

Robot Learning with Task-Parameterized Generative Models

Sylvain Calinon*

Abstract Task-parameterized models provide a representation of movement/behavior that can adapt to a set of task parameters describing the current situation encountered by the robot, such as location of objects or landmarks in its workspace. This paper gives an overview of the *task-parameterized Gaussian mixture model* (TP-GMM) presented in previous publications, and introduces a number of extensions and ongoing challenges required to move the approach toward unconstrained environments. In particular, it discusses its generalization capability and the handling of movements with a high number of degrees of freedom. It then shows that the method is not restricted to movements in task space, but that it can also be exploited to handle constraints in joint space, including priority constraints.

1 Introduction

Robots are provided with an increasing number of sensors and actuators. This trend introduces original challenges in machine learning, where the sample size is often bounded by the cost of data acquisition, thus requiring models capable of handling wide-ranging data. Namely, models that can start learning from a small number of demonstrations, while still being able to continue learning when more data become available.

Robot learning from demonstration is one such field, which aims at providing end-users with intuitive interfaces to transfer new skills to robots. The challenges in robot learning can often be reinterpreted as designing appropriate domain-specific priors that can supply the required generalization capability from small training sets. The position adopted in this paper is that: 1) generative models are well suited for robot learning from demonstration because they can treat recognition, classification, prediction and synthesis within the same framework; and 2) an efficient and versatile prior is to consider that the task parameters describing the current situation (body and workspace configuration encountered by the robot) can be represented as

* Idiap Research Institute, Martigny, Switzerland, e-mail: sylvain.calinon@idiap.ch
This work was in part supported by the DexROV Project through the EC Horizon 2020 programme (Grant #635491).

affine transformations (including frames of reference, coordinate systems or projections).

By providing such structure to the skill generation problem, the role of the experimenter is to provide the robot with a set of candidate frames (list of coordinate systems) that could potentially be relevant for the task. This paper will show that structuring the affine transformations in such way has a simple interpretation, that it can be easily implemented, and that it remains valid for a wide range of skills that a robot can experience.

The task-parameterized Gaussian mixture model (TP-GMM) was presented in [11, 8, 10] for the special case of frames of reference representing rotations and translations in Cartesian space. The current paper discusses the potentials of the approach and introduces several routes for further investigation, aiming at applying the proposed technique to a wider range of affine transformations (directly exploiting the considered application domain), including constraints in both configuration and operational spaces, as well as priority constraints. It also shows that the proposed method can be applied to different probabilistic encoding strategies, including subspace clustering approaches enabling the consideration of high dimension feature spaces. Examples are provided in simulations and on a real robot (transfer of manipulation skills to the Baxter bimanual robot). Accompanying source codes are available at <http://www.idiap.ch/software/pbdl/lib/>.

2 Adaptive models of movements

Task-parameterized models of movements/behaviors refer to representations that can adapt to a set of task parameters describing for example the current context, situation, state of the environment or state of the robot configuration. The *task parameters* can for example refer to the variables collected by the system to describe the position of objects in the environment. The *task parameters* can be fixed during an execution trial or they can vary while the motion is executed. The *model parameters* refer to the variables learned by the system, namely, that are stored in memory (the internal representation of the movement). During reproduction, a new set of *task parameters* (describing the present situation) is used to generate a new movement (e.g., adaptation to new position of objects).

Several denominations have been introduced in the literature to describe these models, such as *task-parameterized* [40, 11] (the denomination that will be used here), *parametric* [49, 26, 29] or *stylistic* [7]. In these models, the encoding of skills usually serve several purposes, including classification, prediction, synthesis and online adaptation. A taxonomy of task-parameterized models is presented in [8], classifying existing methods in three broad categories: 1) Approaches employing M models for the M demonstrations, performed in M different situations, see e.g. [16, 23, 29, 25, 45, 12, 21]; 2) Ap-

proaches employing P models for the P frames of reference that are possibly relevant for the task, see e.g. [32, 13]; 3) Approaches employing a single model whose parameters are modulated by task parameters, see e.g. [49, 26, 20].

In the majority of these approaches, the retrieval of movements from the model parameters and the task parameters is viewed as a standard regression problem. This generality might look appealing at first sight, but it also limits the generalization scope of these models. Our work aims at increasing the generalization capability of task-parameterized models by exploiting the functional nature of the task parameters. The approach arose from the observation that the task parameters in robotics applications can most of the time be related to some form of frames of reference, coordinate systems, basis functions or local projections, whose structure can be exploited to speed up learning and provide the robot with remarkable extrapolation capability.

2.1 Motivation

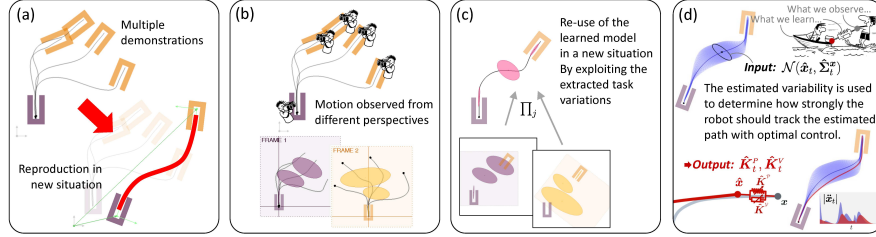


Fig. 1 Illustration of the overall approach (see main text for details). (a) Observation of a task in different situations and generalization to new contexts. Multiple demonstrations provide the opportunity to discern the structure of the task. (b) Probabilistic encoding of continuous movements in multiple coordinate systems. (c) Exploitation of variability and correlation information to adapt the motion to new situations. With cross-situational observations of the same task, the robot is able to generalize the skill to new situations. (d) Computation of the underlying optimal control strategy driving the observed behavior.

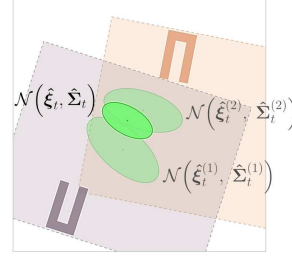
The core of the approach is to represent an observed movement or behavior as a spring-damper system with varying parameters, where a generative model is used to encode the evolution of the attractor, and the variability and correlation information is used to infer the impedance parameters of the system. These impedance parameters figuratively correspond to the stiffness of a spring and to the damping coefficient of a viscous damper, with the difference that they can also be full stiffness and damping matrices. The model shares links with optimal feedback control strategies in which devia-

tions from an average trajectory are corrected only when they interfere with task performance, resulting in a *minimal intervention principle* [43].

In its task-parameterized version, several frames of reference are interacting with each other to describe tracking behaviors in multiple coordinate systems, where statistical analysis from the perspective of each of these observers is used to estimate feedforward and feedback control terms with linear quadratic optimal control. Fig. 1 presents an illustration of the overall approach, which can be decomposed into multiple steps, involving statistical modeling, dynamical systems and optimal control.

2.2 Example with a single Gaussian

Fig. 2 Minimization of the objective function in Eq. (3) composed of a weighted sum of quadratic error terms, whose result corresponds to a product of Gaussians. It is easy to show that $\mathcal{N}(\hat{\xi}_t, \hat{\Sigma}_t)$ corresponds to the Gaussian outcome from the product of the two Gaussians $\mathcal{N}(\hat{\xi}_t^{(1)}, \hat{\Sigma}_t^{(1)})$ and $\mathcal{N}(\hat{\xi}_t^{(2)}, \hat{\Sigma}_t^{(2)})$.



Before presenting the details of the task-parameterized model, the approach is motivated by an introductory example with a single Gaussian.

Two frames will be considered, described respectively at each time step t by $\{\mathbf{b}_{t,1}, \mathbf{A}_{t,1}\}$ and $\{\mathbf{b}_{t,2}, \mathbf{A}_{t,2}\}$, representing the origin of the observer $\mathbf{b}_{t,j}$ and a set of basis vectors $\{\mathbf{e}_1, \mathbf{e}_2, \dots\}$ forming a transformation matrix $\mathbf{A}_{t,j} = [\mathbf{e}_{1,t,j}, \mathbf{e}_{2,t,j}, \dots]$.

A set of demonstrations is observed from the perspective of the two frames. During reproduction, each frame expects the new datapoints to lie within the same range as that of the demonstrations. If $\mathcal{N}(\boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma}^{(1)})$ and $\mathcal{N}(\boldsymbol{\mu}^{(2)}, \boldsymbol{\Sigma}^{(2)})$ are the normal distributions of the observed demonstrations in the first and second frames, the two frames respectively expect the reproduction attempt to lie at the intersection of the distributions $\mathcal{N}(\hat{\xi}_t^{(1)}, \hat{\Sigma}_t^{(1)})$ and $\mathcal{N}(\hat{\xi}_t^{(2)}, \hat{\Sigma}_t^{(2)})$. These distributions can be computed with the linear transformation property of normal distribution as

$$\hat{\xi}_t^{(1)} = \mathbf{A}_{t,1} \boldsymbol{\mu}^{(1)} + \mathbf{b}_{t,1}, \quad \hat{\Sigma}_t^{(1)} = \mathbf{A}_{t,1} \boldsymbol{\Sigma}^{(1)} \mathbf{A}_{t,1}^\top, \quad (1)$$

$$\hat{\xi}_t^{(2)} = \mathbf{A}_{t,2} \boldsymbol{\mu}^{(2)} + \mathbf{b}_{t,2}, \quad \hat{\Sigma}_t^{(2)} = \mathbf{A}_{t,2} \boldsymbol{\Sigma}^{(2)} \mathbf{A}_{t,2}^\top. \quad (2)$$

A trade-off thus needs to be determined during reproduction to concord with the distributions expected by each frame. The objective function can be defined as the weighted sum of quadratic error terms

$$\hat{\xi}_t = \arg \min_{\xi_t} \sum_{j=1}^2 (\xi_t - \hat{\xi}_t^{(j)})^\top \hat{\Sigma}_t^{(j)-1} (\xi_t - \hat{\xi}_t^{(j)}). \quad (3)$$

The above objective can easily be solved by differentiation, providing a point $\hat{\xi}_t$, with an error defined by covariance $\hat{\Sigma}_t$. This estimate corresponds to a product of Gaussians (intersection between the two Gaussians). Fig. 2 illustrates this process for one of the Gaussians of Fig. 1.

3 Task-parameterized Gaussian mixture model (TP-GMM)

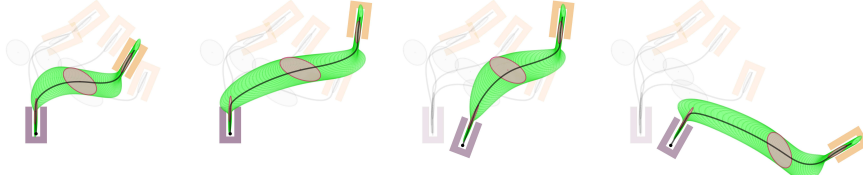


Fig. 3 Generalization capability of task-parameterized Gaussian mixture model. Each graph shows a different situation with increasing generalization complexity. In each graph, the four demonstrations and the associated adapted model parameters are depicted in semi-transparent colors.

TP-GMM is a direct extension of the objective problem presented above, by considering multiple frames and multiple clusters of datapoints (soft clustering via mixture modeling). It probabilistically encodes the relevance of candidate frames, which can change throughout the task. In contrast to approaches such as [33] that aim at extracting a single (most prominent) coordinate system located at the end of a motion segment, the proposed approach allows the superposition and transition of different coordinate systems that are relevant for the task (parallel organization of behavior primitives, adaptation to multiple viapoints in the middle of the movement, modulation based on positions, orientations or geometries of objects, etc.).

Each demonstration $m \in \{1, \dots, M\}$ contains T_m datapoints forming a dataset of N datapoints $\{\xi_t\}_{t=1}^N$ with $N = \sum_m T_m$. The task parameters are represented by P coordinate systems, defined at time step t by $\{\mathbf{b}_{t,j}, \mathbf{A}_{t,j}\}_{j=1}^P$, representing respectively the origin and the basis of the coordinate system.

The demonstrations $\boldsymbol{\xi} \in \mathbb{R}^{D \times N}$ are observed from these different viewpoints, forming P trajectory samples $\mathbf{X}^{(j)} \in \mathbb{R}^{D \times N}$. These samples can be collected from sensors located at the frames, or computed with

$$\mathbf{X}_t^{(j)} = \mathbf{A}_{t,j}^{-1}(\boldsymbol{\xi}_t - \mathbf{b}_{t,j}). \quad (4)$$

The parameters of the proposed *task-parameterized GMM* (TP-GMM) with K components are defined by $\{\pi_i, \{\boldsymbol{\mu}_i^{(j)}, \boldsymbol{\Sigma}_i^{(j)}\}_{j=1}^P\}_{i=1}^K$ (π_i are the mixing coefficients, $\boldsymbol{\mu}_i^{(j)}$ and $\boldsymbol{\Sigma}_i^{(j)}$ are the center and covariance matrix of the i -th Gaussian component in frame j).

Learning of the parameters is achieved by log-likelihood maximization subject to the constraint that the data in the different frames arose from the same source, resulting in an EM process iteratively updating the model parameters until convergence, see [10] for details. Model selection (i.e., determining the number of Gaussians in the GMM) is compatible with techniques employed in standard mixture models (*Bayesian information criterion* [37], *Dirichlet process* [34], small-variance asymptotics [27], etc.). For a movement in Cartesian space with 10 demonstrations and 3 candidate frames, the overall learning process typically takes 1-3 *sec*. The reproduction is much faster and can be computed online (typically below 1 *msec*).

The learned model is then used to reproduce movements in other situations (for new position and orientation of candidate frames). A new GMM with parameters $\{\pi_i, \hat{\boldsymbol{\xi}}_{t,i}, \hat{\boldsymbol{\Sigma}}_{t,i}\}_{i=1}^K$ can thus automatically be generated with

$$\begin{aligned} \mathcal{N}(\hat{\boldsymbol{\xi}}_{t,i}, \hat{\boldsymbol{\Sigma}}_{t,i}) &\propto \prod_{j=1}^P \mathcal{N}(\hat{\boldsymbol{\xi}}_{t,i}^{(j)}, \hat{\boldsymbol{\Sigma}}_{t,i}^{(j)}), \\ \text{with } \hat{\boldsymbol{\xi}}_{t,i}^{(j)} &= \mathbf{A}_{t,j} \boldsymbol{\mu}_i^{(j)} + \mathbf{b}_{t,j}, \quad \hat{\boldsymbol{\Sigma}}_{t,i}^{(j)} = \mathbf{A}_{t,j} \boldsymbol{\Sigma}_i^{(j)} \mathbf{A}_{t,j}^\top, \end{aligned} \quad (5)$$

where the result of the Gaussian product is given by

$$\hat{\boldsymbol{\Sigma}}_{t,i} = \left(\sum_{j=1}^P \hat{\boldsymbol{\Sigma}}_{t,i}^{(j)-1} \right)^{-1}, \quad \hat{\boldsymbol{\xi}}_{t,i} = \hat{\boldsymbol{\Sigma}}_{t,i} \sum_{j=1}^P \hat{\boldsymbol{\Sigma}}_{t,i}^{(j)-1} \hat{\boldsymbol{\xi}}_{t,i}^{(j)}. \quad (6)$$

For computational efficiency, the above equations can be computed with precision matrices instead of covariances.

Several approaches can be used to retrieve movements from the proposed model. An option is to encode both static and dynamic features in the mixture model to retrieve continuous behaviors [51, 39, 22]. An alternative option is to encode time as additional feature in the GMM, and use *Gaussian mixture regression* (GMR) [18] to retrieve continuous behaviors. Similarly, if the evolution of a decay term is encoded instead of time, the system yields a probabilistic formulation of dynamical movement primitives (DMP) [20], see

[11] for details. Fig. 3 presents TP-GMR reproduction results for the example in Fig. 1.

4 Extension to task-parameterized subspace clustering

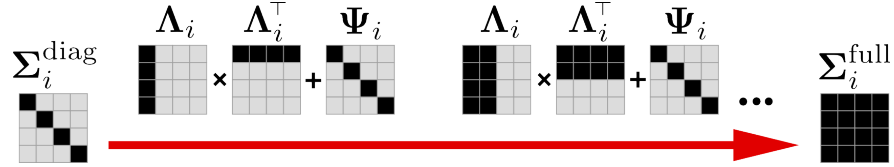


Fig. 4 The mixture of factor analyzers (MFA) covers a wide range of covariance structures for the modeling of the data, from diagonal covariances (*left*) to full covariances (*right*).

Classical model-based clustering will tend to perform poorly in high-dimensional spaces. A simple way of handling this issue is to reduce the number of parameters by considering diagonal covariances instead of full matrices, which corresponds to a separated treatment of each variable. Although common in robotics, such decoupling can be a limiting factor to encode movements and sensorimotor streams, because it follows a strategy that is not fully exploiting principles underlying coordination, motor skill acquisition and action-perception couplings.

The rationale is that diagonal structures are unadapted to motor skill representation because they do not encapsulate coordination information among the control variables. The good news is that a wide range of mixture modeling techniques exist between the encoding of diagonal and full covariances. At the exception of [14] and [47], these techniques have only been exploited to a limited extent in robot skills acquisition. They can be studied as a subspace clustering problem, aiming to group datapoints such that they can be locally projected in subspaces of reduced dimensionality. Such subspace clustering helps the analysis of the local trend of the movement, while reducing the number of parameters to be estimated, and "locking" the most important coordination patterns to efficiently cope with perturbations.

Several possible constraints can be considered, grouped in families such as parsimonious GMM [6], *mixtures of factor analyzers* (MFA) [30] or *mixtures of probabilistic principal component analyzers* [42]. Methods such as MFA provide a simple approach to the problem of high-dimensional cluster analysis with a slight modification of the generative model underlying the mixture of Gaussians to enforce low-dimensional models (i.e., noninvasive regarding the other methods used in the proposed framework). The basic idea of factor analysis (FA) is to reduce the dimensionality of the data while keeping the

observed covariance structure. MFA assumes for each component i a covariance structure of the form $\Sigma_i = \Lambda_i \Lambda_i^\top + \Psi_i$, where $\Lambda_i \in \mathbb{R}^{D \times d}$, known as the *factor loadings matrix*, typically has $d < D$ (providing a parsimonious representation of the data), and a diagonal noise matrix Ψ_i .

Fig. 4 shows that the covariance structure in MFA can span a wide range of covariances.

The TP-GMM presented in Section 3 is fully compatible with the subspace clustering approaches mentioned above. Bayesian nonparametric approaches such as [48] can be used to simultaneously select the number of clusters and the dimension of the subspace in each cluster.

The TP-MFA extension of TP-GMM opens several roads for further investigation. A possible extension is to use tied structures in the covariances to enable the organization and reuse of previously acquired synergies [17]. Another possible extension is to enable deep learning techniques in task-parameterized models. As discussed in [41], the prior of each FA can be replaced by a separate second-level MFA that learns to model the aggregated posterior of that FA (instead of the isotropic Gaussian), providing a hierarchical structure organization where one layer of latent variables can be learned at a time.

5 Extension to minimal intervention control

We showed in [10] that TP-GMM can be used to autonomously regulate the stiffness and damping behavior of the robot, see also Fig. 1-(d). It shares similarities with the solution proposed by Medina *et al.* in the context of risk-sensitive control for haptic assistance [31], by exploiting the predicted variability to form a minimal intervention controller (in task space or in joint space). The retrieved variability and correlation information is exploited to generate safe and natural movements within an optimal control strategy, in accordance to the predicted range of motion to reproduce the task, evaluated for the current situation. TP-GMM is fully compatible with *linear quadratic regulation* (LQR) and *model predictive control* (MPC) [4], providing an approach to learn controllers adapted to the current situation, with feedforward and feedback control commands varying in regard to external task parameters, see [10] for details.

Fig. 5 demonstrates that a TP-GMM with a single Gaussian, combined with an infinite-horizon LQR, can readily be used to represent various behaviors that directly exploit the torque control capability of the robot and the redundancy, both at the level of the task and at the level of the robot kinematic structure.

It is worth noting that each frame in the TP-GMM has an associated sub-objective function as in Eq. (3), which aims at minimizing the discrepancy between the demonstrations and the reproduction attempt. By considering the

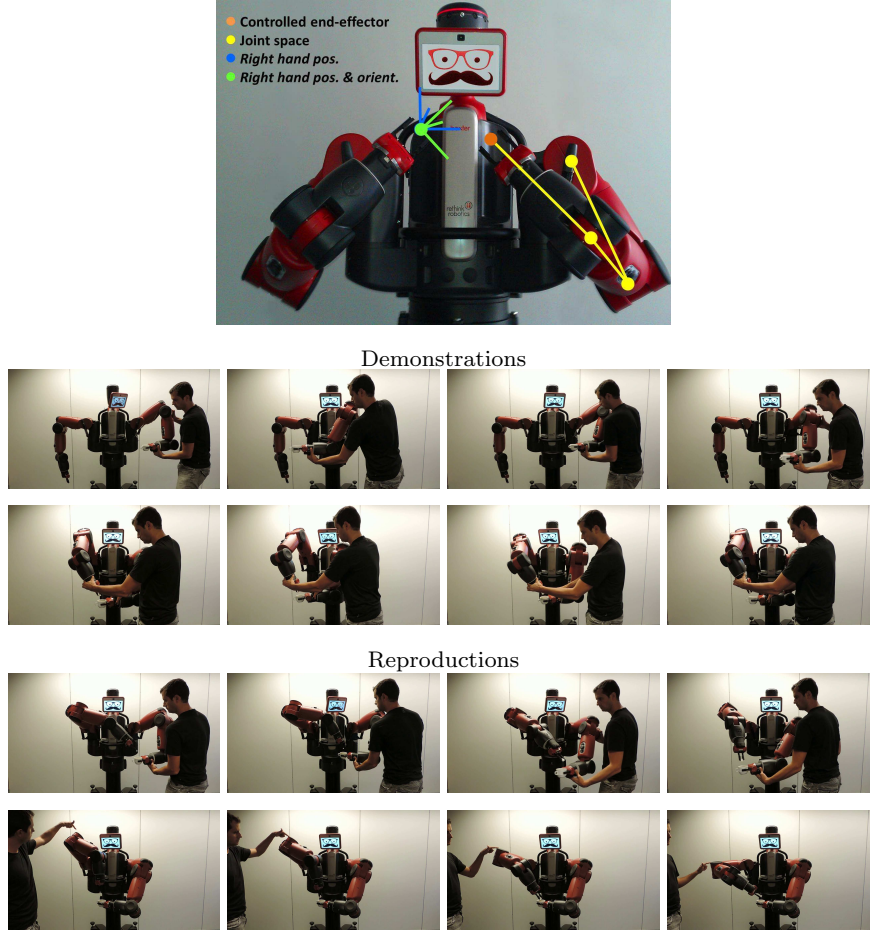


Fig. 5 Learning of two behaviors with the Baxter robot. The taught tasks consist of holding a cup horizontally with one hand, and holding a sugar cube above the cup with the other hand, where the two task primitives can be combined in parallel. The demonstrations are provided in two steps by kinesthetic teaching, namely, by holding the arms of the robot and moving them during the task while the robot compensates for the effect of gravity. This procedure allows the user to move the robot arms without feeling their weight and without feeling the motors in the articulations, while the sensors are used to record position information. Here, the data are recorded in several frames of reference (*top image*). During reproduction, the robot is controlled by following a minimal intervention principle, where the computed feed-forward and feedback control commands result in different levels of stiffness obeying the extracted variation and coordination constraints of the task. *First sequence*: Brief demonstration to show the robot how to hold a cup horizontally. *Second sequence*: Brief demonstration to show how to hold a sugar cube above the cup. *Third sequence*: Manual displacement of the left arm to test the learned behavior (the coordination of the two hands was successfully learned). *Last sequence*: Combination of the two learned task primitives. Here, the user pushes the robot to show that the robot remains soft for perturbations that do not conflict with the acquired task constraints (automatic exploitation of the redundant degrees of freedom that do not conflict with the task).

combination of these sub-objectives in the overall objective, the problem can be viewed as a rudimentary form of *inverse optimal control* (IOC) [1]. This form of IOC does not have external constraints and can be solved analytically, which means that it can provide a controller without exploratory search, at the expense of being restricted to simple forms of objectives (weighted sums of quadratic errors whose weights are learned from the demonstrations). This dual view can be exploited for further research in learning from demonstration, either to bridge action-level and goal-driven imitation, or to initialize the search in IOC.

6 Extension to multimodal data and projection constraints

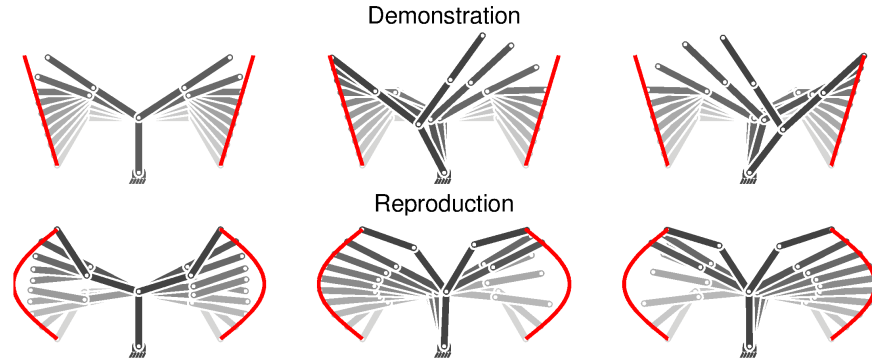


Fig. 6 Illustration of the encoding of priority constraints in a TP-GMM. The top row shows 3 demonstrations with a bimanual planar robot with 5 articulations. The color of the robot changes from light gray to black with the movement. The task consists of tracking two objects with the left and right hands (the path of the objects are depicted in red). In some parts of the demonstrations, the two objects could not be reached, and the demonstrator either made a compromise (*left* graph), or gave priority to the left or right hand (*middle* and *right* graphs). The bottom row shows reproduction attempts for new trajectories of the two objects. Although faced with different situations, the priority constraints are reproduced in a similar fashion as in the corresponding demonstrations.

TP-GMM is not limited to coordinate systems representing objects in Cartesian space. It can be extended to other forms of locally linear transformations or projections, which opens many roads for further research.

The consideration of non-square $\mathbf{A}_{t,j}$ matrices is for example relevant to learn and reproduce soft constraints in both configuration and operational spaces (through Jacobian operators). With a preliminary model of task-parameterized movements, we explored in [9] how a similar approach could

be used to simultaneously learn constraints in joint space and task space. The model also provides a principled way to learn priority constraints in a probabilistic form (through nullspace operators). The different frames correspond in this case to several subspace projections of the same movement, whose relevance is estimated statistically from the demonstrations.

A wide range of motor skills could potentially be adapted to this framework, by exploiting the functional nature of task parameters to build models that learn the local structure of the task from a small number of demonstrations. Indeed, most task parameterization in robot control can be related to some form of frames of reference, coordinate systems or basis functions, where the involvement of the frames can change during the execution of the task, with transformations represented as local linear projection operators (Jacobians for inverse kinematics, kernel matrices for nullspace projections, etc.).

The potential applications are diverse, with an objective that is well in line with the original purpose of motor primitives to be composed together serially or in parallel [15]. Further work is required to investigate in which manner TP-GMM could be exploited to provide a probabilistic view of robotics techniques that are in practice predefined, handled by *ad hoc* solutions, or sometimes inefficiently set as hard constraints. This includes the consideration of soft constraints in both configuration and operational spaces. A wide range of robot skills can be defined in such way, see e.g. the possible tasks described in §6.2.1 of [3]. In humanoids, the candidate frames could for example be employed to learn the constraints of whole-body movements from demonstration or experience, based on the regularities extracted from different subspace projections.

An important category of applications currently attracting a lot of attention concerns the problems requiring priority constraints [44, 28, 19, 50, 36]. With an appropriate definition of the frames and with an initial set of candidate task hierarchies, such constraints can be learned and encoded within a TP-GMM. Here, the probabilistic encoding is exploited to discover, from statistical analysis of the demonstrations, in which manner each subtask is prioritized.

For a controller handling constraints both in configuration and operational spaces, the most common candidate projection operators can be defined as

$$\hat{\mathbf{q}}_{t,i}^{(j)} = \mathbf{I} \quad \mu_i^{(j)} + \mathbf{0} \quad (7)$$

$$\hat{\mathbf{q}}_{t,i}^{(j)} = \mathbf{J}^\dagger(\mathbf{q}_{t-1}) \quad \mu_i^{(j)} + \mathbf{q}_{t-1} - \mathbf{J}^\dagger(\mathbf{q}_{t-1})\mathbf{x}_{t-1} \quad (8)$$

$$\hat{\mathbf{q}}_{t,i}^{(j)} = \mathbf{J}^\dagger(\mathbf{q}_{t-1})\mathbf{A}_t^\circ \quad \mu_i^{(j)} + \mathbf{q}_{t-1} + \mathbf{J}^\dagger(\mathbf{q}_{t-1})[\mathbf{b}_t^\circ - \mathbf{x}_{t-1}] \quad (9)$$

$$\hat{\mathbf{q}}_{t,i}^{(j)} = \mathbf{N}(\mathbf{q}_{t-1}) \quad \mu_i^{(j)} + \mathbf{J}^\dagger(\mathbf{q}_{t-1})\mathbf{J}(\mathbf{q}_{t-1})\mathbf{q}_{t-1} \quad (10)$$

$$\hat{\mathbf{q}}_{t,i}^{(j)} = \mathbf{N}(\mathbf{q}_{t-1})\tilde{\mathbf{J}}^\dagger(\mathbf{q}_{t-1}) \quad \mu_i^{(j)} + \mathbf{q}_{t-1} - \mathbf{N}(\mathbf{q}_{t-1})\tilde{\mathbf{J}}^\dagger(\mathbf{q}_{t-1})\mathbf{x}_{t-1} \quad (11)$$

$$\hat{\mathbf{q}}_{t,i}^{(j)} = \underbrace{\mathbf{N}(\mathbf{q}_{t-1})\tilde{\mathbf{J}}^\dagger(\mathbf{q}_{t-1})\mathbf{A}_t^\circ}_{\mathbf{A}_{t,j}} \quad \mu_i^{(j)} + \underbrace{\mathbf{q}_{t-1} + \mathbf{N}(\mathbf{q}_{t-1})\tilde{\mathbf{J}}^\dagger(\mathbf{q}_{t-1})[\mathbf{b}_t^\circ - \mathbf{x}_{t-1}]}_{\mathbf{b}_{t,j}}, \quad (12)$$

covering a wide range of robotics applications.

Note here that the product of Gaussians is computed in configuration space (\mathbf{q} and \mathbf{x} represent respectively poses in joint space and task space). Eq. (7) describes joint space constraints in a fixed frame. It corresponds to the canonical frame defined by $\mathbf{A}_{t,j} = \mathbf{I}$ (identity matrix) and $\mathbf{b}_{t,j} = \mathbf{0}$. Eq. (8) describes absolute position constraints (in operational space), where \mathbf{J}^\dagger is the Jacobian pseudoinverse used as least-norm inverse kinematics solution. Note that Eq. (8) describes a moving frame, where the task parameters change at each iteration (observation of a changing pose in configuration space). Eq. (9) describes relative position constraints, where the constraint in task space is related to an object described at each time step t by a position \mathbf{b}_t° and an orientation matrix \mathbf{A}_t° in task space. Eq. (10) describes nullspace/priority constraints in joint space, with $\mathbf{N} = \mathbf{I} - \mathbf{J}^\dagger \mathbf{J}$ a nullspace projection operator. Eq. (11) describes absolute position nullspace/priority constraints, where the secondary objective is described in task space (for a point in the kinematic chain with corresponding Jacobian $\tilde{\mathbf{J}}$). Finally, Eq. (12) describes relative position nullspace/priority constraints.

The above equations can be retrieved without much effort by discretizing (with an Euler approximation) the standard inverse kinematics and nullspace control relations that can be found in most robotics textbooks, see e.g. [3].

Fig. 6 presents a TP-GMM example with task parameters taking the form of nullspace bases. The frames are defined by Equations (9) and (12) with two different combinations of nullspaces \mathbf{N} and Jacobians $\tilde{\mathbf{J}}$ corresponding to the left and right arm.

7 Discussion and further work

A potential limitation of the current TP-GMM approach is that it requires the experimenter to provide an initial set of frames that will act as candidate projections/transformations of the data that can potentially be relevant for the task. The number of frames can be overspecified by the experimenter (e.g., by providing an exhaustive list), at the expense of potentially requiring a large number of demonstrations to obtain sufficient statistics to discard the frames that have no role in the task. The demonstrations must also be sufficiently varied, which becomes more difficult as the number of candidate frames increases. The problem *per se* is not different from the problem of selecting the variables that will form the feature vector fed to a learning system. The only difference here is that the initial selection of frames takes the form of affine transformations instead of the initial selection of elements in a feature vector.

In practice, the experimenter selects the list of objects or landmarks in the robot workspace, as well as the locations in the robot kinematic chain that might be relevant for the task, which are typically the end-effectors of

the robot, where tools, grippers or parts in contact with the environment are mounted. It should be noted here that if some frames of reference are missing during reproduction (e.g., when occlusions occur or when frames are collected at different rates), the system is still able to reproduce an appropriate behavior given the circumstance, see [2] for details.

The issue of predefining an initial set of frames of reference is not restrictive when the number of frames remains reasonably low (e.g., when they come from a set of predefined objects tracked with visual markers in a lab setting). However, for perception in more unconstrained environment, the number of frames could potentially grow (e.g., detection of phantom objects), while the number of demonstrations should remain low.

Further work is thus required to detect redundant frames or remove irrelevant frames, as well as to automatically determine in which manner the frames are coordinated with each other and locally contribute to the achievement of the task. A promising route for further investigation is to exploit the recent developments in multilinear algebra and tensor analysis [38, 24] that exploit the multivariate structure of data for statistical analysis and compression without transforming it to a matrix form.

In the proposed task-parameterized framework, the movement is expressed simultaneously in multiple coordinate systems, and is stored as a multidimensional array (tensor-variate data). This opens many roads for further investigation, where multilinear algebra could provide a principled method to simultaneously extract *eigenframes*, *eigenposes* and *eigentrajectories*. Multiway analysis of tensor-variate data could imaginably offer a rich set of data decomposition techniques, which has been demonstrated in computer imaging fields such as face processing [46], video analysis [52], geoscience [35] or neuroimaging [5], but which remains an uncharted territory in robotics and motor skills acquisition.

There are several other encoding methods that can be explored within the proposed task-parameterized approach (e.g., with *hidden Markov models* (HMM), with *Gaussian processes* (GP) or with other forms of trajectory distributions). Indeed, it is worth noting that the approach is not restricted to mixture models and can be employed with other representations as long as a local measure of uncertainty is available.

8 Conclusion

An efficient prior assumption in robot learning from demonstration is to consider that skills are modulated by external task parameters. These task parameters often take the form of affine transformations, whose role is to describe the current situation encountered by the robot (body and workspace configuration). We showed that this structure can be used with different statistical modeling strategies, including standard mixture models and sub-

space clustering. The approach can be used in a wide variety of problems in robotics, by reinterpreting them with a structural relation between the task parameters and the model parameters represented as candidate frames of reference. The rationale is that robot skills can often be related to coordinate systems, basis functions or local projections, whose structure can be exploited to speed up learning and provide robots with better generalization capability. Early promises of the approach were discussed in a series of problems in configuration and operational spaces, including tests on a Baxter robot.

References

1. Abbeel, P., Ng, A.Y.: Apprenticeship learning via inverse reinforcement learning. In: Proc. Intl Conf. on Machine Learning (ICML). Banff, Alberta, Canada (2004)
2. Alizadeh, T., Calinon, S., Caldwell, D.G.: Learning from demonstrations with partially observable task parameters. In: Proc. IEEE Intl Conf. on Robotics and Automation (ICRA), pp. 3309–3314. Hong Kong, China (2014)
3. Antonelli, G.: Underwater Robots. Springer International Publishing (2014). 3rd Edition
4. Astrom, K.J., Murray, R.M.: Feedback Systems: An Introduction for Scientists and Engineers. Princeton University Press, Princeton, NJ, USA (2008)
5. Basser, P.J., Pajevic, S.: A normal distribution for tensor-valued random variables: applications to diffusion tensor MRI. *IEEE Trans. on Medical Imaging* **22**(7), 785–794 (2003)
6. Bouveyron, C., Brunet, C.: Model-based clustering of high-dimensional data: A review. *Computational Statistics and Data Analysis* **71**, 52–78 (2014)
7. Brand, M., Hertzmann, A.: Style machines. In: Proc. ACM Intl Conf. on Computer graphics and Interactive Techniques (SIGGRAPH), pp. 183–192. New Orleans, Louisiana, USA (2000)
8. Calinon, S., Alizadeh, T., Caldwell, D.G.: On improving the extrapolation capability of task-parameterized movement models. In: Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS), pp. 610–616. Tokyo, Japan (2013)
9. Calinon, S., Billard, A.G.: Statistical learning by imitation of competing constraints in joint space and task space. *Advanced Robotics* **23**(15), 2059–2076 (2009)
10. Calinon, S., Bruno, D., Caldwell, D.G.: A task-parameterized probabilistic model with minimal intervention control. In: Proc. IEEE Intl Conf. on Robotics and Automation (ICRA), pp. 3339–3344. Hong Kong, China (2014)
11. Calinon, S., Li, Z., Alizadeh, T., Tsagarakis, N.G., Caldwell, D.G.: Statistical dynamical systems for skills acquisition in humanoids. In: Proc. IEEE Intl Conf. on Humanoid Robots (Humanoids), pp. 323–329. Osaka, Japan (2012)
12. Campbell, C.L., Peters, R.A., Bodenheimer, R.E., Bluethmann, W.J., Huber, E., Ambrose, R.O.: Superpositioning of behaviors learned through teleoperation. *IEEE Trans. on Robotics* **22**(1), 79–91 (2006)
13. Dong, S., Williams, B.: Learning and recognition of hybrid manipulation motions in variable environments using probabilistic flow tubes. *Intl Journal of Social Robotics* **4**(4), 357–368 (2012)
14. Field, M., Stirling, D., Pan, Z., Naghdy, F.: Learning trajectories for robot programming by demonstration using a coordinated mixture of factor analyzers. *IEEE Trans. on Cybernetics* (2015). In Press.

15. Flash, T., Hochner, B.: Motor primitives in vertebrates and invertebrates. *Current opinion in neurobiology* **15**(6), 660–666 (2005)
16. Forte, D., Gams, A., Morimoto, J., Ude, A.: On-line motion synthesis and adaptation using a trajectory database. *Robotics and Autonomous Systems* **60**(10), 1327–1339 (2012)
17. Gales, M.J.F.: Semi-tied covariance matrices for hidden markov models. *IEEE Trans. on Speech and Audio Processing* **7**(3), 272–281 (1999)
18. Ghahramani, Z., Jordan, M.I.: Supervised learning from incomplete data via an EM approach. In: J.D. Cowan, G. Tesauro, J. Alspector (eds.) *Advances in Neural Information Processing Systems (NIPS)*, vol. 6, pp. 120–127. Morgan Kaufmann Publishers, Inc., San Francisco, CA, USA (1994)
19. Hak, S., Mansard, N., Stasse, O., Laumond, J.P.: Reverse control for humanoid robot task recognition. *IEEE Trans. on Systems, Man, and Cybernetics, Part B: Cybernetics* **42**(6), 1524–1537 (2012)
20. Ijspeert, A., Nakanishi, J., Pastor, P., Hoffmann, H., Schaal, S.: Dynamical movement primitives: Learning attractor models for motor behaviors. *Neural Computation* **25**(2), 328–373 (2013)
21. Inamura, T., Toshima, I., Tanie, H., Nakamura, Y.: Embodied symbol emergence based on mimesis theory. *Intl Journal of Robotic Research* **23**(4-5), 363–377 (2004)
22. Khansari-Zadeh, S.M., Billard, A.: Learning stable non-linear dynamical systems with Gaussian mixture models. *IEEE Trans. on Robotics* **27**(5), 943–957 (2011)
23. Kober, J., Wilhelm, A., Oztog, E., Peters, J.: Reinforcement learning to adjust parametrized motor primitives to new situations. *Autonomous Robots* **33**(4), 361–379 (2012)
24. Kolda, T., Bader, B.: Tensor decompositions and applications. *SIAM Review* **51**(3), 455–500 (2009)
25. Kronander, K., Khansari-Zadeh, M.S.M., Billard, A.: Learning to control planar hitting motions in a minigolf-like task. In: *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, pp. 710–717 (2011)
26. Krueger, V., Herzog, D.L., Baby, S., Ude, A., Kragic, D.: Learning actions from observations: Primitive-based modeling and grammar. *IEEE Robotics and Automation Magazine* **17**(2), 30–43 (2010)
27. Kulis, B., Jordan, M.I.: Revisiting k-means: New algorithms via Bayesian non-parametrics. In: *Proc. Intl Conf. on Machine Learning (ICML)*, pp. 1–8. Edinburgh, Scotland, UK (2012)
28. Lober, R., Padois, V., Sigaud, O.: Multiple task optimization using dynamical movement primitives for whole-body reactive control. In: *Proc. IEEE Intl Conf. on Humanoid Robots (Humanoids)*. Madrid, Spain (2014)
29. Matsubara, T., Hyon, S.H., Morimoto, J.: Learning parametric dynamic movement primitives from multiple demonstrations. *Neural Networks* **24**(5), 493–500 (2011)
30. McLachlan, G.J., Peel, D., Bean, R.W.: Modelling high-dimensional data by mixtures of factor analyzers. *Computational Statistics and Data Analysis* **41**(3-4), 379–388 (2003)
31. Medina, J.R., Lee, D., Hirche, S.: Risk-sensitive optimal feedback control for haptic assistance. In: *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, pp. 1025–1031 (2012)
32. Mühlig, M., Gienger, M., Steil, J.: Interactive imitation learning of object movement skills. *Autonomous Robots* **32**(2), 97–114 (2012)
33. Niekum, S., Osentoski, S., Konidaris, G., Chitta, S., Marthi, B., Barto, A.G.: Learning grounded finite-state representations from unstructured demonstrations. *The International Journal of Robotics Research* **34**(2), 131–157 (2015)

34. Rasmussen, C.E.: The infinite Gaussian mixture model. In: S.A. Solla, T.K. Leen, K.R. Mueller (eds.) *Advances in Neural Information Processing Systems (NIPS)*, pp. 554–560. MIT Press (2000)
35. Renard, N., Bourennane, S., Blanc-Talon, J.: Denoising and dimensionality reduction using multilinear tools for hyperspectral images. *IEEE Geoscience and Remote Sensing Letters* **5**(2), 138–142 (2008)
36. Saveriano, M., An, S., Lee, D.: Incremental kinesthetic teaching of end-effector and null-space motion primitives. In: *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, pp. 3570–3575 (2015)
37. Schwarz, G.: Estimating the dimension of a model. *Annals of Statistics* **6**(2), 461–464 (1978)
38. Signoretto, M., Van de Plas, R., De Moor, B., Suykens, J.A.K.: Tensor versus matrix completion: A comparison with application to spectral data. *IEEE Signal Processing Letters* **18**(7), 403–406 (2011)
39. Sugiyama, K., Iwahashi, N., Kashioka, H., Nakamura, S.: Learning, generation, and recognition of motions by reference-point-dependent probabilistic models. *Advanced Robotics* **25**(6-7), 825–848 (2011)
40. Tang, J., Singh, A., Goehausen, N., Abbeel, P.: Parameterized maneuver learning for autonomous helicopter flight. In: *Proc. IEEE Intl Conf. on Robotics and Automation (ICRA)*, pp. 1142–1148 (2010)
41. Tang, Y., Salakhutdinov, R., Hinton, G.: Deep mixtures of factor analysers. In: *Proc. Intl Conf. on Machine Learning (ICML)*. Edinburgh, Scotland (2012)
42. Tipping, M.E., Bishop, C.M.: Mixtures of probabilistic principal component analysers. *Neural Computation* **11**(2), 443–482 (1999)
43. Todorov, E., Jordan, M.I.: Optimal feedback control as a theory of motor coordination. *Nature Neuroscience* **5**, 1226–1235 (2002)
44. Towell, C., Howard, M., Vijayakumar, S.: Learning nullspace policies. In: *Proc. IEEE/RSJ Intl Conf. on Intelligent Robots and Systems (IROS)*, pp. 241–248 (2010)
45. Ude, A., Gams, A., Asfour, T., Morimoto, J.: Task-specific generalization of discrete and periodic dynamic movement primitives. *IEEE Trans. on Robotics* **26**(5), 800–815 (2010)
46. Vasilescu, M.A.O., Terzopoulos, D.: Multilinear analysis of image ensembles: TensorFaces. In: *Computer Vision (ECCV)*, *Lecture Notes in Computer Science*, vol. 2350, pp. 447–460. Springer Berlin Heidelberg (2002)
47. Vijayakumar, S., D’souza, A., Schaal, S.: Incremental online learning in high dimensions. *Neural Computation* **17**(12), 2602–2634 (2005)
48. Wang, Y., Zhu, J.: DP-space: Bayesian nonparametric subspace clustering with small-variance asymptotics. In: *Proc. Intl Conf. on Machine Learning (ICML)*, pp. 1–9. Lille, France (2015)
49. Wilson, A.D., Bobick, A.F.: Parametric hidden Markov models for gesture recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence* **21**(9), 884–900 (1999)
50. Wrede, S., Emmerich, C., Ricarda, R., Nordmann, A., Swadzba, A., Steil, J.J.: A user study on kinesthetic teaching of redundant robots in task and configuration space. *Journal of Human-Robot Interaction* **2**, 56–81 (2013)
51. Zen, H., Tokuda, K., Kitamura, T.: Reformulating the HMM as a trajectory model by imposing explicit relationships between static and dynamic feature vector sequences. *Computer Speech and Language* **21**(1), 153–173 (2007)
52. Zhao, Q., Zhou, G., Adali, T., Zhang, L., Cichocki, A.: Kernelization of tensor-based models for multiway data analysis: Processing of multidimensional structured data. *IEEE Signal Processing Magazine* **30**(4), 137–148 (2013)