Optimality of Index Policies for a Sequential Sampling Problem

David Castañon, Simon Streltsov, and Pirooz Vakili

***Abstract*—Consider the following sequential sampling problem: at each time, a choice must be made between obtaining an independent sample from one of a set of random reward variables or stopping the sampling. Sampling a random variable incurs a random cost at each time. The objective of the problem is to maximize the expected net difference between the largest sample reward obtained before stopping and the accumulated costs incurred while sampling. In this paper, the authors prove that the optimal feedback strategies for this problem are index policies and provide an explicit expression for the optimal expected reward from any state. The problem is motivated by search methods for global optimization problems where the cost of computation is explicitly incorporated into the objective.**

***Index Terms*—Optimal policy, sequential sampling.**

I. Introduction

Consider a sequential sampling problem where, at each time, one may either sample from a set of choices I or decide to stop the sampling. Assume that sampling from choice $i \in I$ yields a real-valued random reward Y_i with known distribution F_i at a random nonnegative cost of C_i with known distribution G_i. Assume also that rewards and costs of sampling at different sampling times are independent. The goal is to determine an optimal sequential sampling policy to maximize the expected total reward defined by

$$\mathrm{E}\left[\max\{Y(0), Y(1)\cdots, Y(N)\} - \sum_{j=1}^{N} C(j)\right]$$

where $Y(0)$ is the initial value (before sampling), N is the stopping time of sampling, and $Y(j)$ and $C(j)$ are, respectively, the reward and cost from the sampling at the jth time interval.

Chow and Robbins [3] consider this problem under the assumption that the set I is a singleton (in a somewhat more general setting); in this case, the problem reduces to determining the optimal stopping policy. They show that the optimal stopping policy is specified by a single threshold value z^* and is defined by

$$\begin{cases} \text{continue sampling}, & \text{if } \max\{Y(0), \cdots, Y(n)\} < z^* \\ \text{stop sampling}, & \text{if } \max\{Y(0), \cdots, Y(n)\} \geq z^*. \end{cases}$$

Moreover, they show that the optimal expected reward is $\max\{Y(0), z^*\}$.

Modeling a class of economic search problems, Weitzman [8] considers a version of the above problem where: 1) I is a finite set; 2) the maximum number of samples allowed from each $i \in I$, is bounded *a priori*; and 3) the maximum total number of samples allowed is equal to the sum of the maximums allowed from each option.

Weitzman shows that the optimal sampling policy is an index policy where the index for each option i, say z_i^*, is identical (in the undiscounted reward case) to that derived in [3] assuming an infinite horizon and that the sole option is sampling from option i. The optimal policy derived by Weitzman is defined as follows. *Stopping rule*: stop sampling if the current maximum reward is higher or equal to z_i^* for all options from which sampling is still allowed; *selection rule*: if the decision is to continue sampling, sample from the option with the highest index z_i^* among all those from which sampling is still allowed.

The similarity of our setting to that of multi-armed bandit problems is apparent (see, for example, [4]). However, the difference in the reward structure renders the results derived in that setting inapplicable in ours.

Our motivation for considering the above sequential sampling problem is to obtain guidelines for deriving and analyzing global optimization algorithms where, in evaluating the performance of the algorithm, the cost of computation is directly taken into account. A finite horizon version of our problem was considered by Tang [6] in the case of partitioned random search for global optimization. Tang [2] considers a case where the maximum number of samples allowed is specified and where I, the number of regions to sample from, is a finite set. This problem does not seem to have a simple index policy solution (Castañon and Tang [6] give a counterexample to the solution offered in [6]).

The setting of this paper can be used more generally, where the set of choices, for example, may include different algorithms or subroutines for solving some of the iterations of an optimization problem. The set I should be viewed as representing all or a subset of available options in the search for a global optimum.

In this paper, we assume that I is a compact set in some appropriate topology (generally a subset of R^n) and that the horizon is infinite. Given the ultimate goal of solving a global optimization problem, we assume the existence of a uniform upper bound $M < \infty$ on all sample rewards (so that $P(Y_i \leq M) = 1$ for all $i \in I$.) The main result of the paper is that under conditions the optimal sampling policy is a pure index policy based on the indexes z_i^* (defined above), and the optimal reward can be expressed in terms of these indexes. Indeed, our results indicate that the optimal sampling policy should sample from one option only, the one with the maximum index.

The problem studied in this paper assumes a complete statistical knowledge about the reward and cost of sampling from each option (knowledge of the distributions F_i and G_i for each $i \in I$.) This assumption is rarely satisfied in practice. Our results can be used to generate adaptive sampling strategies for cases where these distributions are estimated from data, as in Tang [6]. Streltsov and Vakili [7] use the z_i^* indexes (recalculated after each sample) to modify a known global optimization algorithm which is based on a statistical model of the reward and report overall performance improvements.

The rest of the paper is organized as follows. In Section II we define the sampling problem of interest. In Section III, we develop a dynamic programming formulation of this problem, derive a sufficient statistic for control, and give an equivalent formulation of the problem as a minimization of cost problem. In Section IV, we define global indexes and establish the optimality of a candidate index policy and develop an expression for the optimal cost.

II. Mathematical Formulation

Let I be a set of choices. Assume that sampling from $i \in I$ yields a real-valued random reward with known distribution F_i and costs a random (nonnegative) real-valued amount with known distribution

Manuscript received November 18, 1996. This work was supported in part by the NSF under Grant EID-9527422 and the AFOSR under Grant F49620-96-1-0028.

D. Castañon is with the Electrical and Computer Engineering Department, Boston University, Boston, MA 02215 USA.

S. Streltsov and P. Vakili are with the Manufacturing Engineering Department, Boston University, Boston, MA 02215 USA.

Publisher Item Identifier S 0018-9286(99)00596-6.

G_i. We assume that the distributions G_i have finite means for all $i \in I$. Assume that rewards and costs of consecutive samplings are independent for all sampling policies. We consider the following sequential sampling problem.

Given n samples, denoted by $Y(1), \cdots, Y(n)$, and an initial value $Y(0)$, we may make one of two types of decisions: 1) we may sample from choice $i \in I$, in which case the new sample, $Y(n+1)$, has distribution F_i, and a cost of sampling with distribution G_i is incurred, or, 2) we may stop sampling and collect a reward equal to $\max\{Y(0), \cdots, Y(n)\}$.

Let $Y^n = (Y(0), \cdots, Y(n)) \in R^{n+1}$. A policy $\pi = \{\eta_0, \eta_1, \cdots\}$ is defined by a sequence of functions $\eta_n: R^{n+1} \to I \cup \{\sigma\}$ where, $\eta_n(\cdot) = i \in I$ represents the decision to sample from choice i, and $\eta_n(\cdot) = \sigma$ represents the decision to stop.

Let $N = \min\{n: \eta_n(Y^n) = \sigma\}$ denote the stopping instance (let $N = \infty$, if $\eta_n \neq \sigma$ for all $n \geq 0$). Our objective is to find a sequential sampling policy that maximizes the expected total reward defined by

$$R^\pi(y) = \mathrm{E}\left[-\sum_{n=0}^{N-1} C_{\eta_n(Y^n)} + \max\{Y(0), \cdots, Y(N)\}\right]$$

where $Y(0) = y$.

We make the following assumptions.

A. Bounded/Compact/Continuous (BCC) Assumptions

1) There exists a real number $M > 0$ such that $F_i(x) = 1$ when $x > M$ for all $i \in I$. That is, all sample rewards are uniformly bounded by M.
2) I is a compact set in a topology defined on I.
3) The functions $\Psi_1: I \to \{F_i; i \in I\}$ $(\Psi_1(i) = F_i)$ and $\Psi_2: I \to \{G_i; i \in I\}$ $(\Psi_2(i) = G_i)$ are continuous (weak topology on $\{F_i; i \in I\}$ and $\{G_i; i \in I\}$).

Note that parts 2) and 3), i.e., compactness and continuity assumptions, are trivially valid when I is a finite set.

To solve the above problem, we will formulate it as a dynamic programming problem, as described next.

III. Dynamic Programming Formulation

Consider the sequential sampling problem described previously. For $n \geq 0$, define

$$S_n = R^{n+1} \cup \{s\} = \text{the state space at period } n$$
$$U_n = U = I \cup \{\sigma\} = \text{the control space at period } n$$
$$g_n: S_n \times U \to R = \text{the period reward for period } n$$

where s is an absorbing state that is reached when sampling is stopped. and $\eta_n: S_n \to U$, with the stipulation that $\eta_n(s) = \sigma$ for all n for all admissible policies π. Let $X_n \in S_n$ denote the state of the system at stage n. Define the initial state $X_0 = Y(0)$. With this notation, the state evolution is described recursively as

$$X_{n+1} = f_n(X_n, \eta_n)$$

where the function f_n is defined as

$$f_n(X_n, i) = \begin{cases} (X_n, Y(n+1)) \in R^{n+1}, \\ \quad \text{if } i \in I, \text{ where } Y(n+1) \text{ has} \\ \quad \text{distribution } F_i \\ s, \quad \text{if } i = \sigma. \end{cases}$$

The reward associated with period n is defined by $g_n(X_n, i)$, where

$$g(X_n, i) = \begin{cases} -C_i & \text{for } i \in I \\ \max\{Y(0), \cdots, Y(n)\}, \\ & \text{for } X_n = Y^n \in R^{n+1}, i = \sigma \\ 0, & \text{if } X_n = s. \end{cases}$$

The reward associated with a policy π is the expected infinite horizon reward given by

$$R^\pi(y) = \mathrm{E}\left[\sum_{n=0}^{\infty} g_n(X_n, \eta_n(X_n))\right]$$

where $X_0 = y$.

Now, define the statistic

$$Z_n = \Psi(X_n) = \begin{cases} \max\{Y(0), \cdots, Y(n)\}, & \text{if } X_n = Y^n \\ s, & \text{if } X_n = s. \end{cases}$$

Before sampling is stopped, Z_n represents the maximum value observed so far ($Z_n = X_n = s$ after sampling is stopped.) This statistic captures all the relevant past information and is all that is needed to make the next decision. That is to say, Z_n is a sufficient statistic for control for the above dynamic programming problem.

To establish this fact, we first formulate an alternative dynamic programming problem in terms of Z_n. (To avoid cumbersome notation, we will use similar notation as above for the corresponding terms in the following dynamic programming problem.)

For $n \geq 0$, let

$$S = R \cup \{s\} = \text{the state space at period } n$$
$$U = I \cup \{\sigma\} = \text{the control space at period } n$$
$$g: S \times U \to R = \text{the period reward for each stage } n$$

where s and σ are as defined above. Define decision rules $\tilde{\eta}_n: S \to U$, and let $\tilde{\pi}$ denote admissible policies $\{\tilde{\eta}_0, \tilde{\eta}_1, \cdots\}$ satisfying $\tilde{\eta}_n(s) = \sigma$ for all n. For any policy $\tilde{\pi}$, the system dynamics are given by

$$Z_{n+1} = f(Z_n, \tilde{\eta}_n(Z_n))$$

where

$$f(z, i) = \begin{cases} \max\{z, Y(n+1)\}, \\ \quad \text{for } i \in I, \text{ and } Y(n+1) \text{ has} \\ \quad \text{distribution } F_i \\ s, \quad \text{if } i = \sigma. \end{cases}$$

The period reward is defined by

$$g(z, i) = \begin{cases} -C_i, & \text{for } i \in I \\ z, & \text{for } i = \sigma, z \neq s \\ 0, & \text{for } z = s. \end{cases}$$

The reward associated with a policy $\tilde{\pi}$ is given by

$$R^\pi(y) = \mathrm{E}\left[\sum_{n=0}^{\infty} g(Z_n, \tilde{\eta}_n(Z_n))\right]$$

where $Z_0 = y$. Note that the system dynamics and the period reward functions for the second dynamic programming problem are stationary.

Proposition 1: The sequence $\{Z_0, Z_1, \cdots\}$ is a sufficient statistic for control for the original dynamic programming problem.

Proof: Note that the control spaces for the two problems are identical and hence by moving to the sequence $\{Z_0, Z_1, \cdots\}$ no new control constraints are imposed. Moreover, it is simple to verify that for any policy $\pi = \{\eta_0, \eta_1, \cdots\}$, the distribution of Z_{n+1} is uniquely determined by the values of $Z_n = \Psi(X_n)$ and the decision $\eta_n(X_n)$. And, finally, it follows directly from the above definitions that for all policies

$$\mathrm{E}[g_n(X_n, \eta_n(X_n)) | \Psi(X_n) = z, \eta_n(X_n) = i)] = g(z, i).$$

Therefore, the sequence $\{Z_0, Z_1, \cdots\}$ is a sufficient statistic for control for the original dynamic programming problem (see [1, Ch. 10]). □

To simplify notation throughout the rest of this paper, we use π and η_n to denote admissible strategies in the modified optimization problem using the sufficient statistic Z_n. As a last step in this preliminary section, we reformulate the problem of maximization of reward as an equivalent minimization of cost problem. The advantage of this new formulation is that the period costs are always nonnegative, and hence the problem falls into one of the classical cases for which characterizations for the optimal cost and the optimal policy exist.

Let the period cost h: $S \times U \rightarrow R^+$ (we use the same notation g) be defined by

$$h(z,i) = \begin{cases} C_i, & \text{if } i \in I \\ M - z, & \text{if } i = \sigma, z \neq s \\ 0, & \text{if } z = s \end{cases}$$

and consider the problem of determining the optimal policy π to minimize the expected total cost

$$J^\pi = \mathrm{E}\left[\sum_{n=0}^{\infty} h(Z_n, \eta_n(Z_n))\right].$$

It is simple to verify that this minimization of cost problem is equivalent to the above maximization of the reward problem. With this reformulation, therefore, we have an undiscounted infinite-horizon dynamic programming problem with nonnegative period costs.

IV. Optimality of Index Policies

In this section we propose an index policy and show that it is the optimal sampling policy. We begin with defining the indexes which form the basis for this policy.

A. Global Indexes

For $i \in I$, define the function I_i as

$$I_i(z) = E[(Y_i - z)^+] - E[C_i] = \int_z^\infty \overline{F}_i(x)\, dx - c_i$$

where $x^+ = \max\{x, 0\}, \overline{F}_i(x) = 1 - F_i(x)$, and $c_i = E[C_i]$.

$I_i(z)$ represents a *myopic* index associated with choice $i \in I$, given a "current value" z. This index evaluates the one-step implication of the decision to take *one more* sample from choice i when $Z_n = z$. More specifically, it gives the expected incremental reward of taking one more sample from choice $i \in I$ and stopping at the next decision period. ($-I_i(z)$ gives the expected incremental cost.)

It is simple to verify the following.

Lemma 1: For all $i \in I$: 1) I_i is a continuous function; 2) $\lim_{z \to -\infty} I_i(z) = \infty$, and $I_i(z) = -c_i$ for $z \geq M$; and 3) I_i strictly decreases from ∞ to $-c_i$.

Define a *global* index, z_i^*, associated with $i \in I$ as follows: z_i^* is the unique solution of $I_i(z_i^*) = 0$ or, equivalently

$$\int_{z_i^*}^\infty \overline{F}_i(x)\, dx = c_i.$$

In light of Lemma 1, the existence and uniqueness of z_i^* is guaranteed.

The global character of the index z_i^* can be seen from the following result (due to Chow and Robbins [3]). Consider the optimal stopping time problem of sampling from choice i only. Let the stationary policy $\pi_i^* = \{\eta_i, \eta_i, \cdots\}$ be defined by

$$\eta_i(z) = \begin{cases} i, & \text{if } z > z_i^* \\ \sigma, & \text{if } z \geq z_i^*. \end{cases}$$

Then we have the following.

Lemma 2: π_i^* is the optimal stopping policy when the set $I = \{i\}$ contains only one element and the optimal reward function R_i^* is defined by $R_i^*(z) = \max\{z, z_i^*\}$; equivalently, in our minimization problem formulation, the optimal cost function is defined by $J_i^*(z) = M - \max\{z, z_i^*\}$.

Example: In this example, we give a few samples of the index z^*.

1) Consider the discrete r.v. Y defined by $P(Y = 10) = 0.5$ and $P(Y = 2) = 0.5$. It is easy to verify that for $c = 8, z^* = -3.5$, and for $c = 3, z^* = 2.5$.
2) Let Y be an exponential random variable with mean μ. Then, for $c \leq \mu$, $z^* = \mu ln(c/\mu)$ and for $c > \mu, z^* = \mu - c$.
3) Let Y be a normal random variable with mean μ and standard deviation σ. Then $z^*(c; \mu, \sigma) = \mu + \sigma z^*(c/\sigma; 0, 1)$ where $z^*(c/\sigma; 0, 1)$ is the solution to the equation
$$\phi(z) - z(1 - \Phi(z)) = \frac{c}{\sigma}$$
and ϕ and Φ are, respectively, the standard normal density and the distribution. For more details on computing indexes for the normal case, see [5].

B. The Optimal Policy

Let

$$z^* = \max\{z_i^*; i \in I\},$$
$$i^* = \operatorname{argmax}\{z_i^*; i \in I\}.$$

Under BCC assumptions the index z_i^* is a continuous function of i. Then, compactness of I guarantees the existence and uniqueness of z^*, and the existence of i^*.

Now, define η^*: $S = R \cup \{s\} \rightarrow U = I \cup \{\sigma\}$ by

$$\eta^*(z) = \begin{cases} i^*, & z < z^* \\ \sigma, & z \geq z^*, \text{ or } z = s \end{cases}$$

and define the stationary policy $\pi^* = (\eta^*, \eta^*, \cdots)$. Note that $\pi^* = \pi^{i^*}$ as defined above. Therefore, Lemma 2 implies that the cost associated with policy π^*, denoted by J^*, is given by $J^*(z) = M - \max\{z, z^*\}$. Our main result is that the optimal policy and the optimal cost function for the sequential sampling problem are, respectively, π^* and J^*.

Let $\hat{J}$ denote the optimal cost function. Clearly $\hat{J}(s) = 0$ (the same is true of all policies, i.e., $J^\pi(s) = 0$ for all π). Therefore, in what follows, we focus on characterizing $\hat{J}$ on R only. The next proposition provides a partial characterization.

Lemma 3: $\hat{J}(z) = M - z$ for $z \geq z^*, z \in R$.

Proof: Consider any sampling policy $\pi = \{\eta_1, \eta_2, \cdots\}$. The cost of this policy, starting from state $Z_0 = z$, can be written as

$$J^\pi(z) = M - z + \mathrm{E}\left[\sum_{n=0}^{\infty} h'(Z_n, \eta_n(Z_n))\right]$$

where h' is defined by

$$h'(z,i) = \begin{cases} C_i - (Y_i - z)^+, & z \in R, i \in I \\ 0, & \text{if } i = \sigma. \end{cases}$$

In other words, a cost equal to $M - z$ is incurred before sampling begins; all subsequent costs are incremental costs of sampling and there is no cost for stopping. Note that

$$\begin{aligned}\mathrm{E}[h'(Z_n, \eta_n(Z_n))] &= \mathrm{E}[\mathrm{E}[h'(Z_n, \eta_n(Z_n))|Z_n \neq s, \eta_n(Z_n) = i \in I]] \\ &= \mathrm{E}[\mathrm{E}[C_i - (Y_i - Z_n)^+|Z_n \neq s, \eta_n(Z_n) = i \in I]].\end{aligned}$$

Clearly, $Z_n \geq z$ for all n and all sample paths; therefore

$$\begin{aligned}\mathrm{E}[h'(Z_n, \eta_n(Z_n))] &\geq \mathrm{E}[\mathrm{E}[C_i - (Y_i - z)^+|Z_n \neq s, \eta_n(Z_n) = i \in I]] \\ &= \mathrm{E}[\mathrm{E}[-I_i(z)|Z_n \neq s, \eta_n(Z_n) = i \in I]].\end{aligned}$$

From the definition of z^* we have $I_i(z) \leq 0$ for all $i \in I$ and $z \geq z^*$. Thus, $\mathrm{E}[h'(Z_n, \eta_n(Z_n))] \geq 0$ for all n. Hence, for all sampling policies, we have $J^{\pi}(z) \geq M - z$ if $z \geq z^*$. This establishes the lemma, because this lower bound can be achieved by the policy $\eta_0(z) = \sigma$. □

Let J: $R \rightarrow R$ be a real-valued function. Let T be the dynamic programming operator, defined as

$$\begin{aligned} T(J)(z) &= \min\{M - z, \min_{i \in I}\{\mathrm{E}[h(z,i) + J(f(z,i))]\}\} \\ &= \min\{M - z, \min_{i \in I}\{c_i + \mathrm{E}[J(\max(Y_i, z))]\}\} \end{aligned}$$

where Y_i is an independent sample from the distribution F_i.

We have the following result.

Lemma 4: Let J: $R \rightarrow R^+$ be a monotone nonincreasing function. Then, $T(J)$ is also monotone nonincreasing.

Proof: Let $z_1, z_2 \in R$ and $z_1 \geq z_2$. Clearly, $\max(Y_i, z_1) \geq \max(Y_i, z_2)$ for each $i \in I$. Thus,

$$J(\max(Y_i, z_1)) \leq J(\max(Y_i, z_2)).$$

In light of the above equation for T, we have

$$T(J)(z_1) \leq T(J)(z_2)$$

establishing that $T(J)$ is monotone nonincreasing.

Lemma 5: Under BCC assumptions $\hat{J}$ is monotone nonincreasing on R.

Proof: Compactness of I implies compactness of the control space $U = I \ \cup \{\sigma\}$. This compactness and the continuity assumptions of BCC imply that the dynamic programming problem under consideration satisfies the assumptions of the semicontinuous model of [1, Ch. 9], and hence it follows that the iteration $J^{n+1} = T(J^n)$, initialized with the initial condition $J^0(z) \equiv 0$, converges monotonically to the optimal cost $\hat{J}$ (see [1, Proposition 9.17 and Corollary 9.17.2]). From Lemma 4, since the initial condition is monotone nonincreasing, J^n is also monotone nonincreasing for all n, establishing that the limit $\hat{J}$ will also be monotone nonincreasing. □

We can now show the main result of the paper.

Theorem 1: Under BCC assumptions the sampling policy π^* is the optimal policy and the function J^* is the optimal cost function.

Proof: From Lemma 3 we have $\hat{J}(z) = J^*(z)$ for $z \geq z^*$.

Let $z < z^*$. Then, by Lemma 5, the optimal cost is monotone nonincreasing, so

$$\hat{J}(z) \geq \hat{J}(z^*) = M - z^* = J^*(z).$$

On the other hand, since $\hat{J}$ is the minimum cost

$$\hat{J}(z) \leq J^*(z).$$

Hence $\hat{J}(z) = J^*(z)$ for $z < z^*$ and we have shown that $\hat{J} = J^*$.

According to Lemma 2, J^* is the cost function corresponding to policy π^*, therefore, π^* is the optimal policy. □

Remark 1—Vector Sampling: The above result extends to the case of vector sampling (or parallel sampling). Assume that at each stage instead of one sample, d $(d > 1)$ samples are taken (possibly in parallel). For each $\underline{i} = (i_1, \cdots, i_d) \in I^d$, let $C_{\underline{i}}$ represent the cost of taking d samples from choices $i_1, \cdots, i_d$ and let $W_{\underline{i}}$ represent the maximum of the samples taken. The distribution of $W_{\underline{i}}$ is given by $F_{W_{\underline{i}}} = \Pi_{j=1}^d F_{Y_{i_j}}$.

The sequential sampling problem discussed in this paper is equivalent to this problem, posed in terms of collecting samples $W_{\underline{i}}$ given the distribution of $C_{\underline{i}}$ and the derived distributions $F_{W_{\underline{i}}}$. As a result, a consequence of the above theorem is that the optimal sampling scheme is to always sample from the same combination of choices, i.e., the combination that produces the highest index (note that the d choices may not all be the same).

References

[1] D. Bertsekas and S. Shreve, *Stochastic Optimal Control: Discrete Time Case*, Mathematics In Science and Engineering Series, vol. 139. New York: Academic, 1978.

[2] D. A. Castañon and Z. Tang, "A counterexample to adaptive partitioned random search to global optimization," *IEEE Trans. Automat. Contr.*, vol. 41, p. 310, 1996.

[3] Y. Chow and H. Robbins, "A martingale systems theorem and applications," in *Proc. 4th Berkeley Symp. Mathematical Statistics and Probability*, 1961, vol. 1, pp. 93–104.

[4] J. Gittins, *Multi-Armed Bandit Allocation Indices*, 1st ed. New York: Wiley, 1989.

[5] D. B. Rosenfield, "Optimal strategies for selling an asset," *Management Sci.*, vol. 29, no. 9, pp. 1051–1058, 1983.

[6] Z. Tang, "Adaptive partitioned random search to global optimization," *IEEE Trans. Automat. Contr.*, vol. 39, pp. 2235–2244, Nov. 1994.

[7] S. Streltsov and P. Vakili, "A nonmyopic utility function for statistical global optimization," to be published.

[8] M. L. Weitzman, "Optimal search for the best alternative," *Econometrica*, vol. 47, no. 3, pp. 641–654, 1979.

On Regularizing Singular Systems by Decentralized Output Feedback

Dianhui Wang and C. B. Soh

***Abstract*— This paper considers linear time-invariant decentralized singular systems which are either nonregular or, if they are regular, they have impulsive modes. It derives algebraic necessary and sufficient conditions for making a singular system both regular and impulse-free by decentralized output feedback control laws and decentralized proportional-plus-derivative output feedback control laws.**

***Index Terms*—Decentralized control, regularization, singular systems, output feedback.**

I. Introduction

Consider the following decentralized singular systems:

$$\begin{aligned} E\dot{x} &= Ax + \sum_{i=1}^{N} B_i u_i \\ y_i &= C_i x, \qquad i = 1, 2, \cdots, N \end{aligned} \tag{1}$$

where E and A are $n \times n$ real matrices with E singular, x and y_i are state vector and outputs vectors, respectively, B_i and C_i are $n \times m_i$ real matrices and $l_i \times n$ real matrices, respectively.

System (1) is said to be regular if the pencil pair $sE - A$ is regular, i.e., $\det(sE - A)$ is not identically zero. It is well known that regularity of singular systems guarantees the existence and uniqueness of the solutions [1], [2]. Almost all of the given results

Manuscript received March 20, 1996; revised January 1, 1997.

D. Wang is with the Institute of Intelligent Control, Dalian Maritime University, Dalian, Liaoning, 116026 China (e-mail: dhwang@dlmu.edu.cn).

C. B. Soh is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798 Singapore.

Publisher Item Identifier S 0018-9286(99)00574-7.