Abstract

In this paper we study a general stochastic fluid model with a single infinite capacity buffer, where the buffer content can change continuously as well as by instantaneous upward jumps. The continuous as well as the instantaneous change is modulated by an external environment process modelled as a finite state continuous time Markov chain. The Laplace-Stieltjes transform of the steady-state joint distribution of the (buffer content, state of the environment) is determined explicitly in terms of the solutions of a generalized eigen-value problem. The methodology is illustrated by several well-known queueing problems.

1 Introduction

A stochastic fluid queueing system describes the input-output flow of a fluid in a storage device, called a buffer. The rates at which the fluid enters and leaves the buffer depend on a random environment process that is usually chosen to be an irreducible CTMC. The study of fluid flow models is motivated by their various applications. One of these are the real-world systems that deal with the processing of continuous entities such as the ones used in the petroleum and chemical industries. Fluid flow models also provide an important tool for the performance analysis of high-speed data networks, or large-scale production systems where a large number of relatively small jobs are processed. Fluid flow models are also used as models of the asymptotic behavior of queues in heavy traffic.

Most of the classical research on stochastic fluid systems in the area of telecommunication is based on the work of Anick, Mitra and Sondhi [2] which is an extension of the pioneering work of Kosten [4]. They study a model with several identical and independent Markov on-off sources that transmit fluid to an infinite-capacity buffer. The fluid is then
processed at a fixed rate. The limiting distribution of the buffer content process is computed as a solution of a set of ordinary differential equations. The identical input sources facilitate the analysis of the differential equations and the main result provides the system's eigenvalues in explicit form. In [8], [9] Mitra generalizes this model by introducing multiple on-off switching. Most of the work in fluid queues deals with the steady-state distribution of the buffer content. See the survey paper by Kulkarni [5] for an extensive overview of the research in this area.

In this paper we consider a modification of the classical fluid model with an infinite capacity buffer where the buffer content process increases continuously at a rate depending on the state of the environment. In the model we study the buffer content can also have instantaneous upward jumps. A similar stochastic model with jumps is studied in [12]. Sengupta considers a bivariate Markov process \( \{(X(t), I(t)), t \geq 0\} \) where \( X(t) \) increases linearly and has downward jumps. Sengupta shows that the steady state joint distribution of \( \{(X(t), I(t)), t \geq 0\} \) has a matrix-exponential form under some additional assumptions that make the process a natural continuous analog of the Markov chains studied by Neuts [11]. A rate matrix \( T \) associated with the joint distribution is the solution of a non-linear matrix integral equation. The author also provides a new simplified characterization for the waiting time and queue length distributions of the \( G1/PH/1 \) queue with FCFS discipline. In the current paper it is shown that for the fluid model with upward jumps the Laplace-Stieltjes transform of the steady state joint distribution of \( \{(X(t), I(t)), t \geq 0\} \) can be explicitly determined in terms of the solutions of a generalized eigen-value problem. The theory is applied to several well-known queueing problems.

2 The Model

Consider a general stochastic fluid model with a single infinite capacity buffer where the buffer content \( X(t) \) can increase continuously as well as by instantaneous jumps. The change of the fluid in the buffer depends on the state of an external random environment process \( \{I(t), t \geq 0\} \) which is taken to be a stochastic process with a finite state space \( S = \{1, \ldots, N\} \). While \( I(t) = i \) the buffer content increases continuously with rate \( r_i \in (-\infty, \infty) \). The process \( \{I(t), t \geq 0\} \) jumps to any state \( j \in S \) (not necessarily different from \( i \)) with probability \( p_{ij} \). In state \( i \in S \) it will make a jump after an exponential amount of time with mean \( 1/q_i \). When the \( I(t) \) process jumps from state \( i \) to state \( j \) the amount of fluid in the buffer can increase by a lumpsum non-negative random amount with a given c.d.f. \( G_{ij}(y), y \geq 0 \) and mean \( m_{ij} \). Thus, the bivariate Markov process \( \{(X(t), I(t)), t \geq 0\} \) can jump from state \( (x, i) \) to state \( (x + y, j) \) with rate \( q_i p_{ij} dG_{ij}(y), y \geq 0 \).

Let \( R \) denote the diagonal \( N \times N \) matrix with the net input rates \( r_i \) along the diagonal

\[
R := \text{diag}[r_1, \ldots, r_N],
\]
and also define
\[ Q_{ij}(x) := q_ip_{ij}g_{ij}(x), \quad x \geq 0, \ i, j \in S, \ i \neq j, \quad \text{(2.1)} \]
\[ Q_{ii}(x) := q_ip_{ii}g_{ii}(x) - q_i, \quad x \geq 0, \ i \in S. \quad \text{(2.2)} \]

Note that \( \{I(t), t \geq 0\} \) itself is a CTMC on \( S \) with rate matrix \( Q = Q(\infty) \). We assume that \( \{I(t), t \geq 0\} \) is irreducible. Let \( \pi_i := \lim_{t \to \infty} P(I(t) = i) \) be the limiting distribution of the \( I(t) \) process. Then the system is stable if the mean net input rate is negative,
\[ \sum_{i=1}^{N} \pi_i(r_i) + \sum_{j=1}^{N} q_ip_{ij}m_{ij} < 0. \quad \text{(2.3)} \]

Define the matrix \( \Gamma = [\Gamma_{ij}] \) as follows:
\[ \Gamma_{ij} = q_ip_{ij}m_{ij}. \]

Then the stability condition can be written in matrix form as follows
\[ \pi(R + \Gamma)e < 0. \quad \text{(2.4)} \]

We shall assume from now on that this stability condition holds.

Now, denote
\[ F_i(t, x) := P(X(t) \leq x, I(t) = i), \ x, t \geq 0, \ i \in S, \quad \text{(2.5)} \]
and
\[ F_i(x) := \lim_{t \to \infty} P(X(t) \leq x, I(t) = i), \ x \geq 0, \ i \in S. \quad \text{(2.6)} \]

The next theorem gives the differential equations satisfied by
\[ F(x) := [F_1(x), \ldots, F_N(x)], \]
and
\[ \frac{dF}{dx} := \left[ \frac{dF_1}{dx}, \ldots, \frac{dF_N}{dx} \right]. \]

We need the following notation:
\[ S_- := \{ i \in S : r(i) < 0 \}, \ N_- := |S_-|, \]
\[ S_0 := \{ i \in S : r(i) = 0 \}, \ N_0 := |S_0|, \]
\[ S_+ := \{ i \in S : r(i) > 0 \}, \ N_+ := |S_+|. \]
Theorem 2.1  The limiting distribution \( F(x) \) satisfies

\[
\frac{dF}{dx} R = F \ast Q(x). \tag{2.7}
\]

The boundary conditions are given by

\[
F_i(0) = 0, \ i \in S_+,
\]

\[
F(0)Q_{j}(0) = 0, \ j \in S_0,
\]

where \( Q_{j}(0) \) is the \( j \)-th column of \( Q(0) \).

**Proof:** First, consider \( F_j(t,x), \ x > 0, \ j \in S \) and condition on a small time interval of length \( h > 0 \) so that

\[
F_j(t,x) = \sum_{i=1}^{N} \int_{z=0}^{x-r_ih} P(X(t-h) \leq x - r_ih - z, I(t-h) = i)q_i p_{ij} hdG_{ij}(z) + (1 - q_jh)P(X(t-h) \leq x - r_jh, I(t-h) = j) + o(h)
\]

\[
= h \sum_{i=1}^{N} \int_{z=0}^{x-r_ih} F_i(t-h, x - r_ih - z)q_i p_{ij} dG_{ij}(z) + (1 - q_jh)F_j(t-h, x - r_jh) + o(h)
\]

Now let \( t \to \infty \) and rearrange the last equation so that

\[
\frac{F_j(x) - F_j(x - r_jh)}{h} = \sum_{i=1}^{N} \int_{z=0}^{x-r_ih} F_i(x - r_ih - z)q_i p_{ij} dG_{ij}(z) - q_j F_j(x - r_jh) + \frac{o(h)}{h}
\]

From the definition of \( Q(x) \) in (2.1) and (2.2) we have the following jump rates that determine the generator \( dQ(x) \) of the bivariate Markov process \( \{(X(t), I(t)), t \geq 0\} \):

\[
dQ_{ij}(x) = q_i p_{ij} dG_{ij}(x), \ x \geq 0, \ i, j \in S, \ i \neq j; \tag{2.10}
\]

\[
dQ_{ii}(x) = \begin{cases} 
q_i p_{ii} dG_{ii}(x), & \text{if } x > 0, \\
q_i(p_{ii} dG_{ii}(x) - 1) & \text{if } x = 0.
\end{cases} \tag{2.11}
\]

Now, we can write the last equation in the following nice form:

\[
\frac{F_j(x) - F_j(x - r_jh)}{h} = \sum_{i=1}^{N} \int_{z=0}^{x-r_ih} F_i(x - r_ih - z) dQ_{ij}(z) + \frac{o(h)}{h}.
\]

After we let \( h \to 0 \) we get

\[
r_j \frac{dF_j(x)}{dx} = \sum_{i=1}^{N} \int_{z=0}^{x} F_i(x - z) dQ_{ij}(z) = \sum_{i=1}^{N} F_i \ast Q_{ij}(x), \tag{2.12}
\]

which in matrix form becomes Eq. (2.7)

\[
\frac{dF}{dx} R = F \ast Q(x).
\]
From the definition of \( F_i(x) \) we have 
\[
F_i(0) = \lim_{t \to \infty} P(X(t) = 0, I(t) = i), \ i \in S.
\]
Therefore, for states \( i \in S_+ \) with positive net input rates the long-run probabilities \( F_i(0) \) are zero. Thus, the first set of boundary conditions (2.8) is given by 
\[
F_i(0) = 0, \ i \in S_+.
\]
To get the second set of boundary conditions (2.9) we can again apply conditioning on a small time interval of length \( h > 0 \), for \( x = 0, \ j \in S_0 \), as above to obtain 
\[
\sum_{i=1}^{N} \int_{z=0}^{(-r_i h)^+} F_i(t-h, (-r_i h)^+ - z)q_i p_{ij} h dG_{ij}(z) + (1 - q_j h)F_j(t-h, 0) = F_j(t, 0), \ j \in S_0,
\]
where \((\cdot)^+ = \max(\cdot, 0)\). After letting \( t \to \infty \) and using the notation \( dQ_{ij}(z) \) as defined above we get 
\[
\sum_{i=1}^{N} \int_{z=0}^{(-r_i h)^+} F_i((-r_i h)^+ - z)dQ_{ij}(z) = 0, \ j \in S_0.
\]
Now, as \( h \to 0 \) the boundary conditions (2.9) are obtained 
\[
\sum_{i=1}^{N} F_i(0)Q_{ij}(0) = 0, \ j \in S_0.
\]
\[\diamondsuit\]

**Remark:** Equations (2.7)-(2.9) have a simple rate interpretation. For example, if \( r_j > 0 \), then the rate out of the set of states \( \{(y, j), 0 \leq y \leq x\} \) is equal to 
\[
r_j \frac{dF_j(x)}{dx} - F_j * Q_{jj}(x)
\]
and the rate into that set, 
\[
\sum_{i \neq j} F_i * Q_{ij}(x).
\]
Equating the two rates yields Equation (2.12).

We solve the differential equations (2.7) of the above theorem using Laplace Stieltjes Transforms (LST). Note that the LST of \( Q(x) \) is given by 
\[
\tilde{Q}_{ij}(s) = q_i p_{ij} \tilde{G}_{ij}(s), \ i, j \in S, \ i \neq j, \quad (2.13)
\]
\[
\tilde{Q}_{ii}(s) = q_i p_{ii} \tilde{G}_{ii}(s) - q_i, \ i \in S. \quad (2.14)
\]
This implies that $\tilde{Q}(0) = Q$ is the generator matrix of the CTMC $\{I(t), t \geq 0\}$. Taking the Laplace-Stieltjes transform (LST) of both sides of equation (2.7) we get
\[
\tilde{F}(s)(sR - \tilde{Q}(s)) = sF(0)R. \tag{2.15}
\]
Thus, in order to find the LST $\tilde{F}(s)$ we need to know $F(0) = [F_1(0), \ldots, F_N(0)]$. From the boundary conditions (2.8) above we have that $N_+$ components of $F(0)$ are zero and $N_0$ components can be expressed in terms of the remaining $N_-$ from the second set of boundary conditions (2.9). The next theorem is used to determine the remaining $N_- = N - N_+ - N_0$ components of $F(0)$ (cf. Theorem 5 in Loynes [6]).

**Theorem 2.2** Suppose the stability condition in Equation (2.3) is satisfied. Then the generalized eigenvalue problem
\[
(sR - \tilde{Q}(s))\phi = 0 \tag{2.16}
\]
has exactly $N_-$ solutions $(s_1, \phi_1), \ldots, (s_{N_-}, \phi_{N_-})$, with $s_1 = 0$, $\text{Re}(s_i) > 0$, $i = 2, \ldots, N_-$ and $\phi_i \neq 0$. Furthermore, these zeros $s_1, \ldots, s_{N_-}$ lie on or inside the circle in the complex plane with center at $\lambda = \max_{i: r_i \neq 0} -\frac{q_i}{r_i}$ and radius $\lambda$.

**Proof:** is given in the Appendix. \hfill \blacklozenge

Now, let
\[
M(s) := sR - \tilde{Q}(s)
\]
and $M'(s)$ be the derivative of $M(s)$. The next theorem is the main result of this paper. It summarizes the entire set of equations satisfied by $F(0)$.

**Theorem 2.3** The LST row vector $\tilde{F}(s)$ satisfies
\[
\tilde{F}(s)(sR - \tilde{Q}(s)) = sF(0)R, \tag{2.17}
\]
where the unknowns $F(0) = [F_1(0), F_2(0), \ldots, F_N(0)]$ are given by the solution to the $N$ equations
\[
F_i(0) = 0, \quad i \in S_+; \tag{2.18}
\]
\[
F(0)Q_{-,i}(0) = 0, \quad i \in S_0; \tag{2.19}
\]
\[
F(0)R\phi_i = 0, \quad \text{for } i = 2, \ldots, N_-, \tag{2.20}
\]
\[
F(0)Re = \pi(R + \Gamma)e. \tag{2.21}
\]

**Proof:** The equation for $\tilde{F}(s)$ is as derived above, Eq. (2.15). Also, recall the first two sets of equations for $F_i(0)$ when $i \in S_+$ or $i \in S_0$, given in Theorem 2.1, equations (2.8) and (2.9). Since $\tilde{F}(s)$ is analytic for $\text{Re}(s) \geq 0$, it must be the case that for every $i$,
\[
\tilde{F}(s_i)(s_iR - \tilde{Q}(s_i))\phi_i = s_iF(0)R\phi_i = 0. \tag{2.22}
\]
However, for $s_1 = 0$ the last equation is trivially satisfied and we need one additional condition to determine $F(0)$. Thus, we get equations (2.20).

To derive the normalization equation (2.21) first recall that from (2.17) we have

$$\tilde{F}(s)M(s) = sF(0)R.$$  \hspace{1cm} (2.23)

After differentiating equation (2.23) and setting $s = 0$ we get

$$\tilde{F}(0)M'(0) + \tilde{F}'(0)M(0) = F(0)R.$$  \hspace{1cm} (2.24)

Multiplying the last equation from the right by $e$ and noting that $M(0)e = -\tilde{Q}(0)e = 0$, we get

$$F(0)Re = \tilde{F}(0)M'(0)e.$$  \hspace{1cm} 

Since $\tilde{F}(0) = F(\infty) = \pi$ and $M'(0) = R + \Gamma$, we get Equation (2.21).  \hspace{1cm} $\diamond$

### 3 Examples

We illustrate the above methodology with the help of several examples.

#### 3.1 The $M|G|1$ Queue.

Consider a standard $M|G|1$ queue with arrival rate $\lambda$ and service time distribution $G$ with mean $\tau$. Let $X(t)$ be the work content in this system at time $t$. The $\{X(t), t \geq 0\}$ process decreases at rate 1 and jumps up by a random amount with cdf $G$ whenever an arrival occurs. If $X(t)$ becomes zero, it stays zero until an upward jump occurs. Thus we can model this as a fluid process with jumps with a single state environment process. The parameters are as follows: $S = \{1\}$, $q_1 = \lambda$, $p_{1,1} = 1$, $G_{1,1} = G$, $m_{1,1} = \tau$, $r_1 = -1$. Hence $\pi_1 = 1$, and the condition of stability (2.3) reduces to $-1 + \lambda \tau < 0$, or $\rho = \lambda \tau < 1$. This is the standard condition of stability for the $M|G|1$ queue. We get $\tilde{Q}(s) = \lambda(\tilde{G}(s) - 1)$, and Equation (2.17) reduces to the scalar equation

$$\tilde{F}(s)(s - \lambda(1 - \tilde{G}(s))) = sF(0).$$

There is only one boundary condition, given by Equation (2.21). It reduces to $F(0) = 1 - \lambda \tau = 1 - \rho$. Using that we get

$$\tilde{F}(s) = \frac{s(1 - \rho)}{s - \lambda(1 - \tilde{G}(s))}.$$  

This matches the LST of the queueing time in an $M|G|1$ queue, as it must.


3.2 The $BMAP|G|1$ Queue.

Consider a single server queue whose arrival process is given by a $BMAP$ defined by the sequence $\{D_k, k \geq 0\}$. Here $D_0$ has negative diagonal elements and nonnegative off-diagonal elements, $D_k$, $k \geq 1$, are nonnegative, and $D$, defined as

$$D = \sum_{k=0}^{\infty} D_k,$$

is the generator an irreducible CTMC on state space $S = \{1, 2, ..., N\}$. Transitions according to $D_k$ correspond to batch arrivals of size $k$. The service times are iid random variables with cdf $G$ and mean $\tau$. Let $I(t)$ be the phase of the arrival process at time $t$. It is a CTMC on $S$ with generator matrix $D$. Let $\pi$ be the limiting distribution of $I(t)$.

Let $X(t)$ be the work content in the system at time $t$. The $\{X(t), t \geq 0\}$ process decreases at rate 1 and jumps up by a random amount with cdf $G^*k$ whenever a batch arrival of size $k$ occurs. If $X(t)$ becomes zero, it stays zero until an upward jump occurs. Thus we can model the work content process as a fluid process with jumps with the following parameters: The state space of the environment process (the phase of the arrival process) is $S = \{1, 2, ..., N\}$.

We get

$$\tilde{Q}(s) = \sum_{k=0}^{\infty} D_k \tilde{G}^k(s) = D(\tilde{G}(s)),$$

where

$$D(z) = \sum_{k=0}^{\infty} D_k z^k, \quad |z| \leq 1.$$

We also have $R = -I$, since the net rate is $-1$ in every state. The condition of stability is

$$\pi \sum_{k=1}^{\infty} k D_k e \tau < 1.$$ 

Since $S_- = S$, Theorem 2.2 implies that there are $N$ (eigenvalue, eigenvector) pairs $(s_i, \phi_i)$ ($i = 1, 2, ..., N$) satisfying

$$(s I + D(\tilde{G}(s))) \phi = 0,$$

with $s_1 = 0$ and $Re(s_i) > 0$ for $i = 2, 3, ..., N$. Finally, Theorem 2.3 yields the following result for the LST of $F$:

$$\tilde{F}(s) = s F(0) (s I + D(\tilde{G}(s)))^{-1},$$

(3.25)

where the unknown vector $F(0)$ is obtained from the $N$ equations:

$$F(0) \phi_i = 0, \quad i = 2, 3, ..., N.$$

$$F(0) e = 1 - \pi \sum_{k=1}^{\infty} k D_k e \tau.$$ 

This matches with the known results; see Equation (44) in Lucantoni [7].
3.3 The $PH|G|1$ Queue.

The $PH$ arrival process with parameters $(\alpha, T)$ is a special case of the $BMAP$. Here $D_0 = T$, $D_1 = -Te\alpha$ and $D_k = 0$ for all $k > 1$, where $T = [t_{i,j}]_{i,j=1,2,...,N}$ is an invertible generator matrix with non-positive row sums of an irreducible CTMC on state space $S = \{1, 2, ..., N\}$, and $\alpha = [\alpha_1, \alpha_2, ..., \alpha_N]$ is a non-negative vector whose elements add up to one. Equation (3.25) for the LST of $F$ now simplifies to

$$\tilde{F}(s) = sF(0)(sI + T - Te\alpha\hat{G}(s))^{-1},$$

which matches with Equation (5.2.4), page 257 of Neuts [10].

3.4 The $M|G|1$ Queue with an Up-Down Server.

Consider the queueing system of the first example, with the following modification: The server can be up or down. The up times are $\text{Exp}(\beta)$ random variables, and the down times are $\text{Exp}(\alpha)$ random variables. Let $I(t)$ be the state of the server (1 if up, and 0 if down) at time $t$. Thus $\{I(t), t \geq 0\}$ is a CTMC. Let $X(t)$ be the work content at time $t$ in this system. We model $\{(X(t), I(t)), t \geq 0\}$ as a fluid queue with jumps with the following parameters:

$$q_0 = \lambda + \alpha, \quad q_1 = \lambda + \beta$$

$$p_{01} = \frac{\alpha}{\alpha + \lambda}, \quad p_{00} = \frac{\lambda}{\alpha + \lambda}, \quad p_{10} = \frac{\beta}{\lambda + \beta}, \quad p_{11} = \frac{\lambda}{\lambda + \beta}$$

so that

$$Q(x) = \begin{bmatrix} -(\alpha + \lambda(1 - G(x))) & \alpha \\ \beta & -(\beta + \lambda(1 - G(x))) \end{bmatrix},$$

and

$$R = \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$ 

This implies that $S_0 = \{0\}$ and $S_- = \{1\}$. The stability condition is given by

$$\rho = \lambda \tau < \frac{\alpha}{\alpha + \beta} = \pi_1.$$ 

We have

$$\tilde{Q}(s) = \begin{bmatrix} -(\alpha + \lambda(1 - \hat{G}(s))) & \alpha \\ \beta & -(\beta + \lambda(1 - \hat{G}(s))) \end{bmatrix}. $$

Theorem 2.3 yields the following result for the LST of $F(x) = [F_0(x), F_1(x)]$:

$$\tilde{F}(s) = sF(0)(sR - \tilde{Q}(s))^{-1}.$$
The boundary conditions yield the following explicit expression for $F(0) = [F_0(0), F_1(0)]:$

$$F_0(0) = \frac{\beta}{\alpha + \lambda} \left( \frac{\alpha}{\alpha + \beta} - \rho \right), \quad F_1(0) = \frac{\alpha}{\alpha + \beta} - \rho.$$ 

This can be simplified to get

$$\tilde{F}_0(s) = \frac{s \beta (\pi_1 - \rho)}{\alpha \beta - (\alpha + \lambda (1 - \tilde{G}(s))) (\beta + \lambda (1 - \tilde{G}(s)) - s)}, \quad \tilde{F}_1(s) = \frac{s (\pi_1 - \rho) (\alpha + \lambda (1 - \tilde{G}(s)))}{\alpha \beta - (\alpha + \lambda (1 - \tilde{G}(s))) (\beta + \lambda (1 - \tilde{G}(s)) - s)}.$$ 

This matches the results in Gaver [3]. As a special case, when $\beta = 0$, the server is permanently up, and we get the standard queue of the first example. The results match for this special case.

4 Appendix

Proof of Theorem 2.2: We follow the method used in a similar proof in [1].

Note that it follows from Equations (2.13) and (2.14) that we can rewrite the matrix $\tilde{Q}(s)$ in the form

$$\tilde{Q}(s) = \tilde{Q}(s) - Q_d,$$

where

$$\tilde{Q}(s) := [q_{ij} \tilde{G}_{ij}(s)], \quad Q_d := \text{diag}[q_1, \ldots, q_N]. \quad (4.26)$$

Then equation (2.16) has a solution with $\phi \neq 0$ if

$$\det(\tilde{Q}(s) - Q_d - sR) = 0, \quad (4.27)$$

which is equivalent to

$$\det(Q_d^{-1} \tilde{Q}(s) - sQ_d^{-1} R - I) = 0. \quad (4.28)$$

We first assume that for some $\epsilon > 0$ the transformations $\tilde{G}_{ij}(s)$ are analytic for all $s$ with $\text{Re}(s) > -\epsilon$. This holds for distributions with an exponential tail or distributions with finite support.

Let

$$\lambda := \max_{i: r_i \neq 0} \left( -\frac{q_i}{r_i} \right), \quad (4.29)$$

and let

$$C_\delta := \{s : |s - \lambda| = \lambda + \delta\}, \quad (4.30)$$
denote the circle with center at \( \lambda \) and radius \( \lambda + \delta \), where \( 0 < \delta < \epsilon \).

Next, we show that for \( 0 \leq u \leq 1 \) and small \( \delta > 0 \),

\[
\det(uQ_d^{-1}\tilde{Q}(s) - sQ_d^{-1}R - I) \neq 0, \quad s \in C_{\delta}.
\]  

(4.31)

First, in the case \( \text{Re}(s) \geq 0 \) we prove that the matrix \( uQ_d^{-1}\tilde{Q}(s) - sQ_d^{-1}R - I \) is diagonally dominant (with strict dominance in at least one row). This, plus the fact that \( P = |p_{ij}| \) is irreducible, imply (see [13]) that it has a non-zero determinant.

1. Let \( i : r_i \neq 0 \) and denote by \( \lambda_i := -\frac{q_i}{r_i} \) so that \( \lambda := \max_{i : r_i \neq 0} \lambda_i \). Clearly,

\[
|up_i\tilde{G}_{ii}(s) - \frac{sr_i}{q_i} - 1| = |up_i\tilde{G}_{ii}(s) + \frac{s}{\lambda_i} - 1| \geq \left| \frac{s}{\lambda_i} - 1 \right| - up_i\tilde{G}_{ii}(0).
\]

- \( i : r_i < 0 \) or equivalently \( i : \lambda_i > 0 \)

Then we have

\[
\left| \frac{s}{\lambda_i} - 1 \right| = \left| \frac{s - \lambda_i}{\lambda_i} \right| = \frac{|s - \lambda|}{\lambda_i} \geq \frac{|\lambda + \delta|}{\lambda_i} = \frac{|\lambda_i + \delta|}{\lambda_i} = 1 + \frac{\delta}{\lambda_i}.
\]

and therefore

\[
|up_i\tilde{G}_{ii}(s) - \frac{sr_i}{q_i} - 1| \geq 1 + \frac{\delta}{\lambda_i} - up_i\tilde{G}_{ii}(0) > u\tilde{G}_{ii}(0) - up_i\tilde{G}_{ii}(0) = \sum_{j \neq i} up_{ij}
\]

\[
\geq \sum_{j \neq i} |up_{ij}\tilde{G}_{ij}(s)|.
\]

- \( i : r_i > 0 \) or equivalently \( i : \lambda_i < 0 \)

Then it is clear that for \( \text{Re}(s) \geq 0 \) and \( s \in C_{\delta} \)

\[
\left| \frac{s}{\lambda_i} - 1 \right| > 1,
\]

and therefore

\[
|up_i\tilde{G}_{ii}(s) - \frac{sr_i}{q_i} - 1| \geq 1 - up_i\tilde{G}_{ii}(0) \geq u\tilde{G}_{ii}(0) - up_i\tilde{G}_{ii}(0) = \sum_{j \neq i} up_{ij}
\]

\[
\geq \sum_{j \neq i} |up_{ij}\tilde{G}_{ij}(s)|.
\]

2. Let \( i : r_i = 0 \).

Then

\[
|up_i\tilde{G}_{ii}(s) - 1| \geq 1 - up_i\tilde{G}_{ii}(0) \geq u\tilde{G}_{ii}(0) - up_i\tilde{G}_{ii}(0) = \sum_{j \neq i} up_{ij} \geq \sum_{j \neq i} |up_{ij}\tilde{G}_{ij}(s)|.
\]

(4.34)
Next, consider the case $s \in C_\delta$ with $Re(s) < 0$. Note that the determinant of (4.31) is nonzero if and only if $0$ is not an eigenvalue. Therefore, next we study the eigenvalues of $uQ^{-1}_d \hat{Q}(s) - sQ^{-1}_d R - I$ in a neighborhood of $s = 0$. If we write

$$uQ^{-1}_d \hat{Q}(s) - sQ^{-1}_d R - I = P - I - sQ^{-1}_d R + (u - 1)Q^{-1}_d \hat{Q}(s) + Q^{-1}_d \hat{Q}(s) - P,$$

we can see that the above matrix is a perturbation of $P - I$ when $(s, u)$ is close to $(0, 1)$. Since $P$ is assumed to be irreducible, $P - I$ has a simple eigenvalue $0$. Thus, in a neighborhood of $(0, 1)$, there exist differentiable $x(s, u)$ and $\mu(s, u)$ such that

$$(uQ^{-1}_d \hat{Q}(s) - sQ^{-1}_d R - I)x(s, u) = \mu(s, u)x(s, u), \; \mu(0, 1) = 0, \; x(0, 1) = e$$

(4.36)

Differentiating the last equation with respect to $s$ and setting $s = 0$, $u = 1$ in the result gives the following:

$$\left( Q^{-1}_d \frac{d\hat{Q}(0)}{ds} - Q^{-1}_d R \right) e + (P - I) \frac{\partial}{\partial s} x(0, 1) = \frac{\partial}{\partial s} \mu(0, 1)e,$$

(4.37)

After multiplying this equation from the left by $\pi Q_d$, where $\pi := [\pi_1, \ldots, \pi_N]$, $\pi_i = \lim_{t \to \infty} P(I(t) = i)$ we get

$$\pi \left( \frac{d\hat{Q}(0)}{ds} - R \right) e + \pi Q_d(P - I) \frac{\partial}{\partial s} x(0, 1) = \frac{\partial}{\partial s} \mu(0, 1)\pi Q_d e.$$

(4.38)

Note that $\pi$ is the solution to

$$\pi Q_d(P - I) = 0,$$

(4.39)

and also from the definition of $\hat{Q}(s)$ above we have

$$\frac{d\hat{Q}(s)}{ds} = \left[q_ip_{ij} \frac{d\tilde{G}_{ij}(s)}{ds}\right],$$

(4.40)

and therefore

$$\frac{d\hat{Q}(0)}{ds} = [-q_ip_{ij}m_{ij}] = -Q_d P \cdot M,$$

(4.41)

where $P \cdot M$ denotes the Hadamard matrix multiplication. Hence we have

$$\pi (-Q_d P \cdot M - R) e = \frac{\partial}{\partial s} \mu(0, 1)\pi Q_d e.$$

(4.42)

In a similar way differentiation with respect to $u$ leads to the following:

$$\pi Q_d Pe = \frac{\partial}{\partial u} \mu(0, 1)\pi Q_d e,$$

(4.43)

and since

$$\pi Q_d Pe = \sum_{i=1}^{N} \sum_{j=1}^{N} \pi_i q_i p_{ij} = \sum_{j=1}^{N} \pi_j q_j = \pi Q_d e,$$

(4.44)

12
we have
\[ \frac{\partial}{\partial u} \mu(0, 1) = 1. \]  
(4.45)

Therefore, in a neighbourhood of \((0, 1)\) from the last equation and Eq. (4.42) it follows
\[ \mu(s, u) \approx s \frac{\pi (-Q_d P \cdot M - R) e}{\pi Q_d e} + u - 1. \]  
(4.46)

From the stability condition (2.3) we have
\[ \pi (-Q_d P \cdot M - R) e > 0, \]  
(4.47)

so we can conclude that \(\mu(s, u) \neq 0\) for small \(\delta\), \(s \in C_\delta\) with \(\text{Re}(s) < 0\), and \(u\) close to 1, say
\[ 1 - \delta \leq u \leq 1. \]

For \(0 \leq u < 1 - \delta\), it can be shown that \(uQ_d^{-1} \bar{Q}(s) - sQ_d^{-1} R - I\) is diagonally dominant (with strict dominance in at least one row) for \(s \in C_\delta\) with \(\text{Re}(s) < 0\), provided that \(\delta\) is taken small enough so that
\[ 1 - \delta \lambda_i > (1 - \delta) \max_{i,j \in \tilde{S}} \tilde{G}_{ij}(-\delta). \]  
(4.48)

For \(i : r_i < 0\) we have as before
\[ |up_{ii} \tilde{G}_{ii}(s) - \frac{sr_i}{q_i} - 1| \geq 1 + \frac{\delta}{\lambda_i} - up_{ii} \tilde{G}_{ii}(s) > 1 + \frac{\delta}{\lambda_i} - up_{ii} \tilde{G}_{ii}(s) > (1 - \delta) \max_{i,j \in \tilde{S}} \tilde{G}_{ij}(-\delta) - up_{ii} \tilde{G}_{ii}(s) \]
\[ > u \max_{i,j \in \tilde{S}} \tilde{G}_{ij}(-\delta) - up_{ii} \max_{i,j \in \tilde{S}} \tilde{G}_{ij}(-\delta) \geq \sum_{j \neq i} |up_{ij} \tilde{G}_{ij}(s)|. \]

For \(i : r_i \geq 0\) it is obvious that
\[ |up_{ii} \tilde{G}_{ii}(s) - \frac{sr_i}{q_i} - 1| \geq 1 - up_{ii} \tilde{G}_{ii}(s) > 1 - \frac{\delta}{\lambda_i} - up_{ii} \tilde{G}_{ii}(s), \]  
(4.49)

and the rest of the argument is the same as above. This completes the proof of Inequality (4.31).

Let \(f(u)\) denote the number of zeros of \(\det(uQ_d^{-1} \bar{Q}(s) - sQ_d^{-1} R - I)\) inside the circle \(C_\delta\). Then from the Cauchy Theorem of Complex Analysis we have
\[ f(u) = \frac{1}{2\pi i} \int_{C_\delta} \frac{\partial}{\partial s} \det(uQ_d^{-1} \bar{Q}(s) - sQ_d^{-1} R - I) ds. \]  
(4.50)

It is clear that \(f(0) = N_-\), since
\[ \det(-sQ_d^{-1} R - I) = -\prod_{i=1}^{N} \left( \frac{sr_i}{q_i} + 1 \right), \]  
(4.51)
and therefore the zeros are \( s_i = -\frac{q_i}{r_i} \), \( r_i \neq 0 \) and exactly \( N_\ast \) of them have \( \text{Re}(s) \geq 0 \) and thus are inside the circle \( C_\delta \). Since \( f(u) \) is an integer-valued continuous function on \([0, 1]\) it follows that it is constant. Hence, \( f(1) = N_\ast \). As \( \delta \to 0 \), we can conclude that \( \det(sR - \tilde{Q}(s)) \) has exactly \( N_\ast \) zeros inside or on \( C_0 \).

It is clear that \( s = 0 \) satisfies
\[
\det(sR - \tilde{Q}(s)) = 0.
\]

It can be shown that \( s = 0 \) is a simple root of this equation using the irreducibility of \( P \) and the stability condition (2.3). The arguments are the same as in [1] and therefore we skip them.

To finally complete the proof of Theorem 2.2 we have to remove the initial assumption that for some \( \epsilon > 0 \) the transforms \( \tilde{G}_{ij}(s) \) are analytic for all \( s \) with \( \text{Re}(s) > -\epsilon \). To this end, first consider the truncated distributions \( G^K_{ij}(x) \) defined as \( G^K_{ij}(x) = G_{ij}(x) \) for \( 0 \leq x < K \) and \( G^K_{ij}(x) = 1 \) for \( x \geq K \). Then Theorem 2.2 holds for the distributions \( G^K_{ij}(x) \); by letting \( K \) tend to infinity, the result also follows for the original distributions.

**References**


