**IEOR 6711: Stochastic Models I**

**Professor Whitt, Thursday, November 29, 2012**

# Weirdness in CTMC's

"Where's your will to be weird?" – Jim Morrison, *The Doors*

"We are all a little weird. And life is a little weird. And when we find someone whose weirdness is compatible with ours, we join up with them and fall into mutually satisfying weirdnessand call it love – true love." – Robert Fulghum, *True Love*

"Some are born weird, some achieve it, others have weirdness thrust upon them." – Dick Francis, *To the Hilt*

"To spend hundreds and hundreds of hours sitting in front of a computer screen staring at lines of information is pretty tedious. More like a computer programmer. And no matter how cool the Matrix made looking at code seem, computer programmers are even weirder than authors." – Christopher Hopper, *The White Lion Chronicles*

## 1 Weirdness in CTMC's

See Chung [9] or Chapter 15 of Breiman [7] for an early account of the foundations. See Karlin and McGregor [11, 12] for early foundational papers on BD processes. See Asmussen [6], Chapter II and Section II.2 for a recent account. Many books have reasonable accounts. The $W^2$ Lecture Notes on CTMC's focuses on the finite state case, which avoids all the weirdness. The results in the infinite-state case are the same, under appropriate regularity conditions.

For more information about any concept discussed below, see the Wikipedia entry.

### 1.1 The Strong Markov Property

It is important that the Markov property extend to random stopping times. There are complications about the strong Markov property in general Markov processes, but all CTMC's have the strong Markov property.

### 1.2 Instantaneous Transitions: Pure-Jump Processes

This pathology has transitions occurring in 0 time. This is avoided by direct assumption. We directly assume that for each initial state $i$ that the process remains in state $i$ for an exponential length of time with positive finite mean, We then say we are considering a Markov *jump* process.

### 1.3 Explosions and the Minimal Construction

Explosions occur when the process diverges to infinity in finite time. This is possible, even for Markov jump processes. A CTMC is called **regular** if the number of transitions from each state in finite time is finite with probability 1; i.e., a regular CTMC has no explosions. For a pure birth process, divergence to infinity turns out to occur if and only if the expected time to go from any state to infinity is finite.

**Theorem 1.1** (*explosions in a pure birth process*) *A pure birth process is explosive (diverges to infinity in finite time) if and only if the mean time to diverge is finite, i.e.,*

$$\sum_{n=1}^{\infty}(1/\lambda_n) < \infty. \tag{1}$$

**Example 1.1** *Here are simple examples: A pure birth process is explosive is $\lambda_n = cn^2$, $n \geq 1$, but not if $\lambda_n = cn$, $n \geq 1$.*

The CTMC can be well defined even with explosions. The **minimal construction** for a CTMC is based on exponential holding times in each state and a DTMC transition matrix at transition times, but with the added feature that the process is absorbed in an extra "death state" $\Delta$ after infinitely many transitions have taken place. Let $S_n$ be the time of the $n^{\text{th}}$ transition. Let

$$S_\infty \equiv \sup_{n \geq 1} S_n, \tag{2}$$

which is less than or equal it infinity. let the process be defined by

$$X(t) \equiv \Delta \quad \text{if} \quad t \geq S_\infty. \tag{3}$$

The extra state $\Delta$ plays no role in a regular CTMC.

**Theorem 1.2** (*Reuter's condition*) *A pure-jump CTMC is regular if and only if the only nonnegative bounded solution $y$ to the matrix equation*

$$Qy = y \tag{4}$$

*is the 0 vector, with $y_j = 0$ for all $j$.*

See Proposition II.3.3 on p. 47 of Asmussen [6].

**Theorem 1.3** (*Kolmogorov ODE's*) *The Kolmogorov forward and backward ODE's have unique identical solutions for any regular pure-jump CTMC.*

**Theorem 1.4** (*Kolmogorov ODE's for minimal construction*) *The minimal construction yields a solution to the Kolmogorov forward and backward ODE's.*

**Theorem 1.5** (*regularity and recurrence of BD processes*) *An irreducible BD process is recurrent, and thus regular, if and only if*

$$\sum_{n=0}^{\infty}(\lambda_n r_n)^{-1} = \infty, \tag{5}$$

*where $r_0 \equiv 1$ and*

$$r_n \equiv \frac{\lambda_0 \times \lambda_1 \times \cdots \lambda_{n-1}}{\mu_1 \times \mu_2 \times \cdots \mu_n}, \quad n \geq 1. \tag{6}$$

**Corollary 1.1** (*bounded birth rates*) *An irreducible finite-state BD process is always regular if the birth rates are bounded.*

**Theorem 1.6** (*positive recurrence of BD processes*) *An irreducible BD process is positive recurrent if and only if it is recurrent and*

$$\sum_{n=0}^{\infty} r_n < \infty \tag{7}$$

*for $r_n$ defined above. In which case, there is a unique stationary distribution and limiting distribution, which coincide and satisfy*

$$\alpha_j \equiv \lim_{t\to\infty} P(X(t) = j) = \frac{r_j}{\sum_{n=0}^{\infty} r_n}, \quad j \geq 0. \tag{8}$$

**Corollary 1.2** (*finite-state BD processes BD processes*) *An irreducible finite-state BD process is always positive recurrent.*

**Theorem 1.7** (*reversibility*) *The stationary version* (*obtained by taking $P(X(0) = j) = \alpha_j$ for all $j$*) *of a positive recurrent irreducible BD process is time reversible, because*

$$\alpha_i \lambda_i = \alpha_{i+1}\mu_{i+1} \quad \textit{for all} \quad i \geq 0. \tag{9}$$

## 2 Fitting BD Processes to Data

A BD process can be fit to data from any stochastic process that makes all transitions up one or down one. Let $T_i(t)$ be the total time spent in state $i$ in the interval $[0, t]$; Let $A_i(t)$ be the number of transitions up one from state $i$ in the interval $[0, t]$; and Let $D_i(t)$ be the number of transitions down one from state $i$ in the interval $[0, t]$. Then define estimated birth and death rates by

$$\bar{\lambda}_i(t) \equiv \frac{A_i(t)}{T_i(t)} \quad \text{and} \quad \bar{\mu}_i(t) \equiv \frac{D_i(t)}{T_i(t)}. \tag{10}$$

Let the estimated steady-state distribution be

$$\bar{\alpha}_i(t) \equiv \frac{T_i(t)}{t}, \quad i \geq 0. \tag{11}$$

We might say that the BD model fits the data well if the estimated steady-state probability vector $\bar{\alpha}$ agrees closely with the theoretical steady-state distribution based on the estimated birth and death rates, using formula (8). However, **the estimated steady-state probability vector $\bar{\alpha}$ automatically agrees very closely with the theoretical steady-state probability vector based on the estimated birth and death rates.** See $W^2$ [22] and references therein.

In particular, Theorem 1 of [22] shows that $\bar{\alpha}$ coincides exactly with the theoretical steady-state probability vector based on the estimated birth and death rates if the system ends in the same state it starts. Otherwise, there is likelihood-ratio stochastic order (which implies ordinary stochastic order), depending on the ordering of the initial and final states. As illustrated by Corollary 4.1 of [22], it is possible to show, under regularity conditions, that the difference between the two probability vectors goes to 0 as the amount of data increases. Note that this holds **without any model assumptions** beyond having all transitions be up one or down one. In particular, the model need not be Markovian and the behavior could be highly time-dependent. The arrival rate might be highly time-dependent, such as sinusoidal. Nevertheless, a BD model fit to the data will necessarily produce a steady-state distribution that matches the long-run average performance. Of course, the long-run average performance may not match what happens at any particular time.

## 3 First-Passage Times in BD Processes

The first passage time from state $i$ to state $j$, $T_{i,j}$, can be expressed as the sum of the first passage times to neighboring states. If $i < j$ then

$$T_{i,j} = T_{i,i+1} + T_{i+1,i+2} + \cdots + T_{j-1,j} \tag{12}$$

If $i > j$ then

$$T_{i,j} = T_{i,i-1} + T_{i-1,i-2} + \cdots + T_{j+1,j} \tag{13}$$

where the sum in each case is over independent random variables.

The first passage time up to the nearest neighbor has a relatively simple form, because it suffices to consider a finite-state absorbing CTMC; e.g., see Keilson [13]. The first passage time down is more complicated with an infinite state space. **Its Laplace transform can be computed using continued fractions, which can be used to calculate the distribution, using numerical inversion of Laplace transforms**; see Abate and $W^2$[3]. (The numerical inversion algorithm requires computing the Laplace transform for a modest number of complex arguments, e.g., about 50. Those required transform values can in turn be computed by algorithms for calculating (infinite) continued fractions.)

We now give the construction of the Laplace transform recursion for first passage times down. For $i \geq 1$, let $T_i$ be the first passage time down from state $i$ to state $i-1$. Let $\hat{f}_i(s)$ be its Laplace transform, i.e.,

$$\hat{f}_i(s) \equiv E[e^{-sT_i}] = \int_0^\infty e^{-sx} f_{T_i}(x)\, dx. \tag{14}$$

To develop a recursion, we consider the first transition from state $i$. With probability $\lambda_i/(\lambda_i + \mu_i)$, the process moves up to state $i+1$; with probability $\mu_i/(\lambda_i + \mu_i)$, the process moves down to state $i-1$. If the process moves down then the first passage is complete. If the process moves up, then it must move down to $i$ from $i+1$ and then move down from $i$ to $i-1$. Recall that the time of the first transition is independent of the location (basic property of minimum of two independent exponential random variables). Let $L$ be the location of the first transition from $i$ and let $T$ be the time of that transtion. Thus, as in (4.7) of Abate and $W^2$ [3],

$$\begin{aligned} \hat{f}_i(s) &\equiv E[e^{-sT_i}] = P(L = i-1)E[e^{-sT}] + P(L = i+1)\left(E[e^{-sT}]\hat{f}_{i+1}(s)\hat{f}_i(s)\right) \\ &= \left(\frac{\mu_i}{\lambda_i + \mu_i}\right)\left(\frac{\lambda_i + \mu_i}{\lambda_i + \mu_i + s}\right) \\ &\quad \left(\frac{\lambda_i}{\lambda_i + \mu_i}\right)\left(\frac{\lambda_i + \mu_i}{\lambda_i + \mu_i + s}\right)\hat{f}_{i+1}(s)\hat{f}_i(s), \end{aligned} \tag{15}$$

from which we obtain, by simple algebra, the recursive relation

$$\hat{f}_i(s) = \frac{\mu_i}{\lambda_i + \mu_i + s - \lambda_i \hat{f}_{i+1}(s)}. \tag{16}$$

This would be a finite recursion if there were only finitely many states, but the recursion never ends if there are infinitely many states. Nevertheless, for conventional BD processes, we expect a finite limit.

To put this in one of the standard continued fraction (CF) representations, we want no constant factor before $\hat{f}_{i+1}(s)$ in (16). To understand the basic recursion, write out the next step:

$$\hat{f}_i(s) = \frac{\mu_i}{\lambda_i + \mu_i + s - \frac{\lambda_i \mu_{i+1}}{\lambda_{i+1} + \mu_{i+1} + s - \lambda_{i+1}\hat{f}_{i+2}(s)}}. \tag{17}$$

and the step after that

$$\hat{f}_i(s) = \frac{\mu_i}{\lambda_i + \mu_i + s - \frac{\lambda_i \mu_{i+1}}{\lambda_{i+1}+\mu_{i+1}+s-\frac{\lambda_{i+1}\mu_{i+2}}{\lambda_{i+2}+\mu_{i+2}+s-\lambda_{i+2}\hat{f}_{i+3}(s)}}}. \tag{18}$$

To obtain a clean orderly representation, we rewrite the last version as

$$\hat{f}_i(s) = -\frac{1}{\lambda_{i-1}} \left( \frac{-\lambda_{i-1}\mu_i}{\lambda_i + \mu_i + s + \frac{-\lambda_i \mu_{i+1}}{\lambda_{i+1}+\mu_{i+1}+s+\frac{-\lambda_{i+1}\mu_{i+2}}{\lambda_{i+2}+\mu_{i+2}+s-\lambda_{i+2}\hat{f}_{i+3}(s)}}} \right). \tag{19}$$

We then express the result as

$$w_i = c_i \Phi_{n=i}^{\infty} \frac{a_n}{b_n} \quad \text{or} \quad w_i = c_i \left( \frac{a_i}{b_i +} \frac{a_{i+1}}{b_{i+1}+} \frac{a_{i+2}}{b_{i+2}+} \frac{a_{i+3}}{b_{i+3}+} \cdots \right) \tag{20}$$

for

$$c_i \equiv -\frac{1}{\lambda_{i-1}}, \quad a_n \equiv -\lambda_{n-1}\mu_n \quad \text{and} \quad b_n \equiv \lambda_n + \mu_n + s, \quad n \ge i. \tag{21}$$

For a CF (sometimes called a generalized CF, because a CF can be expressed in more than one way), we write

$$w = \Phi_{n=1}^{\infty} \frac{a_n}{b_n} \quad \text{or} \quad w = \frac{a_1}{b_1 +} \frac{a_2}{b_2 +} \frac{a_3}{b_3 +} \frac{a_4}{b_4 +} \cdots \tag{22}$$

There is a relatively simple recursion for calculating the successive approximants due to Euler in 1737, namely,

$$w_n = \frac{P_n}{Q_n}, \tag{23}$$

where $P_0 = 0$, $P_1 = a_1$, $Q_0 = 1$, $Q_1 = b_1$ and

$$P_n = b_n P_{n-1} + a_n P_{n-2} \quad \text{and} \quad Q_n = b_n Q_{n-1} + a_n Q_{n-2}, \quad n \ge 2. \tag{24}$$

**Example 3.1** *A continued fraction for* $\pi$ *is* (22) *with*

$$a_1 = 4, \quad b_1 = 1, \quad a_n = (n-1)^2 \quad \textit{and} \quad b_n = 2 \quad \textit{for all} \quad n \ge 2. \tag{25}$$

*A continued fraction for* $e$ *is* $b_0 + w$ *for* $w$ *in* (23) *with* $b_0 = 2$,

$$a_n = 1 \quad \textit{for} \quad n \ge 1 \quad \textit{and} \quad \{b_k : k \ge 1\} = \{1, 2, 1, 1, 4, 1, 1, 6, 1, 1, 8, 1, 1, 10, 1, 1, 12, \ldots\} \tag{26}$$

*This is sequence* A003417 *in the online encyclopedia of sequences* (OEIS).

For more on integer sequences and stochastic models, see Abate and $W^2$ [4, 5].

# 4 Comparing BD processes

Two BD process can be compared using a sample-path stochastic ordering if the smaller one has lower birth rates and higher death rates. To do the construction, use thinning of a Poisson process whenever the two processes are in the same state $i$. Let potential transitions be generated with a Poisson process having rate $(\lambda_i^{(1)} \bigvee \lambda_i^{(2)}) + (\mu_i^{(1)} \bigvee \mu_i^{(2)})$. Make the upper process have a birth whenever the lower process has a birth; make the lower process have a death whenever the upper process has a death. In that way the two processes each have the given probability law, but the sample paths are ordered w.p.1. This is a variant of the comparisons in $W^2$ [20].

**Theorem 4.1** (*sample path stochastic order for BD processes*) *For $i = 1, 2$, let $\lambda_k^{(i)}$, $k \ge 0$ and $\mu_k^{(i)}$, $k \ge 1$, be birth and death rates for two BD processes. If*

$$\lambda_k^{(1)} \le \lambda_k^{(2)} \quad \textit{and} \quad \mu_k^{(1)} \ge \mu_k^{(2)} \quad \textit{for all} \quad k, \tag{27}$$

*then there exists special versions of these BD processes $\{X^{(i)}(t) : t \ge 0\}$ constructed on the same sample space such that each separately has the correct probability law as a BD process, while*

$$P(X^{(1)}(t) \le X^{(2)}(t) \quad \textit{for all} \quad t \ge 0) = 1. \tag{28}$$

Here is a typical corollary. Let $T_{j,k}^{(i)}$ be the first passage time from state $j$ to state $k$ in BD process $i$. Let $\le_{st}$ denote stochastic order; see Chapter 9 of Ross.

**Corollary 4.1** (*stochastic order of first passage for BD processes*) *Consider two BD processes with birth and death rates ordered as in* (27). *If $j \le k$, then*

$$T_{j,k}^{(2)} \le_{st} T_{j,k}^{(1)}. \tag{29}$$

*If $j \ge k$, then*

$$T_{j,k}^{(1)} \le_{st} T_{j,k}^{(2)}. \tag{30}$$

**Proof.** We establish only (29). Apply Theorem 4.1 to get (28). If $j \le k$, then for that special construction get

$$P(T_{j,k}^{(2)} \le T_{j,k}^{(1)})) = 1. \tag{31}$$

As an immediate consequence, get

$$P(T_{j,k}^{(2)} > a) \le P(T_{j,k}^{(1)} > a) \quad \text{for all} \quad a, \tag{32}$$

which is equivalent to the stated conclusion. Notice that the last statement applies to the distribution of the two processes viewed separately; i.e., it no longer depends on the special construction. ∎

# 5 Spectral Representation

Time-reversible CTMC's such as BD processes have spectral representations where the eigenvalues and eigenvectors are all real. See Chapter 3 of Keilson [13] or see Abate and $W^2$ [2] and references therein. These spectral representations provide explicit representations for the transient transition probabilities and explicit representations for the rate of convergence to the steady-state probability vector.

There is a corresponding algebraic approach to DTMC's too. The Perron-Frobenius theory of positive matrices can be applied. The rough idea is to diagonalize the transition matrix $P$ in DTMC's or $P(t)$ in CTMC's; i.e., we write

$$P = U \Lambda U^{-1}, \tag{33}$$

where the diagonal elements of $\Lambda$ are the eigenvalues of $P$, while $U$ and $U^{-1}$ are made up of the associated eigenvectors. Then

$$P^n = U \Lambda^n U^{-1}, \quad n \ge 1. \tag{34}$$

This structure occurs in a relatively simple form for reversible Markov chains. For simplicity, suppose that the state space is finite of dimension $m$. We now give additional details, following Keilson. First, we consider an irreducible DTMC with transition matrix $P$ and unique stationary probability vector $\pi$, satisfying $\pi = \pi P$. Let $\pi_D$ be the $m \times m$ diagonal matrix with $i^{\text{th}}$ diagonal element $\pi_i$ and all off-diagonal elements 0. Then observe that **the DTMC $P$ is reversible if and only if the $m \times m$ matrix $S \equiv \pi_D P$ is a symmetric matrix**. (A matrix $S$ is symmetric if it equals its transpose, i.e., if $S_{i,j} = S_{j,i}$ for all $i$ and $j$.) We then can apply the finite spectral theorem for real symmetric matrices. The finite-dimensional spectral theorem says that any symmetric matrix whose entries are real can be diagonalized by an orthogonal matrix, i.e., we can write (33) where $U$ and $U^{-1}$ are real and nonsingular with $UU^{-1} = I$ and $\Lambda$ is a diagonal matrix with real eigenvalues. The columns of $U$ are right eigenvector of $P$ while the rows of $U^{-1}$ are left eigenvectors of $P$. We can then apply the Perron-Frobenius theory to conclude that there is one eigenvalue taking the value 1 and all the other eigenvalues satisfy $|\lambda_i| < 1$. We thus have the explicit representations

$$P_{i,j} = \sum_{k=1}^{m} U_{i,k} U^{-1}_{k,j} \lambda_k \quad \text{and} \quad P^n_{i,j} = \sum_{k=1}^{m} U_{i,k} U^{-1}_{k,j} \lambda^n_k = \pi_j + \sum_{k=1}^{m-1} a_{i,j} \lambda^n_k, \tag{35}$$

where $\lambda_k$ are real numbers and, in the last expression, $a_{i,j}$ are real numbers and $|\lambda_k| < 1$, so that the last expression gives an explicit rate of convergence to steady state.

A corresponding story holds for CTMC's. For finite-state CTMC's, that can always be achieved by uniformization (as discussed in §3.4 of the CTMC lecture notes). We can represent a CTMC as a Poisson randomization of a DTMC. We can write the CTMC as

$$X(t) = Y_{N(t)}, \quad t \ge 0, \tag{36}$$

where $\{Y_n : n \ge 0\}$ is a DTMC and $\{N(t) : t \ge 0\}$ is a Poisson process. We let the rate of the Poisson process be $r$, where $r > |Q_{i,i}|$ for all $i$. We let the transition matrix of the DTMC be $P_{i,j} \equiv Q_{i,j}/r$ for all $i \ne j$. That is we let $P = I + r^{-1}Q$. We then have

$$P_{i,j}(t) = \sum_{k=0}^{\infty} P^k_{i,j} \frac{e^{-rt}(rt)^k}{k!}, \tag{37}$$

as in (3.28) of the lecture notes. Hence the spectral representation for the DTMC carries over directly to the CTMC

$$\begin{aligned} P_{i,j}(t) &= \sum_{k=0}^{\infty} P^k_{i,j} \frac{e^{-rt}(rt)^k}{k!} \\ &= \sum_{k=0}^{\infty} \left( \sum_{l=1}^{m} U_{i,l} U^{-1}_{l,j} \lambda^k_l \right) \frac{e^{-rt}(rt)^k}{k!} \\ &= \sum_{l=1}^{m} U_{i,l} U^{-1}_{l,j} \sum_{k=0}^{\infty} \lambda^k_l \frac{e^{-rt}(rt)^k}{k!} \\ &= \sum_{l=1}^{m} U_{i,l} U^{-1}_{l,j} e^{-rt} \sum_{k=0}^{\infty} \frac{(\lambda_l rt)^k}{k!} \\ &= \sum_{l=1}^{m} U_{i,l} U^{-1}_{l,j} e^{-r(1-\lambda_l)t}. \end{aligned} \tag{38}$$

Since $|\lambda_i| \leq 1$ for all $i$ and equals 1 for only one $i$, we have

$$P_{i,j}(t) = \alpha_j + \sum_{l=1}^{m-1} a_{i,j} e^{-r_l t}, \quad t \geq 0, \tag{39}$$

where $r_j > 0$ for all $j$ and $a_{i,j}$ is a real number for all $i$ and $j$.

# 6 Diffusion Approximations in BD Processes

A diffusion process can be regarded as a continuous analog of a BD process. In turn, a BD process can be regarded as a discrete analog of a diffusion process; e.g., see Feller [10]. It is natural to consider diffusion approximations for BD processes; see [8, 17, 19].

# 7 Quasi-Birth-and-Death (QBD) Processes

A QBD process is a CTMC whose rate matrix has a block tri-diagonal form, so that the QBD process is a matrix analog of a BD process; see the book by Latouche and Ramaswami [14].

The books by Neuts [15, 16] discuss far-reaching matrix generalizations of the $M/G/1$ and $GI/M/1$ Markov chains. These are Markov chains that have the same kind of structure as these particular queueing Markov chains, where the elements are replaced by entire matrices. The matrices in block form are upper-triangular or lower-triangular. The QBD is a special case of both. See Section XI.3 of [6] for a quick treatment. See [14] for the theory of QBD processes. See Perry and $W^2$ [18] for an application of QBD processes in fluid approximations based on a stochastic averaging principle.

# References


[1] Abate, J., G. L. Choudhury, G. L., W. Whitt. 1999. An introduction to numerical transform inversion and its application to probability models. in *Computational Probability*, W. Grassman (ed.), Kluwer, Boston, 257–323.

[2] Abate, J., W. Whitt. 1989. Spectral Theory for Skip-Free Markov Chains. *Probability in the Engineering and Information Sciences* 3, 77–88.

[3] Abate, J., W. Whitt. 1999. Computing Laplace Transforms for Numerical Inversion Via Continued Fractions. *INFORMS Journal on Computing* 11, 394–405.

[4] Abate, J., W. Whitt. 2010. Integer Sequences from Queueing Theory. *Journal of Integer Sequences*, 13, article 10.5.5

[5] Abate, J., W. Whitt. 2011. Brownian Motion and the Generalized Catalan Numbers. *Journal of Integer Sequences* 14, article 11.2.6

[6] Asmussen, S. 2003. *Applied Probability and Queues*, second edition, Springer.

[7] Breiman, L. 1968. *Probability*, Addison-Wesley.

[8] Browne, S. and W. Whitt. 1995. Piecewise-Linear Diffusion Processes. In *Advances in Queueing*, J. Dshalalow, ed., CRC Press, Boca Raton, FL, 463–480.

[9] Chung, K. L. 1960. *Markov Chains with Stationary Transition Probabilities.* Springer, Berlin.

[10] Feller, W. 1959. The Birth-and-Death Processes as Diffusion Processes, *J Math. Pures Appl.* 38, 301–395.

[11] Karlin, S. and K. McGregor. 1957a. The Differential Equations of Birth-and-Death Processes, and the Stieltjes Moment Problem *Trans. Amer. Math. Soc.* 85, 489–546.

[12] Karlin, S. and K. McGregor. 1957b. The Classification of Birth and Death Processes *Trans. Amer. Math. Soc.* 85, 366–400.

[13] Keison, J. 1979. *Markov Chain Models – Rarity and Exponentiality*, Applied Math. Sci. 28, Springer, Berlin.

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

[15] Neuts, M. F. 1981. *Matrix-Geometric Solutions in Stochastic Models*, The Johns Hopkins University Press.

[16] Neuts, M. F. 1989. *Structured Stochastic Matrices of M/G/1 Type and their Applications*, Marcel Dekker.

[17] Pang, G., R. Talreja and W. Whitt. 2007. Martingale Proofs of Many-Server Heavy-Traffic Limits for Markovian Queues. *Probability Surveys* 4, 193–267.

[18] Perry, O. and W. Whitt. 2011. An ODE for an Overloaded X Model Involving a Stochastic Averaging Principle. *Stochastic Systems* 1, 59–108.

[19] Stone, C. 1963. Limit theorems for random walks, birth and death processes and diffusions. *Illinois J. Math.* 7, No. 1, 638–660.

[20] Whitt, W. 1981. Comparing Counting Processes and Queues. *Advances in Applied Probability* 13, No. 1, 207–220.

[21] Whitt, W. 2002. *Stochastic-Process Limits*, Springer.

[22] Whitt, W. 2012. Fitting Birth-and-Death Queueing Models to Data. *Statistics and Probability Letters*, 82, 998–1004.