On functional iteration methods for solving nonlinear matrix equations arising in queueing problems

Paola Favati* Beatrice Meini†

Abstract

The problem of the computation of the minimal nonnegative solution G of the nonlinear matrix equation $X = \sum_{i=0}^{+\infty} X^i A_i$ is considered. This problem arises in the numerical solution of M/G/1 type Markov chains, where A_i, $i \geq 0$, are nonnegative $k \times k$ matrices such that $\sum_{i=0}^{+\infty} A_i$ is column stochastic. We analyze classical functional iteration methods, by estimating the rate of convergence, in relation with spectral properties of the starting approximation matrix X_0. Based on these new convergence results, we propose an effective method to choose a matrix X_0, which drastically reduces the number of iterations; the additional cost needed to compute X_0 is much less than the overall savings achieved by reducing the number of iterations.

1 Introduction

Many queueing problems can be modeled by Markov chains of M/G/1 type, that is, Markov chains whose probability transition matrix has the structure

$$P^t = \begin{pmatrix} B_1 & A_0 & & \bigcirc & \\ B_2 & A_1 & A_0 & & \\ B_3 & A_2 & A_1 & A_0 & \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix}, \tag{1}$$

where B_{i+1}, A_i, $i \geq 0$, are $k \times k$ nonnegative matrices such that $\sum_{i=1}^{+\infty} B_i$ and $\sum_{i=0}^{+\infty} A_i$ are column stochastic. One of the major problems in Markov chains is the computation of the probability invariant vector $\boldsymbol{\pi}$, i.e., the solution of

*Istituto di Matematica Computazionale del C.N.R., via S. Maria 46, 56127 Pisa, Italy. E-mail: favati@imc.pi.cnr.it

†Dipartimento di Matematica, via Buonarroti 2, 56127 Pisa, Italy. E-mail: meini@dm.unipi.it

the infinite system $\boldsymbol{\pi} = P^t\boldsymbol{\pi}$, $\|\boldsymbol{\pi}\|_1 = 1$. For the matrices of structure (1), the computation of $\boldsymbol{\pi}$ can be reduced (compare [16]) to the computation of the minimal nonnegative solution G of the nonlinear matrix equation

$$X = \sum_{i=0}^{+\infty} X^i A_i, \tag{2}$$

where X is a $k \times k$ matrix. Indeed, once G is known, the vector $\boldsymbol{\pi}$ can be recovered by means of the Ramaswami formula [17]. Moreover, the knowledge of the matrix G is very important, due to its role in calculating performance measures [16].

Recently, the analysis of numerical methods for computing the solution of the matrix equation (2) has been largely developed. New quadratically convergent methods have been proposed in [12], [13], [11], [3], [4], [5]; classical linearly convergent methods based on functional iterations have been improved and studied in details in [8], [14]. The latter methods, unlike most of the quadratically convergent methods, are very easy to implement. They are based on the recursion

$$X_{n+1} = F(X_n), \quad n \geq 0, \tag{3}$$

where

$$F(X) = K(X)(I - H(X))^{-1}, \tag{4}$$

and $H(X) = \sum_{i=0}^{+\infty} X^i H_i$, $K(X) = \sum_{i=0}^{+\infty} X^i K_i$ are such that $H_i \geq O$, $K_i \geq O$, $i \geq 0$, and $\sum_{i=0}^{+\infty} z^i A_i = z\sum_{i=0}^{+\infty} z^i H_i + \sum_{i=0}^{+\infty} z^i K_i$ [8], [14]. It can be proved that, by suitable choosing the starting approximation X_0, the sequence (3) linearly converges to the solution G of (2); moreover, the rate of convergence is strongly related to the choice of X_0. More precisely, we associate with any function $F(\cdot)$ of (4) a suitable nonnegative $k^2 \times k^2$ matrix $\widehat{R}_F$. In the case where $X_0 = 0$ in [14] it is shown that the asymptotic rate of convergence of $\{X_n\}$ coincides with the spectral radius of $\widehat{R}_F$. Moreover, in the case where X_0 is any column stochastic matrix, it is shown that at each step n, the vector $\hat{\mathbf{e}}_n$, obtained by the column-wise arrangement of the entries of the matrix $G - X_n$, is orthogonal to the nonnegative left eigenvector $\mathbf{v}^T$ associated with the spectral radius of the matrix $\widehat{R}_F$. From this property, it follows that the rate of convergence of X_n is less than or equal to the spectral radius of the matrix M_1 obtained by projecting $\widehat{R}_F$ onto a suitable subspace $\mathcal{V}_1$.

From the above results it follows that, if the matrix $\widehat{R}_F$ is irreducible, then the mean asymptotic rate of convergence of the sequence X_n is:

– equal to the spectral radius of $\widehat{R}_F$, if $X_0 = O$;

– less than or equal to the second largest modulus eigenvalue of $\widehat{R}_F$, if X_0 is a column stochastic matrix.

The better rate of convergence, when X_0 is a stochastic matrix, is mainly due to the fact that both G and X_0 have the eigenvalue 1 and associated left eigenvector $\mathbf{e}^t = (1, \ldots, 1)$.

In this paper, in order to further improve the rate of convergence, we consider starting approximations X_0 that share with G more eigenvalues, and the corresponding left eigenvectors, in addition to the eigenvalue 1 and the eigenvector $\mathbf{e}^t$.

More precisely, let $\lambda_1 = 1, \lambda_2, \ldots, \lambda_h$, $h \leq k$, be eigenvalues of the matrix G such that the corresponding left eigenvectors $\mathbf{w}_1^t = \mathbf{e}^t, \mathbf{w}_2^t, \ldots, \mathbf{w}_h^t$ are linearly independent. Suppose that the matrix X_0 satisfies

$$\mathbf{w}_i^t X_0 = \lambda_i \mathbf{w}_i^t, \ i = 1, \ldots, h, \tag{5}$$

that is, it shares with G the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_h$ and the corresponding left eigenvectors. Then we prove that the rate of convergence of the sequence X_n is bounded from above by $\rho(M_h)$, where $\rho(A)$ denotes the spectral radius of the matrix A and M_h is the matrix obtained by projecting $\widehat{R}_F$ onto a suitable subspace $\mathcal{V}_h$. Moreover, we show that $\rho(M_h) \leq \rho(M_1)$, for any $h \in \{1, \ldots, k\}$. In this way the rate of convergence can be substantially reduced.

Once the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_h$, and the corresponding eigenvectors $\mathbf{w}_1^t, \ldots, \mathbf{w}_h^t$ are known, a matrix X_0 satisfying (5) can simply be obtained by means of the equation $X_0 = V Diag(\lambda_1, \ldots, \lambda_h) W^t$, where W is the $k \times h$ matrix whose columns are the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_h$, and V is the generalized inverse of W^t. When $h = k$, i.e., all the eigenvalues, and the corresponding left eigenvectors, are known, the matrix V coincides with the inverse of W^t. In this case some authors [9], [15] have proposed to approximate G by simply computing $W^{-t} Diag(\lambda_1, \ldots, \lambda_k) W^t$. In practice, due to the roundoff errors generated in the computation, the matrix $\widetilde{G}$ obtained in this way can be a poor approximation of G; we propose to improve this approximation by choosing $X_0 = \widetilde{G}$ and by applying few steps of the functional iteration method (3).

We have performed several numerical experiments. We have first computed eigenvalues and corresponding left eigenvectors, for different values of h. Then a matrix X_0 satisfying (5) has been computed and functional iteration has been applied. From the many results, we have observed a drastic reduction of the number of iterations needed to approximate G when $h > 1$, with respect to the case where $h = 1$, or X_0 is a generic column stochastic matrix. On the other hand, the time needed to approximate the eigenvalues, and the corresponding left eigenvectors, when they are not known, is negligible with respect to the time needed to perform functional iterations.

The paper is organized as follows: in Section 2 we prove the main convergence results, by estimating the rate of convergence, in relation with the starting approximation X_0; in Section 3 we propose a method for computing a matrix X_0 satisfying (5); in Section 4 we present numerical results and comparisons.

2 Convergence rate and starting approximation

In this section we extend the convergence results proved in [14] for functional iteration methods based on the recursion (3), for solving the nonlinear matrix equation (2); these new convergence results suggest how to choose the starting approximation matrix X_0 in order to drastically reduce the number of iterations needed to approximate the solution G.

Let us denote by $\hat{\mathbf{e}}_n$ the vector obtained by column-wise arranging the entries of the matrix $G - X_n$. It can be shown [14] that the error $\hat{\mathbf{e}}_{n+1}$ is related to the error $\hat{\mathbf{e}}_n$ by means of the recursion

$$\hat{\mathbf{e}}_{n+1} = \widehat{R}_n \hat{\mathbf{e}}_n \tag{6}$$

where

$$\widehat{R}_n = \sum_{j=0}^{+\infty} Y_{j,n} \otimes G^j \tag{7}$$

and

$$\begin{aligned} Y_{0,n} &= \textstyle\sum_{i=1}^{+\infty} \left(X_n^{i-1} K_i (I - H(X_n))^{-1} \right)^t \\ Y_{j,n} &= \textstyle\sum_{i=j}^{+\infty} \left(X_n^{i-j} A_{i+1} (I - H(X_n))^{-1} \right)^t, \quad j \geq 1. \end{aligned}$$

Denote by

$$\widehat{R}_F = \sum_{j=0}^{+\infty} Y_j \otimes G^j$$

where

$$\begin{aligned} Y_0 &= \textstyle\sum_{i=1}^{+\infty} \left(G^{i-1} K_i (I - H(G))^{-1} \right)^t \\ Y_j &= \textstyle\sum_{i=j}^{+\infty} \left(G^{i-j} A_{i+1} (I - H(G))^{-1} \right)^t, \quad j \geq 1. \end{aligned}$$

Observe that, if $\lim_n X_n = G$, then $\widehat{R}_F = \lim_n \widehat{R}_n$. In the following we denote by $\mathbf{e}_i$, $i = 1, \ldots, k$, the k-dimensional vector having the i-th entry equal to 1, and the remaining ones equal to 0.

Relation (6) allows us to prove the following convergence result [14]:

Theorem 1 *Let*

$$r = \lim_n \sqrt[n]{\|\hat{\mathbf{e}}_n\|}$$

be the mean asymptotic rate of convergence of the sequence $X_{n+1} = F(X_n)$, where $F(\cdot)$ is defined by (4) and $\|\cdot\|$ is any vector norm. The following results hold:

– if $X_0 = 0$ then $r = \rho(\widehat{R}_F)$;

– if X_0 is a stochastic matrix then $r \leq \rho(T_1^t \widehat{R}_F T_1)$, where T_1 is the $k^2 \times (k^2 - k)$ matrix whose columns complete the set of vectors

$$\frac{1}{\sqrt{k}} (\mathbf{e}_i \otimes \mathbf{e}), \quad i = 1, \ldots, k, \tag{8}$$

to an orthonormal basis of $\mathbf{R}^{k^2}$.

The reduction of the rate of convergence, in the case where X_0 is a column stochastic matrix, is mainly due to the fact that X_0 has the eigenvalue 1 and associated left eigenvector $\mathbf{e}^t$, as well as the matrix G. In fact, from this property it follows that the vector $\hat{\mathbf{e}}_n$ is orthogonal to the nonnegative left eigenvector associated with the spectral radius of the matrix $\widehat{R}_F$.

This fact suggests us to choose a starting approximation X_0 that shares with G some other eigenvalues and the corresponding left eigenvectors, in addition to the eigenvalue equal to 1 and the eigenvector $\mathbf{e}^t$.

Denote by $\Lambda = \{\lambda_1, \ldots, \lambda_k\}$ the set of the eigenvalues of the matrix G, and with $\mathbf{w}_1^t, \ldots, \mathbf{w}_k^t$ the corresponding left eigenvectors, where $\lambda_1 = 1$ and $\mathbf{w}_1^t = \mathbf{e}^t$. Let Λ_h a subset of h eigenvalues of Λ, $1 \leq h \leq k$, such that:

1. $\lambda_1 \in \Lambda_h$;
2. if $\lambda \in \Lambda_h$, then $\bar{\lambda} \in \Lambda_h$, where $\bar{\lambda}$ denotes the complex conjugate of λ;
3. the left eigenvectors corresponding to the eigenvalues in Λ_h are linearly independent.

Throughout, we suppose, without loss of generality, that $\Lambda_h = \{\lambda_1, \ldots, \lambda_h\}$.

Denote by $\mathcal{U}_h$ the subspace generated by the k^2–dimensional vectors $\mathbf{e}_i \otimes \mathbf{w}_j$, $i = 1, \ldots, k$, $j = 1, \ldots, h$, and let $\mathcal{V}_h = \mathcal{U}_h^{\perp}$. Let $S_h = I \otimes E_h$ a $k^2 \times hk$ matrix whose columns are an orthonormal basis of $\mathcal{U}_h$, and let $T_h = I \otimes F_h$ a $k^2 \times (k^2 - hk)$ matrix whose columns are an orthonormal basis of $\mathcal{V}_h$. Then the following result holds:

Theorem 2 *Let X_n, $n \geq 0$, be a convergent sequence generated by $X_{n+1} = F(X_n)$, where $F(\cdot)$ is defined by (4). Denote by*

$$r = \lim_n \sqrt[n]{\|\hat{\mathbf{e}}_n\|}$$

the mean asymptotic rate of convergence, where $\|\cdot\|$ is any vector norm. Let $1 \leq h \leq k$ and $\Lambda_h \subset \Lambda$ satisfying properties 1, 2, 3. Suppose that X_0 verifies

$$\mathbf{w}^t X_0 = \lambda \mathbf{w}^t, \tag{9}$$

for any $\lambda \in \Lambda_h$, where $\mathbf{w}^t$ is the corresponding left eigenvector of G. Then it holds $r \leq \rho(T_h^t \widehat{R}_F T_h)$.

Proof. First observe that $\hat{\mathbf{e}}_n \in \mathcal{V}_h$, for $n \geq 0$. In fact, for $n = 0$, this property follows from the relations $\mathbf{w}^t(G - X_0) = \lambda \mathbf{w}^t - \lambda \mathbf{w}^t = 0$, for any $\lambda \in \Lambda_h$; for $n \geq 1$, the property follows from (6) and from the fact that, if $\mathbf{u} \in \mathcal{U}_h$, then $\mathbf{u}^t \widehat{R}_n \in \mathcal{U}_h$, for any $n \geq 0$.

Let $P = (S_h | T_h)$. From (6) it follows that

$$P^t \hat{\mathbf{e}}_{n+1} = \begin{pmatrix} \mathbf{0} \\ \mathbf{f}_{n+1} \end{pmatrix} = (P^t \widehat{R}_n P)(P^t \hat{\mathbf{e}}_n) = (P^t \widehat{R}_n P) \begin{pmatrix} \mathbf{0} \\ \mathbf{f}_n \end{pmatrix}, \tag{10}$$

where $\mathbf{f}_n = T_h^t \mathbf{e}_n$. Since $S_h^t \widehat{R}_n T_h = O$, it follows that

$$P^t \widehat{R}_n P = \begin{pmatrix} S_h^t \\ - \\ T_h^t \end{pmatrix} \widehat{R}_n \ (S_h | T_h) = \begin{pmatrix} V_n & O \\ T_n & Z_n \end{pmatrix}, \tag{11}$$

where $Z_n = T_h^t \widehat{R}_n T_h$. From (10) and (11), the recurrence (6) can be rewritten in the form

$$\mathbf{f}_{n+1} = Z_n \mathbf{f}_n \tag{12}$$

and

$$r = \lim_n \sqrt[n]{\|\hat{\mathbf{e}}_n\|} = \lim_n \sqrt[n]{\|\mathbf{f}_n\|}.$$

By following the same lines of the proof of Theorem 15 in [14], we easily arrive at $r \leq \rho(T_h^t \widehat{R}_F T_h)$. □

From the above theorem it follows that, by starting with a matrix X_0 sharing with G more than one eigenvalue, the rate of convergence can be reduced. In fact, the following result holds:

Theorem 3 *Under the hypotheses of Theorem 2, the eigenvalues of the matrix $T_h^t \widehat{R}_F T_h$ are given by the set*

$$\bigcup_{i=h+1}^{k} \{\mu : \mu \text{ is eigenvalue of } Q_i\}$$

where $Q_i = \sum_{j=0}^{\infty} \lambda_i^j Y_j$, $i = 1, \ldots, k$.

Proof. Observe that the matrix $Z_F = T_h^t \widehat{R}_F T_h$ can be written as

$$Z_F = (I \otimes F_h^t) \widehat{R}_F (I \otimes F_h) = \sum_{j=0}^{\infty} Y_j \otimes F_h^t G^j F_h.$$

Whence, the matrix Z_F is similar to the matrix $\sum_{j=0}^{\infty} F_h^t G^j F_h \otimes Y_j$, which is similar to $\sum_{j=0}^{\infty} U^j \otimes Y_j$, where U is the Schur canonical form of $F_h^t G F_h$. Since the eigenvalues of $F_h^t G F_h$ are $\lambda_{h+1}, \ldots, \lambda_k$, U is an upper triangular matrix whose diagonal entries are $\lambda_{h+1}, \ldots, \lambda_k$. In this way we find that Z_F is similar to a block upper triangular matrix whose diagonal blocks are the matrices Q_i, $i = h+1, \ldots, k$. □

From the above theorem it follows that:

Corollary 4 *Under the hypotheses of Theorem 2 it holds that*

$$\rho(T_h^t \widehat{R}_F T_h) \leq \rho(T_1^t \widehat{R}_F T_1)$$

for any $1 \leq h \leq k$.

In practice, we are not able to establish in advance the optimal number of eigenvalues/eigenvectors h to be shared in order to obtain the best rate of convergence. The optimal number h is strongly related to the spectral properties of G, and thus to the specific problems.

In the above theorems we have presented conditions on the starting matrix X_0 to increase the rate of convergence of the sequence $X_{n+1} = F(X_n)$, under the hypothesis that X_0 is such that the sequence X_n is convergent. In the case where the matrix X_0 is stochastic, the sequence is obviously convergent. In the more general case where X_0 is not necessarily nonnegative, the local convergence is guaranteed by the classical results on functional iteration methods [18], since $\rho(\widehat{R}_F) < 1$.

3 Choosing the starting approximation

Based on the results of Section 2, in order to find a good starting matrix X_0, we need first to know eigenvalues and left eigenvectors of the matrix G. The eigenvalues $\lambda_1, \ldots, \lambda_k$ of G are given by the zeros of the function

$$\varphi(z) = \det(zI - \sum_{j=0}^{+\infty} A_j z^j) \tag{13}$$

lying in the closed unit circle of the complex plane (compare [9]), and the left eigenvectors are the solutions of the linear systems

$$\mathbf{w}_i^t(\lambda_i I - \sum_{j=0}^{+\infty} A_j \lambda_i^j) = 0, \ i = 1, \ldots, k. \tag{14}$$

The zeros of the function $\varphi(z)$ can be computed by means of any algorithm for approximating the zeros of a polynomial, for example by means of Aberth's method [7], [1]; the left eigenvectors can be computed by solving the above homogeneous linear systems.

In the applications, it may be convenient to approximate only the eigenvalues that can be computed in a few numbers of iterations, and then to compute the corresponding matrix X_0.

Suppose we know a subset $\Lambda_h = \{\lambda_1, \ldots, \lambda_h\}$ of Λ satisfying properties 1, 2, 3 of Section 2; denote by $\mathbf{w}_1^t, \ldots, \mathbf{w}_h^t$ the corresponding left eigenvectors. Let W_h be the $k \times h$ matrix whose columns are the vectors $\mathbf{w}_1, \ldots, \mathbf{w}_h$ and let V_h be a $k \times h$ matrix such that

$$V_h^t W_h = I_h, \tag{15}$$

where I_h is the h dimensional identity matrix. Then, the matrix

$$X_0 = V_h Diag(\lambda_1, \ldots, \lambda_h) W_h^t \tag{16}$$

verifies conditions (9). A simple way to obtain a matrix V_h satisfying (15) is to compute the generalized inverse of W_h^t, say by means of QR or SVD method [10], or by means of the relation $V_h = (W_h W_h^t)^{-1} W_h$, that holds since W_h is a full rank matrix. Since $\lambda \in \Lambda_h$ implies $\bar{\lambda} \in \Lambda_h$, the matrix X_0 obtained in this way is real. If, in addition, V_h is such that

$$V_h Diag(\lambda_1, \ldots, \lambda_h) W_h^t \geq 0, \tag{17}$$

then X_0 is stochastic, since $X_0 \geq 0$ and $\mathbf{w}_1^t X_0 = \mathbf{w}_1^t$, and $\mathbf{w}_1^t = \mathbf{e}^t$.

A matrix V_h satisfying both conditions (15) and (17) can be viewed as a solution of a linear programming problem. However the computation of a solution, if it exists, can require a long time. On the other hand, in many numerical experiments, even if the starting matrix X_0, obtained by means of the generalized inverse, has negative entries after few steps the matrix X_n is nonnegative. Whence, the use of generalized inverse, leads to an advantageous choice of X_0.

Suppose we know all the eigenvalues $\lambda_1, \ldots, \lambda_k$ of G and the corresponding left eigenvectors $\mathbf{w}_1^t, \ldots, \mathbf{w}_k^t$. If the above eigenvectors are linearly independent, then it is well known [9], [15] that an approximation of the matrix G can be obtained by means of the relation

$$G = W_k^{-t} Diag(\lambda_1, \ldots, \lambda_k) W_k^t. \tag{18}$$

However, due to errors generated in the computation, the matrix obtained by calculating $W_k^{-t} Diag(\lambda_1, \ldots, \lambda_k) W_k^t$ can be a not sufficiently accurate approximation of G. In this case, in order to improve the approximation of G, we apply the functional iteration method, starting with $X_0 = W_k^{-t} Diag(\lambda_1, \ldots, \lambda_k) W_k^t$. In most applications, the matrix X_0 is nonnegative and generates a sequence X_n which rapidly converges to G.

4 Numerical results

We have performed numerical experiments in order to test the behavior of the rate of convergence with different starting approximations X_0. Since the matrix X_0 of (16) is not necessarily nonnegative, the matrices X_n could have negative entries and the matrices $I - H(X_n)$ of (4) could be singular, for small values of n. Thus, in order to avoid possibly singular matrix inversions, we use the standard functional iteration method $X_{n+1} = \sum_{i=0}^{+\infty} X_n^i A_i$, $n \geq 0$.

We have chosen matrices X_0 sharing with G one, two, or more eigenvalues and corresponding left eigenvectors. We have stopped iterations when the residual error $||X_n - \sum_{i=0}^{+\infty} X_n^i A_i||_1$ was less than $\epsilon = 10^{-11}$.

We have implemented our method in Fortran 77; the program has been run on a Ultra Sparc Workstation.

We have considered two problems; Problem 1 arises from the mathematical modeling of a metropolitan queueing network (Metaring MAC Protocol) [2],

h	Iterations	Time (s.)
1	2763	197.3
2	34	4.9
4	26	4.4
5	25	4.4
6	20	4.0
8	20	4.0
9	20	4.0
10	20	4.1
12	20	4.1
14	20	4.1
15	19	4.0
16	19	4.0

Table 1: Metaring MAC Protocol

where the blocks A_i have dimension 16; Problems 2 models a teletraffic system as a continuous time QBD Markov chain, where the blocks A_i have dimension 24.

For Problem 1 we have approximated the eigenvalues of the matrix G by means of Aberth's method, applied to the function (13), and the left eigenvectors by solving the linear systems (14). The time needed to approximate all the eigenvalues, and the corresponding left eigenvectors, was 2.5 seconds, and 0.02 seconds, respectively. We have sorted the eigenvalues in such a way that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_k|$, and we have chosen different subsets $\Lambda_h = \{\lambda_1, \ldots, \lambda_h\}$, where h is such that properties 1, 2, and 3 of Section 2 are satisfied. Table 1 reports: the number h of eigenvalues shared by X_0 with G; the number of iterations needed by the functional iteration method to reach the required accuracy; the total time (in seconds) needed to compute the matrix X_0 and to perform the functional iterations. Figure 1 reports the logarithm (to the base 10) of the residual error for $h = 1, 2, 16$.

For this problem, the time needed to reach the required accuracy with the faster functional iteration $X_{n+1} = A_0(I - \sum_{i=1}^{+\infty} X_n^{i-1} A_i)^{-1}$ [14] and $X_0 = I$ is 207.3 seconds, with 2943 iterations. It is worth pointing out the substantial reduction of the number of iterations (and of the total time) when h passes from 1 to 2, and the strong advantage with respect to the faster functional iteration method starting with the identity matrix.

For Problem 2 we have computed the eigenvalues of the matrix G, by approximating the eigenvalues of an auxiliary $2k \times 2k$ matrix, as suggested in [6], and the eigenvectors, by solving the linear systems (14). The time needed was 0.01 seconds. The eigenvalues, which are real and distinct for this problem, are

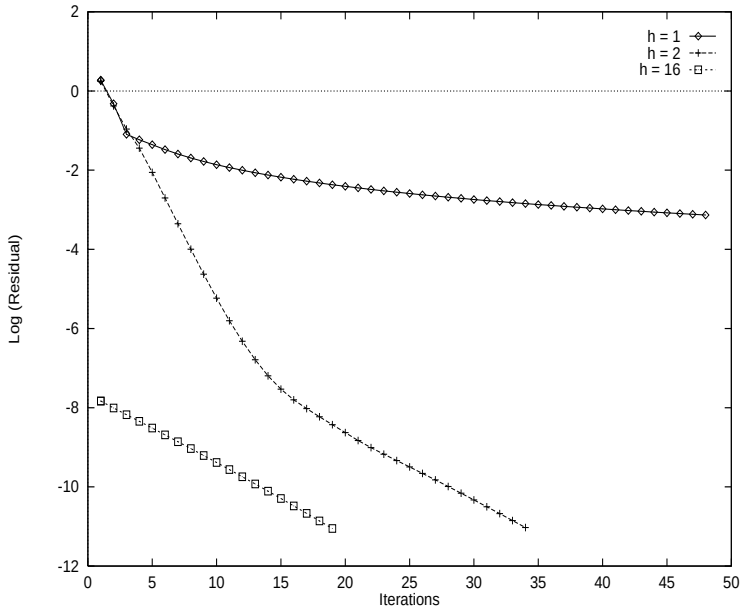

Figure 1: Metaring MAC Protocol

sorted in such a way that $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_k|$. We have chosen X_0 with eigenvalues $\lambda_1, \ldots, \lambda_h$, and left eigenvectors $\mathbf{w}_1^t, \ldots, \mathbf{w}_h^t$, for different values of h in $\{1, \ldots, k\}$. Table 2 reports: the number h of eigenvalues shared by X_0 with G; the number of iterations needed by the functional iteration method to reach the required accuracy; the total time (in seconds) needed to compute the matrix X_0 and to perform the functional iterations.

From Table 2 and Figure 2, we can observe a substantial reduction of the number of iterations for increasing values of h. When $h = 24$, i.e., when all the eigenvalues and the corresponding left eigenvectors are known, the matrix X_0 is a poor approximation of G, that is refined in few steps of functional iteration.

References

[1] O. Aberth. Iteration methods for finding all zeros of a polynomial simultaneously. *Math. Comp.*, 27, 1973.

[2] G. Anastasi, L. Lenzini, and B. Meini. Performance evaluation of a worst case model of the MetaRing MAC protocol with global fairness. *Performance Evaluation*, 29:127–151, 1997.

h	Iterations	Time (s.)
1	83361	167.1
2	67271	134.1
5	39732	79.8
10	16718	33.5
12	11198	24.0
15	4736	9.9
18	340	0.8
24	340	0.8

Table 2: PH/PH/D Queue

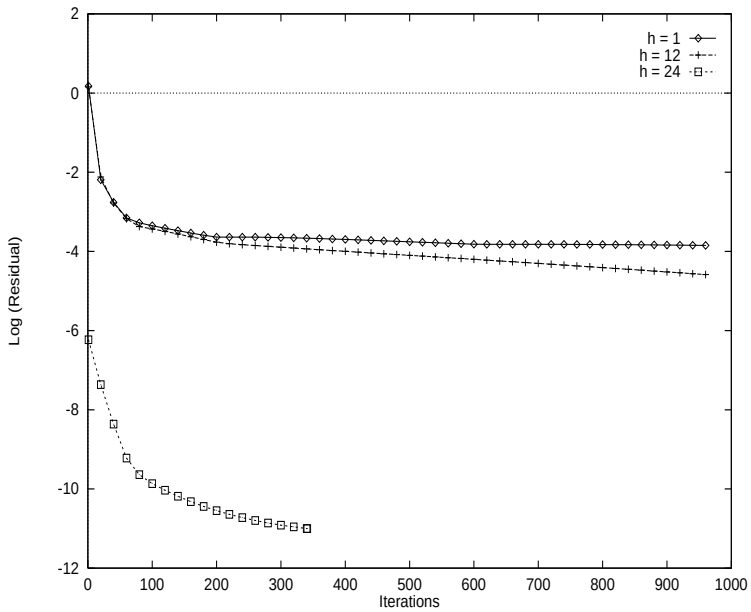

Figure 2: PH/PH/D Queue

[3] D. Bini and B. Meini. On cyclic reduction applied to a class of Toeplitz-like matrices arising in queueing problems. In W. J. Stewart, editor, *Computations with Markov Chains*, pages 21–38. Kluwer Academic Publisher, 1995.

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

[5] D. Bini and B. Meini. Improved cyclic reduction for solving queueing problems. *Numerical Algorithms*, 15:57–74, 1997.

[6] J. N. Daigle and D. M. Lucantoni. Queueing systems having phase-dependent arrival and service rates. In W. J. Stewart, editor, *Numerical solution of Markov Chains*, pages 161–202. Marcel Dekker, New York, 1991.

[7] L. W. Ehrilch. A modified Newton method for polynomials. *Comm. A.C.M.*, 10:107–108, 1967.

[8] P. Favati and B. Meini. Relaxed functional iteration techniques for the numerical solution of M/G/1 type Markov chains. *BIT*, 38, 1998. In printing.

[9] H. R. Gail, S. L. Hantler, and B. A. Taylor. Solutions of the basic matrix equation for M/G/1 and G/M/1 type Markov chains. *Commun. Statist. Stochastic Models*, 10:1–43, 1994.

[10] G.H. Golub and C.F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, Baltimore, 1989.

[11] G. Latouche. Newton's iteration for non-linear equations in Markov chains. *IMA J. of Numerical Analysis*, 14:583–598, 1994.

[12] G. Latouche and V. Ramaswami. A logarithmic reduction algorithm for Quasi-Birth-Death processes. *J. Appl. Probability*, 30:650–674, 1993.

[13] G. Latouche and G.W. Stewart. Numerical methods for M/G/1 type queues. In W. J. Stewart, editor, *Computations with Markov Chains*, pages 571–581. Kluwer Academic Publishers, 1996.

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

[15] G. R. Murthy, M. Kim, and E. J. Coyle. Equilibrium analysis of skip free Markov chains: nonlinear matrix equations. *Commun. Statist. Stochastic Models*, 7:547–571, 1991.

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

[17] V. Ramaswami. A stable recursion for the steady state vector in Markov chains of M/G/1 type. *Commun. Statist. Stochastic Models*, 4:183–188, 1988.

[18] R. S. Varga. *Matrix Iterative Analysis*. Prentice Hall, New Jersey, 1962.