SIAM J. CONTROL OPTIM.
Vol. 38, No. 3, pp. 875–892

© 2000 Society for Industrial and Applied Mathematics

A CLASS OF TEAM PROBLEMS WITH DISCRETE ACTION SPACES: OPTIMALITY CONDITIONS BASED ON MULTIMODULARITY*

PETER R. DE WAAL†‡ AND JAN H. VAN SCHUPPEN†

Abstract. In this paper we discuss a class of team problems with discrete action spaces. We introduce multimodularity into team theory as a natural alternative to convexity in continuous spaces. The main result relates coordinatewise-optimal (cw-optimal) points to the optimal team decision for a class of team problems. The method is based on a characterization of coordinatewise minima of multimodular functions.

Key words. team theory, multimodularity, person-by-person optimality

AMS subject classifications. 93A14, 93E20

PII. S0363012996301452

1. Introduction. In 1955 Marschak introduced in [6] *team problems* as a mathematical model for cooperative decision making. In a team problem there are two or more decision makers or controllers who receive a common reward as the joint result of their decisions. The fact that the decision makers have a common objective sets it apart from the models that are usually encountered in game theory. Team problems differ from ordinary decision problems with one controller, since the controllers may have different information on which they have to base their decision. The role of information in control problems is discussed in Witsenhausen [11, 12] and Ho and Chu [3]. For some examples and a tutorial introduction to team problems see Ho [4].

The applications of team problems were at first found in the area of decision making in organizations (see Marschak [6]). Recently the attention in team theory has acquired a new impulse from the area of load balancing in distributed computer systems (see, for instance, Wang and Morris [10]). The environment of high-performance computer networking provides a typical example of a complex and highly-distributed system for which decentralized control and team theory appear to provide the right framework.

Despite a history of more than forty years, there are not that many fundamental results in team theory. The verification of the optimality of a team strategy, for instance, is equivalent to a minimization over a function space, and this is infeasible without additional assumptions. To the best of our knowledge, there are only two papers in the literature that present conditions for optimality of team strategies. Radner presents in [7] a sufficient condition that guarantees optimality: if the cost is a convex function of the decision variables and the expected cost is locally finite and stationary for a given team strategy, then this strategy is optimal. Stationarity is defined as a first-order property of the conditional expectation of the cost given the different information patterns. In the case of a convex cost function stationarity of a strategy implies that it is person-by-person optimal (pbpo). This means that the

*Received by the editors April 1, 1996; accepted for publication (in revised form) June 8, 1999; published electronically March 8, 2000.

http://www.siam.org/journals/sicon/38-3/30145.html

†CWI, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands (waal@nlr.nl, schuppen@cwi.nl).

‡Current address: National Aerospace Laboratory NLR, P.O. Box 90502, 1006 BM Amsterdam, The Netherlands (waal@nlr.nl). The research of this author was supported by a grant of NWO (The Netherlands Organization for Scientific Research) through NFI.

expected cost cannot be improved by any player alone if the other team members keep using the same strategy. The importance of this result is twofold. First, it provides a way to verify the optimality of a strategy, and second, it suggests an algorithm to search for the optimal strategy. The local finiteness condition of Radner is relaxed by Krainak, Speyer, and Marcus [5] and replaced with a weaker condition. Both results, however, rely on the fact that the cost function is defined on a continuous space and that it is convex. The continuity makes it possible to compare the expected costs of any two team strategies by effectively constructing a randomization of the strategies. The expected cost of the randomized strategy is then a convex function of the randomization weight and the equivalence of local and global optimality for this one-dimensional convex function ensures the optimality of the stationary strategy.

The primary motivation for our research comes from decentralized control in distributed computer systems. These systems consist of a large number of computers that are interconnected by a network, and they allow sharing of resources and processors. Typically one computer is able to generate processes or tasks that can be performed on the computer itself, or they can be delegated to another processor. In the framework of team theory each computer (or more accurately each process scheduler) is a team player that has to decide for each process that is generated locally where the task has to be performed, locally or on another processor. The action space for such a team problem is intrinsically discrete, and it also does not allow a straightforward extension to a continuous action space. This property prohibits applying the results of [7] and [5].

The aim of this paper is to introduce a framework for team problems with a discrete action space and to present preliminary results for the existence and uniqueness of optima. For the cost function we consider a discrete space analogy for convexity, namely multimodularity. The specific results that were obtained for a class of two-person team problems are as follows:

- we present a characterization of the set of pbpo strategies;
- we give a procedure to check, for any pbpo strategy, in which direction to look for the optimal strategy. Not only does this provide us with an efficient search procedure, but it also enables us to check the optimality of a strategy.

The outline of this paper is as follows. In section 2 we introduce the team problem. A special class of team problems is described in section 3 and we present the optimality conditions for this class. For these conditions we rely on some results on the minima of multimodular functions. These results are summarized in the appendix.

2. Team problems.

In this section we introduce our general formulation of the team decision problem. We restrict our attention to a nondynamical team problem.

The following definition of a team problem is based on the definitions of Radner [7] and Krainak, Speyer, and Marcus [5]. We use an underlying probability space $(\Omega, \mathcal{F}, \mathcal{P})$, with Ω the space of elementary events, $\mathcal{F}$ a sigma field of subsets of Ω, and $\mathcal{P}$ a known probability measure on $\mathcal{F}$. We use the letter ω to denote an event, but we do not really distinguish between events and states. In fact we also refer to Ω as the underlying, unobserved, state space, and we also call ω the *state* when it represents an outcome.

A *team* is a set of N decision makers or players. Each player i can choose a decision a_i from a set A_i, the action set. In this paper we assume that the action sets are subsets of $\mathbf{Z}$. Here $\mathbf{Z}$ indicates the set of integers, and $\mathbf{N}$ indicates the set of natural numbers, including 0. If the players choose the action vector $a = (a_1, \ldots, a_N)$ and the state is ω, then a cost $C(a, \omega)$ is incurred. C is a real-valued function that is measurable with respect to the sigma field generated on the product space $(\mathbf{Z}^m \times \Omega)$ by the Borel sets $\mathcal{B}(\mathbf{Z}^m)$ of $\mathbf{Z}^m$ and by the σ-algebra $\mathcal{F}$. On discrete spaces the σ-algebras are not necessary, but they are retained to simplify the notation and to emphasize the analogy with the case of continuous spaces.

Contrary to a conventional optimal decision problem, we assume that each player has his own observation of the underlying event space. This is implemented as follows. We assume that for each player i there exists an observation space Y_i, a given sigma field $\mathcal{Y}_i$ of subsets of Y_i, and a function $h_i : \Omega \to Y_i$ that is $\mathcal{Y}_i$-measurable. If the event ω occurs, then player i will observe $h_i(\omega)$, and thus each function h_i is a random variable on $(\Omega, \mathcal{F}, \mathcal{P})$. We refer to $\mathcal{Y}_i$ as the *information subfield* of player i, and we define $\mathcal{F}^{h_i} = \{h_i^{-1}(A) \mid A \in \mathcal{Y}_i\}$ as the sigma field that is induced by h_i. The decision that player i makes can depend only on its observation, and thus the set of admissible control laws $\mathcal{U}_i$ for player i is defined by the set of $\mathbf{Z}$-valued functions that are $\mathcal{Y}_i$-measurable. We let $\mathcal{U} = \mathcal{U}_1 \times \cdots \times \mathcal{U}_N$ denote the set of admissible team strategies.

Note that under a strategy γ the team action a is by definition a function of the state ω, i.e., $a = \gamma(h(\omega))$ with $h(\omega) = (h_1(\omega), \ldots, h_N(\omega))$. Under a strategy γ the expected cost of the strategy $J(\gamma)$ is now defined as

$$J(\gamma) = E\{C(\gamma(h(\omega)), \omega)\}, \tag{2.1}$$

where E denotes the expectation with respect to $\mathcal{P}$.

DEFINITION 2.1. *A strategy* $\gamma^* \in \mathcal{U}$ *is* optimal *if*

$$J(\gamma^*) \le J(\gamma), \qquad \gamma \in \mathcal{U}. \tag{2.2}$$

The next definition is a variation on the concept of cw-optimality as was introduced in Radner [7]. In that paper a strategy γ is called pbpo if $J(\gamma)$ cannot be improved by changing the strategy for one player alone. The idea that lies behind this definition is that under some extra conditions on the cost function C there exists only one pbpo strategy and this strategy is by the conditions on the cost function then also optimal. As an extra bonus the computation of a pbpo strategy is much easier than for the globally optimal strategy, since the optimization problem in a sense becomes separable. In our model with discrete action spaces we introduce the same concept of pbpo. This is done in a way that is different from [7] and [5], where stationarity is defined by means of the differential of a conditional expectation with respect to the individual decisions. An example of the use of pbpo strategies to determine an optimal solution for a detection problem can be found in [9].

DEFINITION 2.2. *A team strategy* $\overline{\gamma} \in \mathcal{U}$ *is pbpo if* $J(\overline{\gamma}) < \infty$ *and for each player* i, $i = 1, \ldots, N$,

$$E\Big[C(\overline{\gamma}(h(\omega)), \omega) \mid \mathcal{F}^{h_i}\Big] \le E\Big[C(\gamma(h(\omega)), \omega) \mid \mathcal{F}^{h_i}\Big], \qquad (\mathcal{P} - a.s.) \tag{2.3}$$

for all team strategies $\gamma \in \mathcal{U}$ *with*

$$\gamma_j \equiv \overline{\gamma}_j \textit{ for all } j = 1, \ldots, N, \textit{ with } j \ne i.$$

A team strategy is called strictly pbpo if the $\le$ *sign in* (2.3) *is replaced by a strict inequality* ($<$).

Note that the inequality (2.3) is well defined, since both conditional expectations are random variables on the same probability space. In fact, this implies that the

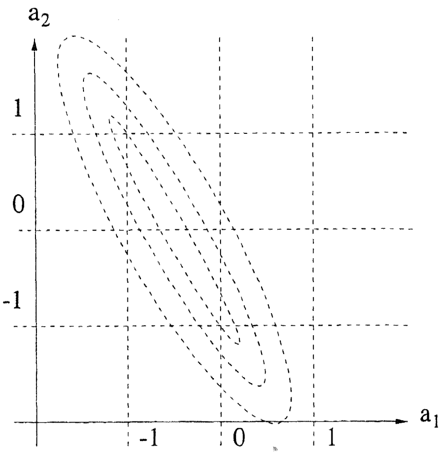

FIG. 2.1. *An example of a continuous convex function restricted to a discrete space.*

inequalities can be replaced by the usual stochastic order (cf. Shaked and Shanthikumar [8, pp. 3, 5]).

If the formulation of the team problem is such that there exists a natural continuation $\overline{C}$ of the cost function from $\mathbf{Z}^N \times \Omega$ to $\mathbf{R}^N \times \Omega$, and $\overline{C}$ is a convex, differentiable and locally finite function, then this continuous version of the problem has a unique optimal solution and this solution is also the only pbpo solution (see Radner [7]). There is no guarantee that restriction to a discrete action space leads to an optimal solution that is in the "neighborhood" of the continuous solution. By neighborhood we mean that the discrete solution is close to the continuous solution, in the usual metric of $\mathbf{R}^N$. Consider for example the cost function

$$C(a_1, a_2, \omega) = (a_1 + 0.5 - \sqrt{3}a_2)^2 + 2(\sqrt{3}(a_1 + 0.5) + a_2)^2 \tag{2.4}$$

for a team problem with only one possible outcome ω, i.e., a deterministic team or an ordinary minimization problem. From the picture of the contour lines as in Figure 2.1, we can see that the continuous minimum is in $(a_1, a_2) = (-0.5, 0)$, while the two integer valued minima are in $(a_1, a_2) = (-1, 1)$ and $(0, -1)$. This example can be modified, however, such that the "discrete" solution is arbitrarily far from the "continuous" solution. Note also that the discrete nature of the problem in this case allows *two* solutions.

In many problems the continuation of C to $\mathbf{R}^N$ may not be as straightforward as in the example. If that is the case, then we might try to construct one. This is where the idea of multimodularity comes in, and it is shown in detail in Hajek [2]. Hajek constructs atoms that span the space $\mathbf{R}^N$. Each atom contains exactly $m + 1$ extreme points, and these points lie in $\mathbf{Z}^N$. The continuation $\overline{C}$ of C is piecewise affine on all the atoms. If the function C is multimodular on $\mathbf{Z}^N$, then the continuation of the function is convex in $\mathbf{R}^N$. Unfortunately this continuation is not differentiable, so the results of Radner [7] and Krainak, Speyer, and Marcus [5] cannot be applied here (see also the example in [7, p. 802]). For a justification of the use of multimodular cost functions see the remarks in [2, p. 546] and Bartroli and Stidham [1]. The discussion of multimodularity and its relation to convexity is presented in the appendix.

3. Solution of a class of team problems. In this section we investigate a special class of team problems. It is intended as an example for team problems with discrete action spaces. It will also serve to indicate the possibilities of using multimodularity properties in solving this kind of team problem. We shall discuss various properties of optimal and cw-optimal strategies. These properties can be used in a procedure to search for the optimal team strategy.

We first need to introduce multimodular functions. For this we consider functions defined on $\mathbf{Z}^m$. We define the vectors $v_0, v_1, \ldots, v_m$ in $\mathbf{Z}^m$ as

$$\begin{aligned} v_0 &= (-1, 0, \ldots, 0), \\ v_1 &= (1, -1, 0, \ldots, 0), \\ v_2 &= (0, 1, -1, 0, \ldots, 0), \\ &\vdots \\ v_{m-1} &= (0, \ldots, 1, -1), \\ v_m &= (0, \ldots, 0, 1), \end{aligned}$$

and we let $\mathcal{V} = \{v_0, v_1, \ldots, v_m\}$. Note that any subset of m vectors of $\mathcal{V}$ is a basis for $\mathbf{Z}^m$, and furthermore we remark that

$$v_0 + v_1 + \cdots + v_m = (0, \ldots, 0). \tag{3.1}$$

DEFINITION 3.1. *A function f on $\mathbf{Z}^m$ for $m \geq 2$ is said to be* multimodular *if for all $z \in \mathbf{Z}^m$,*

$$g(z + v_i) + g(z + v_j) \geq g(z) + g(z + v_i + v_j) \tag{3.2}$$

for any $v_i, v_j \in \mathcal{V}$ and $v_i \neq v_j$.

For a function f on $\mathbf{Z}^m$, $n \in \{1, \ldots, m\}$ and $z \in \mathbf{Z}^m$, we denote the first-order n-difference of f at z as

$$\Delta_n f(z) := f(z + e_n) - f(z), \tag{3.3}$$

where e_n denotes the nth unit vector.

DEFINITION 3.2. *Let f be a real-valued function defined on $\mathbf{Z}^m$. A point $z \in \mathbf{Z}^m$ is called* minimal *for f if $f(z) \leq f(y)$ for all $y \in \mathbf{Z}^m$, $y \neq z$, and it is called* coordinatewise minimal (cw-minimal) *if $f(z) \leq f(z + \lambda e_i)$ for any $i \in \{1, \ldots, m\}$ and any $\lambda \in \mathbf{Z}$, $\lambda \neq 0$. We define a point $z \in \mathbf{Z}^m$ to be* strictly minimal *or* strictly cw-minimal *if these inequalities are replaced by strict inequalities.*

With these definitions we can now introduce the class of team problems that we want to describe. We consider a problem with two players. The underlying event space Ω has three elements, numbered as $\Omega = \{1, 2, 3\}$, and each element occurs with the same probability. We assume that $\mathcal{F}$ is the sigma algebra generated by $\{\{1\}, \{2\}, \{3\}\}$. The action sets for both players are $\mathbf{Z}$.

We assume that the two players have distinctly different information patterns. Player 1 cannot distinguish between events 1 and 2, so $\mathcal{F}^{h_1} = \sigma(\{\{1, 2\}, \{3\}\})$, while player 2 cannot distinguish between events 2 and 3, so $\mathcal{F}^{h_2} = \sigma(\{\{1\}, \{2, 3\}\})$.

For the cost structure we assume the following.

ASSUMPTION 3.3. *For each possible outcome ω, the cost function $C(u_1, u_2, \omega)$ is multimodular as a function of the decision variables $(u_1, u_2) \in \mathbf{Z}^2$.*

In this section we shall discuss the properties of pbpo strategies. For this we shall make use of the classification of cw-minimal points of multimodular functions on $\mathbf{Z}^2$. The following lemma is a direct consequence of the results of Appendix A, but we specifically state it here for easy reference.

LEMMA 3.4. *If $g : \mathbf{Z}^2 \to \mathbf{R}$ is a multimodular function and $y = (y_1, y_2)$ and $z = (z_1, z_2)$ are two distinct strictly cw-minimal points of g, where $z_1 \geq y_1$, then there is some $B > 0$ such that*

1. $(z_1, z_2) = (y_1 + B, y_2 - B)$;
2. *for all b, $0 < b < B$, $(y_1 + b, y_2 - b)$ is also cw-minimal;*
3. *if $g(z) > g(y)$, then the minimum of g cannot be in the set $\{(y_1+b, y_2-b) | b \geq B\}$;*
4. *if both y and z, for some $B \geq 0$, are minimal for g, then $(y_1 + b, y_2 - b)$ is minimal for all $0 \leq b \leq B$;*
5. *if both $g(y_1 + 1, y_1 - 1) \geq g(y_1, y_2)$ and $g(y_1 - 1, y_1 + 1) \geq g(y_1, y_2)$, then (y_1, y_2) is minimal; if the inequalities are strict, then the minimum is also unique.*

Proof. See Appendix A, Lemma A.10. □

Note that (strict) multimodularity of a function does not guarantee that the minimum is unique. It can take the form of a line segment $\{(y_1+z, y_2-z) \mid 0 \leq z \leq B\}$ for some $(y_1, y_2) \in \mathbf{Z}^2$ and $B \geq 0$. There can be only one such segment, however.

We now introduce the notation for a team decision strategy. We already saw in section 2 that the observations of the players can be modelled by functions that are defined on the event space. The observations of player 1 are given by a function $h_1 : \Omega \to \{1, 3\}$ of the state, where

$$h_1(\omega) = \begin{cases} 1 & \text{if } \omega = 1, 2, \\ 3 & \text{if } \omega = 3. \end{cases} \tag{3.4}$$

Similarly, we represent the information pattern of player 2 by a function h_2, which is defined as

$$h_2(\omega) = \begin{cases} 1 & \text{if } \omega = 1, \\ 3 & \text{if } \omega = 2, 3. \end{cases} \tag{3.5}$$

With this definition we can now represent a team decision function as $\gamma = (\gamma_{11}, \gamma_{13}, \gamma_{21}, \gamma_{23}) \in \mathbf{Z}^4$, where γ_{ij} represent the action that γ prescribes when player i gets observation j. Finally, we can now write the expected cost $J(\gamma)$ as a function of the team decision rule γ as

$$J(\gamma_{11}, \gamma_{13}, \gamma_{21}, \gamma_{23}) = \frac{1}{3}C(\gamma_{11}, \gamma_{21}, 1) + \frac{1}{3}C(\gamma_{11}, \gamma_{23}, 2) + \frac{1}{3}C(\gamma_{13}, \gamma_{23}, 3). \tag{3.6}$$

In principle, this makes finding the optimal strategy an optimization problem on $\mathbf{Z}^4$.

The properties that γ has to satisfy for optimality and cw-optimality are summarized in the following lemma.

LEMMA 3.5. *A team decision strategy $\gamma^* = (\gamma_{11}{}^*, \gamma_{13}{}^*, \gamma_{21}{}^*, \gamma_{23}{}^*)$ is minimal if*

$$(\gamma_{11}{}^*, \gamma_{13}{}^*, \gamma_{21}{}^*, \gamma_{23}{}^*) = \underset{(\gamma_{11}, \gamma_{13}, \gamma_{21}, \gamma_{23}) \in \mathbf{Z}^4}{\arg\min} J(\gamma_{11}, \gamma_{13}, \gamma_{21}, \gamma_{23}), \tag{3.7}$$

or, in other words, if γ^* is a minimum for J. A strategy $\gamma = (\overline{\gamma}_{11}, \overline{\gamma}_{13}, \overline{\gamma}_{21}, \overline{\gamma}_{23})$ is strictly pbpo if J is strictly cw-minimal in γ, or

$$J(\overline{\gamma}_{11}, \overline{\gamma}_{13}, \overline{\gamma}_{21}, \overline{\gamma}_{23}) < J(u_{11}, \overline{\gamma}_{13}, \overline{\gamma}_{21}, \overline{\gamma}_{23}) \text{ for all } u_{11} \in \mathbf{Z}, u_{11} \neq \overline{\gamma}_{11}, \tag{3.8}$$

$$J(\overline{\gamma}_{11}, \overline{\gamma}_{13}, \overline{\gamma}_{21}, \overline{\gamma}_{23}) < J(\overline{\gamma}_{11}, u_{13}, \overline{\gamma}_{21}, \overline{\gamma}_{23}) \text{ for all } u_{13} \in \mathbf{Z}, u_{13} \neq \overline{\gamma}_{13}, \tag{3.9}$$

$$J(\overline{\gamma}_{11}, \overline{\gamma}_{13}, \overline{\gamma}_{21}, \overline{\gamma}_{23}) < J(\overline{\gamma}_{11}, \overline{\gamma}_{13}, u_{21}, \overline{\gamma}_{23}) \text{ for all } u_{21} \in \mathbf{Z}, u_{21} \neq \overline{\gamma}_{21}, \tag{3.10}$$

$$J(\overline{\gamma}_{11}, \overline{\gamma}_{13}, \overline{\gamma}_{21}, \overline{\gamma}_{23}) < J(\overline{\gamma}_{11}, \overline{\gamma}_{13}, \overline{\gamma}_{21}, u_{23}) \text{ for all } u_{23} \in \mathbf{Z}, u_{23} \neq \overline{\gamma}_{23}. \tag{3.11}$$

Proof. The first statement is immediate from the definition of optimality. The second statement follows from the fact that, by definition, a team strategy $\overline{\gamma} = (\overline{\gamma}_{11}, \overline{\gamma}_{13}, \overline{\gamma}_{21}, \overline{\gamma}_{23})$ is strictly pbpo if it satisfies the following set of equations:

$$C(\overline{\gamma}_{11}, \overline{\gamma}_{21}, 1) + C(\overline{\gamma}_{11}, \overline{\gamma}_{23}, 2) < C(\gamma_{11}, \overline{\gamma}_{21}, 1) + C(\gamma_{11}, \overline{\gamma}_{23}, 2), \tag{3.12}$$

$$C(\overline{\gamma}_{13}, \overline{\gamma}_{23}, 3) < C(\gamma_{13}, \overline{\gamma}_{23}, 3), \tag{3.13}$$

$$C(\overline{\gamma}_{11}, \overline{\gamma}_{21}, 1) < C(\overline{\gamma}_{11}, \gamma_{21}, 1), \tag{3.14}$$

$$C(\overline{\gamma}_{11}, \overline{\gamma}_{23}, 2) + C(\overline{\gamma}_{13}, \overline{\gamma}_{23}, 3) < C(\overline{\gamma}_{11}, \gamma_{23}, 2) + C(\overline{\gamma}_{13}, \gamma_{23}, 3) \tag{3.15}$$

for all $\gamma_{11}, \gamma_{13}, \gamma_{21}, \gamma_{23} \in \mathbf{Z}$. For instance, inequality (3.12) is immediate from the definition of pbpo and the fact that $\mathcal{F}^{h_1} = \sigma(\{\{1,2\},\{3\}\})$. Since $C(\overline{\gamma}_{13}, \overline{\gamma}_{23}, 3)$ is independent of $\overline{\gamma}_{11}$, (3.12) thus implies that $J(\overline{\gamma}_{11}, \overline{\gamma}_{13}, \overline{\gamma}_{21}, \overline{\gamma}_{23}) < J(\gamma_{11}, \overline{\gamma}_{13}, \overline{\gamma}_{21}, \overline{\gamma}_{23})$ for all $\gamma_{11} \in \mathbf{Z}$. In a similar way, one can prove the cw-minimality of J for the other components of $\overline{\gamma}$. □

In the remainder of this section we shall exploit the special nature of multimodular functions to derive properties of optima and cw-optima. We show how one can search for other (coordinatewise) minima starting from a cw-minimum. The main result is as follows.

THEOREM 3.6. *If $\alpha = (\alpha_{11}, \alpha_{13}, \alpha_{21}, \alpha_{23})$ and $\beta = (\beta_{11}, \beta_{13}, \beta_{21}, \beta_{23})$ are both strictly pbpo strategies, then they have to satisfy either $M\beta^T \leq M\alpha^T$ or $M\beta^T \geq M\alpha^T$ for*

$$M = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ -1 & -1 & 0 & -1 \\ -1 & -1 & -1 & 0 \end{pmatrix}. \tag{3.16}$$

The vector inequality $\leq$ is to be interpreted componentwise, i.e., $z \leq y$ if and only if $z_i \leq y_i$ for all i.

Proof. For any strategy $z = (z_{11}, z_{13}, z_{21}, z_{23})$, the expected cost $J(z)$ for this strategy is

- multimodular in (z_{11}, z_{21}) if (z_{13}, z_{23}) is fixed,
- multimodular in (z_{11}, z_{23}) if (z_{13}, z_{21}) is fixed,
- multimodular in (z_{13}, z_{21}) if (z_{11}, z_{23}) is fixed,
- multimodular in (z_{13}, z_{23}) if (z_{11}, z_{21}) is fixed.

Now assume that α is strictly pbpo. This implies that $\Delta_{11}J(\alpha) > 0$. Using the fact that J is multimodular in (z_{11}, z_{21}) for (z_{13}, z_{23}) fixed, we get that $\Delta_{11}J(\alpha_{11}+m, z_{13}, \alpha_{21}-m, \alpha_{23}) > 0$ for all $m \geq 0$ and all $z_{13} \in \mathbf{Z}$. Note that $\Delta_{11}J(z)$ is independent of z_{13}. Next use the multimodularity of J in (z_{11}, z_{23}) for (z_{13}, z_{21}) fixed to get $\Delta_{11}J(\alpha_{11} + m+n, z_{13}, \alpha_{21}-m, \alpha_{23}-n) > 0$ for all $m, n \geq 0$ and all z_{13}. Finally, use the fact that for any multimodular function f defined on $\mathbf{Z}^2$, $\Delta_i f(x) \leq \Delta_i f(x+e_j)$ for $x \in \mathbf{Z}^2$ and $i, j = 1, 2$, to prove that $\Delta_{11}J(z) > 0$ for all z in

$$\mathcal{G}_{11+} := \{(\alpha_{11}+m+n+i, z_{13}, \alpha_{21}-m+j, \alpha_{23}-n+k) \mid i, j, k, m, n \geq 0, z_{13} \in \mathbf{Z}\}.$$

Specifically, this means that the points in the interior of $\mathcal{G}_{11+}$ cannot be pbpo. With the interior of $\mathcal{G}_{11+}$ we mean the points z with $z_{11} + z_{21} + z_{23} > \alpha_{11} + \alpha_{21} + \alpha_{23}$. Analogously, we can exploit the fact that $\Delta_{11}J(\alpha_{11} - 1, \alpha_{13}, \alpha_{21}, \alpha_{23}) < 0$ to show that $\Delta_{11}J(z) < 0$ for z in

$$\mathcal{G}_{11-} := \{(\alpha_{11}-1-m-n-i, z_{13}, \alpha_{21}+m-j, \alpha_{23}+n-k) \mid i, j, k, m, n \geq 0, z_{13} \in \mathbf{Z}\}.$$

In the same manner, we also find that the fact that α is strictly pbpo implies $\Delta_{13}J(z) > 0$ for all z in

$$\mathcal{G}_{13+} := \{(z_{11}, \alpha_{13} + m + n + i, \alpha_{21} - m + j, \alpha_{23} - n + k) \mid i, j, k, m, n \geq 0, z_{13} \in \mathbf{Z}\},$$

and $\Delta_{13}J(z) < 0$ for z in

$$\mathcal{G}_{13-} := \{(z_{11}, \alpha_{13} - 1 - m - n - i, \alpha_{21} + m - j, \alpha_{23} + n - k) \mid i, j, k, m, n \geq 0, z_{11} \in \mathbf{Z}\}.$$

For the decisions of player 2, we get, in a similar manner, $\Delta_{21}J(z) > 0$ for z in

$$\mathcal{G}_{21+} := \{(\alpha_{11} - m + j, \alpha_{13} - n + k, \alpha_{21} + m + n + i, z_{23},) \mid i, j, k, m, n \geq 0, z_{23} \in \mathbf{Z}\},$$

and $\Delta_{21}J(z) < 0$ for all z in

$$\mathcal{G}_{21-} := \{(\alpha_{11} + m - j, \alpha_{13} + n - k, \alpha_{21} - 1 - m - n - i, z_{23},) \mid i, j, k, m, n \geq 0, z_{23} \in \mathbf{Z}\},$$

and finally $\Delta_{23}J(z) > 0$ for z in

$$\mathcal{G}_{23+} := \{(\alpha_{11} - m + j, \alpha_{13} - n + k, z_{21}, \alpha_{23} + m + n + i,) \mid i, j, k, m, n \geq 0, z_{21} \in \mathbf{Z}\},$$

and $\Delta_{23}J(z) < 0$ for z in

$$\mathcal{G}_{23-} := \{(\alpha_{11} + m - j, \alpha_{13} + n - k, z_{21}, \alpha_{23} - 1 - m - n - i,) \mid i, j, k, m, n \geq 0, z_{21} \in \mathbf{Z}\}.$$

Observe that, by the same reasoning as for $\mathcal{G}_{11+}$, there cannot be other pbpo strategies in the interior of any of the $\mathcal{G}$'s.

Now assume that β is also pbpo and $\beta_{11} > \alpha_{11}$. Since β cannot lie in $\mathcal{G}_{11+}$, this implies that

$$\beta_{11} + \beta_{21} + \beta_{23} \leq \alpha_{11} + \alpha_{21} + \alpha_{23}. \tag{3.17}$$

This proves the second inequality of $M\beta^T \leq M\alpha^T$. Since $\beta_{11} > \alpha_{11}$, it also implies that $\beta_{21} + \beta_{23} \leq \alpha_{21} + \alpha_{23}$. This in itself implies that at least one of the inequalities $\beta_{21} \leq \alpha_{21}$ and $\beta_{23} \leq \alpha_{23}$ hold. We shall prove that, in fact, both inequalities hold. Assume that $\beta_{21} > \alpha_{21}$. Since β is pbpo, it cannot lie in the interior of $\mathcal{G}_{21+}$, and thus $\beta_{11} + \beta_{13} + \beta_{21} < \alpha_{11} + \alpha_{13} + \alpha_{21}$. Because of the assumptions on β_{11} and β_{21}, we conclude that $\beta_{13} < \alpha_{13}$. Now β is also outside the interior of $\mathcal{G}_{13-}$, so $\beta_{13} + \beta_{21} + \beta_{23} \geq \alpha_{13} + \alpha_{21} + \alpha_{23}$. This contradicts $\beta_{13} < \alpha_{13}$ and $\beta_{21} + \beta_{23} \leq \alpha_{21} + \alpha_{23}$. We can thus conclude that both $\beta_{21} \leq \alpha_{21}$ and $\beta_{23} \leq \alpha_{23}$.

Since $\beta_{21} \leq \alpha_{21}$ and β is outside of the interior of $\mathcal{G}_{21-}$, this implies

$$\beta_{11} + \beta_{13} + \beta_{21} \geq \alpha_{11} + \alpha_{13} + \alpha_{21}, \tag{3.18}$$

and this proves the fourth component of the matrix inequality. Analogously to the reasoning above, we conclude from this that $\beta_{13} \geq \alpha_{13}$. This, together with the fact that β is outside of the interior of $\mathcal{G}_{13+}$, implies

$$\beta_{13} + \beta_{21} + \beta_{23} \leq \alpha_{13} + \alpha_{21} + \alpha_{23}, \tag{3.19}$$

the first inequality, and finally $\beta_{23} \leq \alpha_{23}$ leads us to

$$\beta_{11} + \beta_{13} + \beta_{23} \geq \alpha_{11} + \alpha_{13} + \alpha_{23}, \tag{3.20}$$

the third inequality.

If we start with $\beta_{11} < \alpha_{11}$, then this same reasoning gives us $M\beta^T \geq M\alpha^T$. The case where $\beta_{11} = \alpha_{11}$ is treated in the following lemma. □

Note that the proof of Theorem 3.6 does not depend on the particular information patterns of this problem. It is based solely on the additive structure of the expected cost and on the multimodularity of the cost function.

Theorem 3.6 provides us with a characterization of the areas around a known pbpo strategy, where we might find other pbpo strategies. If we want to design a search procedure, then we may want to know in which direction we have to search in the immediate neighborhood of α. By immediate neighborhood of α we refer to those strategies β, with $|\beta_{ij} - \alpha_{ij}| \leq 1$ for all coefficients β_{ij}. The following lemmas provide us with the necessary results.

LEMMA 3.7. *If $\alpha = (\alpha_{11}, \alpha_{13}, \alpha_{21}, \alpha_{23})$ and $\beta = (\beta_{11}, \beta_{13}, \beta_{21}, \beta_{23})$ are both strictly pbpo strategies and $\alpha \neq \beta$, then*

- *if $\alpha_{11} = \beta_{11}$, then $\alpha_{21} = \beta_{21}$,*
- *if $\alpha_{23} = \beta_{23}$, then $\alpha_{13} = \beta_{13}$.*

Proof. Assume that $\alpha_{11} = \beta_{11}$. Define $g(z_{13}, z_{21}, z_{23}) := J(\alpha_{11}, z_{13}, z_{21}, z_{23})$; then

$$g(z_{13}, z_{21}, z_{23}) = \frac{1}{3}\Big[C(\alpha_{11}, z_{21}, 1)\Big] + \frac{1}{3}\Big[C(\alpha_{11}, z_{23}, 2) + C(z_{13}, z_{23}, 3)\Big].$$

It is clear that g is a convex function of z_{21} on $\mathbf{Z}$, and that $\alpha_{21} = \arg\min_{z_{21}\in\mathbf{Z}} g(z_{13}, z_{21}, z_{23})$ is independent of z_{13} and z_{23}. Since both α and β are pbpo, they are cw-minimal for g and thus $\alpha_{21} = \beta_{21}$. The proof for $\alpha_{23} = \beta_{23}$ proceeds analogously. □

Lemma 3.7 does not tell us what happens if $\alpha_{13} = \beta_{13}$ or $\alpha_{21} = \beta_{21}$. It appears that we can construct two strictly pbpo strategies α and β that have some components in common. All these possibilities are summarized in the next lemma.

LEMMA 3.8. *If $\alpha = (\alpha_{11}, \alpha_{13}, \alpha_{21}, \alpha_{23})$ and $\beta = (\beta_{11}, \beta_{13}, \beta_{21}, \beta_{23})$ are two distinct strictly pbpo strategies, then one of the following possibilities holds:*

1. *α and β differ in at least two components,*
2. *$(\alpha_{11}, \alpha_{13}) \neq (\beta_{11}, \beta_{13})$,*
3. *$(\alpha_{21}, \alpha_{23}) \neq (\beta_{21}, \beta_{23})$,*
4. *$(\alpha_{11}, \alpha_{23}) \neq (\beta_{11}, \beta_{23})$.*

Proof. Part 1. Assume that the statement is false and that, for instance, α_{11} and β_{11} are the only two coefficients that are different. This implies $C(\alpha_{11}, \alpha_{21}, 1) + C(\alpha_{11}, \alpha_{23}, 2) < C(\beta_{11}, \alpha_{21}, 1) + C(\beta_{11}, \alpha_{23}, 2) = C(\beta_{11}, \beta_{21}, 1) + C(\beta_{11}, \beta_{23}, 2) < C(\alpha_{11}, \beta_{21}, 1) + C(\alpha_{11}, \beta_{23}, 2)$. This gives a contradiction, since $\alpha_{21} = \beta_{21}$ and $\alpha_{23} = \beta_{23}$.

Parts 2 and 3. Assume $(\alpha_{11}, \alpha_{13}) = (\beta_{11}, \beta_{13})$. Consider the function $g(z_{21}, z_{23})$ defined as $J(\alpha_{11}, \alpha_{13}, z_{21}, z_{23})$, so

$$g(z_{21}, z_{23}) = \frac{1}{3}\Big[C(\alpha_{11}, z_{21}, 1) + C(\alpha_{11}, z_{23}, 2) + C(\alpha_{13}, z_{23}, 3)\Big] =: g_1(z_{21}) + g_2(z_{23}).$$

Since α is a strictly pbpo strategy, $(\alpha_{21}, \alpha_{23})$ must be cw-minimal for g. The function g is multimodular in (z_{21}, z_{23}), and thus g_1 and g_2 are convex functions of z_{21} and z_{23}, respectively. This implies that g has a unique minimum, so $\alpha = \beta$.

Part 4. From Lemma 3.7, we see that $\alpha_{11} = \beta_{11}$ implies $\alpha_{21} = \beta_{21}$ and $\alpha_{23} = \beta_{23}$ implies $\alpha_{13} = \beta_{13}$. □

Now assume that we have found a pbpo strategy α, and we want to check the strategies in the neighborhood of α to see whether they are pbpo. Among the pbpo strategies we can then check the value of the expected cost J for optimality. The neighborhood of a strategy α is the set

$$\{(\beta_{11}, \beta_{13}, \beta_{21}, \beta_{23}) \mid |\beta_{ij} - \alpha_{ij}| \leq 1, \quad i = 1, 2, \ j = 1, 3\}.$$

Note that there are 80 strategies (excluding α) in this set. If we combine the results of Theorem 3.6 and Lemmas 3.7 and 3.8, the following corollary shows how 68 of these strategies can be eliminated.

COROLLARY 3.9. *If $\alpha = (\alpha_{11}, \alpha_{13}, \alpha_{21}, \alpha_{23})$ is a strictly pbpo strategy, then the set of possible pbpo strategies in the neighborhood of α are the strategies of the form $\beta = \alpha \pm \epsilon$ with $\epsilon \in \{(1,0,-1,0), (0,1,0,-1), (1,0,0,-1), (1,0,-1,-1), (1,1,0,-1), (1,1,-1,-1)\}$. Outside of this set there cannot be pbpo strategies in the neighborhood of α.*

Proof. We sketch the proof in three steps.

1. Assume that a pbpo strategy $\beta = \alpha + \epsilon$ is of the form $\epsilon = (1, -1, \epsilon_{21}, \epsilon_{23})$ for any $\epsilon_{21}, \epsilon_{23} \in \{-1, 0, 1\}$. We show that β cannot satisfy $M\beta^T \leq M\alpha^T$, since then the third and fourth matrix inequalities imply that $\alpha_{21} \leq \beta_{21}$ and $\alpha_{23} \leq \beta_{23}$. These again imply via the first and last inequalities that $\alpha_{11} \geq \beta_{11}$ and $\alpha_{13} \geq \beta_{13}$, and this gives a contradiction with the assumption on the signs of the first two coefficients of ϵ. The inequality $M\beta^T \geq M\alpha^T$ gives a similar contradiction. Analogously, we can show that ϵ cannot be of the form $\epsilon = (-1, 1, \epsilon_{21}, \epsilon_{23})$ or $\pm(\epsilon_{11}, \epsilon_{13}, 1, -1)$.

2. Assume that a pbpo strategy $\beta = \alpha + \epsilon$ is of the form $\epsilon = (1, \epsilon_{13}, \epsilon_{21}, 1)$ for any ϵ_{13}, $\epsilon_{21} \in \{-1, 0, -1\}$. If β were to satisfy $M\beta^T \leq M\alpha^T$, then this would imply $\epsilon_{21} \leq 2$, while $M\beta^T \geq M\alpha^T$ would imply $\epsilon_{21} \geq 2$. Both contradictions show that the ϵ cannot be of the proposed form. Similarly, we can show that ϵ cannot be of the forms $\pm(1, \epsilon_{13}, 1, \epsilon_{23})$, $\pm(\epsilon_{11}, 1, 1, \epsilon_{23})$, or $\pm(\epsilon_{11}, 1, \epsilon_{21}, 2)$.

3. Combine 1 and 2 to get the possible candidates of ϵ. □

Note that within this set there are pairs of strategies that differ in exactly one coefficient, and thus of these pairs only one can be strictly pbpo. Furthermore, for any ϵ in the set of Corollary 3.9, both $\alpha + \epsilon$ and $\alpha - \epsilon$ can be pbpo, but at most one of these two strategies can have an expected cost smaller than $J(\alpha)$ (see Lemma A.6). Finally, if it turns out that $J(\alpha + \epsilon) \geq J(\alpha)$ for some ϵ, then the same lemma ensures that $J(\alpha + k\epsilon) \geq J(\alpha)$ for all $k \in \mathbf{N}$, and thus these points cannot be minimal. In immediate consequence of this is the following, which can be proven analogously to part 5 of Lemma 3.4.

COROLLARY 3.10. *If α is a strictly pbpo strategy and for all ϵ as in Corollary 3.9 we have $J(\alpha+\epsilon) \geq J(\alpha)$ and $J(\alpha-\epsilon) \geq J(\alpha)$, then α is minimal. If all the inequalities are strict, then α is the unique optimal strategy.*

This concludes our exploration of this class of team problems. We have developed a check for the optimality of a team strategy, and we have given the description of a procedure to search for the optimal strategy.

4. Conclusions.

In this paper we have discussed team problems with discrete action spaces. Inspired by known results for problems on continuous spaces with convex cost functions, we have introduced multimodularity as a natural abstraction of convexity onto discrete spaces.

In the class of team problems of section 3, we have seen that multimodularity of the cost function translates to properties for the expected cost as a function of the strategy. These properties allow us to check for optimality of a strategy, and they indicate how the complexity of a search for the optimum can be reduced. The example, however, indicates that the complexity is still rather high, and we feel that it must be possible to reduce it even more by exploiting the multimodularity even further. This is a topic for future research.

If we extend the results of section 3 to a model with a larger observation space for both players, then most of the results of the section remain valid. The proof of Theorem 3.6 relies only on the multimodularity of the cost function and not on the structure of the information patterns. This means that it is a straightforward exercise to extend the results of Theorem 3.6 to a larger observation space. In Lemmas 3.7 and 3.8, the particular structure of the information patterns *is* used, and any extension in this direction has to be done on an ad hoc basis.

Extending the team problem to more than two players is not a trivial task. If we try to mimic the proof of Theorem 3.6 for an example with three players, then even for small observation spaces it is not clear if a matrix inequality of the form $M\alpha^T \leq M\beta^T$ will hold and what the form of M will be.

Appendix A. Multimodular functions and optimality.

In this appendix, we introduce the concept of multimodular functions. Furthermore, we define cw-optimality for this class of functions, and we show its relation to ordinary optimality. We present a classification of cw-optimal points, and we specify a procedure to search for the optimum. For a more elaborate introduction to multimodular functions, we refer to Hajek [2].

We consider functions defined on $\mathbf{Z}^m$. We define the vectors $v_0, v_1, \ldots, v_m$ in $\mathbf{Z}^m$ as

$$\begin{aligned} v_0 &= (-1, 0, \ldots, 0), \\ v_1 &= (1, -1, 0, \ldots, 0), \\ v_2 &= (0, 1, -1, 0, \ldots, 0), \\ &\vdots \\ v_{m-1} &= (0, \ldots, 1, -1), \\ v_m &= (0, \ldots, 0, 1), \end{aligned}$$

and we let $\mathcal{V} = \{v_0, v_1, \ldots, v_m\}$. Note that any subset of m vectors of $\mathcal{V}$ is a basis for $\mathbf{Z}^m$, and furthermore we remark that

$$v_0 + v_1 + \cdots + v_m = (0, \ldots, 0). \tag{A.1}$$

DEFINITION A.1. *A function f on $\mathbf{Z}^m$ for $m \geq 2$ is said to be* multimodular *if for all $z \in \mathbf{Z}^m$,*

$$g(z + v_i) + g(z + v_j) \geq g(z) + g(z + v_i + v_j) \tag{A.2}$$

for any $v_i, v_j \in \mathcal{V}$, and $v_i \neq v_j$.

For a function f on $\mathbf{Z}^m$, $n \in \{1, \ldots, m\}$, and $z \in \mathbf{Z}^m$ we denote the first-order n-difference of f at z as

$$\Delta_n f(z) := f(z + e_n) - f(z), \tag{A.3}$$

where e_n denotes the nth unit vector.

DEFINITION A.2. *Let f be a real-valued function defined on $\mathbf{Z}^m$. A point $z \in \mathbf{Z}^m$ is called* minimal *for f if $f(z) \leq f(y)$ for all $y \in \mathbf{Z}^m$, $y \neq z$, and it is called* coordinatewise minimal (cw-minimal) *if $f(z) \leq f(z + \lambda e_i)$ for any $i \in \{1, \ldots, m\}$ and any $\lambda \in \mathbf{Z}$, $\lambda \neq 0$. We define a point $z \in \mathbf{Z}^m$ to be* strictly minimal *or* strictly cw-minimal *if these inequalities are replaced by strict inequalities.*

Note that of course a minimal point is also cw-minimal. The following lemma gives an indication of the properties of cw-optimal points of a multimodular function.

LEMMA A.3. *Let z^* be a strictly cw-minimal point of a multimodular function f, let z be any point in $\mathbf{Z}^m$, and let the coordinates of $z - z^*$ with respect to the bases $\{v_1, \dots, v_m\}$ and $\{v_0, \dots, v_{m-1}\}$ be*

$$z - z^* = k_1 v_1 + k_2 v_2 + \cdots + k_m v_m \tag{A.4}$$

and

$$z - z^* = l_0 v_0 + l_1 v_1 + \cdots + l_{m-1} v_{m-1}, \tag{A.5}$$

respectively.

A. *If $k_i > 0$ for all $i = 1, \dots, m$, then $0 < \Delta_1 f(z - e_1)$, and thus z is not cw-minimal.*

B. *If $k_i < 0$ for all $i = 1, \dots, m$, then $\Delta_1 f(z) < 0$, and z is not cw-minimal.*

C. *If $l_i > 0$ for all $i = 0, \dots, m-1$, then $\Delta_m f(z) < 0$, and z is not cw-minimal.*

D. *If $l_i < 0$ for all $i = 0, \dots, m-1$, then $0 < \Delta_m f(z - e_1)$, and z is not cw-minimal.*

Proof. For statement A we assume, without loss of generality, that $z^* = 0$, and let $z \in \mathbf{Z}^m$ be $z = k_1 v_1 + \cdots + k_m v_m$. From the definition of multimodularity, we get, by taking $v_i = v_0 = (-1, 0, \dots, 0)$:

$$f(u - e_1) - f(u) \geq f(u - e_1 + v_j) - f(u + v_j)$$

for all $u \in \mathbf{Z}^m$ and all $v_j \in \mathcal{V}, v_j \neq v_0$. Since u is arbitrary, we can rewrite this as

$$\Delta_1 f(u) \leq \Delta_1 f(u + v_j), \qquad u \in \mathbf{Z}^m,\ v_j \in \mathcal{V},\ v_j \neq v_0. \tag{A.6}$$

Note that $-e_1 = v_0 = -v_1 - v_2 - \cdots - v_m$, so

$$z - e_1 = (k_1 - 1)v_1 + (k_2 - 1)v_2 + \cdots + (k_m - 1)v_m,$$

where, by assumption, $k_i - 1 \geq 0$ for all i. By repeated application of (A.6), we thus get

$$\Delta_1 f(z - e_1) \geq \Delta_1 f(0) > 0.$$

From $\Delta_1 f(z - e_1) > 0$ follows that z cannot be a cw-minimal point, and this proves A.

For statement B we note that if $z^* = 0$ is a strictly cw-minimal point, then $\Delta_1 f(-e_1) < 0$, and

$$\begin{aligned}
\Delta_1 f(z) &= \Delta_1 f(k_1 v_1 + \cdots + k_m v_m) \\
&= \Delta_1 f(-e_1 + (k_1 + 1)v_1 + \cdots + (k_m + 1)v_m) \\
&\leq \Delta_1 f(-e_1) \\
&< 0,
\end{aligned}$$

so z is not cw-minimal.

For statements C and D, note that $v_m = e_m$, so the proof is analogous to cases A and B by showing that (A.6) now becomes

$$\Delta_m f(u) \geq \Delta_m f(u + v_j), \qquad u \in \mathbf{Z}^m,\ v_j \in \mathcal{V},\ v_j \neq v_m.$$

This concludes the proof. □

To continue with a classification of cw-optimal points, we introduce the following definition of cones and atoms.

DEFINITION A.4. *For $z \in \mathbf{Z}^m$, define the following polyhedral cones:*

$$C_{0+}(z) = \{u \in \mathbf{Z} \mid u = z + k_1 v_1 + \cdots + k_m v_m,\ k_i \in \mathbf{Z},\ k_i > 0\}, \tag{A.7}$$

$$C_{0-}(z) = \{u \in \mathbf{Z} \mid u = z + k_1 v_1 + \cdots + k_m v_m,\ k_i \in \mathbf{Z},\ k_i < 0\}, \tag{A.8}$$

$$C_{m+}(z) = \{u \in \mathbf{Z} \mid u = z + k_0 v_0 + \cdots + k_{m-1} v_{m-1},\ k_i \in \mathbf{Z},\ k_i > 0\}, \tag{A.9}$$

$$C_{m-}(z) = \{u \in \mathbf{Z} \mid u = z + k_0 v_0 + \cdots + k_{m-1} v_{m-1},\ k_i \in \mathbf{Z},\ k_i < 0\}. \tag{A.10}$$

We let $C(z)$ denote the union

$$C(z) = C_{0+}(z) \cup C_{0-}(z) \cup C_{m+}(z) \cup C_{m-}(z). \tag{A.11}$$

From Lemma A.3 we know that if z is a strictly cw-minimal point, then there are no other cw-minimal points in $C(z)$. This means that if we start from a known cw-minimal point z, then we have to search only the complement of $C(z)$ for other possible cw-minimal points. This complement can be characterized by means of a simplicial decomposition of $\mathbf{R}^m$. We now continue with a brief introduction to this decomposition. For a detailed discussion, we refer to Hajek [2].

DEFINITION A.5. *We let Σ denote the set of permutations of $\{0, \ldots, m\}$. Let $\sigma \in \Sigma$ and $z \in \mathbf{Z}^m$. The set $\{u_0, \ldots, u_m\}$ of extreme points of the atom $S(z, \sigma)$ is defined as follows:*

$$\begin{aligned}
u_0 &= z,\\
u_1 &= u_0 + v_{\sigma(1)},\\
u_2 &= u_1 + v_{\sigma(2)},\\
&\ \vdots\\
u_m &= u_{m-1} + v_{\sigma(m)};
\end{aligned}$$

hence $u_0 = u_m + v_{\sigma(0)}$. The atom $S(z, \sigma) \subset \mathbf{R}^m$ is thus the set of convex combinations of $\{u_0, \ldots, u_m\}$: $S(z, \sigma) = \{\sum_{i=0}^m a_i u_i \in \mathbf{Z}^m \mid a_i \in \mathbf{R}_+, \sum_{i=0}^m a_i = 1\}$. We denote $S(z, \sigma) = \langle u_0, \ldots, u_m \rangle$.

Each atom is in fact a simplex, since it contains exactly $m + 1$ extreme points in $\mathbf{Z}^m$. Examples of atoms in two and three dimensions are depicted in Figure A.1. In $\mathbf{R}^2$ the atoms are triangles. In $\mathbf{R}^3$ each atom is bounded by four triangles, and each of the triangles that belong to the same atom share exactly one side.

The atoms allow the following alternative characterization of multimodularity. Every atom S contains exactly $m + 1$ points $\{u_0, \ldots, u_m\}$, so for any function f defined on $\mathbf{Z}^m$, there is a unique affine function $L_S(z)$ that agrees with f on the $m + 1$ extreme points. If f is multimodular, then $L_S(z) \leq f(z)$ for $z \in \mathbf{Z}^m$ (see Hajek [2, Lemma 4.2]). The entire $\mathbf{R}^m$ can be decomposed uniquely into atoms of the form $S(z, \sigma)$, and for a function f defined on $\mathbf{Z}^m$ we can thus uniquely construct a continuous function $\underline{f}$ on $\mathbf{R}^m$ that is piecewise affine on all the atoms $S(z, \sigma)$, $z \in \mathbf{Z}^m$. If f is a multimodular on $\mathbf{Z}^m$, then this $\underline{f}$ is a convex function on $\mathbf{R}^m$.

The next property of multimodular functions will be used a couple of times in this paper, so for this reason we state it here explicitly.

LEMMA A.6. *If f is a multimodular function, then $f(z + ke_i)$ is a convex function of $k \in \mathbf{Z}$ for any $z \in \mathbf{Z}$ and unit vector e_i, $i = 1, \ldots, m$.*

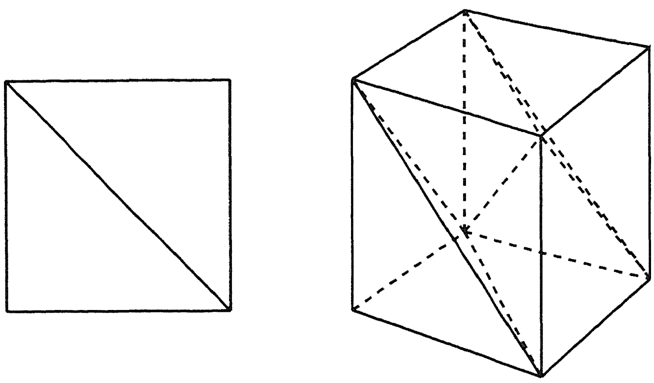

FIG. A.1. *Atoms in two and three dimensions.*

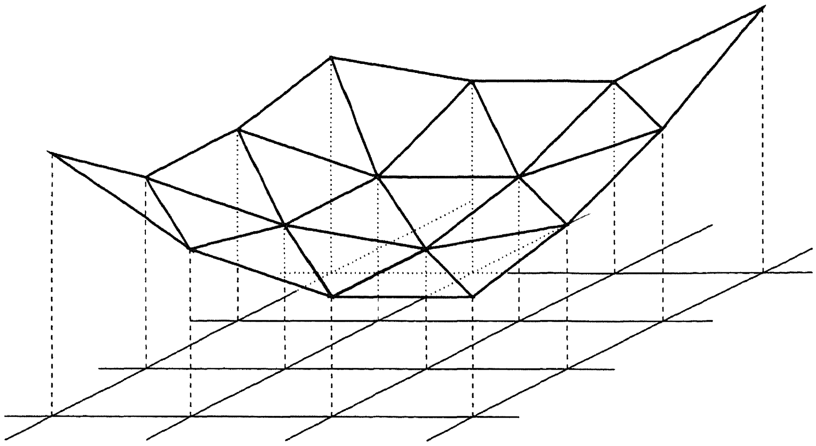

FIG. A.2. *Example of a multimodular function in $\mathbf{Z}^2$.*

Proof. Let $z \in \mathbf{Z}^m$ and e_i be some unit vector. Let S be an atom that contains both z and $z + e_i$. Such an atom exists, since $e_i = v_i + \cdots + v_m$. Using $L_S(z) \le f(z)$ for this atom, we get

$$\Delta_i f(z) = f(z + e_i) - f(z) = L_S(z + e_i) - L_S(z) = L_S(z + 2e_i) - L_S(z + e_i)$$

$$\le f(z + 2e_i) - f(z + e_i) = \Delta_i f(z + e_i).$$

The third equality is due to the fact that L_S is affine. □

In a similar manner, we may conclude that in fact for any vector $v \in \mathcal{V}$ and for any $z \in \mathbf{Z}$, the function $f(z + kv)$ is convex in $k \in \mathbf{Z}$.

For an example of a multimodular function on $\mathbf{Z}^2$, see Figure A.2. The atoms that decompose $\mathbf{R}^2$ were depicted in Figure A.1.

DEFINITION A.7. *For an atom $S(z, \sigma)$, $z \in \mathbf{Z}^m$, $\sigma \in \Sigma$, we define $C_\sigma(z)$ as the polyhedral cone in $\mathbf{Z}^m$:*

$$C_\sigma(z)$$
$$= \Big\{u \in \mathbf{Z}^m \mid u = z + k_1(u_1 - u_0) + \cdots + k_m(u_m - u_0), k_i \in \mathbf{N}, S(z, \sigma) = \langle u_0, \dots, u_m\rangle\Big\}.$$

Define Σ^* *to be the following subset of permutations:*

$$\Sigma^* = \{\sigma \in \Sigma \mid \sigma(0) \neq 0,\ \sigma(0) \neq m,\ \sigma(1) \neq 0,\ \sigma(1) \neq m\},$$

and

$$C_{\Sigma^*}(z) := \bigcup_{\sigma \in \Sigma^*} C_\sigma(z),$$

$$\underline{C}_{\Sigma^*}(z) := \{u \in C_{\Sigma^*}(z) \mid u \neq z + ke_i \text{ for all } k \in \mathbf{Z}, i = 1, \dots, m\}.$$

Finally, we define $P(z)$ *as the plane through* z *that has normal vector* $(1, \dots, 1)$, *i.e.,*

$$P(z) := \{u \in \mathbf{Z}^m | u_1 + u_2 + \cdots + u_m = z_1 + z_2 + \cdots + z_m\}.$$

Note in this definition that for the case of $m = 2$, the set Σ^* is empty, and thus $C_{\Sigma^*}(z)$ and $\underline{C}_{\Sigma^*}(z)$ are also empty.

In the remainder of this section, we shall use these definitions to build a characterization of the set of cw-minimal points. The following lemma gives us the necessary preliminary results.

LEMMA A.8. *For any* $z \in \mathbf{Z}$,

$$C(z) \cap C_{\Sigma^*}(z) = \emptyset, \tag{A.12}$$

$$\mathbf{Z}^m = C(z) \cup C_{\Sigma^*}(z) \cup P(z). \tag{A.13}$$

Proof. First consider (A.12). If $m = 2$, then $C_{\Sigma^*}(z) = \emptyset$ and the result is immediate. For $m > 2$, assume that we have $z \in \mathbf{Z}^m$ and $y \in C_\sigma(z)$ for some $\sigma \in \Sigma^*$. This means that we can write $y - z$ as

$$y - z = k_1(u_1 - u_0) + \cdots + k_m(u_m - u_0)$$

or

$$y - z = k_1 v_{\sigma(1)} + k_2(v_{\sigma(1)} + v_{\sigma(2)}) + \cdots + k_m(v_{\sigma(1)} + v_{\sigma(2)} + \cdots + v_{\sigma(m)}), \tag{A.14}$$

and thus

$$\begin{aligned} y - z &= (k_1 + k_2 + \cdots + k_m)v_{\sigma(1)} \\ &+ (k_2 + k_3 + \cdots + k_m)v_{\sigma(2)} \\ &\vdots \\ &+ k_m v_{\sigma(m)}. \end{aligned} \tag{A.15}$$

Note that all $k_i \geq 0$. Since $\sigma \in \Sigma^*$, we know that $\sigma(j) = 0$ for some $j \neq 0,\ 1$. Using (A.1) and subtracting $(k_j + k_{j+1} + \cdots + k_m)(v_0 + \cdots + v_m)$ from (A.15), we get

$$\begin{aligned} y - z &= (k_1 + \cdots + k_{j-1})v_{\sigma(1)} \\ &+ (k_2 + \cdots + k_{j-1})v_{\sigma(2)} \\ &\vdots \\ &+ k_{j-1}v_{\sigma(j-1)} \\ &- k_j v_{\sigma(j+1)} \\ &- (k_j + k_{j+1})v_{\sigma(j+2)} \\ &\vdots \\ &- (k_j + k_{j+1} + \cdots + k_{m-1})v_{\sigma(m)}. \end{aligned}$$

Recall that $\{v_{\sigma(1)}, \ldots, v_{\sigma(j-1)}, v_{\sigma(j+1)}, \ldots, v_{\sigma(m)}\}$ is a basis for $\mathbf{Z}^m$, so this representation is unique. Since $k_i \geq 0$, we see that y is neither in $C_{0+}(z)$ nor in $C_{0-}(z)$. Analogously, one can prove that $z \notin C_{m+}(z)$ and $z \notin C_{m-}(z)$, and thus $z \notin C(z)$. This proves that $C(z)$ and $C_{\Sigma^*}(z)$ are disjoint.

To prove (A.13), assume that $y \in \mathbf{Z}^m$ and $y \notin C_{\Sigma^*}(z)$. We have to prove that y is in $C(z)$ or $P(z)$. The set $\{S(z, \sigma) \mid z \in \mathbf{Z}^m, \sigma \in \Sigma\}$ forms a partition of $\mathbf{Z}^m$, so there must be a permutation σ of $\{0, \ldots, m\}$, such that $y \in C_\sigma(z)$. Since $y \notin C_{\Sigma^*}(z)$, σ is not in Σ^*, and we must have $\sigma(0) = 0$ or m, or $\sigma(1) = 0$ or m. We shall deal with the case of $\sigma(1) = 0$ first. According to the definition of $C_\sigma(z)$, we can write $y - z$ as

$$y - z = (k_1 + k_2 + \cdots + k_m)v_0 + (k_2 + \cdots + k_m)v_{\sigma(2)} + \cdots + k_m v_{\sigma(m)} \tag{A.16}$$

for some $k_i \geq 0$. We now have to distinguish between the three following cases.

1: $k_1 > 0$. Use (A.1) to show that

$$\begin{aligned} y - z &= -k_1 v_{\sigma(2)} \\ &- (k_1 + k_2)v_{\sigma(3)} \\ &\vdots \\ &- (k_1 + \cdots + k_{m-1})v_{\sigma(m)} \\ &- (k_1 + \cdots + k_m)v_{\sigma(0)}. \end{aligned}$$

Since k_1 is strictly positive, this means that $y \in C_{0-}(z)$.

2: $k_1 = 0$, $\sigma(0) = m$. We first show that $k_m > 0$. Assume that $k_m = 0$. Construct a permutation τ as follows: $\tau(1) = \sigma(2)$, $\tau(2) = \sigma(1) = 0$, $\tau(m) = \sigma(0) = m$, $\tau(0) = \sigma(m)$, and $\tau(i) = \sigma(i)$ for $i = 3, \ldots, m-1$. This makes $\tau \in \Sigma^*$. From (A.16), using $k_1 = k_m = 0$, we get

$$\begin{aligned} y - z &= (k_2 + \cdots + k_m)v_{\tau(1)} \\ &+ (k_2 + \cdots + k_m)v_{\tau(2)} \\ &\vdots \\ &+ k_{m-1}v_{\tau(m-1)}, \end{aligned}$$

which implies that $y \in C_{\Sigma^*}(z)$, and this contradicts $y \notin C_{\Sigma^*}(z)$. We may conclude that the assumption $k_m = 0$ is incorrect, and then it is immediate from (A.16) that $y \in C_{m+}(z)$.

3: $k_1 = 0$ and $\sigma(0) \neq m$. Define the permutation τ by $\tau(1) = \sigma(2)$, $\tau(2) = \sigma(1)$ and $\tau(i) = \sigma(i)$, $i = 3, \ldots, m$. From (A.16), we have

$$\begin{aligned} y - z &= (k_2 + \cdots + k_m)v_{\tau(1)} \\ &+ (k_2 + \cdots + k_m)v_{\tau(2)} \\ &\vdots \\ &+ k_m v_{\tau(m)}, \end{aligned} \tag{A.17}$$

and thus $y \in C_\tau(z)$. Now assume that $\tau(1) \neq m$, then $\tau \in \Sigma^*$, and thus $y \in C_{\Sigma^*}(z)$, which contradicts the assumption that $y \notin C_{\Sigma^*}(z)$. We conclude that $\tau(1) = m$. Since $v_{\tau(1)} = v_m$ and $v_{\tau(2)} = v_0$, we may conclude from (A.17) that the coefficients of $y - z$ sum up to zero, and thus $y \in P(z)$.

The case where $\sigma(1) = m$ is proven analogously, with the roles of v_0 and v_m interchanged. The case where $\sigma(0) = 0$ (or $\sigma(0) = m$) proceeds as follows. Again, use (A.1) by substituting

$$v_{\sigma(1)} + v_{\sigma(2)} + \cdots + v_{\sigma(j)} = -v_{\sigma(j+1)} - \cdots - v_{\sigma(m)} - v_{\sigma(0)} \tag{A.18}$$

into each line of (A.14) to get

$$
\begin{aligned}
y - z \quad = \quad & -(k_1 + \cdots + k_m)v_{\sigma(0)} \\
& -(k_1 + \cdots + k_{m-1})v_{\sigma(m)} \\
& -(k_1 + \cdots + k_{m-2})v_{\sigma(m-1)} \\
& \vdots \\
& -k_1 v_{\sigma(2)}.
\end{aligned} \tag{A.19}
$$

This brings $y - z$ in a form similar to (A.15), with the exception that now all the coefficients of the v_is become negative. The proof concludes analogously to the case where $\sigma(1) = 0$. □

Note that the three sets $C(z)$, $C_{\Sigma^*}(z)$, and $P(z)$ are not mutually disjoint. $C(z)$ and $P(z)$ are disjoint, but $C_{\Sigma^*}(z)$ and $P(z)$ have a nonempty intersection. The immediate consequence of Lemma A.8 is summarized in the following theorem, which is the main result of this section.

THEOREM A.9. *If $z \in \mathbf{Z}^m$ is a strictly cw-minimal point of a multimodular function f, then there can be other cw-minimal points only in $\underline{C}_{\Sigma^*}(z)$ or in the plane $P(z)$.*

Proof. It is immediate from Lemma A.8 that the only other cw-minimal points must lie in $C_{\Sigma^*}(z)$ or in $P(z)$. Actually, we do not need to include the entire set $C_{\Sigma^*}(z)$ as a possibility for other coordinatewise minima. It may contain an axis of the form $\{z + ke_i \mid k \in \mathbf{Z}\}$ for some unit vector e_i, $i = 1, \ldots, m$. Since z is strictly cw-minimal and f is convex along this axis (see Lemma A.6), it is immediately obvious that there cannot be other coordinatewise minima on this axis. □

Theorem A.9 not only gives a characterization of the set of cw-minimal points, but it also enables us to search for the minimum in an efficient manner. In the two-dimensional case, for instance, the theorem means that if $z \in \mathbf{Z}^2$ is strictly cw-minimal, then the possible other coordinatewise minima must lie on the line $\{(z_1 + k, z_2 - k) \mid k \in \mathbf{Z}\}$. For an indication of the implications of Theorem A.9, take a look at Figure A.3. We assume that 0 (the center of the cube) is strictly cw-minimal. The points indicated with a bold filled circle are the points on the unit cube that are in both $\underline{C}_{\Sigma^*}(0)$ and in $P(0)$. The bold open circles are the points of $\underline{C}_{\Sigma^*}(0)$ that are not in $P(0)$. The bold open diamonds are the points $P(0)$ that are not in $\underline{C}_{\Sigma^*}(0)$.

We conclude this appendix with the proof of Lemma 3.4. It summarizes the results of Theorem A.9 for multimodular functions defined on $\mathbf{Z}^2$.

LEMMA A.10. *If $g : \mathbf{Z}^2 \to \mathbf{R}$ is a multimodular function and $y = (y_1, y_2)$ and $z = (z_1, z_2)$ are two distinct strictly cw-minimal points of g where $z_1 \geq y_1$, then*

1. *$(z_1, z_2) = (y_1 + B, y_2 - B)$ for some $B > 0$;*
2. *for all b, $0 < b < B$, $(y_1 + b, y_2 - b)$ is also cw-minimal;*
3. *if $g(z) > g(y)$, then the minimum of g cannot be in the set $\{(y_1 + b, y_2 - b) | b \geq B\}$;*
4. *if both y and $z = (y_1 + B, y_2 - B)$ for some $B \geq 0$ are minimal, then $(y_1 + b, y_2 - b)$ is minimal for all $0 \leq b \leq B$;*
5. *if both $g(y_1 + 1, y_1 - 1) \geq g(y_1, y_2)$ and $g(y_1 - 1, y_1 + 1) \geq g(y_1, y_2)$, then (y_1, y_2) is minimal; if the inequalities are strict, then the minimum is also unique.*

Proof. The proof follows immediately from Theorem A.9. Note that if $m = 2$, then there exist no permutations σ of $\{0, 1, 2\}$ with both $\sigma(0) \neq 0, 1$ and $\sigma(1) \neq 0, 1$, so $C_{\Sigma^*}(z) = \emptyset$ for all $z \in \mathbf{Z}^2$. This means that if z is a strictly cw-minimal point, then

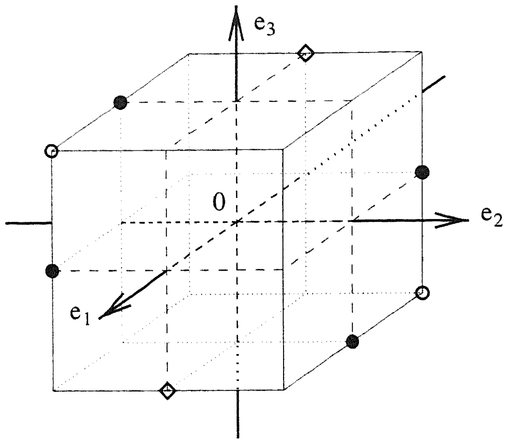

FIG. A.3. *The cone* $C_{\Sigma^*}(0)$ *and the plane* $P(0)$.

the only other cw-minimal points must lie in $P(z) = \{(z_1 + b, z_2 - b) \mid b \in \mathbf{Z}\}$. This proves 1.

To prove 2, note that y strictly cw-minimal implies that $\Delta_1 g(y) > 0$. Since g is multimodular, this implies that $\Delta_1 g(y_1 + b, y_2 - b) > 0$ for $b \geq 0$. In the same manner, $\Delta_1 g(z_1 - 1, z_2) < 0$, and by multimodularity $\Delta_1 g(z_1 - 1 - b, z_2 + b) < 0$ for all $b > 0$. Analogously, one can show that $\Delta_2 g(z_1 - b, z_2 + b) > 0$ and $\Delta_2 g(y_1 + b, y_2 - 1 - b) < 0$ for all $b \geq 0$, and these equalities combined prove 2.

For 3, 4, and 5, note that $f(z) := g(y_1 + z, y_2 - z)$ is a convex function of $z \in \mathbf{Z}$ by the remark below Lemma A.6. □

REFERENCES

[1] M. BARTROLI AND S. STIDHAM, JR., *Multimodular Triangulations*, Tech. report 88-07, Dept. of Oper. Res., University of North Carolina, Chapel Hill, NC, 1988.

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

[3] Y. HO AND K. CHU, *Information structure in dynamic multi-person control problems*, Automatica J. IFAC, 10 (1974), pp. 341–351.

[4] Y.-C. HO, *Team decision theory and information structures*, Proc. IEEE, 68 (1980), pp. 644–654.

[5] 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.

[6] J. MARSCHAK, *Elements for a theory of teams*, Management Sci., 1 (1955), pp. 127–137.

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

[8] M. SHAKED AND J. SHANTHIKUMAR, *Stochastic Orders and their Applications*, Academic Press, San Diego, 1994.

[9] D. TENEKETZIS AND P. VARAIYA, *The decentralized quickest detection problem*, IEEE Trans. Automat. Control, AC-29 (1984), pp. 641–644.

[10] Y.-T. WANG AND R. MORRIS, *Load sharing in distributed systems*, IEEE Trans. Comput., C-34 (1985), pp. 204–217.

[11] H.S. WITSENHAUSEN, *On information structures, feedback and causality*, SIAM J. Control, 9 (1971), pp. 149–160.

[12] H. WITSENHAUSEN, *Separation of estimation and control for discrete time systems*, Proc. IEEE, 59 (1971), pp. 1557–1566.