strikingly superior to the individual-optimum rule; if the dissimilarities are limited to service rates of the servers, the "separable" rule practically always does better than the individual-optimum rule; on the other hand, if the individual servers are all alike in the various queues, but the queues differ in their numbers of servers, then the individual-optimum rule tends to do better than the "separable" rule. From the results derived in [8] and in the present paper, there emerges a common analytical framework for routing studies of voice traffic and data traffic.

## References


[1] R. B. Cooper, *Introduction to Queueing Theory*, 2nd Ed. Amsterdam: North-Holland, 1981.

[2] A. Ephremides, P. Varaiya, and J. Walrand, "A simple dynamic routing problem," *IEEE Trans. Automat. Contr.*, vol. AC-25, no. 4, pp. 690–693, 1980.

[3] R. G. Gallager, "A minimum delay routing algorithm using distributed computation," *IEEE Trans. Commun.*, vol. COM-25, no. 1, pp. 73–85, 1977.

[4] A. Harel and P. Zipkin, "Strong convexity results for queueing systems," *Oper. Res.*, vol. 35, no. 3, pp. 405–418, 1987.

[5] R. A. Howard, *Dynamic Programming and Markov Processes*. Cambridge, MA: M.I.T. Press, 1960.

[6] *IEEE Trans. Commun.*, vol. COM-29. Apr. 1981.

[7] K. R. Krishnan, "Joining the right queue: A Markov decision-rule," in *Proc. 26th IEEE Conf. Decision Contr.*, Los Angeles, CA, Dec. 1987, pp. 1863–1868.

[8] K. R. Krishnan and T. J. Ott, "State-dependent routing for telephone traffic: Theory and results," in *Proc. 25th IEEE Conf. Decision Contr.*, Athens, Greece, Dec. 1986, pp. 2124–2128.

[9] S. Stidham, Jr., "Optimal control of admission to a queueing system," *IEEE Trans. Automat. Contr.*, vol. AC-30, no. 6, pp. 705–713, 1985.

[10] R. Weber, "On the optimal assignment of customers to parallel servers," *J. Appl. Prob.*, vol. 15, pp. 406–413, 1978.

[11] W. Winston, "Optimality of the shortest-processing-time discipline," *J. Appl. Prob.*, vol. 14, pp. 181–189, 1977.

[12] T. Yum and M. Schwartz, "The join-biased-queue rule and its application to routing in computer communication networks," *IEEE Trans. Commun.*, vol. COM-29, pp. 505–511, 1981.


# Large-Scale Convex Optimal Control Problems: Time Decomposition, Incentive Coordination, and Parallel Algorithm

TSU-SHUAN CHANG, XIAO-XUAN JIN, PETER B. LUH, AND XIYI MIAO

*Abstract*—**Based on the time decomposition and incentive coordination, a parallel algorithm is developed for long horizon optimal control problems. This is done by first decomposing the original problem into subproblems with shorter time horizon, and then using the incentive coordination scheme to coordinate the interaction of subproblems. For strictly convex problems, it is proved that the decomposed problem with linear incentive coordination is equivalent to the original problem, in the sense that each optimal solution of the decomposed problem produces one global optimal solution of the original problem and vice versa. In other words, linear incentive terms are sufficient in this case and impose no additional computation burden on the subproblems. The high-level parameter optimization problem is shown to be nonconvex, despite the uniqueness of the optimal solution and the convexity of the original problem. Nevertheless, the high-level problem has no local minimum, even though it is nonconvex. A parallel algorithm based on a prediction method is developed and a numerical example is used to demonstrate the feasibility of the approach.**

Manuscript received February 11, 1987; revised October 26, 1987. Paper recommended by Past Associate Editor, D. A. Castanon. This work was supported in part by NSF Grants ECS-85-12815, ECS-85-04133, and ECS-85-13163.

T.-S. Chang and X.-X. Jin are with the Department of Electrical Engineering and Computer Science, University of California, Davis, CA 95616.

P. B. Luh and X. Miao are with the Department of Electrical and Systems Engineering, University of Connecticut, Storrs, CT 06268.

IEEE Log Number 8931891.

## I. Introduction

Many problems can be formulated as optimal control problems. In principle, they can be solved by using dynamic programming to obtain optimal solutions. However, due to the curse of dimensionality, this approach is in general computationally infeasible (e.g., [7]). For many large-scale problems, the long time horizon adds another dimension of difficulty besides the large number of variables and constraints. In this paper, we consider a decomposition and coordination scheme that decomposes a large problem into manageable subproblems along the time axis. Low-level subproblems can be solved simultaneously by parallel processing, and a corresponding coordination scheme is created. In other words, a hierarchical structure is imposed upon the long horizon problem to facilitate the application of parallel processing, and to improve computational efficiency. Note that state decomposition can be carried out inside each subproblem when needed. In this sense, our approach is complementary to existing results.

The concept of time decomposition and incentive coordination was originally motivated by power system scheduling problems [5]. In [2], the idea was formulated and a parallel algorithm was developed. This paper lays the theoretical foundation for incentive coordination for convex problems. One relevant approach is the target coordination scheme proposed in [3]. Both approaches decompose the long time horizon problem into subproblems according to time axis, but with different coordination schemes. Numerical testing and an application of the target coordination can be found in [4].

The paper is organized as follows. In Section II, the problem and its hierarchical decomposition framework are stated. In Section III, the equivalence of the original strictly convex problem and the decomposed problem with linear incentive coordination is proved. The nonconvexness of the high-level problem is illustrated by a simple example. In Section IV, a parallel algorithm is developed, and a numerical example is presented to demonstrate the feasibility of the parallel algorithm. A short discussion is then given in Section V.

## II. Problem Statement

Consider the following convex optimal control problem:

$$(P): \quad \min_{\substack{u_i,\\ 0\le i<N}} J = g_N(x_N) + \sum_{i=0}^{N-1} g_i(x_i, u_i) \tag{2.1}$$

subject to the system dynamics

$$x_{i+1} = A_i x_i + B_i u_i,\ i = 0, \cdots, N-1 \quad \text{where } x_0 \text{ is given,} \tag{2.2}$$

$$x_i \in R^n, \text{ and } u_i \in R^m. \tag{2.3}$$

We assume that functions $g_i$'s are strictly convex, twice continuously differentiable, and their function values go to infinity when the magnitude of their variables approaches infinity. Thus, the optimal solution exists and is unique.

We now decompose $(P)$ into $M$ subproblems with incentive coordination as in [2], each having the shorter time horizon so that each subproblem can be solved efficiently. Let $T$ denote the number of stages in each subproblem, and assume $N = MT$ for simplicity. We have

$(P-j)$ Subproblem $j$, $j = 1, \cdots, M$:

$$\min_{\substack{u_k\\ (j-1)T\le k<jT}} J_j = \sum_{i=(j-1)T}^{jT-1} g_i(x_i, u_i) + I_j(x_{jT}), \tag{2.4}$$

subject to corresponding system dynamics

$$x_{k+1} = A_k x_k + B_k u_k,\ (j-1)T \le k < jT \tag{2.5}$$

where $x_{(j-1)T}$ is given.

Note that $I_j$ is a function of $x_{jT}$ only, the terminal state of each subproblem. We also set $I_M(x_{MT}) \equiv g_N(x_N)$ for subproblem $(P-M)$. The higher level's problem $(P-H)$ is to select the best coordination terms to minimize the total cost, i.e.,

$$(P-H) \quad \min_{\substack{I_j \\ 1\le j\le M-1}} J, \qquad \text{with } J=\sum_{j=1}^{M} J_j-\sum_{j=1}^{M-1} I_j = g_N(x_N^*)+\sum_{i=0}^{N-1} g_i(x_i^*, u_i^*), \quad (2.6)$$

where $x_i^*$ and $u_i^*$ are solutions of $(P-j)$.

## III. Linear Incentive Coordination

In this section, we shall show that linear incentive terms are sufficient for the convex optimal control problems, that is, we need only to consider

$$I_j(x_{jT})=p_j^T x_{jT}, \qquad j=1,\cdots,M-1. \quad (3.1)$$

We shall assume that the subproblems are solved sequentially. More specifically, we first solve Problem $(P-j)$ with the given initial state $x_{(j-1)T}$ and to find the optimal final state $x_{jT}^*$. The $x_{jT}^*$ is then used as the initial state for Problem $(P-(j+1))$. The assumption will be removed in the next section to develop a parallel algorithm.

Let us denote the set $\{x_{i+1},\cdots,x_j\}$ by $x_{(i,j]}$. Note that if an index is associated with a parenthesis, it is not included in the set. The index associated with a bracket is included. The same notation is used in $u_{(i,j]}$, etc. For a given $p_j$, the optimal solution of $(P-j)$ is denoted by $x^*_{((j-1)T,jT]}(x_{(j-1)T},p_j)$ and $u^*_{[(j-1)T,jT)}(x_{(j-1)T},p_j)$. In the notations used, the superscript * is used to indicate that the designated variable is the optimal solution of the associated problems. The listed variable in the parentheses is then used to indicate which (sub)problem it refers to as to be seen later. Other notations used are as follows.

$$\frac{\partial J}{\partial x}=\left[\frac{\partial J}{\partial x_1}\cdots\frac{\partial J}{\partial x_n}\right]^T, \quad (3.2a)$$

$$\frac{\partial y}{\partial x}=\begin{bmatrix}\frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_1}{\partial x_n}\\ \cdots & & \\ \frac{\partial y_m}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_n}\end{bmatrix}, \quad (3.2b)$$

where

$$y=[y_1\cdots y_m]^T.$$

We can now state the main theoretical result of this paper in Theorem 1. Those who are only interested in the parallel algorithm can skip the proof of the theorem.

*Theorem 1:* Problems $(P)$ and $(P-H)$ are equivalent in the following sense.

a) If $x^*_{(0,N]}$ and $u^*_{(0,N]}$ is the optimal solution of $(P)$, then there exists a $p^*_{[1,M)}$ such that

$$x^*_{((j-1)T,jT]}=x^*_{((j-1)T,jT]}(x^*_{(j-1)T},p_j^*), \quad (3.3a)$$

$$u^*_{[(j-1)T,jT)}=u^*_{[(j-1)T,jT)}(x^*_{(j-1)T},p_j^*), \quad (3.3b)$$

and the $p^*_{[1,M)}$ is an optimal solution of $(P-H)$.

b) Let $p^*_{[1,M)}$ be an optimal solution of $(P-H)$, and denote

$$x^*_{((j-1)T,jT]}\equiv x^*_{((j-1)T,jT]}(x^*_{(j-1)T},p_j^*), \quad (3.4a)$$

$$u^*_{[(j-1)T,jT)}\equiv u^*_{[(j-1)T,jT)}(x^*_{(j-1)T},p_j^*); \quad (3.4b)$$

then $x^*_{(0,N]}$ and $u^*_{[0,N)}$ is the optimal solution of $(P)$.

*Proof:* The theorem is proved by following the lemmas. The result a) comes directly from Lemma 3 and b) comes from Lemmas 4 and 7.

A state trajectory $x_{[0,N]}$ is admissible if there exists an admissible $u_{[0,N)}$ such that $x_{[0,N]}$ satisfies the state dynamics. A pair $(x_i, x_j)$ is called admissible if there exists at least an admissible $x_{[i,j]}$. For a given admissible state trajectory $x_{[0,N]}$, we define subproblem $(T-j)$ as follows.

Problem $(T-j)$, $j=1,\cdots,M-1$:

$$\min_{\substack{u_k\\(j-1)T\le k<jT}} \hat{J}_j(x_{((j-1)T,jT]}, u_{[(j-1)T,jT)})=\sum_{i=(j-1)T}^{jT-1} g_i(x_i,u_i), \quad (3.5a)$$

$$\text{subject to } x_{i+1}=A_i x_i+B_i u_i, \qquad i=(j-1)T,\cdots,jT-1, \quad (3.5b)$$

$$\text{where } x_{(j-1)T} \text{ and } x_{jT} \text{ are given.} \quad (3.5c)$$

Problem $(T-M)$:

$$\min_{\substack{u_k\\(M-1)T\le k<MT}} \hat{J}_M(x_{((M-1)T,MT]}, u_{[(M-1)T,MT)}) = \sum_{i=(M-1)T}^{MT-1} g_i(x_i,u_i)+g_N(x_N) \quad (3.5d)$$

$$\text{subject to } x_{i+1}=A_i x_i+B_i u_i, \qquad i=(M-1)T,\cdots,MT-1 \quad (3.5e)$$

$$\text{where } x_{(M-1)T} \text{ is given} \quad (3.5f)$$

*Lemma 1:* For a given admissible $x_{[0,N]}$, the optimal control and state trajectory of $(T-j)$, $j=1,\cdots,M-1$, always exists, denoted by $u^*_{[(j-1)T,jT)}(x_{(j-1)T},x_{jT})$ and $x^*_{((j-1)T,jT]}(x_{(j-1)T},x_{jT})$, and can be found by solving

$$\frac{\partial g_i}{\partial u_i}+B_i^T\lambda_{i+1}=0, \quad (3.6a)$$

$$\frac{\partial g_i}{\partial x_i}+A_i^T\lambda_{i+1}-\lambda_i=0, \quad i=(j-1)T+1,\cdots,jT-1, \quad (3.6b)$$

$$x_{i+1}=A_I x_i+B_i u_i, \qquad i=(j-1)T,\cdots,jT-1 \quad (3.6c)$$

$$\text{where } x_{(j-1)T} \text{ and } x_{jT} \text{ are given.} \quad (3.6d)$$

*Proof:* Since $(T-j)$ is a strictly convex optimization problem, its solution is unique and can be found from the necessary condition (3.6) derivable from the calculus of variations.

*Lemma 2:* If $(x_{(j-1)T}, x_{jT})$ is admissible, then there exists a $p_j^*$ such that the optimal solution of $(P-j)$ with $x_{(j-1)T}$ as the initial condition is

$$u^*_{[(j-1)T,jT)}(x_{(j-1)T},p_j^*)=u^*_{[(j-1)T,jT)}(x_{(j-1)T},x_{jT}) \quad (3.7a)$$

$$x^*_{((j-1)T,jT]}(x_{(j-1)T},p_j^*)=x^*_{((j-1)T,jT]}(x_{(j-1)T},x_{jT}). \quad (3.7b)$$

*Proof:* For a given $p_j$, the optimal solution of $(P-j)$ is unique and can be found by solving

$$\frac{\partial g_i}{\partial u_i}+B_i^T\lambda_{i+1}=0, \quad (3.8a)$$

$$\frac{\partial g_i}{\partial x_i}+A_i^T\lambda_{i+1}-\lambda_i=0, \qquad i=(j-1)T+1,\cdots,jT-1, \quad (3.8b)$$

$$x_{i+1}=A_i x_i+B_i u_i, \quad i=(j-1)T,\cdots,jT-1, \quad (3.8c)$$

$$\text{where } x_{(j-1)T} \text{ is given,} \quad (3.8d)$$

$$\text{and } p_j-\lambda_{jT}=0. \quad (3.8e)$$

Note that the only difference between (3.6) and (3.8) is that the final state constraint in (3.6d) is replaced by the final condition on $\lambda_{jT}$ in (3.8e). Therefore, $u^*_{[(j-1)T,jT)}(x_{(j-1)T},x_{jT})$ and $x^*_{((j-1)T,jT]}(x_{(j-1)T},x_{jT})$ satisfy (3.8a)-(3.8d). Substituting them into (3.8a)-(3.8d), we can obtain $\lambda_{jT}^*$. By choosing

$$p_j^*=\lambda_{jT}^*, \quad (3.9)$$

(3.8e) will be satisfied. Since the optimal control and state trajectories are unique for convex problems, the $u^*$ and $x^*$ in (3.7) which satisfy the necessary conditions must be the solution.

*Lemma 3:* If $x^*_{(0,N]}$ and $u^*_{[0,N)}$ are the optimal solution of $(P)$, then there exists a $p^*_{[1,M)}$ such that

$$x^*_{((j-1)T,jT]} = x^*_{((j-1)T,jT]}(x^*_{(j-1)T}, p^*_j), \tag{3.10a}$$

$$u^*_{[(j-1)T,jT)} = u^*_{[(j-1)T,jT)}(x^*_{(j-1)T}, p^*_j), \tag{3.10b}$$

and $p_{[1,M)}$ is the optimal solution of $(P-H)$.

*Proof:* From the principle of optimality and the uniqueness of the solution of $(P-j)$, we have

$$x^*_{((j-1)T,jT]}(x^*_{(j-1)T}, x^*_{jT}) = x^*_{((j-1)T,jT]}, \tag{3.11a}$$

$$u^*_{[(j-1)T,jT)}(x^*_{(j-1)T}, x^*_{jT}) = u^*_{[(j-1)T,jT)}. \tag{3.11b}$$

From Lemma 2 and (3.11), there exists $p_{[1,M)}$ such that (3.10) is satisfied. To see that $p^*_{[1,M)}$ is indeed an optimal solution of $(P-H)$, we have

$$J^H(p_{[1,M)}) = J^P(x_{(0,N]}, u_{[0,N)}) \geq J^P(x^*_{(0,N]}, u^*_{[0,N)}) = J^H(p^*_{[1,M)}) \tag{3.12}$$

where

$$x_{((j-1)T,jT]} \equiv x^*_{((j-1)T,jT]}(x_{(j-1)T}, p_j), \tag{3.13a}$$

$$u_{[(j-1)T,jT)} \equiv u^*_{[(j-1)T,jT)}(x_{(j-1)T}, p_j). \tag{3.13b}$$

*Lemma 4:* Let $p^*_{[1,M)}$ be a global optimal solution of $(P-H)$ and denote

$$x^*_{((j-1)T,jT]} \equiv x^*_{((j-1)T,jT]}(x_{(j-1)T}, p^*_j), \tag{3.14a}$$

$$u^*_{[(j-1)T,jT)} \equiv u^*_{[(j-1)T,jT)}(x_{(j-1)T}, p^*_j), \tag{3.14b}$$

then $x^*_{(0,N]}$ and $u^*_{[0,N)}$ are the optimal solution of $(P)$.

*Proof:* Since $p^*_{[1,M)}$ is the global optimal solution of $(P-H)$, we have

$$J^H(p^*_{[1,M)}) \leq J^H(p_{[1,M)}) \qquad \text{for any } p_{[1,M)}. \tag{3.15}$$

For any admissible $x_{(0,N]}$ and $u_{[0,N)}$, there exists a $\bar{p}_j$ such that

$$x_{jT} = x^*_{jT}(x_{(j-1)T}, \bar{p}_j). \tag{3.16}$$

We then have

$$J^P(x_{(0,N]}, u_{[0,N)}) = \sum_{i=1}^{N-1} g_i(x_i, u_i) + g_N(x_N) \tag{3.17a}$$

$$= \sum_{j=1}^{M-1} \left\{ \sum_{i=(j-1)T}^{jT-1} g_i(x_i, u_i) + \bar{p}_j x_{jT} \right\} \tag{3.17b}$$

$$+ \sum_{i=(M-1)T}^{MT-1} g_i(x_i, u_i) + g_N(x_N) - \sum_{j=1}^{M} \bar{p}_j x_{jT}$$

$$\geq \sum_{j=1}^{M-1} J_j(x_{((j-1)T,jT]}(x_{(j-1)T}, \bar{p}_j),$$

$$u_{[(j-1)T,jT)}(x_{(j-1)T}, \bar{p}_j)) + J_M(x_{(M-1)T})$$

$$- \sum_{j=1}^{M-1} \bar{p}_j x^*_{jT}(x_{(j-1)T}, \bar{p}_j) \tag{3.17c}$$

$$= J^H(\bar{p}_{[1,M)}) \geq J^H(p^*_{[1,M)})$$

$$= J^P(x^*_{(0,N]}, u^*_{[0,N)}). \tag{3.17d}$$

Equation (3.17b) is correct since all the $\bar{p}_j$ terms are cancelled by each other. Equation (3.17c) is due to the minimization operation on subproblems embedded in the notations used, (3.17d) is from the definition of $J^H$, and because of (3.15). Note that $J_M$ is a function of $x_{(M-1)T}$ alone. Therefore, the proof is completed.

*Lemma 5:* If all the pairs $(x_{(j-1)T}, x_{jT}) \in R^{2n}$ are admissible for $j = 1, \cdots, M-1$, then the optimal solution of $(P-H)$ is unique and there are no other local minimums.

*Proof:* To prove this lemma, we only need to show that $(P-H)$ has a local minimum if and only if $(P)$ has a local minimum. Since $(P)$ has a unique solution and there are no other local minimums, so does $(P-H)$.

Denote the optimal cost of $(T-j)$, $j = 1, \cdots, M-2$ by $\hat{J}_j(x_{(j-1)T}, x_{jT})$. For notational simplicity, we combine the optimal cost of $(T-(M-1))$ and $(T-M)$ together and call it $\hat{J}_{M-1}(x_{(M-2)T}, x_{(M-1)T})$. The optimal $x^*_{jT}$, $j = 1, \cdots, M-1$ can then be found for any given $x_0$ by solving

$$(P)' \quad \min_{\substack{x_{jT} \\ j=1,\cdots,M-1}} J^P = \sum_{j=1}^{M-1} \hat{J}_j(x_{(j-1)T}, x_{jT}), \quad \text{where } x_0 \text{ is given.} \tag{3.18}$$

Problem $(P)'$ has a unique solution satisfying the following necessary conditions:

$$\frac{\partial J^P}{\partial x_{(M-1)T}} = \left[\frac{\partial \hat{J}_{M-1}(x^*_{(M-2)T}, x_{(M-1)T})}{\partial x_{(M-1)T}}\right]_{x_{(M-1)T}=x^*_{(M-1)T}} = 0, \tag{3.19a}$$

$$\frac{\partial J^P}{\partial x_{jT}} = \left[\frac{\partial \hat{J}_j(x^*_{(j-1)T}, x_{jT})}{\partial x_{jT}} + \frac{\partial \hat{J}_{j+1}(x_{jT}, x^*_{(j+1)T})}{\partial x_{jT}}\right]_{x_{jT}=x^*_{jT}} = 0, \quad j = 1, \cdots, M-2. \tag{3.19b}$$

The necessary condition of $(P-H)$ is

$$\frac{\partial J^H}{\partial p_j} = 0, \qquad j = 1, \cdots, M-1. \tag{3.20}$$

We have to show

$$(3.19) \text{ holds if and only if } (3.20) \text{ holds.} \tag{3.21}$$

For any given $p_j$ and initial condition $x_{(j-1)T}$, the optimal $x_{jT}$ for $(P-j)$ is denoted by $x_{jT}(p_j)$ and can be found by solving [3]

$$(P-j)' \quad \min_{x_{jT}} \hat{J}_j(x_{(j-1)T}, x_{jT}) + p_j x_{jT}, j = 1, \cdots, M-1. \tag{3.22}$$

Note that $(P-(M-1))'$ combines $(P-(M-1))$ and $(P-M)$ together as in the definition of $\hat{J}_{M-1}(x_{(M-2)T}, x_{(M-1)T})$.

The necessary condition of $(P-j)'$ is

$$p_j + \frac{\partial \hat{J}_j(x_{(j-1)T}, x_{jT})}{\partial x_{jT}} = 0, \qquad j = 1, \cdots, M-1. \tag{3.23}$$

By taking the derivative with respect to $p_j$ to (3.23), we have

$$I + \frac{\partial^2 \hat{J}_j(x_{(j-1)}, x_{jT})}{\partial x^2_{jT}} \frac{\partial x_{jT}}{\partial p_j} = 0. \tag{3.24}$$

To find $p^*_{[1,M-1]}$ for $(P-H)$, we solve problem $(P-H)'$:

$$(P-H)' \quad \min_{p[1,M-1]} J^H = \sum_{j=1}^{M-1} \hat{J}_j(x_{(j-1)T}(p_{j-1}), x_{jT}(p_j)), \quad \text{where } x_0(p_0) \equiv x_0. \tag{3.25}$$

The necessary condition is

$$\frac{\partial J^H}{\partial p_{M-1}} = \left(\frac{\partial x_{(M-1)T}}{\partial p_{M-1}}\right)^T \frac{\partial \hat{J}_{M-1}(x_{(M-2)T}, x_{(M-1)T})}{\partial x_{(M-1)T}} = 0 \tag{3.26a}$$

$$\frac{\partial J^H}{\partial p_j} = \sum_{i=j}^{M-2} \left(\frac{\partial x_{iT}}{\partial p_j}\right)^T \left[\frac{\partial \hat{J}_j(x_{(i-1)T}, x_{iT})}{\partial x_{iT}} + \frac{\partial \hat{J}_{j-1}(x_{iT}, x_{(i+1)T})}{\partial x_{iT}}\right] + \left(\frac{\partial x_{(M-1)T}}{\partial p_j}\right)^T \frac{\partial \hat{J}_{M-1}}{\partial x_{(M-1)T}},$$
$$j = 1, \cdots, M-2. \qquad (3.26b)$$

If (3.19) holds, we can see that (3.20) holds just by substituting (3.19) into (3.26). To show that (3.19) holds when (3.20) holds, we first observe from (3.24) that $\partial x_{(M-1)T}/\partial p_{M-1}$ is invertible. If $\partial J^H/\partial p_{M-1} = 0$, then from (3.26a), we see that (3.19a) holds. The proof for $j = 1, \cdots, M-2$ is accomplished by using mathematical induction as follows.

Assume that

$$\frac{\partial J^H}{\partial p_i} = 0 \text{ in (3.20) implies } \frac{\partial J^P}{\partial x_{iT}} = 0 \text{ in (3.19)}$$
$$\text{for all } i = j, \cdots, M-1. \qquad (3.27a)$$

To use the induction method, we have to show that

$$\frac{\partial J^H}{\partial p_{j-1}} = 0 \text{ implies } \frac{\partial J^P}{\partial x_{(j-1)T}} = 0. \qquad (3.27b)$$

From (3.19b), (3.26b), and (3.27a), we have

$$\frac{\partial J^H}{\partial p_{j-1}} = \left(\frac{\partial x_{j-1}}{\partial p_{j-1}}\right)^T \frac{\partial J^P}{\partial x_{j-1}}. \qquad (3.28)$$

Since $\partial x_{j-1}/\partial p_{j-1}$ is invertible from (3.24), (3.27b) holds and the proof is completed.

In Lemma 5, we assume that all the pairs $(x_{(j-1)T}, x_{jT}) \in R^{2n}$ are admissible for $j = 1, \cdots, M-1$. We now remove the constraint, and consider the case in which all the admissible pairs $(x_{(j-1)T}, x_{jT})$ form only a subset of $R^{2n}$.

*Lemma 6:* Define

$$X_{jT}(x_{(j-1)T}) \equiv \{x_{jT} - \bar{A}_j x_{(j-1)T} | (x_{(j-1)T}, x_{jT}) \text{ is admissible}\}, \qquad (3.29a)$$

$$\text{where } \bar{A}_j \equiv A_{jT-1} \cdots A_{(j-1)T}, \qquad (3.29b)$$

$$P_j(x_{(j-1)T}, x_{jT}) \equiv \left\{ p_j^* + \frac{\partial \hat{J}_j(x_{(j-1)T}, x_{jT})}{\partial x_{jT}} \middle| \text{ all } p_j^*\text{'s in Lemma 2} \right\}, \qquad (3.30)$$

then both $X_{jT}$ and $P_j$ are linear spaces. If the dimension of $X_{jT}$ is $r_j$, then the dimension of $P_j$ is $n - r_j$.

*Proof:* From the dynamic equation, we can get

$$x_{jT} - \bar{A}_j x_{(j-1)T} = H_{(j-1)T} u_{(j-1)T} + \cdots + H_{jT-1} u_{jT-1} \qquad (3.31a)$$

where

$$H_i \equiv A_{jT-1} \cdots A_{i+1} B_i, i = (j-1)T, \cdots, jT-2, H_{jT-1} \equiv B_{jT-1}. \qquad (3.31b)$$

Therefore, $X_{jT}$ is a linear space spanned by the columns of $H_i$'s and $r_j$ is the number of linearly independent columns of $H_i$'s.

Let $E^j \equiv [e_i^j, i = 1, \cdots, r_j]$ be an $n \times r_j$ matrix where $e_i^j$'s form an orthonormal basis of $X_{jT}$ and $y^j \equiv [y_i^j, j = 1, \cdots, r_j]^T$ is the representation of an element in $X_{jT}$, that is,

$$E^j y^j = \hat{x}_{jT} - \bar{A}_j x_{(j-1)T}. \qquad (3.32)$$

The relationship between $x_{jT}$ and $p_j$ can be obtained by solving $(P-j)$'s in (3.22) in terms of $y_i^j$. The necessary condition is

$$\left[\frac{\partial x_{jT}}{\partial y^j}\right]^T p_j + \left[\frac{\partial x_{jT}}{\partial y^j}\right]^T \frac{\partial \hat{J}_j(x_{(j-1)T}, x_{jT})}{\partial x_{jT}} = 0 \qquad (3.33a)$$

or

$$(E^j)^T \left\{ p_j + \frac{\partial \hat{J}_j(x_{(j-1)T}, x_{jT})}{\partial x_{jT}} \right\} = 0. \qquad (3.33b)$$

From Lemma 3, there exists at least one solution $p_j$ for (3.33). Since the dimension of $E^j$ is $r_j$, the degree of freedom of $p_j$ is $n - r_j$. In other words, $n - r_j$ out of $n$ components of $p_j$'s are free variables. That is, the dimension of $P_j$ is $n - r_j$.

*Lemma 7:* The global optimal solution $p^*$ of the parameter optimization problem $(P-H)$ may not be unique, and there is no other local minimum.

*Proof:* The proof is similar to that in Lemma 5. Note that the constrained optimization problem below

$$\min_{x_{jT} \in X_{jT}(x_{(j-1)T}) + \bar{A}_j x_{(j-1)T}} \hat{J}_j(x_{(j-1)T}, x_{jT}) \qquad (3.34)$$

is equivalent to the unconstrained problem

$$\min_{y^j \in R^{r_j}} \hat{J}_j(x_{(j-1)T}, (y^j)^T E^j + \bar{A}_j x_{(j-1)T}). \qquad (3.35)$$

From (3.35), we have the necessary condition for $(P-H)$ in terms of $p_{M-1}$:

$$\frac{\partial J^H}{\partial p_{M-1}} = \left[\frac{\partial y^{M-1}}{\partial p_{M-1}}\right]^T \frac{\partial J_{M-1}}{\partial y^{M-1}} = 0. \qquad (3.36)$$

From (3.33b), we have

$$[E^{M-1}]^T \left\{ I + \frac{\partial^2 \hat{J}_{M-1}(x_{(M-2)T}, x_{(M-1)T})}{\partial x^2_{(M-1)T}} \frac{\partial x_{(M-1)T}}{\partial y^{M-1}} \frac{\partial y^{M-1}}{\partial p_{M-1}} \right\} = 0 \qquad (3.37a)$$

or

$$[E^{M-1}]^T = -[E^{M-1}]^T \frac{\partial^2 \hat{J}_{M-1}}{\partial x^2_{(M-1)T}} E^{M-1} \frac{\partial y^{M-1}}{\partial p_{M-1}}. \qquad (3.37b)$$

Since $E^{M-1}$ is a full rank matrix and so is $\partial^2 \hat{J}_{M-1}/\partial x^2_{(M-1)T}$ from (3.24), the matrix $\partial y^{M-1}/\partial p_{M-1}$ must be full rank. From (3.36), we therefore have

$$\frac{\partial J^H}{\partial p_{M-1}} = 0 \quad \text{ if and only if } \frac{\partial \hat{J}_{M-1}}{\partial y^{M-1}} = 0. \qquad (3.38)$$

Following a similar induction approach as in Lemma 5, we can prove the counterpart of (3.21).

Note that all $p_j \in P_j$ induces the same $x_i$'s and $u_i$'s for low-level subproblems; they thus give the same high level cost $J^H$. Also, the correspondence between the set $P_j(x_{(j-1)T}, x_{jT})$ and a $x_{jT} - \bar{A}_j x_{(j-1)T} \in X_{jT}$ is one-to-one for a given $x_{(j-1)T}$. Therefore, the solution for $\partial J^H/\partial p_{M-1} = 0$ in (3.38) may not be unique. All of them give the global minimum cost of $(P-H)$, which is also the global minimum cost of $(P)$.

From Lemma 7, we know that the high-level parameter optimization problem can have more than one optimal solution. However, all of them induce the unique global optimal solution of the original problem, and thus cause no problem to find the original solution. While the original problem is convex and has a unique global optimal solution, we may expect that the high-level problem is also convex. However, this is not the case, as shown in Example 1.

*Example 1:* Let us consider a strictly convex function $\hat{J}_1(x_0, x_T) = e^{x_T^2}$ in (3.22) with $j = 1$ and $n = 1$. It can be shown that

$$\frac{\partial^2 J^H}{\partial p_1^2} = \frac{1 - 2x_T^2}{2(1 + 2x_T^2)^3 e^{x_T^2}}. \qquad (3.39)$$

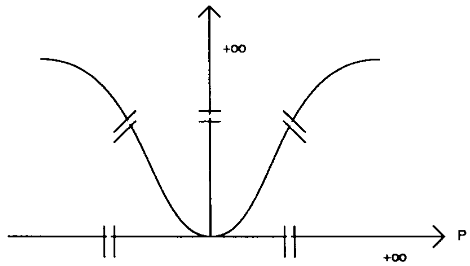

Fig. 1. Nonconvexity of a high-level problem.

In this case, we have

$$\frac{\partial^2 J^H}{\partial p_1^2} < 0, \qquad \text{for } x_T > \frac{1}{\sqrt{2}} \text{ or } x_T < -\frac{1}{\sqrt{2}} \tag{3.40}$$

or equivalently, for $p_1 < -2.3$ and $p_1 > 2.3$.

Therefore, $\hat{J}^H(p_1)$ is concave for $p_1 < -2.3$ and $p_1 > 2.3$. The function $\hat{J}^H(p_1)$ is depicted in Fig. 1. This example clearly illustrates why the optimal solution for the nonconvex high-level problem is still unique.

## IV. A Parallel Algorithm

For clarity, we shall first present a sequential algorithm and then modify it to a parallel algorithm. By using the gradient methods for $(P-H)$, we update the coordination term by

$$p_j^{k+1} = p_j^k - \epsilon_j \frac{\partial J}{\partial p_j}, \qquad j = 1, \cdots, M-1 \tag{4.1}$$

where the superscript represents the number of iteration and $\epsilon_j$ is the step size for subproblem $(P-j)$.

Note that $J$ is not an explicit function of $p_j$. The computation of $\partial J/\partial p_j$ in (4.1) is presented next. For a given set $\{p_j\}_{j=1,\cdots,M-1}$, Problem $(P-j)$'s are solved sequentially. Denote $x_i^*$, $u_i^*$ as the optimal solution found for all $(P-j)$'s. Define

$$\bar{J}_j(x^*_{(j-1)T}) = \sum_{i=(j-1)T}^{jT-1} g_i(x_i^*, u_i^*), \qquad j = 1, \cdots, M-1, \tag{4.2}$$

$$\bar{J}_M(x^*_{(M-1)T}) = g_N(x_N^*) + \sum_{i=(M-1)T}^{MT-1} g_i(x_i^*, u_i^*). \tag{4.3}$$

Since subproblems are evaluated sequentially, the change of $p_j$ affects $\bar{J}_j$ directly, and affects all $\bar{J}_i$'s for $i > j$ indirectly through the change of initial conditions. We have

$$\frac{\partial \bar{J}}{\partial p_j} = \frac{\partial \bar{J}_j}{\partial p_j} + \left(\frac{\partial x^*_{jT}}{\partial p_j}\right)^T \frac{\partial \bar{J}_{j+1}}{\partial x^*_{jT}} + \cdots + \left(\frac{\partial x^2_{jT}}{\partial p_j}\right)^T \left(\frac{\partial x^*_{(M-1)T}}{\partial x^*_{(M-2)T}}\right)^T \cdots \frac{\partial \bar{J}_M}{\partial x^*_{(M-1)T}}. \tag{4.4}$$

Thus, once the following quantities

$$\frac{\partial \bar{J}_j}{\partial p_j}, \frac{\partial x^*_{jT}}{\partial p_j}, \frac{\partial \bar{J}_j}{\partial x^*_{(j-1)T}}, \frac{\partial x^*_{jT}}{\partial x_{(j-1)T}}, x^*_{jT} \tag{4.5}$$

are passed to high level from subproblems, the gradient in (4.4) can be computed at the higher level. Note that $x^*_{jT}$ will be used later in the parallel algorithm.

To compute $\partial \bar{J}_j/\partial p_j$, we use

$$\frac{\partial \bar{J}_j(x^*_{(j-1)T})}{\partial p_j} = \left[\frac{\partial x^*_{jT}}{\partial p_j}\right]^T \frac{\partial \hat{J}_j(x^*_{(j-1)T}, x^*_{jT})}{\partial x^*_{jT}} = \left[\frac{\partial x^*_{jT}}{\partial p_j}\right]^T \frac{\partial g^*_{jT-1}(x^*_{jT-1}, x^*_{jT})}{\partial x^*_{jT}}, \tag{4.6}$$

where

$$g_i^*(x_i, x_{i+1}) = \min_{u_i} g_i(x_i, u_i). \tag{4.7a}$$

subject to $x_{i+1} = A_i x_i + B_i u_i$, with $x_i$ and $x_{i+1}$ given. (4.7b)

Note that $\hat{J}_j(x^*_{(j-1)T}, x^*_{jT})$ is the optimal cost of $(T-j)$ for given $x^*_{(j-1)T}$ and $x^*_{jT}$, and the second equality is obtained from [3, Lemma 2]. Similarly,

$$\frac{\partial \hat{J}_{j+1}(x^*_{jT})}{\partial x^*_{jT}} = \frac{\partial \hat{J}_{j+1}(x^*_{jT}, x^*_{(j+1)T})}{\partial x^*_{jT}} + \left[\frac{\partial x^*_{(j+1)T}}{\partial x^*_{jT}}\right]^T \frac{\partial \hat{J}_{j+1}(x^*_{jT}, x^*_{(j+1)T})}{\partial x^*_{(j+1)T}} \tag{4.8a}$$

$$= \frac{\partial g^*_{jT}(x^*_{jT}, x^*_{jT+1})}{\partial x^*_{jT}} + \left[\frac{\partial x^*_{(j+1)T}}{\partial x^*_{jT}}\right]^T \frac{\partial g^*_{(j+1)T}(x^*_{(j+1)T-1}, x^*_{(j+1)T})}{\partial x^*_{(j+1)T}}. \tag{4.8b}$$

The computation of $\partial x^*_{jT}/\partial p_j$'s are more involved. To see this, let us formulate differentials of the necessary condition in (3.8) around the optimal solution. We have

$$\frac{\partial^2 g_i}{\partial u_i^2}\Delta u_i + \frac{\partial^2 g_i}{\partial x_i \partial u_i}\Delta x_i + B_i^T \Delta\lambda_{i+1} = 0, \tag{4.9a}$$

$$A_i^T \Delta\lambda_{i+1} + \frac{\partial^2 g_i}{\partial x_i \partial u_i}\Delta u_i + \frac{\partial^2 g_i}{\partial x_i^2}\Delta x_i - \Delta\lambda_i = 0, \tag{4.9b}$$

$$\Delta x_{i+1} = A_i \Delta x_i + B_i \Delta u_i, \qquad i = (j-1)T, \cdots, jT, \tag{4.9c}$$

$$\Delta x_{(j-1)T} = 0 \text{ and } \Delta p_j = \Delta\lambda_{jT}. \tag{4.9d}$$

To obtain $\partial x_{jT}/\partial p_j$, we need only to compute the ratio

$$\frac{\partial x_{jT}}{\partial p_j} = \frac{\Delta x_{jT}}{\Delta\lambda_{jT}}. \tag{4.10}$$

We can then compute (4.10) by choosing $\Delta\lambda_{(j-1)T} = 1$ together with $\Delta x_{(j-1)T} = 0$ because (4.9) are linear homogeneous difference equations. Similarly, $\partial x_{jT}/\partial x_{(j-1)T}$ can be obtained.

Since the only linkage between subproblems $(P-j)$ and $(P-(j+1))$ is $x_{jT}$, parallel processing can be achieved if the high level can supply $(P-(j+1))$ with the initial condition $x_{jT}$. Let $x_{jT}^k$ denote the optimal terminal state of $(P-j)$ for the given coordination parameter $p_j^k$ at the $k$th iteration. With the new coordination parameter $p_j^{k+1}$ in (4.1), the optimal terminal state $x_{jT}^{k+1}$ will be changed correspondingly. Let

$$\Delta p_j^k \equiv p_j^{k+1} - p_j^k, \qquad j = 1,\cdots,M-1,$$
$$\text{and } \Delta x_{jT}^k \equiv x_{jT}^{k+1} - x_{jT}^k; \quad (4.11)$$

then we have the following first-order approximation:

$$\Delta x_T^k = \frac{\partial x_T}{\partial p_1}\Delta p_1^k \text{ and}$$
$$\Delta x_{jT}^k = \frac{\partial x_{jT}}{\partial x_{(j-1)T}}\Delta x_{(j-1)T}^k + \frac{\partial x_{jT}}{\partial p_j}\Delta p_j^k. \quad (4.12)$$

The high level can compute (4.12) by using the information contained in (4.5). It can then update the initial condition for $(P-(j+1))$ to achieve the parallel processing according to

$$x_{jT}^{k+1} = x_{jT}^k + \Delta x_{jT}^k. \quad (4.13)$$

The structure of the parallel iterative scheme is shown in Fig. 2. Its feasibility is demonstrated by using Example 2. Note that the high level is a parameter optimization problem when subproblems are solved sequentially. Since it uses the gradient method in high level, its convergence rate is linear in sequential processing. If the step size in the parallel processing is very small, the approximation procedure used in the algorithm is very good. Thus, it is expected that there is not much difference between the convergence rate of the parallel and sequential processing in the case. For this example, typical trajectories for the sequential and parallel processing are computed.

*Example 2:* Consider the example in [3]:

$$J = \frac{5}{2}x_N^2 + \frac{1}{8}x_N^4 + \sum_{i=0}^{15}\left(\frac{1}{2}x_i^2 + u_i^2 + \frac{1}{8}x_i^4\right),$$
$$\text{where } x_{i+1} = -x_i + u_i. \quad (4.14)$$

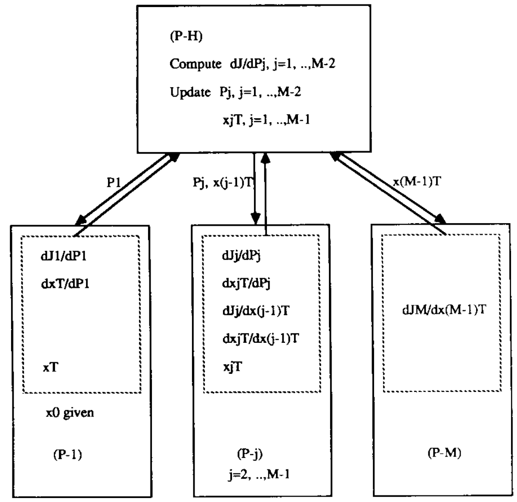


Fig. 2. A parallel algorithm.

The low-level subproblems are solved by using the method of variation of extremals [6, p. 343]. The steepest descent algorithm with fixed step size is used to solve the high-level problem. A typical run is shown as follows.

Step size $\epsilon_j = 0.22$, $\quad j = 1, 2, 3.$

Initial values $p_j = 5$, $\quad j = 1, 2, 3.$

The algorithm terminates if $\frac{\partial J}{\partial p_j} < 10^{-5}$, $\quad j = 1, 2, 3.$

| Count | Cost | $x(1T)$ | $x(2T)$ | $x(3T)$ | $x(4T)$ |
|---|---|---|---|---|---|
| 0 | 2.047575e+02 | −3.743567e+00 | −4.029408e+00 | −4.029408e+00 | −4.029408e+00 |
| 4 | 1.062122e+02 | 3.989132e−01 | 2.975307e−01 | 2.769890e−01 | −1.374960e−01 |
| 8 | 1.054864e+02 | 2.043705e−01 | 6.855832e−03 | −4.600316e−03 | 6.705500e−04 |
| 12 | 1.054861e+02 | 1.985175e−01 | 1.189246e−02 | 9.858966e−04 | 1.004316e−05 |
| 16 | 1.054861e+02 | 1.981958e−01 | 1.232782e−02 | 8.044351e−04 | 2.306958e−05 |
| 20 | 1.054861e+02 | 1.981747e−01 | 1.236148e−02 | 7.742419e−04 | 2.403846e−05 |

The above is the optimal solution from parallel processing.

| | | | | | |
|---|---|---|---|---|---|
| 0 | 3.752306e+02 | −3.743567e+00 | −4.162684e+00 | −4.171398e+00 | −4.171572e+00 |
| 4 | 1.056515e+02 | 2.656210e−01 | −7.055640e−03 | −3.291734e−02 | −2.283414e−01 |
| 8 | 1.054862e+02 | 2.011813e−01 | 1.041555e−02 | 5.289137e−04 | −3.641595e−03 |
| 12 | 1.054861e+02 | 1.983296e−01 | 1.226140e−02 | 7.994780e−04 | −3.755366e−05 |
| 16 | 1.054861e+02 | 1.981814e−01 | 1.235885e−02 | 7.735814e−04 | 2.301377e−05 |
| 18 | 1.054861e+02 | 1.981750e−01 | 1.236303e−02 | 7.718387e−04 | 2.398394e−05 |

The above is the optimal solution for sequential processing.

## V. Discussion

In this paper, we show that linear functions are sufficient for convex optimal control problems with time decomposition and incentive coordination. As a matter of fact, the two-level problem is equivalent to the original problem in the sense as stated in Theorem 1. A parallel algorithm based upon the gradient method is developed and a numerical example is used to illustrate its feasibility.

Intuitively, the concept of incentive coordination seems to be quite general. Thus, we expect that this approach can be extended to more general problems. In fact, we have shown that, under certain technical conditions, a nonconvex optimal control problem is equivalent to a two-level problem with quadratic incentives, in a similar sense as in this paper. The result was reported in [1]. Other ongoing research includes the extension of the current approach to constrained problems, the comparison of incentive coordination and the Lagrangian multiplier method, and numerical justification of the efficiency of the approach. The results will be reported in forthcoming papers.

ACKNOWLEDGMENT

The authors would like to thank Dr. S.-C. Chang and J. Tang for their valuable discussions.

REFERENCES

[1] M. Bromberg, T. S. Chang, and P. B. Luh, "Decomposition and coordination for non-convex optimal control problems with parallel algorithm," in *Proc. 26th IEEE Conf. Decision Contr.*, Los Angeles, CA, Dec. 1987, pp. 1468–1475.
[2] T. S. Chang and P. B. Luh, "On the decomposition and coordination of large scale dynamic control problems," in *Proc. 24th Conf. Decision Contr.*, Fort Lauderdale, FL, Dec. 1985, pp. 1484–1485.
[3] T. S. Chang, S. C. Chang, and P. B. Luh, "A hierarchical decomposition for large scale optimal control problems with parallel processing capability," in *Proc. 1986 Amer. Contr. Conf.*, Seattle, WA, June 1986, pp. 1995–2000.
[4] S. C. Chang, "A hierarchical temporal decomposition approach to long horizon optimal control problems," Ph.D. dissertation, Dep. Elec. Syst. Eng., Univ. Connecticut, Storrs, Dec. 1986.
[5] S. C. Chang, P. B. Luh, and T. S. Chang, "Stackelberg game approach to power system scheduling problems," in *Proc. 1984 Amer. Contr. Conf.*, Boston, MA, June 1984, pp. 1434–1439.
[6] D. E. Kirk, *Optimal Control Theory: An Introduction*. Englewood Cliffs, NJ: Prentice-Hall, 1970.
[7] A. P. Sage and C. C. White, III, *Optimum Systems Control*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1977.

## A New Problem in Control System Design

D.-W. GU AND D. Q. MAYNE

*Abstract*—**A new problem, arising in the design of circuits and control systems, is formulated and an outer approximation algorithm for solving this problem is presented. The problem is characterized by constraints which must be satisfied for all functions in a given subset of a function space.**

### I. INTRODUCTION

The main purpose of this note is to formulate a new optimization problem arising in circuit design and the design of control systems [1], [2] and to present an algorithm for solving this problem. It is well known that many design problems may be expressed as constrained optimization problems. These optimization problems may be conventional, nondifferentiable, semiinfinite [3], [4], or infinite dimensional [5] and various algorithms [3]–[8], [10] have been proposed to solve these problems. The semiinfinite optimization problems model the situation where the design objectives include "hard" constraints on time and frequency response (i.e., constraints which must be satisfied for all times, or all frequencies, in a given interval) and (in the case of robustness to structured perturbations) constraints which must be satisfied for all (plant) parameters in a given subset of $\boldsymbol{R}^m$. In this note we consider design problems which give rise to constraints which must be satisfied for all real functions in a given subset of a function space.

Manuscript received July 23, 1987; revised June 30, 1988 and November 11, 1988. Paper recommended by Associate Editor, R. Patel.
D.-W. Gu is with the Department of Engineering, University of Leicester, Leicester, England.
D. Q. Mayne is with the Department of Electrical Engineering, Imperial College, London, England.
IEEE Log Number 8931902.

To be more specific, consider a single-input single-output system, linear or nonlinear, whose input/output map $f: \mathcal{C} \times \Theta \to \mathcal{C}$ ($\mathcal{C}$ is the class of scalar continuous functions endowed with the sup norm) depends on a design parameter vector $\theta \in \Theta$, which is a compact subset in $\boldsymbol{R}^p$, i.e.,

$$y = f(u, \theta) \tag{1.1}$$

where $u$ is the "input" and $y$ the output. The design problem we consider is the following.

$P_1$: Find a $\theta$, such that

$$y = f(u, \theta) \in \mathcal{C}_y, \qquad \text{for all } u \in \mathcal{C}_u \tag{1.2}$$

where

$$\mathcal{C}_u \triangleq \{u: T \to \boldsymbol{R} | u \text{ is continuous; } u(t) \in [u_1(t), u_2(t)], t \in T\} \tag{1.3}$$

$$\mathcal{C}_y \triangleq \{y: T \to \boldsymbol{R} | y \text{ is continuous; } y(t) \in [y_1(t), y_2(t)], t \in T\} \tag{1.4}$$

$$T \triangleq [0, t_1], \qquad t_1 > 0$$

and $u_1$, $u_2$, $y_1$, $y_2$ are piecewise linear, continuous scalar functions on $T$ satisfying

$$u_1(t) < u_2(t)$$

$$y_1(t) < y_2(t)$$

for all $t \in T$.

This problem is relevant, for example, to the design of logic circuits where it is desired to ensure that the output waveform $y$ is an acceptable approximation to an ideal logic waveform when the input $u$ to the circuit is also acceptable; constraint (1.2) is a formal expression of this objective. It is also relevant to certain problems in control design; consider the case when $u$ represents a disturbance and $y$ the response of the closed-loop system to the disturbance. In this case, $\mathcal{C}_u$ is the class of potential disturbances and $\mathcal{C}_y$ the class of admissible outputs (deviations) due to these disturbances. The constraint $f(u, \theta) \in \mathcal{C}_y$ for all $u \in \mathcal{C}_u$ then represents a performance constraint. Control design objectives also include, of course, stability. While it is possible to formulate input-output stability as a similar constraint (by defining $T$ as the semiinfinite interval $[0, \infty)$ and permitting $\mathcal{C}_y$ to depend on $u$) such an extension would not be computationally useful. Thus, for practical design, $P_1$ should be modified to include stability constraints as well as other conventional constraints arising in control design. For simplicity of presentation, these constraints are omitted in the sequel, permitting concentration on the unique problems arising from the new constraint $\{f(u, \theta) \in \mathcal{C}_y, \text{ for all } u \in \mathcal{C}_u\}$. The algorithm presented in the following sections can easily be modified to include conventional constraints, but this is probably premature since further research is required to improve its performance.

### II. CONCEPTUAL OUTER APPROXIMATION ALGORITHM FOR $P_1$

It is relatively easy to formulate a conceptual algorithm for solving $P_1$. Before doing so it is convenient to introduce a function $\xi: \mathcal{C} \to \boldsymbol{R}$ whose value $\xi(y)$ is zero if and only if $y \in \mathcal{C}_y$. The function $\xi$ is defined by

$$\xi(y) \triangleq \max\{0; y(t) - y_2(t); y_1(t) - y(t) | t \in T\}. \tag{2.1}$$

The function $\eta: \mathcal{C} \times \Theta \to \boldsymbol{R}$, defined by

$$\eta(u, \theta) \triangleq \xi(f(u, \theta)) \tag{2.2}$$

will also be useful. Clearly, $\eta(u, \theta) = 0$ if and only if its "input" $u$ to the system is such that the corresponding output $y = f(u, \theta)$ lies in $\mathcal{C}_y$. Using $\eta$, problem $P_1$ may be reformulated as follows.

$P_2$: Find $\theta \in \Theta$, such that

$$\eta(u, \theta) = 0, \qquad \text{for all } u \in \mathcal{C}_u.$$

Let $F$ denote the feasible set for $P_1$ or $P_2$, i.e.,

$$F \triangleq \{\theta \in \Theta | \eta(u, \theta) = 0, \qquad \text{for all } u \in \mathcal{C}_u\}.$$