Robot Learning with Invariant Hidden Semi-Markov Models

Ajay Kumar Tanwani, Jonathon Lee, Michael Laskey, Sanjay Krishnan, Roy Fox, Ken Goldberg
University of California, Berkeley
{ajay.tanwani, jonathon_lee, mdlaskey, sanjaykrishnan, royf, goldberg}@berkeley.edu

***Abstract*—Generalizing manipulation skills to new situations requires extracting invariant patterns from demonstrations. For example, the robot needs to understand the demonstrations at a higher level while being invariant to the appearance of the objects, geometric aspects of objects such as its position, size, orientation, viewpoint of the observer in the demonstrations. In this paper, we learn a joint probability density function of the demonstrations with invariant formulations of hidden semi-Markov model, and smoothly follow the generated sequence of states with a linear quadratic tracking controller. We present parsimonious and Bayesian non-parametric online learning formulations of the HSMM to exploit the invariant segments (also termed as sub-goals, options or actions) in the demonstrations and adapt the movement in accordance with the external environmental situations such as size, position and orientation of the objects in the environment using a task-parameterized formulation. We show an application of robot learning from demonstrations in picking and placing an object while avoiding a moving obstacle.**

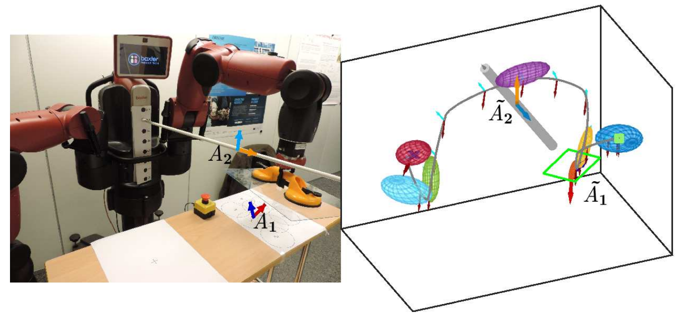

Fig. 1: *(left)* Baxter robot picks the glass plate with a suction lever and places it on the cross after avoiding an obstacle of varying height, *(right)* reproduction for previously unseen object and obstacle position.

I. Introduction

Generative models are widely used in imitation learning to learn the distribution of the data for regenerating new samples from the model [11, 1, 12]. Common examples include probability density function estimation, image regeneration and so on. The focus of this paper is to learn the joint probability density function of the human demonstrations with a **Hidden Semi-Markov Model (HSMM)** in an **unsupervised** manner. HSMMs replace the self-transition probabilities of staying in a state with an explicit model of state duration [16]. This helps to adequately bias the generated motion with longer state dwell times for skill acquisition. We show how the model can be systematically adopted to changing situations such as position/size/orientation of the objects in the environment with a task-parameterized formulation. We combine tools from statistical machine learning and optimal control to segment the demonstrations into different components or sub-goals that are sequenced together to perform manipulation tasks.

In this paper, we unify two formulations of the HSMM for encoding and decoding of real-world robot manipulation tasks under varying environmental situations (see Fig. 2): 1) segmentation and dimensionality reduction simultaneously to impose a parsimonious structure on the covariance matrix and reduce the number of parameters that can be robustly estimated [13], and 2) Bayesian non-parametric formulation of HSMM with Hierarchical Dirichlet process (HDP) for online learning under small variance asymptotics [14]. Our objective is to reduce the number of demonstrations required for learning a new task, while ensuring effective generalization in new environmental situations.

II. Hidden Semi-Markov Models

Hidden Markov models (HMMs) encapsulate the spatio-temporal information by augmenting a GMM with latent states that sequentially evolve over time in the demonstrations. HMM is thus defined as a doubly stochastic process, one with sequence of hidden states and another with sequence of observations/emissions. Spatio-temporal encoding with HMMs can handle movements with variable durations, recurring patterns, options in the movement, or partial/unaligned demonstrations [9, 6, 4, 7]. **Semi-Markov models** relax the Markovian structure of state transitions by relying not only upon the current state but also on the duration/elapsed time in the current state, i.e., the underlying process is defined by a *semi-Markov* chain with a variable duration time for each state. The state duration stay is a random integer variable that assumes values in the set $\{1, 2, \ldots, s^{\max}\}$. The value corresponds to the number of observations produced in a given state, before transitioning to the next state. **Hidden Semi-Markov Models** (HSMMs) associate an observable output distribution with each state in a semi-Markov chain, similar to how we associated a sequence of observations with a Markov chain in a HMM.

To make it concrete, let $\{\boldsymbol{\xi}_t\}_{t=1}^{T}$ denote the sequence of observations with $\boldsymbol{\xi}_t \in \mathbb{R}^D$ collected while demonstrating a manipulation task. The state may represent the visual observation, kinesthetic data such as the pose and the velocities of the end-effector of the human arm, haptic information,

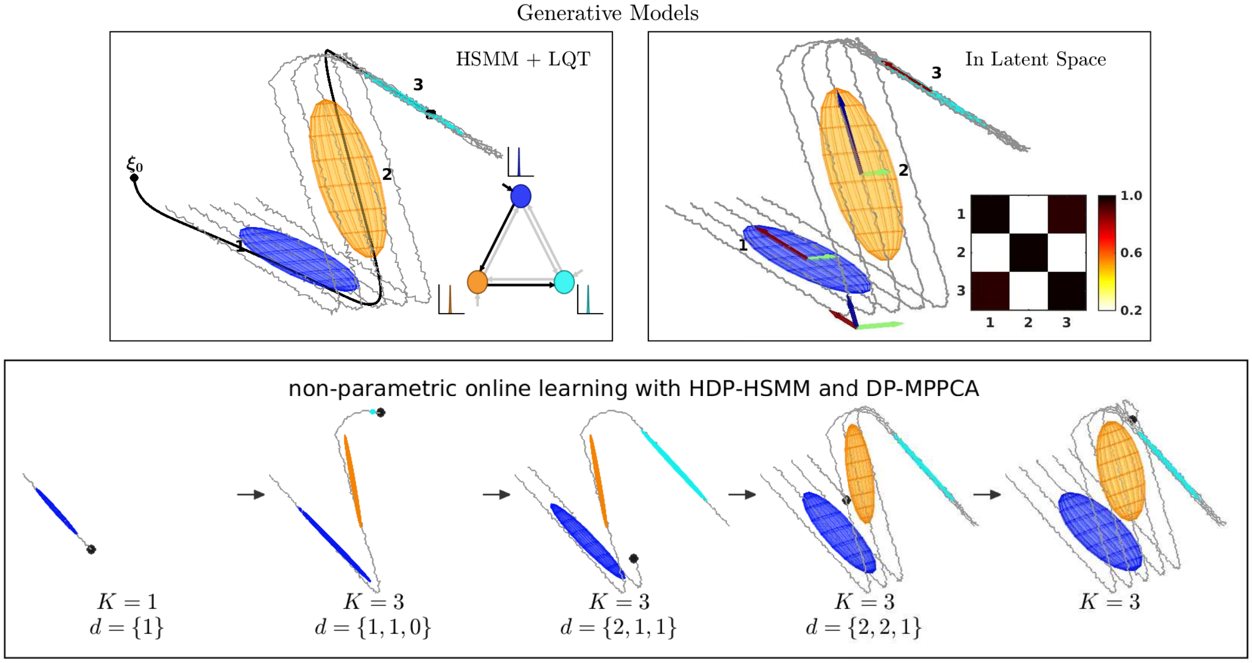

Fig. 2: HSMM formulations for learning on synthetic Z-shaped data: *(top-left)* demonstrations encoded with a hidden semi-Markov model and decoded with a linear quadratic tracking controller, *(top-right)* latent space encoding for parsimonious representation, *(bottom)* Bayesian non-parametric online sequence clustering.

or any arbitrary features defining the task variables of the environment. The observation sequence is associated with a hidden state sequence $\{z_t\}_{t=1}^{T}$ with $z_t \in \{1 \dots K\}$ belonging to the discrete set of K cluster indices. The cluster indices correspond to different segments of the task such as reach, grasp, move etc. The transition between one segment i to another segment j is denoted by the transition matrix $\boldsymbol{a} \in \mathbb{R}^{K \times K}$ with $a_{i,j} \triangleq P(z_t = j | z_{t-1} = i)$. The parameters $\{\mu_j^S, \Sigma_j^S\}$ represent the mean and the standard deviation of staying s consecutive time steps in state j estimated by a Gaussian $\mathcal{N}(s|\mu_j^S, \Sigma_j^S)$. The hidden state follows a multinomial distribution with $z_t \sim \text{Mult}(\boldsymbol{\pi}_{z_{t-1}})$ where $\boldsymbol{\pi}_{z_{t-1}} \in \mathbb{R}^K$ is the next state transition distribution over state z_{t-1} with Π_i as the initial probability, and the observation $\boldsymbol{\xi}_t$ is drawn from the output distribution of state j, described by a multivariate Gaussian with parameters $\{\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j\}$. The overall parameter set for an HSMM is defined by $\left\{\Pi_i, \{a_{i,m}\}_{m=1}^{K}, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i, \mu_i^S, \Sigma_i^S\right\}_{i=1}^{K}$. Parameters $\left\{\Pi_i, \{a_{i,m}\}_{m=1}^{K}, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\right\}_{i=1}^{K}$ are estimated using the EM algorithm for HMMs, and the duration parameters $\{\mu_i^S, \Sigma_i^S\}_{i=1}^{K}$ are estimated empirically from the data after training using the most likely hidden state sequence $\boldsymbol{z}_t = \{z_1 \dots z_T\}$.

Given the learned model parameters $\{\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^{K}$ and a sequence of observations $\{\boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_T\}$, the probability of the hidden state sequence over the next time horizon T_p, i.e., $p(z_t, z_{t+1}, \dots, z_{T_p} \mid \boldsymbol{\xi}_1, \dots, \boldsymbol{\xi}_t)$ is decoded using the Viterbi algorithm with the forward variable of the algorithm. The decoded sequence is combined with a finite-horizon linear quadratic tracking controller for smooth retrieval of robot motion.

A. Scalable HSMMs in Latent Spaces

HSMMs tend to suffer from the well-known curse of dimensionality when few datapoints are available as in the case of learning from human demonstrations. Statistical subspace clustering methods address this challenge by using a parsimonious model to reduce the number of parameters that can be robustly estimated. Contrary to existing methods that impose a parsimonious structure on each covariance matrix separately (diagonal/isotropic matrix, low-rank decomposition), we exploit a technique to share the parameters across the mixture components along the important synergistic directions [5]. The technique associates or ties the covariance matrices of the mixture model with a common latent space, and only uses a diagonal matrix for appropriate scaling of the basis vectors in the latent space. When the covariance matrices of the mixture model share the same set of parameters for the latent feature space, we call the model a *semi-tied* mixture model. The main idea behind semi-tied mixture models is to decompose the covariance matrix $\boldsymbol{\Sigma}_i$ into two terms: a common latent feature matrix $\boldsymbol{H} \in \mathbb{R}^{D \times D}$ and a component-specific diagonal matrix $\boldsymbol{\Sigma}_i^{(\text{diag})} \in \mathbb{R}^{D \times D}$, i.e.,

$$\boldsymbol{\Sigma}_i = \boldsymbol{H} \boldsymbol{\Sigma}_i^{(\text{diag})} \boldsymbol{H}^{\top}. \tag{1}$$

The latent feature matrix encodes the locally important synergistic directions represented by D non-orthogonal basis vectors that are shared across all the mixture components, while the diagonal matrix selects the appropriate subspace of

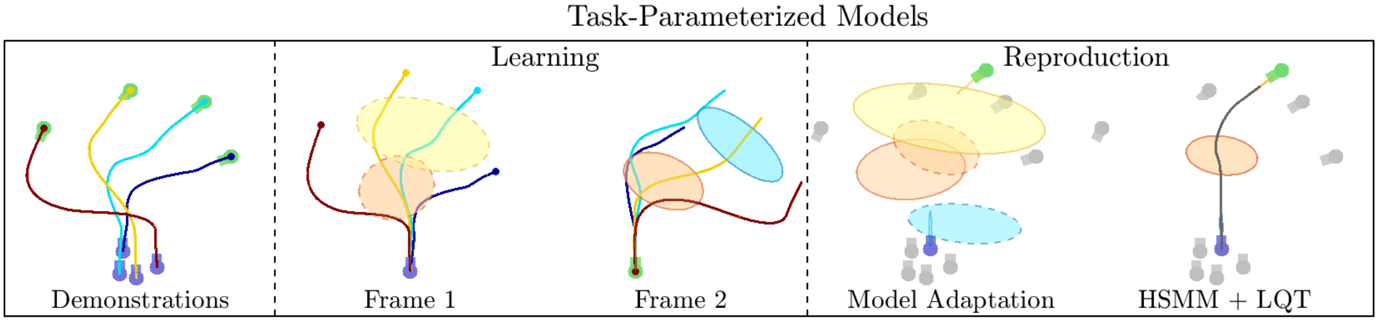

Fig. 3: Task-parameterized formulation of HSMM. The demonstration on left are observed from the coordinate systems that move with the object (starting in purple position and ending in green position in each demonstration) and the generative model is learned in the respective coordinate systems. The model parameters in respective coordinate systems are adapted to the new unseen object positions by computing the products of linearly transformed Gaussian mixture components. The resulting HSMM is combined with LQT for smooth retrieval of manipulation tasks.

each mixture component as convex combination of a subset of the basis vectors of $\boldsymbol{H}$. Note that the eigen decomposition of $\boldsymbol{\Sigma}_i = \boldsymbol{U}_i \boldsymbol{\Sigma}_i^{(\text{diag})} \boldsymbol{U}_i^{\top}$ contains D basis vectors of $\boldsymbol{\Sigma}_i$ in $\boldsymbol{U}_i$. In comparison, semi-tied mixture model gives D globally representative basis vectors that are shared across all the mixture components. The decomposition allows update of the parameters $\boldsymbol{H}$ and $\boldsymbol{\Sigma}_i^{(\text{diag})}$ in closed form with EM updates of HSMM.

B. Bayesian Non-Parametric Online HSMMs

Adapting the HSMM online with large scale streaming data and specifying the number of latent states is a challenging problem. Bayesian non-parametric methods provide flexibility in model selection, however, their widespread use is limited by the computational overhead of existing sampling-based and variational techniques for inference.

We make use of **Small Variance Asymptotic** (SVA) analysis of Bayesian non-parametric HSMM under Hierarchical Dirichlet Process (HDP) [10]. Small variance asymptotic (SVA) analysis implies that the covariance matrix $\boldsymbol{\Sigma}_{t,i}$ of all the Gaussians is set to the isotropic noise σ^2, i.e., $\boldsymbol{\Sigma}_{t,i} \approx \lim_{\sigma^2 \to 0} \sigma^2 \boldsymbol{I}$ in the likelihood function and the prior distribution [8, 2]. The analysis yields simple deterministic models, while retaining the non-parametric nature. For example, SVA analysis of the Bayesian non-parametric GMM leads to the DP-means algorithm [8]. Restricting the covariance matrix to an isotropic/spherical noise, however, fails to encode the coordination patterns in the demonstrations. Consequently, we model the covariance matrix in its intrinsic affine subspace of dimension d_i with projection matrix $\boldsymbol{\Lambda}_i^{d_i} \in \mathbb{R}^{D \times d_i}$, such that $d_i < D$ and $\boldsymbol{\Sigma}_i = \lim_{\sigma^2 \to 0} \boldsymbol{\Lambda}_i^{d_i} \boldsymbol{\Lambda}_i^{d_i \top} + \sigma^2 \boldsymbol{I}$. Under this assumption, we apply the small variance asymptotic limit on the remaining $(D - d_i)$ dimensions to encode the most important coordination patterns while being parsimonious in the number of parameters.

The analysis gives a scalable online sequence clustering algorithm that is non-parametric in the number of clusters and the subspace dimension of each cluster. The resulting algorithm groups the new datapoint in its low dimensional subspace by online inference in a non-parametric mixture of probabilistic principal component analyzers based on Dirichlet process, and captures the state transition and state duration information online in a HDP-HSMM. The cluster assignment and the parameter updates at each iteration minimize the loss function, thereby, increasing the model fitness while penalizing for new transitions, new dimensions and/or new clusters. An interested reader can find more details of the algorithm in [14].

III. Task-Parameterized HSMMs

Task-parameterized models provide a probabilistic formulation to deal with different real world situations by adapting the model parameters in accordance with the external task parameters that describe the environment/configuration/situation, instead of hard coding the solution for each new situation or handling it in an *ad hoc* manner [15, 3, 13]. When a different situation occurs (position/orientation of the object changes), changes in the task parameters/reference frames are used to modulate the model parameters in order to adapt the robot movement to the new situation.

We represent the task parameters with P coordinate systems, defined by $\{\boldsymbol{A}_j, \boldsymbol{b}_j\}_{j=1}^P$, where $\boldsymbol{A}_j$ denotes the orientation of the frame as a rotation matrix and $\boldsymbol{b}_j$ represents the origin of the frame. We assume that the coordinate frames are specified by the user, based on prior knowledge about the carried out task. Typically, coordinate frames will be attached to objects, tools or locations that could be relevant in the execution of a task. Each datapoint $\boldsymbol{\xi}_t$ is observed from the viewpoint of P different experts/frames, with $\boldsymbol{\xi}_t^{(j)} = \boldsymbol{A}_j^{-1}(\boldsymbol{\xi}_t - \boldsymbol{b}_j)$ denoting the datapoint observed with respect to frame j. The parameters of the task-parameterized HSMM are defined by $\theta = \left\{ \{\boldsymbol{\mu}_i^{(j)}, \boldsymbol{\Sigma}_i^{(j)}\}_{j=1}^P, \{a_{i,m}\}_{m=1}^K, \mu_i^S, \Sigma_i^S \right\}_{i=1}^K$, where $\boldsymbol{\mu}_i^{(j)}$ and $\boldsymbol{\Sigma}_i^{(j)}$ define the mean and the covariance matrix of i-th mixture component in frame j. Parameter updates of the task-parameterized HSMM algorithm remain the same as HSMM, except the computation of the mean and the covariance matrix

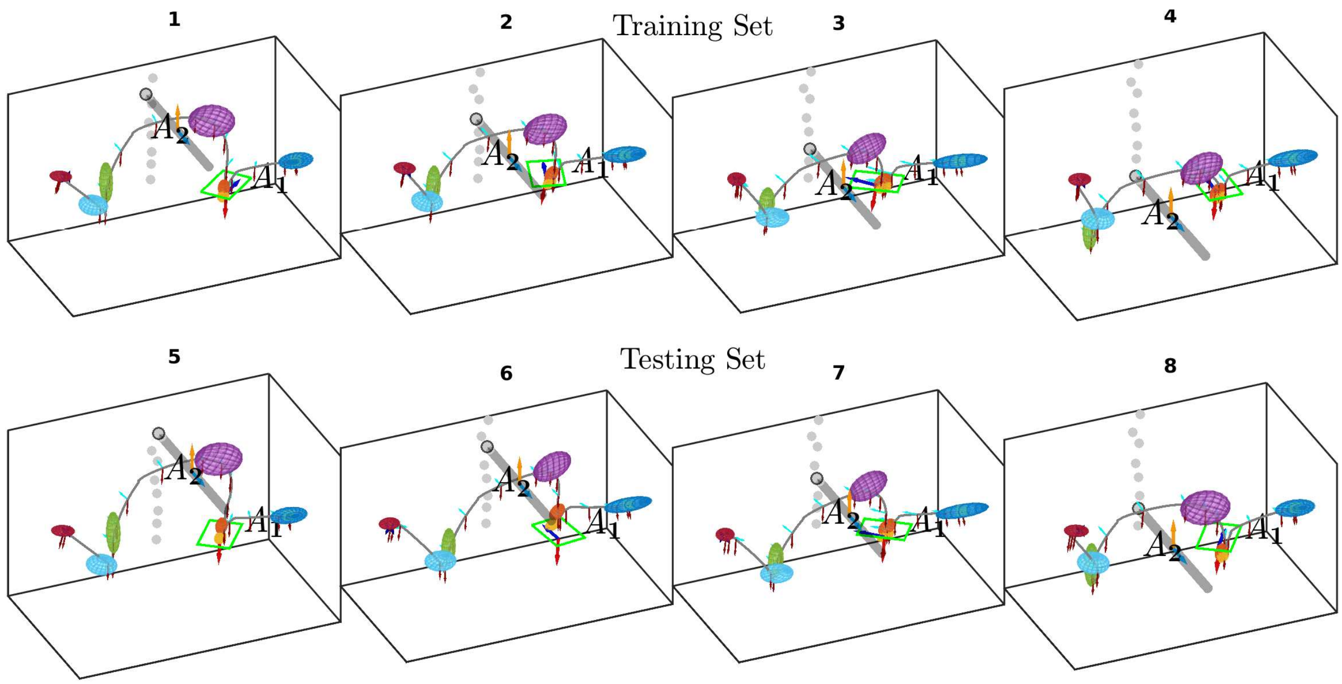

Fig. 4: Task-Parameterized Semi-Tied HSMM performance on pick-and-place with obstacle avoidance task: *(top)* training set reproductions, *(bottom)* testing set reproductions.

is repeated for each coordinate system separately.

In order to combine the output of the experts for an unseen situation represented by the frames $\{\tilde{\boldsymbol{A}}_j, \tilde{\boldsymbol{b}}_j\}_{j=1}^P$, we linearly transform the Gaussians back to the global coordinates with $\{\tilde{\boldsymbol{A}}_j, \tilde{\boldsymbol{b}}_j\}_{j=1}^P$, and retrieve the new model parameters $\{\tilde{\boldsymbol{\mu}}_i, \tilde{\boldsymbol{\Sigma}}_i\}$ for the i-th mixture component by computing the products of the linearly transformed Gaussians (see Fig. 3)

$$\mathcal{N}(\tilde{\boldsymbol{\mu}}_i, \tilde{\boldsymbol{\Sigma}}_i) \propto \prod_{j=1}^{P} \mathcal{N}\left(\tilde{\boldsymbol{A}}_j \boldsymbol{\mu}_i^{(j)} + \tilde{\boldsymbol{b}}_j, \tilde{\boldsymbol{A}}_j \boldsymbol{\Sigma}_i^{(j)} \tilde{\boldsymbol{A}}_j^\top\right). \quad (2)$$

IV. Experiments, Results and Discussion

We now show an application of our work to learn a task from a few human demonstrations. The objective of the task is to place the object in a desired target position by picking it from different initial positions and orientations of the object, while adapting the movement to avoid the obstacle. The setup of pick-and-place task with obstacle avoidance is shown in Fig. 1. The Baxter robot is required to grasp the glass plate with a suction lever placed in an initial configuration as marked on the setup. The obstacle can be vertically displaced to one of the 8 target configurations. We describe the task with two frames, one for the object initial configuration with $\{\boldsymbol{A}_1, \boldsymbol{b}_1\}$ and other for the obstacle $\{\boldsymbol{A}_2, \boldsymbol{b}_2\}$ with $\boldsymbol{A}_2 = \boldsymbol{I}$ and $\boldsymbol{b}_2$ to specify the centre of the obstacle. We collect 8 kinesthetic demonstrations with different initial configurations of the object and the obstacle successively displaced upwards as marked with the visual tags in the figure. Alternate demonstrations $\{1, 3, 5, 7\}$ are used for the training set, while the rest are used for the test set. Each observation comprises of the end-effector Cartesian position, quaternion orientation, linear velocity, and quaternion derivative with $D = 14, P = 2$, and a total of 200 datapoints per demonstration.

During evaluation of the learned task-parameterized HSMM in latent space and the Bayesian non-parametric online learning case, we observe a similar performance of the algorithms compared to the batch case with much less parameters and streaming demonstrations. The robot arm is able to generalize effectively by following a similar pattern to the recorded demonstrations in picking and placing the object (see Fig. 4 for reproductions on training and testing set). The model exploits variability in the observed demonstrations to statistically encode different phases of the task such as reach, grasp, move, place, return. Here, reaching the object and avoiding the obstacle have higher variability in the demonstrations, whereas aligning with the frames for grasping and placing the object have little observed variations in their respective coordinate systems. The imposed structure with task-parameters and HSMM allows us to acquire a new task in a few human demonstrations.

References

[1] Brenna D. Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. A survey of robot learning from demonstration. *Robot. Auton. Syst.*, 57(5):469–483, May 2009. ISSN 0921-8890.

[2] Tamara Broderick, Brian Kulis, and Michael I. Jordan. Mad-bayes: Map-based asymptotic derivations from bayes. In *Proceedings of the 30th International Conference on Machine Learning, ICML 2013, Atlanta, GA, USA, 16-21 June 2013*, pages 226–234, 2013.

[3] S. Calinon. A tutorial on task-parameterized movement learning and retrieval. *Intelligent Service Robotics*, 9(1): 1–29, 2016. doi: 10.1007/s11370-015-0187-9.

[4] Y Ephraim and N Merhav. Hidden Markov processes. *IEEE Transactions on Information Theory*, 48(6):1518–1569, 2002. doi: 10.1109/tit.2002.1003838.

[5] Mark J. F. Gales. Semi-tied covariance matrices for hidden markov models. *IEEE Transactions on Speech and Audio Processing*, 7(3):272–281, 1999.

[6] Zoubin Ghahramani. Hidden markov models. chapter An Introduction to Hidden Markov Models and Bayesian Networks, pages 9–42. World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2002. ISBN 981-02-4564-5. URL http://dl.acm.org/citation.cfm?id=505741.505743.

[7] Sanjay Krishnan, Animesh Garg, Sachin Patil, Colin Lea, Gregory Hager, Pieter Abbeel, and Ken Goldberg. *Transition State Clustering: Unsupervised Surgical Trajectory Segmentation for Robot Learning*, pages 91–110. Springer International Publishing, Cham, 2018. doi: 10.1007/978-3-319-60916-4_6.

[8] Brian Kulis and Michael I. Jordan. Revisiting k-means: New algorithms via bayesian nonparametrics. In *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, pages 513–520, New York, NY, USA, 2012. ACM.

[9] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE*, 77:2:257–285, 1989.

[10] Anirban Roychowdhury, Ke Jiang, and Brian Kulis. Small-variance asymptotics for hidden markov models. In *Advances in Neural Information Processing Systems 26*, pages 2103–2111. Curran Associates, Inc., 2013.

[11] S. Schaal, A. Ijspeert, and A. Billard. Computational approaches to motor learning by imitation. *Philosophical Transaction of the Royal Society of London: Series B, Biological Sciences*, 358(1431):537–547, 2003.

[12] A. K. Tanwani. *Generative Models for Learning Robot Manipulation Skills from Humans*. PhD thesis, Ecole Polytechnique Federale de Lausanne, Switzerland, 2018.

[13] A. K. Tanwani and S. Calinon. Learning robot manipulation tasks with task-parameterized semitied hidden semi-markov model. *Robotics and Automation Letters, IEEE*, 1(1):235–242, 2016. doi: 10.1109/LRA.2016.2517825.

[14] A. K. Tanwani and S. Calinon. Small Variance Asymptotics for Non-Parametric Online Robot Learning. *CoRR*, abs/1610.0, 2016. URL http://arxiv.org/abs/1610.02468. Int. J. of Robotics and Research (condtionally accepted).

[15] A. D. Wilson and A. F. Bobick. Parametric hidden Markov models for gesture recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(9):884–900, 1999.

[16] S.-Z. Yu. Hidden semi-Markov models. *Artificial Intelligence*, 174:215–243, 2010.