# A Composite Energy Function Based Sub-optimal Learning Control Approach for Nonlinear Systems with Tima-varying Parametric Uncertainties

Tan Ying and Xu Jian-xin
Department of Electrical and Computer Engineering
National University of Singapore
Singapore 117576

**Abstract**

In this paper, a novel composite energy function (CEF) is introduced to provide a general framework for incorporating system information along both time and learning repetition horizon. Based on the CEF, learning control is integrated with nonlinear sub-optimal control to enhance control performance for a class of nonlinear system with time-varying parametric uncertainties. Sub-optimal control strategy based on control Lyapunov function (CLF) and Sontag's formula provides a sub-optimal performance as well as stability along time horizon for the nominal part of the nonlinear dynamic system. Learning mechanism tries to learn unknown time-varying parametric uncertainties so as to eliminate uncertain effects. The proposed control scheme achieves asymptotical convergence along learning repetition horizon. At the same time, the boundedness and pointwise convergence of the tracking error along time horizon are also ensured.

## 1 Introduction

Nonlinear optimal control has been an active subject of considerable research work over past few decades. Because of difficulties in solving the nonlinear partial Hamilton-Jacobi-Bellman (HJB) differential equation, sub-optimal control and inverse optimal control based on control Lyapunov function (CLF) have been proposed in the design of nonlinear optimal control [1-4]. However it would be difficult to apply these methods in the presence of system uncertainties and disturbances. In this work, learning control is integrated with sub-optimal control and applied to nonlinear systems with time-varying parametric uncertainties without the prior knowledge of the uncertainty bounds. Under repeatable control environment, the time-varying parametric uncertainties are learnable. A repeatable tracking control environment is specified by the following two factors. 1) repeatable tracking control tasks: finite interval but uniform tracking requirement over the entire interval; 2) deterministic dynamic system with the same initial condition, and possibly subject to time-varying parametric uncertainties and interval uncertainties. Repeatable control problems are well encountered both in motion control and process control areas such as wafer process, batch reactor control, IC welding process, industrial robot control on assembly line, etc. The system repeatability warrants the learnability of the system time-varying parametric uncertainties. By virtue of internal model principle, the learning control mechanism can completely nullify the influence of system uncertainties from tracking performance. Hence learning control can be regarded as another kind of optimal control schemes in the sense of canceling the learnable uncertainties existed in nonlinear systems. By integration, the advantages of both optimal control and learning control can be retained – the "sub-optimal" performance along time horizon and "optimal" performance along learning repetition horizon. A number of deterministic learning control schemes have been proposed [5-8]. To lay the foundation for control integration and rigorous analysis, we introduce a novel composite energy function (CEF) into learning control, which consists of two parts. The first part is a control Lyapunov function (CLF) which facilitates the derivation of sub-optimal control law and assesses system behavior along time horizon for each learning cycle. The second part is a $L^2$ norm of parametric learning error which facilitates the derivation of learning control law and evaluates the learning effect.

This paper is organized as follows. In Section 2, the problem formulation is given. A brief introduction of nonlinear optimal control is presented in Section 3. In Section 4, the CEF and sub-optimal learning are defined and discussed in details. Section 5 provides an illustrative example.

## 2 Problem Formulation

Consider a class of MIMO nonlinear dynamic deterministic system with time-varying uncertainties represented by the following equation (throughout this paper, for the sake of brevity, arguments are sometimes omitted when no confusion is likely to arise)

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x},t) + B(\mathbf{x},t)(\mathbf{u}+\mathbf{d}) \\ &= \mathbf{f}(\mathbf{x},t) + B(\mathbf{x},t)[\mathbf{u}+\Theta(t)\boldsymbol{\xi}(\mathbf{x},t)] \;\; \mathbf{f}(\mathbf{0},\cdot)=\mathbf{0}, \quad (1) \end{aligned}$$

where $\mathbf{x} \in R^n$ is the measurable state vector of the system (1); $\mathbf{u} \in R^m$ is the control input vector, where $m < n$; nonlinear function $(\mathbf{x}, t) : R^n \times R_+ \to R^n$ and $B(\mathbf{x}, t) : R^n \times R_+ \to R^{m \times n}$ are known continuous functions with respect to the argument $\mathbf{x}$ and $t$; $\mathbf{d}$ is the nonlinear uncertainties which can be factorized as $\mathbf{d} = \Theta(t)\boldsymbol{\xi}(\mathbf{x}, t)$, where $\Theta \in R^{m \times n_1}$ is the time-varying uncertainties and $\boldsymbol{\xi}(\mathbf{x}, t) : R^n \times R_+ \to R^{n_1}$ is known nonlinear vector valued function which may in general include the global and local Lipschitzian functions as the subset. Here $n_1$ is a finite integer.

Regarding the uncertain nonlinear system (1), assumption $A_1$ is necessary for a repeatable control environment.

**Assumption 1** *The initial resetting condition $\mathbf{x}_i(0) = \mathbf{x}_d(0)$ is satisfied for all learning cycles where $i$ denotes the $i$-th learning cycle.*

The desired trajectory $\mathbf{x}_d \in R^n$ is continuously differentiable. The control objective is to find an appropriate control input sequence $\mathbf{u}_i$ for the nonlinear system (1) so as to achieve the perfect tracking performance as follows

$$\lim_{i\to\infty} ||\mathbf{e}_i(t)|| = 0 \quad \forall t \in [0,\ T_f] \qquad (2)$$

where $||\cdot||$ is the Euclidean norm for vectors. $\mathbf{e}(t)$ is the tracking error defined as $\mathbf{e}(t) \triangleq \mathbf{x}(t) - \mathbf{x}_d(t)$.

## 3 Nonlinear Optimal Control

First, we consider the nonlinear optimal control for the nominal part of nonlinear uncertain system (1) as follows

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},t) + B(\mathbf{x},t)\mathbf{u} \qquad \mathbf{f}(\mathbf{0},\cdot)=\mathbf{0} \qquad (3)$$

the error dynamics is:

$$\begin{aligned} \dot{\mathbf{e}}(t) &= \dot{\mathbf{x}}(t) - \dot{\mathbf{x}}_d(t) \\ &= \mathbf{f}(\mathbf{x},t) + B(\mathbf{x},t)\mathbf{u}(t) - \dot{\mathbf{x}}_d(t) \\ &= \tilde{\mathbf{f}}(\mathbf{e},t) + \tilde{B}(\mathbf{e},t)\mathbf{u}(t) \qquad (4) \end{aligned}$$

where $\tilde{\mathbf{f}}(\mathbf{e},t) = \mathbf{f}(\mathbf{e}+\mathbf{x}_d,t) - \dot{\mathbf{x}}_d(t)$, $\tilde{B}(\mathbf{e},t) = B(\mathbf{e}+\mathbf{x}_d,t)$. The objective function of optimal control is:

$$J = \inf_{\mathbf{u}(\cdot)} \int_0^{T_f} (q(\mathbf{e}) + \mathbf{u}^T\mathbf{u})dt \qquad (5)$$

where $q(\mathbf{e}) : R^n \to R_+$ is a continuously differentiable class-KR function with the desired solution being a state-feedback control law.

The HJB equation from above system (4) and objective function (5) is

$$q(\mathbf{e}) - \frac{1}{4}V_{\mathbf{e}}^*\tilde{B}\tilde{B}^T V_{\mathbf{e}}^{*T} + V_{\mathbf{e}}^*\tilde{\mathbf{f}}(\mathbf{e},t) = 0 \qquad (6)$$

where $V^*$ is commonly referred to as the *value function* and can be obtained from current error state $\mathbf{e}(t)$ [9]

$$V^*(\mathbf{e}(t)) = \inf_{\mathbf{u}(\cdot)} \int_t^{T_f} [q(\mathbf{e}(\tau)) + \mathbf{u}^T(\tau)\mathbf{u}(\tau)]d\tau \quad (7)$$

If the continuously differentiable solution of (6) exists, the optimal control law is given as

$$\mathbf{u}^* = -\frac{1}{2}\tilde{B}^T V_{\mathbf{e}}^{*T} \qquad (8)$$

The HJB partial equation solves the optimal control problem for every initial condition all at once. However it is difficult or almost impossible to solve HJB equation in general. Sontag's formula [2] provides a stable controller for system (4). The modified Sontag's formula [4] is introduced to provide a sub-optimal solution to nonlinear system (4) as follows

$$\mathbf{u}_{op} = \begin{cases} -\{\frac{V_{\mathbf{e}}\tilde{\mathbf{f}} + \sqrt{(V_{\mathbf{e}}\tilde{\mathbf{f}})^2 + q(\mathbf{e})\boldsymbol{\alpha}\boldsymbol{\alpha}^T}}{\boldsymbol{\alpha}\boldsymbol{\alpha}^T}\}\boldsymbol{\alpha}^T & \boldsymbol{\alpha} \neq \mathbf{0} \\ 0 & \boldsymbol{\alpha} = \mathbf{0} \end{cases} \quad (9)$$

where $\boldsymbol{\alpha} \triangleq V_{\mathbf{e}}\tilde{B}$, $V_{\mathbf{e}} \triangleq \frac{\partial V^T}{\partial \mathbf{e}} \in R^{1\times n}$ and $V$ is an arbitrary control Lyapunov function (CLF) [1] which is a continuously differentiable, positive, and radially unbounded function. For system (4), the existence of CLF implies that $\forall \mathbf{e} \neq \mathbf{0}$,

$$V_{\mathbf{e}}\tilde{B} = 0 \;\Rightarrow V_{\mathbf{e}}\tilde{\mathbf{f}} < 0.. \qquad (10)$$

The existence of CLF is a sufficient condition for stability of system (4).

**Remark 1** *The sub-optimal control law (9) has the sector margin $[1, \infty)$. The existence of sector margin makes the sub-optimal control law (9) robust with respect to certain classes of input disturbance (sector nonlinearities), but impossible to deal with general nonlinear time-varying uncertainties $\mathbf{d} = \Theta(t)\boldsymbol{\xi}(\mathbf{x},t)$ in (1). To overcome this problem, learning control is incorporated to deal with the learnable uncertainties and ensure the stability of the dynamic system.*

**Remark 2** *From the above sub-optimal control law (9), it can be seen that*

$$V_{\mathbf{e}}(\tilde{\mathbf{f}} + \tilde{B}\mathbf{u}_o) \leq 0, \quad (11)$$

*tracking error* $\mathbf{e}$ *of (4) is bounded since* $\dot{V} \leq 0$.

Next, we show that the sub-optimal control law (9) is bounded.

From the definition of the sub-optimal control law (9), $\mathbf{u}_{op}$ is continuous except for the surface $V_{\mathbf{e}}\tilde{B} = \mathbf{0}$. Hence when $||V_{\mathbf{e}}\tilde{B}(\mathbf{e})|| \geq \epsilon$, where $\epsilon$ is a non-infinitesimal constant, $\mathbf{u}_{op}$ is bounded from the boundedness of tracking error.

Assume that at a point $\mathbf{e}_0$, $V_{\mathbf{e}}\tilde{B}|_{\mathbf{e}=\mathbf{e}_0} = 0$ and $V_{\mathbf{e}}\tilde{\mathbf{f}}|_{\mathbf{e}=\mathbf{e}_0} < 0$. Since both $\tilde{\mathbf{f}}$ and $\tilde{B}$ are continuous with respect to all arguments, and $V$ is continuously differentiable, $\forall \epsilon > 0$, $\exists \delta > 0$, such that when $||\mathbf{e} - \mathbf{e}_0|| < \delta$, there are $|V_{\mathbf{e}}\tilde{B}(\mathbf{e})| < \epsilon$, and $V_{\mathbf{e}}\tilde{\mathbf{f}}(\mathbf{e}_0) < 0$. Define $b \triangleq V_{\mathbf{e}}\tilde{\mathbf{f}} \in R$ and $||\boldsymbol{\alpha}|| \triangleq \sqrt{\boldsymbol{\alpha}^T\boldsymbol{\alpha}}$, then

$$\begin{aligned} ||\mathbf{u}_{op}|| &\leq \left|\frac{-|b| + \sqrt{b^2 + q(\mathbf{e})||\boldsymbol{\alpha}||^2}}{||\boldsymbol{\alpha}||^2}\right| \, ||\boldsymbol{\alpha}|| \\ &\leq \left|\frac{-|b| + |b| + \sqrt{q(\mathbf{e})}||\boldsymbol{\alpha}||}{||\boldsymbol{\alpha}||^2}\right| \, ||\boldsymbol{\alpha}|| \\ &\leq \sqrt{q(\mathbf{e})} \end{aligned} \quad (12)$$

## 4 CEF and Sub-optimal Learning Control Approach

The error dynamics of system (1) is:

$$\dot{\mathbf{e}}(t) = \tilde{\mathbf{f}}(\mathbf{e}, t) + \tilde{B}(\mathbf{e}, t)[\mathbf{u} + \Theta\tilde{\boldsymbol{\xi}}(\mathbf{e}, t)], \quad (13)$$

where $\tilde{\boldsymbol{\xi}}(\mathbf{e}, t) = \boldsymbol{\xi}(\mathbf{e} + \mathbf{x}_d, t)$.

Define a non-negative composite energy function (CEF) at the $i - th$ learning cycle as:

$$E_i(t) = V(\mathbf{e}_i(t)) + \frac{1}{2\beta_v}\int_0^t trace[\Delta\Theta_i^T\Delta\Theta_i]d\tau \quad (14)$$

where $V$ can be an arbitrary CLF defined in Section 3 and satisfies (10). By definition, $\Delta\Theta_i \triangleq \Psi_i(t) - \Theta(t)$ is the parametric learning error at $i$-th learning repetition.

**Remark 3** *Composite energy function (CEF) is originated from Lyapunov function and subsequently extended to consecutive learning cycles. There are two parts in CEF. Control Lyapunov function part* $V(\mathbf{e}(t))$ *in CEF not only guarantees the finiteness of system states within a finite interval but also makes learning scheme applicable to general nonlinear system without satisfying global Lipschitz condition. A* $L^2$ *norm of parametric learning errors is also incorporated into CEF. Consequently the learning effect with respect to repeated learning cycles can be evaluated. By demonstrating the asymptotical convergence of CEF along the learning repetition horizon, the perfect tracking will be achieved.*

The integrated control law consists of two parts

$$\begin{aligned} \mathbf{u}_i(\mathbf{e}_i, t) &= \mathbf{u}_{i,o}(\mathbf{e}_i, t) + \mathbf{u}_{i,l}(\mathbf{e}_i, t) \\ \mathbf{u}_{i,l}(\mathbf{e}_i, t) &= -\Psi_i\tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t) \end{aligned} \quad (15)$$

where $\mathbf{u}_{i,o}$ is the sub-optimal control part at $i-th$ learning cycle which comes from (9), $\Psi_i\tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t)$ is the learning control part and is updated as follows

$$\begin{aligned} \Psi_i &= \Psi_{i-1} + \beta_v\boldsymbol{\alpha}_i^T\tilde{\boldsymbol{\xi}}^T(\mathbf{e}_i, t) \\ \boldsymbol{\alpha}_i &= V_{\mathbf{e}_i}\tilde{B}_i \end{aligned} \quad (16)$$

where the positive constant $\beta_v$ is the learning rate. Here the initial value of learning part is set as $\Psi_0(t) = 0 \; \forall t \in [0, \; T_f]$. We consider tracking performance from the first learning cycle $i = 1$. The convergence of proposed control scheme is given by the following theorem.

**Theorem 1** *The learning control law (15, 16) and optimal control law (9) under resetting condition* $A_1$ *guarantee the bounded and pointwise convergence of tracking error and the control signal is norm-bounded over* $[0, \; T_f]$ *when the learning repetition approaches to infinity.*

Proof: The proof consists of three parts. Part A derives the negative definiteness of the difference of the composite energy function (CEF). Part B proves the perfect tracking performance. Norm-boundedness property of the internal control signals is proved in Part C.

*Part A: The difference of CEF*

From (14), the difference of $E_i$ is

$$\begin{aligned} \Delta E_i(t) &\triangleq E_i(t) - E_{i-1}(t) \\ &= V(\mathbf{e}_i(t)) + \frac{1}{2\beta_v}\int_0^t trace[\Delta\Theta_i^T(\Delta\Theta_i)]d\tau \\ &\quad -\frac{1}{2\beta_v}\int_0^t trace[(\Delta\Theta_{i-1})^T(\Delta\Theta_{i-1})]d\tau \\ &\quad -V(\mathbf{e}_{i-1}(t)) \end{aligned} \quad (17)$$

According to assumption 1, control law (15), updating law (16) and property (11)

$$\begin{aligned} V(\mathbf{e}_i(t)) &= \int_0^t V_{\mathbf{e}_i}[\tilde{\mathbf{f}}_i + \tilde{B}(\mathbf{e}_i, \tau)(\mathbf{u}_i + \Theta\tilde{\boldsymbol{\xi}}(\mathbf{e}, \tau))]d\tau \\ &\quad +V(\mathbf{e}_i(0)) \\ &\leq \int_0^t \boldsymbol{\alpha}_i[-\Psi_{i-1}\tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t) + \Theta\tilde{\boldsymbol{\xi}}(\mathbf{e}_i, \tau)]d\tau \\ &\quad -\int_0^t \beta_v\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i^T\tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t)^T\tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t)d\tau \end{aligned}$$

$$\leq \int_0^t \varsigma(\tau) d\tau \quad (18)$$

$$\varsigma(t) \triangleq \boldsymbol{\alpha}_i \Theta \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t) - \boldsymbol{\alpha}_i \Psi_{i-1} \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t) - \beta_v \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i^T \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t)^T \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t). \quad (19)$$

Note the following property of *trace*

$$trace(\mathbf{y}\mathbf{v}^T Q) = \mathbf{y}^T Q \mathbf{v} \quad (20)$$

where $Q \in R^{m\times n}$, $\mathbf{v} \in R^{m\times 1}$ and $\mathbf{y} \in R^{n\times 1}$. According to the learning update law (16)

$$\begin{aligned} & \frac{1}{2\beta_v}\{trace[(\Delta\Theta_i^T \Delta\Theta_i] - trace[\Delta\Theta_{i-1}^T \Delta\Theta_{i-1}]\} \\ = & \frac{1}{2\beta_v} trace[(\Psi_i - \Psi_{i-1})^T (\Psi_i + \Psi_{i-1} - 2\Theta)] \\ = & \boldsymbol{\alpha}_i \Psi_{i-1} \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t) - \boldsymbol{\alpha}_i \Theta \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t) \\ & + \frac{\beta_v}{2} \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i^T \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t)^T \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t) \\ = & -\varsigma(t) - \frac{\beta_v}{2} \boldsymbol{\alpha}_i \boldsymbol{\alpha}_i^T \tilde{\boldsymbol{\xi}}^T \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t). \quad (21) \end{aligned}$$

According to (18) and (21), $\boldsymbol{\alpha}_i \boldsymbol{\alpha}_i^T \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t)^T \tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t) = ||\boldsymbol{\alpha}_i||^2 ||\tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t)||^2$, we have

$$\Delta E_i(t) \leq -\int_0^t \frac{\beta_v}{2} ||\boldsymbol{\alpha}_i||^2 ||\tilde{\boldsymbol{\xi}}(\mathbf{e}_i, t)||^2 d\tau - V(\mathbf{e}_{i-1}(t)) \leq 0. \quad (22)$$

The finiteness of $E_i(t)$ is ensured for any learning cycle provided $E_1(t)$ is finite.

**Remark 4** *From (22), it shows that a larger learning gain $\beta_v$ may improve the convergence rate.*

When $i = 1$, $\Psi_1 = \beta_v \boldsymbol{\alpha}_1^T \tilde{\boldsymbol{\xi}}^T(\mathbf{e}_1, t)$. We prove in the following step that the learning control law (15) and sub-optimal control law (9) can guarantee that the tracking error at the first learning cycle is bounded. The energy function $E_1$ is

$$E_1(t) = V_1(t) + \frac{1}{2\beta_v}\int_0^t trace[\Delta\Theta_1^T \Delta\Theta_1] d\tau. (23)$$

Taking the derivative of the composite energy function

$$\begin{aligned} \dot{E}_1(t) = & V_{\mathbf{e}_1}[\tilde{\mathbf{f}}_1 + \tilde{B}(\mathbf{e}_1, \tau)(\mathbf{u}_{o,1} - \Psi_1 \tilde{\boldsymbol{\xi}}(\mathbf{e}_1, t) \\ & + \Theta \tilde{\boldsymbol{\xi}}(\mathbf{e}_1, t))] + \frac{1}{2\beta_v} trace[\Delta\Theta_1^T \Delta\Theta_1] \\ \leq & -\boldsymbol{\alpha}_1 \Theta \tilde{\boldsymbol{\xi}}(\mathbf{e}, t) - \beta_v \boldsymbol{\alpha}_1 \boldsymbol{\alpha}_1^T \tilde{\boldsymbol{\xi}}^T(\mathbf{e}_1, t) \tilde{\boldsymbol{\xi}}(\mathbf{e}_1, t) \\ & + \boldsymbol{\alpha}_1 \Theta \tilde{\boldsymbol{\xi}}(\mathbf{e}_1, t) + \frac{1}{2\beta_v} trace(||\Theta||^2) \\ & + \frac{\beta_v}{2} \boldsymbol{\alpha}_1 \boldsymbol{\alpha}_1^T \tilde{\boldsymbol{\xi}}^T(\mathbf{e}_1, t) \tilde{\boldsymbol{\xi}}(\mathbf{e}_1, t) \\ \leq & -\frac{\beta_v}{2} ||\boldsymbol{\alpha}_1||^2 ||\tilde{\boldsymbol{\xi}}(\mathbf{e}_1, t)||^2 \\ & + \frac{1}{2\beta_v} trace(||\Theta||^2) \\ \leq & \frac{1}{2\beta_v} trace(||\Theta||^2). \quad (24) \end{aligned}$$

Since $\Theta(t)$ is a continuous function, $\frac{1}{2\beta_v} trace(||\Theta||^2) \leq \theta_0 \ \forall t \in [0, T_f]$. Therefore $E_1(t) \leq \theta_0 T_f$ is bounded in the finite time interval $[0,\ T_f]$. The boundedness of tracking error comes from the boundedness of CLF $V(\mathbf{e}_i(t))$.

**Remark 5** *The tracking error is bounded at each learning cycle so that the proposed control scheme can prevent the finite escape time phenomenon.*

*Part B: Learning convergence*

Since $\Delta E_i(t) \leq 0$, following (22) the CEF at $k$-th learning cycle is:

$$\begin{aligned} E_k(t) = & E_1(t) + \sum_{i=2}^{k} \Delta E_i(t) \\ \leq & E_1(t) - \sum_{i=1}^{k} V(\mathbf{e}_{i-1}(t)). \quad (25) \end{aligned}$$

When $k \to \infty$,

$$\lim_{k\to\infty} E_k(t) \leq E_1(t) - \lim_{k\to\infty} \sum_{i=1}^{k} V(\mathbf{e}_{i-1}(t)). \quad (26)$$

$\forall t \in [0,\ T_f]$, $E_k(t)$ is a non-increasing series with upper bound. Therefore $\lim_{k\to\infty} E_k(t)$ exists. Since

$$\lim_{k\to\infty} E_k(t) \leq E_1(t) - \lim_{k\to\infty} \sum_{i=1}^{k} V(\mathbf{e}_{i-1}(t)), \quad (27)$$

$\sum_{i=1}^{\infty} V(\mathbf{e}_{i-1}(t))$ converges and $\lim_{i\to\infty} V(\mathbf{e}_{i-1}(t)) = 0 \ \forall t \in [0,\ T_f]$ is guaranteed. From the definition of CLF, $V$ is radially unbounded function, therefore,

$$\begin{aligned} & \lim_{i\to\infty} V(\mathbf{e}_{i-1}(t)) = 0 \ \forall t \in [0,\ T_f] \\ \Rightarrow & \lim_{i\to\infty} \mathbf{e}_i(t) = \mathbf{0}. \quad (28) \end{aligned}$$

Hence the tracking error also converges to zero $\forall t \in [0,\ T_f]$. The system state can track the desired trajectory perfectly as the learning repetition approaches to infinity.

*Part C: Norm-bounded control profile*

We consider the total energy of the control authority over the learning cycle defined as the following norm

$$||\mathbf{u}||_{[0,T_f]} \triangleq \int_0^{T_f} ||\mathbf{u}||^2 d\tau. \quad (29)$$

It can be seen from the construction of the control law (15), when the learning repetition is finite, the learning control part is continuous in finite time interval. The boundedness of learning control part is ensured from the boundedness of tracking error. At the same time, since the tracking error is bounded, sub-optimal control part is bounded. Hence the norm-boundedness of control law is ensured for finite learning repetitions.

When learning repetition approaches to infinity, $\int_0^{T_f} trace[\Delta\Theta_\infty^T \Delta\Theta_\infty] d\tau \leq E_\infty(T_f) \leq E_1(T_f) \leq M$ where $M$ is the maximum value of $E_1(t)$. Define $a_{j,k} \triangleq \{\Delta\Theta_\infty\}_{j,k},\ j = 1, \dots, m,\ k = 1, \dots, n_1$,

$$\int_0^{T_f} trace[\Delta\Theta_\infty^T \Delta\Theta_\infty] d\tau = \int_0^{T_f} \sum_{j,k} |a_{j,k}|^2 d\tau \leq M, \quad (30)$$

hence $\int_0^{T_f} |a_{j,k}|^2 d\tau$ is bounded for $j = 1, \dots, m,\ k = 1, \dots, n_1$. Since $L^2 \subset L^1$, $\int_0^{T_f} |a_{j,k}| d\tau$ is also bounded. Define $\psi_{j,k} \triangleq \{\Psi_\infty\}_{j,k},\ \theta_{j,k} \triangleq \{\Theta\}_{j,k}\ j = 1, \dots, m,\ k = 1, \dots, n_1$. The boundedness of $\int_0^{T_f} |\psi_{j,k}| d\tau$ is ensured from the boundedness of $\int_0^{T_f} |a_{j,k}| d\tau$ and $\theta_{j,k}$. Assume that $\max_{t\in[0,\ T_f]} |\theta_{j,k}(t)| \leq M_1$ for $j = 1, \dots, m,\ k = 1, \dots, n_1$,

$$\begin{aligned} \int_0^{T_f} \psi_{j,k}^2 d\tau &\leq \int_0^{T_f} |a_{j,k}|^2 d\tau + 2\int_0^{T_f} |\psi_{j,k}|\ |\theta_{j,k}| d\tau \\ &\quad - \int_0^{T_f} \theta_{j,k}^2 d\tau \\ &\leq \int_0^{T_f} |a_{j,k}|^2 d\tau + 2M_1 \int_0^{T_f} |\psi_{j,k}| d\tau \\ &\quad - \int_0^{T_f} \theta_{j,k}^2 d\tau, \end{aligned} \quad (31)$$

$\int_0^{T_f} \psi_{j,k}^2 d\tau$ is bounded, so is $\int_0^{T_f} ||\Psi_\infty||^2 d\tau$. The boundedness of $||\mathbf{e}(t)||$ leads to the boundedness of $||\tilde{\boldsymbol{\xi}}(\mathbf{e}, t)||$, which is denoted as $\max_{t\in[0,\ T_f]} ||\tilde{\boldsymbol{\xi}}(\mathbf{e}, t)|| \leq M_2 < \infty$. Since $\mathbf{u}_\infty = -\Psi_\infty \tilde{\boldsymbol{\xi}}(\mathbf{e}_\infty, t)$

$$\begin{aligned} ||\mathbf{u}_{\infty,l}||_{[0,\ T_f]} &= \int_0^{T_f} ||\mathbf{u}_{\infty,l}||^2 d\tau \\ &\leq \int_0^{T_f} ||\Psi_\infty||^2 ||\tilde{\boldsymbol{\xi}}(\mathbf{e}_\infty, t)||^2 d\tau \\ &\leq M_2 \int_0^{T_f} ||\Psi_\infty||^2 d\tau, \end{aligned} \quad (32)$$

hence the norm-boundedness of the control input is ensured because the sub-optimal control part is always bounded at every learning cycle. ■

## 5 Illustrative Example

To show the effectiveness of the proposed learning control method, the following chaotic Duffing system is considered.

$$\ddot{y} = -0.1\dot{y} - y^3 + 12cos(t) + u(t) \quad (33)$$

which is a nonlinear time-varying system and shows chaotic behavior when $u(t) = 0$. Here $12cos(t)$ is the unknown parametric uncertainty. The objective is to drive the system state $[x_1(t),\ x_2(t)]^T$ to follow the given trajectory $[sin(3t), 3cos(3t)]^T$. The tracking period is $[0,\ 2\pi]$.

The control Lyapunov function is chosen to be $V = (ce_1 + e_2)^2 + e_1^2$. The objective function of the system (33) is

$$J = \int_0^{2\pi} [k(e_1^2 + e_2^2) + u^2] d\tau \quad (34)$$

where $k$ is a penalty factor which balances the performance between better tracking and less control effort.

**Remark 6** *In real application, the learning cycle can hardly be repeated infinite times. Learning update is ceased when the tracking error reaches the pre-defined error bound, i.e.*

$$\max_{t\in[0,\ T_f]} ||\mathbf{e}_k(t)|| \leq \epsilon_0 \quad (35)$$

In this example , the error bound is chosen to be $\epsilon_0 = 0.005$.

To compare the control performance, the objective function with different weightings are considered. First, for fast tracking error convergence, $k = 1$ is selected. The tracking error profile along learning repetition horizon is shown in Fig. 1, from which the learning effect is evident.

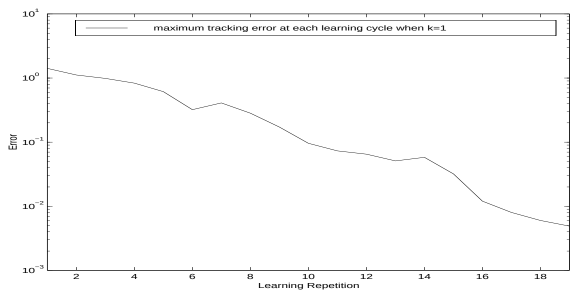


**Figure 1:** The convergence of tracking error along learning repetition horizon when $k = 1$

When less control effort is required, a smaller $k$ is chosen to be 0.0001. The tracking error profile along learning repetition horizon is shown in Fig. 2. It can be seen that, more learning repetitions are needed to reach the prespecified error bound. At the same time, the maximum tracking error is also larger compared with the same learning repetition when $k = 1$.

Now let us compare the control inputs in above two cases. Define $||u_i||_s \triangleq \max_{t\in[0,\ T_f]} |u(t)|$. Different $||u_i||_s$

are shown in Fig. 3 and Fig. 4 for $k = 1$ and $k = 0.0001$ respectively. When $k = 1$, i.e., fast tracking convergence is preferred, the maximum control inputs at the first few learning repetitions are quite large. When $k = 0.0001$, i.e., less control effort is preferred, the maximum control effort can be kept at a lower level less than half of the former case.

In a word, by tuning the value of the penalty factor $k$, we can effectively change the convergence rate of the tracking error and the magnitude of the control effort.

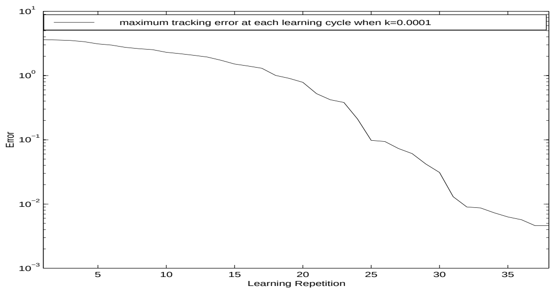


**Figure 2:** The convergence of tracking error along learning repetition horizon when $k = 0.0001$

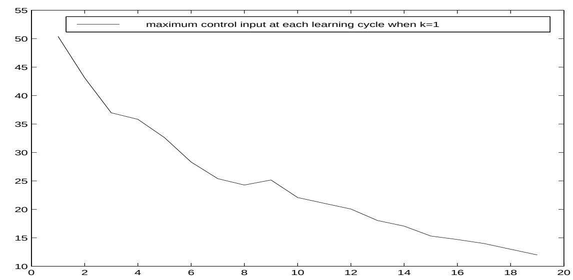


**Figure 3:** The maximum control input along learning repetition horizon when $k = 1$

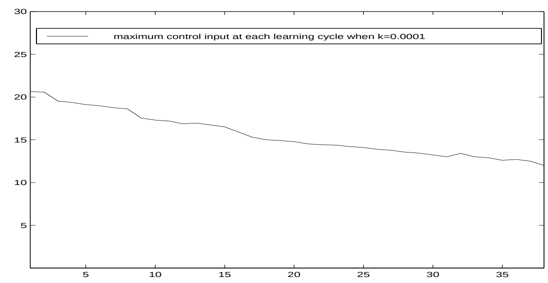


**Figure 4:** The maximum control input along learning repetition horizon when $k = 0.0001$

## 6 Conclusion

A new method of integrating learning control and nonlinear sub-optimal control is proposed in this paper. The idea of the composite energy function (CEF) is introduced, which provides a framework to facilitate the design and analysis of sub-optimal learning control systems. The proposed method can achieve the perfect state tracking with sub-optimal control performance for a class of nonlinear systems with the time-varying parametric uncertainties. Sub-optimal control based on Sontag's formula is utilized to effectively change the convergence rate of the tracking error and the magnitude of the control effort. The sub-optimal control scheme also provides stability margin of the dynamic system so as to enhance the robustness with respect to input disturbance. The learning control part, on the other hand, can effectively handle the unknown time-varying parametric uncertainties. The control performance is enhanced along learning repetition horizon so that perfect tracking can be achieved asymptotically. At the same time the norm-boundedness of control signal is guaranteed . The validity of the proposed scheme has been confirmed by simulation results.

## References

[1] R.Sepulchre, M. Jankovic and P.V. Kokotovic, "Constructive nonlinear control", London:; New York: Springer,c1997, (1997).

[2] Sontag, E. D., 'A Lyapunov-like characterization of asymptotic controllability,' *SIAM J. Contr. Optimiz.*, **21**, 462-471, (1989).

[3] M. Kristic and P.Tsiotras "Inverse optimal stabilization of a rigid spacecraft," *IEEE Trans. Automatic Control* **44**(5), 1042-1049. (1999)

[4] J.A.Primbs, et al., "Nonlinear optimal control: A control Lyapunov function and receding horizon perspective," *Asian Journal of Control* **1**, 14-24. (1999)

[5] S.Arimoto, S. Kawamura and F. Miyazaki, "Bettering operation of robust by learning," *J. Robotic Systems*, vol.1, pp.123–140, (1984).

[6] T.Y.Kuc, J.S.Lee and K.Nan " An iterative learning control theory for a class of nonlinear dynamic systems," *Automatica*, vol.28, pp1215–1221,(1992)

[7] S.S.Saab "A discrete-time learning control algorithm for a class of linear time-invariant," *IEEE Trans. Automat. Control*,vol.46, no.6, pp1138–1142, (1995)

[8] D.W. Wang, "Convergence and robust of discrete time nonlinear systems with iterative learning control," *Automatica*, vol.34, no.11,pp1445–1448,(1998)

[9] A.E. Bryson and Y.H.Ho Applied Optimal Control, Hemisphere Publ. Corp., Washington D.C.. (1975)