Formal Methods in System Design (2019) 54:145–163
https://doi.org/10.1007/s10703-019-00339-8

Statistical verification of PCTL using antithetic and stratified samples

Yu Wang[1] · Nima Roohi[2] · Matthew West[3] · Mahesh Viswanathan[4] · Geir E. Dullerud[5]

Published online: 28 August 2019

Abstract
In this work, we study the problem of statistically verifying Probabilistic Computation Tree Logic (PCTL) formulas on discrete-time Markov chains (DTMCs) with stratified and antithetic samples. We show that by properly choosing the representation of the DTMCs, semantically negatively correlated samples can be generated for a fraction of PCTL formulas via the stratified or antithetic sampling techniques. Using stratified or antithetic samples, we propose statistical verification algorithms with asymptotic correctness guarantees based on sequential probability ratio tests, and show that these algorithms are more sample-efficient than the algorithms using independent Monte Carlo sampling. Finally, the efficiency of the statistical verification algorithm with stratified and antithetic samples is demonstrated by numerical experiments on several benchmarks.

Keywords Markov chains · Temporal logic · Variance reduction · Sequential probability ratio test

✉ Yu Wang
yu.wang094@duke.edu

Nima Roohi
nroohi@ucsd.edu

Matthew West
mwest@illinois.edu

Mahesh Viswanathan
vmahesh@illinois.edu

Geir E. Dullerud
dullerud@illinois.edu

1 Electrical and Computer Engineering, Duke University, Durham, USA

2 Computer Science and Engineering, University of California San Diego, San Diego, USA

3 Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, USA

4 Computer Science, University of Illinois at Urbana-Champaign, Urbana, USA

5 Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, USA

1 Introduction

Statistical verification of temporal specifications has received increasing attention during the past decade [2,6,9,15–17,25,26]. Compared to the symbolic approach, statistical model checkers are usually more scalable to large-scale real-world problems with complicated stochastic behavior [14,21–23,27]. The general idea of statistical verification is to treat the problem of checking a PCTL formula on a probabilistic system as an hypothesis testing problem. By drawing sample behaviors from the underlying probabilistic system, the satisfaction of the specification can be inferred with high confidence.

Currently, most statistical model checking algorithms previously proposed crucially rely on *independent* Monte Carlo sampling. Specifically, the underlying probabilistic system is "simulated" to generate a sample path and a new sample is drawn in the same manner in each round. Consequently, the samples are independent and identically distributed (i.i.d.).

The main thesis of this paper is that the sampling cost for verification can be significantly reduced if the statistical model checker draws *semantically negatively correlated* samples, as opposed to independent samples. Specifically, let us consider the core task of a statistical model checker, namely, to determine if the measure of executions satisfying a property ϕ is greater than some threshold p. For simplicity, assume that the truth of ϕ itself can be determined by a finite prefix of the execution. In such a situation, the model checker draws sample executions, determines how many of the executions satisfy ϕ, and uses this to estimate the measure of paths satisfying ϕ. Thus, each sample can be viewed as a $0/1$-valued random variable X_i (which takes value 1 if the execution satisfies ϕ, and 0 otherwise), whose expectation is estimated by

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

One factor that plays an important role in determining how many samples are needed for the algorithm to be confident in its answer is the variance. Informally, the lower the variance of the estimate, the more likely the estimate is to be close to the actual mean, and therefore, the algorithm requires fewer samples. In general, the variance of the estimate is given by

$$\mathrm{Var}\big[\bar{X}\big] = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}\big[X_i\big] + \frac{2}{n^2}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathrm{Cov}\big[X_i, X_j\big].$$

If the samples are i.i.d., then the covariance is 0, and the variance is given by

$$\mathrm{Var}\big[\bar{X}\big] = \frac{1}{n}\mathrm{Var}\big[X\big].$$

However, as can be seen from the above expression, the variance can be reduced if the samples are *semantically negatively correlated*, i.e.,

$$\sum_{i=1}^{n}\sum_{j=i+1}^{n}\mathrm{Cov}\big[X_i, X_j\big] \le 0.$$

Common techniques to generate such negatively correlated samples with negligible additional computational cost include *stratified sampling* and *antithetic sampling*, which have been popular among the statistics community in improving the accuracy of statistical estimation [7,10,11]. The general idea is to generate samples that are repellent to each other — a sample occupying some region forbids other samples entering the same region. In this

paper, we present statistical model checking algorithms for verifying finite horizon PCTL properties on discrete time Markov chains (DTMC) using semantically negatively correlated samples generated by stratified sampling and antithetic sampling.

To ensure a lucid exposition of the main ideas, we only consider non-nested PCTL formulas of the form $\mathcal{P}_{\sim p}\phi$, where ϕ is a formula without probabilistic operators; in other words, ϕ's truth can be determined on a single path. PCTL formulas in general form with nested probabilistic operators can be handled in the standard manner using the approach proposed in [15–17]. The main contribution of this paper are sequential probability ratio tests that work when samples are drawn using stratified sampling or antithetic sampling, which help reduce the total number of samples needed for a statistical model checker to be confident in its answer.

The results regarding stratified sampling in this paper have appeared in [24] without proofs. This paper extends [24] by including the proofs for the results regarding stratified sampling and providing extra results regarding to antithetic sampling. The rest of the paper is organized as follows. The preliminaries are given in Sect. 2. In Sect. 3, we show that the stratified sampling technique can be used to generate semantically negatively correlated samples for non-nested PCTL formulas, and propose a statistical verification algorithm using stratified samples that is asymptotically more sample-efficient that the ones using i.i.d. samples. The same problem is studied in Sect. 4 with antithetic samples. In Sect. 5, we demonstrate the efficiency of the statistical verification algorithm with stratified samples by numerical experiments on several benchmarks. Finally, we conclude this work in Sect. 6.

2 Preliminaries

We denote the set of natural numbers and real numbers by $\mathbb{N}, \mathbb{R}$. We take the convention that $0^0 = 1$. For $n \in \mathbb{N}$, let $[n] = \{1, 2, \ldots, n\}$. A permutation of $[n]$ is a bijection $\pi : [n] \to [n]$. For $p \in \mathbb{R}^n$, the i^{th} entry of p is denoted by p_i. For $M \in \mathbb{R}^{n\times m}$, the entry in the i^{th} row, j^{th} column of M is denoted by M_{ij}.

2.1 Markov chains

Consider a time-homogeneous discrete-time Markov chain (DTMC) $\mathcal{M}$ of n numbered states with initial state $s \in [n]$ and transition probability matrix M, in which M_{ij} defines the transition probability from i to j. For any $i \in [n]$,

$$\sum_{j=1}^{n} M_{ij} = 1. \tag{1}$$

For a sample path $X = \{X(t)\}_{t\in\mathbb{N}} \subseteq [n]$ of the Markov chain, we can write

$$X(t+1) = f(X(t), E(t)), \quad t \in \mathbb{N}, \tag{2}$$

where $E(t) \sim \mathbf{U}_{[0,1)}$. At each time t, the Markov chain is driven by the random seed $E(t)$ uniformly sampled from the interval $\mathbf{U}_{[0,1]}$, thus (2) induces a surjection from $[0, 1]^T$ to the space of sample paths of the Markov chain $\mathcal{M}$ of length T.

Generally, a discrete-time Markov chain $\mathcal{M}$ can be represented in (2) in multiple ways. In this work, we choose the following representation

$$f(i, e) = \begin{cases} 1, & \text{if } 0 \le e < M_{i1}, \\ j, & \text{if } \sum_{k=1}^{j-1} M_{ik} \le e < \sum_{k=1}^{j} M_{ik}. \end{cases} \tag{3}$$

2.2 PCTL

Probabilistic computational tree logic (PCTL) is commonly used to express probabilistic properties of discrete-time Markov chains. In this paper, we only consider PCTL formulas over finite time horizons. The syntax and semantics of the logic is given below.

Definition 1 (*Syntax*) Let Ω be a given set of atomic propositions. A PCTL formula is generated recursively by

$$\phi ::= \Omega \mid \neg\phi \mid \phi \wedge \psi \mid \mathcal{P}_{\sim p}(\mathcal{X}\phi) \mid \mathcal{P}_{\sim p}(\phi \mathcal{U}_{\leq T} \psi), \tag{4}$$

where $\omega \in \Omega$ is an atomic proposition, $\sim \in \{<, >, \leq, \geq\}$, $p \in (0, 1)$ is a probability threshold, and $T \in \mathbb{N}$ is a time bound.

Remark 1 When $p \in \{0, 1\}$, the PCTL formula reduces to a CTL formula. In this work, we only consider $p \in (0, 1)$. Other common temporal operators can be constructed by composing the temporal logic operators given in Definition 1.

Definition 2 (*Semantics*) Let $L : [n] \to 2^{\Omega}$ be a given labeling function where $[n]$ is the set of states of the Markov chain $\mathcal{M}$. A random sample path starting from the state $s \in [n]$ is denoted by $s(0) = s, s(1), s(2), \ldots$. The semantics of PCTL is defined recursively by

$$\begin{aligned}
&s \models \omega \text{ iff } \omega \in L(s),\\
&s \models \neg\phi \text{ iff } s \not\models \phi,\\
&s \models \phi \wedge \psi \text{ iff } s \models \phi \text{ and } s \models \psi,\\
&s \models P_{\sim p}(\mathcal{X}\phi) \text{ iff } \mathbb{P}\big[s(1) \models \phi\big] \sim p,\\
&s \models \mathcal{P}_{\sim p}(\phi \mathcal{U}_{\leq T} \psi) \text{ iff}\\
&\quad \mathbb{P}\big[\exists t \leq T : s(0) \models \phi, \ldots, s(t-1) \models \phi, s(t) \models \psi\big] \sim p.
\end{aligned} \tag{5}$$

2.3 Statistical verification via hypothesis testing

As mentioned in Sect. 1, we focus on non-nested PCTL formulas of the form $\mathcal{P}_{\sim p}\phi$ in this work. Since ϕ is a formula without probabilistic operators, its correctness can be determined on any sample path of the Markov chain $\mathcal{M}$. Thus, the verification of $\mathcal{P}_{\sim p}\phi$ can be converted to a hypothesis testing problem. Specifically, for any path X of the Markov chain $\mathcal{M}$, we define with a slight abuse of notation that

$$\phi(X) = \begin{cases} 1, & \text{if } X \text{ satisfies } \phi,\\ 0, & \text{otherwise.} \end{cases} \tag{6}$$

Then, checking $\mathcal{P}_{<p}\phi$ is semantically equivalent to the composite hypothesis testing (CHT) problem:

$$\begin{aligned} H_0 &: P_\phi < p,\\ H_1 &: P_\phi \geq p, \end{aligned} \tag{7}$$

where $P_\phi = \mathbb{P}\big[\phi(X)\big]$.

Generally, the CHT problem (7) is challenging. Therefore, as with other literature in this area, we assume some a priori knowledge on the distance $|P_\phi - p|$.

Assumption 1 Let $|P_\phi - p| > \delta$ for some known indifference parameter $\delta > 0$. The interval $(P_\phi - \delta, P_\phi + \delta)$ is called the indifference region.

With Assumption 1, the CHT problem (7) can be solved by statistically testing between $P_\phi \leq p - \delta$ and $P_\phi \geq p + \delta$. By considering the worst case, it suffices to solve a simple hypothesis testing (SHT) problem with two hypothesis H_0' and H_1':

$$\begin{aligned} H_0' &: P_\phi = p - \delta, \\ H_1' &: P_\phi = p + \delta. \end{aligned} \tag{8}$$

In addition, by Assumption 1, the PCTL formulas $\mathcal{P}_{<p}\phi$ and $\mathcal{P}_{\leq p}\phi$, or $\mathcal{P}_{>p}\phi$ and $\mathcal{P}_{\geq p}\phi$, are equivalent. To be concrete, we consider $\mathcal{P}_{<p}\phi$ in the rest of this section; formulas in other forms can be dealt with in similar ways.

The SHT problem (8) can be solved efficiently by sequential probability ratio tests (SPRT) [15–17]. Specifically, let $X_1, X_2, \ldots$ be independent sample paths of $\mathcal{M}$. For a confidence level of Type I error

$$\alpha = \mathbb{P}\big[\text{choose } H_1' \mid P_\phi = p - \delta\big] > 0, \tag{9}$$

and Type II error

$$\beta = \mathbb{P}\big[\text{choose } H_0' \mid P_\phi = p + \delta\big] > 0, \tag{10}$$

the SHT problem (8) is checked by SPRT with

$$\Lambda(X) = \prod_{i=1}^{n} \frac{(p+\delta)^{\phi(X_i)}(1-p-\delta)^{1-\phi(X_i)}}{(p-\delta)^{\phi(X_i)}(1-p+\delta)^{1-\phi(X_i)}}, \tag{11}$$

where $X = (X_1, \ldots, X_n)$. H_0 is accepted if $\Lambda(X) > \frac{\beta}{1-\alpha}$; H_0 is accepted if $\Lambda(X) > \frac{1-\beta}{\alpha}$; otherwise, draw a new sample X_{n+1}.

3 Statistical verification using stratified samples

Stratified sampling is an approach to generate negatively correlated random variables. In this section, we show that by choosing a proper representation of the Markov chain $\mathcal{M}$, the stratified sampling technique can be applied to generate semantically negatively correlated sample paths, reducing the sampling cost for statistically verifying a fraction of PCTL formulas. For a lucid explanation of the main ideas, we focus on non-nested PCTL formulas of the form $\mathcal{P}_{\sim p}\phi$, where ϕ is a formula without probabilistic operators; in other words, ϕ's truth can be determined on a single path. PCTL formulas in general form with nested probabilistic operators can be handled in the standard manner using the compositional approach proposed in [15–17].

3.1 Stratified sampling

The stratified sampling algorithm generates m sample paths simultaneously. Let $[0, 1) = [0, \frac{1}{m}] \cup \ldots \cup [\frac{m-1}{m}, 1)$ be a partition of $[0, 1)$. At each time t, a sample is drawn from each sub-interval. To avoid correlation between steps, we generate a permutation π on $[n]$ uniformly at each time t, and then assign the sub-interval $[\frac{\pi(i)-1}{m}, \frac{\pi(i)}{m})$ to the i^{th} path. The random seeds of the m-stratified sample paths are repellent to each other at each time t, as no more than one of them can occupy the same sub-interval. Accordingly, due to the choice of $f(i, e)$ in (3), the m-stratified sample paths are repellent to each other at each time t, in the sense that the probability of any two of them visiting the same state of the Markov chain

is no more than (and usually less than) that of two independent samples. This is summarized by Definition 3 and Algorithm 1. Compared to i.i.d. samples, the additional computational cost for generating stratified samples is negligible.

Definition 3 $\{X_i\}_{i\in[m]}$ is a set of m-stratified samples if they are generated by Algorithm 1.

Algorithm 1 m-stratified sampling

Require: Number of strata m, number of steps T, and initial state s
$t = 0$
for $i \in [m]$ **do**
  $X_i(0) = s$
end for
for $t = 1, \ldots, T-1$ **do**
  Take π to be a permutation of $[m]$
  for $i \in [m]$ **do**
    Take $E_i \sim \mathbf{U}_{[\frac{\pi(i)-1}{m}, \frac{\pi(i)}{m})}$
    $X_i(t+1) = f(X_i(t), E_i(t))$
  end for
end for
return $\{X_i\}_{i\in[m]}$

3.2 Properties of stratified samples

Now we show that for the PCTL formulas satisfying Assumption 2, the m-stratified samples $\{X_i\}_{i\in[m]}$ are semantically negatively correlated, as stated in Theorem 1. By the syntax of PCTL, ϕ is either of the form $\mathcal{X}\psi$ or $\psi_1 \mathcal{U}_{\leq T} \psi_2$, where ψ_1 and ψ_2 are directly checkable on the states of the Markov chain $\mathcal{M}$. We denote the set of states where ψ holds by

$$V_\psi = \{s \in [n] | \psi \in L(s)\}. \tag{12}$$

Assumption 2 For a PCTL formula of the form $\phi = \psi_1 \mathcal{U}_{\leq T} \psi_2$, we assume $V_{\psi_2} \subseteq V_{\psi_1}$.

Theorem 1 *With Assumption 2, let $\{X_i\}_{i\in[m]}$ be m-stratified samples from Markov chain $\mathcal{M}$ and ϕ be a probabilistic-operator-free PCTL formula with satisfaction probability P_ϕ. Then, for any $i, j \in [m]$,*

(i) $\mathbb{E}\left[\sum_{i=1}^m \phi(X_i)/m\right] = P_\phi$;
(ii) $\mathrm{Cov}\left[\phi(X_i), \phi(X_j)\right] \leq 0$ *for* $i \neq j$.

Proof (i) By Algorithm 1, for the sample X_i with $i \in [n]$, the random seeds $E_i(t)$ are drawn independently and uniformly from $[0, 1]$ for all $t \in \mathbb{N}$. Therefore, we have $\mathbb{P}\left[\phi(X_i)\right] = P_\phi$ for all $i \in [n]$, which immediately gives (i).
(ii) By the syntax of PCTL (1), it suffices to prove the result in two specific cases: 1) $\phi = \psi_1 \mathcal{U}_{\leq T} \psi_2$ and 2) $\phi = \mathcal{X}\psi$, where ψ, ψ_1, ψ_2 are atomic propositions.

(1) $\phi = \psi_1 \mathcal{U}_{\leq T} \psi_2$: Recalling Assumption 2, we have $V_{\psi_2} \subseteq V_{\psi_1}$, where V_ψ is the set of states labeled by ψ. Without loss of generality, we can number the states of the Markov chain $\mathcal{M}$ such that $V_{\psi_1} = [n_1]$ and $V_{\psi_2} = [n_2]$ where $n_1 \geq n_2$.

Claim 1: For any $r_1, r_2 \in [n]$ and any $i \neq j \in [n]$, we have

$$\begin{aligned}&\mathbb{P}\big[X_i(t+1) \in V_{\psi_1} \mid X_j(t+1) \in V_{\psi_1}, X_i(t) = r_1, X_j(t) = r_2\big]\\ &\leq \mathbb{P}\big[X_i(t+1) \in V_{\psi_1} \mid X_i(t) = r_1, X_j(t) = r_2\big].\end{aligned}$$

To prove *Claim 1*, it suffices to show that

$$\begin{aligned}\mathcal{P} = &\mathbb{P}\big[X_i(t+1) \in V_{\psi_1}, X_j(t+1) \in V_{\psi_1} | X_i(t) = r_1, X_j(t) = r_2\big]\\ &- \mathbb{P}\big[X_i(t+1) \in V_{\psi_1} | X_i(t) = r_1, X_j(t) = r_2\big]\\ &\times \mathbb{P}\big[X_i(t+1) \in V_{\psi_1} | X_j(t) = r_1, X_j(t) = r_2\big]\\ &\leq 0. \end{aligned} \tag{13}$$

Let $S = \sum_{k=1}^{n_1} M_{kr_1}$ and $R = \sum_{k=1}^{n_1} M_{kr_2}$, where M is the transition probability matrix. Then we have

$$\begin{aligned}&\mathbb{P}\big[X_i(t+1) \in V_{\psi_1} | X_i(t) = r_1, X_j(t) = r_2\big] = \mathbb{P}\big[E_i(t) \leq S\big] = S,\\ &\mathbb{P}\big[X_j(t+1) \in V_{\psi_1} | X_i(t) = r_1, X_j(t) = r_2\big] = \mathbb{P}\big[E_j(t) \leq R\big] = R.\end{aligned}$$

The two random seeds $E_i(t)$ and $E_j(t)$ are distributed uniformly on $[0,1]^2 - \bigcup_{i=1}^{m}[(i-1)/m, i/m]^2$, as shown in Fig. 1. Without loss of generality, assume $S \leq R$. If $\lfloor mS \rfloor < \lfloor mR \rfloor$, where $\lfloor \cdot \rfloor$ is the floor function, we have

$$\begin{aligned}\mathcal{P} &= \frac{m}{m-1}\Big[SR - \frac{1}{m^2}\lfloor mS \rfloor - \frac{1}{m}(S - \frac{\lfloor mS \rfloor}{m})\Big] - SR\\ &= \frac{S(R-1)}{m-1} \leq 0,\end{aligned} \tag{14}$$

where the equality holds if and only if $S = 0$ or $R = 1$. If $\lfloor mS \rfloor = \lfloor mR \rfloor$, we have

$$\mathcal{P} = \frac{m}{m-1}\left[SR - \frac{1}{m^2}\lfloor mS \rfloor - \left(S - \frac{\lfloor mS \rfloor}{m}\right)\left(R - \frac{\lfloor mR \rfloor}{m}\right)\right] - SR. \tag{15}$$

Since

$$\frac{\partial \mathcal{P}}{\partial S} = \frac{\lfloor mR \rfloor}{m-1} - R \text{ and } \frac{\partial \mathcal{P}}{\partial R} = \frac{\lfloor mS \rfloor}{m-1} - S, \tag{16}$$

the maximum of (15) is achieved when the derivative in (16) is zero, namely when

$$\lfloor mR \rfloor = (m-1)R = \lfloor mS \rfloor = (m-1)S. \tag{17}$$

Plugging (17) back into (15) gives $\mathcal{P} \leq 0$, where the equality holds if and only if $S \in \{0, 1\}$ or $R \in \{0, 1\}$.

Claim 2: For any $r_1, r_2 \in [n]$ and any $i \neq j \in [n]$, we have

$$\begin{aligned}&\mathbb{P}\big[X_i(t+1) \in V_{\psi_1} | X_j(t+1) \in V_{\psi_2}, X_i(t) = r_1, X_j(t) = r_2\big]\\ &\leq \mathbb{P}\big[X_i(t+1) \in V_{\psi_2} | X_i(t) = r_1, X_j(t) = r_2\big].\end{aligned}$$

The proof of *Claim 2* is similar to that of of *Claim 1*. The equality holds if and only if if and only if $\mathbb{P}\big[X_i(t+1) \in V_{\psi_1} | X_i(t)\big] \in \{0, 1\}$ or $\mathbb{P}\big[X_j(t+1) \in V_{\psi_2} | X_j(t)\big] \in \{0, 1\}$.

Now, we show that the events $\phi(X_i)$ and $\phi(X_j)$ are negatively correlated. Assume the initial state $X_i(0) = X_j(0) = r_0 \in V_{\psi_1}$. Otherwise, **(ii)** trivially holds. Let A be the event $X_i(t) \in V_{\psi_1}$ for $t \in [T_1 - 1]$ and $X_i(T_1) \in V_{\psi_2}$. Similarly, let B be the

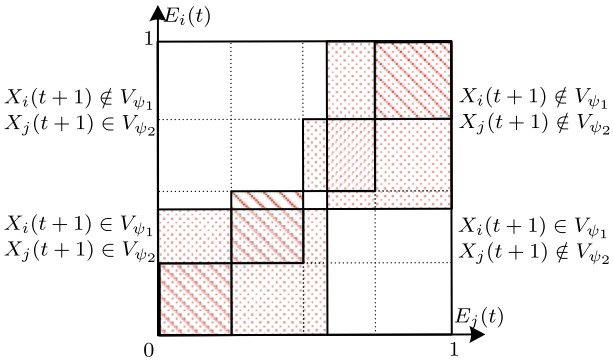


Fig. 1 Joint distribution of two stratified random seeds

event $X_j(t) \in V_{\psi_1}$ for $t \in [T_2 - 1]$ and $X_j(T_2) \in V_{\psi_2}$. By *Claims 1 and 2*, we have for $T_1 > T_2$,

$$\begin{aligned}\mathbb{P}\big[B|A\big] &= \sum \prod_{t\in[T_2]} \mathbb{P}\big[X_j(t) = r_1 \mid X_j(t-1) = r_{t-1}, A\big] \\ &= \sum \prod_{t\in[T_2]} \mathbb{P}\big[X_j(t) = r_1 \mid X_j(t-1) = r_{t-1}, X_i(t) \in V_{\psi_1}, X_i(t-1) \in V_{\psi_1}\big] \\ &\le \sum \prod_{t\in[T_2]} \mathbb{P}\big[X_j(t) = r_1 \mid X_j(t-1) = r_{t-1}\big] = \mathbb{P}\big[B\big], \end{aligned} \tag{18}$$

where $\sum$ stands for $\sum_{r_1\in V_{\psi_1}} \cdots \sum_{r_{T_2-1}\in V_{\psi_1}} \sum_{r_{T_2}\in V_{\psi_2}}$, and the second equality holds because of the Markovian property. The argument for $T_1 \le T_2$ is similar. Therefore, we have $\mathbb{P}\big[\phi(X_j) = 1 \mid \phi(X_i) = 1\big] \le \mathbb{P}\big[\phi(X_j) = 1\big]$. Similarly, we can show $\mathbb{P}\big[\phi(X_j) = 0 \mid \phi(X_i) = 0\big] \le \mathbb{P}\big[\phi(X_j) = 0\big]$. Thus, $\mathrm{Cov}\big[\phi(X_i), \phi(X_j)\big] \le 0$. The equality holds if and only if ϕ is trivially true or false.

(2) $\phi = \mathcal{X}\psi$: Without loss of generality, assume $V_\psi = [l]$ for some $l \in \mathbb{N}$. By the semantics of PCTL, $\phi(X_i) = 1 \iff X_i(1) \in V_\psi \iff E_i(0) \le S = \sum_{j=1}^{l} M_{js}$, where s is the initial state. Thus, by (14) with $R = S$ and $t = 0$, we have $\mathbb{P}\big[\phi(X_i) = 1, \phi(X_j) = 1\big] - S^2 \le 0$ and $\mathbb{P}\big[\phi(X_i) = 0, \phi(X_j) = 0\big] \le (1 - S)^2$, where the equalities hold if and only if $S \in \{0, 1\}$. Thus, $\mathrm{Cov}\big[\phi(X_i), \phi(X_j)\big] \le 0$. The equality holds if and only if ϕ is trivially true or false. □

3.3 Sequential probability ratio test

To implement the SPRT on m-stratified samples $\{X_i\}_{i\in[m]}$, we consider the statistics

$$Y = \sum_{i=1}^{m} \phi(X_i)/m. \tag{19}$$

By Theorem 1, we have $\mathbb{E}\big[Y\big] = P_\phi$. Following the argument in Sect. 2.3, to verify $\mathcal{P}_{<p}\phi$, it suffices to check

$$\begin{aligned} H_0' &: \mathbb{E}\big[Y\big] = p - \delta, \\ H_1' &: \mathbb{E}\big[Y\big] = p + \delta. \end{aligned} \tag{20}$$

In addition, the mean of m-stratified samples within each block are more concentrated than the mean of m independent samples with the same mean,

$$\begin{aligned}\mathrm{Var}\big[Y\big] &= \frac{1}{m^2}\mathrm{Var}\Big[\sum_{j=1}^{m}\phi(X_j)\Big] \\ &= \frac{1}{m}\mathrm{Var}\big[\phi(X_j)\big] + \frac{1}{m}\sum_{k=1,k\neq j}^{m}\mathrm{Cov}\big[\phi(X_j),\phi(X_k)\big] \\ &\leq \frac{1}{m}\mathrm{Var}\big[\phi(X_j)\big]. \end{aligned} \tag{21}$$

Finally, we show that there is no loss of statistical information by considering Y instead of $\{X_i\}_{i\in[m]}$.

Theorem 2 *The joint probability mass function $\pi(x_1,\ldots,x_m)$ of $\phi(X_1),\ldots,\phi(X_m)$ only depends on $\sum_{i=1}^{m}\phi(X_i)$.*

Proof Since the m-stratified samples are symmetric, $\pi(x_1,\ldots,x_m)$ is identical under an arbitrary permutation in the arguments. Therefore, $\pi(x_1,\ldots,x_m)$ only depends on the symmetric polynomials $\sum_{i=1}^{m}x_i$ and $\sum_{i=1}^{m}\sum_{j=1}^{m}x_ix_j$, ..., $x_1,x_2\ldots,x_m$. Each x_i takes value in $\{0,1\}$, so the value of higher order polynomials are determined by $\sum_{i=1}^{m}x_i$, namely the number of 1s in $x_1,\ldots,x_m$. Therefore, $\pi(x_1,\ldots,x_m)$ only depends on $\sum_{i=1}^{m}x_i$. □

Theorem 2 shows that to construct a statistical test for the satisfaction probability P_ϕ of the formula ϕ using m-stratified samples $(X_1,\ldots,X_n)$, it suffices to use the average statistic $Y = \sum_{i=1}^{m}\phi(X_i)/m$. Given the n independent samples $Y^{(n)} = \{Y_1,\ldots,Y_n\} \subseteq \{0,1/m,\ldots,1\}$ of the average statistic Y, we can construct an SPRT algorithm similar to Sect. 2.3,

$$\Lambda'(Y^{(n)}) = \Pi_{i=1}^{n}\frac{\pi_{H_1}(Y^{(n)})}{\pi_{H_0}(Y^{(n)})}, \tag{22}$$

where π_{H_1} and π_{H_0} are the probability mass functions of Y_i under hypotheses H_0 and H_1, respectively.

However, unlike the i.i.d. case in (11), the exact forms of π_{H_1} and π_{H_0} are hard to derive. Therefore, for simplicity, we take an asymptotic approach via the Central Limit Theorem. Let $\nu(Y^{(n)})$ be the empirical distribution given $Y^{(n)}$ and

$$\bar{Y}_n = \frac{1}{n}\sum_{k=1}^{n}Y_k, \quad \sigma_n^2 = \frac{1}{n}\sum_{k=1}^{n}(Y_k-\bar{Y}_n)^2 \tag{23}$$

be the sample mean and sample variance, respectively. Then the Wald statistics converges to the normal distribution $N(0,1)$ for large n:

$$Z_n = \frac{\bar{Y}_n-\theta}{\sigma_n} \to N(0,1), \tag{24}$$

where $\theta = \mathbb{E}\big[Y\big]$ and we assume $\sigma_n \neq 0$ in (24). Therefore, the probability ratio in (22) converges to

$$\Lambda'(Y^{(n)}) \to Ce^{-\frac{2(\bar{Y}_n-p)\delta}{\sigma_n^2}}, \quad n\to\infty, \tag{25}$$

for some normalizing constant C. In practice, this approximation is sufficiently accurate when the number of samples is $n \geq 30$ and $\mathbb{E}\big[Y\big]$ is not close to the end points 0 and 1,

since the convergence of the probability ratio (25) is fast. When $\mathbb{E}[Y]$ is close to 0 or 1, the distribution $\pi(y)$ of Y will become skewed, and the convergence is slower [1,20]. When the number of strata is $m = 1$, the probability ratio (25) is asymptotically equal to (11) in the large sample limit $n \to \infty$.

Using (25), we can construct a sequential hypothesis testing algorithm with an asymptotic confidence guarantee, as shown in Algorithm 2. Noting that using the Wald statistics (24) in SPRT is asymptotically optimal [20], we have the following theorem.

Theorem 3 *The sampling cost of Algorithm 2 is asymptotically no more than that of the SPRT (11) using i.i.d. samples.*

Algorithm 2 SPRT using stratified samples

Require: Number of strata m, probability threshold p, indifference parameter δ, confidence levels $\alpha, \beta > 0$, and minimum number of samples N

$n \leftarrow 0$
$\nu \leftarrow \{0, \ldots, 0\} \in \mathbb{Z}^{m+1}$
while true **do**
  $n \leftarrow n + 1$
  Take m-stratified samples $\{X_{1,n}, \ldots, X_{m,n}\}$
  $Y_n \leftarrow \sum_{i=1}^{m} \phi(X_{i,n})$
  $\nu(Y_n) \leftarrow \nu(Y_n) + 1$
  if $n \geq N/m$ **then**
    $\mu_n \leftarrow \frac{1}{n} \sum_{i=1}^{m+1} \frac{i-1}{m} \nu(i)$
    $\sigma_n^2 \leftarrow \frac{1}{n^2} \sum_{i=1}^{m+1} \left(\frac{i-1}{m}\right)^2 \nu(i) - \frac{\mu_n^2}{n}$
    if $\mu_n - p < -\frac{\sigma_n^2}{2\delta} \ln(\frac{1-\alpha}{\beta})$ **then**
      return H_0
    else if $\mu_n - p > \frac{\sigma_n^2}{2\delta} \ln(\frac{1-\beta}{\alpha})$ **then**
      return H_1
    end if
  end if
end while

Remark 2 Finally, we note that the stratification can be performed over multiple time steps. Take $m = 16$ as an example. As an alternative to taking 16 strata for a single step, we can take 4 strata over 2 consecutive time steps, and generate 16-stratified samples, each from the 4^2 combinations of strata. The SPRT algorithm for stratified samples over multiple time steps is exactly the same as Algorithm 2.

4 Statistical verification using antithetic samples

Antithetic sampling is another approach to generate negatively correlated random variables. In this section, we show that under Assumption 2, the antithetic sampling technique can be employed to generate pairs of semantically negatively correlated sample paths and reduce the sampling cost for statistical verification.

4.1 Antithetic sampling

Let $\{X_+(t)\}_{t\in\mathbb{N}}$ be a sample path of the DTMC $\mathcal{M}$ driven by random seeds $\{E(t)\}_{t\in\mathbb{N}}$. Then its antithetic sample path is driven by $\{1-E(t)\}_{t\in\mathbb{N}}$, as summarized by Definition 4 and Algorithm 3.

Definition 4 $\{X_+, X_-\}$ is a pair of antithetic samples if they are generated by Algorithm 3.

Algorithm 3 Antithetic sampling

Require: Number of steps T and initial state s
1: $t=0$, $X_+(0)=s$, $X_-(0)=s$
2: **for** $t=1,\ldots,T-1$ **do**
3: Take $E\sim\mathbf{U}_{[0,1)}$
4: $X_+(t+1)=f(X_+(t),E(t))$
5: $X_-(t+1)=f(X_-(t),1-E(t))$
6: **end for**
7: **return** $\{X_+, X_-\}$

4.2 Properties of antithetic samples

Similar to Sect. 3.2, for the PCTL formulas satisfying Assumption 2, a pair of antithetic samples $\{X_+, X_-\}$ are semantically negatively correlated, as stated in Theorem 4.

Theorem 4 *With Assumption 2, let $\{X_+, X_-\}$ be a pair of antithetic samples from Markov chain $\mathcal{M}$ and ϕ be a probabilistic-operator-free PCTL formula with satisfaction probability P_ϕ. Then*

(i) $\mathbb{E}\big[\phi(X_+)+\phi(X_-)\big]/2=P_\phi$;
(ii) $\mathrm{Cov}\big[\phi(X_+),\phi(X_-)\big]\le 0$.

Proof (i) The proof is similar to that of Theorem 1 by Algorithm 3. (ii) Similar to the proof of Theorem 1, it suffices to consider the following two cases.

(1) $\phi=\psi_1\mathcal{U}_{\le T}\psi_2$: By Assumption 2, we number the states of the Markov chain in the same way as in the proof of Theorem 1.

Claim 1: For any $r_1, r_2\in[n]$, we have

$$\begin{aligned}&\mathbb{P}\big[X_+(t+1)\in V_{\psi_1}\mid X_-(t+1)\in V_{\psi_1}, X_+(t)=r_1, X_-(t)=r_2\big]\\&\quad\le\mathbb{P}\big[X_+(t+1)\in V_{\psi_1}\mid X_+(t)=r_1, X_-(t)=r_2\big].\end{aligned}$$

To prove *Claim 1*, it suffices to show that

$$\begin{aligned}\mathcal{P}=&\mathbb{P}\big[X_+(t+1)\in V_{\psi_1}, X_-(t+1)\in V_{\psi_1}\mid X_+(t)=r_1, X_-(t)=r_2\big]\\&-\mathbb{P}\big[X_+(t+1)\in V_{\psi_1}\mid X_+(t)=r_1, X_-(t)=r_2\big]\\&\times\mathbb{P}\big[X_+(t+1)\in V_{\psi_1}\mid X_-(t)=r_1, X_-(t)=r_2\big]\\\le&\,0.\end{aligned}\tag{26}$$

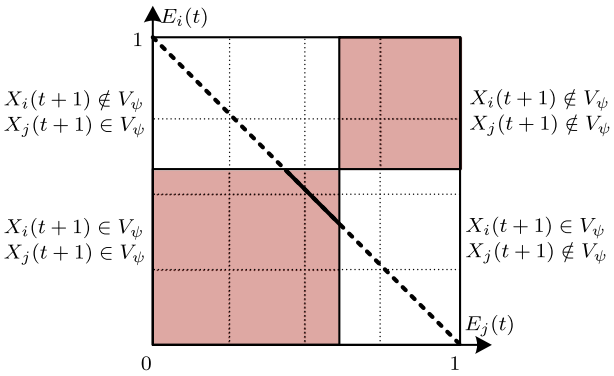


Fig. 2 Joint distribution of two antithetic random seeds

Let $S = \sum_{k=1}^{n_1} M_{kr_1}$ and $R = \sum_{k=1}^{n_1} M_{kr_2}$, where M is the transition probability matrix. Then we have

$$\mathbb{P}\big[X_+(t+1) \in V_{\psi_1} | X_+(t) = r_1, X_-(t) = r_2\big] = \mathbb{P}\big[E_+(t) \le S\big] = S,$$
$$\mathbb{P}\big[X_-(t+1) \in V_{\psi_1} | X_+(t) = r_1, X_-(t) = r_2\big] = \mathbb{P}\big[E_-(t) \le R\big] = R.$$

The two random seeds $E_+(t)$ and $E_-(t)$ are uniformly distributed on the line from (0,1) to (1,0), as shown in Fig. 2. Without loss of generality, assume $S \le R$. In addition, assume $R + S \ge 1$. (The proof for $R + S < 1$ is similar.) Then we have

$$\mathcal{P} = S + R - 1 - SR = -(1 - S)(1 - R) \le 0, \tag{27}$$

where the equalities hold if and only if $S = 1$ or $R = 1$.
(2) $\phi = \mathcal{X}\psi$: The proof is similar to the proof of (2) of Theorem 1

□

4.3 Sequential probability ratio test

To implement the SPRT on antithetic samples $\{X_+, X_-\}$, we consider the statistics

$$Y = \frac{\phi(X_+) + \phi(X_-)}{2}. \tag{28}$$

By Theorem 4, we have $\mathbb{E}\big[Y\big] = P_\phi$, so that verifying $\mathcal{P}_{<p}\phi$ is equivalent to checking the SHT problem as given in (20). In addition, the mean of antithetic samples within each block is more concentrated than the mean of m independent samples with the same mean,

$$\begin{aligned} \mathrm{Var}\big[Y\big] &= \frac{1}{4}\mathrm{Var}\big[\phi(X_+) + \phi(X_-)\big] \\ &= \frac{1}{2}\mathrm{Var}\big[\phi(X_+)\big] + \frac{1}{2}\mathrm{Cov}\big[\phi(X_+), \phi(X_-)\big] \\ &\le \frac{1}{2}\mathrm{Var}\big[\phi(X_+)\big]. \end{aligned} \tag{29}$$

In addition, by the same proof as Theorem 2, we know that Y is a sufficient statistic.

Theorem 5 *The joint probability mass function $\pi(x_+, x_-)$ of $\phi(X_+)$ and $\phi(X_-)$ only depends on $\phi(X_+) + \phi(X_-)$. Thus, to construct a statistical test for the satisfaction probability P_ϕ of the formula ϕ using antithetic samples (X_+, X_-), it suffices to use the average statistic $Y = (\phi(X_+) + \phi(X_-))/2$.*

Given the n independent samples $Y^{(n)} = \{Y_1, \ldots, Y_n\} \subseteq \{0, 1/2, 1\}$ of the average statistic Y, we can construct an SPRT algorithm with asymptotic confidence guarantee in the same way as Sect. 3.3, as shown in Algorithm 4. The asymptotic optimality of the algorithm is stated in Theorem 6.

Theorem 6 *The sampling cost of Algorithm 4 is asymptotically no more than that of the SPRT (11) using i.i.d. samples.*

Algorithm 4 SPRT using antithetic samples.

Require: Probability threshold p, indifference parameter δ, confidence levels $\alpha, \beta > 0$, and minimum number of samples N

$n \leftarrow 0$
$\nu \leftarrow \{0, 0, 0\} \in \mathbb{Z}^3$
while true **do**
  $n \leftarrow n + 1$
  Take antithetic samples $\{X_{+,n}, X_{-,n}\}$
  $Y_n \leftarrow \phi(X_{+,n}) + \phi(X_{-,n})$
  $\nu(Y_n) \leftarrow \nu(Y_n) + 1$
  if $n \geq N/2$ **then**
    $\mu_n \leftarrow \frac{1}{n} \sum_{i=1}^{3} \frac{i-1}{2} \nu(i)$
    $\sigma_n^2 \leftarrow \frac{1}{n^2} \sum_{i=1}^{3} \left(\frac{i-1}{2}\right)^2 \nu(i) - \frac{\mu_n^2}{n}$
    if $\mu_n - p < -\frac{\sigma_n^2}{2\delta} \ln(\frac{1-\alpha}{\beta})$ **then**
      return H_0
    else if $\mu_n - p > \frac{\sigma_n^2}{2\delta} \ln(\frac{1-\beta}{\alpha})$ **then**
      return H_1
    end if
  end if
end while

Remark 3 Finally, we note that the stratified and antithetic sampling techniques can be combined in the following two manners: (i) divide the sample space, specifically the $[0, 1]$ interval, into n strata and pick a pair of antithetic samples within each strata; or (ii) pick n-strata samples for half of the sample space and generate n antithetic samples for the other half.

5 Simulation

In this section, we evaluate Algorithms 2 and 4 on three benchmarks from PRISM [8].
Bounded Retransmission Protocol (BRP) [3,5] is a variant of the Alternating Bit Protocol for sending a large file in `N` chunks. The maximal number of retransmissions for each chunk is `MAX`. We refer to [5] for the details of the protocol and use the PRISM implementation at http://www.prismmodelchecker.org/casestudies/brp.php.
Crowds Protocol (CP) [13,18,19] provides a mechanism for anonymous web browsing by "blending into a crowd". We use the PRISM implementation of the protocol at

http://www.prismmodelchecker.org/casestudies/crowds.php. The size parameters of the models are `TotalRuns`, which is the total number of protocol runs to analyze, and `CrowdSize`, which is the actual number of good crowd members.

EGL Contract Signing Protocol (EGL) [4,12] is a randomized protocol for signing contracts. It provides fair data exchange, where either all participants obtain what they want, or none do. We use the PRISM implementation of the protocol at http://www.prismmodelchecker.org/casestudies/contract_egl.php. The size parameters of the models are `N`, which is the number of pairs of secrets, and `L`, which is the number of bits in each secret.

We implemented Algorithms 2 and 4 in Java and their average running times and sample costs are given in Figs. 3, 4, and 5 with comparisons to existing methods. The sample paths were drawn directly from the PRISM simulator. All experiments were run on Ubuntu 18.04 with i7-8700 CPU 3.2GHz and 16GB memory. We checked the specifications $\mathcal{P}_{<0.39}\mathbf{F}_{[0,99]}(s = 3)$ for BRP, $\mathcal{P}_{<0.15}\mathbf{F}_{[0,99]}(\texttt{observe}_0 > 1)$ for CP, and $\mathcal{P}_{<0.51}\mathbf{F}_{[0,99]}(\neg \texttt{kA} \wedge \texttt{kB})$ for EGL, and refer to https://www.prismmodelchecker.org/casestudies for their meanings. The indifference parameter and the confidence parameters for all the simulations were set to $\alpha = \beta = \delta = 10^{-3}$ for BRP and EGL, and $\alpha = \beta = \delta = 10^{-4}$ for CP. Each simulation setup was repeated for 50 times and error bars corresponding to 95% confidence intervals are shown in Figs. 3, 4 and 5.

We compare the proposed algorithms with five different algorithms in PRISM: the standard SPRT and four symbolic methods. The symbolic methods are different in the way that a transition probability matrix is represented, i.e., multi terminal binary decision diagram (MTBDD), sparse matrix, hybrid (developed to overcome the inefficiencies with MTBDD), and explicit matrix[1]. For all methods, we compare their running time; and for statistical methods, the sample costs are also compared. The running time limit of each algorithm was set to 30 minutes.

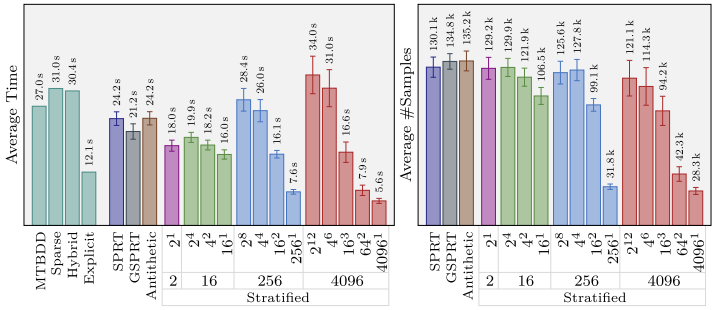
In the simulations, the smallest model has more than 850 000 states and the largest one has more than 135×10^{12} states. As shown in Figs. 3, 4, and 5, compared to symbolic methods, the proposed statistical model checking algorithms scale much better, even when the parameters are set conservatively by $\alpha = \beta = \delta = 10^{-3}$ or $\alpha = \beta = \delta = 10^{-4}$. For the largest models, the symbolic algorithms did not complete, as they ran out of memory for storing the transition probability matrix, even when memory-saving engines such as Sparse or Hybrid were used.

As shown in Figs. 3, 4, and 5, for Algorithm 2, different stratification strategies are evaluated. Specifically, we consider strata sizes $m = 1$ (which we referred to as GSPRT), $m = 2$, $m = 16$, $m = 256$, and $m = 4096$. As discussed in Remark 3, for each m, the stratification is performed for over one or more time steps. For example for $m = 256$, we consider 4 different ways of stratification: 256 strata for 1 time step, denoted by 256^1; 2 strata for 8 consecutive time steps, denoted by 2^8; and similarly for 4^4 and 8^2.

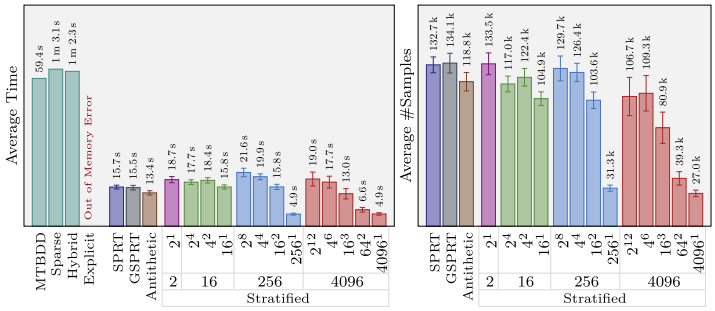
The simulation results show that using more strata in one step results in a smaller number of total samples. Using a well-chosen number of strata can significantly reduce the average number of samples. For example, in Fig. 3a, using 4096^1 strata reduces the average number of samples from about 134 800 to 28 300 (4.76 times smaller). In Fig. 4, the reduction for 4096^1 strata is still about 30%. In Fig. 5d, a significant reduction of 150 times is observed for 4096^1 strata.

Generally, the sample costs decrease as the number of strata increases. For example, increasing the number of strata from 2 to 256 significantly reduces the sample costs in Fig. 3, and similarly for increasing the number of strata from 256 to 4096 in Fig. 4. However,

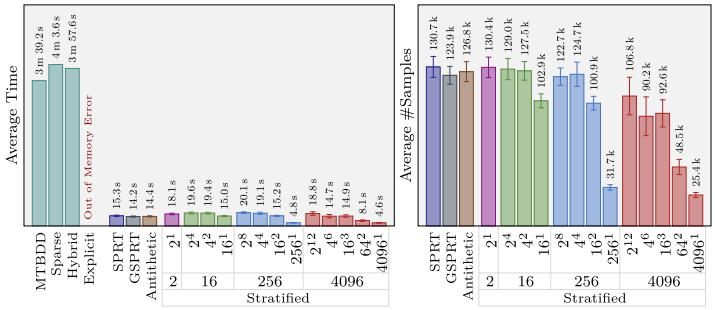
[1] We did not use the 'Exact' engine, since it does not support bounded $\mathcal{U}$ formulas



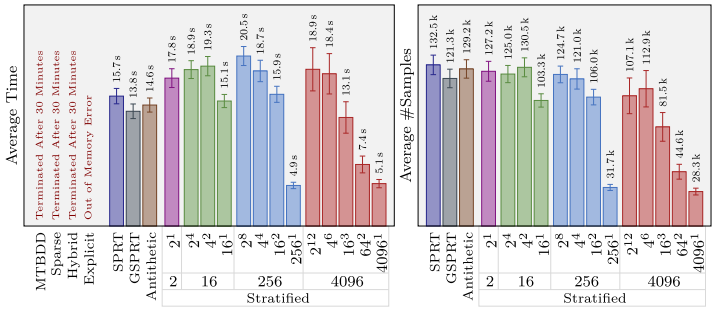
(a) MAX: 15 N: 4,096 States: 864,274 Transitions: 1,179,651



(b) MAX: 20 N: 8,192 States: 2,261,015 Transitions: 3,096,579

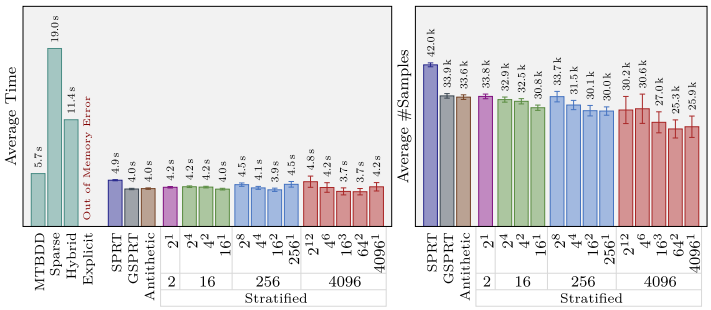


(c) MAX: 64 N: 16,384 States: 13,893,699 Transitions: 19,169,283

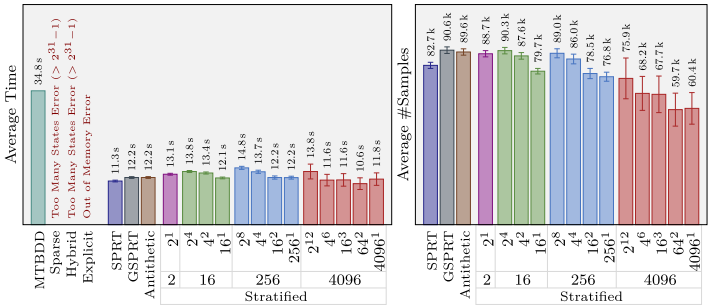


(d) MAX: 256 N: 65,536 States: 219,152,643 Transitions: 303,169,539

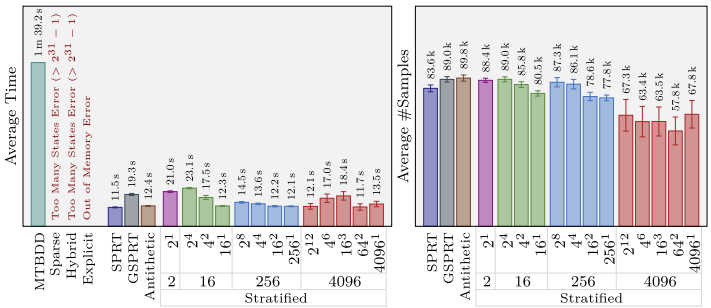
Fig. 3 Bounded retransmission protocol



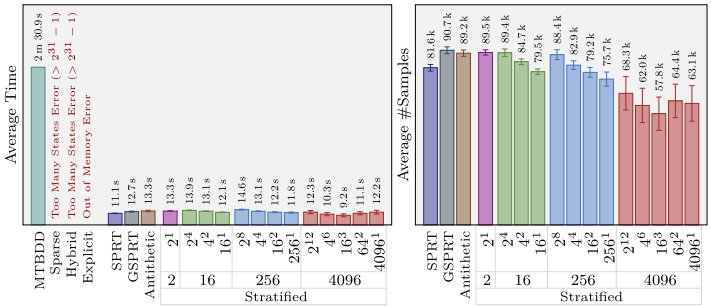
(a) CrowdSize: 20 TotalRuns: 6 States: 10,633,591 Transitions: 38,261,191



(b) CrowdSize: 20 TotalRuns: 11 States: 6,983,580,046 Transitions: 25,611,489,346

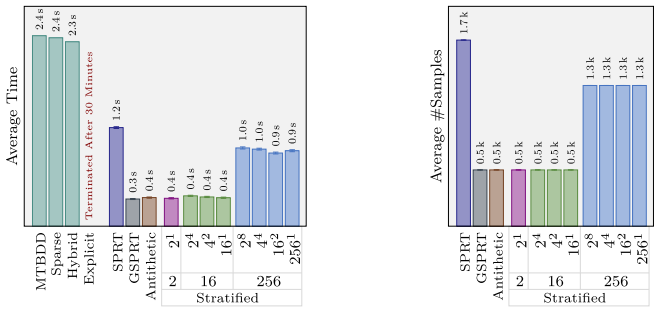


(c) CrowdSize: 20 TotalRuns: 16 States: 862,908,898,831 Transitions: 3,201,427,974,031

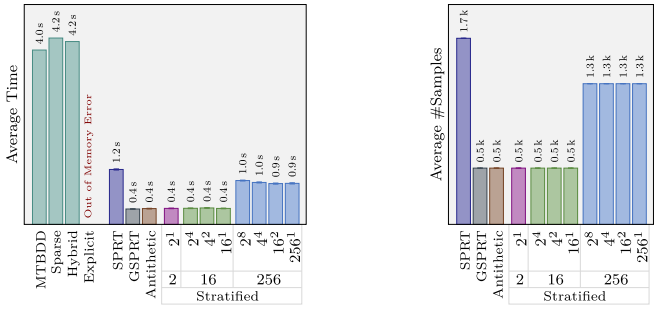


(d) CrowdSize: 20 TotalRuns: 20 States: 20,158,413,809,821 Transitions: 75,297,025,337,821

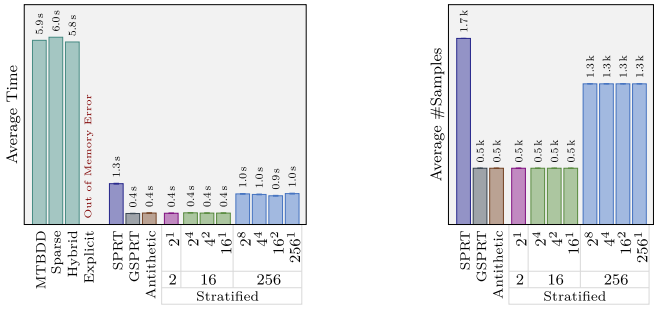
Fig. 4 Crowds protocol



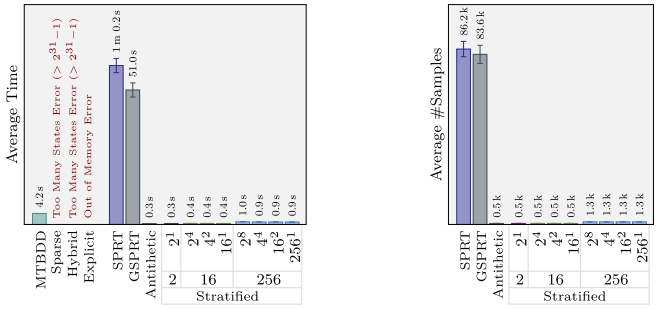
(a) L: 8 N: 8 States: 15,925,246 Transitions: 15,990,781



(b) L: 8 N: 10 States: 317,718,526 Transitions: 318,767,101



(c) L: 8 N: 12 States: 6,090,129,406 Transitions: 6,106,906,621



(d) L: 2 N: 20 States: 135,239,930,216,446 Transitions: 136,339,441,844,221

Fig. 5 EGL contract signing protocol

over-stratification can be harmful as it increases the minimum number of samples N in Algorithms 2 and 4. For example, in Fig. 5 the number of samples needed for SPRT is about 500, so increasing the number of strata to 256 makes N larger than needed. In these cases, the sample costs are exactly N. For better visualization, we do not show results for strata size 4096 in Fig. 5. Finally, the comparison of the results for Algorithm 2 and Algorithm 4 shows that the efficiency of antithetic samples is similar to 2-strata samples.

6 Conclusion

In this work, we discussed the advantage of using stratified and antithetic samples to statistically verify Probabilistic Computation Tree Logic (PCTL) formulas on discrete-time Markov chains (DTMCs). We showed that by properly choosing the representation of the DTMCs, semantically negatively correlated samples can be generated for a fraction of PCTL formulas using the stratified or antithetic sampling techniques. Based on this, we proposed statistical verification algorithms with asymptotic correctness guarantees using stratified or antithetic samples, and demonstrated that these algorithms are more sample-efficient than the algorithms using independent Monte Carlo sampling. The experiments showed that our proposed algorithms use 30%–60% fewer samples (number of strata $\times$ number of blocks of stratified samples) than existing independent-samples methods, for a given confidence level on the benchmark examples.

Acknowledgements This work was supported by NSF CPS Grant 1329991 and AFOSR Grant FA9550-15-1-0059.

References

1. Agresti A, Coull BA (1998) Approximate is better than "exact" for interval estimation of binomial proportions. Am Stat 52(2):119–126
2. Clarke EM, Zuliani P (2011) Statistical model checking for cyber-physical systems. Automated technology for verification and analysis. Springer, Berlin, pp 1–12
3. D'Argenio P, Jeannet B, Jensen H, Larsen K (2001) Reachability analysis of probabilistic systems by successive refinements. In: de Alfaro L, Gilmore S (eds) Proceedings of 1st joint international workshop on process algebra and probabilistic methods, performance modelling and verification (PAPM/PROBMIV'01). Springer, LNCS, vol 2165, pp 39–56
4. Even S, Goldreich O, Lempel A (1985) A randomized protocol for signing contracts. Commun ACM 28(6):637–647
5. Helmink L, Sellink M, Vaandrager F (1994) Proof-checking a data link protocol. In: Barendregt H, Nipkow T (eds) Proceedings of international workshop on types for proofs and programs (TYPES'93). Springer, LNCS, vol 806, pp 127–165
6. Henriques D, Martins JG, Zuliani P, Platzer A, Clarke EM (2012) Statistical model checking for markov decision processes. In: 2012 Ninth international conference on quantitative evaluation of systems, pp 84–93
7. Hermanns H, Nielson F, Jansen DN, Zhang L (2012) Efficient csl model checking using stratification. Log Methods Comput Sci 8:1–18
8. Kwiatkowska M, Norman G, Parker D (2011) Prism 4.0: Verification of probabilistic real-time systems. In: International conference on computer aided verification. Springer, pp 585–591
9. Larsen KG, Legay A (2016) Statistical model checking: past, present, and future. Leveraging applications of formal methods, verification and validation: foundational techniques. Springer, Cham, pp 3–15
10. Liu J (2008) Monte Carlo strategies in scientific computing. Springer, Cham
11. Maginnis PA, West M, Dullerud GE (2016) Variance-reduced simulation of lattice discrete-time markov chains with applications in reaction networks. J Comput Phys 322:400–414
12. Norman G, Shmatikov V (2006) Analysis of probabilistic contract signing. J Comput Secur 14(6):561–589

13. Reiter M, Rubin A (1998) Crowds: anonymity for web transactions. ACM Trans Inf Syst Secur (TISSEC) 1(1):66–92
14. Roohi N, Wang Y, West M, Dullerud GE, Viswanathan M (2017) Statistical verification of the Toyota powertrain control verification benchmark. In: Proceedings of the 20th international conference on hybrid systems: computation and control. ACM, pp 65–70
15. Sen K, Viswanathan M, Agha G (2004) Statistical model checking of black-box probabilistic systems. In: Alur R, Peled DA (eds) computer aided verification. Springer, Berlin, Heidelberg, no. 3114 in Lecture Notes in Computer Science, pp 202–215
16. Sen K, Viswanathan M, Agha G (2005) On statistical model checking of stochastic systems. In: Etessami K, Rajamani SK (eds) Computer aided verification. Springer, Berlin, Heidelberg, no. 3576 in Lecture Notes in Computer Science, pp 266–280
17. Sen K, Viswanathan M, Agha G (2005) Vesta: A statistical model-checker and analyzer for probabilistic systems. In: Second international conference on the quantitative evaluation of systems, 2005, pp 251–252
18. Shmatikov V (2002) Probabilistic analysis of anonymity. In: Proceedings of the 15th IEEE computer security foundations workshop (CSFW'02). IEEE Computer Society Press, pp 119–128
19. Shmatikov V (2004) Probabilistic model checking of an anonymity system. J Comput Secur 12(3/4):355–377
20. Tony Cai T (2005) One-sided confidence intervals in discrete distributions. J Stat Plan Inference 131(1):63–88
21. Wang Y, Roohi N, West M, Viswanathan M, Dullerud GE (2015) A mori-zwanzig and mitl based approach to statistical verification of continuous-time dynamical systems. IFAC-PapersOnLine 48(27):267–273
22. Wang Y, Roohi N, West M, Viswanathan M, Dullerud GE (2015) Statistical verification of dynamical systems using set oriented methods. In: Proceedings of the 18th international conference on hybrid systems: computation and control. ACM, New York, HSCC '15, pp 169–178
23. Wang Y, Roohi N, West M, Viswanathan M, Dullerud GE (2016) Verifying continuous-time stochastic hybrid systems via mori-zwanzig model reduction. In: 2016 IEEE 55th conference on decision and control (CDC), pp 3012–3017
24. Wang Y, Roohi N, West M, Viswanathan M, Dullerud GE (2018) Statistical verification of pctl using stratified samples. IFAC-PapersOnLine 51(16):85–90
25. Younes HLS (2005) Ymer: a statistical model checker. In: Etessami K, Rajamani SK (eds) Computer aided verification. Springer, Berlin, no. 3576 in Lecture Notes in Computer Science, pp 429–433
26. Younes HLS, Simmons RG (2006) Statistical probabilistic model checking with a focus on time-bounded properties. Inf Comput 204(9):1368–1409
27. Zuliani P, Baier C, Clarke EM (2012) Rare-event verification for stochastic hybrid systems. In: Proceedings of the 15th ACM international conference on hybrid systems: computation and control. ACM, New York, HSCC '12, pp 217–226