

# QSAC: Quantum-assisted Secure Audio Communication using Quantum Entanglement, Audio Steganography, and Classical Encryption

Md. Raisul Islam Rifat<sup>a,c</sup>, Md. Mizanur Rahman<sup>b,c</sup>, Md. Abdul Kader Nayon<sup>a,c</sup>, Md Shawmoon Azad<sup>d</sup> and M.R.C. Mahdy<sup>d</sup>

<sup>a</sup>Department of Electrical and Electronic Engineering, Chittagong University of Engineering and Technology, Raozan, Chattogram, 4349, Bangladesh

<sup>b</sup>Department of Computer Science and Engineering, Rajshahi University of Engineering and Technology, Station Rd, 6204, Rajshahi, Bangladesh

<sup>c</sup>Mahdy Research Academy, Bashundhara R/A, Dhaka, 1229, Bangladesh

<sup>d</sup>Department of Electrical and Computer Engineering, North South University, Bashundhara R/A, Dhaka, 1229, Bangladesh

---

## ARTICLE INFO

### Keywords:

Quantum Key Distribution (QKD)  
E91 protocol  
Entanglement  
CHSH inequality  
Steganography  
ChaCha20-Poly1305  
Hashing  
Secure Hash Algorithm (SHA)  
Encryption  
Decryption

---

## ABSTRACT

The emergence of quantum computing possesses a significant security threat to the current classical system, since it has the potential to outperform the current classical computer in some specific task due to their unique principle of operation. This necessitates finding a method that is resistive to quantum computers to securely transfer information. This research addresses these challenges by proposing a novel method combining quantum key distribution (QKD), specifically the E91 protocol with the classical encryption-authentication protocol, i.e. ChaCha20-Poly1305 and concealing information within another message through steganography to securely transfer audio messages. A shared secret key is created between the communicating parties using E91 QKD, which exploits the stellar protection of quantum entanglement against eavesdropping. The shared key is hashed through the SHA-3 hash function to generate a fixed-length, high-entropy symmetric key. The audio message is hidden inside another audio signal through steganography. The steganographic signal is encrypted using ChaCha20-Poly1305 AEAD in order to provide another layer of obfuscation as well as a means to verify integrity. Through rigorous experiment, we validated the robustness of the proposed methodology in both classical and quantum attacks. The processing of secret audio signals of varying duration (00:01:32 to 00:01:36) exhibits consistent encryption results. The encrypted stego audios show high randomness, with an average entropy of 15.9984209099406865, an average UACI of 49.997671320111918, and an average NSCR of 99.99848969721785%. We demonstrated the safety of the shared key using the CHSH inequality test, where in the presence of an eavesdropper the CHSH value is very less than  $2\sqrt{2}$ . In addition, the integrity of the secret audios is also validated through the verification of the authentication tag generated during the encryption process. Our research offers a novel framework for secure audio transmission, combining classical encryption and authentication methods with QKD to enhance confidentiality, integrity, and resilience against eavesdropping, ensuring robust end-to-end security.

---

## 1. Introduction

The simplest, as well as the most important, form of exchange for human beings is verbal communications. The invention of the Telephone by Alexander Graham Bell in 1876 transformed the verbal communication demography - making the transmission of human voice, which are essentially audio signals, over long distances possible. In the subsequent centuries, technological improvements has broadened the scope of audio transmission. With the rise of the Internet, the amount of information exchanged through audio signals increased rapidly. This increase number audio transmission resulted in the development of different encryption techniques and cryptographic algorithms. The most popular among these algorithms are the symmetric AES and the asymmetric RSA encryption schemes. AES [1] is a block cipher scheme that utilizes multiple matrix operation to encrypt and decrypt data using the same key. Without

---

\*Corresponding author

\*\*Principal corresponding author

 jkk@example.in (Md.R.I. Rifat); wjh@example.org (Md.A.K.N.); t.rafeeq@example.in (M.R.C. Mahdy)

 www.jkkrishnan.in (Md.R.I. Rifat); <https://www.university.org> (Md.A.K.N.); www.campus.in (M.R.C. Mahdy)

ORCID(s):

any knowledge of the cipher key, it is computationally infeasible to reverse the operations performed during encryption. RSA [2], on the other hand, utilizes prime number factorization to generate public-private key-pairs for each user that are used to encrypt and decrypt data. The security of the RSA scheme relies on the insane amount of computational power required to obtain the private key from a public through prime number factoring.

However, with the advent of Quantum Computers with increasing number of functional qubits, these classical cryptography and encryption schemes face an imposing threat [3]. In 1996, Lov K. Grover proposed an algorithm to search an unordered database of size  $N$  using  $\sqrt{N}$  quantum queries [4]. Using Grover's algorithm, the number of trials required to brute-force a key of length  $k$  reduces from  $2^k$  to  $2^{k/2}$ . This reduction in number of brute-force trials effectively reduces the security level of symmetric encryption schemes such as AES [5]. For example, the AES-128 encryption scheme with a pre-quantum security level of 128 reduces to a post-quantum security level of 64, which is much easier to brute-force. In 1994, P.W. Shor presented a quantum algorithm that can quickly find the prime factorization of any positive integer  $N$  [6]. As the security of the RSA algorithm relies on the arduousness of prime number factorization to derive private key from public key, it is currently facing an existential threat due to the exponential speed of Shor's algorithm. It is estimated that the time complexity for Shor's algorithm is  $\mathcal{O}(72(\log(N))^3)$ , as opposed to  $\mathcal{O}(N^3)$  for classical computers.

### 1.1. Current Challenges

To address these arising challenges, new research is being done on the field of quantum augmented communication systems. These systems exploit the principles of quantum mechanics to attain secure data transmission. One of the most promising area of research in the field of quantum communications is Quantum Key Distribution (QKD). QKD protocols work by establishing a secure cryptographic key between two users over an insecure channel [7]. QKD employ properties unique to quantum mechanics, such as the no-cloning theorem [8] and the uncertainty principle [9], that ensures the detection of any eavesdropping attempts and thereby guarantees the security of the key. However, QKD itself does not provide security on its own, rather it facilitates the establishment and secure exchange of secret keys that are subsequently used by other cryptographic algorithms and encryption techniques to secure the transmitted information. In addition, the applicability of QKD is restricted by a few practical challenges such as slow key generation, vulnerability to noise, limited operational distance etc. In resemblance, the research being done on the field of steganography heralds the emergence of an effective information concealing technique to obscure sensitive information within seemingly harmless transmission. Steganography is the technique of hiding secret message within another message in such a manner that it is not discernible that a secret message is embedded [10]. However, the security of any steganographic technique is heavily dependent on the strength of cryptographic technique employed to encrypt the data [11]. In summary

- QKD facilitates secure sharing of encryption keys, but secure data transmission requires classical encryption systems.
- In spite of providing improved security, the real-life implementation of QKD is confined due to some practical challenges such as noise vulnerability, slow key generation etc.
- The security of steganography is dependent on the employed encryption scheme. If the encryption scheme is compromised, data security is at risk.
- Hybrid systems combining classical and quantum cryptography methods defend against different types of attacks but it complicates the system design and implementation.

### 1.2. Proposed Solution

The new vulnerabilities in secure data transmission due to quantum computer and the need to oppose them have been the motivation for our research. With the advancement in quantum computing technology, it is imperative to devise innovative cryptographic techniques that can keep the data secure in the prospect of an attack from these powerful computers. It is our aim to develop a vigorous solution for secure data communication by combining the strength of quantum communication with hash function, classical encryption and steganography. Hence, the objective of our experimental work is to investigate the viability and credibility of integrating encrypted steganographic techniques with quantum protocols, specifically QKD, to strengthen the security and resilience of audio communication. This innovative amalgamation of steganography and quantum communication embodies a significant furtherance in the field, providing an exceptional and optimistic approach to secure data transmission.

This paper introduces a novel approach to reinforce the security of digital audio communication by combining the competency of QKD, classical encryption schemes and steganography. A high-level outline of our proposed solution is as follows –

- **Secure Key Generation and Distribution:** Utilizing the E91 QKD protocol proposed by Artur K. Ekert in 1991 [12], a shared cryptographic key with enhanced security against quantum threats is generated.
- **Key Derivation:** Employing the Secure Hashing Algorithm-3 (SHA3-256) [13], the QKD-generated key is hashed to produce a fixed-length, high-entropy key for symmetric encryption.
- **Steganography:** Using Least Significant Bit (LSB) substitution steganography, an audio signal is hidden inside another audio signal [14].
- **Symmetric Encryption with Authentication:** The ChaCha20-Poly1305 AEAD (Authenticated Encryption with Associated Data) algorithm [15] - which combines the stream cipher scheme, ChaCha20 [16] with the message authentication code, Poly1305 [17] - the steganographic audio is encrypted. The encrypted audio is bundled with a message authentication tag, which is used for authentication and integrity verification.

The incorporation of QKD with classical symmetric encryption addresses various security concernment, providing protection against both classical and quantum threats.

### 1.3. Paper Organization

Our paper is organized as follows - Section 2 demonstrates the present state of the field through the review of existing research and their development, laying the foundation for our proposed scheme. Section 3 describes the foundational concepts of our proposed schemes. Section 4 describes the proposed methodology, presenting a detailed piecemeal explanation of our proposed architecture. Section 5 delves into the detailed description of our implementation. In section 6, we present the analysis methods as well as the findings. Section 7 explores deeper into an extensive analysis of the results, construing the findings and excerpting key insights. Finally, section 8 concludes the paper by indicating the direction of future work, highlighting the potential application and extension of our research.

## 2. Related Works

Post-Quantum cryptography (PQC), also known as Quantum resistant Cryptography, is designed to stave off attacks by both classical and quantum computers. In other words, to design a dependable and future-proof data communication system, a blending of Classical Cryptography and Quantum Cryptography is imperative. One such type of blending could be combining some form of symmetric encryption with quantum key distribution (QKD), as explored in this research.

Classical cryptography – based on the hardness of mathematical problems and computation complexity – has been used to secure information and communication for decades. The classical cryptographic techniques include Symmetric Encryption techniques such as Advanced Encryption System(AES), Data Encryption Standard (DES); Asymmetric key algorithms such as Rivest-Shamir-Adleman(RSA), Elliptic Curve Cryptography (ECC) and Hash Functions. These classical cryptography techniques could be the focus of brute force attacks, as they are not capable enough of handling the quantum attacks. To address this vulnerability, the algorithms of Post Quantum Cryptography was proposed – which are strong, reliable and secure against the quantum attacks [18].

E91 and BB84 are the methods of Quantum Key Distribution (QKD) based on the completeness of quantum mechanics. BB84 is the first QKD protocol, relying on the Heisenberg uncertainty principle and the no-cloning theorem. On the other hand, E91 is an entanglement based protocol basically generalized from Bell's theorem. The users (Alice and Bob) can easily detect the eavesdropping attempt checking the Bell's inequality by using this protocol [12]. Among them, E91 is considered as safer and reliable though it faces some technical difficulties [19].

Hash functions are frequently used in modern cryptography in order to ensure authentication of the digital communication system. It takes an arbitrary length input and produces a fixed sized hashed value where any small changes in input results in an alteration of the hashed value. SHA-0, SHA-1, SHA-2, SHA-3 are the dedicated hash functions of the SHA family [20]. Among them, the SHA-3 family is based on permutation having an extendable output function. SHA3-224, SHA3-256, SHA3-384, SHA3-512 are known as cryptographic hash functions. SHA-3

provides resistance to collision, preimage and second preimage attacks which ensures the security attacks such as digital signature generation and verification, key derivation and pseudorandom bit generation [13].

Steganography is a common practice of hiding information. The main purpose of steganography is to secure the information so that it is completely undetectable through human senses [21]. There are many types of steganography using different types of cover objects such as text [22], audio [23][24][25], video, images (2D and 3D) [26] and information matrices [27]. Due to their high level of data transmission rate and redundancy, digital audio signal are highly suitable for steganography [28]. Some of the audio stenographic process include LSB coding, Parity coding, Phase coding, Spread spectrum, Echo hiding. Among them, LSB is comparatively better as it increases robustness against noise addition and MPEG compression, high capacity and simplicity [29][23]. When QKD protocols are combined with this, it improves more data security. using Bernstein-Vazirani algorithm and the BB84 protocol, it generates a secret key to ensure only authorized access to quantum messages. This quantum steganography method increases hidden channel capacity, resists attacks like intercept-resend, and reduces message detectability by using more Bell states and random information distribution [30].

ChaCha, a variant of Salsa20, represent the ChaCha family of stream ciphers, which improves upon the previous Salsa20 design. It increases the resistance to cryptoanalysis by improving the diffusion per round while maintaining performance. It also achieves better software speed compared to Salsa20 in certain platforms due to its efficient design that allows for SIMD operations [16]. ChaCha is a robust alternative to Salsa20, with enhancements that could make it preferable for various cryptographic applications.

The Poly1305 is a specific type of MAC(Message Authentication Protocol) which is a cryptographic tools used to verify the integrity and authenticity of a message [17]. It computes a 16-byte authenticator for variable-length messages using a 32-byte secret key, which includes a 16-byte key derived from some symmetric encryption and a 16-byte additional randomly generated key which helps in high-speed computations and making it suitable for various applications [17]. The probability of successful attack is also very low due to the uniqueness of nonces and the unpredictability of keys. For all the reasons, it is used widely for network protocols and secure communications.

## 2.1. Novel Contributions to Existing Literature Gap

The mentioned works offer valuable contributions, but they all their shortcomings as well. QKD by itself can not provide any data security. Classical cryptography along with hash function as integrity verifier provides good data security, but this method is not competent enough to stave off against quantum attacks. The consolidation of quantum and classical cryptography offers a promising avenue for reinforcing security, as the combined strengths of these two approaches can lead to more robust and resilient cryptographic systems. To that end, we introduce a cryptography scheme combining Quantum Key Distribution and Classical cryptography, added with Steganography technique. This mixture of different types of information securing technique provides a system that is resilient against both classical and quantum attacks while providing efficient data security. Our main contribution can be summarized as follows -

- Proposed an original architecture that successfully combines E91 QKD protocol, ChaCha20-Poly1305 AEAD and audio steganography using LSB.
- Utilized E91 as the key distribution protocol which employs principles of quantum mechanics for secure key exchange.
- Assimilated audio steganography using LSB substitution into the architecture to augment the security and robustness of the overall system.
- Assessed the performance and reliability of the proposed scheme by measuring the security through end-to-end encryption.

## 3. Preliminaries

In this section, we present the primary concepts employed in our methodology. We start with the quantum key distribution (QKD) protocol proposed by Artur Ekert in 1991, commonly known as the E91 key distribution protocol. Protected by Bell's inequality, E91 protocol provides a secure method of sharing generated key pairs. Then we proceed to the SHA3 family of hashing algorithms, specifically the SHA3-256 function, which can produce a unique 256-bit long hexstream for an input of any length. Next, we introduce the ChaCha20-Poly1305 AEAD, which combines the strong yet simple ChaCha20 stream cipher with Poly1305 MAC. Finally, we discuss the LSB steganography algorithm, which is essentially a technique of hiding an audio signal in the least significant bits of another, larger audio.

### 3.1. Ekert 91 Protocol

The E91 protocol is a quantum key distribution (QKD) scheme that uses entangled photon pairs and the violation of Bell's inequality to securely distribute cryptographic keys between two parties, typically referred to as Alice and Bob. The E91 protocol, based on the principles of quantum mechanics and entanglement, allows secure key distribution between two parties, Alice and Bob. The steps are as follows:

1. Creating Entangled Pairs: Either alice or bob or any other third party, Charlie, can initiates the process by generating pairs of strongly entangled quantum particles (Bell states). These states ensure that the measurements on the particles are perfectly correlated. There are four Bell states, which are below:

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), |\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle)$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), |\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$$

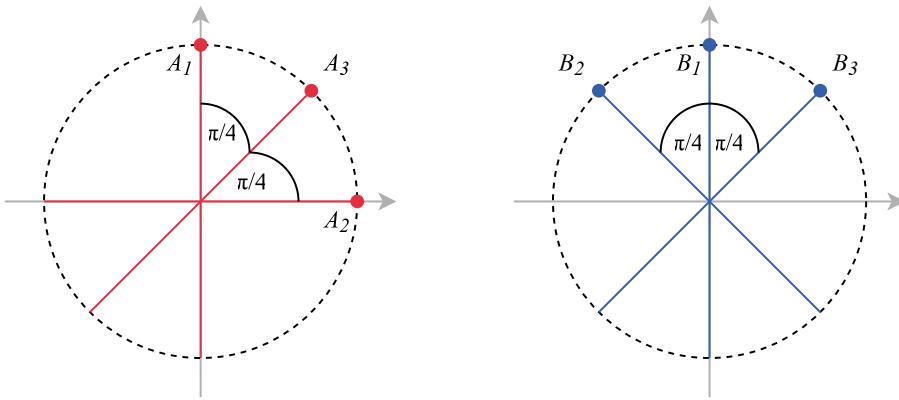
For notions such as  $|01\rangle$  and  $|11\rangle$ , the first digit in brackets refers to the first qubit, and the second digit refers to the second qubit.

2. Sharing the Particles: After generating the entangled pairs, Charlie sends one particle of each pair each to Alice and the other to Bob or if they produced entangled qubits without the help of third party then they will send one particle to another.
3. Performing Measurements: Alice and Bob, located far apart, independently measure their received particles using quantum measurement tools. Their devices are configured with randomly chosen azimuthal angles e.g.

$$\varphi_{A_1} = 0, \quad \varphi_{A_2} = \frac{\pi}{2}, \quad \varphi_{A_3} = \frac{\pi}{4} \quad \text{for Alice};$$

$$\varphi_{B_1} = 0, \quad \varphi_{B_2} = \frac{3\pi}{2}, \quad \varphi_{B_3} = \frac{\pi}{4} \quad \text{for Bob}.$$

The outcomes (0 or 1) are inherently random due to quantum principles. Figure 1 shows the measurement direction:



**Figure 1:** Measurement Direction for Ekert Protocol

4. Comparing Results: Alice and Bob will share their measurement bases over a classical channel. They discard results where their measurement angles differed and keep only those whose bases are aligned. This filtered data forms the raw material for their shared key.
5. Building the Shared key: For the matching bases, Alice and Bob use their correlated outcomes to construct an identical secret key. This key is a binary sequence that is used to securely encrypt/decrypt messages.

6. Checking for Eavesdroppers: The protocol's security depends on the quantum entanglement. If an eavesdropper intercepts the particles, the entanglement is disrupted, altering the expected correlations. By testing a subset of their results, Alice and Bob can detect anomalies. If inconsistencies arise, they abort the key and restart the process.

Alice and Bob can detect the presence of Eve using two steps. The first step is to check if, after measurement, they find that the measurement results differ despite using the same corresponding bases. They can also detect anomalies using the CHSH inequality test. If the measurement result don't violate CHSH inequality, then they can assume that there was an anomaly. The CHSH inequality quantifies correlations between two spatially separated particles measured in different bases. For an entangled pair, if measurements on the first particle are labeled with settings A and a (yielding outcomes  $\pm 1$ ), and those on the second particle use settings B and b (yielding outcomes  $\pm 1$ ), the CHSH parameter S is defined as:

$$S = A(B - b) + a(B + b) \quad (1)$$

Classically, the absolute average value is given by:

$$|\langle S \rangle| \leq 2 \quad \text{or} \quad \pm 2.$$

It is bounded by 2, as one of the terms ( $B \pm b$ ) necessarily cancels out, leaving  $S = \pm 2$ . Expanding S in terms of A, a, B and b results in:

$$|\langle S_1 \rangle| = |\langle AB \rangle - \langle Ab \rangle + \langle aB \rangle + \langle ab \rangle| \leq 2.$$

Remarkably, quantum mechanics predicts violations of these classical bounds. Entangled particles, such as those in Bell states, exhibit stronger correlation due to nonlocal quantum interactions. When measured in specific non-commuting bases the CHSH parameter can reach

$$|\langle S \rangle| = 2\sqrt{2} \approx 2.828 \quad |\langle S \rangle|$$

In our experiment, we expect to obtain a CHSH value close to  $2\sqrt{2}$ ; otherwise, we can claim that there were anomalies/eavesdroppers in the channel.

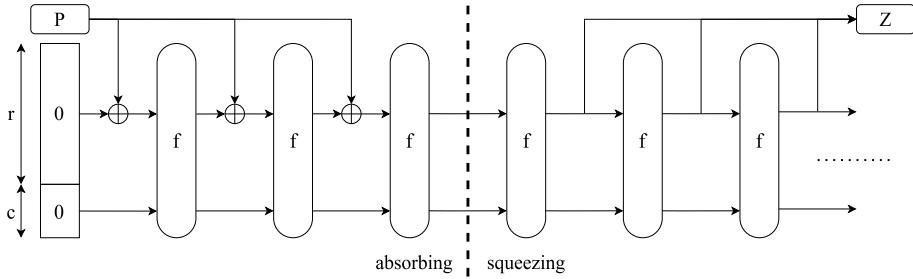
### 3.2. SHA3 – 256 Hash Function

SHA-3 is the latest member [31] of the Secure Hash Algorithm (SHA) family of cryptographic hash functions. Based on the KEECAK algorithm [32], the SHA-3 standard differs fundamentally from its MD-5 [33] like structured predecessors – SHA-1 and SHA-2 [34]. In its essence, the SHA-3 standard is a hash function, that is, for any length of binary data input, there is an output of fixed length. The output is called the *digest* or *hash value*. Depending on the digest length, the four hash functions are called SHA3-224, SHA3-256, SHA3-384 and SHA3-512. The suffixes after the dash indicate the fixed length of the digest, for example, SHA3-256 – which is of interest to our methodology – produces 256 bits long digest. As the digest length is constant irrespective of input length, SHA3 is ideal for integrity and signature verification, as well as password hashing, that is not storing password in clear-text format, rather storing the hash of the password for verification.

SHA-3 works by using the sponge construction [35]. Here, data is "absorbed" into the sponge and the result is "squeezed" out, as shown in Figure 2. For an input bit string  $N$ , a padding function  $pad$ , and a permutation function  $f$  yields the sponge construction  $Z = \text{sponge}[f, pad, r](N, d)$ . Here, the permutation function  $f$  operates on a bit block of width  $b$ , rate  $r$  and yields an output of length  $d$ . From the sponge construct, we have capacity  $c = b - r$  and digest  $Z$  of length  $b$ . The block permutation  $f$  is Keccak-f[1600] that utilizes XOR, AND and NOT operations. It is defined for any power-of-two word size,  $w = 2^l$ . The basic block permutation function consists of  $12 + 2l$  rounds of five steps:  $\theta$  (parity computation),  $\rho$  (bitwise rotation),  $\phi$  (permutation),  $\chi$  (bitwise combination) and  $\iota$  (XOR).

The process of retrieving the digest  $Z$  from input  $N$  is described below –

- The input  $N$  is padded using the padding function,  $pad$ . This yields in a padded bit strength  $P$  the length of which is divisible by  $r$  ( $n = \text{len}(P)/r$ ).
- $P$  is broken into  $n$  consecutive pieces, each piece of length  $r$ , so that  $P \rightarrow P_0, \dots, P_{n-1}$
- State  $S$  is initialized with string of  $b$  zero bits.

**Figure 2:** The sponge construction for SHA-3.

- The input is absorbed into the state: for each block  $P_i$ ,
  - $P_i$  is extended at the end by  $c$  zero bits, so that the length is  $b$ .
  - The resulting bit string is XORed with  $S$ .
  - The block permutation  $f$  is applied on the result, yielding a new state  $S$ .
- $Z$  is initialized with empty string
- while  $\text{length}(Z) < d$ :
  - The first  $r$  bits of  $S$  is appended to  $Z$ .
  - if ( $\text{length}(Z) < d$ ), apply  $f$  to  $S$ , yielding a new  $S$ .
- $Z$  is truncated to  $d$  bits.

For example, for the following binary stream,

$$N = 00100000010001010100000110011000010001001111101010000001100011001111101101011110$$

the digest is 43b3d090d2aebb7790183e58808b71fe890858b5ca3fd308efb2aeba2fde0b71

### 3.3. ChaCha20-Poly1305

ChaCha20-Poly1305 is an authenticated encryption with associated data (AEAD) algorithm that provides a comparatively fast software performance. It is typically faster than AES-GCM (Galois/Counter Mode) in the absence of hardware acceleration [15]. This algorithm combines the ChaCha20 stream cipher [16] and the Poly1305 [17] universal hash family that acts as message authentication code – both of which was developed independently by Daniel J. Bernstein.

A stream cipher is a symmetric key cipher, that is an encryption and decryption algorithm, that combines plaintext data with a pseudorandom keystream [36]. Stream ciphers works by encrypting each plaintext digit one at a time with the corresponding keystream digit, as opposed to block ciphers where blocks of plaintext with a certain length is encrypted with the keystream. In 2008, Bernstein proposed the ChaCha family of stream ciphers, a successor to the Salsa20 stream cipher [37] proposed by Bernstein in 2005. ChaCha was proposed as an alternative to Salsa20 [38] with slightly better performance. Similar to it's predecessor, the initial state of ChaCha is 512-bit long – made up of a 128-bit constant, a 256-bit key, a 64-bit counter and a 64-bit nonce. The 128-bit constant is usually "*expand 32-byte k*". This 512-bit long input is arranged in a  $4 \times 4$  matrix where each entry is a 32-bit word. The matrix arrangement is shown bellow –

$$\begin{bmatrix} \text{Constant}[0] & \text{Constant}[1] & \text{Constant}[2] & \text{Constant}[3] \\ \text{Key}[0] & \text{Key}[1] & \text{Key}[2] & \text{Key}[3] \\ \text{Key}[4] & \text{Key}[5] & \text{Key}[6] & \text{Key}[7] \\ \text{Counter}[0] & \text{Counter}[1] & \text{Nonce}[0] & \text{Nonce}[1] \end{bmatrix} \quad (2)$$

The core operation of ChaCha, and it's predecessor Salsa20, is the quarter-round  $\text{QR}(a, b, c, d)$ . ChaCha uses 4 additions, 4 XORs and 4 rotations to update 4 32-bit state words –  $a, b, c, d$ . The update procedure is as follows

–

```

a += b; d ^= a; d <<<= 16;
c += d; b ^= c; b <<<= 12;
a += b; d ^= a; d <<<= 8;
c += d; b ^= c; b <<<= 7;

```

If the elements of matrix 2 is indexed from 0 to 15

$$\begin{bmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & 11 \\ 12 & 13 & 14 & 15 \end{bmatrix}$$

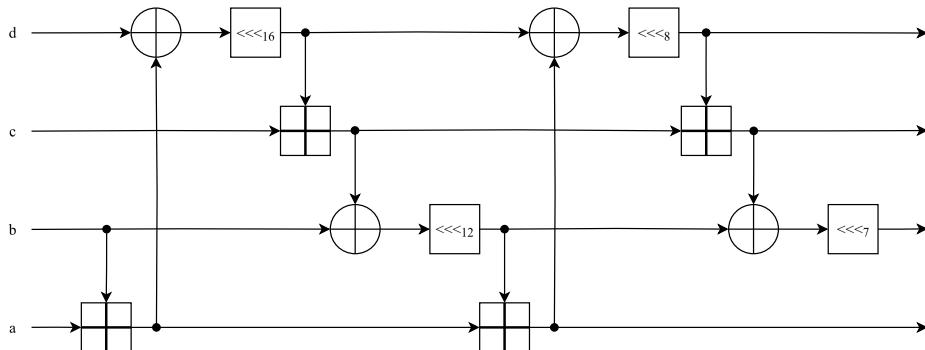
a double round in ChaCha is defined as –

```

// Odd round
QR(0, 4, 8, 12) // column 1
QR(1, 5, 9, 13) // column 2
QR(2, 6, 10, 14) // column 3
QR(3, 7, 11, 15) // column 4
// Even round
QR(0, 5, 10, 15) // diagonal 1 (main diagonal)
QR(1, 6, 11, 12) // diagonal 2
QR(2, 7, 8, 13) // diagonal 3
QR(3, 4, 9, 14) // diagonal 4

```

ChaCha20 uses 10 iterations of the double round – an overall of 20 rounds. Hence the name ChaCha20.



**Figure 3:** The ChaCha quarter-round function.

Each word is updated twice in ChaCha20's quarter rounds, as opposed to Salsa20 where each word is updated only once. This results in an average of 12.5 output bits to change in each quarter round of ChaCha20, while the Salsa20 quarter-round changes 8 output bits. Although the ChaCha quarter round contains the same number of adds, xors, and bit rotates as the Salsa20 quarter round, it is slightly faster because two of the rotate functions are multiples of 8, allowing for a small optimization for x86 and other architectures.

After the 10 iterations, a keystream,  $K$  is generated which is XORed with the message stream,  $M$ , to produce the cipher,  $C$ .

$$C = K \oplus M$$

For decryption, the keystream,  $K$ , is XORed with the cipher,  $C$ , to retrieve the message,  $M$ .

$$M = C \oplus K$$

A hash function is a special mathematical function that can take data of arbitrary size to constant-sized values. The output of a hash function is called hash digest or simply hashes. A unique property of most hash functions is their ability

to generate unique digests for each unique inputs. This makes hash functions invaluable for authentication and data integrity verification. In 2002, Bernstein designed the Poly1305 universal hash function, which was intended to provide a message authentication code to authenticate a message using a shared secret key [39]. The original implementation of Poly1305 was in 2005 when Poly1305 hash was combined as a Carter-Wegman authenticator [40] with AES-128 to authenticate encrypted messages.

Poly1305 provides a 128-bit long authentication tag from a 256-bit long one-time key, and a message of arbitrary length. The key has two portions –  $r$  and  $s$ . It is recommended that the pair  $(r, s)$  to be unique. The first portion,  $r$ , is a 128-bit key produced from a symmetric encryption scheme similar to AES. The later portion,  $s$ , is a 128-bit long integer arranged as a 16-octet little-endian format. It is required that –

- $r[3], r[7], r[11]$  and  $r[15]$  have their top four bits clear (to be in  $\{0, 1, \dots, 15\}$ ).
- $r[4], r[8]$  and  $r[12]$  have their bottom four bits clear (to be in  $\{0, 4, 8, \dots, 252\}$ ).

Each message is required to be accompanied by a 128-bit Nonce generated from the accompanying encryption scheme. The process of producing the authenticator tag from the message and the key are briefly outlined here –

- A variable called the "accumulator" is initialized with zeros
- The message is divided into 128-bit long blocks. Each block is read in little-endian sequence.
- Each block is appended by 1 byte so that each block is 136-bit long. If any of the block is not 136-bit long (usually the last block), the block is padded with zeros to make it of proper length.
- Each 136-bit long message block is multiplied with a clamped value  $r$  and the multiplication result is added to the accumulator.
- The value stored in the accumulator is reduced using  $mod(2^{128})$  and the reduced value is stored in the accumulator.
- When the whole message is converted, the value in the accumulator is added to  $s$ .
- From the result, the 128 least significant bits are formatted in little-endian format to produce the tag.

The ChaCha20-Poly1305 AEAD adopted for Internet Engineering Task Force (IETF) [41] and subsequently as the standard for Transport Layer security (TLS) [42] and the OpenBSD Secure Shell (OpenSSH) [43] is slightly different from the original structures proposed by Daniel J. Bernstein. To be precise, the ChaCha20 symmetric encryption is slightly modified. As opposed to a 64-bit Nonce, the modified ChaCha20 has a 96-bit long Nonce. The modified matrix is shown below –

$$\begin{bmatrix} \text{Constant}[0] & \text{Constant}[1] & \text{Constant}[2] & \text{Constant}[3] \\ \text{Key}[0] & \text{Key}[1] & \text{Key}[2] & \text{Key}[3] \\ \text{Key}[4] & \text{Key}[5] & \text{Key}[6] & \text{Key}[7] \\ \text{Counter}[0] & \text{Nonce}[0] & \text{Nonce}[1] & \text{Nonce}[2] \end{bmatrix} \quad (3)$$

The remaining portions of the ChaCha20 scheme remains unaltered. The same is true for the Poly1305 authenticator. A high level overview of this modified ChaCha20 encryption and Poly1305 authenticator tag generation is outlined below –

- A 256-bit long string is used as key for the ChaCha20 key along with a unique 96-bit long Nonce.
- Using the key and the nonce, the ChaCha20 block function is activated, which produces a 512-bit state.
- The first 256 serialized bits of the 512-bit state is taken as the Poly1305 key – the 128 bit as  $r$  and the next 128 bit as  $s$ .
- The 512-bit state is serialized in little-endian format to produce a keystream.
- The keystream is XORed with the message to produce the ciphertext.

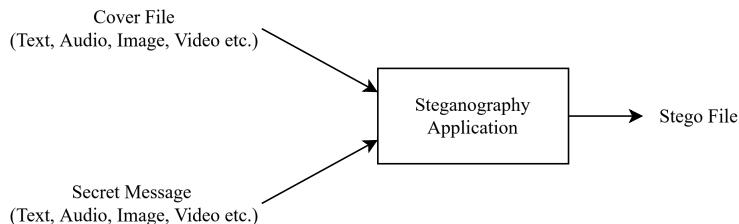
- The Poly1305 function is called with the ciphertext and the previously derived key as input to produce the authenticator tag.

The output from the AEAD is the concatenation of the following –

- A ciphertext which is as long as the plaintext message.
- A 128-bit authenticator tag, generated by the Poly1305 function.

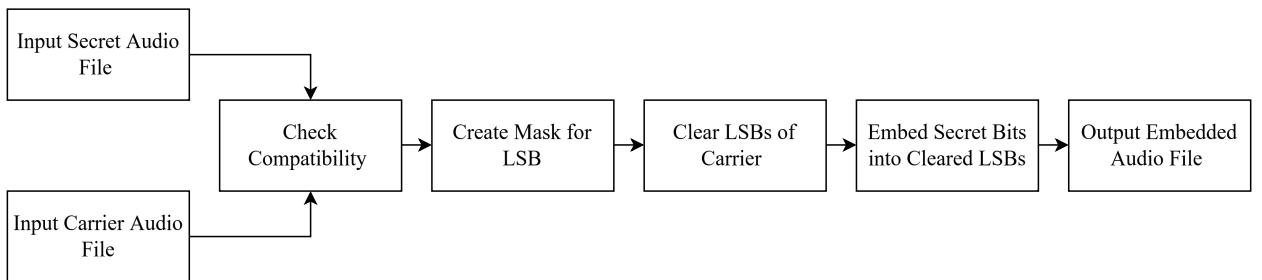
### 3.4. LSB Steganography

Steganography is the process where we don't encrypt the secret message, rather we hide it in another file. Through this process we can hide different types of secret data like text, image, video, audio etc. within cover data which can also be in any different data types.[23][24]



**Figure 4:** Steganography Application Scenario

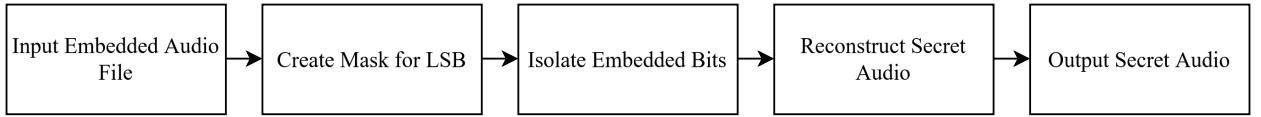
In Audio Steganography secret message is hidden in cover audio but there is no much change in the cover audio because it is use of the psycho-acoustical masking phenomenon of the human auditory system(HAS). There are different methods of steganography. Among them some notables are: LSB Coding, Parity Coding, Phase coding, Spread spectrum, Echo hiding etc.[24] A very popular methodology among them is the LSB(Least Significant Bit) algorithm, which only replaces the least significant bits of the some of the bytes of the cover audio. Due to LSB substitution there is no much changes in the audio also.



**Figure 5:** Block Diagram of LSB Embedding implementation logic.

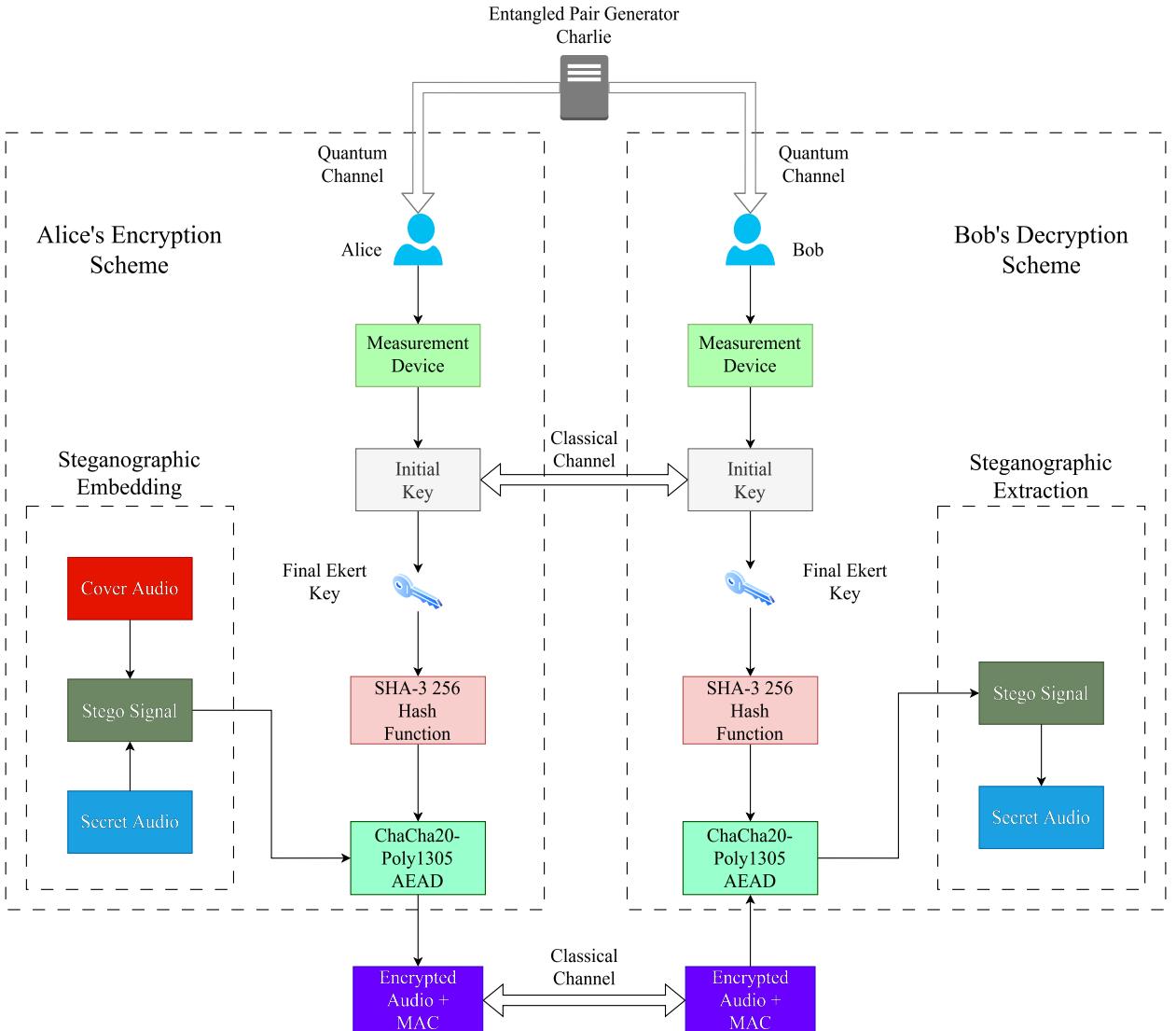
LSB encoding embeds secret audio by replacing the least significant bits of carrier audio frames with bits from the secret audio. First, the carrier and secret audio are checked for compatibility in terms of channels and sample width. The secret audio frames are resized to fit the carrier's capacity. A mask is applied to clear the carrier's LSBs, which are then replaced with shifted bits from the secret audio. The result is a modified carrier audio with embedded data, saved as a new file.

Decoding isolates the embedded bits by applying a mask on the embedded audio and shifting these bits back to their original positions to reconstruct the secret audio. To improve quality, a low-pass filter reduces noise. The reconstructed audio is clipped to valid ranges and saved as the extracted secret.



**Figure 6:** Block Diagram for Steganalysis for extracting the embedded data.

## 4. Methodology



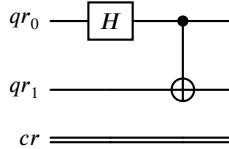
**Figure 7:** A high level overview of the proposed methodology

The architecture that we proposed is a system that integrates E91 QKD protocol with classical cryptographic schemes and lsb steganography to ensure a heavily secure audio data encryption system. The system utilizes the E91 for QKD to generate and exchange the cryptographic keys called the Ekert key, SHA3 to derive a high-entropy 256-bit key from the Ekert key for ChaCha20-Poly1305 encryption and LSB substitution for steganography. The architecture

ensures augmented data security through the combination of classical and quantum techniques. Figure 7 provides a high level overview of our proposed system.

#### 4.1. Initialization

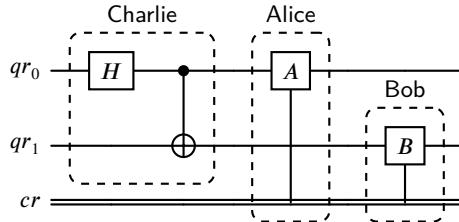
As outlined by Ekert [12], an entangled pair generator, referred to as Charlie, produces pairs of entangled qubits and shares them with Alice and Bob. The circuit shown in Figure 8 is used to produce the entangled qubit pairs. The circuit applied a Hadamard ( $H$ ) gate to put a qubit in superposition and a CNOT ( $\oplus$ ) gate to control the other qubit. Here,  $qr_0$  and  $qr_1$  denotes quantum registers, that is, qubits, while  $cr$  denotes classical register.



**Figure 8:** Singlet circuit for creating entanglement qubit pairs.

#### 4.2. Measurement

Alice and Bob independently create a random set of bases of length  $n$ . Using this set of random bases, Alice and Bob perform measurements on the qubits they received from Charlie. After the completion of the measurements, both parties share their measurement bases over an unsecured Classical Channel. Upon comparison of these measurement bases, the states that lack a common base are discarded. The result is a secret key called the Final Ekert Key that is shared between Alice and Bob. The measurement process is illustrated in Figure 9.



**Figure 9:** Combined circuit for E91 QKD protocol

#### 4.3. Encryption key generation

After the creation of the final Ekert key, Alice and Bob pass the key through the SHA3-256 hashing algorithm to generate another key of fixed length (256 bits) and high entropy. The use of the hashed key instead of the original Ekert key provides an additional layer of security by obscuring the contents of the original key. If an eavesdropper gets hold of the hashed key, he/she will not be able distill the original key as the SHA3 hash function provides a vital protection against reverse calculation.

For 500 singlet states, we may get a key like the following –  
01011000100110100111011100001100111110001011100011000000111011101100000010010111011  
The hash digest of this key is 486340cb54e65a96edc02257b48f3415a0c374f771871a4240ef8097cecdeec2.

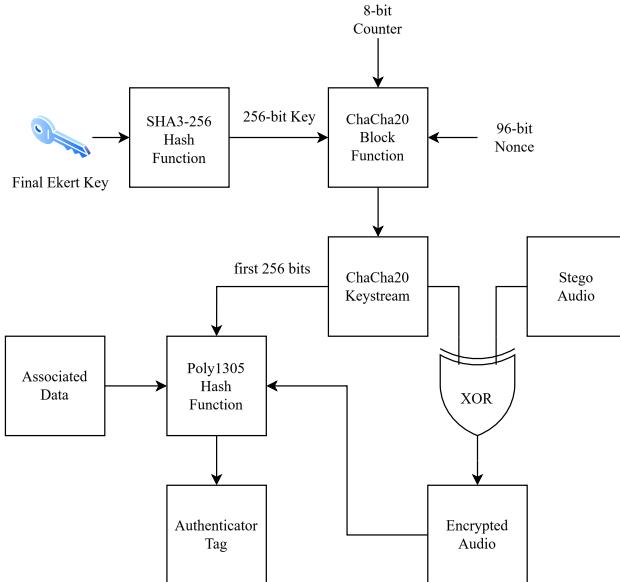
Similarly, for 128 singlet states, we may get a key like the following – 0001001110110110111000101100.  
The hash digest of this key is 9e0ae69c28e4f4e8ba13e1c496fb237284df4e0b6540e168b18bec0cde9ee70f.

#### 4.4. Steganographic embedding

Prior encryption, the audio data is embedded in a cover audio using Steganography. The technique of Least Significant Bit (LSB) substitution is used to produce the stego audio, that is, the audio file created from the cover audio that hides the message audio. First, the message and the cover audio is checked for compatibility. Upon success, the cover audio frames are modified through masking, which clears the LSB of each frame. They are subsequently replaced by the bits of the message audio.

#### 4.5. Encryption and Authenticator generation

The hashed key generated in 4.3 is used to encrypt the stego audio generated in 4.4. The encryption process is accompanied by the generation of an authenticator tag. The ChaCha20-Poly1305 AEAD is used to perform the encryption. The high-level over view of the process is shown in Figure 10.



**Figure 10:** ChaCha20-Poly1305 Encryption and Authenticator tag generation.

The 256-bit key derived from the SHA3-256 hashing algorithm is input along with a 96-bit Nonce and a 8-bit Counter to the ChaCha20 Block function. The function generates a 512-bit state after 20 rounds, which is serialized into a keystream in little-endian format. The first 256-bits of the keystream is input to the Poly1305 hash function along with other associated data. The keystream is XORed with the stego audio to produce the encrypted audio. The encrypted audio is put into the Poly1305 hash function which produces a 128-bit long authenticator tag.

#### 4.6. Decryption and Authentication

The decryption process is the exact reverse of the encryption process. Similar to the encryption process, a 256-bit long key is derived from the shared Ekert key using the SHA3-256 hash function. The key is used to generate the ChaCha20 keystream, similar to the encryption process. The keystream is used in the Poly1305 hash function along with the received encrypted audio to generate a authenticator tag. The generated tag is compared with the received tag to verify the validity and integrity of the received data. Upon validation, the encrypted audio is XORed with the previously generated keystream to recover the stego audio.

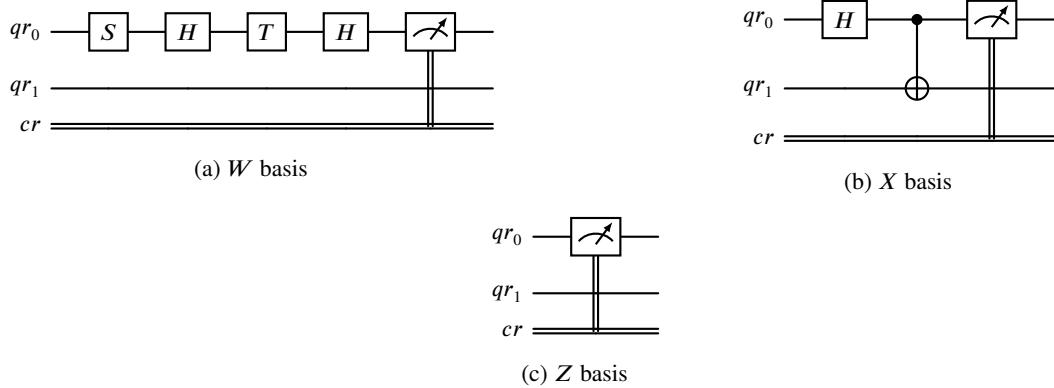
#### 4.7. Steganography extraction

The stego audio recovered in 4.6 is modified to isolate the cover audio and the message audio. As the least significant bits of each frame of the cover audio was replaced by the message audio bits, the LSBs of each frame was extracted in order to reconstruct the binary stream of the message audio. The reconstructed binary stream is modified according to bit-depth to form audio frames. From these frames the message audio is reconstructed.

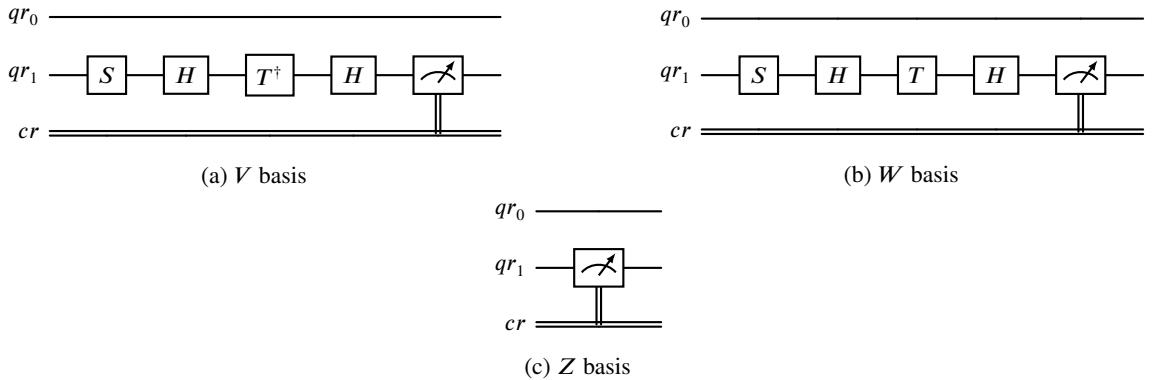
### 5. Implementation Details

In the course of our research, we utilized an amalgamation of powerful tools and libraries in the pursuit of the effective implementation of our proposed scheme. The Qiskit SDK from IBM [44] acted as one of the foundational columns of our research by enabling us to develop and simulate quantum circuits and thus properly implementing the E91 protocol. For our implementation, we specifically used Qiskit v0.39.2. Using Qiskit, we developed the singlet

circuit for entangled pair generation 8, the measurement circuits of Alice and Bob and the creation of the Final Ekert Key. The respective measurement circuits are shown in Figures 11 and 12.



**Figure 11:** Alice's measurement circuits.



**Figure 12:** Bob's measurement circuits.

The key was then zero padded to make it divisible by 8. The zero padded key was converted into a high entropy 256-bit long encryption key using PyCryptodome's [45] hash function – specifically the SHA3-256 function, as stated in 4.3.

For our research, we used a single cover audio to hide two secret message audios using LSB steganography following the process detailed in 4.4 using the computational tools and libraries provided by NumPy [46]. For our cover audio, we chose No Place To Go. The original audio was in .mp3 format. We converted it into .wav format using FFmpeg [47] and stored it as *cover.wav*. The waveform of the cover audio is shown in Figure 13a. For our secret audios, we chose Alone as secret audio 1 and Stylish Deep Electronic as secret audio 2. Similar to the cover audio, the secret audios were converted from .mp3 format to .wav format using FFmpeg and was saved as *secret1.wav* and *secret2.wav*, respectively. Their respective waveforms are shown in Figure 13b and 13c.

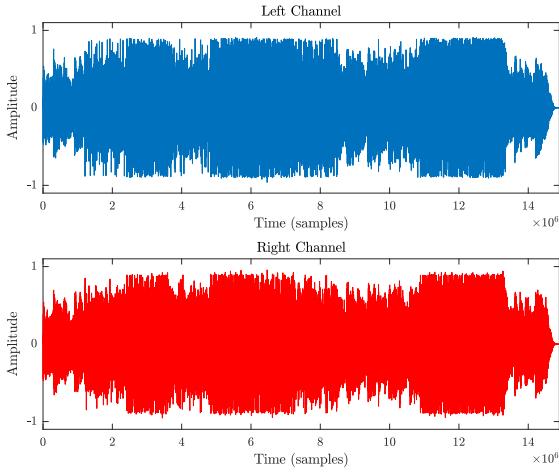
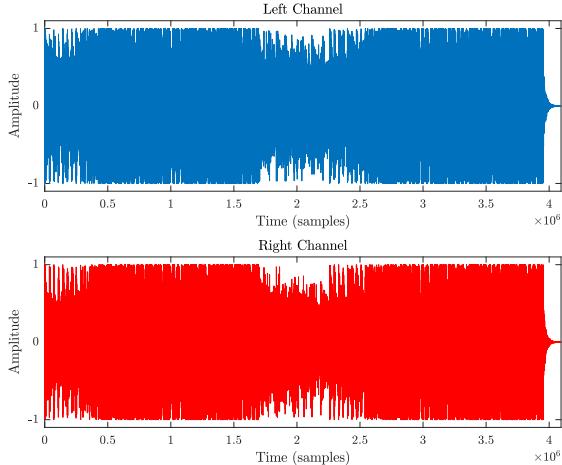
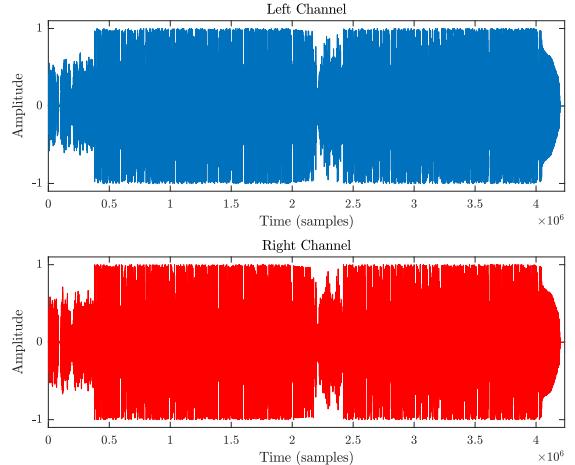
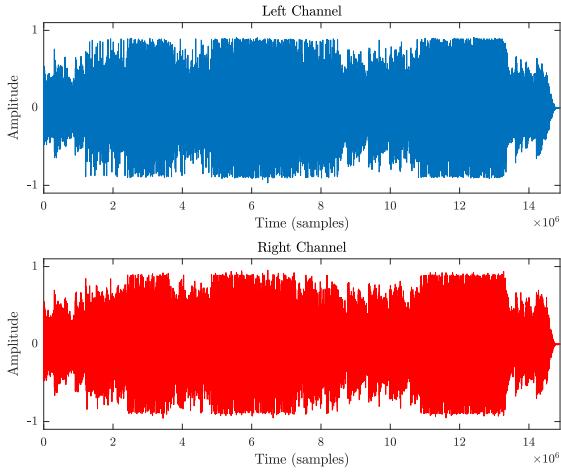
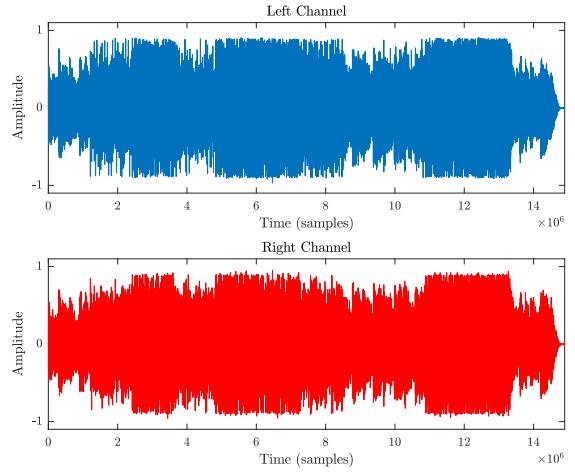
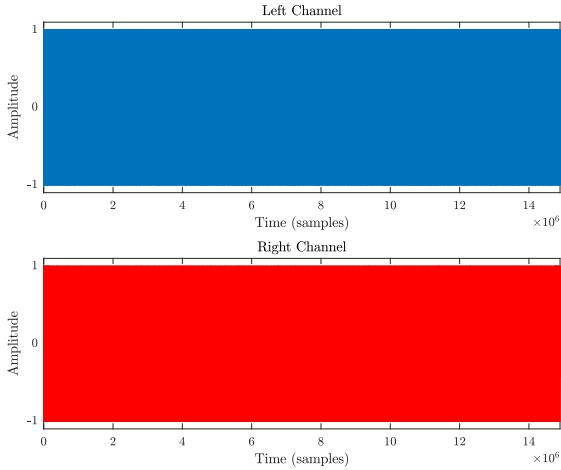
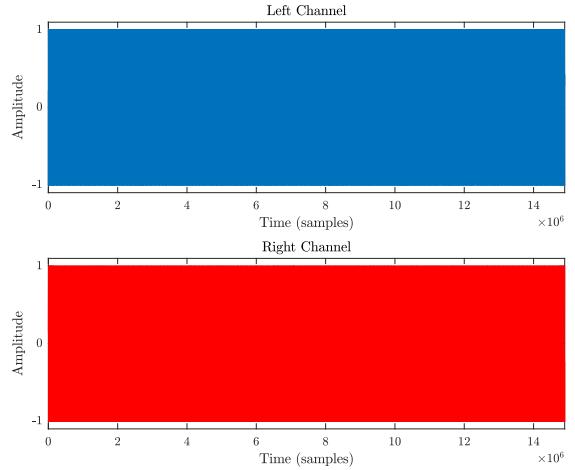
(a) Waveform of Cover audio (*cover.wav*)(b) Waveform of Secret audio 1 (*secret1.wav*).(c) Waveform of Secret audio 2 (*secret2.wav*).

Figure 13: Waveforms of Cover audio and Secret audios.

Using the LSB steganography technique, we embedded *secret1.wav* and *secret2.wav* inside *cover.wav* and saved them as *embedded1.wav* and *embedded2.wav*. Their respective waveforms are shown in Figure 14a and 14b. From visual and auditory analysis of the waveforms, we found imperceptible differences between these three audio signals, indicating effective implementation of LSb steganography.

The encryption key was used as the input key to PyCryptodome's ChaCha20-Poly1305 AEAD function in order to generate the cipher keystream as per the procedure stated in 4.5. Using the keystream, the embedded audios, *embedded1.wav* and *embedded2.wav* was encrypted as ciphertext. From the the ciphertext, the authenticator tag was generated. The nonce used in producing the ciphertext, the ciphertext itself and the corresponding tag was stored in .bin files, namely, *encrypted1.bin* and *encrypted2.bin*. The .bin files were structured in such a way that the 96-bit nonce was followed by the ciphertext, which was in turn followed by the 128-bit tag. The waveforms of the encrypted audios are shown in Figure 15a and 15b.

The encrypted file was received by the intended recipient, who decrypted the file and authenticated it by comparing the received tag and the newly computed tag, as illustrated in 4.6. Upon decryption and successful authentication, the audio was stored in .wav files, namely, *decrypted1.wav* and *decrypted2.wav*, corresponding to *embedded1.wav* and *embedded2.wav*. The waveforms of these decrypted audios are shown in Figure 16a and 16b. Upon analyzing both visually and aurally, we found the decrypted audios, *decrypted1.wav* and *decrypted2.wav* to be indistinguishable from their embedded counterparts, *embedded1.wav* and *embedded2.wav*.

(a) Waveform of Embedded audio 1 (*embedded1.wav*)(b) Waveform of Embedded audio 2 (*embedded2.wav*)**Figure 14:** Waveforms of Embedded audios produced through LSB Steganography.(a) Waveform of Encrypted audio 1 (*encrypted1.wav*)(b) Waveform of Encrypted audio 2 (*encrypted2.wav*)**Figure 15:** Waveforms of the Encrypted audios produced by encrypting the Embedded audios.

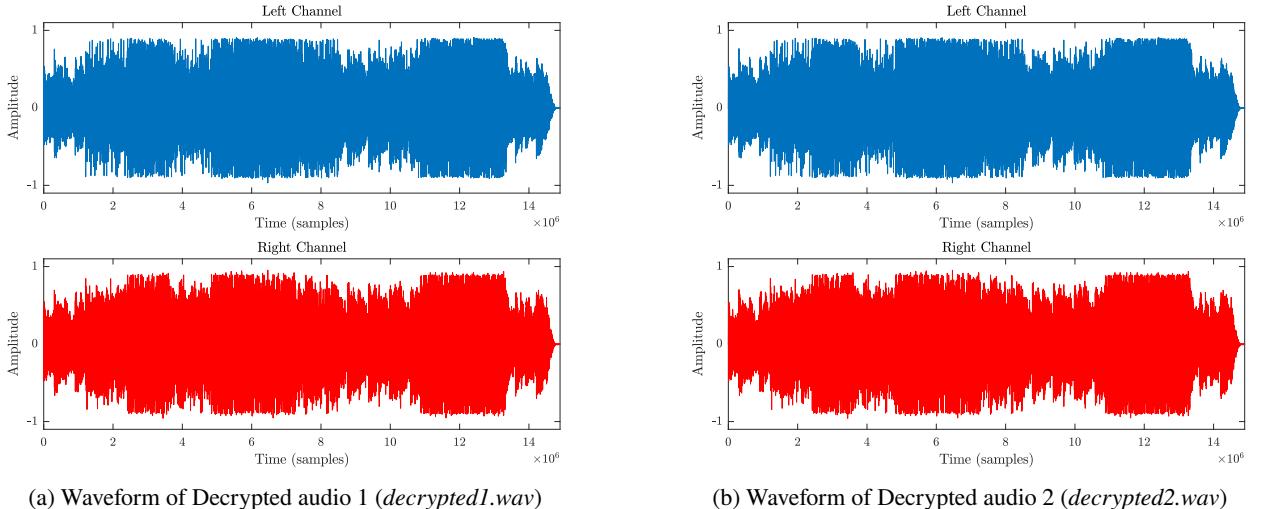
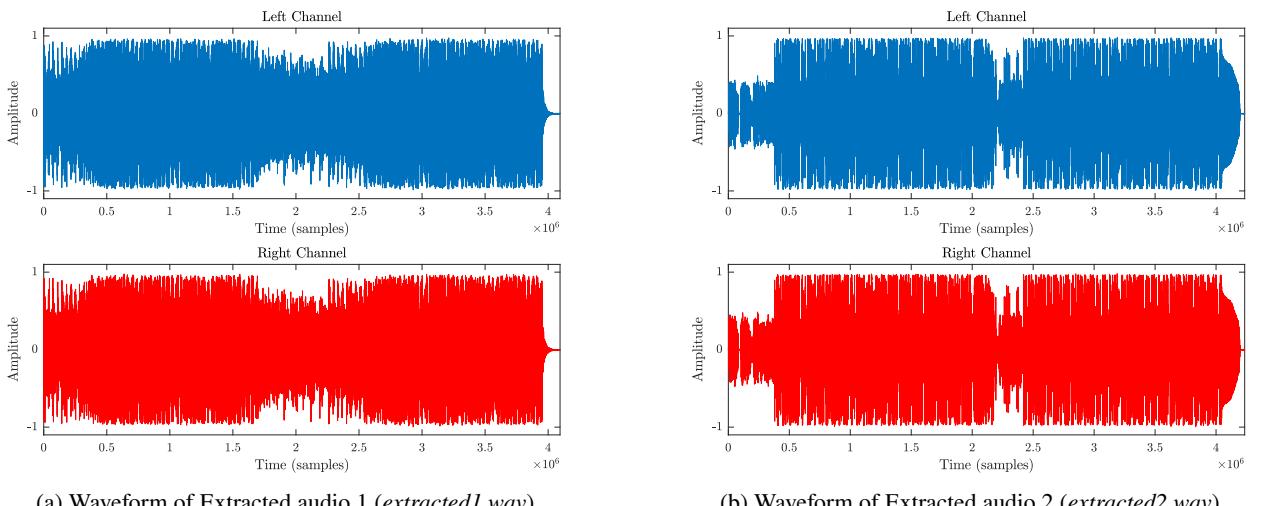
From the decrypted audios, we extracted the secret audio as detailed in 4.7. The extracted audios was stored as *extracted1.wav* and *extracted2.wav*, corresponding to *decrypted1.wav* and *decrypted2.wav*. The waveforms of these audios are shown in Figure 17a and 17b. Visual analysis depicted a slight variation between original secret audios and extracted audios, that is, between 13b and 17a as well as 13c and 17b. This chang was corroborated by the auditory analysis. Nevertheless, these slight variations did not make the extracted audio intelligible, rather, it felt like the amplitude of the audio signals was slightly decreased, proving an effective LSB steganographic extraction.

All the waveforms shown in this section were generated using MATLAB [48].

## 6. Analysis

### 6.1. Analysis of Encryption Properties

With the help of the tools provided by NumPy, we calculated various properties of the encrypted audio signals obtained in the course of implementing our methodology. These properties ranges from entropy of the encrypted signal to sample change rate analysis.

**Figure 16:** Waveforms of the Decrypted audios produced by decrypting the Encrypted audios.**Figure 17:** Waveforms of the Extracted audios produced through LSB extraction.

### 6.1.1. Entropy Analysis

The randomness and unpredictability of a cipher can be measured by the Shannon entropy [49]. In our scenario, the entropy of an encrypted audio signal is an indication of its unpredictability and randomness. The method of calculating entropy is as follows –

$$H(x) = - \sum_{i=1}^n P(x_i) \log_b(P(x_i)) \quad (4)$$

Here,  $H(x)$  is the entropy of the signal while  $P(x_i)$  is the probability of the  $i$ th value of signal  $x$ . For a good encryption technique, it is desirable to have the entropy of the encrypted signal to be as close to 16 as possible (for 16-bit audio signals). Table 1 shows the entropy values of the two encrypted signals, as well as those of the embedded signals.

From the table above, it can be seen that the entropy of the embedded signals are far below 16, almost by a margin of 1.6, while that of the encrypted signals are  $\approx 16$ , indicating proper encryption. In addition, the entropy of the decrypted signals being equal to those of the original signals indicate precise decryption.

**Table 1**

Entropy of original embedded audios before encryption, after encryption and after decryption.

| Audio File              | Original          | Encrypted          | Decrypted         |
|-------------------------|-------------------|--------------------|-------------------|
| Audio 1                 | 9.708523634053172 | 15.998421037422506 | 9.708523634053172 |
| Audio 2                 | 9.792556077074044 | 15.998420782458867 | 9.792556077074044 |
| avg.:15.998420909940687 |                   |                    |                   |

### 6.1.2. Correlation Coefficient Analysis

Correlation coefficient provide an insight into the relationship between two random variables. For this research, we calculated the correlation coefficients between the adjacent sample of the audio signals before and after encryption. The general formula of correlation coefficient is given by the following –

$$r(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (5)$$

Here,  $x$  is the entity with respect to which the correlation coefficient is calculated, which in our case is one sample of both the embedded and encrypted audios, while  $y$  is the entity whose correlation coefficient is being calculated, in this case the adjacent sample of both the embedded and the encrypted audios.  $\bar{x}$  and  $\bar{y}$  are the means of  $x$  and  $y$ , respectively. The correlation coefficients between the embedded and the encrypted audio signals are shown in Table 2.

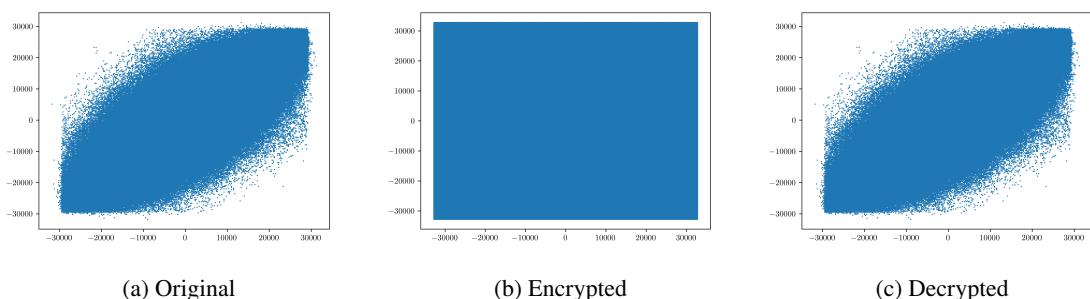
**Table 2**

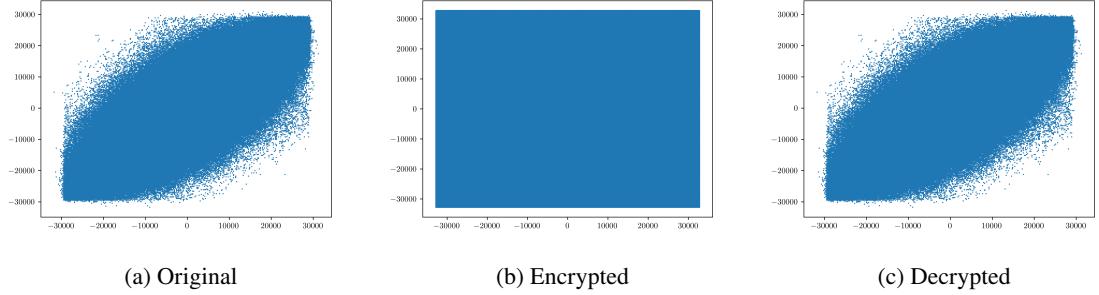
Correlation between adjacent samples of original embedded audios before encryption, after encryption and after decryption.

| Audio File | Original           | Encrypted    | Decrypted          |
|------------|--------------------|--------------|--------------------|
| Audio 1    | 0.543752985460514  | -0.000019216 | 0.543752985460514  |
| Audio 2    | 0.5437499517001024 | 0.0000484704 | 0.5437499517001024 |

From the table above, it is evident that both of the encrypted audio signals have very poor correlation among their corresponding adjacent samples, as both correlation coefficients were well below zero, indicating high randomness of the encrypted signals. Furthermore, the correlation coefficient for the decrypted signals are the same as those of the original signals, implying proper decryption.

Figure 18 and 19 shows the correlation map before encryption, after encryption and after decryption of Audio 1 and Audio 2 produced through steganography, respectively. The blatant difference between Figure 18a and Figure 18b as well as Figure 19a and Figure 19b is a testimony of proper encryption. The similarities between Figure 18a and Figure 18c as well as Figure 19a and Figure 19c indicates proper decryption. Furthermore, the accuracy of the steganography scheme is proved by the fact that Figures 18a, 18c, 19a and 19c are all objectively similar.

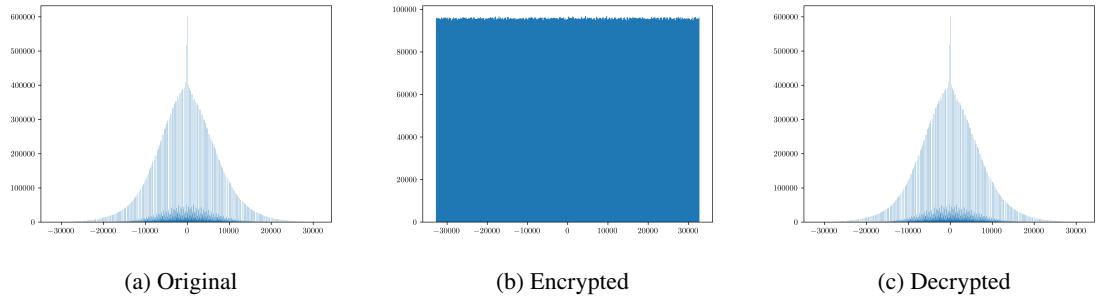
**Figure 18:** Correlation map of Audio 1: (18a) before encryption, (18b) after encryption and (18c) after decryption.



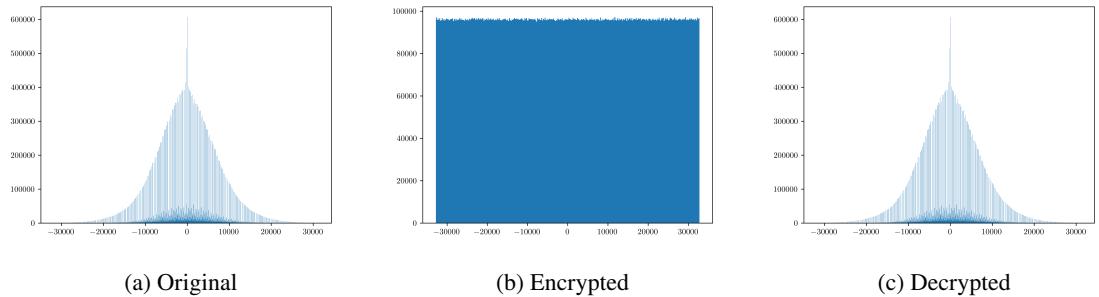
**Figure 19:** Correlation map of Audio 2: (19a) before encryption, (19b) after encryption and (19c) after decryption.

### 6.1.3. Histogram Analysis

Histogram is a visual representation of how data is distributed. For audio signals, Histogram depicts the sample frequency with respect to sample amplitude. For a good encryption, the histogram is expected to be uniformly distributed instead of being converging or diverging. Figure 20 and 21 shows the histogram before encryption, after encryption and after decryption of Audio 1 and Audio 2 produced through steganography, respectively. The histograms



**Figure 20:** Histogram of Audio 1: (20a) before encryption, (20b) after encryption and (20c) after decryption.

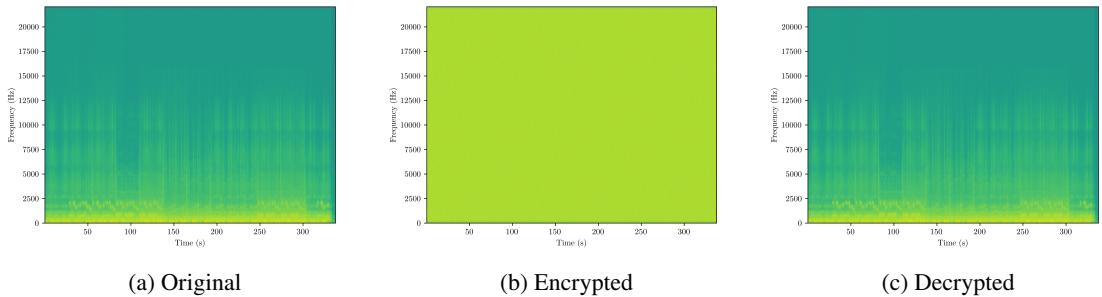


**Figure 21:** Histogram of Audio 2: (21a) before encryption, (21b) after encryption and (21c) after decryption.

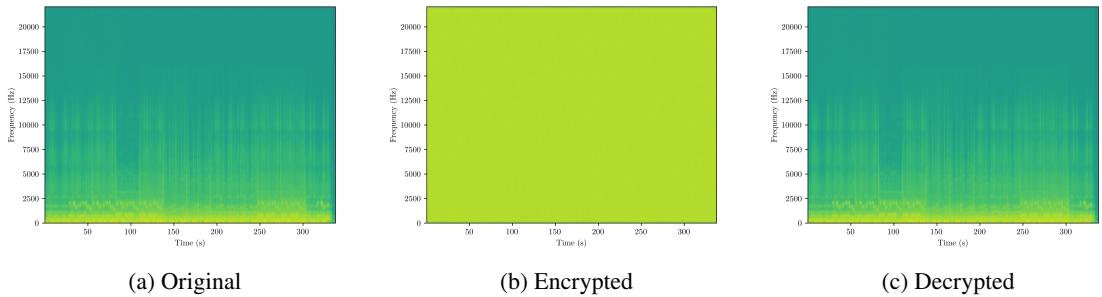
depicted in Figures 20 and 21 demonstrate a significant difference between the embedded and encrypted audio signals, confirming the effectiveness of the encryption process. In contrast, the similarity between the histograms of the embedded and decrypted audio signals (Figures 20a and 20c, and Figures 21a and 21c) verifies the accuracy of the decryption process. Moreover, the visual similarity between the histograms of the embedded and decrypted audio signals provides strong evidence for the reliability of the steganography scheme.

#### 6.1.4. Spectrogram Analysis

Spectrogram is a visual representation of the variation of the spectrum of frequencies of a signal with respect to time. Figure 22 and 23 shows the histogram before encryption, after encryption and after decryption of Audio 1 and Audio 2 produced through steganography, respectively.



**Figure 22:** Spectrogram of Audio 1: (22a) before encryption, (22b) after encryption and (22c) after decryption.



**Figure 23:** Spectrogram of Audio 2: (23a) before encryption, (23b) after encryption and (23c) after decryption.

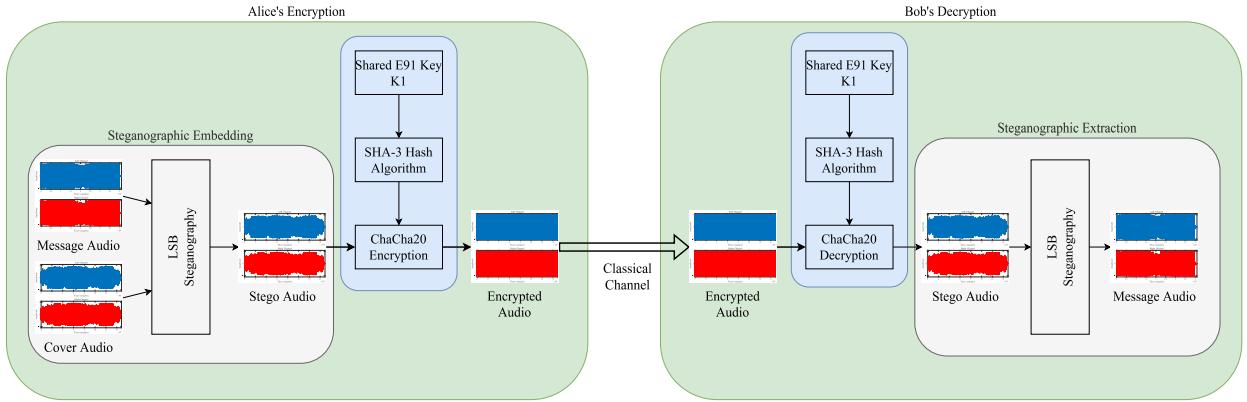
A visual inspection of the spectrograms in Figures 22 and 23 reveals a stark contrast between the embedded and encrypted audio signals, indicating successful encryption. Conversely, the spectrograms of the embedded and decrypted audio signals (Figures 22a and 22c, and Figures 23a and 23c) exhibit a high degree of similarity, confirming the correctness of the decryption process. The striking resemblance between these correlation maps also underscores the efficacy of the steganography scheme, demonstrating its ability to conceal and retrieve the embedded audio signals accurately.

#### 6.1.5. Key Sensitivity Analysis

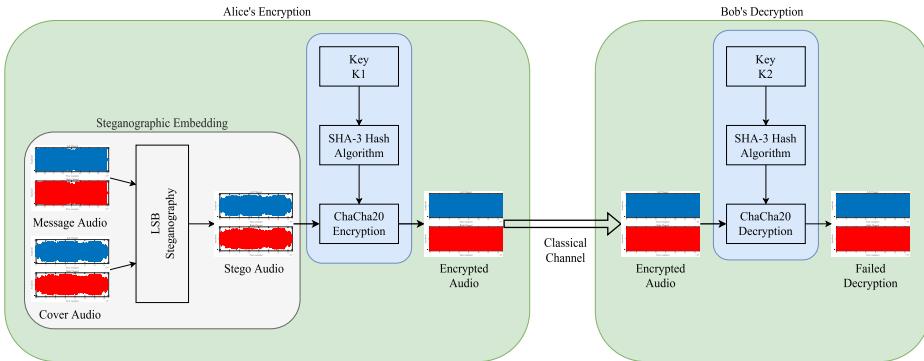
Key Sensitivity is a measure of the response to the slightest change in key of an encryption system. High Key Sensitivity is an indication of the slightest change in encryption key causing huge changes in the encryption and decryption process. To demonstrate Key Sensitivity of the proposed architecture, let us consider two scenarios –

- **Scenario – 1:** The stego-audio is encrypted using K1 key and decrypted using the same key (Figure 24).
- **Scenario – 2:** The stego-audio is encrypted using K1 key but decryption is attempted using K2 key, which differs from K1 by a single bit.(Figure 25)

For instance, let  $K1 = 0000100000100001010001011000011$  and  $K2 = 1000100000100001010001010001011000011$ . The hash output are  $SHA3_{K1} = e57179067dc01f3d3434e4e1447d0368cd716894c71ca9e59da3d7b9a5be6fce$  and  $SHA3_{K2} = a2b702be844ae3adea1fd88c1c6e2452fdd129988fd78f047ebacf5c5d2aa67$ . Due to the mismatch in computed hashdigest, decryption attempt with  $K2$  will fail.



**Figure 24:** Key Sensitivity Analysis Scenario 1: Encryption and Decryption using the same key,  $K1$



**Figure 25:** Key Sensitivity Analysis Scenario 2: Encryption using key  $K1$  and Decryption using key  $K2$

### 6.1.6. Unified Average Changing Intensity (UACI)

Unified Average Changing Intensity is a measure of the average intensity change rate between the original and the modified signal. In this case, UACI measures the average intensity change rate between the embedded audio and the encrypted audio signals. A higher UACI value is preferred, as a higher UACI value indicates a considerable change in the encrypted audio from the original embedded audio. For 16-bit audio signals, UACI is calculated using the following formula –

$$\text{UACI} = \frac{1}{N} \sum_{i=1}^N \left( \frac{|S_i - D_i|}{65535} \right) \times 100 \quad (6)$$

Here,  $S_i$  is the samples of the embedded signal,  $D_i$  is the samples of the encrypted signals and  $N$  is the total number of samples, which is the same in both signals. The calculated values of UACI is shown in Table 3.

**Table 3**

UACI values of audio signals.

| Audio File | UACI               |
|------------|--------------------|
| Audio 1    | 50.003490097286196 |
| Audio 2    | 49.991852542937640 |
| avg.:      | 49.997671320111918 |

It can be gleaned with ease from the table above that the encrypted audios almost do not contain any patterns from the original embedded audios as the UACI values were comparatively high.

### 6.1.7. Number of samples change rate (NSCR)

Number of samples change rate (NSCR) calculates the percentage of the sample values that are in contrast between the original embedded audios and the encrypted audios. Higher NSCR values are an indication of higher number of samples being altered in the course of encryption and thus producing a signal substantially different from the original signal. The following formula is utilized to determine the value of NSCR –

$$\text{NSCR} = \left( \frac{D}{N} \right) \times 100 \quad (7)$$

Here,  $D$  is the number of sample that are distinct between the original embedded signal and the encrypted audio signal and  $N$  is the total number of samples, which is the same in both signals. The NSCR values that we determined is shown in Table 4.

**Table 4**

NSCR values of audio signals.

| Audio File              | NSCR              |
|-------------------------|-------------------|
| Audio 1                 | 99.9985266157303% |
| Audio 2                 | 99.9984527787054% |
| avg.: 99.9984896972179% |                   |

It can be concluded that the encrypted audio signals bear very small resemblance to their original embedded counterpart, as the NSCR values of both signals were almost close to 100%.

## 6.2. Security Analysis

Although our encryption technique provided us with considerably robust encryptions, there exists two different security threat – one in the quantum channel and the other in the classical channel. In the quantum channel, an eavesdropper, Eve, might perform her own measurement on the qubits intended for Alice and Bob. In this process, Eve may gain insight about Alice and Bob's keys. In the classical channel, the encrypted audio suffer from tampering, that is, a malicious party might intercept the encrypted signal and modify it before relaying it to the intended recipient. These scenarios and their solutions are discussed in the subsequent sections.

### 6.2.1. Eavesdropping Analysis

The E91 QKD is protected by Bell's inequality. In the presence of an eavesdropper, Eve, the Clauser – Horne – Shimony – Holt (CHSH) correlation [50] will be violated. This will alert Alice and Bob of Eve's presence, allowing them to discard the now obsolete key and take necessary steps. Here, we present the CHSH correlation values for different key lengths in the presence and absence of eavesdropping. Table 5 displays the CHSH correlation values for different key lengths without eavesdropping, which are close to  $2\sqrt{2} \approx 2.828$ , indicating that the qubits are maximally entangled and thus confirming the absence of an eavesdropper. In contrast, in Table 6 the CHSH correlation value is below 2 indicating that the qubits are not maximally entangled and thus confirming the presence of an eavesdropper.

**Table 5**

CHSH correlation values for different key lengths in the absence of an eavesdropper

| Iterations | Key Length | Number of Mismatched Bits | CHSH  |
|------------|------------|---------------------------|-------|
| 1          | 109        | 0                         | 2.794 |
| 2          | 131        | 0                         | 2.754 |
| 3          | 215        | 0                         | 2.636 |
| 4          | 44         | 0                         | 3.009 |

**Table 6**

CHSH correlation values for different key lengths in the presence of an eavesdropper

| Iterations | Key Length | Number of Mismatched Bits | CHSH  | Eve's Knowledge of Alice's Key | Eve's Knowledge of Bob's Key |
|------------|------------|---------------------------|-------|--------------------------------|------------------------------|
| 1          | 126        | 16                        | 1.229 | 88.1%                          | 96.03%                       |
| 2          | 142        | 12                        | 1.346 | 93.66%                         | 93.66%                       |
| 3          | 243        | 32                        | 1.37  | 93.83%                         | 90.53%                       |
| 4          | 44         | 7                         | 1.48  | 95.45%                         | 84.09%                       |

### 6.2.2. Tampering Analysis

After storing the secret audio in the cover audio using steganography, we encrypted the embedded audio using the ChaCha20-Poly1305 AEAD. The ChaCha20-Poly1305 produces an encrypted file along with an authentication tag, which can be used to validate the authenticity of the encrypted file in the receiver end. If a malicious party modifies the encrypted file en route to the receiver, the authentication tag calculated by the receiver will be different from the tag sent by the sender, alerting the receiver of tampering. This will allow the receiver to take the necessary steps to mitigate the effects of the malicious party. Here, we present two different circumstances –

1. The encrypted files, *encrypted1.bin* and *encrypted2.bin*, reaches their destination without any tampering by a malicious party. A brief depiction of this scenario is presented in Table 9.
2. The encrypted files, *encrypted1.bin* and *encrypted2.bin*, are intercepted by a malicious party, who performs random byte modification on the files and sends the modified files to their intended destination. A brief depiction of this scenario is presented in Table 10.

In the first scenario, the encrypted files will be decrypted and verified successfully. For the second scenario, we simulated the random byte modifications that the malicious party would have done. This resulted in verification failure and in turn decryption failure, alerting the recipient of the actions of the malicious party. Some of these alterations are shown in Table 7 and 8.

**Table 7**Comparison between *encrypted1.bin* and its modified counterpart.

| Original File |   | Modified File |   |
|---------------|---|---------------|---|
| Address       | Hexdata                                   | Address       | Hexdata                                   |
| ...           | .....                                     | ...           | .....                                     |
| 000004a0      | 7f7af670c44b <u>9</u> ffd15b38dc21140951e | 000004a0      | 7f7af670c44b <u>6</u> 2fd15b38dc21140951e |
| ...           | .....                                     | ...           | .....                                     |
| 000005e0      | 0ae68fb6740c4347dc <u>f</u> d2b616643d69c | 000005e0      | 0ae68fb6740c4347dc <u>7</u> c2b616643d69c |
| ...           | .....                                     | ...           | .....                                     |
| 00000ae0      | ea01590d92bac <u>03</u> b8412dce7cf8e7d7  | 00000ae0      | ea01590d92bac <u>06</u> cb8412dce7cf8e7d7 |
| ...           | .....                                     | ...           | .....                                     |
| ...           | .....                                     | ...           | .....                                     |

## 7. Findings and Discussion

Our research strived to illustrate the effectiveness of the combination of E91 QKD, ChaCha20-Poly1305 AEAD and LSB steganography for secure exchange of audio messages. In the preceding sections, we performed different types of analysis to validate our proposed methodology. Figure 14a and 14a provides a visual testimony to our LSB embedding scheme, while Figure 15a and 15b demonstrates the effectivity of our chosen encryption scheme. Figure 16a and 16b illustrated the accuracy of the decryption scheme. Figure 17a and 17b depicts the potency of our LSB extraction scheme. We also measured a number of different evaluation metrics to depict the might of our encryption. Table 1 exhibits the Entropy of the encrypted signals, which were both  $\approx 8$ ; a further evidence to the strength of

**Table 8**Comparison between *encrypted2.bin* and its modified counterpart.

| Original File |                                  | Modified File |                                  |
|---------------|----------------------------------|---------------|----------------------------------|
| Address       | Hexdata                          | Address       | Hexdata                          |
| ...           | .....                            | ...           | .....                            |
| 00000770      | 84fa9cb31654d0215b870be20e14f533 | 00000770      | 84fa9cb31654d0ad5b870be20e14f533 |
| ...           | .....                            | ...           | .....                            |
| 00000a60      | 1d7f2184bd0c5027713aac0e2aa71f64 | 00000a60      | 1d7f21846d0c5027713aac0e2aa71f64 |
| ...           | .....                            | ...           | .....                            |
| 00000ed0      | 76911249292f57364d80912cec5b5f42 | 00000ed0      | 76911249292f57364d17912cec5b5f42 |
| ...           | .....                            | ...           | .....                            |
| ...           | .....                            | ...           | .....                            |

the method of our encryption. Table 2 gives tribute to the high randomness of our encrypted signals and their weak relationship with their original, unencrypted counterparts. Table 3 presents the UACI values while Table 4 construes the NSCR values. For both encrypted files, UACI values were  $\approx 50$ , and NSCR values were  $\approx 100\%$ . The obtained values of UACI and NSCR indicates a strong encryption while keeping the distortion low. Furthermore, we performed the CHSH correlation calculation to demonstrate the circumstance of the qubits not being maximally entangled due to the presence of an eavesdropper. In the absence of an eavesdropper, the qubits are maximally entangled, producing a CHSH value around  $2\sqrt{2}$ , as shown in 5. However, in the presence of an eavesdropper, the qubits will no longer be maximally entangled, resulting in CHSH value that will deviate substantially from  $2\sqrt{2}$ . This condition is shown in Table 6. In addition, we performed authentication tag verification to demonstrate the effect of a malicious party altering the encrypted files en route to their destination. When the encrypted files arrive intact at their destination without modification, the recipient can successfully verify the integrity of the received file and decrypt it, as shown in Table 9. Nonetheless, if the encrypted files were modified prior to reaching their destination, the integrity verification by the recipient will result in a failure, consequently causing the decryption process to fail. This is shown in Table 10.

**Table 9**

Scenario 1

|                  | <i>encrypted1.bin</i>  | <i>encrypted2.bin</i>  |
|------------------|--|--|
| Sender End       | Nonce: 5e4ef364b086ab3e2d1118cb<br>Tag: 416c6c92659047d2f0660c36f58c1fa5 | Nonce: d64f6fe0710eb0324e8cc6<br>Tag: 5977c55fe95a60173ff7476bfad4c29a |
| Receiver End     | Nonce: 5e4ef364b086ab3e2d1118cb<br>Tag: 416c6c92659047d2f0660c36f58c1fa5 | Nonce: d64f6fe0710eb0324e8cc6<br>Tag: 5977c55fe95a60173ff7476bfad4c29a |
| Operation Status | Verification successful!<br>Decryption successful!                       | Verification successful!<br>Decryption successful!                     |

**Table 10**

Scenario 2

|                  | <i>encrypted1.bin</i>  | <i>encrypted2.bin</i>  |
|------------------|--|--|
| Sender End       | Nonce: 5e4ef364b086ab3e2d1118cb<br>Tag: 416c6c92659047d2f0660c36f58c1fa5 | Nonce: d64f6fe0710eb0324e8cc6<br>Tag: 5977c55fe95a60173ff7476bfad4c29a |
| Receiver End     | Nonce: .....<br>Tag: .....   | Nonce: .....<br>Tag: .....   |
| Operation Status | Verification failed!<br>Decryption failed!                               | Verification failed!<br>Decryption failed!                             |

Based on the discussion above, we can observe the following –

1. The proposed technique is effective in encrypting audio signals.
2. The encrypted signals have high entropy and randomness compared to their original counterparts.
3. The incorporation of E91 QKD with ChaCha20-Poly1305 AEAD augmented the security of the proposed technique against potential attacks in both key generation and transmission process.

## 8. Conclusion and Future Work

In our research, we proposed a noble approach of securing digital audio transmission combining the Quantum Key Distribution (QKD) specifically E91 with classical encryption system, secure hash function and audio steganography using LSB. We also introduce ChaCha20-Poly 1305 which is an authenticated encryption scheme enhances secrecy and integrity providing robustness of the overall system. By means of E91 protocol, entangled pair of photons are generated known as Ekert Key send to Alice and Bob. Then SHA generates a random 256 bit key from the Ekert Key for LSB steganography and ChaCha20-Poly1305. One of our primary concern is about quantum attacks in future and the incapability of classical computer or the classical encryption system to deal against those attacks. So, we checked the stability of our proposed signal by illustrating different types of attacks such as - Clauser – Horne – Shimony – Holt (CHSH) inequality and Tampering analysis. Besides, signal quality after encryption and decryption also provides a satisfactory result. However, our system depicts the practical combination of classical cryptographic system and QKD(Quantum Key distribution) also known as Post quantum cryptography alongside with steganography and ChaCha20-Poly1305 resulting in an audio encryption technique which adds an extra layer of protection to keep data safe and intact. Our future work involves using different quantum key distribution such as - B92, COW(Coherent One-Way protocol), SARG04 etc instead of E91 protocol which may bring up a better output.[51] Again, modification of the steganography technique such as - phase steganography can improve the robustness of the hidden data against the attacks. In addition, we can also transmit text, images and video by filtering our proposed algorithm which indicates the diversity of our model. Finally, introducing Quantum error correction techniques [52] will increase the effectiveness of QKD in noisy condition [53] and hence we can improve the key generation and distribution over long distance audio transmission.

## References

- [1] V. Rijmen, J. Daemen, Advanced encryption standard, Proceedings of federal information processing standards publications, national institute of standards and technology 19 (2001) 22.
- [2] R. L. Rivest, A. Shamir, L. Adleman, A method for obtaining digital signatures and public-key cryptosystems, Communications of the ACM 21 (1978) 120–126.
- [3] D. J. Bernstein, T. Lange, Post-quantum cryptography, Nature 549 (2017) 188–194.
- [4] L. K. Grover, A fast quantum mechanical algorithm for database search, in: Proceedings of the twenty-eighth annual ACM symposium on Theory of computing, pp. 212–219.
- [5] J. Daemen, V. Rijmen, The Design of Rijndael, Springer Berlin Heidelberg, 2002.
- [6] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, SIAM Journal on Computing 26 (1997) 14841509.
- [7] R. Alléaume, C. Branciard, J. Bouda, T. Debuisschert, M. Dianati, N. Gisin, M. Godfrey, P. Grangier, T. Länger, N. Lütkenhaus, et al., Using quantum key distribution for cryptographic purposes: a survey, Theoretical Computer Science 560 (2014) 62–81.
- [8] V. Bužek, M. Hillery, Quantum copying: Beyond the no-cloning theorem, Physical Review A 54 (1996) 1844.
- [9] D. Sen, The uncertainty relations in quantum mechanics, Current Science (2014) 203–218.
- [10] D. Kahn, The history of steganography, Springer Berlin Heidelberg, p. 15.
- [11] R. J. Anderson, F. A. Petitcolas, On the limits of steganography, IEEE Journal on selected areas in communications 16 (1998) 474–481.
- [12] A. K. Ekert, Quantum cryptography based on Bell's theorem, Physical Review Letters 67 (1991) 661663.
- [13] M. J. Dworkin, Sha-3 standard: Permutation-based hash and extendable-output functions (2015).
- [14] N. Cvejic, T. Seppanen, Increasing the capacity of lsb-based audio steganography, in: 2002 IEEE Workshop on Multimedia Signal Processing., IEEE, pp. 336–338.
- [15] Y. Nir, A. Langley, ChaCha20 and Poly1305 for IETF Protocols, RFC 7539, 2015.
- [16] D. J. Bernstein, et al., ChaCha, a variant of salsa20, in: Workshop record of SASC, volume 8, Citeseer, pp. 3–5.
- [17] D. J. Bernstein, The poly1305-aes message-authentication code, in: International workshop on fast software encryption, Springer, pp. 32–49.
- [18] S. Sharma, K. Ramkumar, A. Kaur, T. Hasija, S. Mittal, B. Singh, Post-quantum cryptography: A solution to the challenges of classical encryption algorithms, Modern Electronics Devices and Communication Systems: Select Proceedings of MEDCOM 2021 (2023) 23–38.
- [19] L. C. Alvarez, P. C. Caiconte, Comparison and analysis of bb84 and e91 quantum cryptography protocols security strengths, International Journal of Modern Communication Technologies and Research 4 (2016) 265683.
- [20] B. Madhuravani, D. Murthy, Cryptographic hash functions: Sha family, Int J Innov Technol Explor Eng 2 (2013) 326–9.
- [21] M. M. Amin, M. Salleh, S. Ibrahim, M. R. Katmin, M. Shamsuddin, Information hiding using steganography, in: 4th National Conference of Telecommunication Technology, 2003. NCTT 2003 Proceedings., IEEE, pp. 21–25.

- [22] Z.-L. Yang, X.-Q. Guo, Z.-M. Chen, Y.-F. Huang, Y.-J. Zhang, Rnn-stega: Linguistic steganography based on recurrent neural networks, *IEEE Transactions on Information Forensics and Security* 14 (2018) 1280–1295.
- [23] P. Jayaram, H. Ranganatha, H. Anupama, Information hiding using audio steganography—a survey, *The International Journal of Multimedia & Its Applications (IJMA)* Vol 3 (2011) 86–96.
- [24] F. Hemeida, W. Alexan, S. Mamdouh, A comparative study of audio steganography schemes, *International Journal of Computing and Digital Systems* 10 (2021) 555–562.
- [25] F. Djebbar, B. Ayad, K. A. Meraim, H. Hamam, Comparative study of digital audio steganography techniques, *EURASIP Journal on Audio, Speech, and Music Processing* 2012 (2012) 1–16.
- [26] S. Farrag, W. Alexan, Secure 3d data hiding technique based on a mesh traversal algorithm, *Multimedia Tools and Applications* 79 (2020) 29289–29303.
- [27] M. Mashaly, A. El Saied, W. Alexan, A. S. Khalifa, A multiple layer security scheme utilizing information matrices, in: 2019 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), IEEE, pp. 284–289.
- [28] K. U. Singh, A survey on audio steganography approaches, *International Journal of Computer Applications* 95 (2014).
- [29] N. Cvejic, T. Seppanen, Increasing robustness of lsb audio steganography using a novel embedding method, in: International Conference on Information Technology: Coding and Computing, 2004. Proceedings. ITCC 2004., volume 2, IEEE, pp. 533–537.
- [30] S. P. Yalla, S. Hemanth, S. T. Kumar, S. Moulali, S. Dileep, Novel text steganography approach using quantum key distribution: Survey paper, *NeuroQuantology* 20 (2022) 923.
- [31] Hash Functions | CSRC | CSRC — csrc.nist.gov, <https://csrc.nist.gov/projects/hash-functions>, June 22, 2020.
- [32] M. P. Guido Bertoni, Joan Daemen, G. van Assche, The keccak sha-3 submission, 2011. [Accessed 26-12-2024].
- [33] R. Rivest, The MD5 message-digest algorithm, Technical Report, 1992.
- [34] W. Penard, T. Van Werkhoven, On the secure hash algorithm family, *Cryptography in context* (2008) 1–18.
- [35] T. sponge, duplex constructions, Team keccak - guido bertoni, joan daemen, seth hoffert, michaël peeters, gilles van assche and ronny van keer, [https://keccak.team/sponge\\_duplex.html](https://keccak.team/sponge_duplex.html), ??? [Accessed 26-12-2024].
- [36] M. J. Robshaw, Stream ciphers, RSA Laboratories 25 (1995).
- [37] D. J. Bernstein, Salsa20, eSTREAM, ECRYPT Stream Cipher Project, Report 25 (2005) 2005.
- [38] D. J. Bernstein, The chacha family of stream ciphers, DJ Bernsteins webpage: <http://cr.yp.to/chacha.html> (2008).
- [39] D. J. Bernstein, Protecting communications against forgery, *Algorithmic Number Theory*, J. Buhler and P. Stevenhagen (Ed.), to appear (2005).
- [40] J. L. Carter, M. N. Wegman, New hash functions and their use in authentication and set equality, *Journal of computer and system sciences* 22 (1981) 265–277.
- [41] Y. Nir, A. Langley, ChaCha20 and Poly1305 for IETF Protocols, RFC 8439, 2018.
- [42] A. Langley, W.-T. Chang, N. Mavrogiannopoulos, J. Strombergson, S. Josefsson, ChaCha20-Poly1305 Cipher Suites for Transport Layer Security (TLS), RFC 7905, 2016.
- [43] D. Miller, The chacha20-poly1305@ openssh. com authenticated encryption cipher draft-josefsson-ssh-chacha20-poly1305-openssh-00, 2015.
- [44] R. Wille, R. V. Meter, Y. Naveh, Ibms qiskit tool chain: Working with and developing for real quantum computers, 2019 Design, Automation & Test in Europe Conference & Exhibition (DATE) (2019) 1234–1240.
- [45] Pycryptodome Team, Pycryptodome: a Python library for cryptographic primitives and recipes, <https://pypi.org/project/pycryptodome/>, 2022.
- [46] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, T. E. Oliphant, Array programming with NumPy, *Nature* 585 (2020) 357–362.
- [47] J. Newmarch, J. Newmarch, Ffmpeg/libav, Linux sound programming (2017) 227–234.
- [48] MATLAB version 9.0.0.341360 (R2016a), The Mathworks, Inc., Natick, Massachusetts, 2016.
- [49] C. E. Shannon, A mathematical theory of communication, *The Bell system technical journal* 27 (1948) 379–423.
- [50] J. F. Clauser, M. A. Horne, A. Shimony, R. A. Holt, Proposed experiment to test local hidden-variable theories, *Phys. Rev. Lett.* 23 (1969) 880–884.
- [51] a. i. nurhadi, n. r. syambas, quantum key distribution (qkd) protocols: a survey, in: 2018 4th international conference on wireless and telematics (icwt), ieee, pp. 1–5.
- [52] s. j. devitt, w. j. munro, k. nemoto, quantum error correction for beginners, *reports on progress in physics* 76 (2013) 076001.
- [53] k. nagata, t. nakamura, quantum cryptography, quantum communication, and quantum computer in a noisy environment, *international journal of theoretical physics* 56 (2017) 2086–2100.