DIRECTIONAL LEAKAGE AND PARAMETER DRIFT

Morten Hovd [*,1] **Robert R. Bitmead** [**]

[*] *Department of Engineering Cybernetics, Norwegian University of Technology and Science, N-7491 Trondheim, Norway*
[**] *Department of Mechanical and Aerospace Engineering, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0411, USA*

Abstract: A new method for eliminating parameter drift in parameter estimation problems is proposed. Existing methods for eliminating parameter drift work either on a limited time horizon, restrict the parameter estimates to a range that has to be determined *à priori*, or introduce bias in the parameter estimates which will degrade steady state performance. The idea of the new method is to apply leakage only in the directions in parameter space in which the exciting signal is not informative. This avoids the problem of parameter bias associated with conventional leakage.

Keywords: Parameter estimation, parameter drift, leakage

1. INTRODUCTION

Parameter drift is a well known problem in parameter estimation problems where the exciting signal is not sufficiently rich. The parameters estimated will typically converge to a manifold in parameter space on which the observed relationships between the input and output data are well explained. Noise in the observed data will then cause the parameter estimates to drift along this manifold in parameter space.

When the parameter estimates are used in an adaptive controller, the parameter drift will be unproblematic until parameter values are approached for which the model becomes uncontrollable. When this happens, bursting occurs. Bursting involves violent moves in the input as the controller attempts to control a plant which it erroneously thinks is close to uncontrollable, with corresponding large moves in the controlled output. The bursting typically supplies the parameter estimation algorithm with informative data, causing the parameter estimates to move to more reasonable values, from which the parameter drift starts anew.

There is a substantial literature on parameter drift and avoidance of bursting. Common modifications are the use of a deadzone in the parameter update, or the use of leakage. With a deadzone, the parameter updates are stopped as long as the model follows the observed system behaviour with reasonable accuracy. This will reduce the parameter drift, which can only occur when the parameter update is active, but will also reduce the accuracy with which the observed system behaviour can be approximated by the model.

With leakage, on the other hand, one adds an extra term in the parameter update law which drives the parameters towards a particular choice of reference values. Sufficiently strong leakage can eliminate bursting, but will bias the parameter values toward the chosen reference values. A good portion of luck is required for the reference values for the parameters to accurately describe the

[1] To whom correspondence should be addressed. e-mail: morten.hovd@itk.ntnu.no

observed input-output behaviour of the system, and hence leakage typically degrades steady state performance.

In this work, the properties of the exciting signal and the structure of the model is used such that the leakage is applied only in the directions in parameter space in which the input-output data is not informative. Hence, steady state performance is not degraded by the applied leakage.

1.1 *Preliminaries*

This section defines notation and states some established results on parameter estimation. Much of the material in this section is extracted from Ioannou and Sun (1996), where a more comprehensive treatment of parameter estimation and adaptive control can be found. We consider the estimation of model parameters for linear, time-invariant plants. The plant model is expressed as an nth-order difference equation given by

$$\begin{aligned} y_{k+n} + a_{n-1}y_{k+n-1} + \cdots + a_0 y_k &= \\ b_{n-1}u_{n-1} \quad + b_{n-2}u_{n-2} + \cdots + b_0 u & \quad (1) \end{aligned}$$

Obviously, this corresponds to the discrete transfer function representation

$$\begin{aligned} y(z) &= \frac{b_{n-1}z^{n-1} + b_{n-2}z^{n-2} + \cdots + b_0}{z^n + a_{n-1}z^{n-1} + z_{n-2}z^{n-2} + \cdots + a_0} u(z) \\ &= g(z)u(z) \qquad (2) \end{aligned}$$

The unknown parameters are lumped in a parameter vector

$$\theta = [b_0, b_1, \cdots, b_{n-1}, a_0, a_1, \cdots, a_{n-1}]^T \qquad (3)$$

and all input-output signals are collected in the signal vector

$$\begin{aligned} \phi_{k-1} = [&u_{k-n}, u_{k-n+1}, \cdots, u_{k-1}, \\ &-y_{k-n}, -y_{k-n+1}, \cdots, -y_{k-1}]^T \qquad (4) \end{aligned}$$

Thus, Eq. (1) can be compactly expressed as

$$y_k = \phi_{k-1}^T \theta \qquad (5)$$

In a practical situation, measurements will be corrupted by noise. Therefore it is commonplace to filter both y_k and ϕ_{k-1} in the equation above with a common low pass filter. However, in order to keep the presentation simple, we will ignore such filtering in most of the analysis to follow.

Let $\hat{\theta}$ denote an estimate of the parameter vector θ, and $\tilde{\theta} = \theta - \hat{\theta}$ denote the corresponding parameter error vector. The observed model error is denoted e, and is given by

$$e_k = y_k - \phi_{k-1}^T \hat{\theta}_{k-1} = \tilde{\theta}_{k-1}^T \phi_{k-1} \qquad (6)$$

1.2 *The gradient method*

We choose an instantaneous optimization criterion of the form

$$\begin{aligned} J(\hat{\theta}_k) = &\frac{1}{2}(y_k - \phi_{k-1}^T \hat{\theta}_k)^2 \\ &+ \frac{c}{2}(\hat{\theta}_k - \hat{\theta}_{k-1})^T (\hat{\theta}_k - \hat{\theta}_{k-1}) \qquad (7) \end{aligned}$$

Differentiating with respect to $\hat{\theta}_k$ we obtain

$$\begin{aligned} \left(\frac{dJ}{d\hat{\theta}_k}\right) &= -(y_k - \phi_{k-1}^T \hat{\theta}_k)\phi_{k-1} + c(\hat{\theta}_k - \hat{\theta}_{k-1}) \\ &= -e_k \phi_{k-1} + \phi_{k-1}^T(\hat{\theta}_k - \hat{\theta}_{k-1})\phi_{k-1} \\ &\quad + c(\hat{\theta}_k - \hat{\theta}_{k-1}) \qquad (8) \end{aligned}$$

Setting the gradient to zero, and noting that $(cI + \phi_{k-1}\phi_{k-1}^T)^{-1}\phi_{k-1} = \frac{1}{c+\phi_{k-1}^T\phi_{k-1}}\phi_{k-1}$, we obtain the standard equation error parameter estimation algorithm

$$\begin{aligned} \hat{\theta}_k = &\hat{\theta}_{k-1} \\ &+ \frac{1}{c + \phi_{k-1}^T \phi_{k-1}} \phi_{k-1} \left(y_k - \phi_{k-1}^T \hat{\theta}_{k-1} \right) \qquad (9) \end{aligned}$$

We have from Eq. (5) that Eq. (9) in the absence of measurement noise may be expressed as

$$\tilde{\theta}_k = \tilde{\theta}_{k-1} + \frac{1}{c + \phi_{k-1}^T \phi_{k-1}} \phi_{k-1}\phi_{k-1}^T \tilde{\theta}_{k-1} \qquad (10)$$

The matrix $\phi\phi^T$ is clearly singular (of rank 1) at any one time instant. Equation (10), on the other hand, contains $2n$ integrators. It might therefore appear as if there is only one integrator (or linear combination of integrators) that is stabilized by feedback, and that the remaining integrators are left to integrate the noise in the measurement signal. However, if the systematic variation in ϕ is such that the sum of $\phi\phi^T$ over any time interval $[k, k+t]$ is positive definite, there is actually negative feedback around all the integrators in Eq. (10). This results in an asymptotically stable parameter estimation, with the estimated parameters in the absence of measurement noise converging to the true values at steady state.

The requirement that the sum of $\phi\phi^T$ should be positive definite is the familiar *persistent excitation* requirement. The problem of parameter drift and the bursting that results from it arise when the persistent excitation requirement is not met.

1.3 *Least squares parameter estimation algorithm*

The least squares parameter estimation algorithm with exponential data weighting results from recursively minimizing the optimization criterion

$$J_L(\hat{\theta}_k) = \frac{1}{2}\sum_{t=1}^{k}\beta^{k-t}(y_t - \phi_{t-1}^T\hat{\theta}_k)^2 + \frac{1}{2}\beta^k(\hat{\theta}_k - \hat{\theta}_0)\Pi_0^{-1}(\hat{\theta}_k - \hat{\theta}_0) \quad (11)$$

starting from a given initial parameter guess $\hat{\theta}_0$ and a positive definite Π_0. Here β ($0 < \beta \leq 1$) is known as the exponential forgetting factor, and is normally chosen to be slightly less than 1. It is shown in, e.g., Goodwin and Sin (1984) that minimizing Eq. (11) results in the recursive scheme

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \frac{\Pi_{k-2}\phi_{k-1}}{\beta + \phi_{k-1}^T\Pi_{k-2}\phi_{k-1}}\left[y_k - \phi_{k-1}^T\hat{\theta}_{k-1}\right] \quad (12)$$

$$\Pi_{k-1} = \frac{1}{\beta}\left[\Pi_{k-2} - \frac{\Pi_{k-2}\phi_{k-1}\phi_{k-1}^T\Pi_{k-2}}{\beta + \phi_{k-1}^T\Pi_{k-2}\phi_{k-1}}\right] \quad (13)$$

Choosing $\beta = 1$ results in no forgetting of old data. As a result, Π_k will approach zero, and the parameter estimation will essentially turn itself off. A consequence will be that the parameter estimation will be unable to track even slow parameter variations. To avoid this problem, one may choose β slightly smaller than 1. Whereas this makes the tracking of slow parameter variations possible, it also gives room for parameter drift in cases where the exciting signal is not sufficiently rich. It is well know that an excitation signal containing m sinusoids is *sufficiently rich of order* $2m$, i.e., will allow accurate estimation of $2m$ parameters (or $2m$ linear combinations of parameters).

2. DIRECTIONAL LEAKAGE

In this section we will propose a method for performing parameter updates in the subspace of parameter space for which the I/O data is informative, while preventing parameter drift by implementing leakage in the orthogonal subspace. We will first consider the gradient method for parameter estimation, and thereafter consider the least squares method. For both parameter estimation methods, stability is analyzed using averaging, in a manner similar to that of (Bitmead and Johnson, 1987).

2.1 *Directional leakage with the gradient method*

After selection of a basis for the informative subspace, and collecting the basis vectors in a matrix Υ, this is achieved simply by modifying the parameter update equation Eq. (9) as follows

$$\hat{\theta}_k = \hat{\theta}_{k-1} + P\frac{1}{c + \phi_{k-1}^T\phi_{k-1}}\phi_{k-1}\phi_{k-1}^T\left(\theta - \hat{\theta}_{k-1}\right) + rP^{\perp}\left(\theta_{ref} - \hat{\theta}_{k-1}\right) \quad (14)$$

where $0 < r < 1$, $P = \Upsilon(\Upsilon^T\Upsilon)^{-1}\Upsilon^T$ is a projection matrix and $P^{\perp} = I - P$ its orthogonal complement (and $P^{\perp}$ is itself a projection matrix).

We can decompose the dynamics of the parameter estimation described by Eq. (14) into directions driven by input-output data and directions driven by the leakage, by premultiplying both sides of Eq. (14) by P and $P^{\perp}$, respectively.

$$P\hat{\theta}_k = P\left(I - \frac{1}{c + \phi_{k-1}^T\phi_{k-1}}\phi_{k-1}\phi_{k-1}^T\right)\hat{\theta}_{k-1} + P\frac{1}{c + \phi_{k-1}^T\phi_{k-1}}\theta \quad (15)$$
$$= PH_{k-1}\hat{\theta}_{k-1} + PG_{k-1}\theta$$

$$P^{\perp}\hat{\theta}_k = P^{\perp}(1-r)\hat{\theta}_{k-1} + rP^{\perp}\theta_{ref} \quad (16)$$

The dynamics in the directions described by Eq. (16) are trivially asymptotically stable, and have a steady state described by $P^{\perp}(\theta_{ref} - \theta_{\infty}) = 0$. The dynamics described by Eq. (15) require more careful analysis.

Assuming that $\phi_{k-1}\phi_{k-1}^T$ is in the range of P, and that P is of rank $2m$ (where $m < n$), PH_{k-1} has $2(n-m)$ eigenvalues at the origin (the singular directions of P), one eigenvalue inside the unit disk, and the remaining eigenvalues at $+1$. One therefore has to study the evolution of the parameters over several timesteps to conclude about stability. Considering an entire oscillation period for the input signal, i.e. $s = \tau/T$ timesteps, one obtains

$$P\hat{\theta}_{k+s} = \prod_{i=k}^{k+s-1}PH_i\hat{\theta}_k + \sum_{i=k}^{k+s-1}\left(\prod_{l=i+1}^{k+s-1}PH_l\right)PG_i\theta \quad (17)$$

The system signals are repeated every s timesteps, and therefore the stability of the parameter estimation will depend on $\prod_{i=k}^{k+s-1}PH_i$. To simplify notation, denote $\frac{1}{c+\phi_{i-1}^T\phi_{i-1}}$ by ϵ_i and $\phi_{i-1}\phi_{i-1}^T$ by X_i. For sufficiently large c, we have $\epsilon_i \approx \epsilon << 1 \forall i$. Tedious, but straightforward calculation [2] then results in

$$\prod_{i=k}^{k+s-1}H_i = P - P(\epsilon_k X_k + \epsilon_{k+1}X_{k+1} + \cdots + \epsilon_{k+s-1}X_{k+s-1})P + O(\epsilon^2\frac{s(s-1)}{2}) \quad (18)$$

[2] We will require the technical assumption $PX_k = X_kP$, which will be fulfilled by our subsequent choice of P.

Provided $\epsilon\frac{s(s-1)}{2} << 1$, the linear term in Eq. (18) will dominate. Thus, the parameter estimation will be asymptotically stable provided the rank of $P(X_k + X_{k+1} + \cdots + X_{k+s-1})P$ equals the rank of P. Thus, we have to determine which basis to use as the columns of Υ, to ensure that this rank condition holds. Classical results on sufficient richness of signals provide us with the dimension of the informative subspace, i.e., the number of columns in Υ. A straightforward approach is to sum $\phi^T\phi$ over some window of past data with window length of some multiple of s timesteps. The singular vectors corresponding to the $2m$ largest singular values of the resulting matrix can then be chosen as the range space of P. In the ideal case (in the absence of noise) this would lead to a time invariant projection matrix P once the effects of initial conditions have died out.

Example 1. The idea of using directional leakage is illustrated using a simple example with $g(z) = \frac{b}{z+a}$. The input is a unit step signal (corresponding to $\omega = 0 \Rightarrow z = 1$), and is thus sufficiently rich to estimate only one parameter. It is well known that in this situation only the steady state gain, $g(1) = \frac{b}{1+a}$ can be estimated. The true parameters in this example are given by $a = 0.7,\ b = 0.7$. The estimation is simulated in Simulink, and the measurement is corrupted with white noise with power 0.1. The estimation is simulated both with no leakage, and with directional leakage. Initial parameter estimates are $a = 2,\ b = 6$ in both simulations. In the simulations, both input and output are low pass filtered through identical low pass filters $f(z) = \frac{1}{z+0.8}$.

The results without leakage are shown in Fig. 1. The estimation quickly converges to parameter values giving the correct steady state gain (the straight line in the figure), and then starts to drift while staying close to the correct steady state gain.

The results obtained when using directional leakage are shown in Fig. 2. The reference values θ_{ref} are chosen equal to the initial parameters. The parameter estimates converge to a point on the line indicating the correct steady state gain, and no significant parameter drift can be found.

2.2 *Directional leakage with the least squares parameter estimation method*

It was noted above that without exponential forgetting the least squares parameter estimation method will eventually turn itself off. Thus, parameter drift and leakage (whether directional or not) only become relevant when $\beta < 1$.

Introducing directional leakage in Eq. (12) we obtain

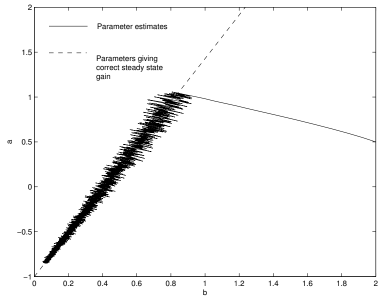

Fig. 1. Results for example 1 when using no leakage.

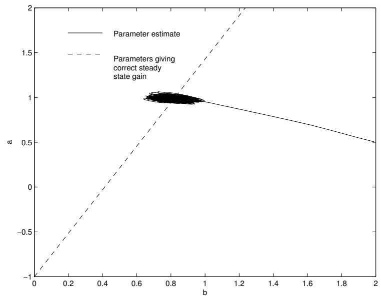

Fig. 2. Results for example 1 when using directional leakage.

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \Bigg\{ P \frac{\Pi_{k-2}\phi_{k-1}}{\beta + \phi_{k-1}^T \Pi_{k-2}\phi_{k-1}} \times \Big[y_k - \phi_{k-1}^T \hat{\theta}_{k-1} \Big] \Bigg\} + P^{\perp}(\theta_{ref} - \theta_{k-1}) \quad (19)$$

whereas Eq. (13) in principle remains unchanged. However, it is clear from Eq. (19) that the matrix Π_{k-2} affects the gain of the parameter estimation. With exponential forgetting, Π_{k-2} may become very large in directions that are poorly excited - and even grow without bounds in directions that are not excited at all. We will therefore impose an upper limit on each of the singular values of Π. Other (and less computationally demanding) methods for keeping Π bounded are discussed in (Goodwin and Sin, 1984).

Proceeding with the analysis for the exponentially weighted least squares, we arrive at an equation mirroring Eq. (17) above. However, we here get that $H_i = I - \frac{1}{\beta + \phi_{k-1}^T \Pi_{k-2}\phi_{k-1}} \Pi_{k-2}\phi_{k-1}\phi_{k-1}^T$. In evaluating $\prod_{i=k}^{k+s-1} PH_i$, we then face the problem that in general Π_{k+m} and $\phi_{k+n}\phi_{k+n}^T$ do not commute. We are therefore unable to arrive at an equally simple expression as in Eq. (18). Instead, we obtain

$$\prod_{i=k}^{k+s-1} H_i = P - P(\epsilon_k \Pi_k X_k + \epsilon_{k+1}\Pi_{k+1}X_{k+1} + \cdots + \epsilon_{k+s-1}\Pi_{k+s-1}X_{k+s-1}) + O((\bar{\epsilon}\bar{\Pi})^2 \frac{s(s-1)}{2}) \quad (20)$$

where here $\epsilon_i = \frac{1}{\beta + \phi_{i-1}^T \Pi_{i-2} \phi_{i-1}}$ whereas $X_i = \phi_{i-1}\phi_{i-1}^T$ as before. The symbols $\bar{\epsilon}$ and $\bar{\Pi}$ represent average values of ϵ_i and Π_i, respectively. We find that for sufficiently small $\bar{\Pi}$, the linear term in Eq. (20) will dominate, here $\bar{\epsilon}$ may well be close to 1. A sufficiently small $\bar{\Pi}$ is obtained by choosing a forgetting factor β close to 1, *and* effectively constraining growth of Π in directions that are insufficiently excited. It should be noted that in the directions in the input-output data that are poorly but persistently excited, Π may grow large unless β is very close to 1 or the singular values of Π are properly constrained. Still, we can conclude from Eq. (20) that given a sufficiently small $\bar{\Pi}$, the parameter estimation will remain asymptotically stable provided $rank(P) = rank(P(\epsilon_k \Pi_k X_k + \epsilon_{k+1}\Pi_{k+1}X_{k+1} + \cdots + \epsilon_{k+s-1}\Pi_{k+s-1}X_{k+s-1})$ Since Π_i is positive definite for all i, this rank condition is essentially equivalent to the rank condition found for the gradient method.

Example 2. The use of directional leakage will next be illustrated used on a problem in adaptive control. The plant is given by

$$g(z) = \frac{2z - 1.8}{z^3 - 2.78z^2 + 2.5715z - 0.79135} \quad (21)$$

The control objective is to track the reference signal

$$r(t) = \sin(\frac{2\pi t}{360}) + \sin(\frac{2\pi t}{45}) \quad (22)$$

Thus, we have five unknown parameters, whereas the reference signal would only be sufficiently rich to identify four parameters. There is also an additive measurement noise, modelled as a normally distributed, zero mean random variable with standard deviation 0.04. The control design is accomplished using adaptive pole placement, using a discrete version of the method described in (Ioannou and Sun, 1996). A sampling interval of 1 is used, and all poles of the closed loop characteristic polynomial are placed at $z = 0.7$. For the parameter estimation, state variable filters for both inputs and outputs equal to $f(z) = \frac{1}{(z-0.8)^3}$ are used. The estimation is done with the least squares method with a forgetting factor $\beta = 0.998$. The singular values of Π are constrained to be no larger than $1e - 4$. The initial parameter estimates correspond to the transfer function.

$$\hat{g}(z, \hat{\theta}) = \frac{2(z - 0.91)}{(z - 0.99)(z - 0.93)(z - 0.86)} \quad (23)$$

Although these initial parameter values would appear to be quite good, initial control performance is appalling. The performance quickly improves, and good performance is obtained after a short initial transient. Luckily, this paper addresses long term rather than initial behavior, and parameter estimation rather than control. Similar control performance is obtained for the simulated time period both without leakage and with directional leakage. Figure 3 may represent the typical control performance for both cases. The results (Figs. 4 - 8) show that the parameter estimates drift when no leakage is applied. When directional leakage is applied, the noise does affect the parameter estimates, but no parameter drift can be seen. The effect of the noise that is seen in the parameter estimates appear have no effect on the control performance. However, the effect of noise may be reduced by reducing the maximum singular value bound for the covariance matrix, and/or using a larger forgetting factor. The forgetting factor of $\beta = 0.998$ corresponds to a time constant for forgetting of 500 samples. This may appear to be somewhat too fast when the signal to be tracked has a sinusoid of period 360.

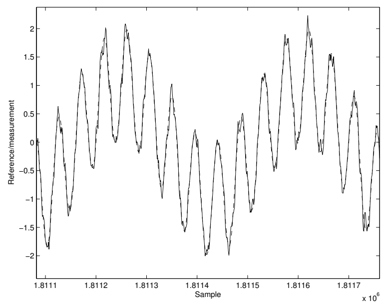

Fig. 3. Typical control performance for Example 2. Solid: noise-corrupted measurement, dashed: reference signal.

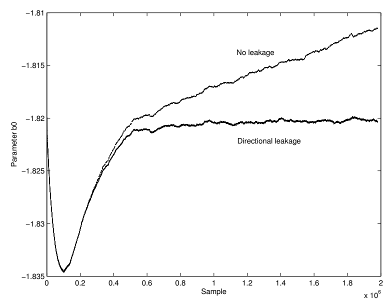

Fig. 4. Estimates of parameter b_0 in Example 2.

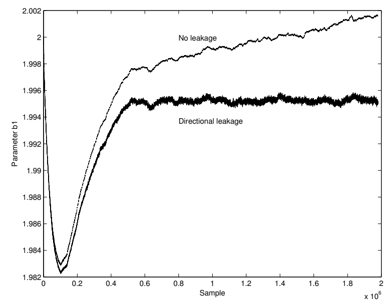

Fig. 5. Estimates of parameter b_1 in Example 2.

3. DISCUSSION AND CONCLUSIONS

This paper has focused on the gradient and least squares methods for parameter estimation. Note

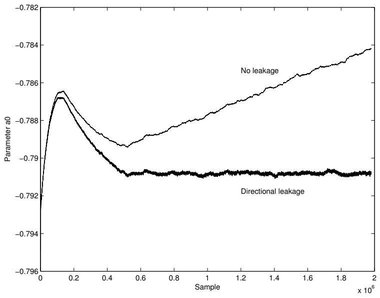

Fig. 6. Estimates of parameter a_0 in Example 2.

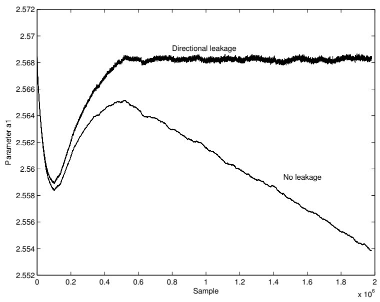

Fig. 7. Estimates of parameter a_1 in Example 2.

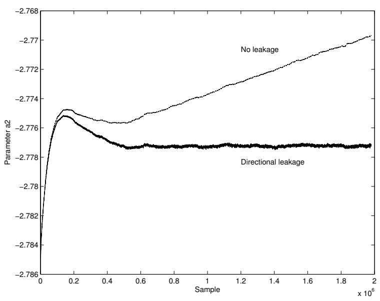

Fig. 8. Estimates of parameter a_2 in Example 2.

that other parameter estimation algorithms, including parameter estimation using the Extended Kalman Filter, suffer from the same parameter drift problems whenever the system is not sufficiently excited and tracking of time-varying parameter values is desired. 'Conventional' leakage is a simple method for eliminating drift, but unfortunately conventional leakage results in bias in the parameter estimates. This paper shows how leakage can be applied only in the directions in parameter space where the plant is not properly excited. This stabilizes the parameter estimation (stops parameter drift) without degrading parameter estimation accuracy in the directions where the input-output data is informative.

It was noted previously that bursting will occur if the parameter estimates approach a hypersurface in parameter space where the model is uncontrollable. Although directional leakage, when properly tuned, makes the parameter estimation asymptotically stable, that does not rule out the possibility that during the initial transient the parameters may encounter such an uncontrollable hypersurface. For problems where this is a serious concern, directional leakage may easily be combined with so-called *parameter projection* (Ioannou and Sun, 1996), which constrains the parameter estimates to stay within an *à priori* defined region of the parameter space where the true model is assumed to lie, and which is also assumed to be free of uncontrollable hypersurfaces.

Thus, despite the advantages of directional leakage noted above, it does not totally eliminate the problem of bursting. Note, however, that such bursting during initial transients may also occur when the exciting signals are persistently exciting, if the initial parameter estimates are ill-chosen. Therefore, any failure to prevent bursting during transient conditions should not be considered a shortcoming of (directional) leakage - it is simply not the problem it is intended to prevent.

The use of a dead-zone (see, e.g., (Egardt, 1979)) in the parameter estimation is a common way of reducing parameter drift. The dead-zone reduces parameter drift by stopping the parameter update when model predictions are close to the physical measurement. Directional leakage, on the other hand, does not stop parameter estimation when the model output is close to the measurement. Since leakage is applied only in the directions in parameter space where the signals are not informative, the asymptotic model accuracy is not affected. Directional leakage should therefore be an attractive alternative to the use of dead-zones.

Acknowledgements The first author got the idea for this paper while attending Miroslav Krstic's Adaptive Control class during a sabbatical at UCSD, and wishes to thank prof. Krstic for his inspiring teaching.

4. REFERENCES

Bitmead, R. R. and C. R. Johnson (1987). *Discrete Averaging Principles and Robust Adaptive Identification.* pp. 237–271. Vol. 25 of *Advances in Control and Dynamic Systems.* Academic Press. New York.

Egardt, B. (1979). *Stability of Adaptive Controllers.* Vol. 20 of *Lecture Notes in Control and Information Sciences.* Springer-Verlag. Berlin, Germany.

Goodwin, C. G. and K. S. Sin (1984). *Adaptive Filtering Prediction and Control.* Prentice-Hall. Englewood Cliffs, New Jersey, USA.

Ioannou, P. A. and J. Sun (1996). *Robust Adaptive Control.* Prentice-Hall. Upper Saddle River, New Jersey, USA.