Preprint version of the article:
S. R. Nekoo, J. A. Acosta, G. Heredia, and A. Ollero, "A PD-Type State-Dependent Riccati Equation with Iterative Learning Augmentation for Mechanical Systems," IEEE/CAA Journal of Automatica Sinica, 2022. doi: 10.1109/JAS.2022.105533

A PD-Type State-Dependent Riccati Equation with Iterative Learning Augmentation for Mechanical Systems

Saeed Rafee Nekoo, José Ángel Acosta, Guillermo Heredia (Member, IEEE) and Anibal Ollero (Fellow, IEEE)

***Abstract*—This work proposes a novel proportional-derivative (PD)-type state-dependent Riccati equation (SDRE) approach with iterative learning control (ILC) augmentation. On the one hand, the PD-type control gains could adopt many useful available criteria and tools of conventional PD controllers. On the other hand, the SDRE adds nonlinear and optimality characteristics to the controller, i.e. increasing the stability margins. These advantages with the ILC correction part deliver a precise control law with the capability of error reduction by learning. The SDRE provides a symmetric-positive-definite distributed nonlinear suboptimal gain $\mathbf{K}(\mathbf{x})$ for the control input law $\mathbf{u} = -\mathbf{R}^{-1}(\mathbf{x})\mathbf{B}^{T}(\mathbf{x})\mathbf{K}(\mathbf{x})\mathbf{x}$. The sub-blocks of the overall gain $\mathbf{R}^{-1}(\mathbf{x})\mathbf{B}^{T}(\mathbf{x})\mathbf{K}(\mathbf{x})$, are not necessarily symmetric positive definite. A new design is proposed to transform the optimal gain into two symmetric-positive-definite gains like PD-type controllers as $\mathbf{u} = -\mathbf{K}_{\mathbf{SP}}(\mathbf{x})\mathbf{e} - \mathbf{K}_{\mathbf{SD}}(\mathbf{x})\dot{\mathbf{e}}$. The new form allows us to analytically prove the stability of the proposed learning-based controller for mechanical systems; and presents guaranteed uniform boundedness in finite-time between learning loops. The symmetric PD-type controller is also developed for the state-dependent differential Riccati equation (SDDRE) to manipulate the final time. The SDDRE expresses a differential equation with a final boundary condition, which imposes a constraint on time that could be used for finite-time control. So, the availability of PD-type finite-time control is an asset for enhancing the conventional classical linear controllers with this tool. The learning rules benefit from the gradient descent method for both regulation and tracking cases. One of the advantages of this approach is a guaranteed-stability even from the first loop of learning. A mechanical manipulator, as an illustrative example, was simulated for both regulation and tracking problems. Successful experimental validation was done to show the capability of the system in practice by the implementation of the proposed method on a variable-pitch rotor benchmark.**

***Index Terms*—SDRE; SDDRE; Symmetric; PD-type; Closed-loop; Iterative Learning Control.**

I. Introduction

THE proportional derivative (PD) control has well-established mathematics, so many well-known tools, assessment and tuning methods, etc. It has been accepted by the industrial platforms which prefer the simplicity of the control law and straightforward analysis. What if someone wants a finite-time PD controller? We intend to gain all the benefits of the PD control with the additional power of nonlinearity and finite-time regulation. The state-dependent Riccati equation (SDRE) has been defined as an optimal control design for nonlinear systems. Optimality, robustness, design flexibility, and systematic procedure are some advantages of the SDRE [1]. The differential form of the SDRE uses a final boundary condition to impose an extra penalty on final states near the end of time, the method is so-called the state-dependent differential Riccati equation (SDDRE) [2]. Both SDRE and SDDRE were widely used in aerospace [3], approximate dynamic programming framework [4], cancer treatment modeling [5], etc. Either of SDRE or SDDRE method provides a suboptimal-distributed symmetric-positive-definite gain $\mathbf{K}(\mathbf{x})$ for the standard form of Riccati control law $\mathbf{u} = -\mathbf{X}(\mathbf{x})\mathbf{x}$ in which $\mathbf{X}(\mathbf{x}) = \mathbf{R}^{-1}(\mathbf{x})\mathbf{B}^{T}(\mathbf{x})\mathbf{K}(\mathbf{x})$. This work proposes a new control structure to transform the distributed gain into two symmetric-positive-definite gains for the error vector and its derivative $\mathbf{u} = -\mathbf{K}_{\mathrm{SP}}(\mathbf{x})\mathbf{e} - \mathbf{K}_{\mathrm{SD}}(\mathbf{x})\dot{\mathbf{e}}$.

The proposed structure possesses many advantages such as independent control of error of the system, PD shape, the structure of a finite-time PD controller, etc.; however, the main objective for the transformation is to release a new design to analytically guarantee the stability of a novel SDRE design, augmented by iterative learning control (ILC) approach. Using PD-like control seems more practical and widely used, and easier for implementation and mathematical derivation. The motivation for using the SDRE controller is the nonlinear optimal structure and finite-time characteristics of the controller. The optimality deviates since the symmetric structure changes the original gain $\mathbf{K}(\mathbf{x})$; nonetheless, the nonlinearity and finite-time control remained as the advantages of the proposed method. Iterative learning control uses previous data to update the next control loop [6]. A proper ILC converges the error of a system towards zero in each control loop [7-9]. One of the main advantages of the ILC is a compensation of the modeling uncertainty by the learning process [10-12]. This is more critical when dynamic modeling is complex, for example,

This Research is funded by the European Commission H2020 Programme under HYFLIERS project contract 779411, AERIAL-CORE project contract number 871479 and the ARTIC (RTI2018-102224-B-I00) project, funded by the Spanish Agencia Estatal de Investigación.

Authors are with the GRVC Robotics Lab., Depto de Ingeniería de Sistemas y Automática, Escuela Técnica Superior de Ingeniería, Universidad de Sevilla, Seville, Spain, emails: saerafee@yahoo.com, (jaar, guiller, aollero)@us.es

the behavior of the flapping wing flying robots [13]; or piezoelectric actuators with hysteretic nonlinearity [14]. Shen and Xu presented an adaptive ILC for systems with randomly varying iteration lengths [15]. Robotics has been an attractive field for the implementation of ILC. Roveda et al. presented the iterative learning control approach with reinforcement peculiar to high-accuracy force tracking in robotics [16]. Oh et al. researched a regulation problem for iterative learning model predictive control [17]. Pan and Yu developed a composite learning robot control possessing guaranteed parameter convergence [18]. The iterative learning control improves trajectory tracking by several trials; Schollig and D'Andrea employed the learning for tracking systems with state and input constraints [19]. The dynamic of a brush-bot is a challenging topic, especially in terms of actuation; hence, one efficient way is to use learning methods. Barrier-certified adaptive control was implemented for the navigation of a brush-bot as a reinforcement learning method [20]. Vision-based learning was used for racing drones in a tracking problem with moving uncertain trajectories [21]. The high level of uncertainty was handled by a deep neural network learning tool. In reinforcement learning, a gradient descent method is a powerful approach for training the models, also capable of modification for computation enhancement [22].

The closed-loop ILCs, with PD- or PID-based structures were reported [23, 24]; however, the ILC was usually employed to skip the unknown closed-loop dynamic in feedforward methods [25, 26]. It was also common to consider ILC based on a core hypothesis that repetitiveness of task and model was satisfied; however, robustness with ILC was shown using a high-order internal model [27-29]. In other words, the learning process helps to deliver an ideal controller without engaging the design with offline dynamic modeling; and using the online data of each step for the performance enhancement of the next loop. The open-loop approach demonstrated the nature of the learning by showing unexpected behavior at primary loops might result in bad trajectories, collisions, saturations, or destruction of the prototypes. So, open-loop learning is preferable for stable systems; i.e. a quadrotor with an internal stable loop [26]. The closed-loop ILC could possess stable primary loops at the beginning. This is useful for dynamically unstable systems. Here in this work, the presented learning-based controller is closed-loop, applicable for a dynamically unstable system. The experimental implementation of this work uses a propeller-type inverted pendulum that, in nature, is unstable at an equilibrium point. So, the PD-type SDDRE generates a primary input signal (stable) for the first loop and the ILC will update and reduce the error in consecutive learning loops.

The stability of the proposed ILC is proved by a new Lyapunov-like candidate with a dynamic-scaling factor. From the control application point of view, the best qualities of the new controller are its PD-like structure and finite-time option. We propose a baseline PD controller that is redesigned to gain robustness via dynamic scaling. The error with this controller will converge to zero at infinity, and in finite-time, although bounded, the error might be too large for the application. Thus, the ILC reduces the error in each iteration based on the update in the previous loop. The proposed design guarantees bounded error, $\varepsilon(k)$, at k-th iteration $\lim_{t\to t_f}|e(t)| \le \varepsilon(k)$ as it is shown in Fig. 1. For an infinite-time case, the stability of the nonlinear PD-like controller for robot manipulators was investigated and it was shown that this form leads to global asymptotic stability [30]. The integrator part of the control also generated semi-global stability for saturated linear PID controllers [31].

The current work uses a gradient descent method as a training rule for updating the learning feedforward control law to propose a symmetric gain SDRE/SDDRE control and combining it with the ILC to form a new finite-time iterative learning approach. The stability proof of the control law is presented. The main advantage of this learning system is to employ a nonlinear (suboptimal) input law besides the ILC to guarantee stable regulation and tracking even from the first loop of the learning. The proposed ILC is also implemented on a variable-pitch (VP) rotor benchmark for experimental validation. The VP platform was controlled by the SDRE [32], though here for the first time, the implementation of a learning-based controller has been presented.

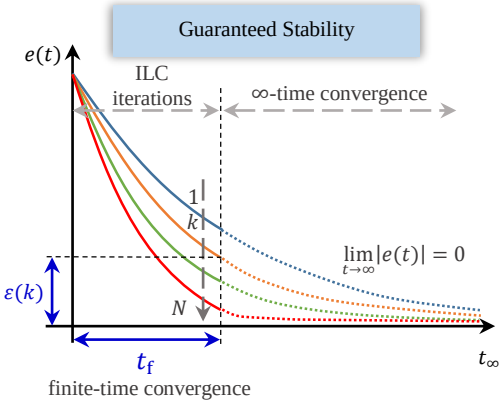

Fig. 1. The explanation of the ILC operation in error reduction in finite-time.

The main contributions of this work are as follows:

C1. Presenting a nonlinear finite-time PD-like controller based on SDDRE with the ILC feedforward compensation to gain more precision.

C2. Transforming the suboptimal-distributed symmetric-positive-definite gain of the SDRE/SDDRE methodology into novel PD-like gains, which is more suitable for mechanical systems.

C3. Proving the stability of the nonlinear controller with ILC.

C4. Introducing a novel convex objective function for the regulation training rule of the gradient descent method.

C5. Guaranteeing uniform boundedness in finite-time between learning loops allows us to use the approach for unstable mechanical systems.

Notations: Let $\lambda_{\max}(\cdot)$ and $\lambda_{\min}(\cdot)$ be maximum and minimum eigenvalues of an arbitrary matrix, respectively. $\mathbb{R}^n$ is denoted as the n-dimensional Euclidean space; $\mathbb{R}^{n\times m}$ is presented as the set of $n \times m$ real matrices. $(\cdot)^T$ shows transpose and $\mathrm{diag}(\cdot)$ means a diagonal matrix.

Section II briefly presents the SDRE/SDDRE controller structure. Iterative control system design has been presented for both regulation and tracking problems in Section III. Learning rules based on the gradient descent method have been expressed in Section IV. Simulations and experiments are presented in Sections V and VI, respectively, and the conclusions in Section VII.

II. The SDRE/SDDRE Control Design

Consider a nonlinear system

$$\dot{\mathbf{x}}(t) = \mathbf{A}\big(\mathbf{x}(t)\big)\mathbf{x}(t) + \mathbf{B}\big(\mathbf{x}(t)\big)\mathbf{u}(t), \tag{1}$$

where $\mathbf{x}(t) \in \mathbb{R}^n$ is a state vector, $\mathbf{u}(t) \in \mathbb{R}^m$ is an input vector. $\mathbf{A}\big(\mathbf{x}(t)\big): \mathbb{R}^n \to \mathbb{R}^{n\times n}$ and $\mathbf{B}\big(\mathbf{x}(t)\big): \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are state-dependent coefficient (SDC) parameterization of a nonlinear system $\dot{\mathbf{x}}(t) = \mathbf{f}\big(\mathbf{x}(t)\big) + \mathbf{g}\big(\mathbf{x}(t), \mathbf{u}(t)\big)$ where $\mathbf{f}\big(\mathbf{x}(t)\big): \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{g}\big(\mathbf{x}(t), \mathbf{u}(t)\big): \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ are piecewise-continuous vector-valued functions that satisfy a local Lipschitz condition. The aim is to minimize the cost function [33]:

$$J(\cdot) = \frac{1}{2}\left(\mathbf{x}^T(t_\mathrm{f})\mathbf{F}\mathbf{x}(t_\mathrm{f}) + \int_0^{t_\mathrm{f}} \{\mathbf{x}^T\mathbf{Q}(\mathbf{x})\mathbf{x} + \mathbf{u}^T\mathbf{R}(\mathbf{x})\mathbf{u}\}\mathrm{d}t\right), \tag{2}$$

and finish a control task in a finite predefined time $t \in [0, t_\mathrm{f}]$. $\mathbf{Q}\big(\mathbf{x}(t)\big): \mathbb{R}^n \to \mathbb{R}^{n\times n}$ and $\mathbf{F} \in \mathbb{R}^{n\times n}$ penalize the states in $t \in [0, t_\mathrm{f})$ and at t_f, respectively (both symmetric positive semi-definite); $\mathbf{R}\big(\mathbf{x}(t)\big): \mathbb{R}^n \to \mathbb{R}^{m\times m}$ penalizes the inputs (symmetric positive definite).

Assumption 1. (Controllability). The nonlinear pair of $\big[\mathbf{A}\big(\mathbf{x}(t)\big), \mathbf{B}\big(\mathbf{x}(t)\big)\big]$ must form a completely controllable parameterization of the original system (1) with the mentioned conditions [1].

Assumption 2. (Observability). The nonlinear pair of $\big[\mathbf{A}\big(\mathbf{x}(t)\big), \mathbf{Q}^{1/2}\big(\mathbf{x}(t)\big)\big]$ must form a completely observable parameterization of the original system (1) with the mentioned conditions [1].

Satisfying Assumptions 1 and 2 was discussed and for more details on checking those conditions, please visit Refs. [34, 35].

The control law representing the SDRE/SDDRE is:

$$\mathbf{u}(t) = -\mathbf{R}^{-1}\big(\mathbf{x}(t)\big)\mathbf{B}^T\big(\mathbf{x}(t)\big)\mathbf{K}\big(\mathbf{x}(t)\big)\mathbf{x}(t), \tag{3}$$

where the symmetric positive-definite suboptimal gain $\big[\mathbf{K}\big(\mathbf{x}(t)\big)\big]_{2n\times 2n}$, is a solution to the state-dependent Riccati equation [1]:

$$\begin{aligned}\mathbf{A}^T(\mathbf{x})\mathbf{K}_{\mathrm{ss}}(\mathbf{x}) + \mathbf{K}_{\mathrm{ss}}(\mathbf{x})\mathbf{A}(\mathbf{x}) \\ - \mathbf{K}_{\mathrm{ss}}(\mathbf{x})\mathbf{B}(\mathbf{x})\mathbf{R}^{-1}(\mathbf{x})\mathbf{B}^T(\mathbf{x})\mathbf{K}_{\mathrm{ss}}(\mathbf{x}) \\ + \mathbf{Q}(\mathbf{x}) = \mathbf{0},\end{aligned} \tag{4}$$

or the state-dependent differential Riccati equation [1]:

$$\begin{aligned}\dot{\mathbf{K}}(\mathbf{x}) = -\{\mathbf{A}^T(\mathbf{x})\mathbf{K}(\mathbf{x}) + \mathbf{K}(\mathbf{x})\mathbf{A}(\mathbf{x}) \\ - \mathbf{K}(\mathbf{x})\mathbf{B}(\mathbf{x})\mathbf{R}^{-1}(\mathbf{x})\mathbf{B}^T(\mathbf{x})\mathbf{K}(\mathbf{x}) \\ + \mathbf{Q}(\mathbf{x})\}, \qquad \mathbf{K}(t_\mathrm{f}) = \mathbf{F}.\end{aligned} \tag{5}$$

In LQR design for linear system or an open-loop two-point boundary value problem, there exists optimality; however, in the SDRE or SDDRE, which is the extension of LQR for nonlinear systems, the following terms appear in the derivation such as $\left(\frac{\partial \mathbf{A}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{x}\right)^T$, $\left(\frac{\partial \mathbf{B}(\mathbf{x})}{\partial \mathbf{x}}\mathbf{x}\right)^T$, etc., for more details Refs. [36-39] could be visited. Therefore, the SDRE is so-called, a suboptimal closed-loop control design. In the LQR, since the **A**, **B**, **Q** and **R** are linear and constant, those derivative terms do not exist and the resultant gain **K** is the optimal answer to the linear system.

An SDRE controller is suitable for cases that need an optimal design when the finishing time is unimportant and the steady-state behavior of the system is more demanded. An SDDRE controller penalizes the states at a final time which is usually less than the conventional time of operation (when the user requires a faster response). The final time is a function of control effort that the designer invests in the platform. More energy results in a faster response; and the fastest response is limited by the actuator's saturations and limits. The other aspect that defines the final time is the amplitude of the error in regulation (point-to-point) control and the length and speed of a trajectory in the tracking case. That means if the error in regulation is big, more time is needed and if the error is small, less time.

- The question is how to determine the final time? Using the SDRE, without penalization of the final boundary condition delivers an infinite-time optimal control solution as a function of error and the convergence-time is limited to the bound of actuators and tuning of the weighting matrices. In that sense, we realize approximately how much time we need for a specific task in regulation between two set points.
- After that how to finish the control task faster? Using the SDDRE with penalizing the final boundary condition. The error at the end of the regulation gets smaller and the amplitude of the input signal decreases which shows an asymptotic behavior in infinite time controllers. Using the finite time approach, the regulation near the end needs to speed up and we see a faster response, which leads to a shorter finishing time. Matrix **F** plays this role and enhances regulation speed.

It should be noted that in SDDRE closed-loop optimal control, at the final time, there is an error but the controller equipped with ILC will reduce it, see Fig. 1. The exact final time with zero error is possible by open-loop two-point boundary value problem [40, 41], etc.; however, using closed-loop optimal control, the exact value without error is impossible.

As a result, the signal of the SDDRE control law increases near the end of the time of operation. The solution to the differential Riccati equation, SDDRE, is a little more complicated than the SDRE (SDRE is an algebraic matrix equation). The more complicated structure was not a limit to the application of this method in control of a super-tanker [42], aircraft [43], wind energy conversion system [44], helicopter [45], etc.

Lemma 1. The nonlinear system (1) with its performance index (2) and necessary conditions based on Assumptions 1 and

2 can be stabilized using control law (3), in which $\mathbf{K}_{ss}(\mathbf{x}(t))$ is the positive-definite solution to the SDRE (4) [46].

Lemma 2. The nonlinear system (1) with its cost function (2) and necessary conditions based on Assumptions 1 and 2 can be stabilized using control law (3), in which $\mathbf{K}(\mathbf{x}(t))$ is the positive-definite solution to the SDDRE (5) with the final condition $\mathbf{K}(t_f) = \mathbf{F}$ [33].

Remark 1. The control law can employ a solution to the SDRE (4) which generates a steady-state suboptimal gain or the SDDRE (5) which releases a time-varying suboptimal gain with the benefit of finite-time control. The stability proofs of Lemmas 1 and 2 used a conventional Lyapunov function $V = \mathbf{x}^T\mathbf{K}\mathbf{x}$. The stability of the proposed PD-type SDRE and ILC law cannot be proved by the common function since it generates extra sign indefinite terms that are difficult to analyze:

$$\dot{V} = -\mathbf{x}^T(\mathbf{K}\mathbf{B}\mathbf{R}^{-1}\mathbf{B}^T\mathbf{K} + \mathbf{Q})\mathbf{x} + \mathbf{x}^T\mathbf{K}\mathbf{B}\mathbf{H} + \mathbf{H}^T\mathbf{B}^T\mathbf{K}\mathbf{x},$$

where $\mathbf{H}$ is a learning part. A different Lyapunov-like function has been considered here, stated in Theorem 1.

The closed-form solution to the SDDRE (5) was reported in Sections 3.3.1 and 3.3.2 of Ref. [33].

III. SDRE/SDDRE Augmented by Iterative Learning

A. Regulation

Consider $\mathbf{q}(t) \in \mathbb{R}^N$ as a generalized coordinate vector of an arbitrary N degree-of-freedom (DoF) mechanical system with an equation of motion:

$$\mathbf{M}(\mathbf{q}(t))\ddot{\mathbf{q}}(t) + \mathbf{C}(\mathbf{q}(t), \dot{\mathbf{q}}(t))\dot{\mathbf{q}}(t) + \mathbf{g}(\mathbf{q}(t)) + \mathbf{D}_f\dot{\mathbf{q}}(t) = \mathbf{u}(t), \quad (6)$$

where $\mathbf{M}(\mathbf{q}(t)): \mathbb{R}^N \to \mathbb{R}^{N\times N}$ is an inertia matrix, $[\mathbf{C}(\mathbf{q}(t), \dot{\mathbf{q}}(t))\dot{\mathbf{q}}(t)]: \mathbb{R}^N \to \mathbb{R}^N$ includes Coriolis and centrifugal terms, $\mathbf{g}(\mathbf{q}(t)): \mathbb{R}^N \to \mathbb{R}^N$ is a gravity vector, $\mathbf{D}_f \in \mathbb{R}^{N\times N}$ is a positive-diagonal matrix (including viscous friction, drag, etc.) and $\mathbf{u}(t) \in \mathbb{R}^N$ is an input vector consisting of $F_i(t)$ (N) for force or $\tau_i(t)$ (Nm) for torque.

Property 1. The inertia matrix $\mathbf{M}(\mathbf{q}(t))$ is a symmetric positive-definite matrix such that $\mathbf{M}_1 \le \mathbf{M}(\mathbf{q}(t)) \le \mathbf{M}_2$ where $\mathbf{M}_1, \mathbf{M}_2 \in \mathbb{R}^{N\times N}$ are constant symmetric positive definite matrices with the same size of $\mathbf{M}(\mathbf{q}(t))$.

Property 2. Defining the Coriolis and centrifugal force matrix through Christoffel's symbols, the matrix $[\dot{\mathbf{M}}(\mathbf{q}(t)) - 2\mathbf{C}(\mathbf{q}(t), \dot{\mathbf{q}}(t))]$ is skew-symmetric and the following holds $\boldsymbol{\eta}^T(t)[\dot{\mathbf{M}}(\mathbf{q}(t)) - 2\mathbf{C}(\mathbf{q}(t), \dot{\mathbf{q}}(t))]\boldsymbol{\eta}(t) = 0$ where $\boldsymbol{\eta}(t) \in \mathbb{R}^N$.

A state vector (including generalized coordinates and the derivative of that) for the system is regarded as $[\mathbf{x}(t)]_{2N\times 1} = [\mathbf{q}^T(t) \quad \dot{\mathbf{q}}^T(t)]^T$, $n = 2N$, and $m = N$. As a result, the state-space equation for the representation of the mechanical system (6) is formed (leaving gravity out of the state-space equation to be compensated by the control law) [39]:

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}(\mathbf{x}(t))\mathbf{x}(t) + \mathbf{B}(\mathbf{x}(t))\mathbf{u}(t)\\ &= \begin{bmatrix}\mathbf{0}_{N\times N} & \mathbf{I}_{N\times N}\\ \mathbf{0}_{N\times N} & -\mathbf{M}^{-1}(\mathbf{x})[\mathbf{C}(\mathbf{x}) + \mathbf{D}_f]\end{bmatrix}\mathbf{x}(t)\\ &+ \begin{bmatrix}\mathbf{0}_{N\times N}\\ \mathbf{M}^{-1}(\mathbf{x})\end{bmatrix}\mathbf{u}(t). \end{aligned} \quad (7)$$

The SDRE is augmented by an iterative learning control law, for the i-th learning loop as:

$$\mathbf{u}^i(t) = \mathbf{u}_S^i(t) + \mathbf{g}^i(\mathbf{q}(t)) + \mathbf{H}^i(t), \quad (8)$$

where $\mathbf{H}^i(t)$ is the learning feedforward term and the feedback part is gravity vector plus SDRE/SDDRE ($\mathbf{u}_S(t)$) control law

$$\mathbf{u}_S(t) = -\mathbf{R}^{-1}(\mathbf{q}, \dot{\mathbf{q}})\mathbf{B}^T(\mathbf{q}, \dot{\mathbf{q}})\mathbf{K}(\mathbf{q}, \dot{\mathbf{q}})[\mathbf{e}^T, \dot{\mathbf{e}}^T]^T, \quad (9)$$

where $\mathbf{e}(t) = \mathbf{q}(t) - \mathbf{q}_d$ and $\dot{\mathbf{e}}(t) = \dot{\mathbf{q}}(t) - \dot{\mathbf{q}}_d$ in which $\mathbf{q}_d$ and $\dot{\mathbf{q}}_d$ are desired position and velocity of states at the final time. The updating term $\mathbf{H}^i(t)$ is a predicted feed-forward part. It will be computed by an iterative learning law at each loop.

Assumption 3. The updated term $\mathbf{H}^i(t)$ of the training rule satisfies the following bound

$$\|\mathbf{H}^i(t)\|_2 \le c_0 + c_1\|\mathbf{e}(t)\|_2 + c_2\|\dot{\mathbf{e}}(t)\|_2, \quad (10)$$

with constants $c_0, c_1, c_2 > 0$ and any positive integer i.

Assumption 3 puts constraints on the updating term $\mathbf{H}(t)$ in the input law; any kind of actuator is physically limited. The input law generates a signal to feed one or more actuators, that might be DC motors, a hydraulics valve, a propeller rotor, thermo-electric valves, etc. The physical interpretability of Assumption 3 could be expressed as an internal bound on a learning term. That means $\mathbf{H}(t)$ has a relatively limited capability of learning concerning the PD-type counterpart of the controller, such that stability is guaranteed during the iterations.

The gain $\mathbf{K}(\mathbf{q}(t), \dot{\mathbf{q}}(t))$ is a $2N \times 2N$ symmetric-positive-definite matrix, a nonlinear optimal solution to the SDRE/SDDRE, partitioned into four square blocks:

$$\mathbf{K}(\mathbf{q}, \dot{\mathbf{q}}) = \left[\begin{array}{c|c}\mathbf{K}_{11}(\mathbf{q}, \dot{\mathbf{q}}) & \mathbf{K}_{12}(\mathbf{q}, \dot{\mathbf{q}})\\ \hline \mathbf{K}_{12}^T(\mathbf{q}, \dot{\mathbf{q}}) & \mathbf{K}_{22}(\mathbf{q}, \dot{\mathbf{q}})\end{array}\right]. \quad (11)$$

Considering the special form of $\mathbf{B}(\mathbf{x}(t))$ for the mechanical system (7), and using Eq. (11), the control law (9) is changed into:

$$\mathbf{u}_S(t) = -\mathbf{X}_1(\mathbf{q}, \dot{\mathbf{q}})\mathbf{e} - \mathbf{X}_2(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{e}}, \quad (12)$$

where

$$[\mathbf{X}_1(\mathbf{q}, \dot{\mathbf{q}})]_{N\times N} = \mathbf{R}^{-1}(\mathbf{q}, \dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})\mathbf{K}_{12}^T(\mathbf{q}, \dot{\mathbf{q}}), \quad (13)$$

$$[\mathbf{X}_2(\mathbf{q}, \dot{\mathbf{q}})]_{N\times N} = \mathbf{R}^{-1}(\mathbf{q}, \dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})\mathbf{K}_{22}(\mathbf{q}, \dot{\mathbf{q}}). \quad (14)$$

Equation (12) is another representation of SDDRE (9) preserving nonlinearity, finite-time, and suboptimal characteristics. The matrix $\mathbf{K}_{12}(\mathbf{q}, \dot{\mathbf{q}})$, in (13), is not symmetric positive definite, hence $\mathbf{X}_1(\mathbf{q}, \dot{\mathbf{q}})$ is not positive definite as well. To prove the stability of the controller in Theorem 1, symmetric positive definiteness of the gains is necessary. We use a transformation to present two sets of symmetric positive definite based on the solution to Riccati equation to preserve

finite-time characteristics and benefits of the SDDRE and at the same time, satisfy the condition for stability proof. The final step is to transform the gains (13) and (14), to symmetric positive-definite matrices

$$\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}}) = \frac{\mathbf{K}_{12}(\mathbf{q},\dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})\mathbf{R}^{-1}(\mathbf{q},\dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})\mathbf{K}_{12}^{T}(\mathbf{q},\dot{\mathbf{q}})}{\|\mathbf{K}_{12}(\mathbf{q},\dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})\|_2},$$
$$\mathbf{K}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}}) = \frac{\mathbf{K}_{22}^{T}(\mathbf{q},\dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})\mathbf{R}^{-1}(\mathbf{q},\dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})\mathbf{K}_{22}(\mathbf{q},\dot{\mathbf{q}})}{\|\mathbf{K}_{22}(\mathbf{q},\dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})\|_2},$$

where multiplying $\mathbf{K}_{12}(\mathbf{q},\dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})$ and $\mathbf{K}_{22}^{T}(\mathbf{q},\dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})$ from left-hand-side makes $\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}})$ and $\mathbf{K}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}})$ symmetric-positive-definite gains, leveled by the division to norms $\|\mathbf{K}_{12}(\mathbf{q},\dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})\|_2$ and $\|\mathbf{K}_{22}(\mathbf{q},\dot{\mathbf{q}})\mathbf{M}^{-1}(\mathbf{q})\|_2$ with respect. It also reshapes the control law (9) to a new symmetric design:

$$\mathbf{u}_{\mathrm{S}}(t) = -\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}})\mathbf{e} - \mathbf{K}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{e}}. \tag{15}$$

So, control law (15) is transformed and is not equal to (12) and the result of this transformation will be used to cancel two terms in the proof of stability in Theorems 1 and 2, using $\dot{\mathbf{q}}^T\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}})\mathbf{e} = \mathbf{e}^T\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}})\dot{\mathbf{q}}$, which is not possible without symmetry of $\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}})$.

Symmetric design and stability analysis: The idea of the symmetric design of the SDRE/SDDRE controller is to reach a PD-like structure to independently control the error and error velocity of a practical system, and analytically prove the stability of the novel ILC controller. The design also preserves the finite-time control capability of the SDDRE. The new structure helps to guarantee the stability of the nonlinear equation of motion (6) of a mechanical system. The advantages of the symmetric design are the optimality and finite-time characteristic of the controller.

Theorem 1. (*Regulation*) Consider the nonlinear mechanical system dynamics given by (6) satisfying Properties 1 and 2 together with the SDDRE dynamic-state feedback (8)-(15) equipped with the suboptimal gain of Eq. (5) and $\mathbf{R} = \left(\frac{\mathbf{R}_0}{r(t)}\right)^{-1}$, $\mathbf{R}_0 = \mathbf{R}_0^T > 0$, set through the scaling-factor dynamics as

$$\dot{r} = \rho_1 r \frac{\left\|2\boldsymbol{\alpha} - \frac{\dot{\gamma}}{\gamma}\boldsymbol{\beta}\right\|_2}{\lambda_{\min}(\boldsymbol{\beta})} + \rho_2(\mathbf{e},\dot{\mathbf{e}},r), \quad r(0) > \|\mathbf{R}_0\|, \tag{16}$$

with $\rho_1 > 1$, any smooth function $\rho_2(\mathbf{e},\dot{\mathbf{e}},r)$ with at most linear growth in r and $\boldsymbol{\alpha}(\mathbf{q},\dot{\mathbf{q}})$, $\boldsymbol{\beta}(\mathbf{q},\dot{\mathbf{q}})$, and $\gamma(\mathbf{q},\dot{\mathbf{q}})$ defined along with the proof. Then, the error trajectories $\mathbf{e}(t)$ are ultimately bounded for any training rule satisfying Assumption 3 for all $0 \le t \le t_{\mathrm{f}}$ and any finite final time t_{f}.

Proof. In the regulation case, the desired velocity and acceleration are zero at the final time, $\ddot{\mathbf{e}}(t) = \ddot{\mathbf{q}}(t)$, $\dot{\mathbf{e}}(t) = \dot{\mathbf{q}}(t)$ and $\mathbf{e}(t) = \mathbf{q}(t) - \mathbf{q}_{\mathrm{d}}$ where $\mathbf{q}_{\mathrm{d}}$ is a constant vector. Substitution of the iterative control law (8) in the equation of motion (6) results in

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + [\mathbf{C}(\mathbf{q},\dot{\mathbf{q}}) + \mathbf{D}_{\mathrm{f}} + \mathbf{K}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}},r)]\dot{\mathbf{q}} + \mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e} = \mathbf{H}. \tag{17}$$

To analyze the stability, a Lyapunov-like candidate is considered

$$V(\mathbf{q},\dot{\mathbf{q}},r) = \frac{1}{2}\{\dot{\mathbf{q}}^T\mathbf{M}(\mathbf{q})\dot{\mathbf{q}} + \mathbf{e}^T\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e}\}. \tag{18}$$

It should be noted that Eq. (17) is a non-homogenous second order differential equation that is updated in each loop by the time-varying term $\mathbf{H}$. $\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)$ is positive definite based on the proposed symmetric structure. Taking time-derivative of (18) provides

$$\dot{V}(\mathbf{q},\dot{\mathbf{q}},r) = \dot{\mathbf{q}}^T\mathbf{M}(\mathbf{q})\ddot{\mathbf{q}} + \frac{1}{2}\dot{\mathbf{q}}^T\dot{\mathbf{M}}(\mathbf{q})\dot{\mathbf{q}} + \mathbf{e}^T\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\dot{\mathbf{e}} + \frac{1}{2}\mathbf{e}^T\dot{\mathbf{K}}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e}, \tag{19}$$

and replacing $\dot{\mathbf{e}} = \dot{\mathbf{q}}$ and $\ddot{\mathbf{q}}$ from (17) into (19) results in

$$\dot{V}(\mathbf{q},\dot{\mathbf{q}},r) = \frac{1}{2}\dot{\mathbf{q}}^T\{\dot{\mathbf{M}}(\mathbf{q}) - 2\mathbf{C}(\mathbf{q},\dot{\mathbf{q}})\}\dot{\mathbf{q}} - \dot{\mathbf{q}}^T[\mathbf{D}_{\mathrm{f}} + \mathbf{K}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}},r)]\dot{\mathbf{q}} - \dot{\mathbf{q}}^T\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e} + \mathbf{e}^T\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\dot{\mathbf{q}} + \frac{1}{2}\mathbf{e}^T\dot{\mathbf{K}}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e} + \dot{\mathbf{q}}^T\mathbf{H}.$$

It can be easily shown that $\dot{\mathbf{q}}^T\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e} = \mathbf{e}^T\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\dot{\mathbf{q}}$ because $\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)$ is a symmetric matrix and that $\dot{\mathbf{M}}(\mathbf{q}) - 2\mathbf{C}(\mathbf{q},\dot{\mathbf{q}})$ is a skew-symmetric matrix (Property 2). Consequently, the derivative of the Lyapunov-like candidate turns into

$$\dot{V}(\mathbf{q},\dot{\mathbf{q}},r) = -\dot{\mathbf{e}}^T[\mathbf{D}_{\mathrm{f}} + \mathbf{K}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}},r)]\dot{\mathbf{e}} + \frac{1}{2}\mathbf{e}^T\dot{\mathbf{K}}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e} + \dot{\mathbf{e}}^T\mathbf{H},$$

in which

$$\dot{\mathbf{K}}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r) = \left\{\left[2\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{K}_{12}\mathbf{M}^{-1})\frac{\mathbf{R}_0}{r}\mathbf{M}^{-1}\mathbf{K}_{12}^T + \mathbf{K}_{12}\mathbf{M}^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathbf{R}_0}{r}\right)\mathbf{M}^{-1}\mathbf{K}_{12}^T\right]\|\mathbf{K}_{12}\mathbf{M}^{-1}\|_2 - \frac{\mathrm{tr}\left\{(\mathbf{K}_{12}\mathbf{M}^{-1})^T\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{K}_{12}\mathbf{M}^{-1})\right\}}{\sqrt{\mathrm{tr}\{(\mathbf{K}_{12}\mathbf{M}^{-1})^T(\mathbf{K}_{12}\mathbf{M}^{-1})\}}}\left(\mathbf{K}_{12}\mathbf{M}^{-1}\frac{\mathbf{R}_0}{r}\mathbf{M}^{-1}\mathbf{K}_{12}^T\right)\right\} / \|\mathbf{K}_{12}\mathbf{M}^{-1}\|_2^2,$$

and could be rewritten as

$$\dot{\mathbf{K}}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r) = \frac{1}{r^2\gamma}\left\{-\dot{r}\boldsymbol{\beta} + r\left(2\boldsymbol{\alpha} - \frac{\dot{\gamma}}{\gamma}\boldsymbol{\beta}\right)\right\}, \tag{20}$$

where we have defined

$$\boldsymbol{\alpha} \coloneqq \frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{K}_{12}\mathbf{M}^{-1})\mathbf{R}_0\mathbf{M}^{-1}\mathbf{K}_{12}^T, \qquad \boldsymbol{\beta} \coloneqq \mathbf{K}_{12}\mathbf{M}^{-1}\mathbf{R}_0\mathbf{M}^{-1}\mathbf{K}_{12}^T,$$
$$\gamma \coloneqq \|\mathbf{K}_{12}\mathbf{M}^{-1}\|_2, \qquad \dot{\gamma} = \frac{\mathrm{tr}\left\{(\mathbf{K}_{12}\mathbf{M}^{-1})^T\frac{\mathrm{d}}{\mathrm{d}t}(\mathbf{K}_{12}\mathbf{M}^{-1})\right\}}{\sqrt{\mathrm{tr}\{(\mathbf{K}_{12}\mathbf{M}^{-1})^T(\mathbf{K}_{12}\mathbf{M}^{-1})\}}}.$$

Recalling the definition of the scaling-factor dynamics given in Theorem 1, Eq. (16), the following bound for derivative

arises

$$\dot{V}(\mathbf{q},\dot{\mathbf{q}},r) \le -\dot{\mathbf{e}}^T\left[\mathbf{D}_\mathrm{f}+\frac{\bar{\mathbf{K}}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}})}{r}\right]\dot{\mathbf{e}}-\frac{c_3}{r}\|\mathbf{e}\|_2^2 - \frac{\rho_2}{2r^2\gamma}\mathbf{e}^T\boldsymbol{\beta}\mathbf{e}+\dot{\mathbf{e}}^T\mathbf{H}, \tag{21}$$

where c_3 is a positive constant, whose existence is guaranteed because the ρ_1-term of (16) dominates $r\left(2\boldsymbol{\alpha}-\frac{\dot{\gamma}}{\gamma}\boldsymbol{\beta}\right)$ of (20), and where for compactness we have defined

$$\bar{\mathbf{K}}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}}) \coloneqq \mathbf{K}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}},1) > 0.$$

By construction, Eq. (16) is linear in $r(t)$, as a result, $r(t)$ is not escaping to infinity in finite time, and consequently $r(t) \in \mathcal{L}_{\infty \mathrm{e}}$ for any fixed t_f, which is the case in SDDRE ("e" stands for extended space). Therefore, for any training rule satisfying Assumption 3, i.e. bounded as in (10) and using Young's inequality, an upper bound for (21) becomes

$$\dot{V}(\mathbf{q},\dot{\mathbf{q}},r) \le -c_4\|\mathbf{e}\|_2^2 - c_5\|\dot{\mathbf{e}}\|_2^2 + \frac{c_0^2}{2}, \tag{22}$$

where the existence of the positive constants c_4, c_5 is guaranteed by the positivity of $\bar{\mathbf{K}}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}})$ and $c_3 > 0$. Additionally, the free function $\rho_2(\mathbf{e},\dot{\mathbf{e}},r)$ must satisfy the constraint $\frac{\mathbf{e}^T\boldsymbol{\beta}\mathbf{e}}{2r^2\gamma}\,|\rho_2(\mathbf{e},\dot{\mathbf{e}},r)| < \lambda_{\min}(\mathbf{D}_\mathrm{f})\,\|\dot{\mathbf{e}}\|_2^2$, e.g. it could be set to zero (its role is explained in Remark 2). On the other hand, recall that $\mathbf{M}(\mathbf{q})$ is positive definite and $r(t) \in \mathcal{L}_{\infty \mathrm{e}}$ for a fixed t_f. Thus, for any fixed finite time t_f, the function (18) is positive definite, and therefore, there always exists a constant $c > 0$ such that (22) yields

$$\dot{V}(\mathbf{q},\dot{\mathbf{q}},r) \le -cV(\mathbf{q},\dot{\mathbf{q}},r) + \frac{c_0^2}{2}. \tag{23}$$

Finally, from (23) for $V \ge \frac{c_0^2}{2c}$ then $\dot{V} \le 0$ and the ultimate boundedness along the trajectories of (17) is guaranteed, for $0 \le t \le t_\mathrm{f}$, i.e. $\mathbf{e}, \dot{\mathbf{e}} \in \mathcal{L}_{\infty e}$. ∎

Corollary 1. Suppose that all the conditions of Theorem 1 hold. Then, if $c_0 = 0$ in (10) of Assumption 3, the error trajectories $\mathbf{e}(t)$ are uniformly ultimately bounded for all $0 \le t \le t_\mathrm{f}$ and any finite final time t_f.

Remark 2. Notice that, the function $\rho_2(\mathbf{e},\dot{\mathbf{e}},r)$ is not needed for the stability proof, and it could be set to zero. This function modulates the size of $r(t)$ in practice. Recall that the mechanism of the scaling factor is to grow and dominate the undesired nonlinearities. Thus, in practice, it is very important to add a stabilizing term to the $r(t)$ dynamics to modulate its size in such a way that we prevent the computer hardware from an overflow in digital implementations. In the simulation section, we discuss thoroughly this and propose a form for the function $\rho_2(\mathbf{e},\dot{\mathbf{e}},r)$ while preserving stability.

Remark 3. Selection of $\mathbf{R}_0$ and $r(0)$ defines the performance of the system. Regarding the definition of $\mathbf{R}(0) = \left(\frac{\mathbf{R}_0}{r(0)}\right)^{-1}$, one should note that $\mathbf{R}(0)$ must be defined as the proper selection of the input weighting matrix for systems. Large values for $\mathbf{R}(0)$ result in an improper performance. $r(0)$ also affects the solution to $r(t)$, and finally, $r(t)$ must be positive.

B. Tracking

Adding and subtracting the desired dynamic to Eq. (6) change the equation of motion to (subscript "d" represents "desired"):

$$\begin{aligned}[\mathbf{M}(\mathbf{q})-\mathbf{M}(\mathbf{q}_\mathrm{d})](\ddot{\mathbf{q}}-\ddot{\mathbf{q}}_\mathrm{d}) &+ [\mathbf{C}(\mathbf{q},\dot{\mathbf{q}})-\mathbf{C}(\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d})](\dot{\mathbf{q}}-\dot{\mathbf{q}}_\mathrm{d}) \\ &+ \mathbf{g}(\mathbf{q}) + \mathbf{D}_\mathrm{f}(\dot{\mathbf{q}}-\dot{\mathbf{q}}_\mathrm{d}) \\ &= \mathbf{u} - \mathbf{S}_\mathrm{d}(\ddot{\mathbf{q}}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d},\mathbf{q}_\mathrm{d}),\end{aligned} \tag{24}$$

where

$$\mathbf{S}_\mathrm{d}(\ddot{\mathbf{q}}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d},\mathbf{q}_\mathrm{d}) = \mathbf{M}(\mathbf{q}_\mathrm{d})\ddot{\mathbf{q}}_\mathrm{d} + \mathbf{C}(\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d})\dot{\mathbf{q}}_\mathrm{d} + \mathbf{D}_\mathrm{f}\dot{\mathbf{q}}_\mathrm{d}. \tag{25}$$

Defining the error and error velocity vector of the system as $\mathbf{e}(t) = \mathbf{q}(t) - \mathbf{q}_\mathrm{d}(t)$ and $\dot{\mathbf{e}}(t) = \dot{\mathbf{q}}(t) - \dot{\mathbf{q}}_\mathrm{d}(t)$, respectively, and substituting ILC law (8) in (24) result in

$$\begin{aligned}\tilde{\mathbf{M}}(\mathbf{q},\mathbf{q}_\mathrm{d})\ddot{\mathbf{e}} + \left[\tilde{\mathbf{C}}(\mathbf{q},\dot{\mathbf{q}},\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d}) + \mathbf{D}_\mathrm{f} + \mathbf{K}_{\mathrm{SD}}(\mathbf{q},\dot{\mathbf{q}})\right]\dot{\mathbf{e}} \\ + \mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}})\mathbf{e} = \mathbf{H} - \mathbf{S}_\mathrm{d}(\ddot{\mathbf{q}}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d},\mathbf{q}_\mathrm{d}),\end{aligned} \tag{26}$$

where $\tilde{\mathbf{M}}(\mathbf{q},\mathbf{q}_\mathrm{d}) = \mathbf{M}(\mathbf{q}) - \mathbf{M}(\mathbf{q}_\mathrm{d})$, and $\tilde{\mathbf{C}}(\mathbf{q},\dot{\mathbf{q}},\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d}) = \mathbf{C}(\mathbf{q},\dot{\mathbf{q}}) - \mathbf{C}(\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d})$.

Property 3. The desired dynamic in trajectory tracking is a copy of the original mechanical structure, i.e. $\mathbf{M}(\mathbf{q})$ and $\mathbf{C}(\mathbf{q},\dot{\mathbf{q}})$. Thus, defining the Coriolis and centrifugal force matrix $\mathbf{C}(\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d})$ through Christoffel's symbols, the property $\boldsymbol{\eta}^T\left[\dot{\tilde{\mathbf{M}}}(\mathbf{q},\mathbf{q}_\mathrm{d}) - 2\tilde{\mathbf{C}}(\mathbf{q},\dot{\mathbf{q}},\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d})\right]\boldsymbol{\eta} = 0$ holds, as in Property 2.

Assumption 4. The predefined trajectories for tracking, $\mathbf{q}_\mathrm{d}(t)$, are uniformly continuous and bounded in $0 \le t \le t_\mathrm{f}$, twice continuously differentiable and with $\dot{\mathbf{q}}_\mathrm{d}(t)$ and $\ddot{\mathbf{q}}_\mathrm{d}(t)$, also bounded. The final time of the trajectory t_f is fixed.

Theorem 2. (*Tracking*) Consider the nonlinear dynamics represented by (24) in the trajectory tracking case, satisfying Properties 1 and 3 together with the SDDRE dynamic-state feedback (8)-(15) with the suboptimal gain of Eq. (5) and $\mathbf{R} = \left(\frac{\mathbf{R}_0}{r(t)}\right)^{-1}$, set through the scaling-factor dynamics (16) with same conditions for ρ_1, $\rho_2(\mathbf{e},\dot{\mathbf{e}},r)$, $\boldsymbol{\alpha}(\mathbf{q},\dot{\mathbf{q}})$, $\boldsymbol{\beta}(\mathbf{q},\dot{\mathbf{q}})$, and $\gamma(\mathbf{q},\dot{\mathbf{q}})$ as in Theorem 1. Then, the error vector $\mathbf{e}(t)$ is ultimately bounded for any tracking training rule satisfying Assumption 3 and for any trajectory under Assumption 4, for all $0 \le t \le t_\mathrm{f}$ and any finite final time t_f.

Proof. The stability proof is very similar to the regulation case and hence we provide only the necessary steps. In this case, the Lyapunov-like candidate is

$$V(\mathbf{q},\dot{\mathbf{q}},\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d},r) = \frac{1}{2}\left\{\dot{\mathbf{e}}^T\tilde{\mathbf{M}}(\mathbf{q},\mathbf{q}_\mathrm{d})\dot{\mathbf{e}} + \mathbf{e}^T\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e}\right\}. \tag{27}$$

Taking time-derivative of (27) provides

$$\begin{aligned}\dot{V}(\mathbf{q},\dot{\mathbf{q}},\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d},r) = \dot{\mathbf{e}}^T\tilde{\mathbf{M}}(\mathbf{q},\mathbf{q}_\mathrm{d})\ddot{\mathbf{e}} + \frac{1}{2}\dot{\mathbf{e}}^T\dot{\tilde{\mathbf{M}}}(\mathbf{q},\mathbf{q}_\mathrm{d})\dot{\mathbf{e}} \\ + \mathbf{e}^T\mathbf{K}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\dot{\mathbf{e}} \\ + \frac{1}{2}\mathbf{e}^T\dot{\mathbf{K}}_{\mathrm{SP}}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e},\end{aligned} \tag{28}$$

and replacing $\ddot{\mathbf{e}}$ from (26) into (28) result in

$$\begin{aligned}\dot{V}(\mathbf{q},\dot{\mathbf{q}},\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d},r) = \frac{1}{2}\dot{\mathbf{e}}^T\left\{\dot{\tilde{\mathbf{M}}}(\mathbf{q},\mathbf{q}_\mathrm{d}) - 2\tilde{\mathbf{C}}(\mathbf{q},\dot{\mathbf{q}},\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d})\right\}\dot{\mathbf{e}} \\ - \dot{\mathbf{e}}^T[\mathbf{D}_\mathrm{f} + \mathbf{K}_\mathrm{SD}(\mathbf{q},\dot{\mathbf{q}},r)]\dot{\mathbf{e}} - \dot{\mathbf{e}}^T\mathbf{K}_\mathrm{SP}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e} \\ + \mathbf{e}^T\mathbf{K}_\mathrm{SP}(\mathbf{q},\dot{\mathbf{q}},r)\dot{\mathbf{e}} + \frac{1}{2}\mathbf{e}^T\dot{\mathbf{K}}_\mathrm{SP}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e} + \dot{\mathbf{e}}^T\mathbf{H}.\end{aligned}$$

Invoking Property 3, the derivative becomes

$$\begin{aligned}\dot{V}(\mathbf{q},\dot{\mathbf{q}},\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d},r) = -\dot{\mathbf{e}}^T\{\mathbf{D}_\mathrm{f} + \mathbf{K}_\mathrm{SD}(\mathbf{q},\dot{\mathbf{q}},r)\}\dot{\mathbf{e}} \\ + \frac{1}{2}\mathbf{e}^T\dot{\mathbf{K}}_\mathrm{SP}(\mathbf{q},\dot{\mathbf{q}},r)\mathbf{e} + \dot{\mathbf{e}}^T\mathbf{H},\end{aligned} \tag{29}$$

where definition of $\dot{\mathbf{K}}_\mathrm{SP}(\mathbf{q},\dot{\mathbf{q}},r)$ is similar to (20). Substituting the scaling-factor $\dot{r}(t)$, Eq. (16), into (29) results in

$$\begin{aligned}\dot{V}(\mathbf{q},\dot{\mathbf{q}},\mathbf{q}_\mathrm{d},\dot{\mathbf{q}}_\mathrm{d},r) \le -\dot{\mathbf{e}}^T\left\{\mathbf{D}_\mathrm{f} + \frac{\bar{\mathbf{K}}_\mathrm{SD}(\mathbf{q},\dot{\mathbf{q}})}{r}\right\}\dot{\mathbf{e}} - \frac{c_3}{r}\|\mathbf{e}\|_2^2 \\ - \frac{\rho_2}{2r^2\gamma}\mathbf{e}^T\boldsymbol{\beta}\mathbf{e} + \dot{\mathbf{e}}^T\mathbf{H}.\end{aligned}$$

that is analogous to (21). Therefore, under the trajectory boundedness Assumption 4, a similar result follows from (21) of Theorem 1, concluding the proof. ■

Remark 4. We would like to underscore that the training rules of the regulation and tracking cases are quite different. Focusing on Assumption 3, while in the regulation case $c_0 = 0$ as considered in Corollary 1, in the tracking case and under Assumption 4, the classical training rule in the literature satisfies Eq. (25) hence $c_1 = c_2 = 0$, in Assumption 3.

IV. Learning Rule

A. Regulation training rule

The gradient descent method is used to define a training rule for updating the iterative learning control $\mathbf{H}^i(t)$. Regarding that the regulation case seeks the minimum error at the final time, a performance index is regarded as

$$J_\mathrm{LR} = \frac{1}{2}\sum_{i=1}^{N_\mathrm{I}}\left\|\alpha\mathbf{H}^i(t) - \frac{1}{\alpha}\mathbf{W}^i(\mathbf{q},\dot{\mathbf{q}})\right\|^2, \tag{30}$$

where N_I is the total number of iterations, J_LR is convex, $0 < \alpha \ll 1$ and $\mathbf{W}^i(\mathbf{q},\dot{\mathbf{q}}) = -\mathbf{K}_\mathrm{SP}^i(\mathbf{q},\dot{\mathbf{q}})[\mathbf{q}^{i-1}(t_\mathrm{f}) - \mathbf{q}_\mathrm{d}] - \mathbf{K}_\mathrm{SD}^i(\mathbf{q},\dot{\mathbf{q}})[\dot{\mathbf{q}}^{i-1}(t_\mathrm{f}) - \dot{\mathbf{q}}_\mathrm{d}]$ in which $\dot{\mathbf{q}}_\mathrm{d}$ and $\mathbf{q}_\mathrm{d}$ are desired velocity and position of generalized coordinates, respectively; $\dot{\mathbf{q}}^i(t_\mathrm{f})$ and $\mathbf{q}^i(t_\mathrm{f})$ are actual ones at a final time of $(i-1)$-th loop since the final error, $\mathbf{q}(t_\mathrm{f})$, of the previous loop is accessible; $\mathbf{K}_\mathrm{SP}^i(\mathbf{q},\dot{\mathbf{q}})$ and $\mathbf{K}_\mathrm{SD}^i(\mathbf{q},\dot{\mathbf{q}})$ are the same symmetric gains in (15). Applying the gradient descent method on (30), the training rule is found:

$$\begin{aligned}\mathbf{H}^i = \mathbf{H}^{i-1} - \beta_\mathrm{R}\frac{\partial J_\mathrm{LR}}{\partial \mathbf{H}^i} \\ = \mathbf{H}^{i-1} \\ - \beta_\mathrm{R}\alpha\left\{\alpha\mathbf{H}^{i-1}(t) - \frac{1}{\alpha}\mathbf{W}^{i-1}(\mathbf{q},\dot{\mathbf{q}})\right\},\end{aligned} \tag{31}$$

in which $0 < \beta_\mathrm{R} < 1$ is a constant-scalar training factor for regulation (point-to-point motion control). The training rule of the regulation, (31), includes the error $\mathbf{e}(t_\mathrm{f})$ and error velocity $\dot{\mathbf{e}}(t_\mathrm{f})$ in $\mathbf{W}^{i-1}(\mathbf{q},\dot{\mathbf{q}})$, therefore the norm of regulation training could be considered bounded as $\|\mathbf{H}(t)\|_2 \le \bar{c}_1\|\mathbf{e}(t_\mathrm{f})\|_2 + \bar{c}_2\|\dot{\mathbf{e}}(t_\mathrm{f})\|_2 \le c_1\|\mathbf{e}(t)\|_2 + c_2\|\dot{\mathbf{e}}(t)\|_2$, stated in Assumption 3.

The gradient descent method converges if ∇J_LR is Lipschitz continuous, the step size is $1/L$ where $\|\nabla J_\mathrm{LR}(v) - \nabla J_\mathrm{LR}(w)\| < L\|v - w\|$ and J_LR is bounded from below. The first condition for $\nabla J_\mathrm{LR} = \alpha\mathbf{H}^{i-1}(t) - \frac{1}{\alpha}\mathbf{W}^{i-1}(\mathbf{q},\dot{\mathbf{q}})$ is satisfied since $\mathbf{W}(\mathbf{q},\dot{\mathbf{q}})$ possesses continuously differentiable gains and $\mathbf{H}(t)$ is also a time-varying vector that follows $\mathbf{W}(\mathbf{q},\dot{\mathbf{q}})$. The selection of step size is a matter of tuning and will be regarded in simulation. The last condition is also satisfied with the convex form of J_LR.

B. Tracking training rule

The performance index of the tracking problem is expressed based on the non-homogenous part of Eq. (24) as [23]:

$$J_\mathrm{LT} = \frac{1}{2}\sum_{i=1}^{N_\mathrm{I}}\left\|\mathbf{H}^i(t) - \mathbf{S}_\mathrm{d}^i(t)\right\|^2, \tag{32}$$

where $\mathbf{S}_\mathrm{d}^i(t)$ is the desired dynamics at i-th loop, Eq. (25). Implementation of the gradient descent method on (32) results in the following training rule:

$$\mathbf{H}^i = \mathbf{H}^{i-1} - \beta_\mathrm{T}\frac{\partial J_\mathrm{LT}}{\partial \mathbf{H}^{i-1}} = \mathbf{H}^{i-1} - \beta_\mathrm{T}\{\mathbf{H}^{i-1} - \mathbf{S}_\mathrm{d}^{i-1}\}, \tag{33}$$

in which $0 < \beta_\mathrm{T} < 1$ is a constant-scalar training factor for the trajectory tracking problem. The training rule of the tracking, (33), includes the desired dynamics in which a bounded trajectory updates the terms. As a result, the norm of tracking training could be considered bounded as $\|\mathbf{H}(t)\|_2 \le c_0$, stated in Assumption 3.

Remark 5. Consider the *continuous counterpart* of the training rule evolution, which is defined through the stable linear (ordinary differential equation) ODE as

$$\dot{\mathbf{H}}(t) = -\beta\mathbf{H}(t) + \beta\boldsymbol{\Phi}, \qquad 0 \le t \le t_\mathrm{f},$$

where $\boldsymbol{\Phi}$ plays the role of $\mathbf{W}^i$. Thus, for any bounded $\boldsymbol{\Phi}$ and $\beta > 0$, the solutions to this ODE can be bounded by a positive constant δ as follows

$$\begin{aligned}|\mathbf{H}(t_\mathrm{f})| \le |\mathbf{H}(0)| + \max_{0\le t\le t_\mathrm{f}}|\boldsymbol{\Phi}|\,(1 - \exp(-\beta t_\mathrm{f})) \\ \le |\mathbf{H}(0)| + \max_{0\le t\le t_\mathrm{f}}|\boldsymbol{\Phi}| \le \delta,\end{aligned}$$

where $|\cdot|$ denotes any norm. This means that, if the controller can maintain the internal stability at each iteration, then $\boldsymbol{\Phi}$ is bounded and so is $\mathbf{H}(t)$, which are the results of the Theorems. Therefore, coming back to the tracking example $\boldsymbol{\Phi}$ reads $\mathbf{S}_\mathrm{d}^i(t)$, that is bounded by a constant.

V. Simulations

A. Regulation

Consider a three-DoF spherical manipulator [47]. The parameters of the manipulator are presented in Table 1. The initial and final points are $A = (0.7,0,0.7)\mathrm{m}$ and $B = (0,-0.4,0.3)\mathrm{m}$ with respect. The final time is 4 seconds and 30

iterations are considered. The following control parameters are set: $\rho_1 = 1.5$, $\beta = 0.3$, $\alpha = 0.01$, $\mathbf{R}_0 = \mathbf{I}_{3\times3}$, $r_0 = 1.1 \times |\mathbf{R}_0|$, $\mathbf{Q} = 10 \times \mathrm{diag}(\mathbf{1}_{1\times3}, \mathbf{2}_{1\times3})$ and $\mathbf{F} = 5 \times \mathbf{Q}$. The $\rho_2(\mathbf{e}, \dot{\mathbf{e}}, r)$ term is defined by:

$$\rho_2(\mathbf{e}, \dot{\mathbf{e}}, r) \coloneqq -\bar{\rho}_2 \frac{2\gamma}{1 + \mathbf{e}^T \boldsymbol{\beta} \mathbf{e}} \lambda_{\min}(\mathbf{D}_f) \|\dot{\mathbf{e}}\|_2^2 \frac{r - r_0}{r}, \qquad (34)$$

where $r_0 \coloneqq r(0) = \|\mathbf{R}_0\|_2 > 1$, $\rho_1 > 1$ and $0 < \bar{\rho}_2 < \frac{27}{4} r_0^2$. The simulation results in successful end-effector error reduction. The configuration of the manipulator is shown in Fig. 2-a and the magnified view of the end-effector in Fig. 2-b. The end-effector error in Cartesian coordinates is also presented in Fig. 3. The end-effector error is defined as $E \coloneqq \sqrt{(x_e - x_d)^2 + (y_e - y_d)^2 + (z_e - z_d)^2}$, where $\{x_e, y_e, z_e\}$(m) is the Cartesian position of the end-effector of the manipulator, and $\{x_d, y_d, z_d\}$(m) is the desired one. Here the role of $\bar{\rho}_2$ is studied in terms of end-effector error. The proposed value is $\bar{\rho}_2 = \frac{27}{4} r_0^2 \sim 8.16$. The change in $\bar{\rho}_2$ affects error and different values were simulated to analyze the trend, presented in Table 2.

To justify the increase near the end of $r(t)$, we should emphasize that the SDDRE is a differential equation with final boundary condition, $\mathbf{K}(t_f) = \mathbf{F}$. So, near the end, we penalize the error to increase the precision, by weighting matrix $\mathbf{F}$. $r(t)$ is a scaling/forcing term that guarantees stability. It manipulates the weighting matrix $\mathbf{R} = \left(\frac{\mathbf{R}_0}{r(t)}\right)^{-1}$. So, since the control law tries to increase the signal near the end, $r(t)$ provides the balance by increasing the amplitude of $\mathbf{R}$. As it was explained in Remark 2, the nature of $r(t)$ in Eq. (34) is growing near the end of the final time and, hence, the extra ρ_2 term decreases that raises at the end (notice the negative sign of the term). This raise at the end directly increases the input weighting matrix $\mathbf{R} = \left(\frac{\mathbf{R}_0}{r(t)}\right)^{-1}$ near the end of the simulation where the error is negligible. This is against the general tuning policy of SDRE/SDDRE control design which states: smaller $\mathbf{R}$ results in less error. So, an increase in $\bar{\rho}_2$ corrects the tuning.

TABLE 1. THE PARAMETERS OF THE SPHERICAL MANIPULATOR.

parameter	value	unit	definition
d_1, d_{c1}	0.6,0.3	m	length and CoM
a_2, a_{c2}	0.5,0.25	m	length and CoM
a_3, a_{c3}	0.4,0.2	m	length and CoM
$m_{1,2,3}, m_p$	2,0.25	kg	mass of links, load
$D_{f,1,2,3}$	0.025	kgm/s	viscous friction
$I_{xx_1}, I_{yy_1}, I_{zz_1}$	$6.25 \times 10^{-4}, 6.25 \times 10^{-4}, 0.06$	kgm^2	moment of inertia, 1st link
$I_{xx_2}, I_{yy_2}, I_{zz_2}$	$0.0417, 6.25 \times 10^{-4}, 6.25 \times 10^{-4}$	kgm^2	moment of inertia, 2nd link
$I_{xx_3}, I_{yy_3}, I_{zz_3}$	$0.0267, 6.25 \times 10^{-4}, 6.25 \times 10^{-4}$	kgm^2	moment of inertia, 3rd link

TABLE 2. END-EFFECTOR ERROR IN VARIOUS VALUES OF $\bar{\rho}_2$ WITH $\rho_1 = 1.5$.

End-effector error (mm)	$\bar{\rho}_2$	End-effector error (mm)	$\bar{\rho}_2$
0.1268	0	0.0127	7
0.1144	0.5	0.0059	8.16
0.1026	1	0.0057	10
0.0629	3	0.0048	50

Comparative results are presented in Fig. 4. The reason for such involved extra ρ_2 term is because it has the bound constraint defined in Theorem 1 to guarantee stability. To see the latter, the derivative of the Lyapunov function is illustrated in Fig. 4 (b) where it can be seen that is negative semi-definite.

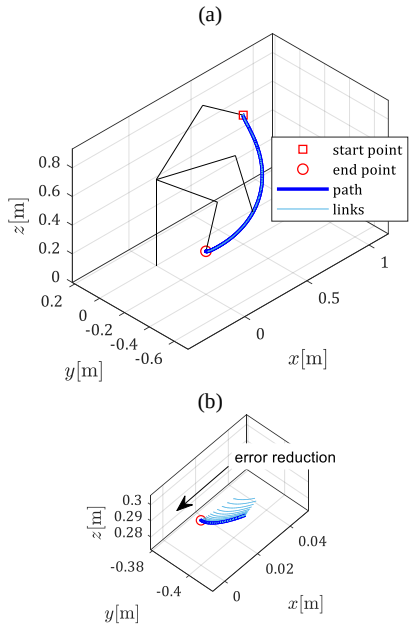

Fig. 2. The configuration of the manipulator; (b) The magnified view of the end-effector error reduction.

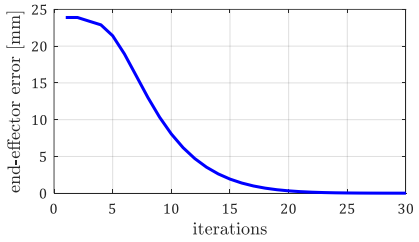

Fig. 3. End-effector errors in iterations.

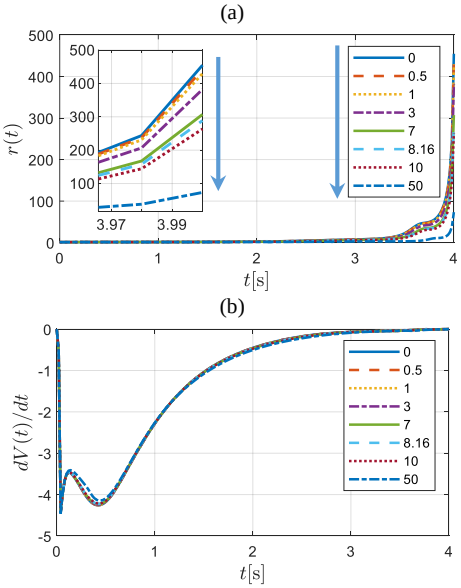

Fig. 4. (a) Forcing value to prove stability; (b) The derivative of Lyapunov function with different $\bar{\rho}_2$.

As it was pointed out in Remark 2 the simplest way to define this extra term is just $\rho_2 = 0$ but in practice, it might be necessary to have a trade-off to modulate its size. This simulation aims to show the design of the control law based on scaling parameter $r(t)$. A series of simulations were also provided to check the capability of the proposed approach in various conditions, Fig. 5-a shows the forcing value, and Fig. 5-b shows the derivative of the Lyapunov function.

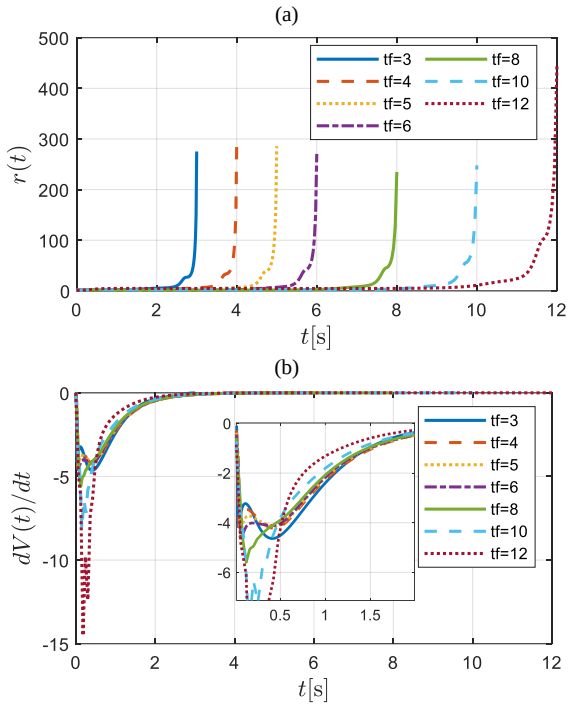

Fig. 5. (a) Forcing value; (b) derivative of Lyapunov function; in a series of different final times.

B. Comparison

To clarify the machinery of the PD-type SDDRE learning controller, a comparison has been done. We compare the proposed method with a PD-based ILC to highlight the effect of finite-time gain; and also, we compare the method with SDDRE and PD without learning capability to highlight the learning effect on error reduction. The error of the proposed method with similar parameters (30 iterations) in Section V-A was gained 0.0059mm. If we use simple PD + gravity control with $\mathbf{K}_\mathrm{p} = \mathbf{I}_{3\times3}$ and $\mathbf{K}_\mathrm{d} = 2 \times \mathbf{I}_{3\times3}$ gains and the same ILC support, the error will increase to 6.8(mm). Also removing the ILC from the control loop, the conventional SDDRE and PD + gravity resulted in 23.8(mm) and 93.2(mm) errors, respectively. Regarding 4(s) simulation time and 1500mm length of the three links, it is hard for the PD to reach the desired condition though the ILC improved its performance by 92%. So, comparing the results, the role of ILC on the closed-loop control is clearer. The ILC reduces the error by each iteration. The controllers, SDDRE or PD + gravity could indeed be tuned and operate properly, but the power of learning in error reduction and feedforward compensation of unpredicted situations is really helpful, especially in experiments.

C. Robustness and uncertainty

Uncertainty reduces the performance of the proposed controller and the ideal situation for the SDRE is to have the model as close as possible to the real platform. The source of uncertainty might be the friction (which is hard to estimate and model), the variation of load, and lack of precision in modeling. To assess the performance of the controller, this subsection is presented. We provide uncertainty to the controller by a change in mass of the load, set on the end-effector of the robot. In this scenario, we do not inform the controller about the load, so $m_\mathrm{p} = 0$, the mass of the load (kg) is zero in the controller, but in the system we set $m_\mathrm{p} = 0.1$kg. This increases the error from 0.0059mm to 1.67mm. Increasing the load more results in failure that shows the weakness towards the uncertainty (the black dotted line in Fig. 6).

Remedy. The SDRE could be modified to show robust characteristics such as incorporating the bounds of uncertainty in the SDC matrices, the addition of correction terms such as sliding mode control [48], or using additional features such as neural networks [49]. Here we add robustness through the definition of an upper bound for load in the controller. The SDC matrices will possess $m_{\mathrm{p,max}}$ and the dynamics might have another value, lower than the bound $m_\mathrm{p} \leq m_{\mathrm{p,max}}$. Considering $m_{\mathrm{p,max}} = 1.5$ and $m_\mathrm{p} = 1.2$, and also increasing the gains $\mathbf{Q} = 100 \times \mathrm{diag}(\mathbf{1}_{1\times3}, \mathbf{2}_{1\times3})$ and $\mathbf{F} = 20 \times \mathbf{Q}$, the end-effector error is gained 0.0054mm (the red dashed line in Fig. 6). The same uncertainty in friction could be addressed by incorporating the bounds in the SDC matrices (the green dashed-dotted line in Fig. 6). The higher error in the initial loop for the uncertainty simulation is due to more load on the end-effector. We emphasize that the focus of the paper is not on robustness and this topic deserves more discussions and a thorough investigation.

Robustness of the proposed SDRE + ILC. There are two mechanisms adding robustness: learning and dynamic scaling. The learning would find larger errors at the final time with disturbances, thus forcing a higher correction in the controller for the next iteration. However, this would only be possible if the core controller maintains the internal stability during each iteration, and this is the role of the scaling factor dynamics, dominating the mismatches and eventually disturbances.

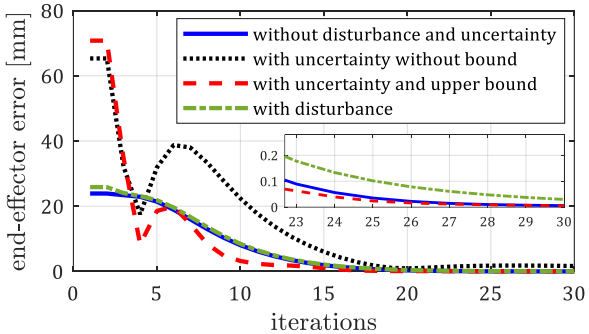

Fig. 6. Comparison of the results in presence of disturbance ad uncertainty with the ideal case.

D. Tracking

The trajectory is a line between $A(0.7,0.3,0.7)$m and $B(-0.651,0.57,0.7)$m to be tracked in 3 seconds, with 10 iterations, using training rule (33). The control gains are selected as $\rho_1 = 1.5, \bar{\rho}_2 = 8.16$, $\beta = 0.5$, $\alpha = 0.01$, $\mathbf{R}_0 = 0.1 \times \mathbf{I}_{3\times3}$, $r_0 = 1.1 \times |\mathbf{R}_0|$, $\mathbf{Q} = 1000 \times \mathrm{diag}(\mathbf{1}_{1\times3}, \mathbf{0.5}_{1\times3})$ and $\mathbf{F} = 10 \times \mathbf{Q}$. The error reduction of the end-effector is presented in Fig. 7. The reduction of error concerning the initial one presented a 25mm-to-0.7mm improvement in the precision of the tracking.

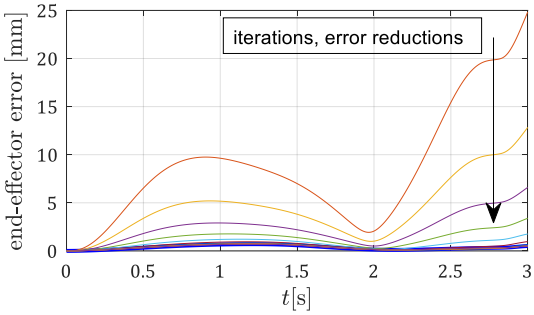

Fig. 7. Error reduction of the system in trajectory tracking case.

VI. Experiment

The experimental implementation of the ILC and SDDRE combination is presented to validate the proposed approach. The contribution of this work concerning Ref. [32] is that here the ILC augmented with SDRE controller has been implemented; nevertheless, in the previous work, only the SDRE was tested. For the first series of experiments, a stationary platform is selected, Fig. 8. The setup is a variable-pitch pendulum, rotating around the center-point of the system, represented by variable $\theta(t)$ (rad). The specifications are expressed in Table 3.

Table 3. The specifications of the variable-pitch benchmark [32].

Para.	value	unit	description
l	0.29	m	Dist. betw. CoM & rotor shaft
R	0.135	m	radius of propeller
I_{yy}	9.354×10^3	kg. m^2	moment of inertia
m	0.870	kg	total weight of the setup
k	3.32×10^{-6}	N. s^2/rad^2	lift constant - thrust factor

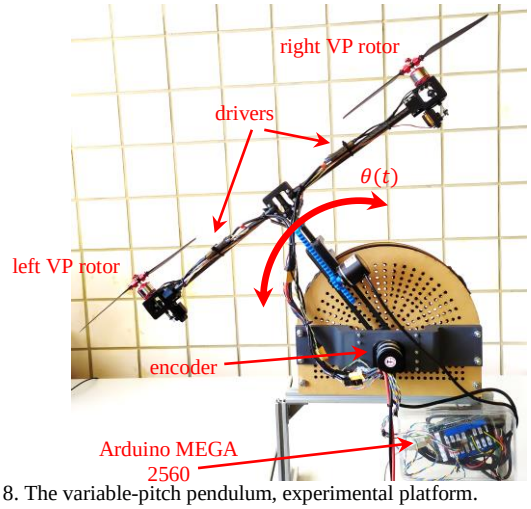

Fig. 8. The variable-pitch pendulum, experimental platform.

The dynamic model of the rotating system is $(mr_{\mathrm{ctr}}^2 + I_{yy})\ddot{\theta}(t) + mgr_{\mathrm{ctr}} \sin\theta(t) + D_{\mathrm{f}}\dot{\theta}(t) = \tau_\theta(t)$, where r_{ctr} (m) is the distance between the center of the pipe and CoM of the quadrotor, m (kg) is the weight of the setup, $g = 9.81$ (m/s^2) presents gravity constant, $D_{\mathrm{f}} = 0.25$ (kgm/s) is the friction and τ_θ(N. m) is the input torque. The state vector is chosen as $\mathbf{x}(t) = [\theta(t), \dot{\theta}(t)]^T$ in which $\dot{\theta}$ (rad/s) presents the angular velocity of the system around the pipe. Considering the dynamics, the state-space representation of the VP is

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \dot{\theta}(t) \\ \dfrac{\tau_\theta(t) - mgr_{\mathrm{ctr}} \sin\theta(t) - D_{\mathrm{f}}\dot{\theta}(t)}{mr_{\mathrm{ctr}}^2 + I_{yy}} \end{bmatrix}.$$

To implement the algorithm, the SDC parameterization matrices are defined $\mathbf{A}(\mathbf{x}(t)) = \begin{bmatrix} 0 & 1 \\ -c_1\left(1 - \frac{x_1^2(t)}{6} + \cdots\right) & -D_{\mathrm{f}}c_2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 0 \\ c_2 \end{bmatrix}$, in which $c_1 = \frac{mgr_{\mathrm{ctr}}}{mr_{\mathrm{ctr}}^2 + I_{yy}}$ and $c_2 = \frac{1}{mr_{\mathrm{ctr}}^2 + I_{yy}}$ [32].

The system initially was set at rest position on the left-hand side of the setup $\theta(0) = -0.6519$(rad), and the final condition was set on the equilibrium point. In the regulation case, the initial and final velocities were set at zero. The weighting matrices were selected as $\rho_1 = 1.5$, $r_0 = 1.1$, $R = 1$, $\mathbf{Q} = \begin{bmatrix} 190 & 0 \\ 0 & 5 \end{bmatrix}$ and $\mathbf{F} = 0.1 \times \mathbf{Q}$. The learning constants are defined as $\beta_{\mathrm{R}} = 0.1$ and $\alpha = 0.01$. The first diagonal element of the $\mathbf{Q}$ penalizes the angular position and the second one, the velocity. The PWM of the brushless DC motors was working with 66% of the power. A number of 250 sampling times were considered for the implementation of each trial and 10 learning loops. The last learning loop took 6.53(s) to finish the operation, see Fig. 9. It can be seen that Fig. 9 shows the time-varying sampling time during the control loop, in different trials. We are not supposed to see any enhancement (concerning the sampling time) by the iterations, the average must be in the range, around 0.03(s).

The error of the system decreased by learning iterations, reduced to zero with 10 loops, Fig. 10. The error was measured directly by the optical encoders, installed on the system for providing the feedback. The amplitude of the input torque increases with each iteration (this will cause an increase at the end of servo signals), and the reason is the usual small raise of the SDDRE caused by finite-time control. The inputs to the servos are also presented in Fig. 11. One of the challenges in VP control is the asymmetric motion of the blades in the opposite direction which caused the deviation zero blade angle in Fig. 11. As a summary, the proposed ILC control with nonlinear optimal control structure of the SDDRE in finite-time was implemented on the VP experimental benchmark and without simplification. The programming was done in MATLAB script thanks to the Arduino package.

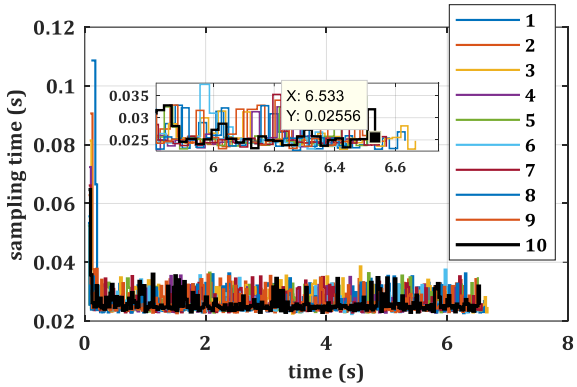

Fig. 9. Time-varying sampling time of the experiment.

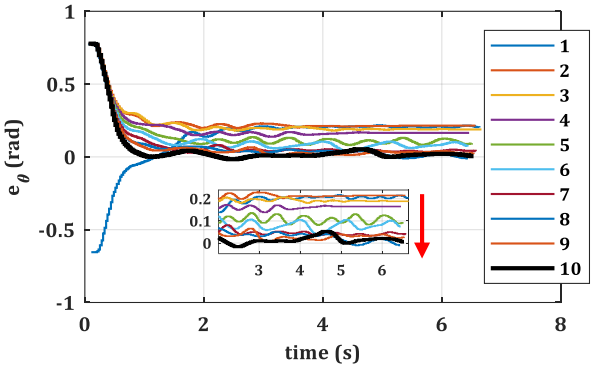

Fig. 10. The error of the system, learning in 10 loops.

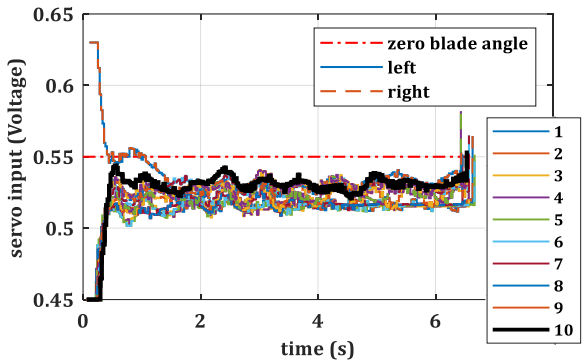

Fig. 11. Servo input of the blades, 10 iterations of learning.

VII. Conclusions

This research proposed a combination of iterative control with the SDRE/SDDRE controller in a symmetric gain structure. The new shape provides the stability proof of the controller for a special case of mechanical systems analytically. The SDRE controller is a suboptimal controller that shares robustness, optimality, systematic structure, and design flexibility in the combination with the ILC. The ILC, on the other hand, reduces the error in each loop trained by the gradient descent method. The advantage of this approach is that the learning system is based on a stable optimal controller. The first loop of iteration does not endanger the control operation. Each consecutive loop enhances performance by reducing the error. The proposed controller has been tested on a real prototype. The HYFLIERS project requires inspection and maintenance in refineries [50]. The safety measurements in refineries are high and not all the learning controllers could be tested in the sites. So, this type of proposed controller with a stable input law, even from the first learning loop, is preferable for such activities. The benchmark was subjected to the learning loops and showed error reduction caused by the ILC and without simplification. This pilot setup prepares the condition for implementation of the method on a VP drone for rotation around a pipe.

References

[1] T. Cimen, "Survey of state-dependent Riccati equation in nonlinear optimal feedback control synthesis," *Journal of Guidance, Control, and Dynamics*, vol. 35, pp. 1025-1047, 2012.
[2] T. Cimen and S. P. Banks, "Global optimal feedback control for general nonlinear systems with nonquadratic performance criteria," *Systems & Control Letters*, vol. 53, pp. 327-346, 2004.
[3] L. Felicetti and G. B. Palmerini, "A comparison among classical and SDRE techniques in formation flying orbital control," in *IEEE Aerospace Conference*, Big Sky, Montana, 2013, pp. 1-12.
[4] A. Heydari and S. N. Balakrishnan, "Fixed-final-time optimal tracking control of input-affine nonlinear systems," *Neurocomputing*, vol. 129, pp. 528-539, 2014.
[5] A. Hamdache, S. Saadi, and I. Elmouki, "Nominal and neighboring-optimal control approaches to the adoptive immunotherapy for cancer," *International Journal of Dynamics and Control*, vol. 4, pp. 346-361, 2016.
[6] S.-R. Oh, Z. Bien, and I. H. Suh, "An iterative learning control method with application to robot manipulators," *IEEE Journal on Robotics and Automation*, vol. 4, pp. 508-514, 1988.
[7] H.-S. Ahn, Y.-Q. Chen, and K. L. Moore, "Iterative learning control: Brief survey and categorization," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 37, pp. 1099-1121, 2007.
[8] D. A. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control," *IEEE control systems magazine*, vol. 26, pp. 96-114, 2006.
[9] Q. Zhu, F. Song, J.-X. Xu, and Y. Liu, "An internal model based iterative learning control for wafer scanner systems," *IEEE/ASME Transactions on Mechatronics*, vol. 24, pp. 2073-2084, 2019.
[10] D. Shen, "Iterative learning control with incomplete information: A survey," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, pp. 885-901, 2018.
[11] J. Zhang, B. Cui, X. Dai, and Z. Jiang, "Iterative learning control for distributed parameter systems based on non-collocated sensors and actuators," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, pp. 865-871, 2019.
[12] J. Wei, Y. Zhang, and H. Bao, "An exploration on adaptive iterative learning control for a class of commensurate high-order uncertain nonlinear fractional order systems," *IEEE/CAA Journal of Automatica Sinica*, vol. 5, pp. 618-627, 2017.
[13] W. He, T. Meng, X. He, and C. Sun, "Iterative learning control for a flapping wing micro aerial vehicle under distributed disturbances," *IEEE transactions on cybernetics*, vol. 49, pp. 1524-1535, 2018.
[14] Y. Jian, D. Huang, J. Liu, and D. Min, "High-precision tracking of piezoelectric actuator using iterative learning control and direct inverse compensation of hysteresis," *IEEE Transactions on Industrial Electronics*, vol. 66, pp. 368-377, 2018.
[15] D. Shen and J.-X. Xu, "Adaptive learning control for nonlinear systems with randomly varying iteration lengths," *IEEE transactions on neural networks and learning systems*, vol. 30, pp. 1119-1132, 2018.
[16] L. Roveda, G. Pallucca, N. Pedrocchi, F. Braghin, and L. M. Tosatti, "Iterative learning procedure with reinforcement for high-accuracy force tracking in robotized tasks," *IEEE Transactions on Industrial Informatics*, vol. 14, pp. 1753-1763, 2018.
[17] S.-K. Oh, B. J. Park, and J. M. Lee, "Point-to-point iterative learning model predictive control," *Automatica*, vol. 89, pp. 135-143, 2018.
[18] Y. Pan and H. Yu, "Composite learning robot control with guaranteed parameter convergence," *Automatica*, vol. 89, pp. 398-406, 2018.
[19] A. Schöllig and R. D'Andrea, "Optimization-based iterative learning control for trajectory tracking," in *European Control Conference*, Budapest, Hungary, 2009, pp. 1505-1510.

[20] M. Ohnishi, L. Wang, G. Notomista, and M. Egerstedt, "Barrier-certified adaptive reinforcement learning with applications to brushbot navigation," *IEEE Transactions on Robotics*, vol. 35, pp. 1186-1205, 2019.
[21] A. Loquercio, E. Kaufmann, R. Ranftl, A. Dosovitskiy, V. Koltun, and D. Scaramuzza, "Deep drone racing: From simulation to reality with domain randomization," *IEEE Transactions on Robotics*, vol. 36, pp. 1-14, 2020.
[22] H. Pan, X. Niu, R.-C. Li, Y. Dou, and H. Jiang, "Annealed gradient descent for deep learning," *Neurocomputing*, vol. 380, pp. 201-211, 2020.
[23] T.-Y. Kuc, K. Nam, and J. S. Lee, "An iterative learning control of robot manipulators," *IEEE Transactions on Robotics and Automation*, vol. 7, pp. 835-842, 1991.
[24] M. Sun and D. Wang, "Closed-loop iterative learning control for non-linear systems with initial shifts," *International Journal of Adaptive Control and Signal Processing*, vol. 16, pp. 515-538, 2002.
[25] C.-W. Chen, S. Rai, and T.-C. Tsao, "Iterative learning of dynamic inverse filters for feedforward tracking control," *IEEE/ASME Transactions on Mechatronics*, vol. 25, pp. 349-359, 2019.
[26] A. P. Schoellig, F. L. Mueller, and R. D’Andrea, "Optimization-based iterative learning for precise quadrocopter trajectory tracking," *Autonomous Robots*, vol. 33, pp. 103-127, 2012.
[27] D. Meng and J. Zhang, "Robust optimization-based iterative learning control for nonlinear systems with nonrepetitive uncertainties," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, pp. 1001-1014, 2021.
[28] D. Meng and J. Zhang, "Design and analysis of data-driven learning control: An optimization-based approach," *IEEE Transactions on Neural Networks and Learning Systems*, 2021.
[29] D. Meng and J. Zhang, "Robust tracking of nonrepetitive learning control systems with iteration-dependent references," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018.
[30] R. Kelly and R. Carelli, "A class of nonlinear PD-type controllers for robot manipulators," *Journal of Robotic Systems*, vol. 13, pp. 793-802, 1996.
[31] J. Alvarez-Ramirez, R. Kelly, and I. Cervantes, "Semiglobal stability of saturated linear PID control for robot manipulators," *Automatica*, vol. 39, pp. 989-995, 2003.
[32] S. R. Nekoo, J. Á. Acosta, G. Heredia, and A. Ollero, "A benchmark mechatronics platform to assess the inspection around pipes with variable pitch quadrotor for industrial sites," *Mechatronics*, vol. 79, p. 102641, 2021.
[33] M. H. Korayem and S. R. Nekoo, "Finite-time state-dependent Riccati equation for time-varying nonaffine systems: Rigid and flexible joint manipulator control," *ISA Transactions*, vol. 54, pp. 125-144, 2015.
[34] T. Cimen, "State-dependent Riccati equation (SDRE) control: A survey," *IFAC Proceedings Volumes*, vol. 41, pp. 3761-3775, Jul. 2008.
[35] S. R. Nekoo, "Tutorial and review on the state-dependent Riccati equation," *Journal of Applied Nonlinear Dynamics*, vol. 8, pp. 109-166, 2019.
[36] Y. Batmani, M. Davoodi, and N. Meskin, "On design of suboptimal tracking controller for a class of nonlinear systems," in *American Control Conference*, Boston, MA, USA, 2016, pp. 1094-1098.
[37] A. Ghaffari, M. Nazari, and F. Arab, "Suboptimal mixed vaccine and chemotherapy in finite duration cancer treatment: state-dependent Riccati equation control," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 37, pp. 45-56, 2015.
[38] A. Wernli and G. Cook, "Suboptimal control for the nonlinear quadratic regulator problem," *Automatica*, vol. 11, pp. 75-84, 1975.
[39] S. R. Nekoo, J. Á. Acosta, and A. Ollero, "Gravity compensation and optimal control of actuated multibody system dynamics," *IET Control Theory & Applications*, vol. 16, pp. 79-93, 2021.
[40] M. H. Korayem, A. Nikoobin, and V. Azimirad, "Maximum load carrying capacity of mobile manipulators: optimal control approach," *Robotica*, vol. 27, pp. 147-159, 2009.
[41] D. E. Kirk, *Optimal control theory: An introduction*: Courier Corporation, 2012.
[42] T. Çimen and S. P. Banks, "Nonlinear optimal tracking control with application to super-tankers for autopilot design," *Automatica*, vol. 40, pp. 1845-1863, 2004.
[43] A. Prach, O. Tekinalp, and D. Bernstein, "Nonlinear aircraft flight control using the forward propagating Riccati equation," in *AIAA Guidance, Navigation, and Control Conference*, San Diego, California, USA, 2016, pp. 1383-1396.
[44] A. Khamis, H. M. Nguyen, and D. S. Naidu, "Nonlinear, optimal control of wind energy conversion systems using differential SDRE," in *Resilience Week*, Philadelphia, PA, 2015, pp. 1-6.
[45] B. Geranmehr, E. Khanmirza, and S. Kazemi, "Trajectory control of aggressive maneuver by agile autonomous helicopter," *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, p. 0954410018755807, 2018.
[46] S. R. Nekoo and B. Geranmehr, "Nonlinear observer-based optimal control using the state-dependent Riccati equation for a class of non-affine control systems," *Journal of Control Engineering and Applied Informatics*, vol. 16, pp. 5-13, 2014.
[47] S. R. Nekoo, "Nonlinear closed loop optimal control: A modified state-dependent Riccati equation," *ISA Transactions*, vol. 52, pp. 285–290, 2013.
[48] S. R. Nekoo, "Digital implementation of a continuous-time nonlinear optimal controller: An experimental study with real-time computations," *ISA transactions*, vol. 101, pp. 346-357, 2020.
[49] M. Xin, S. N. Balakrishnan, and Z. Huang, "Robust state dependent Riccati equation based robot manipulator control," in *Proceedings of the IEEE International Conference on Control Applications*, Mexico City, Mexico, 2001, pp. 369-374.
[50] https://www.oulu.fi/hyfliers/.

Saeed Rafee Nekoo is a senior postdoc researcher at (AICIA) Andalusian Association for Research and Industrial Cooperation, and GRVC research group, University of Seville, Spain. He currently serves as a researcher in aerial robotics and control engineering and is engaged with HYFLIERS (HYbrid FLying-rollIng with-snakE-aRm robot for contact inSpection) H2020 European Research Council project, focusing on design and development of variable-pitch-rotor drones for inspection: design, prototyping, control implementation, and experimentation. Saeed's research interest includes: robotics, nonlinear and optimal control, control engineering, manufacturing, cooperative robotics, flexible joint manipulators, observer, and estimator design, path planning, wheeled mobile robots, control of autonomous underwater vehicles, free-floating space manipulator design, and control, digital implementation of continuous-time nonlinear controllers, and design, manufacturing and control of mechatronics systems, aerial robotics, multirotor UAV and variable-pitch-rotor quadcopter control.

José Ángel Acosta was born in Huelva, Spain. He obtained both the Servo-Electrical and Mechanical Engineering degrees at the University of Huelva, Spain, and the Electrical Engineering degree at the University of Seville, Spain, respectively. He was Marie Curie Control Training Site Fellow as Member of the Centre National de la Recherche Scientifique (CNRS, France) in the Laboratoire des Signaux et Systèmes, Supélec, France, in 2003 and 2005. He obtained the Ph.D. degree in 2004 at the Automatic Control and Systems Engineering Department at the University of Seville and the Ph.D. European Award in 2005. He was nominated for the George S. Axelby Outstanding Paper Award in the IEEE Transactions on Automatic Control journal in 2006. In 1999 he joined that department as Research Assistant, where he is currently Professor and Research

Member of the Automatic Control and Robotics Institute. He has also been a Visitor at the Laboratoire des Signaux & Systèmes (CNRS, France) repeatedly since 2005 and Academic Visitor researching in the Electrical & Electronic Engineering Department as a Member of the Control & Power Group at Imperial College London, UK, in 2008-09-10 and 2011. His research interests are in the field of nonlinear control of dynamical systems with emphasis on electromechanical and robotic systems.

Guillermo Heredia is Full Professor at University of Seville (Spain). He was visiting researcher at the Field Robotics Centre, Carnegie Mellon University (Pittsburgh, USA), and worked for an international automobile manu-facturer (General Motors). He participated as senior researcher in 65 R&D projects (EU, NASA and national projects), leading or co-leading 12 of them, including the FP7 ECSAFEMOBIL and the H2020 AEROBI and HYFLIERS projects. He is author or co-author of more than 100 publications on aerial robotics, aerial manipulation, autonomous vehicles and fault detection and reconfiguration.

Anibal Ollero is full Professor and Head of GRVC at University Seville, and Scientific Advisor of the Center for Aerospace Technologies (CATEC) also in Seville. He has been full professor at the Universities of Santiago and Malaga (Spain) and researcher at the Robotics Institute of Carnegie Mellon University (Pittsburgh, USA) and LAAS-CNRS (Toulouse, France). He authored more than 750 publications, including 9 books and 200 papers in journals and has been editor of 15 books. He has delivered plenaries and keynotes in more than 100 events including IEEE ICRA 2016 and IEEE IROS 2018. He has been supervisor or co-supervisor of 45 PhD Thesis that have received many awards. He led more than 160 research projects, participating in more than 25 projects of the European Research Programmes being coordinator of 7 and associated or deputy coordinator of 3, all of them dealing with unmanned aerial systems and aerial robots. From November 2018 he is running the GRIFFIN ERC-Advanced Grant with the objective of developing a new generation of aerial robots that will be able to glide, flapping the wings, perch and manipulate by maintaining the equilibrium, and from December 2019 he is the coordinator of the H2020-AERiAL-CORE project with the participation of 15 universities, research centers and companies dealing with aerial robotic manipulators and applications to inspection and maintenance He has transferred technologies to more than 20 companies and has been awarded with 25 international research and innovation awards, including the recent Rei Jaume I in New Technologies (Spain), the Challenge 3 of the MBZIRC 2020 International Robotics Competition, the Overall Information and Communication Technologies Innovation Radar Prize 2017 of the European Commission, and has been also elected between the three European innovators of the year being candidate to the European personalities of the year 2017. He is IEEE Fellow "for contributions to the development and deployment of aerial robots". Currently co-chair of the "IEEE Technical Committee on Aerial Robotics and Unmanned Aerial Vehicles", coordinator of the "Aerial Robotics Topic Group" of euRobotics and has been member of the "Board of Directors" of euRobotics until March 2019. He has been also founder and president of the Spanish Society for the Research and Development in Robotics (SEIDROB) until November 2017.