Joint 48th IEEE Conference on Decision and Control and
28th Chinese Control Conference
Shanghai, P.R. China, December 16-18, 2009

WeA05.3

Learning Approaches to the Witsenhausen Counterexample From a View of Potential Games

Na Li, Jason R. Marden and Jeff S. Shamma

Abstract—Since Witsenhausen put forward his remarkable counterexample in 1968, there have been many attempts to develop efficient methods for solving this non-convex functional optimization problem. However there are few methods designed from game theoretic perspectives. In this paper, after discretizing the Witsenhausen counterexample and re-writing the formulation in analytical expressions, we use *fading memory JSFP with inertia*, one learning approach in games, to search for better controllers from a view of potential games. We achieve a better solution than the previously known best one. Moreover, we show that the learning approaches are simple and automated and they are easy to extend for solving general functional optimization problems.

I. INTRODUCTION

A team decision problem consists of a group of decision makers seeking to maximize a common objective that depends on the group's joint decision. The difficulty associated with a team decision problem stems from the fact that each decision maker is making a decision independently in response to incomplete information. Decision makers are allowed to communicate their information to one another within a given information structure; however, such actions bear communication costs. The goal of the team decision problem is to find the optimal policy for the decision makers and the optimal information structure such as to minimize a cost function that incorporates the original objective, the available information, and the communication costs [15], [19].

One example of a team decision problem that has received a significant degree of research attention is the Witsenhausen counterexample (WC). The WC is illustrated in Figure 1 and has the following elements[1]:

- *External Signals*: x, v, independent random variables with finite second moments. In this paper, we assume independent Gaussian random variables, where $x \sim N(0, \delta^2)$; $v \sim N(0, 1)$.
- *Information (Observation)*: x, y, where $y = u_1 + v$. and u_1 is explained below.

[1]It's not exactly the original WC, while it's equivalent to WC [22].
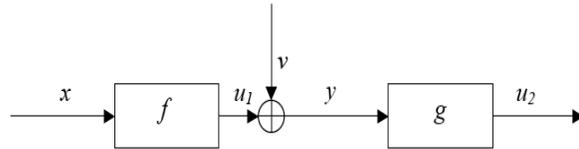
Fig. 1. Information structure of the Witsenhausen Counterexample

- *Decision Variables*: $u_1 = f(x)$, $u_2 = g(y)$, where (f, g) is any pair of Borel functions.
- *Cost objective:*

$$\min_{f, g} J = \mathbf{E}\left[k^2(u_1 - x)^2 + (u_2 - u_1)^2\right] \qquad (1)$$

The goal is for DM_1 to estimate the external signal x and DM_2 to estimate u_1 which is corrupted by the noise signal v. Although the WC involves only two decision makers, it possesses almost all of the main difficulties inherent in any decentralized team decision problem.

The WC is a simple example of a linear-quadratic-Gaussian (LQG) team problem [8]. Before Witsenhausen put forth this counterexample, it was conjectured that in any LQG team problem the optimal controllers are linear. Witsenhausen proved that for some $k > 0$, the WC has an optimal solution that is not of the linear type [22]. Thus he claimed that the conjecture is not necessarily true if the information pattern is not classical.[2] Since then, many researchers have focused on understanding the role of information structures in team decision problems ([1], [8], [10], [11], [20], [21]) and on developing more efficient methods to find improved solutions for the WC ([2], [4], [12], [13]). The work in this paper belongs to the second kind. So far no optimal controller has been found.

Witsenhausen proved that the optimal controllers for the WC have the following properties [22]: (i) if f is optimal, then $\mathbf{E}[f(x)] = 0$, $\mathbf{E}[f(x)^2] \leq 4\delta^2$; (ii) given a fixed $f(x)$ having zero means and variance not exceeding $4\delta^2$, the optimal choice for $g(\cdot)$ is given by

$$g_f^*(y) = \mathbf{E}[f(x)|y] = \frac{\mathbf{E}_x\left[f(x)\phi(y - f(x))\right]}{\mathbf{E}_x\left[\phi(y - f(x))\right]}, \qquad (2)$$

where $\phi(\cdot)$ is the standard Gaussian density. The corresponding payoff function is

$$J^*(f) := J(f, g_f^*) = k^2\mathbf{E}\left[(x - f(x))^2\right] + 1 - I(D_f) \quad (3)$$

[2]If the information pattern is classical, then the information available to earlier decision makers is also available to later decision makers [8], [22].

where

$$I(D_f) = \int \left(\frac{d}{dy} D_f(y) \right)^2 \frac{dy}{D_f(y)} \qquad (4)$$

which is called the *"Fisher Information"* of the random variable y with density $D_f(y)$ being

$$D_f(y) = \int \phi(y - f(x))\phi(x; \ 0, \delta^2)dx \qquad (5)$$

where $\phi(x; \ 0, \ \delta^2)$ is the Gaussian Density with zero mean and variance δ^2.

By utilizing both of the two properties, we can convert the problem of minimizing the cost over a pair of functions (f, g) to the one of minimizing the cost over a single function f. Most of the past attempts to find improved controllers have utilized this idea. Moreover, because of the first property, most work has only investigated functions f that are symmetric about the origin.[3] The main difficulty of finding optimal f is that $J^*(f)$ is not a convex functional, since the fisher information $I(D_f)$ is convex [3] and thus $-I(D_f)$ is concave.

Following the signal scheme suggested in Witsenhausen's original paper [22], many authors have focused on using step functions or functions with other bases to approximate the optimal controllers during the past 40 years. For example, in [9] the authors analyzed the problem from a discrete version; however, the authors were unable to make progress in deriving the optimal solution. Later, it was proven that the discretized Witsenhausen counterexample expressed in discrete form is NP-complete [18]. Table I provides a brief summary of the major advances on solving the WC. For benchmarking, we choose the case when $\sigma = 5$ and $k = 0.2$.

TABLE I

A BRIEF SUMMARY OF MAJOR ADVANCES ON SOLVING THE WC

Solution of $f(x)$ [a]	Total cost J [b]
Optimal affine solution [22] (1968)	0.961852
1-step; by Witsenhausen [22] (1968)	0.404253
1-step; by Bansal and Basar [1] (1987)	0.365015
2-step; by Deng and Ho [4] (1999)	0.190
25-step; by Ho and Lee [12] (2000)	0.1717
2.5-main-step; by Baglietto *et al.* [2] (2001)	0.1701
3.5-main-step; by Lee *et al.* [13] (2001)	0.1673132
3.5-main-step; by our work in this paper (2009)	0.1670790

[a]All $f(x)$ are odd functions, thus "0" is always counted as a breakpoint. "n-step" means that there are n breakpoints in the nonnegative domain of x. "n.5-step" means that $f(x) = 0$ for any $x \in [0, b_1]$, where b_1 is the smallest positive breakpoint.

[b]If the corresponding paper provided J's value for the case of $\sigma = 5$, $k = 0.2$, the value is from the paper. Otherwise, we evaluate J with our numerical methods provided in Section III, whose accuracy is at least gauranteed to be 10^{-6}.

While most of the approaches highlighted in Table I rely on utilizing step functions, [2] demonstrates that the optimal controller $f^*(x)$ is not necessarily strictly piecewise constant but rather slightly sloped. [13] utilized this finding in

[3]There are further discussions on the reason why only symmetric functions need to be considered in [13]-Appendix III.

constructing controllers by first deciding on the number and positions of the main steps and then modifying each main step to be several sub-steps to obtain improved solutions. This explains the meaning of "n-main-step" in Table I.

While several types of numerical methods have been employed to find efficient controllers for the WC, few methods are designed from the perspective of learning in games ([6], [7], [14], [16], [17], [23]). Learning approaches are distributed algorithms designed to find Nash Equilibria in games. In this paper, we discretize the WC and formulate this problem as a game. We utilize the learning algorithm *fading memory joint strategy fictitious play(JSFP) with inertia* [14] to search for an efficient controller. This learning approach provides a controller that improves upon the best known controller in the past 40 years as highlighted in Table I. Furthermore, learning algorithms such as JSFP have alternative advantages over other searching methods as to the issues of complexity, flexibility, as well as generality.

The remainder of the paper is organized as follows. In Section II, we offer some basic background of game theory and introduce the learning approach *fading memory JSFP with inertia*. In Section III, we discretize the WC and derive an analytic formulation for the problem based on the discretizing parameters. We formulate the WC as a potential game, which is a special case of games, and employ the algorithm of *fading memory JSFP with inertia* to find an efficient controller. In Section IV we compare our methods and results with others'. We conclude in Section V.

II. LEARNING APPROACHES IN POTENTIAL GAMES

In this section, we introduce a brief background on the game theoretic concepts and learning approaches in this paper. We refer the readers to [5], [6], [7], [14], [16], [17], [23] for a more comprehensive review.

A. Background of Game Theory

1) Finite Strategic-Form Games: In a finite strategic-form game, there are n-players $\mathcal{N} := \{1, 2, \cdots, n\}$ where each player $i \in \mathcal{N}$ has a finite action set $\mathcal{A}_i$ and a utility function $U_i : \mathcal{A} \to \mathbb{R}$ where $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2 \times \ldots \times \mathcal{A}_n$. Every player i seeks to selfishly maximize their utility. We use G to represent the entire game, i.e. the players, actions sets, and utilities.

For an action profile $a = (a_1, a_2, \ldots, a_n) \in \mathcal{A}$, a_{-i} denotes the action profile of players *other than* player i, i.e., $a_{-i} = (a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$. The definition of pure Nash equilibrium is as follows:

Definition 2.1: $a^* \in \mathcal{A}$ is called a *pure Nash equilibrium* if for all players $i \in \mathcal{N}$, $U_i(a_i^*, a_{-i}^*) = \max_{a_i \in \mathcal{A}_i} U_i(a_i, a_{-i}^*)$.

2) Potential Games: In an identical interest game, all the players have a common utility function, i.e., $U_i(a) = U_g(a)$ for some global utility function $U_g : \mathcal{A} \to \mathbb{R}$. Hence, every identical interest game has at least one pure Nash equilibrium, namely any action profile a that maximizes $U_g(a)$. Potential games are a generalization of identical interest games.

Definition 2.2: A game G is a potential game, iff $\exists$ a function $\Phi : \mathcal{A} \rightarrow \mathbb{R}$ such that for every player i and $\forall\, a_{-i} \in \mathcal{A}_{-i}, \forall\, a_i', a_i'' \in \mathcal{A}_i$,

$$U_i(a_i', a_{-i}) - U_i(a_i'', a_{-i}) = \Phi(a_i', a_{-i}) - \Phi(a_i'', a_{-i}) \quad (6)$$

The function Φ is called a potential for G.

In a potential game, the change in a player's utility resulting from a unilateral change in strategy equals the change in some global potential function. It is easy to verify that any maximum of the potential is a pure Nash equilibrium of the potential game.

3) Repeated Games: In a repeated game, at each time $t = 0, 1, 2, \cdots$, each player $i \in \mathcal{N}$ simultaneously chooses an action $a_i(t) \in \mathcal{A}_i$ and receives the utility $U_i(a(t))$ where $a(t) := (a_1(t), a_2(t), \ldots, a_n(t))$. Each player $i \in \mathcal{N}$ chooses his action $a_i(t)$ at time t according to a probability distribution $p_i^t \in \Delta(\mathcal{A}_i)^4$, which is a function of the information available to the player i up to time t which includes observations from the games played at times $\{0, 1, ..., t-1\}$. In the most general form, a strategy update mechanism for player i takes on the form

$$p_i^t = F_i(a(0), a(1), ..., a(t-1); U_i),$$

meaning that the strategy update mechanism could depend on all past information in addition to the structural form of a player's utility. Different learning algorithms are specified by both the assumption on available information and mechanism by which p_i^t are updated.

B. Fading memory JSFP with inertia

Firstly consider the learning algorithm joint strategy fictitious play (JSFP) with inertia [14]. Define $V_i^{a_i}(t)$ as the average utility player i would have received up to time t if player i selected action a_i at all previous time steps and the actions of the other players remained unchanged

$$V_i^{a_i}(t) := \frac{1}{t} \sum_{\tau=0}^{t-1} U_i(a_i, a_{-i}(\tau)).$$

The average utility admits the following simple recursion:

$$V_i^{a_i}(t) = \frac{t-1}{t} V_i^{a_i}(t-1) + \frac{1}{t} U_i(a_i, a_{-i}(t-1)) \quad (7)$$

Therefore, each player can maintain this average utility vector using minimal computations.

At the time $t = 0$, each player i randomly selects one action from his action set $\mathcal{A}_i$. While at each time $t > 0$, if $a_i(t-1) \in \arg\max_{\bar{a}_i \in \mathcal{A}_i} V_i^{\bar{a}_i}(t)$, player i select the previous action $a_i(t-1)$; otherwise randomly selects any action $a_i(t) \in \arg\max_{\bar{a}_i \in \mathcal{A}_i} V_i^{\bar{a}_i}(t)$ with probability $(1 - \epsilon)$ or selects the previous action $a_i(t) = a_i(t-1)$ with probability ϵ where $\epsilon \in (0, 1)$.

$^4\Delta(\mathcal{A}_i) = \{p_i : p_i \text{ is a probaility distribution on } \mathcal{A}_i, \text{ i.e., } \forall a_i \in \mathcal{A}_i, 0 \le p_i(a_i) \le 0; \& \sum_{a_i \in \mathcal{A}_i} p_i(a_i) = 1.\}.$

In this paper, we consider the learning algorithm fading memory JSFP with inertia. In this setting, each player maintains a weighted average utility, denoted as $\tilde{V}$, as opposed to the true average utility as in JSFP.

$$\tilde{V}_i^{a_i}(t) = (1 - \rho) \sum_{\tau=0}^{t-1} \rho^{t-1-\tau} U_i(a_i, a_{-i}(\tau)) \quad (8)$$

$$\tilde{V}_i^{a_i}(t) = \rho \tilde{V}_i^{a_i}(t-1) + (1 - \rho) U_i(a_i, a_{-i}(t-1)) \quad (9)$$

where $\rho \in [0, 1)$ is referred to as the player's discount factor.

The mechanism of selecting actions in fading memory JSFP with inertia is the same with that in JSFP with inertia. If all players adhere to the prescribed learning rule fading memory JSFP with inertia, then the action profile, $a(t)$, will converge to a pure Nash equilibrium almost surely in all potential games [14]. In this paper, we set the inertia $\epsilon = 0.6$ and discount factor $\rho = 0.8$.

III. A GAME THEORETIC APPROACH TO THE WITSENHAUSEN COUNTEREXAMPLE

In this section, we formulate the WC as a potential game and use the learning approach fading memory JSFP with inertia to find an efficient controller. One approach to formulating the WC as a potential game is to model it as a game between the two decision makers where the actions available to each decision maker are the set of possible control laws. This approach leads to some challenges since the cardinality of each action set is infinite. Hence, we formulate this WC as a potential game in an alternative fashion. We use n-step functions for $f(x)$ and model each interval as a player and the value taken by each interval as the player's action. The number of players and the size of action sets are determined by the way of discretizing the problem. Moreover, in order to reduce the burden of computation, we provide a local utility function for each player which is exactly aligned with the global utility function. For benchmarking, we choose $\sigma = 5, k = 0.2$.

A. Problem Re-formulation

Based on the first property of the optimal controllers for the WC in Section I and Appendix III in [13], we assume that the optimal $f^*(x)$ is symmetric about the origin and hence we only investigate odd functions $f(x)$. We discretize DM_1's nonnegative information and decision variables spaces as follows:

1) Expression for $f(x)$: We just discretize the main ranges of DM_1's information and decision variables domain, i.e. the two intervals $[-\mathrm{range}_b \cdot \delta, \mathrm{range}_b \cdot \delta]$ and $[-\mathrm{range}_a \cdot \delta, \mathrm{range}_a \cdot \delta]$. Moreover, we choose $\mathrm{range}_b = 4, \mathrm{range}_a = 4.5$.

$$f(x) = \begin{cases} a_1 & (0 = b_0 \le x < b_1) \\ a_2 & (b_1 \le x < b_2) \\ \vdots & \vdots \\ a_n & (b_{n-1} \le x < b_n) \\ a_{n+1} & (b_n \le x < b_{n+1} = \infty) \\ -f(-x) & x < 0 \end{cases} \quad (10)$$

in which $b_i = \text{range}_b \cdot \frac{\delta}{n}i$, $i = 0, 1, 2, \cdots, n$, $b_{n+1} = \infty$ and $a_i \in \{a | a = \text{range}_a \cdot \frac{\delta}{m}k, \ k = 0, 1, \cdots, m\}$, $i = 1, 2, \cdots, n+1$. The probabilities at these point intervals are $q_1, \ldots, q_{n+1}$ respectively, where

$$q_i = \int_{b_{i-1}}^{b_i} \phi(s; 0, \delta^2) \ ds = \frac{1}{2}(erf(\frac{b_i}{\sqrt{2}\delta}) - erf(\frac{b_{i-1}}{\sqrt{2}\delta})), \tag{11}$$

and $erf(\cdot)$ is the standard error function.

Remark 3.1: We discretize the two main ranges into intervals with *equal size* to make the problem easier to handle. Note that this approach does not require this intervention.

2) Expression for $g_f^*(y)$: By equation (2), we get

$$g_f^*(y) = \frac{-\sum_{i=1}^{n+1} q_i a_i \phi(y + a_i) + \sum_{i=1}^{n+1} q_i a_i \phi(y - a_i)}{\sum_{i=1}^{n+1} q_i \phi(y + a_i) + \sum_{i=1}^{n+1} q_i \phi(y - a_i)}. \tag{12}$$

Moreover, we have

$$g_f^*(y) = \frac{\sum_{i=1}^{n+1} q_i e^{-a_i^2} (e^{a_i y} - e^{-a_i y}) a_i}{\sum_{i=1}^{n+1} q_i e^{-a_i^2} (e^{a_i y} + e^{-a_i y})}. \tag{13}$$

By writing $e^{a_i y} - e^{-a_i y} = 2\sinh(a_i y)$ and $e^{a_i y} + e^{-a_i y} = 2\cosh(a_i y)$, we get

$$g_f^*(y) = \frac{\sum_{i=1}^{n+1} q_i e^{-a_i^2} a_i \sinh(a_i y)}{\sum_{i=1}^{n+1} q_i e^{-a_i^2} \cosh(a_i y)}. \tag{14}$$

3) Expression for $J(f)$: It is easy to compute the DM_1's cost,

$$k^2 \mathbf{E}\left[(x - f(x))^2\right] = 1 + 2\sum_{i=1}^{n+1} q_i k^2 a_i^2 - \frac{4}{2\pi\delta^2} \sum_{i=1}^{n+1} a_i \left[exp(-\frac{b_{i-1}^2}{2\delta^2}) - exp(-\frac{b_i^2}{2\delta^2})\right]. \tag{15}$$

The expression for $D_f(y)$ is:

$$D_f(y) = \frac{2}{2\pi} e^{-\frac{y^2}{2}} \sum_{i=1}^{n+1} q_i e^{-\frac{a_i^2}{2}} \cosh(a_i y). \tag{16}$$

The expression for $\frac{d}{dy}D_f(y)$ is:

$$\frac{d}{dy}D_f(y) = \frac{2}{2\pi} e^{-\frac{y^2}{2}} \sum_{i=1}^{n+1} q_i e^{-\frac{a_i^2}{2}} a_i \sinh(a_i y) - yD_f(y). \tag{17}$$

Using $J = k^2 \mathbf{E}\left[(x - f(x))^2\right] + 1 - \int(\frac{d}{dy}D_f(y))^2 \frac{dy}{D_f(y)}$ we can evaluate the objective value.

Remark 3.2: In order to calculate the cost J, we need to compute an integration. We choose *"Numerical Gaussian Integration"*, i.e., dividing the main integration interval $\left[-\text{range}_y, \text{range}_y\right]$ into q subintervals, and for each interval, using 16 degree Gaussian Integration. Since $y = u_1 + v$, where $u_1 \in [-\text{range}_a \cdot \delta, \text{range}_a \cdot \delta]$ and $v \sim N(0, 1)$, we choose $\text{range}_y = 5 + \text{range}_a \cdot \delta = 5 + 4.5 \cdot \delta$. Choose proper large r ($r = 200$ in this paper) to ensure the integration error smaller than 10^{-8} over the main integration domain. Using the solution provided in [13] with our numerical gaussian integration, we obtain the same value of total cost which is $J = 0.1673132$. Our numerical method provide a good comparable accuracy.

B. Learning Approaches to the WC

We formulate the WC as a game:
1) *"Players"*: Each interval $[b_{i-1}, b_i)$, $(i = 1, \cdots, n+1)$ can be considered to be a player.
2) *"Action"*: The value $a_i \in \{a | a = \text{range}_a \cdot \frac{\delta}{m}k, \ k = 0, 1, \cdots, m\}$ of $f(x)$ on each interval $[b_{i-1}, b_i)$ is the action taken by each player.
3) *"Utility Function"*: For each player, we have an identical utility function $U(= -J)$ to maximize, which is

$$U = \frac{4}{2\pi\delta^2} \sum_{i=1}^{n+1} a_i \left[exp(-\frac{b_{i-1}^2}{2\delta^2}) - exp(-\frac{b_i^2}{2\delta^2})\right] - 2\sum_{i=1}^{n+1} q_i k^2 a_i^2 + \int(\frac{d}{dy}D_f(y))^2 \frac{dy}{D_f(y)} - 2. \tag{18}$$

Note that the first two summations at the right side of equation (18) can be separated by items depending only on each player's own action a_i. Given this property and in order to reduce the number of calculation, we can regard the utility function as a *potential* of a *potential game* where the local utility function for each player is:

$$U_i(\bar{a}_i, a_{-i}) = \frac{4}{2\pi\delta^2} \bar{a}_i \left[exp(-\frac{b_{i-1}^2}{2\delta^2}) - exp(-\frac{b_i^2}{2\delta^2})\right] - 2\sum_{i=1}^{n+1} q_i k^2 \bar{a}_i^2 + \int(\frac{d}{dy}D_f(y))^2 \frac{dy}{D_f(y)} \tag{19}$$

with the integral part in the right hand side depending on the joint action file of all the players namely $a_1, a_2, \ldots, a_{n+1}$ as shown in equations of (16) and (17).

It is easy to verify that the local utility is perfectly aligned with the original utility. This implies that the Nash equilibria for the potential game are the same as those for the original problem. Note that calculating the local utility is easier and faster than calculating the global utility.

C. Results

Pick $n = 400; m = 450$, and use *"fading memory JSFP with inertia"* , after about 750 steps of iterations, we get the solution shown in Table II. The corresponding cost is $J = 0.1671507$. Figure 2 shows the convergence processes of a_i, $(i = 1, 2, \cdots, n+1)$ and the total cost J.

TABLE II

SOLUTION OF $f(x)$ [a]; $n = 400; m = 450$

$f(x)$	Interval	$f(x)$	Interval
0.00	$0.00 \leq x < 0.70$	13.25	$11.15 \leq x < 12.65$
0.05	$0.70 \leq x < 2.10$	13.30	$12.65 \leq x < 14.10$
0.10	$2.10 \leq x < 3.30$	13.35	$14.10 \leq x < 15.55$
6.45	$3.30 \leq x < 4.25$	13.40	$15.55 \leq x < 16.85$
6.50	$4.25 \leq x < 5.70$	20.25	$16.85 \leq x < 17.65$
6.55	$5.70 \leq x < 7.10$	20.30	$17.65 \leq x < 19.20$
6.60	$7.10 \leq x < 8.55$	20.35	$19.20 \leq x < 20$
6.65	$8.55 \leq x < 9.95$	20.40	$x \geq 20$
13.20	$9.95 \leq x < 11.15$	$-f(-x)$	$x < 0$

[a]Although we discretize the main information domain into 400 intervals, those intervals can choose the same decision value. In this table, we show the intervals together which have the same decision value. We do the same process in Table III.
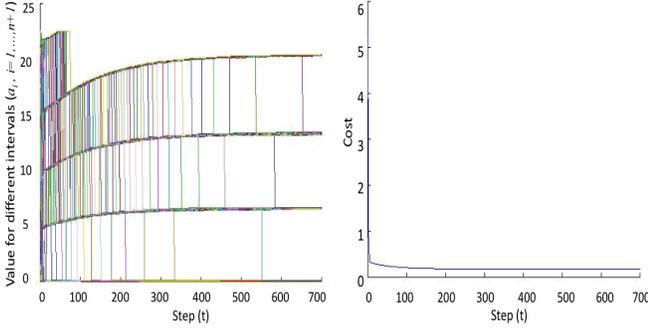
Fig. 2. Convergence processes of a_i and the total cost J in the case where $n = 400$, $m = 450$: each colored plot in the left figure shows one interval's decision value $f(x)$ at each step of iteration. There are totally 400 colored plots in the left figure since we discretize $f(x)$ into 400 steps. The right one shows the convergence of the total cost J.

Fig. 3. Convergence processes of $a_i, (i = 1, 2, \cdots, n + 1)$ in the case where $n = 600$; $m = 675$: the left one shows the first 400 steps of iteration; the right one shows the later 400 steps of iteration. Each colored plot shows one interval's decision value $f(x)$ at each step of iterations. There are totally 600 colored plots since we discretize $f(x)$ into 600 steps.

Change the discretizing parameters to be $n = 600$; $m = 675$, and use *"fading memory JSFP with inertia"*, after about 850 steps of iterations, we get the solution shown in table III. The corresponding cost is $J = 0.1670790$. Figure 3 and

TABLE III
SOLUTION OF $f(x)$; $n = 600$; $m = 675$

$f(x)$	Interval	$f(x)$	Interval
0.00	$0.00 \leq x < 0.467$	13.233	$10.667 \leq x < 11.667$
0.033	$0.467 \leq x < 1.40000$	13.267	$11.667 \leq x < 12.633$
0.067	$1.400 \leq x < 2.333$	13.300	$12.633 \leq x < 13.633$
0.100	$2.333 \leq x < 3.333$	13.333	$13.633 \leq x < 14.600$
6.467	$3.333 \leq x < 4.133$	13.367	$14.600 \leq x < 15.567$
6.500	$4.133 \leq x < 5.100$	13.400	$15.567 \leq x < 16.533$
6.533	$5.100 \leq x < 6.033$	13.433	$16.533 \leq x < 16.867$
6.567	$6.033 \leq x < 7.000$	20.267	$16.867 \leq x < 17.531$
6.600	$7.000 \leq x < 7.933$	20.300	$17.531 \leq x < 18.567$
6.633	$7.933 \leq x < 8.867$	20.333	$18.567 \leq x < 19.600$
6.667	$8.867 \leq x < 9.833$	20.367	$19.600 \leq x < 20$
6.700	$9.833 \leq x < 9.967$	20.400	$x \geq 20$
13.200	$9.967 \leq x < 10.667$	$-f(-x)$	$x < 0$

4 show the convergence processes of a_i $(i = 1, 2, \cdots, n+1)$ and the total cost J respectively. In order to show the convergence dynamics more clearly, we use two separate plots for each process.

Figure 5 plots the two different fs which we obtain by using different algorithm parameters. We can see that in both cases, f is a monotone nondecreasing function with 3.5 main steps, each of which has a slightly sloped piece between discontinuities.

IV. COMPARISON TO THE PREVIOUS METHODS AND RESULTS

In this section, we primarily focus on comparing our results with the previously known best result in [13]. While our approach yields an improved cost over [13], i.e., 0.1670 compared to 0.1673, it is interesting to note that our solutions are relatively close to one another: each f has 3.5 main steps; the corresponding breakpoints and signal levels for the 3.5 main steps in each f are respectively close to each other; and for each main step, there are sub-steps, i.e., each $f(x)$ has a slightly sloped piece between discontinuities.
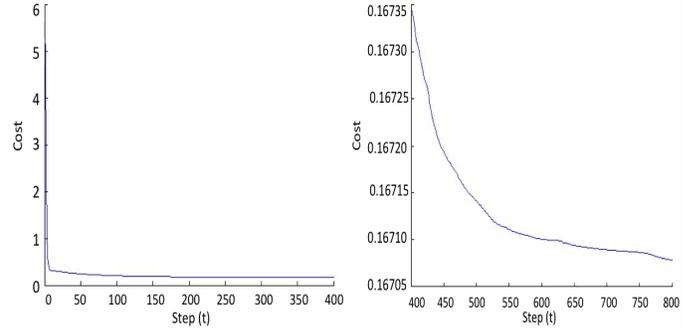
Fig. 4. Convergence process of the total cost J in the case where $n = 600$; $m = 675$: the left one shows the first 400 steps of iteration; the right one shows the later 400 steps of iteration. Note that the ticks for Y-axis are different in the two plots, with the aim of showing the convergence dynamics more clearly.

Furthermore, since the parameters used in our first case (i.e., $n = 400, m = 450$) were chosen to make our solution comparable to the one in [13], we observe that even the breakpoints and signal levels for the sub-steps in each f are respectively close to each other. In Figure 6, we plot the difference between $f(x)$ in Table II and $f(x)$ in [13]. Although we do not know whether the shape of $f(x)$ is optimal, the similarity of the solutions found by two quite different searching methods might not be a coincidence.

Besides improving the best known results to date, the learning algorithms presented in the paper have alternative advantages with regards to complexity when compared to many of the alternative search algorithms proposed in [4], [12], [13]. First, those searching methods for the WC relied on the monotone nondecreasing property of the optimal $f(x)$. This property was proven by Witsenhausen [22] with a certain amount of effort. However, our solutions do not rely on this property. The solutions we achieved verify that $f(x)$ is monotone nondecreasing.

Secondly, some of these algorithms are predicated on information regarding the problem setup; hence, they are not fully automated. Take the *hierarchical search approach* proposed in [13] as an example. During the search process, each main procedure required significant analytic work to
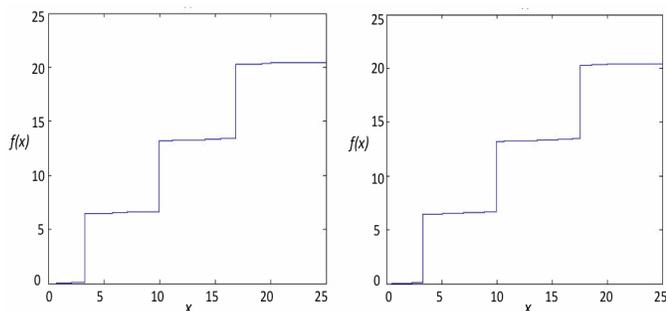
Fig. 5. Solutions of $f(x)$: the left one is the $f(x)$ in table II; the right one is the $f(x)$ in table III.
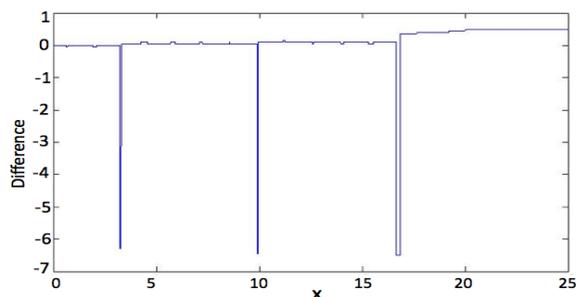
Fig. 6. Difference between $f(x)$ in Table II and that in [13]: $f(x)$ in Table II minus $f(x)$ in [13].

obtain a preference for doing that procedure, such as determining the proper interval values and signaling levels, determining the step number, determining whether and how to insert smaller steps and so on. Hence, their approach requires paying close attention to each procedure. In contrast, our method is automated: once we discretize the problem and setup the learning approach, we wait for the solution to emerge.

Lastly, to some extent learning approaches are easier to extend for solving other optimization problems. Some of other searching methods requires many modifications when being employed to other different problems. Again take the hierarchical search approach as an example. We should design proper hierarchical searching stages according to the properties of different problems; and because each stage requires authors' priori information about the problem, this work is not easy. While, as for learning approaches, we just need to find a reasonable way to model the optimization problem as a game, i.e. set up the corresponding players, strategy space and utility functions. Then we can implement one proper learning approach which has been already designed in game theory to search for a good solution.

V. CONCLUSION

In this paper, we used step functions to approximate the optimal controllers, which is not a new idea; however, we used a different searching method namely learning approaches in game theory to search for such step functions. We demonstrate that it does not only provide us a better result than the previous ones in the past 40 years, but also has

other advantages both for solving the Witsenhausen Counterexample and for solving general optimization problems after comparing with other existing searching methods.

REFERENCES

[1] R. Bansal and T. Basar, "Stochastic Teams with Nonclassical Information Revisited: When is an Affine Law Optimal?" *IEEE Transactions on Automatic Control*, Vol. AC-32, No. 6, pp.554-559, June 1987.
[2] M. Baglietto, T. Parisini and R. Zoppoli, "Numerical Solutions to the Witsenhausen Counterexample by Approximating Networks," *IEEE Transactions on Automatica Control*, Vol. 46, No. 9, pp. 1471-1477, September 2001.
[3] M. L. Cohen, "The Fisher Information and Convexity", *IEEE Transactions on Automatica Control,* Vol. IT-14, pp. 591-592, 1968
[4] M. Deng and Y. C. Ho, "An Ordinal Optimization Approach to Optimal Control Problems(Brief Paper)," *Automatica* 35, pp. 331-338, 1999.
[5] D. Fudenberg and D. K. Levine, "Game Theory", *MIT Press, Cambridge*, MA, 1991
[6] D. Fudenberg and D. Kreps, " Learning mixed equilibria", *Games and Economic Behavior*, Vol. 5, No. 3, pp. 320-367, 1993
[7] D. Fudenberg and D. K. Levine, "The Theory of Learning in Games", *MIT Press, Cambridge*, MA, 1998
[8] Y. C. Ho and K. C. Chu, "Team Decision Theory and Information Structures in Optimal Control Problems–Part I," *IEEE Transactions on Automatic Control*, Vol. AC-17, No. 1, pp. 15-22, February 1972.
[9] Y. C. Ho and T. S. Chang, "Another Look at the Nonclassical Information Structure Problem," *IEEE Transactions on Automatic Control*, Vol. AC-25, No. 3, pp. 537-540, June 1980.
[10] Y. C. Ho and K. C. Chu, "Inforamtion Structure in Dynamic Multi-Person Control Problems," *Automatic*, Vol. 10, pp. 341-351, Pergamon Press, 1974.
[11] Y. C. Ho, "Team Decision Theory and Inforamtion Structures," *Proceedings of the IEEE*, Vol. 68, No. 6, pp. 644-654, June 1980.
[12] Y. C. Ho and J. T. Lee, "Granular Optimization: An Approach to Function Optimization," *Proceeding of the 39th IEEE Conference on Decision and Control*, Sydney, Australia, pp. 103-110, 2000.
[13] J. T. Lee, E. Lau, and Y. C. Ho, "The Witsenhausen Counterexample: A Hierachical Search Approach for Nonconvex Optimization Problems," *IEEE Transactions on Automatic Control*, Vol. 46, No. 3, pp. 382-397, March 2001.
[14] J. R. Marden, G. Arslan, and J. S. Shamma, "Joint Strategy Fictious Play with Inertia for Potential Games," *IEEE Transactions on Automatic Control*, Vol. 54, Issue 2, pp. 208-220, February 2009.
[15] J. Marschak, "Elements for a Theory of Teams," *Management Science*, Vol. 1, No. 2, pp.127-137, Janary 1995.
[16] D. Monderer, "Fictious Play Property for Games with Identical Interests," *Jounal of Economic Theory*, 68, Article No. 0014, pp. 285-265, 1996.
[17] D. Monderer, "Potential Games," *Games and Economic Behavior* 14, Article No. 0044, pp. 124-143, 1996.
[18] C. H. Papadimitriou and J.Tsitsiklis, "Intractable Problem in Control Theory," *SIAM J. Control And Optimization*, Vol. 24, No. 4, pp.639-654, July 1986.
[19] R. Radner, "Team Decision Problems," *The Annals of Mathematical Statistics*, Vol. 33, No. 3, September 1962.
[20] M. Rotkowitz, "Linear Controllers are Uniformly Optimal for the Witsenhausen Counterexample," *IEEE Conference on Decison and Control*, 2006.
[21] M. Rotkowitz, "A Characterization of Convex Problems in Decentralized Control," *IEEE Transactions on Automatic Control*, Vol. 51, No. 2, February 2006.
[22] H. S. Witsenhausen, "A Counterexample in Stochastic Optimum Control," *SIAM J. Control*, Vol. 6, No. 1, pp. 131-147, 1968.
[23] H. P. Young, "Strategic Learning and Its Limits", *Oxford University Press,* Oxford, UK, 2004