SIAM J. MATRIX ANAL. APPL.
Vol. 24, No. 1, pp. 281–291

QUASI-BIRTH-AND-DEATH PROCESSES WITH LEVEL-GEOMETRIC DISTRIBUTION*

TUĞRUL DAYAR† AND FRANCK QUESSETTE‡

Abstract. A special class of homogeneous continuous-time quasi-birth-and-death (QBD) Markov chains (MCs) which possess level-geometric (LG) stationary distribution is considered. Assuming that the stationary vector is partitioned by levels into subvectors, in an LG distribution all stationary subvectors beyond a finite level number are multiples of each other. Specifically, each pair of stationary subvectors that belong to consecutive levels is related by the same scalar, hence the term level-geometric. Necessary and sufficient conditions are specified for the existence of such a distribution, and the results are elaborated in three examples.

Key words. Markov chains, quasi-birth-and-death processes, geometric distributions

AMS subject classifications. 60J27, 65F50, 65H10, 65F05, 65F10, 65F15

PII. S089547980138914X

1. Introduction.

The continuous-time Markov process on the countable state space $\mathcal{S} = \{(l, i) \ : \ l \geq 0, \ 1 \leq i \leq m\}$ with block tridiagonal infinitesimal generator matrix

$$Q = \begin{pmatrix} B_0 & A_0 & & & \\ A_2 & A_1 & A_0 & & \\ & A_2 & A_1 & A_0 & \\ & & \ddots & \ddots & \ddots \end{pmatrix} \tag{1}$$

having blocks that are $(m \times m)$ matrices is called a homogeneous continuous-time quasi-birth-and-death (QBD) Markov chain (MC). The row sums of Q are zero, meaning $(B_0 + A_0)e = 0$ and $(A_0 + A_1 + A_2)e = 0$, where e is a column vector of 1's with appropriate length. The matrices A_0 and A_2 are nonnegative, and the matrices B_0 and A_1 have nonnegative off-diagonal elements and strictly negative diagonals. The first component, l, of the state descriptor vector denotes the level and its second component, i, the phase. In homogeneous QBD MCs, the elements of B_0, A_0, A_1, and A_2 do not depend on the level number.

Neuts has done substantial work in the area of matrix analytic methods for such processes and has written two books [11], [12]. An informative resource that discusses the developments in the area since then is the recent book of Latouche and Ramaswami [9]. The most significant application area of these methods at present is the performance evaluation of communication systems. See, for instance, [13] for several case studies covering application areas from asynchronous transfer mode (ATM) networks to World Wide Web traffic and Transmission Control Protocol/Internet Protocol (TCP/IP) networking.

We assume that the homogeneous continuous-time QBD MC at hand is irreducible and positive recurrent, meaning its steady state probability distribution vector, π

*Received by the editors May 10, 2001; accepted for publication (in revised form) February 6, 2002; published electronically July 9, 2002.
http://www.siam.org/journals/simax/24-1/38914.html

†Department of Computer Engineering, Bilkent University, 06533 Bilkent, Ankara, Turkey (tugrul@cs.bilkent.edu.tr).
‡Lab. PRiSM, Université de Versailles, 45 Avenue des États-Unis, 78035 Versailles Cedex, France (qst@prism.uvsq.fr).

(see [14]), exists. Recall that an MC is said to be positive recurrent if the mean time to return to each state for the first time after leaving it is finite [14, p. 9]. In infinite QBD MCs, this requires that the drift to higher level states be smaller than the drift to lower level states [5, pp. 153–154]. Throughout the paper, we adhere to the convention that probability vectors are row vectors. Being a stationary distribution, π satisfies $\pi Q = 0$ and $\pi e = 1$. Now, let π be partitioned by levels into subvectors π_l, $l \geq 0$, where π_l is of length m. Then π also satisfies the matrix-geometric property [9, p. 142]

$$\pi_{l+1} = \pi_l R \quad \text{for} \quad l \geq 0, \tag{2}$$

where the matrix R of order m records the rate of visit to level $(l+1)$ per unit of time spent in level l. Fortunately, the elements of R for homogeneous QBD MCs do not depend on the level number. Quadratically convergent algorithms for solving QBD MCs appear in [8], [4], [1].

In this paper, we consider a special class of homogeneous continuous-time QBD MCs which possess what we call level-geometric (LG) stationary distribution. To the best of our knowledge, this property has not been explicitly defined before, and hence our "level-geometric" designation. An LG distribution is one that satisfies

$$\pi_{l+1} = \alpha \pi_l \quad \text{for} \quad l \geq L, \tag{3}$$

where $\alpha \in (0,1)$ and L is a finite nonnegative integer. Note that an LG distribution with $L = 0$ is a product-form solution. An LG distribution can be expressed alternatively as

$$\pi_{L+k} = (1-\alpha)\alpha^k a \quad \text{for} \quad k \geq 0, \tag{4}$$

where a is a positive probability vector of length m, with $ae = 1$ when $L = 0$. In an LG distribution, the level is independent of the phase for level numbers greater than or equal to L, and the marginal probability distribution of the levels are given by $\pi_{L+k}e = (1-\alpha)\alpha^k ae$ [9, pp. 295–299] for $k \geq 0$. Throughout the paper, we refer to an LG distribution for which L is the smallest possible nonnegative integer that satisfies (3) as an LG distribution with parameter L. Our motivation is to come up with a solution method for this special class of QBD MCs that does not require R to be computed. We remark that if S_ϵ is the number of iterations required to reach an accuracy of ϵ by the successive substitution algorithm [5, p. 160], then the computation of R with quadratically convergent algorithms takes about $O(\log_2 S_\epsilon)$ iterations (hence, the term quadratically convergent), each of which has a time complexity of $O(m^3)$ floating-point operations. The results that we develop can be extended to the homogeneous discrete-time case without difficulty.

In section 2, we provide background information on the solution of QBD MCs with special structure. In section 3, we give three examples of QBD MCs with LG stationary distribution. In section 4, we specify conditions related to such a distribution and show how it can be computed when it exists. In section 5, we reconsider the three examples of section 3 in light of the new results introduced in section 4. We conclude in section 6.

2. Background material. In this section, an overview of some concepts discussed in [9] and relevant propositions are given. Wherever something has been taken from [9], the appropriate reference to the corresponding page(s) is placed.

Due to the fixed pattern of transitions among levels and within each level, it is not difficult to check the irreducibility of Q. The next proposition is about checking the positive recurrence of Q when Q and $A = A_0 + A_1 + A_2$ are both irreducible. When Q is irreducible but A has multiple irreducible classes, one can resort to the theorem in [9, p. 160]. Note that A is an infinitesimal generator matrix.

PROPOSITION 1. *If Q and A are irreducible, then Q is positive recurrent if and only if $\pi_A(A_0 - A_2)e < 0$, where π_A satisfies $\pi_A A = 0$ and $\pi_A e = 1$* [9, p. 158].

Throughout this paper, we assume that the homogeneous continuous-time QBD MC at hand is irreducible and positive recurrent. Now, let $\rho(R)$ denote the spectral radius of R (i.e., $\rho(R) = \max\{|\lambda| \mid \lambda \in \lambda(R)\}$, where $\lambda(R) = \{\lambda \mid Rv = \lambda v, v \neq 0\}$ is its spectrum). Then, $\rho(R) < 1$ [9, p. 133].

The next proposition specifies necessary and sufficient conditions for the existence of an LG distribution with parameter $L = 0$.

PROPOSITION 2. *The stationary distribution of Q is LG with parameter $L = 0$ if and only if there exists a positive vector a with $ae = 1$ and a positive scalar $\alpha = \rho(R)$ with $\alpha < 1$ such that $a(A_0 + \alpha A_1 + \alpha^2 A_2) = 0$ and $a(B_0 + \alpha A_2) = 0$* [9, pp. 297–298].

This proposition, although very concise and to the point, has two shortcomings. First, it does not indicate how to check for an LG distribution with parameter $L \geq 1$. Second, it requires the solution of a nonlinear system of equations.

The following two propositions indicate the improvement that is obtained in the solution when A_2 and/or A_0 are rank-1 matrices.

PROPOSITION 3. *When A_2 is of rank-1, then $R = -A_0(A_1 + A_0 e b^T)^{-1}$, where $A_2 = cb^T$ and $b^T e = 1$* [9, p. 197]. *Furthermore, π_0 can be computed up to a multiplicative constant using $\pi_0(B_0 + A_0 e b^T) = 0$* [9, p. 236].

Hence, it is relatively simple to compute the stationary distribution when A_2 is of rank-1.

PROPOSITION 4. *When A_0 is of rank-1, then $R = c\xi^T$, where $A_0 = cb^T$, $b^T e = 1$, $\xi^T = -b^T(A_1 + \alpha A_2)^{-1}$, and $\alpha = \xi^T c$ with $\alpha = \rho(R)$* [9, p. 198]. *The stationary subvectors satisfy $\pi_0 = \pi_1 C_0$, where $C_0 = -A_2 B_0^{-1}$, and $\pi_l = \pi_{l+1} C_1$ for $l \geq 1$, where $C_1 = -A_2(A_1 + A_2 e b^T)^{-1}$* [9, p. 236].

COROLLARY 1. *When A_0 is of rank-1, then R is also of rank-1, and $R^2 = \alpha R$ thereby implies $\pi_{l+1} = \alpha \pi_l$ for $l \geq 1$. Hence, Q has an LG distribution with parameter $L \leq 1$.*

The next section elaborates these results with three examples.

3. Examples.

The following examples all have LG distributions, and they aid in understanding the concepts introduced in section 2 and the concepts to be developed in section 4. In order to compactly describe single queueing stations, we use the so-called Kendall notation, which consists of six identifiers separated by vertical bars [5, pp. 13–14]:

$$\text{Arrivals}|\text{Services}|\text{Servers}|\text{Buffersize}|\text{Population}|\text{Scheduling}.$$

Here Arrivals and Services, respectively, characterize the customer arrival and service processes by specifying the interarrival and interservice distributions. For these distributions there are various possibilities, among which are M (i.e., Markovian) for exponential and E_k for k-phase Erlang. Servers gives the number of service-providing entities; Buffersize gives the maximum number of customers in the queueing station, including any in service; Population gives the size of the customer population from which the arrivals are taking place; and Scheduling specifies the employed scheduling strategy. When the Buffersize and/or the Population are omitted, they are assumed to be infinitely large. When the scheduling strategy is omitted, it is assumed to be first come, first served (FCFS).

3.1. Example 1. The first example we consider is a system of two independent queues, where queue 1 is M|M|1 and queue 2 is M|M|1|$m-1$. Queue $i \in \{1,2\}$ has a Poisson arrival process with rate λ_i and an exponential service distribution with rate μ_i. This system corresponds to a QBD process with the level representing the length of queue 1, which is unbounded, and the phase representing the length of queue 2, which can range between 0 and $(m-1)$. We assume $\lambda_1 < \mu_1$. Letting $d = \lambda_1 + \lambda_2 + \mu_1 + \mu_2$, we have $A_0 = \lambda_1 I$, $A_2 = \mu_1 I$,

$$A_1 = \begin{pmatrix} -(d-\mu_2) & \lambda_2 & & & \\ \mu_2 & -d & \lambda_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \mu_2 & -d & \lambda_2 \\ & & & \mu_2 & -(d-\lambda_2) \end{pmatrix},$$

and

$$B_0 = \begin{pmatrix} -(\lambda_1+\lambda_2) & \lambda_2 & & & \\ \mu_2 & -(d-\mu_1) & \lambda_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \mu_2 & -(d-\mu_1) & \lambda_2 \\ & & & \mu_2 & -(\lambda_1+\mu_2) \end{pmatrix}.$$

Q is irreducible, and from Proposition 1 we have

$$A = A_0 + A_1 + A_2 = \begin{pmatrix} -\lambda_2 & \lambda_2 & & & \\ \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 \\ & & & \mu_2 & -\mu_2 \end{pmatrix},$$

which is irreducible, and π_A is the truncated geometric distribution with parameter λ_2/μ_2 [5, p. 84]. Hence, $\pi_A(A_0 - A_2)e = \lambda_1 - \mu_1 < 0$ and Q is positive recurrent. For this example, $\alpha = \lambda_1/\mu_1$, $a_k = \nu^k(1-\nu)/(1-\nu^m)$, $0 \le k \le m-1$, and $L = 0$, where $\nu = \lambda_2/\mu_2$, turn out to be the parameters in (4) that specify an LG distribution.

Recalling that an MC is said to be lumpable with respect to a given partitioning if each block in the partitioning has equal row sums [7, p. 124], we remark that the QBD MC in this example is lumpable, and the lumped chain represents queue 1.

3.2. Example 2. The second example we consider is the continuous-time equivalent of the discrete-time QBD process discussed in [8, pp. 668–669]. The model has 2 phases at each level (i.e., $m = 2$). Assuming that $0 < p < 1$, the process moves from state $(l,1)$, $l \ge 1$, to $(l,2)$ with rate p, and to $(l-1,1)$ with rate $(1-p)$. The process moves from state $(l,2)$, $l \ge 0$, to $(l,1)$ with rate $2p$, and to $(l+1,2)$ with rate $(1-2p)$. Finally, the process moves from state $(0,1)$ to $(0,2)$ with rate 1. All diagonal elements of Q are -1. Hence, we have

$$A_0 = \begin{pmatrix} 0 & 0 \\ 0 & 1-2p \end{pmatrix},\ A_1 = \begin{pmatrix} -1 & p \\ 2p & -1 \end{pmatrix},\ A_2 = \begin{pmatrix} 1-p & 0 \\ 0 & 0 \end{pmatrix},\ B_0 = \begin{pmatrix} -1 & 1 \\ 2p & -1 \end{pmatrix}.$$

Q is irreducible, and from Proposition 1 we have

$$A = A_0 + A_1 + A_2 = \begin{pmatrix} -p & p \\ 2p & -2p \end{pmatrix},$$

which is irreducible, and $\pi_A = (2/3\ 1/3)$. Hence, $\pi_A(A_0 - A_2)e = -1/3 < 0$ and Q is positive recurrent. For this example, $\alpha = (1-2p)/(1-p)$, $a = (1/2\ 1/2)$, and $L = 0$ turn out to be the parameters in (4) that specify an LG distribution. Direct substitution in $\pi Q = 0$ and $\pi e = 1$ confirms this solution.

In this example, Proposition 3 applies with $c = (1-p)e_1$ and $b = e_1$, where e_i is the ith principal axis vector. Hence, $R = (1-2p)e_2^T e/(1-p)$, and $\rho(R) = \alpha$ as expected. Furthermore, $\pi_0 = (1-\alpha)(1/2\ 1/2)$. Note that in this example, Proposition 4 applies as well. The rate matrix is of rank-1 and $\xi = e/(1-p)$. In section 5, we will argue why this example has an LG distribution with parameter $L = 0$ and not $L = 1$. Finally, we remark that this example is also used as a test case in [1].

3.3. Example 3. The third example we consider is the $\mathrm{E}_m|\mathrm{M}|1$ FCFS queue which has an exponential service distribution with rate μ and an m-phase Erlang arrival process with rate $m\lambda$ in each phase [9, pp. 206–208]. The expected interarrival time and the expected service time of this queue are, respectively, $1/\lambda$ and $1/\mu$. We assume $\lambda < \mu$. The queue corresponds to a QBD process with the level representing the queue length (including any in service) and the phase representing the state of the Erlang arrival process. Letting $d = m\lambda + \mu$, we have the $(m \times m)$ matrices $A_0 = m\lambda e_m e_1^T$, $A_2 = \mu I$,

$$A_1 = \begin{pmatrix} -d & m\lambda & & \\ & \ddots & \ddots & \\ & & -d & m\lambda \\ & & & -d \end{pmatrix}, \quad B_0 = \begin{pmatrix} -m\lambda & m\lambda & & \\ & \ddots & \ddots & \\ & & -m\lambda & m\lambda \\ & & & -m\lambda \end{pmatrix}.$$

Q is irreducible, and from Proposition 1 we have

$$A = A_0 + A_1 + A_2 = \begin{pmatrix} -m\lambda & m\lambda & & \\ & \ddots & \ddots & \\ & & -m\lambda & m\lambda \\ m\lambda & & & -m\lambda \end{pmatrix},$$

which is irreducible, and $\pi_A = e^T/m$. Hence, $\pi_A(A_0 - A_2)e = \lambda - \mu < 0$ and Q is positive recurrent. Although the $\mathrm{E}_m|\mathrm{M}|1$ queue does not have an explicit solution, it can be shown by following the formulae in [6, p. 323] that its stationary distribution has an LG distribution with parameter $L = 1$.

In this example, Proposition 4 applies with $c = m\lambda e_m$ and $b = e_1$, implying R is of rank-1, $C_0 = -A_2 B_0^{-1}$, and $C_1 = -A_2(A_1 + \mu e e_1^T)^{-1}$.

The next section builds on the results in section 2 with the aim of coming up with a solution method to compute an LG distribution when it exists.

4. Checking for and computing the LG distribution. The assumption of irreducibility of Q implies that the nonnegative matrix A_0 has at least one positive row sum (see (1)). Since we also have $(B_0 + A_0)e = 0$, it must be that B_0 has nonpositive row sums with at least one negative row sum. Together with the fact that B_0 has nonnegative off-diagonal elements and a strictly negative diagonal, this implies that $-B_0$ is a nonsingular M-matrix and $-B_0^{-1} \geq 0$; see [3].

The next proposition is essential in formulating the results in this section.

PROPOSITION 5. *The sequence of matrices $D_{l+1} = A_1 - A_2 D_l^{-1} A_0$, $l \geq 0$, where $D_0 = B_0$, is well defined. For $l \geq 0$, $-D_l$ is a nonsingular M-matrix, $-D_l^{-1} \geq 0$, and D_l^T denotes the diagonal block at level l after l steps of block Gaussian elimination (GE) on Q^T. Furthermore, $\pi_l = \pi_{l+1} C_l$, where $C_l = -A_2 D_l^{-1} \geq 0$ for $l \geq 0$.*

Proof. Since $-D_0$ is a nonsingular M-matrix, let us show that $-D_1$ is too. It is possible to construct the infinitesimal generator

$$\bar{Q} = \begin{pmatrix} D_0 & A_0 & 0 \\ A_2 & A_1 & s \\ 0 & r^T & \delta \end{pmatrix}$$

so that it is irreducible. Here $s = A_0 e$, r is any nonnegative vector that ensures the irreducibility of $\bar{Q}$, and $\delta = -r^T e$. Now let $X = -\bar{Q}$ and consider the partitioning

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} = \left(\begin{array}{c|cc} -D_0 & -A_0 & 0 \\ \hline -A_2 & -A_1 & -s \\ 0 & -r^T & -\delta \end{array} \right).$$

The negated infinitesimal generator X is an irreducible singular M-matrix [3] by its definition. Therefore, the Schur complement [10, p. 123] S of X_{11}, which is given by

$$S = X_{22} - X_{21} X_{11}^{-1} X_{12} = \begin{pmatrix} -A_1 + A_2 D_0^{-1} A_0 & -s \\ -r^T & -\delta \end{pmatrix},$$

is an irreducible singular M-matrix (see Lemma 1 in [2]). All principal submatrices of an irreducible singular M-matrix except itself are nonsingular M-matrices [3, p. 156]. Hence, $-A_1 + A_2 D_0^{-1} A_0$; that is, $-D_1$ is a nonsingular M-matrix and $-D_1^{-1} \geq 0$. One can similarly show that $-D_l$ is a nonsingular M-matrix and $-D_l^{-1} \geq 0$ for $l > 1$.

Since Q^T is a block tridiagonal matrix, block GE on $Q^T \pi^T = 0$ yields $Z^T \pi^T = 0$ (or equivalently $\pi Z = 0$), where

$$Z = \begin{pmatrix} D_0 & & & \\ A_2 & D_1 & & \\ & A_2 & D_2 & \\ & & \ddots & \ddots \end{pmatrix}, \tag{5}$$

$D_0 = B_0$, and $D_{l+1} = A_1 - A_2 D_l^{-1} A_0$ for $l \geq 0$.

Recalling that $\pi = (\pi_0, \pi_1, \ldots)$ and using $\pi Z = 0$, we obtain $\pi_l D_l + \pi_{l+1} A_2 = 0$, which implies $\pi_l = -\pi_{l+1} A_2 D_l^{-1}$ for $l \geq 0$. That $C_l \geq 0$ for $l \geq 0$ follows from $-D_l^{-1} \geq 0$ and $A_2 \geq 0$. ∎

4.1. Checking for the LG distribution. The form of Z in (5) together with Proposition 5 suggests the next lemma.

LEMMA 1. *If $D_{L+1} = D_L$ for some finite nonnegative integer L, then $D_l = D_L$ for $l > L + 1$, and $\pi_L = \pi_{L+k} C_L^k$ for $k \geq 0$.*

Proof. From Proposition 5 we have $D_{L+1} = A_1 - A_2 D_L^{-1} A_0$ and $D_{L+2} = A_1 - A_2 D_{L+1}^{-1} A_0$. If $D_{L+1} = D_L$, then $D_{L+2} = A_1 - A_2 D_L^{-1} A_0 = D_{L+1} = D_L$. The same argument may be used to show that $D_l = D_L$ for $l > L + 2$. The second part of the lemma follows from its first part and the last part of Proposition 5. ∎

The next theorem states a condition under which one has an LG distribution.

THEOREM 1. *Let L be the smallest finite nonnegative integer for which $D_{L+1} = D_L$. Then the stationary distribution of Q is LG with parameter less than or equal to L.*

Proof. From Lemma 1 and (5), when $D_{L+1} = D_L$, we have

$$Z = \begin{pmatrix} D_0 & & & & \\ A_2 & D_1 & & & \\ & \ddots & \ddots & & \\ & & A_2 & D_{L-1} & \\ & & & Y_L & Z_L \end{pmatrix}, \tag{6}$$

where

$$Y_L = \begin{pmatrix} A_2 \\ 0 \\ 0 \\ \vdots \end{pmatrix} \quad \text{and} \quad Z_L = \begin{pmatrix} D_L & & & \\ A_2 & D_L & & \\ & A_2 & D_L & \\ & & \ddots & \ddots \end{pmatrix}.$$

Since π_l of length m is positive for finite l and unique up to a multiplicative constant with $\lim_{l\to\infty} \pi_l = 0$, the identities $(\pi_L, \pi_{L+1}, \ldots)Z_L = 0$ and $(\pi_{L+1}, \pi_{L+2}, \ldots)Z_L = 0$ obtained from equations $\pi Z = 0$ and (6) together with the recursive structure of Z_L given by

$$Z_L = \begin{pmatrix} D_L & \\ Y_L & Z_L \end{pmatrix}$$

suggest that $\pi_{l+1} = \alpha\pi_l$ for $l \geq L$, where $\alpha \in (0, 1)$. ∎

COROLLARY 2. *When $B_0 = A_1 - A_2 B_0^{-1} A_0$, the stationary distribution of Q is LG with parameter $L = 0$.*

Next we state two lemmas, which will be used in checking for an LG distribution.

LEMMA 2. *If A_1 is irreducible and $A_2 e > 0$, then D_l is irreducible and $C_l > 0$ for $l \geq 1$.*

Proof. From Proposition 5 we have $D_{l+1} = A_1 + C_l A_0$, where $C_l = -A_2 D_l^{-1} \geq 0$ and $l \geq 0$. Since $A_0 \geq 0$ by definition, we obtain $C_l A_0 \geq 0$. Besides, A_1 has nonnegative off-diagonal elements and is assumed to be irreducible. Hence, its sum with the nonnegative $C_l A_0$ will not change the irreducibility, thereby implying irreducible D_{l+1} for $l \geq 0$. Alternatively, D_l, $l \geq 1$, is irreducible. That $-D_l$ is a nonsingular M-matrix from Proposition 5, together with the fact it is irreducible, implies $-D_l^{-1} > 0$ for $l \geq 1$ [3, p. 141]. Since $A_2 \geq 0$ and is assumed to have a nonzero in each row, its product with $-D_l^{-1}$ is positive. Hence, $C_l > 0$ for $l \geq 1$. ∎

LEMMA 3. *If $e^T A_0 > 0$, $A_2 e > 0$, and D_L is irreducible for some finite nonnegative integer L, then D_l is irreducible and $C_l > 0$ for $l \geq L$.*

Proof. When D_L is irreducible and A_2 has a nonzero in each row, we have $C_L > 0$ as in the proof of Lemma 1. Since $A_0 \geq 0$ and is assumed to have a nonzero in each column, we have $C_L A_0 > 0$, thereby implying an irreducible D_{L+1}. The same circle of arguments may be used to show that $C_l > 0$ and D_{l+1} is irreducible for $l > L$. ∎

The next theorem states another condition under which one has an LG distribution.

THEOREM 2. *Let L be the smallest finite nonnegative integer for which C_l is irreducible and $\rho(C_l) = \rho(C_{l+1})$, where $l \geq L$. Then the stationary distribution of Q is LG with parameter L.*

Proof. From Proposition 5 we have $C_l \geq 0$ for $l \geq 0$. If C_l, $l \geq L$, is irreducible, then by the Perron–Frobenius theorem C_l has $\rho(C_l) > 0$ as a simple eigenvalue and a corresponding positive left-hand eigenvector. There are no other linearly independent positive left-hand eigenvectors of C_l [10, p. 673]. From Proposition 5 we also have $\pi_l = \pi_{l+1}C_l$ and $\pi_l > 0$ with $\lim_{l\to\infty} \pi_l = 0$. Multiplying both sides of $\pi_l = \pi_{l+1}C_l$ by $\rho(C_l)$, we obtain $\rho(C_l)\pi_l = (\rho(C_l)\pi_{l+1})C_l$. Since $\rho(C_l)$ is a simple eigenvalue of C_l for $l \geq L$, we must have π_l as its corresponding positive left-hand eigenvector. Therefore, it must also be that $\pi_l = \rho(C_l)\pi_{l+1}$ for $l \geq L$. Since $\rho(C_l) = \rho(C_{l+1})$ for $l \geq L$, we have $\pi_l = \rho(C_L)\pi_{l+1}$, or $\pi_{l+1} = (1/\rho(C_L))\pi_l$ for $l \geq L$. Consequently, Q has an LG distribution with parameter L. □

4.2. Computing the LG distribution. The next theorem gives the value of α in (3) and indicates how π_L can be computed up to a multiplicative constant when one has an LG distribution with parameter L.

THEOREM 3. *If the stationary distribution of Q is LG with parameter L, then $\rho(C_L)\pi_L = \pi_L C_L$, where $\alpha = 1/\rho(C_L)$ and $\pi_L > 0$ in* (3).

Proof. Since Q has an LG distribution with parameter L, from (3) we have $\pi_{L+1} = \alpha\pi_L$, where $\alpha \in (0, 1)$, and $\pi_L > 0$ and $\pi_{L+1} > 0$ with $\lim_{l\to\infty} \pi_l = 0$. That is, for finite L, π_{L+1} is a positive multiple of π_L. Furthermore, from Proposition 5 we have $\pi_L = \pi_{L+1}C_L$, where $C_L \geq 0$. Since π_{L+1} is a positive multiple of π_L, π_L is clearly a positive left-hand eigenvector of C_L and therefore corresponds to the eigenvalue $\rho(C_L)$ [3, p. 28]. Combining the two statements, we obtain $\rho(C_L)\pi_L = \pi_L C_L$, where $\alpha = 1/\rho(C_L)$ and $\pi_L > 0$. □

COROLLARY 3. *When the stationary distribution of Q is LG with parameter less than or equal to L, where $L > 0$, if $\rho(C_L) \neq \rho(C_{L-1})$, then the parameter is L; otherwise the parameter is less than or equal to $L - 1$.*

5. Examples revisited. In this section, we demonstrate the results of the previous section using the three examples introduced in section 3.

5.1. Example 1. For the first example in section 2, D_l^{-1}, $l \geq 0$, is a full matrix, and we have experimentally shown that $D_{l+1} = D_l$ as l approaches infinity. For the particular case of $m = 2$, we have

$$B_0^{-1} = \frac{-1}{\lambda_1(d-\mu_1)} \begin{pmatrix} \lambda_1+\mu_2 & \lambda_2 \\ \mu_2 & \lambda_1+\lambda_2 \end{pmatrix} \quad \text{and} \quad C_0 = -A_2B_0^{-1} = -\mu_1 B_0^{-1},$$

where $d = \lambda_1 + \lambda_2 + \mu_1 + \mu_2$. The correction to A_1 is given by $C_0A_0 = -\lambda_1\mu_1B_0^{-1}$, and therefore

$$D_1 = A_1 + C_0A_0 = \begin{pmatrix} -(d-\mu_2) + \frac{\mu_1(\lambda_1+\mu_2)}{d-\mu_1} & \lambda_2 + \frac{\lambda_2\mu_1}{d-\mu_1} \\ \mu_2 + \frac{\mu_1\mu_2}{d-\mu_1} & -(d-\lambda_2) + \frac{\mu_1(\lambda_1+\lambda_2)}{d-\mu_1} \end{pmatrix} \neq B_0.$$

In a similar manner one can show that $D_{l+1} \neq D_l$ for finite values of l. Hence, Theorem 1 does not apply. However, Lemma 3 applies since A_0 and A_2 are of full-rank and D_0 is irreducible, implying irreducible C_l for $l \geq 0$. Consequently, there is reason to guess that the QBD MC has an LG distribution with parameter $L = 0$ from Theorem 2 and to compute the eigenvalue-eigenvector pair $(\rho(C_0),\pi_0)$ using

Theorem 3. Then the guessed solution can be verified in $\pi Q = 0$. Although this approach will sometimes fail, it works in Example 1 and can be recommended for small values of L.

For $m = 2$, it is not difficult to find, using Theorem 3, that $\rho(C_0) = \mu_1/\lambda_1 > 1$, implying $\alpha = \lambda_1/\mu_1$, and

$$\pi_0 = (1-\alpha)\begin{pmatrix} \frac{1-\nu}{1-\nu^2} & \frac{\nu(1-\nu)}{1-\nu^2} \end{pmatrix},$$

where $\nu = \lambda_2/\mu_2$.

5.2. Example 2. Consider the second example in section 2, for which we have

$$B_0^{-1} = \frac{-1}{1-2p}\begin{pmatrix} 1 & 1 \\ 2p & 1 \end{pmatrix} \quad \text{and} \quad C_0 = -A_2 B_0^{-1} = \frac{1-p}{1-2p}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

Note that C_0 is reducible. The correction to A_1 is given by $C_0 A_0 = (1-p)e_1 e_2^T$, and therefore

$$D_1 = A_1 + C_0 A_0 = \begin{pmatrix} -1 & 1 \\ 2p & -1 \end{pmatrix} = B_0.$$

Hence, in this example, $D_l = D_0$ for $l \geq 1$ from Lemma 1 due to $D_1 = D_0$. From Corollary 2 we conclude that Example 2 has an LG distribution with parameter $L = 0$.

Finally, from Theorem 3 we obtain $\rho(C_0) = (1-p)/(1-2p) > 1$, implying $\alpha = (1-2p)/(1-p)$, and $\pi_0 = (1-\alpha)(1/2\ 1/2)$.

5.3. Example 3. Now consider the third example in section 3, for which we have

$$B_0^{-1} = \frac{-1}{m\lambda}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ & 1 & \cdots & 1 \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix} \quad \text{and} \quad C_0 = -A_2 B_0^{-1} = \frac{\mu}{m\lambda}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ & 1 & \cdots & 1 \\ & & \ddots & \vdots \\ & & & 1 \end{pmatrix}.$$

Note that C_0 is reducible and $\rho(C_0) = \mu/(m\lambda)$, which is not necessarily greater than 1. The correction to A_1 is given by $C_0 A_0 = \mu e e_1^T$, and therefore

$$D_1 = \begin{pmatrix} -m\lambda & m\lambda & & & \\ \mu & -(m\lambda+\mu) & m\lambda & & \\ \vdots & & \ddots & \ddots & \\ \mu & & & -(m\lambda+\mu) & m\lambda \\ \mu & & & & -(m\lambda+\mu) \end{pmatrix} \neq B_0.$$

Noticing that $D_1 = A_1 + \mu e e_1^T$, in which the correction $\mu e e_1^T$ is of rank-1, the Sherman–Morrison formula [10, p. 124] yields

$$D_1^{-1} = A_1^{-1} - \mu\frac{A_1^{-1} e e_1^T A_1^{-1}}{1 + \mu e_1^T A_1^{-1} e}.$$

Letting $\gamma = m\lambda/(m\lambda + \mu)$, we obtain

$$A_1^{-1} = \frac{-1}{m\lambda + \mu} \begin{pmatrix} 1 & \gamma & \gamma^2 & \cdots & \gamma^{m-1} \\ & 1 & \gamma & \cdots & \gamma^{m-2} \\ & & \ddots & \ddots & \vdots \\ & & & 1 & \gamma \\ & & & & 1 \end{pmatrix}, \quad (1 + \mu e_1^T A_1^{-1} e) = \gamma^m,$$

$$\mu(A_1^{-1}e)(e_1^T A_1^{-1}) = \frac{1}{m\lambda + \mu} \begin{pmatrix} 1-\gamma^m & \gamma(1-\gamma^m) & \cdots & \gamma^{m-1}(1-\gamma^m) \\ 1-\gamma^{m-1} & \gamma(1-\gamma^{m-1}) & \cdots & \gamma^{m-1}(1-\gamma^{m-1}) \\ \vdots & \vdots & \ddots & \vdots \\ 1-\gamma & \gamma(1-\gamma) & \cdots & \gamma^{m-1}(1-\gamma) \end{pmatrix},$$

and, after some algebra, $C_1 A_0 = \mu e e_1^T$. Hence, $D_2 = A_1 + C_1 A_0 = D_1$, implying $D_l = D_1$ for $l \geq 2$ from Lemma 1. From Theorem 1 we have an LG distribution with parameter $L \leq 1$. We also remark that the two matrices C_0 and C_1 introduced in Proposition 4 for QBD processes with rank-1 A_0 matrices are given in this example as $C_0 = -\mu D_0^{-1}$ and $C_1 = -\mu D_1^{-1}$. Since $\rho(C_0)$ may be less than 1 and therefore different than $\rho(C_1)$, from Corollary 3 we conclude Example 2 has an LG distribution with parameter $L = 1$.

Regarding the computation of α, for instance, when $m = 2$

$$C_0 = \eta \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad C_1 = \eta \begin{pmatrix} 1+\eta & 1 \\ \eta & 1 \end{pmatrix},$$

where $\eta = \mu/(2\lambda)$. Hence, we have

$$\rho(C_1) = \eta \left(1 + \frac{1}{2}\eta + \sqrt{\eta \left(1 + \frac{1}{4}\eta \right)} \right).$$

Note that $\rho(C_0) \neq \rho(C_1)$. Now, using $\rho(C_1)\pi_1 = \pi_1 C_1$, $\pi_0 = \pi_1 C_0$, and $\pi_1 e/(1-\alpha) + \pi_0 e = 1$, where $\alpha = 1/\rho(C_1)$, we obtain

$$\pi_1 = \left(\frac{(\rho(C_1) - \eta)(\rho(C_1) - 1)}{\rho^2(C_1) + \eta(\rho(C_1) - 1)(2\rho(C_1) - \eta)} \;\; \frac{\eta(\rho(C_1) - 1)}{\rho^2(C_1) + \eta(\rho(C_1) - 1)(2\rho(C_1) - \eta)} \right)$$

and

$$\pi_0 = \left(\frac{\eta(\rho(C_1) - \eta)(\rho(C_1) - 1)}{\rho^2(C_1) + \eta(\rho(C_1) - 1)(2\rho(C_1) - \eta)} \;\; \frac{\eta\rho(C_1)(\rho(C_1) - 1)}{\rho^2(C_1) + \eta(\rho(C_1) - 1)(2\rho(C_1) - \eta)} \right).$$

Normally the computation would be performed numerically for the given parameters of the problem. For $m \geq 3$, we would first compute C_0 and C_1. Then we would obtain the eigenvalue-eigenvector pair $(\rho(C_1), \pi_1)$ from $\rho(C_1)\pi_1 = \pi_1 C_1$ (see Theorem 3). Next we would compute $\pi_0 = \pi_1 C_0$. Finally we would normalize π_0 and π_1 with $\pi_1 e/(1-\alpha) + \pi_0 e$.

6. Conclusion. This paper introduces necessary and sufficient conditions for a homogeneous continuous-time quasi-birth-and-death (QBD) Markov chain (MC) to possess level-geometric (LG) stationary distribution. Furthermore, it discusses how an LG distribution can be computed when it exists. Results that utilize the matrices

A_0, A_1, A_2, and B_0 are given, showing how one can easily check for and compute an LG distribution with parameter $L \leq 1$. The results are elaborated through three examples. Examples 2 and 3, which have been used in the literature as test cases, are shown to possess LG distributions, respectively, with parameters $L = 0$ and $L = 1$. Since the matrices A_0, A_1, A_2, and B_0 that arise in applications are usually sparse, the results developed in this paper may be used before resorting to quadratically convergent algorithms to compute the rate matrix, R.

Acknowledgments. We thank the anonymous referees and Reinhard Nabben for their remarks, which led to an improved manuscript.

REFERENCES

[1] N. Akar and K. Sohraby, *An invariant subspace approach in M/G/1 and G/M/1 type Markov chains*, Comm. Statist. Stochastic Models, 13 (1997), pp. 381–416.

[2] M. Benzi and M. Tůma, *A parallel solver for large-scale Markov chains*, Appl. Numer. Math., 41 (2002), pp. 135–153.

[3] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, SIAM, Philadelphia, 1994.

[4] D. Bini and B. Meini, *On the solution of a nonlinear matrix equation arising in queueing problems*, SIAM J. Matrix Anal. Appl., 17 (1996), pp. 906–926.

[5] B. R. Haverkort, *Performance of Computer Communication Systems: A Model-Based Approach*, John Wiley & Sons, Chichester, England, 1998.

[6] K. Kant, *Introduction to Computer System Performance Evaluation*, McGraw-Hill, New York, 1992.

[7] J. R. Kemeny and J. L. Snell, *Finite Markov Chains*, Van Nostrand, New York, 1960.

[8] G. Latouche and V. Ramaswami, *A logarithmic reduction algorithm for quasi-birth-and-processes*, J. Appl. Probab., 30 (1993), pp. 650–674.

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

[10] C. D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, Philadelphia, 2000.

[11] M. F. Neuts, *Matrix-Geometric Solutions in Stochastic Models. An Algorithmic Approach*, The Johns Hopkins University Press, Baltimore, MD, 1981.

[12] M. F. Neuts, *Structured Stochastic Matrices of M/G/1 Type and Their Applications*, Marcel Dekker, New York, 1989.

[13] A. Ost, *Performance of Communication Systems: A Model-Based Approach with Matrix-Geometric Methods*, Springer-Verlag, Berlin, 2001.

[14] W. J. Stewart, *Introduction to the Numerical Solution of Markov Chains*, Princeton University Press, Princeton, NJ, 1994.