SIAM J. CONTROL OPTIM.
Vol. 45, No. 4, pp. 1359–1368

AN APPROXIMATION ALGORITHM FOR THE DISCRETE TEAM DECISION PROBLEM*

RANDY COGILL† AND SANJAY LALL‡

Abstract. In this paper we study a discrete version of the classical team decision problem. It has been shown previously that the general discrete team decision problem is NP-hard. Here we present an efficient approximation algorithm for this problem. For the maximization version of this problem with nonnegative rewards, this algorithm computes decision rules which are guaranteed to be within a fixed bound of optimal.

Key words. team theory, decentralized control, approximation algorithms

AMS subject classifications. 93A14, 68Q25, 68W25

DOI. 10.1137/050628374

1. Introduction.

The problem of decentralized control arises when several decision makers, each with limited information, must make simultaneous cooperative decisions. The fact that multiple decisions are made on the basis of incomplete and differing observations of the overall system state is what sets decentralized control problems apart from conventional control problems. Such problems arise naturally in applications involving sensing and communication networks. An example is detection by multiple sensors in a sensor network [7]. The job of the sensors is to take measurements of the environment and transmit a minimal amount of information to a fusion center which estimates the state of the environment based on the information received from the sensors. Each sensor is faced with a decentralized decision making problem: based on its limited measurement of the environment, it must decide what information it should send to maximize the probability that the fusion center makes a correct estimate.

Marschak was probably the first to study the problem of decentralized decision making. In his work [5], he introduced *team theory* as a framework for studying decision making problems in organizations. The class of problems considered by team theory, called team decision problems, are analogous in some ways to static games. In these problems, the state of the system is chosen randomly according to some specified probability, and each decision maker partially observes the state. Based on their observations, each decision maker chooses an action. The goal is to choose decision rules which maximize the expected value of a reward which is jointly a function of the system state and all actions.

The later work of Radner [6] presented optimality conditions for a class of team decision problems with reward functions which are concave and differentiable in the decision variables. The main result of that work is that, subject to some technical conditions, *person-by-person optimal* policies are globally optimal for such problems.

*Received by the editors April 1, 2005; accepted for publication (in revised form) April 3, 2006; published electronically October 6, 2006.
http://www.siam.org/journals/sicon/45-4/62837.html
†Department of Electrical Engineering, Stanford University, Stanford, CA 94305 (rcogill@stanford.edu). This author was partially supported by a Stanford Graduate Fellowship.
‡Department of Aeronautics and Astronautics, Stanford University, Stanford, CA 94305 (lall@stanford.edu). This author was partially supported by the Stanford URI Architectures for Secure and Robust Distributed Infrastructures, AFOSR DoD award 49620-01-1-0365.

A set of decision rules is person-by-person optimal if no improvement can be obtained by changing the decision rule for one decision maker while leaving the decision rules for others fixed. A person-by-person optimal policy can be computed by sequentially optimizing the decision rule for each decision maker while leaving the remaining decision rules fixed.

The problems we are interested in have state and action sets which are finite, so the results regarding global optimality of person-by-person optimal policies do not apply. In fact, the problems that we consider may have many person-by-person optimal policies, and there may exist a person-by-person optimal policy which is quite poor compared to the globally optimal policy. Although problems with finite state and action spaces were presented in the original work by Marschak [5], no positive results regarding computation of optimal policies for these problems followed. Tsitsiklis showed in [8] that the general team decision problem with finite state and action spaces is NP-hard, explaining the lack of positive results for this problem. Another recent paper which considers the discrete team decision problem is [9]. In that paper, the authors consider problems with finite state and action spaces and reward functions which are *multimodular* in the decision variables. Multimodularity of functions on discrete spaces is analogous to concavity of functions on continuous spaces. The authors attempt to use multimodularity to extend the results of Radner to discrete problems. They show that necessary and sufficient optimality conditions are not assured by multimodularity, but this property still can be exploited when searching for optimal policies.

Since it is highly unlikely that an algorithm exists which can efficiently compute optimal solutions to the general team decision problem, this leads us to ask if an efficient algorithm exists which can compute acceptable suboptimal solutions. Many problems, although NP-hard, admit efficient *approximation algorithms* which produce solutions guaranteed to be within a fixed bound of optimal [1]. In this paper we present an approximation algorithm for the team decision problem. Specifically, the contributions of this paper are the following:

- For the general team problem with two decision makers, we present an algorithm that runs in $\mathcal{O}(|Y_1||Y_2||U_1||U_2|)$ time. When the objective is to maximize a nonnegative reward function, this algorithm computes a decentralized policy guaranteed to be within a factor of $1/\min\{|U_1|, |U_2|\}$ of optimal. Here, $|Y_1|, |Y_2|$ are the number of states and $|U_1|, |U_2|$ are the number of actions associated with decision makers 1 and 2. Even the special case with $\min\{|U_1|, |U_2|\} = 2$ is still NP-hard, and in this case we obtain a $1/2$ approximation factor.
- We consider the class of team decision problems with two decision makers and multimodular reward functions, as in [9]. We show that this special case is still NP-hard, although the presence of multimodularity leads to a significantly tighter approximation ratio of $1/2 - 1/(2\max\{|U_1|, |U_2|\})$ for our algorithm.

An outline of this paper is as follows. In section 2, we introduce the general discrete team decision problem and provide a proof that this problem is NP-hard. In section 3, we present an approximation algorithm for the general team decision problem and prove an approximation ratio for this algorithm. In section 4, we consider a special class of team decision problems with multimodular cost functions. In this section, we show that this class of problems is still NP-hard, but the approximation ratio for our algorithm is significantly tighter on these problems.

2. Problem formulation and complexity. In this section we will present the Team Decision Problem (TDP), the problem which is the subject of this paper. Before presenting the formal definition, we will briefly describe this problem in words. The problem is essentially a feedback control problem, where the goal is to choose actions in response to observations with the goal of maximizing some reward. The system is in a state described by the variables (y_1, y_2). After observing the system state, we would like to choose a pair of actions (u_1, u_2) to make the reward $c(y_1, y_2, u_1, u_2)$ as large as possible. However, the actions must be chosen in a decentralized manner. That is, action u_1 is chosen based only on an observation of y_1 and action u_2 is chosen based only on an observation of y_2. Our goal is to choose decision rules $\gamma_1 : Y_1 \to U_1$ and $\gamma_2 : Y_2 \to U_2$ to maximize some measure of total reward. A formal definition of the problem is the following.

Team Decision Problem. *Given finite sets Y_1, Y_2, U_1, U_2 and a nonnegative reward function $c : Y_1 \times Y_2 \times U_1 \times U_2 \to \mathbb{Q}_+$, find decision rules $\gamma_i : Y_i \to U_i$, $i = 1, 2$ which maximize the reward*

$$J(\gamma_1, \gamma_2) = \sum_{y_1, y_2} c(y_1, y_2, \gamma_1(y_1), \gamma_2(y_2)).$$

A few words are in order regarding the formulation given above. First of all, this formulation only considers the case in which two decision makers are choosing actions. In general, we have problems where n decision makers are independently choosing actions $u_1, \ldots, u_n$. It turns out, as we will see shortly, that this problem is computationally intractable even in the two decision maker case. This is the simplest special case which exhibits the inherent computational complexity of this problem, and this case is the focus of this paper. It is worth noting that the approach taken in this paper can be extended to problems with more than two decision makers, although the resulting approximation guarantees degrade exponentially as the number of decision makers increases.

In [8], this problem is posed as maximization of the *expected* reward with respect to some probability mass function $p(y_1, y_2)$. While this formulation may relate more naturally to the applications where this problem is of interest, the formulation in terms of expected reward and the formulation in terms of total reward are essentially equivalent. That is, given any instance of the problem of maximizing expected reward, we can easily modify the reward function to obtain an equivalent instance of maximizing total reward. Conversely, any instance of the problem of maximizing total reward is equivalent to an instance of maximizing expected reward where $p(y_1, y_2)$ is uniform. Therefore, we consider the problem of maximizing total reward simply to reduce the amount of required notation.

Also, it is interesting to note that the corresponding centralized problem is trivial. By the centralized problem, we mean the problem of choosing a policy $\gamma : Y_1 \times Y_2 \to U_1 \times U_2$, to maximize the total reward. This problem is solved by simply choosing the (u_1, u_2) which maximizes $c(y_1, y_2, u_1, u_2)$ for each (y_1, y_2). Although the centralized problem is easy, it was shown in [8] that TDP is NP-hard, even when $|U_1| = |U_2| = 2$.

Here we will present a simple new proof of NP-hardness to keep our treatment self contained. We will do this by also showing that the special case of TDP with $|U_1| = |U_2| = 2$, which we refer to as TDP-2, is NP-hard. Our proof of NP-hardness of TDP-2 involves reducing the problem Maximum Cut [1] to TDP-2. The problem Maximum Cut is the following. As input, we are given an undirected graph $G = (V, E)$. The goal is to partition the set of vertices V into two sets S and $\overline{S}$ so that the number of edges crossing from vertices in S to vertices in $\overline{S}$ is maximized. Given that MAXIMUM CUT is NP-hard, we have the following theorem.

THEOREM 2.1. *The problem* TDP-2 *is NP-hard.*

Proof. Consider an arbitrary instance of MAXIMUM CUT specified by an undirected graph $G = (V, E)$. We construct a corresponding instance of TDP-2 as follows. Let $Y_1 = Y_2 = V$. Define c_1 as

$$c_1(y_1, y_2, u_1, u_2) = \begin{cases} |V|^2 & \text{if } y_1 = y_2 \text{ and } u_1 = u_2, \\ 0 & \text{otherwise.} \end{cases}$$

Define c_2 as

$$c_2(y_1, y_2, u_1, u_2) = \begin{cases} 1 & \text{if } (y_1, y_2) \in E \text{ and } u_1 \neq u_2, \\ 0 & \text{otherwise.} \end{cases}$$

The goal is to show that the instance of MAXIMUM CUT is solved by finding an optimal policy for the instance of TDP-2 with reward

$$c(y_1, y_2, u_1, u_2) = c_1(y_1, y_2, u_1, u_2) + c_2(y_1, y_2, u_1, u_2).$$

The optimal reward satisfies $|V|^3 < J(\gamma_1^*, \gamma_2^*)$. The lower bound is achieved by choosing any policy satisfying $\gamma_1 = \gamma_2$. Such policies maximize the component of the reward associated with c_1. As long as G contains one edge, we can always achieve reward strictly greater than $|V|^3$. Moreover, if $\gamma_1(i) \neq \gamma_2(i)$ for some i, then $J(\gamma_1, \gamma_2) \leq |V|^3$. This is because, in this case, the component of the reward associated with c_1 is at most $(|V| - 1)|V|^2$ and the component of the reward associated with c_2 is at most $|V|^2$. Hence $\gamma_1^* = \gamma_2^*$.

Given any decision rules γ_1 and γ_2 satisfying $\gamma_1 = \gamma_2$, we obtain a cut for the instance of MAXIMUM CUT by letting

$$v_i \in \begin{cases} S & \text{if } \gamma_1(v_i) = 1, \\ \overline{S} & \text{if } \gamma_1(v_i) = 0 \end{cases}$$

for all $v_i \in V$. The capacity of this cut is exactly $\frac{1}{2}J(\gamma_1, \gamma_2) - \frac{1}{2}|V|^3$. Hence, any algorithm which computes optimal policies for arbitrary instances of TDP-2 in polynomial time can be easily modified to compute optimal cuts for arbitrary instances of MAXIMUM CUT in polynomial time. ∎

3. An approximation algorithm. Given that TDP is NP-hard, it is highly unlikely that an efficient algorithm for computing optimal policies exists. However, for many NP-hard problems there exist efficient *approximation algorithms* which produce suboptimal solutions which achieve a reward guaranteed to be within some constant factor of optimal. In this section we will give an approximation algorithm for TDP which produces a suboptimal policy in $\mathcal{O}(|Y_1||Y_2||U_1||U_2|)$ operations.

Before presenting our approximation algorithm, we would like to first consider algorithms that produce *person-by-person* optimal policies [6]. A person-by-person optimal policy is a pair of decision rules γ_1, γ_2 for which no improvement can be obtained by modifying γ_1 while leaving γ_2 fixed or by modifying γ_2 while leaving γ_1 fixed. A person-by-person optimal solution can be considered analogous in some ways to a Nash equilibrium. Person-by-person optimal solutions can be computed by choosing initial decision rule γ_1 and γ_2 and alternately maximizing over one decision

TABLE 3.1
For this reward function there is a person-by-person optimal policy with reward $J(\gamma_1, \gamma_2) = 8$. The optimal policy has reward $J(\gamma_1^*, \gamma_2^*) = 400$.

	$y_1 = 0$ $u_1 = 0$	$y_1 = 0$ $u_1 = 1$	$y_1 = 1$ $u_1 = 0$	$y_1 = 1$ $u_1 = 1$
$y_2 = 0$ $u_2 = 0$	2	1	1	2
$y_2 = 0$ $u_2 = 1$	1	100	100	1
$y_2 = 1$ $u_2 = 0$	1	100	100	1
$y_2 = 1$ $u_2 = 1$	2	1	1	2

rule while leaving the other decision rule fixed until no more improvement is obtained. For discrete problems, this process will terminate in a finite number of iterations.

For problems with rewards which are continuous, differentiable, and concave in the variables u_1 and u_2, it was shown that person-by-person optimality implies global optimality [6]. This is probably the best-known result for team decision problems. It is natural, then, to believe that person-by-person optimal policies may be acceptable solutions to discrete team decision problems. The purpose for considering person-by-person optimal solutions here is to show that, in fact, person-by-person optimal solutions may achieve a reward which is arbitrarily far from the globally optimal reward in discrete problems.

Consider the problem instance with the reward function given in Table 3.1. For this instance, consider the policy $\gamma_1(0) = 0$, $\gamma_1(1) = 1$, $\gamma_2(0) = 0$, $\gamma_2(1) = 1$. This policy is person-by-person optimal and achieves a reward of $J(\gamma_1, \gamma_2) = 8$. However, the optimal policy achieves a reward of $J(\gamma_1^*, \gamma_2^*) = 400$. Here, a person-by-person optimal policy is suboptimal by a factor of 50. Of course, we can modify the example to make the reward achieved by the person-by-person optimal as far from the optimal reward as we like. Therefore, algorithms which start with an arbitrary policy and seek person-by-person optimal solutions cannot produce polices which are within a guaranteed bound of optimal.

We will now present our approximation algorithm for TDP. Roughly speaking, this algorithm first constructs a "marginalized" reward function

$$c_1(y_1, u_1) = \sum_{u_2, y_2} c(y_1, y_2, u_1, u_2),$$

then computes a decision rule $\gamma_1 : Y_1 \to U_1$ which is optimal for this marginalized reward. Then, given γ_1, we compute the decision rule $\gamma_2 : Y_2 \to U_2$ which is optimal with respect to γ_1. Unlike an arbitrary person-by-person optimal policy, this policy is guaranteed to be within a constant factor of optimal, where the factor depends on $|U_1|$ and $|U_2|$. To obtain the tightest suboptimality guarantee, we may first compute γ_2 and then γ_1.

ALGORITHM 3.1. *Assume, without loss of generality, that $|U_1| \geq |U_2|$.*

1. *Let $\gamma_1(y_1) \in argmax_{u_1}\{\sum_{u_2} \sum_{y_2} c(y_1, y_2, u_1, u_2)\}$ for all $y_1 \in Y_1$.*
2. *Let $\gamma_2(y_2) \in argmax_{u_2}\{\sum_{y_1} c(y_1, y_2, \gamma_1(y_1), u_2)\}$ for all $y_2 \in Y_2$.*

If $|U_2| > |U_1|$, then we just compute γ_2 first, followed by γ_1. The approximation guarantee for the previous algorithm is given in the following theorem.

THEOREM 3.2. *For* TDP*, Algorithm* 3.1 *produces a policy* γ_1, γ_2 *with value satisfying*

$$J(\gamma_1, \gamma_2) \geq \frac{1}{\min\{|U_1|, |U_2|\}} J(\gamma_1^*, \gamma_2^*).$$

Proof. Without any loss of generality, here we will assume that $|U_1| \geq |U_2|$. Since $\gamma_1(y_1) \in \operatorname{argmax}_{u_1} \sum_{u_2} \sum_{y_2} c(y_1, y_2, u_1, u_2)$, we have

$$\begin{aligned}\sum_{y_2} \sum_{u_2} c(y_1, y_2, \gamma_1(y_1), u_2) &\geq \sum_{y_2} \sum_{u_2} c(y_1, y_2, \gamma_1^*(y_1), u_2) \\ &\geq \sum_{y_2} c(y_1, y_2, \gamma_1^*(y_1), \gamma_2^*(y_2))\end{aligned}$$

for each $y_1 \in Y_1$. Also, $\gamma_2(y_2) \in \operatorname{argmax}_{u_2} \sum_{y_1} c(y_1, y_2, \gamma_1(y_1), u_2)$, so

$$\sum_{y_1} c(y_1, y_2, \gamma_1(y_1), \gamma_2(y_2)) \geq \frac{1}{|U_2|} \sum_{u_2} \left(\sum_{y_1} c(y_1, y_2, \gamma_1(y_1), u_2) \right)$$

for each $y_2 \in Y_2$. Therefore, the reward achieved by policy γ_1, γ_2 satisfies

$$\begin{aligned}\sum_{y_1} \sum_{y_2} c(y_1, y_2, \gamma_1(y_1), \gamma_2(y_2)) &\geq \sum_{y_2} \left(\frac{1}{|U_2|} \sum_{u_2} \sum_{y_1} c(y_1, y_2, \gamma_1(y_1), u_2) \right) \\ &= \frac{1}{|U_2|} \sum_{y_1} \left(\sum_{y_2} \sum_{u_2} c(y_1, y_2, \gamma_1(y_1), u_2) \right) \\ &\geq \frac{1}{|U_2|} \sum_{y_1} \sum_{y_2} c(y_1, y_2, \gamma_1^*(y_1), \gamma_2^*(y_2)). \qquad \square\end{aligned}$$

Although the approximation guarantee degrades as the number of available decisions increases, we must keep in mind that even TDP-2 is NP-hard, and in this case we obtain an approximation factor of 1/2.

It is worth noting that Algorithm 3.1, and consequently the proof of Theorem 3.2, has an alternate interpretation in terms of *randomized* decision rules. In this interpretation, γ_2 is initially set to be the randomized decision rule which selects actions randomly according to a uniform distribution, regardless of the value of y_2. We then choose γ_1 to be the decision rule which is optimal with respect to this randomized γ_2. Finally, the randomized γ_2 is replaced with the decision rule which is optimal with respect to the chosen γ_1. While this interpretation of Algorithm 3.1 is slightly more complex conceptually, this interpretation can be used to simplify the proof of Theorem 3.2.

We can show that the bound is tight for this algorithm, and it is achieved on a simple example with $|U_1| = |U_2| = |Y_1| = |Y_2| = 2$. Consider the reward function given in Table 3.2. Here, ϵ is some arbitrarily small constant which is simply introduced to avoid any ambiguity associated with adding a tie-breaking mechanism into the algorithm. The policy produced by the algorithm, $\gamma_1(0) = 0$, $\gamma_1(1) = 0$, $\gamma_2(0) = 1$, $\gamma_2(1) = 1$, achieves $J(\gamma_1, \gamma_2) = 2 + 2\epsilon$. This is true regardless of which variables we

TABLE 3.2
The bound in Theorem 3.2 is tight for this reward function.

	$y_1 = 0$ $u_1 = 0$	$y_1 = 0$ $u_1 = 1$	$y_1 = 1$ $u_1 = 0$	$y_1 = 1$ $u_1 = 1$
$y_2 = 0$ $u_2 = 0$	1	0	0	1
$y_2 = 0$ $u_2 = 1$	0	1	$1+\epsilon$	0
$y_2 = 1$ $u_2 = 0$	0	1	1	0
$y_2 = 1$ $u_2 = 1$	$1+\epsilon$	0	0	1

choose to marginalize over first. However, the optimal policy $\gamma_1(0) = 0$, $\gamma_1(1) = 1$, $\gamma_2(0) = 0$, $\gamma_2(1) = 1$ achieves $J(\gamma_1^*, \gamma_2^*) = 4 + \epsilon$. Since ϵ can be chosen arbitrarily small, the ratio between the reward achieved by the computed policy and optimal policy can be made arbitrarily close to 1/2. Similar examples can be constructed for larger values of $|U_1|$ and $|U_2|$.

It is worth noting that the policy constructed by Algorithm 3.1 in the previous example is person-by-person optimal. This means that the bound in Theorem 3.2 is still tight even for person-by-person optimal policies obtained by iterating over decision rules using the decision rules produced by Algorithm 3.1 as a starting point.

4. Approximation ratio for multimodular rewards. In this section we consider a class of team decision problems with specially structured rewards and show that our algorithm achieves a tighter approximation ratio on these problems. The problems that we consider here can be thought of as the discrete counterpart of the problems considered by Radner [6]. In particular, Radner considered problems with rewards which are continuous, differentiable, and concave in the decision variables. It was shown that globally optimal policies can be easily computed for these problems. Here we consider rewards which are *multimodular* in discrete decision variables. Multimodularity is an extension of the notion of convexity to functions of discrete variables. Discrete team decision problems with multimodular rewards were studied previously in [9]. The authors attempt to use multimodularity to extend the results of Radner to discrete problems. They show that necessary and sufficient optimality conditions are not assured by multimodularity, but this property still can be exploited when searching for optimal policies. Here we show the class of problems with multimodular rewards is still NP-hard. However, the presence of multimodularity improves the approximation ratio to $1/2 - 1/(2\max\{|U_1|, |U_2|\})$ for the algorithm presented in the previous section. Combining this with the bound proven in the last section, this gives a bound of 1/3 in the worst case which only *improves* as the sizes of U_1 and U_2 increase.

The concept of multimodularity was introduced by Hajek in [2]. In that paper, multimodularity is defined in terms of local properties of a function and the definition we give below is proven later to be a property equivalent to multimodularity. However, here we do not use most of the theory developed for multimodular functions. We use the definition below because it most clearly relates to convexity.

DEFINITION 4.1. *A function $f : \mathbb{Z}^n \to \mathbb{R}$ is said to be* multimodular *iff there exists a concave function $g : \mathbb{R}^n \to \mathbb{R}$ such that $f(x) = g(x)$ for all $x \in \mathbb{Z}^n$.*

It is worth noting that in the existing literature, multimodular functions are typically defined as discrete functions which coincide with *convex* functions. Since we do not know of a corresponding term for discrete functions which coincide with concave functions, we simply refer to these as multimodular as well. Also, throughout this section we consider functions which are defined on some subset $S \subset \mathbb{Z}^n$. When we say that such a function is multimodular, we mean that there is some concave function which coincides with f everywhere on S.

We can now describe the class of reward functions that we consider in this section. Specifically, we consider rewards $c : Y_1 \times Y_2 \times U_1 \times U_2$ such that for each $y_1 \in Y_1$ and $y_2 \in Y_2$, $c(y_1, y_2, u_1, u_2)$ is multimodular in u_1 and u_2. To make this statement precise, we have $U_1 = \{1, \dots, |U_1|\}$ and $U_2 = \{1, \dots, |U_2|\}$. We refer to this class of problems as TDP-M. We can make the following claim regarding the complexity of TDP-M.

THEOREM 4.2. *The problem* TDP-M *is NP-hard.*

Proof. The show this, we just need to show that every instance of TDP-2 is an instance of TDP-M. This is done by showing any function $f : \{0,1\} \times \{0,1\} \to \mathbb{R}$ is multimodular.

For some function $f : \{0,1\} \times \{0,1\} \to \mathbb{R}$, suppose that $f(0,0) + f(1,1) \leq f(0,1) + f(1,0)$. We can define the following piecewise affine continuation on $[0,1] \times [0,1]$:

$$g(x_1, x_2) = \begin{cases} (1-x_1-x_2)f(0,0) + x_1 f(1,0) + x_2 f(0,1) & \text{for } x_1 + x_2 \leq 1, \\ (1-x_2)f(1,0) + (1-x_1)f(0,1) + (x_1+x_2-1)f(1,1) & \text{for } x_1 + x_2 > 1. \end{cases}$$

Verifying that this function is concave on $[0,1] \times [0,1]$ is somewhat tedious, but straightforward. If $f(0,0) + f(1,1) > f(0,1) + f(1,0)$, then we define g as

$$g(x_1, x_2) = \begin{cases} (1 - x_2)f(0,0) + (x_2 - x_1)f(0,1) + x_1 f(1,1) & \text{for } x_1 \leq x_2, \\ (1 - x_1)f(0,0) + (x_1 - x_2)f(1,0) + x_2 f(1,1) & \text{for } x_1 > x_2, \end{cases}$$

which is concave on $[0,1] \times [0,1]$. ∎

Now that we have established that TDP-M is NP-hard, our goal is to show that our algorithm has a tighter approximation ratio for these problems. To ensure the tightest approximation ratio for this class of problems, we make one very minor change to the algorithm.

ALGORITHM 4.3. *Assume, without loss of generality, that* $|U_2| \geq |U_1|$.

1. *Let* $\gamma_1(y_1) \in \text{argmax}_{u_1}\{\sum_{u_2} \sum_{y_2} c(y_1, y_2, u_1, u_2)\}$ *for all* $y_1 \in Y_1$.
2. *Let* $\gamma_2(y_2) \in \text{argmax}_{u_2}\{\sum_{y_1} c(y_1, y_2, \gamma_1(y_1), u_2)\}$ *for all* $y_2 \in Y_2$.

The only difference between this algorithm and Algorithm 3.1 is the order in which the decision rules are computed. As we saw previously, the approximation ratio depends on the number of decisions available to the decision rule which is computed second. For the approximation ratio proven in Theorem 3.2, the quality of the approximation degrades with the number of decisions. For the approximation ratio proven in this section, the quality of the approximation *improves* with the number of decisions.

Before proving approximation ratio for TDP-M, we first need to establish several properties of multimodular functions of a single variable.

LEMMA 4.4. *A function* $f : \mathbb{Z} \to \mathbb{R}$ *is multimodular iff*

$$f(x_2) \geq \frac{x_3 - x_2}{x_3 - x_1} f(x_1) + \frac{x_2 - x_1}{x_3 - x_1} f(x_3)$$

for any $x_1 \leq x_2 \leq x_3$.

Proof. If the inequality is violated for some x_1, x_2, and x_3, then clearly it is violated for any $g : \mathbb{R} \to \mathbb{R}$ which coincides with f. Conversely, if the inequality holds for all $x_1 \le x_2 \le x_3$, then it's easy to verify that the piecewise affine continuation of f,

$$g(x) = (\lceil x \rceil - x) f(\lfloor x \rfloor) + (x - \lfloor x \rfloor) f(\lceil x \rceil),$$

is concave. Here, $\lfloor x \rfloor$ denotes the largest integer less than or equal to x and $\lceil x \rceil$ denotes the smallest integer strictly greater than x. ∎

When proving Theorem 3.2 in the previous section, we used the fact that the sum of a set of nonnegative numbers is an upper bound on the maximum of these numbers. The key to obtaining a tighter approximation ratio for multimodular rewards is to use the fact that the gap between these two quantities can be quite large when the set of numbers corresponds to the values taken by a multimodular function. This is stated precisely in the following theorem.

THEOREM 4.5. *Suppose* $f : \{1, \dots, N\} \to \mathbb{R}_+$ *is multimodular. Then*

$$\sum_{i=1}^{N} f(i) \ge \frac{N-1}{2} f(j)$$

for all $j \in \{1, \dots, N\}$.

Proof. Let k be such that $f(k) \ge f(j)$ for all $j \in \{1, \dots, N\}$. We just need to show that $\sum_{i=1}^{N} f(i) \ge \frac{N-1}{2} f(k)$. Since $f(1) \ge 0$ and $f(N) \ge 0$, Lemma 4.4 gives

$$f(i) \ge \frac{i-1}{k-1} f(k) \quad \text{for all } 1 \le i \le k,$$
$$f(i) \ge \frac{N-i}{N-k} f(k) \quad \text{for all } k \le i \le N.$$

If $k = 1$ or $k = N$, we just have one of these inequalities for all $1 \le i \le N$. Using these inequalities,

$$\begin{aligned}\sum_{i=1}^{N} f(i) &\ge \sum_{i=1}^{k} \frac{i-1}{k-1} f(k) + \sum_{k+1}^{N} \frac{N-i}{N-k} f(k)\\ &= \frac{1}{k-1}\left(\frac{1}{2}k(k-1)\right) f(k) + \frac{1}{N-k}\left(\frac{1}{2}(N-k)(N-k-1)\right) f(k)\\ &= \frac{N-1}{2} f(k). \qquad \square\end{aligned}$$

We can now prove the main result of this section.

THEOREM 4.6. *For* TDP-M*, Algorithm* 4.3 *produces a policy* γ_1*,* γ_2 *with value satisfying*

$$J(\gamma_1, \gamma_2) \ge \left(\frac{1}{2} - \frac{1}{2\max\{|U_1|, |U_2|\}}\right) J(\gamma_1^*, \gamma_2^*).$$

Proof. Without any loss of generality, here we will assume that $|U_2| \ge |U_1|$. In the proof of Theorem 3.2, we use the fact that

$$\sum_{u_2} c(y_1, y_2, \gamma_1^*(y_1), u_2) \ge c(y_1, y_2, \gamma_1^*(y_1), \gamma_2^*(y_2)). \tag{4.1}$$

For any fixed y_1 and y_2, $c(y_1, y_2, \gamma_1^*(y_1), u_2)$ is multimodular in u_2. Therefore, we can use Theorem 4.5 to strengthen inequality (4.1) to

$$\sum_{u_2} c(y_1, y_2, \gamma_1^*(y_1), u_2) \geq \frac{|U_2| - 1}{2} c(y_1, y_2, \gamma_1^*(y_1), \gamma_2^*(y_2)).$$

This gives us the bound

$$\begin{aligned} \sum_{y_1} \sum_{y_2} c(y_1, y_2, \gamma_1(y_1), \gamma_2(y_2)) &\geq \sum_{y_2} \left(\frac{1}{|U_2|} \sum_{u_2} \sum_{y_1} c(y_1, y_2, \gamma_1(y_1), u_2) \right) \\ &\geq \frac{1}{|U_2|} \sum_{y_1} \sum_{y_2} \left(\sum_{u_2} c(y_1, y_2, \gamma_1^*(y_1), u_2) \right) \\ &\geq \left(\frac{1}{2} - \frac{1}{2|U_2|} \right) \sum_{y_1} \sum_{y_2} c(y_1, y_2, \gamma_1^*(y_1), \gamma_2^*(y_2)). \qquad \square \end{aligned}$$

Note that this bound is not tight for the case where $|U_1| = 2$ or $|U_2| = 2$. This is simply because the bound proven in Theorem 4.4 is trivial and never tight when $N = 2$. Of course in this the case, the bound proven in Theorem 3.2 still holds and is tight. With this in mind, the bound in Theorem 4.6 says that the approximation ratio for any instance of TDP-M is at least 1/3, and approaches 1/2 as the sizes of U_1 and U_2 increase.

5. Conclusions. In this paper we presented an approximation algorithm for the discrete team decision problem [8, 9]. We focused on problems involving two decision makers. For the general discrete team decision problem, the approximation ratio for our algorithm depends on the number of actions available to each decision maker. For the case when at least one decision maker chooses between two actions, which is still NP-hard, we have an approximation ratio of 1/2. We then considered a special case of the discrete team decision problem which can be thought of as the discrete counterpart of the problems considered by Radner [6]. We show that this special class of problems is still NP-hard, but the approximation ratio for our algorithm is significantly tighter on these problems.

REFERENCES

[1] M. R. Garey and D. S. Johnson, *Computers and Intractability—A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco, 1979.

[2] B. Hajek, *Extremal splittings of point processes*, Math. Oper. Res., 10 (1985), pp. 543–556.

[3] Y. C. Ho, *Team decision theory and information structures*, Proceedings of the IEEE, 68 (1980), pp. 644–654.

[4] J. Krainak, J. Speyer, and S. Marcus, *Static team problems—Part* I: *Sufficient conditions and the exponential cost criterion*, IEEE Trans. Automat. Control, AC-27 (1982), pp. 839–848.

[5] J. Marschak, *Elements for a theory of teams*, Management Science, 1 (1955), pp. 127–137.

[6] R. Radner, *Team decision problems*, Ann. Math. Statist., 33 (1962), pp. 857–881.

[7] J. Tsitsiklis, *Decentralized Detection*, vol. 2, H. V. Poor and J. B. Thomas, eds., JAI Press, Greenwich, CT, 1993, pp. 297–344.

[8] J. Tsitsiklis and M. Athans, *On the complexity of decentralized decision making and detection problems*, IEEE Trans. Automat. Control, 30 (1985), pp. 440–446.

[9] P. R. De Waal and J. H. Van Schuppen, *A class of team problems with discrete action spaces: Optimality conditions based on multimodularity*, SIAM J. Control Optim., 38 (2000), pp. 875–892.