IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 26, 2018, accepted September 6, 2018, date of publication September 24, 2018, date of current version March 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2871107

Roadmap to Talking Quantum Movies: a Contingent Inquiry

ABDULLAH M. ILIYASU, (Member, IEEE)
Electrical Engineering Department, College of Engineering, Prince Sattam Bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
School of Computing, Tokyo Institute of Technology, Yokohama 226-8502, Japan
School of Computer Science and Engineering, Changchun University of Science and Technology, Changchun 130022, China
e-mail: a.iliyasu@psau.edu.sa

ABSTRACT Quantum computation and quantum information science have transmuted from classroom conjectures and demonstrations usually confined to the laboratory into exciting sub-disciplines with potential for profound impacts in information and communication security, device miniaturization, cosmology, Internet, and so on. Recent developments in both areas have further changed the landscape of future computing and information processing technologies. Encouraged by this, both fields have witnessed interests from professional bodies, notably, the Institute of Electrical and Electronics Engineers (IEEE) via the IEEE P7130 working group on standards for quantum computation as well as similar interests from big technology companies (Google, Microsoft, IBM, etc.) as well as those from national and regional (China, European Union, United Kingdom, Australia, Japan, and so on) governments. With all these activities, presently, interest in and funding for quantum computation and quantum information science are at their highest points ever, thus heralding the inevitable feat of physical realization of quantum computing hardware. Considering the ubiquity and indispensability of imaging and video technologies in our lives today, it is paramount that the utility of any future computing paradigm be measured in terms of its ability to advance and guarantee perpetuity in terms of the efficacy of such technologies. Quantum image processing (QIP), a sub-discipline mainly focused on using quantum computing hardware for the attainment of the aforementioned objectives, has emerged as one of the most promising fields in the quantum computing and quantum information science domains. Its extension to realize quantum movies (or videos), which is the main subject of this disquisition, is an even newer prospect. Consequently, the main objectives of this review are twofold. First, targeting enthusiasts and beginners in the quantum computing field, the treatise provides both an essential background on the rudiments of QIP as well as an exhaustive overview of all literature related to quantum movies and/or videos. Second, for the benefit of those already pursuing research in the QIP sub-discipline, the commentary provides useful insights, allegories, and conjectures to support the integration of other media, such as quantum audio and so on, for the realization of colored talking quantum movies (or video). On the whole, the review is intended to serve as an avenue to invigorate activities tailored toward accelerating the realization of QIP-based quantum technologies.

INDEX TERMS Quantum computation, quantum image processing, quantum information science, quantum movie, quantum video, talking quantum movies, quantum audio.

I. INTRODUCTION

Quantum computing and quantum information science are already living up to their promise to herald the realization of technologies with unrivalled competition in terms of computing speed, tamper-proof security, minimal storage requirements, etc. Evidence of certitude for these capabilities can be traced to the renewed interest and funding in numerous sub-disciplines within both fields.

Recently, these interests have invigorated efforts and investments from the big technology companies to national governments birthing numerous industry-academia partnerships aimed at shaping these inevitable future technologies. Of particular interest in the race for quantum supremacy are: IBM[1] (international business machine)'s collaborative effort with MIT (Massachusetts institute of technology), which is aimed at realizing qubit-wise scalable quantum hardware [1] as well as deployment of such technologies in cloud computing; the Google–NASA (national space agency)–USRA

[1]Readers should refer to page 48 for a detailed list of abbreviations.

(united states space research association) collaboration is exploring the use of quantum computing to tackle problems related to optimization, sampling, search, etc. which are helping to sustain earlier academia-based inquisition into the prospects of quantum machine learning and quantum artificial intelligence [2]. Similarly, Microsoft is involved in an ambitious programme to realize scalable, end-to-end quantum technologies via their promising full-stack quantum computing simulation software [3]. Another game-changer that is certain to impact on the quantum revolution is the Chinese government funded facility whose exploits have led to the establishment of two quantum infrastructure projects. In quantum communication, this effort has produced the ''first long-distance landline quantum communication'' covering distances never reached before [4]. In the emerging sub-discipline of quantum internet, this initiative launched a quantum communication satellite into space and, with that feat, they demonstrated the feasibility of full-scale quantum key distribution [4]. All these efforts have further galvanized interest in deploying quantum technologies in hitherto unchartered areas of cosmology, climatology, etc. [106].

The aspirations and successes enumerated above by both the private and public enterprises (and many others not mentioned here) as well as their collaborations ensure that quantum computing and quantum information science will ''forever alter our economic, academic, and societal landscape [3].'' Quantum technologies will impact on healthcare, environmental systems, communications, banking, smart materials, etc. [106].

In his well-celebrated article [5], Landauer posited that ''information is inevitably tied to some physical embodiment.'' As humans, we have taken this principle to new levels in the manner that we have become accustomed to multi-media (text, image, video, audio) modes of communication as integral parts of our daily lives (via smart, internet-ready devices). The indispensability as well as ubiquity of such communication using sophisticated devices makes it unfathomable to imagine life without them. As we reach the triple-S (speed, security, and size) limits of today's technologies, it is certain that quantum computing will shape future technologies for any and all such communication purposes.

Quantum image processing (QIP) is an emerging sub-discipline that is focused on extending conventional image processing tasks and operations to the quantum computing framework [6]. It is primarily devoted to ''utilizing quantum computing technologies to capture, manipulate, and recover quantum images in different formats and for different purposes [6]''. Due to some of the astounding properties inherent to quantum computation, notably entanglement and parallelism, it is anticipated that QIP technologies will offer capabilities and performances that are, as yet, unrivalled by their traditional equivalents. These improvements could be in terms of computing speed, guaranteed security, and for minimal storage requirements, etc.

In the remainder of this section, we present a brief overview of the varied efforts to link (or utilize) quantum computing to (or in) image processing. Additionally, a summary of the differences between a few of the prominent quantum image representations (QIRs) will be presented.

A. QUANTUM IMAGE PROCESSING

As recounted in [7], retrospectively, the first published material relating quantum mechanics to image processing can be traced to Vlasov's work [8], in 1997, which focused on the ''use of a quantum system to recognize orthogonal images''. This was followed by efforts at using quantum algorithms to search specific patterns in binary images [9] and to detect the posture of certain targets [10]. Following this were efforts to design quantum algorithms based on unstructured picture (data set) in [9] and [10]. Meanwhile, somewhat simultaneously, optics-based descriptions for quantum imaging were already becoming prominent [11], including a more general quantum signal processing framework presented in [12]. However, it is the exploratory work by Beach *et al.* [13], in 2003, that is often credited as the bellwether of what has transmuted into today's description of QIP, albeit their contribution was with a slightly different acronym – quip. Almost simultaneously, this was closely followed by Venegas-Andraca and Bose's Qubit Lattice description for quantum images [14], [72]. Years later, Lattore [15] proposed the Real Ket representation to encode quantum images. The third and final representation, completing the trio of representations that are collectively regarded as the pioneers of today's version of the QIP sub–discipline, is Le *et al.*'s FRQI (flexible representation for quantum images) that was proposed in 2010 [16] and later revised in 2011 [17].

As enumerated in [6], technically, literature relating quantum mechanics to image processing are classified into three groups viz:

a. Optics-based quantum imaging (OQI)
b. Quantum-assisted digital image processing (QDIP)
c. Classically-inspired quantum image processing (QIP)

As further observed in [12] and [18], the former, i.e. Optics-based quantum imaging (in (a)), involves ''investigating the ultimate limits of optical imaging that are allowed by the laws of quantum mechanics'' which is built on the use of techniques in optical imaging and parallel information processing to realize ''image objects with resolutions beyond what is possible using classical optics [12]''. Literature in the quantum-assisted digital image processing (DIP) group (in (b)) are focused on exploiting some of the properties responsible for the potency of quantum computing algorithms to improve some well-known DIP tasks and applications. Finally, the classically-inspired QIP group (in (c)) comprises of studies whose applications derive their inspiration from the expectation that quantum computing hardware will soon be physically realized; and, hence, such research focuses on extending classical image processing (DIP) tasks and applications to the quantum computing framework [6].

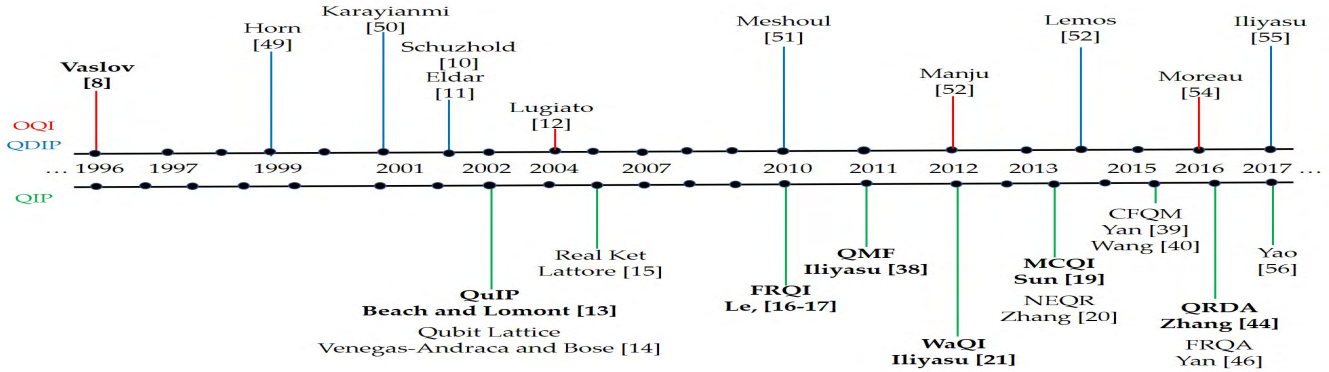

FIGURE 1. Timeline of development of eclectic literature in or relating quantum computing to image processing. Milestone contributions, i.e. work pioneering aspects of the three classes of literature (red for OQI, blue for QDIP and green for QIP) and/or those with high citations are shown highlighted in bold.

Although the approaches listed in groups (a) and (b) are deemed outside the purview of this review, for perspective and comprehensiveness, we present a timeline of notable activities in the growth across all three fields. As presented in Fig. 1, all milestone accomplishments in the three quantum image processing approaches are shown in bold.

Since its conception in 2003, the QIP sub–discipline has witnessed impressive development and is today an active research area being pursued by researchers and groups across China, Japan, Saudi Arabia, Iran, India, Romania, the united Kingdom, etc. Among others, QIP literature are either focused on using quantum mechanical properties of the information carrier in representing images as FRQI, MCQI (multi-channel representation for quantum images) [19], NEQR (novel quantum representation for digital images) [20], etc. states, as well as in information security protocols, such as watermarking [21]–[25], steganography [26], [27], [42], encryption [28], [29], cryptography [30], etc. and other general applications such as image data–base search [31], similarity assessment [32], fidelity measurement [7], image storage and segmentation [33], image scrambling [35], etc. Overall, as mentioned earlier, the objective of QIP is to utilize quantum computing technologies to capture, manipulate, and recover quantum images in different formats and for different purposes. Whereas detailed reviews and surveys on QIP representations, security protocols as well as their applications can be found in [6], [18], [34], and [37], in the sequel the formulation of a few quantum image representations is reviewed with focus on emphasizing their differences and suitability for varied applications. In particular, attention will be given to the FRQI and NEQR quantum image models because most of the quantum movie formats discussed in later sections of the disquisition are built around them.

B. QUANTUM IMAGE REPRESENTATIONS

Within the past decade, a handful of quantum image representations (QIRs) have been proposed to encode an image based on stipulations of quantum mechanics. Of this number, in terms of other quantum image processing protocols and algorithms, including the quantum movies that this review is focused on, three quantum image models have emerged as the most used ones.

In terms of structure and mathematical formulation the three image models – MCQI and NEQR as well as their precursor – Le *et al.*'s FRQI (flexible representation for quantum images) image model – are closely related. More so, in terms of the FRQI and MCQI, which also exhibit the same complexity.

The FRQI model offers a mechanism to encode an image as a normalized state that captures chromatic and spatial details about the image's composition. More than any other QIP image model, the FRQI has played an important role in the development of several strategies to execute quantum image processing tasks [6], [18], [34] and overall growth of the QIP sub-discipline.

On its part, the MCQI model builds on the formulation of the FRQI with focus on encoding the image in terms of R, G and B color channels. Therefore, structurally, unlike the one qubit allocated for chromatic descriptions of an image in the FRQI model, the MCQI requires 3 qubits to encode the chromatic details of an image.

Meanwhile, the NEQR offers a slightly different form of representation since it uses two state sequences to represent the grey level and spatial location of each pixel in an image. Unlike the MCQI quantum image model, but similar to the FRQI model, in the NEQR format, one qubit is assigned for chromatic information about an image. Based on its overall structure, in the NEQR image model, $q + 2n$ qubits are required to fully encode a $2^n \times 2^n$ image with $2q$ grayscale levels. Another important difference between the NEQR model and the more closely aligned FRQI and MCQI formats is in terms of the NEQR's ability to capture different shades of grey. This is attributed to its use of basis states to represent grayscale values, whereas both the FRQI and MCQI models accomplish this using amplitude of probability values [73].

Other QIRs that are built on modifications of the trio of the FRQI, MCQI and NEQR highlighted here include, Caraiman and Manta's CQIR (color quantum image representation) [108], Zhang *et al.'s* QUALPI (quantum representation for log-polar images) [109], Yuan *et al.'s*

TABLE 1. Comparison of prominent QIRs (where *L*, *m*, *n*, etc are parameters associated with dimensions of the images in each model) (table adapted from [18]).

S/No.	QIR	Chromaticity	Complexity	Recovery
1	FRQI [16, 17]	1 angle vector → grayscale	$O(2^{4n})$	Probabilistic
2	MCQI [19]	3 angle vectors → RGB	$O(2^{4n})$	Probabilistic
3	NEQR [20]	qubits sequence → grayscale	$O(qn \cdot 2^{2n})$	Deterministic
4	CQIR [108]	qubit sequence → grayscale	$O(log_2\ L\ \cdot n \cdot 2^{2n})$	Deterministic
5	QUALPI [109]	qubit sequence → grayscale	$O(q(m+n) \cdot 2^{m+n})$	Deterministic
6	SQR [110]	1 angle vector → infrared	$O(2^{2n})$	Probabilistic
7	QSMC-QSNC [33]	1 angle vector → grayscale/RGB	$O(2^{4n})$	Probabilistic
8	NAQSS [111]	1 angle vector → grayscale/RGB	$O(log2^{n}\ \cdot 2^{2n})$	Probabilistic

SQR (radiation energy quantum image model) [110], Li *et al.'s* QSMC-QSNC (M-colors, N-coordinates representation) [33], Li *et al.'s* NAQSS (qubit normal arbitrary quantum superposition state) [111], etc.

Table 1 presents a summary of important features that differentiate some of the main QIRs that have been highlighted. However, we point out that this list is by no means exhaustive. Nevertheless, together with the additional details presented in subsequent sections of this review, sufficient background required to appreciate the developments recorded in the quantum movies (or video) would have been covered. However, readers interested in additional details and differences regarding the QIRs can refer to the detailed survey presented in [18] .

C. QUANTUM MULTIMEDIA PROCESSING

Similar to the growth of traditional image processing and, with it, the emergence of digital videos, in 2011, Iliyasu *et al.* [38] adopted a generalization for realizing quantum hardware to propose a framework to represent and produce movies on quantum computers. Built on the FRQI quantum image model, their proposed quantum movie framework (QMF) was limited in terms of its lack of chromatic exquisiteness and aesthetical candor. Nevertheless, the QMF's structural design, conjectures, nomenclatures, and descriptions have withstood the test of time and they remain the foundations of more recent models for quantum movies. Both Yan *et al.*'s (MCQI-based) [39] and Wang's (NEQR-based) [40] chromatic frameworks for quantum movies (CFQM) relied on the stack-oriented strip formulation of the original QMF [38] to encode the multiple frames that make up the registers encoding the contents of their movies. Furthermore, all the security protocols [41], [42] and applications [39], [40], [43] emanating from the three models of quantum movies relied on manipulations to transform contents of the key, makeup or viewing frames that were originally formulated in Iliyasu's QMF. Moreover, the ability to accomplish security-based and general applications using the three models of quantum movies support the feasibility of their implementation when quantum computing hardware are developed for such purposes.

As precursor and natural extensions to the silent quantum movie models highlighted earlier, the notion and need for quantum audio was first mooted in [38] to support the realization of talking quantum movies. Although still unaccomplished, the necessary first steps, including representations for quantum audio signals have been proposed in [44] and [45] (later revised in [46]). Wang's quantum representation for digital audio (QRDA) [44] holds the title of being the progenitor of quantum audio signal processing (QASP). Using two qubit sequences, the QRDA fuses audio information into quantum mechanical descriptions of quantum information processing. Convinced that the tightly-bounded encoding used in the QRDA could hinder its deployment for certain applications, Yan *et al.* proposed a more intuitive, unbounded and graphic model to represent quantum audio signals called the flexible representation for quantum audio (FRQA) [45], [46]. Formulated to resemble the mathematical rigor of its kindred, the FRQI quantum image model [16], the FRQA provides a conceptual waveform-like description for quantum audio signals. Both quantum audio models (i.e. QRDA and FRQA) support the execution of digital-like audio operations and transformations.

The primary objective of this review is to provide a compendium showcasing some of the gains made in the QIP sub-discipline, notably its offshoot, the quantum movie frameworks (i.e. the QMF and CFQMs), as well as the relatively newer QASP for the benefit of quantum computing and QIP community. Fig. 2 shows the different aspects of QIP and QASP that are covered in the treatise.

As suggested by Fig. 2, this disquisition is broadly divided into three parts. The first part highlights important aspects of every published literature pertaining to QIP-based descriptions for quantum movies. This part is presented in Sections II, III, IV and V, which cover discussions on the QMF, its chromatic extension – the two CFQM color movie models, as well as security protocols and general applications built on the three models.

The second part of our exposition (Sections VI and VII) outlines the progress recorded in quantum audio signal processing (QASP), notably, the rudiments of representing audio content on quantum computers as well as conjectures, allegories, etc. supporting use of the resulting QRDA

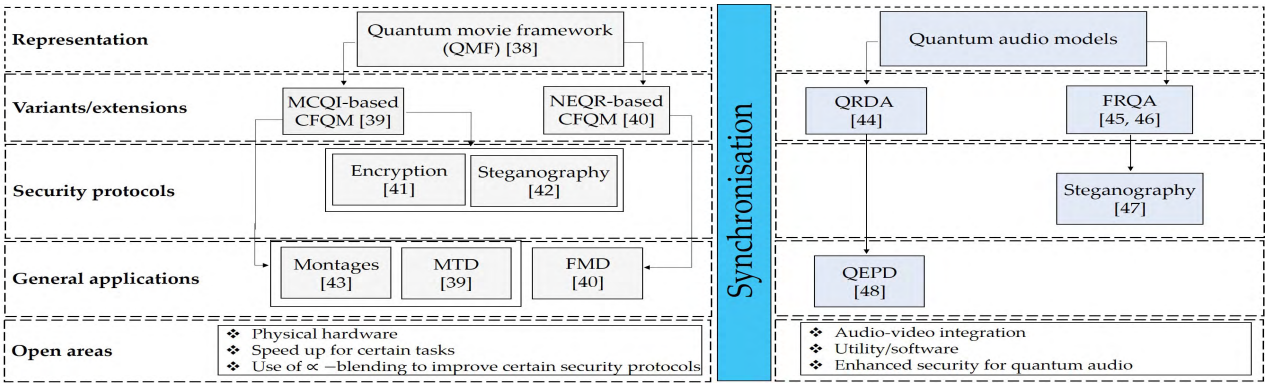

FIGURE 2. Mapping of published literature on quantum movie frameworks and quantum audio models.

and FRQA quantum audio models in different audio signal processing applications. Following that, in Section VII, efforts to utilize both quantum audio models (i.e. QRDA and FRQA) in quantum audio steganography (QAS) [47] and quantum endpoint detection (QEPD) [48] are recounted. Finally, the third part of the review (Sections VIII and IX) presents a brief inquisition regarding the requirements and technologies needed to support the realization of talking color quantum movies. Focus is given to assessing the state of present advances in QIP, QMF, CFQM and QASP to facilitate effective synchronization of quantum movie and audio contents for this purpose.

For enthusiasts and beginners in the quantum computing field, the treatise provides both an essential background on the rudiments of QIP as well as an exhaustive overview of all literature related to the possible evolution of quantum movies and/or videos from concepts engrained in QIP. For the those already pursuing research in the QIP subdiscipline, the commentary provides useful insights, allegories, and conjectures to support the integration of other media such as quantum audio, etc. for the realization of colored talking quantum movies (or video). On the whole, in keeping with momentum and interest in quantum computing generally, and QIP in particular, the review is intended to serve as an avenue to galvanize activities tailored towards accelerating the realization of QIP-based state-of-the-art technologies.

II. FRAMEWORK FOR REPRESENTING AND PRODUCING MOVIES ON QUANTUM COMPUTERS

Quantum computers are expected to harness the strange ability of sub-atomic particles to exist in more than one state at any time. As demonstrated earlier, quantum computers have exhibited huge potential in terms of many sophisticated applications. Consequently, to many, the full-scale realization and/or deployment of quantum computing as complete computing systems or parts of advanced quantum-classical hybrid devices is a matter of when not if.

Motivated by the dominance of TV and movies in today's societies, the huge revenue accruing from these industries, as well as the inevitable role of quantum mechanics in our future computing technologies and the gains recorded in the QIP sub-discipline (a few of which were highlighted in the preceding section), in 2011, Iliyasu *et al.* proposed a theoretical framework to represent and produce movies (or as we shall interchangeably refer to them – quantum videos) on quantum computers [38]. To circumvent violation of some well-established conjectures regarding prohibition of quantum state copying, etc. [57], this scheme comprises of three conceptual devices (viz. the quantum movie disc (QMD), quantum movie player (QMP), and quantum movie reader (QMR)) that are coalesced into one device to meet the generalized requirements for quantum computing, i.e. preparation, manipulation, and measurement. In Iliyasu's quantum movie framework (or simply QMF as it would be referred to henceforth), multiple quantum images are stacked together and encoded as frames of a movie strip and, based on it, the maneuver of content transition from one scene to another is conveyed to the audience. For emphasis, as was clarified in [38], by preparation a QMD composed of a scalable physical quantum state that has been initialized and can exhibit all the inherent properties of a quantum system such as decoherence, entanglemnt, superposition, etc., is implied. Similarly, unless stated otherwise, throughout the remainder of this review, by measurement, non-destructive measurement, whereby the quantum state is recoverable upon the use of appropriate corrections and ancillary information, is implied. Therefore, our conceptual QMR device for reading (or viewing) the content of the movie is formulated with this kind measurement in mind. As stressed in [38], the ''conceit backing the QMF hinges on the assumption that it is possible to realize standalone components to satisfy each of the criteria mentioned earlier (i.e. those required to prepare, manipulate and recover quantum information), and that the components can interact with one another as does the CPU (i.e. central processing unit), keyboard, and display (monitor) units of a typical desktop computer.''

The properties of the aforementioned devices which are the pivots of Iliyasu's QMF and their interaction are highlighted in the sequel.

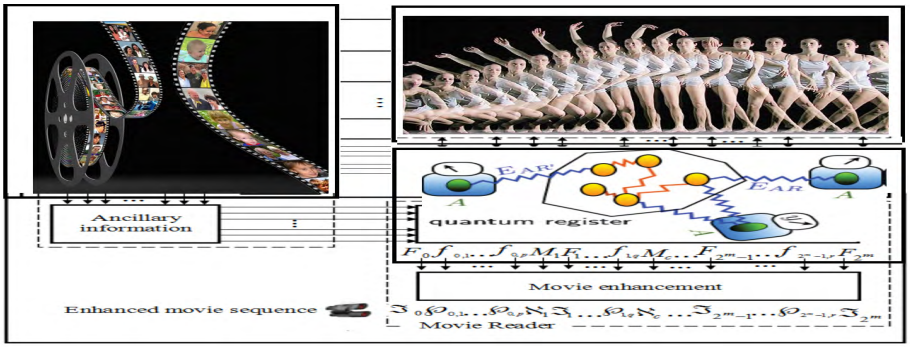

FIGURE 3. Generalization of Iliyasu *et al.*'s quantum movie framework described in [6] and [38] showing the three units of the QMF (Clockwise: the QMD, QMP and QMR).

A. LAYOUT OF QUANTUM MOVIE FRAMEWORK (QMF)

Classically (i.e. on the digital or non-quantum computing paradigm), a movie comprises of a collection of multiple images and every such movie was at some stage a script, i.e. merely a collection (usually in writing) of various predetermined dialogs and instructions required to convey a storyline to the audience [38]. To convey this larger narrative of a movie, four levels of detail are required [58]. At the lowest level, a movie consists of a set of almost identical images called frames, which, at the next level, are grouped into shots. Each shot is delineated by two or more frames that exhibit very little resemblance to each other called keyframes [58]. Consecutive shots are then aggregated into scenes based on their pertinence, such that a scene could have a single shot, and usually all the shots in a single scene have a common background. A sequence of all scenes together composes a movie. Various people are involved at different stages of making a movie. To breathe life into the script, the director translates its content into the various levels of the movie as enumerated earlier on. To accomplish this, he/she relies on different personnel some of which are visible in the movie—the cast—and others (the crew) who although not visible in the final movie are indispensable toward its realization.

Allegorically, authors of the QMF model assumed that these classical terminologies and roles can be extended to their proposed representation and production of movies on quantum computers (i.e. on the QMF [38]). Furthermore, they observed that the inherent nature of quantum computers imposes the need for the services of an additional professional whom they referred to as the *circuitor*. His/her responsibilities complement that of the director, in that, he/she is saddled with choosing appropriate circuit elements to transform each keyframe into a shot (or part thereof), so as to combine with others in order to convey the script to the audience. Alternatively, this role could be added to the director of a quantum movie in addition to his traditionally classical duties.

In line with the aforementioned abstract requirements to represent a movie, a conceptual device or gadget was proposed to encompass each generalized (broad) requirement for quantum information processing, i.e. preparation, manipulation, and measurement [38]. Accordingly, the QMF assumes use of a compact disc-CD (or cassette)-like device to store its most basic information - the keyframes. This proposed device, which we shall refer to as the quantum movie disc or simply QMD, has the "additional capability to prepare, initialize, and store as many keyframes and their ancillary information as required for the various scenes of the movie [6], [38]." Thenceforth, the content of the QMD can only be manipulated using a network of quantum circuit elements. At every measurement layer of such circuits, the transformed version of a keyframe produces a viewing frame: which is a "frame that is marginally different from the frame content at the previous layer [38]." The vastly interconnected circuit network that produces these seamless transitions in the movie is housed in a single device referred to as the quantum movie player (or QMP) in analogy with the classical CD (video compact disc - VCD, digital video decoder - DVD, or video cassette recorder - VCR) player. The combination of two extreme keyframes and the resulting viewing frames that "gradually interpolate between them produces a sequence, which conveys a scene from the movie [6], [38]." The last device required to facilitate the representation and production of a quantum movie is the quantum movie reader (or QMR). As the name suggests, this device measures the contents of the sequence (comprising of the key, makeup, and viewing frames) to retrieve their classical readouts. At appropriate frame transition rates, this sequence creates the impression of continuity as in a movie. Although conceived as separate units, the trio of the QMD, QMP, and QMR, combine to produce the framework to represent and produce movies on quantum computers proposed by Iliyasu *et al.* [38] and reviewed in [6]. An outline of this framework (as highlighted) is presented in Fig. 3.

Succinctly put, as well as pioneering literature in quantum movies (and/or videos), as stated therein, the main contributions of the work in [6] and [38] include provision of:

- allegorical descriptions for movies using quantum mechanical properties inherent to quantum computing hardware;
- an architecture to encode multiple images (keyframes) as a single coherent superposition of the key frames called the strip. This is prepared, stored, and initialized in the QMD;

- theoretical considerations for requirements needed to transform the content of each keyframe to convey part of a movie. These operations transform the keyframe to produce an unbroken sequence of frames needed to depict simple 2D motion operations (such as walking, running, jumping, or movement about a circular path, etc.) according to a movie script;
- a frame transition operation to allow for seamless transition from one frame to another and facilitate classical-like playback/forward operations on quantum video. These operations together with preceding ones are processed by the QMP;
- a mechanism to recover the classical readout of the movie sequence comprising of the keyframes and the new viewing frames realized from their transformation using the QMR.

For clarification, it should be emphasized that the quantum movie representation and production is achieved by assuming concatenation of the three movie components into a single gadget [38]. Furthermore, based on the QMF, the QMD and QMP facilitate capturing the contents of the keyframes as well as their transformation into viewing frames that together combine to represent the bits and pieces of action required to convey the script to the audience [38]. The constraints imposed by the quantum-classical interaction of data, more specifically, quantum entanglement and superposition, ensure that the content as produced and manipulated by the QMD and QMP cannot be readily viewed by the audience. Therefore, the movie reader (QMR) is required to 'decode' these frame (i.e. key, makeup, and viewing) contents in such a way that the earlier constraints are not unsettled, i.e. via non-destructive measurement-based techniques discussed in [38].

We further emphasize that while the components have been designed as standalones, it is assumed that interaction between them is both feasible and mandatory (readers should see Section VIII for this author's insights in this regard). Notwithstanding the interaction, however, the standalone feature of the three movie components guarantees multiple usage of the data stored and processed by the various components.

Later, in other sections of this review, we will highlight efforts made to extend or apply the QMF for other uses and frameworks. Additionally, requirements for realizing efficient movie representation and production including the integration of audio to produce talking quantum movies and likely challenges that may mitigate their realization will be assessed.

In remainder of this section, however, we shall highlight the essential properties of the three devices (QMD, QMP and QMR) that make up the QMF.

B. QUANTUM MOVIE FRAMES, MOVIE STRIP, AND QUANTUM MOVIE DISC (QMD)

A movie comprises of multiple images (called frames) that consist of slight changes from one frame to the next, which combine to produce a single shot [6], [38], [58]. Effectively conveying these changes effectively depends on the appropriate choice of the representation used to encode the content of the keyframe so that it allows the use of specific transformations. In addition, such a representation should also be flexible enough to isolate a smaller region of interest (ROI) within a keyframe so that smaller operations can be depicted as dictated by the movie script. In addition to this ability to isolate a smaller ROI, the representation sought should possess the capability to encode as well as support the manipulation of two additional features: the geometric (position) and visible (color) information of every point in that frame [38]. The flexible representation for quantum images (FRQI) proposed by Le *et al.* [16] and reviewed in [17] exhibits these three broad properties and was considered suficient for representing a single keyframe.

The overall requirements of a movie, however, transcends a single frame encoded using the FRQI image model and so, realizing an effective representation to produce movies on a quantum computer requires a representation capable of capturing the content of each keyframe individually, and a combination of many keyframes as required for each scene of the movie is essential. Like in the QMF [38], unless stated otherwise (such as later in Sections III and IV), throughout this treatise, all references to an FRQI image will imply a keyframe and vice versa. In either case (for the remainder of this section only), a constant background, binary image is implied. In the ensuing discussion, such a binary image will be characterized by two binary levels — black and white. As we shall see later, while restricted in application, such images are of particular interest because they are relatively straightforward to process and, therefore, they provide a useful starting point for recovering contents of the quantum movie. Accordingly, a state $|0\rangle$ shall be assigned to every white point on the 2D grid, while a black point corresponds to a state $|1\rangle$.

Based on these clarifications, we proceed to recount a few definitions considered pertinent to any description of a QMF. This includes clarifications that will assist to distinguish between a keyframe and other types of frames — notably, the makeup frame, which together form the main building block of the quantum movie disc (QMD), the movie strip, and to some extent the entire QMF framework.

Definition 1: A keyframe (i.e. an FRQI quantum image as defined in Equation (1)) is the basic unit of a quantum movie that captures the broad content or information required to convey a single shot (or a part of it) [6], [38].

$$|F_S(n)\rangle = \frac{1}{2^n}\sum_{i=0}^{2^{2^{2n}-1}} \left|c_{S,i}\right\rangle \otimes |i\rangle \tag{1}$$

for $2^n \times 2^n$ movie frames, each an FRQI quantum image whose color component is defined as:

$$\left|c_{S,i}\right\rangle = cos\,\theta_{S,i}\,|0\rangle + Sin\,\theta_{S,i}\,|1\rangle \tag{2}$$

and

$$\theta_{S,i} \in \left[0, \frac{\pi}{2}\right], \quad i = 0, 1, \ldots, 2^{2n} - 1 \quad \text{and } s = 0, 1, \ldots, 2^m - 1 \quad (3)$$

As presented in [6] and [38], the resourcefulness of the QMD lies in the use of quantum circuit elements to interpolate the missing content that connects the shots in the hitherto abstract or summarized version of the movie, i.e. the script. Each shot is characterized by at least a single keyframe, whereas the last one could signal either or both the end of a preceding shot and the start of the present one. In conveying the content of a movie, the transition of its content is gradual both in the change to depict motion (or movement) and in time. When one or more keyframes are set, the motion (as dictated by the movie script) generates the in-between frames, which result in a smooth change (or transition) of the content over time. Moreover, the keyframe representation of the QMF is a simple, yet effective way for summarizing the content of video for content browsing and retrieval [58]. The in-between or missing content realized from each keyframe (called a viewing frame) facilitates a more gradual transition from one keyframe to another. Where a scene in a movie cannot be adequately conveyed by transforming a preceding keyframe, the third type of frame — the makeup frame, is prepared and abutted in the movie sequence. The main difference between a keyframe and a makeup frame is that viewing frames cannot be realized from a makeup frame. In other words, makeup frames do just that — i.e. they make up for missing content within the movie sequence [38]. Based on earlier distinctions between a key and viewing frame, it is underlying to further clarify that, within a movie sequence, a makeup frame can only be preceded or succeeded by a keyframe or another makeup frame. Very often, especially where their distinction from keyframes is essential, we shall refer to makeup frames as $|K\rangle$ (instead of $|F\rangle$ that is now exclusively reserved for keyframes). To conclude, we emphasize that, irrespective of its type (i.e. key, viewing, or makeup), based on the QMF, all the frames are FRQI quantum states as defined in Equation (1). Fig. 4 outlines the schematics for the three types of frames as presented and explained [38], while a quantum register comprising of many key and makeup frames is defined in Definition 2.

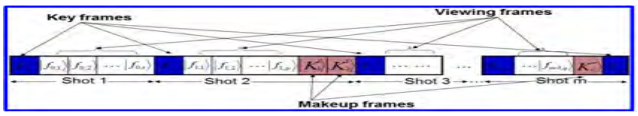

FIGURE 4. ***m*-shots from a movie showing the key, makeup, and viewing frames (figure adapted from [6] and [38]).**

Definition 2: A strip $|S(m,n)\rangle$ representation (first formulated in [38]) is an array comprising 2^m movie frames (key or makeup) each a 2^m FRQI image of $2^n \times 2^n$ dimension constituting a register in the form defined in Equation (4).

$$|S(m,n)\rangle = \frac{1}{2^{\frac{m}{2}}} \sum_{S=0}^{2^{m-1}} |F_S(n)\rangle \otimes |k\rangle \quad (4)$$

where $|k\rangle$ is the position of each image in the strip, m is the number of qubits required to encode the movie frames, $|I_k(n)\rangle$ is a key ($|F>$) or makeup ($|K>$) frame (as defined in Equation (1)) at position $|k\rangle$. Furthermore, as defined in Equation (1), $|c_{k,i}\rangle$ and $|i\rangle$ encode information about the colors and their corresponding positions in frame $|F_k(n)\rangle$. It is trivial to see that $|S(m,n)\rangle$ is a normalized state capable of using only $m+2n+1$ qubits to encode 2^m quantum movie frames (or images).

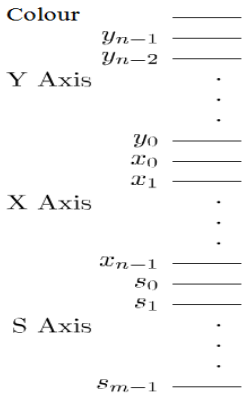

FIGURE 5. Circuit structure to encode the input of a movie strip (figure adapted from [6] and [38]).

As seen in Fig. 5, the depth (in this case relative to the size of the strip) of the strip representation, i.e. the circuit required to capture the input state of the movie strip comprising 2^m key (and makeup) frames increases with the number of such frames [38]. This strip is the input state of the quantum movie disc (QMD) and after its initialization, depending on the location of the strip (S)-axis, the resulting strip could be interpreted as being horizontally- or vertically-oriented [38].

The frame types highlighted here underpin the foremost feature in Iliyasu's QMF, i.e. the strip representation, which fuses the key and makeup frames into a register as defined in Definition 2.

To summarize, each frame (whether key or makeup) is an FRQI state while any abutted combination of multiple key (and makeup) frames — a strip — is best represented as a movie register or strip state.

As discussed in [38], in addition to the 2^m-ending key (and makeup) frames captured and stored (as discussed earlier in this section), auxiliary information that loosely pertains to the color of every pixel in each frame is stored. This ancillary information can be recovered from the script and/or during the preparation of the key (and makeup) frames. In its most simple form, this information conveys chromaticity by specifying whether a pixel is inhabited (i.e. it has a fill) or not which corresponds to states $|1\rangle$ and $|0\rangle$, respectively. Each key (or makeup) frame requires $2^n \times 2^n$ ancillary

information to determine which part is filled (not minding its color) or not. This ancillary information forms the cornerstone on which the dependent ancilla-driven quantum movie reader (QMR) presented later in this section is built.

C. QUANTUM MOVIE PLAYER AND MOTION IN THE QMF

As accentuated earlier, the movie strip representation comprising the key (and makeup) frames encodes information that represents an abstract summary of an entire movie [6], [38]. Therefore, transitions from one key (or makeup) frame to the next are at best abrupt and may not effectively convey the storyline as specified by the script. Hence, there is the need to present a more gradual transition of the content conveyed by two keyframes by way of intercalating the missing content between them as well as abutting all the content for a more harmonious connection between them (i.e. the keyframes and their interpolated content) into a single continuous sequence. The main objective of the quantum movie player (QMP) is to manipulate the content of each keyframe as required to realize viewing frames that capture the required motion or activity in that shot [38]. It is analogues to the central processing unit of the proposed QMF framework wherein the script dictates the use of appropriate circuit elements to manipulate or modify the keyframes (and realize new viewing frames therefrom) as well as makeup frames.

1) SIMPLE MOTION OPERATIONS (SMO)

In any movie representation, depicting the different movements of the various objects that combine to convey the script of the movie to the audience is of paramount importance. In this regard, operations to effectively convey 2D movement of every point (or collectively, every object as an entangled state [59]) in a quantum movie is necessary. For this purpose, the QMF was designed to utilize a multitude of transformations that depict 2D movement of a ROI within a keyframe.

To highlight these operations, we first define the simple motion operation (SMO) in Definition 3.

Definition 3: A simple motion operation (SMO) on an FRQI keyframe is an operation O that shifts every point in that frame as presented in Equation (5).

$$O_{c,2n}\,|\,i\rangle = \left|\,i'\right\rangle, \quad i \in 0, 1, \ldots, y_0, x_0, x_1, \ldots, 2^{2n-1} \tag{5}$$

where $i' = i + c \bmod 2^n$ and $c \in \left\{1, \ldots 2^{2n-1}\right\}$ called the *shift steps* is used to indicate the number of steps every point is shifted.

Definition 3 can be decomposed in terms of SMOs along the horizontal axis as the forward or backward motion operations, which are denoted as M_F^c and M_B^c, respectively; and similarly, along the vertical axis, we obtain the upward and downward motion operations denoted as M_U^c and M_D^c, respectively. The equations and corresponding circuits to execute these operations are presented in Equations (6) through (9) and Fig. 6 respectively.

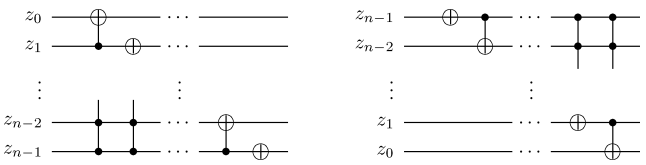

FIGURE 6. Circuits for SMO. Depending on the motion axis = x or y) the circuit on the left is used to accomplish the *M* operations when applied along the x- and y-axis, respectively. Similarly, the circuit on the right is used to accomplish the *M* operations when applied along the x- and y-axis, respectively. (figure adapted from [6] and [38]).

Starting with the forward and backward SMOs, the modified equations are given as

$$M_F^c\,|\,Y\rangle\,\left|X\right\rangle = I\,|\,Y\rangle \otimes O_{c,n}\left|X\right\rangle = |\,Y\rangle\,\left|X + c \bmod 2^n\right\rangle, \tag{6}$$

which shifts every point in the image c steps forward, whereas the SMO shifting the image c steps backward is defined as

$$M_B^c\,|\,Y\rangle\,\left|X\right\rangle = I\,|\,Y\rangle \otimes O_{-c,n}\left|X\right\rangle = |\,Y\rangle\,\left|X - c \bmod 2^n\right\rangle \tag{7}$$

Similarly, along the vertical axis we define the upward and downward SMOs as:

$$M_U^c\,|\,Y\rangle\,\left|X\right\rangle = O_{c,n}\,|\,Y\rangle \otimes I\left|X\right\rangle = |\,Y\rangle\,\left|X + c \bmod 2^n\right\rangle \tag{8}$$

and

$$M_D^c\,|\,Y\rangle\,\left|X\right\rangle = O_{-c,n}\,|\,Y\rangle \otimes I\left|X\right\rangle = |\,Y\rangle\,\left|X - c \bmod 2^n\right\rangle \tag{9}$$

respectively, where I is an identity operation and c specifies the number of positions each pixel in the image is shifted upwards or downwards as the case may be.

From the circuits in Fig. 6, we deduce that the one on the right (in Fig. 6) has $C^{n-1}(\sigma_x)$ controlled-NOT gates on the first layer and $C^{n-2}(\sigma_x)\cdots C(\sigma_x)$ controlled-NOT gates, (where σ_x is the Pauli NOT gate), along the remaining layers. In contrast, the circuit on the left (in the same Fig. 6) has $C(\sigma_x)$, i.e. a NOT gate, on its first layer and $C^{n-2}\,(\sigma_x)\cdots C\,(\sigma_x)\,,\sigma_x$, along its remaining layers. Applied along the vertical axis, the circuits (on the left and right in Fig. 6) produce the M_F^c and M_D^c SMOs, respectively. Similarly, applied along the horizontal axis the same circuits yield the M_B^c and M_U^c SMOs as defined earlier (in Equations (7) and (8)).

Consequently, the number of viewing frames realizable from a key frame depends on:

- The size of the keyframe;
- The size of the ROI targeted by the script; and
- The content of the script, i.e. the specific actions being depicted

As demonstrated in [6] and [38], if used haphazardly, SMOs enervate the content of the script (i.e. a situation referred to as *overflow*), which often erodes or misrepresents

the intended dialogue being conveyed. However, it is envisaged that a well be adept in ensuring that overflows are avoided, or their impact smoothened to depict certain actions such as back and forth movement by an ROI. Readers are referred to [38] where this application has been demonstrated.

The SMO operations reviewed here can be combined with other geometric transformations on quantum images (GTQI) [60], [61] so that more complicated 2D motions (as would be required in a movie) can be realized. As noted in [38], other applications might necessitate the use of color (i.e. CTQI – color transformations on quantum images [61]) operations [62] to accomplish. Definition 4 formalizes the displacement operation that combines two or more SMO operations to accomplish more complex tasks that can displace part or all the content of a keyframe and convey different actions dictated by the movie script [38].

Definition 4: Given a $2^n \times 2^n$ FRQI movie keyframe, the displacement operation to shift the position of every pixel in the entire keyframe (or any part thereof (i.e. an ROI)) c and d steps along the horizontal and vertical axis respectively can be defined as:

$$V_D^{c,d} \left| Y \right\rangle \left| X \right\rangle = \left| Y \pm c \bmod 2^n \right\rangle \otimes \left| X \pm d \bmod 2^n \right\rangle, \quad (10)$$

where D indicates the nature of the final displacements for c and d shifts ($c, d \in \left\{0, 1, \ldots, 2^{2n-1}\right\}$) along the horizontal and vertical axis respectively. Such operations are realized via different combinations of the SMOs defined in Equations (6) through (9).

2) FRAME TO FRAME TRANSITION (FTF)

An important utility of the strip representation proposed in the QMF model [38] is that it allows the simultaneous use of q different sub-circuits each geared toward accomplishing a certain transformation on a preceding frame. Since marginal changes in the content of the resulting frames are expected, seamless movement in ROI can be depicted via multiple movie operations, each using its own sub-circuit. Where $q = 0$, then the transition from one key (viewing or makeup) frame to another, c steps apart, is immediate. To achieve the transition from one frame to the next, we define (in Definition 5) the strip transition operation, T_i^c, which targets the strip (i.e. S-coordinate) only.

Definition 5: The frame-to-frame transition operation T_i^c on a movie strip $|S(m,n)\rangle$ (defined in Equation (4)) is the operation used to shift from one key (or makeup) frame to another in a quantum movie.

$$T_i^c \left| S(m,n) \right\rangle = \frac{1}{2} \sum_{j=0}^{2^m-1} \left| I_j \right\rangle \otimes T_i^c \left| j \right\rangle \quad (11)$$

$$T_i^c \left| S(m,n) \right\rangle = \frac{1}{2} \sum_{j=0}^{2^m-1} \left| I_j \right\rangle \otimes \left| j \pm c \bmod 2^m \right\rangle, \quad (12)$$

where $c \in \left\{0, 1, \ldots, 2^{2n-1}\right\}$ is the shift steps to transit from the n^{th} frame, $|F_n\rangle$, to another frame $|F_{n\pm c}\rangle$, that is c steps up or down the length of the strip.

Transition from one frame to the next requires $c = 1$ shift steps. In addition to aiding smooth transition from frame to frame, the T_i^c operation (in Equation (11)) supports repeating a single frame when the final frame sequence is generated for viewing by the audience. This capability eliminates the need to duplicate the same frame whenever its need elsewhere within the script arises. Additionally, it facilitates the execution of other video processing operations, such as playback (or forward).

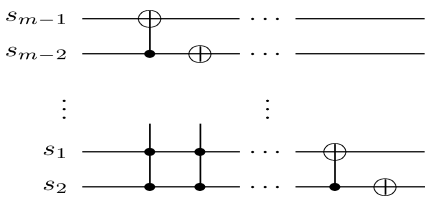

FIGURE 7. QMF circuit to accomplish frame-to-frame transition operation.

Given a script that dictates transformation of the m^{th} keyframe in a movie ($|K_m\rangle$) using q movie operations, then a sequence:

$$SMO \left| F_m \right\rangle = \left| f_{m,0} \right\rangle, \left| f_{m,1} \right\rangle, \cdots \left| f_{m,q} \right\rangle \quad (13)$$

abutting q marginally different viewing frames is obtained after which transition to the next key (or makeup) frame is made. Combining this with other contents of the movie produces a final sequence comprising all the key (and makeup) frames as well as the viewing frames (realized from their transformation of $|K_m\rangle$ as dictated by the movie circuit (i.e. the script)). This sequence, called the movie sequence, $|M\rangle$, is presented in Equation (14).

$$|M\rangle = |F_0\rangle(|f_{0,0}\rangle, |f_{0,1}\rangle, \cdots |f_{0,p}\rangle)|K_1^0\rangle|F_1\rangle, \cdots, \times|F_2^{m-1}\rangle (|f_{2^m-1,0}\rangle, |f_{2^m-1,1}\rangle, \cdots |f_{2^m-1,q}\rangle)|F^{2^m}\rangle \quad (14)$$

It was proven in [38] that the width of the movie circuit ($W(C)$) needed to generate this sequence depends on the number of keyframes (and where required the makeup frames) encoded on the strip and the size of each frame. On its part, the depth of the circuit ($D(C)$) is a function of the output requirements of the movie, in which case it is determined by the number of layers in each sub-circuit. This is in turn dictated by the movie operations necessary to convey the story in the movie script. It is trivial that the cost of preparing a quantum movie $|C|$ is determined by the number of key frames (and makeup frames) required to capture all the major content of the movie, $|m|$; the size of each of the frames $|n|$; the total number of movie operations required to effectively describe the content of the movie, Q; and the number of basic gates required for each operation.

D. QUANTUM MOVIE READER (QMR)

When being written, the dream of every scriptwriter is to see its contents translated into a movie that would convey some storyline to its audience. Similarly, the ultimate objective of

any QMF is to facilitate viewing of its movie content. However, the distinctiveness (i.e. pertaining to its non-copying or cloning of quantum states) of the quantum mechanical framework on which the QMF is built imposes some limitations regarding when, how, and what can be viewed. The aim of the QMR is to provide a platform to recover and observe (or view) the movie contents without unsettling the quantumness of the QMF.

Measurement–based quantum computation (MBQC) [63], [64] was proposed as an ''alternative strategy to the circuit–model and it relies on the effects of measurement on an entangled multi–partite resource state to perform the required computation [38], [63], [64].'' This strategy provides an astute pathway to overcome the perceived shortcomings of the circuit–model of quantum computation and aspects of it ''have been realized experimentally using single–qubit measurements that are considered 'cheap' [38], [63]–[66].''

As disclosed in [38], ''all measurement–based models of quantum computing share the common feature that measurements are not performed on the qubits storing the data.'' This is attributed to the need to ''safeguard the loss of the coherence that is needed for meaningful quantum computation.'' Therefore, ancillary qubits are prepared and interacted with the information carrying qubits, so that the measurement operation is performed on them. By carefully choosing appropriate measurement operations and initial states of the ancillary qubits, the much-sought coherence can be preserved [38], [63].

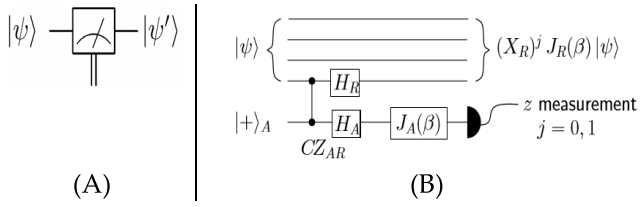

FIGURE 8. (A) Notation for a single qubit projective measurement operation and (B) description of the ancilla-driven measurement operation (figures and explanations in the text are adapted from [6], [63], and [66]).

Fig. 8(A) shows the notation for a single qubit projective (destructive) measurement whence the quantum information is lost immediately its content is observed (or accessed). On its left (i.e. in Fig. 8(B)) is a non-destructive (or ancilla-driven) measurement as used in MBQC, and in this case the ancilla-driven quantum computation (ADQC) model [38], [66].

As at the time Iliyasu's QMF was proposed in 2011, the ADQC [66] was a new type of MBQC, whose adoption as the fulcrum of the QMR of Iliyasu's QMF was considered because it exploits the properties of single–qubit projective measurements as well as the entanglement-based interaction between the ancilla qubit and the register qubit (i.e. strip state encoding the 2^m–ending movie frames). The measurements are performed subject to satisfying some predefined conditions [63]. We remark here that, eventhough the core motivation for using the ADQC still holds, unlike the QMD where efforts to integrate newer (i.e. color) quantum image models have been explored, there has been no effort to redesign the QMR in any other format.

As originally formulated in [63] and used as the core component of the QMR of Iliyasu's QMF [38], the ancilla A is ''prepared and then entangled to a register qubit (in our case of the QMF this corresponds to the single qubit encoding the color information [62]) using a fixed entanglement operator E [38], [63], [66].'' Furthermore, a universal interaction between the ancilla and register is accomplished by using the Controlled-Z (CZ) gate and a swap (S) gate followed by the necessary measurements. It has been established that an ADQC with such an interaction allows the implementation of any computation or universal state preparation [38], [63]–[66]. This is then followed by single qubit corrections on both the ancilla and register qubits.

Like all the measurement-based models, the standard ADQC uses a fully controlled ancilla qubit, which is coupled sequentially to one or at most two qubits of a register via a fixed entanglement operator E [38]. After each coupling, the ancilla is measured in a suitable basis, providing a back action onto the register [66]. This implements both single- and two-qubit operations on the register qubits. Moreover, the ADQC can be viewed as a hybrid of the circuit model, since computation involves a sequence of single and two-qubit gates implemented on a register [38].

The description of Fig. 8(B) as provided via [63] presents the ancilla-driven implementation of a single qubit rotation, $J_R(\beta)$, on a register qubit R, where the initial state of the total register can be a pure state, $|\psi\rangle$, or accordingly for a mixed state. The ancilla and register qubits are first coupled with $EAR = H_{\rm A}\, H_{\rm R}\, CZ_{\rm AR}$. The rotation J($\beta$) is then implemented on the fully controlled ancilla and transferred to the register qubit by measuring the ancilla in the z-basis. The result of the measurement, $j = 0, 1$, determines if an X correction appears on the register qubit [63].

Furthermore, the QMR unit of the QMF model [38] that is being reviewed here, further exploits the preservation of pixel positions in the FRQI representation. This property presupposes that, as a 2D representation, information about each pixel (its 2D grid location and color – i.e. the monochrome or grayscale value) are known beforehand. In other words, based on this property and (in our own case) access to the movie circuit, foreknowledge is assumed beforehand about the position where each pixel in an FRQI quantum image (also a movie frame) resides before and after being transformed. This vital knowledge reduces the amount of information and resources required to recover the transformed image to just the information pertaining to the new color of the pixels (in terms of the presence or absence of a fill in each pixel).

Consequently, by putting it altogether, the under-listed simplifications were utilized in [38] to facilitate recovery of the color of the i^{th} pixel, (c_i) in frame j ($|F_j\rangle$) of a quantum movie.

- An interplay is assumed between the QMD and the QMR. This enables the transfer of the ancillary information about the fill of every pixel as stored in the QMD to the QMR (as discussed earlier). Therefore, [66] simplified the representation for the ancilla states $|+\rangle$ and $|-\rangle$ from $|+\rangle = |0\rangle, |-\rangle = |1\rangle$ for the absence and presence of a fill in that pixel, respectively.
- A universal interaction between the ancilla $|a\rangle$ and register (specifically, the color qubit) is accomplished using the CZ gate and a swap (S) gate and then measured. An ADQC with such an interaction is sufficient for the implementation of any computation or universal state preparation [66], [67], [69]–[71].
- This is then followed by single qubit corrections, $Z^{\dagger}$ and U on the ancillary information and color of the pixel, respectively. The measurement to recover the color of the i^{th} pixel M_i^{θ} returns a value c_i as defined in Equation (2). This is the same as retaining the interaction operation E in the ADQC while replacing the measurement and Pauli corrections with a projective measurement as explained in [67] and [68].
- Finally, adopting the 2D grid position information of every point in a transformed key frame (i.e. an FRQI state), we can determine beforehand the required control-operations needed to recover the transformed keyframe, i.e. the viewing frames.

Using this format and further modifications outlined in [38], the changes in the ancillary information (i.e. whether each pixel in the original image, a keyframe, is inhabited (i.e. having a fill) or not) undergoes can be tracked relative to changes in their colors as they evolve based on the transformations by the movie circuit and/or CTQI operations [62].

While the descriptions presented here would suffice for recovery of the contents of a given pixel, it should be clarified that in recovering the contents of a movie, multiple such measurements are necessary. Definition 6 formalizes our definition of the movie reader as presented in [6] and [38].

Definition 6: A QMF quantum movie reader (QMR) is a system comprising $2^{2n}-1$ ancillary qubits each initialized (as defined earlier in this section) to track the change in the content of every pixel in each frame (key, makeup, or viewing) that forms part of the movie sequence (presented in Equation (15)) based on single-qubit dependent ancilla-driven measurements as described in [38].

The movie enhancement stage of the movie reader is added to show the need to enhance the content of each frame before final display to the audience. This arises because throughout focus has been on color processing tasks in terms of binary ($|0\rangle$ and $|1\rangle$) states. Moreover, this agrees with the intuition in classical image processing where numerous post-processing stages are used for various enhancements to the movie content. At the final stage, integration of numerous classical technologies to recover and enhance the quality of the movie content emanating from the QMF is assumed. The enhancement unit (shown in bottom part of Fig. 3) could also include a sub-unit for tuning the movie sequence to the appropriate frame transition rate (usually 25 frames per second) to be broadcast to the audience. Therefore, this enhanced (i.e. classical readout of the quantum movie) content of the movie can be represented using the sequence in Equation (15).

$$M = \mathfrak{H}_0 \left(t_{0,0}, t_{0,1}, \cdots, t_{0,p}\right) \kappa_1^0 \mathfrak{H}_1, \cdots, \mathfrak{H}_2^{m-1} \times \left(t_{2^m-1,0}, t_{2^m-1,1}, \ldots, t_{2^m-1,q}\right) \mathfrak{H}^{2^m} \quad (15)$$

where $\mathfrak{H}$, t and κ are the *enhanced* key, viewing and makeup frames that are displayed to the audience.

In concluding our review of the pioneering effort to represent and produce movies on quantum computers, we reiterate that while Iliyasu's QMF does not measure with today's state-of-the-art in digital movie and TV production, it presents a modest attempt that advances available quantum image processing (QIP) nomenclature, knowhow, and technologies to descriptions of quantum video and movies.

In doing this, the criteria to realize a quantum system to include preparation, manipulation, and measurement were generalized each as a standalone quantum device supporting the requirements for each criterion. A quantum movie disc (QMD) was proposed for the preparation criteria, wherein, the broad content or keyframes to convey the movie script are encoded, prepared, and initialized as a large superposition of multiple keyframes called a strip [38]. The quantum movie player (QMP) utilizes a set of SMOs, GTQI [60], [61], and CTQI [62] operations to manipulate contents of these keyframes in order to interpolate the missing content between successive keyframes. These operations as required to effectively depict the movements and motions in the scenes and shots of the movie combine to produce the movie circuit. At every measurement layer of the movie circuit, the transformed version of each keyframe produces a viewing frame. Where certain content in a scene cannot be realized by transforming a keyframe, makeup frames are abutted to make up for the hitherto unrepresented content of the movie. The classical sequence comprising of the keyframes, makeup frames, and the interpolated viewing frames in between them are retrieved via ancilla-driven quantum measurement operations in the quantum movie reader (QMR) [38]. At appropriate frame-to-frame transition rates, this sequence conveys the shots and scenes that depict the content of the movie to the audience. Concatenated, these three components together facilitate the proposed framework for quantum movie representation and production, and as enthused in [38], open the door toward using quantum gates and circuits to manipulate quantum information for applications in video processing, animations, etc.

III. COLOR QUANTUM MOVIES

Although it was conceived with in-built sensitivity to the intricacies that pervade information processing on the quantum computing framework, at best, Iliyasu's quantum movie scheme [38] resembles early nineteenth century stages in the developments of contemporary movie production with two limitations pertaining to its usability. First, even without the

pitfalls arising from decoherence and measurement operations to recover the movie contents, based on its formulation, the entire movie is encoded as a monochrome one. Therefore, at best, the resulting movie would be monochrome (or black and white) in nature. Second, since, until recently [44]–[46], there was no such thing as quantum audio; Iliyasu's QMF would produce silent (or mute) movies. For lack of any euphemism, it is incontrovertible that black and white silent movies are not technologies for the 21st century.

To enhance the canonical elegance of the FRQI-based movie scheme in [38], Yan *et al.* proposed a new framework that integrates chromaticity into descriptions of the individual frames (images) and merges that information with the temporal details that tag each frame to the next and binds them all into the quantum register (i.e. a movie strip).

Instead of the FRQI representation that is used to encode the still images that make up the QMF's movie frames, this new effort proposed the use of a chromatic framework for quantum movies (CFQM) by encoding each movie frame as a multi-channel quantum image (MCQI) [19]. The MCQI representation is a color-based extension of the FRQI representation, therefore, its integration will facilitate RGB color modeling into the QMF layout.

Furthermore, Yan's CFQM movie model provides a platform to reformulate the FRQI-based QMF movie scheme in a manner that facilitates the execution of more advanced color transformations, such as frame-to-frame (FTF), color of interest (COI), and sub-block swapping (SBS) operations that are not readily realizable based on the standard FRQI-based movie scheme.

Being a standard extension of Iliyasu's QMF, it is expected that the MCQI – i.e. multi-channel (or colored) movie framework (CFQM) will maintain conformity with earlier structural requirements, definitions and descriptions (in Section II) that now seems widely accepted as the basis for formulation of any QMF. In providing a thorough outline of the CFQM quantum movie framework, in subsequent subsections, we will also highlight all the available color-based QFMs, i.e. including extensions that are based on the FRQI image model as well as those that use the so-called *novel* enhanced quantum representation for digital images (NEQR) [20] to encode information about the movie frames. In addition to accentuating their differences and similarities, their utility in comparison with the pioneer framework for quantum movie production, i.e. Iliyasu's FRQI-based QMF, will be enunciated.

A. MCQI-BASED COLOR QUANTUM MOVIES (CFQM$_M$)

Motivated by the role of color in aesthetics, as well as human inclination towards it, and the human visual system generally, in [19], Sun *et al.* provided an extension for the monochrome FRQI image model to capture color quantum images. They formulated a representation (known as the multi-channel representation for quantum images (MCQI)) that assigns three qubits to encode the R, G, and B channels needed to articulate the chromatic content of an image.

While details of the MCQI including its preparation can be digested from [19], here we formalize it based on the format in [39].

A multi-channel (FRQI) quantum image (MCQI) is defined in format in Equation (16) in terms of its chromatic and geometric information.

$$|I_M(n)\rangle = \frac{1}{2^{n+1}} \sum_{i=0}^{2^{2^{2n}-1}} \left| c^i_{RGB\alpha} \right\rangle \otimes |i\rangle , \tag{16}$$

where the color component of the image (or frame as used in this disquisition) $\left|c^i_{RGB\alpha}\right\rangle$ is further defined in Equation (17).

$$\begin{aligned} |c^i_{RGB\alpha}\rangle = & \cos\theta^i_R|000\rangle + \cos\theta^i_G|001\rangle + \cos\theta^i_B|010\rangle \\ & + \cos\theta^i_\alpha|011\rangle + \sin\theta^i_R|100\rangle + \sin\theta^i_G|101\rangle \\ & + \sin\theta^i_B|110\rangle + \sin\theta^i_\alpha|111\rangle \end{aligned} \tag{17}$$

and $\theta^i_R, \theta^i_G, \theta^i_B \in \left[0, \frac{\pi}{2}\right]$ with θ^i_α set to zero so that the two constant coefficients (i.e. $\cos\theta_\alpha = 1$ and $\sin\theta_\alpha = 0$) carry no information.

As formulated in Equation (16), Sun's MCQI representation provides a platform for direct extension of most FRQI-based QIP applications in terms of colored quantum images [19]. As an example, the QMF (reviewed in earlier sections of this exposition), which was contrived entirely on the FRQI model for encoding the various frames (i.e. key, makeup and viewing) that constitute the movie can be revised to realize colored quantum movies. Whereas most of the earlier definitions and conjectures in Iliyasu's (i.e. FRQI-based) framework for quantum movies still hold, to effectively convey the notion of color quantum movies (or video), it is necessary to reformulate the strip representation in Equation (4) in terms of Sun's MCQI quantum image model.

Furthermore, it was argued in [39] that other motivations for adopting the MCQI image format as the most suitable mechanism to realize the chromatic framework for quantum movies (CFQM$_\mathbf{M}$) include:

- Its flexibility, which supports imposition of additional constrains to confine operations to smaller Regions of Interest (RoI) within a frame (key, makeup or viewing) as required to depict scenes specified by different scripts
- Its adeptness, which provides a platform to capture the location (i.e. position) and RGB color information about every point in each frame

As noted in [38], since the overall requirements of a video transcends a single frame, and in order to accomplish the effective representation for a movie on a quantum computer, a representation capable of capturing the content of each keyframe individually, and a combination of these keyframes as required for each particular scene of the video is indispensable. Therefore, by revising the choice of image model to encode movie frames and adopting Sun's MCQI format [19] for such purposes, while retaining the structural requirements and definitions for building blocks of the quantum movie [38], etc., the description of a color quantum video

(i.e. the CFQM [39]) is formalized in the remainder of this subsection.

Definition 7: A $\text{CFQM}_{\mathbf{M}}$ movie keyframe is an MCQI quantum image (in the form defined in Equation (16)) that captures the broad content from which the additional information required to convey a single shot (or a part of it) in a video is obtained.

Perhaps, the most important proviso of any QMF is the need for a register to capture the multi-frame content of a movie. This stipulates the use of a movie strip, which as a collection of 2^m-ending key and makeup frames, is the plinth to capture information necessary to represent the shots and scenes of a video [38]. All available CFQMs adopt the same descriptions for the strip, as presented in Section II, with the distinction that, in the case of the $\text{CFQM}_{\mathbf{M}}$, each movie frame is an MCQI color quantum image state.

Mathematically, the MCQI-based quantum movie strip ($|S_M(m,n)\rangle$) is obtained by replacing the monochrome FRQI description of the frame ($|F_S(n)\rangle$) in Equation (1) with the multi-channel MCQI representation in Equation (16). Similarly, the circuit layout for the MCQI-based CFQM (or CFQM_{M}) that is presented in Fig. 9 shows the structural representation of a four-frame movie based on the CFQM model. The CFQM_{M} is formalized in Equation (18).

$$|S_M(m,n)\rangle = \frac{1}{2^{\frac{m}{2}}} \sum_{S=0}^{2^{m-1}} |F_M(n)\rangle \otimes |t\rangle \tag{18}$$

where $|F_M(n)\rangle$ is the movie frame, i.e. an MCQI image $|I_M(n)\rangle$ as defined in Equation (17)) and $|t\rangle$ is a time tag that binds each frame to the movie strip.

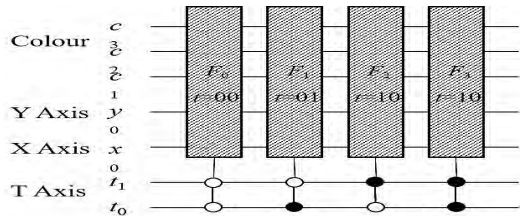

FIGURE 9. Structural representation of a four-frame CFQM_{M} movie [39].

It was observed in Yan *et al.*'s survey on advances in quantum image processing (QIP) [18] that, among others, the flexibility of the FRQI quantum image model (as well as its variants and extensions such as the MCQI) together with the simplicity with which limitless transformations to manipulate its chromatic and spatial content can be executed have seen it emerge as the most preferred quantum image representation (QIR). Many of these applications can be effortlessly extended to the QMF and $\text{CFQM}_{\mathbf{M}}$ quantum movie models.

In addition to this, as a further testimony of the growing interest in QIP and its applications, in the next subsection we will highlight another CFQM that is built on a different image model for the movie frames as well as its handling and manipulation.

B. NEQR-BASED COLOR QUANTUM MOVIES (CFQM_{N})

It was observed in [20] that the FRQI image model is inhibited by two drawbacks related to its accurate retrieval and perceived constraints in terms of color operations. To support the latter claim, Li *et al.* [73] further enthuse that "only a single qubit is used to represent the color information for each pixel". Hence, they further argue that "it is difficult to design quantum circuits for image processing tasks that need to separate pixels of different colors to apply distinct transformations."

Although both claims are disputable, as a solution to these envisaged shortcomings, the so-called *novel* enhanced quantum representation for images (NEQR) was proposed by Zhang *et al.* [20]. Therein, it was proclaimed that, in comparison with the FRQI image model, the NEQR "offers quadratic decrease in terms of requirements for its preparation." Additionally, they surmise that the "NEQR achieves better image compression and retrieval" [20]. Furthermore, [73] assert that, in this model, "images are stored with $(2n+q)$ qubits, where $2n$ qubits describe the pixel position and q qubits describe the grayscale intensity."

To formalize Zhang's NEQR representation, we are told that, supposing a grayscale image has a range 2^q corresponding to a pixel that is encoded using the binary sequence $C_{q-1}^{YX}, C_{q-2}^{YX}, \ldots, C_0^{YX}, C_m^{YX} \in [0,1]\,(m=0,1,\ldots,q-1)\,; f(Y,X) \in \left[0,1,\ldots,2^q-1\right]$, then the representation for a $2^n \times 2^n$ image on the NEQR model is in the form described in Equation (19) [20].

$$|I_N\rangle = \frac{1}{2^n} \sum_{Y=0}^{2^n-1} \sum_{X=0}^{2^{n-1}} |f(Y,X)\rangle\, |Y\rangle\, |X\rangle \tag{19}$$

where chromatic details are represented as basis states of the color qubit and the resulting quantum image is an entangled state fusing the color and position information qubits [20], [73].

Whereas the basic nomenclature and descriptions in Iliyasu's QFM are tenable in Wang's NEQR-based CFQM (or $\text{CFQM}_{\mathbf{N}}$) [40], its formulation of the movie strip is revised (as presented in Equation (20)) to accommodate Zhang's NEQR representation that is used to encode the movie frames.

$$|S_N(m,n)\rangle = \frac{1}{2^{\frac{m}{2}}} \sum_{j=0}^{2^m-1} \left|I_j\right\rangle \otimes |j\rangle \tag{20}$$

where each $\left|I_j\right\rangle$ is an NEQR image as defined in Equation (19) and expanded in the form:

$$\left|I_j\right\rangle = \frac{1}{2^n} \sum_{i=0}^{2^{2n}-1} \left|C_{j,i}\right\rangle \otimes |i\rangle = \frac{1}{2^n} \sum_{i=0}^{2^{2n}-1} \left|C_{q-1}^{j,i} C_{q-2}^{j,i} \cdots C_0^{j,i}\right\rangle |i\rangle, \tag{21}$$

and the color information comprises of 8 qubits encoding the intensity of each pixel as follows:

$$C_s^{j,i} \in \{0,1\}, \quad s=0,1,\ldots,q-1,\ q=24 \tag{22}$$

To date, using Iliyasu's QMF and the extensions to represent the chromatic (i.e. color-based) quantum movies (as reviewed here) a retinue of applications have been proposed for use on the quantum computing paradigm. Important aspects of all these applications are reviewed in Sections IV and V.

IV. SECURITY PROTOCOLS FOR QUANTUM MOVIES

With (FRQI and color-based) quantum movie frameworks (QMFs) taking ground as highlighted in preceding sections, recent literatures have been focused on uses for future quantum movie (or video) algorithms.

In this section, we present an overview of all the quantum movie security protocols that have been published to date (i.e. up to the time this study was completed in June 2018). Throughout, we shall focus on the key elements of each protocol as well as computational or logistical requirements for its execution. We shall, however, omit experimental demonstrations used to validate the performance of such protocols. Instead, where appropriate, we shall refer readers to relevant literature where such details are available.

Learning from the abuses digital content are subjected, one area that seems to be popular with the quantum information science community is multimedia (image, video, audio, text, etc.) security. This interest has extended to quantum movies where security protocols, including steganography [42], [47], encryption and decryption [41] have been proposed. Thus far, a majority of the available efforts in quantum movie security protocols are based on the MCQI-based CFQM movie representation ($\text{CFQM}_{\mathbf{M}}$) as reviewed in the remainder of this section.

A. ENCRYPTION AND DECRYPTION OF COLOR QUANTUM MOVIES

In the area of video processing, the need to handle vast amounts of data securely and quickly imposes constrains on video codec, storage space requirements and network communications [41], [73]. These factors were among the motivations for Yan *et al.*'s study on movie content encryption and decryption on quantum computers [41]. The protocol utilized the MCQI-based CFQM that was reviewed earlier to encode the movie content. Consequently, the notion of a movie consists of a strip composed of multiple frames (as defined in Equation (18) and Fig. 9) each an MCQI multichannel image as defined in Equations (16) and (17). Based on this framework, quantum measurement operations were employed to realize the movie encryption strategy whereas the versatile nature of quantum operations and processes was exploited to accomplish the decryption strategy. These core units of the encryption and decryption protocol are expatiated in the remainder of this subsection.

1) QUANTUM MEASUREMENT OPERATIONS

As observed in [41], basic quantum mechanics specifies that a measurement operation applied on a superposition state $[\alpha\,|0\rangle + \beta\,|1\rangle$ leads to the collapse of this state to produce a result 0 with probability α^2 or 1 with probability of β^2 where $|\alpha^2| + |\beta^2| = 1$ [41], [67]. The color information key (CIK) that is used widely in Yan's encryption and decryption strategy is built based on this simple measurement intuition (which is formalized in Definition 8).

Definition 8: A color information key (CIK) is a sequence of numbers assigned by an encoding rule for the purpose of transforming the chromatic information of an MCQI movie frame [24], [74].

To demonstrate the color information collapse and subsequent generation of the CIK key, we recapitulate the example in Fig. 10, which was originally discussed in [41]. Therein, a purple quantum pixel (top rightmost row in Fig. 10) whose R, G and B (RGB) composition is 128, 64 and 128 respectively was observed (i.e. measured). Upon measurement, this pixel collapses into a state corresponding to black or white pixels with probabilities on each channel determined by the basis of the quantum state measurement. This property was mentioned in the opening of this subsection and further explained in [41] where interested readers can get additional details. Since the CIK key is generated by quantum measurements, it is trivial that its outcome will be updated every time an image (or a frame in the context of a movie) is measured. As a consequence, the CIK key is considered as a 'dynamic' key that is available on the public domain. In total, for an MCQI-based movie content, there are eight possible configurations for a pixel in a post-measurement movie frame. Accordingly, based on the composition of the basic colors in the RGB model, eight indexes (i.e. i varying from 0 to 7) can be assigned to the different pixel colors from black through to white.

Yan *et al.*'s encryption and decryption strategies [41] are built based on the aforementioned configurations, wherein an encoding rule corresponding to the post measurement state of the pixel is generated. For example, the encoding rule for the purple pixel considered earlier is shown in the rightmost column of the post-measurement rows in Fig. 10, while probabilities of the measurements of the pixel in each

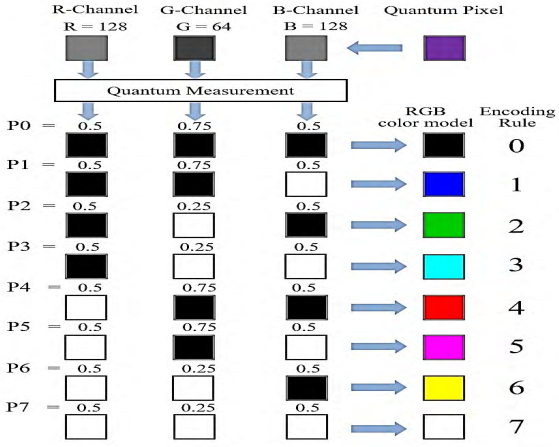

FIGURE 10. An example to show the color collapse in quantum measurement (figure adapted from [41]).

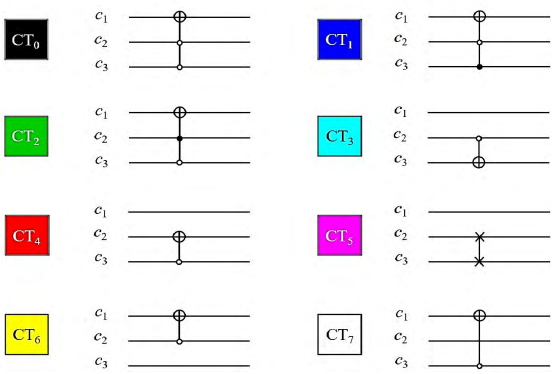

FIGURE 11. Encoding rule and quantum operations corresponding to different elements in CIK key (Figure adapted from [41]).

channel are labeled on the pixel block (i.e. further towards the left in the same figure). It is noteworthy that, since the CIK key is obtained via pixel-wise quantum measurements, then it would have the same length as the number of the pixels in the (post measured) image.

The adept generation of the encoding rule exemplified above serves as the focal point of the remaining encryption and decryption strategies that are reviewed in subsequent subsections.

2) ENCCRYPTION STRATEGY FOR MCQI-BASED COLOR QUANTUM MOVIES

As presented earlier in Section III, like in its forerunner the QMF, in the CFQM, a movie consists of three types of frames: the key, makeup and viewing frames. However, it is the measurement operations on each keyframe ($|F_1\rangle, |F_2\rangle, \cdots |F_m\rangle$) of the CFQM that generates an encoding rule which collectively produce the CIK key. Furthermore, it has been enunciated in this treatise that two or more keyframes (that exbibit very little resemblance to each other) could be used to delineate one shot (this general outline was conveyed earlier in Fig. 4 in Section II). Here, we focus on the 2^m keyframes in shot t, (i.e. $|F_1^t\rangle, |F_2^t\rangle, \cdots |F_m^t\rangle$), (and by assuming multiple copies of these keyframes have been prepared) and the quantum measurements (on each of them), which generates a corresponding color information key $CIK_1, CIK_2, \cdots CIK_m$. Next, utilizing a cogitation of some advanced color transformations that are widely used in QIP, specifically the channel of interest (CoI) and channel swapping (CS) operations [19], a color transformation operation (CTO) is assigned to each of the eight different divisions in the CIK keys such that this ruleset produces a codeword which then determines what operation is executed where. For example, the rulesets emanating from our earlier encoding rule for the purple pixel (in Fig. 10) are executed using the sub-circuits in Fig. 11 where, for brevity, the position and strip axes of the pixel have been omitted.

In the manner encapsulated above, and by further assuming that each frame is a $2^n \times 2^n$ MCQI image, then for a 2^m-ending movie strip (i.e. with m keyframes) m CIK keys will be generated. Descrying this important information in terms

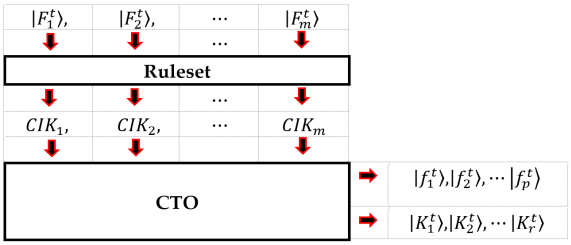

FIGURE 12. Encryption process for the example described in Fig. 10.

of linear algebra, we could interpret it as a row (vector) of an $m \times n^2$ matrix as presented in Equation (23).

The general circuit to implement this procedure on a 2^m-ending movie strip is presented in Fig. 12. For brevity, the half-moon notation [23] is adopted (in these descriptions and throughout our exposition) to denote the bi-level nature (i.e. state 'o' or 'I') of the control-condition that decides the constrains required to confine the CTO operation (CTO_i)) to a specific frame or shot within the movie.

Additionally, since in Yan's MCQI-based CFQM model [39] the number of qubits reserved to encode chromatic information is limited to three, the generated codeword is similarly confined to eight entries. Consequently, the ruleset mapping the choice of CTO operation to the CIK keys can be summarized as presented in Table 2.

$$CIK = R = \begin{pmatrix} CIK_1 \\ CIK_2 \\ \vdots \\ CIK_m \end{pmatrix} = \begin{pmatrix} R_{1,1} & R_{1,2} & \cdots & R_{1,n^2} \\ R_{2,1} & R_{2,2} & \cdots & R_{2,n^2} \\ \vdots & \vdots & \cdots & \vdots \\ R_{m,1} & R_{m,2} & \cdots & R_{m,n^2} \end{pmatrix} \tag{23}$$

TABLE 2. Codeword showing rulesets extracted from Fig. 11 (for the pixels in Fig. 10).

Transformation	Operation	Target channel(s)
CTO_0	Inversion	R
CTO_1	Inversion	G
CTO_2	Inversion	B
CTO_3	Swap	R with G
CTO_4	Swap	R with B
CTO_5	Swap	G with B
CTO_6	Inversion	R with G
CTO_7	Inversion	R with B

At this point, it is important to clarify that while the ruleset presented in Fig. 11 and Table 2 are extracted from the keyframes, the transformation each of them translates to (i.e. the choice of CTO) is determined by the CIK key. Furthermore, the resulting CTO operation is executed on the viewing and makeup frames in the movie shot.

Based on the above descriptions (and with reference to the structural outline of the QMF in Fig. 4) a CIK key will execute CTO operations on a movie segment comprising of content between keyframes $\left|F_1^t\right\rangle$, $\left|F_2^t\right\rangle$, $\cdots \left|F_m^t\right\rangle$ for a shot (t) consisting of p viewing frames $\left|f_1^t\right\rangle$, $\left|f_2^t\right\rangle$, $\cdots \left|f_p^t\right\rangle$ and r makeup frames $\left|K_1^t\right\rangle$, $\left|K_2^t\right\rangle$, $\cdots \left|K_r^t\right\rangle$. Assuming our earlier example (for the purple pixel in Fig. 10) has 3 keyframes, then the entire encryption procedure can be summarized in Fig. 12, while the sequence making up our encrypted movie shot ($|H_t\rangle$) would be in the form presented in Equation (24).

$$|H_t\rangle = |\boldsymbol{F_1^t}\rangle, |f_1^t\rangle, \ldots, |K_1^t\rangle, |\boldsymbol{F_2^t}\rangle, |f_4^t\rangle, \cdots |f_p^t\rangle |K_2^t\rangle, \cdots |K_r^t\rangle, |\boldsymbol{F_m^t}\rangle \quad (24)$$

where bold fonts are used to denote unencrypted frames.

Additionally, by exploiting well-known linear algebraic formulations, the CIK matrix (R) in Equation (23) could be modified to further strengthen the encryption strategy. For example, any two rows or columns could be interchanged or the entire matrix can be transposed, which alters the configuration of matrix R and with it the encrypted movie sequence could be modified. For example, transposing the CIK key used to generate the encrypted sequence in Equation (24) i.e. it is now R^T, would produce a different encrypted movie sequence than the one in Equation (24). It is trivial that, using established properties of matrices, these operations can be reserved, i.e. $(R^T)^T = R$, for the purpose of recovering the unencrypted contents of the movie.

3) DECRYPTION STRATEGY IN MCQI-BASED COLOR QUANTUM MOVIES

The objective of the decryption procedure is to undo the encryption process, so that the original untampered movie content can be recovered. In [41], the decryption process is executed in three steps as enumerated in the sequel.

First, the CIK matrix, R, is recovered by reversing the operation on it. This is similar to the double transpose operation that was mentioned in the latter parts of the preceding subsection. Second, the composition of the CIK keys are extracted from rows of R in Equation (23) (in the form $CIK_1, CIK_2, \cdots CIK_m$) for each keyframe. Third, referring to the codeword (i.e. the one used in the encryption procedure – Table 2), each encoding rule is reversed to produce the inverted CTO operations (CTO_i^{-1}). Applied on the encrypted movie sequence (i.e. in Equation (24)), $CTO_i^{-1}\,|H_t\rangle$ produces the stack of frames that constitute the original quantum movie. It should be clarified that, like its use in the encryption procedure, the reverse CTO operations are applied only on the makeup and viewing frames.

Finally, the structural layout of the QMF as used the encryption and decryption procedures as well as the role of the CIK keys in the video production, encryption and decryption units of Yan *et al.*'s MCQI-based color movie security protocol is presented in Fig. 13.

In concluding our review of Yan's encryption and decryption strategy in [41], we recount the protocol's ability to exploit two simple properties of quantum information processing: measurement and invertibility as tools to extract information needed to encrypt quantum videos as well as recover pristine, unencrypted content of the same videos, respectively.

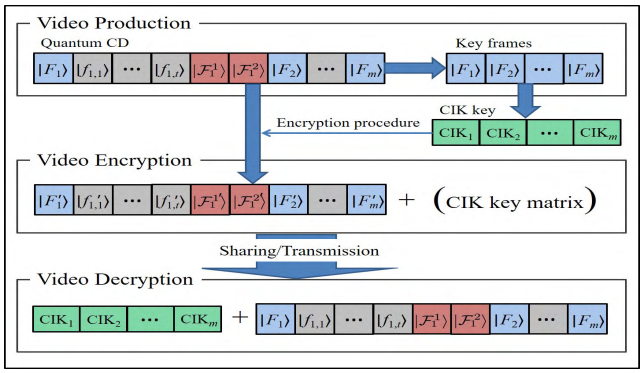

FIGURE 13. Outline of quantum video production, encryption, and decryption procedures used in [41].

Furthermore, the strategy was validated via carefully selected simulation-based experiments that are not reported here. Interested readers should consult [41] for details. Nevertheless, as reported therein, the protocol suffers from the inability to support movies with plane (i.e. black or white) backgrounds. This is attributed to the difficulty in localizing the readouts from measurement operations, which in turn impacts on the ability to generate the encoding rulesets, the codeword, CIK keys and CTO operations. Wang [40] claims that the outcomes from these types of movies produce meaningless encrypted videos. Consequently, there is need to further understand how measurement operations can be effectively utilized to enhance such outcomes.

B. HIGH PAYLOAD STEGANOGRAPHY PROTOCOL FOR COLOR QUANTUM MOVIES

Unlike encryption, steganography is often viewed as a process of achieving security by obscurity [34]. This often entails relying on the secrecy of the design or implementation as part of or the main method for providing security. The primary objective of steganographic (or simply stego) protocols is to mask certain information inside some media (audio, image, video, etc.) to make its presence indiscernible.

Generally speaking, quantum steganography is considered as still being in the early stage of its development, and, even more so, are its application in quantum image processing (QIP). Nevertheless, as a covert means of communication, quantum steganography can be realized by exploiting the physical characteristics of single- or multi-particle quantum states [47].

In terms of quantum multimedia such as images, video, audio, etc. in particular, and QIP in general, available quantum steganography protocols are mainly based on the NEQR image model such as [2], [26], [27], [76]. The Moire pattern was used to execute a recursive embedding procedure in [26], while [27] and [75] focused on using the least significant

bit (LSB) to embed hidden information into the frequency components of an image.

While [34] provides an in-depth overview of different protocols used in quantum image security, here our attention is focused on highlighting the second of only two security protocols available for quantum video (or movies). Consequently, this section is devoted to examining this protocol and its essential units. Incidentally, to date it is the only $\text{CFQM}_{\mathbf{N}}$-based quantum movie security protocol.

1) OUTLINE OF QUANTUM MOVIE STEGO PROTOCOL

We start by clarifying the three assumptions on which the smooth execution of Qu *et al.*'s quantum video stego protocol (in [42]) is built. First, the number of frames in the carrier video (N_C) must outnumber that in the secret video (N_S). Second, the size of the frames in the carrier video (S_C) must exceed those in the secret video (S_S). And third, the number of shots in the carrier video (H_C) must be more than those in the secret video (H_S).

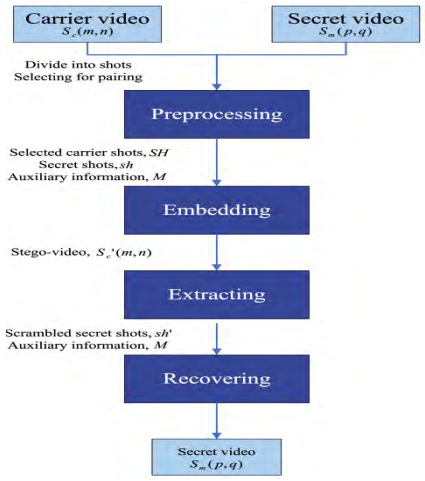

FIGURE 14. Flowchart showing different units of Qu *et al.*'s quantum video stego protocol (figure adapted from [42]).

The flowchart in Fig. 14 illustrates the interconnection between different units of Qu's stego protocol, which are further explained in the remainder of the section. It outlines the preprocessing unit where important information about the movie are retrieved. This is followed by the embedding unit; whose main purpose is to facilitate covert communication by implanting some secret information into the quantum video prior to its sharing in the public domain. Finally, the extraction and recovery units are used to recover the original quantum video.

The first step in Qu's quantum video stego protocol is the retrieval of information about the carrier and secret quantum movie shots, which is carried out in the preprocessing unit or sub-block. The main purpose of this unit is to segment the carrier video ($|S_c(m,n)\rangle$) into different shots according to some predetermined choice of keyframe(s). For this, it is assumed that the Sender and Receiver agree on some key (L), which specifies the keyframes to delineate for further processing in latter stages of the protocol. Furthermore, the absence of a one-to-one mapping between the carrier video shot, ($|S_c(m,n)\rangle$), and the secret video shot, ($|S_m(p,q)\rangle$) guarantees that the content of the selected keyframe(s) is known only by the Sender. Besides that, two additional sets of 'auxiliary' information (M) about the carrier and secret video shots $|S_c(m,n)\rangle$ and $|S_m(p,q)\rangle$ are also retrieved in the preprocessing unit. Specifically, the first set, known as the matching information (M), contains three entries: M_1, which culls some details from the carrier movie shot; M_2, which bookmarks information needed to map selected carrier shots to secret movie shots; and M_3, which cumulates the number of frames in each secret shot [42]. The second type of auxiliary information retrieved during the preprocessing process is the position information, $M_F\,(m,n)$, which relies on the three assumptions made earlier (in the opening remarks of this subsection) regarding the interrelation between the carrier and secret quantum video shots. Specifically, [42] assumed that the size of each frame in the carrier shot ($2^n \times 2^n$) is larger than those that make up the secret shot(s) ($2^q \times 2^q$), where $n > q$. $M_F\,(m,n)$ is a binary sequence that proves crucial in the execution of both the embedding and extraction procedures of Qu's stego protocol. In realizing the stego quantum video, this parameter is used to specify which color channel is utilized as an embedding location.

In the extraction unit, the embedding process is reversed to recover the video content prior to that procedure, albeit itself as an unordered sequence. Finally, a reordering process is used to restore the order of frames that make up the extracted carrier quantum video.

To ensure conformity with the angular formulation of the chromatic content of MCQI image states that encode the movie frames and strip, Qu *et al.* extended the description of the embedding procedure accordingly. In doing so, they supposed that $\theta_X^i \epsilon [0, 90]$ and $G_X^i \epsilon [0, 255]$ as the angular and grayscale values of the i^{th} pixel in carrier quantum video shot, ($|S_c(m,n)\rangle$) (and/or secret quantum video shot, ($|S_m(p,q)\rangle$), where X denotes the color channels, i.e. the R, G and B channels (and where needed, the α blending channel), then the post-embedding color of the i^{th} pixel in the s^{th} stego frame can be computed using Equation (25).

$$\theta_X^{i'} = a + \frac{\varphi_X^i}{90} \times 0.7 \tag{25}$$

where $\theta_X^{i'}$ is the color of the pre-embedding pixel (i.e. θ_X^i) for $\theta_X^i \in [a, b)$; and a and b are the right and left endpoints.

Next, the difference $(\theta_X^{i'} - \theta_X^i)$, called the shifting angle (R), is used as the angle for the channel of interest (CoI) operations that are used to execute the final modification of the chromatic content of the frames that make up the carrier quantum video shot.

To demonstrate execution of the protocol, Qu *et al.* considered, as an example, frames sequences in the pair of carrier and secret video shots in Fig. 15.

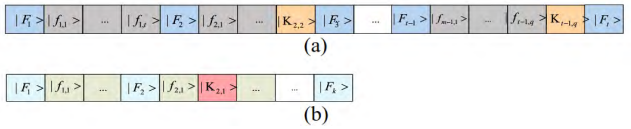

FIGURE 15. (a) Carrier and (b) the secret video shots to demonstrate execution of Qu *et al.*'s quantum movie stego protocol (figure adapted from [42]).

The two sets of auxiliary information position (i.e. M_1 to M_3 and $M_F(m, n)$) are used in the embedding process to modify keyframes of the shot in the carrier quantum video. Following the embedding process, the stego quantum video ($\left|S'_s(m,n)\right\rangle = |V(m,n)\rangle$) as defined in Equation (26) can be safely shared publicly.

$$(|V(m,n)\rangle)=|\boldsymbol{F}^{\boldsymbol{s}}_{\boldsymbol{1}}\rangle,|\boldsymbol{f}_{\boldsymbol{1,1}}\rangle,|\boldsymbol{K}_{\boldsymbol{1,1}}\rangle,|\boldsymbol{f}_{\boldsymbol{1,2}}\rangle,|\boldsymbol{f}_{\boldsymbol{1,3}}\rangle,|\boldsymbol{K}_{\boldsymbol{1,2}}\rangle,|\boldsymbol{K}_{\boldsymbol{1,3}}\rangle,|\boldsymbol{f}_{\boldsymbol{1,4}}\rangle,\cdots|\boldsymbol{f}_{\boldsymbol{1,7}}\rangle,|f_{1,8}\rangle,\cdots|K_{1,9}\rangle,|f_{1,10}\rangle \tag{26}$$

where the bold fonts indicate frames that have been embedded with content from the secret quantum video shot.

2) EMBEDDING UNIT

Like in every information security protocol, the purpose of the embedding unit of Qu's stego quantum video protocol is to infuse a hidden or secret message into publicly shared video such that its presence is concealed. In [42], this process is further divided into two steps. The first step involves embedding the auxiliary information ($M_F(m,n)$) that were extracted during the preprocessing onto the carrier video shot ($|S_c(m,n)\rangle$). This process entails spreading the shot matching information M_1, M_2, and M_3 into a predetermined keyframe in the carrier movie shot $|S_c(m,n)\rangle$. Meanwhile, in the second step of the embedding procedure, the position information $M_F(m,n)$ is used to determine which content of ($|S_m(p,q)\rangle$ to embed onto ($|S_c(m,n)\rangle$). In doing so, a binary sequence is recovered and, using it, a value '1' indicates that the channel is used as an embedding position and a value '0' indicates otherwise, i.e. that the channel is a non-embedding position.

To accomplish the embedding of $|S_{cm}(p,q)\rangle$ into $|S_c(m,n)\rangle$, a scale equating the color angle $\theta^i_X \epsilon [0, 90]$ to corresponding grayscale values $G^i_X \epsilon [0, 255]$ is used such that a value θ^i_X specifies the color angle of channel X (i.e. R, G. B or α) in the i^{th} pixel of the s^{th} stego frame, and, correspondingly, the grayscale value G^i_X in the same i^{th} pixel (in 8 bits).

3) EXTRACTION AND RECOVERY UNITS

Given a stego quantum video shot $\left|S'_s(m,n)\right\rangle = |V(m,n)\rangle$ (in Equation (26)), the extraction process uses two operations to extract the unmasked carrier movie shot, albeit this step itself would produce a disjointed sequence that needs to be reordered.

The first operation retrieves the two sets of auxiliary information: the matching information M_1, M_2, and M_3 and the position information, $M_F(m,n)$. These are extracted from the LSB components of the R, G, and B (and α, when used) channels of the movie frames constituting the stego quantum video ($|V(m,n)\rangle$) as well as the corresponding keyframes in $|S_c(m,n)\rangle$ and $|S_M(p,q)\rangle$.

The second operation is the actual extraction procedure, where the pair of extracted auxiliary information are used to separate $|S_c(m,n)\rangle$ from $|S_m(p,q)\rangle$. Specifically, using M_3, the receiver delineates (the number and location of) frames that make up each shot. Next, using the position information, $M_F(m,n)$, the receiver validates the embedded position within each stego frame. Finally, the frames of the secret shot are mapped to the extracted secret shots.

In terms of the angular descriptions used in the embedding process, the extraction procedure recovers the i^{th} pixel in the s^{th} frame of the stego video shot using Equation (27).

$$\varphi^i_X = \left(\frac{\theta^{i'}_X - a}{0.7}\right) \times 90 \tag{27}$$

where a and $\theta^{i'}_X$ have the same definitions as presented earlier in Equation (25).

The reordering unit is essentially used to reorganize the sequence of frames emanating from the extraction procedure. It utilizes the last unused matching information, i.e. M_2, to reestablish the order of frames that make up the original carrier movie shot.

Interested readers are referred to [42] for additional details regarding implementation of Qu's quantum video stego protocol.

We conclude by echoing the conclusions drawn by the authors of that study, wherein they averred that their protocol takes full advantage of the structural design of the QMF to randomly embed contents of a secret video onto some carrier or host quantum video.

Qu *et al.* (in [42]) also opine that, among others, a possible direction to enhance the performance of this protocol includes integration of aspects of coding theory into the proposed stego protocol.

V. APPLICATIONS FOR COLOR QUANTUM VIDEOS

Despite quantum movie production still being at its infancy, by exploiting many of the astounding properties of quantum mechanics and the adept structural design of the QMF framework for quantum movies, a handful of applications for have been proposed for it. In part, the three applications highlighted in this section were motivated by similar accomplishments in classical (i.e. digital) movie applications and supported by the 'quantumness' inherent to the formulation and structure of the QMF.

Digital movies are natural extensions of image processing, wherein a stack of still images are transitioned (i.e. flickered) at some rate to create the illusion of motion. Over the past decades, the TV and movie industries have evolved from being mainly focused on entertainment into a mega industry worth billions of US Dollars [43]. Similarly, in efforts to meet the ever-increasing expectations of viewers, the need for and use of technologies in applications to stimulate viewership have continued to grow. As noted in [43], recently,

subscription-based television (aka premium or Pay TV) has emerged to broaden the traditional TV screen and provide consumers with access to what they want, when they want it. These new models have behooved the need for multifaceted approaches to tease, stimulate, or convince consumers to watch or buy certain content [43]. Among others, these overtures include trailers, montages, etc. As defined via sources presented in [43], montages are ''film editing techniques that are tailored to convey the passage of time by editing its content into a sequence that condenses space, time, and information.'' Since they are intended to provide a statement regarding an aspect of the movie (a character, event, or plot), montages (together with trailers) provide a form of pre-feature entertainment [43].

As the first application tailored for color quantum movies, the feasibility and capability of present advances in QIP and quantum movie production to support digital-like montage applications was carried out in [43]. In the next subsection, we highlighted some of these outcomes therefrom.

In the field of computer vision, moving target detection (MTD) is a very active research topic [36], [39]. It plays an important role and serve as a crucial step in tasks such as intrusion detection, anomaly behavior detection and so on. With the widespread use of intelligent video surveillance systems and increasing complexity of monitoring conditions, a challenge in MTD is overcoming the random disturbances, which may seriously impede isolation of targets [77]. In digital image processing (DIP), the temporal differencing, background subtraction, and optical flow methods are considered the three most commonly used methods for solving this task [73], [74]. In terms of the computational complexity, detection accuracy, and system robustness, they all have some intrinsic merits and demerits, respectively.

Encouraged by the feats recorded in quantum information science, especially the celebrated speed up and tamper-proof security ascribed to quantum computing as well as recent inroads in the QIP sub-discipline, the possibility of accomplishing moving target and frames motion detection have been explored in [39] and [40], respectively. The details and intricacies related to the execution of these protocols are highlighted as the second application for color quantum movies.

On the whole, as mentioned earlier, this section will focus on providing an overview of the three efforts that have been proposed to use the quantum computing framework to execute movie (or video) related applications.

A. MOVIE MONTAGE OPERATIONS ON QUANTUM VIDEO

As elucidated in [43], the word montage originates from the French word ''monter'', which translates to ''to mount'' or ''cutting'', and in a broader sense, it is a movie editing technique where short clips (shots) are edited into a sequence. Therefore, by ''cutting and assembling the shots in the movie, montages condense space, time, and information.'' The resulting sequence from a montage is therefore tailored towards depicting the passage of time.

Conceptually, Yan's MCQI-based quantum movie montage application is built based on the MCQI CFQM (or as used in Section III, the CFQM_{M}). Structurally, the layout to execute the application is fabricated to support the design in Iliyasu's QMF, which was discussed in Section II. Operationally, the quantum movie montage applications we are reviewing in this section feature the chromatic – color-based operations on CFQM movies that were presented in Section IV as well as other sections of this treatise. These operations include the frame to frame (FTF), channel of interest (CoI) and sub-block swapping (SBS) transformations on contents of the CFQM quantum movies.

The conceit and the interaction (between various units) needed to actualize the montage application is highlighted in the next few subsections.

1) FRAME-TO-FRAME (FTF) OPERATIONS TO CONVEY PSYCHOLOGICAL MONTAGES IN QUANTUM VIDEO

As presented in Section III, we recall the use of a unique time tag, which is integrated into the formulation of each frame based on Yan's MCQI-based CFQM framework. This is used to determine the playing order (time lapse) of the movie clip, scene, etc. By altering the time tag of a frame (or group of frames), the playing order of a movie's content can be manipulated to realize important movie operations such as deletion, playback, and extraction. Based on the formulation of the $\text{CFQM}_{\mathbf{M}}$, framework that was presented in Equation (18), the frame to frame (FTF) operator defined in Equation (28) can be used to transit the t^{th} frame of a $\text{CFQM}_{\mathbf{M}}$ quantum movie ($|F_t(n)\rangle$) to the $(t \pm c)^{\text{th}}$ frame $|F_{t\pm c}(n)\rangle$.

$$\begin{aligned} FTF_c|M(m,n)\rangle &= \frac{1}{2^{\frac{m}{2}}}\sum_{S=0}^{2^{m-1}} |F_t(n)\rangle \otimes FTF_c|t\rangle \\ &= \frac{1}{2^{\frac{m}{2}}}\sum_{S=0}^{2^{m-1}} |F_{t\pm c}(n)\rangle \otimes |t \pm c \bmod 2^m\rangle \quad (28) \end{aligned}$$

where $|t\rangle$ represents the time information that tags each frame to the strip of the CFQM framework (as defined earlier in Equation (18)), $c\epsilon[0, 1, \ldots, 2^m - 1]$ indicates the number of shift steps that move from the t^{th} frame to the $(t \pm c)^{\text{th}}$ frame in the $\text{CFQM}_{\mathbf{M}}$ framework.

In an actual movie, the script comes to life as time elapses. Thus, the storyline is developed and, with it, the causality and depiction of the respective plots and relations between the main characters in the movie become more complicated. Notwithstanding its length, a shot is insufficient to elaborately convey the ideas of the director. Therefore, sometimes playing several shots alternately (e.g. with interludes after each shot) may be necessary to capture the narratives depicted via the transition between the shots. Used in this manner, the FTF operation can be a tool to facilitate abridgments in the movie content, connect shots, and replay certain frames. Therefore, using such an operation in a flexible and purposeful manner can help convey useful content in ways that may not be possible in the original movie.

Use of the FTF operation in depicting psychological montages, such as repeating one or more frames at different instances, so that the sublime inner world of the characters could be exhibited to present their reminiscences, or simply to depict a character's cogitation is one of the applications discussed in [43]. Moreover, the same study observed that use of the FTF operation could eliminate the need to duplicate the same frame when its use is required elsewhere within the movie.

Details of an animated example to demonstrate execution of the psychological montage application can be digested in [43].

2) CHANNEL OF INTEREST (COI) OPERATIONS TO CONVEY COMPARATIVE MONTAGES IN QUANTUM VIDEO

The channel of interest (CoI) operation [78], [79], which by its construction is focused mainly on manipulating the color of some specific channel (R, G, or B), has been successfully implemented in many MCQI-based QIP applications [24], [27] including the color quantum movie encryption and decryption protocol reviewed earlier in Section IV. Here, since the 2^m-ending movie strip is composed of MCQI images, the expediency of utilizing the CoI operation in some advanced MCQI-based quantum movie (i.e. CFQM$_{\mathrm{M}}$) applications, as considered in [43], is reviewed. These applications include those tailored towards manipulations to the chromatic content of a movie, such as the inversion of colors in an image (or parts of it) by manipulating its RGB channel [27]. Such a rendition is presented in the remainder of this subsection. Technically, the CoI_{X} operator shifts the grayscale value of a pre-selected color channel (R, G, or B channel) as defined below.

$$CoI_X = I^{\otimes 2n+3} \otimes \sum_{t\neq 0, t\neq k}^{2^n-1} |t\rangle\langle t| + CT_X \otimes |k\rangle\langle k| \tag{29}$$

where CT_{X} (i.e. the color transformation on channel X) is realized by using the $U_X = C^2R_y(2\theta)$ gate, which is the shifting parameter (or shifting angle) that is formulated in Equation (30).

$$CT_X = U_X \otimes I^{\otimes 2n}, \quad X \in [R, G, B]. \tag{30}$$

Applying that CT_{X} operator on the CFQM$_{\mathbf{M}}$ framework Equation (18) would yield:

$$\begin{aligned} CoI_X |S_m(m,n)\rangle &= CoI_X\Big(\frac{1}{2^{\frac{m}{2}}}\sum_{t=0}^{2^m-1} |F_t(n)\rangle \otimes |t\rangle\Big) \\ &= \frac{1}{2^{\frac{m}{2}}}\sum_{t=0, t\neq k}^{2^m-1} |F_t(n)\rangle \otimes |t\rangle \\ &\quad + \frac{1}{2^{\frac{m}{2}}}\sum_{t=0, t\neq k}^{2^m-1} \left|F'_k(n)\right\rangle \otimes |t\rangle . \end{aligned} \tag{31}$$

It should be noted that the frame $\left|F'_k(n)\right\rangle$ in Equation (31) (i.e. the frame at time $|k\rangle$), is applied on chromatic content from the original frame $|F_k(n)\rangle$ for which the shifting angle θ alters the R, G, or B channel.

As noted in [43], color is a "very important element that is extensively manipulated for different reasons, such as reflecting the theme and/or allure of a movie." Due to the established physiological adaptation of the human visual system to color, varying it at different tones and stages in a movie could convey some special artistic effects to the audience [43]. Furthermore, it was observed that during some frolicsome scenes in a movie, varying various hues produces strong contrast and, so, warm colors should be used in such instances. These interesting allegories formed the basis of the animated examples that depict comparative montage applications on CFQM$_{\mathbf{M}}$ color quantum movies which were discussed in [43]. Interested readers are directed to [43] where further details of the execution of this application can be found.

3) SUB-BLOCK SWAPPING (SBS) OPERATIONS TO CONVEY PARALLEL MONTAGES IN QUANTUM VIDEO

The sub-block blocking (SBS) operation is a class of geometric transformations that target the position information of sub-blocks in the frames of a CFQM$_{\mathbf{M}}$ quantum movie framework. Basically, this operation interchanges sub-blocks in frames from different instants of time within the CFQM$_{\mathbf{M}}$ framework, so that some predetermined manipulation of a scene is achieved [43]. Mathematically, the SBS$_{i,j}$ operation is defined as

$$SBS_{t_i,t_j} = I^{\otimes 3} \otimes |B\rangle\langle B| \otimes I^{\otimes m} + I^{\otimes 3} \otimes \left|B'\right\rangle\left\langle B'\right| \otimes SW_{t_i,t_j} \tag{32}$$

where SW$_{t_i,t_j}$ is a sub-operation that is used to exchange the time information, i.e. t_i and t_j between the two sub-blocks being swapped. It is accomplished in the form:

$$SW_{t_i,t_j} = |t_i\rangle\left\langle t_j\right| + \left|t_i\right\rangle\langle t_j| \sum_{t\neq t_i, t\neq t_j}^{2^m-1} |k\rangle\langle k| \tag{33}$$

Applying the operation SBS$_{t_i,t_j}$ operation on the MCQI-based CQFM framework would yield:

$$\begin{aligned} SBS_{t_i,t_j} |S_m(m,n)\rangle &= SBS_{t_i,t_j}\Big(\frac{1}{2^{\frac{m}{2}}}\sum_{i=0}^{2^m-1} |F_t(n)\rangle \otimes |t\rangle\Big) \\ &= \frac{1}{2^{\frac{m}{2}}}\sum_{i=0}^{2^m-1} \Big(\left|F_t^B(n)\right\rangle + \left|F'^B_t(n)\right\rangle\Big). \end{aligned} \tag{34}$$

In conveying a storyline, a good movie should be capable of captivating its audience. Where the acting or script falls short, graphics and montages can be effective in conveying certain minutia and emotions [43]. Oftentimes, however, to convey certain emotions or plots, advanced redesign of the montage operations is required.

In parallel or cross montages, presentation of two or more storylines in parallel or crosswise is required, and following which they are integrated into a whole plot in order to portray a unified theme. Consequently, by clipping two shots and

combining different parts into one unit, two storylines can evolve together [43]. In this manner, the movie will present to its audience a clear plot that cannot be realized in a single shot. Based on the SBS operation, the content of each frame or different frames could be manipulated to realize these advanced montage applications.

Further discussions on the use of the SBS operations to accomplish the parallel or cross montage application can be appreciated via [43].

While simplistic in design, the montage framework outlined above provides an avenue to use an assemblage of color transformations for rapid editing of movie contents and special effects that present compressed narrations of the movie, i.e. montages.

B. MEASUREMENT-BASED MOVING TARGET DETECTION USING COLOR QUANTUM MOVIES

In [39], Yan *et al.* utilized Yan's MCQI-based quantum movie model (i.e. CFQM in [40]) to propose a precept to detect moving targets in multi-channel colored quantum movies (or video). Similar to its use in the encryption and decryption strategy, the central component of the moving target detection (MTD) technique is the measurement operation. Its use facilitates the execution of the MTD algorithm by tracing the location of a moving object and then trailing its path throughout the 2^m frames that make up the CFQM movie strip.

This section highlights the main contributions of that study by Yan *et al.* (in [39]), which establishes the feasibility of undertaking the moving target detection in multi-channel quantum video.

1) MEASUREMENT OPERATIONS FOR MOVING TARGET DETECTION IN COLOR QUANTUM MOVIES

As reviewed earlier in Section III, like in all quantum systems, measurement operations to recover quantum states are quite tricky to navigate. This is attributed to its impact on the coherence and overall 'quantumness' that is a requirement for meaningful quantum computation. This property as well as the collapse (or loss) of the color information and their role in realizing the encoding rule for Yan *et al.*'s encryption and decryption strategy were examined back in Section III of this disquisition. Moreover, a detailed discussion of the strategy and general overview on QIP security protocols can be found in [6] and [34], respectively.

Our focus at this point is to provide a quick rehash to make the section replete in its content. Back in Section III, the measurement operation was used to isolate and track the motion of a predetermined object as it transits from one frame to the next in a CFQM movie register, i.e. a strip comprising of 2^m frames.

Referring to Fig. 10 (Section III), we saw how the readout for a (purple) pixel was recovered as one of eight possible configurations each with some probability value. Eight indexes 0 through 7 were assigned to each of the possible outcomes. Furthermore, we observed that the basic color (black or white) of the R, G and B channels were unaffected by the measurement operations. Potentially, this property could help in delineating a moving object from its background and, in fact, it is part of the impetus for the study on moving target detection (MTD). This property and its use to detect moving targets are further enumerated in the sequel.

2) GENERATION OF MTD MATRIX AND PROCEDURE FOR DETECTING MOVING OBJECTS

As laid out in [39], suppose our CFQM consists of k movie frames (each an $n \times n$ MCQI quantum image), then the MTD process can be accomplished via the following steps.

- Using the circuit set up in Fig. 16, apply the multiply controlled measurement operation(s) to retrieve a readout of the chromatic information as shown. Knowing the impact of such operations on the content of the movie, it is assumed that sufficient copies of the movie have been prepared and initialized as outlined in many FRQI-based QIP literature, such as [6], [7], [17], [18], and [36]. In this manner, multiple operations such as those in Fig. 16 can be applied to recover the chromatic details of all the frames in a movie strip.

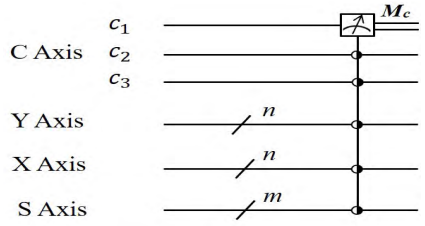

FIGURE 16. Measurements on MCQI-based multi-channel quantum video.

- Since it has been established that observing (i.e. measuring) the information entrapped in a movie leads to a classical readout with (in our case of measurements of chromatic information) a different color for each pixel, then, using the encoding rule described in Section III, we can generate a sequence of n^2 entries emanating from each $n \times n$ movie frame. For a $k-$ ending movie strip, we would realize $k \times n^2$ entries [39]. This intuition was used to generate a $k \times n^2$ matrix known as the MTD matrix, D.
- By traversing the entries of the MTD matrix (D), variations in the composition of elements that constitute the matrix can be determined. These variations delineate an object from its background and, so they could indicate the location of a moving object as it moves from one frame to the next. Put more formally, assuming each element in D is presented as (k, l), then the location (k, i, j) could be used to specify the coordinates of the target in a frame (where $\frac{l}{n} = i\frac{j}{n}$). Consequently, by analyzing the locations of the target in each frame, we can depict a trail showing the path traversed by that object as it journeys from one frame to another throughout the length of the CFQM quantum movie register.

The likelihood of executing this application is attributed to the adept design and structure of the QMF as well as the flexibility of the MCQI image format used to encode each of the frames. These two properties support the use of appropriate constraints to isolate and track the moving target. As demonstrated via the SMO and displacement transformations in Section II, where required, different operations targeting the chromatic and/or geometric content of the movie frames can be used to facilitate more advanced transformations. Interestingly, it was shown in [39] that execution of the MTD procedure (as outlined in the three steps above) requires $O(q.2^{m+4n})$ operations to accomplish, where we assume use of q copies of the quantum movie each consisting of 2^m frames. Additionally, by focusing on the keyframes only, the earlier cost can be reduced to $O(p.\ 2^{4n})$, where it is assumed that the movie has only p keyframes.

To validate the feasibility and performance of the MTD protocol, as is the trend in QIP, Yan *et al.* [39] assembled three interesting simulation-based experiments. Readers are referred to [39] for details of those experiments.

C. FRAMES MOTION DETECTION IN NEQR-BASED COLOR QUANTUM MOVIES

Motivated by the its role in digital image processing (DIP) and computer vision applications, such as in intelligent video surveillance, and closely related to the moving target detection (MTD) protocol reviewed in the preceding section, Wang proposed the use of the NEQR image format [20] (that was reviewed in Section III) to undertake an approach to detect motion in frames of quantum movies.

To recall, in the NEQR-based CFQM quantum movie (also referred to as CFQM_N), the description of movie strip takes the form presented in Equation (19). Based on it, Wang's frame motion detection (FMD) technique [40] is executed in the three steps that are outlined in the remainder of this section.

1) FRAME PARTITIONING OPERATION

This operation is a sort of preprocessing that partitions the movie frames into block-sized sub-blocks.

Given that each CFQM_N quantum movie frame is of $2^n \times 2^n$ dimension, the partition operation generates 2^{2n-6} sub-blocks and it is trivial to see that for such a task $2n-6$ qubits will be required. To simplify matters, a binary sequence (or codeword) corresponding to the location (i.e. an index 0 through 7) of the sub-block is used to address each sub-block. Following this, the contents of each sub-block are exchanged and padded to realize an 8×8 sub-block.

2) COMPARISON OF FRAME BLOCKS

This operation uses the widely used comparator (COMP) operator to determine congruity or similarity between the partitioned sub-blocks in each movie frame. Quantitatively, this is accomplished by computing a qubit-wise pixel difference between each pair of frames, say $|I_1\rangle$ and $|I_2\rangle$, using

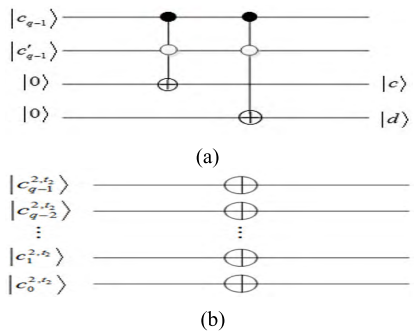

FIGURE 17. Circuits to realize the comparison of sub-blocks (a) quantum comparator, and (b) quantum complement operations (figure and descriptions adapted from [40]).

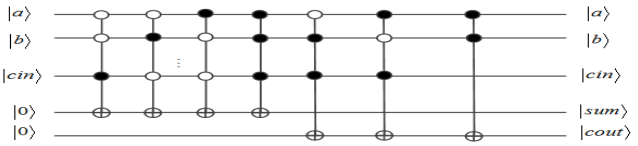

FIGURE 18. Quantum circuit of 1-ADD operation (figure and descriptions adapted from [40]).

Equation (35) below.

$$\left|c'_{q-1} - c_{q-1}\right| \times 2^{q-1} + \left|c'_{q-2} - c_{q-2}\right| \times 2^{q-2} + \cdots \left|c'_0 - c_0\right| \times 2^0 = 2^q - 1 \quad (35)$$

The complexity of the above operation is found to be $2.2^{2n}O(N)$, where $O(N)$ is the complexity of the comparator operation.

3) RESIDUAL INFORMATION CALCULATION

Suppose $\left|B^1_{t_1}\right\rangle$ and $\left|B^1_{t_2}\right\rangle$ are blocks from two video frames, then the residual information (ΔB) for the two frames is computed using Equation (36).

$$|\Delta B\rangle = \frac{1}{2^n} \sum_{y_i \epsilon \{0,1\}, x \epsilon \{0,1\}} |C^{1,t_1} - C^{2,t_2}\rangle |t_{b_y}\rangle |y_2 y_1 y_0\rangle |t_{b_x}\rangle |x_2 x_1 x\rangle \quad (36)$$

where $\left|C^{1,t_1}\right\rangle = \left|c^{2,t_1}_{q-1}, c^{2,t_1}_{q-2} \cdots c^{2,t_1}_0\right\rangle$ and $\left|C^{2,t_1}\right\rangle = \left|c^{2,t_2}_{q-1}, c^{2,t_2}_{q-2} \cdots c^{2,t_2}_0\right\rangle$. The quantum circuit to execute the residual information calculation consists of two layers or sub-circuits. The first comprises of a network of n NOT (inverter) gates (Fig. 17(b)) that are applied on the $\left|c^{i,j}_q\right\rangle$ color qubit of each sub-block in the movie frame producing an output $\left|\overline{c^{2,t_2}}\right\rangle$. This sub-circuit is also referred to as the complement operation and is widely used in QIP [6], [39]. The second layer or sub-circuit of the residual information calculation operation is the quantum adder (q-ADD or ADDER) circuit, which computes the sum of $\left|c^{2,t_2}\right\rangle$ and $\left|\overline{c^{2,t_2}}\right\rangle$.

The 1-ADD circuit in Fig. 18 demonstrates the execution of the larger, more complex q-ADD circuit network (in Fig. 19) for $q = 1$.

Although exiguously implemented and insufficiently validated, the author of [40] claims that the strategy outlined

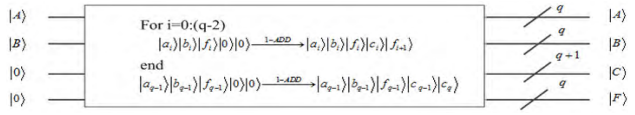

FIGURE 19. Quantum circuit of *q*-ADD operation (figure and descriptions adapted from [40]).

above effectively executes the frame motion detection operation. Readers can refer to [40] to examine the contributions made therein.

D. CONCLUDING REMARKS ON APPLICATIONS FOR COLOR QUANTUM MOVIES

To enhance the canonical elegance of the FRQI-based grayscale movie scheme [38], in [43], chromatic considerations were introduced into descriptions of movies on quantum computers [38]. The expansive collection of color-related transformations available on this new framework as well as the acute use of measurement operations on this new chromatic framework for quantum movies (CFQM) were the building blocks of the applications that were reviewed in this section.

The moving target detection method (in [39]) that was reviewed in this section provides a measurement-based approach that is capable of detecting the location of the moving target in each frame of the quantum video. This facilitated easy and efficient retrieval of the motion trail of an object. It provides a modest attempt to focus on the moving target detection and its applications in quantum video. Meanwhile, on its part, [40] sought to eliminate the need for the measurement operations. Instead, using three stages: frame blocking, comparison of frame blocks and residual calculation, a frame motion detection (FMD) strategy was developed to safeguard the quantum state. Both efforts provide unfledged, but meaningful attempts to introduce information processing techniques to quantum video scenarios.

Integration of color in the new CFQM movie model supported the manipulation and recovery of the different fragments that make up a movie. In [43], an assemblage of these transformations was utilized towards rapid editing of movie contents and special effects to present compressed narrations of the movie, i.e. montages. Specifically, we highlighted the use of a retinue of carefully formulated transformations including the frame-to-frame (FTF), color of interest (COI), and sub-block swapping (SBS) operations were used to demonstrate implementation of psychological, comparative, and parallel (or cross) montage applications that are essential in conveying various dialogues in a movie.

Yan's MCQI-based chromatic frame for quantum movies (CFQM_M) presents a veritable tool to encode both the chromatic and temporal information needed to effectively depict a movie as well as its manipulation for different protocols (as reviewed in Sections II through V). Moreover, this movie format offers improved gear to manipulate the content of a movie for the realization of advanced quantum movie applications in future.

VI. QUANTUM AUDIO SIGNAL PROCESSING (QASP)

In a simplistic description, sound could be considered to consist of digital inscriptions that recreate sound waves, which depict spoken words, conversations, singing, etc. Therefore, an audio signal is a representation of sound, often, within some specified frequency levels.

The first recorded reference to a quantum audio signal in its emerging format as discussed in subsequent parts of this section is traced to Iliyasu *et al.*'s quantum movie framework [38], whence its need and integration into their silent movie was mooted. Since then, two models have been proposed for quantum audio signal processing (QASP). The two models utilize the quantum mechanical structure of quantum computing to encode quantum audio signals as well as mechanisms to undertake essential transformations on those signals. The rudiments of both quantum audio formats are presented and discussed in the remainder of this section.

A. QUANTUM REPRESENTATION FOR DIGITAL AUDIO (QRDA)

Interestingly, the quantum representation for digital audio (QRDA) was inspired by a QIP image presentation, specifically, Zhang *et al.*'s NEQR [20] that was reviewed in Section III. Wang's QRDA [44] went on to serve as the first representation for quantum audio signals.

Structurally, the QRDA captures information about the audio amplitude and temporal information about audio signals as two entangled quantum sequences. In addition to its formulation, Wang provided details of the QRDA preparation and retrieval [44].

An analog audio sample (Fig. 20) can be represented as discretely sampled amplitude and time information $A = [a_0, a_1, \cdots, a_{L-1}]$, where L is the length of the audio $L \in \aleph$ and $a_i \in \{0, 1, \ldots, 2^q - 1\}$. As noted in [43], by emulating this description, the QRDA could be described as a q-bit binary sequence $D_T = D_T^0 D_T^1 \ldots D_T^{q-2} D_T^{q-1}$ to encode the amplitude value (a_T) as presented below.

$$a_T = D_T^0 D_T^1 \ldots D_T^{q-2} D_T^{q-1}, \quad D_T^i \in \{0, 1\}, \; i = 0, 1, \ldots, q-1 \tag{37}$$

where $T = 0, 1, \ldots, L-1$ is the time information.

Similarly, by simplifying this description, a QRDA quantum audio sample could use two qubit sequences to store amplitude and time information, which supports storing the whole digital audio sample as a superposition of the two-qubit sequence [44].

Generally, a QRDA quantum audio sample can be written as [44]:

$$|A\rangle = \frac{1}{\sqrt{2}^l} \sum_{T=0}^{L-1} |D_T\rangle \otimes |T\rangle \tag{38}$$

$$|T\rangle = |t_0 t_1 \cdots t_{t-1}\rangle, \; t_i \in \{0, 1\}$$

$$|D_T\rangle = \left| D_T^0 D_T^1 \ldots D_T^{q-2} D_T^{q-1} \right\rangle, \quad D_T^i \in \{0, 1\} \tag{39}$$

where

$$l = \begin{cases} log_2 l & L > 1 \\ 1 & L = 1 \end{cases} \tag{40}$$

and $|D_T\rangle$ is the binary sequence of the sample amplitude value, $|T\rangle$ is the corresponding time information.

These descriptions indicate that the QRDA quantum audio model requires $(q + l)$ qubits to represent an audio sample with 2^q amplitude range.

Fig. 20 shows a 4-qubit QRDA quantum audio sample and its representative expression in QRDA format. Since the size of the audio sample is $L = 15$, then $l = \lceil log_2 15 \rceil = 4$, which implies that the sample has a binary length of time of 4 bits.

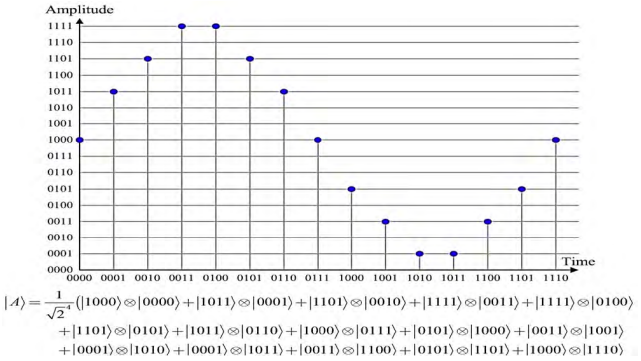

FIGURE 20. Sample audio signal and its representative expression in QRDA (figure adapted from [44]).

1) QRDA AUDIO SIGNAL PREPARATION

The preparation of a QRDA quantum audio sample or signal is executed in two steps as presented in Fig. 21 and the description that follows.

Step 1: The Identity (I) and Hadamard (H) transforms are used to convert the initial state $|0\rangle$ to the intermediate state $|1\rangle$, which is the superposition of all the samples of an empty digital audio state, expressed as U_1 (presented in Equation (41)).

$$U_1 = I^{\otimes q} \otimes H^{\otimes l} \tag{41}$$

Applying the transform U_1 on $|0\rangle^{\otimes q+l}$ produces the state $|\Psi_1\rangle$, which completes the first step of the QRDA quantum audio preparation.

Step 2: In the second step of the preparation procedure, the amplitude values of all samples are stored via L sub-operations. For a sample T, the quantum sub-operation U_T executes the operation as presented in Equation (42).

$$U_T = \left(I \otimes \sum_{i=0, i\neq T}^{2^l-1} |i\rangle\langle i| \right) + \Omega_T \otimes |T\rangle\langle T| \tag{42}$$

where Ω_T denotes the setting operation for sample T defined in Equation (43), which comprises of q quantum oracles as defined in Equation (44).

$$\Omega_T = \bigotimes_{i=0}^{q-1} \Omega_T^i \tag{43}$$

$$|\Psi\rangle_0 = |0\rangle^{\otimes q+l}$$
$$\downarrow \text{Step 1}$$
$$|\Psi\rangle_1 = \frac{1}{\sqrt{2}^l} \sum_{T=0}^{2^l-1} |0\rangle^{\otimes q} |T\rangle$$
$$\downarrow \text{Step 2}$$
$$|\Psi\rangle_2 = \frac{1}{\sqrt{2}^l} \left(\sum_{T=0}^{L-1} \bigotimes_{i=0}^{q-1} \left|D_T^i\right\rangle |T\rangle + \sum_{T=L}^{2^l-1} \bigotimes_{i=0}^{q-1} |0\rangle |T\rangle \right) = |B\rangle$$

FIGURE 21. Procedure for preparation of QRDA quantum audio signal [44].

where

$$\Omega_T^i : |0\rangle \longrightarrow \left|0 \oplus \mathrm{D}_T^i\right\rangle \tag{44}$$

and $\oplus$ is the XOR operation. Specifically, this equation dictates that: if $\mathrm{D}_T^i = 1$, Ω_T^i: $|0\rangle \longrightarrow |1\rangle$ is an l-CNOT operation, i.e. a CNOT gate with l control conditions; otherwise, $\mathrm{D}_T^i = 0$, Ω_T^i: $|0\rangle \longrightarrow |0\rangle$, which is an identity operation.

The final outcome of this step is obtained by applying U_T on the intermediate state $|\Psi_1\rangle$ which produces $|\Psi_2\rangle$ as shown in Fig. 21. Meanwhile, Fig. 22 presents the circuit network for the QRDA quantum audio preparation.

2) QRDA AUDIO SIGNAL PROCESSING

Two or more QRDA quantum audio samples can be manipulated using basic tools of linear algebra, specifically, unitary quantum transformations. This section highlights three basic audio processing operations that can be executed on QRDA quantum audio signals that have been prepared in format described in preceding sections.

- Combination of QRDA audio clips

Oftentimes, the need arises to combine or connect two or more audio clips. On quantum computers, using the QRDA quantum audio format, this task can be accomplished as follows. Assume two QRDA quantum audio clips $|A_X\rangle$ and $A_Y\rangle$ (defined in Equations (45) and (46), respectively) with an amplitude range of 2^q, then the process of combining them as depicted in Fig. 23, can be realized via the operation in Equation (47). This produces an output clip $|A_Z\rangle$ of length 2^{l+1}, for which $l + 1$ qubits are required.

$$\begin{aligned} |A_X\rangle &= \frac{1}{\sqrt{2}^l} \sum_{\mathrm{X}=0}^{2^l-1} |D_X\rangle \otimes |X\rangle \\ |X\rangle &= |x_0 x_1 \cdots x_{l-1}\rangle, \quad x_i \in \{0, 1\} \\ |D_X\rangle &= |D_X^0 D_X^1 \cdots D_X^{q-2} D_X^{q-1}\rangle, \quad D_X^i \in \{0, 1\} \end{aligned} \tag{45}$$

$$\begin{aligned} |A_Y\rangle &= \frac{1}{\sqrt{2}^l} \sum_{\mathrm{Y}=0}^{2^l-1} |D_Y\rangle \otimes |Y\rangle \\ |Y\rangle &= |y_0 y_1 \cdots y_{l-1}\rangle, \quad y_i \in \{0, 1\} \\ |D_X\rangle &= |D_Y^0 D_Y^1 \cdots D_Y^{q-2} D_Y^{q-1}\rangle, \quad D_Y^i \in \{0, 1\} \end{aligned} \tag{46}$$

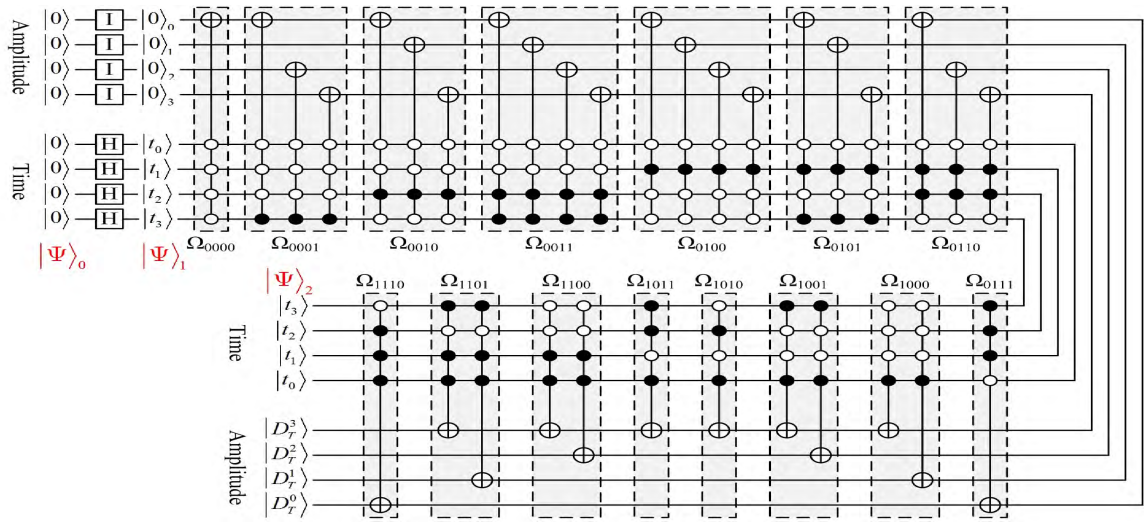

FIGURE 22. Circuit network for preparation of QRDA quantum audio version of the digital audio in Fig. 20 (figure and descriptions adapted from [44]).

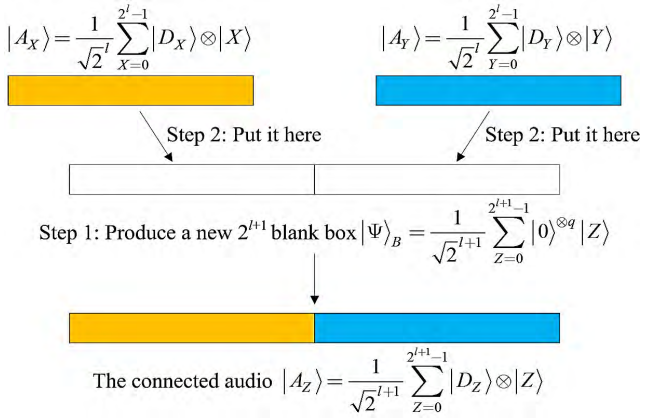

FIGURE 23. Outline of process to combine two QRDA quantum audio clips (figure and descriptions adapted from [44]).

$$
\begin{aligned}
|A_Z\rangle &= \frac{1}{\sqrt{2^l}} \sum_{Z=0}^{2^l-1} |D_Z\rangle \otimes |Z\rangle \\
|Z\rangle &= |z_0 z_1 \cdots z_{l-1}\rangle, \quad z_i \in \{0, 1\} \\
|D_Z\rangle &= |D_Z^0 D_Z^1 \cdots D_Z^{q-2} D_Z^{q-1}\rangle, \quad D_Z^i \in \{0, 1\}
\end{aligned} \tag{47}
$$

Further details pertaining to the execution of the operation in Fig. 23 can be digested via [44]. The outcome of this step shows that $(2l + 1)$ CNOT gates are used to propagate $|A_X\rangle$ and $|A_Y\rangle$ into 2^{l+1} blank boxes $|\Psi\rangle_B$ (Equation (46)), which is similar to combining $|A_X\rangle$ and $|A_Y\rangle$ as required.

$$
|\Psi\rangle_B = \frac{1}{\sqrt{2^l}} \sum_{Z=0}^{2^{l+1}-1} |0\rangle^{\otimes q} |Z\rangle \tag{48}
$$

Assuming the QRDA quantum audio segments are described via an amplitude range and length: $l = g = 2$ then the combined quantum audio is can be realized via the circuit in Fig. 24.

- Mixing of QRDA audio signals

Mixing of audio content is another important digital audio processing task whereby two or more audio clips are mixed into one channel [80]. Audio mixing is a very useful operation in music production as well as speech communication.

Using the quantum QRDA audio format, two clips $|A_X\rangle$ and $|A_Y\rangle$ (as defined earlier in Equations (45) and (46)) can be mixed to produce a mixed audio clip, $|A_M\rangle$ in the form presented in Equation (49).

The q-ADD (or ADDER) operation, whereby $|a\rangle\,|b\rangle \longrightarrow |a\rangle\,|a+b\rangle$ (whose circuit was presented in Fig. 18 and 19 (Section V)) is an integral part of the QRDA quantum audio mixing operation. The addition network takes the form presented in Equation (49).

$$
U_A = \prod_{X=0}^{2^l-1} U_X \tag{49}
$$

where U_X is the q-ADD (or ADDER) operation with 2^l control conditions.

Finally, the QRDA quantum audio mixing operation is executed using the operation:

$$
U_M = U_A \otimes U_{/2} \tag{50}
$$

where $U_{/2}$ was defined in [43] as:

$$
U_{/2} = \bigotimes_{i=q-1}^{0} \mathrm{U}^i \tag{51}
$$

and U^i is a swap operation (or gate).

Putting it altogether, the circuit network that is used to accomplish the mixing of two QRDA quantum audio segments (and an example for the case $l = g = 2$) are

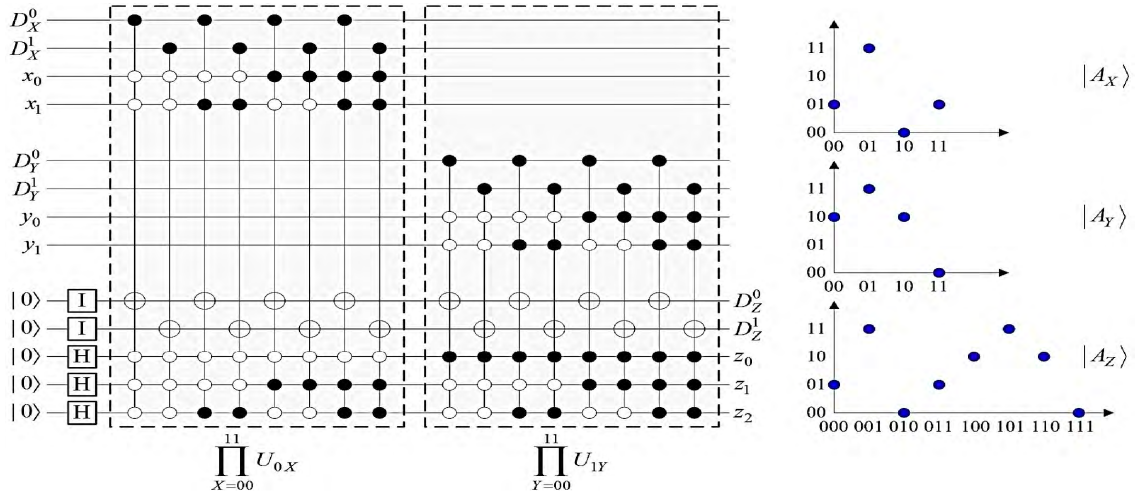

FIGURE 24. Circuit network for combining two QRDA audio quantum signals (b) example to show combination of two audio clips $|A_X\rangle$ and $|A_Y\rangle$ (see text for details) (figure adapted from [44].

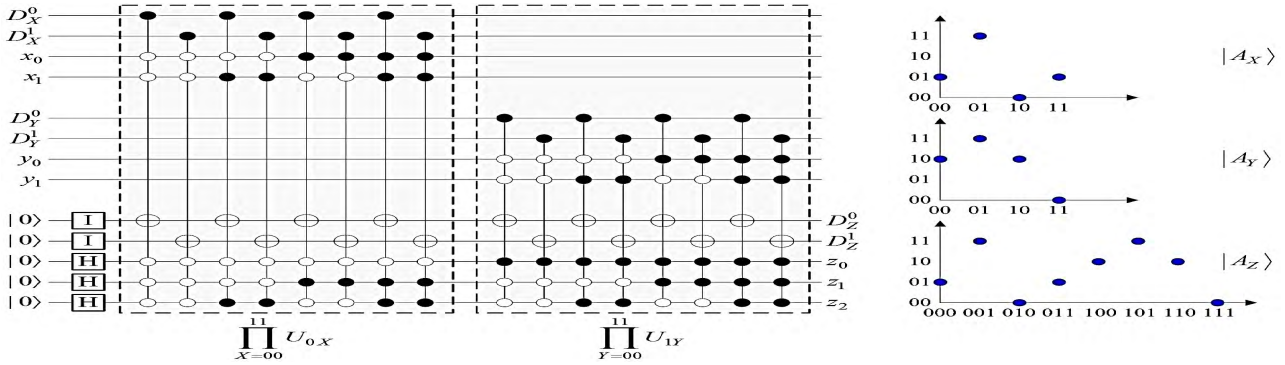

FIGURE 25. Circuit network for mixing two QRDA audio quantum signals (b) example to demonstrate mixing two audio clips $|A_X\rangle$ and $|A_Y\rangle$ (see text for details) (figure adapted from [44]).

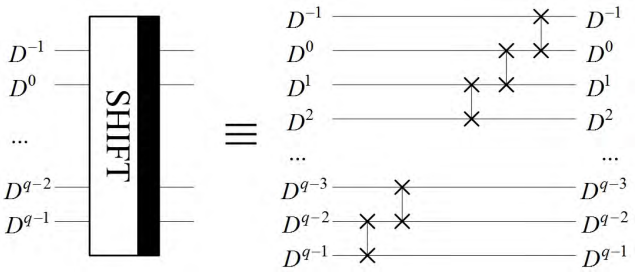

FIGURE 26. Circuit to execute for SHIFT operation in Fig. 25 [44].

presented in Fig. 25. Meanwhile, the detailed circuit for the shift sub-operation in that circuit is presented in Fig. 26.

- Compression of QRDA audio signals

Inspired by Le *et al.*'s use of a quantum image compression (QIC) process [16], [17] to decrease the number of simple quantum gates in the preparation and reconstruction of quantum images, Wang [44] proposed a quantum audio compression (QAC) process that is similarly aimed at simplifying the cost and complexity of the QRDA quantum audio signal preparation procedure. The steps to accomplish this task are highlighted in the sequel.

Definition 9: The process, which decreases the number of simple quantum gates in the preparation and reconstruction of a quantum audio state, is called quantum audio compression (QAC).

Unlike the QIC process used to compress quantum images (in [16] and [17]), which is built entirely on the minimization of Boolean expression (MBE), to increase the audio compression ratio, in [44], Wang explored integrating the differential pulse code modulation (DPCM) [81] into the MBE.

In principle, the DPCM codes the difference between the first sample and the next one. DPCM can compress a signal with memory significantly. For example, 8 bits would be required to code signals in decimal values 248, 249, 252, 255. Using the DPCM, however, their difference signal is 1, 3, 3 and 2-bit long. Audio is a kind of information source with memory. If the previous sample is a, the current sample will be close to a. For example, say, the i^{th} audio sample $S_i = a$, its immediate neighbors S_{i-1} and S_{i+1} would have values close

to a. Therefore, the difference $|S_i - S_{i-1}|$ and $|S_{i+1} - S_i|$ will be smaller than $|a|$.

In order to apply DPCM to QAC, a flag qubit $|f_T\rangle$ is added to QRDA to indicate $|D_T\rangle$ is a difference or an absolute value:

- If $|f_T\rangle = |0\rangle : |D_T\rangle$ is the difference between sample T and sample $T-1$;
- if $|f_T\rangle = |1\rangle : |D_T\rangle$ is absolute value of sample T.

Hence, an QRDA audio can be rewritten in equation (52).

$$\begin{aligned} |A\rangle &= \frac{1}{\sqrt{2}^l} \sum_{\mathrm{T}=0}^{L-1} |D_T\rangle \otimes |T\rangle \otimes |f_T\rangle, \quad f_T \in \{0, 1\} \\ |T\rangle &= |t_0 t_1 \cdots t_{l-1}\rangle, \quad t_i \in \{0, 1\} \\ |D_T\rangle &= |D_T^0 D_T^1 \cdots D_T^{q-2} D_T^{q-1}\rangle, \quad D_T^i \in \{0, 1\} \end{aligned} \tag{52}$$

It can be seen from the above example that the first sample stores the absolute value, whereas other samples store the difference. This itself infers some sort of compression in the audio signal. The QRDA quantum audio compression operation requires $(q + l + 1)$ qubits to represent L audio samples with 2^q amplitude range.

Further details including a few examples on the QRDA quantum audio compression can be obtained in [44].

3) QRDA AUDIO SIGNAL RETRIEVAL

As with other quantum computing systems, the final step of any algorithm is the operation to retrieve or measure the output. In QRDA, simple projective measurements (reviewed earlier in Section II, specifically, Fig. 8) to recover the QRDA quantum audio state. As conceived in [44], every audio sample is retrieved individually and so its amplitude values are recovered.

In concluding, as proclaimed in [44], "quantum audio processing (QAP) is a new branch of QIP," which could potentially offer useful technology to process digital audio information using quantum computing hardware. Wang's effort produced the first quantum representation for quantum audio by fusing audio information into quantum mechanical descriptions of quantum computation. This required the allocation of two qubit sequences to store information about audio sample's location as well as audio related information. Furthermore, operations and circuitry to accomplish the combination, mixing, and compression of QRDA quantum audio samples were presented. This effort in demonstrating the possibility of undertaking quantum audio signal processing (QASP) opened the door for others that followed, notably, the flexible representation for quantum audio, a relatively more recent quantum audio model that claims to improve many shortcomings of the QRDA quantum audio format.

B. FLEXIBLE REPRESENTATION FOR QUANTUM AUDIO (FRQA)

Notwithstanding its innovation, for reasons adduced by Yan *et al.* [45], [46] and reproduced below, the tightly-bounded unipolar encoding strategy used in QRDA may hinder accurate computation and processing operations in quantum audio processing:

- As presented in the preceding section as well as its use in [44], based on the QRDA model for quantum audio processing (QAP), amplitude values at can only represent nonnegative numbers. Hence, some arithmetic operations pertaining to amplitude values are prone to errors. For instance, when there are two amplitude values a_m and a_n (such that $a_m < a_n$), it is virtually impossible to implement and produce a result for the operation $a_m - a_n$.
- As a unipolar representation, the QRDA representation is not formulated to display or determine midrange values of a waveform during processing operations. For instance, as a wave, an audio signal is made up of alternating peak values – called crests and troughs - which could grow, shrink or even cancel out each other when different operations are performed on two different waveforms. Using the formularization in the QRDA quantum audio format, however, it is difficult to execute some operations, more so, since amplitude values are positive (e.g. addition of opposite amplitude values in two waveforms will accumulate to higher values rather than, as expected, the two counteracting or offsetting each other).

Among others, augmenting for these identified lapses in the QRDA were the objectives of the new quantum audio format proposed by Yan *et al.* [45], [46].

Much like the QRDA, Yan *et al.* found motivation for their flexible representation for quantum audio (FRQA) from within the quantum image processing (QIP) sub-discipline. Besides the striking resemblance in nomenclature, the FRQA shares structural and mathematical similarity with the FRQI quantum image model. Consequently, the preparation procedure for the FRQA ectype that of the FRQI.

In many ways, the synergy between the FRQI and FRQA seems ideal, more so, now that advances in both QIP and QAP seem mature enough to support synchronization between quantum movies and audio to produce talking quantum movies. As mentioned in the early stages of this treatise, among others, assessing the maturity of the technologies across both media (images/videos and audio) for this concordance is an objective of our disquisition. This will be further discussed in Section VIII.

In the meantime, for the remainder of this section, attention will be devoted to describing the requirements for preparing, processing, and retrieving quantum audio information based on the FRQA model for quantum audio signal processing (QASP).

First, we start by highlighting some of the utility of the FRQA quantum audio format over the QRDA model.

In Yan *et al.*'s flexible representation of quantum audio, i.e. FRQA, amplitude values are encoded in a bipolar (i.e. both non-negative and negative) manner, i.e. $s_t \in \left\{-2^{q-1}, \ldots, -1, 0, 1, \ldots, 2^{q-1} - 1\right\}$. In this manner, the formalism of binary logic arithmetic provides the FRQA with the adequate tools needed for effective quantum audio processing. Moreover, in digital audio

processing (DAP), the most widespread method to represent bipolar numbers is based on using the notations of two's complement arithmetic. Hence, the extension to encode the amplitude value in quantum audio using the two's complement arithmetic is both analogous and astute. Equation (53) describes the stipulation in that form:

$$S_t = \mathrm{S}_t^0 \mathrm{S}_t^1 \ldots \mathrm{S}_t^{q-1} \mathrm{S}_t^i, \quad i \in 0, 1, \cdots q-1 \tag{53}$$

where $t = 0, 1, \ldots, 2^l - 1$ denotes the time information of a 2l-sized quantum audio signal,$S_t = S_t^0 S_t^1 \ldots S_t^{q-1}$ is the binary sequence encoding the two's complement notation of the amplitude value. Two cases of the binary sequence S_t are listed as follows:

(1) If the amplitude value is non-negative, then $S_t^0 = 0$ and S_t is simply represented as a binary sequence of the value itself.

(2) If the amplitude value is negative, then $S_t^0 = 1$ and S_t is represented by the two's complement mode of its absolute value.

Although one-dimensional in its construction, the FRQA model facilitates descriptions for quantum audio signals in terms amplitude and time components as presented in Equation (54).

$$|A\rangle = \frac{1}{2^{\frac{l}{2}}} \sum_{\mathrm{t}=0}^{2^l-1} |S_t\rangle \otimes |t\rangle \tag{54}$$

where $|S_t\rangle = \left|S_t^0 S_t^1 \ldots S_t^{q-1}\right\rangle$ is the two's complement representation of each amplitude value and $\mathrm{t} \in \{0, 1\}$ is the corresponding temporal information, which the state $|A\rangle$ is deemed normalized, i.e. $||S_t\rangle| = 1$. Based on its formulation, the FRQA model requires $(q + l)$ qubits to represent a quantum audio segment with 2^l samples.

For an L-sized FRQA quantum audio signal, employing l qubits to represent the time information will produce $2^l - L$ audio redundancies [36], [46]. As a default, in the FRQA format, amplitude values of all redundancies are set as $|0\rangle$. Therefore, in the FRQA quantum audio format such a state would be written as presented in Equation (55).

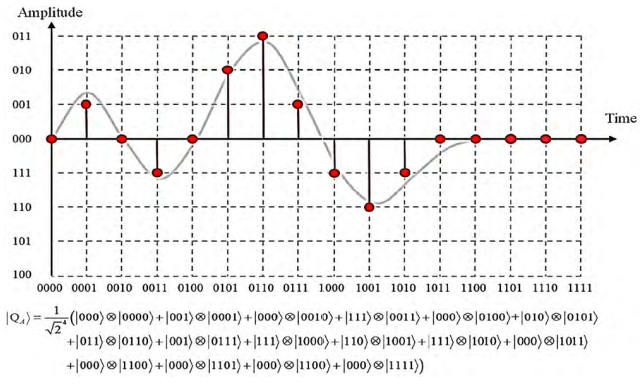

FIGURE 27. Segment of quantum audio signal and its representation in FRQA model (figure adapted from [46]).

Fig. 27 shows a segment of quantum audio signal and its representation in FRQA model. In this example, amplitude values are sampled between -2 and 3, which requires 3 qubits to store. If the size of the audio file is 13, then $l = \lceil log_2 13 \rceil = 4$. Therefore, in total, 7 qubits are required to represent and process this audio segment in the FRQA quantum audio model.

$$\left|A'\right\rangle = \frac{1}{2^{\frac{l}{2}}} \left(\sum_{\mathrm{t}=0}^{L-1} |S_t\rangle \otimes |t\rangle + \sum_{\mathrm{t}=0}^{2^l-1} |0\rangle^{\otimes q} \otimes |t\rangle \right) \tag{55}$$

where

$$L = \begin{cases} \lceil log_2 L \rceil, & L > 1 \\ 1, & L = 1 \end{cases} \tag{56}$$

Compared alongside the only quantum audio format that precedes the FRQA, (i.e. the QRDA), Yan *et al.*'s FRQA is "more intuitive, unbounded, and graphic with the sign (negative or positive) indicating the variation of the waveform [45]," which they claim makes it "conceptually more like the real waveform of digital audio signals [46]." Moreover, it is asserted that "FRQA supports the realization of basic processing operations in an effective manner." For example, (1) FRQA allows two sample values to be accurately added (or mixed in audio parlance) and so, it facilitates the use of quantum circuit elements to handle overflow and wraparound situations that are encountered when size of the resulting audio sequence is greater than the capacity of the register allocated for its storage as well as distortions due to full-scale signal swing (i.e. low and high thresholds); (2) FRQA permits the execution of signal subtraction operation via simple logical addition operations on an augend and addend one of which is inverted (i.e. negated). In this manner, the FRQA offers considerable saving in hardware complexity since the more tedious 'borrow' mechanism associated with the subtraction operation is replaced by a 'carry' procedure in the addition operation.

1) FRQA AUDIO SIGNAL PREPARATION

The quantum mechanical composition of quantum computers imposes a lot of premium on the use unitary operations to transform quantum information. This entails the use of such operations in procedures to prepare quantum registers that encode audio information.

As explained in [45] and [46] and highlighted earlier in this section, in digital audio, the value of each sampled audio point is stored on an assigned q-length binary sequence, i.e. $\left|B_t^0 B_t^1 \cdots B_t^{q-2} B_t^{q-1}\right., B_t^i \in \{0, 1\}$, which indicates that the sequence is equipped to hold the amplitude values from 0 to $2^q - 1$. To convert this sequence into the two's complement system $|S_t\rangle$, we need to perform the value-setting operation which is accomplished via the two sub-operations outlined in the sequel.

First, in the quantization sub-operation, the classical binary representation of B_t is used as reference so that the initialized states in the quantum computer could be transformed into their desired quantum states. The layout on how to

accomplish this has been widely discussed in the literature, notably [77] and [78].

Second, in order to convert a quantized quantum state to its two's complement equivalent, the CNOT gate operation is applied on the Most-Significant-Qubit (MSQb) of the audio sequence (i.e. the leftmost qubit in the sequence encoding the audio sample) to produce the desired amplitude state $|S_t\rangle$.

To illustrate the aforementioned sub-operations, [45] and [46] considered the conversion of a 3-qubit resolution to the amplitude information in quantum audio as presented in Table 3. The resolution values ranging from 0 to 7 are held using a 3-qubit sequence. These values are then binaries and subsequently quantized in two's complement notation for further computation by other quantum audio processing units.

TABLE 3. Mathematical notations for host quantum audio, quantum audio message, and stego quantum audio signal.

	Host audio signal	Quantum audio message	Stego audio signal
FRQA state	$\|Q_H\rangle = \frac{1}{2^{\frac{l}{2}}}\sum_{t=0}^{2^l-1}\|H_I\rangle \otimes \|t\rangle$	$\|Q_M\rangle = \frac{1}{2^{\frac{m}{2}}}\sum_{t=0}^{2^m-1}\|M_I\rangle \otimes \|t\rangle$	$\|Q_S\rangle = \frac{1}{2^{\frac{m}{2}}}\sum_{t=0}^{2^m-1}\|S_I\rangle \otimes \|t\rangle$
Amplitude	$\|H_t\rangle = \|H_t^{q-1}\cdots H_t^0\rangle$	$\|M_t\rangle = \|M_t^{q-1}\cdots M_t^0 M_t^1\rangle$	$\|S_t\rangle = \|S_t^{q-1}\cdots S_t^0 S_t^1\rangle$
Time	$\|t\rangle = \|t_{l-1}\cdots t_1 t_0\rangle$	$\|t\rangle = \|t_{m-1}\cdots t_1 t_0\rangle$	$\|t\rangle = \|t_{l-1}\cdots t_1 t_0\rangle$

As outlined in this example, the value-setting operation Ω_t can be executed using Equations (57) and (58).

$$\Omega_t = \bigotimes_{i=0}^{q-1} \Omega_t^i \tag{57}$$

$$\Omega_t^i = \begin{cases} \overline{|0\rangle \oplus B_t^i}, & i = 0 \\ \overline{|0\rangle \oplus B_t^i}, & i \neq 0 \end{cases} \tag{58}$$

where $\oplus$ is the XOR operation. It is worth clarifying that the operation Ω_t^i works uses an additional CNOT gate to negate the most significant bit only when $B_t^i = 1$; otherwise, it is unchanged. Then, the amplitude value of each sample is set.

In this manner, a unitary transform that encodes the amplitude information by means of two's complement arithmetic is available during the preparation procedure. This procedure itself is accomplished in another two steps as enumerated in the sequel.

First, the initial state $|0\rangle^{\otimes q+l}$ is transformed into an intermediate state $|H\rangle$ via the use of 2-D identity and Hadamard operations (as matrices). This produces the state in Equation (59).

$$|H\rangle = \left(|0\rangle^{\otimes q+l}\right) = \frac{1}{\sqrt{2^l}}\sum_{t=0}^{2^l-1}|0\rangle^{\otimes q+l}|t\rangle \tag{59}$$

The intermediate state (in Equation (59)) is a superposition of all samples in an empty audio file, i.e. one with all amplitude values set as $|0\rangle$.

Second, the value-setting operation Ω_t^i is used to generate the amplitude details of each sample. This requires 2^l sub-operations to execute and control conditions are imposed to restrict the operation itself to different temporal instances, i.e. time. For a given sample, k, this constraint, R_t, helps to transform the intermediate state to the FRQA quantum audio state. It has been observed in [45] and [46] that by using sufficient ancillary qubits, each sub-operation in the second step of the FRQA state preparation can be implemented using $2(l-1)$ Toffoli gates and no more than q CNOT gates. Consequently, complexity-wise, preparation of a 2^l-sized FRQA quantum audio state requires $2^l \times [2(l-1)\times 6 + q] = (12l + q - 12).2^l$ operations to execute. Further details of the preparation procedure can be digested from [45] and [46].

2) FRQA AUDIO SIGNAL PROCESSING

Basic signal operations that propagate more sophisticated applications are the foundations of digital audio processing. The ability to extend similar operations to quantum audio processing is essential to validating the utility of this emerging sub-field of quantum information science. In an effort to accomplish this, based on the FRQA model, Yan *et al.* presented a few basic audio signal operations including signal addition, signal inversion, signal delay, and signal reversal, which are highlighted in this subsection.

- Addition of FRQA quantum audio signals

Signal addition is a fundamental operation in digital audio signal processing. This operation involves the addition of amplitudes of two or more signals at each instant of time to accomplish different audio manipulation tasks, such as echo, reverb addition and active noise reduction. The adeptness of using two's complement system to encode amplitude values in FRQA model facilitates the use of outcomes of arithmetic operations and depicting their waveforms with respect to the midrange values. Consequently, the definition of quantum audio signal addition operation was tailored in terms of this arithmetic advantage [45], [46].

Given two 2^l-sized audio segments $|A_x\rangle$ and $|A_y\rangle$ (presented in Equation (60)-(63)), the signal addition operation produces the output of $|A_z\rangle$ as shown in Equation (60)-(65).

$$|A_X\rangle = \frac{1}{2^{\frac{l}{2}}}\sum_{x=0}^{2^l-1}|S_x\rangle \otimes |t_x\rangle \tag{60}$$

where

$$\begin{aligned} |t_x\rangle &= |t_x^0 t_x^1 \cdots t_x^{l-1}\rangle, \quad t_x^i \in \{0,1\} \\ |S_x\rangle &= |S_x^1 \cdots S_x^{q-2} S_x^{q-1}\rangle, \quad S_x^i \in \{0,1\}, \end{aligned} \tag{61}$$

and

$$|A_y\rangle = \frac{1}{2^{\frac{l}{2}}}\sum_{y=0}^{2^l-1}|S_y\rangle \otimes |t_y\rangle \tag{62}$$

where

$$\begin{aligned} |t_y\rangle &= |t_y^0 t_y^1 \cdots t_y^{l-1}\rangle, \quad t_y^i \in \{0,1\} \\ |S_y\rangle &= |S_y^1 \cdots S_y^{q-2} S_y^{q-1}\rangle, \quad S_y^i \in \{0,1\}, \end{aligned} \tag{63}$$

So,

$$|A_z\rangle = \frac{1}{2^{\frac{l}{2}}} \sum_{z=0}^{2^l-1} |S_z\rangle \otimes |t_z\rangle \tag{64}$$

where

$$\begin{aligned} |t_z\rangle &= |t_z^0 t_z^1 \cdots t_z^{l-1}\rangle, \quad t_z^i \in \{0,1\} \\ |S_z\rangle &= |S_z^1 \cdots S_z^{q-2} S_z^{q-1}\rangle, \quad S_z^i \in \{0,1\}, \end{aligned} \tag{65}$$

On the circuit-model of quantum computation, algorithms are tools dictating the choice, use, and construction of circuit elements to realize advanced operations. The circuit network to implement the addition of two audio signals on the FRQA quantum audio format is presented in Fig. 28.

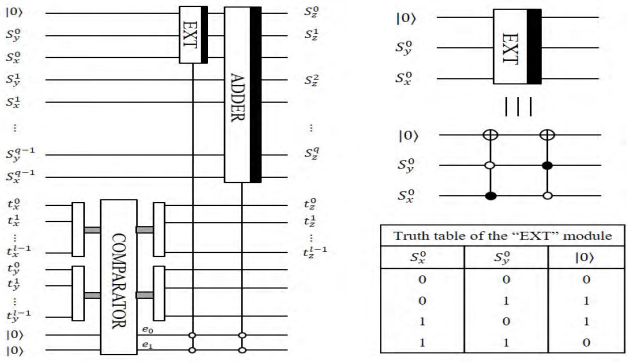

FIGURE 28. Circuit construction to execute the quantum audio signal addition operation (refer to [45] for further details).

As seen from the figure, the FRQA quantum audio signal addition operation consists of three modules: the comparator (COMP), extension (EXT) and adder (q-ADD or ADDER). The COMP module has been extensively highlighted in this treatise, including, most recently, its use in applications for the QRDA quantum audio model (in the previous subsection). Its use here is targeted at the time component of the FRQA quantum audio signal. The 2-controlled extension (EXT) module ensures that overflows resulting from two's complement addition are avoided. The equivalent operation and truth table for the EXT module are presented on the right side of Fig. 28. Finally, the q-ADD (or ADDER) module (also widely discussed in earlier parts of this review) is applied on the amplitude component of the FRQA quantum audio signal.

Formally, the operation (U_A) to accomplish FRQA-based signal addition operation of two quantum audio signals can be written in the form presented in equation (66).

To exemplify the execution of the FRQA quantum audio signal addition operation, we consider the case (where amplitude and range are equal) $l = g = 2$ in Fig. 29, wherein the circuit in (a) is implemented to add the augend (in (b)) to the addend (in (c)) to produce the outcome in (d).

In terms of complexity, the FRQA quantum audio signal addition operation U_A (Equation (66)) requires a network of $24l^2 + 6l + 248q$ operations to execute.

$$U_A = \left|S_x\right\rangle \left|S_y\right\rangle \longrightarrow |S_x\rangle \left|S_x + S_y\right\rangle \tag{66}$$

- Inversion of FRQA quantum audio signals

Inverting a signal is a common and important signal processing operation. A meaningful use of inversion is to transform the signal subtraction to signal addition by means of inverted inputs (audio cancellation is realized as such). In audio signal processing, signal inversion is realized by inverting all amplitude values of an audio signal. When amplitude values are represented by bipolar values, the operation essentially reduces to alterations of positive and negative signs of the signals. In this subsection, the FRQA-based signal inversion operation is formalized.

Assuming that $|A\rangle$ is an FRQA audio sample in the form presented in Equation (60), the signal inversion operation U_I applied on $|A\rangle$ will produce the output of $|A_I\rangle$ in the form shown in Equation (67).

$$|A_I\rangle = \frac{1}{2^{\frac{l}{2}}} \sum_{z=0}^{2^l-1} |S_I\rangle \otimes |t\rangle \tag{67}$$

where

$$\begin{aligned} |t\rangle &= |t_0 t_1 \cdots t_{I-1}\rangle, t_I \in \{0,1\} \\ |S_I\rangle &= |-S_t^0 S_t^1 \cdots S_t^{q-2} S_t^{q-1}\rangle, \quad S_t^i \in \{0,1\}, \end{aligned} \tag{68}$$

As mentioned in [45] and [46], like the two's complement binary arithmetic, an effective way to negate a number is inverting all the qubits and adding '1'. In quantum audio processing (QAP), this procedure is further explained using two steps.

First, an operation U_1^1 is used to invert all qubits in $|S_t\rangle$ as presented in Equation (69) and (70).

$$\begin{aligned} U_1^1 |A\rangle &= \frac{1}{2^{\frac{l}{2}}} U_1^1 \Big(\sum_{z=0}^{2^l-1} |S_I\rangle \otimes |t\rangle\Big) \\ &= \frac{1}{2^{\frac{l}{2}}} \sum_{z=0}^{2^l-1} \left|\overline{S_I}\right\rangle \otimes |t\rangle, \end{aligned} \tag{69}$$

where

$$\begin{aligned} |S_I\rangle &= |S_t^0 S_t^1 \cdots S_t^{q-2} S_t^{q-1}\rangle, \quad S_t^i \in \{0,1\} \\ |\overline{S_I}\rangle &= |\overline{S_t^0 S_t^1 \cdots S_t^{q-2} S_t^{q-1}}\rangle, \quad \overline{S_t^i} \in \{0,1\}, \end{aligned} \tag{70}$$

Then, in the second step, the correction operation is performed by adding '1' to the inverted outcomes from the first step, while neglecting any overflow. Denoting this operation as U_1^2, Equations (71) and (72) formalize the execution of this step of the signal inversion operation.

$$U_1^2 : \left|\overline{S_t}\right\rangle \longrightarrow |S_I\rangle \tag{71}$$

where

$$\begin{aligned} |\overline{S_I}\rangle &= |\overline{S_t^0 S_t^1 \cdots S_t^{q-2} S_t^{q-1}}\rangle, \\ \overline{|S_t^0 S_t^1 \cdots S_t^{q-2} S_t^{q-1} + 1\rangle mod\ 2^q} &= |S_I\rangle \end{aligned} \tag{72}$$

The circuit on the right in Fig. 30 shows the inverted version of the input FRQA quantum audio signal ($|A_z\rangle$)

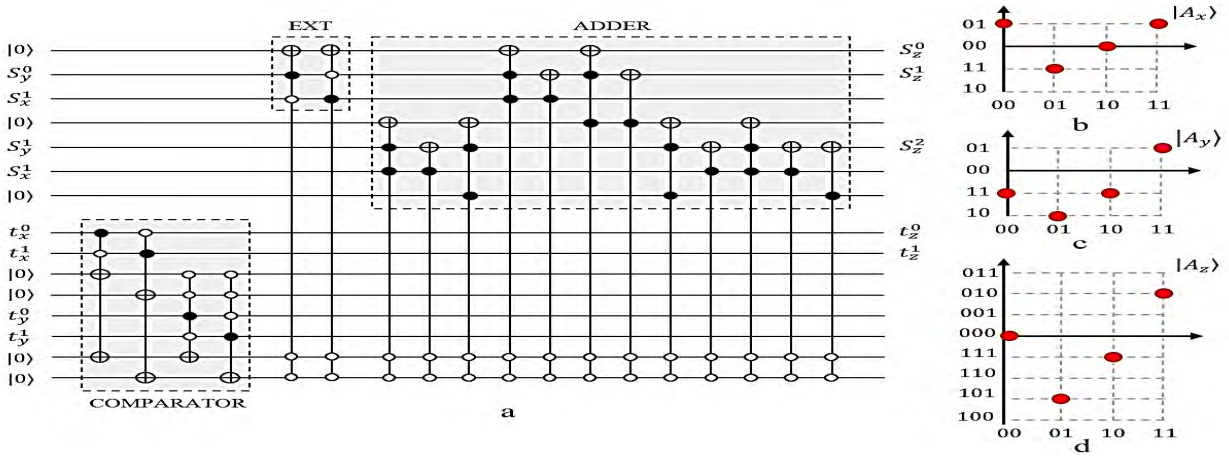

FIGURE 29. Addition of two FRQA quantum audio samples. Clockwise: circuit to implement addition of $|A_x\rangle$ (in (b)) and $|A_y\rangle$ (in (c)) to produce the output $|A_z\rangle$ (in 9d)) (example and figure adapted from [45]).

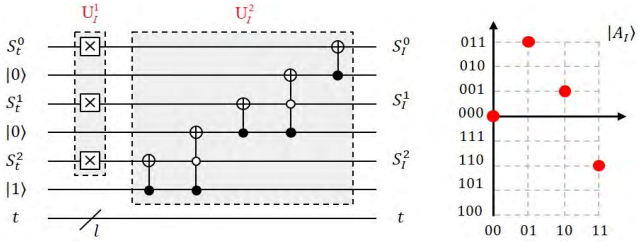

FIGURE 30. Inversion of FRQA quantum audio signal. The input signal $|A_z\rangle$ (from Fig. 29) is inverted using the circuit on the left to produce the inverted signal ($|A_I\rangle$ on the right [45].

(in Fig. 29(a)). It is realized using the FRQA quantum audio signal inversion circuit network on the left (of the same Fig. 30).

Complexity-wise, the U_{I}^1 sub-operation of FRQA quantum audio signal inversion operation requires q CNOT gates to implement, whereas $(q-1)$ Toffoli gates are needed to execute the U_{I}^2 sub-operation. Overall, as presented in [45] and [46], the complexity of the FRQA inversion operation is $7q-6$.

- Delay of FRQA quantum audio signals

Signal delay is the operation that records an input signal and then plays it back after an interval in time [45]. This operation is widely used as an audio effect to create the sound of a repeating and/or decaying echo. In this subsection, basic building blocks to realize the FRQA-based signal delay operation are formalized.

Given an input FRQA quantum audio sample $A(t)$, then $B(t)$ could be considered its delayed version if:

$$A(t) = B\left(t'\right), \quad t' = t + \Delta t, \tag{73}$$

where t denotes the time information of amplitude values, Δt is a fixed interval that specifies the desired delay of the system.

Based on the foregoing, the signal delay operation, U_D, can be formalized using Equations (74) and (75).

$$U_D : A(t) \longrightarrow B\left(t'\right), \quad t' = t + \Delta t, \tag{74}$$

$$B(t') = \begin{cases} 0, & 0 \le t' \le \Delta t - 1 \\ A(t), & \Delta t \le t' \le 2^l - 1. \end{cases} \tag{75}$$

Both the time information t and t' of the original and delayed audio signals are confined in the range $[0, 2^l - 1]$, where, apparently, $\Delta t \le t' = t + \Delta t \le 2^l + \Delta t - 1$ is beyond the interval. For further refinement, the time information can be separated into two parts, i.e. $\left[\Delta t, 2^l - 1\right]$ and $\left[2^l, 2^l + \Delta t - 1\right]$.

Readers are directed to [45] and [46], where additional details of the FRQA quantum audio signal delay operation can be obtained.

In concluding, we note that, as presented in [45] and [46], the circuit network to implement the signal delay operation on FRQA quantum audio samples takes the form reproduced in Fig. 31. The case $l = g = 3$ is presented in Fig. 32 with the signal delayed by 2 time-units. Using the circuit network for FRQA quantum audio delay operation (Fig. 31), the input FRQA quantum audio sample (in Fig. 32(a)) is delayed to the form presented in Fig. 32(b).

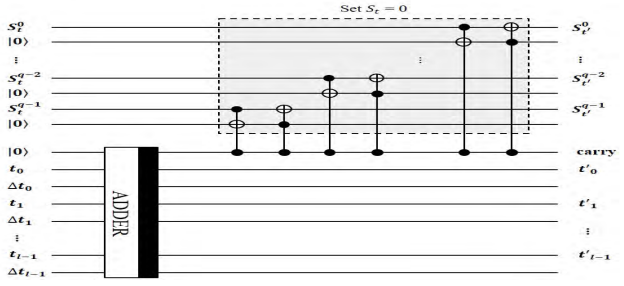

FIGURE 31. Circuit network to implement the FRQA quantum audio signal delay operation (figure and description adapted from [45]).

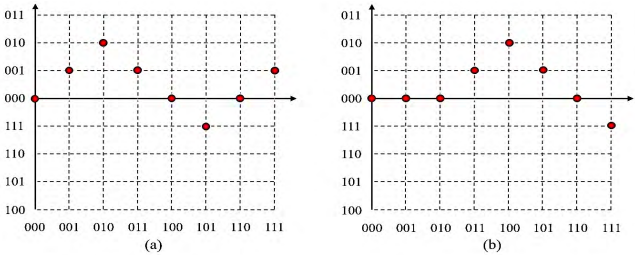

FIGURE 32. (a) Input FRQA quantum audio signal and (b) the delayed version of (a) by 2 units [45].

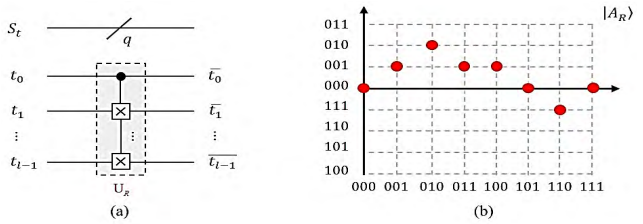

FIGURE 33. (a) Circuit network to execute the FRQA quantum audio signal reversal operation and (b) reversed signal (by 2 time-points) for the sample in Fig. 32(a).

Finally, as noted in [46], implementation of FRQA quantum audio signal delay consists of one q-ADD (or ADDER) module (applied on the time component) and $2q$ Toffoli gates; so, the complexity of this delay operation is $28l + 12q - 12$.

- Reversal of FRQA quantum audio signals

Signal reversal is the process of turning an audio signal bottom-up, such that the end of the signal is heard first and the beginning last. This operation can be used to create interesting sound effects or make small portions of inappropriate language sound incomprehensible when heard. On the FRQA quantum audio model, the audio signal reversal operation can be implemented via the quantum circuit networks highlighted in this subsection.

Given an FRQA quantum audio signal $|A\rangle$ in the form presented in Equation (76).

$$|A_R\rangle = \frac{1}{2^{\frac{l}{2}}} \sum_{z=0}^{2^l-1} |S_I\rangle \otimes \left|\overline{t}\right\rangle \tag{76}$$

where

$$\left|\overline{t}\right\rangle = \left|\overline{t_0 t_1 \cdots t_{I-1}}\right\rangle, \quad \overline{t_i} \in \{0, 1\} \ , \tag{77}$$

the signal reversal operation U_R (applied on $|A\rangle$) produces an output of the form

$$\begin{aligned} U_R|A\rangle &= \frac{1}{2^{\frac{l}{2}}} \sum_{z=0}^{2^l-1} |S_I\rangle \otimes U_R|t\rangle \\ &= \frac{1}{2^{\frac{l}{2}}} \sum_{z=0}^{2^l-1} |S_I\rangle \otimes \left|\overline{t}\right\rangle \end{aligned} \tag{78}$$

The general circuit to execute the signal reversal operation is presented in Fig. 33(a), whereas the output emanating from its application to reserve the FRQA quantum audio signal in Fig. 32(a) is presented in Fig. 33(b). From the FRQA quantum audio reversal circuit network (i.e. Fig. 33(a)), it is glaring that only 1 NOT gate is required to implement this operation. Moreover, using additional control constraints on the time content will support confining the operation to a desired time-period. Some of these instances were demonstrated in [45] and [46] from where readers can obtain further details.

We conclude by noting that, as mentioned in [45] and [46], the complexity of this extended operation to target a specific time-period within the length of the audio signal increases with the number of control qubits. In the worst case, the implementation of FRQA quantum audio signal reversal operation requires an $(l - 1)$-controlled NOT gates, so its complexity is $12l$ - 23.

Motivated by the importance of digital audio signal processing and its role in numerous facets of human endeavor, two models to represent, encode, and process quantum audio signals were proposed in [44]–[46].

Wang's QRDA in [44] uses two entangled sequences to encode amplitude and time components of quantum audio samples. While perceptive, the QRDA was considered susceptible to errors arising from its numerical representation of amplitude values, which also offers a tightly-bounded encoding method for only unipolar (i.e. non-negative) values.

In an effort to address these perceived shortcomings, in [45] and [46], Yan *et al.* proposed a new quantum audio model, the FRQA, which they claim offers a more intuitive, unrestricted model to accurately represent audio content in the quantum computing domain. Structurally, the FRQA protocol encodes amplitude values in two's complement notation so that the arithmetic advantages of doing so can be utilized to facilitate the construction of quantum circuits for amplitude and temporal transformations on the audio content. By exploiting this design, various operations that allow for designing advanced audio processing applications can be proposed [45] , [46].

Aspects of both quantum audio models that were highlighted in this section showcased the potential to accomplish digital-like audio signal processing applications, such as audio signal combination, inversion, addition, reversal, etc. on the quantum computing framework.

VII. APPLICATIONS FOR QUANTUM AUDIO SIGNAL PROCESSING

The characteristics of the physical framework used to embody information are often crucial in deciding the power of a certain computational paradigm.

Quantum computation has emerged as a powerful tool for efficient and secure information processing. As with any effort to process any type of information, as a first step, the requirements for its encoding must be formalized. The QASP studies (i.e. Wang's QRDA and Yan's FRQA) highlighted in the previous section explored the use of quantum mechanical properties, which are at the core of quantum computers, to encode, store, and manipulate audio signals.

On the flexible representation for quantum audio (i.e. FRQA) model [45], [46], two's complement notation was used to encode the audio signal as well as the integration of temporal components into a normalized quantum state. Additionally, the FRQA's structural composition and its interesting similarities with the FRQI quantum image processing model make it flexible and well-suited as a fundamental protocol for quantum audio processing (QAP) applications.

Before then, building on the efforts in [44]–[46], it is natural to explore the use quantum audio models to implement some audio processing applications. In this section, we review the available quantum audio applications, of which one each is built on the FRQA and QRDA models for quantum audio signal processing (QASP).

Based on the FRQA, a quantum audio steganography (QAS) strategy comprising two sub-protocols which utilize the quantum version of the least significant bit (LSB) [82], [85], i.e. the least significant qubit (LSQb), was proposed in [47]. This strategy and its essential parts are highlighted in the next subsection. Then, we review the rudiments of the quantum endpoint detection algorithm that is built on the QRDA model for quantum audio processing [48].

A. PROTOCOL FOR QUANTUM AUDIO STEGANOGRAPHY (QAS)

The term of steganography is derived from the Greek words 'stegos' and 'grafia' which mean 'cover' and 'writing' [82], [85], and, as suggested by the terms, implies that audio steganography (or simply stego) would focus on hiding some secret audio message into a host audio signal for transmission. In the field of traditional information hiding, the classical LSB algorithm is a fundamental method in many security technologies [82]. Convinced by the importance of the LSB algorithm in steganography, Chen *et al.*'s twin quantum audio steganography (QAS) protocols [47] explore its application to enhance secure communication via quantum audio signals. These protocols were designed to provide high imperceptibility (i.e. faithful replication of the original signal) between the host audio and its stego version.

The decisive components of these QAS protocols and their implementation are reviewed in the remainder of this subsection.

B. LSQB AND MSQB-BASED ALGORITHMS FOR INFORMATION SECURITY

We start by restating the expediency of the LSQb algorithm, which motivated its choice for use in Chen's QAS protocols. First proposed in [76], the LSQb is the quantum version of the LSB algorithm that is widely used in watermarking, steganography, and other areas of information security. It has shown potential for use as a basic algorithm to hide a quantum message into a quantum media such as quantum image, quantum video, or quantum audio [26], [42], [47]. As noted in [47], these quantum media are usually represented in the form of a qubit sequence, and the LSQb is the last qubit (of the qubit sequence that is used to encode the amplitude information in the case of audio signal) of the host quantum medium.

Imperceptibility is an important parameter to evaluate the conspicuousness of the hidden audio message that is embedded in the host quantum medium. Although the framework for the classical LSB protocol is widely available in standard information security texts, the basic concepts and procedures of the quantum LSQb-based algorithms (with focus on the QAS) are recounted in the sequel.

First, as originally presented in [47], we highlight the meaning and notations for the three main audio data types: the host or carrier audio signal, the audio message being hidden and the resulting 'stego' audio signal that are used in accomplishing the QAS protocol. These definitions and notations are presented in Table 3 (which itself is reproduced from [47]).

Like in other discussions on steganography (such as [76]), in Chen *et al.'s* protocol in Chen's QAS protocol [47], implementation of a stego procedure involves descriptions of the quantum audio signals: host quantum audio ($|Q_H\rangle$), quantum audio message ($|Q_M\rangle$), and stego quantum audio ($|Q_S\rangle$) which (in [47]) are all encoded as FRQA quantum audio samples (as formulated in Equation (60)). The host quantum audio is the carrier used to hide the secret message in a version of it that is available (i.e. transmitted) in the public domain. A quantum audio message is the secret message that is being embedded and transmitted. Upon modification and the secret message embedded into it, the host quantum audio becomes a stego quantum audio signal that is made publicly accessible. For clarity, the mathematical formulations for these three signals were presented in Table 3, where l and m represent the number of qubits needed to encode the time information of host audio signal and audio message, respectively; whereas q denotes the number of qubits required to encode the respective amplitude information. Based on this, the relationship in Equation (79) would suffice to determine the size of these audio signals.

$$l \geq m + log_2(q) \tag{79}$$

For instance, if the amplitude value is represented by 16 qubits, i.e. $q = 16$, so $log_2(q) = 4$, then $l \geq m+4$ indicates that the number of the qubits (that are used to encode the time information) of the host quantum audio will be 4 qubits more than the quantum audio message.

As presented earlier, the amplitude of each sample in an FRQA audio signal is represented by a qubit sequence. Theoretically, all the qubits (that are used to encode the amplitude information of each sample) in the host quantum audio can be substituted with the qubits (that are used to encode the amplitude information of each sample) in the quantum audio message. However, embedding the qubits encoding the quantum audio message into different positions in the host quantum audio signal will negatively influence the imperceptibility of the stego quantum audio signal. For instance, consider a 10-qubit sequence (i.e. 1010001101) as shown in Fig. 34, where the LSQb is the rightmost qubit and the MSQb is the

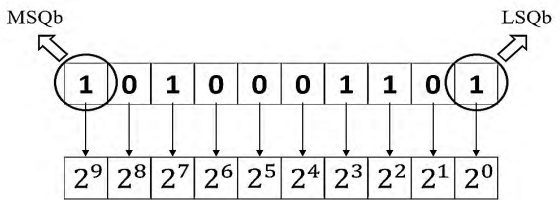

FIGURE 34. LSQb and MSQb in the amplitude of an FRQA quantum audio sample [47].

leftmost qubit. When the embedding operation is performed on the LSQb, 2^0 of amplitude values (of the host quantum audio signal) will be altered, whereas performing the same operation on the MSQb alters 2^9 of amplitude values of the host quantum audio signal will be altered. The latter case will significantly affect content of the host quantum audio signal. This outcome suggests a correlation between the depth (in terms of the more significant qubit) of the position that the audio message is embedded and the imperceptibility between the host and stego quantum audio signals.

In [47], the concept of qubit layer in the LSQb-based QAS algorithm and how it indicates the qubit position (in the host quantum audio) where the quantum audio message will be embedded were discussed. To demonstrate this, as detailed in [47], consider an FRQA quantum audio segment with 13 samples, where the amplitude information of each sample is encoded using 3 qubits, as presented in Fig. 35. The qubits at the same position (depth) along the qubit sequences (that are used to represent the amplitude information of the samples in FRQA quantum audio signal) are located in the same qubit layer. In this manner, in Fig. 35, the first qubit layer is the LSQb layer, whereas the third layer is the MSQb layer. Considering the amplitude information $|110\rangle$ (at time $|1001\rangle$, as an example), the first qubit '1' is the value of the MSQb layer while the last qubit '0' is the value of the LSQb layer. The QAS protocols proposed in [47] focus on hiding information based on choice of related qubit layers.

Like in other standard schemes, the LSQb-based QAS protocol comprises of an embedding and extraction procedure. The purpose of the embedding procedure is to "furtively deliver a hidden message masked in some audible host quantum audio signal" [47].

In the extraction procedure, the aim is to undo the embedding process by recovering the pristine, unadulterated audio signal. In other words, the extraction procedure detaches the hidden message from the stego version of the host quantum audio signal. The basic outline of the embedding operation in Chen *et al.*'s LSQb-based QAS is reproduced in Fig. 35, where the double-headed arrows in the short-dashed rectangular block indicates the exchange of qubits between the host quantum audio signal ($|Q_H\rangle$) and the hidden quantum audio message ($|Q_M\rangle$). In the extraction procedure (presented in Fig. 36), given an initialized qubit sequence (i.e. each qubit is preset as '0'), the extracted audio signal ($|Q_{\text{ex}}\rangle$) is obtained by exchanging the qubits in the initialized audio signal ($|Q_{\text{in}}\rangle$) with the corresponding qubits from the stego audio sample ($|Q_S\rangle$).

1) LSQB PROTOCOL FOR FRQA QUANTUM AUDIO STEGANOGRAPHY

By vivifying a comprehensible use of the least and most significant qubits (i.e. LSQb and MSQb) of an FRQA quantum audio signal, the modus operandi for two QAS protocols to obscurely transmit hidden messages masked in unrestricted quantum audio signals was explored and discussed extensively in [47].

In the proposed LSQb-based stego protocol (i.e. the cLSQ stego protocol), the quantum audio signals are all encoded as FRQA quantum audio signals in the format presented in Equation (60). Each qubit (that is used to encode the amplitude component of each audio sample) in a quantum audio message is embedded into the LSQb of an audio sample in the host quantum audio signal. Therefore, a segment whose amplitude component is encoded using q qubits, requires q samples in the host quantum audio signal to accommodate the amplitude information of one sample of the quantum audio message. Since the embedding procedure is targeted at the LSQb qubit, Chen *et al.* decided to call this approach the

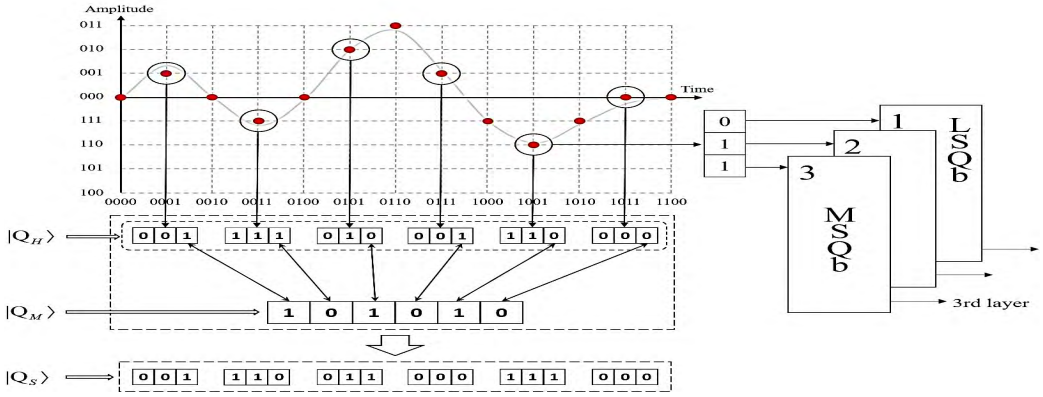

FIGURE 35. Embedding procedure for LSQb-based QAS protocol and qubit layer of an FRQA quantum audio signal (figure and descriptions adapted from [47]).

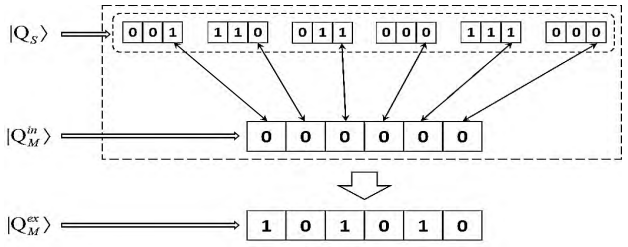

FIGURE 36. Extraction procedure for LSQb-based QAS protocol (figure and descriptions adapted from [47]).

conventional LSQb QAS protocol or simply the cLSQ stego protocol [47].

The layout of the circuit network for embedding a quantum audio message ($|Q_M\rangle$) onto a host quantum audio signal ($|Q_H\rangle$) using the cLSQ protocol is presented in Fig. 37, where $|Q_H\rangle$ consists of amplitude and time components given as $|H_t\rangle$ and $|t_{l-1}\cdots t_1 t_0\rangle$, respectively. Similarly, $|Q_M\rangle$ includes amplitude and time components $|M_t\rangle$ and $|t_{m-1}\cdots t_1 t_0\rangle$, respectively. Both $|Q_H\rangle$ and $|Q_M\rangle$ require the same number of qubits (i.e. q) to encode the amplitude information for each audio sample, whereas l, m, and q were chosen to guarantee conformity with the specifications in Equation (79). The m CNOT gates at the beginning of the circuit (in Fig. 37) are used to ascertain whether the hitherto m qubits used to encode the time component of $|Q_H\rangle$ is equal to those needed to encode the time component of $|Q_M\rangle$ that is being embedded. If they are equal, these m CNOT gates will modify the time component of $|Q_M\rangle$ from $|t_{m-1}\cdots t_1 t_0\rangle$ to $|0\cdots 00\rangle$ after which the swap operation will be used to interchange the qubits in $|Q_H\rangle$ with those in $|Q_M\rangle$ as demonstrated in Fig. 37. Else, the swap operation fails, i.e. it will not work.

As discussed in [47] and mentioned in earlier parts of this subsection, the amplitude component of $|Q_M\rangle$ represented by $\left|M_t^{q-1}\cdots M_t^1 M_t^0\right\rangle$ can be embedded into a corresponding sample in $|Q_H\rangle$. To assign appropriate embedding positions to the q qubits of $|M_t\rangle$ into the samples of $|Q_H\rangle$, the preceding steps of the cLSQ protocol are used. Additionally, the remaining $l-m$ qubits, i.e. $|t_{l-1}\cdots t_{m+1}t_m\rangle$, are used to determine which qubit in $\left|M_t^{q-1}\cdots M_t^1 M_t^0\right\rangle$ should be embedded. For instance, as presented in Fig. 37: if $|t_{l-1}\cdots t_{m+1}t_m\rangle = |0\cdots 00\rangle$, then $\left|H_t^0\right\rangle \rightleftharpoons \left|M_t^0\right\rangle$; if $|t_{l-1}\cdots t_{m+1}t_m\rangle = |1\cdots 00\rangle$, then $\left|H_t^0\right\rangle \rightleftharpoons \left|M_t^1\right\rangle$ is used, and so forth (where $\rightleftharpoons$ indicates the swap operation that is indicated as '$\times$' at the end of each qubit in the circuit). As used in [47], the picked embedding qubit 'PEQ' unit in Fig. 37 is used to retrieve the embedding target qubit, i.e. by determining which qubits in $\left|M_t^{q-1}\cdots M_t^1 M_t^0\right\rangle$ should be swapped with qubits ($\left|H_t^0\right\rangle$) in samples of $|Q_H\rangle$.

The extraction procedure forms the second stage of the cLSQ stego protocol and, like in other steganography applications, it is used to detach the message $|Q_M\rangle$ that is hidden in a stego quantum audio signal ($|Q_S\rangle$), thereby recovering the original audio signal. This process is executed using the circuit in Fig. 38. We recall that, as discussed in [47], a quantum audio message is initialized as $|0\cdots 00\rangle$ (coded as $|Q_{\text{in}}\rangle$) and it has $q+1$ qubits, where the first q qubits represent the amplitude components and the last qubit represents the time components of the audio signal. The desiderata in [47], specifies that a series of m Hadamard gates are used on the lower part of the quantum (audio message) circuit network, so that the time component of $|Q_{\text{in}}\rangle$ is transformed from $|0\cdots 00\rangle$ to $|t_{m-1}\cdots t_1 t_0\rangle$. Mathematically, this operation can be formalized as presented in Equation (80).

$$\underbrace{H\otimes H\otimes\cdots\otimes H}_{m}\left|\underbrace{0\cdots 00}_{m}\right\rangle = \frac{1}{\sqrt{2}^m}\sum_{t=0}^{2^m-1}|t\rangle \tag{80}$$

Like its earlier use in Fig. 37, the 'PEQ' unit is used to retrieve the embedding target qubit by determining which qubits in $|Q_S\rangle$ and $|Q_M\rangle$ should be swapped. The outcome of this step is an extracted audio message ($|Q_{\text{ex}}\rangle$) that has been detached from the stego quantum audio signal, i.e. $|Q_S\rangle$.

2) PSEUDO-MSQB PROTOCOL FOR FRQA QUANTUM AUDIO STEGANOGRAPHY

The MSQb-based protocol reviewed in this sub-section is tailored for embedding the audio message onto the most

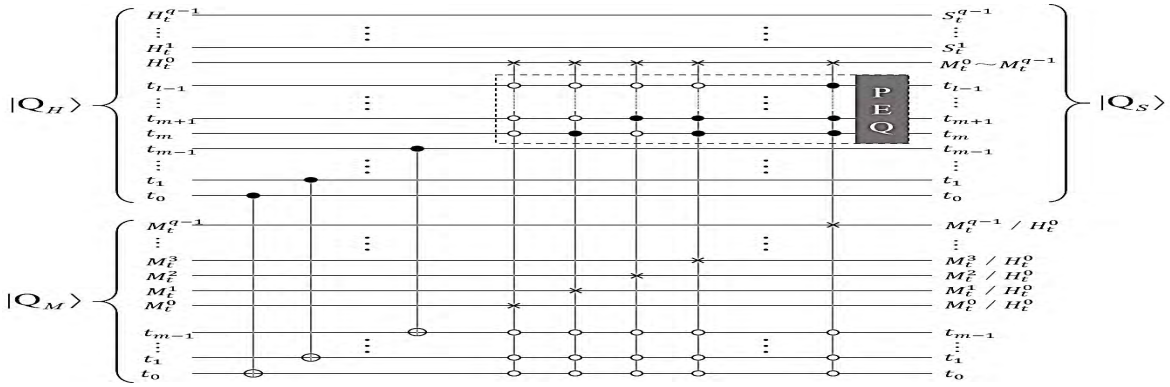

FIGURE 37. Layout of circuit network for the embedding process of the cLSQ stego protocol (figure and descriptions adapted from [47]).

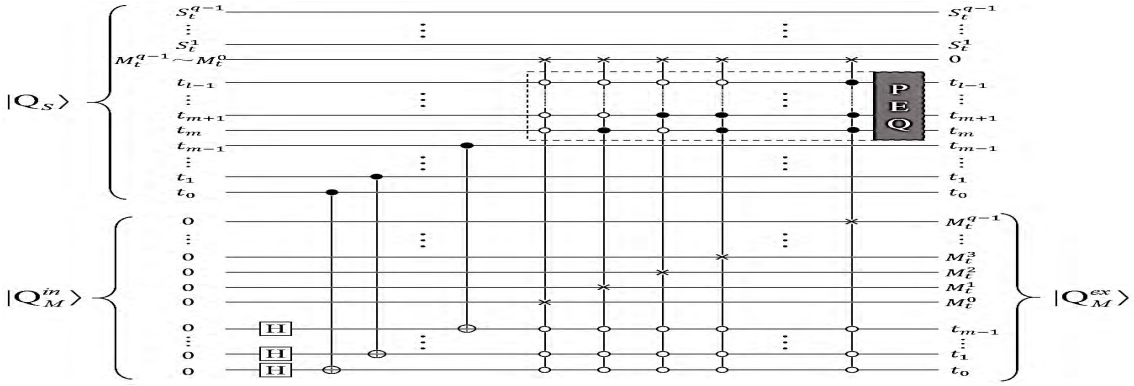

FIGURE 38. Layout of circuit network for the extraction process of the cLSQ stego protocol (figure and descriptions adapted from [47]).

significant qubit of the host quantum audio signal. Its purpose is to enhance the robustness of the resulting stego audio signal. Execution of this protocol mimics its use in digital audio steganography [81], therefore, based on some predefined constraints [47], it is used to modify the choice of the target qubit to the i^{th} qubit layer. This modification is intended to alter the target qubit for the embedding procedure and hence the name pseudo MSQb QAS protocol (or simply the pMSQ stego protocol) was adopted for the second QAS protocol in [47]. However, it has been observed in [47] that increase in the number of qubits required to encode the temporal information, i, (towards m) decreases the imperceptibility, i.e. inconspicuousness, of the quantum audio message in the host audio signal. Furthermore, it has been noted that this "conciliation increases the robustness of the stego audio signal [47]". This trade-off between imperceptibility and robustness enhances the integrity of the stego quantum audio signal that is published in the public domain and as is commonplace in most data hiding schemes. Moreover, [47] has established that when the embedding qubit of the quantum audio message differs from the qubit at the i^{th} layer of the host quantum audio signal, it (the host audio signal) will be changed by $2^i - 1$ after the embedding process. For instance, if the i^{th} layer is the same as the MSQb layer, the stego quantum audio signal will differ significantly from the host audio signal. In other words, the imperceptibility will degrade, which contradicts an important objective of audio signal steganography. This indicates importance of the choice of i^{th} layer in obtaining a reliable stego quantum audio signal. Other requirements dictating the choice of the destination qubit (DQ) [36], [47] and other parameters needed to execute the pMSQ stego QAS protocol are reproduced in Fig. 39.

To determine the embedding position for a host quantum audio signal with amplitude $|H_{q-1}\cdots H_{i+1}H_iH_{i-1}\cdots H_0\rangle$ having value $|H_i\rangle$ at the $(i+1)^{th}$ layer is as follows: first, the variability between DQ in the quantum audio message ($|Q_M\rangle$) and that in the qubit in the i^{th} layer of the host quantum audio signal ($|Q_H\rangle$) is established. If they are the same, then the value of $|H_i\rangle$ remains unchanged, i.e. no operation is

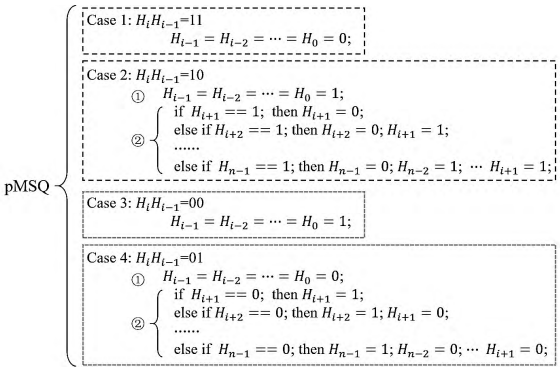

FIGURE 39. Parameters for execution the pMSQ quantum audio stego protocol (figure and descriptions adapted from [47]).

performed on $|Q_H\rangle$. Otherwise, the swap operation is used to exchange between qubits in $|Q_M\rangle$ and $|Q_H\rangle$. Further details of the execution of the pSMQ stego protocol are illustrated in Fig. 40, while in-depth discussions and examples demonstrating its use can be obtained from [47].

As reproduced in Fig. 41, the embedding procedure for the pMSQ stego QAS protocol comprises of two main units: the embedding and destination constraint units. Furthermore, within each of these units to facilitate execution of the 'PEQ' in the embedding unit as well as four sub-units (F1', 'F0', 'H1', and 'H0) in the destination constraint unit. The interoperation between these units to facilitate execution of the pMSQ stego protocol is exhaustively discussed in [47], where readers are referred for the details.

The details of the sub-circuits labeled 'H0' and 'H1' in Fig. 41 are reproduced in Fig. 41(a) and (b), respectively. The qubit sequence $|x_{n-1}\cdots x_ix_0\rangle$ represents the input of the 'H0' sub-unit showing that $2n$ CNOT gates as well as $|0\cdots00\rangle$ ancillary input qubits are required to execute the 'H0' operation. Similarly, the 'H1' sub-unit is used to transform the input sequence $|x_{n-1}\cdots x_ix_0\rangle$ to $|1\cdots11\rangle$. The notations '$\circ$' and '$\bullet$' represent the 0-controlled and 1-controlled constraints that the operations in the 'H0' and 'H1' sub-units

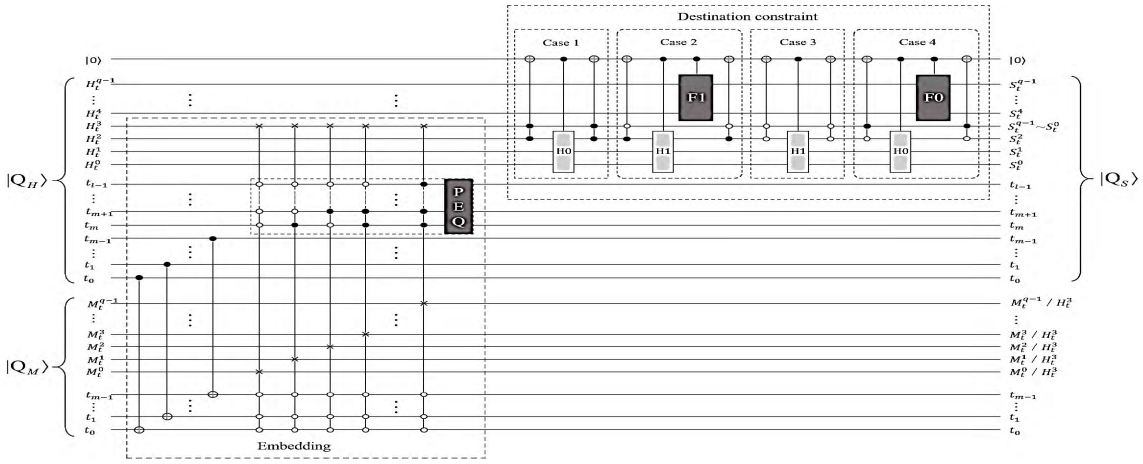

FIGURE 40. Layout of circuit network for the embedding process of the pLSQ stego protocol (figure and descriptions adapted from [47]).

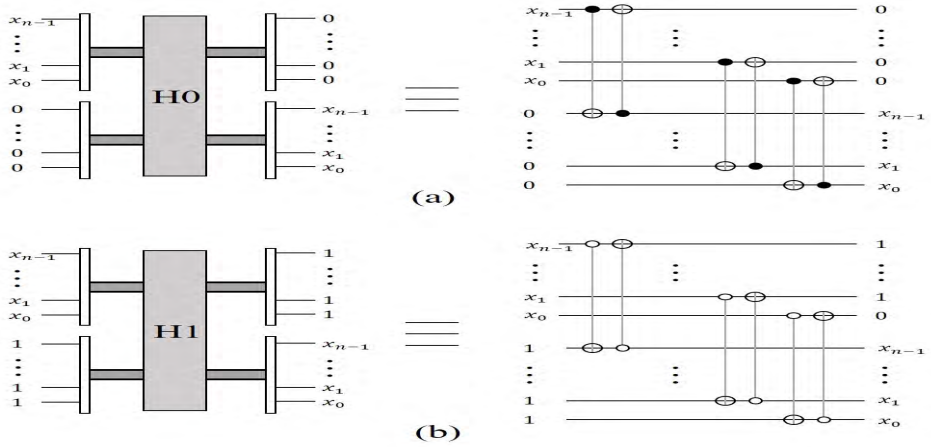

FIGURE 41. Details of circuitry to execute the 'H0' and 'H1' operations in the stego protocol in Fig. 40 (figure and descriptions adapted from [47]).

must satisfy before being executed (i.e. the operations are executed if and only if those 0/1 (zero-one) conditions are satisfied). Similarly, the sub-circuits for implementing the 'F0' and 'F1' operations are reproduced in Fig. 42(a) and (b).

The utility of our proposed pMSQ stego protocol lies in the practicality of the gate execution operations outlined above as well additional details provided in [47].

The procedure to undo the embedding process and detach the audio message, $|Q_M\rangle$, from the stego audio signal, $|Q_S\rangle$ in the pMSQ stego protocol (i.e. the pMSQ stego extraction procedure) is similar to the one in the cLSQ stego protocol that was discussed earlier. Therefore, the same circuit network for the cLSQ protocol extraction procedure would suffice to recover the original audio message, $|Q_{\text{ex}}\rangle$, in the pMSQ stego protocol. Consequently, the circuit network of the extraction procedure for the pMSQ stego protocol (in Fig. 43) resembles the extraction circuit for the cLSQ stego protocol (in Fig. 38). The only difference between these circuit networks is in terms of the control conditions needed to undo the embedding of the audio message, which is attributed to implementing the embedding operation in the more significant layer of the host quantum audio signal (for the pMSQ stego protocol). Further similarities as well as performance evaluation of the two QAS protocols that are discussed in [47].

To conclude, we recount that the study reviewed in this section assessed the plausibility of executing steganography protocols on quantum audio signals. Specifically, two quantum audio steganography (QAS) protocols were proposed based on the least significant qubit (LSQb) algorithm. Utilizing the properties of the FRQA quantum audio model used to encode the host audio signal and audio message being hidden, the circuit networks required to implement the embedding and extraction operations of the two protocols were realized. Simulation-based experiments that were reported in [47] demonstrate the effective implementation of the proposed stego protocols. Furthermore, an analysis of the outcomes attests that both protocols offer promising trade-offs in terms of imperceptibility and robustness.

As at the time the study (i.e. [47]) was completed, among others, the authors claimed that their ongoing and future work were focused on: developing a "suitable trade-off between

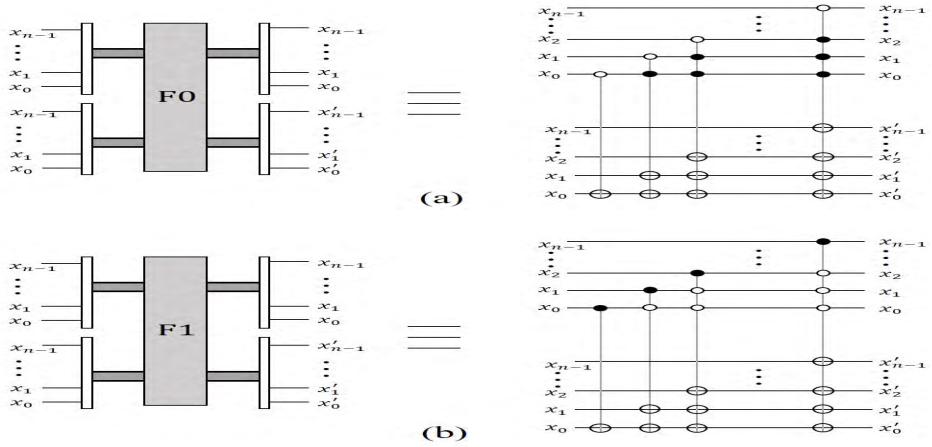

FIGURE 42. Details of circuitry to execute the 'F0' and 'F1' operations in the stego protocol in Fig. 40 (figure and descriptions adapted from [47]).

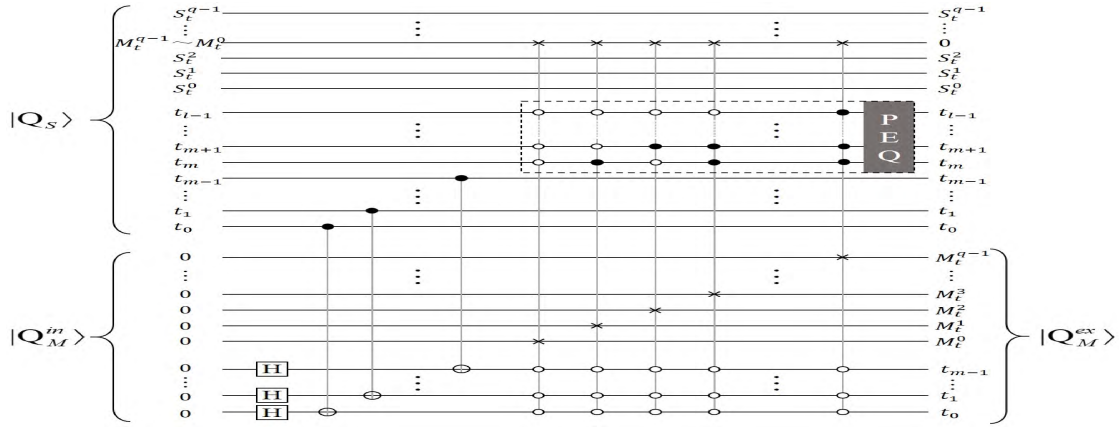

FIGURE 43. Layout of circuit network for the extraction process of the pLSQ stego protocol (figure and descriptions adapted from [47]).

imperceptibility and robustness," which, [47] observed, had "proven difficult to achieve in QAS." This and other directions discussed elsewhere in this treatise present interesting perspectives for readers to explore.

C. ENDPOINT DETECTION IN QUANTUM AUDIO SIGNALS

Speech is arguably the most general means of communication between people. Once there is a common language, comprehension is straightforward.

With further advances in computing technologies, automatic speech recognition systems (ASR) are becoming increasingly popular. ASR facilitates speech to text (S2T), which plays an important role in assisting communication between and with the visually-impaired as well as offer support for emerging human-computer interaction (HCI) interfaces [83]. The utility of ASR imposes new challenges in terms of accuracy and efficiency of the systems, more so considering the large amounts data and training models involved [84].

Ascertaining presence of and quantifying data integrated into mixed media (videos, etc.) is another daunting task. The process of sieving through large quantities of such media to isolate audio content mixed with other types of media and/or isolate it from noisy backgrounds is referred to as endpoint detection (EPD) [85]. The wide-ranging scope of EPD systems suggests its importance in other applications, notably, ASR, HCI, etc. [83], [84].

On the digital computing framework, spectrum entropy computation [84] and double threshold comparison [81] are among the standard approaches used for EPD [48], [85]. Both methods come with intense computational demands.

When the primary objectives of an application involve leveraging large amounts of data efficiently and speedily, quantum computing has been touted as a veritable tool to use [106]. With recent efforts to implement audio processing on quantum computers (QASP), whence the potentials of quantum mechanics can be exploited, in [48], Wang *et al.* explored the implementation of the EPD on the quantum audio processing (QAP) framework. The resulting quantum endpoint detection (QEPD) protocol utilized the QRDA quantum audio model that was reviewed earlier in this treatise, to represent as well as manipulate the quantum audio content. This protocol (i.e. quantum endpoint detection (QEPD)) is highlighted in the remainder of this subsection.

1) QRDA-BASED ENDPOINT DETECTION

The flowchart in Fig. 44 outlines the different stages of Wang *et al.'s* QEPD algorithm [48]. As seen in it, the speech audio input is segmented into frames, a process referred to as windowing. Next, relevant features are extracted from the frames and, finally, thresholds are used to delineate the useable audio content from the noisy backgrounds. Therefore, based on the delineation, start and end points in the speech content can be concatenated. These various stages of the QEPD protocol are further dissertated in the sequel.

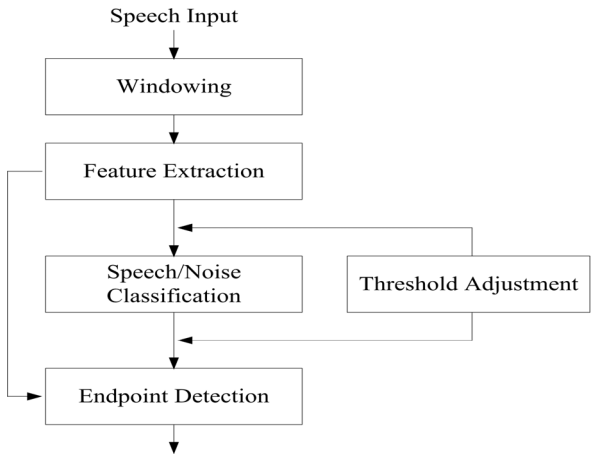

FIGURE 44. Flowchart for execution of Qu *et al.'s* QEPD algorithm [48].

Given an audio sample $|A\rangle$ whose N samples are grouped into frames such that the i^{th} frame is a QRDA quantum audio state $|A_i\rangle$ as defined earlier in Equation (39) – (41), then the resulting frame has samples from $[N\star(i-1)-M\star(i-1)]$ to $[N\star i-M\star(i-1)]$ where M represents the sample adjacent to N.

As outlined earlier, Wang's QEPD algorithm [48] is executed in three main stages. First, a rule $L=[N\star i-M\star(i-1)]$ is used to preselect frames. For example, suppose $N=4$ and $M=2$, then $L=2i-2$, which implies that such an audio segment would have 2 frames and 6 samples where each sample has values between 0 and 2^{q-1}. The second stage of the QEPD involves computing the short-term energy of each frame using Equation (81). The energy value of one frame information (known as short-term energy), E_n, denotes the short-term voice signal, $x_n(m)$ of frame n. Finally, the outcome of the second stage is compared with a threshold.

$$E_n=\sum_{\mathrm{m=1}}^{\mathrm{N-1}} x_n^2(m) \tag{81}$$

where n is an integer and N denotes the length of the audio frame.

In terms of the circuit-model of quantum computing, computation (algorithms) evolves in a 2^n-D Hilbert space of n qubits via the use of unitary transformations as quantum gates. Consequently, execution of the QEPD algorithm could be accomplished using four modules, each a sequence of unitary transformation, as reproduced in Fig. 45 (from [48]) and further explained below.

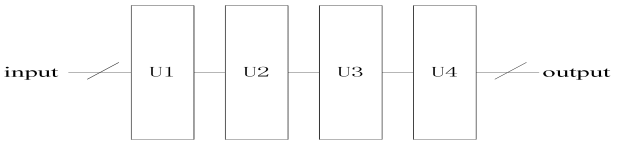

FIGURE 45. Outline of Wang's QEPD algorithm as a sequence of unitary transformations (figure adapted from [48]).

2) SQUARING MODULE OF QEPD

As the name suggests, all audio samples from the first stage are squared in this module. This requires storing the samples in q qubits and, as presented in Fig. 43(a), D_T^0 to D_T^{q-1} qubit locations are allocated for this purpose. D_T^q is an auxiliary qubit used for the squaring operation, which itself requires q swap operations to execute. Fig. 46(a) and (b) show the universal and detailed circuits for squaring operation, respectively.

The squaring operation implemented in this module transforms the QRDA quantum audio input into the new state presented in Equation (82).

$$|B_i\rangle=\frac{1}{\sqrt{2^l}}\sum_{\mathrm{T=0}}^{\mathrm{N-1}}|D_T\rangle\otimes|T\rangle \tag{82}$$

where $\left|D_T\right\rangle=\left|D_T^0D_T^1\cdots D_T^{q-1}D_T^q\right\rangle$ is the squaring sequence for the sample with $\left|D_T^q\right\rangle$ as the ancillary qubit.

3) SHIFTING MODULE OF QEPD

The shifting module implements the feature classification as a superposition of audio content. In executing this operation, all audio samples in a frame are stored in q qubits. The generalized notation and detailed circuit networks for the shifting module are reproduced in Fig. 47(a) and (b) respectively.

As stipulated in the circuit model of quantum computation, the shifting module is the squared QRDA audio content (i.e. the output $|B_i\rangle$) from the squaring module. This is transformed via the shifting operation $U2$ (Fig. 45) to produce an output as presented in Equation (83).

$$|C_i\rangle=\frac{1}{\sqrt{2^l}}\sum_{\mathrm{V=1}}^{q}\sum_{\mathrm{T=0}}^{\mathrm{N-1}}\left|D_T^V\right\rangle \tag{83}$$

where V represents the number of bits and based on Fig. 47(b) shows n 'shadow' qubits using which data can be sampled out, so that n samples are stored into $D_0^qD_1^q\cdots D_n^q$.

4) ADDER MODULE OF QEPD

The adding operation executed in this module has been extensively discussed and used many parts of this disquisition (as the q-Add or ADDER module or operation). Consequently, for brevity, we will overlook the need to repeat most of its discussion again. However, it should be clarified that, although computationally unchanged, as used in the QEPD algorithm [48], the adding operation uses the so-called plane

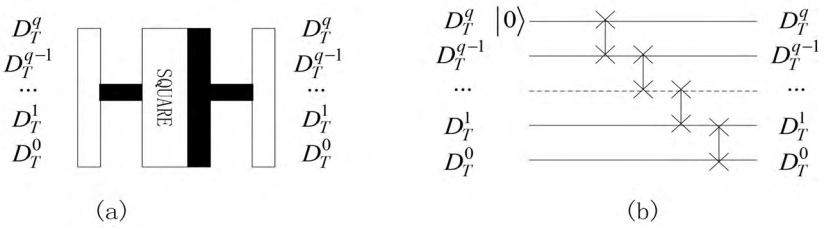

FIGURE 46. Squaring operation used in Wang's QEPD algorithm (figure adapted from [48]).

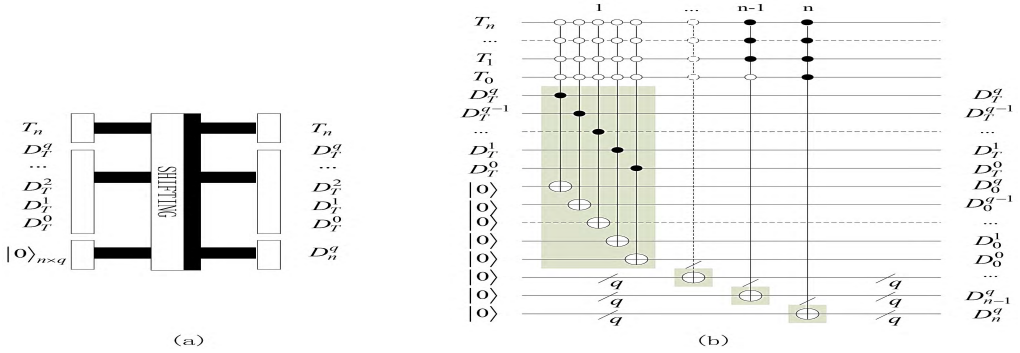

FIGURE 47. Shifting operation of the QEPD algorithm (figure adapted from [48] from where additional details can be obtained).

adder circuit to sum audio samples. Therefore, for 2^n samples, $2^n - 1$ plain adder networks are required.

5) COMPARATOR MODULE OF QEPD

As used in [48], the original nomenclature 'adjudging' module is a misnomer since the main operator used is the comparator (COMP) operator, which has been extensively revised and used in this treatise. Perhaps, the notion of adjudgment arose because, as used in the QEPD algorithm [48], n-bit comparator (or simply n COMP) operators are used. In the n COMP operation, two inputs x and y are compared in terms whether $x > y$, $x = y$ or $x < y$. Using appropriate control conditions, the "adjudging" module of the QEPD uses the n COMP operation to compare if $x \geq y$ or $x < y$. To clarify this, in [48], it was supposed that a qubit z is used to store the results of the computation. If the outcome satisfies the conditions imposed (i.e. $z = 1$), then $x \geq y$; otherwise, $x < y$. The circuit network for the n COMP operator is reproduced in Fig. 48.

Finally, by combining the circuit networks (for the four modules) we obtain the circuit implementation of the QEPD algorithm, which is presented in Fig. 50. It is noteworthy that this extended circuit expands the modular outline presented earlier
in Fig. 45.

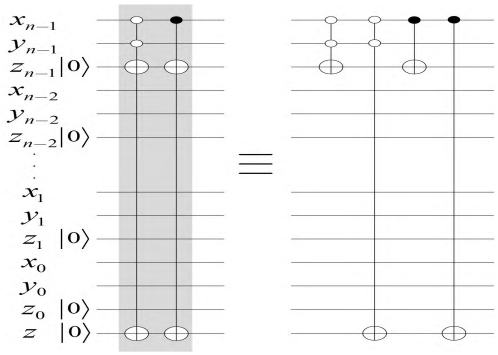

FIGURE 48. Equivalent circuit network for the n COMP operation (figure adapted from [48]).

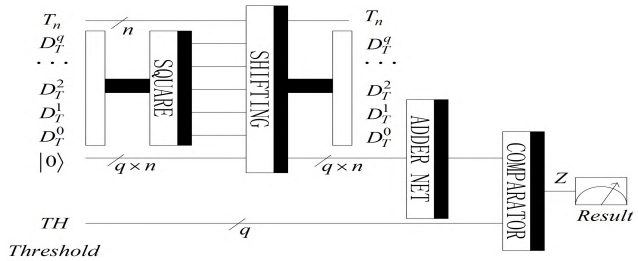

FIGURE 49. Circuit network to exectute the QEPD algorithm (figure adapted from [48]).

VIII. BLUEPRINT FOR REALIZING TALKING COLOR QUANTUM MOVIES

The preceding sections of this review presented a timeline regarding the progress recorded along two emerging sub-topics of quantum image processing, i.e. quantum movie and quantum audio signal processing. The review exhaustively assessed the formulation, requirements, and potentials for both types of media, with the conclusion that, separately, they both have potential roles in future quantum information processing. Learning from the experiences with digital movies and the imminent quest for talking color quantum movies, in this section, we undertake an explorative assessment of the requirements, limitations, and likely impediments

FIGURE 50. Variable-area (left) and variable-density (right) sound-on-film (SOF) systems.

that may mitigate the realization of sound movies on the quantum computing paradigm.

A. CHRONICLE OF SOUND IN DIGITAL MOVIES

Predating today's record player or its recent progeny, the gramophone, is the phonograph whose history extends to the late 1800s. It was designed as a device to transcribe sound via vibrations, themselves as a series of tiny pits on a tinfoil surface mounted on a revolving cylinder [86]. It served as the ''first device to playback recorded audio recordings [86].''

Encouraged by this feat, the fascination to realize sound movies, whence audio and video would be played in sync, was further emboldened. However, it took the persistence of the Warner brothers supported by the invention of electricity to truly provide the needed impetus for coalescing audio and movie contents; thus, bringing talking movies to fruition [87]. Microphones replaced the bulky horns used in Edison's recording device, while electricity the necessary fringe. Electricity provided the fringe to amplify the audio recording as well as support for various stages of the audio-movie synchronization [87].

The next noteworthy contribution in the odyssey of sound movies is the development of Vitaphone (which roughly translates to mean ''living sound'') movies in the early 1920s [88]. In this new device, audio discs were played on a turntable that is physically coupled to a projector motor while the film was being projected [89]. This gave birth the sound-on-disc (SOD) approach to realizing sound movies. SOD systems are a class of sound film processes that use a phonograph or other disc to record or playback sound in sync with movie contents [90].

Early days Vitaphone-equipped theatres had projectors, phonographs, faders, amplifiers and loudspeaker systems [90]. For synchronization, a start mark on the movie strip is aligned with corresponding soundtrack disc on the turntable [90]. The phonograph needle was carefully placed on an arrow scribed on the record's surface [89]. Once projection is started, it rotated the linked turntable, which, theoretically, maintained the synchronization between the audio and projected stack of images [91]. The lines change area (by growing broader or narrower) depending on the magnitude of the signal [92]. Finally, optics-based systems are then used to coalesce the audio content with the video content [92].

While pioneering in the aspiration to realize sound movies, SOD systems, notably, Vitaphone, are vulnerable to over synchronization (which arises when a record is improperly cued up), error correction (this renders entire recordings useless, if, for example, an error related to the choice of audio content, is made), damages to film prints (which leads to loss of sync) [92]. Moreover, simple editing tasks on early Phonographs involved using synchronous motors and switch-triggered playback mechanisms as well as highly complex phonograph-based dubbing systems [92]. Although the system worked, it suffered from poor alignments, and, often, the film reel required adjustments, which could involve removing or adding some movie frames [93].

Phonofilm is the sound-on-film (SOF)'s optical equivalent of the phonograph in SOD systems. It was conceived as a medium to ''record sound directly onto movie content as parallel lines [93].'' These lines photographically recorded electrical waveforms from a microphone, which were translated back into sound waves when the movie was projected [94], [95].

Early SOF technology suffered from poor fidelity and quality. Later improvements, such as the (Radio Corporation of America) RCA-propreited photophone (which made use of ''variable-area'' film exposure systems) included a modulated area corresponding to the waveform of an audio signal, and the Fox Movietone, which uses ''variable-density'' approach for the same purpose provided advanced templates to realizing sound movies [94]–[96]. Figure 50 illustrates the ''variable-area'' and ''variable-density'' SOF systems [96].

On analogue systems, SOF techniques accomplish audio-video synchronization via stereo variable-area (SVA) recording [94]. Here, a two-channel audio signal is recorded as a pair of lines running parallel to the film's pathway [94]. At appropriate frame transition rates, the illusion of motion in sync with audio content is created.

Digitally, SOF techniques use digital technology to digitize the analogue audio content. Subsequently, these digitized audio details can be: (1) embossed between perforations (separating the audio and video content) on the sound side, (2) stored as two redundant strips along the outside edges, or (3) stored on separate audio discs [94], [96].

During recording, endpoints (start and end) are identified for different shots. After this, a timecode sequence is generated at regular intervals and used to facilitate the desired synchronization between audio and video contents [97].

In today's industry-standard, sequence of numeric codes are generated at regular intervals. Specifically, in the SMPTE (society for motion picture and television engineering) timecode – provides a time reference for editing, synchronizing, and identifying film, video, and audio content [97], [98]. Unlike other timecodes, the SMPTE timecode records the absence and presence of a transition in the midpoints of the digital (sawtooth-shaped) waveform to indicate states ''0'' and ''1'', respectively. Intrinsically, the SMPTE timecode is a binary coded decimal (BCD) formulation that encodes information about the hour, minute, second, and frame identification for synchronized movie content [98]. A further 32 bits is allocated for additional user-specified tags [97], [98].

In concluding, we surmise that, while pioneering, SOD-based content editing was cumbersome, which contributed to

its unpopularity and the emergence of SOF systems as the preferred gateway to contemporary sound movies.

Furthermore, today's state-of-the-art audio-video synchronization systems employ dual-system sound (DSS) to sync many audio sources into separately recorded video content [95], [96]. While this adds a layer of complexity, its many upsides include providing consistency, and quality of digital content. On the downside, DSS systems impose additional demands related to simultaneous shoot and edit processes of movie production [96]. Additionally, it requires that audio recording and its synchronization be mapped to the different levels of movie abstraction, i.e. shots, scenes, etc.

B. QUANTUMNESS REQUIRED FOR TALKING COLOR QUANTUM MOVIES

In his now authoritative seminal article, DiVincenzo [57] proposed five broad requirements for physical implementation of quantum computation. Iliyasu's quantum movie framework (QMF) [38] (and its progenies – Yan's MCQI-based [39] and Wang's NEQR-based [40] chromatic frameworks for quantum movies (CFQMs)) are built on a generalization of these criteria as basis to hypothesize on composition of three hypothetical hardware the quantum movie disc (QMD), quantum movie player (QMP) and quantum movie reader (QMR) respectively to prepare, manipulate, and recover contents of quantum movies.

The stimulating perspective towards realizing talking quantum movies that is presented in this section assumes conformity with both established conjectures (i.e. by DiVincenzo [57] and Iliyasu *et al.* [38]). However, the assiduity inherent to the QMF provides evidence of its aggrandisement supported by applications thereof, whereas recent interest in quantum audio models (QAM) suggest it is also picking up; so, like in the persistent success story of digital sound movies (relayed in the preceding subsection), we will focus on requirements to coalesce the separate audio and video content into talking color quantum movies. Nevertheless, throughout, we will refer readers to the rudiments of the subject as well as the encouraging outcomes from published literature (that were also reviewed in earlier sections) in terms of both emerging quantum multimedia (i.e. audio and video) processing. All through, our discussions and excogitation will be predicated on the apparent quantumness required by future quantum movie hardware.

1) COHERENCE AND ENTANGLEMENT

At the heart of the "spooky action at a distance" as well as the "dead-and-aliveneness" that embody the quantumness of quantum systems is quantum coherence [67], [99]. It is a key component in various quantum information protocols as well as the phenomenal advantages offered by quantum tasks over traditional approaches. In their insightful study, Streltsov *et al.* [99] provided evidence suggesting "any degree of coherence with respect to some reference basis can be converted into entanglement via incoherent separations." This rigorously validated preposition forms the background of our ruminations regarding the confounding properties expected of our quantum movie framework (QMF) and quantum audio models (QAM) as well as their effective synchronization. Moreover, it has been established that achieving entanglement via coherence provides a contrivance to generate "maximum bipartite entanglement via incoherent operations and incoherent ancillary information [67], [99]."

Mathematically, given two bipartite subsystems X and Y, incoherent states can be formulated (as presented in Equation (84)) with respect to the reference product states [99].

$$\rho^{XY} = \sum_k p_k \sigma_k^X \otimes \tau_k^Y \tag{84}$$

Here, p_k are probabilities, while states σ_k^X, τ_k^Y are incoherent states on the subsystems X and Y respectively and, as noted in [100], both states are always separable.

Furthermore, for an arbitrary fixed reference basis ($|i\rangle$), the incoherent state can be defined as given in equation (85) [67], [99], [100].

$$\sigma = \sum_i p_i |i\rangle \langle i| \tag{85}$$

Any state that cannot be described in the form presented in Equation (84) is deemed coherent [67], [103] and, unlike in traditional information theory, quantum coherence is basis-dependent [101]–[103]. However, the choice of reference basis to recover the information (i.e. measurement) depends on the physical layout of the quantum system [99], [100].

Coherence ($C(\rho)$) could be deduced in terms of the following outcomes [99], [103].

$$\begin{aligned}
&1 : C(\rho) \geq 0 \quad \text{and } C(\sigma) = 0 \text{ iff} \sigma \in I \\
&2 : C(\rho) \geq (C\Lambda[\rho]) \quad \text{with } \Lambda[\rho] \subseteq I \\
&3 : C(\rho) \geq \sum_l p_l C(\zeta_l) \\
&4 : C(\sum_l p_l \rho_i) \leq \sum_l p_l C(\rho_i)
\end{aligned} \tag{86}$$

where in condition 3, $p_l = Tr\left[K_l \rho K_l^\dagger\right]$ states $\frac{K_l \rho K_l^\dagger}{p_l}$, incoherent Kraus operator K_l obeys $K_l I K_l^\dagger \subseteq I$ and, as noted in [99], [103], conditions 3 and 4 automatically imply condition 2. Overall, systems that satisfy the conditions listed in 1, 2 and 3 are deemed to be monotonically coherent [99], [103], [104].

Adopting the descriptions and conjectures presented in [99] and extending same to our earlier formulation of the movie register, we say that the coherence of an initial state ρ^m of a 2^m frames movie, M, can be converted to entanglement via coherent operations if, by attaching an ancilla sub-network, A, initialized in a reference incoherent state $|0\rangle \langle 0|^A$, the final system ancilla-state is:

$$S = \Lambda^{MA}\left[\rho^M \otimes |0\rangle \langle 0|^A\right], \tag{87}$$

which itself is entangled for some incoherent operation Λ^{MA} [103].

Fig. 51 (adapted from [99]) shows that "any nonzero coherence in the input state of a system S can be converted to entanglement via incoherent operations on S and an incoherent ancilla A [99]."

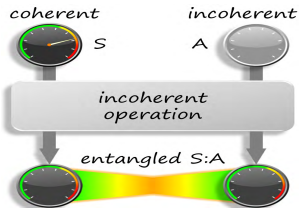

FIGURE 51. Generating entanglement via coherence (figure adapted from [99]).

Alternatively, entanglement can be generated by incoherent operations if the initial state ρ^m is coherent [99], [104]. Furthermore, based on the formulations in Equation (84)-(86) (and further descriptions in [99]), it was validated in [100] that "input coherence and output entanglement are quantitatively equivalent [103], [104]."

Readers should refer to the elegant characterization of entanglement via coherence in [99] as well as discussions in [100], [103] and [104] for further details of the coherent nature envisaged for future quantum movie technologies.

2) MIXED AND SUPERPOSITION STATES

Given a finite-dimensional Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ where $\mathcal{H}_1$ and $\mathcal{H}_2$ are smaller sub-spaces denoting a system and its environment respectively, then for a general $\Psi \in \mathcal{H}$, there is no $\Psi_1 \in \mathcal{H}_1$ that describes the outcome of all measurements in $\mathcal{H}_1$ [67], [105]. Instead, a more general phase-based concept for the state (than normalized elements) is required. Equation (88) formalizes the outcome when we expand Ψ in a basis [101].

$$\begin{aligned}\langle \Psi | o_1 \otimes I | \Psi \rangle &= \sum_{i_1,i_2,j_1,j_2} \overline{\psi}_{j_1,j_2} \psi_{i_1,i_2} \langle j_1 | o_1 | i_1 \rangle \langle j_1 | i_2 \rangle \\ &= \sum_{i_1,i_2,j_1,j_2} \overline{\psi}_{j_1,j_2} \psi_{i_1,i_2} \langle j_1 | o_1 | i_1 \rangle \\ &= tr\mathcal{H}_1 \gamma o_1 \end{aligned} \tag{88}$$

where o_1 is the interaction with $\mathcal{H}_1$, I is the identity operator and tr is the trace operator.

Expanding Ψ in a basis and $(|i_1\rangle)_{i_1 \in I_1}$ to be some orthonormal basis of $\mathcal{H}_1$ (and similarly $(|i_2\rangle)_{i_2 \in I_2}$ for $\mathcal{H}_2$), then:

$$\Psi = \sum_{i_1,i_2} \psi_{i_1,i_2} |i_1\rangle \otimes |i_2\rangle \tag{89}$$

γ in Equation (88) is a positive density matrix operator [105] of $\mathcal{H}_1$, which allows us to omit descriptions in terms of $\mathcal{H}_2$ and, can be formulated as

$$\begin{aligned}\gamma &= \sum_{i_1,i_2,j_1} \psi_{j_1,i_2} \psi_{i_1,j_2} |i_1\rangle \langle j_1| \\ &= tr\mathcal{H}_2(|\Psi\rangle \langle \Psi|) \end{aligned} \tag{90}$$

A density matrix generalizes a "pure" state (i.e. at rank one) and, by spectral theorem [105], a mixed state is described as a convex linear combination of orthogonal pure states [67], [102], [103], [105]. Therefore, given two orthonormal vectors $\psi_1, \psi_2 \in \mathcal{H}$, a mixed state γ_m describes a system with equal probability of being in ψ_1 and ψ_2.

In contrast, for a superposition state ψ_s, using an observable, $|\psi_1\rangle \langle \psi_2|$, measurements produce [104]:

$$\begin{aligned}\gamma_s &= \frac{1}{2}(|\psi_1\rangle + |\psi_2\rangle)(\langle \psi_1| + \langle \psi_2|) \\ &= \frac{1}{2}(|\psi_1\rangle \langle \psi_1| + |\psi_2\rangle \langle \psi_2| + |\psi_1\rangle \langle \psi_2| + |\psi_2\rangle \langle \psi_1|) \end{aligned} \tag{91}$$

As matrices in the basis $\{\psi_1, \psi_2\}$, γ_s and γ_m are defined in Equation (92).

$$\gamma_s = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and } \gamma_m = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tag{92}$$

which further reveals that γ_m has rank 2, while γ_s has rank 1. Using Hermitian unitary operators, such as the Pauli $\sigma_x = \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ matrix [67], we can compute the expectations from both systems.

The chronicle on the long road today's sophisticated digital audio-video synchronization as well as the overview on different approaches to enfetter two or more quantum systems provide useful premise for our similar aspirations on the quantum computing paradigm. These insights provide guidance for our explorations in synchronizing quantum movies and audio media which is presented in the next section.

IX. EXPLORATION ON QMF AND QAM SYNCHRONISATION

Having highlighted the roadmap of audio-video synchronization in present-day digital movies and identified a few quantum mechanical constraints that are essential for similar synchronization in future talking quantum movies (TQM) systems, here, we ruminate on the likely designs for future quantum movie technologies, particularly the requirements to synchronize quantum audio and video contents.

Based on the discussions on digital audio-video synchronization in the preceding section as well as the outlines of the quantum movie framework (QMF) and quantum audio models (QAM) (as reviewed in earlier sections of this treatise), we conject that synchronization of quantum video and audio content will require at least four phases or units (i.e. preparation, integration, synchronization, and measurement), which are allegorized in the remainder of this section. However, our discussion omits the separate procedures to prepare the two quantum media types (audio and video). These have been widely discussed in published literature, notably, those reviewed in earlier sections of this disquisition. Interested readers can consult [6], [16], [17], [18], [20], and [38] for further details.

1) INTEGRATION PHASE OF TQM LAYOUT

In this is the phase of the TQM production, the two separate quantum media (i.e. audio and video) are constricted or coalesced into a single quantum system. Depending on the intended application, this interaction could be used to steer the new system(s) as pure, mixed, superposed, or entangled states [7], [107]. Achieving this seems straightforward since (as explained in subsection VIII.B) each of the aforementioned states can be realized as long as the required coherence is generated and maintained. Furthermore, the integration envisaged could involve the ancillary interactions needed for ancilla-driven (ADQC) and measurement-based (MBQC) models of quantum computation as used in the quantum movie reader in [38] (reviewed in Section II).

2) SYNCHRONIZATION PHASE OF TQM LAYOUT

In this phase of our allegorical layout for TQM production, the hitherto integrated QMF and QAM contents are synced to provide necessary synergy in terms of temporal variability of both the audio and video contents. The time component of both the movie strip and FRQA formulations (refer to Equations (4) and (54)) provide readily available tools to support this symbiotic interaction.

3) TRANSFORMATIONS IN TQM LAYOUT

As is the norm in QIP, our proposed layout for talking color quantum movies (TQM) assumes that once initialized, prepared, interacted, and synced (as inferred in the preceding subsections), our system can only be manipulated via unitary transformations in form of quantum operations and gates. Consequently, the transformations on a TQM layout could be targeted at either or both media along the three core stages of quantum movie production, i.e. preparation, manipulation and recovery [38]. To recall, as recounted in Sections II through VI, these stages were the motivations the development of the QMD, QMP and QMR hypothetical quantum movie devices used by Iliyasu *et al.* in [38]. Similarly, in the TQM framework, at an instant t, either media, M, (with subscripts 'a' for audio and 'v' for movie) or both (indicated as 'b') can be transformed using the operation $T_t^{\mathbf{M},O}$ at the preparation (p), integration (i), synchronization (s), or measurement (m) phases. On the whole, highlighted in the sequel, a total of 11 classes of transformations can be executed via our envisaged TQM layout.

- Preparation - $T_0^{\boldsymbol{a},O}$, $T_0^{\mathbf{v},O}$.
  These are operations to prepare and initialize the quantum audio and quantum movie contents. As formulated in the respective frameworks for the QMF and QAM, and as mentioned earlier in this subsection, based on our explorative TQM layout, preparation-related transformations are the only operations that cannot be executed on both quantum media (audio and video) simultaneously. We note that in both transformations $t = 0$, which signifies that these operations do not directly impact the final movie produced via the TQM.
- Integration - $T_t^{\boldsymbol{a},i}$, $T_t^{\boldsymbol{v},i}$, $T_t^{\boldsymbol{b},i}$.
  At the integration phase of the TQM, either media could be transformed separately as needed to combine the separate media into a coherent quantum state (pure, mixed, entangled, etc.) which is tailored to support further processing at later stages of the TQM. Additionally, these transformations could target both media simultaneously. In addition to transformations focused at coalescing the two media types, other transformations to manipulate the contents of either or both media for movie security-related and general applications (such as the ones reviewed in Sections IV and V of this treatise) could be executed using these three classes of transformations.
- Synchronization - $T_t^{\boldsymbol{a},s}$, $T_t^{\boldsymbol{v},s}$, $T_t^{\boldsymbol{b},s}$.
  Like the transformations in the integration phase, transformations at the syncing phase of the TQM could focus on either media separately or both media simultaneously. Not only are these transformations focused on providing temporal synergy between the contents of the two media, but, depending on the application, they can also be used for other security-related and general applications focused on either media separately or both media simultaneously. For example, following necessary refinements, it is envisaged that many of the security protocols reviewed in earlier sections of this study could be extended to both media simultaneously, i.e. using the class $T_t^{\boldsymbol{b},s}$ class of TQM transformations.
- Measurement - $T_t^{\boldsymbol{a},m}$, $T_t^{\boldsymbol{v},m}$, $T_t^{\boldsymbol{b},m}$.
  The measurement operation entails use of some advanced quantum movie reader QMR (in this case a modified version of the QMR in [38] to accommodate the audio-video synchronization) to recover the final movie sequence that following necessary enhancements produces the movie available to the viewing audience. An interesting feature of the TQM layout is that, with appropriate tuning, the measurements could be restricted to either the audio or video contents of the movie.

4) ENHANCEMENT PHASE OF TQM LAYOUT

In digital movie production, numerous post-production editing takes place to enhance the content presented to the audience. These could be via editing suites to provide the rapid succession of minimally-differing sequential images in animated movies; the visual effects, stunts and choreography expected in action movies; the elicitation of physiological reactions to fear, terror and evil in horror movies; etc. Similarly, in future quantum movies, we envision the need for such phases to enhance the content prior to its screening to the audience. This is more so considering the perplexing and astounding properties, which accompany the contrariety that pervades information processing in the quantum and classical worlds. Consequently, there will be need for an enhancement stage within the TQM layout.

In concluding our allegorical forecast of the layout for future talking color quantum movies, we present in Fig. 52 a diagrammatic overview of the proposed layout. Essentially,

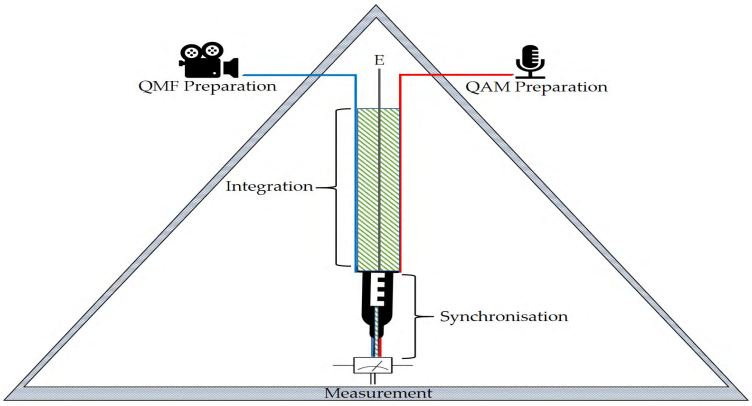

FIGURE 52. Layout for the realization of talking color quantum movies.

the layout combines the four phases outlined earlier into the needle-shaped framework outlining the incidental composition of future technologies to realize talking color quantum movies. The double-layered triangle housing the needle-shaped TQM is intended to depict the separation between the classical and quantum systems as well as the fortification that shields one from the other, which is a prerequisite for quantum computing.

X. OPEN ISSUES PERTAINING QUANTUM MOVIES, QUANTUM AUDIO, AND THEIR SYNCHRONIZATION BASED ON THE PROPOSED TQM

As outlined in the preceding section and in order to support the realization of other protocols (and applications), a few exploratory insights must be availed to showcase likely adjustments to either or both the proposed TQM and/or earlier QMF and QAM models as required to support effective synchronization between quantum audio and video media. In this section, we dwell on a few of these open areas for benefit of the QIP sub-discipline as well as to stimulate interest in the pursuit of QIP-specific technologies and hardware within the larger quantum computing community.

Based on the FRQA quantum audio model, it was mentioned in [45], that the possibility of using quantum-based signal processing operations, such as quantum Fourier transformations [29], [30] and wavelet transformations [33], [56], which have proven well-studied in other quantum information processing applications, should be explored to facilitate advanced manipulation of quantum audio content. Secondly, when the basic quantum audio signal operations are perfected, they could serve as the building block for the realization of more sophisticated algorithms used for specific applications such as quantum audio watermarking, steganography, and encryption. These improvements will facilitate multimedia security during quantum information transmission and communication [43], [107].

Similarly, as lamented in [43], ''constrained by the contrariety permeating quantum and classical computing as well as the inferior state of some quantum computing technologies and lack of physical hardware, the validation of Yan's CFQM movie model and its application in movie montages alongside more advanced digital methods was futile [43].'' Additionally, the impact of some quantum mechanical properties, notably decoherence and error correction, on different stages of the proposed CFQM framework were overlooked for exigent reasons. Both of these shortcomings need to revisited as sophistication of quantum movie technologies improve.

Moreover, it has been observed that ''even within present developments in the QIP sub-discipline, attempts to improve this study by exploiting some tools inherent to MCQI strip representation, which is at the heart of the CFQM framework, could help enhance available operations on the content of the movie.'' For example, the alpha-blending transformation [78] that incorporates the transparency operation could be used in some applications such as watermarking [21] as well as other geometric transformations like the rotation operation [61] to manipulate other movie contents. Additionally, the contrariety pervading the quantum and classical computing make comparisons between the advanced digital movie montages with their quantum versions a daunting undertaking. Hopefully, as quantum computing and QIP technologies mature to a level where resources to encode, manipulate, and play movie contents are available, comprehensive user and expert studies could be integrated into assessments on the performance of quantum movie montages [43].

In terms of our explorative TQM layout (in Section IX), similar to reserving a color framing bit to edit equipment content and confine editing to appropriate fields (in SMPTE timecoded SOF), the plausibility of executing such operations using the restricted variants of Le *et al.*'s color transformations on quantum images (CTQI) [62] or other advanced color operations [43] need to be investigated. Additionally, the tenability and selection of present day technologies that would suite each of the phases envisaged in the TQM layout needs to be assessed and, where required, appropriate modifications need to be adduced and validated.

In concluding, we reiterate that, considering the nascent stage of development of quantum movies, thus far, validation of quantum image processing (QIP)-based protocols has been confined to circuit-based simulataions using classical

computing resources. As argued in many QIP literature (such as [7]), the contrariety pervading the two computing worlds waters down the impact of the successes recorded. Consequently, quantum computing software with QIP dedicated extensions [107] are essential for the development of future quantum movies, more so, the talking color movies that have been the topic of this disquisition.

Earlier in Section II.B, it was mentioned that, a constant background, i.e. binary image is implied for applications of the quantum movie reader (QMR). Although such images are of particular interest and relatively straightforward to process, their applications are restricted. Consequently, to realize more advanced applications, it is important to review the layout of the QMR. Doing so might entail commensurate adjustments to the duo of the QMD and QMP that together make up the QMF in [38] as well as CFQMs in [39] and [40].

Moreover, the practicality of the quantum movie player (QMP) is still not clearly delivered. For example, it is important to clarify the type of interactions with the QMD needed to recover useful quantum information.

The few areas outlined above only abrade numerous shortcomings inherent to present formulations of the QMF as well as modifications needed to demonstrate movie production, including devices and tools to support encoding, manipulation, and recovery of quantum (movie) information and their sought-after synchronization with audio contents to produce talking quantum movies.

Finally, we invite readers to refer to the almost one dozen literature on quantum movie and audio models that formed the principal references that were the focus of this review for further directions regarding the growth of both emerging subtopics. Hopefully, the nominal effort to recount the progress recorded in terms of both the video and audio media types (as highlighted in this disquisition) as well as discussions on the need to reformulate both media types in tandem with developments in realizing physical quantum hardware (and, more recently, quantum software) would invigorate efforts along all directions mentioned.

XI. CONCLUDING REMARKS

Motivated by the long-held promise of quantum computation, the remarkable growth of the quantum image processing (QIP) sub-discipline as well as (recently) renewed interest supported with funding in/for the quantum computation and quantum information science disciplines, the review presented in this study enumerated the separate developments recorded in the quantum movie and, the relatively newer, quantum audio media types as emerging sub-fields of quantum image processing (QIP).

The disquisition presents an exhaustive overview of every literature (up to October 2018) relating the use of quantum computers for processing information encoded as both media (i.e. movies and audio).

In addition to highlighting the utility of the original quantum movie framework proposed by Iliyasu *et al.* [6], [38], in the first part of our treatise, we extended the discussion to cover more recent progenies that extend the QMF in terms MCQI-based [39] and NEQR-based [40] frameworks for color quantum movies (CFQM). Furthermore, recent efforts aimed at using the two CFQM models in accomplishing quantum movie security tasks as well as in protocols for quantum movie montages [43], moving target detection (MTD) [39] and frame motion detection [40] were also dissected.

The second part of the disquisition provided an extensive overview regarding the newer sub-field of quantum audio processing (QAP). In particular, detailed analysis of the only two quantum audio models – the QRDA [44] and FRQA [45], [46] – that are available as well as their applications in audio steganography [47] and quantum endpoint detection [48] were examined.

Finally, in the third part of the commentary, we undertook an explorative inquiry regarding the advances in QIP as well as the quantum movie frameworks and quantum audio models that could support synchronization of the quantum movie and quantum audio contents to realize talking color quantum movies (TQM). The creative layout envisaged for the realization of the TQM was in part motivated by developments in two hitherto unrelated areas. First, drawing parables with the emergence and growth of contemporary (digital) sound movies, we opine that the techniques similar to those used in digital sound-on-frame (SOF) sound movies would be most fitted for coalescing the quantum audio and video models for the purpose of realizing talking color quantum movies. Second, we appraised the core properties of coherence, entanglement, mixed/pure states, and superposition that are at the heart of the quantumness inherent to quantum computing hardware to construe the likely composition of future quantum movie technologies which could support the preparation, integration, synchronization, and measurements phases of any future TQM layout.

The proposed needle-shaped, multi-stage TQM layout (in Fig. 52) consists of eleven categories of unitary transformation groups that could be tailored towards manipulating the contents of quantum video or quantum audio separately or both media types simultaneously along different phases of the TQM.

Among others, the main objectives of the operose efforts invested in this review are twofold. First, targeting enthusiasts and beginners in the quantum computing field, the treatise provides both an essential timeline chronicling advances in QIP sub-discipline as well as an exhaustive overview of all literature related to quantum movies and/or videos. Second, for the benefit of those already pursuing research in or related to the QIP sub-discipline, the commentary provides useful insights, allegories, and conjectures to support the integration of other media such quantum audio, etc. for the realization of colored talking quantum movies (or video).

As emphasized throughout the study, for the most part, the review is intended to serve as an avenue to invigorate activities tailored towards accelerating the realization of QIP-based state-of-the-art technologies.

LIST OF ABBREVIATIONS

IEEE	institute of electrical and electronics engineers
QIP	quantum image processing
IBM	international business machine
MIT	Massachusetts institute of technology
NASA	national space agency
USRA	united states space research association
FRQI	flexible representation for quantum images
OQI	optical quantum imaging
DIP	digital image processing
QDIP	quantum-assisted digital image processing
QIR	quantum image representation
MCQI	multichannel representation for quantum images
NEQR	novel quantum representation for digital audio
QMF	quantum movie framework
CFQM	chromatic framework for quantum movies
QRDA	quantum representation for digital audio
QASP	quantum audio signal processing
FRQA	flexible representation for quantum audio
QAS	quantum audio steganography
QEPD	quantum endpoint detection
QMD	quantum movie disc
QMP	quantum movie player
QMR	quantum movie reader
CPU	central processing unit
VCD	video cassette disc
DVD	digital video decorder
VCR	video cassette recorder
ROI	region of interest
SMO	simple motion operation
GTQI	geometric transformation of quntum images
CTQI	color transformations on quantum images
MBQC	measurement-based quantum computing
ADQC	ancilla-driven quantum computing
$CFQM_M$	MCQI-based chromatic framework for quantum movies
$CFQM_N$	NEQR-based chromatic framework for quantum movies
CIK	color information key
CTO	color transformation operation
LSB	least significant bit
MTD	motion target detection
CT_X	color transformation on channel X
FTF	frame to frame transition
COI	channel of interest
SBS	sub-block swapping
COMP	comparator
ΔB	residual information
q-ADD	quantum adder
DPCM	differential pulse code modulation
MBE	minimization of Boolean expression
QIC	quantum image compression
QAC	quantum audio compression
DAP	digital audio signal processing
MSQb	most significant qubit
EXT	extension
LSQb	least significant qubit
QAS	quantum audio steganography
PEQ	picked embedding qubit
cLSQ	LSQb-based stego protocol
pMSQ	MSQb-based stego protocol
BQ	beginning qubit
DQ	destination qubit
ASR	automatic speech recognition
S2T	speech to text
HCI	human-computer interaction
EPD	endpoint detection (digital)
qEPD	quantum endpoint detection
E_n	short term energy
*n*COMP	*n*-bit comparator
SOD	sound-on-disc
SOF	sound-on-film
RCA	radio corporation of America
SVA	stereo variable area
SMPTE	society for motion picture and television engineering
DSS	dual system sound
QAM	quantum audio model
TQM	talking color quantum movie

ACKNOWLEDGMENT

The author is grateful to Mr Abubakar M. Iliyasu for his effort in typing the initial draft of this paper. Similarly, the author appreciates the assistance of authors and publishers of the articles covered in the review, many of whom not only granted permission to reproduce some of their figures but they were kind enough to help with top-quality versions of such images.

REFERENCES

[1] *IBM Shows Off a Quantum Computing Chip, Via*. Accessed: Dec. 31, 2017. [Online]. Available: http://www.technologyreview.com
[2] *Quantum Computing*. Accessed: Dec. 31, 2017. [Online]. Available: http://www.nas.nasa.gov
[3] *Scalable, End-to-End Technology*. Accessed: Dec. 31, 2017. [Online]. Available: http://www.microfoft.com
[4] *Is China the Leader in Quantum Communications?* Accessed: Dec. 31, 2017.v [Online]. Available: http://www.insidescience.org
[5] R. Landauer, "Information is physical," *Phys. Today*, vol. 44, no. 5, p. 23, 1991.
[6] A. M. Iliyasu, "Towards realising secure and efficient image and video processing applications on quantum computers," *Entropy*, vol. 15, no. 8, pp. 2874–2974, 2013.
[7] A. M. Iliyasu, F. Yan, and K. Hirota, "Metric for estimating congruity between quantum images," *Entropy*, vol. 18, no. 10, p. 360, 2015.
[8] A. Y. Vlasov. (1997). "Quantum computations and images recognition." [Online]. Available: https://arxiv.org/abs/quant-ph/9703010

[9] M. Rastegari. *Quantum Approach to Image Processing*. Accessed: Dec. 31, 2017. [Online]. Available: http://www.umiacs.umd.edu/~mrastega/paper/final3.pdf
[10] R. Schützhold, "Pattern recognition on a quantum computer," *Phys. Rev. A, Gen. Phys.*, vol. 67, no. 6, p. 062311, 2002.
[11] Y. C. Eldar and A. V. Oppenheim, "Quantum signal processing," *IEEE Signal Process. Mag.*, vol. 19, no. 6, pp. 12–32, Nov. 2002.
[12] L. A. Lugiato, A. Gatti, and E. Brambilla, "Quantum imaging," *J. Opt. B, Quantum Semiclassical Opt.*, vol. 4, no. 3, pp. S176–S183, 2002.
[13] G. Beach, C. Lomont, and C. Cohen, "Quantum image processing," in *Proc. 32nd IEEE Conf. Appl. Imagery Pattern Recognit.*, Bellingham, WA, USA, Jul. 2003, pp. 39–44.
[14] S. E. Venegas-Andraca and S. Bose, "Storing, processing, and retrieving an image using quantum mechanics," *Proc. SPIE*, vol. 5105, pp. 137–148, Aug. 2003.
[15] J. I. Latorre. *Image Compression and Entanglement*. Accessed: Dec. 31, 2017. [Online]. Available: http://arxiv.org/abs/quantph/0510031
[16] P. Q. Le, F. Dong, and K. Hirota, "A flexible representation of quantum images for polynomial preparation, image compression, and processing operations," *Quantum Inf. Process.*, vol. 10, no. 1, pp. 63–84, 2012.
[17] P. Q. Le, A. M. Iliyasu, F. Dong, and K. Hirota, "A flexible representation and invertible transformations for images on quantum computers," in *New Advances in Intelligent Signal Processing* (Studies in Computational Intelligence), vol. 372, A. E. Ruano and A. R. Varkonyi-Koczy, Eds. Berlin, Germany: Springer, 2011, pp. 179–202.
[18] F. Yan, A. M. Iliyasu, and S. E. Venegas-Andraca, "A survey of quantum image representations," *Quantum Inf. Process.*, vol. 15, no. 1, pp. 1–35, 2016.
[19] B. Sun, A. M. Iliyasu, F. Yan, F. Dong, K. Hirota, "An RGB multi-channel representation for images on quantum computers," *J. Adv. Comput. Intell. Intell. Inform.*, vol. 17, no. 3, pp. 404–417, 2013.
[20] Y. Zhang, K. Lu, Y. Gao, and M. Wang, "NEQR: A novel enhanced quantum representation of digital images," *Quantum Inf. Process.*, vol. 12, no. 8, pp. 2833–2860, 2013.
[21] A. M. Iliyasu, P. Q. Le, F. Dong, and K. Hirota, "Watermarking and authentication of quantum images based on restricted geometric transformations," *Inf. Sci.*, vol. 186, no. 1, pp. 126–149, 2012.
[22] W.-W. Zhang, F. Gao, B. Liu, Q.-Y. Wen, and H. Chen, "A watermark strategy for quantum images based on quantum Fourier transform," *Quantum Inf. Process.*, vol. 12, no. 2, pp. 793–803, 2013.
[23] A. M. Iliyasu *et al.*, "A two-tier scheme for greyscale quantum image watermarking and recovery," *Int. J. Innov. Comput. Appl.*, vol. 5, no. 2, pp. 85–101, 2013.
[24] F. Yan, A. M. Iliyasu, B. Sun, S. E. Venegas-Andraca, F. Dong, and K. Hirota, "A duple watermarking strategy for multi-channel quantum images," *Quantum Inf. Process.*, vol. 14, no. 5, pp. 1675–1692, 2015.
[25] A. A. A. El-Latif, B. Abd-El-Atty, M. S. Hossain, M. A. Rahman, A. Alamri, and B. B. Gupta, "Efficient quantum information hiding for remote medical image sharing," *IEEE Access*, vol. 6, pp. 21075–21083, 2018.
[26] N. Jiang and L. Wang, "A novel strategy for quantum image steganography based on Moiré pattern," *Int. J. Theor. Phys.*, vol. 54, no. 3, pp. 1021–1032, 2015.
[27] N. Jiang, N. Zhao, and L. Wang, "LSB based quantum image steganography algorithm," *Int. J. Theor. Phys.*, vol. 55, no. 1, pp. 107–123, 2016.
[28] R.-G. Zhou, Q. Wu, M.-Q. Zhang, and C.-Y. Shen, "Quantum image encryption and decryption algorithms based on quantum image geometric transformations," *Int. J. Theor. Phys.*, vol. 52, no. 6, pp. 1802–1817, 2013.
[29] Y.-G. Yang, J. Xia, X. Jia, and H. Zhang, "Novel image encryption/decryption based on quantum Fourier transform and double phase encoding," *Quantum Inf. Process.*, vol. 12, no. 11, pp. 3477–3493, 2013.
[30] Y.-G. Yang, X. Jia, S.-J. Sun, and Q.-X. Pan, "Quantum cryptographic algorithm for color images using quantum Fourier transform and double random-phase encoding," *Inf. Sci.*, vol. 277, pp. 445–457, Sep. 2014.
[31] F. Yan *et al.*, "Quantum image searching based on probability distributions," *J. Quantum Inf. Sci.*, vol. 2, no. 3, pp. 55–60, 2012.
[32] F. Yan, A. M. Iliyasu, P. Q. Le, B. Sun, F. Dong, and K. Hirota, "A parallel comparison of multiple pairs of images on quantum computers," *Int. J. Innov. Comput. Appl.*, vol. 5, no. 4, pp. 199–212, 2013.
[33] H. S. Li, Z. Qingxin, S. Lan, C.-Y. Shen, R. Zhou, J. Mo, "Image storage, retrieval, compression and segmentation in a quantum system," *Quantum Inf. Process.* vol. 12, no. 6, pp. 2269–2290, 2013.
[34] F. Yan, A. M. Iliyasu, and P. Q. Le, "Quantum image processing: A review of advances in its security technologies," *Int. J. Quantum Inf.*, vol. 15, no. 3, p. 1730001, 2017.
[35] N. Jiang, L. Wang, and W.-Y. Wu, "Quantum Hilbert image scrambling," *Int. J. Theor. Phys.*, vol. 53, no. 7, pp. 2463–2484, 2014.
[36] F. Yan, Y Guo, A. M. Iliyasu, Z. Jiang, and H. Yang, "Multi-channel quantum image scrambling," *J. Adv. Comput. Intell. Intell. Inf.*, vol. 20, no. 1, pp. 163–170, 2016.
[37] F. Yan, A. M. Iliyasu, and Z. Jiang, "Quantum computation-based image representation, processing operations and their applications," *Entropy*, vol. 16, no. 10, pp. 5290–5338, 2014.
[38] A. M. Iliyasu, P. Q. Le, F. Dong, and K. Hirota, "A framework for representing and producing movies on quantum computers," *Int. J. Quantum Inf.*, vol. 9, no. 6, pp. 1459–1497, 2011.
[39] F. Yan, A. M. Iliyasu, A. R. Khan, and Z. Jiang, "Measurements-based moving target detection in quantum video," *Int. J. Theor. Phys.* vol. 55, no. 1, pp. 2162–2173, 2016.
[40] S. Wang, "Frames motion detection of quantum video," in *Advances in Intelligent Information Hiding and Multimedia Signal Processing* (Smart Innovation, Systems and Technologies), vol. 64. J.-S. Pan, Eds. 2016, doi: 10.1007/978-3-319-50212-0_18.
[41] F. Yan, A. M. Iliyasu, S. E. Venegas-Andraca, and H. Yang, "Video encryption and decryption on quantum computers," *Int. J. Theor. Phys.*, vol. 54, no. 8, pp. 2893–2904, 2015.
[42] Z. Qu, S. Chen, and S. Ji, "A novel quantum video steganography protocol with large payload based on MCQI quantum video," *Int. J. Theor. Phys.* vol. 56, no. 11, pp. 3543–3561, 2017.
[43] F. Yan, S. Jiao, A. M. Iliyasu, and Z. Jiang, "Chromatic framework for quantum movies and applications in creating montages," *Frontiers Comput. Sci.*, vol. 12, no. 4, pp. 736–748, 2018, doi: 10.1007/s11704-018-7070-8.
[44] J. Wang, "QRDA: Quantum representation of digital audio," *Int. J. Theor. Phys.*, vol. 55, no. 3, pp. 1622–1641, 2016.
[45] F. Yan, Y. Guo, A. M. Iliyasu, and H. Yang. (2017). "Flexible representation and manipulation of audio signals on quantum computers." [Online]. Available: https://arxiv.org/abs/1701.01291
[46] F. Yan, A. M. Iliyasu, Y. Guo, and H. Yang, "Flexible representation and manipulation of audio signals on quantum computers," *Theor. Comput. Sci.*, vol. 752, pp. 71–85, 2018. doi: 10.1016/j.tcs.2017.12.025.
[47] K. Chen, F. Yan, A. M. Iliyasu, and J. Zhao, "Exploring the implementation of steganography protocols on quantum audio signals," *Int. J. Theor. Phys.*, vol. 57, no. 2, pp. 476–494, 2017.
[48] J. Wang, H. Wang, and Y. Song, "Quantum endpoint detection based on QRDA," *Int. J. Theor. Phys.*, vol. 56, no. 10, pp. 3257–3270, 2017.
[49] D. Horn and A. Gottlieb, "Algorithm for data clustering in pattern recognition problems based on quantum mechanics," *Phys. Rev. Lett.*, vol. 88, no. 1, p. 018702, 2001.
[50] N. B. Karayiannis, R. Kretzschmar, H. Richner, Pattern classification based on quantum neural networks: A case study," in *Pattern Recognition, From Classical to Modern Approaches*, S. K. Pal, Eds. Singapore: World Scientific, 2001, pp. 301–328.
[51] S. Meshoul and M. Batouche, "A novel quantum behaved particle swarm optimization algorithm with chaotic search for image alignment," in *Proc. IEEE Congr. Evol. Comput.*, Jul. 2010, pp. 1–6.
[52] A. Manju and M. J. Nigam, "Applications of quantum inspired computational intelligence: A survey," *Artif. Intell. Rev.*, vol. 42, no. 1, pp. 79–156, 2014.
[53] G. B. Lemos, V. Borish, G. D. Cole, S. Ramelow, R. Lapkiewicz, and A. Zeilinger, "Quantum imaging with undetected photons," *Nature*, vol. 512, pp. 409–412, Aug. 2014.
[54] P.-A. Moreau, E. Toninelli, T. Gregory, and M. J. Padgett, "Ghost imaging using optical correlations," *Laser Photon. Rev.*, vol. 12, no. 1, p. 1700143, 2017.
[55] A. M. Iliyasu and C. Fatichah, "A quantum hybrid PSO combined with fuzzy k-NN approach to feature selection and cell classification in cervical cancer detection," *Sensors*, vol. 17, no. 12, p. 2935, 2017.
[56] X. W. Yao *et al.*, "Quantum image processing and its application to edge detection: Theory and experiment," *Phys. Rev. X*, vol. 7, no. 3, p. 031041, 2017.
[57] D. P. DiVincenzo, "The physical implementation of quantum computation," *Fortschritte Phys.*, vol. 48, no. 9, p. 771, 2000.
[58] A. Hampapur, T. Weymouth, and R. Jain, "Digital video segmentation," in *Proc. ACM Multimedia Conf.*, 1994, pp. 357–364.
[59] C. H. Bennett and D. P. DiVincenzo, "Quantum information and computation," *Nature*, vol. 404, p. 247, Mar. 2000.
[60] P. Q. Le, A. M. Iliyasu, F. Dong, and K. Hirota, "Fast geometric transformations on quantum images," *Int. J. Appl. Math.*, vol. 40, no. 3, p. 113, 2010.

[61] P. Q. Le, A. M. Iliyasu, F. Dong, and K. Hirota, "Strategies for designing geometric transformations on quantum images," *Theor. Comput. Sci.*, vol. 412, p. 1406–1418, Mar. 2011.
[62] P. Q. Le, A. M. Iliyasu, F. Dong, and K. Hirota, "Efficient color transformations on quantum images," *J. Adv. Comput. Intell. Intell. Inform.*, vol. 15, no. 6, pp. 698–706, 2011.
[63] J. Anders, D. K. L. Oi, E. Kashefi, D. E. Browne, and E. Andersson, "Ancilla-driven universal quantum computation," *Phys. Rev. A, Gen. Phys.*, vol. 82, no. 2, p. 020301, 2010.
[64] T. B. Pittman, *Quantum Computing With Linear Optics*. Accessed: Aug. 31, 2018. [Online]. Available: https//arxiv.org/abs/quant-ph/0406192
[65] D. Maslov, G. W. Dueck, D. M. Miller, and N. Camille, "Quantum circuit simplification and level compaction," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 27, no. 3, pp. 436–444, Mar. 2008.
[66] J. Anders, E. Andersson, D. E. Browne, E. Kashefi, and D. K. L. Oi, "Ancilla-driven quantum computation with twisted graph states," *Theor. Comput. Sci.*, vol. 430, pp. 51–72, Apr. 2012.
[67] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
[68] S. Caraiman and V. I. Manta, "New applications of quantum algorithms to computer graphics: The quantum random sample consensus algorithm," in *Proc. 6th ACM Conf. Comput. Frontier*, May 2009, pp. 81–88.
[69] A. Barenco *et al.*, "Elementary gates for quantum computation," *Phys. Rev. A, Gen. Phys.*, vol. 52, p. 3457, Nov. 1995.
[70] A. M. Childs, D. W. Leung, and M. A. Nielsen, "Unified derivations of measurement-based schemes for quantum computation," *Phys. Rev. A, Gen. Phys.*, vol. 71, no. 3, p. 032318, 2005.
[71] E. Kashefi, D. K. L. Oi, D. E. Browne, J. Anders, and E. Andersson. (2009). "Twisted graph states for ancilla-driven quantum computation." [Online]. Available: https://arxiv.org/abs/0905.3354v2
[72] S. E. Venegas-Andraca and J. L. Ball, "Processing images in entangled quantum systems," *Quantum Inf. Process.*, vol. 9, no. 1, pp. 1–11, 2010.
[73] P. Li, X. Liu, and H. Xiao, "Quantum image median filtering in the spatial domain," *Quantum Inf. Process.*, vol. 17, no. 3, p. 49, 2018.
[74] B. Bhargava, C. Shi, and S.-Y. Wang, "MPEG video encryption algorithms," *Multimedia Tools Appl.*, vol. 24, no. 1, pp. 57–79, 2004.
[75] S. Wang, J. Sang, X. Song, and X. Niu, "Least significant qubit (LSQb) information hiding algorithm for quantum image," *Measurement*, vol. 73, pp. 352–359, Sep. 2015.
[76] N. Jiang, N. Zhao, and L. Wang, "LSB-based quantum image steganography algorithm," *Int. J. Theor. Phys.*, vol. 55, no. 1, pp. 107–123, 2016.
[77] B. Sun *et al.*, "A multi-channel representation for images on quantum computers using the RGBα color space," in *Proc. IEEE 7th Int. Symp. Intell. Signal Process. (WISP)*, Sep. 2011, pp. 1–6.
[78] B. Sun *et al.*, "Multi-channel information operations on quantum images," *J. Adv. Comput. Intell. Intell. Inform.*, vol. 18, no. 2, pp. 140–149, 2014.
[79] *Recorded Music*. Accessed: Aug. 31, 2018. [Online]. Available: http://en.wikipedia.org/wiki/Audio mixing
[80] C. C. Cutler, "Differential quantization of communication signals," U.S. Patent 2 605 361, Jul. 29, 1952.
[81] M. Nosrati, R. Karimi, and M. Hariri, "Audio steganography: A survey on recent approaches," *World Appl. Program.*, vol. 2, no. 3, pp. 202–205, 2012.
[82] N. Cvejic and T. Seppanen, "Increasing robustness of LSB audio steganography using a novel embedding method," in *Proc. Int. Conf. Inf. Technol., Coding Comput.*, Apr. 2004, pp. 533–537.
[83] J. G. Wilpon, L. R. Rabiner, and T. Martin, "An improved word-detection algorithm for telephone-quality speech incorporating both syntactic and semantic constraints," *AT&T Bell Lab. Tech. J.*, vol. 63, no. 3, pp. 479–498, 1984.
[84] R. Altman, *Sound Theory/Sound Practice*, 2nd ed. Evanston, IL, USA: Routledge, 2015, p. 298.
[85] J.-C. Junqua, B. Reaves, and B. Mak, "A study of endpoint detection algorithms in adverse conditions: Incidence on a DTW and HMM recognizer," in *Proc. Eurospeech*, 1991, pp. 1371–1374.
[86] R. Altman, *Silent Film Sound*, vol. 40. New York, NY, USA: Columbia Univ. Press, 2004, p. 155.
[87] *Audio Engineering History*. Accessed: Aug. 31, 2018. [Online]. Available: http//www.aes.org/aeshc/#AudEngHist
[88] M. Leonard. *The Vitaphone Project Turns 20*. Accessed: Aug. 31, 2018. [Online]. Available: https//indiewire.com/2010/08/lets-hear-it-for-vitaphone-179252
[89] M. Peter. A Resounding 25 Years of Reviving Early Film. Moving Image Archive News. Accessed: Aug. 31, 2018. [Online]. Available: https//imovingimagearchivenews.org/a-resoundingr-25-years-of-reviving-early-film
[90] F. C. Barton, "Victrolac motion picture records," *J. Soc. Motion Picture Eng.*, vol. 18, no. 4, pp. 452–460, 1932. Accessed: Aug. 31, 2018.
[91] F. W. Hoffman, *Encyclopedia of Recorded Sound*. London, U.K.: Taylor & Francis, 2004.
[92] E. W. Kellog, "History of sound motion pictures," *J. Soc. Motion Picture Eng.*, vol. 64, no. 4, pp. 356–374, 1955. Accessed: Aug. 31, 2018.
[93] *Lee de Forest and Phonofilm: Virtual Broadway*. Accessed: Aug. 31, 2018. [Online]. Available: http://www.encylopedia.jrank.org
[94] A. Frayne and H. Wolfe, *Sound Recording*. New York, NY, USA: Wiley, 1949.
[95] C. Poynton, *A Technical Introduction to Digital Video*. New York, NY, USA: Wiley, 1996.
[96] R. E. Uhlig, "Stereophonic photographic soundtracks," *J. Soc. Motion Picture Eng.*, vol. 82, no. 4, pp. 292–295, 1973.
[97] *A Brief History of SMPTE Time Code*. Accessed: Aug. 31, 2018. [Online]. Available: www.horita.com
[98] J. Ratcliff, *Timecode: A User'S Guide*. Evanston, IL, USA: Routledge, 1999.
[99] A. Streltsov, U. Singh, H. S. Dhar, M. N. Bera, and G. Adesso, "Measuring quantum coherence with entanglement," *Phys. Rev. Lett.*, vol. 115, no. 2, p. 020403, 2015.
[100] W. K. Wootters, "Entanglement of formation of an arbitrary state of two qubits," *Phys. Rev. Lett.*, vol. 80, no. 10, p. 2245, 1998.
[101] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," *Nature Photon.*, vol. 5, no. 4, p. 222, 2011.
[102] S. F. Huelga and M. B. Plenio, "Vibrations, quanta and biology," *Contemp. Phys.*, vol. 54, no. 4, pp. 181–207, 2013.
[103] T. Baumgratz, M. Cramer, and M. B. Plenio, "Quantifying coherence," *Phys. Rev. Lett.*, vol. 113, no. 14, p. 140401, 2014.
[104] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Rev. Mod. Phys.*, vol. 81, no. 2, p. 865, 2009.
[105] R. Helling. *The Physics of Density Matrices*. Accessed: Aug. 31, 2018. [Online]. Available: https://www.mathematik.uni-muenchen.de
[106] *UK Quantum Technology Landscape 2016*. Accessed: Aug. 31, 2018. [Online]. Available: http://uknqt.epsrc.ac.uk/files/ukquantumtechnologylandscape2016
[107] H. V. Nguyen *et al.*, "Towards the quantum Internet: Generalised quantum network coding for large-scale quantum communication networks," *IEEE Access*, vol. 5, pp. 17288–17308, 2017.
[108] S. Caraiman and V. Manta, "Image processing using quantum computing," in *Proc. 16th Int. Conf. Syst. Theory, Control Comput. (ICSTCC)*, Oct. 2012, pp. 1–6.
[109] Y. Zhang, K. Lu, Y. Gao, and K. Xu, "A novel quantum representation for log-polar images," *Quantum Inf. Process.*, vol. 12, no. 9, pp. 3103–3126, 2013.
[110] S. Yuan, X. Mao, Y. Xue, L. Chen, Q. Xiong, and A. Compare, "SQR: A simple quantum representation of infrared images," *Quantum Inf. Process.*, vol. 13, no. 6, pp. 1353–1379, 2014.
[111] H.-S. Li, Q. Zhu, R. G. Zhou, L. Song, and X. J. Yang, "Multi-dimensional color image storage and retrieval for a normal arbitrary quantum superposition state," *Quantum Inf. Process.*, vol. 13, no. 4, pp. 991–1011, 2014.

ABDULLAH M. ILIYASU (ABDUL M. ELIAS) received the M.E., Ph.D., and Dr.Eng. degrees in computational intelligence and intelligent systems engineering from the Tokyo Institute of Technology, Japan, where he is presently a Research Faculty with the School of Computing. Concurrently, he is the Principal Investigator and Team Leader of the Computational Intelligence and Intelligent Systems (CIIS) Research Group at the College of Engineering, Prince Sattam Bin Abdulaziz University in the Kingdom of Saudi Arabia. He is also a Professor with the School of Computer Science and Engineering, Changchun University of Science and Technology, China. He has to his credit more than 100 publications traversing the areas of computational intelligence, quantum cybernetics, quantum image processing, quantum machine learning, cyber and information security, hybrid intelligent systems, health informatics, and electronics systems reliability. He can be reached via a.iliyasu@psau.edu.sa.