2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Decentralized control of a multi-agent stochastic dynamic resource allocation problem

Huaning Cai and Andrew E.B. Lim

Abstract— **We study a stochastic system consisting of multiple decentralized control agents who allocate shared system resources in response to requests that arrive stochastically over time. Decentralized agents in our system maximize their own objectives subject to potentially mis-specified models of the way in which system resources are consumed by other agents in the system. We introduce the notion of a transfer price, which is a contract that describes the way in which decentralized agents compensate other agents in the system for using shared resources whenever they satisfy a request. We characterize transfer prices under which there is no efficiency loss relative to the optimal system-wide expected revenue when it is managed by a fully-knowledgeable centralized controller. We also show that optimality of these contracts is insensitive to mis-specification by each of the agents of the behavior of other agents in the system.**

I. INTRODUCTION

In this paper we study a system consisting of multiple decentralized control agents. Each decentralized controller can be thought of as a sales agent who receives requests for products under his/her management from customers who arrive stochastically over time. For each request, the selling agent needs to decide whether it should be accepted or rejected. If a request is rejected, the customer departs. If a request is accepted, the agent receives income from the customer in return for the requested product which is manufactured using inventory from a pool of raw materials that are shared by all selling agents in the system. The cost of raw materials is paid for by the sales agent accepting the request, whose accept/reject decision is made by solving a (decentralized) stochastic optimal control problem, where the goal is to maximize income from accepted requests net the cost of raw materials. The optimal behavior of the decentralized agents and the resulting performance of the aggregate system depends on price of these raw materials, which is described by a *transfer price contract*. The goal of this paper is to understand how transfer prices should be chosen (i.e. how raw materials should be priced) so as to maximize the net revenue of the overall system. One example of a system where this problem arises is an airline alliance where multiple airlines (selling agents) share raw materials (single leg flights) which are packaged and sold as itineraries (products) to customers that arrive stochastically over time.

Several lines of related work are important to mention. Firstly, though the topic of resource allocation is classical (see for example Arrow, Hurwitz and Uzawa [8]), our paper is one of the few that addresses this problem in a stochastic and dynamic setting. Our paper is also related to a line of work on decentralized decision making and control dating back to the papers of Marshak [11], Arrow and Hurwicz [2], Radner [13] in the economics literature, as well as Witsenhausen [18], Ho [6] and others [9], [16] in the systems and control literature (see also the survey paper [7], [12], and the recent papers [1], [3], [14], [15]). From a high level, this literature is concerned with systems in which relevant information is not in the hands of a single centralized decision maker and key distinction is what is meant by information. In [7], [9], [14], [16], [18] different agents see different functions of the aggregate system state (or noisy observations of some subset of it) and the primary concern is describing the aggregate system performance when each agent's policy is constrained to being a function of their observation. In contrast, we assume in this paper that each agent observes the full system state (i.e. system resources) but only knows a subset of the system structure/model parameters (for example, each agent only knows the probability that a product in his/her portfolio is requested but not the probability that other agents will receive a request). See also [10] for related results for a continuous time two-agent single server queueing system. Finally, we also mention the work on decentralized control as related to the management of airline alliances (see [4], [17]. We note, however, that [4] is for deterministic systems while [17] analyzes various transfer contracts but do not address the issue of optimality.

The outline of the paper is as follows. We discuss centralized control of the resource allocation problem in Section II. The decentralized problem in formulated in Section III where stochastic models assumped by each agent, the notion of transfer contracts, and the control problem that each agent solves are introduced. Optimal transfer contracts are characterized in Section IV. One interesting and surprising feature of the optimal contracts is that centralized efficiency can be achieved even if decentralized agents mis-specify the behavior of other agents in their models. Decentralized synthesis of optimal contracts through a message passing algorithm is presented in Section V and convergence of this algorithm is discussed in Section VI. An example related to optimal control of airline alliances is presented in Section VI-B.

This research was supported by the NSF CAREER Award DMI-0348746 and the NSF Grants CMMI-0500503 and CMMI-1031637. The opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Support from the Coleman Fung Chair in Financial Modeling is also acknowledged.

Both authors are with the Department of Industrial Engineering and Operations Research, University of California, Berkeley, CA 94729, USA. huaning@ieor.berkeley.edu, lim@ieor.berkeley.edu

Due to limitations of space, proofs of some of the results have been removed. The interested reader is referred to[5].

II. Centralized control

Consider a stochastic system consisting of a set of agents $\mathscr{I}$, a set of resources $\mathscr{L}$, and a set of bundles $\mathscr{J}$. A bundle $j \in \mathscr{J}$ uses one or more resources, denoted by a requirement vector A_j (where $a_{lj} = 1$ if resource l is required, and zero otherwise), and generates revenue r_j. There is a finite horizon, $\mathscr{T} = \{1, ..., T\}$. Demand $\widetilde{d}_{jt}$ for bundle j arrives randomly over time, and at each time t, we assume there is at most one request where q_{jt} is the probability that the request is for bundle j:

$$\mathbb{P}(\widetilde{d}_{jt} = 1, \widetilde{d}_{j't} = 0, \forall j' \neq j) = q_{jt}, \forall j \in \mathscr{J}.$$

In each time period, there are $|\mathscr{J}| + 1$ possible events, including the possibility of no arrival. It follows that $\sum_j q_{jt} \leq 1$.

Suppose there is a single centralized agent who makes the accept or reject decisions for each request. A request for bundle j is accepted if $\mu_{jt} = 1$ and rejected if $\mu_{jt} = 0$. The set of feasible decisions at time t depends on the available remaining capacities x_t,

$$\mathscr{U}(x_t) = \{\mu(t) \in \{0,1\}^{|\mathscr{J}|} : A_j \mu_{jt} \leq x_t\}.$$

The centralized decision maker's objective is to maximize expected revenue subject to resource constraints:

$$(C) \begin{cases} \max \mathbb{E}\left\{ \sum_{t=1}^{T} \sum_{j \in \mathscr{J}} r_j \mu_{jt} \widetilde{d}_{jt} \right\} \\ \text{subject to:} \\ x_{t+1} = x_t - \sum_{j \in \mathscr{J}} A_j \mu_{jt} \widetilde{d}_{jt}, \quad \forall t, \mu_{jt} \in \mathscr{U}(x_t). \end{cases} \quad (1)$$

The dynamic programming equation for (C) is

$$\begin{cases} V(t,x) = \\ \max_{\mu(t) \in \mathscr{U}(x)} \mathbb{E}\Big\{ \sum_{j \in \mathscr{J}} r_j \mu_{jt} \widetilde{d}_{jt} + V(t+1, x_{t+1}) \;\Big|\; x_t = x \Big\}, \\ V(T+1, x) = 0. \end{cases}$$

Defining

$$\Delta V(t+1, x, A_j) \triangleq V(t+1, x) - V(t+1, x - A_j),$$

it is easy to show that the dynamic programming equations can be written as

$$\begin{cases} V(t,x) = \\ \max_{\mu(t)} \sum_{j \in \mathscr{J}} q_{jt} \mu_{jt} \Big[r_j - \Delta V(t+1, x, A_j) \Big] + V(t+1, x), \\ V(T+1, x) = 0. \end{cases} \quad (2)$$

And it follows that the optimal centralized policy is

$$\mu^*_{jt}(x) = \begin{cases} 1 & \text{If } r_j \geq \Delta V(t+1, x, A_j), \\ 0 & \text{otherwise.} \end{cases} \quad (3)$$

In particular a request for bundle j is only accepted if the generated revenue r_j is no less than the future opportunity cost $\Delta V(t+1, x, A_j)$ of the resources that would be consumed.

Observe that in formulating problem (1), it is implicitly assumed that the centralized decision maker knows all relevant system parameters including the demand probability for each bundle. In many applications, knowledgeable decision makers such as this do not exist. Rather, it is more commonly the case that there are many agents where each manages a subset of the bundles, is knowledgable about the arrival statistics of his/her own bundles but not those of others, and is interested in maximizing his/her own revenue. In such situations, it is not possible to formulate the centralized problem (1).

III. Decentralized control

We now consider the resource allocation problem (1) in a decentralized setting. There are several elements that distinguish the decentralized problem from the centralized case. Firstly, the system consists of multiple agents where each agent i is responsible for the accept/reject decisions for a subset of bundles $\mathscr{J}_i$ from the pool $\mathscr{J}$ and knows the arrival probabilities for items in his/her bundle but not necessarily those in the items in bundles managed by other agents. We note, however, that each agent may attempt to model the demand probabilities for other items managed by other agents but may incorrectly specify these probabilities. Secondly, each agent acts in his/her own self interests and maximizes his/her own objective conditional on local and possibly mis-specified model.

In this section, we formulate agent level dynamics and agent level objectives. We also introduce the notion of transfer contracts which define revenue sharing between agents in the system whenever resources are consumed. Intuitively, revenue sharing imposes a charge on agents who use shared resources which compensates other agents for the loss of these resources. We address two questions: (i) do there exist transfer contracts under which decentralized control can achieve centralized efficiency, and (ii) what is the impact of model mis-specification by local agents on the efficiency of the integrated system.

A. *Stochastic model for Agent i*

In this section, we formulate stochastic models for each of the decentralized agents. We assume the following:

- (A1) Agent i makes accept/reject decisions for the subset of bundles $\mathscr{J}_i \subseteq \mathscr{J}$;
- (A2) All agents can observe the remaining inventory at time t, x_t. Inventory is depleted whenever any agent receives and accepts a request.
- (A3) Agent i knows the probability q_{jt} that an item $j \in \mathscr{J}_i$ from his/her bundle will be requested at time t (i.e. $q_{jt} = 1$) but may not know the request probabilities for items that are managed by other agents or their accept/reject policies. Nevertheless, Agent i may attempt to model the demand for items managed by other agents and their accept/reject policies, but both may be mis-specified. Specifically, if $\mathscr{J}_{i'}$ is the set of bundles managed by

Agent i' ($i' \neq i$), then Agent i may specify the probability q'_{jt} that Agent i' receives a request (i.e. $\widetilde{d}'_{jt} = 1$ for items $j \in \mathscr{J}_{i'}$) and the accept/reject policy μ'_{jt} of Agent i'. However, both q'_{jt} and μ'_{jt} may be mis-specified in that they differ from the true demand probabilities and the actual accept/reject policy which are only known accurately by Agent i'.

Under these assumptions, each agent has a model in which the accept/reject decisions for bundles in his/her portfolio is the decision variable, the probability that he/she receives a request for an item in his/her bundle is accurately specified, but the probability that other agents receive requests and their accept/reject decisions are potentially mis-specified. It follows that Agent i adopts a demand model

$$x_{t+1} = x_t - \sum_{j \in \mathscr{J}_i} A_j \mu_{jt} \widetilde{d}_{jt} - \sum_{i' \neq i} \sum_{j \in \mathscr{J}_{i'}} A_j \mu'_{jt} \widetilde{d}'_{jt},\ t = 1, \cdots, T \quad (4)$$

B. Transfer contracts and revenue sharing

In this section, we introduce the notion of revenue sharing and transfer contracts, which play an important role in coordinating the decentralized system.

A transfer contract is a set of functions $R_{i,i'}(t,x,A_j)$ defined for every pair of agents i and i', time t, inventory level x, and bundle resources A_j. The function $R_{i,i'}(t,\ x,\ A_j)$ specifies the revenue that is transferred by Agent i to i' if Agent i receives a request for resources A_j at time t when the inventory level $x_t = x$ and accepts the request. That is, selling item $j \in \mathscr{J}_i$ costs Agent i

$$\sum_{i' \neq i} R_{i,i'}(t,x,A_j)$$

of which the quantity $R_{i,i'}(t,x,A_j)$ is sent to Agent i'. Conversely, the sale of item $j \in \mathscr{J}_{i'}$ by Agent i' ($i' \neq i$) brings revenue $R_{i',i}(t,x,A_j)$ to Agent i. Transfer contracts $R_{i,i'}(t,x,A_j)$ define a charge that is levied on Agent i by Agent i' for using the quantity A_j of shared inventory. The net impact of transfer contracts is to modify the objective functions and influence the accept/reject policies of each of the agents who are now subject to additional costs and income sources through revenue sharing.

C. Decentralized control

For a given set of revenue transfer contracts, Agent i solves

$$(C_i) \begin{cases} \max\limits_{u_i(t)} \mathbb{E} \sum\limits_{t=1}^{T} \Big\{ \sum\limits_{j \in \mathscr{J}_i} \Big[r_j - \sum\limits_{i' \neq i} R_{i,i'}(t,x,A_j) \Big] \mu_{jt} \widetilde{d}_{jt} \\ + \sum\limits_{i' \neq i} \sum\limits_{j \in \mathscr{J}_{i'}} R_{i',i}(t,x,A_j) \mu'_{jt} \widetilde{d}'_{jt} \Big\} \\ \text{subject to:} \\ x_{t+1} = x_t - \sum_{j \in \mathscr{J}_i} A_j \mu_{jt} \widetilde{d}_{jt} - \sum_{i' \neq i} \sum_{j \in \mathscr{J}_{i'}} A_j \mu'_{jt} \widetilde{d}'_{jt}, \\ \mu_{jt} \in \mathscr{U}(x). \end{cases} \quad (5)$$

The objective function consists of two terms. The quantity

$$r_j - \sum_{i' \neq i} R_{i,i'}(t,x,A_j)$$

is the net revenue received by Agent i when he/she accepts a request for one unit of bundle $j \in \mathscr{J}_i$. It consists of income r_j from the sale net the cash transfers to the other agents. The probability of this event $\mu_{jt} q_{jt}$ is assumed to be accurately specified for items j in the bundle $\mathscr{J}_i$ that it is managing. In the second term, $R_{i',i}(t,x,A_j)$ is the income that is transferred to Agent i from Agent i', as defined by the revenue sharing rule, whenever Agent i' makes a sale. From the perspective of Agent i, the probability that this event occurs, $\mu'_{jt} q'_{jt}$, may be erroneously specified.

Defining

$$\Delta V_i(t+1,x,A_j) \triangleq V_i(t+1,x) - V_i(t+1,x-A_j),$$

we can write the dynamic programming equation for Agent i as,

$$\begin{cases} V_i(t,x;R) \\ = \max\limits_{u_i(t)} \sum\limits_{j \in \mathscr{J}_i} q_{jt} \mu_{jt} \Big[r_j - \sum\limits_{i' \neq i} R_{i,i'}(t,x,A_j) \\ - \Delta V_i(t+1,x,A_j) \Big] + \sum\limits_{i' \neq i} \sum\limits_{j \in \mathscr{J}_{i'}} q'_{jt} \mu'_{jt} \Big[R_{i',i}(t,x,A_j) \\ - \Delta V_i(t+1,x,A_j) \Big] + V_i(t+1,x), \\ V_i(T+1,x;R) = 0. \end{cases} \quad (6)$$

As in Agent i's original problem formulation (5) the probability $q'_{jt} \mu'_{jt}$ that another Agent i' uses the quantity A_j of shared resources by allocating a bundle $j \in \mathscr{J}_{i'}$ it is managing may be mis-specified.

D. Weak duality

Given a contract R, each agent formulates a potentially mis-specified stochastic control problem (5) for an optimal accept/reject policy $\mu_i^*(t) = \{\mu_{jt}^*(t), \forall j \in \mathscr{J}_i\}$. The integrated system is operated under these decentralized policies. Observing that revenue transfers are between agents remain internal to the system, it follows that the expected system revenue under optimal decentralized policies for any arbitrary contract is dominated by the optimal revenue generated by the centralized decision maker (1). More formally, if $\mu_i^*(t)$ is the optimal policy for Agent i under some revenue transfer contract R and let

$$\begin{cases} \hat{V}(t,x;R) \triangleq \mathbb{E} \sum\limits_{s=t}^{T} \sum\limits_{j \in \mathscr{J}} r_j \mu_{js}^* \widetilde{d}_{js} \\ \text{subject to:} \\ x_{s+1} = x_s - \sum\limits_{j \in \mathscr{J}} A_j \mu_{js}^* \widetilde{d}_{jt}, \quad \forall\, t \leq s \leq T, \\ x_t = x \end{cases}$$

denotes the system revenue under decentralized policies, then the following relationship between the expected revenue for the system under decentralized control and the value function for the centralized agent holds.

Proposition 3.1 (Weak Duality): Let R denote an arbitrary transfer contract and $\mu_i^*(t)$ denote the optimal policy for Agent i ($i \in \mathscr{I}$) under R.Then

$$\hat{V}(t,x;R) \leq V(t,x), \quad \forall t,x. \quad (7)$$

where $V(t,x)$ is the value function for the centralized agent (1).

Two questions are immediate:

1) Can contracts R be found under which decentralized agents achieve centralized efficiency, and
2) What is the impact of model mis-specification?

IV. Optimal transfer contracts

We begin by deriving conditions using heuristic arguments that optimal contracts should satisfy. Existence of contracts satisfying these conditions and the optimality of these contracts will then be established.

A. Optimality conditions: Conjecture

Let R be an arbitrary sharing contract and

$$V_1(t,x;R),\cdots,V_{\mathscr{I}}(t,x;R) \tag{8}$$

denote the value functions for each agent obtained by solving the decentralized problems (5) under R, namely $V_i(t,x;R)$ is Agent i's valuation of the shared inventory x at time t.

The contract R defines cash transfers that take place whenever inventory is consumed by one of the agents. In particular, we see from (5) that when bundle $j \in \mathscr{J}_i$ is allocated by Agent i and inventory A_j is consumed, that revenue $R_{i,i'}(t,x;A_j)$ is transferred from Agent i to Agent i'. Intuitively, Agent i compensates Agent i' the amount $R_{i,i'}(t,x;A_j)$ for consuming A_j of the shared resources.

In this light, it is therefore natural to expect that optimal contracts R should be such that the compensation received by Agent i' matches its valuation of the resources that were just consumed

$$R_{i,i'}(t,x,A_j) \triangleq V_{i'}(t+1,x;R) - V_{i'}(t+1,x-A_j;R). \tag{9}$$

It is natural to conjecture that optimal contracts R should satisfy the coupled nonlinear system of equations (8)-(9).

Several issues need to be resolved. Firstly, there is the question of existence. It is by no means clear that a contract R can be found for which the system (8)-(9) is satisfied, and this needs to be established for any further discussion to make sense. Secondly, if existence can be shown, establishing optimality could be slippery because agent dynamics in each of the decentralized problems (5) may be mis-specified, and mis-specification may adversely affect the efficiency of decentralized policies. Finally, there is the question of computing the optimal contracts. Ideally, we would like to construct optimal contracts without decentralized agents having to share information about their own models (e.g. request probabilities) and without having to solve a centralized problem.

B. Solution of conjectured equations

Let

$$\begin{aligned} R_{i,i'}(t,x,A_j) &\triangleq V_{i'}(t+1,x) - V_{i'}(t+1,x-A_j) \\ &= \Delta V_{i'}(t,x,A_j), \end{aligned} \tag{10}$$

where $V_i(t,x)$ is the solution of the recursive equations

$$\begin{cases} V_i(t,x) = \\ \max_{\mu_i} \sum_{j\in\mathscr{J}_i} q_{jt}\mu_{jt}\Big[r_j - \Delta V(t+1,x,A_j)\Big] + V_i(t+1,x), \\ V_i(T+1,x) = 0. \end{cases} \tag{11}$$

Observe that $V(t,x)$ on the RHS of (11) is the value function of the centralized problem (1). Note that the system (10)-(11) gives an explicit construction of the contract R; the recursions (11) are solved for $V_i(t,x)$ from which the contracts (10) are constructed. In this section, we show that the contract (10)-(11) satisfies the conditions (8)-(9).

The following preliminary result is required.

Proposition 4.1: Let $V(t,x)$ denote the value function for the centralized problem (1) and $V_1(t,x),\cdots,V_{|\mathscr{I}|}(t,x)$ be defined in (11). Then

$$V(t,x) = \sum_{i\in\mathscr{I}} V_i(t,x).$$

We skip the detailed proof. The main idea is to exploit the separable structure of the objective function, and apply backward induction argument (see [5]).

We show that (10) is a solution of the system of equations (8)-(9) by showing that the solution $V_i(t,x)$ of (11) is the value function for the decentralized problem (5) under the contract (10)-(11). To see this, observe (by Proposition 4.1) that

$$\Delta V(t+1,x,A_j) = \sum_{i'\neq i} \Delta V_{i'}(t,x,A_j) + \Delta V_i(t+1,x,A_j).$$

This implies that (11) is equivalent to

$$\begin{aligned} V_i(t,x) = \max_{\mu_i} \sum_{j\in\mathscr{J}_i} q_{jt}\mu_{jt}\Big[r_j - \sum_{i'\neq i}\Delta V_{i'}(t,x,A_j) \\ -\Delta V_i(t+1,x,A_j)\Big] + V_i(t+1,x). \end{aligned}$$

When the contract R is given by (10)-(11), this equation can be written as

$$\begin{cases} V_i(t,x) = \max_{u_i} \sum_{j\in\mathscr{J}_i} q_{jt}\mu_{jt}\Big[r_j - \sum_{i'\neq i} R_{i,i'}(t,x,A_j) \\ -\Delta V_i(t+1,x,A_j)\Big] + \sum_{i'\neq i}\sum_{j\in\mathscr{J}_{i'}} q'_{jt}\mu'_{jt}\Big[R_{i',i}(t,x,A_j) \\ -\Delta V_i(t+1,x,A_j)\Big] + V_i(t+1,x), \\ V_i(T+1,x) = 0. \end{cases} \tag{12}$$

Observing that this is nothing but the dynamic programming equation for the decentralized problem (5) under contract (10)-(11), we can now say the following:

- The solution $V_i(t,x)$ of (11) equals the value function for the decentralized problem for Agent i under the contract (10)-(11). It follows that the contract (10)-(11) is a solution of the system of equations (8)-(9).
- The maximizer in the RHS of (11) is also the maximizer in the RHS of the dynamic programming equation (12) under contract (10)-(11). It follows that the maximizer

in (11) defines the optimal accept/reject decision for Agent i under the contract (10)-(11).

- Under transfer contract (10)-(11), the second term in (12) is always zero. It follows that the value function of the decentralized problem (5) as well as the associated optimal accept/reject policy do not depend on Agent i's specification of the demand probabilities q'_{jt} or the accept/reject decisions μ'_{jt} of Agents $i' \neq i$.

We summarize these observations as follows.

Proposition 4.2: Let the transfer contract R be defined by (10)-(11). Then R is a solution of the system of equations (8)-(9). Under this contract, the value function of Agent i's problem (5) is also the solution of (11) and the optimal accept/reject policy for Agent i is

$$\begin{aligned}\mu^*_{jt} &= \begin{cases} 1, & \text{if } r_j \geq \Delta V(t+1,x,A_j), \\ 0, & \text{otherwise.}\end{cases} \\ &= \begin{cases} 1, & \text{if } r_j \geq \sum_{i' \neq i} R_{i,i'}(t,x,A_j) + \Delta V_i(t+1,x,A_j), \\ 0, & \text{otherwise.}\end{cases}\end{aligned}$$

Both the value function $V_i(t,x)$ and the optimal policy $\mu^*_i(t)$ are independent of Agent i's specification of the request probabilities q'_{jt} and the accept/reject policies μ'_{jt} of other agents in his/her model (5).

C. Verification of optimality

Proposition 4.2 tells us that the contract defined in (10)-(11) satisfies the conditions (8)-(9), and that the value functions of the decentralized agents under this contract are independent of mis-specification of the request probabilities and accept/reject policies of other agents. In this section, we show that these contracts are optimal in that they achieve equality in (7), and that the decentralized accept/reject policies are optimal for the integrated system.

Theorem 4.1 (Strong duality): Let R and $V_1(t,x)$, $\cdots$, $V_{|\mathscr{I}|}(t,x)$ denote the solution of the system (10)-(11) and $\mu^*_i(t)$ denote the optimal accept/reject policy for Agent i's problem (5) under this contract. Then

1) R solves the system of equations (8)-(9) and $V_i(t,x)$ equals the value function for Agent i's problem (5) under this contract.
2) The collection of decentralized policies $\{\mu^*_1(t),\cdots,\mu^*_{|\mathscr{I}|}(t)\}$ under contract (10)-(11) is optimal for the centralized problem (1) and the system profit under these decentralized policies equals the optimal profit for the centralized agent;
3) The value function for Agent i and his/her associated optimal policy $\mu^*_i(t)$ are independent of his/her assumptions of the demand probabilities and accept/reject policies of the other agents $i' \neq i$ in the decentralized model (5).

The interested reader can refer to [5] for a detailed proof.

We have already noted that the transfer contract R specifies a price for the shared resource as a function of time and the current inventory level, which relates them to Lagrange multipliers/shadow prices in classical resource allocation problems. From this perspective it is notable that existence of an optimal contract and strong duality does not require convexity in the original problem. (Indeed, our proof of existence is constructive does not use convex analysis but works with the dynamic programming equations associated with the problem).

V. Decentralized synthesis of transfer contracts

Theorem 4.1 characterizes transfer contracts under which decentralized agents optimally choose the centrally optimal policies. We now turn to the question of computing these contracts. One approach is to solve the system of equations (10)-(11) directly, but this can only be done if the demand probabilities q_{jt} of all agents and the value function of the centralized problem $V(t,x)$ are known. This is problematic because agents with sufficient information to solve the centralized problem typically do not exist (this was one of the motivations for studying the decentralized problem). Alternatively, the presence of an agent with the information and capability to solve for the optimal contracts through (10)-(11) directly would actually render decentralized control, and our need to compute the optimal transfer contract, unnecessary. So we refine our question as follows: can the optimal contract be found without an all knowing/centralized agent having to solve the centralized problem, and without decentralized agents having to reveal private model information to other agents?

The following algorithm is motivated by these considerations and can be viewed as an iterative approach to solving the implicit system of equations (8)-(9). (Recall from Theorem 4.1 that the optimal contract is a solution of this system). In each iteration, decentralized agents solve their optimization problems (5) conditional on some suboptimal transfer contract. These contracts are then updated locally by each of the agents and exchanged, and the process repeats. While it is natural to ask whether the algorithm converges and whether the limiting contract is the optimal contract described in Theorem 4.1, which we address in the next section, it is important to recognize that the algorithm can be implemented without an all knowing centralized agent and does not require individual agents to exchange private information about request probabilities; all that is exchanged are updated transfer contracts and the algorithm allows for the possibility that the decentralized problems may be misspecified.

Algorithm 5.1:

- Initialize: Set k = 1, and $R^1_{i,i'}(t,x,A_j) = 0$.
- Step 1: Given k, and $R^k_{i,i'}(t,x,A_j)$
  - Each agent solves his/her own decentralized problem (5) and computes his/her value function $V^k_i(t,x) = V_i(t,x;R^k)$ by solving (6);
  - Stop if a satisfactory level of precision has been reached,

$$\sup_{i,t,x} \left| V^k_i(t,x) - V^{k-1}_i(t,x) \right| \leq \varepsilon;$$

otherwise, each agent updates his/her transfer contract,

$$R_{i',i}^{k+1}(t,x,A_j) = V_i^k(t+1,x) - V_i^k(t+1,x-A_j).$$

  - Each agent communicate the updated transfer contract $R_{i',i}^{k+1}(t,x,A_j)$ to other agents.
- Step 2: Set k to $k+1$ and return to Step 1.

We now address the issue of convergence.

VI. Convergence of decentralized algorithm

Transfer contracts R specify prices that are paid by agents when consuming shared resources and as such are related to Lagrange multipliers/shadow prices in classical resource allocation problems. In this light, Algorithm 5.1 is analogous to a dual type approach for updating prices. We now present results that guarantee convergence of the Algorithm 5.1 to the optimal contract. As in Theorem 4.1 convexity is not required to guarantee convergence, and convergence is robust to agent-level mis-specification of the demand probabilities of other agents.

A. Statement of main results

Theorem 6.1: Let $V_i(t,x)$ be the value function for Agent i's decentralized problem (6), given the optimal contract (10), and the sequence of $\{V_i^k(t,x)\}$ be computed by Algorithm 5.1. Then $\{V_i^k(t,x)\}$ converges strongly to $V_i(t,x)$, namely

$$\lim_{k\to\infty} \|V_i^k(t,x) - V_i(t,x)\| = 0, \quad \forall i \in \mathscr{I}.$$

Suppose $\mu_i^k(t)$ is the optimal policy for Agent i obtained in the k^{th} iteration, and let

$$\begin{cases} \hat{V}^k(t,x) \triangleq \mathbb{E}\sum_{s=t}^{T}\sum_{i\in\mathscr{I}}\sum_{j\in\mathscr{J}_i} r_j \mu_{js}^k \widetilde{d}_{js} \\ \text{subject to:} \\ x_{s+1} = x_s - \sum_{i\in\mathscr{I}}\sum_{j\in\mathscr{J}_i} A_j \mu_{js}^k \widetilde{d}_{js}, \quad \forall\, t \le s \le T, \\ x_t = x \end{cases} \tag{13}$$

denotes the system profit under the joint policy $\{\mu_i^k(t), i \in \mathscr{I}\}$ obtained from the k^{th} iteration of Algorithm 5.1. The following result guarantees convergence of $\hat{V}^k(t,x)$ to the value function $V(t,x)$ of the centralized agent.

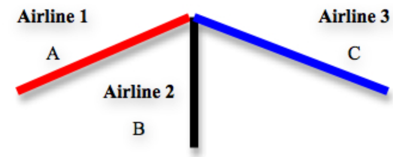
Theorem 6.2: There exist constants $C, M > 0$ such that

$$\|\hat{V}^k(t,x) - V(t,x)\| \le \frac{C(MT)^{k+1}}{M(k+1)!}, \forall k \ge 1.$$

It follows that

$$\lim_{k\to\infty} \|\hat{V}^k(t,x) - V(t,x)\| = 0.$$

The basic idea of the proof is to show that the iteration in 5.1 is a contraction mapping. Due to limitation of space, we skip the detailed proof here, and interested reader can refer to [5].



Airline	Leg (capacity)	Itinerary (fare)
1	A (10)	A ($250), AB ($400)
2	B (10)	B ($250), BC ($400)
3	C (10)	C ($250)

Fig. 1. A three-airline alliance (with three legs and five itineraries)

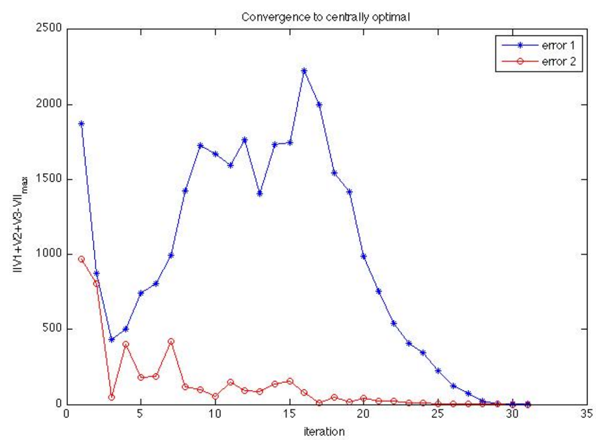


Fig. 2. Convergence of Algorithm 5.1

B. Numerical example

We now illustrate our solution approach with a concrete example from airline alliance revenue management. Figure 1 shows an airline alliance consisting of three agents (airline 1/2/3). The alliance has three resources (flight-leg A/B/C), and markets five bundles (itinerary A/B/C/AB/BC). The arrival probability for itineraries are set such that the network has an overall load factor of 1.33, where the load factor α is defined as the ratio between expected arrival and the initial inventory, $\alpha = \frac{\sum_{ljt} q_{jt} a_{lj}}{\sum_l x_{l1}}$. Lastly the planning horizon has $T = 30$ periods.

Convergence

Figure 2 shows the convergence of Algorithm 5.1, both in terms of error 1 that is evaluated under sup-norm over the entire state space, as well as error 2 that is evaluated only at the initial time and state $(t,x) = (1, 10)$. Figure 3 shows the decomposition of the the centrally optimal value functions into three individual airlines' value functions (note that the value functions decrease monotonically over time).

Transfer contracts and the impact of network topology

Note that our example has a special network structure, such that some airlines are directly connected (e.g. airline-1 and -2), while some are only indirectly connected (e.g. airline-1

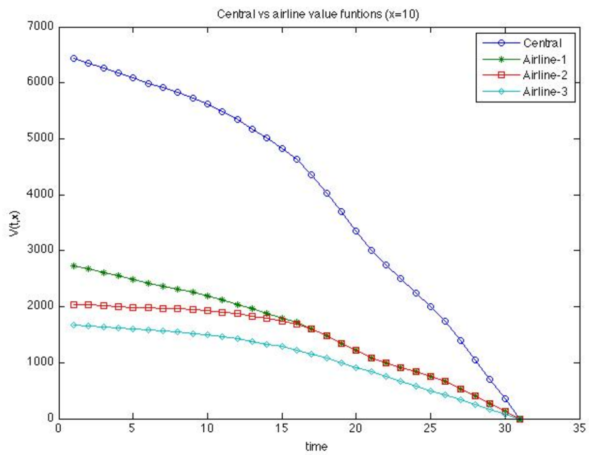


Fig. 3. Central $V(t,x)$ vs individual airlines $V_1(t,x), V_2(t,x), V_3(t,x)$

and -3 via -2). We would like to examine the dependence of the revenue transfers on the underlying network topology. We define a first-order transfer as a payment to a directly connected partner (e.g. airline-1 to -2), and a second-order transfer as a payment to an indirectly connected partner (e.g. airline-1 to -3).

Table I shows the break-down of the net revenue transfer R into first and second order transfers from a simulated sample path of the demand and revenue realization. For example, when airline-3 makes the first sale, \$32 is transferred to airline-2 (first order), while -\$15 is transferred to airline-1 (second order). The most surprising feature is that revenue transfers can take on negative values, meaning that having less inventory can be sometimes beneficial for some airline. For example, airline-1 is always willing to subsidize the sale by taking a negative revenue transfer whenever airline-3 makes a sale (which is second order since airline-1 and -3 are only indirectly connected). The reason is that when airline-3 makes a sale of itinerary C, airline-2 will have less opportunity to sell itinerary BC, thus increasing the opportunity for airline-1 to sell itinerary AB. Clearly, network topology can have a strong impact on the distribution of revenue transfer, and one should exploit such structural properties in designing practical transfer schemes.

Selling Airline	R	1st order R	2nd order R
3	17	32	-15
1	-13	-18	5
2	22	68	-46
3	41	64	-23
3	45	64	-19
2	237	54	183
1	93	120	-27

TABLE I

DECOMPOSED REVENUE TRANSFER FROM A SIMULATED SAMPLE PATH

VII. CONCLUSION

We study a general class of stochastic dynamic resource allocation problems in the decentralized setting, and we characterize a mechanism that is able to coordinate local agents in making centrally optimal decisions. The coordination mechanism consists of a set of transfer contracts that specify prices that each agent need to compensate other agents for the shared resources being consumed. We further derive an iterative algorithm that shows how these transfer contracts can be computed without having agents to reveal their private demand functions and that convergence can be guaranteed without assumptions of convexity.

REFERENCES

[1] D. Adelman and A.J. Mersereau. Relaxations of weakly coupled stochastic dynamic programs. *Operations Research*, 56(3):712–727, 2008.

[2] K. J. Arrow and L. Hurwicz, editors. *Studies in Resource Allocation*. Cambridge University Press, Cambridge, UK, 1977.

[3] D.P. Bersekas. Separable dynamic programming and approximate decomposition methods. *IEEE Transactions on Automatic Control*, 52(2):911–916, 2007.

[4] A. Boyd. Airline alliance. *OR/MS Today*, 1998.

[5] H. Cai and A.E.B. Lim. Decentralized control of stochasic dynamic resource allocation problems. *Working paper*, 2011.

[6] Y.C. Ho. Team decision theory and information structures. *Proceedings of the IEEE*, 68(6):644–654, 1980.

[7] Y.C. Ho. Review of the witsenhausen problem. *Proc.* 47^{th} *IEEE Conference on Decision and Control*, pages 1611–1613, 2009.

[8] L. Hurwicz K. J. Arrow and H. Uzawa, editors. *Studies in Linear and Nonlinear Programing*. Stanford University Press, 1958.

[9] B. Kurtaran. Dynamic two person two objective control problems with delayed sharing information pattern. *IEEE Transactions on Automatic Control*, 22:659–661, 1977.

[10] P. Li, A.E.B. Lim, and J.G. Shanthikumar. Decentralized control of a stochastic pricing-service system. *Under review*, 2010.

[11] J. Marshak. Elements for a theory of teams. *Management Science*, 1:127–137, 1955.

[12] M. Athans N.R. Sandell, P. Varaiya and M.G. Safonov. Survey of decentralized control methods for large scale systems. *IEEE Transactions on Automatic Control*, 23:108–128, 1978.

[13] R. Radner. Team decision problems. *Annals of Mathematical Statistics*, 33(3):857–881, 1962.

[14] A. Rantzer. Dynamic dual decomposition for distributed control. *Working paper, Lund University*, 2009.

[15] M. Rotkowitz and S. Lall. A characterization of convex problems in decentralized control. *IEEE Transactions on Automatic Control*, 51(2):274–286, 2006.

[16] N.R. Sandell and M. Athans. Solution of some nonclassical lqg stochastic decision problems. *IEEE Transactions on Automatic Control*, 19:108–116, 1974.

[17] R.A. Shumsky, C.P. Wright, and H. Groenevelt. Dynamic revenue management in agent alliances. *Transportation Science*, 44(1):15–37, 2010.

[18] H.S. Witsenhausen. A counterexample in stochastic optimum control. *SIAM Journal on Control*, 6(1):138–147, 1968.