

Chapter 2

Dual Control in Big Data Era: An Overview

Peilin Fu

Abstract This paper provides an overview of dual control and its applications in the big data era. Different non-dual and dual controllers as well as their attributes, complexity, and limitations are analyzed. As an example, dual control of a class of discrete-time LQG problems with unknown parameters in both the state and observation equations is discussed in depth. Optimal dual control, open-loop feedback control, active open-loop feedback control via variance minimization approach, and optimal nominal dual control are demonstrated for this type of problems. The optimal nominal dual control, taking into account the effect of future learning, is the best possible (partial) closed-loop feedback control that can be achieved. Applications of dual control in economic systems, manufacturing processes, information retrieval, robotics, etc. are also introduced.

Keywords Dual control • Stochastic control • Dynamic programming • LQG control problem

2.1 Introduction

Most real-world processes are very complex and are not well understood. As such, the control of systems whose dynamics are not completely known is a problem of major theoretical and practical importance. Feldbaum, in his seminal work in 1960s (Feldbaum 1965), pointed out that, when implementing the optimal control strategy for stochastic systems with parameter uncertainty, the controller usually pursues two often conflicting objectives: to drive the system toward a desired state, and to perform active learning to reduce the system's uncertainty. Such a control scheme, which affects not only the states of the system but also the quality of estimation, is known as dual control. In 2000, IEEE Control Systems Society listed the dual control as one of the 25 most prominent subjects in the last century which had significantly impacted the development of control theory.

P. Fu (✉)

Department of Applied Engineering, School of Engineering and Computing,
National University, San Diego, CA, USA

© Springer International Publishing AG 2017

T.-M. Choi et al. (eds.), *Optimization and Control for Systems in the Big-Data Era*,
International Series in Operations Research & Management Science 252,
DOI 10.1007/978-3-319-53518-0_2

Except for a few ideal situations, the optimal dual control cannot be achieved both analytically and numerically. Feldbaum showed that the optimal dual control is the solution to a functional equation known as Bellman equation based on dynamic programming. Solving this functional equation is intractable due to the “curse of dimensionality” inherent in dynamic programming. The two subproblems of stochastic control, estimation and control in most situations are intercoupled. The future uncertainties of the parameters are functions of the control signals applied to the system. The loss function, which has to be minimized with respect to the control signal, thus contains some information of the future observations through the statistics of the observations given the present information (Bar-Shalom and Tse 1974). The efforts in dual control have thus mainly been devoted to developing certain suboptimal solution schemes, such as the certainty equivalence scheme and open-loop feedback control, by bypassing this essential feature of coupling between estimation and control.

The control policies were categorized into the following classes in Bar-Shalom and Tse (1974) according to their information patterns—the availability of past observation and the possible usage of information about the future observation:

- (1) The Open-loop Policy. In this case no measurement knowledge is available for the controller.
- (2) The Feedback Policy. At every time the current information set is available for the computation of the control but no knowledge about the future measurements is available. The open-loop optimal feedback (OLOF) control belongs to the feedback class. It assumes that no observations will be made in the future, the control law is obtained by using the observations already acquired.
- (3) The Closed-Loop Policy. This policy incorporates with the remaining observation program, i.e. the knowledge that the loop will stay closed through the end of the process is fully utilized.

There are two aspects in which the closed-loop policy differs from the feedback policy (Bar-Shalom and Tse 1974).

- (1) Caution: In a stochastic control problem, due to the inherent uncertainties, the controller has to be “cautious” not to increase the effect of the existing uncertainties on the cost. However, the closed-loop controller, since it “knows” that future observations will be available and corrective actions based upon them will be taken, will exercise less “caution.”
- (2) Probing or Active Learning: When the dual effect is present, the control can “help” in learning (estimation) by decreasing the uncertainty about the state. Therefore, the closed-loop control, which takes into account the future observation program and statistics, has the capability of active learning when the dual effect exists. A feedback controller, even though it “learns” by using the measurements, does not actively “help” the learning. This learning can be called, therefore, passive, or accidental, and the corresponding control policy is passively adaptive, as opposed to the closed-loop control which is actively adaptive.

Most resulting suboptimal control laws are of a nature of passive learning, since the function of future active probing of the control is purposely deprived in order to achieve analytical attainability in the solution process. A central problem in dual control, and indeed a key barrier to its development, is to power a control law with the property of active learning.

Prominent features and fundamental properties of dual control have been extensively studied in the literature (Bar-Shalom 1981; Bar-Shalom and Tse 1974; Tse et al. 1973). An analysis of various approximations in dual control was given by Lindoff et al. (1999). Filatov and Unbehauen (2000) developed a bi-criteria approach to cope with the two conflicting goals in dual control. Surveys on dual control can be found in Wittenmark (1975c) and Filatov and Unbehauen (2000).

Li, Qian, and Fu in a series of papers studied the dual control of discrete-time LQG problems with unknown parameters. A variance minimization approach was proposed for discrete-time LQG problems with parameters uncertainty in the observation equation (Li et al. 2002). Minimizing a covariance term at the final stage introduced a feature of active learning for the derived control law. The optimal degree of active learning was determined for achieving an optimality. Fu et al. (2002) further applied the variance minimization approach to discrete-time LQG problems with parameters uncertainty in both the state and observation equations, an optimal open-loop feedback control law with active learning property was developed. The same problem was revisited in Li et al. (2008), in which the optimal nominal dual control was proposed. By exploring the future nominal posterior probabilities, the control law takes into account the function of future learning, thus is the best possible closed-loop feedback control that can be achieved. Some of these results are summarized in Sect. 2.3 as an example of dual control problems.

In Sect. 2.2, classification of controllers is introduced. Different non-dual and dual controllers as well as their attributes, complexity, and limitations are analyzed. As an example, dual control of a class of discrete-time LQG problems with unknown parameters in both the state and observation equations is discussed in depth in Sect. 2.3. Optimal dual control, open-loop feedback control, active open-loop feedback control, and optimal nominal dual control are demonstrated. Section 2.4 provides successful applications of dual control in economic systems, manufacturing processes, information retrieval etc. in the big data era. The paper concludes in Sect. 2.5.

2.2 Classification of Controllers

2.2.1 *Non-dual Controller*

If the performance index only takes into account the previous measurements and does not assume that future information will be available, then the resulting controller will be called non-dual in Feldbaum's terminology. In this situation the control law does not facilitate the identification. The non-dual controllers can

be divided into three classes: certainty equivalence controller, one-step cautious controller, and open-loop optimal feedback controller.

2.2.1.1 Certainty Equivalence Controller

One widely used non-dual approach is developed using the concept of certainty equivalence. The certainty equivalence holds if it is possible to first solve the deterministic problem with known parameters and then obtain the optimal controller for unknown parameters by substituting the true parameter values with the estimated values (Wittenmark 1975c). One well-known class of problems for which the certainty equivalence principle holds is the linear-quadratic-gaussian control problems. In adaptive control there are very few cases where the certainty equivalence principle is applicable. The controller obtained by enforcing the certainty equivalence principle does not take into consideration the fact that the estimated parameters are not equal to the true ones and are inaccurate. Although the simplicity of the control law, it ignores the confidence level of the parameter estimates in deriving the adaptive control scheme. Such a control scheme would result in a control system that is extremely sensitive to stochastic variations.

A method based on process parameter estimation was first described by Kalman (1958) using least squares to determine the unknown parameters in the model. This type of method works well for constant or slow time-varying parameters. Different approximation methods (Hasting-James and Sage 1969; Panuska 1968; Young 1968) have been suggested for the models of maximum likelihood type. Methods using state space models were given in Jenkins and Roy (1966) and Luxat and Lees (1973).

2.2.1.2 One-Step Cautious Controller

Minimizing over a single time period leads to the one-step cautious controller. This controller takes the parameter uncertainties into account, in contrast to the certainty equivalence controller. However, the controller of this type may generate the turn-off phenomenon. If the estimates are very poor, the magnitude of the control signal will become very small. The control is thus unintentionally turned off for some period of time until the noise excites the system in such a way that better estimates are achieved. This makes the one-step cautious controller unsuitable for control of systems with quickly varying parameters.

A one-step minimization where the unknown parameters are modeled by a stochastic process was discussed in Aoki (1967) and Astrom and Wittenmark (1971). The unknown parameters can be first estimated using a Kalman filter, which then give the one-step ahead estimates and covariance matrix based on the current information set. Using a fundamental lemma in stochastic optimal control (Astrom 1970), it is possible to find the control law to solve the one-step minimization problem. The control law clearly shows the influence of the uncertainties of the estimates. Examples with turn-off were given, for instance, by Astrom and Wittenmark (1971).

2.2.1.3 The Open-Loop Feedback Optimal Controller

The open-loop optimal feedback (OLOF) control is derived at distinct time instants under the assumption that no future measurements will be available. Thus an open-loop control sequence is determined. The first step in the control sequence is then used and the performance of the system is measured. Based on the new information (feedback), a new minimization is performed again. In this open-loop feedback approach, the fact that the estimated parameter may not be exact is therefore taken into consideration, but the knowledge of the future observation program is completely ignored. According to the theory of dual control, introduced by Feldbaum (1965), the open-loop feedback control is, from the estimation point of view, passive, since it does not take into account that learning is possible in the future.

Many suboptimal controllers achieved in the current literature are open-loop optimal feedback controllers (Florentine 1962; Tse and Athans 1972; Aoki 1967). Lainiotis, Deshpande, and Upadhyay wrote a series of papers (Deshpande et al. 1973; Lainiotis et al. 1972) on an open-loop feedback optimal approach to the stochastic control of linear systems with unknown parameters. The controller is designed to minimize the average performance-to-go conditioned on the present measurements and past control actions and without any active anticipation of new measurements. The result is a feedback control law similar to the optimal LQG one, but averaged over the space of the unknown parameters. The algorithm is straightforward and easy to implement. It may be generated by computing the average of the specific controllers for some value of the parameters weighted by the a posteriori probability densities which are Gaussian (Deshpande et al. 1973). Casiello and Loparo proved in Casiello and Loparo (1989) that these types of passive control laws are optimal for certain quadratic functionals.

The OLOF controller might be overly cautious because of the assumption that no further measurements will be available to correct for erroneous control actions. The properties of the OLOF controller were further discussed by Bar-Shalom and Sivan (1969) and Tse and Athans (1972).

2.2.2 Dual Controller

If besides the previous measurements the performance index is also considered to be dependent on the future observations, a dual controller will be constructed. In this case, the future uncertainties of the parameters are functions of the control applied to the system. The control law must compromise between the two conflicting tasks: control and identification. The dual controllers can be classified as optimal dual controller and suboptimal dual controller.

2.2.2.1 Optimal Dual Controller

There are very few cases where it is possible to obtain an analytical representation of the optimal dual control law. The imposed assumptions are usually unrealistic. In Gorman and Zaborszky (1968) and Grammaticos and Horowitz (1970), the problem of controller synthesis was considered under the assumption that the entire state is measurable. Moreover, it was assumed that the poles of the system are known while the zeros are unknown. In both cases, it is possible to find the optimal dual control by solving a set of differential or difference equations corresponding to the Riccati equation in the standard linear quadratic case. Sternby (1978) discussed a Markov chain with four states. The transition probabilities are functions of the control. In that particular example it is possible to find the analytical expression of the optimal dual controller.

Some results can be seen in the literature to achieve the optimal dual controller numerically. Florentine (1962) considered a first-order system where the gain is fixed but unknown with a given a priori distribution. The problem was solved by discretizing the state and control. Another numerically solved problem was given by Jacobs and Langdon (1970). The absolute value of the state can be measured through the observation while the sign is unknown. Introducing the probability for the state to be positive, it is possible to derive the corresponding functional equation. A zero-order system with an unknown gain was considered in Astrom and Wittenmark (1971) where the gain was assumed to be described by a known stochastic process. A more general treatment of the problem was given by Griffiths and Loparo (1985).

A variance minimization approach for dual control of discrete-time LQG problems with parameter uncertainty in the observation equation was proposed by Li et al. (2002). Minimizing a covariance term at the final stage introduced a feature of active learning for the derived control law. The optimal degree of active learning was derived for achieving the optimality.

2.2.2.2 Suboptimal Dual Controller

Since it is difficult to determine the optimal dual controllers, much effort has been devoted to finding suboptimal solutions with dual properties. The approaches can be classified as follows (Wittenmark 1975c; Astrom and Wittenmark 1989):

(1) Perturbation signals

Employing a cautious controller can give rise to “turn-off” of the control if the unknown parameters are strongly time-varying. Several ways have been suggested to avoid the turn-off phenomenon. The turn-off is due to a lack of excitement. The perturbation signal, which can be a square-wave or pseudo random signal, etc., can be used to excite the system in order to get good estimation (Wieslander and Wittenmark 1971). The addition of the extra signal will naturally increase the probing loss, but may make it possible to improve the total performance.

(2) Constrained one-step-ahead minimization

Another way to avoid turn-off is to minimize the loss function one-step ahead under certain constraints. The constraints such as limiting the minimum value of the control signal or limiting the variance of the parameter estimates can prevent the control signal from being too small and impose extra probing (Hughes and Jacobs 1974; Alster and Belanger 1974). These controllers have the advantage that the control signal can be easily computed, but the algorithm will contain application-dependent parameters that have to be chosen by the user.

(3) Approximations of the loss function

Suboptimal dual controls can also be obtained by extending the loss function in order to prevent the shortsightedness of the cautious controller (Astrom and Wittenmark 1989). For state space models, one approach is to make a serial expansion of the loss function in the Bellman equation (Gorman and Zaborsky 1968). Such an expansion can be done around the certainty equivalence or the cautious controllers. But due to its computational complexity, this approach has been limited to situations where the control horizon is rather short, usually less than 10.

Another way is to try to solve the two-step minimization problem. The derived suboptimal control has correction terms which depend on the sensitivity functions of the expected future cost, which can avoid the turn-off. But in most cases, it is not possible to get an analytical solution.

(4) Modifications of the loss function

Adding terms that are reflecting quality of the parameter estimate in the loss function can prevent the cautious controller from turning off. Solution proposals (Alster and Langer 1974; Wittenmark 1975b; Milito et al. 1982) have been seen in the literature to incorporate certain variance terms of the state or the innovation process into the objective function in order to force the control to perform active learning. These solution schemes, however, truncate the time horizon into shorter time periods of one stage, prompting a concern of possible myopic behaviors.

2.2.2.3 Optimal Nominal Dual Controller

Although the optimal nominal dual controller is also suboptimal, the author would like to list it as a separate category to distinguish it from other suboptimal dual controllers. The reason is that it is the best possible closed-loop dual control if the optimal dual control cannot be achieved. The optimal nominal dual control was first proposed by Li, Qian and Fu in Li et al. (2008). They pointed out that a major difficulty in solving dual control for discrete-time LQG problems with unknown parameters is that the optimal control cannot be determined when the future posterior probabilities are unknown, while at the same time the future posterior probabilities depend on the control applied at the early stages. In order to break this loop, a possible solution scheme is to derive the relationship between the posterior probability and the control. A control which satisfies a deterministic

version of this relationship is defined as the nominal control. The expected posterior probabilities when applying the nominal control are called nominal future posterior probabilities. Applying the nominal future posterior probabilities generated by the nominal control in the Bellman equation, the effect of future learning can be taken into account. Since in this situation, all the achievable future information is used in terms of its expected value, the control law obtained can be considered to be the best possible closed-loop control law in this sense.

2.3 An Example: LQG Problems with Unknown Parameters

Consider the following class of linear-quadratic stochastic optimal control problems where there exist parameter uncertainties in both the state and the observation equations,

$$(P) \min E \left\{ x'(N)Q(N)x(N) + \sum_{k=0}^{N-1} [x'(k)Q(k)x(k) + u'(k)R(k)u(k)] \mid I^0 \right\}$$

$$\text{s.t. } x(k+1) = A(k, \theta)x(k) + B(k, \theta)u(k) + w(k), \quad k = 0, 1, \dots, N-1$$

$$y(k) = C(k, \theta)x(k) + v(k), \quad k = 1, 2, \dots, N,$$

where $x(k) \in R^n$ is the state, $u(k) \in R^p$ is the control, $y(k) \in R^m$ is the measured output, and I^0 is the initial information set that includes information about the probability distribution of the initial state $x(0)$, the statistics of the random sequences $\{w(k)\}$ and $\{v(k)\}$, and the initial probability distribution of the unknown parameter θ . $\{w(k)\} \in R^n$ and $\{v(k)\} \in R^m$ are the two independent Gaussian white noise sequences with zero mean, and variances σ_w^2 and σ_v^2 , respectively. The random initial state $x(0)$ is assumed to be of Gaussian distribution $N(\hat{x}(0), P(0))$ and is assumed to be independent of the process and observation noises: The quantities $A(k, \theta)$, $B(k, \theta)$, and $C(k, \theta)$ are matrices of appropriate dimensions whose values depend on an unknown parameter θ . It is assumed that θ belongs to a finite set $\Theta = \{\theta_1, \theta_2, \dots, \theta_s\}$ and is a constant over the entire time horizon. The a priori probabilities of the parameter θ are

$$q_i(0, I^0) = P(\theta = \theta_i \mid I^0), \quad i = 1, 2, \dots, s.$$

Furthermore, $\{Q(k)\}$ and $\{R(k)\}$ are sequences of positive semidefinite and positive definite symmetric matrices of appropriate dimensions, respectively. Define the information set at stage k , $k = 0, 1, \dots, N$, to be I^k ,

$$I^k = \{u(0), \dots, u(k-1), y(1), \dots, y(k), I^0\}.$$

The dual control problem for (P) is to find a closed-loop control law,

$$u(k) = f_k(I^k), \quad k = 0, 1, \dots, N-1,$$

such that the expected performance index in (P) is minimized.

Notice that two kinds of uncertainty are involved in (p): irreducible uncertainty caused by Gaussian white noise sequences $\{w(k)\}$ and $\{v(k)\}$, and reducible uncertainty caused by an unknown parameter θ . If there is no parameter uncertainty about θ , the above problem reduces to the conventional linear-quadratic Gaussian stochastic control problem which is not a dual control problem since the control does not have an effect on the system's uncertainty. The certainty equivalence principle then can be applied to determine the optimal control. Note that the certainty equivalence principle may not hold even for some stochastic control problems with only irreducible uncertainty, for example, linear Gaussian systems with an exponential performance criterion (Jacobson 1973).

2.3.1 Optimal Dual Control

Define $\hat{x}_i(k|k)$ to be the state estimate at stage k when assuming $\theta = \theta_i$:

$$\hat{x}_i(k|k) = E \{x(k) | \theta = \theta_i, I^k\}.$$

$\hat{x}_i(k|k)$ can be obtained using the Kalman filters as stated in Casiello and Loparo (1989):

$$\hat{x}_i(k|k) = \hat{x}_i(k|k-1) + F_i(k) [y(k) - C(k, \theta_i)\hat{x}_i(k|k-1)] \quad (2.1)$$

$$\hat{x}_i(k|k-1) = A(k-1, \theta_i)\hat{x}_i(k-1|k-1) + B(k-1, \theta_i)u(k-1) \quad (2.2)$$

$$F_i(k) = P_i(k|k-1)C'(k, \theta_i)[C(k, \theta_i)P_i(k|k-1)C'(k, \theta_i) + \sigma_v^2(k)]^{-1} \quad (2.3)$$

$$P_i(k|k-1) = A(k-1, \theta_i)P_i(k-1|k-1)A'(k-1, \theta_i) + \sigma_w^2 \quad (2.4)$$

$$P_i(k|k) = [I - F_i(k)C(k, \theta_i)]P_i(k|k-1), \quad (2.5)$$

with the initial condition of $\hat{x}_i(0|0) = \hat{x}_i(0)$ and $P_i(0|0) = P(0)$.

Define $q_i(k, I^k)$ to be the posterior probability of model i at stage k ,

$$q_i(k, I^k) = P(\theta = \theta_i | I^k), \quad k = 0, 1, \dots, N-1.$$

The posterior probabilities, $q_i(k, I^k)$, $i = 1, 2, \dots, s$, can be calculated recursively based on the observation (Casiello and Loparo 1989) as follows:

$$q_i(k, I^k) = \frac{L_i(k)}{\sum_{j=1}^s q_j(k-1, I^{k-1})L_j(k)} q_i(k-1, I^{k-1}), \quad k = 1, 2, \dots, N, \quad (2.6)$$

with the initial condition $q_i(0, I^0)$, where

$$L_i(k) = |P_y(k|k-1, \theta_i)|^{-\frac{1}{2}} \exp[-\frac{1}{2} \tilde{y}(k|k-1, \theta_i)' P_y(k|k-1, \theta_i)^{-1} \times \tilde{y}(k|k-1, \theta_i)] \quad (2.7)$$

$$\tilde{y}(k|k-1, \theta_i) = y(k) - C(k, \theta_i) \hat{x}_i(k|k-1) \quad (2.8)$$

$$P_y(k|k-1, \theta_i) = C(k, \theta_i) P_i(k|k-1) C'(k, \theta_i) + \sigma_v^2(k). \quad (2.9)$$

Define for $i = 1, 2, \dots, s$,

$$J_i(k, I^k) = E\{x'(k)Q(k)x(k) + u'(k)R(k)u(k) \mid \theta_i, I^k\}, \quad k = 0, \dots, N-1$$

$$J_i(N, I^N) = E\{x'(N)Q(N)x(N) \mid \theta_i, I^N\}.$$

Then the following is obvious,

$$\begin{aligned} J(k, I^k) &= E\{x'(k)Q(k)x(k) + u'(k)R(k)u(k) \mid I^k\} \\ &= \sum_{i=1}^s q_i(k, I^k) J_i(k, I^k) \quad k = 0, 1, \dots, N-1 \\ J(N, I^N) &= E\{x'(N)Q(N)x(N) \mid I^N\} \\ &= \sum_{i=1}^s q_i(N, I^N) J_i(N, I^N). \end{aligned}$$

By the principle of stochastic dynamic programming, the closed-loop control that minimizes the performance index in problem (P) can be obtained by solving the following recursive relation,

$$\begin{aligned} \min_{u(0)} E \left\{ \sum_{i=1}^s q_i(0, I^0) J_i(0, I^0) \right. \\ + \min_{u(1)} E \left\{ \sum_{i=1}^s q_i(1, I^1) J_i(1, I^1) + \dots \right. \\ + \min_{u(k)} E \left\{ \sum_{i=1}^s q_i(k, I^k) J_i(k, I^k) + \dots \right. \\ + \min_{u(N-1)} E \left[\sum_{i=1}^s q_i(N-1, I^{N-1}) J_i(N-1, I^{N-1}) \right. \\ \left. \left. + \sum_{i=1}^s q_i(N, I^N) J_i(N, I^N) \mid I^{N-1} \right] \dots \mid I^k \right\} \dots \mid I^0 \left. \right\}. \quad (2.10) \end{aligned}$$

In principle, the optimal dual control problem (P) can be solved via (2.10). However, the difficulty and complexity in solving (P) hide deeply behind these seemingly straightforward equations. In fact, in dual control problems, all of the posterior probabilities at later stages are affected by previous controls. The curse of uncertainty of the posterior probabilities in later stages is further compounded by the required expectation operations. Therefore, to derive the cost-to-go functions in stochastic dynamic programming from (2.10) is a formidable task, as long as the posterior probabilities at later stages are previously control-dependent.

2.3.2 Open-Loop Feedback Control

Suppose that future learning will not be performed, the open-loop feedback control can be obtained by fixing all the posterior probabilities in the later stages at $q_i(k, I^k), i = 1, 2, \dots, s$. As a result, the following optimal open-loop feedback control problem is considered at stage k ,

$$\begin{aligned} \min_{u(k)} E \left\{ \sum_{i=1}^s q_i(k, I^k) J_i(k, I^k) + \dots \right. \\ \left. + \min_{u(N-2)} E \left\{ \sum_{i=1}^s q_i(k, I^k) J_i(N-2, I^{N-2}) \right. \right. \\ \left. + \min_{u(N-1)} E \left[\sum_{i=1}^s q_i(k, I^k) J_i(N-1, I^{N-1}) \right. \right. \\ \left. \left. + \sum_{i=1}^s q_i(k, I^k) J_i(N, I^N) | I^{N-1} \right] | I^{N-2} \right\} \dots | I^k \left. \right\}. \quad (2.11) \end{aligned}$$

A controller that uses observations to update online the estimation of the uncertain parameter is said to have a learning feature. The learning policies can be further classified into two types—active learning and passive learning. We can always expect an improving knowledge about the system's uncertainty when future observations are utilized. A controller that takes the future uncertainty reduction is said to have a property of active learning. In return, a controller with an active learning property affects the degree of future uncertainty reduction. To power a control law with a property of active learning is, in general, needed to achieve an optimality in dual control (Griffiths and Loparo 1985). The open-loop feedback control law is a passive scheme that does not possess an active learning feature (as it does not take into account any impact from the future learning) and thus can never be optimal.

2.3.3 Active Open-Loop Feedback Control: Variance Minimization Approach

A degree of success of active learning can be measured by the variance of the final state. Therefore minimizing a variance term of the final state will add a feature of active learning to the derived control law. In this section, we consider a modified problem ($M_a(\mu)$) in which a variance term at the final stage is attached to the performance index of (P),

$$(M_a(\mu)) \min E \left\{ x'(N)Q(N)x(N) + \sum_{k=0}^{N-1} [x'(k)Q(k)x(k) + u'(k)R(k)u(k)] \mid I^0 \right\} \\ + \mu \text{Tr}[Cov(x(N) \mid I^0)] \\ \text{s.t. } x(k+1) = A(k, \theta)x(k) + B(k, \theta)u(k) + w(k) \quad k = 0, 1, \dots, N-1 \\ y(k) = C(k, \theta)x(k) + v(k), \quad k = 1, 2, \dots, N$$

Parameter $\mu \in [0, \infty)$ is a weighting coefficient of active learning. A larger μ implies that more importance has been placed on active learning.

Problem ($M_a(\mu)$) is difficult to be solved directly, since the recursive equations of dynamic programming involve certain nonlinear terms of the state estimates that introduces a nonseparability in the sense of dynamic programming. In order to overcome this difficulty, problem ($M_a(\mu)$) is embedded into a tractable auxiliary problem in which the optimal open-loop feedback control can be found. Solving the auxiliary problem and investigating the relationship between the solution sets of problem ($M_a(\mu)$) and the auxiliary problem, the optimal control of problem ($M_a(\mu)$) can be identified.

Define $S(N) = Q(N) + \mu I$, the performance index of ($M_a(\mu)$) can be written as

$$J = E \left\{ x'(N)S(N)x(N) + \sum_{k=0}^{N-1} [x'(k)Q(k)x(k) + u'(k)R(k)u(k)] \mid I^0 \right\} \\ - \mu E(x(N) \mid I^0)' E(x(N) \mid I^0). \quad (2.12)$$

Let

$$J^I = E \left\{ x'(N)S(N)x(N) + \sum_{k=0}^{N-1} [x'(k)Q(k)x(k) + u'(k)R(k)u(k)] \mid I^0 \right\} \\ J^{II} = E(x(N) \mid I^0).$$

It is easy to see that the performance index in ($M_a(\mu)$), J , is a concave function of J^I and J^{II} ,

$$J(J^I, J^{II}) = J^I - \mu (J^{II})' J^{II}. \quad (2.13)$$

The following auxiliary parametric problem is now constructed for the problem $(M_a(\mu))$ with a fixed multiplier vector $r \in R^n$,

$$\begin{aligned} (A(r, \mu)) \min E \left\{ x'(N)S(N)x(N) + \sum_{k=0}^{N-1} [x'(k)Q(k)x(k) + u'(k)R(k)u(k)] \right. \\ \left. - 2r'x(N) \mid I^0 \right\} \\ \text{s.t. } x(k+1) = A(k, \theta)x(k) + B(k, \theta)u(k) + w(k) \quad k = 0, 1, \dots, N-1 \\ y(k) = C(k, \theta)x(k) + v(k), \quad k = 1, 2, \dots, N. \end{aligned}$$

Theorem 1 *Suppose that $\{u^*(k)\}$ is an optimal open-loop feedback control of problem $(M_a(\mu))$, then $\{u^*(k)\}$ is also an optimal open-loop feedback control of the auxiliary parametric problem $(A(r^*, \mu))$ where r^* satisfies*

$$r^* = \mu E(x(N) \mid I^0) \mid_{\{u^*(k)\}}. \quad (2.14)$$

The implication of Theorem 1 is that any optimal open-loop feedback solution to problem $(M_a(\mu))$ is in the set of optimal open-loop feedback solutions to auxiliary problem $(A(r, \mu))$. Note that the auxiliary problem is strictly convex with respect to $\{u(k)\}$. Thus the optimal open-loop feedback solution to problem $(A(r, \mu))$ is unique for a given r . As a result, if r satisfies the optimality condition in (2.14), then the optimal open-loop feedback control to $(A(r^*, \mu))$ becomes a possible candidate for the optimal open-loop feedback control to $(M_a(\mu))$.

Define for $i = 1, 2, \dots, s$,

$$\begin{aligned} J_i(k, I^k) = E \{ x'(k)Q(k)x(k) + u'(k)R(k)u(k) \mid \theta_i, I^k \}, \quad (2.15) \\ k = 0, \dots, N-1, \end{aligned}$$

$$J_i(N, I^N) = E \{ x'(N)S(N)x(N) - 2r'x(N) \mid \theta_i, I^N \}. \quad (2.16)$$

Then the following is obvious,

$$\begin{aligned} J(k, I^k) = E \{ x'(k)Q(k)x(k) + u'(k)R(k)u(k) \mid I^k \} \\ = \sum_{i=1}^s q_i(k, I^k) J_i(k, I^k) \quad k = 0, 1, \dots, N-1 \end{aligned} \quad (2.17)$$

$$\begin{aligned} J(N, I^N) = E \{ x'(N)S(N)x(N) - 2r'x(N) \mid I^N \} \\ = \sum_{i=1}^s q_i(N, I^N) J_i(N, I^N). \end{aligned} \quad (2.18)$$

Since at stage k , all the posterior probabilities at later stages are unknown, a closed-loop optimal control cannot be computed analytically. Suppose that future

learning is suspended, then the open-loop feedback control can be obtained by fixing all the posterior probabilities at later stages at $q_i(k, I^k)$, $i = 1, 2, \dots, s$. As a result, the following optimal open-loop feedback control problem is considered at stage k ,

$$\begin{aligned} \min_{u(k)} \sum_{i=1}^s q_i(k, I^k) & \left\{ E \left\{ J_i(k, I^k) + \dots \right. \right. \\ & + \min_{u(N-2)} E \{ J_i(N-2, I^{N-2}) \\ & \left. \left. + \min_{u(N-1)} E [J_i(N-1, I^{N-1}) + J_i(N, I^N) | I^{N-1}] | I^{N-2} \} \dots | I^k \right\} \right\}. \quad (2.19) \end{aligned}$$

Define $\lambda = [\lambda_1, \dots, \lambda_s]' = [q_1(0, I^0), \dots, q_s(0, I^0)]'$. Thus the open-loop feedback control problem for $(A(r, \mu))$ at stage 0 is as follows:

$$\begin{aligned} (OFC(\lambda)) \min E & \left\{ \sum_{k=0}^N \left(\sum_{i=1}^s \lambda_i J_i(k) \right) \right\} \\ \text{s.t. } x_i(k+1) & = A_i(k)x_i(k) + B_i(k)u(k) + w(k), \\ & k = 0, 1, \dots, N-1, i = 1, 2, \dots, s \\ y_i(k) & = C_i(k)x_i(k) + v(k), \\ & k = 1, 2, \dots, N, i = 1, 2, \dots, s, \end{aligned}$$

where $A_i(k) = A(k, \theta_i)$, $B_i(k) = B(k, \theta_i)$, $C_i(k) = C(k, \theta_i)$, and $x_i(k)$ and $y_i(k)$ are the state and observation of the i th fictitious system, respectively, when assuming $\theta = \theta_i$. Note that as all the posterior probabilities in the later stages are fixed at $q_i(0, I^0)$, the optimal control to $(OFC(\lambda))$ is the optimal open-loop feedback control to problem $(A(r, \mu))$ at stage 0.

Problem $(OFC(\lambda))$ is a multiple-model formulation with $\lambda_i (= q_i(0, I^0))$, $i = 1, 2, \dots, s$ serving as the weighting coefficients. Let

$$\begin{aligned} X(k) & = [x'_1(k), x'_2(k), \dots, x'_s(k)]' \\ Y(k) & = [y'_1(k), y'_2(k), \dots, y'_s(k)]' \\ \bar{A}(k) & = \text{diag}(A_1(k), A_2(k), \dots, A_s(k)) \\ \bar{B}(k) & = [B'_1(k), B'_2(k), \dots, B'_s(k)]' \\ \bar{C}(k) & = \text{diag}(C_1(k), C_2(k), \dots, C_s(k)) \\ \bar{Q}(k) & = \text{diag}(\lambda_1 Q(k), \lambda_2 Q(k), \dots, \lambda_s Q(k)) \\ \bar{S}(N) & = \text{diag}(\lambda_1 S(N), \lambda_2 S(N), \dots, \lambda_s S(N)) \\ \bar{r} & = [\lambda_1 r', \lambda_2 r', \dots, \lambda_s r']' \end{aligned}$$

$$D_1 = [I_n, I_n, \dots, I_n]'$$

$$D_2 = [I_p, I_p, \dots, I_p]'$$

where *diag* denotes a block diagonal matrix. We thus obtain a compact form for the multi-model formulation,

$$\min E \left\{ X'(N) \bar{S}(N) X(N) + \sum_{k=0}^{N-1} [X'(k) \bar{Q}(k) X(k) + u'(k) R(k) u(k)] \right. \\ \left. - 2\bar{r}' X(N) \mid I^0 \right\} \quad (2.20)$$

$$\text{s.t. } X(k+1) = \bar{A}(k) X(k) + \bar{B}(k) u(k) + D_1 w(k) \quad (2.21)$$

$$k = 0, 1, \dots, N-1$$

$$Y(k) = \bar{C}(k) X(k) + D_2 v(k), k = 1, 2, \dots, N. \quad (2.22)$$

Define

$$\hat{X}(k) = [\hat{x}'_1(k|k), \hat{x}'_2(k|k), \dots, \hat{x}'_s(k|k)]'.$$

The solution to $(OFC(\lambda))$ can be obtained by using dynamic programming. We give the results in the following theorem.

Theorem 2 For a given r , the optimal control of the auxiliary problem $(OFC(\lambda))$ is

$$u^*(k) = -\Gamma_1(k) \hat{X}(k) + \Gamma_2(k) \bar{r} \quad (2.23)$$

where for $k = N-1, N-2, \dots, 1, 0$,

$$\bar{S}(k) = \bar{A}'(k) [\bar{S}(k+1) - T(k+1)] \bar{A}(k) + \bar{Q}(k) \quad (2.24)$$

$$T(k) = \Gamma_1'(k) \bar{B}'(k) [\bar{S}(k+1) - T(k+1)] \bar{A}(k) \quad (2.25)$$

$$G(k) = \bar{B}'(k) [\bar{S}(k+1) - T(k+1)] \bar{B}(k) + R(k) \quad (2.26)$$

$$\Gamma_1(k) = G(k)^{-1} \bar{B}'(k) [\bar{S}(k+1) - T(k+1)] \bar{A}(k) \quad (2.27)$$

$$\Gamma_2(k) = G(k)^{-1} \bar{B}'(k) L'(k+1) \quad (2.28)$$

$$L(k) = L(k+1) [\bar{A}(k) - \bar{B}(k) \Gamma_1(k)] \quad (2.29)$$

with the boundary conditions $T(N) = 0$ and $L(N) = I$.

Recall from Theorem 1 that the optimal open-loop feedback control to problem $(A(r, \mu))$ may also be the optimal open-loop feedback control to problem $(M_a(\mu))$

only when condition (2.14) is satisfied. The following theorem is given to show how to determine parameter r^* at stage 0. Define

$$\Phi(k) = I - \mu H(k) \left\{ \sum_{s=k+1}^{N-1} \prod_{i=s}^{N-1} [\bar{A}(i) - \bar{B}(i)\Gamma_1(i)] \bar{B}(s-1)\Gamma_2(s-1) + \bar{B}(N-1)\Gamma_2(N-1) \right\} H^T(k), \quad (2.30)$$

$$\Psi(k) = \mu H(k) \prod_{i=k}^{N-1} [\bar{A}(i) - \bar{B}(i)\Gamma_1(i)]. \quad (2.31)$$

where $H(k) = [q_1(k, I^k)I_n, q_2(k, I^k)I_n, \dots, q_s(k, I^k)I_n]$.

Theorem 3 *Assume that Φ is invertible. Then the optimal r^* with which the optimal open-loop feedback solution to $(A(r^*, \mu))$ also solves $(M_a(\mu))$ is equal to*

$$r^* = \Phi^{-1}(0)\Psi(0)\hat{X}(0). \quad (2.32)$$

Substitute (2.32) into the control law in (2.23), then the optimal open-loop feedback control of problem $(M_a(\mu))$ at stage 0, $u^*(0)$, can be obtained.

Proceeding to stage k , we can view stage k as the initial stage and $\hat{x}(k)$ as an estimate of the initial state when we consider a truncated dual control problem from stage k to stage N . Based on the principle of optimality and the concept of a rolling horizon, the optimal value of r should be equal to the following using the same derivation scheme as in Theorem 3,

$$r^* = \Phi^{-1}(k)\Psi(k)\hat{X}(k). \quad (2.33)$$

Substitute (2.33) into the control law (2.23), an optimal open-loop feedback control of problem $(M_a(\mu))$ at stage k , $u^*(k)$, can be obtained.

We have derived in the above discussion an optimal open-loop feedback control for problem $(M_a(\mu))$ with a fixed value of μ . The next natural question to be answered is how to determine the value of μ which represents a degree of importance of active learning. Entropy is a measure of uncertainty. Saridis (1988) and Tsai et al. (1992) studied the entropy formulation of optimal and adaptive control problems. We propose in our solution algorithm to assign the value of μ on-line at stage k to be proportional to the entropy of the probability distribution of θ at stage k , i.e.,

$$\mu \propto - \sum_{i=1}^s q_i(k) \ln q_i(k). \quad (2.34)$$

Conceptually, at the first few stages, since there exist parameter uncertainties, more effort will be put in active learning. As time involves, the value of μ will decrease. When the true parameter is identified, the entropy will be equal to zero such that the optimal solution of $(M_a(\mu))$ will converge to the optimal control of problem (P) . The proportional constant that relates the entropy to μ can be determined numerically. Notice that a too large proportional constant may result in a poor control performance due to too much effort was devoted to learning.

2.3.4 Optimal Nominal Dual Control

Minimizing a covariance term at the final stage provides a feature of active learning for the derived control law. The control law obtained, however, is not a closed-loop law but an optimal open-loop feedback control. Under this framework, the impact from the future learning has not been considered.

The key research issues are: (1) what is the best possible (partial) closed-loop control for (2.10), and (2) what is the active learning strategy to achieve this best possible outcome. A major difficulty in solving (2.10) is that the optimal control cannot be determined when the future posterior probabilities are unknown, while at the same time the future posterior probabilities depend on the control applied at the early stages. In order to break this loop, a possible solution scheme is to derive the relationship between the posterior probability and the control. A control which satisfies a deterministic version of this relationship is defined as the nominal control. The expected posterior probabilities when applying the nominal control are called nominal future posterior probabilities. Applying the nominal future posterior probabilities generated by the nominal control instead in (2.10), the effect of future learning can be taken into account. Since in this situation, all the achievable future information is used in terms of its expected value, the control law obtained can be considered to be the best possible closed-loop control law in this sense.

Assume that the current time is k and consider the truncated control problem from stage k to the end of the time horizon. For given $\lambda^t = [\lambda_1^t, \dots, \lambda_s^t]' \in \mathbb{R}_+^s$, $t = k, k+1, \dots, N$, with $\lambda^k = [q_1(k, I^k), q_2(k, I^k), \dots, q_s(k, I^k)]'$, consider the following optimal control problem,

$$\begin{aligned}
 (ONC(\lambda)) \min E & \left\{ \sum_{t=k}^N \left(\sum_{i=1}^s \lambda_i^t J_i(t, I^t) \right) \right\} \\
 \text{s.t. } x_i(t+1) &= A_i(t)x_i(t) + B_i(t)u(t) + w(t), \\
 & t = k, k+1, \dots, N-1, i = 1, 2, \dots, s \\
 y_i(t) &= C_i(t)x_i(t) + v(t), \\
 & t = k+1, k+2, \dots, N, i = 1, 2, \dots, s,
 \end{aligned}$$

where $A_i(t) = A(t, \theta_i)$, $B_i(t) = B(t, \theta_i)$, $C_i(t) = C(t, \theta_i)$, and $x_i(t)$ and $y_i(t)$ are the state and observation of the i th fictitious system, respectively, when assuming $\theta = \theta_i$.

Let

$$\begin{aligned} X(t) &= [x'_1(t), x'_2(t), \dots, x'_s(t)]' \\ Y(t) &= [y'_1(t), y'_2(t), \dots, y'_s(t)]' \\ \bar{A}(t) &= \text{diag}(A_1(t), A_2(t), \dots, A_s(t)) \\ \bar{B}(t) &= [B'_1(t), B'_2(t), \dots, B'_s(t)]' \\ \bar{C}(t) &= \text{diag}(C_1(t), C_2(t), \dots, C_s(t)) \\ \bar{Q}(t, \lambda) &= \text{diag}(\lambda^t_1 Q(t), \lambda^t_2 Q(t), \dots, \lambda^t_s Q(t)) \\ D_1 &= [I_n, I_n, \dots, I_n]' \\ D_2 &= [I_m, I_m, \dots, I_m]' \end{aligned}$$

where *diag* denotes a block diagonal matrix. We can obtain a compact form for the above multi-model formulation of $(ONC(\lambda))$ as follows:

$$\begin{aligned} \min E & \left\{ X'(N) \bar{Q}(N, \lambda) X(N) + \sum_{t=k}^{N-1} [X'(t) \bar{Q}(t, \lambda) X(t) + u'(t) R(t) u(t)] \mid I^k \right\} \\ \text{s.t. } & X(t+1) = \bar{A}(t) X(t) + \bar{B}(t) u(t) + D_1 w(t), \quad t = k, k+1, \dots, N-1 \\ & Y(t) = \bar{C}(t) X(t) + D_2 v(t), \quad t = k+1, k+2, \dots, N. \end{aligned}$$

Define

$$\hat{X}(t) = [\hat{x}'_1(t|t), \hat{x}'_2(t|t), \dots, \hat{x}'_s(t|t)]'.$$

The optimal solution to $(ONC(\lambda))$ can be derived by using dynamic programming,

$$u^*(t) = -\Gamma(t, \lambda) \hat{X}(t), \quad (2.35)$$

where for $t = k, k+1, \dots, N-1$

$$\Gamma(t, \lambda) = -G^{-1}(t, \lambda) \bar{B}'(t) S(t+1, \lambda) \bar{A}(t) \quad (2.36)$$

$$G(t, \lambda) = \bar{B}'(t) S(t+1, \lambda) \bar{B}(t) + R(t) \quad (2.37)$$

$$S(t, \lambda) = \bar{A}'(t) S(t+1, \lambda) \bar{A}(t) + \bar{Q}(t, \lambda) - \Gamma'(t, \lambda) G(t, \lambda) \Gamma(t, \lambda), \quad (2.38)$$

with the boundary condition $S(N, \lambda) = \bar{Q}(N, \lambda)$. Note that the optimal control, $\{u^*(t)\}_{t=k}^{N-1}$, is linear in the augmented state estimation $\hat{X}(t)$ and the feedback gain matrix Γ is nonlinear in λ .

At stage k , the true observation $y(k)$ is known, therefore $\hat{x}_i(k|k)$ can be obtained by the Kalman filter (1) to (5). Since future observations cannot be known in advance, a predicted nominal state trajectory $\{\hat{x}_i^*(t)\}_{t=k+1}^N$ and a predicted nominal observation trajectory $\{\hat{y}_i^*(t)\}_{t=k+1}^N$, can be calculated by setting all random variables at their expected values, i.e.

$$\hat{x}_i^*(t+1) = A_i(t)\hat{x}_i^*(t) + B_i(t)u^*(t), \quad t = k, k+1, \dots, N-1, \quad (2.39)$$

$$\hat{y}_i^*(t) = C_i(t)\hat{x}_i^*(t), \quad t = k+1, k+2, \dots, N, \quad (2.40)$$

with the initial condition $\hat{x}_i^*(k) = \hat{x}_i(k|k)$. For $t = k+1, k+2, \dots, N$, let

$$\hat{X}(t) = [\hat{x}_1^*(t)', \hat{x}_2^*(t)', \dots, \hat{x}_s^*(t)']'.$$

Substituting $\hat{X}(t)$ back into Eq. (2.35), we can close the loop and obtain a predicted nominal control.

Comparing problem $(ONC(\lambda))$ with the closed-loop control problem (2.10) at stage k , it is easy to recognize that if λ_i^t plays the same role as the posterior probabilities $q_i(t, I^t)$ at every stage, the optimal control of problem $(ONC(\lambda))$ is also optimal to problem (P) at stage k . However, those posterior probabilities at the later stages are unattainable. A feasible way is to use the nominal posterior probabilities generated by the nominal control instead. The control law achieved under this framework is referred to as the optimal nominal control to the original problem.

Define for $t = k+1, k+2, \dots, N$

$$\hat{y}^*(t) = \sum_{i=1}^s \lambda_i^t \hat{y}_i^*(t). \quad (2.41)$$

Using the Bayes formula, the predicted nominal posterior probability of mode i at stage k , $i = 1, 2, \dots, s$, satisfies the following recursive equation:

$$\tilde{q}_i(t) = \frac{L_i(t)}{\sum_{j=1}^s \tilde{q}_j(t-1)L_j(t)} \tilde{q}_i(t-1), \quad t = k+1, k+2, \dots, N, \quad (2.42)$$

with the initial condition $q_i(k, I^k)$, where $L_i(t)$ is still the same as given in (7) except

$$\tilde{y}(t|t-1, \theta_i) = \hat{y}^*(t) - \hat{y}_i^*(t). \quad (2.43)$$

It is clear that $\tilde{q}_i(t)$ is a function of $\lambda^k, \lambda^{k+1}, \dots, \lambda^N$.

In order to force the weighting coefficients λ_i^t to be equal to the nominal posterior probability $\tilde{q}_i(t)$ for all $t = k + 1, k + 2, \dots, N$, we construct the following optimization problem at stage k

$$\begin{aligned} \min \quad & \sum_{t=k+1}^N \sum_{i=1}^s (\lambda_i^t - \tilde{q}_i(t))^2 \\ \text{s.t.} \quad & \sum_{i=1}^s \lambda_i^t = 1, \text{ and all } \lambda_i^t \geq 0, t = k + 1, \dots, N. \end{aligned} \quad (2.44)$$

This is a nonlinear programming problem and can be solved by using general nonlinear programming solvers.

2.4 Dual Control in Big Data Era

In the Big Data era, massive amounts of information are generated every day. The high volume, high velocity, and high variety features of Big Data make capturing, managing, analyzing, storing, and retrieving information extremely challenging. In addition, the large-scale interconnected systems such as economic systems, power systems, manufacturing systems, health systems, water distribution systems, biological systems, etc. are complex and rapidly changing. It is not realistic and possible to develop mathematical models precisely to describe the system dynamics. Dual controls with probing features are advantageous in regulating these stochastic systems, especially in two situations: (1) when the time horizon is short and the initial estimates are poor, it is essential to stimulate the systems and rapidly find good estimates before reaching the end of the control horizon; (2) when the parameters of the process are changing very rapidly (Wittenmark 1975a). Some successful applications of dual control are summarized as below.

2.4.1 Economic Systems

Most economic problems are stochastic. There is uncertainty about the present state of the system, uncertainty about the response of the system to policy measures, and uncertainty about future events. For example, in macroeconomics some time series are known to contain more noise than others. Also, policy makers are uncertain about the magnitude and timing of responses to changes in tax rates, government spending, and interest rates. In international commodity stabilization, there is uncertainty about the effects of price changes on consumption (Kendrick 1981). Because of the short time horizon and highly stochastic nature of the parameters in the economic processes, dual controls have been seen in solving economic systems (Bar-Shalom and Wall 1980; Kendrick 1981). Kendrick demonstrated examples of using dual control to solve MacRae problem and a macroeconomic model with measurement error (Kendrick 1981).

2.4.2 Manufacturing Processes

Dual control is also successfully applied in manufacturing processes. The grinding processes in the pulp industry (Allison 1994), where the parameters are changing fairly rapidly and the gain is also changing sign, is probably the first application of dual control to process control. The controller is an active adaptive controller, which consists of a constrained certainty equivalence approach coupled with an extended output horizon and a cost function modification to get probing (Wittenmark 1975a).

Another application of dual control in capital intensive semiconductor manufacturing processes has been seen in Arda Vanli et al. (2011). In such processes, it is often impractical to run large designed experiments and the amount of experimental data available is often not adequate to build sufficiently accurate statistical models or reliably estimating optimal conditions. A dual control approach that simultaneously considers model estimation and optimization objectives is adopted and an adaptive Bayesian response surface model is used. It is shown that by employing the proposed adaptive Bayesian approach one can simultaneously learn the process while not requiring excessive perturbations away from the target level and can achieve faster model estimation than central composite experimental designs.

2.4.3 Automobile Systems

A driver assistance system with a dual control scheme was developed in Saito et al. (2016), which can effectively identify drivers' drowsiness and prevent sleep-related vehicle accidents. The dual control has two purposes: (1) to effect the partial control initiated by the assistance system, preventing lane departure, and (2) enabling the assistance system to judge, through the interaction between the driver and the assistance system, whether the driver recognizes that the vehicle is going to deviate from the lane. The assistance system implements partial control in the event of lane departure and gives the driver the chance to voluntarily take the action needed. If the driver fails to implement the steering action needed within a limited time, the assistance system judges that "the driver's understanding of the given situation is incorrect" and executes the remaining control.

2.4.4 Robotics

Adaptive dual control using neural networks has also been extensively investigated. Neural networks have been used to approximate the unknown functions in the system dynamics of the nonlinear stochastic systems. Such dual control was successfully applied to kinematic control of nonholonomic mobile robots in which the robot dynamic functions are nonlinear with varying uncertain/unknown parameters

(Bugeja et al. 2009). Two schemes are developed in discrete time, and the robot's nonlinear dynamic functions are assumed to be unknown. The Gaussian radial basis function and sigmoidal multilayer perception neural networks are used for function approximation. In each scheme, the unknown network parameters are estimated stochastically in real time, and no preliminary offline neural network training is used. In contrast to other adaptive techniques hitherto proposed in the literature on mobile robots, the dual control laws do not rely on the heuristic certainty equivalence property but account for the uncertainty in the estimates. This results in a major improvement in tracking performance, despite the plant uncertainty and unmodeled dynamics.

2.4.5 Information Retrieval

An Information Retrieval (IR) system consists of a collection of documents and an engine that retrieves documents described by user queries. In large systems, such as the Web, queries are typically too vague, hence an iterative process in which the users refine their queries gradually has to take place. An active learning approach was proposed in Jaakkola and Siegelmann (2001) to reduce the IR users dissatisfactions due to long, tedious repetitive search sessions. The system responds to the initial user's query by successively probing the user for distinctions at multiple levels of abstraction. The system's initiated queries are optimized for speedy recovery and the user is permitted to respond with multiple selections or may reject the query. The information is in each case unambiguously incorporated by the system and the subsequent queries are adjusted to minimize the need for further exchanges. More applications in information retrieval and image retrieval can be seen in Zhang and Chen (2002) and Dagli et al. (2005).

2.5 Conclusions

This overview presents the dual control methods, elaborated from the Feldbaum's seminal work in the 1960s until present. The author and collaborators' research on dual control for a class of discrete-time linear quadratic Gaussian problems with parameter uncertainty in both state and observation equations is summarized to demonstrate different control laws. It is shown that minimizing a covariance term at the final stage introduces a feature of active learning for the derived control law. By exploring the future nominal posterior probabilities, the control law takes into account the function of future learning, thus the best possible closed-loop feedback control can be achieved. Successful applications of dual controls in various areas indicate although cautious, the controller with the probing/active learning feature can help reduce system uncertainties and hence it performs better than the controller with passive or without learning ability.

References

- B.J. Allison, Dual adaptive control of a chip refiner motor load. Ph.D. Thesis, University of British Columbia, 1994
- J. Alster, P.R. Belanger, A technique for dual adaptive control. *Automatica* **10**, 627 (1974)
- J. Alster, P.R.B. Langer, A technique for dual adaptive control. *Automatica*, **10**, 627–634 (1974)
- M. Aoki, *Optimization of Stochastic Systems* (Academic, New York 1967)
- O. Arda Vanli, C. Zhang, B. Wang, An adaptive Bayesian method for semiconductor manufacturing process control with small experimental data sets. *IEEE Trans. Semicond. Manuf.* **24**(3), 418–431 (2011)
- K.J. Astrom, *Introduction to Stochastic Control Theory* (Academic Press, New York, 1970)
- K.J. Astrom, B. Wittenmark, Problems of identification and control. *J. Math. Analysis Appl.* **34**, 90 (1971)
- K.J. Astrom, B. Wittenmark, *Adaptive Control* (Addison Wesley, Reading, 1989)
- Y. Bar-Shalom, Stochastic dynamic programming: caution and probing. *IEEE Trans. Autom. Control* **AC-26**, 1184–1194 (1981)
- Y. Bar-Shalom, R. Sivan, On the optimal control of discrete-time linear systems with random parameters. *IEEE Trans. Autom. Control* **AC-14**, 3–8 (1969)
- Y. Bar-Shalom, E. Tse, Dual effect, certainty equivalence, and separation in stochastic control. *IEEE Trans. Autom. Control* **AC-19**(5), 494–500 (1974)
- Y. Bar-Shalom, K.D. Wall, Dual adaptive control and uncertainty effects in macroeconomic systems optimization. *Automatica* **16**, 147–156 (1980)
- M.K. Bugeja, S.G. Fabri, L. Camilleri, Dual adaptive dynamic control of mobile robots using neural networks. *IEEE Trans. Syst. Man Cybern.* **39**(1), 129–141 (2009)
- F. Casiello, K.A. Loparo, Optimal policies for passive learning controllers. *Automatica* **25**, 757–763 (1989)
- C.K. Dagli, S. Rajaram, T.S. Huang, Combining diversity-based active learning with discriminant analysis in image retrieval, in *Third International Conference on Information Technology and Applications*, vol. 2 (2005), pp. 173–178
- J.G. Deshpande, T.N. Upadhyay, D.G. Lainiotis, Adaptive control of linear stochastic systems. *Automatica* **9**, 107–115 (1973)
- A.A. Feldbaum, *Optimal Control Systems* (Academic, New York, 1965)
- N.M. Filatov, H. Unbehauen, Survey of adaptive dual control methods. *IEE Proc. Control Theory Appl.* **147**, 118–128 (2000)
- J.J. Florentine, Optimal probing, adaptive control of a simple Bayesian system. *J. Electronics Control* **13**, 165 (1962)
- P. Fu, D. Li, F. Qian, Active dual control for linear-quadratic Gaussian system with unknown parameters, in *Proceedings of the 15th World Congress of IFAC*, Barcelona, Spain, July 21–16 (2002)
- D. Gorman, J. Zaborszky, Stochastic optimal control of continuous time systems with unknown gain. *IEEE Trans. Autom. Control* **AC-13**, 630–638 (1968)
- A.J. Grammaticos, B.M. Horowitz, The optimal adaptive control law for a linear plant with unknown input gains. *Int. J. Control* **12**, 337 (1970)
- B.E. Griffiths, K.A. Loparo, Optimal control of jump linear systems. *Int. J. Control.* **42**(4), 791–819 (1985)
- R. Hasting-James, M.W. Sage, Recursive generalized least squares procedure for on-line identification of process parameters. *IEE Proc.* **116**, 2057 (1969)
- D.J. Hughes, O.L.R. Jacobs, Turn-off, escape and probing in nonlinear stochastic control, in *Proceedings of IFAC Symposium Adaptive Control*, Budapest, Hungary, Sept. 1974
- T. Jaakkola, H. Siegelmann, Active information retrieval. *Adv. Neural Inf. Proces. Syst.* **14**, 777–784 (2001)
- D.H. Jacobson, Optimal stochastic linear systems with exponential performance criteria and their relation to deterministic differential games. *IEEE Trans. Autom. Control* **18**, 124–131 (1973)

- O.L.R. Jacobs, S.M. Langdon, An optimal extremum control system. *Automatica* **6**, 297 (1970)
- K. Jenkins, R. Roy, A design procedure for discrete adaptive control systems, Preprints JACC (1966), p. 624
- R.E. Kalman, Design of self-optimizing control system, *Trans. Am. Soc. Mech. Eng.* **80**, 468 (1958)
- D.A. Kendrick, *Stochastic Control for Economic Models* (McGraw-Hill, New York, 1981)
- D. Lainiotis, J.G. Deshpande, T.N. Upadhyay, Optimal adaptive control: a nonlinear separation theorem. *Int. J. Control* **15**, 877–888 (1972)
- D. Li, F. Qian, P. Fu, Variance minimization approach for a class of dual control problems. *IEEE Trans. Autom. Control* **AC-47** 2010–2020 (2002)
- D. Li, F. Qian, P. Fu, Optimal nominal dual control for discrete-time linear-quadratic Gaussian problems with unknown parameters. *Automatica* **44**, 119–127 (2008)
- B. Lindoff, J. Holst, B. Wittenmark, Analysis of approximations of dual control. *Int. J. Adapt. Control Signal Process.* **13**, 593–620 (1999)
- J.C. Luxat, L.H. Lees, Suboptimal control of a class of nonlinear systems. *Int. J. Control* **17**, 965 (1973)
- R. Milito, C. Padilla, R. Padilla, D. Cardorin, An innovations approach to dual control. *IEEE Trans. Autom. Control* **AC-27**, 132–137 (1982)
- V. Panuska, A stochastic approximative method for identification of linear systems using adaptive filter, in *Preprint Joint Automatic Control Conference* (1968)
- Y. Saito, M. Itoh, T. Inagaki, Driver assistance system with a dual control scheme: effectiveness of identifying driver drowsiness and preventing lane departure accidents. *IEEE Trans. Hum. Mach. Syst.* **46**(5), 660–671 (2016)
- G.N. Saridis, Entropy formulation of optimal and adaptive control. *IEEE Trans. Autom. Control* **AC-33**(8), 713–721 (1988)
- J. Sternby, A regulator for time-varying stochastic systems, in *Proceedings of 7th IFAC World Congress*, Helsinki, Finland, June 1978
- Y.A. Tsai, F.A. Casiello, K.A. Loparo, Discrete-time entropy formulation of optimal and adaptive control problems. *IEEE Trans. Autom. Control* **AC-37**(7), 1083–1088 (1992)
- E. Tse, M. Athans, Adaptive stochastic control for a class of linear systems. *IEEE Trans. Autom. Control* **AC-17**, 38–52 (1972)
- E. Tse, Y. Bar-Shalom, L. Meier, Wide-sense adaptive dual control for nonlinear stochastic systems. *IEEE Trans. Autom. Control* **AC-18**, 98–108 (1973)
- B. Wittenmark, Stochastic adaptive control methods: a survey. *Int. J. Control* **21**, 705–730 (1975a)
- B. Wittenmark, An active suboptimal dual controller for systems with stochastic parameters. *Autom. Control Theory Appl.* **3**, 13–19 (1975b)
- B. Wittenmark, Stochastic adaptive control methods: a survey. *Int. J. Control* **21**(5), 705–730 (1975c)
- J. Wieslander, B. Wittenmark, An approach to adaptive control using real-time identification. *Automatica* **7**, 211 (1971)
- P.C. Young, The use of linear regression and related procedure for the identification of dynamic processes. in *Proceedings of Seventh I.E.E.E. Symposium on Adaptive Processes* (1968)
- C. Zhang, T. Chen, An active learning framework for content-based information retrieval. *IEEE Trans. Multimedia* **4**(2), 260–268 (2002)