

Piecewise Linear Expectation Analysis via k -Induction for Probabilistic Programs

TENGSHUN YANG, SKLCS, Institute of Software, University of Chinese Academy of Sciences, China
 HONGFEI FU*, Shanghai Jiao Tong University, China
 JINGYU KE, Shanghai Jiao Tong University, China
 NAIJUN ZHAN, SKLCS, Institute of Software, University of Chinese Academy of Sciences, China
 SHIYANG WU, Shanghai Jiao Tong University, China

Quantitative analysis of probabilistic programs aims at deriving tight numerical bounds for probabilistic properties such as expectation and assertion probability, and plays a crucial role in the verification of probabilistic programs. Along this line of research, most existing works consider numerical bounds over the whole state space monolithically and do not consider piecewise bounds. Clearly, monolithic bounds are either conservative, or not expressive and succinct enough in general. To derive more succinct, expressive and precise numerical bounds for probabilistic properties, we propose a novel approach for synthesizing piecewise linear bounds in this work. To this end, we first show how to extract a piecewise feature w.r.t. a given quantitative property from a probabilistic program using latticed k -induction that captures a wide and representative class of piecewise bound functions. Second, we develop an algorithmic approach to synthesize piecewise linear upper and lower bounds from the piecewise feature, for which we show that the synthesis of piecewise linear bounds can be reduced to bilinear programming. Third, we implement our approach with the bilinear programming solver *Gurobi*. The experimental results indicate that our approach is capable of generating tight or even accurate piecewise linear bounds for an extensive set of benchmarks compared with the state of the art.

ACM Reference Format:

Tengshun Yang, Hongfei Fu, Jingyu Ke, Naijun Zhan, and Shiyang Wu. 2024. Piecewise Linear Expectation Analysis via k -Induction for Probabilistic Programs. 1, 1 (March 2024), 45 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 INTRODUCTION

Probabilistic programming [33, 42, 59] is a programming paradigm that extends classical programming languages with statements such as probabilistic branching and sampling. Probabilistic programs provide a powerful model for randomized algorithms [7], machine learning [15], reliability engineering [16], etc. Therefore, analysis of probabilistic programs is becoming increasingly significant, and attracting more and more attention in recent years.

In this work, we consider the formal analysis of probabilistic programs that aims at deriving guaranteed bounds for probabilistic properties. Unlike the formal analysis for classical programs that mostly focuses on qualitative properties such as proving or refuting assertions, the analysis of probabilistic programs concerns quantitative analysis in many situations, such as expected runtime [1, 31, 37, 38], expected resource consumption [51, 60, 63], sensitivity [2], assertion probabilities [21, 58, 62], and so forth.

*The corresponding author

Authors' addresses: Tengshun Yang, SKLCS, Institute of Software, University of Chinese Academy of Sciences, Beijing, China, yangts@ios.ac.cn; Hongfei Fu, Shanghai Jiao Tong University, Shanghai, China, jt002845@sjtu.edu.cn; Jingyu Ke, Shanghai Jiao Tong University, Shanghai, China, windocotber@gmail.com; Naijun Zhan, SKLCS, Institute of Software, University of Chinese Academy of Sciences, Beijing, China, znj@ios.ac.cn; Shiyang Wu, Shanghai Jiao Tong University, Shanghai, China, sunny1231@sjtu.edu.cn.

Similar to classical programs, semantic equations of probabilistic programs are unsolvable in general. Thus, as an alternative, most of the research focuses on computing numerical bounds on quantitative properties instead of solving semantic equations exactly in the quantitative analysis of probabilistic programs. In the literature, various approaches have been proposed to address this issue, including template-based [17, 18, 20, 34], trace-abstraction-based [57], sampling-based [54], etc. Most of these approaches consider to synthesize a monolithic upper- and/or lower-bound for an underlined quantitative property over the whole state space of a probabilistic program of interest. The disadvantages to synthesize a monolithic linear or polynomial function as a bound are two-fold: First, a monolithic bound is either too conservative (e.g., only very coarse linear bounds exist) or not succinct enough (e.g., although tight monolithic polynomial bounds exist, the tightness usually requires complicated polynomials with higher degree). Second, it may be even worse that no monolithic polynomial bounds exist. A simple example is given below to demonstrate the aforementioned disadvantages.

Example 1.1. Consider the following example, which is a simplified version of the GROWING WALK in Beutner et al. [15]:

GROWING WALK: while $(0 \leq x)$ $\{\{x := x + 1; y := y + x\} [0.5] \{x := -1\}\}$

The example models a simple random walk where the step size x is increased by 1 with 50% probability, and set to -1 with the other half probability. The program terminates when x becomes negative. We intend to analyze the expected value of the total distance, i.e., y , after the program terminates. It can be shown that monolithic linear upper-bound cannot be obtained via inductive synthesis with 1-induction, and monolithic polynomial bounds up to degree of 5 are much more conservative than our piecewise linear bound. Our approach synthesizes a piecewise linear upper bound $[x < 0] \cdot y + [x \geq 0] \cdot (x + y + 2)$, which is actually the exact expected value of y (the tightest bound). With our approach, we can also obtain a piecewise linear lower bound $[x < 0] \cdot y + [x \geq 0] \cdot (x + y + 13/8)$. $\square$

Obviously, piecewise bounds are more accurate than monolithic bounds. However, the synthesis of piecewise bounds over probabilistic programs is not well investigated in the literature. To our best knowledge, a handful relevant work is Batz et al. [11] that proposes an approach to compute piecewise bounds for probabilistic loops. Their work requires an upper bound to be verified as an additional program input and it only return a super-invariant (i.e., a piecewise upper bound) that is sufficient to **verify** the input upper bounds through counterexample-guided inductive synthesis (CEGIS). They start the synthesis by predescribing the piecewise template from the syntactic structure of the given loop and perform the template refinement heuristically when existing template admits no admissible solution. Moreover, there is a limitation in this work that requires all assignments in the program must yield a non-negative results. Another work [5] proposes a data-driven approach that can synthesize piecewise (sub-)invariants. However, their approach needs a suitable list of numerical features that requires prior knowledge of the program or user's assistance.

In this work, we consider how to *automatically* synthesize piecewise linear bounds for linear quantitative properties of probabilistic loops with discrete probabilistic choices. Piecewise linear functions are important in the sense that they constitute a basic class of piecewise numerical bounds, and are widely used in hybrid systems [13, 14], floating point approximation [22, 47, 48], bilinear term approximation [6], etc. A recent work [10] also demonstrates that probabilistic program analysis requires piecewise feature.

Challenges and Gaps. To synthesize piecewise linear bounds for probabilistic programs, there are two key challenges:

- Discover a suitable criterion to partition the state space of a probabilistic program to derive the piecewise structure of a bound function. The criterion needs to be automated, since specifying the piecewise structure by hand requires heavy manual expertise.
- Explore efficient algorithms to synthesize piecewise linear bounds.

We attack the first challenge by exploiting latticed k -induction [12, 45]. The k -induction principle is a powerful proof tactics in the verification of hardwares and softwares [25, 26, 43, 56]. Latticed k -induction [12, 45] extends k -induction to lattices and has application in the formal analysis of probabilistic programs [12], where the operators of latticed k -induction are defined by taking pointwise minimum and maximum of bound functions. In this work, we use these k -induction operators to synthesize piecewise linear bounds by extracting the piecewise pattern from the pointwise minimum and maximum operations.

To address the second challenge, we show that the synthesis of piecewise linear bounds can be achieved by a non-trivial application of Optional Stopping Theorem (OST) and bilinear programming. Optional Stopping Theorem [67, Theorem 10.10] is a fundamental result in martingale theory, while bilinear programming is a special non-convex programming that admits efficient constraint solving algorithms [46].

In summary, our main contributions include the following:

- First, we develop a novel variant of latticed k -induction that is a combination of the k -induction operators and OST. Our variant is non-trivial in two folds. The first is that our variant uses OST instead of fixed point theory that is previously adopted in Batz et al. [12], Lu and Xu [45]. This allows to derive both upper and lower bounds for quantitative properties over probabilistic programs, and does not require a global lower bound of program values (such as requiring non-negativity in Batz et al. [11, 12], Lu and Xu [45]). The second is that the variant requires an extended version of OST as proposed in Wang et al. [64], and we show that the classical OST [67, Chapter 10] does not suffice.
- Second, we propose a novel algorithm for synthesizing piecewise linear bounds w.r.t the latticed k -induction conditions. Our algorithm tackles both upper and lower bounds by a reduction to bilinear programming, and therefore can leverage efficient bilinear programming solvers such as *Gurobi* [35]. A key contribution here is that we connect the derived constraints and the unfolding of probabilistic loop, thus reducing the cumbersome calculation containing pointwise minimize/maximize and addition to the concise calculation of the pre-expectation of loop-free programs.
- Third, we implement a prototype of our approach. Experimental results over an extensive set of benchmarks indicate that our approach is capable of generating tight piecewise linear bounds compared with the state of the art.

Limitations. A major limitation is that our approach could only handle affine probabilistic loops with finite discrete probabilistic choices, and linear return functions. Another limitation is that our approach has the combinatorial explosion when the iteration number k in the k -induction increases, which is an inherent problem in latticed k -induction [12, 45]. However, for small k , our approach works efficiently.

2 PRELIMINARIES

We first review some necessary concepts from probability theory, then present the syntax of the probabilistic loops we consider, and finally define the problem we aim to solve in this work.

$$\begin{aligned}
C &::= \text{skip} \mid x := e \mid x \approx \mu \mid C; C \mid \{C\} [p] \mid \{C\} \mid \text{if } (\varphi) \{C\} \text{ else } \{C\} \\
\varphi &::= e < e \mid \neg \varphi \mid \varphi \wedge \varphi \quad e ::= c \mid x \mid c \cdot e \mid e + e \mid e - e
\end{aligned}$$

Fig. 1. Syntax of probabilistic while loop in the form (1)

2.1 Probability Theory and Martingales

Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where Ω is the sample space, $\mathcal{F}$ is a σ -algebra on Ω and $\mathbb{P} : \mathcal{F} \rightarrow [0, 1]$ is a probability measure on the measurable space $(\Omega, \mathcal{F})$. A *random variable* is an $\mathcal{F}$ -measurable function $X : \Omega \rightarrow \mathbb{R} \cup \{+\infty, -\infty\}$, i.e., a function satisfying that for all $d \in \mathbb{R} \cup \{+\infty, -\infty\}$, $\{\omega \in \Omega : X(\omega) \leq d\} \in \mathcal{F}$. The *expectation* of a random variable X , denoted by $\mathbb{E}(X)$, is the Lebesgue integral of X w.r.t. $\mathbb{P}$, i.e., $\mathbb{E}(X) = \int X d\mathbb{P}$. A filtration of the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is an infinite sequence $\{\mathcal{F}_n\}_{n=0}^\infty$ such that for every n , the triple $(\Omega, \mathcal{F}_n, \mathbb{P})$ is a probability space and $\mathcal{F}_n \subseteq \mathcal{F}_{n+1} \subseteq \mathcal{F}$. A stopping time w.r.t. $\{\mathcal{F}_n\}_{n=0}^\infty$ is a random variable $\tau : \Omega \rightarrow \mathbb{N} \cup \{0, \infty\}$ such that for every $n \geq 0$, the event $\{\tau \leq n\} \in \mathcal{F}_n$, i.e., $\{\omega \in \Omega : \tau(\omega) \leq n\} \in \mathcal{F}_n$. Intuitively, τ is interpreted as the time at which the stochastic process shows a desired behavior. A discrete-time stochastic process is a sequence $\Gamma = \{X_n\}_{n=0}^\infty$ of random variables in $(\Omega, \mathcal{F}, \mathbb{P})$. The process Γ is adapted to a filtration $\{\mathcal{F}_n\}_{n=0}^\infty$, if for all $n \geq 0$, X_n is a random variable in $(\Omega, \mathcal{F}_n, \mathbb{P})$. A discrete-time stochastic process $\Gamma = \{X_n\}_{n=0}^\infty$ adapted to a filtration $\{\mathcal{F}_n\}_{n=0}^\infty$ is a *martingale* (resp. supermartingale, submartingale) if for all $n \geq 0$, $\mathbb{E}(|X_n|) < \infty$ and it holds almost surely that $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$ (resp. $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq X_n$, $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \geq X_n$). See Williams [67] for more details about martingale theory. Applying martingales for probabilistic programs analysis is well-studied [17, 18, 21].

2.2 Affine Probabilistic Loops

In this work, we consider affine probabilistic while loops with the form

$$\text{while } (\varphi) \{C\} \tag{1}$$

where φ is the loop guard and C is the loop body without loops. Furthermore, the loop guard φ and loop body C are generated by the grammar in Figure 1,

where x is a program variable taken from a countable set Vars of variables, $c \in \mathbb{R}$ is a real constant, e is an affine arithmetic expression that involves addition and scalar multiplication, φ is a formula over program variables that is a Boolean combination of arithmetic inequalities, and μ is a predefined probability distribution. In this work, we consider μ to be a finite discrete probability distribution (i.e., distributions with a finite support) such as Bernoulli distribution and discrete uniform distribution. The semantics of most statements, including skip, assignment, sequential composition, conditional, and the while statement, follow their standard meaning in imperative programs. The semantics of a statement $\{C_1\}[p]\{C_2\}$ is a probabilistic choice that flips a coin with bias $p \in [0, 1]$ and executes the statement C_1 if the coin yields head, and C_2 otherwise. The semantics of a statement $x \approx \mu$ samples a value according to the predefined distribution μ and assigns the value to the variable x .

To formally specify the semantics of a probabilistic while loop, we define the notion of program states as follows. A *program state* is a function that maps every program variable to a real number. We denote by S the set of program states, i.e., $S \triangleq \{s \mid s : \text{Vars} \rightarrow \mathbb{R}\}$. Especially, the initial state for the probabilistic while loop is denoted by s^* . The evaluation of an arithmetic expression e and a logical formula φ under a program state s , denoted by $e(s)$ and $\varphi(s)$ respectively, is defined in the standard way.

The semantics of a probabilistic while loop of the form (1) can be interpreted by a discrete-time Markov chain, where the state space is the set of all program states S , and the transition probability

function $\mathbf{P}$ is given by the loop body C and determines the probabilities $\mathbf{P}(s, s')$ for $s, s' \in S$ such that each probability value $\mathbf{P}(s, s')$ specifies the probability that given the current program state s , the next program state is s' after one loop iteration. Note that if the loop guard $\varphi(s)$ evaluates to false, then we treat the program state s as a deadlock state so that $\mathbf{P}(s, s) = 1$ and $\mathbf{P}(s, s') = 0$ for $s \neq s'$ (i.e., when the loop terminates at a program state s , we consider that the loop keeps iterating at this state).

Given the Markov chain of a probabilistic while loop as described above, a *path* is an infinite sequence $\pi = s_0, s_1, \dots, s_n, \dots$ of program states such that $\mathbf{P}(s_n, s_{n+1}) > 0$ for all $n \geq 0$. Intuitively, each s_n corresponds to the state right before the $(n + 1)$ -th loop iteration. A program state s is *reachable* from an initial program state s^* if there exists a path $\pi = s_0, s_1, \dots$ such that $s_0 = s^*$ and $s_n = s$ for some $n \geq 0$, and define $\text{Reach}(s^*)$ as the set of reachable states starting from the initial state s^* . By the standard cylinder construction (see e.g. Baier and Katoen [4, Chapter 10]), the Markov chain with a designated initial program state s^* for the probabilistic loop induces a probability space over paths and reachable states, for which we have a countable state space since we consider probabilistic distributions with finite support. We denote the probability measure in this probability space by $\mathbb{P}_{s^*}$ and its related expectation operator by $\mathbb{E}_{s^*}$.

Given a probabilistic while loop P of the form (1), a *return function* f is a function $f : S \rightarrow \mathbb{R}$ that is used to specify the output of the loop P in the sense that when the loop P terminates at a program state s , then the return value of the loop is given as $f(s)$. We denote by X_f the random variable for the return value of the loop given by a return function f . In this work, we study the following problem:

Piecewise Linear Bound Synthesis. Given a probabilistic loop P of the form (1) and a linear return function f , synthesize *piecewise linear upper and lower bounds* on the expected value of X_f after the execution of P .

3 LATTICED k -INDUCTION OPERATORS

In this section, we review the latticed k -induction operators given in Batz et al. [12], Lu and Xu [45], and propose a new operator dual to the one in Batz et al. [12]. As a theoretical contribution, we show that the operators in Batz et al. [12], Lu and Xu [45] are equivalent, so that we can focus on the one in Batz et al. [12] and our dual operator in the development of our approach. Due to the space constraints, all the proofs in this section are relegated to Appendix A.

To present the latticed k -induction operators, we first have a brief review of lattice theory. Informally, a lattice is a partially ordered set $(E, \sqsubseteq)$ (where E is a set and $\sqsubseteq$ is a partial order on E) equipped with two operations, namely the *meet* operation $\sqcap$ and the *join* operation $\sqcup$. Given two elements $u, v \in E$, the meet $u \sqcap v$ is defined as the infimum of $\{u, v\}$ and dually the join $u \sqcup v$ is defined as the supremum of $\{u, v\}$. A partially ordered set $(E, \sqsubseteq)$ is a *lattice* if for any $u, v \in E$, we have that both $u \sqcap v$ and $u \sqcup v$ exist. Given a lattice $(E, \sqsubseteq)$, we say that an operator $\Phi : E \rightarrow E$ is *monotone* if for all $u, v \in E$, $u \sqsubseteq v$ implies $\Phi(u) \sqsubseteq \Phi(v)$. Throughout this section, we fix a lattice $(E, \sqsubseteq)$ and a monotone operator $\Phi : E \rightarrow E$ for $(E, \sqsubseteq)$.

3.1 k -Induction Operators in Batz et al. [12], Lu and Xu [45]

We first present the k -induction operator in Batz et al. [12], which we call the upper k -induction operator. Informally, the upper k -induction operator simply has a meet operation between the application $\Phi(v)$ of Φ to an element $v \in E$ with a designated element $u \in E$.

Definition 3.1 (*The k -Induction Operator in Batz et al. [12]*). Given any element $u \in E$, the upper k -induction operator Ψ_u w.r.t. u and the aforementioned monotone operator Φ is defined by: $\Psi_u : E \rightarrow E, v \mapsto \Phi(v) \sqcap u$.

The intuition is that the operator Ψ_u pulls $\Phi(v)$ down via the meet with u . In Batz et al. [12], a fundamental result shows that this upper k -induction operator gives an extended form of Park induction [52].

THEOREM 3.2 (PARK INDUCTION FROM k -INDUCTION [12]). *For any $u \in E$ and $k \in \mathbb{N}$, we have that $\Phi(\Psi_u^k(u)) \sqsubseteq u \iff \Phi(\Psi_u^k(u)) \sqsubseteq \Psi_u^k(u)$.*

The upper k -induction operator Φ_u in Definition 3.1 directly has the meet with a designated element $u \in E$. A natural tightening of this operator would be to consider the meet with the input element v itself. This variant is investigated in the work [45]. Formally,

Definition 3.3 (*The k -Induction Operator in Lu and Xu [45]*). The upper k -induction operator Ψ is defined by: $\Psi : E \rightarrow E, v \mapsto \Phi(v) \sqcap v$.

Then we prove the following property of the operator Ψ :

PROPOSITION 3.4. *For any $u \in E$, $\Phi(\Psi^k(u)) \sqsubseteq u \iff \Phi(\Psi^k(u)) \sqsubseteq \Psi^k(u)$.*

Furthermore, by the properties of k -induction operators Ψ_u and Ψ , we can prove the following equivalence theorem which states that the iterated applications of the upper k -induction operators Ψ_u and Ψ on u produce the same sequence of elements.

THEOREM 3.5 (EQUIVALENCE BETWEEN Ψ_u AND Ψ). *For any element $u \in E$, the sequence $\{\Psi_u^k(u)\}_{k \geq 0}$ of elements in E coincides with the sequence $\{\Psi^k(u)\}_{k \geq 0}$. Formally, for any natural number $k \geq 0$, we have that $\Psi_u^k(u) = \Psi^k(u)$.*

From Theorem 3.2, Proposition 3.4 and Theorem 3.5, we have that

$$\Phi(\Psi_u^k(u)) \sqsubseteq u \iff \Phi(\Psi_u^k(u)) \sqsubseteq \Psi_u^k(u) \iff \Phi(\Psi^k(u)) \sqsubseteq \Psi^k(u) \iff \Phi(\Psi^k(u)) \sqsubseteq u.$$

3.2 Dual k -Induction Operators

Below we propose a dual version for the k -induction operator in Batz et al. [12]. The intuition is simply to replace the meet operation in Definition 3.1 with join.

Definition 3.6 (*Dual k -Induction Operator*). Let $u \in E$. The dual k -induction operator Ψ'_u w.r.t. u and the aforementioned monotone operator Φ is defined by: $\Psi'_u : E \rightarrow E, v \mapsto \Phi(v) \sqcup u$.

We call the operator Ψ'_u as the lower k -induction operator. Similar to the case of the upper k -induction operator, one can define a variant that performs the join operation between the function application $\Phi(v)$ with the input element v itself. This variant was examined in Lu and Xu [45], and is restated as follows.

Definition 3.7 (*Dual k -Induction Operator in Lu and Xu [45]*). The lower k -induction operator Ψ' is given by: $\Psi' : E \rightarrow E, v \mapsto \Phi(v) \sqcup v$.

We prove that both of these two lower k -induction operators have the following properties and their equivalence.

PROPOSITION 3.8. *For any element $u \in E$, the lower k -induction operators Ψ'_u and Ψ' have the following properties:*

$$\begin{aligned} \Phi((\Psi'_u)^k(u)) \sqsupseteq u &\iff \Phi((\Psi'_u)^k(u)) \sqsupseteq (\Psi'_u)^k(u) \\ \Phi((\Psi')^k(u)) \sqsupseteq u &\iff \Phi((\Psi')^k(u)) \sqsupseteq (\Psi')^k(u) \end{aligned}$$

THEOREM 3.9 (EQUIVALENCE BETWEEN Ψ'_u AND Ψ'). *For any element $u \in E$, we have that the sequence $\{(\Psi'_u)^k(u)\}_{k \geq 0}$ of elements in E coincides with the sequence $\{(\Psi')^k(u)\}_{k \geq 0}$. Formally, for any natural number $k \geq 0$, we have that $(\Psi'_u)^k(u) = (\Psi')^k(u)$.*

Similarly, from Proposition 3.8 and Theorem 3.9, we have that

$$\Phi((\Psi'_u)^k(u)) \sqsupseteq u \Leftrightarrow \Phi((\Psi'_u)^k(u)) \sqsupseteq (\Psi'_u)^k(u) \Leftrightarrow \Phi((\Psi')^k(u)) \sqsupseteq (\Psi')^k(u) \Leftrightarrow \Phi((\Psi')^k(u)) \sqsupseteq u.$$

4 PIECEWISE BOUNDS VIA LATTICED k -INDUCTION

In this section, we show how one can apply k -induction operators in Section 3 to synthesize piecewise expectation bounds for probabilistic loops and their soundness. We first introduce expectation functions over which we construct concrete k -induction operators, then define upper and lower potential functions, and show their soundness to derive expectation bounds from the properties of k -induction operators and an extended Optional Stopping Theorem [64]. Throughout this section, we fix a probabilistic while loop $P = \mathbf{while}(\varphi)\{C\}$ in the form of (1) and a return function f .

4.1 Expectation Functions

Definition 4.1 (Expectation Functions). An *expectation function* is a function $h : S \rightarrow \mathbb{R}$ that assigns to each reachable program state a real value. The partial order $\leq$ over expectation functions is defined in the pointwise fashion, i.e., $h_1 \leq h_2 \iff \forall s \in S, h_1(s) \leq h_2(s)$. We denote the set of expectation functions by $\mathcal{E}$ and the lattice by $(\mathcal{E}, \leq)$, for which the meet operation $\sqcap$ in the lattice is given by $h_1 \sqcap h_2 := \min\{h_1, h_2\}$, where $\min$ is the pointwise minimum on functions, i.e., $\forall s \in S, \min\{h_1, h_2\}(s) = \min\{h_1(s), h_2(s)\}$, and the join operation $\sqcup$ is given by $h_1 \sqcup h_2 := \max\{h_1, h_2\}$, where $\max$ is the pointwise maximum.

Informally, an expectation function h is that for each reachable program state $s \in S$, the value $h(s)$ is the expected value of return function f after the execution of the while loop P when the loop starts with the program state s . Note that although infinite expected values (i.e., $\infty, -\infty$) theoretically exist, in this work we consider only finite expected values.

It is straightforward to observe that the partially ordered set $(\mathcal{E}, \leq)$ with the meet and join operations defined above is a lattice, but our approach does not rely on the key properties (e.g., fixed point, Park induction, etc.) of lattices. We present expectation functions in terms of lattice only to relate them with the k -induction operators presented in Section 3.

To instantiate the k -induction operators for expectation functions, we further construct the monotone operator for the lattice $(\mathcal{E}, \leq)$. To this end, we first define the notion of pre-expectation as follows, wherein $[\varphi]$ denotes the Iverson-bracket of φ , i.e., $[\varphi](s)$ evaluates to 1 if $s \models \varphi$ and to 0 otherwise. Notice that the random assignment command $x \approx \mu$ (where μ is a discrete distribution) can be written in an iterative style of $\{C_1\} [p] \{C_2\}$.

Definition 4.2 (Pre-expectation [17, 63]). Given an expectation function $h : S \rightarrow \mathbb{R}$. We define its *pre-expectation* across a loop-free program C , $pre_C(h) : S \rightarrow \mathbb{R}$, recursively on the structure of C :

- $pre_C(h) := h$, if $C \equiv \text{skip}$.
- $pre_C(h) := h[x/e]$, if $C \equiv x := e$. Here the substitution $h[x/e]$ is given by $h[x/e](s) = h(s[x/e])$ for any $s \in S$, where $s[x/e](x) = e(s)$ and $s[x/e](y) = s(y)$ for all $y \in \text{Vars} \setminus \{x\}$.
- $pre_C(h) := pre_{C_1}(pre_{C_2}(h))$, if $C \equiv C_1; C_2$.
- $pre_C(h) := p \cdot pre_{C_1}(h) + (1 - p) \cdot pre_{C_2}(h)$, if $C \equiv \{C_1\} [p] \{C_2\}$.
- $pre_C(h) := [\phi] \cdot pre_{C_1}(h) + [\neg\phi] \cdot pre_{C_2}(h)$, if $C \equiv \text{if } (\phi) \{C_1\} \text{ else } \{C_2\}$.

The intuition of pre-expectation is that given an expectation function h , the pre-expectation pre_C computes the expected value $pre_C(h)$ of h after the execution of the command C . With pre-expectation, we then define the monotone operator to be the characteristic function $\bar{\Phi}_f$ of the probabilistic loop P with respect to the return function f as follows.

Definition 4.3 (Characteristic Function [17, 37]). The characteristic function $\bar{\Phi}_f : \mathcal{E} \rightarrow \mathcal{E}$ is defined by $\bar{\Phi}_f(h) := [\neg\varphi] \cdot f + [\varphi] \cdot pre_C(h)$. The monotone operator for the lattice $(\mathcal{E}, \leq)$ is defined as the characteristic function $\bar{\Phi}_f$.

Informally, the characteristic function $\bar{\Phi}_f$ outputs f if the loop guard φ is violated and the loop terminates in the next step, and the pre-expectation of h w.r.t. the loop body C otherwise. It is straightforward to verify the monotonicity of $\bar{\Phi}_f$. In the following, We omit the subscript f in $\bar{\Phi}_f$ if it is clear from the context.

Given the monotone operator, we establish the concrete k -induction operators as follows. Recall that by Theorem 3.5 and Theorem 3.9, we only need to consider the k -induction operators in Definition 3.1 and Definition 3.6.

Definition 4.4 (k -Induction Operators for $(\mathcal{E}, \leq)$). The upper k -induction operator $\bar{\Psi}_h$ is defined by $\bar{\Psi}_h : \mathcal{E} \rightarrow \mathcal{E}, g \mapsto \min\{\bar{\Phi}(g), h\}$, and the lower k -induction operator $\bar{\Psi}'_h$ is given by $\bar{\Psi}'_h : \mathcal{E} \rightarrow \mathcal{E}, g \mapsto \max\{\bar{\Phi}(g), h\}$.

4.2 Potential Functions

We define potential functions as expectation functions that fulfill k -induction conditions. For deriving upper bounds of probabilistic loops, we introduce upper potential functions. For lower bounds, we have lower potential functions.

Definition 4.5 (Potential Functions). Let k be a positive integer. A k -upper potential function is an expectation function h that satisfies the upper k -induction condition $\bar{\Phi}_f(\bar{\Psi}_h^{k-1}(h)) \leq h$, and a k -lower potential function is an expectation function h that satisfies the lower k -induction condition $\bar{\Phi}_f((\bar{\Psi}'_h)^{k-1}(h)) \geq h$.

Below we explore the soundness of potential functions. Since the classical Optional Stopping Theorem [27, 67] (see Appendix B.1) requires bounded changes of the step-wise difference $|X_{n+1} - X_n|$ in a stochastic process $\{X_n\}_{n \geq 0}$, which cannot handle our problem due to the assignment command in the loop body. To address this difficulty, We have sought several extended versions of OST, as proposed in Wang et al. [61, 63, 64], etc. Among which we find the OST variant proposed in Wang et al. [64] can handle our problem. We have copied its proof in Appendix B.2.

THEOREM 4.6 (EXTENDED OST [64]). Let $\{X_n\}_{n=0}^\infty$ be a supermartingale adapted to a filtration $\mathcal{F} = \{\mathcal{F}_n\}_{n=0}^\infty$ and τ be a stopping time w.r.t the filtration $\mathcal{F}$. Suppose there exist positive real numbers b_1, b_2, c_1, c_2, c_3 such that $c_2 > c_3$ and

- (a) Concentration property. For all sufficiently large natural numbers n , it holds that $\mathbb{P}(\tau > n) \leq c_1 \cdot e^{-c_2 \cdot n}$.
- (b) Exponential bound. For every natural number $n \geq 0$, it holds almost-surely that $|X_{n+1} - X_n| \leq b_1 \cdot n^{b_2} \cdot e^{c_3 \cdot n}$.

Then we have that $\mathbb{E}(|X_\tau|) < \infty$ and $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$.

To handle the prerequisite (a) in Theorem 4.6, we consider to fulfill the concentration property by synthesizing difference bounded ranking supermartingales to witness the exponentially-decreasing concentration property (see Chatterjee et al. [18, 19] for details). For the prerequisite (b), we choose a maximum amplifier $c > 0$ such that the absolute value of every program variable x is amplified

by at most c during one loop iteration. The value c can always be taken since the Vars of program variables and the commands in one loop iteration is finite. We guarantee an exponential bound for the stepwise difference by restricting the amplifier c in the loop body with a constant, i.e., $c \leq e^{c_3}$, for some positive value c_3 . The amplifier c can be obtained by simple syntactic check for the loop.

Putting the above properties together, we have our main theorem below. We follow the definitions and notations in Section 2. Recall that we denote s^* as an initial program state, X_f as the random variable for the return value of the loop given a return function f . Furthermore, we denote τ as the random variable of the termination time of the loop, i.e., the number of loop iterations. The theorem states that upper and lower bounds for the expected value of a return function can be derived from upper and lower potential functions.

THEOREM 4.7 (SOUNDNESS OF POTENTIAL FUNCTIONS). *Let k be a positive integer. Suppose that there exist real numbers $c_1 > 0$ and $c_2 > c_3 > 0$ such that (i) The maximum amplifier c satisfies $c \leq e^{c_3}$ and (ii) the termination time random variable τ of P has the concentration property, i.e., $\mathbb{P}(\tau > n) \leq c_1 \cdot e^{-c_2 \cdot n}$. Then the following hold:*

- For any k -upper potential function h , $\mathbb{E}_{s^*}(X_f) \leq \bar{\Psi}_h^{k-1}(h)(s^*) \leq h(s^*)$.
- For any k -lower potential function h , $\mathbb{E}_{s^*}(X_f) \geq (\bar{\Psi}'_h)^{k-1}(h)(s^*) \geq h(s^*)$.

PROOF. We first proof the soundness of upper potential functions. Let s_n be the random vector (random variable) of the program state at the n -th iteration of the probabilistic while loop P , where $s_0 = s^*$ and let $H = \bar{\Psi}_h^{k-1}(h)$. By Definition 4.5 and Theorem 3.2, we obtain that $\forall s \in \text{Reach}(s^*)$, $\bar{\Phi}(H)(s) \leq H(s)$. We define the stochastic process $\{X_n\}_{n=0}^\infty$ by

$$X_n := [s_n \models \varphi] \cdot H(s_n) + [s_n \not\models \varphi] \cdot f(s_n).$$

We first prove that the stochastic process $\{X_n\}$ is a supermartingale. We discuss this in the following two scenarios:

- if $s_n \not\models \varphi$, by the semantics of probabilistic while loop (see Section 2.2), $s_{n+1} = s_n$, and thus $X_{n+1} = X_n$, which satisfies the conditions of supermartingale;
- if $s_n \models \varphi$, we have

$$\begin{aligned} \mathbb{E}_{s^*}[X_{n+1} | \mathcal{F}_n] &= \mathbb{E}_{s^*}[[s_{n+1} \models \varphi] \cdot H(s_{n+1}) + [s_{n+1} \not\models \varphi] \cdot f(s_{n+1})] \\ &= [s_{n+1} \models \varphi] \cdot \mathbb{E}_{s^*}[H(s_{n+1})] + [s_{n+1} \not\models \varphi] \cdot \mathbb{E}_{s^*}[f(s_{n+1})] \\ &\quad \text{(by definition of mathematic expectation)} \\ &= [s_{n+1} \models \varphi] \cdot \text{pre}_C(H)(s_n) + [s_{n+1} \not\models \varphi] \cdot \mathbb{E}_{s^*}[f(s_{n+1})] \\ &\quad \text{(by definition of pre-expectation)} \\ &= \bar{\Phi}(H)(s_n) \\ &\leq H(s_n) \quad \text{(by property of } H) \\ &= X_n \end{aligned}$$

Combining the condition (i), (ii) and H is piecewise linear, we can derive that $\mathbb{E}_{s^*}[X_n] < \infty$ holds. Thus $\{X_n\}$ is a supermartingale.

The condition (a) in Theorem 4.6 depends on the assumption that (ii) P has the concentration property.

Then we prove the condition (b) in Theorem 4.6. We discuss this in the following three scenarios:

- if $s_n \not\models \varphi$, by the semantics of probabilistic while loop (see Section 2.2), we have $s_{n+1} = s_n$, and thus $|X_{n+1} - X_n| = 0$;

- if $s_n \models \varphi$ and if $s_{n+1} \models \varphi$, since the condition (ii), we have that each program variable $x_i (i \in \mathbb{Z}_+)$ and the constant term x_0 at state $s_n (\forall n)$ can be bounded by $K_i \cdot c_i^n$ for some $K_i, c_i (i \in \mathbb{N})$. In addition that H is piecewise linear, we have that $H(s_n) \leq M_n \cdot c^n$ for $M_n > 0$.

$$\begin{aligned}
|X_{n+1} - X_n| &= |H(s_{n+1}) - H(s_n)| \\
&\leq |H(s_{n+1})| + |H(s_n)| \\
&\leq M_n \cdot |c|^n + M_{n+1} \cdot |c|^{n+1} \\
&\leq (M_n + |c| \cdot M_{n+1}) \cdot |c|^n \\
&\leq b_1 \cdot e^{c_3 n}
\end{aligned}$$

- if $s_n \models \varphi$ and if $s_{n+1} \not\models \varphi$, this case is similar with the case of $s_n \models \varphi$ & $s_{n+1} \models \varphi$. We have that

$$\begin{aligned}
|X_{n+1} - X_n| &= |f(s_{n+1}) - H(s_n)| \\
&\leq |f(s_{n+1})| + |H(s_n)| \\
&\leq M_n \cdot |c|^n + M_{n+1} \\
&= M_n \cdot |c|^n + M_{n+1} \cdot e^{0 \cdot n} \\
&\leq b_1 \cdot e^{c_3 n}
\end{aligned}$$

By applying Theorem 4.6, we have that $\mathbb{E}_{s^*}(X_\tau) \leq \mathbb{E}_{s^*}(X_0)$. Since τ is the stopping time, there will be $s_\tau \not\models \varphi$, thus $X_\tau = f(s_\tau) = X_f$. We have $\mathbb{E}_{s^*}(X_f) \leq \mathbb{E}_{s^*}(X_0) = H(s^*)$. The second inequality can be derived directly from the property that $\bar{\Psi}_h^{k-1}(h) \leq h$ holds (see Appendix A.1 and Batz et al. [12]).

The the case of lower potential functions is completely dual to the case of upper potential functions since we can consider the stochastic process $\{-X_n\}$, that is, define the stochastic process by

$$Y_n := [s_n \models \varphi] \cdot (-H(s_n)) + [s_n \not\models \varphi] \cdot (-f(s_n)).$$

The remaining proof is essentially the same. □

5 SYNTHESIZING BOUNDS

In this section, we propose synthesis algorithms following the theoretical results in Section 4. Given a probabilistic while loop P in the form of (1), a linear return function f and a positive integer k (as the parameter in k -induction), our algorithms synthesize piecewise linear upper/lower bounds for the expected value of the return function f with the k -induction conditions. We only present the synthesis of upper bounds, and the case of lower bound is completely dual, as we just need to replace minimum with maximum and $\leq$ with $\geq$.

5.1 A Nutshell of Our Algorithm

Our algorithm consists of the following steps:

- Set up a linear template h with unknown coefficients over program variables. The template h acts as the expectation function that is used to derive the piecewise linear bound $\bar{\Psi}_h^{k-1}(h)$ as in Definition 4.5 and Theorem 4.7.
- Apply the k -induction conditions to obtain the constraints on the unknown coefficients in the template h . For example, applying the upper k -induction condition from Definition 4.5 yields the constraint $\bar{\Phi}_f(\bar{\Psi}_h^{k-1}(h)) \leq h$.
- Transform the constraints from the previous step into a succinct canonical form.

- Solve the constraints in the canonical form from the previous step. This include applying Motzkin's Transposition Theorem to reduce the constraints to bilinear programming and the call of the Gurobi bilinear programming solver.

After solving the canonical constraints, we obtain an instantiation of the unknown coefficients in the template h , and therefore an instantiated expectation function h^* . Thus, we obtain a piecewise linear upper bound $h_{up}^* = \bar{\Psi}_{h^*}^{k-1}(h^*)$.

5.2 Zooming of Our Algorithm

Below we explain the details of our algorithm with the running example in Example 1.1. Consider as input a probabilistic loop P in the form of (1) with an initial program state s^* , a linear return function f and a positive integer k . We assume an affine invariant I at the entry point of the loop, which is a conjunctive linear arithmetic formula over program variables that over-approximates the reachable program states $Reach(s^*)$, i.e., for $s \in Reach(s^*)$, $I(s)$ evaluates to **true** (or $s \models I$). We restrict our attention to program states satisfying the invariant I by replacing the definition for $\leq$ (which is $h_1 \leq h_2 \iff \forall s \in S, h_1(s) \leq h_2(s)$) to $h_1 \leq h_2 \iff \forall s \models I, h_1(s) \leq h_2(s)$, whose soundness follows from the over-approximation of reachable program states by the invariant I . Affine invariants can be obtained via external invariant generators (such as Sankaranarayanan et al. [55]).

Example 5.1. Consider the GROWING WALK in Example 1.1 as input. We take the invariant $I = [-1 \leq x]$ and set $k = 2$. $\square$

Our algorithm includes four steps as follows.

Step 1. Predefining a Linear Template h . Our algorithm establishes a linear template as $h = \mathbf{c}^T \cdot \mathbf{x} + d$ where $\mathbf{x}$ is the vector of program variables, $\mathbf{c}$ is the vector of unknown coefficients and d is the unknown scalar term of the template.

Example 5.2. In the first step, we predefine a linear template $h = a \cdot x + b \cdot y + c$ for the probabilistic loop in Example 5.1, where a, b, c are unknown coefficients. $\square$

Step 2. Deriving Constraints. The second step is to apply the k -induction conditions to the linear template h , and obtain the constraint $\bar{\Phi}_f(\bar{\Psi}_h^{k-1}(h)) \leq h$. To transform the constraint into a simpler form, our algorithm further unrolls the k -induction conditions so that the minimum operations appear at the outermost of the left-hand-side of the inequality. In detail, from the definition of the operator $\bar{\Psi}_h$ (Definition 4.4), the unrolling is reduced to the recursive computation of *pre-expectation* and the pointwise minimum operation. Following the definition of pre-expectation, the unrolling can be done by the following reduction rules for functions $f_1, \dots, f_m, g_1, \dots, g_n$:

- (R1) $\min\{f_1, \dots, f_m\} + \min\{g_1, \dots, g_n\} = \min_{1 \leq i \leq m, 1 \leq j \leq n} \{f_i + g_j\}$;
- (R2) $c \cdot \min\{f_1, \dots, f_m\} = \min\{c \cdot f_1, \dots, c \cdot f_m\}$ for constant $c \geq 0$;
- (R3) $[B] \cdot \min\{f_1, \dots, f_m\} = \min\{[B] \cdot f_1, \dots, [B] \cdot f_m\}$ for predicate B .

By iterative applications of the reduction rules, the constraint $\bar{\Phi}_f(\bar{\Psi}_h^{k-1}(h)) \leq h$ can be transformed into a succinct form with only one minimum operation:

$$\min\{h_1, h_2, \dots, h_m\} \leq h \quad (2)$$

where h is the linear template and each h_i ($i = 1, \dots, m$) is a piecewise expression derived from the unrolling that does not contain the minimum operation.

Instead of directly applying the rules (R1) – (R3), our algorithm employs a more efficient approach to obtain the constraint in the form of (2). We start by unfolding the probabilistic loop from one

arbitrary initial state s^* for at most k times (the choice of this state is unimportant). We make a decision whether we continue to unfold the loop once more at each state we reach (except for the initial state), and this decision process continues until there are no unfolding to be executed or we have already unfolded the loop from the initial state s^* for k times. Each strategy, composed of the choices at each decision step, decides a distinct loop-free program C_d . Let $C_1, \dots, C_m$ be all the loop-free programs generated by the above decision process.

We explore the relationship between the upper (resp. lower) k -induction constraint and the unfolding of the probabilistic loop. We describe it in Proposition 5.3 and implement our automated algorithm based on this proposition. Due to the space limit, the proof is relegated to Appendix C.1.

PROPOSITION 5.3. *The upper (resp. lower) k -induction condition $\bar{\Phi}_f(\bar{\Psi}_h^{k-1}(h)) \leq h$ (resp. $\bar{\Phi}_f(\bar{\Psi}'_h)^{k-1}(h) \geq h$) is equivalent with $\min\{h_1, h_2, \dots, h_m\} \leq h$ (resp. $\max\{h_1, h_2, \dots, h_m\} \geq h$), where each h_i uniquely corresponds to one $C_d \in \{C_1, \dots, C_m\}$ and is equal to $\text{pre}_{C_d}(h)$.*

We demonstrate our Proposition 5.3 through a pedagogical example and a dendrogram. We first show the upper 2-induction constraints considered in the example simplified by symbolic calculation, and then show the relations with the loop unrolling.

Example 5.4. : We consider a simple but general example P_a :

$$\text{while}(\varphi(x)) \{ \{x := a_1x + b_1\} [p] \{x := a_2x + b_2\} \} \quad (3)$$

where $a_i, b_i (i = 1, 2) \in \mathbb{R}$, $p \in [0, 1]$ and $\varphi(x)$ is a linear inequality. We consider to calculate the upper bound for the expected value of f after executing the probabilistic while loop. $\square$

We consider to derive the upper bound for the expected value of f from upper 2-induction condition:

$$\bar{\Phi}_f(\Psi_h(h)) \leq h \quad (4)$$

where $\Psi_h(h) = \min\{\bar{\Phi}_f(h(x)), h(x)\}$, from the Definition 4.4, and the characteristic function $\bar{\Phi}_f$ in this while loop P_a is $[\neg\varphi(x)] \cdot f(x) + [\varphi(x)](p \cdot h(a_1x + b_1) + (1 - p) \cdot h(a_2x + b_2))$.

By unfolding the constraint (4), the left side of the inequality is in the form of $\min\{h_1, h_2, h_3, h_4\}$. In the following, we show all the expressions h_i and illustrate that each $h_i (i = 1, 2, 3, 4)$ represents for the pre-expectation of h across the loop-free programs $C_i \in \{C_1, C_2, C_3, C_4\}$ generated by the above unfolding of the loop (3) within 2 iterations. For ease of understanding, we take $H_1 = [\neg\varphi(a_1x + b_1)] \cdot f(a_1x + b_1) + [\varphi(a_1x + b_1)] \cdot (p \cdot h(a_1(a_1x + b_1) + b_1) + (1 - p) \cdot h(a_2(a_1x + b_1) + b_2))$, which represents for the loop unrolling for once at the state of $a_1x + b_1$, and $H_2 = [\neg\varphi(a_2x + b_2)] \cdot f(a_2x + b_2) + [\varphi(a_2x + b_2)] \cdot (p \cdot h(a_1(a_2x + b_2) + b_1) + (1 - p) \cdot h(a_2(a_2x + b_2) + b_2))$, which represents for the loop unrolling for once at the state of $a_2x + b_2$. Note that if $\varphi(x)$ is not satisfied initially, the loop will terminate and output $f(x)$.

- $h_1 = [\neg\varphi(x)] \cdot f(x) + [\varphi(x)](p \cdot h(a_1x + b_1) + (1 - p) \cdot h(a_2x + b_2))$: unfold the loop for once and will reach two states: $a_1x + b_1$ and $a_2x + b_2$ (if the loop guard $\varphi(x)$ is satisfied), and we do not continue the execution of the loop, i.e., we unfold the loop only for once and obtain a loop-free program C_1 . Notice that h_1 is the pre-expectation of $h(x)$ w.r.t. the program C_1 .
- $h_2 = [\neg\varphi(x)] \cdot f(x) + [\varphi(x)] \cdot (p \cdot h(a_1x + b_1) + (1 - p) \cdot H_2)$: execute the loop once and will reach two states of $a_1x + b_1$ and $a_2x + b_2$ (if the loop guard $\varphi(x)$ is satisfied) respectively.
 - At the state of $a_1x + b_1$ (with the probability p), we do not unfold the loop and obtain the value $h(a_1x + b_1)$;
 - At the state of $a_2x + b_2$ (with the probability $1 - p$), we continue the execution of the loop and we will attain two branches:

- (1) if $\varphi(a_2x + b_2)$ is not satisfied, output the return function $f(a_2x + b_2)$;
- (2) if $\varphi(a_2x + b_2)$ is satisfied, we will arrive at the state of $a_1(a_2x + b_2) + b_1$ with the probability of p and arrive at the state of $a_2(a_2x + b_2) + b_2$ with the probability of $1 - p$, i.e. unfold the loop for twice.

The unfolding process above generates a loop-free program C_2 . h_2 is the pre-expectation of $h(x)$ w.r.t. the program C_2 .

- $h_3 = [\neg\varphi(x)] \cdot f(x) + [\varphi(x)] \cdot (p \cdot H_1 + (1 - p) \cdot h(a_2x + b_2))$: This case is similar with the former one, the only difference is that we choose to continue the execution of the loop at the state of $a_1x + b_1$ and do not unfold the loop at $a_2x + b_2$:
 - At the state of $a_1x + b_1$ (with the probability p), we continue the execution of the loop and we will attain two branches:
 - (1) if $\varphi(a_1x + b_1)$ is not satisfied, output the return function $f(a_1x + b_1)$;
 - (2) if $\varphi(a_1x + b_1)$ is satisfied, we will arrive at the state of $a_1(a_1x + b_1) + b_1$ with the probability of p and arrive at the state of $a_2(a_1x + b_1) + b_2$ with the probability of $1 - p$, i.e. unfold the loop for twice.
 - At the state of $a_2x + b_2$ (with the probability $1 - p$), we do not unfold the loop and obtain the value $h(a_2x + b_2)$.

The above process generates a loop-free program C_3 and h_3 is the pre-expectation of $h(x)$ w.r.t. the program C_3 .

- $h_4 = [\neg\varphi(x)] \cdot f(x) + [\varphi(x)] \cdot (p \cdot H_1 + (1 - p) \cdot H_2)$: At both the states $a_1x + b_1$ and $a_2x + b_2$, we choose to execute the loop for one more time, corresponding to the complete denerogram in figure 2, i.e., we choose to unfold the loop twice at these two states. This generates a loop-free program C_4 and h_4 is the pre-expectation of $h(x)$ w.r.t. the program C_4 .

Recap the constraint of upper 2-induction condition, we only can execute the loop for almost 2 times. Thus here are all the potential cases, which corresponds to each h_i in the set $\{h_1, h_2, h_3, h_4\}$. A denerogram Fig. 2 is attached for this Example 5.4. Moreover, to illustrate the intuition more completely, we supplement the explanation about the upper k -induction condition when $k > 2$. See a simplified dendrogram in the Figure 3 in Appendix C.2 for all the potential cases derived from upper 3-induction condition.

Based on Proposition 5.3, an alternative approach to extract the constraint is by exploring all the unfolded programs within k loop iterations. Notice that the predicates in $pre_{C_d}(h)$ (for each $C_d \in \{C_1, \dots, C_m\}$) are completely determined by the assignments and conditional branches in the loop and the loop guard, while the linear expressions are completely determined by the loop body and the loop iterations. Since these programs $\{C_1, \dots, C_m\}$ are loop-free and structurally similar, we can calculate $pre_{C_d}(h)$ (for each $C_d \in \{C_1, \dots, C_m\}$) at the same time by traversing k -unfolding of program loop once, which reduces the runtime by avoiding excessive and repeatable computations.

Example 5.5. In the GROWING WALK example, we apply 2-induction to synthesize a piecewise linear upper bound. Starting from arbitrary program state $s^* = (x, y)$, we unroll the loop once and arrives at two new program states $(x + 1, x + y + 1)$ and $(-1, y)$. At each new state, we perform the decision process above independently and obtain the following four functions from the template h :

$$\begin{aligned}
 h_1 &= [x < 0] \cdot y + [x \geq 0] \cdot (0.5 \cdot h(x + 1, x + y + 1) + 0.5 \cdot h(-1, y)) \\
 h_2 &= [x < 0] \cdot y + [x \geq 0] \cdot (0.25 \cdot h(-1, y + x) + 0.25 \cdot h(x + 2, 2x + y + 3) + 0.5 \cdot h(-1, y)) \\
 h_3 &= [x < 0] \cdot y + [x \geq 0] \cdot (0.25 \cdot h(-1, y + x) + 0.25 \cdot h(x + 2, 2x + y + 3) + 0.5 \cdot y) \\
 h_4 &= [x < 0] \cdot y + [x \geq 0] \cdot (0.5 \cdot h(x + 1, x + y + 1) + 0.5 \cdot y)
 \end{aligned} \tag{5}$$

Thus, we have the constraint $\forall(x, y) \models I. \min\{h_1, h_2, h_3, h_4\} \leq h$. □

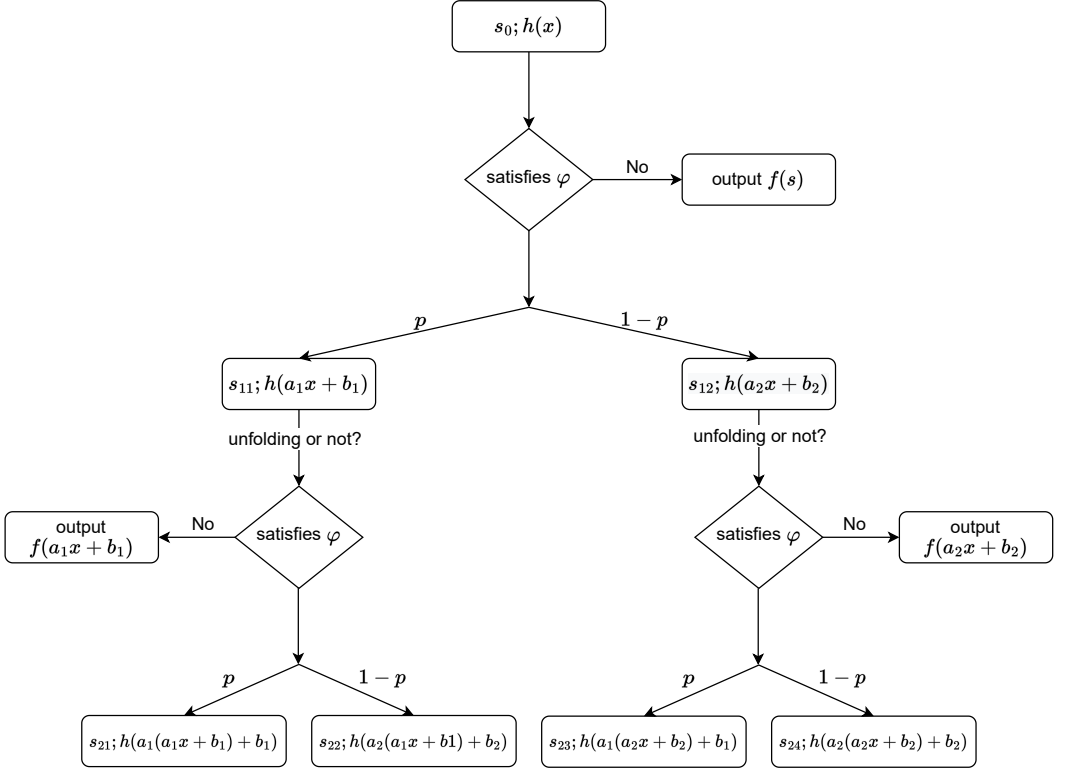


Fig. 2. 2-induction condition for example 5.4

Step 3. Transforming to Canonical Form. In the third step, our algorithm transforms the constraint of the form (2) to a canonical form given by

$$[B_1] \implies \min\{e_{11}, \dots, e_{m1}\} \leq h, \dots, [B_l] \implies \min\{e_{1l}, \dots, e_{ml}\} \leq h \quad (6)$$

where h is the linear template, each $B_j (j \in \{1, \dots, l\})$ is a conjunction of predicates in the program variables that does not contain the unknown coefficients of the template, and each e_{ij} is a monolithic linear expression with the unknown coefficients from the template. The transformation first rewrites the inequality (2) as

$$\min\{\sum_k [B_{1k}] \cdot e_{1k}, \dots, \sum_k [B_{mk}] \cdot e_{mk}\} \leq h \quad (7)$$

where for each $1 \leq i \leq m$, we have that $h_i = \sum_k [B_{ik}] \cdot e_{ik}$ for which each B_{ik} is a predicate without the unknown coefficients and each e_{ik} is a monolithic linear expression with the unknown coefficients. Then, for each conjunctive Boolean combination $B = \bigwedge_{i=1}^m B_{ik_i}$ where each B_{ik_i} is a predicate from the summation $\sum_k [B_{ik}] \cdot e_{ik}$, we have a constraint Ψ_B that $[B] \implies \min_{i=1}^m e_{ik_i} \leq h$, and the transformed inequalities (6) include exactly all those Ψ_B 's. In our algorithm, we remove infeasible Ψ_B (i.e., B is unsatisfiable) by a SMT solver such as Z3 [24] whenever possible.

Example 5.6. Continue with Example 5.2 and Example 5.5. We transform the Eq. (5) into the canonical form. We simply divide the state space of the program into two parts: $[x < 0]$ and $[x \geq 0]$, according to (5). By applying **Step 3** with the removal of unsatisfiable predicates, we have the

following canonical form:

$$[x < 0] \implies \min\{y\} \leq h \quad (8)$$

$$[x \geq 0] \implies \min \left\{ \begin{array}{l} 0.5 \cdot h(x+1, x+y+1) + 0.5 \cdot h(-1, y) \\ 0.25 \cdot h(-1, y+x+1) + 0.25 \cdot h(x+2, 2x+y+3) + 0.5 \cdot h(-1, y) \\ 0.25 \cdot h(-1, y+x+1) + 0.25 \cdot h(x+2, 2x+y+3) + 0.5 \cdot y \\ 0.5 \cdot h(x+1, x+y+1) + 0.5 \cdot y \end{array} \right\} \leq h \quad (9)$$

□

Step 4. Solving Constraints. Below we explain how to solve the canonical constraints from **Step 3**. Note that the canonical form (6) contains minimum operation that renders it non-convex. We first show that the minimum operation can be eliminated by transforming each constraint in the form of (6) into an equivalent formulation via the unsatisfiability of a system of non-strict and strict linear inequalities. Then the derived constraints can be transformed to bilinear constraints by applying Motzkin's Transposition Theorem, which can be solved using off-the-shell bilinear programming solvers such as *Gurobi*. We consider a variant of Motzkin's Transposition Theorem [50] below, which can reduce computation in our problem.

THEOREM 5.7 (MOTZKIN'S TRANSPOSITION THEOREM [50]). *Let*

$$S = \left[\begin{array}{c} \sum_{i=1}^n \alpha_{(1,i)} \cdot x_i + \beta_1 \leq 0 \\ \vdots \\ \sum_{i=1}^n \alpha_{(m,i)} \cdot x_i + \beta_m \leq 0 \end{array} \right] \text{ and } T = \left[\begin{array}{c} \sum_{i=1}^n \alpha_{(m+1,i)} \cdot x_i + \beta_{m+1} < 0 \\ \vdots \\ \sum_{i=1}^n \alpha_{(m+k,i)} \cdot x_i + \beta_{m+k} < 0 \end{array} \right]$$

be systems of linear inequalities where $\alpha_{(1,1)}, \dots, \alpha_{(m+k,n)}$ and $\beta_1, \dots, \beta_{m+k}$ are coefficients taken reals. If S is satisfiable, then $S \wedge T$ is unsatisfiable iff there exist non-negative reals $\lambda_0, \lambda_1, \dots, \lambda_{m+k}$ and at least one coefficient λ_i for $i \in \{m+1, \dots, m+k\}$ is non-zero, such that:

$$0 = \sum_{i=1}^{m+k} \lambda_i \alpha_{(i,1)}, \dots, 0 = \sum_{i=1}^{m+k} \lambda_i \alpha_{(i,n)}, 0 = (\sum_{i=1}^{m+k} \lambda_i \beta_i) - \lambda_0. \quad (10)$$

Theorem 5.7 is first proposed in Chatterjee et al. [20, Theorem 4.5 and Remark 4.6] without proof. We give a complete proof in Appendix C.3. In what follows, we show how Theorem 5.7 can be applied to solve our canonical constraints (6).

- First, we conjunct the affine invariant I to every inequality in the canonical form (6) while removing unsatisfiable predicates, and handle the inequality

$$[I \wedge B_j] \implies \min\{e_{1j}, e_{2j}, \dots, e_{mj}\} \leq h \quad (11)$$

for each $j \in \{1, 2, \dots, l\}$ in (6) separately by Theorem 5.7. In detail, let S_j be the the inequality system $I \wedge B_j$, and let T_j be the collection $\{e_{1j} > h, e_{2j} > h, \dots, e_{mj} > h\}$ of strict linear inequalities from (6). Then we have that the inequality (11) holds for all program states iff $S_j \wedge T_j$ is unsatisfiable. It follows from Theorem 5.7 that $S_j \wedge T_j$ is unsatisfiable iff there exist non-negative real numbers $\lambda_j = (\lambda_{0,j}, \dots, \lambda_{m_j+k_j,j})$ and at least one coefficient λ_{i_j} for $i_j \in \{m_j+1, \dots, m_j+k_j\}$ is non-zero that fulfills the equalities in (10). Notice that (10) is bilinear, since we have two separate groups of variables, namely the group of unknown coefficients in the template and the group of the fresh variables λ_j . Within each group, the relationships are linear.

- Second, our approach collects all those bilinear constraints and puts them together by conjunction, and then solves the whole system of bilinear constraints by off-the-shelf bilinear programming solvers.

Example 5.8. Continue with Example 5.6 and we use the invariant $[x \geq -1]$. We list the bilinear constraints derived from Eq. (9). By substituting $h(x, y)$ with $a \cdot x + b \cdot y + c$, we obtain the following inequalities:

$$\begin{aligned} -x \leq 0 \quad & 0.5(a-b) \cdot x - 0.5b < 0 \quad & 0.75(a-b) \cdot x - (b-0.25a) < 0 \\ & 0.75(a-b) \cdot x + 0.5(b-1) \cdot y + (0.5c - 0.25a - b) < 0 \\ & 0.5(a-b) \cdot x + 0.5(b-1) \cdot y + 0.5(c-a-b) < 0 \end{aligned}$$

Then by applying Motzkin Transposition Theorem, we obtain the following bilinear constraints:

$$\begin{aligned} \exists \lambda_0 \geq 0, \lambda_1 \geq 0, \dots, \lambda_5 \geq 0 \quad & \text{s.t.} \quad (\lambda_2 \neq 0 \vee \lambda_3 \neq 0 \vee \lambda_4 \neq 0 \vee \lambda_5 \neq 0) \wedge \\ 0 = & (-1) \cdot \lambda_1 + 0.5(a-b) \cdot \lambda_2 + 0.75(a-b) \cdot \lambda_3 + 0.75(a-b) \cdot \lambda_4 + 0.5(a-b) \cdot \lambda_5 \wedge \\ 0 = & 0.5(b-1) \cdot \lambda_4 + 0.5(b-1) \cdot \lambda_5 \wedge \\ 0 = & -0.5b \cdot \lambda_2 - (b-0.25a) \cdot \lambda_3 + (0.5c - 0.25a - b) \cdot \lambda_4 + 0.5(c-a-b) \cdot \lambda_5 - \lambda_0 \end{aligned}$$

□

Our algorithm directly calls Gurobi to solve the derived bilinear constraints. The derived constraints over the template coefficients only form a feasible domain (if feasible), and do not specify which solution to the unknown coefficients lead to an optimal upper bound. To synthesize a tight upper bound, we employ a heuristic choice of the objective function that guides the solver towards values that are closer to the tightest possible upper bound. The objective function we employ is of the form $h(\hat{s}) - f(\hat{s})$ for a designated program state $\hat{s}$ of interest and minimize this objective function. With the coefficients in the template h solved from Gurobi, we reconstruct the piecewise linear upper bound by applying the upper k -induction operator $\bar{\Psi}_h^*$ for $k-1$ times, thus obtain $\bar{\Psi}_{h^*}^{k-1}(h^*)$.

Example 5.9. Continue with Example 5.8. We employ the objective function $h-f = a \cdot x + (b-1) \cdot y + c$ when $\hat{s} = (x, y) = (1, 1)$. Therefore, we obtain the candidate $h^* = x + y + 2$. We reconstruct the piecewise upper bound by applying $\bar{\Psi}_h^*$ for once and obtain $[x < 0] \cdot y + [x \geq 0] \cdot (x + y + 2)$. □

We present the pseudocode of our algorithm in Algorithm 1.

6 EXPERIMENTAL RESULTS

In this section, we present the experimental evaluation of our approach. Our experiments were designed to address the following research questions:

- RQ1. How is the ability of our approach to generate piecewise bounds?
- RQ2. How does our approach compare with CEGISPRO2?
- RQ3. How does our approach compare with polynomial bounds?

We summarize our findings as follows:

- Our approach can synthesize tight piecewise bounds on most of our benchmarks, especially on the benchmarks that we cannot obtain linear bounds via 1-induction. Moreover, with increasing k , we can obtain tighter piecewise bounds. The experimental results show that our approaches can derive the exact bound of f , i.e., the tightest upper bound, on some benchmarks (e.g., GEO, COIN, FAIR COIN, etc). We additionally show that on a significant number of benchmarks (e.g., K-GEO, BIN-RAN, GROWING WALK-VARIANT, etc), the piecewise

Algorithm 1: The Algorithm for Our Approach

Input : Probabilistic loop P in the form of (1) and a return function f

Output: Piecewise bounds for the expected value of f after P terminates

Pre-processing:

- Check the prerequisite (i) and (ii) in Theorem 4.7.
- Establish a monolithic linear template h .
- Generate a trivial invariant I by external invariant generators.
- Determine the parameter k and a designated program state $\hat{s}$.

Constraints Extraction: Unfolding the loop within k times and calculate $pre_{C_d}(h)$ for all

$C_d \in \{C_1, \dots, C_m\}$ to obtain the piecewise constraints;

Canonical Form Transformation: Transform the constraints into the form of (6) through an iterative approach and obtain l canonical constraints;

Constraints Solving: $Cons \leftarrow \emptyset$;

for $j \leftarrow 1$ **to** l **do**

Extract the coefficients of the variables from canonical-formed constraints;

Construct bilinear constraints K_j with auxiliary variables λ_j ;

$Cons \leftarrow Cons \cup K_j$;

end

Call Gurobi to solve $Cons$ and obtain the piecewise bound with the solution h^* .

bounds we synthesize are non-trivial (i.e., the program state space S is partitioned into more than $[\varphi]$ and $[-\varphi]$). See more details in *Experimental Results*.

- The most difference between our approach and CEGISPRO2 is that they require an upper /lower bound to be verified as an additional input, while we do not need this input. For a further comparison, we feed the bounds synthesized through our approaches to CEGISPRO2 and observe that on most benchmarks, our piecewise bounds are no worse than theirs. Moreover, our approaches can handle more benchmarks since we do not have the restrictions of non-negativity like theirs.
- On most of our benchmarks, our piecewise linear bounds are significantly tighter and conciser than monolithic polynomial bounds synthesized via 1-induction.

Below we present the details of our experiments. We implement our approach in Python 3.9.12, and use Gurobi libraries in Python for bilinear programming. All experiments are conducted on a machine with Windows 10 (64 bit), Intel(R) Core(TM) i7-9750H CPU 2.60GHz and 16GB of RAM. In our experiments, we do the pre-processing as followed: (1) We check the prerequisite (i) in Theorem 4.7 through the following process: The maximum amplifier c can be chosen as 1 on most of our benchmarks so that the prerequisite can always be satisfied (by assigning c_3 an infinitesimal positive real). For the remaining few benchmarks, we observe that we can always take the right values so that the prerequisites can be satisfied. For example, on the benchmark ST-PETERSBURG, the maximum amplifier c is taken as 2, and we can take $c_3 = \ln 2$ and meanwhile, $c_2 = \ln 4$ to satisfy the prerequisites. We check the prerequisite (ii) in Theorem 4.7 for each benchmark by syntactically checking the transition probabilities in the loop. For example, on the benchmark ST-PETERSBURG, according to the probability parameter $p = \frac{3}{4}$, we take $c_2 = \ln 4$ so that it satisfy the concentration property. (2) Establish a monolithic linear template. (3) We minimize the impact of invariants by deliberately choosing trivial interval-bound invariants that can be directly derived from the loop guard and the increment/decrement statements in the loop body. (4) We set a default initial state $\hat{s}$ that all program variables take the value 1.

6.1 Piecewise Upper Bound Synthesis

Benchmarks. We choose upper-bound benchmarks from existing works [5, 11, 12, 15, 23, 29, 30, 32] that fall into our scope and have the following adaptations. First, for those that do not have linear return functions, we add simple linear return functions. Second, for those whose upper bound that can be handled directly by 1-induction (except for several classical examples: K-GEO, REVBIN, FAIR COIN), we adapt them by reasonable perturbations (such as changing the assignment statement, changing the probability parameters, reducing the continuous distribution to discrete distribution, etc) so that they require $(k > 1)$ -induction. Third, for those whose upper bound that cannot be handled by k -induction with small $k = 1, 2, 3$, we adapt them by reasonable perturbations as above so that they can be handled by $(k > 1)$ -induction, while still cannot be handled by 1-induction. Finally, we relegate examples with either repetitive patterns or cannot be handled by k -induction with small $k = 1, 2, 3$ or can be directly handled by 1-induction to Tables 5 and 6 in Appendix D.1. In this section, We consider 7 original examples and 6 adapted examples adapted from the literature. The examples GEO, K-GEO and EQUAL-PROB-GRID are taken from Batz et al. [11, 12], for which we replace the assertion probability with a linear return function *goal* in EQUAL-PROB-GRID. We consider the benchmark ZERO-CONF-VARIANT adapted from Batz et al. [11], Feng et al. [29]. We revise the original one with assignments and probabilistic parameters in this program, and add a linear return function *curprobe*. The benchmark ST-PETERSBURG VARIANT is taken from Feng et al. [29] where we replace the probability parameter $\frac{1}{2}$ with $\frac{3}{4}$ since the original program does not satisfy the prerequisites in Theorem 4.7. From Bao et al. [5], Chen et al. [23], Feng et al. [30], we consider the benchmarks COIN, MART, REVBIN and FAIR COIN, and revise the assignments, guards on the original benchmarks BIN series so that we obtain a more complex version BIN-RAN. The remaining three examples, EXPECTED TIME, GROWING WALK and its variant, are all adapted from Beutner et al. [15], Gehr et al. [32] by reducing the continuous distributions to discrete distributions.

Experimental Results. We present the experimental results for the synthesis of piecewise upper bounds on these 13 benchmarks in Table 1. In Table 1, t stands for the execution time of our approach, including the parsing from the program input and transforming the k -induction condition constraints into the bilinear problems, and "TO" stands for timeout. "Inv" stands for the invariant, "Solution" stands for the instantiated monolithic linear candidate solved by Gurobi, and "Piecewise Bound" stands for our piecewise results. The time-out threshold of bilinear problem solving via Gurobi is set to 300s. All the numerical data are cut to 10^{-4} precision. Moreover, in most cases, our approach synthesizes these bounds in no more than one minute, and hence is time efficient. On most of our benchmarks, we find that we cannot derive a monolithic linear upper bound directly by simple induction but we can obtain a monolithic linear bound with $k > 1$ -induction, and hence a piecewise linear bound. On some benchmarks, we can obtain tighter bound with increasing k .

Comparison with Other Approaches. As far as we know, the work CEGISPRO2 [11] is one of the handful relevant approaches that synthesizes piecewise upper bounds on the expected value of a return function f . The main difference between CEGISPRO2 and our approach is that CEGISPRO2 requires an upper bound to be verified as an additional program input and it will only return a super-invariant (i.e., a possibly piecewise upper-bound) that is sufficient to **verify** (i.e., globally smaller) the input upper bounds, while we intend to synthesize a tight piecewise upper bound. For a fair comparison, we only compare two simple benchmarks that paired with linear expectations and a given upper bound to be verified: GEO and K-GEO. For the benchmark GEO, the piecewise upper bounds through two methods are the same. For K-GEO, their piecewise results are $[k > N] \cdot y + [k \leq N] \cdot (-k + N + x + y + 1)$, which is consistent with our result over $\mathbb{Z}_{\geq 0}$. While in the scope of real number, our piecewise piecewise upper bound are tighter than theirs. Both of these two approaches synthesize the piecewise upper bounds within 1.0s. To have a richer comparison, we further feed

Table 1. Upper Bounds for the Expected Value of f

Benchmark	f	Inv	$k = 1$	2-induction		3-induction		Piecewise Upper Bound
				t	Solution	t	Solution	
GEO	x	$[0 \leq x]$	-	0.48	$x + 1$	1.41	$x + 1$	$[c > 0] \cdot x + [c \leq 0] \cdot (x + 1)$
K-GEO	y	$[0 \leq x \wedge 0 \leq y$ $k \leq N + 1]$	$-k + N$ $+x + y + 1$	0.79	$-k + N$ $+x + y + 1$	29.65	$-k + N$ $+x + y + 1$	$[k > N] \cdot y +$ $[k \leq N - 1] \cdot (-k + N + x + y + 1) +$ $[N - 1 < k \leq N] \cdot (-0.5k + 0.5N + x + y + 1)$
BIN-RAN	y	$[0 \leq i \leq 11 \wedge$ $0 \leq x \wedge 0 \leq y]$	-	5.56	$0.9x - 21i$ $+y + 233$	TO	-	$[i > 10] \cdot y +$ $[8.5 < i \leq 10] \cdot (0.9x - 21i + y + 233)$ $[i \leq 8.5] \cdot (0.9x - 18.9i + y + 215.1)$
COIN	i	$[0 \leq x \leq 1 \wedge$ $0 \leq y \leq 1 \wedge$ $0 \leq i$	-	3.93	$i + \frac{8}{3}$ or $i + \frac{8}{3}y - \frac{8}{3}x + \frac{8}{3}$	335.93	$i + \frac{8}{3}$	$[x \neq y] \cdot i +$ $[x = y] \cdot (i + \frac{8}{3})$
MART	i	$[0 \leq x]$	-	0.57	$i + 2$	1.52	$i + 2$	$[x \leq 0] \cdot i + [x > 0] \cdot (i + 2)$
GROWING WALK	y	$[-1 \leq x]$	-	0.48	$x + y + 2$	1.52	$x + y + 2$	$[x < 0] \cdot y + [x \geq 0] \cdot (x + y + 2)$
GROWING WALK -VARIANT	y	$[-1 \leq x]$	$x + y + 1$	0.61	$x + y + 1$	23.64	$x + y + 1$	$[x < 0] \cdot y +$ $[0 \leq x < 1] \cdot (0.5x + y + 0.25)$ $+ [x \geq 1] \cdot (x + y)$
EXPECTED TIME	t	$[-1 \leq x \leq 10]$	-	0.64	$12.1925x + t$ $+13.5502$	22.94	$4.4280x + t + 6.2461$	$[x < 0] \cdot t + [0 \leq x < 1] \cdot (t + 1)$ $[1 \leq x < 3.258] \cdot (3.9852x + t + 7.39)$ $[3.258 \leq x < 3.3772] \cdot (4.4280x + t + 6.2461)$ $[3.3772 \leq x] \cdot (3.5867x + t + 9.0874)$
ZERO-CONF -VARIANT	cur	$[0 \leq start \leq 1 \wedge$ $0 \leq est \leq 1]$	-	4.68	$-140.0est +$ $cur + 140.0$	75.08	$cur + 140.0$	$[est > 0] \cdot cur +$ $[start == 0 \wedge est \leq 0] \cdot (cur + 140.0)$ $+ [start \geq 1 \wedge est \leq 0] \cdot (cur + 42.0)$
EQUAL- PROB-GRID	$goal$	$[0 \leq a \leq 10 \wedge$ $0 \leq b \leq 10$ $goal \geq 0]$	-	38.99	$goal + 1.50$	TO	-	$[a > 10 \vee b > 10 \vee goal \neq 0] \cdot goal$ $[a \leq 10 \wedge b \leq 10 \wedge goal = 0] \cdot 1.5$
REVBIN	z	$[x \geq 0]$	$2x + z$	0.65	$2x + z$	28.10	$2x + z$	$[x < 1] \cdot z + [1 \leq x < 2] \cdot (z + x + 1)$ $+ [x \geq 2] \cdot (z + 2x)$
FAIR COIN	i	$[0 \leq x \leq 1 \wedge$ $0 \leq y \leq 1]$	$i - 2y + 2$	4.06	$i + \frac{4}{3}$	27.70	$i + \frac{4}{3}$	$[x > 0 \vee y > 0] \cdot i +$ $[x \leq 0 \wedge y \leq 0] \cdot (i + \frac{4}{3})$
ST-PETERSBURG VARIANT	y	$[0 \leq x \leq 1 \wedge$ $y \leq 0]$	-	0.48	$\frac{3}{2}y$	1.44	$\frac{3}{2}y$	$[x > 0] \cdot y + [x \leq 0] \cdot \frac{3}{2}y$

our benchmarks in Table 1 to CEGISPRO2 paired with the piecewise upper bound synthesized by our approach. We find CEGISPRO2 may have the parsing error that cannot adequately handle piecewise-input. Additionally, it reports failure (violation of non-negativity) on 5 of our benchmarks. By feeding suitable upper bounds (e.g., one segment from the piecewise bounds synthesized via our approaches) for the remaining 6 benchmarks, we find on 4 benchmarks, CEGISPRO2 produce the consistent result with our input on $\mathbb{Z}_{\geq 0}$, while possibly incorrect over $\mathbb{R}$. On BIN-RAN, the results they produce are sophisticated to compare since it produces a bound with numerous segments. On COIN, they do not seem to be able to output a result. The work KIPRO2 [12] verifies an input upper bound by directly applying latticed k -induction which we demonstrate to have a piecewise nature. Thus it is unreasonable to compare our approach with KIPRO2 as we pursue the more general problem of synthesizing piecewise bounds rather than verification. Moreover, the work [5] considers the probabilistic invariant synthesis via data-driven approach, while ours is based on constraint solving and fully automated. The only relevant work with our upper bound synthesis is the exact invariant synthesis. We evaluate our approaches on the common benchmarks and show that on these benchmarks (eg., GEO, COIN, MART, FAIR COIN), the piecewise upper bound we synthesize is actually the exact expected value of the return function f .

Except for the existing tools, we additionally compare our piecewise linear upper bounds with monolithic polynomial bounds synthesized via simple induction (i.e., 1-induction). We implement the polynomial synthesis with Putinar's Positivstellensatz [53] (see Appendix D.2 for more details). For a fair comparison, we generate the polynomial bounds with the same invariant and optimal objective functions for each benchmark, as shown in Table 2. In this table, "Monolithic Linear / Polynomial via Simple Induction" stand for the monolithic linear/polynomial bounds synthesized

Table 2. Comparison with Monolithic Polynomial Bounds: Upper Case

Benchmark	f	Monolithic Linear via Simple Induction ($k = 1$)	Our Approach		Monolithic Polynomial via Simple Induction ($k = 1$)	PCT
			k	Piecewise Upper Bound		
GEO	x	-	3	$[c > 0] \cdot x +$ $[c \leq 0] \cdot (x + 1)$	$1.0000 - 1.9996 * c + 1.0000 * x +$ $0.9996 * c^2 - 0.0002 * x * c + 0.0002 * x * c^2$	0.0%
K-GEO	y	$-k + N +$ $x + y + 1$	3	$[k > N] \cdot y +$ $[k \leq N - 1] \cdot (-k + N + x + y + 1) +$ $[N - 1 < k \leq N] \cdot (-0.5k + 0.5N + x + y + 1)$	$274.1142 - 53.62281 * N - 1.0000 * k +$ $1.0000 * y + 1.0000 * x + 2.7311 * N^2$	2.82%
BIN-RAN	y	-	2	$[i > 10] \cdot y +$ $[8.5 < i \leq 10] \cdot (0.9x - 21i + y + 233)$ $[i \leq 8.5] \cdot (0.9x - 18.9i + y + 215.1)$	$38.6287 + 12.07507 * i - 13.87917 * y -$ $8.2916 * x - 0.5694 * i^2 - 0.1339 * y * i +$ $0.0329 * y^2 - 0.0644 * x * i - 0.1277 * x * y -$ $1.1617 * x^2 - 0.0462 * i^3 + 0.1205 * y * i^2 +$ $0.0116 * y^2 * i - 0.0046 * y^3 + 0.1720 * x * i^2 -$ $0.0005 * x * y * i - 0.0095 * x * y^2 + 0.3601 * x^2 * i$ $- 0.0392 * x^2 * y + 0.3248 * x^3$	62.76%
COIN	i	-	3	$[x \neq y] \cdot i +$ $[x = y] \cdot (i + \frac{5}{3})$	$2.6667 + 1.0000 * i - 0.6381 * y +$ $4.2840 * x - 2.0286 * y^2 - 2.0067 * x * y$ $+ 0.3893 * x^2$	0.0%
MART	i	-	3	$[x \leq 0] \cdot i +$ $[x > 0] \cdot (i + 2)$	$0.0248 + 1.0000 * i +$ $199999.6588 * x + 0.1643 * x^2$	0.0%
GROWING WALK	y	-	3	$[x < 0] \cdot y +$ $[x \geq 0] \cdot (x + y + 2)$	$2.5000 + 1.0000 * y + 1.900 * x$ $- 0.5000 * x^2 + 0.1000 * x^3$	0.0%
GROWING WALK VARIANT	y	$x + y + 1$	3	$[x < 0] \cdot y +$ $[0 \leq x < 1] \cdot (0.5x + y + 0.25)$ $+ [x \geq 1] \cdot (x + y)$	$1.0000 * y - 0.2380 * x + 0.1041 * y^2 -$ $0.0686 * x * y + 0.0951 * x^2 + 0.03558 * x * y^2$ $+ 0.0686 * x^2 * y + 0.1430 * x^3$	5.52%
EXPECTED TIME	t	-	3	$[x < 0] \cdot t + [0 \leq x < 1] \cdot (t + 1)$ $[1 \leq x < 3.258] \cdot (3.9852x + t + 7.39)$ $[3.258 \leq x < 3.3772] \cdot (4.4280x + t + 6.2461)$ $[3.3772 \leq x] \cdot (3.5867x + t + 9.0874)$	$3.1203 + 0.9622 * t + 2.8278 * x +$ $0.0015 * t^2 - 0.01558 * x * t - 0.1397 * x^2 -$ $0.0003 * x * t^2 - 0.0002 * x^2 * t + 0.0025 * x^3$	50.0%
ZERO-CONF -VARIANT	cur	-	3	$[est > 0] \cdot cur +$ $[start == 0 \wedge est \leq 0] \cdot (cur + 140.0)$ $+ [start \geq 1 \wedge est \leq 0] \cdot (cur + 42.0)$	$109.8660 - 0.1357 * cur + 293795.0410 * start +$ $209178.7117 * est + 0.0019 * cur^2 + 0.7202 * start * cur -$ $293865.0570 * start^2 + 1.0313 * est * cur +$ $274251.8886 * est * start - 209283.0750 * est^2$	0.5 %
EQUAL- PROB-GRID	$goal$	-	2	$[a > 10 \vee b > 10 \vee goal \neq 0] \cdot goal$ $[a \leq 10 \wedge b \leq 10 \wedge goal = 0] \cdot 1.5$	$1.6661 + 5.7396 * goal - 9.4857 * 10^{-5} * b +$ $1.5707 * 10^{-5} * a + 0.6003 * goal^2 - 0.6740 * b * goal$ $+ 1.5975 * 10^{-5} * b^2 + 2.2074 * 10^{-5} * a * goal$	0.0%
REVBin	z	$2x + z$	3	$[x < 1] \cdot z +$ $[1 \leq x < 2] \cdot (z + x + 1)$ $+ [x \geq 2] \cdot (z + 2x)$	$1.0000 * z + 2.0000 * x$	0.0%
FAIR COIN	i	$i - 2y + 2$	3	$[x > 0 \vee y > 0] \cdot i +$ $[x \leq 0 \wedge y \leq 0] \cdot (i + \frac{4}{3})$	$1.3333 + 1.0000 * i - 0.4141 * y -$ $0.4141 * x + 1.1743 * i^2 - 2.3486 * y * i +$ $0.2551 * y^2 - 2.3486 * x * i + 3.6820 * x * y$ $+ 0.2551 * x^2$	0.0%
ST-PETERSBURG VARIANT	y	-	3	$[x > 0] \cdot y + [x \leq 0] \cdot \frac{3}{2}y$	$0.0197 + 1.5047 * y + 371727.7656 * x$ $- 0.5028 * x * y - 371727.7734 * x^2$	0.0%

via 1-induction, and value " k " stands for the k -induction principle we apply in this comparison. We compare two results by uniformly taking the grid points of some interesting region (usually a subset of invariant) and evaluate two results, and we compute the percentage of the points that our piecewise upper bound are larger (i.e., not better) than monolithic polynomial, which is shown in the last column "PCT" in Table 2. We show that on all benchmarks but two (BIN-RAN, EXPECTED TIME), our piecewise linear bounds are significantly tighter and conciser than monolithic polynomial bounds. For these two benchmarks, we guess that its exact expected value of f is closer to a piecewise polynomial.

6.2 Piecewise Lower Bound Synthesis

For the lower bounds, we consider the same benchmarks and return function f as in Section 6.1 and use the same global invariant for each benchmark. The experimental results are placed in Table 3. In this table, we only show the piecewise result with at most $k = 3$. We observe that on most of the benchmarks, we can obtain a lower bound via simple induction (i.e., 1-induction), while the piecewise lower bounds we synthesize gets better with increasing k . Only on the benchmark GROWING WALK-VARIANT, we requires ($k > 1$)-induction to synthesize a lower bound.

Table 3. Lower Bounds for the Expected Value of f

Benchmark	f	1-induction		2-induction		3-induction		Piecewise Lower Bound
		t	Solution	t	Solution	t	Solution	
GEO	x	0.11	x	0.47	x	1.44	x	$[c > 0] \cdot x + [c \leq 0] \cdot (x + \frac{3}{4})$
K-GEO	y	0.14	y	0.88	$-\frac{1}{3}k + \frac{1}{3}N + y$	29.49	y	$[k > N] \cdot y + [k \leq N] \cdot (x + y + 0.5)$
BIN-RAN	y	0.31	$-0.5i + y + 5$	5.59	$-0.7241i + y + 7.2414$	TO	-	$\frac{[i > 10] \cdot y}{[1.4 < i \leq 10] \cdot (-0.6517i + y + 0.45x + 7.1397)}$ $\frac{[i \leq 1.4] \cdot (-0.7241i + y + 7.2414)}{[i \leq 1.4] \cdot (-0.7241i + y + 7.2414)}$
COIN	i	0.34	i	3.69	i	351.12	i	$[y \neq x] \cdot i + [y = x] \cdot (i + \frac{13}{8})$
MART	i	0.17	i	0.55	i	1.68	i	$[x \leq 0] \cdot i + [x > 0] \cdot (i + 1.5)$
GROWING WALK	y	0.12	$x + y$	0.48	$x + y$	1.73	$x + y$	$[x < 0] \cdot y + [x \geq 0] \cdot (x + y + \frac{5}{4})$
GROWING WALK -VARIANT	y	-	-	0.61	$y - 1$	23.53	$y - 1$	$\frac{[x < 0] \cdot y + [0 \leq x < 1] \cdot (y + 0.5x - 1) + [1 \leq x < 2] \cdot (y + 0.5x - 1.5) + [2 \leq x] \cdot (y + 0.75x - 2)}{[x < 0] \cdot y + [0 \leq x < 1] \cdot (y + 0.5x - 1) + [1 \leq x < 2] \cdot (y + 0.5x - 1.5) + [2 \leq x] \cdot (y + 0.75x - 2)}$
EXPECTED TIME	t	0.14	$1.1111x + t$	0.66	$1.1728x + t$	23.76	$1.240x + t$	$\frac{[x < 0] \cdot t + [0 \leq x < 1] \cdot (0.124x + t + 0.9)}{[1 \leq x \leq 10] \cdot (1.1284x + t + 1.9116)}$
ZERO-CONF -VARIANT	cur	0.27	cur	4.89	cur	77.81	cur	$\frac{[est > 0] \cdot cur + [start == 0 \wedge est \leq 0] \cdot (cur + 1.9502) + [start \geq 1 \wedge est \leq 0] \cdot (cur + 0.287)}{[est > 0] \cdot cur + [start == 0 \wedge est \leq 0] \cdot (cur + 1.9502) + [start \geq 1 \wedge est \leq 0] \cdot (cur + 0.287)}$
EQUAL- PROB-GRID	$goal$	0.31	$goal$	37.69	$goal$	TO	-	$\frac{[a > 10 \vee b > 10 \vee b < 10 \vee goal \neq 0] \cdot goal}{[a \leq 10 \wedge b == 10 \wedge goal = 0] \cdot 1.5}$
REVBIN	z	0.12	$z + 2x - 2$	0.65	$z + 2x - 2$	28.36	$z + 2x - 2$	$\frac{[x < 1] \cdot z + [1 \leq x < 2] \cdot (z + x) + [x \geq 2] \cdot (z + 2x - 2)}{[x < 1] \cdot z + [1 \leq x < 2] \cdot (z + x) + [x \geq 2] \cdot (z + 2x - 2)}$
FAIR COIN	i	0.30	i	3.66	i	25.93	i	$[x > 0 \vee y > 0] \cdot i + [x = 0 \wedge y = 0] \cdot (i + \frac{5}{4})$
ST-PETERSBURG VARIANT	y	0.12	y	0.47	y	1.47	y	$[x > 0] \cdot y + [x \leq 0] \cdot \frac{11}{8}y$

Comparison with Other Approaches The relevant work [5, 11] require a (possibly piecewise) lower bound (i.e., Pre-expectation E) to be verified as a program input and return a sub-invariant that is sufficient to **verify** the input lower bound, which is the most difference from our work. They produce the results by a proof rule derived from the original OST (see Section 6 & 7 in Batz et al. [11] and Appendix B.1), while we apply an extended OST (see Theorem 4.6), so that some of our benchmarks (e.g., MARTINGALE, EXPECTED TIME) cannot be handled by CEGISPRO2 or EXIST.

Likewise the upper case, we compare our piecewise linear upper bounds with monolithic polynomial bounds synthesized via simple induction (i.e., 1-induction), as shown in Table 4. According to the comparison results "PCT", we observe that on all benchmarks but two (BIN-RAN, EXPECTED TIME), our piecewise linear lower bounds are significantly tighter (i.e., greater) and conciser than monolithic polynomial lower bounds.

7 RELATED WORKS

Our approach uses templates to synthesize piecewise bounds. Most template-based approaches (e.g. Chakarov and Sankaranarayanan [17], Chatterjee et al. [18, 20]) focus on synthesizing monolithic bounds or invariants over probabilistic programs. Compared with these approaches, our approach targets piecewise bounds, and hence is orthogonal.

The work [12] proposes latticed k -induction. The differences with our result are three fold. First, our result solves the automated synthesis of piecewise linear bounds, while the work [12] only determines whether a given (possibly piecewise) bound is an upper bound or not for a probabilistic while loop, and does not solve the automated synthesis of (piecewise) bounds. Second, we use OST as the mathematical backbone, while the work [12] uses fixed point theory. The use of OST

Table 4. Comparison with Monolithic Polynomial Bounds: Lower Case

Benchmark	f	Monolithic Linear via Inductive Synthesis ($k = 1$)	Our Approach		Monolithic Polynomial via Inductive Synthesis ($k = 1$)	PCT
			k	Piecewise Lower Bound		
GEO	x	x	4	$[c > 0] \cdot x +$ $[c \leq 0] \cdot (x + \frac{c}{8})$	$-0.0313 - 0.1902 * c + 1.0478 * x -$ $0.3980 * c^2 + 0.0695 * x * c - 0.0019 * x^2 -$ $0.1595 * x * c^2 + 0.07227 * x^2 * c - 0.0147 * x^3$	0.0%
K-GEO	y	y	3	$[k > N] \cdot y +$ $[k \leq N] \cdot (x + y + 0.5)$	$-221.2813 + 44.6223 * N - 0.7791 * k + 1.0000 * y +$ $0.9281 * x - 2.1922 * N^2 - 0.1043 * x^2$	2.96%
BIN-RAN	y	$-0.5i + y + 5$	2	$[i > 10] \cdot y$ $[1.4 < i \leq 10] \cdot (-0.6517i + y + 0.45x + 7.1397$ $[i \leq 1.4] \cdot (-0.7241i + y + 7.2414)$	$-6.5656 - 17.5555 * i + 19.7273 * y +$ $12.3734 * x + 0.6130 * i^2 + 0.1623 * y * i$ $-0.0377 * y^2 + 0.1105 * x * i + 0.1624 * x * y$ $+1.4912 * x^2 + 0.0530 * i^3 - 0.1521 * y * i^2 -$ $0.0152 * y^2 * i + 0.0058 * y^3 - 0.2255 * x * i^2$ $+0.0002 * x * y * i + 0.0119 * x * y^2 - 0.4682 * x^2 * i$ $+0.0507 * x^2 * y - 0.4160 * x^3$	34.75%
COIN	i	i	4	$[y \neq x] \cdot i + [y = x] \cdot (i + \frac{129}{64})$	$2.6655 + 1.0002 * i - 3622.3830 * y -$ $5419.0667 * x - 0.0001 * i^2 + 0.0007 * y * i +$ $3619.71553 * y^2 - 0.0008 * x * i + 1827.4383 * x * y$ $+3594.2952 * x^2$	2%
MART	i	i	4	$[x \leq 0] \cdot i +$ $[x > 0] \cdot (i + \frac{1}{4})$	$1.0000 * i + 39.9996 * x - 199.9958 * x^2$	1.0%
GROWING WALK	y	$x + y$	4	$[x < 0] \cdot y +$ $[x \geq 0] \cdot (x + y + \frac{13}{8})$	$-0.0004 + 1.0003 * y + 1.3463 * x -$ $0.0001 * y^2 - 0.0010 * x * y - 0.0590 * x^2 -$ $6.6040 * 10^{-5} * x * y^2 + 0.0007 * x^3 * y - 0.0022 * x^3$	0.0%
GROWING WALK VARIANT	y	-	3	$[x < 0] \cdot y +$ $[0 \leq x < 1] \cdot (0.5x + y - 1)$ $+ [1 \leq x < 2] \cdot (0.5x + y - 1.5)$ $+ [2 \leq x] \cdot (0.75x + y - 2)$	$-1.0000 + 1.0000 * y - 0.3903 * x -$ $0.0734 * y^2 + 0.0484 * x * y + 0.4758 * x^2 -$ $0.0250 * x * y^2 - 0.0484 * x^2 * y - 0.0855 * x^3$	0.01%
EXPECTED TIME	t	$1.1111x + t$	3	$[x < 0] \cdot t +$ $[0 \leq x < 1] \cdot (0.124x + t + 0.9)$ $[1 \leq x \leq 10] \cdot (1.1284x + t + 1.9116)$	$-0.0784 + 1.0093 * t + 3.1426 * x -$ $0.0010 * t^2 + 0.0083 * x * t - 0.1576 * x^2 +$ $0.0002 * x * t^2 + 0.0002 * x^2 * t + 0.0043 * x^3$	64.6 %
ZERO-CONF -VARIANT	cur	cur	3	$[est > 0] \cdot cur +$ $[start == 0 \wedge est \leq 0] \cdot (cur + 1.9502)$ $+ [start \geq 1 \wedge est \leq 0] \cdot (cur + 0.287)$	$140.2458 + 1.0098 * cur - 424365.5964 * start -$ $587675.0179 * est - 0.0066 * start * cur +$ $424267.3602 * start^2 - 0.0095 * est * cur -$ $504437.5495 * est * start + 587534.7143 * est^2$	0.64%
EQUAL- PROB-GRID	$goal$	$goal$	2	$[a > 10 \vee b > 10 \vee goal \neq 0] \cdot goal$ $[a \leq 10 \wedge b \leq 10 \wedge goal = 0] \cdot 1.5$	$0.4950 * goal - 0.2020 * goal^2 +$ $0.0053 * b * goal - 0.0011 * a * goal$	0.0 %
REVBIN	z	$2x + z - 2$	3	$[x < 1] \cdot z +$ $[1 \leq x < 2] \cdot (z + x)$ $+ [x \geq 2] \cdot (z + 2x - 2)$	$-2.0000 + 1.0000 * z + 2.0000 * x$	0.0%
FAIR COIN	i	i	4	$[x > 0 \vee y > 0] \cdot i +$ $[x \leq 0 \wedge y \leq 0] \cdot (i + \frac{21}{16})$	$1.0000 * i - 0.3932 * y - 0.39325 * x$ $-0.3153 * i^2 + 0.6305 * y * i - 0.7242 * y^2$ $+0.6305 * x * i - 0.1796 * x * y - 0.7242 * x^2$	0.0%
ST-PETERSBURG VARIANT	y	y	3	$[x > 0] \cdot y + [x \leq 0] \cdot \frac{11}{8}y$	$-0.0017 + 1.0023 * y - 121479.0179 * x$ $-0.0550 * x * y + 121479.0185 * x^2$	0.0%

lifts the non-negativity (or a global lower bound) of program values as required in Batz et al. [12]. Third, the algorithm in Batz et al. [12] is a direction application of SMT solving such as Z3, while we reduce our problem to bilinear programming. The work [45] proposes tightened variants of k -induction operators. We show that these tightened variants are actually equivalent with the original k -induction operator in Batz et al. [12] and the dual k -induction operator proposed in our result.

The work [11] proposes to compute the piecewise bounds for the expected value of post-expectations f in probabilistic loop. The main difference between their work and ours is that they require an input upper bound to be verified as a program input and synthesize a piecewise linear upper bound to verify the input upper bound via counterexample-guided inductive synthesis (CEGIS), while we do not need this additional program input and our approach solves the bounds by bilinear programming rather than CEGIS. Besides, compared with CEGISPRO2, a major advantage is that we do not require global non-negativity and we only need to predefine a monolithic template, instead of a piecewise template. For the synthesis of lower bounds, this work apply a proof rule in Hark et al. [36] derived from the original OST, while our approach apply an extended OST. Thus some benchmarks in our work cannot be handled by CEGISPRO2. The work [5] synthesize exact and sub invariants for probabilistic while loops through data-driven learning and they also capture

piecewise feature for some benchmarks. However, this approach takes as input a suitable list of features that is composed of numerical expressions over program variables. The list of features always contains constructive numerical expressions of program variables and even user-supplied features for more specific invariant, while our approach can capture the piecewise automatically and without any additional knowledge. Apart from Bao et al. [5], Batz et al. [11], several works (e.g. Chen et al. [23], Feng et al. [30], Katoen et al. [39]) aim to synthesize monolithic quantitative sub-invariant for proving lower bounds.

Other approaches [3, 8, 9, 40, 41] focus on moment-based invariants generation and high-order moments derivation for probabilistic programs. Moment-based method is highly efficient and is guaranteed to compute moments exactly. This kind of work can even handle the probabilistic program with non-polynomial expressions and continuous distribution, but they only consider the probabilistic while loop in the form of **while** true { C }, which is a rather restricted form. The work [49] enlarges the theoretical foundation by support of if-statements (thus also guarded loops), but they assume that all variables appearing in if-conditions (loop guards) are finitely valued, hence cannot handle most of our benchmarks. Our approach can handle all the linear forms of loop guards and if-conditions. The work [40] focus on the distribution estimation when the program terminates using statistical approaches, given a concrete initial state, and hence orthogonal. In a similar vein, the works [44, 60] bound higher central moments for running time and other monotonically increasing quantities. However, they are limited to programs with constant size increments.

8 CONCLUSION AND FUTURE WORK

In this work, we propose a novel approach for synthesizing piecewise linear expectation bounds over probabilistic loops. We extract a piecewise pattern from latticed k -induction, establish the theoretical foundation via an existing extension of Optional Stopping Theorem, and solve the bounds by bilinear programming. Experimental results show that our approach is efficient and derives tight piecewise linear bounds over an extensive set of benchmarks. A future work is to explore whether the derived bilinear programming problems can be reduced to semi-definite programming by exploiting recent works in Wang et al. [65, 66]. Another future work is to consider piecewise polynomial bounds for a larger class of quantitative properties via latticed k -induction.

REFERENCES

- [1] Alessandro Abate, Mirco Giacobbe, and Diptarko Roy. 2021. Learning Probabilistic Termination Proofs. In *Computer Aided Verification - 33rd International Conference, CAV 2021, Virtual Event, July 20-23, 2021, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 12760)*, Alexandra Silva and K. Rustan M. Leino (Eds.). Springer, 3–26. https://doi.org/10.1007/978-3-030-81688-9_1
- [2] Alejandro Aguirre, Gilles Barthe, Justin Hsu, Benjamin Lucien Kaminski, Joost-Pieter Katoen, and Christoph Matheja. 2021. A pre-expectation calculus for probabilistic sensitivity. *Proc. ACM Program. Lang.* 5, POPL (2021), 1–28. <https://doi.org/10.1145/3434333>
- [3] Daneshvar Amrollahi, Ezio Bartocci, George Kenison, Laura Kovács, Marcel Moosbrugger, and Miroslav Stankovic. 2022. Solving Invariant Generation for Unsolvable Loops. In *Static Analysis - 29th International Symposium, SAS 2022, Auckland, New Zealand, December 5-7, 2022, Proceedings (Lecture Notes in Computer Science, Vol. 13790)*, Gagandeep Singh and Caterina Urban (Eds.). Springer, 19–43. https://doi.org/10.1007/978-3-031-22308-2_3
- [4] Christel Baier and Joost-Pieter Katoen. 2008. *Principles of model checking*. MIT Press.
- [5] Jialu Bao, Nitesh Trivedi, Drashti Pathak, Justin Hsu, and Subhajit Roy. 2022. Data-Driven Invariant Learning for Probabilistic Programs. In *Computer Aided Verification - 34th International Conference, CAV 2022, Haifa, Israel, August 7-10, 2022, Proceedings, Part I (Lecture Notes in Computer Science, Vol. 13371)*, Sharon Shoham and Yakir Vizel (Eds.). Springer, 33–54. https://doi.org/10.1007/978-3-031-13185-1_3
- [6] Andreas Bärmann, Robert Burlacu, Lukas Hager, and Thomas Kleinert. 2023. On piecewise linear approximations of bilinear terms: structural comparison of univariate and bivariate mixed-integer programming formulations. *J. Glob. Optim.* 85, 4 (2023), 789–819. <https://doi.org/10.1007/S10898-022-01243-Y>
- [7] Gilles Barthe, Marco Gaboardi, Benjamin Grégoire, Justin Hsu, and Pierre-Yves Strub. 2016. Proving Differential Privacy via Probabilistic Couplings. In *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016*, Martin Grohe, Eric Koskinen, and Natarajan Shankar (Eds.). ACM, 749–758. <https://doi.org/10.1145/2933575.2934554>
- [8] Ezio Bartocci, Laura Kovács, and Miroslav Stankovic. 2019. Automatic Generation of Moment-Based Invariants for Prob-Solvable Loops. In *Automated Technology for Verification and Analysis - 17th International Symposium, ATVA 2019, Taipei, Taiwan, October 28-31, 2019, Proceedings (Lecture Notes in Computer Science, Vol. 11781)*, Yu-Fang Chen, Chih-Hong Cheng, and Javier Esparza (Eds.). Springer, 255–276. https://doi.org/10.1007/978-3-030-31784-3_15
- [9] Ezio Bartocci, Laura Kovács, and Miroslav Stankovic. 2020. Mora - Automatic Generation of Moment-Based Invariants. In *Tools and Algorithms for the Construction and Analysis of Systems - 26th International Conference, TACAS 2020, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2020, Dublin, Ireland, April 25-30, 2020, Proceedings, Part I (Lecture Notes in Computer Science, Vol. 12078)*, Armin Biere and David Parker (Eds.). Springer, 492–498. https://doi.org/10.1007/978-3-030-45190-5_28
- [10] Kevin Batz, Tom Jannik Biskup, Joost-Pieter Katoen, and Tobias Winkler. 2023. Programmatic Strategy Synthesis: Resolving Nondeterminism in Probabilistic Programs. *CoRR* abs/2311.06889 (2023). <https://doi.org/10.48550/ARXIV.2311.06889> arXiv:2311.06889
- [11] Kevin Batz, Mingshuai Chen, Sebastian Junges, Benjamin Lucien Kaminski, Joost-Pieter Katoen, and Christoph Matheja. 2023. Probabilistic Program Verification via Inductive Synthesis of Inductive Invariants. In *Tools and Algorithms for the Construction and Analysis of Systems - 29th International Conference, TACAS 2023, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2023, Paris, France, April 22-27, 2023, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 13994)*, Sriram Sankaranarayanan and Natasha Sharygina (Eds.). Springer, 410–429. https://doi.org/10.1007/978-3-031-30820-8_25
- [12] Kevin Batz, Mingshuai Chen, Benjamin Lucien Kaminski, Joost-Pieter Katoen, Christoph Matheja, and Philipp Schröder. 2021. Latticed k-Induction with an Application to Probabilistic Programs. In *Computer Aided Verification - 33rd International Conference, CAV 2021, Virtual Event, July 20-23, 2021, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 12760)*, Alexandra Silva and K. Rustan M. Leino (Eds.). Springer, 524–549. https://doi.org/10.1007/978-3-030-81688-9_25
- [13] Guillaume O. Berger and Sriram Sankaranarayanan. 2023. Counterexample-guided computation of polyhedral Lyapunov functions for piecewise linear systems. *Autom.* 155 (2023), 111165. <https://doi.org/10.1016/J.AUTOMATICA.2023.111165>
- [14] Guillaume O. Berger and Sriram Sankaranarayanan. 2023. Template-Based Piecewise Affine Regression. In *Learning for Dynamics and Control Conference, L4DC 2023, 15-16 June 2023, Philadelphia, PA, USA (Proceedings of Machine Learning Research, Vol. 211)*, Nikolai Matni, Manfred Morari, and George J. Pappas (Eds.). PMLR, 509–520. <https://proceedings.mlr.press/v211/berger23a.html>
- [15] Raven Beutner, C.-H. Luke Ong, and Fabian Zaiser. 2022. Guaranteed bounds for posterior inference in universal probabilistic programming. In *PLDI '22: 43rd ACM SIGPLAN International Conference on Programming Language Design and Implementation, San Diego, CA, USA, June 13 - 17, 2022*, Ranjit Jhala and Isil Dillig (Eds.). ACM, 536–551. <https://doi.org/10.1145/3519939.3523721>

- [16] Michael Carbin, Sasa Misailovic, and Martin C. Rinard. 2013. Verifying quantitative reliability for programs that execute on unreliable hardware. In *Proceedings of the 2013 ACM SIGPLAN International Conference on Object Oriented Programming Systems Languages & Applications, OOPSLA 2013, part of SPLASH 2013, Indianapolis, IN, USA, October 26-31, 2013*, Antony L. Hosking, Patrick Th. Eugster, and Cristina V. Lopes (Eds.). ACM, 33–52. <https://doi.org/10.1145/2509136.2509546>
- [17] Aleksandar Chakarov and Sriram Sankaranarayanan. 2013. Probabilistic Program Analysis with Martingales. In *Computer Aided Verification - 25th International Conference, CAV 2013, Saint Petersburg, Russia, July 13-19, 2013. Proceedings (Lecture Notes in Computer Science, Vol. 8044)*, Natasha Sharygina and Helmut Veith (Eds.). Springer, 511–526. https://doi.org/10.1007/978-3-642-39799-8_34
- [18] Krishnendu Chatterjee, Hongfei Fu, and Amir Kafshdar Goharshady. 2016. Termination Analysis of Probabilistic Programs Through Positivstellensatz's. In *Computer Aided Verification - 28th International Conference, CAV 2016, Toronto, ON, Canada, July 17-23, 2016, Proceedings, Part I (Lecture Notes in Computer Science, Vol. 9779)*, Swarat Chaudhuri and Azadeh Farzan (Eds.). Springer, 3–22. https://doi.org/10.1007/978-3-319-41528-4_1
- [19] Krishnendu Chatterjee, Hongfei Fu, Petr Novotný, and Rouzbeh Hasheminezhad. 2016. Algorithmic analysis of qualitative and quantitative termination problems for affine probabilistic programs. In *Proceedings of the 43rd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2016, St. Petersburg, FL, USA, January 20 - 22, 2016*, Rastislav Bodik and Rupak Majumdar (Eds.). ACM, 327–342. <https://doi.org/10.1145/2837614.2837639>
- [20] Krishnendu Chatterjee, Hongfei Fu, Petr Novotný, and Rouzbeh Hasheminezhad. 2018. Algorithmic Analysis of Qualitative and Quantitative Termination Problems for Affine Probabilistic Programs. *ACM Trans. Program. Lang. Syst.* 40, 2 (2018), 7:1–7:45. <https://doi.org/10.1145/3174800>
- [21] Krishnendu Chatterjee, Petr Novotný, and Dord Zikelic. 2017. Stochastic invariants for probabilistic termination. In *Proceedings of the 44th ACM SIGPLAN Symposium on Principles of Programming Languages, POPL 2017, Paris, France, January 18-20, 2017*, Giuseppe Castagna and Andrew D. Gordon (Eds.). ACM, 145–160. <https://doi.org/10.1145/3009837.3009873>
- [22] Chuangtao Chen, Weikang Qian, Mohsen Imani, Xunzhao Yin, and Cheng Zhuo. 2022. PAM: A Piecewise-Linearly-Approximated Floating-Point Multiplier With Unbiasedness and Configurability. *IEEE Trans. Computers* 71, 10 (2022), 2473–2486. <https://doi.org/10.1109/TC.2021.3131850>
- [23] Yu-Fang Chen, Chih-Duo Hong, Bow-Yaw Wang, and Lijun Zhang. 2015. Counterexample-Guided Polynomial Loop Invariant Generation by Lagrange Interpolation. In *Computer Aided Verification - 27th International Conference, CAV 2015, San Francisco, CA, USA, July 18-24, 2015, Proceedings, Part I (Lecture Notes in Computer Science, Vol. 9206)*, Daniel Kroening and Corina S. Pasareanu (Eds.). Springer, 658–674. https://doi.org/10.1007/978-3-319-21690-4_44
- [24] Leonardo De Moura and Nikolaj Bjørner. 2008. Z3: An Efficient SMT Solver. In *Proceedings of the Theory and Practice of Software, 14th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (Budapest, Hungary) (TACAS'08/ETAPS'08)*. Springer-Verlag, Berlin, Heidelberg, 337–340.
- [25] Leonardo Mendonça de Moura, Harald Rueß, and Maria Sorea. 2003. Bounded Model Checking and Induction: From Refutation to Verification (Extended Abstract, Category A). In *Computer Aided Verification, 15th International Conference, CAV 2003, Boulder, CO, USA, July 8-12, 2003, Proceedings (Lecture Notes in Computer Science, Vol. 2725)*, Warren A. Hunt Jr. and Fabio Somenzi (Eds.). Springer, 14–26. https://doi.org/10.1007/978-3-540-45069-6_2
- [26] Alastair F. Donaldson, Leopold Haller, Daniel Kroening, and Philipp Rümmer. 2011. Software Verification Using k -Induction. In *Static Analysis - 18th International Symposium, SAS 2011, Venice, Italy, September 14-16, 2011. Proceedings (Lecture Notes in Computer Science, Vol. 6887)*, Eran Yahav (Ed.). Springer, 351–368. https://doi.org/10.1007/978-3-642-23702-7_26
- [27] J. L. Doob. 1971. What is a Martingale? *The American Mathematical Monthly* 78, 5 (1971), 451–463. <https://doi.org/10.1080/00029890.1971.11992788> arXiv:<https://doi.org/10.1080/00029890.1971.11992788>
- [28] Gy Farkas. 1894. A Fourier-féle mechanikai elv alkalmazásai. *Mathematikai és Természettudományi Értesítő* 12 (1894), 457–472.
- [29] Shenghua Feng, Mingshuai Chen, Han Su, Benjamin Lucien Kaminski, Joost-Pieter Katoen, and Naijun Zhan. 2023. Lower Bounds for Possibly Divergent Probabilistic Programs. *Proc. ACM Program. Lang.* 7, OOPSLA1 (2023), 696–726. <https://doi.org/10.1145/3586051>
- [30] Yijun Feng, Lijun Zhang, David N. Jansen, Naijun Zhan, and Bican Xia. 2017. Finding Polynomial Loop Invariants for Probabilistic Programs. In *Automated Technology for Verification and Analysis - 15th International Symposium, ATVA 2017, Pune, India, October 3-6, 2017, Proceedings (Lecture Notes in Computer Science, Vol. 10482)*, Deepak D'Souza and K. Narayan Kumar (Eds.). Springer, 400–416. https://doi.org/10.1007/978-3-319-68167-2_26
- [31] Hongfei Fu and Krishnendu Chatterjee. 2019. Termination of Nondeterministic Probabilistic Programs. In *Verification, Model Checking, and Abstract Interpretation - 20th International Conference, VMCAI 2019, Cascais, Portugal, January 13-15, 2019, Proceedings (Lecture Notes in Computer Science, Vol. 11388)*, Constantin Enea and Ruzica Piskac (Eds.). Springer, 468–490. https://doi.org/10.1007/978-3-030-11245-5_22

- [32] Timon Gehr, Sasa Misailovic, and Martin T. Vechev. 2016. PSI: Exact Symbolic Inference for Probabilistic Programs. In *Computer Aided Verification - 28th International Conference, CAV 2016, Toronto, ON, Canada, July 17-23, 2016, Proceedings, Part I (Lecture Notes in Computer Science, Vol. 9779)*, Swarat Chaudhuri and Azadeh Farzan (Eds.). Springer, 62–83. https://doi.org/10.1007/978-3-319-41528-4_4
- [33] Andrew D. Gordon, Thomas A. Henzinger, Aditya V. Nori, and Sriram K. Rajamani. 2014. Probabilistic programming. In *Proceedings of the on Future of Software Engineering, FOSE 2014, Hyderabad, India, May 31 - June 7, 2014*, James D. Herbsleb and Matthew B. Dwyer (Eds.). ACM, 167–181. <https://doi.org/10.1145/2593882.2593900>
- [34] Friedrich Gretz, Joost-Pieter Katoen, and Annabelle McIver. 2013. Prinsys - On a Quest for Probabilistic Loop Invariants. In *Quantitative Evaluation of Systems - 10th International Conference, QEST 2013, Buenos Aires, Argentina, August 27-30, 2013. Proceedings (Lecture Notes in Computer Science, Vol. 8054)*, Kaustubh R. Joshi, Markus Siegle, Mariëlle Stoelinga, and Pedro R. D'Argenio (Eds.). Springer, 193–208. https://doi.org/10.1007/978-3-642-40196-1_17
- [35] Gurobi Optimization, L.: [n.d.]. *Gurobi Optimizer Reference Manual (2023)*. <http://www.gurobi.com>
- [36] Marcel Hark, Benjamin Lucien Kaminski, Jürgen Giesl, and Joost-Pieter Katoen. 2020. Aiming low is harder: induction for lower bounds in probabilistic program verification. *Proc. ACM Program. Lang.* 4, POPL (2020), 37:1–37:28. <https://doi.org/10.1145/3371105>
- [37] Benjamin Lucien Kaminski, Joost-Pieter Katoen, Christoph Matheja, and Federico Olmedo. 2016. Weakest Precondition Reasoning for Expected Run-Times of Probabilistic Programs. In *Programming Languages and Systems - 25th European Symposium on Programming, ESOP 2016, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2016, Eindhoven, The Netherlands, April 2-8, 2016, Proceedings (Lecture Notes in Computer Science, Vol. 9632)*, Peter Thiemann (Ed.). Springer, 364–389. https://doi.org/10.1007/978-3-662-49498-1_15
- [38] Benjamin Lucien Kaminski, Joost-Pieter Katoen, Christoph Matheja, and Federico Olmedo. 2018. Weakest Precondition Reasoning for Expected Runtimes of Randomized Algorithms. *J. ACM* 65, 5 (2018), 30:1–30:68. <https://doi.org/10.1145/3208102>
- [39] Joost-Pieter Katoen, Annabelle McIver, Larissa Meinicke, and Carroll C. Morgan. 2010. Linear-Invariant Generation for Probabilistic Programs: - Automated Support for Proof-Based Methods. In *Static Analysis - 17th International Symposium, SAS 2010, Perpignan, France, September 14-16, 2010. Proceedings (Lecture Notes in Computer Science, Vol. 6337)*, Radhia Cousot and Matthieu Martel (Eds.). Springer, 390–406. https://doi.org/10.1007/978-3-642-15769-1_24
- [40] Andrey Kofnov, Marcel Moosbrugger, Miroslav Stankovic, Ezio Bartocci, and Efstathia Bura. 2022. Moment-Based Invariants for Probabilistic Loops with Non-polynomial Assignments. In *Quantitative Evaluation of Systems - 19th International Conference, QEST 2022, Warsaw, Poland, September 12-16, 2022, Proceedings (Lecture Notes in Computer Science, Vol. 13479)*, Erika Ábrahám and Marco Paolieri (Eds.). Springer, 3–25. https://doi.org/10.1007/978-3-031-16336-4_1
- [41] Andrey Kofnov, Marcel Moosbrugger, Miroslav Stankovic, Ezio Bartocci, and Efstathia Bura. 2023. Exact and Approximate Moment Derivation for Probabilistic Loops With Non-Polynomial Assignments. *CoRR* abs/2306.07072 (2023). <https://doi.org/10.48550/ARXIV.2306.07072> arXiv:2306.07072
- [42] Dexter Kozen. 1981. Semantics of Probabilistic Programs. *J. Comput. Syst. Sci.* 22, 3 (1981), 328–350. [https://doi.org/10.1016/0022-0000\(81\)90036-2](https://doi.org/10.1016/0022-0000(81)90036-2)
- [43] Hari Govind Vadiramana Krishnan, Yakir Vizel, Vijay Ganesh, and Arie Gurfinkel. 2019. Interpolating Strong Induction. In *Computer Aided Verification - 31st International Conference, CAV 2019, New York City, NY, USA, July 15-18, 2019, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 11562)*, Isil Dillig and Serdar Tasiran (Eds.). Springer, 367–385. https://doi.org/10.1007/978-3-030-25543-5_21
- [44] Satoshi Kura, Natsuki Urabe, and Ichiro Hasuo. 2019. Tail Probabilities for Randomized Program Runtimes via Martingales for Higher Moments. In *Tools and Algorithms for the Construction and Analysis of Systems - 25th International Conference, TACAS 2019, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2019, Prague, Czech Republic, April 6-11, 2019, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 11428)*, Tomás Vojnar and Lijun Zhang (Eds.). Springer, 135–153. https://doi.org/10.1007/978-3-030-17465-1_8
- [45] Jia Lu and Ming Xu. 2022. Bisection Value Iteration. In *29th Asia-Pacific Software Engineering Conference, APSEC 2022, Virtual Event, Japan, December 6-9, 2022*. IEEE, 109–118. <https://doi.org/10.1109/APSEC57359.2022.00023>
- [46] Garth P. McCormick. 1976. Computability of global solutions to factorable nonconvex programs: Part I - Convex underestimating problems. *Math. Program.* 10, 1 (1976), 147–175. <https://doi.org/10.1007/BF01580665>
- [47] Antoine Miné. 2004. Relational Abstract Domains for the Detection of Floating-Point Run-Time Errors. In *Programming Languages and Systems, 13th European Symposium on Programming, ESOP 2004, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2004, Barcelona, Spain, March 29 - April 2, 2004, Proceedings (Lecture Notes in Computer Science, Vol. 2986)*, David A. Schmidt (Ed.). Springer, 3–17. https://doi.org/10.1007/978-3-540-24725-8_2
- [48] Antoine Miné. 2006. Symbolic Methods to Enhance the Precision of Numerical Abstract Domains. In *Verification, Model Checking, and Abstract Interpretation, 7th International Conference, VMAI 2006, Charleston, SC, USA, January*

- 8-10, 2006, *Proceedings (Lecture Notes in Computer Science, Vol. 3855)*, E. Allen Emerson and Kedar S. Namjoshi (Eds.). Springer, 348–363. https://doi.org/10.1007/11609773_23
- [49] Marcel Moosbrugger, Miroslav Stankovic, Ezio Bartocci, and Laura Kovács. 2022. This is the moment for probabilistic loops. *Proc. ACM Program. Lang.* 6, OOPSLA2 (2022), 1497–1525. <https://doi.org/10.1145/3563341>
- [50] Theodore Samuel Motzkin. 1936. Beiträge zur Theorie der linearen Ungleichungen. (*No Title*) (1936).
- [51] Van Chan Ngo, Quentin Carbonneaux, and Jan Hoffmann. 2018. Bounded expectations: resource analysis for probabilistic programs. In *Proceedings of the 39th ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI 2018, Philadelphia, PA, USA, June 18-22, 2018*, Jeffrey S. Foster and Dan Grossman (Eds.). ACM, 496–512. <https://doi.org/10.1145/3192366.3192394>
- [52] D. PARK. 1969. Fixpoint Induction and Proofs of Program Properties. *Machine Intelligence* 5 (1969). <https://cir.nii.ac.jp/crid/1573950399497019904>
- [53] Mihai Putinar. 1993. Positive Polynomials on Compact Semi-algebraic Sets. *Indiana University Mathematics Journal* 42, 3 (1993), 969–984. <http://www.jstor.org/stable/24897130>
- [54] Sriram Sankaranarayanan, Aleksandar Chakarov, and Sumit Gulwani. 2013. Static analysis for probabilistic programs: inferring whole program properties from finitely many paths. In *ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI '13, Seattle, WA, USA, June 16-19, 2013*, Hans-Juergen Boehm and Cormac Flanagan (Eds.). ACM, 447–458. <https://doi.org/10.1145/2491956.2462179>
- [55] Sriram Sankaranarayanan, Henny B. Sipma, and Zohar Manna. 2004. Constraint-Based Linear-Relations Analysis. In *Static Analysis, 11th International Symposium, SAS 2004, Verona, Italy, August 26-28, 2004, Proceedings (Lecture Notes in Computer Science, Vol. 3148)*, Roberto Giacobazzi (Ed.). Springer, 53–68. https://doi.org/10.1007/978-3-540-27864-1_7
- [56] Mary Sheeran, Satnam Singh, and Gunnar Stålmarck. 2000. Checking Safety Properties Using Induction and a SAT-Solver. In *Formal Methods in Computer-Aided Design, Third International Conference, FMCAD 2000, Austin, Texas, USA, November 1-3, 2000, Proceedings (Lecture Notes in Computer Science, Vol. 1954)*, Warren A. Hunt Jr. and Steven D. Johnson (Eds.). Springer, 108–125. https://doi.org/10.1007/3-540-40922-X_8
- [57] Calvin Smith, Justin Hsu, and Aws Albarghouthi. 2019. Trace abstraction modulo probability. *Proc. ACM Program. Lang.* 3, POPL (2019), 39:1–39:31. <https://doi.org/10.1145/3290352>
- [58] Toru Takisaka, Yuichiro Oyabu, Natsuki Urabe, and Ichiro Hasuo. 2021. Ranking and Repulsing Supermartingales for Reachability in Randomized Programs. *ACM Trans. Program. Lang. Syst.* 43, 2 (2021), 5:1–5:46. <https://doi.org/10.1145/3450967>
- [59] Jan-Willem van de Meent, Brooks Paige, Hongseok Yang, and Frank Wood. 2018. An Introduction to Probabilistic Programming. *CoRR abs/1809.10756* (2018). arXiv:1809.10756 <http://arxiv.org/abs/1809.10756>
- [60] Di Wang, Jan Hoffmann, and Thomas W. Reps. 2021. Central moment analysis for cost accumulators in probabilistic programs. In *PLDI '21: 42nd ACM SIGPLAN International Conference on Programming Language Design and Implementation, Virtual Event, Canada, June 20-25, 2021*, Stephen N. Freund and Eran Yahav (Eds.). ACM, 559–573. <https://doi.org/10.1145/3453483.3454062>
- [61] Di Wang, Jan Hoffmann, and Thomas W. Reps. 2021. Expected-Cost Analysis for Probabilistic Programs and Semantics-Level Adaption of Optional Stopping Theorems. *CoRR abs/2103.16105* (2021). arXiv:2103.16105 <https://arxiv.org/abs/2103.16105>
- [62] Jinyi Wang, Yican Sun, Hongfei Fu, Krishnendu Chatterjee, and Amir Kafshdar Goharshady. 2021. Quantitative analysis of assertion violations in probabilistic programs. In *PLDI '21: 42nd ACM SIGPLAN International Conference on Programming Language Design and Implementation, Virtual Event, Canada, June 20-25, 2021*, Stephen N. Freund and Eran Yahav (Eds.). ACM, 1171–1186. <https://doi.org/10.1145/3453483.3454102>
- [63] Peixin Wang, Hongfei Fu, Amir Kafshdar Goharshady, Krishnendu Chatterjee, Xudong Qin, and Wenjun Shi. 2019. Cost analysis of nondeterministic probabilistic programs. In *Proceedings of the 40th ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI 2019, Phoenix, AZ, USA, June 22-26, 2019*, Kathryn S. McKinley and Kathleen Fisher (Eds.). ACM, 204–220. <https://doi.org/10.1145/3314221.3314581>
- [64] Peixin Wang, Hongfei Fu, Tengshun Yang, Guanyan Li, and Luke Ong. 2023. Template-Based Static Posterior Inference for Bayesian Probabilistic Programming. *CoRR abs/2307.13160* (2023). <https://doi.org/10.48550/ARXIV.2307.13160> arXiv:2307.13160
- [65] Qiuye Wang, Mingshuai Chen, Bai Xue, Naijun Zhan, and Joost-Pieter Katoen. 2021. Synthesizing Invariant Barrier Certificates via Difference-of-Convex Programming. In *CAV (Lecture Notes in Computer Science, Vol. 12759)*. Springer, 443–466.
- [66] Qiuye Wang, Mingshuai Chen, Bai Xue, Naijun Zhan, and Joost-Pieter Katoen. 2022. Encoding inductive invariants as barrier certificates: Synthesis via difference-of-convex programming. *Information and Computation* 289, Part (2022), 104965.
- [67] David Williams. 1991. *Probability with Martingales*. Cambridge University Press.

A SUPPLEMENTARY MATERIAL FOR SECTION 3

A.1 Properties of the Upper k -Induction Operator in Lu and Xu [45]

LEMMA A.1. *Let Ψ be the k -induction operator in Lu and Xu [45] w.r.t. Φ . Then*

- (1) Ψ is monotonic, i.e., $\forall v_1, v_2 \in E, v_1 \sqsubseteq v_2$ implies $\Psi(v_1) \sqsubseteq \Psi(v_2)$.
- (2) Iterations of Ψ starting from u are descending, i.e.,

$$\dots \sqsubseteq \Psi^k(u) \sqsubseteq \Psi^{k-1}(u) \sqsubseteq \dots \sqsubseteq \Psi(u) \sqsubseteq u$$

And thus we have for all $m < n \in \mathbb{N}$, $\Psi^n(u) \sqsubseteq \Psi^m(u)$.

PROOF. For item 1, observe that if we have $w_1 \sqsubseteq w_2$ and $v_1 \sqsubseteq v_2$, then we have $w_1 \sqcap v_1 \sqsubseteq w_2 \sqcap v_2$.

$$\begin{aligned} \Psi(v_1) &= \Phi(v_1) \sqcap v_1 && \text{(by definition of } \Psi) \\ &\sqsubseteq \Phi(v_2) \sqcap v_2 && \text{(by monotonicity of } \Phi \text{ and above property)} \\ &= \Psi(v_2) && \text{(by definition of } \Psi) \end{aligned}$$

For item 2, we can immediately derived from the definition of Ψ as

$$\begin{aligned} \Psi^k(u) &= \Psi(\Psi^{k-1}(u)) && \text{(by definition of } \Psi^k(u)) \\ &= \Phi(\Psi^{k-1}(u)) \sqcap \Psi^{k-1}(u) && \text{(by definition of } \Psi) \\ &\sqsubseteq \Psi^{k-1}(u) && \text{(by definition of } \sqcap) \end{aligned}$$

□

Proposition 3.4. For any element $u \in E$, we have that

$$\Phi(\Psi^k(u)) \sqsubseteq u \iff \Phi(\Psi^k(u)) \sqsubseteq \Psi^k(u)$$

PROOF. The if-direction is trivial as $\Psi^k(u) \sqsubseteq u$ (by Lemma A.1(2)). For the only-if direction:

$$\begin{aligned} \Psi^k(u) &\sqsupseteq \Psi^{k+1}(u) && \text{(by Lemma A.1(2))} \\ &= \Phi(\Psi^k(u)) \sqcap \Psi^k(u) && \text{(by definition of } \Psi) \\ &= \Phi(\Psi^k(u)) \sqcap \Psi(\Psi^{k-1}(u)) && \text{(by definition of } \Psi^k(u)) \\ &= \Phi(\Psi^k(u)) \sqcap (\Phi(\Psi^{k-1}(u)) \sqcap \Psi^{k-1}(u)) && \text{(by definition of } \Psi) \\ &= (\Phi(\Psi^k(u)) \sqcap \Phi(\Psi^{k-1}(u))) \sqcap \Psi^{k-1}(u) && \text{(by associative law)} \\ &= \Phi(\Psi^k(u)) \sqcap \Psi^{k-1}(u) && \text{(by monotonicity of } \Phi \text{ and Lemma A.1(2))} \\ &\vdots \\ &= (\Phi(\Psi^k(u)) \sqcap \Phi(u)) \sqcap u && \text{(by unfolding } \Psi^k \text{ until } k = 1) \\ &= \Phi(\Psi^k(u)) \sqcap u && \text{(by monotonicity of } \Phi \text{ and Lemma A.1(2))} \\ &= \Phi(\Psi^k(u)) && \text{(by the premise)} \end{aligned}$$

□

A.2 Equivalence between Ψ_u and Ψ

Theorem 3.5. [Equivalence between Ψ_u and Ψ] For any element $u \in E$, the sequence $\{\Psi_u^k(u)\}_{k \geq 0}$ of elements in E coincides with the sequence $\{\Psi^k(u)\}_{k \geq 0}$. In other words, for any natural number $k \geq 0$, we have that $\Psi_u^k(u) = \Psi^k(u)$.

PROOF. Proof by mathematical induction. We denote $X_k = \Psi_u^k(u)$ and $Y_k = \Psi^k(u)$. when $k = 0$, $X_0 = u = Y_0$. When $k = 1$, $X_1 = \Phi(u) \sqcap u = Y_1$, by definition of two operators, respectively.

Now we suppose that $X_k = Y_k$, i.e., $\Psi_u^k(u) = \Psi^k(u)$, and we aim to prove that $\Psi_u^{k+1}(u) = \Psi^{k+1}(u)$.

$$\begin{aligned}
 X_{k+1} &= \Psi_u(\Psi_u^k(u)) && \text{(by definition of } \Psi_u^{k+1}(u) \text{)} \\
 &= \Phi(\Psi_u^k(u)) \sqcap u && \text{(by definition of } \Psi_u \text{)} \\
 Y_{k+1} &= \Psi(\Psi^k(u)) && \text{(by definition of } \Psi^{k+1}(u) \text{)} \\
 &= \Phi(\Psi^k(u)) \sqcap \Psi^k(u) && \text{(by definition of } \Psi \text{)} \\
 &= \Phi(\Psi^k(u)) \sqcap \Psi(\Psi^{k-1}(u)) && \text{(by definition of } \Psi^k(u) \text{)} \\
 &= \Phi(\Psi^k(u)) \sqcap (\Phi(\Psi^{k-1}(u)) \sqcap \Psi^{k-1}(u)) && \text{(by definition of } \Psi \text{)} \\
 &= (\Phi(\Psi^k(u)) \sqcap \Phi(\Psi^{k-1}(u))) \sqcap \Psi^{k-1}(u) && \text{(by associative law)} \\
 &= \Phi(\Psi^k(u)) \sqcap \Psi^{k-1}(u) && \text{(by monotonicity of } \Phi \text{ and Lemma A.1(2))} \\
 &\vdots \\
 &= (\Phi(\Psi^k(u)) \sqcap \Phi(u)) \sqcap u && \text{(by unfolding } \Psi^k \text{ until } k = 1 \text{)} \\
 &= \Phi(\Psi^k(u)) \sqcap u && \text{(by monotonicity of } \Phi \text{ and Lemma A.1(2))}
 \end{aligned}$$

Since we suppose that $\Psi_u^k(u) = \Psi^k(u)$, we obtain that $\Phi(\Psi_u^k(u)) \sqcap u = \Phi(\Psi^k(u)) \sqcap u$, thus we have $\Psi_u^{k+1}(u) = \Psi^{k+1}(u)$, i.e., $X_{k+1} = Y_{k+1}$. $\square$

A.3 Supplementary Properties for the Dual k -Induction Operators Ψ'_u and Ψ

LEMMA A.2. Fix a lattice $(E, \sqsubseteq)$ and a monotone operator Φ . For any element $u \in E$, both of these two dual k -induction operators Ψ'_u and Ψ' have the following properties:

- (1) Ψ'_u (resp. Ψ') is monotone.
- (2) Iterations of Ψ'_u (resp. Ψ') starting from u are ascending, i.e.,

$$u \sqsubseteq \Psi'_u(u) \sqsubseteq \dots (\Psi'_u)^{k-1}(u) \sqsubseteq (\Psi'_u)^k(u) \dots$$

$$u \sqsubseteq \Psi'(u) \sqsubseteq \dots (\Psi')^{k-1}(u) \sqsubseteq (\Psi')^k(u) \dots$$

Thus we have for all $m < n \in \mathbb{N}$, $(\Psi'_u)^m(u) \sqsubseteq (\Psi'_u)^n(u)$ and $(\Psi')^m(u) \sqsubseteq (\Psi')^n(u)$.

PROOF. We only prove the case of dual k -induction operator Ψ'_u , since the proof of the properties of the dual k -induction operator Ψ' is similar with that of Ψ'_u .

For item 1, observe that if we have $w_1 \sqsubseteq w_2$, then we have $w_1 \sqcup u \sqsubseteq w_2 \sqcup u$. Assume that $v_1 \sqsubseteq v_2$

$$\begin{aligned}
 \Psi'_u(v_1) &= \Phi(v_1) \sqcup u && \text{(by definition of } \Psi'_u \text{)} \\
 &\sqsubseteq \Phi(v_2) \sqcup u && \text{(by monotonicity of } \Phi \text{ and above property)} \\
 &= \Psi'_u(v_2) && \text{(by definition of } \Psi'_u \text{)}
 \end{aligned}$$

For item 2, we prove it by mathematical induction. We have $u \sqsubseteq \Psi'_u(u)$ as $\Psi'_u(u) = \Phi(u) \sqcup u$. We then assume that $(\Psi'_u)^k(u) \sqsupseteq (\Psi'_u)^{k-1}(u)$, and we prove that

$$\begin{aligned} (\Psi'_u)^{k+1}(u) &= \Psi'_u((\Psi'_u)^k(u)) && \text{(by definition of } (\Psi'_u)^{k+1}(u)) \\ &\sqsupseteq \Psi'_u((\Psi'_u)^{k-1}(u)) && \text{(by monotonicity of } \Psi'_u \text{ and assumption)} \\ &= (\Psi'_u)^k(u) && \text{(by definition of } (\Psi'_u)^k(u)) \end{aligned}$$

Thus the value sequence is an ascending chain. $\square$

Proposition 3.8. For any element $u \in E$, both of these two dual k -induction operators Ψ'_u and Ψ' have the following properties:

$$\begin{aligned} \Phi((\Psi'_u)^k(u)) \sqsupseteq u &\iff \Phi((\Psi'_u)^k(u)) \sqsupseteq (\Psi'_u)^k(u) \\ \Phi((\Psi')^k(u)) \sqsupseteq u &\iff \Phi((\Psi')^k(u)) \sqsupseteq (\Psi')^k(u) \end{aligned}$$

PROOF. For the case of the dual k -induction operator Ψ'_u :

The if-direction is trivial as $(\Psi'_u)^k(u) \sqsupseteq u$ (by Lemma A.2(2)). For the only-if direction:

$$\begin{aligned} (\Psi'_u)^k(u) &\sqsubseteq (\Psi'_u)^{k+1}(u) && \text{(by Lemma A.2(2))} \\ &= \Psi'_u((\Psi'_u)^k(u)) && \text{(by the definition of } (\Psi'_u)^{k+1}(u)) \\ &= \Phi((\Psi'_u)^k(u)) \sqcup u && \text{(by the definition of } \Psi'_u) \\ &= \Psi((\Psi'_u)^k(u)) && \text{(by the premise)} \end{aligned}$$

For the case of the dual k -induction operator Ψ' :

The if-direction is trivial as $(\Psi')^k(u) \sqsupseteq u$ (by Lemma A.2(2)). For the only-if direction:

$$\begin{aligned} (\Psi')^k(u) &\sqsubseteq (\Psi')^{k+1}(u) && \text{(by Lemma A.2(2))} \\ &= \Psi'((\Psi')^k(u)) && \text{(by the definition of } (\Psi')^{k+1}(u)) \\ &= \Phi((\Psi')^k(u)) \sqcup (\Psi')^k(u) && \text{(by the definition of } \Psi') \\ &= \Phi((\Psi')^k(u)) \sqcup \Psi'((\Psi')^{k-1}(u)) && \text{(by the definition of } (\Psi')^k(u)) \\ &= \Phi((\Psi')^k(u)) \sqcup \Phi((\Psi')^{k-1}(u)) \sqcup (\Psi')^{k-1}(u) && \text{(by the definition of } \Psi') \\ &= (\Phi((\Psi')^k(u)) \sqcup \Phi((\Psi')^{k-1}(u))) \sqcup (\Psi')^{k-1}(u) && \text{(by associate law)} \\ &= \Phi((\Psi')^k(u)) \sqcup (\Psi')^{k-1}(u) && \text{(by monotonicity of } \Phi \text{ and Lemma A.2(2))} \\ &\vdots \\ &= \Phi((\Psi')^k(u)) \sqcup \Psi'(u) && \text{(by unfolding } (\Psi')^k(u) \text{ until } k = 1) \\ &= \Phi((\Psi')^k(u)) \sqcup \Phi(u) \sqcup u && \text{(by definition of } \Psi') \\ &= \Phi((\Psi')^k(u)) \sqcup u && \text{(by monotonicity of } \Phi \text{ and Lemma A.2(2))} \\ &= \Phi((\Psi')^k(u)) && \text{(by the premise)} \end{aligned}$$

$\square$

Theorem 3.9. [Equivalence between Ψ'_u and Ψ'] For any element $u \in E$, we have that the sequence $\{(\Psi'_u)^k(u)\}_{k \geq 0}$ of elements in E coincides with the sequence $\{(\Psi')^k(u)\}_{k \geq 0}$. In other words, for any natural number $k \geq 0$, we have that $(\Psi'_u)^k(u) = (\Psi')^k(u)$.

PROOF. Analogously, we proof it by mathematical induction. $X_k = (\Psi'_u)^k(u)$ and $Y_k = (\Psi')^k(u)$, when $k = 0$, $X_0 = u = Y_0$. When $k = 1$, $X_1 = \Phi(u) \sqcup u = Y_1$, by definition of two dual operators, respectively.

Now we suppose that $X_k = Y_k$, i.e., $(\Psi'_u)^k(u) = (\Psi')^k(u)$, and we aim to prove that $(\Psi'_u)^{k+1}(u) = (\Psi')^{k+1}(u)$.

$$\begin{aligned}
 X_{k+1} &= \Psi'_u((\Psi'_u)^k(u)) && \text{(by definition of } (\Psi'_u)^{k+1}(u) \text{)} \\
 &= \Phi((\Psi'_u)^k(u)) \sqcup u && \text{(by definition of } \Psi'_u \text{)} \\
 Y_{k+1} &= \Psi'((\Psi')^k(u)) && \text{(by definition of } (\Psi')^{k+1}(u) \text{)} \\
 &= \Phi((\Psi')^k(u)) \sqcup (\Psi')^k(u) && \text{(by definition of } \Psi' \text{)} \\
 &= \Phi((\Psi')^k(u)) \sqcup \Psi'((\Psi')^{k-1}(u)) && \text{(by definition of } (\Psi')^k(u) \text{)} \\
 &= \Phi((\Psi')^k(u)) \sqcup (\Phi((\Psi')^{k-1}(u)) \sqcup \Psi'^{k-1}(u)) && \text{(by definition of } \Psi' \text{)} \\
 &= (\Phi((\Psi')^k(u)) \sqcup \Phi((\Psi')^{k-1}(u))) \sqcup \Psi'^{k-1}(u) && \text{(by associative law)} \\
 &= \Phi((\Psi')^k(u)) \sqcup (\Psi')^{k-1}(u) && \text{(by monotonicity of } \Phi \text{ and Lemma A.2(2))} \\
 &\vdots \\
 &= (\Phi((\Psi')^k(u)) \sqcup \Phi(u)) \sqcup u && \text{(by unfolding } (\Psi')^k \text{ until } k = 1 \text{)} \\
 &= \Phi((\Psi')^k(u)) \sqcup u && \text{(by monotonicity of } \Phi \text{ and Lemma A.2(2))}
 \end{aligned}$$

Since we suppose that $(\Psi'_u)^k(u) = (\Psi')^k(u)$, we obtain that $\Phi((\Psi'_u)^k(u) \sqcup u) = \Phi((\Psi')^k(u)) \sqcup u$, thus we have $(\Psi'_u)^{k+1}(u) = (\Psi')^{k+1}(u)$, i.e., $X_{k+1} = Y_{k+1}$. $\square$

B SUPPLEMENTARY MATERIAL FOR SECTION 4

B.1 Classical OST

Optional Stopping Theorem (OST) is a classical theorem in martingale theory that characterizes the relationship between the expected values initially and at a stopping time in a supermartingale. Below we present the classical form of OST.

THEOREM B.1 (OPTIONAL STOPPING THEOREM (OST) [67, CHAPTER 10]). *Let $\{X_n\}_{n=0}^\infty$ be a martingale (resp. supermartingale) adapted to a filtration $\mathcal{F} = \{\mathcal{F}_n\}_{n=0}^\infty$ and τ be a stopping time w.r.t the filtration $\mathcal{F}$. If we have that:*

- $\mathbb{E}(\tau) < \infty$;
- *exists an $M \in [0, \infty)$ such that $|X_{n+1} - X_n| \leq M$ holds almost surely for every $n \leq 0$,*

then it follows that $(|X_\tau|) < \infty$ and $\mathbb{E}(X_\tau) = \mathbb{E}(X_0)$ (resp. $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$).

B.2 Proof of Extended OST (Theorem 4.6)

In this section, we attach the proof for this variant of Optional Stopping Theorem in Wang et al. [64] here.

Theorem 4.6 [Extended OST in Wang et al. [64]] *Let $\{X_n\}_{n=0}^\infty$ be a supermartingale adapted to a filtration $\mathcal{F} = \{\mathcal{F}_n\}_{n=0}^\infty$ and τ be a stopping time w.r.t the filtration $\mathcal{F}$. Suppose there exist positive real numbers b_1, b_2, c_1, c_2, c_3 such that $c_2 > c_3$ and*

- *Concentration property.* For all sufficiently large natural numbers n , it holds that $\mathbb{P}(\tau > n) \leq c_1 \cdot \exp(-c_2 \cdot n)$.

- (b) *Exponential bound.* For every natural number $n \geq 0$, it holds almost-surely that $|X_{n+1} - X_n| \leq b_1 \cdot n^{b_2} \cdot e^{c_3 \cdot n}$.

Then we have that $\mathbb{E}(|X_\tau|) < \infty$ and $\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$.

PROOF. For every $n \in \mathbb{N}_0$,

$$\begin{aligned}
 |X_{\tau \wedge n}| &= \left| X_0 + \sum_{k=0}^{\tau \wedge n - 1} (X_{k+1} - X_k) \right| \\
 &= \left| X_0 + \sum_{k=0}^{\infty} (X_{k+1} - X_k) \cdot \mathbf{1}_{\tau > k \wedge n > k} \right| \\
 &\leq |X_0| + \sum_{k=0}^{\infty} |(X_{k+1} - X_k) \cdot \mathbf{1}_{\tau > k \wedge n > k}| \\
 &\leq |X_0| + \sum_{k=0}^{\infty} |(X_{k+1} - X_k) \cdot \mathbf{1}_{\tau > k}| .
 \end{aligned}$$

Then

$$\begin{aligned}
 &\mathbb{E} \left(|X_0| + \sum_{k=0}^{\infty} |(X_{k+1} - X_k) \cdot \mathbf{1}_{\tau > k}| \right) \\
 &= \mathbb{E}(|X_0|) + \sum_{k=0}^{\infty} \mathbb{E}(|(X_{k+1} - X_k) \cdot \mathbf{1}_{\tau > k}|) \quad (\text{By Monotone Convergence Theorem}) \\
 &= \mathbb{E}(|X_0|) + \sum_{k=0}^{\infty} \mathbb{E}(|X_{k+1} - X_k| \cdot \mathbf{1}_{\tau > k}) \\
 &\leq \mathbb{E}(|X_0|) + \sum_{k=0}^{\infty} \mathbb{E}(b_1 \cdot k^{b_2} \cdot e^{c_3 \cdot k} \cdot \mathbf{1}_{\tau > k}) \quad (\text{by condition (b)}) \\
 &= \mathbb{E}(|X_0|) + \sum_{k=0}^{\infty} b_1 \cdot k^{b_2} \cdot e^{c_3 \cdot k} \cdot \mathbb{P}(\tau > k) \\
 &\leq \mathbb{E}(|X_0|) + \sum_{k=0}^{\infty} b_1 \cdot k^{b_2} \cdot e^{c_3 \cdot k} \cdot c_1 \cdot e^{-c_2 \cdot k} \quad (\text{by condition (a)}) \\
 &= \mathbb{E}(|X_0|) + b_1 \cdot c_1 \cdot \sum_{k=0}^{\infty} k^{b_2} \cdot e^{-(c_2 - c_3) \cdot k} \\
 &< \infty . \quad (\text{by premise})
 \end{aligned}$$

Therefore, by Dominated Convergence Theorem and the fact that $X_\tau = \lim_{n \rightarrow \infty} X_{\tau \wedge n}$ a.s.,

$$\mathbb{E}(X_\tau) = \mathbb{E} \left(\lim_{n \rightarrow \infty} X_{\tau \wedge n} \right) = \lim_{n \rightarrow \infty} \mathbb{E}(X_{\tau \wedge n}) .$$

Finally, the result follows from properties for the stopped process $\{X_{\tau \wedge n}\}_{n \in \mathbb{N}_0}$ that

$$\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0) .$$

□

B.3 Proof of Theorem 4.7

Theorem 4.7. [Soundness of Potential Functions] Let k be a positive integer. Suppose that there exist real numbers $c_1 > 0$ and $c_2 > c_3 > 0$ such that (i) The maximum amplifier c satisfies $c \leq e^{c_3}$ and (ii) the termination time random variable τ of P has the concentration property, i.e., $\mathbb{P}(\tau > n) \leq c_1 \cdot e^{-c_2 \cdot n}$. Then the following hold:

- For any k -upper potential function h , $\mathbb{E}_{s^*}(X_f) \leq \overline{\Psi}_h^{k-1}(h)(s^*) \leq h(s^*)$.
- For any k -lower potential function h , $\mathbb{E}_{s^*}(X_f) \geq (\overline{\Psi}'_h)^{k-1}(h)(s^*) \geq h(s^*)$.

PROOF. We first proof the soundness of upper potential functions. Let s_n be the random vector (random variable) of the program state at the n -th iteration of the probabilistic while loop P , where $s_0 = s^*$ and let $H = \overline{\Psi}_h^k(h)$. By Definition 4.5 and Theorem 3.2, we obtain that $\forall s \in \text{Reach}(s^*)$, $\overline{\Phi}(H)(s) \leq H(s)$. We define the stochastic process $\{X_n\}_{n=0}^\infty$ by

$$X_n := [s_n \models \varphi] \cdot H(s_n) + [s_n \not\models \varphi] \cdot f(s_n).$$

We first prove that the stochastic process $\{X_n\}$ is a supermartingale. We discuss this in the following two scenarios:

- if $s_n \not\models \varphi$, by the semantics of probabilistic while loop (see Section 2.2), $s_{n+1} = s_n$, and thus $X_{n+1} = X_n$, which satisfies the conditions of supermartingale;
- if $s_n \models \varphi$, we have

$$\begin{aligned} \mathbb{E}_{s^*}[X_{n+1} | \mathcal{F}_n] &= \mathbb{E}_{s^*}[[s_{n+1} \models \varphi] \cdot H(s_{n+1}) + [s_{n+1} \not\models \varphi] \cdot f(s_{n+1})] \\ &= [s_{n+1} \models \varphi] \cdot \mathbb{E}_{s^*}[H(s_{n+1})] + [s_{n+1} \not\models \varphi] \cdot \mathbb{E}_{s^*}[f(s_{n+1})] \\ &\quad \text{(by definition of mathematic expectation)} \\ &= [s_{n+1} \models \varphi] \cdot \text{pre}_C(H)(s_n) + [s_{n+1} \not\models \varphi] \cdot \mathbb{E}_{s^*}[f(s_{n+1})] \\ &\quad \text{(by definition of pre-expectation)} \\ &= \overline{\Phi}(H)(s_n) \\ &\leq H(s_n) \quad \text{(by property of } H) \\ &= X_n \end{aligned}$$

Combining the condition (i), (ii) and H is piecewise linear, we can derive that $\mathbb{E}_{s^*}[X_n] < \infty$ holds. Thus $\{X_n\}$ is a supermartingale.

The condition(a) depends on the assumption that (ii) P has the concentration property.

Then we prove the condition (b). We discuss this in the following three scenarios:

- if $s_n \not\models \varphi$, by the semantics of probabilistic while loop (see Section 2.2), $s_{n+1} = s_n$, and thus $|X_{n+1} - X_n| = 0$;
- if $s_n \models \varphi$ and if $s_{n+1} \models \varphi$, since the condition (ii), we have that each program variable $x_i (i \in \mathbb{Z}^+)$ and the constant term x_0 at state $s_n (\forall n)$ can be bounded by $K_i \cdot c_i^n$ for some $K_i, c_i (i \in \mathbb{N})$. In addition that H is piecewise linear, we have that $H(s_n) \leq M_n \cdot c^n$ for $M_n > 0$.

$$\begin{aligned} |X_{n+1} - X_n| &= |H(s_{n+1}) - H(s_n)| \\ &\leq |H(s_{n+1})| + |H(s_n)| \\ &\leq M_n \cdot |c|^n + M_{n+1} \cdot |c|^{n+1} \\ &\leq (M_n + |c| \cdot M_{n+1}) \cdot |c|^n \\ &\leq b_1 \cdot e^{c_3 n} \end{aligned}$$

- if $s_n \models \varphi$ and if $s_{n+1} \not\models \varphi$, this case is similar with the case of $s_n \models \varphi$ & $s_{n+1} \models \varphi$. We have that

$$\begin{aligned}
 |X_{n+1} - X_n| &= |f(s_{n+1}) - H(s_n)| \\
 &\leq |f(s_{n+1})| + |H(s_n)| \\
 &\leq M_n \cdot |c|^n + M_{n+1} \\
 &= M_n \cdot |c|^n + M_{n+1} \cdot e^{0 \cdot n} \\
 &\leq b_1 \cdot e^{c_3 n}
 \end{aligned}$$

By applying Theorem 4.6, we have that $\mathbb{E}_{s^*}(X_\tau) \leq \mathbb{E}_{s^*}(X_0)$. Since τ is the stopping time, there will be $s_\tau \not\models \varphi$, thus $X_\tau = f(s_\tau) = X_f$. We have $\mathbb{E}_{s^*}(X_f) \leq \mathbb{E}_{s^*}(X_0) = H(s^*)$. The second inequality can be derived directly from the property that $\bar{\Psi}_h^k(h) \leq h$ holds (see Appendix A.1 and Batz et al. [12]). $\square$

The case of lower potential functions is completely dual to that of upper potential functions since we can consider the stochastic process $\{-X_n\}$, that is, define the stochastic process by

$$Y_n := [s_n \models \varphi] \cdot (-H(s_n)) + [s_n \not\models \varphi] \cdot (-f(s_n)).$$

The remaining proof is essentially the same. $\square$

C SUPPLEMENTARY MATERIAL FOR SECTION 5

C.1 Proof of Proposition 5.3

We give a proof for Proposition 5.3 in this section.

Proposition 5.3 The upper (resp. lower) k -induction condition $\bar{\Phi}_f(\bar{\Psi}_h^{k-1}(h)) \leq h$ (resp. $\bar{\Phi}_f(\bar{\Psi}_h^{k-1}(h)) \geq h$) is equivalent with $\min\{h_1, h_2, \dots, h_m\} \leq h$ (resp. $\max\{h_1, h_2, \dots, h_m\} \geq h$), where each h_i uniquely corresponds to one $C_d \in \{C_1, \dots, C_m\}$ and is equal to $pre_{C_d}(h)$.

PROOF. We only proof the case of upper k -induction condition, as the lower case is completely dual. Obviously we only need to concentrate on the left side of the constraint: $\bar{\Phi}_f(\bar{\Psi}_h^{k-1}(h)) \leq h$.

We first proof the case of $k = 2$, i.e., $\bar{\Phi}_f(\bar{\Psi}_h^1(h)) \leq h$. Since our syntax of the probabilistic programs is defined in a compositional style (see Fig. 1 in Section 2.2 for more details), we proof by induction on the structure of programs. For simplicity, we denote $pre_C([\Phi])$ by $[\Phi(C)]$, which represent the evaluation of $[\Phi]$ after the execution of C .

- Case $C \equiv \text{skip}$.

$$\begin{aligned}
 &\bar{\Phi}_f(\bar{\Psi}_h(h)) \\
 &= [\neg\varphi] \cdot f + [\varphi] \cdot pre_C(\bar{\Psi}_h(h)) \\
 &= [\neg\varphi] \cdot f + [\varphi] \cdot \bar{\Psi}_h(h) \\
 &= [\neg\varphi] \cdot f + [\varphi] \cdot \min\{\bar{\Phi}_f(h), h\} \\
 &= [\neg\varphi] \cdot f + [\varphi] \cdot \min\{[\neg\varphi] \cdot f + [\varphi] \cdot h, h\} \\
 &= [\neg\varphi] \cdot f + \min\{[\varphi] \cdot h, [\varphi] \cdot h\} \\
 &= [\neg\varphi] \cdot f + [\varphi] \cdot h \\
 &= \bar{\Phi}_f(h)
 \end{aligned}$$

It corresponds to pre-expectation of the loop-free program unfolded with twice (only one program).

- Case $C \equiv x := e$.

$$\begin{aligned}
 & \bar{\Phi}_f(\bar{\Psi}_h(h)) \\
 = & [\neg\varphi] \cdot f + [\varphi] \cdot \text{pre}_C(\bar{\Psi}_h(h)) \\
 = & [\neg\varphi] \cdot f + [\varphi] \cdot \bar{\Psi}_h(h)([x/e]) \\
 = & [\neg\varphi] \cdot f + [\varphi] \cdot \min\{[\neg\varphi] \cdot f + [\varphi] \cdot h([x/e]), h([x/e])\} \\
 = & [\neg\varphi] \cdot f + [\varphi] \cdot \min\{[\neg\varphi]([x/e]) \cdot f([x/e]) + \\
 & \quad [\varphi]([x/e]) \cdot h([x/e])([x/e]), h([x/e])\} \\
 = & \min\{[\neg\varphi] \cdot f + [\varphi \wedge \neg\varphi]([x/e]) \cdot f([x/e]) + \\
 & \quad [\varphi \wedge \varphi]([x/e]) \cdot h([x/e])([x/e]), [\neg\varphi] \cdot f + [\varphi] \cdot h([x/e])\} \\
 = & \min\{[\neg\varphi] \cdot f + [\varphi \wedge \neg\varphi]([x/e]) \cdot f([x/e]) + \\
 & \quad [\varphi \wedge \varphi]([x/e]) \cdot \text{pre}_{C;C}(h), [\neg\varphi] \cdot f + [\varphi] \cdot h([x/e])\}
 \end{aligned}$$

the expressions in the minimize operator correspond to pre-expectation of the two loop-free programs unfolded within twice (one for once, and another for twice).

- Case $C \equiv C_1; C_2$.

$$\begin{aligned}
 & \bar{\Phi}_f(\bar{\Psi}_h(h)) \\
 = & [\neg\varphi] \cdot f + [\varphi] \cdot \text{pre}_C(\bar{\Psi}_h(h)) \\
 = & [\neg\varphi] \cdot f + [\varphi] \text{pre}_{C_1}(\text{pre}_{C_2}(\min\{[\neg\varphi] \cdot f + [\varphi] \cdot \text{pre}_{C_1}(\text{pre}_{C_2}(h)), h\})) \\
 = & [\neg\varphi] \cdot f + [\varphi] \cdot \min\{[\neg\varphi(C_1; C_2)] \cdot \text{pre}_{C_1; C_2}(f) + \\
 & \quad [\varphi(C_1; C_2)] \cdot \text{pre}_{C_1; C_2}(h), \text{pre}_{C_1; C_2}(h)\} \\
 = & \min\{[\neg\varphi] \cdot f + [\varphi \wedge \varphi(C_1; C_2)] \cdot \text{pre}_{C_1; C_2}(f) + \\
 & \quad [\varphi \wedge \neg\varphi(C_1; C_2)] \cdot \text{pre}_{C_1; C_2}(h), \\
 & \quad [\neg\varphi] \cdot f + [\varphi] \cdot \text{pre}_{C_1; C_2}(h)\}
 \end{aligned}$$

the expressions in the minimize operator correspond to pre-expectation of the two loop-free programs unfolded within twice (one for once, and another for twice)

- case $C \equiv \{C_1\}[p]\{C_2\}$.

$$\begin{aligned}
 & \bar{\Phi}_f(\bar{\Psi}_h(h)) \\
 = & [\neg\varphi] \cdot f + [\varphi] \cdot \text{pre}_C(\bar{\Psi}_h(h)) \\
 = & [\neg\varphi] \cdot f + [\varphi] \cdot p \cdot \text{pre}_{C_1}(\bar{\Psi}_h(h)) + [\varphi] \cdot (1 - p) \cdot \text{pre}_{C_2}(\bar{\Psi}_h(h))
 \end{aligned}$$

wherein

$$\begin{aligned}
 \text{pre}_{C_1}(\bar{\Psi}_h(h)) &= \text{pre}_{C_1}(\min\{[\neg\varphi] \cdot f + [\varphi] \cdot (p \cdot \text{pre}_{C_1}(h) + (1 - p) \cdot \text{pre}_{C_2}(h)), h\}) \\
 &= \min\{[\neg\varphi(C_1)] \cdot \text{pre}_{C_1}(f) + [\varphi(C_1)] \cdot \\
 & \quad (p \cdot \text{pre}_{C_1; C_1}(h) + (1 - p) \cdot \text{pre}_{C_1; C_2}(h)), \text{pre}_{C_1}(h)\}
 \end{aligned}$$

and

$$\begin{aligned}
 \text{pre}_{C_2}(\bar{\Psi}_h(h)) &= \text{pre}_{C_2}(\min\{[\neg\varphi] \cdot f + [\varphi] \cdot (p \cdot \text{pre}_{C_1}(h) + (1 - p) \cdot \text{pre}_{C_2}(h)), h\}) \\
 &= \min\{[\neg\varphi(C_2)] \cdot \text{pre}_{C_2}(f) + [\varphi(C_2)] \cdot \\
 & \quad (p \cdot \text{pre}_{C_2; C_1}(h) + (1 - p) \cdot \text{pre}_{C_2; C_2}(h)), \text{pre}_{C_2}(h)\}
 \end{aligned}$$

Thus we have

$$\begin{aligned}
& \bar{\Phi}_f(\bar{\Psi}_h(h)) \\
= & [\neg\varphi] \cdot f + [\varphi] \cdot p \cdot \min\{[\neg\varphi(C_1)] \cdot pre_{C_1}(f) \\
& + [\varphi(C_1)] \cdot (p \cdot pre_{C_1;C_1}(h) + (1-p) \cdot pre_{C_1;C_2}(h)), pre_{C_1}(h)\} + \\
& [\varphi] \cdot (1-p) \cdot \min\{[\neg\varphi(C_2)] \cdot pre_{C_2}(f) \\
& + [\varphi(C_2)] \cdot (p \cdot pre_{C_2;C_1}(h) + (1-p) \cdot pre_{C_2;C_2}(h)), pre_{C_2}(h)\} \\
= & \min\{[\neg\varphi] \cdot f + [\varphi \wedge \neg\varphi(C_1)] \cdot p \cdot pre_{C_1}(f) + [\varphi \wedge \neg\varphi(C_2)] \cdot (1-p) \cdot pre_{C_2}(f) \\
& + [\varphi \wedge \varphi(C_1)] \cdot (p^2 \cdot pre_{C_1;C_1}(h) + p(1-p) \cdot pre_{C_1;C_2}(h)) \\
& + [\varphi \wedge \varphi(C_2)] \cdot ((1-p)p \cdot pre_{C_2;C_1}(h) + (1-p)^2 \cdot pre_{C_2;C_2}(h)), \\
& [\neg\varphi] \cdot f + [\varphi \wedge \neg\varphi(C_1)] \cdot p \cdot pre_{C_1}(f) + \\
& [\varphi \wedge \varphi(C_1)] \cdot (p^2 \cdot pre_{C_1;C_1}(h) + p(1-p) \cdot pre_{C_1;C_2}(h)) + \\
& [\varphi] \cdot (1-p) \cdot pre_{C_2}(h), \\
& [\neg\varphi] \cdot f + [\varphi \wedge \neg\varphi(C_2)] \cdot (1-p) \cdot pre_{C_2}(f) + \\
& [\varphi \wedge \varphi(C_2)] \cdot ((1-p)p \cdot pre_{C_2;C_1}(h) + (1-p)^2 \cdot pre_{C_2;C_2}(h)) + \\
& [\varphi] \cdot p \cdot pre_{C_1}(h), \\
& [\neg\varphi] \cdot f + [\varphi] \cdot p \cdot pre_{C_1}(h) + [\varphi] \cdot (1-p) \cdot pre_{C_2}(h)
\end{aligned}$$

The first expression corresponds to the case that we unfold for twice at each state we reach (after the execution of C_1 and C_2), and the second (resp. third) expression corresponds to the case that we unfold for twice at the state that we reach after the execution of C_1 (resp. C_2) and unfold for once at the state that we reach after the execution of C_2 (resp. C_1). The fourth expression corresponds to the case that we unfold for once at both states, i.e., 1-induction principle.

- case $C \equiv \text{if } (\phi) \{C_1\} \text{ else } \{C_2\}$.

$$\begin{aligned}
& \bar{\Phi}_f(\bar{\Psi}_h(h)) \\
= & [\neg\varphi] \cdot f + [\varphi] \cdot pre_C(\bar{\Psi}_h(h)) \\
= & [\neg\varphi] \cdot f + [\varphi \wedge \phi] \cdot pre_{C_1}(\bar{\Psi}_h(h)) + [\varphi \wedge \neg\phi] \cdot pre_{C_2}(\bar{\Psi}_h(h))
\end{aligned}$$

wherein

$$\begin{aligned}
pre_{C_1}(\bar{\Psi}_h(h)) &= pre_{C_1}(\min\{[\neg\varphi] \cdot f + [\varphi] \cdot ([\phi] \cdot pre_{C_1}(h) + [\neg\phi] \cdot pre_{C_2}(h)), h\}) \\
&= \min\{[\neg\varphi(C_1)] \cdot pre_{C_1}(f) + [\varphi(C_1)] \cdot \\
&\quad ([\phi(C_1)] \cdot pre_{C_1;C_1}(h) + [\neg\phi(C_1)] \cdot pre_{C_1;C_2}(h)), pre_{C_1}(h)\}
\end{aligned}$$

and

$$\begin{aligned}
pre_{C_2}(\bar{\Psi}_h(h)) &= pre_{C_2}(\min\{[\neg\varphi] \cdot f + [\varphi] \cdot ([\phi] \cdot pre_{C_1}(h) + [\neg\phi] \cdot pre_{C_2}(h)), h\}) \\
&= \min\{[\neg\varphi(C_2)] \cdot pre_{C_2}(f) + [\varphi(C_2)] \cdot \\
&\quad ([\phi(C_2)] \cdot pre_{C_2;C_1}(h) + [\neg\phi(C_2)] \cdot pre_{C_2;C_2}(h)), pre_{C_2}(h)\}
\end{aligned}$$

Thus we have

$$\begin{aligned}
 & \bar{\Phi}_f(\bar{\Psi}_h(h)) \\
 = & [\neg\varphi] \cdot f + [\varphi \wedge \phi] \cdot \min\{[\neg\varphi(C_1)] \cdot pre_{C_1}(f) \\
 & + [\varphi(C_1)] \cdot ([\phi(C_1)] \cdot pre_{C_1;C_1}(h) + [\neg\phi(C_1)] \cdot pre_{C_1;C_2}(h)), pre_{C_1}(h)\} + \\
 & [\varphi \wedge \neg\phi] \cdot \min\{[\neg\varphi(C_2)] \cdot pre_{C_2}(f) \\
 & + [\varphi(C_2)] \cdot ([\phi(C_2)] \cdot pre_{C_2;C_1}(h) + [\neg\phi(C_2)] \cdot pre_{C_2;C_2}(h)), pre_{C_2}(h)\} \\
 = & \min\{[\neg\varphi] \cdot f + [\varphi \wedge \phi \wedge \neg\varphi(C_1)] \cdot pre_{C_1}(f) + \\
 & [\varphi \wedge \neg\phi \wedge \neg\varphi(C_2)] \cdot pre_{C_2}(f) + \\
 & [\varphi \wedge \phi \wedge \varphi(C_1) \wedge \phi(C_1)] \cdot pre_{C_1;C_1}(h) + [\varphi \wedge \phi \wedge \varphi(C_1) \wedge \neg\phi(C_1)] \cdot pre_{C_1;C_2}(h) + \\
 & [\varphi \wedge \neg\phi \wedge \varphi(C_2) \wedge \phi(C_2)] \cdot pre_{C_2;C_1}(h) + [\varphi \wedge \neg\phi \wedge \varphi(C_2) \wedge \neg\phi(C_2)] \cdot pre_{C_2;C_2}(h), \\
 & [\neg\varphi] \cdot f + [\varphi \wedge \phi \wedge \neg\varphi(C_1)] \cdot pre_{C_1}(f) + \\
 & [\varphi \wedge \phi \wedge \varphi(C_1) \wedge \phi(C_1)] \cdot pre_{C_1;C_1}(h) + [\varphi \wedge \phi \wedge \varphi(C_1) \wedge \neg\phi(C_1)] \cdot pre_{C_1;C_2}(h) + \\
 & [\varphi \wedge \neg\phi] \cdot pre_{C_2}(h), \\
 & [\neg\varphi] \cdot f + [\varphi \wedge \neg\phi \wedge \neg\varphi(C_2)] \cdot pre_{C_2}(f) + \\
 & [\varphi \wedge \neg\phi \wedge \varphi(C_2) \wedge \phi(C_2)] \cdot pre_{C_2;C_1}(h) + [\varphi \wedge \neg\phi \wedge \varphi(C_2) \wedge \neg\phi(C_2)] \cdot pre_{C_2;C_2}(h) + \\
 & [\varphi \wedge \phi] \cdot pre_{C_1}(h), \\
 & [\neg\varphi] \cdot f + [\varphi \wedge \phi] \cdot pre_{C_1}(h) + [\varphi \wedge \neg\phi] \cdot pre_{C_2}(h)
 \end{aligned}$$

The one-to-one relation is the same as that in the former case (probabilistic choice case).

Then we proof the case of $k > 2$ by mathematical induction. Suppose that the proposition holds when $k = n$, i.e., the upper n -induction condition $\bar{\Phi}_f(\bar{\Psi}_h^{n-1}(h)) \leq h$ is equivalent with $\min\{h_1, h_2, \dots, h_m\} \leq h$, where each h_i uniquely corresponds to one $C_d \in \{C_1, \dots, C_m\}$ and is equal to $pre_{C_d}(h)$, where $\{C_1, \dots, C_m\}$ are all the loop-free programs generated by following the decision process in **Step 2**. Section 5.2 within m unfolding.

Then we proof the case of $n + 1$.

$$\begin{aligned}
 \bar{\Phi}_f(\bar{\Psi}_h^n(h)) &= \bar{\Phi}_f(\bar{\Psi}_h(\bar{\Psi}_h^{n-1}(h))) \\
 &= \bar{\Phi}_f(\min\{\bar{\Phi}_f(\bar{\Psi}_h^{n-1}(h)), h\}) \\
 &= \bar{\Phi}_f(\min\{\min\{h_1, h_2, \dots, h_m\}, h\}) \\
 &= \bar{\Phi}_f(\min\{h_1, h_2, \dots, h_m, h\}) \\
 &= [\neg\varphi] \cdot f + [\varphi] \cdot pre_C(\min\{h_1, h_2, \dots, h_m, h\})
 \end{aligned}$$

Through the same inference on the structure C as above, we show it is equivalent to $\min\{g_1, g_2, \dots, g_M\}$, where $M \geq m + 1$ and each g_i uniquely corresponds to one $C_d \in \{C_1, \dots, C_M\}$ and is equal to $pre_{C_d}(h)$, where $\{C_1, \dots, C_M\}$ are all the loop-free programs generated by following the decision process in **Step 2**. Section 5.2 within $n + 1$ unfolding. Thus the proposition holds when $k = n + 1$. Notice that the operators $\bar{\Phi}_f$ and pointwise min are noncommutative.

By mathematical induction, the proposition holds for $k \geq 2$. $\square$

C.2 Supplementary Material for the Pedagogical Explanation in Step 2

We supplement the case of the upper k -induction condition in Example 5.4. We show all potential cases derived from upper 3-induction condition with a simplified dendrogram in the figure 3. Notice

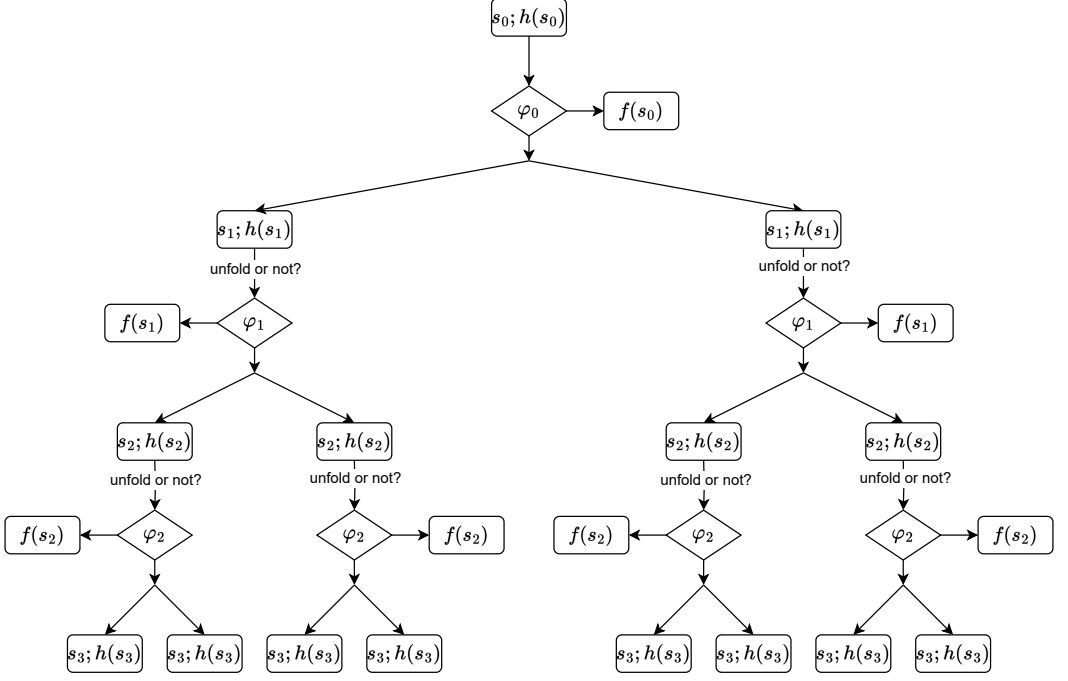


Fig. 3. 3-induction condition example 5.4

that in figure 3, the state $\bar{s}_i$ represents for all the potential states at the i -th iteration of the while loop.

Remarks: If there is only a probabilistic choice structure (or conditional structure) in the loop body, that is, $C = \{C_1\}[p]\{C_2\}$, the total count of the items in the minimize operator satisfies the iteration relation the $C_{k+1} = (C_k + 1)^2 (k \in \mathbb{Z}_+)$. For general loop body, the total count of the items satisfies $C_{k+1} = (C_k + 1)^b (k \in \mathbb{Z}_+)$, where b represents for the count of *branches* when executing the loop body C once from one state.

C.3 Supplementary Material for Step 4 in Section 5.2

Motzkin's Transposition Theorem is a classical theorem that provides a dual characterization for the satisfiability of a system of strict and non-strict inequalities. Below we present the original Motzkin's Transposition Theorem.

THEOREM C.1 (MOTZKIN'S TRANSPOSITION THEOREM [50]). *Given the set of linear, and strict linear, inequalities over real-valued variables $x_1, x_2, \dots, x_n$,*

$$S = \begin{bmatrix} \sum_{i=1}^n \alpha_{(1,i)} \cdot x_i + \beta_1 \leq 0 \\ \vdots \\ \sum_{i=1}^n \alpha_{(m,i)} \cdot x_i + \beta_m \leq 0 \end{bmatrix} \text{ and } T = \begin{bmatrix} \sum_{i=1}^n \alpha_{(m+1,i)} \cdot x_i + \beta_{m+1} < 0 \\ \vdots \\ \sum_{i=1}^n \alpha_{(m+k,i)} \cdot x_i + \beta_{m+k} < 0 \end{bmatrix}$$

in which $\alpha_{(1,1)}, \dots, \alpha_{(m+k,n)}$ and $\beta_1, \dots, \beta_{m+k}$ are real-valued, we have that S and T simultaneously are not satisfiable (i.e., they have no solution in x) if and only if there exist non-negative real numbers $\lambda_0, \lambda_1, \dots, \lambda_{m+k}$ such that either the condition (A_1) :

$$0 = \sum_{i=1}^{m+k} \lambda_i \alpha_{(i,1)}, \dots, 0 = \sum_{i=1}^{m+k} \lambda_i \alpha_{(i,n)}, 1 = (\sum_{i=1}^{m+k} \lambda_i \beta_i) - \lambda_0,$$

or condition (A_2) : at least one coefficient λ_i for i in the range $\{m+1, \dots, m+k\}$ is non-zero and

$$0 = \sum_{i=1}^{m+k} \lambda_i \alpha_{(i,1)}, \dots, 0 = \sum_{i=1}^{m+k} \lambda_i \alpha_{(i,n)}, 0 = (\sum_{i=1}^{m+k} \lambda_i \beta_i) - \lambda_0.$$

In our work, we consider the variant form of Motzkin's Transposition Theorem (see Theorem 5.7) and below we proof it.

Theorem 5.7. [Corollary of Motzkin's Transposition Theorem] Let S and T be the same systems of linear inequalities as that in Theorem C.1. If S is satisfiable, then $S \wedge T$ is unsatisfiable iff there exist non-negative reals $\lambda_0, \lambda_1, \dots, \lambda_{m+k}$ and at least one coefficient λ_i for $i \in \{m+1, \dots, m+k\}$ is non-zero, such that:

$$0 = \sum_{i=1}^{m+k} \lambda_i \alpha_{(i,1)}, \dots, 0 = \sum_{i=1}^{m+k} \lambda_i \alpha_{(i,n)}, 0 = (\sum_{i=1}^{m+k} \lambda_i \beta_i) - \lambda_0.$$

i.e., the condition (A_2) in Theorem C.1.

Before we proof the theorem, we introduce the desired theorem: Farkas's Lemma:

LEMMA C.2 (FARKAS'S LEMMA [28]). Consider the following system of linear inequalities over real-valued variables $x_1, x_2, \dots, x_n$,

$$S = \begin{bmatrix} \alpha_{(1,1)}x_1 & + \dots + & \alpha_{(1,n)}x_n & + \beta_1 & \leq 0 \\ \vdots & & \vdots & & \vdots \\ \alpha_{(m,1)}x_1 & + \dots + & \alpha_{(m,n)}x_n & + \beta_m & \leq 0 \end{bmatrix}$$

When S is satisfiable, it entails a given linear inequality

$$\phi : c_1x_1 + \dots + c_nx_n + d \leq 0$$

if and only if there exist non-negative real numbers $\lambda_0, \lambda_1, \dots, \lambda_m$, such that

$$c_1 = \sum_{i=1}^m \lambda_i \alpha_{(i,1)}, \dots, c_n = \sum_{i=1}^m \lambda_i \alpha_{(i,n)}, d = (\sum_{i=1}^m \lambda_i \beta_i) - \lambda_0$$

Furthermore, S is unsatisfiable if and only if the inequality $1 \leq 0$ can be derived as shown above.

Now we proof the corollary (Theorem 5.7).

PROOF. Proof by contradiction. According to Motzkin's Transposition Theorem, S and T have no solution in x if and only if there exists non-negative real numbers $\lambda_0, \lambda_1, \dots, \lambda_{m+k}$ such that either condition (A_1) or (A_2) is satisfied. We first proof $(\lambda_{m+1} \neq 0) \vee (\lambda_{m+2} \neq 0) \vee \dots \vee (\lambda_{m+k} \neq 0)$.

If it is not satisfied, we assume that $\lambda_{m+1} = \dots = \lambda_{m+k} = 0$. Then we know the condition (A_1) must be satisfied and we have (By applying the assumption $\lambda_{m+1} = \dots = \lambda_{m+k} = 0$):

$$0 = \sum_{i=1}^m \lambda_i \alpha_{(i,1)}, \dots, 0 = \sum_{i=1}^m \lambda_i \alpha_{(i,n)}, \sum_{i=1}^m \lambda_i \beta_i = \lambda_0 + 1 \geq 1,$$

By applying Farkas's Lemma, we have:

$$c_1 = \sum_{i=1}^m \lambda_i \alpha_{(i,1)} = 0, \dots, c_n = \sum_{i=1}^m \lambda_i \alpha_{(i,n)} = 0, d = (\sum_{i=1}^m \lambda_i \beta_i) - \lambda_0 = \lambda_0 + 1 - \lambda_0 = 1,$$

Thus we have:

$$\phi = c_1x_1 + \dots + c_nx_n + d = d = 1 \leq 0$$

if and only if S is not satisfiable, which contradicts the assumption, so the assumption does not hold. We have proved $(\lambda_{m+1} \neq 0) \vee (\lambda_{m+2} \neq 0) \vee \dots \vee (\lambda_{m+k} \neq 0)$.

If condition (A_1) is satisfied, then exists non-negative real numbers $\lambda_0, \lambda_1, \dots, \lambda_{m+k}$ and $(\lambda_{m+1} \neq 0) \vee (\lambda_{m+2} \neq 0) \vee \dots \vee (\lambda_{m+k} \neq 0)$ (what we just prove) such that

$$0 = \sum_{i=1}^{m+k} \lambda_i \alpha_{(i,1)}, \dots, 0 = \sum_{i=1}^{m+k} \lambda_i \alpha_{(i,n)}, 1 = \left(\sum_{i=1}^{m+k} \lambda_i \beta_i \right) - \lambda_0,$$

let $\lambda'_0 = \lambda_0 + 1 \geq 0$ and we can find that it also satisfies the condition (A_2) , that is $A_1 \implies A_2$. Thus, Motzkin's Transposition Theorem can be simplified as: If S is satisfiable, then S and T have no solution in x if and only if there exists non-negative real numbers $\lambda_0, \lambda_1, \dots, \lambda_{m+k}$, such that:

$$((A_1 \vee A_2) \wedge (A_1 \implies A_2)) \iff A_2$$

Thus we prove Theorem 5.7. □

D SUPPLEMENTARY MATERIAL FOR SECTION 6

D.1 Supplementary Experimental Results

In this section, as shown in *Benchmarks* in Section 6.1, we supplement the benchmarks that are either repetitive patterns or cannot be handled by k -induction with small $k = 2, 3$ or can be directly handled by 1-induction. These benchmarks are not concluded in Tables 1 and 3 in the main text, and we show the experimental results of these benchmarks in Tables 5 and 6. These supplementary benchmarks mainly come from Bao et al. [5], Batz et al. [11]. Below we show the details.

Upper bounds. The results of upper bounds are shown in Table 5. For BRP, GRID-SMALL, which are taken from Batz et al. [11], we add a suitable linear return function f and find that they can be directly handled by 1-induction. For BOUNDED-RANDOM-WALK-MULTI-STEP, ZERO-CONF, they cannot be handled by k -induction with small $k = 2, 3$. For the benchmarks in Bao et al. [5], we select the representative benchmarks and do not list the experimental results of the benchmarks that have repetitive patterns, such as, the GEO series (similar with GEO, BIN0, SUM0 (similar with BIN series). For GAMBLER, BIN1, DUELLING, we find that they can be directly handled by 1-induction. For GEOAR, BIN2, we find that they cannot be handled by k -induction with small $k = 2, 3$.

Lower Bounds. We consider the same benchmarks as that in the case of upper bounds. On all but two benchmarks (GEOAR, BIN2), we find that they can be handled by 1-induction. For the benchmarks GEOAR, BIN2, we find that they cannot be handled by k -induction with small $k = 2, 3$.

D.2 Application of Putinar's Positivstellensatz [53]

We recall Putinar's Positivstellensatz below.

THEOREM D.1 (PUTINAR'S POSITIVSTELLENSATZ [53]). *Let V be a finite set of real-valued variables and $g, g_1, \dots, g_m \in \mathbb{R}[V]$ be polynomials over V with real coefficients. Consider the set $\mathcal{S} := \{\mathbf{x} \in \mathbb{R}^V \mid g_i(\mathbf{x}) \geq 0 \text{ for all } 1 \leq i \leq m\}$ which is the set of all real vectors at which every g_i is non-negative. If (i) there exists some g_k such that the set $\{\mathbf{x} \in \mathbb{R}^V \mid g_k(\mathbf{x}) \geq 0\}$ is compact and (ii) $g(\mathbf{x}) > 0$ for all $\mathbf{x} \in \mathcal{S}$, then we have that*

$$g = f_0 + \sum_{i=1}^m f_i \cdot g_i \tag{12}$$

for some polynomials $f_0, f_1, \dots, f_m \in \mathbb{R}[V]$ such that each polynomial f_i is the a sum of squares (of polynomials in $\mathbb{R}[V]$), i.e. $f_i = \sum_{j=0}^k q_{i,j}^2$ for polynomials $q_{i,j}$'s in $\mathbb{R}[V]$.

Table 5. Other Experiment Results: Upper Case

Benchmark	f	Inv	1-induction		2-induction		Piecewise Upper Bound
			Time (s)	Solution	Time (s)	Solution	
BRP-VARIANT	$totalFail$	$[0 \leq sent \leq (toSend + 1) \wedge 0 \leq totalFail]$	0.12	$-\frac{1}{5}sent + \frac{1}{5}toSend + totalFail + \frac{1}{9}$	0.73	$-\frac{1}{5}sent + \frac{1}{5}toSend + totalFail + \frac{1}{9}$	$[sent \geq toSend] \cdot totalFail + [sent < toSend] \cdot (-\frac{1}{5}sent + \frac{1}{5}toSend + totalFail + \frac{1}{9})$
BOUNDED-RW-MULTI-STEP	x	$[0 \leq x \wedge 1 \leq s \leq 5]$	0.46	-	149.80	-	-
GRID-SMALL	$a + b$	$[0 \leq a \leq 11 \wedge 0 \leq b \leq 11]$	0.12	22	1.43	22	$[a < 0 \vee a > 10 \vee b < 0 \vee b > 10] \cdot (a + b) [0 \leq a \leq 10 \wedge 0 \leq b \leq 10] \cdot 22$
ZERO-CONF	cur	$[0 \leq start \leq 1 \wedge 0 \leq est \leq 1]$	0.3	-	6.11	-	-
GAMBLER	i	$[0 \leq x \leq y]$	0.17	$5y - 5x + i$	1.79	$5y - 5x + i$	$[0 < x < y] \cdot (5y - 5x + i) + [x \leq 0 \vee x \geq y] \cdot i$
GEOAR	x	$[0 \leq z]$	0.17	-	0.59	-	-
BIN1	y	$[x \leq 10]$	0.13	$y - 0.5x + 5.0$	0.75	$y - 0.5x + 5.0$	$[x \geq 10] \cdot y + [x < 10] \cdot (y - 0.5x + 5.0)$
BIN2	x	$[0 \leq i]$	0.15	-	0.82	-	-
DUELLING	t	$[0 \leq t \leq 1 \wedge 0 \leq c \leq 1]$	0.22	1	5.7	$0.1t + 0.9$	$[c < 1] \cdot t + [c \geq 1 \wedge t == 1] \cdot (0.45t + 0.5) [c \geq 1 \wedge t == 0] \cdot (0.725t + 0.25)$

Table 6. Other Experiment Results: Lower Case

Benchmark	f	Inv	1-induction		2-induction		Piecewise Lower Bound
			Time (s)	Solution	Time (s)	Solution	
BRP-VARIANT	$totalFail$	$[0 \leq sent \leq (toSend + 1) \wedge 0 \leq totalFail]$	0.12	$-\frac{1}{5}sent + \frac{1}{5}toSend + totalFail$	0.69	$-\frac{1}{5}sent + \frac{1}{5}toSend + totalFail$	$[sent \geq toSend] \cdot totalFail + [sent < toSend] \cdot (-\frac{1}{5}sent + \frac{1}{5}toSend + totalFail)$
BOUNDED-RW-MULTI-STEP	x	$[0 \leq x \wedge 1 \leq s \leq 5]$	0.46	x	149.89	x	$[x \leq 0 \vee s < 1 \vee s > 5] \cdot x [0 < x <= 1 \wedge 1 \leq s \leq 5] \cdot (x + 1) [1 < x \wedge 1 \leq s \leq 5] \cdot (x + s)$
GRID-SMALL	$a + b$	$[0 \leq a \leq 11 \wedge 0 \leq b \leq 11]$	0.12	10	1.60	10	$[a < 0 \vee a > 10 \vee b < 0 \vee b > 10] \cdot (a + b) [0 \leq a \leq 10 \wedge 0 \leq b \leq 10] \cdot 10$
ZERO-CONF	cur	$[0 \leq start \leq 1 \wedge 0 \leq established \leq 1 \wedge 0 \leq cur \leq 100000001]$	0.29	cur	6.08	cur	$[cur < 0 \vee cur >= 100000000 \vee est > 0] \cdot cur [0 \leq cur \leq 100000000 \wedge est \leq 0 \wedge start == 1] \cdot cur [0 \leq cur \leq 100000000 \wedge est \leq 0 \wedge start == 0] \cdot (0.999999999cur + 0.999999999)$
GAMBLER	i	$[0 \leq x \leq y]$	0.16	i	1.61	i	$[0 < x < y] \cdot (i + 1) + [x \leq 0 \vee x \geq y] \cdot i$
GEOAR	x	$[0 \leq z]$	0.13	-	0.49	-	-
BIN1	y	$[x \leq 10]$	0.12	$y - 0.5x + 5.0$	0.78	$y - 0.5x + 5.0$	$[x \geq 10] \cdot y + [x < 10] \cdot (y - 0.5x + 5.0)$
BIN2	x	$[0 \leq i]$	0.17	-	0.87	-	-
DUELLING	t	$[0 \leq t \leq 1 \wedge 0 \leq c \leq 1]$	0.23	$0.084t - 0.083c$	5.77	$t - \frac{3}{2}c$	$[c < 1] \cdot t + [c \geq 1 \wedge t == 1] \cdot \frac{2}{3} [c \geq 1 \wedge t == 0] \cdot (0.5t + \frac{1}{3})$

In our comparison, we utilize the sound form in (12) for witnessing a polynomial g to be non-negative over a semi-algebraic set P for each inductive constraint $\forall x \in P, g(x) \geq 0$. In our comparison, the constraints come from the inductive principle.

E BENCHMARK PROGRAMS

This section contains the benchmark programs evaluated in our work.

E.1 Programs in Tables 1 and 3

This section contains the benchmark programs in the main text.

Example E.1 (Geo).

$$C_{\text{Geo}}: \quad \text{while } (0 \leq c) \{ \\ \quad \{c := 1\} [0.5] \{x := x + 1\} \\ \quad \}$$

Example E.2 (k-geo).

$$C_{\text{k-geo}}: \quad \text{while } (k \leq N) \{ \\ \quad \{k := k + 1; y := y + x; x := 0\} [0.5] \{x := x + 1\} \\ \quad \}$$

Example E.3 (Binomial-random).

$$C_{\text{Bin-ran}}: \quad \text{while } (i \leq 10) \{ \\ \quad \{x := x + 1\} [0.5] \{x := 0\} \\ \quad \{y := y + x; i := i + 1\} [0.9] \{y := y + 1; i := 0\} \\ \quad \}$$

Example E.4 (Coin).

$$C_{\text{Coin}}: \quad \text{while } (x = y) \{ \\ \quad \{x := 0\} [3/4] \{x := 1\} \\ \quad \{y := 0\} [3/4] \{y := 1\} \\ \quad i := i + 1; \\ \quad \}$$

Example E.5 (Martingale).

$$C_{\text{Mart}}: \quad \text{while } (0 < x) \{ \\ \quad \{y := y + x; x := 0\} [0.5] \{y := y - x; x := 2 * x\} \\ \quad i := i + 1; \\ \quad \}$$

Example E.6 (Growing Walk).

$$C_{\text{Growing Walk}}: \quad \text{while } (0 \leq x) \{ \\ \quad \{x := x + 1; y := y + x\} [0.5] \{x := -1\} \\ \quad \}$$

Example E.7 (Growing Walk variant1).

$$C_{\text{Growing Walk1}}: \quad \text{while } (0 \leq x) \{ \\ \quad \{x := x - 1; y := y + x\} [0.5] \{x := 1\} \\ \quad \}$$

Example E.8 (Expected Time).

$$C_{\text{Expected Time}}: \quad \text{while } (0 \leq x) \{ \\ \quad \{x := x - 1; t := t + 1\} [0.9] \{x := 10; t := t + 1\} \\ \quad \}$$

Example E.9 (Bernoulli's St. Petersburg Paradox variant).

$$C_{\text{St. Petersburg1}}: \quad \text{while} (x \leq 0) \{ \\ \quad \{x := 1\} [0.75] \{y := 2 * y\} \\ \quad \}$$

Example E.10 (Zero Conference variant).

$$C_{\text{Zero-Conf-Var}}: \quad \text{while} (established \leq 0 \wedge start \leq 1) \{ \\ \quad \text{if} (start \geq 1) \{ \\ \quad \quad \{start := 0\} [0.3] \{start := 0; established := 1\} \} \\ \quad \text{else} \{ \{curprobe := curprobe + 1\} [0.99] \{start := 1; curprobe := curprobe - 1\} \} \\ \quad \}$$

Example E.11 (Equal Probability Grid Family).

$$C_{\text{Equal-Prob-Grid-Family}}: \quad \text{while} (a \leq 10 \wedge b \leq 10 \wedge goal = 0) \{ \\ \quad \text{if} (b \geq 10) \{ \\ \quad \quad \{goal := 1\} [0.5] \{goal := 2\} \} \\ \quad \text{else} \{ \\ \quad \quad \text{if} (a \geq 10) \{ \\ \quad \quad \quad a := a - 1 \} \\ \quad \quad \text{else} \{ \\ \quad \quad \quad \{a := a + 1\} [0.5] \{b := b + 1\} \} \\ \quad \quad \} \\ \quad \}$$

Example E.12 (RevBin).

$$C_{\text{RevBin}}: \quad \text{while} (1 \leq x) \{ \\ \quad \{x := x - 1; z := z + 1\} [0.5] \{z := z + 1\} \\ \quad \}$$

Example E.13 (Fair Coin).

$$C_{\text{Fair Coin}}: \quad \text{while} (x \leq 0 \wedge y \leq 0) \{ \\ \quad \{x := 0\} [0.5] \{x := 1; i := i + 1\} \\ \quad \{y := 0\} [0.5] \{y := 1; i := i + 1\} \\ \quad \}$$

E.2 Programs in Tables 5 and 6

This section contains the benchmarks in the appendix.

Example E.14 (brp-variant).

$$C_{\text{brp-variant}}: \quad \text{while} (send \leq toSend) \{ \\ \quad \{sent := sent + 1\} [0.9] \{totalFail := totalFail + 1\} \\ \quad \}$$

Example E.15 (grid-small).

$$C_{\text{grid-small}}: \quad \text{while } (a < 10 \wedge b < 10) \{ \\ \quad \{a := a + 1\} [0.5] \{b := b + 1\} \\ \quad \}$$

Example E.16 (Zero Conference).

$$C_{\text{Zero-Conf}}: \quad \text{while } (curprobe < 100000000 \wedge established \leq 0 \wedge start \leq 1) \{ \\ \quad \text{if } (start \geq 1) \{ \\ \quad \quad \{start := 0\} [0.5] \{start := 0; established := 1\} \} \\ \quad \text{else } \{ \{curprobe := curprobe + 1\} [0.999999999] \{start := 1; curprobe := 0\} \} \\ \quad \}$$

Example E.17 (bounded RW multi step).

$$C_{\text{bounded-rw-multi-step}}: \quad \text{while } (0 < x \wedge 1 \leq s \wedge s \leq 5) \{ \\ \quad \{x := x - 1\} \\ \quad [0.5] \\ \quad \{\text{if } (x \leq 1) \{ \\ \quad \quad s := 1 : 1/5 + 2 : 1/5 + 3 : 1/5 + 4 : 1/5 + 5 : 1/5; \\ \quad \quad \}\text{else}\{skip\} \\ \quad \quad x := x + s; \\ \quad \}$$

Example E.18 (Duel Boy).

$$C_{\text{Duel}}: \quad \text{while } (c \geq 1) \{ \\ \quad \text{if } (t > 0) \{ \\ \quad \quad \{c := 0\} [0.5] \{t := 1 - t\} \\ \quad \quad \}\text{else}\{\{c := 0\} [0.75] \{t := 1 - t\}\} \\ \quad \}$$

Example E.19 (Gambler).

$$C_{\text{Gambler}}: \quad \text{while } (0 < x \wedge x < y) \{ \\ \quad \{x := x + 1; i := i + 1\} [0.6] \{x := x - 1; i := i + 1\} \\ \quad \}$$

Example E.20 (GeoAr).

$$C_{\text{GeoAr}}: \quad \text{while } (0 < z) \{ \\ \quad y := y + 1; \\ \quad \{x := x + y\} [0.9] \{z := 0\} \\ \quad \}$$

Example E.21 (Bin1).

$$C_{\text{Bin1}}: \quad \text{while } (x < 10) \{ \\ \qquad \{y := y + 1; x := x + 1\} [0.5] \{x := x + 1\} \\ \quad \}$$

Example E.22 (Bin2).

$$C_{\text{Gambler}}: \quad \text{while } (i > 0) \{ \\ \qquad \{x := x + 1; i := i - 1\} [0.25] \{x := x + y; i := i - 1\} \\ \quad \}$$