A Unified Analysis Tool in Iterative Learning Control: Composite Energy Function

Gijo Sebastian, Ying Tan, Denny Oetomo

***Abstract*— Usually, the convergence analysis of a continuous-time iterative learning control (ILC) algorithm is carried out by using either Contraction Mapping (CM) or Composite Energy Function (CEF) based methods. They are subtly different in terms of types of systems to which it can be applied, the convergence properties, convergence condition and the standing assumptions. Although the link between these two methods has been observed in literature, there is no clear statement to show that CEF-based method can be served as a unified analysis tool. This paper shows this fact by re-proving the well-known CM-based convergence results of P-type ILC algorithms using a novel CEF method.**

I. INTRODUCTION

Iterative Learning Control (ILC) is a control technique that is widely used to "learn" a desired control input for systems performing a repetitive tracking task on a finite-time interval. Over the last three decades, many ILC algorithms have been developed with many successful applications (For more details, see survey papers [1]–[4] and references therein).

Contraction Mapping (CM) and Composite Energy Function (CEF) based methods are widely used for the convergence analysis of an ILC algorithm in a continuous-time domain [5].

CM-based methods were established in the pioneering works of ILC [6]–[10]. Usually, CM-based design is simple and requires relatively less knowledge of the system dynamics [5]. The standing assumption is that the system of interests satisfies so-called global Lipschitz condition (GLC). By using the time-weighted norm, the dynamics of the system can be ignored and the uniform convergence of tracking error or the input error (the error between the desired input and the current input) can be ensured if some convergence condition, which is closely related to the relative degree of the dynamic system, is satisfied.

On the other hand CEF or energy function (EF) based techniques were initially used for cases where the CM-based methods cannot be directly applied , see, for example, [11]–[15]. By requiring more knowledge of the system, such as some stability properties in time-domain, CEF-based ILC design can relax some standing assumptions needed in CM-based methods. For example, the GLC condition can be relaxed [4], [16] or the relative degree requirement [5], [11], [13], [17]. As the performance in time-domain is always incorporated into the CEF, CEF-based ILC works well when feedback laws are used in combination with feed-forward ILC. Without requiring the convergence condition, the CEF method shows that at each time instant, the proposed CEF with the proposed ILC updating law decreases along iteration domain. This leads to the point-wise convergence of tracking error or input error. However extra conditions are needed to show the uniform convergence. Although, these two methods are different, both of them use its own techniques to ignore the dynamics of the system. A monotonic convergence in terms of a time-weighted norm in CM methods is equivalent to the decreasing energy function in CEF methods. Hence route of convergence in terms of supremum norm is unknown in both cases. The differences between these two designs are summarized in Table I.

The CEF-based analysis incorporates information in time-domain and iteration-domain while the CM-based method only focuses on information in iteration-domain. It is intuitively clear that the CEF-based analysis method could be served as a unified analysis tool to show the convergence. However, how to use the convergence condition in the CEF-based method has not been fully investigated. To the best of author's knowledge, how to use CEF method to show the convergence of ILC algorithms, which are designed based on CM method, has not been addressed in the ILC literature, except some preliminary works in [18] where a dual loop ILC was investigated in which the first ILC is based on CM and the second ILC is based on CEF.

This paper proposes two novel CEFs, which can naturally incorporate the convergence condition from CM-based ILC design into the convergence analysis. To illustrate this idea, the well-known P-type feed-forward ILC algorithm for a non-linear multi-input-multi-output square affine system with zero relative degree is used. Two different convergence conditions are considered: one is in terms of output tracking error and the other is in terms of input error. Instead of using standard CM-based analysis, the convergence of both cases can be re-proven by using such novel CEFs. This example demonstrates that the CEF can be treated as a unified analysis tool to show the convergence of any ILC algorithm. As pointed out in [5], as CEFs closely link to Lyapunov functions or energy functions used in time-domain, some analysis tools used in time-domain, such as passivity, input-to-state stability and so on [19] might be easily used in CEF, providing more design freedom for ILC algorithms. It will be an interesting research direction for the future work to investigate on how to balance two design methods: the simplicity from CM and the generality from CEF.

II. PRELIMINARIES

The notations $\mathcal{R}$ and $\mathcal{N}$ represent the set of real numbers and natural numbers respectively. For a given positive number ϵ, the order of magnitude in terms of ϵ is represented by $\mathcal{O}(\epsilon)$. For any vector $\mathbf{x} \in \mathcal{R}^n$, the Euclidean norm, $|\mathbf{x}|$ is defined as $|\mathbf{x}| \triangleq \sqrt{\mathbf{x}^\intercal \mathbf{x}}$. For a given matrix $A \in \mathcal{R}^{n\times n}$, $|A|$ represents the induced matrix norm. For a square matrix A, $\lambda_{min}(A)$ and $\lambda_{max}(A)$ denotes the minimum and maximum eigenvalue of A. The matrix I_n denotes the identity matrix of dimension n. The symbol $i \in \mathcal{N}$ denotes the iteration number and the subscript $(\cdot)_i$ indicates the signal at the i^{th} iteration. For any $j \in \mathcal{N}$, the set of all continuous functions in the interval $[0,\ T_f]$, that is differentiable upto j^{th} order is represented by $\mathcal{C}^j[0,\ T_f]$.

The following definitions of functional norms are used in this paper .

Melbourne School of Engineering, The University of Melbourne, Parkville VIC 3010 Australia, E-mail: gsebastian @student.unimelb.edu.au, {yingt, doetomo}@unimelb.edu.au

TABLE I

SUMMARY OF DIFFERENCES BETWEEN CM-BASED AND CEF-BASED METHODS

	CM-based Methods	**CEF-based Methods**
1	Information in iteration domain is needed. Information in finite time interval is always ignored.	Information in finite-time domain and iteration domain can both appear in the CEF.
2	Need GLC condition to ensure the boundedness of state trajectories.	GLC condition can be relaxed.
3	Uniform convergence is achieved	Pointwise convergence is generally achieved.
4	The knowledge of relative degree of plant is required	CEF can relax the requirement of the relative degree.
5	Always yields a convergence condition to ensure contraction. Hence simple and easy to design.	Do not always provide a convergence condition. Selecting an appropriate CEF is the-state-of-the-art.
6	Commonly deals with output tracking problem	Handles both state tracking and output tracking problem. This can be used for systems with both parametric, norm bounded uncertainties and input, state and output constraints.

Definition 1: For any $\mathbf{x}(\cdot) \in \mathcal{R}^n$ which is defined in $\mathcal{C}[0,\ T_f]$, the supremum norm, $\|\mathbf{x}\|_s : [0, T_f] \times \mathcal{R}^n \to \mathcal{R}_{\geq 0}$ is defined as $\|\mathbf{x}\|_s \triangleq \max_{t\in[0,\, T_f]} |\mathbf{x}(t)|_\infty$, and for any positive constant λ, the λ - norm, $\|\mathbf{x}\|_\lambda : [0, T_f] \times \mathcal{R}^n \to \mathcal{R}_{\geq 0}$ is defined as $\|\mathbf{x}\|_\lambda \triangleq \max_{t\in[0,\, T_f]} e^{-\lambda t} |\mathbf{x}(t)|_\infty$, where $|\mathbf{x}|_\infty = \max_{j\in[1,\ 2,\ \cdots,\ n]} |x^j|$. ◦

Remark 1: It is possible to show that $\|\mathbf{x}\|_\lambda \leq \|\mathbf{x}\|_s \leq e^{\lambda T_f} \|\mathbf{x}\|_\lambda$. This implies that the supremum norm and the λ - norm are equivalent [5, Chapter 2]. These two norms can be used to show the "uniform convergence" of the tracking error in the finite time interval $[0,\ T_f]$. ◦

Definition 2: For any $\mathbf{x}(\cdot) \in \mathcal{R}^n$ which is defined in $\mathcal{C}[0,\ T_f]$, the $\mathcal{L}^2$ norm is defined as $\|\mathbf{x}\|_{\mathcal{L}^2} \triangleq \left(\int_0^{T_f} |\mathbf{x}(\tau)|^2 d\tau\right)^{\frac{1}{2}}$ and for any given positive constant λ, the $\mathcal{L}_e^2$ norm is defined as $\|\mathbf{x}\|_{\mathcal{L}_e^2} \triangleq \left(\int_0^{T_f} e^{-\lambda\tau} |\mathbf{x}(\tau)|^2 d\tau\right)^{\frac{1}{2}}$. ◦

Remark 2: For any $\mathbf{x} \in \mathcal{C}[0,\ T_f]$, its $\mathcal{L}^2$ norm and $\mathcal{L}_e^2$ norm are equivalent. So the convergence in terms of $\mathcal{L}_e^2$ norm indicates the convergence of $\mathcal{L}^2$. ◦

III. PROBLEM FORMULATION

Consider an output tracking problem for a nonlinear multiple-input-multiple-output (MIMO) square[1] affine system, defined for any iteration i:

$$\begin{array}{rcl} \dot{\mathbf{x}}_i &=& \mathbf{f}(\mathbf{x}_i) + G(\mathbf{x}_i)\mathbf{u}_i \\ \mathbf{y}_i &=& \mathbf{h}(\mathbf{x}_i) + D(\mathbf{x}_i)\mathbf{u}_i \end{array} \quad \text{with} \quad \mathbf{x}(0) = \mathbf{x}^0, \tag{1}$$

where $\mathbf{x}_i \in \mathcal{R}^n$, $\mathbf{u}_i \in \mathcal{R}^m$ and $\mathbf{y}_i \in \mathcal{R}^m$ are the state, input and output vector respectively. It is assumed that the nonlinear mappings $\mathbf{f}(\cdot) \in \mathcal{R}^n, G(\cdot) \in \mathcal{R}^{n\times m}$, $\mathbf{h}(\cdot) \in \mathcal{R}^m$ and $D(\cdot) \in \mathcal{R}^{m\times m}$ are globally Lipschitz continuous in its arguments. The matrix $D(\cdot)$ is the direct transmission matrix. The system (1) performs repetitive tracking over a finite time interval $t \in [0,\ T_f]$ where $0 < T_f < \infty$. The following assumptions are needed in a standard ILC design.

Assumption 1: There are positive constants c_f, c_g, c_h and c_d such that for any $\mathbf{z}_1,\ \mathbf{z}_2 \in \mathcal{R}^n$ and $\mathbf{u} \in \mathcal{R}^m$, the following inequalities hold:

$$\begin{array}{rcl} |\mathbf{f}(\mathbf{z}_1) - \mathbf{f}(\mathbf{z}_2)| &\leq& c_f\, |\mathbf{z}_1 - \mathbf{z}_2|\,, \\ |G(\mathbf{z}_1)\mathbf{u} - G(\mathbf{z}_2)\mathbf{u}| &\leq& c_g\, |\mathbf{z}_1 - \mathbf{z}_2|\,, \\ |\mathbf{h}(\mathbf{z}_1) - \mathbf{h}(\mathbf{z}_2)| &\leq& c_h\, |\mathbf{z}_1 - \mathbf{z}_2|\,, \\ |D(\mathbf{z}_1)\mathbf{u} - D(\mathbf{z}_2)\mathbf{u}| &\leq& c_d\, |\mathbf{z}_1 - \mathbf{z}_2|\,. \end{array} \tag{2}$$

□

Remark 3: This is one of the fundamental assumptions for CM-based ILC design [5]. It is possible to relax this assumption as discussed in [4], [20]. ◦

Assumption 2: The direct transmission term $D(\mathbf{x})$ is full rank and the matrix $G(\mathbf{x})$ is full column rank and both matrices are bounded for all $\mathbf{x} \in \mathcal{R}^n$. □

For a given reference input $\mathbf{y}_r(t) \in \mathcal{C}[0,\ T_f]$, the output tracking error is defined as

$$\mathbf{e}(t) \triangleq \mathbf{y}_r(t) - \mathbf{y}(t). \tag{3}$$

Note that the subscript r is used with control signals to indicate that it is related to the reference signals and should not be confused with subscript i that indicate iteration varying system variables. Note that the reference signal is invariant in the iteration-domain.

The simplest first order ILC algorithm for (1) takes the form:

$$\boxed{\mathbf{u}_{i+1}(t) = \mathbf{u}_i(t) + \Gamma(t)\mathbf{e}_i(t),} \quad i = 1, 2, ...,\ \mathbf{u}_1(t) = 0 \tag{4}$$

where $\Gamma(t) \in \mathcal{R}^{m\times m}$ is a learning gain matrix defined for all $t \in [0, T_f]$. Without any loss of generality, the control input is initialized at zero. i.e. at the first iteration $\mathbf{u}_1(t) = 0$.

Let $\mathbf{u}_r \in \mathcal{C}[0,\ T_f]$ be the desired reference input, which is unknown. The "input error" is defined as

$$\delta\mathbf{u}(t) \triangleq \mathbf{u}_r(t) - \mathbf{u}(t). \tag{5}$$

The ILC law (4) for the system (1) can either obtain an *output convergence*, $\mathbf{e}_i \to \mathbf{0}$ or an *input convergence*, $\delta\mathbf{u}_i \to \mathbf{0}$ based on two different *resetting conditions*. These two types of resetting conditions and the corresponding convergence analysis are discussed in the following subsections.

Remark 4: Note that the transmission matrix, $D(\mathbf{x})$ is a square matrix and the system (1) has equal number of inputs and outputs. In such cases there is no basic difference between

[1] A square system has same the dimension for input and output vectors.

the convergence analysis in terms of *output error*, $\mathbf{e}_i$ and *input error* $\delta\mathbf{u}_i$. Differences occur if the dynamic system is a non-square system. If $D(\mathbf{x})$ is full column rank then convergence of the input sequence, $\{\delta\mathbf{u}_i\}$ can be performed. If $D(\mathbf{x})$ is full row rank , then convergence of the output sequence, $\{\mathbf{e}_i\}$ can be performed. ◦

IV. CM-BASED ANALYSIS: A REVISIT

The design and analysis of a standard P-type ILC using CM methods and its essential features for an output tracking problem are reviewed in this section.

A. Output Convergence

The convergence in terms of output error is examined in this subsection, for system (1) satisfying the following resetting condition.

Assumption 3: The system (1) satisfies the following identical initial condition : $(i.i.c: \ 1) \quad \mathbf{x}_i(0) = \mathbf{x}^0, \ \forall i \in \mathcal{N}$. □

Remark 5: Assumption 3 indicates that the dynamic process is initiated from the same initial state $\mathbf{x}^0$ for all iterations. This assumption is needed to show the convergence of the tracking error [5, Chapter 2]. ◦

Theorem 1 achieves uniform convergence of *output error* sequence using CM methods.

Theorem 1: The system (1) with control law (4) under the Assumptions 1, 2 and 3 achieves uniform convergence of the output $\mathbf{y}_i$ to the reference output $\mathbf{y}_r$, if the convergence condition:

$$\max_{\mathbf{x}\in\mathcal{R}^n,\ t\in[0,T_f]} |I_m - D(\mathbf{x})\Gamma(t)|_\infty \leq \rho_1$$

is satisfied for some $0 < \rho_1 < 1$. ■

Proof: The sketch of the proof of the Theorem 1 using CM-based method is summarized as follows

1) Constructing the contraction using input-output relationship using the relative degree, leading to the following relationship:

$$\begin{aligned}\mathbf{e}_{i+1} = &(I_m - D(\mathbf{x}_{i+1})\Gamma)\,\mathbf{e}_i - (\mathbf{h}(\mathbf{x}_{i+1}) - \mathbf{h}(\mathbf{x}_i)) \\ &- (D(\mathbf{x}_{i+1})\mathbf{u}_i - D(\mathbf{x}_i)\mathbf{u}_i)\end{aligned} \tag{6}$$

2) Ignoring dynamics of the state by using Gronwell lemma and the time-weighted norm (λ - norm), the dynamics of the state satisfy the following form:

$$\|\Delta\mathbf{x}_{i+1}\|_\lambda \leq \mathcal{O}_1(\lambda^{-1})\,\|\mathbf{e}_i\|_\lambda \tag{7}$$

where $\Delta\mathbf{x}_{i+1} = \mathbf{x}_{i+1} - \mathbf{x}_i$.

3) Showing the contraction and prove the convergence by selecting a sufficiently large λ.

$$\begin{aligned}\|\mathbf{e}_{i+1}\|_\lambda &\leq \|I_m - D\Gamma\|_s\,\|\mathbf{e}_i\|_\lambda + c_{hd}\,\|\delta\mathbf{x}_i\|_\lambda \\ &\leq \bar{\rho}_1\,\|\mathbf{e}_i\|_\lambda\end{aligned} \tag{8}$$

where $\bar{\rho}_1 = \rho_1 + c_{hd}\mathcal{O}_1(\lambda^{-1})$ and $c_{hd} = c_h + c_d$.

■

Next, the convergence of *input error* sequence,$\delta\mathbf{u}_i$, based on slightly different assumptions is presented.

B. Input Convergence

The convergence in terms of input error $\delta\mathbf{u}_i$ is revisited in this subsection. Assumption on existence of a reference dynamics and a new resetting condition are required for ensuring convergence in terms of control input.

Assumption 4: For any given reference output $\mathbf{y}_r \in \mathcal{C}[0,\ T_f]$, there exist a reference state $\mathbf{x}_r \in \mathcal{C}^1[0,\ T_f]$ and a reference input $\mathbf{u}_r \in \mathcal{C}[0,\ T_f]$ that satisfy the system dynamics:

$$\begin{aligned}\dot{\mathbf{x}}_r &= \mathbf{f}(\mathbf{x}_r) + G(\mathbf{x}_r)\mathbf{u}_r \\ \mathbf{y}_r &= \mathbf{h}(\mathbf{x}_r) + D(\mathbf{x}_r)\mathbf{u}_r,\end{aligned} \tag{9}$$

□

Remark 6: Assumption 4 is called matching condition, which ensures the existence and uniqueness of the desired control input. If there are two control input signals that can generate the same reference output, it is hard to ensure the convergence of the input when output converges. This condition plays an important role when dealing with input saturation [21] ◦

Assumption 5: The system (1) satisfies the following identical initial condition : $(i.i.c: \ 2) \quad \mathbf{x}_i(0) = \mathbf{x}_r(0), \ \forall i \in \mathcal{N}$. □

Remark 7: Assumption 5 indicates that at every iteration the initial state of system is same as that of reference state. This is stronger than Assumption 3 which was used for output convergence. ◦

Theorem 2 achieves uniform convergence of *input error* sequence using CM methods.

Theorem 2: The system (1) with control law (4) under the Assumptions 1, 2, 4 and 5 achieves uniform convergence of the sequence of control input $\{\mathbf{u}_i\}$ to the desired control input $\mathbf{u}_r$, if the convergence condition:

$$\max_{\mathbf{x}\in\mathcal{R}^n,\ t\in[0,T_f]} |I_m - \Gamma(t)D(\mathbf{x})|_\infty \leq \rho_2 < 1$$

is satisfied for some positive ρ_2. ■

Proof: The sketch of proof is similar to that of Theorem 1.

1) Constructing the contraction using input-output relationship from relative degree will lead to the following relationship:

$$\begin{aligned}\delta\mathbf{u}_{i+1} = &(I_m - \Gamma D(\mathbf{x}_i))\,\delta\mathbf{u}_i - \Gamma\,(\mathbf{h}(\mathbf{x}_r) - \mathbf{h}(\mathbf{x}_i)) \\ &- \Gamma\,(D(\mathbf{x}_r) - D(\mathbf{x}_i))\,\mathbf{u}_r\end{aligned} \tag{10}$$

2) Ignoring dynamics of the state by using Gronwell lemma and the time-weighted norm, the dynamics of the state satisfy the following form:

$$\|\delta\mathbf{x}_i\|_\lambda \leq \mathcal{O}_2(\lambda^{-1})\,\|\mathbf{u}_i\|_\lambda \tag{11}$$

where $\delta\mathbf{x}_i = \mathbf{x}_r - \mathbf{x}_i$.

3) Showing the contraction and prove the convergence by selecting a sufficiently large λ.

$$\begin{aligned}\|\mathbf{u}_{i+1}\|_\lambda &\leq \left(\|I_m - \Gamma D\|_s + |\Gamma|\,c_{hd}\mathcal{O}_1(\lambda^{-1})\right)\|\mathbf{e}_i\|_\lambda \\ &\leq \rho_1\,\|\mathbf{e}_i\|_\lambda\end{aligned} \tag{12}$$

where $\bar{\rho}_2 = \rho_2 + |\Gamma|\,c_{hd}\mathcal{O}_2\left(\lambda^{-1}\right) < 1$.

This completes the proof. ■

C. Discussions

The results of Theorem 1 and Theorem 2 are well-known [5]. The following properties of CM-based analysis can be summarized.

1) The convergence is independent of the system mappings $\mathbf{f}(\cdot)$, $G(\cdot)$ and $\mathbf{h}(\cdot)$. It is only depended upon

the choice of Γ that satisfies the convergence condition: $|I_m - D\Gamma|_\infty \leq \rho_1 < 1$ or $|I_m - \Gamma D|_\infty \leq \rho_2 < 1$ for convergence in terms of output or input respectively. Hence the accurate information of transmission matrix,$D(\cdot)$ is not needed for achieving a perfect tracking performance [22].

2) The input-to-output mapping plays an important role (with the concept of relative degree). The dynamics of the state can be ignored by using Gronwell Lemma and the time-weighted norm.
3) GLC condition plays an important role in using Gronwall lemma. Most of the engineered systems do not satisfy this assumption. CM method can not be directly applied if the dynamic system is locally Lipschitz continuous and exhibits some nonlinear phenomenon such as finite escape.
4) Similar results can be obtained when the relative degree of the system is more than zero.

V. CEF-BASED ANALYSIS

Different from CM-based analysis, the CEF-based method has been shown to be a useful analysis tool when the dynamic system does not satisfy GLC. It can also relax the requirement for the relative degree. The concept of CEF is similar to well-known concept of Lyapunov function. For a given CEF, the proposed ILC algorithm has to show that at each time instant, the CEF decreases over iteration, achieving a point-wise convergence. Usually, in the CEF-based analysis, there is no need to have a convergence condition. To the best of authors' knowledge, there is no proof of convergence of P-type ILC (4) using the CEF-based analysis tool.

Next, it will show that how the CEF-based method can handle the convergence condition for the given P-type ILC (4) to obtain the same convergence properties.

It is worthwhile to highlight that re-proving the same result using CEF is not the goal of this paper. The focus here is to show that the CEF-based analysis tool can be served as a unified tool in the design and analysis of an ILC algorithm for various engineering systems. As the complexity of engineering systems increases, a set of tool-kit is needed to provide flexibility and design freedom.

A. Output Convergence

This subsection will re-prove Theorem 1 using a novel CEF. ***Proof***: A novel CEF is proposed as $(\underline{c.e.f-1})$, denoted by $E_i(t)$, with a positive constant λ, $\forall t \in [0, T_f],\ i \in \mathcal{N}$:

$$(\underline{c.e.f-1}): \quad E_i \;=\; \frac{1}{2}e^{-\lambda t}\Delta\mathbf{x}_i^{\intercal}\Delta\mathbf{x}_i + \int_0^t e^{-\lambda\tau}\mathbf{e}_i{}^{\intercal}\mathbf{e}_i d\tau, \tag{13}$$

which is initialized with $\mathbf{x}_0(t) = 0$. Therefore at $i = 1$, $\Delta\mathbf{x}_1 = \mathbf{x}_1$. The CEF, *c.e.f-1*: (13) at the current iteration is the sum of quadratic term of the difference in state variable between the current and the previous iteration (i.e. $\Delta\mathbf{x}_i$) and $\mathcal{L}^2$ norm equivalent of output tracking error at the current iteration which is defined for all $t \in [0,\ T_f]$.

The idea behind the proof is to show that the energy function is non-increasing along the iteration-domain and is bounded. The leads to the convergence of the output error sequence.

1) Non-increasing Energy Function: The difference of CEF, *c.e.f-1*: (13), between two iterations $\Delta E_{i+1} = E_{i+1} - E_i$, i,e.

$$\Delta E_{i+1} = \frac{1}{2}e^{-\lambda t}\left|\Delta\mathbf{x}_{i+1}\right|^2 + \int_0^t e^{-\lambda\tau}\left(\left|\mathbf{e}_{i+1}\right|^2 - \left|\mathbf{e}_i\right|^2\right)d\tau - \frac{1}{2}e^{-\lambda t}\left|\Delta\mathbf{x}_i\right|^2 \tag{14}$$

Using system dynamics (1), it is possible to show that

$$\Delta\dot{\mathbf{x}}_{i+1} = \mathbf{f}(\mathbf{x}_{i+1}) - \mathbf{f}(\mathbf{x}_i) + \left(G(\mathbf{x}_{i+1}) - G(\mathbf{x}_i)\right)\mathbf{u}_i + G(\mathbf{x}_{i+1})\Gamma\mathbf{e}_i \tag{15}$$

Using the resetting condition $i.i.c:1$ from Assumption 3, Assumption 1 and equation (15), the first term in equation (14) can be expanded as:

$$\begin{aligned}
&\frac{1}{2}e^{-\lambda t}\left|\Delta\mathbf{x}_{i+1}\right|^2 \\
&= -\frac{\lambda}{2}\int_0^t e^{-\lambda\tau}\left|\Delta\mathbf{x}_{i+1}\right|^2 d\tau + \int_0^t e^{-\lambda\tau}\Delta\mathbf{x}_{i+1}^{\intercal}\Delta\dot{\mathbf{x}}_{i+1}d\tau \\
&= -\frac{\lambda}{2}\int_0^t e^{-\lambda\tau}\left|\Delta\mathbf{x}_{i+1}\right|^2 d\tau \\
&\quad + \int_0^t e^{-\lambda\tau}\Delta\mathbf{x}_{i+1}^{\intercal}\left(\mathbf{f}(\mathbf{x}_{i+1}) - \mathbf{f}(\mathbf{x}_i)\right)d\tau \\
&\quad + \int_0^t e^{-\lambda\tau}\Delta\mathbf{x}_{i+1}^{\intercal}(G(\mathbf{x}_{i+1})\mathbf{u}_i - G(\mathbf{x}_i)\mathbf{u}_i)d\tau \\
&\quad + \int_0^t e^{-\lambda\tau}\Delta\mathbf{x}_{i+1}^{\intercal}G(\mathbf{x}_{i+1})\Gamma\mathbf{e}_i d\tau \\
&\leq -\left(\frac{\lambda}{2} - c_{fg}\right)\int_0^t e^{-\lambda\tau}\left|\Delta\mathbf{x}_{i+1}\right|^2 d\tau \\
&\quad + g_r\int_0^t e^{-\lambda\tau}\left|\Delta\mathbf{x}_{i+1}\right|\left|\mathbf{e}_i\right| d\tau.
\end{aligned} \tag{16}$$

where $g_r \triangleq \max_{\mathbf{x}\in\mathcal{R}^n, t\in[0,T_f]}|G(\mathbf{x})\Gamma(t)|$ and $c_{fg} = c_f + c_g$. Taking the squared norm of (6), the following relation can be obtained:

$$\left|\mathbf{e}_{i+1}\right|^2 \leq \left|\left(I_m - D(\mathbf{x}_{i+1})\Gamma\right)\mathbf{e}_i\right|^2 + c_{hd}^2\left|\Delta\mathbf{x}_{i+1}\right|^2 + 2c_{hd}\left|\left(I_m - D(\mathbf{x}_{i+1})\Gamma\right)\mathbf{e}_i\right|\left|\Delta\mathbf{x}_{i+1}\right| \tag{17}$$

For the ease of notation, let $\mathrm{P}_1(\mathbf{x}_{i+1}) \triangleq (I_m - D(\mathbf{x}_{i+1})\Gamma)$, $w_1 \triangleq c_{hd}^2$, $w_2 \triangleq 2c_{hd}\max_{\mathbf{x}\in\mathcal{R}^n,\ t\in[0,T_f]}|(I_m - D(\mathbf{x}_{i+1})\Gamma)|$.

Therefore,

$$\left|\mathbf{e}_{i+1}\right|^2 - \left|\mathbf{e}_i\right|^2 \leq -\left(I_m - \left|\mathrm{P}_1(\mathbf{x}_{i+1})\right|^2\right)\left|\mathbf{e}_i\right|^2 + w_1\left|\Delta\mathbf{x}_{i+1}\right|^2 + w_2\left|\mathbf{e}_i\right|\left|\Delta\mathbf{x}_{i+1}\right| \tag{18}$$

If Γ is selected in such a way that:

$$\max_{\mathbf{x}\in\mathcal{R}^n,\ t\in[0,T_f]}\left|I_m - D(\mathbf{x})\Gamma(t)\right|_\infty \leq \rho_1 < 1 \tag{19}$$

is satisfied for some positive ρ_1. This also means that $\lambda_{min}\left(I_m - \mathrm{P}_1^{\intercal}\mathrm{P}_1\right) \leq \rho_1 < 1$ for all $\mathbf{x} \in \mathcal{R}^n$ and $t \in [0, T_f]$.

Using Young's inequality, there exists an $\alpha > 0$ such that:

$$\begin{aligned}
\left|\Delta\mathbf{x}_{i+1}\right|\left|\mathbf{e}_i\right| &= \sqrt{\alpha}\left|\Delta\mathbf{x}_{i+1}\right|\frac{1}{\sqrt{\alpha}}\left|\mathbf{e}_i\right| \\
&\leq \frac{\alpha}{2}\left|\Delta\mathbf{x}_{i+1}\right|^2 + \frac{1}{2\alpha}\left|\mathbf{e}_i\right|^2.
\end{aligned} \tag{20}$$

Finally, substituting equations (16), (18) and (20) back into equation (14) results in

$$\Delta E_{i+1} \leq -\left(\frac{\lambda}{2} - \mathcal{O}_3(\alpha)\right)\int_0^t e^{-\lambda\tau}\left|\Delta \mathbf{x}_{i+1}\right|^2 d\tau - \left(\rho_1 - \mathcal{O}_4(\alpha^{-1})\right)\int_0^t e^{-\lambda\tau}\left|\mathbf{e}_i\right|^2 d\tau - \frac{1}{2}e^{-\lambda t}\left|\Delta \mathbf{x}_i\right|^2 \tag{21}$$

where $\mathcal{O}_3(\alpha) \triangleq c_{fg} + w_1 + \frac{\alpha}{2}(w_2 + g_r)$, $\mathcal{O}_4(\alpha^{-1}) \triangleq \frac{1}{2\alpha}(g_r + w_2)$. For a given ρ_1, there exists a sufficiently large α such that $\rho_1 > \mathcal{O}_4(\alpha^{-1})$. For that α, there exists a $\lambda > 2\mathcal{O}_3(\alpha)$, such that difference of CEF satisfy: $\Delta E_{i+1}(t) \leq 0$, for $t \in [0,\ T_f]$ i.e., composite energy is non-increasing in the iteration-domain.

2) Convergence Property: Firstly *c.e.f-1*: (13), at k^{th} iteration can be written as $E_k(t) = E_1(t) + \sum_{j=2}^{k}\Delta E_k(t)$. This means that $\lim_{k\to\infty} E_k(t)$ exists and

$$\lim_{k\to\infty} E_k(t) \ = \ E_1(t) + \lim_{k\to\infty}\sum_{j=2}^{j}\Delta E_k(t) \ \leq \ E_1(t). \tag{22}$$

as $\Delta E_k(t) \leq 0$ for $k \in \mathcal{N}$, $t \in [0, T_f]$. The series $\{E_k(t)\}_{k\in\mathcal{N}}$ is bounded, if $E_1(t)$ is finite.

From $(\underline{c.e.f-1})$: (13), $E_1(t) = \frac{1}{2}e^{-\lambda t}\left|\mathbf{x}_1\right|^2 + \int_0^t e^{-\lambda\tau}|\mathbf{e}_1|^2 d\tau$ as $\Delta\mathbf{x}_1 = \mathbf{x}_1$. Also the control input is initialized at $\mathbf{u}_1 = 0$. From the system dynamics (1), this leads to $\dot{\mathbf{x}}_1 = \mathbf{f}(\mathbf{x}_1)$ and $\mathbf{y}_1 = \mathbf{h}(\mathbf{x}_1)$. Due to Assumption 1, by invoking the existence and uniqueness theorem [19, Theorem 2.2], there exists a unique solution to $\dot{\mathbf{x}}_1 = \mathbf{f}(\mathbf{x}_1); \mathbf{y}_1 = \mathbf{h}(\mathbf{x}_1)$ with $\mathbf{x}(0) = \mathbf{x}^0$ over a finite interval $t \in [0, T_f]$ and there exists an upper-bound for all the trajectories in the finite time interval. This establishes the uniform boundedness of $\mathbf{x}_1$, $\mathbf{e}_1$ and E_1 in $t \in [0,\ T_f]$.

In addition, $\sum_{j=2}^{k}\Delta E_k(t)$ converges. From the convergence theorem [23], as the sum of the series converges to zero, the series converges, leading to pointwise convergence of $\{E_k(t)\}$, i.e $\lim_{k\to\infty}\Delta E_k(t) = 0, \forall t \in [0,\ T_f]$. Specifically, there exists positive constants $L_\lambda > \frac{\lambda}{2} - \mathcal{O}_3(\alpha) > 0$ and $L_\rho > \rho_1 - \mathcal{O}_4(\alpha^{-1}) > 0$ such that, when $t = T_f$, (22) yields

$$\lim_{i\to\infty}\int_0^{T_f} e^{-\lambda\tau}\left(L_\lambda\left|\Delta\mathbf{x}_i\right|^2 + L_\rho\left|\mathbf{e}_i\right|^2\right)d\tau = 0,$$
$$\lim_{i\to\infty}\|\Delta\mathbf{x}_{i+1}\|_{\mathcal{L}_e^2} = 0 \text{ and } \lim_{i\to\infty}\|\mathbf{e}_i\|_{\mathcal{L}_e^2} = 0 \tag{23}$$

Therefore the sequence $\{\mathbf{e}_i\}_{i\in\mathcal{N}}$ converges to zero in $\mathcal{L}^2[0,\ T_f]$ norm.

Using (17) and Young's inequality relation, there exists some $0 < \alpha_s < 1$, which satisfies $\rho_s = \rho_1 + \alpha_s < 1$ such that

$$\left|\mathbf{e}_{i+1}\right|^2 \leq \rho_s\left|\mathbf{e}_i\right|^2 + \gamma(\alpha_s)\left|\Delta\mathbf{x}_{i+1}\right|^2 \tag{24}$$

for some class-$\mathcal{K}$ function γ. This will ensure that the tracking error will be uniformly bounded over iteration, leading to the uniform convergence of the tracking error. This completes the proof. ■

Remark 8: The construction of CEF is the state-of-the art as finding the appropriate Lyapunov candidate for a dynamic system. Another possibility in CEF is to use $\Delta\mathbf{u}_{i+1} = \mathbf{u}_{i+1} - \mathbf{u}_i$ instead of $\mathbf{e}_i$. Under such a situation, $\int_0^t e^{-\lambda\tau}\Delta\mathbf{u}_{i+1}{}^\intercal\Delta\mathbf{u}_i d\tau$ can be used as the second term in $(\underline{c.e.f-1})$: (13) instead $\int_0^t e^{-\lambda\tau}\mathbf{e}_i{}^\intercal\mathbf{e}_i d\tau$. In a simple first order feed-forward ILC, there is no difference between these two terms. If a higher-order ILC [24] is taken into consideration, then $\Delta\mathbf{u}_i$ is a suitable candidate in the energy function. ○

B. Input Convergence

Similar to Section V-A, Theorem 2 will be re-proved in this section using CEF-based techniques.

Proof: For the proof of convergence of input, a new CEF, $(\underline{c.e.f-2})$, E_i is introduced in this section. For any given $\lambda > 0$, $\forall t \in [0, T_f]$, $i \in \mathcal{N}$, E_i is given by

$$(\underline{c.e.f-2}):\ E_i = \frac{1}{2}e^{-\lambda t}\delta\mathbf{x}_{i-1}^\intercal\delta\mathbf{x}_{i-1} + \int_0^t e^{-\lambda\tau}\delta\mathbf{u}_i^\intercal\delta\mathbf{u}_i d\tau, \tag{25}$$

which is initialized at $\mathbf{x}_0(t) = 0$. Therefore $\delta\mathbf{x}_0 = \mathbf{x}_r$. Note the CEF *c.e.f-2*: (25) at current iteration has a quadratic term of state tracking error from previous iteration.

1) Non-increasing Energy Function: The difference of CEF, *c.e.f-2*: (25),

$$\Delta E_{i+1} = \frac{1}{2}e^{-\lambda t}\left|\delta\mathbf{x}_i\right|^2 + \int_0^t e^{-\lambda\tau}\left(\left|\delta\mathbf{u}_{i+1}\right|^2 - \left|\delta\mathbf{u}_i\right|^2\right)d\tau - \frac{1}{2}e^{-\lambda t}\left|\delta\mathbf{x}_{i-1}\right|^2, \tag{26}$$

It is possible to show that

$$\delta\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}_r) - \mathbf{f}(\mathbf{x}) + (G(\mathbf{x}_r) - G(\mathbf{x}))\mathbf{u}_r + G(\mathbf{x})\delta\mathbf{u} \tag{27}$$

Using Assumption 5 and (27), the first term in *c.e.f-2*: (26) can be written as:

$$\frac{1}{2}e^{-\lambda t}\left|\delta\mathbf{x}_i\right|^2 = -\frac{\lambda}{2}\int_0^t e^{-\lambda\tau}\left|\delta\mathbf{x}_i\right|^2 d\tau + \int_0^t e^{-\lambda\tau}\delta\mathbf{x}_i^\intercal\delta\dot{\mathbf{x}}_i d\tau \leq -\left(\frac{\lambda}{2} - c_{fg}\right)\int_0^t e^{-\lambda\tau}\left|\delta\mathbf{x}_i\right|^2 d\tau + g_0\int_0^t e^{-\lambda\tau}\left|\delta\mathbf{x}_i\right|\left|\delta\mathbf{u}_i\right| d\tau. \tag{28}$$

where $g_0 \triangleq \max_{\mathbf{x}\in\mathcal{R}^n}|G(\mathbf{x})|$. For the ease of notation, let $w_4 \triangleq 2r_c\max_{\mathbf{x}\in\mathcal{R}^n,\ t\in[0,T_f]}|(I_m - \Gamma D(\mathbf{x}_i))|$, $\mathrm{P}_2(\mathbf{x}_i) \triangleq (I_m - \Gamma D(\mathbf{x}_i))$, $w_3 \triangleq r_c^2$, $r_c = |\Gamma|\, c_{hd}$.

Taking the squared norm of (10) and substituting in the expression: $\left|\delta\mathbf{u}_{i+1}\right|^2 - \left|\delta\mathbf{u}_i\right|^2$, results in the following relation:

$$\left|\delta\mathbf{u}_{i+1}\right|^2 - \left|\delta\mathbf{u}_i\right|^2 \leq -\rho_2\left|\delta\mathbf{u}_i\right|^2 + w_3\left|\delta\mathbf{x}_i\right|^2 + w_4\left|\delta\mathbf{x}_i\right|\left|\delta\mathbf{u}_i\right| \tag{29}$$

where ρ_2 satisfies $\lambda_{min}\left(I_m - \mathrm{P}_2^\intercal\mathrm{P}_2\right) \leq \rho_2 < 1$ for all $\mathbf{x} \in \mathcal{R}^n$ and $t \in [0, T_f]$.

Using Young's inequality relation, there exists an $\alpha > 0$ such that:

$$\left|\delta\mathbf{x}_i\right|\left|\delta\mathbf{u}_i\right| \leq \frac{\alpha}{2}\left|\delta\mathbf{x}_i\right|^2 + \frac{1}{2\alpha}\left|\delta\mathbf{u}_i\right|^2. \tag{30}$$

Finally, substituting equations (28), (29) and (30) back into equation (26) results in

$$\Delta E_{i+1} \leq -\left(\frac{\lambda}{2} - \mathcal{O}_5(\alpha)\right)\int_0^t e^{-\lambda\tau}\left|\delta\mathbf{x}_i\right|^2 d\tau - \left(\rho_2 - \mathcal{O}_6(\alpha^{-1})\right)\int_0^t e^{-\lambda\tau}\left|\delta\mathbf{u}_i\right|^2 d\tau - \frac{1}{2}e^{-\lambda t}\left|\delta\mathbf{x}_{i-1}\right|^2 \tag{31}$$

where $\mathcal{O}_5(\alpha) \triangleq c_{fg} + w_3 + \frac{\alpha}{2}(w_4 + g_0)$ and $\mathcal{O}_6(\alpha^{-1}) \triangleq \frac{1}{2\alpha}(g_0 + w_4)$. Similar to the previous case, we can show that $\Delta E_{i+1}(t) \leq 0$ for a suitable choice of λ and α.

Convergence Property: Similar to the analysis and conclusions in convergence property from Section V-A, the pointwise convergence of the sequence can be obtained, followed by the uniform convergence of control input.

Using the same inequality as in (22), the sequence of energy function $(c.e.f - 2)$, $\{E_k(t)\}_{k\in\mathcal{N}}$ is point-wise bounded, if $E_1(t)$ is finite. For this CEF $(c.e.f - 2)$, $\mathbf{x}_0 = 0$ and $\mathbf{u}_1 = 0$ by definition. This yields:

$$E_1(t) = \frac{1}{2}e^{-\lambda t}\left|\mathbf{x}_r\right|^2 + \int_0^t e^{-\lambda\tau}\left|\mathbf{u}_r\right|^2 d\tau \qquad (32)$$

As per Assumption 4, $\mathbf{x}_r$ and $\mathbf{u}_r$ are uniformly continuous functions and is bounded. Hence $E_1(t)$ is finite for all $t \in [0,\ T_f]$.

There exists positive constants $N_\lambda > \frac{\lambda}{2} - \mathcal{O}_6(\alpha) > 0$ and $N_\rho > \rho_2 - \mathcal{O}_6(\alpha^{-1}) > 0$ such that, when $t = T_f$:

$$\lim_{i\to\infty}\int_0^{T_f} e^{-\lambda\tau}\left(N_\lambda\left|\delta\mathbf{x}_{i-1}\right|^2 + N_\rho\left|\delta\mathbf{u}_i\right|^2\right)d\tau = 0,$$
$$\lim_{i\to\infty}\|\delta\mathbf{x}_{i-1}\|_{\mathcal{L}_e^2} = 0 \text{ and } \lim_{i\to\infty}\|\delta\mathbf{u}_i\|_{\mathcal{L}_e^2} = 0 \qquad (33)$$

Therefore the sequence $\{\delta\mathbf{u}_i\}_{i\in\mathcal{N}}$ and $\{\delta\mathbf{x}_i\}_{i\in\mathcal{N}}$ converges in $\mathcal{L}^2[0,\ T_f]$. By using Barbalet lemma in (27), the uniform boundedness of $\delta\mathbf{x}_i$ can be ensured. Hence uniform convergence of $\{\delta\mathbf{x}_i\}$ can be ensured.

The proof of the uniform convergence of the *input error* is similar to the proof of uniform convergence of *output error* using CEF. This completes the proof. ■

The following statements can be summarized from the CEF-based methods.

1) In CEF-based analysis, the effect of system dynamics is neglected which is similar to CM-based analysis.
2) CEF-based analysis naturally leads to point-wise convergence as observed as it is shown that CEF decreases at each time instant. Other techniques are needed to show the uniform convergence such as uniform boundedness of the signals, equicontinuity of the signal (see Ascoli's theorem in [13], [25]) or Barbalat's Lemma [4], [18].
3) Convergence condition is used along with Young's inequality in the convergence analysis of CEF. In future work, how to systematically design a suitable CEF for a given class of systems with available knowledge will be exploited.

As CEF can handle a large class of nonlinear systems without GLC condition and less restrictive requirement of the relative degree [4], [5], this work shows that CEF can be used to re-prove the convergence of P-type ILC, in which the ILC law was designed based on CM method. This indicates that CEF can be served as a unified analysis tool to show the convergence of an ILC algorithm.

VI. Conclusion

The well-known P-type ILC algorithms were designed based on contraction mapping method. By re-visiting and re-proving the convergence of P-type ILC algorithm using a novel CEF method, this paper demonstrates that the CEF-based method can be used as a unified tool in the convergence analysis of any ILC algorithm. Future work will explore new design framework that can balance between the simplicity of CM-based method and generality of CEF-based method.

References

[1] K. L. Moore, "Iterative learning control: An expository overview," in *Applied and computational control, signals, and circuits*. Springer, 1999, pp. 151–214.

[2] H.-S. Ahn, Y. 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, no. 6, p. 1099, 2007.

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

[4] J.-X. Xu, "A survey on iterative learning control for nonlinear systems," *International Journal of Control*, vol. 84, no. 7, pp. 1275–1294, 2011.

[5] J.-X. Xu and Y. Tan, *Linear and nonlinear iterative learning control*. Springer, 2003, vol. 291.

[6] S. Arimoto, S. Kawamura, and F. Miyazaki, "Bettering operation of robots by learning," *Journal of Robotic systems*, vol. 1, no. 2, pp. 123–140, 1984.

[7] J. E. Hauser, "Learning control for a class of nonlinear systems," in *Decision and Control, 1987. 26th IEEE Conference on*, vol. 26. IEEE, 1987, pp. 859–860.

[8] G. Heinzinger, D. Fenwick, B. Paden, and F. Miyazaki, "Robust leaning control," in *Decision and Control, 1989., Proceedings of the 28th IEEE Conference on*. IEEE, 1989, pp. 436–440.

[9] T. Sugie and T. Ono, "An iterative learning control law for dynamical systems," *Automatica*, vol. 27, no. 4, pp. 729–732, 1991.

[10] H.-S. Ahn, C.-H. Choi, and K.-b. Kim, "Iterative learning control for a class of nonlinear systems," *Automatica*, vol. 29, no. 6, pp. 1575–1578, 1993.

[11] S. Arimoto, "Learning control theory for robotic motion," *International Journal of Adaptive Control and Signal Processing*, vol. 4, no. 6, pp. 543–564, 1990.

[12] D. H. Owens, "Iterative learning control-convergence using high gain feedback," in *Proceedings of the 1992 IEEE International Symposium on Intelligent Control*. IEEE, 1992, pp. 455–457.

[13] S. S. Saab, "On the p-type learning control," *IEEE Transactions on Automatic Control*, vol. 39, no. 11, pp. 2298–2302, 1994.

[14] A. Hac, "Learning control in the presence of measurement noise," in *American Control Conference, 1990*. IEEE, 1990, pp. 2846–2851.

[15] C.-J. Chien and J.-S. Liu, "A p-type iterative learning controller for robust output tracking of nonlinear time-varying systems," *International Journal of Control*, vol. 64, no. 2, pp. 319–334, 1996.

[16] G. Sebastian, Y. Tan, D. Oetomo, and I. Mareels, "On feedback-based iterative learning control for nonlinear systems without global lipschitz continuity," in *2018 Annual American Control Conference (ACC)*. IEEE, 2018, pp. 5612–5617.

[17] A. Tayebi, "Adaptive iterative learning control for robot manipulators," *Automatica*, vol. 40, no. 7, pp. 1195–1203, 2004.

[18] Y. Tan, H.-H. Dai, D. Huang, and J.-X. Xu, "Unified iterative learning control schemes for nonlinear dynamic systems with nonlinear input uncertainties," *Automatica*, vol. 48, no. 12, pp. 3173–3182, 2012.

[19] H. K. Khalil, *Nonlinear systems*, 2nd ed. Prentice Hall, Upper Saddle River, NJ 07458, 1996.

[20] Y. Tan, S. P. Yang, and J. X. Xu, "On P-type iterative learning control for nonlinear systems without global Lipschitz continuity condition," in *American Control Conference (ACC), 2015*. IEEE, 2015, pp. 3552–3557.

[21] J.-X. Xu, Y. Tan, and T.-H. Lee, "Iterative learning control design based on composite energy function with input saturation," *Automatica*, vol. 40, no. 8, pp. 1371–1377, aug 2004.

[22] K. L. Moore, *Iterative learning control for deterministic systems*. Springer Science & Business Media, 2012.

[23] I. S. Sokolnikoff, R. M. Redheffer, and J. Avents, "Mathematics of physics and modern engineering," *Journal of The Electrochemical Society*, vol. 105, no. 9, pp. 196C—-196C, 1958.

[24] Z. Bien and K. M. Huh, "Higher-order iterative learning control algorithm," in *IEE Proceedings D (Control Theory and Applications)*, vol. 136. IET, 1989, pp. 105–112.

[25] S. Arimoto, *Mathematical Theory of Learning with Applications to Robot Control*. Boston, MA: Springer US, 1986, pp. 379–388.