Honda Research Institute EU
Honda Research Institute Europe GmbH
https://www.honda-ri.de/

Online Learning of Feed-Forward Models for Task-Space Variable Impedance Control

Michael J. Mathew, Saif Sidhik, Mohan Sridharan, Morteza Azad, Akinobu Hayashi, Jeremy Wyatt

2019

Preprint:

This is an accepted article published in IEEE-RAS International Conference on Humanoid Robots. The final authenticated version is available online at: https://doi.org/10.1109/Humanoids43949.2019.9035063 Copyright 2019 IEEE

CONFIDENTIAL. Limited circulation. For review only.

Online Learning of Feed-Forward Models for Task-Space Variable Impedance Control

Michael J Mathew*[1], Saif Sidhik*[1], Mohan Sridharan[1], Morteza Azad[1], Akinobu Hayashi[2], Jeremy Wyatt[1]

***Abstract*— During the initial trials of a novel manipulation task, humans tend to keep their arms considerably stiff in order to reduce the effects of any unforeseen disturbances on the ability to perform the task accurately. After a few repetitions, humans reduce and adapt the stiffness of their arms without any significant reduction in task performance. Research in human motor control strongly indicates that humans learn and continuously revise internal models of manipulation tasks to support such adaptive behaviour. These internal models help predict future states of the task, anticipate necessary control actions, and adapt impedance quickly to match task requirements. Drawing inspiration from these findings, we propose a novel framework that supports the online learning of a time-independent forward model of a manipulation task from a small number of examples. The measured inaccuracies in the predictions of this model are used to dynamically update the forward model and modify the impedance parameters of a feedback controller during task execution. Furthermore, our framework includes a hybrid force-motion controller that enables the robot to be compliant in particular directions (if required) while adapting the impedance in other directions. These capabilities are illustrated and evaluated on continuous contact tasks such as polishing a board and stirring porridge.**

I. Motivation

Robot manipulation in dynamically changing environments is a challenging open problem. Consider a robot arm stirring porridge. As the robot performs this task, the resistance offered by the porridge varies as a function of the viscosity of the fluid. In other tasks such as polishing a table, the robot has to adapt its stiffness to follow a desired trajectory on the surface of the table in the presence of frictional forces. The robot also has to maintain a sufficient level of force along the surface normal while offering some compliance, e.g., when the table is tilted or moved up.

Existing work on robot manipulation is unable to achieve the desired behavior in such tasks. These methods learn or compute stiffness values based on large labeled training datasets or comprehensive knowledge of the domain dynamics, impose unrealistic assumptions or hardware requirements, or use a state representation that makes its computationally expensive to estimate the stiffness parameters. Our framework makes a significant departure from existing literature by drawing inspiration from research in human (and animal) motor control. This research indicates that when performing a new manipulation task, humans tend to use higher arm stiffness in order to follow the desired trajectory accurately in the presence of unforeseen external disturbances. With sufficient experience, humans tend to adapt their stiffness and perform the task accurately with much lower stiffness. This behavior is achieved by building internal models of the task dynamics to predict the configurations of the object and the hand (i.e., *task state*), and the forces, during task execution [1], [2], [3], [4]. Studies in psycho-physics also indicate that learning to vary stiffness is crucial in human manipulation [5], [6]. The proposed framework makes the following key contributions:

- A forward model of any given manipulation task is learned from a small number of examples and revised incrementally in real-time during task execution.
- Impedance parameters are defined as a state-dependent, time-independent property in task space, with the measured error in the state predicted by the forward models being used for the online adaptation of these parameters to the task at hand.
- A hybrid force-motion controller is introduced contextually separate the directions in which the arm has to be stiff or compliant based on the task(s) at hand.

We illustrate and evaluate these capabilities in the context of three challenging continuous contact tasks: (i) pulling a combination of springs of different stiffness to a particular height and moving it in a desired motion pattern; (ii) polishing boards with different surface friction based on a desired motion pattern on the surface; and (iii) stirring porridge based on a desired motion pattern in the presence of changing viscosity. We show that the proposed formulation supports learning and generalization of the forward models from limited training examples, and rapid adaptation of impedance parameters for different related tasks and environments.

The remainder of the paper is organized as follows. We discuss related work in Section II and describe our framework in Section III. Experimental results are discussed in Section IV, and the conclusions are in Section V.

II. Related Work

We motivate the need for the proposed framework by reviewing related work in robot manipulation. Some methods that are more closely related to our framework are also discussed at the appropriate point later in the paper.

Many existing methods for robot manipulation use machine learning algorithms to compute suitable values of the impedance parameters for the tasks at hand [7], [8], [9], [10], [11], [12], [13]. These methods typically either represent the desired stiffness profiles as a time series or as a task-specific

* Authors contributed equally to this work
[1] University of Birmingham, Birmingham, B15 2TT, UK, [mjm522, sxs1412, m.sridharan, m.azad, j.l.wyatt]@bham.ac.uk
[2]Honda Research Institute Europe GmbH, Offenbach/Main, Germany, akinobu.hayashi@honda-ri.de

Manuscript 138 submitted to 2019 IEEE-RAS International Conference on Humanoid Robots (Humanoids). Received July 7, 2019.

policy, resulting in the need for large labeled training datasets or comprehensive knowledge of the robot's dynamics and other mathematical models; it is often difficult to provide such training datasets or accurate domain knowledge. Research in classical control has also developed many methods for adapting stiffness to achieve precise movement, e.g., hybrid force control [14], parallel force control [15], and impedance control [16]. These approaches require accurate knowledge of the system's dynamics and precise feedback schemes, which is difficult to provide in practical domains. Other methods have explored varying the stiffness from the perspective of the object being manipulated [17], [18], [19]. These methods have predominantly been designed for grasping tasks and require accurate analytic models of the object; it is challenging to provide such analytic models in dynamic domains. Many methods also make unrealistic assumptions such as quasi-static action, zero slippage, and point contacts, and require explicit representation of intrinsic parameters such as friction, mass, and coefficient of restitution, which are non-trivial to estimate [20].

A robot can vary its stiffness using a *variable impedance controller*, e.g., by changing the impedance parameters to match the desired motion profile [16]. It is possible to achieve accurate motion and better rejection of perturbing forces by using a higher impedance (i.e. stiffness), but being very stiff expends more energy and makes it difficult to be *compliant* to external forces. Existing variable impedance control methods are time-dependent and/or include joint space parameters as a part of the state description [7], [21]. Such dependence results in the task model and task execution being out of sync in the presence of unforeseen (force) perturbances, limiting the ability to adapt the impedance as needed.

Forward models have been used widely in the literature to perform manipulation tasks. They are generally used to predict the behaviour of the robot [22], [23] and/or the objects being manipulated [24]. The main challenge in the development of such forward models is the selection of relevant state features of the task to successfully learn a policy that can predict the forces from the current state.

We will discuss some of these methods again after describing our framework and the experimental results.

III. Approach

Figure 1 provides an overview of our framework. The human designer either provides a desired motion pattern (i.e., profile) or moves the robot arm along the desired motion pattern once for any given task. The robot learns a forward model of the task from a small number of trials while attempting to achieve the target motion profile. The learned model is used to predict the force experienced in the next state. This prediction determines a feed-forward term in the control command. The prediction error (measured during task execution) revises the forward model and the gain (i.e., impedance) parameters of a feedback (PD) controller that provides the feed-forward term in the control command. A hybrid force-motion controller separately controls force along the direction(s) in which compliance is desired. We describe individual components of the framework below.

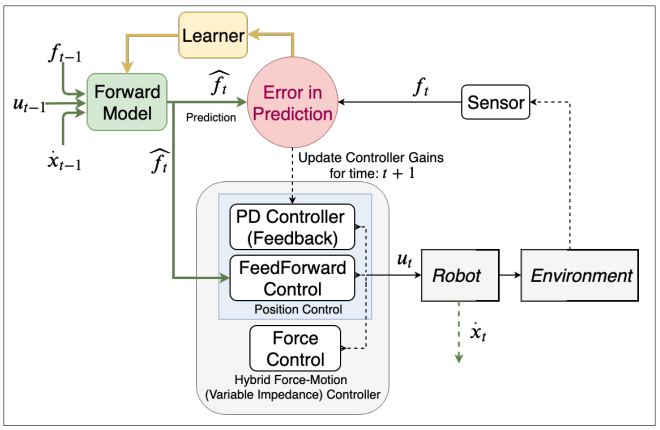


Fig. 1: Block diagram of proposed framework

A. Basic Formulation

We formulate impedance control in the task space of the robot. Task (i.e., Cartesian) space controllers have task-specific parameters and their use is independent of the type of robot manipulator. They also provide an elegant abstraction of multiple equivalent joint space trajectories into a corresponding task-space motion profile. Task-space controllers are typically designed as a *mass-spring-damper* system with the objective of making the robot behave like a spring attached between the end-effector tip and the motion *way-point*. Shaping the *inertia* of a robot to behave like a mass-spring-damper system is tricky and requires accurate measurement of the external forces acting on the robot [25]. Since measuring forces at every acting point is impractical, it is difficult to modulate the inertia tensor without leading to incorrect impedance behaviour. In practice, the desired impedance behaviour is limited to designing *stiffness* and *damping* parameters of the controller while keeping the natural inertia unchanged, resulting in a compliance control problem [25]. Since arbitrarily varying stiffness and damping parameters may result in unstable robot behaviour [26], we use empirically estimated bounds for stiffness and damping parameters in all our experiments.

In the basic formulation, the forward model maps the control command and measured force to a predicted force to be experienced in the resultant state. The force vector typically includes frictional forces along the plane of motion and the force along the surface normal. Later, we describe our hybrid force-motion controller which has a PD controller with fixed gains to control the force along the direction in which compliance is desired (surface normal in current implementation). Our basic formulation thus has a motion controller with a feed-forward term and a feedback term, which is known to reduce the number of training samples for variable impedance control [27]. The corresponding controller equation is:

$$u_t = \mathbf{K_t^p} \Delta x_t + \mathbf{K_t^d} \Delta \dot{x}_t + k_t \tag{1}$$

Manuscript 138 submitted to 2019 IEEE-RAS International Conference on Humanoid Robots (Humanoids). Received July 7, 2019.

where u_t is the control command to the robot (i.e., desired task space force) at time t, $\mathbf{K_t^p}$ and $\mathbf{K_t^d}$ are the (positive definite) stiffness and damping matrices of the feedback controller; k_t is the feed-forward term provided by the forward model; and Δx and $\Delta \dot{x}$ are the errors in the end-effector position and velocity. During task execution, the forward model's predictions determine k_t, and the prediction error controls $\mathbf{K_t^p}$ and $\mathbf{K_t^d}$ (Section III-C).

B. Learning and Using the Forward Model

For ease of explanation, assume that the robot has been provided the target motion profile, i.e., a sequence of points in task space, for a given task. The forward model is learned over a few trials as the robot attempts to follow the desired profile. To create a time-independent forward model, our framework uses a Gaussian Mixture Model (GMM) that is fit over points of the form $p = [S_{t-1}, f_t]$, where S_t can be any combination of features that uniquely represent the robot's state for the task, and f_t is the force felt at the end-effector at time t. S_t can contain information about end-effector position (x_t), velocity ($\dot{x}_t$), forces (f_t), etc. We explored two representations for S_t. The first is of the form $S_t = [\dot{x}_t, f_t]$. The second, motivated by studies of motor control [4], is of the form $S_t = [\dot{x}_t, f_t, u_t]$ where, u_t is the task space control command. This is similar to the "efferent copy" mechanism in animal motor control, where a copy of movement-producing signals are used by internal forward models to predict the effects of actions.

To incrementally, continuously, and quickly update the GMM's parameters (and create new components when needed) during task execution, we used a variant of GMM called Incremental GMM (IGMM) [28], [29], [30]. IGMM internally uses a variant of the Expectation-Maximisation (EM) algorithm to fit the model and maximize the following likelihood function:

$$L(\theta) = p(\mathbf{X}|\theta) = \prod_{n=1}^{T} p(X_n|\theta) = \prod_{n=1}^{T} \left[\sum_{j=1}^{M} p(X_n|j)p(j) \right] \tag{2}$$

where $\theta = (\mu_j, \sigma_j, p_j)$ for $j = 1...M$ are the parameters of the M components of the GMM. $\mathbf{X} = (X_1, ..., X_T)$ represents the points to be fit, with $X_t = [S_{t-1}, f_t]$. Each point contains information about the *previous* end-effector state, along with the *current* force. So, when the learned model is used for prediction during task execution, the force for the next time instant ($(f_{t+1}|S_t)$) is predicted as a function of the robot's current state using Gaussian Mixture Regression (GMR) [31].

C. Varying Feedback Gains

Many manipulation tasks can be completed successfully using a very high stiffness ($\mathbf{K}^p_{max}$), but this would make the robot less compliant and expend more energy. Also, if the robot had to perform a task (e.g., pulling the spring action) in free-space, accurate trajectory following can be achieved with a much lower stiffness ($\mathbf{K}^p_{free}$).

If the learned forward model is accurate, the feed-forward term should cancel out the external forces, reducing motion to that in free-space. Similar to human behaviour with a familiar manipulation task, the corresponding feedback gains can be closer to $\mathbf{K}^p_{free}$. The feedback gains at each step (K_t) of the controller (Equation 1) are given by:

$$\mathbf{K}^p_t = \mathbf{K}^p_{free} + F(e_{pred,t-1})(\mathbf{K}^p_{max} - \mathbf{K}^p_{free}) \tag{3}$$

where $e_{pred,t}$ is the forward model's prediction error at time t, and $F(x) : x \rightarrow [0, 1]$. With this formulation, the robot will be more compliant (stiff) when the model's predictions are more accurate (inaccurate). The damping term is updated using the known constraint of the damping factor for an over-damped system:

$$\mathbf{K}^d_t = \sqrt{\frac{\mathbf{K}^p_t}{4}} \tag{4}$$

D. Hybrid Force-Motion Control for Compliance

Consider a robot polishing a planar surface has to maintain contact force along the surface normal while following the desired motion profile on the surface. If the robot experiences unexpected (e.g., frictional) forces along the plane, the robot has to become stiffer to follow the motion profile accurately (and revise the forward model). However, if the surface is raised suddenly, the formulation described above would predict incorrect forces and increase stiffness, with the robot trying to push the surface down, damaging itself and/or the surface. Thus, some manipulation tasks require compliance along some directions when an unexpected force is experienced. We use a hybrid force-motion controller to intuitively separate the 'compliant' and 'stiff' directions. Such controllers define *artificial* constraints on the robot's degrees of freedom. These constraints that specify (through a control law) the desired values for the velocities in the k directions of motion, and for the forces in the remaining $6-k$ directions for contact reaction. Through direct force control along the directions in which compliance is desired, and motion control for the other directions, the robot can maintain the required normal force while following the trajectory on the surface. These directions are currently defined manually for each task and considered contextual information about the task. The revised controller equation is:

$$u_t = \mathbf{K_t^p}\Delta x_t + \mathbf{K_t^d}\Delta \dot{x}_t + k_t + u_{fc} \tag{5}$$

where u_{fc} specifies part of the command signal produced by the direct force control.

IV. Experimental Setup and Results

We experimentally evaluated the following hypotheses about the capabilities of our framework:

- **H1:** Using feed-forward model along with stiffness adaption improves trajectory tracking performance.
- **H2:** Adding efferent copy as the input to the forward model creates a better model of task dynamics, resulting in improved trajectory tracking.
- **H3:** Online learning of forward model supports adaptation to new and changing environments.

Manuscript 138 submitted to 2019 IEEE-RAS International Conference on Humanoid Robots (Humanoids). Received July 7, 2019.

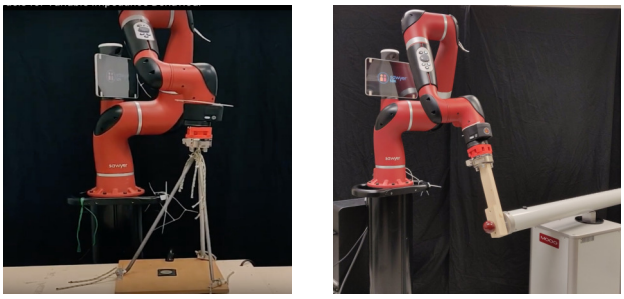


Fig. 2: **Left:** (Non-linear spring pulling) the objective is to pull a combination of springs to a desired height and then along a desired motion pattern; **Right:** (Porridge stirring) the Sawyer is is attached to a Moog HapticMaster which emulates an environment whose viscosity increases as it moves (in X-Z plane).

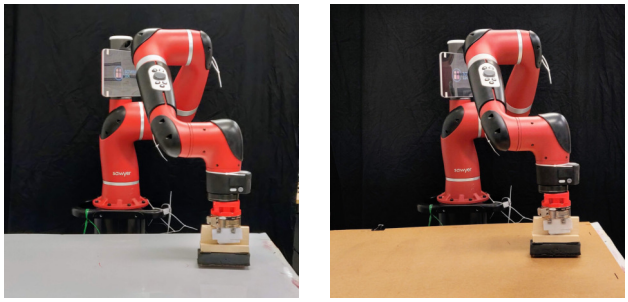


Fig. 3: (Board-polishing) **Left:** Surface 1 with low friction; **Right:** Surface 2 with higher friction.

H1 tests the effectiveness of using the forward model and subsequently adapting stiffness based on the accuracy of the model; **H2** compares the choices of the feature vector (Section III-A); and **H3** assesses if the framework can adapt to new environments. The *root mean square* (RMS) measure is used to quantify the error in achieving the desired motion profile, and suitable plots provide a qualitative indication of performance. Since our approach is significantly different from popular approaches for such manipulation tasks (e.g., based on deep learning), we do not provide an experimental comparison; the obvious advantages of our approach are discussed in Section II and in Section V.

We used three tasks to evaluate the hypotheses, as stated in Section I; we henceforth refer to them as "nonlinear spring pulling", "board polishing" and "porridge stirring". We used a 7-DoF Sawyer robot for our experiments; a high-quality video can be viewed online[1]. The forward model is learned with the feature vector $p = [S_{t-1}, f_t]$, and $S_t = [\dot{x}_t, f_t]$ except for testing **H2** (see Section III-B). The forward model learns the probability distribution of the feature vectors; GMR conditions on S_t to predict f_{t+1}.

The first task involved pulling springs (attached to end effector) to a particular height and then moving along a desired trajectory. This experiment was conducted first with a linear spring (as proof-of-concept) and then with a non-linear spring (see Figure 2), which is significantly more challenging due to its non-linear force response to extension. When the

[1]https://youtu.be/hbzZuO1xal8

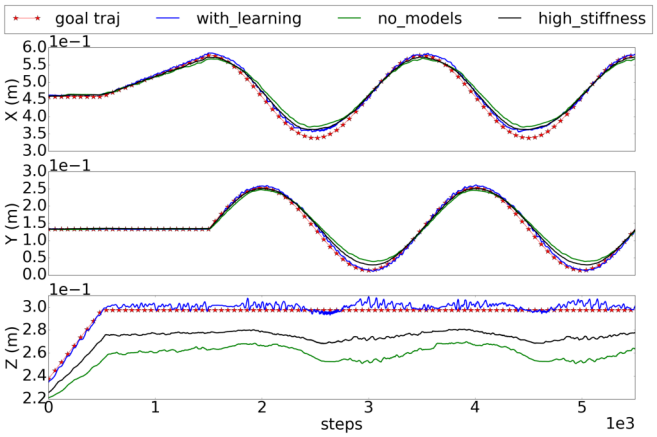


Fig. 4: **(Non-linear spring pulling)** Position tracking.

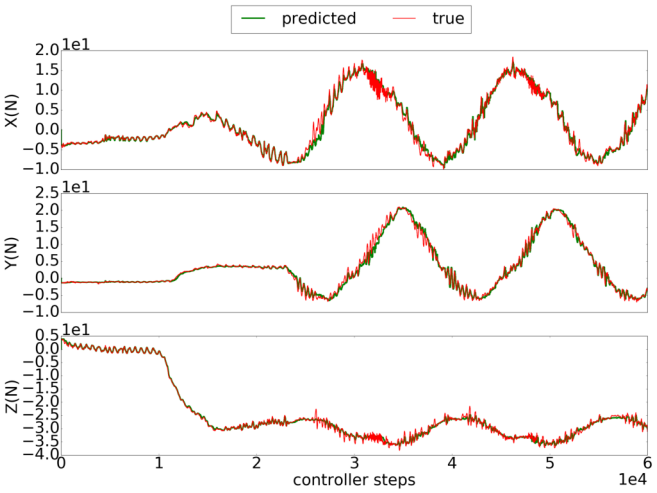


Fig. 5: **(Non-linear spring pulling)** Force prediction.

end-effector moves, it now experiences different changes in force in different directions. The baselines for comparison used constant low impedance and constant high impedance. The low impedance parameters were chosen to be enough to move the end-effector along the desired trajectory in the absence of springs ($\mathbf{K}^p_{free}$), and the high impedance parameters ($\mathbf{K}^p_{max}$) were chosen to be sufficient for pulling the spring in the absence of the forward model.

We conducted multiple trials with the forward model being learned in the first trial and then improved in subsequent trials. The results are shown in Figures 4– 5. Figure 5 shows that the prediction accuracy of the forward model improves significantly over the trials. Figure 4 shows that the accuracy of position tracking and Table I quantifies the performance. Accuracy is better with the forward-model than when it is not used. Performance improves further when the impedance parameters are updated online; just using the feed-forward term is not enough to perform tasks that involve unexpected forces acting on the system (unless the model is perfect). The proposed framework improves the model online and adapts impedance parameters to the accuracy of the model.

To further evaluate **H1**, we explored the board polishing task; the robot had to wipe a surface of unknown friction coefficient by moving its end-effector (a whiteboard eraser)

Manuscript 138 submitted to 2019 IEEE-RAS International Conference on Humanoid Robots (Humanoids). Received July 7, 2019.

	X	Y	Z
no models	0.017	0.015	0.038
high stiffness	0.012	0.00929	0.023
with learning	0.010	0.0064	0.0040

TABLE I: **(Non-linear spring pulling)** Trajectory tracking along the three axes.

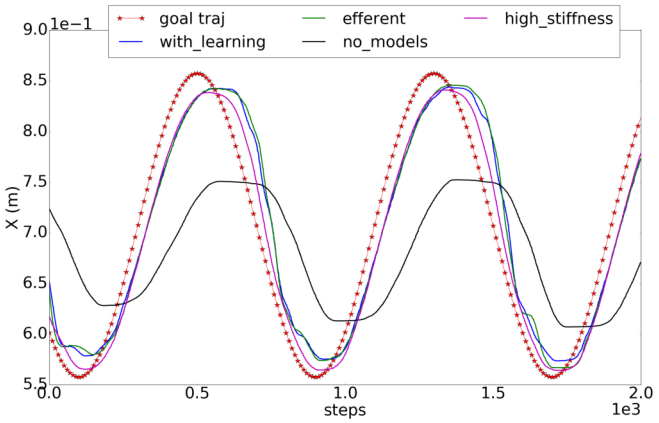


Fig. 6: **(Board polishing)** Surface 1 position tracking. Red: target; Pink: constant stiffness $\mathbf{K}^p_{free}$; Black: constant stiffness $\mathbf{K}^p_{max}$; Blue: adapting impedance without efferent copy; Green: adapting impedance with efferent copy.

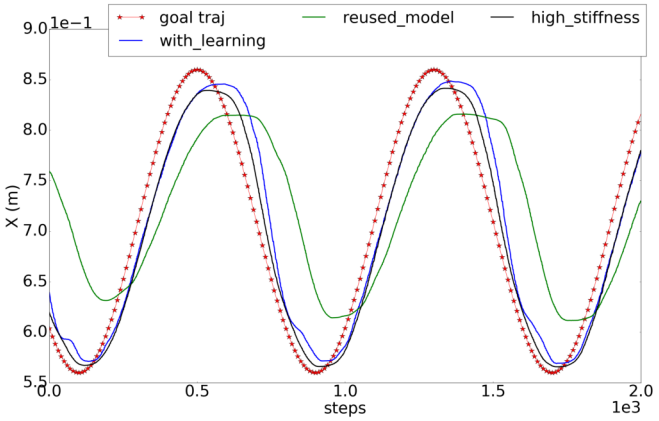
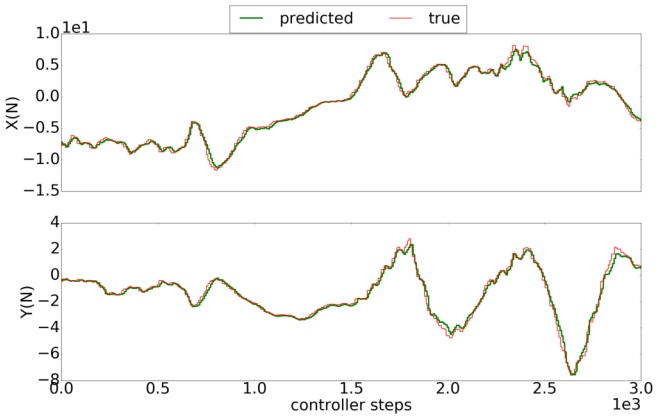


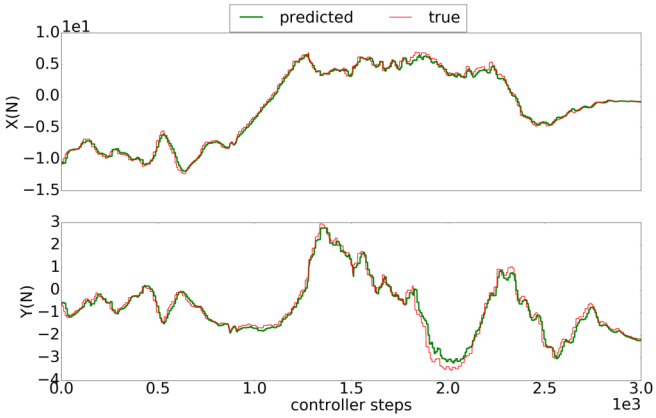
Fig. 7: **(Board polishing)** Surface 2 position tracking. Red: target; pink: constant stiffness $\mathbf{K}^p_{max}$; Green: surface 1 model without adaptation; Blue: online adaptation of previous model.

on the surface along a given trajectory while applying $10N$ downward force. Here, a hybrid force-motion controller offers compliance along the surface normal (the z-axis). While moving, the robot experiences frictional forces hampering the smooth motion of the robot, which it has to learn to predict. The initial model is learned by making the robot wipe the surface following a trajectory (an epicycle) that is considerably different from the one it has to follow during task execution (sine wave).

Figure 6 show that in the absence of the forward model, the robot is unable to follow the desired trajectory since it does not know the interaction forces. The use of the feed-forward model improves tracking performance, with a further improvement when online impedance adaptation is used. The performance of the framework is comparable with that of a high stiffness controller while requiring much lower impedance parameters (Figure 9a). The average RMS error of trajectory are summarized in Table II. These results support the validity of **H1**.



(a) Surface 1



(b) Surface 2

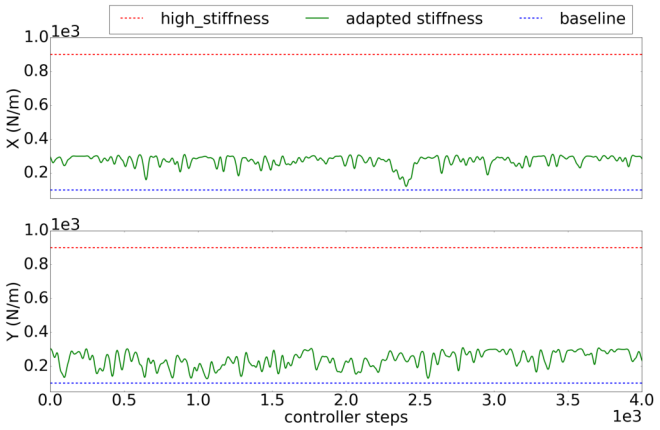
Fig. 8: **(Board polishing)** Force prediction.

	X	Y
with no models	0.091	0.054
using high stiffness	0.027	0.007
with efferent	0.038	0.008
with learning	0.036	0.008

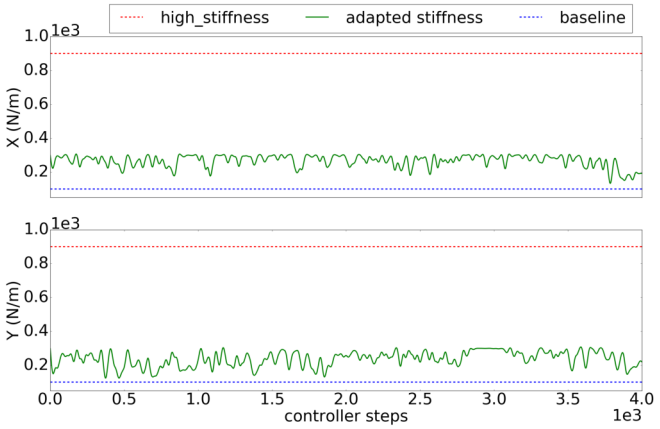
TABLE II: **(Board Polishing)** Surface 1: trajectory tracking errors along different axes.

Next, to evaluate **H2**, we conducted trials of board polishing task with the efferent copy in the feature vector of the forward model. Results indicate that there is no significant improvement in performance. We believe this is because the forward model is able to obtain enough information for force prediction from the current end-effector velocity and forces, making the information encoded in the efferent copy redundant. This observed performance, and the fact that adding dimensions to the state-space makes the learning

Manuscript 138 submitted to 2019 IEEE-RAS International Conference on Humanoid Robots (Humanoids). Received July 7, 2019.



(a) Surface 1



(b) Surface 2

Fig. 9: **(Board polishing)** Stiffness adaptation

	X	Y
high stiffness	0.027	0.007
with reused model	0.084	0.023
improving reused model	0.035	0.006

TABLE III: **(Board Polishing)** Surface 2: trajectory tracking errors along different axes.

more computationally demanding, led us to *not* use the efferent copy in the framework for subsequent experimentation.

To evaluate **H3**, we focused on the ability to generalize and adapt. The results with the board polishing task indicates that the framework generalizes across different trajectories since the model was learned using a trajectory different from that used during task execution. This is a key advantage of learning the forward models in task-space instead of joint-space. Next, adaptability to new forces was tested by performing the same (board polishing) task using a surface with significant different friction. Figures 7, 8b, and 9b show that when the learned model was used for this surface without online improvement, the robot was not able to follow the trajectory accurately. However, if the learned model is revised during task execution, it quickly achieves performance similar to that with the first surface. The RMS errors in trajectory tracking are summarised in Table III. This

	X	Y
no models	0.031	0.035
high stiffness	0.012	0.014
just forward model	0.037	0.026
forward model with stiffness adaption	0.014	0.009

TABLE IV: **(Porridge stirring)** Trajectory tracking errors

capability of the framework to generalise to different surfaces and trajectories is the critical advantage of using a *task-space, time-independent* variable impedance control framework.

Next, the adaptability of the framework to environments that change during task performance is tested with the porridge stirring task. The viscosity of porridge changes as it is stirred. For the experiment on a real robot, the changes in viscosity are emulated on a MOOG HapticMaster [32]. Specifically, the viscosity (damping factor) of the environment is increased continuously (in X-Z plane) until it reaches a maximum predefined value. The end-effector of the Sawyer is attached to the end-effector of the HapticMaster—see Figure 2. The Sawyer has to move its end-effector along a predefined motion trajectory while learning to adapt to the increasing viscous resistance from the environment.

Figure 10 indicates that using constant impedance parameters is not enough for tracking the desired trajectory (shown in red) accurately, even with the maximum allowed stiffness $\mathbf{K}^p_{max}$ (black). Similarly, using the feed-forward term predicted by the forward model alone without impedance adaptation is insufficient for such dynamically changing environments (shown in green), owing to the dynamics of the environment that are unknown to the robot. However, when coupled with the online impedance adaptation, the robot is able to follow the trajectory more accurately (shown in pink). The trajectory tracking errors during task execution (see Table IV) further establishes the effectiveness of the framework. Next, Figure 11 indicates that the force predictions (by forward model) are much more accurate towards the end of the task when the environment finally becomes static as the change in damping saturates. The model has (by then) learned to accurately predict the effects of that damping factor on force. As a result, the impedance parameters take smaller values, as shown in Figure 12. Note that the stiffness values are still lower than the high stiffness value ($\mathbf{K}^p_{max}$).

We also ran trials that examined the ability of the hybrid force-motion controller to provide compliance along the surface normal with the board polishing task. Specifically, we ran trials in which the board was moved up and/or tilted during task execution. We (qualitatively) observed that the robot was able to provide compliance along the surface normal in each such trial. We do not include any quantitative results of these experiments.

V. Discussion and Conclusions

Variable impedance control is vital for reliable and safe robot manipulation. Learning impedance parameters directly is difficult and would require large training data. In this paper, we proposed a framework inspired by findings in

Manuscript 138 submitted to 2019 IEEE-RAS International Conference on Humanoid Robots (Humanoids). Received July 7, 2019.

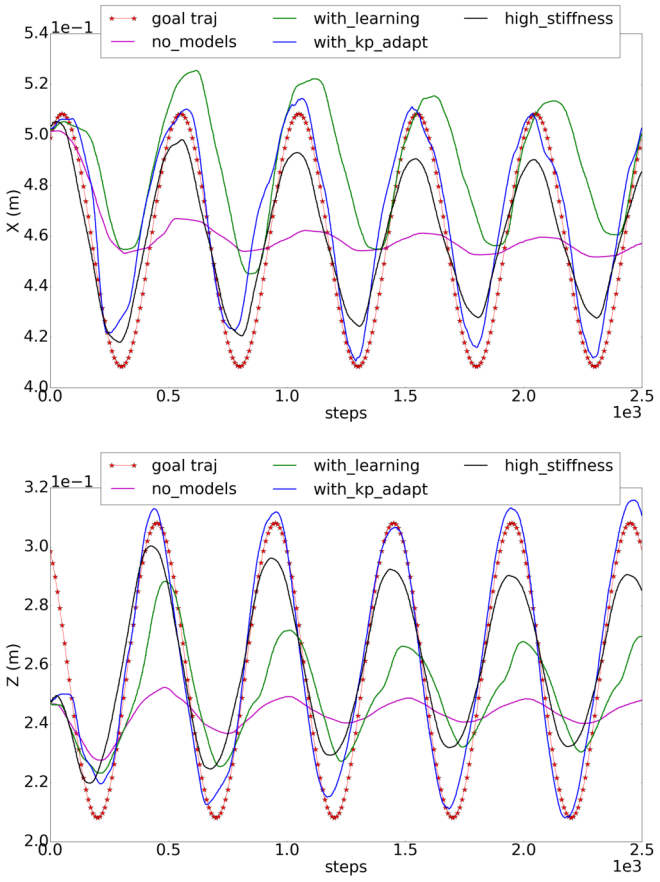


Fig. 10: **(Porridge stirring)** Position tracking (X-Z axes)

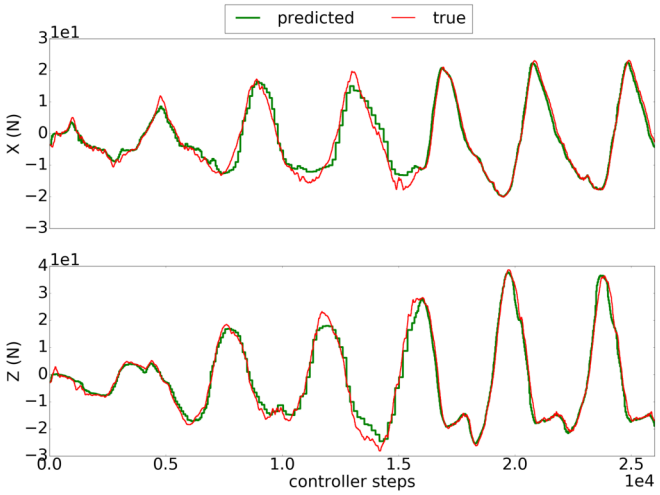


Fig. 11: **(Porridge stirring)** Force prediction

human motor control. Our framework incrementally and continuously learns a time-independent task-space forward model of any given manipulation task online, using the model to predict interaction forces. The prediction error is used to revise the forward model and the impedance parameters for feedback control.

The framework was tested on three distinct tasks: non-linear spring pulling, board polishing, and porridge stirring. Other studies have used models created using knowledge of domain mechanics to make predictions about robot and

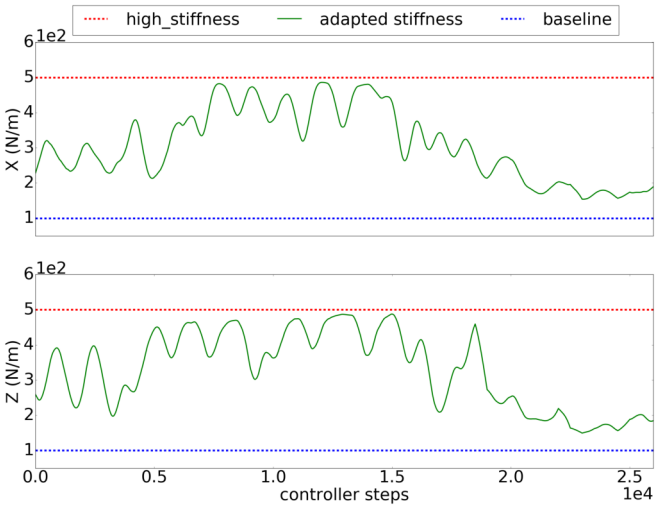


Fig. 12: **(Porridge stirring)** Stiffness adaptation

object motions [24], [33]. However, they make unrealistic assumptions such as point contacts, friction cone approximations and no slippage unlike our method. We started with a linear spring and the corresponding experiments provided insights on why the forward model (by itself) will not suffice, especially in the presence of previously unseen forces. We then introduced the task-space variable impedance feedback controller. One of the techniques proposed in [34] learns a state-dependent stiffness but the learning relies heavily on multiple demonstrations through an expensive special-purpose hardware device.

A key challenge (in our framework) was to choose suitable feature vectors to representing task state. We explored two distinct representations as described in Section III. Experimental results indicated that the use of the efferent copy as input to the forward model does not produce a marked improvement (over the other representation) but this hypothesis deserves further exploration in the context of other tasks. We did notice that the use of the end effector velocity instead of position as a feature in the state vector helped generalize across different trajectories and surfaces. Existing approaches such as [7] and [21] would fail to provide such generalization due to the explicit dependence on time for representing force and impedance parameters. Experimental results thus also demonstrate the advantages of learning forward models in the task space in comparison with existing work that learns forward models in the joint space [22], [8], [9]. In addition, a crucial difference in comparison with other approaches for variable impedance control [18], [19], [17] is that we do not use analytic models that require accurate knowledge of the system dynamics.

The porridge stirring task was the most challenging of the tasks we explored in this paper. It is unlike many other tasks considered by existing work in that it involved dynamic changes during task execution. Also, the factor that changes, i.e., the viscosity of the environment, opposes motion, the forces that the robot experiences are a function of the velocity rather than position. At the same time, this task provides

Manuscript 138 submitted to 2019 IEEE-RAS International Conference on Humanoid Robots (Humanoids). Received July 7, 2019.

a good controlled environment to emulate a dynamically changing environment. The successful stiffness adaptation with improved performance in this task thus provides strong evidence about the capability of our framework. Some existing approaches [22], [23] attempt to learn an action-effect correlation, usually from demonstrations provided by an expert or from experience obtained during trials [35], [36]. However, these methods requires explicit mathematical representations of the task, robot, and/or the objects involved.

Overall, the experimental evaluation provided promising results strongly indicating the ability of the proposed framework to learn forward models and achieve task-space variable impedance control in diffrent continuous contact tasks. Future work will further examine the capabilities of this framework on other robots performing different tasks, and explore the extension of this framework to tasks that require the manipulator to make and break contacts.

REFERENCES

[1] J. R. Flanagan, M. C. Bowman, and R. S. Johansson, "Control strategies in object manipulation tasks," *Current opinion in neurobiology*.

[2] M. Kawato, "Internal models for motor control and trajectory planning," *Current opinion in neurobiology*.

[3] R. S. Johansson and K. J. Cole, "Sensory-motor coordination during grasping and manipulative actions," *Current opinion in neurobiology*, vol. 2, no. 6, pp. 815–823, 1992.

[4] R. Shadmehr and J. W. Krakauer, "A computational neuroanatomy for motor control," *Experimental brain research*, vol. 185, no. 3, pp. 359–381, 2008.

[5] D. W. Franklin, E. Burdet, K. P. Tee, R. Osu, C.-M. Chew, T. E. Milner, and M. Kawato, "Cns learns stable, accurate, and efficient movements using a simple algorithm," *Journal of neuroscience*, vol. 28, no. 44, pp. 11 165–11 173, 2008.

[6] E. Burdet, R. Osu, D. W. Franklin, T. E. Milner, and M. Kawato, "The central nervous system stabilizes unstable dynamics by learning optimal impedance," *Nature*, vol. 414, no. 6862, p. 446, 2001.

[7] J. Buchli, F. Stulp, E. Theodorou, and S. Schaal, "Learning variable impedance control," *IJRR-2011*, vol. 30, no. 7, pp. 820–833, 2011.

[8] M. Kalakrishnan, L. Righetti, P. Pastor, and S. Schaal, "Learning force control policies for compliant manipulation," in *IROS-2011*. IEEE, 2011, pp. 4639–4644.

[9] M. Kalakrishnan, J. Buchli, P. Pastor, M. Mistry, and S. Schaal, "Learning, planning, and control for quadruped locomotion over challenging terrain," *IJRR-2011*, vol. 30, no. 2, pp. 236–258, 2011.

[10] D. J. Braun, F. Petit, F. Huber, S. Haddadin, P. Van Der Smagt, A. Albu-Schäffer, and S. Vijayakumar, "Robots driven by compliant actuators: Optimal control under actuation constraints," *IEEE Transactions on Robotics*, vol. 29, no. 5, pp. 1085–1101, 2013.

[11] M. Howard, D. J. Braun, and S. Vijayakumar, "Transferring human impedance behavior to heterogeneous variable impedance actuators," *IEEE Transactions on Robotics*, vol. 29, no. 4, pp. 847–862, 2013.

[12] A. Kupcsik, M. Deisenroth, J. Peters, L. Ai Poh, V. Vadakkepat, and G. Neumann, "Model-based contextual policy search for data-efficient generalization of robot skills," 2017.

[13] H. van Hoof, T. Hermans, G. Neumann, and J. Peters, "Learning robot in-hand manipulation with tactile features," in *IEEE-Humanoids-2015*, 2015.

[14] J. K. Salisbury, "Active stiffness control of a manipulator in cartesian coordinates," in *Decision and Control including the Symposium on Adaptive Processes, 1980 19th IEEE Conference on*.

[15] S. Chiaverini and L. Sciavicco, "The parallel approach to force/position control of robotic manipulators," *IEEE TRA*.

[16] N. Hogan, "Impedance control: An approach to manipulation," in *IEE-ACC-1984*.

[17] M. Li, H. Yin, K. Tahara, and A. Billard, "Learning object-level impedance control for robust grasping and dexterous manipulation," in *ICRA-2014*.

[18] S. A. Schneider and R. H. Cannon, "Object impedance control for cooperative manipulation: Theory and experimental results," *IEEE TRA*, vol. 8, no. 3, pp. 383–394, 1992.

[19] T. Wimböck, C. Ott, A. Albu-Schäffer, and G. Hirzinger, "Comparison of object-level grasp controllers for dynamic dexterous manipulation," *IJRR*, vol. 31, no. 1, pp. 3–23, 2012.

[20] M. Kopicki, S. Zurek, R. Stolkin, T. Moerwald, and J. L. Wyatt, "Learning modular and transferable forward models of the motions of push manipulated objects," *Autonomous Robots*.

[21] M. Deniša, A. Gams, A. Ude, and T. Petrič, "Learning compliant movement primitives through demonstration and statistical generalization," *IEEE/ASME transactions on mechatronics*, vol. 21, no. 5, pp. 2581–2594, 2015.

[22] K. J. A. Kronander, "Control and learning of compliant manipulation skills," 2015.

[23] B. Huang, M. Li, R. L. De Souza, J. J. Bryson, and A. Billard, "A modular approach to learning manipulation strategies from human demonstration," *Autonomous Robots*, vol. 40, no. 5, pp. 903–927, 2016.

[24] Y. Fan, W. Gao, W. Chen, and M. Tomizuka, "Real-time finger gaits planning for dexterous manipulation," *IFAC-2017*, vol. 50, no. 1, pp. 12 765–12 772, 2017.

[25] T. Wimbock, C. Ott, and G. Hirzinger, "Impedance behaviors for two-handed manipulation: Design and experiments," in *ICRA-2007*.

[26] K. J. A. Kronander, "Control and learning of compliant manipulation skills," p. 204, 2015. [Online]. Available: http://infoscience.epfl.ch/record/210787

[27] N. T. Alberto, M. Mistry, and F. Stulp, "Computed torque control with variable gains through Gaussian process regression," in *IEEE-RAS 2014*, 2014.

[28] M. Song and H. Wang, "Highly efficient incremental estimation of Gaussian mixture models for online data stream clustering," in *SPIE Conference Series*, ser. SPIE Conference-2005, K. L. Priddy, Ed., vol. 5803, March 2005, pp. 174–183.

[29] W. Ahmad, "Incremental learning of gaussian mixture models."

[30] P. Engel and M. Heinen, "Incremental learning of multivariate gaussian mixture models," in *BSAI*. Springer, 2010.

[31] H. G. Sung, "Gaussian mixture regression and classification," Ph.D. dissertation, Rice University, 2004.

[32] O. S. Haptics, *Moog Haptic Master Manual*, 2008 (accessed May 9, 2019). [Online]. Available: https://www.h3dapi.org/modules/mediawiki/index.php/MOOG_FCS_HapticMaster

[33] C. K. Liu, "Dextrous manipulation from a grasping pose," in *ACM-TOG-2009*, vol. 28, no. 3. ACM, 2009, p. 59.

[34] K. Kronander and A. Billard, "Learning compliant manipulation through kinesthetic and tactile human-robot interaction," *IEEE transactions on haptics*, vol. 7, no. 3, pp. 367–380, 2014.

[35] S. Levine and V. Koltun, "Guided policy search," in *ICML-2013*, 2013, pp. 1–9.

[36] A. Kupcsik, M. P. Deisenroth, J. Peters, A. P. Loh, P. Vadakkepat, and G. Neumann, "Model-based contextual policy search for data-efficient generalization of robot skills," *Artificial Intelligence*, vol. 247, pp. 415–439, 2017.