Extremal control of Wiener processes

Wittenmark, Björn; Evans, Robin

Published in:
IEEE Xplore

DOI:
10.1109/CDC.2002.1185109
10.1109/CDC.2002.1185109

2002

Proceedings of the 41st IEEE
Conference on Decision and Control
Las Vegas, Nevada USA, December 2002

Extremal control of Wiener model processes

Björn Wittenmark
email: bjorn@control.lth.se
Lund Institute of Technology
Box 118, SE-221 00 Lund, Sweden

Robin J. Evans
email: r.evans@ee.mu.oz.au
University of Melbourne
Melbourne, VIC 3010, Australia

Abstract

Extremal control of Wiener type processes is considered. These models consists of a linear part followed by a static nonlinearity. We will consider nonlinearities having one extremum point. The purpose is to keep the output of the process as close as possible to the extremum point. The main problem in the control of this kind of processes is the non-uniqueness of the inverse of the nonlinearity. This causes problems, e.g., in the estimation of the states of the process and the identification in the adaptive case. An one-step-ahead controller combined with a probabilistic estimator is proposed and analyzed.

1. Introduction

There are many application where it is of interest to position the process output at an optimum or extremum point. A typical situation is combustion engines where the emission and the efficiency depend on the inputs to the motor such as fuel and air/fuel ratio. Other examples are control of grinding processes, water turbines, and wind mills where it is of interest to operate the system as close as possible to the extremum point.

This problem has been studied over a long period of time and there are solutions to some of the problems occurring in extremum control. See, for instance, Draper and Li (1951), Jacobs and Langdon (1970), Keviczsky and Haber (1974), Keviczsky *et al.* (1979), Sternby (1980a), Sternby (1980b), Dumont and Åström (1988), Wellstead and Scotson (1990), Scotson and Wellstead (1990), Wittenmark (1993), Allison (1994), Wittenmark and Urquhart (1995), Navarro and Zarrop (1996), Krstić and Wang (1997), and Krstić and Wang (2000). In many of the earlier references the static optimization problem has mainly been discussed.

The problem of extremum control can be approached in several ways. Among the first approaches was the introduction of perturbation signals. A perturbation signal is then used to get information about the local gradient of the nonlinearity. See Sternby (1980b) and Krstić and Wang (2000). The perturbation signal method is usually only used to find a constant value of the input and/or to be able to follow a varying operating point. The process will then behave as an open loop system around the extremum point. The perturbation signal method has the advantage that it requires very little information about the process. On the other hand the convergence of the system and the steady state performance are not very good, especially in presence of noise.

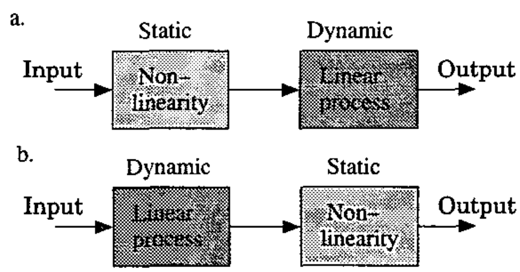

Figure 1: a. Hammerstein model, b. Wiener model

A second approach is to use more advanced optimization methods. If the nonlinearity is a known function the optimal constant input might be computed directly. The drawback is that the static nonlinearity and the open loop gain of the process have to be known. Further, the performance at the extremum point is still as if it were an open loop system. This implies that if the open loop dynamics of the process is slow then the convergence and the recovery after a disturbance will be slow.

There are different classifications of nonlinear systems. Two different classes of systems are shown in Figure 1. The first class of models is called *Hammerstein models* where the nonlinearity is at the input of a linear dynamic subsystem. In Wittenmark (1993) and Wittenmark and Urquhart (1995) different dynamic controllers for the Hammerstein model together with adaptive schemes are investigated. The Hammerstein models have the advantage that the models are linear in the parameters, which makes it easy to estimate the parameters of the model.

In the second class of models, *Wiener models*, the system has a linear part followed by a nonlinearity. In this paper we will discuss extremum control of Wiener models. To start with we will assume that the processes are known and the parameter estimation problem will only be briefly discussed at the end of the paper.

An one-step-ahead controller based on the separation prin-

0-7803-7516-5/02/$17.00 ©2002 IEEE

ciple together with a probabilistic estimator of the linear output is discussed and analyzed in this paper. The paper is organized in the following way. The problem is formulated in Section 2 and different controllers are discussed in Section 3. Some of the controllers require that the output or the states of the linear part of the process are known and different ways to make estimators are discussed in Section 4. An example is used in Section 5 to illustrate the behavior of the different control schemes. Section 6 contains a discussion of estimating the unknown parameters of Wiener models. Finally, some conclusions are given in Section 7.

2. Problem formulation

We assume that the process is a Wiener model where the linear part is described by the known discrete-time system

$$\begin{aligned} z(k) + a_1 z(k-1) + a_2 z(k-2) + \cdots + a_n z(k-n) \\ = b_0 u(k-d) + \cdots + b_{n-d+1} u(k-n-1) \\ + e(k) + c_1 e(k-1) + \cdots + c_n e(k-n) \end{aligned} \tag{1}$$

where $u(k)$ is the input signal, $z(k)$ the output of the linear part, and $e(k)$ is Gaussian distributed white noise with zero mean and standard deviation σ. The model can also be written in polynomial form

$$A(q)z(k) = B(q)u(k) + C(q)e(k) \tag{2}$$

where q is the forward shift operator and $\deg A = \deg C = n$ and $\deg B = n - d$. Further, A and C are monic, i.e. the coefficient of the largest power of q is equal to one. The parameter d is the time delay in the system. The nonlinearity is described as a quadratic function of the form

$$y(k) = h(z(k)) = \gamma_0 + \gamma_1 z(k) + \gamma_2 z(k)^2$$

with $\gamma_2 \neq 0$. Other types of nonlinearities can also be assumed. However, we assume, at least close the optimum point, that the nonlinearity can be described by a quadratic function. For simplicity we assume that the extremum point is a minimum, i.e. $\gamma_2 > 0$. The minimum of $y(k)$ is obtained for

$$z(k) = z_0 = -\frac{\gamma_1}{2\gamma_2}$$

Independent of the value of $z(k)$ the output can never be below the value y_0, where the minimum of $y(k)$ is

$$y_0 = \gamma_0 - \frac{\gamma_1^2}{4\gamma_2}$$

The control signal, $u(k)$ is allowed to be a function of the process output $y(k)$ and previous inputs and outputs. In the derivation of some of the controllers we will also assume that the control signal may be a function of the outputs of the linear system or its state, i.e. of $z(j)$, $j \leq k$. The estimation of the states is discussed in Section 4. The purpose of the control is to keep the output $y(k)$ as close as possible to the

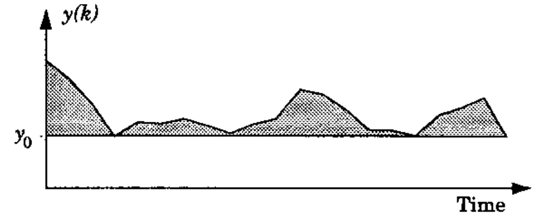

Figure 2: The purpose of the control of the Wiener model is to minimize the indicated area.

optimum point y_0. The loss function is formally expressed as

$$\min_{u(k)} \mathrm{E}(y(k) - y_0)$$

i.e. we want to minimize the area indicated in Figure 2.

3. Control strategies

Three different control strategies will now be discussed.

3.1 Static controller

Assume that there is no noise acting on the system and that the input to (2) is constant then $z(k) = z_0$ if

$$u_0 = \frac{A(1)}{B(1)} z_0 = -\frac{\gamma_1 A(1)}{2\gamma_2 B(1)} \tag{3}$$

The controller (3) can be regarded as a one step minimization of the quadratic nonlinearity opposed to the use of the gradient method when using a perturbation signal. Using (3) on the system (2) gives after a transient

$$z(k) = \frac{A(1)B(q)}{B(1)A(q)} z_0 + \frac{C(q)}{A(q)} e(k) = z_0 + v(k)$$

The second equality follows since z_0 is constant. This implies that the mean value of z is equal to z_0 but the variation around z_0 is determined by the open loop noise dynamics C/A. The output y will thus deviate from the desired value y_0. The variable z is a Gaussian process but the output y is a non-central χ^2 distribution. If the open loop system has slow dynamics then the convergence of z will be slow at the startup or after the noise process has driven z away from its desired value.

Using (3) gives

$$\begin{aligned} y(k) &= \gamma_0 + \gamma_1(z_0 + v(k)) + \gamma_2(z_0 + v(k))^2 \\ &= y_0 + (\gamma_1 + 2\gamma_2 z_0)v(k) + \gamma_2 v(k)^2 \end{aligned}$$

This implies that

$$\mathrm{E}(y(k) - y_0) = \gamma_2 \sigma_v^2$$

where σ_v^2 is the variance of the process $v(k)$, which is the same as the open loop variance of the process, i.e. the controller gives the correct mean value, but the stochastic part of the system is not influenced.

3.2 Prediction using the true linear output

Assume that $z(k)$ is measurable. One way to obtain a good control of the system is to minimize the variance of z around the value z_0. With z available this is essentially the problem of predicting $z(k+d)$ where d is the time delay of the system, see Åström and Wittenmark (1997). To make the prediction we introduce the identity

$$q^d C(q) = A(q)F(q) + G(q)$$

where F is monic and $\deg F = d$ and $\deg G = n-1$ and d is the prediction horizon. Then

$$z(k+d) = F(q)e(k) + \frac{1}{C(q)}\left(B(q)F(q)u(k) + G(q)z(k)\right) \tag{4}$$

The controller that minimizes the variance of $z(k)$ around z_0 is given by

$$u(k) = -\frac{G(q)}{B(q)F(q)}z(k) + \frac{C(1)}{B(1)F(1)}z_0 \tag{5}$$

The controller (5) will keep z as close as possible to z_0 and will also make $y(k)$ close to its optimal value y_0. In the case when $d=1$ then $F=1$ and $G=C-A$ and the controller (5) becomes

$$\begin{aligned} u(k) &= \frac{A(q)-C(q)}{B(q)}z(k) + \frac{C(1)}{B(1)}z_0 \\ &= \frac{A(q)-C(q)}{B(q)}z(k) - \frac{\gamma_1 C(1)}{2\gamma_2 B(1)} \end{aligned}$$

Using (5) gives

$$\begin{aligned} z(k) &= z_0 + F(q)e(k-d+1) \\ &= z_0 + e(k) + f_1 e(k-1) + \cdots + f_{d-1}e(k-d+1) \end{aligned}$$

which gives the output

$$\begin{aligned} y(k) = {} & \gamma_0 + \gamma_1 z_0 + \gamma_2 z_0^2 + \gamma_1 F(q)e(k-d+1) \\ & + 2\gamma_2 z_0 F(q)e(k-d+1) \\ & + \gamma_2 (F(q)e(k-d+1))^2 \end{aligned}$$

Further

$$\mathrm{E}(y(k) - y_0) = \gamma_2(1 + f_1^2 + \cdots + f_{d-1}^2)\sigma^2$$

Since $F(q)$ is the first $d-1$ coefficients of the series expansion of $C(q)/A(q)$ it follows that the average loss per step is lower when (5) is used than when (3) is used.

3.3 Prediction using the estimated linear output

The controller (5) is an idealized controller since the output (or the state) of the linear part cannot be measured. An obvious modification is to assume that the separation principle holds and replace $z(k)$ with the estimated value $\hat{z}(k)$, i.e. to use a certainty equivalence controller. The estimation of $z(k)$ is assumed to be based on measurements of $y(k)$ and previous values of the measured outputs and inputs. The control law based on the estimated linear output is

$$u(k) = -\frac{G(q)}{B(q)F(q)}\hat{z}(k) + \frac{C(1)}{B(1)F(1)}z_0 \tag{6}$$

There are obvious difficulties in the estimation of z, especially since the nonlinear part of the system has a non-unique inverse. See Section 4.

3.4 An example

A simple example is used to illustrate the problem formulation and the different controllers. Assume that the process is described by

$$z(k) + az(k-1) = bu(k-1) + e(k)$$

The output nonlinearity is given by

$$y(k) = \gamma_0 + \gamma_1 z(k) + \gamma_2 z(k)^2$$

The controller (5) is

$$u(k) = \frac{2\gamma_2 a z(k) - \gamma_1}{2\gamma_2 b}$$

This controller gives the expected loss per step equal to

$$V_{pred} = \gamma_2 \sigma^2 \tag{7}$$

The constant controller (3) is

$$u_0 = -\frac{\gamma_1(1+a)}{2\gamma_2 b}$$

Assuming that the linear part is stable, i.e. $|a| < 1$, then the expected loss per step using this controller is

$$V_{const} = \frac{\gamma_2\sigma^2}{(1-a^2)} \tag{8}$$

which is larger than (7).

4. The state estimation problem

The controller (6) requires an estimate of the output of the linear part of the process based on the output from the nonlinearity. This constitutes a nonlinear estimation problem. There are different ways of approaching this problem. One approach is to use the extended Kalman filter. Another way is to try to utilize the structure of the process and approach the estimation of $z(k)$ in some other ways.

The first obvious choice for the estimation would be the extended Kalman filter or some higher order nonlinear filter. All these filters are based on series expansions around some nominal point, which in this case should be the extremum point. The extended Kalman filter will then have zero gain at the interesting point and some other approach is required to obtain a good estimator.

4.1 Probabilistic estimator of the linear output

A new estimator will now be derived by fully utilizing the structure of the process. Since there is not any measurement noise in $y(k)$ in (1) and the nonlinearity is quadratic we can solve for the value of $z(k)$ using $y(k)$. The two possible solutions for the output of the linear system is

$$z_{1,2}(k) = \frac{1}{2\gamma_2}\left(-\gamma_1 \pm \sqrt{\gamma_1^2 + 4\gamma_2(y(k) - \gamma_0)}\right)$$

Using (4) we introduce the following one-step-ahead predictor for the output of the linear system

$$\hat{z}(k|k-1) = \frac{B(q)}{C(q)}u(k-1) + \frac{G(q)}{C(q)}\hat{z}(k-1) \tag{9}$$

Notice that this is a valid predictor for arbitrary d. We now assume that we have an estimate of the probabilities $p_1(k-1)$ and $p_2(k-1)$ that the linear output at time $k-1$ are $z_1(k-1)$ and $z_2(k-1)$, respectively. An algorithm to obtain the estimate at time k is the following

1. Use $y(k-1)$ to compute the two possible values of z at time $k-1$ giving $z_1(k-1)$ and $z_2(k-1)$.
2. Predict one step ahead using (9) with $\hat{z}(k-1) = \hat{z}_i(k-1)$. (For previous values of $\hat{z}(k-1)$ the previous estimates are used.) This results in $\hat{z}_1(k|k-1)$ and $\hat{z}_2(k|k-1)$, respectively. The prediction error has the normal distributed frequency function $f_n(z)$
3. The probabilities for being in the two states can now be updated using the following equations

$$\begin{aligned} p_1(k) &= \alpha_{norm}[p_1(k-1)f_n(z_1(k) - \hat{z}_1(k|k-1)) \\ &\quad + p_2(k-1)f_n(z_1(k) - \hat{z}_2(k|k-1))] \\ p_2(k) &= \alpha_{norm}[p_1(k-1)f_n(z_2(k) - \hat{z}_2(k|k-1)) \\ &\quad + p_2(k-1)f_n(z_2(k) - \hat{z}_1(k|k-1))] \end{aligned} \tag{10}$$

   where α_{norm} is a normalization factor making the sum of $p_1(k)$ and $p_2(k)$ equal to one.
4. The estimate $\hat{z}(k)$ is chosen as the $z_i(k)$ which has the largest probability $p_i(k)$.

The probability based estimator uses the previous measurement of y to compute two possible values of z. These two outputs of the linear system are then predicted into the future using a slightly modified predictor. The two predicted values represent two possible outcomes for z at time k and these are compared with the two possibilities that are obtained from the present measurement. The probabilities for the two outcomes are calculated and the one of the possible present values of $z_i(k)$ that has the highest probability is chosen as the estimate of z at time k. The estimator has the advantage that it cannot be unstable. It is only selecting one of two possible values. There is, however, a possibility that the estimator may sometimes choose the wrong value. An estimator of the form (10) was discussed in Jacobs and Langdon (1970) and Sternby (1980a).

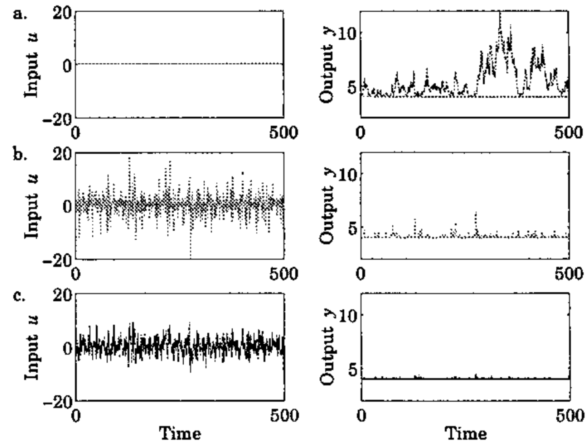

Figure 3: The left diagrams show the input u and the right diagrams the output y together with the minimum value y_0 when using different controllers: a. (3); b. (6); and c. (5).

It is straightforward to show that $p_1(k) = p_2(k) = 0.5$ independent of $p_1(k-1)$ and $p_2(k-1)$ when the constant controller (3) is used. This implies that the constant controller doesn't give any information about which of the values $z_i(k)$ to choose. Another possibility to determine the estimate is to use the mean value, i.e.

$$\hat{z}(k) = p_1(k)z_1(k) + p_2(k)z_2(k) \tag{11}$$

A controller based on this estimate will move the control signal towards u_0, which implies that there will be less excitation of the process compared when the largest probability is used for the decision. Compare the discussion in Section 5.

5. An example

The example in Section 3.4 will be used to illustrate the properties of the different controllers. The following numerical values will be used in the simulations: $a = -0.99$, $b = 0.1$, $\sigma^2 = 0.1$, $\gamma_0 = 12$, $\gamma_1 = -4$, and $\gamma_2 = 0.5$. The system is simulated using the the constant controller (3), one-step-ahead prediction controller using the true linear output (5), and the one-step-ahead prediction controller using the estimated linear output (6) when the estimation is carried out using the probabilistic estimator in Section 4.1.

Figure 3 shows the input u and the output y when using the three controllers. It is clearly seen that output y deviates much more from the optimum y_0 when the constant controller is used compared to the other two controllers. This is also seen in Figure 4 showing the accumulated loss function

$$V(k) = \sum_{i=1}^{k} (y(k) - y_0)$$

when the different controllers are used. The loss when using (5) corresponds well with the theoretical loss per step (7), while the loss when using the constant controller is less in the simulation than what is predicted by (8). Longer simulations are required to obtain a better agreement with the theoretical loss per step since the process in this case has a very low frequency behavior. (The pole of the open loop system is close to the unit circle.) The use of the estimated linear output gives a loss that is 2.3 times larger, while the constant controller gives a loss that is 28 times larger than the loss when using feedback from the true linear output. Simulations also verify that using the mean value estimator (11) gives a much worse performance than when using the most probable value given by (11). The accumulated loss is then 12.5 times the loss when using the feedback from the true linear output. The properties of the estimator is seen in Figure 5. The curves show the estimated and true linear output when the controller (6) is used. There is a quite good estimation of the true linear output but the estimator picks the wrong solution in about 35% of the cases. This is due to the noise and that the process is operating close to the optimum where it is difficult for the estimator to distinguish between the influences of the noise and of the control signal.

The special case $a = -1$, $b = 1$, $\gamma_0 = \gamma_1 = 0$, and $\gamma_2 = 1$ is discussed in Jacobs and Langdon (1970) and Sternby (1980a). The optimal controller is numerically derived in Jacobs and Langdon (1970) by using dynamic programming. The average loss per step is shown to be $2.2\sigma^2$.

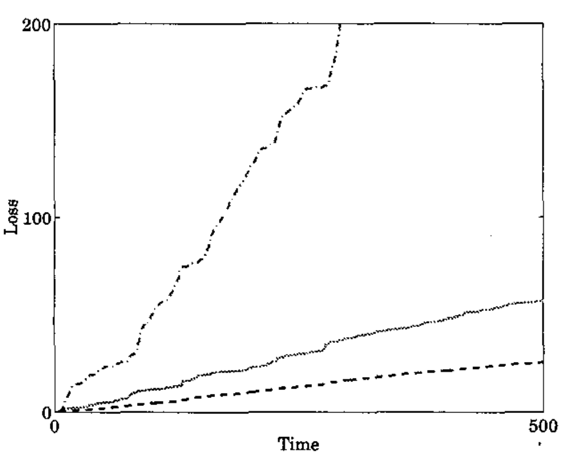

Figure 4: The accumulated loss function when using the three controllers (3) (dash-dotted), (5) (dashed), and (6) (full).

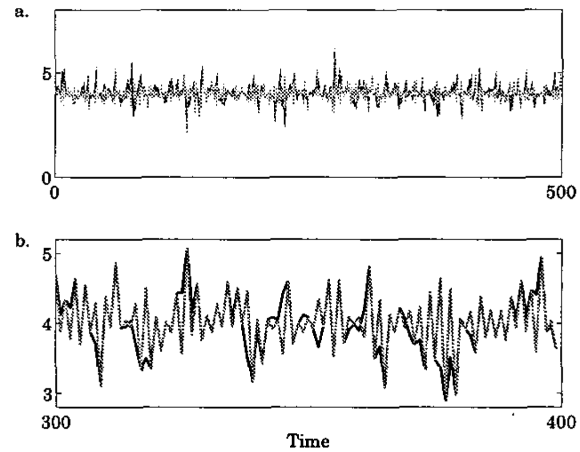

Figure 5: The true (black) and estimated state z (grey) when using the controller (6). Curve b. show an enlarged part of curve a.

A suboptimal controller controller based on an approximate least squares estimation of the linear output and a two-step loss function is proposed in Sternby (1980a) and gives an average loss $2.64\sigma^2$ per step. The proposed probabilistic estimator and the certainty equivalence controller suggested in this paper gives a loss of $2.42\sigma^2$ per step, which is 10% larger than the optimal loss, but better than earlier proposed suboptimal controllers. The improvement is probably due to the perturbation introduced by choosing the most probable value of $z_i(k)$.

6. Unknown process parameters

In the analysis and simulations so far we have assumed that the process and its parameters are fully known. In practical applications this is not the case. In many situations the parameters of the process will change depending on the environment in which the process is working. This is the case, for instance, for combustion engines.

One way to circumvent the lack of knowledge about the process is to use a method that does not depend on the process parameters. The use of a perturbation signal only relies on the assumption that the nonlinearity is a concave or convex function. The input signal is changed based on the estimate of the gradient. The dynamics of the process and the noise, however, have a heavy influence on the estimation of the gradient. The phase lag introduced by the process dynamics can be compensated for as suggested in Sternby (1980b), Krstić and Wang (1997), and Krstić and Wang (2000).

An alternative way is to estimate the parameters of the process on-line or off-line. The estimation of the parameters of the Wiener models is discussed, for instance, in Wigren (1993), Boutayeb and Darouach (1995), Zhu (1999), and Hagenblad and Ljung (2000). Several of the estimation methods are essentially based on the idea of the extended Kalman filter, which is not suitable for extremum seeking control. Boutayeb and Darouach (1995) and Zhu (1999) consider the more difficult case with a non-unique inverse of the nonlinearity that is discussed in this report.

7. Summary

Extremum control of Wiener model processes has been discussed. For known processes there are several possibilities to obtain good control of the process. A crucial part of the controller is the estimation of the output of the linear part of the process. Several types of estimators have been discussed and most of the estimators have the drawback that they have a singular point at the optimum point of the process and this is where we want to keep the process. The method proposed to avoid this problem is to use a probabilistic based estimator that selects between two possible values of the output based on previous measurements and input signals. The combination of this estimator and a prediction controller has the advantage that especially close to the optimum it is insensitive to the accuracy of the estimates.

The controllers discussed have been based on the assumption that the separation principle is valid, which implies that the true linear output can be replaced by its estimate. The behavior when using the constant controller also indicates that perturbation signals should be introduced to improve the performance of the closed loop system. This implies that the controller should have a dual property, which ensures that that the control action is a compromise between making good control and obtaining good estimates of the linear output. Dual control is discussed in, for instance, Åström and Wittenmark (1995) and Filatov and Unbehauen (2000).

Acknowledgments

The research was mainly done while the first author was visiting University of Melbourne with support from CSSIP, Australia. Support for the research has also been obtained from the Royal Physiographic Society, Sweden.

8. References

Allison, B. J. (1994): *Dual adaptive control of a chip refiner motor load*. PhD thesis, University of Brittish Columbia.

Åström, K. J. and B. Wittenmark (1995): *Adaptive Control*, 2nd edition. Addison-Wesley.

Åström, K. J. and B. Wittenmark (1997): *Computer Controlled Systems*, 3rd edition. Prentice Hall, Englewood Cliffs, N. J.

Boutayeb, M. and M. Darouach (1995): "Recursive identification for MIMO Wiener-Hammerstein model." *IEEE Transaction on Automatic Control*, **AC-40**, pp. 287–291.

Draper, C. S. and Y. T. Li (1951): "Principles of optimalizing control systems and an application to the internal combustion engine." In Oldenburger, Ed., *Optimal and Self-Optimizing Control*, pp. 415–429. MIT Press, 1966. Original paper published in *American Society of Mechanical Engineers*, September, 1951.

Dumont, G. and K. J. Åström (1988): "Wood chip refiner control." *IEEE Control Systems Magazine*, **8:2**, pp. 38–43.

Filatov, N. M. and H. Unbehauen (2000): "Survey of adaptive dual control methods." *IEEE Proc. Control Theory Appl.*, **147**, pp. 118–128.

Hagenblad, A. and L. Ljung (2000): "Maximum likelihood estimation of Wiener models." In *Proc. 39th IEEE Conf. on Decision and Control*, pp. 2417–2418. Sydney, Australia.

Jacobs, O. L. R. and S. M. Langdon (1970): "An optimal extremal control system." *Automatica*, **6**, pp. 297–301.

Keviczsky, L. and R. Haber (1974): "Adaptive dual extremum control by Hammerstein models." In *IFAC Symposium on Stochastic Control*, pp. 333–341. Budapest.

Keviczsky, L., I. Vajk, and J. Hetthéssy (1979): "A self-tuning extremal controller for the generalized Hammerstein model." In *IFAC Symposium on Identification and System Parameter Estimation*, pp. 1147–1151. Darmstadt.

Krstić, M. and H.-H. Wang (1997): "Design and stability analysis of extremum seeking feedback for general nonlinear systems." In *Proceedings of the 36th IEEE Conference on Decision and Control*, vol. 2, pp. 1743–1748. San Diego.

Krstić, M. and H.-H. Wang (2000): "Stability of extremum seeking feedback for general nonlinear dynamic systems." *Automatica*, **36**, pp. 595–601.

Navarro, L. B. and M. B. Zarrop (1996): "Dynamic extremum control." In *Proceedings of the 1996 IEE Colloquium on Adaptive Control*. London.

Scotson, P. G. and P. E. Wellstead (1990): "Self-tuning optimization of spark ignition automotive engines." *IEEE Control Systems Magazine*, **10:3**, pp. 94–101.

Sternby, J. (1980a): "Adaptive control of extremum systems." In Unbehauen, Ed., *Methods and Applications in Adaptive Control*, number 24 in Lecture Notes in Control and Information Sciences, pp. 151–160. Springer-Verlag, Berlin, FRG.

Sternby, J. (1980b): "Extremum control systems – An area for adaptive control." In *Preprints Joint American Control Conference*. Paper WA2-A.

Wellstead, P. E. and P. G. Scotson (1990): "Self-tuning extremum control." *Control Theory and Applications, IEE Proceedings D*, **137:3**, pp. 165–175.

Wigren, T. (1993): "Recursive prediction error identification using the nonlinear Wiener model." *Automatica*, **29**, pp. 1011–1025.

Wittenmark, B. (1993): "Adaptive control of a stochastic nonlinear system: An example." *Int. J. Adaptive Control and Signal Processing*, **7**, pp. 327–337.

Wittenmark, B. and A. Urquhart (1995): "Adaptive extremal control." In *Preprints 34th Conference on Decision and Control*, pp. 1639–1644. New Orleans.

Zhu, Y. C. (1999): "Parametric Wiener model identification for control." In *Proceedings of the 14th IFAC World Congress*, vol. H. Beijing. Paper 3a-02:1.