# Gaussian Process Inverse Dynamics Learning for Variable Stiffness Actuator Control

Yongping Pan<sup>1</sup>, Zhigang Zou<sup>2</sup>, Weibing Li<sup>2</sup>, Chenguang Yang<sup>3</sup>, and Haoyong Yu<sup>4</sup>

*Abstract*— The control of variable stiffness actuators (VSAs) is challenging because they exhibit highly nonlinear characteristics and are difficult to model accurately. In this study, we propose a machine learning-based tracking control approach combining Gaussian process (GP) learning and low-gain feedback control for VSAs subjected to unknown dynamics, where the GP model learns the inverse dynamics of agonistic-antagonistic (AA)-VSAs to feedforward control and provides the model fidelity by the predicted variance for the online adjustment of feedback control gains. It is shown that the tracking error is uniformly ultimately bounded and exponentially converges to a small ball under a given probability. Experiments on an AA-VSA named qbmove Advanced have validated the superiority of the proposed method with respect to tracking accuracy and generalization.

# I. INTRODUCTION

Variable stiffness actuators (VSAs) are complicated mechatronic devices whose position and stiffness can be controlled by two motors separately or jointly [1]. They are popular in applications to human-robot interaction and scenes that require adjusting natural dynamics. Controlling VSAs is challenging mainly due to their strong nonlinearities and difficulties in accurate modeling [2]. Model-based feedback control approaches for VSAs but have two major drawbacks: 1) They require exact VSA models that are rarely available in practice; 2) they achieve good control accuracy at the cost of stiffening the physical dynamics, which violates the purpose of introducing compliant actuators [3]. Iterative learning control (ILC) can achieve high-accuracy control without exact VSA models and high feedback gains [4], but their task generalization is limited due to the nature of ILC [5].

Machine learning-based control can improve task generalization while maintaining the advantages of ILC to some extent and attracted some attention for VSAs in recent years, e.g., see [6]–[10]. Guo et al. [6] proposed an adaptive neural network (NN) control method based on feedback linearization for a serial VSA. This method considers system uncertainties and variable loads but requires the higher-order time derivatives of system states that are difficult to obtain in practice. Tran et al. [7] and Liu et al. [8] proposed adaptive NN backstepping

\*This work was supported in part by the Fundamental Research Funds for the Central Universities, Sun Yat-sen University, China, under Grant No. 23]gzy004 (*Corresponding author: Yongping Pan*).

<sup>1</sup>Y. Pan is with the School of Advanced Manufacturing, Sun Yat-sen University, Shenzhen 518100, China panyongp@mail.sysu.edu.cn

<sup>2</sup>Z. Zou and W. Li are with the School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China <code>zouzhg3@mail2.sysu.edu.cn</code>; <code>liwb53@mail.sysu.edu.cn</code>

<sup>3</sup>C. Yang is with the Department of Computer Science, University of Liverpool Liverpool, L69 3BX, UK cyang@ieee.org

<sup>4</sup>H. Yu is with the Department of Biomedical Engineering, National University of Singapore, Singapore 117583, Singapore bieyhy@nus.edu.sg

controllers for an electrohydraulic elastic robot with variable stiffness and a pneumatic artificial muscle-driven robot with variable stiffness, respectively. Mitrovic et al. [9] introduced locally weighted projection regression (LWPR) to learn the dynamic model and the stochastic properties of an agonisticantagonistic (AA)-VSA for optimal impedance control. However, this method has controllability problems due to hardware bandwidth and the difficulty of scaling to multiple degrees of freedom (DoFs) due to computational challenges. Knezevic et al. [10] applied NN and LWPR to design two feedforward controllers for an AA-VSA named qbmove.

Gaussian process (GP) regression is a supervised learning technique broadly applied to robotics due to universal approximation ability and generalization under small training data [11]. This paper applies GP to develop a stable tracking controller consisting of GP feedforward and low-gain feedback actions for AA-VSAs subjected to unknown dynamics, where the GP model is applied to learn the inverse dynamics of AA-VSAs for feedforward control, and the feedback control gains are adapted according to the predicted variance so as to enhance performance. It is shown that the tracking error is uniformly ultimately bounded (UUB) and exponentially converges to a ball under a given probability. The proposed GP controller is compared experimentally with a pure feedback controller and an ILC on an AA-VSA named qbmove Advanced [12].

Notations:  $\mathbb{R}$ ,  $\mathbb{R}^+$ ,  $\mathbb{R}^n$ , and  $\mathbb{R}^{m \times n}$  denote the spaces of real numbers, positive real numbers, real *n*-vectors, and real  $m \times n$ matrices, respectively,  $I_n$  is a  $n \times n$  identity matrix,  $\mathcal{N}(\boldsymbol{\mu}, K)$ is a multivariate Gaussian distribution with a mean vector  $\boldsymbol{\mu} \in \mathbb{R}^m$  and a covariance matrix  $K \in \mathbb{R}^{m \times m}$ ,  $\mathcal{GP}(\boldsymbol{\mu}, k(\boldsymbol{x}, \boldsymbol{x}'))$ is a GP with a mean  $\boldsymbol{\mu} \in \mathbb{R}$  and a kernel function  $k : \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ ,  $\|\boldsymbol{x}\|$  is the Euclidean norm of  $\boldsymbol{x} \in \mathbb{R}^n$ ,  $\max\{\cdot\}$  is the maximum operator, and the probability of a probabilistic event  $\Pi$  is written as  $P\{\Pi\}$ , where *n* and *m* are positive integers. Note that *t* denotes the continuous time and an epoch for the continuous-time and discrete-time cases, respectively.

#### **II. PROBLEM FORMULATION**

Consider an AA-VSA mechanism with a link moving in the vertical plane, where its dynamic model is given by [4]

$$M\ddot{q} + D\dot{q} + g(q) = \frac{\partial E(q, \boldsymbol{\theta})^T}{\partial q} + \tau_{\text{un}}, \qquad (1)$$

$$J\ddot{\boldsymbol{\theta}} + B\dot{\boldsymbol{\theta}} - \frac{\partial E(q,\boldsymbol{\theta})^T}{\partial\boldsymbol{\theta}} = \boldsymbol{\tau}_{\rm m}$$
(2)

in which  $q(t) \in \mathbb{R}$  is the joint angular position,  $\theta(t) \in \mathbb{R}^2$  is the motor angular position,  $M \in \mathbb{R}^+$  is the inertia of the shaft,  $J \in \mathbb{R}^{2 \times 2}$  is the inertia of the motors,  $B \in \mathbb{R}^{2 \times 2}$  is the viscous



Fig. 1. A block diagram of the proposed GP controller for AA-VSAs.

friction at the motor side,  $D \in \mathbb{R}^+$  is the viscous friction at the link side,  $g(q) \in \mathbb{R}$  is the gravity term,  $\tau_{un} \in \mathbb{R}$  collects model uncertainties, such as nonlinear friction and coupling dynamics,  $E(q, \theta) \in \mathbb{R}$  is the elastic potential of the system, and  $\tau_m \in \mathbb{R}^2$  is the input torque from the motors.

To simplify the model, we introduce two assumptions [4]. First, the motor dynamics (2) is ignored such that the motor position  $\theta$  can be used as an effective control input. Second, there exists a nonlinear function  $T : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$  satisfying  $\frac{\partial E(q,\theta)^T}{\partial q} = T(q-r,d)$  to model the elastic joint torque, where  $\theta = [r+d, r-d]^T$  is a coordinate transformation,  $r \in \mathbb{R}$  is a motor reference position, and  $d \in \mathbb{R}$  is a parameter for stiffness adjustment. For AA-VSAs, d is the joint co-contraction level, which is treated to be quasi-static here. Let  $\tau := T(q-r, d)$ . Then, one gets a simplified AA-VSA model

$$M\ddot{q} + D\dot{q} + g(q) = \tau + \tau_{\rm un} \tag{3}$$

where r can be regarded as a new control input. A control scheme shown in Fig. 1 will be designed for (3) in Sec. III.

Define a position tracking error  $e(t) := q(t) - q_d(t)$ , where  $q_d(t) \in \mathbb{R}$  is a joint reference position. We aim to design a suitable control input r for the system (3) such that e is as small as possible under a quasi-static d. Also, r should not change the inherent physical stiffness of (3) beyond a specific limit. That is, given a constant  $\zeta \in \mathbb{R}^+$ , denoting the maximum stiffness variation, the closed-loop stiffness  $\frac{\partial T(q-r,d)}{\partial q}$  should be within a neighborhood of the radius  $\zeta$  of the open-loop stiffness expressed by [3]

$$\left|\frac{\partial T(q-r,d)}{\partial q}\right|_{q\equiv r} - \frac{\partial T(q-\psi(q),d)}{\partial q}\Big|_{q\equiv q_*}\right| \le \zeta \quad (4)$$

where  $\psi(q) \in \mathbb{R}$  is the feedback control part,  $\frac{\partial T(q-\psi(q),d)}{\partial q}$  is the open-loop stiffness of the system (3), and  $q_* \in \mathbb{R}$  is such a value that satisfies  $\psi(q_*) = q_*$ .

*Remark 1*. The above two assumptions have been used in the ILC method of AA-VSAs in [4]. The function T(q - r, d) contains the compliance information of AA-VSAs, and the closed-loop stiffness  $\frac{\partial T(q-r,d)}{\partial q}$  can be adjusted by designing the control input r and setting the adjustment parameter d.

#### **III. GAUSSIAN PROCESS CONTROL DESIGN**

## A. Gaussian Process Modeling

Consider a training data set  $\mathcal{D} = \{X, y\}$  generated by  $y_i = f(x_i) + \eta$  with  $f : \mathbb{R}^n \to \mathbb{R}$ , where  $X := [x_1, x_2, \cdots, x_N]^T \in \mathbb{R}^{N \times n}$ ,  $y := [y_1, y_2, \ldots, y_N]^T \in \mathbb{R}^N$ ,  $\eta \sim \mathcal{N}(0, \sigma_n^2)$  is a noisy signal with the standard deviation  $\sigma_n \in \mathbb{R}^+$ , N is the number of data, and n is an input dimension. Let  $\mathcal{GP}(\mu, k(x, x'))$  be an approximation of f, denoted by  $f \sim \mathcal{GP}(\mu, k(x, x'))$ . Given a test input  $x^* \in \mathbb{R}^n$ , the prediction of f is determined by the mean  $\mu \in \mathbb{R}$  and variance  $\sigma \in \mathbb{R}^+$  as follows:

$$\mu(f|\boldsymbol{x}^*, \mathcal{D}) = K_{\boldsymbol{x}^*X}^T (K_{XX} + I\sigma_n^{-2})^{-1} \boldsymbol{y},$$
(5)

$$\sigma(f|\boldsymbol{x}^*, \mathcal{D}) = K_{\boldsymbol{x}^*\boldsymbol{x}^*} - K_{\boldsymbol{x}^*\boldsymbol{X}}^T (K_{XX} + I\sigma_n^2)^{-1} K_{\boldsymbol{x}^*\boldsymbol{X}}$$
(6)

with  $K_{XX} := K(X, X) \in \mathbb{R}^{N \times N}$ ,  $K_{\boldsymbol{x}^* \boldsymbol{x}^*} := K(\boldsymbol{x}^*, \boldsymbol{x}^*) \in \mathbb{R}$ , and  $K_{\boldsymbol{x}^* X} := K(\boldsymbol{x}^*, X) \in \mathbb{R}^{1 \times N}$ , in which K(X, X) is a covariance matrix given by

$$K(X,X) := \begin{bmatrix} k(\boldsymbol{x}_1, \boldsymbol{x}_1) & \cdots & k(\boldsymbol{x}_1, \boldsymbol{x}_N) \\ \vdots & \ddots & \vdots \\ k(\boldsymbol{x}_N, \boldsymbol{x}_1) & \cdots & k(\boldsymbol{x}_N, \boldsymbol{x}_N) \end{bmatrix}.$$
 (7)

The kernel function k, which represents the correlation between two points, is frequently chosen as a Gaussian type

$$k(\boldsymbol{x}, \boldsymbol{x}') = \phi^2 \exp[-(\boldsymbol{x} - \boldsymbol{x}')\Lambda(\boldsymbol{x} - \boldsymbol{x}')/2]$$

with  $x, x' \in \mathbb{R}^n$ , where  $\Lambda \in \mathbb{R}^{n \times n}$  and  $\phi \in \mathbb{R}^+$  are hyperparameters that can be obtained by optimizing the likelihood function during training. The Gaussian kernel function has the advantages of powerful modeling, good numerical stability, and simple hyperparameters selection, but the disadvantages of weak interpretability and slow calculation speed [13].

Consider the dynamic model (3) with M, D, and g(q) being uncertain. A prior estimation of (3) is given by

$$\hat{\tau} = M\ddot{q} + D\dot{q} + \hat{g}(q) \tag{8}$$

where  $\hat{M} \in \mathbb{R}^+$ ,  $\hat{D} \in \mathbb{R}^+$ , and  $\hat{g}(q) \in \mathbb{R}$  are estimates of M, D, and g(q), respectively, which satisfy the following property.

Property 1. There exist constants  $m_1, m_2, c_d \in \mathbb{R}^+$  that satisfy  $m_1 \leq \hat{M} \leq m_2$  and  $\hat{D} \leq c_d$ .

When there is no the prior knowledge of (3), set  $\hat{M} = c_{\rm m}$ ,  $\hat{D} = 0$ , and  $\hat{g} = 0$ , in which  $c_{\rm m} \in \mathbb{R}^+$  is a constant. Then, a GP model is trained offline with the data set  $\mathcal{D} = \{p_i, \tilde{\tau}_i\}_{i=1}^N$  with  $p_i := [q, \dot{q}]^T \in \mathbb{R}^2$  to learn the unknown part of (3), where  $\tilde{\tau}_i \in \mathbb{R}$  is obtained by subtracting (8) from (3) as follows:

$$\tilde{\tau}_i = M\ddot{q} + D\dot{q} + \tilde{g}(q) - \tau_{\rm un} \tag{9}$$

with  $\tilde{M} := M - \hat{M}$ ,  $\tilde{D} := D - \hat{D}$ , and  $\tilde{g} := g - \hat{g}$ . The GP model only uses q and  $\dot{q}$  as inputs, whereas the acceleration  $\ddot{q}$  is inferred from q and  $\dot{q}$  to avoid the immeasurable  $\ddot{q}$  [14]. Note that  $p_i$  and  $\tilde{\tau}_i$  denote the *i*th data in  $\mathcal{D}$ , generated by a real AA-VSA described by (3) with a special control input r, and the subscript i is omitted below. Next, we give the bound of a modeling error  $\Delta := \mu(\tilde{\tau}|p, \mathcal{D}) - \tilde{\tau} \in \mathbb{R}$ .



Fig. 2. A qbmove Advanced actuator equipped with a link load fixed on the base for experiments, where the link load can move in the vertical plane.

Lemma 1 [15, Th. 6]. For the uncertainty  $\tilde{\tau}$  given by (9) with the bounded reproducing kernel Hilbert space norm  $\|\tilde{\tau}\|_k < \infty$ on any compact set  $\Omega \subset \mathbb{R}^2$ ,  $\Delta$  is bounded by

$$P\left\{\left|\mu(\tilde{\tau}|\boldsymbol{p},\mathcal{D})-\tilde{\tau}\right|\leq\beta\sigma^{\frac{1}{2}}(\tilde{\tau}|\boldsymbol{p},\mathcal{D})\right\}\geq\delta$$

with  $\boldsymbol{p} \in \Omega$ , where  $\delta \in (0, 1)$  is a probability value, and

$$\beta = \sqrt{2 \|\tilde{\tau}\|_k^2 + 300\gamma \ln^3 \left(\frac{N+1}{1-\delta}\right)} \in \mathbb{R}^+,$$
  
$$\gamma = \max_{p_1,\dots,p_N \in \Omega} \frac{1}{2} \log |I_N + \sigma_n^{-2} K_{XX}| \in \mathbb{R}$$

with  $X = [\mathbf{p}_1, \dots, \mathbf{p}_N]^T \in \mathbb{R}^{N \times 2}$ .

*Remark 2.* This study assumes that the function T(q - r, d) can be known a priori. Then, the control torque  $\tau$  in (3) can be obtained by the substitution of the real control input r. When using a GP model to predict the elastic torque  $\tilde{\tau}$ , r needs to be obtained by solving  $\tau = T(q - r, d)$ , which can be calculated offline to reduce the computational burden during control. The above assumption facilitates the GP control design, and model uncertainties resulting from the inaccuracy of T(q - r, d) can be included in the lumped uncertainty  $\tau_{un}$  and learned by the GP model. Hence, the inaccuracy of T(q - r, d) does not affect the learning of the unknown part in (3).

#### B. Gaussian Process Tracking Control

Compared with NN learning, GP possesses two distinctive features [16]: 1) It is more convenient to achieve knowledge acquisition without the stringent condition termed persistent excitation, where its learning capacity depends on the number of stored data; 2) it provides a fidelity measure of the learned model via the predicted variance, which can be utilized to improve control robustness. Analogously to (6), let  $\sigma_p(q) :=$  $\sigma(\tilde{\tau}|q, D) \in \mathbb{R}^+$  and  $\sigma_d(\dot{q}, q) := \sigma(\tilde{\tau}|\dot{q}, q, D) \in \mathbb{R}^+$ . Define variable feedback gains  $K_p(\sigma_p) := \sigma_p(q) + K_{pc} \in \mathbb{R}^+$  and  $K_d(\sigma_d) := \sigma_d(\dot{q}, q) + K_{dc} \in \mathbb{R}^+$  with  $K_{pc}, K_{dc} \in \mathbb{R}^+$  being constants. Then, three assumptions are given as follows.

Assumption 1. The desired trajectory  $q_d$  and velocity  $\dot{q}_d$  are bounded by  $|q_d| \leq c_{d1}, |\dot{q}_d| \leq c_{d2}$  with  $c_{d1}, c_{d2} \in \mathbb{R}^+$ .



Fig. 3. Example profiles of the training data set  $\mathcal{D}$  consisting of p and  $\tilde{\tau}$ .



Fig. 4. The fidelity of the learned GP model via the predicted variance, where the gray region represents the 95% confidence intervals.

Assumption 2. The variable feedback gains  $K_p$  and  $K_d$  are continuous and bounded by

$$k_{p1} \le K_{p}(\sigma_{p}) \le k_{p2},$$
  
$$k_{d1} \le K_{d}(\sigma_{d}) \le k_{d2}$$

 $\forall q, \dot{q} \in \mathbb{R} \text{ with } k_{p1}, k_{p2}, k_{d1}, k_{d2} \in \mathbb{R}^+ \text{ some constants.}$ 

Assumption 3 [4]. The second control goal expressed by (4) in Sec. II restricts the use of high feedback gains, i.e., if

$$\left|\frac{\partial\psi(q)}{\partial q}\right|_{q\equiv q_*}\right| \leq \zeta \left|\frac{\partial T(q-\psi(q),d)}{\partial q}\right|_{q\equiv q_*}\Big|^{-1}$$
(10)

then the constraint (4) holds.

Consider the system (3) and a GP model trained by the data set  $\mathcal{D}$  with (9). The control torque  $\tau$  given by

$$\tau = \underbrace{M\ddot{q}_{d} + D\dot{q}_{d} + \hat{g}(q) + \mu(\tilde{\tau}|\boldsymbol{p}_{d}, \mathcal{D})}_{\tau_{\rm ff}} \underbrace{-K_{\rm d}(\sigma_{\rm d})\dot{e} - K_{\rm p}(\sigma_{\rm p})e}_{\tau_{\rm fb}}$$
(11)

with  $p_d := [q_d, \dot{q}_d]^T$  and  $\mathbf{e} := [e, \dot{e}]^T$ , which is composed of a feedback control term  $\tau_{\text{fb}} \in \mathbb{R}$  and a feedforward control term  $\tau_{\text{ff}} \in \mathbb{R}$  containing the learned GP torque  $\mu(\tilde{\tau}|\mathbf{p}_d, \mathcal{D})$  and the prior model knowledge (8) [see Fig. 1]. To explain the stability and convergence of the system (3) under the control law (11), we adopt a Lyapunov function candidate [16]

$$V(\mathbf{e}) = \frac{1}{2}\hat{M}\dot{e}^2 + \int_0^e z K_{\mathbf{p}}(\sigma_{\mathbf{p}}(z))dz + \epsilon \hat{M}e\dot{e} \qquad (12)$$

where  $\epsilon \in \mathbb{R}^+$  is a small coefficient, and  $z \in \mathbb{R}$  is an integral variable. To show that (12) is positive-definite and bounded, we use *Property 1*, *Assumption 2* and make some transformations to obtain that there exist constants  $c_1, c_2 \in \mathbb{R}^+$  such that

$$c_1 \|\mathbf{e}\| \le V(\mathbf{e}) \le c_2 \|\mathbf{e}\|. \tag{13}$$



Fig. 5. Control results by the proposed GP controller, the PII controller with high gains, and the 1st and 10th iterations of the ILC controller.



Fig. 6. Control results of the proposed GP controller under the change of the desired trajectory  $q_d$  to show generalization.

With *Lemma 1*, there exist constants  $c_3, \varepsilon \in \mathbb{R}^+$  to get

$$P\{V(\mathbf{e}) \le -c_3 V(\mathbf{e}) + \varepsilon\} \ge \delta.$$
(14)

Now, applying [17, Lemma 2] to the results of (13) and (14), one gets that the closed-loop system (3) with (11) is stable in the sense that the tracking error e is UUB and exponentially converges to a ball with a given probability  $\delta$ .

*Remark 3. Assumption 3* shows that the inherent mechanical behavior of AA-VSAs can be preserved by using a low-gain controller. Thus, the control gains  $K_p$  and  $K_d$  should be kept low. As we assume that there is no prior knowledge about the gravity term g(q), the feedforward action  $\tau_{\rm ff}$  in (11) does not depend on q so does not affect Assumption 3.

## **IV. EXPERIMENTAL STUDIES**

This section implements the GP controller (11) to a qbmove Advanced actuator with load [see Fig. 2], in which the control torque  $\tau$  is transferred into the position input r as discussed in *Remark 2*. The qbmove Advanced actuator is a bidirectional AA-VSA that reaches the best performance in the qbmove family [12], where its compliant mechanism is implemented by two motors connected at the output shaft by nonlinear springs, and its actuation torque  $\tau$  in (3) is explicitly expressed by

$$\tau = -k_{\theta 1} \sinh(a_1(q - r - d)) - k_{\theta 2} \sinh(a_2(q - r + d))$$
(15)

where  $a_1 = 8.9995$ ,  $a_2 = 8.9989$ ,  $k_{\theta 1} = 0.0026$ , and  $k_{\theta 2} = 0.0011$  are elastic parameters from the data sheet.

The training data set  $\mathcal{D}$  contains  $q, \dot{q}$ , and  $\tilde{\tau}$ , in which q is obtained by the actuator encoder, and  $\dot{q}$  is estimated through the low-pass filtering of the measured q. As the prior knowledge of the model (3) is unavailable, set  $\hat{M} = 0.001$ ,  $\hat{D} = 0$ , and  $\hat{q} = 0$  in (9). Then, the uncertainty  $\tilde{\tau}$  is calculable by (9) and (15). Note that the model uncertainty caused by the inaccuracy of (15) can be included in the lumped uncertainty  $\tau_{un}$  and learned by the GP model. The data collection is sampled at a frequency of 1 kHz, the data number N is set to be 2000, and the hyperparameters  $\Lambda = [0.5\ 0; 0\ 2.30]$  and  $\phi = 0.0085$ . Note that choosing N needs to consider the modeling error  $\Delta$ , GP input dimension n, and computational burden. After n is fixed, N should be chosen carefully as a small N may lead to a large  $\Delta$ , while a large N may also increase  $\Delta$  as optimizing the length-scale parameter  $\phi$  of GPs becomes computationally intractable [18]. Fig. 3 illustrates  $\mathcal{D}$  under the above input signal r. The fidelity of the learned GP model via the predict+ed variance is verified in Fig. 4, and hence, the GP prediction mean  $\mu(\tilde{\tau}|\boldsymbol{p}_{d}, \mathcal{D})$  can be used to compensate for the unknown part in (3) in the GP controller (11).

The feedback gains of the proposed controller are parameterized as  $K_p(\sigma_p) = 100\sigma_p + 0.004$  and  $K_d(\sigma_d) = 20\sigma_d + 0.001$ . The open-loop stiffness  $\frac{\partial T(q-\psi(q),d)}{\partial q}$  in (4) is calculated as 0.033 N.m/rad by (15) and the rule in [3]. Then, the maximum stiffness variation  $\zeta$  in (4) is calculated as 0.005 N.m/rad by (10) under the proposed controller. Two position-based VSA controllers are chosen as baselines, including the proportional integral-integral (PII) and position-based ILC controllers in [4]. The control gains of the PII controller are set higher than those of the proposed controller to make the tracking performance comparable, where its maximum stiffness variation  $\zeta$  becomes 0.04 N.m/rad, much higher than that of the proposed controller. The ILC controller with the iteration number being set as 10 is compared similarly to the above.

To show the tracking and softness preservation abilities of the proposed controller, the stiffness parameter d is set to 0 rad, implying the minimum stiffness. We use a desired trajectory

$$q_{\rm d}(t) = 0.074t^5 - 0.393t^4 + 0.589t^3, t \in [0, 2),$$

which is a smooth path from 0 to 0.7854 rad with the terminal time  $t_f = 2$  s. Control results by the above three controllers are depicted in Fig. 5. The proposed controller achieves guaranteed tracking accuracy, but the ILC controller at the first iteration and the PII controller show much larger tracking errors *e* than the proposed controller. Note that the link gravity is considered in this case, so the tracking accuracy of the PII controller is not as good as that in [4] without considering link gravity. The ILC controller performs worse than the GP controller at the beginning but achieves higher tracking accuracy after several iterations [see the 10th iteration of ILC in Fig. 5] at the cost of repeating the control task. To evaluate the generalization of the proposed controller, we keep the experimental setup the same as before but set a different desired trajectory  $q_d$ . Control results by the proposed controller are depicted in Fig. 6, which show high tracking accuracy without retraining under the change of  $q_d$ . From our experiments, the PII controller still shows poor tracking accuracy even with high feedback gains when  $q_d$  changes, resulting in a large stiffness variation  $\zeta$ ; the ILC controller achieves good tracking accuracy after several iterations, but it needs to retrain iteratively when  $q_d$  changes. Experimental trajectories by the PII and ILC controllers are not depicted as they are similar to those in Fig. 5. In sharp contrast, the proposed controller achieves guaranteed tracking accuracy, preserves the actuator softness due to the low feedback gains, and possesses favorable generalization.

## V. CONCLUSIONS

This paper has developed a machine learning-based tracking controller that combines GP feedforward and low-gain feedback actions for AA-VSAs, where a GP model is applied to learn the inverse actuator dynamics and provide model fidelity, and feedback gains are adapted according to model fidelity to enhance the tracking performance. Experiments on the qbmove Advanced actuator have validated that the proposed controller achieves guaranteed tracking accuracy, preserves mechanical behavior, and has better generalization than stateof-the-art control approaches. Future work would consider the time-varying stiffness of AA-VSAs, the extension of the proposed method to the multi-DoF case, and the combination of online learning and adaptive control to improve the tracking performance and applicability [19]–[21].

## REFERENCES

- R. Ham, T. Sugar, B. Vanderborght, K. Hollander, and D. Lefeber, "Compliant actuator designs," *IEEE Robot. Autom. Mag.*, vol. 16, no. 3, pp. 81–94, Sep. 2009.
- [2] F. Angelini, C. Della Santina, M. Garabini, M. Bianchi, and A. Bicchi, "Control architecture for human-like motion with applications to articulated soft robots," *Front. Robot. AI*, vol. 7, Sep. 2020, Art. No. 117.
- [3] C. Della Santina, M. Bianchi, G. Grioli, F. Angelini, M. Catalano, M. Garabini, and A. Bicchi, "Controlling soft robots: Balancing feedback and feedforward elements," *IEEE Robot. Autom. Mag.*, vol. 24, no. 3, pp. 75–83, Sep. 2017.
- [4] F. Angelini, C. D. Santina, M. Garabini, M. Bianchi, G. M. Gasparri, G. Grioli, M. G. Catalano, and A. Bicchi, "Decentralized trajectory tracking control for soft robots interacting with the environment," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 924–935, Aug. 2018.

- [5] F. Angelini, R. Mengacci, C. D. Santina, M. G. Catalano, and G. Grioli, "Time generalization of trajectories learned on articulated soft robots," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 3493–3500, Apr. 2020.
- [6] Z. Guo, Y. Pan, T. Sun, Y. Zhang, and X. Xiao, "Adaptive neural network control of serial variable stiffness actuators," *Complexity*, vol. 2017, pp. 1–9, Aug. 2017.
- [7] D.-T. Tran, M.-N. Nguyen, and K. K. Ahn, "RBF neural network based backstepping control for an electrohydraulic elastic manipulator," *Appl. Sci.*, vol. 9, Apr. 2019, Art. No. 11.
- [8] G. Liu, N. Sun, D. Liang, Y. Chen, T. Yang, and Y. Fang, "Neural network-based adaptive command filtering control for pneumatic artificial muscle robots with input uncertainties," *Control Eng. Pract.*, vol. 118, pp. 104 960–104 971, Jan. 2022.
- [9] D. Mitrovic, S. Klanke, and S. Vijayakumar, "Learning impedance control of antagonistic systems based on stochastic optimization principles," *Int. J. Robot. Res.*, vol. 30, no. 5, pp. 556–573, Dec. 2011.
  [10] K. Nikola, L. Branko, and K. Jovanovic, "Feedforward control ap-
- [10] K. Nikola, L. Branko, and K. Jovanovic, "Feedforward control approaches to bidirectional antagonistic actuators based on learning," in *Advances in Service and Industrial Robotics*. Cham, Switzerland: Springer, 2020, pp. 337–345.
- [11] C. E. Rasmussen and C. K. I. Williams, Gaussian Processes for Machine Learning. Cambridge, MA, USA: MIT Press, 2006.
- [12] C. Della Santina, C. Piazza, G. M. Gasparri, M. Bonilla, M. G. Catalano, G. Grioli, M. Garabini, and A. Bicchi, "The quest for natural machine motion: An open platform to fast-prototyping articulated soft robots," *IEEE Robot. Autom. Mag.*, vol. 24, no. 1, pp. 48–56, Mar. 2017.
- [13] G. Pillonetto, F. Dinuzzo, T. Chen, G. De Nicolao, and L. Ljung, "Kernel methods in system identification, machine learning and function estimation: A survey," *Automatica*, vol. 50, no. 3, pp. 657–682, 2014.
- [14] B. Valencia-Vidal, E. Ros, I. Abadía, and N. R. Luque, "Bidirectional recurrent learning of inverse dynamic models for robots with elastic joints: a real-time real-world implementation," *Front. Neurorobot.*, vol. 17, pp. 1–15, Jun. 2023.
- [15] N. Srinivas, A. Krause, S. M. Kakade, and M. W. Seeger, "Informationtheoretic regret bounds for Gaussian process optimization in the bandit setting," *IEEE Trans. Inf. Theory*, vol. 58, no. 5, pp. 3250–3265, May 2012.
- [16] T. Beckers, D. Kulić, and S. Hirche, "Stable Gaussian process based tracking control of Euler–Lagrange systems," *Automatica*, vol. 103, pp. 390–397, May 2019.
- [17] T. Ravichandran, D. Wang, and G. Heppler, "Stability and robustness of a class of nonlinear controllers for robot manipulators," in *Proc. Am. Control Conf.*, Boston, MA, USA, 2004, pp. 5262–5267.
- [18] E. Rueckert, M. Nakatenus, S. Tosatto, and J. Peters, "Learning inverse dynamics models in O(n) time with LSTM networks," in *Proc. IEEE Int. Conf. Humanoid Robot.*, Birmingham, UK, 2017, pp. 811–816.
- [19] W. Li, Z. Li, Y. Liu, and Y. Pan, "Learning robot inverse dynamics using sparse online gaussian process with forgetting mechanism," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatron.*, Sapporo, Japan, 2022, p. 638–643.
- [20] L. Deng, W. Li, and Y. Pan, "Data-efficient gaussian process online learning for adaptive control of multi-dof robotic arms," *IFAC-PapersOnLine*, vol. 55, no. 2, pp. 84–89, 2022.
- [21] L. Deng, Z. Li, and Y. Pan, "Sparse online gaussian process impedance learning for multi-dof robotic arms," in *Proc. IEEE Int. Conf. Adv. Robot. Mechatron.*, Chongqing, China, 2021, pp. 199–206.