

Strategic Bidding of Private Information for Dynamic LQ Networks under Moral Hazard ^{*}

Yasuaki Wasa ^{*} Kenko Uchida ^{*}

^{*} *Waseda University, Tokyo, 169-8555 JAPAN (e-mail:
wasa@aoni.waseda.jp, kuchida@waseda.jp).*

Abstract: This paper investigates the whole system behavior caused by the influence of the agents' strategic behavior while utilizing their individual control and private information for a dynamic linear-quadratic (LQ) network in the presence of a principal-agent relationship. The principal aims at integrating agents' behavior into the socially optimal one based on private information bid by the agents. To avoid a moral hazard on agents' controls, the principal must give a reward to the agents. The reward induces the agents to choose their controls achieving the social objective under the true private information case, but the reward cannot prevent the strategic bidding of the agents' private information. Under this situation, the case is considered that all the agents minimize their net cost composed of their own private cost and the reward from the principal, which is called the strategic bidding problem under moral hazard. Then, the strategic bidding problem is formulated and the optimal design of the problem is analytically derived. Their effectiveness and limitations are also discussed through a simulation.

Keywords: Strategic bidding; principal-agent problems; moral hazard; LQ dynamic game; incentives.

1. INTRODUCTION

Social infrastructure systems have human/machine-in-the-loop controls. Such humans and machines are regarded as agents. The economically rational agent, which is called homo economicus in economics, intends to minimize his own individual cost reflecting his objective. In the social infrastructure systems, each agent has a different goal. To integrate all the agents' strategic behavior rationally and impartially, it is required to introduce an integration mechanism, which is managed by a principal. In economics (see, e.g., Bolton and Dewatripont (2005)), the agents' strategic behavior is divided into two parts: adverse selection (hidden information) and moral hazard (hidden action). Without an appropriate incentive from the principal to the agents, the system behavior led by such strategically selfish behavior of the agents differs greatly from the socially optimal one which is the principal's objective.

The typical example of control problems in the presence of a principal-agent (PA) relationship is in power network systems. In the literature of electricity markets, the strategic bidding problem of private information for competitive agents has been addressed since the late 1990s. Wen and David (2001) propose an optimal bidding strategy for competitive power suppliers in the early days. Steeger et al. (2014) make recent surveys of the strategic bidding problem for hydro-electric producers in day-ahead electricity market. Ghamkhari et al. (2017) point out that one of the unsolved problems in electricity markets is to solve an optimal bidding strategy efficiently. Towards the

realization of smart society projects all over the world, it is expected that the strategic bidding problem will occur in a variety of social infrastructure systems including energy management, traffic, and cyber-physical systems.

In this paper, a PA-type *dynamic* linear-quadratic (LQ) network is considered. The network is constituted by a principal and multiple heterogeneous agents. In the engineering literature of strategic bidding in dynamical systems, Berger and Schweppe (1989) proposed pioneering problem formulation in electricity markets. For the last several years, in order to reform the electricity regulation markets, dynamic incentive mechanisms based on mechanism design are proposed in Tanaka et al. (2012), Taylor et al. (2013) and Murao et al. (2015, 2018). Murao et al. (2015) handle a strategic bidding problem with a dynamic LQ network system, but they do not discuss the performance analysis of the strategic bidding. Wasa et al. (2019b) investigate an optimal design problem for strategic bidding of private parameters without moral hazard in a dynamic LQ network. Many kinds of researches including these five papers focus on only adverse selection without moral hazard. Hence, many academic challenges in strategic bidding problems still remain.

Regarding the moral hazard problem in engineering, there are relatively few papers (Chen and Zhu (2018); Venkatasubramaniam and Gupta (2019); Wasa et al. (2018, 2019a)). The incentives, sometimes called rewards, in moral hazard problems lead the agents' controls to a socially optimal state. The incentive synthesis problem under moral hazard is motivated by the PA problem in contract theory (Holmstrom and Milgrom (1987); Bolton

^{*} This work was partially supported by Japan Science and Technology Agency, CREST Grant Number JPMJCR15K2.

and Dewatripont (2005); Sannikov (2008); Cvitanic et al. (2018)). Chen and Zhu (2018) apply a PA-type moral hazard problem to a differential game between single-principal and single-agent in dynamical cybersecurity management. Wasa et al. (2018) propose an optimal control and incentive synthesis under moral hazard between single-principal and multi-agents in dynamic electricity regulation markets. However, Chen and Zhu (2018) and Wasa et al. (2018, 2019a) do not consider strategic bidding problems of private information under moral hazard. Our objective in this paper is to consider the case integrating both adverse selection and moral hazard. Venkitasubramaniam and Gupta (2019) consider both adverse selection and moral hazard between single-principal and single-agent in static systems, whereas our problem in this paper addresses a strategic bidding problem between a principal and two-agents in dynamical LQ model.

This paper extends the framework proposed in Wasa et al. (2019b) to a strategic bidding model under moral hazard by merging the concept of the strategic bidding problem presented in Wasa et al. (2019b) and the idea for moral hazard problem proposed in Wasa et al. (2018). Specially, private parameter bidding problems of the agents' cost function are handled. The objective is to analytically seek equilibrium for the agents' strategic bidding problem under moral hazard on agents' control. First, an optimal reward is designed to avoid moral hazard action of agents' controls in a PA type dynamic LQ network, following the approach in Wasa et al. (2018). Under the situation, all the agents minimize their own costs without the principal's intervention and an optimal bidding strategy of the agents is characterized. The effectiveness, implementation and limitations of the proposed mechanism are also discussed through a simulation.

2. DYNAMIC LQ NETWORK MODEL IN PRINCIPAL-AGENT RELATIONSHIP

2.1 Network Model

This paper considers an LQ dynamic network model. The network is managed by two types of participants: *Principal* and *Agent*. The principal harmonizes the agents' individual behavior to achieve a socially optimal state by providing each agent with an appropriate pricing information on the basis of their private information bid strategically. Each agent executes suitable control according to the price. Following Wasa et al. (2019b), this paper considers a two-agent model so as to avoid complex technical discussions. The problem can be extended to any number of agents. The state evolution of agent i ($i = 1, 2$) obeys

$$\dot{x}_i = A_i x_i + B_i u_i, \quad x_i(t_0) = x_{0i}, \quad (1)$$

with the individual cost:

$$J_i = \int_{t_0}^{t_f} [(\theta_i - C_i x_i)^\top Q_i (\theta_i - C_i x_i) + u_i^\top R_i u_i] dt, \quad (2)$$

where Q_i and R_i are positive-definite symmetric matrices. Agent i can observe his own state $x_i(t)$ online and has private information $\theta_i(t)$, $t \in [t_0, t_f]$ *a priori*. The private information $\theta_i(t)$ is continuous functions of time t . The system parameters and cost parameters $A_i(t), B_i(t), C_i(t), Q_i(t), R_i(t)$, $t \in [t_0, t_f]$, $i = 1, 2$ in

(1) and (2) are continuous functions of t and common information to all the participants in advance.

The state evolution of the principal, which indicates various imbalance outcomes on the network, obeys

$$\dot{x}_0 = A_0 x_0 + A_{01} x_1 + A_{02} x_2, \quad x_0(t_0) = x_{00}, \quad (3)$$

with the imbalance cost

$$J_0 = \int_{t_0}^{t_f} x_0^\top Q_0 x_0 dt, \quad (4)$$

where Q_0 is a semi-positive-definite symmetric matrix. The principal can observe all the states $x_i(t)$, $i = 0, 1, 2$, online. All the system and cost parameters in (3) and (4) are common information to all the participants as well.

2.2 Social Optimal Control Problem

To evaluate the efficiency over the network, the social cost is defined by

$$I_s = J_0 + J_1 + J_2, \quad (5)$$

where the collective dynamics for the state $x = [x_0^\top, x_1^\top, x_2^\top]^\top$ with the controls $u = [u_1^\top, u_2^\top]^\top$ is described as

$$\dot{x} = Ax + Bu, \quad x(t_0) = [x_{00}^\top, x_{01}^\top, x_{02}^\top]^\top, \quad (6)$$

$$A = \begin{bmatrix} A_0 & A_{01} & A_{02} \\ 0 & A_1 & 0 \\ 0 & 0 & A_2 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ B_1 & 0 \\ 0 & B_2 \end{bmatrix}.$$

Using (6), the social cost (5) is rewritten as

$$I_s = \int_{t_0}^{t_f} (x^\top C^\top Q C x - 2\theta^\top Q C x + \theta^\top Q \theta + u^\top R u) dt,$$

where $\theta = [0^\top, \theta_1^\top, \theta_2^\top]^\top$,

$$C = \begin{bmatrix} I & 0 & 0 \\ 0 & C_1 & 0 \\ 0 & 0 & C_2 \end{bmatrix}, \quad Q = \begin{bmatrix} Q_0 & 0 & 0 \\ 0 & Q_1 & 0 \\ 0 & 0 & Q_2 \end{bmatrix}, \quad R = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}.$$

From the definitions, the matrices Q and R are symmetric positive-definite matrices and I is the identity matrix. The social optimal control problem solved by the principal is to find an optimal control minimizing (5). This social optimal control problem is one of the standard LQ control problems. Each agent bids his private information $\theta_i(t)$, $t \in [t_0, t_f]$. Then, the agents' optimal controls are given by

$$u_i^* = -\frac{1}{2} R_i^{-1} B_i^\top p_i, \quad i = 1, 2, \quad (7)$$

where p_i is the adjoint variable (afterward it is called (shadow) price), and the price vector $p = [p_0^\top, p_1^\top, p_2^\top]^\top$ obeys the adjoint equation

$$\dot{p} = -A^\top p - 2C^\top Q C x + 2C^\top Q \theta, \quad p(t_f) = 0. \quad (8)$$

The optimal pair (x, p) of state and price can be determined uniquely by solving the two point boundary value (TPBV) problem given by the forward equation (6) and the backward equation (8) (see Bryson and Ho (1975) and Ascher et al. (1995)). Ideally, the optimal control $u^* = [u_1^{*\top}, u_2^{*\top}]^\top$ is implemented as follows. If agent i bids his private information θ_i , the principal solves the TPBV problem offline and determines the price p_i , which is optimal from the point of view of the social cost (5). The principal dispatches the price p_i to each agent i and the agent i performs his control (7) in a distributed fashion.

From Bryson and Ho (1975) and Ascher et al. (1995), the price as well as the control can be realized in a state

feedback form with the backward sweep formula for the TPBV problem:

$$p_i = 2P_i x + 2g_i, \quad i = 1, 2, \quad (9)$$

where $P = [P_0^\top \ P_1^\top \ P_2^\top]^\top$, $P_i = [P_{i0} \ P_{i1} \ P_{i2}]$, $i = 0, 1, 2$, is the solution to the Riccati equation:

$$\dot{P} + PA + A^\top P - PBR^{-1}B^\top P + C^\top QC = 0, \quad P(t_f) = 0, \quad (10)$$

and $g = [g_0^\top, g_1^\top, g_2^\top]^\top$ obeys the linear equation:

$$\dot{g} = -(A^\top - PBR^{-1}B^\top)g + C^\top Q\theta, \quad g(t_f) = 0. \quad (11)$$

Here the implementation is discussed to realize the state feedback laws for the price (9) and the control (7). All the model information (all participants' system and weighting variables) except the private information θ_i , $i = 1, 2$ is available to all the participants in common. Any disturbance is not considered in the model. Hence, the principal can calculate the price and the controls in the open-loop form by solving the TPBV problem defined by (6), (7) and (9)–(11) and in the state feedback form as well, as the principal gets the private information bid by all the agents and the state information of all participants' dynamics.

3. OPTIMAL REWARD DESIGN PROBLEM UNDER MORAL HAZARD

3.1 Problem Formulation

As pointed as Wasa et al. (2019b), the agents have a potential to behave strategic actions and the principal cannot achieve the social cost minimization without an appropriate incentive. To incentivize an agent's behavior, the principal uses a reward functional $W_i(x; u, \theta_i, h_i)$, $i = 1, 2$ with a parameter h_i .¹ By combining the reward functional W_i with the individual cost functionals J_0 , J_1 and J_2 , the agent's net cost functional I_i and the principal's net cost functional I are defined as

$$I_i(x; u, \theta_i, h_i) = J_i(x; u) + W_i(x; u, \theta_i, h_i), \quad i = 1, 2, \quad (12)$$

$$I(x; u, \theta, h) = J_0 - W_1(x; u, \theta_1, h_1) - W_2(x; u, \theta_2, h_2), \quad (13)$$

where $h := [h_1^\top, h_2^\top]^\top$. Then, our reward design problem under moral hazard is formulated as follows: Given an initial condition $x(t_0) = x$ and private information $\theta = [0^\top, \theta_1^\top, \theta_2^\top]^\top$, the principal finds an optimal parameter $h = [h_1^\top, h_2^\top]^\top$ that implements an optimal control $u = [u_1^\top, u_2^\top]^\top$ in an admissible set, such that

$$\min_{u=[u_1^\top, u_2^\top]^\top, h=[h_1^\top, h_2^\top]^\top} I(x; u, \theta, h) \quad (14)$$

$$\text{subject to } I_i(x; u_i, u_{-i}, \theta_i, h_i) = \min_{v_i} I_i(x; v_i, u_{-i}, \theta_i, h_i), \quad i = 1, 2, \quad (15)$$

$$I_i(x; u_i, u_{-i}, \theta_i, h_i) \leq 0, \quad i = 1, 2, \quad (16)$$

where the subscript notation $\cdot_{-i} = \cdot_2$ if $i = 1$, and vice versa.² The optimal reward functional W_i and the agents' optimal controls are obtained from the solution to

¹ This approach is motivated by the reward-functional formulation in Sannikov (2008) and Cvitanic et al. (2018), which are investigated in the context of contract theory. Wasa et al. (2018) reveal that the reward design for the initial and terminal cost can be independent of that for the transient cost. Therefore, this paper focuses on the reward functional for only the transient cost, which is given by (18).

² Since the control parameter u is composed of (u_i, u_{-i}) , $i = 1, 2$, we sometimes use $I_i(x; u_i, u_{-i}, \theta_i, h_i)$ in place of $I_i(x; u, \theta_i, h_i)$.

this problem. Assuming the private information (θ_1, θ_2) is reported in all sincerity, constraint (15) claims that the optimal reward functional incentivizes each agent to adopt the optimal control that minimizes its own net cost. This constitutes a Nash equilibrium together with the other agents' control. Constraint (16) ensures a prescribed level of each agent's net cost, which is set as 0 in this paper. The above formulation (14)–(16) is an application of the moral-hazard problem, particularly the PA problem, in contract theory (e.g. see Holmstrom and Milgrom (1987)); using the terminology of contract theory, (15) is called the incentive-compatibility constraint and (16) is the individual-rationality constraint.

3.2 Solution

In this section, the case is considered that all the agents report truthful private parameters (θ_1, θ_2) to the principal. Following Sannikov (2008) and Wasa et al. (2018), an appropriate reward functional and agents' control policy satisfying the constraints (15) and (16) can be obtained straightforwardly. This paper focuses only on the primary results of the problem (14)–(16) in Wasa et al. (2018). See Wasa et al. (2018) for the mathematical assumptions and the rigorous proofs.

First, to solve the reward design problem (14)–(16), a form of the reward functional W_i with h_i is specified. Suppose that there is a Nash equilibrium (u_1^*, u_2^*) defined by

$$u_i^* = \arg \min_{v_i} I_i(x; v_i, u_{-i}^*, \theta_i, h_i), \quad i = 1, 2. \quad (17)$$

for a pair of reward functionals (W_1, W_2) . Then, consequently, given an initial condition $x(t_0) = x$ and optimal controls u_{-i}^* of the other agent, the reward functional W_i of the agent i ($i = 1, 2$) with arbitrary implementable control u_i has the form:

$$\begin{aligned} W_i(x; u_i, u_{-i}^*, \theta_i, h_i) \\ = - \int_{t_0}^{t_f} \left[h_i^\top(t, x(t)) (Ax(t) + B_i u_i^* + B_{-i} u_{-i}^* - \dot{x}(t)) \right. \\ \left. + (\theta_i - C_i x_i(t))^\top Q_i (\theta_i - C_i x_i(t)) + u_i^{*\top} R_i u_i^* \right] dt \end{aligned} \quad (18)$$

along the dynamics $\dot{x}(t) = Ax(t) + B_i u_i + B_{-i} u_{-i}^*$, $t \in [t_0, t_f]$. By substituting (18) into (12) and (15), the optimality of the obtained reward functional (18) can be verified. Then, by taking the price parameter p_i (9) depending on time t and state x and the parameter

$$h_i = [h_{i0}^\top, h_{i1}^\top, h_{i2}^\top]^\top, \quad h_{ij} = \begin{cases} p_i & \text{if } j = i \\ 0 & \text{otherwise} \end{cases}, \quad i = 1, 2, \quad (19)$$

which is provided by the principal,³ the corresponding optimal control for u_{-i}^* is obtained as

$$\begin{aligned} u_i^* = \arg \min_{u_i} [(\theta_i - C_i x_i)^\top Q_i (\theta_i - C_i x_i) \\ + u_i^\top R_i u_i + p_i^\top (A_i x_i + B_i u_i)] \\ = - (1/2) R_i^{-1} B_i^\top p_i^\top. \end{aligned} \quad (20)$$

Note that each agent's control policy (20) with (9) is a function of state x , i.e. a Markov control policy. Thanks

³ The parameter h_i in (18) is the principal's design parameter fulfilling the constraints (15) and (16). From Sannikov (2008) and Wasa et al. (2018), the principal generally has all manner of options of h_i including (19) and Proposition 1 holds for all the candidates of h_i . Meanwhile, solving the specific optimization problem (14) defined by (13), the principal consequently selects (19).

to the control structure and necessary conditions for Markov Nash equilibrium (u_i^*, u_{-i}^*) of the controls (20), the form of the reward functional (18) is acquired. Hence, by using (18), Proposition 1 is obtained regarding the relationship between the parameters (h_1, h_2) with (19) and the constraints (15) and (16).

Proposition 1. (a) For the reward functionals (18) with parameters (h_1, h_2) based on (19), a pair of controls (u_1, u_2) is a Nash equilibrium if and only if the control has the form defined by (20).

(b) For the reward functionals (18) with reward parameters (h_1, h_2) and the Nash equilibrium (u_1^*, u_2^*) obtained by (20), the constraint (16) is fulfilled.

Proof. See Wasa et al. (2018).

From this result, as long as all the agents report truthful private parameters (θ_1, θ_2) to the principal, the reward functional given by (18) guarantees the incentive-compatibility constraint (15) and the individual-rationality constraint (16) on the agent's input u_i . Moreover, the Nash equilibrium (20) and the reward functional (18) with the parameter (h_1, h_2) defined by (19) and (9) achieve social cost minimization (Wasa et al. (2018)).

4. STRATEGIC BIDDING PROBLEM UNDER MORAL HAZARD

4.1 Problem Formulation

When each agent bids his own private information $\theta_i(t)$, $t \in [t_0, t_f]$, the principal can determine an optimal reward functional led by (18) with (19) for the bidden private information (θ_1, θ_2) and the social optimal control preventing agents' moral hazard control behavior for (θ_1, θ_2) is guaranteed as above. Meanwhile, each agent can bid his private information *strategically*. Let us denote by $\bar{\theta}_i(t)$, $t \in [t_0, t_f]$ the true private information of agent i , ($i = 1, 2$). As long as each agent implements (20), the prescribed level of agent's net cost and the social optimal control for (θ_1, θ_2) are guaranteed. The input constituted by Nash equilibrium (20) is independent of the strategic bidding information θ_i . Afterward, it is assumed that each agent i always implements his control led by (20) and the strategic bidding problem of his information θ_i is handled.

The main objective of the strategic bidding problem is to find a Nash equilibrium in a parameter space of the (θ_1, θ_2) . To describe the strategic bidding problem clearly, the state x , the price p , and the parameter g in the backward sweep are determined dependently on the bid parameter (θ_1, θ_2) from the foregoing discussion. From (11), g is determined by

$$\dot{g} = -(A - BR^{-1}B^T P)^T g + \begin{bmatrix} 0 & 0 \\ C_1^T Q_1 & 0 \\ 0 & C_2^T Q_2 \end{bmatrix} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix},$$

$$g(t_f) = 0. \quad (21)$$

For the optimal control (7) with (9) in the state feedback form, the state equation (6) is also described by

$$\dot{x} = (A - BR^{-1}B^T P)x - BR^{-1}Bg,$$

$$x(t_0) = [x_{00}^T, x_{01}^T, x_{02}^T]^T. \quad (22)$$

By combining (2), (9), (12), (18), (19) and (20), the agent's net cost I_i , $i = 1, 2$ is given by

$$I_i^*(\theta_1, \theta_2) = \int_{t_0}^{t_f} \left[(\bar{\theta}_i - C_i x_i)^T Q_i (\bar{\theta}_i - C_i x_i) - (\theta_i - C_i x_i)^T Q_i (\theta_i - C_i x_i) \right] dt$$

$$= \int_{t_0}^{t_f} \left[\bar{\theta}_i^T Q_i \bar{\theta}_i - 2(C_i x_i)^T Q_i (\bar{\theta}_i - \theta_i) - \theta_i^T Q_i \theta_i \right] dt, \quad (23)$$

where x in (23) and g depend on (θ_1, θ_2) and obey (22) and (21), respectively. As Q_i is defined as a positive-definite matrix, the private information θ_i minimizing the cost (23), particularly $-\theta_i^T Q_i \theta_i$, takes a value on the boundary in parameter space. As long as all the agents implement Nash equilibrium (20), all the agents' estimated (enormous) costs with bid information (θ_1, θ_2) are shared with the principal and the agents will make a large profit.

Then, although the principal's objective is to minimize the social cost for evaluating the efficiency of the network, the principal will pay enormous money and the network is unstabilized at its worst. From the point of view of the budget, as the incentive cost W_i led by (18) is born by the principal, it is realistic not to exceed a prescribed level of the principal's budget k . Hence, a revised strategic bidding problem adding the inequality constraint on the principal's net cost

$$I^*(\theta_1, \theta_2) \leq k \quad (24)$$

is reformulated. The constraint (24) is an integral constraint. Therefore, given a Lagrange multiplier $\rho \geq 0$, which is independent of time t , the constraint (24) is equivalent to

$$\rho(I^*(\theta_1, \theta_2) - k) = 0, \quad \rho \geq 0 \quad (25)$$

(e.g., see Liberzon (2011)). From (9), (13) and (18)–(20), the cost functional I^* is given by

$$I^*(\theta_1, \theta_2) = \int_{t_0}^{t_f} \left[x_0^T Q_0 x_0 + \sum_{i=1,2} (\theta_i - C_i x_i)^T Q_i (\theta_i - C_i x_i) + (Px + g)^T BR^{-1}B^T (Px + g) \right] dt. \quad (26)$$

The feasible space of the bidding parameter θ_i of the agent i ($i = 1, 2$) is limited to the space satisfying the budget constraint (24). Hence, it is assumed that each agent i minimizes his own net cost I_i^* subject to the principal's budget constraint (24). Considering the necessary condition of the constrained nonlinear optimization problems, the agent i modifies his objective function to the sum of the original net cost and the penalty functional on I^* led by (25), i.e.,

$$I_i^\dagger(\theta_1, \theta_2) := I_i^*(\theta_1, \theta_2) + \rho_i(I^*(\theta_1, \theta_2) - k), \quad \rho_i > 1, \quad (27)$$

along the dynamics (21) and (22).⁴ The parameter ρ_i in (27) can be regarded as a penalty parameter. The objective function (27) can be interpreted as balancing the agents' individual rationality and the principal's budget for the suitable incentive by using ρ_i and k . Substitute (23) and (26) into (27), then

$$I_i^\dagger(\theta_1, \theta_2) = \int_{t_0}^{t_f} L_i(x, g, \bar{\theta}_i, \theta, \rho_i, k, t) dt, \quad (28)$$

⁴ The constraint $\rho_i > 1$ instead of $\rho_i \geq 0$ is led by the necessary condition in Theorem 2. See Section 4.2 for more details.

where

$$\begin{aligned}
& L_i(x, g, \bar{\theta}_i, \theta, \rho, k, t) \\
&= (\bar{\theta}_i - C_i x_i)^\top Q_i (\bar{\theta}_i - C_i x_i) + \rho_i x_0^\top Q_0 x_0 \\
&\quad + (\rho_i - 1)(\theta_i - C_i x_i)^\top Q_i (\theta_i - C_i x_i) \\
&\quad + \rho_i (\theta_{-i} - C_{-i} x_{-i})^\top Q_{-i} (\theta_{-i} - C_{-i} x_{-i}) \\
&\quad + \rho_i (Px + g)^\top BR^{-1} B^\top (Px + g) - \frac{\rho_i k}{t_f - t_0} \\
&= \bar{\theta}_i^\top Q_i \bar{\theta}_i - 2\bar{\theta}_i^\top Q_i C_i x_i + x_i^\top C_i^\top Q_i C_i x_i \\
&\quad + \theta^\top Q^{\rho_i} \theta - 2\theta^\top Q^{\rho_i} Cx + x^\top C^\top Q^{\rho_i} Cx \\
&\quad + \rho_i (Px + g)^\top BR^{-1} B^\top (Px + g) - \frac{\rho_i k}{t_f - t_0}, \quad (29)
\end{aligned}$$

$$Q^{\rho_1} := \rho_1 Q - \begin{bmatrix} 0 & 0 & 0 \\ 0 & Q_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad Q^{\rho_2} := \rho_2 Q - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & Q_2 \end{bmatrix}.$$

Denote by Θ_i the set of the admissible bid information θ_i of the agent i , $i = 1, 2$. As each agent has his own cost for evaluating the strategic bidding of θ_i , it is ideal that agent i expects to find his bid information θ_i minimizing his individual cost (28) subject to dynamic equations (21) and (22) for an information θ_{-i} bid by agent $-i$. Then, the strategic bidding problem under moral hazard is defined as follows: given a principal's maximum budget k and the corresponding penalty parameters $\rho_1 > 1$ and $\rho_2 > 1$, the agents form a Nash equilibrium $(\theta_1^\dagger, \theta_2^\dagger)$ in $\Theta_1 \times \Theta_2$ satisfying the criterion

$$I_1^\dagger(\theta_1^\dagger, \theta_2^\dagger) \leq I_1^\dagger(\theta_1^\dagger, \theta_2), \quad \forall \theta_2 \in \Theta_2, \quad (30a)$$

$$I_2^\dagger(\theta_1^\dagger, \theta_2^\dagger) \leq I_2^\dagger(\theta_1, \theta_2^\dagger), \quad \forall \theta_1 \in \Theta_1, \quad (30b)$$

under the condition that the backward equation (21) and the forward equation (22) are implemented.

Here, the feasible space of the parameters (ρ_1, ρ_2) is limited in order to be consistent with the purpose of the optimal reward design problem under moral hazard for control (u_1, u_2) in Section 3. To be concrete, the agent's net cost $I_i^*(\theta_1^\dagger, \theta_2^\dagger)$ with the Nash equilibrium $(\theta_1^\dagger, \theta_2^\dagger)$ constituted by (30) must satisfy the individual-rationality constraint (16). As a result, the principal determines the maximum budget k minimizing the principal's net cost (social cost) $I^*(\theta_1^\dagger, \theta_2^\dagger)$ in the range of the parameters (ρ_1, ρ_2) satisfying $I_i^*(\theta_1^\dagger, \theta_2^\dagger) \leq 0$, $i = 1, 2$. As the resulting cost $I^*(\theta_1^\dagger, \theta_2^\dagger)$ is equivalent to k , the constraint (25) holds.

4.2 Solution

Let us seek $(\theta_1^\dagger, \theta_2^\dagger)$ constituted by (30). The strategic bidding problem formulated in Section 4.1 is an open-loop LQ Nash game with the strategies θ_i , $i = 1, 2$, and has the dynamic constraints (21) and (22) defined with a TPBV condition. As Q_i is a symmetric positive-definite matrix and $\rho_i > 1$, the cost I_i^\dagger led by (28) and (29) is strictly convex with respect to θ_i . Then, Theorem 2 is obtained.

Theorem 2. Suppose that, for $\rho_i > 1$, $i = 1, 2$, there is an (open-loop Nash) equilibrium $(\theta_1^\dagger, \theta_2^\dagger)$ led by (30) and let us set $k := I^*(\theta_1^\dagger, \theta_2^\dagger)$. Then, given parameters $\rho_i > 1$, (I) the strategic bidding problem defined by (21)–(23) and (26)–(30) has a unique equilibrium $(\theta_1^\dagger, \theta_2^\dagger)$ in $\Theta_1 \times \Theta_2$; (II) the unique equilibrium $(\theta_1^\dagger, \theta_2^\dagger)$ is given by

$$\theta_i^\dagger = C_i x_i - \frac{1}{\rho_i - 1} C_i \mu_{ii}, \quad i = 1, 2, \quad (31)$$

where μ_{ii} is an element of $\mu^i = [\mu_{i0}^\top, \mu_{i1}^\top, \mu_{i2}^\top]^\top$, $i = 1, 2$, and μ^i is determined as a unique solution of the TPBV (linear) equations led by $\theta_i = \theta_i^\dagger$, $i = 1, 2$:

$$\dot{x} = (A - A_B P)x - A_B g, \quad x(t_0) = [x_{00}^\top, x_{01}^\top, x_{02}^\top]^\top, \quad (32a)$$

$$\begin{aligned} \dot{g} &= -(A - A_B P)^\top g + A_C x \\ &\quad - \frac{1}{\rho_1 - 1} A_C^1 \mu^1 - \frac{1}{\rho_2 - 1} A_C^2 \mu^2, \quad g(t_f) = 0, \end{aligned} \quad (32b)$$

$$\begin{aligned} \dot{\lambda}^i &= -(A - A_B P)^\top \lambda^i - A_C \mu^i - 2(\rho_i P^\top A_B)g \\ &\quad - 2(\rho_i (A_C^0 + P^\top A_B P) + A_C^i)x + 2c_{\bar{\theta}}^i, \\ \lambda^i(t_f) &= 0, \quad i = 1, 2, \end{aligned} \quad (32c)$$

$$\begin{aligned} \dot{\mu}^i &= (A - A_B P)\mu^i + A_B \lambda^i - 2(\rho_i A_B)g - 2(\rho_i A_B P)x, \\ \mu^i(t_0) &= 0, \quad i = 1, 2, \end{aligned} \quad (32d)$$

where $A_B = BR^{-1}B^\top$, $A_C = A_C^1 + A_C^2$,

$$\begin{aligned}
A_C^1 &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & C_1^\top Q_1 C_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad A_C^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & C_2^\top Q_2 C_2 \end{bmatrix}, \\
A_C^0 &= \begin{bmatrix} Q_0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad c_{\bar{\theta}}^1 = \begin{bmatrix} 0 \\ C_1^\top Q_1 \bar{\theta}_1 \\ 0 \end{bmatrix}, \quad c_{\bar{\theta}}^2 = \begin{bmatrix} 0 \\ 0 \\ C_2^\top Q_2 \bar{\theta}_2 \end{bmatrix}.
\end{aligned}$$

Proof. Since the open-loop Nash equilibrium is considered and the parameters (ρ_1, ρ_2) are fixed, the necessary condition for $(\theta_1^\dagger, \theta_2^\dagger)$ can be derived by using the maximum principle (Bryson and Ho (1975); Liberzon (2011)). From Basar and Olsder (1999), the necessary condition is given as the TPBV problem composed of (21) and (22) and their adjoint equations (32c) and (32d) with the equilibrium (31). Thus, by using the existence and uniqueness theorem of the TBPV problem for linear differential equations (Ascher et al., 1995, Theorem 3.26b) with $\rho_i > 1$, the TBPV problem (32) has a unique solution and the equilibrium $(\theta_1^\dagger, \theta_2^\dagger)$ is given by (31) and (32). This completes the proof.

Theorem 2 indicates that the necessary (and sufficient) condition of the Nash equilibrium (30) led by the minimization problem with (27) is given as the unique solution of the TPBV problem (32) under $I^*(\theta_1^\dagger, \theta_2^\dagger) = k$. However, Theorem 2 does not theoretically find the principal's maximum budget k and the corresponding parameters ρ_1 and ρ_2 satisfying the budget constraint $I^*(\theta_1^\dagger, \theta_2^\dagger) = k$ and the individual-rationality constraint $I_i^*(\theta_1^\dagger, \theta_2^\dagger) \leq 0$. This theoretical issue including convergence properties remains as one of the future works and it will be discussed in a separate paper. In Section 5, a heuristic computational method will be proposed to seek suitable parameters k and ρ_i , $i = 1, 2$.

5. SIMULATION

In this section, the quantitative influence of the proposed strategic bidding under moral hazard is analyzed through a simulation. Each agent i ($i = 1, 2$) has the local dynamics obeying $\dot{x}_i = u_i$, $x_i(0) = x_{0i} = 1$ with the individual strategic cost $\int_0^{15} [10(\theta_i - x_i)^2 + u_i^2] dt$ and his actual cost $J_i = \int_0^{15} [10(\bar{\theta}_i - x_i)^2 + u_i^2] dt$, that is $t_0 = 0$, $t_f = 15$,

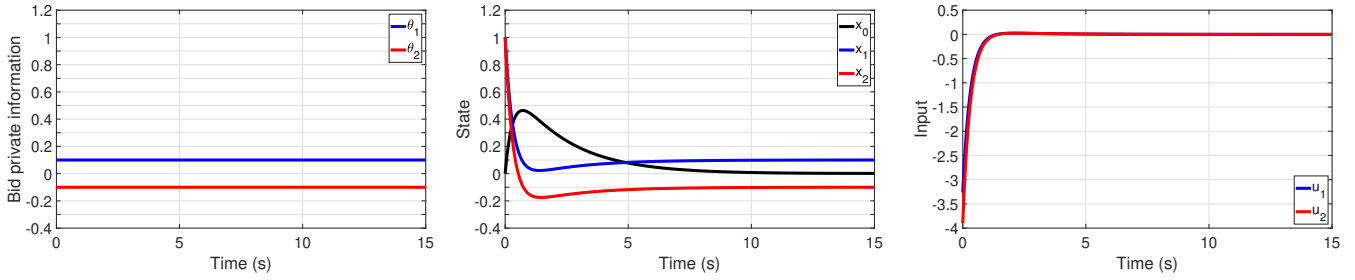


Fig. 1. Time evolution of θ , x and u for social optimal control with true private information ($\theta_1 = \bar{\theta}_1 = 0.1$, $\theta_2 = \bar{\theta}_2 = -0.1$). The resulting costs are $I^* = 7.3580$, $I_1^* = I_2^* = 0$, $W_1 = -2.8370$ and $W_2 = -4.1246$.

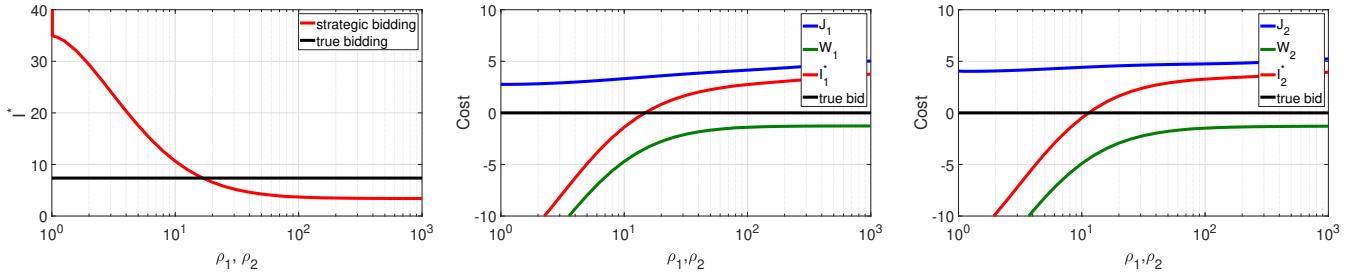


Fig. 2. Costs (I^* , J_i and I_i^*) and incentive (W_i) for strategic bidding of private information at each $\rho_1 = \rho_2 > 1$.

$A_i = 0$, $B_i = 1$, $C_i = 1$, $Q_i = 10$, $R_i = 1$ and the agent's true private information $\bar{\theta}_1 = 0.1$, $\bar{\theta}_2 = -0.1$. The principal's dynamics obey $\dot{x}_0 = x_1 + x_2$, $x_0(0) = x_{00} = 0$ with the cost $J_0 = \int_0^{15} x_0^2 dt$, which is $A_{00} = 0$, $A_{01} = 1$, $A_{02} = 1$ and $Q_0 = 1$. Although there might be a variety of options as the performance analysis of the strategic bidding, the case of $\rho_1 = \rho_2 > 1$ is considered.

Before discussing the strategic bidding problem in Section 4, the case of the social optimal control is verified. The case is also the solution of the optimal reward design problem under moral hazard with true private information, based on Proposition 1 in Section 3. Then, the time evolution of θ , x and u is shown in Fig. 1 and the resulting costs are $I^* = 7.3580$ and $I_1^* = I_2^* = 0$. From Fig. 1, Proposition 1 holds.

To verify the performance of the strategic bidding based on Theorem 2, a parameter ρ_i in the range of $1 < \rho_i \leq 1000$ is selected and the TPBV problem (32) for the fixed parameter ρ_i is solved. Fig. 2 shows the costs $I^*(\theta_1^\dagger, \theta_2^\dagger)$ and $I_i^*(\theta_1^\dagger, \theta_2^\dagger)$ and the corresponding information for $(\theta_1^\dagger, \theta_2^\dagger)$ given by the solution of the TPBV problem (32) with (31) in the range of $1 < \rho_i \leq 1000$. We see from Fig. 2 that I^* is a monotonically-decreasing function for ρ_i and I_i^* is a monotonically-increasing function. To satisfy the individual-rationality constraint $I_i^* \leq 0$, $i = 1, 2$, let us denote by ρ_i^* the parameter ρ_i such that $I_i^*(\theta_1^\dagger, \theta_2^\dagger) = 0$. From the center and right figures in Fig. 2, $\rho_1^* = 14.8$ and $\rho_2^* = 11.3$. Hence, as the principal finds $k = I^*(\theta_1^\dagger, \theta_2^\dagger)$ minimizing I^* from $1 < \rho_1 = \rho_2 \leq \rho^* := \min\{\rho_1^*, \rho_2^*\}$, we obtain $k^* = I^*(\theta_1^\dagger, \theta_2^\dagger) = 9.7146$ at $\rho^* = 11.3$.

Figs. 3 and 4 show θ_i , x and u given by the solution of the TPBV (32) with (31) for $\rho_i = \rho^*$ and $\rho_i = 1000$, respectively. We see from Fig. 3 that, compared with Fig. 1, x_i , $i = 1, 2$ goes to θ_i but x_0 converges to 0.2 not 0. This is caused by the influence of the agents' strategic

bidding taking the individual rationality constraint into account. Hence, the results in Fig. 3 give the limitations of the strategic bidding case. Fig. 4 indicates that the principal sets a relatively small k taking a sufficient large ρ and prioritizes the minimization of J_0 . As a result, the agents receive a small reward W_i and pays a large cost due to the gap between x_i and θ_i . However, the individual rationality does not hold and the agents will not agree with the proposed mechanism from the principal.

6. CONCLUSION

This paper has investigated the whole system behavior caused by the influence of the agents' strategic behavior while utilizing their individual control and private information. In particular, we have formulated a strategic bidding problem under moral hazard for a dynamic LQ network in the presence of a principal-agent relationship. Then, we have analytically derived an optimal design of the strategic bidding problem and discuss their effectiveness and limitations through a simulation. The resulting controls and strategic bidding give the construction of a (near-)optimal agreement between the principal and the agents while satisfying the individual-rationality constraint.

The problem formulation considered in this paper can be regarded as a special case of strategic bidding problems combining moral hazard and adverse selection. To achieve such a strategic bidding problem, Nash equilibrium on both (θ_1, θ_2) and (u_1, u_2) under (W_1, W_2) should be sought. Another issue of the proposed approach is implementation in real systems. These are our next challenge. As the standard problem under moral hazard uses a stochastic system, another future direction is to extend our problem formulation to a stochastic model, e.g. LQG systems. Due to risk-neutral agents and principal, it is expected to directly solve the problem by using the mathematical technique in Wasa et al. (2018) and this paper.

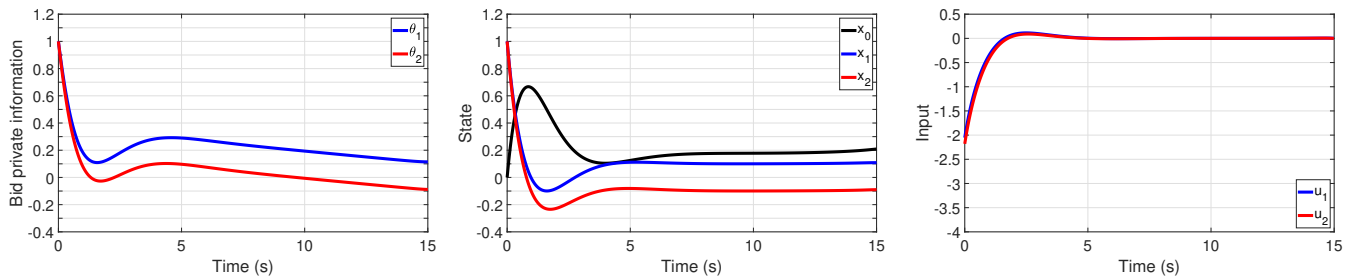


Fig. 3. Time evolution of θ , x and u for strategic bidding of private information with $\rho_1 = \rho_2 = \rho^* = 11.4$. The resulting costs are $I^* = 9.7146$, $I_1^* = -0.9046$, $I_2^* = -0.0180$, $W_1 = -4.2710$ and $W_2 = -4.4637$.

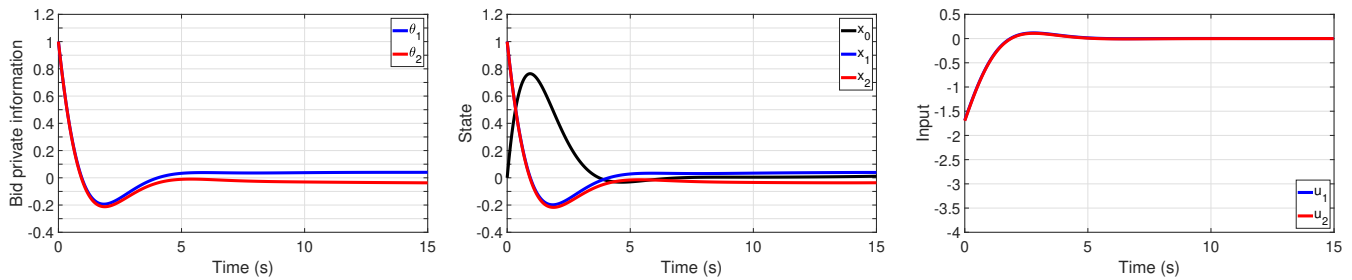


Fig. 4. Time evolution of θ , x and u for true private information with $\rho_1 = \rho_2 = 1000$. The resulting costs are $I^* = 3.3973$, $I_1^* = 3.7660$, $I_2^* = 3.9416$, $W_1 = -1.2703$ and $W_2 = -1.2925$.

REFERENCES

- Ascher, U.M., Mattheij, R.M.M., and Russell, R.D. (1995) *Numerical Solution of Boundary Value Problems for Ordinary Differential Equations*. SIAM.
- Berger, A.W., and Schweppe, F.C. (1989) Real time pricing to assist in load frequency control. *IEEE Trans. Power Systems*, 4(3):920–926.
- Basar, T., and Olsder, G.J. (1999) *Dynamic Noncooperative Game Theory*. 2nd Ed., SIAM.
- Bolton, P., and Dewatripont, M. (2005) *Contract Theory*. The MIT Press.
- Bryson Jr., A.E., and Ho, Y.-C. (1975) *Applied Optimal Control* (Revised Printing). Taylor & Francis.
- Chen, J., and Zhu, Q. (2018) A differential game approach to dynamic contract design for systemic cyber risk management under asymmetric information. *Proc. 56th Annual Allerton Conference on Communication, Control, and Computing*, 575–582.
- Cvitanić, J., Possamaï, D., and Touzi, N. (2018) Dynamic programming approach to principal-agent problems. *Finance and Stochastics*, 22(1):1–37.
- Ghamkhari, M., Sadeghi-Mobarakeh, A., and Mohsenian-Rad, H. (2017) Strategic bidding for producers in nodal electricity markets: A convex relaxation approach. *IEEE Trans. Power Syst.*, 32(3):2324–2336.
- Holmstrom, B., and Milgrom, P. (1987) Aggregation and linearity in the provision of intertemporal incentives. *Econometrica*, 55(2):303–328.
- Liberzon, D. (2011) *Calculus of Variations and Optimal Control Theory: A Concise Introduction*, Princeton University Press.
- Murao, T., Hirata, K., and Uchida, K. (2015) Optimal bidding and worst case pricing under dynamic integration mechanism for LQG power networks. *Proc. 2015 European Control Conference*, 3315–3320.
- Murao, T., Hirata, K., Okajima, Y., and Uchida, K. (2018) Real-time pricing for LQG power networks with independent types: A dynamic mechanism design approach. *European Journal of Control*, 39:95–105.
- Sannikov, Y. (2008) A continuous-time version of the principal-agent problem. *Review of Economic Studies*, 75(3):957–984.
- Steeger, G., Barroso, L.A., and Rebennack, S. (2014) Optimal bidding strategies for hydro-electric producers: A literature survey. *IEEE Trans. Power Syst.*, 29(4):1758–1766.
- Tanaka, T., Cheng, A.Z.W., and Langbort, C. (2012) A dynamic pivot mechanism with application to real time pricing in power systems. *Proc. 2012 American Control Conference*, 3705–3711.
- Taylor, J.A., Nayyar, A., Callaway, D.S., and Poolla, K. (2013) Consolidated dynamic pricing of power system regulation. *IEEE Trans. Power Syst.*, 28(4):4692–4700.
- Venkitasubramaniam, P., and Gupta, V. (2019) Data-driven contract design. *Proc. 2019 American Control Conference*, 2283–2288.
- Wasa, Y., Hirata, K., and Uchida, K. (2018) Contract theory approach to incentivizing market and control design. arXiv:1709.09318.
- Wasa, Y., Hirata, K., and Uchida, K. (2019a) Optimal agency contract for incentive and control under moral hazard in dynamic electric power networks. *IET Smart Grid*, 2(4):594–601.
- Wasa, Y., Murao, T., Tanaka, T., and Uchida, K. (2019b) Strategic bidding of private information for principal-agent type dynamic LQ networks. *Proc. 17th European Control Conference*, 3383–3389.
- Wen, F., and David, A.K. (2001) Optimal bidding strategies and modeling of imperfect information among competitive generators. *IEEE Trans. Power Syst.*, 16(1):15–21.