# A Dynamic Principal-Agent Model with Hidden Information: Sequential Optimality through Truthful State Revelation[1]

**Hao Zhang**

*Marshall School of Business, University of Southern California, Los Angeles, CA 90089*

**Stefanos Zenios**

*Graduate School of Business, Stanford University, Stanford, CA 94305*

**Abstract**

This paper proposes a general framework for a large class of multi-period principal-agent problems. In this framework, a principal has a primary stake in the performance of a system, but delegates its control to an agent. The underlying system is a Markov Decision Process where the state of the system can only be observed by the agent but the agent's action is observed by both parties. The paper develops a dynamic programming algorithm to derive optimal long-term contracts for the principal. The principal indirectly controls the underlying system by offering the agent a menu of continuation utility vectors along public information paths; the agent's best response, expressed in his choice of continuation utilities, induces truthful state revelation and results in actions that maximize the principal's expected payoff. This problem is significant to the Operations Research community for it can be framed as the problem of optimally designing the reward structure of a Markov Decision Process with hidden states, and has many applications of interest as discussed in the paper.

**August 2004, Revised January 2007**

[1]This research was supported by NSF grant SBER-9982446. Hao Zhang wishes to express his gratitude to Mahesh Nagarajan for reading through various drafts of the paper with many helpful comments. Stefanos Zenios wishes to thank the Decision Sciences group at the London Business School for its hospitality. We are also indebted to four referees and the associate editor for their invaluable suggestions.

# 1 Introduction

Many interesting managerial problems involve two decision makers in a complex dynamic environment coupled with information asymmetry. In medical treatment, the disease state is private information observable by the physician but not by the medical insurer paying the physician for his services. In supply chain management, inventory levels of a downstream firm are often not observed by an upstream firm which can potentially implement a dynamic pricing mechanism. An entrepreneur is better informed about the state of new product development than the investor providing the funds for development. Similarly, in drug discovery the scientists are better informed about the status of the experimentation than the strategic planner determining the resources that will be allocated to pursue the experiments. A human resource manager is less well informed about the state of the employees' knowledge than the employees themselves. In all these examples, the party who observes the private information can also influence its evolution through its choice of actions. These actions can be observed by the other party. This paper proposes a general framework for analyzing this type of problems and develops a general solution. It is not only of theoretical interest but can also shed light on significant managerial problems.

**Background of the Topic.** When two parties engage in a business relationship, their interests are usually not perfectly aligned and information asymmetry can further exacerbate the tension between them. The principal-agent model is a stylish framework for studying such a problem, which has been extensively investigated by economists since 1970s, for example Holmstrom (1979). The party who has the bargaining power to design the contract terms is referred to as "the principal" ("she" in this paper) and the other party "the agent" ("he"). It is usually assumed that the agent possesses private information unobservable by the principal. There are two basic classes of models: *hidden action* (or *moral hazard*) and *hidden information* (or *adverse selection*). In the former, the two parties have access to the same information initially but the agent's response cannot be observed by the principal; in the latter, the agent has private information to begin with but his action can be observed. In both cases, the principal faces the problem of providing incentives for the agent to take the desired action, but the solution methods turn out to be drastically different.

Single-period principal-agent models and multi-period models with simple information structures are by far the most studied, discussed in many textbooks, such as Bolton and Dewatripont (2005), Fudenberg and Tirole (1991), Mas-Colell, Whinston and Green (1995) and Salanié (1997). Multi-period models with dynamic information structures are less understood. Fudenberg, Holmstrom, and Milgrom (1990) is one of the first papers to study a dynamic principal-agent model with an underlying stochastic process. Their model assumes hidden actions and they propose a set of assumptions to break optimal long-term contracts into easily computable short-term ones. Plambeck and Zenios (2000) streamlines the model with an underlying Markov Decision Process and provides a dynamic programming solution.

This paper also studies a model with underlying Markov Decision Process as the Plambeck and Zenios model, but with different cost/reward structure and with hidden information (states) instead of hidden actions. The model captures the basic features of the examples presented at the beginning of the introduction. In this framework, a principal has a primary stake in the performance of a Markov Decision Process, but delegates its control to an agent. She must design a mechanism to reimburse the agent for his efforts, but is constrained by the information that is publicly available. As the principal cannot observe the state of the system, she may pay the agent contingent on the observable action history. However, as we will see in the paper, this is not the best solution.

The topic of this paper can also be viewed from another perspective. Markov Decision Processes provide one of the fundamental decision models in Operations Research. In these processes, a system changes from state to state as time passes and the transition at the end of a period depends on the action chosen in that period. While this model is mature with well developed theories, as in Puterman (1994), it is centered on a single decision maker, and the reward structure which drives the optimal strategy is exogenously given. In many applications of interest, the system's reward structure is designed by an intelligent party that has a stake in the performance of that system and there often exists information asymmetry.

**More Literature Review.** As mentioned earlier, there is vast literature on principal-agent problems. However, only a limited part of it is dedicated to truly dynamic models. In the passages below, we briefly review the existing literature related to our work.

An influential paper that tackles a multi-period game with asymmetric information is Abreu, Pearce, and Stacchetti (1990) (*hereafter APS*). The methodology developed in APS provides some important components behind our approach and is briefly reviewed here. Their paper studies a repeated $N$-player game: in a single period, each player takes a private action, these actions together generate a public random signal, and the players' payoffs depend on their own action and the signal; the single-period game is repeated infinitely and payoffs are discounted over time. It is shown that any profile of the players' total payoffs in equilibrium can be constructed from a profile of their first period actions and a profile of their future (continuation) payoffs contingent on the first period signal. Furthermore, the set of total-payoff profiles is equal to the set of continuation-payoff profiles, which can be found by the value-iteration method in dynamic programming via a convergent sequence of payoff sets.

In a fashion analogous to APS, our approach will focus on the set of continuation payoffs for the principal and agent in each period and state, and will develop the optimal contract through a dynamic programming algorithm applied to the set of continuation payoffs. Three other papers exemplify this approach: Fernandes and Phelan (2000) considers an infinite-horizon consumption model where a risk-averse agent owns a private dynamic endowment process and the principal, a central planner, provides monetary transfers to the agent. The principal tries to minimize the expected total transfers while maintaining the agent's expected utility at a given level. Doepke and Townsend (2005) studies a Markov decision process with both hidden states and hidden actions. The analysis is more complex than that of Fernandes and Phelan because of the combination of hidden states and hidden actions, but the approach to address hidden states is similar. The exact solution is complicated and they resort to an approximation algorithm with finitely many possible continuation utility values. Cole and Kocherlakota (2001) extends the APS model to include hidden states as well. At the end of each period, the players update their beliefs about future states of other players according to a public signal. An important assumption in the paper is that the belief of each player only depends on his current state and not any private information in the past.

Some well-known models can be considered as special cases of a dynamic principal-agent model, such as multi-period models with independent or constant states, and models with special parameter structures so that it is sufficient to consider local incentive constraints. Baron and Besanko (1984) provides an early example of the latter, and recently Battaglini (2005) solves a two-state-infinite-horizon model with special parameter structure. Zhang (2004) shows that these special cases can be derived from the general framework proposed in this paper.

There has been a growing body of literature on information asymmetry in the management science and operations research field, especially in supply chain management. But the published works are either in the single-period setting or in the multi-period setting with special information structures. Sample works are Corbett (2001), Ha (2001) and Ozer and Wei (2006). Cachon (2003) and Chen (2003) provide a review of this literature. Among the few papers addressing dynamic models, Ding, Jia and Tang (2003) considers a principal-agent model with hidden states and hidden actions and use Markov Decision Process techniques to find a best stationary contract.

**Characteristics of the Paper.** Our model integrates the physical structure of Markov Decision Processes with the information structure of the principal-agent paradigm. One challenge of the dynamic hidden-information principal-agent problem is the immense space of possible mechanisms that can be designed by the principal. We show a dynamic revelation principle: it is sufficient to consider revelation contracts. This type of contract asks the agent to report the hidden states, and his actions and payments are prescribed contingent on the reports. It is a generalization of the well-known single-period revelation principle; see Myerson (1981) for an early formulation. Doepke and Townsend (2005) proves a similar result for the class of communication games (where the two parties can exchange messages at certain points in time). In this paper, we show this result for more general long-term contracts.

Another challenge of the problem is that the optimal mechanisms may utilize the entire information history and therefore may appear too complicated to analyze. To address this problem, we derive the optimal contract by backward induction. That is, we break the problem into a series of constrained optimization problems, each of which finds an optimal action and payment menu in the current period and the optimal utility-to-go from the next period. The principal and agent's expected utility-to-go at the beginning of any period are vector valued: since the principal cannot observe the underlying state, any remaining contract will generate an expected future utility for each party in each potential state. The principal's maximum utility vector as a function of the agent's utility vector at the beginning of each period is called the efficient utility frontier in that period. Although backward induction has also been used in the papers mentioned above, explicitly characterizing the efficient frontiers is unique in this paper. As our results show, these frontiers possess nice properties such as polyhedral convexity/concavity which enables a hyperplane representation and computational geometry solution.

Our paper extends the existing literature in the following ways: it allows general long-term contracts, treats randomization explicitly, obtains sharper characterizations of the optimal solution, and provides an algorithm to find the exact solution. It should be pointed out, though, that some of our treatments and some properties of the optimal solution are enabled by the assumption that the agent is risk neutral toward monetary transfers (or, has quasi-linear utility function). This is a more restrictive assumption than that made in some of the previous papers.

**Organization.** The main part of the paper is organized as follows. In Section 2, we describe the model environment, long-term contracts and principal's problem. In Section 3, we establish the sufficiency of revelation contracts and present a reformulation of the principal's problem. We also discuss some extensions and potential applications of the model in this section. We solve the principal's problem in Section 4 by a dynamic programming algorithm, based on properties of the efficient utility frontiers. A computational geometry implementation is also presented. In Section 5, we focus on the two-state special case and discuss two numerical examples. We conclude with Section 6 on future research directions. Appendix A presents a road map of the paper, highlighting the main results and their connections. The reader may wish to consult this diagram as he or she proceeds through the paper. Technical proofs and additional materials for the computational geometry implementation are given in Appendices B and C.

# 2 Dynamic Principal-Agent Problem and Long-Term Contracts

**Underlying System and Information Structure.** The model has $T$ periods indexed by $t = 1, ..., T$. Within each period $t$ the following events take place: First, the agent privately observes the state of a Markov Decision Process, denoted by $x_t$; the state set is finite and denoted by $X = \{1, 2, \cdots, |X|\}$, with slight abuse of notation. Next, the agent takes a publicly observable action $a_t$ and incurs a cost $c_{x_t}(a_t)$; the action set is also finite, and is denoted by $A = \{1, 2, \cdots, |A|\}$. Toward the end of the period, the principal receives a reward $r_{x_t}(a_t)$. She then pays the agent $s_t$ depending on publicly observable and verifiable information. Finally, the transition to state $x_{t+1}$ occurs. The transition probability $\Pr(x_{t+1} = y | x_t = x, a_t = a)$ is denoted by $p_{x,y}(a)$ and we define a row vector $\boldsymbol{p}_x(a) = (p_{x,1}(a), ..., p_{x,|X|}(a))$ for any $x \in X$ and $a \in A$. The cost and reward structure and the transition probabilities are known to both parties. The distribution of the initial state $x_1$ is also publicly known, given by the probabilities $\beta_{x_1}$ with $\sum_{x_1 \in X} \beta_{x_1} = 1$. The beginning of period $t$ after the agent observes the state $x_t$ is called *time* $t$. The end of period $T$ is called the terminal time, or, time $T + 1$. The sequence of events and information structure is illustrated in Figure 1.

In the model, some information is public, such as the actions and payments, and the rest is private (only known by the agent), such as the system state. The *history of public information* $\boldsymbol{h}^t$ is a vector of all the public information up to time $t$. By default, $\boldsymbol{h}^1 = \varnothing$. The *history of full information* $\boldsymbol{\omega}^t$ is a vector of all the public and private information up to time $t$ (including $x_t$). We adopt a standard assumption from the contract theory literature:

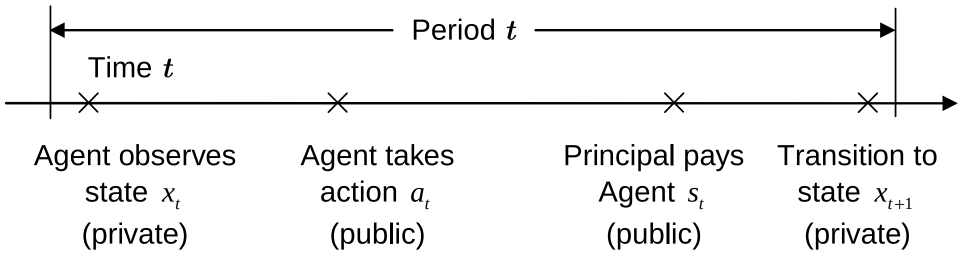


Figure 1: Sequence of Events.

**Assumption** The principal cannot infer the state $x_t$ from the reward $r_{x_t}(a_t)$.

If this assumption is violated, the incentive issue caused by information asymmetry is eliminated and there is no need for strategic interactions as studied in this paper. The assumption can be justified in at least two situations: In one situation (such as the classic seller-buyer problem, where the seller is the principal and the payments are from the agent to the principal), the reward $r(a_t)$ (the seller's cost) is independent of the state $x_t$ (the buyer's type) and therefore the state cannot be inferred from the reward. In another situation, the reward is an intangible value, for instance, the welfare or well-being of a patient in a health care setting. Even though the reward may depend on the state (the patient's health status), it cannot be physically measured or verified and therefore cannot reveal the state credibly.

We assume the principal and agent's utilities are additively separable across time periods and they are both risk-neutral toward monetary transfers. The discounted future utilities of the two parties at time $t$ are given by:

$$\text{Principal:} \qquad \sum_{\tau=t}^{T} \delta^{\tau-t}(r_{x_\tau}(a_\tau) - s_\tau) + \delta^{T+1-t}\pi_{x_{T+1}}, \tag{1}$$

$$\text{Agent:} \qquad \sum_{\tau=t}^{T} \delta^{\tau-t}(s_\tau - c_{x_\tau}(a_\tau)) + \delta^{T+1-t}u_{x_{T+1}}, \tag{2}$$

where the discount factor $\delta \in [0,1]$, and $\pi_{x_{T+1}}$ and $u_{x_{T+1}}$ denote the two parties' terminal utilities. We abbreviate $(x_1, \cdots, x_t)$ as $\boldsymbol{x}^t$, $(a_1, \cdots, a_t)$ as $\boldsymbol{a}^t$, and $(s_1, \cdots, s_t)$ as $\boldsymbol{s}^t$. Evidently, the vector $(\boldsymbol{x}^{T+1}, \boldsymbol{a}^T, \boldsymbol{s}^T) = (x_1, a_1, s_1, \cdots, x_T, a_T, s_T, x_{T+1})$ contains the most important public and private information and we call it the *history of essential information*.

**General Long-Term Contracts.** The principal is trying to design a contract that induces the agent to take desired actions. Before we provide a rigorous formulation of the principal's problem, a definition for contract is necessary. "Contract" is a synonym of "*payment scheme*" or "*mechanism*". We propose the following general definition:

**Definition 1** *A **long-term contract** $\sigma$ is an agreement between the principal and agent that specifies their public activities in periods* 1 *through* $T$*, including but not limited to the agent's actions* $a_t$ *and the principal's payments* $s_t$*, contingent on all public information available up to any given point in time.*

A long-term contract can be defined recursively as well:

**Definition 2** *A time-*$t$ ***continuation contract*** $\sigma_t(\boldsymbol{h}^t)$ *specifies the principal and agent's public activities in period* $t$ *after public history* $\boldsymbol{h}^t$ *(including but not limited to action* $a_t$ *and payment* $s_t$*), and a set of time-*$(t+1)$ *continuation contracts* $\sigma_{t+1}(\boldsymbol{h}^t, h_t)$*, one for each possible realization of period-*$t$ *public information* $h_t$*. A **long-term contract** is a time-*1 *continuation contract.*

These definitions reveal a main challenge of the problem, i.e., the principal is designing a game instead of playing an existing one. It is not hard to recognize that the space of long-term contracts is enormous and the possibilities are impossible to exhaust.

As standard in the literature, we make the following assumptions for long-term contracts: (1) The principal can make full commitment to follow the contract and will not renegotiate with the agent during its execution. As known in contract theory, it will hurt the principal in ex ante if she cannot make full commitment, because then the set of agent's future utilities that can be credibly provided by the principal is more limited, which reduces the principal's choices. (2) The agent can quit at the beginning of any period if continuation of the contract will not provide him at least his reservation utility. This captures many real world scenarios where the agent has the freedom to stop an undesirable relationship with the principal, for example, in the seller-buyer situation. (3) The physical characteristics of the underlying system cannot be altered by a contract. That is, in each period, the agent must take a public action $a_t$, only $a_t$ can affect the hidden state $x_t$, the state transition is fully governed by $p_{x_t,x_{t+1}}(a_t)$, and so on.

After the principal designs the contract $\sigma$ and the agent agrees on it, the contract execution starts. The agent's strategy, denoted by $\xi$, will be contingent on the history of full information. The strategy must be feasible in the sense that it must obey the contract terms. In addition, under $\sigma$ and $\xi$, the agent's expected future utility from any time $t$ must be well-defined. It leads to the following condition on the agent's strategies:

**Definition 3** *Given a long-term contract $\sigma$, the agent's strategy $\xi$ is **well-defined** (with respect to $\sigma$) if the pair $(\sigma, \xi)$ induces a probability space of full information history $\boldsymbol{\omega}^{T+1}$'s.*

Note that the distribution of the initial state $x_1$ is governed by the underlying system and not by the $(\sigma, \xi)$ pair. This condition guarantees that the essential information history $(x_1, a_1, s_1, \cdots, x_T, a_T, s_T, x_{T+1})$ is a random vector defined on the full information space. Naturally, any meaningful and unambiguous strategy of the agent should be well-defined. Therefore, in this paper, we will only consider well-defined strategies and omit the phrase "well-defined".

**The Principal's Problem.** Let $\Omega^t(\sigma, \xi) = \{\text{full information history } \boldsymbol{\omega}^t \text{ under } (\sigma, \xi)\}$, and $\widetilde{\pi}_t$ and $\widetilde{u}_t$ denote the two parties' discounted future utilities (1) and (2) respectively, which are random variables as well. Then, the principal's problem can be formulated as:

$$\max_{\sigma,\xi} E\left(\widetilde{\pi}_1 | \sigma, \xi\right) \tag{3}$$

$$\text{s.t.} \quad E\left(\widetilde{u}_t | \boldsymbol{\omega}^t, \sigma, \xi\right) \geq 0, \ \boldsymbol{\omega}^t \in \Omega^t(\sigma, \xi), t = 1, \cdots, T \tag{4}$$

$$E\left(\widetilde{u}_t | \boldsymbol{\omega}^t, \sigma, \xi\right) \geq E\left(\widetilde{u}_t | \boldsymbol{\omega}^t, \sigma, \xi'\right), \ \boldsymbol{\omega}^t \in \Omega^t(\sigma, \xi), t = 1, \cdots, T, \text{ any } \xi' \tag{5}$$

Constraints (4) are called *participation* (or *individual rationality, IR*) constraints, which match the agent's expected utility with his reservation utility (normalized to 0) at any time $t$. Constraints (5) are called *incentive compatibility* (*IC*) constraints, assuming that the agent will choose a strategy to maximize his expected future utility at any time $t$. Since these constraints are imposed in every period, the more precise names are *sequential incentive compatibility (SIC)* and *sequential individual rationality (SIR)* constraints.

Depending on when the agent compares the continuation contract with his outside offer, there are two types of IR constraints: the *ex ante* ones, where the agent can only leave before he observes the state; and the *ex post* ones, where he can quit after the observation. Ex post IR constraints are more restrictive from the principal's point of view. The IR constraints (4) are ex post ones since $\boldsymbol{\omega}^t$ includes $x_t$ by definition. We will focus on this type of IR constraints in this paper, but the analysis also applies to the ex ante case.

**Examples of Long-Term Contracts.** To shed more light on long-term contracts, we now discuss some examples, starting from the most intuitive one:

**Example 1** *A* ***deterministic action-based contract*** *can be denoted by* $\sigma = \{s_1(\boldsymbol{a}^1), s_2(\boldsymbol{a}^2), \cdots, s_T(\boldsymbol{a}^T)\}_{\boldsymbol{a}^T \in A^T}$*, or recursively,* $\sigma_t(\boldsymbol{a}^{t-1}) = \{s_t(\boldsymbol{a}^{t-1}, a_t), \sigma_{t+1}(\boldsymbol{a}^{t-1}, a_t)\}_{a_t \in A}$*. In such a contract, the payment* $s_t$ *in period* $t$ *depends solely on the agent's action history* $\boldsymbol{a}^t$*.*

Action-based contracts are appealing because of their simple structure. However, finding the best action-based contract is not a simple task and it is suboptimal in general. The following contract is more general:

**Example 2** *A* ***deterministic revelation contract*** *is denoted by* $\sigma = \{a_1(\widehat{\boldsymbol{x}}^1), s_1(\widehat{\boldsymbol{x}}^1); \cdots; a_T(\widehat{\boldsymbol{x}}^T), s_T(\widehat{\boldsymbol{x}}^T)\}_{\widehat{\boldsymbol{x}}^T \in X^T}$*, or* $\sigma_t(\widehat{\boldsymbol{x}}^{t-1}) = \{a_t(\widehat{\boldsymbol{x}}^{t-1}, \widehat{x}_t), s_t(\widehat{\boldsymbol{x}}^{t-1}, \widehat{x}_t), \sigma_{t+1}(\widehat{\boldsymbol{x}}^{t-1}, \widehat{x}_t)\}_{\widehat{x}_t \in X}$ *recursively. Under such a contract, in any period* $t$*, the agent makes a state report* $\widehat{x}_t$*, then takes the prescribed action* $a_t(\widehat{\boldsymbol{x}}^t)$*, and is paid* $s_t(\widehat{\boldsymbol{x}}^t)$*. Both the period-t action and payment depend on the report history* $\widehat{\boldsymbol{x}}^t$*. A revelation contract is called* ***truthful*** *if the agent's best response strategy is to report the state truthfully.*

Under an arbitrary revelation contract, the state reporting need not be truthful. An equivalent way to state that a revelation contract is truthful is that truthful reporting is sequentially incentive compatible. A revelation contract is more general than an action-based contract, because it allows the agent to take the same action yet receive different payments in different states. (But it can be shown that in the single period case, a revelation contract cannot do any better than an action-based contract.)

This type of contract looks artificial, but is actually not. Consider a single period revelation contract $\{a(\widehat{x}), s(\widehat{x})\}$ that requires the agent to report the state and then assigns an action and payment pair $(a, s)$ to him according to the report. The same result can be achieved if the principal proposes a menu of $(a, s)$ pairs and lets the agent freely choose one from them. The agent would choose a pair depending on the true state $x$, which has the same effect as reporting the state under a revelation contract.

The deterministic revelation contracts can be further generalized by randomization:

**Example 3** *A* ***randomized revelation contract*** *is denoted by* $\sigma = \{\theta_1(\widehat{x}_1, a_1), s_1(\widehat{x}_1, a_1); \cdots; \theta_T(\widehat{x}_T, a_T|\widehat{\boldsymbol{x}}^{T-1}, \boldsymbol{a}^{T-1}), s_T(\widehat{x}_T, a_T|\widehat{\boldsymbol{x}}^{T-1}, \boldsymbol{a}^{T-1})\}_{\widehat{\boldsymbol{x}}^T \in X^T, \boldsymbol{a}^T \in A^T}$*, or recursively,* $\sigma_t(\widehat{\boldsymbol{x}}^{t-1}, \boldsymbol{a}^{t-1}) = \{\theta_t(\widehat{x}_t, a_t|\widehat{\boldsymbol{x}}^{t-1}, \boldsymbol{a}^{t-1}), s_t(\widehat{x}_t, a_t|\widehat{\boldsymbol{x}}^{t-1}, \boldsymbol{a}^{t-1}), \sigma_{t+1}(\widehat{\boldsymbol{x}}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t, a_t)\}_{\widehat{x}_t \in X, a_t \in A}$*. Under such a contract, in any period* $t$ *after the report and action history* $(\widehat{\boldsymbol{x}}^{t-1}, \boldsymbol{a}^{t-1})$*, the following events take place: the agent reports the state* $\widehat{x}_t$*; he takes an action* $a_t$ *determined by a public random variable with probability mass function* $\theta_t(\widehat{x}_t, a_t|\widehat{\boldsymbol{x}}^{t-1}, \boldsymbol{a}^{t-1})$*, where* $\sum_{a_t \in A} \theta_t(\widehat{x}_t, a_t|\widehat{\boldsymbol{x}}^{t-1}, \boldsymbol{a}^{t-1}) = 1$ *for each* $\widehat{x}_t$*; and the principal pays the agent* $s_t(\widehat{x}_t, a_t|\widehat{\boldsymbol{x}}^{t-1}, \boldsymbol{a}^{t-1})$*.*

A deterministic revelation contract is a special case of a randomized one by restricting the probabilities to be 0 or 1. As will be shown in the next section, we can restrict our attention to these randomized revelation contracts without loss of generality. Later in Subsection 4.4, we will see the necessity to allow randomization. For convenience, we will often drop the "hat" sign from the reported states when it is clear from the context.

# 3 Revelation Principle and Principal's Problem

In this section, we will establish the sufficiency of revelation contracts, and reformulate the principal's problem under such contracts. We will also show that the sequential individual rationality constraints are unnecessary: only the individual rationality constraints in period one are required. At the end, we will discuss some simple extensions and possible applications of the model.

First, it is necessary to introduce notation for the two parties' expected utilities under a randomized revelation contract. If the agent reports the states truthfully, the two parties' expected future utilities from time $t$ given $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$ can be computed recursively as follows:

$$\begin{aligned}\pi_t(x_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) = \sum_{a_t\in A}\theta_t(x_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})\{&r_{x_t}(a_t) - s_t(x_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) + \\ &\delta\sum_{x_{t+1}\in X}p_{x_t,x_{t+1}}(a_t)\pi_{t+1}(x_{t+1}|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t, a_t))\},\ t = 1,\cdots,T \quad (6)\\ u_t(x_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) = \sum_{a_t\in A}\theta_t(x_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})\{&s_t(x_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) - c_{x_t}(a_t) + \\ &\delta\sum_{x_{t+1}\in X}p_{x_t,x_{t+1}}(a_t)u_{t+1}(x_{t+1}|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t, a_t))\},\ t = 1,\cdots,T \quad (7)\\ \pi_{T+1}(x_{T+1}|\boldsymbol{x}^{T}, \boldsymbol{a}^{T}) = \pi_{x_{T+1}}, \qquad u_{T+1}(x_{T+1}|\boldsymbol{x}^{T}, \boldsymbol{a}^{T}) = u_{x_{T+1}}&. \quad (8)\end{aligned}$$

If the agent misreports state $x_t$ as $\widehat{x}_t$ (and reports truthfully after period $t$), he will trigger the part of contract $(\theta_t(\widehat{x}_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}), s_t(\widehat{x}_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}), \sigma_{t+1}(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t, a_t))_{a_t\in A}$ and receive expected future utility

$$\begin{aligned}\widehat{u}_t(x_t, \widehat{x}_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) = \sum_{a_t\in A}\theta_t(\widehat{x}_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})\{&s_t(\widehat{x}_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) - c_{x_t}(a_t) + \\ &\delta\sum_{x_{t+1}\in X}p_{x_t,x_{t+1}}(a_t)u_{t+1}(x_{t+1}|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t, a_t))\}. \quad (9)\end{aligned}$$

Clearly, $u_t(x_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) = \widehat{u}_t(x_t, x_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$. The subscripts $x_t$ in $c_{x_t}(\cdot)$, $r_{x_t}(\cdot)$ and $\boldsymbol{p}_{x_t}(\cdot)$ always refer to the actual state.

Next, we present a main result:

**Theorem 1 (Dynamic Revelation Principle)** *For any long-term contract $\sigma$ coupled with the agent's best response strategy $\xi^*$, there exists a truthful revelation contract $\sigma^*$ which generates the same expected total utilities for the principal and agent as $(\sigma, \xi^*)$ does.*

The proof of the theorem is in Appendix B. The basic idea is to construct a truthful revelation contract $\sigma^*$ that induces the same marginal distribution of $(\boldsymbol{x}^{T+1}, \boldsymbol{a}^T)$ and the same marginal distribution of $\boldsymbol{s}^T$ conditional on $(\boldsymbol{x}^{T+1}, \boldsymbol{a}^T)$ as those induced by $(\sigma, \xi^*)$. The theorem applies to continuation contracts as well.

The theorem implies that, without loss of optimality, the principal's problem (3)-(5) can be recast as one of finding a revelation contract to maximize the principal's expected total utility:

$$\max_{\text{revelation contract } \sigma} \sum_{x_1 \in X} \beta_{x_1} \pi_1(x_1) \tag{10}$$

$$\text{s.t. } u_t(x_t | \boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) \geq 0, \qquad \boldsymbol{x}^t \in X^t, \boldsymbol{a}^{t-1} \in A^{t-1}, t = 1, \cdots, T, \tag{11}$$

$$u_t(x_t | \boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) \geq \widehat{u}_t(x_t, \widehat{x}_t | \boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}), \;\; \boldsymbol{x}^t \in X^t, \boldsymbol{a}^{t-1} \in A^{t-1}, \widehat{x}_t \in X, t = 1, \cdots, T. \tag{12}$$

$$\sum_{a_t \in A} \theta_t(x_t, a_t | \boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) = 1, \theta_t(x_t, a_t | \boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) \geq 0, \;\; \boldsymbol{x}^t \in X^t, \boldsymbol{a}^t \in A^t, t = 1, \cdots, T. \tag{13}$$

Constraints (11) replace the sequential individual rationality constraints (4). Constraints (12) replace the sequential incentive compatibility constraints (5), which ensure that truth-telling is the best response for the agent in every period. It is sufficient to consider the agent's one-period deviations in (12): using backward induction, one can show that if the agent cannot benefit from one-period deviations, he cannot benefit from multi-period deviations as well. Constraints (13) define the feasible range of the probabilities.

The next lemma shows a convenient result which says the SIR constraints (11) can be reduced to the first period IR constraints, i.e., $u_1(x_1) \geq 0$, for $x_1 \in X$.

**Lemma 1 (Redundancy of SIR)** *In the principal's problem, the sequential individual ra-*

*tionality constraints can be replaced by the first period individual rationality constraints without loss of optimality.*

**Immediate Extensions of the Model.** Problem (10)-(13) can be extended in multiple ways. Here, we present five straightforward extensions, which can be easily verified after the solution method is discussed in the next section. (1) For presentation convenience, the physical structure of the model is assumed time-independent, i.e., the functions $c_{x_t}(a_t)$, $r_{x_t}(a_t)$, and $p_{x_t,x_{t+1}}(a_t)$ are independent of $t$. But the method developed in the next section applies to time-dependent systems as well; (2) The model assumes the payments are from the principal to the agent, but the opposite case (as the seller-buyer example) can be accommodated with a simple change of signs of the rewards $r(\cdot)$, costs $c(\cdot)$ and payments $s(\cdot)$; (3) The principal may have the flexibility to contract with the agent for less than $T$ periods. This can be incorporated by defining a stopping action $a = 0$ with "transition probabilities" $\boldsymbol{p}_x(0) = \mathbf{0}$ for all $x$; (4) It is possible that in some situations, some states can be publicly observed and contracted upon. This can be handled by removing incentive compatibility constraints for those observable states; (5) The principal's objective function can be generalized to $\sum_{x_1 \in X} \beta_{x_1}[\pi_1(x_1) + \lambda u_1(x_1)]$, with $\lambda \in [0, 1]$. At the two extremes, $\lambda = 0$ represents a self-interested principal who maximizes her expected utility; $\lambda = 1$ represents a benevolent principal who maximizes the expected social welfare.

**Applications.** The model represents a general framework, which has many possible applications. Below are a few examples:

(1) *Dynamic Pricing with Changing Customer Types.* A firm sells a non-durable product (or service) in multiple periods. Each customer has a type $i_t$ that affects his or her utility in period $t$ and evolves according to a Markov Decision Process with transition probabilities $p(i_{t+1}|i_t, q_t)$ where $q_t$ is the purchasing quantity (or quality) in period $t$. The firm can observe customers' purchase decisions but not their types. The firm's objective is to design a pricing mechanism to maximize its profit.

(2) *Purchasing Contracts with Inventory Consideration.* A manufacturer wants to design

a multi-period contract with a retailer, knowing that the retailer carries inventory to maximize its own profit. In each period $t$, the retailer observes its initial inventory $I_t$ and places an order $u_t$ to the manufacturer. The inventory process is Markov with transition probabilities $p(I_{t+1}|I_t, u_t)$ determined by the random customer demand in period $t$. The distribution of customer demand is public information. The manufacturer can observe the retailer's orders but not its inventory level. What is the optimal contract for the manufacturer?

(3) *Health Care Contracts*. A payer (Medicare, private insurer) designs a contract with a provider (hospital, independent physician practice, etc.). In each period $t$, the provider observes its patients' health status $x_t$ and provides service $a_t$. Then $x_t$ can be defined as the state of a Markov Decision Process with transition probabilities $p(x_{t+1}|x_t, a_t)$. Existing information systems that facilitate transactions between payers and providers permit the payer to observe $a_t$ but not $x_t$. The objective of the payer is to maximize the patient's utility or the social welfare (if the payer is a government agency), or else to maximize the payer's own net profit.

(4) *Product Development Contracts*. An investor considers investing in a new venture. The state of the venture $s_t$ follows a Markov process with transition probabilities $p(s_{t+1}|s_t, e_t)$, where $e_t$ is management's effort in period $t$. The investor can quit in any state and receive a reward that depends on the unobservable state. What is the investor's optimal investment (and exit) strategy?

# 4 Finding Optimal Revelation Contracts

Theorem 1 enables us to focus on the set of truthful revelation contracts without loss of generality. In this section, we will develop a dynamic programming algorithm to find optimal revelation contracts and investigate their properties. We will first simplify the notation in Subsection 4.1. In Subsection 4.2, we will show that at any time $t$ there exist continuation revelation contracts that dominate others in the Pareto sense—no other contracts can generate the same utility vector for the agent and higher utility vector for the principal. These continuation contracts form an efficient frontier at time $t$ and can be obtained recursively. This result paves the way for a dynamic programming algorithm, presented in Subsection 4.3. In the course of the analysis, we will see that truthful revelation contracts can be constructed state by state independently; each state problem can be further decomposed into a collection of state-action problems. This decomposition into a hierarchy of state problems made up by state-action problems will streamline the construction of the efficient frontiers. We will investigate the properties of the efficient frontiers in Subsection 4.4. The most significant one is that of polyhedral concavity, which enables the construction of the efficient frontiers using hyperplanes and polytopes in Subsection 4.5. As a by-product, it will become clear why randomization is necessary. Subsection 4.5 is mainly intended for readers who are interested in the implementation of the algorithm in practice and can be skimmed otherwise.

## 4.1 Notation Simplification and Continuation Utilities

As introduced in Example 3, a time-$t$ randomized revelation contract can be represented as $\sigma_t(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) = \{\theta_t(x_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}), s_t(x_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}), \sigma_{t+1}(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t, a_t)\}_{x_t\in X, a_t\in A}$. If there is no need to emphasize the history $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$, it can be suppressed from the notation. We can further consolidate notation by moving the reported state $x_t$ to the subscript and dropping the time index $t$ whenever it is clear from the context. The consequence is a much simplified notation for a time-$t$ randomized revelation contract: $\sigma_t = \{\theta_x(a), s_x(a), \sigma_{t+1,x}(a)\}_{x\in X, a\in A}$. We call the collection $(\theta_x(a), s_x(a), \sigma_{t+1,x}(a))_{a\in A}$ given state $x$ a *(randomized) state-menu* with respect to $x$; and the part $(s_x(a), \sigma_{t+1,x}(a))$ given $x$ and $a$ a *state-action-option* with respect to $(x, a)$. In addition, the state $x_{t+1}$ will be denoted by $y$.

If the continuation revelation contract $\sigma_t$ is truthful, the two parties' expected future utilities $\pi_t(x_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$ and $u_t(x_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$, defined in (6) and (7), are called their *time-$t$ continuation utilities* and abbreviated as $\pi_x(\sigma_t)$ and $u_x(\sigma_t)$ respectively. It is convenient to define the column vectors $\boldsymbol{\pi}(\sigma_t) = (\pi_x(\sigma_t))_{x\in X}$ and $\boldsymbol{u}(\sigma_t) = (u_x(\sigma_t))_{x\in X}$ and view that the continuation contract $\sigma_t$ generates the pair of continuation utility vectors $(\boldsymbol{u}, \boldsymbol{\pi})(\sigma_t)$. Then,

(6) and (7) can be rewritten using the simplified notation as:

$$\pi_x(\sigma_t) = \sum_{a\in A} \theta_x(a) \left\{ r_x(a) - s_x(a) + \delta \boldsymbol{p}_x(a) \boldsymbol{\pi}(\sigma_{t+1,x}(a)) \right\}, \tag{14}$$

$$u_x(\sigma_t) = \sum_{a\in A} \theta_x(a) \left\{ s_x(a) - c_x(a) + \delta \boldsymbol{p}_x(a) \boldsymbol{u}(\sigma_{t+1,x}(a)) \right\}. \tag{15}$$

Notice that we write the matrix multiplication of a row vector $\boldsymbol{p}$ and a column vector $\boldsymbol{u}$ as $\boldsymbol{pu}$. In subsequent discussions, it will also be convenient to work with the social welfare function, defined as $\phi_x(\sigma_t) = u_x(\sigma_t) + \pi_x(\sigma_t)$, and the social welfare vector $\boldsymbol{\phi}(\sigma_t) = \boldsymbol{u}(\sigma_t) + \boldsymbol{\pi}(\sigma_t)$. It is equivalent to say that the continuation contract $\sigma_t$ generates the pair of continuation utility vectors $(\boldsymbol{u}, \boldsymbol{\phi})(\sigma_t)$.

## 4.2 Existence of Efficient Frontiers and Sequential Efficiency

According to the problem formulation (10)-(13) and the recursive formulas (6)-(9), a continuation contract $\sigma_t$ plays a role in the principal's problem only through the utility vector pairs $(\boldsymbol{u}_t, \boldsymbol{\pi}_t)$ generated by it. Suppose two continuation contracts $\sigma_t$ and $\sigma_t^\dagger$ generate the same utility vector $\boldsymbol{u}_t$ for the agent but different utility vectors $\boldsymbol{\pi}_t$ and $\boldsymbol{\pi}_t^\dagger$ for the principal. If $\boldsymbol{\pi}_t \geq \boldsymbol{\pi}_t^\dagger$, by formulation (10)-(13), the utility pair $(\boldsymbol{u}_t, \boldsymbol{\pi}_t)$ (or contract $\sigma_t$) dominates the pair $(\boldsymbol{u}_t, \boldsymbol{\pi}_t^\dagger)$ (or contract $\sigma_t^\dagger$) from the principal's perspective yet there is no difference from the agent's point of view. But if neither $\boldsymbol{\pi}_t \geq \boldsymbol{\pi}_t^\dagger$ nor $\boldsymbol{\pi}_t^\dagger \geq \boldsymbol{\pi}_t$, it is unclear which utility pair is better. Fortunately, in that situation, there always exists a third utility pair $(\boldsymbol{u}_t, \boldsymbol{\pi}_t^*)$ that dominates the first two pairs and is generated by a truthful revelation contract, as shown below:

**Lemma 2 (Dominant Continuation Contracts)** *Consider two truthful revelation contracts:* $\sigma_t = \{\theta_x(a), s_x(a), \sigma_{t+1,x}(a)\}_{x\in X, a\in A}$ *and* $\sigma_t^\dagger = \{\theta_x^\dagger(a), s_x^\dagger(a), \sigma_{t+1,x}^\dagger(a)\}_{x\in X, a\in A}$. *Suppose* $\boldsymbol{u}(\sigma_t) = \boldsymbol{u}(\sigma_t^\dagger) = \boldsymbol{u}_t$, $\pi_x(\sigma_t) > \pi_x(\sigma_t^\dagger)$ *for* $x \in S$, $\pi_x(\sigma_t) < \pi_x(\sigma_t^\dagger)$ *for* $x \in S^\dagger$, *and* $\pi_x(\sigma_t) = \pi_x(\sigma_t^\dagger)$ *otherwise. Then, the contract*

$$\sigma_t^* = \begin{cases} \theta_x(a), s_x(a), \sigma_{t+1,x}(a), \ \textit{if } x \in S, \\ \theta_x^\dagger(a), s_x^\dagger(a), \sigma_{t+1,x}^\dagger(a), \ \textit{if } x \in S^\dagger, \\ \textit{either one of the above}, \quad \textit{otherwise} \end{cases}$$

*is a truthful revelation contract which satisfies* $\boldsymbol{u}(\sigma_t^*) = \boldsymbol{u}_t$ *and* $\boldsymbol{\pi}(\sigma_t^*) = \max\{\boldsymbol{\pi}(\sigma_t), \boldsymbol{\pi}(\sigma_t^\dagger)\}$.

*Proof.* Using the simplified notation, the IC constraints (12) for $\sigma_t$ can be rewritten as follows: for any $x \in X$,

$$\begin{aligned} &\sum_{a \in A} \theta_x(a) \left\{ s_x(a) - c_x(a) + \delta \boldsymbol{p}_x(a) \boldsymbol{u}(\sigma_{t+1,x}(a)) \right\} \\ \geq &\sum_{a \in A} \theta_{\widehat{x}}(a) \left\{ s_{\widehat{x}}(a) - c_x(a) + \delta \boldsymbol{p}_x(a) \boldsymbol{u}(\sigma_{t+1,\widehat{x}}(a)) \right\}, \ \widehat{x} \in X. \end{aligned} \tag{16}$$

Using the fact $\boldsymbol{u}(\sigma_t) = \boldsymbol{u}_t$, or more precisely,

$$u_{t,x} = \sum_{a \in A} \theta_x(a) \left\{ s_x(a) - c_x(a) + \delta \boldsymbol{p}_x(a) \boldsymbol{u}(\sigma_{t+1,x}(a)) \right\}, \ x \in X, \tag{17}$$

we can transform the above IC constraints as follows (with the roles of $x$ and $\widehat{x}$ switched):

$$u_{t,\widehat{x}} - u_{t,x} \geq \sum_{a \in A} \theta_x(a) \left\{ c_x(a) - c_{\widehat{x}}(a) + \delta \left[ \boldsymbol{p}_{\widehat{x}}(a) - \boldsymbol{p}_x(a) \right] \boldsymbol{u}(\sigma_{t+1,x}(a)) \right\}, \ \widehat{x} \in X. \tag{18}$$

That is, any state-menu $(\theta_x(a), s_x(a), \sigma_{t+1,x}(a))_{a \in A}$ satisfies (18) given $x$.

Similarly, since $\sigma_t^\dagger$ is truthful and $\boldsymbol{u}(\sigma_t^\dagger) = \boldsymbol{u}_t$, each state-menu $(\theta_x^\dagger(a), s_x^\dagger(a), \sigma_{t+1,x}^\dagger(a))_{a \in A}$ satisfies (18) given $x$ as well (replacing $\theta$ and $\sigma_{t+1}$ by $\theta^\dagger$ and $\sigma_{t+1}^\dagger$). The third contract $\sigma_t^*$ is clearly a revelation contract. The way it is constructed does not change the agent's utility vector or the incentive compatibility of any of the state-menus. So, $\sigma_t^*$ must be truthful and satisfy $\boldsymbol{u}(\sigma_t^*) = \boldsymbol{u}_t$ and $\boldsymbol{\pi}(\sigma_t^*) = \max\{\boldsymbol{\pi}(\sigma_t), \boldsymbol{\pi}(\sigma_t^\dagger)\}$. ■

The proof is presented here instead of Appendix B because some of its expressions provide insights that facilitate the decomposition of the principal's problem. Specifically, constraints (18) state that a state-menu of a truthful revelation contract intended for state $x$ should not be too attractive under another state $\widehat{x}$. The constraints enable us to decompose the problem of finding a continuation contract that maximizes the principal's continuation utility vector given the agent's utility vector $\boldsymbol{u}_t$ into $|X|$ independent state-menu problems. More specifically, given $\boldsymbol{u}_t$ and for each state $x$, we find a state-menu $(\theta_x(a), s_x(a), \sigma_{t+1,x}(a))_{a \in A}$ to maximize the principal's utility, subject to constraints (18). Combining these optimal state-menus, we obtain a continuation contract that attains the highest achievable utility vector $\boldsymbol{\pi}_t^*$ for the principal while maintaining the agent's utility at $\boldsymbol{u}_t$. The state-menu problem will be formally defined in the next subsection.

The concept of efficient frontier follows from Lemma 2.

**Definition 4** *The set of time-$t$ continuation truthful revelation contracts is denoted by $\Sigma_t^{TR}$. The time-$t$ **agent's utility set** is defined as $\boldsymbol{U}_t \triangleq \{\boldsymbol{u}(\sigma_t) : \sigma_t \in \Sigma_t^{TR}\}$. The time-$t$ **efficient utility frontier** is defined as:*

$$\boldsymbol{\pi}_t^*(\boldsymbol{u}_t) = \max\{\boldsymbol{\pi}(\sigma_t) : \sigma_t \in \Sigma_t^{TR} \text{ and } \boldsymbol{u}(\sigma_t) = \boldsymbol{u}_t\},\ \ \boldsymbol{u}_t \in \boldsymbol{U}_t, \tag{19}$$

*and the time-$t$ **efficient social welfare frontier** is defined as:*

$$\boldsymbol{\phi}_t^*(\boldsymbol{u}_t) = \max\{\boldsymbol{\phi}(\sigma_t) : \sigma_t \in \Sigma_t^{TR} \text{ and } \boldsymbol{u}(\sigma_t) = \boldsymbol{u}_t\},\ \ \boldsymbol{u}_t \in \boldsymbol{U}_t. \tag{20}$$

In the definition, the max operation is taken component-wise. The $x$-th components of $\boldsymbol{\pi}_t^*(\boldsymbol{u}_t)$ and $\boldsymbol{\phi}_t^*(\boldsymbol{u}_t)$ will be denoted by $\pi_{t,x}^*(\boldsymbol{u}_t)$ and $\phi_{t,x}^*(\boldsymbol{u}_t)$ respectively. The two types of efficient frontiers are equivalent—a utility vector pair $(\boldsymbol{u}_t, \boldsymbol{\pi}_t)$ is on the efficient utility frontier if and only if $(\boldsymbol{u}_t, \boldsymbol{\phi}_t)$ is on the efficient social welfare frontier, thanks to the identity:

$$\boldsymbol{\phi}_t^*(\boldsymbol{u}_t) = \boldsymbol{\pi}_t^*(\boldsymbol{u}_t) + \boldsymbol{u}_t. \tag{21}$$

We call a contract $\sigma_t$ *efficient* if the utility vector pair $(\boldsymbol{u}, \boldsymbol{\pi})(\sigma_t)$ or $(\boldsymbol{u}, \boldsymbol{\phi})(\sigma_t)$ lies on the corresponding efficient frontier. It turns out the social welfare frontiers are more convenient to analyze and thus are our focus in the subsequent development.

Lemma 2 shows the existence of efficient frontiers. The result can be extended further to establish that an efficient contract is sequentially efficient. As a consequence, efficient frontiers can be constructed recursively.

**Theorem 2 (Sequential Efficiency)** *If $\sigma_t \in \Sigma_t^{TR}$ is efficient, for any public history $\boldsymbol{h}^\tau$, $\tau > t$, that occurs with non-zero probability, the continuation contract $\sigma_\tau(\boldsymbol{h}^\tau)$ is also efficient.*

## 4.3 Dynamic Programming Formulation

Sequential efficiency is derived from the principal's problem by ignoring the sequential individual rationality constraints. Combining Lemma 1 with Theorem 2, we can recast the principal's problem in the dynamic programming fashion. Define $\boldsymbol{U}_{T+1} = \{\boldsymbol{u}_{T+1}\}$ and $\boldsymbol{\phi}^*_{T+1}(\boldsymbol{u}_{T+1}) = \boldsymbol{\phi}_{T+1}$, where $(\boldsymbol{u}_{T+1}, \boldsymbol{\phi}_{T+1})$ is the terminal utility/social welfare vector pair. We have the following algorithm:

1. For $t = T$ to 1: obtain the time-$t$ efficient social welfare frontier $\boldsymbol{\phi}^*_t : \boldsymbol{U}_t \to \mathbb{R}^{|X|}$ from the time-$(t+1)$ efficient social welfare frontier $\boldsymbol{\phi}^*_{t+1} : \boldsymbol{U}_{t+1} \to \mathbb{R}^{|X|}$;
2. Solve the principal's ex ante problem: $\max\{\boldsymbol{\beta}\,(\boldsymbol{\phi}^*_1(\boldsymbol{u}_1) - \boldsymbol{u}_1) : \boldsymbol{u}_1 \in \boldsymbol{U}_1, \boldsymbol{u}_1 \geq \boldsymbol{0}\}$.

The first step is an iteration step, which is the core of the algorithm. The second step is a one-shot optimization problem to be solved at time 1, taking into account the first period IR constraints and the distribution of the initial state. Below, we formulate the iteration step precisely.

As discussed in the last subsection, a continuation contract $\sigma_{t+1}$ plays a role in the principal's problem only through the utility vector pairs $(\boldsymbol{u}_{t+1}, \boldsymbol{\phi}_{t+1})$ generated by it. In addition, since we only need to consider efficient continuation contracts, $\boldsymbol{\phi}_{t+1}$ can be determined from the known efficient social welfare frontier $\boldsymbol{\phi}^*_{t+1}(\cdot)$ through $\boldsymbol{\phi}_{t+1} = \boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1})$. Thus, given $\boldsymbol{\phi}^*_{t+1}(\cdot)$, a randomized state-menu $(\theta_x(a), s_x(a), \sigma_{t+1,x}(a))_{a\in A}$ reduces to $(\theta_x(a), s_x(a), \boldsymbol{u}_{t+1,x}(a))_{a\in A}$, and the *randomized state-menu problem* for state $x$ can be defined as:

$$\phi^*_{t,x}(\boldsymbol{u}_t) \triangleq \max_{\theta(a), \boldsymbol{u}_{t+1}(a) \in \boldsymbol{U}_{t+1}} \sum_{a\in A} \theta(a)\left\{r_x(a) - c_x(a) + \delta \boldsymbol{p}_x(a) \boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1}(a))\right\} \tag{22}$$

$$\text{s.t. } u_{t,\widehat{x}} - u_{t,x} \geq \sum_{a\in A} \theta(a)\left\{c_x(a) - c_{\widehat{x}}(a) + \delta[\boldsymbol{p}_{\widehat{x}}(a) - \boldsymbol{p}_x(a)]\boldsymbol{u}_{t+1}(a)\right\}, \ \widehat{x} \in X \tag{23}$$

$$\sum_{a\in A} \theta(a) = 1, \text{ and } \theta(a) \geq 0, \ a \in A \tag{24}$$

We call the set of parameters $\boldsymbol{u}_t$ making problem (22)-(24) feasible the problem's *(feasible) parameter set*, denoted $\boldsymbol{U}_{t,x}$. Clearly, the optimal values of the variables $\theta(a)$ and $\boldsymbol{u}_{t+1}(a)$

depend on the state $x$. The above representation highlights the recursive nature of the principal's problem: In period $t$ and state $x$, the principal must choose a randomization over the agent's action $a$ and continuation utility vector $\boldsymbol{u}_{t+1}(a)$ to maximize the time-$t$ continuation social welfare while providing an exogenously chosen continuation utility vector $\boldsymbol{u}_t$ for the agent. Payments $s_x(a)$ do not appear in the formulation since they are fully captured by $\boldsymbol{u}_t$ according to (17) and can be easily recovered when needed. The formulation reveals an important characterization of the optimal contracts: the commitment to future payments plays a vital role in providing incentives for truth-telling.

For comparison, we consider the first-best case when all information is public. In that case, the principal can dictate what the agent should do and therefore she will enforce the highest continuation social welfare vector (while compensating the agent with his reservation utility). It leads to a deterministic non-constrained dynamic programming formula:

$$\phi^*_{t,x} = \max_{a\in A} r_x(a) - c_x(a) + \delta \boldsymbol{p}_x(a)\boldsymbol{\phi}^*_{t+1}. \tag{25}$$

## 4.4 Properties of Efficient Frontiers

Randomized state-menus consist of state-action-options. Thus, the state-menu problem can be further decomposed into problems of finding the efficient state-action-options. The *state-action-option (SAO) problem* is defined as follows: for a given agent's utility vector $\boldsymbol{u}_t$ and a given state-action pair $(x, a)$, choose continuation utility vector $\boldsymbol{u}_{t+1}$ to solve

$$\phi^*_{t,x}(\boldsymbol{u}_t|a) \triangleq \max_{\boldsymbol{u}_{t+1}\in \boldsymbol{U}_{t+1}} r_x(a) - c_x(a) + \delta \boldsymbol{p}_x(a)\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1}) \tag{26}$$

$$\text{s.t. } u_{t,\widehat{x}} - u_{t,x} \geq c_x(a) - c_{\widehat{x}}(a) + \delta\left[\boldsymbol{p}_{\widehat{x}}(a) - \boldsymbol{p}_x(a)\right]\boldsymbol{u}_{t+1},\ \widehat{x}\in X. \tag{27}$$

The parameter set of the problem is denoted by $\boldsymbol{U}_{t,x}(a)$.

The solution to the randomized state-menu problem (22)-(24) can be obtained by convexification over the set of optimal objective functions $\phi^*_{t,x}(\boldsymbol{u}_t|a)$, $a \in A$, of the state-action-option problems in the following way:

$$\phi^*_{t,x}(\boldsymbol{u}_t) = \max_{\{\theta(a)\geq 0, \boldsymbol{u}'_t(a)\in \boldsymbol{U}_{t,x}(a)\}} \left\{ \begin{array}{c} \sum_{a\in A}\theta(a)\phi^*_{t,x}(\boldsymbol{u}'_t(a)|a): \\ \sum_{a\in A}\theta(a)\boldsymbol{u}'_t(a) = \boldsymbol{u}_t, \sum_{a\in A}\theta(a) = 1 \end{array} \right\}. \tag{28}$$

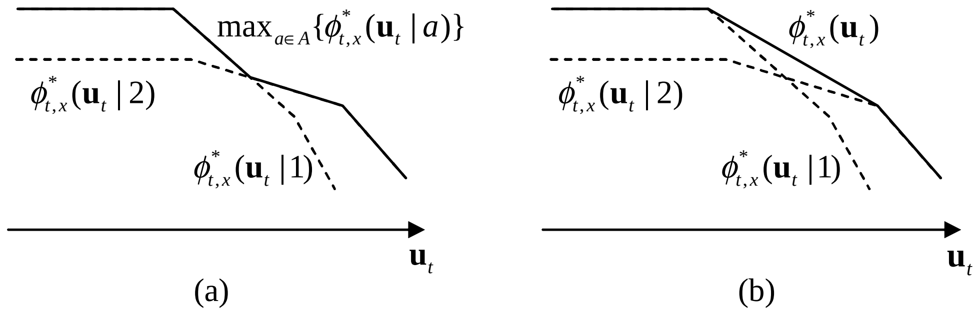


Figure 2: Solving the state-menu problem through the state-action-option problems: (a) A deterministic solution; (b) A randomized solution.

The relationship between $\phi^*_{t,x}(\boldsymbol{u}_t)$ and $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ is illustrated in Figure 2(b) in the two-action case. Figure 2(a) demonstrates that it is in general suboptimal to use deterministic state-menus, by solving $\max_{a\in A}\{\phi^*_{t,x}(\boldsymbol{u}_t|a)\}$, and randomization among state-action-options is necessary.

Next, we show that the efficient frontiers have an important property, as defined below:

**Definition 5** *A* ***set*** $\boldsymbol{U} \subset \mathbb{R}^n$ *is* ***polyhedral convex*** *if there exist a matrix* $\boldsymbol{A}$ *and a vector* $\boldsymbol{b}$ *such that* $\boldsymbol{U} = \{\boldsymbol{u} \subset \mathbb{R}^n : \boldsymbol{A}\boldsymbol{u} \leq \boldsymbol{b}\}$. *A* ***function*** $\phi : \boldsymbol{U} \to \mathbb{R}$ *is* ***polyhedral convex (concave)*** *if the domain* $\boldsymbol{U}$ *is polyhedral convex and there exist finitely many row vectors* $\boldsymbol{h}^i$ *and scalars* $d^i$ *such that* $\phi(\boldsymbol{u}) = \max(\min)_{i=1,\cdots,I}\{\boldsymbol{h}^i\boldsymbol{u} + d^i\}$.

Since the SAO problem (26)-(27) has a linear structure, it is not surprising that if $\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1})$ is polyhedral concave, this property will be preserved in $\phi^*_{t,x}(\boldsymbol{u}_t|a)$. In addition, convexification over $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ will not destroy this property as well. Let the convex hull of set $\boldsymbol{U}$ be denoted by *conv*$(\boldsymbol{U})$ and the *hypograph* of function $\phi : \boldsymbol{U} \to \mathbb{R}$ be defined as *hypo*$(\phi) = \{(\boldsymbol{u}, t) \in \boldsymbol{U} \times \mathbb{R} : t \leq \phi(\boldsymbol{u})\}$. We have the following results:

**Theorem 3 (Polyhedral Convexity/Concavity)** *(a) The parameter set* $\boldsymbol{U}_{t,x}(a)$ *of the state-action-option problem (26)-(27) is polyhedral convex and the optimal objective function* $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ *is polyhedral concave in* $\boldsymbol{u}_t$. *(b) The parameter set* $\boldsymbol{U}_{t,x}$ *and the optimal* *objective function $\phi^*_{t,x}(\boldsymbol{u}_t)$ of the randomized state-menu problem (22)-(24) satisfy: $\boldsymbol{U}_{t,x} = conv\left(\cup_{a\in A}\boldsymbol{U}_{t,x}(a)\right)$ and $hypo(\phi^*_{t,x}(\boldsymbol{u}_t)) = conv(\cup_{a\in A} hypo(\phi^*_{t,x}(\boldsymbol{u}_t|a)))$. Consequently, $\boldsymbol{U}_{t,x}$ is polyhedral convex and $\phi^*_{t,x}(\boldsymbol{u}_t)$ is polyhedral concave. (c) The time-t agent's utility set $\boldsymbol{U}_t = \cap_{x\in X}\boldsymbol{U}_{t,x}$ and is polyhedral convex.*

This theorem provides the foundation for the next subsection in which we will develop an algorithm to construct efficient frontiers by hyperplanes and polytopes. Before that, we show that the agent's utility space has a one-degree redundancy which can be exploited in any calculation of the efficient frontiers.

**Lemma 3 (Redundancy of Agent's Utility Space)** *The efficient social welfare frontier satisfies $\boldsymbol{\phi}^*_t(\boldsymbol{u}_t) = \boldsymbol{\phi}^*_t(\boldsymbol{u}_t + \lambda\mathbf{1})$ for any $\lambda \in \mathbb{R}$.*

The proof can be directly obtained from the problem formulation (22)-(24) and is omitted. Intuitively, in any period, the principal can use a constant payment to alter the agent's utility in all states simultaneously. Since both parties are risk-neutral over monetary payments, this will not change the agent's relative incentives or the social welfare.

This redundancy can be resolved by shifting our attention from the agent's *absolute utility vector* $\boldsymbol{u}_t = (u_{t,1}, \cdots, u_{t,|X|})$ to a *relative utility vector* $\Delta\boldsymbol{u}_t = (\Delta u_{t,1}, \cdots, \Delta u_{t,|X|})$ with respect to certain state $x$, where $\Delta u_{t,\widehat{x}} = u_{t,\widehat{x}} - u_{t,x}$ for $\widehat{x} \neq x$ and $\Delta u_{t,x}$ is undefined (this non-conventional treatment can facilitate the switching among different relative forms). The efficient social welfare function is still $\boldsymbol{\phi}^*_t(\Delta\boldsymbol{u}_t) = \boldsymbol{\phi}^*_t(\boldsymbol{u}_t)$, with $|X|$ components. When it is clear from the context, we can drop the sign "$\Delta$" from a relative utility vector.

This dimensional reduction reduces the agent's utility set $\boldsymbol{U}_t$ to a bounded polytope, because of the following constraints for each pair of states $x$ and $\widehat{x}$:

$$u_{t,\widehat{x}} - u_{t,x} \geq \sum_{a\in A}\theta_x(a)\left\{c_x(a) - c_{\widehat{x}}(a) + \delta[\boldsymbol{p}_{\widehat{x}}(a) - \boldsymbol{p}_x(a)]\boldsymbol{u}_{t+1,x}(a)\right\} \tag{29}$$

$$u_{t,\widehat{x}} - u_{t,x} \leq \sum_{a\in A}\theta_{\widehat{x}}(a)\left\{c_x(a) - c_{\widehat{x}}(a) + \delta[\boldsymbol{p}_{\widehat{x}}(a) - \boldsymbol{p}_x(a)]\boldsymbol{u}_{t+1,\widehat{x}}(a)\right\}. \tag{30}$$

Both inequalities follow from the incentive compatibility constraints (23) (the subscripts $x$ and $\widehat{x}$ under $\theta$ and $\boldsymbol{u}_{t+1}$ are added for clarity); the latter inequality also requires switching $x$ and $\widehat{x}$.

## 4.5 Implementation by Computational Geometry

We have reformulated the principal's problem using dynamic programming, based on efficient social welfare frontiers. The iteration step consists of problems at two levels: the randomized state-menu problem (22)-(24) and the state-action-option problem (26)-(27). A main challenge posed by these problems is that the parameter $\boldsymbol{u}_t$ is drawn from an infinite set $\boldsymbol{U}_t$, the agent's time-$t$ utility vector set, and we need to solve these problems for every $\boldsymbol{u}_t \in \boldsymbol{U}_t$. Theorem 3 suggests a possible remedy to overcome this difficulty: both optimal objective functions $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ and $\phi^*_{t,x}(\boldsymbol{u}_t)$ can be represented by a finite number of hyperplanes and hence can be determined in finite time. In this subsection, we outline such an implementation. The goal is to break the iteration step into standard computational geometry problems. Our treatment is just one possible way of implementation and its efficiency depends largely on the techniques used for solving those standard problems. Due to space limitation, more details of the implementation are provided in Appendix C. For more in-depth coverage of computational geometry, we refer the reader to Berg et al (2000) and Boissonnat and Yvinec (1998).

The iteration step for period $t$ can be divided into five stages: (1) For each state-action pair $(x, a)$, represent the objective function in (26) by a set of hyperplanes in terms of $\boldsymbol{u}_{t+1}$; (2) Simplify the SAO problem (26)-(27) by a change of variables; (3) Solve the simplified SAO problem by projecting the underlying polytope of the objective function along proper axes; (4) For each state $x$, construct the function $\phi^*_{t,x}(\boldsymbol{u}_t)$ and the set $\boldsymbol{U}_{t,x}$ as the convex hulls of $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ and $\boldsymbol{U}_{t,x}(a)$, $a \in A$, respectively; (5) Construct the agent's utility set $\boldsymbol{U}_t$ as the intersection of $\boldsymbol{U}_{t,x}$, $x \in X$.

The first three stages correspond to the state-action-option problem (26)-(27). The last

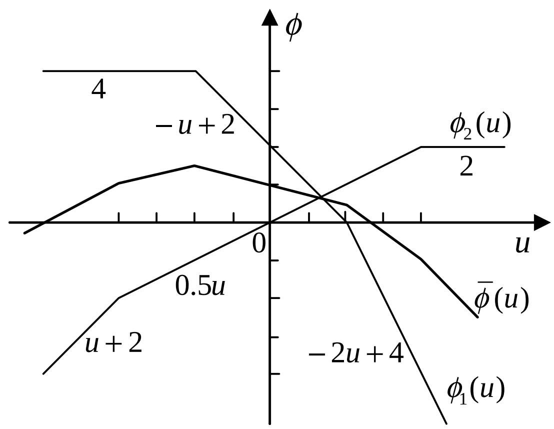


Figure 3: Average of Two Polyhedral Concave Functions. $\phi_1(u)$ is formed from hyperplanes $\phi = 4$, $\phi = -u + 2$, and $\phi = -2u + 4$; $\phi_2(u)$ is formed from $\phi = u + 2$, $\phi = 0.5u$, and $\phi = 2$; and $\overline{\phi}(u) = 0.5\phi_1(u) + 0.5\phi_2(u)$ can be represented by hyperplanes $\phi = 0.5u + 3$, $\phi = 0.25u + 2$, $\phi = -0.25u + 1$, $\phi = -0.75u + 2$, and $\phi = -u + 3$.

two stages correspond to the state-menu problem (22)-(24), involving standard operations such as finding the convex hull and the intersection of a set of polytopes. In what follows, we elaborate on the first three stages.

**Stage 1: Representing the Objective Function in (26) by Hyperplanes.** Assume the hyperplane representation of the efficient frontier $\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1})$ is already known as: $\phi^*_{t+1,y}(\boldsymbol{u}_{t+1}) = \min_{i=1,\cdots,I_y} \left(\boldsymbol{h}^i_y \boldsymbol{u}_{t+1} + d^i_y\right)$, for $y \in X$ and a set of integers $\{I_y\}$ (recall that $y$ denotes $x_{t+1}$). The immediate task we face is to represent the objective function of the SAO problem, $r_x(a) - c_x(a) + \delta \boldsymbol{p}_x(a)\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1})$ or essentially $\boldsymbol{p}_x(a)\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1})$, by a set of hyperplanes. We call the smallest set of hyperplanes that completely determine a polyhedral concave function the *defining hyperplanes* of that function. As illustrated in Figure 3 in the two-dimensional case, finding the defining hyperplanes is not a trivial task.

In Appendix C, we show how to transform this problem into an existing problem in computational geometry by exploring the duality between hyperplanes and points. The *dual* of a hyperplane $\{(\boldsymbol{u}, \phi) : \phi = \boldsymbol{h}\boldsymbol{u} + d\}$, abbreviated as $\phi = \boldsymbol{h}\boldsymbol{u} + d$, is the point $(\boldsymbol{h}, d)$, and vice versa. The problem of finding the defining hyperplanes of $\boldsymbol{p}_x(a)\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1})$ is equivalent to the problem of finding a convex hull in the dual space.

At the end of this stage, for any given $x$ and $a$, we can represent the objective function in (26) by its defining hyperplanes:

$$r_x(a) - c_x(a) + \delta \boldsymbol{p}_x(a)\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1}) = \min_{i=1,\cdots,I_x(a)} \boldsymbol{h}^i_x(a)\boldsymbol{u}_{t+1} + d^i_x(a). \tag{31}$$

**Stage 2: Change of Variables.** The SAO problem can be simplified by the change of variables $w_{t,x} = -c_x(a) + \delta \boldsymbol{p}_x(a)\boldsymbol{u}_{t+1},\ x \in X$, or

$$\boldsymbol{w}_t = -\boldsymbol{c}(a) + \delta \boldsymbol{P}(a)\boldsymbol{u}_{t+1} \tag{32}$$

in the matrix form. The constraints (27) reduce to: $u_{t,\widehat{x}} - u_{t,x} \geq w_{t,\widehat{x}} - w_{t,x},\ \widehat{x} \in X$. Assume the transition matrices $\boldsymbol{P}(a)$, $a \in A$, have full ranks and hence $\boldsymbol{P}(a)^{-1}$ exist (the degenerate case is discussed in the appendix). For each action $a \in A$, relation (32) defines a one-to-one affine mapping between $\boldsymbol{u}_{t+1}$ and $\boldsymbol{w}_t$. Thus, the agent's utility vector set $\boldsymbol{U}_{t+1} = \{\boldsymbol{u}_{t+1} : \boldsymbol{A}\boldsymbol{u}_{t+1} \leq \boldsymbol{b}\}$ is mapped to

$$\boldsymbol{W}_t(a) = \{\boldsymbol{w}_t : \left[\boldsymbol{A}\boldsymbol{P}(a)^{-1}\right] \boldsymbol{w}_t \leq \delta \boldsymbol{b} - \boldsymbol{A}\boldsymbol{P}(a)^{-1}\boldsymbol{c}(a)\}. \tag{33}$$

Also, the defining hyperplanes of the objective function (31) is transformed from the $(\boldsymbol{u}_{t+1}, \phi)$-space to the $(\boldsymbol{w}_t, \phi)$-space:

$$\phi'_{t,x}(\boldsymbol{w}_t|a) \triangleq \min_{i=1,\cdots,I_x(a)} \left(\delta^{-1}\boldsymbol{h}^i_x(a)\boldsymbol{P}(a)^{-1}\right) \boldsymbol{w}_t + \left(\delta^{-1}\boldsymbol{h}^i_x(a)\boldsymbol{P}(a)^{-1}\boldsymbol{c}(a) + d^i_x(a)\right). \tag{34}$$

As a result, the SAO problem (26)-(27) can be simplified to:

$$\phi^*_{t,x}(\boldsymbol{u}_t|a) = \max_{\boldsymbol{w}_t \in \boldsymbol{W}_t(a)} \{\phi'_{t,x}(\boldsymbol{w}_t|a) : u_{t,\widehat{x}} - u_{t,x} \geq w_{t,\widehat{x}} - w_{t,x},\ \widehat{x} \in X\},$$
$$\text{or} \quad \phi^*_{t,x}(\boldsymbol{u}_t|a) = \max_{\boldsymbol{w}_t \in \boldsymbol{W}_t(a)} \{\phi'_{t,x}(\boldsymbol{w}_t|a) : \boldsymbol{u}_t \geq \boldsymbol{w}_t\} \tag{35}$$

using the relative forms of $\boldsymbol{u}_t$ and $\boldsymbol{w}_t$ with respect to state $x$.

**Stage 3: Solving (35) by Projection.** The state-action-option problem (26)-(27) boils down to (35), which describes a simple relationship: call a vector $(\boldsymbol{w}^d_t, 0)$ in the $(\boldsymbol{w}_t, \phi)$-space a *feasible direction* if $\boldsymbol{w}^d_t \geq \boldsymbol{0}$ and $\|\boldsymbol{w}^d_t\| = 1$; if we translate $\phi'_{t,x}(\boldsymbol{w}_t|a)$ along all feasible directions, the outer contour of the traces forms $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ (renaming $\boldsymbol{w}_t$ to $\boldsymbol{u}_t$). Figure 4 illustrates

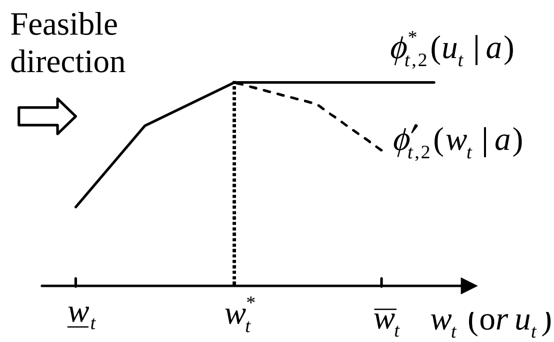


Figure 4: From $\phi'_{t,x}(\boldsymbol{w}_t|a)$ to $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ by projection along the $w_{t,x}$-axes.

a two-dimensional example in state 2, where $\phi^*_{t,2}(u_t|a) = \max_{w_t \in [\underline{w}_t, \overline{w}_t]} \left\{ \phi'_{t,2}(w_t|a) : u_t \geq w_t \right\}$, $u_t = u_{t,1} - u_{t,2}$ and $w_t = w_{t,1} - w_{t,2}$. An algorithm to solve problem (35) can be found in the appendix.

Finally, a few remarks about the time complexity of our algorithm are in order. It crucially depends on the time complexity of solving each standard computational geometry problem, which varies according to the techniques chosen. Some standard algorithms such as the various convex hull algorithms do not have polynomial-time complexity though they work very well for small-dimensional problems. Based on that, we expect the algorithm developed in this paper to be practical for small-sized problems—with small state and action sets, but to become computationally intractable for large scale problems. The development of computationally efficient algorithms that would apply to large scale problems is a fruitful avenue for future research.

# 5 Two-State Case and Numerical Examples

The special case with two states is the simplest case yet it still preserves the main features of the general model. Many insights obtained from the two-state case can be generalized to the general case. Furthermore, it is easy to implement and illustrate. In this section, we first present the two-state problem and then discuss two numerical examples. The main goal is to illustrate some of the complex trade-offs underlying the design of an optimal contract. The examples are necessarily artificial because a full exploration of a more realistic example would be beyond the scope of this paper and is left as a topic for future research.

**The Two-State Problem.** Using relative utilities and defining $u_t = u_{t,1} - u_{t,2}$, $u_{t+1} = u_{t+1,1} - u_{t+1,2}$, we can rewrite the state-action-option problem (26)-(27) as follows: for state 1,

$$\phi^*_{t,1}(u_t|a) = \max_{u_{t+1} \in U_{t+1}} r_1(a) - c_1(a) + \delta \boldsymbol{p}_1(a) \boldsymbol{\phi}^*_{t+1}(u_{t+1}) \tag{36}$$

$$\text{s.t. } u_t \leq c_2(a) - c_1(a) + \delta \left[p_{11}(a) - p_{21}(a)\right] u_{t+1}, \tag{37}$$

and for state 2,

$$\phi^*_{t,2}(u_t|a) = \max_{u_{t+1} \in U_{t+1}} r_2(a) - c_2(a) + \delta \boldsymbol{p}_2(a) \boldsymbol{\phi}^*_{t+1}(u_{t+1}) \tag{38}$$

$$\text{s.t. } u_t \geq c_2(a) - c_1(a) + \delta \left[p_{11}(a) - p_{21}(a)\right] u_{t+1}. \tag{39}$$

Note that we use the same definition $u_t = u_{t,1} - u_{t,2}$ in both problems so that the functions $\phi^*_{t,1}(u_t|a)$, $\phi^*_{t,2}(u_t|a)$, $\phi^*_{t,1}(u_t)$ and $\phi^*_{t,2}(u_t)$ can be depicted in the same chart.

The computational geometry implementation discussed in the last section can be significantly simplified in the two-state case. The main stages have been illustrated in previous figures: Figure 3 illustrates how to obtain $\boldsymbol{p}_x(a)\boldsymbol{\phi}^*_{t+1}(u_{t+1})$ from $\boldsymbol{\phi}^*_{t+1}(u_{t+1})$; Figure 4 shows how to compute $\phi^*_{t,x}(u_t|a)$ from $\phi'_{t,x}(w_t|a)$; and Figure 2(b) shows how to obtain $\phi^*_{t,x}(u_t)$ from $\phi^*_{t,x}(u_t|a)$'s. The principal's one-shot optimization problem at time 1 can be transformed into a single-variable optimization problem as well. In the two-state case, one can also focus on the breakpoints rather than the defining hyperplanes of the utility functions (a breakpoint is where the slope of a piecewise-linear function changes).

**Numerical Examples.** We present two numerical examples, assuming two states, two actions, and two periods. The parameters for the examples and the optimal contracts are summarized in Figure 5. The upper left table in the figure provides the model parameters, the upper right table presents the optimal action strategies and optimal social welfare functions for the first-best case with no hidden information (these provide a useful benchmark for our discussion of the results), and the two trees at the bottom summarize the optimal long-term contracts for the two examples when states are hidden.

First, let's describe the convention used in the trees. Round nodes denote state reports while diamond nodes denote actions prescribed by the contract. Randomization over actions is indicated by multiple branches spreading out of a diamond node, and the italicized numbers next to the branches provide the probabilities. The numbers in circles represent reported states, while the numbers in square brackets represent single-period payments made by the principal.

**Example Parameters**

| | I. $\delta=1, \beta_1=0.15$ | | II. $\delta=1, \beta_1=0.5$ | |
|---|---|---|---|---|
| | Action | | Action | |
| **State** | **1** | **2** | **1** | **2** |
| **1** | 2, 2<br>0.5 | 0, 3<br>0.8 | 0, 0<br>0.9 | 2, 1<br>0.25 |
| **2** | 5, 0<br>0.9 | 3, 0<br>1 | 1, 2<br>0.15 | -1, 0<br>0.75 |

**First-Best Case of Observable-State System**

| | Example I | | | | Example II | | | |
|---|---|---|---|---|---|---|---|---|
| | Action | | Social Welfare | | Action | | Social Welfare | |
| Period | $a^*_{t,1}$ | $a^*_{t,2}$ | $\phi^*_{t,1}$ | $\phi^*_{t,2}$ | $a^*_{t,1}$ | $a^*_{t,2}$ | $\phi^*_{t,1}$ | $\phi^*_{t,2}$ |
| **1** | 1 | 1 | 5.5 | 7.3 | 2 | 1 | 3.25 | 2.15 |
| **2** | 1 | 1 | 2 | 5 | 2 | 1 | 2 | 1 |

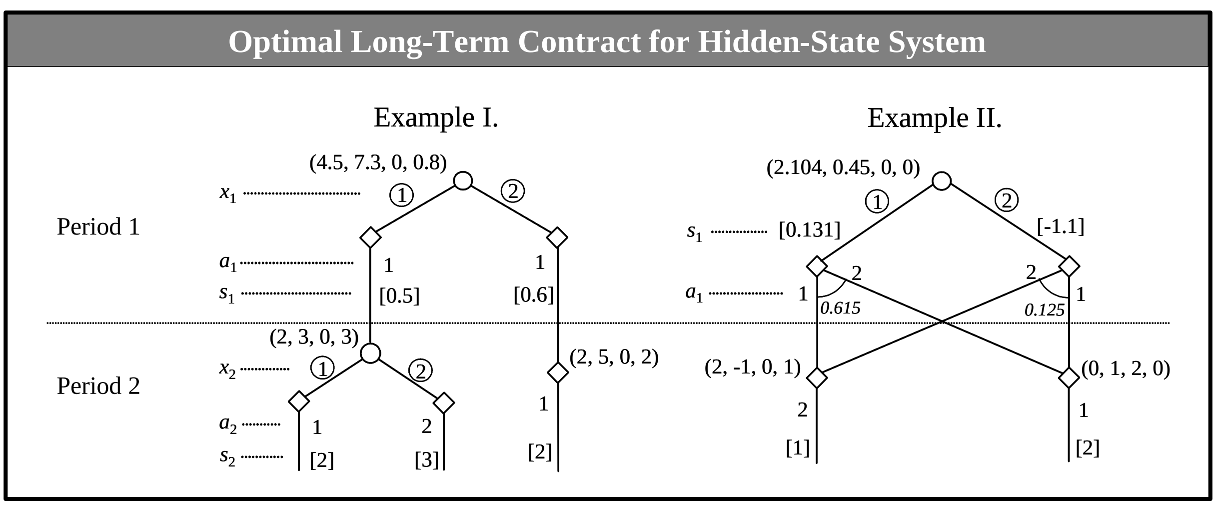


Figure 5: Parameters and Results for Two-State Examples. The triple in each cell of the parameter table gives $(r_x(a) - c_x(a), c_x(a), p_{x1}(a))$. Each quadruple in a tree provides the continuation social welfare vector and agent's utility vector at the beginning of a period, i.e., $(\phi^*_{t,1}, \phi^*_{t,2}, u_{t,1}, u_{t,2})$.

**Example I.** In this example, if all information is public, the first-best contract is to enforce action 1 in all circumstances, which creates the highest social welfare vector $\boldsymbol{\phi}_1 = (5.5, 7.3)$ at time 1 (recall the formula (25) for the first-best case). But if the states are hidden, to enforce the first-best action plan, the principal must pay the agent a premium to ensure his incentive compatibility and participation. It turns out the principal has to offer the agent a continuation utility vector $\boldsymbol{u}_1 = (0, 1.2)$ at time-1. As a result, her own time-1 continuation utility vector is reduced to $\boldsymbol{\pi}_1 = (5.5, 6.1)$. The agent's surplus beyond his reservation utility (due to his information advantage) is called the agent's *information rent*, which represents a monetary transfer from the principal to the agent and reduces the principal's utility. Thus, although the first-best action plan maximizes the social welfare, it also leaves significant information rent for the agent when states are hidden. Lowering the agent's information rent by a more sophisticated contract will no doubt result in social welfare losses. To maximize the principal's utility under hidden states, an optimal contract must strike a balance between these opposite forces.

In our example, the optimal contract operates as follows: In period 1, the agent reports the state. If the report is 1, he must take action 1 and receive a payment 0.5, followed by further differentiation in period 2; If the report is 2, he should also take action 1 but will be paid 0.6 and face no further differentiation in period 2. This optimal contract is not unique (it is derived by adding ex-post participation constraints at time 2). Only the agent's relative utility at time 2 really matters, due to the redundancy in his utility space.

Notice that the agent's time-1-state-2 information rent is reduced from 1.2 to 0.8. The positive information rent for state 2 at time 1 suggests that the main incentive issue in the first period is the agent's temptation to misreport state 2 as state 1. The decline of the information rent implies that this incentive problem has been alleviated. This is achieved by the design of the time-2 continuation contracts. The transition probabilities $p_{11}(1) = 0.5 < p_{21}(1) = 0.9$ implies that in the first period, the time-2 relative utility $u_{2,1} - u_{2,2}$ is more valuable to the agent in state 2 than in state 1. Thus, in order to discourage the agent from misreporting the first period state $x_1$, the time-2 continuation contract following $x_1 = 2$ should provide higher $u_{2,1} - u_{2,2}$ (which is $-2$ in the example) than the one following $x_1 = 1$ (which is $-3$). But on the other hand, the continuation contract following state $x_1 = 1$ deviates from the first-best action plan and causes a social welfare loss. The resulting principal's continuation utility vector at time 1 is $\boldsymbol{\pi}_1 = (4.5, 6.5)$, as opposed to $(5.5, 6.1)$ if the first-best action plan is enforced. The assumption of a small $\beta_1$ (0.15 in the example) is useful now, which downplays the social welfare loss in state $x_1 = 1$. One can see that when $4.5\beta_1 + 6.5(1 - \beta_1) < 5.5\beta_1 + 6.1(1 - \beta_1)$, or $\beta_1 > 0.286$, the social welfare loss will dominate the benefit of information rent reduction and the first-best action plan will become optimal even when states are hidden.

This example demonstrates how the principal can use her commitment to future payments together with a judicially designed revelation mechanism to extract some information rent from the agent and achieve outcomes close to first-best. Notice that the state report in the first period is necessary. As the agent is asked to take the same action in period 1, if the contract were solely based on the action history, the principal could not have acquired any useful information about the first period.

**Example II.** This example demonstrates an instance where randomization is needed. In this example, the optimal contract requires state revelation in period 1, followed by different randomizations among first period actions, and the action in period 2 depends on the public information realized in the first period. Notice that in the figure, we specify the first period payments before actions, because there is a degree of freedom in the payments and only the expectation of the first period payments matters.

Again, the contract uses the revelation mechanism to extract the agent's surplus. This reduces the overall social welfare but maximizes the principal's utility. Further, contrasting the welfare losses in example II to those in example I, one notices that the losses in the second example are more substantial and extend to most of the periods and states. This result is driven by the underlying cost and probability structure of the system. There is a significant misalignment between the two parties' interests in example II: the agent prefers action 1 in state 1 and action 2 in state 2 due to the cost differences, which is exactly opposite to the first-best action plan.

The two examples highlight some features of the optimal long-term contracts: they strike a balance between social welfare maximization (increasing the size of the pie) and information rent minimization (decreasing the agent's share of the pie); revelation contracts offer the principal the right degree of flexibility as opposed to simpler contracts in which the payments are based solely on the agent's actions; and randomization may be necessary to achieve optimality.

# 6 Conclusion

In this paper, we have proposed a general framework for a large class of dynamic principal-agent problems with hidden information. We have developed a dynamic programming algorithm to derive optimal long-term contracts for the principal. The principal indirectly controls the underlying system by choosing the agent's continuation utility vectors along public information paths. It induces truthful state revelation and results in actions that maximize the principal's expected payoff. This problem is significant to the Operations Research community and has many applications of interest.

Due to space limitation, the paper only discusses a general solution of the model. Lack of structure in the parameters in the general case results in optimal contracts with complex characteristics, which may be hard to implement. Problems in reality often possess strong properties which can be successfully exploited to arrive at implementable optimal contracts. For instance, in the manufacturer-retailer application discussed in Section 3, a higher inventory level in one period should result in (weakly) higher initial inventory in the next period, and it should cost the manufacturer more to produce more products. An important topic is therefore to identify conditions under which the optimal long-term contracts admit simple structures. A useful tool is the theory of supermodularity and complementarity, as developed in Topkis (1998). Indeed, the existing contract theory literature has focused on problems with various special structures. It is left for future research to examine whether the general solution developed in this paper can shed new light on the known special cases and help the investigation of new ones.

A related topic is to study special contracts with simple formats, such as linear contracts. These contracts are sub-optimal in general but may perform quite well in certain circumstances and, more importantly, they are closer to the contracts used in reality.

The analysis in the paper assumes that the principal can commit to future payment plans. The solution does not apply when the principal cannot make full commitment to a long-term contract and thus the contract is prone to renegotiation. An interesting extension of this paper is to study situations where renegotiation is permitted or only short-term contracts are possible.

The dynamic programming algorithm suggests that the infinite horizon problem is also solvable and the existence of an optimal solution can be established by contraction mappings on the space of continuation utilities. The infinite horizon problem can be of considerable theoretical and practical interest if the solution is less cumbersome than the finite horizon problem.

In conclusion, there are significant research opportunities for the operations research community in the area of dynamic principal-agent problems and contract theory in general. As discussed above, this study suggests several important extensions. We postulate that under suitable circumstances, the general framework and approach developed in this paper can provide a useful benchmark for future research works.

# Appendix A: Road Map of the Paper

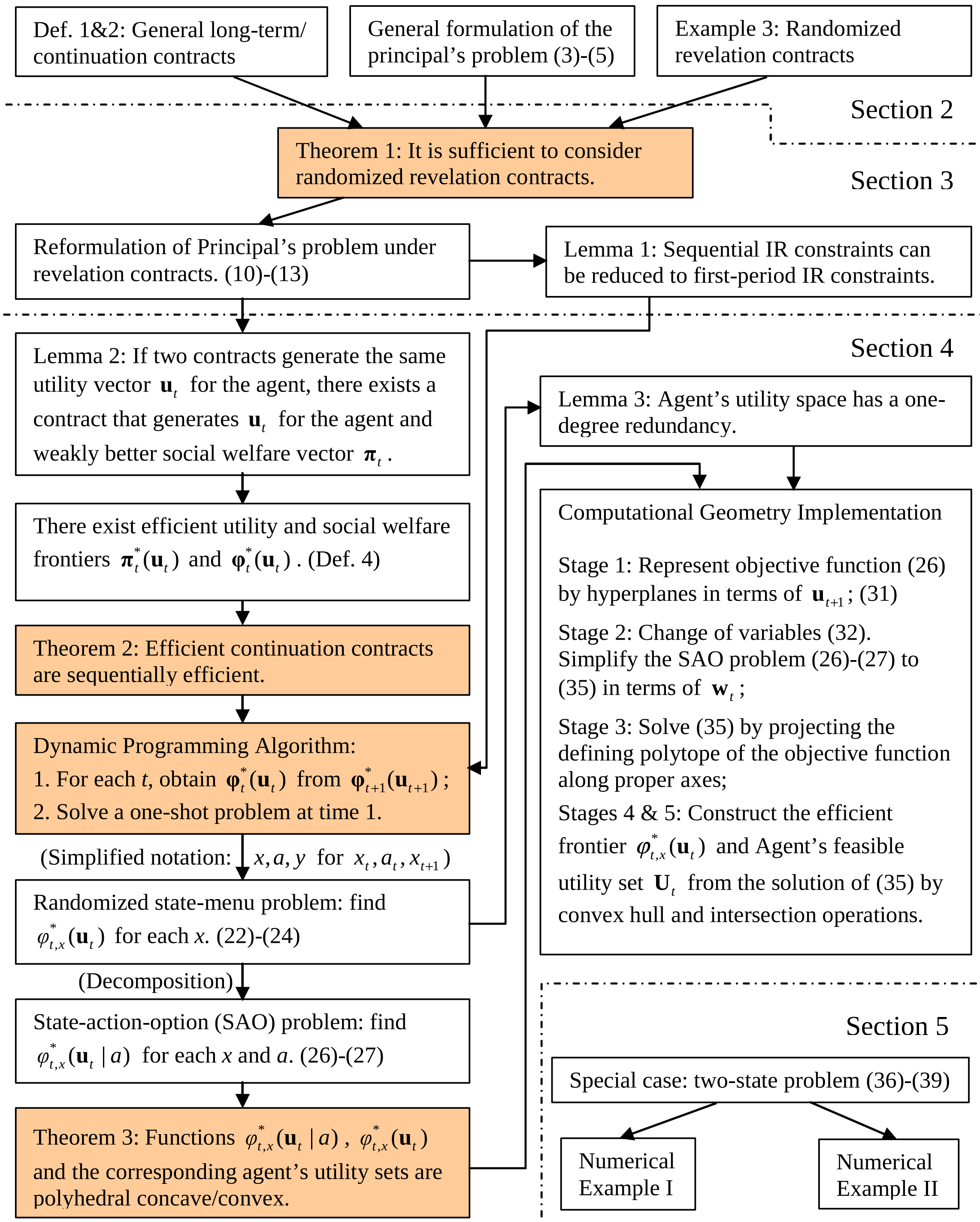


Figure 6: Road Map of the Paper.

# Appendix B: Proofs of Lemmas and Theorems

Proof of Theorem 1. The proof is by construction. Consider a long-term contract $\sigma$ coupled with the agent's best response strategy $\xi^*$. Since $\xi^*$ is well-defined, the $(\sigma, \xi^*)$ pair induces a cumulative distribution function $F(\cdot)$ of the random stream $(x_1, a_1, s_1, x_2 \cdots, a_T, s_T, x_{T+1})$. Notice that since $X$ and $A$ are finite sets, $dF(x_t)$ and $dF\,(a_t)$ represent the probability masses at $x_t \in X$ and $a_t \in A$ respectively. Further, the underlying system requires $dF(x_1) = \beta_{x_1}$ for $x_1 \in X$.

We construct the desired truthful revelation contract $\sigma^*$ below, in three steps:

**1. Constructing a feasible $\sigma^*$.** For any period $t$, define:

$$\theta^*_t(x_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) \triangleq dF(a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t),$$
$$s^*_t(x_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) \triangleq \int_{\mathbb{R}} s_t dF(s_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t, a_t).$$

Since $\int_A dF(a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t) = 1$, the resulting contract $\sigma^* = \{\theta^*_1(x_1, a_1), s^*_1(x_1, a_1); \cdots; \theta^*_T(x_T, a_T|\boldsymbol{x}^{T-1}, \boldsymbol{a}^{T-1}), s^*_T(x_T, a_T|\boldsymbol{x}^{T-1}, \boldsymbol{a}^{T-1})\}_{\boldsymbol{x}^T \in X^T, \boldsymbol{a}^T \in A^T}$ is a legitimate revelation contract. If the agent reports truthfully under $\sigma^*$, it will induce the same marginal distribution of $(\boldsymbol{x}^{T+1}, \boldsymbol{a}^T)$ and the same marginal distribution of $\boldsymbol{s}^T$ conditional on $(\boldsymbol{x}^{T+1}, \boldsymbol{a}^T)$ as those induced by $(\sigma, \xi^*)$.

**2. Comparing $\sigma^*$ with $\sigma$.** We show that if $\sigma^*$ is truthful, it will generate the same expected future utilities for the two parties as $\sigma$ does. If the agent reports truthfully, his time-$t$ continuation utility, as given in (7), should satisfy:

$$\begin{aligned} u_t(x_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) \;=\; & \int_{\mathbb{R}} s_t dF(s_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t) - \int_A c_{x_t}(a_t) dF(a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t) + \\ & \delta \int_{A \times X} u_{t+1}(x_{t+1}|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t, a_t) dF(a_t, x_{t+1}|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t), \end{aligned}$$

where we have used the fact that $p_{x_t, x_{t+1}}(a_t) dF(a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t) = dF(a_t, x_{t+1}|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t)$, as required by the underlying system. On the other hand, the agent's expected future utility under the original contract $\sigma$ following history $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t)$ is given by:

$$\begin{aligned} & E\left(\sum_{\tau=t}^{T} \delta^{\tau-t}(s_\tau - c_{x_\tau}(a_\tau)) + \delta^{T+1-t} u_{x_{T+1}} | \boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t\right) \\ = \; & E\left(s_t | \boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t\right) - E\left(c_{x_t}(a_t) | \boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t\right) + \\ & \delta E\left(E\left(\sum_{\tau=t+1}^{T} \delta^{\tau-t-1}(s_\tau - c_{x_\tau}(a_\tau)) + \delta^{T-t} u_{x_{T+1}} | \boldsymbol{x}^{t}, \boldsymbol{a}^{t}, x_{t+1}\right) | \boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t\right), \end{aligned}$$

which is equal to $u_t(x_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$. That is, if the agent reports truthfully under $\sigma^*$, his expected future utility following history $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t)$ will be the same as that under the original contract $\sigma$. The same can be shown for the principal's expected future utilities.

**3. Showing $\sigma^*$ truthful.** This is done by backward induction. Express $\sigma^*$ in its recursive form as a collection of continuation contracts. Suppose the continuation contracts $\sigma^*_{t+1}(\boldsymbol{x}^t, \boldsymbol{a}^t)$ are truthful for all $(\boldsymbol{x}^t, \boldsymbol{a}^t)$. Consider the continuation contract $\sigma^*_t(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) = \left\{\theta^*_t(x_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}), s^*_t(x_t, a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}), \sigma^*_{t+1}(\boldsymbol{x}^t, \boldsymbol{a}^t)\right\}_{x_t \in X, a_t \in A}$ for a given $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$. If the true state is $x_t$ but the agent reports $\widehat{x}_t$ (and truth-telling after time $t+1$, since $\sigma^*_{t+1}(\cdot)$ are truthful), his time-$t$ expected future utility, as computed in (9), should satisfy:

$$\begin{aligned} \widehat{u}_t(x_t, \widehat{x}_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) = & \int_{\mathbb{R}} s_t dF(s_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t) - \int_A c_{x_t}(a_t) dF(a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t) + \\ & \delta \int_A \sum_{x_{t+1} \in X} p_{x_t, x_{t+1}}(a_t) u_{t+1}(x_{t+1}|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t, a_t) dF(a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t). \end{aligned}$$

If $\sigma^*_t(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$ is not truthful, there must exist states $x_t$ and $\widehat{x}_t$ such that $u_t(x_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) < \widehat{u}_t(x_t, \widehat{x}_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$. We show that this is inconsistent with the original contract. To that end, it suffices to show that the above expected utility $\widehat{u}_t(x_t, \widehat{x}_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$ can be achieved by the agent under the original contract $\sigma$ as well.

Call the full information history $\boldsymbol{\omega}^t$ under contract $\sigma$ *compatible* with the partial history $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t)$ if it includes $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t)$. For any such $\boldsymbol{\omega}^t$, there exists a corresponding full information history $\widehat{\boldsymbol{\omega}}^t$ which is the same as $\boldsymbol{\omega}^t$ except that its last component is $\widehat{x}_t$ instead of $x_t$. There is a one-to-one correspondence between $\boldsymbol{\omega}^t$ and $\widehat{\boldsymbol{\omega}}^t$. Consider a modified strategy $\widehat{\xi}$ of the agent which is the same as $\xi^*$ except that after each compatible $\boldsymbol{\omega}^t$, following the part of $\xi^*$ that originally follows the corresponding $\widehat{\boldsymbol{\omega}}^t$. In other words, after a compatible history $\boldsymbol{\omega}^t$, the agent pretends that the history is $\widehat{\boldsymbol{\omega}}^t$. Then, under the modified strategy $\widehat{\xi}$, the distribution of the essential information after the partial history $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t)$ satisfies: (1) The conditional joint distribution of $a_t$ and $s_t$ is given by $F(a_t, s_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t)$, because the time-$t$ continuation contracts are executed as if the partial history is $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t)$; (2) The conditional joint distribution of $a_t$ and $x_{t+1}$ is given by $p_{x_t,x_{t+1}}(a_t)dF(a_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t)$, not $dF(a_t, x_{t+1}|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t)$, because the actual state transition originates from the true state $x_t$, not $\widehat{x}_t$; (3) The joint distribution of $(a_{t+1}, s_{t+1}, x_{t+2}, \cdots, a_T, s_T, x_{T+1})$ conditional on $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t, a_t, x_{t+1})$ is given by $F(\cdot|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t, a_t, x_{t+1})$, because the time-$(t+1)$ continuation contracts are executed as if the partial history is $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, \widehat{x}_t, a_t, x_{t+1})$ and the state transitions after time $t+1$ originate from $x_{t+1}$.

We can see that under the modified strategy $\widehat{\xi}$, the agent's expected future utility after the partial history $(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}, x_t)$ will be exactly $\widehat{u}_t(x_t, \widehat{x}_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$. Thus, $u_t(x_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1}) < \widehat{u}_t(x_t, \widehat{x}_t|\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$ implies that the agent's original strategy $\xi^*$ can be improved by $\widehat{\xi}$, which contradicts the assumption that $\xi^*$ is his best response strategy. Therefore, the continuation revelation contract $\sigma_t^*(\boldsymbol{x}^{t-1}, \boldsymbol{a}^{t-1})$ must be truthful.

The above argument also applies to period $T$ (noticing $u_{T+1}(x_{T+1}|\boldsymbol{x}^{T-1}, \boldsymbol{a}^{T-1}, \widehat{x}_T, a_T) = u_{x_{T+1}}$ and there is no further $(a_{T+1}, s_{T+1}, x_{T+2}, \cdots)$). This forms the basic step of the induction and thus completes the proof of the theorem. ■

Proof of Lemma 1. For simplicity, we only prove the result for the two-period case, but the argument can be easily extended to the general case. Suppose a revelation contract $\{\theta_1(x_1, a_1), s_1(x_1, a_1); \theta_2(x_2, a_2|x_1, a_1), s_2(x_2, a_2|x_1, a_1)\}_{\boldsymbol{x}^2\in X^2, \boldsymbol{a}^2\in A^2}$ satisfies the IR constraints in period 1 but violates the ones in period 2. Construct another contract $\{\theta_1(x_1, a_1), s_1^\dagger(x_1, a_1);\ \theta_2(x_2, a_2|x_1, a_1), s_2^\dagger(x_2, a_2|x_1, a_1)\}_{\boldsymbol{x}^2\in X^2, \boldsymbol{a}^2\in A^2}$ by postponing a part of the period-1 payments to period 2 as follows:

$$\begin{aligned} s_1^\dagger(x_1, a_1) &= s_1(x_1, a_1) - \delta d_{x_1}, \qquad x_1 \in X, a_1 \in A \\ s_2^\dagger(x_2, a_2|x_1, a_1) &= s_2(x_2, a_2|x_1, a_1) + d_{x_1}, \quad \boldsymbol{x}^2 \in X^2, \boldsymbol{a}^2 \in A^2 \end{aligned}$$

for a set of constants $d_{x_1}$. Straightforward algebra shows that the time-1 utility vectors stay the same, so the IR and IC constraints still hold in the first period under the new contract. By choosing $d_{x_1}$ large enough we can satisfy the period-2 IR constraints as well. Incentive compatibility in period 2 is intact because the agent's utility is elevated by a constant $d_{x_1}$ in all state $x_2$. Therefore, without changing the continuation utilities for the two parties at time 1, we can always transform a truthful revelation contract that satisfies the first period IR constraints into a truthful revelation contract that satisfies the SIR constraints. ■

PROOF OF THEOREM 2. The proof is by contradiction. Suppose that for some period $\tau > t$ and history $\boldsymbol{h}^\tau$, the continuation contract $\sigma_\tau$ is inefficient. By Lemma 2 and the definition of efficiency, there exists an efficient truthful revelation contract $\sigma^*_\tau$ such that $\boldsymbol{u}(\sigma^*_\tau) = \boldsymbol{u}(\sigma_\tau)$ and $\boldsymbol{\pi}(\sigma^*_\tau) \geq \boldsymbol{\pi}(\sigma_\tau)$ with ">" for at least one state. The principal will be better off by replacing $\sigma_\tau$ with $\sigma^*_\tau$. This does not affect the agent's incentives in any period but improves the principal's time-$t$ continuation utilities. It contradicts the efficiency of $\sigma_t$. ■

PROOF OF THEOREM 3. The proof is by induction. Suppose the theorem holds for period $t+1$. Specifically, suppose the time-$(t+1)$ efficient social welfare frontier satisfies: $\boldsymbol{U}_{t+1} = \{\boldsymbol{u}_{t+1} : \boldsymbol{A}\boldsymbol{u}_{t+1} \leq \boldsymbol{b}\}$ for some matrix $\boldsymbol{A}$ and vector $\boldsymbol{b}$; for every $y \in X$ ($y$ stands for $x_{t+1}$), $\phi^*_{t+1,y}(\boldsymbol{u}_{t+1}) = \min_{i=1,\cdots,I_y}\left(\boldsymbol{h}^i_y \boldsymbol{u}_{t+1} + d^i_y\right)$ for some vectors $\boldsymbol{h}^i_y$ and scalars $d^i_y$. The hypograph of $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ is given by $\mathcal{H}^*_{t,x}(a) = \{(\boldsymbol{u}_t, \phi_t) : \phi_t \leq \phi^*_{t,x}(\boldsymbol{u}_t|a)\}$ and define the *augmented hypograph* of $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ as

$$\widetilde{\mathcal{H}}^*_{t,x}(a) = \left\{\begin{array}{ll} (\boldsymbol{u}_t, \boldsymbol{u}_{t+1}, \phi_t, \boldsymbol{\phi}_{t+1}) : & \phi_t \leq r_x(a) - c_x(a) + \delta \boldsymbol{p}_x(a)\boldsymbol{\phi}_{t+1}; \\ \quad u_{t,\widehat{x}} - u_{t,x} \geq c_x(a) - c_{\widehat{x}}(a) + \delta\left[\boldsymbol{p}_{\widehat{x}}(a) - \boldsymbol{p}_x(a)\right]\boldsymbol{u}_{t+1}, \ \widehat{x} \in X; & \\ \quad \phi_{t+1,y} \leq \boldsymbol{h}^i_y \boldsymbol{u}_{t+1} + d^i_y, \quad i = 1, \cdots, I_y, \ y \in X; & \boldsymbol{A}\boldsymbol{u}_{t+1} \leq \boldsymbol{b} \end{array}\right\}.$$

Clearly, $\mathcal{H}^*_{t,x}(a) = \{(\boldsymbol{u}_t, \phi_t) : \exists(\boldsymbol{u}_t, \boldsymbol{u}_{t+1}, \phi_t, \boldsymbol{\phi}_{t+1}) \in \widetilde{\mathcal{H}}^*_{t,x}(a)\}$ and hence $\mathcal{H}^*_{t,x}(a)$ is the projection of $\widetilde{\mathcal{H}}^*_{t,x}(a)$ to the $(\boldsymbol{u}_t, \phi_t)$-space. Since $\widetilde{\mathcal{H}}^*_{t,x}(a)$ is a polyhedral convex set, its projection is also polyhedral convex. Thus, part (a) of the theorem follows: $\boldsymbol{U}_{t,x}(a)$ is polyhedral convex and $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ is polyhedral concave in $\boldsymbol{u}_t$.

Part (b) follows from equation (28), which implies that $\boldsymbol{U}_{t,x} = conv\left(\cup_{a\in A}\boldsymbol{U}_{t,x}(a)\right)$ and $hypo(\phi^*_{t,x}(\boldsymbol{u}_t)) = conv(\cup_{a\in A} hypo(\phi^*_{t,x}(\boldsymbol{u}_t|a)))$. Convexification preserves polyhedral convexity/concavity, and hence the desired properties are passed on to $\boldsymbol{U}_{t,x}$ and $\phi^*_{t,x}(\boldsymbol{u}_t)$. Part (c) follows from the following fact: when constraints (23)-(24) for all $x \in X$ are combined together, they constitute a partition of the constraint set for a truthful revelation contract, and therefore the set of agent's utility vectors $\boldsymbol{U}_t = \cap_{x\in X} \boldsymbol{U}_{t,x}$.

This completes the induction step. The basic step, to show that $\boldsymbol{U}_T$ is polyhedral convex and $\phi^*_{T,x}(\boldsymbol{u}_T)$ is polyhedral concave, follows from the fact that $\boldsymbol{u}_{T+1}$ and $\boldsymbol{\phi}_{T+1}$ in the definition of $\widetilde{\mathcal{H}}^*_{T,x}(a)$ are exogenously fixed and thus $\widetilde{\mathcal{H}}^*_{T,x}(a)$ is a polyhedral convex set. Therefore, the theorem is proved. $\blacksquare$

# Appendix C: Supplement to Computational Geometry Implementation

In this appendix, we provide supplementary materials for Subsection 4.5. For more discussions on computational geometry and polytopes, see Berg et al (2000), Boissonnat and Yvinec (1998), and Fukuda and Weibel (2005).

**Introduction to Polytopes.** A *polytope* can be represented in two equivalent ways: it is the convex hull of a finite number of points in $\mathbb{R}^d$, i.e., $P = conv\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_n\} = \{\sum_{i=1}^n \lambda_i \boldsymbol{v}_i : \sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0, i = 1, \cdots, n\}$, which can be generalized to $P = \{\sum_{i=1}^n \lambda_i \boldsymbol{v}_i : \sum_{i=1}^k \lambda_i = 1, \lambda_i \geq 0, i = 1, \cdots, n\}$ for some $k \leq n$; and it is the intersection of a finite number of closed half-spaces, i.e., $P = \{\boldsymbol{v} \in \mathbb{R}^d : \boldsymbol{A}\boldsymbol{v} \leq \boldsymbol{b}\}$.

A subset $F$ is a *face* of $P$ if there is a vector $\boldsymbol{a}$ and scalar $b$ such that $\boldsymbol{a}\boldsymbol{v} \leq b$ for all $\boldsymbol{v} \in P$ and $F = P \cap \{\boldsymbol{v} \in \mathbb{R}^d : \boldsymbol{a}\boldsymbol{v} = b\}$. A $d$-dimensional polytope has $d$ types of (proper) faces, with dimensions $d-1$, $d-2$, $\cdots$, and 0 respectively. A $(d-1)$-dimensional face is called a *facet* and a 0-dimensional face is called a *vertex*. (Hyperplanes and points correspond to the facets and vertices of a polytope.) Every $k$-face is the intersection of some $(k+1)$-faces, for $k \leq d-2$. A $k$-face and a $(k+1)$-face are *incident* if the former lies on the latter. A data structure to represent a polytope completely is an *incidence graph*, which has $d$ levels and records the pairwise incidence relationships. Each node of the graph represent a face and an arc links two incident nodes. The *upper bound theorem* states that if a $d$-polytope has $n$ vertices (or facets), the total number of faces is at most $O(n^{\lfloor d/2 \rfloor})$, as well as the size of the incidence graph.

The two representations of a polytope are dual to each other, but conversions between them are non-trivial problems, namely the *facet-enumeration* (or *convex hull*) problem and the *vertex-enumeration* problem, respectively. A generic convex hull algorithm for $n$ points takes $O(n^{\lfloor d/2 \rfloor})$ time. (From one of the two representations plus the adjacency information, we can construct the complete incidence graph in $O(n^{\lfloor d/2 \rfloor})$ time.) There are algorithms available on the Internet, such as Qhull and Fukuda's cdd program. On a 600 MHz Pentium 3 computer, Qhull can compute a typical 5-d convex hull of 6,000 points in 12 seconds and a typical 8-d convex hull of 120 points in 15 seconds. Many real-world problems are often very different from the worst output cases and some can be solved with high dimensions well above ten, some even over sixty.

**Stage 1 of the Implementation.** It is known that (see for example Berg et al (2000)) the lower envelope of hyperplanes $\{\phi = \boldsymbol{h}^i \boldsymbol{u} + d^i\}_{i=1,\cdots,I}$ corresponds to the lower convex hull of the points $\{(\boldsymbol{h}^i, d^i)\}_{i=1,\cdots,I}$, as illustrated in Figure 7. More precisely, a hyperplane $\phi = \boldsymbol{h}^{i_k} \boldsymbol{u} + d^{i_k}$ defines a facet of the lower envelope of $\{\phi = \boldsymbol{h}^i \boldsymbol{u} + d^i\}_{i=1,\cdots,I}$ if and only if the point $(\boldsymbol{h}^{i_k}, d^{i_k})$ defines a vertex of the lower convex hull of $\{(\boldsymbol{h}^i, d^i)\}_{i=1,\cdots,I}$. Note that a standard definition of the dual of $\phi = \boldsymbol{h}\boldsymbol{u} + d$ is the point $(\boldsymbol{h}, -d)$; we drop the negative sign to make it more natural without changing any essential result.

Any polyhedral concave function can be represented by the lower envelope of a set of hyperplanes, the defining hyperplanes. Define the *sum* of two point sets $\{(\boldsymbol{h}_1^i, d_1^i)\}_{i=1,\cdots,I_1}$ and $\{(\boldsymbol{h}_2^i, d_2^i)\}_{i=1,\cdots,I_2}$ as $\{(\boldsymbol{h}_1^{i_1} + \boldsymbol{h}_2^{i_2}, d_1^{i_1} + d_2^{i_2})\}_{i_1=1,\cdots,I_1; i_2=1,\cdots,I_2}$. The following result shows that finding the defining hyperplanes of $\boldsymbol{p}_x(a)\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1})$ is equivalent to finding a convex hull in the dual space:

**Lemma 4 (Transformation)** *Suppose* $\phi^*_{t+1,y}(\boldsymbol{u}_{t+1}) = \min_{i=1,\cdots,I_y} \left(\boldsymbol{h}_y^i \boldsymbol{u}_{t+1} + d_y^i\right),\ y \in X.$

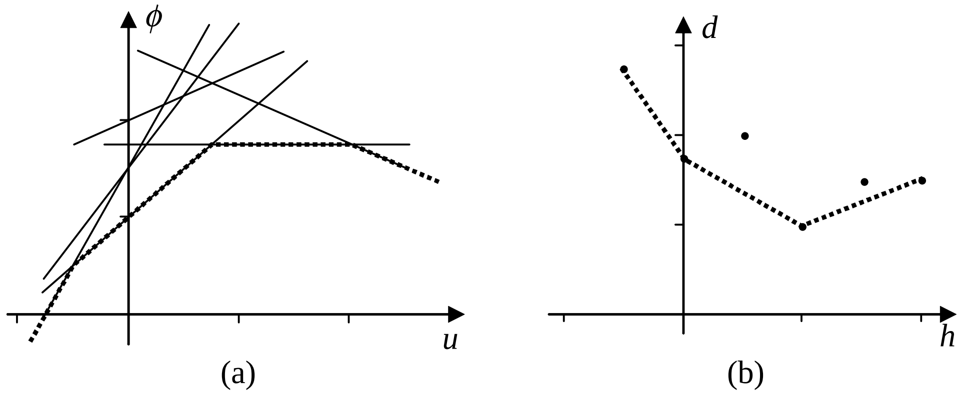


Figure 7: Duality between hyperplanes and points. (a) The lower envelope of hyperplanes; (b) The lower convex hull of points.

*The lower envelope defining the function $\boldsymbol{\lambda}\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1})$, with $\boldsymbol{\lambda} \geq \boldsymbol{0}$, corresponds to the lower convex hull of the sum of $\{(\lambda_y \boldsymbol{h}^i_y, \lambda_y d^i_y)\}_{i=1,\cdots,I_y}$, $y \in X$.*

*Proof.* By definition of $\phi^*_{t+1,y}(\boldsymbol{u}_{t+1})$ and because $\boldsymbol{\lambda} \geq \boldsymbol{0}$, we have

$$\begin{aligned}\boldsymbol{\lambda}\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1}) &= \sum_{y\in X} \lambda_y \min_{i=1,\cdots,I_y} \{\boldsymbol{h}^i_y \boldsymbol{u}_{t+1} + d^i_y\} \\ &= \min_{i_y=1,\cdots,I_y, y\in X} \{(\sum_{y\in X} \lambda_y \boldsymbol{h}^{i_y}_y)\boldsymbol{u}_{t+1} + \sum_{y\in X} \lambda_y d^{i_y}_y\}.\end{aligned}$$

Hence, $\boldsymbol{\lambda}\boldsymbol{\phi}^*_{t+1}(\boldsymbol{u}_{t+1})$ is the lower envelope of $\{\phi = (\sum_{y\in X} \lambda_y \boldsymbol{h}^{i_y}_y)\boldsymbol{u}_{t+1} + \sum_{y\in X} \lambda_y d^{i_y}_y\}_{i_y=1,\cdots,I_y,y\in X}$. These hyperplanes are dual to the points $\{(\sum_{y\in X} \lambda_y \boldsymbol{h}^{i_y}_y, \sum_{y\in X} \lambda_y d^{i_y}_y)\}_{i_y=1,\cdots,I_y,y\in X}$. The rest follows from the result of Berg et al. ■

Although finding the convex hull of a set of points is a standard problem, it may be inefficient when there are too many points—we have $\prod_{y\in X} I_y$ dual points here. A more efficient way to find the convex hull of the sum of $\{(\lambda_y \boldsymbol{h}^i_y, \lambda_y d^i_y)\}_{i=1,\cdots,I_y}$, $y \in X$, is to treat each point set $\{(\lambda_y \boldsymbol{h}^i_y, \lambda_y d^i_y)\}_{i=1,\cdots,I_y}$ as a polytope vertex set and compute the Minkowski sum of the set of polytopes. The *Minkowski sum* of polytopes is still a polytope, and is defined as $P_1 + P_2 = \{\boldsymbol{v} \in \mathbb{R}^n : \boldsymbol{v} = \boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_1 \in P_1, \boldsymbol{v}_2 \in P_2\}$ for polytopes $P_1$ and $P_2$. Fukuda (2004) proposes an algorithm to find the Minkowski sum of polytopes in time linear to the number of vertices of the resulting polytope.[2]

**Stage 2 of the Implementation.** If we have a degenerate case with $\text{rank}(\boldsymbol{P}(a)) = r < |X|$ for some action $a$, we can assume without loss of generality that $\boldsymbol{P}(a)$ admits the form $\begin{pmatrix} \boldsymbol{B} \\ \boldsymbol{D}\boldsymbol{B} \end{pmatrix}$ where $\boldsymbol{B}$ is a $r \times |X|$ matrix with full rank and $\boldsymbol{D}$ is a $(|X|-r) \times (|X|-r)$ matrix. Correspondingly, the cost vector can be split as $\boldsymbol{c}(a) = \begin{pmatrix} \boldsymbol{c}_B(a) \\ \boldsymbol{c}_D(a) \end{pmatrix}$. Then, the vector $\boldsymbol{w}_t$ takes the form $\begin{pmatrix} \boldsymbol{w}_t^B \\ \boldsymbol{w}_t^D \end{pmatrix}$ where $\boldsymbol{w}_t^B = -\boldsymbol{c}_B(a) + \delta \boldsymbol{B}\boldsymbol{u}_{t+1}$ (a degenerate version of (32)) and $\boldsymbol{w}_t^D = \boldsymbol{D}\boldsymbol{w}_t^B + (\boldsymbol{D}\boldsymbol{c}_B(a) - \boldsymbol{c}_D(a))$. The objective function (31) can be represented in terms of $\boldsymbol{w}_t^B$ by projecting its defining polytope from the $(\boldsymbol{u}_{t+1}, \phi)$-space to the $(\boldsymbol{w}_t^B, \phi)$-space, using an algorithm similar to the one presented in stage 3 below. Similarly, the polytope $\boldsymbol{A}\boldsymbol{u}_{t+1} \le \boldsymbol{b}$ can be projected from the $\boldsymbol{u}_{t+1}$-space to the $\boldsymbol{w}_t^B$-space to form the degenerate version of $\boldsymbol{W}_t(a)$.

To help understand the formulation (35) better, we verify that the transformed agent's utility sets $\boldsymbol{W}_t(a)$ and the defining hyperplanes of $\phi'_{t,x}(\boldsymbol{w}_t|a)$, as in (33) and (34), preserve the one-degree redundancy implied by Lemma 3. According to the lemma, the set $\boldsymbol{U}_{t+1} = \{\boldsymbol{u}_{t+1} : \boldsymbol{A}\boldsymbol{u}_{t+1} \le \boldsymbol{b}\}$ must satisfy $\boldsymbol{A}(\boldsymbol{u}_{t+1} + \lambda \boldsymbol{1}) \le \boldsymbol{b}$, for any $\lambda \in \mathbb{R}$ and $\boldsymbol{u}_{t+1} \in \boldsymbol{U}_{t+1}$. It follows that $\boldsymbol{A}\boldsymbol{1} = \boldsymbol{0}$. The transition matrices $\boldsymbol{P}(a)$ have row sums equal to 1, and hence $\boldsymbol{P}(a)^{-1}\boldsymbol{1} = \boldsymbol{P}(a)^{-1}\boldsymbol{P}(a)\boldsymbol{1} = \boldsymbol{1}$. Thus, $[\boldsymbol{A}\boldsymbol{P}(a)^{-1}]\,\boldsymbol{1} = \boldsymbol{0}$ and the set $\boldsymbol{W}_t(a)$ does have the desired redundancy. Similarly, the lemma also implies that each defining hyperplane $\phi = \boldsymbol{h}_x^i(a)\boldsymbol{u}_{t+1} + d_x^i(a)$ in (31) satisfies $\boldsymbol{h}_x^i(a)\boldsymbol{1} = 0$. It follows that $\left[\boldsymbol{h}_x^i(a)\boldsymbol{P}(a)^{-1}\right]\boldsymbol{1} = 0$ and hence the defining hyperplanes of $\phi'_{t,x}(\boldsymbol{w}_t|a)$ also have the desired redundancy.

**Stage 3 of the Implementation.** Berg et al (2000) provides an algorithm to augment an existing convex hull when a new vertex is added. It can be modified to solve the problem (35) by projection of the defining polytope $P'$ of $\phi'_{t,x}(\boldsymbol{w}_t|a)$: Color a facet of $P'$ *red* if it can be "seen" along all feasible directions (i.e., the normal vector of the facet is strictly positive) and *blue* otherwise. Since any face of $P'$ is the intersection of the facets which contain it, a

[2] Hao Zhang would like to thank Thomas McCormick for suggesting this paper.

face with a lower dimension can be colored accordingly: it is *red* if it is the intersection of red facets only, *blue* of blue facets only, or *purple* of red and blue facets. Then the defining polytope $P$ of $\phi^*_{t,x}(\boldsymbol{u}_t|a)$ is an unbounded polytope consisting of all the red faces of $P'$ and unbounded faces formed by projecting all purple faces of $P'$ along proper $w_{t,x}$-axes.

**Time Complexity.** In this last part, we briefly discuss the time complexity of the various stages. For each state $x$ and action $a$, stage 1 involves computing a Minkowski sum of $|X|$ polytopes in the dual space. The number of vertices of the resulting polytope in the dual space equals the number of facets in the primal space, which is given by $I_x(a)$ in expression (31). This stage requires solving $O(I_x(a))$ linear programs of the same size and thus takes time linear in $I_x(a)$, according to Fukuda (2004). In stage 2, transforming the $I_x(a)$ hyperplanes requires time linear in $I_x(a)$ as well. The projection in stage 3 can be done in $O(I_x(a)^{\lfloor |X|/2 \rfloor})$ time, which is the time to generate the complete incidence graph from the $I_x(a)$ hyperplanes. Suppose the output polytope of this stage has $n^v_x(a)$ vertices. In stage 4, for each state $x$, we compute the convex hull of $|A|$ polytopes, which can be done in $O((\sum_{a \in A} n^v_x(a))^{\lfloor |X|/2 \rfloor})$ time. In stage 5, we find the intersection of $|X|$ polytopes. If the polytopes are expressed by half-spaces, this problem is equivalent to removing redundant members of a set of linear inequalities. If the input polytopes have $n^h_x$ facets, $x \in X$, the intersection can be found by solving $\sum_{x \in X} n^h_x$ linear programs. Finally, we would like to mention that the worst-case bound of an algorithm may not be a good measure for its actual performance—the simplex algorithm in linear programming is a famous example.

# References


[1] Abreu, D., D. Pearce, and E. Stacchetti. 1990. Towards a Theory of Discounted Repeated Games with Imperfect Monitoring. *Econometrica* **58**: 1041-1064.

[2] Baron, D., and D. Besanko. 1984. Regulation and Information in a Continuing Relationship. *Information Economics and Policy* **1**: 267-302.

[3] Battaglini, M. 2005. Long-Term Contracting with Markovian Consumers. *American Economic Review* **95**: 637-658.

[4] Berg, M.de, M.van Kreveld, M. Overmars, and O. Schwarzkopf. 2000. *Computational Geometry: Algorithms and Applications*. Berlin: Springer-Verlag.

[5] Boissonnat, J-D., and M. Yvinec. 1998. *Algorithmic Geometry*. Cambridge, UK: Cambridge University Press.

[6] Bolton, P., and M. Dewatripont. 2005. *Contract Theory*. Cambridge, MA: MIT Press.

[7] Cachon, G. 2003. Supply Chain Coordination with Contracts. *Handbooks in Operations Research and Management Science: Supply Chain Management.* S. Graves and T. de Kok (Eds.). North Holland.

[8] Cachon, G., and M. Lariviere. 2001. Contracting to Assure Supply: How to Share Demand Forecasts in a Supply Chain. *Management Science* **47**: 629-646.

[9] Chen, F. 2003. Information Sharing and Supply Chain Coordination. *Supply Chain Management: Design, Coordination and Operations*. A.G. de Kok and S.C. Graves (Eds.). Amsterdam, Netherlands: Elsevier.

[10] Cole, H., and N. Kocherlakota. 2001. Dynamic Games with Hidden Actions and Hidden States. *Journal of Economic Theory* **98**: 114-126.

[11] Corbett, C. 2001. Stochastic Inventory Systems in a Supply Chain with Asymmetric Information: Cycle Stocks, Safety Stocks, and Consignment Stock. *Operations Research* **49**: 487-500.

[12] Ding, Y., R. Jia, and S. Tang. 2003. Dynamic Principal Agent Model Based on CMDP. *Mathematical Methods of Operations Research* **58**: 149-157.

[13] Doepke, M., and R.M. Townsend. 2005. Dynamic Mechanism Design with Hidden Income and Hidden Actions. *Journal of Economic Theory* forthcoming.

[14] Fernandes, A., and C. Phelan. 2000. A Recursive Formulation for Repeated Agency with History Dependence. *Journal of Economic Theory* **91**: 223-247.

[15] FUDENBERG, D., B. HOLMSTROM, AND P. MILGROM. 1990. Short-term contracts and long-term agency relationships. *Journal of Economic Theory* **51**: 1-31.

[16] FUDENBERG, D., AND J. TIROLE. 1991. *Game Theory*. Cambridge, MA: MIT Press.

[17] FUKUDA, K., AND CH. WEIBEL. 2005. Computing All Faces of the Minkowski Sum of $\mathcal{V}$-Polytopes. Working Paper.

[18] FUKUDA, K. 2004. From the Zonetope Construction to the Minkowski Addition of Convex Polytopes. *Journal of Symbolic Computation* **38**: 1261-1272.

[19] HA, A.Y. 2001. Supplier-Buyer Contracting: Asymmetric Cost Information and Cutoff Level Policy for Buyer Participation. *Naval Research Logistics* **48**: 41-64.

[20] HOLMSTROM, B. 1979. Moral Hazard and Observability. *Bell Journal of Economics* **10**: 74-91.

[21] HOWARD, R. 1960. *Dynamic Programming and Markov Processes*. New York: MIT and John Wiley and Sons.

[22] MAS-COLELL, A., M. WHINSTON, AND J. GREEN. 1995. *Microeconomic Theory*. Oxford: Oxford University Press.

[23] MYERSON, R. 1981. Optimal Auction Design. *Mathematics of Operations Research* **6**: 58-73.

[24] PLAMBECK, E., AND S. ZENIOS. 2000. Performance-Based Incentives in a Dynamic Principal-Agent Model. *Manufacturing and Service Operations Management* **2**: 240-263.

[25] PUTERMAN, M.L. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. New York: John Wiley & Sons.

[26] OZER, O., AND W. WEI. 2006. Strategic Commitments for an Optimal Capacity Decision Under Asymmetric Forecast Information. *Management Science* **52**: 1238-1257.

[27] SALANIE, B. 1997. *The Economics of Contracts*. Cambridge, MA: MIT Press.

[28] TOPKIS, D.M. 1998. *Supermodularity and Complementarity*. Princeton, NJ: Princeton University Press.

[29] ZHANG, H. 2004. *A Dynamic Principal-Agent Model with Hidden State.* Unpublished PhD Thesis, Stanford University, Stanford, CA.