On Latouche-Ramaswami's Logarithmic Reduction Algorithm for Quasi-birth-and-death Processes

Qiang Ye *

March 8, 2001

Abstract

In this paper, we consider Latouche-Ramaswami's logarithmic reduction algorithm for solving quasi-birth-and-death models. We shall present some theoretical properties concerning convergence of the algorithm and discuss numerical issues arising in finite precision implementations. In particular, we shall present a numerically more stable implementation together with its rounding error analysis. Numerical examples will be given to demonstrate the higher accuracy achieved by the refined implementation.

1 Introduction

A large number of queueing systems can be modeled by two-dimensional Markov chains called *quasi-birth-and-death processes* (QBD) (see [16] for examples of QBD). A QBD process has a matrix-geometric stationary distribution [22] which is defined through the minimal nonnegative solution of the matrix equation of the form

$$R = A_0 + RA_1 + R^2A_2 \tag{1}$$

or through the related equation

$$G = A_2 + A_1G + A_0G^2 \tag{2}$$

where (A_0, A_1, A_2) are three $n \times n$ nonnegative matrices that define the transition matrix of the QBD process and they are such that $A \equiv \Sigma_{k=0}^{2} A_k$ is stochastic [16]. Numerically solving the above equations play an important role in the applications of QBD models and various numerical methods have been developed to solve (1), (2) or some more general nonlinear matrix equations [22]. They range from linearly convergent fixed point iterations to quadratically convergent iterations such as Newton's method (see [2, 3, 4, 9, 11, 14, 15, 17, 18, 20] and the references contained therein). One of the most efficient ones for (1) or (2) is the algorithm of Latouche and Ramaswami [15], which possesses a quadratic convergence property at a cost more comparable to other linear convergent algorithms. The Latouche-Ramaswami algorithm involves inverting certain matrices

*Department of Mathematics, University of Kentucky, Lexington, KY 40506-0027. E-mail: qye@ms.uky.edu.

at each iteration, and relatively large roundoff errors may incur if the matrices for inversion are ill-conditioned. Because the algorithm is based on partial sum approximation of an infinite series expression for the solution, any roundoff error made at early steps for computing partial sums carries forward to future approximations without any mechanism for correcting such errors. Hence the accuracy of the final approximation is limited by roundoff errors accumulated from early stage of the iterations.

It appears that the ultimate accuracy to which the final solution can be obtained by the Latouche-Ramaswami algorithm depends on condition numbers of certain matrices in the middle of the process. We note however that the matrices defined at each iteration are related to a certain QBD process, which is derived from and is well determined by the original QBD process independent of such condition numbers. Indeed, we shall demonstrate that such a dependence of the algorithm on the condition numbers of the middle matrices is a consequence of the use of numerical methods for inverting a general matrix rather than the intrinsic physical characteristic of the solution. Specifically, we shall use in this paper the fact that the matrices involved in the inversion are *diagonally dominant M-matrices* which, if defined properly, can be inverted entrywise to high relative accuracy independent of any condition number [1]. Thus, by properly defining the matrices involved, each iteration of the algorithm can be implemented in a numerically more stable way to produce a more accurate solution. We remark that this approach has proved to be very effective in producing more accurate solution for fixed point based iterations in [23] and it is possible to generalize it to more general problems such as those of [5, 7, 21].

The paper is organized as follows. In Section 2, we shall briefly describe an entrywise forward stable algorithm developed in [1] for inverting a diagonally dominant M-matrix, which generalizes the GTH-algorithm [13] for stochastic matrices. In Section 3, we present the Latouche-Ramaswami algorithm and some theoretical results. Then in Section 4 we derive a refined implementation of the Latouche-Ramaswami algorithm and present an error analysis to demonstrate its improved accuracy. Finally in Section 5, we give some numerical examples.

Notation: For $m \times n$ matrices $B = (b_{ij})$ and $C = (c_{ij})$, we denote by $|B|$ the matrix of entries $|b_{ij}|$. We write $B \geq C$ if $b_{ij} \geq c_{ij}$ for all i and j. Throughout, we let $\mathbf{e}$ denote the column vector of all ones, i.e.,

$$\mathbf{e} = (1, 1, \cdots, 1)^T.$$

We use $fl(z)$ to denote the computed result of the expression z in a floating point arithmetic, which is also often denoted by $\hat{z}$.

2 Diagonally Dominant M-Matrix

Many matrices arising in stochastic models belong to the class of diagonally dominant M-matrices. Such matrices have better theoretical and numerical behavior and can be, for example, inverted entrywise to high relative accuracy (see [1]). In this section, we briefly discuss such an algorithm, which will be used in later sections.

A matrix $A = (a_{ij})$ is called a *diagonally dominant M-matrix* (**DDM**) if it has non-positive off-diagonal entries and has non-negative row sums [6] , i.e.,

$$a_{ij} \le 0, \quad \text{for } i \neq j, \quad \text{and } v \equiv A\mathbf{e} \ge 0.$$

Let $v = [v_1, v_2, \cdots, v_n]^T = A\mathbf{e}$, which is the row sums (or the diagonally dominant part) of A. It is easy to see that the diagonal entries of A are given by

$$a_{ii} = \Sigma_{j\neq i}(-a_{ij}) + v_i \ge 0. \tag{3}$$

Thus, such a matrix is determined by its off-diagonal entries and the row sums $v = (v_1, \cdots, v_n)^T$. Indeed, this is a more appropriate way to represent such matrices.

Definition 1 *Let $P = (p_{ij})$ be an $n \times n$ nonnegative matrix and $v = (v_i)$ be a nonnegative n-vector. We use (P, v) to represent the unique diagonally dominant M-matrices (**DDM**) A of the form $A = D - P$ that satisfies $A\mathbf{e} = v$, where D is a diagonal matrix. We write $A = (P, v)$.*

In this representation, the off-diagonal entries of A are those of $-P$ and the diagonal entries are given by (3). Note that the diagonal entries of P play no role in defining A and can be arbitrary.

One advantage of this new representation is that the solution to $Ax = b$ (with $b \ge 0$) is well determined by the data in (P, v) and can be computed more accurately. Namely, if A is represented in $A = (P, v)$, then the Gaussian elimination for A can be carried out in this representation and consequently involves no subtraction. This generalizes the idea in the GTH-algorithm for stochastic matrices [13, 19]. The following algorithm was developed in [1] and called a GTH-like Algorithm.

Algorithm 1: GTH-like Algorithm for Solving $Ax = b$.
Step 1: LU factorization
For $k = 1, 2, \cdots, n-1$,
the pivot $\alpha_k = v_k^{(k)} + \sum_{j=k+1}^{n} p_{kj}^{(k)}$.
$p_{ij}^{(k+1)} = p_{ij}^{(k)} + \frac{p_{ik}^{(k)} p_{kj}^{(k)}}{\alpha_k}$ for $i, j > k$, $i \neq j$.
$v_j^{(k+1)} = v_j^{(k)} + \frac{v_k^{(k)}}{\alpha_k} p_{jk}^{(k)}$ for $j > k$.
Step 2: Solving Ly=b
$y_1 = b_1/\alpha_1$
For $k = 2, 3, \cdots n$,
$y_k = (b_k + \sum_{j=1}^{k-1} p_{k,j}^{(j)} y_j)/\alpha_k$
Step 3: Solving Ux=y
$x_n = y_n$
For $k = n-1, n-2, \cdots, 1$
$x_k = y_k + (\sum_{j=k+1}^{n} p_{kj}^{(k)} x_j)/\alpha_k$

The above algorithm has essentially the same computational flops as the standard Gaussian elimination. Yet, the computational results could be significantly more accurate as no subtraction is involved. This is demonstrated by a rounding error analysis, as given in the next theorem [1].

Theorem 1 *Suppose Algorithm 1 is carried out in a floating point arithmetic with the machine precision ϵ to solve the linear system $Ax = b$, and the input data p_{ij}, v_i, b_i $(i, j = 1, 2, \cdots, n)$ are floating-point numbers. Then the computed solution $\widehat{x}$ satisfies*

$$|x - \widehat{x}| \leq (\phi(n)\epsilon + O(\epsilon^2))x. \tag{4}$$

where $\phi(n) = \frac{2(n+2)(n+3)(2n+5)}{3}$.

Each entry of the solution has a relative acuracy of order ϵ. This compares with the standard Gaussian elimination with partial pivoting whose solution has a normwise accuracy dependent on the condition number of A.

In our error analysis, we assume the following standard model for roundoff errors in basic matrix computations [10, p.66]

$$fl(x + y) = x + y + f \quad \text{with} \quad |f| \leq \epsilon|x + y| \tag{5}$$

$$fl(Ax) = Ax + f \quad \text{with} \quad |f| \leq \epsilon n|A||x| + O(\epsilon^2). \tag{6}$$

where $A \in R^{n \times n}$, $x, y \in R^n$ are floating point matrices and ϵ is the machine roundoff unit. Recall that $fl(z)$ denotes the computed result of some expression z in a floating point arithmetic. It is also easy to show that

$$fl(y + Ax) = y + Ax + f \quad \text{with} \quad |f| \leq \epsilon(|y| + (n+1)|A||x|) + O(\epsilon^2), \tag{7}$$

which we will use in later analysis.

3 Latouche-Ramaswami's Algorithm

We first describe the following quadratically convergent iteration for solving (2) due to Latouche and Ramaswami [15].

Algorithm 2: Latouche-Ramaswami's Algorithm.
$H_0 = (I - A_1)^{-1}A_0$;
$L_0 = (I - A_1)^{-1}A_2$;
$G_0 = L_0$;
$T_0 = H_0$;
For $i = 0, 1, 2, \cdots$ until convergence
  $U_i = H_iL_i + L_iH_i$
  $H_{i+1} = (I - U_i)^{-1}H_i^2$;
  $L_{i+1} = (I - U_i)^{-1}L_i^2$;
  $G_{i+1} = G_i + T_iL_{i+1}$,
  $T_{i+1} = T_iH_{i+1}$,

The algorithm implements the following expression of the minimal non-negative solution

$$G = \Sigma_{j=0}^{\infty}(H_0 H_1 \cdots H_{j-1})L_j. \tag{8}$$

and approximates G at step $i-1$ by the partial sum

$$G_i = \Sigma_{j=0}^{i}(H_0 H_1 \cdots H_{j-1})L_j$$

If the QBD is either positive recurrent or transient, then G_i converges to G quadratically. One potential numerical difficulty is caused by inverting $I - U_i$ in the algorithm. A standard inversion algorithm (e.g. the Guassian elimination with partial pivoting) usually produces a backward stable solution; but the relative error in $H_{i+1} = (I - U_i)^{-1}H_i^2$ depends on the condition number of $I - U_i$. Although we know that $\lim U_i = 0$ and therefore $I - U_i$ is asymptotically well-conditioned, at the early stage of iterations (smaller i), U_i needs not be small and thus the roundoff errors caused by inverting $I - U_i$ may be large at the beginning. Unfortunately, the algorithm itself has no error correction mechanism and such roundoff errors carry forward to all subsequent G_i. Thus the accuracy of the final solution is limited by the roundoff errors incurred at the beginning. This is the numerical issue that we shall consider in the next section. Here, we present some convergence properties of Algorithm 1 and discuss when there might be a slow convergence.

In studying the matrix equation (2), it is helpful to consider the level-reversed process as defined by (A_2, A_1, A_0) and the corresponding equation

$$\widehat{G} = A_0 + A_1\widehat{G} + A_2\widehat{G}^2, \tag{9}$$

i.e. $\widehat{G} = A_2\widehat{G}^2 + A_1\widehat{G} + A_0$. It is well-known that

1. If the QBD is positive recurrent, G is stochastic and $\widehat{G}$ is substochastic.

2. If the QBD is transient, $\widehat{G}$ is stochastic and G is substochastic.

3. If the QBD is null-recurrent, both G and $\widehat{G}$ are stochastic.

For Algorithm 1, it is also well known that $H_i + L_i$ is stochastic (see the derivation [15]). Furthermore, it can be deduced from [15] that the QBD as defined by $(H_i, 0, L_i)$ has G^{2^i} and $\hat{G}^{2^i}$ as the corresponding solutions of (2) and (9) respectively, i.e.,

$$G^{2^i} = L_i + H_i G^{2^{i+1}}, \text{ and } \widehat{G}^{2^i} = L_i\widehat{G}^{2^{i+1}} + H_i. \tag{10}$$

The property that $H_i + L_i$ is stochastic leads to the following identity concerning the approximation error as well as stochastic measure of G_i.

Theorem 2 *If the QBD is recurrent (positive recurrent or null recurrent), we have*

$$\mathbf{e} - G_i\mathbf{e} = H_0 H_1 \cdots H_i \mathbf{e}$$

and hence $\|G - G_i\|_\infty = \|H_0 H_1 \cdots H_i\|_\infty$.

Proof From the definition of G_i, we have

$$\begin{aligned}
\mathbf{e} - G_i\mathbf{e} &= \mathbf{e} - \Sigma_{j=0}^{i}(H_0H_1\cdots H_{j-1})L_j\mathbf{e} \\
&= \mathbf{e} - \Sigma_{j=0}^{i}(H_0H_1\cdots H_{j-1})L_j\mathbf{e} - \\
&\quad -\Sigma_{j=0}^{i}(H_0H_1\cdots H_{j-1})H_j\mathbf{e} + \Sigma_{j=1}^{i+1}(H_0H_1\cdots H_{j-1})\mathbf{e} \\
&= -\Sigma_{j=0}^{i}(H_0H_1\cdots H_{j-1})(L_je + H_j\mathbf{e} - \mathbf{e}) + (H_0H_1\cdots H_i)\mathbf{e} \\
&= H_0H_1\cdots H_i\mathbf{e}
\end{aligned}$$

where we have used $L_j\mathbf{e} + H_j\mathbf{e} - \mathbf{e} = 0$. In the case of recurrent QBD, G is stochastic and $G \geq G_i$. Thus

$$\|G - G_i\|_\infty = \|(G - G_i)\mathbf{e}\|_\infty = \|H_0H_1\cdots H_i\mathbf{e}\|_\infty = \|H_0H_1\cdots H_i\|_\infty.$$

□

This result characterizes the dependence of the convergence of G_i on H_i. In particular

$$\frac{\|G - G_i\|_\infty}{\|G - G_{i-1}\|_\infty} \leq \|H_i\|_\infty.$$

We next discuss convergence properties of H_i.

Lemma 1 *Assume that the QBD and $A = A_0 + A_1 + A_2$ are irreducible and the QBD is positive recurrent. Let g^T be the stationary probability vector of G. Then*

$$G = \mathbf{e}g^T + G_1 \quad \text{with } \rho(G_1) < 1. \tag{11}$$

Furthermore, for any $k \geq 1$,

$$G^k = \mathbf{e}g^T + G_1^k. \tag{12}$$

Proof By Theorem 7.2.1 of [16, p.152], G after a suitable permutation can be written as

$$G = \begin{pmatrix} G_{11} & 0 \\ G_{12} & G_{22} \end{pmatrix}$$

where G_{11} is stochastic and irreducible and G_{22} is lower triangular with zeros on the diagonal. Here G_{22} could possibly be empty, in which case G itself is irreducible. Then, G_{11} and hence G have exactly one eigenvalue with modulus 1. Note that g^T and $\mathbf{e}$ are left and right eigenvectors of G with $g^T\mathbf{e} = 1$. Hence, using an eigenvalue decomposition (e.g. the Jordan canonical form), we have $G = \mathbf{e}g^T + G_1$ and $\rho(G_1) < 1$. Furthermore, it follows from $g^TG_1 = g^T(G - \mathbf{e}g^T) = 0$ and $G_1\mathbf{e} = (G - \mathbf{e}g^T)\mathbf{e} = 0$ that

$$G^2 = \mathbf{e}g^T\mathbf{e}g^T + \mathbf{e}g^TG_1 + G_1\mathbf{e}g^T + G_1^2 = \mathbf{e}g^T + G_1^2.$$

Similarly, the more general (12) can be proved by an induction. □

Lemma 2 *Let α^T be the stationary probability vector of $A = A_0 + A_1 + A_2$ and let α_i^T be defined recursively by (assuming it is defined)*

$$\alpha_0^T = \frac{\alpha^T(I - A_1)}{\alpha^T A_0 \mathbf{e} + \alpha^T A_2 \mathbf{e}}, \quad \alpha_i^T = \frac{\alpha_{i-1}^T(I - U_{i-1})}{\alpha_{i-1}^T H_{i-1}^2 \mathbf{e} + \alpha_{i-1}^T L_{i-1}^2 \mathbf{e}}.$$

Then α_i^T is the stationary probability vector of $H_i + L_i$.

Proof It is easy to verify the statement for $i = 0$. Suppose it is true for some i, i.e. $\alpha_i^T(H_i + L_i) = \alpha_i^T$. Then $\alpha_i^T(H_i^2 + L_i^2) = \alpha_i^T(I - H_i L_i - L_i H_i) = \alpha_i^T(I - U_i)$. Therefore, $\alpha_{i+1}^T \mathbf{e} = 1$ and

$$\alpha_{i+1}^T(H_{i+1} + L_{i+1}) = \frac{\alpha_i^T(I - U_i)(I - U_i)^{-1}(H_i^2 + L_i^2)}{\alpha_i^T H_i^2 \mathbf{e} + \alpha_i^T L_i^2 \mathbf{e}} = \alpha_{i+1}^T.$$

This completes the proof. □

Theorem 3 *Assume that the QBD and A are irreducible and the QBD is positive recurrent. We have the following.*

$$\alpha_i^T H_i = \alpha_i^T L_i \widehat{G}^{2^i} = \alpha_i^T \widehat{G}^{2^i} \left(I + \widehat{G}^{2^i}\right)^{-1} \longrightarrow 0, \tag{13}$$

and

$$\alpha_i^T = (1 - \alpha_i^T H_i \mathbf{e}) g^T + \alpha_i^T H_i (I + G_1^{2^i}) \longrightarrow g^T. \tag{14}$$

Proof From the second of (10), we have

$$\alpha_i^T L_i \widehat{G}^{2^{i+1}} + \alpha_i^T H_i = \alpha_i^T \widehat{G}^{2^i} = \alpha_i^T (L_i + H_i) \widehat{G}^{2^i}.$$

Then,

$$\alpha_i^T H_i (I - \widehat{G}^{2^i}) = \alpha_i^T L_i \widehat{G}^{2^i} (I - \widehat{G}^{2^i})$$

which together with the fact that $\widehat{G}$ is substochastic imply

$$\alpha_i^T H_i = \alpha_i^T L_i \widehat{G}^{2^i} = \alpha_i^T \widehat{G}^{2^i} - \alpha_i^T H_i \widehat{G}^{2^i} \tag{15}$$

Thus we have $\alpha_i^T H_i (I + \widehat{G}^{2^i}) = \alpha_i^T \widehat{G}^{2^i}$ and hence (13).

Similarly, From the first of (10), we have

$$\alpha_i^T L_i (I - G^{2^i}) = \alpha_i^T H_i G^{2^i} (I - G^{2^i}).$$

Using Lemma 1, we obtain

$$\alpha_i^T L_i (I - G_1^{2^i}) - \alpha_i^T L_i \mathbf{e} g^T = \alpha_i^T H_i G_1^{2^i} (I - G_1^{2^i})$$

where we note that $G^{2^i}(I - G^{2^i}) = (\mathbf{e} g^T + G_1^{2^i})(I - \mathbf{e} g^T - G_1^{2^i}) = G_1^{2^i}(I - G_1^{2^i})$. Hence, inverting $I - G_1^{2^i}$ and using $g^T(I - G_1^{2^i}) = g^T$, we have

$$\alpha_i^T L_i = \alpha_i^T H_i G_1^{2^i} + \alpha_i^T L_i \mathbf{e} g^T \tag{16}$$

Thus,

$$\alpha_i^T = \alpha_i^T H_i + \alpha_i^T H_i = \alpha_i^T H_i (I + G_1^{2^i}) + (1 - \alpha_i^T H_i \mathbf{e}) g^T.$$

This proves (14). □

The theorem shows that $g^T H_i = \alpha_i^T H_i + (g^T - \alpha_i^T) H_i$ converges quadratically and the convergence at initial stage could be slow if $\rho(\hat{G})$ is near 1 or if $\hat{G}^{2^i}$ (for some small i) has an eigenvalue near -1 (which results in a large $\left(I + \widehat{G}^{2^i}\right)^{-1}$). Clearly, if all entries of g are nonzero, H_i converges in the same way. We next consider another factor (the initial state) that affects the number of iterations.

Theorem 4 *For all $i \geq 1$, we have*

$$\alpha_i^T L_i \mathbf{e} - \alpha_i^T H_i \mathbf{e} = \alpha_i^T L_i^2 \mathbf{e} - \alpha_i^T H_i^2 \mathbf{e} = \frac{\alpha_{i-1}^T L_{i-1}^2 \mathbf{e} - \alpha_{i-1}^T H_{i-1}^2 \mathbf{e}}{\alpha_{i-1}^T L_{i-1}^2 \mathbf{e} + \alpha_{i-1}^T H_{i-1}^2 \mathbf{e}}, \tag{17}$$

and

$$\alpha_0^T L_0 \mathbf{e} - \alpha_0^T H_0 \mathbf{e} = \alpha_0^T L_0^2 \mathbf{e} - \alpha_0^T H_0^2 \mathbf{e} = \frac{\alpha^T A_2 \mathbf{e} - \alpha^T A_0 \mathbf{e}}{\alpha^T A_2 \mathbf{e} + \alpha^T A_0 \mathbf{e}}. \tag{18}$$

Proof From $\alpha_{i-1}^T (L_{i-1}^2 + U_{i-1} + H_{i-1}^2) = \alpha_{i-1}^T (L_{i-1} + H_{i-1})^2 = \alpha_{i-1}^T$, it follows that $\alpha_{i-1}^T L_{i-1}^2 = \alpha_{i-1}^T (I - U_{i-1}) + \alpha_{i-1}^T H_{i-1}^2$. Then

$$\alpha_i^T L_i^2 \mathbf{e} = \frac{\alpha_{i-1}^T L_{i-1}^2 (I - U_{i-1})^{-1} L_{i-1}^2 \mathbf{e}}{\alpha_{i-1}^T H_{i-1}^2 \mathbf{e} + \alpha_{i-1}^T L_{i-1}^2 \mathbf{e}} = \frac{\alpha_{i-1}^T L_{i-1}^2 \mathbf{e} + \alpha_{i-1}^T H_{i-1}^2 (I - U_{i-1})^{-1} L_{i-1}^2 \mathbf{e}}{\alpha_{i-1}^T H_{i-1}^2 \mathbf{e} + \alpha_{i-1}^T L_{i-1}^2 \mathbf{e}}. \tag{19}$$

Using $\alpha_i^T L_i^2 \mathbf{e} = (\alpha_i^T - \alpha_i^T H_i) L_i \mathbf{e}$ and $\alpha_i^T L_i^2 \mathbf{e} = \alpha_i^T L_i (\mathbf{e} - H_i \mathbf{e})$, we obtain

$$\alpha_i^T L_i^2 \mathbf{e} = \alpha_i^T L_i \mathbf{e} - \frac{1}{2} \alpha_i^T U_i \mathbf{e}. \tag{20}$$

Similarly, we can prove

$$\alpha_i^T H_i^2 \mathbf{e} = \frac{\alpha_{i-1}^T H_{i-1}^2 \mathbf{e} + \alpha_{i-1}^T H_{i-1}^2 (I - U_{i-1})^{-1} L_{i-1}^2 \mathbf{e}}{\alpha_{i-1}^T H_{i-1}^2 \mathbf{e} + \alpha_{i-1}^T L_{i-1}^2 \mathbf{e}}. \tag{21}$$

and

$$\alpha_i^T H_i^2 \mathbf{e} = \alpha_i^T H_i \mathbf{e} - \frac{1}{2} \alpha_i^T U_i \mathbf{e}. \tag{22}$$

Now combining (19) with (21) and (20) with (22) we obtain (17). Clearly, (18) can be proved similarly.

□

For the positive recurrent case, it follows from $\alpha_i^T L_i^2 \mathbf{e} + \alpha_i^T H_i^2 \mathbf{e} \leq 1$ that $\alpha_i^T L_i^2 \mathbf{e} - \alpha_i^T H_i^2 \mathbf{e}$ and hence $\alpha_i^T L_i \mathbf{e} - \alpha_i^T H_i$ increase monotonically to 1. By considering $\alpha_i^T L_i \mathbf{e} - \alpha_i^T H_i \mathbf{e}$ as an indication of convergence of H_i, (17) suggests the need of more iterations if the initial $\alpha^T A_2 \mathbf{e} - \alpha^T A_0 \mathbf{e}$ is small.

In the extreme case of a null-recurrent QBD, $\alpha^T A_2 \mathbf{e} - \alpha^T A_0 \mathbf{e} = 0$, $\alpha_i^T L_i \mathbf{e} - \alpha_i^T H_i \mathbf{e}$ remains to be 0 for all i. Noting that $\alpha_i^T L_i \mathbf{e} + \alpha_i^T H_i \mathbf{e} = 1$, we have $\alpha_i^T L_i \mathbf{e} = \alpha_i^T H_i \mathbf{e} = \frac{1}{2}$ (which has also been obtained in [12]). So H_i does not converge to 0.

Theorem 5 *Assume that the QBD and A are irreducible and the QBD is null-recurrent. Let g^T and $\hat{g}^T$ be the stationary probability vectors of G and $\widehat{G}$ respectively and let $G = \mathbf{e}g^T + G_1$ and $\widehat{G} = \mathbf{e}\hat{g}^T + \widehat{G}_1$. Then,*

$$\alpha_i^T H_i \longrightarrow \frac{1}{2}\hat{g}^T, \quad \alpha_i^T \longrightarrow u^T \equiv \frac{1}{2}(g^T + \hat{g}^T). \tag{23}$$

Proof We first note that the proof of (16) is valid for the null-recurrent case and we have

$$\begin{aligned} \alpha_i^T L_i &= \alpha_i^T H_i G_1^{2^i} + \alpha_i^T L_i \mathbf{e} g^T \\ &= \alpha_i^T G_1^{2^i} - \alpha_i^T L_i G_1^{2^i} + \frac{1}{2} g^T, \end{aligned}$$

which implies

$$\alpha_i^T L_i = \alpha_i^T G_1^{2^i}(I + G_1^{2^i})^{-1} + \frac{1}{2}g^T \longrightarrow \frac{1}{2}g^T.$$

On the other hand, since the QBD is null-recurrent, $\widehat{G}$ is stochastic. Therefore, by the same reason,

$$\alpha_i^T H_i = \alpha_i^T \widehat{G}_1^{2^i}(I + \widehat{G}_1^{2^i})^{-1} + \frac{1}{2}\hat{g}^T \longrightarrow \frac{1}{2}\hat{g}^T.$$

Thus, $\alpha_i^T = \alpha_i^T L_i + \alpha_i^T H_i \longrightarrow \frac{1}{2}(g^T + \hat{g}^T)$. □

Now, assume that all entries of g and $\hat{g}$ are nonzero. Then $u > \frac{1}{2}\hat{g}$. Writing $U = diag(u)$ (i.e. the diagonal matrix with $U\mathbf{e} = u$), we have

$$\mathbf{e}^T U H_i U^{-1} = u^T H_i U^{-1} \longrightarrow \frac{1}{2}\hat{g}^T U^{-1} < \mathbf{e}^T.$$

In particular, for some $\delta > 0$ and sufficiently large i, $\|UH_iU^{-1}\|_1 \le \max(\hat{g}^T U^{-1}/2) + \delta < 1$. Thus, $H_0 H_1 \cdots H_i$ converges to 0 linearly and so is $G_i - G$. We remark that the convergence property of null-recurrent QBDs has recently been studied independently by Guo [12]. There, under some mild assumptions, it is shown that G_i converges precisely at the rate of 1/2.

We finally mention an interesting implication of (15). For the case $i = 0$, we have $\alpha_0^T H_0 = \alpha_0^T L_0 \widehat{G}$. Then, if v is an eigenvector of $\widehat{G}$ corresponding to $\rho(G)$, then $\alpha_0^T H_0 v = \rho(G)\alpha_0^T L_0 v$ i.e.

$$\rho(G) = \frac{\alpha_0^T H_0 v}{\alpha_0^T L_0 v} = \frac{\alpha^T A_0 v}{\alpha^T A_2 v}.$$

In another word, $\frac{\alpha_0^T A_0 v}{\alpha_0^T A_2 v}$ appears to be a measure of how close the QBD is to being null-recurrent.

We also note that our derivation of these rsults are purely algebraic; but for at least some of them, probabilistic arguments should be possible.

4 Refined Implementation and Error Analysis

In this section, we present a refined implementation of the Latouche-Ramaswami Algorithm. We first observe that Algorithm 2 involves addition operations (of positive numbers) except at the steps involving inverting $I - U_i$ or $I - A_1$. Thus the only potential numerical instability lies at the inversion steps. We address this issue now.

Theorem 6 *For Algorithm 2, $I - U_i$ is a DDM matrix and $(I - U_i)\mathbf{e} = H_i^2\mathbf{e} + L_i^2\mathbf{e} \geq 0$, i.e. in the notation of Definition 1, $I - U_i$ is represented as*

$$I - U_i = (U_i, v_i), \text{ with } v_i \equiv H_i^2\mathbf{e} + L_i^2\mathbf{e}.$$

Proof First, we note that $(H_i + L_i)^2 = H_i^2 + L^2 + H_iL_i + L_iH_i = H_i^2 + L^2 + U_i$. Since $H_i + L_i$ is stochastic, we have $\mathbf{e} = (H_i + L_i)^2\mathbf{e} = (H_i^2 + L_i^2)\mathbf{e} + U_i\mathbf{e}$, which leads to $(I - U_i)\mathbf{e} = H_i^2\mathbf{e} + L_i^2\mathbf{e}$. Clearly $U_i \geq 0$. Therefore $I - U_i$ is a DDM matrix and $I - U_i = (U_i, v_i)$. □

Now, computations of U_i and v_i involve additions only and therefore the data used to represent $I - U_i$ is each computed to the machine precision. Thus, using Algorithm 1, we can invert $I - U_i$ with entrywise accuracy to the machine precision. Similarly $I - A_1$ can be inverted accurately using the representation $I - A_1 = (A_1, A_0e + A_2e)$. This leads to the following refined implementation of Algorithm 2.

Algorithm 3: Refined Implementation of Latouche-Ramaswami Algorithm.
$v_0 = A_0\mathbf{e} + A_2\mathbf{e}$;
Solve $M_0H_0 = A_2$ by Algorithm 1 with $M_0 = (A_1, v_0)$;
Solve $M_0L_0 = A_0$ by Algorithm 1 with $M_0 = (A_1, v_0)$;
$G_0 = L_0$;
$T_0 = H_0$;
For $i = 0, 1, 2, \cdots$ until convergence
  $v_i = H_i^2\mathbf{e} + L_i^2\mathbf{e}$;
  $U_i = H_iL_i + L_iH_i$
  Solve $M_{i+1}H_{i+1} = H_i^2$ by Algorithm 1 with $M_{i+1} = (U_i, v_i)$;
  Solve $M_{i+1}L_{i+1} = L_i^2$ by Algorithm 1 with $M_{i+1} = (U_i, v_i)$;
  $G_{i+1} = G_i + T_iL_{i+1}$,
  $T_{i+1} = T_iH_{i+1}$

Each step of iteration uses the inverse of M_{i+1} twice but we only need to compute the LU factorization (Algorithm 1) once. Also, For positive recurrent queues, G is stochastic and, once computed, we can obtain the rate matrix R through $R = A_0(I - A_1 - A_0G)^{-1}$ (see [16, p.166]). Again, $I - A_1 - A_0G$ is a DDM and since $(I - A_1 - A_0G)e = A_2e$, it has the representation

$$I - A_1 - A_0G = (A_1 + A_0G, A_2e)$$

which involves additions only. Therefore, R can be computed from G entrywise to the machine precision.

We next present an error analysis to demonstrate the higher accuracy achieved.

Lemma 3 *Let $\widehat{H}_i$, $\widehat{L}_i$ and $\widehat{G}_i$ be the computed H_i, L_i and G_i of Algorithm 3 respectively. We have that $\widehat{H}_i \geq 0$ and $\widehat{L}_i \geq 0$ and hence $\widehat{G}_i$ is monotonically increasing.*

The proof simply follows from the fact that Algorithm 3 involves only addition operations of nonnegative numbers. The following result shows that the stochastic property of $H_i + L_i$ is preserved in finite precision.

Lemma 4 *Let $\widehat{H}_i$ and $\widehat{L}_i$ be the computed H_i and L_i of Algorithm 3. Then*

$$|(\widehat{H}_i + \widehat{L}_i)\mathbf{e} - \mathbf{e}| \leq (\psi(n)\epsilon + O(\epsilon^2))\mathbf{e}.$$

where $\psi(n) = \frac{2}{3}(n+2)(n+3)(2n+5) + 2n.$

Proof For $i \geq 0$, let $\widehat{U}_i$ and $\widehat{v}_i$ be the computed U_i and v_i and $\widehat{R}_i, \widehat{S}_i$ be the computed $\widehat{H}_i^2$ and $\widehat{L}_i^2$ respectively at step i. Then, using (6) and (5), we have

$$\widehat{v}_i = fl(\widehat{R}_i\mathbf{e} + \widehat{S}_i\mathbf{e}) = \widehat{R}_i\mathbf{e} + \widehat{S}_i\mathbf{e} + f, \tag{24}$$

where $|f| \leq 2n\epsilon(\widehat{R}_i\mathbf{e} + \widehat{S}_i\mathbf{e})$.

Now, let $M_{i+1} = (\widehat{U}_i, \widehat{v}_i)$, $X_{i+1} = M_{i+1}^{-1}\widehat{R}_i$ and $Y_{i+1} = M_{i+1}^{-1}\widehat{S}_i$. Note that $\widehat{H}_{i+1}$ and $\widehat{L}_{i+1}$ are the computed solutions to $M_{i+1}H_{i+1} = \widehat{S}_i$ and $M_{i+1}L_{i+1} = \widehat{T}_i$ respectively, while X_{i+1} and Y_{i+1} are respectively the exact solutions. Hence, by applying Theorem 2.1 to columns, we have

$$|X_{i+1} - \widehat{H}_{i+1}| \leq (\phi(n)\epsilon + O(\epsilon^2))X_{i+1}.$$

and

$$|Y_{i+1} - \widehat{L}_{i+1}| \leq (\phi(n)\epsilon + O(\epsilon^2))Y_{i+1}.$$

On the other hand, by the definition of M_{i+1}, we have $M_{i+1}\mathbf{e} = \widehat{v}_i$ or $M_{i+1}^{-1}\widehat{v}_i = \mathbf{e}$. Then, combining it with (24), we have

$$\begin{aligned}(X_{i+1} + Y_{i+1})\mathbf{e} &= M_{i+1}^{-1}(\widehat{R}_i + \widehat{S}_i)\mathbf{e} \\ &= M_{i+1}^{-1}\widehat{v}_i - M_{i+1}^{-1}f = \mathbf{e} - \tilde{f}\end{aligned}$$

where $|\tilde{f}| = |M_{i+1}^{-1}f| \leq 2n\epsilon M_{i+1}^{-1}(\widehat{R}_i\mathbf{e} + \widehat{S}_i\mathbf{e}) = 2n\epsilon(X_{i+1} + Y_{i+1})\mathbf{e}$. Thus,

$$\begin{aligned}|(\widehat{H}_{i+1} + \widehat{L}_{i+1})\mathbf{e} - \mathbf{e}| &\leq |(\widehat{H}_{i+1} - X_{i+1})\mathbf{e}| + |(\widehat{L}_{i+1} - Y_{i+1})\mathbf{e}| + |(X_{i+1} + Y_{i+1})\mathbf{e} - \mathbf{e}| \\ &\leq (\phi(n)\epsilon + O(\epsilon^2))(X_{i+1} + Y_{i+1})\mathbf{e} + |\tilde{f}| \\ &\leq (\psi(n)\epsilon + O(\epsilon^2))\mathbf{e}.\end{aligned}$$

This proves the cases $i \geq 1$ of the theorem. The case $i = 0$ can be proved similarly. □

The near stochastic property of $\widehat{H}_i + \widehat{T}_i$ improves the accuracy of the final approximation. To see this, let us consider again the positive recurrent case, in which the solution G is stochastic. The following theorem demonstrates that at convergence (i.e. when T_i is sufficiently small), the computed solution $\widehat{G}_i$ will be stochastic to the order of the machine precision.

Theorem 7 *Let $\widehat{G}_i$, $\widehat{T}_i$, $\widehat{H}_i$ and $\widehat{L}_i$ be the computed G_i, T_i, H_i and L_i, respectively, of Algorithm 3. Then*

$$|\mathbf{e} - \widehat{G}_i\mathbf{e}| \le \widehat{T}_i\mathbf{e} + (\zeta(n,i)\epsilon + O(\epsilon^2))\mathbf{e}$$

where $\zeta(n,i) = \frac{2}{3}(i+1)(n+2)(n+3)(2n+5) + (i+2)^2(n+1)$.

Proof Using (7), we have $\widehat{G}_i = fl(\widehat{G}_{i-1} + \widehat{T}_{i-1}\widehat{L}_i) = \widehat{G}_{i-1} + \widehat{T}_{i-1}\widehat{L}_i + F_i$, where $|F_i| \le ((n+1)\epsilon + O(\epsilon^2))\widehat{G}_i$. Thus

$$\widehat{G}_i = \widehat{L}_0 + \widehat{T}_0\widehat{L}_1 + \cdots + \widehat{T}_{i-1}\widehat{L}_i + F_1 + \cdots + F_i$$

Write $E_i = F_1 + \cdots + F_i$. Then $|E_i| \le (i(n+1)\epsilon + O(\epsilon^2))\widehat{G}_i$, where we note that $\widehat{G}_i$ is monotonic. On the other hand, using (6),

$$\widehat{T}_i = fl(\widehat{T}_{i-1}\widehat{H}_i) = \widehat{T}_{i-1}\widehat{H}_i + F_i^{(2)} \tag{25}$$

with $|F_i^{(2)}| \le (n\epsilon + O(\epsilon^2))\widehat{T}_i$. Furthermore, by Lemma 4, $\widehat{H}_i\mathbf{e} \le (1+O(\epsilon))\mathbf{e}$ and hence $\widehat{T}_i\mathbf{e} \le (1+O(\epsilon))\widehat{T}_{i-1}\mathbf{e} + O(\epsilon)\widehat{T}_i\mathbf{e}$. This implies $\widehat{T}_i\mathbf{e} \le (1+O(\epsilon))T_{i-1}\mathbf{e}$ and thus $\widehat{T}_i\mathbf{e} \le (1+O(\epsilon))\mathbf{e}$. Using the expression of $\widehat{G}_i$ and Lemma 4, we also have that $\widehat{G}_i\mathbf{e} \le (1+O(\epsilon))(\mathbf{e} + \widehat{T}_0\mathbf{e} + \cdots + \widehat{T}_{i-1}\mathbf{e})$. Hence,

$$\begin{aligned}
\widehat{G}_i\mathbf{e} &= \widehat{L}_0\mathbf{e} + \widehat{T}_0\widehat{L}_1\mathbf{e} + \cdots + \widehat{T}_{i-1}\widehat{L}_i\mathbf{e} + E_i\mathbf{e} \\
&= \widehat{L}_0\mathbf{e} + \Sigma_{j=0}^{i-1}(\widehat{T}_j\widehat{L}_{j+1}\mathbf{e} + \widehat{T}_j\widehat{H}_{j+1}\mathbf{e} - \widehat{T}_j\mathbf{e}) - \Sigma_{j=0}^{i-1}\widehat{T}_j\widehat{H}_{j+1}\mathbf{e} + \Sigma_{j=0}^{i-1}\widehat{T}_j\mathbf{e} + E_i\mathbf{e} \\
&= \widehat{L}_0\mathbf{e} + \widehat{T}_0\mathbf{e} + \Sigma_{j=0}^{i-1}\widehat{T}_j(\widehat{L}_{j+1}\mathbf{e} + \widehat{H}_{j+1}\mathbf{e} - \mathbf{e}) + \Sigma_{j=1}^{i-1}(\widehat{T}_j - \widehat{T}_{j-1}\widehat{H}_j)\mathbf{e} - \widehat{T}_{i-1}\widehat{H}_i\mathbf{e} + E_i\mathbf{e}.
\end{aligned}$$

Noting that $\widehat{T}_0 = \widehat{H}_0$, it follows from (25) and Lemma 4 that

$$\begin{aligned}
|\mathbf{e} - \widehat{G}_i\mathbf{e}| &\le \widehat{T}_{i-1}\widehat{H}_i\mathbf{e} + (\psi(n)\epsilon + O(\epsilon^2))(\mathbf{e} + \widehat{T}_0\mathbf{e} + \widehat{T}_1\mathbf{e} + \cdots + \widehat{T}_{i-1}\mathbf{e}) + \\
&\quad + (n\epsilon + O(\epsilon^2))(\widehat{T}_1\mathbf{e} + \cdots + \widehat{T}_{i-1}\mathbf{e}) + (i(n+1)\epsilon + O(\epsilon^2))\widehat{G}_i\mathbf{e} \\
&\le \widehat{T}_{i-1}\widehat{H}_i\mathbf{e} + (\zeta(n,i)\epsilon + O(\epsilon^2))\mathbf{e}.
\end{aligned}$$

□

The coefficient $\zeta(n,i)$ in the bound depends on n^3, which is inherited from the worst case bound of Theorem 1 and seems to be far too pessimistic in practice (see our numerical examples and those in [1]). On the other hand, the dependence of $\zeta(n,i)$ on i arises in the summation of i terms in the series G_i. This effect can be reduced by implementing a group wise update of the partial sum, namely, updating G_i only when the accumulated T_iL_{i+1} is sufficiently large. However, given the fast convergence of the algorithm (i.e. small i), we have not found this necessary in practice.

5 Numerical Examples

In this section, we present some numerical examples to demonstrate the high accuracy achieved by the refined algorithm. All testing is carried out in MATLAB with the machine precision $\epsilon \approx 10^{-16}$.

Example 1: This example is a QBD process with two phases, which is used by Latouche and Ramaswami [15]. From the state $(i,1)$, $i \geq 1$, the chain moves to $(i,2)$ with probability p and to $(i-1,1)$ with probability $1-p$; from the state $(i,2)$, $i \geq 0$, the chain moves to $(i,1)$ with probability $2p$ and to $(i+1,2)$ with probability $1-2p$. The QBD is always positive recurrent with

$$G = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}.$$

We compare the solution error $\|G - G_i\|_\infty$ and the residual error $\tau_i \equiv \|A_2 + A_1 G_i + A_0 G_i^2 - G_i\|$ for the direct implementation (Algorithm 2) and the refined implementation (Algorithm 3). We present the results in Table 1 for a range of value of p. The termination criterion is set to be $\|T_i L_{i+1}\|_\infty < 10^{-16}$. The two implementations take the same number of iterations.

	Algorithm 2		Algorithm 3	
p	$\|G - G_i\|_\infty$	τ_i	$\|G - G_i\|_\infty$	τ_i
10^{-2}	$1e-14$	$1e-16$	$1e-16$	0
10^{-4}	$7e-13$	0	0	0
10^{-6}	$5e-11$	$1e-16$	$3e-16$	$2e-16$
10^{-8}	$7e-09$	0	$2e-16$	$2e-16$
10^{-10}	$3e-07$	$2e-16$	$1e-15$	$1e-16$
10^{-12}	$6e-05$	$2e-16$	$2e-16$	0
10^{-14}	$2e-03$	0	$6e-16$	0
10^{-16}	$2e-01$	$1e-16$	$1e-16$	$1e-16$

Table 1: Example 1: $\tau_i = \|A_2 + A_1 G_i + A_0 G_i^2 - G_i\|_\infty$

From the table, we see that the refined implementation (Algorithm 3) leads to a solution with both the solution error and the residual error at the level of the machine precision. For the direct implementation (Algorithm 2), as p decreases, the solution error increases steadily. However, the residual error $\|A_2 + A_1 G_i + A_0 G_i^2 - G_i\|$ is in the order of the machine precision, which was originally observed by Latouche and Ramaswami [15]. This is probably true in general and due to the backward stability of the standard algorithm; but we do not have an error analysis demonstrating this for Algorithm 2 (see [23] for an analysis of residual errors for other algorithms). With this, we note that G_i obtained by Algorithm 2 nearly satisfies the matrix equation but it may be far from being stochastic.

Example 2: This example is a continuous QBD model of a teletraffic system taken from Daigle and Lucantoni [8]. It is defined by 24×24 blocks A'_0, A'_1 and A'_2 in which A'_0, A'_2 are diagonal such that $A'_0 = 192\rho_d I$, $(A'_2)_{jj} = 192(1 - j/24)$ for $0 \leq j \leq 23$, and A'_1 is tridiagonal such that $(A'_1)_{j,j+1} = ar(M-j)/M$ (for $0 \leq j \leq 22$) and $(A'_1)_{j,j-1} = jr$ (for $1 \leq j \leq 23$). The corresponding discrete model equation (2) is given by $A_0 = -{A'_1}^{-1} A'_2$, $A_1 = 0$ and $A_2 = -{A'_1}^{-1} A'_0$ (in our notation). We use the parametric values $r = 1/300$, $a = 18.244$ and a few values of M for our testing. The stopping tolerance is again set as $\|T_i L_{i+1}\|_\infty < 10^{-16}$.

We present the results in Table 2 for a range of M. We compare the stochastic measure $\|e - G_i e\|_\infty$, and the residual error τ_i. A behavior similar to Example 1 is observed. In particular, the refined implementation produces G_i that satisfies to the machine precision both the matrix equation and the stochastic property. On the other hand, the solution by the standard implementation may be far from being stochastic.

	Algorithm 2		Algorithm 3	
M	$\|\mathbf{e} - G_i\mathbf{e}\|_\infty$	τ_i	$\|\mathbf{e} - G_i\mathbf{e}\|_\infty$	τ_i
64	$1e-12$	$1e-16$	$6e-16$	$3e-16$
256	$2e-11$	$2e-16$	$4e-16$	$3e-16$
1024	$6e-11$	$2e-16$	$6e-16$	$3e-16$
4096	$1e-10$	$1e-16$	$6e-16$	$3e-16$
16384	$1e-09$	$2e-16$	$5e-16$	$3e-16$
65536	$2e-08$	$1e-16$	$5e-16$	$3e-16$

Table 2: Example 2: (for $M = 65536$) $\tau_i = \|A_2 + A_1 G_i + A_0 G_i^2 - G_i\|_\infty$

To show the convergence behavior and its relation to the stochastic property of $L_i + H_i$, we also present in Figure 1 convergence plots of $\delta_i = \|\mathbf{e} - G_i\mathbf{e}\|_\infty$ (solid line for Alg. 2 and +-line for Alg. 3) and the stochastic measure $\|\mathbf{e} - (L_i + H_i)\mathbf{e}\|_\infty$ (dotted line for Alg. 2 and dash-dotted line for Alg. 3). We see that initially (first 10 steps), the convergence is nearly linear. The stochastic property of $L_i + H_i$ is graduately lost for Algorithm 2, which eventually limits the convergence of G_i. For Alg. 3, $\|\mathbf{e} - (L_i + H_i)\mathbf{e}\|_\infty$ remains at the machine precision level (see Lemma 4) and that ensures $G - G_i$ converges to the machine precision.

References

[1] S.Alfa, J.Xue and Q.Ye, *Accurate computation of the smallest eigenvalue of a diagonally dominant M-matrix,* Math. Comp. (to appear).

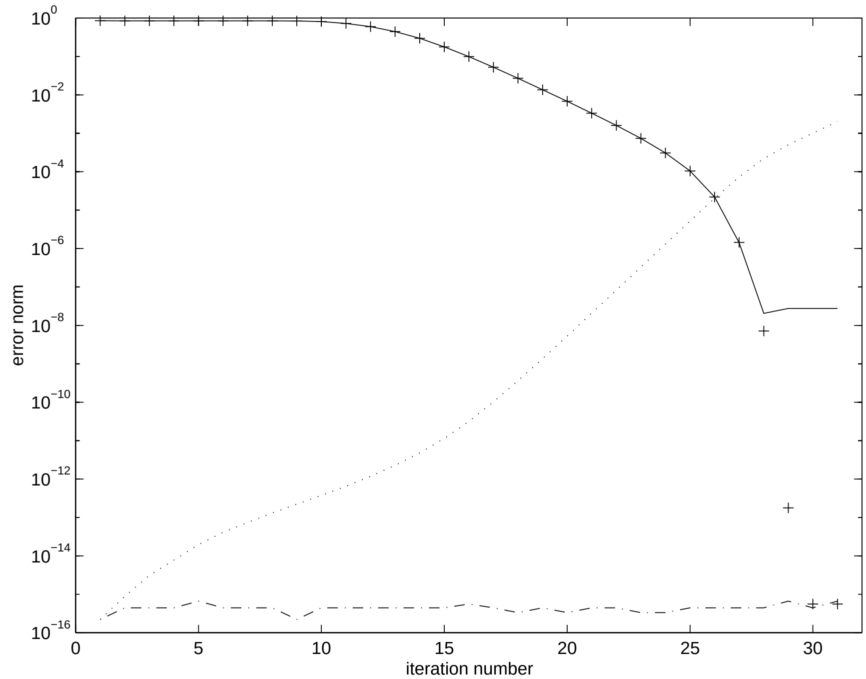
[2] Z.-Z. Bai, *A class of iteration methods based on the Moser formula for nonlinear equations in Markov chains,* Linear Alg. Appl., 266(1997):219-241.

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

[4] D.A. Bini and B. Meini, *Improved cyclic reduction for solving queueing problems,* Numer. Alg. 5(1997):57-74.

[5] D. Bini, B. Meini and V. Ramaswami, *Analyzing M/G/1 paradigms through QBDs: the role of the block structure in computing the matrix G*, in Advances in Algorithmic Methods for Stochastic Models - Proceedings of MAM3, G. Latouche and P.G. Taylor (Editors), 2000, Notable Publications Inc. NJ. pp. 73-86.

Figure 1: Example 2: solid: $\|\mathbf{e} - G_i\mathbf{e}\|_\infty$ for Alg. 2; +: $\|\mathbf{e} - G_i\mathbf{e}\|_\infty$ for Alg. 3; dot: $\|\mathbf{e} - (L_i + H_i)\mathbf{e}\|_\infty$ for Alg. 2; dash-dot: $\|\mathbf{e} - (L_i + H_i)\mathbf{e}\|_\infty$ for Alg. 3



[6] A. Berman and R. Plemmons, *Nonnegative Matrices in Mathematical Science,* Academic Press, New York,1979.

[7] L. Bright and P. G. Taylor, *Calculating the equilibrium distribution in level dependent quasi-birth-and-death processes*, Comm. Statist. Stochastic Models, 11(1995):497-525,

[8] J. Daigle and D. Lucantoni, *Queueing systems having phase-dependent arrival and service rates*, In *Numerical Solution of Markov Chains*, ed. W.J. Stewart, Marcel Dekker, New York, pp. 161-202.

[9] P. Favati and B. Meini, *Relaxed functional iteration techniques for the numerical solution of $M/G/1$ type Markov chains*, BIT, 38(1998): 510-526.

[10] G. Golub and C. Van Loan, *Matrix Computations,* 2nd ed., The Johns Hopkins University Press, Baltimore, MD,1989.

[11] C.-H. Guo, *On the numerical solution of a nonlinear matrix equation in Markov chains,* Linear Alg. Appl., 288(1999):175-186.

[12] C.-H. Guo, *Convergence analysis of the Latouche-Ramaswami algorithm for null recurrent quasi-birth-death processes,* Preprint.

[13] W.K. Grassmann, M.J. Taksar and D.P. Heyman, *Regenerative analysis and steady-state distributions for Markov chains,* Oper. Res., 33(1985):1107-1116.

[14] G. Latouche, *Newton's iteration for nonlinear equations in Markov chains*, IMA J. Numer. Anal. 14(1994):583-598.

[15] G. Latouche and V. Ramaswami, *A logarithmic reduction algorithm for quasi-birth-death processes,* J. Appl. Prob. 30(1993):650-674.

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

[17] D. M. Lucantoni and V. Ramaswami , *Efficient algorithms for solving the nonlinear matrix equations arising in phase type queues*, Comm. Statist. Stochastic Models, 1(1985):29-51.

[18] B. Meini, *New convergence results on functional iteration techniques for the numerical solution of M/G/1 type Markov chains,* Numer. Math. 78(1997):39-58.

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

[20] V. Ramaswami, *Nonlinear matrix equations in applied probability - Solution techniques and open problems*, SIAM Review, 30(1988):256-263.

[21] V. Ramaswami, and P. G. Taylor , *Some properties of the rate operators in level dependent quasi-birth-and-death processes with a countable number of phases*, Comm. Statist. Stochastic Models, 12(1996):143-164.

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

[23] Q. Ye, *High accuracy algorithms for solving nonlinear matrix equations in queueing models*, in Advances in Algorithmic Methods for Stochastic Models - Proceedings of MAM3, G. Latouche and P.G. Taylor (Editors), 2000, Notable Publications Inc. NJ. pp. 401-415.