A SURVEY IN STOCHASTIC AND DECENTRALIZED CONTROL

Submitted by :

NAMAN AGGARWAL

Roll Number 160010058

Supervisors:

Prof. Ankur Kulkarni

Department of Systems and Control Engineering

Prof. Avijit Chatterjee

Department of Aerospace Engineering

INDIAN INSTITUTE OF TECHNOLOGY, BOMBAY

November 2019

CERTIFICATE

This seminar report entitled **A SURVEY IN STOCHASTIC AND DECENTRALIZED CONTROL** by NAMAN AGGARWAL is approved by me for submission for fulfilment of requirements of BTP Phase-I. The report was further certified that to the best of my knowledge, represents the work carried out by the student.

Date :

Signature and Name of Guide

ACKNOWLEDGEMENT

I would like to express my deep sense of gratitude to **Prof. Ankur Kulkarni**, for his invaluable help and guidance during the course of this project. I am highly indebted to him for constantly encouraging me and giving me new directions and ideas to work on.

NAMAN AGGARWAL

November 2019

Indian Institute of Technology, Bombay

Abstract

This report is a survey in stochastic and decentralized control. We consider control and estimation problems where the sensor signals and the actuator signals are transmitted to various subsystems over a network. In contrast to traditional control and estimation problems, here the observation and control packets may be lost or delayed. The unreliability of the underlying communication network is modelled stochastically by assigning probabilities to the successful transmission of packets. This requires a novel theory which generalizes classical control/estimation paradigms. We conduct a literature survey to understand the foundations and applications of such a novel theory.

Contents

Chapter 1

Organization of the Report

This is a survey report divided into the discussion of three primary references [1], [3], and [5] from chapters 2 to 4. The theme of this survey report is estimation and control over lossy networks, decentralized control and related areas.

Chapter 2

Control and Estimation over Lossy Networks

Wireless networks are inherently less reliable and secure than their wired counterparts. These two factors limit the penetration of wireless technology in many application contexts. Issues of communication delay, data loss and time synchronization play critical roles. In particular, communication and control are tightly coupled and cannot be addressed independently. Specific questions that arise are the following. What is the amount of data loss that the control loop can tolerate while reliably performing it's task? Can communication protocols be designed to satisfy this constraint? As stated in [1], the goal of this study is to provide some first steps in answering such questions by examining the basic system-theoretic implications of using unreliable networks for control.

In our study, we consider the following protocols for packet networks communication: TCP-like, and UDP-like. We want to study the effect of data losses due to the unreliability of the network links under these two general protocol abstractions. Accordingly, we model the arrival of both observations and control packets as random processes whose parameters are related to the characteristics of the communication channel.

2.1 Problem Formulation

Consider the following linear stochastic system with intermittent observation and control packets:

$$
\begin{aligned}
&x_{k+1} = Ax_k + Bu_k^a + w_k \\
&u_k^a = \nu_k u_k^c \\
&y_k = \gamma_k C x_k + v_k
\end{aligned}
$$

where u_k^a is the control input to the actuator, u_k^c is the control input computed by the controller, (x_o, w_k, v_k) are Gaussian, uncorrelated, white with mean $(\overline{x_o}, 0, 0)$ and covariance (P_o, Q, R) respectively, and (ν_k, γ_k) are i.i.d. Bernoulli random variables with $P(\gamma_k = 1) = \overline{\gamma}$ and $P(\nu_k = 1) = \overline{\nu}$. The stochastic variable ν_k models the packet loss between the controller and the actuator: if the packet is correctly delivered then $u_k^a = u_k^c$, otherwise if it is lost then the actuator does nothing, i.e., $u_k^a = 0$. The stochastic variable γ_k models the packet loss between the sensor and the controller: if the packet is delivered then $y_k = Cx_k + v_k$, while if the packet is lost the controller reads pure noise, i.e., $y_k = 0$. We modelled the packet arrival processes are Bernoulli i.i.d. to make our analysis tractable. We can also model it as a Markovian process, which is more realistic, but at the cost of making our analysis quite involved and complicated.

2.2 Communication Protocols: TCP-like and UDP-like

Packet networks communication channels typically use one of two fundamentally different protocols: TCP-like (Transmission Control) or UDP-like (User Datagram). In the first case there is acknowledgement of received packets, while in the second case no-feedback is provided on the communication link.

The controller which is decentralized from the system receives an observation (y_k, γ_k) at time instant k, and sends a control signal computed by the algorithm, u_k^c. The control signal reaches the system with a probability defined by the Bernoulli i.i.d. random variable ν_k, and the actuator exercises control action, $u_k^a (= \nu_k u_k^c)$. For TCP-like communication protocol, the receiver, which in this case is the plant/system, sends an acknowledgment back to the sender communicating if it has received the signal or it. For UDP-like protocols, there is no acknowledgement to the sender from the receiver's end. In this light, let us define the following information sets (at the controller):

$I_k = F_k := \{y^k, \gamma^k, \nu^{k-1}\}$, TCP-like
$I_k = G_k := \{y^k, \gamma^k\}$, UDP-like
where
$y^k = \{y_1, y_2, ..., y_k\}$
$\gamma^k = \{\gamma_1, \gamma_2, ..., \gamma_k\}$
$\nu^k = \{\nu_1, \nu_2, ..., \nu_{k-1}\}$

The set F corresponds to the information provided under an acknowledgement-based communication protocols (TCP-like) in which successful or unsuccessful packet delivery at the receiver is acknowledged to the sender within the same sampling time period. The set G corresponds to the information available at the controller under communication protocols in which the sender receives no feedback about the delivery of the transmitted packet to the receiver (UDP-like).

2.3 Optimal Control

Consider also the following cost function:

$J_N(u^{N-1}, \overline{x}_o, P_o) = \mathbb{E}[x_N^{'} W_N x_N + \sum_{k=0}^{N-1}(x_k^{'} W_k x_k + \nu_k u_k^{'} U_k u_k) | u^{N-1}, \overline{x}_o, P_o]$
where $u^{N-1} = \{u_1, u_2, ..., u_{N-1}\}$

Since the controller is constrained to act according to the information it has, we seek the optimal control sequence, u^{*N-1} as a function of the available information such that $u_k = g_k(I_k)$. We have,
$J_N^*(\overline{x}_o, P_o) = \min_{u_k = g_k(I_k)} J_N(u^{N-1}, \overline{x}_o, P_o)$

2.3.1 Optimal Estimation

We start by defining the following variables,
$\hat{x}_{k|k} := \mathbb{E}[x_k | I_k]$
$\hat{e}_{k|k} = x_k - \hat{x}_{k|k}$
$\hat{P}_{k|k} := \mathbb{E}[e_{k|k} e_{k|k}^{'} | I_k]$

We can use the resursive formulae from Baye's Law (derivation is standard, taught in class) as mentioned below to obtain $\pi_{k+1} = T(\pi_k, I_{k+1})$, where π_k is a probability distribution over the state-space defined as, $\pi_k(x_k) = P(X_k = x_k | I_k)$,

$\pi_{k+1}(x_{k+1}) = P(X_{k+1} = x_{k+1} | I_{k+1})$

2.3.1.1 Estimator Design under TCP-like Protocols

Equations for the given system can be derived in the same way as done in case of a standard Kalman Filter using the Bayesian recursive formula. We have,

$\hat{x}_{k+1|k} = \mathbb{E}[x_{k+1} | F_k] = A\,\mathbb{E}[x_k | F_k] + \nu_k B u_k$
$e_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k} = A e_{k|k} + w_k$
$P_{k+1|k} = \mathbb{E}[e_{k+1|k} e_{k+1|k}^{'} | \nu_k, F_k] = A P_{k|k} A^{'} + Q$

where the independence of w_k and F_k , and the requirement that u_k is a deterministic function of F_k , are used. Since y_{k+1} , γ_{k+1} , w_k and F_k are independent, the correction step is given by

$\hat{x}_{k+1|k+1} = \hat{x}_{k|k+1} + \gamma_{k+1} K_{k+1}(y_{k+1} - C\hat{x}_{k+1|k})$
$e_{k+1|k+1} = x_{k+1} - \hat{x}_{k+1|k+1} = (I - \gamma_{k+1} K_{k+1} C) e_{k+1|k} - \gamma_{k+1} K_{k+1} v_{k+1}$
$P_{k+1|k+1} = P_{k+1|k} - \gamma_{k+1} K_{k+1} C P_{k+1|k}$ where, $K_{k+1} = P_{k+1|k} C^{'} (C P_{k+1|k} C^{'} + R)^{-1}$

2.3.1.2 Estimator Design under UDP-like Protocols

We derive the equations for the optimal estimator using similar arguments as in the standard Kalman filtering equations. The innovation step is given by,

$$\hat{x}_{k+1|k} = \mathbb{E}[x_{k+1}|G_k] = \mathbb{E}[Ax_k + \nu_k Bu_k + w_k|G_k] = A\hat{x}_{k|k} + \overline{\nu} Bu_k$$
$$e_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k} = Ae_{k|k} + (\nu_k - \overline{\nu})Bu_k + w_k$$
$$P_{k+1|k} = \mathbb{E}[e_{k+1|k}e^{'}_{k+1|k}|G_k] = AP_{k|k}A^{'} + \overline{\nu}(1-\overline{\nu})Bu_k u^{'}_k B^{'} + Q$$

where we used the independence and zero-mean of w_k , $(\nu_k - \overline{\nu})$, and G_k, and the fact that u_k is a deterministic function of the information set G_k. Note how under UDP-like communication, differently from TCP-like, the error covariance $P_{k+1|k}$ depends explicitly on the control input u_k . This is the main difference with control feedback systems under TCP-like protocols.

The correction step is the same as for the TCP case,
$\hat{x}_{k+1|k+1} = \hat{x}_{k|k+1} + \gamma_{k+1}K_{k+1}(y_{k+1} - C\hat{x}_{k+1|k})$
$P_{k+1|k+1} = P_{k+1|k} - \gamma_{k+1}K_{k+1}CP_{k+1|k}$ where, $K_{k+1} = P_{k+1|k}C^{'}(CP_{k+1|k}C^{'} + R)^{-1}$

2.3.2 Optimal Control under TCP-like protocols

We follow a dynamic programming approach based on the cost-to-go iterative procedure for the derivation of the optimal feedback control law and the corresponding value for the objective function. We define the optimal value function as follows,

$$V_N(x_N) = \mathbb{E}[x^{'}_N W_N x_N|F_N]$$
$$V_k(x_k) = \min_{u_k} \mathbb{E}[x^{'}_k W_k x_k + \nu_k u^{'}_k U_k u_k + V_{k+1}(x_{k+1})|F_k] \text{ where } k = N-1, ..., 1$$

Using dynamic programming theory, we can show that $J^{*}_N = V_o(x_o)$. The value function $V_k(x_k)$ under TCP-like protocols for the given system can be written as:
$V_k(x_k) = \mathbb{E}[x^{'}_k S_k x_k|F_k] + c_k$, $k = 0, 1, ...N-1$
where the matrix S_k and scalar c_k can be computed recursively as follows:
$S_k = A^{'}S_{k+1}A + W_k - \overline{\nu}A^{'}S_{k+1}B(B^{'}S_{k+1}B + U_k)^{-1}B^{'}S_{k+1}A$
$c_k = \text{trace}((A^{'}S_{k+1}A + W_k - S_k)P_{k|k}) + \text{trace}(S_{k+1}Q) + \mathbb{E}[c_{k+1}|F_k]$
with initial values $S_N = W_N$ and $c_N = 0$

Proof
We make use of the following lemma in the derivation to follow.

Lemma 2.3.1 The following facts are true,

(a) $\mathbb{E}[(x_k - \hat{x}_k)\hat{x}_k^{'}] = \mathbb{E}[e_{k|k}\hat{x}_k^{'}|I_k] = 0;$
(b) $\mathbb{E}[x_k S x_k^{'}|I_k] = \hat{x}_k S \hat{x}_k^{'} + \mathrm{trace}(SP_{k|k}), \forall S \geq 0;$
(c) $\mathbb{E}[\mathbb{E}[g(x_{k+1})|I_{k+1}]|I_k] = \mathbb{E}[g(x_{k+1})|I_k], \forall g(.)$

Using Lemma 2.3.1, we can prove the following two results which we'll use in the main derivation,
$\mathbb{E}\left[x_{k+1}' S x_{k+1}|\ \mathcal{I}_k\right] = \mathbb{E}\left[x_k' A' S A x_k|\mathcal{I}_k\right] + \bar{\nu} u_k' B' S B u_k + 2\bar{\nu} u_k' B' S A \hat{x}_{k|k} + \mathrm{trace}(SQ)$
where both the independence of ν_k, w_k and x_k, and the zero-mean property of w_k are exploited. The previous expectation holds true for both the information sets, i.e., $I_k = F_k$ or $I_k = G_k$. Also,
$\mathbb{E}\left[e_{k|k}' T e_{k|k}|\mathcal{I}_k\right] = \mathrm{trace}\left(T\mathbb{E}\left[e_{k|k} e_{k|k}'|\mathcal{I}_k\right]\right) = \mathrm{trace}\left(TP_{k|k}\right), \quad \forall T \geq 0$

The proof employs an induction argument. The claim is clearly true for $k = N$ with the choice of parameters, $S_N = W_N$ and $c_N = 0$. We suppose that the claim is true for $k+1$ i.e. $V_{k+1}\left(x_{k+1}\right) = \mathbb{E}\left[x_{k+1}' S_{k+1} x_{k+1}|\mathcal{F}_{k+1}\right] + c_{k+1}$. The value function at time step k is the following:

$V_k\left(x_k\right) = \min_{u_k} \mathbb{E}\left[x_k' W_k x_k + \nu_k k_k' U_k u_k + V_{k+1}\left(x_{k+1}\right)|\mathcal{F}_k\right]$
$V_k\left(x_k\right) = \min_{u_k} \mathbb{E}\left[x_k' W_k x_k + \nu_k u_k' U_k u_k + |\mathcal{F}_k\right] + \mathbb{E}\left[\mathbb{E}\left[x_{k+1}' S_{k+1} x_{k+1} + c_{k+1}|\mathcal{F}_{k+1}\right]|\mathcal{F}_k\right]$
$V_k\left(x_k\right) = \min_{u_k} \mathbb{E}\left[x_k' W_k x_k + \nu_k u_k' U_k u_k + x_{k+1}' S_{k+1} x_{k+1} + c_{k+1}|\mathcal{F}_k\right]$
And hence,
$V_k\left(x_k\right) = \mathbb{E}[x_k' W_k x_k + x_k' A' S_{k+1} A x_k|\mathcal{F}_k] + \mathrm{trace}\left(S_{k+1} Q\right) + \mathbb{E}\left[c_{k+1}|\mathcal{F}_k\right]$
$+\, \bar{\nu} \min_{u_k}(u_k'\left(U_k + B' S_{k+1} B\right) u_k + 2u_k' B' S_{k+1} A \hat{x}_{k|k})$

where we used Lemma 2.3.1(c) to get the third equality. The value function is a quadratic function of the input, therefore, the minimizer can be simply obtained by solving $\frac{\partial V_k}{\partial u_k} = 0$. The optimal feedback is, thus, a simple linear function of the estimated state,
$u_k^* = -\left(B' S_{k+1} B + U_k\right)^{-1} B' S_{k+1} A \hat{x}_{k|k} = L_k \hat{x}_{k|k}$
If we substitute the optimal control action u_k^*, to get $V_k(x_k)$, and use Lemma 2.3.1(b) we get,
$V_k(x_k) = \mathbb{E}\left[x_k' W_k x_k + x_k' A' S_{k+1} A x_k - \bar{\nu} x_k' A' S_{k+1} B(U_k + B' S_{k+1} B)^{-1} B' S_{k+1} A x_k|\mathcal{I}_k\right] + \mathrm{trace}(S_{k+1} Q) + \mathbb{E}\left[c_{k+1}|\mathcal{I}_k\right] + \bar{\nu}\mathrm{trace}(A' S_{k+1} B\left(U_k + B' S_{k+1} B\right)^{-1} B' S_{k+1} P_{k|k})$

From the above expression, we see that our claim is valid, if the recursion on S_k and c_k is as given earlier. Since, $J_N^*\left(\bar{x}_0, P_0\right) = V_0\left(x_0\right)$, from the observation that $V_k(x_k) = \mathbb{E}[x_k^{'} S_k x_k|F_k] + c_k$ $\forall$k $\in \{0, 1, ...N-1\}$, it follows that the cost function for the optimal LQG using TCP-like protocols is given by,
$J_N^* = \bar{x}_0' S_0 \bar{x}_0 + \mathrm{trace}\left(S_0 P_0\right) + \sum_{k=0}^{N-1} \mathrm{trace}\left(S_{k+1} Q\right)$
$+ \sum_{k=0}^{N-1} \mathrm{trace}\left(\left(A' S_{k+1} A + W_k - S_k\right) \mathbb{E}_\gamma\left[P_{k|k}\right]\right)$

where we used the fact that $\mathbb{E}\left[x_0^* S_0 x_0\right] = \bar{x}_0^* S_0 \bar{x}_0 + \operatorname{trace}(S_0 P_0)$, and $\mathbb{E}_\gamma[\cdot]$ explicitly indicates that the expectation is calculated with respect to the arrival sequence $\{\gamma_k\}$.

It is important to remark that the error covariance matrices $\left\{P_{k|k}\right\}_{k=0}^{N}$ are stochastic since they depend on the sequence $\{\gamma_k\}$. Moreover, since the matrix $P_{k+1|k+1}$ is a nonlinear function of the previous time step matrix covariance $P_{k|k}$ (see Section 2.3.1.1), the exact expected value of these matrices, $\mathbb{E}_\gamma\left[P_{k|k}\right]$, cannot be computed analytically, as shown in [2]. However, they can be bounded by computable deterministic quantities, as shown in [2] from which state the following lemma:

Lemma 2.3.2 The expected error covariance matrix $\mathbb{E}_\gamma\left[P_{k|k}\right]$ satisfies the following bounds:
$\widetilde{P}_{k|k} \leq \mathbb{E}_\gamma\left[P_{k|k}\right] \leq \widehat{P}_{k|k} \quad \forall k \geq 0$
where the matrices $\widehat{P}_{k|k}$ and $\widetilde{P}_{k|k}$ can be computed as follows:

$$\widehat{P}_{k+1|k} = A\widehat{P}_{k|k-1}A' + Q - \bar{\gamma}A\widehat{P}_{k|k-1}C'\left(C\widehat{P}_{k|k-1}C' + R\right)^{-1}C\widehat{P}_{k|k-1}A'$$
$$\widehat{P}_{k|k} = \widehat{P}_{k|k-1} - \bar{\gamma}\widehat{P}_{k|k-1}C'\left(C\widehat{P}_{k|k-1}C' + R\right)^{-1}C\widehat{P}_{k|k-1}$$
$$\widetilde{P}_{k+1|k} = (1-\bar{\gamma})A\widetilde{P}_{k|k-1}A' + Q$$
$$\widetilde{P}_{k|k} = (1-\bar{\gamma})\widetilde{P}_{k|k-1}$$
where the initial conditions are $\widehat{P}_{0|0} = \widetilde{P}_{0|0} = P_0$.

Proof

The argument is based on the observation that the matrices $P_{k+1|k}$ and $P_{k|k}$ are concave and monotonic functions of $P_{k|k-1}$. The proof is offered in [2] and is thus omitted.

From this lemma, it follows that also the minimum achievable cost J_N^*, cannot be computed analytically, but can bounded using the bounds on the expected error covariance matrix $\mathbb{E}_\gamma\left[P_{k|k}\right]$ from Lemma 2.3.2.

2.3.2.1 Summary of Results: TCP-like Protocols

(a) The separation principle still holds for TCP-like communication, since the optimal estimator (error covariance of the estimator), is independent of the control input u_k.

(b) The optimal estimator gain K_k is time-varying and stochastic since it depends on the past observation arrival sequence $\{\gamma_j\}_{j=1}^{k}$.

(c) The optimal control input is a linear function of the estimated state $\hat{x}_{k|k}$ i.e. $u_k = L_k\hat{x}_{k|k}$ and the optimal gain L_k is independent of the process sequences $\{\nu_k, \gamma_k\}$.

2.3.3 Optimal Control under UDP-like protocols

In [1], the authors show that the optimal LQG controller, under UDP-like communication protocols, is in general not a linear function of the state estimate. Consequently, estimation and controller design cannot be treated independently. The authors have a constructed a simple counter-example using a scalar system to demonstrate this.

2.4 Concluding Remarks

In this survey (based on [1]), we have analyzed the LQG control problem in the case where both observation and control packets may be lost during transmission over a communication channel. This situation arises frequently in distributed systems where sensors, controllers and actuators reside in different physical locations and have to rely on data networks to exchange information. We studied the LQG control problem under two classes of protocols: TCP-like and UDP-like. In TCP-like protocols,acknowledgements of successful transmissions of control packets are provided to the controller, while in UDP-like protocols, no such feedback is provided.

For TCP-like protocols, we have shown that the optimal control is a linear function of the state and that the separation principle holds. As a consequence, controller design and estimator design are decoupled under TCP-like protocols. However, unlike standard LQG control with no packet loss, the gain of the optimal observer does not converge to a steady state value. Rather, the optimal observer gain is a time-varying stochastic function of the packet arrival process. In analyzing the infinite horizon problem in [1], the authors have shown that the infinite horizon cost is bounded if and only if arrival probabilities $\bar{\gamma}, \bar{\nu}$ exceed a certain threshold. Thus, the underlying communication channel must be sufficiently reliable in order for LQG optimal controllers to stabilize the plant.

UDP-like protocols present a much more complex problem. We saw that the lack of acknowledgement of control packets results in the failure of the separation principle. Estimation and control are now intimately coupled. The LQG optimal control is, in general, nonlinear in the estimated state.

Chapter 3

Optimal Sensor Transmission Energy Allocation for Linear Control Over a Packet Dropping Link with Energy Harvesting

Wireless sensors have become more powerful, affordable and compact, and are thus used in many areas. Sensors are often connected in remote places and cannot be connected to reliable power sources. Thus, sensors are often powered by batteries and can only use a limited amount of energy for sensing, processing and communicating information. Hence, the communication links are unreliable and information might be lost in a random manner. It is therefore an important task to study the eects of such unreliable communication channels on ltering and control over sensor networks with energy constraints. One way to help overcome the limitations of limited battery capacities is to use energy harvesting as sensors are often placed in an environment where energy can be harvested using solar panels, wind mills or other technical devices. The harvested energy can then be used for data transmission or be stored in the battery for future use.

[3] extends [4] to the case of a closed loop control system with a packet dropping link between the sensor and the controller at the receiver, and a feedback channel that can be prone to intermittent losses. We study the optimal energy allocation policy at the transmitter and the optimal control design at the receiver such that a nite-time horizon LQG control cost is minimized.

3.1 System Model and Problem Formulation

3.1.1 Plant Model and Sensor

The plant is modeled as a simple linear system with state $x_k \in R^n$, process noise $w_k \in R^n$ (i.i.d. Gaussian noise with zero mean and covariance matrix $M = \mathbb{E}[w_k w_k^{'}] \geq 0$), and a control input $u_k \in R^p$, that is $x_{k+1} = Ax_k + Bu_k + w_k$. The initial state x_o is also Gaussian with mean $\overline{x}_o$, and covariance $\overline{P}_o$, and A, B are matrices with appropriate dimensions.

The sensor produces a noisy measurement of the state $y_k = Cx_k + v_k$ where $y_k \in R^q$, and $v_k \in R^q$ is assumed to be i.i.d. Gaussian noise (independent of x_o and w_k) with zero

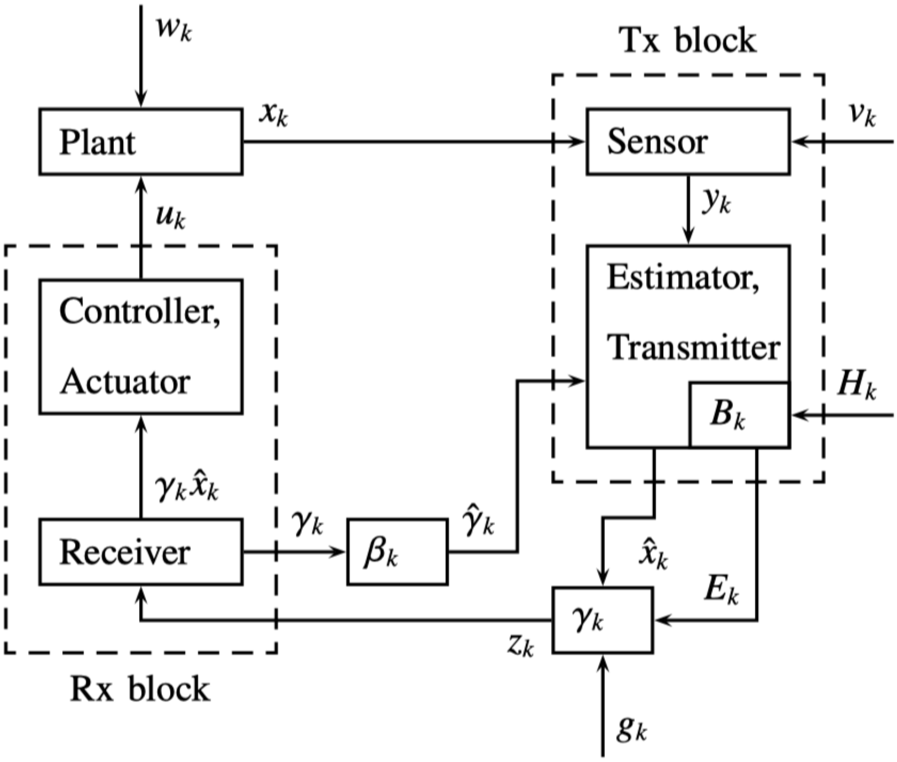

Figure 3.1: Scheme of system model from [3]

mean and covariance matrix $N = \mathbb{E}[v_k v_k^{'}] > 0$.

3.1.2 State Estimator at the Transmitter

Assume the sensor is smart with computational capability, and the sensor transmitter forwards a state estimate to the remote estimator/controller. The sensor measurements are used at the transmitter to estimate the current state based on the information set $I_k = \{\hat{x}_0, y_l, \hat{\gamma}_{l-1} : 1 \leq l \leq k\}$, where $\hat{\gamma}_l$ is the channel feedback acknowledgment, which will be discussed in detail in Section 3.1.5. The estimate is given by

$$\begin{aligned}
\hat{x}_k &:= \hat{x}_{k|k} = \mathbb{E}\left\{x_k | I_k\right\} \\
&= \hat{x}_{k|k-1} + K_k \left(y_k - C\hat{x}_{k|k-1}\right) \\
\hat{x}_{k+1|k} &= \mathbb{E}\left\{x_{k+1} | I_k\right\} = A\hat{x}_{k|k} + B\hat{u}_k
\end{aligned}$$

where $\hat{u}_k = \mathbb{E}[u_k | I_k]$

I_k^c is the information structure at the decentralized controller, and hence $u_k = g_k(I_k^c)$. $\hat{u}_k$ is the conditional expectation of u_k which is a function of I_k^c, with respect to the information structure at the transmitter, I_k.

The matrix K_k should be chosen such that it minimizes the error covariance matrix of the state estimation error, which depend on the feedback communication channel. In case the acknowledgements from the receiver regarding whether the transmitted state estimate has been received or not (ACK/NACK), are always received perfectly, the error covariance matrices follow the standard Riccati dierence equations, discussed further in Section 3.2. However, in case the ACK/NACK feedback is dropped intermittently, the separation principle no longer holds. This leads to a non-standard form of the error covariance matrices, see Section 3.3.

3.1.3 Energy Harvester and Battery Dynamics

The transmitter has a rechargeable battery equipped with an energy harvester, that can gather energy from the environment. The amount of energy available to be harvested at time slot k, denoted by H_k and modeled by an i.i.d. process, is stored in the battery and can be used for data transmission. We assume that the energy used for computational purposes at the transmitter are negligible compared to the amount of energy required for transmission. This is particularly true if data is transmitted over a wireless channel to a receiver that is a long distance away. The amount of stored energy in the battery at time k, B_k, evolves according to

$B_{k+1} = \min\left\{B_k - E_k + H_{k+1}; \quad \bar{B}\right\}$

with $0 \leq B_o \leq \overline{B}$ and where $\overline{B}$ is the battery capacity, and E_k is the energy used for transmission during k-th slot.

3.1.4 Forward Communication Channel

A wireless, packet dropping communication channel is used to transmit the state estimate $\hat{x}_k$ to the receiver such that the estimate is either exactly received ($\gamma_k = 1$) or completely lost due to corrupted data or substantial delay ($\gamma_k = 0$), where γ_k is the Bernoulli random variable modeling the packet loss process. The received signal is $z_k = \gamma_k \hat{x}_k$. The probability of successfully transmitting the packet is

$\mathbb{P}\left(\gamma_k = 1 | g_k, E_k\right) := h\left(g_k E_k\right)$

where g_k is the time-varying wireless fading channel gain and E_k is the transmission energy for transmitting the packet at k. $h : [0, \infty] \rightarrow [0, 1]$ is monotonically increasing and continuous.

We assume that the channel gain g_k is described by an i.i.d. process, independent of the energy harvesting process H_k, and known to the transmitter. Based on the channel gain g_k, and the current battery level B_k, the transmitter nds an optimal energy allocation policy $\{E_k\}$ in order to minimize a suitable nite horizon control cost. The details of this optimal energy allocation scheme will be provided later in this section.

3.1.5 Erroneous Feedback Communication Channel

After receiving z_k , the receiver sends an acknowledgment to the transmitter via a packet dropping feedback channel such that the received acknowledgement is of the form

$$\hat{\gamma}_k = \begin{cases} \gamma_k & \text{if } \beta_k = 1 \\ 2 & \text{if } \beta_k = 0 \end{cases}$$

where the β_k is a Bernoulli random variable indicating if the ACK/NACK packet has been received with $P(\beta_k = 0) = \eta \in [0, 1]$. In case no ACK/NACK is received, the transmitter receives the feedback signal $\hat{\gamma}_k = 2$ indicating the packet drop.

3.1.6 Estimator/Controller and Actuator in the Receiver block

The controller in the receiver block has access to the information set $T_k^{\mathrm{c}} := \{\hat{x}_0^{\mathrm{c}}, z_l, \gamma_l : 1 \leq l \leq k\}$. Since the estimates from the transmitter are dropped intermittently, the state estimate at the Rx block, $\hat{x}_k^{\mathrm{c}} = E\left[x_k | T_k^{\mathrm{c}}\right]$, is

$$\hat{x}_k^{\mathrm{c}} = \gamma_k \hat{x}_k + (1 - \gamma_k)\left(A\hat{x}_{k-1}^{\mathrm{c}} + Bu_{k-1}\right)$$

The task of the controller is to design an optimal control sequence $\{u_k\}$ based on the information pattern I_k^c such that k a suitable nite horizon control cost is minimized. It is assumed that the link between the Rx block and the plant is lossless, such that the correct control signal u_k is applied to the plant.

3.1.7 Optimal Transmission Energy and Control Policy Design

The aim is to nd the optimal transmission energy allocation policy $\mathbf{E}^{N-1*}$ and the optimal control policy $\mathbf{u}^{N-1*}$, that jointly minimize the nite horizon LQG control cost

$$J\left(\mathbf{u}^{\mathrm{N}-1}, \mathbf{E}^{\mathbf{N}-1}, \bar{x}_0, \bar{P}_0\right) = \mathbb{E}\left\{x_N^{\mathrm{T}} Q x_N + \sum_{k=0}^{N-1}\left(x_k^{\mathrm{T}} Q x_k + u_k^{\mathrm{T}} S u_k\right)\right\}$$

where $\mathbf{u}^{N1} = \{u_0, u_1, ..., u_{N1}\}$, and $\mathbf{E}^{N1} = \{E_0, E_1, ..., E_{N1}\}$, and the dependence of the cost on the mean and the variance of the initial state is explicitly shown. We will show in Section 3.2, that if the feedback channel is perfect, the separation principle holds and the design of the optimal control input u_k and the optimal transmission energy E_k can be separated. However, in the case of imperfect channel feedback, the optimal choices of u_k and E_k depend on each other, as shown in Section 3.3.

Note: We focus on the scenario where the transmitter designs the optimal energy allocation policy and the receiver designs the optimal control policy. It will be seen below that in the case of perfect channel feedback, the transmitter can design the optimal energy allocation policy due to the separation principle. In the case of imperfect channel feedback, the energy allocation policy design at the transmitter becomes strictly sub-optimal.

3.2 Perfect Feedback

In case the channel feedback link is perfect, the control of the closed loop system follows well known principles. After clarifying the dynamics of the error covariance matrices at the estimator and the controller, it will be shown that the separation principle holds. Since we only consider a nite horizon problem, the quadratic control cost is always bounded and a stability proof is not needed.

3.2.1 Error Covariance Matrices at the Transmitter and Receiver

The estimator at the Tx block calculates an estimate of the system state via a Kalman lter based on the information set $I_k := \{\hat{x}_0, y_l, \gamma_{l-1} : 1 \leq l \leq k\}$. Since it is assumed that the acknowledgements are received without faults, the estimator has perfect knowledge of the state estimate at the controller and hence the applied control input, $\hat{u}_k = u_k$. The error covariance matrices at the transmitter are

$$P_{k|k} = \mathbb{E}\left\{\left(x_k - \hat{x}_{k|k}\right)\left(x_k - \hat{x}_{k|k}\right)^{\mathrm{T}} |\boldsymbol{I}_k\right\}$$
$$P_{k+1} := P_{k+1|k} = \mathbb{E}\left\{\left(x_{k+1} - \hat{x}_{k+1|k}\right)\left(x_{k+1} - \hat{x}_{k+1|k}\right)^{\mathrm{T}} |\boldsymbol{I}_k\right\}$$

With $e_{k|k} = x_k - \hat{x}_{k|k}$ and $e_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k} = Ae_{k|k} + w_k$, this yields

$$P_{k+1} = \mathbb{E}\left\{\left(Ae_{k|k} + w_k\right)\left(Ae_{k|k} + w_k\right)^{\mathrm{T}}\right\} = AP_{k|k}A^{\mathrm{T}} + M$$

Further, choosing $K_k = P_{k|k-1}C^{\mathrm{T}}\left(CP_{k|k-1}C^{\mathrm{T}} + N\right)^{-1}$ leads to the minimal error covariance matrix after updating $P_{k|k}$ in the standard form

$$P_{k|k} = P_{k-1} - P_{k-1}C^{\mathrm{T}}\left(CP_{k-1}C^{\mathrm{T}} + N\right)^{-1}CP_{k-1}$$

The initial covariance matrix is given by $P_o = \overline{P}_o$. Since the current state estimate is intermittently unavailable at the Rx block, it is replaced by $\hat{x}_k^c$. The corresponding estimation error covariance matrix $P_k^{\mathrm{c}} := \mathbb{E}\left\{(x_k - \hat{x}_k^{\mathrm{c}})(x_k - \hat{x}_k^{\mathrm{c}})^{\mathrm{T}} |T_k^{\mathrm{c}}\right\}$ is

$$P_k^{\mathrm{c}} = \gamma_k P_k + (1 - \gamma_k)\left(AP_{k-1}^{\mathrm{c}}A^{\mathrm{T}} + M\right)$$

For simplicity, we assume $P_o^c = \overline{P}_o$.

3.2.2 Separation Principle

In case the acknowledgements are received without error, the control input u_k is perfectly known at the transmitter. Hence, the control signal estimate $\hat{u}_k$ in the system equations is replaced by u_k . Thus, the estimation error is independent of the control input as can be veried from Section 3.2.1 (see also [1]). Clearly, the separation principle holds in this case. This implies, that the tasks of obtaining the optimal Kalman ltered state estimate $\hat{x}_k$, $\hat{x}_k^c$, calculating the optimal control input u_k^* at the controller, and computing the optimal energy allocation E_k^* at the transmitter can be carried out separately. Consequently, implementing the Kalman lters as discussed above is optimal.

3.2.3 LQG Controller

Since the separation principle holds, the control cost can be minimized by solely optimizing over u_k while keeping E_k xed. Further, the optimal controller is linear and has the form $u_k^* = L_k \hat{x}_k^{\mathrm{c}} = -\left(B^{\mathrm{T}} G_{k+1} B + S\right)^{-1} B^{\mathrm{T}} G_{k+1} A \hat{x}_k^{\mathrm{c}}$
with $G_k = Q + A^{\mathrm{T}} G_{k+1} A - A^{\mathrm{T}} G_{k+1} B \left(B^{\mathrm{T}} G_{k+1} B + S\right)^{-1} B^{\mathrm{T}} G_{k+1} A$ and the condition $G_N = Q$.

3.2.4 Optimal Energy Allocation Policy

Due to the separation principle, the optimal energy allocation policy at the transmitter can be obtained by minimizing

$\min_{0 \leq E_k \leq B_k \vee k} \sum_{k=0}^{N-1} \mathbb{E}\left\{\operatorname{tr}\left(P_k^{\mathrm{c}}\right) | E_k\right\}$

This problem can be cast as a stochastic control problem where (g_k, B_k) forms the state process and E_k forms the control action. A finite horizon dynamic programming algorithm can be used to solve the corresponding backward Bellman dynamic programming equation to minimize the above cost.
$V_k\left(g_k, B_k\right) = \min_{0 \leq E_k \leq B_k} \left\{\operatorname{tr}\left(P_k^{\mathrm{c}}\right) | E_k + \mathbb{E}\left\{V_{k+1}\left(g_k, B_k\right)\right\}\right\}$
with the terminal condition $V_N\left(g_N, B_N\right) = \operatorname{tr}\left(P_N^{\mathrm{c}}\right) | B_N$.

3.3 Imperfect Acknowledgements

In this section, it will be shown that in case of a packet dropping channel feedback link, the separation principle does not hold. Hence, the optimal energy allocation policy, state estimation algorithm and controller design are not independent of one another.

3.3.1 Assumed State Estimate

Similar to the case of perfect feedback discussed in Section 3.2, the current state estimate at the Rx block is given by $\hat{x}_k^{\mathrm{c}} = \gamma_k \hat{x}_k + (1-\gamma_k)\left(A \hat{x}_{k-1}^{\mathrm{c}} + B u_{k-1}\right)$. The calculation of the estimate at the Tx block depends on the knowledge of the applied input signal u_k , which is not directly known by the estimator, but calculated by the controller. Hence, the estimator at the transmitter also has to estimate the current state estimate used by the controller to calculate u_k by using the information of the imperfect feedback channel:

$$
\hat{x}_k^{\mathrm{ce}} = (1-\beta_k)\left(h(g_k E_k)\hat{x}_k + (1-h(g_k E_k))\left(A\hat{x}_{k-1}^{\mathrm{ce}} + B\hat{u}_{k-1}\right)\right)
+ \beta_k\left(\gamma_k \hat{x}_k + (1-\gamma_k)\left(A\hat{x}_{k-1}^{\mathrm{ce}} + B\hat{u}_{k-1}\right)\right)
$$

In case an acknowledgment was received, the information of the acknowledgment is used. Otherwise, the state estimate at the controller is estimated using the package dropping probability of the forward channel.

3.3.2 Estimation Error Covariance Matrices and Kalman Filter

Note that $\hat{u}_k$ denotes the assumed control input at the transmitter. For simplicity, it will be assumed that the control input is a stationary policy of the form $u_k = L_k \hat{x}_k^{\mathrm{c}}$. Hence, the term $\hat{u}_k$ can be substituted by $L_k \hat{x}_k^{\mathrm{ce}}$. Since the transmitter only has knowledge of $\hat{u}_{k-1}$ and $\hat{x}_k^{\mathrm{ce}}$ instead of u_k and $\hat{x}_k^{\mathrm{c}}$, the estimation error $e_{k+1|k} = Ae_{k|k} + BL_k e_k^{\mathrm{e}} + w_k$ now depends on the error $e_k^{\mathrm{e}} := \hat{x}_k^{\mathrm{c}} - \hat{x}_k^{\mathrm{ce}}$. Defining $P_k^{\mathrm{e}} := \mathbb{E}\left\{e_k^{\mathrm{e}}\left(e_k^{\mathrm{e}}\right)^{\mathrm{T}}\right\}$, it follows

$$
P_{k+1} = AP_{k|k}A^{\mathrm{T}} + M + BL_k P_k^{\mathrm{e}} L_k^{\mathrm{T}} B^{\mathrm{T}} + BL_k \mathbb{E}\left\{e_k^{\mathrm{e}}\left(e_{k|k}\right)^{\mathrm{T}}\right\} A^{\mathrm{T}} + A\mathbb{E}\left\{e_{k|k}\left(e_k^{\mathrm{e}}\right)^{\mathrm{T}}\right\} L_k^{\mathrm{T}} B^{\mathrm{T}}
$$

Choosing again $K_k = P_{k|k-1}C^{\mathrm{T}}\left(CP_{k|k-1}C^{\mathrm{T}} + N\right)^{-1}$ in leads to the error covariance matrix after updating the estimate $P_{k|k}$ at the estimator as in the perfect feedback case. However, note that the K_k and the error covariance matrices P_{k+1} and $P_{k|k}$ now depend on the controller matrix L_k and mixed terms such as $\mathbb{E}\left\{e_k^{\mathrm{e}}\left(e_{k|k}\right)^{\mathrm{T}}\right\}$.

In case of imperfect acknowledgements the separation principle does not hold since the estimate depends on the controller matrix L_k . Further, the error covariance matrices $P_k, P_k^{\mathrm{c}}, P_k^{\mathrm{e}}$ and $P_k^{\mathrm{ce}} := \mathbb{E}\left\{e_k^{\mathrm{ce}}\left(e_k^{\mathrm{ce}}\right)^{\mathrm{T}}\right\}$ depend on each other, hindering an exact analysis. Since the separation principle does not hold, it is not optimal to design the estimator, the LQR controller and the energy allocation policy separately as done in Section 3.2 for the case of perfect channel feedback.

3.4 Concluding Remarks

This paper studied a closed loop control system where a sensor runs a local Kalman filter and sends it's state estimate to the receiver block consisting of the controller/actuator over a packet dropping link that has a time-varying packet loss probability due to a randomly time-varying fading channel gain. The transmitter is powered by a nite battery and can harvest a random amount of energy from its environment. The receiver sends an ACK/NACK feedback, which may also be lost intermittently. The objective is to design a jointly optimal sensor transmission energy allocation and optimal control design policy for minimizing a nite horizon LQG control cost.

In the case of perfect channel feedback, it is seen that the separation principle holds. Hence, the Kalman filters (at the transmitter and receiver) and a linear controller are optimal and the transmission energy allocation policy that minimizes the sum of the expected estimation error covariance over a nite time horizon can be obtained by standard dynamic programming techniques. In the case of erroneous channel feedback, the separation principle no longer holds. Hence, the optimal estimator and controller design, and the optimal energy allocation policy are in general coupled and nonlinear, and suboptimal designs are the only solution.

Chapter 4

Decentralized Control via Dynamic Stochastic Prices: The ISO Problem

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 benet accrued from its consumption or the negative of its generation cost. The social welfare, the sum of all these utilities, is the total benet 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 linear quadratic Gaussian (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.

4.1 System Model and ISO Problem

We consider a smart grid of M agents, each of which may act as a producer, consumer, or both, i.e., a prosumer, evolving over a time interval $t = 0, 1, ..., T-1$.
Agents are modelled as stochastic dynamical systems. The state $x_i(t)$ of agent i at t is known to it, and evolves as $x_i(t+1) = f_i(x_i(t), u_i(t), w_i(t), w_c(t), t)$, where f_i describes the dynamics of the agent i. The initial condition $x_i(0)$ can be random. The special case of particular interest is where $u_i(t)$ is a scalar that denotes the amount of electricity consumed (negative if supplied) from the grid by agent i at time t, but we allow it to be a vector of several commodities being produced and consumed. $\{w_c(\omega, t), 0 \le t \le T-1\}$ (defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$) is a "private uncertainity" (e.g. temperature of city) that affects and is observed by all the agents, while $\{w_i(\omega, t), 0 \le t \le T-1\}$ is a "private uncertainity" that specifically affects only

agent i, e.g., wind at a particular wind farm, and is known causally only to agent i. We suppose without loss of generality that $w_c(\cdot)$ and $w_i(\cdot)$ for $1 \leq i \leq M$ are all independent identically distributed (i.i.d.) processes with each random variable uniformly distributed on some alphabet W and that are also independent of each other. The function $f_i(\cdot, t)$ allows us to model dependent and time-varying stochastic process uncertainities as being formed out of these primitive i.i.d. Unif(W) uncertainities.

Consumption and 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_c^t. Agent i's choice has to satisfy the local capacity constraints $(F_i\,(w_i^t, w_c^t, t)\, u_i(t) \leq g_i\,(w_i^t, w_c^t, t)) + \sum_{s=0}^{t-1} J_i\,(w_i^t, w_c^t, s, t)\, u_i(s)$ and $h_i\,(w_i^t, w_c^t, t, u_i(t)) \leq 0$ for each t, almost surely.

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 a delay in performing a task resulting from inadequate purchase of electricity etc. $e_i\,(u_i, t)$ can be thought of as a one-step tax on the use of u_i units of electricity at time t.

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 can allow for more general linear vector constraints of the form: $\sum_{i=1}^{M} K_i(t) u_i(t) = d(t)$.

Knowledge available to the agents and the ISO: In the general case, we assume that all agents and the ISO know the alphabet W. Each agent also knows its own one-step cost function and it's own constraints. 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.

The *Independent System Operator (ISO)* solicits electricity purchase/sale bids from the agents in each time slot $t = \{0, 1, ..., T-1\}$, 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 nite horizon total of the one-step costs incurred by all the agents plus any taxes, $\mathbb{E} \sum_{i=1}^{M+1} \sum_{t=0}^{T-1} \left[c_i\left(x_i(t), u_i(t), t\right) + e_i\left(u_i(t), t\right)\right]$. 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 $t = \{0, 1, ..., T-1\}$, consisting of all the private uncertainties and the common uncertainties, as well as the random initial conditions of all the M agents.

Goal of social welfare maximization: Let F_t be the σ-algebra generated by all the noises upto 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 F_t , we call it the full state information case. *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,
$\min \mathbb{E} \sum_{t=0}^{T-1} \sum_{i=1}^{M} \left[c_i\left(x_i(t), u_i(t), t\right) + e_i\left(u_i(t), t\right)\right]$
such that $x_i(t+1) = f_i\left(x_i(t), u_i(t), w_i(t), w_c(t), t\right)$;
with capacity constraints $h_i\left(w_i^t, w_c^t, t, u_i(t)\right) \leq 0$, $F_i\left(w_i^t, w_c^t, t\right) u_i(t) \leq g_i\left(w_i^t, w_c^t, t\right) + \sum_{s=0}^{t-1} J_i\left(w_i^t, w_c^t, s, t\right) u_i(s)$ and $\sum_{i=1}^{M} K_i(t) u_i(t) = d(t)$ 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, it is not a standard one since the agents do even know the quantities involved in the optimization problem.

4.2 Stochastic Dynamic Agents: Bid-Price Iterations

Denote the combined state of the system by $x(t) := (x_1(t), x_2(t), ..., x_M(t))$, and the combined actions by $u(t) := (u_1(t), u_2(t), ..., u_M(t))$.

A tree visualization of the system randomness, as in Fig. 1, is helpful. Suppose that w(t) assumes only nitely 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 $w(0, \omega), w(1, \omega), ..., w(s-1, \omega)$ corresponding to a specic node v at the level s.

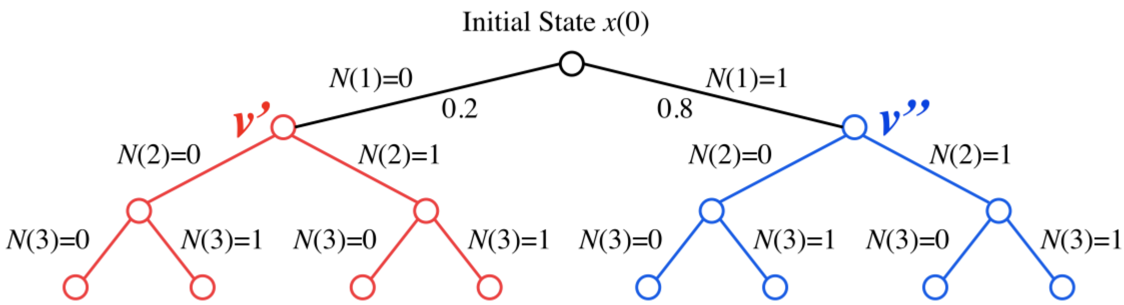

Figure 4.1: Tree visualization of i.i.d. uncertainty for a system evolving over three bid times, where the uncertainty values are binary, either 0 or 1. Each sequence of noise realizations corresponds to a specic node v at the level s. Note that the red and blue subtrees with roots at nodes v' and v'' 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 toward 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 ${}^{k}_{s}$ 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 sub-trees of nodes rooted at level $s+1$ are identical, and simply perform these announcements for a tree of depth Ts, with the understanding that these bids apply to the sub-tree 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 $(u^k_{i,s}(s), u^k_{i,s}(s+1), ..., u^k_{i,s}(T-1))$ in response to the price stochastic process ${}^{k}_{s}$ by solving the following problem, dubbed Agent i's Problem:

$\min_{u_i \in U_i} \mathbb{E}\left[\sum_{t=s}^{T-1} \left[c_i\left(x_i(t), u_i(t), t\right) + e_i\left(u_i(t), t\right) \ +\lambda^k_s(t)^T K_i(t) u_i(t)\right] |\mathcal{F}_{i,s}\right]$
where U_i is the constraint set and,
$\mathcal{F}_{i,t} := \sigma(x_i(0), w_i(0), w_i(1), \ldots, w_i(t-1), w_c(0), w_c(1), \ldots w_c(t-1)$ denotes the sigma-algebra generated by agent i's observations upto 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. Guided by the "excess consumption function" $\sum_{i=1}^{M} K_i(t) u^k_{i,s}(t) - d(t)$, it raises or lowers prices to satisfy the general linear constraint, as follows:
$\lambda^{k+1}_s(t) = \lambda^k_s(t) + \frac{1}{k}\left[\sum_{i=1}^{M} K_i(t) u^k_{i,s}(t) - d(t)\right]; \ s \le t \le T-1$

The averaged allocations of consumption/generation: At time s, after the prices have converged, i.e.,
$\lambda^\star_s(t) := \lim_{k\to\infty} \lambda^k_s(t)$ for $s \le t \le T-1$
the ISO announces the allocations at the current time s as the limit, $u^\star_{i,s}(s) := \lim_{k\to\infty} \bar{u}^k_{i,s}(s)$ of the following average of the iterates of the bids for time s,

$\bar{u}^k_{i,s}(s) = \frac{\sum_{s=1}^{k-1} s^\theta}{\sum_{s=1}^{k} s^\theta} \bar{u}^{k-1}_{i,s}(s) + \frac{k^\theta}{\sum_{s=1}^{k} s^\theta} u^k_{i,s}(s)$, with $\bar{u}^0_{i,s}(s) = u^0_{i,s}(s)$.

The above bid-price solution, with price updates as defined above, bid updates determined as the optimal solution of individual agent's problem, and allocations at each t given by $\lim_{k\to\infty} \bar{u}^k$ where $\overline{u}^k$ is obtained as the averaged version of u_k, achieves the maximum social welfare that could have been attained in the full state information case, when the cost functions satisfies certain assumptions as mentioned in [5].

It is shown in [5] that when the agents are all LQG systems (linear dynamics affected by Gaussian noise and have quadratic costs), one can dramatically simplify the bids and the bidding process. Even though agents have private uncertainties, each bid consists

Algorithm 1: Stochastic Dynamic Agents With Private Uncertainties.

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

$k = 0$

repeat

Each agent i solves the problem

$$\text{Min} \quad \mathbb{E} \sum_{t=s}^{T-1} [c_i(x_i(t), u_i(t), t) + e_i(u_i(t), t) + \lambda^k(t)^T K_i(t) u_i(t)],$$

with initial condition $x_i(s)$ to obtain the optimal $\{u_{i,s}^k(t), s \le t \le T-1\}$ subject to (17), (18), and submits it to the ISO.

The ISO declares new prices for $s \le t \le T-1$, as

$$\lambda^{k+1}(t) = \lambda^k(t) + \frac{1}{k}\left[\sum_{i=1}^{M} K_i(t) u_i^k(t) - d(t)\right].$$

$k \to k+1$

until $\lambda^k(t)$ converges a.s. to $\lambda^\star(t)$ for $s \le t \le T-1$.

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

end for

Figure 4.2: Pseudo-code for the bid-price iteration

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 rst step of the prices and consumptions/generations at each time t is implemented. 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 satises energy balance and other constraints.

4.3 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.

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. The agents are all presumed to be “price takers”. Examining this in a broader context is an important issue and is an ongoing research area.

References

[1] L. Schenato, B. Sinopoli, M. Franceschetti, K. Poolla, and S. Sastry, “Foundations of control and estimation over lossy networks,” Proc. IEEE, vol. 95, no. 1, pp. 163–187, 2007.

[2] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations", IEEE Trans. Autom. Contr., vol. 49, no. 9, pp. 1453–1464, Sep. 2004.

[3] Steffi Korn, and Subhrakanti Dey, "Optimal Sensor Transmission Energy Allocation for Linear Control Over a Packet Dropping Link with Energy Harvesting", IEEE Conf. on Decision and Control, pp. 1199-1204, Dec. 2015.

[4] M. Nourian, A. Leong, and S. Dey, “Optimal energy allocation for Kalman ltering over packet dropping links with imperfect acknowledgments and energy harvesting constraints,” IEEE Trans. Autom. Control, vol. 59, no. 8, pp. 2128–2143, 2014.

[5] Rahul Singh, P.R. Kumar, and Le Xie, "Decentralized Control via Dynamic Stochastic Prices: The Independent System Operator Problem", IEEE Trans. Autom. Contr., vol. 63, no. 10, pp. 3206-3220, Oct. 2018.