A NEW ITERATION APPROACH TO SOLVE A CLASS OF 
FINITE-HORIZON CONTINUOUS-TIME NONAFFINE 
NONLINEAR ZERO-SUM GAME

XIN ZHANG¹,², HUAGUANG ZHANG¹,², XINGYUAN WANG³ AND YANHONG LUO¹,²

¹College of Information Science and Engineering 
Northeastern University 
No. 11, Lane 3, Wenhua Road, Heping District, Shenyang 110819, P. R. China
²Key Laboratory of Integrated Automation for the Process Industry 
Ministry of Education 
Shenyang 110004, P. R. China 
jackie_zx@yahoo.com.cn; hgzhang@ieee.org; neuluo@gmail.com
³School of Electronic and Information Engineering 
Dalian University of Technology 
No. 2, Linggong Road, Ganjingzi District, Dalian 116024, P. R. China 
wangxy@dlut.edu.cn

Received September 2009; revised February 2010

ABSTRACT. In this paper, a new iteration approach is derived to solve the optimal strategies for quadratic zero-sum game of finite-horizon continuous-time nonaffine nonlinear system. Through iteration algorithm between two sequences which are a sequence of state trajectories of linear quadratic zero-sum games and a sequence of corresponding Riccati differential equations, the optimal strategies for the nonaffine nonlinear zero-sum game are given. Under very mild conditions of local Lipschitz continuity, the convergence of approximating linear time-varying sequences is proved. A numerical example is given to demonstrate the convergence and effectiveness of the proposed approach.

Keywords: Zero-sum game, Nonaffine nonlinear system, Riccati equation, Approximation theory, Iteration algorithm

1. Introduction. Game theory is concerned with the study of decision making in situations where two or more rational opponents are involved under conditions of conflicting interests, which has been widely investigated by many authors [3-8,10,21-24]. The two-player zero-sum game with a quadratic performance index plays an important role in the game theory. One player tries to minimize the performance index while the other tries to maximize it.

Much work has been contributed to this minimax problem under the frameworks of linear quadratic zero-sum games [1,2,18-20,28]. In [1], Al-Tamimi et al. applied the heuristic dynamic programming and dual heuristic dynamic programming structures to solve a discrete-time linear quadratic zero-sum game problem in which the state and action spaces are continuous. Then, they designed the optimal strategies for the discrete-time linear quadratic zero-sum game without knowing the system dynamical matrices by the model-free Q-learning approach [2]. A class of continuous-time affine nonlinear quadratic zero-sum game problem was researched by Wei et al. in [21]. Abu-Khalaf et al. studied the affine nonlinear zero-sum game problem which is converted from solving $H_{\infty}$ control of nonlinear systems with constrained input system by policy iteration in [16] and used neural networks to solve it in [17]. It is worthy of mentioning that most of the above
discussions are focused on the linear or affine nonlinear zero-sum game problems. However, many applications of practical zero-sum game have the nonlinear structure which is nonaffine in control inputs. To the best knowledge of the authors, it is rare to solve the optimal strategies for the nonaffine nonlinear zero-sum game, which is the main motivation of this paper.

For the general nonaffine nonlinear control system, the direct feedback linearization is difficult because the input does not appear linearly. Therefore, many researches have introduced neural networks, fuzzy logic and other approximation approaches to design the controller [9-15, 25-27]. Furthermore, in the nonaffine nonlinear zero-sum game problem, it is very difficult to solve the Hamilton-Jacobi-Issue (HJI) equation which is a generalization of the Hamilton-Jacobi-Bellman (HJB) partial differential equation.

In this paper, we transform the nonaffine nonlinear zero-sum game into an equivalent sequence of linear quadratic zero-sum games. This method is to change the state trajectory of the nonaffine nonlinear quadratic zero-sum game into a sequence of linear time-varying equations. Moreover, the corresponding HJI equation is transformed into a sequence of Riccati differential equations with constant coefficients. It is simpler and more effectively to compute the Riccati differential equations than to solve the HJI equation. Under very mild conditions of local Lipschitz continuity, it can be proved that the sequence of linear quadratic zero-sum game can arbitrarily closely approximate to the nonaffine nonlinear zero-sum game by iteration sequentially.

The remainder of this paper is organized as follows. Section 2 introduces a class of finite-horizon continuous-time nonaffine nonlinear zero-sum game that we want to research in this paper. In Section 3, an iteration algorithm is presented. In Section 4, the convergence proof of iteration algorithm is given and the design procedure of optimal strategies is summarized. In Section 5, an example is given to demonstrate the convergence and effectiveness of the proposed approach. Finally, some conclusions are given in Section 6.

2. Problem Statement and Preliminaries. Considered a continuous-time nonaffine nonlinear zero-sum game described by the state equation

\[ \dot{x}(t) = f(x(t), u(t), w(t)), \quad x(t_0) = x_0 \]

with the finite-horizon performance index function

\[
V(x_0, u, w) = \frac{1}{2} x^T(t_f) F(x(t_f)) x(t_f) \\
+ \frac{1}{2} \int_{t_0}^{t_f} \left[ x^T(t) Q(x(t)) x(t) + u^T(t) R(x(t)) u(t) - w^T(t) S(x(t)) w(t) \right] dt
\]

where, \( x(t) \in \mathbb{R}^n \) is the state, \( x(t_0) \in \mathbb{R}^n \) is the initial state, \( t_f \) is the terminal time, the control input \( u(t) \) takes values in convex and compact set \( U \subset \mathbb{R}^m \), and \( w(t) \) takes values in convex and compact set \( W \subset \mathbb{R}^m \). \( u(t) \) seeks to minimize the performance index function \( V(x_0, u, w) \) while \( w(t) \) seeks to maximize it. The state-dependent weighting matrices \( F(x(t)), Q(x(t)), R(x(t)), S(x(t)) \) are with suitable dimensions and \( F(x(t)) \geq 0, Q(x(t)) \geq 0, R(x(t)) > 0, S(x(t)) > 0 \). In this paper, \( x(t), u(t) \) and \( w(t) \) sometimes are described as \( x, u \) and \( w \) for brevity. Our objective is to find the optimal strategies for the above nonaffine nonlinear zero-sum game.

In the nonaffine nonlinear zero-sum game problem, nonlinear functions are implicit function with respect to controller inputs. It is very hard to obtain the optimal strategies satisfying (1) and (2). For practical purpose one may just as well be interested in finding
a near optimal or an approximate optimal policy. Therefore, we present a iteration algorithm to deal with this problem. In the next section, the nonaffine nonlinear zero-sum game is transformed into an equivalent sequence of linear quadratic zero-sum games which can use the linear quadratic zero-sum game theory directly.

3. An Iteration Algorithm for Nonaffine Nonlinear Zero-Sum Game. Using a factored form to represent the system (1), we can get

\[ \dot{x}(t) = A(x(t))x(t) + B(x(t), u(t))u(t) + E(x(t), w(t))w(t), \quad x(t_0) = x_0 \]  

(3)

where \( A: \mathbb{R}^n \to \mathbb{R}^{n \times n} \) is a nonlinear matrix-valued function of \( x \), \( B: \mathbb{R}^n \times \mathbb{R}^{m_1} \to \mathbb{R}^{n \times m_1} \) is a nonlinear matrix-valued function of both the state \( x \) and control input \( u \), \( E: \mathbb{R}^n \times \mathbb{R}^{m_2} \to \mathbb{R}^{n \times m_2} \) is a nonlinear matrix-valued function of both the state \( x \) and control input \( w \).

We use the following sequence of linear time-varying differential equations to approximate the state Equation (3)

\[ \dot{x}^i(t) = A \left( x^{i-1}(t) \right) x^i(t) + B \left( x^{i-1}(t), u^{i-1}(t) \right) u^i(t) + E \left( x^{i-1}(t), w^{i-1}(t) \right) w^i(t) \]

\[ x^i(t_0) = x_0, \quad i \geq 0 \]  

(4)

with the corresponding quadratic performance index function

\[ V^i(x_0, u, w) = \frac{1}{2} x^{iT}(t_f)F \left( x^{i-1}(t_f) \right)x^i(t_f) + \frac{1}{2} \int_{t_0}^{t_f} \left[ x^{iT}(t)Q \left( x^{i-1}(t) \right)x^i(t) + u^{iT}(t)R \left( x^{i-1}(t) \right)u^i(t) \right. \]

\[ \left. -w^{iT}(t)S \left( x^{i-1}(t) \right)w^i(t) \right] dt, \quad i \geq 0. \]  

(5)

where the superscript \([i]\) represents the iteration process. For the first approximation \( i = 0 \), we assume that the initial values \( x^{[0]-1}(t) = x_0 \), \( u^{[0]-1}(t) = 0 \) and \( w^{[0]-1}(t) = 0 \). Obviously, for the \( i \)th iteration, \( A \left( x^{[i-1]}(t) \right), B \left( x^{[i-1]}(t), u^{[i-1]}(t) \right), E \left( x^{[i-1]}(t), u^{[i-1]}(t) \right) \), \( F \left( x^{[i-1]}(t_f) \right), Q \left( x^{[i-1]}(t) \right), R \left( x^{[i-1]}(t) \right) \) and \( S \left( x^{[i-1]}(t) \right) \) are time-varying functions which don’t depend on \( x^{[i]}(t) \), \( u^{[i]}(t) \) and \( w^{[i]}(t) \). Hence, each approximation problem in (4) and (5) is the linear quadratic zero-sum game problem which can be solved by the existing classical linear quadratic zero-sum game theory.

The corresponding Riccati differential equation of each linear quadratic zero-sum game can be expressed as

\[ \dot{P}^i(t) = -Q \left( x^{[i-1]}(t) \right) - P^i(t)A \left( x^{[i-1]}(t) \right) - A^T \left( x^{[i-1]}(t) \right)P^i(t) \]

\[ + P^i(t)B \left( x^{[i-1]}(t), u^{[i-1]}(t) \right)R^{-1} \left( x^{[i-1]}(t) \right)B^T \left( x^{[i-1]}(t), u^{[i-1]}(t) \right) \]

\[ -E \left( x^{[i-1]}(t), w^{[i-1]}(t) \right)S^{-1} \left( x^{[i-1]}(t) \right)E^T \left( x^{[i-1]}(t), w^{[i-1]}(t) \right)P^i(t), \quad i \geq 0 \]  

(6)

where \( P^i \in \mathbb{R}^{n \times n} \) is real, symmetric and nonnegative definite matrix.

**Assumption 3.1.** It is assumed that \( S \left( x^{[i-1]}(t) \right) > \hat{S}^i \), where the threshold value \( \hat{S}^i \) is defined as \( \hat{S}^i = \inf \{ S^i(t) > 0, \ (6) \text{ does not have a conjugate point on } [0, t_f] \} \).

If Assumption 3.1 is satisfied, the game admits the optimal strategies given by

\[ u^i(t) = -R^{-1} \left( x^{[i-1]}(t) \right)B^T \left( x^{[i-1]}(t), u^{[i-1]}(t) \right)P^i(t)x^{i}(t), \quad i \geq 0 \]  

\[ w^i(t) = S^{-1} \left( x^{[i-1]}(t) \right)E^T \left( x^{[i-1]}(t), w^{[i-1]}(t) \right)P^i(t)x^{[i]}(t), \quad i \geq 0 \]  

(7)
where \( x^{[i]}(t) \) is the corresponding optimal state trajectory generated by

\[
\dot{x}^{[i]}(t) = \left[ A \left( x^{[i-1]}(t) \right) - B \left( x^{[i-1]}(t), u^{[i-1]}(t) \right) R^{-1} \left( x^{[i-1]}(t) \right) B^T \left( x^{[i-1]}(t), u^{[i-1]}(t) \right) P^{[i]}(t) \right. \\
\left. + E \left( x^{[i-1]}(t), w^{[i-1]}(t) \right) S^{-1} \left( x^{[i-1]}(t) \right) E^T \left( x^{[i-1]}(t), w^{[i-1]}(t) \right) P^{[i]}(t) \right] x^{[i]}(t)
\]

\[
x^{[i]}(t_0) = x_0.
\]

By using the iteration between sequences (6) and (8) sequentially, the limit of the solution of approximating sequence (4) will converge to the unique solution of system (1), and the sequences of optimal strategies (7) will converge, too. The convergence of iteration will be analysed in the next section. Notice that the factored form in (3) need not be unique. The approximating linear time-varying sequences will converge whatever the representation of \( A(x(t)), B(x(t), u(t)) \) and \( E(x(t), w(t)) \).

**Remark 3.1.** For the fixed finite interval \([t_0, t_f]\), if \( S(x^{[i-1]}(t)) > \tilde{S}^{[i]} \), the Riccati differential Equation (6) has no conjugate point on \([t_0, t_f] \). It means that \( V^{[i]}(x_0, u, w) \) is strictly concave in \( w \). Otherwise, since \( V^{[i]}(x_0, u, w) \) is quadratic and \( R(x(t)) > 0 \), \( F(x(t)) \geq 0 \), \( Q(x(t)) \geq 0 \), it follows that \( V^{[i]}(x_0, u, w) \) is strictly convex in \( u \). Hence, the linear quadratic zero-sum game (4) with the performance index function (6) exists a unique saddle point which are the optimal strategies.

4. **Convergence Analysis.** In this section, the convergence of the algorithm described in Section 3 will be proved. It requires that

1. The sequence \( \{x^{[i]}(t)\} \) converges on \( C([t_0, t_f]; \mathbb{R}^n) \), which means that the limit of the solution of approximating sequence (4) converges to the unique solution of system (1).
2. The sequences of optimal strategies \( \{u^{[i]}(t)\} \) and \( \{w^{[i]}(t)\} \) converge on \( C([t_0, t_f]; \mathbb{R}^{m_1}) \) and \( C([t_0, t_f]; \mathbb{R}^{m_2}) \), respectively.

For simplicity, the approximating sequence (4) is rewritten as

\[
\dot{x}^{[i]}(t) = A(x^{[i-1]}(t))x^{[i]}(t) + H(x^{[i-1]}(t), u^{[i-1]}(t))x^{[i]}(t) + G(x^{[i-1]}(t), w^{[i-1]}(t))x^{[i]}(t)
\]

\[
x^{[i]}(t_0) = x_0, \quad i \geq 0
\]

where \( H(x^{[i-1]}, u^{[i-1]}) \triangleq -B(x^{[i-1]}(t), u^{[i-1]}(t))R^{-1}(x^{[i-1]}(t))B^T(x^{[i-1]}(t), u^{[i-1]}(t))P^{[i]}(t) \),

\( G(x^{[i-1]}, w^{[i-1]}) \triangleq E(x^{[i-1]}(t), w^{[i-1]}(t))S^{-1}(x^{[i-1]}(t))E^T(x^{[i-1]}(t), w^{[i-1]}(t))P^{[i]}(t) \).

The optimal strategies for the zero-sum game are rewritten as

\[
u^{[i]}(t) = C(x^{[i-1]}(t), w^{[i-1]}(t))x^{[i]}(t)
\]

\[
w^{[i]}(t) = D(x^{[i-1]}(t), w^{[i-1]}(t))x^{[i]}(t)
\]

\[
i \geq 0
\]

where \( C(x^{[i-1]}, u^{[i-1]}) \triangleq -R^{-1}(x^{[i-1]}(t))B^T(x^{[i-1]}(t), u^{[i-1]}(t))P^{[i]}(t), D(x^{[i-1]}, w^{[i-1]}) \triangleq S^{-1}(x^{[i-1]}(t))E^T(x^{[i-1]}(t), w^{[i-1]}(t))P^{[i]}(t) \).

**Assumption 4.1.** \( B(x, u), E(x, w), R^{-1}(x), S^{-1}(x), F(x) \) and \( Q(x) \) are bounded and Lipschitz continuous in their arguments \( x, u \) and \( w \), thus satisfying

\( B1 \) \( ||B(x, u)|| \leq b, \quad ||E(x, u)|| \leq e \)

\( B2 \) \( ||R^{-1}(x)|| \leq r, \quad ||S^{-1}(x)|| \leq s \)

\( B3 \) \( ||F(x)|| \leq f, \quad ||Q(x)|| \leq q \)

\( \forall x \in \mathbb{R}^n, \quad \forall u \in \mathbb{R}^{m_1}, \quad \forall w \in \mathbb{R}^{m_2}, \) and for finite positive numbers \( b, e, r, s, f \) and \( q \).

Define \( \Phi^{[i-1]}(t, t_0) \) is the transition matrix generated by \( A^{[i-1]}(t) \). It is well-known that

\[
||\Phi^{[i-1]}(t, t_0)|| \leq \exp \left[ \int_{t_0}^t \mu(A(x^{[i-1]}(\tau)))d\tau \right], \quad t \geq t_0
\]

(11)
where $\mu(A)$ is the measure of matrix $A$, $\mu(A) = \lim_{h \to 0+} \frac{\|I+hA\|_1}{h}$. We use the following Lemma to get an estimate for $\Phi^{(i-1)}(t, t_0) - \Phi^{(i-2)}(t, t_0)$.

**Lemma 4.1.** [25] Suppose that

(A1) $\mu(A(x)) \leq \mu_0$ for some constant $\mu_0$ and for all $x$

(A2) $\|A(x_1) - A(x_2)\| \leq \alpha \|x_1 - x_2\|$, $\forall x_1, x_2 \in \mathbb{R}^n$, $\alpha > 0$, then

$\Phi^{(i-1)}(t, t_0) - \Phi^{(i-2)}(t, t_0) \leq \alpha e^{\mu_0(t-t_0)}(t-t_0) \sup_{s \in [t_0, t]} \|x^{(i-1)}(s) - x^{(i-2)}(s)\|.$

The following Lemma is about the solution of the Riccati differential Equation (6), which is the basis for proving the convergence.

**Lemma 4.2.** Let Assumption 4.1 hold, the solution of Riccati differential Equation (6) satisfies

(1) $P[t](t)$ is Lipschitz continuous.

(2) $P[t](t)$ is bounded, if linear time-varying system (4) is controllable.

**Proof:** Firstly, let us prove $P[t](t)$ is Lipschitz continuous. Transform (6) into the form of matrix differential equation as

$$
\begin{bmatrix}
\dot{\lambda}[i](t) \\
\bar{X}[i](t)
\end{bmatrix} =
\begin{bmatrix}
-A(x^{[i-1]}(t)) & -Q(x^{[i-1]}(t)) \\
\Xi & A(x^{[i-1]}(t))
\end{bmatrix}
\begin{bmatrix}
\lambda[i](t) \\
X[i](t)
\end{bmatrix},
\begin{bmatrix}
\lambda[i](t_f) \\
X[i](t_f)
\end{bmatrix} =
\begin{bmatrix}
F(t_f) \\
I
\end{bmatrix}
$$

where $\Xi = B(x^{[i-1]}(t), u^{[i-1]}(t))R^{-1}(x^{[i-1]}(t))B^T(x^{[i-1]}(t), u^{[i-1]}(t)) - E(x^{[i-1]}(t), w^{[i-1]}(t)) \times S^{-1}(x^{[i-1]}(t))E^T(x^{[i-1]}(t), w^{[i-1]}(t)).$ Thus, the solution $P[i](t)$ of the Riccati differential Equations (6) becomes

$$
P[i](t) = \lambda[i](t) \left( X[i](t) \right)^{-1}
$$

If Assumption 4.1 is satisfied, such that $A(x), B(x, u), E(x, w), R^{-1}(x), S^{-1}(x), F(x)$ and $Q(x)$ are Lipschitz continuous, then, $X[i](t)$ and $\lambda[i](t)$ are Lipschitz continuous. Furthermore, it is easy to verify that $(X[i](t))^{-1}$ also satisfies the Lipschitz condition. Hence, $P[i](t)$ is Lipschitz continuous.

Next, we prove that $P[i](t)$ is bounded.

If the linear time varying system (4) is controllable, there must exist $\hat{u}[i](t), \bar{w}[i](t)$ such that $x(t_1) = 0$ at $t = t_1$. We define $\hat{u}[i](t), \bar{w}[i](t)$ as

$$
\hat{u}[i](t) = \begin{cases}
\hat{u}[i](t), & t \in [0, t_1) \\
0, & t \in [t_1, \infty)
\end{cases}
$$

$$
\bar{w}[i](t) = \begin{cases}
\bar{w}[i](t) = S^{-1}(x^{[i-1]}(t))E^T(x^{[i-1]}(t), w^{[i-1]}(t))P[i](t)x[i](t), & t \in [0, t_1) \\
0, & t \in [t_1, \infty)
\end{cases}
$$

where $\hat{u}[i](t)$ is any control strategy to make $x(t_1) = 0$, $\bar{w}[i](t)$ is defined as the optimal strategy. When $t \geq t_1$, let $\hat{w}[i](t)$ and $\bar{w}[i](t)$ be 0, the state $x(t)$ will still hold at 0.

The optimal performance index function $V^*[i](x_0, u, w)$ described as

$$
\begin{aligned}
V^*[i](x_0, u, w) = & \frac{1}{2} x^{[i-1]}(T_f)F(x^{[i-1]}(T_f))x^{[i]}(T_f) + \frac{1}{2} \int_{t_0}^{T_f} \left[ x^{[i]}(t)Q(x^{[i-1]}(t))x^{[i]}(t) \\
& + u^{*[i]}(t)R(x^{[i-1]}(t))u^{*[i]}(t) - w^{*[i]}(t)S(x^{[i-1]}(t))w^{*[i]}(t) \right] dt
\end{aligned}
$$

where $u^{*[i]}(t)$ and $w^{*[i]}(t)$ are the optimal strategies. $V^*[i](x_0, u, w)$ is minimized by $u^{*[i]}(t)$ and maximized by $w^{*[i]}(t)$. For the linear system, $V^*[i](x_0, u, w)$ can be expressed as

$$
V^*[i](x_0, u, w) = \frac{1}{2} x^{[i]}(T_f)P[i](t)x^{[i]}(t).
$$

Since $x^{[i]}(t)$ is arbitrary, if $V^*[i](x_0, u, w)$ is bounded, then $P[i](t)$ is bounded. Next we discuss the boundedness of $V^*[i](x_0, u, w)$ in two cases.
Case (1) $t_1 < t_f$,
\[
V^{*\{i\}}(x_0, u, w) \leq \frac{1}{2} x^{[i]^T(t_f)} F(x^{[i-1]}(t_f)) x^{[i]}(t_f) + \frac{1}{2} \int_{t_0}^{t_f} \left[ x^{[i]^T(t)} Q(x^{[i-1]}(t)) x^{[i]}(t) + u^{[i]^T(t)} R(x^{[i-1]}(t)) u^{[i]}(t) - w^{[i]^T(t)} S(x^{[i-1]}(t)) w^{[i]}(t) \right] dt \\
= \frac{1}{2} \int_{t_0}^{t_1} \left[ x^{[i]^T(t)} Q(x^{[i-1]}(t)) x^{[i]}(t) + u^{[i]^T(t)} R(x^{[i-1]}(t)) u^{[i]}(t) - w^{[i]^T(t)} S(x^{[i-1]}(t)) w^{[i]}(t) \right] dt
\]
\[
= V^{*\{i\}}_{t_1}(x) < \infty
\]

Case (2) $t_1 \geq t_f$,
\[
V^{*\{i\}}(x_0, u, w) \leq \frac{1}{2} x^{[i]^T(t_f)} F(x^{[i-1]}(t_f)) x^{[i]}(t_f) + \frac{1}{2} \int_{t_0}^{t_f} \left[ x^{[i]^T(t)} Q(x^{[i-1]}(t)) x^{[i]}(t) + u^{[i]^T(t)} R(x^{[i-1]}(t)) u^{[i]}(t) - w^{[i]^T(t)} S(x^{[i-1]}(t)) w^{[i]}(t) \right] dt \\
\leq \frac{1}{2} \int_{t_0}^{\infty} \left[ x^{[i]^T(t)} Q(x^{[i-1]}(t)) x^{[i]}(t) + u^{[i]^T(t)} R(x^{[i-1]}(t)) u^{[i]}(t) - w^{[i]^T(t)} S(x^{[i-1]}(t)) w^{[i]}(t) \right] dt \\
= \frac{1}{2} \int_{t_0}^{t_1} \left[ x^{[i]^T(t)} Q(x^{[i-1]}(t)) x^{[i]}(t) + u^{[i]^T(t)} R(x^{[i-1]}(t)) u^{[i]}(t) - w^{[i]^T(t)} S(x^{[i-1]}(t)) w^{[i]}(t) \right] dt
\]
\[
= V^{*\{i\}}_{t_1}(x) < \infty
\]

From (14) and (15), we can know that $V^{*\{i\}}_{t_1}(x)$ has an upper bound independent of $t_f$. Hence, $P^{[i]}(t)$ is bounded.

According to Lemma 4.2, $P^{[i]}(t)$ is bounded and Lipschitz continuous. If Assumption 4.1 is satisfied, then $C(x, u), D(x, w), H(x, w)$ and $G(x, w)$ are bounded and Lipschitz continuous in their arguments, thus satisfying

\[
(B4) \quad \|C(x, u)\| \leq \delta_1, \quad \|D(x, w)\| \leq \sigma_1
\]
\[
(B5) \quad \|C(x_1, u_1) - C(x_2, u_2)\| \leq \delta_2 \|x_1 - x_2\| + \delta_3 \|u_1 - u_2\|,
\]
\[
\|D(x_1, w_1) - D(x_2, w_2)\| \leq \sigma_2 \|x_1 - x_2\| + \sigma_3 \|w_1 - w_2\|
\]
\[
(B6) \quad \|H(x, u)\| \leq \xi_1, \quad \|G(x, w)\| \leq \xi_1
\]
\[
(B7) \quad \|H(x_1, u_1) - H(x_2, u_2)\| \leq \xi_2 \|x_1 - x_2\| + \xi_3 \|u_1 - u_2\|
\]
\[
\|G(x_1, w_1) - G(x_2, w_2)\| \leq \xi_2 \|x_1 - x_2\| + \xi_3 \|w_1 - w_2\|
\]
\[
\forall x \in \mathbb{R}^n, \forall u \in \mathbb{R}^{m_1}, \forall w \in \mathbb{R}^{m_2}, \text{ and for finite positive numbers } \delta_i, \sigma_i, \xi_i, \zeta_i, i = 1, 2, 3.
\]

**Theorem 4.1.** Consider the system (1) of the nonaffine nonlinear zero-sum game with the performance index function (2), the approximating sequences (4) and (5) can be introduced. $F(x(t)) \geq 0, Q(x(t)) \geq 0, R(x(t)) > 0$, the terminal time $t_f$ is specified. Let Assumption 4.1, assumptions (A1), (A2) hold and $S(x(t)) > \tilde{S}$, for small enough $t_f$ or $x_0$, the limit of the solution of approximating sequence (4) converge to the unique solution of system (1) on $C([t_0, t_f]; \mathbb{R}^n)$. Meanwhile, the approximating sequences of optimal strategies given by (7) also converge on $C([t_0, t_f]; \mathbb{R}^{m_1})$ and $C([t_0, t_f]; \mathbb{R}^{m_2})$ if

\[
\|\Psi(t)\| < 1
\]
where

\[
\Psi(t) = \begin{bmatrix}
\psi_1 & \psi_2 & \psi_3 \\
\psi_4 & \psi_5 & \psi_6 \\
\psi_7 & \psi_8 & \psi_9 \\
\end{bmatrix}
\]

\[
\psi_1(t) = \frac{\xi_1 + \mu_0}{\xi_1 + \epsilon_1} x_0 \left\| e^{(\xi_1 + \epsilon_1)(t-t_0)} - \frac{\xi_1 + \mu_0}{\xi_1 + \epsilon_1} \right\| \|x_0\| e^{\mu_0(t-t_0)},
\]

\[
\psi_2(t) = \frac{\xi_1 + \mu_0}{\xi_1 + \epsilon_1} x_0 \left\| e^{\mu_0(t-t_0)} (e^{(\xi_1 + \epsilon_1)(t-t_0)} - 1) \right\|
\]

\[
\psi_3(t) = \frac{\xi_1 + \mu_0}{\xi_1 + \epsilon_1} x_0 \left\| e^{\mu_0(t-t_0)} (e^{(\xi_1 + \epsilon_1)(t-t_0)} - 1) \right\|
\]

\[
\psi_4(t) = \delta_1 \psi_1(t) + \delta_2 \left\| x_0 \right\| e^{(\xi_1 + \epsilon_1)(t-t_0)},
\]

\[
\psi_5(t) = \delta_1 \psi_2(t) + \delta_3 \left\| x_0 \right\| e^{(\xi_1 + \epsilon_1)(t-t_0)},
\]

\[
\psi_6(t) = \delta_1 \psi_3(t),
\]

\[
\psi_7(t) = \sigma_1 \psi_1(t) + \sigma_2 \left\| x_0 \right\| e^{(\xi_1 + \epsilon_1)(t-t_0)},
\]

\[
\psi_8(t) = \sigma_1 \psi_2(t),
\]

\[
\psi_9(t) = \sigma_1 \psi_3(t) + \sigma_3 \left\| x_0 \right\| e^{(\xi_1 + \epsilon_1)(t-t_0)},
\]

\[
\tilde{S} = \max\{\tilde{S}^{[i]}\}.
\]

**Proof:** The approximating sequence (9) is an inhomogeneous differential equation, whose solution can be given by

\[
x^{[i]}(t) = \Phi^{[i-1]}(t, t_0) x^{[i]}(t_0)
\]

\[
+ \int_{t_0}^{t} \Phi^{[i-1]}(t, s) \left[ H \left(x^{[i-1]}(s), u^{[i-1]}(s)\right) + G \left(x^{[i-1]}(s), w^{[i-1]}(s)\right) \right] x^{[i]}(s) ds.
\]

Then,

\[
\left\| x^{[i]}(t) \right\| \leq \left\| \Phi^{[i-1]}(t, t_0) \right\| \left\| x^{[i]}(t_0) \right\|
\]

\[
+ \int_{t_0}^{t} \left\| \Phi^{[i-1]}(t, s) \right\| \left\| H \left(x^{[i-1]}, u^{[i-1]}\right) \right\| ds
\]

\[
+ \left\| G \left(x^{[i-1]}, w^{[i-1]}\right) \right\| \left\| x^{[i]}(s) \right\| ds.
\]

According to inequality (11) and assuming (B6) hold, we can obtain

\[
e^{-\mu_0 t} \left\| x^{[i]}(t) \right\| \leq e^{-\mu_0 t_0} \left\| x_0 \right\| + \int_{t_0}^{t} \left( \xi_1 + \xi_1 \right) e^{-\mu_0 s} \left\| x^{[i]}(s) \right\| ds.
\]

On the basis of Gronwall-Bellman's inequality

\[
\left\| x^{[i]}(t) \right\| \leq \left\| x_0 \right\| e^{(\mu_0 + \xi_1)(t-t_0)}
\]

which is bounded by a small time interval \( t \in [t_0, t_f] \) or small \( x_0 \).
From (17) we have
\[ x_i(t) - x_{i-1}(t) = \left[ \Phi[i-1](t, t_0) - \Phi[i-2](t, t_0) \right] x_0 \]
\[ + \int_{t_0}^{t} \Phi[i-1](t, s) H (x[i-1], u[i-1]) \left[ x_i(s) - x_{i-1}(s) \right] ds \]
\[ + \int_{t_0}^{t} \Phi[i-1](t, s) G (x[i-1], w[i-1]) \left[ x_i(s) - x_{i-1}(s) \right] ds \]
\[ + \int_{t_0}^{t} \Phi[i-1](t, s) \left[ H (x[i-1], u[i-1]) - H (x[i-2], u[i-2]) \right] x_{i-1}(s) ds \]
\[ + \int_{t_0}^{t} \Phi[i-1](t, s) \left[ G (x[i-1], w[i-1]) - G (x[i-2], w[i-2]) \right] x_{i-1}(s) ds \]
\[ + \int_{t_0}^{t} \left[ \Phi[i-1](t, s) - \Phi[i-2](t, s) \right] H (x[i-2], u[i-2]) x_{i-1}(s) ds \]
\[ + \int_{t_0}^{t} \left[ \Phi[i-1](t, s) - \Phi[i-2](t, s) \right] G (x[i-2], w[i-2]) x_{i-1}(s) ds. \] (21)

Consider supremum to both sides of (21) and let 
\[ \beta[i](t) = \sup_{s \in [t_0, t]} \left\| x_i(s) - x_{i-1}(s) \right\|, \]
\[ \gamma[i](t) = \sup_{s \in [t_0, t]} \left\| u_i(s) - u_{i-1}(s) \right\|, \eta[i](t) = \sup_{s \in [t_0, t]} \left\| w_i(s) - w_{i-1}(s) \right\|. \]

By using (11), (B6), (B7) and Lemma 4.1, we can get
\[ \beta[i](t) \leq \alpha \left\| x_0 \right\| e^{\rho_0(t-t_0)} (t - t_0) \beta[i-1](t) + (\zeta_1 + \xi_1) \int_{t_0}^{t} e^{\rho_0(s-t)} \beta[i](s) ds \]
\[ + \left\| x_0 \right\| e^{\rho_0(t-t_0)} \int_{t_0}^{t} e^{(\zeta_1 + \xi_1)(s-t_0)} \left[ \zeta_2 \beta[i](s) + \zeta_3 \gamma[i-1](s) \right] ds \]
\[ + \left\| x_0 \right\| e^{\rho_0(t-t_0)} \int_{t_0}^{t} e^{(\zeta_1 + \xi_1)(s-t_0)} \left[ \xi_2 \beta[i](s) + \xi_3 \eta[i-1](s) \right] ds \] (22)
\[ + \alpha \zeta_1 \left\| x_0 \right\| e^{\rho_0(t-t_0)} \int_{t_0}^{t} e^{(\zeta_1 + \xi_1)(s-t_0)} (t - s) \beta[i-1](s) ds \]
\[ + \alpha \zeta_1 \left\| x_0 \right\| e^{\rho_0(t-t_0)} \int_{t_0}^{t} e^{(\zeta_1 + \xi_1)(s-t_0)} (s - t) \beta[i-1](s) ds. \]

Combining the similar terms, we have
\[ \beta[i](t) \leq \psi_1(t) \beta[i-1](t) + \psi_2(t) \gamma[i-1](t) + \psi_3(t) \eta[i-1](t) \] (23)

where \( \psi_1(t) - \psi_3(t) \) are described in (16).

Similarly, from (10), we can get
\[ u[i](t) - u[i-1](t) = C (x[i-1], u[i-1]) \left[ x_i(t) - x_{i-1}(t) \right] \]
\[ + \left[ C (x[i-1], u[i-1]) - C (x[i-2], u[i-2]) \right] x_{i-1}(t) \]
\[ w[i](t) - w[i-1](t) = D (x[i-1], w[i-1]) \left[ x_i(t) - x_{i-1}(t) \right] \]
\[ + \left[ D (x[i-1], w[i-1]) - D (x[i-2], w[i-2]) \right] x_{i-1}(t). \] (24)

According to (B4), (B5) and (20), we have
\[ \gamma[i](t) \leq \psi_4(t) \beta[i-1](t) + \psi_5(t) \gamma[i-1](t) + \psi_6(t) \eta[i-1](t) \]
\[ \eta[i](t) \leq \psi_7(t) \beta[i-1](t) + \psi_8(t) \gamma[i-1](t) + \psi_9(t) \eta[i-1](t) \] (25)
where \( \psi_3(t) \) and \( \psi_6(t) \) are shown in (16).

Then, combining (23) and (25), we have

\[ \Theta^{[i]}(t) \leq \Psi(t)\Theta^{[i-1]}(t) \]  

(26)

where \( \Theta^{[i]}(t) = \begin{bmatrix} \beta^{[i]}(t) \\ \gamma^{[i]}(t) \\ \eta^{[i]}(t) \end{bmatrix} \), \( \Psi(t) = \begin{bmatrix} \psi_1 & \psi_2 & \psi_3 \\ \psi_4 & \psi_5 & \psi_6 \\ \psi_7 & \psi_8 & \psi_9 \end{bmatrix} \).

By induction, \( \Theta^{[i]} \) satisfies

\[ \Theta^{[i]}(t) \leq \Psi^{i-1}(t)\Theta^{[1]}(t) \]  

(27)

which implies that \( x^{[i]}(t), u^{[i]}(t) \) and \( w^{[i]}(t) \) are Cauchy sequences in Banach spaces \( C([t_0, t_f]; \mathbb{R}^n) \), \( C([t_0, t_f]; \mathbb{R}^m_1) \) and \( C([t_0, t_f]; \mathbb{R}^m_2) \), respectively. If \( \| \Psi(t) \| < 1 \), then, the state sequence \( \{x^{[i]}(t)\} \) converges on \( C([t_0, t_f]; \mathbb{R}^n) \), and the sequences of optimal strategies \( \{u^{[i]}\} \) and \( \{w^{[i]}\} \) also converge on \( C([t_0, t_f]; \mathbb{R}^m_1) \) and \( C([t_0, t_f]; \mathbb{R}^m_2) \) on \([t_0, t_f]\). It means that \( x^{[i]}(t) = x^{[i]}(t), u^{[i]}(t) = u^{[i]}(t), w^{[i]}(t) = w^{[i]}(t) \) when \( i \to \infty \).

Hence, system (1) has a unique solution on \([t_0, t_f]\) which is given by the limit of the solution of approximating sequence (4). \( \square \)

Based on the iteration algorithm described in Theorem 4.1, the design procedure of optimal strategies for nonlinear nonaffine zero-sum game is summarized as follows.

1. Give \( x_0 \), maximum iteration times \( i_{\text{max}} \) and approximation accuracy \( \varepsilon \).
2. Use a factored form to represent the system as (3).
3. Set \( i = 0 \). Let \( x^{[i]}(t) = x_0, w^{[i]}(t) = 0 \) and \( w^{[i]}(t) = 0 \). Compute the corresponding matrix-valued functions \( A(x_0), B(x_0, 0), E(x_0, 0), F(x_0), Q(x_0), R(x_0) \) and \( S(x_0) \).
4. Compute \( x^{[0]}(t) \) and \( P^{[0]}(t) \) according to differential Equations (6) and (8) with \( x(t_0) = x_0, P(t_f) = F(x_f) \).
5. Set \( i = i + 1 \). Compute the corresponding matrix-valued functions \( A(x^{[i-1]}(t)), B(x^{[i-1]}(t), w^{[i-1]}(t)), E(x^{[i-1]}(t), w^{[i-1]}(t)), F(x^{[i-1]}(t)), Q(x^{[i-1]}(t)), R(x^{[i-1]}(t)) \) and \( S(x^{[i-1]}(t)) \).
6. Compute \( x^{[i]}(t) \) and \( P^{[i]}(t) \) by (6) and (8) with \( x(t_0) = x_0, P(t_f) = F(x_f) \).
7. If \( \| x^{[i]}(t) - x^{[i-1]}(t) \| < \varepsilon \), go to Step 9; otherwise, go to Step 8.
8. If \( i > i_{\text{max}} \), then go to Step 9; else, go to Step 5.
9. Stop.

5. \textbf{Numerical Example.} We now show the power of our iterative algorithm for finding optimal strategies for the nonaffine nonlinear zero-sum game.

In the following, we introduce an example about a control system has the form (1) with control input \( u \), subject to disturbance \( w \) and a performance index \( V \). The control input \( u \) is required to minimize the performance index \( V \). If the disturbance has a great effect on the system, the disturbance single \( w \) has to maximize the performance index \( V \). The conflicting design can guarantee the optimality and strong robustness of the system at the same time. This is a zero-sum game problem, which can be described by the state equations

\begin{align*}
\dot{x}_1(t) &= -2x_1(t) + x_2^2(t) - x_1(t)u(t) + u^2(t) - 3x(t)w(t) + 5w^2(t) \\
\dot{x}_2(t) &= 5x_1^2(t) - 2x_2(t) + x_2^3(t) + w^2(t) + w^2(t)
\end{align*}

(28)

Define the finite-horizon performance index function as form (2), where \( F = 0.01I_{2 \times 2}, Q = 0.01I_{2 \times 2}, R = 1 \) and \( S = 1 \), where \( I \) is an identity matrix. Clearly, (28) is not affine in \( u(t) \) and \( w(t) \), has the control-nonaffine nonlinear structure. Therefore, consider representing the system (28) in the factored form \( A(x(t))x(t), B(x(t), u(t))u(t) \) and...
The optimal strategies designs given by Theorem 4.1 can now be applied to (3) with dynamics (29).

\[
E(x(t), w(t)) = \begin{bmatrix}
2 & x_2(t) \\
5x_1(t) & -2 + x_2(t)
\end{bmatrix}, \quad B(x(t), u(t)) = \begin{bmatrix}
x_1(t) + u(t) \\
u(t)
\end{bmatrix}
\]

\[
E(x(t), w(t)) = \begin{bmatrix}
-3x_1(t) + 5w(t) \\
w(t)
\end{bmatrix}.
\]

The optimal strategies designs given by Theorem 4.1 can now be applied to (3) with dynamics (29).

Figure 1. The state trajectory $x_1(t)$ of each iteration

The initial state vectors are chosen as $x_0 = [0.7, 0]^T$ and the terminal time is set to $t_f = 5$. Let us define the required error norm between the solutions of the linear time-vary differential equations by $\|x[i]^0(t) - x[i-1]^0(t)\| < \varepsilon = 0.005$, which needs to be satisfied if convergence is to be achieved. The factorization is given by (29). Implementing the iteration algorithm described in Section 4, it just needs six sequences to satisfy the required bound, $\|x[6]^0(t) - x[5]^0(t)\| = 0.0032$. With the increasing of iterative times, the approximation error will reduce obviously. When the iterative times $i = 25$, the approximation error just is $5.1205e-010$.

Define the maximum iterative times $i_{\text{max}} = 25$. Figure 1 represents the convergence curves of the state trajectory of each linear quadratic zero-sum game. It can be seen that the sequence is obviously convergent. Magnification of state trajectories are given in the figure, which shows that the error will be smaller as the iteration times becomes bigger. The trajectories of control input $u(t)$ and disturbance input $w(t)$ of each iteration are also convergent by iteration sequentially, which are shown in Figure 2. The approximate optimal strategies $u^*(t)$ and $w^*(t)$ are obtained by the last iteration. Substituting the approximate optimal strategies $u^*(t)$ and $w^*(t)$ into the system of the zero-sum game (28), we can get the state trajectory. The norm of the error between this state trajectory and the state trajectory of the last iteration just is 0.0019, which has proved that the approximating iteration approach proposed in this paper is highly effective.
A NEW ITERATION APPROACH TO SOLVE A CLASS OF FCNN ZERO-SUM GAME

Figure 2. The trajectories of $u(t)$ and $w(t)$ of each iteration

6. Conclusion. In this paper, a general class of nonaffine nonlinear zero-sum game has been considered. A new iteration method was introduced to design the optimal strategies by iteration between a sequence of state trajectories and a sequence of Riccati differential equations. The convergence of the approximating linear time-varying sequences has been proved under the local Lipschitz continuity condition. The simulation results have successfully demonstrated the effectiveness of the proposed design method. Future work will focus on the situation that the saddle point of the zero-sum game does not exist.

Acknowledgment. This work was supported by the National Natural Science Foundation of China (60774048, 50977008, 60821063, 61034005).

REFERENCES


