Available online at www.sciencedirect.com

ScienceDirect

Information Sciences 178 (2008) 1558–1597

INFORMATION SCIENCES
AN INTERNATIONAL JOURNAL

www.elsevier.com/locate/ins

Software testing processes as a linear dynamic system

Kai-Yuan Cai [a,b,*], Zhao Dong [c], Ke Liu [c]

[a] *Department of Automatic Control, Beijing University of Aeronautics and Astronautics, Beijing 100083, China*
[b] *State Key Laboratory of Virtual Reality Technology and Systems, Beijing, China*
[c] *Institute of Applied Mathematics, Academy of Mathematics and System Sciences, Chinese Academy of Sciences, Beijing 100080, China*

Received 26 January 2007; received in revised form 19 October 2007; accepted 22 October 2007

Abstract

Software testing is essential for software reliability improvement and assurance, and the processes of software testing are intrinsically dynamic. However they are seldom investigated in a mathematically rigorous manner. In this paper a theoretical study is presented to examine the dynamic behavior of software testing. More specifically, a set of simplifying assumptions is adopted to formulate and quantify the software testing processes. The mathematical formulae for the expected number of observed software failures are rigorously derived, the bounds and trends of the expected number of observed software failures are analyzed, and the variance of the number of observed software failures is examined. On the other hand, it is demonstrated that under the simplifying assumptions, the software testing processes can be treated as a linear dynamic system. This suggests that the software testing processes could be classified as linear or non-linear, and there be intrinsic link between software testing and system dynamics.

Keywords: Software testing; Software reliability; Linear dynamic system

1. Introduction

Software testing is carried out in every software engineering project. Similar to software process management, it serves as one of the most effective apparatus for software reliability improvement and assurance. The importance of software testing is undoubted. It is estimated that software testing consumes almost half of software development resources [1]. However software testing is claimed to be the least understood part in the software development process [25]. What coverage criteria should be adopted for white-box testing? When should detected defects be removed from the software under test? To what extent could software testing help to guarantee software correctness? How much additional testing efforts might be required to reduce the failure rate of the software under test to 10^{-6} failure/h? How much improvement in software reliability could be gained with additional testing of 500 h? How should 1000 tests be scheduled so that the current reliability of the software under test can be best estimated? Could a unified theory be developed for software testing that

* Corresponding author. Tel.: +86 10 82317328; fax: +86 10 82338414.
E-mail address: kycai@buaa.edu.cn (K.-Y. Cai).

0020-0255/$ - see front matter
doi:10.1016/j.ins.2007.10.024

guides the selection of software test cases to achieve a quantitative software reliability goal given a priori? The answers to these questions remain obscure. A systematic review of empirical studies on software testing techniques concludes that the amount of knowledge of software testing is limited [17]. This raises a natural question: what are the underlying causes for the obscure answers for various questions concerning software testing?

A useful approach to revealing the underlying causes is to do empirical studies. This can be observed in various empirical or experimental studies that are documented in the existing literature [4,6,14,16,24]. A general procedure for a conventional empirical study is as follows:

(1) Select factors or theories whose effects are of interest to be examined.
(2) Select a few subject software processes or systems on which the effects of the factors or theories are to be examined.
(3) Build up the practical scenarios or experimental platforms from which data of interest are to be generated and/or collected.
(4) Analyze the collected data to discover possible theories or examine hypothesized theories.

The above procedure was demonstrated in a recent study of regression testing [23], in which four factors (test suite granularity, test input grouping, regression testing technique, interaction) were considered and four hypotheses, each involving a single factor, were examined. Two subject programs were adopted, each comprising nine consecutive versions.

There is no doubt that empirical studies may generate decisive insights that help to improve software testing practice and enhance software reliability. However the procedures that empirical studies follow are rather ambiguous. The conclusions drawn from empirical studies are seldom accurate or rigorous. Different empirical studies may lead to different or even conflicting conclusions. This led Briand et al [2] to conclude that although many principles and techniques could be reused, software quality did not follow universal laws and quality models must be developed *locally*, wherever needed. Replication is a challenge to empirical studies [9]. Actually, this should not be surprising. Software testing processes are not completely repeatable and are subject to large variations. Software systems operate under diverse operating environments including networking environments without clear-cut boundaries. There is no guarantee that all major factors are properly taken into account in a single empirical study. Therefore, empirical studies are not enough to reveal all the underlying causes for the obscure answers for various questions concerning software testing.

As a complementary apparatus to empirical studies, theoretical studies that simplify and formulate software testing processes in a mathematically rigorous manner should be carried out. An advantage of theoretical studies over empirical studies is that conclusions drawn from theoretical studies can be rigorously proved, as long as the corresponding assumptions are valid. Suppose that software testing can be described in terms of a set of simplifying assumptions, what rigorous conclusions can be drawn for software testing? The main motivation of this paper is to present a theoretical study that examines this question in particular, and reveals the underlying causes for the obscure answers for various questions concerning software testing in general. Consequently, the nature of software reliability testing should be better understood.

More specifically, the purpose of this theoretical study is twofold. On the one hand, during software testing, test cases are selected and applied to the software under test one by one. Detected defects are removed and the software under test is subject to change or evolution. Dynamic behavior can be observed in the test case selection process and the test case execution process, as well as in the software evolution process. Then what is the nature of the dynamics of software testing? Or is there any intrinsic connection between the dynamic behavior of software testing and those of linear or non-linear dynamic systems [10,18]? This question will be answered in Section 4.

On the other hand, as software testing proceeds with more and more software defects being detected and removed, the reliability of the software under test tends to grow. Numerous software reliability growth models have been proposed since 1970s [3,20,28]. In these models various sets of assumptions are taken for the software reliability growth behavior and the process of software test case selection is seldom described in an explicit manner. For example, in the Jelinski–Moranda model, it is assumed that the times between successive observed failures are independent and exponentially distributed. In the Goel–Okumoto NHPP model, it is simply assumed that the cumulative number of observed failures follows a non-homogeneous Poisson process.

In Refs. [19,29], it is directly assumed that the number of defects remaining in the software under test satisfies a stochastic linear differential equation. Although various empirical studies have been devoted to validating or invalidating these models, it is not clear if related assumptions taken in these models can be justified or disproved in a mathematically rigorous manner. A recent empirical study even argues that software reliability behavior is chaotic instead of stochastic [11]. A major cause for controversies of this kind is that test case selection mechanisms are not formulated in an explicit manner, and it is not clear how software testing leads to software reliability improvement in quantitative terms. Therefore, it is highly desirable to examine the following question: how software reliability growth behavior may look like if the software testing process follows a specific mechanism? Most parts of the rest of this paper are devoted to this question.

In order to serve the purpose of the theoretical study, a set of simplifying assumptions for software test case selections is adopted in this paper. With these assumptions, the mathematical formulae for the mean and variance of the cumulative number of observed failures can be derived and the corresponding behavior can be examined in a mathematically rigorous manner. In this way, the software testing process can be formulated as a dynamic system as shown in Section 4. The state of the dynamic system or the software testing process is a vector whose entries are defined in a mathematically rigorous manner and can be interpreted as the expected ratios of failures that have not been revealed by an arbitrary sequence of actions starting with various distinct testing actions. Since the paper is aimed to present a theoretical study for software testing and reliability and to better understand the nature of software reliability improvement process, no empirical or experimental results are included. However, simulation results are presented in Section 7 to illustrate theoretical results and visually judge how software reliability growth behavior may look like if the software test case selection processes are explicitly formulated. Of course, this by no means suggests that empirical or experimental studies be less important for software testing and reliability. How the conclusions drawn from the theoretical study presented in the paper can be applied to assess or improve software reliability behavior in practice is left to future investigation.

The rest of this paper is organized as follows. Section 2 reviews the simplifying model of software reliability testing that was proposed in our previous work [5]. Section 3 calculates the expected number of observed software failures for the simplifying model. Section 4 shows that the corresponding software reliability testing process can be modeled as a linear dynamic system. Section 5 analyzes the bounds and trends of the expected number of observed software failures. The variance of the number of software failures is examined in Section 6. Illustrative examples are presented in Section 7 to demonstrate the dynamic behavior of software testing. Concluding remarks are contained in Section 8.

2. Model assumptions

Suppose the input domain or the test suite, denoted C, of the software under test is partitioned into m disjoint or non-disjoint classes, $C_1, C_2, \ldots, C_m$, $C = \bigcup_{j=1}^{m} C_j$. In the Markov usage model based testing [26,27], test cases are selected from the test suite or $C_1, C_2, \ldots, C_m$ one by one and applied to the software under test, following the assumption that the selection transitions among the different classes $C_1, C_2, \ldots, C_m$ are governed by a Markov chain. This testing approach has drawn a lot of attention in recent years [13,22]. However several drawbacks are associated with it. First, software defect removal mechanisms are not explicitly considered. Software defects may be removed immediately upon being detected. They may remain un-removed until a given number of tests are executed or a given number of defects are revealed. They may also remain un-removed during testing. Different removal mechanisms have different impacts on software reliability behavior and this has extensively been examined in software reliability modeling [3,20,28]. Second, two commonly adopted reliability measures, the times between successive observed failures and the total number of observed failures up to given instants of time, are seldom analyzed in a mathematically rigorous manner. Third, the continuous-time domain is avoided. Test cases are selected and executed one by one and it is true that the effects of software testing on software reliability can be modeled in the discrete-time domain. However executions of different test cases take different CUP or calendar times and it should be more appropriate to combine the discrete-time domain with the continuous-time domain. The software reliability behavior should be examined in the discrete-time domain as well as in the continuous-time domain.

In order to fit the requirement of software reliability testing or to quantify the effects of software testing on software reliability improvements, the Markov usage model based testing should be extended in at least two aspects. First, the continuous-time properties should be taken into account in the test case selection process. The corresponding Markov chains may be extended to Markov processes. Second, the software defect removal mechanism should be given. Consequently, a set of simplifying assumptions was proposed for software reliability testing in our previous work [5]. It is assumed that the number of test cases selected and applied follows a Poisson process. Further, one and only one failure-causing defect is removed upon a failure being observed.

More specifically, the simplifying assumptions are as follows [5]:

(1) The input domain or the given test suite, C, of the software under test comprises m classes of test cases, $C_1, C_2, \ldots, C_m$, which may or may not be disjoint; that is, $C = \bigcup_{j=1}^{m} C_j$; $C_1, C_2, \ldots, C_m$ do not change in the course of software testing.

(2) The software under test contains N defects at the beginning of testing.

(3) Each test case picked up by an action or from a class may or may not reveal a failure; let

$$Z_i = \begin{cases} 1 & \text{if the } i\text{th action } A_i \text{ reveals a failure} \\ 0 & \text{otherwise} \end{cases}$$

(4) Upon a failure being revealed, the execution of the current test case terminates; one and only one failure-causing defect is removed immediately from the software under test, and no new defects are introduced.

(5) A next test case is selected and executed after the current action is finished; the sequence $\{A_1, A_2, \ldots, A_i, A_{i+1}, \ldots\}$ forms a Markov chain with

$$\Pr\{A_{i+1} = l | A_i = k\} = p_{kl}$$

(6) During the time interval $[0, t)$ a total of $H(t) + 1$ test cases are selected; the first one is taken at the beginning of software testing and $H(t)$ forms a Poisson process with parameter λ, or

$$\Pr\{H(t) = k\} = \frac{(\lambda t)^k}{k!} e^{\lambda t}; \quad k = 0, 1, 2, \ldots$$

where λ is referred to as the testing intensity; each testing action including the first one takes an exponentially distributed length of time with parameter λ.

(7) The first i actions or test cases detect M_i defects and the testing process during the time interval $[0, t]$ detects $M(t)$ defects; that is,

$$M_i = \sum_{k=1}^{i} Z_k \quad \text{with } M_0 = 0$$

$$M(t) = \sum_{k=0}^{H(t)} Z_k \quad \text{with } M(0) = 0, \quad Z_0 = 0$$

(8) The probability of a test case revealing a failure is proportional to the number of defects remaining in the software under test,

$$\Pr\{Z_i = 1 | A_i = j, M_{i-1} = k\} = (N - k)\theta_j$$

$$\Pr\{Z_i = 0 | A_i = j, M_{i-1} = k\} = 1 - (N - k)\theta_j; \quad i = 1, 2, \ldots; \quad k = 0, 1, \ldots, N - 1$$

(9) $\{M_1, M_2, \ldots\}$ and $\{A_1, A_2, \ldots\}$ are conditionally independent of each other as follows:

$$\Pr\{M_1, A_2 | A_1\} = \Pr\{M_1 | A_1\} \Pr\{A_2 | A_1\}$$

and for $i > 1$,

$$\Pr\{M_i, A_{i+1} | M_{i-1}, \ldots, M_1, A_i, A_{i-1}, \ldots, A_1\}$$
$$= \Pr\{M_i | M_{i-1}, \ldots, M_1, A_i, A_{i-1}, \ldots, A_1\} \Pr\{A_{i+1} | M_{i-1}, \ldots, M_1, A_i, A_{i-1}, \ldots, A_1\}$$

$$\Pr\{M_i | M_{i-1}, \ldots, M_1, A_i, A_{i-1}, \ldots, A_1\} = \Pr\{M_i | M_{i-1}, A_i\}$$

$$\Pr\{A_{i+1} | M_{i-1}, \ldots, M_1, A_i, A_{i-1}, \ldots, A_1\} = \Pr\{A_{i+1} | A_i\}$$

(10) The process $\{M_i; i = 0, 1, 2, \ldots\}$ is independent of the Poisson process $\{H(t), t \geqslant 0\}$; more accurately, it holds

$$\Pr\{M_0 = 0, M_1 = k_1, \ldots, M_i = k_i | H(t) = i\} = \Pr\{M_0 = 0, M_1 = k_1, \ldots, M_i = k_i\}$$

(11) The process $\{A_i; i = 0, 1, 2, \ldots\}$ is independent of the Poisson process $\{H(t), t \geqslant 0\}$; more accurately, it holds

$$\Pr\{A_1 = j_1, A_2 = j_2, \ldots, A_i = j_i | H(t) = i\} = \Pr\{A_1 = j_1, A_2 = j_2, \ldots, A_i = j_i\}$$

(12) The first testing action is selected according to the probability distribution $\{p_1, p_2, \ldots, p_m\}$, that is, $\Pr\{A_1 = j\} = p_j; j = 1, 2, \ldots, m$.

(13) The software testing process terminates till all the N defects are removed.

Here we note that each class comprises a number of test cases. When a test case is selected from a given class, we usually suppose that it is selected from the given class at random. However no specific selection mechanism is further given in the above assumptions. An action actually comprises several subactions including *selection, initialization, execution, termination,* and *check*, and the software testing process can be described as follows. At the beginning of software testing the first action or test case is selected. This is finished instantaneously. Then the required test initialization is conducted. In general, the test initialization for an action may include setting initial state for the software under test, removing a failure-causing defect that was detected by the last action, and so on. The test initialization is finished instantaneously. This is followed by execution of the selected action or test case that takes an exponentially distributed length of time with parameter λ. The execution of the action then terminates. The termination is finished instantaneously. After termination, a test oracle or the tester checks or decides if a failure is revealed. The check is finished instantaneously. This completes the first action. Then the next action starts with another test case being selected, followed by the required test initialization and so on. $A(t) = j$ means two things. First, a test case was selected from C_j prior to time t; second,the selected test case is being conducted at time t.

Here we note that several assumptions presented in the above are not realistic in practice. For example, Assumption (4) states that a failure-causing defect is removed immediately upon the failure being revealed. It also states that one and only one failure-causing defect is removed for each revealed failure without new defects being introduced. In practice, new defects may be introduced while detected defects are removed and this has been taken into account in a number of software reliability models [3]. Assumption (8) states that all the remaining defects are equally detectable. This may not be true for software testing. Despite their drawbacks, the above simplifying assumptions are adopted for two reasons. First, although these assumptions are not most realistic in practice, simulation results presented in our previous work [5] show that the reliability growth behaviors delivered in accordance with these simplifying assumptions coincide with those delivered in accordance with more realistic or generalized assumptions after a short initial period of software testing. The simplifying assumptions really capture some of essential features of the software testing process. Further, Section 4 will show that assumption (8) is not essential for treating the software testing process as a linear dynamic system. Second, as a starting point to investigate the software reliability growth behavior from the perspective of software test case selection in a mathematically rigorous manner, it is reasonable to adopt a set of simplifying assumptions for the sake of mathematical tractability. More realistic assumptions can be adopted to extend or generalize the results presented in this paper in the future investigations.

3. Expected number of observed software failures

In our previous work we showed that $\{M_i, i = 1, 2, \ldots\}$ is not a Markov chain in general [5]. Consequently, $\{M(t); t \geqslant 0\}$, or the cumulative number of observed failures up to time t, is not a Markov process in general. This implies that the NHPP assumption taken in numerous software reliability growth models is not acceptable in theory. In this section we derive the exact formulae for M_i and $M(t)$ in a mathematically rigorous manner. For this purpose, we adopt the following notations.

$$\mathbf{P}=\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & & & \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix} \quad \mathbf{p_0}=(p_1,p_2,\ldots,p_m)$$

$$\Psi=\begin{bmatrix} \theta_1 p_{11} & \theta_2 p_{12} & \cdots & \theta_m p_{1m} \\ \theta_1 p_{21} & \theta_2 p_{22} & \cdots & \theta_m p_{2m} \\ \vdots & & & \\ \theta_1 p_{m1} & \theta_2 p_{m2} & \cdots & \theta_m p_{mm} \end{bmatrix}=\mathbf{P}\Theta \quad \Theta=\begin{bmatrix} \theta_1 & 0 & \cdots & 0 \\ 0 & \theta_2 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \theta_m \end{bmatrix}$$

$$\mathbf{H}=\begin{bmatrix} (1-\theta_1)p_{11} & (1-\theta_2)p_{12} & \cdots & (1-\theta_m)p_{1m} \\ (1-\theta_1)p_{21} & (1-\theta_2)p_{22} & \cdots & (1-\theta_m)p_{2m} \\ \vdots & & & \\ (1-\theta_1)p_{m1} & (1-\theta_2)p_{m2} & \cdots & (1-\theta_m)p_{mm} \end{bmatrix}=\mathbf{P}(\mathbf{I}-\Theta) \quad \mathbf{I}=\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$\alpha(k,i)=E[M_k|A_k=i]\Pr\{A_k=i\}$$

$$\beta(k,i)=\Pr\{A_k=i\}; \quad i=1,2,\ldots,m$$

$$\alpha_k=(\alpha(k,1),\alpha(k,2),\ldots,\alpha(k,m))$$

$$\beta_k=(\beta(k,1),\beta(k,2),\ldots,\beta(k,m))=(p_1,p_2,\ldots,p_m)\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & & & \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}^{k-1}=\mathbf{p_0}\mathbf{P}^{k-1}$$

Recall the assumptions presented in Section 2, we note that $\mathbf{P}$ denotes the probability transition matrix of the test case selection process, and $\mathbf{p_0}$ is the initial probability distribution that describes how probably the first test case is selected from one of the m classes. $\beta(k,i)$ represents the probability that the kth test case is selected from the ith classes. However the physical interpretations for other mathematical notions are not obvious.

Proposition 3.1. *Under the assumptions of Section 2, it holds that*

$$\alpha_{k+1}=\alpha_k\mathbf{H}+N\beta_k\Psi=\alpha_k\mathbf{P}(\mathbf{I}-\Theta)+N\mathbf{p_0}\mathbf{P}^k\Theta=N\sum_{j=0}^{k}\mathbf{p_0}\mathbf{P}^{k-j}\Theta(\mathbf{P}(\mathbf{I}-\Theta))^j; \quad k\geqslant 0$$

Proof. See Appendix. $\square$

Proposition 3.1 suggests that the row vector α_k can be calculated in a recursive way. This is important to obtain a closed-form expression for the expected cumulative number of observed failures, as shown in the following proposition.

Proposition 3.2. *Under the assumptions of Section 2, it holds that*

$$EM_k=N\mathbf{p_0}(\mathbf{I}-(\mathbf{I}-\Theta)[\mathbf{P}(\mathbf{I}-\Theta)]^{k-1})\bullet\mathbf{1}^{\mathrm{T}}$$

$$EM(t)=N-N\mathbf{p_0}\,\mathrm{e}^{-\lambda t[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]}\bullet\mathbf{1}^{\mathrm{T}}$$

where $\bullet$ denotes inner product of two vectors, and $\mathbf{1}$ denotes the row vector of each entry being 1 of appropriate dimension, with $\mathbf{1}^{\mathrm{T}}$ being the transpose of $\mathbf{1}$.

Proof. See Appendix. □

The above proposition demonstrates that the expected cumulative number of observed failures grows in an exponential way. This implies that there may be an intrinsic connection between the software testing process and a linear dynamic system.

4. Linear dynamic systems

4.1. Linear models

A linear dynamic system can be induced from Proposition 3.2 as follows: Let

$$\mathbf{K} = -\lambda[\mathbf{I} - (\mathbf{I} - \Theta)\mathbf{P}]$$

$$\boldsymbol{X} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}$$

Then $e^{-\lambda t[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]}\bullet\, \mathbf{1}^{\mathrm{T}}$ is the solution to the following system of linear differential equations with the initial state $\mathbf{X}(0) = \mathbf{1}^{\mathrm{T}}$,

$$\dot{\boldsymbol{X}}(t) = \mathbf{K}\boldsymbol{X}(t)$$

From Proposition 3.2 we have

$$E\left[\frac{N - M(t)}{N}\right] = (p_1, p_2, \ldots, p_m)\begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_m(t) \end{pmatrix}$$

Note that $E\left[\frac{N-M(t)}{N}\right]$ represents the expected ratio of failures that have not been revealed by time t. $x_i(t)$ should be interpreted as the expected ratio of failures that have not been revealed by an arbitrary sequence of actions starting with action i by time t. There are two kinds of probabilistic uncertainty associated with a sequence of actions. First, it is not certain how many actions are applied during the time interval $(0, t)$. Second, it is not certain what actions the sequence comprises and in what order these actions are applied.

Overall, we obtain a linear system for the software testing process defined in Section 2,

$$\dot{\boldsymbol{X}}(t) = \mathbf{K}\boldsymbol{X}(t) \quad \text{with } \boldsymbol{X}(0) = \mathbf{1}^{\mathrm{T}}$$
$$v(t) = \mathbf{c}\mathbf{X}(t)$$

where

$$\mathbf{c} = \mathbf{p}_0 = (p_1, p_2, \ldots, p_m)$$
$$v(t) = E\left[\frac{N - M(t)}{N}\right]$$

Accordingly, the software testing process defined in Section 2 can be classified as linear, and it is reasonable to expect there are non-linear software testing processes too. A notable feature of the linear dynamic system is that the corresponding states are not defined a priori. However, they are interpreted in a reasonable manner as explained in the last paragraph. There are unexpected and intrinsic links between software testing and system dynamics, which have not been revealed in the literature. The theory of linear or non-linear control [10,18] may be adopted to guide or improve the software testing processes by adjusting the matrix $\mathbf{K}$ on-line. Without extra control, the testing process defined in Section 2 follows $\dot{\boldsymbol{X}}(t) = \mathbf{K}\boldsymbol{X}(t)$ to evolve. The matrix $\mathbf{K}$ determines the nature of the dynamic behavior of the software testing process. A noticeable feature of $\mathbf{K}$ is that it is independent of the parameter N. This implies that the initial number of software defects is not an intrinsic factor for characterizing the dynamics of the software testing process. The extra control effect can take place as $\mathbf{K}$ is adjusted on-line by updating parameters λ, Θ and/or $\mathbf{P}$. The corresponding control problem may be formulated in the setting of model predictive control [12]. We leave this problem to future investigations.

Here it is important to identify the difference between the above linear model and those models proposed in Refs. [19,29]. Refs. [19,29] *assign* a linear stochastic differential equation *directly* to the number of remaining defects. No test case selection mechanisms are involved. It is not clear under what test case selection mechanisms the stochastic differential equation is valid. On the other hand, the linear model presented above is *derived* from the process of test case selection in a mathematically rigorous manner. The test case selection mechanism is described as a Markovian chain, and the linear model is theoretically valid under the set of simplifying assumptions presented in Section 2. Another related model is that of Congussu et al. [7,8], which is concerned with the management aspects of software testing. However the linear model presented in this paper is concerned with the technological aspects of software testing.

Software testing processes other than the one defined in Section 2 can also induce a linear dynamic system. Suppose that at the beginning of software testing the software under test contains N defects that can be divided into two classes: N_1 defects, each of which has detection rate of $\theta_i^{(1)}$ by action A_i; and N_2 defects, each of which has detection rate of $\theta_i^{(2)}$ by action A_i. Suppose there are two observers. The first one observes those failures triggered by the first class of defects but is not capable of observing those failures triggered by the second class of defects. The number of failures observed by the first observer up to time t is $M^{(1)}(t)$. The second one observes those failures triggered by the second class of defects but is not capable of observing those failures triggered by the first class of defects. The numbers of failures observed by the first and second observers up to time t are $M^{(1)}(t)$ and $M^{(2)}(t)$, respectively. If a failure is triggered by one defect of the first class as well as by one defect of the second class, then one and only one failure-causing defect of the first class is removed (refer to Assumption (4) of Section 2). Consequently, either $M^{(1)}(t)$ or $M^{(2)}(t)$ is increased by one, but it is impossible that both $M^{(1)}(t)$ and $M^{(2)}(t)$ are increased simultaneously. Then $M(t) = M^{(1)}(t) + M^{(2)}(t)$ is the total number of failures observed up to time t. Note that the behavior of $M^{(1)}(t)$ induces the following linear system

$$\dot{X}^{(1)}(t) = \mathbf{K}^{(1)}X^{(1)}(t) \quad \text{with } \mathbf{K}^{(1)} = -\lambda[\mathbf{I} - (\mathbf{I} - \Theta^{(1)})\mathbf{P}]$$

whereas the behavior of $M^{(2)}(t)$ induces the following linear system

$$\dot{X}^{(2)}(t) = \mathbf{K}^{(2)}X^{(2)}(t) \quad \text{with } \mathbf{K}^{(2)} = -\lambda[\mathbf{I} - (\mathbf{I} - \Theta^{(2)})\mathbf{P}]$$

where

$$\Theta^{(1)} = \begin{bmatrix} \theta_1^{(1)} & 0 & \cdots & 0 \\ 0 & \theta_2^{(1)} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \theta_m^{(1)} \end{bmatrix} \quad \Theta^{(2)} = \begin{bmatrix} \theta_1^{(2)} & 0 & \cdots & 0 \\ 0 & \theta_2^{(2)} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \theta_m^{(2)} \end{bmatrix}$$

In this way an aggregate linear dynamic system can be obtained to describe the behavior of $M(t)$

$$\begin{bmatrix} \dot{\mathbf{X}}^{(1)}(t) \\ \dot{\mathbf{X}}^{(2)}(t) \end{bmatrix} = \begin{bmatrix} \mathbf{K}^{(1)} & 0 \\ 0 & \mathbf{K}^{(2)} \end{bmatrix} \begin{bmatrix} \mathbf{X}^{(1)}(t) \\ \mathbf{X}^{(2)}(t) \end{bmatrix}$$

Similarly, the N defects can be divided into more than two classes and a corresponding linear dynamic system can be induced too. It is *not* essential to assume that defects are equally detectable to define a linear testing process. At most $N \times m$ state variables are sufficient to characterize the induced linear dynamic system of the testing process with m distinct actions for the software under test that contains N defects at the beginning of software testing. Recall that the initial number of software defects, N, is not important for characterizing the dynamics of the software testing process if all the N defects are equally detectable. However the matrix

$$\begin{bmatrix} \mathbf{K}^{(1)} & 0 \\ 0 & \mathbf{K}^{(2)} \end{bmatrix}$$

does reflect the assumption that the software contains two distinct classes of defects, each of which comprises equally detectable defects. Therefore, it is *not* the initial number of software defects, but the number of distinct classes of equally detectable defects that defines an intrinsic parameter of the dynamics of the software testing process. This implies the number of distinct classes of equally detectable defects can serve as a control variable. The defect injection mechanisms can affect the dynamics of the software testing process since they may change the number of distinct classes of equally detectable defects. On the other hand, different test suite may lead to different classifications of the software defects.

4.1.1. Special cases

This subsection considers two special cases of $\mathbf{P}$. First, let $\mathbf{P} = \mathbf{I}$. That is, if an action is selected, then it is always selected and no other actions are selected during the testing process. The probability distribution $\{p_1, p_2, \ldots, p_m\}$ determines which action is selected at the beginning of software testing. Then

$$\mathbf{K} = -\lambda[\mathbf{I} - (\mathbf{I} - \Theta)\mathbf{P}] = -\lambda\Theta = -\lambda\,\mathrm{diag}[\theta_1, \theta_2, \ldots, \theta_m]$$
$$\mathbf{X}(t) = [\mathrm{e}^{-\lambda\theta_1 t}, \mathrm{e}^{-\lambda\theta_2 t}, \ldots, \mathrm{e}^{-\lambda\theta_m t}]^{\mathrm{T}}$$
$$v(t) = \sum_{i=1}^{m} p_i \mathrm{e}^{-\lambda\theta_i t}$$

That is, the state variables or the expected ratios of failures that have not been revealed by an arbitrary sequence of actions starting with various distinct actions by time t decay exponentially as software testing proceeds.

For the second special case, let

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 1 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

That is, during software testing, all actions are consecutively selected. Suppose the current test case is selected from C_k, $k < m$, then the next test case is selected from C_{k+1}. If the current test case comes from C_m, then the next test case belongs to C_1. We have

$$\mathbf{K} = -\lambda \begin{bmatrix} 1 & -(1-\theta_1) & 0 & 0 & \cdots & 0 \\ 0 & 1 & -(1-\theta_2) & 0 & \cdots & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & 0 & \cdots & -(1-\theta_{m-1}) \\ -(1-\theta_m) & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

The m eigenvalues of $\mathbf{K}$ are

$$\sigma(\mathbf{K})_k = -\lambda + \lambda\left[\prod_{j=1}^{m}(1-\theta_j)\right]^{\frac{1}{m}}\left[\cos\left(\frac{2k\pi}{m}\right) + \mathrm{img}\left(\sin\left(\frac{2k\pi}{m}\right)\right)\right]; \quad k = 1, \ldots, m$$

where *img* denotes the imaginary part of a complex variable. This implies that, if m is odd, then $\mathbf{K}$ has one and only one real eigenvalue $-\lambda + \lambda\left[\prod_{j=1}^{m}(1-\theta_j)\right]^{\frac{1}{m}}$. If m is even, then $\mathbf{K}$ has two and only two real eigenvalues $-\lambda + \lambda\left[\prod_{j=1}^{m}(1-\theta_j)\right]^{\frac{1}{m}}$ and $-\lambda - \lambda\left[\prod_{j=1}^{m}(1-\theta_j)\right]^{\frac{1}{m}}$. Therefore for all m, $\mathbf{X}(t)$ approaches zero in an exponential way as $t \to \infty$.

4.1.2. Eigenvalues and dynamic behavior

The rest of this paper is still stuck to the software testing process defined in Section 2. Since $x_i(t)$ is interpreted as the expected ratio of failures that have not been revealed by an arbitrary sequence of actions starting with action i by time t, a natural question is whether $x_i(t) \to 0$ as $t \to \infty$. The answer is affirmative and can be formulated in a mathematically rigorous manner. For the matrix $\mathbf{K}$, according to the theory of matrix [15], there exists a similarity transformation $\mathbf{T}$ such that

$$\mathbf{TKT}^{-1} = \lambda \,\mathrm{diag}\left[\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_s, \widetilde{\mathbf{J}}_1, \widetilde{\mathbf{J}}_2, \ldots, \widetilde{\mathbf{J}}_r\right] = \lambda \mathbf{J}$$

where

$$\mathbf{J}_i = \begin{bmatrix} \mu_i & 1 & 0 & \cdots & 0 \\ 0 & \mu_i & 1 & \cdots & 0 \\ & \ddots & \ddots & & \\ 0 & & \ddots & & 1 \\ 0 & & & & \mu_i \end{bmatrix}; \quad i = 1, \ldots, s$$

$$\widetilde{\mathbf{J}}_i = \begin{bmatrix} \mathbf{D}_i & \mathbf{I}_2 & 0 & \cdots & 0 \\ 0 & \mathbf{D}_i & \mathbf{I}_2 & \cdots & 0 \\ & \ddots & \ddots & & \\ 0 & & \ddots & \mathbf{I}_2 & \\ 0 & & & & \mathbf{D}_i \end{bmatrix}; \quad i = 1, \ldots, r$$

$$\mathbf{D}_i = \begin{bmatrix} \delta_i & \varepsilon_i \\ -\varepsilon_i & \delta_i \end{bmatrix} \quad \mathbf{I}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

and δ_i and ε_i are constants uniquely determined by the matrix $\mathbf{K}/\lambda$. $\mathbf{T}$ transforms $\mathbf{K}$ into its Jordan form.

Note that $\mathbf{T}^{-1}\mathbf{T} = \mathbf{TT}^{-1} = \mathbf{I}$, we have

$$\lambda \mathbf{J} = \mathbf{TKT}^{-1} = \mathbf{T}(\lambda[(\mathbf{I} - \Theta)\mathbf{P} - \mathbf{I}])\mathbf{T}^{-1} = \lambda[\mathbf{T}(\mathbf{I} - \Theta)\mathbf{PT}^{-1} - \mathbf{I}]$$

Therefore the matrix $\mathbf{T}$ also transforms $(\mathbf{I} - \Theta)$ $\mathbf{P}$ into its Jordan form. That is,

$$\mathbf{T}(\mathbf{I} - \Theta)\mathbf{PT}^{-1} = \mathrm{diag}\left[\mathbf{J}'_1, \mathbf{J}'_2, \ldots, \mathbf{J}'_s, \widetilde{\mathbf{J}}'_1, \widetilde{\mathbf{J}}'_2, \ldots, \widetilde{\mathbf{J}}'_r\right] = \mathbf{J}'$$

where

$$\mathbf{J}'_i = \begin{bmatrix} \mu_i + 1 & 1 & 0 & \cdots & 0 \\ 0 & \mu_i + 1 & 1 & \cdots & 0 \\ & \ddots & \ddots & & \\ 0 & & & \ddots & 1 \\ 0 & & & & \mu_i + 1 \end{bmatrix}; \quad i = 1, \ldots, s$$

$$\widetilde{\mathbf{J}}'_i = \begin{bmatrix} \mathbf{D}_i + \mathbf{I}_2 & \mathbf{I}_2 & 0 & \cdots & 0 \\ 0 & \mathbf{D}_i + \mathbf{I}_2 & \mathbf{I}_2 & \cdots & 0 \\ & & \ddots & \ddots & \\ 0 & & & \ddots & \mathbf{I}_2 \\ 0 & & & & \mathbf{D}_i + \mathbf{I}_2 \end{bmatrix}; \quad i = 1, \ldots, r$$

as a result of $\mathbf{J} = \mathbf{J}' - \mathbf{I}$.

Lemma 4.1. *Let* $\mathbf{K} = -\lambda[\mathbf{I} - (\mathbf{I} - \Theta)\mathbf{P}]$. *It holds that*

(1) *The following statements are equivalent*:
  (a) υ *is an eigenvector of* $\mathbf{K}$;
  (b) υ *is an eigenvector of* $\mathbf{K}/\lambda$;
  (c) υ *is an eigenvector of* $(\mathbf{I} - \Theta)\mathbf{P}$.
(2) *The following statements are equivalent*:
  (a) μ *is an eigenvalue of* $\mathbf{K}$;
  (b) μ/λ *is an eigenvalue of* $\mathbf{K}/\lambda$;
  (c) $\mu/\lambda + 1$ *is an eigenvalue of* $(\mathbf{I} - \Theta)\mathbf{P}$.
(3) *If* μ *is an eigenvalue of* $\mathbf{K}$, *then* e^{μ} *is an eigenvalue of* $e^{\mathbf{K}}$.
(4) *If* μ *is an eigenvalue of* $(\mathbf{I} - \Theta)\mathbf{P}$, *then* e^{μ} *is an eigenvalue of* $e^{(\mathbf{I}-\Theta)\mathbf{P}}$.

Proof. (1) and (2) are trivial. Here we only show the validity of (3). The validity of (4) can be proved similarly. Suppose υ be an eigenvector of $\mathbf{K}$ with respect to μ. Then

$$\lambda[(\mathbf{I} - \Theta)\mathbf{P} - \mathbf{I}]\upsilon = \mu\upsilon$$

$$e^{\lambda[(\mathbf{I}-\Theta)\mathbf{P}-\mathbf{I}]}\upsilon = \sum_{n=0}^{\infty} \frac{1}{n!}[\lambda[(\mathbf{I} - \Theta)\mathbf{P} - \mathbf{I}]]^n\upsilon = \sum_{n=0}^{\infty} \frac{1}{n!}\mu^n\upsilon = e^{\mu}\upsilon. \qquad \square$$

From Lemma 4.1 we see that the eigenvalues and eigenvectors of $\mathbf{K}$ can be represented in terms of the eigenvalues and eigenvectors of $(\mathbf{I} - \Theta)\mathbf{P}$, respectively.

Lemma 4.2. *Let* μ *be the maximal eigenvalue of* $(\mathbf{I} - \Theta)\mathbf{P}$. *Then* μ *is real and*

$$1 - \max_{1\leqslant i\leqslant m} \sum_{j=1}^{m} \theta_j p_{ij} = \min_{1\leqslant i\leqslant m}\left(1 - \sum_{j=1}^{m} \theta_j p_{ij}\right) \leqslant \mu \leqslant \max_{1\leqslant i\leqslant m}\left(1 - \sum_{j=1}^{m} \theta_j p_{ij}\right) = 1 - \min_{1\leqslant i\leqslant m} \sum_{j=1}^{m} \theta_j p_{ij}$$

Proof. See Appendix. $\square$

The above proposition gives the lower and upper bounds for μ. Here we should note that

(1) $\sum_{j=1}^{m}\theta_j p_{ij}$ is the average defect detection rate;
(2) The maximal eigenvalue of $(\mathbf{I} - \Theta)\mathbf{P}$ is less than 1; and
(3) The maximal eigenvalue of $\mathbf{K} = -\lambda[\mathbf{I} - (\mathbf{I} - \Theta)\mathbf{P}]$ lies in $\left[-\max_{1\leqslant i\leqslant m}\sum_{j=1}^{m}\theta_j p_{ij}, -\min_{1\leqslant i\leqslant m}\sum_{j=1}^{m}\theta_j p_{ij}\right]$.

Proposition 4.1. *For the induced linear system of the software testing process defined in Section 2, it holds that* $\mathbf{X}(t) \to 0$ *exponentially as* $t \to \infty$.

Proof. See Appendix. $\square$

The above proposition implies that the ratio of failures that have not been revealed eventually approach zero as software testing proceeds. In other words, all defects remaining in the software under test will be removed eventually. This conclusion is intuitively obvious. However no mathematically rigorous proof was presented in the existing literature.

5. Analysis of the expected number of observed software failures

This section is concerned with the bounds and evolving trends of the expected number of the observed software failures. Recall that

$$EM_k = \alpha_k \bullet \mathbf{1}^{\mathrm{T}} = N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^j \bullet \mathbf{1}^{\mathrm{T}}$$

For the case of $\theta_1 = \theta_2 = \cdots = \theta_m = \theta \neq 0$, we have

$$EM_k = N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^j \bullet \mathbf{1}^{\mathrm{T}} = N\sum_{j=0}^{k-1}(1-\theta)^j\theta\mathbf{p_0}\mathbf{P}^{k-1} \bullet \mathbf{1}^{\mathrm{T}} = N[1-(1-\theta)^k]$$

The following two propositions determine the lower and upper bounds of the expected number of the observed software failures.

Proposition 5.1. *Suppose $\theta_1 \geqslant \theta_2 \geqslant \cdots \geqslant \theta_{m-1} \geqslant \theta_m$ Then it holds that*

$$\frac{\theta_m}{\theta_1}N(1-(1-\theta_1)^k) \leqslant EM_k \leqslant \frac{\theta_1}{\theta_m}N(1-(1-\theta_m)^k)$$

And

$$\frac{\theta_m}{\theta_1}N(1-\mathrm{e}^{-\theta_1\lambda t}) \leqslant EM(t) \leqslant \frac{\theta_1}{\theta_m}N(1-\mathrm{e}^{-\theta_m\lambda t})$$

Proof. See Appendix. □

Here we note that the term $N(1-(1-\theta_m)^k)$ is just EM_k in the case of $\theta_1 = \theta_2 = \cdots = \theta_m$. The lower and upper bounds given in the above proposition are intuitively understandable. However the following proposition presents tighter bounds for EM_k and $EM(t)$. This is because $\theta_m \leqslant \sum_{l=1}^{m}\theta_l p_l \leqslant \theta_1$.

Proposition 5.2. *Suppose $\theta_1 \geqslant \theta_2 \geqslant \cdots \geqslant \theta_{m-1} \geqslant \theta_m$. Then it holds*

$$N - N\left(1-\sum_{l=1}^{m}\theta_l p_l\right)(1-\theta_m)^{k-1} \leqslant EM_k \leqslant N - N\left(1-\sum_{l=1}^{m}\theta_l p_l\right)(1-\theta_1)^{k-1}\frac{\mathrm{e}^{-\theta_1\lambda t}}{1-\theta_1}$$
$$\leqslant \frac{N-EM(t)}{N\left(1-\sum_{l=1}^{m}p_l\theta_l\right)} \leqslant \frac{\mathrm{e}^{-\theta_m\lambda t}}{1-\theta_m}$$

Proof. See Appendix. □

In order to analyze the evolving trend of $EM(t)$, denote

$$EM(t) = g(t)$$

Proposition 5.3. *Suppose that* $\mathbf{p_0}\mathbf{P} = \mathbf{p_0}$. *Then*

$$g''(t) < 0$$

Proof. See Appendix. □

The above proposition demonstrates that if the software testing begins with the equilibrium distribution as its initial distribution, then the growth rate of the expected number of observed software failures, determined by $\frac{\mathrm{d}EM(t)}{\mathrm{d}t}$, will monotonically decrease with the testing time t. Note that $\frac{\mathrm{d}EM(t)}{\mathrm{d}t}$ has seldom been investigated in the existing literature.

Proposition 5.3 is concerned with the case of equilibrium distribution. In the case that the initial distribution is not the equilibrium distribution, the behavior of $\frac{\mathrm{d}EM(t)}{\mathrm{d}t}$ is not theoretically obvious. However we can consider a special case as shown in the following proposition.

Proposition 5.4. *Suppose that $\theta_1 = \theta_2 = \cdots = \theta_m = \theta \neq 0$. Then*

$$g''(t) < 0$$

Proof. See Appendix. □

The above proposition demonstrates that if all the m actions are equally capable of detecting defects, then $\frac{\mathrm{d}EM(t)}{\mathrm{d}t}$ will monotonically decrease with the testing time t. This special case can be treated as the case that there is only one single action in total.

6. Analysis of the variance of the number of observed software failures

The variance of the number of observed software failures, $E[M(t) - EM(t)]^2$, is seldom analyzed in the literature. This might be due to that the behavior of the variance is hard to be analyzed. However the variance can be treated as a measure of the stability of the software testing process and thus should draw our attention. Ideally, the variance should be as small as possible, which implies that the testing process tends to be stable.

Proposition 6.1. *It holds that*

$$0 \leqslant E(M(t))^2 - (EM(t))^2 \leqslant 2N^2\mathbf{p}_0\,\mathrm{e}^{-\lambda[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]t} \bullet \mathbf{1}^{\mathrm{T}}\Pr\{M(t) \geqslant EM(t)\} \leqslant 2N^2\mathbf{p}_0\,\mathrm{e}^{-\lambda[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]t} \bullet \mathbf{1}^{\mathrm{T}}$$

Proof. See Appendix. □

The above proposition implies that, in general, the variance of the number of observed number of failure is upper bounded by a convex combination of exponential functions and approaches zero eventually. Since the mean of the number of observed failures tends to grow steadily as software testing proceeds, the above proposition also confirms that the NHPP assumption that has been extensively adopted in software reliability modeling (refer to Section 1) is theoretically false.

Remark

(1) From the above proposition we have

$$0 \leqslant E\left[\frac{M(t) - EM(t)}{N}\right]^2 \leqslant 2E\left[\frac{N - M(t)}{N}\right]\Pr\{M(t) \geqslant EM(t)\}$$

(2) If $\theta_1 \geqslant \theta_2 \geqslant \cdots \geqslant \theta_m$, then

$$\begin{aligned} E\left[\frac{M(t) - EM(t)}{N}\right]^2 &\leqslant 2\mathbf{p}_0\mathrm{e}^{-\lambda[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]t} \bullet \mathbf{1}^{\mathrm{T}} = 2\mathbf{p}_0\,\mathrm{e}^{-\lambda\mathbf{I}}\,\mathrm{e}^{\lambda(\mathbf{I}-\Theta)\mathbf{P}t} \bullet \mathbf{1}^{\mathrm{T}} \leqslant 2\mathbf{p}_0\,\mathrm{e}^{-\lambda\mathbf{I}}\,\mathrm{e}^{\lambda(1-\vartheta_m)\mathbf{P}t} \bullet \mathbf{1}^{\mathrm{T}} \\ &= 2\mathbf{p}_0\,\mathrm{e}^{-\lambda[\mathbf{I}-(1-\theta_m)\mathbf{P}]t} \bullet \mathbf{1}^{\mathrm{T}} = 2\,\mathrm{e}^{-\theta_m t} \end{aligned}$$

Proposition 6.2. *For the case of* $\theta_1 = \theta_2 = \cdots = \theta_m = \theta < \frac{1}{2}$, *it holds that*

$$E(M(t))^2 - (EM(t))^2 = N[\mathrm{e}^{-\lambda\theta t} - \mathrm{e}^{-2\lambda\theta t}]$$

Proof. See Appendix. □

That is, for the special case $\theta_1 = \theta_2 = \cdots = \theta_m = \theta < \frac{1}{2}$, the variance of the number of observed failures tends to decay in an exponential manner.

Proposition 6.3. *For the software testing process defined in Section 2, it holds that*

$$\begin{aligned} E[M_k^2|A_k = i]\Pr\{A_k = i\} = \sum_{l=1}^{m}&[N\theta_i p_{li} + (2N-1)\theta_i p_{li}E[M_{k-1}|A_{k-1} = l] \\ &+ (1 - 2\theta_i)p_{li}E[M_{k-1}^2|A_{k-1} = l]]\Pr\{A_{k-1} = l\} \end{aligned}$$

For all $k \geqslant 1$.

Proof. See Appendix. □

The above formula can be simplified in notation. Let

$$\gamma(k,i) = E[M_k^2|A_k = i]\Pr\{A_k = i\}; \quad i = 1,2,\ldots,m; \quad k \geqslant 1$$

For $k = 1$, $\gamma(1,i) = E[M_1^2|A_1 = i]\Pr\{A_1 = i\} = E[M_1|A_1 = i]\Pr\{A_1 = i\} = \alpha(1,i)$. For $k = 2$,

$$\begin{aligned}
E[M_2^2|A_2 = i]\Pr\{A_2 = i\} &= \sum_{l=1}^{m}[N\theta_i p_{li} + (2N-1)\theta_i p_{li} E[M_1|A_1 = l] + (1-2\theta_i)p_{li}E[M_1^2|A_1 = l]]\Pr\{A_1 = l\} \\
&= \sum_{l=1}^{m}[N\theta_i p_{li} + (2N-1)\theta_i p_{li} E[M_1|A_1 = l] + (1-2\theta_i)p_{li}E[M_1|A_1 = l]]\Pr\{A_1 = l\} \\
&= \sum_{l=1}^{m} N\theta_i p_{li}\Pr\{A_1 = l\} + \sum_{l=1}^{m}[(2N-1)\theta_i p_{li} + (1-2\theta_i)p_{li}]\alpha(1,l)
\end{aligned}$$

We have for $k \geqslant 2$

$$\begin{aligned}
(\gamma(k,1),\gamma(k,2),\ldots,\gamma(k,m)) &= (\beta(k-1,1),\beta(k-1,2),\ldots,\beta(k-1,m))\begin{bmatrix} N\theta_1 p_{11}, N\theta_2 p_{12},\ldots,N\theta_m p_{1m} \\ N\theta_1 p_{21}, N\theta_2 p_{22},\ldots,N\theta_m p_{2m} \\ \vdots \\ N\theta_1 p_{m1}, N\theta_2 p_{m2},\ldots,N\theta_m p_{mm}\end{bmatrix} \\
&+ (2N-1)(\alpha(k-1,1),\alpha(k-1,2),\ldots,\alpha(k-1,m))\begin{bmatrix} \theta_1 p_{11}, \theta_2 p_{12},\ldots,\theta_m p_{1m} \\ \theta_1 p_{21}, \theta_2 p_{22},\ldots,\theta_m p_{2m} \\ \vdots \\ \theta_1 p_{m1}, \theta_2 p_{m2},\ldots,\theta_m p_{mm}\end{bmatrix} \\
&+ (\gamma(k-1,1),\gamma(k-1,2),\ldots,\gamma(k-1,m))\begin{bmatrix} (1-2\theta_1)p_{11},(1-2\theta_2)p_{12},\ldots,(1-2\theta_m)p_{1m} \\ (1-2\theta_1)p_{21},(1-2\theta_2)p_{22},\ldots,(1-2\theta_m)p_{2m} \\ \vdots \\ (1-2\theta_1)p_{m1},(1-2\theta_2)p_{m2},\ldots,(1-2\theta_m)p_{mm}\end{bmatrix}
\end{aligned}$$

Further let

$$\gamma_k = (\gamma(k,1),\gamma(k,2),\ldots,\gamma(k,m))$$

$$\mathbf{H}_{\langle 2\rangle} = \begin{bmatrix} (1-2\theta_1)p_{11} & (1-2\theta_2)p_{12} & \cdots & (1-2\theta_m)p_{1m} \\ (1-2\theta_1)p_{21} & (1-2\theta_2)p_{22} & \cdots & (1-2\theta_m)p_{2m} \\ \vdots & & & \\ (1-2\theta_1)p_{m1} & (1-2\theta_2)p_{m2} & \cdots & (1-2\theta_m)p_{mm}\end{bmatrix} = \mathbf{P}(\mathbf{I} - 2\Theta)$$

We obtain

$$\gamma_k = N\beta_{k-1}\boldsymbol{\Psi} + (2N-1)\alpha_{k-1}\boldsymbol{\Psi} + \gamma_{k-1}H_{\langle 2\rangle}$$

EM_k^2 can be determined as follows:

$$EM_k^2 = \sum_{i=1}^{m} E[M_k^2|A_k = i]\Pr\{A_k = i\} = \sum_{i=1}^{m}\gamma(k,i) = \gamma_k \bullet \mathbf{1}^{\mathrm{T}}$$

Proposition 6.4. *For the software testing process defined in Section 2, it holds that*

$$\gamma_k = \sum_{j=1}^{k}\left[N\mathbf{p}_0\mathbf{P}^{k-j}\Theta + (2N-1)N\sum_{l=0}^{k-j-1}\mathbf{p}_0\mathbf{P}^{k-j-1-l}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^l\mathbf{P}\Theta\right][\mathbf{P}(\mathbf{I}-2\Theta)]^{j-1}$$

for $k \geqslant 3$.

Proof. See Appendix. □

The above proposition suggests that the row vector γ_k can be determined in an explicit form. This is important to obtain the following proposition and to examine the behavior of higher moments of the number of observed failures.

Proposition 6.5. *For the software testing process defined in Section 2, it holds that*

$$E[M_k^n, A_k = i] = \sum_{l=1}^{m} N\theta_i p_{li}\Pr\{A_{k-1} = l\} + \sum_{s=1}^{n-1}(NC_n^s - C_n^{s-1})\sum_{l=1}^{m}\theta_i p_{li}E[M_{k-1}^s, A_{k-1} = l]$$
$$+ \sum_{l=1}^{m}(1-n\theta_i)p_{li}E[M_{k-1}^n, A_{k-1} = l]$$

For $k \geqslant N+1$, where $C_n^s = \frac{n!}{s!(n-s)!}$.

Proof. See Appendix. □

7. Illustrative examples

Example 7.1. Suppose

$$\Theta = \begin{bmatrix} \theta_1 & 0 & 0 \\ 0 & \theta_2 & 0 \\ 0 & 0 & \theta_3 \end{bmatrix} = \begin{bmatrix} 0.001 & 0 & 0 \\ 0 & 0.0002 & 0 \\ 0 & 0 & 0.0001 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.3 & 0.2 \\ 0.4 & 0.4 & 0.2 \\ 0.3 & 0.2 & 0.5 \end{bmatrix}$$

$$\mathbf{p}_0 = (p_1 \quad p_2 \quad p_3) = (0.3 \quad 0.4 \quad 0.3)$$

$$\lambda = 0.1 \quad N = 1000$$

That is, we suppose the software under test contains 1000 defects at the beginning of software testing and is subject to three distinct classes of test cases. Class 1 has the highest defect detection rate with the value being 0.001, class 2 has the second highest defect detection rate with the value being 0.0002, and class 3 has the lowest defect detection rate with the value being 0.0001. The first action selects a test case from class 1 with probability 0.3, selects a test case from class 2 with probability 0.4, or selects a test case from class 3 with probability 0.3. After the first test case being executed, for various reasons, the subsequent actions select test cases from the three classes in accordance with a Markovian chain whose transition probability matrix is **P**. One of the reasons is that in practice the exact values of the defect detection rates are not known, otherwise test cases of class 1 (with the highest defect detection rate) should always be selected. Each selected test case takes a random time being executed, and the average time is $1/\lambda = 10$ s. We are interested in how the cumulative number of defects which are detected and removed will grow as software testing proceeds. To this end, we need to obtain the mathematical formulae of $EM(t)$, $\frac{d(EM(t))}{dt}$, and $\mathrm{var}(M(t))$.

We have

$$\mathbf{K} = \begin{bmatrix} -0.050050 & 0.029970 & 0.019980 \\ 0.039992 & -0.060008 & 0.019996 \\ 0.029997 & 0.019998 & -0.050005 \end{bmatrix}$$

The three eigenvalues of $\mathbf{K}$ are

$$\begin{aligned} \lambda(\mathbf{K})_1 &= -0.00005015495102 \\ \lambda(\mathbf{K})_2 &= -0.07000728758326 \\ \lambda(\mathbf{K})_3 &= -0.09000555746572 \end{aligned}$$

All of them are negative reals, which correspond to three distinct actions, respectively. For the linear system $\dot{X}(t) = \mathbf{K}X(t)$ with $X(0) = \mathbf{1}^{\mathrm{T}}$, the state variables are determined as follows:

$$\begin{aligned} x_1(t) &= 0.9994005164\,\mathrm{e}^{-0.00005015495102t} + 0.0002041047200\,\mathrm{e}^{-0.07000728758326t} \\ &\quad + 0.0003953788991\,\mathrm{e}^{-0.09000555746572t} \\ x_2(t) &= 1.000289791\,\mathrm{e}^{-0.00005015495102t} + 0.0002043498961\,\mathrm{e}^{-0.07000728758326t} \\ &\quad - 0.0004941413474\,\mathrm{e}^{-0.09000555746572t} \\ x_3(t) &= 1.000559855\,\mathrm{e}^{-0.00005015495102t} - 0.0005103975466\,\mathrm{e}^{-0.07000728758326t} \\ &\quad - 0.00004945786501\,\mathrm{e}^{-0.09000555746572t} \end{aligned}$$

It should be noticed that each state is dominated by its first term, corresponding to the eigenvalue $\lambda(\mathbf{K})_1 = -0.00005015495102$ or the effect of one of the three actions. Fig. 7.1 shows the behavior of $\mathbf{X}(t)$, which demonstrates a straight-line like trajectory. Note that various values of p_{ij} are at the same order of magnitude, whereas θ_1 is much greater than θ_2 and θ_3, we may argue that it is the action corresponding to θ_1 that dominates the behavior of $\mathbf{X}(t)$. In other words, the test cases of class 1 are most important since they have the highest defect detection rate.

The expected number of observed software failures is expressed

$$\begin{aligned} EM(t) &= 1000 - 1000.104028\,\mathrm{e}^{-0.00005015495102t} + 0.01014788953\,\mathrm{e}^{-0.07000728758326t} \\ &\quad + 0.09388022872\,\mathrm{e}^{-0.09000555746572t} \end{aligned}$$

Obviously, $EM(t)$ is dominated by the eigenvalue $\lambda(\mathbf{K})_1 = -0.00005015495102$ and can be approximated as

$$EM(t) \approx 1000 - 1000.104028\,\mathrm{e}^{-0.00005015495102t} \approx 1000(1 - \mathrm{e}^{-0.00005015495102t})$$

This coincides with the argument that the action corresponding to θ_1 dominates the behavior of $\mathbf{X}(t)$. Fig. 7.2 shows the exact behavior of $EM(t)$. It takes about 2200 s to detect and remove the first 100 defects, 13,900 s to detect and remove the first 500 defects, and 46,000 s to detect and remove the first 900 defects. Much more time is required to detect and remove the remaining 100 defects.

Further, $\frac{\mathrm{d}(EM(t))}{\mathrm{d}t}$ is expressed as

$$\begin{aligned} \frac{\mathrm{d}(EM(t))}{\mathrm{d}t} &= 0.05016016854\,\mathrm{e}^{-0.00005015495102t} - 0.0007104262207\,\mathrm{e}^{-0.07000728758326t} \\ &\quad - 0.008449742321\,\mathrm{e}^{-0.09000555746572t} \end{aligned}$$

Fig. 7.3 shows the behavior of $\frac{\mathrm{d}(EM(t))}{\mathrm{d}t}$. At the very beginning of software testing, $\frac{\mathrm{d}(EM(t))}{\mathrm{d}t}$ grows as software testing proceeds. That is, the cumulative number of defects which are detected and removed grows very fast on average. Soon after the beginning, $\frac{\mathrm{d}(EM(t))}{\mathrm{d}t}$ becomes as a monotonically decreasing function, which eventually approaches zero. In other words, the growth trend of the cumulative number of defects which are detected and removed begins to slow down, and disappears eventually. However if $EM(t) \approx 1000(1 - \mathrm{e}^{-0.000005015495102t})$ were taken, then $\frac{\mathrm{d}(EM(t))}{\mathrm{d}t}$ would be a monotonically decreasing function. Recall that Proposition 5.4 asserts that if all the m actions are equally capable of detecting defects, then $\frac{\mathrm{d}EM(t)}{\mathrm{d}t}$ will monotonically decrease with the testing time t. This implies that it is the effects of the other actions corresponding to θ_2 and θ_3 that make

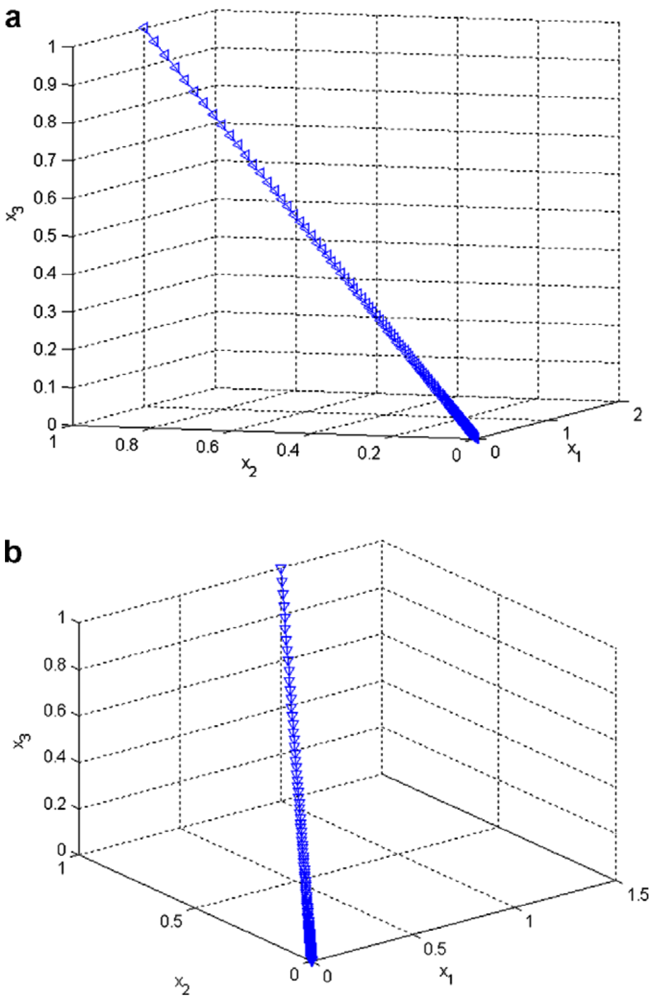

Fig. 7.1. Behavior of the states of the linear dynamic system for Example 7.1.

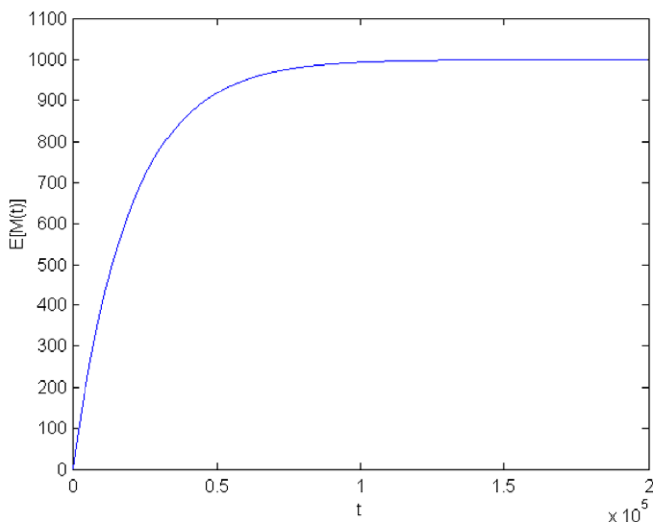

Fig. 7.2. Behavior of $EM(t)$ for Example 7.1.

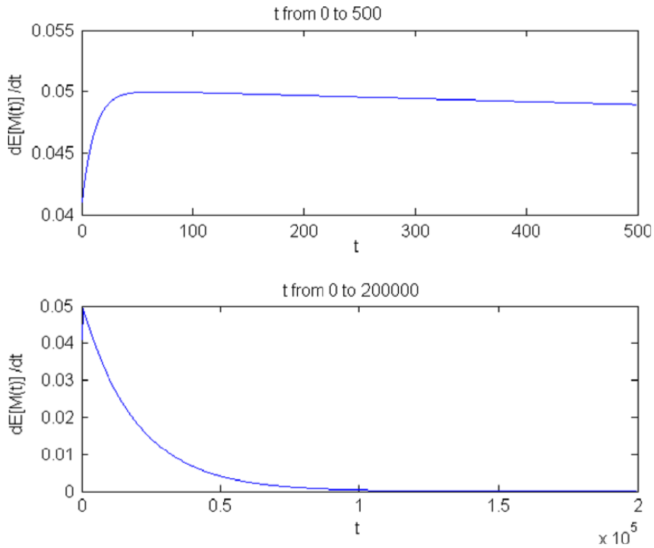

Fig. 7.3. The behavior of $\frac{\mathrm{d}(EM(t))}{\mathrm{d}t}$ for Example 7.1.

$\frac{\mathrm{d}(EM(t))}{\mathrm{d}t}$ behave as a monotonically increasing function at the very beginning of software testing. As software testing proceeds, more and more defects are detected and removed, the effects of the actions corresponding to θ_2 and θ_3 tend to diminish in comparison with the effect of the action corresponding to θ_1. However this does not mean that any of the state variables $x_1(t)$, $x_2(t)$ and $x_3(t)$ would diminish very fast. Actually, these state variables each approach zero in a comparable rate as demonstrated in Fig. 7.1. The state variable $x_i(t)$ represents the effect of a sequence of actions starting with action i. It does not represent the total effects of the single action, action i. These observations are important to clarify the roles of test cases of distinct classes.

In order to demonstrate the behavior of the variance of $M(t)$, it is important to note that $E(M(t))^2$ can be exactly expressed in a finite number of terms and thus so can the variance of $M(t)$. Recall that from Section 6,

$$EM_k^2 = \gamma_k \bullet \mathbf{1}^{\mathrm{T}}$$

Then

$$\begin{aligned}
E(M(t))^2 &= \sum_{k=0}^{\infty} EM_k^2 \frac{(\lambda t)^k}{k!} \mathrm{e}^{-\lambda t} = EM_1^2\left(\lambda t \mathrm{e}^{-\lambda t}\right) + EM_2^2 \frac{(\lambda t)^2}{2!} \mathrm{e}^{-\lambda t} + \sum_{k=3}^{\infty} EM_k^2 \frac{(\lambda t)^k}{k!} \mathrm{e}^{-\lambda t} \\
&= N\mathbf{p}_0\Theta \bullet \mathbf{1}^{\mathrm{T}}\left(\lambda t \mathrm{e}^{-\lambda t}\right) + EM_2^2 \frac{(\lambda t)^2}{2!} \mathrm{e}^{-\lambda t} + \sum_{k=3}^{\infty} EM_k^2 \frac{(\lambda t)^k}{k!} \mathrm{e}^{-\lambda t} \\
&= N\lambda t \mathrm{e}^{-\lambda t} \sum_{i=1}^{m} p_i\theta_i + EM_2^2 \frac{(\lambda t)^2}{2!} \mathrm{e}^{-\lambda t} \\
&\quad + \sum_{k=3}^{\infty} \sum_{j=1}^{k} \left\{ N\mathbf{p}_0\mathbf{P}^{k-j}\Theta + (2N-1)N \sum_{l=0}^{k-j-1} \mathbf{p}_0\mathbf{P}^{k-j-1-l}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^l\mathbf{P}\Theta \right\} \\
&\quad \times [\mathbf{P}(\mathbf{I}-2\Theta)]^{j-1} \bullet \mathbf{1}^{\mathrm{T}} \frac{(\lambda t)^k}{k!} \mathrm{e}^{-\lambda t} \\
&= N\lambda t \mathrm{e}^{-\lambda t} \sum_{i=1}^{m} p_i\theta_i + EM_2^2 \frac{(\lambda t)^2}{2!} \mathrm{e}^{-\lambda t} + N\mathbf{p}_0 \sum_{k=3}^{\infty} \sum_{j=1}^{k} \mathbf{P}^{k-j}\Theta[\mathbf{P}(\mathbf{I}-2\Theta)]^{j-1} \bullet \mathbf{1}^{\mathrm{T}} \frac{(\lambda t)^k}{k!} \mathrm{e}^{-\lambda t} \\
&\quad + (2N-1)N\mathbf{p}_0 \sum_{k=3}^{\infty} \sum_{j=1}^{k} \sum_{l=0}^{k-j-1} \mathbf{P}^{k-j-1-l}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^l\mathbf{P}\Theta[\mathbf{P}(\mathbf{I}-2\Theta)]^{j-1} \bullet \mathbf{1}^{\mathrm{T}} \frac{(\lambda t)^k}{k!} \mathrm{e}^{-\lambda t} \\
&= N\lambda t \mathrm{e}^{-\lambda t} \sum_{i=1}^{m} p_i\theta_i + EM_2^2 \frac{(\lambda t)^2}{2!} \mathrm{e}^{-\lambda t} + S_1 + S_2
\end{aligned}$$

where (with the observation that $\mathbf{P}$ is invertible in the example)

$$\begin{aligned}
S_1 &= N\mathbf{p}_0 \sum_{k=3}^{\infty} \sum_{j=1}^{k} \mathbf{P}^{k-j}\Theta[\mathbf{P}(\mathbf{I}-2\Theta)]^{j-1} \bullet \mathbf{1}^{\mathrm{T}} \frac{(\lambda t)^k}{k!}\,\mathrm{e}^{-\lambda t} \\
&= N\mathbf{p}_0\,\mathrm{e}^{-\lambda t} \sum_{k=3}^{\infty} \sum_{j=0}^{k-1} \mathbf{P}^{k-1-j}\Theta[\mathbf{P}(\mathbf{I}-2\Theta)]^{j} \bullet \mathbf{1}^{\mathrm{T}} \frac{(\lambda t)^k}{k!} \\
&= N\mathbf{p}_0\,\mathrm{e}^{-\lambda t} \sum_{k=3}^{\infty} \frac{(\lambda t)^k}{k!} \sum_{j=1}^{k-1} \mathbf{P}^{k-1-j}\Theta[\mathbf{P}(\mathbf{I}-2\Theta)]^{j} \bullet \mathbf{1}^{\mathrm{T}} + N\mathbf{p}_0\,\mathrm{e}^{-\lambda t} \sum_{k=3}^{\infty} \frac{(\lambda t)^k}{k!} \mathbf{P}^{k-1}\Theta \bullet \mathbf{1}^{\mathrm{T}} \\
&= N\mathbf{p}_0\,\mathrm{e}^{-\lambda t} \sum_{k=3}^{\infty} \frac{(\lambda t)^k}{k!} \sum_{j=1}^{k-1} \mathbf{P}^{k-1-j}\Theta\mathbf{P}[(\mathbf{I}-2\Theta)\mathbf{P}]^{j-1}(\mathbf{I}-2\Theta) \bullet \mathbf{1}^{\mathrm{T}} + N\mathbf{p}_0\,\mathrm{e}^{-\lambda t}\mathbf{P}^{-1} \sum_{k=3}^{\infty} \frac{(\lambda t)^k}{k!} \mathbf{P}^{k}\Theta \bullet \mathbf{1}^{\mathrm{T}} \\
&= N\mathbf{p}_0\,\mathrm{e}^{-\lambda t} \sum_{k=3}^{\infty} \frac{(\lambda t)^k}{k!} \sum_{j=1}^{k-1} \mathbf{P}^{k-1-j}\Theta\mathbf{P}[(\mathbf{I}-2\Theta)\mathbf{P}]^{j-1}(\mathbf{I}-2\Theta) \bullet \mathbf{1}^{\mathrm{T}} \\
&\quad + N\mathbf{p}_0\,\mathrm{e}^{-\lambda t}\mathbf{P}^{-1}\left(\mathrm{e}^{\lambda t\mathbf{P}} - \mathbf{I} - \lambda t\mathbf{P} - \frac{(\lambda t\mathbf{P})^2}{2}\right)\Theta \bullet \mathbf{1}^{\mathrm{T}}
\end{aligned}$$

And

$$\begin{aligned}
S_2 &= (2N-1)N\mathbf{p}_0 \sum_{k=3}^{\infty} \sum_{j=1}^{k} \sum_{l=0}^{k-j-1} \mathbf{P}^{k-j-1-l}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^{l}\mathbf{P}\Theta[\mathbf{P}(\mathbf{I}-2\Theta)]^{j-1} \bullet \mathbf{1}^{\mathrm{T}} \frac{(\lambda t)^k}{k!}\,\mathrm{e}^{-\lambda t} \\
&= (2N-1)N\mathrm{e}^{-\lambda t}\mathbf{p}_0 \sum_{k=3}^{\infty} \frac{(\lambda t)^k}{k!} \sum_{j=0}^{k-1} \left\{ \sum_{l=0}^{k-j-2} \mathbf{P}^{k-j-2-l}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^{l}\mathbf{P}\Theta \right\} [\mathbf{P}(\mathbf{I}-2\Theta)]^{j} \bullet \mathbf{1}^{\mathrm{T}} \\
&= (2N-1)N\mathrm{e}^{-\lambda t}\mathbf{p}_0 \sum_{k=3}^{\infty} \frac{(\lambda t)^k}{k!} \sum_{j=1}^{k-1} \left\{ \sum_{l=0}^{k-j-2} \mathbf{P}^{k-j-2-l}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^{l}\mathbf{P}\Theta \right\} [\mathbf{P}(\mathbf{I}-2\Theta)]^{j} \bullet \mathbf{1}^{\mathrm{T}} + (2N-1)N\mathrm{e}^{-\lambda t}\mathbf{p}_0 \\
&\quad \times \sum_{k=3}^{\infty} \frac{(\lambda t)^k}{k!} \sum_{l=0}^{k-2} \mathbf{P}^{k-2-l}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^{l}\mathbf{P}\Theta \bullet \mathbf{1}^{\mathrm{T}}
\end{aligned}$$

Therefore

$$\begin{aligned}
E(M(t))^2 &= \sum_{k=0}^{\infty} EM_k^2 \frac{(\lambda t)^k}{k!}\,\mathrm{e}^{-\lambda t} \\
&= N\mathbf{p}_0(\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}) \bullet \mathbf{1}^{\mathrm{T}} \lambda t\,\mathrm{e}^{-\lambda t} + \frac{(\lambda t)^2}{2!}\,\mathrm{e}^{-\lambda t} EM_2^2 \\
&\quad + N\mathbf{p}_0\,\mathrm{e}^{-\lambda t}\mathbf{P}^{-1}\left(\mathrm{e}^{\lambda t\mathbf{P}} - \mathbf{I} - \lambda t\mathbf{P} - \frac{(\lambda t\mathbf{P})^2}{2}\right)\Theta \bullet \mathbf{1}^{\mathrm{T}} \\
&\quad + N(2N-1)\mathrm{e}^{-\lambda t}\mathbf{p}_0 \underset{k=3}{\overset{\infty}{m}} \frac{(\lambda t)^k}{k!} \left(\sum_{j=0}^{k-2} \mathbf{P}^{k-j-2}\Theta\mathbf{P}[(\mathbf{I}-\Theta)\mathbf{P}]^{j}\Theta \bullet \mathbf{1}^{\mathrm{T}} \right) \\
&\quad + N\mathrm{e}^{-\lambda t}\mathbf{p}_0 \sum_{k=3}^{\infty} \frac{(\lambda t)^k}{k!} \left(\sum_{j=1}^{k-1} \mathbf{P}^{k-j-1}\Theta\mathbf{P}[(\mathbf{I}-2\Theta)\mathbf{P}]^{j-1}(\mathbf{I}-2\Theta) \bullet \mathbf{1}^{\mathrm{T}} \right) \\
&\quad + N(2N-1)\mathrm{e}^{-\lambda t}\mathbf{p}_0 \sum_{k=3}^{\infty} \frac{(\lambda t)^k}{k!} \left(\sum_{j=1}^{k-1} \left(\sum_{l=0}^{k-j-2} \mathbf{P}^{k-j-l-2}\Theta\mathbf{P}[(\mathbf{I}-\Theta)\mathbf{P}]^{l}\Theta\mathbf{P} \right) [(\mathbf{I}-2\Theta)\mathbf{P}]^{j-1} \right) \\
&\quad \times (\mathbf{I}-2\Theta) \bullet \mathbf{1}^{\mathrm{T}}
\end{aligned}$$

where

$$EM_2^2 = \gamma_2 \bullet \mathbf{1}^{\mathrm{T}} = [N\mathbf{p}_0\mathbf{P}\Theta + (2N-1)N\mathbf{p}_0\Theta\mathbf{P}\Theta + N\mathbf{p}_0\Theta\mathbf{P}(\mathbf{I}-2\Theta)] \bullet \mathbf{1}^{\mathrm{T}}$$

Note that the matrices $\mathbf{P}$ and $(\mathbf{I}-\Theta)\mathbf{P}$ can be transformed via the similarity matrices $\mathbf{T}_{\mathbf{P}}$ and $\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}$, respectively, into the diagonal forms as follows:

$$\begin{aligned}
&\mathbf{P} = \mathbf{T}_{\mathbf{P}}\mathbf{J}_{\mathbf{P}}\mathbf{T}_{\mathbf{P}}^{-1} \\
&(\mathbf{I}-\Theta)\mathbf{P} = \mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}\mathbf{J}_{(\mathbf{I}-\Theta)\mathbf{P}}\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}^{-1}
\end{aligned}$$

where $\mathbf{J}_{\mathbf{P}} = \mathrm{diag}\{\lambda_{\mathbf{P}1}, \lambda_{\mathbf{P}2}, \lambda_{\mathbf{P}3}\}$, $\quad \mathbf{J}_{(\mathbf{I}-\Theta)\mathbf{P}} = \mathrm{diag}\{\lambda_{(\mathbf{I}-\Theta)\mathbf{P}1}, \lambda_{(\mathbf{I}-\Theta)\mathbf{P}2}, \lambda_{(\mathbf{I}-\Theta)\mathbf{P}3}\}$. In this way

$$\begin{aligned}
\sum_{k=3}^{\infty}\frac{(\lambda t)^k}{k!}\left(\sum_{j=0}^{k-2}\mathbf{P}^{k-j-2}\Theta\mathbf{P}[(\mathbf{I}-\Theta)\mathbf{P}]^j\right) &= \sum_{k=3}^{\infty}\frac{(\lambda t)^k}{k!}\left(\sum_{j=0}^{k-2}\left[\mathbf{T}_{\mathbf{P}}\mathbf{J}_{\mathbf{P}}\mathbf{T}_{\mathbf{P}}^{-1}\right]^{k-j-2}\Theta\mathbf{P}\left[\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}\mathbf{J}_{(\mathbf{I}-\Theta)\mathbf{P}}\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}^{-1}\right]^j\right) \\
&= \sum_{k=3}^{\infty}\frac{(\lambda t)^k}{k!}\left(\sum_{j=0}^{k-2}\mathbf{T}_{\mathbf{P}}\mathbf{J}_{\mathbf{P}}^{k-j-2}\mathbf{T}_{\mathbf{P}}^{-1}\Theta\mathbf{P}\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}\mathbf{J}_{(\mathbf{I}-\Theta)\mathbf{P}}^{j}\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}^{-1}\right) \\
&= \sum_{k=3}^{\infty}\frac{(\lambda t)^k}{k!}\mathbf{T}_{\mathbf{P}}\mathbf{J}_{\mathbf{P}}^{k-2}\left(\sum_{j=0}^{k-2}\mathbf{J}_{\mathbf{P}}^{-j}\mathbf{T}_{\mathbf{P}}^{-1}\Theta\mathbf{P}\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}\mathbf{J}_{(\mathbf{I}-\Theta)\mathbf{P}}^{j}\right)\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}^{-1}
\end{aligned}$$

Denote

$$\mathbf{T}_{\mathbf{P}}^{-1}\Theta\mathbf{P}\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}} = [d_{st}]_{3\times 3}$$

Then

$$\sum_{j=0}^{k-2}\mathbf{J}_{\mathbf{P}}^{-j}\mathbf{T}_{\mathbf{P}}^{-1}\Theta\mathbf{P}\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}\mathbf{J}_{(\mathbf{I}-\Theta)\mathbf{P}}^{j} = \sum_{j=0}^{k-2}\mathrm{diag}\left\{\lambda_{\mathbf{P}1}^{-j}, \lambda_{\mathbf{P}2}^{-j}, \lambda_{\mathbf{P}3}^{-j}\right\}[d_{st}]_{3\times 3}\mathrm{diag}\left\{\lambda_{(\mathbf{I}-\Theta)\mathbf{P}1}^{j}, \lambda_{(\mathbf{I}-\Theta)\mathbf{P}2}^{j}, \lambda_{(\mathbf{I}-\Theta)}^{\mathbf{P}3j}\right\}$$

Consider one of the entries of the matrix $\sum_{j=0}^{k-2}\mathrm{diag}\left\{\lambda_{\mathbf{P}1}^{-j}, \lambda_{\mathbf{P}2}^{-j}, \lambda_{\mathbf{P}3}^{-j}\right\}[d_{st}]_{3\times 3}\mathrm{diag}\left\{\lambda_{(\mathbf{I}-\Theta)\mathbf{P}1}^{j}, \lambda_{(\mathbf{I}-\Theta)\mathbf{P}2}^{j}, \lambda_{(\mathbf{I}-\Theta)}^{\mathbf{P}3j}\right\}$, which is $\sum_{j=0}^{k-2}\lambda_{\mathbf{P}s}^{-j}d_{st}\lambda_{(\mathbf{I}-\Theta)\mathbf{P}t}^{j}$. We have

$$\sum_{j=0}^{k-2}\lambda_{\mathbf{P}s}^{-j}d_{st}\lambda_{(\mathbf{I}-\Theta)\mathbf{P}t}^{j} = d_{st}\frac{1-\left(\frac{\lambda_{(\mathbf{I}-\Theta)\mathbf{P}t}}{\lambda_{\mathbf{P}s}}\right)^{k-1}}{1-\left(\frac{\lambda_{(\mathbf{I}-\Theta)\mathbf{P}t}}{\lambda_{\mathbf{P}s}}\right)}$$

Therefore each entry of the matrix $\sum_{k=3}^{\infty}\frac{(\lambda t)^k}{k!}\mathbf{T}_{\mathbf{P}}\mathbf{J}_{\mathbf{P}}^{k-2}\left(\sum_{j=0}^{k-2}\mathbf{J}_{\mathbf{P}}^{-j}\mathbf{T}_{\mathbf{P}}^{-1}\Theta\mathbf{P}\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}\mathbf{J}_{(\mathbf{I}-\Theta)\mathbf{P}}^{j}\right)\mathbf{T}_{(\mathbf{I}-\Theta)\mathbf{P}}^{-1}$ can be expressed in a finite number of terms. Eventually, we have

$$\begin{aligned}
\mathrm{var}(M(t)) &= E(M(t))^2 - (EM(t))^2 \\
&\approx -0.01\,\mathrm{e}^{-0.070007t} - 20.163\,\mathrm{e}^{-0.070015t} - 0.00010298\,\mathrm{e}^{-0.14001t} - 0.0088135\,\mathrm{e}^{-0.18001t} \\
&\quad + 1000\,\mathrm{e}^{-0.000050155t} - 0.09\,\mathrm{e}^{-0.090006t} + 187.78\,\mathrm{e}^{-0.090056t} - 0.0019054\,\mathrm{e}^{-0.16001t} \\
&\quad + 20.298\,\mathrm{e}^{-0.070057t} + 999210\,\mathrm{e}^{-0.0001003t} - 187.68\,\mathrm{e}^{-0.090011t} - 1000200\,\mathrm{e}^{-0.00010031t}
\end{aligned}$$

The above formula is obtained with the help of Matlab and thus subject to approximation. Fig. 7.4 shows the behavior of the variance of $M(t)$, which demonstrates as a convex function. It grows fast at the beginning of software testing. Then it reaches the peak around $k = 2500$ (noticing that each action takes $1/\lambda = 10$ s on average and the peak appears at about $t \approx 25{,}000$ s). After the peak, the variance of $M(t)$ drops down gradually and eventually approaches zero. Comparing Fig. 7.4 with Fig. 7.2, it is evident that $EM(t) \neq \mathrm{var}(M(t))$. This coincides with the empirical observation that $M(t)$ does not follow a non-homogeneous Poisson process [4]. The shape of the curve depicted in Fig. 7.4 also coincides with that of the estimated variance of the number of remaining defects that was obtained from an empirical dataset (Fig. 5 in Ref. [29]).

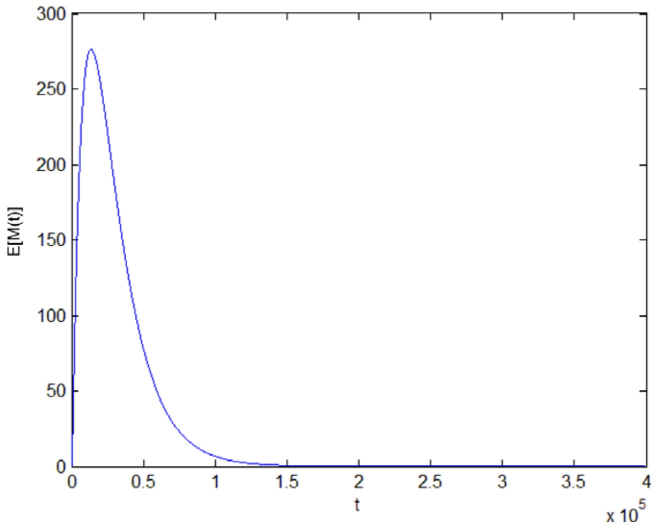

Fig. 7.4. Behavior of the variance of $M(t)$ for Example 7.1.

Example 7.2. The setting of this example is identical to that of Example 7.1 except

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

That is, suppose an action selects a test case from class 1, the next action must select a test case from class 2, and then a test case of class 3 must be selected. After the selected test case of class 3 is executed, a test case of class 1 must be selected again. This process of test case selection is repeated again and again. As in Example 7.2, we are interested in the behavior of $M(t)$. In particular, we need to examine if different $\mathbf{P}$ leads to dramatically different patterns of the behavior of $M(t)$.

We have

$$\mathbf{K} = \begin{bmatrix} -0.10000 & 0.09990 & 0 \\ 0 & -0.10000 & 0.09998 \\ 0.09999 & 0 & -0.10000 \end{bmatrix}$$

The three eigenvalues of $\mathbf{K}$ are,

$$\lambda(\mathbf{K})_1 = -0.00004334144943$$
$$\lambda(\mathbf{K})_2 = -0.14997832927528 + \mathrm{img}(0.08656500558220)$$
$$\lambda(\mathbf{K})_3 = -0.14997832927528 - \mathrm{img}(0.08656500558220)$$

That is, $\mathbf{K}$ has one real eigenvalue and two cognate complex eigenvalues, where *img* denotes the imaginary part of a complex eigenvalue.

Figs. 7.5–7.8 are the counterparts to Figs. 7.1–7.4, respectively, which show the behavior of the corresponding $\mathbf{X}(t)$, $EM(t)$, $\frac{\mathrm{d}(EM(t))}{\mathrm{d}t}$, and $\mathrm{var}(M(t))$. They demonstrate similar patterns as those observed in Figs. 7.1–7.4, although the corresponding $\mathbf{P}$ is essentially different from that of Example 7.1 (in the sense of eigenvalues). The underlying cause for the emergence of the similar patterns under different $\mathbf{P}$ is not clear at this stage and should be investigated in the future.

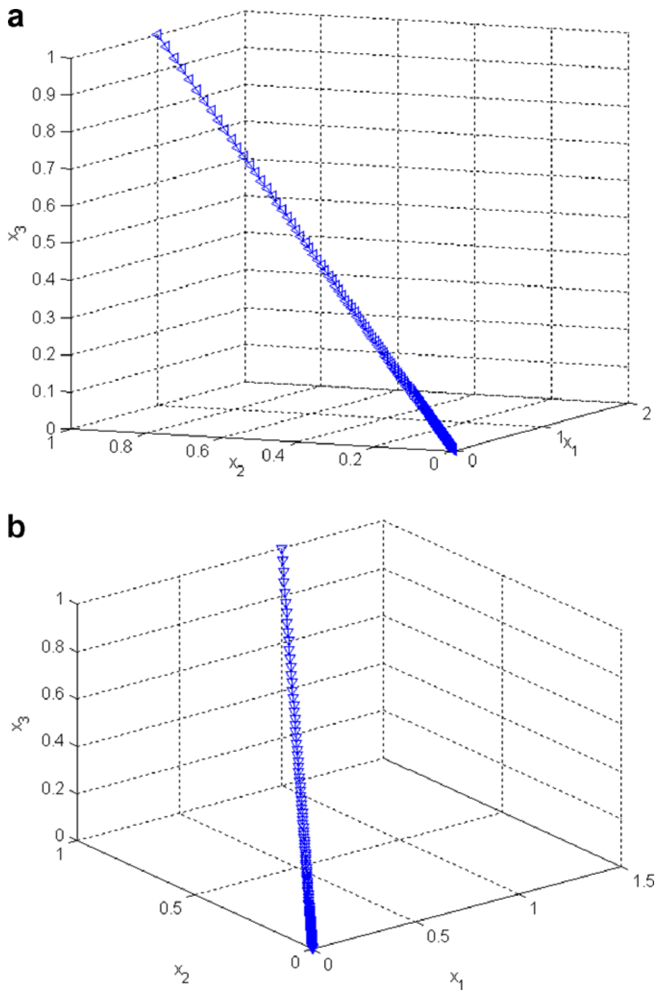

Fig. 7.5. Behavior of the states of the linear dynamic system for Example 7.2.

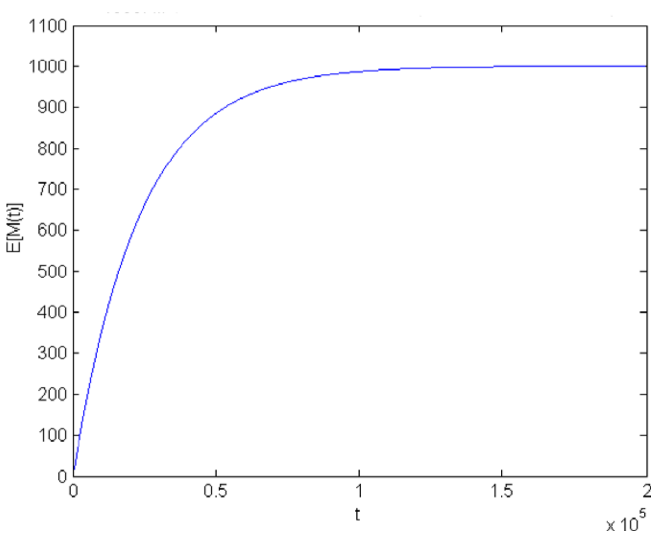

Fig. 7.6. Behavior of $EM(t)$ for Example 7.2.

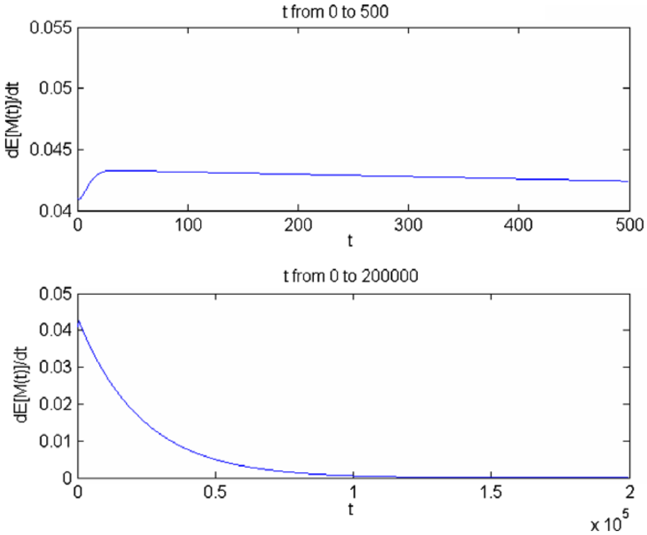

Fig. 7.7. The behavior of $\frac{\mathrm{d}(EM(t))}{\mathrm{d}t}$ for Example 7.2.

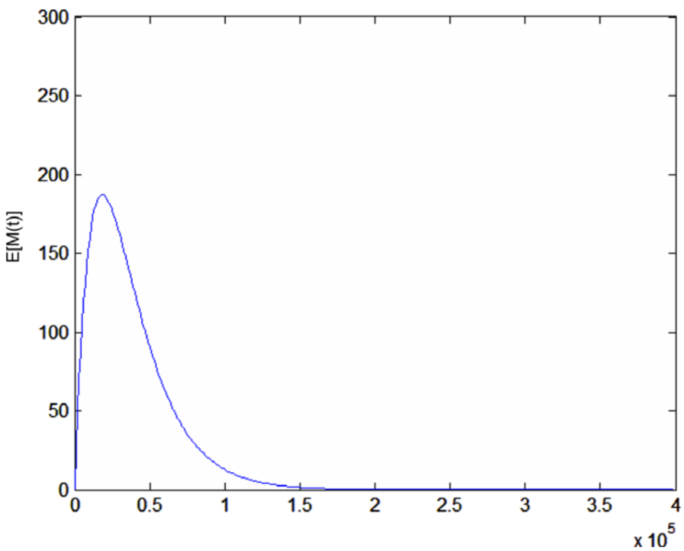

Fig. 7.8. Behavior of the variance of $M(t)$ for Example 7.2.

8. Concluding remarks

Software testing is essential for software reliability improvement and assurance. The processes of software testing are intrinsically dynamic since test cases are dynamically selected one by one and the software under test is subject to changes due to defect removals. However it is not clear how to formulate the software testing processes with quantitative reliability concern in a mathematically rigorous manner and what the dynamics of the software testing processes may demonstrate. In order to partially address these questions, in the preceding sections we adopt a set of simplifying assumptions for software testing and analyze the resulting software testing process. In particular, the mathematical formulae for the expected number of observed software failures are rigorously derived, the bounds and trends of the expected number of observed software failures are analyzed, and the variance of the number of observed software failures is examined. On the other hand, it is demonstrated that under the simplifying assumptions, the software testing process can be treated as a linear dynamic system. This suggests that various software testing processes could be classified as linear or non-linear, and software testing might be investigated from the perspective of system dynamics.

Many research topics deserve further investigation. For example, how can software reliability be predicted by using the simplifying assumptions? How can the linear system model be identified from the observed testing data? How can the software testing process be controlled with the linear system model? What software testing processes are non-linear? All these questions look challenging.

Acknowledgements

The simulation results were obtained with the help of Lin Tie and Quan Quan. The mathematical results were refined with the help of Ping Cao and Fang Liu. The valuable comments of the anonymous reviewers led the authors to clarify the motivation of this paper and to improve the readability of this paper. Cai was partially supported by the National Science Foundation of China and MicroSoft Research Asia (Grant Nos. 60633010 and 60474006) and the 863 Programme of China (Grant No. 2006AA01Z174). Dong was partially supported by the National Science Foundation of China (Grant No. 10671197) and the 973 Programme of China (Grant No. 2006CB805900). Liu was partially supported by the National Science Foundation of China (Grant Nos. 60674082 and 70221001).

Appendix

Proof of Proposition 3.1. For the case of $k > N$, it holds that

$$
\begin{aligned}
E[M_k|A_k=i] &= \sum_{j=0}^{N} j\Pr\{M_k=j|A_k=i\} = \sum_{j=1}^{N} \frac{j\Pr\{M_k=j,A_k=i\}}{\Pr\{A_k=i\}} \\
&= \frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{N} j\Pr\{M_k=j,A_k=i\} = \frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{N}\sum_{l=1}^{m} j\Pr\{M_k=j,A_k=i,A_{k-1}=l\} \\
&= \frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{N}\sum_{l=1}^{m} j[\Pr\{M_k=j,A_k=i;M_{k-1}=j,A_{k-1}=l\} \\
&\quad + \Pr\{M_k=j,A_k=i;M_{k-1}=j-1,A_{k-1}=l\}] \\
&= \frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{N}\sum_{l=1}^{m} j[\Pr\{M_k=j,A_k=i|M_{k-1}=j,A_{k-1}=l\}\Pr\{M_{k-1}=j,A_{k-1}=l\} \\
&\quad + \Pr\{M_k=j,A_k=i|M_{k-1}=j-1,A_{k-1}=l\}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}] \\
&= \frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{N}\sum_{l=1}^{m} j[(1-(N-j)\theta_i)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\} \\
&\quad + (N-j+1)\theta_i p_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}] \\
&= \frac{1}{\Pr\{A_k=i\}}\sum_{l=1}^{m}\sum_{j=1}^{N}[\big(j(1-N\theta_i)+j^2\theta_i\big)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\} \\
&\quad + \big((N+1)j-j^2\big)\theta_i p_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]
\end{aligned}
$$

Note that the four terms in the above brackets are determined as follows:

$$\begin{aligned}
&\sum_{j=1}^{N} j(1-N\theta_i)p_{li}\Pr\{M_{k-1}=j, A_{k-1}=l\}\\
&\quad=\sum_{j=1}^{N} j(1-N\theta_i)p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=(1-N\theta_i)p_{li}E\{M_{k-1}|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\sum_{j=1}^{N} j^2\theta_i p_{li}\Pr\{M_{k-1}=j, A_{k-1}=l\}\\
&\quad=\sum_{j=1}^{N} j^2\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=\theta_i p_{li}E\left\{[M_{k-1}]^2|A_{k-1}=l\right\}\Pr\{A_{k-1}=l\}\\
&\sum_{j=1}^{N} j(N+1)\theta_i p_{li}\Pr\{M_{k-1}=j-1, A_{k-1}=l\}\\
&\quad=\sum_{j=1}^{N} j(N+1)\theta_i p_{li}\Pr\{M_{k-1}=j-1|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=\sum_{j=0}^{N-1} (j+1)(N+1)\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=\sum_{j=0}^{N-1} j(N+1)\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\qquad+\sum_{j=0}^{N-1} (N+1)\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=\sum_{j=0}^{N} j(N+1)\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\qquad-N(N+1)\theta_i p_{li}\Pr\{M_{k-1}=N|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\qquad+\sum_{j=0}^{N} (N+1)\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\qquad-(N+1)\theta_i p_{li}\Pr\{M_{k-1}=N|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=(N+1)\theta_i p_{li}E\{M_{k-1}|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\qquad+(N+1)\theta_i p_{li}\Pr\{A_{k-1}=l\}\\
&\qquad-(N+1)^2\theta_i p_{li}\Pr\{M_{k-1}=N|A_{k-1}=l\}\Pr\{A_{k-1}=l\}
\end{aligned}$$

and

$$\begin{aligned}
\sum_{j=1}^{N} j^2\theta_i p_{li}\Pr\{M_{k-1}=j-1, A_{k-1}=l\} &= \sum_{j=0}^{N-1} (j+1)^2\theta_i p_{li}\Pr\{M_{k-1}=j, A_{k-1}=l\}\\
&= \sum_{j=0}^{N} (j+1)^2\theta_i p_{li}\Pr\{M_{k-1}=j, A_{k-1}=l\}\\
&\quad-(N+1)^2\theta_i p_{li}\Pr\{M_{k-1}=N, A_{k-1}=l\}\\
&= \theta_i p_{li}E\left\{[M_{k-1}]^2|A_{k-1}=l\right\}\Pr\{A_{k-1}=l\}\\
&\quad+2\theta_i p_{li}E\{M_{k-1}|A_{k-1}=l\}\Pr\{A_{k-1}=l\}+\theta_i p_{li}\Pr\{A_{k-1}=l\}\\
&\quad-(N+1)^2\theta_i p_{li}\Pr\{M_{k-1}=N, A_{k-1}=l\}
\end{aligned}$$

In this way

$$\sum_{j=1}^{N}\sum_{l=1}^{m}[\big(j(1-N\theta_i)+j^2\theta_i\big)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\}+\big((N+1)j-j^2\big)\theta_i p_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]$$
$$=\sum_{l=1}^{m}[N\theta_i p_{li}+(1-\theta_i)p_{li}E[M_{k-1}|A_{k-1}=l]]\Pr\{A_{k-1}=l\}$$

Therefore

$$E[M_k|A_k=i]\Pr\{A_k=i\}=\sum_{l=1}^{m}[N\theta_i p_{li}+(1-\theta_i)p_{li}E[M_{k-1}|A_{k-1}=l]]\Pr\{A_{k-1}=l\}$$

For the case of $k\leqslant N$, it holds that

$$\begin{aligned}
E[M_k|A_k=i]&=\sum_{j=0}^{k}j\Pr\{M_k=j|A_k=i\}=\sum_{j=1}^{k}\frac{j\Pr\{M_k=j,A_k=i\}}{\Pr\{A_k=i\}}\\
&=\frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{k}\sum_{l=1}^{m}j\Pr\{M_k=j,A_k=i,A_{k-1}=l\}\\
&=\frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{k-1}\sum_{l=1}^{m}j\Pr\{M_k=j,A_k=i,A_{k-1}=l\}\\
&\quad+\frac{1}{\Pr\{A_k=i\}}\sum_{l=1}^{m}k\Pr\{M_k=k,A_k=i,A_{k-1}=l\}\\
&=\frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{k-1}\sum_{l=1}^{m}j[\Pr\{M_k=j,A_k=i,M_{k-1}=j,A_{k-1}=l\}\\
&\quad+\Pr\{M_k=j,A_k=i,M_{k-1}=j-1,A_{k-1}=l\}]\\
&\quad+\frac{1}{\Pr\{A_k=i\}}\sum_{l=1}^{m}k\Pr\{M_k=k,A_k=i,M_{k-1}=k-1,A_{k-1}=l\}
\end{aligned}$$

This implies that

$$\begin{aligned}
E[M_k|A_k=i]&=\sum_{j=0}^{k}j\Pr\{M_k=j|A_k=i\}\\
&=\frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{k-1}\sum_{l=1}^{m}j[\Pr\{M_k=j,A_k=i|M_{k-1}=j,A_{k-1}=l\}\Pr\{M_{k-1}=j,A_{k-1}=l\}\\
&\quad+\Pr\{M_k=j,A_k=i|M_{k-1}=j-1,A_{k-1}=l\}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]\\
&\quad+\frac{1}{\Pr\{A_k=i\}}\sum_{l=1}^{m}k\Pr\{M_k=k,A_k=i|M_{k-1}=k-1,A_{k-1}=l\}\Pr\{M_{k-1}=k-1,A_{k-1}=l\}\\
&=\frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{k-1}\sum_{l=1}^{m}j[(1-(N-j)\theta_i)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\}\\
&\quad+(N-j+1)\theta_i p_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]\\
&\quad+\frac{1}{\Pr\{A_k=i\}}\sum_{l=1}^{m}k(N-k+1)\theta_i p_{li}\Pr\{M_{k-1}=k-1,A_{k-1}=l\}\\
&=\frac{1}{\Pr\{A_k=i\}}\sum_{l=1}^{m}k(N-k+1)\theta_i p_{li}\Pr\{M_{k-1}=k-1,A_{k-1}=l\}\\
&\quad+\frac{1}{\Pr\{A_k=i\}}\sum_{l=1}^{m}\sum_{j=1}^{k-1}[\big(j(1-N\theta_i)+j^2\theta_i\big)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\}\\
&\quad+\big((N+1)j-j^2\big)\theta_i p_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]
\end{aligned}$$

Note that the four terms in the above brackets are determined as follows:

$$\begin{aligned}
&\sum_{j=1}^{k-1} j(1-N\theta_i)p_{li}\Pr\{M_{k-1}=j, A_{k-1}=l\}\\
&\quad=\sum_{j=1}^{k-1} j(1-N\theta_i)p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=(1-N\theta_i)p_{li}E[M_{k-1}|A_{k-1}=l]\Pr\{A_{k-1}=l\}\\
&\sum_{j=1}^{k-1} j^2\theta_i p_{li}\Pr\{M_{k-1}=j, A_{k-1}=l\}\\
&\quad=\sum_{j=1}^{k-1} j^2\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=\theta_i p_{li}E[(M_{k-1})^2|A_{k-1}=l]\Pr\{A_{k-1}=l\}\\
&\sum_{j=1}^{k-1} j(N+1)\theta_i p_{li}\Pr\{M_{k-1}=j-1, A_{k-1}=l\}\\
&\quad=\sum_{j=1}^{k-1} j(N+1)\theta_i p_{li}\Pr\{M_{k-1}=j-1|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=\sum_{j=0}^{k-2} (j+1)(N+1)\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=\sum_{j=0}^{k-2} j(N+1)\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\qquad+\sum_{j=0}^{k-2} (N+1)\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=\sum_{j=0}^{k-1} j(N+1)\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\qquad-(k-1)(N+1)\theta_i p_{li}\Pr\{M_{k-1}=k-1|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\qquad+\sum_{j=0}^{k-1} (N+1)\theta_i p_{li}\Pr\{M_{k-1}=j|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\qquad-(N+1)\theta_i p_{li}\Pr\{M_{k-1}=k-1|A_{k-1}=l\}\Pr\{A_{k-1}=l\}\\
&\quad=(N+1)\theta_i p_{li}E[M_{k-1}|A_{k-1}=l]\Pr\{A_{k-1}=l\}+(N+1)\theta_i p_{li}\Pr\{A_{k-1}=l\}\\
&\qquad-k(N+1)\theta_i p_{li}\Pr\{M_{k-1}=k-1|A_{k-1}=l\}\Pr\{A_{k-1}=l\}
\end{aligned}$$

and

$$\begin{aligned}
\sum_{j=1}^{k-1} j^2\theta_i p_{li}\Pr\{M_{k-1}=j-1, A_{k-1}=l\}&=\sum_{j=0}^{k-2} (j+1)^2\theta_i p_{li}\Pr\{M_{k-1}=j, A_{k-1}=l\}\\
&=\sum_{j=0}^{k-1} (j+1)^2\theta_i p_{li}\Pr\{M_{k-1}=j, A_{k-1}=l\}\\
&\quad-k^2\theta_i p_{li}\Pr\{M_{k-1}=k-1, A_{k-1}=l\}\\
&=\theta_i p_{li}E[(M_{k-1})^2|A_{k-1}=l]\Pr\{A_{k-1}=l\}\\
&\quad+2\theta_i p_{li}E[M_{k-1}|A_{k-1}=l]\Pr\{A_{k-1}=l\}\\
&\quad+\theta_i p_{li}\Pr\{A_{k-1}=l\}-k^2\theta_i p_{li}\Pr\{M_{k-1}=k-1, A_{k-1}=l\}
\end{aligned}$$

In this way

$$\sum_{j=1}^{k-1}\sum_{l=1}^{m}[\big(j(1-N\theta_i)+j^2\theta_i\big)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\}+\big((N+1)j-j^2\big)\theta_ip_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]$$

$$\begin{aligned}&=\sum_{l=1}^{m}[(1-N\theta_i)p_{li}E[M_{k-1}|A_{k-1}=l]+\theta_ip_{li}E[(M_{k-1})^2|A_{k-1}=l]+(N+1)\theta_ip_{li}E[M_{k-1}|A_{k-1}=l]\\&\quad+(N+1)\theta_ip_{li}-(N+1)k\theta_ip_{li}\Pr\{M_{k-1}=k-1|A_{k-1}=l\}-\theta_ip_{li}E[(M_{k-1})^2|A_{k-1}=l]\\&\quad-2\theta_ip_{li}E[M_{k-1}|A_{k-1}=l]-\theta_ip_{li}+k^2\theta_ip_{li}\Pr\{M_{k-1}=k-1|A_{k-1}=l\}]\Pr\{A_{k-1}=l\}\\&=\sum_{l=1}^{m}[N\theta_ip_{li}+(1-\theta_i)p_{li}E[M_{k-1}|A_{k-1}=l]+\big(k^2-k-Nk\big)\theta_ip_{li}\Pr\{M_{k-1}=k-1|A_{k-1}=l\}]\Pr\{A_{k-1}=l\}\end{aligned}$$

Therefore

$$\begin{aligned}E[M_k|A_k=i]&=\frac{1}{\Pr\{A_k=i\}}\sum_{l=1}^{m}k(N-k+1)\theta_ip_{li}\Pr\{M_{k-1}=k-1,A_{k-1}=l\}\\&\quad+\frac{1}{\Pr\{A_k=i\}}\sum_{j=1}^{k-1}\sum_{l=1}^{m}[\big(j(1-N\theta_i)+j^2\theta_i\big)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\}\\&\quad+\big((N+1)j-j^2\big)\theta_ip_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]\\&=\frac{1}{\Pr\{A_k=i\}}\sum_{l=1}^{m}[N\theta_ip_{li}+(1-\theta_i)p_{li}E[M_{k-1}|A_{k-1}=l]]\Pr\{A_{k-1}=l\}\end{aligned}$$

In summary, for all $k>0$ ($k\leqslant N$ or $k>N$), it always holds that

$$E[M_k|A_k=i]\Pr\{A_k=i\}=\sum_{l=1}^{m}[N\theta_ip_{li}+(1-\theta_i)p_{li}E[M_{k-1}|A_{k-1}=l]]\Pr\{A_{k-1}=l\}$$

Or

$$\alpha_{k+1}=\alpha_k\mathbf{H}+N\beta_k\Psi=\alpha_k\mathbf{P}(\mathbf{I}-\Theta)+N\mathbf{p_0}\mathbf{P}^k\Theta=N\sum_{j=0}^{k}\mathbf{p_0}\mathbf{P}^{k-j}\Theta(\mathbf{P}(\mathbf{I}-\Theta))^j;\quad k\geqslant 0.\qquad\square$$

Proof of Proposition 3.2

First we note that

$$EM_k=\sum_{i=1}^{m}E[M_k|A_k=i]\Pr\{A_k=i\}=\sum_{i=1}^{m}\alpha(k,i)=\alpha_k\bullet\mathbf{1}^{\mathrm{T}}$$

It holds

$$\begin{aligned}EM_k&=N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^j\bullet\mathbf{1}^{\mathrm{T}}=N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}(\mathbf{I}-(\mathbf{I}-\Theta))[\mathbf{P}(\mathbf{I}-\Theta)]^j\bullet\mathbf{1}^{\mathrm{T}}\\&=N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}[\mathbf{P}(\mathbf{I}-\Theta)]^j\bullet\mathbf{1}^{\mathrm{T}}-N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}(\mathbf{I}-\Theta)[\mathbf{P}(\mathbf{I}-\Theta)]^j\bullet\mathbf{1}^{\mathrm{T}}\\&=N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}\mathbf{H}^j\bullet\mathbf{1}^{\mathrm{T}}-N\sum_{j=0}^{k-2}\mathbf{p_0}\mathbf{P}^{k-j-2}[\mathbf{P}(\mathbf{I}-\Theta)]^{j+1}\bullet\mathbf{1}^{\mathrm{T}}-N\mathbf{p_0}(\mathbf{I}-\Theta)[\mathbf{P}(\mathbf{I}-\Theta)]^{k-1}\bullet\mathbf{1}^{\mathrm{T}}\\&=N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}\mathbf{H}^j\bullet\mathbf{1}^{\mathrm{T}}-N\sum_{j=0}^{k-2}\mathbf{p_0}\mathbf{P}^{k-1-(j+1)}\mathbf{H}^{j+1}\bullet\mathbf{1}^{\mathrm{T}}-N\mathbf{p_0}(\mathbf{I}-\Theta)[\mathbf{P}(\mathbf{I}-\Theta)]^{k-1}\bullet\mathbf{1}^{\mathrm{T}}\\&=N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}\mathbf{H}^j\bullet\mathbf{1}^{\mathrm{T}}-N\sum_{j=1}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}\mathbf{H}^j\bullet\mathbf{1}^{\mathrm{T}}-N\mathbf{p_0}(\mathbf{I}-\Theta)[\mathbf{P}(\mathbf{I}-\Theta)]^{k-1}\bullet\mathbf{1}^{\mathrm{T}}\\&=N\mathbf{p_0}\mathbf{P}^{k-1}\bullet\mathbf{1}^{\mathrm{T}}-N\mathbf{p_0}(\mathbf{I}-\Theta)[\mathbf{P}(\mathbf{I}-\Theta)]^{k-1}\bullet\mathbf{1}^{\mathrm{T}}\end{aligned}$$

Therefore

$$EM_k = N\mathbf{p_0}\left(\mathbf{I} - (\mathbf{I} - \Theta)[\mathbf{P}(\mathbf{I} - \Theta)]^{k-1}\right) \bullet \mathbf{1}^{\mathrm{T}}$$

On the other hand,

$$\begin{aligned}
EM(t) &= \sum_{k=0}^{\infty} EM_k \,\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!} = \sum_{k=1}^{\infty} N\mathbf{p_0}\left(\mathbf{I} - (\mathbf{I} - \Theta)[\mathbf{P}(\mathbf{I} - \Theta)]^{k-1}\right) \bullet \mathbf{1}^{\mathrm{T}}\,\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!} \\
&= N\mathbf{p_0} \bullet \mathbf{1}^{\mathrm{T}}\left(1 - \mathrm{e}^{-\lambda t}\right) - \sum_{k=1}^{\infty} N\mathbf{p_0}(\mathbf{I} - \Theta)[\mathbf{P}(\mathbf{I} - \Theta)]^{k-1} \bullet \mathbf{1}^{\mathrm{T}}\,\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!} \\
&= N\left(1 - \mathrm{e}^{-\lambda t}\right) - \sum_{k=1}^{\infty} N\mathbf{p_0}[(\mathbf{I} - \Theta)\mathbf{P}]^k \bullet \mathbf{1}^{\mathrm{T}}\,\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!} = N\mathbf{p_0} \bullet \mathbf{1}^{\mathrm{T}} - N\mathbf{p_0}\,\mathrm{e}^{-\lambda t}\,\mathrm{e}^{\lambda t(\mathbf{I}-\Theta)\mathbf{P}} \bullet \mathbf{1}^{\mathrm{T}} \\
&= N - N\mathbf{p_0}\,\mathrm{e}^{-\lambda t[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]} \bullet \mathbf{1}^{\mathrm{T}}. \qquad \square
\end{aligned}$$

Proof of lemma 4.2

From Ref. [21], we know that for any nonnegative matrix $\mathbf{L} = [l_{ij}]_{m\times m}$ with $l_i = \sum_{j=1}^{m} l_{ij} > 0$, $i = 1, 2, \ldots, m$, it holds that

$$\min_{1\leqslant i\leqslant m}\left(\frac{1}{l_i}\sum_{j=1}^{m} l_{ij}l_j\right) \leqslant \mu^{(\mathbf{L})} \leqslant \max_{1\leqslant i\leqslant m}\left(\frac{1}{l_i}\sum_{j=1}^{m} l_{ij}l_j\right)$$

where $\mu^{(\mathbf{L})}$ denotes the maximal eigenvalue of the matrix $\mathbf{L}$. Now let $\mathbf{L} = (\mathbf{I} - \Theta)\mathbf{P}$, we have

$$\min_{1\leqslant i\leqslant m}\frac{1}{1-\theta_i}\sum_{j=1}^{m}(1-\theta_i)p_{ij}r_j \leqslant \mu \leqslant \max_{1\leqslant i\leqslant m}\frac{1}{1-\theta_i}\sum_{j=1}^{m}(1-\theta_i)p_{ij}r_j$$

where $r_j = \sum_{k=1}^{m}\left(1-\theta_j\right)p_{jk} = 1 - \sum_{k=1}^{m}\theta_j p_{jk} = 1 - \theta_j$. In this way

$$\frac{1}{1-\theta_i}\sum_{j=1}^{m}(1-\theta_i)p_{ij}r_j = \sum_{j=1}^{m}p_{ij}\left(1-\theta_j\right) = 1 - \sum_{j=1}^{m}\theta_j p_{ij}. \qquad \square$$

Proof of Proposition 4.1

For any initial value $\mathbf{X}(0)$ which is non-negative, we have

$$\begin{aligned}
\mathbf{X}(t) &= \mathrm{e}^{\mathbf{K}t}\mathbf{X}(0) = \sum_{n=0}^{\infty}\frac{t^n}{n!}\mathbf{K}^n\mathbf{X}(0) = \mathbf{T}^{-1}\mathbf{T}\sum_{n=0}^{\infty}\frac{t^n}{n!}\mathbf{K}^n\mathbf{T}^{-1}\mathbf{T}\mathbf{X}(0) = \mathbf{T}^{-1}\sum_{n=0}^{\infty}\frac{t^n}{n!}\mathbf{T}\mathbf{K}^n\mathbf{T}^{-1}\mathbf{T}\mathbf{X}(0) \\
&= \mathbf{T}^{-1}\left(\mathbf{I} + \sum_{n=1}^{\infty}\frac{t^n}{n!}\mathbf{T}\mathbf{K}^n\mathbf{T}^{-1}\right)\mathbf{T}\mathbf{X}(0) = \mathbf{T}^{-1}\left(\mathbf{I} + \sum_{n=1}^{\infty}\frac{t^n}{n!}\mathbf{T}\mathbf{K}^{n-1}\mathbf{T}^{-1}\mathbf{T}\mathbf{K}\mathbf{T}^{-1}\right)\mathbf{T}\mathbf{X}(0) \\
&= \mathbf{T}^{-1}\left(\mathbf{I} + \sum_{n=1}^{\infty}\frac{t^n}{n!}\mathbf{T}\mathbf{K}^{n-1}\mathbf{T}^{-1}(\lambda\mathbf{J})\right)\mathbf{T}\mathbf{X}(0) = \mathbf{T}^{-1}\left(\mathbf{I} + \frac{\lambda t}{1!}\mathbf{J} + \sum_{n=2}^{\infty}\frac{t^n}{n!}\mathbf{T}\mathbf{K}^{n-2}\mathbf{K}\mathbf{T}^{-1}(\lambda\mathbf{J})\right)\mathbf{T}\mathbf{X}(0) \\
&= \mathbf{T}^{-1}\left(\mathbf{I} + \frac{\lambda t}{1!}\mathbf{J} + \sum_{n=2}^{\infty}\frac{t^n}{n!}\mathbf{T}\mathbf{K}^{n-2}\mathbf{T}^{-1}\mathbf{T}\mathbf{K}\mathbf{T}^{-1}(\lambda\mathbf{J})\right)\mathbf{T}\mathbf{X}(0) = \mathbf{T}^{-1}\sum_{n=0}^{\infty}\frac{t^n}{n!}(\lambda\mathbf{J})^n\mathbf{T}\mathbf{X}(0) = \mathbf{T}^{-1}\mathrm{e}^{\lambda\mathbf{J}t}\mathbf{T}\mathbf{X}(0)
\end{aligned}$$

where $\mathbf{T}$ is the required similarity transformation. Therefore

$$\mathbf{T}\mathbf{X}(t) = \mathrm{e}^{\lambda\mathbf{J}t}\mathbf{T}\mathbf{X}(0)$$

Denote

$$\mathbf{Y}(t) = \mathbf{T}\mathbf{X}(t)$$

We obtain

$$\mathbf{Y}(t) = \mathrm{e}^{\lambda \mathbf{J} t}\mathbf{Y}(0)$$

From (2) of Lemma 4.1, suppose μ is an eigenvalue of $(\mathbf{I} - \Theta)\mathbf{P}$, then $\lambda(\mu - 1)$ is an eigenvalue of $\mathbf{K}$, and $\mu - 1$ is an eigenvalue of $\mathbf{J}$. Further, from Lemma 4.2, the real part of an arbitrary eigenvalue of $\mathbf{J}$ is negative. In this way $\mathbf{Y}(t) \to 0$ as $t \to \infty$. This implies $\mathbf{X}(t) \to 0$ exponentially as $t \to \infty$. □

Proof of Proposition 5.1

First we have

$$\begin{aligned} EM_k &= N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^j \bullet \mathbf{1}^{\mathrm{T}} \geqslant N\sum_{j=0}^{k-1}\mathbf{p_0}\mathbf{P}^{k-j-1}\Theta[\mathbf{P}(1-\theta_1)]^j \bullet \mathbf{1}^{\mathrm{T}} \\ &= N\sum_{j=0}^{k-1}(1-\theta_1)^j\mathbf{p_0}\mathbf{P}^{k-j-1}\Theta\mathbf{P}^j \bullet \mathbf{1}^{\mathrm{T}} \geqslant N\sum_{j=0}^{k-1}(1-\theta_1)^j\theta_m\mathbf{p_0}\mathbf{P}^{k-j-1}\mathbf{P}^j \bullet \mathbf{1}^{\mathrm{T}} = N\sum_{j=0}^{k-1}(1-\theta_1)^j\theta_m \\ &= \frac{\theta_m}{\theta_1}N\left(1-(1-\theta_1)^k\right) \end{aligned}$$

Note that the quantity $N(1 - (1 - \theta_1)^k)$ represents the expected number of failures observed in the first k testing actions if each action has probability of θ_1 revealing a failure.

Similarly, we can obtain

$$EM_k \leqslant \frac{\theta_1}{\theta_m}N\left(1-(1-\theta_m)^k\right)$$

On the other hand, from the fact that

$$EM(t) = \sum_{k=0}^{\infty}EM_k\,\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!}$$

We obtain

$$\frac{\theta_m}{\theta_1}N\left(1-\mathrm{e}^{-\theta_1\lambda t}\right) \leqslant EM(t) \leqslant \frac{\theta_1}{\theta_m}N\left(1-\mathrm{e}^{-\theta_m\lambda t}\right). \qquad \square$$

Proof of Proposition 5.2

First from Proposition 3.2 we have

$$\begin{aligned} EM_k &= N\mathbf{p_0}\left(\mathbf{I}-(\mathbf{I}-\Theta)[\mathbf{P}(\mathbf{I}-\Theta)]^{k-1}\right) \bullet \mathbf{1}^{\mathrm{T}} = N - N\mathbf{p_0}(\mathbf{I}-\Theta)[\mathbf{P}(\mathbf{I}-\Theta)]^{k-1} \bullet \mathbf{1}^{\mathrm{T}} \\ &\geqslant N - N\mathbf{p_0}(\mathbf{I}-\Theta)(1-\theta_m)^{k-1} \bullet \mathbf{P}^{k-1} \bullet \mathbf{1}^{\mathrm{T}} = N - N\mathbf{p_0}(\mathbf{I}-\Theta)(1-\theta_m)^{k-1} \bullet \mathbf{1}^{\mathrm{T}} \\ &= N - N\left(1-\sum_{l=1}^{m}\theta_l p_l\right)(1-\theta_m)^{k-1} \end{aligned}$$

Similarly,

$$EM_k \leqslant N - N\left(1-\sum_{l=1}^{m}\theta_l p_l\right)(1-\theta_1)^{k-1}$$

Then from

$$EM(t)=\sum_{k=0}^{\infty}EM_k\,\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!}$$

We obtain

$$\begin{aligned}EM(t)&\geqslant\sum_{k=0}^{\infty}\left[N-N\left(1-\sum_{l=1}^{m}\theta_l p_l\right)(1-\theta_m)^{k-1}\right]\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!}\\&=N-N\left(1-\sum_{l=1}^{m}\theta_l p_l\right)\sum_{k=0}^{\infty}(1-\theta_m)^{k-1}\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!}=N-\frac{N\left(1-\sum_{l=1}^{m}\theta_l p_l\right)}{1-\theta_m}\sum_{k=0}^{\infty}(1-\theta_m)^{k}\mathrm{e}^{-\lambda t}\frac{(\lambda t)^k}{k!}\\&=N-N\left(1-\sum_{l=1}^{m}\theta_l p_l\right)\mathrm{e}^{-\lambda t}\frac{\mathrm{e}^{\lambda t(1-\theta_m)}}{1-\theta_m}=N-N\left(1-\sum_{l=1}^{m}\theta_l p_l\right)\frac{\mathrm{e}^{-\lambda t\theta_m}}{1-\theta_m}\end{aligned}$$

In a similar way, we have

$$EM(t)\leqslant N-N\left(1-\sum_{l=1}^{m}\theta_l p_l\right)\frac{\mathrm{e}^{-\lambda t\theta_1}}{1-\theta_1}$$

Therefore,

$$\frac{\mathrm{e}^{-\theta_1\lambda t}}{1-\theta_1}\leqslant\frac{N-EM(t)}{N\left(1-\sum_{l=1}^{m}p_l\theta_l\right)}\leqslant\frac{\mathrm{e}^{-\theta_m\lambda t}}{1-\theta_m}.\qquad\square$$

Proof of Proposition 5.3

From Proposition 3.2, it holds that

$$g(t)=N-N\mathbf{p_0}\,\mathrm{e}^{\mathbf{K}t}\bullet\mathbf{1}^{\mathrm{T}}$$

Then

$$g'(t)=\frac{\mathrm{d}}{\mathrm{d}t}g(t)=-N\mathbf{p_0}\,\mathrm{e}^{\mathbf{K}t}\mathbf{K}\bullet\mathbf{1}^{\mathrm{T}}$$
$$g''(t)=\frac{\mathrm{d}^2}{\mathrm{d}t^2}g(t)=N\mathbf{p_0}\mathbf{K}\,\mathrm{e}^{\mathbf{K}t}\mathbf{K}\bullet\mathbf{1}^{\mathrm{T}}$$

Note that

$$\mathbf{K}\bullet\mathbf{1}^{\mathrm{T}}=-\lambda[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]\bullet\mathbf{1}^{\mathrm{T}}=-\lambda\vartheta$$

where

$$\vartheta=\begin{pmatrix}\theta_1\\\theta_2\\\vdots\\\theta_m\end{pmatrix}$$

In this way

$$g''(t)=\lambda-N\mathbf{p_0}\mathbf{K}\,\mathrm{e}^{\mathbf{K}t}\mathbf{K}\bullet\mathbf{1}^{\mathrm{T}}=N\lambda\mathbf{p_0}\mathbf{K}\,\mathrm{e}^{\mathbf{K}t}\vartheta=-N\lambda^2\mathbf{p_0}[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]\mathrm{e}^{\mathbf{K}t}\vartheta$$

Since $\mathbf{p_0}\mathbf{P}=\mathbf{p_0}$, we have

$$\begin{aligned}g''(t)&=-N\lambda^2\mathbf{p_0}[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]\mathrm{e}^{\mathbf{K}t}\vartheta=-N\lambda^2\mathbf{p_0}[\mathbf{I}-\mathbf{P}+\Theta\mathbf{P}]\mathrm{e}^{\mathbf{K}t}\vartheta=-N\lambda^2[\mathbf{p_0}-\mathbf{p_0}\mathbf{P}+\mathbf{p_0}\Theta\mathbf{P}]\mathrm{e}^{\mathbf{K}t}\vartheta\\&=-N\lambda^2\mathbf{p_0}\Theta\mathbf{P}\,\mathrm{e}^{\mathbf{K}t}\vartheta<0.\qquad\square\end{aligned}$$

Proof of Proposition 5.4

Since $\theta_1 = \theta_2 = \cdots = \theta_m = \theta \neq 0$, we have

$$\vartheta = \theta \mathbf{1}^{\mathrm{T}}$$
$$\mathbf{K} \bullet \mathbf{1}^{\mathrm{T}} = -\lambda\theta\mathbf{1}^{\mathrm{T}}$$

Then

$$\begin{aligned} g''(t) &= -N\lambda^2 \mathbf{p_0}[\mathbf{I} - (\mathbf{I} - \Theta)\mathbf{P}]\mathrm{e}^{\mathbf{K}t}\vartheta = -N\lambda^2 \mathbf{p_0}[\mathbf{I} - \mathbf{P} + \Theta\mathbf{P}]\mathrm{e}^{\mathbf{K}t}\theta \bullet \mathbf{1}^{\mathrm{T}} = -N\lambda^2\theta \mathbf{p_0}[\mathbf{I} - \mathbf{P} + \Theta\mathbf{P}]\sum_{i=0}^{\infty}\frac{t^i}{i!}\mathbf{K}^i \bullet \mathbf{1}^{\mathrm{T}} \\ &= -N\lambda^2\theta \mathbf{p_0}[\mathbf{I} - \mathbf{P} + \Theta\mathbf{P}]\sum_{i=0}^{\infty}\frac{t^i}{i!}(-\lambda\theta)^i \bullet \mathbf{1}^{\mathrm{T}} = -N\lambda^2\theta \mathbf{p_0}[\mathbf{I} - \mathbf{P} + \Theta\mathbf{P}]\mathrm{e}^{-\lambda\theta t} \bullet \mathbf{1}^{\mathrm{T}} \\ &= -N\lambda^2\theta \mathbf{p_0}\left[\mathbf{1}^{\mathrm{T}} - \mathbf{1}^{\mathrm{T}} + \Theta \bullet \mathbf{1}^{\mathrm{T}}\right]\mathrm{e}^{-\lambda\theta t} = -N\lambda^2\theta^2\,\mathrm{e}^{-\lambda\theta t}\mathbf{p_0} \bullet \mathbf{1}^{\mathrm{T}} = -N\lambda^2\theta^2\,\mathrm{e}^{-\lambda\theta t} < 0. \qquad \square \end{aligned}$$

Proof of Proposition 6.1

It holds that

$$\begin{aligned} E[M(t) - EM(t)]^2 &= E(M(t))^2 - (EM(t))^2 = E[(M(t) - EM(t))(M(t) + EM(t))] \\ &\leqslant E\Big[(M(t) - EM(t))(M(t) + EM(t))\chi_{\{M(t) \geqslant EM(t)\}}\Big] \leqslant 2NE\Big[(M(t) - EM(t))\chi_{\{M(t) \geqslant EM(t)\}}\Big] \\ &= 2N\Big[EM(t)\chi_{\{M(t) \geqslant EM(t)\}} - E\Big(EM(t)\chi_{\{M(t) \geqslant EM(t)\}}\Big)\Big] \\ &= 2N\Big[EM(t)\chi_{\{M(t) \geqslant EM(t)\}} - EM(t)\Pr\{M(t) \geqslant EM(t)\}\Big] \\ &\leqslant 2N\Big[NE\chi_{\{M(t) \geqslant EM(t)\}} - EM(t)\Pr\{M(t) \geqslant EM(t)\}\Big] \\ &= 2N[N\Pr\{M(t) \geqslant EM(t)\} - EM(t)\Pr\{M(t) \geqslant EM(t)\}] \\ &= 2N[N - EM(t)]\Pr\{M(t) \geqslant EM(t)\} \end{aligned}$$

where

$$\chi_{\{M(t) \geqslant EM(t)\}}(\omega) = \begin{cases} 1 & \text{if } \omega \in \{M(t) \geqslant EM(t)\} \\ 0 & \text{otherwise} \end{cases}$$

Recall that

$$EM(t) = N - N\mathbf{p}_0\,\mathrm{e}^{-\lambda[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]t} \bullet \mathbf{1}^{\mathrm{T}}$$

Then

$$\begin{aligned} E[M(t) - EM(t)]^2 &\leqslant 2N[N - EM(t)]\Pr\{M(t) \geqslant EM(t)\} = 2N^2\left[\mathbf{p}_0\,\mathrm{e}^{-\lambda[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]t} \bullet \mathbf{1}^{\mathrm{T}}\right]\Pr\{M(t) \geqslant EM(t)\} \\ &\leqslant 2N^2\mathbf{p}_0\,\mathrm{e}^{-\lambda[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]t} \bullet \mathbf{1}^{\mathrm{T}}. \qquad \square \end{aligned}$$

Proof of Proposition 6.2

From $\theta_1 = \theta_2 = \cdots = \theta_m = \theta$ and Proposition 5.2, we have

$$EM_k = N\left(1 - (1 - \theta)^k\right)$$

and

$$\begin{aligned}EM_k^2 &= N\theta + \theta(2N-1)(EM_{k-1}) + (1-2\theta)\left(EM_{k-1}^2\right)\\ &= N\theta + \theta(2N-1)N\left(1-(1-\theta)^{k-1}\right) + (1-2\theta)\left(EM_{k-1}^2\right)\\ &= 2\theta N^2 - N\theta(2N-1)(1-\theta)^{k-1} + (1-2\theta)\left(EM_{k-1}^2\right)\end{aligned}$$

Denote

$$\begin{aligned}U_k &= 2\theta N^2 - N\theta(2N-1)(1-\theta)^k\\ U_0 &= N\theta\end{aligned}$$

Then for $2\theta < 1$,

$$\begin{aligned}EM_k^2 &= U_{k-1} + (1-2\theta)\left(EM_{k-1}^2\right) = U_{k-1} + (1-2\theta)\left[U_{k-2} + (1-2\theta)\left(EM_{k-2}^2\right)\right]\\ &= U_{k-1} + (1-2\theta)U_{k-2} + (1-2\theta)^2\left(EM_{k-2}^2\right) = \sum_{l=1}^{k} U_{l-1}(1-2\theta)^{k-l}\\ &= \sum_{l=1}^{k}(1-2\theta)^{k-l}\left[2\theta N^2 - N\theta(2N-1)(1-\theta)^{l-1}\right]\\ &= \sum_{l=1}^{k}(1-2\theta)^{k-l}\left(2\theta N^2\right) - \sum_{l=1}^{k} N\theta(2N-1)(1-2\theta)^{k-l}(1-\theta)^{l-1}\\ &= N^2 - N^2(1-2\theta)^k - N\theta(2N-1)\sum_{l=1}^{k}(1-2\theta)^{k-l}(1-\theta)^{l-1}\end{aligned}$$

Note that

$$\begin{aligned}\sum_{l=1}^{k}(1-2\theta)^{k-l}(1-\theta)^{l-1} &= \frac{(1-2\theta)^k}{1-\theta}\sum_{l=1}^{k}\left(\frac{1-\theta}{1-2\theta}\right)^l = \frac{(1-2\theta)^k}{1-\theta}\left[\frac{1-\theta-\frac{(1-\theta)^{k+1}}{(1-2\theta)^k}}{-\theta}\right]\\ &= \frac{1}{\theta}\left[(1-\theta)^k - (1-2\theta)^k\right]\end{aligned}$$

We arrive at

$$\begin{aligned}EM_k^2 &= N^2 - N^2(1-2\theta)^k - N\theta(2N-1)\frac{1}{\theta}\left[(1-\theta)^k - (1-2\theta)^k\right]\\ &= N^2 - N^2(1-2\theta)^k - N(2N-1)\left[(1-\theta)^k - (1-2\theta)^k\right]\\ &= N^2 + N(N-1)(1-2\theta)^k - N(2N-1)(1-\theta)^k\end{aligned}$$

Recall that

$$EM(t) = N - N\mathbf{p_0}\,\mathrm{e}^{-\lambda t[\mathbf{I}-(\mathbf{I}-\Theta)\mathbf{P}]} \bullet \mathbf{1}^{\mathrm{T}}$$

We have

$$\begin{aligned}EM(t) &= N - N\,\mathrm{e}^{-\lambda\theta t}\\ (EM(t))^2 &= N^2 - 2N^2\,\mathrm{e}^{-\lambda\theta t} + N^2\,\mathrm{e}^{-2\lambda\theta t}\end{aligned}$$

Further,

$$\begin{aligned}E(M(t))^2 &= \sum_{k=0}^{\infty} EM_k^2\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t} = \sum_{k=0}^{\infty}\left(N^2 + N(N-1)(1-2\theta)^k - N(2N-1)(1-\theta)^k\right)\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}\\ &= N^2 + N(N-1)\mathrm{e}^{-\lambda t}\,\mathrm{e}^{\lambda(1-2\theta)t} - N(2N-1)\mathrm{e}^{-\lambda t}\,\mathrm{e}^{\lambda(1-\theta)t} = N^2 + N(N-1)\mathrm{e}^{-2\lambda\theta t} - N(2N-1)\mathrm{e}^{-\lambda\theta t}\end{aligned}$$

In this way

$$\begin{aligned} E(M(t))^2 - (EM(t))^2 &= N^2 + N(N-1)\mathrm{e}^{-2\lambda\theta t} - N(2N-1)\mathrm{e}^{-\lambda\theta t} - N^2 + 2N^2\,\mathrm{e}^{-\lambda\theta t} - N^2\,\mathrm{e}^{-2\lambda\theta t} \\ &= N\left[\mathrm{e}^{-\lambda\theta t} - \mathrm{e}^{-2\lambda\theta t}\right]. \qquad \square \end{aligned}$$

Proof of Proposition 6.3

For $k \geqslant N+1$,

$$\begin{aligned} E\left[M_k^2|A_k = i\right]\Pr\{A_k = i\} &= \sum_{j=0}^{N} j^2 \Pr\{M_k = j, A_k = i\} \\ &= \sum_{j=1}^{N}\sum_{l=1}^{m} j^2[\Pr\{M_k = j, A_k = i; M_{k-1} = j, A_{k-1} = l\} \\ &\quad + \Pr\{M_k = j, A_k = i; M_{k-1} = j-1, A_{k-1} = l\}] \\ &= \sum_{j=1}^{N}\sum_{l=1}^{m} j^2[(1-(N-j)\theta_i)p_{li}\Pr\{M_{k-1} = j, A_{k-1} = l\} \\ &\quad + (N-j+1)\theta_i p_{li}\Pr\{M_{k-1} = j-1, A_{k-1} = l\}] \\ &= \sum_{l=1}^{m}\sum_{j=1}^{N}[\left(j^2(1-N\theta_i) + j^3\theta_i\right)p_{li}\Pr\{M_{k-1} = j, A_{k-1} = l\} \\ &\quad + \left((N+1)j^2 - j^3\right)\theta_i p_{li}\Pr\{M_{k-1} = j-1, A_{k-1} = l\}] \end{aligned}$$

Note that

$$\begin{aligned} &(1-N\theta_i)\sum_{j=1}^{N} j^2 p_{li}\Pr\{M_{k-1} = j, A_{k-1} = l\} \\ &\qquad = (1-N\theta_i)p_{li}E\left[M_{k-1}^2|A_{k-1} = l\right]\Pr\{A_{k-1} = l\} \\ &\theta_i p_{li}\sum_{j=1}^{N} j^3 \Pr\{M_{k-1} = j, A_{k-1} = l\} \\ &\qquad = \theta_i p_{li}E\left[M_{k-1}^3|A_{k-1} = l\right]\Pr\{A_{k-1} = l\} \\ &(N+1)\theta_i p_{li}\sum_{j=1}^{N} j^2 \Pr\{M_{k-1} = j-1, A_{k-1} = l\} \\ &\qquad = (N+1)\theta_i p_{li}\sum_{j=0}^{N-1}(j+1)^2\Pr\{M_{k-1} = j, A_{k-1} = l\} \\ &\qquad = (N+1)\theta_i p_{li}\sum_{j=0}^{N}\left(j^2+2j+1\right)\Pr\{M_{k-1} = j, A_{k-1} = l\} \\ &\qquad\quad - (N+1)\theta_i p_{li}(N+1)^2\Pr\{M_{k-1} = N, A_{k-1} = l\} \\ &\qquad = (N+1)\theta_i p_{li}(E\left[M_{k-1}^2|A_{k-1} = l\right] + 2E[M_{k-1}|A_{k-1} = l] \\ &\qquad\quad + 1 - (N+1)^2\Pr\{M_{k-1} = N|A_{k-1} = l\})\Pr\{A_{k-1} = l\} \end{aligned}$$

and

$$\begin{aligned}\theta_i p_{li}\sum_{j=1}^{N} j^3\Pr\{M_{k-1}=j-1,A_{k-1}=l\} &= \theta_i p_{li}\sum_{j=0}^{N-1}(j+1)^3\Pr\{M_{k-1}=j,A_{k-1}=l\}\\ &= \theta_i p_{li}\sum_{j=0}^{N}\left(j^3+3j^2+3j+1\right)\Pr\{M_{k-1}=j,A_{k-1}=l\}\\ &\quad -\theta_i p_{li}(N+1)^3\Pr\{M_{k-1}=N,A_{k-1}=l\}\\ &= \theta_i p_{li}(E\left[M_{k-1}^3|A_{k-1}=l\right]+3E\left[M_{k-1}^2|A_{k-1}=l\right]+3E[M_{k-1}|A_{k-1}=l]\\ &\quad +1-(N+1)^3\Pr\{M_{k-1}=N|A_{k-1}=l\})\Pr\{A_{k-1}=l\}\end{aligned}$$

In this way we have

$$\begin{aligned}&E\left[M_k^2|A_k=i\right]\Pr\{A_k=i\}\\ &\quad=\sum_{j=1}^{N}\sum_{l=1}^{m}[\left(j^2(1-N\theta_i)+j^3\theta_i\right)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\}\\ &\quad\quad+\left((N+1)j^2-j^3\right)\theta_i p_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]\\ &\quad=\sum_{l=1}^{m}[(1-N\theta_i)p_{li}E\left[M_{k-1}^2|A_{k-1}=l\right]+\theta_i p_{li}E\left[M_{k-1}^3|A_{k-1}=l\right]\\ &\quad\quad+(N+1)\theta_i p_{li}(E\left[M_{k-1}^2|A_{k-1}=l\right]+2E[M_{k-1}|A_{k-1}=l]+1\\ &\quad\quad-(N+1)^2\Pr\{M_{k-1}=N|A_{k-1}=l\})-\theta_i p_{li}(E\left[M_{k-1}^3|A_{k-1}=l\right]+3E\left[M_{k-1}^2|A_{k-1}=l\right]\\ &\quad\quad+3E[M_{k-1}|A_{k-1}=l]+1-(N+1)^3\Pr\{M_{k-1}=N|A_{k-1}=l\})]\Pr\{A_{k-1}=l\}\end{aligned}$$

Straightforward derivation of the right-hand side of the above formula leads to

$$\begin{aligned}E\left[M_k^2|A_k=i\right]\Pr\{A_k=i\} &= \sum_{l=1}^{m}[N\theta_i p_{li}+(2N-1)\theta_i p_{li}E[M_{k-1}|A_{k-1}=l]\\ &\quad+(1-2\theta_i)p_{li}E\left[M_{k-1}^2|A_{k-1}=l\right]]\Pr\{A_{k-1}=l\}\end{aligned}$$

For $1\leqslant k\leqslant N$,

$$\begin{aligned}E\left[M_k^2|A_k=i\right]\Pr\{A_k=i\} &= \sum_{j=0}^{k} j^2\Pr\{M_k=j|A_k=i\}\Pr\{A_k=i\}\\ &= \sum_{j=1}^{k-1}\sum_{l=1}^{m} j^2[\Pr\{M_k=j,A_k=i;M_{k-1}=j,A_{k-1}=l\}\\ &\quad+\Pr\{M_k=j,A_k=i;M_{k-1}=j-1,A_{k-1}=l\}]\\ &\quad+\sum_{l=1}^{m}k^2\Pr\{M_k=k,A_k=i;M_{k-1}=k-1,A_{k-1}=l\}\\ &= \sum_{j=1}^{k-1}\sum_{l=1}^{m} j^2[(1-(N-j)\theta_i)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\}\\ &\quad+(N-j+1)\theta_i p_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]\\ &\quad+\sum_{l=1}^{m}k^2(N-k+1)\theta_i p_{li}\Pr\{M_{k-1}=k-1,A_{k-1}=l\}\\ &= \sum_{l=1}^{m}\sum_{j=1}^{k-1}[\left(j^2(1-N\theta_i)+j^3\theta_i\right)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\}\\ &\quad+\left((N+1)j^2-j^3\right)\theta_i p_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]\\ &\quad+\sum_{l=1}^{m}k^2(N-k+1)\theta_i p_{li}\Pr\{M_{k-1}=k-1,A_{k-1}=l\}\end{aligned}$$

Note that

$$
\begin{aligned}
&(1-N\theta_i)p_{li}\sum_{j=1}^{k-1} j^2\Pr\{M_{k-1}=j,A_{k-1}=l\}\\
&\quad=(1-N\theta_i)p_{li}E\left[M_{k-1}^2|A_{k-1}=l\right]\Pr\{A_{k-1}=l\}\\
&\theta_i p_{li}\sum_{j=1}^{k-1} j^3\Pr\{M_{k-1}=j,A_{k-1}=l\}\\
&\quad=\theta_i p_{li}E\left[M_{k-1}^3|A_{k-1}=l\right]\Pr\{A_{k-1}=l\}\\
&(N+1)\theta_i p_{li}\sum_{j=1}^{k-1} j^2\Pr\{M_{k-1}=j-1,A_{k-1}=l\}\\
&\quad=(N+1)\theta_i p_{li}\sum_{j=0}^{k-2}(j+1)^2\Pr\{M_{k-1}=j,A_{k-1}=l\}\\
&\quad=(N+1)\theta_i p_{li}\sum_{j=0}^{k-1}\left(j^2+2j+1\right)\Pr\{M_{k-1}=j,A_{k-1}=l\}\\
&\qquad-(N+1)\theta_i p_{li}k^2\Pr\{M_{k-1}=k-1,A_{k-1}=l\}\\
&\quad=(N+1)\theta_i p_{li}(E\left[M_{k-1}^2|A_{k-1}=l\right]+2E[M_{k-1}|A_{k-1}=l]\\
&\qquad+1-k^2\Pr\{M_{k-1}=k-1|A_{k-1}=l\})\Pr\{A_{k-1}=l\}
\end{aligned}
$$

and

$$
\begin{aligned}
\theta_i p_{li}\sum_{j=1}^{k-1} j^3\Pr\{M_{k-1}=j-1,A_{k-1}=l\}&=\theta_i p_{li}\sum_{j=0}^{k-2}(j+1)^3\Pr\{M_{k-1}=j,A_{k-1}=l\}\\
&=\theta_i p_{li}\sum_{j=0}^{k-1}\left(j^3+3j^2+3j+1\right)\Pr\{M_{k-1}=j,A_{k-1}=l\}\\
&\quad-\theta_i p_{li}k^3\Pr\{M_{k-1}=k-1,A_{k-1}=l\}=\theta_i p_{li}(E\left[M_{k-1}^3|A_{k-1}=l\right]\\
&\quad+3E\left[M_{k-1}^2|A_{k-1}=l\right]+3E[M_{k-1}|A_{k-1}=l]\\
&\quad+1-k^3\Pr\{M_{k-1}=k-1|A_{k-1}=l\})\Pr\{A_{k-1}=l\}
\end{aligned}
$$

Therefore we arrive at

$$
\begin{aligned}
E\left[M_k^2|A_k=i\right]\Pr\{A_k=i\}&=\sum_{l=1}^{m}\sum_{j=1}^{k-1}[\left(j^2(1-N\theta_i)+j^3\theta_i\right)p_{li}\Pr\{M_{k-1}=j,A_{k-1}=l\}\\
&\quad+\left((N+1)j^2-j^3\right)\theta_i p_{li}\Pr\{M_{k-1}=j-1,A_{k-1}=l\}]\\
&\quad+\sum_{l=1}^{m}k^2(N-k+1)\theta_i p_{li}\Pr\{M_{k-1}=k-1,A_{k-1}=l\}\\
&=\sum_{l=1}^{m}(1-N\theta_i)p_{li}E\left[M_{k-1}^2|A_{k-1}=l\right]\Pr\{A_{k-1}=l\}\\
&\quad+\sum_{l=1}^{m}\theta_i p_{li}E\left[M_{k-1}^3|A_{k-1}=l\right]\Pr\{A_{k-1}=l\}\\
&\quad+\sum_{l=1}^{m}(N+1)\theta_i p_{li}(E\left[M_{k-1}^2|A_{k-1}=l\right]\\
&\quad+2E[M_{k-1}|A_{k-1}=l]+1-k^2\Pr\{M_{k-1}=k-1|A_{k-1}=l\})\Pr\{A_{k-1}=l\}\\
&\quad-\sum_{l=1}^{m}\theta_i p_{li}(E\left[M_{k-1}^3|A_{k-1}=l\right]+3E\left[M_{k-1}^2|A_{k-1}=l\right]+3E[M_{k-1}|A_{k-1}=l]\\
&\quad+1-k^3\Pr\{M_{k-1}=k-1|A_{k-1}=l\})\Pr\{A_{k-1}=l\}\\
&\quad+\sum_{l=1}^{m}k^2(N-k+1)\theta_i p_{li}\Pr\{M_{k-1}=k-1,A_{k-1}=l\}
\end{aligned}
$$

Or

$$E\left[M_k^2|A_k=i\right]\Pr\{A_k=i\}=\sum_{l=1}^{m}[N\theta_i p_{li}+(2N-1)\theta_i p_{li}E[M_{k-1}|A_{k-1}=l]$$
$$+(1-2\theta_i)p_{li}E\left[M_{k-1}^2|A_{k-1}=l\right]]\Pr\{A_{k-1}=l\}.\qquad\square$$

Proof of Proposition 6.4

Recall that

$$\gamma_k=N\beta_{k-1}\Psi+(2N-1)\alpha_{k-1}\Psi+\gamma_{k-1}H_{\langle 2\rangle}$$

Denote

$$V_k=N\beta_k\Psi+(2N-1)\alpha_k\Psi$$
$$V_0=N\mathbf{p}_0\Psi$$

We have

$$\begin{aligned}\gamma_k&=V_{k-1}+\gamma_{k-1}H_{\langle 2\rangle}=V_{k-1}+\left(V_{k-2}+\gamma_{k-2}H_{\langle 2\rangle}\right)H_{\langle 2\rangle}=V_{k-1}+V_{k-2}H_{\langle 2\rangle}+V_{k-3}H_{\langle 2\rangle}^2+\cdots+V_0H_{\langle 2\rangle}^{k-1}\\&=\sum_{j=1}^{k}V_{k-j}H_{\langle 2\rangle}^{j-1}=\sum_{j=1}^{k}\left[N\beta_{k-j}\Psi+(2N-1)\alpha_{k-j}\Psi\right]H_{\langle 2\rangle}^{j-1}\\&=\sum_{j=1}^{k}\left[N\mathbf{p}_0\mathbf{P}^{k-j}\Theta+(2N-1)\alpha_{k-j}\mathbf{P}\Theta\right][\mathbf{P}(\mathbf{I}-2\Theta)]^{j-1}\end{aligned}$$

Since

$$\alpha_k=N\sum_{j=0}^{k-1}\beta_{k-1-j}\Psi\mathbf{H}^j=N\sum_{j=0}^{k-1}\mathbf{p}_0\mathbf{P}^{k-j-1}\Theta[\mathbf{P}(\mathbf{I}-\Theta)]^j$$

We arrive at

$$\gamma_k=\sum_{j=1}^{k}\left[N\mathbf{p}_0\mathbf{P}^{k-j}\Theta+(2N-1)N\sum_{l=0}^{k-j-1}\mathbf{p}_0\mathbf{P}^{k-j-1-l}\Theta[\mathbf{P}(\mathbf{I}-\Theta)\mathbf{P}]^l\mathbf{P}\Theta\right][\mathbf{P}(\mathbf{I}-2\Theta)]^{j-1}.\qquad\square$$

Proof of Proposition 6.5

For $k\geqslant N+1$, and parameter $\mu>0$,

$$\begin{aligned}E\left[\left(\mathrm{e}^{\mu M_k}-1\right),A_k=i\right]&=\sum_{j=0}^{N}(\mathrm{e}^{\mu j}-1)\Pr\{M_k=j,A_k=i\}\\&=\sum_{j=1}^{N}(\mathrm{e}^{\mu j}-1)\Pr\{M_k=j,A_k=i\}\\&=\sum_{j=1}^{N}\sum_{l=1}^{m}(\mathrm{e}^{\mu j}-1)(\Pr\{M_k=j,A_k=i;M_{k-1}=j,A_{k-1}=l\}\\&\quad+\Pr\{M_k=j,A_k=i;M_{k-1}=j-1,A_{k-1}=l\})\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=1}^{N} \sum_{l=1}^{m} (e^{\mu j} - 1)((1 - (N-j)\theta_i)p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\} \\
&\quad + (N - j + 1)\theta_i p_{li} \Pr\{M_{k-1} = j-1, A_{k-1} = l\}) \\
&= \sum_{j=1}^{N} \sum_{l=1}^{m} (e^{\mu j} - 1)(1 - (N-j)\theta_i)p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\} \\
&\quad + \sum_{j=1}^{N} \sum_{l=1}^{m} (e^{\mu j} - 1)(N - j + 1)\theta_i p_{li} \Pr\{M_{k-1} = j-1, A_{k-1} = l\} \\
&= \sum_{j=1}^{N} \sum_{l=1}^{m} (e^{\mu j} - 1)(1 - (N-j)\theta_i)p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\} \\
&\quad + \sum_{j=0}^{N-1} \sum_{l=1}^{m} \left(e^{\mu(j+1)} - 1\right)(N - j)\theta_i p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\} \\
&= \sum_{j=0}^{N} \sum_{l=1}^{m} (e^{\mu j} - 1)(1 - (N-j)\theta_i)p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\} \\
&\quad + \sum_{j=0}^{N} \sum_{l=1}^{m} \left(e^{\mu(j+1)} - 1\right)(N - j)\theta_i p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\} \\
&= \sum_{j=0}^{N} \sum_{l=1}^{m} \left(e^{\mu(j+1)}(N-j)\theta_i + e^{\mu j}(1 - (N-j)\theta_i) - 1\right)p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\}
\end{aligned}
$$

Let

$$f_{j,i}(\mu) = e^{\mu(j+1)}(N-j)\theta_i + e^{\mu j}(1 - (N-j)\theta_i) - 1$$

Then

$$E\left[\left(e^{\mu M_k} - 1\right), A_k = i\right] = \sum_{j=0}^{N} \sum_{l=1}^{m} f_{j,i}(\mu) p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\}$$

Note that

$$
\begin{aligned}
\frac{\partial^n}{\partial \mu^n} E\left[\left(e^{\mu M_k} - 1\right), A_k = i\right] &= \sum_{j=0}^{N} \sum_{l=1}^{m} \frac{\partial^n}{\partial \mu^n} f_{j,i}(\mu) p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\} \\
&= \sum_{j=0}^{N} \sum_{l=1}^{m} f_{j,i}^{(n)}(\mu) p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\}
\end{aligned}
$$

where for positive integer $n \geqslant 1$,

$$
\begin{aligned}
f_{j,i}^{(n)}(\mu) &= \frac{\partial^n}{\partial \mu^n} f_{j,i}(\mu) = (j+1)^n (N-j)\theta_i e^{\mu(j+1)} + j^n (1 - (N-j)\theta_i) e^{\mu j} \\
f_{j,i}^{(n)}(0) &= (j+1)^n (N-j)\theta_i + j^n (1 - (N-j)\theta_i) = [(j+1)^n - j^n](N-j)\theta_i + j^n \\
&= \sum_{s=0}^{n-1} C_n^s j^s (N-j)\theta_i + j^n = \left[N \sum_{s=0}^{n-1} C_n^s j^s - \sum_{s=0}^{n-1} C_n^s j^{s+1}\right]\theta_i + j^n \\
&= \left[N \sum_{s=0}^{n-1} C_n^s j^s - \sum_{s=1}^{n} C_n^{s-1} j^s\right]\theta_i + j^n = \left[\sum_{s=1}^{n-1} \left(N C_n^s - C_n^{s-1}\right) j^s + N - n j^n\right]\theta_i + j^n
\end{aligned}
$$

In this way,

$$
\begin{aligned}
E[M_k^n, A_k = i] &= \frac{\partial^n}{\partial \mu^n} E\left[\left(\mathrm{e}^{\mu M_k} - 1\right), A_k = i\right]|_{\mu=0} \\
&= \sum_{j=0}^{N} \sum_{l=1}^{m} f_{j,i}^{(n)}(0) p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\} \\
&= \sum_{j=0}^{N} \sum_{l=1}^{m} \left(\left[\sum_{s=1}^{n-1} \left(N C_n^s - C_n^{s-1} \right) j^s + N - n j^n \right] \theta_i + j^n \right) p_{li} \Pr\{M_{k-1} = j, A_{k-1} = l\} \\
&= \sum_{l=1}^{m} p_{li} \left(\sum_{j=0}^{N} \sum_{s=1}^{n-1} \left(N C_n^s - C_n^{s-1} \right) j^s \theta_i \Pr\{M_{k-1} = j, A_{k-1} = l\} \right. \\
&\quad + \sum_{j=0}^{N} N \theta_i \Pr\{M_{k-1} = j, A_{k-1} = l\} - \sum_{j=0}^{N} n j^n \theta_i \Pr\{M_{k-1} = j, A_{k-1} = l\} \\
&\quad \left. + \sum_{j=0}^{N} j^n \Pr\{M_{k-1} = j, A_{k-1} = l\} \right) \\
&= \sum_{l=1}^{m} p_{li} \left(\sum_{s=1}^{n-1} \left(N C_n^s - C_n^{s-1} \right) \theta_i \sum_{j=0}^{N} j^s \Pr\{M_{k-1} = j, A_{k-1} = l\} \right. \\
&\quad \left. + N \theta_i \Pr\{A_{k-1} = l\} - n \theta_i E\left[M_{k-1}^n, A_{k-1} = l\right] + E\left[M_{k-1}^n, A_{k-1} = l\right] \right) \\
&= \sum_{l=1}^{m} p_{li} \left(\sum_{s=1}^{n-1} \left(N C_n^s - C_n^{s-1} \right) \theta_i E\left[M_{k-1}^s, A_{k-1} = l\right] + N \theta_i \Pr\{A_{k-1} = l\} \right. \\
&\quad \left. + (1 - n\theta_i) E\left[M_{k-1}^n, A_{k-1} = l\right] \right) \\
&= \sum_{l=1}^{m} N \theta_i p_{li} \Pr\{A_{k-1} = l\} + \sum_{s=1}^{n-1} \left(N C_n^s - C_n^{s-1} \right) \sum_{l=1}^{m} \theta_i p_{li} E\left[M_{k-1}^s, A_{k-1} = l\right] \\
&\quad + \sum_{l=1}^{m} (1 - n\theta_i) p_{li} E\left[M_{k-1}^n, A_{k-1} = l\right]. \qquad \square
\end{aligned}
$$

References

[1] R.V. Binder, Testing Object-Oriented Systems: Models, Patterns, and Tools, Addison-Wesley, 2000.

[2] L.C. Briand, J. Wust, H. Lounis, Replicated case studies for investigating quality factors in object-oriented design, Empirical Software Engineering 6 (2001) 11–58.

[3] K.Y. Cai, Software Defect and Operational Profile Modeling, Kluwer Academic Publishers, 1998.

[4] K.Y. Cai, Software reliability experimentation and control, Journal of Computer Science and Technology 21 (5) (2006) 697–707.

[5] K.Y. Cai, Z. Dong, K. Liu, C.G. Bai, A mathematical modeling framework for software reliability testing, International Journal of General Systems 36 (4) (2007) 399–463.

[6] K.Y. Cai, B. Gu, H. Hu, Y.C. Li, Adaptive software testing with fixed-memory feedback, Journal of Systems and Software 80 (2007) 1328–1348.

[7] J.W. Cangussu, R.A. DeCarlo, A.P. Mathur, A formal model for the software test process, IEEE Transactions on Software Engineering 27 (8) (2002) 782–796.

[8] J.W. Cangussu, R.A. DeCarlo, A.P. Mathur, Using sensitivity analysis to validate a state variable model of the software test process, IEEE Transactions on Software Engineering 29 (5) (2003) 430–443.

[9] A. Cater-Steel, M. Toleman, T. Rout, Addressing the changes of replication of surveys in software engineering research, in: Proceedings of the International Symposium on Empirical Software Engineering, 2005, pp. 204–213.

[10] C.T. Chen, Linear System Theory and Design, CBS College Publishing, 1984.

[11] S. Dick, C.L. Bethel, A. Kandel, Software-reliability modeling: the case for deterministic behavior, IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Human 37 (1) (2007) 106–119.

[12] R. Findeisen, F. Allgower, An introduction to nonlinear model predictive control, http://www.wfw.wtb.tue.nl/benelux2002/nmpc.pdf.

[13] W.J. Gutjahr, Software dependability evaluation based on Markov usage models, Performance Evaluation 40 (2000) 199–222.

[14] K. Henningsson, C. Wohlin, Assuring fault classification agreement – an empirical evaluation, in: Proceedings of 2004 International Symposium on Empirical Software Engineering, 2004, pp. 95–104.

[15] Lin Huang, Linear Algebra in the Theory of Systems and Control, Science Press, 1984 (Theorem 3.3.5, p. 110).

[16] A. Janes, M. Scotto, W. Pedrycz, B. Russo, M. Stefanovic, G. Succi, Identification of defect-prone classes in telecommunication software systems using design metrics, Information Sciences 176 (2006) 3711–3734.

[17] N. Juristo, A.M. Moreno, S. Vegas, Reviewing 25 years of testing technique experiments, Empirical Software Engineering 9 (2004) 7–44.

[18] H.K. Khalil, Nonlinear Systems, third ed., Prentice-Hall, 2001.

[19] C.H. Lee, Y.T. Kim, D.H. Park, S-shaped software reliability growth models derived from stochastic differential equations, IIE Transactions 36 (2004) 1193–1199.

[20] M.R. Lyu (Ed.), Handbook of Software Reliability Engineering, McGraw-Hill, 1996.

[21] H. Minc, Nonnegative Matrices, Wiley, New York, 1988, pp. 11–19, 24–36.

[22] J.H. Poore, G.H. Walton, J.A. Whittaker, A constraint-based approach to the representation of software usage models, Information and Software Technology 42 (2000) 825–833.

[23] G. Rothermel, S. Elbaum, A.G. Malishevsky, P. Kallakui, X. Qiu, On test suite composition and cost-effective regression testing, ACM Transactions on Software Engineering and Methodology 13 (3) (2004) 277–331.

[24] D.I.K. Sjoberg, J.E. Hannay, O. Hansen, V.B. Kampenes, A. Karahasanovic, N.K. Liborg, A.C. Rekdal, A survey of controlled experiments in software engineering, IEEE Transactions on Software Engineering 31 (9) (2005) 733–753.

[25] J.A. Whittaker, What is software testing? And why is it so hard? IEEE Software (January/February) (2000) 70–79.

[26] J.A. Whittaker, J.H. Poore, Markov analysis of software specifications, ACM Transactions on Software Engineering and Methodology 2 (1993) 93–106.

[27] J.A. Whittaker, M.G. Thomason, A Markov chain model for statistical software testing, IEEE Transactions on Software Engineering 20 (10) (1994) 812–824.

[28] M. Xie, Software Reliability Modeling, World Scientific, 1991.

[29] S. Yamada, M. Kimura, H. Tanaka, S. Osaki, Software reliability measurement and assessment with stochastic differential equations, IEICE Transactions, Fundamentals E77-A (1994) 109–115.