Fast Simulation for Slow Paths in Markov Models

Daniël Reijsbergen¹, Pieter-Tjerk de Boer¹, Werner Scheinhardt¹, and Boudewijn Haverkort¹,²

¹Center for Telematics & Information Technology, University of Twente, Enschede, The Netherlands
²Embedded Systems Institute, Eindhoven, The Netherlands
{d.p.reijsbergen, p.t.deboer, w.r.w.scheinhardt, b.r.h.m.haverkort}@utwente.nl

Abstract—Inspired by applications in the context of stochastic model checking, we are interested in using simulation for estimating the probability of reaching a specific state in a Markov chain after a large amount of time τ has passed. Since this is a rare event, we apply importance sampling. We derive approximate expressions for the sojourn times on a given path in a Markov chain conditional on the sum exceeding τ, and use those expressions to construct a change of measure. Numerical examples show that this change of measure performs very well, leading to high precision estimates in short simulation times.

I. INTRODUCTION

Stochastic model checking is an increasingly important tool to support the design process of a variety of systems. The systems are modelled using a formalism like Petri nets, Markov reward models (MRMs), etc., and properties of these models are then verified [3]. Increasingly, these properties are stochastic in nature, and they often involve events that are hopefully rare, such as system failures.

Many methods for verifying such properties are known, but in the case of complex stochastic systems, statistical model checking using simulation is often the only feasible method. However, in order to efficiently simulate rare events, special techniques are needed. In a recent paper [5], an importance-sampling-based rare-event simulation method was developed for estimating probabilities of events of the form “absorption before a specific time” in a broad class of absorbing continuous-time Markov chains (CTMCs).

In the present work, we consider the opposite event, namely “absorption after a specific amount of time has passed”. While rarely studied in the rare event simulation literature, it is particularly motivated by MRMs where the event of interest is “collecting sufficient reward before absorption”. Both problems can be shown to be equivalent [2].

For estimating such probabilities in general CTMCs, we envision a two-step approach: in the first step of each simulation run, the simulator samples a path (i.e., a sequence of states) through the chain, and in the second step it samples the sojourn times for that path. The present paper presents work in progress about the second subproblem: the probability of interest (i.e., that the sum of the sojourn times of a given path in a CTMC exceeds some threshold) is known in closed form [1], but its numerical evaluation is computationally expensive. Therefore, we derive an efficient importance sampling simulation algorithm for it, drawing sojourn times from a distribution that closely resembles the conditional distribution given the rare event of interest.

The rest of this paper is organised as follows. In Section II we study the conditional distribution of the sojourn times. In Section III we briefly introduce importance sampling simulation, and describe our algorithm. The good performance of our algorithm is illustrated experimentally in Section IV, and Section V provides some conclusions.

II. CONDITIONAL SOJOURN TIMES

As noted above, we assume that a path ϕ through the Markov chain is already given, consisting of n states \( x_1, \ldots, x_n \); only the sojourn times in the states on this path are unknown, but the rates of the states are given as \( q_1, \ldots, q_n \), some of which may be identical. This path itself can be seen as an absorbing Markov chain on its own, as depicted in Figure 1. We now proceed to analyze the behaviour of the sojourn times \( T_j \) in the individual states \( j \) of this path, conditional on absorption occurring after some time bound \( \tau \). The results of this section will be used in Section III to obtain an efficient simulation algorithm.

The probability density of the sojourn time \( T_j \) is given by \( f_j(x) = q_j e^{-q_j x} \), but we are interested in the distribution of \( T_j \) conditional on occurrence of the event \( T > \tau \), where \( T \triangleq \sum_{j=1}^{n} T_j \). Considering without loss of generality \( j = 1 \), we condition on the value of \( T_1 \) to find

\[ P(T_1 > t | T > \tau) = \int_t^{\infty} \frac{f_1(t_1)}{P(T > \tau)} P(T - T_1 > \tau - t_1) \, dt_1 \]

and hence

\[
 f_1(t | T > \tau) = \begin{cases} \frac{f_1(t)}{P(T > \tau)} P(T - T_1 > \tau - t) & \text{if } t < \tau, \\ \frac{f_1(t)}{P(T > \tau)} & \text{otherwise}. \end{cases}
\]

This expression contains the probability \( P(T > \tau) \) which we are trying to estimate. Therefore our goal is now to obtain insight into the behaviour of \( f_1(t | T > \tau) \) for large \( \tau \) so we can construct a good approximation in the next section.

![Figure 1. Path ϕ, seen as a Markov birth process.](image-url)
We start by making (1) explicit for a two-state path \( \phi = (x_1, x_2) \) with rates \( q_1 \) and \( q_2, q_1 \neq q_2 \). Then

\[
\begin{aligned}
f_1(t|T > \tau) &= \begin{cases} 
  \frac{q_1 e^{-(q_1 - q_2)t}}{q_1 - q_2} & \text{if } t < \tau, \\
  \frac{q_1 e^{-q_1 t}}{q_1 - q_2} + \frac{q_2 e^{-(q_2 - q_1)\tau}}{q_2 - q_1} & \text{otherwise.}
\end{cases}
\end{aligned}
\]  

(2)

Note that the same expression holds for \( f_2(t|T > \tau) \) after interchanging \( q_1 \) and \( q_2 \).

The shape of this function for \( t > \tau \) is always exponential with rate \( q_1 \). However, the shape of the part where \( t < \tau \) depends on the parameter setting, where we distinguish three cases. For \( q_1 > q_2 \), this part is still negative exponential albeit with a different parameter, namely \( q_1 - q_2 \). However, for \( q_1 < q_2 \), this part is positive exponential, again with parameter \( q_1 - q_2 \). In between, as \( q_1 \) and \( q_2 \) become equal, this part approaches a constant. This can be seen in Figure 2.

In Figure 3 the time bound \( \tau \) was increased sixfold, illustrating the limit behaviour of the system for large \( \tau \). For \( q_1 > q_2 \) we see that the probability mass right of \( \tau \) vanishes, so we can approximate the function (2) by a simple exponential density with rate \( q_1 - q_2 \).

It is also interesting to observe how the expected share of the burden of consuming \( \tau \) time units is distributed over the states. One easily derives the following from (2):

\[
\mathbb{E}(T_1|T > \tau) \sim \begin{cases} 
  \tau/2 & \text{if } q_1 < q_2 \\
  (q_2 - q_1)^{-1} & \text{if } q_1 > q_2.
\end{cases}
\]

with \( \sim \) meaning that the ratio of left- and right-hand side goes to 1 as \( \tau \to \infty \). This is illustrated in Figure 4. We see that when the rates differ, almost all of the time \( \tau \) is typically spent in the state with the lowest rate, while the time spent in the other state tends to a constant.

These core observations do not just hold for two-state paths but for any path \( \phi \). Denote the lowest rate by \( \beta_1 \) and the second-lowest by \( \beta_2 \), and let \( r_i \) be the number of times rate \( \beta_i \) occurs on the path. Then, in the limit for large \( \tau \), a state \( i \) whose rate \( q_i \neq \beta_1 \) will contribute only an exponentially distributed amount of time with the bounded mean \((q_i - \beta_1)^{-1}\).

If \( r_1 = 1 \), then the single state with rate \( \beta_1 \) will account for an amount that has an asymmetric Laplace distribution peaking at \( t = \tau \) with rates \( \beta_2 - \beta_1 \) on the left side and \( \beta_1 \) on the right side. If there are \( r_1 > 1 \) states with rate \( \beta_1 \), then the expected contribution of each of these states is \( \tau/r_1 \), and the conditional sojourn time in each state has an exponential distribution with rate \( \beta_1 \) to the right of \( \tau \), but a polynomial density with degree \( r_1 - 2 \) left of \( \tau \).

III. SIMULATION

We now proceed to construct an efficient simulation estimator for our probability of interest, namely \( P(T > \tau) \) given a path \( \phi \). The standard simulation estimator for this is

\[
\hat{P} = \frac{1}{N} \sum_{i=1}^{N} \hat{1}_{1_{i,j} > \tau},
\]

where \( t_{ij} \) is the sampled sojourn time (with density \( f_j(t) \)) for state \( j \) in the \( i \)th simulation run, \( N \) is the number of simulation runs, and \( 1 \) the indicator function. This approach is very inefficient when the target event is rare. A remedy is importance sampling [4], where the samples \( t_{ij} \) are drawn from a different density \( f_j^*(t) \) and weighted by a likelihood ratio:

\[
\hat{P}^* = \frac{1}{N} \sum_{i=1}^{N} \prod_{j=1}^{n} \frac{f_j(t_{ij})}{f_j^*(t_{ij})} \hat{1}_{1_{i,j} > \tau},
\]

Figure 2. \( f_1(t|T > \tau) \) for different parameter values \( q_1 \) and \( q_2 \), with \( \tau = 5 \). Solid line: \( q_1 = 2.4, q_2 = 2 \). Dotted line: \( q_1 = 2.2, q_2 = 2.2 \). Dashed line: \( q_1 = 2, q_2 = 2.4 \).

Figure 3. Same choices for the values of the parameters \( q_1 \) and \( q_2 \) as in Figure 2, but with \( \tau = 30 \).

Figure 4. Expected sojourn times as a function of \( \tau \).
Since the exact calculation of $f_j(t | T > \tau)$ is problematic in general, we propose to use the following approximation instead, inspired by the findings in the previous section:

$$f_j^*(t) = \begin{cases} (q_j - \beta_1) \cdot e^{-(q_j - \beta_1)t} & \text{if } q_j > \beta_1 \\ r_1 / \tau \cdot e^{-r_1/\tau t} & \text{if } q_j = \beta_1 \text{ and } r_1 = 1 \\ g(t | \beta_1, \beta_2) & \text{otherwise}, \end{cases} \tag{5}$$

with $\beta_i$ and $r_i$ defined as before, and $g(t | \beta_1, \beta_2)$ is given by the r.h.s. of (2) with each $q_i$ replaced by $\beta_i$.

In practical applications where the Markov chain is not a pure-birth Markov process, the above algorithm for each simulation run $i$ should be preceded by a phase in which the path itself (i.e., the set of states) is sampled, possibly also using importance sampling (cf. the two-phase approach discussed in Section I).

### IV. Simulation Results

In this section we empirically demonstrate the effectiveness of the method. Throughout this section, all results are based on $10^6$ simulation runs. We compare standard Monte Carlo (MC) simulation using (3) to our importance sampling (IS) approach using (4) and (5). In the first two examples, we also give the true values for $p$, computed directly using the Erlang and hypoexponential distribution functions.

In Table I, we consider a two-state path $\phi$ with unequal rates. We see that the method works well; the very slow increase of the relative error (r.e.) (defined as 1.96 times the sample standard deviation of the estimator $\hat{p}$ (or $\hat{p}^*$) divided by the sample mean $\hat{p}$ (or $\hat{p}^*$) itself) as $\tau$ becomes bigger, suggests the relative error is in fact upper-bounded.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$\hat{p}$</th>
<th>MC-r.e.</th>
<th>$\hat{p}^*$</th>
<th>IS-r.e.</th>
<th>true</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2.52E-4</td>
<td>0.1235</td>
<td>2.417E-4</td>
<td>0.0047</td>
<td>2.417E-4</td>
</tr>
<tr>
<td>7</td>
<td>8.0E-6</td>
<td>0.6929</td>
<td>4.71E-6</td>
<td>0.0054</td>
<td>4.73E-6</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>—</td>
<td>8.947E-8</td>
<td>0.0060</td>
<td>8.93E-8</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>—</td>
<td>8.3E-89</td>
<td>0.0078</td>
<td>8.3E-89</td>
</tr>
</tbody>
</table>

Table I

Simulation results, $q_1 = 2$, $q_2 = 2.4$.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$\hat{p}$</th>
<th>MC-r.e.</th>
<th>$\hat{p}^*$</th>
<th>IS-r.e.</th>
<th>true</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>2.03E-4</td>
<td>0.1375</td>
<td>2.011E-4</td>
<td>0.0058</td>
<td>2.004E-4</td>
</tr>
<tr>
<td>7</td>
<td>3.0E-6</td>
<td>1.1316</td>
<td>3.372E-6</td>
<td>0.0070</td>
<td>3.363E-6</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>—</td>
<td>6.29E-9</td>
<td>0.0279</td>
<td>6.3E-9</td>
</tr>
</tbody>
</table>

Table II

Simulation results, $q_1 = q_2 = 2.2$.

In Table II, we set $q_1 = q_2$. The results are still good but somewhat less accurate, which can be explained by the poor resemblance between $f_j^*(t)$ and $f_j(t | T > \tau)$ for $\tau < \tau$.

Finally, in Table III we show results for a path with 50 states and 25 different rates. Note that direct calculation of the true probability is not numerically feasible in this case. Having demonstrated that the method works well for the pure-birth processes for which it was intended (as the second step of the two-step approach), we also give an example derived from a Markov-reward model involving an M/M/5/5 queue.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$\hat{p}$</th>
<th>MC-r.e.</th>
<th>$\hat{p}^*$</th>
<th>IS-r.e.</th>
<th>true</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>2.092E-2</td>
<td>0.0134</td>
<td>2.097E-2</td>
<td>0.0051</td>
<td>—</td>
</tr>
<tr>
<td>20</td>
<td>1.4E-5</td>
<td>0.5238</td>
<td>1.727E-5</td>
<td>0.0070</td>
<td>—</td>
</tr>
<tr>
<td>100</td>
<td>0</td>
<td>—</td>
<td>2.19E-39</td>
<td>0.0180</td>
<td>—</td>
</tr>
</tbody>
</table>

Table III

Simulation results, $q_i = \lceil \frac{i+1}{5} \rceil$, $i = 1, \ldots, 50$.

### V. Conclusion and Future Work

We have found explicit results and useful approximations for the conditional distribution of sojourn times on a given path in a Markov chain, given that their sum exceeds a bound. The resulting expressions are relatively simple and yield insight into how this rare event typically happens. Based on these insights we have constructed an importance sampling change of measure and shown that performs well. Future research will first focus on more general Markov chains, where we need to identify the most probable paths leading to the rare event, after which we can apply the method presented here to those paths. Also it will be interesting to consider (rare) events in which both a time- and reward-bound play a role.

### REFERENCES


