Hierarchical Structures for Economic Repetitive Control

Kirti D. Mishra[1], James Reed[1], Maxwell Wu[2], Kira Barton[2], and Chris Vermillion[1]

***Abstract*— For many emerging repetitive control applications such as wind and marine energy generation systems, gait-cycle following in legged locomotion, remote sensing, surveillance, and reconnaissance, the primary objective for repetitive control (RC) is optimization of a cycle cost such as the lap-averaged power generated and metabolic cost of locomotion, as opposed to the classical requirement of tracking a known reference trajectory by the system output. For this newer class of applications, only a range of reference trajectories suitable for cyclic operation is known a priori, the range potentially encapsulating various operational constraints, and as part of repetitive control, it is desired that over a number of operation cycles, the cycle cost, or the economic metric, is optimized. With this underlying motivation, a hierarchical solution is presented, wherein the inner loop includes a classical repetitive controller that tracks a reference trajectory of known period, and the outer loop iteratively learns the desired reference trajectory using a combination of the system and cost function models and the measured cycle cost. This approach results in optimum steady-state cyclic operation. A steepest descent type algorithm is used in the outer loop, and via Lyapunov-like arguments, the existence of tuning parameters resulting in robust and optimal steady-state cyclic operation is discussed. Appropriate guidelines for parameter tuning are presented, and the proposed method is numerically validated using an example of an inverted pendulum.**

I. INTRODUCTION

A. Motivation

A wide class of engineered systems has benefited from the theory of, and design procedures for, repetitive control (RC) [1]. In particular, note the first application to a power-supply control system [2]. A common theme shared by the research efforts cited in [1] is the *a-priori* knowledge of the desired reference trajectory to be tracked by the system output. While many systems are characterized by a clearly defined reference signal (for instance, consider the application of contouring control in multi-axis motion systems [3]), there exist a host of systems for which the optimal reference signal is not at all obvious. Consider, for example, the applications of legged robots, exoskeletons, and powered prosthetic limbs, wherein repetitive gait following with optimized metabolic cost or reduced mechanical effort is desired, and the task of tuning the parameters (performed manually) is laborious and time-consuming [4]. For numerous energy generation systems such as wind turbines, airborne wind and marine hydrokinetic energy systems that operate cyclically to generate power, and robotic systems such as autonomous vehicles that are required to repetitively follow approximate closed-shaped paths for surveillance and monitoring in minimum time, the process of tuning the desired reference trajectory manually can be very difficult. In these applications, the power generated, the mission time required for surveillance, and metabolic costs are examples of *economic* metrics to be maximized or minimized during cyclic operation.

For the aforementioned applications, the repetitive control problem should include the relevant economic objective function, and repetitive controllers should be designed to maximize or minimize the objective function of interest over a number of system operation cycles. However, this requirement is not accommodated by the existing framework of RC [5], as discussed in [6] (which focused on bipedal walking; however, the overarching conclusion that existing RC is not tailored toward economic metrics is applicable to a large span of applications). This need motivates the primary contribution of the current study. In addition, the requirement of *economic operation* in these systems typically translates into two other requirements of (i) continuous operation and (ii) flexible cycle times. The need for continuous system operation emerges from the fact that an offline phase in between cycles, potentially required for resetting the system to some initial condition, as is the case in iterative learning control (ILC) [7], will result in reduced productivity, and reduced cycle times indicate a further increase in this productivity. In order to achieve these reduced cycle times, or cycle-time optimality, starting from a sub-optimal cycle-time, the cycle time should be allowed to be flexible during operation.

B. Relevant Literature

The class of systems characterized in the preceding paragraph has been the focus of many research efforts lately, with control methodologies being adopted from both the iterative learning control (ILC) framework [8], [9], [10] and the model predictive control (MPC) framework [11], [12], [13]. The choice of an ILC framework in [8], [9], [10] is an interesting one, as the assumption of an initial condition reset is central to the classical formulation of ILC [5], [7]. Despite the challenges involved in using ILC for control of continuously operating systems (see for instance [14]), useful methods are proposed in [8], without any theoretical guarantees, and in [9], with theoretical results that rely on the somewhat artificial assumption of an initial condition reset. While a flexible-time ILC formulation is presented in [10], the proposed approach is again restricted to systems satisfying the assumption of an initial condition reset. This is especially problematic since the requirement of reduced cycle-times for economic operation necessitates that a system

*This work was supported by NSF award number 1727779.

[1]Kirti D. Mishra, James Reed, and Chris Vermillion are with the Department of Mechanical & Aerospace Engineering, North Carolina State University, Raleigh, NC, USA - 276957001 kdmishra@ncsu.edu

[2] Maxwell Wu and Kira Barton are with the Department of Mechanical Engineering, University of Michigan, MI, USA - 48109

finishes a cycle of operation *faster* than it started, which in turn requires that the initial conditions be different from one cycle to the next.

The solutions presented in [11], [12], [13], which adapt tools from the MPC community for use in economic repetitive control, share the idea of constructing safe terminal sets using historic system data from previous trials. Similar to the research efforts cited in the preceding paragraph, the assumption of an initial condition reset is posited in [12]. In contrast to [12], this assumption is relaxed in [11]. In order to implement these MPC-based approaches, a good process model and computational capabilities to perform real-time optimization is required. Also, the required computational effort increases exponentially for nonlinear systems, as a nonlinear optimization problem is required to be solved in real-time [15], which is too massive to handle for many onboard microprocessors across different application domains.

C. Proposed Solution

Motivated by the need for a new framework of repetitive control (RC) that can accommodate economic metrics, and the idea of economic ILC presented in [8], a hierarchical method for economic repetitive control is presented in the current study. The method is represented using a block-diagram in Fig. 1A, wherein a hierarchical structure consisting of (i) a repetitive control system in the inner loop and (ii) an iterative learning method for learning economically optimal reference trajectories in the outer loop. As is clear from Fig. 1A, the inner loop is required to follow the reference trajectory generated by the outer loop and is therefore executed in the time domain, whereas the outer loop iteratively learns the most profitable reference trajectories using trial-to-trial feedback of the measured system variables and the economic cost and is therefore executed in the iteration domain.

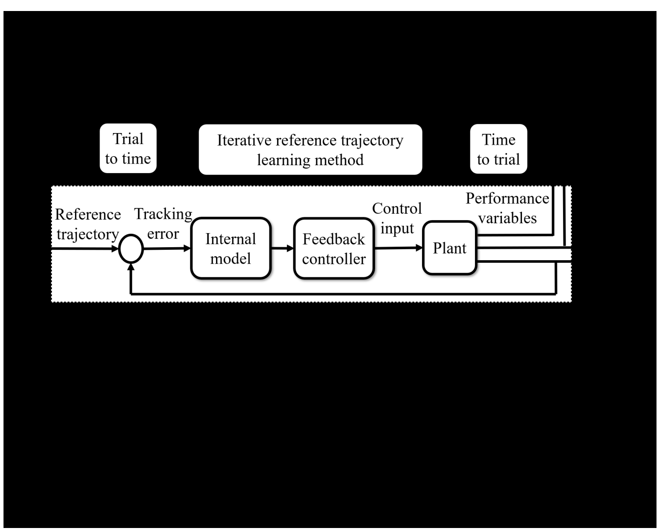

Fig. 1. (A) A block diagram representation of, and (B) a representative evolution of reference and system output trajectories in, economic repetitive control systems

The outer loop, which is expanded in Fig. 1A, ultimately updates the reference trajectory based on the observed economic cost and cycle tracking error. Specifically, an update is triggered (via the "Reference trajectory update check" block) whenever the cycle tracking error falls below a specified threshold, the number of cycles executed for a particular reference trajectory exceeds a specified threshold, or a combination of the two conditions. Instances where the trigger is based entirely on tracking error are referred to as *patient* economic RC, whereas instances where the trigger is forced based on a specified number of cycles are referred to as *restless* economic RC. When triggered, the reference trajectory is updated based on the performance over the most recent cycle corresponding to the previous reference trajectory, referred to as the *trial*. A typical evolution of the reference trajectory and system output in time is shown in Fig. 1B, from where the relationship of a cycle and a trial may be noted. In Fig. 1B, the *intervals* of cyclic operation are separated by the vertical dotted lines. It is important to note that in ILC, a trial is defined by the duration of the desired (trial-invariant) reference trajectory, and even for systems with the trial-varying trial duration [16], the existence of an offline phase clearly differentiates any two consecutive trials. The continuous nature of the applications considered in this study requires a clear economic-operation-centric definition of a trial.

There are three major contributions of this work: (i) a systematic framework for addressing repetitive control of systems with the primary objective of optimizing an economic metric, (ii) a procedure for judiciously updating the reference trajectory in the outer loop of the proposed hierarchical structure, and (iii) theoretical results that guarantee that the parameters of the corresponding algorithm can be tuned to asymptotically achieve economically optimal operation. In addition to these, guidelines are discussed for designing the inner loop, and the proposed method is validated numerically for a nonlinear system - the inverted pendulum.

II. ECONOMIC REPETITIVE CONTROL

In this section, the method of economic RC is detailed for nonlinear and linear systems. For the outer loop, an approach similar to iterative solution algorithms for static optimization problems is used here. An a-priori known range of reference trajectories to which the optimum solution belongs forms the set of candidate solutions, and iterative reference trajectory update laws must be designed that mimic the mentioned class of static optimization solution methods. For the inner loop, repetitive controllers are designed such that all of the candidate reference trajectories' solutions can be tracked.

A. Problem formulation

The class of (potentially nonlinear) systems considered in this study is described in (1), wherein $x_p \in \mathbb{R}^{n_{x_p}\times 1}, u \in \mathbb{R}^{n_u \times 1}, z \in \mathbb{R}^{n_z\times 1}, y \in \mathbb{R}^{n_y \times 1}$ denote the system state, control input, performance variables, and system output respectively:

$$\dot{x}_p = f(x_p, u),\ z = g(x_p, u),\ y = h(x_p). \tag{1}$$

The functions f, g, and h are assumed to be smooth. Additionally, it is further assumed that the solution is unique and exists for all time, a bare minimum requirement for analysis and control of steady-state system operation. The cost function J is a function of the performance variable z, which is assumed to be measured. In turn, z is a function of (a subset of) the state variable x_p and control input u. For the system in (1), assumptions on invertibility and uncertainty are posited in A1 and A2.

A1. The system in (1) is assumed to be invertible, i.e., for a given desired reference trajectory r, $\exists$ *unique* steady-state solutions x_d and u_d such that $x_p = x_d, u = u_d$, and $y = r$ which satisfy (1). Similar to r, x_d and u_d are also functions with the same (fundamental and higher order) periodicity as r.

A2. The functions f, g, and h are known to belong to the uncertain sets $\mathcal{F}, \mathcal{G}$, and $\mathcal{H}$, respectively, known only approximately for control design, the approximations being $\hat{f}, \hat{g}$, and $\hat{h}$ respectively.

In addition, we allow the desired reference r, a periodic function of time, to be parameterized by $\theta^r \in \mathbb{R}^{n_r \times 1}$. For instance, if $r = a_r \sin \omega t, \theta_r = [a_r \ \omega]^T$, and $n_r = 2$. Such parameterization of the reference trajectory can reduce the number of variables that are adjusted by the outer loop.

As the goal of economic RC here is optimal (steady-state) cyclic performance, the idea of a *cycle* is central to this study. In contrast to the traditional RC literature, wherein a cycle is typically defined to be the same as the time-invariant period of the reference trajectory to be tracked, owing to the requirement of flexible cycle times here, as described in Section I, the cycle-time itself is time-varying. In order to ensure that the comparison of economic performance for any two operational cycles is fair, a cycle in this study is defined using a non-dimensional path parameter ρ that spans the range [0 1], similar to [8], [9]. The parameter ρ is equal to 0 when the system state x is at the *start line*, and as x traverses through a closed orbit (a periodic solution) in the state space, the desired trajectory x_d being one such closed orbit, ρ increases from 0 to 1, and is equal to 1 when x returns to the start line [12]. The start line for a closed-orbit solution is defined as the line joining the center of the closed orbit to some arbitrarily fixed point on it.

B. Outer loop design

As part of outer loop design in economic RC, the goal is to develop optimization algorithms and associated update rules for θ^r such that the economic cost J, which is computed over one cycle of system operation, is optimized in real-time, ultimately resulting in economically optimum steady-state cyclic operation. For an operation cycle with the initial conditions $x_p(t_0^{cyc})$ for the system in (1), where t_0^{cyc} denote the time at which the operation cycle starts, the cost function is a function of the commanded reference θ^r and $x_p(t_0^{cyc})$, as these uniquely determine the performance variables z over the cycle being analyzed.

As the cumulative tracking error over the cycle, or cycle tracking error, approaches zero asymptotically, the initial condition $x_p(t_0^{cyc})$ of a cycle is uniquely determined by the steady state solution x_d, which in turn, uniquely depends on θ^r (per A1). As will be discussed shortly, it is useful to express the economic cost function as described in (2), wherein the term δ diminishes to zero as the cycle tracking error approaches zero. So, as $t \to \infty$, $\delta \to 0$. In order to arrive at (2), the cost function is assumed to be smooth, allowing for a Taylor series expansion of J about θ^r and $x_d(t_0^{cyc})$.

$$J(\theta^r, x_p(t_0^{cyc})) = J(\theta^r, x_d(t_0^{cyc})) + \delta(\theta^r, x_d(t_0^{cyc}) - x_p(t_0^{cyc})) \tag{2}$$

In (2), as $x_d(t_0^{cyc})$ is uniquely determined by θ^r, it is reasonable to drop the second argument of J during steady state and, consequently, the cost function with only one argument should be understood to be the *steady-state cost function*.

It is noteworthy that if the system (1) and economic cost J are accurately known, the steady-state optimization problem, which is a nonlinear programming problem, can be solved offline. However, for practical applications of interest, the system and the cost are only partially known at best. Therefore, there is a need to include trial-to-trial feedback from the system during cyclic operation. These arguments lead into the following set of assumptions for the system and economic cost.

A3. $\exists$ a convex set Θ_R such that the cost function $J(\theta^r)$ is convex in Θ_R, where the unique optimizer is denoted as $\theta_r^\star \in \Theta_R$.

A4. The cost function available for outer loop design, which is denoted as J', is also convex in Θ_R with the unique minimizer $\hat{\theta}_r^\star$, and $|\theta_r^\star - \hat{\theta}_r^\star| < \Delta^\star$. Here, $\Delta^\star$ is not known a priori.

In A3, convexity of the steady-state cost function $J(\theta^r)$ is assumed, and in A4, the cost function used for control design is characterized.

Remark 1 *The economic cost will typically be a nonlinear function of the system state x_p and control input u, and the exact dependence $J(\theta^r)$ during steady state operation depends on the system* (1) *and the repetitive controller of the inner loop. As long as the inner loop converges to a steady state, potentially with non-zero tracking error (see Remark 3), the mapping $J(\theta^r)$ is well defined. The mapping $J(\theta^r)$ and, in particular, $J(\theta^\star)$, is a function of the inner loop that includes the system and the repetitive controller to be designed. The method presented and formally discussed for the case of zero steady-state tracking error in this section easily extends to the case of non-zero steady-state tracking error.*

The *reference trajectory update check* and *reference trajectory update law* blocks, the tuning guidelines for the proposed method, and details of the method used here for computing the cost function gradient are described next.

1) Reference trajectory update check: As briefly alluded to earlier, an approach similar to iterative solution algorithms for static nonlinear optimization is used here for designing the outer loop. More specifically, the economic cost to be optimized, a static (in time) nonlinear function, is optimized iteratively using the same principle as that used by iterative solution algorithms such as steepest descent with the key difference that trial-to-trial feedback is used to compensate for uncertainty in the system model and/or cost function model. Central to any iterative approach is the definition of an iteration or a trial. In general, after the desired reference trajectory parameter θ^r is updated, there will exist initial transients over one or more cycles, as the system output seeks to track the newly defined trajectory. In order to conduct a meaningful reference trajectory at the outer loop, it is essential to wait until these transients have sufficiently died off before triggering an update; however, there exists a tradeoff between waiting many iterations for all transients to settle and aggressively adjusting the reference trajectory with the goal of converging quickly to the optimal trajectory.

To balance these two goals, a reference trajectory update is performed whenever one of the following two conditions are satisfied, (i) $\int_0^1 \|e(\rho)\|^2 d\rho \leq \overline{e}_c$, where ρ denotes the path parameter defined earlier, and $\overline{e}_c$ represents a threshold on tracking error below which a reference trajectory change will be triggered, or (ii) $n_{\text{wait}} \geq n_{\max}$, where n_{wait} represents the number of cycles that have occurred since the last reference trajectory update, and $n_{\max}$ represents a specified number of cycles beyond which a reference trajectory change will be triggered. We will refer to the first condition as a *patient RC trigger* (owing to the fact that we are waiting for the error to fall below a threshold) and the latter condition as a *restless RC trigger* (owing to the fact that we are forcing a reference trajectory change, irrespective of the error). In Fig. 1B, a trial is defined according to the patient RC trigger, which may be noted from the fact that the first two intervals are three cycles long whereas the third interval is only two cycles long.

Typically, the threshold $\overline{e}_c$ will be small for economic RC systems with zero steady-state tracking error, but not equal to zero, implying that the measured cost J_m^j after the j^{th} trial is equal to the cost $J(\theta^r, x_p(t_0^{cyc}))$ in (2), wherein δ represents a *transient cost* that vanishes as $\overline{e}_c \to 0$. If the repetitive controller of the inner loop results in a quick convergence of the tracking error to zero, a small value of $\overline{e}_c$ may be used. However, if the convergence is slow, a larger $\overline{e}_c$ may be required for reasonable convergence times, the larger value resulting in more frequent updates. In such a setting, the magnitude of $\overline{e}_c$ should be tuned to ensure such that the cost improves before every reference trajectory update, thereby ensuring that the closed loop orbit (periodic solution) in the state space is *closer* to the true optimal closed loop orbit.

2) Reference trajectory update method: After every trial, the reference trajectory is updated as per the steepest descent type algorithm of (3), wherein j denotes the trial index, $\Gamma \in \mathbb{R}^{n_r \times n_r}$ denotes the *step length matrix*, and $\hat{q}_j$ denotes the estimated gradient, which is used to establish a suitable descent direction for optimization, the computation of which is described shortly. Here, $\Gamma = \texttt{diag}(\gamma_1 \ \dots \ \gamma_{n_r})$.

$$\theta_{j+1}^r = \theta_j^r - \Gamma \hat{q}_j \tag{3}$$

There are three distinct phases in the iterative solution proposed here for optimization of the cost function. During the first phase, at the start of this iterative process, only the gradient computed using the system and cost function models is available. A minimum of two trials are required in this phase to create an estimate of the cost function gradient using the measured cycle cost values. Let $q_j = \partial J / \partial \theta^r |_{\theta^r = \theta_j^r}$, $q_j' = \partial J' / \partial \theta^r |_{\theta^r = \theta_j^r}$, and q_j^m denote, at $\theta^r = \theta_j^r$, the true gradient, the gradient computed using the system and cost function models, and the gradient estimated using the measurements J_j^m of the economic cost, respectively. During the second phase, the gradients q_j' and q_j^m are used in combination for iterative optimization. During the third phase, θ^r is in between $\theta_r^\star$ and $\hat{\theta}_r^\star$ (see A4), and the gradients q_j' and q_j^m provide opposite directions for descent. Based on these considerations, the cost function gradient $\hat{q}_j$ is estimated using (4), wherein $0 < \alpha_j < 1$.

$$\hat{q}_j = \begin{cases} \alpha_j q_j' + (1-\alpha_j) q_j^m, & \text{If sign } q_j' = \text{sign } q_j^m \\ q_j^m, & \text{otherwise} \end{cases} \tag{4}$$

Regarding initialization of iterative learning, θ_0^r should be chosen to be different than $\hat{\theta}_r^\star$, as the gradient $q_0' = 0$ at $\hat{\theta}_r^\star$, implying that the search cannot continue, and for practical applications, $\hat{\theta}_r^\star \neq \theta_r^\star$.

In what follows, it will be shown that the steepest-descent type algorithm of (3) and (4) with the patient RC trigger can be tuned effectively to *robustly* optimize the steady-state operation of nonlinear repetitive systems with uncertain models and uncertain economic cost functions. These arguments easily extend to economic RC with the restless RC trigger. Let $\underline{\gamma} \triangleq \min\{\gamma_1 \ \dots \ \gamma_{n_r}\}$.

Theorem 1. *There exists $\underline{\gamma}, \overline{e}_c$, $n_{\max}$, and a sequence of weighing parameters α_j such that starting from any $\theta_0^r \in \Theta_R$, as $j \to \infty$, the economic cost $J(\theta_j^r)$ of the closed-loop system defined by* (1)-(4) *converges to* $J(\theta_r^\star)$.

Proof: The proof consists of two parts. First, using a Lyapunov-like function $\delta\theta_j^{rT} \delta\theta_j^r$ ($\delta\theta_j^r \triangleq \theta_r^\star - \theta_j^r$) and convexity of J in Θ_R, robust convergence will be shown possible for a sufficiently small $\underline{\gamma}$ and some sequence α_j for which $\exists$ some j' such that $\forall j > j'$, the sign of q_j^m can be ensured to be the same as the sign of q_j. Second, it will be shown that the condition on the sign of q_j^m can always be ensured for sufficiently small $\overline{e}_c$ and $\underline{\gamma}$.

Let $V_j = \delta\theta_j^{rT} \delta\theta_j^r$ denote the (Lyapunov-like) distance function between θ_j^r and $\theta_r^\star$. To analyze the evolution of $\delta\theta_j^r$ in the trial-domain, the discrete-time derivative of V_j or the difference of V_j and V_{j+1} is evaluated along an arbitrary closed-loop trajectory of (1)-(4), as shown in (5) and (6).

$$V_{j+1} = (\theta_r^\star - \theta_{j+1}^r)^T (\theta_r^\star - \theta_{j+1}^r) \tag{5}$$

$$V_{j+1} = V_j + 2\underline{\gamma}(\theta_r^\star - \theta_j^r)^T \hat{q}_j + \underline{\gamma}^2 \hat{q}_j^T \hat{q}_j \tag{6}$$

As $\underline{\gamma} \to 0$, the term $\underline{\gamma}^2 \hat{q}_j^T \hat{q}_j \ll 2\underline{\gamma}(\theta_r^\star - \theta_j^r)^T \hat{q}_j$ for any given θ_j^r (no matter how small), implying that the relative values of V_j and V_{j+1} may be evaluated by analyzing $2\underline{\gamma}(\theta_r^\star - \theta_j^r)^T \hat{q}_j$. As an assumption for this first part, let the sign of $\hat{q}_j$ be the same as that of q_j $\forall j > j'$ for some j'. Owing to this assumption, and A3 - $\theta_r^\star \in \Theta_R$ and J is convex in Θ_R, the term $2\underline{\gamma}(\theta_r^\star - \theta_j^r)^T \hat{q}_j \prec 0$, where $\prec$ denotes the element-wise inequality operator, as long as $\theta_j^r \in \Theta_R$. Now consider $j = 0$, as $\theta_0^r \in \Theta_R$, $V_1 < V_0$ which, in turn, implies that $\theta_1^r \in \Theta_R$. Extending this argument through $j \to \infty$, $\theta_j^r \to \theta_r^\star$.

For the second part, in order to show that $\exists\ j'$ such that $\forall\ j > j'$, the sign of $\hat{q}_j$ can be ensured to be the same as the sign of q_j, consider that as $\overline{e}_c \to 0$, $J(\theta^r, x_p(t_0^{cyc})) \to J(\theta^r, x_d(t_0^{cyc})) = J(\theta^r)$ in (2). In addition, if $\underline{\gamma}$ is small (not necessarily $\underline{\gamma} \to 0$, i.e. arbitrarily small), implying that θ_j^r and θ_{j-1}^r are close, using a simple finite-difference method, the measured cost function values $J^m_{\theta_j^r}$, $J^m_{\theta_{j-1}^r}$, and $j' = 2$, q_j^m is easily computed such that the sign of q_j^m is the same as the sign of q_j. Now let the sequence of weighing parameters $\alpha_j = \underline{\alpha}$, and $\underline{\alpha} \to 0$. From (4), it can be noted that under such a choice of α_j, the sign of $\hat{q}_j$ is the same as the sign of $q_j\ \forall\ j > j'$. $\square$

Remark 2 *As the reference trajectory parameter θ_j^r approaches the true optimal $\theta_r^\star$, the cost function gradient reduces in magnitude, implying an increasingly small change in θ_j^r. In the limit, the change is so small that the cycle tracking error asymptotically goes to zero as $\theta_j^r \to \theta_r^\star$. More generally, the cycle tracking error goes to a minimum steady-state value (see Remark 1).*

3) Tuning guidelines: During the first phase, the gradient q_j' should be used with smaller values of the step lengths $\gamma_i\ \ldots\ \gamma_{n_r}$. On the one hand, the smaller step lengths allow simple computation of q_j^m via the finite-difference method. On the other hand, these prevent larger, and potentially incorrect, changes in θ_j^r. The nominal models of the system and cost function are recommended to be used for tuning $\gamma_1\ \ldots\ \gamma_{n_r}$ via simulations. Specifically, by using a θ_0^r sufficiently different from $\hat{\theta}_r^\star$, it can be ensured that the sign of q_j' (the gradient computed using uncertain models) and q_j are the same for θ_j^r close to θ_0^r (see A3). The implication here is that (sufficiently small) $\gamma_1\ \ldots\ \gamma_{n_r}$ may be tuned using the simulated cycle costs.

During the second and third phases, where the gradient q_j^m computed via measurements is available, higher values of $\gamma_1\ \ldots\ \gamma_{n_r}$ may be used. In these phases, it is desirable to progressively weight q_j^m higher as $j \to \infty$. Consequently, a potential choice for $\alpha_j = \alpha^{j-j'}$ with $\alpha < 1$. While the desired value of α is application dependent, a good initial value for further tuning is 0.5, which (at $j = j' + 1$) weighs q_j' and q_j^m equally.

The value of $\overline{e}_c$ required for successful economic RC is heavily dependent on the inner loop, and more specifically, on its tracking performance. As was discussed in the second part of the proof of Theorem 1, $\overline{e}_c$ should be chosen such that δ in (2) is small, ensuring that the sign of the true gradient q_j is estimated accurately. If the tracking error of the inner loop converges quickly to a non-zero steady-state value, $\overline{e}_c$ may be set slightly larger than this value. If the convergence is slow, a higher value of $\overline{e}_c$ may be used, and the tuning of $\overline{e}_c$ should proceed using models or experiments. The true gradient, q_j, should first be estimated by setting $\overline{e}_c$ close to the non-zero steady-state tracking error and setting $\gamma_1\ \ldots\ \gamma_{n_r}$ to small values. Once the sign of the estimated true gradient, q_j, is known, using a few more trials with small $\gamma_1\ \ldots\ \gamma_{n_r}$, the size of $\overline{e}_c$ should be progressively increased while ensuring that the sign of q_j^m is the same as this sign. This tuning process must be repeated for a few randomly chosen θ_0^r, and the minimum of all such tuned $\overline{e}_c$ should be retained. It is noted here that for economic RC systems characterized by non-zero steady-state tracking error of the inner loop (see Remark 1), $\overline{e}_c$ should always be set higher than the non-zero steady-state cycle tracking error, else the static optimization problem being solved here iteratively will be ill-posed in the sense that every initial candidate solution θ_0^r will be a local optimum.

4) Cost function gradient estimation: In addition to the tuning parameters discussed in the preceding paragraphs, successful implementation of economic RC requires accurate computation of the gradients q_j' and q_j^m. Using the models for the system and cost function, the gradient q_j' is proposed to be computed numerically using sensitivity simulations with respect to $\theta_j^r = \theta^r$, and the method of finite difference. As it may be challenging to simulate nonlinear systems online, a wide range of simulations may be performed offline, and the corresponding economic cost can stored in the form of a map. In this method, the system is simulated until it reaches steady-state operation. Online estimation of the gradient of a cost function has been a focus of study for many research efforts, as gradients plays an important role in many fields of control theory [17], [18]. While a wide range of techniques ranging from simple finite-difference-based to those using time-varying Kalman filters are available for this purpose [19], in this study, a method that uses n_m past measurements of the cycle cost $J_j^m,\ J_{j-1}^m\ \ldots J_{j-n_m+1}^m$ for computing a local polynomial approximation of the cost function J is used, the approximation subsequently being used to evaluate the gradient q_j^m at $\theta_j^r = \theta^r$.

C. Inner loop design

Output tracking of periodic reference trajectories is an established sub-field of control theory. A brief discussion on repetitive controllers for trajectory tracking is included here for nonlinear and linear systems. The mathematical formulation of nonlinear repetitive control falls under the class of nonlinear output regulation or servomechanism problems [20], wherein the asymptotic tracking or rejection of (not necessarily periodic) signals generated by exo-systems are studied. Isidori & Byrnes introduced the much celebrated idea of zero error manifolds in [20], and derived the necessary condition for the existence of a solution to the nonlinear regulation problem. Per this necessary condition, every state and (feedforward) input trajectory (x_d, u_d) on this zero error manifold satisfies the nonlinear *regulator equations*. A feedback (linear) controller that locally stabilizes this zero error manifold, i.e. $x - x_d$, is guaranteed to ensure asymptotic convergence of the tracking error to zero.

For linear systems, an internal model is included in the closed-loop (see Fig. 1) and the feedback controller is designed such that the resulting closed-loop is asymptotically stable. The internal model here is a function of the periodicity of the reference trajectory and, thus, for the application of economic RC, the feedback controller should stabilize the set of closed-loop systems resulting from a range of possible time periods of the reference trajectory.

III. ECONOMIC RC OF AN INVERTED PENDULUM

The method of economic RC described in Section II is validated numerically here using an inverted pendulum example, the governing equation for which is described in (7), wherein $I_o, \phi, m,$ and L denote the effective inertia, rotation angle, mass, and length of the pendulum respectively, and g, b and τ denote the damping coefficient, acceleration due to gravity, and control input torque respectively. The rotation angle ϕ is measured from the vertical downwards position – the stable equilibrium of a simple pendulum.

$$I_o\ddot{\phi} = -mgL\sin\phi - b\dot{\phi} + \tau \tag{7}$$

The inverted pendulum example has been used widely in the controls literature. In addition to being theoretically interesting, it is reflective of the basic characteristics of several dynamical systems, including for example bipedal walking motion [21].

As part of control objectives, it is desired that the inverted pendulum oscillates about the unstable equilibrium $\phi = \pi$ with the amplitude of oscillation ϕ_o and cycle time T such that $\underline{\phi}_o < \phi_o < \overline{\phi}_o$ and $\underline{T} < T < \overline{T}$. An economic cost that penalizes the distance between the system trajectories of (7) and an *unknown* desired reference trajectory is to be minimized in steady-state operation.

A model of the cost function that measures the distance between the system trajectories of (7) and a decent guess of this desired reference trajectory is used for computing q_j'. Typically, such a guess is known from experience. Using the procedure of harmonic balance described in Section II, the internal model, i.e. the coefficients of the (truncated) Fourier series representation of the desired state-space trajectory $\hat{x}_d$ and the corresponding feed-forward control input $\hat{u}_d$, is computed for a discretized 2-d grid in the space of ϕ_o and T and is stored in the form of a look-up table as part of the economic RC implementation. The linear controller K in (**??**) is designed using the linear quadratic regulator (LQR) technique. Through the relative weighting in the LQR cost function, the rate of convergence of the system state $x = [\phi\ \dot{\phi}]^T$ to the desired trajectory x_d may be manipulated.

An instance of the simulation results is shown in Figs. 2-4, wherein the economic RC method with the patient RC trigger described in Section II is demonstrated to be effective in iteratively optimizing the cycle cost over time for the ongoing example of inverted pendulum oscillations. Figs. 2-4 represent results corresponding to one of many simulations with randomly chosen system parameters belonging to an uncertainty range of 20% that were performed for validation. In addition, via these figures, the role of the gradient estimated via measurements of the cycle cost is also highlighted. More specifically, two cases are compared: (i) $\alpha_j = 0\ \forall\ j$ and (ii) $\alpha_j = \alpha^{j-j'}$. In Fig. 2, the unknown true optimal trajectory in the state-space and a guess of it are represented by blue and red solid lines respectively, the system trajectories are represented by the yellow and green solid lines for the cases $\alpha_j = 0$ and $\alpha_j = \alpha^{j-j'}$ respectively, and the converged cycle trajectories for the two cases are represented by the dashed red and blue lines respectively. For both the cases, the same initial condition is used, which is shown by the black marker in Fig. 2, and the system trajectories correspond to a time duration of 175 s. During this duration, the reference trajectory parameter θ_j^r is updated when the cycle tracking error is computed to be less than $\overline{e}_c = 1$, and the trajectory of θ_j^r in the plane of $T - \phi_o$ is shown in Fig. 3, wherein the contour plot of the unknown true cost function J is used to indicate the progress made during cyclic operation. As can be noted from Fig. 2 (and Fig. 3), the commanded reference trajectory is updated such that the system trajectory (green) for the case $\alpha_j = \alpha^{j-j'}$ converges to the desired economically optimal trajectory (blue).

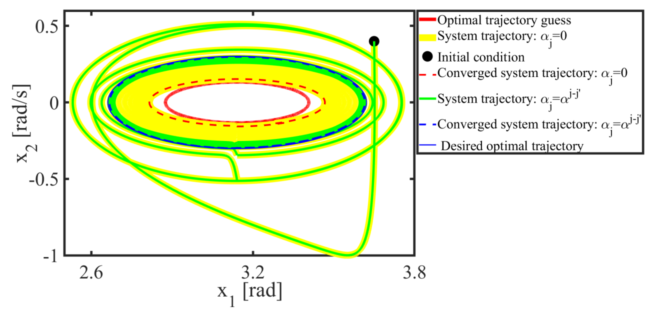

Fig. 2. Economic repetitive motion in the state space.

The two cases being compared in Fig. 2 are use the same initial reference parameter value $\theta_0^r = [\phi_0^o\ \ T_0]^T$ with $\phi_0^o = 36.5$ deg and $T_0 = 7.5$ s, which is represented by the black circular marker in Fig. 3. In this figure, the unknown true optimal value, $\theta_r^\star = [27.5 \text{ deg}\ \ 10 \text{ s}]^T$, and the guess of it, $\hat{\theta}_r^\star = [15.25 \text{ deg}\ \ 12.5 \text{ s}]^T$, are represented by the blue and red circular markers respectively, and the trajectories of θ_j^r are represented by red asterisk and blue square markers for the cases $\alpha_j = 0$ and $\alpha_j = \alpha^{j-j'}$ respectively. It is noted here that θ_0^r is chosen sufficiently far from $\hat{\theta}_r^\star$, per the discussion regarding initialization presented before (4). Specifically, choosing θ_0^r close to $\hat{\theta}_r^\star$ may result in the convergence of θ_j^r during the first phase to the red marker on the contour plot in Fig. 3 or, alternatively, a greater distance allows for the true descent direction to be figured out during cyclic operation.

As briefly alluded to earlier in Section II, the cost function model ensures reasonable (sub-optimal) progress during the first phase or, in general, whenever the gradient q_j^m estimated using measurements is not very accurate. For the results shown, $\alpha = 0.5$ and $j' = 4$, implying that $\forall\ j \geq 5$, the computed gradient q_j^m is used in conjunction with q_j' as described in (4), and at $j = 5$, the two gradients are weighted equally. It may also be noted in this figure that that the commanded amplitude angle ϕ_o and the cycle time

T remain within the bounds of $\underline{\phi}^o = 10$ deg, $\overline{\phi}^o = 45$ deg and $\underline{T} = 5$ s, $\overline{T} = 15$ s respectively.

The evolution of cycle cost with respect to time is represented by the blue cross and red circular markers in Fig. 4 for the cases $\alpha_j = \alpha^{j-j'}$ and $\alpha_j = 0$ respectively. It may be verified from this figure that (i) the cost reduces initially for the case $\alpha_j = 0$, but as θ_j^r approaches $\hat{\theta}_r^\star$, the cost starts to increase, and (ii) the cost improvement is (almost) monotonic for the case $\alpha_j = \alpha^{j-j'}$.

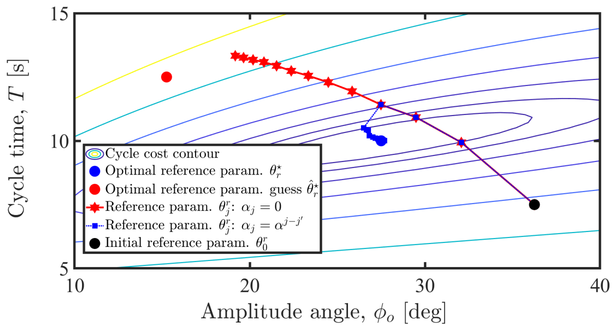

Fig. 3. Optimization of the reference trajectory parameter θ_j^r.

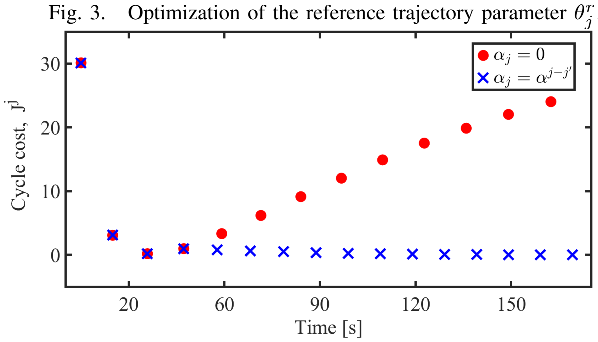

Fig. 4. Evolution of the cycle cost

IV. CONCLUSIONS

Motivated by the need for a repetitive control (RC) framework and associated design methods that result in economically optimal steady-state cyclic operation, this study presents a hierarchical solution with the inner loop using the conventional repetitive controllers for tracking a periodic reference, and the outer loop using a steepest-descent type iterative algorithm that updates the reference trajectory such that the optimum cyclic operation results. It has been shown that the proposed algorithm can be tuned to result in robust convergence of the reference trajectory to the economically optimal one, and appropriate tuning guidelines are discussed. The economic RC method is numerically validated for the example of an inverted pendulum, and the results presented demonstrate that the developed RC method is effective in iteratively learning and tracking the optimal reference trajectory.

REFERENCES

[1] E. Kurniawan, Z. Cao, O. Mahendra, and R. Wardoyo, "A survey on robust repetitive control and applications," in *2014 IEEE International Conference on Control System, Computing and Engineering (ICCSCE 2014)*, 2014, pp. 524–529.

[2] T. Inoue, M. Nakano, T. Kubo, S. Matsumoto, and H. Baba, "High accuracy control of a proton synchrotron magnet power supply," *IFAC Proceedings Volumes*, vol. 14, no. 2, pp. 3137–3142, 1981, 8th IFAC World Congress on Control Science and Technology for the Progress of Society, Kyoto, Japan, 24-28 August 1981.

[3] C. Hu, B. Yao, Z. Chen, and Q. Wang, "Adaptive robust repetitive control of an industrial biaxial precision gantry for contouring tasks," *IEEE Transactions on Control Systems Technology*, vol. 19, no. 6, pp. 1559–1568, 2011.

[4] A. M. Simon, K. A. Ingraham, N. P. Fey, S. B. Finucane, R. D. Lipschutz, A. J. Young, and L. J. Hargrove, "Configuring a powered knee and ankle prosthesis for transfemoral amputees within five specific ambulation modes," *PloS one*, vol. 9, no. 6, p. e99387, 2014.

[5] R. W. Longman, "Iterative learning control and repetitive control for engineering practice," *International Journal of Control*, vol. 73, no. 10, pp. 930–954, 2000.

[6] R. Longman and K. Mombaur, *Investigating the Use of Iterative Learning Control and Repetitive Control to Implement Periodic Gaits*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 189–218.

[7] D. A. Bristow, M. Tharayil, and A. G. Alleyne, "A survey of iterative learning control," *IEEE Control Systems Magazine*, vol. 26, no. 3, pp. 96–114, 2006.

[8] M. Cobb, M. Wu, K. Barton, and C. Vermillion, "Flexible-time economic iterative learning control: A case study in airborne wind energy," in *2019 IEEE 58th Conference on Decision and Control (CDC)*, 2019, pp. 5580–5586.

[9] M. K. Cobb, K. Barton, H. Fathy, and C. Vermillion, "Iterative learning-based path optimization for repetitive path planning, with application to 3-d crosswind flight of airborne wind energy systems," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 4, pp. 1447–1459, 2020.

[10] M. J. Wu, M. Cobb, C. Vermillion, and K. Barton, "A flexible-time iterative learning control framework for linear, time-based performance objectives," in *2020 American Control Conference (ACC)*, 2020, pp. 4792–4797.

[11] M. Brunner, U. Rosolia, J. Gonzales, and F. Borrelli, "Repetitive learning model predictive control: An autonomous racing example," in *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 2017, pp. 2545–2550.

[12] U. Rosolia and F. Borrelli, "Learning model predictive control for iterative tasks. a data-driven control framework," *IEEE Transactions on Automatic Control*, vol. 63, no. 7, pp. 1883–1896, 2018.

[13] N. Scianca, U. Rosolia, and F. Borrelli, "Learning model predictive control for periodic repetitive tasks," in *2020 European Control Conference (ECC)*, 2020, pp. 29–34.

[14] K. L. Moore, "A non-standard iterative learning control approach to tracking periodic signals in discrete-time non-linear systems," *International Journal of Control*, vol. 73, no. 10, pp. 955–967, 2000.

[15] Y. Long, L. Xie, and S. Liu, "Nontracking type iterative learning control based on economic model predictive control," *International Journal of Robust and Nonlinear Control*, vol. 30, no. 18, pp. 8564–8582, 2020.

[16] P. M. Sammons, D. Hoelzle, and K. Barton, "Time-scale transformed iterative learning control for a class of nonlinear systems with uncertain trial duration," *IEEE Transactions on Control Systems Technology*, vol. 28, no. 5, pp. 1972–1979, 2020.

[17] M. Krstić and H.-H. Wang, "Stability of extremum seeking feedback for general nonlinear dynamic systems," *Automatica*, vol. 36, no. 4, pp. 595–601, 2000.

[18] D. Krishnamoorthy, E. Jahanshahi, and S. Skogestad, "Feedback real-time optimization strategy using a novel steady-state gradient estimate and transient measurements," *Industrial & Engineering Chemistry Research*, vol. 58, no. 1, pp. 207–216, 2019.

[19] J. J. Ryan and J. L. Speyer, "Peak-seeking control using gradient and hessian estimates," in *Proceedings of the 2010 American Control Conference*, 2010, pp. 611–616.

[20] A. Isidori and C. I. Byrnes, "Output regulation of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 35, no. 2, pp. 131–140, 1990.

[21] M. Srinivasan, "Fifteen observations on the structure of energy-minimizing gaits in many simple biped models," *Journal of the Royal Society Interface*, vol. 8, no. 54, pp. 74–98, 2011.