

Cooperative stochastic differential games with information adaptation

David Yeung^{1,a,*} and Ovanes Petrosyan^{2,b}

¹*Center of Game Theory, St Petersburg State University, St Petersburg, 198504, Russia;*

SRS Consortium for Advanced Study in Cooperative Dynamic Games, Hong Kong Shue Yan University, Hong Kong

²*Faculty of Applied Mathematics and Control Process, St Petersburg State University, St Petersburg, 198504, Russia*

E-mail: ^adwkyeung@hksyu.edu, ^bpetrosian.ovanes@yandex.ru

**Corresponding author*

This paper develops a new class of stochastic differential games in which the state dynamics and the players' payoff structures in the future are unknown. Information about the future will be revealed stage by stage. Adapting to the new environment players revise their strategies accordingly. A new technique is derived for solving the optimal cooperative solution involving information adaptation. A novel theorem for the derivation of a Payoff Distribution Procedure (PDP) leading to a subgame consistent solution for this class of games is also presented.

Keywords: Stochastic Differential Games; Payoff Distribution Procedures; Information Adaptation.

1. Introduction

Technology development, new knowledge discovery, natural disasters, economic changes, scientific advancement and climate change often lead to outcomes that are unforeseeable. In reality, many events in the considerably far future are intrinsically unknown. Therefore, one cannot even obtain the probability distributions of these events. This paper develops a new class of cooperative stochastic differential games with intrinsic unknowns in the state dynamics and the players' future payoff structures. Information about the future will be revealed stage by stage. Adapting to the new information the players revise their strategies accordingly. In dynamic cooperative games, a stringent condition on cooperation and agreement is required: In the solution, the optimality principle must remain optimal throughout the game, at any instant of

time along the optimal state trajectory determined at the outset. This condition is known as subgame consistency (see Yeung and Petrosyan [1-2]).

This is the first time that stochastic differential games are expanded to accommodate intrinsic unknowns. It extends the Petrosyan [3] analysis to a framework with stochastic state dynamics with more general payoff structures. The present analysis also covers the Petrosyan and Yeung [4] game in the case where precise information on the probabilities of the unknown future events can be obtained. Of course, allowing the existence of intrinsic unknowns represents an expanded paradigm of the Petrosyan and Yeung [4] work with more realistic characteristics. Finally, the paper develops two novel mathematical theorems. First, a special class of stochastic dynamic programming technique is developed to obtain the expected joint payoff under cooperation with intrinsic unknowns. Second, a novel theorem for the derivation of a Payoff Distribution (PDP) for a subgame consistent solution for this class of games is presented.

2. Game Formulation and Joint Payoff Maximization

Consider the n – person cooperative stochastic differential game $\Gamma(x_0, t_0, T)$ with prescribed duration $[t_1, T]$ and initial state $x(t_1) = x^0 \in R^m$. The game duration can be partitioned into η time intervals -- $[t_1, t_2), [t_2, t_3), \dots, [t_\eta, T]$. The payoff structures and the state dynamics are not the same in these time intervals due to unpredictable future events. At initial time t_1 , the state dynamics and the payoff structures of the players in the intervals $[t_1, t_2)$ and $[t_2, t_3)$ are known. In addition, the best guess of the payoffs of the players covering the time interval $[t_3, T]$ are also known at initial time t_1 . When the game reaches time t_2 , new information about the players' payoff structures and the state dynamics in the interval $[t_3, t_4)$, and the best guesses of the payoffs of the players covering the interval $[t_4, T]$ become known. When the game reaches time t_k , for $k \in \{1, 2, \dots, \eta - 2\}$, new information about the players' payoff structures and the state dynamics in the interval $[t_{k+1}, t_{k+2})$, and the best guesses of the payoffs of the players covering the interval $[t_{k+2}, T]$ become known. The terminal payoff of player i at time T is known to be equal to $q^i(T, x)$.

Given that precise information are known for two time intervals, the game becomes a compressed game with 2 time intervals. Consider the case where the players agree to cooperate and maximized their expected joint payoff. In

particular, the compressed cooperative game in the time interval $[t_k, t_{k+2})$, for $k \in \{1, 2, \dots, \eta - 2\}$ can be expressed as maximizing the joint expected payoff

$$E \left\{ \sum_{j=1}^n \left[\int_{t_k}^{t_{k+1}} g^{k(j)}(s, x, u_1, u_2, \dots, u_n) ds + \int_{t_{k+1}}^{t_{k+2}} g^{k+1(j)}(s, x, u_1, u_2, \dots, u_n) ds \right] + \hat{W}^{(t_k)}(t_{k+2}, x(t_{k+2})) \right\}, \quad (1)$$

subject to state dynamics $dx = f^k(s, x, u_1, u_2, \dots, u_n) ds + \sigma^k(s, x) dz(s)$ in the time interval $[t_k, t_{k+1})$, and $dx = f^{k+1}(s, x, u_1, u_2, \dots, u_n) ds + \sigma^{k+1}(s, x) dz(s)$ in the time interval $[t_{k+1}, t_{k+2})$,

$$\text{for } k \in \{1, 2, \dots, \eta - 1\}, \quad (2)$$

where $\hat{W}^{(t_k)}(t_{k+2}, x)$ is the best guess of the expected joint payoff covering the time interval $[t_{k+2}, T]$.

A schematic diagram depicting the information flow and the formation of compressed games in the time interval $[t_k, t_{k+2})$, for $k \in \{1, 2, \dots, \eta - 1\}$, is provided in Figure 1 below.

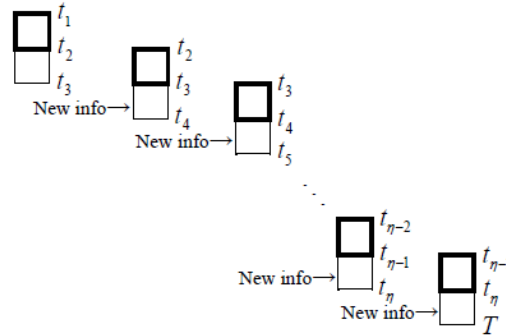


Fig. 1 Information Flow and Formation of Compressed Games

A theorem characterizing the optimal solution of the joint expected payoff problem of the cooperative game (1)-(2) can be derived as:

Theorem 1. An n -tuple of control strategies $\{u_i^* = \psi_i^{(k)k^*}(t, x) \in U^i ; i \in N \text{ and } t \in [t_{k-1}, t_k)\}$ and $\{u_i^* = \psi_i^{(k)k+1^*}(t, x) \in U^i ; i \in N \text{ and } t \in [t_k, t_{k+1})\}$ provides an optimal solution to the stochastic control problem (1)-(2) if there

exist functions $W^{(k)}(t, x) : [t_k, t_{k+2}] \times R^m \rightarrow R$, satisfying the partial differential equations

$$-W_t^{(k)}(t, x) - \frac{1}{2} \sum_{h, \zeta=1}^m \Omega^{h\zeta}(t, x) W_{x^h x^\zeta}^{(k)}(t, x) = \max_{u_1, u_2, \dots, u_n} \left\{ \sum_{j=1}^n g^{(k)j}(t, x, u_1, u_2, \dots, u_n) + W_x^{(k)}(t, x) f^k(t, x, u_1, u_2, \dots, u_n) \right\},$$

for $t \in [t_k, t_{k+1})$; (3)

$$-W_t^{(k)}(t, x) - \frac{1}{2} \sum_{h, \zeta=1}^m \Omega^{h\zeta}(t, x) W_{x^h x^\zeta}^{(k)}(t, x) = \max_{u_1, u_2, \dots, u_n} \left\{ \sum_{j=1}^n g^{(k+1)j}(t, x, u_1, u_2, \dots, u_n) + W_x^{(k)}(t, x) f^{k+1}(t, x, u_1, u_2, \dots, u_n) \right\},$$

for $t \in [t_{k+1}, t_{k+2})$; (4)

$$W^{(k)}(t_{k+2}, x) = \hat{W}^{(k)}(t_{k+2}, x), \text{ for } k \in \{1, 2, \dots, \eta - 1\}.$$

Proof. This result for $t \in [t_k, t_{k+1})$ and that for $t \in [t_{k+1}, t_{k+2})$ follow readily from the result of stochastic dynamic programming (see Basar and Olsder [5]) and therefore provide an optimal solution to the stochastic control problem (1)-(2).

Note that for the compressed control problem in the interval $[t_k, t_{k+2})$ information updating occurs at time instant t_{k+1} and the players will adapt themselves to a new compressed control problem for the time interval $[t_{k+1}, t_{k+3})$ and revise their optimal control in $[t_{k+1}, t_{k+2})$. We use X_t^* to denote the set of realizable values of $x^*(t)$ at time t generated by (3.4). The term $x_t^* \in X_t^*$ is used to denote an element in X_t^* . We use the terms $x^*(t)$ and x_t^* interchangeably in case where there is no ambiguity.

3. Subgame Consistent Payoff Distribution

To ensure that cooperation can be sustainable we have to seek a subgame consistent solution for cooperative stochastic differential games with intrinsic unknowns. At time $\tau \in [t_k, t_{k+1})$, we use $\xi^{(k)i}(\tau, x_\tau^*)$ denote the agreed-upon imputation of payoff to player i under cooperation over the time interval $[\tau, T]$.

We first consider the derivation of a Payoff Distribution Procedure (PDP) leading to a subgame consistent solution. In the compressed cooperative game in the time $[t_k, t_{k+2})$ according to the solution generated by the agreed-upon optimality principle. Following Yeung and Petrosyan [1] we define the procedure $B_i^{(k)k}(s, x_s^*)$ for $s \in [t_k, t_{k+1})$ and $B_i^{(k)k+1}(s, x_s^*)$ for $s \in [t_{k+1}, t_{k+2})$ which brings about an imputation to player i such that:

$$\xi^{(k)i}(t, x^*) = E \left\{ \int_t^{t_k} B_i^{(k)k}(s, x^*(s)) ds + \int_{t_k}^{t_{k+1}} B_i^{(k)k+1}(s, x^*(s)) ds + \hat{W}^{(t_k)}(t_{k+2}, x(t_{k+2})) \right\}, \text{ for } i \in N \text{ and } k \in \{1, 2, \dots, \eta\}, \quad (5)$$

where $x^*(s)$ is the solution path derived from Theorem 2.

One can readily verify that the payoff distribution procedure $B_i^{(k)k}(s, x_s^*)$ for $s \in [t_k, t_{k+1})$ and $B_i^{(k)k+1}(s, x_s^*)$ for $s \in [t_{k+1}, t_{k+2})$ for $i \in N$ and $k \in \{1, 2, \dots, \eta\}$. Hence the agreed upon optimality principle is satisfied and subgame consistency follows. A theorem yielding a PDP that leads to a subgame consistent solution is provided below.

Theorem 2. A PDP with an instantaneous payment at time $s \in [t_k, t_{k+1})$:

$$B_i^{(k)k}(s, x_s^*) = -\xi_s^{(k)i}(s, x_s^*) - \frac{1}{2} \sum_{h, \zeta=1}^m \Omega^{h\zeta}(s, x_s^*) \xi_{x_t^* x_t^*}^{(k)i}(s, x_s^*) - \xi_{x_t^*}^{(k)i}(s, x_s^*) f^k[s, x_s^*, \psi_1^{(k)k*}(s, x_s^*), \psi_2^{(k)k*}(s, x_s^*), \dots, \psi_n^{(k)k*}(s, x_s^*)], \quad (6)$$

for $i \in N$, $x_s^* \in X_s^*$ and $s \in [t_k, t_{k+1})$; with the condition that

$$B_i^{(k)k+1}(s, x_s^*) = -\xi_s^{(k+1)i}(s, x_s^*) - \frac{1}{2} \sum_{h, \zeta=1}^m \Omega^{h\zeta}(s, x_s^*) \xi_{x_t^* x_t^*}^{(k+1)i}(s, x_s^*), - \xi_{x_t^*}^{(k+1)i}(s, x_s^*) f^{k+1}[s, x_s^*, \psi_1^{(k)k+1*}(s, x_s^*), \psi_2^{(k)k+1*}(s, x_s^*), \dots, \psi_n^{(k)k+1*}(s, x_s^*)] \text{ for } i \in N, x_s^* \in X_s^* \text{ and } s \in [t_{k+1}, t_{k+2}); \quad (7)$$

yields imputation vector $\xi_i^{(k)}(\tau, x_\tau^*)$, for $\tau \in [t_k, t_{k+1})$.

Proof. The instantaneous payment in (6) satisfies the condition of subgame consistency in the time interval $[t_k, t_{k+1})$ proved in Yeung and Petrosyan [1]. The instantaneous payment in (7) satisfies the condition of subgame consistency

in the time interval $[t_{k+1}, t_{k+2})$ proved in Yeung and Petrosyan [1]. Hence Theorem 2 yields a PDP that leads to a subgame consistent solution of the cooperative stochastic differential game with intrinsic unknowns (1)-(2).

Note that the payoff distribution procedure in (6)-(7) would lead to the realization of $\xi^{(k)i}(t, x_i^*)$, for $t \in [t_k, t_{k+1})$ and $\xi^{(k+1)i}(t, x_i^*)$, for $t \in [t_{k+1}, t_{k+2})$. Hence a subgame consistent solution is obtained.

In the compressed game in the interval $[t_k, t_{k+1})$, new information will be revealed at time t_{k+1} . Therefore only $B_i^{(k)k}(s, x_s^*)$ for $i \in N$ will be distributed in the interval $[t_{k+1}, t_{k+2})$. Adaptation will occur as new information is revealed and a new compressed control problem in the interval $[t_{k+1}, t_{k+3})$ will appear. The payoff distribution $B_i^{(k+1)k}(s, x_s^*)$ for $i \in N$ in the new problem in the interval $[t_{k+1}, t_{k+3})$ will be distributed in the interval $[t_{k+2}, t_{k+3})$.

4. Conclusion

This paper develops a new class of stochastic differential games with intrinsic unknowns in the state dynamics and the players' future payoff structures. New techniques involving information adaptation is derived for solving the optimal solution of cooperative stochastic differential games. A novel theorem for the derivation of a Payoff Distribution Procedure (PDP) leading to a subgame consistent solution for this class of games is also presented. The game represents an extension of Petrosyan [3] analysis with stochastic dynamics and more general payoff structures. The present analysis provides an expanded framework to the Petrosyan and Yeung [4] game with the more realistic element of intrinsic unknowns. Further research on differential games with the realistic elements of intrinsic unknown future events is expected.

References

1. D.W.K. Yeung, L.A. Petrosyan: Subgame Consistent Cooperative Solutions in Stochastic Differential Games, *Journal Of Optimization Theory And Applications*: 120, 651-666 (2004).
2. D.W.K. Yeung, L.A. Petrosyan: Subgame Consistent Economic Optimization: An Advanced Cooperative Dynamic Game Analysis, (Boston: Birkhäuser, 2012).
3. O. Petrosyan: Looking Forward Approach in Cooperative Differential Games, forthcoming in *International Game Theory Review*, **18**, (2016).

4. L.A. Petrosyan, D.W.K. Yeung: Subgame-consistent Cooperative Solutions in Randomly-furcating Stochastic Differential Games, *International Journal of Mathematical and Computer Modelling*, 45, 1294-1307 (2007).
5. T. Basar and Olsder, G.J.: *Dynamic Noncooperative Game Theory*, 2nd edn. (Academic Press, London, 1995).