Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-18 (2016) 962–967

An extremum seeking approach to sampled-data iterative learning control of continuous-time nonlinear systems ⋆

Sei Zhen Khong ∗ **Dragan Nešić** ∗∗ **Miroslav Krstić** ∗∗∗

∗ *Institute for Mathematics and its Applications, The University of Minnesota, Minneapolis, MN 55455, USA. (e-mail: `szkhong@umn.edu`)*
∗∗ *Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, VIC 3010, Australia (e-mail: `dnesic@unimelb.edu.au`)*
∗∗∗ *Department of Mechanical and Aerospace Engineering, University of California, San Diego, CA 92093-0411, USA. (e-mail: `krstic@ucsd.edu`)*

Abstract: Iterative learning control (ILC) of continuous-time nonlinear plants with periodic sampled-data inputs is considered via an extremum seeking approach. ILC is performed without exploiting knowledge about any plant model, whereby the input signal is constructed recursively so that the corresponding plant output tracks a prescribed reference trajectory as closely as possible on a finite horizon. The ILC is formulated in terms of a non-model-based extremum seeking control problem, to which local optimisation methods such as gradient descent and Newton are applicable. Sufficient conditions on convergence to a neighbourhood of the reference trajectory are given.

Keywords: Iterative learning control, extremum seeking, gradient optimisation methods, nonlinear time-varying systems

1. INTRODUCTION

Iterative learning control (ILC) is a learning based method for achieving the objective of tracking a prescribed trajectory. It performs the same task multiple times in succession with respect to iteratively updated control inputs while improving the tracking performance by learning from previous executions (Moore, 1993; Moore et al., 1992; Xu and Tan, 2003). ILC is known to achieve good performance in the presence of repeating disturbances and certain model uncertainty due to its iteratively learning feature. For instance, ILC has been applied to motion planning in gait modelling (Srinivasan and Ruina, 2006) and stroke rehabilitation (Freeman et al., 2012, 2015), where the repeating nature of the patients' tasks is exploited to improve performance and respond to physiological changes by learning from past clinical trials.

The vast majority of optimisation-based ILC methods in the literature rely on prior knowledge about the models. For example, the updating control laws as well as convergence of the ILC methods in Gunnarsson and Norrlöf (2001); Owens and Hätönen (2005) depend on the precise knowledge of the nominal model. Owens. et al. (2009) assumes the modelling uncertainty is multiplicative and bounded and proposes a robust monotone gradient-based scheme for ILC of linear time-invariant (LTI) systems. The case where an LTI model is subject to noisy disturbances is investigated in Schoellig et al. (2012) with a consolidated model-based Kalman filter and convex optimisation approach to ILC. A primal barrier method to ILC of LTI systems is proposed in Mishra et al. (2011). It relies on the availability of knowledge about the gradient and Hessian of the quadratic cost function, which in turn is dependent on the models.

This paper proposes an *extremum-seeking* based framework in which periodic sampled-data iterative learning control of continuous-time time-varying nonlinear systems on finite horizon is performed without relying on knowledge of the models. The control laws are restricted to being periodic sampled data signals (piecewise constant in time), as is implementable in digital control with a zero-order hold device. Extremum seeking is used to locate an optimum without exploiting knowledge about the underlying mathematical models describing the dynamics of the systems (Ariyur and Krstić, 2003; Zhang and Ordóñez, 2011). Difficulty in modelling complex nonlinear systems may give rise to the unavailability of such knowledge. Several approaches to extremum seeking control can be found in the literature. They include the adaptive control methods (Krstić and Wang, 2000; Tan et al., 2006; Nešić et al., 2010; Ghaffari et al., 2012; Guay and Dochain, 2015), nonlinear programming methods (Teel and Popović, 2001), and global optimisation methods (Khong et al., 2013a,b). Stochastic approximation based methods are considered in Khong et al. (2015) in the presence of output measurement noise.

⋆ This research was supported in part by the Institute for Mathematics and its Applications with funds provided by the National Science Foundation and the Australian Research Council.

10.1016/j.ifacol.2016.10.292

Optimisation-based extremum seeking algorithms are applied to ILC in this paper. It is shown that the proposed ILC framework is amenable to a broad range of local optimisation methods. This allows the complexity of implementation and convergence speed of the algorithms to be taken into consideration in the control synthesis phase. For instance, Newton-based methods can be employed if a quadratic convergence rate is sought after within a neighbourhood of the minimiser at the expense of heavier computational load than first-order methods. Furthermore, the convergence properties of ILC can be analysed using standard tools from the field of optimisation. Sufficient conditions for ultimately bounded asymptotic stability of local minima is demonstrated, where the cost function is defined as the distance between the system output and the reference trajectory.

The work presented in this paper has strong relevance to the theory of optimal control, which is an important area of systems theory in which the problem of finding a control law for a given system to achieve a certain optimality criterion is tackled. When the dynamics are linear and the cost is quadratic in the state and input, the problem reduces to the so-called the linear quadratic regulation, which can be constructed by solving a Ricatti equation (Kalman, 1960). In other cases, the optimal control law can be derived using the Pontryagin's minimum principle or by solving the Hamilton-Jacobi-Bellman equation, the former of which is a necessary condition for optimality and the latter also sufficient (Liberzon, 2012). In all of the cases above, a model of the system is required. A first attempt at employing extremum seeking for obtaining an optimal control law in the absence of a system model can be found in Frihauf et al. (2013), where linear discrete-time systems with quadratic costs are considered. Other works on optimal control of unknown discrete-time systems include the neural networks based (Dierks et al., 2009) and the reinforced learning based (Yang and Jagannathan, 2012). In Khong (2014), the problem of computing periodic sampled-data control laws satisfying saturation constraints for nonlinear continuous-time systems and generic cost functions is considered via an extremum seeking approach.

The paper is structured as follows. The problem of sampled-data ILC of nonlinear systems is formulated in the next section. An extremum seeking approach to ILC based on local optimisation methods is introduced in Section 3. Gradient based methods which are compatible with the extremum seeking framework are described in Section 4. Concluding remarks are provided in Section 5.

Some basic notation is introduced here. A function $\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class-$\mathcal{K}$ (denoted $\gamma \in \mathcal{K}$) if it is continuous, strictly increasing, and $\gamma(0) = 0$. If γ is also unbounded, then $\gamma \in \mathcal{K}_\infty$. A continuous function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is of class-$\mathcal{KL}$ if for each fixed t, $\beta(\cdot, t) \in \mathcal{K}$ and for each fixed s, $\beta(s, \cdot)$ is decreasing to zero (Khalil, 2002). Given any subset $\mathcal{X}$ of $\mathbb{R}^m$ and a point $x \in \mathbb{R}^m$, define the distance of x from $\mathcal{X}$ as $\|x\|_{\mathcal{X}} := \inf_{a \in \mathcal{X}} \|x - a\|_2$.

2. ITERATIVE LEARNING CONTROL

2.1 Reference tracking

The class of nonlinear, possibly infinite-dimensional, systems that we consider in this paper is introduced in this section. We begin with the following notational definitions. The natural and real numbers are denoted by $\mathbb{N}$ and $\mathbb{R}$, respectively. The Euclidean norm is denoted $|\cdot|$. Let $\mathcal{X}$ be a Banach space whose norm is denoted $\|\cdot\|_{\mathcal{X}}$. Let $T > 0$ denote the finite length of the time horizon of interest. Define

$$\mathscr{X} := \{x : [0, T] \to \mathcal{X} \mid x \text{ is measurable}\}$$

and

$$\mathscr{U} := \{u : [0, T] \to [a, b] \mid u \text{ is measurable}\}.$$

where the compact interval $[a, b]$ with $a, b \in \mathbb{R}$ denotes the input space of interest. This is motivated by the ubiquity of control input saturation constraints in physical systems (Khalil, 2002).

Definition 1. Let the state of a dynamical system be represented by $x \in \mathscr{X}$. The input to the system is denoted by $u \in \mathscr{U}$. Let $x(\cdot, x_0, u)$ denote the state of the dynamical system starting at $x(0) = x_0 \in \mathcal{X}$ with input $u \in \mathscr{U}$ that is Lipschitz continuous in x_0 and u. The output of the system is given by

$$y(t) = h(x(t, x_0, u)) \quad \forall t \in [0, T],$$

where $h : \mathcal{X} \to \mathbb{R}$ is a locally Lipschitz function.

An example satisfying the definition above is given by the following nonlinear time-varying finite-dimensional state-space system

$$\begin{aligned} \dot{x} &= f(t, x, u) \quad x(0) = x_0 \\ y &= h(x), \end{aligned} \tag{1}$$

where $f : \mathbb{R} \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ and $h : \mathbb{R}^n \to \mathbb{R}$ are locally Lipschitz functions in each argument.

Let $r : [0, T] \to \mathbb{R}$ denote the reference trajectory.

Definition 2. Given a system described in Definition 1 with initial condition $x_0 \in \mathcal{X}$, define the cost function $J_r : \mathscr{U} \to \mathbb{R}_{\geq 0}$ to be

$$J_r(u) = \|r - y\|,$$

where $\|\cdot\|$ denotes the $\mathscr{L}_2$ norm

$$\|z\| := \left(\int_{t=0}^{T} \|z(t)\|_2^2 \, dt \right)^{\frac{1}{2}}$$

and $y : [0, T] \to \mathbb{R}$ is the output of the system with respect to the initial condition x_0 and input $u \in \mathscr{U}$.

Note that by Definition 1, J_r is Lipschitz continuous in the sense that there exists an $L > 0$ such that

$$\begin{aligned} |J_r(u) - J_r(u')| &\leq L\|u - u'\|_\infty \\ &:= L \operatorname*{ess\,sup}_{t \in [0,T]} |u(t) - u'(t)| \end{aligned}$$

for all $u, u' \in \mathscr{U}$. Denote by $h > 0$ a real number by which T is divisible, i.e. there exists an $K \in \mathbb{N}$ such that $Kh = T$. An h-periodic sampled-data control law $u : [0, T) \to [a, b]$ is given by

$$u(t) := u_i, \quad t \in [(i-1)h, ih) \tag{2}$$

for $i = 1, 2, \ldots, K$ and some $u_1, \ldots, u_K \in [a, b]$. Denote by $\mathscr{U}_h$ the set of all such control laws. Since J_T is continuous and $\mathscr{U}_h$ is a compact subset of $\{u : \mathbb{R}_{\geq 0} \to \mathbb{R} \mid u \text{ is measurable}\}$ with respect to the metric induced by $\|\cdot\|_\infty$, it follows that J_T admits a global minimum over $\mathscr{U}_h$ by the extreme value theorem (Rudin, 1976, Thm. 4.16). The *finite-horizon sampled-data optimal reference tracking* problem can thus be stated as

$$u^*(h) := \arg\min_{u \in \mathscr{U}_h} J_r(u). \tag{3}$$

Note that the number of decision variables in the optimisation problem above is K. In the case where $\mathcal{X}$ is a finite-dimensional space, say $\mathbb{R}^n$ as in (1), the initial condition x_0 can also be incorporated in the above optimisation formulation if n is known a priori. That is, both the input and the initial condition are decision variables to be chosen to minimise the difference between the system output and the reference trajectory. The extremum-seeking approach introduced in the following section can accommodate such a reformulation straightforwardly, at the expense of increasing the problem complexity by n variables.

An alternative formulation, in contrast to the sampled-data control law introduced in (2), is to parameterise the control input with a finite number of basis functions and select the minimising control signal over the set of parameters. Commonly used basis functions include Bezier polynomials of finite order, finite sums of sinusoids with variable frequencies and magnitudes, and piecewise linear functions with variable slopes. These formulations all boil down to choosing a finite number of variables for optimal reference tracking and can also be approached using the extremum-seeking framework introduced next in the paper. Mathematically, let $\psi_i : [0, T] \to \mathbb{R}$, $i = 1, \ldots, K$ be measurable functions and

$$\mathscr{U}_\psi := \left\{ u : [0, T] \to \mathbb{R} \;\middle|\; u = \sum_{i=1}^{K} \theta_i \psi_i,\ \theta_i \in [a, b] \right\}.$$

The finite-horizon reference tracking problem can be posed as

$$u := \arg\min_{u \in \mathscr{U}_\psi} J_r(u) = \arg\min_{\theta_i \in [a,b]} J_r(u).$$

2.2 Iterative learning control for reference tracking

Iterative learning control (ILC) is a recursive learning based algorithm for solving the reference tracking problem (Moore, 1993; Moore et al., 1992; Xu and Tan, 2003). The basic idea of ILC is to use previous iteration information to update the control signal for the next iteration so that the optimal control is found to within some tolerance after a sufficiently large number of ILC trials. Within the ILC formulation, it is possible to use temporal information that would be non-causal in standard control provided it is generated on a previous iteration. Indeed, if such information is not present the control law is equivalent to feedback action. A common ILC algorithm takes the following form:

$$u_{j+1}(t) = u_j(t) + \Upsilon(r, y_j)(t) \quad t \in [0, T],$$

where $j = 0, 1, \ldots$ denotes the iteration number and Υ is a possibly non-causal operator. By running this algorithm for a system satisfying Definition 1, the limit of u_j is expected to converge to a neighbourhood of u^* in (3).

Different ways of exploiting information from past iterations in the update via Υ results in different ILC algorithms. A more general form of ILC is to include information from a finite number s of previous iterations in the computation of the current iteration control, i.e. $u_j, u_{j-1}, \ldots, u_{j-s+1}$ and $y_j, y_{j-1}, \ldots, y_{j-s+1}$. This is well known — termed higher-order ILC — and by having $s > 1$ the control law in this paper has this structure. Higher-order ILC has long history and in general it is not known how to decide when this gives advantages over $s = 1$ or what value of s should be used. Besides being applicable to infinite-dimensional nonlinear time-varying systems (cf. Definition 1), the extremum-seeking based ILC proposed in the next section employs optimisation algorithms as part of its operation, and is hence naturally a higher-order ILC scheme. By the same token, its convergence properties can be analysed using well-known methods from optimisation.

3. AN EXTREMUM SEEKING FRAMEWORK

The majority of the ILC literature is focused on discrete-time finite-dimensional linear systems (Moore, 1993; Bristow et al., 2006; Ahn et al., 2007). Such systems are often analysed using the lifted-system approach and super vector notation, i.e. by considering the time series as a vector. The discrete-time controllers and analysis methods do not readily have natural counterparts in the continuous-time setting. On the contrary, the proposed extremum-seeking-control framework in this section is applicable to the nonlinear time-varying ILC problem introduced in the previous section and is built upon the well-studied area of operations research.

It is first demonstrated that the sampled-data optimal reference tracking problem in (3) can be transformed into a problem of static optimisation. To this end, the following operations are useful. Given a vector $v \in \mathbb{R}^K$, where K satisfies $Kh = T$, define the demultiplexer $D : \mathbb{R}^K \to \mathcal{U}_h$ by

$$\begin{aligned} D(v) &= u \\ u(t) &= v_i \quad t \in [(i-1)h, ih). \end{aligned} \tag{4}$$

Similarly, given a $z \in \mathcal{U}_h$, define the multiplexer $M : \mathcal{U}_h \to \mathbb{R}^K$ by

$$\begin{aligned} v &= M(u) \\ v_i &= u((i-1)h) \quad i = 1, \ldots, K. \end{aligned} \tag{5}$$

The demultiplexer and multiplexer are useful for analytically relating the behaviour of the plant with the optimisation method.

Given a system Σ satisfying Definition 1 with initial condition $x_0 \in \mathcal{X}$, reference trajectory $r : [0, T] \to \mathbb{R}$, and $\theta_j \in \mathbb{R}^K$, let $Q : \mathbb{R}^K \to \mathbb{R}$ be defined as

$$\begin{aligned} Q(\theta_j) := J_r(u_j) &= \|r - y_j\| \\ &= \|r - h(x(\cdot, x_0, u_j))\|, \end{aligned} \tag{6}$$

where $u_j := D(\theta_j) \in \mathcal{U}_h$, for $j = 0, 1, \ldots$. Note that Q is locally Lipschitz continuous. The process above transforms the problem of ILC to one of static optimisation, to which a broad array of local and global algorithms is applicable.

In particular, θ_{j+1} can be designed iteratively based on some update relationship

$$\theta_{j+1} = \Gamma(J_r(u_j), \dots, J_r(u_{j-s}), \theta_j, \dots, \theta_{j-s})$$

and applied to Σ via the demultiplexer D. Figure 1 illustrates the extremum seeking based approach to ILC. Notice that the demultiplexer accepts as its input a vector of real numbers from the optimisation method and outputs a corresponding control input signal to the plant.

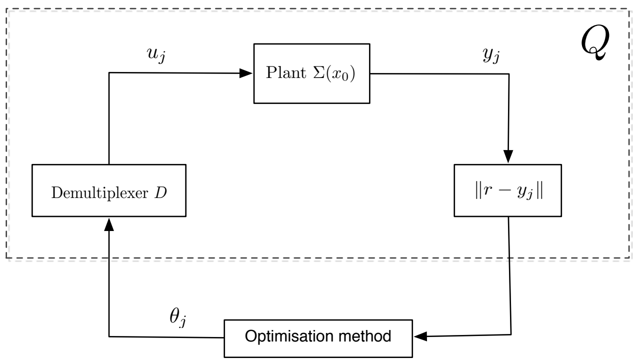

Fig. 1. An extremum seeking based framework for iterative learning control.

Consider the following static function optimisation problem:

$$z^* := \min_{\theta \in \Omega} Q(\theta), \tag{7}$$

where $Q : \mathbb{R}^K \to \mathbb{R}$ is continuously differentiable and

$$\Omega := \left\{ u \in \mathbb{R}^K \mid u_i \in [a_i, b_i] \subset \mathbb{R}, i = 1, 2, \dots, K \right\}$$

is compact. Assume that the Jacobian $\nabla Q = 0$ in a nonempty, compact set $\mathcal{C} \subset \mathbb{R}^K$, i.e. Q achieves its minimum on $\mathcal{C}$. A local optimisation method Γ may generate its update according to

$$\theta_{j+1} = \Gamma(\theta_j, \dots, \theta_{j-1}, z_j, \dots, z_{j-s}),$$

where $z_i = Q(\theta_i)$. A broad class of gradient-based optimisation algorithms can be shown to satisfy the following assumption provided that they are properly initialised. Some of these algorithms are described in Section 4.

Assumption 3. The optimisation method Γ, when applied to (7), possesses the following properties:

(i) Denote by $\{\theta_j\}_{j=0}^{\infty} \subset \Omega \subset \mathbb{R}^K$ the output sequence Γ generates based on inputs to Γ, $\{z_j\}_{j=1}^{\infty}$, where $z_j := Q(\theta_j)$. Γ is causal in the sense that the output θ_N at any time $N \in \mathbb{N}$ is determined based only on θ_j and z_j for $j = 0, 1, \dots, N-1$, i.e. the past probe values to Q and the corresponding measurements.

(ii) There exists a class-$\mathcal{KL}$ function β such that for any initial point $\theta_0 \in \Omega$ and some $\delta \geq 0$,

$$\|\theta_j(\theta_0)\|_{\mathcal{C}} \leq \beta(\|\theta_0\|_{\mathcal{C}}, j) + \delta \quad \forall j \geq 0. \tag{8}$$

Note that Assumption 3(ii) states that the sequence θ_j converges asymptotically to a δ-neighbourhood of $\mathcal{C}$. The convergence result of the proposed ILC in Figure 1 is in order. In the following, recall the demultiplexer D and multiplexer M defined in (4) and (5) respectively.

Theorem 4. Given a nonlinear plant Σ with initial condition $x_0 \in \mathcal{X}$ described in (1) and a reference trajectory $r : [0, T] \to \mathbb{R}$, the feedback interconnection shown in Figure 1 with the optimisation method Γ satisfying Assumption 3 has the following convergence property: there exists a class-$\mathcal{KL}$ function β such that for any initial point $\theta_0 \in \Omega$ and for some $\delta \geq 0$,

$$\|M(u_j)\|_{\mathcal{C}} \leq \beta(\|M(u_0)\|_{\mathcal{C}}, j) + \delta \quad \forall j \geq 0, \tag{9}$$

where $Q : \Omega \subset \mathbb{R}^K \to \mathbb{R}$ is as defined in (6), whose Jacobian is zero on $\mathcal{C} \subset \Omega$.

Proof. Note that by the setup of the feedback interconnection in Figure 1, $u_j = D(\theta_j)$ for all $j = 0, 1, \dots$. By applying Γ to Q, it follows from Assumption 3 that there exists a class-$\mathcal{KL}$ function β such that for any initial point $\theta_0 \in \Omega$ and some $\delta \geq 0$,

$$\|\theta_j\|_{\mathcal{C}} \leq \beta(\|\theta_0\|_{\mathcal{C}}, j) + \delta \quad \forall j \geq 0.$$

This is equivalent to (9) via the relationship $M(u_j) = \theta_j$, as claimed.

Theorem 4 gives sufficient conditions under which the iterations on u_j are convergent to a δ-neighbourhood of local minima of the function Q defined in (6), which measures the size of the tracking error. The next section describes certain optimisation algorithms that satisfy Assumption 3. Tradeoffs between the speed of convergence and computational complexity can be taken into account by the user when choosing an optimisation method for implementation in the proposed extremum-seeking based iterative learning control framework.

4. GRADIENT OPTIMISATION METHODS

Gradient optimisation methods generally employ estimations about the (first or higher order) derivatives of the cost function in their updates. This section describe some of the optimisation algorithms that can be used in the extremum seeking based framework introduced in the previous section. In particular, we mention two of the most well-known methods (Boyd and Vandenberghe, 2004; Polak, 1997) in the field of optimisation which satisfy Assumption 3. They are (i) the *gradient descent method*:

$$\theta_{j+1} = \theta_j - \lambda_j \nabla Q(\theta_j), \tag{10}$$

where λ_i denotes the step size which can be computed by, say, the Armijo method (Polak, 1997, Alg. 1.3.3) and (ii) the *Newton's method*:

$$\theta_{j+1} = \theta_j - \lambda \nabla^2 Q(\theta_j)^{-1} \nabla Q(\theta_j), \tag{11}$$

where $\nabla Q(\cdot)$ and $\nabla^2 Q(\cdot)$ denote, respectively, the Jacobian and Hessian of Q. It can be readily seen that the gradient and Newton methods satisfy the causality property in Assumption 3(i). The following result can be found in Polak (1997); Boyd and Vandenberghe (2004).

Proposition 4.1. Suppose $Q : \Omega \to \mathbb{R}$ is twice Lipschitz continuously differentiable and strictly convex on $\mathcal{S} \subset \Omega$, whereby there exist $\underline{M}, \bar{M} \in \mathbb{R}$ such that

$$\underline{M} I \leq \nabla^2 Q(\theta) \leq \bar{M} I \quad \text{for all } \theta \in \mathcal{S}.$$

Furthermore, suppose there exists a minimiser $\theta^* \in \mathcal{S}$ such that $\nabla Q(\theta^*) = 0$. Let $\{\theta_j\}_{j=0}^{\infty}$ be the sequence generated by the gradient or Newton method with respect to Q. Then there exists a class-$\mathcal{KL}$ function β such that for any $\theta_0 \in \mathcal{S}$,

$$\|\theta_j - \theta^*\|_2 \leq \beta(\|\theta_0 - \theta^*\|_2, j) \quad \forall j \geq 0. \tag{12}$$

Note that the rate of convergence for the gradient descent method is linear while that for Newton is quadratic when the input lies within a sufficiently small neighbourhood of the minimiser.

Proposition 4.1 states the convergence conditions for the gradient descent and Newton method when exact values of the Jacobian $\nabla Q(\theta_j)$ and Hessian $\nabla^2 Q(\theta_j)$ are known. In practice, they need to be estimated from several past measurements. This can be achieved by using the Euler methods, trapezoidal method, or the more sophisticated Runge-Kutta methods (Press et al., 2007); see Figure 2.

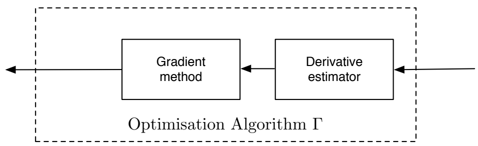

Fig. 2. An extremum seeking controller paradigm involving a derivative estimator and a gradient method.

To be more specific, let the initial output of the optimisation method be $\theta_0 := \tau_0$. As determined by the derivative estimator, the following length-p sequence of step commands $\{\theta_j\}_{j=0}^{p-1}$ can be used to probe Q along the desired directions:

$$(\tau_0 + d_1(\tau_0), \ldots, \tau_0 + d_p(\tau_0)), \tag{13}$$

where $d_i : \Omega \to \mathbb{R}^K, i = 1, \ldots, p$ denote the dither signals. The corresponding outputs of Q are then collected by the derivative estimator to numerically approximate the Jacobian $\nabla Q(\theta_0)$. Exploiting this information, the optimisation algorithm can then update its next probing point τ_1, and the series of steps described above can be repeated to generate $\{\theta_j\}_{j=p}^{2p-1}$. Note that to obtain an estimate of $\nabla Q(\cdot)$, p is required to be at least $K + 1$ in general.

Suppose the use of the derivative estimates (by contrast to their precise values) in Figure 2 introduces a bounded additive error term in the update of the gradient and Newton methods:

$$\begin{aligned} \theta_{j+1} &= \theta_j - \lambda_j \nabla Q(\theta_j) + e_1(j, \theta_j) \quad \text{and} \\ \theta_{j+1} &= \theta_j - \lambda \nabla^2 Q(\theta_j)^{-1} \nabla Q(\theta_j) + e_2(j, \theta_j), \end{aligned} \tag{14}$$

where

$$\begin{aligned} \|e_1(j, \theta_j)\|_2 &\le l_1 + q_1 \alpha(\|\theta_j\|_{\mathcal{C}}) \quad \text{and} \\ \|e_2(j, \theta_j)\|_2 &\le l_2 + q_2 \alpha(\|\theta_j\|_{\mathcal{C}}), \end{aligned} \tag{15}$$

for some $l_1, l_2, q_1, q_2 \ge 0$. This can be ensured by appropriately choosing the step sizes in the approximation methods. It follows from the non-vanishing perturbation results for discrete-time systems in Cruz-Hernández et al. (1999) that for sufficiently small l and q, the gradient/Newton-based extremum seeking controller in Figure 2 satisfies the ultimately bounded asymptotic stability Assumption 3(ii). In particular, there exist a class-$\mathcal{K}$ function α and a class-$\mathcal{KL}$ function β such that

$$\|\theta_j\|_{\mathcal{C}} \le \beta(\|\theta_0\|_{\mathcal{C}}, j) + \alpha(l) \quad \forall j = 0, 1, \ldots.$$

Furthermore, if there exists an $\alpha_d \in \mathcal{K}$ such that the dither signals in (13) satisfy for each $i = 1, \ldots, p$,

$$\|d_i(\theta)\|_2 \le \alpha_d(\|\theta\|_{\mathcal{C}}),$$

then it follows that the step size used in estimating the derivatives converges to zero as θ_j approaches the minimising set $\mathcal{C}$. This implies that the magnitudes of the error terms e_1 and e_2 in (15) tend to zero as $j \to \infty$, i.e. $l_1 = l_2 = 0$. That is to say the perturbations are vanishing and the extremum seeking controller is asymptotically stable as in Assumption 3(ii) with $\delta = 0$ (Cruz-Hernández et al., 1999). Note that an extremum-seeking based iterative learning control scheme described in Figure 1 using either the gradient descent or Newton methods together with a derivative estimator in Figure 2 naturally results in an higher-order update law of the form

$$u_{j+1}(t) = u_j(t) + \Upsilon(r, y_j, y_{j-1}, \ldots, y_{j-s})(t) \quad t \in [0, T].$$

Though not discussed here, many other optimisation algorithms such as coordinate descent and quasi-Newton methods can also be used within the extremum seeking based iterative learning control framework. Usual tradeoffs between speed of convergence, computational complexity, and required smoothness on the objective functions can be taken into consideration while selecting an optimisation algorithm for a class of problems.

5. CONCLUSIONS

This paper proposes an extremum-seeking scheme for sampled-data iterative learning control of nonlinear time-varying systems with the objective of locating an input that gives rise to the plant output that tracks a prescribed reference trajectory. The proposed scheme, as is standard in the extremum seeking literature, does not rely on the underlying model for the dynamics of the plant. Optimisation algorithms for which convergence is well understood can be used off-the-shelf within the extremum seeking framework, from which convergence of iterative learning control can be concluded. Future research directions may involve investigating optimising over an infinite-dimensional space of input signals, such as the class of smooth functions.

REFERENCES

Ahn, H.S., Chen, Y.Q., and Moore, K.L. (2007). Iterative learning control: brief survey and categorization. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 37(6), 1099–1121.

Ariyur, K.B. and Krstić, M. (2003). *Real-Time Optimization by Extremum Seeking Control.* Wiley-Interscience.

Boyd, S. and Vandenberghe, L. (2004). *Convex Optimization.* Cambridge University Press.

Bristow, D., Tharayil, M., and Alleyne, A. (2006). A survey of iterative learning control. *IEEE Control Systems Magazine*, 26(3), 96–114.

Cruz-Hernández, C., Alvarez-Gallegos, J., and Castro-Linares, R. (1999). Stability of discrete nonlinear systems under nonvanishing perturbations: application to a nonlinear model-matching problem. *IMA Journal of Mathematical Control and Information*, 16, 23–41.

Dierks, T., Thumati, B.T., and Jagannathan, S. (2009). Optimal control of unknown affine nonlinear discrete-time systems using offline-trained neural networks with proof of convergence. *Neural Networks*, 22(5-6), 851–860.

Freeman, C., Rogers, E., Burridge, J., Hughes, A.M., and Meadmore, K. (2015). *Iterative Learning Control for Electrical Stimulation and Stroke Rehabilitation.* Springer.

Freeman, C., Rogers, E., Hughes, A.M., Burridge, J., and Meadmore, K. (2012). Iterative learning control in health care: Electrical stimulation and robotic-assisted upper-limb stroke rehabilitation. *IEEE Control Systems*, 32(1), 18–43.

Frihauf, P., Krstić, M., and Başar, T. (2013). Finite-horizon LQ control for unknown discrete-time linear systems via extremum seeking. *European Journal of Control*, 19, 399–407.

Ghaffari, A., Krstić, M., and Nešić, D. (2012). Multivariable Newton-based extremum seeking. *Automatica*, 48, 1759–1767.

Guay, M. and Dochain, D. (2015). A time-varying extremum-seeking control approach. *Automatica*, 61, 356363.

Gunnarsson, S. and Norrlöf, M. (2001). On the design of ILC algorithms using optimization. *Automatica*, 37, 2011–2016.

Kalman, R. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME, Journal of Basic Engineering*, 82, 34–45.

Khalil, H.K. (2002). *Nonlinear Systems*. Prentice Hall, 3rd edition.

Khong, S.Z. (2014). Non-model based finite-horizon sampled-data optimal control via global extremum seeking. In *IEEE Multi-Conference on Systems and Control*, 1400-1405.

Khong, S.Z., Nešić, D., Manzie, C., and Tan, Y. (2013a). Multidimensional global extremum seeking via the DIRECT optimisation algorithms. *Automatica*, 49(7), 1970–1978.

Khong, S.Z., Nešić, D., Tan, Y., and Manzie, C. (2013b). Unified frameworks for sampled-data extremum seeking control: global optimisation and multi-unit systems. *Automatica*, 49(9), 2720–2733.

Khong, S.Z., Tan, Y., Manzie, C., and Nešić, D. (2015). Extremum seeking of dynamical systems via gradient descent and stochastic approximation methods. *Automatica*, 56, 44–52.

Krstić, M. and Wang, H.H. (2000). Stability of extremum seeking feedback for general nonlinear dynamic systems. *Automatica*, 36(4), 595–601.

Liberzon, D. (2012). *Calculus of variations and optimal control theory: a concise introduction*. Princeton University Press.

Mishra, S., Topcu, U., and Tomizuka, M. (2011). Optimization-based constrained iterative learning control. *IEEE Transactions on Control Systems Technology*, 19(6), 1613–1621.

Moore, K.L. (1993). *Iterative Learning Control for Deterministic Systems*. Spinger.

Moore, K.L., Dahleh, M., and Bhattacharyya, S.P. (1992). Iterative learning control: A survey and new results. *Journal of Intelligent and Robotic Systems*, 9(5), 563–594.

Nešić, D., Tan, Y., Moase, W.H., and Manzie, C. (2010). A unifying approach to extremum seeking: adaptive schemes based on estimation of derivatives. In *Proceedings of IEEE Conference on Decision and Control*, 4625–4630.

Owens, D.H. and Hätönen, J. (2005). Interative learning control — an optimization paradigm. *Annual Reviews in Control*, 29, 57–70.

Owens., D.H., Hatonen, J.J., and Daley, S. (2009). Robust monotone gradient-based discrete-time iterative learning control. *International journal of robust and nonlinear control*, 19, 634–661.

Polak, E. (1997). *Optimization — Algorithms and Consistent Approximations*. Applied Mathematical Sciences 124. Spinger.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (2007). *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, 3rd edition.

Rudin, W. (1976). *Principles of Mathematical Analysis*. McGraw-Hill, 3rd edition.

Schoellig, A.P., Mueller, F.L., and DAndrea, R. (2012). Optimization-based iterative learning for precise quadrocopter trajectory tracking. *Autonomous Robots*, 33(1-2), 103–127.

Srinivasan, M. and Ruina, A. (2006). Computer optimization of a minimal biped model discovers walking and running. *Nature*, 439, 72–75.

Tan, Y., Nešić, D., and Mareels, I.M.Y. (2006). On non-local stability properties of extremum seeking controllers. *Automatica*, 42, 889–903.

Teel, A.R. and Popović, D. (2001). Solving smooth and nonsmooth multivariable extremum seeking problems by the methods of nonlinear programming. In *Proceedings of American Control Conference*, volume 3, 2394–2399.

Xu, J.X. and Tan, Y. (2003). *Linear and nonlinear iterative learning control*. Spinger.

Yang, Q. and Jagannathan, S. (2012). Reinforcement learning controller design for affine nonlinear discrete-time systems using online approximators. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 42(2), 377–390.

Zhang, C. and Ordóñez, R. (2011). *Extremum-Seeking Control and Applications: A Numerical Optimization-Based Approach*. Advances in Industrial Control. Spinger.