Matrix-Analytic Methods in Stochastic Models

Proceedings of the Tenth International Conference on Matrix-Analytic Methods in Stochastic Models

Edited by:
Sophie Hautphenne, Małgorzata O'Reilly, Federico Poloni

February 13 - 15, 2019
Hobart, Australia

MAM 10

The Tenth International Conference on Matrix-Analytic Methods in Stochastic Models

13-15 February 2019, University of Tasmania, Hobart, Australia

Keynote speakers: Azam Asanjarani Søren Asmussen Jevgenijs Ivanovs Giang Nguyen Zbigniew Palmowski and Phil Pollett

Sponsors: Andrew Bassom Peter Taylor

Aust MS

ACEMS

AMSI
AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

Proceedings

Tenth International Conference on Matrix-Analytic Methods in Stochastic Models

Editors

SOPHIE HAUTPHENNE
The University of Melbourne
Victoria 3010, Australia

MAŁGORZATA O'REILLY
School of Natural Sciences
University of Tasmania
Tasmania 7005, Australia

FEDERICO POLONI
Department of Computer Science
University of Pisa
56127 Pisa, Italy

ISBN: 978-0-646-99707-0 (Print version, soft cover)
ISBN: 978-0-646-99825-1 (Electronic version)

Printed in 2019, Hobart, Australia.

Note from the editors

Over the years, matrix-analytic models have proved to be successful in providing performance measures for a large number of real-world systems. In their corresponding computational methods, known as matrix-analytic methods, algorithmic issues are investigated in detail and the probabilistic interpretation of the proposed numerical procedures plays a major role. These methods have been developed initially in the context of queueing models and have given rise to the theory of quasi-birth-and-death processes and of skip-free Markov chains, both belonging to the class of structured Markov chains. More recently, matrix-analytic methods have been extended further for stochastic fluid queues, branching processes, and Markov-modulated Brownian motion.

The *Tenth International Conference on Matrix-Analytic Methods in Stochastic Models* (MAM10) was held at the University of Tasmania in Hobart from the 13^{th} to the 15^{th} of February 2019, continuing the established tradition of previous fruitful MAM conferences in Flint (1995), Winnipeg (1998), Leuven (2000), Adelaide (2002), Pisa (2005), Beijing (2008), New York (2011), Calicut (2014), and Budapest (2016).

The MAM10 conference was sponsored by the Australian Mathematical Sciences Institute (AMSI), the Australian Mathematical Society (AustMS), the Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS), as well as by Peter Taylor (Australian Laureate Fellow, at the University of Melbourne, Director of ACEMS) and Andrew Bassom (Head of Discipline – Mathematics, University of Tasmania).

MAM conferences aim to bring together researchers working on the theoretical, algorithmic and methodological aspects of these methods and the applications of such mathematical research across a broad spectrum of fields, which includes computer science and engineering, telephony and communication networks, electrical and industrial engineering, operations research, management science, financial and risk analysis, bio-statistics, and evolution.

This book forms the Proceedings of MAM10, and contains 34 abstracts, including 7 extended abstracts. Each extended abstract was reviewed anonymously by two members of the program committee. Keynote talks were given by Azam Asanjarani, Søren Asmussen, Jevgenijs Ivanovs, Giang Nguyen, Zbigniew Palmowski, and Phil Pollett.

We thank our steering committee and program committee members, and all other people who helped in the organization of the conference. We thank the University of Tasmania for hosting the conference, and our sponsors for their financial support. We thank all the authors who contributed to the abstracts, the reviewers, and all the participants. We thank Odyseusz Zawalski for the design of the poster, the MAM logo, and the cover of this book. We thank Kelly Carpenter for the administrative support.

Sophie Hautphenne, Małgorzata O'Reilly, and Federico Poloni

Organising Committee:

Małgorzata O'Reilly, Conference Chair, University of Tasmania, Australia
Sophie Hautphenne, Program Co-Chair, The University of Melbourne, Australia
Federico Poloni, Program Co-Chair, University of Pisa, Italy
Mark Fackrell, The University of Melbourne, Australia
Barbara Holland, University of Tasmania, Australia
Michael Brideson, University of Tasmania, Australia

Steering Committee:

Attahiru S. Alfa, University of Manitoba, Canada, and University of Pretoria, South Africa
Guy Latouche, Universite Libre de Bruxelles, Belgium
Miklós Telek, Technical University of Budapest, Hungary
Peter Taylor, The University of Melbourne, Australia
Qi-Ming He, University of Waterloo, Canada
V. Ramaswami, Statmetrics, LLC, United States

Program Committee:

Søren Asmussen, Aarhus University, Denmark
Nigel Bean, The University of Adelaide, Australia
Peter Braunsteins, The University of Melbourne, Australia
Peter Buchholz, Technische Universität Dortmund, Germany
Giuliano Casale, Imperial College London, UK
Srinivas Chakravarthy, Kettering University, United States
Tuğrul Dayar, Bilkent University, Turkey
Mark Fackrell, The University of Melbourne, Australia
Qi-Ming He, University of Waterloo, Canada
Andras Horvath, Università di Torino, Italy
Gábor Horváth, Budapest University of Technology and Economics, Hungary
Udo Krieger, IEEE, Germany
Barbara Margolius, Cleveland State University, United States
Stefano Massei, École Polytechnique Fédérale de Lausanne, Switzerland
Beatrice Meini, University of Pisa, Italy
Masakiyo Miyazawa, Tokyo University of Science, Japan
Yoni Nazarathy, The University of Queensland, Australia
Giang Nguyen, The University of Adelaide, Australia
V. Ramaswami, Statmetrics, LLC, United States
Rotislav Razumchik, Russian Academy of Sciences, Russia
Alexander Rumyantsev, Russian Academy of Sciences, Russia
Evgenia Smirni, College William & Mary, United States
Aviva Samuelson, University of Tasmania, Australia
Mark Squillante, IBM, United States
Tetsuya Takine, Osaka University, Japan
Peter Taylor, The University of Melbourne, Australia
Miklós Telek, Technical University of Budapest, Hungary
Erik van Doorn, University of Twente, Netherlands
Benny van Houdt, University of Antwerp, Belgium
Eleni Vatamidou, University of Lausanne, Switzerland
Yiqiang Zhao, Carleton University, Canada

Table of Contents

Soohan Ahn and Beatrice Meini *Matrix equations in Markov modulated Brownian motion: theoretical properties and numerical solution* . . . 9

Aregawi Kiros Abera, Małgorzata O'Reilly, Barbara Holland, Mark Fackrell and Mojtaba Heydar *Decision support model for patient admission scheduling problem with random arrivals and departures* . . . 10

Azam Asanjarani *Bursty Markovian Arrival Processes* . . . 15

Søren Asmussen *On some fixed-point problems connecting branching and queueing* . . . 16

Nigel Bean, Giang Nguyen, Bo Friis Nielsen and Oscar Peralta *A fluid flow model with RAP components* . 17

Nigel Bean, Giang T. Nguyen and Federico Poloni *A new algorithm for time-dependent first-return probabilities of a fluid queue* . . . 18

Dario Bini, Stefano Massei, Beatrice Meini and Leonardo Robol *Matrix analytic methods for reflected random walks with stochastic restarts* . . . 23

Peter Braunsteins and Sophie Hautphenne *Extinction in lower Hessenberg branching processes with countably many types* . . . 24

Peter Braunsteins and Sophie Hautphenne *The probabilities of extinction in a branching random walk on a strip* 25

Pavithra Celeste R. and Deepak T.G. *On Fisher Information of Some Functions of Phase Type Variates* . . 26

Jiahao Diao, Tristan Stark, David Liberles, Małgorzata O'Reilly and Barbara Holland *Model for the evolution of the family of gene duplicates* . . . 29

Qi-Ming He, Mark Fackrell and Peter Taylor *Characterization of the Boundary of the Set of Matrix-Exponential Distributions with Only Real Poles* . . . 30

Mojtaba Heydar and Małgorzata O'Reilly *Markovian decision-support model for patient-to-ward assignment problem in a random environment* . . . 31

Gábor Horváth, Illés Horváth and Miklós Telek *High order low variance matrix-exponential distributions* . . 33

Gábor Horváth, Illés Horváth, Salah Almousa and Miklós Telek *Numerical Inverse Laplace Transformation by concentrated ME distributions* . . . 37

Jevgenijs Ivanovs *One-sided Markov additive processes: the three fundamental matrices and the scale function* 41

Sarah James, Nigel Bean and Jonathan Tuke *Modelling intensive care units using quasi-birth-death processes* 42

Kayla Javier and Brian Fralix *An exact analysis of a class of Markovian Bitcoin models* . . . 43

Guy Latouche *Nearly-completely decomposable Markov modulated fluid queues* . . . 44

Barbara Margolius *Eulerian Numbers and an Explicit Formula for a Random Walk Generating Function* . . 45

Barbara Margolius and Małgorzata O'Reilly *Asymptotic periodic analysis of cyclic stochastic fluid flows with time-varying transition rates* . . . 48

Stefano Massei and Sophie Hautphenne *A new numerical method for computing the quasi-stationary distribution of subcritical Galton-Watson branching processes* . . . 49

Giang Nguyen and Oscar Peralta-Gutierrez *Rate of strong convergence of stochastic fluid processes to Markov-modulated Brownian motion* . . . 50

Zbigniew Palmowski *One-sided Markov additive processes: exit problems and related topics* . . . 51

Phil Pollett *Quasi stationarity* . . . 52

Aviva Samuelson, Małgorzata O'Reilly and Nigel Bean *Construction of algorithms for discrete-time quasi-birth-and-death processes through physical interpretation* . . . 53

Julia Shore, Barbara Holland, Jeremy Sumner, Kay Nieselt and Peter Wills *Substitution matrices recapitulate amino acid specificity of aaRS phylogenies* . . . 58

Matthieu Simon *SIR epidemics with stochastic infectious periods* . . . 59

Tristan Stark, Małgorzata O'Reilly, Barbara Holland and David Liberles *Models for the evolution of gene-duplicates: Applications of Phase-Type distributions* . . . 60

Tristan Stark, Małgorzata O'Reilly and Barbara Holland *Models for the evolution of microsatellites* . . . 61

Jeremy Sumner and Michael Baake *The Markov embedding problem: a new look from an algebraic perspective* 62

Jeremy Sumner and Julia Shore *Algebraic constraints on the transition probability matrices produced from Lie-Markov models* . . . 63

Venta Terauds and Jeremy Sumner *Maximum likelihood rearrangement distance for circular genomes* . . . 64

Max Wurm, Andrew Baird, Nigel Bean, Sean Connolly, Ariella Helfgott and Giang Nguyen *Polyp fiction: A stochastic fluid model for the Adaptive Bleaching Hypothesis* . . . 65

Matrix equations in Markov modulated Brownian motion: theoretical properties and numerical solution

Soohan Ahn
Department of Statistics, The University of Seoul
163 Seoulsiripdaero, Dongdaemun-gu, Seoul
02504, South Korea.
sahn@uos.ac.kr

Beatrice Meini
Dipartimento di Matematica, Università di Pisa
Largo B. Pontecorvo 5, 56127 Pisa, Italy
beatrice.meini@unipi.it

ABSTRACT

A Markov modulated Brownian motion(MMBM) is a substantial generalization of the classical Brownian Motion and is obtained by allowing the Brownian parameters to be modulated by an underlying Markov chain of environments [2]. As with Brownian Motion, the stationary analysis of the MMBM becomes easy once the distributions of the first passage time between levels are determined. However, in the MMBM those distributions cannot be obtained explicitly, and we need efficient numerical methods to compute them. In particular, in [2], the computation of the distribution is ultimately reduced to solving a quadratic matrix equation (QME). In relation to this, Ahn and Ramaswami [1] derived a nonsymmetric algebraic Riccati equation(NARE) and proved that the distributions can be obtained by using the minimal nonnegative solution of the equation.

The contribution of this talk is twofold. From one hand we provide an algebraic connection between the QME and the NARE, more specifically we show that the NARE can be obtained by means of a linearization of a quadratic matrix polynomial associated with the QME. On the other hand, we discuss the doubling algorithms such as the structure-preserving doubling algorithm(SDA, [3]) and alternating-directional doubling algorithm(ADDA, [5]) which are used for finding the minimal nonnegative solution of the NARE. These algorithms are quadratically convergent except for the null-recurrent case of the MMBM. To improve the speed of convergence of the doubling algorithms, we introduce a shifted NARE by applying the shift technique, which was investigated by Guo, Iannazzo, and Meini [4]. We observe that the convergence of the doubling algorithms is accelerated and also quadratic even in the null-recurrent case when they are applied to the shifted NARE, as claimed by Guo, Iannazzo, and Meini. Numerical examples show that the algorithm applying ADDA to the shifted NARE is superior to the other doubling algorithms in comparison. This also holds when compared to Nguyen and Poloni's quadratically convergent algorithm [6] that is based on the quadratic matrix equation obtained by Asmussen.

Keywords : Markov modulated Brownian motion, first passage time distribution, doubling algorithm, quadratic convergence, shifted nonsymmetric algebraic Riccati equation.

References

[1] S. Ahn and V. Ramaswami. A quadratically convergent algorithm for first passage time distributions in the markov modulated brownian motion. *Stochastic Models*, 33(1):59–96, 2017.

[2] S. Asmussen. Stationary distributions for fluid flow models with or without brownian noise. *Stochastic Models*, 11:1–20, 1995.

[3] D. A. Bini, B. Iannazzo, and B. Meini. *Numerical solution of algebraic Riccati equations*. Siam, Philadelphia, 2012.

[4] C. H. Guo, B. Iannazzo, and B. Meini. On the doubling algorithm for a shifted nonsymmetric algebraic riccati equation. *SIAM J. Matrix Anal. Appl.*, 63:109–129, 2007.

[5] X. X. Guo, W. W. Lin, and S. F. Xu. A structure-preserving doubling algorithm for nonsymmetric algebraic riccati equation. *Numer. Math.*, 103:393–412, 2006.

[6] G. T. Nguyen and F. Poloni. Componentwise accurate brownian motion computations using cyclic reduction. *arXiv:1605.01482[math.PR]*, 2015.

MAM10 2019, Hobart, Australia

Decision support model for the patient admission scheduling problem with random arrivals and departures

Aregawi K. Abera [*]
University of Tasmania
TAS 7001, Australia
aregawi.abera@utas.edu.au

Małgorzata M. O'Reilly [*†]
University of Tasmania
TAS 7001, Australia
malgorzata.oreilly@utas.edu.au

Barbara R. Holland
University of Tasmania
TAS 7001, Australia
barbara.holland@utas.edu.au

Mark Fackrell [*†]
The University of Melbourne
VIC 3010, Australia
fackrell@unimelb.edu.au

Mojtaba Heydar [*]
University of Tasmania
TAS 7001, Australia
mojtaba.heydar@utas.edu.au

ABSTRACT

Introduction. The patient admission scheduling (PAS) problem is a class of scheduling problems that must be handled by the managers of the hospital admission systems. The problem arises when patients arriving at the hospital need to be allocated to beds in an optimal manner, subject to the availability of beds and the needs of patients.

The PAS problem in a dynamic context, as analysed in Ceschia and Schaerf [2] and Lusby et al. [6], considers a scenario in which random arrivals and unknown departures of patients are gradually revealed over the planning horizon. The problem was formulated as an integer programming model, and various procedures for computing the optimal solution were proposed. Ceschia and Schaerf [2] developed a metaheuristic algorithm based on simulated annealing and neighborhood search. Lusby et al. [6] developed an adaptive large neighbourhood search procedure to solve the problem.

Although the arrivals and departures of patients are in general random, the models in [2, 6] assumed deterministic inputs such as a fixed length of stay for each patient, and a fixed number of arrivals at the start of each day. Here, we build on the analysis in Lusby et al. [6], and develop a model for the PAS problem in a dynamic context, which captures the random dynamics of the flow of the patients.

Our aim here is to develop an improved mathematical model to solve the PAS problem in a dynamic environment with random arrivals and departures. At the start of each day we record new information about the registered patients, newly arrived patients and future arrivals (including emergency patients and scheduled arrivals), and then determine an optimal assignment of patients to beds. Our goal is to provide a decision support tool for the patient scheduling process to be used by hospital administrators and planners.

Notation. We use similar notation to Demeester et al. [3] and Turhan and Bilgen [8] for the parameters and variables of our model, with some minor changes.

- Patients are classified into three groups, admitted patients, planned patients, and emergency patients. Admitted patients are patients that are successfully admitted to the hospital, and allocated to a bed. Planned patients have not been admitted to the hospital as yet, but have pre-determined admission dates, denoted by d_p^{plan}. Emergency patients are patients whose admission date is equal to their registration date, that is, $d_p^{plan} = d_p^{reg}$, since their admission cannot be postponed and is unplanned.

- Patients are denoted by p, with $p \in \mathcal{P}$, where $\mathcal{P}$ is the set of all patients. Also let $\mathcal{M} \subset \mathcal{P}$ be the set of all male patients, and $\mathcal{F} \subset \mathcal{P}$ be the set of all female patients. Patients have the following properties: admission date and a discharge date, age and gender, required treatment, and room preference.

- Days are denoted by d, with $d \in \mathcal{D}$, where $\mathcal{D} = \{0, 1, \ldots, D\}$ is the set of all days in the planning period of the time horizon. Further, let $d_p \in \mathcal{D}_p$ be the admission day of patient p, where $\mathcal{D}_p = \{d_p^{plan}, \cdots, d_p^{max}\} \subseteq \mathcal{D}$ is the set of acceptable days for patient p to be admitted to the hospital.

- The length of stay (LoS) of patient p is denoted by L_p. This is a random variable recording the number of days patient p will stay in the hospital till he/she gets discharged. We assume L_p takes values $\ell_p = 0, 1, \ldots, \ell_p^{max}$, for some positive integer ℓ_p^{max}.

- A hospital consists of different wards. Typically, each ward is specialized in treating one kind of pathology

[*]Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers.

[†]We would like to thank the Australian Research Council for funding this research through Linkage Project LP140100152.

MAM10 2019, Hobart, Australia

such as cardiovascular diseases, oncology, or dermatology, which is considered as the major specialism of the ward as in Demeester et al. [3]. Wards can also perform other treatments as minor specialisms. Wards are denoted by $\mathcal{W}_i$, $i = 1, \ldots, W$, where W is the total number of wards in the hospital. Wards can support one or more specialisms S_u, $u = 1, \ldots, S$, where S is the total number of specialisms. We write

$$S_u \sim \mathcal{W}_i \tag{1}$$

when specialism S_u is available in ward $\mathcal{W}_i$, and

$$S(p) = S_u \tag{2}$$

when patient p requires specialism S_u. Patients admitted to ward $\mathcal{W}_i$ may have to be in a particular age range, between some minimum $a(\mathcal{W}_i)$ and maximum $A(\mathcal{W}_i)$.

- Rooms are denoted by r, with $r \in \mathcal{R} = \{1, \ldots, R\}$, where R is the total number of rooms in the hospital. We write

$$r \in \mathcal{W}_i \tag{3}$$

when room r is in ward $\mathcal{W}_i \subset \mathcal{R}$. A room can be described by its age policy, gender policy and by its special features, such as the presence of oxygen, nitrogen, telemetry or television. A room may support one or more different specialisms S_u, depending on the room features. Rooms have a specified gender policy, which is one of the following; male only M, female only F, depends on the gender of the first patient SG (same-gender policy), or all genders are allowed N. It is preferable to not assign male and female patients to the same room at the same time. We denote by $\mathcal{R}^M, \mathcal{R}^F, \mathcal{R}^{SG}, \mathcal{R}^N \subset \mathcal{R}$ the sets of all rooms with policies M, F, SG, N, respectively.

- The capacity of room r is denoted by κ_r. This is the total number of beds in room r.

- Assignment σ is the collection of decisions $x_{p,r,d}(\sigma)$ and $y_{p,d}(\sigma)$ defined as,

$$x_{p,r,d}(\sigma) = \begin{cases} 1 & \text{if patient } p \text{ is assigned} \\ & \text{to room } r \text{ on day } d \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

$$y_{p,d}(\sigma) = \begin{cases} 1 & \text{if patient } p \text{ is} \\ & \text{admitted on day } d \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

and note that $y_{p,d}(\sigma) = \mathbf{1}\{d = d_p(\sigma)\}$, where $\mathbf{1}\{.\}$ is an indicator function.

- In order to calculate the violation of gender policy, we define the presence of male, female and both patients in room r on day d as follows,

$$m_{r,d}(\sigma) = \begin{cases} 1 & \text{if there is at least one male} \\ & \text{patient in room } r \text{ on day } d, \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

$$f_{r,d}(\sigma) = \begin{cases} 1 & \text{if there is at least one female} \\ & \text{patient in room } r \text{ on day } d, \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

$$b_{r,d}(\sigma) = \begin{cases} 1 & \text{if both genders are present} \\ & \text{in room } r \text{ on day } d, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

- Given patient p, the required features of a room for allocation are grouped into two categories, needed room feature (NRF), and preferred room features (PRF). Given feature j of some room r, we write

$$NRF_j(p,r)(\sigma) = \begin{cases} 1 & \text{if the needed room feature} \\ & \text{is provided} \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Similarly, we write

$$PRF_j(p,r)(\sigma) = \begin{cases} 1 & \text{if the prefered room feature} \\ & \text{is provided} \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

- Transfer means relocating a patient from one room to another during their stay. As it is described by Demeester et al. [3], transfers can be planned or unplanned, the latter should be avoided if possible. As an example of a planned transfer, a patient might be transferred from surgery to an intensive care unit, and after recovery they might be transferred to another ward. An unplanned transfer could be due to a shortage of resources such as beds or rooms. The transfer of patient p from room r to another room r^* on day d is recorded using variable

$$t_{p,r,r^*,d}(\sigma) = \mathbf{1}\{x_{p,r,d-1}(\sigma) = 1, x_{p,r^*,d}(\sigma) = 1, r^* \neq r\}. \tag{11}$$

That is, $t_{p,r,r^*,d}(\sigma) = 1$ when patient p was transferred from room r to room $r^* \neq r$ on day d, and $t_{p,r,r^*,d}(\sigma) = 0$ otherwise.

- Let $Q_{r,d}(\sigma)$ be the event that a gender conflict is observed in room r on day d, given assignment σ. Also, define the random variable $b_{r,d}(\sigma)$ such that $b_{r,d}(\sigma) = 1$ if the event $Q_{r,d}(\sigma)$ occurs, and $b_{r,d}(\sigma) = 0$ otherwise. That is,

$$b_{r,d}(\sigma) = \mathbf{1}\{Q_{r,d}(\sigma)\}. \tag{12}$$

Then the mean value of $b_{r,d}(\sigma)$ is equal to the probability of the event $Q_{r,d}(\sigma)$ occurring, with

$$E(b_{r,d}(\sigma)) = P_r(Q_{r,d}(\sigma)). \tag{13}$$

Denote by $A_{m,d}(\sigma)$ and $A_{f,d}(\sigma)$ the events that all males have left the room before day d, and that all females have left the room before day d, respectively. $\mathcal{F}_{r,d}$ is the set of all female patients assigned to room r on day d, that is $\mathcal{F}_{r,d} = \{p \in \mathcal{F} : x_{p,r,d}(\sigma) = 1\}$, and $\mathcal{M}_{r,d}$ is the set of all male patients assigned to room r on day d, which is $\mathcal{M}_{r,d} = \{p \in \mathcal{M} : x_{p,r,d}(\sigma) = 1\}$. Then,

$$\begin{aligned} 1 - P_r(Q_{r,d}(\sigma)) &= P_r(A_{m,d}(\sigma)) + P_r(A_{f,d}(\sigma)) \\ &\quad - P_r(A_{m,d}(\sigma) \cap A_{f,d}(\sigma)) \\ &= \prod_{\mathcal{M}_{r,d}} x_{p,r,d}(\sigma) P_r(L_p < d - d_p(\sigma)) \\ &\quad + \prod_{\mathcal{F}_{r,d}} x_{p,r,d}(\sigma) P_r(L_p < d - d_p(\sigma)) \\ &\quad - \prod_{\mathcal{M}_{r,d} \cup \mathcal{F}_{r,d}} x_{p,r,d}(\sigma) P_r(L_p < d - d_p(\sigma)). \end{aligned} \tag{14}$$

- Let $Z_{p,r,d}(\sigma)$ be a random variable such that, given assignment σ, $Z_{p,r,d}(\sigma) = 1$ if patient p is in room r on day d, and $Z_{p,r,d}(\sigma) = 0$ otherwise.

- Let $Y_{r,d}(\sigma) = \sum_{p\in\mathcal{P}} Z_{p,r,d}(\sigma)$ be a random variable recording the number of patients in room r on day d, given assignment σ, and $(E(Y_{r,d}(\sigma)) - \kappa_r)$ be the expected excess in room r on day d. We then have

$$\begin{aligned} E(Y_{r,d}(\sigma)) &= E\left(\sum_{p\in\mathcal{P}} Z_{p,r,d}(\sigma)\right) \\ &= \sum_{p\in\mathcal{P}} E\left(Z_{p,r,d}(\sigma)\right) \\ &= \sum_{p\in\mathcal{P}} P_r\left(Z_{p,r,d}(\sigma) = 1\right) \\ &= \sum_{p\in\mathcal{P}} x_{p,r,d}(\sigma) P_r(L_p \geq d - d_p(\sigma)). \end{aligned} \quad (15)$$

- We define the following cost functions, which we later use as coefficients in the objective function. Let $c_{p,r,d}$ be the cost of assigning patient p to a room r on day d. Let $c^{(T)}_{p,r,r^*,d}$ be the cost of transferring patient p from room r to room r^* on day d, with $c^{(T)}_{p,r,r,d} = 0$. Let $c^{(G)}_{r,d}$ be the penalty incurred for the violation of gender policy in room r on day d. Let $c^{(O)}_{r,d}$ be the penalty incurred when the capacity κ_r of room r is exceeded on day d. Let $c^{(De)}_{p,d}$ be the penalty incurred for the admission delay of patient p on day d.

Using the parameters and variables mentioned above, we now construct a stochastic integer programming model with suitable constraints due to patients medical needs and age, room capacity, and gender policy, similar to Lusby et al. [6], with suitable modifications. These include hard constraints that must be met and soft constraints that can be violated when necessary, but which are subject to cost penalties.

Hard constraints. For a solution to be feasible, it has to satisfy the following set of hard constraints (16)-(21). First, we set the room capacity constraints,

$$\sum_{p\in P} x_{p,r,d}(\sigma) \quad \leq \quad \widehat{\kappa}_r, \quad \forall r \in \mathcal{R}, \forall d \in \mathcal{D}, \quad (16)$$

where $\widehat{\kappa}_r \geq \kappa_r$ is some maximum allowed threshold for the total number of patients in room r, after taking into account an overstay risk.

Next, patient p should be assigned to ward $\mathcal{W}_i$ that is suited for the patient's age, denoted A_p. The minimum age limit $a(\mathcal{W}_i)$ and maximum age limit $A(\mathcal{W}_i)$ allowed in ward $\mathcal{W}_i$ should be respected. Therefore,

$$\begin{aligned} x_{p,r,d}(\sigma)\mathbf{1}\{r \in \mathcal{W}_i\} \quad &\leq \quad \mathbf{1}\{a(\mathcal{W}_i) \leq A_p \leq A(\mathcal{W}_i)\}, \\ &\forall p \in \mathcal{P}, r \in \mathcal{R}, d \in \mathcal{D}. \end{aligned} \quad (17)$$

Furthermore, a patient p should be assigned to a ward W_i with a suitable specialism S_u, for some u. Therefore,

$$\begin{aligned} x_{p,r,d}(\sigma)\mathbf{1}\{r \in \mathcal{W}_i, S(p) = S_u\} \quad &\leq \quad \mathbf{1}\{S_u \sim \mathcal{W}_i\}, \\ &\forall p \in \mathcal{P}, r \in \mathcal{R}, d \in \mathcal{D}. \end{aligned} \quad (18)$$

Additionally, the medical treatment of a patient p may require that he/she is assigned to a room r with special equipment or other features required for the treatment. That is, when making decision $x_{p,r,d}(\sigma) = 1$ we must have r such that $NRF_j(p, r) = 1$, when patient p requires room feature j. Therefore,

$$x_{p,r,d}(\sigma) \quad \leq \quad \mathbf{1}\{NRF_j(p, r) = 1\}, \quad \forall p \in \mathcal{P}, r \in \mathcal{R}, d \in \mathcal{D}. \quad (19)$$

Also, patients have to be admitted within the planning horizon, and so

$$\sum_{d\in\mathcal{D}_p} y_{p,d}(\sigma) = 1, \quad \forall p \in \mathcal{P}. \quad (20)$$

Moreover, if patient p is admitted on day $\bar{d}$, the patient must appear in some room r the following $\ell_p^{max} - 1$ nights, which gives,

$$\begin{aligned} \sum_{r\in R} x_{p,r,d}(\sigma) \quad &\geq \quad y_{p,\bar{d}}(\sigma), \\ &\forall p \in \mathcal{P}, d = \bar{d}, \ldots, \bar{d} + \ell_p^{max} - 1, \bar{d} \in \mathcal{D}_p. \end{aligned} \quad (21)$$

Soft constraints. The set of soft constraints (22)-(25) corresponds to desirable conditions that do not have to be met, but are subject to penalties.

Ideally, patients should be allocated as per their gender to an appropriate room r with its specified gender policy. We evalute the presence of a female patient $f_{r,d}(\sigma)$ in room r on day d is using

$$f_{r,d}(\sigma) \quad \geq \quad x_{p,r,d}(\sigma), \quad \forall p \in \mathcal{F}, \forall r \in \mathcal{R}^{SG}, \forall d \in \mathcal{D}, \quad (22)$$

and the presence of a male patient $m_{r,d}(\sigma)$ using

$$m_{r,d}(\sigma) \quad \geq \quad x_{p,r,d}(\sigma), \quad \forall p \in \mathcal{M}, \forall r \in \mathcal{R}^{SG}, \forall d \in \mathcal{D}. \quad (23)$$

To determine when both genders are present $b_{r,d}(\sigma)$ we use the following constraint,

$$b_{r,d}(\sigma) \quad \geq \quad m_{r,d}(\sigma) + f_{r,d}(\sigma) - 1, \quad \forall r \in \mathcal{R}^{SG}, \forall d \in \mathcal{D}. \quad (24)$$

The transfer of patients is handled using the following constraint,

$$\begin{aligned} t_{p,r,r^*,d}(\sigma) \quad &\geq \quad x_{p,r,d}(\sigma) - x_{p,r,d-1}(\sigma), \\ &\forall p \in P, \forall r \in R, \forall d = 2, \ldots, \mathcal{D}. \end{aligned} \quad (25)$$

Some other desirable conditions could also be considered. A patient who asked for a single room, in case of lack of single rooms should preferably be assigned to a twin room.

In addition to major medical treatment, a patient p may need to undergo other minor medical treatments within department W_i in a room r with special equipment to treat the patient assigned, which requires some minor specialism S_ℓ for some suitable ℓ.

A patient p may prefer a room r with features that in some degree correspond to the specialism that is required to treat the patient's clinical condition. That is, it is preferable to have $PRF_j(p,r) = 1$, when patient p prefers room feature j.

Objective function. We define the *stochastic* objective function as the total *expected* cost incurred over the planning horizon $\mathcal{D} = \{0, 1, \ldots, D\}$, and write it as a sum of the following cost components. The first component captures the cost of assigning patient $p \in \mathcal{P}$ to room $r \in \mathcal{R}$ on day $d \in \mathcal{D}$. The second component calculates the cost of transferring patient $p \in \mathcal{P}$ from room $r \in \mathcal{R}$ to another room $r^* \in \mathcal{R}$ on day $d \in \mathcal{D}$. The third component determines the penalty incurred for the violation of gender policy in room $r \in \mathcal{R}$ on day $d \in \mathcal{D}$. The fourth component determines the penalty incurred when the capacity κ_r of room $r \in \mathcal{R}$ is exceeded on day $d \in \mathcal{D}$. The fifth component computes the penalty incurred when the admission of patient $p \in \mathcal{P}$ is delayed beyond the maximum acceptable admission day d_p^{max} on day $d \in \mathcal{D}$. The resulting expression is stated as,

$$\begin{aligned}\min_{\sigma} \Bigg\{ & \sum_{p\in\mathcal{P}} \sum_{d\in\mathcal{D}} \sum_{r\in\mathcal{R}} c_{p,r,d} \times x_{p,r,d}(\sigma) \times Pr(L_p \geq d - d_p(\sigma)) \\ & + \sum_{p\in\mathcal{P}} \sum_{d\in\mathcal{D}} \sum_{r\in\mathcal{R}} c^{(T)}_{p,r,r^*,d} \times t_{p,r,r^*,d}(\sigma) \times Pr(L_p \geq d - d_p(\sigma)) \\ & + \sum_{d\in\mathcal{D}} \sum_{r\in\mathcal{R}} c^{(G)}_{r,d} \times P_r(Q_{r,d}(\sigma)) \\ & + \sum_{d\in\mathcal{D}} \sum_{r\in\mathcal{R}} c^{(O)}_{r,d} \times \left(\frac{\max\{0,\ E(Y_{r,d}(\sigma)) - \kappa_r\}}{\widehat{\kappa}_r - \kappa_r} \right) \\ & + \sum_{p\in\mathcal{P}} c^{(De)}_{p,d} \times \sum_{d\in\mathcal{D}} \left(\frac{d - d_p^{plan}}{d_p^{max} - d_p^{plan}} \right) \times y_{p,d}(\sigma) \quad \Bigg\}, \end{aligned} \tag{26}$$

where $\max\{0,\ E(Y_{r,d}(\sigma)) - \kappa_r\}$ is the expected number of patients in room r on day d *above the capacity* of room r, given assignment σ.

Random arrivals and departures. In order to model the random departures, we assume that the random variable L_p that records the LoS of the type-p patient, and takes values $\ell_p = 0, 1, \ldots, \ell_p^{max}$, for some positive integer ℓ_p^{max}, follows a discrete phase-type distribution in Latouche and Ramaswami [5, Chapter 2] and Neuts [7] with parameters that depend on p.

That is, we consider a discrete-time Markov chain with state space $\mathcal{V} = \{0, 1, \ldots, \ell_p^{max}\}$, where ℓ_p^{max} is an absorbing state, and one-step transition probability matrix $\mathbf{P}$ given by

$$\mathbf{P}^* = \begin{bmatrix} \mathbf{P} & \mathbf{p} \\ \mathbf{0} & 1 \end{bmatrix}, \tag{27}$$

for some matrix $\mathbf{P} = [P_{i,j}]_{i,j=0,1,\ldots,\ell_p^{max}-1}$ and (column) vector $\mathbf{p} = [p_{i,\ell_p^{max}}]_{j=0,1,\ldots,\ell_p^{max}-1}$, and the initial distribution (row) vector $\boldsymbol{\tau} = [\tau_i]_{i=0,1,\ldots,\ell_p^{max}-1}$.

We then assume that the random variable L_p follows discrete phase-type distribution with parameters $\boldsymbol{\tau}$ and $\mathbf{P}$, which models time till absorption in the above chain,

$$L_p \sim PH(\boldsymbol{\tau}, \mathbf{P}), \tag{28}$$

which gives, for $\ell_p = 0, 1, \ldots, \ell_p^{max}$,

$$Pr(L_p = \ell_p) = \boldsymbol{\tau}\mathbf{P}^{\ell_p}\mathbf{p}, \tag{29}$$

$$Pr(L_p \leq \ell_p) = 1 - \boldsymbol{\tau}\mathbf{P}^{\ell_p}\mathbf{1}, \tag{30}$$

where $\mathbf{1}$ is a (column) vector of ones of appropriate size. We use these expressions in order to evaluate the first two components of the objective function in (26).

In order to include the random arrivals that may occur during the planning horizon, we apply an approach similar to Kumar et al. [4]. We simulate random arrivals (from a suitable distribution) multiple times, resulting in a number of possible solutions. We then compare the different solutions by running simulations over some long time period, and then choose the preferred solution.

For example, suppose that the arrivals of patients (emergency or scheduled) occur according to a Poisson process with rate λ_p per day, for type-p patient, for all $p \in \mathcal{P}$, where patient type is determined by their medical needs, age and gender. We generate the random arrivals of emergency patients in the time horizon $[0, D]$, using standard simulation methods. As one possibility, for each patient type p, assume that $\lceil D\lambda_p \rceil$ arrivals have occurred during the time interval $[0, D]$, and then draw the random arrival times from a discrete uniform distribution on $\{0, 1, \ldots, D\}$. We then add the set of such generated patients to the problem, and solve it using the model in (26), treating these patients as registered patients, and so, patients that are known to the system.

Solution approach. We use simulation in order to generate random inputs for our model, and apply metaheuristic algorithms, similar to [6], including greedy search, adaptive neigbourhood search and simulated annealing, to solve the stochastic integer program. In our algorithm, we set the initial solution to be the optimal solution of the algorithm in [6], and compare our results with those of Lusby et al. [6]. The results of the application of our model will be reported in [1].

1. REFERENCES

[1] A. K. Abera, M. M. O'Reilly, B. R. Holland, M. Fackrell, and M. Heydar. On the decision support model for the patient admission scheduling problem with random arrivals and departures: Solution approach. In preparation.

[2] S. Ceschia and A. Schaerf. Modeling and solving the dynamic patient admission scheduling problem under uncertainty. *Artificial Intelligence in Medicine*, 56(3):199–205, 2012.

[3] P. Demeester, W. Souffriau, P. De Causmaecker, and G. Vanden Berghe. A hybrid tabu search algorithm for automatically assigning patients to beds. *Artificial Intelligence in Medicine*, 48(1):61–70, 2010.

[4] A. Kumar, A. M. Costa, M. Fackrell, and P. G. Taylor. A sequential stochastic mixed integer programming model for tactical master surgery scheduling. *European Journal of Operational Research*, 270(2):734–746, 2018.

[5] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modelling*. ASA SIAM, 1999.

[6] R. M. Lusby, M. Schwierz, T. M. Range, and J. Larsen. An adaptive large neighborhood search procedure applied to the dynamic patient admission scheduling problem. *Artificial Intelligence in Medicine*, 74:21–31, 2016.

[7] M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models: an Algorithmic Approach*. Dover Publications Inc., 1981.

[8] A. M. Turhan and B. Bilgen. Mixed integer programming based heuristics for the patient admission scheduling problem. *Computers and Operations Research*, 80:38–49, 2017.

Bursty Markovian Arrival Processes

Keynote speaker

Azam Asanjarani
The University of Auckland
38 Princess Street
Auckland, New Zealand
azam.asanjarani@auckland.ac.nz

ABSTRACT

We consider stationary Markovian Arrival Processes (MAPs) where both the squared coefficient of variation of inter-event times and the asymptotic index of dispersion of counts are greater than unity:

$$c^2 = \frac{\mathrm{Var}(T_n)}{\mathbb{E}^2\,[T_n]} \geq 1, \qquad d^2 := \lim_{t\to\infty} \frac{\mathrm{Var}(N(t))}{\mathbb{E}[N(t)]} \geq 1.$$

We refer to such MAPs as *bursty*. The simplest bursty MAP is a Hyperexponential Renewal Process (H-renewal process). Applying Matrix analytic methods (MAM), we establish further classes of MAPs as Bursty MAPs: the Markov Modulated Poisson Process (MMPP), the Markov Transition Counting Process (MTCP) and the Markov Switched Poisson Process (MSPP). Of these, MMPP has been used most often in applications, but as we illustrate, MTCP and MSPP may serve as alternative models of bursty traffic. Hence understating MTCPs, MSPPs, and MMPPs and their relationships is important from a data modelling perspective. We establish a duality in terms of first and second moments of counts between MTCPs and a rich class of MMPPs which we refer to as slow-MMPPs (modulation is slower than the events).

MAM10 2019, Hobart, Australia

On some fixed-point problems connecting branching and queueing

Keynote speaker

Søren Asmussen
Aarhus University, Denmark

ABSTRACT

Connections between branching and queueing have a long history. A classical case is the M/G/1 queue where one can view the children of a customer as the say N customers arriving during his service time S. This leads to the fixed-point equation

$$R \stackrel{d}{=} S + \sum_{i=1}^{N} R_i \tag{1}$$

for the busy period R, where $R_1, R_2, \ldots$ are i.i.d. and independent of (S, N). Similar equations occur in other branching process connections. A quick application is to note that the queue is stable if and only if the corresponding branching process is subcritical, which immediately gives the $\rho < 1$ criterion. Equations similar have been used in recent work ([5], [9]) related to the Google page rank algorithm to derive tail asymptotics of R under regular variation (RV); for the busy period, the RV asymptotics has earlier been studied in [6] and [10]. We present a simple random walk argument from [1] to give a short proof of these results as well as certain extensions. Motivated from a multiclass queueing model originating from [4], also a multivariate version of (1) is studied under RV conditions.

Following [2], we also consider preemptive-repeat LIFO queues where the time R in system is related to the fixed-point equation

$$R(s) \stackrel{d}{=} T \wedge s + \mathbf{1}(T \le s)\big(R + R^*(s)\big) \tag{2}$$

where T is the interarrival time and $R(s)$ the time-in-system of a customer with service time s. Using again a branching connection gives a highly non-standard stability condition for the M/G/1 case. However, for GI/G/1 equation (2) does not have the correct interpretation, and we present a matrix-analytic approach that lead to an algorithm for finding the stability region for PH/G/1. The approach indeed uses a connection to a (multitype) branching process, but meets the difficulty that the offspring distribution is not explicit. For somewhat related models, MAM have earlier been used in [7] and [3].

1. REFERENCES

[1] S. Asmussen and S. Foss (2018) Regular variation in a fixed-point point problem for single-and multiclass queues and branching processes. To appear in *Branching and Applied Probability. Papers in honour of Peter Jagers. Advances in Applied Probability* **50A**. `arXiv:1709.05140v2`

[2] S. Asmussen and P.W. Glynn (2017) On preemptive-repeat LIFO queues. *QUESTA* **87**, 1–22. DOI 10.1007/s11134-017-9532-3.

[3] D.A. Bini, G. Latouche and B. Meini (2003) Solving nonlinear matrix equations arising in tree-like stochastic processes. *Linear Algebra and its Applications* **366**, 39–64

[4] P. Ernst, S. Asmussen and J. Hasenbein (2018) Stability and tail asymptotics in a multiclass queue with state dependent arrival rates. `arXiv:1609.03999`, accepted for *QUESTA*

[5] P.R. Jelenković and M. Olvera-Cravioto (2010) Information ranking and power laws on trees. *Adv. Appl. Probab.* **42**, 577–604.

[6] A. de Meyer and J.L. Teugels (1980) On the asymptotic behaviour of the distributions of the busy period and service time in M/G/1. *J. Appl. Probab.* **17**, 802–813.

[7] Q.-M He and A.S. Alfa (1998) The MMAP[K]/PH[K]/1 queues with a last-come-first-served preemptive service discipline *QUESTA* **29**, 269–291.

[8] M.F. Neuts (1974). The Markov renewal branching process. *Mathematical methods in queueing theory* (Proc. Conf., Western Michigan Univ., Kalamazoo, Mich., 1973). *Lecture Notes in Economics and Mathematical Systems* **98**, 1–21. Springer.

[9] Y. Volkovich and N. Litvak (2010) Asymptotic analysis for personalized web search. *Adv. Appl. Probab.* **42**, 577–604.

[10] B. Zwart (2001) Tail asymptotics for the busy period in the GI/G/1 queue. *Math. Oper. Res.* **26**, 485–493.

MAM10 2019, Hobart, Australia

A fluid flow model with RAP components

Nigel Bean
School of Mathematical Sciences
The University of Adelaide
Adelaide, Australia
nigel.bean@adelaide.edu.au

Giang Nguyen
School of Mathematical Sciences
The University of Adelaide
Adelaide, Australia
giang.nguyen@adelaide.edu.au

Bo Friis Nielsen
DTU Compute
Technical University of Denmark
Lyngby, Denmark
bfni@dtu.dk

Oscar Peralta
School of Mathematical Sciences
The University of Adelaide
Adelaide, Australia
oscar.peraltagutierrez@adelaide.edu.au

ABSTRACT

The class of matrix–exponential distributions (ME) constitutes an algebraic generalisation of the class of phase–type distributions (PH). Similarly, the Rational arrival process (RAP) generalises the Markovian arrival process (MAP) in an algebraic sense, however, their probabilistic constructions are considerably different. For the MAP, the driving process is a Markov jump process with finite state space, while it is a piecewise deterministic Markov process (PDMP) for the RAP; see [1]. This approach was further studied in [2] to define a class of quasi–birth and death processes (QBD) with RAP components, an algebraic generalisation of the QBD. In this talk we provide a generalisation of the fluid flow model in a similar direction. More specifically, we consider the process

$$V_t = \int_0^t \mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{U}\} - \mathbb{1}\{\boldsymbol{A}(s) \in \mathfrak{D}\} ds,$$

where $\mathfrak{U}$ and $\mathfrak{D}$ are some affine spaces in $\mathbb{R}^n$ and $\mathbb{R}^m$ (respectively), and $\{\boldsymbol{A}(t)\}_{t\geq 0}$ is a PDMP with state space $\mathfrak{U}\cup\mathfrak{D}$ with the following local characteristics. The process $\{\boldsymbol{A}(t)\}_{t\geq 0}$ evolves between jumps according to the system of differential equations

$$\frac{d\boldsymbol{A}(t)}{dt} = \begin{cases} \boldsymbol{A}(t)\boldsymbol{C}_+ - \boldsymbol{A}(t)\boldsymbol{C}_+\boldsymbol{e}\cdot\boldsymbol{A}(t) & \text{if} \quad \boldsymbol{A}(t)\in\mathfrak{U} \\ \boldsymbol{A}(t)\boldsymbol{C}_- - \boldsymbol{A}(t)\boldsymbol{C}_-\boldsymbol{e}\cdot\boldsymbol{A}(t) & \text{if} \quad \boldsymbol{A}(t)\in\mathfrak{D}, \end{cases}$$

for some real square matrices $\boldsymbol{C}_+$ and $\boldsymbol{C}_-$ of appropiate dimensions. At each $\boldsymbol{a} \in \mathfrak{U}\cup\mathfrak{D}$, $\{\boldsymbol{A}(t)\}_{t\geq 0}$ has a jump intensity $\lambda(\boldsymbol{a})$ given by

$$\lambda(\boldsymbol{a}) = \begin{cases} \boldsymbol{a}\boldsymbol{D}_{+-}\boldsymbol{e} & \text{if} \quad \boldsymbol{A}(t)\in\mathfrak{U} \\ \boldsymbol{a}\boldsymbol{D}_{-+}\boldsymbol{e} & \text{if} \quad \boldsymbol{A}(t)\in\mathfrak{D}. \end{cases}$$

Given that a jump happens at $\boldsymbol{a} \in \mathfrak{U}\cup\mathfrak{D}$, it will land in $\dfrac{\boldsymbol{a}\boldsymbol{D}_{+-}}{\boldsymbol{a}\boldsymbol{D}_{+-}\boldsymbol{e}} \in \mathfrak{D}$ if $\boldsymbol{a}\in\mathfrak{U}$, or in $\dfrac{\boldsymbol{a}\boldsymbol{D}_{-+}}{\boldsymbol{a}\boldsymbol{D}_{-+}\boldsymbol{e}} \in \mathfrak{U}$ if $\boldsymbol{a}\in\mathfrak{D}$, where $\boldsymbol{D}_{+-}$ and $\boldsymbol{D}_{-+}$ are some real matrices of appropiate dimensions. Due to the similiarities of this construction to the one corresponding to the RAP in [1], we call the process $\{V_t, \boldsymbol{A}(t)\}_{t\geq 0}$ a **Fluid RAP** (FRAP). We study some distributional properties of the process $\{V_t\}_{t\geq 0}$, such as first passage probabilities and the stationary distribution of its queue, and relate them to classic results of fluid flow models. Finally, we discuss some similarities and differences between the techniques that were needed to study the FRAP and the ones commonly used for fluid flow models in the literature.

1. REFERENCES

[1] S. Asmussen and M. Bladt. Point processes with finite-dimensional conditional probabilities. *Stochastic Processes and their Applications*, 82(1):127–142, July 1999.

[2] N. G. Bean and B. F. Nielsen. Quasi-Birth-and-Death Processes with Rational Arrival Process Components. *Stochastic Models*, 26(3):309–334, Aug. 2010.

MAM10 2019, Hobart, Australia

A new algorithm for time-dependent first-return probabilities of a fluid queue

N. G. Bean
School of Mathematical Sciences
The University of Adelaide
Adelaide, Australia
nigel.bean@adelaide.edu.au

G. T. Nguyen
School of Mathematical Sciences
The University of Adelaide
Adelaide, Australia
giang.nguyen@adelaide.edu.au

F. Poloni
Dipartimento di Informatica
Università di Pisa
Pisa, Italy
federico.poloni@unipi.it

1. INTRODUCTION

Let $(\mathcal{X}, \varphi) = \{X(t), \varphi(t)\}_{t\geq 0}$ be a fluid queue, where φ is the environment, modelled as a continuous-time Markov chain with state space $\mathcal{S}$ and generator $Q \in \mathbb{R}^{N\times N}$, $N = |\mathcal{S}|$, and $\mathcal{X}$ is the fluid level with $dX(t)/dt = c_{\varphi_t}$ for $t \geq 0$. One key quantity in the steady-state analysis of a fluid queue is its first return matrix Ψ with entries

$$\Psi_{ij} = \mathbb{P}[\tau < \infty,\, \varphi(\tau) = j \in \mathcal{S}_- \mid \varphi(0) = i \in \mathcal{S}_+], \qquad (1)$$

where $\tau = \min\{t > 0 \colon X(\tau) = X(0)\}$ is the first return time, $\mathcal{S}_+ = \{i \in \mathcal{S} : c_i > 0\}$ and $\mathcal{S}_- = \{i \in \mathcal{S} : c_i < 0\}$.

Similarly, for its transient analysis one is interested in computing the time-dependent matrix $\Psi(t)$ with elements

$$[\Psi(t)]_{ij} = \mathbb{P}[\tau < t,\, \varphi(\tau) = j \in \mathcal{S}_- \mid \varphi(0) = i \in \mathcal{S}_+].$$

To compute $\Psi(t)$, one of the most popular methods in the literature is not to evaluate it directly, but to determine its Laplace-Stieltjes transform $\hat{\Psi}(s)$ in the complex plane [2, 7], then relying on algorithms for the inverse transform [1]. This approach works well in practice, but has two drawbacks. First, it requires working with complex arithmetic to compute results that are real positive quantities; second, its results are not accurate to full machine precision, due to intrinsic inaccuracies in these inverse transforms.

A different, direct algorithm not based on Laplace transforms can be obtained with small modifications to [5], where the authors focus on the busy period, $\Psi(t)\mathbf{1}$. Following their technique, one can also obtain an algorithm for $\Psi(t)$. To the best of our knowledge, this algorithm has no probabilistic interpretation; its proof is based on algebraic verification that the resulting function satisfies the Kolmogorov equations.

In this work, we propose another algorithm, which works directly on probability matrices and has a direct physical interpretation. Moreover, it is essentially *subtraction-free*, i.e., it requires only sums and products of positive quantities. Subtraction-free algorithms have already been proposed for various similar tasks, see, e.g., [9, 3, 8]. Their main advantage is that, using the subtraction-free property often one can prove excellent stability properties, obtaining a forward relative error of the order of the machine precision. (Note that the algorithm obtained from [5] is *almost* subtraction-free, but it still performs subtractions between the rates c_i.)

We explain in Section 2 the recurrence on which our algorithm is based, then explore in Section 3 how to truncate the infinite sum appearing in it. In Section 4 we compute the complexity of the algorithm and compare it to that of competing methods. Finally, in Section 5 we present some experiments for comparison.

MAM10 2019, Hobart, Australia

2. THE NEW ALGORITHM

For simplicity, in this short exposition we restrict ourselves to the case in which no rate c_i equals 0, so $\mathcal{S} = \mathcal{S}_+ \cup \mathcal{S}_-$. Suppose that the fluid is uniformized, i.e., φ is replaced by its uniformized discrete-time Markov chain $\varphi^d = \{\varphi^d(t)\}_{t\geq 0}$ with transition matrix $P = I + \lambda^{-1}Q$, for a suitable $\lambda > 0$. Note that the matrix $\Psi(t)$ remains unchanged after the uniformization.

Let $\{t_i\}_{i\in\mathbb{N}}$ be the sequence of Poisson epochs of φ^d, $t_0 = 0$, and set for brevity $X_i = X(t_i)$ and $\varphi^d_i = \varphi^d(t_i)$. Define for $n \in \mathbb{N}^+$ and $t \in \mathbb{R}^{>0}$ the quantity $m_n = \arg\min_{\ell=1,2,\dots,n} X_\ell$, and two matrices $\Psi^+_n(t)$ and $\Psi^-_n(t)$ with entries

$$[\Psi^+_n(t)]_{ij} = \mathbb{P}[X_{m_n} > X(t) > X(0),\, \varphi^d(t) = j \in \mathcal{S}_- \mid \varphi^d(0) = i \in \mathcal{S}_+,\, t_n < t \leq t_{n+1}],$$
$$[\Psi^-_n(t)]_{ij} = \mathbb{P}[X_{m_n} > X(0) > X(t),\, \varphi^d(t) = j \in \mathcal{S}_- \mid \varphi^d(0) = i \in \mathcal{S}_+,\, t_n < t \leq t_{n+1}],$$

respectively, from which it follows that $\Psi^+_0(t) = \Psi^-_0(t) = 0$.

The first non-trivial result is the following.

LEMMA 1. *The matrices $\Psi^\pm_n(t)$ are independent of t.*

PROOF. This follows from a rescaling argument, we shall prove $\Psi^+_n(t) = \Psi^+_n(1)$ for any $t \in \mathbb{R}^{>0}$. The same argument then holds for $\Psi^-_n(t)$.

Consider an arbitrary but fixed time $t > 0$. Conditioned on there being n Poisson events in $[0, t]$, the epochs t_k, $k = 1, \dots, n$, are uniformly distributed in $[0, t]$, which implies that the epochs t_k/t, $k = 1, \dots, n$, are uniformly distributed in $[0, 1]$. Corresponding with each sequence of transition times $\{\widehat{t}_0, \dots, \widehat{t}_n\}$ and an associated sequence of states $\{\widehat{\varphi}^d_0, \dots, \widehat{\varphi}^d_n\}$, is a unique sample path $\widehat{x}_{[0,t]}$ of the fluid level

$\mathcal{X}$ over $[0,t]$. Similarly, corresponding to $\{\widehat{t}_0/t,\dots,\widehat{t}_n/t\}$ and the same state sequence $\{\widehat{\varphi}_0^d,\dots,\widehat{\varphi}_n^d\}$, is a unique sample path $\widehat{y}_{[0,1]}$ of the rescaled process $\{1/tX(s/t)\}_{s\in\mathbb{R}\geq 0}$ on $[0,1]$.

As the distribution of the states is determined by P alone (and not by t), the probability density associated with $\widehat{x}_{[0,t]}$ is the same as the probability density associated with the rescaled sample path $\widehat{y}_{[0,1]}$. Thus,

$$\begin{aligned}[\Psi_n^+(t)]_{ij} &= \mathbb{P}[X_{m_n} > X(t) > X(0),\, \varphi^d(t) = j \in \mathcal{S}_- \mid \\ &\qquad \varphi^d(0) = i \in \mathcal{S}_+,\, t_n < t < t_{n+1}] \\ &= \mathbb{P}[X(t_{m_n}/t) > X(1) > X(0),\, \varphi^d(1) = j \in \mathcal{S}_- \mid \\ &\qquad \varphi^d(0) = i \in \mathcal{S}_+,\, t_n/t < 1 < t_{n+1}/t] \\ &= \mathbb{P}[X(t^*_{m^*_n}) > X(1) > X(0),\, \varphi^d(1) = j \in \mathcal{S}_- \mid \\ &\qquad \varphi^d(0) = i \in \mathcal{S}_+,\, t^*_n < 1 < t^*_{n+1}],\end{aligned}$$

where $\{t_i^*\}_{i\in\mathbb{N}}$ is a sequence of epochs of a Poisson process of the same rate λ and $m_n^* = \arg\min_{\ell=1,2,\dots,n} X(t_\ell^*)$. The last equality follows from the fact that conditioned on there being n events in $[0,1]$, $t_i^*, i = 1,\dots,n$ are uniformly distributed on $[0,1]$. □

Next, we show that $\Psi(t)$ can be computed using $\Psi_n^-(t)$ alone. To that end, observe that conditioned on there being n Poisson events in $[0,t]$ and $\varphi(0) \in \mathcal{S}_+$, there are four sets of exhaustive and mutually exclusive sample paths, those with

1. $X_{m_n} > X(t) > X(0)$, which contribute to $\Psi_n^+(t)$ — these have not returned to level $X(0)$ by time t and therefore do not contribute to $\Psi(t)$;
2. $X_{m_n} > X(0) > X(t)$, which contribute to $\Psi_n^-(t)$ — these have returned to level $X(0)$ at some time τ such that $t_n < \tau < t$, and contribute to $\Psi(t)$;
3. $X(t) > X_{m_n} > X(0)$ — these have not returned to $X(0)$ by time t and thus do not contribute to $\Psi(t)$;
4. $X(0) > X_{m_n}$, these have returned to $X(0)$ at some time τ such that $\tau < t_{m_n}$ — they contribute to $\Psi(t)$.

(For the first two sets, it is not possible that $\varphi^d(t) \in \mathcal{S}_+$.) Note also that in the second set there are n Poisson events in $(0,\tau)$, and in the fourth set there are k, for $k = 1,\dots,n-1$, events in $(0,\tau)$. By Lemma 1, we abbreviate $\Psi_n^\pm(t)$ to $\Psi_n^\pm$, and have the following.

Lemma 2. *We have*

$$\Psi(t) = \sum_{n=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^n}{n!}\sum_{k=1}^{n}\Psi_k^-. \tag{2}$$

Proof. Conditioning on the number of events n in $[0,t]$ and on the index k of the last event before τ, we have

$$\begin{aligned}[\Psi(t)]_{ij} &= \sum_{n=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^n}{n!}\sum_{k=1}^{n}\mathbb{P}[t_k < \tau < t_{k+1},\, \varphi^d(\tau) = j \in \mathcal{S}_- \mid \\ &\qquad \varphi^d(0) = i \in \mathcal{S}_+, t_n < t < t_{n+1}] \\ &= \sum_{n=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^n}{n!}\left(\Psi_n^-(t) + \sum_{k=1}^{n-1}\Psi_k^-(t_{k+1})\right) \\ &= \sum_{n=0}^{\infty} e^{-\lambda t}\frac{(\lambda t)^n}{n!}\sum_{k=1}^{n}\Psi_k^-(t),\end{aligned}$$

where the last equality follows from Lemma 1. □

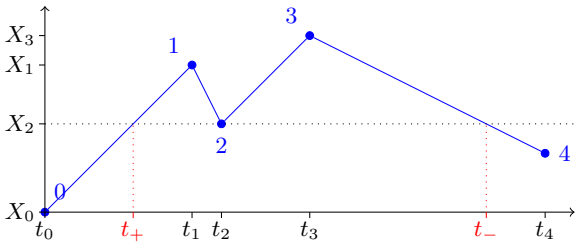


Figure 1: A sample path in $\Psi_3^+(t_4)$ with $m_3 = 2$.

Lemma 3. *Setting $\Psi_n = \Psi_n^+ + \Psi_n^-$, we have*

$$[\Psi_n^+]_{ij} = \frac{c_i}{c_i + |c_j|}[\Psi_n]_{ij}, \quad [\Psi_n^-]_{ij} = \frac{|c_j|}{c_i + |c_j|}[\Psi_n]_{ij}. \tag{3}$$

Proof. The sample paths counted in $[\Psi_n(t_{n+1})]_{ij} = [\Psi_n]_{ij}$ are those in which $X_{m_n} > \max(X_0, X_{n+1})$, and, in addition, the fluid level increases with rate $c_i > 0$ in (t_0, t_1) and decreases with rate $c_j < 0$ in (t_n, t_{n+1}). We shall show, that under these conditions the probability that $X_0 > X_{n+1}$ is $|c_j|/(c_i + |c_j|)$; this implies the thesis.

A sample path counted in $[\Psi_n(t_{n+1})]_{ij}$ is determined uniquely by (i) the sequence of states $\{\varphi_0^d,\dots,\varphi_n^d\}$, and (ii) the i.i.d. increments $t_1 - t_0, t_2 - t_1, \dots, t_{n+1} - t_n$, independent from the states. If we fix the values of a sequence of states, and of the increments $t_2 - t_1, \dots, t_n - t_{n-1}$ (*excluding* the first and last), this determines uniquely the shape of the sample path, apart from the lengths of the first and last legs. In particular, the times $t_+ \in (0, t_1]$ and $t_- \in [t_n, t_{n+1})$ at which $X(t_+) = X(t_-) = X_{m_n}$ are uniquely determined by this choice, and they are independent from $t_1 - t_0$ and $t_{n+1} - t_n$. (See Figure 1 for an example that shows how the various quantities are defined.)

Recall that we have conditioned on $X_{m_n} > X_0$, i.e., the length of $t_1 - t_0 \sim \mathrm{Exp}(\lambda)$ exceeds $t_1 - t_+$; by the memoryless property, $t_+ - t_0$ is also $\mathrm{Exp}(\lambda)$. The corresponding difference in level is $X_{m_n} - X_0 = X(t_+) - X_0 = c_i(t_+ - t_0) \sim \mathrm{Exp}(\lambda/c_i)$. Analogously, $X_{m_n} - X_{n+1} = X(t_-) - X_{n+1} = |c_j|(t_{n+1} - t_-) \sim \mathrm{Exp}(\lambda/|c_j|)$, which is independent from $X_{m_n} - X_0$. The probability that $X_0 > X_{n+1}$ equals the probability that $X_{m_n} - X_0 < X_{m_n} - X_{n+1}$, which is thus $(\lambda/c_i)/(\lambda/c_i + \lambda/|c_j|) = |c_j|/(c_i + |c_j|)$. □

Our main result is the following recursion, which, together with (3), allows one to determine $\Psi_n^\pm$.

Theorem 4. *The following recurrence holds.*

$$\Psi_1 = P_{+-}, \quad \textit{and, for all } n > 1,$$

$$\Psi_n = P_{++}\Psi_{n-1}^- + \sum_{m=2}^{n-1}\Psi_{m-1}^+ P_{-+}\Psi_{n-m}^- + \Psi_{n-1}^+ P_{--}.$$

Proof. We consider the signs of the rate of the fluid before and after the event at time t_{m_n}.

At time t_{m_n}, if the rate transitions from $\mathcal{S}_+$ to $\mathcal{S}_+$, there cannot be other transitions before t_{m_n}, because otherwise $X_{m_n - 1} < X_{m_n}$, which would contradict the choice of m_n; hence $m_n = 1$. Furthermore, the interval (t_1, t) contains $n-1$ Poisson epochs, and $X(t) < X_1 < \min(X_2,\dots,X_{m_n})$; but, up to a relabelling of the epochs, this is exactly the definition of Ψ_{n-1}^-. So the probability of this case is $P_{++}\Psi_{n-1}^-$.

Symmetrically, if the rate goes from $\mathcal{S}_-$ to $\mathcal{S}_-$, then it must be the last, $m_n = n$, otherwise $X_{m_n+1} < X_{m_n}$, and it comes after a path in Ψ^+_{n-1}; this produces the term $\Psi^+_{n-1}P_{--}$.

In the case the rate goes from from $\mathcal{S}_+$ to $\mathcal{S}_-$, then there cannot be events before nor after t_{m_n}. Hence it must be the case that $m_n = n = 1$. So in Ψ_1 only we have a term P_{+-}, corresponding to the probability of this single transition.

Finally, if the rate changes from $\mathcal{S}_-$ to $\mathcal{S}_+$, then there must be at least one transition before and after it, so $2 \leq m_n \leq n-1$. In $(0, t_{m_n})$ we have $m_n - 1$ events and observe a path in $\Psi^+_{m_n-1}$; in (t_{m_n}, t) we have $n - m_n$ events and observe a path in $\Psi^-_{n-m_n}$. This produces the summands $\Psi^+_{m_n-1}P_{-+}\Psi^-_{n-m_n}$, for each possible value of m_n. □

3. FINITE APPROXIMATION

In a numerical algorithm, we need to truncate or approximate the sum (2), which has an infinite number of terms. We propose two different procedures to do this.

Algorithm A1. Note that $B_n = \sum_{k=1}^n \Psi_k^-$ is a non-decreasing sequence, and $\lim_{n\to\infty} B_n = \Psi(\infty) = \Psi$. Generalizing the approach in [5], whenever n is sufficiently large we can replace B_n with Ψ, which may be computed directly with various algorithms (see, e.g., [6]). More specifically, if n' is the smallest integer such that $(\Psi - B_{n'}) \sum_{n>n'} e^{-\lambda t}\frac{(\lambda t)^n}{n!} < \varepsilon$, then we can approximate $\Psi(t)$ with

$$\Psi'(t) = \sum_{n=0}^{n'} e^{-\lambda t}\frac{(\lambda t)^n}{n!}B_n + \Psi \sum_{n=n'+1}^{\infty} e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \qquad (4)$$

with error bounded by

$$\begin{aligned}\Psi'(t) - \Psi(t) &= \sum_{n=n'+1}^{\infty} (\Psi - B_n)e^{-\lambda t}\frac{(\lambda t)^n}{n!} \\ &\leq \sum_{n=n'+1}^{\infty} (\Psi - B_{n'})e^{-\lambda t}\frac{(\lambda t)^n}{n!} < \varepsilon.\end{aligned}$$

Algorithm A2. Another truncation strategy consists in swapping the order of summation in (2), obtaining

$$\Psi(t) = \sum_{k=1}^{\infty} \Psi_k^- \sum_{n=0}^{k-1} e^{-\lambda t}\frac{(\lambda t)^n}{n!}.$$

This sum can be truncated when Ψ_k is sufficiently small. If we stop the computation of the coefficients at the same index n' as above, we get

$$\Psi''(t) = \sum_{k=1}^{n'} \Psi_k^- \sum_{n=0}^{k-1} e^{-\lambda t}\frac{(\lambda t)^n}{n!} \leq \Psi(t). \qquad (5)$$

Note that $\Psi''(t) \leq \Psi(t) \leq \Psi'(t)$, so this method produces an explicit inclusion interval for each entry of $\Psi(t)$.

The quantities $\sum_{n=0}^{k-1} e^{-\lambda t}\frac{(\lambda t)^n}{n!}$ and $\sum_{n=n'+1}^{\infty} e^{-\lambda t}\frac{(\lambda t)^n}{n!}$ are related to the upper and lower incomplete Gamma functions by classical identities [4, Eqn. (8.69)], and they can be computed exactly with the routines included in the mathematical libraries of many languages.

4. COMPLEXITY ANALYSIS

We compare the complexities of various algorithms to compute $\Psi(t)$. The primitives required by all of them are linear algebra operations between matrices whose dimensions are either $|\mathcal{S}_+|$ or $|\mathcal{S}_-|$. The most general way to bound the cost of each of these operations is $O(N^3)$ floating point operations (flops). We shall use this expression in all the costs in the following.

Algorithms A1, A2: computing the recursion in Thm. 4, and then using (4) or (5), respectively, which give approximations from above and from below. They require $O(N^3(n')^2)$ flops: we compute n' steps of the recurrence in Thm. 4, and each step requires $O(n')$ linear algebra operations. The value of n' required to obtain a prescribed accuracy varies with t. We show in Section 5 (and, especially, in Fig. 2) how the two values are related in an example. We often consider A1 and A2 together because, once one computes the recursion in Thm. 4, both estimates can be obtained with a minimal time overhead.

Algorithm BST: the algorithm obtained by modifying the approach suggested in [5] (named after the initials of its authors). It is based on filling certain triangular arrays, each element of which is a matrix of size $|\mathcal{S}| \times |\mathcal{S}_-|$ in our modified version. There are r such arrays, where $1 \leq r \leq |\mathcal{S}_-|$ is the number of *distinct* negative rates, and each of them has size $O(n')^2$. Filling in each element requires $O(1)$ linear algebra operations, hence the total complexity of this algorithm is $O(N^3 r(n')^2)$ flops, which is a factor of r more than A1 or A2.

Algorithm LST: the algorithm based on inverse Laplace-Stieltjes transforms (LSTs). Our tests used the Euler algorithm in [1] for inversion, which requires evaluating the transform $\hat{\Psi}(s)$ in several nodes. The number of nodes recommended in [1] is $2d+1$, where d is the number of (decimal) digits of precision required. With ε equal to the machine precision u, $d = 16$. Each evaluation was carried out with a number of steps h of Newton's method [7, Alg. 4]; each step requires $O(1)$ linear algebra operations. This gives a total cost of $O(N^3 dh)$ flops. Note, h is related to the drift δ of the fluid by $\varepsilon = O(\delta^{2^h})$, due to quadratic convergence results for Newton's methods.

Several points. There is another important factor that works against LST. Often, one needs to evaluate $\Psi(t)$ for m different values of t. In A1/A2, the most expensive part is computing the recursion; but it is sufficient to do it once, with a truncation criterion n' determined by the largest of the values t. Thus, the total cost to compute $\Psi(t_1), \ldots, \Psi(t_m)$ is $O(N^3(n')^2 + N^2 n' m)$ flops for A1 and A2, where the second terms comes from applying (4) and/or (5) m times. For BST, a similar analysis holds; the complexity is $O(N^3 r(n')^2 + N^2 n' m)$ flops.

In contrast, we do not know of an established numerical scheme in the literature that allows one to compute inverse Laplace transforms at multiple points, so our best estimate for LST is m times the cost for one point, i.e., $O(N^3 dhm)$ flops. This fact makes A1 and A2 a very convenient choice when multiple evaluations are required.

A further remark is that a more careful analysis of the N^3 term for the linear algebra cost would show that A1, A2, BST scale better than LST when the matrix Q is sparse, or when one among $|\mathcal{S}_+|$ and $|\mathcal{S}_-|$ is much smaller than the other.

5. NUMERICAL EXPERIMENTS

We ran some tests to compare the various numerical algorithms. Where not specified otherwise, the truncation

Algorithm	$t = 0.1$	$t = 1.1$	$t = 9.9$
BST	6.6×10^{-16}	5.9×10^{-16}	1.2×10^{-15}
LST	6.1×10^{-11}	4.2×10^{-11}	4.1×10^{-11}
A1	6.5×10^{-16}	1.9×10^{-16}	3.8×10^{-16}
A2	2.8×10^{-16}	2.4×10^{-16}	8.4×10^{-16}

Table 1: Relative errors obtained with the various algorithms in a simple example.

Algorithm	$t = 0.1$	$t = 1.1$	$t = 9.9$	$t = 15$	$t = 0{:}15$
BST	0.17	0.39	3.54	6.11	6.11
LST	0.06	0.09	0.07	0.12	5.49
A1 & A2	0.08	0.09	0.26	0.64	0.64
n'	17	47	182	243	243

Table 2: CPU times (seconds) obtained with the various algorithms.

threshold ε in A1, A2, and BST is equal to the machine precision $\mathtt{u} \approx 2 \cdot 10^{-16}$.

Relative errors. We generated a toy model with $|\mathcal{S}_+| = 2, |\mathcal{S}_-| = 3$ with the MATLAB instructions

```
rng('default');
T = rand(N); T = T - diag(T*ones(N,1));
C = blkdiag(diag(0.2 * rand(Nplus,1)), ...
                -diag(rand(Nminus,1)));
```

which produced a model with a negative drift of -0.096.

We evaluated for each computed value $\tilde{X}$ the relative forward error $\|\tilde{X} - X\|_\infty / \|X\|_\infty$. Here X is an 'exact' reference value obtained by running both LST and A1 with higher working precision, as well as a higher number of nodes for LST, and checking that their results coincide up to the machine precision $\mathtt{u}$.

Table 1 confirms that LST (with normal working precision and the default number of nodes $d = 16$) can obtain only an accuracy of the order of 10^{-11}. These results are consistent with the theory in [1, Sect. 7], which predicts an accuracy (for well-behaved functions) of $0.6d$ decimal digits when working with $2d + 1$ nodes and d decimal digits of precision. Further tests show that this is the typical behaviour, at least for small values of $|\mathcal{S}|$, although computing these reference values is extremely slow.

Running times. We generated a more challenging test, with the same code but $|\mathcal{S}_+| = 10, |\mathcal{S}_-| = 11$. Table 2 shows the CPU times obtained for several values of t and, in the last column, for when 100 values equispaced in $[0, 15]$ are computed simultaneously. The results show that LST requires a constant time to compute each time sample, irrespective of t, while the other algorithms get slower as t increases. On the other hand, these algorithms can compute the 100 time samples of $\Psi(t)$ basically with no overhead w.r.t. the cost of computing the one with the largest t. This confirms the theory in Section 4.

To see how the time to compute the recursion in Thm. 4 varies with t and with the accuracy required, we plot in Fig. 2 these times, for various values of ε. The reason for the plateau observed for $\varepsilon = 10^{-5}$ is, at $t \approx 11$ the first return has already been observed w.p. $1 - 10^{-5}$, so $\Psi(t)$ coincides with $\Psi = \Psi(\infty)$ at that level of accuracy. Fig. 3 shows the complementary CDF of the first return time observed when starting from the state with the largest positive rate. Observing the two plots, one sees that the values of t that require a high CPU time correspond to the 'tail' of the distribution of $\Psi(t)$, and are needed only for a fine analysis at high levels of accuracy.

Note that these algorithms do not require a negative drift to work; the main effect of the drift is that $\Psi(t)$ and Ψ_n^- decay more slowly with t and n respectively, thus increasing the computational time.

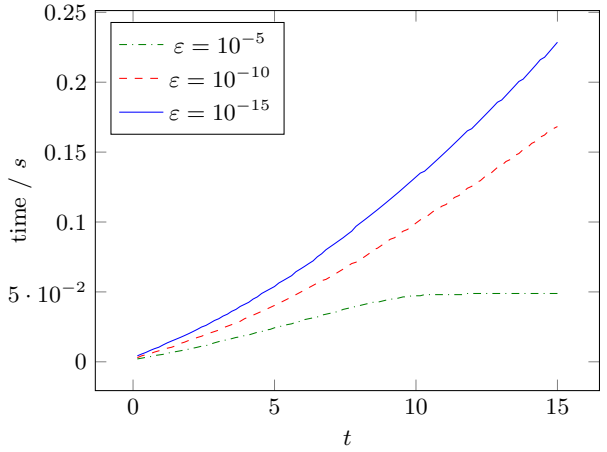


Figure 2: CPU times required by the new algorithm to compute the quantities Ψ_n to various required accuracies for $\Psi(t)$.

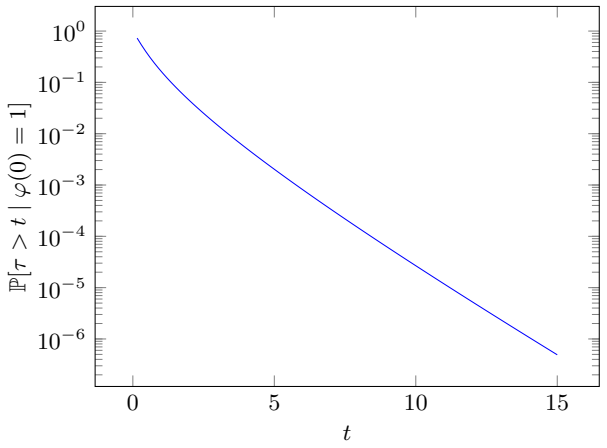


Figure 3: Complementary CDF (survival function) of the first return time starting from the state with the largest positive rate

6. CONCLUSIONS

We introduced a subtraction-free algorithm to compute the time-dependent first return probabilities of a fluid queue, $\Psi(t)$. Experimentally, this algorithm is faster than that obtained by generalizing the results in [5], and also faster and more accurate than the one based on the inverse LST, when one computes several values of $\Psi(t)$ simultaneously. Its main drawback is that the CPU time required increases when one is interested in computing very accurately the values of $\Psi(t)$ within the very tail of the distribution.

7. REFERENCES

[1] J. Abate and W. Whitt. A unified framework for numerically inverting Laplace transforms. *INFORMS J. Comput.*, 18(4):408–421, 2006.

[2] S. Ahn and V. Ramaswami. Efficient algorithms for transient analysis of stochastic fluid flow models. *J. Appl. Probab.*, 42(2):531–549, 2005.

[3] A. S. Alfa, J. Xue, and Q. Ye. Accurate computation of the smallest eigenvalue of a diagonally dominant M-matrix. *Math. Comp.*, 71(237):217–236, 2002.

[4] G. B. Arfken and H. J. Weber. *Mathematical methods for physicists*. Elsevier Academic press, sixth international edition, 2005.

[5] N. Barbot, B. Sericola, and M. Telek. Distribution of busy period in stochastic fluid models. *Stoch. Models*, 17(4):407–427, 2001.

[6] N. G. Bean, M. M. O'Reilly, and P. G. Taylor. Algorithms for return probabilities for stochastic fluid flows. *Stoch. Models*, 21(1):149–184, 2005.

[7] N. G. Bean, M. M. O'Reilly, and P. G. Taylor. Algorithms for the Laplace-Stieltjes transforms of first return times for stochastic fluid flows. *Methodol. Comput. Appl. Probab.*, 10(3):381–408, 2008.

[8] G. T. Nguyen and F. Poloni. Componentwise accurate fluid queue computations using doubling algorithms. *Numer. Math.*, 130(4):763–792, 2015.

[9] C. A. O'Cinneide. Entrywise perturbation theory and error analysis for Markov chains. *Numer. Math.*, 65(1):109–120, 1993.

MATRIX ANALYTIC METHODS FOR REFLECTED RANDOM WALKS WITH STOCHASTIC RESTARTS

Dario A. Bini
Dipartmento di Matematica,
Università di Pisa
L.go Bruno Pontecorvo, 5
56127, Pisa, Italy
bini@dm.unipi.it

Stefano Massei
EPFL
MA B2 515 (Bâtiment MA)
CH-1015 Lausanne,
Switzerland
stefano.massei@epfl.ch

Beatrice Meini
Dipartmento di Matematica,
Università di Pisa
L.go Bruno Pontecorvo, 5
56127, Pisa, Italy
meini@dm.unipi.it

Leonardo Robol
ISTI
Area della ricerca CNR
Via G. Moruzzi, 1
56124, Pisa, Italy
leonardo.robol@isti.cnr.it

ABSTRACT

We consider some Markov processes involving infinite state spaces, e.g., Quasi-Birth-and-Death (QBD) on $\mathbb{N}$ or $\mathbb{N}^2$. Often their transition probability matrices have Toeplitz or block Toeplitz structure, and the boundary conditions are encoded by low-rank corrections with finite support. Finding their steady state probability distribution can be recasted to finding minimal solutions to the quadratic matrix equation

$$AX^2 + BX + C = X,$$

where the matrices A, B, C are obtained from the blocks of the multilevel Toeplitz matrix. The state-of-the-art algorithms to solve these equations require to perform matrix operations involving these matrices. When the Markov process is defined on $\mathbb{N}^2$, the matrices A, B, C are semi-infinite. Recently, a computational framework for handling these problems without truncating the dimension of the state space has been proposed in [2]. This is achieved by approximating this kind of matrices as the sum of a banded semi-infinite Toeplitz plus a low-rank correction with finite support.

We propose an extension of this framework which allows to deal with more general situations such as processes involving restart events, where the process can move to level (or phase) 0 at any moment with a certain positive probability. This is motivated by the need for modeling processes that can incur in unexpected failures like computer system reboots. Algebraically, this gives rise to corrections with infinite support that can not be treated using the tools currently available in the literature. We present a theoretical analysis of an enriched space that, combined with appropriate algorithms, enables the numerical treatment of these problems. We demonstrate that the new approach considerably improves the method proposed in [2] as well, since there are cases where the coefficients A, B, C live in the space described there, but the solution X cannot be approximated without considering low-rank corrections with infinite support [1]. We show that this class of problem, previously numerical intractable, can now be handled in this framework.

We test our implementation on some case studies that confirm the applicability of the method.

1. REFERENCES

[1] D. Bini, S. Massei, B. Meini, and L. Robol. On quadratic matrix equations with infinite size coefficients encountered in qbd stochastic processes. *Numerical Linear Algebra with Applications*, 2018.

[2] D. A. Bini, S. Massei, and B. Meini. Semi-Infinite Quasi-Toeplitz Matrices with Applications to QBD Stochastic Processes. *Mathematics of Computation*, 2018.

MAM10 2019, Hobart, Australia

Extinction in lower Hessenberg branching processes with countably many types

Peter Braunsteins
The University of Melbourne
Grattan Street, Parkville, Victoria, 3010
braunsteins.p@unimelb.edu.au

Sophie Hautphenne
The University of Melbourne
Grattan Street, Parkville, Victoria, 3010
sophiemh@unimelb.edu.au

ABSTRACT

We consider a class of branching processes with countably many types which we refer to as *Lower Hessenberg branching processes*. These are multitype Galton-Watson processes with typeset $\mathcal{X} = \{0, 1, 2, \dots\}$, in which individuals of type i may give birth to offspring of type $j \leq i+1$ only. For this class of processes, we study the set S of fixed points of the progeny generating function. In particular, we highlight the existence of a continuum of fixed points whose minimum is the global extinction probability vector $\boldsymbol{q}$ and whose maximum is the partial extinction probability vector $\tilde{\boldsymbol{q}}$. In the case where $\tilde{\boldsymbol{q}} = \mathbf{1}$, we derive a global extinction criterion which holds under second moment conditions, and when $\tilde{\boldsymbol{q}} < \mathbf{1}$ we develop necessary and sufficient conditions for $\boldsymbol{q} = \tilde{\boldsymbol{q}}$.

MAM10 2019, Hobart, Australia

The probabilities of extinction in a branching random walk on a strip

Peter Braunsteins
The University of Melbourne
Melbourne
Australia
braunsteins.p@unimelb.edu.au

Sophie Hautphenne
The University of Melbourne
Melbourne
Australia
sophiemh@unimelb.edu.au

ABSTRACT

We consider a class of multitype Galton-Watson branching processes with countably infinite type set $\mathcal{X}$ whose mean progeny matrices have a block lower Hessenberg form. For these processes, we derive partial and global extinction criteria. Our approach involves embedding a finite-type explosive Galton-Watson process in a varying environment in the original infinite-type process, and then establishing asymptotic relationships between the two processes. We study the probability of extinction in sets of types $A \subseteq \mathcal{X}$, $\boldsymbol{q}(A)$. In particular, we develop conditions for $\boldsymbol{q}(A)$ to be different from the global and partial extinction probability vectors. We present an iterative method to compute the vectors $\boldsymbol{q}(A)$, and investigate their location in the set of fixed points of the progeny generating vector.

1. REFERENCES

[1] P. Braunsteins, G. Decrouez, and S. Hautphenne. A pathwise approach to the extinction of branching processes with countably many types. *Stochastic Processes and their Applications*, 2018. In Press.

[2] P. Braunsteins and S. Hautphenne. Extinction in lower hessenberg branching processes with countably many types. arXiv:1706.02919, 2018.

[3] P. Braunsteins and S. Hautphenne. The probabilities of extinction in a branching random walk on a strip. arXiv:1805.07634, 2018.

[4] S. Hautphenne, G. Latouche, and G. Nguyen. Extinction probabilities of branching processes with countably infinitely many types. *Advances in Applied Probability*, 45(4):1068–1082, 2013.

MAM10 2019, Hobart, Australia

On Fisher Information of Some Functions of Phase Type Variates

Pavithra Celeste R
Department of Mathematics
Indian Institute of Space Science and Technology
Thiruvananthapuram
pavithraceleste.17@res.iist.ac.in

T. G. Deepak
Department of Mathematics
Indian Institute of Space Science and Technology
Thiruvananthapuram
deepak@iist.ac.in

ABSTRACT

Bladt *et.al.* [2] introduced a method for obtaining FIM(Fisher Information Marix) for PH(Phase Type) class using the Expectation- Maximisation (EM) algorithm. In this article, we attempt to find out the *Fisher Information* for some functions of PH variates. We discuss the following cases : (i) when the function $g(X)$, of the PH variate X, is differentiable for all $X = x$ and either the derivative at x is strictly positive or negative, (ii) when the derivative of g is continuous and non zero for all but finite number of values of x and for every real number y, there exists $n = n(y)$ inverses and, (iii) when $Y = g(X)$ and g is invertible only in a finite interval and at each point y the function is having countable number of inverses. The FIM for the finite support PH variates, which comes under case (iii), introduced by Ramaswami and Viswanath [7], is computed using the EM algorithm.

Keywords: Finite support phase type distributions; Fisher information; EM algorithm; Functions of PH variates .

1. INTRODUCTION

Phase type (PH) distributions introduced by Neuts [5] form a dense family of distributions (in the metric of weak convergence of distributions) on $[0, \infty)$ and have found lot of applications in the area of applied probability. A PH distribution can be regarded as the distribution of the time until absorption in a finite state Markov chain with one absorbing state into which absorption is certain. In the recent past, many research papers have been appeared to study, the models which are governed by PH distributed time. Apart from their denseness property that makes them versatile as models, there are many other motivations for using PH distributions in statistical models. The most important ones come from their connection with Markov chains and matrix theory. While the former is offering much simplicity in various conditioning arguments occurred in the model analysis, the latter is helping us to develop more accurate and faster algorithms that make many models involving PH distributions computationally tractable.

Phase type distributions, however, can be used for modelling only non-negative random variables, and they have an infinite support. But, in practice, there are many random variables for which the distributions have finite support. Even though, theoretically, the denseness property of the PH class allow us to fit such distributions by a suitable PH distribution, it may require a representation of large order so that practically the fit may not render a realistic approximation. In order to cater to the needs of this situation, Ramaswami and Viswanath [7] introduced a new class of distributions derived from PH distributions called, phase type distributions with finite support (FSPH). Since these distributions are also based on phase type distributions, they bear a strong connection to Markov chains. Their densities are of the matrix exponential type giving thereby the ability to bring to bear all the tools of matrix computations.

Fisher information or more commonly called Fisher information matrix (FIM) plays a key role in uncertainty calculation and in other aspects of estimation for a wide range of statistical applications. It essentially describes the amount of information that the data provide about unknown parameters. The expected FIM, that is the expectation of the square of the gradient of the incomplete data log-likelihood functions at the maximum likelihood estimates (MLEs), can be used to calculate the Cramer-Rao lower bound and asymptotic distribution of the MLE. The role of FIM in the computation of the asymptotic distribution of the MLE enable us to use it effectively in testing of hypothesis and in the construction of confidence regions for the unknown parameters. In addition to compute the Fisher information contained in the sample observations assumed by a single variate, the same can be computed for the sample taken from a process.

The expectation-maximization (EM) algorithm introduced by Dempster, Laird and Rubin [3] is a well-known method to compute the MLE iteratively from the observed data. It is applied to problems in which the observed data is incomplete or the log-likelihood function of the observed data is too difficult to be solved to get the MLE directly from it. It provides a sequence of results obtained from the simple complete data log-likelihood function, and hence avoids calculations from the complicated incomplete data log-likelihood function. One of the major criticisms of the EM approach was that it cannot be directly used to obtain the FIM for the observed data since the EM algorithm does not automatically produce an estimate of the covariance matrix

MAM10 2019, Hobart, Australia

of the MLE. However, Oakes [6] provided a simple explicit formula for the matrix of second derivatives of the criterion function (the conditional expectation of the complete data log-likelihood given the observed data) invoked by the EM algorithm.

Estimation and statistical inference for PH distributions are of considerable importance when taking into consider its role in different areas of application. Asmussen, Nerman, and Olsson [1] was the first to develop a general approach to ML estimation of continuous PH distributions. Bladt, Esparza, and Nielsen [2] suggested an alternative method of calculating matrix exponentials and related integrals, appearing in the E-step in the EM algorithm, in order to speed up the execution of the algorithm considerably well for small and medium sized data sets. Apart from this, the main contributions of Bladt, Esparza, and Nielsen [2] are that of proposing methods for calculating the FIM for discrete and continuous PH distributions using formulae that are related to both EM algorithm and the Newton-Raphson approach.

In this paper, we discuss the computation of FIM of some functions of PH variates by the EM approach. We follow the method adopted by Bladt, Esparza, and Nielsen [2] - where they considered FIM for PH class - for computing FIM of some functions of PH variates .

2. FIM

We consider the following three types of functions of a PH variate X.

2.1 Case 1

Let the function $Y = g(X)$ be differentiable for all $X = x$ and the derivative g' at x is either strictly positive or negative. In this case the function g will be invertible. So, if we are given n observations $y_1, y_2, \ldots, y_n$ from Y, then by taking the inverses of these observations *i.e,* $g^{-1}(y_1), g^{-1}(y_2), \ldots, g^{-1}(y_n)$, we get n observations from $g^{-1}(Y) = X$, a PH variate. Hence we can compute the FIM of Y using that of X, which is already known. For example, the FIM of $logPH$ distribution, introduced by Ahn *et. al.* [8], can be computed in this manner.

2.2 Case 2

Let the derivative of g be continuous and non zero for all but finite number of values of x and for every real number there exist a finite number of inverse values.*i.e,* for every real number y, there exist a positive integer $n = n(y)$ and real numbers $x_1, x_2, \ldots, x_n$ such that, $g[x_k] = y,\ \ g'[x_k] \neq 0,\ \ k = 1, 2, \ldots, n(y)$.
This case can be illustrated with the example, $Y = |X - k|$ where k is any positive real number.

2.3 Case 3

Here we consider the case when $Y = g(X)$, where g is, in general, not an invertible function. Assume that in a finite interval g is invertible and at each point y, the function is having countable number of inverses, which are given by $h(y) + l(n), n = 0, 1, 2, \ldots$ where h is the inverse of g in the finite interval and l is a linear function of n. So once we are given a sample from Y, using the relation $X = h(Y) + l(n)$ we can obtain observations from a PH variate. Also, if F is the distribution function of X, then that of Y is,

$$\tilde{F}(y) = \sum_{n=0}^{\infty} [F(h(y) + l(n)) - F(l(n))].$$

Now using this distribution function and the MLE of the $PH-$variate X, we obtain the MLE as well as the FIM of Y.

As an example, we have considered two functions of PH variates namely $FSPH$ variate and *sine* function. In Oakes[6], the FIM is given as,

$$\begin{aligned} \frac{\partial^2 L(\theta; y)}{\partial \theta^2} &= \left\{ \frac{\partial^2 Q(\hat{\theta}/\theta)}{\partial \hat{\theta^2}} + \frac{\partial^2 Q(\hat{\theta}/\theta)}{\partial \theta \partial \hat{\theta}} \right\}_{\hat{\theta}=\theta} \\ \text{where } Q(\hat{\theta}/\theta) &= E_\theta \left(l_f(\hat{\theta}; x)/y \right). \end{aligned}$$

For finding the FIM of an $FSPH$ variate of order p, let $\theta = (\alpha_1, \alpha_2, \cdots$
$\alpha_{p-1}, t_1, T_{12}, \cdots T_{1p}, T_{21}, t_2, T_{23} \cdots T_{2p}, \cdots T_{p1}, T_{p2} \cdots T_{p,p-1}, t_p)$ be the parameter vector of order $p-1+p^2$ for the $FSPH_1(\alpha, T)$ distribution. Note that here we take $\alpha_p = 1 - \sum_{j=1}^{p-1} \alpha_j$ and $T_{ii} = -\sum_{\substack{j=1 \\ j \neq i}}^{p} T_{ij} - t_i$. Since the $FSPH$ distribution is being derived from the PH distribution, from the likelihood function of the latter, we get,

$$\begin{aligned} Q(\hat{\theta}/\theta) &= \sum_{i=1}^{p} \log(\hat{\alpha_i}) \sum_{k=1}^{n} \hat{B_i}^k + \sum_{i=1}^{p} \sum_{\substack{j=1 \\ j \neq i}}^{p} \sum_{k=1}^{n} \log(\hat{T_{ij}}) \hat{N_{ij}}^k \\ &\quad - \sum_{i=1}^{p} \sum_{j=1 j \neq i}^{p} \sum_{k=1}^{n} \hat{T_{ij}} \hat{Z_i}^k + \sum_{i=1}^{p} \sum_{k=1}^{n} \log(\hat{t_i}) \hat{N_i}^k \\ &\quad - \sum_{i=1}^{p} \sum_{k=1}^{n} \hat{t_i} \hat{z_i}^k \end{aligned}$$

where $\hat{B_i}^k, \hat{N_i}^k, \hat{N_{ij}}^k$ and $\hat{Z_i}^k$ are the estimates of the sufficient statistics B_i, N_i, N_{ij} and Z_i respectively related to the k^{th} sample observation, and B_i, N_i, N_{ij} and Z_i are given by,

- B_i, the number of trajectories that start in phase i, $i = 1, 2, \cdots, p$.
- N_i, the number of trajectories for which absorption occurs from phase i, $i = 1, 2, \cdots, p$.
- N_{ij}, the number of transitions that occur from phase i to phase j, $1 \leq i,\ j \leq p,\ i \neq j$.
- Z_i, the total sojourn time in phase i for all the n trajectories combined, for $i = 1, 2, \cdots, p$.

Then by substituting $\alpha = \sum_{j=1}^{p-1} \alpha_j e_j^T + \left(1 - \sum_{j=1}^{p-1} \alpha_j\right) e_p^T$ and using the FIM of PH distribution, we get the elements of the FIM for $FSPH_1(\alpha, T)$ distribution of order p as follows: For $i, j = 1, 2, \cdots p-1$, the $(i, j)^{th}$ element is given by

$$\frac{\partial U_i}{\partial \alpha_j} - \frac{\partial U_p}{\partial \alpha_j};$$

for $m = 1, 2, \cdots p-1$ and $i, j = 1, 2 \cdots p$, the $(ip-1+j, m)^{th}$ element is given by

$$\frac{\partial U_m}{\partial T_{ij}} - \frac{\partial U_p}{\partial T_{ij}} \text{ if } i \neq j,\ \frac{\partial U_m}{\partial t_i} - \frac{\partial U_p}{\partial t_i} \text{ if } i = j;$$

the $(m, ip-1+j)^{th}$ element is given by

$$\frac{\partial V_{ij}}{\partial \alpha_m} - \frac{\partial V_{ii}}{\partial \alpha_m} \text{ if } i \neq j, \ \frac{\partial W_i}{\partial \alpha_m} - \frac{\partial V_{ii}}{\partial \alpha_m} \text{ if } i = j;$$

and for $i, j, m, n = 1, 2 \cdots p$, the $(ip-1+j, mp-1+n)^{th}$ element is given by

$$\frac{\partial V_{ij}}{\partial T_{mn}} - \frac{\partial V_{ii}}{\partial T_{mn}} \text{ if } i \neq j, m \neq n, \qquad \frac{\partial V_{ij}}{\partial t_m} - \frac{\partial V_{ii}}{\partial t_m} \text{ if } i \neq j, m = n,$$
$$\frac{\partial W_i}{\partial T_{mn}} - \frac{\partial V_{ii}}{\partial T_{mn}} \text{ if } i = j, m \neq n, \qquad \frac{\partial W_i}{\partial t_m} - \frac{\partial V_{ii}}{\partial t_m} \text{ if } i = j, m = n,$$

where,

$$\begin{aligned} U_i &= \textstyle\sum_{k=1}^{n} \frac{e_i' e^{Ty_k} (I-e^T)^{-1} t}{f(y_k)} \\ W_i &= \textstyle\sum_{k=1}^{n} \frac{\alpha e^{Ty_k} (I-e^T)^{-1} e_i}{f(y_k)} \\ S_i &= -\textstyle\sum_{\substack{j=1 \\ j \neq i}}^{p} \hat{T_{ij}} - \hat{t_i} \\ \text{and } V_{ij} &= \textstyle\sum_{k=1}^{n} \frac{e_j' M^*(y_k, \alpha, T) e_i}{f(y_k)}, \end{aligned}$$

with,

$$\begin{aligned} M^*(y, \alpha, T) &= (I-e^T)^{-1} \ M(y, \alpha, T) + (I-e^T)^{-1} \\ & \quad M(1, \alpha, T) \ (I-e^T)^{-1} e^{Ty} \\ \text{and } M(y, \alpha, T) &= \int_0^y e^{T(y-u)} \ t\alpha e^{Tu} \ du. \end{aligned}$$

As an another example we have considered the case of $Y = sinX$, where X is a $PH(\alpha, T)$ variate, in which case we get the distribution function of Y as,

$$\begin{aligned} F(y) &= \left\{ \alpha T^{-1} (I - e^{2\pi T})^{-1} \left[e^{\pi T} (e^{\pi T} - e^{-\sin^{-1} yT}) + \right. \right. \\ & \quad \left. \left. e^{\sin^{-1} yT} - I \right] t \right\} \delta_{0 \leq y \leq 1} + \left\{ \alpha T^{-1} (I - e^{2\pi T})^{-1} \right. \\ & \quad \left. e^{\pi T} \left[e^{(\pi - \sin^{-1} y)T} - e^{\sin^{-1}(-y)T} \right] t \right\} \delta_{-1 \leq y \leq 0} \end{aligned}$$

where δ is the indicator function.
Proceeding as in the above case, we can compute the FIM of Y.

The Fisher information matrix is widely used in optimal experimental design, machine learning, physics, computational neuroscience etc. It is used in the formulation of test statistics as well as finding the confidence region of parameters under consideration. The inverse of the FIM gives the asymptotic variances and covariances of the maximum likelihood estimates of the parameters. The FIM, in general, is not invertible and the invertibility has to be verified for each model. Since the same phase type distribution has different representations, we can expect the non-singularity for the representation with minimum number of parameters. The three cases, that are discussed in this paper covers almost all classes of distributions which can be obtained from the class of PH- distributions. Many of them find applications in various fields including insurance and finance. For example, LogPH class distributions exhibits a heavy tail and has some nice tail properties that are well aligned with the extreme value theory. Since fitting heavy tailed loss data with parametric distributions being an important analytical task in the insurance field, LogPH, which can fit the whole data in a straightforward manner plays an important role in insurance field. In addition, some classes of distributions, which come under the above discussed cases, have denseness properties (denseness of $FSPH$ and $LogPH$ are proven). So the statistical inference of those classes can offer the properties of rich classes of distributions without separate analyses. Hence the parameter estimation and the computation of FIM play an important role in the statistical inference of many distributions which are having practical applications.

3. REFERENCES

[1] S. Asmussen,O. Nerman and M. Olsson. Fitting phase-type distributions via the EM algorithm. *Scand. J. Statist.* **1996**, *23*, 419-441.

[2] M. Bladt, L. J. Esparza and B. F. Nielsen. Fisher information and statistical inference for phase-type distributions. *J. Appl. Prob. Spec.* **2011**, *48A*, 277-293.

[3] A. P. Dempster,N. M. Laird and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Statis. Soc.: Series B (Methodological).* **1977**,*39*, 1-38.

[4] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling* , SIAM & ASA: Philadelphia, 1999.

[5] M. F. Neuts. *Matrix Geometric Solutions in Stochastic Models*, Johns Hopkins University Press: Baltimore & London, 1981.

[6] D. Oakes. Direct calculation of the information matrix via the EM. *J. R. Statis. Soc.: Series B (Statistical Methodology).* **1999**, *61*(2), 479-482.

[7] V. Ramaswami, and N. C.Viswanath. Phase type distributions with finite support. *Stoch. Models.* **2014**, *30*, 576-597.

[8] Soohan Ahn, Joseph H.T Kim and Vaidyanathan Ramaswami. A new class of models for heavy tailed distributions in finance and insurance risk. *Insurance: Mathematics and Economics.* **2012**, *51*, 43-52

Model for the evolution of the family of gene duplicates

Jiahao Diao
University of Tasmania
Churchill Avenue
Hobart, Tasmania
jiaha.diao@utas.edu.au

Tristan L. Stark
Temple University
1900 N 12th Street
Philadelphia, USA
tristan.stark@temple.edu

David A. Liberles
Temple University
1900 N 12th Street
Philadelphia, USA
tuf77157@temple.edu

Małgorzata M. O'Reilly *†
University of Tasmania
Churchill Avenue
Hobart, Tasmania
malgorzata.oreilly@
utas.edu.au

Barbara R. Holland *
University of Tasmania
Churchill Avenue
Hobart, Tasmania
barbara.holland@
utas.edu.au

ABSTRACT

Consider a Markov model for the evolution of the family of genes proposed in [1, Section 10], in which the state (n, m) records the number copies $n = 0, 1, 2, \ldots$, and the number of *redundant* copies $m = 0, 1, , \ldots, n$ of a gene in the family. By redundant copies we mean copies whose loss will not result in the loss of the functions of the gene when these are preseved by some other genes in the family.

The transition rates between the state of the Markov model in [1, Section 10] depend on the following key processes: duplication of a gene; loss of one copy of a gene; one copy acquiring a new function (neofunctionalisation); and a loss of some regulatory region in one of the genes which leads to a number of genes required to fulfill some function (subfunctionalisation).

The aim of this work is to develop suitable expressions for the transition rates function. As an initial inspiration of our analysis, we consider a Markov model with a more detailed state, represented as a matrix, where rows corresponds to the various genes in the family, and columns to their functions. That is, state is a binary matrix $\mathbf{A} = [A_{ij}]_{i=1,\ldots,M;j=1,\ldots,Z}$, where M is the number of genes in the family, and Z the number of functions, such that $A_{ij} = 1$ when function j is performed by some regulatory region of gene i, and 0 otherwise.

Although we suspect that the model with the detailed representation will not be tractable for developing analytic solutions (except in very small cases) we propose to use it as a simulation model to give insight into suitable transition rates in the reduced state space representation.

1. REFERENCES

[1] A. I. Teufel, J. Zhao, M. O'Reilly, L. Liu, and D. A. Liberles. On mechanistic modeling of gene content evolution: birth-death models and mechanisms of gene birth and gene retention. *Computation*, 2(3):112–130, 2014.

*We would like to thank the Australian Research Council for funding this research through Discovery Project DP180100352.
†Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers.

MAM10 2019, Hobart, Australia

Characterization of the Boundary of the Set of Matrix-Exponential Distributions with Only Real Poles

Qi-Ming He
Department of Management Sciences, University of Waterloo
200 University Avenue West
Waterloo, Ontario, Canada
N2L 3G1
q7he@uwaterloo.ca

Mark Fackrell
School of Mathematics and Statistics
University of Melbourne
Melbourne, Vic 3010, Australia
fackrell@unimelb.edu.au

Peter Taylor
School of Mathematics and Statistics
University of Melbourne
Melbourne, Vic 3010, Australia
taylorpg@unimelb.edu.au

ABSTRACT

We study the boundary of the set of matrix-exponential distributions with m real poles. We show that the boundary consists of three parts: two parts in some lower dimensional spaces and a curved part. The curved part is characterized as the union of convex hulls of solutions from $m-2$ curves each generated by three Coxian distributions. The characterization of the boundary leads to algorithms for constructing the set of matrix-exponential distributions.

1. REFERENCES

[1] Bean N, Fackrell M, Taylor P (2008) Characterization of matrix-exponential distributions. Stochastic Models 24 (3): 339-363.

[2] Dehon M, Latouche G (1982) A geometric interpretation of the relations between the exponential and generalized Erlang distributions. Advances in Applied Probability 14: 885-897.

[3] Fackrell M (2003) Characterization of Matrix-exponential Distributions. PhD thesis, School of Applied Mathematics, University of Adelaide, South Australia, 2003. Available at http://thesis.library.adelaide.edu.au/public/adt-SUA20051207.123257/.

[4] Neuts MF (1981) Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach. The John Hopkins University Press: Baltimore.

[5] O'Cinneide CA (1990) Characterization of phase-type distributions. Stochastic Models 6: 1-57.

MAM10 2019, Hobart, Australia

Markovian decision-support model for patient-to-ward assignment problem in a random environment

Mojtaba Heydar [†]
University of Tasmania
TAS 7001, Australia
mojtaba.heydar@utas.edu.au

Małgorzata M. O'Reilly [*†]
University of Tasmania
TAS 7001, Australia
malgorzata.oreilly@utas.edu.au

ABSTRACT

The purpose of this research is to model the decision making process in a hospital in which patients are allocated to beds in different wards, according to their needs, priorities, as well as availability of the resources. The problem analysed here is a generalisation of the advanced patient admission scheduling (APAS) problem studied in [1, 2, 6], in which patients are admitted and alocated to the hospital capacitated resources [2] over some planning horizon.

We construct a Markov Decision Process (MDP), in which, at the start of the planning period, arrived patients are assigned to suitable wards in a manner that optimises the performance measures of interest. We build on the MDP model in [7], in which decisions are made at the time of arrival or departure of a patient. Here, we assume that the decisions are made at the start of the time period of some fixed length, and so the allocation involves several arrived patients.

Assume that the information about the system at the start of the planning period is given as state $s = ([n_{k,i}]_{\mathcal{K}\times\mathcal{I}}, [q_i]_{1\times\mathcal{I}})$, where $n_{k,i}$ is the number of type-i patients in ward k, with $i \in \mathcal{I}$, $k \in \mathcal{K}$, and q_i is the number of newly arrived type-i patients (or the expected number of type-i patients to arrive during the current period), who are yet to be assigned to the wards.

Here, $\mathcal{I} = \{1, 2, \ldots, I\}$ is the set of all patient types, where a type may correspond to the medical needs of the patient, their priority, their age, gender, and other relevant features. We assume that the set $\mathcal{I}$ is ordered according to the priorities, so that 1 and I represents the least and the most severe patients, respectively. Furthermore, $\mathcal{K}$ is the set of the wards in the hospital system. There are $\mathcal{B}_k$ beds in ward $k \in \mathcal{K}$.

[*]We would like to thank the Australian Research Council for funding this research through Linkage Project LP140100152.
[†]ARC Centre of Excellence for Mathematical and Statistical Frontiers.

MAM10 2019, Hobart, Australia

Suppose that type-i patients arrive according to a Poisson process with the rate λ_i. Let $Q_i(t)$ be the random variable recording the number of arrivals of type-i patients during the time period of length t. The distribution of $Q_i(t)$ is Poisson, $Q_i(t) \sim Poi(\lambda_i t)$.

The newly assigned type-i patient will stay in ward k for the duration of time $LoS_{k,i}$, referred to as the length of stay, according to some discrete Phase-type distribution [3, 4], $PH(\boldsymbol{\gamma}^{(k,i)}, \mathbf{P}^{(k,i)})$. Let $Z_{k,i}(t)$ be the random variable recording the number of departures of type-i patients from ward k during the time period of length t. The distribution of $Z_{k,i}(t)$ is binomial, $Z_{k,i}(t) \sim Bin(n_{k,i}, p_{k,i}(t))$, where

$$p_{k,i}(t) = Pr(LOS_{k,i} \leq t) = 1 - \boldsymbol{\gamma}^{(k,i)} \left(\mathbf{P}^{(k,i)}\right)^t \mathbf{1},$$

for all $t = 0, 1, 2, \ldots$, where $\mathbf{1}$ is a column vector of ones of appropriate size.

After observing state $s = ([n_{k,i}]_{\mathcal{K}\times\mathcal{I}}, [q_i]_{1\times\mathcal{I}})$, a decision $a = ([x_{k,i}]_{\mathcal{K}\times\mathcal{I}}, [y_{k,\ell,i}]_{\mathcal{K}\times\mathcal{K}\times\mathcal{I}})$ is made about where to assign the patients, which involves assigning $x_{k,i}$ type-i patients to ward k, and transfering $y_{k,\ell,i}$ type-i patients from ward k to ℓ. The resulting post-decision state is some $\bar{s} = ([\bar{n}_{k,i}]_{\mathcal{K}\times\mathcal{I}})$, with

$$\bar{n}_{k,i} = n_{k,i} + x_{k,i} + \sum_{k\in\mathcal{K}} y_{k,\ell,i} - \sum_{\ell\in\mathcal{K}} y_{k,\ell,i}.$$

Following the decision a, the process transitions from state s to some state $s' = \left([n'_{k,i}]_{\mathcal{K}\times\mathcal{I}}, [q'_i]_{1\times\mathcal{I}}\right)$ with $n'_{k,i} = \bar{n}_{k,i} - z_{k,i}$, according to the probability given by

$$\Pr\{s' |\ (s,a)\} = \prod_i \Pr(Q_i(t) = q'_i) \prod_{k,i} \Pr(Z_{k,i}(t) = z_{k,i}).$$

The above model describes the evolution of the system that depends on decision making process as well as random arrivals and departures in a given time period of length t. We formulate suitable value function and constraints, and consider the corresponding infinite-horizon MDP, in which the aim is to optimise the long-run discounted total mean cost, given discount factor α.

Since solving large-scale infinite-horizon MDPs is a challenging task, we consider suitable approximation techniques, such as Approximate Dynamic Programming (ADP)[5], in order to find an approximation that represent the value function of the MDP.

1. REFERENCES

[1] Y. Gocgun and M. L. Puterman. Dynamic scheduling with due dates and time windows: an application to chemotherapy patient appointment booking. *Health care management science*, 17(1):60–76, 2014.

[2] P. J. Hulshof, M. R. Mes, R. J. Boucherie, and E. W. Hans. Patient admission planning using approximate dynamic programming. *Flexible services and manufacturing journal*, 28(1-2):30–61, 2016.

[3] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modelling, 1st edition. Chapter 2: PH Distributions*. ASA SIAM, 1999.

[4] M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models: an Algorithmic Approach, Chapter 2: Probability Distributions of Phase Type*. Dover Publications Inc., 1981.

[5] W. B. Powell. *Approximate Dynamic Programming: Solving the curses of dimensionality*. John Wiley & Sons, 2007.

[6] S. Samiedaluie, B. Kucukyazici, V. Verter, and D. Zhang. Managing patient admissions in a neurology ward. *Operations Research*, 65(3):635–656, May-June 2017.

[7] E. Trainer, M. M. O'Reilly, M. Fackrell, P. G. Taylor, A. Tirdad, and M. Heydar. Stochastic model for patient-bed asignment problem with random arrivals and departures. *In preparation*.

High order low variance matrix-exponential distributions*

Gábor Horváth
Dept. of Networked Systems and Services
Budapest University of Technology and Economics
Magyar Tudósok Körútja 2,
1117 Budapest, Hungary
ghorvath@hit.bme.hu

Illés Horváth
MTA-BME Information Systems Research Group
Magyar Tudósok Körútja 2,
1117 Budapest, Hungary
pollux@math.bme.hu

Miklós Telek
Dept. of Networked Systems and Services
Budapest University of Technology and Economics
Magyar Tudósok Körútja 2,
1117 Budapest, Hungary
telek@hit.bme.hu

ABSTRACT

This paper presents matrix-exponential (ME) distributions, whose squared coefficient of variation (SCV) is very low. Currently there is no symbolic construction available to obtain the most concentrated ME distributions, and the numerical optimization-based approaches to construct them have many pitfalls too. We present a numerical optimization-based procedure which avoids numerical issues.

Keywords: Non-negative matrix-exponential functions, Matrix-exponential distributions, Numerical optimization, Coefficient of variation

1. INTRODUCTION

Highly concentrated matrix exponential functions play an important role in many research fields, for example, they turned out to be essential for numerical inverse Laplace transform methods as well [6].

The least varying phase type (PH) distribution of order N is known to be the Erlang distribution [1] with SCV=$1/N$ (defined as $\frac{\mu_0\mu_2}{\mu_1^2}-1$, where $\mu_i, i=0,1,2$, are the moments of the distribution). The least varying ME distribution for order N much less known. It is known that for order 2 the class of ME distributions is identical to the class of PH distributions, and it is also known that there exists order 3 ME distribution with SCV=0.200902 $< 1/3$, but it is still only a conjecture that this is the least varying order 3 ME distribution. Concentrated ME distributions are provided in [2] up to order 17 and in [5] up to order 47. These preliminary results indicate that the minimal SCV of order N ME distributions tends to be less than $2/N^2$. In this work, we propose numerical procedures by which much higher order concentrated ME distributions can be computed and based on that we refine the dependence of the minimal SCV on the order.

*This work is supported by the OTKA K-123914 project

MAM10 2019, Hobart, Australia

2. CONCENTRATED ME DISTRIBUTIONS

Definition 1. *Order N* ME functions *(referred to as ME(N)) are given by*

$$f(t) = \underline{\alpha} e^{\mathbf{A}t}(-\mathbf{A})\mathbb{1}, \tag{1}$$

where $\underline{\alpha}$ is a real row vector of size N, $\mathbf{A}$ is a real matrix of size $N \times N$ and $\mathbb{1}$ is the column vector of ones of size N, and $\underline{\alpha}$ is such that $\underline{\alpha}\mathbb{1} > 0$.

If $f(t) \geq 0, \forall t \geq 0$ and $\underline{\alpha}\mathbb{1} = 1$ then $f(t)$ is the probability density function of a ME distribution.

According to (1), vector $\underline{\alpha}$ and matrix $\mathbf{A}$ define the matrix exponential function. We refer to the pair $(\underline{\alpha}, \mathbf{A})$ as *matrix representation* in the sequel.

An ME distribution is said to be concentrated when its squared coefficient of variation

$$SCV(f(t)) = \frac{\mu_0\mu_2}{\mu_1^2} - 1, \tag{2}$$

is low. In (2), μ_i denotes the ith moment, defined by $\mu_i = \int_{t=0}^{\infty} t^i f(t)dt$ for $i = 0, 1, 2$. SCV is insensitive to multiplication and scaling, i.e. $SCV(f(t)) = SCV(c_1 f(c_2 t))$.

Although matrix-exponential functions have been used for many decades, there are still many questions open regarding their properties. Such an important question is how to decide efficiently if a matrix-exponential function is non-negative $\forall t > 0$. In general, $f(t) \geq 0, \forall t > 0$ does not necessarily hold for given $(\underline{\alpha}, \mathbf{A})$ parameters, unless it has been constructed to be always non-negative. In this paper, we are going to restrict our attention to such a special construction, the exponential-cosine square functions.

For the least varying ME(N) distributions only conjectures are available for $N \geq 3$ [2]. According to the current conjecture for odd N, the most concentrated ME(N) distribution belongs to a special subset of ME(N) given by the definition below.

Definition 2. *The set of* exponential cosine-square *functions of order n has the form*

$$f^+(t) = e^{-t} \prod_{i=1}^{n} \cos^2\left(\frac{\omega t - \phi_i}{2}\right). \tag{3}$$

An exponential cosine-square function is defined by $n+1$ parameters: ω and ϕ_i for $i = 1, \ldots, n$. An exponential cosine-square function is a matrix exponential function. Although

the representation in (3), which we refer to as the *cosine-square representation*, is not a matrix representation, [5, Appendix A] presents the associated matrix representation of size $N = 2n + 1$. Consequently, the set of exponential cosine-square functions of order n is a special subset of ME(N) (where $N = 2n + 1$) which, by construction, is non-negative. The SCV of an exponential cosine-square function is a complicated function of the parameters, whose minimum does not exhibit a closed analytic form. That is why we have resorted to the following numerical problem. For a given odd order $N = 2n + 1$, we are looking for efficient numerical methods for finding the ω and ϕ_i $(i = 1, \ldots, n)$ parameters which result in a low SCV. For efficient numerical minimization of the SCV for $N > 47$ (i.e., $n > 23$) we need

i) an accurate computation of the SCV based on the parameters with low computational cost and

ii) an efficient optimization procedure with low computational cost.

In this paper we present a method that addresses i) in Section 3, and one that addresses ii) in Section 4.

3. EFFICIENT COMPUTATION OF THE SQUARED COEFFICIENT OF VARIATION

To evaluate the objective function of the optimization, namely the SCV, we need efficient methods to compute μ_0, μ_1 and μ_2. Deriving the μ_i parameters based on (3) is difficult (for large N). Hence we propose to compute them based on a different representation.

3.1 The hyper-trigonometric representation

The following theorem defines the *hyper-trigonometric form* of the exponential cosine-square functions and provides a recursive procedure to obtain its parameters from $\omega, \phi_i, i = 1, \ldots, n$.

THEOREM 1. *An order $N = 2n + 1$ exponential cosine-square function can be transformed to a hyper-trigonometric representation of form*

$$f^+(t) = c^{(n)} \cdot e^{-t} + e^{-t} \sum_{k=1}^{n} a_k^{(n)} \cos(k\omega t) + e^{-t} \sum_{k=1}^{n} b_k^{(n)} \sin(k\omega t), \tag{4}$$

where $c^{(n)} = \frac{1}{2} a_0^{(n)}$ and the coefficients $a_k^{(n)}$, $b_k^{(n)}$ are calculated recursively:

- *for $n = 1$:*

$$a_0^{(1)} = 1, \quad b_0^{(1)} = 0, \quad a_1^{(1)} = \frac{1}{2}\cos\phi_1, \quad b_1^{(1)} = \frac{1}{2}\sin\phi_1, \tag{5}$$

- *for $k > n, n \geq 1$:*

$$a_k^{(n)} = b_k^{(n)} = 0,$$

- *for $k = 0, n \geq 1$:*

$$a_0^{(n)} = \frac{1}{2} a_0^{(n-1)} + \frac{1}{2} a_1^{(n-1)} \cos\phi_n + \frac{1}{2} b_1^{(n-1)} \sin\phi_n, \tag{6}$$

$$b_0^{(n)} = 0, \tag{7}$$

- *for $1 \leq k \leq n, n \geq 2$*

$$a_k^{(n)} = \frac{1}{2} a_k^{(n-1)} + \frac{1}{2} \frac{a_{k-1}^{(n-1)} + a_{k+1}^{(n-1)}}{2} \cos\phi_n + \frac{1}{2} \frac{b_{k+1}^{(n-1)} - b_{k-1}^{(n-1)}}{2} \sin\phi_n, \tag{8}$$

$$b_k^{(n)} = \frac{1}{2} b_k^{(n-1)} + \frac{1}{2} \frac{b_{k-1}^{(n-1)} + b_{k+1}^{(n-1)}}{2} \cos\phi_n + \frac{1}{2} \frac{a_{k-1}^{(n-1)} - a_{k+1}^{(n-1)}}{2} \sin\phi_n. \tag{9}$$

The hyper-trigonometric representation makes it possible to express the Laplace transform (LT) and the moments μ_i in a simple and compact way.

COROLLARY 1. *The LT and the $\mu_i, i = 0, 1, 2$ moments of the exponential cosine-square function are given by*

$$f^*(s) = \frac{c^{(n)}}{1+s} + \sum_{k=1}^{n} \frac{a_k^{(n)}(1+s) + b_k^{(n)} k\omega}{(1+s)^2 + (k\omega)^2}, \tag{10}$$

and

$$\mu_0 = c^{(n)} + \sum_{k=1}^{n} \frac{a_k^{(n)} + b_k^{(n)} k\omega}{1 + (k\omega)^2}, \tag{11}$$

$$\mu_1 = c^{(n)} + \sum_{k=1}^{n} \frac{a_k^{(n)} + 2b_k^{(n)} k\omega - a_k^{(n)}(k\omega)^2}{(1 + (k\omega)^2)^2},$$

$$\mu_2 = 2c^{(n)} + \sum_{k=1}^{n} \frac{2a_k^{(n)} + 6b_k^{(n)} k\omega - 6a_k^{(n)}(k\omega)^2 - 2b_k^{(n)}(k\omega)^3}{(1 + (k\omega)^2)^3}.$$

3.2 Numerical computation of the moments

Theorem 1 together with Corollary 1 provides a very efficient explicit method to compute the SCV based on the parameters $\omega, \phi_i, i = 1, \ldots, n$.

There is one numerical issue that has to be taken care of when applying this numerical procedure with floating point arithmetic for large values of n. To evaluate the SCV, coefficients $a_k^{(n)}, b_k^{(n)}, c^{(n)}$ need to be obtained from the ω and $\phi_i, i = 1, \ldots, n$ parameters. The recursion defined in Theorem 1 involves multiplications between bounded numbers (sine and cosine always fall into $[-1, +1]$), which is beneficial from the numerical stability point of view, but subtractions are unfortunately also present, leading to loss of precision. To overcome this loss of precision, we introduced increased precision floating point arithmetic both in our Mathematica and C++ implementations[1]. Mathematica can quantify the precision loss, enabling us to investigate this issue experimentally. According to Figure 1, the number of accurate decimal digits lost when evaluating the SCV from the ω, ϕ_i parameters (computed by the `Precision` function of Mathematica), denoted by L_n, is nearly linear and can be approximated by

$$L_n \approx 1.487 + 0.647n. \tag{12}$$

[1]In C++ we used to `mpfr` library for multi-precision computations

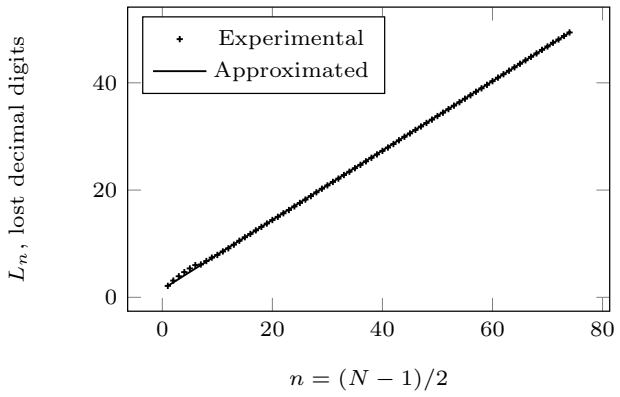


Figure 1: The precision loss while computing the SCV

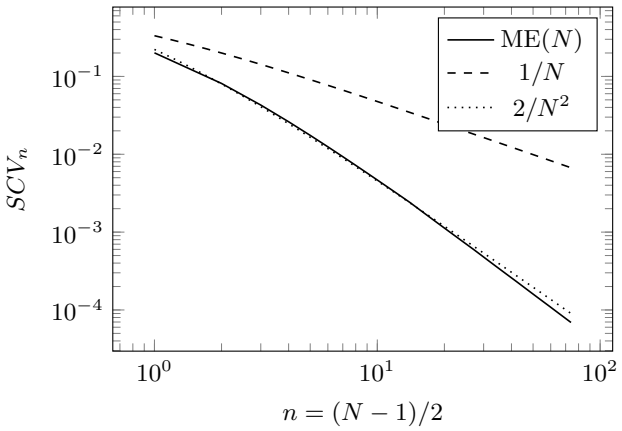


Figure 2: The minimal SCV of the exponential cosine-square functions as the function of n in log-log scale

In the forthcoming numerical experiments we have set the floating point precision to $L_n + 16$ decimal digits to obtain an accuracy of results up to 16 decimal digits, and this precision setting eliminated all numerical issues.

It is important to note that the high precision is needed only to calculate the $a_k^{(n)}, b_k^{(n)}, c^{(n)}$ coefficients and the SCV itself. Representing parameters ω, ϕ_i themselves does not need extra precision, and the resulting exponential cosine-square function $f(t)$ can be evaluated with machine precision as well (in the range of our interest, $n \leq 1000$).

A basic pseudo-code of the computation of the SCV with the indications where high precision is needed is provided by Algorithm 1.

Algorithm 1 Pseudo-code for the computation of the SCV

1: **procedure** COMPUTESCV($n, \omega, \phi_i, i = 1, \ldots, n$)
2: Compute the required precision, L_n, from (12)
3: Convert $\omega, \phi_i, i = 1, \ldots, n$ to $L_n + 16$ digits precision
4: Calculate $a_k^{(n)}, b_k^{(n)}, c^{(n)}, k = 1, \ldots, n$, recursively by Theorem 1 (**high precision**)
5: Calculate moments μ_0, μ_1, μ_2 according to (11) (**high precision**)
6: Calculate $SCV = \frac{\mu_0 \mu_2}{\mu_1^2} - 1$ (**high precision**)
7: Convert SCV to machine precision
8: **return** SCV
9: **end procedure**

4. MINIMIZING THE SQUARED COEFFICIENT OF VARIATION

Given the size of the representation $N = 2n + 1$, the $f^+(t)$ function providing the minimal SCV is obtained by minimizing (2) subject to ω and $\phi_i, i = 1, \ldots, n$. The form of the SCV does not allow a symbolic solution, and its numerical optimization is challenging too. The surface to optimize has many local optima, hence simple gradient descent procedures failed to find the global optimum and are sensitive to the initial guess.

4.1 Optimizing the parameters

In the numerical optimization of the parameters, we had success with evolutionary optimization methods, in particular with *evolution strategies*. The results introduced in [5] were obtained by one of the simplest evolution strategies, the Rechenberg method [7]. In [5], it was the high computational demand of the numerical integration needed to obtain the SCV and its reduced accuracy that prevented the optimization for $N > 47$ ($n > 23$).

However, computing the SCV based on the hyper-trigonometric representation using the results of Section 3.1 allows us to evaluate the moments orders of magnitudes faster and more accurately, enabling the optimization for higher n values. With the Rechenberg method ([7], also referred to as (1+1)-ES in the literature) it is possible to obtain low SCV values relatively quickly for orders as high as $n = 125$, but these values are suboptimal in the majority of cases.

With more advanced evolution strategies the optimal SCV can be approached better. Our implementation supports the covariance matrix adoption evolution strategy (CMA-ES [3]), and one of its variants, the BIPOP-CMA-ES with restarts [4]. Starting from a random initial guess, we got very low SCV values much quicker with the CMA-ES than with the (1+1)-ES with similar suboptimal minimum values (cf. Fig. 4). The limit of applicability of CMA-ES is about $n = 180$. The best solution (lowest SCV for the given order), however, was always provided by the BIPOP-CMA-ES method, although it is by far the slowest among the three methods we studied. In fact, we believe that BIPOP-CMA-ES returned the global optimum for $n = 1, \ldots, 74$, and we investigate the properties of those solutions in the next sections. For $n > 74$, we can still compute low SCV functions with the BIPOP-CMA-ES method, but its computation time gets to be prohibitive, and we are less confident about the global minimality of the results.

For our particular problem, the running time, T, and the quality of the minimum, Q (how low the SCV is), obtained by the different optimization methods can be summarized as follows

$$T_{\text{CMA-ES}} < T_{(1+1)\text{-ES}} << T_{\text{BIPOP-CMA-ES}},$$
$$Q_{\text{CMA-ES}} \sim Q_{(1+1)\text{-ES}} < Q_{\text{BIPOP-CMA-ES}}.$$

4.2 Properties of the minimal SCV solutions

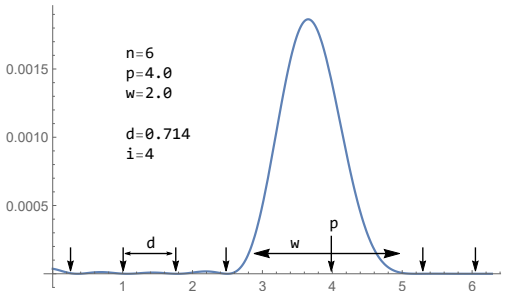


Figure 3: The spike and the zeros of $f^+(t)$

The minimal SCV values obtained by the BIPOP-CMA-ES optimization, which we consider as being optimal for $n = 1, \ldots, 74$, are depicted in Figure 2. Apart from the minimal SCV values of the exponential cosine-square functions, Figure 2 also plots $1/N$ and $2/N^2$, for comparison. The $SCV = 1/N$ is known to be the minimal SCV value of phase-type (PH) distributions of order N [1], which form a subset in the set of ME distributions by assuming positive off-diagonal elements for $\mathbf{A}$ and nonnegative elements for $\underline{\alpha}$. The $2/N^2$ curve is reported to be the approximate decay rate in [5], up to $n = 23$ ($N = 47$).

Figure 2 indicates that the SCV decreases much faster than $1/N$ and a bit faster than $2/N^2$. Indeed, $2/N^2$ is a good approximation up to $n = 23$, but the decay seems to decrease below $2/N^2$ for $n > 23$.

5. HEURISTIC OPTIMIZATION WITH 3 PARAMETERS

According to the previously discussed approach the number of parameters to optimize increases with n. This drawback limits the applicability of the general optimization procedures to about $n \leq 74$ in case of BIPOP-CMA-ES and about $n \leq 180$ in case of the basic CMA-ES. By these n values the optimization procedure takes several days to terminate on our average PC clocked at 3.4 GHz.

While the $f^+(t)$ function obtained this way for $n = 180$ have an extremely low ($\approx 10^{-5}$) SCV already, some applications might benefit from ME distributions with even lower SCV. To overcome this limitation we developed a suboptimal heuristic procedure, that aims to obtain low SCV for a given large order n.

Our heuristic procedure has to optimize only three parameters, independent of the order n. The procedure is based on the assumption that the location of the spike in the $(0, 2\pi)$ cycle of the cosine-squared function plays the most important role in the SCV, and the exact values of the ϕ_k parameters are less important, the only important feature is that the cosine-squared terms characterized by the ϕ_k parameters should suppress $f^+(t)$ uniformly in the $(0, 2\pi)$ cycle – apart from the spike (cf. Figure 3).

Based on this assumption we set the ϕ_k parameters of the cosine-squared terms equidistantly. This way the position of the spike (p) and its width (w) inside the $(0, 2\pi)$ interval completely define the ϕ_k values for a given order n.

The distance of the ϕ_k parameters (d) and the number of ϕ_k parameters before the spike (i) can be computed from p and w by

$$d = \frac{2\pi - w}{n}, \quad i = \left\lfloor \frac{p - w/2}{d} + \frac{1}{2} \right\rfloor, \tag{13}$$

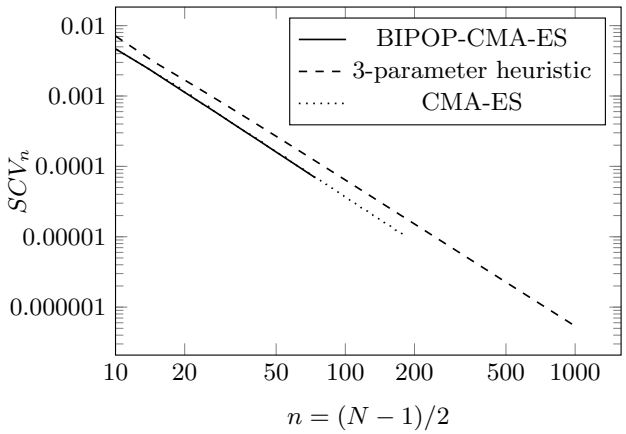


Figure 4: The minimal and the heuristic SCV as a function of order n in log-log scale

and for $k = 1, \ldots, n$ the ϕ_k parameters are

$$\phi_k + \pi = \begin{cases} (k - 1/2)d & \text{if } k \leq i, \\ (k - 1/2)d + w & \text{if } k > i. \end{cases} \tag{14}$$

Figure 4 depicts the SCV obtained by the heuristic procedure for large n values, compared with the outputs of the highly accurate BIPOP-CMA-ES and the faster CMA-ES optimization procedures. Figure 4 suggests that the heuristic optimizations remains very close to the minimum also for larger n values and the SCV obtained by the heuristic optimization maintains its polynomial decay between $n^{-2.1}$ and $n^{-2.2}$.

6. REFERENCES

[1] D. Aldous and L. Shepp. The least variable phase type distribution is Erlang. *Stochastic Models*, 3:467–473, 1987.

[2] T. Éltető, S. Rácz, and M. Telek. Minimal coefficient of variation of matrix exponential distributions. In *2nd Madrid Conference on Queueing Theory, Madrid, Spain (July 2006)*, 2006.

[3] N. Hansen. The cma evolution strategy: a comparing review. In *Towards a New Evolutionary Computation*, pages 75–102. Springer, 2006.

[4] N. Hansen. Benchmarking a bi-population cma-es on the bbob-2009 function testbed. In *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers*, pages 2389–2396. ACM, 2009.

[5] I. Horváth, O. Sáfár, M. Telek, and B. Zámbó. Concentrated matrix exponential distributions. In *European Workshop on Performance Engineering*, pages 18–31. Springer, 2016.

[6] I. Horváth, Z. Talyigás, and M. Telek. An optimal inverse Laplace transform method without positive and negative overshoot – an integral based interpretation. *Electronic Notes in Theoretical Computer Science*, 337:87 – 104, 2018.

[7] I. Rechenberg. Evolutionsstrategien. In *Simulationsmethoden in der Medizin und Biologie*, pages 83–114. Springer, 1978.

Numerical Inverse Laplace Transformation by concentrated matrix exponential distributions *

Gábor Horváth
Dept. of Networked Systems and Services
Budapest University of Technology and Economics
Magyar Tudósok Körútja 2, 1117
Budapest, Hungary
ghorvath@hit.bme.hu

Illés Horváth
MTA-BME Information Systems Research Group
Magyar Tudósok Körútja 2, 1117
Budapest, Hungary
pollux@math.bme.hu

Salah Al-Deen Almousa
Dept. of Networked Systems and Services
Budapest University of Technology and Economics
Magyar Tudósok Körútja 2, 1117
Budapest, Hungary
almousa@hit.bme.hu

Miklós Telek
Dept. of Networked Systems and Services
Budapest University of Technology and Economics
Magyar Tudósok Körútja 2, 1117
Budapest, Hungary
telek@hit.bme.hu

ABSTRACT

We present effective numerical inverse Laplace transformation (ILT) method which belongs to the Abate–Whitt framework and exhibits some of the best properties among all the procedures of the framework. E.g., the proposed ILT method does not generate overshoot and undershoot (upward/downward jump exceeding the jump of the original function), numerically stable and gradually improving.

Keywords: numerical inverse Laplace transformation, Abate–Whitt framework, concentrated matrix exponential distribution.

1. INTRODUCTION

There are plenty of numerical inverse Laplace transformation methods published in the literature (for a relatively recent survey we refer to [6]). Among these methods one of the most widely applied and well characterized subset is Abate–Whitt framework defined in [1]. This framework implicitly defines function families in which various optimizations can be performed in order to obtain efficient inverse Laplace transformation methods.

We propose a procedure which is based on the most general function family of the Abate–Whitt framework (referred to as Class III in [1]) where we adopt a restriction that the inverse Laplace transformation should be non-overshooting.

It turns out that matrix exponential (ME) distribution applied in the Abate–Whitt framework ensures non-overshooting inverse Laplace transformation. In [5] low order inverse Laplace transformation is applied, using concentrated ME (CME) distributions, which were available at that time [2, 4]. Recent improvements in the computation of CME distributions [3] allow us to extend the numerical ILT method also to high orders. In this work we present the first numerical experiences about the ILT method based on high order CME distributions.

*This work is supported by the OTKA K-123914 project

MAM10 2019, Hobart, Australia

2. INVERSE LAPLACE TRANSFORMATION AND THE ABATE–WHITT FRAMEWORK

For a real or complex valued function $h(t)$ the Laplace transform is defined as

$$h^*(s) = \int_{t=0}^{\infty} e^{-st} h(t) \mathrm{d}t. \tag{1}$$

and the inverse transform problem is to find an approximate value of h at point T (i.e., $h(T)$) based on $h^*(s)$.

REMARK 1. *We assume that* $\int_{t=0}^{\infty} e^{-st} h(t) \mathrm{d}t$ *is finite for* $\mathrm{Re}(s) > 0$ *thus* $h^*(s)$ *is well-defined by* (1) *for* $\mathrm{Re}(s) > 0$.

REMARK 2. *We assume that* $h(t)$ *is real in this work. As a result,* $h^*(\bar{s}) = \bar{h}^*(s)$ *and* $h^*(\bar{s}) + h^*(s) = 2\mathrm{Re}(h^*(s))$.

Among the wide range of inverse Laplace transformation methods, we restrict our attention to the Abate–Whitt framework which we summarize below.

2.1 The Abate–Whitt framework

The idea is to approximate h by a finite linear combination of the transform values, via

$$h(T) \approx h_n(T) := \sum_{k=1}^{n} \frac{\eta_k}{T} h^* \left(\frac{\beta_k}{T} \right), \; T > 0, \tag{2}$$

where the nodes β_k and weights η_k are complex numbers, which depend on n, but not on the transform h^* or the time argument T. This framework was introduced and investigated by Abate and Whitt in [1]. When $h(t)$ in (1) is real valued it can be approximated by the real part of the weighted transform values:

$$\mathrm{Re}(h(T)) \approx \mathrm{Re}(h_n(T)) = \sum_{k=1}^{n} \mathrm{Re} \left(\frac{\eta_k}{T} h^* \left(\frac{\beta_k}{T} \right) \right).$$

In the special case when there is a complex conjugate pair among the nodes and weights (that is, $\eta_i = \bar{\eta}_j$ and $\beta_i = \bar{\beta}_j$) then

$$\frac{\eta_i}{T} h^* \left(\frac{\beta_i}{T}\right) + \frac{\eta_j}{T} h^* \left(\frac{\beta_j}{T}\right) = 2\mathrm{Re}\left[\frac{\eta_i}{T} h^* \left(\frac{\beta_i}{T}\right)\right].$$

For numerical comparisons, we consider two classic algorithms of the Abate–Whitt framework: the Gaver–Stehfest method and the Euler method, which are investigated also in [1]. These two methods approximate $h(T)$ by $h_n(T)$, where $h_n(T)$ has the form (2) with weights η_k and nodes β_k, $k = 1, 2, \ldots n$ as follows.

Gaver–Stehfest method (for even n)

$$\beta_k = k\ln(2), \quad \text{for } 1 \le k \le n,$$

$$\eta_k = (-1)^{n/2+k} \ln(2) \sum_{j=\lfloor (k+1)/2 \rfloor}^{\min(k,n/2)} \frac{j^{n/2+1}}{(n/2)!} \binom{n/2}{j} \binom{2j}{j} \binom{j}{k-j}, \quad \text{for } 1 \le k \le n,$$

where $\lfloor x \rfloor$ is the greatest integer less than or equal to x.

Euler method (for odd n)

$$\beta_k = \frac{(n-1)\ln(10)}{6} + \pi i(k-1), \quad 1 \le k \le n$$
$$\eta_k = 10^{(n-1)/6}(-1)^{k-1}\xi_k, \quad 1 \le k \le n$$

where

$$\xi_1 = \frac{1}{2}$$
$$\xi_k = 1, \quad 2 \le k \le (n+1)/2$$
$$\xi_n = \frac{1}{2^{(n-1)/2}}$$
$$\xi_{n-k} = \xi_{n-k+1} + 2^{-(n-1)/2} \binom{(n-1)/2}{k} \quad \text{for } 1 \le k < (n-1)/2.$$

REMARK 3. *The set of real valued functions $\sum_k \eta_k e^{-\beta_k t}$ with potentially complex valued coefficients has the following real representations.*

Class I *If both η_k and β_k are real then $\sum_k \eta_k e^{-\beta_k t}$ is a real representation.*

Class II *If η_k is real and β_k is complex then*

$$\mathrm{Re}\left(\sum_k \eta_k e^{-\beta_k t}\right) = \sum_k \eta_k e^{-b_k t}\cos(\omega_k t)$$

is its real representation, where $\beta_k = b_k + i\omega_k$.

Class III *If both η_k and β_k are complex then*

$$\mathrm{Re}\left(\sum_k \eta_k e^{-\beta_k t}\right) = \sum_k a_k e^{-b_k t}\cos(\omega_k t + \phi_k)$$

is its real representation, where $\beta_k = b_k + i\omega_k$ and a_k, ϕ_k are real and obtained from the real and imaginary parts of η_k [4].

The Gaver–Stehfest method falls into Class I, the Euler method falls into Class II, the proposed ME distribution based method (described in detail in Section 3.2) falls into Class III.

For $\mathrm{Re}(\beta_k) > 0$, we can reformulate the inverse Laplace transformation methods of the Abate–Whitt framework as

$$h_n(T) = \frac{1}{T}\sum_{k=1}^{n} \eta_k h^*\left(\frac{\beta_k}{T}\right) = \frac{1}{T}\sum_{k=1}^{n}\eta_k \int_0^\infty e^{-\frac{\beta_k}{T}t} h(t)\mathrm{d}t$$
$$= \int_0^\infty h(t) f_T^n(t)\mathrm{d}t, \tag{3}$$

where the numerical approximation of the Laplace inverse at point T is obtained as the integral of the original function, $h(t)$, with

$$f_T^n(t) = \frac{1}{T}\sum_{k=1}^{n}\eta_k e^{-\frac{\beta_k}{T}t}. \tag{4}$$

If $f_T^n(t)$ was the Dirac impulse function at point T then the Laplace inversion would be perfect, but depending on the order of the approximation (n), the applied inverse transformation method (weights η_k, nodes β_k) and the time point (T), function $f_T^n(t)$ only approximates the Dirac impulse function with a given accuracy.

REMARK 4. *$f_T^n(t)$ is a scaled version of*

$$f_1^n(t) = \sum_{k=1}^{n}\eta_k e^{-\beta_k t} \tag{5}$$

because, according to (4),

$$f_T^n(t) = \frac{1}{T} f_1^n\left(\frac{t}{T}\right). \tag{6}$$

3. MATRIX EXPONENTIAL DISTRIBUTIONS

The class of matrix exponential distributions of order N, denoted ME(N), contains random variables with pdf of the form

$$f_X(t) = -\alpha \mathbf{A} e^{\mathbf{A}t}\mathbf{1}, \qquad t \ge 0, \tag{7}$$

where α is a row vector of length N, $\mathbf{A}$ is a matrix of size $N \times N$ and $\mathbf{1}$ is a column vector of ones of size N. As $f_X(t)$ is a pdf, $f_X(t)$ is non-negative for $t \ge 0$.

Assuming that $\mathbf{A}$ is diagonalizable, with spectral decomposition $\mathbf{A} = \sum_{i=1}^{n} u_i \lambda_i v_i$, the pdf can be written as

$$f_X(t) = \sum_{i=1}^{n} \underbrace{-\alpha \mathbf{A} u_i v_i \mathbf{1}}_{c_i} e^{\lambda_i t} = \sum_{i=1}^{n} c_i e^{\lambda_i t}, \tag{8}$$

where $\lambda_1, \ldots, \lambda_n$ are eigenvalues of $\mathbf{A}$. Comparing (8) and (5) indicates that ME distributions with diagonalizable matrix $\mathbf{A}$ can be used in the place of $f_1^n(t)$.

REMARK 5. *Due to the non-negativity of $f_X(t)$, the integral in* (3) *results in an inverse Laplace transformation without overshoot.*

3.1 Concentrated ME distributions

Concentrated ME(N) distributions with low coefficient of variation has been calculated in [4] up to $N = 47$ and in [3] for up to $N = 2001$, using the following form (for odd N):

$$f_{\mathrm{ME}}(t) = c\,\mathrm{e}^{-\lambda t} \prod_{i=0}^{(N-1)/2} \cos^2(\omega t - \phi_i) \tag{9}$$

with real values of c, λ, ω and $\phi_1, \ldots, \phi_{(N-1)/2}$ obtained from numerical optimization.

3.2 ME distribution-based inverse Laplace transformation

In order to apply the CME distributions for inverse Laplace transformation (9) needs to be rewritten in a form consistent with (5):

$$f_{\mathrm{ME}}(t) = c\,\mathrm{e}^{-\lambda t} \prod_{i=0}^{(N-1)/2} \cos^2(\omega t - \phi_i) = \sum_{i=1}^{N} \eta_i e^{-\beta_i t}$$

where $n = (N+1)/2$, η_1, β_1 are real, and the values $\beta_2, \ldots, \beta_n$ have positive imaginary parts. For the details of this transformation, see the Appendix of [4].

4. NUMERICAL COMPARISON WITH THE ME BASED METHOD

In order to investigate the properties of the considered inverse Laplace transformation methods (Euler, Gaver–Stehfest (Gaver in short), Concentrated ME based (CME)), we performed a set of numerical inverse Laplace transformations for the 6 functions of Table 1 using our Matlab implementation, where we applied standard double precision floating point The arithmetic for CME and 100 digit precision arithmetic with the Matlab Symbolic Math Toolbox for Euler and Gaver. Numerical properties of the 6 test functions are rather similar; we demonstrate them using mainly the test function $\lfloor t \rfloor \mod 2$.

$h(t)$	e^{-t}	$\sin t$	$\mathbb{1}(t>1)$
$h^*(s)$	$\frac{1}{1+s}$	$\frac{1}{s^2+1}$	$\frac{1}{s}e^{-s}$
$h(t)$	$\mathbb{1}(t>1)e^{1-t}$	$\lfloor t \rfloor$	$\lfloor t \rfloor \mod 2$
$h^*(s)$	$\frac{e^{-s}}{1+s}$	$\frac{1}{s}\frac{1}{e^s-1}$	$\frac{1}{s}\frac{1}{e^s+1}$

Table 1: Set of test functions

Figures 1 and 2 investigate the dependency of the Gaver and the Euler methods on the order. The Gaver method fails to follow the alternating feature of the original function for low order ($n = 10$). It produces a smooth curve with some overshoot for medium order ($n = 50$), and reaches its limit of numerical stability, despite using 100 digit precision, at $n = 64$. The Euler method for low order ($n = 11$) follows the alternating feature of the original function for longer; it produces a smooth curve with more dominant overshoot for medium order ($n = 51$), and reaches its limit of numerical stability, despite using 100 digit precision, at $n = 101$.

In comparison, the inverse Laplace transformation obtained by the CME method is depicted in Figure 3. The CME method does not produce overshoot at any order. Similar to the Euler method, the CME method follows the alternating feature of the original function for low order ($n = 10$). It produces a smooth curve for low, medium ($n = 50$) and high ($n = 500$) orders using double precision arithmetic. The accuracy of the inverse Laplace transformation continuously increases with the order.

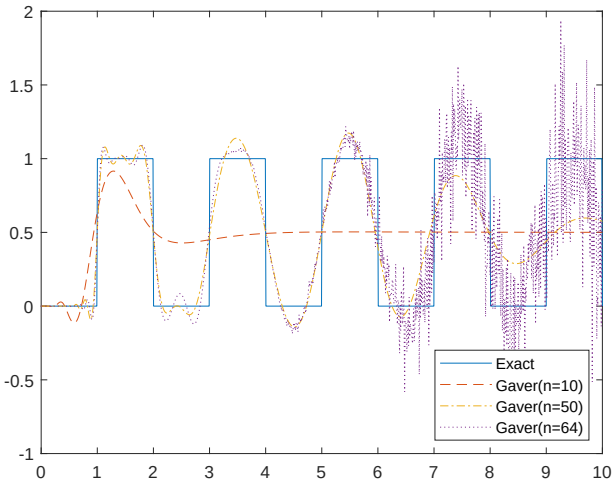


Figure 1: $h(t) = \lfloor t \rfloor \mod 2$ with Gaver method

Figures 4 and 5 compare the methods for $h(t) = \lfloor t \rfloor$ with low and medium orders, while Figures 6 and 7 compare the methods for $h(t) = \lfloor t \rfloor \mod 2$ with the same orders. In each case, the benefit of the non-overshooting inverse Laplace transformation is dominant. Especially, the figures with medium orders indicate the uncertainties coming from overshooting inverse Laplace transformation using the more alternating Euler method.

The sharpest increase of the Euler and the CME methods are similar for the same orders. Approximating discontinuity, the Euler method provides a bit sharper increase/decrease than the CME method at a given order, but at the cost of significant overshoots before and after the discontinuity.

5. REFERENCES

[1] J. Abate and W. Whitt. A unified framework for numerically inverting laplace transforms. *INFORMS Journal on Computing*, 18(4):408–421, Fall 2006.

[2] T. Éltető, S. Rácz, and M. Telek. Minimal coefficient of variation of matrix exponential distributions. In *2nd Madrid Conference on Queueing Theory, Madrid, Spain (July 2006)*, 2006.

[3] G. Horváth, I. Horváth, and M. Telek. High order low variance matrix-exponential distributions. *Submitted to The Tenth International Conference on Matrix-Analytic Methods in Stochastic Models (MAM10)*, 2019.

[4] I. Horváth, O. Sáfár, M. Telek, and B. Zámbó. *Concentrated Matrix Exponential Distributions*, pages 18–31. Springer International Publishing, Cham, 2016.

[5] I. Horváth, Z. Talyigás, and M. Telek. An optimal inverse Laplace transform method without positive and negative overshoot – an integral based interpretation. *Electronic Notes in Theoretical Computer Science*, 337:87 – 104, 2018.

[6] Q. Wang and H. Zhan. On different numerical inverse laplace methods for solute transport problems. *Advances in Water Resources*, 75:80 – 92, 2015.

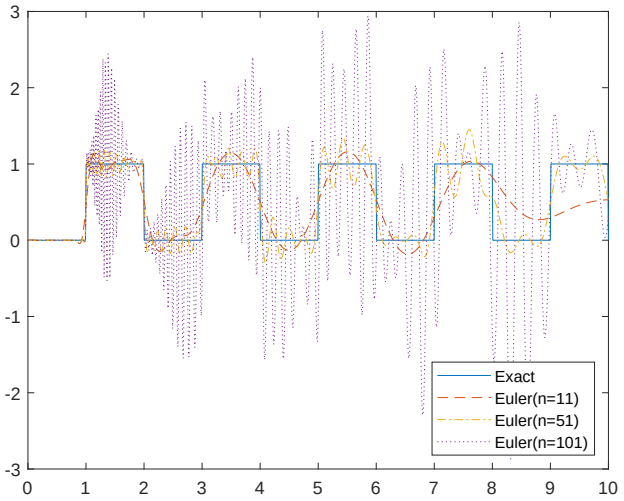


Figure 2: $h(t) = \lfloor t \rfloor \mod 2$ **with Euler method**

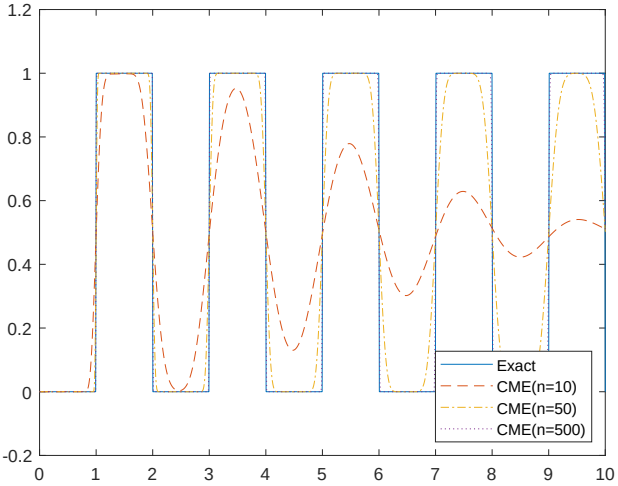


Figure 3: $h(t) = \lfloor t \rfloor \mod 2$ **with CME method**

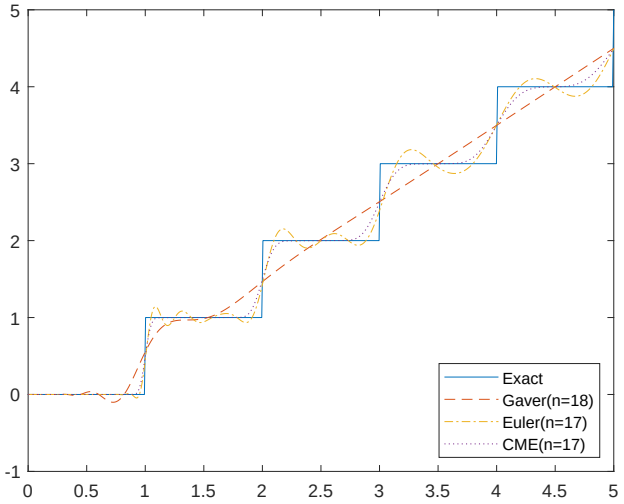


Figure 4: $h(t) = \lfloor t \rfloor$ **with low orders**

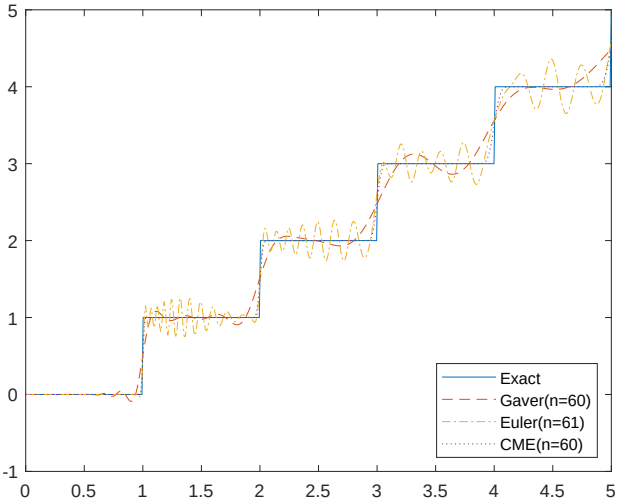


Figure 5: $h(t) = \lfloor t \rfloor$ **with medium orders**

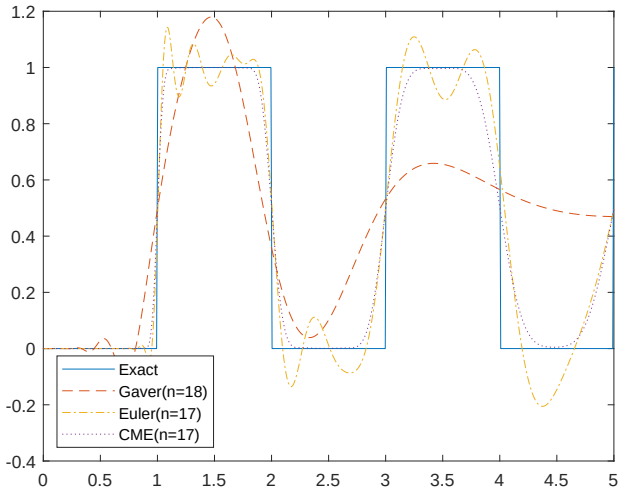


Figure 6: $h(t) = \lfloor t \rfloor \mod 2$ **with low orders**

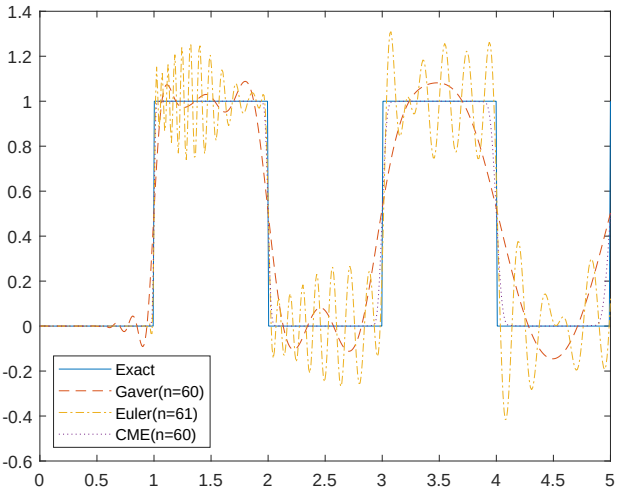


Figure 7: $h(t) = \lfloor t \rfloor \mod 2$ **with medium orders**

One-sided Markov additive processes: the three fundamental matrices and the scale function

Keynote speaker

Jevgenijs Ivanovs
Aarhus University, Denmark

ABSTRACT

Traditional matrix-analytic models can be viewed as Markov additive processes with level component living on a lattice and some special behaviour at the boundary. The basic assumption that the level process is skip-free in one direction allows for an in-depth analysis of such models. It turns out that a somewhat similar analysis is possible in the continuous-time non-lattice case when assuming that the level process is jump-free in one direction. Such model corresponds to a Lévy process modulated by a Markov chain with additional jumps at phase switching epochs, where all jumps are of the same sign.

This tutorial will focus on some basic theory for one-sided Markov additive processes highlighting the links to the traditional matrix-analytic models [3]. In particular, we will consider the three fundamental matrices: the right-, the left-solution of a certain matrix equation, and the linking matrix of expected occupation times at zero. Moreover, we will construct the so-called scale (matrix-valued) function [4], and discuss some applications of the theory to insurance risk [1]. If time allows we will also look at the underlying spectral theory based on the so-called generalized Jordan chains corresponding to analytic matrix functions [2].

Prior knowledge of Lévy processes (Lévy-Khintchine formula) will not be assumed, but will render this tutorial easier to follow.

1. REFERENCES

[1] H. Albrecher and J. Ivanovs (2013). A risk model with an observer in a Markov environment. *Risks* 1(3), p. 148–161.

[2] B. D'Auria, J. Ivanovs, O. Kella, and M. Mandjes (2010). First passage of a Markov additive process and generalized Jordan chains. *Journal of Applied Probability* 47(4), p. 1048–1057.

[3] J. Ivanovs, G. Latouche and P.G. Taylor (2018). One-sided Markov additive processes with lattice and non-lattice increments. In preparation.

[4] J. Ivanovs and Z. Palmowski (2012). Occupation densities in solving exit problems for Markov additive processes and their reflections. *Stochastic Processes and their Applications* 122(9), p. 3342–3360.

MAM10 2019, Hobart, Australia

Modelling intensive care units using quasi-birth-death processes

Sarah James
University of Adelaide
North Terrace, Adelaide, 5005
South Australia, Australia
sarah.james@adelaide.edu.au

Nigel Bean
University of Adelaide
North Terrace, Adelaide, 5005
South Australia, Australia
nigel.bean@adelaide.edu.au

Jonathan Tuke
University of Adelaide
North Terrace, Adelaide, 5005
South Australia, Australia
simon.tuke@adelaide.edu.au

ABSTRACT

An intensive care unit (ICU) is a crucial and limited resource in a hospital which is affected by uncertainty and variability. ICUs provide high level care to critically ill patients. This includes burns, cardiothoracic, general medicine and surgical, spinal, and trauma cases, as well as patients who require life support. ICUs have the highest nurse to patient ratios in the hospital and often operate close to capacity. Insufficient resources in the ICU causes many negative effects both within the ICU and in other connected departments [1].

Queueing theory has been used to model the bed occupancy in ICUs for the last 20 years, with a particular focus on using $M/M/\cdot$ and $M/PH/\cdot$ queueing models [1]. These queueing models assume that the arrival process (patient arrivals) is a Poisson process, the service times (length of stay) follow an exponential or Phase-Type distribution, and most crucially that the arrival process and service times are independent of each other. These assumptions are ideal for researchers as many problems become tractable and the analysis of such queueing models is relatively straightforward. However, several studies have revealed that patients may be refused entry into the ICU due to capacity issues [2], and that there may be some dependence structure between the arrival process and the service times in an ICU [3]. Without independence between the arrival process and service times, all standard queueing models become invalid.

Our research analyses the patient flow of the ICU at the Royal Adelaide Hospital, Adelaide, South Australia, Australia. The Royal Adelaide Hospital (RAH) is the largest hospital in Adelaide, South Australia. Prior to November 2017, the RAH was located on North Terrace, Adelaide, containing 680 beds and an ICU capacity of 42 beds. The dataset used in this research was obtained from the ICU at the RAH and contains de-identified information on 7124 patients who arrived to the ICU from 1 October 2014 to 30 November 2016. This dataset was also used in the analysis to demonstrate the dependence structure between patient arrivals and patient length of stay [3].

Varney *et al.* [3] provided an engineering-style method to provide a reasonable model of bed occupancy in the ICU at the RAH. In this talk, we aim to provide a more principled approach to modelling bed occupancy in ICUs using quasi-birth-and-death processes (QBDs). Although $PH/PH/\cdot$ queues are examples of QBDs, they maintain the independence between the arrival process and the service times. By allowing the phases of the arrival process and the service times to interact with each other, QBDs provide the flexibility to model a queueing system with dependence between the arrival process and the service times.

However, little research has been conducted in fitting non-specified queueing models to data from queueing systems. In particular, no theory has been developed to statistically fit arbitrary QBD models to bed occupancy data. In this talk, we will describe the approaches we have taken to fit suitable QBDs to the RAH ICU data.

1. REFERENCES

[1] J. Bai, A. Fügener, J. Schoenfelder, and J. O. Brunner. Operations research in intensive care unit management: a literature review. *Health care management science*, 21(1):1–24, 2018.

[2] M. A. Metcalfe, A. Sloggett, and K. McPherson. Mortality among appropriately referred patients refused admission to intensive-care units. *The Lancet*, 350(9070):7–11, 1997.

[3] J. Varney, N. Bean, and M. Mackay. The self-regulating nature of occupancy in ICUs: stochastic homoeostasis. *Health care management science*, pages 1–20, 2018.

MAM10 2019, Hobart, Australia

An Exact Analysis of a Class of Markovian Bitcoin Models

Kayla Javier
School of Mathematical and Statistical Sciences
Clemson University
Martin Hall, Box 340975
Clemson, South Carolina, USA, 29634
kjavier@clemson.edu

Brian Fralix
School of Mathematical and Statistical Sciences
Clemson University
Martin Hall, Box 340975
Clemson, South Carolina, USA, 29634
bfralix@clemson.edu

ABSTRACT

Recently, in the work of Göbel et al [2], two continuous-time Markov chain models were introduced in order to model interactions between a small pool of miners, and a larger collection of miners, within the Bitcoin blockchain. The first model—Model 1—represents the case where the pool behaves in an honest manner, while the second model—Model 2—attempts to model what can happen when the pool behaves dishonestly, and follows a selfish-mining strategy.

Our first result is a new derivation of the stationary distribution (up to a normalizing constant) of Model 1, and we further build on the results of [2] by showing that the normalizing constant can be expressed in closed-form. Furthermore, the Laplace transforms of the transition functions are shown to be just as tractable, if we assume that the larger collection of miners has the same information about the blockchain as the smaller pool at time zero.

Our next set of results show that the stationary distribution of Model 2—when it exists—can be calculated exactly as well, without having to truncate the state space. A similar statement can be made about the Laplace transforms of the transition functions of Model 2 as well, if we further assume that the larger collection of miners has the same information about the blockchain as the smaller pool at time zero.

The methods we use to study both models, particularly Model 2, make use of ideas that often appear in the theory of matrix-analytic methods, except our analysis makes use of the recently-developed random-product technique introduced in [1].

1. REFERENCES

[1] Buckingham, P., and Fralix, B. (2015) Some new insights into Kolmogorov's criterion, with applications to hysteretic queues. *Markov Processes and Related Fields* **21**, 339-368.

[2] Göbel, J., Keeler, H.P., Krzesinski, A.E., and Taylor, P.G. (2016). Bitcoin blockchain dynamics: the selfish-mine strategy in the presence of propagation delay. *Performance Evaluation* **104**, 23-41.

MAM10 2019, Hobart, Australia

Nearly-completely decomposable Markov modulated fluid queues

Sarah Dendievel
Ghent University
SMACS Research Group
Sint-Pietersnieuwstraat 41
9000 Gent, Belgium
Sarah.Dendievel@UGent.be

Guy Latouche
Université libre de Bruxelles
Faculté des sciences, CP212
Boulevard du Triomphe 2
1050 Bruxelles, Belgium
latouche@ulb.ac.be

Yuanyuan Liu
Central South University
School of Mathematics and Statistics
Changsha, Hunan, 410083, China
liuyy@csu.edu.cn

Yingchun Tang
Southern University of Science and Technology
Department of Mathematics
Shenzhen, Guangdong, 518055, China
1060864320@qq.com

ABSTRACT

We consider a Markov-modulated fluid queue for which the environment is nearly-completely decomposable: the set of phases is partitioned into subsets in such a way that transition rates between different subsets are order of magnitudes smaller than transition rates between phases in the same subset. Technically, the generator is of the form

$$Q(\varepsilon) = Q^* + \varepsilon G, \tag{1}$$

where ε is a is a small number and Q^* is a block-diagonal matrix, so that the phase process is split into separate processes when $\varepsilon = 0$. The matrix G is such that $Q(\varepsilon)$ is an irreducible generator when $\varepsilon > 0$.

Each subset may be thought of as driving its own fluid queue and the whole process behaves sometimes like one fluid queue, sometimes like another. This is a model for a system with very slowly varying parameters, and we are interested in analysing the stationary distribution of the whole system.

If each of the isolated fluid queues is positive recurrent, then the whole system is positive recurrent as well, and we show that its stationary distribution is approximated, for $\varepsilon \to 0$, by a properly weighted mixture of the stationary distributions of the isolated fluid queues. A similar result holds for finite, discret-time Markov chains, as shown in Simon and Ando [5], Courtois [1] and Schweitzer [4] among others. In the recent paper Jiang *et al.* [2], the authors analyse the case of a countable state space and give a recent bibliography.

One may imagine that the whole system is positive recurrent when $\varepsilon > 0$, while some of the isolated fluid queues are transient on their own. In that case, the result mentioned above may not apply, since such transient isolated fluid queues do not have a stationary distribution. Instead, we show that some of the probability mass disappears from sight as ε tends to 0. To make the presentation sufficiently simple, we assume that there are only two subsets of phases, corresponding respectively to a positive recurrent and to a transient isolated fluid queues.

Latouche and Schweitzer [3] bears some similarity with the fluid queue analysed here. Our present results, however, are more detailed and more general. In particular, we show that the usual Riccati equation for the matrix Ψ has three (sub)stochastic solutions for $\varepsilon = 0$, only one of which is physically relevant as a limit for $\varepsilon \to 0$.

1. REFERENCES

[1] P. Courtois. *Decomposability. Queueing and Computer Systems Applications.* Academic Press, New York, 1977.

[2] S. Jiang, Y. Liu, and Y. Tang. A unified perturbation analysis framework for countable Markov chains. *Linear Algebra and its Applications*, 529:413–440, 2017.

[3] G. Latouche and P. J. Schweitzer. A Markov modulated, nearly completely decomposable M/M/1 queue. In W. J. Stewart, editor, *Computations with Markov chains*, pages 39–48. Springer, 1995.

[4] P. J. Schweitzer. Perturbation series expansions of nearly completely-decomposable Markov chains. In O. Boxma, J. Cohen, and H. Tijms, editors, *Teletraffic Analysis and Computer Performance Evaluation*, pages 319–328, Amsterdam, 1986. Elsevier, North–Holland.

[5] H. A. Simon and A. Ando. Aggregation of variables in dynamic systems. *Econometrica*, 29:111–138, 1961.

MAM10 2019, Hobart, Australia

Eulerian Numbers and an Explicit Formula for a Random Walk Generating Function

Barbara Margolius
Cleveland State University
Cleveland, Ohio 44115-2214, United States
b.margolius@csuohio.edu

ABSTRACT

Consider a level independent quasi-birth-death process (QBD) [3][1] with time-varying periodic rates, and N phases in each level. The infinitesimal generator for such a system is given by

$$\mathbf{Q}(t) = \begin{bmatrix} \mathbf{A}_{00}(t) & \mathbf{A}_1(t) & \cdot & \cdot \\ \mathbf{A}_{-1}(t) & \mathbf{A}_0(t) & \mathbf{A}_1(t) & \cdot \\ \cdot & \ddots & \ddots & \ddots \end{bmatrix}$$

where the blocks $\mathbf{A}_i(t)$ are $N \times N$ matrices whose components are periodic functions representing transition rates. When the system is ergodic, there is an asymptotic periodic solution to the system of differential equations

$$\frac{\partial}{\partial t}\mathbf{p}_n(t) = \mathbf{p}_{n-1}(t)\mathbf{A}_1(t) + \mathbf{p}_n(t)\mathbf{A}_0(t) + \mathbf{p}_{n+1}(t)\mathbf{A}_{-1}(t)$$

for $n > 0$ and boundary condition

$$\frac{\partial}{\partial t}\mathbf{p}_0(t) = \mathbf{p}_0(t)\mathbf{A}_{00}(t) + \mathbf{p}_1(t)\mathbf{A}_{-1}(t)$$

The vectors $\mathbf{p}_n(t)$ are of length N. The generating function for the asymptotic periodic probabilities is given by

$$P_z(t) = \int_{t-1}^{t} \mathbf{p}_0(u)\left(\mathbf{A}_{00}(u) - \mathbf{A}_0(u) - z^{-1}\mathbf{A}_{-1}(u)\right)\Phi_z(u,t)du \times \left(\mathbf{I} - \Phi_z(t-1,t)\right)^{-1}$$

where $\Phi_z(s,t)$ is the generating function for the corresponding unbounded process, that is, the random walk process[2]. For this generating function, the coefficients on z^n are matrices. The (i,j)th component of the coefficient on z^n is the probability of a transition from phase i to phase j and up n levels.

$\Phi_z(s,t)$ solves the differential equation

$$\frac{\partial}{\partial t}\Phi_z(s,t) = \Phi_z(s,t)\mathbf{A}(z,t)$$

where

$$\mathbf{A}(z,t) = z\mathbf{A}_1(t) + \mathbf{A}_0(t) + z^{-1}\mathbf{A}_{-1}(t).$$

If we were working with a constant rate process, $\Phi_z(s,t)$ would be the matrix exponential $\mathrm{e}^{\mathbf{A}(z)(t-s)}$.

In general, an explicit formula for $\Phi_z(s,t)$ is not known. In this paper, we consider several cases where such a formula is available and explore the combinatorial interpretation of the resulting expressions. In particular, we study a single-server pre-emptive priority queue in which the Eulerian numbers appear[4] and a QBD with Erlang arrivals in which roots of unity play a role[5].

Pre-emptive Priority Queue with Finite Buffer

In the case of the pre-emptive priority queue with finite buffer space for class-2 customers, out state space is given by $(X(t), J(t))$ where $X(t)$ gives the number of class-1 customers and $J(t)$ gives the number of class-2 customers at time t. Associated with this queue-length process, is a two-dimensional random walk where the level, $X(t)$ is unbounded, and the phase, $J(t)$ ranges from zero to K where K is the size of the buffer. Let $\lambda_i(t)$ give the arrival rate of class i-customers and $\mu_i(t)$ give the service rate of class-i customers. Since class-2 customers are served only when there are no class-1 customers in the system, $\mu_2(t)$ is not part of the expression for the random walk generating function.

Let $a_k(s,t)$ give the probability of k class-2 arrivals during the time-interval $(s,t]$, so that

$$a_k(s,t) = \frac{\left(\int_s^t \lambda_2(u)du\right)^k}{k!}\mathrm{e}^{-\int_s^t \lambda_2(u)du}.$$

Define the random walk generating function for a particle moving to the right at rate $\lambda_1(t)$ and to the left at rate $\mu_1(t)$ as

$$\beta_z(s,t) = \exp\left\{\int_s^t \lambda_1(u)du(z-1) + \int_s^t \mu_1(u)du(z^{-1}-1)\right\} = \sum_{n=-\infty}^{\infty} \Pr\{X(t) = n+k | X(s) = k\}z^n.$$

MAM10 2019, Hobart, Australia

Then $\Phi_z(s,t)$ is given by

$$\Phi_z = \beta_z \begin{bmatrix} a_0 & a_1 & a_2 & \dots & \dots & a_{N-1} & a_{>N-1} \\ 0 & a_0 & a_1 & \ddots & \dots & a_{N-2} & a_{>N-2} \\ 0 & 0 & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & a_0 & \dots & a_{N-i} & a_{>N-i} \\ 0 & 0 & 0 & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & a_0 & a_{>0} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

where the dependence on s and t is suppressed in the notation.

We are also interested in an explicit expression for $(I - \Phi_z(t-1,t))^{-1}$. This too is available for the pre-emptive priority queue with finite buffer. We consider transition rates which are periodic with period 1. Note that for such transition rates, the integral of the rate from $t-1$ to t is equal to the average value of the rate, so, for example, $\int_{t-1}^{t} \lambda_1(u)du = \bar{\lambda}_1$. We may express $(I - \Phi_z(t-1,t))^{-1}$ in terms of these average rates.

Remark: Note that it is not true for general quasi-birth-death processes with time-varying periodic rates that $(I - \Phi_z(t-1,t))^{-1}$ does not depend on t; however, it is true that the determinant of $\Phi_z(t-1,t)$ does not depend on t.

Recall that

$$(I-\Phi_z(t-1,t))^{-1} = \sum_{n=0}^{\infty} \Phi_z^n(t-1,t) = \sum_{n=0}^{\infty} \Phi_z(t-1,t+n-1).$$

For the Poisson probabilities giving the probability of k class-2 arrivals during an interval of length n, we have

$$a_k(t-1,t+n-1) = \frac{\left(\int_{t-1}^{t+n-1} \lambda_2(u)du\right)^k}{k!} \mathrm{e}^{-\int_{t-1}^{t+n-1} \lambda_2(u)du} = \frac{\bar{\lambda}_2^k n^k}{k!} \mathrm{e}^{-\bar{\lambda}_2 n}.$$

The $(j, j+k)$ component of the matrix generating function $(I - \Phi_z)^{-1}$ is given by

$$[(I - \Phi_z)^{-1}]_{j,j+k} = \sum_{n=1}^{\infty} \frac{\bar{\lambda}_2^k n^k}{k!} \mathrm{e}^{-\bar{\lambda}_2 n} \beta_z^n(0,1) = \frac{\bar{\lambda}_2^k}{k!} \sum_{n=1}^{\infty} n^k \mathrm{e}^{-\bar{\lambda}_2 n} \beta_z^n(0,1).$$

The Carlitz identity for the Eulerian polynomials[4] is

$$\sum_{n=0}^{\infty} (n+1)^k t^n = \frac{S_k(t)}{(1-t)^{k+1}}$$

where $S_k(t)$ is the kth Eulerian polynomial. The kth Eulerian polynomial provides a generating function for the number of permutations of length k with a given number of descents. Applying this identity, we have

$$[(I - \Phi_z)^{-1}]_{j,j+k} = \mathrm{e}^{-\bar{\lambda}_2} \frac{\bar{\lambda}_2^k}{k!} \frac{\beta_z(0,1) S_k(\beta_z(0,1)\mathrm{e}^{-\bar{\lambda}_2})}{(1 - \mathrm{e}^{-\bar{\lambda}_2}\beta_z(0,1))^{k+1}}.$$

We can use this result to obtain asymptotic estimates for the probability distribution at time t within the period for the pre-emptive priority queue with finite buffer. For an ergodic QBD with time-varying periodic rates, we will obtain periodic estimates.

We will also provide a combinatorial interpretation of the 2-dimensional generating function given by $(I - \Phi_z)^{-1}$ in terms of a sequence of random walks.

$K-$Erlang arrivals and exponential departures

For a random walk with $K-$Erlang arrivals and exponential departures, the matrices $\mathbf{A}_1(t)$, $\mathbf{A}_0(t)$ and $\mathbf{A}_{-1}(t)$ are given by

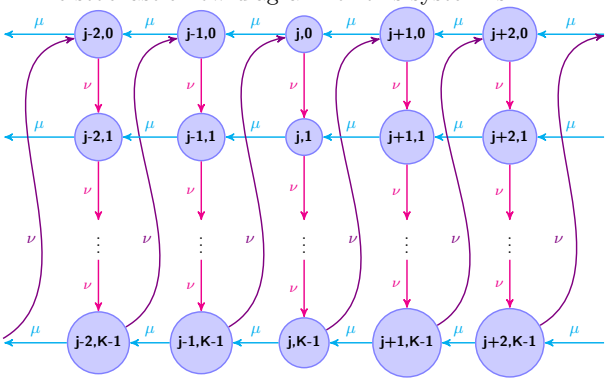
$$\mathbf{A}_1(t) = \begin{bmatrix} \cdot & \cdots & \cdot & \nu(t) \\ \cdot & \cdots & \cdots & \cdot \\ \vdots & \vdots & \vdots & \vdots \\ \cdot & \cdots & \cdots & \cdot \end{bmatrix}$$

$$\mathbf{A}_0(t) = \begin{bmatrix} -\nu(t) - \mu(t) & \cdot & \cdot & \cdot \\ \nu(t) & -\nu(t) - \mu(t) & \cdot & \cdot \\ \cdot & \ddots & \ddots & \cdot \\ \cdot & \cdot & \nu(t) & -\nu(t) - \mu(t) \end{bmatrix}$$

and

$$\mathbf{A}_{-1}(t) = \begin{bmatrix} \mu(t) & \cdot & \cdot \\ \cdot & \ddots & \cdot \\ \cdot & \cdot & \mu(t) \end{bmatrix}.$$

The stochastic flow diagram for this system is



Define

$$\mathbf{A}_z(t) = \mathbf{A}_1(t)z + \mathbf{A}_0(t) + \mathbf{A}_{-1}(t)z^{-1}.$$

Then the eigenvalues of $\mathbf{A}_z(t)$ are given by

$$\epsilon_\ell(t) = \omega_K^{-\ell} \nu(t) z^{1/K} - \mu(t) - \nu(t) + \mu(t) z^{-1}$$

for $\ell = 0, 1, \dots, K-1$ with

$$\omega_K = \mathrm{e}^{-\frac{2\pi i}{K}} = \cos\left(\frac{2\pi}{K}\right) - i \sin\left(\frac{2\pi}{K}\right).$$

The diagonalization of the matrix $\mathbf{A}_z(t)$ is

$$\mathbf{A}_z(t) = HD(t)H^{-1}.$$

The eigenvectors do not depend on t. This makes the matrix function $\Phi_z(s,t)$ particularly easy to compute. It is given in

terms of the exponential of the integrals of the diagonals so we have

$$\Phi_z(s,t) = H \begin{bmatrix} \mathrm{e}^{\int_s^t \epsilon_0(u)du} & 0 & \cdots & 0 \\ 0 & \mathrm{e}^{\int_s^t \epsilon_1(u)du} & 0 & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mathrm{e}^{\int_s^t \epsilon_{K-1}(u)du} \end{bmatrix} H^{-1}.$$

Note that $\Phi_z(s,r)\Phi_z(r,t) = \Phi_z(s,t)$, that is

$$\Phi_z(s,r)\Phi_z(r,t) = H \begin{bmatrix} \mathrm{e}^{\int_s^r \epsilon_0(u)du} & 0 & \cdots & 0 \\ 0 & \mathrm{e}^{\int_s^r \epsilon_1(u)du} & 0 & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mathrm{e}^{\int_s^r \epsilon_{K-1}(u)du} \end{bmatrix} \times H^{-1} H \times \begin{bmatrix} \mathrm{e}^{\int_r^t \epsilon_0(u)du} & 0 & \cdots & 0 \\ 0 & \mathrm{e}^{\int_r^t \epsilon_1(u)du} & 0 & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \mathrm{e}^{\int_r^t \epsilon_{K-1}(u)du} \end{bmatrix} H^{-1} = \Phi_z(s,t)$$

The normalized eigenvector corresponding to the ℓth eigenvalue, $\epsilon_\ell(t) = \omega_K^{-\ell}\nu(t)z^{1/K} - \mu(t) - \nu(t) + \mu(t)z^{-1}$, is

$$v_\ell = \frac{1}{\sqrt{K}} \begin{bmatrix} \omega_K^{0\ell} z^{\frac{K-1}{K}} \\ \omega_K^{1\ell} z^{\frac{K-2}{K}} \\ \vdots \\ \omega_K^{(K-2)\ell} z^{\frac{1}{K}} \\ \omega_K^{(K-1)\ell} z^0 \end{bmatrix}.$$

Define $\mathbf{B}$ as a diagonal matrix with diagonal:

$$\begin{bmatrix} z^{\frac{K-1}{K}} & z^{\frac{K-2}{K}} & \cdots & z^0 \end{bmatrix}.$$

Let H be the matrix whose columns are the normalized eigenvectors. We can write H in terms of a matrix of roots of unity and the diagonal matrix B. Let

$$\Omega = \begin{bmatrix} \omega_K^0 & \omega_K^0 & \cdots & \omega_K^0 \\ \omega_K^0 & \omega_K^1 & \cdots & \omega_K^{K-1} \\ \vdots & \ddots & \ddots & \vdots \\ \omega_K^0 & \omega_K^{K-1} & \cdots & \omega_K^{(K-1)^2} \end{bmatrix},$$

then

$$H = \frac{1}{\sqrt{K}} B\Omega,$$

and H^{-1} is then given by

$$H^{-1} = \frac{1}{\sqrt{K}} \overline{\Omega} B^{-1}$$

where $\overline{\Omega}$ is the complex conjugate of the matrix Ω.

For general K, an explicit formula for the (m,j) component of $\Phi_z(s,t)$ is

$$[\Phi_z(s,t)]_{m,j} = \frac{z^{\frac{j-m}{K}} \mathrm{e}^{\int_s^t \left(-\mu(u)-\nu(u)+\mu(u)z^{-1}\right)du}}{K} \sum_{\ell=0}^{K-1} \mathrm{e}^{\int_s^t \omega_K^\ell \nu(u) z^{1/K} du} \omega_K^{\ell(j-m)}$$

$$= \mathrm{e}^{-\int_s^t (\nu(u)+\mu(u)-\mu(u)z^{-1})du} \sum_{n=1}^{\infty} \frac{\left(\int_s^t \nu(u)du\right)^{Kn-j+m}}{(Kn-j+m)!} z^n$$

The (m,j) component depends on $j-m$, the distance from the diagonal and not on j and m separately. $\Phi_z(s,t)$ is a Toeplitz matrix.

The structure of the $\Phi_z(t-1,t)$ matrix makes it particularly simple to compute the matrix $(I - \Phi_z(t-1,t))^{-1}$:

$$(I - \Phi_z(t-1,t))^{-1} = H\mathrm{diag}\left[\frac{1}{1-\mathrm{e}^{\epsilon_i}}\right] H^{-1}.$$

This result readily yields formulas suitable for asymptotic analysis, that is, asymptotic in the level of the process.

Summary

Explicit formulas are seldom available for the random walk generating functions associated with level independent QBDs with time-varying periodic rates. In the special circumstances where they are available, they can be exploited to better understand the combinatorial properties of the process, and to facilitate asymptotic (in the level) analysis of the distribution. It is hoped that insights gained from processes for which explicit formulas are known may be extended to more general QBDs with time-varying periodic rates.

1. REFERENCES

[1] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modelling, 1st edition. Chapter 2: PH Distributions.* ASA SIAM, 1999.

[2] B. H. Margolius. The matrices R and G of matrix analytic methods and the time-inhomogeneous periodic quasi-birth-and-death process. *Queueing Systems*, 60(1-2):131–151, 2008.

[3] M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models: an Algorithmic Approach, Chapter 2: Probability Distributions of Phase Type.* Dover Publications Inc., 1981.

[4] T. K. Petersen. *Eulerian Numbers.* Birkhauser, 2015.

[5] L. Takacs. Transient behavior of queueing processes with erlang input. *Transactions of the American Mathematical Society*, 100(1):1–28, 1961.

Asymptotic periodic analysis of cyclic stochastic fluid flows with time-varying transition rates

Barbara Margolius
Cleveland State University
Cleveland, Ohio 44115-2214, United States
b.margolius@csuohio.edu

Małgorzata M. O'Reilly *†
University of Tasmania
TAS 7001, Australia
malgorzata.oreilly@utas.edu.au

ABSTRACT

Consider a cyclic stochastic fluid model $\{(\widehat{X}(t), J(t)) : t \geq 0\}$ with level variable $\widehat{X}(t) \geq 0$ and phase variable $J(t) \in \mathcal{S}$, driven by a continuous-time Markov chain $\{J(t) : t \geq 0\}$ with a time-varying generator $T(t) = [\mathcal{T}(t)_{ij}]_{i,j\in\mathcal{S}}$ and cycle of length 1 such that $T(t) = T(t+1)$ for all $t \geq 0$.

Assume real-valued rates $c_{J(t)} \in \mathbb{R}$ such that when $\widehat{X}(t) > 0$ then the level $\widehat{X}(t)$ changes at rate $c_{J(t)}$, and when $\widehat{X}(t) = 0$ then the level $\widehat{X}(t)$ changes at rate $\max\{0, c_{J(t)}\}$.

Let $\mathcal{S}_0 = \{i \in \mathcal{S} : c_i = 0\}$, $\mathcal{S}_+ = \{i \in \mathcal{S} : c_i > 0\}$, $\mathcal{S}_- = \{i \in \mathcal{S} : c_i < 0\}$, and partition the generator $T(t)$ according to $\mathcal{S} = \mathcal{S}_0 \cup \mathcal{S}_+ \cup \mathcal{S}_-$, with

$$T(t) = \begin{bmatrix} T_{00}(t) & T_{0+}(t) & T_{0-}(t) \\ T_{+0}(t) & T_{++}(t) & T_{+-}(t) \\ T_{-0}(t) & T_{-+}(t) & T_{--}(t) \end{bmatrix}.$$

Consider the asymptotic periodic distribution of the model corresponding to the observations at time points $t \in [0, 1)$ within the cycle, which consists of density $\pi(t; x)$, $t \in [0, 1)$, $x > 0$, partitioned according to $\mathcal{S} = \mathcal{S}_+ \cup \mathcal{S}_- \cup \mathcal{S}_0$,

$$\pi(t; x) = \begin{bmatrix} \pi(t; x)_+ & \pi(t; x)_- & \pi(t; x)_0 \end{bmatrix}, \qquad (1)$$

such that for all $t \in [0, 1)$ and $i \in \mathcal{S}$,

$$\pi(t; x)_i = \frac{\partial}{\partial x} \sum_{n=0}^{\infty} P(X(t+n) \leq x, J(t+n) = i), \qquad (2)$$

and mass at zero, also partitioned according to $\mathcal{S} = \mathcal{S}_+ \cup \mathcal{S}_- \cup \mathcal{S}_0$,

$$p(t) = \begin{bmatrix} \mathbf{0} & p_-(t) & p_0(t) \end{bmatrix}, \qquad (3)$$

such that for all $t \in [0, 1)$ and $i \in \mathcal{S}_- \cup \mathcal{S}_0$,

$$p(t)_i = \sum_{n=0}^{\infty} P(X(t+n) = 0, J(t+n) = i). \qquad (4)$$

The aim of this research is to derive theoretical expressions and algorithms for the asymptotic periodic distribution of the model. This work is an extension of the results in Margolius and O'Reilly [1], where the model was introduced.

1. REFERENCES

[1] B. H. Margolius and M. M. O'Reilly. The analysis of cyclic stochastic fluid flows with time-varying transition rates. *Queueing Systems*, 82(1-2):43–73, 2016.

*We would like to thank the Australian Research Council for funding this research through Linkage Project LP140100152.
†ARC Centre of Excellence for Mathematical and Statistical Frontiers.

MAM10 2019, Hobart, Australia

A new numerical method for computing the quasi-stationary distribution of subcritical Galton-Watson branching processes

Sophie Hautphenne
University of Melbourne
Parkville Victoria 3010
Melbourne, Australia
sophiemh@unimelb.edu.au

Stefano Massei
École polytechnique fédérale
de Lausanne
Lausanne 1015, Switzerland
stefano.massei@epfl.ch

ABSTRACT

Branching processes [1] describe the dynamics of a population of individuals which reproduce and die independently, according to some specific probability distributions. More precisely, we assume that any individual has a unit lifetime, at the end of which it might give birth to one or more offsprings simultaneously. This is encoded into the probability generating function $P(z) := \sum_{j\in\mathbb{N}} p_j z^j$ where p_j is the probability of generating j individuals. These kind of processes are known in the literature as *Galton-Watson processes*.

We consider populations that are certain to become extinct, yet appear to be stationary over any reasonable time scale. More precisely, we are interested in characterizing the quasi-stationary distribution of the process, i.e., the asymptotic distribution of the population size, conditional on its survival.

Yaglom [2] proved that if $m := P'(1) \in (0,1)$ then the quasi stationary distribution exists and its probability generating function $G(z) := \sum_{j\in\mathbb{N}} g_j z^j$ solves the Schröder functional equation

$$G(P(z)) = mG(z) + 1 - m, \qquad \forall z \in [0,1]. \tag{1}$$

We study the link between the regularity of $P(z)$ and that of $G(z)$ and we propose a strategy for solving (1).

In the case where $P(z)$ and $G(z)$ are analytic on a disc of radius $r > 1$, we rewrite (1) as

$$\int_{|t|=r} \frac{G(t)}{t - P(z)} dt = mG(z) + 1 - m, \qquad \forall |z| \le r. \tag{2}$$

The discretization of (2) leads to a numerical method that is capable to find arbitrary accurate approximations of the coefficients of $G(z)$. Moreover, we point out the (numerical) low-rank structure that appears in the discretized problem, and we show how to exploit it in the proposed procedure. Numerical tests confirm the nice scalability of the computational cost with respect to the accuracy of the approximation of $G(z)$. Finally, we discuss the extension of the technique to multi-type processes and how to deal with the curse of dimensionality.

1. REFERENCES

[1] T. E. Harris. *The theory of branching processes*. Courier Corporation, 2002.

[2] A. M. Yaglom. Certain limit theorems of the theory of branching random processes. In *Doklady Akad. Nauk SSSR (NS)*, volume 56, page 3, 1947.

MAM10 2019, Hobart, Australia

Rate of strong convergence of stochastic fluid processes to Markov-modulated Brownian motion

Keynote speaker

Giang Nguyen
School of Mathematical Sciences
The University of Adelaide
Adelaide, Australia
giang.nguyen@adelaide.edu.au

Oscar Peralta-Gutierrez
School of Mathematical Sciences
The University of Adelaide
Adelaide, Australia
oscar.peraltagutierrez@
adelaide.edu.au.edu.au

ABSTRACT

In [2], the authors constructed a sequence of stochastic fluid processes and showed that it converges weakly to a Markov-modulated Brownian motion $\{(X_t, \varphi_t)\}_{t\geq 0}$.

Here, we construct another sequence of stochastic fluid processes, with different characteristics to the ones considered in [2], and show that it converges strongly to $\{(X_t, \varphi_t)\}_{t\geq 0}$. We also show that the rate of this almost sure convergence is proportional to $n^{-1/2}\log n$.

When reduced to the special case of standard Brownian motion, our convergence rate is an improvement over that obtained by [1], which is proportional to $n^{-1/2}(\log n)^{5/2}$.

1. REFERENCES

[1] L. G. Gorostiza and R. J. Griego. Rate of convergence of uniform transport processes to Brownian motion and application to stochastic integrals. *Stochastics*, 3(1-4):291–303, 1980.

[2] G. Latouche and G. T. Nguyen. The morphing of fluid queues into Markov-modulated Brownian motion. *Stochastic Systems*, 5(1):62–86, 2015.

MAM10 2019, Hobart, Australia

One-sided Markov additive processes: exit problems and related topics

Keynote speaker

Zbigniew Palmowski
Wrocław University of Science and Technology

ABSTRACT

This tutorial continues the story of Markov additive processes presented by Jevgenijs Ivanovs, focusing on exit problems and possible applications in applied probability intimately related with queueing theory, risk theory and financial mathematics.

At the beginning, we discuss main ideas presented in [2] and [3] corresponding to exit identities and potential measures. Later, following [4] and [6], we consider the seminal Wiener-Hopf factorization. We will also mention new models based on random observations (see [5]) and on the killing that depends on the state of the Markov additive processes (see [1]).

At the end of the tutorial we show extended list of possible applications of the above mentioned theory.

1. REFERENCES

[1] I. Czarna, A. Kaszubowski, S. Li and Z. Palmowski (2018). Fluctuation identities for omega-killed Markov additive processes and dividend problem. Submitted for publication, https://arxiv.org/abs/1806.08102.

[2] J. Ivanovs and Z. Palmowski (2012). Occupation densities in solving exit problems for Markov additive processes and their reflections. *Stochastic Processes and their Applications* 122(9), p. 3342–3360.

[3] J. Ivanovs (2014). Potential measures of one-sided Markov additive processes with reflecting and terminating barriers. *Journal of Applied Probability* 51(4), 2014, p. 1154–1170.

[4] J. Ivanovs (2017). Splitting and time reversal for Markov additive processes. *Stochastic Processes and Their Applications* 127(8), p. 2699–2724.

[5] J. Ivanovs and H. Albrecher (2017). Strikingly simple identities relating exit problems for Lévy processes under continuous and Poisson observations. *Stochastic Processes and Their Applications* 127(2), p. 643–656.

[6] P. Klusik and Z. Palmowski (2014). A note on Wiener-Hopf factorization for spectrally negative Markov Additive processes. *Journal of Theoretical Probability* 27, p. 202–219.

MAM10 2019, Hobart, Australia

Quasi stationarity

Keynote speaker

Phil Pollett
The University of Queensland

ABSTRACT

Many biological systems are predicted to "die out", yet they persist over any reasonable time scale. This phenomenon, termed quasi stationarity, is apparent in many biological systems: a population may persist over many years; an infection may become endemic in a population; a chronically ill patient may survive for a long period of time. I will review several approaches to modelling this behaviour.

MAM10 2019, Hobart, Australia

Construction of algorithms for discrete-time quasi-birth-and-death processes through physical interpretation

Aviva Samuelson *
School of Natural Sciences
University of Tasmania
TAS 7001, Australia,
aviva.samuelson@utas.edu.au

Małgorzata M. O'Reilly * †
School of Natural Sciences
University of Tasmania
TAS 7001, Australia,
malgorzata.oreilly@utas.edu.au

Nigel G. Bean *
School of Mathematical Sciences
University of Adelaide
SA 5005, Australia,
nigel.bean@adelaide.edu.au

ABSTRACT

We apply physical interpretations to construct algorithms for the key matrix $\mathbf{G}$ in discrete-time quasi-birth-and-death (dtQBD) and its z-transform $\mathbf{G}(z)$, motivated by the work on stochastic fluid models (SFMs) in [13]. In this methodology, we first write a summation expression for $\mathbf{G}(z)$ by considering a physical interpretation similar to that of an algorithm in [13]. Next, we construct the corresponding iterative scheme, and prove its convergence to $\mathbf{G}(z)$.

In particular, here we consider the physical interpretation of Algorithm 1 for $\mathbf{\Psi}(s)$ in [13], and use a similar physical interpretation for $\mathbf{G}(z)$ partitioned into three sections, each expressed in terms of matrices analogous to block matrices in the fluid generator $\mathbf{Q}(s)$ in stochastic fluid models.

1. INTRODUCTION

We consider stochastic fluid models (SFMs) and quasi-birth-and-death processes (QBDs), which are key processes in the literature of matrix-analytic methods (MAMs). We exploit the similarities between them in order to apply analogous physical interpretations to construct new expressions and algorithms for the key matrix $\mathbf{G}$ in discrete-time QBDs (dtQBDs) and its z-transform $\mathbf{G}(z)$.

Many expressions in the theory of SFMs, including those for the matrix $\mathbf{\Psi}(s)$, have underlying physical interpretations, which are obtained by deconstructing sample paths into various components, and then writing corresponding expressions in terms of fluid generator $\mathbf{Q}(s)$ [9, 11, 12, 13].

Here, we apply a physical interpretation and conditioning similar to that used in the construction of [13, Algorithm 1] for matrix $\mathbf{\Psi}(s)$ in SFMs to derive an expression and algorithm for the matrix $\mathbf{G}(z)$.

The main algorithms for QBDs were generated with physical interpretations where the set of included sample paths were partitioned, according to the iteration count, by the maximum level reached. The linearly-convergent algorithms then had a linearly increasing maximum level, while the quadratically convergent algorithms had a geometrically increasing maximum level.

Such a partitioning was not entirely appropriate when algorithms were being developed for SFMs. Instead, in [11, Sections 3.1-3.5], sample paths included in the matrix $\mathbf{\Psi}(s)$ were partitioned according to a *key* level, y, and the behaviours on the required three sample path components were carefully controlled: starting at level 0 and reaching level y, leaving level y before returning to level y, and starting at level y until reaching level 0 for the first time. Each of these components can be expressed in terms of the fluid generator $\mathbf{Q}(s)$.

To construct an analogous algorithm for $\mathbf{G}(z)$, we use a similar partitioning principle and three sample path components to construct the iterations for $\mathbf{G}(z)$. We express each component in terms of the matrix $\mathbf{M}(z)$, the dtQBD-equivalent of the fluid generator $\mathbf{Q}(s)$.

2. PRELIMINARIES

2.1 Discrete-time QBDs

Consider a dtQBD, denoted $\{X_t : t = 0, 1, 2, \ldots\}$, on a two-dimensional state space $\{(n, i) : n \geq 0, 1 \leq i \leq m\}$, with level variable n and phase variable i, and the one-step transition probability matrix

$$\mathbf{P} = \begin{bmatrix} \mathbf{B} & \mathbf{A}_+ & \mathbf{0} & \mathbf{0} & \cdots \\ \mathbf{A}_- & \mathbf{A}_0 & \mathbf{A}_+ & \mathbf{0} & \cdots \\ \mathbf{0} & \mathbf{A}_- & \mathbf{A}_0 & \mathbf{A}_+ & \cdots \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_- & \mathbf{A}_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{1}$$

where matrices $\mathbf{B}, \mathbf{A}_+, \mathbf{A}_-, \mathbf{A}_0$ are square matrices of order m such that, for all $i, j \in \{1 \leq i \leq m\}$,

$$[\mathbf{B}]_{ij} = P(X_{t+1} = (0, j) \mid X_t = (0, i)), \tag{2}$$
$$[\mathbf{A}_+]_{ij} = P(X_{t+1} = (n+1, j) \mid X_t = (n, i)), \tag{3}$$
$$[\mathbf{A}_-]_{ij} = P(X_{t+1} = (n-1, j) \mid X_t = (n, i)), \tag{4}$$
$$[\mathbf{A}_0]_{ij} = P(X_{t+1} = (n, j) \mid X_t = (n, i)). \tag{5}$$

Discrete-time QBDs have been used to analyse a variety of real-life situations including modelling embedded queues [16] and maintenance [14]. The analytical expressions for the

*Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers.

†The second author would like to thank the Australian Research Council for funding this research through Linkage Project LP140100152.

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

key quantities in the transient and stationary analysis of this model have led to powerful algorithms. The derivations of both the analytic expressions and subsequent algorithms appear in [1, 6, 16].

We now define a key building block used in the construction of the algorithm in later sections.

DEFINITION 1. *For complex number z inside the unit disk, let the matrices $\mathbf{M}_+(z) = [M_{+ij}(z)]_{1\le i,j\le m}$ and $\mathbf{M}_-(z) = [M_{-ij}(z)]_{1\le i,j\le m}$ be given by*

$$\begin{aligned}\mathbf{M}_+(z) &= \sum_{n=0}^{\infty}(\mathbf{A}_0 z)^n\mathbf{A}_+ z = (\mathbf{I}-\mathbf{A}_0 z)^{-1}\mathbf{A}_+ z,\\ \mathbf{M}_-(z) &= \sum_{n=0}^{\infty}(\mathbf{A}_0 z)^n\mathbf{A}_- z = (\mathbf{I}-\mathbf{A}_0 z)^{-1}\mathbf{A}_- z.\end{aligned}$$

For an irreducible discrete-time QBD, $(\mathbf{I}-\mathbf{A}_0 z)^{-1}$ exists [17].

The entry $[\mathbf{M}_+(z)]_{ij}$ is the probability generating function (PGF) of the time taken for the process to reach level $(n+1)$ for the first time and do so in phase j, whilst avoiding level $(n-1)$, given the process starts in level $n>0$ in phase i at time zero.

The entry $[\mathbf{M}_-(z)]_{ij}$ is the PGF of the time taken for the process to reach level $(n-1)$ for the first time and do so in phase j, whilst avoiding level $(n+1)$, given the process starts in level $n>0$ in phase i at time zero.

The particular quantity we consider in this paper is the matrix $\mathbf{G}(z)$. Before defining $\mathbf{G}(z)$, we define τ as the time taken to first reach level $(n-1)$. Then the (i,j)-th entry of the matrix $\mathbf{G}(z)$ is defined

$$[\mathbf{G}(z)]_{ij} = E[z^\tau I\{\tau<\infty, X_\tau=(n-1,j)\}|X_0=(n,i)], \quad (6)$$

where $[\mathbf{G}(z)]_{ij}$ records the PGF of the time taken for the process to reach level $n-1$ for the first time and do so in phase j, given the process starts in level n at phase i. Note ,$I\{\cdot\}$ denotes the indicator function throughout the paper.

For $0\le z\le 1$, the matrix $\mathbf{G}(z)$ is the minimal non-negative solution [17] to

$$\mathbf{G}(z) = \mathbf{A}_- z + \mathbf{A}_0 z\mathbf{G}(z) + \mathbf{A}_+ z(\mathbf{G}(z))^2. \quad (7)$$

Transforming this into a fixed-point equation by using the recommended iteration by Neuts [17], and then representing in terms of $\mathbf{M}_+(z)$ and $\mathbf{M}_-(z)$, we get

$$\mathbf{G}(z) = \mathbf{M}_-(z) + \mathbf{M}_+(z)(\mathbf{G}(z))^2 \quad (8)$$

$$\text{or} \quad \mathbf{G}(z) = (\mathbf{I}-\mathbf{M}_+(z)\mathbf{G}(z))^{-1}\mathbf{M}_-(z). \quad (9)$$

2.2 SFMs

Consider a SFM, denoted $\{(\varphi(t), X(t)) : t\ge 0\}$, with phase variable $\varphi(t)\in\mathcal{S}=\{1,\ldots,n\}$ and level variable $X(t)$ with lower bound at zero, such that:

- the phase process $\{\varphi(t) : t\ge 0\}$ is an irreducible continuous-time Markov chain (CTMC) with state space $\mathcal{S}$ and generator $\mathbf{T}=[T_{ij}]_{i,j\in\mathcal{S}}$,
- the level variable $X(t)$ changes at rate $dX(t)/dt = c_{\varphi(t)}$ at time t whenever $X(t)>0$, and at rate $\max\{c_{\varphi(t)},0\}$ whenever $X(t)=0$.

SFMs have been used in the analysis of a variety of real-life situations, including telecommunications systems [19], risk assessment [7], power generation systems [10] and congestion control [18].

The stationary and transient analysis of SFMs and powerful algorithms for the numerical evaluations of various performance measures can be found in [2, 3, 4, 5, 11, 13, 20].

We now define the key building blocks in SFMs used in the construction of the algorithms in [10, 11, 12, 13, 21]. Let $\mathcal{S}_+=\{i\in\mathcal{S}: c_i>0\}$, $\mathcal{S}_-=\{i\in\mathcal{S}: c_i<0\}$, $\mathcal{S}_0=\{i\in\mathcal{S}: c_i=0\}$. Block matrices $\mathbf{Q}_{++}(s)$ and $\mathbf{Q}_{--}(s)$ in the fluid generator $\mathbf{Q}(s)$ introduced in [12] are analogous to the building blocks $\mathbf{M}_+(z)$ and $\mathbf{M}_-(z)$ in QBDs, respectively.

DEFINITION 2. *For $s\in\mathbb{C}$ with $\mathbb{R}(s)\ge 0$, let*

$$\mathbf{Q}(s) = \begin{bmatrix}\mathbf{Q}_{++}(s) & \mathbf{Q}_{+-}(s)\\ \mathbf{Q}_{-+}(s) & \mathbf{Q}_{--}(s)\end{bmatrix}, \quad (10)$$

where

$$\begin{aligned}\mathbf{Q}_{++}(s) &= \mathbf{C}_+^{-1}[\mathbf{T}_{++} - s\mathbf{I} - \mathbf{T}_{+0}(\mathbf{T}_{00}-s\mathbf{I})^{-1}\mathbf{T}_{0+}],\\ \mathbf{Q}_{--}(s) &= \mathbf{C}_-^{-1}[\mathbf{T}_{--} - s\mathbf{I} - \mathbf{T}_{-0}(\mathbf{T}_{00}-s\mathbf{I})^{-1}\mathbf{T}_{0-}],\\ \mathbf{Q}_{+-}(s) &= \mathbf{C}_+^{-1}[\mathbf{T}_{+-} - \mathbf{T}_{+0}(\mathbf{T}_{00}-s\mathbf{I})^{-1}\mathbf{T}_{0-}],\\ \mathbf{Q}_{-+}(s) &= \mathbf{C}_-^{-1}[\mathbf{T}_{-+} - \mathbf{T}_{-0}(\mathbf{T}_{00}-s\mathbf{I})^{-1}\mathbf{T}_{0+}].\end{aligned}$$

The physical interpretation is reliant on the following definitions of the in-out fluid $h(t)$ and the first time for $h(t)$ to hit some $y>0$. For any $t\ge 0$, define the random variable

$$h(t) = \int_{u=0}^{t}|c_{\varphi(u)}|du, \quad (11)$$

interpreted as the total amount of fluid that has entered or exited the buffer $X(\cdot)$ during the time interval $[0,t]$, and referred to as the *in-out fluid* [12] of the process $X(\cdot)$. Also, for any $y>0$, define the random variable

$$\omega(y) = \inf\{t>0 : h(t)=y\}, \quad (12)$$

interpreted as the first time at which the in-out fluid of the process $X(\cdot)$ reaches y.

Subsequently, the physical interpretation of $[e^{\mathbf{Q}_{++}(s)y}]_{ij}$ as shown in [12] and extended in [21] is the Laplace-Stieltjes Transform (LST) of the distribution of time for the in-out fluid to reach y for the first time and do so when the process is in phase $j\in\mathcal{S}_+$ whilst avoiding phases in $\mathcal{S}-$, given that the in-out fluid starts at 0 and the process starts in phase $i\in\mathcal{S}_+$.

Further, let $\theta(x)=\inf\{t>0 : X(t)=x\}$ be the first passage time to level x. For $i\in\mathcal{S}_+$, $j\in\mathcal{S}_-$, and $s\in\mathbb{C}$, where $\mathbb{R}(s)\ge 0$, $[\mathbf{\Psi}(s)]_{ij}$ is given by the conditional expectation

$$[\mathbf{\Psi}(s)]_{ij} = E[e^{s\theta(x)}I\{\theta(x)<\infty, \varphi(\theta(x))=j\}|X(0)=x, \varphi(0)=i]. \quad (13)$$

The physical interpretation of $[\mathbf{\Psi}(s)]_{ij}$ is the LST of the time taken for the process to hit level x for the first time and does so in phase j, given the process starts from level x whilst avoiding levels below x.

3. LOWEST-TROUGH ALGORITHM

We construct a lowest-trough algorithm (LT) for the matrix $\mathbf{G}(z)$ by considering the physical interpretation of the sample path corresponding to Algorithm 1 for matrix $\mathbf{\Psi}(s)$ in [13]. Algorithm 1 partitions all relevant sample paths according to the lowest-trough observed in any sample path, corresponding to $\mathbf{\Psi}(s)$, occurring at level y.

In this section, we first summarise Algorithm 1 in [13] by stating the iterative scheme, the integral equation equivalent

to the iterative expression, and the physical interpretation of the expression. Next, we apply a similar physical interpretation to derive a summation equation for $\mathbf{G}(z)$. Further, we construct an iterative scheme resulting from that equation, denoting by $\mathbf{G}_n^{LT}(z)$ the matrix calculated in the n-th iteration of the scheme. Finally, we let the corresponding algorithm be called the LT Algorithm, and prove its convergence to $\mathbf{G}(z)$.

3.1 Algorithm 1 for $\Psi(s)$ in SFMs.

Algorithm 1 in [13] is based upon the following iterative scheme, which continues until an appropriate stopping criterion is met.

1. Let $\mathbf{\Psi}_0(s) = \mathbf{0}$.
2. For $n = 0, 1, 2, \ldots$, let $\mathbf{\Psi}_{n+1}(s)$ be the unique solution to

$$\begin{aligned}&\mathbf{Q}_{++}(s)\mathbf{\Psi}_{n+1}(s) + \mathbf{\Psi}_{n+1}(s)\mathbf{Q}_{--}(s)\\ &\quad = -\mathbf{Q}_{+-} - \mathbf{\Psi}_n(s)\mathbf{Q}_{-+}(s)\mathbf{\Psi}_n(s).\end{aligned} \tag{14}$$

For $s \geq 0$, the above iterative scheme converges to the minimal non-negative solution of the Riccati equation for $\mathbf{\Psi}(s)$ [11, 13].

As described in [13], equation (14) is equivalent to

$$\begin{aligned}\mathbf{\Psi}_{n+1}(s) &= \int_{y=0}^{\infty} e^{\mathbf{Q}_{++}(s)y}(\mathbf{Q}_{+-}(s)\\ &\quad +\mathbf{\Psi}_n(s)\mathbf{Q}_{-+}(s)\mathbf{\Psi}_n(s))e^{\mathbf{Q}_{--}(s)y}dy,\end{aligned} \tag{15}$$

which has the following physical interpretation, after defining Φ_n as the set of sample paths contributing to $\mathbf{\Psi}_n$.

Each sample path contributing to $[\mathbf{\Psi}_{(n+1)}(s)]_{ij}$ has three stages outlined below and depicted in Figure 1.

1. Given the process starts at level 0 in phase $i \in \mathcal{S}_+$, the process remains in some phases in $\mathcal{S}_+ \cup \mathcal{S}_0$ until the process reaches level y in phase $i_1 \in \mathcal{S}_+$ whilst avoiding any transitions into $\mathcal{S}_-$. The LST corresponding to this stage is $[e^{\mathbf{Q}_{++}(s)y}]_{ii_1}$.
2. Given the process starts at level y in phase $i_1 \in \mathcal{S}_+$, the process:
   - Either makes a transition from phase i_1 to $i_2 \in \mathcal{S}_-$ with instantaneous LST rate $[\mathbf{Q}_{+-}(s)]_{i_1i_2}$.
   - Or, the process leaves level y in phase $i_1 \in \mathcal{S}_+$, before returning to level y in some phase $i_3 \in \mathcal{S}_-$ along a path in Φ_n with LST $[\mathbf{\Psi}_n(s)]_{i_1i_3}$.
     Then the process makes a transition from phase i_3 to phase $i_4 \in \mathcal{S}_+$ with instantaneous LST rate $[\mathbf{Q}_{-+}(s)]_{i_3i_4}$.
     The process then leaves level y in phase $i_4 \in \mathcal{S}_+$, before returning to level y in some phase i_2 along a path in Φ_n with LST $[\mathbf{\Psi}_n(s)]_{i_4i_2}$.

   The LST rate corresponding to this stage is $([\mathbf{Q}_{+-}(s) + \mathbf{\Psi}_n(s)\mathbf{Q}_{-+}(s)\mathbf{\Psi}_n(s)]_{i_1i_2})$.
3. Given the process starts from level y in phase $i_2 \in \mathcal{S}_-$, the process remains in some phases in $\mathcal{S}_- \cup \mathcal{S}_0$ until the process drains to level 0 and does so in phase j, whilst avoiding any transitions into $\mathcal{S}_+$. The LST of the time taken to complete this stage is $[e^{\mathbf{Q}_{--}(s)y}]_{i_2j}$.

By integrating over all possible y, all possible sample paths are captured in (15).

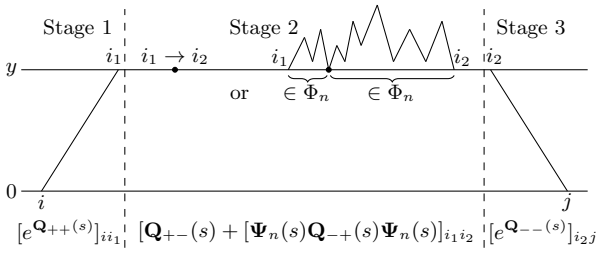


Figure 1: A sample path for $\mathbf{\Psi}_{n+1}(s)$ from equation (15) for a particular y.

3.2 Lowest-trough algorithm for $\mathbf{G}(s)$ in dtQBDs

We now construct the LT Algorithm using a similar physical interpretation to that of Algorithm 1 in [13], as described in Section 3.1 above. Let $\mathbf{G}_0(z) = 0$ and define Ω_n^{LT}, $n \geq 0$, as the set of sample paths contributing to $\mathbf{G}_n(z)$.

Suppose that the process starts from level 1 in phase i and first reaches level 0 in phase j. Then for each sample path contributing to $\mathbf{G}_n(z)$ the following three stages must occur, also depicted in Figure 2.

1. Given the process starts at level 1 in phase i, the process reaches level k in some phase i_1 without making a downwards transition. The PGF corresponding to this stage is $[\mathbf{M}_+(z)^{k-1}]_{ii_1}$.
2. Given the process starts from level k in phase i_1, the process:
   - Either makes a transition to level $(k-1)$ in phase i_2, without reaching leve $(k+1)$, with PGF $[\mathbf{M}_-(z)]_{i_1i_2}$.
   - Or, the process first reaches level $(k+1)$ for the first time whilst avoiding level $(k-1)$, and does so in some phase i_3. The corresponding PGF is $[\mathbf{M}_+(z)]_{i_1i_3}$.
     Next, given the process starts from level $(k+1)$ in phase i_3, it reaches level k for the first time and does so in some phase i_4 along a path Ω_n^{LT}. The corresponding PGF is $[\mathbf{G}_n(z)]_{i_3i_4}$.
     The process must repeat the above at least one more time ending at level k in some phase i_5 under the same restrictions. The corresponding PGF is $[\sum_{\ell=2}^{\infty}(\mathbf{M}_+(z)\mathbf{G}_n(z)))^{\ell}]_{i_1i_5}$. Finally, the process makes a transition from level k in phase i_5 to level $(k-1)$ in phase i_2.

   The PGF corresponding to this stage is $[\mathbf{M}_-(z) + \sum_{\ell=2}^{\infty}(\mathbf{M}_+(z)\mathbf{G}_n(z)))^{\ell}\mathbf{M}_-(z)]_{i_1i_2}$.
3. Finally, given the process starts from level $(k-1)$ in phase i_2, the process first drains to level 0 in phase j without making an upwards transition. The PGF corresponding with this stage is $[\mathbf{M}_-(z)^{k-1}]_{i_2j}$.

By stages 1 – 3 above and partitioning on $k \geq 1$, we obtain the following expression,

$$\begin{aligned}\mathbf{G}_{n+1}(z) &= \sum_{k=1}^{\infty}\mathbf{M}_+(z)^{k-1}\left(\mathbf{I} + \sum_{\ell=2}^{\infty}(\mathbf{M}_+(z)\mathbf{G}_n(z)))^{\ell}\right)\mathbf{M}_-(s)^k\\ &= \sum_{k=1}^{\infty}\mathbf{M}_+(z)^{k-1}((\mathbf{I} - \mathbf{M}_+(z)\mathbf{G}_n(z))^{-1}\\ &\quad -\mathbf{M}_+(z)\mathbf{G}_n(z))\mathbf{M}_-(z)^k.\end{aligned} \tag{16}$$

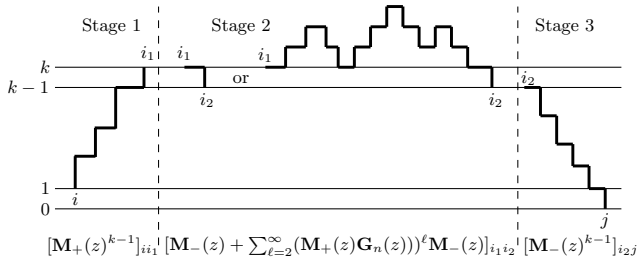


Figure 2: A sample path for $\mathbf{G}_{n+1}^{LT}(z)$ from equation (16) for a particular k.

We use [15, Theorem 3] below, where $\rho(\cdot)$ represents the spectral radius of a given matrix.

LEMMA 1. *Equation*

$$\mathbf{X} \;=\; \mathbf{AXB}+\mathbf{C}, \tag{17}$$

for appropriately sized matrices $\mathbf{A}$, $\mathbf{B}$ *and* $\mathbf{C}$, *has the unique solution given by*

$$\mathbf{X} \;=\; \sum_{k=0}^{\infty}\mathbf{A}^k\mathbf{C}\mathbf{B}^k \tag{18}$$

if and only if $\rho(\mathbf{A})\rho(\mathbf{B})<1$.

After applying Lemma 1, we express equation (16) as the iterative scheme

$$\mathbf{G}_0^{LT}(z)=\mathbf{0}, \tag{19}$$

$$\begin{aligned}&\mathbf{G}_{n+1}^{LT}(z)-\mathbf{M}_+(z)\mathbf{G}_{n+1}^{LT}(z)\mathbf{M}_-(z) \qquad (20)\\ &\quad=\Big((\mathbf{I}-\mathbf{M}_+(z)\mathbf{G}_n^{LT}(z))^{-1}-\mathbf{M}_+(z)\mathbf{G}_n^{LT}(z)\Big)\mathbf{M}_-(z),\end{aligned}$$

to construct the LT Algorithm and then show its convergence to $\mathbf{G}(z)$.

LEMMA 2. $\mathbf{G}_n^{LT}(z)$ *converges to* $\mathbf{G}(z)$ *as* $n\to\infty$.

Proof: Firstly, let Ω be the set of sample paths that contribute to $\mathbf{G}(z)$ and recall that Ω_n^{LT} is the set of sample paths that contribute to the n-th iteration of $\mathbf{G}_n^{LT}(z)$. The physical interpretation of $\mathbf{G}_n^{LT}(z)$ is the PGF of the time taken to traverse paths in $\Omega_n^{LT}\subseteq\Omega$, and so $\mathbf{0}\le\mathbf{G}_n^{LT}(z)\le\mathbf{G}(z)$.

Now, from the physical interpretation, for all n, the sample paths contributing to $\mathbf{G}_n^{LT}(z)$, Ω_n^{LT} must also contribute to $\mathbf{G}(z)$, that is $\Omega_n^{LT}\subseteq\Omega$. Now, consider an arbitrary sample path in Ω. That path must either have a single peak at some level $k\ge1$ or a minimum trough at some level k. As such that path is counted by the n-th iteration of Ω_n^{LT} for all n sufficiently large to allow for the necessary sample path components. Since the sample path was chosen arbitrarily, then all sample paths in Ω are contained within Ω_n^{LT} for some n. The result follows. ■

4. NUMERICAL EXAMPLE

Consider the six-phase dtQBD version of Example 1 in [8], with

$$\mathbf{A}_+=\begin{bmatrix}0.0151 & 0.3021 & 0 & 0 & 0 & 0\\ 0 & 0.0151 & 0.3021 & 0 & 0 & 0\\ 0 & 0 & 0.0151 & 0 & 0 & 0\\ 0 & 0 & 0 & 0.0151 & 0.3021 & 0\\ 0 & 0 & 0 & 0 & 0.0151 & 0\\ 0 & 0 & 0 & 0 & 0 & 0.0151\end{bmatrix}, \tag{21}$$

Algorithm 1 LT Algorithm for calculating $\mathbf{G}(z)$

Input: $\mathbf{A}_-$, $\mathbf{A}_0$, $\mathbf{A}_+$
Set a real $\epsilon>0$, $z\in Re>0$.
Set:
$\mathbf{M}_+(z)=(\mathbf{I}-\mathbf{A}_0z)^{-1}\mathbf{A}_+z$,
$\mathbf{M}_-(z)=(\mathbf{I}-\mathbf{A}_0z)^{-1}\mathbf{A}_-z$, and
$\mathbf{G}_n^{LT}(z)=\mathbf{0}$.
while $||\mathbf{G}_{n+1}^{LT}(z)-\mathbf{G}_n^{LT}(z)||_\infty>\epsilon$ **do**
  Compute:
  $C=((\mathbf{I}-\mathbf{M}_+(z)\mathbf{G}_n^{LT}(z))^{-1}-\mathbf{M}_+(z)\mathbf{G}_n^{LT}(z))\mathbf{M}_-(z)$
  Solve:
  $X-\mathbf{M}_+(z)X\mathbf{M}_-(z)=C$
  Set:
  $\mathbf{G}_n^{LT}(z)=X$
end while
Output: $\mathbf{G}(z)\approx\mathbf{G}_n^{LT}(z)$

Work Count:
One matrix inversion, two matrix products, and solving the Sylvester equation, for approximately $61m^3$ floating point operations per iteration.

$$\mathbf{A}_0=\begin{bmatrix}0.6344 & 0.0302 & 0 & 0 & 0 & 0\\ 0.0302 & 0.6042 & 0.0302 & 0 & 0 & 0\\ 0 & 0.0302 & 0 & 0.0302 & 0 & 0\\ 0 & 0 & 0.0302 & 0.6042 & 0.0302 & 0\\ 0 & 0 & 0 & 0.0302 & 0 & 0.0302\\ 0 & 0 & 0 & 0 & 0.0302 & 0.0302\end{bmatrix}, \tag{22}$$

and

$$\mathbf{A}_-=\begin{bmatrix}0.0181 & 0 & 0 & 0 & 0 & 0\\ 0 & 0.0181 & 0 & 0 & 0 & 0\\ 0 & 0 & 0.0181 & 0.9063 & 0 & 0\\ 0 & 0 & 0 & 0.0181 & 0 & 0\\ 0 & 0 & 0 & 0 & 0.0181 & 0.9063\\ 0.9063 & 0 & 0 & 0 & 0 & 0.0181\end{bmatrix}. \tag{23}$$

We set the stopping criterion to $\epsilon=10^{-12}$.

The algorithm achieved the desired precision in the production of the matrix

$$\mathbf{G}=\begin{bmatrix}0.7831 & 0.0149 & 0.0016 & 0.1084 & 0.0015 & 0.0905\\ 0.6538 & 0.0492 & 0.0030 & 0.1889 & 0.0018 & 0.1033\\ 0.0533 & 0.0016 & 0.0183 & 0.9180 & 0.0002 & 0.0087\\ 0.7426 & 0.0015 & 0.0016 & 0.1270 & 0.0022 & 0.1252\\ 0.0650 & 0.0001 & 0.0000 & 0.0040 & 0.0182 & 0.9126\\ 0.9489 & 0.0002 & 0.0000 & 0.0017 & 0.0006 & 0.0485\end{bmatrix}, \tag{24}$$

LT algorithm converges in 60 iterations with an average time of 0.015 seconds on a Dell OptiPlex 7450 AIO.

5. CONCLUSION

We constructed a linearly-convergent lowest-trough algorithm for $\mathbf{G}(z)$ by applying physical interpretation analogous to that of Algorithm 1 for $\mathbf{\Psi}(s)$ in [13].

Future work includes using similar methodology to construct further algorithms through their physical interpretation. That is, we partition sample paths according to some key level k, derive the corresponding iterative schemes, and numerically compare them with the existing algorithms.

6. REFERENCES

[1] S. Ahn, J. Jeon, and V. Ramaswami. Steady state analysis of finite fluid flow models using finite qbds. *Queueing Systems*, 49(3-4):223–259, 2005.

[2] S. Ahn and V. Ramaswami. Fluid flow models and queues — a connection by stochastic coupling. *Stochastic Models*, 19(3):325–348, 2003.

[3] S. Ahn and V. Ramaswami. Transient analysis of fluid flow models via stochastic coupling to a queue. *Stochastic Models*, 20(1):71–101, 2004.

[4] S. Ahn and V. Ramaswami. Efficient algorithms for transient analysis of stochastic fluid flow models. *Journal of Applied Probability*, 42(2):531–549, 2005.

[5] S. Asmussen. Stationary distributions for fluid flow models with or without Brownian noise. *Stochastic Models*, 11(1):21–49, 1995.

[6] S. Asmussen and M. Pihlsgård. Transient properties of many-server queues and related qbds. *Queueing Systems*, 46(3-4):249–270, 2004.

[7] A. Badescu, L. Breuer, A. da Silva Soares, G. Latouche, M. A. Remiche, and D. Stanford. Risk processes analyzed as fluid queues. *Scandinavian Actuarial Journal*, 2005(2):127–141, 2005.

[8] N. Bean, G. Latouche, and P. Taylor. Physical interpretations for quasi-birth-and-death process algorithms. *Accepted*, 2018.

[9] N. G. Bean and M. M. O'Reilly. Spatially-coherent uniformization of a stochastic fluid model to a quasi-birth-and-death process. *Performance Evaluation*, 70(9):578–592, 2013.

[10] N. G. Bean, M. M. O'Reilly, and J. E. Sargison. A stochastic fluid flow model of the operation and maintenance of power generation systems. *IEEE Transactions on Power Systems*, 25(3):1361–1374, 2010.

[11] N. G. Bean, M. M. O'Reilly, and P. G. Taylor. Algorithms for return probabilities for stochastic fluid flows. *Stochastic Models*, 21(1):149–184, 2005.

[12] N. G. Bean, M. M. O'Reilly, and P. G. Taylor. Hitting probabilities and hitting times for stochastic fluid flows. *Stochastic processes and their applications*, 115(9):1530–1556, 2005.

[13] N. G. Bean, M. M. O'Reilly, and P. G. Taylor. Algorithms for the Laplace–Stieltjes transforms of first return times for stochastic fluid flows. *Methodology and Computing in Applied Probability*, 10(3):381–408, 2008.

[14] B. de Jonge and E. Jakobsons. Optimizing block-based maintenance under random machine usage. *European Journal of Operations Research*, 265(2):703–709, 2018.

[15] P. Lancaster. Explicit solutions of linear matrix equations. *SIAM review*, 12(4):544–566, 1970.

[16] G. Latouche and V. Ramaswami. *Introduction to matrix analytic methods in stochastic modeling*, volume 5. Society for Industrial Mathematics, 1999.

[17] M. F. Neuts. *Matrix-geometric solutions in stochastic models: an algorithmic approach*. Dover Publications, 1981.

[18] M. M. O'Reilly. Multi-stage stochastic fluid models for congestion control. *European Journal of Operations Research*, 238:514–526, 2014.

[19] M. M. O'Reilly and Z. Palmowski. Loss rates for stochastic fluid models. *Performance Evaluation*, 70(9):593–606, 2013.

[20] V. Ramaswami. Matrix analytic methods for stochastic fluid flows. In *ITC16: International Teletraffic Congress*, pages 1019–1030, 1999.

[21] A. Samuelson, M. M. O'Reilly, and N. G. Bean. On the generalised reward generator for stochastic fluid models: New riccati equation for $\mathbf{\Psi}$. *Stochastic Models*, 33:495–523, 2017.

Substitution matrices recapitulate amino acid specificity of aaRS phylogenies

Julia Shore
University of Tasmania
Churchill Avenue
Sandy Bay, Tasmania
julia.shore@utas.edu.au

Barbara Holland
University of Tasmania
Churchill Avenue
Sandy Bay, Tasmania
barbara.holland@utas.edu.au

Jeremy Sumner
University of Tasmania
Churchill Avenue
Sandy Bay, Tasmania
jeremy.sumner@utas.edu.au

Kay Nieselt
Universitat Tubingen
Geschwister-Scholl-Platz
Tubingen, Germany
kay.nieselt@uni-tuebingen.de

Peter Wills
University of Auckland
Auckland 1010
New Zealand
p.wills@auckland.ac.nz

ABSTRACT

Aminoacyl-tRNA synthase (aaRS) are a set of 20 enzymes essential in the biological process of gene expression. For most life forms, they can be divided by their chemical properties into two categories: Class I and Class II [3, 4, 1]. Our analysis aimed to find ways of testing the hypothesis that aaRS enzymes came into existence at the same time as the genetic code and had a role in how it was determined.

The methods for this analysis brought about ways of building a family of stochastic rate matrices from a phylogenetic tree which was then fit to empirical data [2]. It was found that for a given tree, the set of Markov matrices that could be generated formed a closed set under matrix multiplication and addition.

The results of the analysis found that trees which took into account aaRS class fit data better than randomly generated trees. Other chemical properties of interest, particularly polarity of amino acids, were used to build trees and it was found that some types of tree generally fit the data better than ones that only took aaRS class into account. However, it was found in these cases that also including aaRS class into the analysis improved the trees even more.

For further analysis, an F-test was developed to compare the matrices generated by two nested trees to see if the fit improvement of their respective matrices was statistically significant. This analysis resulted in confirming that aaRS class did add significant improvement to trees that took into account other chemical properties.

1. REFERENCES

[1] G. Caetano-Anollés, M. Wang, and D. Caetano-Anollés. Structural phylogenomics retrodicts the origin of the genetic code and uncovers the evolutionary impact of protein flexibility. *PLoS One*, 8(8):e72225, 2013.

[2] S. Q. Le and O. Gascuel. An improved general amino acid replacement matrix. *Molecular biology and evolution*, 25(7):1307–1320, 2008.

[3] P. O'Donoghue and Z. Luthey-Schulten. On the evolution of structure in aminoacyl-trna synthetases. *Microbiology and Molecular Biology Reviews*, 67(4):550–573, 2003.

[4] Y. I. Wolf and E. V. Koonin. On the origin of the translation system and the genetic code in the rna world by means of natural selection, exaptation, and subfunctionalization. *Biology Direct*, 2(1):14, 2007.

MAM10 2019, Hobart, Australia

SIR epidemics with stochastic infectious periods

Matthieu Simon
University of Melbourne
School of Mathematics and Statistics
Victoria 3010, Australia
matthieus@unimelb.edu.au

ABSTRACT

SIR epidemic models describe the spread of an infectious disease in a closed homogeneously mixing population subdivided into three classes: the susceptibles, the infected individuals and the removed cases. The class of susceptibles contains the healthy individuals who can become infected. When contaminated, a susceptible is contagious for a random duration, called the infectious period. During this period, he can transmit the disease to the susceptibles, independently of the other infected individuals. Then he becomes a removed case and plays no further role in the spread of the epidemic.

In this work, we consider the case of an SIR epidemic in which the infectious periods are represented by an arbitrary absorbing Markov process. When an individual gets infected, a version of this process starts to govern his period: the rate at which the individual makes contaminations is a function of the state occupied by his infection process, and the removal occurs as soon as it is absorbed.

We use a martingale approach to determine the exact final epidemic outcome, that is, the state of the population when there are no more infected individuals. We derive the distribution of two statistics: the ultimate number of susceptibles and the final severity, a measure of the total cumulative cost due to all infected individuals that emerged during the course of the epidemic.

Next, we consider a particular case where the infection processes are represented by Markov-modulated fluid flows, the contamination rate and severity functions depending on both the level and the phase. This provides a flexible and tractable class of models that allow to incorporate different features in the description of the disease. The continuous variations in the infection mechanism may reflect, for instance, the periods when an individual undergoes a treatment and those when his illness is worsening. We use techniques from matrix-analytic methods to refine the results in this setting. We also provide a closed expression for the basic reproduction number, one of the most widely used estimators of the virulence of the epidemic.

MAM10 2019, Hobart, Australia

Models for the evolution of gene-duplicates: Applications of Phase-Type distributions

Tristan L. Stark
Temple University
1900 N 12th Street
Philadelphia, USA
tristan.stark@temple.edu

Małgorzata M. O'Reilly *†
University of Tasmania
Churchill Avenue
Hobart, Tasmania
malgorzata.oreilly@utas.edu.au

Barbara Holland *
University of Tasmania
Churchill Avenue
Hobart, Tasmania
barbara.holland@utas.edu.au

David A. Liberles *
Temple University
1900 N 12th Street
Philadelphia, USA
daliberles@temple.edu

ABSTRACT

Genomes typically contain thousands of genes that perform different functions. Rather than all arising independently, these genes group into families that share a common evolutionary origin. Gene families arise because of gene duplication, which is thought to be one of the major sources of evolutionary novelty [4, 6]. However, genes are only preserved in genomes if they have a function that is maintained by selection, and at the time of origin a duplicated gene is identical to another gene with the same functions. A duplicate gene can suffer one of several fates [1, 2]: it may be lost (nonfunctionalisation) - effectively destroyed by mutation - leaving its partner to carry out all the gene's functions; both copies may be retained but with complementary and reduced functionality (subfunctionalisation); or it may acquire a new function (neofunctionalisation) that comes to be protected by selection.

In 2017 Stark et. al [5] developed and analyzed a mathematical model for the fate of a pair of duplicated genes. In this model it is assumed that each gene can carry out a number of functions that are controlled by different regulatory regions (e.g. different transcription factor binding sites might activate the gene in different pathways). All the functions are assumed to be protected by selection. Immediately following duplication the genes are both able to perform all of the functions. Over time mutations (modelled as a Poisson process) are able to knock-out regulatory regions or the coding region of different genes. A mutation to the coding region of a gene will inactivate all of it's functions. Eventually the genes will meet one of two fates: either one gene will be lost (nonfunctionalisation) or both genes will be retained but with complementary functions (subfunctionalisation). The innovation in Stark et. al (2017) was to express this stochastic process as an absorbing state Markov chain and to recognize that the results from the rich literature of Phase-Type distributions [3] could be applied to give analytic solutions for the time to absorption into different states.

In this talk we will introduce the model above and also discuss some initial results that extend the framework introduced in [5] to cases where:

- There are $n > 2$ gene duplicates.
- New functions can arise due to mutation in an existing regulatory region.
- New functions can arise due to de novo creation of a regulatory region.

*We would like to thank the Australian Research Council for funding this research through Discovery Project DP180100352.

†ARC Centre of Excellence for Mathematical and Statistical Frontiers.

MAM10 2019, Hobart, Australia

1. REFERENCES

[1] A. Force, M. Lynch, F. B. Pickett, A. Amores, Y.-l. Yan, and J. Postlethwait. Preservation of duplicate genes by complementary, degenerative mutations. *Genetics*, 151(4):1531–1545, 1999.

[2] M. W. Hahn. Distinguishing among evolutionary models for the maintenance of gene duplicates. *Journal of Heredity*, 100(5):605–617, 2009.

[3] G. Latouche and V. Ramaswami. *Introduction to matrix analytic methods in stochastic modeling*, volume 5. Siam, 1999.

[4] S. Ohno. Evolution by gene duplication springer. *New York*, 1970.

[5] T. L. Stark, D. A. Liberles, B. R. Holland, and M. M. O'Reilly. Analysis of a mechanistic markov model for gene duplicates evolving under subfunctionalization. *BMC evolutionary biology*, 17(1):38, 2017.

[6] J. S. Taylor and J. Raes. Duplication and divergence: the evolution of new genes and old ideas. *Annu. Rev. Genet.*, 38:615–643, 2004.

Models for the evolution of microsatellites

Tristan L. Stark
Temple University
1900 N 12th Street
Philadelphia, USA
tristan.stark@temple.edu

Małgorzata M. O'Reilly *†
University of Tasmania
Churchill Avenue
Hobart, Tasmania
malgorzata.oreilly@utas.edu.au

Barbara Holland *
University of Tasmania
Churchill Avenue
Hobart, Tasmania
barbara.holland@utas.edu.au

ABSTRACT

A microsatellite, or simple sequence repeat, is a strand of DNA which repeats a motif of length 1–6 nucleotides [1]. For example, we may have the string of nucleotides ATATATATAT, which is the motif AT repeated 5 times. Microsatellites undergo a mutation process, slipped-strand mispairing, which leads to a change in the number of repeats, at a rate which is orders of magnitude higher than the rate for other forms of mutation. Most existing microsatellite models focus on treating slipped-strand mispairing alone, however, it has been observed [2, 3, 4] that mutation rate varies between interrupted and uninterrupted repeat sequences. This leads to an interaction effect between point mutations (which introduce interruptions) and slipped-strand mispairing. The aim of this project is to develop new models for microsatellite evolution which treat slipped-strand mispairing together with point mutation in a biologically realistic manner, and to fit these models to data in order to better understand the dynamics of microsatellite evolution.

The first part of the project is theoretical. We introduce an absorbing level-dependent quasi-birth and death process to model the evolution of microsatellite sequences, with the levels tracking the conventionally-modelled repeat number, and the phases tracking the extent of interruption in the repeat sequence. This model is developed at the level of individual microsatellite sequences, and is then extended to a population-level model by the introduction of a Poisson process to model the birth of microsatellites. We then derive a transient distribution for the population-level process, which can be fit to some empirical data to estimate, in particular, the extent of the slowdown conferred by interruptions in the repeat sequence, among other biologically meaningful parameters.

*We would like to thank the Australian Research Council for funding this research through Discovery Project DP180100352.

†ARC Centre of Excellence for Mathematical and Statistical Frontiers.

MAM10 2019, Hobart, Australia

1. REFERENCES

[1] H. Ellegren. Microsatellite mutations in the germline: implications for evolutionary inference. *Trends in Genetics*, 16(12):551–558, 2000.

[2] A. Estoup, C. Tailliez, J. Cornuet, and M. Solignac. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris (Apidae). Molecular Biology and Evolution*, 12(6):1074–1084, 1995.

[3] P. Jarne and P. Lagoda. Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution*, 11(10):424–429, 1996.

[4] L. Wissler, L. Godmann, and E. Bornberg-Bauer. Evolutionary dynamics of simple sequence repeats across long evolutionary time scale in genus *Drosophila. Trends in Evolutionary Biology*, 4(1):e7, 2012.

The Markov embedding problem: a new look from an algebraic perspective

Michael Baake*
Bielefeld University
Bielefeld, Germany
mbaake@math.uni-bielefeld.de

Jeremy Sumner†
University of Tasmania
Hobart, Tasmania
jsumner@utas.edu.au

ABSTRACT

Recall that a Markov matrix is said to be "embeddable" if it is expressible as a matrix exponential e^{Qt} where Q is a Markov generator. In general this is known to be a hard characterisation problem: Davies [2] provides a useful summary of what is known in general, with a complete result known only in the 2×2 case [3].

Our particular perspective is to address the problem in a model-specific fashion where we restrict to classes of Markov matrices satisfying particular algebraic conditions (for example, symmetric transition probabilities). In this way, we show that significant progress can be made for particular models and our key observation is that there is, in almost all cases, a clear algebraic relation between an embeddable Markov matrix and its generator. Namely, the generator must belong to the *centralizer* of the Markov matrix (and vice versa).

To illustrate, we will present the complete solution to the (model-specific) embedding problem for so-called *equal-input* models. This is an important foundational model class in the context of phylogenetic and molecular substitution models. We will also explain the role that complex eigenvalues play in producing exceptional examples. This is work in preparation [1].

1. REFERENCES

[1] M. Baake and J. Sumner. The Markov embedding problem: a new look from an algebraic perspective. *In preparation*, 2018.

[2] E. B. Davies. Embeddable Markov Matrices. *Electron. J. Probab.*, 15:1474–1486, 2010.

[3] J. F. C. Kingman. The imbedding problem for finite markov chains. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, 1(1):14–24, 1962.

*Research supported by German Research Foundation (DFG), within the SPP 1590

†Research supported by Australian Research Project DP180102215

MAM10 2019, Hobart, Australia

Algebraic constraints on the transition probability matrices produced from Lie-Markov models

Jeremy Sumner*
University of Tasmania
Hobart, Tasmania
jsumner@utas.edu.au

Julia Shore†
University of Tasmania
Hobart, Tasmania
julia.shore@utas.edu.au

ABSTRACT

Over the last 5 years+ our research group has explored a class of continuous-time Markov chains we like to refer to as "Lie-Markov models" [3]. Each model in this class has the pleasing property of producing substitution matrices that are closed under matrix multiplication. To date, our primary motivation and focus has been on applications to phylogenetics and model of molecular substitution. However the general concept of multiplicative closure of a Markov model ushers in a multitude of only partially resolved mathematical questions.

In particular, there is a close relationship to these models and the well-known algebraic construction of Lie groups and Lie algebras. In that field, the central dogma states that a given Lie group, which is generally a non-linear set, is more easily analysed by passing to its associated Lie algebra (its, linear, tangent space). The Lie group is then recoverable by applying the exponential map to the Lie algebra (technically, the connected component to the identity is recovered in this way).

In the context of the Lie-Markov models, the situation is reversed. A Lie-Markov model is defined by choosing a set of generator matrices that form a Lie algebra. The transition probability matrices are then obtained by applying the exponential map (as it standard in Markov chain theory). However, this construction leaves open the question of what is the algebraic connection between the structure of the resulting transition probability matrices and the linear constraints that define the Lie algebra of generator rate matrices. A notable result follows: the constraints on the transition probabilities are non-linear if and only if the generator matrices span a linear subspace that is closed under commutators (as is required for a Lie algebra) but is *not* closed under matrix multiplication.

Somewhat surprisingly, with notable exceptions (e.g. [1]), only a few previous authors have constructive solutions to this problem in the context of general Lie group theory.

In this talk, we will explain the relevance of the required algebraic structures, explore various illuminating examples, and present an algorithmic approach to finding a solution for general Lie-Markov models.

This is work in preparation [2].

*Research supported by Australian Research Project DP150100088

†Research supported by Australian Research Council Research Training Program scholarship

1. REFERENCES

[1] W. A. de Graaf. *Computation with linear algebraic groups*. Chapman and Hall/CRC, 2017.

[2] J. Shore and J. Sumner. Finding the Lie group associated to a Lie algebra of matrices. *In preparation.*

[3] J. G. Sumner, J. Fernández-Sánchez, and P. D. Jarvis. Lie markov models. *Journal of Theoretical Biology*, 298:16–31, 2012.

MAM10 2019, Hobart, Australia

Maximum likelihood rearrangement distance for circular genomes

Venta Terauds *
University of Tasmania
Churchill Avenue
Hobart, Tasmania
venta.terauds@utas.edu.au

Jeremy Sumner *
University of Tasmania
Churchill Avenue
Hobart, Tasmania
jeremy.sumner@utas.edu.au

ABSTRACT

Phylogenetic modelling attempts to recover the evolutionary relationships between present-day biological organisms. Typical input to phylogenetic methods is genomic data such as DNA or amino acid sequences. There are many techniques available, but the predominant modern approach is to model sequence evolution (be it DNA, amino acids, or other) as a continuous-time Markov chain on a finite state space and then use likelihood (or a Bayesian approach) to estimate model parameters and the most likely evolutionary history.

In contrast to the point mutation focus of much phylogenetic modelling, genome rearrangement models compare genomes with identifiably similar content, such as genes or other large scale genomic units, and focus on differences in structure, such as the order that these units appear in the genome. Under these models, evolution occurs via rearrangement of the genomic units.

The calculation of evolutionary distance via models of genome rearrangement has an inherent combinatorial complexity. Various algorithms and estimators have been used to address this, however many of these set quite specific conditions for the underlying model.

We discuss a technique, first presented in in [1] and [2], which applies representation theory to calculate evolutionary distance between circular genomes as a maximum likelihood estimate (MLE) of time elapsed. In [3], the implementation of the technique was explored and it was shown that it may be applied to models with any choice of rearrangements and relative probabilities thereof. We give the results of some initial MLE calculations for various models, and show that one may predict the existence or otherwise of an MLE, for a given pair of genomes, without needing to calculate the entire likelihood function.

*This work was supported by Australian Research Council Discovery Early Career Research Award DE130100423 to JS and by use of the Nectar Research Cloud, a collaborative Australian research platform supported by the National Collaborative Research Infrastructure Strategy.

MAM10 2019, Hobart, Australia

1. REFERENCES

[1] S. Serdoz, A. Egri-Nagy, J. Sumner, B. R. Holland, P. D. Jarvis, M. M. Tanaka, and A. R. Francis. Maximum likelihood estimates of pairwise rearrangement distances. *J. Theoret. Biol.*, 423:31–40, 2017.

[2] J. G. Sumner, P. D. Jarvis, and A. R. Francis. A representation-theoretic approach to the calculation of evolutionary distance in bacteria. *J. Phys. A*, 50(33):335601, 14, 2017.

[3] V. Terauds and J. G. Sumner. Maximum likelihood estimates of rearrangement distance: Implementing a representation-theoretic approach. *Bull Math Biol*, 2018.

Polyp fiction: A stochastic fluid model for the Adaptive Bleaching Hypothesis

Max Wurm
School of Mathematical Sciences
The University of Adelaide
SA 5005, Australia
max.wurm@adelaide.edu.au

Andrew Baird
James Cook University
Townsville, Australia
andrew.baird@jcu.edu.au

Nigel Bean
School of Mathematical Sciences
The University of Adelaide
SA 5005, Australia
nigel.bean@adelaide.edu.au

Sean Connolly
James Cook University
Townsville, Australia
sean.connolly@jcu.edu.au

Ariella Helfgott
Environmental Change Institute
Oxford University
OX1 2JD, UK
ariella.helfgott@ouce.ox.ac.uk

Giang Nguyen
School of Mathematical Sciences
The University of Adelaide
SA 5005, Australia
giang.nguyen@adelaide.edu.au

ABSTRACT

Coral reefs have a wide variety of functions in ocean ecosystems. The process of coral bleaching is a sign of distress, in which the coral polyps eject their algal resident. However, some believe that bleaching could be an adaptive response designed to aid in coral's resilience. This interpretation is known as the Adaptive Bleaching Hypothesis (ABH). Coral-algal symbiosis is crucial to the survival of reef corals, and the ABH posits that bleaching encourages changes in these symbioses and thus may allow coral to be better adapted to the stressful environment, should they subsequently be re-colonised by a more suitable clade of algae.

We consider a stochastic fluid model for the ABH, inspired by the work of Helfgott *et al.* [1], to model the algal density in an abstracted coral 'unit'. We consider the phase process as representing the dominant clade of algae residing in the coral and the level process to represent the algal density, which is a proxy for the rate of production of energy for the coral.

The phase process can change organically due to better-adapted clades growing faster and becoming dominant, or by bleaching, after which other clades may or may not colonise the vacant space. The level input/output rates correspond to algal growth (increasing density), algal decay (decreasing density) and bleaching (rapidly decreasing density).

Our primary aim is to model coral mortality through bleaching, and so we propose that under an algal density threshold z, our coral 'unit' is not receiving enough energy to sustain itself. We derive the Laplace-Stietljes transforms of times spent below this density threshold, and times to reach the threshold. Thus, we can approximate the time until mortality in our coral 'unit' by the time taken for the coral to reach this threshold as a result of bleaching, and subsequently exhaust its energy supplies by spending too long below the threshold.

1. REFERENCES

[1] A. Helfgott, N. Bean, S. Connolly, and A. Baird. A stochastic fluid model of the adaptive bleaching hypothesis on the Great Barrier Reef. Unpublished, 2009.

MAM10 2019, Hobart, Australia

ISBN 978-0-646-99707-0