Markov chains with perturbed rates to absorption: Theory and application to model repair

Alexander Gouberman, Markus Siegle, Bharath Tati

Informatik III, Universität der Bundeswehr München

Abstract

This work investigates properties of continuous-time Markov chains with absorbing states, where there is no restriction on the structure of the set of transient states. The paper studies the behaviour of the hitting probabilities, as the rates from transient to absorbing states are perturbed by a common factor. New results about monotonicity and asymptotic limits of those hitting probabilities are established. The theoretical findings are applied to an instance of a model repair problem on a fairly general class of Markov chains, which appears frequently when modelling technical systems subject to ageing and failure.

Keywords: Markov Chain, Perturbation, Hitting Probability, Monotonicity, Asymptotics, CSL Model Checking, Model Repair

1. Introduction

This paper investigates properties of continuous-time Markov chains (CTMC) with at least two absorbing states, no other recurrent classes, and a set of transient states. There is no restriction on the set of transient states, i.e. the transient states may be partitioned into more than one communication class. In particular, we study such CTMCs where all rates into the absorbing states are multiplied by a small perturbation factor. We analyze in detail the behaviour of the hitting probabilities (also known as trapping probabilities) – seen as functions of the perturbation factor – from the individual transient states to the different absorbing states. As a major result, we show that the

Email addresses: `alexander.gouberman@unibw.de` (Alexander Gouberman), `markus.siegle@unibw.de` (Markus Siegle), `bharath.tati@unibw.de` (Bharath Tati)

individual hitting probabilities are not necessarily monotonic with respect to the perturbation factor, but, interestingly, their enveloping functions are indeed monotonic. Our further attention is on the limit of the hitting probabilities as the perturbation factor goes to zero. We show how for a given transient state as initial state, this limit can be calculated from the stationary distribution among transient states (when transitions to absorbing states are ignored) and the rates from transient to absorbing states. We also show that the limits of the hitting probabilities for all transient states belonging to the same bottom communication class[1] (within the set of transient states) coincide. In order to prove these facts, we use properties of M-matrices and the concept of the Drazin inverse, which allows us to perform the analysis in terms of simpler algebraic manipulations. The so-called exitpoint Markov chain, a DTMC associated with the transient states of the original CTMC, which characterizes absorbing paths, also plays an important role in our argumentations.

It is important to point out that the methods presented in this paper cannot only be applied to Markov chains which are absorbing a priori, but to general Markov chains, where one is interested in the question of eventually hitting certain sets of good or bad states.

For us, the question of monotonicity of hitting probabilities arose while we developed an algorithm to solve a particular instance of model repair. In general, the model repair problem is to fix a system (or rather a model thereof) in case it does not satisfy some desirable property. Earlier work on model repair of probabilistic systems can be found, e.g., in [1, 2, 3]. We are interested in model repair problems arising in the context of CTMCs labelled with state properties, where requirements are expressed with the help of continuous stochastic logic (CSL) [4], a temporal logic that has become very popular and can be automatically checked with tools such as PRISM [5] or Storm [6]. In Sec. 5 of this paper, we look at a special (but still quite general and widely applicable) class of Markov chains, representing systems that pursue a mission while being subject to ageing and risky activities, which might lead to failure. This has numerous applications, for instance in reliability and risk analysis of spaceborne systems. For this class of Markov chain, we look at a typical time-bounded reach-avoid requirement, asserting that with high probability the system should be continuously operational

[1] A communication class is called bottom if there are no transitions leaving the class.

for a given minimum time span, eventually terminating its mission successfully. This requirement is expressed by the CSL until operator with lower time bound and lower probability bound. For the case that the requirement is violated, the paper proposes an algorithm how to repair the model, by uniformly reducing certain subsets of its transition rates. Depending on the case at hand, either one or two reduction factors are employed. It is exactly the monotonicity property derived in the earlier sections which ensures that our model repair strategy will be always successful, i.e. it is shown that the proposed algorithm will always succeed in repairing the model.

Related work: Absorbing Markov chains, also known as lossy Markov chains, have received much attention for a long time, see e.g. [7] for an early paper. One of the interesting issues about them is their quasi-stationary distribution, i.e. the kind of equilibrium attained after a long time, provided that absorption has not yet happened. From a different point of view, an absorbing Markov chain can also be seen as a phase-type distribution [8, 9], a powerful class of probability distributions frequently used for the fitting of traffic traces [10], where the matching of moments [11] and finding canonical representations [12] are prime concerns. In that context, the focus is on the distribution of the time to absorption, and there is usually only a single absorbing state. However, our concern is neither the quasi-stationary distribution nor the time to absorption, but the hitting behaviour in case of perturbation of the absorption rates. To the best of our knowledge, the question of monotonicity of hitting probabilities, in the context of perturbed absorbing Markov chains with more than one absorbing state, has not received any attention in the literature, apart from our recent conference paper [13]. However, the present paper generalizes the results of [13] since it drops the restrictive assumption that the transient states of the CTMC form a single communication class.

Among related research, it is important to mention the large body of work on parametric Markov chains and parametric Markov decision processes, which has varying goals and focus. In such parametric models, transition probabilities or transition rates are commonly given by polynomials or rational functions of real-valued parameters. Approaches to parameter synthesis can be found, e.g., in [14, 15] and the recent paper [16], where strategies have been proposed to find valid parameter values in a multi-dimensional search space. A method for analyzing parametric Markovian models was described in [17], building on the pioneering work [18]. It computes a closed-form rational function for the desired reachability probability, following a

state-elimination strategy, and is also able to deal with rewards and nondeterminism. These algorithms are implemented in several tools such as PARAM [19], the parametric extension of PRISM [5] and PROPhESY [20]. Improved implementation strategies for those algorithms have recently been presented in [21]. Our focus in this paper is different: Rather than synthesizing satisfying parameter values or calculating reachability probabilities in symbolical form, we wish to show some general properties of a class of Markov chains with perturbed rates to absorption, those properties being interesting on their own and useful for, among others, certain problems of model repair.

Structure of the paper: Sec. 2 introduces the necessary terminology and notation, Sec. 3 derives the main monotonicity result for the perturbed Markov chains, and Sec. 4 is devoted to the analysis of the asymptotic limit of the hitting probabilities as the perturbation factor goes to zero. In Sec. 5, an algorithm for the model repair problem is presented and its correctness is proven with the help of the results from Sec. 3 and 4. In order to illustrate the algorithm, that section also contains an application example from the area of software version release management. Conclusions and future work are discussed in Sec. 6.

2. Preliminaries

Positivity. For a matrix $A = (A_{ij}) \in \mathbb{R}^{n\times m}$ write $A \geq 0$ if $A_{ij} \geq 0$ for all i, j and $A \gg 0$ if $A_{ij} > 0$ for all i, j. A matrix $A \in \mathbb{R}^{n\times n}$ is a (i) *Z-matrix* if $A_{ij} \leq 0$ for $i \neq j$, (ii) an *L-matrix* if it is a Z-matrix with positive diagonal ($A_{ii} > 0$ for all i) and (iii) an *M-matrix* if $A = sI - B$ for some $B \geq 0$ and $s \geq \rho(B)$ where $\rho(B)$ is the spectral radius of B. An M-matrix A is nonsingular if and only if $A = sI - B$ for some $B \geq 0$ and $s > \rho(B)$ or equivalently if A is a Z-matrix and every eigenvalue of A has (strictly) positive real part [22, Chapter 6]. A nonsingular M-matrix A is an L-matrix and it holds $A^{-1} \geq 0$. If A is an irreducible nonsingular M-matrix then $A^{-1} \gg 0$ [22, Theorem 6.2.7].

For a square matrix $A \in \mathbb{R}^{n\times n}$ its i-th row is *weakly diagonally dominant* if $|A_{ii}| \geq \sum_{j\neq i} |A_{ij}|$ and *strongly diagonally dominant* if $|A_{ii}| > \sum_{j\neq i} |A_{ij}|$. The matrix A is called weakly (resp. strongly) diagonally dominant if all its rows are weakly (resp. strongly) diagonally dominant. A matrix $A \in \mathbb{R}^{n\times n}$ is *weakly chained diagonally dominant* if A is weakly diagonally dominant and for each row i there is a path $i =: i_0 \to i_1 \to \cdots \to i_l$ in the directed graph

of $|A| = (|A_{ij}|)_{ij}$ (meaning that $|A_{i_k,i_{k+1}}| > 0$ for all $k = 0, \dots, l-1$) such that the i_l-th row of A is strongly diagonally dominant. Any weakly chained diagonally dominant matrix is nonsingular. For us of interest is the following important theorem: a matrix A is a weakly chained diagonally dominant L-matrix if and only if A is a nonsingular weakly diagonally dominant M-matrix [23, Theorem 2.24].

Drazin inverse. For $A \in \mathbb{C}^{n\times n}$ denote by $\mathcal{N}(A)$ and $\mathcal{R}(A)$ the nullspace and range of A, by $\operatorname{ind}(A) := \min\{k \in \mathbb{N} \mid \mathcal{N}(A^k) = \mathcal{N}(A^{k+1})\} < \infty$ the *index* of A and by $A^D \in \mathbb{C}^{n\times n}$ the Drazin inverse of A, i.e. the unique matrix satisfying $A^{\nu+1}A^D = A^\nu$ for $\nu = \operatorname{ind}(A)$, $A^DAA^D = A^D$ and $AA^D = A^DA$. If A is invertible (i.e. $\operatorname{ind}(A) = 0$) then $A^D = A^{-1}$ and if A is nilpotent then $A^D = 0$. If $\operatorname{ind}(A) \le 1$ then it also holds that $AA^DA = A$ (since in this case A is group invertible and the group inverse coincides with A^D). For any matrix $A \in \mathbb{C}^{n\times n}$ we can decompose $\mathbb{C}^n = \mathcal{N}(A^\nu) \oplus \mathcal{R}(A^\nu)$ where $\nu := \operatorname{ind}(A)$ and the matrices AA^D and $I - AA^D$ are the corresponding (spectral) projections to $\mathcal{R}(A^\nu)$ along $\mathcal{N}(A^\nu)$ resp. vice versa. For any projection P it holds $P^D = P$. If $A, B \in \mathbb{C}^{n\times n}$ commute then $(AB)^D = B^DA^D$. For a matrix $A \in \mathbb{C}^{n\times n}$ we also denote the spectral projection as $A^\pi := I - AA^D$ [2].

Stability. A matrix $A \in \mathbb{C}^{n\times n}$ is *stable* if every eigenvalue of A has strictly negative real part or equivalently if e^{At} converges to 0 as $t \to \infty$. Every stable matrix A is nonsingular and $\int_0^\infty e^{At}\,dt = A^{-1}$. A matrix $A \in \mathbb{C}^{n\times n}$ is *semistable* if $\operatorname{ind}(A) \le 1$ (i.e. A is group invertible) and the non-zero eigenvalues of A have strictly negative real part [26, p. 87]. Equivalently, A is semistable if and only if e^{At} converges as $t \to \infty$ and in this case the limit is given by $\lim_{t\to\infty} e^{At} = I - AA^D = A^\pi$.

Markov chains. Let $\mathbf{1} \in \mathbb{R}^n$ denote the column vector with values $\mathbf{1}_i = 1$. A matrix $P \in \mathbb{R}^{n\times n}$ is *stochastic* if $P \ge 0$ and $P\mathbf{1} = \mathbf{1}$. A matrix $Q \in \mathbb{R}^{n\times n}$ is a *generator* if $Q_{ij} \ge 0$ for all $i \ne j$ and $Q\mathbf{1} = 0$. There are several ways to

[2] For any matrix A the matrix A^π is the unique matrix P satisfying $AP = PA$, $P^2 = P$, AP is nilpotent and $A + P$ is invertible. A is group invertible if and only if $AP = 0$. The Drazin inverse A^D and the spectral projection A^π are closely related to each other: any of the matrices A, A^D and A^π can be computed from the other two. See [24] and [25] for more information and further references.

convert a stochastic matrix into a generator and vice versa[3]. Any stochastic matrix P defines a discrete-time Markov chain (DTMC) with P as its transition matrix and we sometimes refer to P as the DTMC. Similarly, any generator Q defines a continuous-time Markov chain (CTMC) with Q as the generator matrix of its transition function $P(t) = e^{Qt}$ and we sometimes refer to Q as the CTMC. Every generator Q is semistable and the limiting matrix $\lim_{t\to\infty} e^{Qt} = I - QQ^D = Q^\pi$ is also called the *ergodic projection* of Q. The i-th row of Q^π is the limiting distribution of the corresponding CTMC with initial state i. The convex hull of the rows of Q^π forms a simplex in $\mathbb{R}^{1\times n}$ and its extreme points are the limiting distributions corresponding to the recurrent classes of the CTMC.
Throughout this paper, we are mostly interested in the distributions of CTMCs and DTMCs (or equivalently in properties of generator matrices and stochastic matrices) and less in their trajectory behaviour, so that most of the time we omit the concrete stochastic processes and underlying probability spaces. An exception is made when we talk about the probabilistic interpretation of a certain stochastic matrix.

3. Perturbed Hitting Probabilities and Monotonicity

3.1. Setting

Consider a CTMC with m absorbing states (and no other recurrent classes) and n transient states[4]. Then the generator $Q \in \mathbb{R}^{(n+m)\times(n+m)}$ of the Markov chain can be decomposed as $Q = Q_1 + Q_2$ with

$$Q_1 = \begin{pmatrix} E & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad Q_2 = \begin{pmatrix} D & F \\ 0 & 0 \end{pmatrix} \tag{1}$$

where $F \in \mathbb{R}^{n\times m}$ comprises the rates for transitions from transient states to absorbing states and $D := -\operatorname{diag}(F\mathbf{1}) \in \mathbb{R}^{n\times n}$ is the diagonal matrix

[3]If Q is a generator and D a non-singular matrix such that $P := DQ + I \geq 0$ then P is stochastic. The uniformization and the embedding of a generator correspond to choosing D as a suitable diagonal matrix.

[4] Note that for the purpose of analysis and computation of hitting probabilities for an arbitrary CTMC it is enough to restrict one's consideration to the class of absorbing CTMCs. For applications, where one is given a set of goal states and a disjoint set of taboo states, $m = 2$ absorbing states are sufficient: one can lump the goal states into one absorbing state and the taboo states including the (remaining) recurrent classes into a second absorbing state.

comprising the negative row sums of F on its diagonal. The matrix $E \in \mathbb{R}^{n\times n}$ contains as its off diagonal entries the rates for transitions between the transient states and E has row sums 0. In particular, E is a generator matrix of some CTMC, namely the CTMC that arises from Q by removing the absorbing states (including the transitions leading into them).

Consider the ergodic projection $Q^\pi = \lim_{t\to\infty} e^{Qt} \in \mathbb{R}^{(n+m)\times(n+m)}$ of Q and the canonical decomposition $Q^\pi = RL$ into a matrix $R \in \mathbb{R}^{(n+m)\times m}$ which contains the hitting probabilities into the recurrent classes of Q and $L \in \mathbb{R}^{m\times(n+m)}$ which contains the stationary distributions of Q. Since we supposed that Q is absorbing and has m absorbing states, it follows that Q has m recurrent classes, each consisting of a single state. Therefore L and R are of the form

$$L = \begin{pmatrix} 0 & I \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} \widetilde{R} \\ I \end{pmatrix} \quad \text{so that} \quad Q^\pi = \begin{pmatrix} 0 & \widetilde{R} \\ 0 & I \end{pmatrix} \tag{2}$$

where $I \in \mathbb{R}^{m\times m}$ is the identity matrix, 0 a zero matrix (of appropriate size) and $\widetilde{R} \in \mathbb{R}^{n\times m}$ comprises the hitting probabilities from each of the n transient states into each of the m recurrent classes.

3.2. Hitting Probabilities and the associated exitpoint DTMC

In this section, we first establish an explicit expression for the hitting probabilities $\widetilde{R}$ of an absorbing CTMC with generator Q. We then associate to the CTMC Q a particular DTMC $\widehat{D}$ (the exitpoint DTMC) and prove some relationships between properties of $\widehat{D}$ and Q. This is slightly technical but it will be helpful in deducing the limiting probabilities in Sec. 4. Let us first show the following

Proposition 3.1. The matrix $-(E + D)$ is a nonsingular M-matrix.

Proof. In order to prove that $A := -(E + D)$ is a nonsingular M-matrix it is enough to show that A is a weakly chained diagonally dominant L-matrix. First, observe that since $-E$ is a weakly diagonally dominant Z-matrix (because E is a generator) it follows that A is also a weakly diagonally dominant Z-matrix (because D is diagonal and $D \leq 0$). Moreover, A is an L-matrix since its diagonal entries $A_{ii} = -Q_{ii}$, $i = 1, \ldots, n$ are positive because $Q_{ii} \leq 0$ and $Q_{ii} \neq 0$ because i is not an absorbing state for Q. Now let us also show that A is weakly chained diagonally dominant. For this, let $i \in \{1, \ldots, n\}$. Since Q is absorbing, there is a path $i =: i_0 \to i_1 \to$

$\cdots \to i_l \to i_{l+1}$ ($l \geq 0$) in the directed graph of Q (i.e. $Q_{i_k,i_{k+1}} > 0$ for all $k = 0, \ldots, l$) with $i_{l+1} \in \{n+1, \ldots, n+m\}$ and $i_k \in \{1, \ldots, n\}$ for all $k = 0, \ldots, l$. Then $i = i_0 \to i_1 \to \cdots \to i_l$ is a path in the directed graph of $-A$ because $-A_{i_k,i_{k+1}} = Q_{i_k,i_{k+1}} > 0$ for $k = 0, \ldots, l-1$. Since we also have $Q_{i_l,i_{l+1}} > 0$ it follows from $-Q_{i_l,i_l} = \sum_{j \in \{1,\ldots,n+m\} \setminus \{i_l,i_{l+1}\}} Q_{i_l,j} + Q_{i_l,i_{l+1}}$ that

$$|A_{i_l,i_l}| = -Q_{i_l,i_l} > \sum_{j \in \{1,\ldots,n\} \setminus \{i_l\}} Q_{i_l,j} = \sum_{j \in \{1,\ldots,n\} \setminus \{i_l\}} |A_{i_l,j}|.$$

In other words, the i_l-th row of A is strongly diagonally dominant and we have found a path from i to i_l in the directed graph of A. □

Remark 3.2. An alternative proof of Proposition 3.1 (without relying on properties of weakly chained diagonally dominant matrices) can be given as follows. The communication equivalence relation on Q partitions $\{1, \ldots, n+m\}$ into $k+m$ equivalence classes $S_1, \ldots, S_k, \{1\}, \ldots, \{m\}$ with $0 \leq k \leq n$. Observe that the transient classes $S_1, \ldots, S_k$ of Q are precisely the communication classes of E. Writing Q in canonical form as a corresponding block (upper) triangular matrix, the submatrix $E + D$ also has block triangular form and each of its diagonal blocks forms an irreducible subgenerator[5]: if S_i is a recurrent class of E then there is a transition in Q from S_i to some absorbing state (so that D contains at least one negative entry in the diagonal block of E corresponding to S_i) whereas if S_i is a transient class of E then there is a transition in E from S_i to some other class S_j. It follows by [13, Corollary 3.3] that each diagonal block of $-(E + D)$ is a nonsingular M-matrix. Therefore, $-(E + D)$ is a nonsingular M-matrix, because each eigenvalue of $-(E + D)$ is an eigenvalue of some of its diagonal blocks and has therefore positive real part.

Proposition 3.3. The hitting probabilities are given by

$$\widetilde{R} = -(E + D)^{-1} F.$$

Proof. Recall that since Q is semistable its ergodic projection Q^π is given by $Q^\pi = I - QQ^D$. In the following, we are going to compute Q^D. Instead of the decomposition $Q = Q_1 + Q_2$ as in Eq. (1) consider the decomposition

[5] A square matrix A is a subgenerator if $A_{ij} \geq 0$ for $i \neq j$ and $\sum_j A_{ij} \leq 0$ for all i.

$Q = B + N$ where

$$B := \begin{pmatrix} E+D & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad N := \begin{pmatrix} 0 & F \\ 0 & 0 \end{pmatrix}.$$

Note that $NB = 0$ and N is nilpotent of index 2 (since $N \neq 0$ and $N^2 = 0$). With this decomposition we can apply [27, Corollary 2.3] (or [28, Corollary 2.1 (iv)]) which results in

$$Q^D = (B+N)^D = B^D + (B^D)^2 N.$$

It follows that

$$\begin{aligned} Q^\pi &= I - QQ^D = I - (BB^D + NB^D + B(B^D)^2 N + N(B^D)^2 N) \\ &= I - BB^D(I + B^D N) \end{aligned}$$

where we have applied that $NB^D = NB^D BB^D = NB(B^D)^2 = 0$ since $NB = 0$. Since $E + D$ is invertible by Proposition 3.1 we have

$$B^D = \begin{pmatrix} (E+D)^{-1} & 0 \\ 0 & 0 \end{pmatrix}.$$

It follows

$$\begin{aligned} Q^\pi &= \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \left(\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} + \begin{pmatrix} 0 & (E+D)^{-1}F \\ 0 & 0 \end{pmatrix} \right) \\ &= \begin{pmatrix} 0 & -(E+D)^{-1}F \\ 0 & I \end{pmatrix}. \end{aligned}$$

Thus when comparing with Q^π as in Eq. (2) we get $\widetilde{R} = -(E+D)^{-1}F$. □

Similarly to the matrix $(E + D)^{-1}F$ from Proposition 3.3 let us also consider the matrices

$$\widehat{D} := (E+D)^{-1}D \quad \text{and} \quad \widehat{E} := (E+D)^{-1}E. \tag{3}$$

These two matrices play a key role in the following sections. In Remark 3.5 and Proposition 3.8 we describe in which way these matrices are related to the CTMC with generator Q. But let us first state the following

Lemma 3.4. The matrix $\widehat{D}$ is stochastic and $-\widehat{E} = \widehat{D} - I$ is a generator.

Proof. The identity $\widehat{E} + \widehat{D} = I$ is clear. Since $-(E + D)$ is a nonsingular M-matrix by Proposition 3.1, it is in particular inverse-positive, meaning that $-(E+D)^{-1} \geq 0$ and from $D \leq 0$ it follows $\widehat{D} \geq 0$. From $E\mathbf{1} = 0$ we get $(E + D)\mathbf{1} = D\mathbf{1}$ and thus $\widehat{D}\mathbf{1} = (E + D)^{-1}D\mathbf{1} = \mathbf{1}$ from which we conclude that $\widehat{D}$ is stochastic (and therefore $-\widehat{E} = \widehat{D} - I$ a generator). □

Remark 3.5. The matrix $-(E + D)^{-1}$ has the following probabilistic interpretation [7, Eq. (2.2)]. Let $(X_t)_{t\geq 0}$ be a CTMC (as a stochastic process) with generator Q and state space $\{1, \ldots, n+m\}$ and for every $i = 1, \ldots, n+m$ denote by $\mathbb{P}_i$ and $\mathbb{E}_i$ the probability measure resp. the expectation with i as initial state. For a transient state j denote by T_j the total sojourn time of $(X_t)_{t\geq 0}$ in state j (until absorption) so that $T_j = \int_0^\infty \mathbb{1}_{\{X_t=j\}}\, dt$ where $\mathbb{1}_A$ denotes the indicator of an event A. Then for every transient state i it holds

$$\mathbb{E}_i(T_j) = \int_0^\infty \mathbb{E}_i(\mathbb{1}_{\{X_t=j\}})\, dt = \int_0^\infty \mathbb{P}_i(X_t = j)\, dt = \int_0^\infty P_{ij}(t)\, dt$$

where $P(t) = e^{Qt}$ is the transition function of $(X_t)_{t\geq 0}$. From the block decomposition Eq. (1) and $\int_0^\infty e^{(E+D)t}\, dt = -(E + D)^{-1}$ (the matrix $E + D$ is stable[6]) it follows for all $i, j = 1, \ldots, n$

$$\left(-(E + D)^{-1}\right)_{ij} = \mathbb{E}_i(T_j). \tag{4}$$

In particular, the hitting probabilities $\widetilde{R}_{ik}$ from i into an absorbing state k can be represented by $\widetilde{R}_{ik} = \sum_j \mathbb{E}_i(T_j)F_{jk}$. Eq. (4) also provides a probabilistic interpretation of the matrix $\widehat{D}$: each component $\widehat{D}_{ij} = \mathbb{E}_i(T_j)(-D_{jj})$ is the expected total sojourn time in state j weighted by the rate $-D_{jj} = \sum_k F_{jk}$ to absorption from j (see Figure 1). In particular, in $\widehat{D}$ there is a transition from i to j (i.e. $\widehat{D}_{ij} > 0$) if and only if j is reachable from i in Q (i.e. $\mathbb{E}_i(T_j) > 0$) and there is a direct transition from j to some absorbing state (i.e. $-D_{jj} > 0$). In probabilistic terms, a transition from i to j in $\widehat{D}$ represents the event B_j that the CTMC $(X_t)_{t\geq 0}$ with initial state i gets absorbed from state j. In other words, $\widehat{D}_{ij}$ is the probability of the set of paths from i to some absorbing state where the absorption takes place from j. Indeed, since B_j represents the event to find the system in state j at some point in time t

[6]Since $-(E + D)$ is a nonsingular M-matrix its eigenvalues have positive real parts.

and from there to get absorbed with rate $\sum_{k=1}^{m} F_{jk} = -D_{jj}$ we get

$$P_i(B_j) = \int_0^\infty P_{ij}(t)(-D_{jj})\,dt = \mathbb{E}_i(T_j)(-D_{jj}) = \widehat{D}_{ij}.$$

Remark 3.6. Proposition 3.3 as well as the identities $(-(E+D)^{-1})_{ij} = \mathbb{E}_i(T_j)$ and $\widehat{D}_{ij} = \mathbb{P}_i(B_j)$ in the previous remark can also be proven by means of a uniformization of the CTMC $(X_t)_{t\geq 0}$ to a DTMC with transition probability matrix $\frac{1}{\mu}Q + I = \begin{pmatrix} U & V \\ 0 & I \end{pmatrix}$ with $U := \frac{1}{\mu}(E+D) + I$ and $V := \frac{1}{\mu}F$ for some $\mu \in (0,\infty)$ large enough (so that $P \geq 0$) [7]. To deduce the conclusion of Proposition 3.3 one can appeal to the well known facts that (i) Q and P have the same absorbing states, (ii) the absorption probabilities are the same and (iii) that they can be computed by [29, Theorem 3.3.7] as $\widetilde{R} = (I-U)^{-1}V = -(E+D)^{-1}F$. In our proof of Proposition 3.3, we follow a different strategy to directly compute the Drazin inverse Q^D and the spectral projection $Q^\pi = I - QQ^D$. This proof is of rather algebraic nature and it can also be applied in a similar way in more general setups, such as for certain non-uniformizable absorbing Markov chains on a countably infinite state space or even certain Markov processes on more general state spaces [30].

Definition 3.7. For a generator Q we refer to the DTMC with transition matrix $\widehat{D}$ as the *exitpoint DTMC* associated with Q.

In [13, Lemma 3.4] it was shown that if E is irreducible (i.e. Q has a single transient class) then $\widehat{D}$ has a single recurrent class, namely $\{j \mid D_{jj} < 0\}$ (but $\widehat{D}$ need not be irreducible). The following proposition generalizes this result to our setup (i.e. for a general transient class structure of Q) and hereby also provides additional information on the state classification of $\widehat{D}$. One can write down a proof in matrix-theoretic terms similar to the proof of the result in [13] by decomposing $\widehat{D} = (E+D)^{-1}D$ into block triangular form. But this may get quite cumbersome and possibly unclear. In contrast, we decided to use the information given in Remark 3.5 since this allows us to provide a more illustrative proof in purely graph-theoretic terms[8].

[7] This was suggested to us by an anonymous reviewer.

[8] In this graph-theoretic proof we employ probabilistic terminology for the state clas-

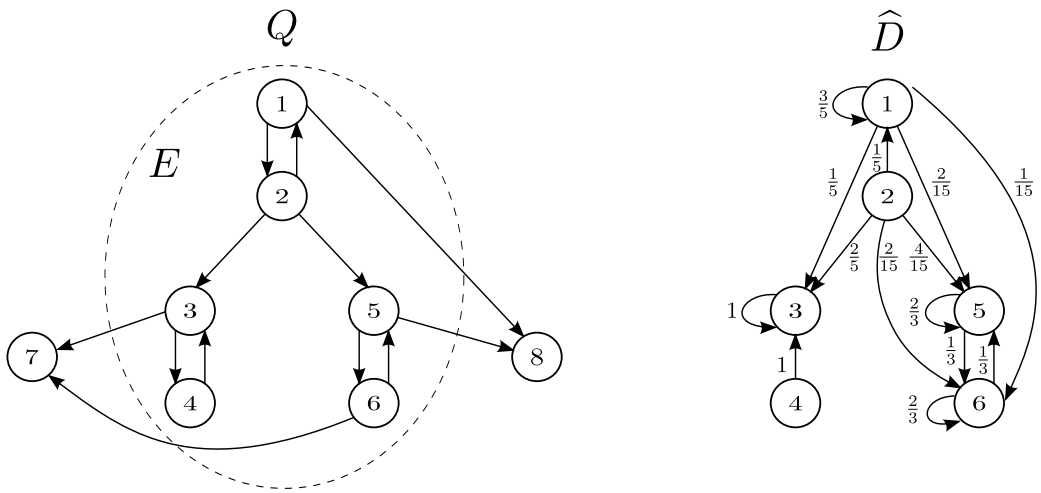

Figure 1: Left: A CTMC with generator Q and two absorbing states. The transient states of Q form a CTMC on their own with generator E. All the rates for transitions are assumed to be identical, e.g. 2.0. Right: The exitpoint DTMC with transition matrix $\widehat{D} = (E + D)^{-1}D$.

Proposition 3.8. The exitpoint DTMC with transition matrix $\widehat{D}$ and the CTMC with generator $-\widehat{E} = \widehat{D} - I$ have the same state classification. Furthermore, it holds that:

(i) A state j is recurrent for $\widehat{D}$ if and only if j is recurrent for E and there is a transition in Q from j to some absorbing state ($D_{jj} < 0$). Equivalently, j is transient for $\widehat{D}$ if and only if j is transient for E or if there is no transition in Q from j to some absorbing state.

(ii) Each communication class of $\widehat{D}$ is contained in a communication class of E. In other words, the communication classes of E split into communication classes of $\widehat{D}$. In more detail,

- each transient class of E splits into transient classes of $\widehat{D}$ and
- each recurrent class of E splits into a single recurrent class of $\widehat{D}$ and transient classes of $\widehat{D}$ each consisting of a single state.

sification (transient and recurrent states) instead of corresponding graph-theoretic terminology ((bottom) strongly connected components).

Proof. Note that the DTMC that arises as the uniformization of the CTMC with generator $-\widehat{E}$ (with uniformization rate 1) coincides with the DTMC with transition matrix $\widehat{D}$. Therefore, both processes give rise to the same state classification.

(i) "$\Leftarrow$": Let j be a state that is recurrent for E and such that there is a transition in Q from j to some absorbing state. Assume that j is transient for $\widehat{D}$. Then there is state k recurrent for $\widehat{D}$ and a path from j to k in $\widehat{D}$. This implies that there is a path from j to k in E. Since j is recurrent for E it follows that k and j belong to the same recurrence class in E. Since there is a path from k to j in Q and a transition in Q from j to some absorbing state it follows that there is a transition from k to j in $\widehat{D}$. Therefore, k and j belong to the same communication class for $\widehat{D}$ and therefore j is recurrent for $\widehat{D}$, contradiction.

"$\Rightarrow$": Let j be recurrent for $\widehat{D}$. There is a state k_0 that is recurrent for E and reachable from j in E. Then there is a state k in the recurrence class of k_0 for E from which there is a transition in Q to some absorbing state. In particular, there is a path from j to k in Q. Hence, there is a transition from j to k in $\widehat{D}$. Since j is recurrent in $\widehat{D}$ it follows that k is also recurrent in $\widehat{D}$ and both states belong to the same communication class in $\widehat{D}$. Therefore, there is also a path from k to j in $\widehat{D}$ which implies that there is a path from k to j in Q and a transition from j in Q to some absorbing state. Now, since k and j communicate in Q they also communicate in E and since k is recurrent for E it follows that j is recurrent for E.

(ii) If two states i and j are communicating in $\widehat{D}$, then j is reachable from i in E and i is reachable from j in E, hence i and j are communicating in E. Therefore, each communication class of $\widehat{D}$ is contained in a communication class of E. By (i), every recurrent state for $\widehat{D}$ is also recurrent for E. Therefore, every transient class for E does not contain any of the recurrent classes for $\widehat{D}$ and hence splits into transient classes for $\widehat{D}$.

Assume that there is a recurrent class R for E that contains two disjoint recurrent classes $\widehat{R}_1$ and $\widehat{R}_2$ for $\widehat{D}$. There are states $j_1 \in \widehat{R}_1$ and $j_2 \in \widehat{R}_2$ from which there is a transition in Q to some absorbing state. Since $j_1, j_2 \in R$, they communicate in E and hence in Q. This implies that they also communicate in $\widehat{D}$, contradiction. Therefore, each recurrent class of E contains exactly one recurrent class of $\widehat{D}$. Now let us also show that each transient class $\widehat{T}$ of $\widehat{D}$ that is contained in a recurrent class R of E consists only of a single state. Let $i \in \widehat{T}$ and assume that there is $i' \in \widehat{T}$ with $i \neq i'$. Since i

and i' communicate in $\widehat{D}$, there is a path from i' to i in $\widehat{D}$. But this implies that there is a transition in Q from i to some absorbing state. Let $\widehat{R}$ denote the (unique) recurrent class for $\widehat{D}$ that is contained in R and let $j \in \widehat{R}$. Since R is a recurrent class of E and $j, i \in R$, there is a path in E from j to i. Now since there is a transition in Q from i to some absorbing state it follows that there is a transition in $\widehat{D}$ from j to i. Hence, i is recurrent for $\widehat{D}$, contradiction. $\square$

By Proposition 3.8, we can order the states in such a way that the canonical block decomposition of E provides also a block decomposition of $\widehat{E}$ such that both matrices are of the form

$$E = \begin{pmatrix} E_{00} & E_{01} & \dots & E_{0r} \\ & E_1 & & \\ & & \ddots & \\ & & & E_r \end{pmatrix} \quad \text{and} \quad \widehat{E} = \begin{pmatrix} \widehat{E}_{00} & \widehat{E}_{01} & \dots & \widehat{E}_{0r} \\ & \widehat{E}_1 & & \\ & & \ddots & \\ & & & \widehat{E}_r \end{pmatrix}$$

where r is the number of recurrent classes of E (and of $\widehat{E}$). Let T denote the set of transient states of E and R_k the k-th recurrent class of E for $k = 1, \dots, r$. The matrix E_k, $k = 1, \dots r$ comprises the rates for transitions between states in R_k (hence E_k is an irreducible generator), whereas E_{0k} comprises the rates for transitions from states in T to T ($k = 0$) or to R_k ($k = 1, \dots, r$). Let $\widehat{T}$ denote the set of transient states of $-\widehat{E}$ and $\widehat{R}_k$ the k-th recurrent class of $-\widehat{E}$, $k = 1, \dots, r$. Recall that R_k can be decomposed into $R_k = \widehat{R}_k \cup \widehat{T}_k$ where $\widehat{T}_k := \widehat{T} \cap R_k$ comprises those states that are transient for $-\widehat{E}$ and contained in R_k. (Note that $\widehat{T} = T \cup \bigcup_{i=1}^{r} \widehat{T}_k$.) Hence, the states in R_k can be further reordered such that the matrix $\widehat{E}_k$ can be written in the block form

$$\widehat{E}_k = \begin{pmatrix} 0 & \widehat{E}_k^* \\ 0 & \widehat{E}_k' \end{pmatrix} \tag{5}$$

where $-\widehat{E}_k'$ is the irreducible generator that comprises the rates for transitions between states in $\widehat{R}_k$ and $-\widehat{E}_k^*$ comprises the rates for transitions from $\widehat{T}_k$ to $\widehat{R}_k$. The upper left block of $\widehat{E}_k$ is 0 because there are no transitions of $-\widehat{E}$ between states in $\widehat{T}_k$.

These joint block representations of E and $\widehat{E}$ allow us also to relate the stationary behaviour of the DTMC with transition matrix $\widehat{D}$ (or equivalently

of the CTMC with generator $-\widehat{E}$) to that of the CTMC with generator E. Consider the ergodic projections $E^\pi = I - EE^D$ of E and $\widehat{E}^\pi = I - \widehat{E}\widehat{E}^D$ of $-\widehat{E}$ and for better readability in the following denote these by $\Pi := E^\pi$ and $\widehat{\Pi} := \widehat{E}^\pi$. Then with respect to the above given ordering of the states we can jointly decompose Π and $\widehat{\Pi}$ into

$$\Pi = \begin{pmatrix} 0 & \Pi_{01} & \dots & \Pi_{0r} \\ & \Pi_1 & & \\ & & \ddots & \\ & & & \Pi_r \end{pmatrix} \quad \text{and} \quad \widehat{\Pi} = \begin{pmatrix} 0 & \widehat{\Pi}_{01} & \dots & \widehat{\Pi}_{0r} \\ & \widehat{\Pi}_1 & & \\ & & \ddots & \\ & & & \widehat{\Pi}_r \end{pmatrix}$$

Since E_i is irreducible ($i = 1, \dots, r$), it has a unique stationary distribution $\pi_i \gg 0$ so that $\Pi_i = \mathbf{1}\pi_i$ has equal rows π_i. Similarly, since $-\widehat{E}_i$ has only one recurrent class (it need not be irreducible), there is a unique stationary distribution $\widehat{\pi}_i \geq 0$ for $-\widehat{E}_i$ of the form $\widehat{\pi}_i = \begin{pmatrix} 0 & \widehat{\pi}_i' \end{pmatrix}$ where $\widehat{\pi}_i' \gg 0$ so that $\widehat{\Pi}_i = \mathbf{1}\widehat{\pi}_i = \begin{pmatrix} 0 & \mathbf{1}\widehat{\pi}_i' \end{pmatrix}$. Let us also decompose $D = \operatorname{diag}(D_0, D_1, \dots, D_r)$ into corresponding diagonal blocks D_k (each D_k is diagonal).

Example: For the generators E and $-\widehat{E} = \widehat{D} - I$ from Figure 1, their block matrix representations are immediately observable from the figure. Their ergodic projections $\Pi = E^\pi = I - EE^D$ and $\widehat{\Pi} = (-\widehat{E})^\pi = \widehat{E}^\pi = I - \widehat{E}\widehat{E}^D$ (the latter being the same as the limiting matrix of the exitpoint DTMC $\widehat{D}$) are given by

$$\Pi = \begin{pmatrix} 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \quad \text{and} \quad \widehat{\Pi} = \begin{pmatrix} 0 & 0 & \frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Proposition 3.9. It holds

$$\widehat{\Pi} = \Pi W$$

where $W := \operatorname{diag}(0, W_1, \dots, W_r)$ and $W_k := \frac{1}{\|\pi_k D_k\|_1}(-D_k)$. In other words, since W is diagonal, $\widehat{\Pi}$ arises from Π by suitably weighting the columns of Π. The identity $\widehat{\Pi} = \Pi W$ reads blockwise as

$$\widehat{\Pi}_k = \Pi_k W_k \quad \text{and} \quad \widehat{\Pi}_{0k} = \Pi_{0k} W_k.$$

Proof. Since Π and $\widehat{\Pi}$ are the ergodic projections of E resp. of $-\widehat{E}$ it follows $\Pi E = 0$ and $\widehat{\Pi}\widehat{E} = 0$. From $\Pi E = 0$ we can also deduce that $\Pi D\widehat{E} = 0$ since

$$\Pi D\widehat{E} = \Pi D(E+D)^{-1}E = \Pi(E+D)(E+D)^{-1}E = \Pi E = 0.$$

Let us compare the matrices $\widehat{\Pi}$ and ΠD. Consider first the diagonal blocks $\widehat{\Pi}_k$ of $\widehat{\Pi}$ for $k = 1, \ldots, r$. From $\widehat{\Pi}\widehat{E} = 0$ and $\Pi D\widehat{E} = 0$ it follows $\widehat{\Pi}_k\widehat{E}_k = 0$ and $\Pi_k D_k \widehat{E}_k = 0$. From $\Pi_k = \mathbf{1}\pi_k$ and $\widehat{\Pi}_k = \mathbf{1}\widehat{\pi}_k$ we deduce that $\pi_k D_k \widehat{E}_k = 0$ and $\widehat{\pi}_k\widehat{E}_k = 0$. Consider the decomposition of $\widehat{E}_k$ as in Eq. (5) and correspondingly decompose $\widehat{\pi}_k = \begin{pmatrix} 0 & \widehat{\pi}'_k \end{pmatrix}$ where $\widehat{\pi}'_k \gg 0$ and $D_k = \operatorname{diag}(D^*_k, D'_k)$. Then $D^*_k = 0$ because from such transient states for $\widehat{D}$ there is no transition in Q to an absorbing state. From $D_k < 0$ it follows that $D'_k < 0$. Since $-\widehat{E}'_k$ is irreducible, the Perron-Frobenius theorem tells us that every positive solution $x > 0$ of $x(-\widehat{E}'_k) = 0$ is a positive multiple of $\widehat{\pi}'_k$. From $\pi_k D_k \widehat{E}_k = 0$ it follows that $\pi'_k D'_k \widehat{E}'_k = 0$ and since $\pi'_k(-D'_k) > 0$ (because $\widehat{\pi}'_k \gg 0$ and $-D'_k > 0$) we deduce that

$$\pi_k(-D_k) = \begin{pmatrix} 0 & \pi'_k(-D'_k) \end{pmatrix} = \begin{pmatrix} 0 & c_k\widehat{\pi}'_k \end{pmatrix} = c_k\widehat{\pi}_k$$

for some $c_k \in \mathbb{R}$, $c_k > 0$. Taking 1-norms on both sides we deduce that $\|\pi_k D_k\|_1 = c_k\|\widehat{\pi}_k\|_1 = c_k$ and therefore $\widehat{\pi}_k = \frac{1}{\|\pi_k D_k\|_1}\pi_k(-D_k)$. Setting $W_k := \frac{1}{\|\pi_k D_k\|_1}(-D_k)$ we observe that $\widehat{\pi}_k = \pi_k W_k$ and hence $\widehat{\Pi}_k = \Pi_k W_k$.
Now let us consider the off-diagonal blocks $\widehat{\Pi}_{0k}$ of $\widehat{\Pi}$ for $k = 1, \ldots, r$. Let us show that $\widehat{\Pi}_{0k}\mathbf{1} = \Pi_{0k}\mathbf{1}$ by a probabilistic argument. Denote by T the set of transient states of E and for $i \in T$ set $p_{ik} := (\Pi_{0k}\mathbf{1})_i$ and $\widehat{p}_{ik} := (\widehat{\Pi}_{0k}\mathbf{1})_i$. The value p_{ik} is the probability that the CTMC with generator E and initial state i eventually hits the k-th recurrent class R_k of E. This value p_{ik} coincides with the probability that the CTMC with generator Q and initial state i hits the set R_k (which is a transient class for Q) without getting into any of the absorbing states of Q before hitting R_k (almost surely). By Remark 3.5 the value $\widehat{p}_{ik}$ is the probability that the DTMC with transition matrix $\widehat{D}$ performs a single transition from i to some state in R_k. Moreover, since a transition from i to some state j of $\widehat{D}$ represents the set of paths in Q of the form $i \to \cdots \to j \to a$ where a is some absorbing state of Q in the sense that both events have the same probability, it follows that $\widehat{p}_{ik}$ is the probability that Q hits R_k from i (without getting absorbed before hitting R_k). In other words, we have shown that $\widehat{p}_{ik} = p_{ik}$ as required. Now from $\Pi^2 = \Pi$ and

$\widehat{\Pi}^2 = \widehat{\Pi}$ together with $\widehat{\Pi}_{0k}\mathbf{1} = \Pi_{0k}\mathbf{1}$ and $\widehat{\pi}_k = \pi_k W_k$ it follows that

$$\widehat{\Pi}_{0k} = \widehat{\Pi}_{0k}\widehat{\Pi}_k = \widehat{\Pi}_{0k}\left(\mathbf{1}\widehat{\pi}_k\right) = \left(\Pi_{0k}\mathbf{1}\right)\left(\pi_k W_k\right) = \Pi_{0k}\Pi_k W_k = \Pi_{0k} W_k$$

for each $k = 1, \ldots, r$. Setting $W := \operatorname{diag}(W_0, W_1, \ldots, W_r)$ (where W_0 is an arbitrary $|T| \times |T|$-matrix, e.g. $W_0 := 0$) it follows that $\widehat{\Pi} = \Pi W$. $\square$

3.3. Perturbation of rates for transitions to absorbing states

Let us now scale the rates of the transitions of Q to absorption with a small factor $\varepsilon > 0$, i.e. consider the family of generators $Q^\varepsilon := Q_1 + \varepsilon Q_2$ with Q_1 and Q_2 as in Sec. 3.1. The generator εQ_2 can be regarded as an additive perturbation to the generator Q_1 and Q^ε as a generator of some perturbed Markov chain. Note that the number of recurrent classes of Q^ε is m for $\varepsilon > 0$ (i.e. the m absorbing states) and it is strictly larger than m for $\varepsilon = 0$ (the m absorbing states plus the recurrent classes of E (which are transient for $\varepsilon > 0$)). When we substitute Q_2 by εQ_2 for $\varepsilon > 0$ then by Proposition 3.3 we get that the hitting probabilities $\widetilde{R}(\varepsilon)$ are also perturbed by ε and given by

$$\widetilde{R}(\varepsilon) = -(E + \varepsilon D)^{-1}\varepsilon F.$$

Note that for any $\varepsilon > 0$ the matrix $E + \varepsilon D$ is nonsingular while for $\varepsilon = 0$ it is singular. The matrix function $(E + \varepsilon D)^{-1}$ is the restriction of the *generalized resolvent* $(E + \lambda D)^{-1}$ (which is defined for all those $\lambda \in \mathbb{C}$ for which $E + \lambda D$ is nonsingular) to the positive real line $(0, \infty)$.

Figure 2 shows an example of such a perturbed CTMC and the hitting probabilities $\widetilde{R}(\varepsilon)_{i2}$ from all transient states $i = 1, \ldots, 6$ to the absorbing state 8 (transitions to state 8 correspond to the 8-th column of $Q^\varepsilon \in \mathbb{R}^{8\times 8}$ and to the 2-nd column of $\widetilde{R}(\varepsilon) \in \mathbb{R}^{6\times 2}$, see Eq. (2)). The functions $\widetilde{R}(\varepsilon)_{i2}$ are rational in ε and they were computed and plotted by Wolfram Mathematica. These rational functions can also be computed by a parametric CTMC model in the PRISM model checker [5].

In the following, we are going to analyze the behaviour of the hitting probabilities $\widetilde{R}(\varepsilon)$. While for a fixed transient state i the hitting probability $\widetilde{R}(\varepsilon)_{ik}$ into the absorbing state k is not necessarily monotonic, we prove the monotonicity of their enveloping functions $\max_i \widetilde{R}(\varepsilon)_{ik}$ and $\min_i \widetilde{R}(\varepsilon)_{ik}$ for every absorbing state k. Following Campbell [26], it turns out to be helpful to consider the transformations as in Eq. (3) and define for $\varepsilon > 0$ the matrix

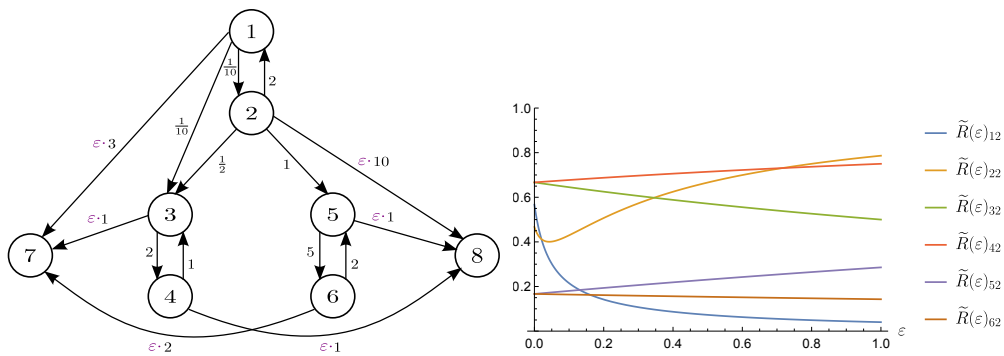

Figure 2: Left: an absorbing Markov chain with rates to absorption scaled by $\varepsilon > 0$. Right: the corresponding hitting probabilities $\widetilde{R}(\varepsilon)_{i2}$ into state 8 (the second column of $\widetilde{R}(\varepsilon) \in \mathbb{R}^{6\times 2}$) for $i = 1, \ldots, 6$. The function $\widetilde{R}(\varepsilon)_{22}$ is non-monotonic, but the enveloping functions $\max_i \widetilde{R}(\varepsilon)_{i2}$ and $\min_i \widetilde{R}(\varepsilon)_{i2}$ are monotonic.

functions [9]

$$\widehat{E}_\varepsilon := (E + \varepsilon D)^{-1} E \quad \text{and} \quad \widehat{D}_\varepsilon := (E + \varepsilon D)^{-1} D.$$

Before establishing the monotonicity of the enveloping functions, we first state some helpful facts and identities involving the matrices $\widehat{E}_\varepsilon$ and $\widehat{D}_\varepsilon$.

Lemma 3.10. (i) For any $\varepsilon > 0$ and $\delta > 0$ the matrix $\widehat{E}_\varepsilon + \delta \widehat{D}_\varepsilon$ is invertible and

$$\widehat{E}_\delta = (\widehat{E}_\varepsilon + \delta \widehat{D}_\varepsilon)^{-1} \widehat{E}_\varepsilon \quad \text{and} \quad (\widehat{E}_\varepsilon + \delta \widehat{D}_\varepsilon)^{-1} = \widehat{E}_\delta + \varepsilon \widehat{D}_\delta.$$

(ii) For $0 < \varepsilon \leq \delta$ it holds

$$\widehat{E}_\varepsilon + \delta \widehat{D}_\varepsilon \geq I \quad \text{and} \quad \widehat{E}_\varepsilon + \varepsilon \widehat{D}_\varepsilon = I.$$

(iii) For $\varepsilon > 0$ and $\delta > 0$ the matrices $\widehat{E}_\varepsilon$, $\widehat{E}_\varepsilon^D$, $\widehat{D}_\varepsilon$, $\widehat{D}_\varepsilon^D$, $\widehat{E}_\delta$, $\widehat{E}_\delta^D$, $\widehat{D}_\delta$ and $\widehat{D}_\delta^D$ commute pairwise.

(iv) For $\varepsilon > 0$ the matrix $\varepsilon \widehat{D}_\varepsilon$ is stochastic and $-\widehat{E}_\varepsilon = \varepsilon \widehat{D}_\varepsilon - I$ is a generator.

[9] It is possibly more natural to define $\widehat{D}_\varepsilon$ by $(E + \varepsilon D)^{-1}(\varepsilon D)$ (i.e. substituting εD for D in Eq. (3)) but we decided to stay with the notation as in [26].

Proof. (i) Since $E+\varepsilon D$ is invertible for any $\varepsilon > 0$ it follows that $\widehat{E}_\varepsilon + \delta\widehat{D}_\varepsilon = (E+\varepsilon D)^{-1}(E+\delta D)$ is also invertible and its inverse is given by

$$(\widehat{E}_\varepsilon + \delta\widehat{D}_\varepsilon)^{-1} = (E+\delta D)^{-1}(E+\varepsilon D) = \widehat{E}_\delta + \varepsilon\widehat{D}_\delta.$$

The other identity follows from

$$\begin{aligned}\widehat{E}_\delta &= (E+\delta D)^{-1}E = (E+\delta D)^{-1}(E+\varepsilon D)(E+\varepsilon D)^{-1}E \\ &= ((E+\varepsilon D)^{-1}(E+\delta D))^{-1}(E+\varepsilon D)^{-1}E = (\widehat{E}_\varepsilon + \delta\widehat{D}_\varepsilon)^{-1}\widehat{E}_\varepsilon.\end{aligned}$$

(ii) The identity $\widehat{E}_\varepsilon + \varepsilon\widehat{D}_\varepsilon = I$ is clear. Since $D \le 0$ and $-(E+\varepsilon D)$ is a nonsingular M-matrix by Proposition 3.1 it follows that $(E+\varepsilon D)^{-1} \le 0$ and hence $\widehat{D}_\varepsilon \ge 0$. Thus, if $\delta \ge \varepsilon$ then $\widehat{E}_\varepsilon + \delta\widehat{D}_\varepsilon \ge \widehat{E}_\varepsilon + \varepsilon\widehat{D}_\varepsilon = I$.
(iii) Since $\widehat{E}_\varepsilon + \varepsilon\widehat{D}_\varepsilon = I$ by (ii) it follows that the four ε-matrices $\widehat{E}_\varepsilon$, $\widehat{E}^D_\varepsilon$, $\widehat{D}_\varepsilon$ and $\widehat{D}^D_\varepsilon$ commute pairwise. By (i), $\widehat{E}_\delta$ is expressible in terms of $\widehat{E}_\varepsilon$ (as a power series in $\widehat{E}_\varepsilon$ for a fixed δ). Therefore, $\widehat{E}_\delta$ commutes with the ε-matrices and it then follows that $\widehat{E}^D_\delta = \widehat{E}^D_\varepsilon(\widehat{E}_\varepsilon + \delta\widehat{D}_\varepsilon)$, $\widehat{D}_\delta$ and $\widehat{D}^D_\delta$ also commute with the ε-matrices.
(iv) This follows from Lemma 3.4. □

Remark 3.11. As in Remark 3.5, the entries in $\varepsilon\widehat{D}_\varepsilon$ for a fixed $\varepsilon > 0$ can be interpreted in terms of the expected total sojourn time in transient states of the CTMC with generator $Q^\varepsilon = Q_1 + \varepsilon Q_2$. In addition, observe that the state classification of $\varepsilon\widehat{D}_\varepsilon$ (and hence of $-\widehat{E}_\varepsilon$) does not depend on $\varepsilon > 0$. In fact, by Proposition 3.8 the state classification of $\varepsilon\widehat{D}_\varepsilon$ depends only on the state classification of Q^ε and the latter does not depend on $\varepsilon > 0$.

We are now ready to establish the monotonicity of the enveloping functions. This will allow us to compare the hitting probabilities $\widetilde{R}(\varepsilon)$ for different perturbation values ε.

Theorem 3.12. Fix an absorbing state $1 \le k \le m$ and consider the perturbed hitting probabilities $\widetilde{R}(\varepsilon)_{ik}$ from all transient states $1 \le i \le n$ to k for $\varepsilon > 0$. Define the functions

$$M_k(\varepsilon) := \max_{i=1,\dots,n} \widetilde{R}(\varepsilon)_{ik} \quad \text{and} \quad m_k(\varepsilon) := \min_{i=1,\dots,n} \widetilde{R}(\varepsilon)_{ik} \quad .$$

Then $M_k(\varepsilon)$ is monotonically increasing and $m_k(\varepsilon)$ is monotonically decreasing.

Proof. Consider the k-th column of $\widetilde{R}$. We show that for all transient states $i \in \{1, \dots, n\}$, for all $\delta > 0$ and for all $0 < \varepsilon \leq \delta$ the i-th component $\widetilde{R}(\varepsilon)_{ik}$ is a convex combination of all the $\widetilde{R}(\delta)_{jk}$, $j = 1, \dots, n$. It then follows that

$$m_k(\delta) = \min_j \widetilde{R}(\delta)_{jk} \leq \widetilde{R}(\varepsilon)_{ik} \leq \max_j \widetilde{R}(\delta)_{jk} = M_k(\delta).$$

Since these inequalities hold for an arbitrary transient state i it follows that

$$m_k(\delta) \leq m_k(\varepsilon) = \min_i \widetilde{R}(\varepsilon)_{ik} \leq \max_i \widetilde{R}(\varepsilon)_{ik} = M_k(\varepsilon) \leq M_k(\delta)$$

and the conclusion follows. So let $\delta > 0$ and $0 < \varepsilon < \delta$. Then

$$\begin{aligned} \widetilde{R}(\varepsilon) = -(E + \varepsilon D)^{-1} \varepsilon F &= (E + \varepsilon D)^{-1}(E + \delta D)(E + \delta D)^{-1}(-\varepsilon F) \\ &= (\widehat{E}_\varepsilon + \delta \widehat{D}_\varepsilon) \frac{\varepsilon}{\delta} \widetilde{R}(\delta). \end{aligned}$$

Now note that $P(\varepsilon, \delta) := (\widehat{E}_\varepsilon + \delta \widehat{D}_\varepsilon)\frac{\varepsilon}{\delta}$ is stochastic. Indeed, since $0 < \varepsilon \leq \delta$ we get by Lemma 3.10(ii) that $\widehat{E}_\varepsilon + \delta \widehat{D}_\varepsilon \geq I \geq 0$ and thus $P(\varepsilon, \delta) \geq 0$ and moreover $P(\varepsilon, \delta)\mathbf{1} = \frac{\varepsilon}{\delta}\widehat{E}_\varepsilon \mathbf{1} + \varepsilon \widehat{D}_\varepsilon \mathbf{1} = \mathbf{1}$ where we applied that $\varepsilon \widehat{D}_\varepsilon$ is stochastic by Lemma 3.10(iv) and $\widehat{E}_\varepsilon \mathbf{1} = (E + \varepsilon D)^{-1} E \mathbf{1} = 0$ since $E\mathbf{1} = 0$. Therefore $\widetilde{R}(\varepsilon)_{ik} = \sum_j P(\varepsilon, \delta)_{ij} \widetilde{R}(\delta)_{jk}$ is a convex combination of all the $\widetilde{R}(\delta)_{jk}$ for any i. □

Remark 3.13.

1. We recall that in contrast to the enveloping functions $m_k(\varepsilon)$ and $M_k(\varepsilon)$, a fixed component function $\widetilde{R}(\varepsilon)_{ik}$ of $\widetilde{R}(\varepsilon)$ need not be increasing or decreasing (Figure 2).

2. The lower bound $\widehat{E}_\varepsilon + \delta \widehat{D}_\varepsilon \geq I$ from Lemma 3.10(ii) provides additional information on the behaviour of the hitting probabilities $\widetilde{R}(\varepsilon)$: from $\widetilde{R}(\varepsilon) = (\widehat{E}_\varepsilon + \delta \widehat{D}_\varepsilon)\frac{\varepsilon}{\delta}\widetilde{R}(\delta)$ and $\widetilde{R}(\delta) \geq 0$ it follows that $\widetilde{R}(\varepsilon) \geq \frac{\varepsilon}{\delta}\widetilde{R}(\delta)$ for $0 < \varepsilon \leq \delta$. In other words, any component of $\frac{1}{\varepsilon}\widetilde{R}(\varepsilon)$ is decreasing in ε and by differentiating it we deduce that $\frac{d}{d\varepsilon}\widetilde{R}(\varepsilon) \leq \frac{1}{\varepsilon}\widetilde{R}(\varepsilon)$ for each $\varepsilon > 0$. The inequality $\widetilde{R}(\varepsilon) \geq \frac{\varepsilon}{\delta}\widetilde{R}(\delta)$ for $0 < \varepsilon \leq \delta$ also implies that for any $\delta > 0$ the graph of $\widetilde{R}(\varepsilon)_{ik}$ in the interval $(0, \delta]$ is always above the line connecting the origin with the point $(\delta, \widetilde{R}(\delta)_{ik})$.

4. Asymptotic Limit of Hitting Probabilities

In this section we analyze the limiting behaviour of the perturbed hitting probabilities $\widetilde{R}(\varepsilon)$, i.e. we establish the limit of $\widetilde{R}(\varepsilon)$ as $\varepsilon \to 0$. We begin with the following

Lemma 4.1. Fix $\varepsilon > 0$. Then

(i) $\widehat{E}_\varepsilon \widehat{E}_\varepsilon^D$ and $\widehat{D}_\varepsilon \widehat{E}_\varepsilon^D$ do not depend on ε.

(ii) $\operatorname{ind}(\widehat{E}_\varepsilon) = 1$ and thus $\widehat{E}_\varepsilon \widehat{E}_\varepsilon^D \widehat{E}_\varepsilon = \widehat{E}_\varepsilon$.

(iii) $\varepsilon \widehat{D}_\varepsilon (I - \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D) = I - \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D$

(iv) $(I + \varepsilon \widehat{E}_\varepsilon^D \widehat{D}_\varepsilon)^{-1} \widehat{E}_\varepsilon^D = \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D$

Proof. (i) can be found in [26, Theorem 3.1.2, p. 36].
(ii) By Lemma 3.10(iv), $-\widehat{E}_\varepsilon$ is a generator matrix. Such matrices are semistable and singular and hence $\operatorname{ind}(\widehat{E}_\varepsilon) = 1$. Therefore, $\widehat{E}_\varepsilon^D$ is the group inverse of $\widehat{E}_\varepsilon$ and thus $\widehat{E}_\varepsilon \widehat{E}_\varepsilon^D \widehat{E}_\varepsilon = \widehat{E}_\varepsilon$.
(iii) By applying Lemma 3.10(ii, iii) and this lemma (ii) we compute

$$\varepsilon \widehat{D}_\varepsilon (I - \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D) = (I - \widehat{E}_\varepsilon)(I - \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D) = I - \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D - \widehat{E}_\varepsilon + \widehat{E}_\varepsilon^2 \widehat{E}_\varepsilon^D = I - \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D.$$

(iv) From Lemma 3.10(ii, iii) and this lemma (ii) it follows that

$$\widehat{E}_\varepsilon (I + \varepsilon \widehat{E}_\varepsilon^D \widehat{D}_\varepsilon) = \widehat{E}_\varepsilon (I + \widehat{E}_\varepsilon^D (I - \widehat{E}_\varepsilon)) = \widehat{E}_\varepsilon + \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D - \widehat{E}_\varepsilon^2 \widehat{E}_\varepsilon^D = \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D.$$

Since $\widehat{E}_\varepsilon$ and $\widehat{D}_\varepsilon$ commute it follows that

$$(I + \varepsilon \widehat{E}_\varepsilon^D \widehat{D}_\varepsilon)^{-1} \widehat{E}_\varepsilon^D = (\widehat{E}_\varepsilon (I + \varepsilon \widehat{E}_\varepsilon^D \widehat{D}_\varepsilon))^D = (\widehat{E}_\varepsilon \widehat{E}_\varepsilon^D)^D = \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D$$

where in the last step we used that $\widehat{E}_\varepsilon \widehat{E}_\varepsilon^D$ is a projection. □

In order to establish the limit of $\widetilde{R}(\varepsilon)$ as $\varepsilon \to 0$ we show that $\widetilde{R}(\varepsilon)$ can be extended to an analytic function on $(-\varepsilon_0, \infty)$ for some $\varepsilon_0 > 0$ and establish its power series expansion at $\varepsilon = 0$. For this purpose, we first state the following

Proposition 4.2. The generalized resolvent $(E+\varepsilon D)^{-1}$ satisfies

$$(E+\varepsilon D)^{-1} = \left((I-\widehat{\Pi})(\widehat{E}_\delta+\varepsilon\widehat{D}_\delta)^{-1}+\widehat{\Pi}\frac{\delta}{\varepsilon}\right)(E+\delta D)^{-1} \tag{6}$$

where $\delta > 0$ is arbitrary and $\widehat{\Pi} := I - \widehat{E}_\delta \widehat{E}_\delta^D$ (independent of δ).

Proof. First write

$$\begin{aligned}(E+\varepsilon D)^{-1} &= (E+\varepsilon D)^{-1}(E+\delta D)(E+\delta D)^{-1}\\ &= ((E+\delta D)^{-1}(E+\varepsilon D))^{-1}(E+\delta D)^{-1}\\ &= (\widehat{E}_\delta+\varepsilon\widehat{D}_\delta)^{-1}(E+\delta D)^{-1}.\end{aligned}$$

Decompose $(\widehat{E}_\delta+\varepsilon\widehat{D}_\delta)^{-1}$ with respect to the projection $\widehat{\Pi} = I - \widehat{E}_\delta\widehat{E}_\delta^D$:

$$(\widehat{E}_\delta+\varepsilon\widehat{D}_\delta)^{-1} = (I-\widehat{\Pi})(\widehat{E}_\delta+\varepsilon\widehat{D}_\delta)^{-1}+\widehat{\Pi}(\widehat{E}_\delta+\varepsilon\widehat{D}_\delta)^{-1}.$$

Simplify the right hand side as required by applying Lemma 3.10(i) which gives

$$\widehat{\Pi}(\widehat{E}_\delta+\varepsilon\widehat{D}_\delta)^{-1} = \widehat{\Pi}(\widehat{E}_\varepsilon+\delta\widehat{D}_\varepsilon) = \delta\widehat{D}_\varepsilon\widehat{\Pi} = \frac{\delta}{\varepsilon}\widehat{\Pi}$$

where in the last two steps we also applied Lemma 3.10(iii) and Lemma 4.1(ii, iii). □

Remark 4.3. In [26, Proof of Theorem 4.2.1, p. 80] one can also find the Laurent series expansion at 0 of the generalized resolvent $(E+\varepsilon D)^{-1}$ which takes the form

$$(E+\varepsilon D)^{-1} = \left(\widehat{E}_\delta^D\sum_{k=0}^{\infty}(-\widehat{E}_\delta^D\widehat{D}_\delta)^k\varepsilon^k+\widehat{D}_\delta^D\widehat{\Pi}\frac{1}{\varepsilon}\right)(E+\delta D)^{-1}.$$

This expansion can be also deduced from Eq. (6) by using the identities $\widehat{D}_\delta^D\widehat{\Pi} = \delta\widehat{\Pi}$ (which follows from Lemma 4.1(iii)) together with $(I-\widehat{\Pi})(\widehat{E}_\delta+\varepsilon\widehat{D}_\delta)^{-1} = \widehat{E}_\delta^D(I+\varepsilon\widehat{E}_\delta^D\widehat{D}_\delta)^{-1}$ (see proof of Theorem 4.4) and its Neumann series expansion.

In the following, we are going to apply the preceding proposition in order to establish the power series expansion at 0 of the hitting probabilities $\widetilde{R}(\varepsilon)$ which can then be used to compute their asymptotic behaviour as $\varepsilon \to 0$.

Theorem 4.4. The hitting probabilities $\widetilde{R}(\varepsilon)$ are given for $\varepsilon > 0$ by

$$\widetilde{R}(\varepsilon) = \left(\widehat{E}_\delta^D (I + \varepsilon \widehat{M})^{-1} \varepsilon + \widehat{\Pi}\delta\right) \cdot \left(\frac{1}{\delta}\widetilde{R}(\delta)\right) \tag{7}$$

where $\delta > 0$ is arbitrary and $\widehat{M} := \widehat{E}_\delta^D \widehat{D}_\delta$ (independent of δ).

By expanding the term $(I + \varepsilon\widehat{M})^{-1} = \sum_{n=0}^{\infty} (-\widehat{M})^n \varepsilon^n$ into its Neumann series (which converges for all $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$ for some $\varepsilon_0 > 0$ small enough), we observe that the right-hand side of Eq. (7) can be extended to a power series that converges for $\varepsilon \in (-\varepsilon_0, \varepsilon_0)$. This provides an extension of $\widetilde{R}(\varepsilon)$ to an analytic function on $(-\varepsilon_0, \infty)$.

Proof. From Proposition 4.2 we have

$$\begin{aligned}
\widetilde{R}(\varepsilon) &= -(E + \varepsilon D)^{-1} \varepsilon F \\
&= -\left((I - \widehat{\Pi})(\widehat{E}_\delta + \varepsilon \widehat{D}_\delta)^{-1} + \widehat{\Pi}\frac{\delta}{\varepsilon}\right)(E + \delta D)^{-1} \varepsilon F \\
&= \left((I - \widehat{\Pi})(\widehat{E}_\delta + \varepsilon \widehat{D}_\delta)^{-1} \varepsilon + \widehat{\Pi}\delta\right)\frac{1}{\delta}\widetilde{R}(\delta).
\end{aligned}$$

We show that

$$(I - \widehat{\Pi})(\widehat{E}_\delta + \varepsilon \widehat{D}_\delta)^{-1} = \widehat{E}_\delta^D (I + \varepsilon \widehat{E}_\delta^D \widehat{D}_\delta)^{-1}.$$

For this purpose, note that

$$\begin{aligned}
\widehat{E}_\delta^D (I + \varepsilon \widehat{E}_\delta^D \widehat{D}_\delta)^{-1} &= \widehat{E}_\delta^D \widehat{E}_\delta \widehat{E}_\delta^D (I + \varepsilon \widehat{E}_\delta^D \widehat{D}_\delta)^{-1} \\
&= \widehat{E}_\delta^D \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D (I + \varepsilon \widehat{E}_\varepsilon^D \widehat{D}_\varepsilon)^{-1} \\
&= \widehat{E}_\delta^D \widehat{E}_\varepsilon \widehat{E}_\varepsilon \widehat{E}_\varepsilon^D = \widehat{E}_\delta^D \widehat{E}_\varepsilon
\end{aligned}$$

where we used Lemma 4.1(i, iv, ii). In order to show $\widehat{E}_\delta^D \widehat{E}_\varepsilon = (I - \widehat{\Pi})(\widehat{E}_\delta + \varepsilon \widehat{D}_\delta)^{-1}$ we show that $I - \widehat{\Pi} = \widehat{E}_\delta^D \widehat{E}_\varepsilon (\widehat{E}_\delta + \varepsilon \widehat{D}_\delta)$:

$$\begin{aligned}
\widehat{E}_\delta^D \widehat{E}_\varepsilon (\widehat{E}_\delta + \varepsilon \widehat{D}_\delta) &= \widehat{E}_\varepsilon (\widehat{E}_\delta^D \widehat{E}_\delta + \varepsilon \widehat{E}_\delta^D \widehat{D}_\delta) = \widehat{E}_\varepsilon (\widehat{E}_\varepsilon^D \widehat{E}_\varepsilon + \varepsilon \widehat{E}_\varepsilon^D \widehat{D}_\varepsilon) \\
&= \widehat{E}_\varepsilon^D \widehat{E}_\varepsilon (\widehat{E}_\varepsilon + \varepsilon \widehat{D}_\varepsilon) = \widehat{E}_\varepsilon \widehat{E}_\varepsilon = I - \widehat{\Pi}
\end{aligned}$$

where we applied Lemma 3.10(ii, iii) and Lemma 4.1(i). Finally, the desired identity in Eq. (7) follows. □

Corollary 4.5. The componentwise limit of $\widetilde{R}(\varepsilon)$ as $\varepsilon \to 0$ exists and is given by

$$\widetilde{R}(0+) := \lim_{\varepsilon \to 0} \widetilde{R}(\varepsilon) = \Pi N F$$

where N is the diagonal matrix with block structure $N = \operatorname{diag}(0, \frac{1}{c_1} I, \ldots, \frac{1}{c_r} I)$ and $c_k := \|\pi_k D_k\|_1$.

Proof. Letting $\varepsilon \to 0$ in Eq. (7) then as $(I + \varepsilon \widehat{M})^{-1}$ is bounded in a neighborhood of 0 the componentwise limit $\widetilde{R}(0+)$ of $\widetilde{R}(\varepsilon)$ as $\varepsilon \to 0$ exists and is given by $\widetilde{R}(0+) = \widehat{\Pi} \widetilde{R}(1)$ where we have chosen $\delta := 1$ and $\widehat{\Pi} = I - \widehat{E}_1 \widehat{E}_1^D$. By Proposition 3.9 it holds $\widehat{\Pi} = \Pi W$ and moreover $W = -ND = -DN$ (N and D are diagonal). Therefore $\widetilde{R}(0+) = \Pi N D (E + D)^{-1} F$. Since each diagonal block of N is a multiple of the identity and $\Pi E = 0$ it follows by looking at the joint block structures of Π, N and E that $\Pi N E = \Pi E N = 0$. We conclude that

$$\widetilde{R}(0+) = \Pi N D (E + D)^{-1} F = (\Pi N E + \Pi N D)(E + D)^{-1} F = \Pi N F. \quad \square$$

Example: In order to provide some more intuition to the expression of $\widetilde{R}(0+)$ in Corollary 4.5 let us return to the example from Figure 2 and compute the limiting probabilities $\widetilde{R}(0+)_{i2}$ into state 8 (second column of $\widetilde{R}(0+)$) in a more intuitive way. The Markov chain E of the transient states $1, \ldots, 6$ of Q has two recurrent classes $\{3, 4\}, \{5, 6\}$ and one transient class $\{1, 2\}$. The ergodic projection $\Pi = I - EE^D = \lim_{t \to \infty} e^{Et}$ containing the limiting distributions for the states $1, \ldots, 6$ of E can be computed to

$$\Pi = \begin{pmatrix} 0 & 0 & \frac{4}{15} & \frac{8}{15} & \frac{2}{35} & \frac{1}{7} \\ 0 & 0 & \frac{1}{5} & \frac{2}{5} & \frac{4}{35} & \frac{2}{7} \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{2}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{2}{7} & \frac{5}{7} \\ 0 & 0 & 0 & 0 & \frac{2}{7} & \frac{5}{7} \end{pmatrix}$$

For the recurrent states $i = 3, 4$ the limits $\widetilde{R}(0+)_{i2}$ are equal and given by

$$\widetilde{R}(0+)_{32} = \widetilde{R}(0+)_{42} = \frac{\frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1}{\frac{1}{3} \cdot (1 + 0) + \frac{2}{3} \cdot (0 + 1)} = \frac{2}{3}.$$

Similarly, for the recurrent states $i = 5, 6$ we get

$$\widetilde{R}(0+)_{52} = \widetilde{R}(0+)_{62} = \frac{\frac{2}{7} \cdot 1 + \frac{5}{7} \cdot 0}{\frac{2}{7} \cdot (0 + 1) + \frac{5}{7} \cdot (2 + 0)} = \frac{1}{6}.$$

In order to compute the limiting probabilities $\widetilde{R}(0+)_{i2}$ for the transient states $i = 1, 2$ first observe that for small $\varepsilon > 0$ the transitions from i to an absorbing state in Q can be neglected as $\varepsilon \to 0$. Hence, we only have to take into account the absorbing transitions from recurrent states $j = 3, \dots 6$ and the limiting probability to find the system in state j (when started in i) from which it potentially may get absorbed. As an example, for $i = 1$ we get

$$\widetilde{R}(0+)_{12} = \left(\frac{4}{15} + \frac{8}{15}\right) \cdot \frac{2}{3} + \left(\frac{2}{35} + \frac{1}{7}\right) \cdot \frac{1}{6} = \frac{17}{30} \approx 0.567$$
$$\widetilde{R}(0+)_{22} = \left(\frac{1}{5} + \frac{2}{5}\right) \cdot \frac{2}{3} + \left(\frac{4}{35} + \frac{2}{7}\right) \cdot \frac{1}{6} = \frac{7}{15} \approx 0.467.$$

5. Application to Model Repair

In this section, we study a class of systems which is characterized as follows: The system performs some regular task, but every now and then it needs to enter some special situation (such as service, update, repair, refueling, ...) during which time it is vulnerable and may fail. The system should remain active for at least some minimum time t, and the goal is that it should eventually terminate successfully (instead of failing). Modelled as an absorbing CTMC, the general structure of such systems is shown in Figure 3. The model has two absorbing states, labelled *success* and *failed*. The transient states are either *regular* or *vulnerable* states, with transitions between them, and for convenience we define the abbreviation $work = regular \vee vulnerable$. Another characteristic is that the system suffers from ageing, i.e. it moves from higher levels to lower levels, where typically vulnerable states will be visited more frequently with growing age and the failure rates also increase with growing age. Concrete examples of such systems are:

- A spacecraft which needs to survive at least a minimum mission time. At certain intervals it needs to perform a docking manoeuvre which is dangerous. With growing age, the docking manoeuvre may become even more risky.

- A submarine on a secret mission which needs to surface every now and then, during which time it is vulnerable because it can be detected more easily.

- A system for Software Version Release Management (SVRM), described in more detail below in Sec. 5.2.

Note that, in the generic CTMC depicted in Figure 3, depending on the modelled application, the "diagonal" transitions (like $2 \to 3$) need not exist, and that in some applications additional transitions (like $1 \to 4$) might make sense, but this does not affect the validity of our approach.

The general requirement for such systems can be expressed with the help of the logic CSL [4] as the following time-bounded Until formula, with lower probability bound b and lower time bound t:

$$\Phi = P_{\geq b}(work\, \mathcal{U}^{\geq t}\, success) \tag{8}$$

This requirement states that the probability should be at least b that the system spends at least t time units in the *regular* or *vulnerable* states (i.e. it does not get absorbed for at least t time units) before eventually entering the *success* state. This implies, of course, that the absorbing *failed* state should never be entered.

Note that, in order to apply the method described in this section, the CTMC under investigation does not have to be absorbing a priori. It could be a non-absorbing CTMC with some of its states labelled by *regular*, *vulnerable* or *success*, where some states should satisfy the requirement of Eq. (8). For analyzing (and potentially repairing) the CTMC, the behaviour after entering a *success* state or otherwise entering any state not satisfying *work* (i.e. $regular \vee vulnerable$) is irrelevant, thus those states can be made absorbing, leading to a situation as depicted in Figure 3.

5.1. An algorithm for model repair

We consider a CTMC with generator Q, with n transient states and $m = 2$ absorbing states $\{failed, success\}$ which are both reachable from all the transient states. As explained in the previous section, the CTMC represents a system that performs useful work for some time (all transient states carry the label *work*) and – if all goes well – eventually finishes by moving to the state *success*. We consider the CSL time-bounded Until formula given above in Eq. (8) with lower probability bound $0 < b < 1$ and lower time bound $t > 0$. For now we assume that it is required that each transient state should satisfy Φ. In a concrete application, though, it may be sufficient that only a proper subset of the transient states, e.g. the initial state, satisfies Φ, which makes the process of model repair easier.

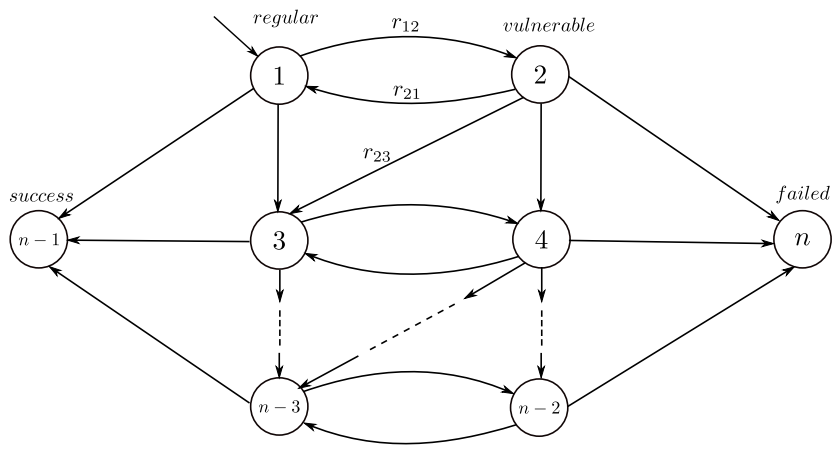

Figure 3: General structure of considered class of Markov chains

If some of the transient states violate requirement Φ, the system should be "repaired", i.e. modified according to some strategy. Among the many possible approaches to model repair, such as adding / removing states or transitions, we advocate a scheme where the structure of the CTMC remains untouched, but transition rates may be reduced. The rationale behind rate reduction is that in most real systems, slowing down a process (a processor, a machine, etc.) is possible, while acceleration may not be feasible. However, rate reduction still leaves many degrees of freedom. For example, each transition could be reduced by its individual reduction factor, which could lead to good solutions but would open a possibly huge multidimensional search space. Therefore we restrict ourselves further by only allowing for common reduction factors applied to sets of transitions.

Basically, for a transient state s, there can be two reasons for violating requirement Φ:

(1) The hitting probability from s to state *success* is too low (in other words, the hitting probability to state *failed* is too high).

(2) The hitting probability to state *success* is high enough, but the time to absorption (starting from s) is too short.

We propose a general solution which takes into account (1) and (2) and is guaranteed to lead to a solution for all transient states.

(I) We first try to deal with both (1) and (2) at the same time by applying the common reduction factor $0 < \eta \leq 1$ to all transitions from the

transient class to state *failed*. As η is reduced, the probability of getting absorbed in state *success* can be made arbitrarily close to 1, and at the same time the system will become "slower", since the exit rates of the transient states are reduced. Depending on the case at hand, it may be possible to find some $0 < \eta \leq 1$ such that Φ will be satisfied for all transient states, in which case we are done. But it is also possible that no such η exists (since the system goes to absorption too early, even though some rates were reduced), in which case we need to proceed.

Note that, in practice, it can be checked in advance whether step (I) will lead to a solution. For this it is enough to set $\eta = 0$, thereby making the hitting probability to *success* equal to 1, and then check whether all transient states satisfy Φ in the thus modified CTMC (i.e. whether they satisfy the time bound).

(IIa) If step (I) was not successful, we first concentrate on the time-unbounded problem, i.e. we deal exclusively with issue (1). The weakened requirement for this step is $\Phi' = P_{>b}(\mathit{work}\ \mathcal{U}\ \mathit{success})$, where the time bound has been removed and the probability bound has been changed from $\geq b$ to $> b$. As shown in [31], one can always find a common reduction factor η_{ut} (applied simultaneously to all transitions from transient states to state *failed*) such that all transient states satisfy the time-unbounded requirement Φ'. After step (IIa) we always move to (IIb).

(IIb) In this final step, we deal with issue (2). We keep factor η_{ut} fixed and return to the original time-bounded requirement Φ. We now introduce a second common reduction factor $0 < \varepsilon \leq 1$ to all transitions from transient states to all absorbing states (*failed* and *success*). The purpose is to slow down the system, such that absorption before t becomes less likely. It is essential that this perturbation by factor ε does not cause the hitting probabilities, which were already fixed in step (IIa), to violate the bound b. This is where we need Theorem 3.12 from Sec. 3.3, which guarantees that during this slow-down the hitting probabilities are preserved in the admissible range.

Proposition 5.1. The procedure described in steps (I), (IIa) and (IIb) solves the model repair problem for the given requirement $\Phi = P_{\geq b}(\mathit{work}\ \mathcal{U}^{\geq t}\ \mathit{success})$, for all transient states.

Proof. The goal is that all transient states s should satisfy $P_{\geq b}(\mathit{work}\ \mathcal{U}^{\geq t}\ \mathit{success})$,

where b and t are fixed. As already observed in [31], it is clear that one can find a solution for the corresponding time-unbounded problem, i.e. one can find a reduction factor η_{ut} s.t. $s \models P_{>b}(work\ \mathcal{U}\ success)$ for all $0 < \eta \leq \eta_{ut}$, or in other words, that $b < Pr^{\eta_{ut}}(s, work\ \mathcal{U}\ success)$ (the superscript indicates the probability measure for the Markov chain modified according to (IIa)). Keeping η_{ut} fixed, we know by Theorem 3.12, shown in Sec. 3.3, that for all transient states s of the CTMC in which all rates to absorption are further reduced by the common reduction factor $0 < \varepsilon \leq 1$ according to (IIb) the following inequality also holds:

$$b < Pr^{\eta_{ut},\varepsilon}(s, work\ \mathcal{U}\ success)$$

As $\varepsilon \to 0$, the right hand side converges to some value $p_s \geq b$ (Corollary 4.5), and this limit is the same for all transient states s belonging to the same bottom communication class of E. Since $\min_s Pr^{\eta_{ut},\varepsilon}(s, work\ \mathcal{U}\ success)$ (taken over all transient states s) is decreasing in ε it actually follows that $p_s > b$ for all s. Now, for any Markov chain it trivially holds that

$$Pr(s, work\ \mathcal{U}\ success) = Pr(s, work\ \mathcal{U}^{<t}\ success) + Pr(s, work\ \mathcal{U}^{\geq t}\ success).$$

We apply this to the Markov chain modified by both reduction factors η_{ut} and ε and combine it with the previous inequality:

$$\begin{aligned} b &< Pr^{\eta_{ut},\varepsilon}(s, work\ \mathcal{U}\ success) \\ &= Pr^{\eta_{ut},\varepsilon}(s, work\ \mathcal{U}^{<t}\ success) + Pr^{\eta_{ut},\varepsilon}(s, work\ \mathcal{U}^{\geq t}\ success). \end{aligned}$$

Since as $\varepsilon \to 0$ the first term of the sum vanishes, the second term of the sum converges to p_s. From $p_s > b$ it follows that there is $\varepsilon > 0$ for which the second term is $\geq b$ for all transient states s. □

5.2. Model repair of a SVRM system

In order to show the effectiveness and usability of our model repair algorithm, we now consider as a use-case the problem of Software Version Release Management (SVRM) from the software engineering domain. The SVRM model is given in Figure 4, which is a concretisation of Figure 3. It consists of 8 states, two of which are absorbing. According to the setting in Sec. 3 that means $n = 6$ and $m = 2$, where the transient states corresponding to the generator matrix E are $\{1, 2, 3, 4, 5, 6\}$. These transient states form $k = 3$ communication classes, namely $\{1, 2\}$, $\{3, 4\}$ and $\{5, 6\}$. Only $r = 1$ of

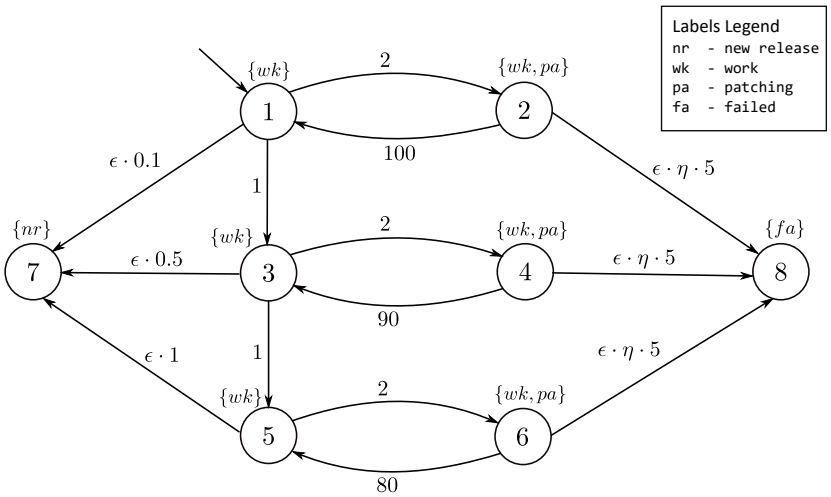

Figure 4: SVRM model

these classes is recurrent (i.e. bottom) for E, namely the class $\{5, 6\}$, whereas the set of states $\{1, 2, 3, 4\}$ is transient for E. In Figure 4, states are labeled with atomic propositions whose meaning is explained in the legend. At the time of a new release, the software system starts in state 1. The transition chain $1 \to 3 \to 5$ corresponds to ageing of the software. In states $2, 4$ and 6, patching takes place, which may cause a failure, represented by the absorbing state 8. From the non-patching states, it is possible that a new release of the software is ready for issue, which will take the system to the absorbing state 7 which marks the end of the current software cycle. The basic time unit for this model is one month, so for example the mean sojourn time (per visit) of state 1, which has exit rate 3.1, is 0.323 months, which is approximately 10 days. Starting from state 1, the mean time to absorption of the model is 1.862 months (approx. 56 days), and the hitting probability in the desired state 7 is 81%. Note that the the mean duration of *patching* goes up as the system ages (the exit rates of states $2, 4, 6$ decrease, in that order), and the failure probability during patching also increases with growing age (from $\frac{5}{105}$ via $\frac{5}{95}$ to $\frac{5}{85}$), which is what operators of software systems experience in practice. Figure 4 also depicts the two reduction factors, η and ε, which are employed by the model repair algorithm. In real life, factor η can be influenced by making the patching process less error-prone, and factor ε could be decreased by again improving the patching process and at the same time reducing the issue rate of new releases. In practice, the latter two effects

could be jointly achieved by reassigning software developers from the new release team to the patch development team (which would delay the issuing of new releases and at the same time improve the quality of the patches). It should be pointed out that the model in Figure 4 represents only a single cycle of the software, i.e. after installing a new release the system will restart in state 1.

Distribution and installation of a new release is expensive, so this should not take place too often, i.e. not too soon after having started a new cycle. Therefore, the user property that needs to be verified on the model is given in Eq. (9).

$$\Phi = P_{\geq b}(work\, \mathcal{U}^{\geq t}\, newrelease) \tag{9}$$

This user property seeks that the probability should be at least b that the system will be working beyond the time point t before a new release is issued (which implies that the system should never move to the $failed$ state).

For this example, it is sufficient that the starting state 1 satisfies requirement (9), because state 1 represents the beginning of a software release cycle. Therefore we now perform model repair for state 1. For this purpose, we set the time-bound as $t = 1$ and the probability bound as $b = 0.85$. The actual probability of state 1 to satisfy Φ is found to be 0.58835. So, in order to make the probability to reach the bound b, according to step (I) of our algorithm, one needs to reduce η. But as $\eta \to 0$, the maximum probability that can be obtained is only 0.765, which is less than required. Therefore, the two step procedure explained above needs to be deployed: For the untimed requirement Φ', in order to make the probability for state 1 higher than 0.85, the value of η_{ut} in step (IIa) of the algorithm is computed as $\eta_{ut} \approx 0.751$. Thereafter, in step (IIb), by fixing $\eta = \eta_{ut}$, and then reducing reduction factor ε, the satisfying value for the original time-bounded requirement Φ is found to be $\varepsilon = 0.178$, which solves the model repair problem for state 1. In Figure 5, probability curves for state 1 and requirement Φ are plotted against reduction factor ε (all for fixed $\eta_{ut} = 0.751$), to illustrate the behavior for different time-bounds. One can see that the red curve, related to time-bound $t = 1$, indeed crosses the 0.85 level at the value $\varepsilon = 0.178$.

Having successfully performed model repair for state 1, we also wish to show that our algorithm is able to perform model repair for *all* transient states of this type of model (even though this is not required by this SVRM application). Therefore, we will now use our algorithm to repair the model for the full set of transient states $\{1, 2, 3, 4, 5, 6\}$, such that they all satisfy

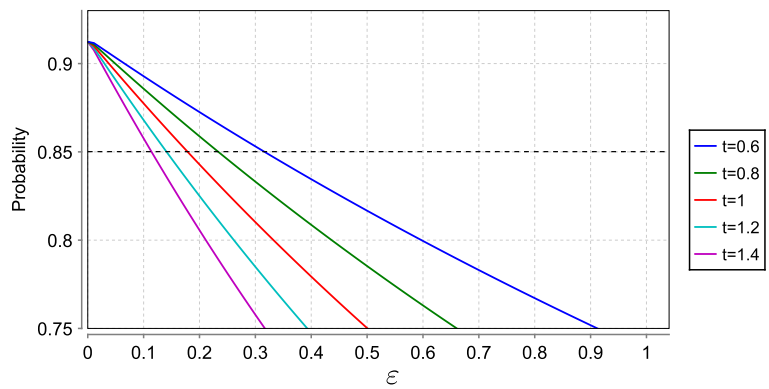

Figure 5: Probability of state 1 to satisfy Φ depending on ε for different values of time-bound t

State	Timed Φ original prob.	Timed Φ at $\eta = 0.821$
1	0.58835	0.61464
2	0.56291	0.59307
3	0.41342	0.42977
4	0.39524	0.41474
5	0.29980	0.31146
6	0.28584	0.30009

Table 1: Probabilities for each state to satisfy Φ for $t = 1$ (the goal is to exceed probability bound $b = 0.3$)

requirement Φ given in Eq. (9). Let the value for the time and probability bounds be $t = 1$ (i.e. one month) and $b = 0.3$ (this latter value would, of course, not be of practical interest). The results of this verification are given in the second column of Table 1. From those values, we can see that states 5 and 6 violate the requirement, since their satisfaction probabilities are below 0.3, so the model should be repaired. According to step (I) of our algorithm, the strategy is to reduce η, and Figure 6 shows how the probabilities to satisfy Φ increase as η is reduced. The satisfying solution is found to be $\eta = 0.821$, and the repaired probabilities for each state are shown in the third column of Table 1.

As a second scenario, in order to demonstrate the two step model repair procedure, the same property Φ is now verified with the modified time and

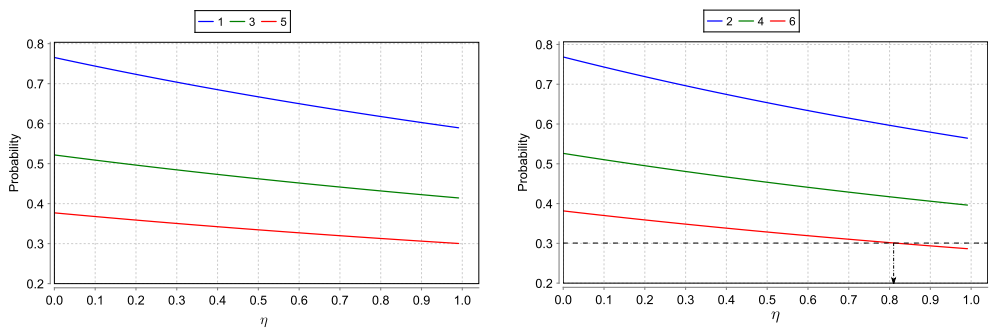

Figure 6: Probability of different states to satisfy Φ for $t = 1$ while reducing factor η

probability bounds again given as $t = 1$ and $b = 0.85$. This is the same problem as earlier, but this time the goal is to make all states from the set $\{1, 2, 3, 4, 5, 6\}$ satisfying. The actual probabilities are already computed and are shown in the second column of Table 1. Applying step (I) of our algorithm, it turns out that by reducing only the factor η, the probability of state 6 reaches a maximum of 0.382 and state 5 reaches a maximum of 0.377 as $\eta \to 0$ (which can also be observed from Figure 6). Therefore, step (I) does not lead to a solution. So we need to apply step (IIa), which means we determine the satisfying η_{ut} value for the untimed property Φ'. This value is found to be $\eta_{ut} = 0.605$, and the probabilities for each state are given in the third column of Table 2. Note that this also helped to move the probabilities to satisfy the original time-bounded requirement Φ in the right direction (column four of Table 2), although not far enough. Keeping this η_{ut} fixed, we then proceed with step (IIb), where we find that reducing the factor ε to the value 0.0835 will cause all the six states to satisfy Φ (see last column of Table 2 and Figure 7).

As Prop. 5.1 states, the algorithm presented in Sec. 5.1 is guaranteed to find a feasible solution, such that Φ is satisfied for all transient states. However, this solution is not unique, and the values of η (or η_{ut} and ε) determined by the algorithm are not claimed to be optimal in any way. In Figure 8, the satisfying region is shown, where the solution found by our algorithm is marked by the black dot very close to its boundary[10] (note that the border curve is not linear). As we can see, for this particular combination

[10]In fact, it is assured by the algorithm that the solution lies within the satisfying region and is very close to its boundary, as determined by a precision predefined by the user.

State	Untimed Φ' ($\eta = 1$)	Untimed Φ' ($\eta_{ut} = 0.605$)	Timed Φ ($\eta_{ut} = 0.605, \varepsilon = 1$)	Timed Φ ($\eta_{ut} = 0.605, \varepsilon = 0.0835$)
1	0.81059	0.87600	0.64911	0.90008
2	0.77199	0.85028	0.63284	0.89818
3	0.86885	0.91504	0.45108	0.87390
4	0.82312	0.88529	0.44030	0.87214
5	0.89473	0.93207	0.32665	0.85185
6	0.84210	0.89811	0.31877	0.85010

Table 2: Probabilities for each state to satisfy Φ', resp. Φ for $t = 1$ (the goal is to exceed probability bound $b = 0.85$)

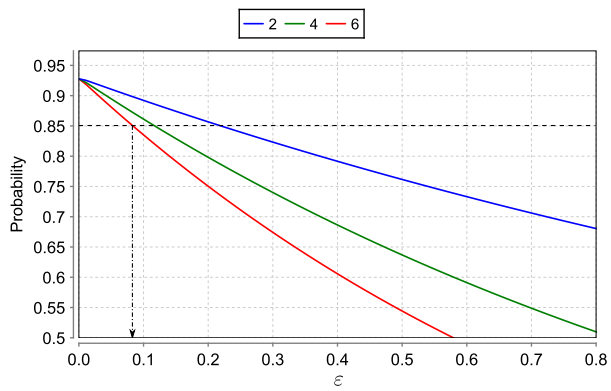

Figure 7: Probability for patching states to satisfy Φ while reducing ε (for fixed $\eta_{ut} = 0.605$)

of probability bound $b = 0.85$ and time bound $t = 1$, even reducing ε alone would lead to a solution. However, for other combinations (e.g. $b = 0.9, t = 1$), reducing only ε would not lead to a solution. In general, exploring the satisfying region for a given model repair problem is an interesting topic for future work.

We close this section with a note on the computational aspects: All calculations made in Sec. 5 (except Fig. 8, which was prepared using Wolfram Mathematica), were made with the help of the model checker PRISM [5]. Satisfying values for the parameters η, η_{ut} and ε were found by binary search in the interval $(0, 1]$ with a predefined precision of 10^{-4}, similarly as in [32]. That means, the models were re-evaluated for different values of the parameters, using PRISM's standard algorithms for solving linear systems (in the

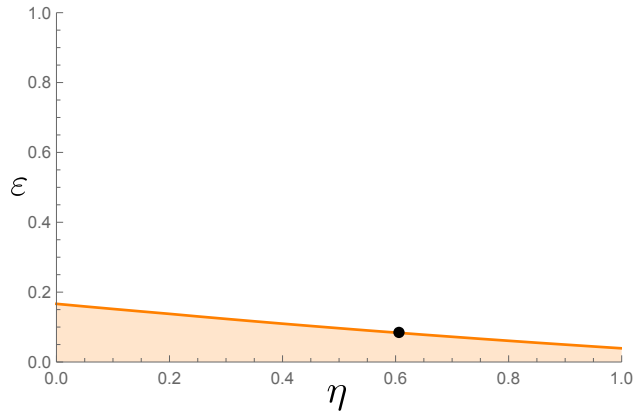

Figure 8: Region where requirement Φ is satisfied (shaded area)

case of η_{ut}) and uniformization (for η and ε). Finding the reduction factors with the specified precision, and even computing all the graphs in the figures, is a matter of a few seconds on a standard laptop. This approach scales well to much larger models as long as PRISM (or an alternative tool) can handle the system of equations. Using the parametric analysis methods mentioned in the introduction (such as [17, 19]), one could only tackle the linear system (i.e. not the time-bounded case which requires uniformization), and for large models that approach suffers from the symbolic blow-up of the resulting rational function.

6. Conclusion and Future Work

From a theoretical point of view, this paper has closed a gap in the literature about absorbing Markov chains, where all rates to absorption are perturbed by a common factor ε. The paper has studied the absorption probabilities from transient states to a particular absorbing state, seen as a function of the perturbation factor ε. It has been found that the absorption probability from a certain transient state is not necessarily monotonic, as an illustrating example has demonstrated. However, if we do not concentrate on a single transient state but look at at the envelope for all transient states, i.e. the minimum or maximum of the hitting probabilities over all transient states to a particular absorbing state, that function is indeed monotonic. This fact has been proven in Theorem 3.12, the central result of this paper. This generalizes the results of the earlier conference paper [13] in that the restrictive assumption that all transient states need to be in the same communication

class has been dropped. In addition, the paper has also established important results about the exitpoint Markov chain and the asymptotic limit of the hitting probabilities.

As a practical contribution, the paper has proposed an algorithm for repairing a CTMC model in case it violates a requirement given in the form of a CSL Until formula with lower time bound and lower probability bound. This algorithm employs up to two reduction factors, one for biasing the hitting probability, and the other one for slowing down the absorption process in order to meet the time bound. With the help of the theoretical results from the earlier sections, it has been shown that the algorithm will always find a solution to this type of model repair problem. The article has presented a class of CTMC models where such an algorithm can be useful, and it has discussed in depth a particular case study from the domain of software version release management.

Future work on the theoretical side will address the question of how the theory in Sec. 3 and 4 could be extended if also transitions between transient states are perturbed. On the practical side, we are planning to develop our model repair algorithm further. As proven, the presented algorithm always finds a feasible solution, but the future goal will be to find a solution which is optimal with respect to some criterion yet to be defined. As we saw in Figure 8, there can be a whole satisfying region, and this picture would become even more differentiated if one allowed more than two reduction factors to be employed, which opens up an interesting search field.

References

[1] E. Bartocci, R. Grosu, P. Katsaros, C. Ramakrishnan, S. Smolka, Model Repair for Probabilistic Systems, in: TACAS'11, Springer LNCS 6605, 2011, pp. 326–340.

[2] T. Chen, E. M. Hahn, T. Han, M. Kwiatkowska, H. Qu, L. Zhang, Model Repair for Markov Decision Processes, in: Proc. 7th Int. Symp. Theor. Aspects of Software Engineering (TASE), IEEE CS Press, 2013, pp. 85–92.

[3] S. Pathak, E. Ábrahám, N. Jansen, A. Tacchella, J.-P. Katoen, A Greedy Approach for the Efficient Repair of Stochastic Models, in: NASA Formal Methods - 7th Int. Symposium, 2015, pp. 295–309.

[4] C. Baier, B. Haverkort, H. Hermanns, J.-P. Katoen, Model-Checking Algorithms for Continuous-Time Markov Chains, IEEE Transactions on Software Engineering 29 (6) (2003) 524–541.

[5] M. Kwiatkowska, G. Norman, D. Parker, PRISM 4.0: Verification of Probabilistic Real-time Systems, in: Proc. 23rd International Conference on Computer Aided Verification (CAV'11), Springer LNCS 6806, 2011, pp. 585–591.

[6] C. Dehnert, S. Junges, J.-P. Katoen, M. Volk, A Storm is Coming: A Modern Probabilistic Model Checker, in: Proc. Computer Aided Verification (CAV'17), Springer LNCS 10427, 2017, pp. 592–600.

[7] J. N. Darroch, E. Seneta, On Quasi-Stationary Distributions in Absorbing Continuous-Time Finite Markov Chains, J. Appl. Probab. 4 (1) (1967) 192–196.

[8] C. A. O'Cinneide, Characterization of Phase-Type Distributions, Communications in Statistics. Stochastic Models 6 (1) (1990) 1–57.

[9] G. Latouche, V. Ramaswami, Introduction to Matrix Analytic Methods in Stochastic Modeling, Society for Industrial and Applied Mathematics, 1999.

[10] A. Horvath, M. Telek, PhFit: A General Phase-Type Fitting Tool, in: T. Field, et al. (Eds.), TOOLS 2002, Springer LNCS 2324, 2002, pp. 82–91.

[11] A. Bobbio, A. Horváth, M. Telek, Matching Three Moments with Minimal Acyclic Phase Type Distributions, Stochastic Models 2-3 (21) (2005) 303–326.

[12] G. Horvath, M. Telek, On the Canonical Representation of Phase Type Distributions, Performance Evaluation 66 (8) (2009) 396 – 409, selected papers of the Fourth European Performance Engineering Workshop (EPEW) 2007.

[13] A. Gouberman, M. Siegle, B. Tati, Perturbation of CTMC Trapping Probabilities with Application to Model Repair, in: Proc. 11th EAI Int. Conf. on Performance Evaluation Methodologies and Tools (Valuetools 2017), ACM, 2017, pp. 36–43.

[14] T. Han, J. Katoen, A. Mereacre, Approximate Parameter Synthesis for Probabilistic Time-Bounded Reachability, in: Proceedings of the 29th IEEE Real-Time Systems Symposium, RTSS, 2008, pp. 173–182.

[15] M. Ceska, F. Dannenberg, M. Z. Kwiatkowska, N. Paoletti, Precise Parameter Synthesis for Stochastic Biochemical Systems, in: Computational Methods in Systems Biology - 12th International Conference, CMSB, 2014, pp. 86–98.

[16] T. Quatmann, C. Dehnert, N. Jansen, S. Junges, J.-P. Katoen, Parameter Synthesis for Markov Models: Faster Than Ever, in: ATVA 2016, Springer LNCS 9938, 2016, pp. 50–67.

[17] E. M. Hahn, H. Hermanns, L. Zhang, Probabilistic Reachability for Parametric Markov Models, International Journal on Software Tools for Technology Transfer 13 (1) (2011) 3–19.

[18] C. Daws, Symbolic and Parametric Model Checking of Discrete-Time Markov Chains, in: Theoretical Aspects of Computing - ICTAC 2004, Springer LNCS 3407, 2005, pp. 280–294.

[19] E. M. Hahn, H. Hermanns, B. Wachter, L. Zhang, PARAM: A Model Checker for Parametric Markov Models, in: CAV, Springer LNCS 6174, 2010, pp. 660–664.

[20] C. Dehnert, S. Junges, N. Jansen, F. Corzilius, M. Volk, H. Bruintjes, J.-P. Katoen, E. Ábrahám, PROPhESY: A PRObabilistic ParamEter SYnthesis Tool, in: CAV, Springer LNCS 9206, 2015, pp. 214–231.

[21] P. Gainer, E. M. Hahn, S. Schewe, Accelerated Model Checking of Parametric Markov Chains, CoRR abs/1805.05672. `arXiv:1805.05672`.

[22] A. Berman, R. Plemmons, Nonnegative Matrices in the Mathematical Sciences, Computer Science and Applied Mathematics, Academic Press, 1979.

[23] P. Azimzadeh, A fast and stable test to check if a weakly diagonally dominant matrix is a nonsingular M-matrix, Mathematics of Computation (2018).

[24] S. Roch, B. Silbermann, Continuity of generalized inverses in Banach algebras, Studia Mathematica 136 (3) (1999) 197–227.

[25] J. J. Koliha, A generalized Drazin inverse, Glasgow Mathematical Journal 38 (3) (1996) 367–381.

[26] S. L. Campbell, Singular systems of differential equations, Pitman, London, 1980.

[27] N. C. González, Additive perturbation results for the Drazin inverse, Linear Algebra and its Applications 397 (0) (2005) 279 – 297.

[28] R. E. Hartwig, G. Wang, Y. Wei, Some additive results on Drazin inverse, Linear Algebra and its Applications 322 (1 - 3) (2001) 207 – 217.

[29] J. Kemeny, J. Snell, Finite Markov Chains, University series in undergraduate mathematics, Van Nostrand, 1960.

[30] J. Koliha, T. D. Tran, The Drazin inverse for closed linear operators and the asymptotic convergence of C_0-semigroups, Journal of Operator Theory (2001) 323–336.

[31] B. Tati, M. Siegle, Parameter and Controller Synthesis for Markov Chains with Actions and State Labels, in: É. André, G. Frehse (Eds.), 2nd International Workshop on Synthesis of Complex Parameters (SynCoP'15), Vol. 44 of OpenAccess Series in Informatics (OASIcs), Dagstuhl, Germany, 2015, pp. 63–76.

[32] B. Tati, M. Siegle, Rate reduction for state-labelled Markov chains with upper time-bounded CSL requirements, in: T. Brihaye et al. (Ed.), Proc. Cassting Workshop on Games for the Synthesis of Complex Systems and 3rd Int. Workshop on Synthesis of Complex Parameters, Open Publ. Assoc., El. Proc. in Theor. Computer Science Vol. 220, 2016, pp. 77–89.