Safety of Stochastic Systems: An Analytic and Computational Approach

Rafael Wisniewski [a] , Manuela L. Bujorianu [b],

[a] *Section of Automation & Control, Aalborg University, 9220 Aalborg East, Denmark*

[b] *Maritime Safety Research Center, Department of Naval Architecture, Ocean Marine Engineering, University of Strathclyde, Scotland, UK*

Abstract

We refine the concept of stochastic reach avoidance for a general class of Markov processes introducing a threshold of p for the reaching probability. This new problem is called p-safety, and it aims to ensure that the given process reaches a forbidden set before leaving its 'working' state space with a probability of less than p. In the situation when an initial probability measure characterizes the initial states, a variant of p-safety is put forward. We call this form of safety weak p-safety. In this work, we characterize both p-safety and weak p-safety and show how to compute them. We employ semi-definite programming to compute p-safety and linear programming to compute weak p-safety. To get to this point, we use certificates of positivity of polynomials translated into the sum of squares and the Bernstein forms.

Key words: safety analysis; stochastic systems; Markov models; optimisation problems; polynomial methods; moment method.

1 Introduction

Safety verification plays an essential role as the instrument of analyzing whether a system works according to the specification requirements.

Usually, a system is said to be safe if it does not violate any system constraints. This notion of stochastic safety has been studied using the concept of barrier certificates (see [21], [34], [25] and the references therein). In this paper, we advance our analytical and computational studies of p-safety initiated in [8]. The concept of p-safety is ultimately related to the notion of risk. Indeed, risk is defined as the product of the probability of a failure, loss, or injury (p-safety) and its cost. There is an extensive body of work on qualitative risk analysis and its application. For instance, [1] conducts risk analysis for a drilling operation, including the probability and consequences of potential accidents scenarios. Specifically, Bayesian network is used to assess the probability of blowout. [17] discusses the computation of collision probability between space-borne objects.

In the probabilistic setting, the concept of p-safety is defined at the confluence of two research streams. One stream focuses on the characterization of the stochastic reach-avoidance problem [28], which is a specialization of stochastic reachability. The second research direction originates in safety engineering and is related to dynamic barrier management. Dynamic barrier management within the overall risk management framework is related to adopting an overall approach to safety [22]. The effective barriers are firstly created to prevent or reduce the impact of accidents. Afterwards, these are continuously monitored to predict and control the risks. In control engineering, the concept of safety barriers (barrier certificates) has been combined with Lyapunov stability theory, in order to control a system with constraints [33], [23] and [31].

In our framework, the objective of p-safety analysis is to classify the initial states according to their significance in the reach-avoid probability computation. This idea can also be related to the hazard identification, which is the first step in the risk assessment process. Hazard identification aims to estimate if any particular item (control action, state, decision) could have the potential to cause harm. In our case, the 'hazard items' are the initial states that lead to an unsafe region with probability bigger than p.

Previously, the series of papers [21], and [34] have devel-

Email addresses: raf@es.aau.dk (Rafael Wisniewski), luminita.bujorianu@strath.ac.uk (Manuela L. Bujorianu).

oped the mathematical apparatus to tackle p-safety comprised stochastic barrier certificates and their stochastic characterization using martingale theory. In [21], a barrier function was proposed for a diffusion process and a switched diffusion process. Later in [34], an optimisation problem was presented for computing p-safety for the switched diffusion and piece-wise deterministic Markov processes. The primary tool for formulating this optimisation was the barrier certificate from [21] combined with the Dirichlet problem's solution. Also related to this work is [2]. It considers a stochastic hybrid system in discrete time with Markov policies. It defines two safety problems of determining the set of initial states for which the process stays safe with a probability p for a given and for some policy. This work proposes dynamic programming methods for solving the two problems.

For the first time, in this paper, we present quantitative analytical characterizations of barrier certificates for general Markov processes. Moreover, optimisation algorithms are developed to approximate the p-safety probabilities.

Many practical applications within robotics, manufacturing, energy production, transportation, to name a few, require the use of Markov models. For safety verification, instead of computing the reachable states, a feasible approach is to use barrier certificates. There is a strong necessity to understand how the certificates are used for the computation of the reach probabilities. Answers to this problem are developed in this paper.

We use the duality between super-martingales and stochastic Lyapunov functions. The latest ones are known as excessive/super-regular functions in the context of probabilistic potential theory. Changing the focus from the martingale theory to potential theory opens a new avenue which makes it possible to characterize the p-safety as an optimisation problem. In short, the potential theory is an analytic tool for studying Markov processes [6]. The part of this theory which we use is the Hunt balayage theorem. It characterizes the p-safety as the infimum of a specific cone of excessive functions - the excessive functions are viewed as the barrier certificates from [21].

We examine two forms of p-safety: p-strong safety and weak p-safety, which we introduced before in [35]. We study the following configuration: a forbidden (unsafe) subset U of a state-space S and the set of initial conditions A. In strong safety, we aim at finding the largest probability that a process starts (deterministically) at a point of A and reaches U before it leaves the state space S. Weak safety is defined similarly, but it allows the use of different initial distributions of the process. First, we provide an analytical characterization of the reach probabilities (or safety functions) using stochastic barrier certificates. Based on the characterizations of both definitions of safety, we provide algorithms for computing safety. The novelty emerges from the fruitful combination of the analytical characterizations provided by the probabilistic potential theory and optimisation. In particular, safety is translated into semi-definite programming. To this end, we employ the sum of squares [20]; whereas, weak safety is converted into linear programming. For this purpose, we use Bernstein forms [15].

The significance of the coupling between potential theory and optimisation is prodigious. It opens new research avenues where analytical characterizations are translated into scalable algorithmic methods, not only for safety, but also for stability and other performance criteria. To that end, we address a broad class of Markov processes - the right continuous Markov processes. This class contains popular processes encountered in control engineering such as diffusion processes, switched diffusion processes, piece-wise deterministic Markov processes [10], and stochastic hybrid systems [7].

The paper is organized as follows. To keep the article self-contained, we have recalled some instrumental concepts from stochastic processes in Section 2. The concepts of safety are introduced in Section 3. The reach-avoidance problem is formulated in Section 4, and it is solved using the super-martingale characterization in Section 5. In Section 6, probabilistic potential theory, specifically Hunt balayage theorem, is employed for formulating an abstract optimisation. Subsequently, in Section 7, the optimisation is re-formulated as a semi-definite programming. A numerical example of p-safety computation for switching diffusion is provided. Section 8 is devoted to the analytic characterization of weak p-safety. Again, the potential theory is shown to be fruitful for the derivation of abstract optimisation, this time for computing weak safety. This optimisation is, in Section 9, re-formulated using Bernstein forms as linear programming. Subsequently, a numerical example of computing weak p-safety for a Brownian motion is given.

Notations

$\mathbb{R}_+ \equiv \{x \in \mathbb{R} |\ x \geq 0\}$ and $\mathbb{Z}_+ \equiv \{x \in \mathbb{Z} |\ x \geq 0\}$. Let $Q(x)$ be a predicate of a variable x. We will use the notation $[Q(x)]$ instead of $\{x \in X |\ Q(x)\}$ if the set X is implicitly known. Occasionally, we write "Q on a set S", it means that $Q(x)$ holds for all $x \in S$. For example, a function $f > 0$ on S means $f(x) > 0$ for all $x \in S$. For two functions f and g, $(f \wedge g)(x) \equiv \min\{f(x), g(x)\}$, and $(f \vee g)(x) \equiv \max\{f(x), g(x)\}$. The Borel sigma-algebra on a topological space $\mathcal{Y}$ is denoted by $\mathcal{B}(\mathcal{Y})$. For a set $A \in \mathcal{B}(\mathcal{Y})$, I_A denotes the indicator function of A. The complement of a set A is denoted by A^c, its closure by $\mathrm{cl}(A)$, its boundary by ∂A, and its interior by $\mathrm{int}(A)$. We say that a set is a domain if it is open and connected.

We say that a subset $\mathcal{K}$ of a vector space is a positive cone if for any $h_1, h_2 \in \mathcal{K}$, and any $\alpha \geq 0$ the following conditions hold:

(1) $h_1 + h_2 \in \mathcal{K}$, and
(2) $\alpha h_1 \in \mathcal{K}$,
(3) $\mathcal{K} \cap (-\mathcal{K}) = \{0\}$.

2 Background

In this section, we recollect some instrumental concepts from stochastic processes, herein the notions of different generators.

Specifically, we study a special class of Markov processes, namely (Borel) right processes [7]. We consider such a Markov process $(X_t) \equiv (X_t)_{t\geq 0}$ on the underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in a Borel space $\mathcal{Y}$ [1]. We associate a family of probabilities $(\mathbb{P}^y) \equiv (\mathbb{P}^y)_{y\in\mathcal{Y}}$ on $\mathcal{Y}$ with the property $\mathbb{P}^y[X_0 = y] = 1$; they are called *Wiener probabilities.* The expectation with respect to $\mathbb{P}^y$ is denoted $\mathbb{E}^y$. To $(\mathbb{P}^y)$, we associate a transition semigroup $(p_t) \equiv (p_t)_{t\geq 0}$ corresponding to the transition probability kernels $p_t(y, A) = \mathbb{P}^y[X_t \in A]$. The action of the kernel p_t on the Banach space $\mathcal{B}_b(\mathcal{Y})$ of bounded measurable real-valued functions $f : \mathcal{Y} \to \mathbb{R}$ is defined by

$$p_t f(y) \equiv \int_{\mathcal{Y}} f(x) p_t(y, dx) = \mathbb{E}^y f(X_t).$$

For $\alpha > 0$, the resolvent V_α is the Laplace transform of transition probabilities (p_t), i.e., $V_\alpha f = \int_0^\infty e^{-\alpha t} p_t f dt$, [24]. In the theory of Markov processes, there exists the Hille-Yosida characterization that provides the equivalence of the following three descriptions of a Markov process: by the transition semigroup, by the resolvent and by the generator, which will be discussed in Subsection 2.1.

For a measurable set B, *the first hitting time* τ_B associated to this set, is

$$\tau_B := \inf\{t \geq 0 | X_t \in B\};$$

whereas, *the first exit time* from B is $\zeta_B = \tau_{B^c}$ (i.e., the first hitting time of the complement of B).

We will often use the notion of a stopped process. For stopping τ, the *stopped process* (X_t^τ) is

$$X_t^\tau \equiv \begin{cases} X_t & \text{if } t \leq \tau \\ \delta & \text{if } t > \tau \end{cases}$$

where δ is an absorbing (cemetery) point added to $\mathcal{Y}$.

[1] $\mathcal{Y}$ is a Borel subset of a complete separable metric space. An example of such a space is $\mathbb{R}^n$ with the standard Euclidean distance

2.1 Generators, Super-martingales, Super-regular, Excessive Functions

Let $(\mathcal{F}_t) \equiv (\mathcal{F}_t)_{t\geq 0}$ be a filtration. We assume that (X_t) is adapted to $(\mathcal{F}_t)$. We recall, a real-valued process (X_t) is a *martingale* (with respect to $(\mathcal{F}_t)$) if $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ for $t > s$ and *super-martingale* if $\mathbb{E}[X_t|F_s] \leq X_s$ for $t > s$. A process (X_t) is a *local (super-) martingale* if there exists a sequence $(T_n)_{n\in\mathbb{N}}$ of stopping times (with respect to $(\mathcal{F}_t)$) such that $T_n \to \infty$ pointwise and the stopped process $(X_t^{T_n})$ is a (super-) martingale.

To a process (X_t), we associate a function-cone whose elements $h : \mathbb{R}^n \to \mathbb{R}$ satisfy the following condition: for each h in this cone, the resulting process (h_t) with $h_t \equiv h(X_t)$ is a local super-martingale. The reason for insisting on the super-martingale property is that we are consequently able to estimate the upper bound of the expected value of (h_t). Specifically, $\mathbb{E}[h_t] \leq \mathbb{E}[h_s]$ for $t \geq s$. This observation gives rise to a profound super-martingale inequality [3]:

$$c\mathbb{P}[\sup_{[a,b]} X_t \geq c] \leq \mathbb{E}[X_a] + \mathbb{E}[-X_b \vee 0]. \qquad (1)$$

Using the transition operator semigroup, one can define the *infinitesimal generator* $\mathcal{G}$ associated to a Markov process, as the derivative at $t = 0$ of the transition semigroup with respect to the sup norm of the Banach space $\mathcal{B}_b(\mathcal{Y})$. Let $\mathcal{DG} \subset \mathcal{B}_b(\mathcal{Y})$ be the set of functions f for which this derivative (denoted by $\mathcal{G}f$) exists. In most cases, the operator semigroup can be itself characterized by its infinitesimal generator. When $\mathcal{DG}$ is large enough, the infinitesimal generator captures the law of the whole dynamics of a Markov process and provides a tool to study its properties.

The infinitesimal generator admits a couple of extensions: the *weak generator*, the *characteristic operator*, and the *extended generator*. The reason for defining different concepts of generators is to increase the domains of the generators to the detriment of having more abstract theory. Specifically, the domain of the infinitesimal generator is smaller than the weak generator, which is again smaller than the domain of the characteristic generator. Whereas, the extended generator has the largest domain. The *weak generator* is defined using the same formula as the strong generator, but considering the point-wise convergence. Later in the paper, in Theorem 15, we will use *the characteristic operator*, Definition 7.5.1 [19]. Denote by $A_k \downarrow x$ a sequence of open sets $\{A_k|\ k \in \mathbb{N}\}$ with $A_{k+1} \subset A_k$, and $\bigcap_{k\in\mathbb{N}} A_k = \{x\}$. The characteristic operator $\mathcal{A}$ is defined by

$$\mathcal{A}f(y) = \lim_{A_k \downarrow x} \frac{\mathbb{E}^y[f(X_{\tau_{A_k}})] - f(y)}{\mathbb{E}^y[\tau_{A_k}]}.$$

Another generalisation is *the extended generator*, which will be the main object of study in this work. We define the *extended (full) generator* $\mathcal{L}$ following [10]. The domain of the extended generator, denoted by $\mathcal{DL}$, is the set of measurable function $h : \mathcal{Y} \to \mathbb{R}$ having the property that there is a measurable function $g : \mathcal{Y} \to \mathbb{R}$ such that the function $t \mapsto g(X_t)$ is almost surely $\mathbb{P}^y$ integrable for each $y \in \mathcal{Y}$, and the process (C_t^h) given by

$$C_t^h \equiv h(X_t) - h(X_0) - \int_0^t g(X_s)ds \qquad (2)$$

is a local martingale (or a martingale in the case of the full generator) with respect to $(\mathcal{F}_t)$. We write $\mathcal{L}h = g$ and call $(\mathcal{DL}, \mathcal{L})$, or even $\mathcal{L}$, an *extended generator*. Note that if N is a measurable set such that $\mathbb{P}^y[\lambda\{t|X_t \in N\} = 0] = 1$ for all $y \in \mathcal{Y}$, where λ is the Lebesgue measure on $\mathbb{R}$, then g may be altered on N without changing the validity of (2). Therefore, the map $h \mapsto g$ is not unique and the extended generator $(\mathcal{DL}, \mathcal{L})$ is a multi-valued operator.

The extended generators of many interesting processes in control have been characterised; herein, diffusion processes, their generalisations jump diffusion processes and switching diffusion processes, also piecewise-deterministic Markov processes.

It is instrumental to understand how the properties of a generator are related to the process $(h(X_t))$ provided that (C_t^h) in (2) is a local martingale. Let $\mathcal{L}$ be the extended generator of (X_t). A measurable function $h \in \mathcal{DL}$ is called a *super-regular* function if $\mathcal{L}h \leq 0$. Notice that if $\mathcal{L}h$ is a polynomial then verifying if h is super-regular boils down to the application of a certificate of positivity [13], e.g., by means of the sum of squares [20] or Bernstein forms [15]. This property will be exposed later in the paper.

Now, (C_t^h) in (2) being a local martingale implies that the process $(h(X_t))$ becomes a local super-martingale whenever h is a super-regular. This result will be instrumental throughout the paper.

Proposition 1 (Th.4.1 [11]) *Let (X_t) be a Markov process with the extended generator $\mathcal{L}$. For a function h, the process (h_t) with $h_t = h(X_t)$ is a local super-martingale if h is a super-regular.*

For the right Markov processes, the super-regular functions can be characterized using the transition semigroup. They coincide with the so-called excessive functions. These play the role of Lyapunov functions for stochastic processes. We say that a non-negative measurable function h is *excessive* [10] if the following two conditions are satisfied:

(1) $p_t h \leq h$ for all $t \geq 0$, and

(2) $\lim_{t\searrow 0} p_t h = h$ (pointwise).

We shall denote the *cone of excessive functions* by $\mathcal{E}_X$.

In general, any excessive function (in the domain of the generator) is super-regular. The opposite result, i.e., any super-regular function is excessive, has been proven for standard and right Markov processes [32].

3 Concepts of Safety

Suppose that S and U are two measurable sets in $\mathcal{B}(\mathcal{Y})$ with $U \subset S$. We think about the set S as the state space, and U as a set representing a dangerous situation, for example, a failure of machinery. We want to compute the probability that the process (X_t) will be, in the future, in a dangerous state. Strictly speaking, we strive to determine the probability that (X_t) reaches U at some time without leaving S. The above statement can be further formalized using the hitting time τ_U of the set U, and the first exit time ζ_S from S. We will study the probability that the sample paths visit U before leaving S, which we write $\mathbb{P}^y[\tau_U < \zeta_S]$. It is natural to think that if $\mathbb{P}^y[\tau_U < \zeta_S]$ is bigger than a certain threshold p, then the state y is considered unsafe. We will examine safety in an infinite time-horizon. The study of the safety in the finite time horizon T can be reduced to the case of the infinite time horizon using the time-space extension of the process [35].

Definition 2 *A state $y \in S$ is (strongly) p-safe if*

$$\mathbb{P}^y[\tau_U < \zeta_S] \leq p. \qquad (3)$$

A state that does not satisfy (3) *is called p-unsafe.*

The definitions of strong p-safety, or for short p-safety, is intimately connected with the property of the following *safety function*, which is called capacitor function (or condenser potential) in the mathematical literature [9]

$$P(y) \equiv P(y; U, S) \equiv \mathbb{P}^y[\tau_U < \zeta_S]. \qquad (4)$$

We can also express the safety function using the indicator function as

$$P(y) = \mathbb{E}^y[I_U(X_{\tau_{U\cup S^c}})], \qquad (5)$$

where $\tau_{U\cup S^c}$ is the first hitting time of $U \cup S^c$.

We extend the safety function to act on Borel sets. For $A \in \mathcal{B}(\mathcal{Y})$, we define

$$P(A) \equiv P(A; U, S) \equiv \sup_{y\in A} P(y).$$

Subsequently, the definition of a p-safe state can be extended to a p-safe set.

Definition 3 *A Borel subset $A \subset S$ is p-safe if all points $y \in A$ are p-safe, or in other words, if*

$$P(A) = \sup\{\mathbb{P}^y[\tau_U < \zeta_S] | \ y \in A\} \le p.$$

For an arbitrary initial measure μ_0, we define the following *safety measure*, which is the action of μ_0 on P

$$(\mu_0 P)(A) \equiv \int_A P(y)\mu_0(dy), \forall A \in \mathcal{B}(\mathcal{Y}). \quad (6)$$

In the next definition, we combine the probability of hitting the forbidden set U with the probability of taking a specific initial value.

Definition 4 *An initial measure μ_0 on A is p-safe if*

$$(\mu_0 P)(A) \le p.$$

We say that the initial measure μ_0 is p-safe if

$$(\mu_0 P) \equiv (\mu_0 P)(S) \le p.$$

For a given initial probability measure μ_0, we will refer to p-safety of μ_0 as weak p-safety (without explicitly referring to μ_0). We will come back to the problem weak p-safety in Section 8. In the next sections, we will address p-safety.

4 Problem Formulation

Each of the definitions in Section 3 creates an intriguing theoretical and practical problem of numerically determining it. Specifically in this paper, for a given measurable set S (the state space of the process (X_t)), we want to solve the reach-avoidance problem, i.e., to identify numerical algorithms to compute the probability $P(A; U, S)$ that (X_t) reaches a Borel set U of S without leaving the set S provided that X_0 belongs to another Borel subset A of S.

Using the hitting time τ_U of U, and the exit time ζ_S from S (recall $\zeta_S = \tau_{S^c}$), we will study the probability that the sample paths starting in A visit U before leaving S.

Problem 5 *We aim to compute*

$$P(A; U, S) = \sup\{\mathbb{P}^y[\tau_U < \zeta_S] | \ y \in A\}.$$

To exemplify this problem, let us consider two cases: a Markov chain when the state space is discrete (finite or countable), and a diffusion process.

Example 6 (Markov chain, Section III.b [29]) *We study a Markov chain with the family $\{p_{yz}\}$ of transition probabilities from the state y to the state z. For a subset S, we define its boundary as follows*

$$\delta S \equiv \{z \in S^c | p_{yz} \ne 0 \text{ for some } y \in S\}.$$

We take the target set U to be a singleton in S, i.e., $U = \{z\}$. We let the initial set A also to be a singleton in S, $A = \{y\}$. The safety problem formulated above reads for the discrete case as the problem of finding the probability that the Markov chain, starting at y, hits z before reaching δS (when δS is nonempty). Therefore, we aim to compute

$$P(\{y\}; \{z\}, S) = \mathbb{P}^y[\tau_z < \tau_{\delta S}]. \quad (7)$$

It is known that the probability $P(\{y\};\{z\},S)$ is a solution of a boundary value problem for a discrete Laplacian [29], which we address next. The discrete Laplacian for a Markov chain is defined as

$$\Delta f(y) \equiv \sum_x (f(y) - f(x)) p_{yx}$$

for all $f : S \to \mathbb{R}$. In the matrix form, $\Delta = I - \mathcal{P}$, where $\mathcal{P} = [p_{yx}]$ is the stochastic matrix and I is the identity matrix. For a typical random walk on a graph, p_{yx} is usually equal to $1/d_y$ (where d_y is the degree of y) when x is adjacent to y, and 0 otherwise. Then $P(y) = P(\{y\}; \{z\}, S)$ is the solution of the following Dirichlet problem

$$\begin{aligned} \Delta P(y) &= 0 \text{ if } y \in S \setminus \{z\} \text{ and} \\ P(z) &= 1, \\ P(w) &= 0 \text{ if } w \in \delta S. \end{aligned}$$

Example 7 (Brownian motion, Example 9.1.3 [19]) *Consider a Brownian motion (B_t) on $\mathcal{Y} = \mathbb{R}^n$. The Laplace operator Δ is defined by*

$$\Delta f \equiv \sum \frac{\partial^2 f}{\partial x_i^2}$$

for all twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$. The characteristic operator of (B_t) is $\mathcal{A} = \frac{1}{2}\Delta$. Also in this example, we let the initial set A be a singleton $\{y\}$; whereas, U is an arbitrary open subset of S such that $\partial U \cap \partial S = \emptyset$. Then $P(y) = P(\{y\}; U, S)$ is the solution of the following Dirichlet problem [19]

$$\begin{aligned} \Delta P &= 0 \quad \text{on} \quad S \setminus U, \\ P &= 1 \quad \text{on} \quad \partial U, \\ P &= 0 \quad \text{on} \quad \partial S. \end{aligned}$$

Problem 8 *For an initial measure μ_0, we strive to compute*

$$P(\mu_0, A; U, S) = (\mu_0 P)(A).$$

In the following sections, we will show how to solve Problems 5 and 8.

5 Barrier Certificates - Super-martingale Characterization

We show that a function h such that $h(X_t)$ is a local super-martingale can be used to estimate the upper bound of $P(A; U, S)$. In the following proposition, we make use of the stopped process $X_t^{\zeta_S}$ of (X_t) with respect to the exit time ζ_S.

Proposition 9 *Let $A, U, S \in \mathcal{B}(\mathbb{R}^n)$ with S bounded, A and U two subsets of S and $\mathrm{cl}(A) \cap \mathrm{cl}(U) = \emptyset$. Consider a cadlag process*[2] *(X_t). Suppose that there is a continuous non-negative function $h : S \to \mathbb{R}_+$ such that $h_t^{\zeta_S} \equiv h(X_t^{\zeta_S})$ is a local super-martingale. Then*

$$H_U \cdot P(A; U, S) \leq H_A, \tag{8}$$

where $H_A \equiv \sup\{h(y) |\ y \in A\}$, $H_U \equiv \inf\{h(y) |\ y \in U\}$.

An intuitive interpretation of the inequality (8) is that the probability of the process (X_t) being unsafe decreases with the gap between the values of the function h on A and U. Later in the paper, Theorems 15 and Theorem 17 will provide a tight bound of $P(A; U, S)$ for a diffusion process and an arbitrary right process, respectively. Before continuing with the proof of the proposition, we will give an example of a function h.

Example 10 *We continue with Example 7 of a Brownian motion on the plane. We denote the closed disk centered at c with radius r by $D_c(r)$. We suppose that $S = D_{(0,0)}(10)$, $U = D_{(0,0)}(1)$, and A is the the annulus with center at 0, internal radius 5 and external radius 10, i.e., $A = D_{(0,0)}(10) \setminus \mathrm{int}(D_{(0,0)}(5))$.*

Suppose $h(x) = 10^2 - x_1^2 - x_2^2$. We show that h is super regular,

$$\mathcal{A}f = -2 < 0,$$

where $\mathcal{A}$ is the characteristic operator of the Brownian motion on the plane. By Proposition 1, if h a super regular function, $(h(X_t))$ is a local super-martingale. Hence, by (8)

$$P(A; U, S) \leq \frac{10^2 - 5^2}{10^2 - 1^2}.$$

[2] (X_t) is a cadlag if its paths $t \mapsto X_t$ are right-continuous with left limits everywhere with probability one.

PROOF. [Proposition 9] From the outset, we observe that $\tau_U(\omega) < \zeta_S(\omega)$ is equivalent to the existence of t_ω such that $I_U(X_{t_\omega}(\omega)) = 1$, and implies that $h_{t_\omega}^{\zeta_S}(\omega) \geq H_U$. Hence,

$$\begin{aligned} P(x) &= \mathbb{P}[\max\{I_U(X_t^{\zeta_S})\}_{t\geq 0} = 1 | X_0 = x] \\ &\leq \mathbb{P}\left[\sup\{h_t^{\zeta_S}\}_{t\geq 0} \geq H_U | X_0 = x\right]. \end{aligned}$$

We fix $\bar{t} > 0$ and consider the sequence (T_n) of stopping times in the definition of a local super-martingale. The process h_t is cadlag, since h is continuous and X_t is càdlàg. We use the super-martingale inequality

$$c\mathbb{P}[\sup\{h_t^{\zeta_S \wedge T_n}\}_{t\in[0,\bar{t}]} \geq c] \leq \mathbb{E}[h_0] = h_0,$$

where again $h_0 = h(X_0)$. Since $\bar{t}$ is arbitrary and $T_n \to \infty$, after substituting $c = H_U$, we arrive at

$$H_U \cdot P(x) \leq h_0. \tag{9}$$

Hence, taking supreme over $X_0 = x \in A$ on both sides of inequality (9), we conclude that

$$H_U \cdot P(A; U, S) \leq H_A.$$

□

Subsequently, we define the notion of a stochastic barrier function.

Definition 11 *We say that a continuous function $h : \mathcal{Y} \to \mathbb{R}_+$ is a super-martingale barrier function for a process (X_t) and a triple (A, U, S) of subsets of $\mathcal{Y}$ if*

(1) $h_t^{\zeta_S}$ is a local super-martingale, and
(2) $\inf\{h(u) |\ u \in U\} \geq \sup\{h(a) |\ a \in A\}$.

In the next proposition, we list properties of the set of all barrier functions.

Proposition 12 *Let $\mathcal{C}_B$ be the set of all super-martingale barrier functions for a process (X_t) and a triple (A, U, S), where S is bounded.*

(I) The set $\mathcal{C}_B$ is a positive cone that contains constant functions.
(II) If $h^1, h^2 \in \mathcal{C}_B$ then $h^1 \wedge h^2 \in \mathcal{C}_B$.

(III) If $\mathcal{C}_B \neq \emptyset$ then there exists a function $h \in C_B$ and $p \in [0,1]$ such that
(a) $h \geq 1$ on U,
(b) $h \leq p$ on A.

PROOF. Part (I) of the proposition follows directly from the definition of a super-martingale. For part (II), we make an observation that for $i \in \{1,2\}$

$$h^1(a) \wedge h^2(a) \leq h^i(a) \leq h^i(u) \text{ for all } (a,u) \in A \times U;$$

hence, $h_1(a) \wedge h_2(a) \leq h_1(u) \wedge h_2(u)$.

Furthermore by monotonicity of the conditional expectation

$$\mathbb{E}[(h^1 \wedge h^2)(X_t)|\mathcal{F}_s] \leq \mathbb{E}[h^i_t|\mathcal{F}_s] \leq h^i_t.$$

For part (III), pick an $f \in \mathcal{C}_B$. Let $\underline{f}_U \equiv \inf\{f(y)|\ y \in U\}$; it is well defined as f is continuous. We define $h \equiv f/\underline{f}_U$, and conclude that by part (I), $h \in \mathcal{C}_B$, and h satisfies conditions (a), and (b) with $p = \sup\{h(y)|\ y \in A\}$. Observe that $p \in [0,1]$, as h is non-negative, and $1 = \inf\{h(b)|\ b \in U\} \geq \sup\{h(a)|\ a \in A\}$.□

We define a partial order on $\mathcal{C}_B$ by

$$h^1 \succeq h^2 \Leftrightarrow \exists \alpha \in \mathcal{C}_B \text{ such that } h^1 = h^2 + \alpha.$$

From (II) in Proposition 12, we conclude that $(\mathcal{C}_b, \succeq)$ is a meet-semilattice, i.e., each two-element subset has a greatest lower bound.

We combine Propositions 9 and 12 in the following corollary.

Corollary 13 *Let $p \in [0,1]$. If there exists a continuous function $h : \mathcal{Y} \to \mathbb{R}_+$ such that $(h^{\zeta_S}_t)$ is a local supermartingale, and*

(a) $h \geq 1$ on U,
(b) $h \leq p$ on A.

are satisfied then $P(A;U,S) \leq p$.

For an excessive function h, h_t is a supermartingale. Hence, the condition in Corollary 13 of $h^{\zeta_S}_t$ being a local supermartingale can be substituted by h being superregular function, and the conclusion of the corollary holds.

Corollary 14 *Let $p \in [0,1]$. If there exists a superregular function $h : \mathcal{Y} \to \mathbb{R}_+$ such that (a) and (b) of Proposition 13 are satisfied then*

$$P(A;U,S) \leq p.$$

We put forward the following idea. We search among all barrier functions h and find the one with the smallest ratio H_A/H_U, where H_A is the supremum of h on the set A and H_U is the infimum of h on U. In the remainder of this section, we will demonstrate that this idea works for the diffusion processes. Whereas in the next section, we will show that it can be generalized to the right processes.

In the next theorem, we specialize the results to diffusion processes. Specifically, by Ch. 9 in [19] if $\mathcal{A}$ is the characteristic generator of a diffusion process, then the probability that (X_t) reaches U before leaving S solves the following Dirichlet problem

$$\begin{aligned} \mathcal{A}P(y) &= 0 \quad \text{for } y \in S \setminus U, \\ P(y) &= 1 \quad \text{for } y \in \partial U, \\ P(y) &= 0 \quad \text{for } y \in \partial S. \end{aligned}$$

The next theorem states that there is an optimisation scheme for finding the probability $P(A;U,S)$.

Theorem 15 *Let S be a bounded subset of $\mathcal{Y} = \mathbb{R}^n$, A, U be two disjoint closed subsets of S. Let $S \setminus U$ be a domain with a smooth boundary. Let (X_t) be a diffusion process with the characteristic operator $\mathcal{A}$. Suppose that*

$$p^* = \inf p \tag{10}$$

subject to $(p,h) \in \mathcal{C} \subseteq [0,1] \times \mathcal{DL}$ defined by: $(p,h) \in \mathcal{C}$ if and only if

(1) $\mathcal{A}h(y) \leq 0$ for all $y \in S \setminus U$,
(2) $h(y) \geq 0$ for all $y \in S$,
(3) $p \geq h(y)$ for all $y \in A$,
(4) $1 \leq h(y)$ for all $y \in U$.

Then

$$P(A;U,S) = p^*.$$

PROOF. By Corollary 14, if $(p,h) \in \mathcal{C}$ then $P(A;U,S) \leq p^*$. It is enough to show that $P(A;U,S) \geq p^*$. To this end, we use Theorem 24.5 in [12] and Theorem 9.2.5 in [19]; the probability $P(y) = \mathbb{P}^y[\tau_U < \zeta_S]$ solves the following boundary value problem

$$\begin{aligned} \mathcal{A}P(y) &= 0 \quad \text{for } y \in S \setminus U, \\ P(y) &= 1 \quad \text{for } y \in \partial U, \\ P(y) &= 0 \quad \text{for } y \in \partial S. \end{aligned}$$

Since $P(A;U,S) = \sup_{x\in A} P(x)$, and conditions 1) to 4) are satisfied, $(P(x), p^*) \in \mathcal{C}$. Hence, $P(A;U,S) \geq p^*$. □

The importance of Theorem 15 is that assuming A, U, and S semi-algebraic (sub-level sets of some polynomials) and compact, and using Putinar's positivstellensatz (positive polynomial theorem) [18], (10) can be solved by means of semidefinite programming. Specifically, Putinar's theorem provides algebraic conditions formulated in terms of the sum of squares to determine if a polynomial is positive on a semi-algebraic set. We will clarify this aspect in Section 7.

Example 16 *Consider the following stochastic differential equation on $\mathcal{Y} = \mathbb{R}^n$*

$$dX_t = f(X_t)dt + \sigma(X_t)dB_t, \tag{12}$$

where (B_t) is the Brownian motion with values in a Euclidean space $\mathbb{R}^l$, and the maps $f : \mathbb{R}^n \to \mathbb{R}^n$, $\sigma : \mathbb{R}^n \to \mathbb{R}^n \times \mathbb{R}^l$ are Lipschitz continuous.

The characteristic operator $\mathcal{A}$ is given as follows: For any differentiable function $h : \mathbb{R}^n \to \mathbb{R}$

$$\mathcal{A}h =< \nabla h, f > + \frac{1}{2}\mathrm{tr}\left(\sigma\sigma^T D^2 h\right),$$

where $\mathrm{tr}()$ stands for the trace of a matrix, $< \nabla h, f >= \sum \frac{\partial h}{\partial x_i} f_i$ and $D^2 h = [\frac{\partial^2 h}{\partial x_i \partial x_j}]$ is the Hessian of the function h.

Assuming that $S \setminus U$ is compact and semialgebraic of the form

$$S \setminus U = \{x \in \mathbb{R}^n |\ g_i(x) \geq 0,\ i = 1 \dots k\},$$

where each g_i is a real polynomial in n variables.

Then the condition (1) in Theorem 15 boils down to

$$-\mathcal{A}h - \sum_{i=1}^{k} s_i g_i$$

is a sum of squares of polynomials for some sum of squares s_i $(i = 1, \dots k)$.

In the rest of the paper, we will generalise Theorem 15 to an arbitrary right process using potential theory.

6 Barrier Certificates - Potential Theory Characterisation

We have categorized the reach-avoidance problem by employing super-martingales. Subsequently, if the process had a characteristic generator, we were able to formulate this categorization using super-regular functions. On the other hand, if a function is super-regular, then it is excessive, recall definition in Section 2.1. We will show that potential theory provides a characterization of the reach-avoidance problem in terms of excessive functions.

We state the main result of this section.

Theorem 17 *We consider a right process, equipped with its extended generator $\mathcal{L}$ with the domain $\mathcal{DL}$. Then*

$$P(A; U, S) = \inf p \tag{13}$$

subject to $(p, h) \in [0, 1] \times \mathcal{DL}$ such that

(1) $h(y) \geq 0$ for all $y \in S$,
(2) $\mathcal{L}h(y) \leq 0$ for all $y \in S$,
(3) $p \geq h(y)$ for all $y \in A$,
(4) $1 \leq h(y)$ for all $y \in U$.

The proof of Theorem 17 follows from several results, which we will present next. In a following subsection, we will show how to use this theorem to compute p-safety.

6.1 Proof of the main result

In a nutshell, Theorem 17 leans upon Hunt's balayage theorem, Theorem 49.5 in [27]. Hunt's theorem provides the characterization of the hitting distributions in terms of excessive functions. It was also used in solving the Dirichlet problem via the balayage method.

At the outset, we define the set $\mathcal{E}_X^S$ of functions on $\mathcal{Y}$ excessive with respect to the restriction of the underlying stochastic process (X_t) to a set $S \in \mathcal{B}(\mathcal{Y})$.

Definition 18 (Excessive functions on a set) *We say that $f \in \mathcal{E}_X^S$ if and only if for all $y \in S$ and $t \geq 0$, $\mathbb{E}^y f(X_t^{\zeta_S}) \leq f(y)$, and $\lim_{t\searrow 0} \mathbb{E}^y f(X_t^{\zeta_S}) = f(y)$. We call $\mathcal{E}_X^S$ the cone of excessive functions restricted to S.*

In other words, $\mathcal{E}_X^S = \mathcal{E}_{X^{\zeta_S}}$, recall that $(X_t^{\zeta_S})$ is the stopped process of (X_t) with respect to the exit time ζ_S.

We define the set of *potential barrier functions*

$$\mathcal{K} \equiv \{h \in \mathcal{E}_X^S |\ h \geq 1 \text{ on } U\}. \tag{14}$$

In order to avoid topological complications in the proof of our main result, the following assumption is in force.

Assumption 19 *The set N of irregular points of U, i.e., those points $y \in U$ for which $\mathbb{P}^y\{\tau_U > 0\} = 1$ is a polar set, in the sense that $\mathbb{P}^y[\tau_N < \zeta_s] = 0$ for all $y \in S$.*

For example, Assumption 1 holds for Brownian motion and diffusion processes for which the diffusion coefficient matrix has a bounded inverse and the drift coefficient satisfies the Novikov condition, Sec. 9.2 in [19].

Theorem 20 *Let (X_t) be a right process. Suppose A, U, $S \in \mathcal{B}(\mathcal{Y})$, and A and U are subsets of S. Furthermore, Assumption 19 is satisfied.*

Then

$$P(A;U,S) = \inf_{h\in\mathcal{K}} \sup_{x\in A} h(x),$$

where $\mathcal{K}$ is the set of potential barrier functions.

PROOF. The proof consists of two steps:

(1) Show: $P(A;U,S) = \sup_{x\in A} \inf_{h\in\mathcal{K}} h(x)$,
(2) Show: $\sup_{x\in A} \inf_{h\in\mathcal{K}} h(x) = \inf_{h\in\mathcal{K}} \sup_{x\in A} h(x)$.

Step 1. At the outset, consider the réduite (or *reduced function*) of $f : S \to \mathbb{R}_+$, Sec. 5.5.6 in [7], given by

$$R(f) \equiv \inf\{h \in \mathcal{E}_X^S |\ h \geq f\}.$$

For the set U and $v \in \mathcal{E}_X^S$, we define the reduced function (réduite) of v on U by

$$R_U(v) \equiv R(I_U v).$$

Specifically for $v = 1 : y \mapsto 1$,

$$\begin{aligned} R_U 1(y) &= \inf\{h(y) |\ h \in \mathcal{E}_X^S,\ h \geq 1 \text{ on } U\} \\ &= \inf_{h\in\mathcal{K}} h(x). \end{aligned}$$

But by the Hunt balayage theorem,

$$P(y) = R_U 1(y). \tag{15}$$

Hence,

$$P(A;U,S) = \sup_{x\in A} \inf_{h\in\mathcal{K}} h(x)$$

Step 2. We define the 'value function' $H : \mathcal{K} \times A \to \mathbb{R}_+$ by

$$H(h,y) := h(y).$$

It remains to show that

$$\sup_{y\in A} \inf_{h\in\mathcal{K}} H(h,y) = \inf_{h\in\mathcal{K}} \sup_{y\in A} H(h,y).$$

First, we notice that

$$\sup_{y\in A} \inf_{h\in\mathcal{K}} H(h,y) \leq \inf_{h\in\mathcal{K}} \sup_{y\in A} H(h,y).$$

always hold.

In the remainder of the proof, we will show the opposite inequality. To this end, we use the concept of balayage $B_U(v)$ of $R_U(v)$, i.e.,

$$B_U(v) = \sup_{\alpha>0} \alpha V_\alpha(R_U(v)),$$

where $(V_\alpha)_{\alpha>0}$ is the resolvent corresponding to the transition probabilities (p_t) of the process (X_t). Here, the definition of $B_U(v)$ is less important. Rather, we will use the following properties of the balayage:

(1) $B_U^S(v) \in \mathcal{E}_X^S$,
(2) $R_U^S(v) \geq B_U^S(v)$,
(3) $B_U^S(v) = R_U^S(v)$ on $S \setminus U$,
(4) $B_U^S(v) = R_U^S(v) = v$ on $U \setminus N_v$, where N_v is a negligible subset of U with $N_v \subseteq N$, and N is the set of irregular points of U.

We consider the following two cases:

(i) The set of irregular point N is empty. In this case, the balayage and the reduced function coincide.

(ii) The set of irregular point N is nonempty, but it is a polar set (according to the Assumption 19).

First, we observe that the set N does not have any contribution to the safety measure, i.e.,

$$P(A;U \setminus N,S) = P(A;U,S).$$

For this N, define the set $\mathcal{K}_1 \equiv \{h \in \mathcal{E}_X^S |\ h \geq 1 \text{ on } U \setminus N\}$. Specifically, for the case (i), this set coincides with $\mathcal{K}$. For both cases, $B_U(1) \in \mathcal{K}_1$, since corresponding negligible set N_v for $v \equiv 1$ on S is a subset of N.

We compute

$$\begin{aligned} P(A;U \setminus N,S) &= \sup_{x\in A} \inf_{h\in\mathcal{K}_1} h(x) \leq \inf_{h\in\mathcal{K}_1} \sup_{x\in A} h(x) \\ &\leq \sup_{x\in A} B_U(1)(x) = P(A;U,S). \end{aligned}$$

But $P(A;U \setminus N,S) = P(A;U,S)$. Hence,

$$P(A;U,S) = \inf_{h\in\mathcal{K}_1} \sup_{y\in A} h(y).$$

Since $\mathcal{K} \subseteq \mathcal{K}_1$,

$$\begin{aligned} \inf_{h\in\mathcal{K}} \sup_{y\in A} h(y) \leq \inf_{h\in\mathcal{K}_1} \sup_{y\in A} h(y) &= P(A;U,S) \\ &= \sup_{y\in A} \inf_{h\in\mathcal{K}} h(y). \end{aligned}$$

We conclude that

$$P(A;U,S) = \inf_{h\in\mathcal{K}} \sup_{y\in A} h(y).$$

□

The importance of the last theorem is that it allows to formulate an optimisation problem. To this end, we articulate the following proposition.

Proposition 21 *Let* $\mathcal{P} \equiv \{p \in \mathbb{R} | \ \exists h \in \mathcal{K}, \forall y \in A, \ p \geq h(y)\}$. *Then*

$$\inf \mathcal{P} = \inf_{h \in \mathcal{K}} \sup_{y \in A} h(y).$$

PROOF. Let $a = \inf_{h \in \mathcal{K}} \sup_{y \in A} h(y)$, and define

$$\mathcal{Q} \equiv \{(p, h) \in \mathbb{R} \times \mathcal{K} | \ \forall y \in A, \ p \geq h(y)\},$$

and notice that for all $(p, h) \in Q$

$$p \geq \inf_{h \in \mathcal{K}} \sup_{y \in A} h(y).$$

Hence, $\inf \mathcal{P} \geq \inf_{h \in \mathcal{K}} \sup_{y \in A} h(y)$.

On the other hand, for any sufficiently small $\epsilon > 0$, there is $(h_\epsilon, y_\epsilon) \in \mathcal{K} \times A$ such that $h_\epsilon(y_\epsilon) = a + \epsilon$. Furthermore, $a + \epsilon \in \mathcal{P}$ and $a + \epsilon \geq \inf \mathcal{P}$. Since ϵ is arbitrary small, $a \geq \inf \mathcal{P}$. □

PROOF. [Theorem 17] The proof of Theorem 17 follows from Theorem 20 and Proposition 21.

7 Computation of p-safety

In this section, we will show how to transform Theorem 17 into semi-definite optimisation. To this end, we will use polynomial certificates of positivity.

Suppose that real polynomials are dense in $\mathcal{DL}$. To illustrate, for diffusion processes and switching diffusion processes, the C^2 functions are dense in the domain $\mathcal{DL}$ of the extended generator. On the other hand, by Stone-Weierstrass theorem, on a compact set S polynomials are dense in the set of all continuous functions. There are also available results in [26] that if there exists $c > 0$ such that $\int e^{2c|x|} \mu(dx) < \infty$, where $|x| = \sum_{j=1}^{k} |x_j|$, then the polynomials are dense in the space $L^2(\mathbb{R}^n; \mu)$. Consequently, optimisation defined in Theorem 20 can be formulated as the sum of squares programming. From the outset, we say that a basic semi-algebraic set B is generated by a family of polynomials $\mathcal{F} = \{g_1, \dots, g_m\}$ if

$$B = G(\mathcal{F}) \equiv \{x \in \mathbb{R}^n | \ g_1(x) \geq 0, \dots, g_m(x) \geq 0\},$$

and the quadratic module generated by $\mathcal{F}$ is

$$Q(\mathcal{F}) \equiv \left\{ s_0 + \sum_{i=1}^{m} g_i s_i | s_0, s_1, \dots, s_m \in \Sigma^2[X] \right\},$$

where $\Sigma^2[X]$ is the cone of the sum of squares of polynomials (SOS). We suppose that there is $g \in Q(\mathcal{F})$ such that $[g(x) \geq 0]$ is compact. In this setup, Putinar's Positivstellensatz [14] pronounces: If a polynomial p is positive on a compact basic semi-algebraic set B then $p \in Q(\mathcal{F})$.

Let S, A, and U be compact basic semi-algebraic sets, i.e., for a finite family of polynomials $\mathcal{F}_i$, $i \in \{S, \ A, \ U\}$, $S = G(\mathcal{F}_S)$, $A = G(\mathcal{F}_A)$, and $U = G(\mathcal{F}_U)$. The application of Theorem 17 together with the Putinar's Positivstellensatz results in the following sum of squares programming:

Find the minimum of p such that

$$h \in Q(\mathcal{F}_S) \tag{16a}$$
$$-\mathcal{L}h \in Q(\mathcal{F}_S) \tag{16b}$$
$$p - h \in Q(\mathcal{F}_A) \tag{16c}$$
$$h - 1 \in Q(\mathcal{F}_U). \tag{16d}$$

7.1 Numerical Study

We illustrate Theorem 17 in an example of a switching diffusion process, for short SDP. To this end, we recall that an SDP is a hybrid process, whose continuous states evolve as specified by stochastic differential equations (SDEs) and the jumps between them is triggered by a continuous time Markov chain.

Definition 22 (SDP [7]) *A switching diffusion process is a collection*

$$(n, Q, (f, \sigma), \nu_0, \{\lambda_{ij} | \ (i, j) \in Q \times Q\}),$$

where

- Q *is a finite discrete state space;*
- *for* $n \in \mathbb{N}$, $\mathcal{Y} = Q \times \mathbb{R}^n$ *is the SDP (hybrid) state space;*
- $f : \mathcal{Y} \to \mathbb{R}^n$ *is the drift term;*
- $\sigma : \mathcal{Y} \to \mathbb{R}^{n \times m}$ *is the diffusion term;*
- $\nu_0 : \mathcal{B}(\mathcal{Y}) \to [0, 1]$ *is an initial probability measure on* $(\mathcal{Y}, \mathcal{B}(\mathcal{Y}))$*;*
- $\lambda_{i,j} : \mathbb{R}^n \to \mathbb{R}$ *are the state-dependent transition rates with*

$$\lambda_{i,j}(x) \geq 0 \text{ for } x \in \mathbb{R}^n \text{ and } i \neq j,$$
$$\lambda_{i,i}(x) = -\sum_{j \in Q, i \neq j} \lambda_{i,j}(x) \text{ for all } x \in \mathbb{R}^n \text{ and } i \in Q.$$

The execution of SDP is a two component process (q_t, X_t) with values in $\mathcal{Y}$ that satisfies the SDE (17) and the transition probabilities (18)

$$dX_t = f(q_t, X_t)dt + \sigma(q_t, X_t)dB_t, \tag{17}$$

where B_t is the m-dimensional Brownian motion, (q_0, X_0) has distribution ν_0, and

$$\mathbb{P}[q_{t+\delta} = j|\ q_t = i,\ X_s,\ q_s, s \le t] = \lambda_{i,j}(X_t)\delta + o(\delta) \tag{18}$$

for $i \neq j$.

In our specific example, we consider $\mathcal{Y} = \{0, 1\} \times \mathbb{R}^2$, the drift

$$f(0, x) = \begin{bmatrix}1 & 1.4\end{bmatrix}^{\mathrm{T}},\quad f(1, x) = \begin{bmatrix}1.4 & 1\end{bmatrix}^{\mathrm{T}},$$

the diffusion term

$$\sigma(0, x) = \sigma(1, x) = 0.5I,$$

where I is the identity matrix of size 2, and the transition rates $\lambda_{0,1} = \lambda_{1,0} = 10$.

Let $D_c(r)$ denote the closed disk centered at c and with radius r. We suppose that the state space S, the set A of initial states and the forbidden set U are as follows

$$S = \{0, 1\} \times D_{(0,0)}(10),\ A = \{0, 1\} \times D_{(0,0)}(1),$$
$$U = \{0, 1\} \times D_{(5,5)}(1).$$

For the computation of p-safety, it will be instrumental to use the well-known expression of infinitesimal generator of associated to SDP [4]. This encapsulates a part corresponding to the diffusion component and another part associated to the switching part. Explicitly, for any function $h : \mathcal{Y} \to \mathbb{R}$ with $h(i, \cdot) \in C^2(\mathbb{R}^2)$, $i \in \{0, 1\}$, the generator $\mathcal{L}$ is defined by

$$\begin{aligned}\mathcal{L}h(i, x) \equiv\ & \frac{1}{2}\mathrm{tr}(\sigma(i, x)\sigma^{\mathrm{T}}(i, x)\mathrm{D}^2 h(i, x)) \\ & + \langle f(i, x), \nabla h(i, x)\rangle \\ & + \lambda_{i,i+1}(x)(h(i + 1, x) - h(i, x)),\end{aligned}$$

where $\mathrm{tr}(\cdot)$ stands for the trace, ∇h is the gradient and $\mathrm{D}^2 h$ is the Hessian of $h(i, \cdot)$, and $i+1$ is to be understood modulo 2. Concretely,

$$\begin{aligned}\mathcal{L}h_0(x) = 0.125 &\left(\frac{\partial^2 h_0}{\partial x_1^2} + \frac{\partial^2 h_0}{\partial x_2^2}\right) + \frac{\partial h_0}{\partial x_1} + 1.4\frac{\partial h_0}{\partial x_2} \\ &+ 10(h_1(x) - h_0(x)), \\ \mathcal{L}h_1(x) = 0.125 &\left(\frac{\partial^2 h_1}{\partial x_1^2} + \frac{\partial^2 h_1}{\partial x_2^2}\right) + 1.4\frac{\partial h_1}{\partial x_1} + \frac{\partial h_1}{\partial x_2} \\ &+ 10(h_0(x) - h_1(x)).\end{aligned}$$

To compute p-safety, we use Yalmip optimisation toolbox for Matlab. The code is available on `https://github.com/SecureProject/Safety`). The disks $D_{(0,0)}(10) = [g_1 \ge 0]$, $D_{(0,0)}(1) = [g_2 \ge 0]$ and $D_{(5,5)}(1) = [g_3 \ge 0]$ are defined by the polynomials

$$\begin{aligned}g_1(X_1, X_2) &= 10^2 - X_1^2 - X_2^2, \\ g_2(X_1, X_2) &= 1 - X_1^2 - X_2^2, \\ g_3(X_1, X_2) &= 1 - (X_1 - 5)^2 - (X_2 - 5)^2.\end{aligned}$$

We write $h_i(\cdot) := h(i, \cdot)$. We use the following instance of (16)

- $\mathrm{sos}(h_0 - s_1 * g_1)$,
- $\mathrm{sos}(h_1 - s_2 * g_1)$,
- $\mathrm{sos}(-\mathcal{L}h_0 - s_3 * g_1)$,
- $\mathrm{sos}(-\mathcal{L}h_1 - s_4 * g_1)$,
- $\mathrm{sos}(p - h_0 - s_5 * g_2)$,
- $\mathrm{sos}(p - h_1 - s_6 * g_2)$,
- $\mathrm{sos}(h_0 - 1 - s_7 * g_3)$,
- $\mathrm{sos}(h_1 - 1 - s_8 * g_3)$,

where s_k for $k \in \{1, \dots, 8\}$ are unknown SOS (polynomials in $\Sigma^2[X]$), and sos stands for an SOS constraint.

The result of running the numerical example is $P(A; U, S) = 0.28$. The influence of the transition rates on the safety of SDP can be studied by testing different values of $\lambda_{i,i+1}$. In the extreme situation of no switches between the diffusion processes, i.e., for $\lambda_{i,i+1} = 0$, $P(A; U, S) = 0.27$.

8 Weak p-safety

In the last part of Section 3, we have defined the concept of weak p-safety. In this section, we will show how to compute it, i.e., how to compute the smallest number p such that the process (X_t) is weak p-safe. Specifically, we regard the situation when not all of the states in the set A are equally probable, but rather, there exists an initial probability measure μ_0 with $\mathrm{supp}\mu_0 \subseteq S$.

Problem 23 *We want to compute the probability that the process (X_t), with the initial distribution μ_0 of X_0 hits a subset U of S without leaving the set S. In other words, for a given initial measure μ_0, we strive to compute*

$$\langle \mu_0, P\rangle = (\mu_0 P)(\mathcal{Y}) = \int_S P(y)\mu_0(dy),$$

where P is the safety function defined in (4) *(in the last equality, we have used that the support of μ_0 is a subset of S).*

We formulate the main result of this section, a solution to Problem 23, which shows that the weak safety can be computed employing the following optimisation.

Theorem 24 *Suppose that $\mathcal{L}$ is the extended generator of a right process (X_t) with the initial distribution μ_0. Suppose that*

$$p^* = \sup_{\mu} \mu(U) \tag{19}$$

subject to probability measures μ on S that satisfy

$$\langle \mu_0, f \rangle \geq \langle \mu, f \rangle$$

for all $f \in \mathcal{DL}$ such that

(1) $f \geq 0$ on S,
(2) $\mathcal{L}f \leq 0$ on S.

Then

$$\langle \mu_0, P \rangle = p^*.$$

The proof follows from a number of steps, which will present next. Subsequently, we will illustrate how to use the theorem to compute weak p-safety.

8.1 Proof of the main result

In the reminding part of this section, we will prove Theorem 24. To this end, we follow [5], and on the set of probability measures on S, we define the *balayage order* with respect to the cone $\mathcal{E}_X^S$ of excessive functions (restricted to S), see Definition 18,

$$\nu_1 \vdash \nu_2 \Leftrightarrow \langle \nu_1, f \rangle \geq \langle \nu_2, f \rangle, \forall f \in \mathcal{E}_X^S.$$

The next proposition shows how to evaluate the weak p-safety utilizing the balayage order.

Theorem 25 *The weak safety measure can be computed from*

$$\langle \mu_0, P \rangle = \sup_{\mu_0 \vdash \mu} \mu(U). \tag{20}$$

PROOF. We use the reduced function introduced in the beginning of the proof of Theorem 20

$$R(g)(y) = \inf\{h(y) | h \in \mathcal{E}_X^S, g \leq f\}. \tag{21}$$

Furthermore, we have shown in (15) that $P(y) = R_U 1(y)$, but $R_U 1(y) = R(I_U)(y)$. Therefore,

$$\langle \mu_0, P \rangle = \langle \mu_0, R(I_U) \rangle$$

From [5], it is known that for a given probability measure ν on S, the reduced function satisfies the following relation

$$\langle \nu, R_S(g) \rangle = \sup_{\nu \vdash \mu} \langle \mu, g \rangle. \tag{22}$$

From (21) and (22), the following formula is deduced

$$\langle \mu_0, P \rangle = \sup_{\mu_0 \vdash \mu} \langle \mu_0, I_U \rangle = \sup_{\mu_0 \vdash \mu} \mu(U). \tag{23}$$

□

PROOF. [Theorem 24] A super-regular function is excessive for right Markov processes. As a consequence, of Theorem 25, after unfolding the definitions of the balayage order and the cone of excessive functions we obtain the desired conclusion. □

9 Computation of weak p-safety

In this section, we will show how to compute weak p-safety employing Bernstein forms, i.e., polynomials represented in the Bernstein polynomial basis.

We suppose that $\mathcal{Y} = \mathbb{R}^n$, and the closure of the interior of S is S itself. Furthermore, we assume that S is partitioned by a finite family $\mathcal{S}$ of n-simplices (simplices of the dimension n) such that

(1) $S = \bigcup_{\sigma \in \mathcal{S}} \sigma$, and
(2) $\sigma_1 \cap \sigma_2$ is a face of both σ_1 and σ_2, $\forall \sigma_1, \sigma_2 \in \mathcal{S}$.

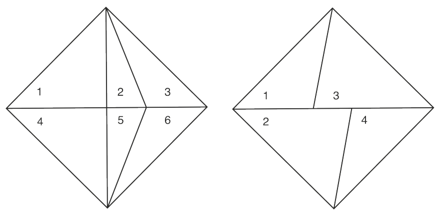

Fig. 1. The state-space S is a diamond. The partitioning by simplices $\{1, \ldots, 6\}$ to the right satisfies Conditions (1) and (2); whereas, the partitioning to the left does not satisfy Condition (2). For example, the intersection of simplices 1 and 2 is not a face of simplex 2.

We suppose that the set of polynomials is dense in $\mathcal{DL}$. We will represent polynomials in *Bernstein basis*

$$p = \sum_{|\alpha|=D} b_\alpha(p, D, \sigma) B_\alpha(D, \sigma),$$

where $|\alpha| = \sum_{i=0}^{n} \alpha_i$, $b_\alpha(p, D, \sigma)$ are the Bernstein coefficients, and $B_\alpha(D, \sigma)$ are the Bernstein basis polynomials. The coefficients and the basis depend not only on the polynomial p and the degree D, but also on the specific simplex σ. To see this, we recall that the *Bernstein basis polynomials* are defined by

$$B_\alpha(D,\sigma) = \binom{D}{\alpha}\lambda^\alpha,$$

where $\lambda = (\lambda_0, \dots \lambda_n)$ are the *barycentric coordinates*, $\lambda^\alpha = \prod \lambda_i^{\alpha_i}$, $\binom{D}{\alpha} = \frac{D!}{\alpha_0! \dots \alpha_n!}$, and the barycentric coordinates are functions of points in $\mathbb{R}^n$. Suppose that the simplex σ is given by $n+1$ affinely independent point $\sigma_0, \dots, \sigma_n \in \mathbb{R}^n$. Since $x = \lambda_0\sigma_0 + \dots + \lambda_n\sigma_n$, and $\sum_{i=0}^n \lambda_i = 1$, we have

$$\lambda = \begin{bmatrix} \sigma_0 & \dots & \sigma_n \\ 1 & \dots & 1 \end{bmatrix}^{-1} \begin{bmatrix} x \\ 1 \end{bmatrix}. \tag{24}$$

We have chosen to represent polynomials in Bernstein basis because there is a straightforward way to verify whether they are non-negative on a simplex.

Theorem 26 (Bernstein Theorem [15]) *Suppose that a polynomial p of degree d is non-negative. Then there is a degree $D \geq d$ such that the coefficients $b_\alpha(\sigma)$ are all none-negative.*

The choice of sufficiently large degree D is often necessary to certify positivity of a polynomial. To this end, we employ *degree elevation.* The bounds on the degree necessary to certify positivity with Bernstein coefficients are provided in [15]. Suppose p is a polynomial of degree d positive on a standard simplex σ. If

$$D > \frac{d(d-1)}{2} \frac{\max_{|\alpha|=d} |b_\alpha(p,d,\sigma)|}{m},$$

where m is the minimum of p over σ, then Bernsten coefficients certify positivity of p. Specifically, [16] provides an example of $p(x) = 5x^2 - 4x + 1$ positive on $\sigma = [-1, 1]$, which has to be elevated to at least degree $D = 21$ to be certified for positivity. In such a situation, [15] proposes to use a subdivision of the partitioning $\mathcal{S}$.

Suppose a polynomial p is represented in both the Bernstain basis of degree D and degree $D' > D$

$$\begin{aligned} p(x) &= \sum_{|\alpha|=D'} b_\alpha(p,D',\sigma)B_\alpha(D',\sigma) \\ &= \sum_{|\gamma|=D} b_\gamma(p,D,\sigma)B_\gamma(D,\sigma). \end{aligned}$$

As a consequence, the coefficients are related by

$$b_\gamma(p,D',\sigma) = \sum_{\substack{\gamma-\alpha\geq 0 \\ |\gamma-\alpha|=D'-D}} a_\alpha^\gamma b_\alpha(p,D',\sigma), \tag{25}$$

where

$$a_\alpha^\gamma = \binom{D'}{\alpha}\binom{D'-D}{\gamma-\alpha}\binom{D}{\gamma}^{-1}.$$

To ease the notation, we use the lexicographic order and collect the coefficients $b_\alpha(p,D,\sigma)$, $|\alpha| = D$, in the vector $b(p,D,\sigma)$. Consequently, the vector $b(p,D,\sigma)$ has $N_D = \binom{D+n}{n}$ entries. Similarly, we collect the Bernstein polynomials $B_\alpha(D,\sigma)$ in the vector $B(D,\sigma)$.

To formulate the next statement, we recall the definition of a standard simplex. The *standard simplex* $\Delta(n)$ is the following subset of $\mathbb{R}^n$

$$\Delta(n) \equiv \{Y \in \mathbb{R}^n |\ Y \geq 0 \text{ and } \sum_{i=1}^n Y_i = 1\}.$$

A polynomial can be represented in Bernstein basis with respected to an arbitrary simplex in $\mathcal{S}$. Nonetheless, for any simplex $\sigma \in \mathcal{S}$, there is a linear isomorphism $T(D,\sigma)$ such that [30]

$$b(p,D,\sigma) = T(D,\sigma)b(p,D,\Delta(n)). \tag{26}$$

The extended generator $\mathcal{L}$ in Theorem 24 acting on polynomials give rise to the linear operator $L(D,\sigma)$ acting on the vector $b(p,D,\sigma)$ of Bernstein coefficients. It is defined by

$$L(D,\sigma)b(p,D,\sigma) = b(\mathcal{L}p,D,\sigma) \tag{27}$$

for all polynomials p in $\mathcal{DL}$. Notice that $L(D,\sigma)$ is well-defined since $\mathcal{L}$ is a linear operator.

The following lemma will be instrumental.

Lemma 27 *Let f be a real polynomial on S, and S be partitioned by a finite family of simplices $\mathcal{S}$.*

Then $f \geq 0$ on S, and $-\mathcal{L}f \geq 0$ on S if and only if there is a degree D such that

$$T(D,\sigma)b(f,D,\Delta(n)) \geq 0 \textit{ for all } \sigma \in \mathcal{S}, \tag{28}$$

and

$$-T(D,\sigma)L(D,\Delta(n))b(f,D,\Delta(n)) \geq 0 \textit{ for all } \sigma \in \mathcal{S}. \tag{29}$$

In (28) and (29), we have used the convention $v \geq 0$ in $\mathbb{R}^{N_D}$ meaning that each entry $v_i \geq 0$.

PROOF.

By Theorem 26, there exists a degree D such that

$$b(f, D, \sigma) \geq 0 \text{ for all } \sigma \in \mathcal{S} \tag{30}$$

and

$$-b(\mathcal{L}f, D, \sigma) \geq 0 \text{ for all } \sigma \in \mathcal{S}. \tag{31}$$

We combine (30) with (26) to conclude the inequality in (28). To prove, the inequality in (29), we observe

$$\begin{aligned} b(\mathcal{L}f, D, \sigma) &= T(D, \sigma) b(\mathcal{L}f, D, \Delta(n)) \\ &= T(D, \sigma) L(D, \Delta(n)) b(f, D, \Delta(n)). \end{aligned}$$

□

We define the *Bernstein moments* (of degree D) of a measure μ on a simplex σ by

$$Y_\alpha(\mu, D, \sigma) \equiv \int_\sigma B_\alpha(D, \sigma) d\mu.$$

We collect the moments $Y_\alpha(\mu, D, \sigma)$ in a vector of moments $Y(\mu, D, \sigma)$. Specifically, the vector of the Bernstein moments of the initial measure μ_0 is

$$Y_0(D) \equiv Y(\mu_0, D, \sigma) = \int_S B(D, \sigma) d\mu_0.$$

However, not all vectors Y are vectors of Bernstein moments on a simplex σ.

Lemma 28 *Suppose $S \subset \mathbb{R}^n$ is closed, and $\sigma \subset S$ is an n-simplex. Let $(Y(D))_{D \in \mathbb{Z}_+}$ be a sequence of vectors with $Y(D) \in \mathbb{R}^{N_D}$. There exists a probability measure μ on S such that*

$$\int_S B(D, \sigma) d\mu = Y(D)$$

for all $D \in \mathbb{Z}_+$ if and only if

$$Y(D) \in \Delta(N_D). \tag{32}$$

Furthermore, for all $|\alpha| = D$,

$$Y_\alpha(D) = \binom{D+1}{\alpha} \sum_{i=0}^{n} \binom{1}{e_i} \binom{D}{\alpha + e_i}^{-1} Y_{\alpha+e_i}(D+1), \tag{33}$$

where e_i is the vector of zeros in all entries except the entry $i+1$ where it is 1.

PROOF. Since

$$\sum_{|\alpha|=D} B_\alpha(D, \sigma) = 1,$$

we have

$$1 = \int_S d\mu = \sum_{|\alpha|=D} \int_S B_\alpha(D, \sigma) d\mu = \langle \mathbb{1}, Y(\mu, D, \sigma) \rangle, \tag{34}$$

where $\mathbb{1}$ is the vector of entries 1. The above equality shows (32).

Let $p = B_\alpha(D, \sigma)$ for some $|\alpha| = D$. Subsequently, by integrating (25) on S for $D' = D + 1$,

$$\begin{aligned} Y_\alpha(\mu, D, \sigma) &= \sum_{|\gamma|=D+1} b_\gamma(\mu, D+1, \sigma) Y_\gamma(\mu, D+1, \sigma) \\ &= \sum_{i=0}^{n} b_{\alpha+e_i}(\mu, D+1, \sigma) Y_{\alpha+e_i}(\mu, D+1, \sigma) \\ &= \sum_{i=0}^{n} \binom{D+1}{\alpha} \binom{1}{e_i} \binom{D}{\alpha + e_i}^{-1} Y_{\alpha+e_i}(\mu, D+1, \sigma). \end{aligned}$$

The set S is closed; hence, by Riesz-Haviland theorem [14, Theorem 3.1], there is a finite Borel measure μ_D such that $Y(D)$ is a vector of Bernstein moments of μ_D if and only if $\langle Y(D), F \rangle \geq 0$ for all $F \geq 0$. This is equivalent to $Y(D) \geq 0$. Combining it with (34) gives $Y(D) \in \Delta(N_D)$. Since, $Y(D)$s are related by (33), μ can be chosen such that $\mu = \mu_D$ for all $D \in \mathbb{Z}_+$. □

It will also be instrumental to recall the definition of a dual cone. Let $\mathcal{C}$ be a cone in $\mathbb{R}^N$, the dual cone $\mathcal{C}^*$ of $\mathcal{C}$ is

$$\mathcal{C}^* \equiv \{Y \in \mathbb{R}^N | \ \langle Y, F \rangle \geq 0 \text{ for all } F \in \mathcal{C}\}.$$

For a subset $\mathcal{C} \subset \mathbb{R}^N$ and a vector $Y \in \mathbb{R}^N$, we write $Y + \mathcal{C} \equiv \{Y + Z | \ Z \in \mathcal{C}\}$.

We are ready to state the main result of this section.

Theorem 29 *Let S be partitioned by a finite family of simplices $\mathcal{S}$. Suppose that (q_k) is a sequence of polynomials converging point-wise to the indicator function I_U that is bounded on S, i.e., there is c such that $|q_k(x)| < c$ for $x \in S$ and $k \in \mathbb{N}$.*

Let μ_0 be the initial distribution, with its vector of Bernstein moments

$$Y_0(D) \equiv Y(\mu_0, D, \Delta(n)).$$

We define the following objects:

- *The cone* $\mathcal{C}(D)$ *in* $\mathbb{R}^{N_D}$ *given by:*

$$\mathcal{C}(D) \equiv \bigcap_{\sigma\in\mathcal{S}} \{F \in \mathbb{R}^{N_D} |\, TF \geq 0, \\ -TL(D,\Delta(n))F \geq 0\},$$

where $T \equiv T(D,\sigma)$*, and* $L(D,\Delta(n))$ *is given in* (27).

- *The polyhedron* $\mathcal{D}(D)$ *given by:*

$$\mathcal{D}(D) \equiv (Y_0(D) - \mathcal{C}(D)^*) \cap \Delta(N_D).$$

- *The sequence:*

$$p_k \equiv \sup_{D\in\mathbb{Z}_+} \sup_{Y\in\mathcal{D}(D)} \langle Y, b(q_k, D, \Delta(n))\rangle \quad (35)$$

Then

$$\langle \mu_0, P\rangle = \lim_{k\to\infty} p_k.$$

PROOF. We denote by F the vector of Bernstein coefficients $b(f, D, \Delta(n))$ of f. Employing (30), we write $f \geq 0$ and $\mathcal{L}f \leq 0$ on S if and only if there exists $D \in \mathbb{Z}_+$ such that

$$F \in \mathcal{C}(D).$$

Explicitly, observing that $Y_0(D)$ is the vector of the Bernstein moments of the initial measure μ_0, $\langle \mu_0 - \mu, f\rangle \geq 0$ for $f \geq 0$ on S pronounces $\langle Y_0(D) - Y, F\rangle \geq 0$ for all $F \in \mathcal{C}(D)$, where $Y \equiv Y(\mu, D, \Delta(n))$ are the Bernstein moments of μ.

From Lemma 28, μ is a probability measure if and only if its moments are in the standard simplex, $Y(D) \in \Delta(N_D)$, and (33) holds. Combining all the above properties of the moments of μ, we have $Y(D) \in \mathcal{D}(D)$.

We notice that for any probability measure μ

$$\begin{aligned}\mu(U) &= \int_S I_u d\mu = \lim_{k\to\infty} \int_S q_k d\mu \quad (36)\\ &= \lim_{k\to\infty} \sum_{|\alpha|=1} b_\alpha(q_k, D, \Delta(n)) \int_{\Delta(n)} B_\alpha(D,\Delta(n))d\mu \\ &= \lim_{k\to\infty} \langle b(q_k, D, \Delta(n)), Y(\mu, D, \Delta(n))\rangle,\end{aligned}$$

where the second equality follows form Lebesgue's dominated convergence theorem. Let $p^* = \sup_\mu \mu(U)$ subject to the constraints in Theorem 24. For any ϵ there is a measure μ^ϵ such that

$$\mu^\epsilon(U) + \epsilon \geq p^* \geq \mu^\epsilon(U). \quad (37)$$

Furthermore by (36), for any ϵ' there exists N such that for $k > N$

$$|\mu^\epsilon(U) - \langle b(q_k, D_k, \Delta(n)), Y(\mu^\epsilon, D_k, \Delta(n))\rangle| \leq \epsilon', \quad (38)$$

where $D_k \equiv D(q_k)$ is the degree of the polynomial q_k.

From (37) and (38), we have

$$\langle b(q_k, D_k, \Delta(n)), Y(\mu^\epsilon, D_k, \Delta(n))\rangle + \epsilon + \epsilon' \geq p^* \quad (39)$$

and

$$p^* \geq \langle b(q_k, D_k, \Delta(n)), Y(\mu^\epsilon, D_k, \Delta(n))\rangle - \epsilon'. \quad (40)$$

From the discussion in the beginning of the proof, for any $Y(D)$ in $\mathcal{D}(D)$, there is a probability measure μ that satisfies the constraints in Theorem 24. Furthermore, from (39) and (40), we conclude that for any $\epsilon > 0$ and $\epsilon' > 0$ there is $N > 0$ and $Y \in \mathcal{D}(D)$ such that for $k > N$

$$\begin{aligned}\langle b(q_k, D_k, \Delta(n)), Y\rangle + \epsilon + \epsilon' &\geq p^* \\ &\geq \langle b(q_k, D_k, \Delta(n)), Y\rangle - \epsilon'.\end{aligned}$$

Since ϵ and ϵ' can be made arbitrarily small, the conclusion of the theorem follows. □

Theorem 29 comprises an algorithm for the computation of weak p-safety. The core of the algorithm is the linear program (35).

Corollary 30 *Since the polyhedral set* $\mathcal{D}(D,\sigma)$ *is compact and the the optimisation in* (35) *is linear,* (35) *is equivalent to*

$$p_k = \sup_{D\in Z_+} \max_{Y\in V(\mathcal{D}(D))} \langle Y, b(q_k, D, \Delta(n))\rangle,$$

where $V(\mathcal{D}(D))$ *is the set of vertices of the polyhedron* $\mathcal{D}(D)$*.*

Remark 31 *We use Theorem 29 is the following way. We pick a sufficiently good approximation* q_k *of the indicator function* I_U*. We choose a sufficiently large degree* D *then*

$$\langle \mu_0, P\rangle \approx \max_{Y\in V(\mathcal{D}(D))} \langle Y, b(q_k, D, \Delta(n))\rangle.$$

9.1 Numerical Example

The computation of p-safety in Section 8 has been based on an readily available optimisation toolbox. The situation with weak p-safety is more involved as there is no designated toolbox supporting the automatic conversion from the weak safety problem statement to linear programming discussed in Theorem 29 and Remark 31. This will be the subject of our future work. Nonetheless, the next example will illustrate the concept developed in this chapter for one-dimensional Brownian motion. From Example 7, the infinitesimal generator is

$$\mathcal{L}f = \frac{1}{2}f''.$$

We consider the state space $S = [0,1] \subset \mathcal{Y} = \mathbb{R}$. We suppose that the initial measure μ_0 corresponds to uniform distribution on the interval $[0,0.1]$, the forbidden set $U = [0.2,1]$.

To compute weak p-safety (the Matlab code is available on `https://github.com/SecureProject/Safety`). For manipulating polyhedral sets, we have used bensolve toolbox for Matlab.

9.1.1 *Approximation of I_U and Bernstein moments Y_0*

On the interval $[0,1]$, the Bernstein basis of degree D is of the form

$$B_m(x) \equiv B_m(D)(x) = \binom{D}{m} x^m (1-x)^{D-m}.$$

For a real-valued function f defined and bounded on the interval $[0,1]$, let $\hat{B}_D(f)$ be the Bernstein polynomial of degree D that approximates f on $[0,1]$

$$\hat{B}_D(f) := \sum_{m=0}^{D} B_m(x) f\left(\frac{m}{D}\right),$$

and the Bernstein coefficients of $\hat{B}_D(f)$ are $f\left(\frac{m}{D}\right)$.

Therefore, the sequence q_k in Theorem 29 corresponds to the indicator function $I_{[0.2,1]}$,

$$q_k(x) = \sum_{m=0}^{k} \binom{k}{m} x^m (1-x)^{k-m} I_{[0.2,1]}\left(\frac{m}{k}\right).$$

The approximation only makes sense for a large number k; nonetheless, for the sake of illustrating the method, we instantiate the example for $k = 3$,

$$q_3 = x^3,$$

and the vector of Bernstein coefficient is $b(q_3) = \begin{bmatrix} 0 & 1 & 1 & 1 \end{bmatrix}^{\mathrm{T}}$.

Next, we compute the Bernstein moments $Y_0 \equiv Y_0(D)$ of the initial measure μ_0

$$Y_0(m) = \int_0^1 B_m(x)\mu_0(dx) = 10\int_0^{0.1} B_m(x)dx$$

Specifically, for $D = 3$, $Y_0 = \begin{bmatrix} 0.86 & 0.13 & 0.01 & 0.00 \end{bmatrix}^{\mathrm{T}}$.

9.1.2 *Cone $\mathcal{C}(D)$ and its dual $\mathcal{C}(D)^*$*

At the outset, we define

$$f(x) = \sum_{m=0}^{D} F_m B_m(x)$$

and represent

$$\mathcal{L}f(x) = \frac{1}{2}\sum_{m=0}^{D} F_m B''_m(x)$$

in Bernstein basis for $D = 3$,

$$\begin{aligned} \mathcal{L}f(x) = {} & (3F_0 - 6F_1 + 3F_2)x \\ & + (3F_1 - 6F_2 + 3F_3)(1-x). \end{aligned} \tag{41}$$

In (41), we have used that

$$B_m(D)'(x) = D\left(B_{m-1}(D-1)(x) - B_{m-}(D-1)(x))\right).$$

As a consequence, the cone $\mathcal{C}(3)$ in Theorem 29 is

$$\mathcal{C}(3) = \{F \in \mathbb{R}^4 |\ AF \geq 0\}$$

with the matrix A given by

$$A = \begin{bmatrix} I \\ -\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix} \end{bmatrix},$$

where I is 4×4 identity matrix. The dual cone to $\mathcal{C}(3)$ is

$$\mathcal{C}(3)^* = \{Z \in \mathbb{R}^4 |\ BZ \geq 0\},$$

where the matrix B is

$$B = \begin{bmatrix} 3 & 2 & 1 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 0 \end{bmatrix}.$$

The dual cone is computed by changing h-representation of $\mathcal{C}(3)$ given by supporting hyper-planes to v-representation defined by

$$\mathcal{C}(3) = \{B^{\mathrm{T}}Y \mid Y \geq 0\}.$$

and observing that

$$\mathcal{C}(3)^* = \{Z \in \mathbb{R}^4 \mid \langle Z, B^{\mathrm{T}}Y\rangle \geq \ \forall\, Y \geq 0\}$$
$$= \{Z \in \mathbb{R}^4 \mid \langle BZ, Y\rangle \geq 0 \ \forall\, Y \geq 0\} = \{Z \in \mathbb{R}^4 \mid BZ \geq 0\}.$$

9.1.3 Approximation of weak p-safety

We approximate weak p-safety by

$$\max\langle Y, b(q_3)\rangle$$

subject to

$$B(Y_0 - Y) \geq 0,\ 1 \geq Y \geq 0 \text{ and } \mathbb{1}^{\mathrm{T}}Y = 1.$$

The above linear optimisation gives 0.14 for Bernstein moments of the measure μ^*,

$$Y^* = \begin{bmatrix}0.860 & 0.133 & 0.004 & 0.003\end{bmatrix}^{\mathrm{T}}.$$

For the degree $D = 20$, the weak p-safety is approximated by 0.25.

10 Conclusion

The main result of this work is two-fold. Firstly, we have analytically characterized two concepts of safety: p-safety and weak p-safety. The first concept is the probability of hitting a forbidden state before reaching the desired shape when it starts in the specified initial condition. The second notion has randomized the initial state by requiring that the initial state be chosen randomly according to an initial distribution. Secondly, we have translated the theoretical findings to optimization problems. Upon solving them, p-safety and weak p-safety can be calculated. We have provided computational examples for both forms of safety to explain better the methods developed in the paper better and allow the usage of the code for future research.

Our future adventures are to extend our results to the problem of selecting policies such that a process is kept p-safe. We intend to develop a toolbox for weak p-safety. To this end, we need to develop algorithms for efficient triangulation of the state space, develop algorithms aiming the computation on Bernstein forms. Another avenue is devoted to the application of the method for leakage detection in water networks.

Acknowledgements

The work of the first author has been supported by the Poul Due Jensens Fond in the Project SWIft. The second author wishes to acknowledge and thank for the financial support from Maritime Safety Research Center research sponsors DNV-GL and Royal Caribbean Cruise Ltd.

References

[1] Quantitative risk analysis of offshore drilling operations: A bayesian approach. *Safety science*, 57:108–117, 2013.

[2] Alessandro Abate, Maria Prandini, John Lygeros, and Shankar Sastry. Probabilistic reachability and safety for controlled discrete time stochastic hybrid systems. *Automatica (Oxford)*, 44(11):2724–2734, 2008.

[3] L. Arnold. *Stochastic differential equations: theory and applications.* Wiley, 1974.

[4] Nicholas A. Baran, George Yin, and Chao Zhu. Feynman-Kac formula for switching diffusions: connections of systems of partial differential equations and stochastic differential equations. *Adv. Difference Equ.*, pages 2013:315, 13, 2013.

[5] J.-M. Bismut. Potential theory in optimal stopping and alternatinc processes. *Stochastic Control Theory and Stochastic Differential Systems*, 16:285–293, 1979.

[6] R. M. Blumenthal and R. K. Getoor. *Markov processes and potential theory.* Pure and Applied Mathematics, Vol. 29. Academic Press, New York-London, 1968.

[7] L.M. Bujorianu. *Stochastic Reachability Analysis of Hybrid Systems.* Communications and Control Engineering. Springer London, London, 2012.

[8] M. L. Bujorianu and R. Wisniewski. New insights on p-safety of stochastic systems. In *2019 IEEE 58th Conference on Decision and Control (CDC)*, pages 4433–4438, 2019.

[9] K. L. Chung and R. K. Getoor. The condenser problem. *Ann. Probab.*, 5(1):82–86, 02 1977.

[10] M. H. A. Davis. *Markov models and optimization.* Chapman & Hall, 1993.

[11] J. L. Doob. Semimartingales and subharmonic functions. *Trans. Amer. Math. Soc.*, 77:86–121, 1954.

[12] O. Kallenberg. *Foundations of modern probability.* Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[13] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM Journal on Optimization*, 11(3):796–817, March 2001.

[14] J. B. Lasserre. *Moments, positive polynomials and their applications*, volume 1 of *Imperial College Press Optimization Series.* Imperial College Press, London, 2010.

[15] R. Leroy. Certificates of positivity in the simplicial bernstein basis. *hal.archives-ouvertes*, hal-00589945:1–35, 2011.

[16] Tobias Leth. *Polynomials in the Bernstein Basis and Their Use in Stability Analysis.* PhD thesis, 2017. PhD supervisor: Prof. Rafał Wisniewski, Aalborg University Assistant PhD supervisor: Assoc. Prof. Christoffer Sloth, Aalborg University.

[17] Aleksander A Lidtke, Hugh G Lewis, and Roberto Armellin. Impact of high-risk conjunctions on active debris removal target selection. *Advances in space research*, 56(8):1752–1764, 2015.

[18] M. Marshall. *Positive polynomials and sums of squares*, volume 146 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2008.

[19] B. Øksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer-Verlag, 5th edition, 2000.

[20] P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Math. Program.*, 96(2, Ser. B):293–320, 2003. Algebraic and geometric methods in discrete optimization.

[21] S. Prajna, A. Jadbabaie, and G. J. Pappas. A framework for worst-case and stochastic safety verification using barrier certificates. *IEEE Transactions on Automatic Control*, 52(8):1415–1428, Aug 2007.

[22] Stephen Prajna and Anders Rantzer. Convex programs for temporal verification of nonlinear dynamical systems. *SIAM J. Control Optim.*, 46(3):999–1021, 2007.

[23] M.Z. Romdlony and B. Jayawardhana. Stabilization with guaranteed safety using control lyapunovbarrier function. *Automatica*, 66:39–47, 2016.

[24] Th. G. Kurtz S. N. Ethier. *Markov Processes : Characterization and Convergence*. Wiley, New York, N.Y, 2005.

[25] C. Santoyo, M. Dutreix, and S. Coogan. Verification and control for finite-time safety of stochastic systems via barrier functions. In *2019 IEEE Conference on Control Technology and Applications (CCTA)*, pages 712–717, 2019.

[26] B. Schmuland. Dirichlet forms with polynomial domain. *Math. Japon.*, 37(6):1015–1024, 1992.

[27] M. Sharpe. *General theory of Markov processes*, volume 133 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA, 1988.

[28] S. Summers and J. Lygeros. Verification of discrete time stochastic hybrid systems: A stochastic reach-avoid decision problem. *Automatica*, 46(12):1951 – 1961, 2010.

[29] R. Syski. Probabilistic methods in applied mathematics. In A. T. Bharucha-Reid, editor, *Potential Theory for Markov Chains*, pages 214–275. Academic Press, 1973.

[30] Leth T. *Polynomials in the Bernstein Basis and their use in stability theory*. PhD thesis, Aalborg University, 2017.

[31] L. Wang, A. D. Ames, and M. Egerstedt. Safety barrier certificates for collisions-free multirobot systems. *IEEE Transactions on Robotics*, 33(3):661–674, 2017.

[32] T. Watanabe. On the equivalence of excessive functions and superharmonic functions in the theory of Markov processes, I. *Proceedings of the Japan Academy*, 38(7.S1):397 – 401, 1962.

[33] P. Wieland and F. Allgöwer. Constructive safety using control barrier functions. *IFAC Proceedings Volumes*, 40(12):462–467, 2007.

[34] R. Wisniewski and M. L. Bujorianu. Stochastic safety analysis of stochastic hybrid systems. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 2390–2395, 2017.

[35] R. Wisniewski, M. L. Bujorianu, and C. Sloth. p-safe analysis of stochastic hybrid processes. *IEEE Transactions on Automatic Control*, pages 1–16, 2020.