



# Gaussian Process-Based Learning Control of Underactuated Balance Robots With an External and Internal Convertible Modeling Structure

**Feng Han<sup>1</sup>**

Mem. ASME

Department of Mechanical Engineering,  
New York Institute of Technology,  
Old Westbury, NY 11568  
e-mail: fhnyit@yahoo.com

**Jingang Yi**

Fellow ASME

Department of Mechanical and  
Aerospace Engineering,  
Rutgers, The State University of New Jersey,  
Piscataway, NJ 08854  
e-mail: jgyi@rutgers.edu

*External and internal convertible (EIC) form-based motion control is one of the effective designs of simultaneous trajectory tracking and balance for underactuated balance robots. Under certain conditions, the EIC-based control design is shown to lead to uncontrolled robot motion. To overcome this issue, we present a Gaussian process (GP)-based data-driven learning control for underactuated balance robots with the EIC modeling structure. Two GP-based learning controllers are presented by using the EIC property. The partial EIC (PEIC)-based control design partitions the robotic dynamics into a fully actuated subsystem and a reduced-order underactuated subsystem. The null-space EIC (NEIC)-based control compensates for the uncontrolled motion in a subspace, while the other closed-loop dynamics are not affected. Under the PEIC- and NEIC-based, the tracking and balance tasks are guaranteed, and convergence rate and bounded errors are achieved without causing any uncontrolled motion by the original EIC-based control. We validate the results and demonstrate the GP-based learning control design using two inverted pendulum platforms.* [DOI: 10.1115/1.4065937]

## 1 Introduction

An underactuated balance robot possesses fewer control inputs than the number of degrees-of-freedom (DOFs) [1,2]. Motion control of underactuated balance robots requires both the trajectory tracking of the actuated subsystem and balance control of the unactuated, unstable subsystem [3–5]. Inverting the nonminimum phase unactuated nonlinear dynamics brings additional challenges in causal feedback control design. Several modeling and control methods have been proposed for these robots and their applications [4–10]. Orbital stabilization method was used for balancing underactuated robots [1,11–13], with applications to bipedal robot [14] and cart-inverted pendulum [1]. Energy shaping-based control was also designed for underactuated balance robots [15,16]. One feature of those methods is that the achieved balance-enforced trajectory is not unique and cannot be prescribed explicitly [1,11–13]. In Refs. [5] and [17], a simultaneous trajectory tracking and balance control of underactuated balance robots was proposed by using the property of the external and internal convertible (EIC) form of the robot dynamics. The EIC-based control has been demonstrated as one of the effective approaches to achieve fast convergence with guaranteed performance.

The above-mentioned control designs require an accurate model of robot dynamics, and the control performance would deteriorate under model uncertainties or external disturbances. Machine learning-based methods provide an efficient tool for robot modeling and control [18,19]. In particular, Gaussian process (GP) regression is an effective learning approach that generates nearly analytical structure and bounded prediction errors [7,19–21]. Development of GP-based performance-guaranteed control for underactuated balance robots has been reported in Refs. [4], [20], and [22]. In Ref. [4], the control design was conducted in two steps. A GP-based inverse dynamics controller for unactuated subsystem to achieve balance and a model predictive control (MPC) was used to simultaneously track the given reference trajectory and estimate the balance equilibrium manifold (BEM). The GP prediction uncertainties were incorporated into the control design to enhance the control robustness. The work in Ref. [5] followed the sequential control design in the EIC-based framework, and the controller was adaptive to the prediction uncertainties. The training data were selected to reduce the computational complexity.

This work takes advantage of the structured GP modeling approach in Refs. [5] and [7] and presents an integration of EIC-based control with GP models. We first present the conditions under which uncontrolled motions exist under the original EIC-based control design for underactuated balance robots. We identify these conditions and design the stable GP-based learning control with the properly selected nominal robot dynamic model. Two different controllers, called partial- and null-space-EIC (i.e., PEIC- and NEIC), are presented to improve the closed-loop performance. The PEIC-based control constructs a virtual inertia matrix to reshape the dynamics coupling

The 2024 IEEE International Conference on Robotics and Automation.

<sup>1</sup>Corresponding author.

Contributed by the Dynamic Systems Division of ASME for publication in the JOURNAL OF DYNAMIC SYSTEMS, MEASUREMENT, AND CONTROL. Manuscript received November 29, 2023; final manuscript received June 17, 2024; published online August 17, 2024. Assoc. Editor: Vaibhav Srivastava.

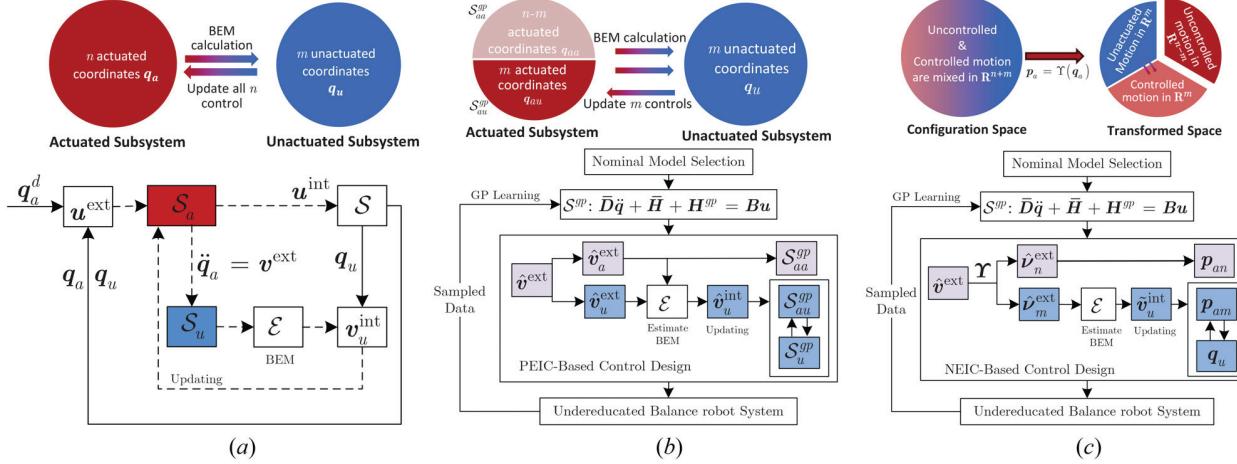
between the actuated and unactuated subsystems. The EIC-induced uncontrolled motion is eliminated, and the robotic system behaves as a combined fully actuated subsystem and a reduced-order unactuated subsystem. Alternatively, the compensation effect in the NEIC-based control is applied to the uncontrolled coordinates in the null space, while the other part of the stable system motion stays unchanged. The PEIC- and NEIC-based controls achieve guaranteed robust performance with a fast convergence of the closed-loop tracking errors.

The control tasks considered in this work include both the trajectory tracking for the actuated subsystem and platform balance for the unstable subsystem. The interconnection between these two subsystems lies in implicit dynamic relationship that needs to be estimated in real time. The control problem considered here distinguishes from the work in literature. Most existing approaches, such as orbital stabilization and energy shaping, focus on stabilization only, that is, the trajectory of the actuated subsystem is not prescribed, and the main control task is to stabilize the unstable subsystem. The main contribution of this work lies in the new GP-based learning control of underactuated balance robots using the EIC structural properties. Compared with the approaches in Refs. [5] and [17], this work reveals underlying design properties and limitations of the original EIC-based control for underactuated balance robots. Compared with the work in Refs. [4] and [23], the proposed method takes advantage of the attractive EIC modeling properties for control design and does not use MPC that requires high computational demands. Compared with other learning control methods such as reinforcement learning, the proposed control integrates the robot's dynamics property (i.e., EIC structure) and the GP-based model learning. By integrating physics knowledge into model learning, we identify the conditions for nominal model selection, and the proposed control is designed with guaranteed performance. This paper is an extension of the previous conference submission [24] with new design, analysis, and experiments. Particularly, the NEIC-based control design and experiments were not presented in Ref. [24].

The rest of the paper is outlined as follows. We introduce the EIC-based control and present the problem statement in Sec. 2. Section 3 presents the GP-based robot dynamics. The PEIC- and NEIC-based controls are presented in Sec. 4. The stability analysis is discussed in Sec. 5. The experimental results are presented in Sec. 6, and finally Sec. 7 summarizes the concluding remarks.

## 2 External and Internal Convertible-Based Robot Control and Problem Statement

**2.1 Robot Dynamics and External and Internal Convertible-Based Control.** We consider an underactuated balance robot with  $(n+m)$  DOFs,  $n, m \in \mathbb{N}$ , and the generalized



**Fig. 1** Illustrative diagrams for (a) the original EIC-based control, (b) the PEIC-control design, and (c) the NEIC-based control design. The top row shows the general idea for the control design, and the bottom row illustrates the information flow in the design. In (a), the dashed line indicates the design flow, and the solid line indicates the control flow.

coordinates are denoted as  $\mathbf{q} \in \mathbb{R}^{n+m}$ . The robot dynamics is expressed as

$$\mathcal{S} : \mathbf{D}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{G}(\mathbf{q}) = \mathbf{B}\mathbf{u} \quad (1)$$

where  $\mathbf{D}(\mathbf{q})$ ,  $\mathbf{C}(\mathbf{q}, \dot{\mathbf{q}})$ , and  $\mathbf{G}(\mathbf{q})$  are the inertia matrix, Coriolis, and gravity matrix, respectively.  $\mathbf{B}$  denotes the input matrix, and  $\mathbf{u} \in \mathbb{R}^n$  is the control input. The coordinates are partitioned as  $\mathbf{q} = [\mathbf{q}_a^T \mathbf{q}_u^T]^T$ , with actuated coordinate  $\mathbf{q}_a \in \mathbb{R}^n$  and unactuated coordinate  $\mathbf{q}_u \in \mathbb{R}^m$ . We focus on the case  $n \geq m$ , and without loss of generality, we assume that  $\mathbf{B} = [\mathbf{I}_n \mathbf{0}]^T$ , where  $\mathbf{I}_n \in \mathbb{R}^n$  is the identity matrix with dimension  $n$ . The robot dynamic model in Eq. (1) is rewritten as

$$\mathcal{S}_a : \mathbf{D}_{aa}\ddot{\mathbf{q}}_a + \mathbf{D}_{au}\ddot{\mathbf{q}}_u + \mathbf{H}_a = \mathbf{u} \quad (2a)$$

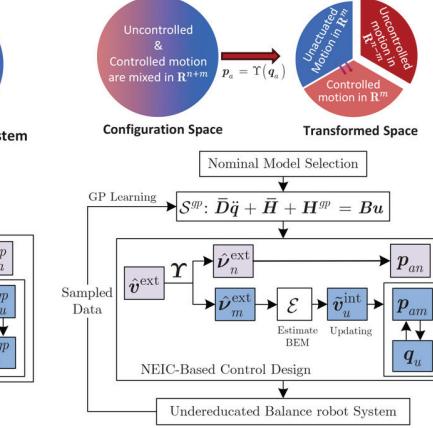
$$\mathcal{S}_u : \mathbf{D}_{ua}\ddot{\mathbf{q}}_a + \mathbf{D}_{uu}\ddot{\mathbf{q}}_u + \mathbf{H}_u = \mathbf{0} \quad (2b)$$

for actuated ( $\mathcal{S}_a$ ) and unactuated ( $\mathcal{S}_u$ ) subsystems, respectively. Subscripts “ $aa$  ( $uu$ )” and “ $ua$  ( $au$ )” indicate the variables related to the actuated (unactuated) coordinates and coupling effects, respectively. For presentation convenience, we introduce  $\mathbf{H} = \mathbf{C}\dot{\mathbf{q}} + \mathbf{G}$ ,  $\mathbf{H}_a = \mathbf{C}_a\dot{\mathbf{q}} + \mathbf{G}_a$ , and  $\mathbf{H}_u = \mathbf{C}_u\dot{\mathbf{q}} + \mathbf{G}_u$ , and the dependence of  $\mathbf{D}$ ,  $\mathbf{C}$ , and  $\mathbf{G}$  on  $\mathbf{q}$  and  $\dot{\mathbf{q}}$  is dropped. Subsystems  $\mathcal{S}_a$  and  $\mathcal{S}_u$  are also referred to as the external and internal subsystems, respectively [4,17].

The control goal is to steer actuated coordinate  $\mathbf{q}_a$  to follow a given desired trajectory  $\mathbf{q}_a^d$  for  $\mathcal{S}_a$ , while the unactuated, unstable subsystem  $\mathcal{S}_u$  is balanced at unknown equilibrium  $\mathbf{q}_u^e$ . Therefore, we need to estimate  $\mathbf{q}_u^e$  in real time to achieve simultaneously trajectory tracking (for  $\mathcal{S}_a$ ) and platform balance (for  $\mathcal{S}_u$ ). It is noted that not all arbitrary trajectories can be followed given the underactuated dynamics and balance requirement. Such a property has been explicitly discussed for the autonomous bikebot example in Ref. [25]. In this work, we assume that the given trajectory  $\mathbf{q}_a^d$  is well planned and the control exists. In this work, we assume that the given trajectory  $\mathbf{q}_a^d$  is well planned and the control exists. Designing and planning feasible trajectory  $\mathbf{q}_a^d$  is out of the scope of this work.  $\mathbf{q}_a^d$

The original EIC-based control design is considered in two steps [5,17]. As shown in the top figure in Fig. 1(a), the first step is to identify and estimate the unknown equilibrium  $\mathbf{q}_u^e$  under an external trajectory tracking control. With the estimated  $\mathbf{q}_u^e$ , the external control design is updated with simultaneously trajectory tracking and balancing tasks. Following such a concept, we first designs external input  $\mathbf{u}^{\text{ext}}$  to follow  $\mathbf{q}_a^d$  by temporarily neglecting  $\mathcal{S}_u$ , namely,

$$\mathbf{u}^{\text{ext}} = \mathbf{D}_{aa}\mathbf{v}^{\text{ext}} + \mathbf{D}_{au}\ddot{\mathbf{q}}_u + \mathbf{H}_a \quad (3)$$



where  $v^{\text{ext}} = \mathbf{q}_u^d - \mathbf{k}_{d1}\dot{\mathbf{q}}_a - \mathbf{k}_{p1}\mathbf{e}_a$  is the auxiliary input under which the tracking error  $\mathbf{e}_a = \mathbf{q}_a - \mathbf{q}_u^d$  converges to the origin, and  $\mathbf{k}_{p1}, \mathbf{k}_{d1}$  are diagonal matrices with positive elements. Assuming that  $\mathbf{u}^{\text{ext}}$  is applied to  $\mathcal{S}$  and  $\mathcal{S}_a$  follows  $\mathbf{q}_a^d$ ,  $\mathbf{q}_u$  should keep balance around its equilibrium, which is however unknown. Then, BEM is introduced and used to capture the equilibrium of  $\mathbf{q}_u$  under  $\ddot{\mathbf{q}}_a = v^{\text{ext}}$ , namely,

$$\mathcal{E} = \{\mathbf{q}_u : \Gamma(\mathbf{q}_u; v^{\text{ext}}) = \mathbf{0}, \dot{\mathbf{q}}_u = \ddot{\mathbf{q}}_u = \mathbf{0}\} \quad (4)$$

where  $\Gamma(\mathbf{q}_u; v^{\text{ext}}) = \mathbf{D}_{uu}\ddot{\mathbf{q}}_u + \mathbf{D}_{ua}v^{\text{ext}} + \mathbf{H}_u$ .  $\mathbf{q}_u^e$  is obtained by inverting  $\Gamma_0 = \Gamma(\mathbf{q}_u; v^{\text{ext}})|_{\dot{\mathbf{q}}_u=\ddot{\mathbf{q}}_u=0} = \mathbf{0}$ . Obtaining  $\mathbf{q}_u^e$  requires accurate system dynamics and needs to invert the nonminimum phase dynamics  $\mathcal{S}_u$ , which is challenging for noncausal control design.

To stabilize  $\mathbf{q}_u$  onto  $\mathcal{E}$ , the  $\mathbf{q}_a$  motion is updated as

$$v^{\text{int}} = -\mathbf{D}_{ua}^+(\mathbf{H}_u + \mathbf{D}_{uu}v_u^{\text{int}}) \quad (5)$$

where  $\mathbf{D}_{ua}^+ = (\mathbf{D}_{ua}^T\mathbf{D}_{ua})^{-1}\mathbf{D}_{ua}^T$  is the generalized inverse of  $\mathbf{D}_{ua}$ ,  $v_u^{\text{int}} = \mathbf{q}_u^e - \mathbf{k}_{d2}\dot{\mathbf{q}}_u - \mathbf{k}_{p2}\mathbf{e}_u$  is the auxiliary control that drives error  $\mathbf{e}_u = \mathbf{q}_u - \mathbf{q}_u^e$  toward zero, and  $\mathbf{k}_{p2}, \mathbf{k}_{d2}$  are diagonal matrices with positive elements. The final control is obtained by replacing  $v^{\text{ext}}$  in Eq. (3) with  $v^{\text{int}}$  in Eq. (5), that is,

$$\mathbf{u}^{\text{int}} = \mathbf{D}_{aa}v^{\text{int}} + \mathbf{D}_{au}\ddot{\mathbf{q}}_u + \mathbf{H}_a \quad (6)$$

where  $v^{\text{int}}$  is used as the virtual control input in  $\mathcal{S}_u$ , that is, under  $\ddot{\mathbf{q}}_a = v^{\text{int}}$ ,  $\ddot{\mathbf{q}}_u = v_u^{\text{int}}$ .

Figure 1(a) illustrates the above sequential EIC-based control design. It has been shown in Ref. [17] that the control  $\mathbf{u}^{\text{int}}$  guarantees both  $\mathbf{e}_a$  and  $\mathbf{e}_u$  convergence to a neighborhood of the origin exponentially if the high-order approximation terms of the closed-loop systems are affine with error  $\mathbf{e}$ . Therefore, the EIC-based control achieves trajectory tracking for  $\mathcal{S}_a$  and balancing task for  $\mathcal{S}_u$  simultaneously.

**2.2 Motion Property Under External and Internal Convertible-Based Control.** Control design (5) uses a mapping from low-dimensional ( $m$ ) to high-dimensional ( $n$ ) spaces (i.e.,  $n \geq m$ ). Under control (6) with properly selected control gains, it has been shown in Ref. [17] that there exists a finite time  $T > 0$ , and for small number  $\varepsilon > 0$ ,  $\|\mathbf{q}_u(t) - \mathbf{q}_u^e(t)\| < \varepsilon$  for  $t > T$ . Therefore, given the negligible error, we obtain  $\mathbf{D}_{ua}(\mathbf{q}_a, \mathbf{q}_u) \approx \mathbf{D}_{ua}(\mathbf{q}_a, \mathbf{q}_u^e)$ .

For  $\mathcal{S}$  in Eq. (2), if  $\text{rank}(\mathbf{D}_{au}) = m$  for all  $\mathbf{q}$ , applying singular value decomposition (SVD) to  $\mathbf{D}_{ua}$  and  $\mathbf{D}_{ua}^+$ , we obtain

$$\mathbf{D}_{ua} = \mathbf{U}\Lambda\mathbf{V}^T, \quad \mathbf{D}_{ua}^+ = \mathbf{V}\Lambda^+\mathbf{U}^T \quad (7)$$

where  $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m] \in \mathbb{R}^{m \times m}$  and  $\mathbf{V} \in \mathbb{R}^{n \times n}$  are unitary orthogonal matrices.  $\Lambda = [\Lambda_m \mathbf{0}] \in \mathbb{R}^{m \times n}$ ,  $\Lambda^+ = [\Lambda_m^{-1} \mathbf{0}]^T \in \mathbb{R}^{n \times m}$  and  $\Lambda_m = \text{diag}(\sigma_1, \dots, \sigma_m)$  with singular values  $\sigma_i > 0$ ,  $i = 1, \dots, m$ . We partition  $\mathbf{V}$  into the block matrix  $\mathbf{V} = [V_m \mathbf{V}_n]$ ,  $\mathbf{V}_m \in \mathbb{R}^{n \times m}$  and  $\mathbf{V}_n \in \mathbb{R}^{n \times (n-m)}$ . Since  $\text{rank}(\mathbf{D}_{au}) = m$ , the null space of  $\mathbf{D}_{ua}$  is  $\ker(\mathbf{D}_{ua}) = \text{span}(\mathbf{V}_n)$ .

Column vectors of matrix  $\mathbf{V}$  serve as a complete set of basis in  $\mathbb{R}^n$ , and we introduce a coordinate transformation  $\mathbf{Y} : \mathbf{x} \mapsto \mathbf{V}^T \mathbf{x}$  for  $\mathbf{x} \in \mathbb{R}^n$ . Clearly,  $\mathbf{Y}$  is a linear, time-varying, smooth map. Applying  $\mathbf{Y}$  to  $\mathbf{q}_a$  and  $v^{\text{ext}}$ , we have

$$\mathbf{p}_a = \mathbf{V}^T \mathbf{q}_a, \quad v^{\text{ext}} = \mathbf{V}^T v^{\text{ext}} \quad (8)$$

where  $\mathbf{p}_a = [\mathbf{p}_{am} \mathbf{p}_{an}^T]^T$ ,  $v^{\text{ext}} = [(\mathbf{v}_m^{\text{ext}})^T \ (\mathbf{v}_n^{\text{ext}})^T]^T$ , and  $\mathbf{p}_{am}, \mathbf{v}_m^{\text{ext}} \in \mathbb{R}^m$ ,  $\mathbf{p}_{an}, \mathbf{v}_n^{\text{ext}} \in \mathbb{R}^{n-m}$ . Note that  $[\mathbf{p}_a^T \ \mathbf{q}_u^T]^T$  still serves as a complete set of generalized coordinates for  $\mathcal{S}$ . Using the new coordinate  $\mathbf{p}_a$ , we have the following motion property under the original EIC-based control for  $\mathcal{S}$ , and the proof is given in Appendix A1.

**LEMMA 1.** For  $\mathcal{S}$  in Eq. (2), if  $\text{rank}(\mathbf{D}_{au}) = m$  holds for  $\mathbf{q}$  and all  $n$  control inputs appear in  $\mathcal{S}_u$  dynamics (through  $\ddot{\mathbf{q}}_a$ ), under the EIC-based control (6), the BEM in Eq. (4) is associated with only  $\mathbf{v}_m^{\text{ext}}$ , and robot dynamics can be written into

$$\mathcal{S}_{\text{EIC}} : \ddot{p}_{ai} = -\frac{\mathbf{u}_i^T(\mathbf{H}_u + \mathbf{D}_{uu}v_u^{\text{int}})}{\sigma_i}, \quad i = 1, \dots, m \quad (9a)$$

$$\ddot{p}_{aj} = 0, \quad j = m+1, \dots, n \quad (9b)$$

$$\ddot{\mathbf{q}}_u = v_u^{\text{int}} \quad (9c)$$

No control input appears for coordinates in  $\ker(\mathbf{D}_{ua})$  as shown in Eq. (9b), and only  $m$  actuated coordinates in  $\text{span}(\mathbf{V})$  are under active control, as shown in Eq. (9a). The results in Lemma 1 reveal the motion property of  $\mathcal{S}$  under the original EIC-based control design. The uncontrolled motion happens to a special set of underactuated balance robots under the conditions in Lemma 1. If the unactuated motion is only related to  $m$  (out of  $n$ ) control inputs, the motion (9b) vanishes, and the EIC-based control works well. In Ref. [5], the EIC-based control worked properly for the rotary inverted pendulum with  $n = m = 1$ . In Refs. [4] and [25], the EIC-based control also worked well for the bikebot with  $n = 2$  (planar motion) and  $m = 1$  (roll motion) but the roll motion depends on steering control only, that is, no velocity control, and therefore, does not satisfy the condition for Lemma 1. We will show an example of the three-link inverted pendulum platform that demonstrates the uncontrolled motion under the original EIC-based control in Sec. 6.

With the above-discussed motion property under the EIC-based control, we consider the following problem.

**Problem Statement:** The goal of robot control is to design an enhanced EIC-based learning control to drive the actuated coordinate  $\mathbf{q}_a$  to follow a given profile  $\mathbf{q}_a^d$  and simultaneously the unactuated coordinate  $\mathbf{q}_u$  to be stabilized on the estimated  $\mathbf{q}_u^e$ . The uncontrolled motion presented in Lemma 1 should be avoided for robot dynamics (2).

### 3 Gaussian Process-Based Robot Dynamics Model

We build a GP-based robot dynamics model that will be used for control design in Sec. 4.

**3.1 Gaussian Process-Based Robot Dynamics Model.** To keep it self-contained, we briefly review the GP regression model. We consider a multivariate continuously smooth function  $y = f(\mathbf{x}) + w$ ,  $\mathbf{x}_i \in \mathbb{R}^{n_x}$ , where  $w \in \mathbb{R}$  is the zero-mean Gaussian noise and  $n_x$  is the dimension of  $\mathbf{x}$ . Denote the training data as  $\mathbb{D} = \{\mathbf{X}, \mathbf{Y}\} = \{\mathbf{x}_i, y_i\}_{i=1}^N$ , where  $\mathbf{X} = \{\mathbf{x}_i\}_{i=1}^N$ ,  $\mathbf{Y} = \{y_i\}_{i=1}^N$ , and  $N \in \mathbb{N}$  is the number of the data point. The GP model is trained by maximizing posterior probability  $p(\mathbf{Y}; \mathbf{X}, \Theta)$  over the hyperparameters  $\Theta$ , that is,  $\Theta$  is obtained by solving

$$\min_{\Theta} -\log(\mathbf{Y}; \mathbf{X}, \Theta) = \min_{\Theta} -\frac{1}{2}\mathbf{Y}^T \mathbf{K}^{-1} \mathbf{Y} - \frac{1}{2}\log \det(\mathbf{K})$$

where  $\mathbf{K} = (K_{ij}), K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j) = \sigma_f^2 \exp(-(1/2)(\mathbf{x}_i - \mathbf{x}_j)^T \mathbf{W}(\mathbf{x}_i - \mathbf{x}_j)) + \vartheta^2 \delta_{ij}$ ,  $\mathbf{W} = \text{diag}\{W_1, \dots, W_{n_x}\} > 0$ ,  $\delta_{ij} = 1$  for  $i = j$ , and  $\Theta = \{\mathbf{W}, \sigma_f, \vartheta\}$  are hyperparameters.

The GP agent builds the joint distribution of new measurement  $\mathbf{x}^*$  and the training data as

$$\begin{bmatrix} \mathbf{Y} \\ \mathbf{y} \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \mathbf{K} & \mathbf{k}^T \\ \mathbf{k} & k^* \end{bmatrix}\right) \quad (10)$$

where  $\mathbf{k} = \mathbf{k}(\mathbf{x}^*, \mathbf{X})$  and  $k^* = \mathbf{k}(\mathbf{x}^*, \mathbf{x}^*)$ , and  $\mathcal{N}(\mu, \Sigma)$  denotes the Gaussian distribution with mean  $\mu$  and variance  $\Sigma$ . The mean value and variance for input  $\mathbf{x}^*$  are

$$\mu(\mathbf{x}^*) = \mathbf{k}^T \mathbf{K}^{-1} \mathbf{Y}, \quad \Sigma(\mathbf{x}^*) = k^* - \mathbf{k} \mathbf{K}^{-1} \mathbf{k}^T \quad (11)$$

We integrate the GP regression with a nominal model. For  $\mathcal{S}$  in Eq. (1), we first build a nominal model

$$\mathcal{S}^n : \bar{\mathbf{D}}\ddot{\mathbf{q}} + \bar{\mathbf{H}} = \mathbf{B}\mathbf{u} \quad (12)$$

where  $\bar{\mathbf{D}}$  and  $\bar{\mathbf{H}}$  are the nominal inertia and nonlinear matrices, respectively. Generally, the nominal dynamic model does not hold for the data sampled from the physical robot systems. The GP models are built to capture the difference between  $\mathcal{S}^n$  and  $\mathcal{S}$ , namely,

$$\mathbf{H}^e = \mathbf{D}\ddot{\mathbf{q}} + \mathbf{H} - \bar{\mathbf{D}}\ddot{\mathbf{q}} - \bar{\mathbf{H}} = \mathbf{B}\mathbf{u} - \bar{\mathbf{D}}\ddot{\mathbf{q}} - \bar{\mathbf{H}}$$

We build GP models to estimate  $\mathbf{H}^e = [(\mathbf{H}_a^e)^\top (\mathbf{H}_u^e)^\top]^\top$ , where  $\mathbf{H}_a^e$  and  $\mathbf{H}_u^e$  are for  $\mathcal{S}_a$  and  $\mathcal{S}_u$ , respectively. The training data  $\mathbb{D} = \{\mathbf{X}, \mathbf{Y}\}$  are sampled from  $\mathcal{S}$  as  $\mathbf{X} = \{\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{q}}\}$  and  $\mathbf{Y} = \{\mathbf{H}^e\}$ .

The GP predicted mean and variance are denoted as  $(\mu_i(\mathbf{x}), \Sigma_i(\mathbf{x}))$  for  $\mathbf{H}_i^e, i = a, u$ . The GP-based robot dynamics models  $\mathcal{S}_a^{\text{gp}}$  and  $\mathcal{S}_u^{\text{gp}}$  are given as

$$\mathcal{S}_a^{\text{gp}}: \bar{\mathbf{D}}_{aa}\ddot{\mathbf{q}}_a + \bar{\mathbf{D}}_{au}\ddot{\mathbf{q}}_u + \mathbf{H}_a^{\text{gp}} = \mathbf{u} \quad (13a)$$

$$\mathcal{S}_u^{\text{gp}}: \bar{\mathbf{D}}_{ua}\ddot{\mathbf{q}}_a + \bar{\mathbf{D}}_{uu}\ddot{\mathbf{q}}_u + \mathbf{H}_u^{\text{gp}} = \mathbf{0} \quad (13b)$$

where  $\mathbf{H}_i^{\text{gp}} = \bar{\mathbf{H}}_i + \mu_i(\mathbf{x}), i = a, u$ . The GP-based model prediction error is

$$\Delta = \begin{bmatrix} \Delta_a \\ \Delta_u \end{bmatrix} = \begin{bmatrix} \mu_a(\mathbf{x}) - \mathbf{H}_a^e \\ \mu_u(\mathbf{x}) - \mathbf{H}_u^e \end{bmatrix} \quad (14)$$

To quantify the GP prediction error, the following property for  $\Delta$  is obtained directly from Theorem 6 in Ref. [26].

LEMMA 2. *Given training dataset  $\mathbb{D}$ , if the kernel function  $k(\mathbf{x}_i, \mathbf{x}_j)$  is chosen such that  $\mathbf{H}_a^e$  for  $\mathcal{S}_a$  has a finite reproducing kernel Hilbert space norm  $\|\mathbf{H}_a^e\|_k < \infty$ , for given  $0 < \eta_a < 1$*

$$\Pr\left\{ \|\Delta_a\| \leq \|\kappa_a^\top \Sigma_a^{1/2}(\mathbf{x})\| \right\} \geq \eta_a \quad (15)$$

where  $\Pr\{\cdot\}$  denotes the probability of an event,  $\kappa_a \in \mathbb{R}^n$ , and its  $i$ th entry is  $\kappa_{ai} = \sqrt{2\|\mathbf{H}_{a,i}^e\|_k^2 + 300\zeta_i \ln^3((N+1)/(1-\eta_a^{1/n}))}$ ,  $\zeta_i = \max_{\mathbf{x}, \mathbf{x}' \in \mathcal{X}} (1/2) \ln |1 + \vartheta_i^{-2} k_i(\mathbf{x}, \mathbf{x}')|$ . A similar conclusion holds for  $\Delta_u$  with  $0 < \eta_u < 1$ .

**3.2 Nominal Model Selection.** The nominal model plays an important role in the EIC control. We consider the following conditions for choosing the nominal model  $\mathcal{S}^n$  to overcome the uncontrolled motion under the learning control.

$\mathcal{C}_1$ :  $\bar{\mathbf{D}} = \bar{\mathbf{D}}^\top$  is positive definite,  $\|\bar{\mathbf{D}}\| \leq d$ ,  $\|\bar{\mathbf{H}}\| \leq h$ , where constants  $0 < d, h < \infty$ ;

$\mathcal{C}_2$ :  $\text{rank}(\bar{\mathbf{D}}_{aa}) = n$ ,  $\text{rank}(\bar{\mathbf{D}}_{uu}) = m$ ; and

$\mathcal{C}_3$ : nonconstant kernel of  $\bar{\mathbf{D}}_{ua}$ .

With  $\mathcal{C}_1$  and  $\mathcal{C}_2$ , the generalized inversions of  $\bar{\mathbf{D}}_{aa}$ ,  $\bar{\mathbf{D}}_{uu}$ , and  $\bar{\mathbf{D}}_{au}$  exist, which are used to compute the auxiliary controls. We can select  $\bar{\mathbf{D}} = \bar{\mathbf{D}}^\top$  to ensure  $\bar{\mathbf{D}}_{au} = \bar{\mathbf{D}}_{ua}^\top$ . To see the requirement of  $\mathcal{C}_3$ , we rewrite  $\mathbf{q}_a = \sum_{i=1}^n p_{ai} \mathbf{v}_i$ . By Eq. (9), under the updated control  $\mathbf{v}^{\text{int}}$ ,  $\ddot{\mathbf{q}}_a = \sum_{i=1}^m \ddot{p}_{ai} \mathbf{v}_i + \sum_{i=m+1}^n \ddot{p}_{ai} \mathbf{v}_i$ , where  $\mathbf{v}_i$  is the  $i$ th column of  $\mathbf{V}$ . Note that the part  $\sum_{i=m+1}^n \ddot{p}_{ai} \mathbf{v}_i$  of  $\mathcal{S}_a$  dynamics is free of control if  $\mathbf{V}$  is constant. Although  $\mathbf{q}_a$  is stabilized on  $\mathbf{q}_a^e$ ,  $\mathbf{q}_a$  converges to  $\mathbf{q}_a^d$  only in an  $m$ -dimensional subspace and the other  $(n-m)$  dimensional motion uncontrolled. If the system is stable, the uncontrolled motion cannot be fixed in the configuration space throughout the entire control process. Therefore, a nonconstant kernel  $\bar{\mathbf{D}}_{ua}$  is needed.

Conditions  $\mathcal{C}_1$ – $\mathcal{C}_3$  provide sufficient nominal model selection criteria. The commonly used nominal model in Refs. [5] and [7] is  $\bar{\mathbf{D}}\ddot{\mathbf{q}} = \mathbf{B}\mathbf{u}$  with  $\bar{\mathbf{H}} = \mathbf{0}$ . The constant nominal model is used in Ref. [7] as the system is fully actuated. It is not difficult to satisfy the nominal model conditions in practice. First, the nonlinear term is canceled by feedback linearization, and  $\bar{\mathbf{H}} = \mathbf{0}$  can be used. Matrix  $\bar{\mathbf{D}}$  captures the robots' inertia property. The mass and length of robot links are usually available or can be measured. Meanwhile, the

dynamics coupling for revolute joints shows up in the inertia matrix as trigonometric functions of the relative joint angles. Therefore, the diagonal elements can be filled with mass or inertia estimates, and the off-diagonal entries can be constructed with trigonometric functions multiplying inertia constants.

#### 4 Gaussian Process-Enhanced External and Internal Convertible-Based Control

In this section, we propose two enhanced controllers using the GP model  $\mathcal{S}^{\text{gp}}$ , i.e., PEIC- and NEIC-based control. The PEIC-based control aims to eliminate uncontrolled motion under the original EIC-based control by reassigning the dynamics coupling, while the NEIC-based control directly manages the uncontrolled motion in a transformed space; see Figs. 1(b) and 1(c).

**4.1 Robust Auxiliary Control.** With  $\mathcal{S}^{\text{gp}}$ , we incorporate the variance from  $\mathcal{S}_a^{\text{gp}}$  into tracking control as

$$\hat{\mathbf{v}}^{\text{ext}} = \ddot{\mathbf{q}}_a^d - \hat{\mathbf{k}}_{p1}\mathbf{e}_a - \hat{\mathbf{k}}_{d1}\dot{\mathbf{e}}_a \quad (16)$$

where  $\hat{\mathbf{k}}_{p1} = \mathbf{k}_{p1} + k_{n1}\Sigma_a$  and  $\hat{\mathbf{k}}_{d1} = \mathbf{k}_{d1} + k_{n2}\Sigma_a$  are control gains with parameters  $k_{n1}, k_{n2} \geq 0$ . The variance of GP prediction  $\Sigma_a$  captures the uncertainty in robot dynamics and is updated online with sensor measurements.

Given the GP-based dynamics, the BEM is estimated by solving the following optimization problem rather than by inverting the system dynamics:

$$\hat{\mathbf{q}}_u^e = \arg \min_{\mathbf{q}_u} \|\Gamma_0(\mathbf{q}_u; \hat{\mathbf{v}}^{\text{ext}})\| \quad (17)$$

The balance control is then designed as

$$\hat{\mathbf{v}}_u^{\text{int}} = \ddot{\mathbf{q}}_u^e - \hat{\mathbf{k}}_{p2}\hat{\mathbf{e}}_u - \hat{\mathbf{k}}_{d2}\dot{\mathbf{e}}_u \quad (18)$$

where  $\hat{\mathbf{e}}_u = \mathbf{q}_u - \hat{\mathbf{q}}_u^e$  is the unactuated subsystem tracking error relative to the estimated BEM. Similar to  $\hat{\mathbf{k}}_{p2}, \hat{\mathbf{k}}_{d2}, \hat{\mathbf{k}}_{p2} = \mathbf{k}_{p2} + k_{n3}\Sigma_u$  and  $\hat{\mathbf{k}}_{d2} = \mathbf{k}_{d2} + k_{n4}\Sigma_u$  depend on  $\Sigma_u$  with the parameters by  $k_{n3}, k_{n4} \geq 0$ .

Let  $\Delta\mathbf{q}_u^e = \mathbf{q}_u^e - \hat{\mathbf{q}}_u^e$  denote the BEM estimation error, and the actual BEM is  $\mathbf{q}_u^e = \hat{\mathbf{q}}_u^e + \Delta\mathbf{q}_u^e$ . The control design based on actual BEM should be  $\hat{\mathbf{v}}_u^{\text{int}} = \ddot{\mathbf{q}}_u^e - \hat{\mathbf{k}}_{p2}\mathbf{e}_u - \hat{\mathbf{k}}_{d2}\mathbf{e}_u$ , and therefore we have

$$\mathbf{v}_u^{\text{int}} = \hat{\mathbf{v}}_u^{\text{int}} - \Delta\mathbf{v}_u^{\text{int}}$$

where  $\Delta\mathbf{v}_u^{\text{int}} = \Delta\ddot{\mathbf{q}}_u^e + \hat{\mathbf{k}}_{p2}\Delta\mathbf{q}_u^e + \hat{\mathbf{k}}_{d2}\Delta\dot{\mathbf{q}}_u^e$ . There are two sources causing the BEM estimation error. First, the learned dynamics  $\mathcal{S}_u^{\text{gp}}$  deviates from the actual one due to the prediction error  $\Delta_u$ . Therefore, the exact BEM solution using  $\mathcal{S}_u^{\text{gp}}$  deviates from that obtained in Eq. (4). Second, there exist differences between the BEM solved from  $\mathcal{S}_u^{\text{gp}}$  and that obtained from Eq. (17) due to the optimization algorithm. Given the bounded GP prediction error and limited optimization error, it is reasonable to assume that  $\Delta\mathbf{q}_u^e$  is bounded. Because of the bounded Gaussian kernel function, the GP prediction variances are also bounded, i.e.,

$$\|\Sigma_a(\mathbf{x})\| \leq (\sigma_a^{\max})^2, \|\Sigma_u(\mathbf{x})\| \leq (\sigma_u^{\max})^2 \quad (19)$$

where  $\sigma_a^{\max} = \max_i(\sigma_{fai}^2 + \vartheta_{ai}^2)^{1/2}$ ,  $\sigma_u^{\max} = \max_i(\sigma_{fui}^2 + \vartheta_{ui}^2)^{1/2}$ , and  $\sigma_f$  and  $\vartheta$  are the hyperparameters in each channel. Furthermore, we require the control gains to satisfy the following bounds:

$$k_{i1} \leq \lambda(\hat{\mathbf{k}}_{i1}) \leq k_{i3}, \quad k_{i2} \leq \lambda(\hat{\mathbf{k}}_{i2}) \leq k_{i4}, \quad i = p, d$$

for constants  $k_{pj}, k_{dj} > 0, j = 1, \dots, 4$ , where  $\lambda(\cdot)$  denotes the eigenvalue operator.

The control design should follow the guidelines: (1) the  $\mathbf{p}_{am}$  and  $\mathbf{q}_u$  dynamics are preserved (since they are stable under the original

EIC-based control), and (2) the uncontrolled motion (in  $\mathcal{S}_a^{\text{gp}}$ ) is either eliminated or under active control. The second requirement also implies that the motion of  $\mathbf{q}_u$  should depend on only  $m$  control inputs. To see this, solving  $\ddot{\mathbf{q}}_a$  from  $\mathcal{S}_a^{\text{gp}}$  and plugging it into  $\mathcal{S}_u^{\text{gp}}$  yields

$$(\bar{\mathbf{D}}_{uu} - \bar{\mathbf{D}}_{ua}\bar{\mathbf{D}}_{aa}^{-1}\bar{\mathbf{D}}_{au})\ddot{\mathbf{q}}_u + \mathbf{H}_u^{\text{gp}} - \bar{\mathbf{D}}_{ua}\bar{\mathbf{D}}_{aa}^{-1}\mathbf{H}_a^{\text{gp}} = -\bar{\mathbf{D}}_{ua}\bar{\mathbf{D}}_{aa}^{-1}\mathbf{u}$$

Note that  $\bar{\mathbf{D}}_{ua} \in \mathbb{R}^{m \times n}$ ,  $\bar{\mathbf{D}}_{aa}^{-1} \in \mathbb{R}^{n \times n}$ , and  $\mathbf{q}_u$  is overactuated given  $n = \dim(\mathbf{u}) \geq m = \dim(\mathbf{q}_u)$ . If  $\mathbf{q}_u$  depends on the same number of control inputs,  $(n - m)$  column vectors in  $\bar{\mathbf{D}}_{ua}\bar{\mathbf{D}}_{aa}^{-1}$  should be zero. Thus, the EIC-based control is applied between the same number of actuated and unactuated coordinates. The uncontrolled motion is avoided.

**4.2 Partial External and Internal Convertible-Based Control Design.** The control design  $\hat{\mathbf{v}}^{\text{int}}$  in Eq. (5) updates the input  $\mathbf{v}_n^{\text{ext}}$ , and  $\ddot{\mathbf{q}}_a$  acts as a virtual control to steer  $\mathbf{q}_u$  to  $\mathbf{q}_u^e$ . The  $\mathcal{S}_u$  dynamics is rewritten into

$$\ddot{\mathbf{q}}_u = -\mathbf{D}_{uu}^{-1}\mathbf{H}_u - \mathbf{D}_{uu}^{-1}\mathbf{D}_{ua}\ddot{\mathbf{q}}_a$$

where  $\mathbf{q}_u$  is overactuated with respect to  $\ddot{\mathbf{q}}_a$ . We instead reallocate the coupling between  $\mathbf{q}_a$  and  $\mathbf{q}_u$  and assign  $m$  control inputs for the unactuated subsystem; see Fig. 1(b). To achieve such a goal, we partition the actuated coordinates as  $\mathbf{q}_a = [\mathbf{q}_{aa}^T \mathbf{q}_{au}^T]^T$ ,  $\mathbf{q}_{au} \in \mathbb{R}^m$ ,  $\mathbf{q}_{aa} \in \mathbb{R}^{n-m}$ , and  $\mathbf{u} = [\mathbf{u}_a^T \mathbf{u}_u^T]^T$ . The  $\mathcal{S}_u^{\text{gp}}$  dynamics in Eq. (13) is rewritten as

$$\begin{bmatrix} \bar{\mathbf{D}}_{aa}^a & \bar{\mathbf{D}}_{aa}^u & \bar{\mathbf{D}}_{au}^a \\ \bar{\mathbf{D}}_{aa}^u & \bar{\mathbf{D}}_{aa}^u & \bar{\mathbf{D}}_{au}^u \\ \bar{\mathbf{D}}_{au}^a & \bar{\mathbf{D}}_{au}^u & \bar{\mathbf{D}}_{uu} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{q}}_{aa} \\ \ddot{\mathbf{q}}_{au} \\ \ddot{\mathbf{q}}_u \end{bmatrix} + \begin{bmatrix} \mathbf{H}_{aa}^{\text{gp}} \\ \mathbf{H}_{au}^{\text{gp}} \\ \mathbf{H}_u^{\text{gp}} \end{bmatrix} = \begin{bmatrix} \mathbf{u}_a \\ \mathbf{u}_u \\ \mathbf{0} \end{bmatrix} \quad (20)$$

where all block matrices are in proper dimensions. We rewrite Eq. (20) into three groups as

$$\mathcal{S}_{aa}^{\text{gp}} : \bar{\mathbf{D}}_{aa}^a \ddot{\mathbf{q}}_{aa} + \mathbf{H}_{aa}^a = \mathbf{u}_a \quad (21a)$$

$$\mathcal{S}_{au}^{\text{gp}} : \bar{\mathbf{D}}_{aa}^u \ddot{\mathbf{q}}_{au} + \bar{\mathbf{D}}_{au}^u \ddot{\mathbf{q}}_u + \mathbf{H}_{au}^u = \mathbf{u}_u \quad (21b)$$

$$\mathcal{S}_u^{\text{gp}} : \bar{\mathbf{D}}_{au}^u \ddot{\mathbf{q}}_{au} + \bar{\mathbf{D}}_{uu} \ddot{\mathbf{q}}_u + \mathbf{H}_{uu} = \mathbf{0} \quad (21c)$$

where  $\mathbf{H}_{aa}^a = \bar{\mathbf{D}}_{aa}^a \ddot{\mathbf{q}}_{aa} + \bar{\mathbf{D}}_{au}^a \ddot{\mathbf{q}}_u + \mathbf{H}_{aa}^{\text{gp}}$ ,  $\mathbf{H}_{au}^u = \bar{\mathbf{D}}_{aa}^u \ddot{\mathbf{q}}_{au} + \bar{\mathbf{D}}_{au}^u \ddot{\mathbf{q}}_u + \mathbf{H}_{au}^{\text{gp}}$ , and  $\mathbf{H}_{uu} = \bar{\mathbf{D}}_{au}^u \ddot{\mathbf{q}}_{au} + \mathbf{H}_{uu}^{\text{gp}}$ . Apparently,  $\mathcal{S}_u^{\text{gp}}$  is virtually independent of  $\mathcal{S}_{aa}^{\text{gp}}$ , and the dynamics coupling exists only between  $\mathcal{S}_u^{\text{gp}}$  and  $\mathcal{S}_{au}^{\text{gp}}$ .

Let  $\hat{\mathbf{v}}^{\text{ext}}$  in Eq. (16) be partitioned into  $\hat{\mathbf{v}}_a^{\text{ext}}$  and  $\hat{\mathbf{v}}_u^{\text{ext}}$  corresponding to  $\mathbf{q}_{aa}$  and  $\mathbf{q}_{au}$ , respectively.  $\hat{\mathbf{v}}_a^{\text{ext}}$  is directly applied to  $\mathcal{S}_{aa}^{\text{gp}}$ , and  $\hat{\mathbf{v}}_u^{\text{ext}}$  is updated for balance control purpose. As aforementioned, the condition to eliminate the uncontrolled motion in  $\mathcal{S}_a$  is that  $\mathbf{q}_u$  only depends on  $m$  inputs. The task of driving  $\mathbf{q}_u$  to  $\mathbf{q}_u^e$  is assigned to  $\mathbf{q}_{au}$  coordinates only. With this observation, the PEIC-based control takes the form of  $\hat{\mathbf{u}}^{\text{int}} = [\hat{\mathbf{u}}_a^T \hat{\mathbf{u}}_u^T]^T$  with

$$\hat{\mathbf{u}}_a = \bar{\mathbf{D}}_{aa}^a \hat{\mathbf{v}}_a^{\text{ext}} + \mathbf{H}_{aa}^a, \quad \hat{\mathbf{u}}_u = \bar{\mathbf{D}}_{au}^u \hat{\mathbf{v}}_u^{\text{int}} + \bar{\mathbf{D}}_{au}^u \ddot{\mathbf{q}}_u + \mathbf{H}_{au}^u \quad (22)$$

where  $\hat{\mathbf{v}}^{\text{int}} = -(\bar{\mathbf{D}}_{au}^u)^{-1}(\mathbf{H}_{uu} + \bar{\mathbf{D}}_{uu} \hat{\mathbf{v}}_u^{\text{int}})$ . Clearly, the unactuated subsystem only depends on  $\hat{\mathbf{u}}_u$  (or  $\mathbf{q}_{au}$ ) under the PEIC design as illustrated in Fig. 1(b). The following lemma presents the qualitative assessment of the PEIC-based control, and the proof is given in Appendix A2.

**LEMMA 3.** If conditions  $\mathcal{C}_1$  to  $\mathcal{C}_3$  are satisfied and  $\mathcal{S}_u^{\text{gp}}$  is stable under the EIC-based control design,  $\mathcal{S}_u^{\text{gp}}$  is stable under the PEIC-based control  $\hat{\mathbf{u}}^{\text{int}}$ .

**4.3 Null-Space External and Internal Convertible-Based Control Design.** Besides the PEIC-based control, we propose an alternative method in which the control input for  $\mathbf{p}_{an}$  is explicitly designed. Noting that  $\mathbf{p}_{an} \in \text{span}(\mathbf{V}_m)$  and  $\mathbf{p}_{an} \in \ker(\bar{\mathbf{D}}_{ua}) = \text{span}(\mathbf{V}_n)$ , subspaces  $\text{span}(\mathbf{V}_m)$  and  $\text{span}(\mathbf{V}_n)$  are orthogonal, and the motion of  $\mathbf{p}_{an}$  is independent of  $\mathbf{p}_{am}$ . Therefore, a compensation is designed in  $\text{span}(\mathbf{V}_m)$  for  $\mathbf{p}_{an}$ , which leaves the motion in  $\text{span}(\mathbf{V}_m)$  unchanged. Based on this observation, the NEIC-based control takes the form

$$\hat{\mathbf{u}}^{\text{int}} = \bar{\mathbf{D}}_{aa} \hat{\mathbf{v}}_a^{\text{int}} + \bar{\mathbf{D}}_{au} \ddot{\mathbf{q}}_u + \mathbf{H}_a^{\text{gp}} \quad (23)$$

where  $\hat{\mathbf{v}}_a^{\text{int}} = \hat{\mathbf{v}}^{\text{int}} + \tilde{\mathbf{v}}_{an}$ ,  $\tilde{\mathbf{v}}_{an} = \mathbf{V}_n \mathbf{v}_n$ ,  $\hat{\mathbf{v}}^{\text{int}} = -\bar{\mathbf{D}}_{aa}^+(\mathbf{H}_a^{\text{gp}} + \bar{\mathbf{D}}_{uu} \hat{\mathbf{v}}_u^{\text{int}})$ ,  $\mathbf{v}_n$  is the control design that drives  $\mathbf{p}_{ai}$  to  $\mathbf{p}_{ai}^d$ ,  $i = m + 1, \dots, n$ , and  $\mathbf{p}_a^d = \mathbf{Y}(\mathbf{q}_a^d)$  is transformed reference trajectory. The design of  $\mathbf{v}_n$  drives  $\mathbf{e}_a$  to the origin in  $\ker(\bar{\mathbf{D}}_{ua})$ . A straightforward yet effective design of  $\mathbf{v}_n$  can be  $\mathbf{v}_n = \alpha \hat{\mathbf{v}}_u^{\text{ext}}$ , where  $\alpha > 0$ . Compared to the PEIC-based control,  $\mathbf{p}_{an}$  plays the similar role of  $\mathbf{q}_{aa}$  coordinates. In the new coordinate, the  $\mathbf{q}_u$  is associated with  $\mathbf{p}_{an}$  only.

The following result gives the property of the NEIC-based control, and the proof is given in Appendix A3.

**LEMMA 4.** For  $\mathcal{S}$ , if  $\mathcal{S}_u^{\text{gp}}$  satisfies conditions  $\mathcal{C}_1$  to  $\mathcal{C}_3$  and  $\mathcal{S}_u^{\text{gp}}$  is stable under the original EIC-based control,  $\mathcal{S}_u^{\text{gp}}$  under the NEIC-based control  $\hat{\mathbf{v}}_a^{\text{int}}$  is also stable. Meanwhile,  $\mathcal{S}_u^{\text{gp}}$  is unchanged compared to that under the EIC-based control.

The proofs of Lemmas 3 and 4 show that the inputs  $\hat{\mathbf{u}}_a^{\text{int}}$  and  $\hat{\mathbf{u}}_u^{\text{int}}$  follow the control design guidelines. Both the PEIC- and NEIC-based controllers preserve the structured form of the EIC design. Figures 1(b) and 1(c) illustrate the overall flowchart of the PEIC- and NEIC-based control design, respectively. To take advantage of the EIC-based structure, we follow the design guideline to make sure that motion of unactuated coordinates only depends on  $m$  inputs in configuration space (PEIC-based control) or transformed space (NEIC-based control). The input  $\mathbf{v}_n^{\text{ext}}$  is re-used for uncontrolled motion under the NEIC-based control. The PEIC-based control assigns the balance task to a partial group of the actuated coordinates.

## 5 Control Stability Analysis

**5.1 Closed-Loop Dynamics.** To investigate the closed-loop dynamics, we consider the GP prediction error and the BEM estimation error. The GP prediction error in Eq. (14) is extended to  $\Delta_{aa}$ ,  $\Delta_{au}$ , and  $\Delta_u$  for  $\mathbf{q}_{aa}$ ,  $\mathbf{q}_{au}$ , and  $\mathbf{q}_u$  dynamics, respectively. Under the PEIC-based control, the dynamics of  $\mathcal{S}$  becomes

$$\begin{aligned} \ddot{\mathbf{q}}_{aa} &= \hat{\mathbf{v}}_a^{\text{ext}} - (\bar{\mathbf{D}}_{aa}^a)^{-1} \Delta_{aa} \\ \ddot{\mathbf{q}}_{au} &= -(\bar{\mathbf{D}}_{ua}^u)^{-1} (\mathbf{H}_{un} + \bar{\mathbf{D}}_{uu} \hat{\mathbf{v}}_u^{\text{int}}) - (\bar{\mathbf{D}}_{aa}^u)^{-1} \Delta_{au} \\ \ddot{\mathbf{q}}_u &= \hat{\mathbf{v}}_u^{\text{int}} - \bar{\mathbf{D}}_{uu}^{-1} [\Delta_u - \bar{\mathbf{D}}_{ua}^u (\bar{\mathbf{D}}_{aa}^u)^{-1} \Delta_{au}] \end{aligned}$$

Obtaining BEM with Eq. (17) under  $(\ddot{\mathbf{q}}_{aa}, \hat{\mathbf{v}}_u^{\text{ext}})$  is equivalent to inverting Eq. (21c). Thus,  $\hat{\mathbf{v}}_u^{\text{ext}} = -(\bar{\mathbf{D}}_{ua}^u)^{-1} \mathbf{H}_{un} |_{\mathbf{q}_{aa} = \hat{\mathbf{q}}_a^e, \ddot{\mathbf{q}}_u = \ddot{\mathbf{q}}_u^e = \mathbf{0}}$ . Substituting the above equation into the  $\mathbf{q}_{au}$  dynamics yields  $\ddot{\mathbf{q}}_{au} = \hat{\mathbf{v}}_u^{\text{ext}} + \mathbf{O}_{au}$ , where  $\mathbf{O}_{au} = -(\bar{\mathbf{D}}_{ua}^u)^{-1} \bar{\mathbf{D}}_{uu} \hat{\mathbf{v}}_u^{\text{int}} - (\bar{\mathbf{D}}_{aa}^u)^{-1} \Delta_{au} + \mathbf{o}_1$  and  $\mathbf{o}_1$  denotes the higher order terms.

Defining the total error  $\mathbf{e}_q = [\mathbf{e}_a^T \mathbf{e}_u^T]^T$  and  $\mathbf{e} = [\mathbf{e}_q^T \dot{\mathbf{e}}_q^T]^T$ , the closed-loop error dynamics becomes

$$\dot{\mathbf{e}} = \underbrace{\begin{bmatrix} \mathbf{0} & \mathbf{I}_{n+m} \\ -\hat{\mathbf{k}}_p & -\hat{\mathbf{k}}_d \end{bmatrix}}_A \begin{bmatrix} \mathbf{e}_q \\ \dot{\mathbf{e}}_q \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{0} \\ \mathbf{O}_{\text{tot}} \end{bmatrix}}_{\mathbf{o}_1} = \mathbf{Ae} + \mathbf{O}_1 \quad (24)$$

with  $\mathbf{O}_{\text{tot}} = [\mathbf{O}_a^T \mathbf{O}_u^T]^T$ ,  $\mathbf{O}_a = [\mathbf{O}_{aa}^T \mathbf{O}_{au}^T]^T$ ,  $\mathbf{O}_{aa} = -(\bar{\mathbf{D}}_{aa}^a)^{-1} \Delta_{aa}$ ,  $\mathbf{O}_u = -\bar{\mathbf{D}}_{uu}^{-1} [\Delta_u - \bar{\mathbf{D}}_{ua}^u (\bar{\mathbf{D}}_{aa}^u)^{-1} \Delta_{au}] - \Delta_{au}^{\text{int}}$ ,  $\hat{\mathbf{k}}_p = \text{diag}(\hat{\mathbf{k}}_{p1}, \hat{\mathbf{k}}_{p2})$ , and  $\hat{\mathbf{k}}_d = \text{diag}(\hat{\mathbf{k}}_{d1}, \hat{\mathbf{k}}_{d2})$ .

Because of bounded  $\bar{\mathbf{D}}$ , there exist constants  $0 < d_{a1}, d_{a2}, d_{u1}, d_{u2} < \infty$  such that  $d_{a1} \leq \|\bar{\mathbf{D}}_{aa}\| \leq d_{a2}$  and  $d_{u1} \leq \|\bar{\mathbf{D}}_{uu}\| \leq d_{u2}$ . The perturbation terms are further bounded as

$$\begin{aligned} \|\mathbf{O}_a\| &= \left\| - \begin{bmatrix} \mathbf{0} \\ (\bar{\mathbf{D}}_{ua}^u)^{-1} \bar{\mathbf{D}}_{uu} \hat{\mathbf{v}}_u^{\text{int}} \end{bmatrix} - (\bar{\mathbf{D}}_{aa}^a)^{-1} \Delta_{aa} + \begin{bmatrix} \mathbf{0} \\ \mathbf{o}_1 \end{bmatrix} \right\| \\ &\leq \frac{d_{u2}}{\sigma_1} \|\hat{\mathbf{v}}_u^{\text{int}}\| + \frac{1}{d_{a1}} \|\Delta_{aa}\| + \|\mathbf{o}_1\|, \end{aligned}$$

and

$$\begin{aligned} \|\mathbf{O}_u\| &= \left\| -\bar{\mathbf{D}}_{uu}^{-1} \left( \Delta_u - \bar{\mathbf{D}}_{ua}^u (\bar{\mathbf{D}}_{aa}^u)^{-1} \Delta_{au} \right) - \Delta_{au}^{\text{int}} \right\| \\ &\leq \frac{1}{d_{u1}} \|\Delta_u\| + \frac{\sigma_m}{d_{u1} d_{a1}} \|\Delta_{au}\| + \|\Delta_{au}^{\text{int}}\| \end{aligned}$$

The perturbation  $\mathbf{o}_1$  is due to approximation, and  $\Delta \mathbf{v}_u^{\text{int}}$  is the control difference by the BEM calculation with the GP prediction. They are both assumed to be affine with  $\mathbf{e}$ , i.e.,

$$\|\mathbf{o}_1\| \leq c_1 \|\mathbf{e}\| + c_2, \quad \|\Delta \mathbf{v}_u^{\text{int}}\| \leq c_3 \|\mathbf{e}\| + c_4 \quad (25)$$

with  $0 < c_i < \infty$ ,  $i = 1, \dots, 4$ . From Eq. (19), we have  $\|\boldsymbol{\kappa}_a^T \boldsymbol{\Sigma}_a^{1/2}\| \leq \sigma_a^{\max} \|\boldsymbol{\kappa}_a\|$  and  $\|\boldsymbol{\kappa}_u^T \boldsymbol{\Sigma}_u^{1/2}\| \leq \sigma_u^{\max} \|\boldsymbol{\kappa}_u\|$ . Thus, for  $0 < \eta = \eta_a \eta_u < 1$ , we can show that

$$\Pr\{\|\mathbf{O}_1\| \leq d_1 + d_2 \|\mathbf{e}\| + l_{u1} \|\boldsymbol{\kappa}_u\| + l_{a1} \|\boldsymbol{\kappa}_a\|\} \geq \eta \quad (26)$$

where  $d_1 = c_2 + (1 + (d_{u2}/\sigma_1)) c_4$ ,  $d_2 = c_1 + (d_{u2}/\sigma_1) c_3$ ,  $l_{a1} = ((\sigma_a^{\max} (d_{u1} + \sigma_m))/d_{u1} d_{a1})$ , and  $l_{u1} = \sigma_u^{\max}/d_{u1}$ .

To obtain the closed-loop dynamics under the NEIC-based control, plugging the NEIC-based control into  $\mathcal{S}_u^{\text{gp}}$ , we obtain

$$\ddot{\mathbf{p}}_{am} = -\Lambda_m^{-1} \mathbf{U}^T (\mathbf{H}_u^{\text{gp}} + \bar{\mathbf{D}}_{uu} \hat{\mathbf{v}}_u^{\text{int}}) - \Lambda_m^{-1} \mathbf{U}^T \Delta_u - \mathbf{V}_m^T \bar{\mathbf{D}}_{aa}^{-1} \Delta_a \quad (27a)$$

$$\ddot{\mathbf{p}}_{an} = \mathbf{v}_n^{\text{ext}} - \mathbf{V}_n^T \bar{\mathbf{D}}_{aa}^{-1} \Delta_a \quad (27b)$$

$$\ddot{\mathbf{q}}_u = \hat{\mathbf{v}}_u^{\text{int}} - \bar{\mathbf{D}}_{uu}^{-1} (\Delta_u - \bar{\mathbf{D}}_{ua} \bar{\mathbf{D}}_{aa}^{-1} \Delta_a) \quad (27c)$$

To obtain the error dynamics, we take advantage of the definition of BEM. From Eq. (A3), we have  $\mathbf{v}_a^{\text{ext}} = -\Lambda_m^{-1} \mathbf{U}^T \mathbf{H}_u^{\text{gp}} \Big|_{\mathbf{q}_u = \hat{\mathbf{q}}_u^e, \dot{\mathbf{q}}_u = \ddot{\mathbf{q}}_u^e} = \mathbf{0}$ . Then, we rewrite Eq. (27a) into

$$\begin{aligned} \ddot{\mathbf{p}}_{am} &= -\Lambda_m^{-1} \mathbf{U}^T \mathbf{H}_u^{\text{gp}} \Big|_{\mathbf{q}_u = \hat{\mathbf{q}}_u^e, \dot{\mathbf{q}}_u = \ddot{\mathbf{q}}_u^e} + \mathbf{o}_2 - \Lambda_m^{-1} \mathbf{U}^T \bar{\mathbf{D}}_{uu} \hat{\mathbf{v}}_u^{\text{int}} \\ &\quad - \Lambda_m^{-1} \mathbf{U}^T \Delta_u - \mathbf{V}_m^T \bar{\mathbf{D}}_{aa}^{-1} \Delta_a = \mathbf{v}_m^{\text{ext}} + \mathbf{O}_m \end{aligned} \quad (28)$$

where  $\mathbf{o}_2$  is the residual that contains higher order terms.  $\mathbf{O}_{am} = \mathbf{o}_2 - \Lambda_m^{-1} \mathbf{U}^T \bar{\mathbf{D}}_{uu} \hat{\mathbf{v}}_u^{\text{int}} - \Lambda_m^{-1} \mathbf{U}^T \Delta_u - \mathbf{V}_m^T \bar{\mathbf{D}}_{aa}^{-1} \Delta_a$  denotes the total perturbations.

The  $\mathcal{S}_u^{\text{gp}}$  dynamics keeps the same form as that in the PEIC-based control. We write the error dynamics under the NEIC-based control as

$$\ddot{\mathbf{e}}_{am} = -\hat{\mathbf{k}}_{p1} \mathbf{e}_{am} - \hat{\mathbf{k}}_{d1} \mathbf{e}_{am} + \mathbf{O}_{am} \quad (29a)$$

$$\ddot{\mathbf{e}}_{an} = -\hat{\mathbf{k}}_{p1} \mathbf{e}_{an} - \hat{\mathbf{k}}_{d1} \mathbf{e}_{an} + \mathbf{O}_{an} \quad (29b)$$

$$\ddot{\mathbf{e}}_u = -\hat{\mathbf{k}}_{p2} \mathbf{e}_u - \hat{\mathbf{k}}_{d2} \mathbf{e}_u + \mathbf{O}_u \quad (29c)$$

where  $\mathbf{e}_{am} = \mathbf{p}_{am} - \mathbf{p}_{am}^d$ ,  $\mathbf{e}_{an} = \mathbf{p}_{an} - \mathbf{p}_{an}^d$ , and  $\mathbf{O}_{an} = -\mathbf{V}_n^T \bar{\mathbf{D}}_{aa}^{-1} \Delta_a$ . Applying inverse mapping  $\boldsymbol{\Upsilon}^{-1}$  to Eqs. (29a) and (29b), the error dynamics in  $\mathbf{q}$  is obtained as

$$\mathcal{S}_{e,\text{NEIC}} : \dot{\mathbf{e}} = \mathbf{A} \mathbf{e} + \mathbf{O}_2 \quad (30)$$

where  $\mathbf{O}_2$  is the transformed perturbations of  $[\mathbf{O}_{an}^T \mathbf{O}_{am}^T \mathbf{O}_u^T]^T$ . Following the same steps to obtain Eq. (26), we have

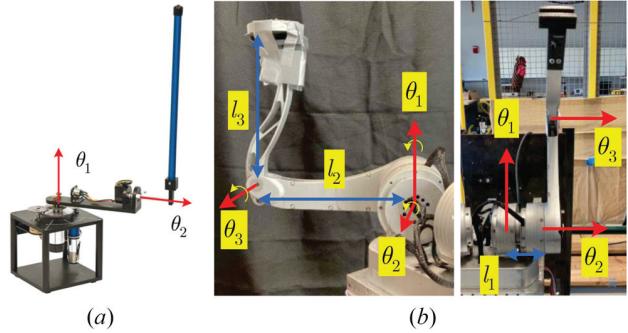
$$\Pr\{\|\mathbf{O}_2\| \leq d_1 + d_2 \|\mathbf{e}\| + l_{u2} \|\boldsymbol{\kappa}_u\| + l_{a2} \|\boldsymbol{\kappa}_a\|\} \geq \eta \quad (31)$$

where  $l_{u2} = \sigma_{u,\max}((\sigma_1 + d_{u1})/\sigma_1 d_{u1})$ , and  $l_{a2} = \sigma_{a,\max}((\sigma_m + d_{u1})/d_{a1} d_{u1})$ .

**5.2 Stability Results.** To show the stability, we consider the Lyapunov function candidate  $V = \mathbf{e}^T \mathbf{P} \mathbf{e} \geq 0$ , where positive definite matrix  $\mathbf{P} = \mathbf{P}^T$  is the solution of

$$\mathbf{A}_0^T \mathbf{P} + \mathbf{P} \mathbf{A}_0 + \mathbf{Q} = \mathbf{0}, \quad \mathbf{A}_0 = \begin{bmatrix} \mathbf{0} & \mathbf{I}_{n+m} \\ -\mathbf{k}_p & -\mathbf{k}_d \end{bmatrix} \quad (32)$$

for given positive definite matrix  $\mathbf{Q} = \mathbf{Q}^T$ , where  $\mathbf{A}_0$  is the constant part of  $\mathbf{A}$  in Eq. (24) and does not depend on variances  $\boldsymbol{\Sigma}_a$  or  $\boldsymbol{\Sigma}_u$ .  $\mathbf{k}_p = \text{diag}(\mathbf{k}_{p1}, \mathbf{k}_{p2})$  and  $\mathbf{k}_d = \text{diag}(\mathbf{k}_{d1}, \mathbf{k}_{d2})$ .



**Fig. 2** (a) A Furuta pendulum. The base link joint  $\theta_1$  is actuated, and the inverted pendulum link joint  $\theta_2$  is unactuated. (b) A three-link inverted pendulum with actuated joints  $\theta_1$  and  $\theta_2$  and unactuated joint  $\theta_3$ . The rotation axis of link 3 is perpendicular to that of link 2 and link 3.

We denote the corresponding Lyapunov function candidates for the NEIC- and PEIC-based controls as  $V_1$  and  $V_2$ , respectively. The stability results are summarized as follows with the proof given in Appendix A4.

**THEOREM 1.** For robot dynamics (2), using the GP-based model (13) that satisfies conditions C1–C3, under the PEIC- and NEIC-based control, the Lyapunov function under each controller satisfies

$$\Pr\{\dot{V}_i \leq -\gamma_i V_i + \rho_i + \varpi_i\} \geq \eta, \quad i = 1, 2 \quad (33)$$

and the error  $\mathbf{e}$  converges to a small ball around the origin, where  $\gamma_i$  is the convergence rate,  $\rho_i$  and  $\varpi_i$  are the perturbation terms, and  $0 < \eta = \eta_a \eta_u < 1$ .

## 6 Experimental Results

Two inverted pendulum platforms are used to conduct experiments to validate the control design. The results from each platform demonstrate different aspects of the control design.<sup>2</sup>

### 6.1 Two Degree-of-Freedom Rotary Inverted Pendulum

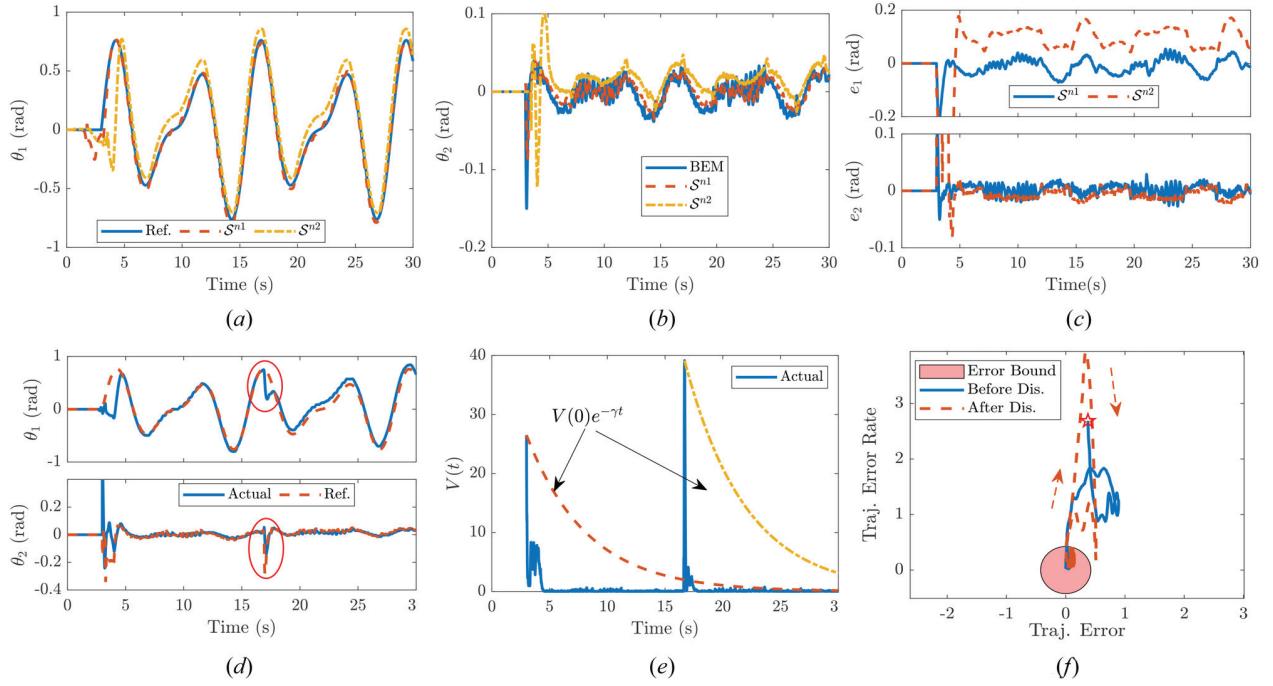
Figure 2(a) shows a 2DOF rotary inverted pendulum that was fabricated by Quanser Inc., Markham, ON, Canada. The base joint ( $\theta_1$ ) is actuated by a DC motor, and the inverted pendulum joint ( $\theta_2$ ) is unactuated, i.e.,  $n = m = 1$ . We use this platform to illustrate the original EIC-based control and also compare the performance under different nominal models and controllers. The robot dynamic model is given in Ref. [27] and is also found in Appendix B1.

Since  $m = n = 1$ , there is no uncontrolled motion when the original EIC-based control is applied. Therefore, either a constant or time-varying nominal model would work for the GP-based learning control. We created the following two nominal models:

$$\begin{aligned} \mathcal{S}^{n_1} : \bar{\mathbf{D}}_1 &= \frac{1}{100} \begin{bmatrix} 5 & -2c_2 \\ -2c_2 & 2 \end{bmatrix}, \quad \bar{\mathbf{H}}_1 = \begin{bmatrix} 0 \\ -s_2 \end{bmatrix} \\ \mathcal{S}^{n_2} : \bar{\mathbf{D}}_2 &= \frac{1}{100} \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \bar{\mathbf{H}}_2 = \mathbf{0} \end{aligned}$$

where  $c_i = \cos \theta_i$ ,  $s_i = \sin \theta_i$  for angle  $\theta_i$ ,  $i = 1, 2$ . The training data were sampled and obtained by applying control input  $u = \mathbf{k}^T [\theta_1 - \theta_1^t \ \theta_2 - \theta_2^t \ \dot{\theta}_1 - \dot{\theta}_1^t \ \dot{\theta}_2]^T$ , where  $\mathbf{k} \in \mathbb{R}^{4 \times 1}$  and  $\theta_1^t$  was the combination of sinusoidal waves with different amplitudes and frequencies. We chose this input to excite the system, and the gain  $\mathbf{k}$  was selected without the need to balance the platform. It is difficult to guarantee that the system is fully excited. However, we changed the frequency of sinusoidal waves and obtained the motion data around the target trajectory.

<sup>2</sup>The video of the experiment is available at <https://www.youtube.com/watch?v=ZOYb0UW3KS8>



**Fig. 3 Experiment results with guaranteed performance:** (a) arm rotation angle, (b) pendulum rotation angle, (c) tracking control error under GP-based control, (d) pendulum motion profile, (e) profile of Lyapunov function, and (f) trajectory error motion. At  $t = 17$  s, an impact disturbance is applied. The dashed arrow in (f) indicates the direction in which the error grows after disturbance is applied.

We trained the GP regression models using a total of 500 data points randomly selected from a large dataset. We designed the control gains as  $\hat{k}_{p1} = 10 + 50\Sigma_a$ ,  $\hat{k}_{d1} = 3 + 10\Sigma_a$ ,  $\hat{k}_{p2} = 1000 + 500\Sigma_u$ , and  $\hat{k}_{d2} = 100 + 200\Sigma_u$ . The variances  $\Sigma_a$  and  $\Sigma_u$  were updated online with new measurements in real time. The reference trajectory was  $\theta_1^d = 0.5 \sin t + 0.3 \sin 1.5t$  rad. The control was implemented at 400 Hz in MATLAB/SIMULINK real-time system. Both the velocity and acceleration are needed for control design and GP training and prediction. To reduce the influence of measurement noise on control design, BEM estimation, and GP agent training, a sliding window was used to filter the velocity measurement online. The acceleration was obtained through real-time differentiation. The same technique was also used for the three-link inverted pendulum in Sec. 6.2.

Figures 3(a) and 3(b) show the tracking of  $\theta_1$  and balance of  $\theta_2$  under the EIC-based control. With either  $S^{n_1}$  or  $S^{n_2}$ , the base link joint  $\theta_1$  closely followed the reference trajectory  $\theta_1^d$ , and the pendulum link joint  $\theta_2$  was stabilized around its equilibrium  $\theta_2^e$  as well. The tracking error was reduced further, and the pendulum closely followed the small variation under  $S^{n_1}$ . With  $S^{n_2}$ , the tracking errors became large when the base link changed rotation direction; see Fig. 3(c) at  $t = 10, 17$ , and  $22$  s. Both the time-varying and constant nominal models worked for the EIC-based learning control.

Table 1 further lists the tracking errors (mean and one standard deviation) under both GP models. For comparison purposes, we also conducted additional experiments to implement the original EIC-based control and the GP-based MPC design in Ref. [4]. The tracking and balance errors under the EIC-based learning control with model  $S^{n_1}$  are the smallest. In particular, with the time-varying model  $S^{n_1}$ , the mean values of tracking errors  $e_1$  and  $e_2$  were reduced

by 75% and 65%, respectively, in comparison with those under the original EIC-based control. Compared with the MPC method in Ref. [4], the tracking errors with nominal model  $S^{n_2}$  are at the same level.

Figure 3(d) shows the control performance with nominal model  $S^{n_1}$  under disturbance. At  $t = 17$  s, an impact disturbance (by manually pushing the pendulum link) was applied, and the joint angles changed rapidly with  $\Delta\theta_1 = 0.7$  rad and  $\Delta\theta_2 = 0.3$  rad. The control gains increased ( $\hat{k}_{p2} = 1215$ ,  $\hat{k}_{d2} = 143$ ) to respond to the disturbance. As a result, the pendulum motion tracked the BEM closely and maintained the pendulum balance after the impact disturbance. Figure 3(e) shows the calculated Lyapunov function candidate  $V(t)$  and its envelope (i.e.,  $V(t) = V(0)e^{-\gamma t}$ ,  $\gamma = 0.1898$ ) during the experiment. Figure 3(f) shows the error trajectory in the  $\|e_q\| - \|\dot{e}_q\|$  plane. The solid/dashed line shows the error trajectory before/after impact disturbance. The tracking error converged quickly into the error bound. After the disturbance was applied at  $t = 17$  s, both the Lyapunov function and errors grew dramatically. As the control gains increased, the errors quickly converged back to the estimated bound again.

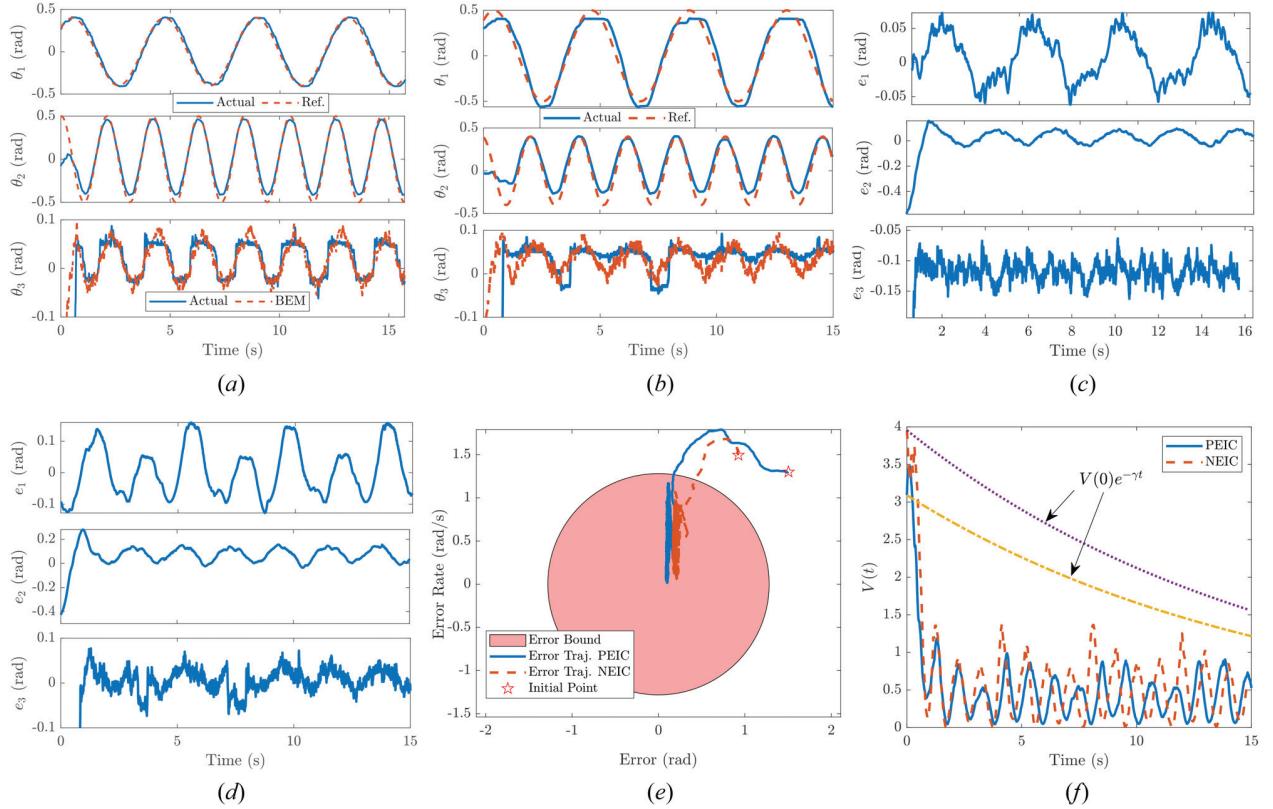
**6.2 Three Degree-of-Freedom Rotary Inverted Pendulum.** Figure 2(b) for a 3DOF inverted pendulum with two actuated joints ( $\theta_1$  and  $\theta_2$ ) and one unactuated joint ( $\theta_3$ ), namely,  $n = 2, m = 1$ . The physical model of the robot dynamics was obtained using the Lagrangian method and is given in Appendix B2. All controllers were implemented at an updating frequency of 200 Hz through the Robot Operating System. The time-varying nominal model was selected as

$$\bar{D} = \begin{bmatrix} 0.15 & 0.025c_2 & 0.025c_3 \\ 0.025c_2 & 0.15 & 0.05c_{2-3} \\ 0.025c_3 & 0.05c_{2-3} & 0.1 \end{bmatrix}, \quad \bar{H} = \begin{bmatrix} 0 \\ 0.2c_2 \\ 0.1s_3 \end{bmatrix}$$

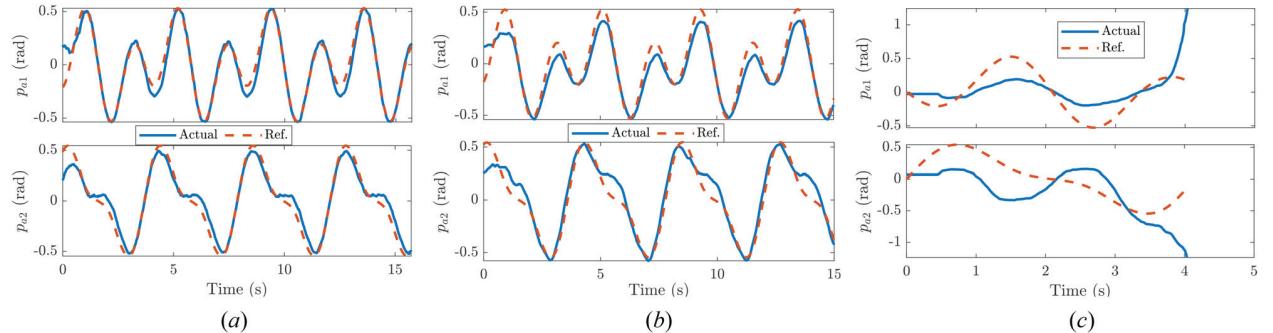
where  $c_{i\pm j} = \cos(\theta_i \pm \theta_j)$ . The control gains were  $\hat{k}_{p1} = 15I_2 + 20\Sigma_a$ ,  $\hat{k}_{d1} = 3I_2 + 10\Sigma_a$ ,  $\hat{k}_{p2} = 25 + 20\Sigma_u$ , and  $\hat{k}_{d2} = 5.5 + 10\Sigma_u$ , where GP variances  $\Sigma_a$  and  $\Sigma_u$  were updated online in real-time. The

**Table 1 Tracking errors comparison under various controllers ( $\times 10^{-1}$  rad)**

	$S^{n_1}$	$S^{n_2}$	GP-based MPC [4]	Physical EIC
$ e_1 $	$0.24 \pm 0.17$	$0.96 \pm 0.34$	$0.87 \pm 0.52$	$1.09 \pm 0.40$
$ e_2 $	$0.09 \pm 0.05$	$0.09 \pm 0.39$	$0.07 \pm 0.06$	$0.26 \pm 0.15$



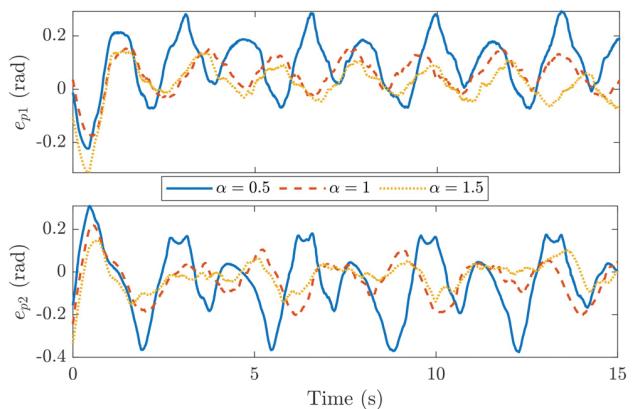
**Fig. 4** Experiment results with the 3DOF inverted pendulum: (a) and (b) Motion profiles under the PEIC- and NEIC-based control, (c) and (d) tracking errors under the PEIC- and NEIC-based control, (e) error trajectory in the  $\|e_q\|$ - $\|\dot{e}_q\|$  plane, and (f) comparison of the estimated Lyapunov function profile with the actual one



**Fig. 5** Motion profiles of robotic leg in transformed space under (a) PEIC-based control, (b) NEIC-based control, and (c) EIC-based control

reference trajectory was chosen as  $\theta_1^d = 0.5 \sin 1.5t$  and  $\theta_2^d = 0.4 \sin 3t$  rad.

For the PEIC-based control, we chose  $q_{aa} = \theta_1$  and  $q_{au} = \theta_2$ , and the NEIC-based control was  $v_n = \hat{v}_n^{\text{ext}}$ . Figure 4 shows the experimental results under the PEIC- and NEIC-based control. Under both controllers, the actuated joints ( $\theta_1$  and  $\theta_2$ ) followed the given reference trajectories ( $\theta_1^d$  and  $\theta_2^d$ ) closely, and the unactuated joint ( $\theta_3$ ) was balanced around the BEM ( $\theta_3^d$ ) as shown in Figs. 4(a) and 4(b). The pendulum link motion displayed a similar pattern for both controllers. However, the tracking error  $e_1$  under the PEIC-based control (i.e., from  $-0.05$  to  $0.05$  rad) was much smaller than that under the NEIC-based control (i.e., from  $-0.15$  to  $0.15$  rad); see Figs. 4(c) and 4(d). The balance task in the PEIC-based control was assigned to joint  $\theta_2$ , and joint  $\theta_1$  is viewed as virtually independent of  $\theta_2$  and  $\theta_3$ . Joint  $\theta_1$  achieved almost-perfect tracking control regardless of the errors for  $\theta_2$  and  $\theta_3$ . The compensation effect in the null space appeared in the entire configuration space, and any



**Fig. 6** The tracking errors in coordinate  $p_a$  under the NEIC-based control with various  $\alpha$  values

**Table 2 Statistical analysis of tracking performance (mean and standard deviation for errors) under different controllers**

	$ e_1 $ (rad)	$ e_2 $ (rad)	$ e_3 $ (rad)	$\ e\ $	$\int u^T u dt$
PEIC (GP)	$0.0302 \pm 0.0178$	$0.0566 \pm 0.0685$	$0.1182 \pm 0.0160$	$0.1343 \pm 0.0166$	5.7659
NEIC (GP, $\alpha = 0.5$ )	$0.1395 \pm 0.0946$	$0.1166 \pm 0.0512$	$0.0303 \pm 0.0209$	$0.2001 \pm 0.0770$	5.9022
NEIC (GP, $\alpha = 1.0$ )	$0.0651 \pm 0.0416$	$0.0756 \pm 0.0481$	$0.0195 \pm 0.0152$	$0.1101 \pm 0.0499$	5.7089
NEIC (GP, $\alpha = 1.5$ )	$0.0376 \pm 0.0302$	$0.0792 \pm 0.0482$	$0.0207 \pm 0.0169$	$0.0972 \pm 0.0470$	5.7305
PEIC (model)	$0.2168 \pm 0.1165$	$0.2398 \pm 0.1649$	$0.0179 \pm 0.0140$	$0.3587 \pm 0.1307$	5.7978
NEIC (model, $\alpha = 1.0$ )	$0.1374 \pm 0.0922$	$0.1237 \pm 0.0597$	$0.0455 \pm 0.0385$	$0.2095 \pm 0.0769$	5.8452

motion error in the unactuated joints affected the motion of all actuated joints. Similar to the previous example, Fig. 4(e) shows the error trajectory profile in the  $\|e_q\|$ - $\|\dot{e}_q\|$  plane. Figure 4(f) shows the Lyapunov function profiles under the PEIC- and NEIC-based controls.

Figure 5 shows the motion of the actuated coordinate in the transformed coordinate  $p_a$  under various controllers. Under the PEIC- and NEIC-based controls, the  $p_a$  variables followed the reference profile  $p_a^d$  as shown in Figs. 5(a) and 5(b). Figure 5(c) shows the motion profile under the original EIC-based control. In the first 2 s, joint  $\theta_3$  followed the BEM under the EIC-based control, and  $p_{a1}$  coordinates displayed a similar motion pattern. However,  $p_{a2}$  coordinate showed diverge behavior and led to a failure completely. Therefore, as analyzed previously, the system became unstable under the EIC-based control though conditions  $C_1$  to  $C_3$  were satisfied.

In NEIC-based control,  $v_n$  drives the uncontrolled motion variable to its reference trajectory. To further reduce the tracking error, we can increase  $\alpha$  values. Figure 6 shows the experiment results of the  $p_a$  error profiles under various  $\alpha$  values varying from 0.5 to 1.5. With a large  $\alpha$  value, the tracking error of the actuated coordinates was reduced. Table 2 further lists the steady-state errors (in joint angles) under the NEIC-based control with various  $\alpha$  values, the PEIC-based control and the physical model-based control design. Under the NEIC-based control with  $\alpha = 0.5$ , the system was stabilized; when increasing  $\alpha$  values to 1 and 1.5, the mean tracking errors were reduced 50% and 70% for  $\theta_1$ , respectively, and 40% for  $\theta_2$ . Since control input  $v_n$  did not affect the balance task of the unactuated subsystem, the tracking errors for  $\theta_3$  maintained the same level. It is of interest that the control effort (i.e., last column in Table 2) only shows a slight increase with large  $\alpha$  values.

**6.3 Discussion.** For the rotary pendulum example, we have  $n = m$ , and the null space  $\ker(D_{au})$  vanishes. The compensation effect is no longer needed by the NEIC-based control, i.e.,  $\tilde{v}_a^{\text{int}} = \tilde{v}^{\text{int}}$  and  $\tilde{u}^{\text{int}} = \tilde{D}_{aa}\tilde{v}_a^{\text{int}} + \tilde{D}_{au}\tilde{q}_u + \tilde{H}_a^{\text{gp}} = u^{\text{int}}$ . In this case, the PEIC- and NEIC-based controls are degenerated to the EIC-based control. For the 3DOF inverted pendulum, the control inputs  $u_1$  and  $u_2$  act on  $\theta_3$  joints through  $\ddot{\theta}_1$  and  $\ddot{\theta}_2$ . Therefore, as shown in Lemma 1, the uncontrolled motion exists since all controls show up in  $\mathcal{S}_u$  dynamics. This observation explains why the original EIC-based control failed to balance the three-link inverted pendulum. If the  $\mathcal{S}_u$  dynamics is related to  $m$  control inputs (through  $\ddot{q}_a$ ) for  $n > m$  such as the bikebot dynamics in Refs. [4] and [25], only  $m$  external controls were updated, and the EIC-based control worked well without any uncontrolled motion.

For the PEIC-based control, the robot dynamics were partitioned into  $\mathcal{S}^{\text{gp}} = \{\mathcal{S}_{aa}^{\text{gp}}, \{\mathcal{S}_{au}^{\text{gp}}, \mathcal{S}_u^{\text{gp}}\}\}$ , which contains a fully actuated system  $\mathcal{S}_{aa}^{\text{gp}}$  and a reduced-order underactuated system  $\{\mathcal{S}_{au}^{\text{gp}}, \mathcal{S}_u^{\text{gp}}\}$ . The EIC-based control is applied to  $\mathcal{S}_{au}^{\text{gp}}$  and  $\mathcal{S}_u^{\text{gp}}$  only. The dynamics of  $q_u$  in general does not depend on any specific  $m$  actuated coordinates, since the mapping  $\Upsilon$  is time-varying across different control cycles. In the NEIC-based control design,  $p_{am}$  and  $q_u$  become an underactuated subsystem, and  $p_{an}$  is fully actuated.

In practice, no specific rules are defined to select  $q_{au}$  out of  $q_a$  coordinates, and therefore, there are a total of  $C_n^m = n!/(m!(n-m)!)$  options to select different coordinates. We take advantage of such a property to optimize tracking performance for selected coordinates. In the 3DOF pendulum case, we assigned the balance task of  $\theta_3$  to  $\theta_2$  motion. The length of link 1 was only 0.09 m and was much shorter than the length of link 2 (0.23 m). The coupling effect between  $\theta_2$  and  $\theta_3$  was much stronger than that between  $\theta_1$  and  $\theta_3$ ; see  $D_{13}$  and  $D_{23}$  in Appendix B2. Thus, it was efficient to use the motion of  $\theta_2$  as a virtual control input to balance  $\theta_3$ . When implementing the PEIC-based controller with  $q_{au} = \theta_1$ , the system cannot achieve the desired performance and becomes unstable. We also implemented the proposed controller with the physical model. The control errors are listed in Table 2. Compared with the learning-based controllers, the model-based control resulted in larger errors. Since the mechanical frictions and other unstructured effects were not considered, the physical model might not capture and reflect the accurate robot dynamics. The results confirmed the advantages of the proposed learning-based control approaches.

The unique feature of the proposed control lies in integration of the robot's inherent dynamics property (EIC structure) and the GP-based model learning, compared with other learning-based control approach [18,22]. By integrating physics knowledge into model learning, we identified the conditions for nominal model selection. The overall model learning and control design framework forms a white-box-like, physics knowledge involved control, which differs from the reinforcement learning-based policy search approach [18]. The solution also has the potential to further incorporate the bounded GP prediction error for a robust control [4].

## 7 Conclusion

This paper presented a new learning-based modeling and control framework for underactuated balance robots. The proposed design was an extension and improvement of the EIC-based control with GP-enabled robot dynamics. The proposed new robot controllers preserved the structural design of the original EIC-based control and achieved both tracking and balance tasks. The PEIC-based control reshaped the coupling between the actuated and unactuated coordinates. The robot dynamics was transferred into a fully actuated subsystem and one reduced-order underactuated balance subsystem. The NEIC-based control compensated for uncontrolled motion in a subspace. We validated and demonstrated the new control design on two experimental platforms and confirmed that stability and balance were guaranteed. The comparison with the physical model-based EIC control and the MPC design confirmed superior performance in terms of the error bound. Extension of the GP-based learning control design for highly underactuated balance robots is one of the ongoing research directions.

## Funding Data

- U.S. National Science Foundation (NSF) (Award No. CNS-1932370; Funder ID: 10.13039/100000001).

## Data Availability Statement

No data, models, or code were generated or used for this paper.

## Nomenclature

$\mathbf{e}_a, \mathbf{e}_u, \mathbf{e}$	tracking, balance, and overall errors
$\mathbf{p}_a, \mathbf{v}^{\text{ext}}$	transformed $\mathbf{q}_a$ and $\mathbf{v}^{\text{ext}}$ in $p$ coordinates
$\mathbf{p}_{am}, \mathbf{p}_{an}$	controlled and uncontrolled coordinates
$\mathbf{q}_a, \mathbf{q}_u$	coordinates for actuated and unactuated subsystems
$\mathbf{q}_{aa}, \mathbf{q}_{au}$	partitioned actuated coordinates in $(n-m)$ - and $m$ -dimensions
$\mathbf{q}_u^e, \hat{\mathbf{q}}_u^e$	actual and estimated BEMs
$\mathcal{S}$	robot dynamics
$\mathcal{S}^n, \mathcal{S}^{\text{gp}}$	nominal and GP-based robot dynamics
$\mathbf{u}^{\text{int}}, \hat{\mathbf{u}}^{\text{int}}, \tilde{\mathbf{u}}^{\text{int}}$	EIC-, PEIC-, NEIC-based control inputs
$\mathbf{v}^{\text{ext}}, \mathbf{v}^{\text{int}}$	trajectory tracking and balanced-embedded control inputs
$\mathbf{v}_u^{\text{int}}$	BEM stabilization control input
$\hat{\mathbf{v}}_a^{\text{ext}}, \hat{\mathbf{v}}_u^{\text{ext}}$	trajectory tracking control inputs for $\mathbf{q}_{aa}$ and $\mathbf{q}_{au}$
$\hat{\mathbf{v}}_a^{\text{int}}$	control input for $\mathbf{p}_{am}$
$\gamma, r$	convergence rate and error bound
$\Delta_a, \Delta_u$	estimation errors of actuated and unactuated dynamics

## Appendix A: Proofs

**A1 Proof of Lemma 1.** The system dynamics  $\mathcal{S}$  under control  $\mathbf{u}^{\text{ext}}$  is

$$\ddot{\mathbf{q}}_a = \mathbf{v}^{\text{ext}}, \quad \ddot{\mathbf{q}}_u = -\mathbf{D}_{uu}^{-1}(\mathbf{D}_{ua}\mathbf{v}^{\text{ext}} + \mathbf{H}_u) \quad (\text{A1})$$

When  $\text{rank}(\mathbf{D}_{au}) = m$  holds for  $\mathbf{q}$ , the SVD in Eq. (7) exists and all  $m$  singular values are great than zero, i.e.,  $\sigma_i > 0$ . Thus,  $\ker(\mathbf{D}_{au}) = \mathbf{V}_n$  contains  $(n-m)$  column vectors. Plugging Eq. (7) into Eq. (A1) and considering the coordinate transformation, we obtain

$$\ddot{\mathbf{p}}_a = \mathbf{v}^{\text{ext}}, \quad \ddot{\mathbf{q}}_u = -\mathbf{D}_{uu}^{-1}(\mathbf{U}\Lambda_m\mathbf{v}_m^{\text{ext}} + \mathbf{H}_u) \quad (\text{A2})$$

where  $\mathbf{U}\Lambda\mathbf{V}^T\mathbf{v}^{\text{ext}} = \mathbf{U}\Lambda_m\mathbf{v}_m^{\text{ext}}$  is used based on the fact that  $\Lambda \in \mathbb{R}^{m \times m}$  is a rectangular diagonal matrix.

Given the definition of  $\mathcal{E}$ ,  $\mathbf{q}_u^e$  is obtained by solving the algebraic equation  $\Gamma_0(\mathbf{q}_u; \mathbf{v}^{\text{ext}}) = \mathbf{0}$ . We substitute  $\mathbf{D}_{ua}(\mathbf{q}_u^e)$  with  $\mathbf{D}_{ua}(\mathbf{q}_u)$  in  $\Gamma_0$ , and therefore, using Eq. (7),  $\Gamma_0 = \mathbf{0}$  is rewritten into

$$\mathbf{A}_m\mathbf{v}_m^{\text{ext}} + \mathbf{U}^T\mathbf{H}_u^{\text{gp}}|_{\mathbf{q}_u = \mathbf{q}_u^e, \dot{\mathbf{q}}_u = \ddot{\mathbf{q}}_u = \mathbf{0}} = \mathbf{0} \quad (\text{A3})$$

The BEM  $\mathcal{E}$  depends only on  $\mathbf{v}_m^{\text{ext}}$ , that is, the control effect in  $\ker(\mathbf{D}_{ua})$  is not used when obtaining the BEM.

Furthermore, since all controls show up in  $\mathcal{S}_u$  dynamics, the control inputs should be updated, and the EIC-based control in Eq. (6) exists. We substitute Eqs. (7) and (6) into  $\mathcal{S}_u$  dynamics and obtain

$$\ddot{\mathbf{q}}_a = \mathbf{v}^{\text{ext}} = -\mathbf{D}_{ua}^+(\mathbf{H}_u + \mathbf{D}_{uu}\mathbf{v}_u^{\text{int}}) = -\mathbf{V}\Lambda^+\mathbf{U}^T(\mathbf{H}_u + \mathbf{D}_{uu}\mathbf{v}_u^{\text{int}})$$

Multiplying the above equation on both sides with  $\mathbf{V}^T$  and considering Eq. (8),  $\mathcal{S}$  under the EIC-based control becomes Eq. (9), and the  $(n-m)$  coordinates are free of control.

**A2 Proof of Lemma 3.** Under input  $\mathbf{u}_u$ ,  $\ddot{\mathbf{q}}_{au} = \mathbf{v}_{au}^{\text{int}}$ , we solve  $\ddot{\mathbf{q}}_u$  by Eq. (21c)

$$\begin{aligned} \ddot{\mathbf{q}}_u &= -\mathbf{D}_{uu}^{-1}[\bar{\mathbf{D}}_{ua}\mathbf{v}_{au}^{\text{int}} + \mathbf{H}_{un}] \\ &= -\mathbf{D}_{uu}^{-1}[\mathbf{H}_{un} - \bar{\mathbf{D}}_{ua}^+(\bar{\mathbf{D}}_{ua})^{-1}(\mathbf{H}_{un} + \bar{\mathbf{D}}_{uu}\hat{\mathbf{v}}_u^{\text{int}})] = \hat{\mathbf{v}}_u^{\text{int}} \end{aligned}$$

Clearly, the unperturbed subsystem  $\mathcal{S}_u^{\text{gp}}$  remains the same as that under the original EIC-based control. With the designed control,  $\mathbf{q}_a$  dynamics is unchanged, and  $\ddot{\mathbf{q}}_{aa} = \hat{\mathbf{v}}_a^{\text{ext}}$  holds regardless of  $\hat{\mathbf{v}}_u^{\text{int}}$ . For  $\mathbf{q}_{aa}$  and  $\mathbf{q}_{au}$ , we obtain  $\ddot{\mathbf{q}}_{aa} = \hat{\mathbf{v}}_a^{\text{ext}}$  and  $\ddot{\mathbf{q}}_{au} = \hat{\mathbf{v}}_u^{\text{int}}$ . The relationship in

Eq. (9) indicates that if the unactuated subsystem dynamics is written into  $\ddot{\mathbf{q}}_u = \mathbf{v}_u^{\text{int}}$ , the dynamics  $\ddot{\mathbf{q}}_a$  under the transformation  $\mathbf{Y}$  must contain the portion (9a). Similarly, we obtain

$$\mathcal{S}_{\text{PEIC}}: \ddot{\mathbf{p}}_{ai} = -\frac{\mathbf{u}_i^T(\mathbf{H}_u^{\text{gp}} + \bar{\mathbf{D}}_{uu}\hat{\mathbf{v}}_u^{\text{int}})}{\sigma_i}, \quad i = 1, \dots, m \quad (\text{A4a})$$

$$\ddot{\mathbf{p}}_{aj} = \mathbf{v}_j^T\hat{\mathbf{v}}_a^{\text{int}}, \quad j = m+1, \dots, n \quad (\text{A4b})$$

$$\ddot{\mathbf{q}}_u = \hat{\mathbf{v}}_u^{\text{int}} \quad (\text{A4c})$$

where  $\hat{\mathbf{v}}_a^{\text{int}} = [(\hat{\mathbf{v}}_a^{\text{ext}})^T, (\hat{\mathbf{v}}_u^{\text{int}})^T]^T$ . Since  $\hat{\mathbf{v}}_a^{\text{int}}$  is not obtained in the way as in Eq. (5), i.e.,  $\hat{\mathbf{v}}_a^{\text{int}} \notin \ker(\bar{\mathbf{D}}_{ua})$ ,  $\mathbf{v}_{m+j}^T\hat{\mathbf{v}}_a^{\text{int}} \neq 0$  and  $\mathbf{p}_{an}$  is under active control. Meanwhile,  $\mathbf{v}_{m+j}^T\hat{\mathbf{v}}_a^{\text{int}}$  drives  $\mathbf{q}_a \rightarrow \mathbf{q}_a^d$  in  $\ker(\bar{\mathbf{D}}_{au})$ , given that  $\hat{\mathbf{v}}_a^{\text{ext}}$  and  $\hat{\mathbf{v}}_u^{\text{int}}$  are designed to drive  $\mathbf{q}_a \rightarrow \mathbf{q}_a^d$ . Therefore, if the unperturbed system under the original EIC-based control is stable, it is also stable under the PEIC-based control.

**A3 Proof of Lemma 4.** Under the NEIC-based control input, the  $\mathcal{S}_u^{\text{gp}}$  becomes

$$\ddot{\mathbf{q}}_a = \hat{\mathbf{v}}_a^{\text{int}} = \mathbf{v}^{\text{int}} + \mathbf{v}_a^{\text{ext}} = -\bar{\mathbf{D}}_{ua}^+(\mathbf{H}_u^{\text{gp}} + \bar{\mathbf{D}}_{uu}\hat{\mathbf{v}}_u^{\text{int}}) + \mathbf{V}_n\mathbf{v}_n \quad (\text{A5})$$

Plugging above equation into  $\mathcal{S}_u^{\text{gp}}$ , we obtain

$$\begin{aligned} \ddot{\mathbf{q}}_u &= -\bar{\mathbf{D}}_{uu}^{-1}(\bar{\mathbf{D}}_{ua}\ddot{\mathbf{q}}_a + \mathbf{H}_u^{\text{gp}}) = -\bar{\mathbf{D}}_{uu}^{-1}[-\bar{\mathbf{D}}_{ua}\bar{\mathbf{D}}_{ua}^+(\mathbf{H}_u^{\text{gp}} \\ &\quad + \bar{\mathbf{D}}_{uu}\hat{\mathbf{v}}_u^{\text{int}}) + \bar{\mathbf{D}}_{ua}\mathbf{V}_n\mathbf{v}_n + \mathbf{H}_u^{\text{gp}}] = \hat{\mathbf{v}}_u^{\text{int}} - \bar{\mathbf{D}}_{uu}^{-1}\bar{\mathbf{D}}_{ua}\mathbf{V}_n\mathbf{v}_n \end{aligned}$$

Using the SVD form of  $\bar{\mathbf{D}}_{ua}$  in Eq. (7) and  $\Lambda\mathbf{V}^T\mathbf{V}_n \equiv \mathbf{0}$ , the above equation is further simplified as

$$\ddot{\mathbf{q}}_u = \hat{\mathbf{v}}_u^{\text{int}} - \bar{\mathbf{D}}_{uu}^{-1}\mathbf{U}\Lambda\mathbf{V}^T\mathbf{V}_n\mathbf{v}_n = \hat{\mathbf{v}}_u^{\text{int}} \quad (\text{A6})$$

Clearly,  $\mathcal{S}_u^{\text{gp}}$  dynamics is unchanged compared to Eq. (9).

We further apply the transformation  $\mathbf{Y}$  to  $\mathbf{q}_a$  and SVD to  $\bar{\mathbf{D}}_{ua}^+$ . The  $\mathcal{S}_u$  dynamics (A5) and (A6) become

$$\mathcal{S}_{\text{NEIC}}: \ddot{\mathbf{p}}_{ai} = -\frac{\mathbf{u}_i^T(\mathbf{H}_u^{\text{gp}} + \bar{\mathbf{D}}_{uu}\hat{\mathbf{v}}_u^{\text{int}})}{\sigma_i}, \quad i = 1, \dots, m \quad (\text{A7a})$$

$$\ddot{\mathbf{p}}_{aj} = \nu_{nj}, \quad j = m+1, \dots, n \quad (\text{A7b})$$

$$\ddot{\mathbf{q}}_u = \hat{\mathbf{v}}_u^{\text{int}} \quad (\text{A7c})$$

Compared to Eq. (9), we add control  $\hat{\mathbf{v}}_a^{\text{ext}}$  to drive  $\mathbf{q}_a \rightarrow \mathbf{q}_a^d$  in the subspace  $\ker(\bar{\mathbf{D}}_{ua})$ . Therefore, if the system (9) is stable, Eq. (A4) is also stable, as the  $\mathbf{p}_{am}$  and  $\mathbf{q}_u$  dynamics are unchanged.

**A4 Proof for Theorem 1.** We present the stability proof for the PEIC- and NEIC-based controls using the Lyapunov method.

*PEIC-Based Control:* Plugging Eq. (24) into  $V_1 = V$  and considering Eq. (32), we obtain  $\dot{V}_1 = \mathbf{e}^T(\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A})\mathbf{e} + 2\mathbf{e}^T\mathbf{P}\mathbf{O}_1 = -\mathbf{e}^T\mathbf{Q}\mathbf{e} + \mathbf{e}^T\mathbf{Q}_\Sigma\mathbf{e} + 2\mathbf{e}^T\mathbf{P}\mathbf{O}_1$ , where  $\mathbf{Q}_\Sigma = (\mathbf{A} - \mathbf{A}_0)^T\mathbf{P} + \mathbf{P}(\mathbf{A} - \mathbf{A}_0)$ . The bounded variance leads to the bounded eigenvalue of matrix  $\mathbf{Q}_\Sigma$ . Given the fact that  $\mathbf{Q}_\Sigma = \mathbf{Q}_\Sigma^T$ , the eigenvalues of  $\mathbf{Q}_\Sigma$  are real numbers.

Noting that  $\mathbf{Q}_\Sigma$  is bounded and  $\mathbf{P}, \mathbf{Q}$  are constant, the perturbation term  $\mathbf{O}_1$  is bounded as shown in Eq. (26). Then,  $\dot{V}_1$  is rewritten as

$$\begin{aligned} \dot{V}_1 &\leq -[\lambda_{\min}(\mathbf{Q}) - \lambda_{\max}(\mathbf{Q}_\Sigma)]\|\mathbf{e}\|^2 + 2\lambda_{\max}(\mathbf{P})\|\mathbf{e}\|(d_1 \\ &\quad + d_2\|\mathbf{e}\|) + 2\lambda_{\max}(\mathbf{P})\|\mathbf{e}\|(l_{u1}\|\mathbf{k}_u\| + l_{a1}\|\mathbf{k}_a\|) \\ &= -[\lambda_{\min}(\mathbf{Q}) - \lambda_{\min}(\mathbf{Q}_\Sigma) - 2d_2\lambda_{\max}(\mathbf{P})]\|\mathbf{e}\|^2 \\ &\quad + 2(d_1 + \omega_1)\lambda_{\max}(\mathbf{P})\|\mathbf{e}\| \end{aligned}$$

where  $\omega_1 = l_{u1}\|\mathbf{k}_u\| + l_{a1}\|\mathbf{k}_a\|$  denotes the uncertainties related to GP prediction errors.  $\lambda_{\min}(\cdot)$  and  $\lambda_{\max}(\cdot)$  denote the smallest and greatest eigenvalues of a matrix, respectively. Considering  $\lambda_{\min}(\mathbf{P})\|\mathbf{e}\|^2 \leq V_1 \leq \lambda_{\max}(\mathbf{P})\|\mathbf{e}\|^2$ , we define

$$\gamma_1 = \frac{\lambda_{\min}(\mathbf{Q}) - \lambda_{\max}(\mathbf{Q}_\Sigma) - 2d_2\lambda_{\max}(\mathbf{P})}{\lambda_{\max}(\mathbf{P})}$$

$\rho_1 = 2d_1\lambda_{\max}(\mathbf{P})\|\mathbf{e}\|$ ,  $\varpi_1 = 2\omega_1\lambda_{\max}(\mathbf{P})\|\mathbf{e}\|$ . With the bounded perturbations  $\rho_1$  and  $\omega_1$ , the closed-loop system dynamics can be shown stable in probability as  $\Pr\{V_1 \leq -\gamma_1 V_1 + \rho_1 + \varpi_1\} > \eta$ . Taking further analysis, we obtain a nominal estimation of the error convergence as  $\Pr\{\dot{V}_1 \leq V_1(0)e^{-\gamma_1 t}\} > \eta$  and the error bound estimation  $\Pr\{\|\mathbf{e}\| \leq r_1\} > \eta$  with  $r_1 = (2d_1\lambda_{\max}(\mathbf{P})) / (\lambda_{\min}(\mathbf{Q}) - \lambda_{\max}(\mathbf{Q}_\Sigma) - 2d_2\lambda_{\max}(\mathbf{P}))$ .

**NEIC-Based Control:** Without the loss of generality, we select  $\mathbf{v}_n = \mathbf{V}_n^T \hat{\mathbf{v}}^{\text{ext}}$ . We take  $V_2 = V$  as the Lyapunov function candidate for  $\mathcal{S}_{\text{e,NEIC}}$ . If the control gains are the same as that in the PEIC-based control and  $\alpha = 1$  for compensation effect,  $\gamma_2 = \gamma_1$ . We choose control gains properly such that  $\gamma_2 > 0$ . The system can be shown stable as  $\Pr\{\dot{V}_2 \leq -\gamma_2 V_2 + \rho_2 + \varpi_2\} > \eta$ , where  $\rho_2 = 2d_1\lambda_{\max}(\mathbf{P})\|\mathbf{e}\|$ ,  $\varpi_2 = 2\omega_2\lambda_{\max}(\mathbf{P})\|\mathbf{e}\|$ , and  $\omega_2 = l_{a2}\|\mathbf{k}_u\| + l_{d2}\|\mathbf{k}_d\|$  is defined same as  $\omega_1$  containing the GP prediction uncertainties. A nominal estimation of error convergence and final error bound can also be obtained.

To show  $\gamma_i > 0$ ,  $i = 1, 2$ , the control gains should be properly selected. With a small predefined error limit as a stop criterion in BEM estimation,  $c_i$  values can be shown as  $c_i \ll 1$ . Given the explicit form,  $d_i$  are estimated for  $\mathbf{A}_0$  and  $\mathbf{Q}, \mathbf{P}$  is obtained by solving Eq. (32). The matrix  $\mathbf{Q}_\Sigma$  depends on the control gains associated with the reduction variance. Since the variance is bounded, we design  $\mathbf{k}_{ni}$  such that  $\lambda_{\max}(\mathbf{Q}_\Sigma)$  satisfies the inequality  $\lambda_{\min}(\mathbf{Q}) - \lambda_{\max}(\mathbf{Q}_\Sigma) - 2d_2\lambda_{\max}(\mathbf{P}) > 0$  and then  $\gamma_1 > 0$ . Thus, the stability is obtained.

## Appendix B: Dynamics Model of Underactuated Balance Robots

**B1 Rotary Inverted Pendulum.** The dynamics model for the rotary pendulum is in the form of Eq. (1) with  $q_a = \theta_1$  and  $q_u = \theta_2$ . The model parameters are  $\mathbf{B} = [1 \ 0]^T$  and

$$\begin{aligned} D_{aa} &= C(m_p l_p^2 + 0.25m_p l_p^2 s_2^2 + J_r) \\ D_{au} &= D_{ua} = -0.5Cm_p l_p l_r c_2, \quad D_{uu} = C(J_p + 0.25m_p l_p^2) \\ H_a &= C(0.5m_p l_p^2 \dot{\theta}_1 \dot{\theta}_2 s_2 c_2 + 0.5m_p l_p l_r \dot{\theta}_2^2 s_2 \\ &\quad + d_r \dot{\theta}_1 + k_g^2 k_r k_m \dot{\theta}_1 / R_m) + K_g k_r \dot{\theta}_2 \\ H_u &= C(d_p \dot{\theta}_2 - 0.25m_p l_p^2 c_2 s_2 \dot{\theta}_2^2 - 0.5m_p l_p g s_2) \end{aligned}$$

where  $l_r, J_r$ , and  $d_r$  are the length, mass inertia, and viscous damping coefficient of the base link,  $l_p, J_p$ , and  $d_p$  are corresponding parameters of the pendulum,  $m_p$  is the pendulum mass,  $g$  is the gravitational constant, and  $k_r, k_m, K_g, R_m$ , and  $C$  are robot constant. The values of these parameters can be found in Ref. [27]. The control input is the motor voltage, i.e.,  $u = V_m$ .

**B2 Three-Link Inverted Pendulum.** The model parameters in Eq. (1) are

$$\begin{aligned} D_{11} &= (m_3(l_3^2 + 0.25l_3^2) + 0.25m_2l_2^2 - 0.5m_3l_3^2 c_3^2 - m_3l_2l_3s_3)c_2^2 \\ &\quad + (0.5m_3s_3l_3^2 - m_3l_2l_3)s_2c_3c_2 + 0.25m_3c_3^2l_3^2 \\ &\quad + (0.25m_1 + m_2 + m_3)l_1^2 + J_1 \\ D_{12} &= D_{21} = -(m_3l_2 + 0.5m_2l_2)l_1s_2 - 0.5m_3l_1l_3c_{2+3} \\ D_{13} &= D_{31} = 0.5m_3l_1l_3c_{2+3} \\ D_{22} &= J_2 + (m_3 + 0.25m_2)l_2^2 + 0.25m_3l_3^2 - m_3l_2l_3s_3 \\ D_{23} &= D_{32} = (0.25l_3 - 0.5l_2s_3)m_3l_3, \quad D_{33} = J_3 + 0.25m_3l_3^2 \\ G_1 &= 0, \quad G_2 = -(0.5m_2 + m_3)c_2l_2g + 0.5m_3l_3s_{2+3}g \\ G_3 &= -0.5m_3l_3s_{2+3}g \end{aligned}$$

where  $m_i, l_i$ , and  $J_i$  are the mass, length, and mass inertia of each link, and  $s_{i+j} = \sin(\theta_i + \theta_j)$ . Matrix  $\mathbf{C}$  is obtained as  $C_{ij} = \sum_{k=1}^3 c_{ijk} \dot{\theta}_k$ , where Christoffel symbols  $c_{ijk} = \frac{1}{2}((\partial D_{ij}/\partial \theta_k) + (\partial D_{ik}/\partial \theta_j) - (\partial D_{jk}/\partial \theta_i))$ . The physical parameters are  $m_1 = 0.7 \text{ kg}$ ,  $m_2 = 1.3 \text{ kg}$ ,  $m_3 = 0.3 \text{ kg}$ ,  $l_1 = 0.065 \text{ m}$ ,  $l_2 = 0.23 \text{ m}$ ,  $l_3 = 0.25 \text{ m}$ ,  $J_1 = 0.0008 \text{ kg m}^2$ ,  $J_2 = 0.005 \text{ kg m}^2$ , and  $J_3 = 0.003 \text{ kg m}^2$ .

## References

- [1] Kant, N., and Mukherjee, R., 2020, "Orbital Stabilization of Underactuated Systems Using Virtual Holonomic Constraints and Impulse Controlled Poincaré Maps," *Syst. Control Lett.*, **146**, p. 104813.
- [2] Han, F., and Yi, J., 2023, "On the Learned Balance Manifold of Underactuated Balance Robots," *IEEE International Conference on Robotics and Automation (ICRA)*, London, UK, May 29–June 2, pp. 12254–12260.
- [3] Han, F., Jelvani, A., Yi, J., and Liu, T., 2022, "Coordinated Pose Control of Mobile Manipulation With an Unstable Bikebot Platform," *IEEE/ASME Trans. Mechatron.*, **27**(6), pp. 4550–4560.
- [4] Chen, K., Yi, J., and Song, D., 2023, "Gaussian-Process-Based Control of Underactuated Balance Robots With Guaranteed Performance," *IEEE Trans. Rob.*, **39**(1), pp. 572–589.
- [5] Han, F., and Yi, J., 2021, "Stable Learning-Based Tracking Control of Underactuated Balance Robots," *IEEE Rob. Autom. Lett.*, **6**(2), pp. 1543–1550.
- [6] Turrisi, G., Capotondi, M., Gaz, C., Modugno, V., Oriolo, G., and Luca, A. D., 2022, "On-Line Learning for Planning and Control of Underactuated Robots With Uncertain Dynamics," *IEEE Rob. Autom. Lett.*, **7**(1), pp. 358–365.
- [7] Beckers, T., Kulić, D., and Hirche, S., 2019, "Stable Gaussian Process Based Tracking Control of Euler–Lagrange Systems," *Automatica*, **103**, pp. 390–397.
- [8] Chen, K., and Yi, J., 2015, "On the Relationship Between Manifold Learning Latent Dynamics and Zero Dynamics for Human Bipedal Walking," *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Hamburg, Germany, Sept. 28–Oct. 2, pp. 971–976.
- [9] Grizzle, J. W., Chevallereau, C., Sinnen, R. W., and Ames, A. D., 2014, "Models, Feedback Control, and Open Problems in 3D Bipedal Robotic Walking," *Automatica*, **50**(8), pp. 1955–1988.
- [10] Han, F., Huang, X., Wang, Z., Yi, J., and Liu, T., 2022, "Autonomous Bikebot Control for Crossing Obstacles With Assistive Leg Impulsive Actuation," *IEEE/ASME Trans. Mechatron.*, **27**(4), pp. 1882–1890.
- [11] Shiriaev, A. S., Perram, J. W., and Canudas-de-Wit, C., 2005, "Constructive Tool for Orbital Stabilization of Underactuated Nonlinear Systems: Virtual Constraints Approach," *IEEE Trans. Autom. Control*, **50**(8), pp. 1164–1176.
- [12] de Wit, C. C., Espiau, B., and Urrea, C., 2002, "Orbital Stabilization of Underactuated Mechanical Systems," *IFAC Proc. Vol.*, **35**(1), pp. 527–532.
- [13] Maggiore, M., and Consolini, L., 2013, "Virtual Holonomic Constraints for Euler–Lagrange Systems," *IEEE Trans. Autom. Control*, **58**(4), pp. 1001–1008.
- [14] Chevallereau, C., Grizzle, J. W., and Shih, C.-L., 2009, "Asymptotically Stable Walking of a Five-Link Underactuated 3-D Bipedal Robot," *IEEE Trans. Rob.*, **25**(1), pp. 37–50.
- [15] Fantoni, I., Lozano, R., and Spong, M. W., 2000, "Energy Based Control of Pendubot," *IEEE Trans. Autom. Control*, **45**(4), pp. 725–729.
- [16] Xin, X., and Kanedai, M., 2005, "Analysis of the Energy-Based Control for Swinging Up Two Pendulums," *IEEE Trans. Autom. Control*, **50**(5), pp. 679–684.
- [17] Getz, N., 1995, "Dynamic Inversion of Nonlinear Maps With Applications to Nonlinear Control and Robotics," Ph.D. thesis, Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA.
- [18] Lambert, N. O., Schindler, C. B., Drew, D. S., and Pister, K. S. J., 2021, "Nonholonomic Yaw Control of an Underactuated Flying Robot With Model-Based Reinforcement Learning," *IEEE Rob. Autom. Lett.*, **6**(2), pp. 455–461.
- [19] Beckers, T., and Hirche, S., 2022, "Prediction With Approximated Gaussian Process Dynamical Models," *IEEE Trans. Autom. Control*, **67**(12), pp. 6460–6473.
- [20] Lederer, A., Yang, Z., Jiao, J., and Hirche, S., 2023, "Cooperative Control of Uncertain Multiagent Systems Via Distributed Gaussian Processes," *IEEE Trans. Autom. Control*, **68**(5), pp. 3091–3098.
- [21] Deisenroth, M., and Ng, J. W., 2015, "Distributed Gaussian Processes," *International Conference on Machine Learning*, Lille, France, July 6–11, pp. 1481–1490.
- [22] Helwa, M. K., Heins, A., and Schoellig, A. P., 2019, "Provably Robust Learning-Based Approach for High-Accuracy Tracking Control of Lagrangian Systems," *IEEE Rob. Autom. Lett.*, **4**(2), pp. 1587–1594.
- [23] Chen, K., Yi, J., and Song, D., 2019, "Gaussian Processes Model-Based Control of Underactuated Balance Robots," *International Conference on Robotics and Automation (ICRA)*, Montreal, QC, Canada, May 20–24, pp. 4458–4464.
- [24] Han, F., and Yi, J., 2024, "Gaussian Process-Enhanced, External and Internal Convertible (EIC) Form-Based Control of Underactuated Balance Robots," *Proceedings of the IEEE International Conference on Robotics and Automation*, Yokohama, Japan, May 13–17, pp. 8937–8933.
- [25] Wang, P., Han, F., and Yi, J., 2023, "Gyroscopic Balancer-Enhanced Motion Control of an Autonomous Bikebot," *ASME J. Dyn. Syst., Meas., Control*, **145**(5), p. 101002.
- [26] Srinivas, N., Krause, A., Kakade, S. M., and Seeger, M. W., 2012, "Information-Theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting," *IEEE Trans. Inf. Theory*, **58**(5), pp. 3250–3265.
- [27] Apkarian, J., Karam, P., and Levis, M., 2011, *Instructor Workbook: Inverted Pendulum Experiment for Matlab/Simulink Users*, Quanser Inc., Markham, ON, Canada.