Abstract—In an earlier paper [1], we introduced the notion of safety control of stochastic discrete event systems (DESs), modeled as controlled Markov chains. Safety was specified as an upper bound on the components of the state probability distribution, and the class of irreducible and aperiodic Markov chains were analyzed relative to this safety criterion. Under the assumption of complete state observations, (i) we identified the set of all state-feedback controllers that enforce the safety specification, for all safe initial probability distributions, and (ii) for any given state-feedback controller, we constructed the maximal invariant safe set. In this paper we extend the work reported in [1] in several ways: (i) Safety is specified in terms of both upper and lower bounds; (ii) We consider a larger class of Markov chains that includes reducible and periodic chains; (iii) We present a more general iterative algorithm for computing the maximal invariant safe set, which is quite flexible in its initialization; (iv) We obtain an explicit upper bound for the number of iterations needed for the algorithm to terminate.

Index Terms—Stochastic discrete event system, Markov chain, Safety specification, Reliability

I. INTRODUCTION

The study of safety control for non-stochastic discrete event systems (DESs) has its origins in the pioneering work of Ramadge-Wonham [12], and has been subsequently extended by other researchers [9]. A non-stochastic DES is typically modeled as a state machine or an automaton which evolves in response to occurrence of events. The supervisory controller dynamically disables certain events so as to achieve the specified qualitative control objective. The safety control objective is typically specified in terms of a set of forbidden states that the system must avoid (or, alternatively, as a set of forbidden event sequences), and a controller enforcing a safety specification ensures that the forbidden states are never visited.

The state machine model of non-stochastic DESs is naturally extended to obtain the Markov chain model of stochastic DESs by associating a probability with each state transition. Prior work on control of stochastic DESs focuses primarily on quantitative control objectives, i.e., on optimal control, where a controller that optimizes a certain performance measure is computed [2], [8]. The optimal control of stochastic systems with state constraints has also been studied in [3], [4], [6]. In this framework, a constraint can be imposed, at least implicitly, on the set of states that the controlled system may visit.
The formalism of probabilistic languages was introduced in [7], aiming to facilitate the study of the qualitative behaviors of stochastic DESs, and the control of such behaviors was studied in [10]. We refer the reader to the bibliography in [7] for other formalisms for modeling qualitative behaviors of stochastic discrete event systems, and their control. There has been some work in the computer science literature on verification of probabilistic systems (see for example [5]), but the focus here is very different. The objective in this work is to verify whether certain properties of the deterministic behavior hold almost surely (i.e., with probability one). In this context, a property is deemed to hold as long as the behaviors violating the property occur with probability zero. This problem can be reduced to that of verification of properties of a deterministic system. By contrast, in our setting, we are interested in verifying whether certain properties of the stochastic behavior hold at all instances of time. Also, whether there exists a state-feedback control enforcing the desired properties of the stochastic behavior.

We introduced the notion of safety control of stochastic DESs in [1], by naturally extending concepts from the non-stochastic setting. A safety specification in the non-stochastic setting may be viewed as a binary valued vector equal in size to the number of states. A state is deemed forbidden if the corresponding entry in that vector is zero. If we represent the states visited under the action of the supervisory control by a binary valued vector, whose zero entries correspond to states that are not visited, then we say that the controller meets the safety control objective if the vector of state visits is dominated by the vector representing the safety specification. In extending this concept to stochastic DESs in [1], we represented the safety specification as a unit-interval valued vector which is meant to impose an upper bound on the components of the state probability distributions of the system under control, thereby naturally generalizing the notion of safety found in the non-stochastic setting. In other words the requirement is that the probability of visiting any given state, at each time step, should not exceed the corresponding bound provided by the safety specification. Thus, a state probability distribution is called safe if it is within the bounds of the safety specification. For example, for a financial portfolio, a natural constraint is that the probability of becoming bankrupt be bounded above by a certain number.

In [1], we embarked on the study of safety control by focusing on the class of irreducible and aperiodic Markov chains. Under the assumption of complete state observations, (i) we identified the set of all state-feedback controllers that enforce the safety specification, for all safe initial probability distributions, and (ii) for any given state-feedback controller, we constructed the maximal invariant safe set (MISS), i.e., those safe initial probability distributions that result in a state trajectory which remains safe at all time steps. In this paper we extend the work reported in [1] in several ways: (i) We allow for safety specifications that impose both an upper as well as a lower bound on the state probability distributions of the system under control. In other words, safe behavior requires that the probability of visiting any given state must lie in the interval specified by the two bounds, at each time step. (ii) We consider a larger class of Markov chains that includes reducible and periodic chains. (iii) We present a more general iterative algorithm for computing the MISS, which is quite flexible in its initialization, i.e., the safe set of probability distributions chosen to start the iterations. (iv) We obtain an explicit upper bound for the number of iterations needed for the algorithm to terminate.

As in [1], we assume complete state observations. We first obtain necessary and sufficient conditions for the trajectory of the controlled system to meet the safety specification at all time steps, for any initial state probability distribution that is safe. These conditions are expressed in terms of a set of $2n$ linear inequalities, where $n$ is the number of states of the Markov chain. Thus the complexity of the criterion is polynomial in the number of states of the chain.

Next, we characterize the MISS, for a given state feedback controller. We show that the Markov chain having a safe invariant probability distribution is necessary and sufficient for the MISS to be non-empty. For Markov chains whose invariant probability distributions are all safe, we provide an algorithm for computing the MISS. The algorithm is guaranteed to terminate in a finite number of iterations, as long as it is initialized on a set containing all the invariant probability distributions of the Markov chain. It also follows from this algorithm that the MISS is a polyhedron, and hence can be represented by finite number of linear inequalities. Furthermore, we provide an upper bound on the number of iterations required for the algorithm to terminate. In order to simplify the algebra, this upper bound is derived for the class of chains that have a unique invariant probability distribution. However, the method extends to more general chains. Lastly, we include a running example to illustrate the results of the paper.

The significance of our setting with respect to the current literature on stochastic control is the consideration of the entire behavior of the stochastic system in evaluating its performance. For a stochastic system modeled by a controlled Markov chain, the control policy is often decided relative to the long-run average behavior, thus ignoring the transient fluctuation of the state probability distribution as it converges to its stationary value. However, for decision makers that are highly sensitive to the system’s performance at each time step (e.g., short-term investment business, or small business that can only survive in the environment of limited supply-demand variation), both the long and short-term behavior of the system are of concern. When using a stationary policy that minimizes the long run average cost, the transient behavior of the system is more influenced by the initial probability distribution than by the policy itself. Our paper addresses this issue by introducing the safety concept, which specifies an allowable range for the state probability distribution at each time step. As a result, the entire behavior of the system is taken into consideration in evaluating its performance.

Two non-trivial examples (Example 3.2 & 4.3) are included to illustrate these ideas. These examples show the potential of our results in applications to systems modeled by controlled Markov chains. As we point out in these examples, the combination of a dual control objective that combines the safety specification with a functional to optimize can sometimes lead
to very economic solutions, e.g., for maintenance problems in manufacturing systems.

The rest of the paper is organized as follows: Section II is a brief introduction to controlled Markov chains. In Section III we characterize the class of safety enforcing state feedback controllers. In Section IV we present an algorithm for computing the MISS for a given state feedback controller, prove that it terminates in finitely many iterations, and derive an upper bound on the number of the required iterations. Section V concludes the paper.

II. NOTATION AND PRELIMINARIES

A controlled Markov chain is represented by the tuple $(S, U, \{P(u)\}_{u \in U}, \pi_0)$, where $S$ is a finite state space composed of $n$ states, i.e., $S = \{1, \ldots, n\}$, $U$ is a finite set of control inputs, or actions, and for each $u \in U$, $P(u) \in [0,1]^{n \times n}$ is a probability transition matrix on $S$. Also, $x_k$ denotes the state variable at time $k$, and $\pi_0$ is the probability distribution of $x_0$. Let $\Pi$ and $\Pi_U$ denote the sets of probability distributions on the state space and action space, respectively. Assuming a complete observation of states, we define the *history space* $H_k$, for each time $k \geq 0$, by $H_0 = S$ and

$$H_{k+1} = H_k \times U \times S = (S \times U)^k \times S, \quad k \geq 0.$$  

In other words the sequence of states visited and actions taken in the past, called the *history*, is an element of the history space. At each time $k$, the history is the information available to the controller, and consequently, an *admissible control policy*, is a sequence $\{\mu_k \mid k \geq 0\}$ of functions $\mu_k : H_k \rightarrow \Pi_U$. The set of all admissible policies is denoted by $\Sigma$. Certain classes of admissible policies are of special interest. A policy $\mu$ is called *Markov* if each $\mu_k$ depends only on $x_k$. A Markov policy is called stationary, if $\mu_k$ does not depend on the time $k$, i.e., if there exists a function $\varphi : S \rightarrow \Pi_U$ such that $\mu_k \equiv \varphi$, for all $k$. In other words, a stationary policy corresponds to a state feedback controller. We denote the class of stationary policies by $\Sigma_S$. If $\varphi$ is a stationary policy, then the resulting state process $\{x_k\}$ is Markov with transition probability matrix $P_{\varphi}$. This can be obtained from $\{P(u)\}_{u \in U}$ as follows. Note that for each $i \in S$, $\varphi(i) = \{\varphi_u(i)\}_{u \in U}$ is a probability distribution vector on $U$. Then $P_{\varphi}$ can be defined as:

$$[P_{\varphi}]_{ij} := \sum_{u \in U} \varphi_u(i) [P(u)]_{ij}.$$  

Let $\pi_k \in \Pi$ denote the state probability distribution at time $k$ and suppose the chain is under the control of a stationary policy $\varphi$. Then the system moves to the next state probability distribution $\pi_{k+1} \in \Pi$ according to the rule $\pi_{k+1} = \pi_k P_{\varphi}$.

A probability distribution $\pi^* \in \Pi$ is said to be an *invariant probability distribution* of $P$ if $\pi^* P = \pi^*$, and $\Pi \subseteq \Pi$ is said to be an *invariant set of probability distributions* of $P$ if $\pi \in \Pi$ implies $\pi P \in \Pi$, or equivalently, $\Pi P \subseteq \Pi$. A pair of states $i$ and $j$ are said to communicate, and this is denoted by $i \leftrightarrow j$, if $[P^k]_{ij} > 0$ and $[P^l]_{ji} > 0$ for some $k, \ell \in \mathbb{N}$. It follows that $\leftrightarrow$ is an equivalence relation on $S$ and the equivalence classes of states are called *communicative classes*. A state $i \in S$ is said to be *recurrent* if $\sum_k [P^k]_{ii} = \infty$. The period of a state $i \in S$, denoted by $d(i)$ is defined as the greatest common divisor of the set $\{k > 0 \mid [P^k]_{ii} > 0\}$. A state $i \in S$ is said to be *aperiodic* if $d(i) = 1$. We also define the period $d$ of $P$ as the least common multiple of the periods $d(i)$ of the individual states. If $P$ has period $d$, then $P^d$ is aperiodic.

For $\pi \in \Pi$, $\lim_{k \to \infty} \pi P^k$, if it exists, is called a *limiting probability distribution* of $P$. A limiting distribution is always an invariant distribution. It is well known that a necessary and sufficient condition for a transition matrix $P$ to possess a unique limiting distribution $\lim_{k \to \infty} \pi P^k$, for any $\pi \in \Pi$, is that $S$ has exactly one communicative class of recurrent states and that this class has period $1$ [13].

III. CONTROLLERS THAT ENFORCE SAFETY

We start with appropriate definitions of safety.

Definition 3.1: A *safety specification* is represented by a pair of vectors $\bar{b}, \bar{b} \in [0,1]^n$, satisfying $\sum_i b_i < 1 < \sum_i \bar{b}_i$. We let

$$\Pi(\bar{b}, \bar{b}) := \{\pi \in \Pi \mid b_i \leq \pi_i \leq \bar{b}_i, \forall i \in S\}$$

denote the set of state probability distributions which meet the safety specification, and refer to the elements of $\Pi(\bar{b}, \bar{b})$ as *safe*. On the other hand, a pair $(P, \pi)$ is said to be *safe* if the resulting state probability distribution meets the safety specification at all time steps, i.e., if

$$\pi P^k \in \Pi(\bar{b}, \bar{b}), \quad \text{for all } k \geq 0. \quad (1)$$

Also, a set $\Pi_0 \subseteq \Pi$, is *safe relative to a policy* $\varphi \in \Sigma_S$, if (1) holds with $P = P_{\varphi}$, for all $\pi \in \Pi_0$.

Remark 3.1: Since $\sum_i \pi_i = 1$, for each $\pi \in \Pi$, $\Pi(\bar{b}, \bar{b})$ is empty or is a singleton, unless $\sum_i b_i < 1$ and $\sum_i \bar{b}_i > 1$.

Example 3.1: Consider a single machine which operates in either of its two states, namely, “up”, called state 1, and “down”, called state 2. Suppose the probability that the machine maintains its current state at the next step is given by $p$ (resp., $q$) if the current state is up (resp., down). Then the state space of the machine is given by $S = \{1, 2\}$, and the state transition matrix by

$$P = \begin{bmatrix} p & 1 - p \\ 1 - q & q \end{bmatrix}.$$  

The entries of the state transition matrix can be controlled at any given state by adjusting the rate of usage, and the rate of maintenance. While in the up state, $p$ is an increasing function of the rate of maintenance, and a decreasing function of the rate of usage. While in the down state, $q$ is a decreasing function of the rate of maintenance, and it does not depend on the intensity of usage (since the machine is not used in its down state).

Suppose it is desired that at each step the machine is never down with probability more than 25%, but that it is down with probability at least 10% (for maintenance). Then, according to the previous discussion, the safety specification is represented by $\bar{b} = [1 \frac{1}{4}]$ and $\bar{b} = [0 \frac{1}{10}]$. Later in this section, we derive conditions on $p$ and $q$ for the machine to meet the safety control objective.
Example 3.2: An engineer is in charge of the maintenance scheduling task for an Automated Manufacturing System (AMS) composed of several subsystems such as assembly stations, robots, and computer control systems. To simplify the maintenance work of this complex system, the engineer categorizes all its components into three types: $E$ (electrical), $M$ (mechanical), and $L$ (lubricant), and contracts with three companies for the associated maintenance work. The operation of the AMS is then classified into four states: $E$, $M$, $L$, and $G$. If state $E$ ($M$, or $L$) is reached, the AMS’s electrical (mechanical, or lubricant) components need maintenance and the associated company is called. If the system is in $G$ (good) state, no maintenance work is needed. Let $S = \{1, 2, 3, 4\}$, where 1, 2, 3, and 4 represent the states $E$, $M$, $L$, and $G$, respectively. Suppose that empirical data shows that the probability transition matrix of the AMS’s operation can be expressed as

$$P = \begin{bmatrix}
    a & \frac{2}{3}(1-a-e) & \frac{1}{3}(1-a-e) & e \\
    \frac{2}{3}(1-b-f) & b & \frac{1}{3}(1-b-f) & f \\
    \frac{1}{3}(1-c-g) & \frac{1}{3}(1-c-g) & c & g \\
    \frac{1}{3}(1-d) & \frac{1}{3}(1-d) & \frac{1}{3}(1-d) & d
\end{bmatrix},$$

and that each contracted company can perform either basic or advanced maintenance defined by the maintenance cost limit. Suppose that these costs are 100 (300), 500 (1000), and 50 (200) units of dollars for basic (advanced) electrical, mechanical, and lubricant maintenance, respectively. If the transition probabilities $a$, $b$, $c$, $d$, $e$, $f$, and $g$, for basic (advanced) maintenance are $0.2 (0.05), 0.2 (0.05), 0.2 (1.0), 0.7 (0.7), 0.5 (0.9), 0.5 (0.9)$, and $0.4 (0.9)$ respectively, then the optimal policy to minimize the long-run average maintenance cost is (see e.g. [11] for details) to perform the basic maintenance when the AMS is in state $E$ or $M$, and to perform the advanced maintenance when the AMS is in state $L$. As a result, the transition probability matrix corresponding to this optimal policy becomes

$$P = \begin{bmatrix}
    0.2 & 0.2 & 0.1 & 0.5 \\
    0.2 & 0.2 & 0.1 & 0.5 \\
    0 & 0 & 0.1 & 0.9 \\
    0 & 0 & 0.1 & 0.7
\end{bmatrix},$$

and has an invariant probability distribution

$$\pi^* = [0.1125, 0.1125, 1.0000, 0.6750].$$

Suppose that a margin of $\pm 20\%$ of the stationary probability distribution is decided upon as the safety range, namely,

$$b = \pi^* \times (1 + 20\%) = [0.135, 0.135, 0.12, 0.81]$$

$$\bar{b} = \pi^* \times (1 - 20\%) = [0.09, 0.09, 0.08, 0.54].$$

Later in this section, we use Theorem 3.1 below to verify that the optimal controller in (2) enforces safety, thus reducing the maintenance cost.

In what follows, we obtain necessary and sufficient conditions on $P_\varphi$ so that the state probability distribution of the controlled Markov chain under a stationary policy $\varphi \in \Sigma_\varphi$ meets the safety specification at all time steps, for any initial safe probability distribution, i.e.,

$$\pi \in \Pi(b, \bar{b}), \text{ then } \pi P_\varphi^k \in \Pi(b, \bar{b}), \text{ for all } k \geq 0.$$  

This is equivalent to the condition

$$\left[ \pi \in \Pi(b, \bar{b}) \right] \implies \left[ \pi P_\varphi^k \in \Pi(b, \bar{b}) \right],$$

i.e., $\Pi(b, \bar{b})$ is invariant under $P_\varphi$.

Let $\rho^{(j)} = [\rho_1^{(j)}, \rho_2^{(j)}, \ldots, \rho_n^{(j)}]^T$ denote the $j$th column of $P_\varphi$. Let $\sigma_j$ be a permutation of $\{1, \ldots, n\}$ that arranges the entries of $\rho^{(j)}$ in decreasing order, i.e.,

$$\rho_{\sigma_j(1)}^{(j)} \geq \rho_{\sigma_j(2)}^{(j)} \geq \cdots \geq \rho_{\sigma_j(n)}^{(j)}, \quad \forall j \in S.$$  

For each $j \in S$, define $\underline{\ell}_j$ and $\overline{\ell}_j$ to be the smallest integers in $\{1, \ldots, n\}$ such that

$$\sum_{i=1}^{\overline{\ell}_j} b_{\sigma_j(i)} + \sum_{i=\overline{\ell}_j+1}^{n} b_{\sigma_j(i)} \geq 1 \quad (3a)$$

$$\sum_{i=1}^{\overline{\ell}_j} b_{\sigma_j(i)} + \sum_{i=\overline{\ell}_j+1}^{n} b_{\sigma_j(i)} \leq 1 \quad (3b)$$

By Definition 3.1, (3a) [(3b)] holds if $\overline{\ell}_j = n$ [$\overline{\ell}_j = n$], and fails if $\overline{\ell}_j = 0$ [$\overline{\ell}_j = 0$]. Thus, smallest such $\overline{\ell}_j$ and $\underline{\ell}_j$ do exist, and moreover, by (3a) and (3b), for each $j \in S$, we have:

$$\sum_{i=1}^{\overline{\ell}_j} b_{\sigma_j(i)} + \sum_{i=\overline{\ell}_j+1}^{n} b_{\sigma_j(i)} < 1 \leq \sum_{i=1}^{\overline{\ell}_j} b_{\sigma_j(i)} + \sum_{i=\overline{\ell}_j+1}^{n} b_{\sigma_j(i)}$$

$$\sum_{i=1}^{\overline{\ell}_j} b_{\sigma_j(i)} + \sum_{i=\overline{\ell}_j+1}^{n} b_{\sigma_j(i)} > 1 \geq \sum_{i=1}^{\overline{\ell}_j} b_{\sigma_j(i)} + \sum_{i=\overline{\ell}_j+1}^{n} b_{\sigma_j(i)}.$$  

Therefore, if we define

$$\kappa_j := 1 - \sum_{i=1}^{\overline{\ell}_j} b_{\sigma_j(i)} - \sum_{i=\overline{\ell}_j+1}^{n} b_{\sigma_j(i)}$$

$$\overline{\kappa}_j := 1 - \sum_{i=1}^{\overline{\ell}_j} b_{\sigma_j(i)} - \sum_{i=\overline{\ell}_j+1}^{n} b_{\sigma_j(i)},$$

then (3a) and (3b) can be expressed as:

$$\underline{\kappa}_j \leq \kappa_j \leq \overline{\kappa}_j,$$  

$$\underline{\kappa}_j \leq \overline{\kappa}_j < \overline{\ell}_j < b_{\sigma_j(\overline{\ell}_j)}.$$  

$$\underline{\kappa}_j \leq \overline{\kappa}_j < b_{\sigma_j(\overline{\ell}_j)}.$$  

$$\underline{\kappa}_j \leq \overline{\kappa}_j < b_{\sigma_j(\overline{\ell}_j)}.$$
The following theorem provides a necessary and sufficient criterion for a given feedback controller to be safety enforcing, when starting from an arbitrary safe initial probability distribution.

**Theorem 3.1:** It holds that

\[ \pi P_\varphi \in \Pi (\overline{b}, \overline{b}) \, , \, \forall \pi \in \Pi (\overline{b}, \overline{b}) \, , \]

if and only if, for all \( j \in S \),

\[ \pi_j P_\varphi (j) = \sum_{i=1}^{\overline{T}_j-1} b_{\sigma_j(i)} P_\varphi (i) + \sum_{i=\overline{T}_j+1}^{\overline{t}_j-1} b_{\sigma_j(i)} P_\varphi (i) \leq \overline{b}_j \, (6a) \]

\[ \pi_j P_\varphi (j) = \sum_{i=1}^{\overline{T}_j-1} b_{\sigma_j(i)} P_\varphi (i) + \sum_{i=\overline{T}_j+1}^{\overline{t}_j-1} b_{\sigma_j(i)} P_\varphi (i) \geq \overline{b}_j \, (6b) \]

**Proof:** To show necessity, for each \( j \in S \), define \( \pi_j \) and \( \overline{\pi}_j \) by

\[ \pi_j (j) := \begin{cases} b_k & \text{if } k = \sigma_j(i), \ 1 \leq i < \overline{T}_j \\ \pi_j & \text{if } k = \overline{T}_j \\ \overline{\pi}_j & \text{otherwise} \end{cases} \, (7) \]

and

\[ \overline{\pi}_j (j) := \begin{cases} b_k & \text{if } k = \sigma_j(i), \ 1 \leq i < \overline{T}_j \\ \overline{\pi}_j & \text{if } k = \overline{T}_j \\ \overline{\pi}_j & \text{otherwise} \end{cases} \, (8) \]

By (4), (7), and (8), we obtain \( \sum_k \pi_j (j) = \sum_k \overline{\pi}_j (j) = 1 \), i.e., \( \pi_j, \overline{\pi}_j \in \Pi \), for all \( j \in S \). Furthermore, it follows by (5) that \( \pi_j, \overline{\pi}_j \in \Pi (\overline{b}, \overline{b}) \), for all \( j \in S \).

Since \( \pi_j, \overline{\pi}_j \in \Pi (\overline{b}, \overline{b}) \), for all \( j \in S \), it follows from the hypothesis of the theorem that

\[ \{ \pi_j P_\varphi, \overline{\pi}_j P_\varphi \} \subset \Pi (\overline{b}, \overline{b}) \, , \, \forall j \in S \, , \]

or equivalently,

\[ \overline{b}_j \leq \pi_j P_\varphi (j) \leq \overline{b}_j \, , \, \overline{b}_j \geq \overline{\pi}_j P_\varphi (j) \geq \overline{b}_j \, , \, \forall j \in S \, . \]

Observing that \( \pi_j P_\varphi (j) \leq \overline{b}_j \) is equivalent to (6a) and that \( \pi_j P_\varphi (j) \geq \overline{b}_j \) is equivalent to (6b), necessity follows.

To show sufficiency, suppose (6a) holds. Let \( \pi \in \Pi (\overline{b}, \overline{b}) \), and fix an arbitrary \( j \in S \). We have,

\[ \pi_j P_\varphi (j) = \sum_{i=1}^{n} \pi_{\sigma_j(i)} P_\varphi (i) \]

\[ \leq \overline{b}_j \, (6b) \]

where the first inequality follows from the assumption that \( \pi \in \Pi (\overline{b}, \overline{b}) \), and from the definition of \( \sigma_j \) which implies that \( \rho_{\sigma_j(i)} - \rho_{\sigma_j(i)} \geq 0 \), for \( i \leq \overline{T}_j - 1 \), and \( \rho_{\sigma_j(i)} - \rho_{\sigma_j(i)} \leq 0 \), for \( i \geq \overline{T}_j + 1 \), and the second inequality follows from the hypothesis. Similarly, using (6b) we obtain \( \pi_j P_\varphi (j) \geq \overline{b}_j \).

**Remark 3.2:** It follows from Theorem 3.1 that the problem of verifying whether a given state feedback controller can enforce a given safety specification for an arbitrary initial safe state is polynomially decidable, and requires the verification of \( 2n \) inequalities given by (6).

Note that it is also possible to solve this problem using a linear programming formulation. However, we prefer the “explicit” solution over the “implicit” one based on linear-programming for two reasons. First, Theorem 3.1 shows that the problem has such a structure that it admits a solution based on solving a set of inequalities, which provides an explicit bound for the complexity. The second reason is that this approach provides an insight as to why a given safe set is invariant under the system transition matrix.

**Remark 3.3:** Theorem 3.1 can be used to characterize the class of all safety enforcing controllers, i.e., the set of all \( P_\varphi \) such that \( \Pi (\overline{b}, \overline{b}) P_\varphi \subset \Pi (\overline{b}, \overline{b}) \). Thus, given a safety specification, we can determine whether a safety enforcing controller exists, by searching for a corresponding state transition matrix \( P_\varphi \) that satisfies (6). We explore this further in Example 3.3.

**Remark 3.4:** Note that in light of the definition of \( \pi_j \) and \( \overline{\pi}_j \) given in (7)–(8), it follows that (6) simplifies to

\[ \pi_j P_\varphi (j) \leq \overline{b}_j \, \text{and} \, \overline{\pi}_j P_\varphi (j) \geq \overline{b}_j \, , \, \forall j \in S \, . \, (9) \]

**Remark 5.5:** By Theorem 3.1, if the state transition matrix, \( P_\varphi \), with \( \varphi \in \Sigma_S \) satisfies (6), then \( \Pi (\overline{b}, \overline{b}) \) is invariant under \( P_\varphi \). Since \( \Pi (\overline{b}, \overline{b}) \) is also closed and convex, then \( P_\varphi \) possesses a safe invariant probability distribution. This follows by Theorem 4.1 below, which shows that the existence of a safe invariant probability distribution of \( P_\varphi \) is both necessary and sufficient for the existence of a nonempty set of safe probability distributions, which is invariant under \( P_\varphi \).

**Example 3.3:** We continue with the analysis of the single machine model introduced in Example 3.1. Since \( \overline{b} = [1 \ \frac{3}{4}] \) and \( \overline{b} = [0 \ \frac{1}{4}] \), we need to concentrate only on the second component of the state probability distribution. If \( p \geq 1 - q \), then \( \sigma_2 = (2, 1) \), and we obtain \( \overline{T}_2 = \overline{t}_2 = 2 \). It follows from (7) that

\[ \overline{\pi} (2) = \left[ \begin{array}{c} \frac{3}{4} \\ \frac{1}{4} \end{array} \right] \, ; \, \pi (2) = \left[ \begin{array}{c} \frac{9}{10} \\ \frac{1}{10} \end{array} \right] \, . \]
Thus, (9) yields

\[ \frac{3}{4} (1-p) + \frac{1}{4} q \leq \frac{1}{4} ; \quad \frac{3}{10} (1-p) + \frac{1}{10} q \geq \frac{1}{10}, \]

or equivalently, \([3p - q \geq 2] \land [9p - q \leq 8]\).

If \( p \leq 1 - q, \) then \( \sigma_2 = (1, 2), \) \( t_2 = t_2 = 1, \) and (7) yields

\[ \pi^{(2)} = \left[ \frac{9}{10} \right] ; \quad \pi^{(2)} = \left[ \frac{3}{4} \right]. \]

From (9), we obtain,

\[ \frac{9}{10} (1-p) + \frac{1}{10} q \leq \frac{1}{4} ; \quad \frac{3}{4} (1-p) + \frac{1}{4} q \geq \frac{1}{4}, \]

or equivalently, \([18p - 2q \geq 13] \land [15p - 5q \leq 13].\)

Thus, a state feedback safety enforcing controller with state transition matrix \( P_\varnothing = \left[ \frac{1}{1-p}, \frac{p}{q} \right] \) must satisfy:

\[ \left( p \geq 1 - q \right) \land \left( 3p - q \geq 2 \right) \land \left( 9p - q \leq 8 \right) \]
\[ \lor \left( p \leq 1 - q \right) \land \left( 18p - 2q \geq 13 \right) \land \left( 15p - 5q \leq 13 \right). \]

**Example 3.4:** We continue the analysis of Example 3.2. We use Theorem 3.1 and Remark 3.4 to check if the optimal policy meets the safety objective. We obtain \( \sigma_1 = (1, 2, 3), \) for \( i = 1, 2, 3, \) and \( \sigma_2 = (3, 4, 2, 1). \) Consequently,

\[ \pi^{(1)} = \begin{bmatrix} 0.09 & 0.09 & 0.12 & 0.7 \end{bmatrix} \rho^{(1)} = 0.106 \geq b_1, \]
\[ \pi^{(2)} = \begin{bmatrix} 0.09 & 0.09 & 0.12 & 0.7 \end{bmatrix} \rho^{(2)} = 0.106 \geq b_2, \]
\[ \pi^{(3)} = \begin{bmatrix} 0.09 & 0.09 & 0.12 & 0.7 \end{bmatrix} \rho^{(3)} = 0.1 \geq b_3, \]
\[ \pi^{(4)} = \begin{bmatrix} 0.135 & 0.135 & 0.08 & 0.65 \end{bmatrix} \rho^{(4)} = 0.662 \leq b_4. \]

and

\[ \pi^{(1)} = \begin{bmatrix} 0.135 & 0.135 & 0.08 & 0.65 \end{bmatrix} \rho^{(1)} = 0.119 \leq b_1, \]
\[ \pi^{(2)} = \begin{bmatrix} 0.135 & 0.135 & 0.08 & 0.65 \end{bmatrix} \rho^{(2)} = 0.119 \leq b_2, \]
\[ \pi^{(3)} = \begin{bmatrix} 0.135 & 0.135 & 0.08 & 0.65 \end{bmatrix} \rho^{(3)} = 0.1 \leq b_3, \]
\[ \pi^{(4)} = \begin{bmatrix} 0.09 & 0.09 & 0.12 & 0.7 \end{bmatrix} \rho^{(4)} = 0.688 \leq b_4. \]

We conclude that \( \Pi(b, \bar{b}) \) is an invariant set of probability distributions under the optimal policy.

**IV. Maximal Invariant Safe Set**

When the condition of Theorem 3.1 fails, characterizing the set of initial probability distributions that are safe relative to a policy \( \varnothing \in \Sigma_S, \) amounts to computing the MISS of \( P_\varnothing. \) As we establish in this section, the computation can be performed iteratively and terminates in a finite number of steps.

Given \( P_\varnothing, \) with \( \varnothing \in \Sigma_S, \) we denote the class of invariant sets of safe states by

\[ \mathcal{J}(b, \bar{b}) := \{ \Pi \subseteq \Pi(b, \bar{b}) \mid \pi P_\varnothing \in \Pi, \forall \pi \in \Pi \} \]
\[ = \{ \Pi \subseteq \Pi(b, \bar{b}) \mid \pi_0 = \Pi \Rightarrow \pi_0 P_\varnothing^k \in \Pi, \forall k \}. \]

It is obvious that \( \mathcal{J}(b, \bar{b}) \) is closed under intersections, and its unique minimal element is the empty set. Similarly, \( \mathcal{J}(b, \bar{b}) \) is closed under unions and hence possesses a unique maximal element, denoted by \( \Pi_\varnothing. \) The following theorem provides a criterion for verifying that \( \Pi_\varnothing \) is nonempty.

**Theorem 4.1:** Given a stationary policy \( \varnothing \in \Sigma_S, \) let \( \Pi_\varnothing \subseteq \Pi(b, \bar{b}) \) be the MISS relative to \( P_\varnothing. \) Then \( \Pi_\varnothing \) is nonempty if and only if \( P_\varnothing \) has an invariant probability distribution that lies in \( \Pi(b, \bar{b}). \)

**Proof:** If \( \Pi(b, \bar{b}) \), then its topological closure as well as its convex hull are also elements of \( \mathcal{J}(b, \bar{b}) \) (this follows from the fact that \( \Pi(b, \bar{b}) \) itself is closed and convex). Therefore, \( \Pi_\varnothing \) must be closed and convex. Thus, if \( \Pi_\varnothing \neq \emptyset, \) then by Schauder’s fixed point theorem it must contain an invariant probability distribution \( \pi^* \) of \( P_\varnothing, \) and necessity follows. To establish sufficiency, suppose \( \pi^* \in \Pi(b, \bar{b}) \) is an invariant probability distribution of \( P_\varnothing. \) Then, \( \{ \pi^* \} \subseteq \Pi_\varnothing, \) i.e., \( \pi^* \in \Pi_\varnothing, \) which proves that \( \Pi_\varnothing \) is not empty.

**A. Iterative Computation of \( \Pi_\varnothing \)**

Consider the following algorithm:

**Algorithm 4.2:** Let \( N_0 \subseteq \Pi(b, \bar{b}) \) and for \( k = 1, 2, \ldots, \) set

\[ N^{(0)} := N_0 \]
\[ N^{(k)} := \{ \pi \in \Pi(b, \bar{b}) \mid \pi P_\varnothing \in N^{(k-1)} \}. \]

Note that

\[ N^{(k)} = \{ \pi \mid \pi P_\varnothing \in N^{(k-1)} \}, \quad \pi P_\varnothing \in P_\varnothing, \quad 1 \leq \ell < k \}

The behavior of Algorithm 4.2 depends on the initial set \( N_0. \) Note that if \( N^{(0)} \supseteq N^{(1)} \supseteq N^{(2)} \), then it follows that \( N^{(k)} \supseteq N^{(k-1)} \supseteq N^{(k-2)} \), for all \( k \in \mathbb{N}. \) Also if \( N^{(k-1)} \) is closed and convex, so is \( N^k. \) Thus if \( N_0 = \Pi(b, \bar{b}) \), then \( \{ N^{(k)} \} \to \ell \) is a non-increasing sequence of closed, convex sets. It is evident in this case that \( \Pi_\varnothing \subseteq N^{(k)}, \) for all \( k \in \mathbb{N}, \) and that \( \bigcap_{k \in \mathbb{N}} N^{(k)} \) is invariant under \( P_\varnothing. \) Therefore,

\[ \bigcap_{k \in \mathbb{N}} N^{(k)} = \Pi_\varnothing. \]

On the other hand, if \( N_0 \not\supseteq \Pi(b, \bar{b}) \), then since by invariance \( N^1 \supseteq N^0, \) it follows that \( \{ N^{(k)} \} \to \ell \) is an non-decreasing sequence of elements of \( \mathcal{J}(b, \bar{b}). \) However, in this case, the algorithm is not guaranteed to converge to \( \Pi_\varnothing. \) Additional hypotheses are needed on the chain for this to happen.

For the remaining of the paper, we fix the topology to be the subspace topology relative to the Euclidean space \( \{ z \in \mathbb{R}^n \mid \sum_i z_i = 1 \}. \) This, of course, affects the definition of open subsets of \( \Pi. \) In this topology, \( \pi \in \Pi(b, \bar{b}) \) is an interior point of \( \Pi(b, \bar{b}) \) if and only if

\[ \left\{ \begin{array}{l} \pi_i < b_i, \quad \text{if } b_i \neq 1, \\
\pi_i > \bar{b}_i, \quad \text{if } b_i \neq 0, \end{array} \right. \forall i \in S. \]

Note that (11) is equivalent to the requirement that, for some \( \varepsilon > 0, \)

\[ b_i \leq (1 - \varepsilon) \pi_i \leq \bar{b}_i - \varepsilon, \quad \forall i \in S. \]

We use the characterization of interior points in (12) later in the paper, without elaborating further.

In order to show that Algorithm 4.2, terminates in finitely many iterations we need the following assumption.

**Assumption 4.1:** Let \( d \) be the period of \( P_\varnothing, \) and suppose that the stationary policy \( \varnothing \in \Sigma_S, \) the set of invariant probability distributions of \( P_\varnothing, \) denoted by \( \Pi^*, \) lie in the interior of \( \Pi(b, \bar{b}). \)
Assumption 4.1 implies that the state probability distribution will eventually lie inside the safe set. The issue remains as to what is the set of initial probability distributions so that under the evolution of $P_{\varphi}$ it is invariant. Hence, the finite termination of Theorem 4.2 remains inside the safe set. (This set is what we denote as $\Pi_{\varphi}$.) Clearly, by Theorem 4.1, Assumption 4.1 guarantees that $\Pi_{\varphi}$ is nonempty. The finite termination of Algorithm 4.2 is the topic of the following Theorem.

Theorem 4.3: Let Assumption 4.1 hold and suppose that $\Pi^*$ is contained in $N_0$, the interior of $N_0$. Then there exists a finite integer $k \in \mathbb{N}$ such that $N^{(k+1)} = N(k) = \Pi_{\varphi}$.

Proof: We argue by contradiction. If the iteration does not terminate in a finite number of steps, then there exists a sequence $\{x_n\} \subset \Pi(b, \overline{b})$ such that $x_{n+1} = f(x_n)$, for all $n \geq 0$. Let $x_n = (x_{n1}, \ldots, x_{nn})$ and $x_{n+1} = (x_{n+11}, \ldots, x_{n+1n})$.

Consider the set $V = \{x_n \in \mathbb{R}^n : x_n \neq 0\}$. If $V$ is compact, then $x_n$ converges to some point $x_\infty$. Assume $x_\infty = (0, \ldots, 0)$.

Let $x_{n+1} = f(x_n) = (f_{n1}, \ldots, f_{nn})$, for all $n \geq 0$. Let $x_{n+1} = (x_{n1}, \ldots, x_{nn})$ and $x_{n+1} = (x_{n+11}, \ldots, x_{n+1n})$.

Thus, (13) implies

$$\left\{ \pi(k)^{P_{\varphi}^k}, \pi(k)^{P_{\varphi}^{k+1}} \right\} \subset N_0, \quad \forall k \in \mathbb{N}. \quad (14)$$

Let $\pi$ be any limit point of $\{\pi(\ell)\}_{\ell=1}^{\infty}$. Since $P_{\varphi}$ is aperiodic, $P_{\varphi}^{\ell}$ tends to a limit $\pi$, as $\ell \to \infty$. Thus, $\pi P_{\varphi}^{\ell} \to \pi_0$, and $\pi P_{\varphi}^{\ell+1} \to \pi_1$, as $\ell \to \infty$, for some $\pi_0, \pi_1 \in \Pi$. However, since $\{\pi_0, \pi_1\} \subset \Pi^*$, it follows by the hypothesis of the theorem that $\{\pi_0, \pi_1\} \subset N_0$. Since $\pi P_{\varphi}^{\ell} \to \pi_0$, using a triangle inequality we deduce that there exist integers $\ell_0$ and $m_0$ such that $\{\pi_0 P_{\varphi}^{\ell_0}, \pi_1 P_{\varphi}^{\ell_0+1}\} \subset N_0$, for all $\ell \geq \ell_0$, and $m \geq m_0$. Choose $m' \geq m_0$ such that $\ell' = \ell_0 + m'$, and set $k' = k + m'.d$. It follows that $\pi P_{\varphi}^{k'} \in N_0$ and $\pi P_{\varphi}^{k+1} \in N_0$, which contradicts (14). Thus, the iteration terminates at some finite $k \in \mathbb{N}$. It is clear that $\Pi(k) \in \mathbb{N}$, and hence $\Pi(k) \subset \Pi_\varphi$.

To show that $\Pi(k) \supset \Pi_\varphi$, let $\pi \in \Pi_\varphi$ be arbitrary. Then, $\pi P_{\varphi}^{\ell} \to \pi_\varphi$, as $\ell \to \infty$, for some finite collection $\{\pi_\varphi, \rho = 0, \ldots, d-1\}$, which deduces that, for some $k_0 \in \mathbb{N}, \pi P_{\varphi}^{k_0} \in N_0$, for all $k \geq k_0$. Therefore, by (10), $\pi \in N(k)$, for all $k$ sufficiently large.

If we select $N_0 = \Pi(b, \overline{b})$, it follows from Theorem 4.3 that $\Pi_{\varphi}$ is a polyhedron. This is summarized in the following corollary.

Corollary 4.4: Let Assumption 4.1 hold, and let $k$ be as in Theorem 4.3. Then $\Pi_{\varphi}$ is a polyhedron consisting of the probability distributions $\pi \in \Pi$ given by

$$b \leq \pi P_{\varphi}^k \leq \overline{b}, \quad \forall k \leq k.$$

The key reason behind the finite termination of the algorithm in Theorem 4.3 is the uniform asymptotic stability of the set $\Pi^*$ under the iterates of the map $P_{\varphi}$. In fact this result holds in a rather more generic setting as the following Theorem shows.

Theorem 4.5: Let $f: \mathbb{R}^n \to \mathbb{R}^n$ be a continuous map, and $K \subset \mathbb{R}^n$ a closed, bounded set. Suppose that some compact set $G \subset K^\circ$ is asymptotically stable under the action of $f$ on $\mathbb{R}^n$, and that its region of attraction contains the set $K$. With $W \subset K$ an open neighborhood of $G$, let $N(0) := W$ and for $k = 1, 2, \ldots$, let $N(k) := f^{-1}(N(k-1)) \cap K$. Then, there exists an integer $\hat{k}$ such that $N(\hat{k}) = N(\hat{k}+1)$, and $N(\hat{k})$ is the maximal $f$-invariant subset of $K$.

Proof: Arguing by contradiction, if the iteration does not terminate in finitely many steps, then as in Theorem 4.3, we deduce that there exists a sequence $\{x_n\} \subset K$, for which it holds $f(x_n), f^{k+1}(x_n) \in W$, for all $k \in \mathbb{N}$. Since $G$ is asymptotically stable, there exists an open set $V$, satisfying $G \subset V$, and such that

$$f(x) \in V, \quad \forall x \in V, \quad \forall k \in \mathbb{N}. \quad (15)$$

Let $x$ be a limit point of $\{x_n\}$ in $K$ (such a limit point exists by the compactness of $K$). Also, let $\{x_n\} \subset N(k)$ be a subsequence over which $x_n \to x$. By the asymptotic stability of $G$, there exists $\ell' \in \mathbb{N}$, such that $x_n \in V$, Hence, by the continuity of $f$, $f^{\ell'}(x_n) \in V$, for all $k_n$ sufficiently large. Pick such a $k_n > \ell'$, and use (15) to conclude that $f^{\ell'}(x_n) = f^{k_n+\ell'}(f^{k_n}(x_n)) \in W$.

Similarly $f^{k_n+1}(x_n) \in W$, thus arriving at a contradiction. Hence the iteration must terminate at some $k \in \mathbb{N}$. It is clear that $N(k) \subset K$, and that $N(k)$ is invariant under the action of $f$. To show that $N(k)$ is maximal, suppose that for some $x \in K$, we have $f(x) \in K$, for all $k \in \mathbb{N}$. By the asymptotic stability of $G$, $f(x) \in V$, for some $k'$, and using (15), we obtain $f(x) \in W$, for all $k \geq k'$. Thus $x \in N(k)$, for all $k \geq k'$, which implies $x \in N(k)$, thus proving that $N(k)$ is maximal.

B. Upper Bound on the Iteration Steps

In what follows, we compute an upper bound on the number of steps required for Algorithm 4.2 to terminate under Assumption 4.1. The upper bound on the iteration is useful in numerical implementations of the algorithm. Since a numerical computation uses a rational approximation of a real number, it may never reach a fixed point even when a finite termination is assured. In such a case, the upper bound on the required number of steps can be used to terminate the numerical computation.

In order to keep the algebra simple, we assume that $P_{\varphi}$ has a unique communicative class of aperiodic recurrent states. Then, under Assumption 4.1, the unique invariant probability distribution $\pi^*$ of $P_{\varphi}$ lies in the interior of $\Pi(b, \overline{b})$.

Consider the set

$$\Pi_\varepsilon := (1 - \varepsilon)\pi^* + \varepsilon\Pi.$$

where $\pi^*$ is the invariant probability distribution of $P_{\varphi}$. We show in the following lemma that if $\varepsilon$ is chosen appropriately, then $\Pi_\varepsilon \in \mathbb{N}(b, \overline{b})$.

Lemma 4.6: Let Assumption 4.1 hold. Suppose $P_{\varphi}$ has a unique communicative class of aperiodic recurrent states and denote its unique invariant probability distribution by $\pi^*$. Let $\varepsilon_0$ be a number satisfying:

$$\varepsilon_0(1 - \pi^*_b) \leq \pi_i - \pi^*_i, \quad \text{and} \quad \varepsilon_0\pi^*_i \leq \pi_i - b_i, \quad \forall i \in S,$$

or equivalently,

$$\varepsilon_0 \leq \min_{i \in S} \left\{ \frac{\pi_i - b_i}{\pi^*_b - \pi_i}, \frac{\pi_i - b_i}{\pi^*_i - \pi_b} \right\}.$$
Then,
$$P_{\varphi} \in \mathcal{J}(b, \overline{b}), \quad \forall \varepsilon \in [0, \varepsilon_0].$$

**Proof:** Since $\pi^* \in \Pi(b, \overline{b})$, we have
$$0 \leq \overline{b}_i \leq \pi^*_i \leq \overline{b}_i \leq 1.$$ 

Hence, it follows that
$$0 \leq \frac{\overline{b}_i - \pi^*_i - k}{\pi^*_i - \overline{b}_i} \leq 1,$$
which implies $0 \leq \varepsilon_0 \leq 1$, i.e., $\varepsilon_0$ is well defined.

Next we establish that $P_{\varphi} \in \Pi(b, \overline{b})$. Since $0 \leq \varepsilon \leq \varepsilon_0$, it follows from the definition that,
$$\varepsilon(1 - \pi^*_i) \leq \overline{b}_i - \pi^*_i, \quad \forall i \in S.$$ 

This is equivalent to
$$(1 - \varepsilon)\pi^*_i + \varepsilon \leq \overline{b}_i, \quad \forall i \in S,$$
and since $\pi^*_i \leq 1$, it implies
$$(1 - \varepsilon)\pi^*_i + \varepsilon \pi \leq \overline{b}_i, \quad \forall \pi \in \Pi.$$ 

(16)

On the other hand, we have $\varepsilon \pi^* \leq \pi^* - \overline{b}$ which combined with (16) yields
$$\overline{b} \leq (1 - \varepsilon)\pi^* + \varepsilon \pi \leq \overline{b}, \quad \forall \pi \in \Pi,$$
or, equivalently, that
$$P_{\varphi} = (1 - \varepsilon)\pi^* + \varepsilon \Pi \in \Pi(b, \overline{b}).$$

It remains to show that $P_{\varphi} \in \Pi_{\varepsilon} \subset \Pi_{\varepsilon}$. But since $\pi^* P_{\varphi} = \pi^*$ and $P_{\varphi} \subseteq \Pi$ then, for each $\pi \in \Pi$,
$$[(1 - \varepsilon)\pi^* + \varepsilon \Pi] P_{\varphi} = (1 - \varepsilon)\pi^* + \varepsilon \pi P_{\varphi} \in (1 - \varepsilon)\pi^* + \varepsilon \Pi.$$ 

This completes the proof.

Suppose Assumption 4.1 holds. Let $R$ denote the single communicating class of recurrent states of $P_{\varphi}$, and let $n_R$ denote its cardinality. Let $R$ be the restriction of $P_{\varphi}$ on $R$. The smallest positive integer $\gamma(R)$ such that $|R^{\gamma(R)}|_{ij} > 0$, for all $i, j \in R$, is upper bounded by $\gamma(R) \leq n_R^2 = 2n_R + 2$ [13, Theorem 2.9]. The integer $\gamma(R)$ is called the index of primitivity of $R$. We use $\gamma = \gamma(P_{\varphi})$ to denote the smallest positive integer such that $[P_{\varphi}^{\gamma}]_{ij} > 0$, for all $i \in S, j \in R$.

An upper bound on the index of primitivity, $\gamma(P_{\varphi})$, depending only on the dimension of $P_{\varphi}$ and the cardinality of the recurrent class $R$, is given in the following lemma.

**Lemma 4.7:** Let Assumption 4.1 hold, and suppose $P_{\varphi}$ has a unique communicative class $R$ of aperiodic recurrent states, of cardinality $n_R$. Then
$$\gamma(P_{\varphi}) \leq n + n_R^2 - 3n_R + 2.$$

**Proof:** Reordering the states in $S$, we write $P_{\varphi}$ in block-matrix form as
$$P_{\varphi} = \begin{pmatrix} R & 0 \\ C & T \end{pmatrix},$$
where $R$ is the restriction of $P_{\varphi}$ on $R$. Then
$$P_{\varphi}^k = \begin{pmatrix} R^k & 0 \\ Q(k) & T^k \end{pmatrix},$$
where $Q(k) = \sum_{\ell=1}^{k} T^{\ell - 1} C R^{k-\ell}$.

Let $q_i, i = 1, \ldots, n - n_R$, denote the row vector of dimension $n - n_R$ whose $i^{th}$ element is equal to 1 and the rest of the elements equal to 0. We claim that, for each $i$, $q_i T^{\ell - 1} C$ is not identically zero, for all $\ell = 1, \ldots, n - n_R$ if this is not the case, then, by the Caley-Hamilton Theorem, $T^{\ell - 1} C$ has a zero row, for all $\ell \in \mathbb{N}$. It follows that $Q(k)$ has a zero row, for all $k \in \mathbb{N}$, contradicting the fact that the rows of $Q(k)$ converge, as $k \to \infty$, to the invariant probability distribution of the recurrent states, which is positive.

Since, by definition, $\gamma(R)$ is the smallest positive integer such that all entries of $R^{\gamma(R)}$ are positive (we denote this by $R^{\gamma(R)} > 0$), we have $R^{\gamma(R)+k} > 0$, for all $k \in \mathbb{N}$ and $i, j \in \{1, \ldots, n - n_R\}$. Let $n = \gamma(R) + n - n_R$. Then, for all $i, j \in \{1, \ldots, n - n_R\}$,
$$Q_{ij}(n) = \sum_{\ell=1}^{n-n_R} T^{\ell - 1} C R^{\ell - 1} C_{ij} > 0.$$ 

It follows that $\gamma(P_{\varphi}) \leq n + n_R^2 - 3n_R + 2$. 

**Remark 4.1:** The upper bound on the index of primitivity $\gamma(P_{\varphi})$ in Lemma 4.7 is sharp, as the following example shows.

**Lemma 4.8:** Let the assumptions of Lemma 4.7 hold. With $\gamma = \gamma(P_{\varphi})$, define
$$\rho := \min_{i \in S, j \in R} \{ [P_{\varphi}^{\gamma}]_{ij} \}.$$ 

Let $\lambda_0$ be the smallest non-negative real number satisfying
$$(1 - \lambda_0) \pi^*_i \leq \rho, \quad \forall i \in R,$$
or, equivalently,
$$\lambda_0 = \max_{i \in \mathbb{R}} \left(1 - \frac{p_i}{\pi^*_i}\right).$$

Then $P_{\varphi}^\rho$ is contractive over $\Pi - \pi^*$, in the sense
$$\rho P_{\varphi}^\rho \subseteq \lambda_0 (\Pi - \pi^*),$$ 

(17)
or, equivalently,
$$\forall \pi \in \Pi : \pi P_{\varphi}^\rho \subseteq \lambda_0 (\Pi - \pi^*) + \pi^*.$$ 

**Proof:** First note that since $\rho > 0$ and $\pi^*_i \geq 0$, for each $i \in \mathbb{R}$, then
$$\lambda_0 = \max_{i \in \mathbb{R}} \left(1 - \frac{p_i}{\pi^*_i}\right) < 1.$$ 

Next, it is enough to show that for $\pi \in \Pi$,
$$\pi P_{\varphi} - (1 - \lambda_0) \pi^* \subseteq \lambda_0 \Pi.$$ 

(18)
Since \( \pi_i^* = 0 \) for \( i \notin \mathcal{R} \), it follows that for \( i \notin \mathcal{R} \), \((\pi P^*_\varphi)_i \geq 0 = (1 - \lambda_0)\pi_i^* \). On the other hand, for \( i \in \mathcal{R} \),
\[
(\pi P^*_\varphi)_i \geq (\pi \rho^{n \times n})_i = \rho \geq (1 - \lambda_0)\pi_i^*,
\]
where \( \rho^{n \times n} \) is a matrix of size \( n \times n \) with each entry having the value \( \rho \). Thus we obtain,
\[
\pi P^*_\varphi - (1 - \lambda_0)\pi^* \geq 0 \quad \text{for all } i.
\]
Next, note that
\[
\sum_{i=1}^{n} ((\pi P^*_\varphi)_i - (1 - \lambda_0)\pi_i^*) = \sum_{i=1}^{n} (\pi P^*_\varphi)_i - (1 - \lambda_0) \sum_{i=1}^{n} \pi_i^* = 0 = (1 - \lambda_0) = \lambda_0. \quad \text{(19)}
\]
Thus, (18) follows from (19) and (20).

The following corollary follows by applying (17) repeatedly.

**Corollary 4.9:** Let the assumptions of Lemma 4.7 hold, and let \( \varepsilon_0, \gamma \) and \( \lambda_0 \) be as in Lemmas 4.7–4.8. Then,
\[
(\Pi - \pi^*) P^*_\varphi \subset \lambda_0^k (\Pi - \pi^*), \quad \forall k \geq 0.
\]
or equivalently,
\[
\pi P^*_\varphi^k \in \lambda_0^k (\Pi - \pi^*) + \pi^* \quad \forall \pi \in \Pi.
\]
A theorem providing an upper bound on the iterations of Algorithm 4.2 follows.

**Theorem 4.10:** Suppose the assumptions of Lemma 4.7 hold. Let \( \varepsilon_0, \gamma \) and \( \lambda_0 \) be as in Lemmas 4.7–4.8, and let \( k_0 \) be a positive integer satisfying
\[
\lambda_0^k 0 \leq \varepsilon_0. \quad \text{(21)}
\]
Consider Algorithm 4.2 with \( N_0 \geq \Pi_{\varepsilon_0} \). Then, \( N^{(\gamma k_0 + 1)} = N^{(\gamma k_0)} = \Pi_{\varphi} \).

**Proof:** If the above iteration does not terminate in \( \gamma k_0 \) steps, then there exists \( \pi \in \Pi^{(\gamma k_0 + 1)} \Delta \Pi^{(\gamma k_0)} \). This means that \( \pi P^*_\varphi \) and \( \pi P^*_\varphi^{k+1} \) cannot both belong to \( N_0 \). Since
\[
\lambda_0^k (\Pi - \pi^*) = \varepsilon_0 \left( \frac{\lambda_0^k \Pi}{\varepsilon_0} \right) + \left( 1 - \frac{\lambda_0^k}{\varepsilon_0} \right) \pi^* - \pi^*
\]
\[
\subseteq \pi^* + \varepsilon_0 (\Pi - \pi^*),
\]
then, using Corollary 4.9 and (21), we obtain
\[
\pi P^*_\varphi^k \in \pi^* + \lambda_0^k (\Pi - \pi^*)
\]
\[
\subseteq \pi^* + \varepsilon_0 (\Pi - \pi^*) \subseteq N_0 \quad \forall k \geq k_0,
\]
leading to a contradiction. The assertion \( N^{(\gamma k_0)} = \Pi_{\varphi} \) follows from Theorem 4.3.

**Example 4.1:** To illustrate our results, we consider a 3-state Markov chain with state transition matrix
\[
P_{\varphi} = \begin{bmatrix} .65 & .25 & .10 \\ .15 & .65 & .20 \\ .15 & .50 & .35 \end{bmatrix}.
\]
Since \( p_{ij} > 0 \), for all \( i, j \in S = \{1, 2, 3\} \), it follows that \( P_{\varphi} \) is irreducible and aperiodic. Also, its unique invariant probability distribution is given by \( \pi^* = [.3 \quad .5 \quad .2] \). Consider a lower bound constraint \( b = [.1 \quad .2 \quad .1] \), and an upper bound constraint \( b = [.5 \quad .8 \quad .4] \). Then \( \pi^* \in \Pi(b, \bar{b}) \), and by Theorem 4.1, \( \Pi_{\varphi} \neq \emptyset \). We first check whether \( \Pi(b, \bar{b}) \in \mathcal{J}(b, \bar{b}) \) by verifying the conditions of (9).

Using the notation \( \tilde{b}_{\varphi_j} := [\tilde{b}_{\varphi_j(1)} \tilde{b}_{\varphi_j(2)} \tilde{b}_{\varphi_j(3)}]^T \), for \( j = 1, 2, 3 \), and similarly for \( \tilde{b}_{\varphi_j} \), we have
\[
\sigma_1 = (1, 2, 3), \quad \tilde{b}_{\varphi_1} = [.5 \quad .8 \quad .4], \quad \tilde{b}_{\varphi_1} = [.1 \quad .2 \quad .1]
\]
\[
\sigma_2 = (2, 3, 1), \quad \tilde{b}_{\varphi_2} = [.8 \quad .4 \quad .5], \quad \tilde{b}_{\varphi_2} = [.2 \quad .1 \quad .1]
\]
\[
\sigma_3 = (3, 2, 1), \quad \tilde{b}_{\varphi_3} = [.4 \quad .8 \quad .5], \quad \tilde{b}_{\varphi_3} = [.1 \quad .2 \quad .1].
\]
Continuing the calculation, we obtain
\[
\tilde{b}_1 = 2, \quad \tilde{b}_1 = 2, \quad \pi^{(1)} = [.5 \quad .4 \quad .1], \quad \pi^{(1)} = [.1 \quad .5 \quad .4]
\]
\[
\tilde{b}_2 = 1, \quad \tilde{b}_2 = 2, \quad \pi^{(2)} = [.8 \quad .1 \quad .4], \quad \pi^{(2)} = [.2 \quad .3 \quad .5]
\]
\[
\tilde{b}_3 = 2, \quad \tilde{b}_3 = 3, \quad \pi^{(3)} = [.4 \quad .5 \quad .1], \quad \pi^{(3)} = [.1 \quad .4 \quad .5].
\]
Then,
\[
\pi^{(1)} = .325 + .06 + .015 = .4 \leq \tilde{b}_1 = .5
\]
\[
\pi^{(2)} = .065 + .075 + .06 = .2 \geq \tilde{b}_2 = .8
\]
\[
\pi^{(3)} = .13 + .15 + .125 = .305 \leq \tilde{b}_3 = .2
\]
Thus it follows from Theorem 3.1 that \( \Pi_{\varphi} = \Pi(b, \bar{b}) \).

Now consider another pair of bounds, \( b = [2 \quad .3 \quad .1] \) and \( \bar{b} = [.5 \quad .5 \quad .4] \). We obtain
\[
\sigma_1 = (1, 2, 3), \quad \tilde{b}_{\varphi_1} = [.5 \quad .5 \quad .4], \quad \tilde{b}_{\varphi_1} = [.2 \quad .3 \quad .1]
\]
\[
\tilde{b}_1 = 2, \quad \tilde{b}_1 = 2, \quad \pi^{(1)} = [.5 \quad .4 \quad .1], \quad \pi^{(1)} = [.2 \quad .4 \quad .1].
\]

Thus,
\[
\pi^{(1)} = .325 + .06 + .015 = .4 \leq \tilde{b}_1 = .5
\]
\[
\pi^{(2)} = .13 + .06 + .06 = .25 \geq \tilde{b}_2 = .2
\]
It follows that \( \Pi_{\varphi} \neq \Pi(b, \bar{b}) \).

We next compute the set \( \Pi_{\varphi} \) by using a “top-down” method that starts with \( N^{(0)} = \Pi(b, \bar{b}) \) and iterates “downward” to converge to \( \Pi_{\varphi} \). Since the invariant probability distribution \( \pi^* = [.3 \quad .5 \quad .2] \) lies on the boundary of the safe set (note that \( \pi^* = \bar{b}_2 = .5 \)), Assumption 4.1 is not satisfied and so a finite termination is not a priori assured. Under a certain quantization, the numerical implementation of the iteration terminates in eight steps, as shown in Figure 1, where the first diagram is that of \( N^{(0)} = \Pi(b, \bar{b}) \), and the last diagram is that of \( N^{(7)} = \Pi_{\varphi} \).

**Example 4.2:** The process of a machine’s operation that deteriorates due to parts aging and revives by the choice of performing maintenance is often modeled by a three-state (Good, Old, Down) controlled Markov chain. Table I lists the table of parameters with respect to this model where state Good, Old, and Down is represented by G, O, and D respectively. If
\[
\begin{bmatrix} c_{11} & c_{12} & c_{21} & c_{22} & c_{31} & c_{32} \end{bmatrix} = [0 \quad 30 \quad 5 \quad 40 \quad 120 \quad 80]
\]
and
\[
\begin{bmatrix} a_{11} & a_{12} & a_{21} & a_{22} & r \end{bmatrix} = [.8 \quad .1 \quad .5 \quad .2 \quad .1],
\]
Thus, we set $\pi^*$ to be in $\Pi(b, \bar{b})$, and thus by Theorem 4.1, $\Pi_{\varphi}$ $\neq$ $\emptyset$. Also, we observe that $\Pi(b, \bar{b}) P \not\subseteq \Pi(b, \bar{b})$ (e.g., $[1, 0, 0] \in \Pi(b, \bar{b})$ but $[1, 0, 0] P \notin \Pi(b, \bar{b})$), hence the controller induced by the optimal policy does not meet the safety objective. The MISS can be calculated by Algorithm 4.2 setting $N^{(0)}$ $= \Pi(b, \bar{b})$. The calculation shows that the iteration terminates in eight steps. Namely, if we let $\rho^{(j)}$ be the $j$th column of $P^j$, then we can express

$$\Pi_{\varphi} = \{ \pi \in \Pi(b, \bar{b}) \mid \forall i = 1, \ldots, 7, b_j \leq \pi \rho^{(j)},$$

for $j = 1$ and $\pi \rho^{(j)} \leq \bar{b}_j$, for $j = 2, 3 \}$.

This result can be verified by the observation that in the 8th iteration

$$P^8 = \begin{bmatrix}
.4584 & .4562 & .0854 \\
.3925 & .5056 & .1019 \\
.4255 & .4809 & .0936
\end{bmatrix}.$$ 

The fact that the $(i, j)^{th}$ element of $P^k$ is greater than $b_i$ and less than $\bar{b}_j$ for $i, j \in \{1, 2, 3\}$ implies that $\Pi P^k \subset \bar{I}(b, \bar{b})$ for $k \geq 8$.

### V. Conclusion

The paper continues our earlier work on safety control of Markov chains [1]. A safety specification is given as a unit-interval valued vector that imposes lower and upper bounds on the state probability distribution. Such a specification may be used to impose a reliability constraint on the behavior of a stochastic discrete event system. Under the assumption of complete observations, we obtain explicit conditions on the state transition matrix corresponding to a stationary control policy, so that safety is enforced for any safe initial state. Next we determine the MISS of a given state feedback controller. This is done under the assumption that the set of invariant probability distributions of the chain is safe. The computation is iterative and terminates in a finite number of steps. As a result, the MISS is polyhedral. An explicit upper bound on the number of steps needed for the termination of the iterative computation is obtained for the case when the chain has a unique invariant probability distribution.

Next we identify some future research directions. Although the results of the paper enable seeking a safety enforcing controller by exhaustive search, a more algorithmic method should be developed for this purpose. Assumption 4.1 gives a sufficient condition for the finite termination of the iterative computation of the MISS. Further work is needed to relax this assumption, and develop ways to compute an invariant safe set (that is not necessarily maximal) when a finite termination of the iterative computation is not assured. Further, more general qualitative constraints should be formulated and studied. Also, the topic of optimal control of Markov chains subject to qualitative constraints such as safety is of interest. Finally, the theory needs to be extended to account for models with partial state observations.

### TABLE I

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$G$</td>
<td>Do nothing</td>
<td>$a_{11}$</td>
<td>$1-a_{11}$</td>
<td>$c_{11}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Do Maintain</td>
<td>$1-a_{12}$</td>
<td>$a_{12}$</td>
<td>$0$</td>
<td></td>
</tr>
<tr>
<td>$O$</td>
<td>Do nothing</td>
<td>$a_{21}$</td>
<td>$1-a_{21}$</td>
<td>$c_{21}$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Do Maintain</td>
<td>$1-a_{22}$</td>
<td>$a_{22}$</td>
<td>$0$</td>
<td></td>
</tr>
<tr>
<td>$D$</td>
<td>Do nothing</td>
<td>$0$</td>
<td>$0$</td>
<td>$1$</td>
<td>$c_{31}$</td>
</tr>
<tr>
<td></td>
<td>Do Repair</td>
<td>$1-6r$</td>
<td>$5r$</td>
<td>$r$</td>
<td>$c_{32}$</td>
</tr>
</tbody>
</table>

The optimal control policy that minimizes the long-run average cost is to perform maintenance when the system is in the $O$ or $D$ states (see [11] for details). When the system is in $G$ state, no maintenance should be done. The transition matrix under this policy is

$$P = \begin{bmatrix}
0 & .8 & .2 \\
.8 & .2 & 0 \\
.4 & .5 & .1
\end{bmatrix},$$

and has the stationary probability distribution

$$\pi^* = [.4235 \ .4824 \ .0941].$$

If the quality of the machine’s output is our top concern, the machine should be in the Good state as often as possible. Thus, we set $\bar{b}_1 = 1$ and $\bar{b}_2 = \bar{b}_3 = 0$. If we let $\bar{b}_1 = (1-10\%)\pi^*_1$, $\bar{b}_2 = (1+10\%)\pi^*_2$, and $\bar{b}_3 = (1+10\%)\pi^*_3$, then $\bar{b} = [.3811 \ 0 \ 0]$ and $\bar{b} = [.1 \ .5306 \ .1035]$. Clearly, $\pi^* \in \Pi(\bar{b}, \bar{b})$. The calculation shows that the iteration terminates in eight steps. Namely, if we let $\rho^{(j)}$ be the $j$th column of $P^j$, then we can express
REFERENCES