Decentralized Control via Dynamic Stochastic Prices: The Independent System Operator Problem

Rahul Singh, Student Member, IEEE, P. R. Kumar, Fellow, IEEE, and Le Xie, Senior Member, IEEE

Abstract—A smart grid connects several electricity consumers/producers, e.g., wind/solar/storage farms, fossil-fuel plants, industrial/commercial loads, or load-serving aggregators, all modeled as stochastic dynamical systems. In each time period, each consumes/supplies some electrical energy. Each such agent’s utility is the benefit accrued from its consumption or the negative of its generation cost. The social welfare, the sum of all these utilities, is the total benefit accrued from all consumption minus the total cost of generation. The Independent System Operator is charged with maximizing the social welfare subject to total generation equalling consumption in each time period, but without the agents revealing their system states, dynamic models, utility functions or uncertainties. It has to announce prices after interacting with agents via bid-price interactions. This paper examines the case where the agents respond in a compliant price-taking manner.

It is shown that there is an iterative bid-price interaction, where agents respond to price announcements by complying to the requirement to announce their optimal responses according to their true stochastic model, or, in the case where the agents are LQG systems, according to deterministic versions of their true stochastic models, that leads to the same global maximum value of social welfare attainable if all agents had pooled their information. In the important LQG case, the bid-price iteration is dramatically simple, exchanging only real-valued vectors of future prices and consumptions/generations at each time step. Agents need not even know of the existence of other agents. DC Power Flow Equations can also be incorporated.

The results may be of broader interest via-a-vis general equilibrium theory of economics for stochastic dynamic agents.


I. INTRODUCTION

In the electricity grid, the power generated should equal the power consumed at all times, neglecting line losses. Unlike other commodities, electricity cannot be stored in the grid. This task of balancing them, and in the most economical way, is entrusted to the Independent System Operator (ISO) in deregulated electricity markets [1].

Energy is increasingly obtained from uncertain, dynamically varying renewables such as wind/solar. In the future, besides controlling traditional sources such as coal plants, demand will also be continually controlled to some extent to balance generation and consumption. Both generators and loads are dynamical systems that may possess thermal or other inertias. Both are also subject to uncertainties, such as wind, cloud cover, ambient temperature, or customer traffic. The problem we address is how the ISO can perform its task in the emerging scenario where loads and generators are stochastic dynamic systems.

The primary mechanism for coordinating entities is through time-varying stochastic prices. However, entities will need to know the probability distribution of future prices to plan their optimal consumption/generation over time. Future prices in turn depend on the dynamics of all agents and their future uncertainties. However, an entity is generally unaware of the dynamics or uncertainties of other agents or how they will respond, and may not wish to explicitly reveals the details of its dynamic model or utility function to the ISO. The problem we address is that of choosing generations/consumptions to achieve the same social welfare optimum that could have been attained with full state and model information of all agents, but through only a constrained price-bid framework between agents and the ISO.

The role played by price in coordinating agents was initiated by Walras [2] in general equilibrium theory. Arrow and Debreu [3]–[5] showed that a correct choice of prices for commodities ensures that for quasi-concave utility functions, a system of individual entities, where each optimizes its own response given prices, results in a systemwide Pareto optimal Subsequently, Arrow, Block and Hurwicz [6] showed that the prices can be discovered by Walrasian tatonnement [2] under appropriate conditions such as gross substitutes. Their theory extends to allow for uncertainty by simply considering each good under a different random state of nature as a different good, as shown by Arrow [7]. Subsequently, Radner [8] has shown the existence of prices corresponding to an equilibrium even if different agents have different random observations of the uncertainties. However, if agents communicate, the problem is more complicated. Radner emphasizes tractability, since schemes may require un-
bounded computation.

The idea of employing prices to perform this task in the electrical power domain was introduced by Caramanis, Bohn and Schweppe [9] and Bohn, Caramanis and Schweppe [10]. Hogan [11] further elaborated the detailed implementation of a locational marginal price-based electricity market operation. Fundamentally based on a static dispatch with no uncertainty, today’s electricity market design and corresponding price signal are simply not designed for achieving social welfare optimality for dynamic and uncertain generators and loads. The current market mechanism requires participants to make decoupled bids for separate time intervals. In the day-ahead market, a generator has to bid a price-generation curve for the 8am-9am slot, another separate curve for the 9am-10am slot, and so on, for each hour of the next day. However, generators have ramping constraints, such as 50 MW/hour, which give rise to inter-temporal constraints between different time slots. These are typically handled by ad hoc out-of-market (OOM) merit order measures [12]. The bidding procedure fundamentally does not allow a generator to bid a time function even though that is critical to its operation. Even more challengingly, in the real-time market, the bidding process does not allow a participant to optimize with respect to stochastic process models of uncertain resources. There have been many studies on the potential problems associated with this market design, such as unnecessarily price volatility [13], network externalities [14], and lack of investment signals [15]. While the deterministic and static approach to approximating the underlying dynamic and stochastic power system may be practically appealing without much loss of optimality in conventional systems, it is not so for demand response and intermittent renewables [16].

The ultimate holy grail is to determine a bidding scheme for the strategic case where agents may be untruthful. This paper takes a first step by examining in the simpler case, where agents are complaint and price-takers, whether we can (i) determine how to compute what is the “price” in real-time in the case of dynamic stochastic agents as the system evolves; (ii) determine how compliant agents (i.e., price-takers) and ISO should interact; (iii) determine whether this could be achieved in a computationally tractable manner; and (iv) if so, for what classes of models. The next phase to be addressed in a future paper is to determine whether and how this price-taking solution can be incentive-compatibly approximated in the case where the agents are strategic and may lie.

The paper is organized as follows. Section II presents the main results. Section III formulates the ISO problem. Section IV presents a preliminary modification to a straightforward bid-price iteration to satisfy energy balance. The case of general stochastic dynamic agents is addressed in Section V. The case where all agents are LQG systems is addressed in Section VI. The LQG case with DC power flow constraints or other linear constraints is treated in Section VII. Section VIII presents the results of illustrative simulations, concluding in Section IX.

II. The main results

(i) For a system where different agents are subject to different private uncertainties, a bid-price iteration scheme between the agents and the ISO, repeated at every time instant, where the ISO employs an averaging procedure to determine allocations, does achieve the global maximum of the social welfare that would have been attainable were all agents pooling their information (Theorem 2). Such bid-price interaction between the agents and the ISO, or some form of “two-way exploration” by the ISO, appears to be needed to be able to assign optimal consumptions/generations, since even exchanging supply and demand functions and knowledge of optimal prices, without any two-way interaction, is inadequate even in a static deterministic system to determine optimal allocation of consumptions/generations or even assure energy balance, when utility functions are concave but not necessarily strictly concave. (Example 1).

This result is related to the work of Radner [8], which extends the work of Arrow [7] and Debreu [17] to the case of general information structures. Under conditions on the preference functions of agents similar to those investigated by Arrow-Debreu, he shows that there exist equilibrium prices under which all agents can choose optimal actions, with the resulting optimal actions attaining balance between supply and demand. Radner does not specifically address how these equilibrium prices are to be obtained. In the second portion of Radner’s work in [8], he formulates the problem where agents can send messages to each other during the evolution of the system. In our work, we aim not at the particular equilibrium that would result from the private information structure, but at the social welfare optimal solution that could be attained if all agents had pooled their information and made it available to all. To achieve this, we specifically delve into what sorts of “messages” (using the language of Radner) need to be exchanged between agents in order to achieve this particular social welfare optimal solution. Of specific interest is how this can be done through the restrictive medium of an ISO, and not by arbitrary bilateral exchanges of information between agents, with the interaction between the ISO and the agents being price announcements and consumption/generation bids. This problem involves the problem of recovering primal optimal solution from dual optimal prices. We exploit recent results from [18] that employ averaging of iterates to obtain a primal solution that is feasible, in this case, satisfying energy balance, and optimal. This requires an interactive scheme of exploratory price announcements by the ISO, and the agents responding to the ISO in a price-taking fashion. We thereby provide a solution that involves a bid-price interaction between agents and the ISO at each time step and attains the goal of maximizing social welfare. We note that we do not make a “non-satiation” assumption, as in Radner.

(ii) In the case where all agents are linear quadratic Gaussian (LQG) systems with private uncertainties, and
are price-takers compliant to the requirement to bid with respect to a deterministic model of their systems, there is a simple bid-price iteration scheme, where each agent’s bid is only a simple function of time, which attains the maximal social welfare attainable when all agents pool their information (Theorem 3). This is a substantial simplification since agents don’t have to bid for all future uncertainty realizations. Also, it does not require any subsequent averaging of the bids by the ISO. We should note that this problem falls outside work on team theory since unlike in team theory the agents do not know the models of the dynamics of other agents or their utility functions. Also, almost sure constraints, such as balancing, are not typically considered in team problems. In his work examining general information structures, Radner emphasizes that in the case of general information structures the solutions may require unbounded computation. The LQG problem may therefore be of interest as a special case of the simplified model. This potentially provides an implementable scheme. This is illustrated through simulation examples in Section VIII.

Remark 1: The time horizon $T$ could be 96, corresponding to one day of 15 minute slots in the real-time market. However, there is considerable flexibility to incorporate other scenarios. For example, one can model the risk-limited dispatch of [19] where purchases of forward energy are made for blocks of time, with blocks getting shorter as operations approach real time. The uniform distribution over $W$ is not essential; any distribution is fine and it could even be time varying, as long as it is known to all agents; $w_i(t)$ could be the “innovations process” of the uncertainty process of agent $i$.

Consumption/generation constraints. Let $w_i^t := (w_i(0), w_i(1), ..., w_i(t))$ denote the past of $w_i$, and similarly define $w_i^t(\omega)$ for the case where the power flows can be modeled through additional linear constraints if desired. Agent $i$’s choice has to satisfy the local capacity constraints $F_i(w_i^t, w_i^t, t)u_i(t) \leq g_i(w_i^t, w_i^t(t)) + \sum_{s=0}^{t-1} J_i(w_i^t, w_i^t(s), s,t)u_i(s)$ and $h_i(w_i^t, w_i^t, t, u_i(t)) \leq 0$ for each $t$, almost surely. The affineness of the former constraints in past $u_i$’s can be used to model, for example, constraints on the rate of ramping up of coal plant output, the dependence on $t$ allows for seasonality, and the dependence on $w_i, w_c$ allows random effects on capacity. The one-step cost function of an agent $i$, $1 \leq i \leq M$, denoted $c_i(x_i, u_i(t))$ or its negative, the one-step utility function $-c_i(x_i, u_i(t))$, is a function of its state and action, in period $t$. For producers, this could be the cost of labor or coal. For consumers, this could represent the cost incurred due to the high temperature of a house/business facility, or due to a delay in performing a task resulting from inadequate purchase of electricity, or the negative of some benefit of the electricity usage.

Pigovian taxes may be imposed on negative externalities such as pollution [20], e.g., a carbon tax, with one-step tax $e_i(x_i, u_i)$. By allowing $e_i(x_i, u_i)$ to be positive/negative, we can also allow for cross-subsidies to be incorporated. Energy balance should be maintained in each period, i.e., $\sum_{i=1}^{M} u_i(t) = 0$ for all $t = 0, 1, ..., T - 1$ almost surely. We will allow even more general linear vector constraints: $\sum_{i=1}^{M} K_i(t)u_i(t) = d(t)$ almost surely, which can be used to capture other constraints, e.g., DC power flow constraints as is done in Section VII, or budget balance under linear cross-subsidies. As earlier, the constraints are to hold almost surely.

Remark 2: Beyond balancing generation and consumption, the ISO also needs to model the electrical transmission network governed by algebraic equations based on Kirchoff’s laws that have to be satisfied by the voltage and current magnitudes and phase angles. They impose constraints on $\{u_i(t), 1 \leq i \leq M\}$. They can be modeled through additional linear constraints if one employs an approximation to the AC Power Flow equations called DC Power Flow equations. For linear levies, $e_i(x_i, u_i(t)) := e_i(x_i, u_i(t))$, ISO budget balance, $\sum_{i=1}^{M} e_i(x_i, u_i(t))u_i(t) = 0$, can be incorporated as an additional linear constraint, if desired.
Knowledge available to the agents and the ISO. In the general case, we assume that all agents and the ISO know the alphabet $\mathcal{W}$. Each agent also knows its own one-step cost function and its own constraints. In the special case of LQG systems in Section VI we will only assume that each agent only knows its own LQG system.

Remark 3: These are severe constraints on the information available to the ISO. It does not know the states, dynamic models, or utility functions of individual agents. Agents may be averse to disclosing information for competitive reasons or to ensure privacy.\(^1\) Moreover, even if all agents were willing to share all their information with the ISO, it would be such an intractably large amount of information, amounting to a complete state of the world, that the ISO would not be able to handle it with acceptable complexity and delay anyway.

The Independent System Operator (ISO) solicits electricity purchase/sale bids from the agents in each time slot $t = 0, 1, \ldots, T - 1$, and announces prices. Our model allows for agents and ISO to iterate on the bids. Once the price iterations have converged, the ISO declares the market clearing prices, and the electrical energy to be consumed/generated by the agents.

Bidding schemes allow the ISO and agents to reach a solution for prices and generations/consumptions. Depending on the assumptions made about the system model, there will be different bidding schemes. An example is the following. Consider time $s$. The ISO announces a price sequence for current and future times $s \leq t \leq T - 1$, or future events, to all agents. Agent $i$ bids the amount of electricity it is willing to purchase/generate at the current and future times $s \leq t \leq T - 1$, or at future events, at the prices indicated by the ISO. After collecting the bids, the ISO updates the prices. An iteration of price updates followed by bid updates continues till the prices and the bids converge, and then the ISO announces the allocations of generations/consumptions to agents for the current period $s$. This entire process can be repeated in each discrete time slot $s$ in real-time.

Total system operating cost, the negative of the social welfare, is the sum of the expected value of the finite horizon total of the one-step costs incurred by all the agents plus any taxes, $\mathbb{E} \sum_{s=1}^{M+1} \sum_{t=0}^{T-1} [c_i(x_i(t), u_i(t), t) + c_i(u_i(t), t)]$. It is the total electricity generation cost plus any taxes assessed minus the utility provided to the consumers. The expectation above is taken with respect to the combined uncertainty or “noise” process $w(t) := (w_c(t), w_1(t), w_2(t), \ldots, w_M(t))$ for $t = 0, 1, \ldots, T - 1$, consisting of all the private uncertainties and the common uncertainties, as well as the random initial conditions of all the $M$ agents.

Remark 4: The utility of a load is the “benefit” that the load accrues from the consumed power. The utility of a generator is the negative of its cost of generation.

The total of all agents’ utilities, called social welfare, is therefore the benefit of the power consumed minus the cost of generating it.

**Goal of social welfare maximization:** Let $\mathcal{F}_t$ be the $\sigma$-algebra generated by all the noises up to time $t$, private and common, as well as all initial conditions. This represents the complete information available to all agents. If $u(t)$ is allowed to be adapted to $\mathcal{F}_t$, we call it the full state information case. Now we come to the stringent goal of this paper. We would like to determine bidding schemes that attain the same maximum value of the social welfare as could be attained in the full state information case.

The resulting ISO Problem is:

$$\min \mathbb{E} \sum_{t=0}^{T-1} \sum_{i=1}^{M} [c_i(x_i(t), u_i(t), t) + c_i(u_i(t), t)]$$

such that $x_i(t+1) = f_i(x_i(t), u_i(t), w_i(t), w_c(t), t)$ a.s.;

with capacity constraints $h_i(w_i^1, w_c^1, t, u_i(t)) \leq 0$ a.s.,

$$F_i(w_i^1, w_c^1, t)u_i(t) \leq g_i(w_i^1, w_c^1, t)$$

$$+ \sum_{s=0}^{t-1} J_i(w_i^1, w_c^1, s, t)u_i(s) \text{ a.s.,}$$

$$\sum_{i=1}^{M} K_i(t)u_i(t) = d(t) \text{ a.s. for } 0 \leq t \leq T - 1.$$ 

The central issue is the following: How should the ISO determine pricing and allocations to dynamic stochastic agents so that the overall system is as optimal as it could be in the full state information case, even though agents do not know each other’s states, constraints, dynamic models or cost functions, and neither does the ISO? Due to the lack of knowledge of other agents’ dynamic models or cost functions, this problem falls outside of usual stochastic control. Though written as an optimization problem in (1-21), it is not a standard one since the agents do even know the quantities involved in the optimization problem.

Remark 5: A more general formulation of the problem would take into consideration the power flows in the electrical network connecting the nodes. These power flows are determined by the AC Power Flow Equations [21], an approximation of which leads to the so-called DC Power Flow equations which are linear equations that are typically used in today’s market clearing models [21]. These power flows can be incorporated by adding a set of linear constraints as in (21) at each node. The results of the paper extend to this situation. For brevity we only illustrate how this generalization proceeds in the special case of LQG systems in Section VII.

Remark 6: There have been many efforts since the deregulation of the electricity sector on a market-based framework to clear the system. Today’s locational marginal price-based nodal market design is based on seminal work in [10], [11]. This has been followed up by a large body of literature focusing on designing a transmission pricing
mechanism in support of an efficient market [22] [23]. The naive belief that deregulation of electricity industry would simply work was critically re-assessed following the Enron crisis and lack of long-term investment incentives [24] [25]. There has been pioneering work on game theoretic approaches to modeling the market power issues in the electricity market [26] [27]. With increasing penetration of stochastic resources, there have been efforts at designing a market bidding mechanism that achieves the social welfare optimum. Ilic et al. [28] have proposed a two-layered approach that internalizes individual constraints of market participants while allowing the ISO to manage the spatial complexity. References [29], [30] contain some heuristic approaches. Reference [31] applies progressive hedging to deal with uncertainties on the progression of stochastic resources, there have been efforts to theoretic approaches to modeling the market power issues [24] [25]. There has been pioneering work on game Enron crisis and lack of long-term investment incentives [22] [23].

From (ii), Slater’s Condition: There exists a feasible is compact, and (5,6,7) has an optimal solution.

Even if prices are correctly determined, it turns out that the difficulty through an example, and then present a correct prices need not bid solutions that lead to energy balance. This necessitates modifications to the bidding mechanism in support of an efficient market [22] [23]. The naive belief that deregulation of electricity industry would simply work was critically re-assessed following the Enron crisis and lack of long-term investment incentives [24] [25]. There has been pioneering work on game theoretic approaches to modeling the market power issues in the electricity market [26] [27]. With increasing penetration of stochastic resources, there have been efforts at designing a market bidding mechanism that achieves the social welfare optimum. Ilic et al. [28] have proposed a two-layered approach that internalizes individual constraints of market participants while allowing the ISO to manage the spatial complexity. References [29], [30] contain some heuristic approaches. Reference [31] applies progressive hedging to deal with uncertainties on the progression of stochastic resources, there have been efforts to theoretic approaches to modeling the market power issues [24] [25]. There has been pioneering work on game Enron crisis and lack of long-term investment incentives [22] [23].

IV. A PRELIMINARY ISSUE IN PRICE-BASED COORDINATION

Even if prices are correctly determined, it turns out that agents bidding in a price-taking manner in response to the correct prices need not bid solutions that lead to energy balance. This necessitates modifications to the bidding mechanism in support of an efficient market [22] [23]. The naive belief that deregulation of electricity industry would simply work was critically re-assessed following the Enron crisis and lack of long-term investment incentives [24] [25]. There has been pioneering work on game theoretic approaches to modeling the market power issues in the electricity market [26] [27]. With increasing penetration of stochastic resources, there have been efforts at designing a market bidding mechanism that achieves the social welfare optimum. Ilic et al. [28] have proposed a two-layered approach that internalizes individual constraints of market participants while allowing the ISO to manage the spatial complexity. References [29], [30] contain some heuristic approaches. Reference [31] applies progressive hedging to deal with uncertainties on the progression of stochastic resources, there have been efforts to theoretic approaches to modeling the market power issues [24] [25]. There has been pioneering work on game Enron crisis and lack of long-term investment incentives [22] [23].

Consider the simple situation where all generators and consumers are static and deterministic:

\[
\begin{align*}
\min_{u_1, u_2, \ldots, u_M} & \sum_{i=1}^{M} [c_i(u_i) + e_i(u_i)], \\
\text{subject to:} & \quad F_i u_i \leq g_i, h_i(u_i) \leq 0 \quad \text{for } 1 \leq i \leq M, \\
\text{and} & \quad \sum_{i=1}^{M} K_i u_i = d. 
\end{align*}
\]

**Assumption 1:**
(i) \(c_i(\cdot), e_i(\cdot), h_i(\cdot)\) are convex, \(\{u_i : F_i u_i \leq g_i, h_i(u_i) \leq 0\}\) is compact, and (5,6,7) has an optimal solution.
(ii) Slater’s Condition: There exists a feasible \(\bar{u}\), satisfying \(h_i(\bar{u}) < 0\) in \(\text{RelInt}(\text{Dom}(c_i)) \cap \text{RelInt}(\text{Dom}(e_i))\).

Dualizing only the constraint (7), and denoting \(u := (u_1, u_2, \ldots, u_M)\), yields, respectively, the Lagrangian, Dual function, and optimal reward of the Dual Problem:

\[
\begin{align*}
\mathcal{L}(u, \lambda) & := \sum_{i=1}^{M} [c_i(u_i) + e_i(u_i) + \lambda^T K_i u_i] - \lambda^T d, \\
D(\lambda) & := \min_{u : F_i u_i \leq g_i, h_i(u_i) \leq 0 \ \forall i} \mathcal{L}(u, \lambda), \\
J^* & := \max_{\lambda} D(\lambda) = D(\lambda^*). 
\end{align*}
\]

From (ii), \(J^*\) is also the optimal cost of the Primal (5).

Since \(D(\lambda)\) can be decomposed agent-by-agent as

\[
D(\lambda) = \sum_{i=1}^{M} \min_{u_i : \mathbb{R}^n} \left[ c_i(u_i) + e_i(u_i) + \lambda^T K_i u_i \right] - \lambda^T d, 
\]

the ISO can conceivably simply announce the “optimal price” \(\lambda^*\) per unit of power as that which attains the max in (8), along with the additional levy \(e_i(u_i)\) on agent \(i\). Each agent \(i\) can then respond with either its generation \(-u_i^*\) or consumption \(u_i^*\) that minimizes its “net” disutility \(c_i(u_i) + e_i(u_i) + \lambda^* u_i\) over (6). The ISO can finally announce the generation/consumption allocations to the agents.

Since agents do not disclose their cost functions, there needs to be a price discovery process, as in a Walrasian auction [32]. The ISO’s price needs to be reduced/increased according to whether the agents’ response results in excess total generation/consumption. We consider the following iterative bid-price process:

\[
\begin{align*}
\lambda^{k+1} &= \lambda^k + \frac{1}{K} \sum_{i=1}^{M} K_i u_i^k - d, \\
u_i^{k+1} &= \arg\min_{u_i : \mathbb{R}^n} [c_i(u_i) + e_i(u_i) + (\lambda^{k+1})^T K_i u_i]. 
\end{align*}
\]

This iteration of prices\(^2\) and bids is a subgradient algorithm that converges to an optimal price for the Dual under Assumption 1 [33].

However, the recovery of optimal generations/consumptions from optimal price is problematic.

Example 1 (Counterexample to generation/consumption recovery from optimal price): Consider one generator and one load. The generator’s cost of producing \(-u_1\) units of energy is \(-\frac{2}{3} u_1\), with \(u_1\) restricted to \([-1, 0]\), and it is assessed a Pigovian tax \(-\frac{1}{3} u_1\). The load’s utility from consuming \(u_2\) units of energy is \(\log(1 + u_2)\) with \(u_2\) restricted to \([0, 2]\). Energy should be balanced. The social welfare problem is:

\[
\begin{align*}
\min & \quad -\frac{2}{3} u_1 - \frac{1}{10} u_1 - \log(1 + u_2) \\
\text{Subject to:} & \quad -1 \leq u_1 \leq 0, 0 \leq u_2 \leq 2, u_1 + u_2 = 0.
\end{align*}
\]

The optimal solution is \((u_1^*, u_2^*) = (-1, 1)\).

The Dual function of price \(\lambda\) is

\[
D(\lambda) = \min_{1 \leq u_1 \leq 0} \left[ -\frac{1}{2} u_1 + \lambda u_1 \right] + \lambda u_2).
\]

The minimizers and minimum, \((u_1(\lambda), u_2(\lambda), D(\lambda))\), are

\[
\left\{
\begin{array}{ll}
(0, \min(\frac{1}{\lambda} - 1, 2), 1 - \lambda + \log \lambda) & \text{if } \lambda < \frac{1}{2}, \\
(\text{any point in } [0, 1], \frac{1}{\lambda} - 1, 1 - \lambda + \log \lambda) & \text{if } \lambda = \frac{1}{2}, \\
(1, -\lambda + \frac{1}{2} + \log \lambda) & \text{if } \frac{1}{2} < \lambda \leq 1, \\
(1, 0, -\lambda + \frac{1}{2}) & \text{if } 1 < \lambda,
\end{array}
\right.
\]

The optimal solution of the Dual is \(\lambda^* = \frac{1}{2}\).

However, when the price \(\lambda^* = \frac{1}{2}\) is announced by the ISO, the generator can bid \(-u_1 = 0\) since any point in

\(^2\)The gain \(\frac{1}{\delta}\) can be replaced by \(\frac{1}{\delta^2}\) for \(\frac{1}{2} < \delta \leq 1\) with \(\alpha > 0\) in Sections IV and V.
[0, 1] is optimal. The load’s bid is $u_2 = 1$, and there will not be balance between generation and consumption. □

Therefore one cannot determine the optimal bids from agents’ responses to the optimal prices. However, they can be obtained from the iterations of the bidding process by taking weighted averages of previous bids [18].

**Theorem 1 (Determining optimal bids by generators and loads [18]):** Consider the bid-price iteration scheme (9, 10) for the problem (5, 6, 7) under Assumption 1.

(i) Then $\lambda^k \to \lambda^*$, the optimal price.

(ii) Let $\theta \geq 0$. Suppose the ISO recursively averages the obtained bids as follows:

$$\bar{u}_i^k = \frac{\sum_{s=1}^{k-1} s^\theta \bar{u}_i^{s-1} + k^\theta}{\sum_{s=1}^{k} s^\theta} u_i^k; \quad \bar{u}_i^0 = u_i^0. \quad (11)$$

Then $\bar{u}_i^k \to u_i^*$ which is optimal for (5). □

A larger $\theta$ weights more recent values of the iterates for $u_i$, while $\theta = 0$ takes a plain average.

**Example 2 (Continued):** Choosing $\theta = 2$, one obtains:

$$\{\lambda_k\} : 0, 0.1, 0.6667, 0.5416, \ldots \to \frac{1}{2},$$

$$\{u^k\} : \left[\begin{array}{c}
-1 \\
0
\end{array}\right], \left[\begin{array}{c}
0.9412 \\
0.1176
\end{array}\right], \left[\begin{array}{c}
-0.9898 \\
0.4133
\end{array}\right],$$

$$\left[\begin{array}{c}
0 - 0.9972 \\
0.7263
\end{array}\right], \left[\begin{array}{c}
-0.9990 \\
0.9526
\end{array}\right], \ldots \to \left[\begin{array}{c}
-1 \\
1
\end{array}\right]. \quad \Box$$

This example illustrates the important point that knowledge of the optimal price alone is not sufficient to determine optimal generations/consumptions of agents. In fact the information gathered during the very process of iterative bidding is itself important. The journey is as important as the destination. The convergence rate of averaging methods is a topic of active research, cf. [18].

**Remark 7:** This difficulty in price-based coordination in the absence of strong convexity is also addressed in the work of Culioli and Cohen [34]. They propose an auxiliary problem principle that appends a differentiable strongly convex function, an approach which subsumes many variants. Employing a quadratic auxiliary function, it yields a variant that requires each agent to add an additional term $(u_i - u_i^k)^2$ to the criterion minimized in (10) at each iterate. It therefore requires asking the agents to bid according to some artificially specified cost. In contrast, our approach allows them to make their natural bids for generation/consumption based on the price announcement they have heard, but shifts the modification instead to the ISO side which is required at the end to simply allocate weighted averages of previous bids. Importantly, this however entails no loss for the agents since the weighted averages are guaranteed to be optimal responses for them at the converged price.

V. STOCHASTIC DYNAMIC AGENTS: BID-PRICE ITERATIONS

We now turn to the general problem of interest where the agents are stochastic dynamic systems. Denote the combined state of the system by $x(t) := (x_1(t), x_2(t), \ldots, x_M(t))$, and the combined actions by $u(t) := (u_1(t), u_2(t), \ldots, u_M(t))$.

A tree visualization of the system randomness, as in Fig. 1, is helpful. Suppose that $w(t)$ assumes only finitely many values. We can then construct an uncertainty tree of depth $T$, with the root node corresponding to the initial system state at level 0, and each sequence of transpired noises $\{w(0, \omega), w(1, \omega), \ldots, w(s-1, \omega)\}$ corresponding to a specific node $v$ at the level $s$. Note that the red and blue subtrees with roots at nodes $v'$ and $v''$ at the same level are identical due to the i.i.d. nature of the noise.

At each time $s$, a sequence of iterative tentative price announcements by the ISO for each node at or below the current node at level $s$, followed by tentative bids by all agents for such nodes responding optimally to the price announcement, takes place, until they converge. At each iteration, the ISO revises the tentative price announcement to drive the “excess consumption” at each node towards zero, and agents respond optimally according to their own cost-to-go function. This iteration of tentative prices and tentative bids continues till the prices converge. At that point the ISO announces and agents consume/generate the weighted average amount they bid for the particular node occupied at time $s$. The system then moves forward to time $s + 1$, arriving at a random node at level $s + 1$ according to $w(s)$, and the entire process is repeated. This is in the same fashion as Model Predictive Control.

**Bid-Price Iteration**

At each time $s$, the agents and the ISO engage in a Bid-Price Iteration as follows:

**k-th Bid Update at time $s$:** Suppose that the ISO has declared a price stochastic process $\lambda^k_s$ at time $s$ that associates a price with each node downstream of the
current node at level \((s+1)\) that the system is at, where \(k\) is an index that we will use for iteration. Since neither the ISO nor the agents know exactly which node they are at due to the existence of private noises, they take advantage of the fact that all subtrees of nodes rooted at level \(s+1\) are identical, and simply perform these announcements for a tree of depth \(T-s\), with the understanding that these bids apply to the subtree rooted at wherever the current node is among the nodes at level \(s+1\). In the Bid Update, each agent \(i\), in response, chooses its bid stochastic process \((u_{i,s}^k(s), u_{i,s}^k(s+1), \ldots, u_{i,s}^k(T-1))\) in response to the price stochastic process \(\lambda_s^k\) by solving the following problem, dubbed Agent \(i\)'s Problem, 
\[
\min_{u_i \text{ s.t. } (2.3)} \mathbb{E} \left[ \sum_{t=s}^{T-1} [c_i(x_i(t), u_i(t), t) + e_i(u_i(t), t) + \lambda^k(t)K_i(t)u_i(t)] | \mathcal{F}_{i,t} \right], \tag{12}
\]
where \(\mathcal{F}_{i,t} := \sigma(x_i(0), u_i(1), \ldots, w_i(t - 1), u_i(1), \ldots, w_i(t - 1))\) denotes the sigma-algebra generated by agent \(i\)'s observations up to time \(t\). It thereby generates a consumption/ generation for each node at a lower level in the subtree with root at the present node that the system is at.

**(K+1)-th Price Update at time \(s\):** The ISO updates the price stochastic process in response to the agents’ bids. (We note that in (12), the variables \(x_i(s+1), x_i(s+2), \ldots, x_i(T-1)\) are not the actual states of the agent at those times, but are the variables for a hypothetical problem that needs to be solved by the agent at time \(s\) in order to determine its bid at time \(s\). This hypothetical problem takes as input the price stochastic process, including its law, as announced by the ISO). Guided by the “excess consumption function” \(\sum_{i=1}^{M} K_i(t)u_i^k(t) - d(t)\), it raises or lowers prices to satisfy the general linear constraint, as follows:
\[
\lambda_{s+1}^k(t) = \lambda_{s}^k(t) + \frac{1}{k} \left( \sum_{i=1}^{M} K_i(t)u_i^k(t) - d(t) \right), \quad s \leq t \leq T-1. \tag{13}
\]

**The averaged allocations of consumption/generation:** At time \(s\), after the prices have converged, i.e.,
\[
\lambda_{s}^k(t) := \lim_{k \to \infty} \lambda_{s}^k(t) \text{ for } s \leq t \leq T-1, \tag{14}
\]
the ISO announces the allocations at the current time \(s\) as the limit
\[
u_{i,s}^*(s) := \lim_{k \to \infty} \bar{u}_{i,s}^k(s), \tag{15}
\]
of the following average of the iterates of the bids for time \(s\),
\[
\bar{u}_{i,s}^k(s) = \frac{1}{s} \sum_{\theta = 1}^{s} u_{i,s}^k(\theta) = \frac{1}{s} \sum_{s=1}^{k} s^\theta u_{i,s}^k(s) + \frac{k^\theta - 1}{\theta} u_{i,s}^0(s), \tag{16}
\]
with \(u_{i,s}^0(s) = u_{i,s}^0(s)\).

This is presented in Algorithm 1.

**Assumption 2:**
(i) There is an optimal solution of (1) with finite cost.
(ii) \(\sum_{t=0}^{T-1} c_i(x_i(t), u_i(t), t), \sum_{t=0}^{T-1} e_i(u_i(t), t), \text{ and } h_i(w_i^t, w_i^t, u_i(t))\) are convex in \(w_i^T, w_i^t, u_i(t)\) for each noise sequence \(w_i^T, w_i^t\), expanded out recursively and written in terms of \(w_i^{-1}\) and \(w_t\).
(iii) For each fixed noise sequence \(w_i^t, w_i^t\), there exists a feasible \(\bar{u}\) satisfying \(h_i(w_i^t, w_i^t, u_i(t)) < 0\) in \(\text{RelInt(Dom}(e_i)) \cap \text{RelInt(Dom}(e_i))\) for \(1 \leq i \leq M, 0 \leq t \leq T-1\).

**Theorem 2:** The above bid-price solution, with price updates (13), bid updates determined as the optimal solution of (12), and allocations at each \(t\) given by \(\lim_{k \to \infty} \bar{u}_{i,s}^k\) where \(\bar{u}_{i,s}^k\) is obtained as the averaged version of \(u_{i,s}^k\) as in (11), achieves the maximum social welfare that could have been attained in the full state information case, when the cost functions satisfy Assumption 2.

**Proof:** First we analyze how the above Iterative Bidding Scheme functions in the case of common uncertainties, i.e., \(w_i(t) \equiv w_i(t)\), and there are no private noises \(w_i\). In this case it will turn out that we need to conduct the bid-price iteration only at time 0.

Let us suppose that \(x(0)\) is fixed, without loss of generality.

For simplicity of exposition only we suppose that the noise processes \(w_i(t), w_i(t)\) assume only finitely many values, allowing them to be represented by a tree as in Fig. 1. Let \(\delta(v)\) denote the depth \(t\) of a node \(v = (w(0,0), w(1,1), \ldots, w(t-1,1))\) in the tree of Fig. 1; it is the time \(t\) associated with the node \(v\). Given a Markov policy \(\pi\) that takes action \(u = \pi(x)\) when the state \(x(s) = x\), and a noise realization \(\nu\), one can determine the resulting state \(x(t)\) by recursively applying the system function \(x(s + 1) = f(x(s), u(s), w(s), s)\) for

**Algorithm 1:** Stochastic Dynamic Agents with Private Uncertainties

```plaintext
for bidding times \(s = 0\) to \(T - 1\) do
    \(k = 0\)
    repeat
        Each agent \(i\) solves the problem
        \[
        \min_{u_i} \mathbb{E} \left[ \sum_{t=s}^{T-1} [c_i(x_i(t), u_i(t), t) + e_i(u_i(t), t) + \lambda^k(t)K_i(t)u_i(t)] | \mathcal{F}_{i,t} \right],
        \]
        with initial condition \(x_i(s)\) to obtain the optimal \(\{u_i^k(s), s \leq t \leq T - 1\}\) subject to (17,18), and submits it to the ISO.
        The ISO declares new prices for \(s \leq t \leq T - 1\), as
        \[
        \lambda_i^{k+1}(t) = \lambda_i^k(t) + \frac{1}{k} \left( \sum_{i=1}^{M} K_i(t)u_i^k(t) - d(t) \right).
        \]
        \(k \rightarrow k + 1\)
    until \(\lambda_i^k(t)\) converges a.s. to \(\lambda_i^*(t)\) for \(s \leq t \leq T - 1\).
```

ISO computes \(\bar{u}_{i,s}^k(s)\) as in (16), and announces generations/consumptions \(u_{i,s}^k(s) := \lim_{k \to \infty} \bar{u}_{i,s}^k(s)\).

end for
$s = 0, 1, \ldots, t - 1$. One can thereby also determine the action $u(t) = \pi(x(t), t)$. Thus, for a given Markov policy $\pi$, one has a mapping $v \mapsto (x, u)$ that specifies the state reached, and the action taken, at each node $v$. Since the policy is Markov, this mapping has to satisfy the consistency condition that two nodes that are associated with the same state $x$, and that are at the same level (and hence correspond to the same time $t$), have to take the same action. Now let us consider a more general tree policy $\sigma$ that specifies an action $\sigma(v)$ to be taken in each node $v$ without requiring this consistency condition. The class of tree policies is therefore more general than the class of Markov policies. Since the class of Markov policies contains an optimal policy, so does the class of tree policies.

Denote by $\sigma^v := \{u(0), \ldots, u(t)\}$ the sequence of actions taken in the preceding $t + 1$ steps, where $t$ denotes the depth of node $v$. The state $x(t)$ corresponding to $v$ is thereby determined by $(v, \sigma^v)$. Let $\sigma_i(v) := u_i(t)$ and $\sigma^v_i := \{u_i(0), \ldots, u_i(t)\}$ denote the corresponding actions by agent $i$. The problem (1) can then be written equivalently as the following optimization problem, with a decision variable $\sigma(v)$ associated with each node:

$$\begin{align*}
\min & \sum_{i=1}^{M} \sum_v p_v [c_i(v, \sigma^v) + e_i(\sigma^v)] \\
F_i(v)\sigma_i(v) & \leq g_i(v) + \sum_{\{v' : \delta(v') < \delta(v)\}} J_i(\delta(v'), v)\sigma_i(v'), \\
h_i(\sigma_i(v), v) & \leq 0, \quad (17) \\
\text{such that} & \sum_{i=1}^{M} K_i(v)\sigma_i(v) = d(v), \forall v. \quad (18)
\end{align*}$$

Under Assumption 2, the convex programming problem has no duality gap. Let $\lambda(v)$ be the Lagrange multiplier for the constraint $\sum_{i=1}^{M} K_i(v)\sigma_i(v) = d(v)$, and define the vector $\lambda := \{\lambda(v)\}$. We obtain,

$$\mathcal{L}(\sigma, \lambda) := \sum_{i=1}^{M} \sum_v p_v [c_i(v, u^v) + e_i(u^v) + \lambda(v)^T K_i(v)\sigma_i(v)] - \sum_v p_v \lambda(v)^T d(v).$$

The process $\lambda(v)$ is the “price process”. Each agent submits a bid for each possible future realization $v$ of the noise process, while the ISO specifies a price at each $v$. Now the rest of the proof for the case of common uncertainties parallels the deterministic one of Theorem 1.

The key point to notice is that all the optimal actions and prices corresponding to each node can be determined at time 0. After that, since the uncertainties are all common and all agents can observe the uncertainty, all agents know exactly which node at level $t + 1$ the noise process is at at each time $t$. Hence each agent knows which action to apply at each time as the system evolves over time.

Now we turn to the case with private uncertainties. In the private uncertainty case also, all the optimal actions and prices corresponding to each node can be determined at time 0. They know what actions to apply at time 0, and the ISO knows what price to assign at time 0. However, the main problem is that at future times, the agents do not know which node at level $t + 1$ the noise process is at at time $t \geq 1$ due to the presence of private uncertainties. Hence they do not know what action to apply at future times $t \geq 1$.

However, at each future time $t \geq 1$ we can simply consider the state at that time as a fresh initial condition, and reconduct the bidding process, which then yields the optimal actions and prices at that time $t$. Thus the Bid-Price Iteration repeated at each time, with only the initial action and price implemented at each time, leads to an optimal solution for the system with private uncertainties.

The major drawback of this algorithm is that it is exponentially complex in $T$ due to the number of states in the tree, even if each $w(t)$ is finite valued. One may note the following gradation of complexity in the ISO Problem.

1) In the case where agents have private uncertainties, this is a complex iteration consisting of prices and bids for all future noise states that has to moreover be repeated at each time during the evolution of the system.

2) If the agents have only a common uncertainty, then as the proof of Theorem 2 shows, the complex iteration needs to be conducted only at the initial time.

3) If the system is a deterministic system, then the bids or price updates simplify to open-loop time-functions of future consumption/generation or prices, respectively. The reason is that there being no noise, future noise states are singletons. Moreover as in the common information case, of which this is special case, the bid-price iteration needs to be conducted only at the outset at time 0.

Remark 8: In the deterministic static and deterministic dynamic cases, the problem of bidding between consumers and generators has been extensively studied. Reference [35] studies the robustness and stability of a tattonment process that decides electricity prices for a single time horizon, and [36] takes a penalty function approach to design an iterative algorithm that yields optimal operation point and simultaneously regulates frequency, also for an essentially static set-up. Li, Chen, and Low [37] consider a bidding problem between several households and a single utility company for the deterministic dynamic case that features a price update involving the derivative of the cost of obtaining electricity from the wholesale market. Iterative algorithms similar to ours for the deterministic dynamic case are proposed in [38], [39] for particular models of dynamic systems, e.g., involving batteries. The common information case is handled in the work of Arrow [7].
VI. THE ISO PROBLEM FOR LQG AGENTS

We will now show that when the agents are all LQG systems, one can dramatically simplify the bids and the bidding process. Even though agents have private uncertainties, each bid consists of only a simple time-function, just as in the deterministic case. The only difference is that the bidding needs to be carried out at each time $t$. Similar to Model Predictive Control, only the first step of the prices and consumptions/generations at each time $t$ is implemented. This bidding scheme appears practically feasible with bid periods separated by minutes. Another simplifying feature is that the ISO need not average the bids. The bids of agents converge at each time instant without averaging to a feasible solution that satisfies energy balance and other constraints.

The $M$ agents have linear dynamics affected by Gaussian noise and have quadratic costs. Initial conditions and noises are Gaussian: $x_i(0) \sim N(0, \Sigma_i, 0)$ and $w_i(t) \sim w(0, P_i, t)$, and independent of all others. The cost functions of agents, are quadratic, with $Q_i \geq 0$ and $R_i > 0$. The ISO Problem is:

$$\text{Min } \mathbb{E} \sum_{t=0}^{T-1} \sum_{i=1}^{M} [x_i^T(t)Q_i x_i(t) + w_i^T(t)R_i w_i(t)]$$

where $x_i(t + 1) = A_i x_i(t) + B_i u_i(t) + w_i(t)$, \hspace{1cm} (19)

$$\sum_{i=1}^{M} u_i(t) = 0 \text{ a.s., for } t = 0, 1, \ldots, T - 1.$$ \hspace{1cm} (20)

The case of time-varying systems is entirely analogous.

Agents have no knowledge even of each other’s presence. Agent $i$ does not know the value of $M$, the number of agents, the matrices $\{A_j, B_j, Q_j, R_j, \Sigma_j, 0, P_j, \}$ and other agents, the realizations of their state processes $\{x_j(t), j \neq i\}$ or noises $\{w_j(t), j \neq i\}$.

The $k$-th iterate of the bid function submitted at time $s$ specifying the quantity of electricity that agent $i$ is willing to purchase at times $t = s, s + 1, \ldots, T - 1$ is not a function of the outcomes of the noise sequence $\{w(t), t > s\}$. It is simply a vector $(u_i^{s,k}(s), u_i^{s,k}(s + 1), \ldots, u_i^{s,k}(T - 1))$ comprised of $T - s + 1$ entries. The same is also true for prices. The ISO just specifies a vector $(\lambda_i^{s,k}(s), \lambda_i^{s,k}(s + 1), \ldots, \lambda_i^{s,k}(T - 1))$ of $T - s$ entries.

The key to showing the existence of such a simple bidding scheme lies in utilizing the certainty equivalence property of LQG systems [40].

**Algorithm 2**: ISO Problem with LQG Agents

for bidding times $s = 0$ to $T - 1$

$k = 0$

Initialize $\{\lambda_i^0(t) : s \leq t \leq T - 1\}$ arbitrarily.

repeat

Each agent $i$ solves the problem (22) for a deterministic system (23) with initial condition $x_i^0(s_i) = x_i(s), \text{ where } x_i(s)$ is the state of the $i$-th agent at time $s$, and submits the optimal values, denoted $u_i^k(s), \text{ for } s \leq t \leq T - 1 \text{ to the ISO}$. ISO updates the prices according to (25,26).

Increment $k$ by 1.

until $u_i^k(s) \text{ converges to } u_i^*(s)$,

Implement $(u_i^1(s), u_i^2(s), \ldots, u_i^M(s))$

end for

The iterative bidding scheme is illustrated in Fig. 2 and Algorithm 2. We bring attention to the following two assumptions that are formally stated in Theorem 3 below:

**Assumption 1**: The agents are assumed to be compliant in the sense that they conform to the described scheme.

**Assumption 2**: The scheme specifically requires that agents employ a deterministic version of their system for computing their bids.

This iterative bid-price scheme achieves the same optimal social welfare for the LQG ISO Problem attainable under full-state information, as described below.

**Theorem 3**: Consider the overall system comprised of the $M$ agents (20) for $i = 1, 2, \ldots, M$, where $x_i(0) \sim N(0, \Sigma_i, 0)$, $w_i(t) \sim w(0, P_i, t)$, and all random variables $\{x_i(s) : 1 \leq i \leq M\}, \{w_i(t) : 1 \leq i \leq M, 0 \leq t \leq T - 1\}$ are independent. Let $u(t) := (u_1(t), u_2(t), \ldots, u_M(t))$. Let $F_i := \sigma(\{x_i(0) : 1 \leq i \leq M\}, \{w_i(\tau) : 1 \leq i \leq M, 0 \leq \tau < t\})$ be the $\sigma$-algebra generated by all the uncertainties up to time $t$. It is desired to minimize the overall social welfare cost (19) subject to the balancing constraints (21), over all $u(t) : 0 \leq t \leq T - 1$ that are adapted to $F_i$.

Consider the following bid-price scheme where the price vector $(\lambda_i^0(s), \lambda_i^0(s + 1), \ldots, \lambda_i^0(T - 1))$ with real-
valued entries is initialized arbitrarily at each time \( s = 0, 1, \ldots, T - 1 \).

At time \( s \), in response to the \((k-1)\)-th price vector iterate \( (\lambda_s^k(s), \lambda_{s+1}^k(s), \ldots, \lambda_{s+T-1}^k) \) with real-valued entries, announced by the ISO, agent \( i \) announces the optimal open-loop sequence \( u_i^{s,k}(s), u_i^{s,k}(s+1), \ldots, u_i^{s,k}(T-1) \) for the following deterministic Linear Quadratic Regulator (LQR) problem:

\[
\min_{\{u_i^{s,k}(t)\}_{t=s}^{T-1}} \sum_{t=s}^{T-1} \left[ w_i^{s,k}(t)^T Q_i w_i^{s,k}(t) + u_i^{s,k}(t)^T R_i u_i^{s,k}(t) + \lambda_s^k(t) u_i^{s,k}(t) \right]
\]

\[
\text{s.t. } x_i^{s,k}(t+1) = A_i x_i^{s,k}(t) + B_i u_i^{s,k}(t), \quad s \leq t \leq T - 1,
\]

with initial condition \( x_i^{s,k}(s) := x_i(s) \).

(22)

The ISO then adjusts the price vector \( (\lambda_s^{k+1}(s), \lambda_s^{k+1}(s+1), \ldots, \lambda_s^{k+1}(T-1)) \) as:

\[
\lambda_s^{k+1}(t) = \lambda_s^k(t) + \alpha_k \sum_{i=1}^{M} u_i^{s,k}(t), \quad s \leq t \leq T - 1.
\]

(25)

where \( \{\alpha_k\} \) is any sequence of positive numbers satisfying

\[
\alpha_k > 0, \lim_k \alpha_k = 0, \quad \sum_{k=0}^{\infty} \alpha_k = +\infty.
\]

(26)

At time \( s \), the iterations in \( k \) are continued till the price iterations \( (\lambda_s^k(s), \lambda_s^k(s+1), \ldots, \lambda_s^k(T-1)) \) converge to \( (\lambda_s^*(s), \lambda_s^*(s+1), \ldots, \lambda_s^*(T-1)) \). Denote the corresponding limit of the input sequence of agent \( i \) by \( (u_i^{*}(s), u_i^{*}(s+1), \ldots, u_i^{*}(T-1)) \).

The price at time \( s \) is then set to \( \lambda_s^*(s) \) and each agent \( i \) applies the input \( u_i^*(s) \). This is repeated at time \( s+1 \).

Then the sequence \( \{u_i^*(t) : 0 \leq t \leq T - 1\} \) attains the minimum of the cost (19) over all control laws adapted to \( F_i \) that satisfy the balancing constraint (21).

**Proof:** Let

\[
X := (X_1, X_2, \ldots, X_M), U := (U_1, U_2, \ldots, U_M),
\]

\[
A := \text{diag}(A_1, A_2, \ldots, A_M), B := \text{diag}(B_1, B_2, \ldots, B_M),
\]

\[
Q = \text{diag}(Q_1, Q_2, \ldots, Q_M), R = \text{diag}(R_1, R_2, \ldots, R_M),
\]

and consider the following deterministic constrained LQR problem, with no noise, and featuring energy balance:

\[
\min_{\{u_i^*(t)\}_{t=0}^{T}} \sum_{t=0}^{T} [X^T(t)QX(t) + U^T(t)RU(t)], \text{ subject to } (27)
\]

\[
X(t+1) = AX(t) + BU(t); X(0) = x(0) \quad \text{and } (21).
\]

(28)

Since the state is affine in \( U \), after substituting for the states, we have a positive definite quadratic programming problem with equality constraints. The Karush-Kuhn-Tucker matrix is nonsingular (Section 10.1 of [41]) since \( R_i > 0 \), and so there are unique \( U^*, \lambda^* \) optimal for the primal and dual, respectively. The Dual function is a differentiable concave quadratic function, and the subgradient method is actually a gradient method that converges under non-summability of step-sizes, without even requiring square summability (Section 2.5 of [42]). The iterative bids \( U_i^k \) are affine functions of the prices \( \lambda^k \). The limiting prices yield bids that satisfy balancing. Hence this deterministic problem can be solved by the Bid-Price iteration involving only time-functions of prices and consumptions/generations between the agents and the ISO to obtain the optimal inputs \( U(t) \) for \( 0 \leq t \leq T - 1 \), as noted for the case of deterministic problems in Section V.

However, at the particular time \( s = 0 \) with \( x_{i,0}^k(0) = x_i(0) \), the Bid-Price iteration (22,23,24) and (25) is the same as the Bid-Price iteration in Section IV with the simplification that the constraint (5) is absent, and \( K_i = 1, d = 0 \) in (7). Hence the end result of Algorithm 2 at time \( s = 0 \) is the optimal action for (27,28),

\[
u(0) = U(0).\]

(29)

Now note that due to energy balance, no matter how the first \((M-1)\) agents choose their consumptions/generations, agent \( M \)'s choice is forced to be

\[
U_M(t) = - \sum_{i=0}^{M-1} U_i(t) \text{ for all } t,
\]

(30)

due to the energy balance constraint. Hence one can substitute for \( U_M(t) \) and obtain an equivalent standard, i.e., unconstrained, deterministic LQR problem featuring only \((M-1)\) inputs \( U_{\text{reduced}} := (U_1, U_2, \ldots, U_{M-1}) \), where there is no energy balance constraint:

\[
\min_{\{u_i^*(t)\}_{t=0}^{T}} \sum_{t=0}^{T} [X^T(t)QX(t) + U^T_{\text{reduced}}(t)R_{\text{reduced}}U_{\text{reduced}}(t)],
\]

(31)

subject to \( X(t+1) = AX(t) + B_{\text{reduced}}U_{\text{reduced}}(t) \), (32) the deterministic reduced unconstrained LQR problem.

For this problem (31,32), which is just a standard unconstrained LQR Problem, the optimal solution is given by linear feedback \( U_{\text{reduced}}(0) = \Gamma_{\text{reduced}}(0)X(0) \), where \( \Gamma_{\text{reduced}}(\cdot) \) is the optimal feedback gain.

Noting that \( U_M \) is linear in \( U_{\text{reduced}} \), we deduce that for the full system (27,28) with all \( M \) agents, the optimal solution for the deterministic constrained LQR problem with the energy balance constraint, is \( U(0) = \Gamma(0)x(0) \), where \( \Gamma(\cdot) \) is the optimal feedback gain obtained from \( \Gamma_{\text{reduced}} \) through (30).

Now consider the corresponding reduced unconstrained stochastic LQR problem where there is white Gaussian noise in the state equations (28):

\[
\min_{\{u_i^*(t)\}_{t=0}^{T}} \sum_{t=0}^{T} [x^T(t)Qx(t) + u^T_{\text{reduced}}(t)R_{\text{reduced}}u_{\text{reduced}}(t)],
\]

(33)

with \( x(t+1) = Ax(t) + B_{\text{reduced}}u_{\text{reduced}}(t) + w(t) \).

(34)

By Certainty Equivalence [40], the same linear feedback gain as in the deterministic reduced LQR problem is also optimal. In particular, \( u(0) = \Gamma(0)x(0) \) continues to be
optimal at time $t = 0$. Thus $u(0)$ given by (29) is optimal for (33,34).

However, reduced unconstrained stochastic LQG problem (33,34) is equivalent to unreduced constrained LQG problem (19,20,21), and so the same $u(0)$ is optimal.

Thus the Bid-Price iteration scheme determines the optimal actions for the agents at time 0. Our scheme (22,23,24) for the LQG problem repeats such a Bid-Price scheme iteration at each time $s = 0,1,\ldots,T-1$. Each $x(s)$ can be regarded as an initial state for a subsequent system re-started at time $s$, and the above argument shows that the actions $u(s)$ that it results in for the agents at all times $s$ are also optimal, completing the proof. □

Concerning the convergence rate of the bid-price iterates, it may be noted that in this primal strictly positive definite quadratic problem with linear equality constraints (27,28), the dual is a strictly negative definite quadratic in prices. Hence standard results on convergence of the gradient method for quadratics apply for the price iterations [42], and the convergence rate for the bids follows since they are simply affine functions of the prices. For example, one may even use a small enough constant step size $\alpha^k \equiv \alpha > 0$, in which case the iterates converge linearly, i.e., geometrically.

The result of this section extends to LQG systems where each agent $i$ only has noisy observations $y_i(t) = A_i x_i(t) + n_i(t)$, where $n_i$ are independent and Gaussian.

We note that the LQG modeling constitutes a rich framework that has been successfully used in control with soft quadratic costs used to capture preferences and constraints. This is illustrated in the example treated in Section VIII.

Remark 9: Team problems have been extensively studied, e.g., [43]–[45], but it should be noted that those formulations do not apply here since agents don’t know the system dynamics or models of other agents. Moreover, typically, almost sure constraints such as balancing are not treated. Even setting these issues aside, and supposing that the models are known to all and assuming there are no almost sure constraints, the ISO Problem with private measurements would still lie at the core of decentralized stochastic control with a non-classical information structure [29], [44]–[49], since agents’ observations are influenced by the unknown actions of other agents, as happens here because the price announcements that constitute part of an agent’s observations are dependent on unknown actions taken by other agents. This poses serious difficulties potentially leading to intractability even for LQG systems, as shown by Witsenhausen’s counterexample of a two-stage problem [46]. So there are two questions that one may ponder: (i) Given the prices, how should the agents act? (ii) What are the right prices and how are they determined? Concerning the first question, the linear in state (for given prices) control law may be explained by the fact that given the prices, the agents are conditionally independent of each other. Yuksel [50] has examined such conditional independence conditions on agents’ observations, actions and states, and addressed when tractable solutions can be found. Concerning the second question, one can view the price announcements as a form of “signaling.” In decentralized stochastic control [45], [49], [51], controllers can generally signal some private information to other agents over a “channel” which may even be the physical plant itself. The roles of observation, signaling [44], and the trade-off between communication and control are evident from Witsenhausen’s counterexample [46].

VII. INCORPORATING ADDITIONAL LINEAR CONSTRAINTS: THE DC OPTIMAL POWER FLOW EQUATIONS

As noted in Remark 3, the bid-price iterations can be extended to encompass any additional linear constraints, such as those arising from the DC Power Flow Equations. The only difference is that there are several prices, one for each constraint, that each agent needs to incorporate in choosing its actions.

Theorem 4: Consider a system consisting of $M$ agents, where each agent $i$’s system is a Linear Gaussian System:

$$x_i(t+1) = A_i x_i(t) + B_i u_i(t) + w_i(t) \text{ a.s.}$$

Agent $i$ has a quadratic cost (negative utility):

$$\min \mathbb{E} \left( \sum_{t=0}^{T-1} [x_i^T(t)Q_i x_i(t) + u_i^T(t)R_i u_i(t)] \right).$$

There are $N$ linear constraints that need to be satisfied:

$$\sum_{i=1}^{M} \gamma_{i,n} u_i(t) = 0 \text{ for } 1 \leq n \leq N, t = 0,1,\ldots,T-1, \text{ a.s.}$$

Neither ISO nor agents know the number $M$ or the dynamics/costs/states/noises of other agents.

Consider the following Bid-Multiple Price Iteration. At each time $s = 0,1,\ldots,T-1$, at each iterate $k$, in response to prices $\{\lambda_{n,s}^k(t) : s \leq t \leq T-1\}$, announced by the ISO, agent $i$ solves the deterministic LQR problem:

$$\min \sum_{t=s}^{T-1} [x_i^T(t)Q_i x_i(t) + u_i^T(t)R_i u_i(t) + \sum_{n=1}^{N} \lambda_{n,s}^k(t) u_i(t)],$$

with $x_i(t) = x_i(s)$, determines the optimal $\{u_{n,s}^k(t) : s \leq t \leq T-1\}$, and communicates this sequence to the ISO. Upon receiving the bids at iterate $k$ from all the agents at time $s$, the ISO updates the $N$ price sequences:

$$\lambda_{n,s}^{k+1}(t) = \lambda_{n,s}^k(t) + \alpha^k \left( \sum_{i=1}^{M} \gamma_{i,n} u_{i,s}^k(t) \right),$$

for $1 \leq n \leq N$ and $s \leq t \leq T-1$, with the step-sizes satisfying (26). The multiple iterations converge, and let $\{\lambda_{n,s}^* : s \leq t \leq T-1\}$ denote the limit. Correspondingly let $\{u_{n,s}^* : s \leq t \leq T-1\}$ denote the limits of the bids by the agents. At each time $s$, agent $i$ applies $u_i(s) = u_{i,s}^*$. Then this Bid-Multiple Price Iteration yields the maximum social welfare that would have been attainable under full state information, under the multiple constraints.

Proof: The proof parallels the single constraint case. □
VIII. SIMULATION EXAMPLES

In the following, we use the space conditioning example from [52] for thermal inertial load agents. Let $S_1, S_2, S_3$ be sets of conditioning facilities (loads), conventional generators, and renewable suppliers, respectively, and let $i \in S_1, j \in S_2, k \in S_3$. The dynamics of the temperature $x_i(t)$ of the $i$-th facility is given by (35), where $x_i^O(t)$ is the outside temperature at time $t$, $\epsilon = e^{-\tau/TC}$ = “factor of inertia”, $TC = 2.5$ hours = time-constant of the system, $\tau$ = time duration between control epochs, which is the same as the inter-bid duration, $\eta = 2.5 = \text{thermal conversion efficiency}$, and $A = 0.14kW/{}^\circ F$ = overall thermal conductivity. With $x_i^O(t)$ the desired facility temperature, the cost incurred is a quadratic in the temperature deviation. For fossil-fuel generators, the unit-time conventional generation cost curves [53] for supplying energy are quadratic in generation $u_j$. We replace hard constraints on ramp-rates $|u_j(t) - u_j(t-1)|$ by a quadratic penalty, with $J_3$ below chosen so that the hard bounds are met, with state given by (36). For a renewable energy facility $k$, $B_k$ denotes its buffer capacity, $w_k(t)$ stochastic wind/solar energy, and $x_i(t)$ the renewable energy level satisfying (37). Its operating cost is constant. The resulting ISO Problem (1) is

$$\min \mathbb{E}\left\{ \sum_{i \in S_1} \sum_{t=0}^{T-1} (x_i(t) - x_i^O(t))^2 + \sum_{j \in S_2} J_{j,1}u_j(t) + J_{j,2}u_j^2(t) + J_{j,3}u_j(t) - x_j(t))^2 \right\}$$

such that $\sum_{M} u_j(t) = 0$, for $t = 1, 2, \ldots, T-1$,

$$x_i(t+1) = \epsilon x_i(t) + (1-\epsilon) \left( x_i^O(t) + \frac{\eta}{A} u_i(t) \right), \quad \text{(35)}$$

$$x_j(t+1) = u_j(t), \quad \text{(36)}$$

$$x_k(t+1) = \min\{ x_k(t) - u_k(t) + w_k(t), B_k \}. \quad \text{(37)}$$

We will compare the performance of the proposed Stochastic Dynamic Optimal Bid-Price Iteration scheme of Sections V or VI, called "Optimal" below, with the currently followed Static Dispatch scheme of Section IV used in dynamic situations as explained in Section I, under which the agents perform separate and uncoupled bid-price iterations at each time $t$ to optimize the static cost $c(x(t), u(t))$ incurred at that time $t$.

**Bidding with LQG Systems:** A day is divided into twelve $\tau = 2$ hour bid-slots, so $\epsilon = 0.4493$. There are only thermal loads, and wind-farms which have a cost function $\frac{1}{2}x^2(t)$ and with infinite storage capacity $B$. Outside temperatures and available wind power are modeled as i.i.d. and normal. (This is only a first step towards modeling the uncertainty, and other types of distributions can potentially be similarly explored). Variance of wind energy is 1 unit for all $t$. The scenario is described in Table I. At the beginning of day, the thermal loads have temperature of $70^\circ F$, while wind-farms have 100 units of energy. The price vector is projected at each update onto a large compact set, and, at termination, the bid vector is projected onto the hyperplane $\sum u_i = 0$.

Figs. 3-5 compare performance of the two schemes as the number of bid-price updates, the number of agents connected to the grid, and variance of wind energy process, are varied. Figs. 6 and 7 show how the Optimal scheme is able to attain better social welfare for scale 2.

<table>
<thead>
<tr>
<th>Outside Temp.</th>
<th>Desired Temp.</th>
<th>Wind Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>55 65 70 75 80 80 80 80 70 70 70 0</td>
<td>80 80 80 80 60 80 80 80 80 80 80 80</td>
<td>30 60 90 120 150 180 210 240 270 300 330 360</td>
</tr>
</tbody>
</table>

**TABLE I:** Mean outside and desired thermal load temperatures (in $^\circ F$), and mean wind power for the 12 periods.

---

Fig. 3: Cost, i.e., negative social welfare, vs. number of Bid-Price iterations with five thermal loads and two windfarms.

Fig. 4: Cost as number of users is scaled linearly by $i$, with ratio of thermal loads to windfarms held constant at $5/2$, and with $= 15 + 5i$ bid-price iterations at each time $t$.

Fig. 5: Cost vs. wind variance with five thermal loads and two windfarms, with 30 Bid-Price iterations.

**Bidding in Tree Scenario:** The time-horizon is 2 and time duration between two bids is 5 hours, roughly coinciding with morning (7 am)/12 noon, giving $\epsilon = 0.1353$. Table II
So that savings achieved by Optimal Scheme is three entries, while they are scalar at time \( t = 1 \) while thermal loads disutility are power generations for Scenario 1, are shown in Table III. 

For fossil plants, \( J_1 = 0.1, J_2 = 0.01, J_3 = 0.1 \). Windfarms incur no operational cost. Bid/price vectors at \( t = 0 \) have three entries, while they are scalar at time \( t = 1 \). 

TABLE II: The only stochasticity is wind availability at time 1, with possible realizations \( w_1, w_2 \) with respective probabilities \( P_1, P_2 \). \(|S_1|/|S_2|/|S_3| \) are the relative numbers of thermal loads, fossil plants and windmills. 

Figures 8-10 compare the costs averaged over multiple wind realizations of the two policies under various scenarios, for the two schemes. Thermal loads are allowed to become energy producers, while wind-farm operators are allowed to store energy in case there is excess energy supply in the market, showcasing potential prosumer behavior in energy markets. The particular prices and power generations for Scenario 1, are shown in Table III. 

TABLE III: Prices, power generation and cost savings.

<table>
<thead>
<tr>
<th>Case</th>
<th>( T^1(1), T^2(1) )</th>
<th>( T^1(2), T^2(2) )</th>
<th>( W_1, W_2 )</th>
<th>( P_1, P_2 )</th>
<th>( S_1/S_2/S_3 )</th>
<th>( B )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>30.40 60.80</td>
<td>5.0 0.5 0.5</td>
<td>7/1/1</td>
<td>30</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Case 2</td>
<td>40.80 60.90</td>
<td>10.0 0.95 0.05</td>
<td>4/1/1</td>
<td>40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

IX. Concluding Remarks

The ISO problem gives rise to a problem in general equilibrium theory that is complicated by the facts that agents have private uncertainties but one wants to attain not the price equilibrium that would hold naturally under the corresponding private information structures, but the social welfare optimal solution that could be attained were all agents to pool all their observations, with the further restriction that this is to be accomplished through the medium of an ISO that can only interact through price announcements and consumption/generation bids with agents. This problem of maximizing the social welfare of a collection of distributed dynamic stochastic agents is more complex than decentralized stochastic control since agents do not know the dynamical equations or utility functions of others, and also by the almost sure constraints.

We have exhibited iterative bidding schemes that attain the performance that could have been attained by a centralized control policy that is aware of the dynamics, utilities, uncertainties and states of all agents, under appropriate compactness-convexity or LQG assumptions. The ISO critically exploits the information obtained during the iterative bid-price process to determine the optimal prices and generation/consumption allocations. In the LQG case, the bid-price iteration is particularly simple and tractable. It yields the optimal stochastic dynamic locational marginal prices.

The social-welfare optimality can potentially result in significant economic benefits in energy markets with deep renewable penetration.

The results may be of interest vis-a-vis general equilib-
rium theory for its treatment of systems with stochastic dynamic agents.

The agents are all presumed to be “price takers.” Examining this in a broader context is an important issue and is the subject of a future work.

ACKNOWLEDGMENT

The authors thank Pravin Varaiya for identifying a significant error in an earlier version of the paper.

REFERENCES


Rahul Singh received the B.E. degree in electrical engineering from Indian Institute of Technology, Kanpur, India, in 2009, the M.Sc. degree in Electrical Engineering from University of Notre Dame, South Bend, IN, in 2011, and the Ph.D. degree in electrical and computer engineering from the Department of Electrical and Computer Engineering Texas A&M University, College Station, Tx, in 2015.

He is currently a Postdoctoral Associate at the Laboratory for Information Decision Systems (LIDS), Massachusetts Institute of Technology. His research interests include decentralized control of large-scale complex cyberphysical systems, operation of electricity markets with renewable energy, and scheduling of networks serving real time traffic.

P. R. Kumar B. Tech. (IIT Madras, ‘73), D.Sc. (Washington University, St. Louis, ’77), was a faculty member at UMBC (1977-84) and Univ. of Illinois, Urbana-Champaign (1985-2011). He is currently at Texas A&M University. His current research is focused on stochastic systems, energy systems, wireless networks, security, automated transportation, and cyberphysical systems.

He is a member of the US National Academy of Engineering, The World Academy of Sciences, and a Foreign Fellow of the Indian National Academy of Engineering. He was awarded a Doctor Honoris Causa by ETH, Zurich. He has received the IEEE Field Award for Control Systems, the Donald P. Eckman Award of the AACC, Fred W. Ellersick Prize of the IEEE Communications Society, the Outstanding Contribution Award of ACM SIGMOBILE, the Infocom Achievement Award, and the SIGMOBILE Test-of-Time Paper Award. He is a Fellow of IEEE and ACM Fellow. He was Leader of the Guest Chair Professor Group on Wireless Communication and Networking at Tsinghua University, is a D. J. Gandhi Distinguished Visiting Professor at IIT Bombay, and an Honorary Professor at IIT Hyderabad. He was awarded the Distinguished Alumnus Award from IIT Madras, the Alumni Achievement Award from Washington Univ, and the Daniel Drucker Eminent Faculty Award from the College of Engineering at the Univ. of Illinois.

Le Xie (S’05-M’10-SM’16) received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2004, the M.Sc. degree in engineering sciences from Harvard University, Cambridge, MA, in 2005, and the Ph.D. degree in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, in 2009.

He is currently an Associate Professor with the Department of Electrical and Computer Engineering, Texas A&M University, College Station.

His research interest includes modeling and control of large-scale complex systems, smart grid application with renewable energy resources, and electricity markets.