Decentralized decision-making in a large team with local information

Paat Rusmevichientong [a] and Benjamin Van Roy [b,*]

[a] *Stanford University, Stanford, CA 94305, USA*
[b] *Management Science and Engineering, Terman Engineering Center 427, Stanford University, Stanford, CA 94305-4023, USA*

Abstract

We study a problem involving a team of agents each associated with a node in a chain. Each agent makes a decision that influences only his own cost and those of adjacent agents. Prior to making his decision, each agent observes only the cost structure associated with nodes that can be reached by traversing no more than r arcs. Decisions are selected without any coordination, with the common objective of minimizing average cost among agents. We consider such decisions decentralized since agents act based on different information. Cost incurred by an optimal centralized strategy, in which a single decision-maker has access to all information and dictates all decisions, is employed as a performance benchmark. We show that, to maintain a certain level of performance relative to optimal centralized strategies, decentralized deterministic strategies require r to be proportional to the number of agents. This means that the amount of information accessible to any agent should be proportional to the total number of agents. Stochastic strategies, on the other hand, decentralize more gracefully—the amount of information required by each agent is independent of the total number of agents.
© 2003 Elsevier Science (USA). All rights reserved.

---

* Corresponding author.
  E-mail addresses: paatrus@stanford.edu (P. Rusmevichientong), bvr@stanford.edu (B. Van Roy).

Nomenclature

B	upper bound on cost functions;	$\widetilde{\Phi}^{\delta}_{d,\mathcal{U}}$	set of δ-stochastic strategies;
d	number of nodes in the chain;	$\Phi_{d,\mathcal{U},r}$	set of decentralized deterministic strategies of radius r;
δ	probability that a decision is chosen randomly under δ-stochastic strategies;	$\widetilde{\Phi}_{d,\mathcal{U},r}$	set of decentralized stochastic strategies of radius r;
$E[\ ]$	expectation of a random variable;	ϕ	strategy;
f	overall cost function;	ϕ^*	optimal (deterministic) strategy;
f_i	cost function at node i;	ϕ^{δ}	optimal δ-stochastic strategy;
f_i^r	vector of cost functions that are within the neighborhood of radius r around f_i;	$\phi^{\delta,r}$	decentralized stochastic strategy of radius r;
J_i	cost-to-go function at node i under the optimal policy;	$\bar{\phi}^{\delta,r}$	r-lookahead strategy;
$\tilde{J}_i$	approximation to J_i under the decentralized stochastic strategy;	$\phi_i(f)$	decision at node i for the problem instance f under the strategy ϕ ($\phi_i^*(f)$, $\phi_i^{\delta}(f)$, $\phi_i^{\delta,r}(f)$, and $\bar{\phi}_i^{\delta,r}(f)$ are defined analogously);
$\tilde{J}_{i,l}$	cost-to-go function generated during the computation of $\tilde{J}_i$;	r	radius of decentralization;
M^{δ}	set of functions that define δ-stochastic strategies;	T^i	dynamic programming operator at node i;
$\Pr\{\ \}$	probability of an event;	$T^i_{\pi_i}$	dynamic programming operator at node i under the mapping π_i;
π_i	function that defines the stochastic decision at node $i+1$ under the optimal δ-stochastic strategy;	u_i	decision at node i;
		u_i^*	optimal decision at node i;
		$\mathcal{U}$	set of available decisions at each node;
$\tilde{\pi}_i$	function that defines the decision at node $i+1$ under the decentralized stochastic strategy;	$\mathcal{U}^d$	set of available decisions for each problem instance;
		W_i	random variable that is uniformly distributed in $[0,1]$;
$\Psi_{d,\mathcal{U},B}$	class of problem instances;	$(\Omega, \mathcal{F}, \mathcal{P})$	probability triple;
$\Phi_{d,\mathcal{U}}$	set of deterministic strategies;	$\|\cdot\|_s$	span-seminorm;
$\widetilde{\Phi}_{d,\mathcal{U}}$	set of stochastic strategies;	$\|\cdot\|_{\infty}$	maximum norm;
		$\underline{\underline{\mathcal{D}}}$	equality in distribution.

1. Introduction

For several decades, the subject of decentralized decision-making has intrigued economists, operations researchers, and engineers. Today, interest in the topic is amplified by its necessity. Centralized decisions are not an option in the face of magnified complexity and miniaturized time scales posed by modern organizations and engineering systems. The

Internet and social networks with their ever increasing size and interconnectivity pose two popular examples.

Team theory—as introduced by Marschak (1955) and Marschak and Radner (1972)—presents a general framework for decentralized decision-making that accounts for limited exchange of information. A team is defined as a collection of agents, each of whom makes a decision based on information available to him—which may differ from that available to other agents. Agents share as their goal minimization of a common cost function.

Unfortunately, even for simple problems where optimal centralized decisions are readily derived, computation of optimal decentralized decisions can be intractable, as formally established by the work of Tsitsiklis and Athans (1985) and Papadimitriou and Tsitsiklis (1986). This impediment motivates more restrictive formulations with reduced computational requirements.

Two early papers by Radner offer interesting cases where team decision problems simplify. The first (Radner, 1959) shows that optimal decentralized decisions can be generated via linear programming if cost is a known convex polyhedral function of decisions. This procedure is efficient when the number of linear pieces making up the cost function is reasonably small. A second paper (Radner, 1962) treats the case of a quadratic cost function. Coefficients for linear terms of the cost function together with all observations are drawn from a Gaussian prior. The main result establishes that each agent's optimal decision is a linear function of his observations. Extensions of this result to settings where agents possess differing Gaussian priors have been explored by Başar (1985).

Though the Gaussian model captures a broad array of relevant phenomena, complex organizations and engineering systems of contemporary interest generally fall beyond its scope. Such systems do nevertheless exhibit special structure. However, instead of taking the form of Gaussian distributions and linear relationships, the special structure arises from limited interdependencies among agents. Each agent shares information with and influences costs of "nearby" agents but not agents that are "far away." As suggested by the colloquial terms in the preceding sentence, influence between agents and commonality of information is governed by proximity. In a physical system this may correspond to physical distance, while in a social network proximity may relate to the "degrees of separation" between two individuals.

Regardless of the context, proximity between agents can be represented abstractly in terms of an undirected graph $\mathcal{G}$ with vertices $\mathcal{V} = \{1, \dots, d\}$ and edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. The distance $\rho(i, j)$ between two nodes $i, j \in \mathcal{V}$ is defined to be the minimal number of arcs needed to form a path connecting i and j. Each node $i \in \mathcal{V}$ corresponds to an agent and therefore a decision u_i. We take this graph to identify two structural characteristics associated with a given decision problem. First, cost decomposes according to

$$f(u) = \sum_{(i,j) \in \mathcal{E}} f_{ij}(u_i, u_j).$$

Second, each ith agent observes a cost function f_{jk} if and only if $\rho(i, j) \leqslant r$ and $\rho(i, k) \leqslant r$. Hence, the graph limits information available to agents as well as the scope of their influence.

Graphs are capable of capturing structure inherent in many complex systems, and exploring how this structure may reduce requirements of information and computation poses a promising research direction. Indeed, graphical models have been exploited to this end in nonserial dynamic programming (Bertelé and Brioschi, 1972) and probabilistic inference (see, e.g., (Jensen, 1996; Pearl, 1988)). In this paper, as a starting point to exploring benefits enabled by graphical structure in decentralized-decision making, we study problems for which interdependencies are represented by the simplest form of graph—a chain.

In decentralized decision-making, there is a dividing line between situations in which agents can and in which they cannot *coordinate*. We associate the term *coordination* with an ability to select decision strategies in a centralized manner, though decisions generated by the strategies upon observation of problem data may entail only decentralized processing. In other words, the agents are allowed to communicate, even to collude, with other agents when they form a decision strategy; however, the decision generated from such strategy may only use partial information. The focus of this paper is on the case where coordination is prohibited. This situation is of interest when one considers the management of large networks. In such contexts, coordination may incur substantial overhead in communication costs. Practical strategies typically do not involve coordination. Instead, each agent makes his decision based on observed problem data, without regard to strategies employed by other agents. We model the absence of coordination by requiring agents to share a common strategy. In other words, if two agents face the same problem data relative to their respective locations in the chain, they make the same decision.

Another dividing line separates deterministic from stochastic strategies. We will consider both and demonstrate significant advantages associated with the latter in the face of decentralization. It is worth commenting on the form of stochastic strategies we will consider. One approach to defining a stochastic strategy involves allowing each agent to generate random numbers that influence only his own decision. We consider a broader class of stochastic strategies in which different agents can make use of common random numbers. In particular, random numbers—along with other problem data—are associated with particular nodes in the chain, and any agent observing data at a particular node can also observe random numbers associated with that node. One may think of these random numbers as information that is available at a node but irrelevant to the decision problem at hand.

In the next section, we formulate the main problem considered in the paper—that of decentralized decision-making in a chain without coordination. In this context, our results capture the following facts:

(1) To maintain a certain level of performance relative to optimal centralized strategies, decentralized deterministic strategies require r to be proportional to the number of agents. In particular, the worst-case performance difference is bounded below by $c(1/4 - r/d)$, where d is the length of the chain, and c is some positive constant.

(2) Stochastic strategies decentralize more gracefully. In particular, for any $p < 1$, optimal decentralized stochastic strategies offer performance differing from that of optimal centralized strategies by no more than c/r^p, for some positive scalar c, which is independent of d.

These results constitute the main contributions of the paper. They provide new insights into the role of information in the decision-making process in a team environment. Our results are different from other works in the literature in our characterization and use of proximity between agents in a team. Most of the prior research assume that an information structure is specified *a priori* and focus primarily on finding an optimal decentralized controller.

The proof of the positive result involves showing that a particular decentralized stochastic strategy, which is constructed in Section 3, delivers promised performance. The proof is provided in Section 4. The strategy amounts to a computational procedure that is fortuitously efficient, so as a by-product of our analysis, it is established that the performance identified by the theorem is delivered by an efficient algorithm.

In Section 5, we consider the case where coordination is allowed. In this context, it is easy to show that the performance of optimal decentralized deterministic strategies differ from that of an optimal centralized strategy by no more than c/r, for some positive scalar c, independent of problem size. This performance appears quite positive in light of the negative result concerning decentralized decision-making in the absence of coordination. However, as we will further discuss in Section 5, coordination is sometimes impractical or impossible. In such cases, randomization enables comparable performance in the absence of coordination.

It is worth noting that our results relate to other interesting and quite separate lines of research. One involves the study of games in which a team of decision-makers who are unable to coordinate face a single adversary (von Stengel and Koller, 1997). In our context, the adversary selects cost-structure, while constrained by the topology of a graph. Our formulation brings a new sort of special structure to such problems, and our results offer means for a team to devise effective counter-measures against the adversary.

Another related line of research, involving minimum cost network-flow problems, studies how changes in costs and constraints in one part of a network impact decisions in another (Granot and Veinott Jr., 1985). The main result here is that "influence" diminishes as "distance" increases. This is the same intuitive notion that underlies our use of proximity in graphs as a basis for decentralization.

Finally, our results relate to the work of Billard and Pasquale (1995) on resource sharing in distributed computer systems, where multiple users attempt to access a fixed set of resources. The authors show that reasonable system performance can be achieved with limited communication among users. They showed that there is a threshold for the level of communication below which decision quality deteriorates. Our work can be viewed as an extension of this idea, and it is distinguished by its use of special structure based on proximity to characterize communication constraints.

2. Cost structure and decision strategies

We consider a problem entailing the selection of a decision from a finite set $\mathcal{U}^d = \mathcal{U} \times \cdots \times \mathcal{U}$. An element $u \in \mathcal{U}^d$ can be thought of as a vector $u = (u_1, \ldots, u_d)$ with

$u_i \in \mathcal{U}$ for each i. The objective is to minimize a cost function $f : \mathcal{U}^d \mapsto \Re$ that decomposes according to

$$f(u) = \frac{1}{d} \sum_{i=1}^{d-1} f_i(u_i, u_{i+1}),$$

for some sequence of functions $f_1, \ldots, f_{d-1}$. An optimal decision $u \in \mathcal{U}^d$ is one that attains the minimum of f in $\mathcal{U}^d$.

Each *instance* of the aforementioned optimization problem is characterized by a triplet $P = (d, \mathcal{U}, f)$ where d is the number of decision variables, $\mathcal{U}$ is a set whose Cartesian product $\mathcal{U}^d$ forms the decision space, and $f = (f_1, \ldots, f_{d-1})$ is a sequence of $d-1$ functions, with each f_i mapping $\mathcal{U} \times \mathcal{U}$ to $\Re$. We will denote by $\Psi_{d,\mathcal{U},B}$ a *class* of problem instances, defined by

$$\Psi_{d,\mathcal{U},B} = \left\{ (d, \mathcal{U}, f) \mid \|f_i\|_\infty \leqslant B \ \forall i \right\}.$$

Note that each element $(d, \mathcal{U}, f)$ of $\Psi_{d,\mathcal{U},B}$ is distinguished only by f, and therefore we will often refer to elements of $\Psi_{d,\mathcal{U},B}$ simply in terms of f. For example, we will use a statement $f \in \Psi_{d,\mathcal{U},B}$ as shorthand for $(d, \mathcal{U}, f) \in \Psi_{d,\mathcal{U},B}$. Abusing notation somewhat, we will also use f as though it were a function defined by

$$f(u) = \frac{1}{d} \sum_{i=1}^{d-1} f_i(u_i, u_{i+1}).$$

A *strategy* is a mapping $\phi : \Psi_{d,\mathcal{U},B} \to \mathcal{U}^d$ that, for each problem instance $f \in \Psi_{d,\mathcal{U},B}$, generates a decision $\phi(f) \in \mathcal{U}^d$. Let $\Phi_{d,\mathcal{U}}$ denotes the collection of strategies available for optimization problems over $\mathcal{U}^d$. A strategy $\phi^* : \Psi_{d,\mathcal{U},B} \to \mathcal{U}^d$ associated with the class of problems $\Psi_{d,\mathcal{U},B}$ is optimal if

$$f\big(\phi^*(f)\big) = \min_{u \in U} f(u),$$

for every $f \in \Psi_{d,\mathcal{U},B}$.

2.1. Decentralized strategies

In general, selecting a component u_i of a decision that minimizes the cost function f entails consideration of every component function f_i. In other words, for each optimal strategy $\phi^* \in \Phi_{d,\mathcal{U}}$, $\phi_i^*(f)$ may depend on $f_1, f_2, \ldots, f_{d-1}$—that is, optimal strategies are *centralized*. We consider the possibility of *decentralizing* the decision-making process and the consequences of decentralization on cost.

In a decentralized strategy each decision is determined from cost functions within a *neighborhood*. Let f_i^r denote a vector of cost functions that are within the *neighborhood of radius r* around f_i, i.e.,

$$f_i^r = (f_{(i-r)\vee 1}, \ldots, f_i, \ldots, f_{(i+r)\wedge(d-1)}).$$

We define the space $\Phi_{d,\mathcal{U},r}$ of *decentralized strategies of radius r* by

$$\Phi_{d,\mathcal{U},r} = \left\{ \phi \in \Phi_{d,\mathcal{U}} \mid \forall f \in \Psi_{d,\mathcal{U},B}, \phi_i(f) = \phi_i\left(f_i^r\right), \right.$$
$$\left. \text{and } \phi_i(f) = \phi_j(f) \text{ if } f_i^r = f_j^r \right\},$$

where the condition $\phi_i(f) = \phi_i(f_i^r)$ means that the decision $\phi_i(f)$ only depends on f_i^r. We require that $\phi_i(f) = \phi_j(f)$ whenever $f_i^r = f_j^r$. This condition means that the decision at each node depends only on the cost functions in its neighborhood, not on its relative position in the chain. In a large network where it is impractical to keep track of the network configuration, this requirement is quite reasonable since it is not practical to have a decision strategy that depends on the location of each node in the network. Moreover, this assumption yields a decision strategy that can be implemented efficiently since the decision at each node only depends on the cost functions within its vicinity. We should note that this requirement effectively precludes coordination among nodes—an issue that will be further discussed in Section 5.

Since decentralized strategies base decisions on partial cost information, a new optimality criterion that takes into account uncertainty of the unavailable information is required. We adopt a worst-case objective:

$$\max_{f \in \Psi_{d,\mathcal{U},B}} f\left(\phi(f)\right).$$

The optimization problem faced is then

$$\min_{\phi \in \Phi_{d,\mathcal{U},r}} \max_{f \in \Psi_{d,\mathcal{U},B}} f\left(\phi(f)\right).$$

For decentralized strategies to be effective in large problems, one would hope that a small loss of optimality is incurred when the "radius of observation" r is sufficiently large and that the required radius is independent of the problem size d. Unfortunately, this is not the case. As captured by the following theorem, which is proved in Section 4, the radius required for ϵ-optimality grows linearly with problem size. In the theorem, we assume that $d \geqslant 4$ and $4r < d$. In general, we are often interested in the case when $r \ll d$; thus, this assumption does not impose any practical constraint.

Theorem 1. *For any $|\mathcal{U}| \geqslant 2$ and $B > 0$ there exists a positive scalar c such that*

$$\min_{\phi \in \Phi_{d,\mathcal{U},r}} \max_{f \in \Psi_{d,\mathcal{U},B}} \left(f\left(\phi(f)\right) - f\left(\phi^*(f)\right)\right) \geqslant c\left(\frac{1}{4} - \frac{r}{d}\right),$$

for all $d \geqslant 4$ and $4r < d$.

It is worth mentioning that our performance measure is quite conservative. It views the problem instance confronted by a decision-maker as being chosen by an *adversary* that knows exactly the decision strategy to be deployed. A less pessimistic formulation might construct a probability distribution over problem instances and require optimization of an expectation:

$$\min_{\phi \in \Phi_{d,\mathcal{U},r}} E\left[f\left(\phi(f)\right)\right],$$

where the expectation is taken over a random cost function f drawn from $\Psi_{d,\mathcal{U},B}$. A reasonable conjecture would be that—for certain classes of distributions—the loss of optimality associated with decentralization is attenuated relative to the worst-case formulation. This may be true. However, rather than characterizing classes of distributions for which decentralized decision-making is effective, we take a direction that offers stronger justification for decentralization. In particular, we leave the choice of problem instance in the hands of an adversary with complete knowledge of the decision-maker's strategy, but generalize the formulation to allow use of *stochastic* strategies in which decisions are influenced by random events predictable neither to the adversary nor the decision-maker. It turns out that—even in the adversarial setting—stochastic strategies decentralize gracefully, as will be discussed in the following section.

2.2. Stochastic strategies

A *stochastic decision* $U = (U_1, U_2, \ldots, U_d)$ is a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$, taking on values in $\mathcal{U}^d$. For the purpose of evaluating stochastic decisions, we generalize our objective to one involving an expectation $E[f(U)]$. An optimal stochastic decision minimizes this expectation. A *stochastic strategy* ϕ is a mapping that, for each problem instance $f \in \Psi_{d,\mathcal{U},B}$ and random sample $\omega \in \Omega$, generates a stochastic decision $\phi(f, \omega) = (\phi_1(f, \omega), \ldots, \phi_d(f, \omega))$. We will often suppress notation indicating the dependence on ω and view each decision $\phi_i(f) = \phi_i(f, \omega)$ as a random variable. We denote the set of stochastic strategies by $\widetilde{\Phi}_{d,\mathcal{U}}$, and a stochastic strategy $\phi^* \in \widetilde{\Phi}_{d,\mathcal{U}}$ is optimal if

$$E\big[f\big(\phi^*(f)\big)\big] = \min_{\phi \in \widetilde{\Phi}_{d,\mathcal{U}}} E\big[f\big(\phi(f)\big)\big],$$

for all $f \in \Psi_{d,\mathcal{U},B}$.

The space of *decentralized stochastic strategies of radius r*, denoted by $\widetilde{\Phi}_{d,\mathcal{U},r}$, is defined by

$$\widetilde{\Phi}_{d,\mathcal{U},r} = \big\{ \phi \in \widetilde{\Phi}_{d,\mathcal{U}} \mid \forall f \in \Psi_{d,\mathcal{U},B}, \ \phi_i(f) = \phi_i\big(f_i^r\big),$$
$$\phi_j(f) \perp \phi_k(f) \ \forall \, |j - k| > 2r,$$
$$\text{and } \phi_i(f) \stackrel{\mathcal{D}}{=} \phi_j(f) \text{ if } f_i^r = f_j^r \big\},$$

where $\phi_j(f) \perp \phi_k(f)$ denotes independence between random variables $\phi_j(f)$ and $\phi_k(f)$. This condition precludes the possibility that agents further than distance $2r$ apart make use of common information. The condition $\phi_i(f) = \phi_i(f_i^r)$ requires that for each ω, $\phi_i(f, \omega)$ only depends on f_i^r. Finally, the condition $\phi_i(f) \stackrel{\mathcal{D}}{=} \phi_j(f)$ means that the random variables $\phi_i(f)$ and $\phi_j(f)$ have the same distribution whenever $f_i^r = f_j^r$. Note that this definition generalizes that employed for decentralized deterministic strategies—the deterministic counterpart can be viewed as a special case in which $|\Omega| = 1$.

We set as our objective minimization of a worst-case criterion:

$$\max_{f \in \Psi_{d,\mathcal{U},B}} E\big[f\big(\phi(f)\big)\big].$$

The optimization problem is therefore

$$\min_{\phi \in \widetilde{\Phi}_{d,\mathcal{U},r}} \max_{f \in \Psi_{d,\mathcal{U},B}} E\big[f\big(\phi(f)\big)\big].$$

It turns out that stochastic strategies decentralize more gracefully than their deterministic counterparts. The following theorem—which is the main result of this paper—characterizes the rate at which performance converges to that of an optimal centralized strategy:

Theorem 2. *For any $p < 1$, B, and $\mathcal{U}$, there exists a scalar c such that*

$$\min_{\phi \in \widetilde{\Phi}_{d,\mathcal{U},r}} \max_{f \in \Psi_{d,\mathcal{U},B}} \big(E\big[f\big(\phi(f)\big)\big] - f\big(\phi^*(f)\big)\big) \leqslant \frac{c}{r^p},$$

for all d and r.

One interesting implication of this theorem is that the radius of observation r required for ϵ-optimality no longer depends on the problem size d. We describe in the next section a decentralized stochastic strategy that achieves the level of performance promised by Theorem 2. This strategy is used as the basis for a constructive proof of the theorem, which is provided in Section 4.

3. A decentralized stochastic strategy

Rather than describing at once all characteristics of our proposed decentralized stochastic strategy, let us present and motivate its features piecemeal via a sequence of simpler strategies. In particular, we define in the following subsections several related strategies:

(1) a centralized deterministic strategy;
(2) a centralized stochastic strategy;
(3) and finally, a decentralized stochastic strategy.

3.1. A centralized deterministic strategy

Recall that a decision $u = (u_1, \ldots, u_d)$ is optimal if it attains the minimum in

$$\min_{u \in \mathcal{U}^d} \frac{1}{d} \sum_{i=1}^{d-1} f_i(u_i, u_{i+1}).$$

The additive structure of our cost function enables dynamic programming as an effective approach to computing optimal decisions. In particular, optimal decisions $u_1, \ldots, u_d$ can be generated based on cost-to-go functions $J_0, \ldots, J_d$, each mapping $\mathcal{U}$ to $\Re$. These functions are computed via a dynamic programming recursion:

$$J_d(u_d) = 0, \quad \forall u_d \in \mathcal{U},$$

$$J_i(u_i) = \min_{u_{i+1} \in \mathcal{U}} \left(\frac{1}{d} f_i(u_i, u_{i+1}) + J_{i+1}(u_{i+1}) \right), \quad \forall u_i \in \mathcal{U}, \ i = 0, \dots, d-1,$$

where $f_0(u_0, u_1) = 0$. (We have introduced here an inconsequential decision u_0 and cost function f_0 solely for notational convenience.) Optimal decisions can then be generated sequentially by letting u_0 be any element of $\mathcal{U}$ and, for each $i = 0, \dots, d-1$, letting u_{i+1} be a decision attaining the minimum in

$$\min_{u_{i+1} \in \mathcal{U}} \left(\frac{1}{d} f_i(u_i, u_{i+1}) + J_{i+1}(u_{i+1}) \right).$$

For concreteness, let us denote by ι an arbitrary fixed element of $\mathcal{U}$, and let $u_0 = \iota$. Let us define $\pi_i : \mathcal{U} \mapsto \mathcal{U}$ as a mapping from a decision u_i to one that attains the minimum in the above expression. Then, the equation

$$u_{i+1} = \pi_i(u_i), \quad \forall i = 0, \dots, d-1,$$

together with the initial condition $u_0 = \iota$ defines a sequence of optimal decisions.

We have provided a method—based on dynamic programming—for computing optimal decisions given a problem instance. It is straightforward to translate this into a strategy. In particular, an optimal centralized deterministic strategy ϕ^* can be defined as one that generates decisions according to the method we have provided. Note that this strategy is *not* decentralized. Decisions are generated sequentially, so each ith decision u_i is influenced by "upstream" decisions $u_{i-1}, u_{i-2}, \dots, u_1$. Furthermore, the selection of each ith decision u_i relies on the cost-to-go function J_i, which is computed based on "downstream" cost functions $f_i, f_{i+1}, \dots, f_{d-1}$. In the coming subsections, we develop a decentralized stochastic variant of this strategy.

3.2. A centralized stochastic strategy

The deterministic strategy introduced in the previous section generates decisions iteratively according to

$$u_{i+1} = \pi_i(u_i), \quad \forall i = 0, \dots, d-1.$$

One mechanism for introducing randomness into the decision process involves redefining each π_i to be a function of two variables: a decision $\mathcal{U}$ and a random number in $[0, 1]$. In particular, we will consider a stochastic decision $U = (U_1, \dots, U_d)$ generated by letting $U_0 = \iota$ and

$$U_{i+1} = \pi_i(U_i, W_i), \quad \forall i = 0, \dots, d-1, \tag{1}$$

for some sequence of functions $\pi_0, \dots, \pi_{d-1}$, where $W_0, \dots, W_{d-1}$ are independent random variables, each uniformly distributed in $[0, 1]$.

We will consider only functions π_i from a special class. For each $\delta \in (0, 1)$, we define a set M^δ to consist of functions $\pi : \mathcal{U} \times [0, 1] \mapsto \mathcal{U}$ that take on the form

$$\pi(U, W) = \begin{cases} \bar{\pi}(U) & \text{if } 0 \leqslant W < 1 - \delta, \\ \text{the } k\text{th element of } \mathcal{U} & \text{if } 1 - \delta + \dfrac{(k-1)\delta}{|\mathcal{U}|} \leqslant W < 1 - \delta + \dfrac{k\delta}{|\mathcal{U}|}, \end{cases}$$

for some function $\bar{\pi} : \mathcal{U} \mapsto \mathcal{U}$, where the set $\mathcal{U}$ is assumed to be ordered in some predetermined way, and $\pi(U, 1)$ is defined to be the $|\mathcal{U}|$th element of $\mathcal{U}$. Note that if W is uniformly distributed in $[0, 1]$, a function $\pi \in M^\delta$ assigns a (dominant) probability $1 - \delta$ to a selection of a distinguished decision $\bar{\pi}(U)$. The remaining probability δ is distributed evenly among all decisions. Hence, the decision follows a deterministic strategy with probability $1 - \delta$ and is chosen in a completely random way with probability δ.

A δ-stochastic strategy ϕ is a mapping that, for each problem instance $f \in \Psi_{d,\mathcal{U},B}$, generates a decision $\phi(f)$ according to Eq. (1) with each π_i being an element of M^δ. The class of δ-stochastic strategies will be denoted by $\widetilde{\Phi}^\delta_{d,\mathcal{U}}$. A strategy $\phi^\delta \in \widetilde{\Phi}^\delta_{d,\mathcal{U}}$ is an optimal δ-stochastic strategy if it attains the minimum in

$$\min_{\phi \in \widetilde{\Phi}^\delta_{d,U}} \max_{f \in \Psi_{d,\mathcal{U},B}} E\big[f\big(\phi(f)\big)\big].$$

We will establish in Section 4.2 that the loss in performance incurred when using an optimal δ-stochastic strategy ϕ^δ in place of an optimal deterministic strategy ϕ^* grows at most linearly in δ.

Analogous to the deterministic case, given a problem instance $f \in \Psi_{d,U,B}$, an optimal δ-stochastic decision $\phi^\delta(f)$ can be computed via dynamic programming. In our new context, cost-to-go functions are generated according to

$$J_d(u) = 0, \quad \forall u \in \mathcal{U},$$
$$J_i(u) = \min_{\pi \in M^\delta} E\left[\frac{1}{d} f_i\big(u, \pi(u, W_i)\big) + J_{i+1}\big(\pi(u, W_i)\big)\right], \quad \forall u \in \mathcal{U}, \ i = 0, \ldots, d-1,$$

where the expectation is taken over W_i. Given the cost-to-go functions, each π_i is chosen to attain the minimum in

$$\min_{\pi \in M^\delta} E\left[\frac{1}{d} f_i\big(u, \pi(u, W_i)\big) + J_{i+1}\big(\pi(u, W_i)\big)\right],$$

simultaneously for all u. Then, the sequence of decisions $\phi^\delta_1(f), \ldots, \phi^\delta_d(f)$ can be generated iteratively by letting $\phi^\delta_0(f) = \iota$ and

$$\phi^\delta_{i+1}(f) = \pi_i\big(\phi^\delta_i(f), W_i\big), \quad \forall i = 0, \ldots, d-1.$$

Note that—similar to the deterministic context—stochastic strategies generated via dynamic programming are *not* decentralized. The iterative selection process induces a dependence of future decisions on preceding ones, while use of cost-to-go functions makes the choice of a decision rely on all future cost information.

Let us close this subsection by introducing some notation that will enable streamlined discussion of ideas we have developed. Let $T^i : \Re^{|\mathcal{U}|} \to \Re^{|\mathcal{U}|}$ be defined by

$$\big(T^i H\big)(u) = \min_{\pi \in M^\delta} E\left[\frac{1}{d} f_i\big(u, \pi(u, W_i)\big) + H\big(\pi(u, W_i)\big)\right], \quad \forall H \in \Re^{|\mathcal{U}|}, \ u \in \mathcal{U}.$$

Also, for any $\pi \in M^\delta$, let $T^i_\pi : \Re^{|\mathcal{U}|} \to \Re^{|\mathcal{U}|}$ be defined by

$$\big(T^i_\pi H\big)(u) = E\left[\frac{1}{d} f_i\big(u, \pi(u, W_i)\big) + H\big(\pi(u, W_i)\big)\right], \quad \forall H \in \Re^{|\mathcal{U}|}, \ u \in \mathcal{U}.$$

Given this notation, we see that the cost-to-go functions are related via

$$J_i = T^i J_{i+1}, \quad \forall i = 0, \ldots, d - 1.$$

Once cost-to-go functions $J_0, \ldots, J_d$ are computed, decisions are generated iteratively according

$$\phi_{i+1}^{\delta}(f) = \pi_i\big(\phi_i^{\delta}(f), W_i\big), \quad \forall i = 0, \ldots, d - 1,$$

where $\phi_0^{\delta}(f) = \iota$, and each $\pi_i \in M^{\delta}$ is chosen to satisfy

$$T_{\pi_i}^i J_{i+1} = T^i J_{i+1}.$$

The expected cost, $E[f(\phi^{\delta}(f))]$, is given by

$$E\big[f\big(\phi^{\delta}(f)\big)\big] = T_{\pi_0}^0 J_1 = J_0.$$

3.3. A decentralized stochastic strategy

In this section, we will formulate a decentralized version of ϕ^{δ}, which will be denoted by $\phi^{\delta,r}$. To motivate the definition of $\phi^{\delta,r}$, let us briefly review why ϕ^{δ} is not decentralized:

(1) Use of cost-to-go functions makes the choice of a decision rely on all future cost information. In other words, the computation of J_i depends on $f_i, f_{i+1}, \ldots, f_{d-1}$. As a result, the mapping π_{i-1} that defines the decision $\phi_i^{\delta}(f)$ depends on $f_i, f_{i+1}, \ldots, f_{d-1}$.
(2) The iterative selection process induces a dependence of future decisions on preceding ones, that is, $\phi_i^{\delta}(f)$ depends on $\phi_{i-1}^{\delta}(f), \ldots, \phi_1^{\delta}(f)$.

Our decentralized stochastic strategy $\phi^{\delta,r}$ alleviates these dependencies.

First, in order to limit the dependence on future cost information to r stages, we consider an approximation to J_i, denoted by $\tilde{J}_i$, which is generated according to $\tilde{J}_i = \tilde{J}_{i,0}$, where

$$\tilde{J}_{i,r} = 0, \qquad \tilde{J}_{i,l} = T^{i+l} \tilde{J}_{i,l+1}, \quad \forall l = 0, \ldots, r - 1.$$

Note that computation of $\tilde{J}_i$ relies only on $f_i, f_{i+1}, \ldots, f_{i+r-1}$. The hope is that—if r is sufficiently large—$\tilde{J}_i$ provides a good approximation to J_i. Given such approximations, we choose a sequence $\tilde{\pi}_0, \ldots, \tilde{\pi}_{d-1} \in M^{\delta}$ to satisfy

$$T_{\tilde{\pi}_i}^i \tilde{J}_{i+1} = T^i \tilde{J}_{i+1}, \quad \forall i = 0, 1, \ldots, d - 1.$$

Note that each $\tilde{\pi}_i$ is determined by the cost functions $f_i, f_{i+1}, \ldots, f_{i+r}$.

One might imagine generating decisions according to

$$\phi_{i+1}^{\delta,r}(f) = \tilde{\pi}_i\big(\phi_i^{\delta,r}(f), W_i\big), \quad \forall i = 0, \ldots, d - 1,$$

with $\phi_0^{\delta,r}(f) = \iota$. However, each such decision would then depend on decisions at all preceding stages. To limit this dependence to r stages we let each decision $\phi_i^{\delta,r}(f)$ be generated according to $\phi_i^{\delta,r}(f) = \phi_{i,0}^{\delta,r}(f)$, where

$$\phi_{i,l+1}^{\delta,r}(f) = \tilde{\pi}_{i+l}\big(\phi_{i,l}^{\delta,r}(f), W_{i+l}\big), \quad \forall l = -r, \ldots, -1,$$

with $\phi_{i,-r}^{\delta,r}(f) = \iota$.

Indeed, $\phi^{\delta,r}$ is a decentralized strategy (with radius r). To see this, consider the process of generating each $\phi_i^{\delta,r}(f)$. It is clear that the distribution of $\phi_i^{\delta,r}(f)$ is completely determined from the mappings $\tilde{\pi}_{i-r}, \tilde{\pi}_{i-r+1}, \ldots, \tilde{\pi}_{i-1}$, which depend only on cost functions $f_{i-r}, \ldots, f_i, \ldots, f_{i+r}$. In addition, for any i, j with $i > j + r$, the decision $\phi_i^{\delta,r}(f)$ makes use of the random numbers $W_{i-r}, \ldots, W_{i-1}$, while the decision $\phi_j^{\delta,r}(f)$ uses $W_{j-r}, \ldots, W_{j-1}$. Since W_i's are independent, $\phi_i^{\delta,r}(f)$ is independent of $\phi_j^{\delta,r}(f)$. The situation when $j > i + r$ is analogous.

Before we proceed with the performance analysis, let us briefly discuss the choice of decentralized stochastic strategies proposed in this paper. One approach to defining a stochastic strategy involves allowing each agent to generate random numbers that influence only his own decision. This approach would imply independence of decisions made by each agent. However, the decisions generated by each agent under an optimal strategy are dependent. To allow for this dependency, our formulation associates a random numbers with nodes in the chain. When an agent observes cost-relevant data at a node, he also observes its random number. This random number might be thought of as cost-irrelevant data that can be observed by agents within the vicinity. The use of common random numbers induces dependencies among decisions made by different agents. Nevertheless, each agent's decision depends only on information within its vicinity. We will establish in the next section that—given appropriate choices of δ and $r - \phi^{\delta,r}$ satisfies the properties promised by Theorem 2.

4. Performance analysis

In this section, we analyze performance of decentralized strategies. We start in Section 4.1 by proving Theorem 1—our negative result concerning decentralized deterministic strategies. The remaining sections pursue an analysis of the stochastic decentralized strategy proposed in the previous section, showing that it does indeed exhibit properties promised by Theorem 2. This constitutes a constructive proof of the theorem. The analysis is presented in stages. In Section 4.2, we start by arguing that the performance of an optimal δ-stochastic strategy ϕ^δ is close to that of an optimal centralized strategy ϕ^*. Then, in Section 4.3, the decentralized stochastic strategy $\phi^{\delta,r}$ is shown to perform almost as well as ϕ^δ, provided that r is sufficiently large. The proof of Theorem 2 is almost a direct consequence, as discussed in Section 4.4.

4.1. Performance of decentralized deterministic strategies (Proof of Theorem 1)

Recall that Theorem 1 identifies performance limitations of decentralized deterministic strategies. Formally, the theorem states that for any $|\mathcal{U}| \geqslant 2$ and $B > 0$, there exists a positive scalar c such that

$$\min_{\phi \in \Phi_{d,\mathcal{U},r}} \max_{f \in \Psi_{d,\mathcal{U},B}} \left(f\big(\phi(f)\big) - f\big(\phi^*(f)\big) \right) \geqslant c\left(\frac{1}{4} - \frac{r}{d}\right),$$

for all $d \geqslant 4$ and $4r < d$.

To prove the theorem, it suffices to consider the case of $\mathcal{U} = \{0, 1\}$. Let $f^* : \mathcal{U} \times \mathcal{U} \mapsto [-B, B]$ be defined by

$$f^*(u, v) = \begin{cases} B & \text{if } u = v, \\ -B & \text{if } u \neq v. \end{cases}$$

Consider a problem instance $f = (f_1, \ldots, f_{d-1}) \in \Psi_{d,\mathcal{U},B}$ such that $f_i = f^*$ for all i. For this problem, the optimal decision is either $(0, 1, 0, \ldots)$ or $(1, 0, 1, \ldots)$ with an associated cost of $-B(d-1)/d$.

Let us now consider decisions generated by an arbitrary strategy $\phi \in \Phi_{d,\mathcal{U},r}$. Since $f_i = f^*$ for all i, it follows from the definition of $\Phi_{d,\mathcal{U},r}$ that

$$\phi_i(f) = \phi_j(f), \quad r + 1 \leqslant i, j \leqslant d - 1 - r,$$

which implies that

$$f\big(\phi(f)\big) \geqslant \frac{(d - 2r - 2)B}{d} - \frac{B(2r + 1)}{d},$$

where $(d - 2r - 2)B/d$ corresponds to the cost associated with decisions indexed $r + 1, \ldots, d - 1 - r$, while $-B(2r + 1)/d$ represents a lower bound on costs associated with decisions indexed $1, \ldots, r, d - r, \ldots, d$. Therefore,

$$f\big(\phi(f)\big) - f\big(\phi^*(f)\big) \geqslant \left(\frac{d - 2r - 2}{d} - \frac{2r + 1}{d} + \frac{d - 1}{d} \right) B = 4B \left(\frac{1}{2} - \frac{1}{d} - \frac{r}{d} \right)$$

$$\geqslant 4B \left(\frac{1}{4} - \frac{r}{d} \right),$$

where the last inequality follows from the fact that $d \geqslant 4$. Hence,

$$\max_{f \in \Psi_{d,\mathcal{U},B}} \big(f(\phi(f)) - f(\phi^*(f)) \big) \geqslant 4B \left(\frac{1}{4} - \frac{r}{d} \right).$$

Since $\phi \in \Phi_{d,\mathcal{U},r}$ is arbitrary,

$$\min_{\phi \in \Phi_{d,\mathcal{U},r}} \max_{f \in \Psi_{d,\mathcal{U},B}} \big(f(\phi(f)) - f(\phi^*(f)) \big) \geqslant 4B \left(\frac{1}{4} - \frac{r}{d} \right).$$

Let us highlight the intuitive reason for poor performance of decentralized deterministic strategies and briefly discuss possible remedies. This discussion will hopefully motivate the choice of decentralized stochastic strategy considered in this paper. The primary reason for poor performance is the requirement that each agent's decision depends only on cost functions within its vicinity, i.e., $\phi_i(f) = \phi_j(f)$ whenever $f_i^r = f_j^r$. Given the particular cost function used in the proof, this condition implies that the decision at each node, except for those near the boundary, must be the same. This leads to poor performance in the face of our cost function, which is optimized when adjacent agents make different decisions. This situation underscores the pitfalls associated with deterministic strategies in the absence of coordination. One solution to this shortcoming is to allow coordination among agents. However, when coordination is impractical or impossible, an alternative solution is called for.

The decentralized stochastic strategy presented in Section 3.3 offers such an alternative. By appropriate randomization, this strategy avoids the situation where every agent makes

the same decision. In some sense, the fact that nearby agents use common random numbers enables a form of coordination that reduces the probability that adjacent agents select the same decision. This fact will be employed in the upcoming performance analysis, which establishes the desirable performance bound offered by Theorem 2.

4.2. Performance of δ-stochastic strategies

As captured formally by the following lemma, the performance of optimal δ-stochastic strategies degrades at most linearly as δ grows.

Lemma 1.

$$\max_{f \in \Psi_{d,U,B}} \left(E\left[f\left(\phi^\delta(f) \right) \right] - f\left(\phi^*(f) \right) \right) \leqslant 4\delta B.$$

Proof. Fix $f \in \Psi_{d,\mathcal{U},B}$, and let ϕ^* be an optimal strategy. Consider a δ-stochastic strategy that generates decisions using a sequence $\pi_0, \ldots, \pi_{d-1} \in M^\delta$ defined by

$$\pi_i(U, W) = \begin{cases} \phi_{i+1}^*(f) & \text{if } 0 \leqslant W < 1 - \delta, \\ k\text{th element of } \mathcal{U} & \text{if } 1 - \delta + \dfrac{(k-1)\delta}{|\mathcal{U}|} \leqslant W < 1 - \delta + \dfrac{k\delta}{|\mathcal{U}|}, \end{cases}$$

where $\pi_i(U, 1)$ is defined to the $|\mathcal{U}|$th element of $\mathcal{U}$. Hence, the decisions are random variables generated according to

$$U_{i+1} = \pi_i(U_i, W_i), \quad \forall i = 0, \ldots, d - 1,$$

with $U_0 = \iota$. Since each $\pi_i(U_i, W_i)$ depends only on W_i, the decisions $U_1, \ldots, U_d$ are independent random variables. We therefore have $(U_i, U_{i+1}) = (\phi_i^*(f), \phi_{i+1}^*(f))$ with probability at least $(1 - \delta)^2$, which implies that

$$E\left| f_i\left(\phi_i^*(f), \phi_{i+1}^*(f) \right) - f_i(U_i, U_{i+1}) \right| \leqslant 2B\left(1 - (1 - \delta)^2 \right) \leqslant 4\delta B,$$

where the inequalities hold because $\| f_i \|_\infty \leqslant B$ and $1 - (1 - \delta)^2 \leqslant 2\delta$. Since f is arbitrary and ϕ^δ attains the optimum within the class of δ-stochastic strategies, it follows that

$$\max_{f \in \Psi_{d,U,B}} \left(E\left[f\left(\phi^\delta(f) \right) \right] - f\left(\phi^*(f) \right) \right) \leqslant 4\delta B. \quad \square$$

4.3. Performance losses from decentralizing δ-stochastic strategies

In this section, we will show that as r grows, the performance of $\phi^{\delta,r}$ becomes close to that of ϕ^δ in a sense stated formally in the following lemma.

Lemma 2.

$$\max_{f \in \Psi_{d,U,B}} \left(E\left[f\left(\phi^{\delta,r}(f) \right) \right] - E\left[f\left(\phi^\delta(f) \right) \right] \right) \leqslant \left(\frac{2B}{\delta} + 4B|\mathcal{U}|^2 \right) \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r.$$

Instead of comparing directly the performance of $\phi^{\delta,r}$ with that of ϕ^{δ}, we will first introduce an *r-lookahead* strategy, denoted by $\bar{\phi}^{\delta,r}$. This is a variant of the centralized stochastic strategy ϕ^{δ} that can be viewed as an intermediate step in the transition from the centralized strategy ϕ^{δ} to the decentralized strategy $\phi^{\delta,r}$. The strategy $\bar{\phi}^{\delta,r}$ shares common features with $\phi^{\delta,r}$, though it makes use of centralized information as does ϕ^{δ}. After studying *r-lookahead* strategies in Section 4.3.1, we relate in Section 4.3.2 such strategies to the decentralized strategy $\phi^{\delta,r}$ in order to prove Lemma 2.

4.3.1. *r-lookahead strategies and their performance*

Recall that the decentralized strategy $\phi^{\delta,r}$ entails computation of a sequence $\tilde{J}_0, \ldots, \tilde{J}_d$ of approximate cost-to-go functions. In particular, for each i, $\tilde{J}_i$ is obtained from the following recursion:

$$\tilde{J}_{i,r} = 0, \qquad \tilde{J}_{i,l} = T^{i+l}\tilde{J}_{i,l+1}, \quad \forall l = 0, \ldots, r-1,$$

where we set $\tilde{J}_i = \tilde{J}_{i,0}$. Mappings $\tilde{\pi}_0, \ldots, \tilde{\pi}_{d-1}$ are then chosen to satisfy

$$T^i_{\tilde{\pi}_i}\tilde{J}_{i+1} = T^i\tilde{J}_{i+1}, \quad \forall i = 0, 1, \ldots, d-1.$$

A sequence of decisions $\phi^{\delta,r}_{i,-r}(f), \ldots, \phi^{\delta,r}_{i,0}(f)$ is then computed according to

$$\phi^{\delta,r}_{i,l+1}(f) = \tilde{\pi}_{i+l}\big(\phi^{\delta,r}_{i,l}(f), W_{i+l}\big), \quad \forall l = -r, \ldots, -1,$$

initialized with $\phi^{\delta,r}_{i,-r}(f) = \iota$, and the decision $\phi^{\delta,r}_i(f)$ is set to be equal to $\phi^{\delta,r}_{i,0}(f)$.

An *r-lookahead* strategy entails computation of the same approximate cost-to-go functions $\tilde{J}_0, \ldots, \tilde{J}_d$ and mappings $\tilde{\pi}_0, \ldots, \tilde{\pi}_{d-1}$. However, the *r*-lookahead strategy $\bar{\phi}^{\delta,r}$ differs from the decentralized strategy in the way that these mappings are used to generate decisions. In particular, $\bar{\phi}^{\delta,r}$ generates decisions according to

$$\bar{\phi}^{\delta,r}_{i+1}(f) = \tilde{\pi}_i\big(\bar{\phi}^{\delta,r}_i(f), W_i\big), \quad i = 0, 1, \ldots, d,$$

initialized with $\bar{\phi}^{\delta,r}_0(f) = \iota$. Note that $\bar{\phi}^{\delta,r}$ is *not* decentralized since each decision depends on all of its predecessors.

When r becomes large, the performance of $\bar{\phi}^{\delta,r}$ becomes close to that of the centralized stochastic strategy ϕ^{δ}, as captured in the following lemma.

Lemma 3.

$$\max_{f \in \Psi_{d,U,B}} \big(E\big[f\big(\bar{\phi}^{\delta,r}(f)\big)\big] - E\big[f\big(\phi^{\delta}(f)\big)\big]\big) \leqslant \frac{2B}{\delta}\left(1 - \frac{\delta}{|\mathcal{U}|}\right)^r.$$

We will prove this lemma in two stages. First, we will show that the approximate cost-to-go functions $\tilde{J}_i$ is close to the true cost-to-go function J_i, in some suitably defined sense. We then use this fact to prove that the two strategies deliver similar performance.

As a metric for the proximity between J_i and $\tilde{J}_i$ we will use the span-seminorm $\|\cdot\|_s$. This seminorm is defined by

$$\|H\|_s = \max_{u \in \mathcal{U}} H(u) - \min_{u \in \mathcal{U}} H(u).$$

for all $H \in \Re^{|\mathcal{U}|}$. It is not hard to verify that for any scalar $\alpha \in \Re$,

$$\|\alpha H\|_s = |\alpha| \|H\|_s,$$

and for H_1, $H_2 \in \Re^{|\mathcal{U}|}$,

$$\|H_1 + H_2\|_s \leqslant \|H_1\|_s + \|H_2\|_s.$$

Moreover, if $\|H\|_s = 0$, then $H = Ce$ for some scalar $C \in \Re$. We have the following result whose proof is given in Appendix A.

Lemma 4. *For each i,*

$$\left\| J_i - \tilde{J}_i \right\|_s \leqslant \frac{2B}{\delta d} \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r.$$

The key ideas underlying the proof of Lemma 4 are a contraction property associated with T_i and a bound on the cost-to-go function J_i. As shown in Appendix A, the contraction property implies that

$$\left\| J_i - \tilde{J}_i \right\|_s \leqslant \|J_{i+r}\|_s \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r.$$

The maximum cost at each time step is of order B/d and J_{i+r} denotes the expected cost-to-go starting from time $i + r$; thus, we might expect that $\|J_{i+r}\|_s$ is of order $(d - i - r) \times 2B/d$, where $(d - i - r)$ denotes the remaining time after $i + r$. However, the randomization in our stochastic strategies induces mixing, which attenuates future costs. This allows us to obtain a sharper bound on $\|J_{i+r}\|_s$ that is independent of i:

$$\|J_{i+r}\|_s \leqslant \frac{2B}{d} \left(1 + (1 - \delta) + \cdots + (1 - \delta)^{d-i-r} \right) \leqslant \frac{2B}{\delta d}.$$

The lemma follows from this bound.

The proximity between J_i and $\tilde{J}_i$ suggests that the mappings π_i and $\tilde{\pi}_i$ that are generated from J_i and $\tilde{J}_i$, respectively, should exhibit some form of similarity. This notion is captured by the following lemma, whose proof is provided in Appendix B.

Lemma 5. *For H_1, $H_2 \in \Re^{|\mathcal{U}|}$, let $\theta_1, \theta_2 \in M^\delta$ satisfy*

$$T_{\theta_1}^i H_1 = T^i H_1 \quad and \quad T_{\theta_2}^i H_2 = T^i H_2.$$

Then,

$$\left\| T_{\theta_1}^i H_1 - T_{\theta_2}^i H_1 \right\|_\infty \leqslant \|H_1 - H_2\|_s.$$

When applied in a context where H_1 is a true cost-to-go function and H_2 is an approximation, this result shows that the performance of the greedy policy θ_2 that is derived from H_2 approximates that of an optimal policy, provided that H_2 is close to H_1 (in a span-seminorm sense). Let us now prove Lemma 3, which almost directly follows from Lemmas 4 and 5.

Proof of Lemma 3. Fix $f \in \Psi_{d,U,B}$. Let $\phi^\delta(f) = (\phi_1^\delta(f), \ldots, \phi_d^\delta(f))$ and $\bar{\phi}^{\delta,r}(f) = (\bar{\phi}_1^{\delta,r}(f), \ldots, \bar{\phi}_d^{\delta,r}(f))$ denote the decisions under the strategies ϕ^δ and $\bar{\phi}^{\delta,r}$, respectively. Let J_i and $\bar{J}_i$ denote the cost-to-go functions under $\phi^\delta(f)$ and $\bar{\phi}^{\delta,r}(f)$, respectively. These cost-to-go functions are related via the following dynamic programming recursion:

$$J_i = T_{\pi_i}^i J_{i+1} \quad \text{and} \quad \bar{J}_i = T_{\tilde{\pi}_i}^i \bar{J}_{i+1},$$

where we initialize $J_d = \bar{J}_d = 0$. Therefore,

$$
\begin{aligned}
\left\| J_i - \bar{J}_i \right\|_\infty &= \left\| T_{\pi_i}^i J_{i+1} - T_{\tilde{\pi}_i}^i \bar{J}_{i+1} \right\|_\infty \\
&\leqslant \left\| T_{\pi_i}^i J_{i+1} - T_{\tilde{\pi}_i}^i J_{i+1} \right\|_\infty + \left\| T_{\tilde{\pi}_i}^i J_{i+1} - T_{\tilde{\pi}_i}^i \bar{J}_{i+1} \right\|_\infty \\
&\leqslant \left\| T_{\pi_i}^i J_{i+1} - T_{\tilde{\pi}_i}^i J_{i+1} \right\|_\infty + \left\| J_{i+1} - \bar{J}_{i+1} \right\|_\infty,
\end{aligned}
$$

where the last inequality follows from a standard result in dynamic programming that the operator $T_{\tilde{\pi}_i}^i$ is non-expansive under the maximum norm (see, e.g., (Bertsekas, 1995)). Recall that the mappings π_i and $\tilde{\pi}_i$ are chosen to satisfy

$$T_{\pi_i}^i J_{i+1} = T^i J_{i+1} \quad \text{and} \quad T_{\tilde{\pi}_i}^i \tilde{J}_{i+1} = T^i \tilde{J}_{i+1},$$

where $\tilde{J}_{i+1}$ is the approximate cost-to-go functions computed under the r-lookahead strategy. It follows from Lemma 4 and 5 that

$$\left\| T_{\pi_i}^i J_{i+1} - T_{\tilde{\pi}_i}^i J_{i+1} \right\|_\infty \leqslant \left\| J_{i+1} - \tilde{J}_{i+1} \right\|_s \leqslant \frac{2B}{\delta d} \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r.$$

Therefore,

$$\left\| J_i - \bar{J}_i \right\|_\infty \leqslant \frac{2B}{\delta d} \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r + \left\| J_{i+1} - \bar{J}_{i+1} \right\|_\infty.$$

Using the fact that $J_d = \bar{J}_d = 0$, the above recursion implies that

$$\left\| J_0 - \bar{J}_0 \right\|_\infty \leqslant \frac{2B}{\delta} \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r.$$

Note that

$$E\left[f\left(\phi^\delta(f) \right) \right] = T_{\pi_0}^0 J_1 = J_0 \quad \text{and} \quad E\left[f\left(\bar{\phi}^{\delta,r}(f) \right) \right] = T_{\tilde{\pi}_0}^0 \bar{J}_1 = \bar{J}_0,$$

therefore,

$$\left(E\left[f\left(\bar{\phi}^{\delta,r}(f) \right) \right] - E\left[f\left(\phi^\delta(f) \right) \right] \right) \leqslant \frac{2B}{\delta} \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r.$$

Since $f \in \Psi_{d,\mathcal{U},B}$ is arbitrary,

$$\max_{f \in \Psi_{d,U,B}} \left(E\left[f\left(\bar{\phi}^{\delta,r}(f) \right) \right] - E\left[f\left(\phi^\delta(f) \right) \right] \right) \leqslant \frac{2B}{\delta} \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r. \qquad \square$$

4.3.2. Performance of the decentralized strategy (Proof of Lemma 2)

In this section, we provide a proof of Lemma 2, which relies heavily on Lemma 3 and similarities between $\bar{\phi}^{\delta,r}$ and $\phi^{\delta,r}$. The proof also makes use of the following lemma, which relates distributions of decisions under the two strategies. The proof of this lemma is given in Appendix C, and it exploits mixing of the stochastic decision processes associated with $\bar{\phi}^{\delta,r}$ and $\phi^{\delta,r}$.

Lemma 6. *For $f \in \Psi_{d,\mathcal{U},B}$ and $x, y \in \mathcal{U}$,*

$$\left| \Pr\left\{ \left(\phi_i^{\delta,r}(f), \phi_{i+1}^{\delta,r}(f) \right) = (x, y) \right\} - \Pr\left\{ \left(\bar{\phi}_i^{\delta,r}(f), \bar{\phi}_{i+1}^{\delta,r}(f) \right) = (x, y) \right\} \right| \leqslant 4(1 - \delta)^r,$$

for all i.

The above lemma tells us that the decision-processes generated by the two strategies exhibit almost the same distribution. This suggests that expected costs should also be close. The proof of Lemma 2, provided below, combines this insight with the Lemma 1.

Proof. Fix $f \in \Psi_{d,\mathcal{U},B}$. For any i, we have

$$E\left[f_i\left(\phi_i^{\delta,r}(f), \phi_{i+1}^{\delta,r}(f) \right) \right] = \sum_{(x,y) \in \mathcal{U} \times \mathcal{U}} f_i(x, y) \Pr\left\{ \left(\phi_i^{\delta,r}(f), \phi_{i+1}^{\delta,r}(f) \right) = (x, y) \right\},$$

and

$$E\left[f_i\left(\bar{\phi}_i^{\delta,r}(f), \bar{\phi}_{i+1}^{\delta,r}(f) \right) \right] = \sum_{(x,y) \in \mathcal{U} \times \mathcal{U}} f_i(x, y) \Pr\left\{ \left(\bar{\phi}_i^{\delta,r}(f), \bar{\phi}_{i+1}^{\delta,r}(f) \right) = (x, y) \right\}.$$

Thus, it follows from Lemma 6 that

$$\left| E\left[f_i\left(\phi_i^{\delta,r}(f), \phi_{i+1}^{\delta,r}(f) \right) \right] - E\left[f_i\left(\bar{\phi}_i^{\delta,r}(f), \bar{\phi}_{i+1}^{\delta,r}(f) \right) \right] \right|$$
$$\leqslant 4(1 - \delta)^r \sum_{(x,y) \in \mathcal{U} \times \mathcal{U}} \left| f_i(x, y) \right| \leqslant 4B|\mathcal{U}|^2 (1 - \delta)^r,$$

where the last inequality follows from the fact that $\| f_i \|_\infty \leqslant B$. Therefore,

$$\left| E\left[f\left(\phi^{\delta,r}(f) \right) \right] - E\left[f\left(\bar{\phi}^{\delta,r}(f) \right) \right] \right| \leqslant 4B|\mathcal{U}|^2 (1 - \delta)^r.$$

Since $f \in \Psi_{d,U,B}$ is arbitrary, it follows that

$$\max_{f \in \Psi_{d,U,B}} \left| E\left[f\left(\phi^{\delta,r}(f) \right) \right] - E\left[f\left(\bar{\phi}^{\delta,r}(f) \right) \right] \right| \leqslant 4B|\mathcal{U}|^2 (1 - \delta)^r$$

$$\leqslant 4B|\mathcal{U}|^2 \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r.$$

We know from Lemma 3 that

$$\max_{f \in \Psi_{d,U,B}} \left(E\left[f\left(\bar{\phi}^{\delta,r}(f) \right) \right] - E\left[f\left(\phi^{\delta}(f) \right) \right] \right) \leqslant \frac{2B}{\delta} \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r,$$

from which it follows that

$$\max_{f \in \Psi_{d,U,B}} \left(E\left[f\left(\phi^{\delta,r}(f) \right) \right] - E\left[f\left(\phi^{\delta}(f) \right) \right] \right) \leqslant \left(\frac{2B}{\delta} + 4B|\mathcal{U}|^2 \right) \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r. \quad \square$$

4.4. Proof of Theorem 2

The proof of Theorem 2 makes use of the following lemma whose proof is given in Appendix D. This result will be used to establish the desired upper bound on the worst-case performance of decentralized stochastic strategies.

Lemma 7. *For any $a > 0$ and $0 < p < 1$, there exists a scalar c' such that*

$$r^p \left(1 - \frac{a}{r^p} \right)^r \leqslant \frac{c'}{r^p},$$

for all $r \geqslant 1$.

Proof of Theorem 2. It follows from Lemmas 1 and 2 that

$$\max_{f \in \Psi_{d,U,B}} \left(E\left[f\left(\phi^\delta(f) \right) \right] - f\left(\phi^*(f) \right) \right) \leqslant 4\delta B$$

and

$$\max_{f \in \Psi_{d,U,B}} \left(E\left[f\left(\phi^{\delta,r}(f) \right) \right] - E\left[f\left(\phi^\delta(f) \right) \right] \right) \leqslant \left(\frac{2B}{\delta} + 4B|\mathcal{U}|^2 \right) \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r.$$

Therefore,

$$\max_{f \in \Psi_{d,U,B}} \left(E\left[f\left(\phi^{\delta,r}(f) \right) \right] - f\left(\phi^*(f) \right) \right) \leqslant 4\delta B + \left(\frac{2B}{\delta} + 4B|\mathcal{U}|^2 \right) \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r.$$

Since $\delta \in (0, 1)$ can be chosen arbitrarily, let $\delta = r^{-p}$. It follows that

$$\min_{\phi \in \widetilde{\Phi}_{d,\mathcal{U},r}} \max_{f \in \Psi_{d,U,B}} \left(E\left[f\left(\phi(f) \right) \right] - f\left(\phi^*(f) \right) \right)$$

$$\leqslant \frac{4B}{r^p} + \left(2Br^p + 4B|\mathcal{U}|^2 \right) \left(1 - \frac{1}{|\mathcal{U}|r^p} \right)^r.$$

It follows from Lemma 7 (with $a = 1/|\mathcal{U}|$) that

$$\left(2Br^p + 4B|\mathcal{U}|^2 \right) \left(1 - \frac{1}{|\mathcal{U}|r^p} \right)^r \leqslant \frac{c''}{r^p},$$

for some constant c''. The desired result follows. $\quad\square$

5. Allowing coordination

In order to preclude cooperation, our definition of decentralized deterministic strategies requires that $\phi_i(f) = \phi_j(f)$ whenever $f_i^r = f_j^r$. This condition implies that each agent's decision does not depend on its relative position in the chain. Theorem 1 identifies an undesirable loss in performance introduced by decentralization when strategies are required to be deterministic and coordination is forbidden. In particular, to maintain a given level of error, r is required to be proportional to d. The situation is much improved

by the introduction of stochastic strategies. Theorem 2 establishes that under the stochastic strategy, the radius of decentralization r is independent of d. In particular, the loss of performance due to decentralization is bounded for any $p < 1$ by c/r^p, where c is a scalar whose value does not depend on d.

Let us consider a situation in which coordination is allowed, but strategies are required to be deterministic. In particular, the strategy space is

$$\overline{\Phi}_{d,\mathcal{U},r} = \left\{ \phi \in \Phi_{d,\mathcal{U}} \mid \forall f \in \Psi_{d,\mathcal{U},B}, \; \phi_i(f) = \phi_i\big(f_i^r\big) \right\}.$$

The following theorem provides a bound offering a performance guarantee superior to that associated with decentralized stochastic strategies.

Theorem 3. *For any $\mathcal{U}$ and B,*

$$\min_{\phi \in \overline{\Phi}_{d,\mathcal{U},r}} \; \max_{f \in \Psi_{d,\mathcal{U},B}} \big(f\big(\phi(f)\big) - f\big(\phi^*(f)\big) \big) \leqslant \frac{2B}{r}.$$

Proof. Without loss of generality, let us assume that $d = kr$ for some integer k. The same argument applies for general d. Let $f \in \Psi_{d,\mathcal{U},B}$ be given, and let $f^i : \mathcal{U}^r \mapsto \Re$ be defined by

$$f^i(u_{ir+1}, \ldots, u_{ir+r}) = \sum_{l=1}^{r-1} f_{ir+l}(u_{ir+l}, u_{ir+l+1}),$$

for $i = 0, \ldots, k-1$. Also, let $\bar{u}^i = (\bar{u}_{ir+1}, \ldots, \bar{u}_{ir+r}) \in \mathcal{U}^r$ be defined by

$$f^i\big(\bar{u}^i\big) = \min_{u \in \mathcal{U}^r} f^i(u), \quad \forall i = 0, \ldots, k-1.$$

Consider a strategy $\bar{\phi} \in \overline{\Phi}_{d,\mathcal{U},B}$ such that $\bar{\phi}(f) = (\bar{u}^0, \ldots, \bar{u}^{k-1}) \in \mathcal{U}^d$. If $\phi^*(f) = (u_1^*, \ldots, u_d^*)$ denotes the optimal decision for this problem instance, then

$$
\begin{aligned}
f\big(\bar{\phi}(f)\big) &= \frac{1}{d} \sum_{i=0}^{k-1} f^i\big(\bar{u}^i\big) + \frac{1}{d} \sum_{i=1}^{k-1} f_{ir}(\bar{u}_{ir}, \bar{u}_{ir+1}) \\
&= \frac{1}{d} \sum_{i=0}^{k-1} \sum_{l=1}^{r-1} f_{ir+l}(\bar{u}_{ir+l}, \bar{u}_{ir+l+1}) + \frac{1}{d} \sum_{i=1}^{k-1} f_{ir}(\bar{u}_{ir}, \bar{u}_{ir+1}) \\
&\leqslant \frac{1}{d} \sum_{i=0}^{k-1} \sum_{l=1}^{r-1} f_{ir+l}\big(u_{ir+l}^*, u_{ir+l+1}^*\big) + \frac{1}{d} \sum_{i=1}^{k-1} f_{ir}(\bar{u}_{ir}, \bar{u}_{ir+1}) \\
&= f\big(\phi^*(f)\big) + \frac{1}{d} \sum_{i=1}^{k-1} f_{ir}(\bar{u}_{ir}, \bar{u}_{ir+1}) - f_{ir}\big(u_{ir}^*, u_{ir+1}^*\big) \\
&\leqslant f\big(\phi^*(f)\big) + \frac{2B}{r}
\end{aligned}
$$

where the last inequality follows from the fact that $k = d/r$, and $\|f_i\|_\infty \leqslant B$ for all i. Since $\bar{\phi} \in \overline{\Phi}_{d,\mathcal{U},r}$ and $f \in \Psi_{d,\mathcal{U},B}$ is arbitrary, it follows that

$$\min_{\phi \in \overline{\Phi}_{d,\mathcal{U},r}} \max_{f \in \Psi_{d,\mathcal{U},B}} \left(f\big(\phi(f)\big) - f\big(\phi^*(f)\big) \right) \leqslant \frac{2B}{r}. \qquad \square$$

The above result suggests that coordination—rather than randomization—is a desirable way to go. The loss in performance diminishes at the rate of $1/r$, which is faster than that associated with decentralized stochastic strategies that do not coordinate. As illustrated in the constructive proof provided above, this is achieved simply by partitioning agents into groups of r consecutive nodes. However, as we will now discuss, coordination is often impractical. In making a case for this claim, let us consider the more general case of a network, rather than a chain. One might imagine situations where the network topology changes over time, as agents are introduced or deleted and connections are created or destroyed. At the time when a decision is required, the topology might not be known to all agents. Formation of appropriate groups for coordinated decision-making, however, entails either centralized knowledge of this topology or iterative protocols that can require time proportional with the size of the network. Such requirements become prohibitive when networks are large.

In contrast, decentralized stochastic strategies offer comparable performance without coordination. Each agent needs only information associated with his local vicinity. Such strategies may therefore provide more practical means to effective decentralization.

6. Concluding remarks

We have studied the problem of decentralized decision-making in a chain, in the absence of coordination. Our main results identify advantages of stochastic over deterministic strategies. In particular, the amount of information required by each agent in order to guarantee a certain level of performance is independent of problem size, when appropriate stochastic strategies are employed. With deterministic strategies, on the other hand, information requirements can grow proportionately with the total number of agents.

A by-product of our main result is an efficient algorithm that generates stochastic strategies that deliver promised performance. In fact, the analysis was constructive, involving verification that strategies generated by this algorithm do indeed satisfy the error bound.

There are several interesting directions in which the results of this paper might be extended. One involves generalizing the results to encompass networks of arbitrary topology. Another interesting direction is to study the implications of such results in stochastic networks that are controlled over time, such as queuing networks. The topic of decentralized control in dynamic systems has been studied for quite some time (see, e.g., (Sandell Jr. et al., 1978) for a survey), but there has generally been a lack of effective methods. Even relatively simple problems have proven to be computationally intractable. Whether limited interdependencies—as captured by graphical structure—can reduce the complexity of such problems poses an intriguing research question.

Acknowledgments

Appendix A. Proof of Lemma 4

In this section, we will show that the approximate cost-to-go function $\tilde{J}_i$ is close to the true cost-to-go function J_i. The metric that we will employ is the span-seminorm. Recall that

$$\|H\|_s = \max_{u \in \mathcal{U}} H(u) - \min_{u \in \mathcal{U}} H(u),$$

for all $H \in \Re^{\mathcal{U}}$. The proof of Lemma 4 requires the following result which shows that the operator T^i is a contraction with respect to the span-seminorm.

Lemma 8. *For any* $H_1, H_2 \in \Re^{|\mathcal{U}|}$,

$$\|T^i H_1 - T^i H_2\|_s \leqslant \left(1 - \frac{\delta}{|\mathcal{U}|}\right) \|H_1 - H_2\|_s.$$

Proof. From the definition of $\|\cdot\|_s$, it follows that for any $H \in \Re^{|\mathcal{U}|}$ and $C \in \Re$,

$$\|H\|_s = \|H + Ce\|_s,$$

where e denotes a vector of all ones. Fix $\bar{u} \in \mathcal{U}$. Then,

$$\begin{aligned}
\|T^i H_1 - T^i H_2\|_s &= \|T^i H_1 - T^i H_2 - (H_1(\bar{u}) - H_2(\bar{u}))e\|_s \\
&= \|T^i(H_1 - H_1(\bar{u})e) - T^i(H_2 - H_2(\bar{u})e)\|_s,
\end{aligned}$$

since $T^i(H + Ce) = T^i H + Ce$ for any $C \in \Re$. Hence, we can assume without loss of generality that $H_1(\bar{u}) = H_2(\bar{u}) = 0$. Let $\theta_1, \theta_2 \in M^\delta$ be defined by

$$T^i_{\theta_1} H_1 = T^i H_1 \quad \text{and} \quad T^i_{\theta_2} H_2 = T^i H_2.$$

It follows from the definition of T^i that for any $u \in \mathcal{U}$,

$$\left(T^i H_1\right)(u) = E\left[\frac{1}{d} f_i\left(u, \theta_1(u, W_i)\right) + H_1\left(\theta_1(u, W_i)\right)\right] \leqslant E\left[\frac{1}{d} f_i\left(u, \theta_2(u, W_i)\right) + H_1\left(\theta_2(u, W_i)\right)\right].$$

Similarly,

$$\left(T^i H_2\right)(u) = E\left[\frac{1}{d} f_i\left(u, \theta_2(u, W_i)\right) + H_2\left(\theta_2(u, W_i)\right)\right] \leqslant E\left[\frac{1}{d} f_i\left(u, \theta_1(u, W_i)\right) + H_2\left(\theta_1(u, W_i)\right)\right],$$

and therefore,

$$E\left[(H_1 - H_2)\left(\theta_1(u, W_i)\right)\right] \leqslant \left(T^i H_1 - T^i H_2\right)(u) \leqslant E\left[(H_1 - H_2)\left(\theta_2(u, W_i)\right)\right].$$

Since $H_1(\bar{u}) = H_2(\bar{u}) = 0$ and $\theta_2 \in M^\delta$, it follows that

$$\left(T^i H_1 - T^i H_2\right)(u) \leqslant E\left[(H_1 - H_2)\left(\theta_2(u, W_i)\right)\right] = \sum_{v \neq \bar{u}} \Pr\{\theta_2(u, W_i) = v\}(H_1 - H_2)(v)$$

$$\leqslant \left(\sum_{v \neq \bar{u}} \Pr\{\theta_2(u, W_i) = v\}\right) \max_{v \in \mathcal{U}}(H_1 - H_2)(v)$$

$$\leqslant \left(1 - \frac{\delta}{|\mathcal{U}|}\right) \max_{v \in \mathcal{U}}(H_1 - H_2)(v).$$

Since $u \in \mathcal{U}$ is arbitrary,

$$\max_{u \in \mathcal{U}}\left(T^i H_1 - T^i H_2\right)(u) \leqslant \left(1 - \frac{\delta}{|\mathcal{U}|}\right) \max_{u \in \mathcal{U}}(H_1 - H_2)(u).$$

A similar argument shows that

$$\min_{u \in \mathcal{U}}\left(T^i H_1 - T^i H_2\right)(u) \geqslant \left(1 - \frac{\delta}{|\mathcal{U}|}\right) \min_{u \in \mathcal{U}}(H_1 - H_2)(u).$$

The last two inequalities imply that

$$\left\| T^i H_i - T^i H_2 \right\|_s = \max_{u \in \mathcal{U}}\left(T^i H_1 - T^i H_2\right)(u) - \min_{u \in \mathcal{U}}\left(T^i H_1 - T^i H_2\right)(u)$$

$$\leqslant \left(1 - \frac{\delta}{|\mathcal{U}|}\right)\left(\max_{u \in \mathcal{U}}(H_1 - H_2)(u) - \min_{u \in \mathcal{U}}(H_1 - H_2)(u)\right)$$

$$= \left(1 - \frac{\delta}{|\mathcal{U}|}\right)\|H_1 - H_2\|_s. \qquad \square$$

Another property making the span-seminorm a desirable instrument for analysis of our problem at hand is that $\|J_i\|_s$ is uniformly bounded over i, and that the bound is inversely proportional with d.

Lemma 9. *For each i,*

$$\|J_i\|_s \leqslant \frac{2B}{\delta d}.$$

Proof. Let $\bar{u}, \underline{u} \in \mathcal{U}$ be chosen to satisfy

$$J_i(\bar{u}) = \max_{u \in \mathcal{U}} J_i(u) \quad \text{and} \quad J_i(\underline{u}) = \min_{u \in \mathcal{U}} J_i(u).$$

Since $J_i = T^i J_{i+1} = T^i_{\pi_i} J_{i+1}$,

$$J_i(\bar{u}) = \left(T^i_{\pi_i} J_{i+1}\right)(\bar{u}) = E\left[\frac{1}{d} f_i\left(\bar{u}, \pi_i(\bar{u}, W_i)\right) + J_{i+1}\left(\pi_i(\bar{u}, W_i)\right)\right],$$

and

$$J_i(\underline{u}) = \left(T^i_{\pi_i} J_{i+1}\right)(\underline{u}) = E\left[\frac{1}{d} f_i\left(\underline{u}, \pi_i(\underline{u}, W_i)\right) + J_{i+1}\left(\pi_i(\underline{u}, W_i)\right)\right].$$

Because $\|f_i\|_\infty \leqslant B$,

$$\left|E\left[f_i\left(\bar{u}, \pi_i(\bar{u}, W_i)\right)\right] - E\left[f_i\left(\underline{u}, \pi_i(\underline{u}, W_i)\right)\right]\right| \leqslant 2B,$$

and therefore,

$$\|J_i\|_s = J_i(\bar{u}) - J_i(\underline{u}) \leqslant \frac{2B}{d} + \left|E\left[J_{i+1}\left(\pi_i(\bar{u}, W_i)\right)\right] - E\left[J_{i+1}\left(\pi_i(\underline{u}, W_i)\right)\right]\right|$$

$$\leqslant \frac{2B}{d} + \left|\sum_{x \in \mathcal{U}}\left(\Pr\{\pi_i(\bar{u}, W_i) = x\} - \Pr\{\pi_i(\underline{u}, W_i) = x\}\right) J_{i+1}(x)\right|.$$

Since $\pi_i \in M^\delta$, there exist $\bar{v}, \underline{v} \in \mathcal{U}$ such that

$$\Pr\{\pi_i(\bar{u}, W_i) = x\} = \begin{cases} 1 - \delta + \dfrac{\delta}{|\mathcal{U}|} & \text{if } x = \bar{v}, \\[2mm] \dfrac{\delta}{|\mathcal{U}|} & \text{otherwise}, \end{cases}$$

and

$$\Pr\{\pi_i(\underline{u}, W_i) = x\} = \begin{cases} 1 - \delta + \dfrac{\delta}{|\mathcal{U}|} & \text{if } x = \underline{v}, \\[2mm] \dfrac{\delta}{|\mathcal{U}|} & \text{otherwise}. \end{cases}$$

Therefore,

$$\sum_{x \in \mathcal{U}} \left(\Pr\{\pi_i(\bar{u}, W_i) = x\} - \Pr\{\pi_i(\underline{u}, W_i) = x\} \right) J_{i+1}(x) = (1-\delta)\left(J_{i+1}(\bar{v}) - J_{i+1}(\underline{v}) \right),$$

which implies that

$$\|J_i\|_s \leqslant \frac{2B}{d} + (1-\delta)\left| J_{i+1}(\bar{v}) - J_{i+1}(\underline{v}) \right| \leqslant \frac{2B}{d} + (1-\delta)\|J_{i+1}\|_s.$$

Since $J_d = 0$, we have

$$\|J_i\|_s \leqslant \frac{2B}{d}\left(1 + (1-\delta) + (1-\delta)^2 + \cdots + (1-\delta)^{d-i-1} \right) \leqslant \frac{2B}{\delta d}. \qquad \square$$

The proof of Lemma 4 is a direct consequence of Lemmas 8 and 9, and it is given below.

Proof. It follows from the definition of J_i and $\tilde{J}_i$ that

$$J_i = T^i \ldots T^{i+r-1} J_{i+r} \quad \text{and} \quad \tilde{J}_i = T^i \ldots T^{i+r-1} \tilde{J}_{i,r},$$

where $\tilde{J}_{i,r} = 0$. By repeated applications of Lemma 8, we have

$$\left\| J_i - \tilde{J}_i \right\|_s \leqslant \left(1 - \frac{\delta}{|\mathcal{U}|} \right)^r \|J_{i+r}\|_s,$$

and the desired conclusion follows from Lemma 9. $\square$

Appendix B. Proof of Lemma 5

The proof Lemma 5 makes use of the following auxiliary results. The first shows that the expectations we are taking induce contractions of the span-seminorm.

Lemma 10. *For $H \in \Re^{|\mathcal{U}|}$ and $u \in \mathcal{U}$, let $Q : M^\delta \to \Re$ be defined by*

$$Q(\theta) = E\left[H\left(\theta(u, W) \right) \right],$$

where the random variable W is uniformly distributed in $[0, 1]$. *Then,*

$$\|Q\|_s = (1-\delta)\|H\|_s \leqslant \|H\|_s.$$

Proof. Let $\bar{u}, \underline{u} \in \mathcal{U}$ be chosen to satisfy

$$H(\bar{u}) = \max_{u \in \mathcal{U}} H(u) \quad \text{and} \quad H(\underline{u}) = \min_{u \in \mathcal{U}} H(u).$$

Choose $\bar{\theta}, \underline{\theta} \in M^\delta$ such that

$$
\Pr\{\bar{\theta}(u, W) = x\} = \begin{cases} 1 - \delta + \dfrac{\delta}{|\mathcal{U}|} & \text{if } x = \bar{u}, \\[2mm] \dfrac{\delta}{|\mathcal{U}|} & \text{otherwise,} \end{cases}
$$

and

$$
\Pr\{\underline{\theta}(u, W) = x\} = \begin{cases} 1 - \delta + \dfrac{\delta}{|\mathcal{U}|} & \text{if } x = \underline{u}, \\[2mm] \dfrac{\delta}{|\mathcal{U}|} & \text{otherwise.} \end{cases}
$$

Then,

$$
\begin{aligned}
\|Q\|_s &= \max_{\theta \in M^\delta} E\big[H\big(\theta(u, W)\big)\big] - \min_{\theta \in M^\delta} E\big[H\big(\theta(u, W)\big)\big] = E\big[H\big(\bar{\theta}(u, W)\big)\big] - E\big[H\big(\underline{\theta}(u, W)\big)\big] \\
&= \sum_{x \in \mathcal{U}} \big(\Pr\{\bar{\theta}(u, W) = x\} - \Pr\{\underline{\theta}(u, W) = x\}\big) H(x) = (1 - \delta)\big(H(\bar{u}) - H(\underline{u})\big) \\
&= (1 - \delta)\|H\|_s. \qquad \square
\end{aligned}
$$

The next lemma establishes a general result about the span-seminorm.

Lemma 11. *Let $\mathcal{S}$ be a finite space. For any $g_1, g_2 : \mathcal{S} \to \Re$, if $x_1^*, x_2^* \in \mathcal{S}$ satisfy*

$$
g_1\big(x_1^*\big) = \min_{x \in \mathcal{S}} g_1(x) \quad \text{and} \quad g_2\big(x_2^*\big) = \min_{x \in \mathcal{S}} g_2(x),
$$

then

$$
0 \leqslant g_1\big(x_2^*\big) - g_1\big(x_1^*\big) \leqslant \|g_1 - g_2\|_s.
$$

Proof. Note that

$$
\begin{aligned}
g_1\big(x_2^*\big) - g_1\big(x_1^*\big) &= g_1\big(x_2^*\big) - g_2\big(x_2^*\big) - g_1\big(x_1^*\big) + g_2\big(x_2^*\big) \\
&= g_1\big(x_2^*\big) - g_2\big(x_2^*\big) - g_1\big(x_1^*\big) + g_2\big(x_1^*\big) - g_2\big(x_1^*\big) + g_2\big(x_2^*\big) \\
&= \big((g_1 - g_2)\big(x_2^*\big) - (g_1 - g_2)\big(x_1^*\big)\big) + g_2\big(x_2^*\big) - g_2\big(x_1^*\big) \\
&\leqslant \|g_1 - g_2\|_s,
\end{aligned}
$$

where the last inequality follows from the definition of $\|\cdot\|_s$ and the fact that

$$
g_2\big(x_2^*\big) = \min_{x \in \mathcal{S}} g_2(x) \leqslant g_2\big(x_1^*\big). \qquad \square
$$

Proof of Lemma 5. Fix $u \in \mathcal{U}$. Let $Q^1, Q^2 : M^\delta \to \Re$ be defined by

$$
Q^1(\theta) = \big(T_\theta^i H_1\big)(u) = E\left[\frac{1}{d} f_i\big(u, \theta(u, W_i)\big) + H_1\big(\theta(u, W_i)\big)\right]
$$

and

$$
Q^2(\theta) = \big(T_\theta^i H_2\big)(u) = E\left[\frac{1}{d} f_i\big(u, \theta(u, W_i)\big) + H_2\big(\theta(u, W_i)\big)\right].
$$

Thus,

$$
\big(Q^1 - Q^2\big)(\theta) = E\big[(H_1 - H_2)\big(\theta(u, W_i)\big)\big],
$$

and it follows from Lemma 10 that

$$\left\| Q^1 - Q^2 \right\|_s \leqslant \|H_1 - H_2\|_s.$$

Since

$$Q^1(\theta_1) = \min_{\theta \in M^\delta} Q^1(\theta) \quad \text{and} \quad Q^2(\theta_2) = \min_{\theta \in M^\delta} Q^2(\theta),$$

it follows from Lemma 11 that

$$0 \leqslant \left(T^i_{\theta_2} H_1\right)(u) - \left(T^i_{\theta_1} H_1\right)(u) = Q^1(\theta_2) - Q^1(\theta_1) \leqslant \left\| Q^1 - Q^2 \right\|_s \leqslant \left\| H_1 - H_2 \right\|_s,$$

and since $u \in \mathcal{U}$ is arbitrary,

$$\left\| T^i_{\theta_1} H_1 - T^i_{\theta_2} H_1 \right\|_\infty \leqslant \|H_1 - H_2\|_s. \qquad \square$$

Appendix C. Proof of Lemma 6

The proof Lemma 6 makes use of the following result.

Lemma 12. *Let $\{\theta_k\}_{k \geqslant 0}$ be a sequence of mappings in M^δ, and let $\{W_k\}_{k \geqslant 0}$ be a sequence of independent random variables, each uniformly distributed in $[0, 1]$. Consider sequences of random variables $\{X_k\}_{k \geqslant 0}$, $\{Y_k\}_{k \geqslant 0}$, and $\{Z_k\}_{k \geqslant 1}$ defined by*

$$X_{k+1} = \theta_k(X_k, W_k), \quad k \geqslant 0,$$

and

$$Y_{k+1} = \theta_k(Y_k, W_k), \qquad Z_{k+2} = \theta_{k+1}(Z_{k+1}, W_{k+1}), \quad k \geqslant 0,$$

where X_0, Y_0, and Z_1 are any random variables in $\mathcal{U}$. Then, for $u, v \in \mathcal{U}$,

$$\left| \Pr\{(X_r, X_{r+1}) = (u, v)\} - \Pr\{(Y_r, Z_{r+1}) = (u, v)\} \right| \leqslant 4(1 - \delta)^r,$$

for all r.

Proof. Let a stopping time N_1 be defined by

$$N_1 = \inf\{k > 0 : Y_k = X_k\}.$$

It follows that on $\{N_1 \leqslant r\}$, $(Y_r, Y_{r+1}) = (X_r, X_{r+1})$ almost surely. Thus, for any $u, v \in \mathcal{U}$,

$$\begin{aligned}
\Pr\{(Y_r, Y_{r+1}) = (u, v)\} &= \Pr\{(Y_r, Y_{r+1}) = (u, v), \; N_1 \leqslant r\} + \Pr\{(Y_r, Y_{r+1}) = (u, v), \; N_1 > r\} \\
&= \Pr\{(X_r, X_{r+1}) = (u, v), \; N_1 \leqslant r\} + \Pr\{(Y_r, Y_{r+1}) = (u, v), \; N_1 > r\}.
\end{aligned}$$

Since

$$\Pr\{(X_r, X_{r+1}) = (u, v), \; N_1 \leqslant r\} = \Pr\{(X_r, X_{r+1}) = (u, v)\} - \Pr\{(X_r, X_{r+1}) = (u, v), \; N_1 > r\},$$

it follows that

$$\left| \Pr\{(X_r, X_{r+1}) = (u, v)\} - \Pr\{(Y_r, Y_{r+1}) = (u, v)\} \right| \leqslant 2 \Pr\{N_1 > r\}.$$

Note that

$$\Pr\{N_1 > r\} = \Pr\{X_1 \neq Y_1, \dots, X_r \neq Y_r\} = \prod_{l=0}^{r-1} \Pr\{X_{l+1} \neq Y_{l+1} \mid X_s \neq Y_s, \; 1 \leqslant s \leqslant l\}.$$

Since $\theta_l \in M^\delta$, it follows that

$$\theta_l(U, W) = \begin{cases} \bar{\theta}_l(U) & \text{if } 0 \leqslant W < 1 - \delta, \\ k\text{th element of } \mathcal{U} & \text{if } 1 - \delta + \dfrac{(k-1)\delta}{|\mathcal{U}|} \leqslant W < 1 - \delta + \dfrac{k\delta}{|\mathcal{U}|}, \end{cases}$$

where $\bar{\theta}_l$ is some function in $\mathcal{U}$. Recall that

$$X_{l+1} = \theta_l(X_l, W_l) \quad \text{and} \quad Y_{l+1} = \theta_l(Y_l, W_l).$$

Thus, if $W_l > 1 - \delta$, then $X_{l+1} = Y_{l+1}$ regardless of the value of X_l and Y_l. Therefore,

$$\Pr\{X_{l+1} = Y_{l+1} \mid X_s \neq Y_s, \ 1 \leqslant s \leqslant l\} \geqslant \delta,$$

which implies that

$$\Pr\{N_1 > r\} = \prod_{l=0}^{r-1} \Pr\{X_{l+1} \neq Y_{l+1} \mid X_s \neq Y_s, \ 1 \leqslant s \leqslant l\} \leqslant (1 - \delta)^r.$$

Hence,

$$\left| \Pr\big\{(X_r, X_{r+1}) = (u, v)\big\} - \Pr\big\{(Y_r, Y_{r+1}) = (u, v)\big\} \right| \leqslant 2(1 - \delta)^r.$$

Now, consider a stopping time N_2 defined by

$$N_2 = \inf\{k > 0 : \ Y_{1+k} = Z_{1+k}\}.$$

Then, on $\{N_2 \leqslant r\}$, $(Y_r, Y_{r+1}) = (Y_r, Z_{r+1})$ almost surely. Using exactly the same argument as above, we find that

$$\left| \Pr\big\{(Y_r, Y_{r+1}) = (u, v)\big\} - \Pr\big\{(Y_r, Z_{r+1}) = (u, v)\big\} \right| \leqslant 2(1 - \delta)^r.$$

An application of a triangle inequality yields

$$\left| \Pr\big\{(X_r, X_{r+1}) = (u, v)\big\} - \Pr\big\{(Y_r, Z_{r+1}) = (u, v)\big\} \right| \leqslant 4(1 - \delta)^r. \qquad \square$$

Proof of Lemma 6. It follows from the definition of $\bar{\phi}^{\delta,r}$ that there are mappings $\tilde{\pi}_0, \ldots, \tilde{\pi}_{d-1} \in M^\delta$ such that

$$\bar{\phi}_{k+1}^{\delta,r}(f) = \tilde{\pi}_k\big(\bar{\phi}_k^{\delta,r}(f), W_k\big), \quad \forall k = 0, \ldots, d-1,$$

where $\bar{\phi}_0^{\delta,r}(f) = \iota$, and $W_0, \ldots, W_{d-1}$ are independent random variables, each uniformly distributed in $[0, 1]$. Under the decentralized stochastic strategy, the decision $\phi_i^{\delta,r}(f) = \phi_{i,0}^{\delta,r}(f)$ is generated from the equation:

$$\phi_{i,l+1}^{\delta,r}(f) = \tilde{\pi}_{i+l}\big(\phi_{i,l}^{\delta,r}(f), W_{i+l}\big), \quad \forall l = -r, \ldots, -1,$$

where $\phi_{i,-r}^{\delta,r}(f) = \iota$. Analogously, the decision $\phi_{i+1}^{\delta,r}(f) = \phi_{i+1,0}^{\delta,r}(f)$ is generated from the equation:

$$\phi_{i+1,l+1}^{\delta,r}(f) = \tilde{\pi}_{i+1+l}\big(\phi_{i+1,l}^{\delta,r}(f), W_{i+1+l}\big), \quad \forall l = -r, \ldots, -1,$$

where $\phi_{i+1,-r}^{\delta,r}(f) = \iota$. Let sequences of random variables $\{X_k\}_{k \geqslant 0}$, $\{Y_k\}_{k \geqslant 0}$, and $\{Z_k\}_{k \geqslant 1}$ be defined by

$$X_k = \bar{\phi}_{i-r+k}^{\delta,r}(f), \quad \forall k \geqslant 0,$$

and

$$Y_k = \phi_{i,-r+k}^{\delta,r}(f), \qquad Z_{k+1} = \phi_{i+1,-r+k}^{\delta,r}(f), \quad \forall k \geqslant 0.$$

Then, for any $x, y \in \mathcal{U}$,

$$\Pr\big\{\big(\bar{\phi}_i^{\delta,r}(f), \bar{\phi}_{i+1}^{\delta,r}(f)\big) = (x, y)\big\} = \Pr\big\{(X_r, X_{r+1}) = (x, y)\big\},$$

and

$$\Pr\big\{\big(\phi_i^{\delta,r}(f), \phi_{i+1}^{\delta,r}(f)\big) = (x, y)\big\} = \Pr\big\{(Y_r, Z_{r+1}) = (x, y)\big\}.$$

The desired result follows directly from Lemma 12 (with $\theta_k = \tilde{\pi}_{i-r+k}$). $\quad \square$

Appendix D. Proof of Lemma 7

It suffices to show that

$$\lim_{r \to \infty} \frac{r^p(1 - a/r^p)^r}{1/r^p} = 0.$$

Since

$$\lim_{x \to 0} \frac{\ln(1+x)}{x} = 1,$$

it follows that

$$1 = \lim_{r \to \infty} \frac{\ln(1 - a/r^p)}{-a/r^p} = \lim_{r \to \infty} \frac{\ln(1 - a/r^p)^r}{-ar^{1-p}}.$$

Thus, there exists N such that whenever $r > N$,

$$\frac{1}{2} \leqslant \frac{\ln(1 - a/r^p)^r}{-ar^{1-p}} \leqslant \frac{3}{2}.$$

Note that

$$\ln\left(\frac{r^p(1 - a/r^p)^r}{1/r^p}\right) = 2p \ln r + \ln\left(1 - \frac{a}{r^p}\right)^r = 2p \ln r - ar^{1-p} \frac{\ln(1 - a/r^p)^r}{-ar^{1-p}}.$$

Thus, for $r > N$,

$$\ln\left(\frac{r^p(1 - a/r^p)^r}{1/r^p}\right) \leqslant 2p \ln r - \frac{a}{2} r^{1-p},$$

which implies that

$$\lim_{r \to \infty} \ln\left(\frac{r^p(1 - a/r^p)^r}{1/r^p}\right) = -\infty.$$

Therefore,

$$\lim_{r \to \infty} \frac{r^p(1 - a/r^p)^r}{1/r^p} = 0.$$

References

Başar, T., 1985. An equilibrium theory for multiperson decision making with multiple probabilistic models. IEEE Trans. Automat. Control 30 (2).

Bertelé, U., Brioschi, F., 1972. Nonserial Dynamic Programming. Academic Press, New York, NY.

Bertsekas, D.P., 1995. Dynamic Programming and Optimal Control. Vols. I and II. Athena Scientific, Belmont, MA.

Billard, E., Pasquale, J., 1995. Adaptive coordination in distributed systems with delayed information. IEEE Trans. Systems Man Cybern. 25 (4), 546–554.

Granot, F., Veinott, A.F. Jr., 1985. Substitutes, complements and ripples in network flows. Math. Oper. Res. 10 (3).

Jensen, F.V., 1996. An Introduction to Bayesian Networks. Springer-Verlag, New York.

Marschak, J., 1955. Elements for a theory of teams. Manage. Sci. 1, 127–137.

Marschak, J., Radner, R., 1972. Economic Theory of Teams. Yale Univ. Press, New Haven, NJ.

Papadimitriou, C.H., Tsitsiklis, J.N., 1986. Intractable problems in control theory. SIAM J. Control Optim. 24 (4).

Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Kaufmann, San Mateo, CA.

Radner, R., 1959. The application of linear programming to team decision problems. Manage. Sci. 5 (2), 143–150.

Radner, R., 1962. Team decision problems. Ann. Math. Statist. 33, 857–881.

Sandell, N.R. Jr., Varaiya, P.P., Athans, M., Safonov, M.G., 1978. Survey of decentralized control methods for large scale systems. IEEE Trans. Automat. Control 23 (2).

Tsitsiklis, J.N., Athans, M., 1985. On the complexity of decentralized decision making and detection problems. IEEE Trans. Automat. Control 30 (5).

von Stengel, B., Koller, D., 1997. Minmax equilibria in team games. Games Econ. Behav. 21, 309–321.