The interface between sequential experimental design and dual control

L. Pronzato[a], *C. Kulcsár*[b] *and E. Walter*[b]

[a]Laboratoire I3S, CNRS-URA 1376, bât. 4, 250 rue A. Einstein, Sophia Antipolis, 06560 Valbonne, France

[b]Laboratoire des Signaux et Sytèmes, CNRS-ESE, plateau de Moulon, 91192 Gif-sur-Yvette, France

The links between experimental design and control are generally presented as twofold: (i) the experimental conditions to be chosen may contain the input signals to be applied on the system, (ii) the experiments may be chosen sequentially, which results in a dynamical control problem even if the system is purely static. The situation considered here involves deeper interactions between experimental design and control, because the experiment is designed to optimise the quadratic distance between the response of a system with unknown parameters and some reference trajectory.

1 Experimental design for parameter estimation

Assume that a vector of observations $\boldsymbol{y} = (y(1), \ldots, y(N))$, with $y(k) = \eta(\bar{\theta}, k) + \epsilon(k)$, is available to estimate the model parameters $\theta \in \mathbb{R}^p$. Here $\bar{\theta}$ denotes the unknown true value of these parameters, and the measurement errors $\epsilon(k)$ are assumed to be independently identically distributed (i.i.d.), with a known distribution. The maximum likelihood estimator is then given by $\hat{\theta}_{ML} = \arg\max_{\theta \in \mathbb{R}^p} \pi(\boldsymbol{y}|\theta)$, where $\pi(\boldsymbol{y}|\theta)$ is the conditional probability density of $\boldsymbol{y}$ given θ (the likelihood function). When the errors are normally distributed $\mathcal{N}(0, \sigma^2)$, $\hat{\theta}_{ML}$ coincides with the least-squares estimator $\hat{\theta}_{LS} = \arg\min_{\theta \in \mathbb{R}^p} \sum_{k=1}^{N} (y(k) - \eta(\theta, k))^2$, which will be simply denoted by $\hat{\theta}$ in what follows. An important issue is the precision of the estimation. Under rather general conditions, $\hat{\theta}$ is asymptotically normally distributed $\mathcal{N}(\bar{\theta}, M_F^{-1}(\bar{\theta}))$, with $M_F(\theta)$ the Fisher information matrix, given in the normal case by

$$M_F(\theta) = \frac{1}{\sigma^2} \sum_{k=1}^{N} \frac{\partial \eta(\theta, k)}{\partial \theta} \frac{\partial \eta(\theta, k)}{\partial \theta^T} .$$

A precise estimation thus means a "small" matrix $M_F^{-1}(\theta)$, which is the basic idea for experimental design, see, e.g., Fedorov (1972).

2 First links between experimental design and control

2.1 Design of input sequences for a dynamic system

Consider for example an ARX system defined by

$$y(k) = \bar{a}y(k-1) + \bar{b}u(k-1) + \epsilon(k)\,,\ k = 1, \ldots, N\,,\ y(1) = 0\,,$$

where the parameters to be estimated are $\theta = (a, b)^T$, and where the errors $\epsilon(k)$ are i.i.d. $\mathcal{N}(0, \sigma^2)$. The Fisher information matrix is then given by

$$M_F(\theta) = \frac{1}{\sigma^2}\sum_{k=2}^{N+1}\begin{bmatrix} y_m^2(k-1) + \sigma^2\frac{1-a^{2k}}{1-a^2} & -u(k-1)y_m(k-1) \\ -u(k-1)y_m(k-1) & u^2(k-1) \end{bmatrix},$$

with $y_m(k) = ay_m(k-1) + bu(k-1)$ and $y_m(k) = 0$ for $k \leq 1$. The problem is to choose the sequence $u(1), \ldots, u(N)$ in order to maximise $\det M_F(\theta)$ (what is called D-optimal design), with a power constraint $\frac{1}{N}\sum_{k=1}^{N} u^2(k) \leq p_{\max}$. This corresponds to an optimal control problem, with a rather unusual control objective related to the precision on the parameter estimates. See Walter and Pronzato (1994) for details on this example and e.g. Zarrop (1979) for a survey on optimal input design.

2.2 Sequential design on a static system

Consider now the following example of a static system: $y(k) = \exp(-\bar{\theta}\xi^k) + \epsilon(k)$, where $\bar{\theta}$ is scalar, $\xi^k \in \mathbb{R}$ denotes the experimental condition of the kth observation, and the measurement errors are i.i.d. $\mathcal{N}(0, \sigma^2)$. The Fisher information matrix then reduces to the scalar

$$M_F(\theta) = \frac{1}{\sigma^2}\sum_{k=1}^{N}(\xi^k)^2\exp(-2\theta\xi^k)\,.$$

Designing a D-optimal experiment then means choosing the ξ^k's that maximise $M_F(\theta)$, which yields $\xi^1 = \xi^2 = \ldots = \xi^N = 1/\theta$. Since the true value of θ is unknown, a prior nominal value is classically used at this design stage (which is then called *local design*). The experiment can also be performed sequentially, see, e.g., Chernoff (1975), as indicated on Figure 1.

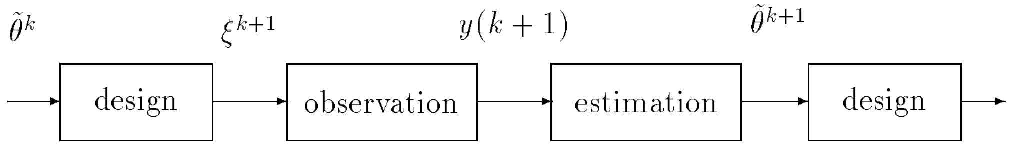

Figure 1. Sequential experimental design.

Although the system under study is static, ξ^{k+1} depends on previous observations $y(1), \ldots, y(k)$, which makes sequential design a dynamic process. The most widely used approach consists in choosing ξ^{k+1} optimally for some estimated parameter $\tilde{\theta}^k$, which is called *Forced Certainty Equivalence* (FCE) in the control literature, see Section 3. Note that the choice of $\tilde{\theta}^k$ is crucial. If it is estimated from $y(k)$ only, there is no accumulation of information and thus no convergence. A classical choice is $\tilde{\theta}^k = \frac{1}{k}\sum_{i=1}^k \hat{\theta}^i$, with $\hat{\theta}^i$ estimated from $y(i)$ only. It has been shown in Pronzato *et al.* (1993) that in this case $\tilde{\theta}^k$ does *not* converge to $\bar{\theta}$. In order to obtain convergence, $\tilde{\theta}^k$ should be estimated from all previous observations $y(1), \ldots, y(k)$.

A possible alternative to face the dependence in θ of the optimal experiment is to resort to an average optimality criterion. For instance, an ELD-optimal experiment maximises $E_\theta\{\log\det M_F(\theta)\}$, with the expectation E_θ computed for a prior distribution $\pi(\theta)$. For the static example considered here, if $N=2$, $\xi^1 = \xi^2 = \frac{1}{E_\theta\{\theta\}}$. This approach can also be implemented sequentially. Assume first that there is no constraint on ξ^1 and ξ^2. One then first applies $\xi^1 = \frac{1}{E_\theta\{\theta\}}$, observes $y(1)$ and computes the posterior distribution $\pi(\theta|y(1))$. At the second step, one applies $\xi^2 = \frac{1}{E_{\theta|y(1)}\{\theta\}}$. Although all the information available is taken into account when choosing ξ^2, such a procedure is not optimal (it might be called passively adaptive, see Section 3). Indeed, ξ^1 was chosen as if it corresponded to the last observation to be performed. The optimal policy corresponds to $\xi^{1*} = \arg\max_{\xi^1} E_{y(1)}\{h(\xi^1, y(1))\}$, where $h(\xi^1, y(1)) = \max_{\xi^2} E_{\theta|y(1)}\{\log\det M_F(\theta)\}$. If a constraint $\xi^2 \geq \xi^1$ is imposed, a cautious design $\xi^{1*} < \frac{1}{E_\theta\{\theta\}}$ is then obtained (Pronzato *et al.*, 1993).

More generally, the optimal sequential policy corresponds to the solution of a stochastic control problem, involving an imbedded sequence of optimisations and maximisations, which is extremely difficult to perform, see, e.g., Kulcsár *et al.* (1994) for the joint problem of choosing the input sequence and the measurement times in a simple biomedical example. Stochastic control is considered in a more general setup in the next section.

3 stochastic optimal control

Consider a SISO dynamical system defined by the following state-space equations:

$$\boldsymbol{x}(k+1) = \boldsymbol{f}(\boldsymbol{x}(k), u(k), \boldsymbol{v}(k))\,,\ y(k) = g(\boldsymbol{x}(k), \epsilon(k))\,,\ k = 1, \ldots, N\,, \tag{1}$$

where $\boldsymbol{v}(k)$ and $\epsilon(k)$ denote perturbations with known distributions. The control objective, to be minimised with respect to the input sequence $\boldsymbol{u}_0^{N-1} = (u(0), \ldots, u(N-1))^T$ takes the additive form $E\{\sum_{k=0}^{N-1} j(\boldsymbol{x}(k+1), u(k))|\mathcal{I}^0\}$, where $\mathcal{I}^0$ is the prior information on the system (the prior density of $\boldsymbol{x}^0$ and the distribution of the perturbations), and $E\{.\}$ denotes the expectation with respect to all random variables. The information at stage k will be denoted by $\mathcal{I}^k$, i.e. $\mathcal{I}^k = \{\mathcal{I}^0, \boldsymbol{y}_1^k, \boldsymbol{u}_0^{k-1}\}$.

The optimal input sequence is then obtained as the solution of the following stochastic dynamic programming problem:

$$\begin{aligned}
\min\nolimits_{u(0)} \quad & [E_{\boldsymbol{x}(1),y(1)}\{j(\boldsymbol{x}(1), u(0)) + \ldots \min_{u(k)}[E_{\boldsymbol{x}(k+1),y(k+1)}\{j(\boldsymbol{x}(k+1), u(k)) + \ldots \\
\min\nolimits_{u(N-2)} \quad & [E_{\boldsymbol{x}(N-1),y(N-1)}\{j(\boldsymbol{x}(N-1), u(N-2)) + \\
\min\nolimits_{u(N-1)} \quad & [E_{\boldsymbol{x}(N)}\{j(\boldsymbol{x}(N), u(N-1))|\mathcal{I}^{N-1}\}]|\mathcal{I}^{N-2}\}] \ldots |\mathcal{I}^k\}] \ldots |\mathcal{I}^0\}]\,.
\end{aligned}$$

This problem is well-known to be extremely difficult to solve, even in most simple situations, so that suboptimal approaches are classically used. The most frequent one corresponds to FCE, for which $u(k)$ is obtained via the solution of the optimisation problem $\min_{\boldsymbol{u}_k^{N-1}} \sum_{i=k}^{N-1} j(\hat{\boldsymbol{x}}(i+1)|\mathcal{I}^k, u(i))$, where $\hat{\boldsymbol{x}}(i+1)|\mathcal{I}^k$ is an estimate of $\boldsymbol{x}(i+1)$ obtained from the information $\mathcal{I}^k$. In *Open-Loop-Feedback-Optimal* (OLFO) control, $u(k)$ is obtained by the solution of the problem $\min_{\boldsymbol{u}_k^{N-1}} E\{\sum_{i=k}^{N-1} j(\boldsymbol{x}(i+1), u(i))|\mathcal{I}^k\}$, where $\boldsymbol{u}_k^{N-1}$ is a function of $\mathcal{I}^k$ only. This policy is *quasi-adaptive* (Bertsekas, 1976), in the sense that it performs better than the open-loop policy without feedback obtained by using the prior information $\mathcal{I}^0$ instead of $\mathcal{I}^k$ at each iteration. FCE control is not quasi-adaptive, see Thau and Witsenhausen (1966), but sometimes much simpler to compute than OLFO control.

Consider now the special case where the system is linear,

$$\left\{ \begin{array}{rcl} \boldsymbol{x}(k+1) & = & \boldsymbol{A}\boldsymbol{x}(k) + \boldsymbol{b}u(k) + \boldsymbol{v}(k)\,, \\ y(k) & = & \boldsymbol{c}^T\boldsymbol{x}(k) + \epsilon(k)\,, \end{array} \right. \tag{2}$$

with $\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}$ known, and gaussian independent perturbations $\boldsymbol{v}(k)$ and $\epsilon(k)$. The cost is quadratic, $j(\boldsymbol{x}(k+1), u(k)) = \boldsymbol{x}^T(k+1)\boldsymbol{Q}_{k+1}\boldsymbol{x}(k+1) + r_k u^2(k)$, with $\boldsymbol{Q}_k$ non-negative definite and r_k positive for all k. The control has then no *dual effect* (it is *neutral*), beacuse the inputs do not influence the accuracy of the estimation of $\boldsymbol{x}$, obtained by Kalman filtering. The system has also the *separation property*: the optimal value $u^*(k)$ of $u(k)$ only depends on $\mathcal{I}^k$ through the estimate $\hat{\boldsymbol{x}}(k)$. Moreover, the *Certainty Equivalence Principle* holds, and $u^*(k) = \phi(\hat{\boldsymbol{x}}(k))$, with $\phi(.)$ the same function (here a linear feedback) as would be used if $\boldsymbol{x}(k)$ were known.

These properties do no longer hold when $\boldsymbol{A}, \boldsymbol{b}$ or $\boldsymbol{c}$ contain unknown parameters θ. The control has then a *dual effect*: it influences the trajectory followed by the system as well as the accuracy of the estimation of $\boldsymbol{x}$ and θ. We then get a stochastic control problem for a system similar to (1), with an extended state $\boldsymbol{x}_c$ now including the state variables of the system (2) as well as the unknown parameters in $\boldsymbol{A}, \boldsymbol{b}$ and $\boldsymbol{c}$.

Classical suboptimal policies, like FCE, OLFO, cautious control... are *passively adaptive*: they take into account the past information but not the influence of a present action on future uncertainty. Few *active* policies, i.e. policies that take into account their dual effect, have been suggested in the literature, see Kulcsár (1995) for a survey. In general, either they are very complex, or they rely on a two-step ahead horizon. On the other hand, the prediction of future precision is the cornerstone of experimental design for parameter estimation, and it can be used to obtain simple suboptimal active strategies.

Consider for instance the case of a FIR model, defined by $y(k) = \sum_{i=1}^n \theta_i u(k-i) + \epsilon(k)$, and a control objective given by

$$E\{\sum_{i=0}^{N-1} q_{i+1}(y(i+1) - \gamma_{i+1})^2 + r_i u^2(i)|\mathcal{I}^0\}\,, \tag{3}$$

where $(\gamma_i)_i$ defines a known reference trajectory and the q_i's and r_i's are known weighting factors. At step k, we approximate the future posterior density $\pi(\theta|\boldsymbol{y}_1^i)$, $i > k$, by the density of the normal distribution $\mathcal{N}(\hat{\theta}^k, \boldsymbol{C}(i))$, where $\hat{\theta}^k$ and $\boldsymbol{C}(i)$ respectively denote the estimated value of the parameters at step k and their covariance matrix at step i, as obtained by recursive least-squares. The input at step k is then obtained by solving

$$\min_{\boldsymbol{u}_k^{N-1}} \sum_{i=k}^{N-1} (q_{i+1}\sigma^2 + q_{i+1}(\boldsymbol{u}_{i-n}^i)^T \hat{\theta}^k - \gamma_{i+1})^2 + q_{i+1}(\boldsymbol{u}_{i-n}^i)^T \boldsymbol{C}(i)(\boldsymbol{u}_{i-n}^i) + r_i u^2(i) \, .$$

The dual effect is taken into account through the dependence of $\boldsymbol{C}(i)$ in $\boldsymbol{u}_k^{i-1}$. The passive OLFO policy uses $\boldsymbol{C}(k)$ instead of $\boldsymbol{C}(i)$ for all $i > k$. FCE control ignores the uncertainty on θ and behaves like if the covariance matrices above were equal to the null matrix. Several examples of FIR models are presented in Kulcsár *et al.* (1995a), which illustrate the superiority of the dual approach advocated here over passive policies.

ARX and linear state-space models require more attention. At step k, we define a nominal trajectory based on $\hat{\theta}^k$ and predict the precision on $\hat{\theta}^i, i > k$ through the expected information matrix, evaluated at $\hat{\theta}^k$, see Pronzato *et al.* (1995a, 1995b), Kulcsár (1995) and Kulcsár *et al.* (1995b) for details and examples.

An important feature of active policies is that they are influenced by the knowledge that some observations will be censored (through the values of the matrices $\boldsymbol{C}(i), i > k$), while passive policies are not. The low complexity of the proposed suboptimal dual approach, together with the fact that all future horizon is taken into account, permits to face the joint design problem of the input sequence and measurement times. For any fixed number of observations, we can optimise their location by a DETMAX-like algorithm (Mitchell, 1974): each evaluation of the criterion for a given set of measurement times amounts to the determination of the suboptimal dual input sequence for the corresponding sampling schedule. A simpler suboptimal approach is based on the evaluation of the cost obtained for a fixed passive control strategy, e.g. FCE. The following example corresponds to

$$\boldsymbol{x}(k+1) = \begin{pmatrix} 1 & 0 \\ \bar{\theta}_1 & \bar{\theta}_2 \end{pmatrix} \boldsymbol{x}(k) + \begin{pmatrix} 1 \\ 1 \end{pmatrix} u(k) \, , \quad y(k) = (1 \;\; \bar{\theta}_3)\boldsymbol{x}(k) + \epsilon(k) \, ,$$

where $\boldsymbol{x}(0) = \boldsymbol{0}$, θ is *a priori* normally distributed $\mathcal{N}(\hat{\theta}^0, \boldsymbol{C}(0))$, with $\hat{\theta}^0 = (-0.9, -0.6, 0.5)^T$ and $\boldsymbol{C}(0) = \mathrm{diag}(0.04, 0.01, 0.03)$. The perturbations $\epsilon(k)$ are i.i.d. $\mathcal{N}(0, 0.04)$. The control objective is given by (3), with $N = 12$, $q_i = 1$ and $r_i = 0.001$ for all i. Figure 2 presents the envelopes of the output responses obtained for 100 random realisations. The continuous line gives the reference trajectory, the dotted and dashed lines respectively correspond to FCE and dual controls. Two different measurement times are considered: $t = 2$ in Figure 2(a), and $t = 8$ in Figure 2(b), which was found to be optimal for the approach above. Other examples are presented in Kulcsár *et al.* (1995b).

References

[1] D.P. Bertsekas, *Dynamic Programming and Stochastic Control*, Academic Press, New York (1976).

[2] H. Chernoff, Approaches in sequential design of experiments, In J.N. Srivastava (ed.), *A Survey of Statistical Design and Linear Models*, pages 67–90. North Holland, Amsterdam (1975).

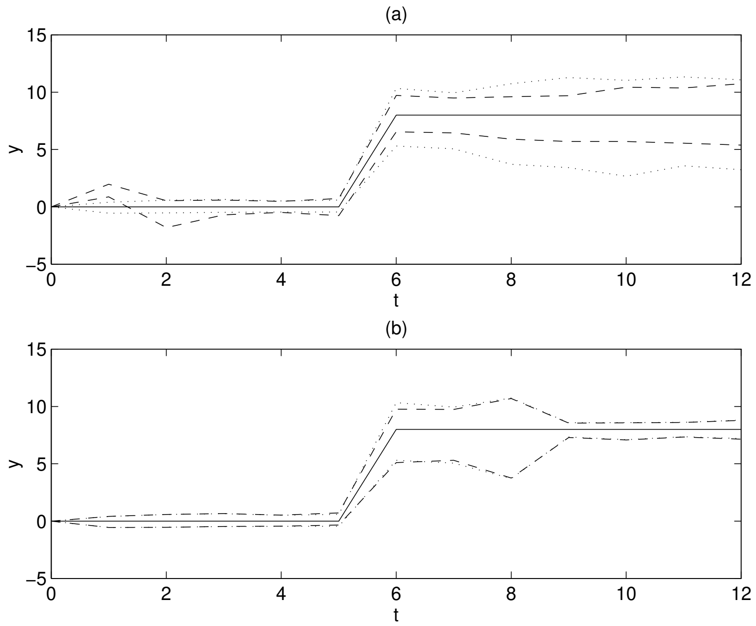

Figure 2. Envelopes of the output responses for the FCE (dotted line) and dual (dashed line) controls for two different measurement times.

[3] V.V. Fedorov, *Theory of Optimal Experiments*, Academic Press, New York (1972).

[4] C. Kulcsár, *Planification d'expériences et commande duale*, Ph.D. Thesis, Université Paris XI Orsay, Feb. 1995.

[5] C. Kulcsár, L. Pronzato, and E. Walter, Optimal experimental design and therapeutic drug monitoring. *International Journal of Bio-Medical Computing*, **36**:95–101 (1994).

[6] C. Kulcsár, L. Pronzato, and E. Walter, A dual control policy for linearly parameterized systems, in *Proc. 3rd European Control Conf.*, Rome, Sept. 95, **1**:55–60 (1995a).

[7] C. Kulcsár, L. Pronzato, and E. Walter, Experimental design for the control of linear state-space systems, submitted to the 1996 IFAC World Congress (1995b).

[8] T.J. Mitchell, An algorithm for the construction of "D-optimal" experimental designs, *Technometrics*, **16**:203–210 (1974).

[9] L. Pronzato, E. Walter and C. Kulcsár, A dynamical system approach to sequential design, in W.G. Müller, H.P. Wynn and A.A. Zhigljavsky (eds.), *Proc. MODA 3*, Petrodvorets, Russia, May 1992, Physica Verlag, Heidelberg, pp. 11–24 (1993).

[10] L. Pronzato, C. Kulcsár and E. Walter, An actively adaptive control policy based on the prediction of posterior densities, *Tatra Mountains Math. Pub.*, to appear (1995a).

[11] L. Pronzato, C. Kulcsár and E. Walter, An actively adaptive control policy for linear models, *IEEE Trans. Automatic Control*, accepted (1995b).

[12] F.E. Thau and H.S. Witsenhausen, A comparison of closed-loop and open-loop optimum systems, *IEEE Trans. Automatic Control*, **11**:619–621 (1966).

[13] E. Walter and L. Pronzato, *Identification de modèles paramétriques à partir de données expérimentales*, Masson, Paris (1994).

[14] M.B. Zarrop, *Optimal Experiment Design for Dynamic System Identification*, Springer Verlag, Heidelberg (1979).