A Quantum-Statistical Approach Toward Robot Learning by Demonstration

Sotirios P. Chatzis, Dimitrios Korkinof, and Yiannis Demiris, *Senior Member, IEEE*

***Abstract*—Statistical machine learning approaches have been at the epicenter of the ongoing research work in the field of robot learning by demonstration over the past few years. One of the most successful methodologies used for this purpose is a Gaussian mixture regression (GMR). In this paper, we propose an extension of GMR-based learning by demonstration models to incorporate concepts from the field of quantum mechanics. Indeed, conventional GMR models are formulated under the notion that all the observed data points can be assigned to a distinct number of model states (mixture components). In this paper, we reformulate GMR models, introducing some quantum states constructed by superposing conventional GMR states by means of linear combinations. The so-obtained quantum statistics-inspired mixture regression algorithm is subsequently applied to obtain a novel robot learning by demonstration methodology, offering a significantly increased quality of regenerated trajectories for computational costs comparable with currently state-of-the-art trajectory-based robot learning by demonstration approaches. We experimentally demonstrate the efficacy of the proposed approach.**

***Index Terms*—Quantum statistics, robot learning by demonstration, statistical machine learning.**

I. Introduction

ROBOT learning by demonstration has been an active research topic in the field of robotics during the past few years, encompassing methods by which a robot can learn new skills by simple observation of a human teacher; in the same way, humans learn new skills by imitation [1]–[3]. As a result, learning by demonstration alleviates the need of programming a robot how to perform a task, which can be rather tedious and expensive, and can speed up reinforcement learning techniques, since it significantly reduces the search space of the learning algorithm, while, by making robots more user-friendly, it increases the appeal of applying robots to real-life environments.

To effect these goals, robot learning by demonstration combines methods from diverse research areas, such as machine learning, computer vision, and human–robot interaction. Current approaches toward learning by demonstration can be divided into two broad categories: low-level *trajectory encoding* and high-level decomposition of a skill into action units, which is usually referred to as*symbolic encoding*. In this paper, we focus on the former approach toward learning by demonstration. The first step toward the implementation of a successful trajectory-based learning by demonstration algorithm is to choose the right variables to encode the movements under consideration; usually, representations in the joint space, task space, or torque space are considered [4]. Subsequently, coming up with methods capable of successfully extracting and modeling the underlying patterns in the demonstrated motions comes to the fore, being the crucial factor that eventually determines the effectiveness of a developed learning by demonstration algorithm.

Several researchers have considered the application of statistical machine learning algorithms as the effective means to facilitate extraction of the trajectory patterns underlying a set of demonstrated skills. Indeed, one of the most popular trends of work in the field of trajectory-based robot learning by demonstration consists in the investigation of the utility of a Gaussian mixture regression (GMR) [5]. GMR has been shown to be very successful in encoding demonstrations, extracting their underlying constraints, and reproducing smooth generalized motor trajectories through a Gaussian mixture model (GMM) trained by means of the expectation–maximization (EM) algorithm [6], while imposing considerably low-computational costs [7]–[10]. GMR-based approaches toward learning by demonstration rely on the postulation of a GMM to encode the covariance relations between different variables (either in the task space, or in the robot joints space). If the correlations vary significantly between regions, then each local region of the state space visited during the demonstrations will need a few Gaussians to encode these local dynamics.

One interesting and worthwhile to mention variant of GMR is a locally weighted projection regression (LWPR) [11]. This variant of GMR is specially tailored to the needs of robot learning by demonstration applications with high-dimensional input signals possibly including redundant and irrelevant input dimensions. LWPR, as well asreward-weighted LWPR variants [12], has been shown to yield similar or inferior performance compared with GMR under several experimental setups entailing low-dimensionality input signals (see, e.g., [13]), but seem to perform considerably better in applications where high-dimensional input signals are being modeled.

Recently, several researchers have considered the application of concepts from quantum mechanics in the field of machine learning [14]. The main notion behind these studies consists of the generalization of the probability distribution of the

Manuscript received October 30, 2011; revised April 5, 2012; accepted May 29, 2012. This paper was recommended for publication by Associate Editor A. Ijspeert and Editor D. Fox upon evaluation of the reviewers' comments. This work was supported in part by the European Union Seventh Framework Program through the ALIZ-E Project under Grant 248116.

The authors are with the Department of Electrical and Electronic Engineering, Imperial College London, London, SW7 2BT, U.K. (e-mail: s.chatzis@imperial.ac.uk; d.korkinof10@imperial.ac.uk; y.demiris@imperial.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TRO.2012.2203055

postulated models by introducing a *density matrix*, which is a self-adjoint positive-semidefinite matrix of trace one. Indeed, it has been shown (see, e.g., [15]) that the basic probability rule of quantum mechanics, widely known as the *Born Rule*, which gives rise to the concept of generalized probability by introduction of a density matrix, is closely related to spectral clustering and other machine learning algorithms based on spectral theory.

In [16], a combination of the margin maximization scheme with a probabilistic modeling approach is presented, facilitated by incorporating the concepts of quantum detection and estimation theory [17]. In [18], a quantum Markov random field was proposed, based on the utilization of quantum statistics techniques, such as the concept of the density matrix; the method was successfully applied to image restoration. More recently, a quantum-statistical-mechanical extension of the GMM was presented in [19]; the proposed method was based on the representation of the model marginal likelihood as a function of a density matrix, and was shown to outperform conventional GMMs in an image segmentation task. Finally, a novel regard toward variational Bayesian (VB) inference [20] was proposed in [21], and was shown to outperform conventional VB inference algorithms when applied to a latent Dirichlet allocation [22], one of the most popular probabilistic graphical models for topic-based document retrieval.

Motivated by these results, in this paper, we introduce a novel method for GMR, inspired by concepts from the field of quantum statistics. Indeed, conventional GMR models are formulated under the notion that all the observed data points can be assigned to a distinct number of model states (mixture components). In this paper, we reformulate GMR models, introducing some quantum states constructed by superposing conventional GMR states by means of linear combinations. To affect this goal, we reformulate the expression of the likelihood of conventional GMR models into a special form (diagonal) density matrix, and we further show that this matrix can be generalized into a more generic, nondiagonal form. Based on this novel formulation, a quantum-mechanical-inspired expression of the conditional predictive distribution of the GMR model is eventually derived, and applied to yield a quantum-statistical approach toward robot learning by demonstration. We illustrate the efficacy of the proposed approach by considering a number of demanding robot learning by demonstration scenarios, and we compare its performance with state-of-the-art trajectory-based robot learning by demonstration methodologies.

The remainder of this paper is organized as follows. In Section II, GMR as applied to robot learning by demonstration is introduced in a concise manner. In Section III, we provide a brief review of concepts from the field of quantum information processing, with a special focus on density matrices, which provide a quantum-inspired extension of conventional probability, and the related calculus. In Section IV, we derive the proposed quantum Gaussian mixture regression (QGMR) model, and we elaborate on its application to robot learning by demonstration. In Section V, the experimental evaluation of the proposed algorithm is performed. The final section concludes this paper.

II. Gaussian Mixture Regression for Robot Learning by Demonstration

In this section, we provide a brief overview of GMR as applied in the context of robot learning by demonstration.

Let us define a trajectory as a set of position data (in joint space or task space) sequentially appearing over time. GMR can be used to retrieve smooth generalized trajectories with associated predictive variances expressing the variations of the trajectory variables. Indeed, following the standard setting of all regression algorithms, GMR consists in modeling the conditional expectation of a set of response variables $\dot{\boldsymbol{\beta}} \in \mathbb{R}^{p_2}$ given a set of predictor variables $\boldsymbol{\beta} \in \mathbb{R}^{p_1}$, by exploiting the information available in a set of training observations $\{\boldsymbol{\beta}_j, \dot{\boldsymbol{\beta}}_j\}_{j=1}^{N}$. In the case of trajectory-based robot learning by demonstration, the predictor variable $\boldsymbol{\beta}$ might represent the current position of the moving end-effector, with the response variable $\dot{\boldsymbol{\beta}}$ being the velocity that must be adopted by the end-effector for the next time step, in order to comply with the learnt trajectory.

To effect these goals, contrary to most of the traditional regression methodologies, GMR does not approximate the regression function in a direct fashion. In contrast, GMR postulates a GMM to model the *joint* probability distribution of the considered response and predictor variables ($\dot{\boldsymbol{\beta}}$ and $\boldsymbol{\beta}$), i.e., it considers a model of the form

$$p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^{K}) = \sum_{i=1}^{K} \pi_i \mathcal{N}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \qquad (1)$$

where $\boldsymbol{\pi} = (\pi_i)_{i=1}^{K}$ are the prior weights of the mixture component densities, and $\mathcal{N}(\cdot|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ is a Gaussian with mean $\boldsymbol{\mu}_i$ and covariance matrix $\boldsymbol{\Sigma}_i$. The postulated GMM (1) is trained by means of the EM algorithm [23], using a set of training data corresponding to a number of trajectories obtained by human demonstrators. Then, using the obtained GMM $p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^{K})$, GMR retrieves a generalized trajectory by estimating at each time step the conditional expectation $\mathbb{E}[\dot{\boldsymbol{\beta}}|\boldsymbol{\beta}; \boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^{K}]$.

Let us express the means $\boldsymbol{\mu}_i$ of the component densities of the postulated GMM (1) in the form

$$\boldsymbol{\mu}_i = \begin{bmatrix} \boldsymbol{\mu}_i^{\boldsymbol{\beta}} \\ \boldsymbol{\mu}_i^{\dot{\boldsymbol{\beta}}} \end{bmatrix} \qquad (2)$$

where $\boldsymbol{\mu}_i^{\boldsymbol{\beta}}$ is the mean of the variable $\boldsymbol{\beta}$, and $\boldsymbol{\mu}_i^{\dot{\boldsymbol{\beta}}}$ is the mean of the variable $\dot{\boldsymbol{\beta}}$. Let us also introduce the notation

$$\boldsymbol{\Sigma}_i = \begin{bmatrix} \boldsymbol{\Sigma}_i^{\boldsymbol{\beta}} & \boldsymbol{\Sigma}_i^{\boldsymbol{\beta}\dot{\boldsymbol{\beta}}} \\ \boldsymbol{\Sigma}_i^{\dot{\boldsymbol{\beta}}\boldsymbol{\beta}} & \boldsymbol{\Sigma}_i^{\dot{\boldsymbol{\beta}}} \end{bmatrix} \qquad (3)$$

for the covariance matrices of the model component densities. Then, it is easy to show that, based on (1) and the assumptions (2) and (3), the conditional probability $p(\dot{\boldsymbol{\beta}}|\boldsymbol{\beta}; \boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^{K})$ of the response variables $\dot{\boldsymbol{\beta}}$, given the predictor variables $\boldsymbol{\beta}$ and the postulated GMM, yields [24]

$$p(\dot{\boldsymbol{\beta}}|\boldsymbol{\beta}; \boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^{K}) = \mathcal{N}(\dot{\boldsymbol{\beta}}|\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) \qquad (4)$$

where

$$\hat{\boldsymbol{\mu}} = \sum_{i=1}^{K} \phi_i(\boldsymbol{\beta})[\boldsymbol{\mu}_i^{\dot{\boldsymbol{\beta}}} + \boldsymbol{\Sigma}_i^{\dot{\boldsymbol{\beta}}\boldsymbol{\beta}}(\boldsymbol{\Sigma}_i^{\boldsymbol{\beta}})^{-1}(\boldsymbol{\beta} - \boldsymbol{\mu}_i^{\boldsymbol{\beta}})] \tag{5}$$

$$\hat{\boldsymbol{\Sigma}} = \sum_{i=1}^{K} \phi_i^2(\boldsymbol{\beta})[\boldsymbol{\Sigma}_i^{\dot{\boldsymbol{\beta}}} - \boldsymbol{\Sigma}_i^{\dot{\boldsymbol{\beta}}\boldsymbol{\beta}}(\boldsymbol{\Sigma}_i^{\boldsymbol{\beta}})^{-1}\boldsymbol{\Sigma}_i^{\boldsymbol{\beta}\dot{\boldsymbol{\beta}}}] \tag{6}$$

and

$$\phi_i(\boldsymbol{\beta}) = \frac{\pi_i \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\mu}_i^{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_i^{\boldsymbol{\beta}})}{\sum_{k=1}^{K} \pi_k \mathcal{N}(\boldsymbol{\beta}|\boldsymbol{\mu}_k^{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_k^{\boldsymbol{\beta}}).} \tag{7}$$

Based on (4), predictions under the GMR approach can be obtained by taking the conditional expectations $\mathbb{E}(\dot{\boldsymbol{\beta}}|\boldsymbol{\beta}; \boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^K)$, i.e.,

$$\hat{\dot{\boldsymbol{\beta}}} \triangleq \mathbb{E}(\dot{\boldsymbol{\beta}}|\boldsymbol{\beta}; \boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^K) = \hat{\boldsymbol{\mu}}. \tag{8}$$

As we observe, a significant merit of GMR consists in the fact that it provides a full predictive distribution; thus, a predictive variance

$$\mathbb{V}(\dot{\boldsymbol{\beta}}|\boldsymbol{\beta}; \boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^K) = \hat{\boldsymbol{\Sigma}}$$

is available at any position of the end-effector. Therefore, GMR offers a model-estimated measure of predictive uncertainty not only at specific positions, but continuously along the generated trajectories.

One of the most significant advantages of GMR-based robot learning by demonstration can be traced to the prediction generation procedure: Contrary to most discriminative regression algorithms (e.g., support vector machines [25] and Gaussian processes [26]), the computational time required for trajectory reproduction using (5) does not increase with the number of demonstrations provided to the robot, which is a particularly important property for lifelong learning robots. Indeed, the available model training data provided by the employed human demonstrators are processed in only an off-line fashion, to obtain the estimates of the model parameters by means of the EM algorithm. This way, prediction generation under GMR reduces to the estimation of a simple weighted sum of linear models; as a result, the regression phase is processed very quickly, which is advantageous because the reproduction of smooth trajectories is fast enough to be used at any appropriate time by the robot.

Finally, apart from the GMM model estimation, another significant issue that always has to be addressed is data-driven selection of the appropriate number of GMM component densities. The number of component densities in the postulated GMMs is significant for the performance of GMR-based trajectory learning, as it determines the compromise for GMR between having an accurate estimation of the response and having a smooth response (bias-variance tradeoff). Optimal model size *(order)* selection for finite mixture models is an important but very difficult problem which has not been completely resolved. Indeed, a number of approaches have been proposed for this purpose, including likelihood ratio test statistics, information criteria, Bayesian-based information criteria, and classification-based information criteria [23]. In this paper, we consider application of the main and oldest Bayesian-based model selection criterion, namely the Bayesian information criterion (BIC) of Schwarz [27]. The BIC model selection criterion as applied to a GMR-fitted GMM used for trajectory-based robot learning by demonstration consists of the determination of the number of model component densities which minimizes the metric

$$\mathcal{L} \stackrel{\triangle}{=} -2\sum_{n=1}^{N} \log p(\{\boldsymbol{\beta}_n, \dot{\boldsymbol{\beta}_n}\}_{n=1}^N|\boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^K) + d\log N \tag{9}$$

where d is the total number of model parameters, hence a function of the number of mixture component densities K, and N is the number of available model training data points. BIC has been shown not to underestimate the required number of mixture components, asymptotically, and to provide consistent model order estimators under certain conditions [28].

III. Quantum Statistics and the Density Matrix

Probability is the main concept in the field of classical statistics. This is also the main point where quantum statistics depart from classical statistics: in quantum statistics, probabilities are replaced by density matrices. A density matrix is a self-adjoint positive-semidefinite matrix and its trace is one. Indeed, it is easy to show that the conventional probabilities used in classical statistics can be expressed as density matrices of a special form *(diagonal)*. For example, let us consider a classical system comprising q states. Let us also introduce the notation $\boldsymbol{\varpi} = (\varpi_i)_{i=1}^q$, with the ϖ_i being the (prior) probability of occurrence of the ith system state. We also denote as $\{\boldsymbol{e}_i\}_{i=1}^q$ a set of basis vectors, such that $\boldsymbol{e}_i = (e_{ik})_{k=1}^q$ with

$$e_{ik} = \begin{cases} 1, & \text{if } i = k \\ 0, & \text{if } i \neq k \end{cases}$$

Based on these assumptions, the density matrix for this postulated classical system can be defined as

$$\Phi \stackrel{\triangle}{=} \begin{bmatrix} \varpi_1 & 0 & \dots & 0 \\ 0 & \varpi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \varpi_q \end{bmatrix} = \sum_{i=1}^{q} \varpi_i \boldsymbol{e}_i \boldsymbol{e}_i^T. \tag{10}$$

In quantum statistics, the concept of probability is extended by introducing nondiagonal elements in the density matrix Φ of the treated system. The states of a system in quantum statistics are defined by the unit vectors $\boldsymbol{u}$; the matrix $\boldsymbol{u}\boldsymbol{u}^T$ corresponding to a state vector $\boldsymbol{u}$ is called dyad and has trace one, i.e., $\mathrm{tr}(\boldsymbol{u}\boldsymbol{u}^T) = \boldsymbol{u}^T\boldsymbol{u} = 1$. On this basis, the density matrix Φ generalizes the concept of probability distribution and can be defined as a mixture of dyads in the sense

$$\Phi = \sum_{i=1}^{q} \varpi_i \boldsymbol{u}_i \boldsymbol{u}_i^T \tag{11}$$

where ϖ_i is the (prior) probability of the system state represented by the dyad of $\boldsymbol{u}_i$. Under this construction, the probability assigned to the unit vector $\boldsymbol{u}$ and its associated dyad $\boldsymbol{u}\boldsymbol{u}^T$ yields

$$p(\boldsymbol{u}) = \mathrm{tr}(\Phi \boldsymbol{u}\boldsymbol{u}^T) = \boldsymbol{u}^T \Phi \boldsymbol{u}. \tag{12}$$

Equation (12) is widely known within the quantum mechanics community as the *Born rule*. Note that the unit vectors and their associated dyads in quantum systems have a straightforward natural interpretation. Consider, for example, a quantum system with four "pure" states; then, a state vector of the form $\boldsymbol{u} = (\frac{1}{2}, 0, \frac{\sqrt{3}}{2}, 0)$ represents the mixture of the first "pure" state and the third "pure" state of the system, with probabilities $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$, and $\left(\frac{\sqrt{3}}{2}\right)^2 = \frac{3}{4}$, respectively.

IV. Quantum Gaussian Mixture Regression

A. Using Quantum Gaussian Mixture Models to Obtain a Regression Algorithm

Let us reconsider the case of GMR and the associated GMM postulated to represent the joint distribution of the modeled predictor and response variables. Let us consider a K-component postulated GMM, as in (1). We introduce the following matrices:

$$\boldsymbol{\Psi} \stackrel{\triangle}{=} -\begin{bmatrix} \log\pi_1 & 0 & \dots & 0 \\ 0 & \log\pi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \log\pi_K \end{bmatrix} \tag{13}$$

and

$$\boldsymbol{\Omega}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}) \stackrel{\triangle}{=} \begin{bmatrix} p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\theta}_1) & 0 & \dots & 0 \\ 0 & p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\theta}_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\theta}_K) \end{bmatrix} \tag{14}$$

where $\boldsymbol{\theta}_i \stackrel{\triangle}{=} \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}$, and $p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\theta}_i)$ is the conditional joint probability of the predictor and response variables given that they are emitted from the ith component density of the postulated GMM, i.e.,

$$p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\theta}_i) = \mathcal{N}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i). \tag{15}$$

Under this scheme, the likelihood $p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^K)$ of the postulated model, given by (1), can be equivalently expressed in the form [19]

$$p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^K) = \frac{\mathrm{tr}[\exp(-\boldsymbol{H}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}))]}{\int_{\boldsymbol{\omega}} \mathrm{tr}[\exp(-\boldsymbol{H}(\boldsymbol{\omega}))]\mathrm{d}\boldsymbol{\omega}} \tag{16}$$

where

$$\boldsymbol{H}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}) = \boldsymbol{\Psi} - \log\boldsymbol{\Omega}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}). \tag{17}$$

The exponential $\exp(\boldsymbol{A})$ of a matrix $\boldsymbol{A}$ is defined as

$$\exp(\boldsymbol{A}) \stackrel{\triangle}{=} \sum_{\rho=0}^{\infty} \frac{1}{\rho!}\boldsymbol{A}^{\rho} \tag{18}$$

and the logarithm $\log(\boldsymbol{A})$ is given by

$$\log(\boldsymbol{A}) \stackrel{\triangle}{=} -\sum_{\rho=1}^{\infty} \frac{1}{\rho}(\boldsymbol{I} - \boldsymbol{A})^{\rho}. \tag{19}$$

On the basis of the aforementioned construction, we have managed to express the likelihood (1) of the postulated GMM employed by the GMR model as a function of the density matrix of a quantum system with a special form *(having a diagonal density matrix)*. Then, based on the relevant discussions of Section III, one may generalize the results of (16) to the case of a general (symmetric) density matrix so as to obtain a quantum-statistical approach toward GMR-based trajectory learning by demonstration. Specifically, let us extend the diagonal matrix $\boldsymbol{\Psi}$ to a symmetric $K \times K$ matrix as follows:

$$\boldsymbol{\Psi} \stackrel{\triangle}{=} -\begin{bmatrix} \log\pi_1 & \gamma & \dots & \gamma \\ \gamma & \log\pi_2 & \dots & \gamma \\ \vdots & \vdots & \ddots & \vdots \\ \gamma & \gamma & \dots & \log\pi_K \end{bmatrix} \tag{20}$$

where γ is a hyperparameter related to the prior probability of a "mixed" model state (comprising whichever two "pure" model states).Then, based on (17), the symmetric $K \times K$ matrix $\boldsymbol{H}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}})$ of the model yields

$$\begin{aligned} \boldsymbol{H}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}) &= -\sum_{k=1}^{K}\sum_{k'=1}^{K} \xi_{kk'}\boldsymbol{\Delta}_{kk'} \\ &= -\begin{bmatrix} \log\{\pi_1 p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\theta}_1)\} & \dots & \gamma \\ \gamma & \dots & \gamma \\ \vdots & \ddots & \vdots \\ \gamma & \dots & \log\{\pi_K p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\theta}_K)\} \end{bmatrix} \end{aligned} \tag{21}$$

where the coefficients $\xi_{kk'}$ are defined as

$$\xi_{kk'} \stackrel{\triangle}{=} \begin{cases} \log\{\pi_k p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\theta}_k)\}, & \text{if } k = k' \\ \gamma, & \text{if } k \neq k' \end{cases}$$

and the $\boldsymbol{\Delta}_{kk'}$ are $K \times K$ matrices, the (l, l')th elements of which are defined as

$$(\boldsymbol{\Delta}_{kk'})_{ll'} \stackrel{\triangle}{=} \delta(k - l)\delta(k' - l') \tag{22}$$

where $\delta(\cdot)$ is the Kronecker's delta function.

A GMM with likelihood expression of the form (16), where the density matrices $\boldsymbol{H}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}})$ are given by the generalized expression (21), is usually referred to as a quantum Gaussian mixture model (QGMM) [19]. An issue of this formulation of QGMMs is that the integral in the denominator (regularization constant) of their likelihood (16) is difficult to compute analytically. To alleviate these issues, typically, the following approximation is adopted [19]:

$$p(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}|\boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^K) = \frac{\mathrm{tr}[\exp(-\boldsymbol{H}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}}))]}{\mathrm{tr}[\exp(-\boldsymbol{\Psi})]}. \tag{23}$$

Apparently, a GMR model the employed GMM of which is of QGMM type departs from the classical definition of the GMR model, providing a quantum-statistical approach toward GMR. The entailed density matrices of the so-derived model include some quantum effects and are based on states constructed by

superposing the model states corresponding to the mixture components of a classical GMR model.

B. Proposed Algorithm

Definition 1: We define the QGMR model as a GMR model that employs a GMM with likelihood of the form (23), with the density matrices $\boldsymbol{H}(\boldsymbol{\beta}, \dot{\boldsymbol{\beta}})$ given by the generalized expression (21).

To develop a trajectory-based robot learning by demonstration algorithm based on the QGMR framework, we have to 1) provide an algorithm for model training using a set of human-generated demonstrations and 2) derive the expression of the conditional density of the response variables given the estimated joint distribution model.

1) Model Training: Let us consider a model fitting dataset comprising N samples $\{\boldsymbol{y}_j\}_{j=1}^N \triangleq \{[\boldsymbol{\beta}_j; \dot{\boldsymbol{\beta}}_j]\}_{j=1}^N$. Then, likelihood maximization for the postulated model (23) can be shown to yield the following estimators for the model parameters [19]

$$\pi_i \approx \frac{\sum_{j=1}^N \psi_{ij}}{N} \tag{24}$$

$$\boldsymbol{\mu}_i = \frac{\sum_{j=1}^N \psi_{ij} \boldsymbol{y}_j}{\sum_{j=1}^N \psi_{ij}} \tag{25}$$

and

$$\boldsymbol{\Sigma}_i = \frac{\sum_{j=1}^N \psi_{ij} (\boldsymbol{y}_j - \boldsymbol{\mu}_i)(\boldsymbol{y}_j - \boldsymbol{\mu}_i)^T}{\sum_{j=1}^N \psi_{ij}} \tag{26}$$

where

$$\psi_{ij} \triangleq \frac{\mathrm{tr}[\boldsymbol{\Delta}_{ii} \exp(-\boldsymbol{H}(\boldsymbol{\beta}_j, \dot{\boldsymbol{\beta}}_j))]}{\mathrm{tr}[\exp(-\boldsymbol{H}(\boldsymbol{\beta}_j, \dot{\boldsymbol{\beta}}_j))]} \tag{27}$$

and the matrices $\boldsymbol{\Delta}_{ii}$ are given by (22).

2) Predictive Density: Having obtained the estimators of the QGMR model parameters, we can now proceed to the derivation of the model predictive density, that is the conditional density $p(\dot{\boldsymbol{\beta}}|\boldsymbol{\beta}; \boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^K)$. For this purpose, we rely on the adoption of the classical expression

$$p(\dot{\boldsymbol{\beta}}|\boldsymbol{\beta}; \boldsymbol{\pi}, \{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^K) = \mathcal{N}(\dot{\boldsymbol{\beta}}|\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\Sigma}}) \tag{28}$$

for the predictive density of our model, where the predictive mean is given by

$$\hat{\boldsymbol{\mu}} = \sum_{i=1}^K \phi_i(\boldsymbol{\beta}) [\boldsymbol{\mu}_i^{\dot{\boldsymbol{\beta}}} + \boldsymbol{\Sigma}_i^{\dot{\boldsymbol{\beta}}\boldsymbol{\beta}} (\boldsymbol{\Sigma}_i^{\boldsymbol{\beta}})^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_i^{\boldsymbol{\beta}})] \tag{29}$$

and the predictive covariance reads

$$\hat{\boldsymbol{\Sigma}} = \sum_{i=1}^K \phi_i^2(\boldsymbol{\beta}) [\boldsymbol{\Sigma}_i^{\dot{\boldsymbol{\beta}}} - \boldsymbol{\Sigma}_i^{\dot{\boldsymbol{\beta}}\boldsymbol{\beta}} (\boldsymbol{\Sigma}_i^{\boldsymbol{\beta}})^{-1} \boldsymbol{\Sigma}_i^{\boldsymbol{\beta}\dot{\boldsymbol{\beta}}}] \tag{30}$$

with the $\boldsymbol{\mu}_i^{\boldsymbol{\beta}}$, $\boldsymbol{\mu}_i^{\dot{\boldsymbol{\beta}}}$, $\boldsymbol{\Sigma}_i^{\boldsymbol{\beta}}$, $\boldsymbol{\Sigma}_i^{\boldsymbol{\beta}\dot{\boldsymbol{\beta}}}$, $\boldsymbol{\Sigma}_i^{\dot{\boldsymbol{\beta}}\boldsymbol{\beta}}$, and $\boldsymbol{\Sigma}_i^{\dot{\boldsymbol{\beta}}}$ defined as in (2) and (3), and the estimates of $\boldsymbol{\mu}_i$ and $\boldsymbol{\Sigma}_i$ given by (25) and (26), respectively.

To determine the model state weight values $\phi_i(\boldsymbol{\beta})$ in (29) and (30) for the QGMR model, we first consider the quantum GMM of the predictor variables

$$p(\boldsymbol{\beta}|\boldsymbol{\pi}, \{\boldsymbol{\mu}_i^{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_i^{\boldsymbol{\beta}}\}_{i=1}^K) = \frac{\mathrm{tr}\left[\exp(-\boldsymbol{H}(\boldsymbol{\beta}))\right]}{\mathrm{tr}\left[\exp(-\boldsymbol{\Psi})\right]} \tag{31}$$

where

$$\boldsymbol{H}(\boldsymbol{\beta}) = -\begin{bmatrix} \log\left\{\pi_1 p(\boldsymbol{\beta}|\boldsymbol{\mu}_1^{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_1^{\boldsymbol{\beta}})\right\} & \dots \gamma \\ \gamma & \dots \gamma \\ \vdots & \vdots \\ \gamma \dots & \log\{\pi_K p(\boldsymbol{\beta}|\boldsymbol{\mu}_K^{\boldsymbol{\beta}}, \boldsymbol{\Sigma}_K^{\boldsymbol{\beta}})\} \end{bmatrix} \tag{32}$$

and $\boldsymbol{\Psi}$ is defined in (20). Based on this model, selection of the values of $\phi_i(\boldsymbol{\beta})$ is conducted by setting them equal to the *responsibilities*of the quantum GMM (31) of the predictor variables, i.e.,

$$\phi_i(\boldsymbol{\beta}) = \frac{\mathrm{tr}\left[\boldsymbol{\Delta}_{ii} \exp(-\boldsymbol{H}(\boldsymbol{\beta}))\right]}{\mathrm{tr}\left[\exp(-\boldsymbol{H}(\boldsymbol{\beta}))\right]}. \tag{33}$$

V. Experimental Evaluation

In this section, we provide a thorough experimental evaluation of the QGMR algorithm, in a series of applications dealing with robot learning by demonstration, and compare its performance with state-of-the-art methods in the field. Our source codes have been developed in MATLAB R2010b, and were run on a Macintosh platform with an Intel Core i7 2.67 GHz CPU, and 4 GB RAM, running Mac OS X 10.6. The model-estimated values generated in MATLAB were sent in real time to the NAO robot by means of a MATLAB–NAO communication protocol written in Python. Therefore, the results reported here are obtained from the actual robot and not by means of simulation.

In our experiments, we employ a humanoid robotic platform, namely the NAO robot (academic edition), a humanoid robot with 27 degrees of freedom [29]. The predictor variable $\boldsymbol{\beta}$ used by the considered models is the position vector of the robot joints, whereas the response variable $\dot{\boldsymbol{\beta}}$ is the velocity vector of the robot joints, that is, the velocity that should be imposed on the robot joints so as to remain on the learnt trajectory.

Our approach is compared against two popular state-of-the-art methods for robot learning by demonstration, namely GMR [10], and the local Gaussian process regression (LGPR) method of [30]. The latter method clusters the input space into smaller subsets, and fits a dedicated Gaussian process regression model for each one of these subspaces. As such, it shares similarities with the GMR and QGMR methods, which also divide the input space into subspaces and postulate different regression models on each one of these subspaces. We utilize several performance metrics for our comparisons, selected on the basis of the individual characteristics of each experiment. Regarding model order selection for the GMR and QGMR methods, we repeat our experiments for various numbers of model states Q, and detect the values yielding the lowest generalization error rates. All our experiments are conducted with the hyperparameter γ of the QGMR model density matrix set equal to $\gamma = 2$, a value heuristically determined to work well for our model.

TABLE I
DATASETS DETAILS

Task	One-shot learning dataset		Multi-shot learning dataset		Units
	#Data points	#Dimensions	#Data points	#Dimensions	
Blocking	445	11	2175	8	rad
Ph. Education	355	5	849	5	rad
Lazy figure 8	242	6	717	5	rad
Lazy figure 8 end-effector	316-337	3	690-712	3	cm

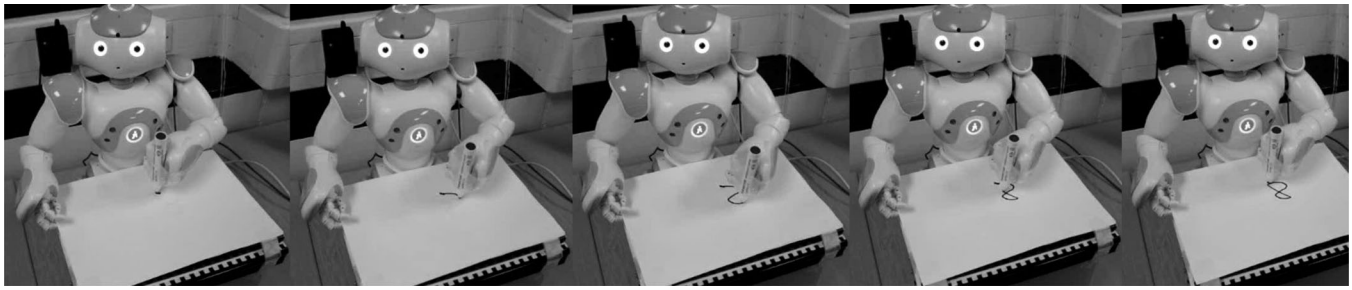

Fig. 1. NAO robot during the L8 experiment.

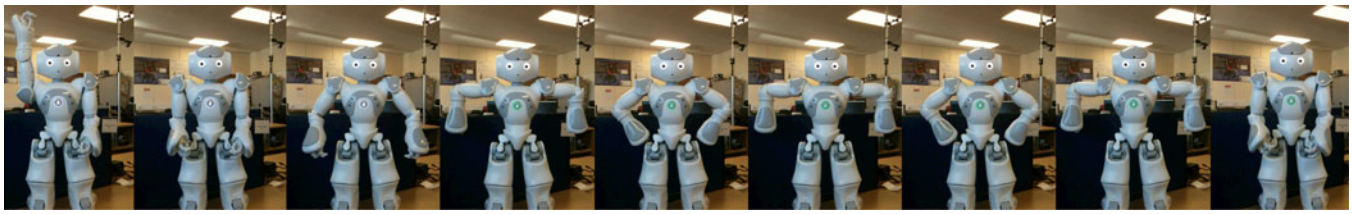

Fig. 2. Communicative gesture for the violation "Blocking."

In an attempt to account for the effect of bad local optima where the EM algorithm might get trapped into in cases of poor model initialization, all our experiments have been executed multiple times, each time with different k-means initializations for the training algorithms of the evaluated models. Means and standard deviations of the performance of the compared algorithms over the executed multiple runs are provided, and the statistical significance of these results is assessed. Finally, we would like to underline that, in all our experiments, we have ensured that *the total number of model component densities is at least one order of magnitude less than the* number of available training data points (see also Table I and Figs. 4 and 6). This is a good method to ward off the possibility of model overfitting [23].

A. One-Shot Learning

In the following set of experiments, we evaluate the ability of the proposed model to learn and reproduce a skill from a single demonstration. We briefly describe the experiments below; more details concerning the used datasets are provided in Table I.

1) *Lazy figure 8:* In this experiment, we evaluate the considered methods in terms of their applicability in teaching a robot by demonstration how to draw a complex figure. The considered figure comprises a *lazy figure 8* (L8) (see Fig. 1). The L8 generation task is a classical benchmark for pattern generation methodologies [31], [32]. From the first impression, the task appears to be trivial, since an 8 figure can be interpreted as the superposition of a sine on the horizontal direction and a cosine of half the sine's frequency on the vertical direction. A closer inspection though will reveal that in reality, this seemingly innocent task entails surprisingly challenging stability problems, which come to the fore especially when using very limited model training datasets. The dataset used consists of joint angle data from drawing three consecutive *L8s*.
2) *Upper body motion:* In the case of upper body motion, our experiments involve a higher number of joints, thus further increasing the dimensionality and, consequently, the complexity of the addressed problem. We examine learning and reproduction of a communicative gesture used by Basketball officials, with potential applicability in the case of a robotic referee. We have chosen a gesture that poses a challenge on the learning by demonstration algorithm in terms of the implied motion complexity, namely, the sign concerning the violation "blocking"[1] (see Fig. 2).
3) *Lower body motion:* Finally, we examine an experimental case involving movement of the lower robot body, simulating a lower abdominal muscle exercise (see Fig. 3). This is one of the scenarios under investigation of the ALIZ-E EU FP7 project (aliz-e.org), where robots are used as companions to diabetic and obese children in pediatric ward settings over extended time periods, and learn along with

[1] Also referred to as "traveling."

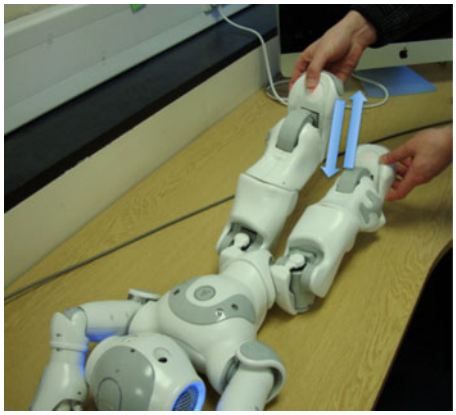

Fig. 3. Ph. education exercise for the lower abdominal muscles.

the children various sensorimotor activities (e.g., dance, games, and physical exercises) so that they can practice and improve together.

In this set of experiments, we wish to investigate the overall accuracy of reproduction of the learned trajectories. Therefore, we choose as our performance metric the mean square error (MSE) calculated over the whole length of the obtained trajectories. Especially in the case of the L8s experiment, we also compute the end-effector MSE so as to assess the accuracy of the eventual reproduction of the figures. The training trajectories are presented to the NAO robot by means of kinesthetics;[2] during this procedure, joint position sampling is conducted, with the sampling rate equal to 10 Hz. The number of joints actively participating in each experiment varies according to the specification of the performed motion types.

In our experiments, we use the training sequences obtained from human demonstrators (through kinesthetics) without further preprocessing. As such, our datasets also contain information pertaining to joints with minor contribution to the learned movements, thus further increasing the difficulty of the tasks. According to our experimental scenario, during the testing phase, the evaluated algorithms are initialized at points obtained by adding uniformly distributed noise $\mathcal{U}(0,1)$ to the initial points of the training sequences, and the algorithms are executed to regenerate the (rest of the) learnt trajectories.

The error means and standard deviations resulting from 20 independent, but common for both methods, random initializations of the GMR and QGMR models are presented in Fig. 4. The best mean MSE for each method and the associated standard deviation, along with the LGPR error, are presented in Table II. In Table II, we have calculated the mean MSE along with its standard deviation for each value of the number of states, and presented the best result. We have also applied the Student-t statistical test on the obtained results to establish the statistical significance of our findings. The outcome of this statistical test is presented in Table IV; based on our findings, we can definitively deduce that, in every experiment, there is a statistically significant difference between the two main evaluated methods, namely, the GMR and the proposed QGMR algorithm.

[2]Manually moving the robot's arms and recording the joint angles.

Additionally, from Fig. 4, we observe that GMR suffers from higher volatility and higher MSE errors than the QGMR approach, which achieves not only better but also more consistent results. Finally, we observe that an insufficient number of states are translated in slightly higher MSEs for both the evaluated methods, and also in higher model performance volatility. From Table II, we can conclude that the mean error obtained by the GMR is from twice as high to approximately one order of magnitude higher than the QGMR. LGPR yielded competitive results; however, the results were clearly inferior to our approach in most cases.

As previously mentioned, especially in the case of the *L8*s experiment, we are interested to evaluate the end-effector error. This result is obtained by reproducing the original demonstration, as well as the predicted data generated by both the GMR and QGMR methods. Specifically, the NAO robot is given the joint angles of the demonstration and the model-predicted data. The end-effector positions are recorded and the resulting trajectories obtained by using the GMR and QGMR-generated predictions are then separately aligned with the end-effector data from the original demonstration. This way, we ensure fairness for both methods. We have chosen a case where both methods' performance is neither the best nor the worst, for which the corresponding end-effector MSE results are presented in Table III.

B. Multishot Learning

In this experimental case, we use multiple demonstrations of each task so as to capture the variability of the human action, and evaluate our model's ability to generalize learned trajectories. More specifically, the tasks in question are the same ones described in the one-shot experimental scenario, namely, the *L8*, Ph. education exercise, and Blocking communicative gesture. The training trajectories are again presented to the NAO robot by means of kinesthetics; during this procedure, joint position sampling is conducted, with the sampling rate now equal to 20 Hz. For each of the three tasks, we have recorded four demonstrations and used three for training and one for testing purposes. Due to the temporal variations observed in the demonstrations, we have preprocessed the sequences using dynamic time warping (DTW) [33], a method first used in speech recognition for signal alignment, combined with a low-pass filter to smooth the resulting trajectories. In Table I, we present some details concerning the number of points and the dimensionality of each dataset. It should be noted that the number of points for the multishot experiments is considerably higher compared with the one-shot scenario due to the higher sampling frequency as well as the oversampling that occurs as a result of the alignment of the trajectories.

The error metric used in this case is the MSE, as we are again interested in the overall accuracy of movement reproduction. We would like to emphasize that, in our experiments,comparison of the reproduced trajectories to the demonstrated trajectories has been conducted using the time-aligned trajectories obtained by application of DTW. As such, the calculated MSE statistics provide a genuine assessment of the performance difference between the evaluated methods, not affected by reproduction delays.

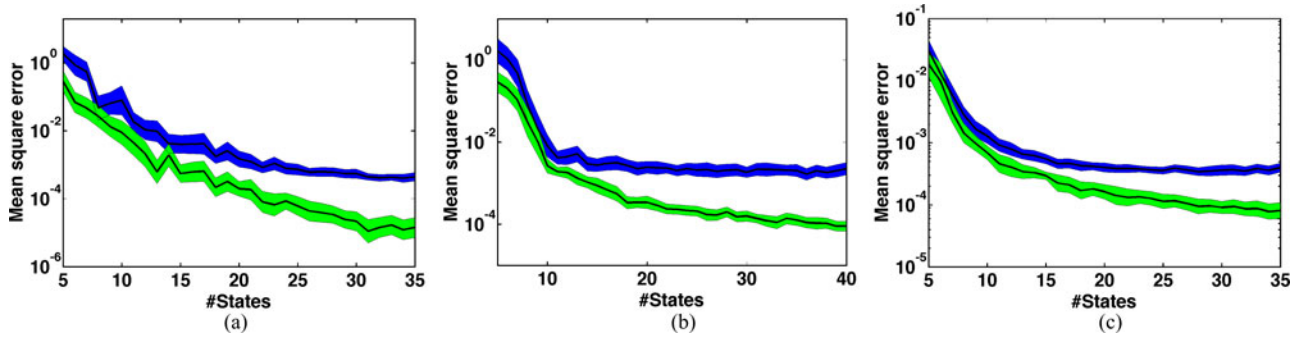

Fig. 4. One-shot learning experiments: Mean and standard deviation of the MSE over the executed repetitions as a function of model size. Blue: GMR, Green: QGMR. (a) Blocking communicative gesture experiment. (b) Ph. education exercise experiment. (c) Drawing L8s experiment.

TABLE II
ONE-SHOT AND MULTISHOT LEARNING EXPERIMENTS: BEST MEAN MSE RESULTS FOR ALL EVALUATED METHODS

	One-shot learning MSE			Multi-shot learning MSE		
Task	GMR	LGPR	QGMR	GMR	LGPR	QGMR
Blocking	$3.8 \cdot 10^{-4}$ ($\pm 1 \cdot 10^{-4}$)	$8.5 \cdot 10^{-5}$	$1.1 \cdot 10^{-5}$ ($\pm 1.6 \cdot 10^{-5}$)	0.0215 ($\pm$0.0024)	0.034303	0.0205 ($\pm$0.0019)
Ph. Education	0.0017 ($\pm 9.7 \cdot 10^{-4}$)	$10.3 \cdot 10^{-5}$	$9.5 \cdot 10^{-5}$ ($\pm 3.5 \cdot 10^{-5}$)	0.0927($\pm$0.0671)	0.047290	0.0075($\pm$0.0029)
Lazy figure 8s	$3.7 \cdot 10^{-4}$ ($\pm 6.1 \cdot 10^{-5}$)	$2.4 \cdot 10^{-4}$	$6.9 \cdot 10^{-5}$ ($\pm 3.4 \cdot 10^{-5}$)	0.0064 ($\pm 9.8 \cdot 10^{-4}$)	0.005985	0.0036 ($\pm 1.3 \cdot 10^{-4}$)

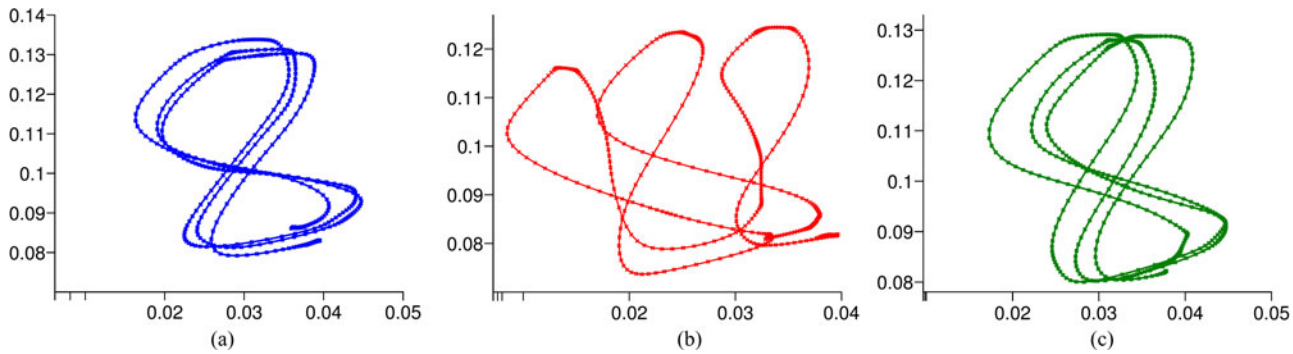

Fig. 5. Multishot learning experiments: Visual representation of the end-effector data from both the original demonstration and the GMR and QGMR methods. (a) Original demonstration. (b) GMR. (c) QGMR.

TABLE III
END-EFFECTOR MSES FOR THE L8 EXPERIMENT

Task	One-shot learning		Multi-shot learning		Units
	GMR	QGMR	GMR	QGMR	
Joint angles error	$3.5 \cdot 10^{-4}$	$6 \cdot 10^{-5}$	0.0059	0.0035	rad
End-effector error	0.0206	0.0032	0.2925	0.0571	cm

Additionally, especially for the needs of the *L8* experiment,we have also considered the end-effector error in a manner similar to the one-shot scenario (see Table III). This way, we are able to definitively show that the demonstrations are not distorted by the preprocessing, that the proposed method yields better end-effector error, and that the proposed QGMR approach is able to successfully reproduce the demonstration in occasions where the GMR fails to do so. The visual result of the end-effector data is presented in Fig. 5, where we can clearly see that the QGMR result is closer to the original demonstration, and the GMR performs poorly in the reproduction of the *L8* figures.

We have calculated the mean and standard deviation of the errors from 100 repetitions of the training and testing procedures with common initialization for the GMR and the QGMR methods in each repetition. The obtained results are presented in Fig. 6. In Table II, we present the best MSE errors obtained by all the evaluated methods. To assess the statistical significance of our findings, in Table IV, we show the results of the Student-t test for the number of model states that yields optimal performance. As we observe, compared with our approach, the GMR method achieves an optimal error two times higher and one order of magnitude higher in the *L8* and Ph. education experiments, respectively. Fig. 6 also reveals lower errors and higher consistency of the results. As far as the "Blocking" experiment is concerned, the optimal error results are much closer for the two evaluated methods. However, the QGMR method again achieves lower and more consistent errors, especially for a higher number of states. Moreover, the Student-t test rejects, with very high certainty, the null hypothesis that the error values of the two methods follow the same distribution. Similar to the one-shot learning scenario, GMR is unable to successfully reproduce any of the learned trajectories. Regarding LGPR, we observed competitive performance, which, however, is clearly inferior to our approach.

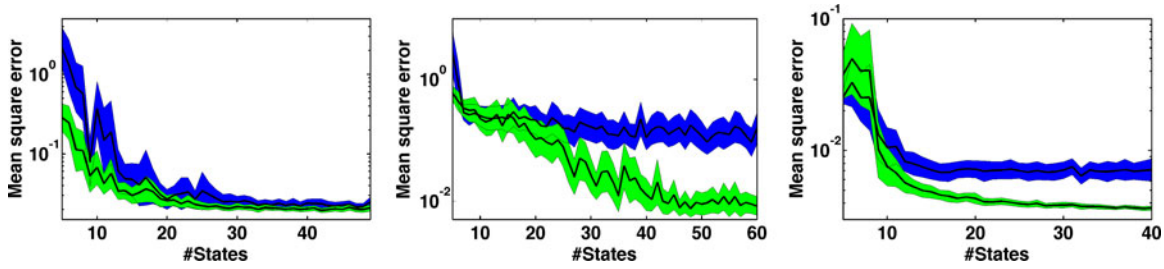

Fig. 6. Multishot learning experiments: Mean and standard deviation of the MSE over the executed repetitions as a function of model size. Blue: GMR, Green: QGMR. (a) "Blocking" communicative gesture experiment. (b) Ph. education exercise experiment. (c) Drawing L8s experiment.

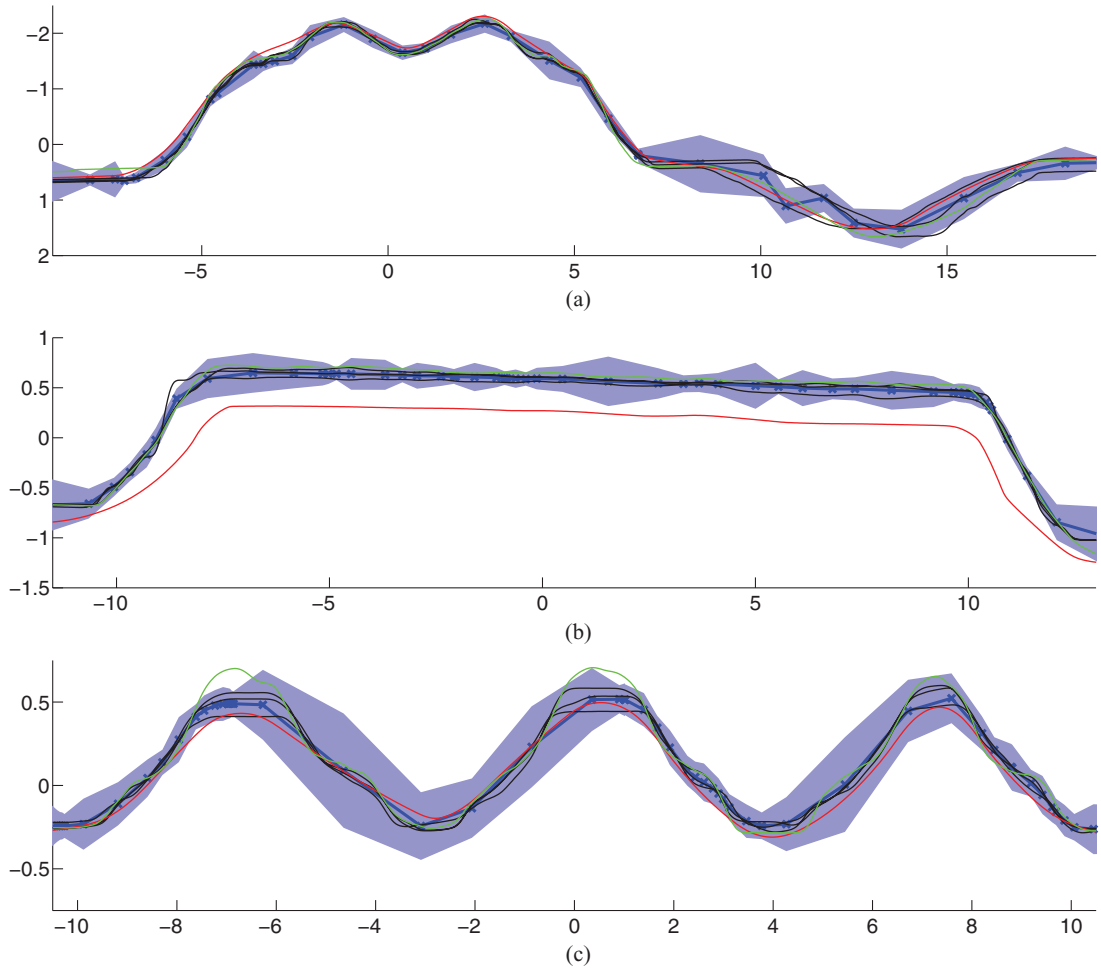

Fig. 7. Goodness of fit graph: Black: Training set; Green: Testing set; Red: GMR; and Blue: QGMR. (a) "Blocking" communicative gesture. (b) Ph. education experiment. (c) L8s experiment.

TABLE IV
STATISTICAL SIGNIFICANCE RESULTS FROM THE STUDENT-T TEST

Task	One-shot		Multi-shot	
	Null hypothesis	p-value	Null hypothesis	p-value
Blocking	rejected	$2.5 \cdot 10^{-22}$	rejected	$4.8 \cdot 10^{-6}$
Ph. Education	rejected	$3.7 \cdot 10^{-32}$	rejected	$5.2 \cdot 10^{-17}$
Lazy Figure 8s	rejected	$1.9 \cdot 10^{-53}$	rejected	$1.8 \cdot 10^{-42}$

Obtained p-values below 10^{-2} indicate high statistical significance.

In conclusion, in Fig. 7, we present a graphical representation of the fit of the model to the data, where we depict the three training sequences (black), the testing sequence (green), the GMR-predicted data (red), and the means and standard deviations of the QGMR model. As all trajectories are of high dimensionality, this graph was obtained by effectively reducing the data dimensions to $D = 2$, by application of the Karhunen–Loeve transform. In order to calculate the corresponding covariance matrices of the model in this low-dimensional space, we sampled from the distributions $\{N(\mu_j, \Sigma_j)\}_{j=1}^{Q}$, where Q is the number of model states, and subsequently found the covariance matrices of the low-dimensional projections of the sampled data. We observe that the QGMR model fits the data very well, which is not always the case for the GMR method. We also observe

that in those segments where the demonstration trajectories differ from each other, there is also a high uncertainty of the model fit.

VI. Conclusion

In this paper, we presented a quantum-statistical approach toward trajectory-based robot learning by demonstration. The proposed approach is based on an extension of conventional GMR formulations, effected by introducing the concept of quantum states, which can be constructed by superposing conventional GMR model states in a linear fashion. To derive our model, we reformulated the expression of the likelihood of conventional GMR models into a special form (diagonal) density matrix, and we further showed that this matrix can be generalized into a more generic, nondiagonal form.

The so-obtained quantum GMR model was applied to yield a quantum-statistical approach toward robot learning by demonstration, and its efficacy was illustrated by considering a number of demanding robot learning by demonstration scenarios, with its performance being compared with state-of-the-art robot learning by demonstration methodologies. As we showed, our method allows for a significant performance increase, while imposing computational requirements similar to its alternatives, since prediction under all these approaches eventually reduces to a sum of linear regression models. Based on our results, we can definitively conclude that the proposed approach is especially suitable for learning complex demonstration trajectories, under both a sparse one-shot and a multishot learning setting. The MATLAB implementation of the QGMR method shall be made available through the website of the authors.

References

[1] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," in *Handbook of Robotics*, B. Siciliano and O. Khatib, Eds. Secaucus, NJ: Springer-Verlag, 2008, pp. 1371–1394.

[2] B. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robot. Auton. Syst.*, vol. 57, no. 5, pp. 469–483, 2009.

[3] Y. Demiris and B. Khadhouri, "Hierarchical attentive multiple models for execution and recognition (HAMMER)," *Robot. Auton. Syst.*, vol. 54, pp. 361–369, 2006.

[4] K. Yamane and Y. Nakamura, "Dynamics filter-concept and implementation of online motion generator for human figures," *IEEE Trans. Robot. Autom.*, vol. 19, no. 3, pp. 421–432, Jun. 2003.

[5] Z. Ghahramani and M. I. Jordan, "Supervised learning from incomplete data via an EM approach," in *Advances in Neural Information Processing System*, vol. 6. San Mateo, CA: Morgan Kaufmann, 1994, pp. 120–127.

[6] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statist. Soc., B*, vol. 39, no. 1, pp. 1–38, 1977.

[7] S. Calinon and A. Billard, "Statistical learning by imitation of competing constraints in joint space and task space," *Adv. Robot.*, vol. 23, no. 15, pp. 2059–2076, 2009.

[8] S. Calinon, F. D'halluin, E. Sauser, D. Caldwell, and A. Billard, "Learning and reproduction of gestures by imitation: An approach based on hidden Markov model and Gaussian mixture regression," *IEEE Robot. Autom. Mag.*, vol. 17, no. 2, pp. 44–54, Jun. 2010.

[9] D. Lee and Y. Nakamura, "Mimesis scheme using a monocular vision system on a humanoid robot," in *Proc. IEEE Int. Conf. Robot. Autom.*, Apr. 2007, pp. 2162–2168.

[10] S. Calinon, F. Guenter, and A. Billard, "On learning, representing and generalizing a task in a humanoid robot," *IEEE Trans. Syst., Man, Cybern., Part B*, vol. 37, no. 2, pp. 286–298, Apr. 2007.

[11] S. Vijayakumar, A. D'souza, and S. Schaal, "Incremental online learning in high dimensions," *Neural Comput.*, vol. 17, no. 2, pp. 2602–2634, 2005.

[12] J. Peters and S. Schaal, "Learning to control in operational space," *Int. J. Robot. Res.*, vol. 27, no. 2, pp. 197–212, 2008.

[13] T. Cederborg, M. Li, A. Baranes, and P.-Y. Oudeyer, "Incremental local online Gaussian mixture regression for imitation learning of multiple tasks," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Oct. 2010, pp. 267–274.

[14] M. K. Warmuth and D. Kuzmin, "Bayesian generalized probability calculus for density matrices," *J. Mach. Learning*, vol. 78, no. 1/2, pp. 63–101, 2009.

[15] L. Wolf, "Learning using the Born rule," Mass. Inst. Technol., Cambridge, MA, Tech. Rep. TR-2006-036, 2006.

[16] K. Crammer and A. Globerson, "Discriminative learning via semidefinite probabilistic models," in *Proc. 22nd Annu. Conf. Uncertainty Artif. Intell.*, 2006, pp. 98–105.

[17] C. W. Helstrom, "Quantum detection and estimation theory," *J. Statist. Phys.*, vol. 1, no. 2, pp. 231–252, 1969.

[18] K. Tanaka and T. Horiguchi, "Probabilistic, iterated and quantum-iterated computational methods in gray-level image restoration," *Interdisciplinary Inf. Sci.*, vol. 8, no. 1, pp. 33–50, 2002.

[19] K. Tanaka and K. Tsuda, "A quantum-statistical-mechanical extension of Gaussian mixture model," in *Proc. Int. Workshop Statist.-Mech. Informat.* (J. Phys.: Conf. Ser.), 2007, vol. 95, pp. 1–9.

[20] S. Chatzis, D. Kosmopoulos, and T. Varvarigou, "Signal modeling and classification using a robust latent space model based on t distributions," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 949–963, Mar. 2008.

[21] I. Sato, K. Kurihara, S. Tanaka, H. Nakagawa, and S. Miyashita, "Quantum annealing for variational Bayes inference," in *Proc. 25th Conf. Uncertainty Artif. Intell.*, 2009, pp. 479–486.

[22] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learning Res.*, vol. 3, pp. 993–1022, 2003.

[23] G. McLachlan and D. Peel, *Finite Mixture Models (Wiley Series in Probability and Statistics)*. New York: Wiley-Interscience, 2000.

[24] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer, 2006.

[25] V. N. Vapnik, *Statistical Learning Theory*. New York: Wiley, 1998.

[26] C. E. Rasmussen and C. K. I. Williams, *Gaussia Processes for Machine Learning*. Cambridge, MA: MIT Press, 2006.

[27] G. Schwarz, "Estimating the dimension of a model," *Ann. Statist.*, vol. 6, pp. 461–464, 1978.

[28] B. G. Leroux, "Consistent estimation of a mixing distribution," *Ann. Statist.*, vol. 20, pp. 1350–1360, 1992.

[29] Aldebaran robotics: The NAO robot. (2012). [Online]. Available: http://www.aldebaran-robotics.com/en/Nao.php

[30] M. Schneider and W. Ertel, "Robot learning by demonstration with local Gaussian process regression," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2010, pp. 255–260.

[31] B. Pearlmutter, "Gradient calculation for dynamic recurrent neural networks: A survey," *IEEE Trans. Neural Netw.*, vol. 6, no. 5, pp. 1212–1228, Sep. 1995.

[32] P. Zegers and M. K. Sundareshan, "Trajectory generation and modulation using dynamic neural networks," *IEEE Trans. Neural Netw.*, vol. 14, no. 3, pp. 520–533, May 2003.

[33] C. S. Myers and L. R. Rabiner, "A comparative study of several dynamic time-warping algorithms for connected word recognition," *Bell Syst. Tech. J.*, vol. 60, no. 7, pp. 1389–1409, Sep. 1981.

Sotirios P. Chatzis received the M.Sc. (five-year diploma) degree in electrical and computer engineering in 2005 and the Ph.D. degree in machine learning in 2008, both from the National Technical University of Athens, Athens, Greece. His Ph.D. research was supported by the Bodossaki Foundation, Greece, and the Greek Ministry for Economic Development, where he was awarded the Dean's scholarship for Ph.D. studies, being the best performing Ph.D. student in his class.

He currently is a Postdoctoral Researcher with the Department of Electrical and Electronic Engineering, Imperial College London, London, U.K. Over the past five years, he has first-authored 28 journal papers in the most prestigious journals of his research field. His major research interests comprise machine learning theory and methodologies with a special focus on hierarchical Bayesian models, reservoir computing, robot learning by demonstration, copulas, quantum statistics, Bayesian nonparametrics, and artificial creativity.

Dimitrios Korkinof received the M.Eng. degree in electrical and computer engineering from the Aristotle University of Thessaloniki, Thessaloniki, Greece. He graduated third in academic achievement and received a British equivalent of the first-class (Hons.) degree in 2010. He is currently working toward the Ph.D. degree with the Department of Electrical and Electronic Engineering of Imperial College London, London, U.K., where he is researching aspects of statistical machine learning theory with applications in robotics. In the first two years of his Ph.D. studies, he published 4 papers in prestigious journals of his research field.

His current academic interests include Bayesian and quantum statistics, stochastic processes, Bayesian nonparametrics and random measure theory, with particular focus on computer vision, action recognition, and robot learning by demonstration.

Yiannis Demiris (SM'08) received the B.Sc. (Honours) degree in artificial intelligence and computer science and the Ph.D. degree in intelligent robotics from the Department of Artificial Intelligence, University of Edinburgh, Edinburgh, U.K., in 1994 and 1999, respectively.

He is currently a Reader (Associate Professor) with the Imperial College London, London, U.K. He has significant expertise in cognitive systems, assistive robotics, multirobot systems, robot–human interaction, and learning by demonstration, in particular in action perception and learning. His research is funded by the U.K.'s Engineering and Physical Sciences Research Council, the Royal Society, BAE Systems, and the European Union Seventh Framework Program through projects ALIZ-E and EFAA, both addressing novel machine learning approaches to human–robot interaction. Additionally, the group collaborates with the BBC's Research and Development Department on the "Learning Human Action Models" project.

Dr. Demiris has guest edited special issues of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, specifically on Learning by Observation, Demonstration, and Imitation, and of the *Adaptive Behavior Journal on Developmental Robotics*. He has organized six international workshops on Robot Learning, BioInspired Machine Learning, Epigenetic Robotics, and Imitation in Animals and Artifacts, was the Chair of the IEEE International Conference on Development and Learning in 2007, as well as the Program Chair of the ACM/IEEE International Conference on Human–Robot Interaction in 2008. He is a member of the Institute of Engineering and Technology of Britain.