Received: 28 April 2022 | Revised: 13 September 2022 | Accepted: 20 September 2022

IET Quantum Communication

DOI: 10.1049/qtc2.12050

ORIGINAL RESEARCH

Second quantised information distance

Songsong Dai

School of Electronics and Information Engineering, Taizhou University, Taizhou, Zhejiang, China

Correspondence

Songsong Dai, School of Electronics and Information Engineering, Taizhou University, Taizhou 318000, Zhejiang, China.
Email: ssdai@tzc.edu.cn

Funding information

National Natural Science Foundation of China, Grant/Award Number: 62006168; Natural Science Foundation of Zhejiang Province, Grant/Award Number: LQ21A010001

Abstract

The Kolmogorov complexity of a string is the minimum length of a programme that can produce that string. Information distance between two strings based on Kolmogorov complexity is defined as the minimum length of a programme that can transform either string into the other one, both ways. The second quantised Kolmogorov complexity of a quantum state is the minimum average length of a quantum programme that can reproduce that state. In this paper, a second quantised information distance is defined based on the second quantised Kolmogorov complexity. It is described as the minimum average length of a transformation quantum programme between two quantum states. This distance's basic properties are discussed. A practical analogue of quantum information distance is also developed based on quantum data compression.

KEYWORDS

indeterminate length quantum string, Kolmogorov complexity, quantum data compression, quantum information distance, quantum Kolmogorov complexity, second quantised Kolmogorov complexity

1 | INTRODUCTION

In 1998, Bennett et al. [1] introduced a universal information distance based on Kolmogorov complexity [2–4]. Kolmogorov complexity, $K(x)$, of a string x is defined as the length of a shortest programme that can produce x. Then information distance, $d(x, y)$, of two strings x and y is defined as the length of a shortest programme that can transform x into y and vice versa.

Information distance is a universal measure of the difficulty of the transformation between two strings. A similar question is given two quantum states, how difficult to transform either state into the other one, both ways. Quantum computer provides more accurate description of quantum objects than classical computer. Therefore, we are interested in finding a universal measure of the difficulty of the transformation on a universal quantum computer.

There are many different definitions of the shortest description of quantum states, that is, quantum Kolmogorov complexities. Vitányi [5] defined the quantum Kolmogorov complexity of a state as the length of its shortest classical description. Berthiaume et al. [6] defined the quantum Kolmogorov complexity of a state as the length of its shortest quantum description. Building upon their quantum Kolmogorov complexities [5, 6], Dai [7, 8] developed two quantum information distances, which are, respectively, based on classical and quantum descriptions. However, some quantum states have indeterminate lengths because of their superposition. Rogers and Vedral [9, 10] defined the second quantised Kolmogorov complexity of a state as the minimum average length of its quantum description on a universal quantum computer. The physical interpretation is described as the minimum energy required to describe that state [9, 10]. In this paper, a second quantised information distance based on second quantised Kolmogorov complexity is defined. It is described as the minimum average length of a shortest quantum programme that can transform either state into the other one, both ways. Our research framework is shown in Figure 1.

The rest of the paper is organised as follows: In Section 2, we review classical Kolmogorov complexity and information distance. Then we review the existing definitions of quantum information distances in Section 3. In Section 4, we recall the concepts of indeterminate and determinate length states and second quantised Kolmogorov complexity. In Section 5, we define the second quantised information distance and discuss its basic properties, its physical interpretation and its relationship to previous quantum information distance. In Section 6, we develop a practical analogue of quantum information

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *IET Quantum Communication* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

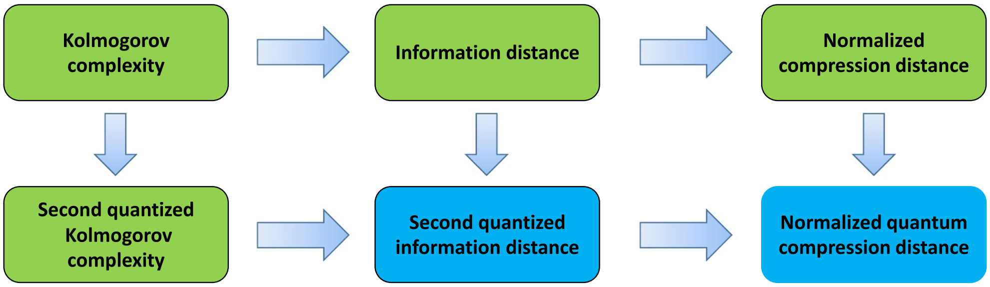

FIGURE 1 A summary on the relative concepts

distance based on lossless quantum data compression. Discussions and conclusions are briefly summed up in Section 7.

2 | FROM KOLMOGOROV COMPLEXITY TO NORMALISED COMPRESSION DISTANCE

We briefly detail here some concepts of Kolmogorov complexity and its derivations. For more details, see Li and Vitányi's book [11].

Let T be a universal Turing machine; the **Kolmogorov complexity** [2–4], $K(x)$, of a string x with respect to T is the length of a shortest programme to compute x, that is,

$$K_T(x) = \min\{l(p) : T(p) = x\}. \tag{1}$$

where $l(p)$ is the number of bits in the programme p.

Furthermore, **conditional Kolmogorov complexity**, $K(x|y)$, is the length of the shortest programme to compute x given y, that is,

$$K_T(x|y) = \min\{l(p) : T(p, y) = x\}. \tag{2}$$

Building upon conditional Kolmogorov complexity, **information distance** [1] of two strings x and y is the length of the shortest programme that generates x from y as well as generating y from x, that is,

$$d_T(x, y) = \min\{l(p) : T(p, y) = x \text{ and } T(p, x) = y\}. \tag{3}$$

An essential feature of these concepts is that they do not matter which machine is used to define them. They are invariant, that is, for any Turing machine Q, there exists a constant c_Q (dependent only on Q) such that

$$K_T(x) \le K_Q(x) + c_Q, \tag{4}$$

$$K_T(x|y) \le K_Q(x|y) + c_Q, \tag{5}$$

$$d_T(x, y) \le d_Q(x, y) + c_Q. \tag{6}$$

for any strings x and y.

Moreover, Li et al. [12] defined the **normalised information distance** as

$$e_T(x, y) = \frac{\max\{K_T(x|y), K_T(y|x)\}}{\max\{K_T(y), K_T(y)\}}. \tag{7}$$

However, Kolmogorov complexity is not computable. Therefore, Li et al. [12] developed a practical analogue of the normalised information distance based on real-world compressors, called the **normalised compression distance**, that is,

$$NCD(x, y) = \frac{C(x, y) - \min\{C(x), C(y)\}}{\max\{C(x), C(y)\}}, \tag{8}$$

where function $C(x)$ is the size of the compression of a string x.

After that, the normalised compression distance has successfully been applied to bioinformatics [13, 14], music clustering [15, 16], language trees [12], plagiarism detection [17] and question answering [18].

3 | PREVIOUS WORK

In this section, some of the previous work on quantum Kolmogorov complexity and quantum information distance are described. First, we recall the concept of indeterminate and determinate length quantum states, which have been discussed in the lossless quantum data compression [19, 20]. Next, we review the concepts of quantum Kolmogorov complexity [6] and quantum information distance [7]. Then we review the concept of second quantised Kolmogorov complexity [9].

3.1 | Indeterminate and determinate length states

A quantum bit or qubit is a unit length vector in the Hilbert space $\mathbb{C}^2$. We pick the standard orthonormal basis $\{|0\rangle, |1\rangle\}$, where $|0\rangle = (1, 0)$ and $|1\rangle = (0, 1)$. Thus, a qubit $|\psi\rangle$ may be written as a linear superposition of the basis states:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

26328925, 2023, 1, Downloaded from https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/qtc2.12050 by National Health And Medical Research Council, Wiley Online Library on [04/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

where $\alpha, \beta \in \mathbb{C}$ and $|\alpha|^2 + |\beta|^2 = 1$.

Analogously, an $n-$ qubit is a unit vector in the n-fold tensor space $\mathcal{H}_n := \mathbb{C}^2 \otimes \cdots \otimes \mathbb{C}^2$. An $n-$ qubit $|\varphi\rangle$ can be expressed as a superposition of the orthonormal basis $n-$ qubit states $|00\cdots 0\rangle$ through $|11\cdots 1\rangle$:

$$|\varphi\rangle = \sum_{x\in\{0,1\}^n} \alpha_x |x\rangle$$

where $\alpha_x \in \mathbb{C}$ and $\sum_{x\in\{0,1\}^n} |\alpha_x|^2 = 1$.

Since this paper wants to allow superpositions of different lengths, we consider the Hilbert space $\mathcal{H}_{\{0,1\}^*}$ defined as

$$\mathcal{H}_{\{0,1\}^*} := \bigoplus_{n=0}^{\infty} \mathcal{H}_n.$$

The classical finite binary strings $\{0,1\}^*$ are identified with the basis vectors in $\mathcal{H}_{\{0,1\}^*}$.

In classical information theory, codewords in a variable-length code have determinate lengths, for example, the codewords 0, 10, 11 of the code {0, 10, 11} have lengths 1, 2, 2, respectively, whereas codewords in a quantum variable-length code have indeterminate lengths because of the superposition.

Definition 1 [19]. A quantum state

$$|\varphi\rangle = \sum_i \alpha_i |i\rangle \tag{9}$$

has an indeterminate length if there exists j and k such that $|\alpha_j| > 0$ and $|\alpha_k| > 0$ but $l(j) \neq l(k)$. We say $|\varphi\rangle$ has a determinate length if it does not have an indeterminate length.

For example, the quantum state

$$\frac{1}{\sqrt{2}}(|10\rangle + |110011\rangle) \tag{10}$$

is an indeterminate length (or variable length) quantum state. It is a superposition of two strings of different lengths.

Boströem and Felbinger [20] defined two concepts of the length of an indeterminate length quantum state.

Definition 2 [20]. Let $|\varphi\rangle = \sum_i \alpha_i |i\rangle$ be an indeterminate length quantum state. The base length of $|\varphi\rangle$ is the length of the longest component of its superposition

$$L(|\varphi\rangle) = \max_{|\alpha_i|>0} l(i). \tag{11}$$

The average length of $|\varphi\rangle$ is the expected length of its superposition

$$\widehat{L}(|\varphi\rangle) = \sum_i |\alpha_i|^2 l(i). \tag{12}$$

Obviously, we have $L(|\varphi\rangle) \geq \widehat{L}(|\varphi\rangle)$.

For example, the indeterminate length quantum state

$$\frac{1}{\sqrt{2}}(|1011\rangle + |110011\rangle) \tag{13}$$

has base length 6 and average length 5.

Two quantum states can be concatenated by simply appending the components.

Definition 3 [19, 20]. Let

$$|\varphi\rangle = \sum_i \alpha_i |i\rangle, \tag{14}$$

$$|\phi\rangle = \sum_j \beta_j |j\rangle \tag{15}$$

be two quantum states. Their concatenation is

$$|\zeta\rangle = \sum_{i,j} \alpha_i \beta_j |ij\rangle. \tag{16}$$

Lemma 1 *(additivity of base and average lengths). If $|\zeta\rangle$ is the concatenation of $|\varphi\rangle$ and $|\phi\rangle$, then*

(1) $L(|\zeta\rangle) = L(|\varphi\rangle) + L(|\phi\rangle)$;
(2) $\widehat{L}(|\zeta\rangle) = \widehat{L}(|\varphi\rangle) + \widehat{L}(|\phi\rangle)$.

Proof 1 If i ($|\alpha_i| > 0$) is the longest component of $|\varphi\rangle$, j ($|\beta_j| > 0$) is the longest component of $|\phi\rangle$. It is obvious that ij ($|\alpha_i\beta_j| > 0$) is the longest component of $|\zeta\rangle$. Thus $L(|\zeta\rangle) = L(|\varphi\rangle) + L(|\phi\rangle)$.

(2) Since $\sum_i |\alpha_i|^2 = \sum_j |\beta_j|^2 = 1$, we have

$$\begin{aligned}
\widehat{L}(|\zeta\rangle) &= \sum_{ij} |\alpha_i\beta_j|^2 l(ij) = \sum_i \left(\sum_j |\alpha_i\beta_j|^2 (l(i) + l(j)) \right) \\
&= \sum_i \left(\sum_j |\alpha_i\beta_j|^2 l(i) + \sum_j |\alpha_i\beta_j|^2 l(j) \right) \\
&= \sum_i \left(\sum_j |\beta_j|^2 |\alpha_i|^2 l(i) + |\alpha_i|^2 \widehat{L}(|\phi\rangle) \right) \\
&= \sum_i \left(|\alpha_i|^2 l(i) + |\alpha_i|^2 \widehat{L}(|\phi\rangle) \right) \\
&= \sum_i |\alpha_i|^2 l(i) + \sum_i |\alpha_i|^2 \widehat{L}(|\phi\rangle) \\
&= \widehat{L}(|\varphi\rangle) + \widehat{L}(|\phi\rangle).
\end{aligned} \tag{17}$$

26328925, 2023, 1, Downloaded from https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/qtc2.12050 by National Health And Medical Research Council, Wiley Online Library on [04/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Example 1 Two quantum states are given as

$$|\varphi\rangle = \frac{1}{\sqrt{2}}(|10\rangle + |1001\rangle), \tag{18}$$

$$|\phi\rangle = \frac{1}{\sqrt{2}}(|11\rangle + |011\rangle) \tag{19}$$

$|\varphi\rangle$ has base length 4 and average length 3, and $|\phi\rangle$ has base length 3 and average length 2.5. Their concatenation is

$$|\zeta\rangle = \frac{1}{2}(|1011\rangle + |10011\rangle + |100111\rangle + |1001011\rangle). \tag{20}$$

Then $|\zeta\rangle$ has base length 7 and average length 5.5.

3.2 | Quantum Kolmogorov complexity and quantum information distance

Both base length and average length are used in quantum Kolmogorov complexity. Berthiaume et al. [6] defined the quantum Kolmogorov complexity as the minimal base length of its shortest quantum description on a universal quantum Turing machine U

$$QK_U(|\varphi\rangle) = \min\{L(|\phi\rangle) : U(|\phi\rangle) = |\varphi\rangle\}. \tag{21}$$

Remark 1 *In this paper $U(|\phi\rangle) = |\varphi\rangle$ also includes the approximation case, that is, for any k*

$$\text{Fidelity}\left(U\left(\left|\phi\right\rangle, |1^k\rangle\right), |\varphi\rangle\right) \geq f(k) \tag{22}$$

where 1^k denotes the string consisting of k 1's and $f(k)$ is a fidelity function converging to 1 such as $1 - \frac{1}{k}$ and $1 - \frac{1}{2^k}$. This means that U converts $|\phi\rangle$ to $|\varphi\rangle$ with fidelity converging to 1.

Then the **conditional quantum Kolmogorov complexity** is defined as

$$QK_U(|\varphi\rangle||\phi\rangle) = \min\{L(|\chi\rangle) : U(|\phi\rangle, |\chi\rangle) = |\varphi\rangle\}. \tag{23}$$

Based on Berthiaume et al.'s conditional quantum Kolmogorov complexity, Dai [7] defined the **quantum information distance** as the minimal base length of a shortest programme that converts $|\varphi\rangle$ to $|\phi\rangle$ as well as converts $|\phi\rangle$ to $|\varphi\rangle$ on U

$$QD_U(|\varphi\rangle, |\phi\rangle) = \min\{L(|\chi\rangle) : U(|\phi\rangle, |\chi\rangle) = |\varphi\rangle \text{ and } U(|\varphi\rangle, |\chi\rangle) = |\phi\rangle\}. \tag{24}$$

These concepts are invariant [6, 7]. Let T be a universal quantum Turing machine, then for any quantum Turing machine Q, there exists a constant c_Q (dependent only on Q) such that

$$QK_T(|\varphi\rangle) \leq QK_Q(|\varphi\rangle) + c_Q, \tag{25}$$

$$QK_T(|\varphi\rangle||\phi\rangle) \leq QK_Q(|\varphi\rangle||\phi\rangle) + c_Q, \tag{26}$$

$$QD_T(|\varphi\rangle, |\phi\rangle) \leq QD_Q(|\varphi\rangle, |\phi\rangle) + c_Q. \tag{27}$$

for any strings $|\phi\rangle$ and $|\varphi\rangle$.

Another highly related theoretical work is due to Gács's quantum algorithmic entropy [21]. In application, quantum Kolmogorov complexities have been used. Until now, quantum Kolmogorov complexities have some applications, including communication complexity theory [22], quantum key distribution (QKD) [23], bounded quantum memory [24] and quantum data compression [25].

3.3 | Second quantised Kolmogorov complexity

The descriptions in quantum Kolmogorov complexity [6] and quantum information distance [7] have determinate lengths. Rogers and Vedral [9] use indeterminate length quantum states to define **second quantised Kolmogorov complexity**, which is described as the minimum average length of its quantum description on a quantum machine U

$$\widehat{QK}_U(|\varphi\rangle) = \min\left\{\widehat{L}(|\phi\rangle) : U(|\phi\rangle) = |\varphi\rangle\right\}. \tag{28}$$

Then the **conditional second quantised Kolmogorov complexity** is defined as

$$\widehat{QK}_U(|\varphi\rangle||\phi\rangle) = \min\left\{\widehat{L}(|\chi\rangle) : U(|\phi\rangle, |\chi\rangle) = |\varphi\rangle\right\}. \tag{29}$$

The second quantised Kolmogorov complexity is also invariant.

Theorem 1 *(Invariance). [9] There is a reference universal quantum Turing machine T* such that for any quantum Turing machine Q there exists a constant c_Q depending only on U such that for any strings $|\varphi\rangle$ and $|\phi\rangle$,

$$\widehat{QK}_T(|\varphi\rangle||\phi\rangle) \leq \widehat{QK}_Q(|\varphi\rangle||\phi\rangle) + c_Q \tag{30}$$

4 | SECOND QUANTISED INFORMATION DISTANCE

We define the second quantised information distance based on the conditional second quantised Kolmogorov complexity.

4.1 | Definition and properties

Definition 4 (Second quantised information distance with respect to U). The second quantised information distance SD_U between quantum states $|\varphi\rangle$ and $|\phi\rangle$ with respect to a quantum Turing machine U is

26328925, 2023, 1, Downloaded from https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/qtc2.12050 by National Health And Medical Research Council, Wiley Online Library on [04/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$\mathrm{SD}_U(|\varphi\rangle, |\phi\rangle) = \min\Big\{\widehat{L}(|\chi\rangle)\colon U(|\phi\rangle, |\chi\rangle) = |\varphi\rangle \text{ and } U(|\varphi\rangle, |\chi\rangle) = |\phi\rangle\Big\}. \quad (31)$$

This distance is defined as the minimal average length of a shortest programme that converts $|\varphi\rangle$ to $|\phi\rangle$ as well as converts $|\phi\rangle$ to $|\varphi\rangle$ on the machine U.

Theorem 2 *(Relationship of conditional complexity and SD). For any quantum Turing machine* Q and strings $|\varphi\rangle$ and $|\phi\rangle$,

$$\widehat{QK}_U(|\varphi\rangle, |\phi\rangle) \geq \mathrm{SD}_U(|\varphi\rangle, |\phi\rangle) \geq \max\Big(\widehat{QK}_U(|\varphi\rangle||\phi\rangle), \widehat{QK}_U(|\phi\rangle||\varphi\rangle)\Big). \quad (32)$$

Proof This is easily obtained from their definitions.

Lemma 2 *(Fidelity of composition). [6] For any density matrices* σ_1, σ_2 and σ_3, *if Fidelity*$(\sigma_1, \sigma_2) \geq 1 - \epsilon_1$ *and Fidelity* $(\sigma_2, \sigma_3) \geq 1 - \epsilon_2$, *then Fidelity*$(\sigma_1, \sigma_3) \geq 1 - 2\epsilon_1 - 2\epsilon_2$

The second quantised information distance is also invariant.

Theorem 3 ***(Invariance).*** *There is a reference universal Turing machine T such that for any quantum Turing machine* Q *and strings* $|\varphi\rangle$ *and* $|\phi\rangle$,

$$\mathrm{SD}_T(|\varphi\rangle, |\phi\rangle) \leq \mathrm{SD}_U(|\varphi\rangle, |\phi\rangle) + c_Q \quad (33)$$

c_Q is a constant depending only on Q.

Proof This is based on the existence of a reference universal quantum Turing machine.

For any strings $|\varphi\rangle$ and $|\phi\rangle$, assume that $|\chi\rangle$ is a minimum average length string such that

$$\mathrm{Fidelity}\Big(Q\Big(|\chi\rangle, |\varphi\rangle, 1^k\Big), |\phi\rangle\Big) \geq 1 - \frac{1}{2^{k+2}}, \quad (34)$$

$$\mathrm{Fidelity}\Big(Q\Big(|\chi\rangle, |\phi\rangle, 1^k\Big), |\varphi\rangle\Big) \geq 1 - \frac{1}{2^{k+2}} \quad (35)$$

Since T is a reference universal Turing machine with the description corresponding to c_Q, T can simulate Q with arbitrary accuracy, and then we have

$$\mathrm{Fidelity}\Big(T\Big(|\chi\rangle, |\varphi\rangle, 1^k\Big), Q\Big(|\chi\rangle, |\varphi\rangle, 1^k\Big)\Big) \geq 1 - \frac{1}{2^{k+2}}, \quad (36)$$

$$\mathrm{Fidelity}\Big(T\Big(|\chi\rangle, |\phi\rangle, 1^k\Big), Q\Big(|\chi\rangle, |\varphi\rangle, 1^k\Big)\Big) \geq 1 - \frac{1}{2^{k+2}} \quad (37)$$

By Lemma 2, we have

$$\mathrm{Fidelity}\Big(U\Big(|\chi\rangle, |\varphi\rangle, 1^k\Big), |\phi\rangle\Big) \geq 1 - \frac{1}{2^k}, \quad (38)$$

$$\mathrm{Fidelity}\Big(U\Big(|\chi\rangle, |\phi\rangle, 1^k\Big), |\varphi\rangle\Big) \geq 1 - \frac{1}{2^k} \quad (39)$$

Then $\mathrm{SD}_T(|\varphi\rangle, |\phi\rangle) \leq \widehat{(}|\chi\rangle) + c_Q$. This completes the proof.

4.2 | Relationship of quantum information distance and second quantised information distance

First, quantum information distance [7] and second quantised information distance have the following relationship.

Theorem 4 ***(Relationship of QD and SD).*** *For any quantum Turing machine* Q and strings $|\varphi\rangle$ and $|\phi\rangle$,

$$\mathrm{SD}_U(|\varphi\rangle, |\phi\rangle) \leq QD_U(|\varphi\rangle, |\phi\rangle). \quad (40)$$

Proof Because for any $|\chi\rangle$ with $U(|\phi\rangle, |\chi\rangle) = |\varphi\rangle$ and $U(|\varphi\rangle, |\chi\rangle) = |\phi\rangle$, we have $\widehat{L}(|\chi\rangle) \leq L(|\chi\rangle)$.

Then we come to the problem: why we need a new information distance between quantum states. What is the difference between the two information distances? We give an example to illustrate the difference. Consider the state $|x\rangle = \sqrt{\frac{n-1}{n}}|0\rangle + \frac{1}{\sqrt{n}}|n\rangle$. Then for some large n, $L(|x\rangle)$ is big but $\widehat{L}(|x\rangle)$ is small. And especially, if n is incompressible, then $\mathrm{QK}_U(|x\rangle)$ is big but $\mathrm{SK}_U(|x\rangle)$ is small. With this idea of thought, it is possible that $\mathrm{QD}_U(|0\rangle, |x\rangle)$ is big but $SD_U(|0\rangle, |x\rangle)$ is small. This means that for some strings $|\varphi\rangle$ and $|\phi\rangle$, their quantum information distance is big but their second quantised information distance is small. So the second quantised information distance is necessary for measuring the difference between quantum states.

4.3 | Second quantised information distance as minimum conversion energy

Rallan and Vedral [26] gave a physical interpretation of the average length of indeterminate length states. $\widehat{L}(|\varphi\rangle)$ is described as the energy within this state. Along the way, Rogers and Vedral [10] described the second quantised Kolmogorov complexity of a state as the minimum energy required to describe that state.

Based on these interpretations, we give a physical interpretation of the second quantised information distance, which can be described as the minimum energy of conversion cost. The second quantised information distance $\mathrm{SD}_T(|\varphi\rangle, |\phi\rangle)$ between

26328925, 2023, 1, Downloaded from https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/qtc2.12050 by National Health And Medical Research Council, Wiley Online Library on [04/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$|\varphi\rangle$ and $|\phi\rangle$ is viewed as the minimum energy requirement of conversion between two quantum strings.

4.4 | Normalised quantum information distance and normalised quantum compression distance

Following the same reasoning for information distance, quantum information distances are not computable. In classical information distance theory, Li et al. [12] developed the normalised compression distance, in which the classical Kolmogorov complexity is approximated using real-world compressors such as gzip or GenCompress. In this section, we develop the quantum compression distance where the quantum Kolmogorov complexity is approximated using quantum data compression [19, 20].

First, we give the normalised versions of quantum information distance [7] and second quantised information distance.

Definition 5 The normalised distance $f_1(|\varphi\rangle, |\phi\rangle)$ of two pure quantum states $|\varphi\rangle, |\phi\rangle \in \mathcal{H}_{\{0,1\}^*}$ with respect to U is:

$$f_1(|\varphi\rangle,|\phi\rangle) = \frac{\max\{QK_U(|\varphi\rangle||\phi\rangle), QK_U(|\phi\rangle||\varphi\rangle)\}}{\max\{QK_U(|\varphi\rangle), QK_U(|\phi\rangle)\}} \quad (41)$$

Definition 6 The normalised distance $f_2(|\varphi\rangle, |\phi\rangle)$ of two pure quantum states $|\varphi\rangle, |\phi\rangle \in \mathcal{H}_{\{0,1\}^*}$ with respect to U is:

$$f_2(|\varphi\rangle,|\phi\rangle) = \frac{\max\left\{\widehat{QK}_U(|\varphi\rangle||\phi\rangle), \widehat{QK}_U(|\phi\rangle||\varphi\rangle)\right\}}{\max\left\{\widehat{QK}_U(|\varphi\rangle), \widehat{QK}_U(|\phi\rangle)\right\}} \quad (42)$$

A set S of quantum states is prefix free [19, 20] if any two (not necessarily distinct) quantum states $|\varphi\rangle$ and $|\phi\rangle$ in S are not prefixes of one another.

A lossless quantum code C [19, 20] is an isometric linear map from $\mathcal{H}$ into a closed prefix-free subspace $\mathcal{H}' \subset \mathcal{H}_{\{0,1\}^*}$. A number of quantum data compressions have been discussed [27–30].

Based on the lossless quantum code C, the normalised compression distances between quantum states are given as follows:

Definition 7 Let C be a lossless quantum code. The normalised distance $d_1(|\varphi\rangle, |\phi\rangle)$ of two quantum states $|\varphi\rangle$ and $|\phi\rangle$ with respect to C is:

$$d_1(|\varphi\rangle,|\phi\rangle) = \frac{L(C(|\varphi\phi\rangle)) - \min\{L(C(|\varphi\rangle)), L(C(|\phi\rangle))\}}{\max(L(C(|\varphi\rangle)), L(C(|\phi\rangle)))} \quad (43)$$

Definition 8 Let C be a lossless quantum code. The normalised distance $d_2(|\varphi\rangle, |\phi\rangle)$ of two quantum states $|\varphi\rangle$ and $|\phi\rangle$ with respect to C is:

$$d_2(|\varphi\rangle,|\phi\rangle) = \frac{\widehat{L}(C(|\varphi\phi\rangle)) - \min\left\{\widehat{L}(C(|\varphi\rangle)), \widehat{L}(C(|\phi\rangle))\right\}}{\max\left(\widehat{L}(C(|\varphi\rangle)), \widehat{L}(C(|\phi\rangle))\right)} \quad (44)$$

Let $\mathcal{E} = \{p_i, |\varphi_i\rangle\}$ be an ensemble of quantum states in a Hilbert space; the expected base length of compression of C is [29]

$$\overline{L}(C(\mathcal{E})) = \sum_i p_i L(C(|\varphi_i\rangle)) \quad (45)$$

Based on this the expected base length of compression and the normalised quantum compression distances between two ensembles of quantum states are given as follows:

Definition 9 Let C be a lossless quantum code. The normalised distance $d(\mathcal{E},\mathcal{F})$ of two ensembles $\mathcal{E} = \{p_i, |\varphi_i\rangle\}$ and $\mathcal{F} = \{q_i, |\phi_i\rangle\}$ with respect to C is:

$$d(\mathcal{E},\mathcal{F}) = \frac{\overline{L}(C(\mathcal{E}\otimes\mathcal{F})) - \min\{\overline{L}(C(\mathcal{E})), \overline{L}(C(\mathcal{F}))\}}{\max(\overline{L}(C(\mathcal{E})), \overline{L}(C(\mathcal{F})))} \quad (46)$$

Based on this the expected base length of compression of C is optimal [29] if for any other code C',

$$\overline{L}(C(\mathcal{E})) \le \overline{L}(C'(\mathcal{E})) \quad (47)$$

Suppose we have two different ensembles $\mathcal{E} = \{p_i, |\varphi_i\rangle\}$ and $\mathcal{F} = \{q_i, |\phi_i\rangle\}$, which have optimal codes C_E and C_F, respectively. As the codes are prefix free, we may concatenate them to obtain a code $C_E C_F$ for $\mathcal{E}\otimes\mathcal{F}$. Then

$$\overline{L}(C_E \circ C_F(\mathcal{E}\otimes\mathcal{F})) = \overline{L}(C_E(\mathcal{E})) + \overline{L}(C_F(\mathcal{F})) \quad (48)$$

Let C_{EF} be an optimal code for $\mathcal{E}\otimes\mathcal{F}$, then

$$\overline{L}(C_{E\circ F}(\mathcal{E}\otimes\mathcal{F})) \le \overline{L}(C_E \circ C_F(\mathcal{E}\otimes\mathcal{F})) \quad (49)$$

Lemma 3 *If* $C_E C_F$ is also an optimal code for $\mathcal{E}\otimes\mathcal{F}$, then $d(\mathcal{E},\mathcal{F}) = 1$

Proof If $\overline{L}(C_E(\mathcal{E})) \ge \overline{L}(C_F(\mathcal{F}))$ then

26328925, 2023, 1, Downloaded from https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/qtc2.12050 by National Health And Medical Research Council, Wiley Online Library on [04/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$
\begin{aligned}
&d(\mathcal{E},\mathcal{F})\\
&=\frac{\overline{L}(C_E(\mathcal{E}))+\overline{L}(C_F(\mathcal{F}))-\min\{\overline{L}(C_E(\mathcal{E})),\overline{L}(C_F(\mathcal{F}))\}}{\max(\overline{L}(C_E(\mathcal{E})),\overline{L}(C_F(\mathcal{F})))}\\
&=\frac{\overline{L}(C_E(\mathcal{E}))+\overline{L}(C_F(\mathcal{F}))-\overline{L}(C_F(\mathcal{F}))}{\overline{L}(C_E(\mathcal{E}))}\\
&=\frac{\overline{L}(C_E(\mathcal{E}))}{\overline{L}(C_E(\mathcal{E}))}\\
&=1.
\end{aligned}
\tag{50}
$$

Similarly, we have $d(\mathcal{E},\mathcal{F})=1$ for $\overline{L}(C_E(\mathcal{E}))<\overline{L}(C_F(\mathcal{F}))$.

5 | CONCLUSION

Classical information distance based on Kolmogorov complexity is a good tool for describing the difference between two individual strings. In the quantum setting, although a quantum information distance [7] has been developed based on Berthiaume et al.'s quantum Kolmogorov complexity [6], this description of a state used the length of the longest component of its superposition. In this paper, by using the minimum average length of an indeterminate length quantum state, we defined the second quantised information distance with an intuitive physical interpretation. It is based on Rogers and Vedral's second quantised Kolmogorov complexity [9, 10]. We also developed a practical analogue of quantum information distance based on lossless quantum data compression.

There are a number of applications of classical information distance. We hope that quantum information distances will lead to a similar wide variety of applications in quantum information processing. We present our views on some potential applications.

1) In 2009, Miyadera and Imai [23] formulated a security criterion for the QKD by using the quantum Kolmogorov complexity. QKD plays a major role in quantum communication networks [31]. QKD algorithms are now also being taught in undergraduate studies [32]. As future work, we can study the security and privacy of quantum information and communication technologies by using the second quantised quantum Kolmogorov complexity and second quantised information distance.
2) Classical information distances are used frequently in image processing [33–35]. In recent years, increasing numbers of researchers turned their attention to quantum image processing [36, 37]. It is very applicable to study the application of quantum information distance in quantum image processing.

ACKNOWLEDGEMENT

This work was supported by the National Science Foundation of China under Grant No. 62006168 and Zhejiang Provincial Natural Science Foundation of China under Grant No. LQ 21A010001.

CONFLICT OF INTEREST

The author has declared no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analysed in this study.

ORCID

Songsong Dai https://orcid.org/0000-0003-2283-9459

REFERENCES

1. Bennett, C.H., et al.: Information distance. IEEE Trans. Inf. Theor. 44(4), 1407–1423 (1998). https://doi.org/10.1109/18.681318
2. Solomonoff, R.: A formal theory of inductive inference, part I. Inf. Control 7(1), 1–22 (1964). https://doi.org/10.1016/s0019-9958(64)90223-2
3. Kolmogorov, A.: Three approaches to the quantitative definition of information. Probl. Inf. Transm. 1(1), 1–7 (1965)
4. Chaitin, G.: On the length of programs for computing finite binary sequences: statistical considerations. J. ACM 16(1), 145–159 (1969). https://doi.org/10.1145/321495.321506
5. Vitányi, P.M.B.: Quantum Kolmogorov complexity based on classical descriptions. IEEE Trans. Infor. Theor 47(6), 2464–2479 (2001). https://doi.org/10.1109/18.945258
6. Berthiaume, A., Van Dam, W., Laplante, S.: Quantum Kolmogorov complexity. J. Comput. Syst. Sci. 63(2), 201–221 (2001). https://doi.org/10.1006/jcss.2001.1765
7. Dai, S.: Quantum information distance. Int. J. Quan. Inform. 19(6), 2150031 (2021). https://doi.org/10.1142/s0219749921500313
8. Dai, S.: Quantum information distance based on classical descriptions. Quan. Machine Intelligence 4(1), 7 (2022). https://doi.org/10.1007/s42484-022-00064-2
9. Rogers, C., Vedral, V.: The second quantized quantum Turing machine and Kolmogorov complexity. Mod. Rev. Phys. Lett. B 22(12), 1203–1210 (2008). https://doi.org/10.1142/s021798490801464x
10. Rogers, C., Vedral, V., Nagarajan, R.: The second quantized Kolmogorov complexity. Int. J. Quant. Inf. 6(4), 907–928 (2008). https://doi.org/10.1142/s021974990800375x
11. Li, M., Vitányi, P.M.B.: An Introduction to Kolmogorov Complexity and its Applications, 4th ed. Springer-Verlag, New York (2019)
12. Li, M., et al.: The similarity metric. IEEE Trans Inf Theor 50(12), 3250–3264 (2004). https://doi.org/10.1109/tit.2004.838101
13. Li, M., et al.: An information-based sequence distance and its application to whole mito chondrial genome phylogeny. Bioinformatics 17(2), 149–154 (2001). https://doi.org/10.1093/bioinformatics/17.2.149
14. Li, M.: Information distance and its applications. Int. J. Found Comput. Sci. 18(04), 669–681 (2007). https://doi.org/10.1142/s0129054107004905
15. Cilibrasi, R.L., Vitányi, P.M.B., Wolf, R.: Algorithmic clustering of music based on string compression. Comput. Music J. 28(4), 49–67 (2004). https://doi.org/10.1162/0148926042728449
16. Cilibrasi, R.L., Vitányi, P.M.B.: Clustering by compression. IEEE Trans. Inf. Theo.r 51(4), 1523–1545 (2005). https://doi.org/10.1109/tit.2005.844059
17. Chen, X., et al.: Shared information and program plagiarism detection. IEEE Trans. Inf. Theor. 50(7), 1545–1551 (2004). https://doi.org/10.1109/tit.2004.830793
18. Zhang, X., et al.: New information distance measure and its application in question answering system. J. Comput. Sci. Technol. 23(4), 557–572 (2008). https://doi.org/10.1007/s11390-008-9152-9
19. Schumacher, B., Westmoreland, M.: Indeterminate-length quantum coding. Phys. Rev. 64(4), 042304 (2001). https://doi.org/10.1103/physreva.64.042304
20. Bostroem, K., Felbinger, T.: Lossless quantum data compression and variable-length coding. Phys. Rev. 65(3), 032313 (2002). https://doi.org/10.1103/physreva.65.032313
21. Gács, P.: Quantum algorithmic entropy. J. Phys. Math. Gen. 34(35), 6859–6880 (2001). https://doi.org/10.1088/0305-4470/34/35/312

26328925, 2023, 1, Downloaded from https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/qtc2.12050 by National Health And Medical Research Council, Wiley Online Library on [04/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

22. Mora, C., Briegel, H., Kraus, B.: Quantum Kolmogorov complexity and its applications. Int. J. Quant. Inf. 5(05), 12–750 (2006). https://doi.org/10.1142/s0219749907003171
23. Miyadera, T., Imai, H.: Quantum Kolmogorov complexity and quantum key distribution. Phys. Rev. 79(1), 012324 (2009). https://doi.org/10.1103/physreva.79.012324
24. Miyadera, T.: Quantum Kolmogorov complexity and bounded quantum memory. Phys. Rev. 83(4), 042316 (2011). https://doi.org/10.1103/physreva.83.042316
25. Rogers, C., Nagarajan, R.: Lossless quantum data compression and quantum Kolmogorov complexity. Int J Quant Inf 9(02), 637–650 (2011). https://doi.org/10.1142/s0219749911007848
26. Rallan, L., Vedral, V.: Energy requirements for quantum data compression and 1-1 coding. Phys Rev A, Phys Rev A 68(4), 042309 (2003). https://doi.org/10.1103/physreva.68.042309
27. Hayashi, M., Matsumoto, K.: Quantum universal variable-length source coding. Phys. Rev. 66(2), 022311 (2002). https://doi.org/10.1103/physreva.66.022311
28. Hayden, P., Jozsa, R., Winter, A.: Trading quantum for classical resources in quantum data compression. J. Math. Phys. 43(9), 4404–4444 (2002). https://doi.org/10.1063/1.1497184
29. Mueller, M., Rogers, C., Nagarajan, R.: Lossless quantum prefix compression for communication channels that are always open. Phys. Rev. 79(1), 012302 (2009). https://doi.org/10.1103/physreva.79.012302
30. Ahlswede, R., Cai, N.: On lossless quantum data compression with a classical helper. IEEE Trans Inf Theor 50(6), 1208–1219 (2004). https://doi.org/10.1109/tit.2004.828071
31. Liu, R., et al.: Towards the industrialisation of quantum key distribution in communication networks: a short survey. IET Quant Comm 3(3), 151–163 (2022). https://doi.org/10.1049/qtc2.12044
32. Bloom, Y., et al.: Quantum cryptography simplified undergraduate experiment and simulation. Physics 4(1), 104–123 (2022). https://doi.org/10.3390/physics4010009
33. Cohen, A.R., Vitányi, P.M.B.: Normalized compression distance of multisets with applications. IEEE Trans. Pattern Anal. Mach. Intell. 37(8), 1602–1614 (2014). https://doi.org/10.1109/tpami.2014.2375175
34. Guha, T., Ward, R.K.: Image similarity using sparse representation and compression distance. IEEE Trans. Multimed. 16(4), 980–987 (2014). https://doi.org/10.1109/tmm.2014.2306175
35. Pinho, A.J., Ferreira, P.J.S.G.: 'Image similarity using the normalized compression distance based on finite context models. In: 18th IEEE International Conference on Image Processing, pp. 1993–1996 (2011)
36. Yan, F., Venegas-Andraca, S.E.: Quantum Image Processing. Springer, Berlin (2020)
37. Elaraby, A.: Quantum medical images processing foundations and applications. IET Quant Comm, 1–13 (2022). https://doi.org/10.1049/qtc2.12049

How to cite this article: Dai, S.: Second quantised information distance. IET Quant. Comm. 4(1), 17–24 (2023). https://doi.org/10.1049/qtc2.12050

26328925, 2023, 1, Downloaded from https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/qtc2.12050 by National Health And Medical Research Council, Wiley Online Library on [04/10/2024]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License