To appear in *Advanced Robotics*
Vol. 00, No. 00, April 2019, 1–19

FULL PAPER

Data-efficient Learning of Robotic Clothing Assistance using Bayesian Gaussian Process Latent Variable Model

Nishanth Koganti[a*], Tomohiro Shibata[b], Tomoya Tamei[c], and Kazushi Ikeda[a]

[a] *Graduate School of Science and Technology, Nara Institute of Science and Technology, Japan*;
[b] *Graduate School of Life Sciences and Systems Engineering, Kyushu Institute of Technology, Japan*;
[c] *Mathematical and Data Science Center, Kobe University, Japan*

(*v3.0 released April 2019*)

Motor-skill learning for complex robotic tasks is a challenging problem due to the high task variability. Robotic clothing assistance is one such challenging problem that can greatly improve the quality-of-life for the elderly and disabled. In this study, we propose a data-efficient representation to encode task-specific motor-skills of the robot using Bayesian nonparametric latent variable models. The effectivity of the proposed motor-skill representation is demonstrated in two ways: 1) through a real-time controller that can be used as a tool for learning from demonstration to impart novel skills to the robot and 2) by demonstrating that policy search reinforcement learning in such a task-specific latent space outperforms learning in the high-dimensional joint configuration space of the robot. We implement our proposed framework in a practical setting with a dual-arm robot performing clothing assistance tasks.

Keywords: Clothing Assistance; Gaussian Processes; Latent Variable Models; Policy Search Reinforcement Learning; Learning from Demonstration

Index

1. Introduction

Recently, there has been a tremendous increase in the elderly population of the world leading to a shortage of caregivers and therapists. To address this issue, we need a greater presence of ICT and assistive robotics. One of the major problems that arise due to aging is the loss of motor functions to perform dexterous tasks such as putting on clothes. Patients with bone and muscle related diseases find it difficult to move their arms beyond a certain range and are not able to wear clothes by themselves. Similarly, patients with conditions like Alzheimers disease face difficulty in performing fine movements such as putting on buttons and are dependent on

*Corresponding author. Email: nishanth-k@is.naist.jp

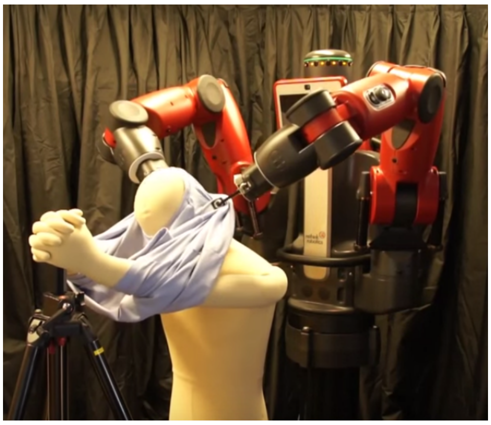


Figure 1. Clothing assistance setting: Dual-arm robot clothing soft-mannequin with T-shirt.

the caregiver. Although in most cases, the patient is not entirely dependent on the caregiver and can perform the tasks requiring some assistance. Service robots could ideally perform this task.

On the other hand, an area of active research in the domain of robotics is motor-skills learning that enables robots to perform complex tasks [1, 2]. However, existing methods require a large number of interactions to learn the optimal behavior which is not suitable in practical scenarios such as in the field of assistive robotics. Clothing assistance is one such task which is a necessity in the daily life of the elderly and disabled people. Design of a practical framework involves cloth state estimation in real-time and a learning framework that can detect and adapt to various failure scenarios. A promising approach is to formulate robotic clothing assistance as a reinforcement learning (RL) problem wherein the robot learns to recover from failure scenarios and adapt to new settings from experience.

An RL framework for clothing assistance was proposed by Tamei *et al.* [3] where a dual-arm robot was the agent, and a mannequin was used as the subject. Policy search is performed in a kinematic space which is considerably high dimensional for a 7 degree of freedom (DOF) dual-arm robot. The policy was represented using via-points extracted using a minimum jerk criterion. To ensure tractable learning time, policy update was done using the finite difference policy gradient algorithm applied to a single via-point of a single joint in each robot arm. This severely constrained the generalization capability to very different environmental settings such as significant changes in the subject's posture or using a different clothing article.

In this study, we propose an efficient representation of motor-skills that relies on the use of Bayesian Gaussian Process Latent Variable Model (BGPLVM) [4]. BGPLVM is capable of learning a data-efficient latent space for clothing tasks performed by a dual-arm robot. We present two applications with the BGPLVM latent space as shown in Figure 2. Firstly, a user-friendly interface to impart novel motor-skills to a bulky dual-arm robot. We present a real-time controller with input from the latent space that can be used as a user-friendly tool for LfD. Second, we propose a novel RL framework where the BGPLVM latent space is used as a search space in combination with an unbiased policy gradient algorithm, Policy learning by Weighted Exploration with Returns (PoWER) [5]. We demonstrate that the learned latent space generates robot trajectories that maintain task space constraints required for clothing tasks. We apply our proposed method in a practical setting of robotic clothing assistance as shown in Figure 1. The experimental results indicate a promising representation with reinforcement learning that can be used for robotic tasks with complex motor-skills.

The rest of the paper is structured as follows. Section 2 provides an overview of related studies. The proposed framework is presented in Section 3. Section 4 includes experimental results and Section 6 concludes the paper with directions for future work.

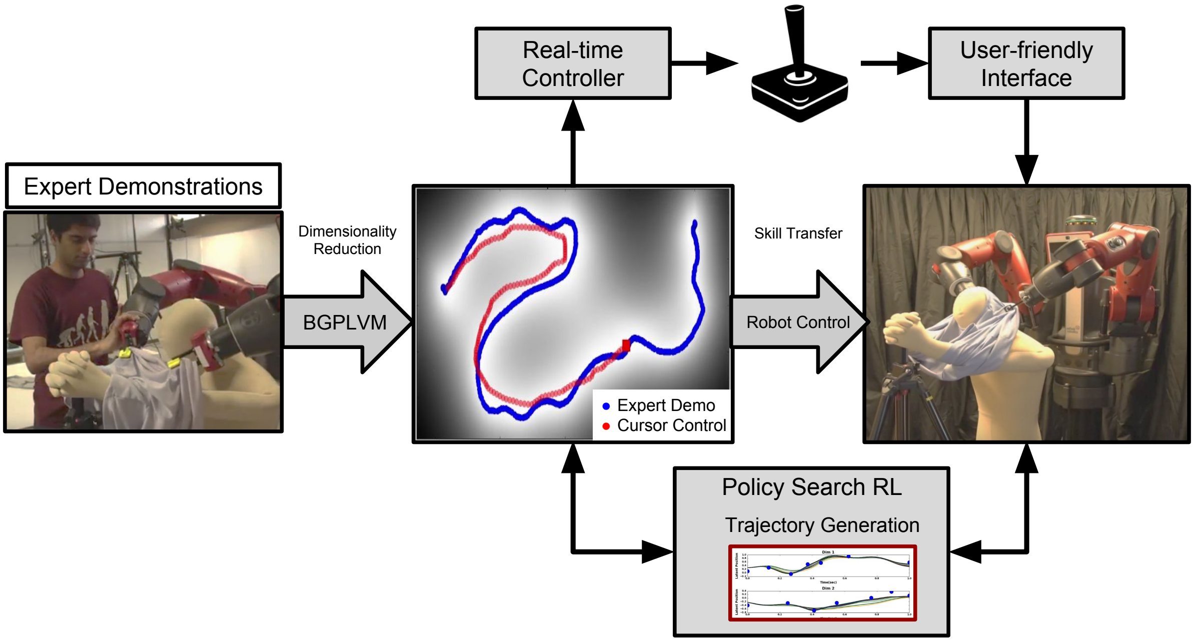


Figure 2. Overview of proposed framework presented as a flowchart. Figure on left shows human demonstration of clothing assistance used to learn a low-dimensional latent space (figure in middle) which is used for controlling the robot (figure on right) in two different ways.

2. Related Work

In this section, we present the related work in three subcategories: application of Latent Variable Models (LVM) to robotics, motor-skills learning, and robotic clothing assistance.

2.1 *Latent Variable Models for Robotics*

Robotic tasks have high-dimensional observations obtained using noisy sensors. One approach is to use dimensionality reduction and obtain a low-dimensional manifold that efficiently captures the task representation as well as variability in different settings. Principal Component Analysis (PCA) [6] is commonly used with applications including robot localization [7], LfD in humanoid robots [8], robotic hand grasp planing [9] and EMG-based robot arm control [10]. With the increasing popularity of deep learning, there have been several studies that use neural network based autoencoders in various robotics applications. Watter *et al.* [11] used variational autoencoders and a locally linear approximation of nonlinear dynamical systems to generate high-dimensional image trajectories given current observations. Van Hoof *et al.* [12] proposed the use of autoencoders to learn a low dimensional feature space of high dimensional tactile and visual information. Linear models such as PCA are not suitable to handle the nonlinear dynamics of clothing articles. On the other hand, methods such as variational autoencoders are capable of learning complex non-linear mappings but are not sample-efficient.

There have been several studies that use Gaussian process latent variable model (GPLVM) and their extensions to perform dimensionality reduction in various settings. Shon *et al.* [13] proposed an LVM formulation to learn a shared latent space between human joints and humanoid degrees of freedom which enabled robotic imitation of the human poses. Ko *et al.* [14] extended Bayes filters using Gaussian process (GP) based observation, prediction models and applied it to various robotic applications. Wang *et al.* [15] proposed a dynamical extension to GPLVM and applied it to predict human intention during a human-robot interaction task. Nakamura *et al.* [16] performed segmentation of human motion capture data using a hidden semi Markov model where the emission distributions were modeled using a GP. Koskinopoulou *et al.* [17]

performed LfD using GPLVM where the mapping is learned using human demonstrations. Some of these studies use GPLVM which relies on a Maximum-A-Posteriori (MAP) estimate of the latent space, and so the models tend to overfit to the training data and do not generalize well to drastic changes in the task settings.

2.2 *Motor-skill Learning using RL and LVMs*

In recent years, RL has been successfully applied to various robotic tasks [1]. Several studies have proposed motor skills learning specifically for cloth manipulation to handle the inherent non-rigidity. Doumanoglou *et al.* [18] formulated a Partially Observable Markov Decision Process (POMDP) framework for cloth unfolding along with the use of random forests for cloth classification. Huang *et al.* [19] used depth and appearance features for detecting graspable regions and generated trajectories through a warp function to bring clothes to the desired configuration. Yang *et al.* [20] proposed a deep learning based framework for autonomous cloth folding. They relied on a convolutional autoencoder for task-specific feature extraction from raw images and a time-delayed neural network to learn the dynamics of cloth folding.

Balaguer *et al.* [21] proposed a reinforcement learning framework that exploits the dynamics of clothing articles to perform a high momentum folding task. They used a motion capture system for real-time cloth state estimation as the clothing article did not undergo much occlusion. Another closely related application is the manipulation of deformable objects such as ropes and soft objects. Monso *et al.* [22] proposed a probabilistic motion planning framework for cloth separation by formulating the problem as a POMDP to handle uncertainty during manipulation. They defined a low-dimensional state representation to ensure fast and efficient learning of the task. Lee *et al.* [23] proposed a force-based manipulation framework for applications such as knot tying. They rely on the use of non-rigid registration to modify expert demonstrations to the current environmental setting. Hu *et al.* [24] proposed a visual-servo based control framework using online GP regression where uninformative observations are removed for real-time control.

Few studies have handled sample-efficiency explicitly which is crucial for our assistive robotics task. Deisenroth *et al.* [25] proposed a model-based RL framework that relies on a GP transition model along with explicit incorporation of uncertainty in long-term predictions. This framework suffers from computational intractability for high-dimensional applications. RL usually suffer from the curse of dimensionality and a possible solution is the use of dimensionality reduction (DR). Some studies use DR as a preprocessing step and perform RL in the reduced search space [26]. Others inherently combine DR and RL wherein the dimensionality reduction is motivated by the rewards obtained during the learning phase [27, 28]. However, these studies either use linear models for DR limiting the modeling capability or rely on a MAP estimate of the latent space which tends to overfit to the training data.

2.3 *Robotic Clothing Assistance*

Recently, there have been several promising studies that tackle this challenging problem. For people with cognitive impairments, dressing assistance mainly involves providing social cues rather than physical support. Burleson *et al.* [29] proposed a framework to detect abnormal dressing states by tracking clothing articles that have fiducial markers. Orr *et al.* [30] developed a multi-agent system that relied on sensor fusion to provide recommendations for clothing when the users were about to leave their house. Klee *et al.* [31] proposed a clothing assistance framework to communicate and coordinate with a human to complete clothing tasks.

Conducting clothing assistance experiments and future real-world implementations need to be safe and efficient. One approach to sidestep this problem is to develop a realistic simulator and use this as a testbed. Clegg *et al.* [32] developed a framework to synthesize dressing motion performed by animated human characters. Erickson *et al.* [33] relied on simulations of dressing a human hand in a shirt to infer the forces felt by the human from the end effector forces of

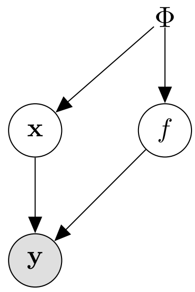


Figure 3. Graphical model of BGPLVM [4]. Φ are the model hyper parameters that need to be optimized.

the robot. They generated large amounts of data and used it to train deep neural networks for force estimation. Yu *et al.* [34] used haptic information obtained from a simulation of dressing assistance to train a classifier that could predict failure scenarios in a real-world implementation of the same task.

Yamazaki *et al.* [35] proposed an end-to-end framework for bottom-dressing that relies on proprioceptive and visual information to detect failure scenarios. Colomé *et al.* [36] relied on a friction-based complaint controller to perform a scarf wrapping task. Kapusta *et al.* [37] used hidden Markov models to detect failure scenarios for the task where a linear actuator clothes human subjects with a hospital gown. Gao *et al.* [38] proposed a path optimization approach that relies on robot proprioceptive information for assisting human users with dressing. These studies have handled various aspects of the problem. However, they do not handle the scenarios where there is a tight coupling between the human and clothing article. Furthermore, the emphasis was not motor-skills learning, and they usually perform point-to-point motion planning.

In our previous work, we demonstrated the applicability of BGPLVM for real-time cloth state estimation [39, 40]. We performed multi-view learning using Manifold Relevance Determination (MRD) [41] and perform accurate human-cloth relationship estimation from very high-dimensional and noisy point cloud data. In this study, we propose the use of BGPLVM to learn a low-dimensional latent space for encoding motor-skills. The focus of this study is to design an efficient action representation for RL of clothing assistance. The advantage of BGPLVM is that it relies on variational inference to learn a posterior distribution on the latent space rather than a MAP estimate as in GPLVM [42]. This avoids overfitting to the training data, thereby, improving the generalization capability of the model to unseen environmental settings. We further explore various feature representations and their effect on the resultant latent space, specific to clothing assistance tasks. We implement our framework on a practical setting of clothing assistance as formulated by Tamei *et al.* [3] that involves tight coupling between human and the clothing article along with high variability in policy depending on the task settings.

3. Methods

In this section, we present our formulation to encode motor-skills and a generic learning framework that relies on this formulation. Section 3.1 provides the mathematical formulation of BGPLVM. Section 3.2 provides the application of BGPLVM to encode motor-skills for clothing assistance task. Finally, Section 3.3 describes the reinforcement learning framework which relies on BGPLVM latent space to perform policy search.

3.1 *Bayesian Nonparametric Dimensionality Reduction*

BGPLVM is a dimensionality reduction technique proposed by Titsias *et al.* [4] derived from the generative model shown in Figure 3. The model assumes that the high-dimensional observed data, $\mathbf{Y} = \{\mathbf{y}_i \in \mathrm{R}^D\}_{n=1}^N$, is generated from low-dimensional latent inputs $\mathbf{X} = \{\mathbf{x}_i \in \mathrm{R}^q\}_{n=1}^N$

through a noisy process,

$$\mathbf{y}_i = f(\mathbf{x}_i) + \epsilon, \ \epsilon \sim \mathcal{N}(\mathbf{0}, \beta^{-1}\mathbf{I}) \tag{1}$$

where β is the inverse variance for noise random variable ϵ. The mapping function f is modeled using a GP which has several useful properties. GP performs data-efficient learning as it can learn informative mappings with few samples and capable of learning complex mappings using non-linear kernel functions. The conditional likelihood for the generative model is given by:

$$p(\mathbf{Y}|\mathbf{X}, \Phi) = \prod_{d=1}^{D} \mathcal{N}(\mathbf{Y}_{:,d}|\mathbf{0}, \mathbf{K} + \beta^{-1}I) \tag{2}$$

where $\mathbf{Y}$ is the observed data and $\mathbf{X}, \Phi$ are the unknown latent points, hyperparameters for the GP mapping that need to be inferred. The conditional likelihood is factorized with respect to the output dimensionality D with independent GP mapping for each output dimension d. $\mathbf{K}$ is the kernel matrix constructed from the latent points. In the generative model, the latent positions need to be marginalized out for having a purely Bayesian treatment,

$$\begin{aligned} p(\mathbf{X}) &= \prod_{n=1}^{N} \mathcal{N}(\mathbf{x}_n|\mathbf{0}, \mathbf{I}), \\ p(\mathbf{Y}|\Phi) &= \int p(\mathbf{Y}|\mathbf{X}, \Phi)p(\mathbf{X})d\mathbf{X} \end{aligned} \tag{3}$$

Variational inference is used to compute a tractable lower bound for the marginalization thereby inferring a posterior distribution on the latent space which avoids the problem of overfitting.

For performing the automatic model selection of the latent space dimensionality, the Automatic Relevance Determination (ARD) Kernel [43] can be used in the GP mapping:

$$k_{\text{ard}}(\mathbf{x}_i, \mathbf{x}_j) = \sigma_{ard}^2 \exp\left(-\frac{1}{2}\sum_{k=1}^{q} \alpha_q (x_{i,k} - x_{j,k})^2\right) \tag{4}$$

The ARD weights α_q describe the relevance of each dimension and zero weight indicates complete irrelevance. Maximizing the marginal likelihood w.r.t. these weights allows the inference of latent space dimensionality. The inference for unseen test data can now be performed through a Bayesian formulation instead of relying on a MAP estimation of the latent space. The predictive distribution is given by the ratio of two marginal likelihoods, both of which can be approximated using the variational inference technique:

$$p(\mathbf{y}^*|\mathbf{Y}) = \frac{\int p(\mathbf{y}^*, \mathbf{Y}|\mathbf{x}^*, \mathbf{X})p(\mathbf{x}^*, \mathbf{X})d\mathbf{X}d\mathbf{x}^*}{\int p(\mathbf{Y}|\mathbf{X})p(\mathbf{X})d\mathbf{X}} \tag{5}$$

3.2 *Motor-skills Representation using BGPLVM*

Motor-skills for robotic clothing assistance is given by joint angle trajectories of a dual-arm robot that lie in a high-dimensional space. The robot also has to maintain several task space constraints such as coupling with a clothing article along with safe human-robot interaction as shown in Figure 2. To address these problems, we propose the use of BGPLVM for learning a low-dimensional latent space through a nonlinear mapping to the kinematic space. There are several motivating factors for this choice. The Bayesian treatment avoids overfitting to the

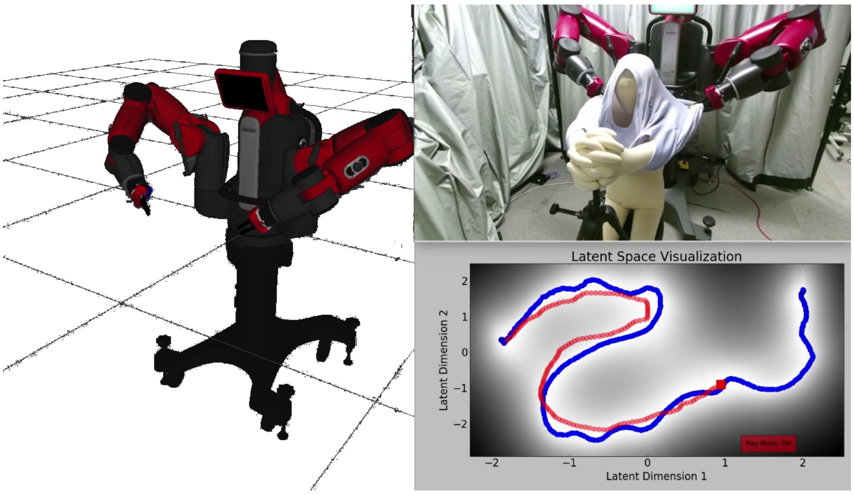


Figure 4. Interface for imparting motor-skills. Left: Simulator to plan trajectory, Right Top: Robot performing trajectory. Right Bottom: BGPLVM latent space to control robot.

training data, and the use of an ARD kernel in the GP mapping leads to the inference of the inherent dimensionality to encode motor-skills.

In this section, we present the formulation used to apply BGPLVM to clothing assistance skills. We consider the clothing task where a dual-arm robot dresses a soft mannequin in a T-shirt which is initially resting on the mannequin's arms. The training dataset is given by human demonstration through kinesthetic movement while controlling the robot under gravity compensation mode. The BGPLVM model learns a mapping from the low dimensional latent space to the robot kinematic space such that a trajectory of points in the latent space generates a trajectory on the dual-arm robot. The GP mapping leads to data-efficient learning, thereby requiring few demonstrations, possibly one, for task generalization.

The latent features learned by BGPLVM depend on the input dataset provided. We consider two alternate representations i.e. 1) kinematic representation (JA) given by joint angles of 7 DoF dual-arm robot $D_K = 14$ and 2) task space representation (EE) given by the end-effector pose of both arms with Cartesian position $P_X, P_Y, P_Z \in \mathrm{R}^3$ and orientation with quaternion representation $O_X, O_Y, O_Z, O_\omega \in \mathrm{R}^4$ forming a 14-dimensional space $D_T = 14$. We set the dimension of the latent space as $q = 10$ for all our experiments. However, the dimensionality is eventually inferred through the training of the ARD kernel weights as explained in Section 3.1.

There can be several types of failure scenarios when the robot performs clothing tasks. To recover from these failures, not only is the trajectory of the robot important, but also the speed of execution. Typically such trajectories are provided to the robot through kinesthetic demonstrations where a person holds the robot arms and physical moves them. Imparting these skills through kinesthetic movement can be difficult for inexperienced users especially when the robot is bulky and with specific kinematic constraints. Inexperienced users could impart noisy demonstrations which could lead to suboptimal performance and in some cases unsafe movements as well.

To address this problem, we implemented a user interface that uses BGPLVM. Our interface involves interaction with a BGPLVM latent space trained using expert demonstrations displayed in a 2D space. A real-time controller is implemented as shown in Figure 12 that maps the cursor coordinates from the BGPLVM latent space to a pose of dual arm 7-DoF robot. This interface can be used as a tool for LfD where the necessary clothing skills are imparted to the robot by using cursor control over the latent space. This can be a user-friendly interface where the user need not physically interact with the robot and can control the robot through intuitive interfaces such as a mouse or even a touchpad.

Real-time implementation of the controller was designed using the Robot Operating System

(ROS) software framework. A pipeline was formed where cursor coordinates in the latent space were mapped through BGPLVM to generate a robot joint angle pose which was provided as input to the low-level controllers of the robot. Using this interface, a path traced in the latent space converts to a trajectory performed on the dual-arm robot in real-time. The implementation of the latent space controller along with the user-friendly interface is provided as an open source repository at `https://github.com/ShibataLab/cloth_assist_framework` for further reference. An example scenario for this interface is care-givers imparting motor-skills to assistive robots in a real-world health care facility.

3.3 *Latent Space Reinforcement Learning*

In this section, we formulate a policy search framework in the BGPLVM latent space. The objective is to learn a high-dimensional robot trajectory for performing the clothing task on an unseen posture of the mannequin by searching within this latent space. Firstly, a dataset of successful clothing assistance trajectories is used to train a latent space that encodes the motor-skills. Each of the trajectories is now transformed into a sequence of points in the latent space forming latent space trajectories. The policy search is performed using PoWER [5] which is a commonly used policy search algorithm. PoWER can be considered reward-weighted recombination of previous experience. It has an Expectation-Maximization like formulation where the policy is given by a weighted summation of basis functions with state-dependent exploration noise $\pi(\mathbf{a}_t|\mathbf{s}_t, t) = \theta^T \mu(\mathbf{s}, t) + \epsilon(\mu(\mathbf{s}, t))$ and the policy parameters are updated as follows:

$$\theta' = \theta + \mathrm{E}\left\{\sum_{t=1}^{T} \mathbf{W}(\mathbf{s}_t, t) Q^{\pi}(\mathbf{s}_t, \mathbf{a}_t, t)\right\}^{-1} \mathrm{E}\left\{\sum_{t=1}^{T} \mathbf{W}(\mathbf{s}_t, t)\ \epsilon_t\ Q^{\pi}(\mathbf{s}_t, \mathbf{a}_t, t)\right\} \tag{6}$$

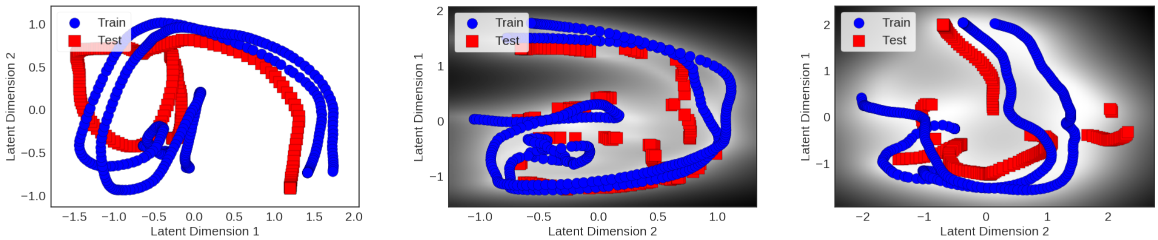
where θ are the policy parameters, $W(\mathbf{s}_t, t)$ is the basis function matrix and ϵ_t is a time-dependent exploration noise term. $Q^{\pi}(\mathbf{s}_t, \mathbf{a}_t, t)$ is the value function which is approximated using an unbiased estimator, $\hat{Q}^{\pi}(\mathbf{s}, \mathbf{a}, t) = \sum_{\tilde{t}=t}^{T} r(\mathbf{s}_{\tilde{t}}, \mathbf{a}_{\tilde{t}}, \tilde{t})$.

We consider Dynamic Movement Primitives (DMP) [44] as policy representation. It is the combination of a point attractor dynamical system, and a non-linear forcing term $(f(s))$ learned using Locally Weighted Regression (LWR) [45]:

$$\begin{aligned} \tau \ddot{x} &= K(g - x) - D\dot{x} + (g - x_0) f, \\ f(s) &= \frac{\sum_i w_i \psi_i(s) s}{\sum_i \psi_i(s)}, \text{ where } \tau \dot{s} = -\alpha s \end{aligned} \tag{7}$$

where K, D are the coefficients for the point attractor system, x_0, g are the initial and goal positions. The non-linear forcing term is described using another canonical system (s) given by a weighted summation (w_i) of basis functions $(\psi_i(s))$. The weight parameters, w_i, are treated as the policy parameters in this study where these weights need to be estimated to generate the desired robot trajectory.

We train a DMP on the latent points corresponding to one of the training trajectory, and the LWR weight coefficients are used as policy parameters which are modified to generalize to unseen environmental settings. The cost function for policy improvement is designed in the high-dimensional action space. In the current setting, we obtain a demonstration for the unseen posture and consider this as an optimized trajectory that needs to be learned by the DMP controller. The optimized trajectory is efficiently encoded by via-points extracted using the minimum jerk criterion [46] that the robot needs to pass through. The cost function is given by



(a) PCA, Latent Space (b) GPLVM, Latent Space (c) BGPLVM, Latent Space

Figure 5. Comparison of latent spaces learned by LVMs: a) PCA, b) GPLVM and c) BGPLVM. The axes are given by the most significant dimension ordered using either eigen-values or the ARD kernel weights.

the sum of all errors between the current policy and the desired via-points:

$$R(\tau) = \sum_{i=1}^{n_{\text{dims}}} \sum_{j=1}^{n_{\text{via}}} \|V_{i,j} - x_i^{recons}(t_{i,j})\|^2 \quad (8)$$

where $R(\tau)$ is the total reward for trajectory τ, $V_{i,j}$ is the j^{th} via-point of i^{th} dimension and $x_i^{recons}(t)$ is the value at time t for i^{th} dimension of reconstructed trajectory. The number of via-points n_{via} was estimated such that the reconstruction error for the generated trajectory is lower than a user-specified threshold.

4. Results

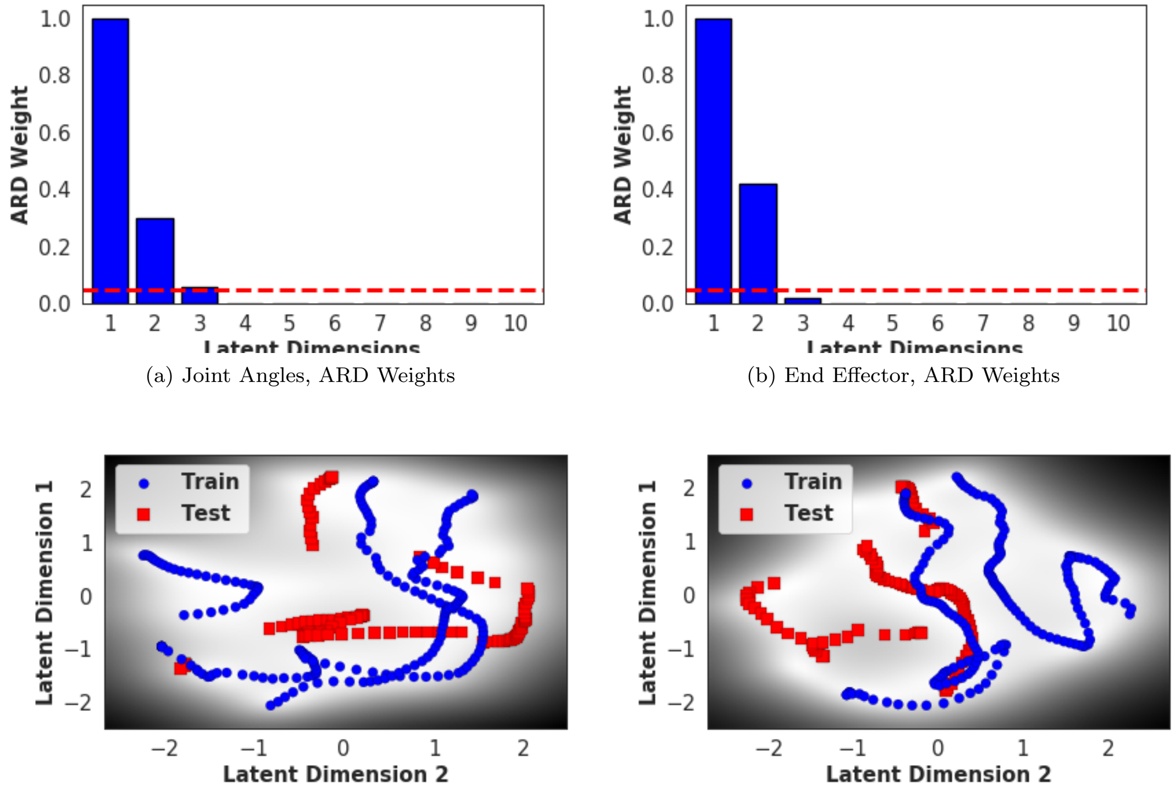
In this section, we present the performance of our proposed framework in a practical setting of robotic clothing assistance. The experimental setup includes Baxter research robot, a soft mannequin as the subject and a T-shirt as the clothing article. We consider the performance of the clothing and unclothing tasks of the T-shirt. These tasks follow different dynamics and running a clothing demonstration backward for unclothing usually leads to failure. The evaluation dataset contains demonstrations performed by three experienced users interacting with the Baxter robot. For each demonstrator, clothing and unclothing demonstrations were recorded for 6 different postures of the mannequin wherein the pair of shoulder elevation and head elevations were as follows $\{(65^o, 30^o), (70^o, 30^o), (75^o, 30^o), (70^o, 45^o), (75^o, 45^o), (80^o, 45^o)\}$. These postures cover the entire range for which the robot can successfully perform clothing tasks thereby spanning all feasible postures.

4.1 *Comparison of Latent Variable Models*

In this experiment, we inspect the motor-skills, i.e. latent features learned by BGPLVM and evaluate the predictive performance in comparison to other LVMs such as Principal Component Analysis (PCA) [6] and GPLVM [42]. PCA is a linear dimensionality reduction technique where the observations $\mathbf{Y}$ are assumed to be generated from the following generative process:

$$\mathbf{Y} = \mathbf{W}\mathbf{X} \quad (9)$$

where $\mathbf{X}$ is the unobserved latent position and $\mathbf{W} \in \mathrm{R}^{D\times d}$ comprises of a set of orthonormal basis vectors that maximizes the scatter of latent points. On the other hand, GPLVM has a similar generative process to BGPLVM given in Equation 1. However, the unknown latent positions $\mathbf{X}$ are evaluated using Maximum-A-Posteriori (MAP) estimation along with jointly maximizing the hyperparameters. The objective function for optimizing the latent positions $\mathbf{X}$ is derived from



(a) Joint Angles, ARD Weights

(b) End Effector, ARD Weights

(c) Joint Angles, ARD Weights

(d) End Effector, ARD Weights

Figure 6. Comparison of ARD weights and latent space for different feature representations

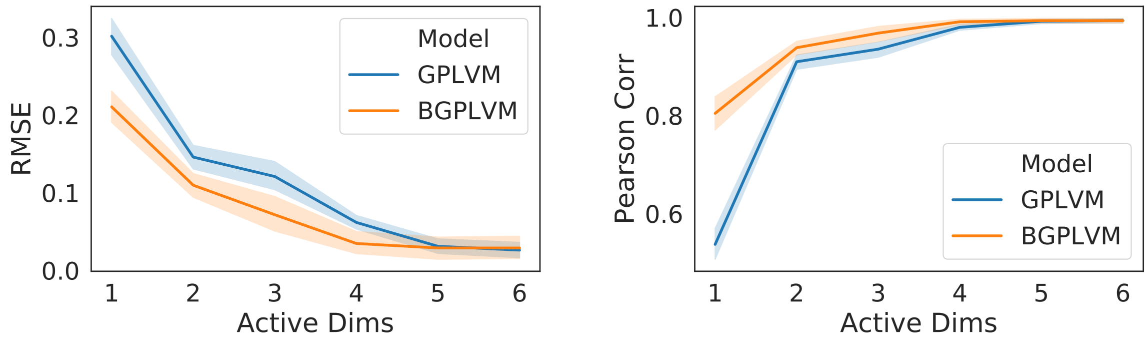
the conditional log-likelihood given in Equation 2:

$$L = -\frac{DN}{2}\ln(2\pi) - \frac{D}{2}\ln|\mathbf{K}| - \frac{1}{2}\mathrm{tr}(\mathbf{K}^{-1}\mathbf{Y}\mathbf{Y}^{\mathrm{T}}) \tag{10}$$

Firstly, we consider the dataset of demonstrations from the evaluation dataset for both clothing and unclothing tasks. For each task, six-fold cross-validation was performed where demonstrations for four postures were used as training data and for two postures were used as test data in each fold. BGPLVM models were trained for both the kinematic space (JA) and task space (EE) representations. The ARD weights for BGPLVM on training resulted in the varying number of active latent dimensions for different feature representations as shown in Figure 6. For the clothing task, the JA representation resulted in three active latent dimensions on average and two dimensions for the EE representation. This implies that the encoded motor-skills varies depending on the input observations and can be an essential design consideration depending on the task.

It was observed that the latent space varied for different LVMs. Latent spaces learned for the clothing task, and JA representation is shown in Figure 5. The grayscale background for GPLVM and BGPLVM indicate the predictive variance of the GP mapping which can be considered as a measure of uncertainty of the model. In general, all the LVMs resulted in latent trajectories for the training data which correspond to the demonstrations. There is a considerable difference in the predictive variance where BGPLVM is more confident than the GPLVM model assigning lower predictive variance over a larger region and hence effectively using the latent space.

BGPLVM can be used to automatically infer the dimensionality of the latent space through the ARD kernel weights. This can be very useful for complex tasks such as clothing assistance



(a) Normalized RMSE (b) Pearson Correlation

Figure 7. Comparison of LVMs as a dependence on number of active dimensions used for reconstruction.

where it is difficult to estimate the inherent dimensionality given high dimensional trajectories of the robot performing the task. Although GPLVM also uses an ARD kernel, the reconstruction error remains high on using only the active dimensions as shown in Figure 7. This indicates that Bayesian inference proves useful in jointly learning the model hyperparameters along with optimizing the latent positions. On the other hand, MAP inference used by GPLVM leads to poor learning of the kernel hyperparameters such as the ARD relevance weights. The EE representation required more latent dimensions in comparison to JA representation, this could indicate that the variability to adapt to different postures could not be captured in the EE representation.

The predictive performance of BGPLVM was evaluated by comparing the reconstruction error with two other LVM, i.e. PCA and GPLVM. For each demonstrator, six-fold cross-validation was performed where four demonstrations were used as training data. The remaining two demonstrations (Test Pose) along with demonstrations from an unseen demonstrator (Test Demo) were used as test data. To evaluate generalization capability, we evaluated reconstruction error given by comparing the input data and the reconstructed data from the latent points corresponding to the input,

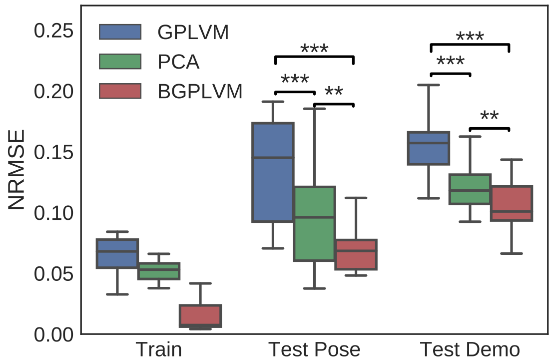
$$\begin{aligned} \mathrm{Err} &= \|y_{\mathrm{org}} - y_{\mathrm{pred}}\|, \\ y_{\mathrm{pred}} &= f_{\mathrm{model}}(f^{-1}_{\mathrm{model}}(y_{\mathrm{org}})) \end{aligned} \tag{11}$$

where y_{org} is an input sample from dataset, y_{pred} is the predicted value after reconstruction and f_{model} is the forward mapping from latent space to observation space.

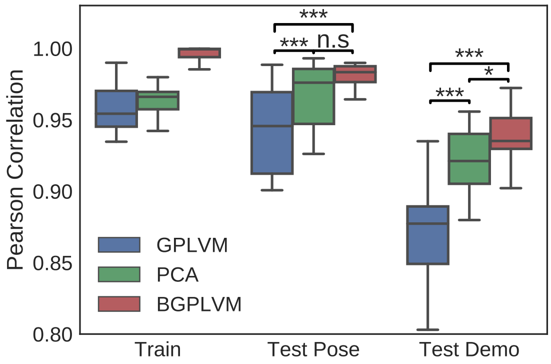
The results for reconstruction error are provided in Figure 8. We evaluated Normalized Root Mean Square Error (NRMSE) and Pearson correlation as the metrics. The Wilcoxon signed rank sum test [47] was used to assess the statistical significance and the p-value was evaluated for a one-sided test. It can be seen that BGPLVM has the best predictive performance which can be considered as a measure of generalization capability of BGPLVM latent space to unseen environmental settings.

4.2 *Evaluation of Latent Space Controller*

In this experiment, we evaluate the latent space controller for imparting motor-skills to the robot. For the evaluation, BGPLVM models were trained on kinesthetic demonstrations of performing clothing and unclothing task for a given posture of the mannequin. The model is used as an interface to control the robot and reproduce the task by performing cursor control on the latent space of the demonstration. An instance of performing cursor control to generate a different



(a) Normalized RMSE



(b) Pearson Correlation

Figure 8. Comparison of generalizability by LVMs over evaluation dataset a) Normalized RMS error, b) Pearson Correlation

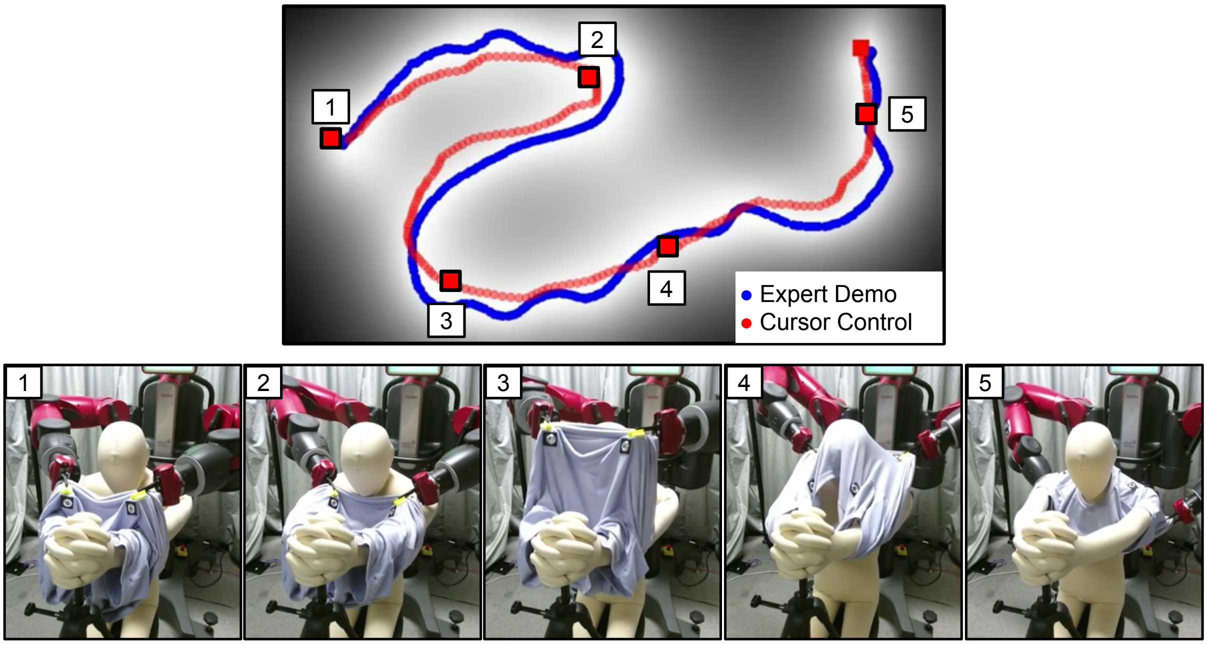


Figure 9. Example of using latent space controller to perform clothing assistance trajectory. Red trajectory indicates path traced by user with cursor control.

trajectory is shown in Figure 9. It can be seen that tracing a trajectory different from the expert demonstration used to train the model still leads to the successful execution of the clothing assistance trajectory. Furthermore, the latent trajectory for test data that minimizes reconstruction error could be discontinuous as shown in Figure 5 but tracing a smooth trajectory with cursor control still leads to successful execution. A video demonstration with the exploration of the latent dimensions is also available at `https://youtu.be/bul5lq44ru8`.

The latent dimensions learned by BGPLVM capture a specific aspect of the clothing motor-skills. For the joint angle scenario, the most significant dimension captured the horizontal motion of the arms along the mannequin while maintaining the constraints for clothing. The second dimension captured various vertical motions of pulling up the T-shirt in the beginning and pulling it down along the torso at the end. The third dimension captured variations in joint configurations across the demonstrations and could explain the constraint of safe human-robot interaction. This makes the interface intuitive to the users for planning a desired alteration to the robot trajectory. The robot trajectories generated corresponding to a trajectory along each

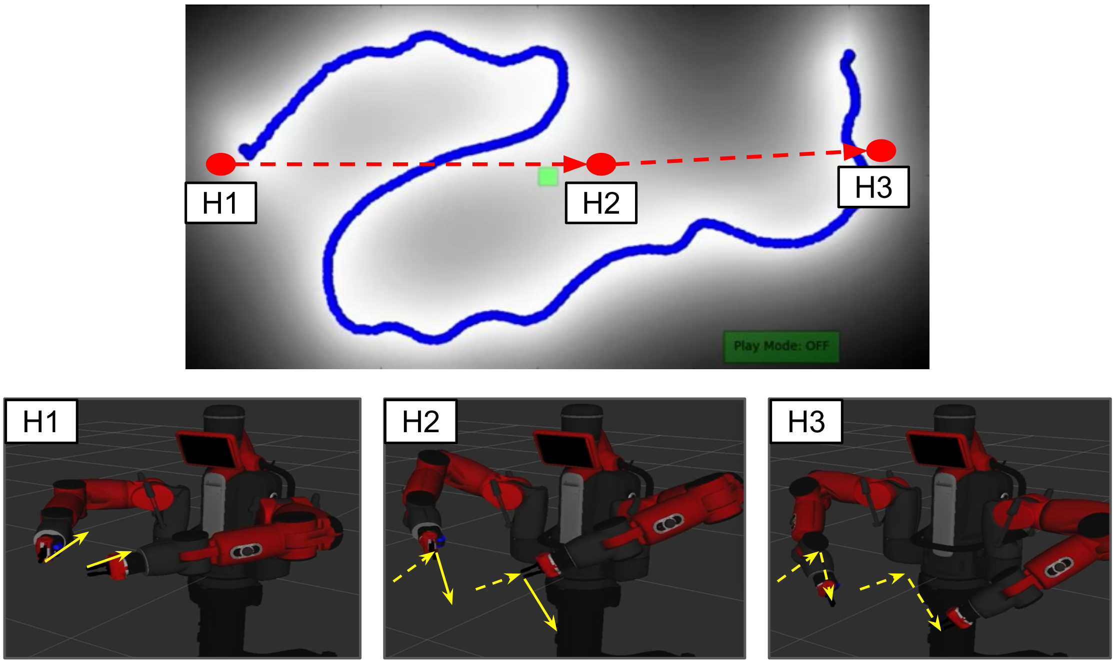


Figure 10. Robot trajectory generated corresponding to horizontal latent dimension.

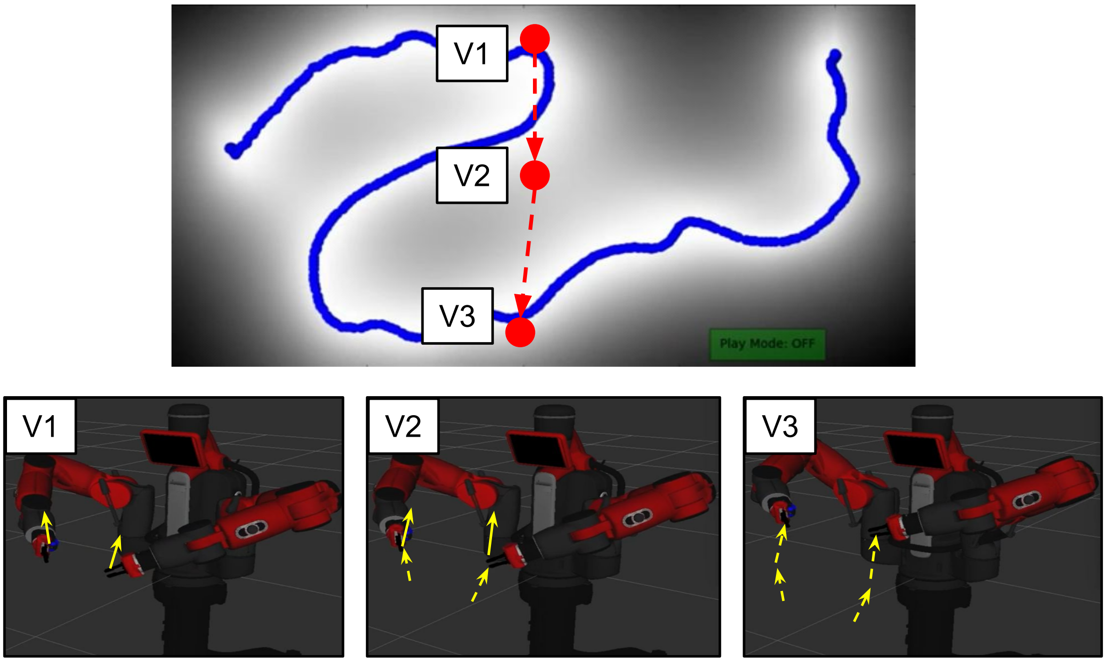
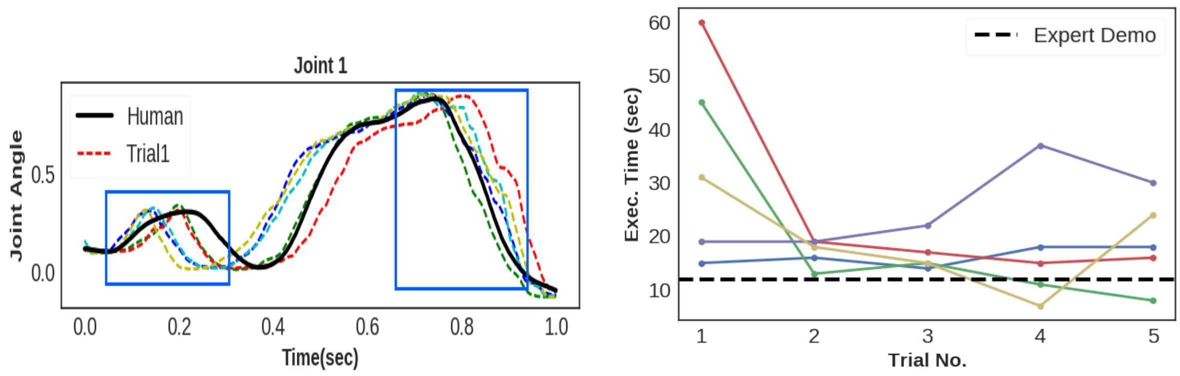


Figure 11. Robot trajectory generated corresponding to vertical latent dimension.

latent dimension is shown in Figures 10,11.

Five subjects without prior experience of interacting with the robot were asked to use the interface to reproduce the tasks. The joint angles along with proprioceptive information were recorded while the interface was used. In general, the subjects were able to reproduce the clothing demonstration even when the latent trajectory was different from the training latent points. The subjects had fast learning curves with the execution time for performing the demonstration decreasing drastically within five trials of using the interface as shown in Figure 12a. There were also instances where the execution was slower as the subject felt that the T-shirt was stuck on



(a) Trajectory Variability

(b) Execution Time

Figure 12. Evaluation of latent space controller: a) Variation in dynamics of performing clothing task using latent space controller, b) Execution time of five novice users interacting with the latent space controller over five trials.

the mannequin's head. This indicates that the controller can be used to modulate the dynamics of the task depending on the current setting. A video demonstration of an inexperienced subject interacting with the real-time controller is available at `https://youtu.be/bul5lq44ru8`.

The subjects also learned to modulate the dynamics of performing the task wherein crucial parts were performed slowly and parts without much human-cloth interaction being performed quickly. Figure 12b shows the modulation in the dynamics of performing the task for one of the subjects. The time normalized trajectories for a single joint of the robot is shown for five trials of using the interface along with the original kinesthetic demonstration. For the regions indicated in the blue squares, it can be seen that the trajectories through the interface vary drastically from the original demonstration. This indicates that novel motor-skills for performing complex tasks can be imparted by inexperienced users as well.

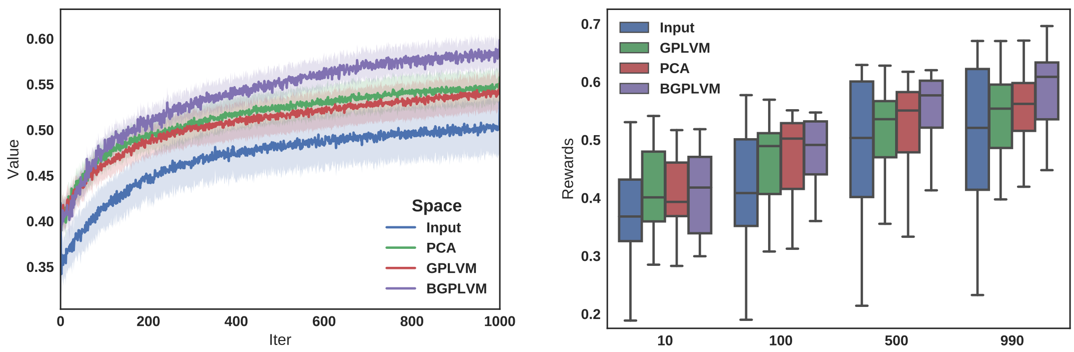
4.3 *Latent Space Reinforcement Learning*

In this section, we demonstrate the effectivity of performing policy search in the BGPLVM latent space by comparison with search in the high-dimensional action space and latent spaces learned using PCA and GPLVM. PoWER algorithm was implemented with the same state and reward representations, and only the policy search space varied depending on the scenario. For each demonstrator, six-fold cross-validation was performed where three demonstrations of the clothing task were used as training data for learning latent spaces. The latent space of each LVM was used as a search space for reinforcement learning to learn the policy corresponding to the remaining three unseen demonstrations. Policy search was run for 1000 iterations where 20 rollouts with the best overall rewards were used to update the policy in each iteration. The initial policy was generated from DMP that is fit on a latent space trajectory of one of the human demonstrations used to train the latent space.

The performance of using BGPLVM as a search space was exhaustively evaluated by considering the complete dataset. The average learning curves over the entire dataset are shown in Figure 13. It can be seen that BGPLVM scenario outperforms all other search spaces. This indicates that BGPLVM captures the necessary motor-skills most efficiently and also has the best generalizability in comparison to other LVMs.

5. Discussion

In Section 4, we demonstrated the applicability of using BGPLVM to learn meaningful latent representations for the task of robotic clothing assistance. The results in Section 4.1 indicate



(a) Learning Curves (b) Reward Box Plots

Figure 13. Comparison of different search spaces for policy search shown as a bar plot over the evaluation dataset.

that non-linear mapping and Bayesian inference are necessary for the most efficient encoding of motor-skills. PCA relies on a linear model thereby constraining its capability. GPLVM has performance similar to PCA as it relies on a MAP estimate of the latent space and it seems to be overfitting to the training data. Overfitting of the GP hyperparameters could lead to worse performance on using a non-linear mapping. GPLVM could achieve better performance over PCA on the careful tuning of the hyperparameter values. On the other hand, BGPLVM relies on full Bayesian inference where a posterior distribution on the latent space is learned. This leads to superior performance as a model of the appropriate complexity can be learned automatically, and the advantage of using a non-linear mapping can also be utilized. This validates our hypothesis that encoding motor-skills using Bayesian nonparametric latent variable models is a useful parametrization for learning.

The results in Section 4.2 indicate that BGPLVM is also able to learn interpretable latent features which are very useful for skill-transfer. In the coming years, there will be an increasing presence of assistive robots in healthcare facilities, and it is essential to design user-friendly interfaces for the caregivers to adapt to such technologies. The interface presented in this study could have several extensions making it suitable for other assistive tasks. It could include the pose of the assisted person which can be obtained using skeleton tracking. Section 4.3 presents the feasibility of using BGPLVM latent space as a search space for reinforcement learning. Policy search is performed using the PoWER algorithm which has been shown to perform well in several applications. Currently, dimensionality reduction (DR) and policy search (RL) are performed independently, and the latent features do not adapt to the rewards obtained during policy search. Luck et al. [27] proposed an Expectation Maximization (EM) formulation which inherently combines DR with policy search. However, they relied on a linear mapping from the latent space to observation space. Interesting future work will be to combine policy search with BGPLVM through such an EM-like formulation.

Recently, there has been an increasing emphasis on Deep Learning which is capable of solving complex tasks. However, these methods usually fail when presented with small datasets. The primary advantage of using a nonparametric method is its efficiency in the low data regime. This is crucial for assistive robotics due to the difficulty in collecting large datasets. BGPLVM is a memory-based method and has prohibitive training costs. The computational complexity to train a model scales with the size of the training dataset (n) and the number of inducing points used to approximate the dataset (m), i.e. $O(nm^2)$ [4]. However, both the applications of using BGPLVM proposed in this study relies on the inference of high-dimensional joint angles given a latent position. The latent position is specified from a human in LfD and through policy search in RL. During test inference, the training dataset remains fixed and so precomputed values of the kernel matrix can be used to reduce time-complexity to $O(n)$ which makes it much smaller than training instance and suitable for real-time implementation. The focus of this study is on the policy or action representation used for learning clothing assistance skills. A trajectory-centric reward function is used to perform the policy update. This method can be combined along with our previous work [39, 40] which focuses on efficient state estimation to formulate an entire RL framework. The human-cloth relationship estimated by our framework can be used as the state representation. A reward function defined in this state space will be used to perform policy update in the BGPLVM action space proposed in this study.

Our proposed framework could also be extended to other tasks that involve complex motor-skills and require data-efficiency for practical implementation. Related tasks such as dressing of other clothing articles, cloth folding, non-rigid object manipulation and so on can be potential applications. The primary requirement to apply our framework is expert demonstrations which can be used to learn efficient latent space representations. A limitation of the current study is that a soft mannequin is used as a subject. However, a concurrent study [48] evaluated the usability of performing clothing assistance using the latent space controller with ten subjects and was able to demonstrate its applicability. This framework can be extended towards a practical scenario of robotic dressing assistance implemented in a nursing care facility. A shared mobile and dual-arm dressing robot can be made accessible to about 20 residents of the care home. End-to-end dressing tasks can be performed by integrating related methods such as cloth extremity [49] and cloth grasp point detection [50]. Factors such as sociability and personalization to each user will also be considered in the implementation.

6. Conclusion

In this study, we have presented the use of BGPLVM as a representation for encoding motor-skills to perform clothing assistance task. We demonstrated the applicability of this framework as an intuitive and user-friendly tool for LfD. The experimental results indicate our method as a promising approach for learning in combination with RL. There can be several extensions to this work. A latent space representation can be learned from multiple observation spaces using Manifold Relevance Determination (MRD) [41]. Our future work will be to learn models that incorporate visual information about the relationship between human and cloth. The long term goal is to develop a data-efficient policy search RL framework that unifies learning the latent space simultaneously with policy learning. We will also evaluate our learning framework on human subjects along with a real-world implementation of RL.

Acknowledgement

This work was supported in part by ImPACT Program of Council for Science, Technology, and Innovation (Cabinet Office, Government of Japan) 2015-PM07-36-01. This work was also supported in part by the Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science (No. 16H01749, No. 16H06534 and No. 17H05863) and the NAIST Big-Data Project.

References

[1] Kober J, Bagnell JA, Peters J. Reinforcement Learning in Robotics: A Survey. The International Journal of Robotics Research. 2013;32(11):1238–1274.
[2] Deisenroth MP, Neumann G, Peters J, et al.. A Survey on Policy Search for Robotics. Foundations and Trends® in Robotics. 2013;2(1–2):1–142.
[3] Tamei T, Matsubara T, Rai A, Shibata T. Reinforcement Learning of Clothing Assistance with a Dual-Arm Robot. In: IEEE-RAS International Conference on Humanoid Robots (Humanoids). IEEE. 2011. p. 733–738.
[4] Titsias MK, Lawrence ND. Bayesian Gaussian Process Latent Variable Model. In: AISTATS. Vol. 9. 2010. p. 844–851.

[5] Kober J, Peters JR. Policy Search for Motor Primitives in Robotics. In: Advances in Neural Information Processing Systems. 2009. p. 849–856.
[6] Jolliffe I. Principal Component Analysis. Wiley Online Library. 2002.
[7] Brunskill E, Roy N. SLAM using Incremental Probabilistic PCA and Dimensionality Reduction. In: IEEE International Conference on Robotics and Automation (ICRA).. IEEE. 2005. p. 342–347.
[8] Calinon S, Guenter F, Billard A. On Learning, Representing, and Generalizing a Task in a Humanoid Robot. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics). 2007;37(2):286–298.
[9] Ciocarlie M, Goldfeder C, Allen P. Dimensionality Reduction for Hand-independent Dexterous Robotic Grasping. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).. IEEE. 2007. p. 3270–3275.
[10] Artemiadis PK, Kyriakopoulos KJ. EMG-based Control of a Robot Arm using Low-dimensional Embeddings. IEEE Transactions on Robotics. 2010;26(2):393–398.
[11] Watter M, Springenberg J, Boedecker J, Riedmiller M. Embed to control: A locally linear latent dynamics model for control from raw images. In: Advances in neural information processing systems. 2015. p. 2746–2754.
[12] van Hoof H, Chen N, Karl M, van der Smagt P, Peters J. Stable Reinforcement Learning with Autoencoders for Tactile and Visual Data. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE. 2016. p. 3928–3934.
[13] Shon AP, Grochow K, Hertzmann A, Rao RP. Learning Shared Latent Structure for Image Synthesis and Robotic Imitation. Advances in Neural Information Processing Systems. 2006;18:1233.
[14] Ko J, Fox D. Learning GP-BayesFilters via Gaussian Process Latent Variable Models. Autonomous Robots. 2011;30(1):3–23.
[15] Wang Z, Mülling K, Deisenroth MP, Ben Amor H, Vogt D, Schölkopf B, Peters J. Probabilistic Movement Modeling for Intention Inference in Human-Robot Interaction. The International Journal of Robotics Research. 2013;32(7):841–858.
[16] Nakamura T, Nagai T, Mochihashi D, Kobayashi I, Asoh H, Kaneko M. Segmenting Continuous Motions with Hidden Semi-markov Models and Gaussian Processes. Frontiers in neurorobotics. 2017; 11:67.
[17] Koskinopoulou M, Piperakis S, Trahanias PE. Learning from Demonstration Facilitates Human-Robot Collaborative Task Execution. In: ACM/IEEE International Conference on Human Robot Interaction (HRI). IEEE Press. 2016. p. 59–66.
[18] Doumanoglou A, Kargakos A, Kim TK, Malassiotis S. Autonomous active recognition and unfolding of clothes using random decision forests and probabilistic planning. In: Robotics and automation (icra), 2014 ieee international conference on. IEEE. 2014. p. 987–993.
[19] Huang SH, Pan J, Mulcaire G, Abbeel P. Leveraging appearance priors in non-rigid registration, with application to manipulation of deformable objects. In: Intelligent robots and systems (iros), 2015 ieee/rsj international conference on. IEEE. 2015. p. 878–885.
[20] Yang PC, Sasaki K, Suzuki K, Kase K, Sugano S, Ogata T. Repeatable folding task by humanoid robot worker using deep learning. IEEE Robotics and Automation Letters. 2017;2(2):397–403.
[21] Balaguer B, Carpin S. Combining imitation and reinforcement learning to fold deformable planar objects. In: Intelligent robots and systems (iros), 2011 ieee/rsj international conference on. IEEE. 2011. p. 1405–1412.
[22] Monsó P, Alenyà G, Torras C. Pomdp approach to robotized clothes separation. In: Intelligent robots and systems (iros), 2012 ieee/rsj international conference on. IEEE. 2012. p. 1324–1329.
[23] Lee AX, Lu H, Gupta A, Levine S, Abbeel P. Learning force-based manipulation of deformable objects from multiple demonstrations. In: IEEE International Conference on Robotics and Automation (ICRA). IEEE. 2015. p. 177–184.
[24] Hu Z, Sun P, Pan J. Three-Dimensional Deformable Object Manipulation Using Fast Online Gaussian Process Regression. IEEE Robotics and Automation Letters. 2018;3(2):979–986.
[25] Deisenroth MP, Fox D, Rasmussen CE. Gaussian Processes for Data-efficient Learning in Robotics and Control. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2015;37(2):408–423.
[26] Bitzer S, Howard M, Vijayakumar S. Using Dimensionality Reduction to Exploit Constraints in Reinforcement Learning. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE. 2010. p. 3219–3225.
[27] Luck KS, Neumann G, Berger E, Peters J, Amor HB. Latent Space Policy Search for Robotics. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE. 2014. p.

1434–1440.
[28] Zhang C, Zhang H, Parker LE. Feature Space Decomposition for Effective Robot Adaptation. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE. 2015. p. 441–448.
[29] Burleson W, Lozano C, Ravishankar V, Rowe J, Mahoney E, Mahoney D. Assistive dressing system: A capabilities study for personalized support of dressing activities for people living with dementia. Iproceedings. 2015;1(1):e13.
[30] Orr C, Nugent C, Wang H, Zheng H. A multi-agent approach to assist with dressing in a smart environment. In: ehealth 360. Springer. 2017. p. 220–223.
[31] Klee SD, Ferreira BQ, Silva R, Costeira JP, Melo FS, Veloso M. Personalized assistance for dressing users. In: Social robotics. Springer. 2015. p. 359–369.
[32] Clegg A, Tan J, Turk G, Liu CK. Animating human dressing. ACM Transactions on Graphics (TOG). 2015;34(4):116.
[33] Erickson Z, Clegg A, Yu W, Turk G, Liu CK, Kemp CC. What does the person feel? learning to infer applied forces during robot-assisted dressing. In: Robotics and automation (icra), 2017 ieee international conference on. IEEE. 2017. p. 6058–6065.
[34] Yu W, Kapusta A, Tan J, Kemp CC, Turk G, Liu CK. Haptic simulation for robot-assisted dressing. In: Robotics and automation (icra), 2017 ieee international conference on. IEEE. 2017. p. 6044–6051.
[35] Yamazaki K, Oya R, Nagahama K, Okada K, Inaba M. Bottom Dressing by a Life-sized Humanoid Robot provided Failure Detection and Recovery Functions. In: IEEE/SICE International Symposium on System Integration (SII). IEEE. 2014. p. 564–570.
[36] Colomé A, Planells A, Torras C. A Friction-model based Framework for Reinforcement Learning of Robotic Tasks in Non-rigid Environments. In: IEEE International Conference on Robotics and Automation (ICRA). IEEE. 2015. p. 5649–5654.
[37] Kapusta A, Yu W, Bhattacharjee T, Liu CK, Turk G, Kemp CC. Data-driven Haptic Perception for Robot-Assisted Dressing. In: IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). IEEE. 2016. p. 451–458.
[38] Gao Y, Chang HJ, Demiris Y. Iterative Path Optimisation for Personalised Dressing Assistance using Vision and Force Information. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE. 2016. p. 4398–4403.
[39] Koganti N, Ngeo JG, Tomoya T, Ikeda K, Shibata T. Cloth Dynamics Modeling in Latent Spaces and its Application to Robotic Clothing Assistance. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE. 2015. p. 3464–3469.
[40] Koganti N, Tamei T, Ikeda K, Shibata T. Bayesian Nonparametric Learning of Cloth Models for Real-Time State Estimation. IEEE Transactions on Robotics. 2017;33(2):916–931.
[41] Damianou AC, Titsias MK, Lawrence ND. Variational Inference for Latent Variables and Uncertain Inputs in Gaussian Processes. The Journal of Machine Learning Research. 2016;17(1):1425–1486.
[42] Lawrence N. Probabilistic Non-linear Principal Component Analysis with Gaussian Process Latent Variable Models. Journal of Machine Learning Research. 2005;6(Nov):1783–1816.
[43] Rasmussen C, Williams C. Gaussian Processes for Machine learning. The MIT Press, Cambridge. 2006;10.
[44] Ijspeert AJ, Nakanishi J, Hoffmann H, Pastor P, Schaal S. Dynamical Movement Primitives: Learning Attractor Models for Motor Behaviors. Neural Computation. 2013;25(2):328–373.
[45] Schaal S, Atkeson CG. Constructive Incremental Learning from only Local Information. Neural Computation. 1998;10(8):2047–2084.
[46] Wada Y, Kawato M. A Theory for Cursive Handwriting based on the Minimization Principle. Biological Cybernetics. 1995;73(1):3–13.
[47] Wilcoxon F. Individual Comparisons by Ranking Methods. Biometrics Bulletin. 1945;1(6):80–83.
[48] Joshi RP, Koganti N, Shibata T. A Framework for Robotic Clothing Assistance by Imitation Learning. Advanced Robotics. 2019;Under Review.
[49] Ngo G, Gaurav V, Shibata T. Clothing Extremity Identification Using Convolutional Neural Network Regressor. In: IEEE International Conference on Informatics, Electronics & Vision (ICIEV) and International Conference on Imaging, Vision & Pattern Recognition (icIVPR). IEEE. 2018. p. 436–441.
[50] Saxena K, Labuguen R, Joshi RP, Koganti N, Shibata T. A Study on Human-Robot Collaboration for Table-setting Task. In: IEEE International Conference on Robotics and Biomimetics (ROBIO). IEEE. 2017. p. 183–188.