International Journal of Theoretical Physics, Vol. 41, No. 5, May 2002 (© 2002)

Characterization of Sequential Quantum Machines

Daowen Qiu[1]

Received September 9, 2001

We further investigate some properties of sequential quantum machines (SQMs) and introduce so-called quantum sequential machines (QSMs). In particular, the equivalence between SQMs and QSMs is also presented. We give a counterexample to answer an open problem proposed by S. Gudder recently.

KEY WORDS: quantum computation; stochastic sequential machines; sequential quantum machines.

1. INTRODUCTION

As theoretical models of quantum computation, quantum automata and more complicated quantum machines have become an important research field (Bertoni and Carpentieri, 2001; Gudder, 1999, 2000; Moore and Crutchfield, 2000; and see Gruska, 1999, pp. 151–192 for the details). Recently, Gudder (2000) considered sequential quantum machines (SQMs), which may be looked as a quantum variant of stochastic sequential machines (SSMs) (Paz, 1971). As is well known, an important result on SSMs is that two SSMs with n and n' states, respectively and the same input and output alphabets are equivalent if and only if they are $(n + n' - 1)$-equivalent (see Theorem 2.7 in Paz, 1971). So Gudder (2000) proposed an open problem of whether it also holds for SQMs. More exactly, let $\mathcal{M}$ and $\mathcal{M}'$ be SQMs with n and n' states, respectively and the same input and output alphabets. Is it true that $\mathcal{M}$ and $\mathcal{M}'$ are equivalent (i.e., $p_{\mathcal{M}}(v \mid u) = p_{\mathcal{M}'}(v \mid u)$ for all words u, v) if and only if $p_{\mathcal{M}}(v \mid u) = p_{\mathcal{M}'}(v \mid u)$ for all words u, v with length not bigger than $n + n' - 1$? (See Gudder, 2000, p. 2159.) In this paper, a negative answer is given.

In Section 2, we first recall the definition of SQMs and then discuss the relation between the extension of their transition amplitude functions and the transition operators. Afterwards, we in Section 3 introduce quantum sequential machines (QSMs), which may be more analogous to SSMs formally than SQMs. In particular, we prove that the classes of SQMs and QSMs are exactly equivalent to

[1] State Key Laboratory of Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University, 100084 Beijing, People's Republic of China; e-mail: qiudw@tsinghua.edu.cn.

each other (Theorem 2). Finally a counterexample is presented in Section 4 and therefore the above problem proposed by Gudder is solved.

2. SEQUENTIAL QUANTUM MACHINES

Sequential quantum machines (SQMs) was considered by Gudder (2000), and there so-called factorizable and strongly factorizable SQMs were also proposed. In this section, we discuss the relation between the extension of transition amplitude function δ and the transition operator U that describes the evolution of states.

A SQM is a 5-tuple $\mathcal{M} = (S, s_0, I, O, \delta)$, where S is a finite set of internal states, $s_0 \in S$ is the start state, I and O are finite input and output alphabets, respectively, and $\delta : I \times S \times O \times S \to \mathbf{C}$ is a *transition amplitude function* satisfying

$$\sum_{y,t} \delta(x, s, y, t)\delta(x, s', y, t)^* = \delta_{s,s'} \tag{1}$$

for every $x \in I$, $s, s' \in S$. The symbol $*$ stands for complex conjugation and $\delta_{s,s'}$ is the Kronecker delta. $\delta(x, s, y, t)$ is interpreted as the transition amplitude that SQM $\mathcal{M}$ prints y and enters state t after scanning x in the current state s. In fact, there is a natural extension (see Proposition 1 (ii)) of δ to $I^* \times S \times O^* \times S$ by letting

$$\begin{aligned}
&\delta(x_1 \cdots x_m, s, y_1 \cdots y_m, t)\\
&\quad = \sum_{s_1,\ldots,s_{m-1}} \delta(x_1, s, y_1, s_1)\delta(x_2, s_1, y_2, s_2) \cdots \delta(x_m, s_{m-1}, y_m, t)
\end{aligned} \tag{2}$$

and $\delta(u, s, v, t) = 0$ for $|u| \neq |v|$, where $|u|$ and $|v|$ denote the length of words u and v, and I^* and O^* represent the sets of all words over I and O, respectively. Then we have

$$\sum_{y_1,\ldots,y_m,t} \delta(x_1 \cdots x_m, s, y_1 \cdots y_m, t)\delta(x_1 \cdots x_m, s', y_1 \cdots y_m, t)^* = \delta_{s,s'}. \tag{3}$$

Proof: By utilizing (1) repeatedly we obtain that

$$\begin{aligned}
&\sum_{y_1,\ldots,y_m,t} \delta(x_1 \cdots x_m, s, y_1 \cdots y_m, t)\delta(x_1 \cdots x_m, s', y_1 \cdots y_m, t)^*\\
&\quad = \sum_{y_1,\ldots,y_m,t} \left(\sum_{s_1,\ldots,s_{m-1}} \delta(x_1, s, y_1, s_1)\delta(x_2, s_1, y_2, s_2) \cdots \delta(x_m, s_{m-1}, y_m, t) \right)\\
&\qquad \times \left(\sum_{s'_1,\ldots,s'_{m-1}} \delta(x_1, s', y_1, s'_1)\delta(x_2, s'_1, y_2, s'_2) \cdots \delta(x_m, s'_{m-1}, y_m, t) \right)^*\\
&\quad = \sum_{y_1,\ldots,y_m,t} \sum_{s_1,\ldots,s_{m-1}} \sum_{s'_1,\ldots,s'_{m-1}} \delta(x_1, s, y_1, s_1)\delta(x_1, s', y_1, s'_1)^* \cdot \delta(x_2, s_1, y_2, s_2)\\
&\qquad \times \delta(x_2, s'_1, y_2, s'_2)^* \cdots \delta(x_m, s_{m-1}, y_m, t)\delta(x_m, s'_{m-1}, y_m, t)^*
\end{aligned}$$

$$= \sum_{y_1,\ldots,y_{m-1}} \sum_{s_1,\ldots,s_{m-1}} \sum_{s_1',\ldots,s_{m-1}'} \delta(x_1, s, y_1, s_1)\delta(x_1, s', y_1, s_1')^* \cdot \delta(x_2, s_1, y_2, s_2)$$

$$\times \delta(x_2, s_1', y_2, s_2')^* \cdots \sum_{y_m,t} \delta(x_m, s_{m-1}, y_m, t)\delta(x_m, s_{m-1}', y_m, t)^*$$

$$= \sum_{y_1,\ldots,y_{m-2}} \sum_{t_1} \sum_{s_1,\ldots,s_{m-2}} \sum_{s_1',\ldots,s_{m-2}'} \delta(x_1, s, y_1, s_1)\delta(x_1, s', y_1, s_1')^*$$

$$\cdots \delta(x_{m-1}, s_{m-2}, y_{m-1}, t_1)\delta(x_{m-1}, s_{m-2}', y_{m-1}, t_1)^*$$

$$= \cdots = \sum_{y_1,t_{m-1}} \delta(x_1, s, y_1, t_{m-1})\delta(x_1, s', y_1, t_{m-1})^*$$

$$= \delta_{s,s'}. \qquad \square$$

Now we turn to dealing with the *transition operator* U that characterizes the evolution of states. For convenience, we identify S and O^* with two orthonormal bases for some complex Hilbert spaces H_S and H_{O^*}, respectively. That is to say, H_{O^*} may be looked as a closed subspace spanned by O^*. Furthermore, H_{O^*} is isomorphic with

$$K = \mathbf{C} \oplus H_O \oplus \otimes^2 H_O \oplus \cdots \oplus \otimes^n H_O \oplus \cdots$$

where H_O is a finite dimensional Hilbert space whose basis vectors correspond to the symbols in O. For any $x_1 \cdots x_m \in I^*$, operator $U(x_1 \cdots x_m) : H_S \otimes H_{O^*} \to H_S \otimes H_{O^*}$ is defined by letting

$$U(x_1 \cdots x_m)s \otimes v = \sum_{y_1,\ldots,y_m,t} \delta(x_1 \cdots x_m, s, y_1 \cdots y_m, t)t \otimes vy_1 \cdots y_m \tag{4}$$

and extending to $H_S \otimes H_{O^*}$ by linearity and closure. A linear operator T on interior product space H is called an isometry, if $\|T\varphi\| = \|\varphi\|$ for any $\varphi \in H$. Then we have the following proposition.

Proposition 1. (i) $U(x_1 \cdots x_m)$ *is an isometry on* $H_S \otimes H_{O^*}$ *if and only if (3) holds.* (ii) $U(x_1 \cdots x_m) = U(x_m) \cdots U(x_1)$ *if and only if (2) holds.*

Proof: The proof of (i) is similar to Lemma 3.1 in Gudder (2000), so we omit it and just prove (ii). If (2) holds, then according to (4) we have that

$$U(x_1 \cdots x_m)S \otimes v$$

$$= \sum_{y_1,\ldots,y_m,t} \sum_{s_1,\ldots,s_m} \delta(x_1, s, y_1, s_1) \cdots \delta(x_m, s_{m-1}, y_m, t)t \otimes y_1 \cdots y_m$$

$$
\begin{aligned}
&= \sum_{y_1,\ldots,y_{m-1}} \sum_{s_1,\ldots,s_{m-1}} \delta(x_1, s, y_1, s_1) \cdots \delta(x_{m-1}, s_{m-2}, y_{m-1}, s_{m-1}) \\
&\quad \times U(x_m) s_{m-1} \otimes v y_1 \cdots y_{m-1} \\
&= \sum_{y_1,\ldots,y_{m-2}} \sum_{s_1,\ldots,s_{m-2}} \delta(x_1, s, y_1, s_1) \cdots \delta(x_{m-2}, s_{m-3}, y_{m-2}, s_{m-2}) \\
&\quad \times U(x_m) U(x_{m-1}) s_{m-2} \otimes v y_1 \cdots y_{m-2} \\
&= \cdots = \sum_{y_1} \sum_{s_1} \delta(x_1, s, y_1, s_1) \cdot U(x_m) \cdots U(x_2) s_1 \otimes v y_1 \\
&= U(x_m) \cdots U(x_1) s \otimes v.
\end{aligned}
$$

Conversely, since $\{s \otimes v : s \in S, v \in O^*\}$ is a pairwise orthogonal set, by (4) it is easy to follow (2) from the above proof. □

Definition 1. Let $\mathcal{M} = (S, s_0, I, O, \delta)$ be a SQM. Then the probability of the machine printing word $y_1 \cdots y_m \in O^*$, having been inputted the word $x_1 \cdots x_m \in I^*$ is defined as

$$
p_{\mathcal{M}}(y_1 \cdots y_m \mid x_1 \cdots x_m) = \sum_s |\langle U(x_1 \cdots x_m) s_0 \otimes \varepsilon, s \otimes y_1 \cdots y_m \rangle|^2.
$$

From (4) and (2) it follows that

$$
p_{\mathcal{M}}(y_1 \cdots y_m \mid x_1 \cdots x_m) = \sum_s |\delta(x_1 \cdots x_m, s_0, y_1 \cdots y_m, s)|^2 \tag{5}
$$

$$
= \sum_s \left| \sum_{s_1,\ldots,s_{m-1}} \delta(x_1, s_0, y_1, s_1) \cdots \delta(x_m, s_{m-1}, y_m, s) \right|^2 . \tag{6}
$$

3. QUANTUM SEQUENTIAL MACHINES

In this section, we introduce a class of more intuitive quantum machines, namely *quantum sequential machines* (QSMs) and show the equivalence between SQMs and QSMs.

Definition 2. A QSM is 5-tuple $\mathcal{M} = (S, \eta_{i_0}, I, O, \{A(y \mid x) : y \in O, x \in I\})$, where $S = \{s_1, s_2, \ldots, s_n\}$ is a finite set of internal states; $\eta_{i_0} = (0 \cdots 1 \cdots 0)^{\mathrm{T}}$ is a degenerate stochastic column vector of n dimension, that is, the i_0th entry is 1; I and O are input and output alphabets, respectively; $A(y \mid x)$ is an $n \times n$ matrix satisfying

$$
\sum_{y \in O} A(y \mid x) A(y \mid x)^{\dagger} = \mathbf{I} \tag{7}
$$

for any $x \in I$, where the symbol † denotes Hermitian conjugate operation and **I** is unit matrix.

Suppose that $A(y \mid x) = [a_{ij}(y \mid x)]$, then $a_{ij}(y \mid x)$ and $|a_{ij}(y \mid x)|^2$ are the transition amplitude and probability that $\mathcal{M}$ prints y and enters state s_j, having been in state s_i and scanned x, respectively. Therefore, the probability that $\mathcal{M}$ prints y and enters some states after scanning x in the current state s_{i_0}, is $\sum_j |a_{i_0 j}(y \mid x)|^2$. Usually, in the current state s_i, upon receiving a word $x_1 \cdots x_m$, $\mathcal{M}$ first scans x_1 and then prints y_1, while entering some state; under this state, after scanning x_2, $\mathcal{M}$ prints y_2 and enters some state again; according to this process, $\mathcal{M}$ does not stop until all input symbols have been scanned. So

$$\sum_{k_1,\ldots,k_{m-1}} a_{ik_1}(y_1 \mid x_1) a_{k_1 k_2}(y_2 \mid x_2) \cdots a_{k_{m-1} j}(y_m \mid x_m)$$

denoted by $a_{ij}(y_1 \cdots y_m \mid x_1 \cdots x_m)$, is the transition amplitude that $\mathcal{M}$ prints $y_1 \cdots y_m$ and enters state s_j after scanning $x_1 \cdots x_m$ step-by-step in the current state s_i. Denote

$$A(y_1 \cdots y_m \mid x_1 \cdots x_m) = A(y_1 \mid x_1) \cdots A(y_m \mid x_m)$$

then by induction one can easily get that indeed

$$A(y_1 \cdots y_m \mid x_1 \cdots x_m) = [a_{ij}(y_1 \cdots y_m \mid x_1 \cdots x_m)].$$

The probability of the above QSM $\mathcal{M}$ printing the word $y_1 \cdots y_m$ having been inputted the word $x_1 \cdots x_m$ is naturally defined as

$$p_{\mathcal{M}}(y_1 \cdots y_m \mid x_1 \cdots x_m) = \sum_j |a_{i_0 j}(y_1 \cdots y_m \mid x_1 \cdots x_m)|^2. \tag{8}$$

By a direct calculation it follows that

$$\begin{aligned} & p_{\mathcal{M}}(y_1 \cdots y_m \mid x_1 \cdots x_m) \\ &= \eta_{i_0}^{\mathrm{T}} A(y_1 \cdots y_m \mid x_1 \cdots x_m) A(y_1 \cdots y_m \mid x_1 \cdots x_m)^{\dagger} \eta_{i_0} \qquad (9) \\ &= \left\| A(y_1 \cdots y_m \mid x_1 \cdots x_m) \eta_{i_0} \right\|^2 \qquad (10) \end{aligned}$$

where the symbol T denotes the transpose operation.

Definition 3. Two machines (SQMs or QSMs) $\mathcal{M}_1$ and $\mathcal{M}_2$ with the same input and output alphabets are called equivalent if $p_{\mathcal{M}_1}(v \mid u) = p_{\mathcal{M}_2}(v \mid u)$ for any input–output pair (u, v).

Theorem 2. *For any SQM $\mathcal{M}$ there is a corresponding QSM $\mathcal{M}'$ with the same input and output alphabets, such that $\mathcal{M}$ and $\mathcal{M}'$ are equivalent, and vice versa.*

Proof: Let SQM $\mathcal{M} = (S, s_{i_0}, I, O, \delta)$, where $S = \{s_1, \ldots, s_n\}$. Then we can construct a QSM $\mathcal{M}' = (S, \eta_{i_0}, I, O, \{A(y \mid x) : y \in O, x \in I\})$, where $A(y \mid x) = [a_{ij}(y \mid x)]$ and $a_{ij}(y \mid x) = \delta(x, s_i, y, s_j)$. First we check that $A(y \mid x)$ satisfies (7). Since δ satisfies (1), we have that

$$\begin{aligned}
&\sum_y A(y \mid x) A(y \mid x)^\dagger \\
&= \sum_y \begin{bmatrix} \delta(x, s_1, y, s_1) & \cdots & \delta(x, s_1, y, s_n) \\ \vdots & \vdots & \vdots \\ \delta(x, s_n, y, s_1) & \cdots & \delta(x, s_n, y, s_n) \end{bmatrix} \begin{bmatrix} \delta(x, s_1, y, s_1) & \cdots & \delta(x, s_1, y, s_n) \\ \vdots & \vdots & \vdots \\ \delta(x, s_n, y, s_1) & \cdots & \delta(x, s_n, y, s_n) \end{bmatrix}^\dagger \\
&= \sum_y \begin{bmatrix} \sum_k |\delta(x, s_1, y, s_k)|^2 & \cdots & \sum_k \delta(x, s_1, y, s_k)\delta(x, s_n, y, s_k)^* \\ \vdots & \vdots & \vdots \\ \sum_k \delta(x, s_n, y, s_k)\delta(x, s_1, y, s_k)^* & \cdots & \sum_k |\delta(x, s_n, y, s_k)|^2 \end{bmatrix} \\
&= \mathbf{I}.
\end{aligned}$$

Next by induction we aim to prove that

$$A(y_1 \mid x_1) \cdots A(y_m \mid x_m) = [b_{ij}] \tag{11}$$

where $b_{ij} = \delta(x_1 \cdots x_m, s_i, y_1 \cdots y_m, s_j)$. Therefore, we have that

$$A(y_1 \cdots y_m \mid x_1 \cdots x_m) = [\delta(x_1 \cdots x_m, s_i, y_1 \cdots y_m, s_j)]. \tag{12}$$

Indeed, if $m = 2$, then

$$\begin{aligned}
&A(y_1 y_2 \mid x_1 x_2) = A(y_1 \mid x_1) A(y_2 \mid x_2) \\
&= \begin{bmatrix} \delta(x_1, s_1, y_1, s_1) & \cdots & \delta(x_1, s_1, y_1, s_n) \\ \vdots & \vdots & \vdots \\ \delta(x_1, s_n, y_1, s_1) & \vdots & \delta(x_1, s_n, y_1, s_n) \end{bmatrix} \begin{bmatrix} \delta(x_2, s_1, y_2, s_1) & \cdots & \delta(x_2, s_1, y_2, s_n) \\ \vdots & \vdots & \vdots \\ \delta(x_2, s_n, y_2, s_1) & \vdots & \delta(x_2, s_n, y_2, s_n) \end{bmatrix} \\
&= \begin{bmatrix} \sum_i \delta(x_1, s_1, y_1, s_i)\delta(x_2, s_i, y_2, s_1) & \cdots & \sum_i \delta(x_1, s_1, y_1, s_i)\delta(x_2, s_i, y_2, s_n) \\ \vdots & \vdots & \vdots \\ \sum_i \delta(x_1, s_n, y_1, s_i)\delta(x_2, s_i, y_2, s_1) & \cdots & \sum_i \delta(x_2, s_n, y_2, s_1)\delta(x_2, s_i, y_2, s_n) \end{bmatrix} \\
&= \begin{bmatrix} \delta(x_1 x_2, s_1, y_1 y_2, s_1) & \cdots & \delta(x_1 x_2, s_1, y_1 y_2, s_n) \\ \vdots & \vdots & \vdots \\ \delta(x_1 x_2, s_n, y_1 y_2, s_1) & \cdots & \delta(x_1 x_2, s_n, y_1 y_2, s_n) \end{bmatrix} = [\delta(x_1 x_2, s_i, y_1 y_2, s_j)].
\end{aligned}$$

Suppose that it holds for the case $m - 1$, then

$$\begin{aligned}
&A(y_1 \cdots y_m \mid x_1 \cdots x_m) = (A(y_1 \mid x_1) \cdots A(y_{m-1} \mid x_{m-1})) A(y_m \mid x_m) \\
&= \begin{bmatrix} \delta(x_1 \cdots x_{m-1}, s_1, y_1 \cdots y_{m-1}, s_1) & \cdots & \delta(x_1 \cdots x_{m-1}, s_1, y_1 \cdots y_{m-1}, s_n) \\ \vdots & \vdots & \vdots \\ \delta(x_1 \cdots x_{m-1}, s_n, y_1 \cdots y_{m-1}, s_1) & \cdots & \delta(x_1 \cdots x_{m-1}, s_n, y_1 \cdots y_{m-1}, s_n) \end{bmatrix}
\end{aligned}$$

$$\times \begin{bmatrix} \delta(x_m, s_1, y_m, s_1) & \cdots & \delta(x_m, s_1, y_m, s_n) \\ \vdots & \vdots & \vdots \\ \delta(x_m, s_n, y_m, s_1) & \cdots & \delta(x_m, s_n, y_m, s_n) \end{bmatrix}$$

$$= \begin{bmatrix} \delta(x_1 \cdots x_m, s_1, y_1 \cdots y_m, s_1) & \cdots & \delta(x_1 \cdots x_m, s_1, y_1 \cdots y_m, s_n) \\ \vdots & \vdots & \vdots \\ \delta(x_1 \cdots x_m, s_n, y_1 \cdots y_m, s_1) & \cdots & \delta(x_1 \cdots x_m, s_n, y_1 \cdots y_m, s_n) \end{bmatrix}$$

$$= [\delta(x_1 \cdots x_m, s_i, y_1 \cdots y_m, s_j)].$$

So (11) and (12) hold. By combining (5) and (8) with (12) we obtain that

$$\begin{aligned} p_{\mathcal{M}}(y_1 \cdots y_m \mid x_1 \cdots x_m) &= \sum_i \left|\delta\left(x_1 \cdots x_m, s_{i_0}, y_1 \cdots y_m, s_i\right)\right|^2 \\ &= \left\| A(y_1 \cdots y_m \mid x_1 \cdots x_m)^\dagger \eta_{i_0} \right\|^2 \\ &= p_{\mathcal{M}'}(y_1 \cdots y_m \mid x_1 \cdots x_m). \end{aligned}$$

Conversely, suppose that QMS $\mathcal{M} = (S, \eta_{i_0}, I, O, \{A(y \mid x) : y \in O, x \in I\})$ where $S = \{s_1, \ldots, s_n\}$, then we can construct a SQM $\mathcal{M}' = (S, s_{i_0}, I, O, \delta)$, such that

$$\delta(x, s_i, y, s_j) = a_{ij}(y \mid x)$$

for any $x \in I$, $y \in O$. The rest of the process is similar to the preceding proof, so we omit it and, therefore, complete the proof of lemma. □

4. A COUNTEREXAMPLE

It is our main purpose in this section to construct a counterexample outlining the differences between SSMs and SQMs (or QSMs), and particularly answering the question proposed by Gudder (2000). Let us recall that problem again. Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be SQMs or QSMs with n_1 and n_2 states, respectively, and the same input and output alphabets. Is it true that $\mathcal{M}_1$ and $\mathcal{M}_2$ are equivalent if $p_{\mathcal{M}_1}(v \mid u) = p_{\mathcal{M}_2}(v \mid u)$ for all input–output pair (u, v) with length not bigger than $n_1 + n_2 - 1$? It follows that the answer is negative from the following example.

Example 1. Let $I = \{0, 1\}$, $O = \{a, b, c\}$. Suppose that

$$\mathcal{M}_1 = (\{s_0\}, \eta, I, O, \{A(y \mid x) : y \in O, x \in I\})$$

and

$$\mathcal{M}_2 = (\{s_1, s_2\}, \eta_1, I, O, \{A(y \mid x) : y \in O, x \in I\})$$

where $A_1(a \mid 0) = A_1(b \mid 0) = \sqrt{2}/2$, $A_1(c \mid 0) = 0$, $A_1(c \mid 1) = 1$, $A_1\,(a \mid 1) = A_1(b \mid 1) = 0$, $\eta = 1$; and

$$A_2(a \mid 0) = \frac{1}{4}\begin{bmatrix} \sqrt{2}-\sqrt{2}i & 1+\sqrt{3}i \\ -1+\sqrt{3}i & \sqrt{2}+\sqrt{2}i \end{bmatrix},$$

$$A_2(b \mid 0) = A_2(a \mid 0)^* = \frac{1}{4}\begin{bmatrix} \sqrt{2}+\sqrt{2}i & 1-\sqrt{3}i \\ -1-\sqrt{3}i & \sqrt{2}-\sqrt{2}i \end{bmatrix},$$

$$A_2(c \mid 0) = A_2(a \mid 1) = A_2(b \mid 1) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

$$A_2(c \mid 1) = \frac{1}{2}\begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}.$$

We first check that $\mathcal{M}_1$ and $\mathcal{M}_2$ are exactly QSMs, that is, $A_1(y \mid x)$ and $A_2(y \mid x)$ satisfy (7):

$$\sum_y A_1(y \mid 0)A_1(y \mid 0)^\dagger = \frac{\sqrt{2}}{2} \times \frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} \times \frac{\sqrt{2}}{2} = 1;$$

$$\sum_y A_1(y \mid 1)A_1(y \mid 1)^\dagger = A_1(c \mid 1)A_1(c \mid 1)^\dagger = 1;$$

$$\begin{aligned}
\sum_y & A_2(y \mid 0)A_2(y \mid 0)^\dagger \\
&= \frac{1}{16}\begin{bmatrix} \sqrt{2}-\sqrt{2}i & 1+\sqrt{3}i \\ -1+\sqrt{3}i & \sqrt{2}+\sqrt{2}i \end{bmatrix}\begin{bmatrix} \sqrt{2}+\sqrt{2}i & -1-\sqrt{3}i \\ 1-\sqrt{3}i & \sqrt{2}-\sqrt{2}i \end{bmatrix} \\
&\quad + \frac{1}{16}\begin{bmatrix} \sqrt{2}+\sqrt{2}i & 1-\sqrt{3}i \\ -1-\sqrt{3}i & \sqrt{2}-\sqrt{2}i \end{bmatrix}\begin{bmatrix} \sqrt{2}-\sqrt{2}i & -1+\sqrt{3}i \\ 1+\sqrt{3}i & \sqrt{2}+\sqrt{2}i \end{bmatrix} \\
&= \frac{1}{16}\begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix} + \frac{1}{16}\begin{bmatrix} 8 & 0 \\ 0 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
\sum_y A_2(y \mid 1)A_2(y \mid 1)^\dagger &= \frac{1}{4}\begin{bmatrix} 1+i & 1-i \\ 1-i & 1+i \end{bmatrix}\begin{bmatrix} 1-i & 1+i \\ 1+i & 1-i \end{bmatrix} \\
&= \frac{1}{4}\begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\end{aligned}$$

So $\mathcal{M}_1$ and $\mathcal{M}_2$ are QSMs. Next we show that $p_{\mathcal{M}_1}(v \mid u) = p_{\mathcal{M}_2}(v \mid u)$ for all words u, v with $|u| = |v| \le 2$. Because of

$$\begin{aligned}
A_2(a \mid 0)^\dagger A_2(b \mid 0)^\dagger &= \frac{1}{4}\begin{bmatrix} \sqrt{2}+\sqrt{2}i & -1-\sqrt{3}i \\ 1-\sqrt{3}i & \sqrt{2}-\sqrt{2}i \end{bmatrix} \\
&\quad \times \frac{1}{4}\begin{bmatrix} \sqrt{2}-\sqrt{2}i & -1+\sqrt{3}i \\ 1+\sqrt{3}i & \sqrt{2}+\sqrt{2}i \end{bmatrix} \\
&= \frac{1}{8}\begin{bmatrix} 3-\sqrt{3}i & -\sqrt{2}(1+i) \\ \sqrt{2}(1-i) & 3+\sqrt{3}i \end{bmatrix}, \\
A_2(b \mid 0)^\dagger A_2(a \mid 0)^\dagger &= (A_2(a \mid 0)^\dagger A_2(b \mid 0)^\dagger)^* \\
&= \frac{1}{8}\begin{bmatrix} 3+\sqrt{3}i & -\sqrt{2}(1-i) \\ \sqrt{2}(1+i) & 3-\sqrt{3}i \end{bmatrix}, \\
A_2(a \mid 0)^\dagger A_2(a \mid 0)^\dagger &= \frac{1}{8}\begin{bmatrix} 2-2i & -\sqrt{2}(1+\sqrt{3}i) \\ \sqrt{2}(1-\sqrt{3}i) & -2-2i \end{bmatrix}, \\
A_2(b \mid 0)^\dagger A_2(b \mid 0)^\dagger &= (A_2(a \mid 0)^\dagger A_2(a \mid 0)^\dagger)^* \\
&= \frac{1}{8}\begin{bmatrix} 2+2i & -\sqrt{2}(1-\sqrt{3}i) \\ \sqrt{2}(1+\sqrt{3}i) & -2+2i \end{bmatrix}, \\
A_2(c \mid 1)^\dagger A_2(c \mid 1)^\dagger &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},
\end{aligned}$$

$$\begin{aligned}
&A_2(a \mid 0)^\dagger A_2(c \mid 1)^\dagger \\
&\quad = \frac{1}{8}\begin{bmatrix} (2\sqrt{2}-1+\sqrt{3})-(1+\sqrt{3})i & -(1+\sqrt{3})+(2\sqrt{2}-\sqrt{3}+1)i \\ (2\sqrt{2}+1-\sqrt{3})-(1+\sqrt{3})i & 1+\sqrt{3}-(2\sqrt{2}+\sqrt{3}-1)i \end{bmatrix}, \\
&A_2(c \mid 1)^\dagger A_2(a \mid 0)^\dagger \\
&\quad = \frac{1}{8}\begin{bmatrix} (2\sqrt{2}+1+\sqrt{3})+(1-\sqrt{3})i & 2\sqrt{2}-(1+\sqrt{3})+(1-\sqrt{3})i \\ (1-\sqrt{3})+(2\sqrt{2}-1-\sqrt{3})i & -1+\sqrt{3}-(2\sqrt{2}+1+\sqrt{3})i \end{bmatrix}, \\
&A_2(b \mid 0)^\dagger A_2(c \mid 1)^\dagger \\
&\quad = \frac{1}{8}\begin{bmatrix} -(1+\sqrt{3})-(2\sqrt{2}+1-\sqrt{3})i & -(1+\sqrt{3})+(2\sqrt{2}-1+\sqrt{3})i \\ (1+\sqrt{3})+(2\sqrt{2}-1+\sqrt{3})i & (2\sqrt{2}+1-\sqrt{3})+(1+\sqrt{3})i \end{bmatrix},
\end{aligned}$$

$$A_2(c \mid 1)^\dagger A_2(b \mid 0)^\dagger = \frac{1}{8}\begin{bmatrix} 1-\sqrt{3}-(2\sqrt{2}-1-\sqrt{3})i & -1+\sqrt{3}+(2\sqrt{2}+1+\sqrt{3})i \\ (2\sqrt{2}+1+\sqrt{3})+(-1+\sqrt{3})i & (2\sqrt{2}-1-\sqrt{3})+(-1+\sqrt{3})i \end{bmatrix},$$

we obtain that

$$p_{\mathcal{M}_2}(ab \mid 00) = \|A_2(b \mid 0)^\dagger A_2(a \mid 0)^\dagger \eta_1\|^2 = \frac{1}{4} = p_{\mathcal{M}_1}(ab \mid 00),$$

$$p_{\mathcal{M}_2}(ba \mid 00) = \|A_2(a \mid 0)^\dagger A_2(b \mid 0)^\dagger \eta_1\|^2 = \frac{1}{4} = p_{\mathcal{M}_1}(ba \mid 00),$$

$$p_{\mathcal{M}_2}(aa \mid 00) = \|A_2(a \mid 0)^\dagger A_2(a \mid 0)^\dagger \eta_1\|^2 = \frac{1}{4} = p_{\mathcal{M}_1}(aa \mid 00),$$

$$p_{\mathcal{M}_2}(bb \mid 00) = \|A_2(b \mid 0)^\dagger A_2(b \mid 0)^\dagger \eta_1\|^2 = \frac{1}{4} = p_{\mathcal{M}_1}(bb \mid 00),$$

$$p_{\mathcal{M}_2}(cc \mid 11) = \|A_2(c \mid 1)^\dagger A_2(c \mid 1)^\dagger \eta_1\|^2 = 1 = p_{\mathcal{M}_1}(cc \mid 11),$$

$$p_{\mathcal{M}_2}(ac \mid 01) = \|A_2(c \mid 1)^\dagger A_2(a \mid 0)^\dagger \eta_1\|^2 = \frac{1}{2} = p_{\mathcal{M}_1}(ac \mid 01),$$

$$p_{\mathcal{M}_2}(ca \mid 10) = \|A_2(a \mid 0)^\dagger A_2(c \mid 1)^\dagger \eta_1\|^2 = \frac{1}{2} = p_{\mathcal{M}_1}(ca \mid 10),$$

$$p_{\mathcal{M}_2}(bc \mid 01) = \|A_2(c \mid 1)^\dagger A_2(b \mid 0)^\dagger \eta_1\|^2 = \frac{1}{2} = p_{\mathcal{M}_1}(bc \mid 01),$$

$$p_{\mathcal{M}_2}(cb \mid 10) = \|A_2(b \mid 0)^\dagger A_2(c \mid 1)^\dagger \eta_1\|^2 = \frac{1}{2} = p_{\mathcal{M}_1}(cb \mid 10),$$

and $p_{\mathcal{M}_1}(v \mid u) = p_{\mathcal{M}_2}(v \mid u) = 0$ for the rest cases with $|u| = |v| = 2$. For $|u| = |v| = 1$, it is easy to check that $p_{\mathcal{M}_1}(v \mid u) = p_{\mathcal{M}_2}(v \mid u)$, and we omit the details. However, we shall show that $p_{\mathcal{M}_2}(cabb \mid 1000) \neq p_{\mathcal{M}_1}(cabb \mid 1000)$. Indeed, we have that

$$\begin{aligned} &A_2(b \mid 0)^\dagger A_2(b \mid 0)^\dagger A_2(c \mid 0)^\dagger A_2(c \mid 1)^\dagger \\ &= \frac{1}{8}\begin{bmatrix} 2+2i & -\sqrt{2}(1-\sqrt{3}i) \\ \sqrt{2}(1+\sqrt{3}i) & -2+2i \end{bmatrix} \\ &\quad \times \frac{1}{8}\begin{bmatrix} (2\sqrt{2}-1+\sqrt{3})-(1+\sqrt{3})i & -(1+\sqrt{3})+(2\sqrt{2}-\sqrt{3}+1)i \\ (2\sqrt{2}+1-\sqrt{3})-(1+\sqrt{3})i & 1+\sqrt{3}-(2\sqrt{2}+\sqrt{3}-1)i \end{bmatrix} \\ &= \frac{1}{32}\begin{bmatrix} 3\sqrt{2}+2\sqrt{3}+\sqrt{6}-2 & 2\sqrt{3}-\sqrt{6}-\sqrt{2}-2 \\ \sqrt{6}+2\sqrt{3}-\sqrt{2}+2 & 3\sqrt{2}-2\sqrt{3}-\sqrt{6}-2 \end{bmatrix} \end{aligned}$$

$$+\frac{1}{32}\begin{bmatrix}(\sqrt{6}+2\sqrt{3}+\sqrt{2}-2)i & (3\sqrt{2}-2\sqrt{3}+\sqrt{6}+2)i \\ (3\sqrt{2}+2\sqrt{3}-\sqrt{6}+2)i & (2\sqrt{3}-\sqrt{6}+\sqrt{2}+2)i\end{bmatrix}.$$

So we obtain that

$$\begin{aligned}
&p_{\mathcal{M}_2}(cabb \mid 1000) \\
&\quad= \|A_2(b \mid 0)^\dagger A_2(b \mid 0)^\dagger A_2(c \mid 0)^\dagger A_2(c \mid 1)^\dagger \eta_1\|^2 \\
&\quad= \left\| \frac{1}{32}\begin{pmatrix} 3\sqrt{2}+2\sqrt{3}+\sqrt{6}-2+(\sqrt{6}+2\sqrt{3}+\sqrt{2}-2)i \\ \sqrt{6}+2\sqrt{3}-\sqrt{2}+2+(3\sqrt{2}+2\sqrt{3}-\sqrt{6}+2)i \end{pmatrix} \right\|^2 \\
&\quad= \frac{8+\sqrt{6}+\sqrt{2}}{64},
\end{aligned}$$

but

$$p_{\mathcal{M}_1}(cabb \mid 1000) = \left| \frac{\sqrt{2}}{2} \times \frac{\sqrt{2}}{2} \times \frac{\sqrt{2}}{2} \times 1 \right|^2 = \frac{1}{8}.$$

Therefore, the answer for the problem proposed by S. Gudder is negative.

Certainly, there is positive example as follows.

Example 2. Let $I = \{0\}$ and $O = \{a, b\}$. Suppose that

$$\mathcal{M}_1 = (\{s\}, \eta, I, O, \{A(y \mid x) : y \in O, x \in I\})$$

and

$$\mathcal{M}_2 = (\{s_1, s_2\}, \eta_1, I, O, \{A(y \mid x) : y \in O, x \in I\})$$

where $\eta = 1$, $A_1(a \mid 0) = 1/2$, $A_1(b \mid 0) = \sqrt{3}/2$; and

$$A_2(a \mid 0) = \frac{\sqrt{2}}{4}\begin{bmatrix}1 & 1 \\ 1 & -1\end{bmatrix}, \quad A_2(b \mid 0) = \frac{\sqrt{6}}{4}\begin{bmatrix}1 & 1 \\ 1 & -1\end{bmatrix}.$$

Then one can check that $p_{\mathcal{M}_1}(v \mid u) = p_{\mathcal{M}_2}(v \mid u)$ for all input–output pair (u, v). That is to say, $\mathcal{M}_1$ and $\mathcal{M}_2$ are equivalent.

However, we do not know yet what are the sufficient and necessary conditions for the equivalence between two SQMs or QSMs.

ACKNOWLEDGMENTS

This work was supported by the National Foundation for Distinguished Young Scholars (Grant No. 69725004), the National Key Project for Basic Research

(Grant No. 1998030509), the National Foundation of Natural Sciences (Grant No. 69823001) of China.

REFERENCES

Bertoni, A. and Carpentieri, M. (2001). Analogies and differences between quantum and stochastic automata. *Theoretical Computer Sciences* **262**, 69–81.

Gruska, J. (1999). *Quantum Computing*, McGraw-Hill, London.

Gudder, S. (1999). Quantum automata: An overview. *International Journal of Theoretical Physics* **38**, 2261–2282.

Gudder, S. (2000). Quantum computers. *International Journal of Theoretical Physics* **39**, 2151–2177.

Moore, C. and Crutchfield, P. (2000). Quantum automata and quantum grammars. *Theoretical Computer Sciences* **237**, 275–306.

Paz, A. (1971). *Introduction to Probabilistic Automata*, Academic Press, New York.