Multi-agent Optimal Control Problems and Variational Inequality Based Reformulations

George Leitmann, Stefan Pickl, and Zhengyu Wang

Abstract The multi-agent optimal control problem involves a decision process with multiple agents, where each agent solves an optimal control problem with the individual cost functional and strategy set, and the cost functional is dependent on all the other agents' state and/or control variables. Here the "agent" can be understood as a true decision maker, or as an abstract optimization criterion. The strategy sets, along with admissible control set, are often described by a system of parameterized ordinary differential/difference equations (the state dynamic) or partial differential equations, and in realistic settings they may be dependent on the rivals's variables due to, for example, certain constraints from the common resources. This chapter describes the multi-agent optimal control problem, and studies the reformulation of a system of differential equations constrained by parameterized variational inequalities, along with some initial and/or boundary conditions. This reformulation presents differential equations, variational inequalities, and equilibrium conditions in a systematic way, and is advantageous since it can be treated as a system of differential algebraic equations, for which abundant theory and algorithms are available.

1 Optimal Control Problems

In a multi-agent optimal control problem each agent solves an optimal control problem that is dependent on the rivals' states and decisions. Let us begin the study with the single-agent case: the standard optimal control problem.

G. Leitmann
Mechanical Engineering Department, University of California, Berkeley, Berkeley, USA
e-mail: gleit@berkeley.edu

S. Pickl (✉)
Department of Computer Science, Core Competence Center for Operations Research, Universität der Bundeswehr München, Munich, Germany
e-mail: stefan.pickl@unibw.de

Z. Wang
Department of Mathematics, Nanjing University, Nanjing, China
e-mail: zywang@nju.edu.cn

J. Haunschmied et al. (eds.), *Dynamic Games in Economics*,
Dynamic Modeling and Econometrics in Economics and Finance 16,
DOI 10.1007/978-3-642-54248-0_10,

For more details on optimal control problems and on the multi-agent extension we refer to the basic books by Leitmann (1976, 1981).

1.1 Problem Description

Let the terminal time $T > 0$ and the initial point $x^0 \in \mathbb{R}^n$ be given, let $U \subset \mathbb{R}^n$ be an open bounded set, $\Xi \in \mathbb{R}^m$ be convex and closed, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $\{W(s)\}_{0 \leq t \leq T}$ be a d-dimensional Brownian motion, let $\mathcal{A}$ be a subset of all progressive measurable stochastic processes $u(\cdot) : [0, T] \times \Omega \to \Xi$. Given the following four functions, of which the first two constitute the dynamic and the last two give the cost functional:

- $f : [0, T] \times \mathbb{R}^n \times \Xi \to \mathbb{R}^n$ (drift term),
- $\sigma : [0, T] \times \mathbb{R}^n \times \Xi \to \mathbb{R}^{n \times d}$ (diffusion term),
- $\varphi : [0, T] \times \bar{U} \times \Xi \to \mathbb{R}$ (running cost),
- $\psi : [0, T] \times \bar{U} \times \Xi \to \mathbb{R}$ (terminal cost).

For a $t \geq 0$ and for every $u(\cdot) \in \mathcal{A}$ and $(s, x) \in [t, T] \times \bar{U}$, the *state dynamic* is a stochastic differential equation (SDE for short) which is to find an Itô process $x(s)$ satisfying:

$$\begin{cases} dx(s) = f(s, x(s), u(s))ds + \sigma(s, x(s), u(s))dW(s) \\ x(t) = x \end{cases} \tag{1}$$

for $s \in (t, T]$, where a $u(\cdot) \in \mathcal{A}$ is called the *control*, and $x(\cdot)$ is called the *state*. We define the *cost functional* $J : [0, T] \times \mathbb{R}^n \times \mathcal{A} :\to \mathbb{R}$ by:

$$J\big(t, x, x(\cdot), u(\cdot)\big) := \mathbb{E}\left\{\int_t^T \varphi\big(s, x(s), u(s)\big)ds + \psi\big(T, x(T)\big)\right\}, \tag{2}$$

where $\mathbb{E}$ means the expectation over the statistics of $\{W(s)\}$. Denote $J(x(\cdot), u(\cdot)) = J(0, x^0, x(\cdot), u(\cdot))$ for simplicity if the state $x(\cdot)$ starts from x^0 at $t = 0$. Then the optimal control problem is just to find a pair $(x(\cdot), u(\cdot))$ minimizing $J(x(\cdot), u(\cdot))$ under the constraint given by the SDE (1):

$$\begin{aligned} &\min J\big(x(\cdot), u(\cdot)\big) \\ &\text{s.t. } dx(s) = f\big(s, x(s), u(s)\big)ds + \sigma\big(s, x(s), u(s)\big)dW(s) \\ &\qquad x(0) = x^0. \end{aligned} \tag{3}$$

For the SDE (1), one of the problems of the most interest is its solvability. For the details on this issues we refer to Øksendal (2003). Here we just mention the conditions required in part for guaranteeing the existence and the uniqueness of the strong solution with continuous paths of (1) for any choice of $u(\cdot)$:

$$\left\| f(t,x,u) - f(s,y,u) \right\|_2 + \left\| \sigma(t,x,u) - \sigma(s,y,u) \right\|_F \le C\big(\|x-y\|_2 + |t-s|\big)$$
$$\left\| f(t,x,u) \right\|_2 + \left\| \sigma(t,x,u) \right\|_F \le C\big(1 + \|x\|_2\big),$$

where $\|\cdot\|_F$ denotes the Frobenius norm, $f(s,x,u)$ and $\sigma(s,x,u)$ are assumed in $C^0([0,T]\times\mathbb{R}^n\times\Xi)$ and $f(\cdot,\cdot,u)$ and $\sigma(\cdot,\cdot,u)$ are in $C^1([0,T]\times\mathbb{R}^n)$ for every $u\in\Xi$, and $C\ge 0$ is a constant, $u\in\Xi$, $x,y\in\mathbb{R}^n$ and $t,s\in[0,T]$ are arbitrary.

1.2 Hamilton–Jacobi–Bellman Equation

Define the *value function*:

$$\begin{aligned} v(t,x) := \min\ & J\big(t,x,x(\cdot),u(\cdot)\big) \\ \text{s.t. } & dx(s) = f\big(s,x(s),u(s)\big)ds + \sigma\big(s,x(s),u(s)\big)dW(s) \\ & x(t) = x. \end{aligned}$$

Denote $\chi(t,x,u) = \frac{1}{2}\|\sigma(t,x,u)\|_F^2$, and denote

$$H(t,x,u,\nabla v,\Delta v) = \chi(t,x,u)\Delta v(t,x) + \big\langle f(t,x,u), \nabla v(t,x)\big\rangle + \varphi(t,x,u). \quad (4)$$

Suppose that $H(t,x,u,\nabla v,\Delta v)$ is continuously differentiable in u. Then the optimal control problem (3) can be reformulated as the Hamilton–Jacobi–Bellman equation (HJB equation for short):

$$\frac{\partial v(t,x)}{\partial t} + \min_{u\in\Xi} H\big(t,x,u,\nabla v(x,t),\Delta v(t,x)\big) = 0, \quad (5)$$

along with the terminal condition $v(T,x) = \psi(T,x)$, where $\Delta v(t,x)$ and $\nabla v(t,x)$ denote the Laplacian and the gradient of v in x, respectively. Normally, the HJB equation does not have a classic solution, for this one has to use another notion of solution: viscosity solution (refer to Fleming and Rishel 1975, for example).

1.3 Constrained Hamilton System

For the deterministic case: $\sigma(t,x,u)\equiv 0$, we introduce the costate variable $p(t) = \nabla v(t,x(t))$. Then the Hamiltonian defined in (4) reads:

$$H(t,x,u,p) = f(t,x,u)^T p + \varphi(t,x,u), \quad (6)$$

and by simple calculus we obtain the following Hamilton system from the HJB equation:

$$\begin{cases} \dot{p}(t) = -\nabla_x H(t, x(t), u(t), p(t)) \\ \dot{x}(t) = \nabla_p H(t, x(t), u(t), p(t)) \\ u(t) \in \arg\min\{H(t, x(t), z, p(t)), \text{ s.t. } z \in \Xi\} \\ x(0) = x^0 \text{ and } p(T) = \nabla_x \psi(T, x(T)), \end{cases} \tag{7}$$

where $\nabla_x \psi(T, x)$ denotes the gradient of $\psi(t, x)$ with respect to x.

1.4 VI Based Reformulations

Variational inequality (VI for short) is a powerful model for characterizing the optimal condition of optimization problems in a general setting (Facchinei and Pang 2003). Given a closed and convex subset $\Omega \subseteq \mathbb{R}^m$ and a mapping $G : \Omega \to \mathbb{R}$. Then by the minimum principle, we know that a local minimizer x^* of $G(\cdot)$ over the feasible domain Ω must satisfy the variational inequality (VI for short) of the following form

$$\left(x - x^*\right)^T \nabla G\left(x^*\right) \geq 0, \quad \forall x \in \Omega. \tag{8}$$

We remind us that Ξ is assumed convex and closed set. Here we allow us an abuse of the notation u: it means the control variable in the general case and means also a local minimizer of $H(t, x, \cdot, \nabla v, \Delta v)$ in some specific cases, which can be readily distinguished in the context. Then by the minimum principle, we know that u satisfies the VI

$$(z - u)^T \nabla_u H(t, x, u, \nabla v, \Delta v) \geq 0 \quad \forall z \in \Xi, \tag{9}$$

where $\nabla_u H(t, x, u, \nabla v, \Delta v)$ denote the gradient of $H(t, x, u, \nabla v, \Delta v)$ in u. We denote by $\mathrm{SOL}(\Xi, \nabla_u H(t, x, \cdot, \nabla v, \Delta v))$ the solution set of the above VI. Further known is that if moreover $\nabla_u H$ is convex in u, then a solution of the VI is just a global minimizer of H.

Now we arrive at the position for reformulating the HJB equation as the following PDE, which is constrained by a VI

$$\begin{cases} \frac{\partial v(t,x)}{\partial t} + H(t, x, u, \nabla v(t, x), \Delta v(t, x)) = 0 \\ u \in \mathrm{SOL}(\Xi, \nabla_u H(t, x, \cdot, \nabla v(t, x), \Delta v(t, x)) \\ v(T, x) = \psi(T, x). \end{cases}$$

It is well known that $u \in \mathrm{SOL}(\Xi, \nabla_u H(t, x, \cdot, \nabla v, \Delta v))$ if and only if

$$u = \mathrm{Pr}_{\Xi}\left(u - \nabla_u H(t, x, u, \nabla v, \Delta v)\right),$$

where $\Pr_{\Xi}(\cdot)$ denotes the projection onto Ξ. Then the HJB equation can further be reformulated as the PDE constrained by a system of algebraic equations

$$\begin{cases} \frac{\partial v(t,x)}{\partial t} + H(t, x, u, \nabla v(t, x), \Delta v(t, x)) = 0 \\ u = \Pr_{\Xi}(u - \nabla_u H(t, x, u, \nabla v(t, x), \Delta v(t, x))) \\ v(T, x) = \psi(T, x). \end{cases} \tag{10}$$

Note that the projection operation often leads to the nonsmoothness of the algebraic system in the above hybrid system.

For the constrained Hamilton system (7), the VI formulation gives the following system:

$$\begin{cases} \dot{p}(t) = -\nabla_x H(t, x(t), u(t), p(t)) \\ \dot{x}(t) = \nabla_p H(t, x(t), u(t), p(t)) \\ u(t) = \Pr_{\Xi}(u(t) - \nabla_u H(t, x(t), u(t), p(t))) \\ x(0) = x^0 \text{ and } p(T) = \nabla_x \psi(T, x(T)), \end{cases} \tag{11}$$

where $H(t, x, u, p)$ is defined by (6). This is a system of ordinary differential equations constrained by a parameterized VI, called *differential variational inequality* (DVI for short). For a comprehensive treatment of the DVI, we refer to Pang and Stewart (2008).

The system (11) usually has no classic solution, and we have to seek the weak solution $(x(t), p(t), u(t))$, where x and p are absolutely continuous and u is integrable on $[0, T]$ such that $\forall 0 \le s \le t \le T$:

$$x(t) - x(s) = \int_s^t \nabla_p H\big(\tau, x(\tau), u(\tau), p(\tau)\big) d\tau,$$

and

$$p(t) - p(s) = -\int_s^t \nabla_x H\big(\tau, x(\tau), u(\tau), p(\tau)\big) d\tau,$$

and $u(t) = \Pr_{\Xi}(u(t) - \nabla_u H(t, x(t), u(t), p(t)))$ holds for almost all $t \in [0, T]$.

2 Multi-agent Optimal Control Problems

2.1 Problem Description

The multi-agent optimal control problem involves a decision process with multiple agents, where each agent solves an optimal control problem with his own cost functional and admissible control set. Each agent's cost functional is, and its admissible control set may be, dependent on all the other agents' state and control variables. Such a problem is also referred as the Nash equilibrium problem, where the agent is usually called as *player*.

Denote by $x_\nu \in \mathbb{R}^{n_\nu}$ and $u_\nu \in \mathbb{R}^{m_\nu}$ the ν-th player's state and control variables, respectively. The control is also called as strategy, action or decision. Collectively write $x = (x_\nu)_{\nu=1}^N \in \mathbb{R}^n$, $u = (u_\nu)_{\nu=1}^N \in \mathbb{R}^m$, $x_{-\nu} = (x_{\nu'})_{\nu' \neq \nu} \in \mathbb{R}^{n-n_\nu}$ and $u_{-\nu} = (u_{\nu'})_{\nu' \neq \nu} \in \mathbb{R}^{m-m_\nu}$, where $n = \sum_{\nu=1}^N n_\nu$ and $m = \sum_{\nu=1}^N m_\nu$. When we emphasize the ν-th player's state and strategy variables, we use the block form $x = (x_\nu, x_{-\nu})$ and $u = (u_\nu, u_{-\nu})$ to represent x and u, respectively. For the ν-th player, we denote

- admissible control set (the strategy set) by

$$\Xi_\nu(u_{-\nu}) = \big\{u_\nu | h_\nu(u_\nu) \le 0, g(u_\nu, u_{-\nu}) \le 0\big\},$$

  where $h_\nu(\cdot) : \mathbb{R}^{n_\nu} \to \mathbb{R}^{l_\nu}$ and $g(\cdot, u_{-\nu}) : \mathbb{R}^{n_\nu} \to \mathbb{R}^\ell$;
- the state dynamic by

$$\begin{cases} dx_\nu(t) = f_\nu(t, x_\nu(t), u_\nu(t))dt + \sigma_\nu(t, x_\nu(t), u_\nu(t))dW(t) \\ x_\nu(0) = x_\nu^0, \end{cases} \tag{12}$$

  where $x_\nu^0 \in \mathbb{R}^{n_\nu}$ is the initial state, $f_\nu : [0, T] \times \mathbb{R}^{n_\nu} \times \Xi \to \mathbb{R}^n$ is the drift term, $\sigma_\nu : [0, T] \times \mathbb{R}^{n_\nu} \times \Xi \to \mathbb{R}^{n_\nu \times d_\nu}$ is the diffusion term;
- the cost functional by

$$J_\nu\big(x(\cdot), u(\cdot)\big) := \mathbb{E}\bigg\{\int_0^T \varphi_\nu\big(t, x(t), u(t)\big)dt + \psi_\nu\big(T, x_\nu(T)\big)\bigg\}, \tag{13}$$

  where $\psi_\nu : [0, T] \times \mathbb{R}^{n_\nu} \to \mathbb{R}$ and $\varphi_\nu : [0, T] \times \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, and $T > 0$ is the terminal time.

Writing

$$J_\nu\big(x(\cdot), u(\cdot)\big) = J_\nu\big(x_\nu(\cdot), x_{-\nu}(\cdot), u_\nu(\cdot), u_{-\nu}(\cdot)\big),$$

the solution (or called the equilibrium point) of the multi-agent optimal control problem is a state-control pair $(x^*(\cdot), u^*(\cdot))$ satisfying: for fixed $x^*_{-\nu}(\cdot)$ and $u^*_{-\nu}(\cdot)$, $(x^*_\nu(\cdot), u^*_\nu(\cdot))$ is a solution of the following optimal control problem

$$\begin{aligned} &\min\ J_\nu\big(x_\nu(\cdot), x^*_{-\nu}(\cdot), u_\nu(\cdot), u^*_{-\nu}(\cdot)\big) \\ &\text{s.t.}\ \ dx_\nu(t) = f_\nu\big(t, x_\nu(t), u_\nu(t)\big)dt + \sigma_\nu\big(t, x_\nu(t), u_\nu(t)\big)dW(t) \\ &\qquad x_\nu(0) = x_\nu^0 \\ &\qquad u_\nu(t) \in \Xi_\nu\big(u^*_{-\nu}(t)\big) \quad \text{for almost all } t \in [0, T]. \end{aligned} \tag{14}$$

Note that $\Xi_\nu(\cdot)$ is a set-valued mapping given by the shared constraint $g(u_\nu, u_{-\nu}) \le 0$, namely, the ν-th player's strategy set is dependent on its rivals' states and controls. Without the shared constraint, the strategy set Ξ_ν is constant, and then the problem (14) reduces to the standard dynamic Nash equilibrium problem.

Here we make the following blanket assumptions on the convexity of the strategy set.

Assumption 1 Suppose that all the components of h_ν and g are convex for any ν.

2.2 Reformulation of System of HJB Equations

Define the value function for the ν-th player:

$$
\begin{aligned}
v_\nu(t,x) := \min\ & J_\nu\big(x_\nu(\cdot), x^*_{-\nu}(\cdot), u_\nu(\cdot), u^*_{-\nu}(\cdot)\big) \\
\text{s.t.}\ \ & dx_\nu(s) = f_\nu\big(s, x_\nu(s), u_\nu(s)\big)ds + \sigma_\nu\big(s, x_\nu(s), u_\nu(s)\big)dW(s) \\
& x_\nu(t) = x_\nu \\
& u_\nu(s) \in \Xi_\nu(u^*_{-\nu}(s)) \quad \text{for almost all } s \in [t,T].
\end{aligned}
$$

Denote $\chi_\nu(t,x,u) = \frac{1}{2}\|\sigma_\nu(t,x,u)\|_F^2$, and denote

$$
\begin{aligned}
& H_\nu(t,x,u,\nabla_{x_\nu} v_\nu, \Delta_{x_\nu} v_\nu) \\
&\quad = \chi_\nu(t,x_\nu,u_\nu)\Delta_{x_\nu} v_\nu(t,x) + \big\langle f_\nu(t,x_\nu,u_\nu), \nabla_{x_\nu} v_\nu(t,x)\big\rangle + \varphi_\nu(t,x,u), \quad (15)
\end{aligned}
$$

$\Delta_{x_\nu} v_\nu(t,x)$ and $\nabla_{x_\nu} v_\nu(t,x)$ denote the Laplacian and the gradient of v_ν in x_ν, respectively. We suppose in our setting that $H_\nu(t,x,u,\nabla_{x_\nu} v_\nu, \Delta_{x_\nu} v_\nu)$ is continuously differentiable in u_ν. Write $v_\nu(T,x) = v_\nu(T,x_\nu,x_{-\nu})$. Then the HJB equation (5) for the problem (14) has the following form:

$$
\begin{cases}
\frac{\partial v_\nu(t,x)}{\partial t} + \min_{u_\nu \in \Xi_\nu(u_{-\nu})} H_\nu(t,x,u,\nabla_{x_\nu} v_\nu, \Delta_{x_\nu} v_\nu) = 0, \\
v_\nu(T,x_\nu,x_{-\nu}) = \psi_\nu(T,x_\nu).
\end{cases} \quad (16)
$$

Applying the VI formulation (9) to characterize the optimality of the minimization in (16), it gives

$$
\begin{cases}
\frac{\partial v_\nu(t,x)}{\partial t} + H_\nu(t,x,u,\nabla_{x_\nu} v_\nu, \Delta_{x_\nu} v_\nu) = 0, \\
u_\nu \in \mathrm{SOL}(\Xi_\nu(u_{-\nu}), \nabla_{u_\nu} H_\nu(t,x,u,\nabla_{x_\nu} v_\nu, \Delta_{x_\nu} v_\nu)) \\
v_\nu(T,x_\nu,x_{-\nu}) = \psi_\nu(T,x_\nu).
\end{cases} \quad (17)
$$

Now we have for each player one partial differential equation, which is parameterized by the rivals' states and controls, and is subject to the parameterized VI. We are going to collect all such equations into one system, whose solution may give an equilibrium state of the multi-agent optimal control. Denote

$$
\Xi(u) = \prod_{\nu=1}^N \Xi_\nu(u_{-\nu}), \quad \big(\mathrm{R}^m \rightrightarrows \mathbb{R}^m\big)
$$

here we mention that Ξ is a set-valued mapping. Collecting all the value functions, we have the value function profile:

$$
V(t,x) = \big(v_\nu(t,x)\big)_{\nu=1}^N, \quad \big([0,T] \times \mathbb{R}^n \to \mathbb{R}^N\big)
$$

which is to be computed. Collecting all the terminal payoff, we have the profile:

$$\Psi(t,x) = \big(\psi_\nu(t,x_\nu)\big)_{\nu=1}^N. \quad \big([0,T]\times\mathbb{R}^n \to \mathbb{R}^N\big)$$

Collecting the functions defining the HJB equations, we have

$$F(t,x,u,V) = \big(H_\nu(t,x,u,\nabla_{x_\nu} v_\nu, \Delta_{x_\nu} v_\nu)\big)_{\nu=1}^N,$$

and collecting the functions defining the parameterized VIs, we have

$$G(t,x,u,V) = \big(\nabla_{u_\nu} H_\nu(t,x,u,\nabla_{x_\nu} v_\nu, \Delta_{x_\nu} v_\nu)\big)_{\nu=1}^N.$$

Then by concentrating the HJB equations of the form (17) and the parameterized VIs for all the players, we have the following system

$$\begin{cases} \frac{\partial V(t,x)}{\partial t} + F(t,x,u,V) = 0, \\ u \in \mathrm{SOL}(\Xi(u), G(t,x,u,V)) \\ V(T,x) = \Psi(T,x). \end{cases} \tag{18}$$

Here $u \in \mathrm{SOL}(\Xi(u), G(t,x,u,V))$ is meant given t, x, V fixed, it holds

$$(z-u)^T G(t,x,u,V) \geq 0, \quad \forall z \in \Xi(u).$$

This is just a *quasi* variational inequality (QVI for short). Then (18) is a system of partial differential equations constrained by a QVI.

Because of the complex structure of $\Xi(u)$, it is hard to analyze the solvability and the convergence of numerical algorithms for solving (18). Here we try to propose a VI-based formulation, instead of the quasi one. Denote

$$\Xi = \big\{u \in \mathbb{R}^m | h_\nu(u_\nu) \leq 0, g(u_\nu, u_{-\nu}) \leq 0\big\}. \tag{19}$$

Assumption 1 ensures that Ξ is closed and convex. The following lemma states that the solvability of the VI implies the solvability of the quasi VI.

Lemma 1 (Facchinei et al. 2007) *For any fixed t, x and V, we have*

$$\mathrm{SOL}\big(\Xi, G(t,x,\cdot,V)\big) \subseteq \mathrm{SOL}\big(\Xi(u), G(t,x,\cdot,V)\big).$$

This lemma justifies the VI-based reformulation of the multi-agent optimal control problem:

$$\begin{cases} \frac{\partial V(t,x)}{\partial t} + F(t,x,u,V) = 0, \\ u \in \mathrm{SOL}(\Xi, G(t,x,u,V)) \\ V(T,x) = \Psi(T,x). \end{cases}$$

Moreover, using the projection formulation of the VI, we equivalently rewrite the above system of HJB equations as the following form, which is a system of partial differential equations, along with the boundary value conditions and the algebraic constraints:

$$\begin{cases} \frac{\partial V(t,x)}{\partial t} + F(t,x,u,V) = 0, \\ u = \mathrm{Pr}_{\Xi}(u - G(t,x,u,V)) \\ V(T,x) = \Psi(T,x). \end{cases} \tag{20}$$

Note that the algebraic constraints is defined by a system of equations that is nonsmooth, as the projection operator is nonsmooth.

2.3 Reformulation of Hamilton System

For the deterministic case: $\sigma_\nu(t, x_\nu, u_\nu) \equiv 0$ for $\nu = 1, \ldots, N$, we introduce the costate variable $p_\nu = \nabla_{x_\nu} v_\nu(t, x)$. Then for the ν-th player the Hamiltonian reads

$$H_\nu(t,x,u,p_\nu) = \langle f_\nu(t,x_\nu,u_\nu), p_\nu \rangle + \varphi_\nu(t,x,u),$$

and we have the following constrained Hamilton system

$$\begin{cases} \dot{p}_\nu(t) = -\nabla_{x_\nu} H_\nu(t,x,u,p_\nu) \\ \dot{x}_\nu(t) = \nabla_{p_\nu} H_\nu(t,x,u,p_\nu) \\ u_\nu(t) \in \mathrm{SOL}(\Xi_\nu(u_{-\nu}), \nabla_{u_\nu} H_\nu(t,x,\cdot,u_{-\nu},p_\nu)), \\ x_\nu(0) = x_\nu^0 \text{ and } p_\nu(T) = \nabla_{x_\nu} \psi_\nu(T, x(T)), \end{cases} \tag{21}$$

where we write $H_\nu(t,x,u,p_\nu) = H_\nu(t,x,u_\nu,u_{-\nu},p_\nu)$ for emphasizing the dependence of the mapping $H_\nu(t,x,u,p_\nu)$ on the rivals' control variables $u_{-\nu}$, and where $\nabla_{x_\nu} \psi_\nu(t,x_\nu)$ denotes the gradient of $\psi_\nu(t,x_\nu)$ with respect to x_ν.

Collectively write

$$G(t,x,u,p) = \left(\nabla_{u_\nu} H_\nu(t,x,u,p_\nu)\right)_{\nu=1}^N$$

and

$$\Gamma\big(x(0), p(0), x(T), p(T)\big) = \begin{pmatrix} x_\nu(0) - x_\nu^0 \\ p_\nu(T) - \nabla_{x_\nu} \psi_\nu(T, x(T)) \end{pmatrix}_{\nu=1}^N .$$

Concatenating (21) for $\nu = 1, \ldots, N$, we can formulate the multi-agent optimal control problem as the following differential quasi VI:

$$\begin{aligned} \dot{p}(t) &= \left(-\nabla_{x_\nu} H_\nu(t,x,u,p_\nu)\right)_{\nu=1}^N, \\ \dot{x}(t) &= \left(\nabla_{p_\nu} H_\nu(t,x,u,p_\nu)\right)_{\nu=1}^N, \\ u(t) &\in \mathrm{SOL}\big(\Xi(u), G(t,x,\cdot,p)\big) \\ 0 &= \Gamma\big(x(0), p(0), x(T), p(T)\big). \end{aligned} \tag{22}$$

Again, Lemma 1 implies a VI-based reformulation of the multi-agent optimal control problem:

$$
\begin{aligned}
\dot{p}(t) &= \big(-\nabla_{x_\nu} H_\nu(t,x,u,p_\nu)\big)_{\nu=1}^N,\\
\dot{x}(t) &= \big(\nabla_{p_\nu} H_\nu(t,x,u,p_\nu)\big)_{\nu=1}^N,\\
u(t) &\in \mathrm{SOL}\big(\Xi, G(t,x,\cdot,p)\big)\\
0 &= \Gamma\big(x(0),p(0),x(T),p(T)\big),
\end{aligned} \tag{23}
$$

where Ξ is defined by (19). Moreover, we use the projection operator to reformulate the system (23) into the following system of differential algebraic equations:

$$
\begin{aligned}
\dot{p}(t) &= \big(-\nabla_{x_\nu} H_\nu(t,x,u,p_\nu)\big)_{\nu=1}^N,\\
\dot{x}(t) &= \big(\nabla_{p_\nu} H_\nu(t,x,u,p_\nu)\big)_{\nu=1}^N,\\
u(t) &= \mathrm{Pr}_\Xi\big(u - G(t,x,u,p)\big)\\
0 &= \Gamma\big(x(0),p(0),x(T),p(T)\big).
\end{aligned} \tag{24}
$$

Write $\varphi_\nu(t,x,u) = \varphi_\nu(t,x_\nu,x_{-\nu},u_\nu,u_{-\nu})$. Suppose for any $\nu = 1,\dots,N$ that $\psi_\nu(T,\cdot)$ and each components of h_ν and $g(\cdot,u_{-\nu})$ are convex, and suppose that $\varphi_\nu(t,\cdot,x_{-\nu},\cdot,u_{-\nu})$ and each component of $\nabla_{p_\nu} H_\nu(t,x,u,p_\nu)$ are convex and continuously differentiable for any fixed $x_{-\nu}$ and $u_{-\nu}$, suppose that $\nabla_{p_\nu} H_\nu(t,x,u,p_\nu)$ is linear with respect to (x_ν,u_ν). Here we call (x,u) as a *feasible pair* of the multi-agent optimal control problem if $u \in \Xi$ and $\dot{x}_\nu(t) = \nabla_{p_\nu} H_\nu(t,x,u,p_\nu)$ for $\nu = 1,\dots,N$. Then we can show that

Theorem 1 *Let (x^*,u^*) be a weak solution of (23). Then (x^*,u^*) is a solution of the multi-agent optimal control problem in the following sense: for any feasible pair (x,u), we have for $\nu = 1,\dots,N$:*

$$
J_\nu\big(x_\nu(\cdot),x^*_{-\nu}(\cdot),u_\nu(\cdot),u^*_{-\nu}(\cdot)\big) \ge J_\nu\big(x^*_\nu(\cdot),x^*_{-\nu}(\cdot),u^*_\nu(\cdot),u^*_{-\nu}(\cdot)\big).
$$

Proof The details of the proof can be found in Chen and Wang (2013b). □

3 Approximation of Variational Inequality

The systems (20) and (24) concern the projection equation $u = \mathrm{Pr}_\Xi(u - G(t,x,u,V))$ and $u = \mathrm{Pr}_\Xi(u - G(t,x,u,p))$, respectively, which may have no solution, or have multiple (possibly infinitely many) solutions. Finding a solution of the systems involves solving optimization problems without standard constraint qualifications at each grid. Let $G : \mathbb{R}^m \times \mathbb{R}^k \to \mathbb{R}^m$ be given for defining the concerned parameterized VI, where the parameter vector is taken in the space of $\mathbb{R}^k$. Denote

$$
\Phi(u,\alpha) = u - \mathrm{Pr}_\Xi\big(u - G(u,\alpha)\big),
$$

where α is a parameter, and we are interested in finding for a given parameter α a vector u satisfying

$$\Phi(u,\alpha) = 0. \tag{25}$$

Here we propose a regularized smoothing method to find a solution of (25). Our main idea is to replace $\Phi(u,\alpha)$ by the following regularized and smoothing function

$$\Phi_{\lambda,\mu}(u,\alpha) = \int_R \big[u - \mathrm{Pr}_\Xi\big(u - G(u,\alpha) - \lambda u - \mu s e\big)\big]\rho(s)ds, \tag{26}$$

where $\lambda > 0$ and $\mu > 0$ are the regularization and smoothing parameters. The integration is performed componentwise with $e = (1, 1, \ldots, 1)^T$ and $\rho(\cdot)$ is a density function with

$$\kappa = \int_R |s|\rho(s)ds < \infty.$$

Suppose that $G(\cdot,\alpha)$ is monotone for any fixed α. Then when $\mu = 0$, the regularized system

$$\Phi_{\lambda,0}(u,\alpha) := u - \mathrm{Pr}_\Xi\big(u - G(u,\alpha) - \lambda u\big) = 0$$

has a unique solution u for any fixed α, but $\Phi_{\lambda,0}$ and u may not be differentiable with respect to (t,x). To overcome the non-smoothness of the projection operator, we adopt the smoothing approximation. The regularized smoothing function $\Phi_{\lambda,\mu}(u,\alpha)$ has the following properties

$$\big\|\Phi_{\lambda,0}(u,\alpha) - \Phi(u,\alpha)\big\|_2 \le \lambda\|u\|_2$$

and

$$\big\|\Phi_{\lambda,\mu}(u,\alpha) - \Phi_{\lambda,0}(u,\alpha)\big\|_2 \le \kappa\sqrt{m}\mu.$$

For fixed $\alpha \in \mathbb{R}^k$, $\lambda > 0$ and $\mu > 0$ the mapping $\Phi_{\lambda,\mu}(\cdot,\alpha)$ is continuously differentiable and the system

$$\Phi_{\lambda,\mu}(u,\alpha) = 0 \tag{27}$$

has a unique solution $u_{\lambda,\mu}(\alpha)$, which is continuously dependent on α. For $\lambda, \mu \downarrow 0$ (λ, μ chosen in an appropriate way—see also the second point in the summary) we approximate the solution of (25).

Namely, we approximate (20) by the following differential algebraic system

$$\begin{cases} \frac{\partial V(t,x)}{\partial t} + F(t,x,u,V) = 0, \\ \Phi_{\lambda,\mu}(t,x,u,V) = 0 \\ V(T,x) = \Psi(T,x), \end{cases} \tag{28}$$

and approximate (24) by the following differential algebraic system

$$
\begin{aligned}
\dot{p}(t) &= \big(-\nabla_{x_\nu} H_\nu(t,x,u,p)\big)_{\nu=1}^{N},\\
\dot{x}(t) &= \big(\nabla_{p_\nu} H_\nu(t,x,u,p)\big)_{\nu=1}^{N},\\
0 &= \Phi_{\lambda,\mu}(t,x,u,p)\\
0 &= \Gamma\big(x(0),p(0),x(T),p(T)\big).
\end{aligned}
\tag{29}
$$

We mention four points on this methodology.

- Finding an equilibrium point of the multi-agent optimal control problem is of the great practical importance, which is quite hard because the problem is coupled by the optimization problems, dynamical systems and the side constraints. Existing methods normally can not treat the generalized case: the problem with coupled strategy sets (see for example Krabs et al. 2000; Krabs 2005; Krabs and Pickl 2010). The methodology proposed here is promising since it reformulates the multi-agent optimal control problem as a differential algebraic equation, for which abundant theory and algorithms can be utilized. This new approach will be extended in the future.
- The convergence of the solution of the approximating system to the original one is of the most interest. Suitable notions of convergence, for example the Γ-convergence, have to be carefully selected. The convergence may be considerably dependent on the dependence between λ and μ, different dependence defines different regularized smoothing system, and therefore the different system of differential algebraic equations. Now we are in the position to touch the next point.
- Smoothing approximation and regularization have been studied extensively in solving the static VI (Facchinei and Pang 2003). However, to the best of our knowledge, the impact of the dependence between the smoothing and regularization parameters on the convergence behavior has not been studied. An example can be found in Chen and Wang (2013a, 2013b), which shows that for different relations of the two parameters λ, μ, the solution $u_{\lambda,\mu}(\alpha)$ of (27) can be divergent, or convergent to different solutions of the original projection equation (25). For the system (27), if $\mu = o(\lambda)$ is taken, then $u_{\lambda,\mu}(\alpha)$ is convergent to the least norm element of the solution set of (25). Note that finding the least norm solution is significant since it can provide a stable solution (refer to Chen and Wang 2013b, for more details).
- Our methodology is variational, which employs the comparison of solutions in a neighborhood of the optimal one to derive the necessary conditions, and to obtain candidate of the optimal solutions. Of course we need to impose additional conditions for ensuring the optimality. A different approach, namely the direct method, is also available, which offers global optima by using coordinate transformations instead of comparison techniques (Leitmann 1962). This method can also be applied to a class of differential games (Leitmann 1976). It is our aim to combine these destinguished approaches in the future.

- As mentioned before the multi-agent optimal control problem involves a certain decision process with multiple agents, where each agent solves an optimal control problem with the individual cost functional and strategy set. As a specialty the cost functional itself is dependent on all the other agents' state and/or control variables.

  In a forthcoming contribution we would like to apply this specific model to decision problems in the context of complex aviation management processes. It is our aim to apply the gained algorithms to the solution of concrete decision problems which occur in this innovative context of Operations Research.

Acknowledgements This work was supported by the *Bayerisches Hochschulzentrum für China* which made it possible that Zhengyu Wang visited in 2013 for a research stay the Universität der Bundeswehr München. We are thankful for their new scientific exchange program.

References

Chen, X., & Wang, Z. (2013a). Convergence of regularized time-stepping methods for differential variational inequalities. *SIAM Journal on Optimization*, *23*, 1647–1671.

Chen, X., & Wang, Z. (2013b). Differential variational inequality approaches to dynamic games with shared constraints. *Mathematical Programming Series A*. doi:10.1007/s10107-013-0689-1.

Facchinei, F., & Pang, J.-S. (2003). *Finite-dimensional variational inequalities and complementarity problems*. Berlin: Springer.

Facchinei, F., Fischer, A., & Piccialli, V. (2007). On generalized Nash games and variational inequalities. *Operations Research Letters*, *35*, 159–164.

Fleming, W., & Rishel, R. (1975). *Deterministic and stochastic optimal control*. Berlin: Springer.

Krabs, W. (2005). A dynamical method for the calculation of Nash-equilibria in n-person games. *Journal of Difference Equations and Applications*, *11*, 919–932.

Krabs, W., & Pickl, S. (2010). *Dynamical systems: stability, controllability and chaotic behavior*. Berlin: Springer.

Krabs, W., Pickl, S., & Scheffran, J. (2000). Optimization of an n-person game under linear side conditions. In *Optimization, dynamics, and economic analysis. Essays in honor of Gustav Feichtinger*. Heidelberg: Physica-Verlag.

Leitmann, G. (1962). *Optimization techniques, with applications to aerospace systems*. San Diego: Academic Press.

Leitmann, G. (1976). *Multicriteria decision making and differential games*. New York: Plenum.

Leitmann, G. (1981). *The calculus of variations and optimal control*. Heidelberg: Springer.

Øksendal, B. (2003). *Universitext*. *Stochastic differential equations: an introduction with applications* (6th ed.). Berlin: Springer.

Pang, J.-S., & Stewart, D. E. (2008). Differential variational inequalities. *Mathematical Programming*, *113*, 345–424.