the maximum capacity, a high degree of modification is invoked to enable the embedding. For example, Level 1 achieved high SNR, but it had low capacity. Level 3 produced a high capacity but with the same embedding settings, it achieved lower SNR at low capacities. The maximum capacity will be restricted if the method uses a low degree of modification, and vice versa. However, in both cases of high and low capacities, any single level method will lose on performance.

To understand the adaptivity potential of the multiple-level method, another experiment was carried out where the method of [37] was selected as an example for its simplicity. The method originally modifies 4 LSBs per sample in a linear fashion. In the modified version, three levels were established. The first level modifies 1 LSB bit, the second level modifies 3 LSBs and the third level modifies 4 LSBs as per the original. Therefore, Level 1 produced a total embedding rate of 44.1 Kbps, Level 2 produced 132.3 Kbps and Level 3 produced 176.4 Kbps. The capacity of the modified version was the same, and the robustness was approximately similar. Fig. 9 shows the difference between the three-level version of the method and the original one. In Fig. 9, the adaptive version achieved better rates in 7 out of 10 payloads. In more detail, around (12 to 5) dBs of SNR was saved by introducing Level 1 and Level 2, respectively. Moreover, such an improvement in performance was accomplished without any cost in the trade-off among the requirements. Therefore, it is concluded that the adaptive multi-level concept achieved better performance by breaking down the degree of modification to three levels and utilizing horizontal space.

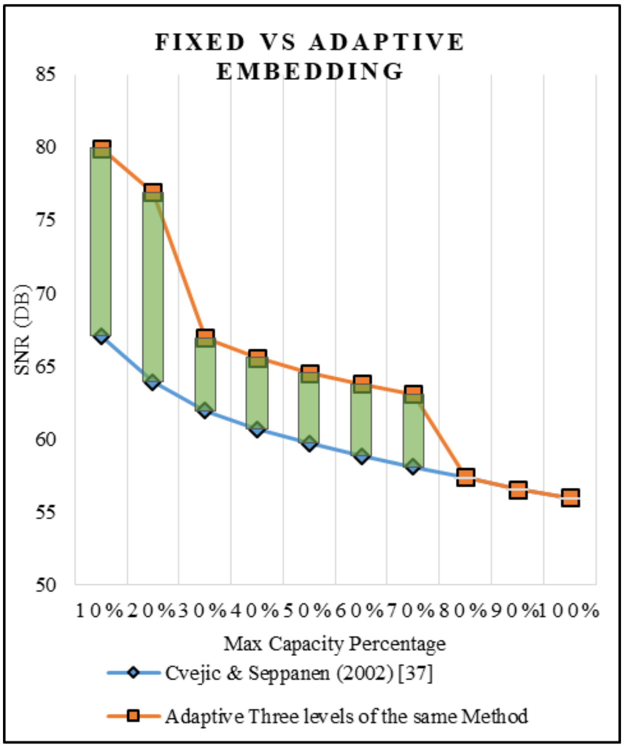


FIGURE 9. SNR charts of the fixed embedding method and the adaptive multi-level.

E. DYNAMIC SECURITY EVALUATION

Two experiments were used to capture the effect of message allocation on the security of embedding in cover under-loading scenarios. The real threat that dynamic security attempts to minimize is when the steganalysis method can differentiate embedded and clear signals uncertainly. In this case, the distribution density of errors can be used to assist the main steganalyzer.

The visual error distribution and the SegSNR spike captured the message allocation over the cover signals. However, in each experiment, different settings were used.

1) VISUAL ERROR DISTRIBUTION

The goal of this experiment was to visually capture the error concentration in cover under-loading scenarios. In this experiment, the capacity ratio was set to 50%, where each method embedded 50% of its maximum capacity on a unified cover. Then, the error was calculated after embedding by subtracting the stego signals of the original cover. A more dispersed error indicates lower error density and therefore better transparency. Fig. 10 shows the error distribution among the methods at 50% capacity. Most of the existing audio steganography method yielded similar results as presented in

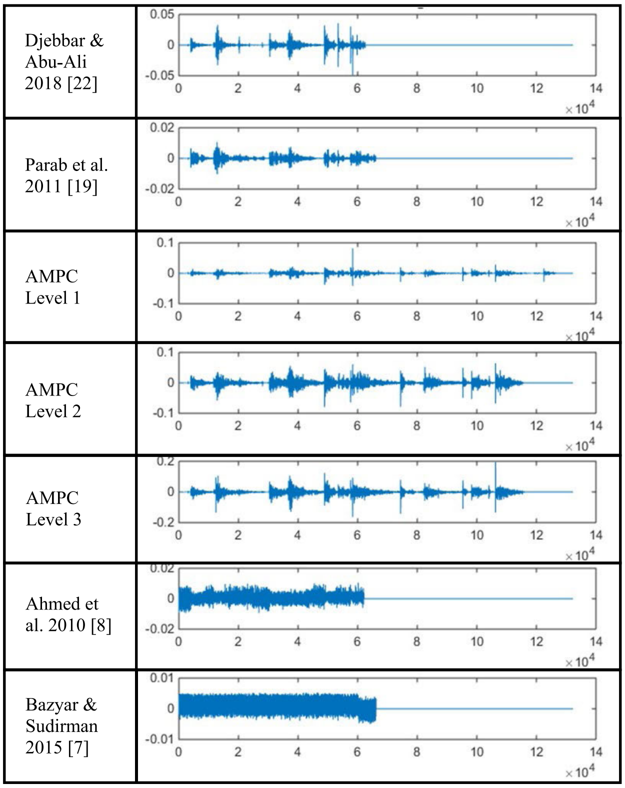


FIGURE 10. Error distribution over the length of the cover signal at 50% of maximum capacity for each method.

this experiment because they followed a sequential embedding approach that visits audio data units one-by-one for embedding. Moreover, even selective-based methods, which claim an un-sequential embedding, failed to choose dynamic criteria to spread the message over the cover. Fig. 10 shows that in all existing methods, the embedding ratio determined the cover signal ratio to be used for embedding. On the other hand, AMPC applied a fair distribution of error on the signal regardless of the embedding ratio, which is a result of the *ESF*. The main disadvantage of error condensing methods is the ability of steganalysis methods to distinguish the stego signal from the clean and embedded parts, which could compromise not only the message presence, but even the approximate size of the hidden message too.

2) SegSNR SPIKE

In this experiment, the goal was to observe the location of the spike between the embedding and the clean parts. Segmental SNR was used to show the effect of error distribution on the quality and noise difference in stego signals. In this experiment, the cover and stego signals were divided into equal frame sizes with a 50% overlap, and the SNR of each frame was calculated between the cover and the stego. Then, the resulting ordered SNRs were saved into a vector. Next, the vector was divided into distinct groups. A finer level of analysis requires a larger number of groups. However, in this research, frame sizes were set to 256 with a 50% overlap and 10 groups. Finally, the average of each group was calculated to represent the SNR in that segment. Two capacity ratios were used, which are 20% and 35% of the maximum capacity of each method for the selected cover. Fig. 11(a) and Fig. 11(b) show the SegSNR for the included methods at 20% and 35% of the maximum capacity of each method.

The results in Fig. 11 are similar to those in Fig. 10 in which both charts showed a dependence of the error distribution on the embedding ratio in the existing time-domain and phase coding methods. For the existing methods in Fig. 11(a),

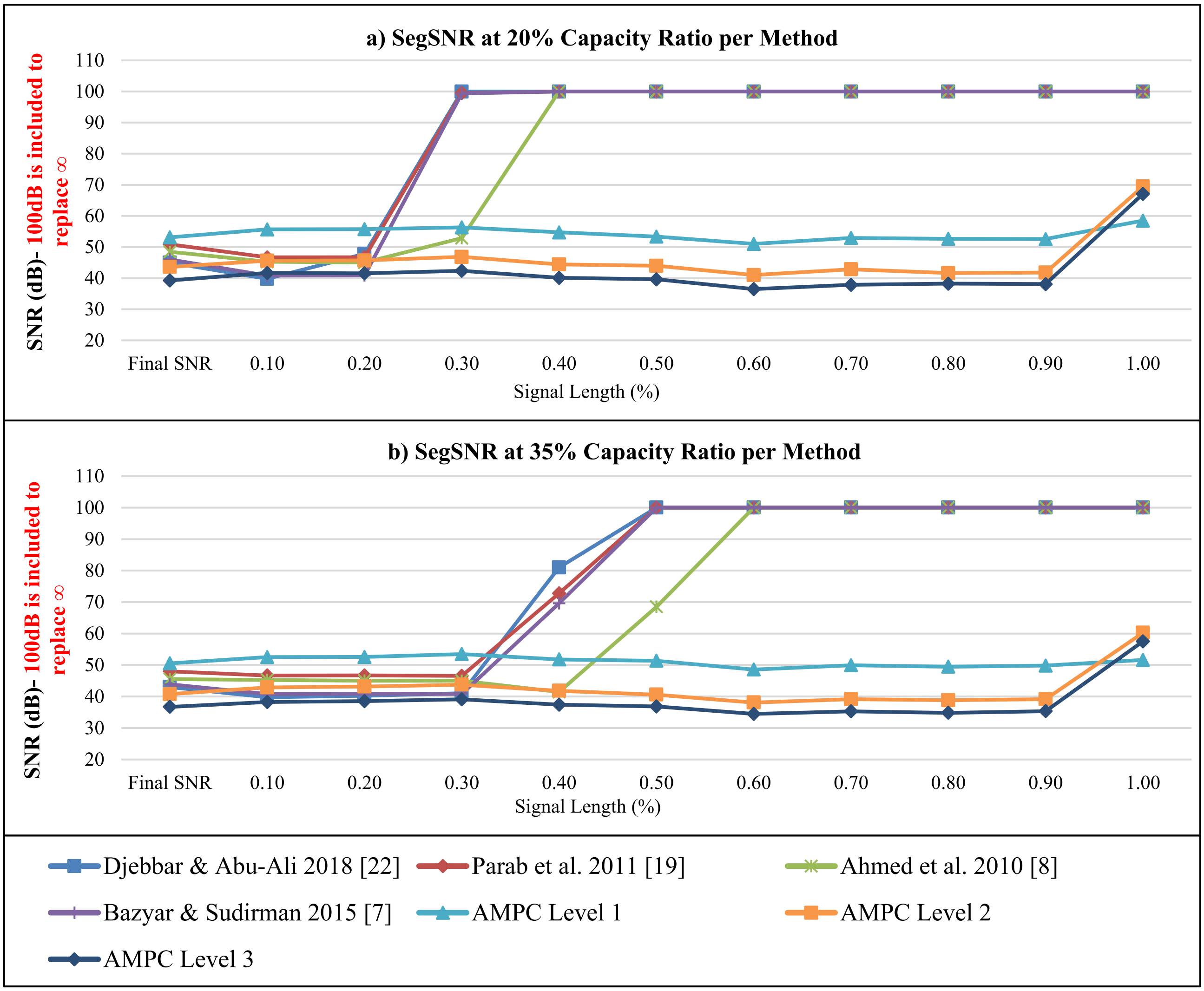
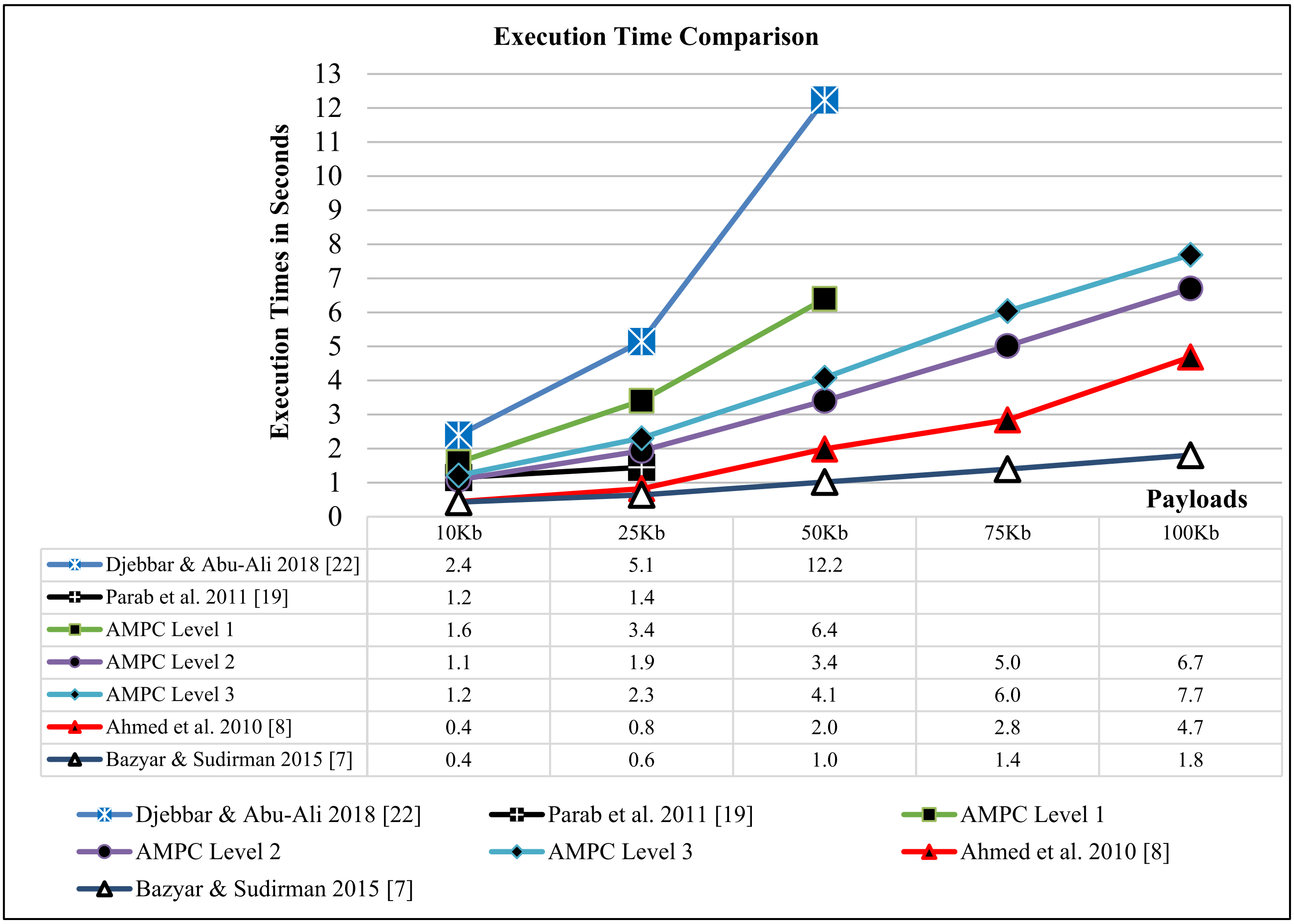


FIGURE 11. SegSNR Spike at A) 20% and B) 35% embedding ratio per method.



	10Kb	25Kb	50Kb	75Kb	100Kb
Djebbar & Abu-Ali 2018 [22]	2.4	5.1	12.2		
Parab et al. 2011 [19]	1.2	1.4			
AMPC Level 1	1.6	3.4	6.4		
AMPC Level 2	1.1	1.9	3.4	5.0	6.7
AMPC Level 3	1.2	2.3	4.1	6.0	7.7
Ahmed et al. 2010 [8]	0.4	0.8	2.0	2.8	4.7
Bazyar & Sudirman 2015 [7]	0.4	0.6	1.0	1.4	1.8

FIGURE 12. Execution time comparison.

the chart showed an early spike at 20% of the signal length, while in Fig.11 (b), the spike moved to 35% of the signal length to follow the embedding ratio. On the other hand, in AMPC which enabled *ESF*, the location of the spike was not affected by the capacity ratio. The fair error distribution eliminated the chances of error hotspots formation in underloading scenarios.

It is noticed that in [8], a late spike was attained due to the threshold in place. In general, AMPC provides a better association between the message and the cover in under-loading scenarios because the error was well distributed dynamically. As a result of such a modification, the message size must be known at the receiver to retrieve the message. Therefore, the message size must be shared beforehand or can be shared in the stego signal by allocating a segment of the stego signal for message size sharing as in [38].

F. EXECUTION TIME COMPARISON

This experiment compared the execution time of the AMPC to the related methods. An audio signal of 10-second length was selected alongside 5 different payloads, where the payloads were selected based on the capacity of the included methods. The experiments were conducted using an Intel Core i5-4590 workstation. Fig. 12 shows the comparison results. Due to the capacity differences, some methods were not able to embed all payloads. Based on the results, it is fair to report that time-domain methods are generally faster than phase coding methods due to the multiple calls of FFT, IFFT and the Unwrap functions during phase coding run times. The method in [19] was the fastest of the phase coding methods and the method in [22] achieved the highest execution time. While the AMPC showed moderate execution time.

Based on this experiment, the main limitation of AMPC could be its design complexity, as it includes three levels that require unique parameter settings and implementations.

V. CONCLUSION

In this research, three general issues were discussed, namely the imbalanced performance, a lack of adaptivity and a lack of dynamic security. Moreover, it is noticed that existing phase coding methods suffer from low capacity and high retrieval error rates in LSB-based methods. Thus, several solutions were formulated in the proposed Adaptive Multi-Level Phase Coding (AMPC) method. A balanced performance was targeted by designing multiple embedding levels with variable performances of each aspect, where the parameter settings were carried out empirically. The adaptivity solution completed the previous solution by creating a ladder of degree-of-modification and enabling better utilization of horizontal space. Moreover, the fair payload distribution provided the solution to achieve better dynamic security. One of the main highlights in this research is the proposed method of data injection, called the Interval Centering Quantization (ICQ). The ICQ was formulated to enable an accurate selection of the degree of modification, increase the number of bits per modification and reduce the chance of retrieval errors due to floating-point errors.

The comparative results show that AMPC achieved the highest robustness levels against AWGN and Lossy Compression signal attacks. Also, AMPC achieved better transparency and capacity when compared with existing phase coding methods. Therefore, it can be concluded that AMPC obtained better performance balance than the existing methods. Also, AMPC was able to achieve an embedding rate of 33 Kbps at 35 dB of SNR, exceeding the existing record in phase coding of 24 Kbps at 32 dB of SNR. Moreover, AMPC maintained a high level of adaptivity that significantly improved its performance. Furthermore, the fair payload distribution showed a better distribution of error than sequential embedding methods in both visual error distribution and SegSNR tests.

REFERENCES

[1] D. M. Ballesteros L and J. M. Moreno A, "Highly transparent steganography model of speech signals using efficient wavelet masking," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9141–9149, 2012.

[2] Q. Liu, A. H. Sung, and M. Qiao, "Temporal derivative-based spectrum and Mel-Cepstrum audio steganalysis," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 3, pp. 359–368, Sep. 2009.

[3] W. Zeng, H. Ai, and R. Hu, "A novel steganalysis algorithm of phase coding in audio signal," in *Proc. 6th Int. Conf. Adv. Lang. Process. Web Inf. Technol. (ALPIT)*, Aug. 2007, pp. 261–264.

[4] F. Djebbar and B. Ayad, "A new steganalysis method to detect information hiding in speech," in *Proc. 13th Int. Wireless Commun. Mob. Comput. Conf. (IWCMC)*, Jun. 2017, pp. 1879–1884.

[5] F. Djebbar, B. Ayad, K. A. Meraim, and H. Hamam, "Comparative study of digital audio steganography techniques," *EURASIP J. Audio, Speech, Music Process.*, vol. 2012, Oct. 2012, Art. no. 25.

[6] F. Djebbar, B. Ayad, H. Hamam, and K. Abed-Meraim, "A view on latest audio steganography techniques," in *Proc. Int. Conf. Innov. Inf. Technol.*, Apr. 2011, pp. 409–414.

[7] M. Bazyar and R. Sudirman, "A new method to increase the capacity of audio steganography based on the LSB algorithm," *J. Technol.*, vol. 74, no. 6, pp. 49–53, Apr. 2015.

[8] M. A. Ahmed, M. L. M. Kiah, B. B. Zaidan, and A. A. Zaidan, "A novel embedding method to increase capacity and robustness of low-bit encoding audio steganography technique using noise gate software logic algorithm," *J. Appl. Sci.*, vol. 10, no. 1, pp. 59–64, 2010.

[9] A. H. Ali, L. E. George, A. A. Zaidan, and M. R. Mokhtar, "High capacity, transparent and secure audio steganography model based on fractal coding and chaotic map in temporal domain," *Multimed. Tools Appl.*, vol. 77, no. 23, pp. 31487–31516, Dec. 2018.

[10] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 3, pp. 920–935, Sep. 2011.

[11] S. Ahani, S. Ghaemmaghami, and Z. J. Wang, "A sparse representation-based wavelet domain speech steganography method," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 23, no. 1, pp. 80–91, Jan. 2015.

[12] K. Gopalan, "A unified audio and image steganography by spectrum modification," in *Proc. IEEE Int. Conf. Ind. Technol.*, Feb. 2009, pp. 1–5.

[13] K. Gopalan, "Audio steganography by modification of cepstrum at a pair of frequencies," in *Proc. 9th Int. Conf. Signal Process. (ICSP)*, Oct. 2008, pp. 2178–2181.

[14] M. Nathan, N. Parab, and K. T. Talele, "Audio steganography using spectrum manipulation," in *Proc. Technol. Syst. Manage.*, vol. 145, Jan. 2011, pp. 152–159.

[15] H. B. Dieu and N. X. Huy, "An improved technique for hiding data in audio," in *Proc. 4th Int. Conf. Digit. Inf. Commun. Technol. Appl. (DICTAP)*, May 2014, pp. 149–153.

[16] F. Djebbar and B. Ayad, "Audio steganograpgy by phase modification," in *Proc. 8th Int. Conf. Emerg. Secur. Inf., Syst. Technol.*, 2014, pp. 31–35.

[17] F. Djebbar, B. Ayad, K. Abed-Meraim, and H. Hamam, "Unified phase and magnitude speech spectra data hiding algorithm," *Secur. Commun. Netw.*, vol. 6, no. 8, pp. 961–971, Aug. 2013.

[18] E. Rivas, "Fourier phase domain steganography: Phase bin encoding via interpolation," *Proc. SPIE*, vol. 6579, May 2007, Art. no. 65790W.

[19] N. Parab, M. Nathan, and K. T. Talele, "Audio steganography using differential phase encoding," in *Proc. Technol. Syst. Manag.*, 2011, pp. 146–151.

[20] S. K. Moon and R. D. Raut, "Application of data hiding in audio-video using anti forensics technique for authentication and data security," in *Proc. IEEE Int. Advance Comput. Conf. (IACC)*, Feb. 2014, pp. 1110–1115.

[21] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, nos. 3–4, pp. 313–336, 1996.

[22] F. Djebbar and N. Abu-Ali, "Lightweight noise resilient steganography scheme for Internet of Things," in *Proc. IEEE Glob. Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–6.

[23] B. B. Zaidan, A. A. Zaidan, H. A. Karim, and N. N. Ahmad, "A new digital watermarking evaluation and benchmarking methodology using an external group of evaluators and multi-criteria analysis based on 'large-scale data'," *Softw. Pract. Exp.*, vol. 47, no. 10, pp. 1365–1392, Oct. 2017.

[24] K. U. Singh, "A survey on audio steganography approaches," *Int. J. Comput. Appl.*, vol. 95, no. 14, pp. 7–14, Jan. 2014.

[25] M. Kaur and M. Juneja, "A new LSB embedding for 24-bit pixel using Multi-Layered bitwise XOR," in *Proc. Int. Conf. Inventive Comput. Technol. (ICICT)*, vol. 2, Aug. 2016, pp. 1–5.

[26] A. Delforouzi and M. Pooyan, "Adaptive digital audio steganography based on integer wavelet transform," *Circuits, Syst. Signal Process.*, vol. 27, no. 2, pp. 247–259, Apr. 2008.

[27] A. Westfeld, "F5—A steganographic algorithm: High capacity despite better steganalysis," in *Proc. 4th Int. Work. Inf. Hiding*, Apr. 2001, pp. 289–302.

[28] A. Kaur, M. K. Dutta, K. M. Soni, and N. Taneja, "Localized & self adaptive audio watermarking algorithm in the wavelet domain," *J. Inf. Secur. Appl.*, vol. 33, pp. 1–15, Apr. 2017.

[29] X. Dong, M. F. Bocko, and Z. Ignjatovic, "Data hiding via phase manipulation of audio signals," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, May 2016, pp. 377–380.

[30] S. W. Smith, "Fourier transform properties," in *The Scientist and Engineer's Guide to Digital Signal Processing*. San Diego, CA, USA: California Technical Pub., 1997, ch. 10, pp. 185–208.

[31] F. Djebbar, H. Hamamy, K. Abed-Meraimz, and D. Guerchix, "Controlled distortion for high capacity data-in-speech spectrum steganography," in *Proc. 6th Int. Conf. Intell. Inf. Hiding Multimedia Signal Process. (IIHMSP)*, Oct. 2010, pp. 212–215.

[32] H. N. Xuan and D. H. Ba, "An efficient method for hiding data in audio," in *Proc. Int. Conf. Adv. Technol. Commun. (ATC)*, Oct. 2015, pp. 167–171.

[33] E.-T. B. Abdelsatir, N. C. Debnath, and H. Abushama, "A multilayered scheme for transparent audio data hiding," in *Proc. IEEE/ACS 12th Int. Conf. Comput. Syst. Appl. (AICCSA)*, Nov. 2015, pp. 1–6.

[34] H. Kumar and A. Taluja, "Enhanced LSB technique for audio steganography," in *Proc. 3rd Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT)*, Jul. 2012, pp. 1–4.

[35] T.-S. Wu, H.-Y. Lin, W.-C. Hu, and Y.-S. Chen, "Audio watermarking scheme with dynamic adjustment in mute period," *Expert Syst. Appl.*, vol. 38, no. 6, pp. 6787–6792, Jun. 2011.

[36] M. L. M. Kiah, B. B. Zaidan, A. A. Zaidan, A. M. Ahmed, and S. H. Al-bakri, "A review of audio based steganography and digital watermarking," *Int. J. Phys. Sci.*, vol. 6, no. 16, pp. 3837–3850, 2011.

[37] N. Cvejic and T. Seppanen, "Increasing the capacity of LSB-based audio steganography," in *Proc. IEEE Work. Multimed. Signal Process. (MMSP)*, Dec. 2002, pp. 336–338.

[38] A. A. Alsabhany, F. Ridzuan, and A. H. Azni, "A hybrid method for data communication using encrypted audio steganography," *Adv. Sci. Lett.*, vol. 23, no. 5, pp. 4896–4900, Nov. 2017.

AHMED ABDULJABBAR ALSABHANY received the bachelor's degree in computer science from Yarmouk University, Jordan, and the M.Sc. degree in computer science from Al al-Bayt University (AABU), Jordan, in 2015. He is currently pursuing the Ph.D. degree with Universiti Sains Islam Malaysia (USIM), Malaysia. His research interests include information security, information hiding, audio steganography, and audio steganalysis. He has published a number of articles in these areas in journals and conferences.

FARIDA RIDZUAN received the degree (Hons.) in computer science (majoring in industrial modeling and computing) from the Universiti Teknologi Malaysia (UTM), the M.Sc. degree in discrete mathematics and its applications from the University of Essex, U.K., and the Ph.D. degree from Curtin University, Australia. She is currently a Senior Lecturer with the Information Security and Assurance (ISA) Programme, Faculty of Science and Technology (FST), Universiti Sains Islam Malaysia (USIM), Malaysia. She is currently carrying out research on information security specifically on audio and text steganography, encryption, and web spam. She has published numerous academic articles in international journals related with computer security. She also has a very large interest in research related to problems modeling, algorithms, optimization, and ontology. She has won several innovation awards.

A. H. AZNI received the bachelor's degree in computer information systems from Bradley University, IL, USA, in 1998, the M.Sc. degree in digital communication from Monash University at Clayton, Australia, in 2002, and the Ph.D. degree in computer science (wireless security) from the Universiti Teknikal Malaysia Melaka (UTeM), in 2014. From May 2003 to May 2007, she was with the Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak (UNIMAS). She joined Universiti Sains Islam Malaysia (USIM), in May 2007, and has been appointed as the Deputy Dean of the Center for Graduate Studies. She has many experiences in presenting research talks and articles at national and international conferences. She has also published tremendous articles in highly esteemed journals. Her research interests include wireless security, the IoTs, and cryptography.