Quantum Hidden Markov Models based on Transition Operation Matrices

Michał Cholewa Piotr Gawron Przemysław Głomb
30 March 2015

Abstract

In this work, we extend the Quantum Markov chains proposition [S. Gudder. Quantum Markov chains. J. Math. Phys., 49(7), 2008] to propose Quantum Hidden Markov Models (QHMMs). For that, we use the notions of Transition Operation Matrices (TOM) and Vector States, which are an extension of classical stochastic matrices and probability distributions. Our main result is the Mealy QHMM formulation and proofs of algorithms needed for application of this model: Forward for general case and Viterbi for a restricted class of QHMMs.

1 Introduction

The classical Hidden Markov Model was introduced as a method of modelling signal sources observed in noise. It is now extensively used, e.g. in speech and gesture recognition or biological sequence analysis. Their popularity is a result of their rich structure, which is able to model wide variety of problems and effective algorithms that facilitate their application. The HMM is characterized by three fundamental problems [10]: given a sequence of symbols of length $T$, $O = (o_1, o_2, \ldots, o_T)$, and a HMM parametrized by $\lambda$,

1. Compute the $P(O|\lambda)$, probability of the sequence $O$ given a HMM with parameters $\lambda$.

2. Select the sequence of states $N_T = (n_0, n_1, \ldots, n_T)$ that maximizes the probability $P(O|\lambda, N_T)$; in other words the most likely state sequence in HMM $\lambda$ that produces $O$.

3. Adjust the model parameters $\lambda$ to maximize $P(O|\lambda)$. 

The above problems are solved, respectively, the Forward, Viterbi and Baum-Welch algorithms. The effectiveness of those algorithms is based on optimized procedure of computation, which uses a ‘trellis’: a two dimensional lattice structure of observations and states. This formulation is based on the Markov property of model evolution and reduces the complexity from exponential $O(TN^T)$ to polynomial $O(N^2T)$, where $T$ is the number of observations and $N$ the number of model states [10].

Depending on the formulation, there are two definitions of a Hidden Markov Model: Mealy and Moore. In the former, the probability of next state $n_{t+1}$ depends both on the current state $n_t$ and the generated output symbol $o_t$. In the latter, the symbol generation is independent from state switch, i.e. $P(S(t + 1) = S_i | o_t = o, S(t) = S_j) = P(S(t + 1) = S_i | S(t) = S_j)$. While the expressive power of Moore and Mealy models is the same, i.e. process can be realized with Moore model if and only if it is realizable by Mealy model, the minimal model order for the realization is lower in Mealy models [15]. In this work we focus only on Mealy models.

1.1 Related work

In this work we follow the scheme proposed by Gudder in [5] and extend it in order to construct Quantum Hidden Markov Models (QHMMs). Gudder introduced the notions of Transition Operation Matrices and Vector States, which give an elegant extension of classical stochastic matrices and probability distributions. These notions allows to define Markov processes that exhibit both classical and quantum behaviour.

Below we review two areas of research most closely related to our work: open quantum walks and Hidden quantum Markov models.

Open quantum walks In recent years a new sub-field of quantum walks has emerged. In series of papers [9, 3, 12, 13, 2, 14] Attal, Sabot, Sinasky, and Petrucione introduced the notion of Open Quantum Walks. Theorems for limit distributions of open quantum random walks were provided in [6]. In [1] the average position and the symmetry of distribution in the $SU(2)$ Open Quantum Walk is studied.

The notion of open quantum walks is generalised to quantum operations of any rank in [8] and analysed in [11]. In first of these two papers the notion of mean first passage time for a generalised quantum walk is introduced and studied for class of walks on Apollonian networks. In the second paper a central limit theorem for reducible and irreducible open quantum walks is provided.
Quantum hidden Markov models  Hidden quantum Markov models were introduced in [7]. The construction provided there by the authors is different from ours. In their work the hidden quantum Markov model consists of a set of quantum operations associated with emission symbols. The evolution of the system is governed by the application of quantum operations on a quantum state. The sequence of emitted symbols defines the sequence of quantum operations being applied on the initial state of the hidden quantum Markov model.

1.2 Our contribution

In this work we propose a Quantum Hidden Markov model formulation using the notions of Transition Operation Matrices. We focus on Mealy models, for which we derive first the Forward algorithm in general case, then the Vitterbi algorithm, for models restricted to those with sub-TOMs elements are trace-monotonicity preserving quantum operations.

The paper is organised as follows: in Section 2 we collect the basic mathematical objects and their properties, in Section 3 we define Quantum Hidden Markov Models and provide Forward na Viterbi algorithm for these models, and finally in Section 4 we conclude.

2 Transition Operation Matrices

In what follows we provide basic elements of quantum information theory and summarize definitions and properties of objects introduced by Gudder in [5].

2.1 Quantum theory

Let $\mathcal{H}$ be a complex finite Hilbert space and $\mathcal{L}(\mathcal{H})$ be the set of linear operators on $\mathcal{H}$. We also denote set of positive operators on $\mathcal{H}$ as $\mathcal{P}^+(\mathcal{H})$ and the set of positive semi-definite operators on $\mathcal{H}$ as $\mathcal{P}(\mathcal{H})$.

Definition 1. A linear operator $\rho \in \mathcal{P}(\mathcal{H})$ is called a quantum state if $\text{tr} \rho = 1$. Set of quantum states is denoted by $\Omega(\mathcal{H})$.

Definition 2. A linear operator $\rho \in \mathcal{P}(\mathcal{H})$ is called sub-normalized [4] quantum state if $\text{tr} \rho \leq 1$. Set of sub-normalized quantum states is denoted by $\Omega_\leq(\mathcal{H})$.

Definition 3. A linear map $\Phi \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2))$ is called positive map if, for every $\rho \in \mathcal{P}(\mathcal{H}_1)$, $\Phi(\rho) \in \mathcal{P}(\mathcal{H}_2)$.  

3
Definition 4. A linear map $\Phi \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2))$ is called completely positive (CP) if for any complex Hilbert space $\mathcal{H}_3$, the map $\Phi \otimes 1 \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1 \otimes \mathcal{H}_3), \mathcal{L}(\mathcal{H}_2 \otimes \mathcal{H}_3))$ is positive.

Definition 5. A linear map $\Phi \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2))$ is called trace preserving if $\text{tr}(\Phi(\rho)) = \text{tr}\rho$ for every $\rho \in \mathcal{L}(\mathcal{H}_1)$.

Definition 6. A linear map $\Phi \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2))$ is called trace non-increasing if $\text{tr}(\Phi(\rho)) \leq \text{tr}\rho = 1$ for every quantum state $\rho \in \Omega(\mathcal{H}_1)$.

Definition 7. A linear map $\Phi \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2))$ is called a quantum operation if it is completely positive and trace non-increasing.

Definition 8. A linear map $\Phi \in \mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2))$ is called a quantum channel if it is completely positive and trace preserving.

Definition 9. By quantum measurement we call a mapping from a finite set $\Theta$ of measurement outcomes to subset of set of measurement operators $\mu : \Theta \to \mathcal{P}(\mathcal{H})$ such that $\sum_{a \in \Theta} \mu(a) = 1$.

We will call a measurement $\mu$ trivial if $|\Theta| = 1$.

With each measurement $\mu$ we associate non-negative functional $p : \Theta \to \mathbb{R}_+ \cup \{0\}$ which maps measurement outcome $a$ for given positive operator $\rho$ and measurement $\mu$ to non-negative real number in the following way $p(a) = \text{tr} \mu(a) \rho$. If $\text{tr}\rho = 1$, for given $\rho$ and $\mu$ the value of $p$ can be interpreted as probability of obtaining measurement outcome $a$ in quantum state $\rho$.

If $\rho$ is a sub-normalized state the trivial measurement $\mu : a_e \mapsto 1$ measures the probability $p(a_e) = \text{tr}\rho$ that the state $\rho$ exists. One should note that this kind of measurement commutes with any other measurement and thus does not disturb the quantum system.

2.2 Transition Operation Matrices

The core object of the Gudder’s scheme is Transition Operation Matrix (TOM) which generalizes the idea of stochastic matrix.

Definition 10 (Transition Operation Matrix). Let $\mathcal{H}_1$, $\mathcal{H}_2$ denote two finite dimensional Hilbert spaces and $\Omega(\mathcal{H}_1)$, $\Omega(\mathcal{H}_2)$ denote sets of quantum states acting on those spaces respectively.

A TOM is a matrix in form $\mathcal{E} = \{E_{ij}\}_{i,j=1}^{M,N}$, where $E_{ij}$ is completely positive map in $\mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2))$ such that for every $j$ and $\rho \in \Omega(\mathcal{H}_1)$ $\sum_i E_{ij}(\rho) \in \Omega(\mathcal{H}_2)$.
Alternatively one can say that $E = \{E_{ij}\}_{i,j=1}^{M,N}$ is a TOM if and only if for every column $j$ $\sum_i E_{ij}$ is a quantum channel (Completely Positive Trace Preserving map). A simple implication of this definition is that each $E_{ij}$ is CP-TNI mapping.

Note that in this definition TOM has four parameters:

- size of matrix “output” (number of rows) — $M$,
- size of matrix “input” (number of columns) — $N$,
- “input” Hilbert space — $\mathcal{H}_1$,
- “output” Hilbert space — $\mathcal{H}_2$.

The set of TOMs we will denote as $\Gamma^{M,N}(\mathcal{H}_1, \mathcal{H}_2)$.

**Definition 11 (Sub Transition Operation Matrix).** Let $\mathcal{H}_1$, $\mathcal{H}_2$ denote two finite dimensional Hilbert spaces, $\Omega(\mathcal{H}_1)$ denotes set of quantum states acting on the first space and $\Omega^\leq(\mathcal{H}_2)$ denotes set of sub-normalised quantum states acting on the second Hilbert space.

A sub-TOM is a matrix in form $E = \{E_{ij}\}_{i,j=1}^{M,N}$, where $E_{ij}$ is completely positive map in $\mathcal{L}(\mathcal{L}(\mathcal{H}_1), \mathcal{L}(\mathcal{H}_2))$ such that for every $j$ and $\rho \in \Omega(\mathcal{H}_1)$. $\sum_i E_{ij}(\rho) \in \Omega^\leq(\mathcal{H}_2)$.

**Definition 12 (Quantum Markov chain).** Let a TOM $E = \{E_{ij}\}_{i,j=1}^{M,N}$ be given. Quantum Markov chain is a finite directed graph $G = (E, V)$ labelled by $E_{ij}$ for $e \in E$ and by zero operator for $e \notin E$.

**Definition 13 (Vector state).** Vector state is a column vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)^T$ such that $\alpha_i \in \Omega^\leq(\mathcal{H})$ are sub-normalized quantum states and $\sum_{i=1}^N \alpha_i \in \Omega(\mathcal{H})$. We will denote the set of vector states as $\Delta^N(\mathcal{H})$.

**Definition 14 (Subnomralized vector state).** Subnomralized vector state is a column vector $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)^T$ such that $\alpha_i \in \Omega^\leq(\mathcal{H})$ are sub-normalized quantum states and $\sum_{i=1}^N \alpha_i \in \Omega^\leq(\mathcal{H})$. $W$ will denote a set of sub-normalized vector states as $\Xi^N(\mathcal{H})$.

Applying TOM $E$ on a vector state $\alpha$ produces vector state $\beta = E(\alpha)$ where $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)^T$, $\alpha_i \in \Omega^\leq(\mathcal{H}_1)$, where $\beta = (\beta_1, \beta_2, \ldots, \beta_M)^T$, $\beta_i \in \Omega^\leq(\mathcal{H}_2)$, and $E \in \Gamma^{M,N}(\mathcal{H}_1, \mathcal{H}_2)$, and in the following way $\beta_i = \sum_{j=1}^N E_{ij}(\alpha_j)$.

**Remark 1.** Product of two TOMs is a TOM.

**Remark 2.** Product of two sub-TOMs is a sub-TOM.
Product of TOMs that have same parameters is associative \( \text{i.e.} (\mathcal{E} \mathcal{F}) \mathcal{G} = \mathcal{E}(\mathcal{F} \mathcal{G}) \) and \( (\mathcal{E} \mathcal{F})(\alpha) = \mathcal{E}(\mathcal{F}(\alpha)) \).

For any square TOM \( \mathcal{E} \in \Gamma^{M,M}(\mathcal{H}, \mathcal{H}) \) whose both input and output Hilbert spaces are equal one can define integer exponent. By \( \mathcal{E}^{(n)}(\alpha) \) we understand a TOM that is equivalent to \( n \)-fold application of TOM \( \mathcal{E} \). Vector state \( \mathcal{E}^{(n)}(\alpha) \) is computed from the following recursive relation \( \mathcal{E}^{(n)} = \mathcal{E} \mathcal{E}^{(n-1)}(\alpha) \) and \( \mathcal{E}^{(0)}(\alpha) = \alpha \).

### 3 Quantum Hidden Markov Model

In order to explain the idea of QHMM we can form following analogy. QHMM might be understood as a system consisting of a particle that can has an internal quantum state \( \rho \in \Omega \leq \) and it occupies a classical state \( S_i \). This particle hops from one classical state into another passing trough a quantum operation associated with a sub-TOM element \( \mathcal{P}^{V_k}_{S_j,S_i} \). With each transition a symbol \( V_k \) is emitted from the system.

We will now define the classical and quantum version of the Mealy Hidden Markov Model.

#### 3.1 Mealy HMM and QHMM

**Definition 15** (Mealy Hidden Markov Model). Let \( S \) and \( \mathcal{V} \) be set of states and an alphabet respectively. Mealy HMM is specified by tuple \( \lambda = (S, \mathcal{V}, \Pi, \pi) \), where:

- \( \pi \in [0, 1]^N \) is a stochastic vector representing initial states, where \( \pi_i \) is probability that initial states is \( S_i \);
- \( \Pi \) is a mapping \( \mathcal{V} \rightarrow \Psi \in \mathbb{R}^{N,N} \), where \( \Psi \) is sub-stochastic matrix, such that \( \Pi^\Sigma := \sum_i^M \Pi^V_i \in \mathbb{R}^{N,N} \) is stochastic matrix. The element \( \Pi^o_{i,j} \) is \( p(S(t+1) = S_j, o_t = o|S(t) = S_i) \), that is, probability of going from state \( i \) to \( j \) while generating the output \( o \).

Denote set of all finite strings as \( O \) and let \( o = o_1o_2, \ldots o_T \in O \). Let then \( P : O \rightarrow [0, 1] \) be string probabilities, defined as \( P(o) = p(o(1) = o_1, o(2) = o_2, \ldots, o(T) = o_T) \). Of course function \( P \) satisfies \( P(O) = 1 \) and \( \sum_{oA \in O} P(oA) = P(o) \), where \( oA \) is concatenation of strings \( o \) and \( oA \).

The string probabilities generated by Mealy HMM \( (S, \mathcal{V}, \Pi, \pi) \) are given by

\[
P(o) = \sum_{i=1}^N \alpha_i, \text{ where } \alpha = \Pi^{o_T}\Pi^{o_{T-1}} \ldots \Pi^{o_1} \pi.
\]
Definition 16 (Mealy Quantum Hidden Markov Model). Let $\mathcal{S}$ and $\mathcal{V}$ be set of states and an alphabet respectively. Mealy QHMM is specified by tuple $\lambda = (\mathcal{S}, \mathcal{V}, \mathcal{P}, \pi)$, where:

- $\pi \in \Delta^N(\mathcal{H})$ is an initial vector state;
- $\mathcal{P}$ is a mapping $\mathcal{V} \to \Xi^{N,N}(\mathcal{H}, \mathcal{H})$ such that $\mathcal{P}^S := \sum_{i=1}^M \mathcal{P}_i^{V_i} \in \Gamma^{N,N}(\mathcal{H}, \mathcal{H})$ is a TOM, with $\mathcal{P}_i^{V_i}$ being value of $\mathcal{P}$ for $V_i$.

A graphical representation of three-state two-symbol Mealy QHMM is presented in Fig. 1.

3.1.1 Properties of QHMMs

Remark 3. For $\dim \mathcal{H} = 1$ QHMM reduces to classical HMM. In this case TOMs reduce to stochastic matrices, sub-TOMs to sub-stochastic matrices, the vector states to probability vectors, sub-vector states to sub-normalized probability vectors.

3.1.2 Forward algorithm for Mealy QHMM

With each Mealy QHMM we can associate a mapping $\varphi : \mathcal{V}^* \to \Omega_{\leq}(\mathcal{H})$.

Given a sequence $O = (o_1, o_2, \ldots, o_T)$ and Mealy QHMM $\lambda$ one can compute resulting sub-normalized quantum state $\rho_{O|\lambda}$.

Let us consider sub-normalized vector states $\alpha_T \in \Xi^M(\mathcal{H})$ such that

$$\alpha_T = \mathcal{P}^{o_T} \cdots \mathcal{P}^{o_2} \mathcal{P}^{o_1}(\pi),$$

(1)

then $\rho_{O|\lambda} := \varphi(O) = \sum_{i=1}^M \alpha_{T,i}$. Equation (1) we call the Forward algorithm for QHMMs. It allows us to formulate the following theorem:

Theorem 1. Let $\mathcal{V}^*$ be the set of sequences over alphabet $\mathcal{V}$. Let $\mathcal{V}^T \subset \mathcal{V}^*$ be a set of all sequences of length $T$. For any QHMM $\lambda$ we have $\sum_{O \in \mathcal{V}^T} \rho_{O|\lambda} \in \Omega(\mathcal{H})$.

Proof:

We will proceed with induction on $T$:

(*) For $T = 1$

$$\sum_{i=1}^M \sum_{o \in \mathcal{V}} \mathcal{P}^o(\pi)$$

(2)

But $\mathcal{P}_1^* = \sum_{o \in \mathcal{V}} \mathcal{P}^o$ is TOM, therefore $\sum_{i=1}^M \mathcal{P}(\pi) \in \Omega(\mathcal{H})$. 

7
The alphabet consists of two symbols $V_1, V_2$. Figure 1: Graphical representation of three-state Mealy QHMM $\lambda$, whose alphabet consists of two symbols $V_1, V_2$.

\[ \lambda = \left\{ \{S_1, S_2, S_3\}, \{V_1, V_2\}, \{V_1 \rightarrow P_{V_1}V_2 \rightarrow P_{V_2}\}, \pi_{S_1}, \pi_{S_2}, \pi_{S_3} \right\} \]

\[ P_{V_1} = \begin{bmatrix} P_{V_1}^{S_1S_1} & P_{V_1}^{V_1V_1} & P_{V_1}^{V_1S_1} \\ P_{V_1}^{S_2S_1} & P_{V_1}^{V_1V_2} & P_{V_1}^{V_1S_2} \\ P_{V_1}^{S_3S_1} & P_{V_1}^{V_1V_3} & P_{V_1}^{V_1S_3} \end{bmatrix}, \quad P_{V_2} = \begin{bmatrix} P_{V_2}^{S_1S_1} & P_{V_2}^{V_2V_1} & P_{V_2}^{V_2S_1} \\ P_{V_2}^{S_2S_1} & P_{V_2}^{V_2V_2} & P_{V_2}^{V_2S_2} \\ P_{V_2}^{S_3S_1} & P_{V_2}^{V_2V_3} & P_{V_2}^{V_2S_3} \end{bmatrix} \]

\[ \text{(**)} \quad \text{For } T = n+1 \]

\[ \sum_{i=1}^{M} \sum_{o \in V^{n+1}} P^{o_{n+1}} \cdots P^{o_{2}} P^{o_{1}}(\pi) = \sum_{i=1}^{M} \sum_{o \in V} P^{o} \sum_{o \in V^{n}} P^{o_{n+1}} \cdots P^{o_{2}} P^{o_{1}}(\pi) \]

From (\text{**)}, we have $P_{n}^{*} = \sum_{o_{1}, \ldots, o_{n} \in V^{n}} P^{o_{n}} \cdots P^{o_{2}} P^{o_{1}}$ is a TOM, so

\[ \sum_{i=1}^{M} \sum_{o \in V^{n+1}} P^{o_{n+1}} \cdots P^{o_{2}} P^{o_{1}}(\pi) = \sum_{i=1}^{M} \sum_{o \in V} P^{o} P_{n}^{*}(\pi) \]
But $\sum_{o \in V} P^o = P^1_1$, so we have

$$\sum_{i=1}^{M} \sum_{o \in V^{n+1}} P^{o_{n+1}} \ldots P^{o_2} P^{o_1}(\pi) = \sum_{i=1}^{M} P^*_i P^*_n(\pi). \quad (3)$$

Now from (1) $X = P^*_1 P^*_n$ is a TOM, therefore

$$\sum_{i=1}^{M} X(\pi) \in \Omega(\mathcal{H}). \quad (4)$$

Note that the result of the algorithm application is the state $\rho_{O|\lambda}$. As in classical case the result of the Forward algorithm is the probability, a straightforward extension would be to compute $\text{tr} \rho_{O|\lambda}$. That way, however, we would lose the quantum information contained in that state.

### 3.1.3 Viterbi algorithm for Mealy QHMM

We are given a QHMM $\lambda$ with set of states $S = \{S_1, S_2, \ldots, S_{|S|}\}$ and an alphabet of symbols $V = \{V_1, V_2, \ldots, V_{|V|}\}$. We denote $P^k_{ij} = P^V_{S_i S_j}$.

We have a sequence of length $T$, $O = (o_1, o_2, \ldots, o_T)$, of symbols from alphabet $V$, $o_i \in V$.

A Mealy QHMM emits symbols on transition from one state to the next. For our sequence $O$ we index corresponding QHMM states by $n_i$, i.e. $n_0$ is the initial state (before the emission of the first symbol), and $n_i, i \geq 1$ is the state after emission of the symbol $o_i$, $n_i \in S$.

We denote the set of partial sequences of state indexes as $N_k = \{(n_0, n_1, \ldots, n_k) : n_j \in S, j = 0, 1, \ldots, k\}$, where $k \leq T$. A set beginning with $n_0$ and ending after $k$ steps with $S_i$ we denote $N^S_k = \{(n_0, n_1, \ldots, n_{k-1}, n_k = S_i) : n_j \in S, j = 0, 1, \ldots, k\} \subset N_k$.

**Theorem 2.** Let $O$ be a given sequence of emissions from $V$. Let $\lambda = (S, V, P, \pi)$ be a Mealy QHMM satisfying

$$\forall_{n_i, n_j \in S, o \in O} \forall_{\alpha, \beta \in \Omega(\mathcal{H})} \text{tr} \alpha > \text{tr} \beta \implies \text{tr} P^o_{n_i, n_j}(\alpha) > \text{tr} P^o_{n_i, n_j}(\beta) \quad (5)$$

i.e. all sub-TOMs elements are trace-monotonicity preserving quantum operations.

We define $w \in N^S_k$ to be a sequence of states ending with $S_i$. A sub-normalized state associated with $w$ and sequence $O$ is $B_w \in \Omega(\mathcal{H})$ defined
as \( B_w = \mathcal{P}^0_{n_k,n_{k-1}} \mathcal{P}^0_{n_{k-1},n_{k-2}} \cdots \mathcal{P}^0_{n_1,n_0}(\pi_{n_0}) \). The sub-normalized state that maximizes trace over set of all \( B_w \)s with \( w \in N^S_k \) is

\[
A_k,S_i = \text{argmax} \{ B_w : w \in N^S_k \} \tag{6}
\]

Then the following holds

\[
\text{tr} A_k,S_i = \max_{n_{k-1} \in S} \text{tr} \mathcal{P}^0_{n_k,n_{k-1}} (A_{k-1,n_{k-1}}). \tag{7}
\]

Proof:

Let us denote

\[
w^*_k,S_i = (n^*_0, \ldots, n^*_k, n_k = S_i) \in N^S_k \tag{8}
\]

as the sequence of states maximizing trace of \( B_w \), so that

\[
\text{tr} A_k,S_i = \text{tr} B_{w^*_k,S_i}. \tag{9}
\]

We now have

\[
\text{tr} A_k,S_i = \max_{w \in N^S_k} \text{tr} B_w = \max_{n_0, \ldots, n_{k-1}, n_k = S_i} \text{tr} \mathcal{P}^0_{n_k,n_{k-1}} \mathcal{P}^0_{n_{k-1},n_{k-2}} \cdots \mathcal{P}^0_{n_1,n_0}(\pi_{n_0}) \tag{10}
\]

Obviously

\[
\text{tr} A_k,S_i = \text{tr} \mathcal{P}^0_{n^*_k,S_i} \mathcal{P}^0_{n^*_{k-1},n^*_k} \cdots \mathcal{P}^0_{n^*_1,n^*_0}(\pi_{n^*_0}) \tag{11}
\]

We will now prove that for \( n^*_k = S_i \)

\[
w^*_k,S_i = (n^*_0, \ldots, n^*_k, n_k) \implies w^*_{k-1,n^*_k} = (n^*_0, \ldots, n^*_{k-1}) \tag{12}
\]

Let us assume that it is not true. That would mean, that

\[
w^*_{k-1,n^*_k} = (l^*_0, \ldots, l^*_{k-2}, n^*_{k-1}) \neq (n^*_0, \ldots, n^*_{k-2}, n^*_{k-1}).
\]

Of course

\[
\text{tr} B_{(l^*_0, \ldots, l^*_{k-2}, n^*_{k-1})} > \text{tr} B_{(n^*_0, \ldots, n^*_{k-2}, n^*_{k-1})}
\]

From this, and 5, we have

\[
\forall n_k, S,y_k \in \{1,2,\ldots,|V|\} \text{ tr} \mathcal{P}^0_{n_k,n^*_{k-1}} (B_{(l^*_0, \ldots, l^*_{k-2}, n^*_{k-1})}) > \text{tr} \mathcal{P}^0_{n_k,n^*_{k-1}} (B_{(n^*_0, \ldots, n^*_{k-2}, n^*_{k-1})}),
\]

that leads to

\[
w^*_k,S_i \neq (n^*_0, \ldots, n^*_{k-1}, n^*_k)
\]
which is a contradiction. That proves that implication (12) holds.

Then, for \( n_k = S_i \)

\[
\text{tr} A_{k,S_i} = \text{tr} B_{w^*_{k,S_i}} = \text{tr} \mathcal{P}^{n_k}_{n_k, n_k} (B_{w^*_{k,n_k}}) = \text{tr} \mathcal{P}^{n_k}_{n_k, n_k} (A_{k-1,n_k})
\]

\[
= \max_{n_{k-1} \in S} \text{tr} \mathcal{P}^{n_k}_{n_k} (A_{k-1,n_k}) .
\]

(13)

Remark 4. It can be easily seen that (5) holds iff quantum operation \( \mathcal{P}^y_{n_j,n_i} \) is of form \( c \cdot \Phi \), where \( c \in [0,1) \) and \( \Phi \) is a quantum channel (CP-TP map).

From Theorem 2 we immediately derive the Viterbi algorithm for Mealy QHMMs conditioned with (5) that computes most likely sequence of states for a given sequence \( O \).

Initialization:

\[ A_{0,S_i} = \pi_{S_i} \]

(14)

Computation for step number \( k \):

\[
\forall S_i \in S, k \in \{1, \ldots, T\} \ n^*_{k-1}(S_i) = \arg \max_{n_{k-1} \in S} \text{tr} \mathcal{P}^{n_k}_{n_k, n_{k-1}} (A_{k-1,n_{k-1}}) \]

(15)

\[
\forall S_i \in S, k \in \{1, \ldots, T\} \ A_{k,S_i} = \mathcal{P}^{n_k}_{n_k} (A_{k-1,n_k}(S_i)) ,
\]

(16)

Termination:

\[ n^*_T = \arg \max_{S_i \in S} \text{tr} A_{T,S_i} \]

(17)

The most probable state sequence is \((n_0^*, \ldots, n_T^*)\), with resulting state being \( A_{T,n_T^*} \), with probability given by \( \text{tr} A_{T,n_T^*} \).

In case when (5) does not apply, one can resort to exhaustive search over all state sequences. As a result of the multitude of possible quantum operations the behaviour of the Quantum Hidden Markov Model can be markedly different than its classical counterpart. This is similar to the relation of quantum and classical Markov models [8].

4 Conclusions

We have introduced a new model of Quantum Hidden Markov Models based on the notions of transition operation matrices and vector states. We have shown that for a subclass of QHMMs and emission sequences the modified Viterbi algorithm can be used to calculate the most likely sequence of internal
states that lead to a given emission sequence. We have also proposed a formulation of the Forward algorithm that is applicable for general QHMMs. Because of the fact that the structure of Quantum Hidden Markov Models is more complicated than their classical counterparts, the most likely sequence of states leading to a given emissions sequence has to be calculated using extensive search.

We believe that proposed model can find applications in modelling systems that posses both quantum and classical features.

Acknowledgements

We would like to thank Z. Puchała and Ł. Pawela for fruitful discussions about subject of this paper. Research of Piotr Gawron was supported by the Grant N N516 481840 financed by Polish National Science Centre. Research of Przemysław Głomb was supported by Polish National Science Centre grant number DEC-2011/03/D/ST6/03753. Research of Michał Cholewa was supported by the means provided by National Science Centre, based on decision no DEC-2012/07/N/ST6/03656.

References

[1] Clement Ampadu. Averaging in su (2) open quantum random walk. *Chinese Physics B*, 23(3):030302, 2014.

[2] Stephane Attal, Francesco Petruccione, Christophe Sabot, and Ilya Sinayskiy. Open quantum random walks. *Journal of Statistical Physics*, 147(4):832–852, 2012.

[3] Stephane Attal, Francesco Petruccione, and Ilya Sinayskiy. Open quantum walks on graphs. *Physics Letters A*, 376(18):1545–1548, 2012.

[4] Valerio Cappellini, Hans-Jürgen Sommers, and Karol Życzkowski. Subnormalized states and trace-nonincreasing maps. *Journal of Mathematical Physics*, 48(5):052110–052110–19, 2007.

[5] Stanley Gudder. Quantum Markov chains. *Journal of Mathematical Physics*, 49(7):072105, 2008.

[6] Norio Konno and Hyun Jae Yoo. Limit theorems for open quantum random walks. *Journal of Statistical Physics*, 150(2):299–319, 2013.
[7] Alex Monras, Almut Beige, and Karoline Wiesner. Hidden Quantum Markov Models and non-adaptive read-out of many-body states. *Applied Mathematical and Computational Sciences*, 3:93, 2011.

[8] Łukasz Pawela, Piotr Gawron, Jarosław Adam Miszczak, and Przemysław Sadowski. Generalized open quantum walks on apollonian networks. *arXiv preprint arXiv:1407.1184*, 2014.

[9] Francesco Petruccione and Ilya Sinayskiy. Open quantum random walks and the open quantum pascal triangle. In *International Conference on Quantum Information*, page QMH2. Optical Society of America, 2011.

[10] Lawrence Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.

[11] Przemysław Sadowski and Łukasz Pawela. Central limit theorem for reducible and irreducible open quantum walks. *arXiv preprint arXiv:1412.5086*, 2014.

[12] Ilya Sinayskiy and Francesco Petruccione. Efficiency of open quantum walk implementation of dissipative quantum computing algorithms. *Quantum Information Processing*, 11(5):1301–1309, 2012.

[13] Ilya Sinayskiy and Francesco Petruccione. Properties of open quantum walks on Z. *Physica Scripta*, 2012(T151):014077, 2012.

[14] Ilya Sinayskiy and Francesco Petruccione. Open quantum walks: a short introduction. In *Journal of Physics: Conference Series*, volume 442, page 012003. IOP Publishing, 2013.

[15] Bart Vanluyten, Jan C Willems, and Bart De Moor. Equivalence of state representations for hidden markov models. *Systems & Control Letters*, 57(5):410–419, 2008.