Online Learning of an Open-Ended Skill Library for Collaborative Tasks

Dorothea Koert[1], Susanne Trick[1], Marco Ewerton[1], Michael Lutter[1], and Jan Peters[1,3]

Abstract— Intelligent robotic assistants can potentially improve the quality of life for elderly people and help them maintain their independence. However, the number of different and personalized tasks render pre-programming of such assistive robots prohibitively difficult. Instead, to cope with a continuous and open-ended stream of cooperative tasks, new collaborative skills need to be continuously learned and updated from demonstrations. To this end, we introduce an online learning method for a skill library of collaborative tasks that employs an incremental mixture model of probabilistic interaction primitives. This model chooses a corresponding robot response to a human movement where the human intention is extracted from previously demonstrated movements. Unlike existing batch methods of movement primitives for human-robot interaction, our approach builds a library of skills online, in an open-ended fashion and updates existing skills using new demonstrations. The resulting approach was evaluated both on a simple benchmark task and in an assistive human-robot collaboration scenario with a 7DoF robot arm.

I. INTRODUCTION

The expected demographic change is a major challenge for society as a significantly growing elderly population will require substantially increased assistance [1]. Intelligent robot assistants could improve quality of life, assist in cooperative tasks and help the elderly maintain their partial independence while staying in their own homes. However, such cooperative robot assistants need to be able to adapt to individual needs and a multitude of tasks at hand, rendering pre-programming of all possible tasks prohibitively difficult in practice.

An intuitive way for non-expert users to teach personalized skills to the robot is required for which Learning from Demonstration (LfD) is considered a promising approach [2]. In particular, personal cooperative robots require the ability to learn multiple different tasks and adapt them to varying contexts. Hereby, the robot should be able to learn to distinguish between different human intentions and react with the matching interaction patterns. However, human motions exhibit a high variability [3]. To model and react to this variability in human motions, a probabilistic approach to cooperative skill learning is needed, which takes the variations in human motions during demonstrations and execution time into account. Additionally, the ability to incorporate

[1]Intelligent Autonomous Systems, TU Darmstadt, Germany
[3]MPI for Intelligent Systems, Tuebingen
{koert,ewerton,lutter,peters} @
ias.tu-darmstadt.de, ,
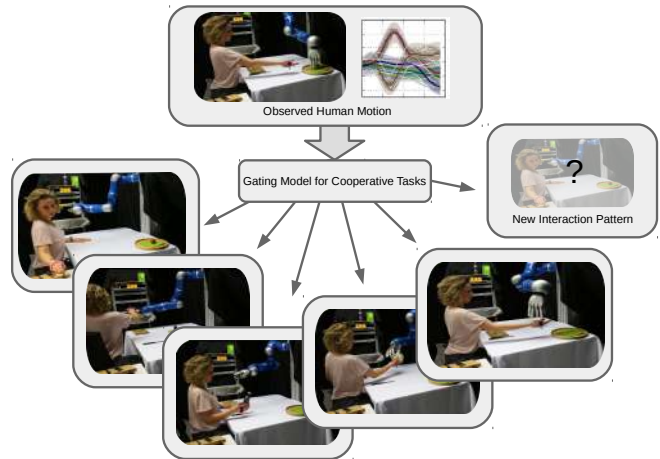susanne_gabriele.trick @ stud.tu-darmstadt.de

Fig. 1. Intelligent robot assistants should learn multiple personalized cooperative tasks from a continuous and open-ended stream of new demonstrations. To this end, we propose a novel approach for online and open-ended learning of a mixture model of probabilistic interaction primitives. In particular, our approach updates existing cooperative tasks from new demonstrations and extends the collaborative skill library for new tasks when needed. Hereby, our model chooses a robot response to an observed human motion based on prior demonstrations while considering variance in the demonstrations as well as coupling between human and robot motions.

new demonstrations in an online and open-ended fashion is desirable.

In this paper, we propose an online learning method for collaborative skills that employs an incremental mixture model of probabilistic interaction primitives for online learning of collaborative skills from demonstrations. Probabilistic interaction movement primitives (Interaction ProMPs) [4] are used as a representation of cooperative skills, that can capture correlations between human and robot movements as well as the inherent variance. While Interaction ProMPs have already been used in scenarios with multiple contexts [5], to the best of the author's knowledge no method for online and open-ended learning of multiple Interaction ProMPs from demonstrations has been proposed. However, for personalized robot assistants it is crucial to open-endedly learn new tasks and continuously update existing cooperative skills with new demonstrations. In particular, in such an open-ended scenario the total number of cooperative tasks cannot be known beforehand and thus needs to be extended during the learning process. In this paper, this problem is tackled with an online learning method to update and extend a library of cooperative skills. This library allows inferring the human intention from previous demonstrations and is used to choose the appropriate robot response to a human motion or request more demonstrations in case of high uncertainty. The employed probabilistic movement representation is capable of

representing the abundant variance in demonstrations and can adapt to variation in human motions during execution. The resulting approach is able to distinguish between multiple different interaction tasks, to update existing skills with new demonstrations and, if necessary, to extend the interaction library for new tasks. In contrast to prior work on learning a Mixture of Interaction primitives [5], our new approach does not rely solely on demonstrations which are available at the first training time but can integrate new demonstrations and tasks into the collaborative skill library over multiple training sessions.

The rest of this paper is organized as follows: First, we discuss related work in Section II. Next, in Section III, we provide a short overview on the existing approach of Batch Learning for a Mixture of Interaction Primitives and afterwards introduce our novel approach for Online Open-Ended Learning of a Mixture of Interaction Primitives. In Section IV we evaluate this new approach on 2D trajectory data and on a collaborative scenario where a robot assists a human in making a salad. Finally, we conclude with Section V and discuss ideas for future work.

II. RELATED WORK

Learning cooperative tasks between humans and robots from demonstration is a popular approach as it enables also non-expert users to teach personalized skills to robots [6], [7], [8], [9]. When learning from demonstrations, the concept of movement primitives offers a lower dimensional representation of trajectories [10], [11], [12], [8]. In particular, Probabilistic Interaction Movement Primitives (Interaction ProMPs) [13], [4], [9] offer a probabilistic representation to model inherent correlations in the movements of two actors, such as human and robot, from coupled demonstrated trajectories.

However, to achieve a personalized cooperative robot it is desirable to learn multiple different cooperative tasks and decide on their activation depending on the context or on the human intention [14], [15], [16]. To this end, an approach that deploys Gaussian Mixture Models (GMMs) and Expectation Maximization to learn multiple Interaction ProMPs from unlabeled demonstrations has been introduced [5]. This approach considers batch data, i.e. assuming the availability of all data points during training. This limits its application to settings where the number of tasks does not change after training and no new demonstration trajectories need to be integrated. Moreover, such batch learning prevents scalability as the computation time and memory requirements become infeasible for large skill libraries or datasets [17]. Various approaches outside the human-robot interaction (HRI) scope have been addressing these problems.

Initially, the machine learning communities have proposed incremental learning approaches for Gaussian Mixture Models. Some approaches propose updating a GMM with complete new model component datasets [18] or assume the incoming data points to be time-coherent [19]. Incremental Gaussian Mixture Model learning introduced a way to continuously learn a GMM from an incoming data stream while not fixing the number of total components beforehand [20], [21]. Another two-level approach introduces methods for splitting and merging of GMM components [22].

Updating of robotic movement representations online from new demonstrations has also been used for incremental learning of extensions of GMMs for gesture imitation [17], updating Gaussian Processes from demonstrations and thereby reducing the movement variance [23] or incremental updating of task-parameterized Gaussian Mixture Models [24].

While all these works focus on updating multiple existing movement representations, in a long-term setting adding new tasks is also important. Approaches that also add new components when needed have been proposed in the context of online updating of task-parameterized semi-tied hidden semi-Markov models for manipulation tasks [25], learning full-body movements [26], a bootstrapping cycle for automatic extraction of primitives from complex trajectories [27] or robot table tennis [28]. However, while we draw inspiration from the aforementioned related work, in an HRI scenario it is additionally desirable to consider the inherent coupling and variance in human and robot motions in the demonstrations.

III. INCREMENTAL INTERACTION PRIMITIVES

In this Section we introduce a new approach to continuously learn and update multiple cooperative skills from demonstrations.

Here, demonstrations are given in form of coupled human and robot trajectories $d_n = \{\tau_n^h, \tau_n^r\}$, where τ_n^h can e.g. be a sequence of human wrist positions and τ_n^r can e.g. be a sequence of robot joint positions. To learn multiple cooperative tasks from these demonstrations in an online open-ended fashion we introduce a model that is inspired by the Mixture of Experts architecture [29] and consists of two intertwined parts. On the one hand, we use the human trajectories from the demonstrations to train and update a gating model, which will later be used to decide between different cooperative tasks. In addition, we train probabilistic models to generate appropriate robot response trajectories. Here, we deploy Interaction ProMPs [4], as they are able to capture the inherent correlation in robot and human motions from the demonstrations. Figure 2 summarizes our approach to achieve training of this mixture model in an online and open-ended fashion.

In the following, we briefly describe the previously proposed batch-based, stationary Mixture of Interaction ProMPs in Section III-A. Next, we present our novel approach to learn a mixture model of Probabilistic Movement Primitives in an online and open-ended fashion in Section III-B. Finally, in Section III-C we show how the obtained library of multiple interaction ProMPs and the corresponding gating model can be deployed in an HRI scenario.

A. Batch Learning for Mixture of Interaction ProMPs

Probabilistic Movement Primitives (ProMPs) [12] represent demonstrated movements in the form of distributions over trajectories. In order to obtain this distribution, the trajectories are first approximated by a linear combination

of basis functions ϕ. More precisely, a joint position q_t at time step t can be represented as

$$q_t = \phi_t^T w + \epsilon, \qquad (1)$$

where ϕ_t contains N basis functions ϕ evaluated at time step t, w is a weight vector and ϵ is a zero-mean Gaussian noise. The choice of basis functions ϕ depends on the type of demonstrated movements.

The weight vector w for each demonstrated trajectory is computed with Ridge Regression. For multiple recorded demonstrated trajectories, a Gaussian distribution over the weight vectors $p(w) = \mathcal{N}(\mu_w, \Sigma_w)$ can then be obtained with Maximum Likelihood Estimation. Since the number N of basis functions is usually much lower than the number of time steps of recorded trajectories, the distribution $p(w)$ can be seen as a compact representation of the demonstrated movements, which accounts for variability in the execution. In particular, ProMPs offer a representation that allows for operations from probability theory to specify goal or via-points, correlate different degrees of freedom via conditioning and combine different primitives through blending [12].

An Interaction ProMP [4] is a ProMP that uses a distribution over the trajectories of at least two interacting agents. The demonstrations are now given in the form of a stacked vector for the observed and the controlled agent $q = [q^o, q^c]^T$, where q^o denotes the demonstrated trajectories for the observed agent and q^c denotes the demonstrated trajectories of the controlled agent. Respectively, the weight vector is also represented in an augmented form $\bar{w} = [w_o^T, w_c^T]^T$. Given a set of demonstrations, a distribution over multiple stacked weight vectors can be obtained just as previously described such that $p(\bar{w}) = \mathcal{N}(\mu_{\bar{w}}, \Sigma_{\bar{w}})$. Given a sequence $\mathcal{D}$ of positions of the observed agent (e.g. human), Interaction ProMPs provide methods to infer a corresponding (most likely) trajectory of the controlled agent (robot) [4].

The previously proposed batch learning for Mixture of Interaction ProMPs [5] is an extension to Interaction ProMPs that allows to learn several different interaction patterns from unlabeled demonstrations by applying Gaussian Mixture Models (GMMs), where each mixture component represents one interaction pattern. The Mixture of Interaction Primitives is hereby learned from batch data and the number of components needs to be fixed beforehand. In the case of K different interaction patterns, the distribution over the weight vectors $\bar{w}$ is

$$p(\bar{w}) = \sum_{k=1}^{K} p(k)p(\bar{w}|k) = \sum_{k=1}^{K} \alpha_k \, \mathcal{N}(\bar{w}|\mu_k, \Sigma_k), \quad (2)$$

where α_k is the $k-th$ mixture weight that can be prior (if not learned) or posterior (if learned from given data), μ_k is the mean and Σ_k the covariance matrix of the k-th component.

The parameters of the GMM are hereby learned in the weight space using the Expectation Maximization (EM) algorithm. However, since this approach assumes that all demonstrations are available at the learning time the number of components K remains fixed after learning. This means
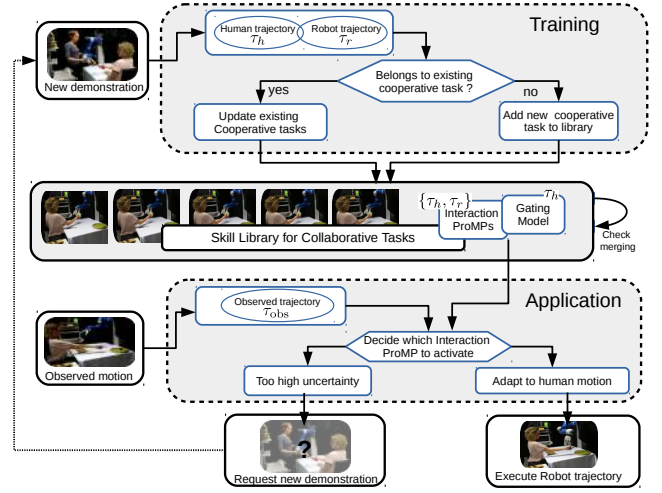

Fig. 2. We introduce a novel approach for online and open-ended learning of a mixture model for cooperative tasks. During training, demonstrations are given in form of trajectories of a human demonstrator τ^h and corresponding trajectories of a robot arm τ^r, that are obtained via kinesthetic teaching. From these demonstrations, we update or extend our library for cooperative tasks that consists of a gating model and multiple corresponding Interaction ProMPs. During runtime, the gating model decides on activation of particular Interaction ProMPs that we subsequently adapt to the variance in the observed motion. If the gating model is too uncertain about activation of Interaction ProMPs the robot can request more demonstrations.

that the parameters of the GMM need to be computed using all the previous demonstrations in case a new demonstration is available. Moreover, a GMM with a fixed number of mixture components cannot cope with new interaction patterns.

B. Online Open-Ended Mixture of Interaction ProMPs

We propose a new method to achieve online learning of cooperative tasks in an open-ended fashion. Hereby, demonstrations are given in the form of robot and human trajectories $\{\tau^r, \tau^h\}$. First, we compute a corresponding representation with weight vectors as introduced in Section III-A. Here, we consider that the human trajectory is of dimensions $D_h \times T$ where D_h is the degree of freedom of the observations (e.g. in case of observing the wrist position $D_h = 3$) and T is the number of time steps and the robot trajectory is of dimensions $D_r \times T$ where D_r is the degrees of freedom of the robot (e.g. $D_r = 7$ in case of a 7DoF robot arm). For N basis functions ϕ we compute the matrix $\Phi = [\phi_0, ...\phi_t, ..., \phi_T]$ with dimension $N \times T$. In this work, Gaussian basis functions evenly spaced along the time axis are an appropriate choice due to the stroke-based movements. We then compute the weight vectors as a lower dimensional representation of the trajectories where we first compute the weight vectors for each dimension $\tilde{w}$ as

$$\begin{bmatrix} \tilde{w}_1^h \\ ... \\ \tilde{w}_{D_h}^h \\ \tilde{w}_1^r \\ ... \\ \tilde{w}_{D_r}^r \end{bmatrix} = (\Phi\Phi^T + \beta I)^{-1} \Phi \begin{bmatrix} \tau_h \\ \tau_r \end{bmatrix}, \qquad (3)$$

where β is a factor for Ridge Regression and I is an identity matrix. In experimental evaluation, we found that

Algorithm 1:

input: $\boldsymbol{\Sigma}_{init}^g = \boldsymbol{I}, T_{\text{nov}}, v_{\min}$;
while new data $\boldsymbol{\tau}_n^h, \boldsymbol{\tau}_n^r$ **do**
    compute $\boldsymbol{w}_n^h, \boldsymbol{w}_n^r$ from $\boldsymbol{\tau}_n^h, \boldsymbol{\tau}_n^r$ Eq (3),(4);
    compute $p(\boldsymbol{w}_n^h|k)$ $\forall k$ according to Eq (5).;
    if $p(\boldsymbol{w}_n^h|k) < T_{\text{nov}}$ $\forall k$ **then**
        add new component Eq. (9);
        k++;
    else
        compute $p(k|\boldsymbol{w}_n^h)$ $\forall k$;
        update model parameters $\forall k$ Eq.(6), (7), (8);
        if $v_k > v_{\min}$ **then**
            check merge and if any merge Eq. (10);
        end
    end
end

normalizing the trajectory data within a fixed range before transforming it into the weight space yields overall better result. Subsequently, we compute the stacked weight vectors

$$\boldsymbol{w}_h = [\tilde{\boldsymbol{w}}_1^h, ..., \tilde{\boldsymbol{w}}_{D_h}^h] \quad \text{and}$$
$$\boldsymbol{w}_r = [\tilde{\boldsymbol{w}}_1^r, ..., \tilde{\boldsymbol{w}}_{D_r}^r]. \tag{4}$$

From these demonstrations, now represented in form of $\{\boldsymbol{w}^r, \boldsymbol{w}^h\}$, we learn the two intertwined parts of our model: The gating model that decides on the cooperative tasks based on human motions and multiple corresponding Interaction ProMPs that can subsequently generate a corresponding robot response. For the gating model we train a Gaussian Mixture Model (GMM) only on the weights of the human trajectories $\boldsymbol{w}^h$, as at runtime only the human motion will be observed when the system needs to decide on the particular cooperative task and the response of the robot. In parallel to the gating model, the corresponding Interaction ProMPs are trained with the augmented weight vector $\bar{\boldsymbol{w}}$ of human and robot trajectories to model the correlations in the motions.

We assume that new training data needs to be integrated continuously and that we do not know beforehand the number of different collaborative tasks that might be shown to the robot during long-term training. To this end, we use Incremental Gaussian Mixture Models [20] to achieve the continuous integration of new demonstrations. Here, we update the gating model and the parameters of the Interaction ProMPs in an Expectation Maximization fashion.

In the Expectation step we compute the responsibilities λ_{kn} of the existing cooperative task k for a new demonstration $\{\boldsymbol{w}_n^h, \boldsymbol{w}_n^r\}$, that is the probability of a new demonstration to belong to an already known cooperative task

$$\lambda_{kn} := p(k|\boldsymbol{w}_n^h) = \frac{p(\boldsymbol{w}_n^h|k)p(k)}{p(\boldsymbol{w}_n^h)} = \frac{\alpha_k \mathcal{N}(\boldsymbol{w}_n^h|\boldsymbol{\mu}_k^g, \boldsymbol{\Sigma}_k^g)}{\sum_{j=1}^K \alpha_j \mathcal{N}(\boldsymbol{w}_n^h|\boldsymbol{\mu}_j^g, \boldsymbol{\Sigma}_j^g)}, \tag{5}$$

where $\boldsymbol{\mu}_k^g$ and $\boldsymbol{\Sigma}_k^g$ are respectively the mean and covariance matrix of the k-th component of the gating and α_k are the mixture component weights. In the Maximization step, we use the responsibilities to recursively update the

parameters of the gating model as well as the parameters of the already learned Interaction ProMPs. For each already learned Interaction ProMP k we first compute

$$v_k = v_k + 1, \qquad s_k = s_k + \lambda_{kn},$$
$$\gamma_k = \frac{\lambda_{kn}}{s_k}, \qquad \tilde{\gamma}_k = \gamma_k + \exp(-s_k)\lambda_{kn}, \tag{6}$$

where v_k is the age of the k-th component and s_k represents the trajectories the component already modeled well. We then update the parameters of the gating model

$$\boldsymbol{\mu}_k^g = \boldsymbol{\mu}_k^g + \gamma_k(\boldsymbol{w}_n^h - \boldsymbol{\mu}_k^g),$$
$$\boldsymbol{C}_k^g = (1 - \tilde{\gamma}_k)\boldsymbol{C}_k^g + \tilde{\gamma}_k(\boldsymbol{w}_n^h - \boldsymbol{\mu}_k^g)(\boldsymbol{w}_n^h - \boldsymbol{\mu}_k^g)^T -$$
$$(\tilde{\gamma}_k - \gamma_k)(\boldsymbol{\mu}_k^g - \boldsymbol{\mu}_k^{g,old})(\boldsymbol{\mu}_k^g - \boldsymbol{\mu}_k^{g,old})^T,$$
$$\alpha_k = \frac{s_k}{\sum_{j=1}^K s(j)}, \tag{7}$$

where the formulas correspond to the formulas in the incremental GMM [20], except that we introduce $\tilde{\gamma}_k$ to achieve that during the first demonstrations the covariance is shifted faster away from the (possibly wrong) initialization. Additionally, we compute the updated parameters of the corresponding Interaction ProMPs

$$\boldsymbol{\mu}_k^e = \boldsymbol{\mu}_k^e + \gamma(\bar{\boldsymbol{w}}_n - \boldsymbol{\mu}_k^e),$$
$$\boldsymbol{C}_k^e = (1 - \tilde{\gamma}_k)\boldsymbol{C}_k^e + \tilde{\gamma}_k(\bar{\boldsymbol{w}}_n - \boldsymbol{\mu}_k^e)(\bar{\boldsymbol{w}}_n - \boldsymbol{\mu}_k^e)^T -$$
$$(\tilde{\gamma}_k - \gamma_k)(\boldsymbol{\mu}_k^e - \boldsymbol{\mu}_k^{e,old})(\boldsymbol{\mu}_k^e - \boldsymbol{\mu}_k^{e,old})^T, \tag{8}$$

where $\boldsymbol{\mu}_k^e$ is the mean of the k-th Interaction ProMP and $\boldsymbol{\Sigma}_k^e$ is the covariance matrix of the k-th Interaction ProMP. Whenever $p(\boldsymbol{w}_n^h|k)$ is below a threshold T_{nov} for all existing K components we initialize a new component with

$$\boldsymbol{\mu}_{K+1}^g = \boldsymbol{w}_n^h, \qquad \boldsymbol{\Sigma}_{K+1}^g = \boldsymbol{\Sigma}_{init}^g,$$
$$\boldsymbol{\mu}_{K+1}^e = \boldsymbol{w}_n, \qquad \boldsymbol{\Sigma}_{K+1}^e = \boldsymbol{\Sigma}_{init}^e,$$
$$v_k = 1, \qquad s_k = 1, \tag{9}$$

If a component has reached a certain age $v_k > v_{min}$ we check also for merging of components to ensure that no unnecessary components are maintained. Therefore, we compute the probability of the mean of a cluster j to belong to a cluster i as $p(\boldsymbol{\mu}_j|i) = \mathcal{N}(\boldsymbol{\mu}_j|\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i)$ and decide on merging if $p(\boldsymbol{\mu}_j|i) > T_{\text{nov}}$. Once we decide on candidates i, j for merging we recompute the joined mean and covariance

$$\boldsymbol{\mu}_{ij} = \frac{s_i\boldsymbol{\mu}_i + s_j\boldsymbol{\mu}_j}{s_i + s_j}$$
$$\boldsymbol{\Sigma}_{ij} = \frac{s_i^2\boldsymbol{\Sigma}_i + s_j^2\boldsymbol{\Sigma}_j + (s_i\boldsymbol{\mu}_i + s_j\boldsymbol{\mu}_j)s_i\boldsymbol{\mu}_i + s_j\boldsymbol{\mu}_j^T}{(s_i + s_j)^2}$$
$$- \boldsymbol{\mu}_{ij}\boldsymbol{\mu}_{ij}^T. \tag{10}$$

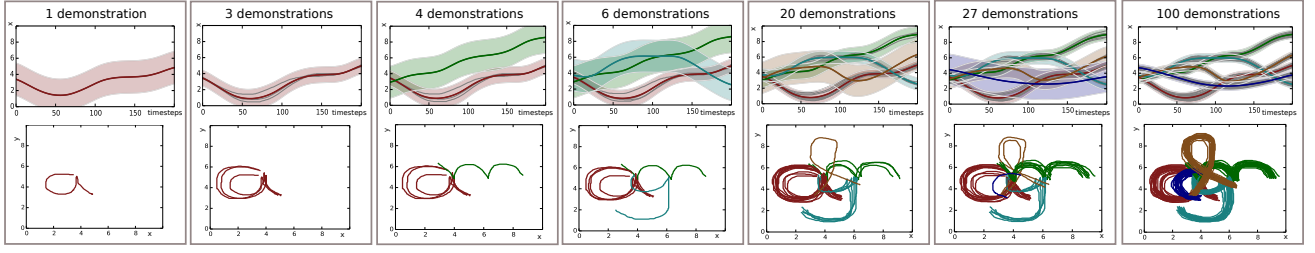Algorithm 1 summarizes our approach for online learning of a gating model and multiple Interaction ProMPs.

Fig. 3. We evaluate our approach first on a task of learning ProMPs of multiple hand-drawn letter trajectories when the demonstrations are provided incrementally and no batch data are stored. The intermediate results during the training of the ProMP library are shown in the upper row while the accumulated demonstrations are shown in the lower row. In the upper row, the shaded area represents two times the standard deviation while the solid lines show the mean and the demonstrated trajectories are shown as gray lines. Here, our approach successfully updates existing components with new demonstrations and adds new components when required. In particular, for more demonstrations per letter, the influence of the initial covariance matrix diminishes and the components covariance converges to the covariance of the demonstrations.

C. A Skill Library for Collaborative Tasks

To demonstrate the use of the learned probabilistic mixture model for cooperative tasks we assume we are now observing the human and obtain an observation $\boldsymbol{w}_*^h$. To determine the most probable cluster given the observations we need to model the posterior of the cluster given the observation

$$p(k|\boldsymbol{w}_*^h) = \lambda_{k*} \tag{11}$$

where λ_{k*} is the responsibility of the k-th cluster for the observation $\boldsymbol{w}_*^h$ as defined in Equation (5).

For a given observation $\boldsymbol{w}_*^h$ we can now infer the most likely Interaction ProMP k^* using our probabilistic gating model

$$k^* = \arg\max_k \quad p(k|\boldsymbol{w}_*^h). \tag{12}$$

If the responsibility of all components is smaller than the novelty threshold T_{nov} the robot does not execute a response but asks the user for new demonstrations that get subsequently included in the library as described in Section III-B. Otherwise, we condition the chosen Interaction ProMP on the observed trajectory to infer the corresponding robot response. Therefore, the observation $\boldsymbol{o}$ is used to obtain a posterior distribution over the weights. The posterior is again Gaussian with mean $\boldsymbol{\mu}_{new}$ and covariance matrix $\boldsymbol{\Sigma}_{new}$

$$\boldsymbol{\Lambda} = \boldsymbol{\Sigma}_{\boldsymbol{k}^*} \boldsymbol{H}_t (\boldsymbol{\Sigma}_{\boldsymbol{o}} + \boldsymbol{H}_t^{\mathrm{T}} \boldsymbol{\Sigma}_{\boldsymbol{k}^*} \boldsymbol{H}_t)^{-1}$$
$$\boldsymbol{\mu}_{new} = \boldsymbol{\mu}_{\boldsymbol{k}^*} + \boldsymbol{\Lambda}(\boldsymbol{o} - \boldsymbol{H}_t^{\mathrm{T}} \boldsymbol{\mu}_{\boldsymbol{k}^*})$$
$$\boldsymbol{\Sigma}_{new} = \boldsymbol{\Sigma}_{\boldsymbol{k}^*} - \boldsymbol{\Lambda} \boldsymbol{H}_t \boldsymbol{\Sigma}_{\boldsymbol{k}^*}, \tag{13}$$

where $\boldsymbol{\Sigma}_{\boldsymbol{o}} = \boldsymbol{I}\sigma_o$ is the observation noise and $\boldsymbol{H}_t$ is the observation matrix as defined in [4]. More details can be found in [9]. To obtain a corresponding robot motion we execute the mean robot trajectory of this posterior.

IV. EXPERIMENTAL EVALUATION

We evaluate our approach on 2D trajectory data and on a collaborative scenario with a 7DoF robot arm. For both, we show the qualitative applicability and evaluate the quantitative convergence w.r.t. to a baseline. In addition, we demonstrate that the proposed approach can learn personalized libraries for collaborative tasks for different persons and report successful task completion via the decision accuracy of our gating model.
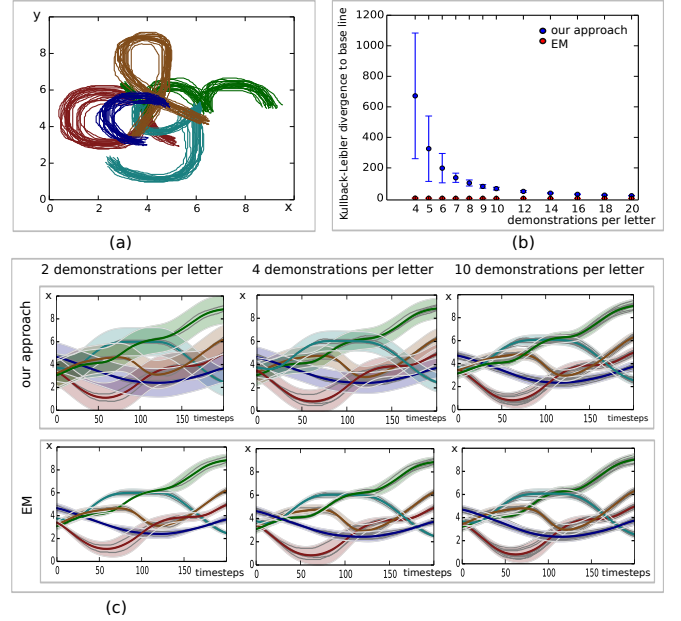

Fig. 4. (a) The demonstrations in the first experiment are given in the form of multiple hand-drawn letter trajectories. (b) We compare our approach against an EM approach, where for both we compute the KL-divergence to a baseline solution from labeled data. For increasing number of samples per letter our approach converges against the EM solution, while additionally being able to continuously integrate new demonstrations. (c) For fewer samples per letter the covariance of the components in our approach is governed by the initial covariance. For increasing number of samples per letter the covariances converge to the underlying data covariances and result in comparable results to the EM approach.

A. 2D trajectory data

For the 2D trajectory data experiment, demonstrations are given in the form of multiple hand-drawn letters, as illustrated in Figure 4 (a). Here, we learn a library of Probabilistic movement primitives in an incremental fashion. The system never has access to the whole training dataset at once, but only one new unlabeled demonstration is provided at each update step. The general procedure is shown in Figure 3, where the upper row shows the x-dimension of the learned library and the lower row the accumulated demonstrations. Initially, a single "a" is demonstrated and the first skill is added, with the initial covariance $\boldsymbol{\Sigma}_{init}$. Afterwards, additional "a"s are demonstrated, recognized and used to update the mean and covariance of the corresponding cluster. Once a new demonstration, i.e. the letter "m", is recognized to not

Fig. 5. (a) We evaluate our approach in a collaborative scenario where a robot (A) assists a human (B) in making a salad. Hereby the robot can hand over the board (D) the dressing (C), a tomato (F), the bowl (E) or assist with a standup motion. (b) The demonstrations are recorded as human and robot trajectories where the robot is moved in kinesthetic teaching mode and the wrist trajectory is tracked with a motion capturing system. (c) We again compare the KL-divergence of both our approach and an EM approach to a baseline from labeled data. For an increasing number of demonstrations per task our approach converges to the EM solution but requires less recomputation and memory as new demonstrations arrive.
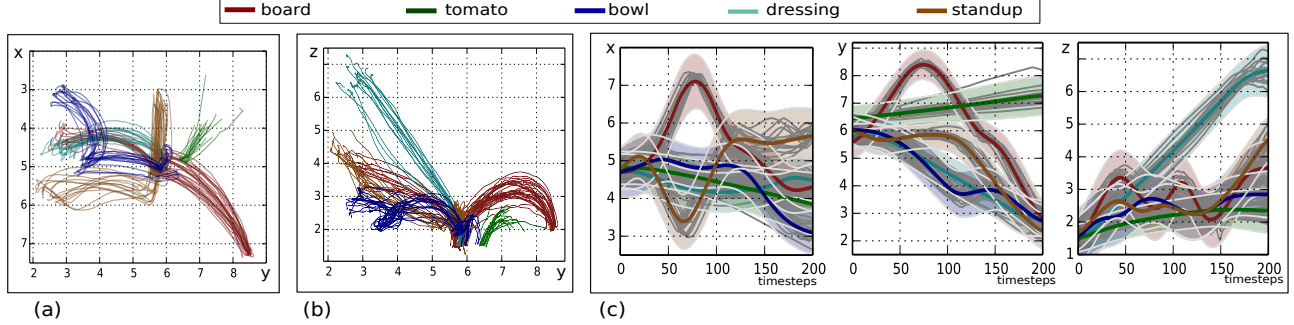

Fig. 6. For a human test subject (subject 1) we recorded 15 demonstrations per task. The trajectories are shown from a top-down view (a) and from a front view (b). On the demonstrations, we train a collaborative task library, where we do not provide all demonstrations as batch data but incrementally add more demonstrations. (c) shows the resulting gating model, which corresponds to the human trajectory part of the Interaction ProMPs. The shaded area represents two times the standard deviation while the solid lines show the mean. The demonstrated trajectories are depicted as gray lines.

belong to the existing cluster a new cluster is generated. With an increasing number of samples, the variance converges to the variance of the demonstrations as the impact of the initialization covariance decreases. The final skill library consists of five clusters representing the different letters. Please note that in this experiment $\boldsymbol{\mu}_k^g, \boldsymbol{\Sigma}_k^g = \boldsymbol{\mu}_k^e, \boldsymbol{\Sigma}_k^e$. We evaluate the approach in a collaborative setting later in Section IV-B. To demonstrate that the library learned with our new approach using the incremental processing of demonstrations converges to the solution of EM with batch learning, we compare the resulting skill libraries first qualitatively as shown in Figure 4 (c) and quantitatively using the Kullback-Leibler Divergence to a baseline as shown in Figure 4 (b). Qualitatively speaking, our approach (Figure 4 (c), upper row) represents all different letters as individual clusters and the trajectory means of the mixture model components match the means learned with EM in batch mode (Figure 4 (c), bottom row). While for fewer samples per letter the trajectory covariances learned with our approach are dominated by $\boldsymbol{\Sigma}_{init}$, with increasing number of samples per letter the trajectory covariances converge to the covariances of the EM solution as the influence of the initial covariance decreases. The same behavior can also be observed in the quantitative comparison. Hereby we compute the Kullback-Leibler (KL) divergence of our approach and EM to a baseline, computed with Maximum Likelihood estimation from labeled data. The KL-divergence of our approach is averaged over 100 trials, where the order of demonstrations is randomly permuted. In the batch EM case, we provided the method with the correct number of components, while in our approach the
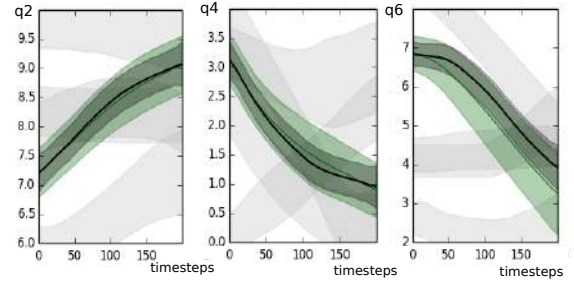

Fig. 7. Once trained with demonstrations our model can subsequently be used to produce a corresponding robot response to an observed human trajectory. Therefore, first the gating model decides which of the Interaction Primitives (light gray) to activate (green). The activated primitive is subsequently adapted to the variance in the observed human trajectory via conditioning (dark gray). The plots show joints q2, q4 and q6 of the robot arm where the shaded area represents two times the standard deviation while the solid lines show the mean.

algorithm had to find the correct number of components by itself. Figure 4 (b) shows that the KL-divergence between the solution of our approach and the baseline is large for fewer samples and decreases with increasing number of samples. The high variance in the KL-divergence for few samples is expected as the KL-divergence is sensitive to the entropy of the ground truth model, which clearly depends on the selected demonstrations. The variance also shrinks as the entropy converges for multiple samples. The experiments show that our novel approach achieves results comparable to those of an EM approach. However, in contrast to EM, our approach does not require all data in batch mode but can incrementally learn and update its models from new demonstrations in an online and open-ended fashion.

B. Learning Cooperative Tasks with a Robotic Arm

In this experiment, the proposed approach is tested in a collaborative scenario, where a robot is supposed to assist an elderly person in making a salad. The robot assists the person by first observing and recognizing the human action and second determining, adapting and executing an assistive response based on prior demonstrations. For the salad scenario, shown in Figure 5 (a), five different cooperative tasks are required, namely:

- Board: The robot hands over the cutting board after the human grasped the knife.
- Tomato: The robot passes the tomato when the human reaches for the tomato.
- Bowl: The robot passes the salad bowl when the human reaches for the bowl.
- Dressing: The robot gets the salad dressing from the shelf after the human reached for the dressing.
- Standup: The robot supports the standup motion

Each of the cooperative tasks is demonstrated separately and multiple times. The robot response is shown using kinesthetic teaching, while the human action is recorded using motion capturing markers on the wrist. This teaching procedure is shown for the bowl task in Figure 5 (b).

In an initial experiment, 15 demonstrations are recorded for every task with a human test subject (subject 1). The resulting human trajectories are shown in top-down and front view in Figure 6 (a), (b). From the demonstrations our approach learns a library for cooperative skills consisting of a gating model and multiple corresponding Interaction ProMPs. Hereby, the demonstrations are not provided as batch data but incrementally and the data are not stored. An example of the gating model for the human trajectories (which corresponds to the human part of the Interaction ProMPs) is shown in Figure 6 (c). The five different skill clusters are clearly visible. Figure 5 (c) shows that similarly to the letter experiment, the averaged KL-divergence w.r.t. to the ground truth solution learned from labeled data decreases with the number of demonstrations per task and is for more demonstrations per task comparable to the EM solution, while it requires fewer computations and memory.

In the interactive setting, the robot determines and adapts its response to the human movement based on prior demonstrations. Such an adapted robot response is shown for the bowl task in Figure 5 (b) in the bottom. The adaptation of the robot response is achieved by conditioning the Interaction ProMP on the observed human wrist trajectory as described in Section III-C. An example of such an adaptation can be seen in Figure 7, for the tomato task.

To further demonstrate the applicability and robustness of the proposed approach, we conducted more experiments with different subjects and identical hyperparameters. For each subject individual demonstrations are recorded and a corresponding personalized skill library is incrementally learned. To evaluate the performance, the classification accuracy for recognizing the correct cooperative task is evaluated by k-fold cross-validation. For subject 1 we use a training set of 10 demonstrations per task and test on 5 demonstrations per task, while all other subjects use 4 demonstrations per task for training and 1 demonstration per task for testing. The classification results averaged over 100 test and train sets are shown in Table 1. The first value corresponds to the percentage of successful classifications, the second to the percentage of wrong classifications (such as classifying tomato as bowl) and the third to the percentage of classifications as unknown. The results reveal that even though our approach works well for six of the subjects the classification accuracies for the other 4 subjects vary between tasks. In particular, for the board task subject 3, 7 and 9 have some movements with a high variance to the training set that are therefore classified as unknown. However, the classification as unknown does not yield a wrong robot response and the robot would only ask for a new demonstration. Only for subject 10 the robot misclassifies the stand-up skill. In addition to the classification accuracy, Table 2 shows the number of learned components for the individual subjects to provide some insights about the personalized skill libraries. Here, depending on a variance of the subject's movement a single skill can be represented by multiple clusters since we used the same hyperparameters for all subjects and did not tune them individually. Additional clusters do not cause wrong classifications but can lead to unknown classification as shown for subjects 3, 7 and 9. The results for subject 10 illustrate that the wrong classifications for the standup task for this subject were caused by too few learned clusters.

In all the experiments we assumed a fixed observation time of the human motion, after which the gating model decides on a particular cooperative task. For future work we consider to replace this fixed observation time with a more flexible observation time to allow for temporal correlation between human and robot motions.

V. CONCLUSIONS

In this paper, we introduce a novel approach to learn a mixture model of probabilistic interaction primitives in an online and open-ended fashion. In contrast to prior work which focussed on batch learning of Mixture of Interaction Primitives, our approach is able to update existing interaction primitives continuously from new data and extend a cooperative tasks library with new interaction patterns when needed. Experimental evaluation on a collaborative scenario with a 7DoF robot arm showed that our approach is able to learn multiple different collaborative tasks from unlabeled training data and generate corresponding robot motions, based on prior demonstrations. Evaluations with 10 human subjects showed that our approach successfully learned a personalized collaborative library for the majority of subjects.

However, since the experiments on different subjects indicate that motion data do not work equally well for all subjects and tasks, we are currently investigating how to include other modalities such as gaze direction or voice commands in our gating model. Another interesting line for future work is to train a cooperative tasks library across subjects to

TABLE I

CLASSIFICATION ACCURACY

	board	tomato	dressing	standup	bowl
subject1	1.0	1.0	1.0	1.0	0.99 (0|0.01)
subject2	1.0	1.0	1.0	1.0	1.0
subject3	0.85 (0|0.15)	1.0	1.0	0.72 (0|0.28))	1.0
subject4	1.0	1.0	1.0	1.0	1.0
subject5	1.0	1.0	0.99 (0.01|0)	1.0	1.0
subject6	1.0	1.0	1.0	1.0	1.0
subject7	0.77 (0|0.23)	1.0	0.88 (0|0.12)	1.0	1.0
subject8	1.0	1.0	1.0	1.0	1.0
subject9	0.84 (0|0.16)	1.0	1.0	1.0	1.0
subject10	1.0	1.0	1.0	0.77 (0.23|0)	1.0

TABLE II

NUMBER OF CLUSTERS AFTER TRAINING

	4	5	6	7
subject1	0	1.0	0	0
subject2	0	1.0	0	0
subject3	0	0.23	0.77	0
subject4	0	1.0	0	0
subject5	0.01	0.72	0.23	0.04
subject6	0	1.0	0	0
subject7	0	0.85	0.15	0
subject8	0	1.0	0	0
subject9	0	0.78	0.22	0
subject10	0.23	0.77	0	0

achieve better transfer to new subjects. In particular, we are investigating how the number of demonstrations per new subject can be reduced when reusing prior demonstrations from other subjects. Moreover, since for now the Interaction ProMPs in the cooperative tasks library are solely learned from demonstrations an important component for future work is to enrich and improve the trajectories of the robot, for example, by using reinforcement learning.

REFERENCES

[1] K. Linz and S. Stula, "Demographic change in europe-an overview," *Observatory for Sociopolitical Developements in Europe*, vol. 4, no. 1, pp. 2–10, 2010.

[2] S. Schaal, "Is imitation learning the route to humanoid robots?" *Trends in cognitive sciences*, vol. 3, no. 6, pp. 233–242, 1999.

[3] D. A. Rosenbaum, *Human motor control*. Academic press, 2009.

[4] G. Maeda, M. Ewerton, R. Lioutikov, H. B. Amor, J. Peters, and G. Neumann, "Learning interaction for collaborative tasks with probabilistic movement primitives," in *Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on*. IEEE, 2014, pp. 527–534.

[5] M. Ewerton, G. Neumann, R. Lioutikov, H. B. Amor, J. Peters, and G. Maeda, "Learning multiple collaborative tasks with a mixture of interaction primitives," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1535–1542.

[6] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and autonomous systems*, vol. 57, no. 5, pp. 469–483, 2009.

[7] A. Billard, S. Calinon, R. Dillmann, and S. Schaal, "Robot programming by demonstration," in *Springer handbook of robotics*. Springer, 2008, pp. 1371–1394.

[8] D. Vogt, S. Stepputtis, S. Grehl, B. Jung, and H. B. Amor, "A system for learning continuous human-robot interactions from human-human demonstrations," in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*. IEEE, 2017, pp. 2882–2889.

[9] G. J. Maeda, G. Neumann, M. Ewerton, R. Lioutikov, O. Kroemer, and J. Peters, "Probabilistic movement primitives for coordination of multiple human–robot collaborative tasks," *Autonomous Robots*, vol. 41, no. 3, pp. 593–612, Mar 2017. [Online]. Available: https://doi.org/10.1007/s10514-016-9556-2

[10] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, "Dynamical movement primitives: learning attractor models for motor behaviors," *Neural computation*, vol. 25, no. 2, pp. 328–373, 2013.

[11] S. Calinon, F. Guenter, and A. Billard, "On learning, representing, and generalizing a task in a humanoid robot," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 2, pp. 286–298, 2007.

[12] A. Paraschos, C. Daniel, J. Peters, and G. Neumann, "Using probabilistic movement primitives in robotics," *Autonomous Robots*, pp. 1–23, 2018.

[13] H. B. Amor, G. Neumann, S. Kamthe, O. Kroemer, and J. Peters, "Interaction primitives for human-robot cooperation tasks," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 2831–2837.

[14] C. Pérez-D'Arpino and J. A. Shah, "Fast target prediction of human reaching motion for cooperative human-robot manipulation tasks using time series classification," in *Robotics and Automation (ICRA), 2015 IEEE International Conference on*. IEEE, 2015, pp. 6175–6182.

[15] D. Lee, C. Ott, and Y. Nakamura, "Mimetic communication model with compliant physical contact in human-humanoid interaction," *The International Journal of Robotics Research*, vol. 29, no. 13, pp. 1684–1704, 2010.

[16] G. Konidaris, S. Kuindersma, R. Grupen, and A. Barto, "Robot learning from demonstration by constructing skill trees," *The International Journal of Robotics Research*, vol. 31, no. 3, pp. 360–375, 2012.

[17] S. Calinon and A. Billard, "Incremental learning of gestures by imitation in a humanoid robot," in *Proceedings of the ACM/IEEE international conference on Human-robot interaction*. ACM, 2007, pp. 255–262.

[18] A. Ahmed and E. Xing, "Dynamic non-parametric mixture models and the recurrent chinese restaurant process: with applications to evolutionary clustering," in *Proceedings of the 2008 SIAM International Conference on Data Mining*. SIAM, 2008, pp. 219–230.

[19] O. Arandjelovic and R. Cipolla, "Incremental learning of temporally-coherent gaussian mixture models," *Society of Manufacturing Engineers (SME) Technical Papers*, pp. 1–1, 2006.

[20] P. M. Engel and M. R. Heinen, "Incremental learning of multivariate gaussian mixture models," in *Brazilian Symposium on Artificial Intelligence*. Springer, 2010, pp. 82–91.

[21] R. C. Pinto and P. M. Engel, "A fast incremental gaussian mixture model," *PloS one*, vol. 10, no. 10, p. e0139931, 2015.

[22] A. Declercq and J. H. Piater, "Online learning of gaussian mixture models-a two-level approach." in *VISAPP (1)*, 2008, pp. 605–611.

[23] G. Maeda, M. Ewerton, T. Osa, B. Busch, and J. Peters, "Active incremental learning of robot movement primitives," in *Conference on Robot Learning (CORL)*, 2017.

[24] J. Hoyos, F. Prieto, G. Alenya, and C. Torras, "Incremental learning of skills in a task-parameterized gaussian mixture model," *Journal of Intelligent & Robotic Systems*, vol. 82, no. 1, pp. 81–99, 2016.

[25] I. Havoutis, A. K. Tanwani, and S. Calinon, "Online incremental learning of manipulation tasks for semi-autonomous teleoperation," 2016.

[26] D. Kulić, C. Ott, D. Lee, J. Ishikawa, and Y. Nakamura, "Incremental learning of full body motion primitives and their sequencing through human motion observation," *The International Journal of Robotics Research*, vol. 31, no. 3, pp. 330–345, 2012.

[27] A. Lemme, R. F. Reinhart, and J. J. Steil, "Self-supervised bootstrapping of a movement primitive library from complex trajectories," in *Humanoid Robots (Humanoids), 2014 14th IEEE-RAS International Conference on*. IEEE, 2014, pp. 726–732.

[28] K. Muelling, J. Kober, and J. Peters, "Learning table tennis with a mixture of motor primitives," in *Humanoid Robots (Humanoids), 2010 10th IEEE-RAS International Conference on*. IEEE, 2010, pp. 411–416.

[29] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, "Adaptive mixtures of local experts," *Neural computation*, vol. 3, no. 1, pp. 79–87, 1991.