2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Synchronized Optimization With Prescribed Performance for High-order Strict-feedback System

Yuxiang Zhang[1], Xiaoling Liang[1], Dongyu Li*[2], Shuzhi Sam Ge[1], and Tong Heng Lee[1]

Abstract— **This paper investigates synchronized optimization with prescribed performance for the strict-feedback system with time-synchronized convergence property, which is the highly essential performance desired in various real-world high-precision control applications. The prescribed performance is considered to keep the state-variables within a predefined region during the control period to meet the required system performance. To consider optimization performance while also concurrently attaining the time-synchronized properties simultaneously of each backstepping subsystem, optimized backstepping is utilized to establish the learning framework; wherein the norm-normalized sign function is appropriately incorporated in each backstepping subsystem, which generates the decomposition of the optimal system control and gradient term of the cost function with appropriate time-synchronized control items and unknown independently learning parts to be approximated with neural networks. With this decomposition design, the learning objective is transformed to adaptively explore the optimal control parameter in the admissible policy region. By additionally employing the adaptive dynamic programming technique, actor-critic method, and gradient-constrained method, the solution of the Hamilton-Jacobi-Bellman equation is iteratively approximated while the learnable parameter stays within the predefined region. The work here has the outcome of time-synchronized convergence which surpasses the usual typical developments in this class of problems considered. The proposed method is verified with the vehicle platoon problem to show its effectiveness in that the system preserves special properties of time-synchronized stability and control while optimizing the overall system control.**

I. INTRODUCTION

High-precision performance optimization and control have raised great attention and are being applied to various areas, which further require the specific performance in the finite- or fixed-time convergence, (fixed) time-synchronized convergence, constrained performance, prescribed performance, optimized performance, and so on [1]. Generally speaking, the standalone performance is not adequate for the demand of accomplishing multiple complex specific tasks, the constrained and synchronized optimization is studied in this work, which essentially considers the prescribed performance and time-synchronized convergence.

The practical systems are usually safety-critical systems, which require system performance to satisfy certain limits, and are denoted as the safe zone/region. To satisfy the safety requirement of the system, the constraints on the system state-variables are usually considered. Regarding the constrained control methods, the Barrier Lyapunov function (BLF) can effectively keep the system state-variables stayed within the desired zone by keeping the BLF bounded in the close-loop [2], [3]. The prescribed performance control (PPC) is a more direct way to constrain the state-variables and realize prescribed tracking performance [4]. Model Predictive Control (MPC) is another effective method to consider system constraints explicitly by solving the open-loop optimization problem in the predictive horizon with the consideration of system tracking performance and control inputs, which usually require substantial online calculation [5].

The important concepts of time-synchronized stability (TSS) and fixed time-synchronized stability (FTSS), also termed as a simultaneous-arrival-to-origin (SATO) convergence, specifically require all state-variables of the system converge to the origin at the same time [6], which have been well-investigated and recognized as the critical demand in the control design of various practical systems, e.g. synchronously motion and cooperatively work in the multi-agent cooperative missions, the synchronization control for shift control of inverse automated manual transmission to rejects jerk [7], adaptive reference-free trajectory planning of autonomous vehicles under multi-scenario driving [8], synchronized control governor for autonomous vehicle motion control to reduce the chattering [9], etc.

The time-synchronized property is innovated by the core analytical methodology of the formally defined norm-normalized sign function when compared with the classical sign function [10]. The characteristic of the norm-normalized sign function is that each output element contains information on the whole input vector, which enables the overall system control to guide all system state-variables converge time-synchronously [11]. This ultimate time-synchronized convergence property is also additionally notable for attaining superior system performance with an outcome of a shorter and smoother output trajectory; less chattering on the output trajectory; and also lower energy consumption; which can be obtained after the optimization of control inputs. To consider optimization performance on system control while also concurrently attaining the important time-synchronized properties simultaneously, reinforcement learning (RL) and

[1]Y. Zhang, X. Liang, S. S. Ge, and T. H. Lee are with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore, 117576.

[2]D. Li is with the School of Cyber Science and Technology, Beihang University, Beijing 100191, China, also from Tianmushan Laboratory, Hangzhou, 310023 P. R. China. dongyuli@buaa.edu.cn

This research/project is partially supported by the Ministry of Education, Singapore, under its Research Centre of Excellence award to the Institute for Functional Intelligent Materials (I-FIM, project No. EDUNC-33-18-279-V12), Tianmushan Laboratory Research Project TK-2023-B-010 and TK-2023-C-020, the National Nature Science Foundation of China No. 62273256, Key Laboratory of Industrial Internet of Things & Networked Control, Ministry of Education under Project 2022FF04.

adaptive dynamic programming (ADP) that are always regarded as effective learning-based sequential control methods can be employed by satisfying Bellman optimality and solving the Hamilton-Jacobi-Bellman (HJB) equations [12]. Due to the inherent nonlinearity of the HJB equation, the unique solution is usually impossible and rather complicated to be directly obtained, which requires neural networks (NNs) as an approximation for iteratively solving [13].

From a deeper perspective, synchronized optimization with prescribed performance and time-synchronized convergence property for high-order system control is an important requirement, which is not yet well established in current research. In this work, for the high-order strict feedback system, the prescribed performance is considered to keep the state-variables within a predefined region during the control period to meet the required system performance. To consider optimization performance on system control while also concurrently attaining the important time-synchronized properties simultaneously of each backstepping subsystem, optimized backstepping is utilized that considers the learning-based method with ADP and NNs approximation to solve the HJB equation with iterations. In each backstepping subsystem, control is decomposed into the nonlinear control law and NN-based unknown part to be optimized via iterative learning. With this decomposition, the learning part of reinforcement learning is transformed to adaptively explore the optimal control parameter in the admissible policy region. By designing the specific form of the control law with the norm-normalized sign function into the learning framework, the system preserves special properties of time-synchronized stability and control while optimizing the overall system control. The proposed method is applied to the vehicle platoon problem, which has verified the effectiveness of the optimization and constrained control with the time-synchronized property.

Notation: $\mathbb{R}$ and $\mathbb{R}^n$ are the set of real numbers and the n dimensional Euclidean space, respectively. $\|\cdot\|$ denotes the vector norm or matrix norm in $\mathbb{R}^n$. $\mathrm{diag}(x_1, x_2, ..., x_n)$ is a diagonal matrix with diagonal elements x_1, x_2,..., x_n. $x \prec y$ denotes that vector y is element-wisely greater and no less than the vector x, respectively.

II. Synchronized Optimization with Prescribed Performance

Consider the strict-feedback system with the following form as

$$\begin{cases} \dot{x}_i = f_i(\bar{x}_i) + g_i(\bar{x}_i)x_{i+1}, \\ \dot{x}_n = f_n(\bar{x}_n) + g_n(\bar{x}_n)u, \end{cases} \tag{1}$$

where $\bar{x}_i = (x_1, x_2, ..., x_i)$, $x_i \in \mathbb{R}^m$ denotes the system state-variable, $u \in \mathbb{R}^m$ is the control input.

Assumption 1: [14] *The system dynamics $f_i(\bar{x}_i) \in R^m$ and $g_i(\bar{x}_i) \in \mathbb{R}^{m\times m}$ are second-order differentiable. $g_i(\bar{x}_i)$ satisfy* $\det(g_i(\bar{x}_i)) \neq 0$.

Let the reference signal $y_d(t)$ be a sufficiently smooth function. Define the tracking error for the first backstepping step as $z_1 = x_1 - y_d$. The control objective is to design the synchronized optimization for the strict-feedback system with prescribed performance, such that the control law in each backstepping subsystem is optimized under the reinforcement learning framework with the preserved time-synchronized property. The diagram of the proposed algorithm is shown in Fig.1.

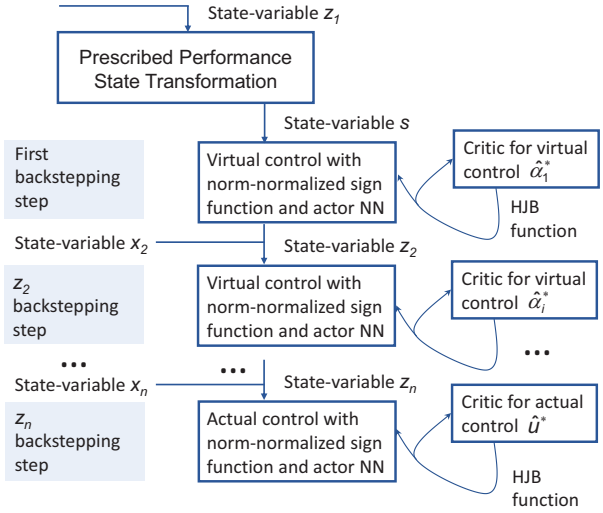

Fig. 1. The diagram of the proposed algorithm

A. Prescribed performance state transformation

With consideration of the prescribed performance, system control is designed to ensure the constraints on state-variable $z_1(t)$ as

$$\underline{\eta}\rho(t) \leq z_1(t) \leq \overline{\eta}\rho(t) \tag{2}$$

where $\underline{\eta}$, $\overline{\eta}$ are positive constants, $\rho(t)$ is a time-varying strictly decreasing and positive function.

In order to achieve the prescribed performance, the constraint in (2) is then transformed as unconstrained control, in which $\tilde{s}(t)$ is the transformed error

$$\tilde{s}(t) = \Phi^{-1}(z_1(t)/\rho(t)) \tag{3}$$

where $\Phi^{-1}(\cdot)$ is the inverse function of $\Phi(\cdot)$, and $\Phi(\cdot)$ is smooth, strictly increasing, denoted as Φ for clarification. Design the transformation function with the following form [15]

$$\Phi(\tilde{s}(t)) = \frac{\overline{\eta}e^{\tilde{s}(t)} - \underline{\eta}e^{-\tilde{s}(t)}}{e^{\tilde{s}(t)} + e^{-\tilde{s}(t)}} \tag{4}$$

Then, (3) is changed into

$$\tilde{s}(t) = \Phi^{-1}\left(\frac{z_1(t)}{\rho(t)}\right) = \frac{1}{2}\ln\frac{\underline{\eta} + \Phi}{\overline{\eta} - \Phi} \tag{5}$$

Further, reformat the error signal as

$$s(t) = \tilde{s}(t) - \frac{1}{2}\ln(\underline{\eta}/\overline{\eta}) - \varrho(\tilde{s}(t)) \tag{6}$$

where $\frac{1}{2}\ln(\underline{\eta}/\overline{\eta})$ is used to keep that $s(t) \in \left[-\frac{1}{2}\ln(\underline{\eta}/\overline{\eta}), \frac{1}{2}\ln(\underline{\eta}/\overline{\eta})\right]$, $\varrho(\tilde{s}(t))$ is used to relax the

limitation of nonzero initial of $\dot{s}(0)$ and $\ddot{s}(0)$, which can be expressed as [16]

$$\varrho(s(t)) = (\dot{s}(0)t + \frac{1}{2}\ddot{s}(0)t^2)e^{-t^2} \tag{7}$$

B. Time-synchronized Optimization Design

Synchronized optimization is formulated with the preliminaries of the well-defined FTSS/TSS with norm-normalized sign function from the earlier work in [6], [11]. The learning mechanism of synchronized optimization is constructed with the optimized backstepping and the time-synchronized control methods, which will be stated in the following.

Step 1: The time derivative of error dynamics of $s(t)$

$$\begin{aligned}\dot{s}(t) =&\dot{\tilde{s}} - \dot{\varrho}(\tilde{s})\\ =&\varpi\left(\dot{z}_1 - \frac{\dot{\rho}(t)z_1}{\rho(t)}\right) - \dot{\varrho}(\tilde{s})\\ =&\varpi\left(f_1(\bar{x}_1) + g_1(\bar{x}_1)x_2 - \dot{y}_d - \frac{\dot{\rho}(t)z_1}{\rho(t)}\right) - \dot{\varrho}(\tilde{s})\\ =&f_s(\bar{x}_1) + g_s(\bar{x}_1)x_2\end{aligned} \tag{8}$$

where $f_s(\bar{x}_1) = \varpi(f_1(\bar{x}_1) - \dot{y}_d - (\dot{\rho}(t)z_1)/\rho(t)) - \dot{\varrho}(\tilde{s})$, $g_s(\bar{x}_1) = \varpi g_1(\bar{x}_1)$, $\varpi = \frac{1}{2\rho(t)}\left(\frac{1}{\Phi+\eta} - \frac{1}{\Phi-\overline{\eta}}\right)$.

To formulate the optimization problem in subsystem s with the optimized virtual control α_1^*, design optimal cost function as

$$J^*(s) = \int_t^\infty (s^T K_s s + \alpha_1^{*T} K_{\alpha_1}\alpha_1^*)\mathrm{d}\tau, \tag{9}$$

where $K_s = \mathrm{diag}(k_{s_1}, k_{s_2}, ..., k_{s_m})$ and $K_{\alpha_1} = \mathrm{diag}(k_{\alpha_{11}}, k_{\alpha_{12}}, ..., k_{\alpha_{1m}})$ are positive definite matrices.

The optimal gradient term of cost function $J^*(s)$ and the optimal virtual control α_1^* can be obtained by solving the HJB function [17], which can be expressed as

$$\begin{aligned}H\left(s, \alpha_1^*, \frac{\partial J^*(s)}{\partial s}\right) =\ & s^T K_s s + \alpha_1^{*T} K_{\alpha_1}\alpha_1^*\\ &+ \frac{\partial J^*(s)}{\partial s}^T\left(f_s(\bar{x}_1) + g_s(\bar{x}_1)\alpha_1^*\right)\end{aligned} \tag{10}$$

For simplification, denote $J_s^* \triangleq \partial J^*(s)/\partial s \in \mathbb{R}^m$, $H\big(s, \alpha_1^*, J_s^*\big) \triangleq H_s$. The unique solution of $H_s = 0$ can be obtained by calculating $\partial H_s/\partial\alpha_1^* = 0$, which means the optimal gradient term of cost function J_s^* and the optimal virtual control α_1^* satisfies

$$\alpha_1^* = -\frac{1}{2}K_{\alpha_1}^{-1} g_s^T(\bar{x}_1)J_s^*. \tag{11}$$

Following the time-synchronized control methodology, the estimation of the optimized fixed time-synchronized control (TSC) virtual control law can be expressed as

$$\hat{\alpha}_1^* = g_s^{-1}(\bar{x}_1)(-f_s(\bar{x}_1) - \sigma_s - \hat{h}_{a_1}) \tag{12}$$

where $\sigma_s \in \mathbb{R}^m$ is a switching variable that contains M_s and is triggered by s to avoid the singularity problem. $\hat{h}_{a_1}$ is the unknown learning part to be iterated. M_s, σ_s and $\hat{h}_{a_1}$ are designed as

$$M_s = \kappa_{11}\mathrm{sig}_\mathrm{n}^{p_1}(s) + \kappa_{12}\mathrm{sig}_\mathrm{n}^{p_2}(s), \tag{13}$$

$$\sigma_s = \begin{cases} M_s & \text{if } \|s\| > \varepsilon,\\ \gamma_1 s + \gamma_2 \mathrm{sig}_\mathrm{n}^4(s), & \text{if } \|s\| \le \varepsilon,\end{cases} \tag{14}$$

$$\hat{h}_{a_1} = \hat{b}_{a_1}\,\mathrm{sig}_\mathrm{n}^{p_1}(s) \tag{15}$$

where κ_{11}, κ_{12} are positive constants, p_1, p_2 are defined as $p_1 = p_{11}^*/p_{12}^* > 1$, $0 < p_2 = p_{21}^*/p_{22}^* < 1$, with positive odd integers p_{11}^*, p_{12}^*, p_{21}^*, p_{22}^*, ε is a small positive constant, $\hat{h}_{a_1}$ is designed specifically to preserve the time-synchronized property, which contains the learnable control parameter $\hat{b}_{a_1}$ that is defined as the output of Actor NNs. The forms of γ_1 and γ_2 are designed to ensure the continuity of the switching algorithm on σ_s and $\dot{\sigma}_s$. The detailed forms can be found in [6], [11], which are omitted here.

Remark 1: *The decomposed form with the nonlinear design method for the design of the optimized control $\hat{\alpha}_1^*$ in (12) can naturally provide the required admissible appropriate initial policy as well as suitable learning region [14], [17], [18], which enables the policy is TSC during the whole learning period [6], [11].*

Correspondingly, with (11), the estimation of optimal gradient term of value function J_s^* yields

$$\hat{J}_s^* = 2G_s^{-1}\big(f_s(\bar{x}_1) + \sigma_s + \hat{h}_{c_1}\big) \tag{16}$$

where $G_s = g_s(\bar{x}_1)K_{\alpha_1}^{-1}g_s^T(\bar{x}_1)$, $\hat{h}_{c_1} = \hat{b}_{c_1}\,\mathrm{sig}_\mathrm{n}^{p_1}(s) \in \mathbb{R}^m$ is the independent of the learning part $\hat{h}_{a_1}$ in (12). $\hat{b}_{c_1}$ is the learnable control parameter that is output by Critic NNs. Substituting (12) and (16) into (10), the estimation of H_s can be represented as

$$\hat{H}_s = s^T K_s s + \hat{\alpha}_1^{*T} K_{\alpha_1}\hat{\alpha}_1^* + \hat{J}_s^{*T}\big(-\sigma_s - \hat{h}_{a_1}\big) \tag{17}$$

The learning objective for subsystem s is to optimize the learning parts $\hat{h}_{a_1}$ and $\hat{h}_{c_1}$ and constrained the parameters $\hat{b}_{a_1}$ and $\hat{b}_{c_1}$ in a specific range enabling to iteratively find the approximated solution of the optimal system control. The constraints on the learnable region of parameters $\hat{b}_{a_1}$ and $\hat{b}_{c_1}$ will be analyzed with the Lyapunov synthesis.

Remark 2: *For the reason there is no explicit control law regarding the unique solution of the nonlinear HJB function with the nonlinear model, the iteration-based method is required to approximate the numerical solution with NNs. The decomposed design enables the ultimate control to satisfy some predefined characteristic, e.g. safe and time-synchronized, which does not influence the relationship between the optimal control policy and the optimal gradient term of the cost function. The approximation $\hat{h}_a$ and $\hat{h}_c$ represent the residual unknown parts between the expected optimized system control and the classic fixed TSC law. The learning part $\hat{h}_a$ and $\hat{h}_c$ with the specific form preserve the time-synchronized property during the learning process.*

Step i ($i = 2, 3, ..., n-1$): Define the tracking error for subsystem z_i as $z_i = x_i - \alpha_{i-1}$. The time derivative of error dynamics of z_i is

$$\dot{z}_i = f_i(\bar{x}_i) + g_i(\bar{x}_i)x_{i+1} - \dot{\alpha}_{i-1} \tag{18}$$

Formulate the optimization problem in subsystem z_i with the optimized virtual control α_i^*, define optimal cost function as

$$J^*(z_i) = \int_t^\infty (z_i^T K_{z_i} z_i + \alpha_i^{*T} K_{\alpha_i} \alpha_i^*) \mathrm{d}\tau, \tag{19}$$

where $K_{z_i} = \mathrm{diag}(k_{z_{i1}}, k_{z_{i2}}, ..., k_{z_{im}})$ and $K_{\alpha_i} = \mathrm{diag}(k_{\alpha_{i1}}, k_{\alpha_{i2}}, ..., k_{\alpha_{im}})$ are positive definite matrices.

Following the same way in subsystem z_1, the estimation of the optimized fixed TSC virtual control $\hat{\alpha}_i^*$ and the optimal gradient term of cost function $\hat{J}_{z_i}^*$ in subsystem z_i can be expressed as

$$\hat{\alpha}_i^* = g_i^{-1}(\bar{x}_i)\big(- f_i(\bar{x}_i) - \sigma_{z_i} - \hat{h}_{a_i} - \zeta_i + \dot{\hat{\alpha}}_{i-1}\big) \tag{20}$$

$$\hat{J}_{z_i}^* = 2G_{z_i}^{-1}\big(f_i(\bar{x}_i) + \sigma_{z_i} + \hat{h}_{c_i} + \zeta_i - \dot{\hat{\alpha}}_{i-1}\big) \tag{21}$$

where the switching variable σ_{z_i}, M_{z_i} (that is included in σ_{z_i} and includes control parameters κ_{i1} and κ_{i2}), and $\hat{h}_{a_i}$ can be designed following the forms defined in (13)-(15) in the first backstepping step by replacing s with z_i. ζ_i can be expressed as

$$\zeta_i = \begin{cases} g_s s & i = 2 \\ g_{i-1} z_{i-1} & i = 3, ..., n-1 \end{cases} \tag{22}$$

Using (20) and (21) to estimate the corresponding HJB function $H\left(z_i, \alpha_i^*, J_{z_i}^*\right)$ as

$$\begin{aligned} \hat{H}\left(z_i, \hat{\alpha}_i^*, \hat{J}_{z_i}^*\right) = z_i^T K_{z_i} z_i + \hat{\alpha}_i^{*T} K_{\alpha_i} \hat{\alpha}_i^* + \hat{J}_{z_i}^T \\ \times \left(- \sigma_{z_i} - \hat{h}_{a_i} - \zeta_i\right) \end{aligned} \tag{23}$$

Step n: Define the tracking error for subsystem z_n as $z_n = x_n - u$. The time derivative of error dynamics of z_n is

$$\dot{z}_n = f_n(\bar{x}_n) + g_n(\bar{x}_n)u - \dot{\alpha}_{n-1} \tag{24}$$

To formulate the optimization problem in subsystem z_n with the optimized system control u^*, design optimal cost function as

$$J^*(z_n) = \int_t^\infty (z_n^T K_{z_n} z_n + u^{*T} K_u u^*) \mathrm{d}\tau, \tag{25}$$

where $K_{z_n} = \mathrm{diag}(k_{z_{n1}}, k_{z_{n2}}, ..., k_{z_{nm}})$ and $K_u = \mathrm{diag}(k_{u_1}, k_{u_2}, ..., k_{u_m})$ are positive definite matrices.

Following the same way in the previous subsystems, the estimation of the optimized fixed TSC system control $\hat{u}^*$ and the optimal gradient term of cost function $\hat{J}_{z_n}^*$ in the subsystem z_n can be expressed as

$$\hat{u}^* = g_n^{-1}(\bar{x}_n)\big(- f_n(\bar{x}_n) - \sigma_{z_n} - \hat{h}_{a_n} - g_{n-1} z_{n-1} + \dot{\hat{\alpha}}_{n-1}\big) \tag{26}$$

$$\hat{J}_{z_n}^* = 2G_{z_n}^{-1}\big(f_n(\bar{x}_n) + \sigma_{z_n} + \hat{h}_{c_n} + g_{n-1} z_{n-1} - \dot{\hat{\alpha}}_{n-1}\big) \tag{27}$$

where the switching variable σ_{z_n}, M_{z_n} (that is included in σ_{z_n} and includes control parameters κ_{n1} and κ_{n2}), and $\hat{h}_{a_n}$ can be designed following the forms defined in (13)-(15) in the first backstepping step by replacing s with z_n.

Using (26) and (27) to estimate the corresponding HJB function $H\left(z_n, u^*, J_{z_n}^*\right)$ as

$$\begin{aligned} \hat{H}\left(z_n, \hat{u}^*, \hat{J}_{z_n}^*\right) = z_n^T K_{z_n} z_n + \hat{u}^{*T} K_u \hat{u}^* + \hat{J}_{z_n}^{*T} \\ \times \left(- \sigma_{z_n} - \hat{h}_{a_n} - g_{n-1} z_{n-1}\right) \end{aligned} \tag{28}$$

The learning objective for subsystem z_2 to z_n is similar to subsystem s, the learning parts $\hat{h}_{a_i}$ and $\hat{h}_{c_i}$ $(i = 2, .., n)$ are optimized and the parameters $\hat{b}_{a_i}$ and $\hat{b}_{c_i}$ are constrained in a specific range to find the approximated solution of the optimal system control. The constraints on the learnable region of parameters $\hat{b}_{a_i}$ and $\hat{b}_{c_i}$ will be analyzed with the Lyapunov synthesis.

C. Reinforcement Learning Algorithm Design

In this section, the iterative algorithm that contains policy evaluation and policy improvement procedures is designed to realize the learning objective, which optimizes the overall system control with attained FTSS property. The procedure of the iterative algorithm is established under the actor-critic framework as shown in Fig.1. To realize this process that solves the rather complicated HJB equation in each backstepping step, the evaluation signals are defined with RL methodology and utilized for policy improvement. The residual signals are defined as follows for policy evaluation, and we take the subsystem s for example

$$\tilde{e}_{\hat{H}_s} := \hat{H}_s - H_s, \tag{29}$$

$$\tilde{e}_{\hat{\alpha}_1^*} := \hat{\alpha}_1^* + \frac{1}{2} K_{\alpha_1}^{-1} g_s^T(\bar{x}_1) \hat{J}_s^*, \tag{30}$$

The updating step can be described as updating the estimation of the optimal gradient term $\hat{J}_s^*$ towards the HJB equation, then updating the estimation of the optimized policy $\hat{\alpha}_1^*$ towards the extremum point of HJB equation.

Rewrite the residual error $\tilde{e}_{\hat{H}_s}$ in vector form as $\tilde{e}_{\hat{\mathbf{H}}_s} = [\tilde{e}_{\hat{\mathbf{H}}_{s_1}}, ..., \tilde{e}_{\hat{\mathbf{H}}_{s_m}}]^T$ to consider possible different calculation ranges of each element, in which the i-th element $\tilde{e}_{\hat{\mathbf{H}}_{s_i}} = k_{si} s_i^2 + k_{\alpha_{1i}} \hat{\alpha}_{1i}^{*2} + \hat{J}_{si}^*(-\sigma_{si} - \hat{h}_{a_{1i}})$, and $\sigma_s = [\sigma_{s1}, \sigma_{s2}, ..., \sigma_{sm}]^T$, $\hat{h}_{a_1} = [\hat{h}_{a_{11}}, \hat{h}_{a_{12}}, ..., \hat{h}_{a_{1m}}]^T$. The constraints on the $\hat{b}_{a_1}$ and $\hat{b}_{c_1}$ satisfy

$$\Phi_{a_1} \triangleq \{\hat{b}_{a_1} \in \mathbb{R} \mid \psi_j(\hat{b}_{a_1}) \leq 0,\ j = 1, 2\}, \tag{31}$$

$$\Phi_{c_1} \triangleq \{\hat{b}_{c_1} \in \mathbb{R} \mid \chi_j(\hat{b}_{c_1}) \leq 0\,,\ j = 1, 2\}\,, \tag{32}$$

where $\psi_j(\hat{b}_{a_1})$ is the j-th element in $\psi(\hat{b}_{a_1})$ and $\chi_j(\hat{b}_{c_1})$ is the j-th element in $\chi(\hat{b}_{c_1})$. Denote $M_{1,\min} := \kappa_{1,\min} s^T \mathrm{sig}_{\mathrm{n}}^{p_1}(s) + \kappa_{2,\min} s^T \mathrm{sig}_{\mathrm{n}}^{p_2}(s)$ and $M_{1,\max} := \kappa_{1,\max} s^T \mathrm{sig}_{\mathrm{n}}^{p_1}(s) + \kappa_{2,\max} s^T \mathrm{sig}_{\mathrm{n}}^{p_2}(s)$ with appropriate constants $\kappa_{1,\min} > 0$, $\kappa_{2,\min} > 0$, $\kappa_{1,\max} > 0$, $\kappa_{2,\max} > 0$.

$$\psi(\hat{b}_{a_1}) = \begin{bmatrix} -M_s - \hat{h}_{a_1} + M_{1,\min} \\ M_s + \hat{h}_{a_1} - M_{1,\max} \end{bmatrix}, \tag{33}$$

$$\chi(\hat{b}_{c_1}) = \begin{bmatrix} -M_s - \hat{h}_{c_1} + M_{1,\min} \\ M_s + \hat{h}_{c_1} - M_{1,\max} \end{bmatrix}, \tag{34}$$

The error signals for the updating of $\hat{b}_{a_1}$ in the estimation of optimal system control input $\hat{\alpha}_1^*$ and $\hat{b}_{c_1}$ in estimation of

optimal gradient term $\hat{J}_s^*$ are designed as

$$e_{\hat{b}_{c_1}} = \mathrm{Proj}(-\gamma_{c_1}\Gamma_{c_1}\tilde{e}_{\hat{\mathbf{H}}_s}) \tag{35}$$

$$e_{\hat{b}_{a_1}} = \mathrm{Proj}(-\Theta_{a_1}\tilde{E}_{a_1}) \tag{36}$$

where

$$\Gamma_{c_1} = \left[\frac{\partial\tilde{e}_{\hat{\mathbf{H}}_{s_1}}}{\partial\hat{b}_{c_1}}, ..., \frac{\partial\tilde{e}_{\hat{\mathbf{H}}_{s_m}}}{\partial\hat{b}_{c_1}}\right] \in \mathbb{R}^{1\times m}, \Theta_{a_1} = [\gamma_{c_1}, \gamma_{a_1}],$$

$\gamma_{\{c_1,a_1\}} \in \mathbb{R}^+, \tilde{E}_{a_1} = [\Gamma_{c_1}^T\tilde{e}_{\hat{\mathbf{H}}_s}, \tilde{e}_{b_1}]^T \in \mathbb{R}^{2\times 1}, \tilde{e}_{b_1} = \hat{b}_{a_1} - \hat{b}_{c_1},$

The updating equation with the consideration of constraints is designed with the gradient projection method [14], which is used to confine all elements of the learning part vector to be inside the constraint sets Φ_{c_1} and Φ_{a_1}. If some updating direction of the learning part vectors are going to be outside the constraint sets Φ_{c_1} and Φ_{a_1}, these elements are bounded by the projection operator after updating. If the elements of the updating direction of the learning part vectors are inside the constraint sets Φ_{c_1} and Φ_{a_1}, these elements will maintain their value.

The property of the iterative algorithm can be stated in the following main result.

Theorem 1: *Consider the strict-feedback system governed by (1), the prescribed performance is satisfied with state transformation, then design and approximate the optimal system control inputs with the TSC law form as (12) in the first backstepping step, as (20) in $z_i (i = 2, ..., n-1)$ backstpping step and as (26) in z_n backstepping step, and approximate the optimal value of the term J_s^* as (16) in z_1 subsystem, $J_{z_i}^*$ as (21) in z_i subsystem, $J_{z_n}^*$ as (27) in z_n subsystem. With the same design method of updating the equation on the unknown part shown in (35)-(36), the error signals for each backstepping step can be designed as*

$$e_{\hat{b}_{c_i}} = \mathrm{Proj}(-\gamma_{c_i}\Gamma_{c_i}\tilde{e}_{\hat{\mathbf{H}}_{(\cdot)}}) \tag{37}$$

$$e_{\hat{b}_{a_i}} = \mathrm{Proj}(-\Theta_{a_i}\tilde{E}_{a_i}) \tag{38}$$

with γ_{c_i}, Γ_{c_i}, Θ_{a_i}, $\tilde{E}_{a_i}$ that follows the design method in the first backstepping step, $\tilde{e}_{\hat{\mathbf{H}}_{(\cdot)}}$ is $\tilde{e}_{\hat{\mathbf{H}}_s}$ or $\tilde{e}_{\hat{\mathbf{H}}_{z_i}}$. The estimation of the optimal system control input and the associated gradient term of the cost function in each subsystem can eventually satisfy Bellman optimality via learning iteratively. The closed-loop system is FTSS, and the time-synchronized optimized control performance can be guaranteed.

Proof: The main parts of proofing will be stated as follows: Firstly, consider the Lyapunov function candidate $V_z = \frac{1}{2}s^Ts + \frac{1}{2}\sum_{i=2}^n z_i^Tz_i = V_s + \sum_{i=2}^n V_{z_i}$. We prove that the state-variables s and z_i converge in fixed time and synchronously. Calculating the time derivative of V_z and substituting the control laws in all the subsystems

$$\begin{aligned}\dot{V}_z =& \ s^T\big(-\sigma_s - \hat{h}_{a_1} + g_sz_2\big) + z_2^T\big(-\sigma_{z_2} - \hat{h}_{a_2} - g_ss \\ &+ g_2z_3\big) + \sum_{i=3}^{n-1} z_i^T\big(-\sigma_{z_i} - \hat{h}_{a_i} - g_{i-1}z_{i-1} + g_iz_{i+1}\big) \\ &+ z_n^T\big(-\sigma_{z_n} - \hat{h}_{a_n} - g_{n-1}z_{n-1}\big) \\ =& \ s^T\big(-\sigma_s - \hat{h}_{a_1}\big) + \sum_{i=2}^n z_i^T\big(-\sigma_{z_i} - \hat{h}_{a_i}\big) \\ =& -2^{\frac{1+p_1}{2}}(\kappa_{11} + \hat{b}_{a_1})V_s^{\frac{1+p_1}{2}} - 2^{\frac{1+p_2}{2}}\kappa_{12}V_s^{\frac{1+p_2}{2}} \\ &+ \sum_{i=2}^n\big(-2^{\frac{1+p_1}{2}}(\kappa_{i1} + \hat{b}_{a_i})V_{z_i}^{\frac{1+p_1}{2}} - 2^{\frac{1+p_2}{2}}\kappa_{i2}V_{z_i}^{\frac{1+p_2}{2}}\big)\end{aligned}$$

With this development, the convergence of state-variables and the finite setting time are therefore guaranteed.

The (37) and (38) are derived with the subsequent learning procedures of policy evaluation and policy improvement under the generalized policy iteration frame. The overall system control is optimized by obtaining the solution of the HJB function at each subsystem, which means the defined policy evaluation signals are expected to $\tilde{e}_{\hat{H}_s}, \tilde{e}_{\hat{H}_{z_i}} \to 0$ and $\tilde{e}_{\hat{\alpha}_i^*}, \tilde{e}_{\hat{u}^*} \to 0$. The policy improvement updates the estimation of the optimal gradient term $\hat{J}_s^*$, $\hat{J}_{z_i}^*$ that pushes the estimation $\hat{H}_s \to H_s = 0$, $\hat{H}_{z_i} \to H_{z_i} = 0$; and the updating step on the estimation of the optimized control $\hat{\alpha}_i^*$, $\hat{u}^*$ towards/holds as the extremum point of current residual error $\tilde{e}_{\hat{H}_s} = 0$ and $\tilde{e}_{\hat{H}_{z_i}} = 0$, subsequently. Take subsystem s for example, $\tilde{e}_{\hat{\alpha}_1^*}$ can be calculated as

$$\tilde{e}_{\hat{\alpha}_1^*} = \frac{\hat{h}_{a_1} - \hat{h}_{c_1}}{g_s(\bar{x}_1)} = \frac{\hat{b}_{a_1} - \hat{b}_{c_1}}{g_s(\bar{x}_1)}\mathrm{sig}^{p_1}(s) \tag{39}$$

Design Lyapunov function candidates V_c and V_a that are equivalent to $\tilde{e}_{\hat{H}_s} = 0$, $\tilde{e}_{\hat{H}_{z_i}} = 0$ and $\tilde{e}_{\hat{\alpha}_1^*} = 0$, $\tilde{e}_{\hat{u}^*} = 0$, which can be expressed as the following form

$$V_o = V_c + V_a, \tag{40}$$

where

$$V_c = \frac{1}{2}\tilde{e}_{\hat{\mathbf{H}}_s}^T\tilde{e}_{\hat{\mathbf{H}}_s} + \sum_{i=2}^n\frac{1}{2}\tilde{e}_{\hat{\mathbf{H}}_{z_i}}^T\tilde{e}_{\hat{\mathbf{H}}_{z_i}}, \tag{41}$$

$$V_a = \sum_{i=1}^n\frac{1}{2}\left(\hat{b}_{a_i} - \hat{b}_{c_i}\right)^2. \tag{42}$$

The time derivative of V_c and V_a with (37) and (38) is

$$\begin{aligned}\dot{V}_c =& \tilde{e}_{\hat{\mathbf{H}}_s}^T\Gamma_{c_1}e_{\hat{b}_{c_1}} + \sum_{i=2}^n\tilde{e}_{\hat{\mathbf{H}}_{z_i}}^T\Gamma_{c_i}e_{\hat{b}_{c_i}} \le -\gamma_{c_1}\tilde{e}_{\hat{\mathbf{H}}_s}^T\Gamma_{c_1}\Gamma_{c_1}^T\tilde{e}_{\hat{\mathbf{H}}_s} \\ &- \sum_{i=2}^n\gamma_{c_i}\tilde{e}_{\hat{\mathbf{H}}_{z_i}}^T\Gamma_{c_i}\Gamma_{c_i}^T\tilde{e}_{\hat{\mathbf{H}}_{z_i}} \le 0\end{aligned} \tag{43}$$

$$\begin{aligned}\dot{V}_a =& \sum_{i=1}^n\left(\frac{\partial V_a}{\partial\hat{b}_{a_i}}e_{\hat{b}_{a_i}} + \frac{\partial V_a}{\partial\hat{b}_{c_i}}e_{\hat{b}_{c_i}}\right) = \sum_{i=1}^n\frac{\partial V_a}{\partial\hat{b}_{a_i}}\left(e_{\hat{b}_{a_i}} - e_{\hat{b}_{c_i}}\right) \\ \le& -\gamma_{a_i}\sum_{i=1}^n\left(\hat{b}_{a_i} - \hat{b}_{c_i}\right)^2 \le 0\end{aligned} \tag{44}$$

With this, and with the gradient projection method, the closed-loop system is FTSS, and the time-synchronized optimized control performance can be guaranteed. The required proof is thus completed. ■

Following is a pertinent remark related to the proof above:

Remark 3: *The settings of $\kappa_{i1,\min}$, $\kappa_{i2,\min}$, $\kappa_{i1,\max}$, $\kappa_{i2,\max}$ in $M_{i,\min}$ $M_{i,\max}$ decide the range of the learning*

parts $\hat{b}_{a_i}$ and $\hat{b}_{c_i}$, which finally influence the property of optimization and convergence with the designed system control. The constraint sets on $\hat{b}_{a_i}$, $\hat{b}_{c_i}$ ensure that the control stays in the admissible policy region that satisfies

$$-D_{Z,1} \leq \dot{V}_z \leq D_{Z,2}$$

where $D_{Z,1}$ and $D_{Z,2}$ are designed boundaries when κ_{i1} and κ_{i2} are set as the min and max values, respectively. Thus the work here has the outcome of time-synchronized convergence which surpasses the usual typical developments in this class of problems considered.

III. Simulation

In this section, the autonomous vehicle control problem is considered to verify the effectiveness of the proposed synchronized optimization algorithm. The dynamics of each vehicle i-th in the platoon as the following form [19]

$$\dot{p}_i = v_i,\ \dot{v}_i = a_i,\ \dot{a}_i = -1/\tau_i a_i + u_i, \tag{45}$$

where p_i, v_i and a_i denote the position, velocity and acceleration of i-th vehicle, respectively; τ_i denotes the time constant parameter; u_i is the control input of the vehicle i.

Design $y_{d,i} = p_{i-1} - d_{s,i}$. $d_{s,i}$ is the expected spacing of the vehicles. The initial positions are set as -1m, -13m, -25m, -36m, -46m, and -56m, respectively. The initial velocities are 10 m/s, and the initial accelerations are zero. In the time-synchronized control (TSC) and time-synchronized optimization control (TSOC), control parameters are set as the same: κ_{11}=10, κ_{21}=5, κ_{31}=5, κ_{12}=2, κ_{22}=2, κ_{32}=2, $p_1 = 1.1$, $p_2 = 0.65$. In TSOC, the k_{1s}, k_{2s}, k_{3s} increase with 0.01 from 1 to 10 at each 0.01s time-step. k_{1c}, k_{2c}, k_{3c} decrease with 0.01 from 3 to 0.1 at each 0.01s time-step. The parameters $\gamma_c = 0.018$ and $\gamma_a = 0.01$. The control performance of the TSC and TSOC are compared as shown in Fig. 2. The trajectories of state-variables z_1, z_2 and z_3 are shown in Figs. 2 (a)-(e), (f)-(g) and (h)-(i), respectively. As shown in Figs. 2 (a)-(e), the state-variables s_1 to s_5 are constrained within the prescribed region. The control performance is better with the proposed time-synchronized optimization control with less fluctuation and control inputs. Time-synchronized optimization control can optimize the system control in a manner that each state-variable element converges and arrives at the origin synchronously in each backstepping step, which satisfies the requirement of high-precision control problems.

IV. Conclusion

Based on the methodology of time-synchronized optimized control, this paper has proposed and investigated synchronized optimization with prescribed performance for the strict-feedback system. To effectively achieve such high-performance control, prescribed performance control and optimized backstepping, and adaptive dynamic programming are utilized to establish the learning framework thein the norm-normalized sign function is appropriately incorporated in each backstepping subsystem. With the decomposition

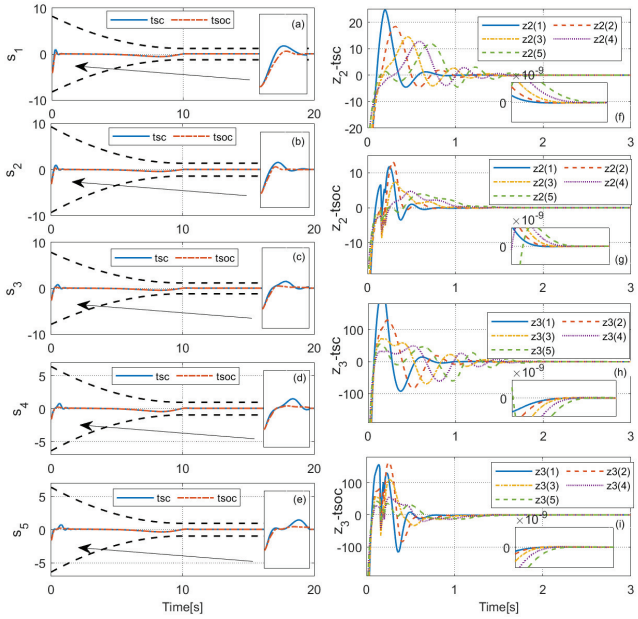

Fig. 2. Comparison of the trajectories of the state-variables z_1, z_2, and z_3 with TSC and TSOC.

design, the learning objective can be transformed to adaptively explore the optimal control parameter in the admissible policy region, which is analyzed via Lyapunov synthesis. The effectiveness of the proposed method has been verified with the vehicle platoon problem. The results show that the overall system control is optimized and prescribed performance is ensured while attaining special time-synchronized properties during the learning process. In the future, we will apply this method to more complex and abundant synchronized optimization problems among various types of unmanned vehicle systems.

References

[1] Y. Chen, Y.-J. Liu, S. Tong, C. P. Chen, and L. Liu, "Adaptive intelligent controller design-based iss modular approach for uncertain nonlinear systems with time-varying full-state constraints," *IEEE Transactions on Artificial Intelligence*, vol. 2, no. 4, pp. 352–361, 2021.

[2] X. Liang, Y. Zhang, S. S. Ge, B. V. E. How, and D. Li, "Dynamic control for lng carrier with output constraints," *IET Control Theory & Applications*, vol. 16, no. 7, pp. 729–740, 2022.

[3] Y. Zhang, X. Liang, D. Li, S. S. Ge, B. Gao, H. Chen, and T. H. Lee, "Barrier lyapunov function-based safe reinforcement learning for autonomous vehicles with optimized backstepping," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[4] R. Ji, S. S. Ge, and D. Li, "Saturation-tolerant prescribed control for nonlinear systems with unknown control directions and external disturbances," *IEEE Transactions on Cybernetics*, 2023.

[5] H. Chu, D. Meng, S. Huang, M. Tian, J. Zhang, B. Gao, and H. Chen, "Autonomous high-speed overtaking of intelligent chassis using fast iterative model predictive control," *IEEE Transactions on Transportation Electrification*, 2023.

[6] D. Li, S. S. Ge, and T. H. Lee, "Time-synchronized control for disturbed systems," in *Time-Synchronized Control: Analysis and Design*. Springer, 2022, pp. 61–99.

[7] B. Gao *et al.*, "Gear ratio optimization and shift control of 2-speed i-amt in electric vehicle," *Mechanical Systems and Signal Processing*, vol. 50, pp. 615–631, 2015.

[8] Y. Zhang, B. Gao, L. Guo, H. Guo, and H. Chen, "Adaptive decision-making for automated vehicles under roundabout scenarios using optimization embedded reinforcement learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 12, pp. 5526–5538, 2020.

[9] Y. Zhang, X. Liang, S. S. Ge, B. Gao, and H. Chen, "Manoeuver planning, synchronized optimization and boundary motion control for autonomous vehicles under cut-in scenarios," *Nonlinear Dynamics*, vol. 111, no. 8, pp. 6923–6939, 2023.

[10] S. S. Ge, T. H. Lee, and C. J. Harris, *Adaptive neural network control of robotic manipulators*. World Scientific, 1998, vol. 19.

[11] D. Li, H. Yu, K. P. Tee, Y. Wu, S. S. Ge, and T. H. Lee, "On time-synchronized stability and control," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021.

[12] S. Bhasin, R. Kamalapurkar, M. Johnson, K. G. Vamvoudakis, F. L. Lewis, and W. E. Dixon, "A novel actor–critic–identifier architecture for approximate optimal control of uncertain nonlinear systems," *Automatica*, vol. 49, no. 1, pp. 82–92, 2013.

[13] G. Wen, C. P. Chen, and S. S. Ge, "Simplified optimized backstepping control for a class of nonlinear strict-feedback systems with unknown dynamic functions," *IEEE Transactions on Cybernetics*, vol. 51, no. 9, pp. 4567–4580, 2020.

[14] Y. Zhang, X. Liang, D. Li, S. S. Ge, B. Gao, H. Chen, and T. H. Lee, "Adaptive safe reinforcement learning with full-state constraints and constrained adaptation for autonomous vehicles," *IEEE Transactions on Cybernetics*, 2023.

[15] H. Ma, Q. Zhou, H. Li, and R. Lu, "Adaptive prescribed performance control of a flexible-joint robotic manipulator with dynamic uncertainties," *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 12 905–12 915, 2021.

[16] Z. Gao, Y. Zhang, and G. Guo, "Prescribed-time control of vehicular platoons based on a disturbance observer," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 9, pp. 3789–3793, 2022.

[17] K. G. Vamvoudakis and F. L. Lewis, "Online actor–critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.

[18] H. Modares, F. L. Lewis, and M.-B. Naghibi-Sistani, "Integral reinforcement learning and experience replay for adaptive optimal control of partially-unknown constrained-input continuous-time systems," *Automatica*, vol. 50, no. 1, pp. 193–202, 2014.

[19] H. Chu, L. Guo, Y. Yan, B. Gao, and H. Chen, "Self-learning optimal cruise control based on individual car-following style," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 10, pp. 6622–6633, 2020.