Asian Journal of Control, Vol. 12, No. 3, pp. 377–391, May 2010
Published online in Wiley InterScience (www.interscience.wiley.com) DOI: 10.1002/asjc.217

IMPROVED CONTROLLABILITY TEST FOR DEPENDENT SIPHONS IN S^3PR BASED ON ELEMENTARY SIPHONS

Daniel Y. Chao

ABSTRACT

When siphons in a flexible manufacturing system (FMS) modeled by an ordinary Petri net (OPN) become unmarked, the net gets deadlocked. To prevent deadlocks, some control places and related arcs are added to strict minimal siphons (SMS) so that no siphon can be emptied. For large systems, it is infeasible to add a monitor to every SMS since the number of SMS or control elements grows exponentially with respect to the size of a Petri net. To tackle this problem, Li and Zhou propose to add control nodes and arcs for only elementary siphons. The rest of siphons, called dependent ones, may be controlled by adjusting control depth variables of elementary siphons associated with a dependent siphon after the failure of two tests. First, they test a Marking Linear Inequality (MLI); if it fails, then they perform a Linear Integer Programming (LIP) test which is an NP-hard problem. This implies that the MLI test is only sufficient, but not necessary. We propose a sufficient and necessary test for adjusting control depth variables in an S^3PR to avoid the sufficient-only time-consuming linear integer programming (LIP) test (NP-complete problem) required previously for some cases. We theoretically prove the following: i) no need for LIP test for Type II siphons; and ii) Type I strongly n-dependent ($n>2$) siphons being always marked. As a result, the total time complexity to check controllability of all strongly dependent siphons is no longer exponential but reduced to linear if all siphons are of Type I. The total time complexity is $O(|\Pi_E||\Pi_D|)$ (order of the product of total number of elementary siphons and total number of dependent siphons) if all siphons are of Type II. A well-known S^3PR example has been illustrated to show the advantages.

Key Words: Petri nets, siphons, controllability, FMS, S^3PR.

I. INTRODUCTION

A flexible manufacturing systems (FMS) consists of a set of *working processes* (*WP*) competing for resources. A *WP* models a sequence of operations to manufacture a product. The circular wait for resources can bring the system into a deadlock [1–3] where some *WP* can never finish.

A Petri net model is constructed for an FMS. The analysis of this PN model is conducted and system properties are claimed. Liveness in Flexible Manufacturing Systems (FMS) modeled by ordinary Petri nets (OPN) is closely related to emptiable siphons. A siphon (resp. trap) S is a set of places where tokens can leak out (resp. inject in) into (resp. from) another set of places called complementary set $[S]$ of the siphon (resp. trap). Thus, these tokens stay either in S or $[S]$. S and $[S]$ together form the support of a so-called P-invariant. The total number of tokens S and $[S]$ is conservative. Once an

Manuscript received February 19, 2009; revised June 14, 2009; accepted February 5, 2010.

Daniel Y. Chao is with the National Chengchi University, Department of Management Information Systems, No. 64, Sec. 2, Chihnan Rd., Taipei, Taiwan, R.O.C. (e-mail: yuhyaw@gmail.com).

emptiable siphon is found, output transitions of places in the siphon can never be fired. Hence the net is not live and has deadlocks.

To prevent a siphon S from becoming empty of tokens, we often add a control place V_S and some control arcs so that $[S]$ plus V_S form part of the support of a new P-invariant. By controlling the initial number of tokens [denoted by $\boldsymbol{M}_0(V_S)$] in V_S, we can limit the maximal of tokens leaking from S into $[S]$. We say that S is invariant-controlled.

The number of SMS or control elements grows in general exponentially with respect to the size of a Petri net. Hence for large systems, it is impractical to add a monitor to each SMS. Unlike other techniques, Li and Zhou [1–5] divide SMS into two kinds: elementary and dependent. A T-vector $\boldsymbol{\eta}$ is associated with each SMS S so that $\boldsymbol{\eta}(t_i)$ is the number of tokens gained or lost from S by firing transition t_i once. A dependent siphon S_0 depends on elementary siphons $S_1, S_2, \ldots, S_k$ if $\boldsymbol{\eta}_0 = a_1\boldsymbol{\eta}_1 + a_2\boldsymbol{\eta}_2 + \cdots + a_k\boldsymbol{\eta}_k$. If all a_i $(i = 1, 2, 3, \ldots, k)$ are positive, then S_0 strongly depends on $S_1, S_2, \ldots,$ and S_k, otherwise if some a_i are negative, then S_0 is a weakly dependent siphon. The T-vectors for elementary siphons are mutually independent.

Li and Zhou [1, 4] add control nodes and arcs for only elementary siphons greatly reducing the number of control nodes and arcs. As a result, for complex systems, it is essential to apply the concept of elementary siphons to add monitors; the number of which is linear to the size of the nets modeling the systems.

After the failure of two tests, control depth variables of elementary siphons associated with a dependent siphon are adjusted to satisfy a Marking Linear Inequality (MLI).

First, the above MLI is tested; if it fails, then a Linear Integer Programming (LIP) test is performed, which takes exponential time due to the LIP's non-polynomial complexity.

Thus, the MLI test is only a sufficient (rather than a necessary) test. Li and Zhou [5–7] further improve the above sufficient MLI test so that a dependent siphon that previously failed the MLI test may now satisfy the new sufficient test to avoid the LIP test.

We will develop a sufficient and necessary test, better than that of Theorem 1 of [1], so that in some cases where the test in Theorem 1 fails, the new test succeeds, thus avoiding the time-consuming LIP. Once it has failed, there is no need for the LIP test, as with Li and Zhou's new approaches in [5–7], since the new test is necessary for the controllability.

We categorize siphons into two types and show that type II dependent siphons need no LIP test. Furthermore, we will show that strongly type I dependent siphons need no control if they depend on more than two (*i.e.*, $n > 2$) elementary siphons. Thus, even the above MLI test can be avoided. These results are the first of their kind as explained below.

This is significant since the number of dependent siphons is exponential to the size of the net, even though that of elementary siphons is linear. Thus, the time to verify against the MLI for all dependent siphons is exponential as for previous approaches, however, the number of dependent siphons with $n < 3$ is polynomial. As a result, the total time complexity is reduced from exponential to polynomial.

Further, for an $n = 2$ dependent siphon, a simple algebraic test is both sufficient and necessary to determine whether control depth variables need to be adjusted. Thus, the time-consuming LIP test is completely eliminated. Thus, among all strongly dependent siphons, we need to apply the polynomial-time new MLI test to only $n = 2$ type I strongly dependent siphons with no LIP test.

The rest of the paper is organized as follows: Section II presents the basis (S^3PR, elementary siphons, and characteristic T-vectors) to understand the paper. Section III motivates the reader by presenting some simple examples. The results are proved and generalized in Section IV. A well-known S^3PR example has been illustrated to show the advantages in Section V. Finally, Section VI concludes the paper.

II. PRELIMINARIES

A marked Petri Net (PN) is defined by a quadruple $N = (P, T, \boldsymbol{F}, \boldsymbol{M}_0)$, where P is the set of places, T is the set of transitions, $\boldsymbol{F} : (P \times T) \cup (T \times P) \to Z^+$ (the set of nonnegative integers) is the flow relation, and $\boldsymbol{M}_0 : P \to Z^+$ is the net initial marking assigned to each place $p \in P$, $\boldsymbol{M}_0(p)$ tokens. In the special case that the flow relation $\boldsymbol{F}$ maps onto $\{0, 1\}$; the Petri net is said to be ordinary (otherwise, general). The incidence matrix of N is a matrix $[\mathbf{N}] : P \times T \to Z$ (the set of integers) indexed by P and T such that $[\mathbf{N}](p, t) = \boldsymbol{F}(t, p) - \boldsymbol{F}(p, t)$ where $\boldsymbol{F}(p, t)$ is the weight of the arc from place p to its output transition t, and $\boldsymbol{F}(t, p)$ is the weight of the arc from transition t to its output place p.

The set of input (resp. output) transitions of a place p is denoted by ${}^\bullet p$ (resp. $p^\bullet$). Similarly, the set of input (resp. output) places of a transition t is denoted by ${}^\bullet t$ (resp. $t^\bullet$). Finally, an ordinary PN such that (s.t.) $\forall t \in T$, $|t^\bullet| = |{}^\bullet t| = 1$, is called a State Machine (*SM*).

Given a marking $\boldsymbol{M}$, a transition t is enabled if $\forall p \in {}^\bullet t$, $\boldsymbol{M}(p) \geq \boldsymbol{F}(p, t)$, and this is denoted by $\boldsymbol{M}[t>$. Firing an enabled transition t results in a new marking

$\boldsymbol{M}_1$, which is obtained by removing $\boldsymbol{F}(p,t)$ tokens from each place $p \in {}^\bullet t$, and placing $\boldsymbol{F}(t,p')$ tokens in each place $p' \in t^\bullet$ moving the system state from $\boldsymbol{M}_0$ to $\boldsymbol{M}_1$. Repeating this process, it reaches $\boldsymbol{M}'$ by firing a sequence $\sigma = t_1, t_2, \ldots, t_k$ of transitions. $\boldsymbol{M}'$ is said to be reachable from $\boldsymbol{M}_0$; *i.e.*, $\boldsymbol{M}_0[\boldsymbol{\sigma}>\boldsymbol{M}']$.

A transition $t \in T$ is live under $\boldsymbol{M}_0$ *iff* $\forall \boldsymbol{M} \in R(N,\boldsymbol{M}_0)$, $\exists \boldsymbol{M}' \in R(N,\boldsymbol{M})$, t is firable under $\boldsymbol{M}'$. A PN is *live* under $\boldsymbol{M}_0$ *iff* $\forall t \in T$, t is live under $\boldsymbol{M}_0$. A Petri net is said to be deadlockfree, if at least one transition is enabled at every reachable marking.

A *P-vector* is a column vector $\boldsymbol{L}: P \to Z$ indexed by P and a *T-vector* is a column vector $\boldsymbol{J}: T \to Z$ indexed by T, where Z is the set of integers. For economy of space, we use $\sum \boldsymbol{L}(p)p$ (resp. $\sum \boldsymbol{J}(t)t$) to denote a P (resp. T)-vector.

P-vector $\boldsymbol{I}$ is a P-invariant $\boldsymbol{I}$ *iff* $\boldsymbol{I} \neq \boldsymbol{0}$ and $\boldsymbol{I}^T \cdot [\boldsymbol{N}] = \boldsymbol{0}^T$. $\|\boldsymbol{I}\| = \{p \in P | \boldsymbol{I}(p) \neq 0\}$ is called the support of $\boldsymbol{I}$. A P-invariant is said to be minimal if its support is not a strict superset of that of another, and the greatest common divisor of its elements is one.

For a Petri net $(N,\boldsymbol{M}_0)$, a non-empty subset S (resp. τ) of places is called a siphon (resp. trap) if ${}^\bullet S \subseteq S^\bullet$ (resp. $\tau^\bullet \subseteq {}^\bullet\tau$), *i.e.*, every transition having an output (resp. input) place in S has an input (resp. output) place in S (resp. τ). A transition in $S^\bullet$ is called a sink transitions of S. S is called an empty siphon at $\boldsymbol{M}_0$ if $\boldsymbol{M}_0(S) = \sum_{p \in S} \boldsymbol{M}_0(p) = 0$. A *minimal siphon* does not contain a siphon as a proper subset. It is called a strict minimal siphon (SMS), denoted by S, if it does not contain a trap. A siphon is said to be controlled if it is always marked.

Property 1 ([*1*]). If $\boldsymbol{I}$ is a P–invariant of N, then given an initial marking $\boldsymbol{M}_0$, $\forall \boldsymbol{M} \in R(N,\boldsymbol{M}_0)$, $\boldsymbol{I} \cdot M = \boldsymbol{I} \cdot \boldsymbol{M}_0$.

An initially marked $\|\boldsymbol{I}\|$ can never become empty of tokens. The union of a set of $\|\boldsymbol{I}\|$ forms another $\|\boldsymbol{I}'\|$. An emptiable siphon (or SMS) S can be obtained by deleting a complimentary set of places, denoted by $[S]$, from the union.

Property 2 (*Property in* [*1*]). For a given SMS S in an S^3PRN, $S \cup [S]$ is the support of a P-invariant of N.

Property 3 (*Corollary 3 in* [*1*]). Let S be a strict redundant SMS w.r.t. elementary siphons $S_1, S_2, \ldots, S_n$, and in an S^3PR. We have $[S_0] = [S_1] \cup [S_2] \cup \ldots \cup [S_n]$.

Definition 1. $N' = (P', T', \boldsymbol{F}')$ is called a subnet of N where $P' \subseteq P$, $T' \subseteq T$, $\boldsymbol{F}' = \boldsymbol{F} \cap ((P' \times T') \cup (T' \times P'))$. A net N is strongly connected iff for every node pair (n_i, n_j), $n_i, n_j \in P \cup T$, there is a directed path from n_i to n_j. A subnet $N_i = (P_i, T_i, \boldsymbol{F}_i)$ of N is generated by $X = P_i \cup T_i$, if $\boldsymbol{F}_i = \boldsymbol{F} \cap (X \times X)$. It is an I–subnet, denoted by N_I, of N if $T_i = {}^\bullet P_i$. N_I is the I–subnet [the subnet derived from $(S, {}^\bullet S)$] of an SMS S. Note that $S = P(N_I)$; S is the set of places in N_I.

2.1 S^3PR

The following definitions are adapted from [1]. The reader can refer to [1] for more details of the S^3PR model.

Definition 2 ([*1*]). A simple sequential process (S^2P) is a net $N = (P \cup \{p^0\}, T, \boldsymbol{F})$ where: (1) $P \neq \emptyset$, $p^0 \notin P$ (p^0 is called the process idle or initial or final operation place); (2) N is strongly connected state machine (SM) and (3) every circuit of N contains the place p^0.

Transitions in $p^{0\bullet}$ and ${}^\bullet p^0$ are called source and sink transitions respectively.

Definition 3 ([*1*]). A simple sequential process with resources (S^2PR), also called a working processes (WP), is a net $N = (P \cup \{p^0\} \cup P_R, T, \boldsymbol{F})$ so that (1) the subnet generated by $X = P \cup \{p^0\} \cup T$ is an S^2P; (2) $P_R \neq \emptyset$ and $P \cup \{p^0\} \cap P_R = \emptyset$; (3) $\forall p \in P, \forall t \in {}^\bullet p, \forall t' \in p^\bullet$, $\exists r_p \in P_R, {}^\bullet t \cap P_R = t'^\bullet \cap P_R = \{r_p\}$; (4) The two following statements are verified: $\forall r \in P_R, a)^{\bullet\bullet} r \cap P = r^{\bullet\bullet} \cap P \neq \emptyset$; b) ${}^\bullet r \cap r^\bullet = \emptyset$; (5) ${}^{\bullet\bullet}(p^0) \cap P_R = (p^0)^{\bullet\bullet} \cap P_R = \emptyset$. $\forall P \in P$, p is called an operation place, $\forall r \in P_R$, r is called a resource place, $H(r) = {}^{\bullet\bullet} r \cap P$ denotes the set of holders of r (operation places that use r). Any resource r is associated with a minimal P-invariant whose support is denoted by $\rho(r) = \{r\} \cup H(r)$.

Definition 4 ([*1*]). A system of $S^2PR(S^3PR)$ is defined recursively as follows: (1) An S^2PR is defined as an S^3PR; (2) Let $N_i = (P_i \cup P_i^0 \cup P_{Ri}, T_i, \boldsymbol{F}_i)$, $i \in \{1,2\}$ be two S^3PR so that $(P_1 \cup P_1^0) \cap (P_2 \cup P_2^0) = \emptyset$. $P_{R1} \cap P_{R2} = P_C (\neq \emptyset)$ and $T_1 \cap T_2 = \emptyset$. The net $N = (P \cup P^0 \cup P_R, T, \boldsymbol{F})$ resulting from the composition of N_1 and N_2 via P_C (denoted by $N_1 o N_2$) defined as follows: (1) $P = P_1 \cup P_2$; (2) $P^0 = P_1^0 \cup P_2^0$; (3) $P_R = P_{R1} \cup P_{R2}$; (4) $T = T_1 \cup T_2$ and (5) $\boldsymbol{F} = \boldsymbol{F}_1 \cup \boldsymbol{F}_2$ is also an S^3PR. A path (resp. circuit, subnet) $\boldsymbol{\Gamma}$ (resp. $\boldsymbol{c}$, $\boldsymbol{N}'$) in $\boldsymbol{N}$ is called a resource path (resp. circuit, subnet) if $\forall \boldsymbol{p} \in \Gamma(\boldsymbol{c}, \boldsymbol{N}'), \boldsymbol{p} \in \boldsymbol{P_R}$.

An S^3PR is composed of some state machines (with choices) holding and releasing some common resources. Fig. 1(a) shows an example of S^3PR (solid part) and its controlled model (including dashed part) respectively. We construct an SMS based on the concept of handles.

Roughly speaking, a "handle" is an alternate disjoint path between two nodes. A PT-handle starts

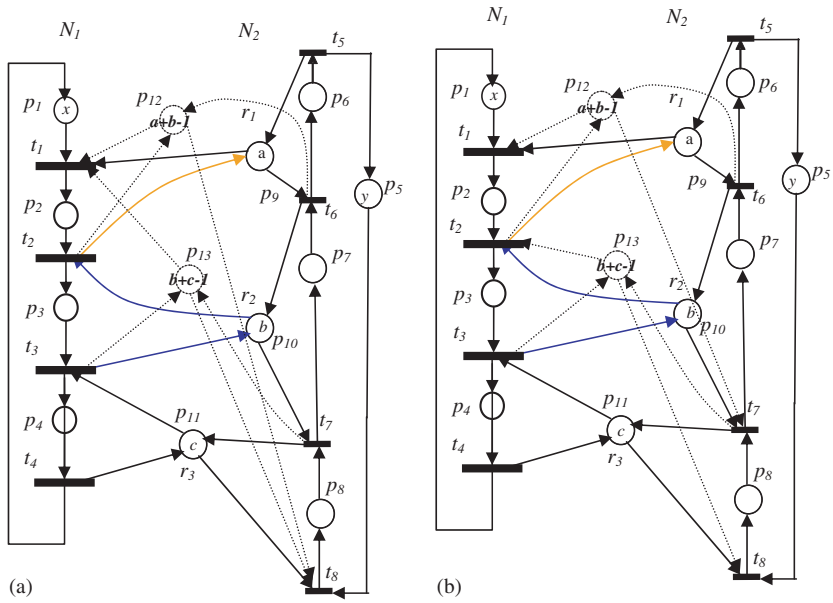

Fig. 1. (a) Example elementary-siphon approach and (b) example alternative approach.

with a place and ends with a transition while a TP-handle starts with a transition and ends with a place.

A handle H to a subnet N' is a directed path from a node n_s in N' to a node n_e in N'; any other node in H is not in N'. In a XY-handle (X, $Y=T$ or P), $n_s \in X$ and $n_e \in Y$. A virtual handle is a handle with only two nodes. We [8] constructed an SMS using the following:

Property 4 ([*8*])**.** (1) N_I is strongly connected. (2) A subnet N' is an I subnet of a minimal siphon iff N' is maximal in the sense that each handle H in N' is a *PP*- or *TP*- or virtual *PT*- handle and there are none of *PP*-, *TP*-, and virtual *PT*- handles to N'; (3) $P(N')$ is an SMS iff there is a nonvirtual *PT*-handle to N'', which is a subnet of N' without any *TP*-handles.

There is a circuit in every N_I since it is strongly connected. Such a circuit c is called a core circuit containing at least two resources since deadlock occurs due to mutual waiting among resources.

The following procedure is based on this property:

Handle-Construction Procedure [8]. Given a core circuit c: (1) add all *PP*'-handles (of the form $[r_1 t_1 p_1 t_2 p_2 \ldots p_{n-1} t_n r_2]$, p_i an operation place, $i=1,2,\ldots,n-1$, $r_1 \in c$ and $r_2 \in c$) to c. The resulting core circuit is called an expanded c^e; (2) add all *PP*- and *TP* handles [that are part of I-subnet of an $\rho(r)$] to c^e to form υ; (3) $P(\upsilon)$ is an SMS if it does not contain an $\rho(r)=H\{r\}\cup\{r), r \in P(\upsilon)$; 4) $P(\upsilon')$ (the set of places in υ') is an SMS if it does not contain any other minimal siphon.

Example. For the net in Fig. 1(a), first find core circuit $c_1=[p_9 t_6 p_{10} t_2 p_9]$. Second add TP-handles (starting from a transition and ending at a place) $[t_2 p_3 t_3 p_{10}]$ and $[t_6 p_6 t_5 p_9]$ to get υ' and $S_1=P(\upsilon')=S_1=\{p_9, p_{10}, p_3, p_6\}$, $S_2=\{p_{10}, p_{11}, p_4, p_7\}$ for $c_2=[p_{10} t_7 p_{11} t_3 p_{10}]$. c_1 and c_2 form a compound circuit (denoted by $\boldsymbol{c}_1 \mathbf{o} \boldsymbol{c}_2$), from which we can synthesize a third $S_3=\{p_4, p_6, p_9, p_{10}, p_{11}\}$, called a compound siphon.

In [9], we propose a polynomial time algorithm to find elementary siphons in a graphical fashion, where we also show that an SMS can be synthesized from a strongly connected resource subnet in an S^3PR. We further prove that each elementary siphon can be synthesized from an elementary resource circuit, while a strongly dependent siphon can be synthesized from a compound resource circuit, which consists of a number of elementary resource circuits $c_1, c_2, \ldots, c_n$ such that $c_i \cap c_{i+1}=\{r_i\}$, $r_i \in P_R$ (*i.e.*, c_i and c_{i+1} intersects at a resource place r_i).

2.2 Elementary siphons and characteristic T-vectors

This section defines elementary, dependent siphons and characteristic T-vectors.

Definition 5 ([*1*])**.** Let $\Omega \subseteq P$ be a subset of places of N. P–vector λ_Ω is called the characteristic P-vector of Ω iff $\forall p \in \Omega$, $\lambda_\Omega(p)=1$; otherwise $\lambda_\Omega(p)=0$. $\boldsymbol{\eta}$ is called

the characteristic T-vector of Ω, if $\boldsymbol{\eta}^T = \lambda_\Omega^T \cdot [\boldsymbol{N}]$, where $[\boldsymbol{N}]$ is the incidence matrix and '·' means a vector or matrix multiplication.

Physically, the firing of a transition t where $[\boldsymbol{\eta}(t)>0, \boldsymbol{\eta}(t)=0$, and $\boldsymbol{\eta}(t)<0]$ increases, maintains and decreases the number of tokens in S, respectively.

Definition 6 ([4]). Let $N=(P,T,\boldsymbol{F})$ be a net with $|P|=m$, which has k siphons $S_1, S_2, \ldots, S_k$, m, $k \in Z^+$, where $Z^+=\{0,1,2,\ldots\}$. Define $[\lambda]_{k\times m} = [\lambda_1|\lambda_2|\cdots|\lambda_k]^T$ and $[\boldsymbol{\eta}]_{k\times n} = [\boldsymbol{\eta}_1|\boldsymbol{\eta}_2|\cdots|\boldsymbol{\eta}_k]^T$. $[\lambda]([\boldsymbol{\eta}])$ is called the characteristic $P(T)$-vector matrix $[\lambda]([\boldsymbol{\eta}])$ of the siphons in N. Let $\boldsymbol{\eta}_{S\alpha}, \boldsymbol{\eta}_{S\beta}, \ldots,$ and $\boldsymbol{\eta}_{S\gamma}(\{\alpha,\beta,\ldots,\gamma\} \subseteq \{1,2,\ldots,k\})$ be a linear independent maximal set of matrix $[\boldsymbol{\eta}]$. Then $\Pi_E = \{S_\alpha, S_\beta, \ldots, S_\gamma\}$ is called a set of elementary siphons. $S \notin \Pi_E$ is called a strongly dependent siphon if $\boldsymbol{\eta}_S = \sum_{Si\in\Pi E} a_i \boldsymbol{\eta}_{Si}$ where $a_i \geq 0$. $S \notin \Pi_E$ is called a weakly dependent siphon if $\exists$ non-empty A, $B \subset \Pi_E$, such that $A \cap B = \emptyset$ and $\eta_S = \sum_{Si\in A} a_i \boldsymbol{\eta}_{Si} - \sum_{Sj\in B} b_j \boldsymbol{\eta}_{Sj}$ where $a_i>0$ and $b_j>0$.

In [1], a strongly dependent siphon is also called a strict redundant one. Li and Zhou [1] propose to find elementary siphons by constructing the characteristic P-vector (resp. T-vector)-vector matrix $[\boldsymbol{\lambda}]$ (resp. $[\boldsymbol{\eta}]$) of the siphons in N followed by finding linearly independent vectors in $[\boldsymbol{\lambda}]$ (resp. $[\boldsymbol{\eta}]$). The siphons corresponding to these independent vectors are the elementary siphons in the net system.

Note that Definition 6 and the above calculation of linearly independent vectors do not assume N to be an S^3PR and are applicable to arbitrary nets.

An example is shown in Fig. 1(a): $S_1=\{p_9, p_{10}, p_3, p_6\}$, $S_2=\{p_{10}, p_{11}, p_4, p_7\}$, and $S_3=\{p_9, p_{10}, p_6, p_{11}, p_4\}$. $\boldsymbol{\eta}_1=[-1\ \ 1\ \ 0\ \ 0\ \ 0\ \ 1\ \ -1\ \ 0]^{\mathrm{T}}$, $\boldsymbol{\eta}_2=[0\ \ -1\ \ 1\ \ 0\ \ 0\ \ 0\ \ 1\ \ -1]^{\mathrm{T}}$, $\boldsymbol{\eta}_3=[-1\ \ 0\ \ 1\ \ 0\ \ 0\ \ 1\ \ 0\ \ -1]^{\mathrm{T}}$. It is easy to see that $\boldsymbol{\eta}_3=\boldsymbol{\eta}_1+\boldsymbol{\eta}_2$ (no negative term) and S_1 and S_2 (resp. S_3) are elementary (resp. strongly dependent) siphons. Fig. 4(a) shows an example of weakly dependent siphon. Its controlled model is shown in Fig. 4(b). Table I below lists the four S and their $\boldsymbol{\eta}$, where $\boldsymbol{\eta}_4=\eta_1+\eta_2-\eta_3$.

2.3 Control policy

This subsection explains the idea of the control policy necessary to underlie the theory presented in Section V. A control policy involves three factors: for each monitor place V_S, (1) its input and output arcs; (2) its initial marking; and (3) the siphon it controls. The following two lemmas are helpful to deal with factor (1).

In Fig. 1(b), we add a control place $V_{S1}=p_{12}$ and the associated arcs for S_1 so that we can consider V_{S1} as another shared resource since the structure involved is similar. V_{S1} plus its holder set of places form the support $\{V_{S1}, p_2, p_7\}=\{V_{S1}\} \cup H(V_{S1})$ of a new P-invariant. Comparing with the support $S_1 \cup [S_1]$ of a P-invariant, we define the controller region $[V_{S1}]=H(V_{S1})$.

Lemma 1. Let (N, M_0) be an ordinary Petri net (PN) system, S an SMS, Monitor V_S with $\boldsymbol{M}_0(V_S)=\boldsymbol{M}_0(S)-1$ is added to S such that V_S and $H(V_S)$ form the support of a new minimal P-invariant $\boldsymbol{I}_S$ associated with S, where $\forall p \in \|\boldsymbol{I}_S\|$, $\boldsymbol{I}_S(p)=1$; $\forall p \in P \backslash \|\boldsymbol{I}_S\|$, $\boldsymbol{I}_S(p)=0$, and $\boldsymbol{M} \in R(N, \boldsymbol{M}_0)$. 1) $\boldsymbol{M}([S])+\boldsymbol{M}(S)=\boldsymbol{M}_0(S)$. 2) $\boldsymbol{M}([V_S])+\boldsymbol{M}(V_S)=\boldsymbol{M}_0(V_S)$. 3) If S is never empty, then $\boldsymbol{M}([S]) \leq \boldsymbol{M}_0(S)-1$. 4) $\boldsymbol{M}([V_S]) \leq \boldsymbol{M}_0(V_S)$.

Proof. 1) $S \cup [S]$ is the support of a P-invariant $\boldsymbol{I}_S$. By the assumption, $\forall p \in \|\boldsymbol{I}_S\|$, $\boldsymbol{I}_S(p)=1$; $\forall p \in P \backslash \|\boldsymbol{I}_S\|$, $\boldsymbol{I}_S(p)=0$. Based on Property 1, $\boldsymbol{I}_S^T \cdot \boldsymbol{M} = \boldsymbol{M}_0(S) \Rightarrow \boldsymbol{M}([S])+\boldsymbol{M}(S)=\boldsymbol{M}_0(S)$. 2) The proof is similar to that for 1). 3) S cannot be empty; hence, $\boldsymbol{M}(S) \geq 1 \Rightarrow \boldsymbol{M}([S]) \leq \boldsymbol{M}_0(S)-1$. 4) $\boldsymbol{M}([V_S])=\boldsymbol{M}_0(V_S)-\boldsymbol{M}(V_S)$ from 2) $\Rightarrow \boldsymbol{M}([V_S]) \leq \boldsymbol{M}_0(V_S)$, since $\boldsymbol{M}(V_S) \geq 0$. □

Lemma 2. Let S be an SMS in a marked $S^3PR(N, \boldsymbol{M}_0)$. Monitor V_S with $\boldsymbol{M}_0(V_S)=\boldsymbol{M}_0(S)-1$ is added to S such that V_S and $H(V_S)$ form the support of a new P-invariant. If siphon S is never empty, then $[S] \subseteq [V_S]$.

Proof. Assume $[S] \supset [V_S] \Rightarrow \boldsymbol{M}([S])>\boldsymbol{M}([V_S]) \Rightarrow \boldsymbol{M}_{max}([S])>\boldsymbol{M}_{max}([V_S])=\boldsymbol{M}_0(V_S)=\boldsymbol{M}_0(S)-1$. It is possible that $M([S]\backslash[V_S])=1$ when $\boldsymbol{M}([S])=\boldsymbol{M}_{max}([S])$. Thus, $\boldsymbol{M}_{max}([S])=\boldsymbol{M}([S]\backslash[V_S])+\boldsymbol{M}_{max}([V_S])=\boldsymbol{M}_0(S) \Rightarrow \boldsymbol{M}(S)=0$ (empty siphon) — contradiction. □

Based on part 1 of Lemma 1 (or Lemma 1.1), we have

$$\boldsymbol{M}_0(S_1)=\boldsymbol{M}(S_1)+\boldsymbol{M}([S_1]) \tag{1}$$

Equation (1) implies that $\boldsymbol{M}_{max}([S_1])=\boldsymbol{M}_0(S_1)$ that happens when $\boldsymbol{M}(S_1)=0$. To avoid unmarked siphons, we set $\boldsymbol{M}(S_1) \geq 1$, implying that $\boldsymbol{M}_{max}([S_1])=\boldsymbol{M}_0(S_1)-1$ or $\boldsymbol{M}_0(S_1)-1 \geq \boldsymbol{M}([S_1])$.

Based on part 2 of Lemma 1 (or Lemma 1.2) and Lemma 2, we have

$$\boldsymbol{M}_0(V_{S1}) = \boldsymbol{M}(V_{S1})+\boldsymbol{M}([V_{S1}])=\boldsymbol{M}(V_{S1}) +\boldsymbol{M}([V_{S1}]\backslash[S_1])+\boldsymbol{M}([S_1]), \tag{2}$$

or equivalently,

$$\boldsymbol{M}([S_1]) = \boldsymbol{M}_0(V_{S1})-\boldsymbol{M}(V_{S1})-\boldsymbol{M}([V_{S1}]\backslash[S_1])$$

$\boldsymbol{M}_{max}([S_1])$ occurs when $\boldsymbol{M}(V_{S1})=\boldsymbol{M}([V_{S1}]\backslash[S_1])=0$; *i.e.*, $\boldsymbol{M}_{max}([S_1])=\boldsymbol{M}_0(V_{S1})=\boldsymbol{M}_0(S_1)-1$. To disturb

Table I. Four SMS in Fig. 4(a) and their η. $\eta_4=\eta_1+\eta_2-\eta_3$.

S	η	Set of places
S_1	$+t_2-t_4+t_8-t_9$	$\{p_4, p_{12}, p_{13}, p_{14}, p_{15}\}$
S_2	$+t_1-t_3+t_7-t_{10}$	$\{p_5, p_{11}, p_{14}, p_{15}, p_{16}\}$
S_3	$+t_2-t_3-t_4+t_7$	$\{p_4, p_{11}, p_{14}, p_{15}\}$
S_4	$+t_1+t_8-t_9-t_{10}$	$\{p_5, p_{12}, p_{13}, p_{14}, p_{15}, p_1\}$

the controller region the least, we should allow $\boldsymbol{M}([S_1])$ to reach its maximum; thus setting $\boldsymbol{M}_0(V_{S1})=\boldsymbol{M}_0(S_1)-1$. In general, $\boldsymbol{M}_0(V_{S1})=\boldsymbol{M}_0(S_1)-\xi_{S1}$, where $\xi_{S1}\geq 1$ is the control depth variable. ξ_{S1} is adjusted to be greater than 1 if some dependent siphons are not controlled. As a result, $\boldsymbol{M}_{max}([S_1])$ is less than $\boldsymbol{M}_0(S_1)-1$ and the controller region is more disturbed causing more states lost.

In [9], M_0 for each control place is set to $\boldsymbol{M}_0(p)=\boldsymbol{M}_0(S)-1$; S is said to be limit controlled since $\boldsymbol{M}_{min}(S)=1$ or $\xi_S=1$.

Definition 5. S is said to reach its limit state when $\boldsymbol{M}(S)=1$; it is limit-controlled iff it is able to reach its limit state but not able to reach unmarked state; *i.e.*, $\xi_S=1$ or $\boldsymbol{M}_{min}(S)=1$.

Based on Lemma 2, we should set $[V_S]=[S]$ to keep the disturbed region as small as possible. To do so, we have $V_S^\bullet={}^\bullet[S]\backslash[S]^\bullet$ and ${}^\bullet V_S=[S]^\bullet\backslash{}^\bullet[S]$ where ${}^\bullet[S]=\{{}^\bullet x|x\in[S]\}$ and $[S]^\bullet=\{x^\bullet|x\in[S]\}$.

Li and Zhou defined $B_S=[V_S]\backslash[S]$ in [8]; they rearranged the output control arcs so that $|B_S|\geq 0$.

Definition 7. Let $(N,\boldsymbol{M}_0)$ be a marked S^3PR. $N=(P\cup P_0\cup P_R, T, \boldsymbol{F})$. It is a disturbanceless (resp. rearrangement) control model if $\forall V_S$, $B_S=\emptyset$ (resp. $|B_S|\geq 0$). It is an SMSless control model, if $\forall V_S$, $V_S^\bullet\subseteq P_0^\bullet$.

The disturbanceless model disturbs the original or uncontrolled model less than the traditional (called the SMSless approach) one in [1] where the support of the new P-invariant associated with V_S covers $[S]\cup\{V_S\}$ as a proper subset. Therefore, the disturbanceless model may reach more states. However, it may create new SMS while the traditional one does not. This paper is mainly concerned with the SMSless approach in [1].

III. MOTIVATION

After the failure of two tests, Li & Zhou adjust control depth variables ξ_S of elementary siphons associated with a dependent siphon to satisfy a Marking Linear Inequality (MLI) as follows:

Theorem 1 (*Theorem 1 in* [*1*])**.** Let $(N_0,\boldsymbol{M}_0)$ be a net system and $S_0, S_1, S_2, \ldots,$ and S_n be its SMS. Assume that S_0 is a strict dependent SMS w.r.t. elementary siphons $S_1, S_2, \ldots,$ and S_n where $\boldsymbol{\eta}_0=\sum_{i=1}^n(a_i\boldsymbol{\eta}_i)$. S_0 is controlled if 1) N_0 is extended by n additional control places $V_{S0}, V_{S1}, V_{S2}, \ldots, V_{Sn}$ such that $S_1, S_2, \ldots$ and S_n are controlled and 2) if $\boldsymbol{M}_0(S_0)>\sum_{i=1}^n(a_i\boldsymbol{M}_0(S_i)-$ $-a_i\xi_{Si})$ where ξ_{Si} is the control depth variable for S_i.

First, they test the above inequality; if it fails, then they perform the following Linear Integer Programming (LIP) test (NP-complete problem):

Theorem 2 (*Theorem 5 in* [*1*])**.** Let $(N_0,\boldsymbol{M}_0)$ be a marked S^3PR and $S_1, S_2, \ldots, S_n$ be the elementary siphons of N_0. By the method stated in Definition 8 in [1], add control places to make elementary siphons controlled. The extended net system is denoted by $(N_1,\boldsymbol{M}_1)$. Let $\theta=\{\boldsymbol{I}_i|i=1,2,\ldots,m\}$ be the set of minimal P-invariants of N_1 and S be an SMS of N_0. S can never be emptied if $\boldsymbol{M}_0(S)>max\{\sum_{p\in[S]}\boldsymbol{M}(p)|\boldsymbol{M}\in R(N_1\boldsymbol{M}_{01})\}$ where $\max\{\sum_{p\in[S]}\boldsymbol{M}(p)\}$ is obtained by linear integer programming (LIP)

$$max \qquad \{\sum_{p\in[S]}\boldsymbol{M}(p)\}$$

subject to

$$\boldsymbol{I}_1^T\cdot\boldsymbol{M}=\boldsymbol{I}_1^T\cdot\boldsymbol{M}_0$$
$$\boldsymbol{I}_2^T\cdot\boldsymbol{M}=\boldsymbol{I}_2^T\cdot\boldsymbol{M}_0$$
$$\ldots\ldots$$
$$\boldsymbol{I}_m^T\cdot\boldsymbol{M}=\boldsymbol{I}_m^T\cdot\boldsymbol{M}_0$$

In [5], Li & Zhou further improve the test in Theorem 1 by establishing more general conditions under which a dependent siphon can be always marked. It is based on a linear programming problem (LPP), where the marking of a place can hold real (instead of integer) numbers, thus avoiding the NP-complete problem. However the improved condition is only sufficient, but not necessary. That is, even if the condition fails, the dependent siphon may remain unmarked. To

decide whether to adjust control depth variables, the LIP in Theorem 2 may still need to be performed.

We will develop a better (sufficient and necessary) test than that in Theorem 1 so that one adjusts control depth variables if and only if the new test fails. This avoids the time-consuming integer programming test completely whether the new test fails or not.

Definition 8. Let A be a set of operation places. $R(A)=\{r|p\in A, p\in H(r)\}$, where $H(r)$ denotes the set of holders of r (operation places that use r). Let B be a set of resource places, $\boldsymbol{M}(B)=\sum_{r\in B}\boldsymbol{M}(r)$. Let $\boldsymbol{I}_V$ (resp. $\boldsymbol{I}_S$) be the minimal P-invariant associated with control place V (resp. siphon S). $[V]=\|\boldsymbol{I}_V\|\backslash\{V\}$ and complementary siphon $[S]=\|\boldsymbol{I}_S\|\backslash S$.

In Fig. 1(a), $[S_1]=\{p_2, p_7\}$, $[S_2]=\{p_3, p_8\}$, $[S_3]=\{p_2, p_3, p_7, p_8\}$, $[V_{S1}]=\{p_2, p_7, p_8\}$, $[V_{S2}]=\{p_2, p_3, p_8\}$, and the MLI: $\boldsymbol{M}_0(S_3)>\boldsymbol{M}_0(S_1)-\xi_{S1}+\boldsymbol{M}_0(S_2)-\xi_{S2}$. Let $\boldsymbol{M}_0(p_9)=a$, $\boldsymbol{M}_0(p_{10})=b$, and $\boldsymbol{M}_0(p_{11})=c$. We say S_i $(i=1,2)$ reaches its limit state when $\boldsymbol{M}(S_i)=1$; it is limit-controlled if it is able to reach its limit state but not able to reach empty status. Note that when $[V_{Si}]=[S_i]$, S_i is limit-controlled if $\xi_{Si}=1$.

Note that output control arcs from monitors p_{12} and p_{13} end at source transitions [t_1 and t_8, respectively in Fig. 1(a)] of the processes, rather than sink transitions [$t_2\in S_1^\bullet$ and $t_7\in S_2^\bullet$, respectively in Fig. 1(b)] of siphons S_1 and S_2, respectively. This is to avoid new SMS generation [*i.e.*, $S_c=\{p_{12}, p_{13}, p_3, p_7\}$ in Fig. 1(b)] that may lead to deadlocks [*i.e.*, when $b=1$ in Fig. 1(b)] if no monitor is added.

Consider the case: $b=\boldsymbol{M}_0(r_2)=2$. Note that $\boldsymbol{M}_0(V_{S1})=a+b-1$, $\boldsymbol{M}_0(V_{S2})=b+c-1$, and $\boldsymbol{M}_0(S_3)=a+b+c=\boldsymbol{M}_0(S_1)-\xi_{S1}+\boldsymbol{M}_0(S_2)-\xi_{S2}=((a+b)-1)+((b+c)-1)$($\xi_{S1}=\xi_{S2}=1$, both S_1 and S_2 are limit-controlled); thus S_3 does not satisfy the MLI. Thus, the time-consuming integer programming test must be started when Li and Zhou's method is used. However, S_3 is controlled and the net is live if $\boldsymbol{M}_0(r_2)<a+c+2$. This discrepancy arises from the fact that the MLI in Theorem 1 assumes that $[V_{Si}]=[S_i]$, while in [1], $[V_{Si}]\supset[S_i]$ $(i=1,2)$.

To understand this, a theoretical analysis is performed below. Set $b=b_1+b_2$, where $b_1=\boldsymbol{M}(p_3)$ and $b_2=\boldsymbol{M}(p_7)$. First we explore the condition under which S_3 is emptied. In order to empty S_3, all tokens in p_9 and p_{11} must go to p_2 and p_8, respectively, and all tokens in p_{10} must distribute to p_3 and p_7. Thus,

$$\boldsymbol{M}(p_2)=a,\quad \boldsymbol{M}(p_8)=c,\quad \boldsymbol{M}(V_{S1})=\boldsymbol{M}(V_{S2})=0,$$

$$\boldsymbol{M}_0(V_{S1})=c+b_2+a=a+b-\xi_{S1}\Rightarrow b_2=b-\xi_{S1}-c$$
$$\Rightarrow b_1=c+\xi_{S1}$$
$$\boldsymbol{M}_0(V_{S2})=c+b_1+a=b+c-\xi_{S2}\Rightarrow b_1=b-\xi_{S2}-a$$
$$\Rightarrow b_2=a+\xi_{S2}$$

Adding the two equations, we have $\boldsymbol{M}_0(r_2)=b=c+a+\xi_{S1}+\xi_{S2}\geq c+a+2$ since $\xi_{S1}\geq 1$ and $\xi_{S2}\geq 1$. $b=c+a+2$ is the condition to empty S_3 when $\xi_{S1}=\xi_{S2}=1$, and the condition for S_3 to be marked (or controlled) is

$$b<c+a+2. \tag{3}$$

Note that to empty S_3, $b=c+a+2$, $b_1=c+1$ and $b_2=a+1$ when $[V_{S1}]\supset[S_1]$ $(i=1,2)$, versus $b=2$, $b_1=1$ and $b_2=1$ when $[V_{S1}]=[S_1]$ $(i=1,2)$. b is increased by

$$c+a=\boldsymbol{M}_0(R([V_{S1}]\backslash[S_1]))+\boldsymbol{M}_0(R([V_{S2}]\backslash[S_2])).$$

Physically, $[V_{Si}]$, $i=1,2$, covers more places than $[S_i]$ due to the movement of output nodes of control arcs to output, called source, transitions of idle places. As a result, when all tokens in $[V_{Si}]$ $(i=1,2)$ are used to trap tokens, some of $[V_{Si}]$ are in $[V_{Si}]\backslash[S_i]$ $(i=1,2)$ and fewer tokens are in $[S_3]$ than the case when $[V_{Si}]=[S_i]$.

This reduces the number of tokens in V_{Si} to trap the tokens in S_3. To compensate for this, we increase $\boldsymbol{M}_0(V_{Si})$ via increasing b by Δb_i (called compensation factor); $\Delta b=\Delta b_1+\Delta b_2=\Delta V_{S1}+\Delta V_{S2}=c+a$. Thus b is increased to $c+a+2$ (from 2) to empty S_3. And neither S_1 nor S_2 can be limit-controlled since $\boldsymbol{M}(S_1)\geq\boldsymbol{M}(p_3)=b_1=c+1>1$ and $\boldsymbol{M}(S_2)\geq\boldsymbol{M}(p_7)=b_2=a+1>1$. Thus, it seems that the MLI can now be modified to $\boldsymbol{M}_0(S_3)>(\boldsymbol{M}_0(S_1)-(\xi_{S1}+c))+(\boldsymbol{M}_0(S_2)-(\xi_{S2}+a))$.

To extend to more general cases, we should consider $\boldsymbol{M}_0(R(([V_{Si}]\cap[S_3])\backslash[S_i]))$ rather than $\boldsymbol{M}_0(R([V_{Si}]\backslash[S_i]))$. This is because in order to empty S_3, tokens in V_{Si} may not need to be trapped in $[V_{Si}]\backslash[S_3]$. For instance, in Fig. 2, set $\boldsymbol{M}_0(p_7)=a$, $\boldsymbol{M}_0(p_8)=b$, $\boldsymbol{M}_0(p_9)=c$, $\boldsymbol{M}_0(p_{10})=d$, and $\boldsymbol{M}_0(p_{11})=e$. $S_1=\{p_7, p_8, p_3, p_2'\}$, $S_2=\{p_8, p_9, p_4, p_3'\}$, $S_3=\{p_7, p_8, p_9, p_4, p_2'\}$, $[S_1]=\{p_2, p_3'\}$, $[S_2]=\{p_3, p_4'\}$, and $[S_3]=\{p_2, p_3', p_3, p_4'\}$44. S_1 and S_2 (resp. S_3) are elementary (resp. dependent) siphons and $\boldsymbol{\eta}_3=\boldsymbol{\eta}_1+\boldsymbol{\eta}_2$. $[V_{S1}]=\{p_2, p_3', p_4', p_5', p_6'\}$, $[V_{S2}]=\{p_2, p_3, p_4', p_5', p_6'\}$,

$\boldsymbol{M}_0(R(([V_{S1}]\cap[S_3])\backslash[S_1]))=\boldsymbol{M}_0(R(\{p_4'\}))=\boldsymbol{M}_0(p_9)=c$ and $\boldsymbol{M}_0(R(([V_{S2}]\cap[S_3])\backslash[S_2]))=\boldsymbol{M}_0(R(\{p_2\}))=\boldsymbol{M}_0(p_7)=a$. Thus, it remains that $b=c+a+2$ rather than $b=c+a+d+e+2$. This is because when S_3 is empty, the markings of p_{10} and p_{11} may remain at their initial ones.

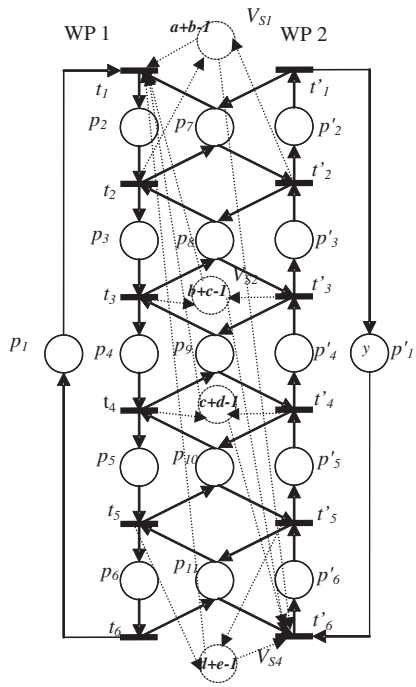

Fig. 2. Example 2.

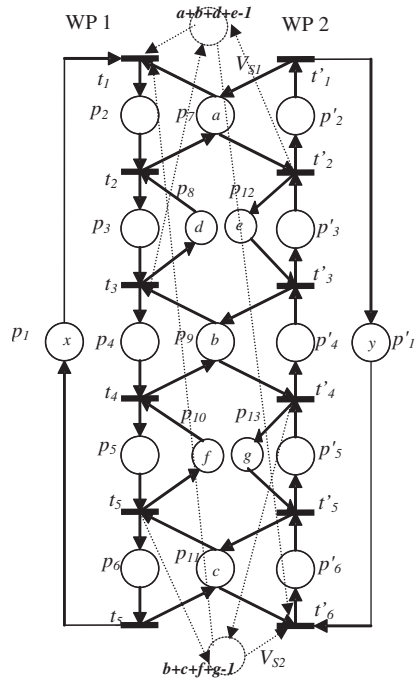

Fig. 3. Example 3.

Another example is shown in Fig. 3. Set $\boldsymbol{M}_0(p_7)=a$, $\boldsymbol{M}_0(p_9)=b$, $\boldsymbol{M}_0(p_{11})=c$, $\boldsymbol{M}_0(p_8)=d$, $\boldsymbol{M}_0(p_{10})=f$, $\boldsymbol{M}_0(p_{12})=e$, and $\boldsymbol{M}_0(p_{13})=g$. $S_1=\{p_7, p_8, p_9, p_{12}, p_4, p'_2\}$, $[S_1]=\{p_2, p_3, p'_3, p'_4\}$; $S_2=\{p_9, p_{10}, p_{11}, p_{13,}, p_6, p'_4\}$, $[S_2]=\{p_4, p_5, p'_5, p'_6\}$; $S_3=\{p_7, p_8, p_9, p_{10}, p_{11}, p_{12}, p_{13,}, p_6, p'_2\}$, $[S_3]=\{p_2, p_3, p_4, p_5, p'_3, p'_4, p'_5, p'_6\}$. S_1 and S_2 (resp. S_3) are elementary (resp. dependent) siphons and $\boldsymbol{\eta}_3=\boldsymbol{\eta}_1+\boldsymbol{\eta}_2$. $[V_{S1}]=\{p_2, p_3, p'_3, p'_4, p'_5, p'_6\}$, $[V_{S2}]=\{p_2, p_3, p_4, p_5, p'_5, p'_6\}$, $\boldsymbol{M}_0(R(([V_{S1}]\cap[S_3])\backslash[S_1]))=\boldsymbol{M}_0(R(\{p'_5, p'_6\}))=\boldsymbol{M}_0(p_{11})+\boldsymbol{M}_0(p_{13})=c+g$ and $\boldsymbol{M}_0(R(([V_{S2}]\cap[S_3])\backslash[S_2]))=\boldsymbol{M}_0(R(\{p_2, p_3\}))=\boldsymbol{M}_0(p_7)+\boldsymbol{M}_0(p_8)=a+d$. Thus, $b=c+a+d+g+2$ rather than $b=c+a+2$. Note that, unlike that in Fig. 2, each of S_1 and S_2 contains a resource place (p_8 and p_{10}, respectively) that is not shared, but used by a single WP.

In summary, we have the following:

Observation 1. Let $(N_0, \boldsymbol{M}_0)$ be a marked S^3PR and S_3 a dependent siphon w.r.t. elementary siphons S_1 and S_2 such that $\boldsymbol{\eta}_3=\boldsymbol{\eta}_1+\boldsymbol{\eta}_2$. By Definition 8 in [1], add 2 control places such that S_1 and S_2 are controlled with control depth variables ξ_{S1} and ξ_{S2}, respectively. S_3 can never be emptied if $\boldsymbol{M}_0(S_3)>(\boldsymbol{M}_0(S_1)-\mu_{S1}-\xi_{S1})+(\boldsymbol{M}_0(S_2)-\mu_{S2}-\xi_{S2})$ where

$$\mu_{S1}=\boldsymbol{M}_0(R(([V_{S1}]\cap[S_3])\backslash[S_1])),$$

$$\mu_{S2}=\boldsymbol{M}_0(R(([V_{S2}]\cap[S_3])\backslash[S_2]))$$

$$\textit{and} \quad \boldsymbol{M}_0(S_i)-\mu_{Si}\geq\xi_{Si}\geq 1, \quad i=1,2.$$

In order for $\boldsymbol{M}([S_3])$ to reach maximal, tokens from S_3 should all be trapped in $[S_3]$; *i.e.*, there are no tokens trapped in places except for $[S_3]$; hence $\boldsymbol{M}([V_{S3}]\backslash[S_3])=0$. However this cannot be extended to S_1 and S_2 such that $\boldsymbol{M}([V_{S1}]\backslash[S_1])=0$ and $\boldsymbol{M}([V_{S2}]\backslash[S_2])=0$ (assumed in Theorem 1) since some tokens in $[S_3]$ fall in $[V_{S1}]\backslash[S_1]$ and $[V_{S2}]\backslash[S_2]$, respectively; *i.e.*, $\boldsymbol{M}(([V_{S1}]\cap[S_3])\backslash[S_1])\neq 0$ and $\boldsymbol{M}(([V_{S2}]\cap[S_3])\backslash[S_2])\neq 0$ and the controllability in Theorem 1 must be modified considering these two terms. We will prove and generalize this in Section IV.

Note that the presence of μ_{Si}, if positive, makes the MLI test in Theorem 1 inaccurate and hence induces the subsequent LIP test in Theorem 2. $\mu_{Si}>0$, if S_1 and S_2 are vertically stacked as shown in Figs. 1–3 (*i.e.*, sink transitions of S_1 and S_2 are in the same processes). Such a dependent siphon $S_3=S_1oS_2$ is called a Type I dependent siphon. On the other hand, $\mu_{Si}=0$, if S_1 and S_2 are horizontally stacked when sink transitions of S_1 and S_2 are in different processes. In this case, there is no need for LIP test since the MLI test is accurate. Such a dependent siphon $S_3=S_1oS_2$ is called a Type II dependent siphon. In the sequel, we will assume all dependent siphons belong to Type I.

IV. THEORY

We first propose the basic theory below to decide whether a siphon is dependent.

Definition 9. An n-dependent siphon is a dependent siphon depending on n elementary siphons.

To further explore the controllability for an n-dependent siphon, $n>2$, specific cases of $n=3$ and $n=4$ will be presented. From which, a general theorem is proposed to conclude that any n-dependent siphon, $n>2$ is already controlled and needs no monitor if every elementary siphon is limit-controlled.

Theorem 3. Let $(N_0, \boldsymbol{M}_0)$ be a net system and S_0 be a dependent SMS w.r.t. elementary siphons $S_1, S_2, \ldots, S_n, S_{n+1}S_{n+2}, \ldots,$ and S_{n+m} where

$$\boldsymbol{\eta}_{S0} = \sum_{i=1}^{n}(a_i\eta_{S_i}) - \sum_{j=1}^{m}(b_{n+j}\eta_{S_{n+j}})$$

Then

(1) $\forall S \in \{S_0, S_1, S_2, \ldots, S_n, S_{n+1}S_{n+2}, \ldots, S_{n+m}\}$,

$\boldsymbol{\eta}_S = -\boldsymbol{\eta}_{[S]}$ (characteristic T -vector of the complementary set of siphon S equals the negative of that of S).

(2) $\lambda_{[S0]} = a_1\lambda_{[S1]} + a_2\lambda_{[S2]} + \cdots + a_n\lambda_{[Sn]} - b_{n+1}\lambda_{[Sn+1]} - b_{n+2}\lambda_{[Sn+2]} - \cdots - b_{n+m}\lambda_{[Sn+m]}$, where $a_i, b_j \in R$ (set of real numbers), $i \in \{1, 2, \ldots, n\}$ and $j \in [1, 2, \ldots m]$ (characteristic P-vectors of the complementary sets of siphon $S_0, S_1, S_2, \ldots, S_n, S_{n+1}, S_{n+2}, \ldots, S_{n+m}$ follow the same equation as that of the corresponding characteristic T-vectors).

(3) $\boldsymbol{M}([S_0]) = a_1\boldsymbol{M}([S_1]) + a_2\boldsymbol{M}([S_2]) + \cdots + a_n\boldsymbol{M}([S_n]) - b_{n+1}\boldsymbol{M}([S_{n+1}]) - b_{n+2}\boldsymbol{M}([S_{n+2}]) - \cdots - b_{n+m}\boldsymbol{M}([S_{n+m}])$, $\boldsymbol{M} \in R(N, \boldsymbol{M}_0)$

(total tokens in the complementary sets of siphon $S_0, S_1, S_2, \ldots, S_n, S_{n+1}S_{n+2}, \ldots, S_{n+m}$ follow the same equation as that of the corresponding characteristic T-vectors). (4)

Proof.

(1) $S \cup [S] = S_R \cup (\cup_{r \in SR} H(r))$ is the support of a P-invariant $\boldsymbol{I}$ based on Property 2 and $S \cap [S] = \emptyset$. Note that $S_R = S \cap P_R$. $\forall p \in S \cup [S]$, $\boldsymbol{I}(p) = 1$; otherwise, $\boldsymbol{I}(p) = 0$. Thus,

$$\boldsymbol{I} = \lambda_S + \lambda_{[S]} \cdot \boldsymbol{I}^T \cdot [\boldsymbol{N}] = \lambda_S^T \cdot [\boldsymbol{N}] + \lambda_{[S]}^T \cdot [\boldsymbol{N}] = 0$$

(By the definition of P-invariant)

$$\Rightarrow \boldsymbol{\eta}_S = -\boldsymbol{\eta}_{[S]}.$$

(2) Based on equations $\eta_{S_0} = \sum_{i=1}^{n}(a_i\eta_{S_i}) - \sum_{j=1}^{m}(b_{n+j}\eta_{S_{n+j}})$, the fact that $\boldsymbol{\eta}_S = -\boldsymbol{\eta}_{[S]}$ and $\boldsymbol{\eta}_S^T = \lambda_S^T \cdot [\boldsymbol{N}]$, we have

$$\begin{aligned}
\boldsymbol{\eta}_{[S0]} &= a_1\boldsymbol{\eta}_{[S1]} + a_2\boldsymbol{\eta}_{[S2]} + \cdots + a_n\boldsymbol{\eta}_{[Sn]} \\
&\quad - b_{n+1}\boldsymbol{\eta}_{[Sn+1]} - b_{n+2}\boldsymbol{\eta}_{[Sn+2]} - \cdots \\
&\quad - b_{n+m}\boldsymbol{\eta}_{[Sn+m]} \\
&\Rightarrow \lambda_{[S0]}^T \cdot [\boldsymbol{N}] = a_1\lambda_{[S1]}^T \cdot [\boldsymbol{N}] + a_2\lambda_{[S2]}^T \cdot [\boldsymbol{N}] + \cdots \\
&\quad + a_n\lambda_{[Sn]}^T \cdot [\boldsymbol{N}] - b_{n+1}\lambda_{[Sn+1]}^T \cdot [\boldsymbol{N}] \\
&\quad - b_{n+2}\lambda_{[Sn+2]}^T \cdot [\boldsymbol{N}] - \cdots - b_{n+m}\lambda_{[Sn+m]}^T \cdot [\boldsymbol{N}] \\
&\Rightarrow (\lambda_{[S0]} - a_1\lambda_{[S1]} - a_2\lambda_{[S2]} - \cdots - a_n\lambda_{[Sn]} \\
&\quad + b_{n+1}\lambda_{[Sn+1]} + b_{n+2}\lambda_{[Sn+2]} + \cdots \\
&\quad + b_{n+m}\lambda_{[Sn+m]})^T \cdot [N] = \mathbf{0} \text{ (a vector with all components being 0).}
\end{aligned}$$

If $\boldsymbol{\zeta} = \lambda_{[S0]} - a_1\lambda_{[S1]} - a_2\lambda_{[S2]} - \cdots - a_n\lambda_{[Sn]} + b_{n+1}\lambda_{[Sn+1]} + b_{n+2}\lambda_{[Sn+2]} + \cdots + b_{n+m}\lambda_{[Sn+m]} \neq \mathbf{0}$, then ζ is a P-invariant. However, all places in $[S_0]$, $[S_1]$, $[S_2]$, …, and $[S_{n+m}]$ are not marked in the initial marking of N and hence the union of $[S_0]$, $[S_1]$, $[S_2]$, …, and $[S_{n+m}]$ cannot be the support of a P-invariant. This implies that

$$\begin{aligned}
\boldsymbol{\zeta} = \mathbf{0} \Rightarrow \lambda_{[S0]} &= a_1\lambda_{[S1]} + a_2\lambda_{[S2]} + \cdots + a_n\lambda_{[Sn]} \\
&\quad - b_{n+1}\lambda_{[Sn+1]} - b_{n+2}\lambda_{[Sn+2]} \\
&\quad - \cdots - b_{n+m}\lambda_{[Sn+m]}.
\end{aligned}$$

(3) Multiplying both sides of the equation in (2) by $\boldsymbol{M}^T$, we have

$$\begin{aligned}
&\lambda_{[S0]} \cdot \boldsymbol{M}^T \\
&= a_1\lambda_{[S1]} \cdot \boldsymbol{M}^T + a_2\lambda_{[S2]} \cdot \boldsymbol{M}^T + \cdots + a_n\lambda_{[Sn]} \cdot \boldsymbol{M}^T \\
&\quad - b_{n+1}\lambda_{[Sn+1]} \cdot \boldsymbol{M}^T - b_{n+2}\lambda_{[Sn+2]} \cdot \boldsymbol{M}^T - \cdots \\
&\quad - b_{n+m}\lambda_{[Sn+m]} \cdot \boldsymbol{M}^T \\
&\Rightarrow \boldsymbol{M}([S_0]) = a_1\boldsymbol{M}([S_1]) + a_2\boldsymbol{M}([S_2]) + \cdots
\end{aligned}$$

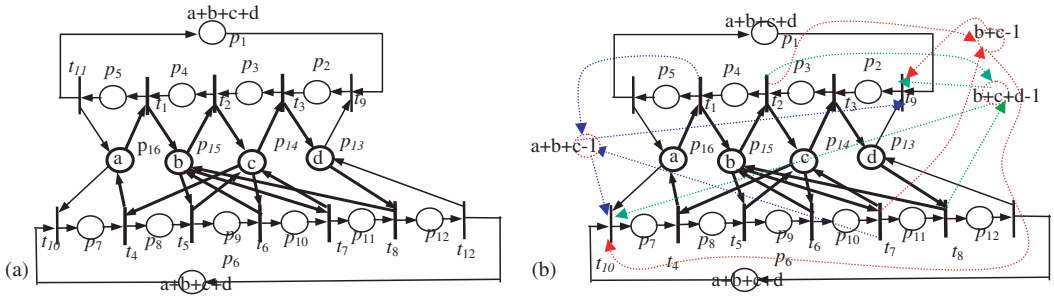

Fig. 4. (a) Example weakly dependent siphon [2] and (b) Controlled model of that in Fig. 4(a).

$$+a_n\boldsymbol{M}([S_n])-b_{n+1}\boldsymbol{M}([S_{n+1}]$$
$$-b_{n+2}\boldsymbol{M}([S_{n+2}])-\cdots-b_{n+m}\boldsymbol{M}([S_{n+m}]).$$

□

For instance, in Fig. 4(a), $\boldsymbol{M}([S_4])=\boldsymbol{M}([S_1])+\boldsymbol{M}([S_2])-\boldsymbol{M}([S_3])$. In the sequel, we will consider only strongly dependent siphons. Furthermore, as indicated in [1], when S_0 is a strongly dependent siphon in an S^3PR, all $a_i=1$. Note that in an S^3PR, an SMS can be synthesized from a strongly connected resource subnet and any strongly dependent siphon corresponds to a compound circuit where the intersection between any two elementary circuits is at most a resource place [8].

We will derive the controllability for $S_0=S_1oS_2o\ldots oS_n$ (denoting that S_0 strongly depends on $S_1, S_2\ldots$, and S_n). In order to compute the exact MLI to avoid the subsequent linear integer programming test, we need to find the exact $\boldsymbol{M}_{max}([S_0])$.

Lemma 3. Let $S_0=S_1oS_2o\ldots oS_n$.

(1) $\boldsymbol{M}([S_0])=\boldsymbol{M}([S_1])+\boldsymbol{M}([S_2])+\cdots+\boldsymbol{M}([S_n])$. (5)
(2) $[S_0]\subseteq[V_{S1}]\cup[V_{S2}]\cup\ldots\cup[V_{Sn}]$
(3) If $\boldsymbol{M}_{max}([S_0])$ occurs, then $\forall S_i$, $\boldsymbol{M}([V_{Si}]\backslash[S_0])=0$; *i.e.*, there are no tokens in places outside $S_0\cup[S_0]$

Proof.

(1) It follows from Theorem 3.3 and that all $a_i=1$.
(2) From [1] or Property 3, we have

$$[S_0]=[S_1]\cup[S_2]\cup\ldots\cup[S_n],$$

since each place p appears in $[S_0]$ only once, there is only one $[S_j]$ among all $[S_i]$ that contains p. This equation together with the fact that $[S_i]\subseteq[V_{Si}]$ [Otherwise, it may be that $\boldsymbol{M}([S_i]\backslash[V_{Si}])=\boldsymbol{M}_0(S_i)$ and S_i is empty] lead to

$$[S_0]\subseteq[V_{S1}]\cup[V_{S2}]\cup\ldots\cup[V_{Sn}]$$

(3) $\boldsymbol{M}([V_{S1}]\cup[V_{S2}]\cup\ldots\cup[V_{Sn}])$

$$=\boldsymbol{M}(([V_{S1}]\cup[V_{S2}]\cup\ldots\cup[V_{Sn}])\backslash[S_0])$$
$$+\boldsymbol{M}([S_0])$$
$$\Rightarrow\boldsymbol{M}([S_0])=\boldsymbol{M}([V_{S1}]\cup[V_{S2}]\cup\ldots\cup[V_{Sn}])$$
$$-\boldsymbol{M}(([V_{S1}]\cup[V_{S2}]\cup\ldots\cup[V_{Sn}])\backslash[S_0])$$
$$\Rightarrow\boldsymbol{M}_{max}([S_0])=\boldsymbol{M}_{max}([V_{S1}]\cup[V_{S2}]$$
$$\cup\ldots\cup[V_{Sn}])-\boldsymbol{M}_{min}(([V_{S1}]\cup[V_{S2}]$$
$$\cup\ldots\cup[V_{Sn}])\backslash[S_0])$$
$$=\boldsymbol{M}_{max}([V_{S1}]\cup[V_{S2}]\cup\ldots\cup[V_{Sn}]),$$

where we have set

$$\boldsymbol{M}_{min}(([V_{S1}]\cup[V_{S2}]\cup\ldots\cup[V_{Sn}])\backslash[S_0])=0,$$

which implies that

$$\forall S_i,\quad \boldsymbol{M}([V_{Si}]\backslash[S_0])=0.$$

Thus, in order to reach $\boldsymbol{M}_{max}([S_0])$, it must be that $\forall S_i$, $\boldsymbol{M}([V_{Si}]\backslash[S_0])=0$. □

Theorem 4. Let $S_0=S_1oS_2o\ldots oS_n$ such that $\boldsymbol{\eta}_o=\sum_{i=1}^{n}(a_i\boldsymbol{\eta}_i)$. Then

(1) $\boldsymbol{M}_{max}([S_0])=\sum_{i=1}^{n}(\boldsymbol{M}_0(V_{Si})-\mu_{Si})=\sum_{i=1}^{n}\boldsymbol{M}_0(V_{Si})-\sum_{i=1}^{n}\mu_{Si}=\sum_{i=1}^{n}(\boldsymbol{M}_0(S_i)-(\xi_{Si}+\mu_{Si}))$ (maximum tokens of $[S_0]$ equals the sum of initial marking of each S_i minus the sum of compensation factors and control depth variables). For all i, $\boldsymbol{M}_0(V_{Si})-\mu_{Si}$ should be greater than zero.
(2) S_0 can never be emptied iff

$$\boldsymbol{M}_0(S_0)>\sum_{i=1}^{n}(\boldsymbol{M}_0(S_i)-(\xi_{Si}+\mu_{Si}))$$
$$=\sum_{i=1}^{n}(\boldsymbol{M}_0(S_i)-\xi_{Si})-\sum_{i=1}^{n}\mu_{Si}$$

(initial marking of S_0 is reduced by the sum of compensation factors to make S_0 controlled). (6)

Proof. (1) Since $[V_{Si}] \supseteq [S_i]$, the controller region $[V_{Si}]$ can be separated into two: $[S_i]$ and $[V_{Si}]\backslash[S_i]$. The latter can be further divided into $[V_{Si}]\backslash[S_0]$ and $([V_{Si}]\cap[S_0])\backslash[S_i]$; *i.e.*,

$$[V_{Si}] = ([V_{Si}]\backslash[S_0]) \cup (([V_{Si}]\cap[S_0])\backslash[S_i]) \cup [S_i].$$

Similarly, the marking of the controller region $[V_{Si}]$ is the sum of that of the above three subregions:

$$\boldsymbol{M}([V_{Si}]) = \boldsymbol{M}([V_{Si}]\backslash[S_0]) + \boldsymbol{M}(([V_{Si}]\cap[S_0])\backslash[S_i]) + \boldsymbol{M}([S_i])$$

Rearranging the terms, we have

$$\boldsymbol{M}([S_i]) = \boldsymbol{M}([V_{Si}]) - \boldsymbol{M}([V_{Si}]\backslash[S_0]) - \boldsymbol{M}(([V_{Si}]\cap[S_0])\backslash[S_i]) \quad (7)$$

$\boldsymbol{M}_{max}([S_i])$ occurs when $\boldsymbol{M}([V_{Si}]\backslash[S_0]) = \boldsymbol{M}_{min}([V_{Si}]\backslash[S_0]) = 0$ and $\boldsymbol{M}([V_{Si}]) = \boldsymbol{M}_{max}([V_{Si}]) = \boldsymbol{M}_0(V_{Si})$. Thus,

$$\boldsymbol{M}_{max}([S_i]) = \boldsymbol{M}_0(V_{Si}) - \mu_{Si}$$

where we have set $\mu_{Si} = \boldsymbol{M}(([V_{Si}]\cap[S_0])\backslash[S_i])$ so that the compensation effect occurs the most when $\boldsymbol{M}([V_{Si}]) = \boldsymbol{M}_0(V_{Si})$ and $\boldsymbol{M}(R(S_0)) = 0$.

Substituting the above $\boldsymbol{M}([S_i])$ into (5), we have

$$\boldsymbol{M}_{max}([S_0]) = \sum_{i=1}^{n}(\boldsymbol{M}_0(V_{Si}) - \mu_{Si}) = \sum_{i=1}^{n}(\boldsymbol{M}_0(S_i) - (\xi_{Si} + \mu_{Si})), \quad (8)$$

We are now ready to prove 2):
($\leftarrow$) When Eq. (8) holds, $\boldsymbol{M}([S_0]) < \boldsymbol{M}_0(S_0)$, which implies $\boldsymbol{M}(S_0) \geq 1$ and S_0 can never be emptied. ($\rightarrow$) Assume contrarily that $\boldsymbol{M}_0(S_0) \leq \sum_{i=1}^{n} a_i(\boldsymbol{M}_0(S_i) - (\xi_{Si} + \mu_{Si}))$. Then by (8), $\boldsymbol{M}_{max}([S_0]) \geq \boldsymbol{M}_0(S_0)$ and S_0 is emptied — contradiction. □

This theorem clearly shows that S_0 is easier to be controlled in a Type I n-dependent siphon than a Type II one by an amount of $\sum_{i=1}^{n} \mu_{Si}$. Note that $n = 2$ in Observation 1 is a degenerate case of that in the above theorem.

In the sequel, we will deal with special cases of S_0 being a 3-, and a 4-dependent siphon, respectively. We will infer a general formula and show that S_0 is always controlled and needs no monitor for $n > 2$. For 2-dependent siphon case, we will verify the result in (4) and thus prove it theoretically.

In Fig. 2, let $S_1 = \{p_7, p_8, p_3, p_2'\}$, $S_2 = \{p_8, p_9, p_4, p_3'\}$, and $S_3 = \{p_9, p_{10}, p_5, p_4'\}$. For $S_0 = S_1 o S_2 o S_3$, apply the same method as the examples in Section III (*i.e.*, setting $b = b_1 + b_2$, where $b_1 = \boldsymbol{M}(p_3)$ and $b_2 = \boldsymbol{M}(p_3')$ and $c = c_1 + c_2$, where $c_1 = \boldsymbol{M}(p_4)$, $c_2 = \boldsymbol{M}(p_4')$. To empty S_0, all tokens in p_7 and p_{10} must go to p_2 and p_5' respectively, all tokens in p_8 must distribute to p_3 and p_3', and all tokens in p_9 must distribute to p_4 and p_4'). Thus,

$$\mu_{S1} = d + c_2 = \boldsymbol{M}(p_5') + \boldsymbol{M}(p_4'),$$
$$\mu_{S2} = d + a = \boldsymbol{M}(p_5') + \boldsymbol{M}(p_2),$$
$$\mu_{S3} = a + b_1 = \boldsymbol{M}(p_2) + \boldsymbol{M}(p_3).$$

Summing the above three equations, we have

$$\mu_{S1} + \mu_{S2} + \mu_{S3} = (a + d + b_1 + c_2) + (d + a) = \boldsymbol{M}_0(V_{S2}) + (d + a)$$

where $\mu_{Si} = \boldsymbol{M}(([V_{Si}]\cap[S_0])\backslash[S_i])$, $\boldsymbol{M}(p_2) = a$, and $\boldsymbol{M}(p_5') = d$.

Similarly, for $S_0 = S_1 o S_2 o S_3 o S_4$ ($S_1 - S_3$ are defined above and $S_4 = \{p_{10}, p_{11}, p_6, p_5'\}$)

$$\mu_{S1} = e + c_2 + d_2,$$
$$\mu_{S2} = a + e + d_2,$$
$$\mu_{S3} = a + e + b_1,$$
$$\mu_{S4} = a + b_1 + c_1,$$

Summing the above four equations, we have

$$\mu_{S1} + \mu_{S2} + \mu_{S3} + \mu_{S4} = (a + b_1 + c_2 + d_2 + e) + (a + b_1 + c_1 + d_2 + e) + (e + a) = \boldsymbol{M}_0(V_{S2}) + \boldsymbol{M}_0(V_{S3}) + (e + a)$$

where $\mu_{Si} = \boldsymbol{M}(([V_{Si}]\cap[S_0])\backslash[S_i])$, $\boldsymbol{M}(p_2) = a$, $\boldsymbol{M}(p_3) = b_1$, $\boldsymbol{M}(p_3') = b_2$, $\boldsymbol{M}(p_4) = c_1$, $\boldsymbol{M}(p_4') = c_2$, $\boldsymbol{M}(p_5) = d_1$, $\boldsymbol{M}(p_5') = d_2$, and $\boldsymbol{M}(p_6') = e$.

In general, for $S_0 = S_1 o S_2 \ldots o S_{n-1} o S_n [R(S_1) = \{r_1, r_2\}, \quad R(S_2) = \{r_2, r_3\}, \ldots R(S_n) = \{r_n, r_{n+1}\}]$, we have, as will be proved in Theorem 5,

$$\mu_{S1} + \mu_{S2} + \cdots + \mu_{Sn} = \boldsymbol{M}_0(V_{S2}) + \boldsymbol{M}_0(V_{S3}) + \cdots + \boldsymbol{M}_0(V_{Sn-1}) + (\boldsymbol{M}_0(r_1) + \boldsymbol{M}_0(r_{n+1})) \quad (9)$$

Substituting (9) into (6),

$$\begin{aligned}\boldsymbol{M}_0(S_0) &> \sum_{i=1}^{n}(\boldsymbol{M}_0(S_i)-\xi_{Si})-\sum_{i=1}^{n}\mu_{Si}\\ &= \sum_{i=1}^{n}(\boldsymbol{M}_0(S_i)-\xi_{Si})-(\boldsymbol{M}_0(V_{S2})+\boldsymbol{M}_0(V_{S3})\\ &\quad +\cdots+\boldsymbol{M}_0(V_{Sn-1})+(\boldsymbol{M}_0(r_1)+\boldsymbol{M}_0(r_{n+1}))\\ &= \sum_{i=1}^{n}(\boldsymbol{M}_0(S_i)-\xi_{Si})-(\sum_{i=2}^{n-1}(\boldsymbol{M}_0(S_i)-\xi_{Si})\\ &\quad +\boldsymbol{M}_0(r_1)+\boldsymbol{M}_0(r_{n+1}))\\ &= (\boldsymbol{M}_0(S_1)-\xi_{S1})+(\boldsymbol{M}_0(S_n)-\xi_{Sn})\\ &\quad -(\boldsymbol{M}_0(r_1)+\boldsymbol{M}_0(r_{n+1}))\end{aligned}$$

where we have set $\boldsymbol{M}_0(V_{Si})=\boldsymbol{M}_0(S_i)-\xi_{Si}$.

Rearranging the terms, we have

$$\begin{aligned}&\boldsymbol{M}_0(S_0)+(\boldsymbol{M}_0(r_1)+\boldsymbol{M}_0(r_{n+1}))+(\xi_{S1}+\xi_{Sn})\\ &\quad >\boldsymbol{M}_0(S_1)+\boldsymbol{M}_0(S_n)\end{aligned} \tag{10}$$

For $n=2$ and $\xi_{S1}=\xi_{Sn}=1$, (10) becomes

$$\begin{aligned}&\boldsymbol{M}_0(S_0)+(\boldsymbol{M}_0(r_1)+\boldsymbol{M}_0(r_{n+1}))+(1+1)\\ &\quad >\boldsymbol{M}_0(S_1)+\boldsymbol{M}_0(S_2).\end{aligned}$$

Now making use of the fact that $\boldsymbol{M}_0(S_0)=a+b+c$, $\boldsymbol{M}_0(r_1)=a$, $\boldsymbol{M}_0(r_{n+1})=c$, $\xi_{S1}=\xi_{Sn}=1$, $\boldsymbol{M}_0(S_1)=a+b$, $\boldsymbol{M}_0(S_2)=b+c$, we have

$$a+b+c+a+c+2>a+b+b+c.$$

After algebraic simplification, we have $c+a+2>b$, which is the same as (3).

When $n>2$, $\boldsymbol{M}(S_0)\geq\boldsymbol{M}_0(S_1)+\boldsymbol{M}_0(S_n)$, and the above inequality holds since $(\boldsymbol{M}_0(r_1)+\boldsymbol{M}_0(r_{n+1}))+(\xi_{S1}+\xi_{Sn})>0$. Thus, an n-dependent siphon S_0 is controlled if $n>2$ and needs no monitor.

Note that in the above, we have only N_1 and N_2 so that $b=b_1+b_2$. Similarly, $[V_{Si}]=([V_{Si}]\cap P_1)\cup([V_{Si}]\cap P_2)$; *i.e.*, each of $[V_{Si}]$ and μ_{Sj} and can be divided into two terms: one in the N_1 side and another in the N_2 side; *i.e.*, $[V_{Si}]=[V_{Si}^1]+[V_{Si}^2]$ and $\mu_{Sj}=\mu_{Sj}^1+\mu_{Sj}^2$. In general, we have $N_{i1}, N_{i2},\ldots,N_{iq}$ and $N_{k1}, N_{k2},\ldots,N_{kv}$ so that $b=b_{i1}+b_{i2}+\ldots+b_{iq}+b_{k1}+b_{k2}+\ldots+b_{kv}$. In the sequel, we derive (9) by assuming that resources are shared between N_1 and N_2 and all siphons are Type I to simplify the presentation. The above generalization case can be proved in a similar manner.

Theorem 5. Let $S_0=S_1oS_2o\ldots oS_n$ such that $\boldsymbol{\eta}_0=\sum_{i=1}^{n}\boldsymbol{\eta}_i$ Then $\mu_{S1}+\mu_{S2}+\ldots+\mu_{Sn}=\boldsymbol{M}_0(V_{S2})+\boldsymbol{M}_0(V_{S3})+\ldots+\boldsymbol{M}_0(V_{Sn})+(\boldsymbol{M}_0(r_1)+\boldsymbol{M}_0(r_{n+1}))$.

Proof. Define $[S_j]^i=[S_j]\cap P_i$ (places of $[S_j]$ in P_i), $\boldsymbol{M}^i(S_j)=\boldsymbol{M}(S_j\cap P_i)$ (tokens in both S_i and P_i), and $\mu_{Sj}^i=\boldsymbol{M}((([V_{Si}]\cap P_i)\cap[S_0])\backslash[S_i])$ (portion of μ_{Sj} that are in N_i), where P_i is the set of places in N_i defined in Definition 4.

In order to empty S_0, all tokens in r_1 (resp. r_{n+1}) must distribute in $[S_1]^2$ (resp. $[S_n]^1$) so that

$$\boldsymbol{M}_0(r_1)=\boldsymbol{M}^2([S_1])\ [\text{resp. } \boldsymbol{M}_0(r_{n+1})=\boldsymbol{M}^1([S_n])].$$

We first compute μ_{Sj}^1, $j=1,2,\ldots,n-1$,

$$\begin{aligned}\mu_{S1}^1 &= \boldsymbol{M}^1([S_2])+\boldsymbol{M}^1([S_3])+\cdots+\boldsymbol{M}^1([S_{n-1}])\\ &\quad +\boldsymbol{M}^1([S_n])\\ \mu_{S2}^1 &= \boldsymbol{M}^1([S_3])+\boldsymbol{M}^1([S_4])+\cdots+\boldsymbol{M}^1([S_n])\\ \mu_{S3}^1 &= \boldsymbol{M}^1([S_4])+\cdots+\boldsymbol{M}^1([S_n])\\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ \mu_{Sn-1}^1 &= \boldsymbol{M}^1([S_n]).\\ \mu_{Sn}^1 &= 0\end{aligned}$$

Summing the above terms, we have

$$\begin{aligned}\sum_{j=1}^{n}\mu_{Sj}^1 &= \boldsymbol{M}^1([S_2])+2\boldsymbol{M}^1([S_3])+\cdots+(n-2)\\ &\quad \times\boldsymbol{M}^1([S_{n-1}])+(n-1)\boldsymbol{M}^1([S_n]).\end{aligned} \tag{11}$$

Similarly, for the N_2 side, we have

$$\begin{aligned}\sum_{j=1}^{n}\mu_{Sj}^2 &= \boldsymbol{M}^2([S_{n-1}])+2\boldsymbol{M}^2([S_{n-2}])+\cdots+(n-2)\\ &\quad \times\boldsymbol{M}^2([S_2])+(n-1)\boldsymbol{M}^2([S_1])\end{aligned} \tag{12}$$

Summing the terms in Eqs. (11) & (12), we have, after rearranging the terms,

$$\begin{aligned}&\left(\sum_{j=1}^{n}\mu_{Sj}^1+\sum_{j=1}^{n}\mu_{Sj}^2\right)\\ &\quad =[\boldsymbol{M}^2([S_1])+\boldsymbol{M}^2([S_2])+\boldsymbol{M}^1([S_2])+\boldsymbol{M}^1([S_3])\\ &\qquad +\boldsymbol{M}^1([S_4])+\cdots+\boldsymbol{M}^1([S_n])]+[\boldsymbol{M}^2([S_1])\\ &\qquad +\boldsymbol{M}^2([S_2])+\boldsymbol{M}^2([S_3])+\boldsymbol{M}^1([S_3])+\boldsymbol{M}^1([S_4])\\ &\qquad +\cdots+\boldsymbol{M}^1([S_n])]+\cdots+[\boldsymbol{M}^2([S_1])+\boldsymbol{M}^2([S_2])\end{aligned}$$

$$
\begin{aligned}
&+\cdots+\boldsymbol{M}^2([S_{n-3}])+\boldsymbol{M}^2([S_{n-2}])+\boldsymbol{M}^2([S_{n-1}])\\
&+\boldsymbol{M}^1([S_{n-1}])+\boldsymbol{M}^1([S_n])]+\boldsymbol{M}^2([S_1])\\
&+\boldsymbol{M}^1([S_n])\\
=&\boldsymbol{M}([V_{S2}])+\boldsymbol{M}(V_{S3})+\cdots+\boldsymbol{M}([V_{Sn-1}])\\
&+\boldsymbol{M}_0(r_1)+\boldsymbol{M}_0(r_{n+1})=\boldsymbol{M}_0(V_{S2})+\boldsymbol{M}_0(V_{S3})\\
&+\cdots+\boldsymbol{M}_0(V_{Sn-1})+\boldsymbol{M}_0(r_1)+\boldsymbol{M}_0(r_{n+1}),
\end{aligned}
$$

where

$$
\begin{aligned}
\boldsymbol{M}([V_{S2}])=&\boldsymbol{M}_0(V_{S2})=\boldsymbol{M}^2([S_1])+\boldsymbol{M}^2([S_2])\\
&+\boldsymbol{M}^1([S_2])+\boldsymbol{M}^1([S_3])+\boldsymbol{M}^1([S_4])\\
&+\cdots+\boldsymbol{M}^1([S_n]),\\
\boldsymbol{M}([V_{S3}])=&\boldsymbol{M}_0(V_{S3})=\boldsymbol{M}^2([S_1])+\boldsymbol{M}^2([S_2])\\
&+\boldsymbol{M}^2([S_3])+\boldsymbol{M}^1([S_3])+\boldsymbol{M}^1([S_4])\\
&+\cdots+\boldsymbol{M}^1([S_n]),\\
&\cdots\cdots\\
\boldsymbol{M}([V_{Sn-1}])=&\boldsymbol{M}_0(V_{Sn-1})=\boldsymbol{M}^2([S_1])+\boldsymbol{M}^2([S_2])\\
&+\cdots+\boldsymbol{M}^2([S_{n-3}])+\boldsymbol{M}^2([S_{n-2}])\\
&+\boldsymbol{M}^1([S_{n-1}])+\boldsymbol{M}^2([S_{n-1}])+\boldsymbol{M}^1([S_n]),\\
\boldsymbol{M}_0(r_1)=&\boldsymbol{M}^2([S_1]),\ \text{ and }\ \boldsymbol{M}_0(r_{n+1})=\boldsymbol{M}^1([S_n]).
\end{aligned}
$$

This proves (9). □

The reader may refer to the sentences from the second paragraph behind the proof of Theorem 4 for the specific cases of 3- and 4-dependent siphons to understand the above the proof.

This theorem expresses the sum of compensation factors in terms of known quantities (initial markings). Note that this theorem holds only if $a_1=a_2=\ldots=a_n$, or when the dependent siphon is a compound one.

Remarks. Note that (1) Theorem 5, like Theorem 4 for n-dependent siphons $n>2$, holds only when all $a_i=1$ and all $b_j=1$. When $\mu_{Si}=0\ \forall i\in[1,2,\ldots n]$ and $S_0=S_1oS_2o\ldots oS_n$, there is no need for the time-consuming integer programming test after we perform the MLI test by adjusting control depth variables in an S^3PR. However, if some $\mu_{Si}\neq 0$ for $n>2$, then the dependent siphon is already controlled.

Total Time Complexity. Case (1): Only Type I siphons exist. The worst total time complexity is $O(n^2)$ since only 2-dependent siphon needs to check the new MLI based on Theorem 4.2 and there are at worst $O(n^2)$ 2-dependent siphons, where $n$is the total number of resource places in the net (recall that each SMS in an S^3PR must contain at least two resource places). In practice, Type I strongly 2-dependent occurs between adjacent resource places shared between two processes and there are linear number of 2-dependent siphons as shown in [10]. As a result, the total time complexity to check controllability of all strongly dependent siphons is reduced from exponential to linear. Case (2): Only Type II siphons exist. The time complexity to verify the MLI of a dependent siphon is $O(|\Pi_E|)$ since there are $O(|\Pi_E|)$ terms on the right hand side of the MLI (see the inequality in Theorem (1). They are $|\Pi_D|$ dependent siphons. As a result, the total time complexity is $O(|\Pi_E||\Pi_D|)$, where $|\Pi_E|$ is total number of elementary siphons and $|\Pi_D|$ total number of dependent siphons.

V. S^3PR EXAMPLE

This section compares the proposed approach with the LIP one in [1] based on the well-known S^3PR example. The layout [5] of the flexible manufacturing cell is shown in Fig. 3 in [1] and the Petri net model of the system is shown in Fig. 5.

The net system is an S^3PR and contains deadlocks. There are six corresponding elementary or basic siphons synthesized from six resource circuits using the handle-construction procedure and 12 strongly dependent siphons as shown in Tables I and II in [8] respectively. For example, $S_3$is a dependent SMS w.r.t. to S_4 and S_{18}. To apply the elementary-siphon approach to this net system, we first add six control places V_{S1}, V_{S4}, V_{S10}, V_{S16}, V_{S17}, and V_{S18}which correspond to six elementary siphons S_1, S_4, S_{10}, S_{16}, S_{17}, and S_{18}, respectively.

Among the 12 dependent siphons, eight (S_2, S_3, S_7, S_8, S_9, S_{12}, S_{14}, S_{15}) of them are $n=2$ dependent siphons that need to do the new MLI test and only S_{15} does not satisfy the old MLI when the corresponding control depth variables are 1.

Li and Zhou show that only S_{11}, S_{13}, and S_{15} (wasting time to check S_{11} and S_{13}) do not meet the old MLI and prove that the three SMSs can be controlled through LIP test based on Theorem 2. Note $S_{11}=S_1oS_{16}oS_{17}$, $S_{13}=S_1oS_{16}oS_{18}$ ($n=3$, 3 elementary siphons), and $S_{15}=S_1oS_{16}$ ($n=2$, 2 elementary siphons). Based on the discussion on (10), S_{11} and S_{13} are already controlled and need no monitors since $n=3>2$. For S_{15}, $S_1\cap S_{16}=\{p_{26}\}$, $a=\boldsymbol{M}_0(p_{21})=1$, $b=\boldsymbol{M}_0(p_{26})=2$, and $c=\boldsymbol{M}_0(p_{22})=1$. S_{15} is already

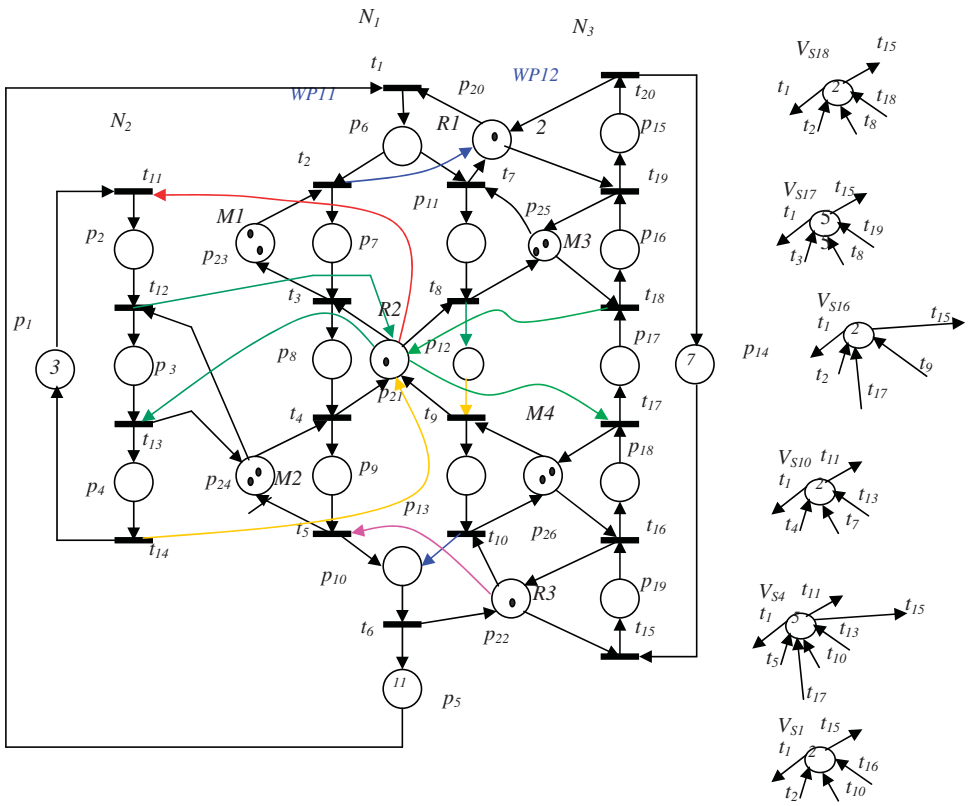

Fig. 5. An S3PR in [1] on the left and its elementary control model on the right.

controlled since $b<c+a+2=4$ by (3). This example illustrates the advantage of our test by avoiding LIP or LPP test.

Remarks. On reconsidering the six control places computed for the well-known S^3PR example, the proposed method cannot deal with redundant control places. Uzam [12] *et al.* propose a redundancy test (the first of its kind) for the liveness enforcing supervisors (LES) of an FMS. When the redundancy test [12] is carried out for this particular example, BFT (back to front; *i.e.*, V_{S18}, V_{S17}, V_{S16}, V_{S10}, V_{S4}, and V_{S1} in that order) and FTB (front to back; *i.e.*, V_{S1}, V_{S4}, V_{S10}, V_{S16}, V_{S17}, and V_{S18} in that order) tests indicate that the control place V_{S17} shown in Fig. 5(b) is redundant (*i.e.*, the system remains live after removing V_{S17}). Moreover, when V_{S17} is removed from the controlled model of Fig. 5(b), the controlled model can even reach more good states than the one (6287) shown in Fig. 5(b). In other words the controlled model obtained by the control places V_{S1}, V_{S4}, V_{S10}, V_{S16}, and V_{S18} is live and can reach 6331 good states. However, it requires reachability analysis to test liveness which takes exponential time complexity negating the advantage of the proposed polynomial approach. It is interesting to find polynomial approaches to remove redundant monitors. Also it is still much less than the maximally permissive one. This is due to the larger controller region (by making control arcs to end at source transitions of processes) causing the original uncontrolled model to be more disturbed. As a matter of the fact, the number of good states provided by a liveness-enforcing supervisor is considered as a kind of quality measure within the literature. A comparison for this benchmark example among the different methods available within the literature in this respect can be seen from [13]. This is to say that it is also necessary to improve this quality measure in addition to computational complexity. It is interesting to extend the proposed approach to the maximally permissive control policy.

VI. CONCLUSION

We have improved the sufficiency test in the elementary siphon approach by Li & Zhou [1] for the special case when $\eta_3=\eta_1+\eta_2$ so that if the modified MLI is satisfied, there is no need for the

ensuing time-consuming linear integer programming test. We have further generalized it to the case where $\eta_0 = \eta_1 + \eta_2 + \cdots + \eta_n$. The MLI needs to be modified by adding a constant μ_{Si} to each control depth variable ξ_{Si} as shown in Theorem 4.2. When $\mu_{Si} = 0$, the MLI is the same as that in [1, 4]; such siphons are called type II ones and there is no need for LIP test.

In addition, we have derived a general formula to show that S_0 is always controlled and needs no monitor for $n > 2$. We have also extended the theory to weakly dependent siphons (to be reported in a future paper) and showed that weakly dependent siphons have similar controllability for both $n = 2$ and $n > 2$ cases. As a result, we need only verify controllability for $n = 2$ based on Theorem 4.2. Therefore, it takes linear time complexity compared with the exponential one in [1].

This paper is both theoretically and practically important. To control an FMS, it reduces the complexity from exponential to polynomial since only $n = 2$ strongly dependent type I ($\mu_{Si} \neq 0$) siphons need to be verified against our new MLI test; the number of which is polynomial. We further prove that our new MLI test is both sufficient and necessary, much better than the only sufficient MLI or LPP (linear programming problem) in [1, 4–7]. This eliminates the LIP completely. Future work should apply to very large systems to enjoy the important theory developed in this paper.

REFERENCES

1. Li, Z. W. and M. C. Zhou, "Elementary siphons of Petri nets and their application to deadlock prevention in flexible manufacturing systems," *IEEE Trans. Syst. Man Cybern. Part A*, Vol. 34, No. 1, pp. 38–51 (2004).
2. Zhang, W., R. P. Judd, and P. Deering, "Necessary and sufficient conditions for deadlocks in flexible manufacturing systems based on a digraph model," *Asian J. Control*, Vol. 6, No. 2, pp. 217–228 (2004).
3. Zhang, W. and R. P. Judd, "Deadlock avoidance for flexible manufacturing systems with choices based on digraph circuit analysis," *Asian J. Control*, Vol. 9, No. 2, pp. 111–120 (2007).
4. Li, Z. W. and M. C. Zhou, "Clarifications on the definitions of elementary siphons in Petri nets," *IEEE Trans. Syst. Man Cybern. Part A*, Vol. 36, No. 6, pp. 1227–1229 (2006).
5. Li, Z. W. and M. C. Zhou, "Control of elementary and dependent siphons in Petri nets and their application," *IEEE Trans. Syst. Man Cybern. Part A*, Vol. 38, No. 1, pp. 133–148 (2008).
6. Li, Z. W. and M. Zhao, "On controllability of dependent siphons for deadlock prevention in generalized Petri nets," *IEEE Trans. Syst. Man Cybern. Part A*, Vol. 38, No. 2, pp. 369–384 (2008).
7. Li, Z. W. and M. C. Zhou, "On siphon computation for deadlock control in a class of Petri nets," *IEEE Trans. Syst. Man Cybern. Part A*, Vol. 38, No. 3, pp. 667–679 (2008).
8. Chao, D. Y., "Computation of elementary siphons in Petri nets for deadlock control," *Comput. J.*, Vol. 49, No. 4, pp. 470–479 (2006).
9. Chao, D. Y., "Incremental approach to computation of elementary siphons for arbitrary S^3PR," *IET Contr. Theory Appl.*, Vol. 2, No. 2, pp. 168–179 (2008).
10. Chao, D. Y., "An incremental approach to extract minimal bad siphons," *J. Inf. Sci. Eng.*, Vol. 23, No. 1, pp. 203–214 (2007).
11. Li, Z. W. and M. C. Zhou, "Two-stage method for synthesizing liveness-enforcing supervisors for flexible manufacturing systems using Petri nets," *IEEE Trans. Ind. Inform.*, Vol. 2, No. 4, pp. 313–325 (2006).
12. Uzam, M., Z. W. Li, and M. C. Zhou, "Identification and elimination of redundant control places in Petri net based liveness enforcing supervisors of FMS," *Int. J. Adv. Manuf. Technol.*, Vol. 35, No. 1–2, pp. 150–168 (2009).
13. Li, Z. W. and M. C. Zhou, *Deadlock Resolution in Automated Manufacturing Systems: a Novel Petri Net Approach*, Springer, London (2009).

D. Y. Chao received the Ph.D. degree from electrical engineering and computer science from the University of California, Berkeley in 1987. From 1987–1988, he worked at Bell Laboratories. Since 1988, he joined the computer and information science department of New Jersey Institute. Since 1994, he joined the MIS department of NCCU as an associate professor. Since February, 1997, he has been promoted to a full professor. He is now working on the optimal control of flexible manufacturing systems.