

SOLVING MATRIX POLYNOMIALS BY NEWTON'S METHOD WITH EXACT LINE SEARCHES

JONG HYEON SEO¹ AND HYUN-MIN KIM^{2†}

¹DEPARTMENT OF MATHEMATICS, PUSAN NATIONAL UNIVERSITY, SOUTH KOREA

E-mail address: hyeonni94@hanmail.net

²DEPARTMENT OF MATHEMATICS, PUSAN NATIONAL UNIVERSITY, SOUTH KOREA

E-mail address: hyunmin@pnu.edu

ABSTRACT. One of well known and much studied nonlinear matrix equations is the matrix polynomial which has the form $P(X) = A_0X^m + A_1X^{m-1} + \cdots + A_m$, where A_0, A_1, \dots, A_m and X are $n \times n$ complex matrices. Newton's method was introduced a useful tool for solving the equation $P(X) = 0$. Here, we suggest an improved approach to solve each Newton step and consider how to incorporate line searches into Newton's method for solving the matrix polynomial. Finally, we give some numerical experiment to show that line searches reduce the number of iterations for convergence.

1. INTRODUCTION

Nonlinear matrix equations often occur in applications and modeling of scientific problems. In this work we specially consider one of the nonlinear matrix equations which is called the matrix polynomial. A matrix polynomial can be defined by

$$P(X) = A_0X^m + A_1X^{m-1} + \cdots + A_m = 0, \quad (1)$$

where $A_0, A_1, \dots, A_m, X \in \mathbb{C}^{n \times n}$. If $A_0 = I$ where I is an $n \times n$ identity matrix, then we say that $P(X)$ is monic. Matrix polynomials appear in the theory of differential equations, system theory, network theory, stochastic theory and other areas [8], [9], [10], [11]. A matrix S satisfying the equation $P(S) = 0$ is called a solvent, or more precisely, a right solvent of $P(X)$ to distinguish it from a left solvent, which is a solution of the related matrix equation

$$X^m A_0 + X^{m-1} A_1 + \cdots + A_m = 0.$$

In the quadratic case ($m = 2$), Davis [1], [2] considered Newton's Method and Higham and Kim [5], [6] incorporated the exact line searches into Newton's method and developed functional iterations. For solving matrix polynomials Newton's method was considered by

2000 *Mathematics Subject Classification.* 65F30, 65H10.

Key words and phrases. matrix polynomial, solvent, Newton's method, line searches.

[†] Corresponding author: This author was supported for two years by Pusan National University Research Grant.

Kratz and Stickel [7] and Berlloulli's iteration was suggested by Dennis, Jr., Traub and Weber [3], [4].

Much motivation for studying the matrix polynomial comes from the polynomial eigenvalue problems

$$P(\lambda)\mathbf{v} = (A_0\lambda^m + A_1\lambda^{m-1} + \cdots + A_m)\mathbf{v} = 0. \quad (2)$$

If $P(\lambda_0)$ is singular, λ_0 is called a polynomial eigenvalue and a vector \mathbf{v}_0 corresponding λ_0 is called a polynomial eigenvector, or more precisely, a right polynomial eigenvector of $P(\lambda)$ to distinguish it from a left polynomial eigenvector, which is a vector corresponding λ_0 of the related equation

$$\mathbf{v}^T P(\lambda) = 0^T.$$

In this work, we extend the existence theory of solvent of the quadratic matrix equation which suggested by Xu and Lu [12] to the matrix polynomial (1) and improve Newton step using Schur decomposition. Also we show how to incorporate exact line searches into Newton's method for solving matrix polynomials.

2. THEORY

We deal primarily with the monic case for theoretical progress. First, we introduce the conditions of existence of solvents by spectral data of a given the equation $P(\lambda)$. The next result show how solvents can be constructed from eigenpairs of the polynomial eigenvalue problems.

Theorem 2.1. [3, Lem. 4.1] If a polynomial eigenvalue problem (2) has n linearly independent eigenvectors, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n$ corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_n$, then $Q\Lambda Q^{-1}$ is a solvent to the matrix polynomial (1), where $Q = [\mathbf{v}_1, \cdots, \mathbf{v}_n]$ and $\Lambda = \text{diag}(\lambda_1, \cdots, \lambda_n)$.

The following theorem gives information on the number of solvents of $P(X)$.

Theorem 2.2. [6, Thm. 2.1] Suppose $P(\lambda)$ in (2) has p distinct eigenvalues $\{\lambda_i\}_{i=1}^p$, with $n \leq p \leq mn$, and the corresponding set of p eigenvectors $\{\mathbf{v}_i\}_{i=1}^p$ satisfies the Haar condition. Then there are at least $\binom{p}{n}$ different solvents of $P(X)$, and exactly this many if $p = mn$, which are given by

$$S = Q \text{diag}(\mu_i) Q^{-1}, \quad Q = [\mathbf{q}_1, \cdots, \mathbf{q}_n],$$

where the eigenpairs $(\mu_i, \mathbf{q}_i)_{i=1}^n$ are chosen from among the eigenpairs $(\lambda_i, \mathbf{v}_i)_{i=1}^p$ of $P(\lambda)$.

However, the next example shows that even if we can find the spectral data for associated polynomial eigenvalue problem it is not always possible to construct the solvent.

Example 2.3. Let

$$P_M(X) \equiv X^2 + X + \begin{bmatrix} -6 & -5 \\ 0 & -6 \end{bmatrix} = 0.$$

Then, $P_M(X) = 0$ has two solvents

$$\begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} -3 & -1 \\ 0 & -3 \end{bmatrix}.$$

But there is only one eigenvector associative with $P_M(\lambda)\mathbf{v} = 0$, namely $(1, 0)^T$ for $\lambda = 2$ or $\lambda = -3$. Thus it is impossible to construct a solvent having the form $U \text{daig}\{2, -3\}U^{-1}$, since the corresponding eigenvectors are not linearly independent.

Xu and Lu [12] provided an existence theorem of solvents to the equation $AX^2 + BX + C = 0$, where $A = I$, B and C are simultaneously triangularizable. We now extend these results to obtain some information of solvents from the restricted matrix polynomials, and reify the Theorem 2.1.

Theorem 2.4. *Let $P(X) = A_0X^m + A_1X^{m-1} + \dots + A_m$ be a matrix polynomial with $A_p = UR_pU^*$ for $p = 0, 1, \dots, m$. (U is unitary matrix and R_p 's are upper triangular matrices). Then, if for $i, j = 1, \dots, n$, $s_{i,j}$ is a solution of scalar coefficient polynomial $f_{i,j}(t) = \sum_{p=0}^m [R_p]_{ij}t^{m-p}$ and $i > j$ implies that $f_{j,j}(s_{i,i}) \neq 0$, then there exists a solvent S of $P(X) = 0$ having a form of $S = URU^*$ for some upper triangular matrix R .*

Proof. Define

$$N(X) \equiv \sum_{p=0}^m R_p X^{m-p}.$$

Then, we can easily verify that R is a solvent of $N(X) = 0$ implies that URU^* is a solvent of $P(X) = 0$. Therefore it is sufficient to prove that there exists a solvent of $N(X) = 0$.

For a associated polynomial eigenvalue problem $N(\lambda)$,

$$\det(N(\lambda)) = \prod_{j=1}^n f_{j,j}(\lambda).$$

Thus the set of all solutions of $f_{j,j}(\lambda) = 0$ for $j = 1, \dots, n$ is also the that of eigenvalues of $N(\lambda)$. For each chosen eigenvalue $\lambda_i = s_{i,i}$ ($i = 1, 2, \dots, n$), a right eigenvector \mathbf{x}_i can be obtained by solving the system

$$N(\lambda_i)\mathbf{x}_i = \begin{bmatrix} f_{1,1}(\lambda_i) & f_{1,2}(\lambda_i) & \cdots & \times & \times \\ 0 & f_{2,2}(\lambda_i) & \ddots & \times & \times \\ & 0 & \ddots & & \vdots \\ & & \ddots & f_{n-2,n-1}(\lambda_i) & \times \\ & & & f_{n-1,n-1}(\lambda_i) & \times \\ & & & 0 & f_{n,n}(\lambda_i) \end{bmatrix} \begin{bmatrix} x_1^{(i)} \\ x_2^{(i)} \\ \vdots \\ x_i^{(i)} \\ \vdots \\ x_n^{(i)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Notice that $f_{1,1}(\lambda_i), \dots, f_{i-1,i-1}(\lambda_i)$ are not zero by the hypothesis.

Since the matrix $N(\lambda_i)$ is upper triangular matrix, we may find a solution of the system by backward substitution. First put $x_k^{(i)} = 0$ for $k = i + 1, \dots, n$. (It is trivial solution of the k -th

linear equation. Next if $k = i$, then i -th linear equation has the form $0 \cdot x_i^{(i)} = 0$. Since we can choose $x_i^{(i)}$ arbitrary, take $x_i^{(i)} = 1$ for convenience. At last if $k < i$, then we can easily get $x_k^{(i)}$ by solving the k -th linear equation, since $f_k(\lambda_i) \neq 0$. Therefore for each $i = 1, 2, \dots, n$, the eigenvector \mathbf{x}_i has the form

$$\mathbf{x}_i = [x_1^{(i)} \quad x_2^{(i)} \quad \cdots \quad x_{i-1}^{(i)} \quad 1 \quad 0 \quad \cdots \quad 0],$$

where $x_k^{(i)}$ ($k = 1, 2, \dots, i-1$) is the solution of the linear equations such that

$$f_{k,k}(\lambda_i) \cdot x_k^{(i)} = -(f_{k,k+1}(\lambda_i) \cdot x_{k+1}^{(i)} + \cdots + f_{k,n}(\lambda_i) \cdot x_n^{(i)}).$$

The set of eigenvectors \mathbf{x}_i ($i = 1, \dots, n$) chosen by above method is linearly independent and that makes us possible to construct a nonsingular vector matrix such that

$$Q = \left[\begin{array}{c|c|c|c|c} \mathbf{x}_1 & \mathbf{x}_2 & \mathbf{x}_3 & \cdots & \mathbf{x}_n \end{array} \right] = \left[\begin{array}{c|c|c|c|c} 1 & \times & \times & \cdots & \times \\ & 1 & \times & \cdots & \times \\ & & 1 & \ddots & \times \\ & & & \ddots & \vdots \\ & & & & 1 \end{array} \right].$$

By the Theorem 2.1, the upper triangular matrix R which is a solvent of $N(X) = 0$ can be obtained by

$$R = Q \operatorname{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\} Q^{-1}.$$

□

The next corollary is directly induced.

Corollary 2.5. *Let a matrix polynomial of which the coefficient matrices are all upper triangular (or lower triangular) be given and the additional hypothesis of Theorem 2.4 hold. Then there exists a solvent.*

Now our attention may be moved to the numerical methods for solving matrix polynomials.

3. NEWTON'S METHOD

Newton's method is the most well-known and powerful numerical method for solving nonlinear equations. For this reason it is a natural approach that we apply Newton's method to solve the matrix polynomial (1). Newton's method for solving matrix polynomials was introduced by Kratz and Stickel [7]. In this section we will suggest an improved Newton step computationally using the Schur decomposition and show how to incorporate exact line searches when solving matrix polynomials by Newton's method.

We now construct the iterative method which has second order converges. Consider the nonlinear matrix equations

$$G(X) = 0, \tag{3}$$

where $G : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n}$. Define $H_k \in \mathbb{C}^{n \times n}$ as the solution of the linear equation $G(X_k) + G'(X_k)H_k = 0$, where the linear operator $G'(X)H : \mathbb{C}^{n \times n} \rightarrow \mathbb{C}^{n \times n}$ is the Fréchet derivative of G at X in the direction H . Then Newton's method for the nonlinear matrix equation (3) can be defined by

$$\left. \begin{aligned} X_0 & \text{ given,} \\ G(X_k) + G'(X_k)H_k &= 0 \\ X_{k+1} &= X_k + H_k \end{aligned} \right\}, \quad k = 0, 1, 2, \dots$$

Thus, each step of Newton's method involves finding the solution $H \in \mathbb{C}^{n \times n}$ of linear equation

$$G'(X)H = -G(X). \quad (4)$$

For the matrix polynomial $P(X) = 0$ the equation (4) can be explicitly represented by

$$P'(X)H = D_X(H) = \sum_{i=1}^m \left\{ \left(\sum_{j=0}^{m-i} A_j X^{m-(j+i)} \right) H X^{i-1} \right\} = -P(X). \quad (5)$$

We now have a natural question: when is the Fréchet derivative D_X regular, both at a solvent and an iterate, so that (5) has a unique solution? Unfortunately we do not have remarkable condition. Kratz [7] answered the questions in specific case and presented the Newton method to find solvents of the matrix polynomial (1) and proved that the algorithm converges quadratically near a simple solvent.

Finding the solvents of the matrix polynomials by Newton's method can be regarded as solving the linear equations (5). The general approach solving (5) is using vec function and Kronecker product.

Using vec operator $\text{vec}(D_X(H))$ is represented by

$$\begin{aligned} \text{vec}(D_X(H)) &= \text{vec}(B_1 H + \dots + B_m H X^{m-1}) \\ &= \text{vec}(B_1 H I) + \dots + \text{vec}(B_m H X^{m-1}) \\ &= I \otimes B_1 \text{vec}(H) + \dots + (X^T)^{m-1} \otimes B_m \text{vec}(H) \\ &= (I \otimes B_1 + \dots + (X^T)^{m-1} \otimes B_m) \text{vec}(H) \\ &= \left(\sum_{i=1}^m (X^T)^{i-1} \otimes B_i \right) \text{vec}(H) \\ &= \left\{ \sum_{i=1}^m \left((X^T)^{i-1} \otimes \sum_{j=0}^{m-i} A_j X^{m-(j+i)} \right) \right\} \text{vec}(H) \\ &= D \text{vec}(H) \end{aligned} \quad (6)$$

where

$$B_p = A_0 X^{m-p} + A_1 X^{m-(p+1)} + \dots + A_p, \quad p = 1, 2, \dots, m, \quad (7)$$

and

$$\mathcal{D} = \sum_{i=1}^m \left((X^T)^{i-1} \otimes \sum_{j=0}^{m-i} A_j X^{m-(j+i)} \right).$$

By the equality (6), the linear equation $D_X(H) = -P(X)$ is changed by the $n^2 \times n^2$ linear system which is

$$\mathcal{D} \operatorname{vec}(H) = \operatorname{vec}(-P(X)). \quad (8)$$

Note that D_X is regular if and only if \mathcal{D} is a nonsingular matrix, and then $\inf_{\|H\|=1} \|D_X(H)\| = \min_{\|H\|=1} \|\mathcal{D} \operatorname{vec}(H)\| > 0$ [7, Lem. 1]. If D_X is regular, then the linear equation $D_X(H) = -P(X)$ has a unique solution H , where $\operatorname{vec}(H)$ solves the $n^2 \times n^2$ linear system $\mathcal{D} \operatorname{vec}(H) = \operatorname{vec}(-P(X))$. It seems to be reasonable, however it is nonsense from the viewpoint of numerical analysis. In the Newton step, we need to reduce the system size of the equation (8) to $n \times n$. Here is the useful algorithm using Schur decomposition. Given $X \in \mathbb{C}^{n \times n}$, compute the Schur decomposition of X

$$Q^* X Q = R,$$

where Q is unitary and R is upper triangular. Then, substituting this into (5) transforms the system to

$$D_R(H') = B_1 H' + B_2 H' R + \cdots + B_m H' R^{m-1} = C \quad (9)$$

where $B_i = \sum_{j=0}^{m-i} A_j X^{m-j+i}$, $H' = H Q$ and $C = -P(X) Q$. So, taking the vec function both sides of (9) makes a linear system such that

$$\operatorname{vec}(D_R(H')) = \tilde{\mathbf{D}} \operatorname{vec}(H) \quad (10)$$

where the matrix $\tilde{\mathbf{D}} \in \mathbb{C}^{n^2 \times n^2}$ is given by

$$\tilde{\mathbf{D}} = \sum_{i=1}^m ((R^T)^{i-1} \otimes B_i). \quad (11)$$

If we define $\tilde{D}_{ij} = \sum_{k=1}^m [R^{k-1}]_{ji} B_k$, then $\tilde{\mathbf{D}}$ in (11) is represented by

$$\tilde{\mathbf{D}} = \begin{pmatrix} \tilde{D}_{11} & & & \\ \tilde{D}_{21} & \tilde{D}_{22} & & \\ \vdots & & \ddots & \\ \tilde{D}_{n1} & \cdots & \cdots & \tilde{D}_{nn} \end{pmatrix}.$$

Since \tilde{D} is a block lower triangular matrix, using the block forward substitution, the equation (10) is changed to n linear systems with size $n \times n$ such that

$$\begin{aligned} \mathbf{h}'_1 &= \tilde{D}_{11}^{-1} \mathbf{c}_1 \\ \mathbf{h}'_2 &= \tilde{D}_{22}^{-1} (\mathbf{c}_2 - \tilde{D}_{21} \mathbf{h}'_1) \\ &\vdots \\ \mathbf{h}'_n &= \tilde{D}_{nn}^{-1} (\mathbf{c}_n - \tilde{D}_{n1} \mathbf{h}'_1 - \cdots - \tilde{D}_{n,n-1} \mathbf{h}'_{n-1}), \end{aligned}$$

where \mathbf{h}'_i and \mathbf{c}_i are i -th columns of H' and C , respectively.

Algorithm 3.1. Given matrices R, B_1, B_2, \dots, B_m which is defined by (7) and (9) and positive integers i, j ($i \leq j$ and $i, j \leq n$), the following algorithm computes \tilde{D}_{ij}

```

for  $r = 1 : m$ 
   $\tilde{D}_{ij} \leftarrow \tilde{D}_{ij} + R^{r-1}(j, i)B_r$ 
end

```

Algorithm 3.2 (Solving the equation $D_X(H) = -P(X)$ by Schur decomposition). $A_0, \dots, A_m, X \in \mathbb{C}^{n \times n}$ is given and $X = QRQ^*$ is the Schur decomposition of X . This algorithm finds $H \in \mathbb{C}^{n \times n}$ which is the solution of $D_X(H) = -P(X)$ in (5)

```

 $B_m \leftarrow A_0, R_1 \leftarrow I$ 
for  $i = 1 : m - 1$ 
   $B_{m-i} \leftarrow B_{m-i+1}X + A_i$ 
   $R_{i+1} \leftarrow R_i R$ 
end
 $C \leftarrow -(B_1 X + A_m)Q$ 
for  $i = 1 : n$ 
   $\mathbf{d} \leftarrow 0$ 
  for  $j = 1 : i - 1$ 
    Use algorithm (3.1) to compute  $\tilde{D}_{ij}$ 
     $\mathbf{d} \leftarrow \mathbf{d} + \tilde{D}_{ij}H(:, j)$ 
  end
  Use algorithm (3.1) to compute  $\tilde{D}_{ii}$ 
   $H(:, i) \leftarrow \tilde{D}_{ii}^{-1}(C(:, i) - \mathbf{d})$ 
end
 $H \leftarrow HQ^*$ 

```

Let the matrix polynomial (1) have only real coefficient matrices. Although the desired solvent is real above method may required complex arithmetic. We will use the technique of Schur algorithm to desire an algorithm for computation a solvent that use only real arithmetic if the given coefficient matrices are real.

Let $X \in \mathbb{R}^{n \times n}$ be given and $X = Q\widehat{R}Q^T$ be a real Schur decomposition with

$$\widehat{R} = \begin{bmatrix} R_{11} & \cdots & \cdots & R_{1p} \\ & R_{22} & & \vdots \\ & & \ddots & \vdots \\ & & & R_{pp} \end{bmatrix} \quad (12)$$

where R_{ii} is 1×1 or 2×2 matrices. Substituting $X = Q\widehat{R}Q^T$ into (5) transforms the system to

$$D_R(H') = B_1H' + B_2H'\widehat{R} + \cdots + B_mH'\widehat{R}^{m-1} = \widehat{C} \quad (13)$$

where $B_i = \sum_{j=0}^{m-i} A_j X^{m-j+i}$, $H' = HQ$, and $\widehat{C} = -P(X)Q$. By taking the vec function both sides of the equation (13), we have the linear system given by

$$\widehat{D}\text{vec}(\widehat{H}) = \text{vec}(\widehat{C}) \quad (14)$$

with

$$\widehat{D} = \begin{pmatrix} \widehat{D}_{11} & & & \\ \widehat{D}_{21} & \widehat{D}_{22} & & \\ \vdots & & \ddots & \\ \widehat{D}_{p1} & \cdots & \cdots & \widehat{D}_{pp} \end{pmatrix}$$

where

$$\widehat{D}_{ij} = \begin{cases} (I)^{\Psi(i)} \otimes B_1 + \sum_{k=2}^m (R_{ji}^{k-1})^T \otimes B_k & \text{if } i == j \\ \sum_{k=2}^m (R_{ji}^{k-1})^T \otimes B_k & \text{else.} \end{cases}$$

and $\Psi(i)$ is the number of columns of block matrix R_{ii} in (12).

By quasi-block-forward substitution, we have p pieces of linear system such that

$$\widehat{D}_{ii}\widehat{\mathbf{h}}_i = \widehat{\mathbf{c}}_i - \sum_{k=1}^{i-1} \widehat{D}_{ik}\widehat{\mathbf{h}}_k, \quad (15)$$

where $\begin{bmatrix} \widehat{\mathbf{h}}_1 \\ \vdots \\ \widehat{\mathbf{h}}_p \end{bmatrix} = \text{vec}(\widehat{H})$, $\begin{bmatrix} \widehat{\mathbf{c}}_1 \\ \vdots \\ \widehat{\mathbf{c}}_p \end{bmatrix} = \text{vec}(\widehat{C})$, and $\widehat{\mathbf{c}}_i, \widehat{\mathbf{h}}_i \in \mathbb{C}^{\Psi(i)n}$ for $i = 1, 2, \dots, p$. Note

that \widehat{D}_{ii} 's are either $2n \times 2n$ or $n \times n$ matrix, the index of element $[R_{ij}]_{11}$ ($i \leq j$) of \widehat{R} is $(\sum_{k=1}^{i-1} \Psi(k) + 1, \sum_{k=1}^{j-1} \Psi(k) + 1)$ and the size of R_{ij} is $\Psi(i) \times \Psi(j)$.

Algorithm 3.3. Given matrices $\widehat{R}, B_1, B_2, \dots, B_m$ which is defined by (13) and (7) and positive integers i, j ($i \leq j$ and $i, j \leq p$), the following algorithm computes \widehat{D}_{ij}

if $i == j$

$\widehat{D}_{ij} \leftarrow I_2 \otimes B_1$ (I_2 is 2×2 identity matrix.)

else

$\widehat{D}_{ij} \leftarrow O_{n \times \Psi(i), n \times \Psi(j)}$ (O_{nm} is $n \times m$ zero matrix.)

for $r = 2 : m$
 $\widehat{D}_{ij} \leftarrow \widehat{D}_{ij} + (R_{ji}^{r-1})^T \otimes B_r$
end

Algorithm 3.4 (Solving the equation $D_X(H) = -P(X)$ by real Schur decomposition). $A_0, \dots, A_m, X \in \mathbb{R}^{n \times n}$ is given and $X = Q\widehat{R}Q^T$ is the real Schur decomposition of X in (13). This algorithm find $H \in \mathbb{R}^{n \times n}$ which is the solution of $D_X(H) = -P(X)$ in (5)

$B_m = A_0, \widehat{R}_0 = I$
for $i = 1 : m - 1$
 $B_{m-i} \leftarrow B_{m-i+1}X + A_i$
 $\widehat{R}_i \leftarrow \widehat{R}_{i-1}\widehat{R}$
end
 $\widehat{C} \leftarrow -(B_1X + A_m)Q$
 $pos \leftarrow 0, i \leftarrow 1$
while $pos < n$
if $\widehat{R}(pos + 2, pos + 1) == 0$ **then** $\Psi(i) \leftarrow 1$
else $\Psi(i) \leftarrow 2$
end
 $pos \leftarrow \Psi(i) + pos$
 $i \leftarrow i + 1$
end
 $pos \leftarrow 1$
for $i = 1 : p$ (p is the maximum of domain of Ψ .)
 $\mathbf{d} \leftarrow O_{n \times \Psi(i), 1}$
for $j = 1 : i - 1$
Use algorithm (3.3) to compute \widehat{D}_{ij}
 $\mathbf{d} \leftarrow \mathbf{d} + \widehat{D}_{ij}$
end
Use algorithm (3.3) to compute \widehat{D}_{ii}
 $\widehat{\mathbf{h}}_i \leftarrow \widehat{D}_{ii}^{-1} [\text{vec}\{C(:, pos : pos + \Psi(i) - 1)\} - \mathbf{d}]$
if $\Psi(i) == 1$
 $H(:, pos) \leftarrow \widehat{\mathbf{h}}_i$
else
 $H(:, pos) \leftarrow \widehat{\mathbf{h}}_i(1 : n, 1)$
 $H(:, pos + 1) \leftarrow \widehat{\mathbf{h}}_i(n + 1, 2n, 1)$
end
 $pos \leftarrow pos + \Psi(i)$
end
 $H = HQ^T$

We now show how to incorporate exact line searches when solving a matrix polynomial (1) by Newton's method. For implementation we need to find the exact expansion of $P(X + tH)$.

Let μ, ν be the permutations of $1, 2, \dots, n$, and integer $m(1 \leq m \leq n)$ be given. Then the relation $\mu \sim_m \nu$ given by $\mu(i) > m$ if and only if $\nu(i) > m$ and $\mu(i) \leq m$ if and only if $\nu(i) \leq m$ for all $i = 1, 2, \dots, n$ is an equivalent relation. So we can define the equivalent class of permutations μ 's by $[\mu]_m$. Note that the number of $[\mu]_m$ is $n!/m!(n-m)!$. Let $r_1 = r_2 = \dots = r_m = X \in \mathbb{C}^{n \times n}$, $r_{m+1} = r_{m+2} = \dots = r_n = Y \in \mathbb{C}^{n \times n}$ then the function $\Phi_{X,Y} : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{C}^{n \times n}$ can be defined by

- (1) $\Phi_{X,Y}[0, 0] \equiv I$,
- (2) $\Phi_{X,Y}[m, n-m] \equiv \sum_{[\mu]_m} r_{[\mu]_m(1)} r_{[\mu]_m(2)} \cdots r_{[\mu]_m(n)}$.

The function $\Phi_{X,Y}$ is the sum of the products of all repeated permutations of X and Y . By using the notation of Φ , we can describe the expansions of $(X + tH)^k$ for $k = 0, 1, 2, \dots, m$

$$\begin{aligned} A_m &= A_m \Phi_{XH}[0, 0], \\ A_{m-1}(X + tH) &= A_{m-1}X + tA_{m-1}H, \\ &= A_{m-1}\Phi_{X,H}[1, 0] + tA_{m-1}\Phi_{X,H}[0, 1], \\ A_{m-2}(X + tH)^2 &= A_{m-2}X^2 + A_{m-2}(XH + HX) + t^2A_{m-2}H^2 \\ &= A_{m-2}\Phi_{X,H}[2, 0] + tA_{m-2}\Phi_{X,H}[1, 1] + t^2A_{m-2}\Phi_{X,H}[0, 2], \\ &\vdots \\ A_{m-k}(X + tH)^m &= A_{m-k}\Phi_{X,H}[m, 0] + tA_{m-k}\Phi_{X,H}[m-1, 1] + \cdots + t^m\Phi_{X,H}[0, m]. \end{aligned}$$

Therefore we obtain the expansion formula

$$(X + tH)^k = \sum_{i=0}^k t^i \Phi_{X,H}[k-i, i] \quad (16)$$

and easily verify

$$(\Phi_{X,H}[i, j])^* = \Phi_{X^*, Y^*}[i, j].$$

By using the formula (16) and equality $D_X(H) + P(X) = 0$, $P(X + tH)$ can be represented by

$$\begin{aligned} P(X + tH) &= \sum_{i=0}^m \sum_{j=0}^i t^j \Phi_{X,H}[i-j, j] \\ &= P(X) + tD_X(H) + \sum_{i=2}^m \sum_{j=2}^i t^j \Phi_{X,H}[i-j, j] \\ &= (1-t)P(X) + \sum_{i=2}^m \sum_{j=2}^i t^j \Phi_{X,H}[i-j, j]. \end{aligned}$$

Finally the merit function $p(t)$ can be obtained by

$$\begin{aligned}
p(t) &= \|P(X + tH)\|_2^F \\
&= \text{trace} \left((1-t)P(X)^* + \sum_{i=2}^m \sum_{j=2}^i \Phi_{X^*, Y^*}[i-j, j] \right) \\
&\quad \left((1-t)P(X) + \sum_{i=2}^m \sum_{j=2}^i \Phi_{X, Y}[i-j, j] \right) \\
&= (1-t)^2 \|P(X)\|_F^2 + \cdots + t^{2m} \Phi_{X^*, Y^*}[0, m] \Phi_{X, Y}[0, m] \\
&= (1-t)^2 \|P(X)\|_F^2 + \cdots + t^{2m} \|H^m\|_F^2.
\end{aligned}$$

Since $p'(0) = -2\|P(X)\|_F^2 < 0$ and leading coefficient of $p'(t)$ is positive, p' has a real zero in positive real number, and this zero corresponds to a minimum or a point of inflection of p . If $p(t)$ has the minimum for large t , then the Newton step with line search may have numerically a bad effect on convergence. So we must restrict our attention to the appropriate interval $(0, k]$. Because $t = 1$ corresponds to a pure Newton step, k must be greater than 1. Thus we define t by

$$p(t) = \min_{x \in (0, k]} p(x) \text{ with } k > 1.$$

Roughly speaking, exact line searches means finding scalar t that makes $X + tH$ closer point in direction of H from solvent.

4. NUMERICAL EXPERIMENTS AND CONCLUSIONS

In this section we show and compare some experimental results using Newton's method with and without line searches. Our experiments were done in MATLAB. Iterations for Newton's method with and without exact line searches are terminated when the residual $P(X_k)$ is of the same order of magnitude as the round error in computing it, namely when the relative residual $\rho(X_k)$ satisfies

$$\rho(X_k) = \frac{\|fl(P(X_k))\|_F}{\|A_0\|_F \|X_k\|_F^m + \cdots + \|A_m\|_F} \leq nu, \quad (17)$$

where $u = 2^{-53} \simeq 1.1 \times 10^{-16}$ is unit round off. MATLAB codes for Newton's method has an option to choose whether to use exact line searches. Suggested examples are in Kratz [7] solved by pure Newton's method. We will compare the results with exact line searches.

Consider a matrix equation

$$P_1(X) = X^3 + \begin{bmatrix} -6 & 6 \\ -3 & -15 \end{bmatrix} X^2 + \begin{bmatrix} 2 & -42 \\ 21 & 65 \end{bmatrix} X + \begin{bmatrix} 18 & -66 \\ 33 & 81 \end{bmatrix} = 0. \quad (18)$$

We obtained two solvents such that

$$S_1 = \begin{bmatrix} 4 & -2 \\ 1 & 7 \end{bmatrix}, \quad S_2 = \begin{bmatrix} 0 & -2 \\ 1 & 3 \end{bmatrix}.$$

with starting matrices $218I_2$ and $-218I_2$, respectively. Figure 1 shows Newton's method with exact line searches converges faster than Newton's method without exact line searches for the two starting matrices.

Here, we consider the matrix differential equation

$$\mathbf{y}^{(4)} + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \mathbf{y}^{(2)} + \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix} \mathbf{y}' + \begin{bmatrix} -20 & 2 & 1 \\ 2 & -20 & 0 \\ 1 & 0 & -20 \end{bmatrix} \mathbf{y} = 0.$$

Such equations occur in connection with vibrating systems. The characteristic polynomial is given by

$$P_2(X) = X^4 + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} X^2 + \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{bmatrix} X + \begin{bmatrix} -20 & 2 & 1 \\ 2 & -20 & 0 \\ 1 & 0 & -20 \end{bmatrix} = 0. \quad (19)$$

Figure 2 shows the convergence of Newton's method with and without exact line searches in $P_2(X) = 0$ with starting matrices $24I$ and $-24I$.

Next, we choose starting matrices

$$X_0 = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 \\ x_7 & x_8 & x_9 \end{pmatrix}, \quad -100 \leq x_1, \dots, x_6 \leq 100$$

with an equally spaced grid of 100 random points (x_1, \dots, x_9) . Figure 3 shows how many times a solvent is produced within 30, 50, and 100 iterations with and without exact line searches. For Newton's method with and without exact line searches, exact line searches give

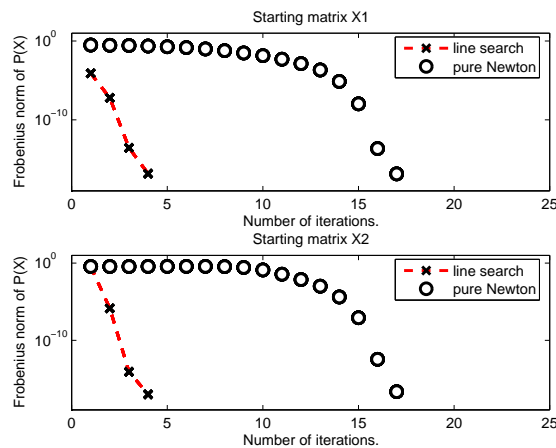


FIGURE 1. Convergence for example (18) with and without Newton's method.

the result in more frequent convergence. Convergence of Newton’s methods with and without exact line searches is obtained to four different solvents depending on the starting matrix.

The matrix polynomial in (1) arises in some applications, for example, the stochastic problems and the polynomial eigenvalue problems. Specially, the polynomial eigenvalue problems can be solved by the $n \times n$ standard eigenvalue problems, if we can find a solvent of associated matrix polynomials.

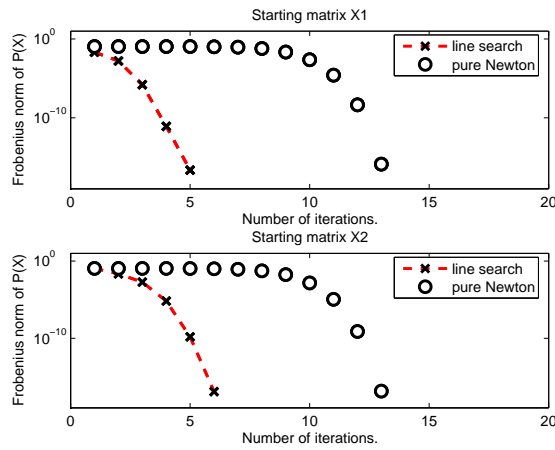


FIGURE 2. Convergence for example (19) with and without Newton’s method.

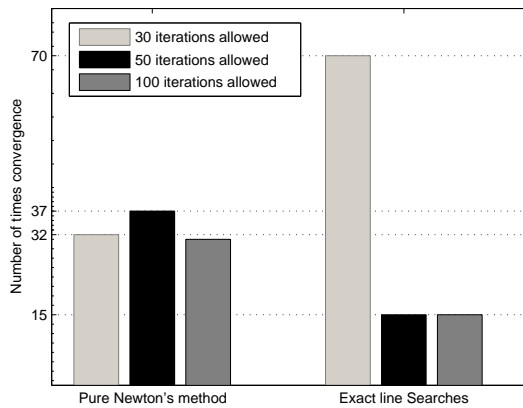


FIGURE 3. Number of times convergence obtained for problem (19) with 100 different arbitrary starting matrices.

We described existence theorems for solvents of the matrix polynomials (1) using spectral theory for the polynomial eigenvalue problems. And we had constructed the solvents for some special matrix polynomial. We also derived Newton's method and improved this algorithm by Schur decomposition and incorporated exact line searches into Newton's method. Finally, we experimented with some examples. Newton's method with exact line searches frequently reduced the number of iterations.

REFERENCES

- [1] George J. Davis, *Numerical solution of a quadratic matrix equation*, SIAM J. Sci. Stat. Comput., **2** (1981), 164–175.
- [2] George J. Davis, *An algorithm to compute solvents of the matrix equation $AX^2 + BX + C = 0$* , ACM Trans. Math. Software, **9** (1983), 246–254.
- [3] J. E. Dennis, Jr., J. F. Traub, and R. P. Weber, *The algebraic theory of matrix polynomials*, Numerical Algorithms, **13** (1976), 831–845.
- [4] J. E. Dennis, Jr. and Robert B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1983.
- [5] N. J. Higham and H.-M. Kim, *Numerical analysis of a quadratic matrix equation*, IMA J. Numer. Anal., **20** (2000), 499–519.
- [6] Nicholas J. Higham and Hyun-Min Kim, *Solving a quadratic matrix equation by Newton's method with exact line searches*, SIAM J. Matrix Anal. Appl., **23** (2001), 303–316.
- [7] W. Kratz and E. Stickel, *Numerical solution of matrix polynomial equations by Newton's method*, IMA J. Numer. Anal., **7** (1987), 355–369.
- [8] P. Lancaster, I. Gohberg and L. Rodman, *Matrix Polynomials*, Academic press, New York, 1982.
- [9] P. Lancaster and M. Tismenetsky, *The Theory of Matrices with Applications, 2nd ed.*, Academic press, New York, 1985.
- [10] Peter Lancaster, *Lambda-matrices and Vibrating Systems*, Pergamon Press, Oxford, 1996.
- [11] Guy Latouche and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*, Society for Industrial and Applied Mathematics, 1999.
- [12] Hong guo Xu and Lin zhang Lu, *Properties of a quadratic matrix equation and the solution of the continuous-time algebraic Riccati equation*, Linear Algebra Appl., **222** (1995), 127–145.

ON THE BAYES ESTIMATOR OF PARAMETER AND RELIABILITY FUNCTION OF THE ZERO-TRUNCATED POISSON DISTRIBUTION

ANWAR HASSAN^{1†}, PEER BILAL AHMAD², AND M. ISHAQ BHATTI³

¹DEPARTMENT OF STATISTICS, UNIVERSITY OF KASHMIR, SRINAGAR, INDIA

E-MAIL address: anwar.hassan2007@gmail.com, anwar_hassan2007@hotmail.com

²DEPARTMENT OF STATISTICS, UNIVERSITY OF KASHMIR, SRINAGAR, INDIA

E-MAIL address: peerbilal@yahoo.co.in

³FINANCIAL STATISTICS GROUP, DEPARTMENT OF ECONOMICS AND FINANCE, SCHOOL OF BUSINESS LA TROBE UNIVERSITY, AUSTRALIA

E-MAIL address: i.bhatti@latrobe.edu.au

ABSTRACT: In this paper Bayes estimator of the parameter and reliability function of the zero-truncated Poisson distribution are obtained. Furthermore, recurrence relations for the estimator of the parameter are also derived. Monte Carlo simulation technique has been made for comparing the Bayes estimator and reliability function with the corresponding maximum likelihood estimator (MLE) of zero-truncated Poisson distribution.

1. INTRODUCTION

The Poisson distribution is defined by

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!}; x = 0, 1, 2, \dots \quad (1)$$

The distribution (1) has been obtained by S.D Poisson [1] as a limiting case of the Binomial distribution, for some reasons Newbold [2], Jensen [3] and David [4] preferred to give credit to De Moivre [5] rather than to S.D. Poisson for discovering of Poisson distribution. The distribution is so important among the discrete distributions that even Fisher, once remarked ‘Among discontinuous distributions’, the Poisson series is of the first importance. Johnson, Kotz and Kemp [6] have discussed the genesis of Poisson distribution in detail. Ahmad and Roohi [7] have discussed the characterization of the Poisson distribution. Roohi and Ahmad [8] studied the inverse ascending factorial moments and estimation of the parameter of hyper-Poisson distribution using negative moments. The Poisson distribution has been described as playing a “similar role with respect to discrete distribution to that of the normal for absolutely continuous distribution”.

The commonest form of truncation is the omission of the zero class, because the observational apparatus becomes active only when at least one event occurs. The distribution (1) can be truncated at $x = 0$ and is defined

2000 *Mathematics Subject Classification.* 62E15.

Key words and phrases. Zero-truncated Poisson distribution, reliability function, Bayes estimator, recurrence relation, Monte Carlo simulation, maximum likelihood estimator.

† Corresponding author.

$$P_1(X = x) = \frac{e^{-\lambda} \lambda^x}{x!(1 - e^{-\lambda})}; \quad x=1,2,\dots \quad (2)$$

This is usually called the positive Poisson distribution. Cohen [9] called it a conditional Poisson distribution. The truncated distribution (2) has been first considered by David and Johnson [10]. In particular, they derived the maximum likelihood estimate (MLE) of λ and its asymptotic variance, and discussed the efficiency of the estimation by moments. Plackett [11] provided a similar estimate of λ for distribution (2), to show that it is highly efficient, and to estimate its sampling variance. Murakami [12] also discussed the maximum likelihood estimators based on censored samples from truncated Poisson distributions. Tate and Goen [13] obtained minimum variance unbiased estimation. Cohen ([14] [15]) provided the estimation of the model (2) from the sample that are truncated on the right. A brief list of authors and their works can be seen in Johnson and Kotz [16], Johnson, Kotz and Kemp [6] and Consul [17].

Bayesian estimation is a likelihood based style of inference that incorporates prior information on the unknown variables. ML estimates are equivalent to the modes of the Bayesian posterior distribution, when the prior distribution for the unknown variables is flat. However, the goal of a Bayesian analysis is generally not just a point estimate like the posterior mode (mean or median), but a representation of the entire distribution for the unknown parameter(s) (Gelman, Carlin, Stern, Rubin, [18] 1995, page 301).

Kyriakoussis and Papadopoulos [19] (1993) derived the Bayes estimators of the probability of success and reliability function of the zero-truncated binomial and negative binomial distributions. In this paper we have made an attempt to obtain Bayes estimator of the parameter and reliability function of the zero-truncated Poisson distribution. Furthermore, recurrence relations for the estimator of the parameter are also derived. Monte Carlo simulation technique has been made for comparing the Bayes estimator and reliability function with the corresponding maximum likelihood estimator (MLE) of zero-truncated Poisson distribution.

2. BAYESIAN ESTIMATOR OF PARAMETER OF ZERO-TRUNCATED POISSON DISTRIBUTION

Let x_1, x_2, \dots, x_n be a random sample from (2). The likelihood function is given by

$$L = \frac{\bar{e}^{-n\lambda} \lambda^{\sum x_i} (1 - \bar{e}^{-\lambda})^{-n}}{\prod_{i=1}^n x_i!} = c \bar{e}^{-n\lambda} \lambda^y (1 - \bar{e}^{-\lambda})^{-n} \quad (3)$$

where $y = \sum x_i$ and $c = \prod_{i=1}^n x_i!$

Regarding the parameter λ in (2), as a random variable, a natural conjugate of its prior distribution is the gamma distribution, given as

$$g(\lambda/\alpha, \beta) = \frac{\alpha^\beta}{\Gamma(\beta)} \bar{e}^{\alpha\lambda} \lambda^{\beta-1}, \quad \alpha, \beta > 0, \quad \lambda > 0 \quad (4)$$

Using Bayes theorem, the posterior distribution of λ from (3) and (4) can be shown to be

$$\begin{aligned} \prod(\lambda/y) &= \frac{L g(\lambda/\alpha, \beta)}{\int_0^\infty L g(\lambda/\alpha, \beta) d\lambda} \\ &= \frac{\bar{e}^{(\alpha+n)\lambda} \lambda^{(y+\beta)-1} (1 - \bar{e}^{-\lambda})^{-n}}{\int_0^\infty \bar{e}^{(\alpha+n)\lambda} \lambda^{(y+\beta)-1} (1 - \bar{e}^{-\lambda})^{-n} d\lambda} \end{aligned} \quad (5)$$

Under squared error loss function, the Bayes estimator of λ is the posterior mean

$$\begin{aligned} p^*(y, n) &= \int_0^\infty \lambda \prod(\lambda/y) d\lambda \\ &= \frac{\int_0^\infty \bar{e}^{(n+\alpha)\lambda} \lambda^{(y+\beta)} (1 - \bar{e}^{-\lambda})^{-n} d\lambda}{\int_0^\infty \bar{e}^{(n+\alpha)\lambda} \lambda^{(y+\beta)-1} (1 - \bar{e}^{-\lambda})^{-n} d\lambda} \end{aligned} \quad (6)$$

Using identity

$$(1-z)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} z^k, \quad |z| < 1$$

and the relation

$$\int_0^\infty \bar{e}^{at} t^{b-1} dt = \Gamma(b)/a^b, \quad a, b > 0, \quad t > 0$$

Where

$$\Gamma(b) = \int_0^\infty \bar{e}^{-t} t^{b-1} dt$$

we obtain,

$$\int_0^\infty \bar{e}^{(n+\alpha)\lambda} \lambda^{(y+\beta)} (1 - \bar{e}^{-\lambda})^{-n} d\lambda = \int_0^\infty \bar{e}^{(n+\alpha)\lambda} \lambda^{(y+\beta)} \sum_{k=0}^{\infty} \binom{n+k-1}{k} (\bar{e}^{-\lambda})^k d\lambda$$

$$\begin{aligned}
&= \sum_{k=0}^{\infty} \binom{n+k-1}{k} \int_0^{\infty} \bar{e}^{-(n+\alpha+k)\lambda} \lambda^{(y+\beta)} d\lambda \\
&= \Gamma(y+\beta+1) \sum_{k=0}^{\infty} \binom{n+k-1}{k} \frac{1}{(\alpha+n+k)^{y+\beta+1}}
\end{aligned} \tag{7}$$

$$= \Gamma(y+\beta+1) \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha+k)^{y+\beta+1}} \tag{8}$$

and similarly,

$$\int_0^{\infty} \bar{e}^{(n+\alpha)\lambda} \lambda^{y+\beta-1} (1-\bar{e}^{\lambda})^{-n} d\lambda = \Gamma(y+\beta) \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha+k)^{y+\beta}} \tag{9}$$

Substituting (8) and (9) into (6) and using relations $\Gamma(b+1) = b\Gamma b$ and

$$k \binom{k-1}{n-1} = n \binom{k-1}{n} + n \binom{k-1}{n-1}, \tag{10}$$

we get

$$\begin{aligned}
p^*(y, n, \alpha, \beta) &= \frac{(y+\beta) \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha+k)^{y+\beta+1}}}{\sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{(k+\alpha)}{(\alpha+k)^{y+\beta+1}}} \\
&= \frac{(y+\beta) \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha+k)^{y+\beta+1}}}{n \sum_{k=n+1}^{\infty} \binom{k-1}{n} \frac{1}{(\alpha+k)^{y+\beta+1}} + (n+\alpha) \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha+k)^{y+\beta+1}}}
\end{aligned} \tag{11}$$

or

$$p^*(y, n) = \frac{(y+\beta)M(y, n)}{nM(y, n+1) + (n+\alpha)M(y, n)} \tag{12}$$

$$y = n, n+1, \dots, \quad n = 1, 2, \dots$$

where

$$M(y, n) = \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha+k)^{y+\beta+1}} \tag{13}$$

and

$$M(y, n+1) = \sum_{k=n+1}^{\infty} \binom{k-1}{n} \frac{1}{(\alpha+k)^{y+\beta+1}} \tag{14}$$

After simplification (12) becomes

$$p^*(y, n) = \frac{(y + \beta)}{\{(n + \alpha) + n B(y, n)\}} \quad (15)$$

$y = n, n + 1, \dots, \quad n = 1, 2, \dots$
where

$$B(y, n) = \frac{M(y, n + 1)}{M(y, n)} \quad (16)$$

3. RECURRENCE RELATIONS

In order to obtain a recurrence relation for $p^*(y, n)$, first we need recurrence relations for the numbers $M(y, n)$ and $B(y, n)$, which are obtained by following two lemmas:

LEMMA1: The numbers $M(y, n)$, satisfy the recurrence relations:

$$M(y, n + 1) = \frac{1}{n} M(y - 1, n) - \frac{(n + \alpha)}{n} M(y, n) \quad (17)$$

$y = n, n + 1, n + 2, \dots, \quad n = 1, 2, \dots$
with initial condition

$$M(y, 1) = \sum_{k=1}^{\infty} \frac{1}{(\alpha + k)^{y+\beta+1}}, \quad y = 1, 2, 3, \dots \quad (18)$$

Proof: From the relation (13), we have

$$\begin{aligned} M(y - 1, n) &= \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha + k)^{(y-1)+\beta+1}} \\ &= \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{(\alpha + k)}{(\alpha + k)^{y+\beta+1}} \end{aligned} \quad (19)$$

Using the relation (10) then (3.3) becomes

$$M(y - 1, n) = n \sum_{k=n+1}^{\infty} \binom{k-1}{n} \frac{1}{(\alpha + k)^{y+\beta+1}} + (n + \alpha) \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha + k)^{y+\beta+1}}$$

From (13) and (14) we have

$$M(y - 1, n) = n M(y, n + 1) + (n + \alpha) M(y, n) \quad (20)$$

from which, we have (17). Also from (13) for $n=1$ we have the relation (18)

Remark 1: Since, α is a positive integer and $\beta > 0$, from (13) we have

$$M(y, 1) = \sum_{K=1}^{\infty} \frac{1}{(\alpha + K)^{y+\beta+1}}$$

$$\begin{aligned}
&= \frac{1}{y + \beta} \sum_{k=1}^{\infty} \left\{ \frac{1}{(\alpha + K)^{y+\beta}} - \frac{1}{(\alpha + K)^{y+\beta}} \right\} \\
&\leq \frac{1}{y + \beta} \sum_{m=\alpha+1}^{\infty} \left\{ \frac{1}{(m-1)^{y+\beta}} - \frac{1}{m^{y+\beta}} \right\} \\
&= \frac{1}{y + \beta} \cdot \frac{1}{(\alpha + 1 - 1)^{y+\beta}}
\end{aligned}$$

We also have

$$M(y, 1) \geq \frac{1}{(\alpha + 1)^{y+\beta+1}}$$

consequently the series $M(y, 1)$ exists and from (17) by mathematical induction we conclude that the series $M(y, n)$ also exists.

Remark 2: Combining the relations (12), (13) and (14) we get that

$$\begin{aligned}
B(y, n) &= \frac{\frac{1}{n}M(y-1, n) - \frac{1}{n}(n + \alpha)M(y, n)}{M(y, n)} \\
&= \left[\frac{M(y-1, n)}{M(y, n)} - (\alpha + n) \right] / n
\end{aligned} \tag{21}$$

and

$$\begin{aligned}
p^*(y, n) &= (y + \beta)M(y, n) / \left[(n + \alpha)M(y, n) + n \left(\frac{1}{n}(y-1, n) - \frac{1}{n}(n + \alpha)M(y, n) \right) \right] \\
&= \frac{(y + \beta)M(y, n)}{M(y-1, n)}
\end{aligned} \tag{22}$$

LEMMA 2: The numbers $B(y, n)$, satisfy the recurrence relations:

$$B(y, n + 1) = \frac{[nB(y, n) + (n + \alpha)]B(y-1, n)}{(n + 1)B(y, n)} - \frac{[(n + 1) + \alpha]}{(n + 1)} \tag{23}$$

$n = 1, 2, \dots, \quad y = n, n + 1, \dots$

with initial conditions

$$B(y, 1) = \frac{M(y-1, 1)}{M(y, 1)} - (1 + \alpha) \tag{24}$$

Proof: From the relation (14) and the recurrence relation (17), we get

$$\begin{aligned}
B(y, n + 1) &= \frac{M(y, n + 2)}{M(y, n + 1)} \\
&= \frac{\frac{1}{(n+1)}M(y-1, n+1) - \frac{1}{n+1}(n+1 + \alpha)M(y, n+1)}{M(y, n+1)}
\end{aligned}$$

$$= \frac{1}{(n+1)} \frac{M(y-1, n+1)/M(y, n) - [(n+1) + \alpha]B(y, n)}{B(y, n)} \quad (25)$$

We also have,

$$\begin{aligned} B(y-1, n) &= \frac{M(y-1, n+1)}{M(y-1, n)} = \frac{M(y-1, n+1)/M(y, n)}{M(y-1, n)/M(y, n)} \\ &= \frac{M(y-1, n+1)/M(y, n)}{n M(y, n+1)/M(y, n) + (n + \alpha)} \end{aligned}$$

From (16) and (20) we have

$$B(y-1, n) = \frac{M(y-1, n+1)/M(y, n)}{n B(y, n) + (n + \alpha)} \quad (26)$$

or

$$M(y-1, n+1)/M(y, n) = [n B(y, n) + (n + \alpha)]B(y-1, n) \quad (27)$$

Substituting (26) into (25) we obtain (23). Using the relation (22) for $n=1$ we easily obtain the initial conditions (23).

Theorem 1: The Bayes estimator of the parameter λ satisfies the recurrence relation:

$$p^*(y, n+1) = \frac{[(y + \beta) - (n + \alpha)]p^*(y, n)}{[(y - 1 + \beta) - (n + \alpha)]p^*(y - 1, n)} \quad (28)$$

$y = n, n+1, \dots, n=1, 2, 3, \dots$

with initial conditions

$$p^*(y, 1) = \frac{(y + \beta)M(y, 1)}{M(y - 1, 1)} \quad (29)$$

Proof: From the relation (15) we have

$$p^*(y, n+1) = \frac{(y + \beta)}{(n + 1 + \alpha) + (n + 1)B(y, n+1)} \quad (30)$$

Substituting (23) into (30) and using (15) we obtain (28), after some algebraic manipulations. From the relation (22) for $n = 1$ we easily get (29).

4. BAYES ESTIMATOR OF THE RELIABILITY FUNCTION OF ZERO-TRUNCATED POISSON DISTRIBUTION

The Bayes estimator $R^*(t)$, for $R(t) = P(X > t)$, where the random variable X has the distribution (2), is given by

$$R^*(t) = E[R(t)/x_1, x_2, \dots, x_n]$$

$$\begin{aligned}
& \int_0^{\infty} \mathbf{R}(t) \bar{e}^{(n+\alpha)\lambda} \lambda^{(y+\beta)-1} (1 - \bar{e}^\lambda)^{-n} d\lambda \\
&= \frac{\int_0^{\infty} \bar{e}^{(n+\alpha)\lambda} \lambda^{(y+\beta)-1} (1 - \bar{e}^\lambda)^{-n} d\lambda}{\int_0^{\infty} \bar{e}^{(n+\alpha)\lambda} \lambda^{(y+\beta)-1} (1 - \bar{e}^\lambda)^{-n} d\lambda}
\end{aligned} \tag{31}$$

where

$$\mathbf{R}(t) = \sum_{x=[t]+1}^{\infty} \frac{\bar{e}^\lambda \lambda^x}{x!} (1 - \bar{e}^\lambda)^{-1}, \quad [t], \text{ the integer part of } t. \tag{32}$$

Making similar computations, as for $p^*(y, n)$ we get

$$\mathbf{R}^*(t) = \sum_{x=[t]+1}^{\infty} \frac{\int_0^{\infty} \mathbf{R}(t) \bar{e}^{(\alpha+n+1)\lambda} \lambda^{(y+\beta+x)-1} (1 - \bar{e}^\lambda)^{-(n+1)} d\lambda}{\int_0^{\infty} \bar{e}^{(n+\alpha)\lambda} \lambda^{(y+\beta)-1} (1 - \bar{e}^\lambda)^{-n} d\lambda} \tag{33}$$

Using the identity

$$(1 - z)^{-n} = \sum_{k=0}^{\infty} \binom{n+k-1}{k} z^k, \quad |z| < 1$$

we obtain

$$\int_0^{\infty} \bar{e}^{(\alpha+n+1)\lambda} \lambda^{(y+\beta+n)-1} (1 - \bar{e}^\lambda)^{-(n+1)} d\lambda = \Gamma(y + \beta + x) \sum_{k=n+1}^{\infty} \binom{k-1}{n} \frac{1}{(\alpha + k)^{(y+\beta+x)}} \tag{34}$$

Similarly,

$$\int_0^{\infty} \bar{e}^{(n+\alpha)\lambda} \lambda^{(y+\beta)-1} (1 - \bar{e}^\lambda)^{-n} d\lambda = \Gamma(y + \beta) \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha + k)^{y+\beta}} \tag{35}$$

Using (35) and (34) in (33), we get

$$\mathbf{R}^*(t) = \sum_{x=[t]+1}^{\infty} \frac{\Gamma(y + \beta + x)}{x! \Gamma(y + \beta)} \frac{\sum_{k=n+1}^{\infty} \binom{k-1}{n} \frac{1}{(\alpha + k)^{y+\beta+x}}}{\sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha + k)^{y+\beta}}} \tag{36}$$

Using the relation (10), we get

$$\mathbf{R}^*(t) = \sum_{x=[t]+1}^{\infty} \frac{\Gamma(y + \beta + x)}{x! \Gamma(y + \beta)} \frac{\sum_{k=n+1}^{\infty} \binom{k-1}{n} \frac{(\alpha + k)}{(\alpha + k)^{y+\beta+x+1}}}{\sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{(\alpha + k)}{(\alpha + k)^{y+\beta+1}}}$$

$$\begin{aligned}
 &= \sum_{x=[t]+1}^{\infty} \frac{\Gamma(y+\beta+x)}{x!\Gamma(y+\beta)} \left[\frac{(n+1) \sum_{k=n+2}^{\infty} \binom{k-1}{n+1} \frac{1}{(\alpha+k)^{y+x+\beta+1}} + (n+1+\alpha) \sum_{k=n+1}^{\infty} \binom{k-1}{n}}{n \sum_{k=n+1}^{\infty} \binom{k-1}{n} \frac{1}{(\alpha+k)^{y+\beta+1}} + (n+\alpha) \sum_{k=n}^{\infty} \binom{k-1}{n-1}} \right. \\
 &\qquad \qquad \qquad \left. \frac{1}{(\alpha+k)^{y+x+\beta+1}} \right] \\
 &= \sum_{x=[t]+1}^{\infty} \frac{\Gamma(y+\beta+x)}{\Gamma(y+\beta)} \left[\frac{(n+1)M(y+x, n+2) + (n+1+\alpha)M(y+x, n+1)}{(n+\alpha)M(y, n) + nM(y, n+1)} \right] \quad (37)
 \end{aligned}$$

Where

$$M(y, n) = \sum_{k=n}^{\infty} \binom{k-1}{n-1} \frac{1}{(\alpha+k)^{y+\beta+1}}$$

Or

$$R^*(t) = 1 - \sum_{x=0}^{[t]} \frac{1}{x!} \frac{\Gamma(y+\beta+x)}{\Gamma(y+\beta)} \left[\frac{(n+1)M(y+x, n+2) + (n+1+\alpha)M(y+x, n+1)}{(n+\alpha)M(y, n) + nM(y, n+1)} \right] \quad (38)$$

5. COMPUTER SIMULATION AND CONCLUSION

In order to compare the estimators, Monte Carlo Simulations were performed on 1000 samples for each simulation. The following steps summarize the simulation,

- 1) A value is generated from a gamma distribution with specified parameters α and β .
- 2) Based on the realization from the gamma distribution a sample of size $n=8$ or 30 is generated from the zero-truncated Poisson distribution.
- 3) The estimates of the parameter and reliability function are computed from the generated sample, and the estimates and their squared error losses were stored.
- 4) Steps 1-3 were repeated 1000 times.
- 5) Average values and root mean square errors (RMSE's) of the estimates are computed over the 1000 samples.

Tables 1-4 show some of the results. In comparing the estimators the root mean square error criterion will be used, namely the estimator with the smallest RMSE is the best estimator. The reliability function was evaluated arbitrarily at times 1, 2 and 3. Two sample sizes of $n=8, 30$ were utilized in the simulation.

TABLE 1. Average values and RMSE's for the estimators of the zero- truncated Poisson.

Gamma prior with $\alpha = 1$ and $\beta = 1$, Sample Size $n=8$

Parameter						
True Value		Bayes Ave. RMSE		MLE Ave. RMSE		RMSE ratio MLE/Bayes
4.0832		4.0776	3.6190	4.0784	3.6200	1.0003
Reliability						
Time	Exact Value	Bayes Ave. RMSE		MLE Ave. RSME		RMSE ratio MLE/Bayes
1	4.4816	4.4724	3.6230	4.4695	3.6264	1.0009
2	4.3879	4.3769	3.6392	4.3779	3.6408	1.0004
3	4.2785	4.2681	3.6523	4.2699	3.6542	1.0005

TABLE 2. Average values and RMSE's for the estimators of the zero- truncated Poisson.

Gamma prior with $\alpha = 2$ and $\beta = 5$, Sample Size $n=8$

Parameter						
True Value		Bayes Ave. RMSE		MLE Ave. RMSE		RMSE ratio MLE/Bayes
2.1904		2.1912	1.9665	2.1792	1.9708	1.0022
Reliability						
Time	Exact Value	Bayes Ave. RMSE		MLE Ave. RSME		RMSE ratio MLE/Bayes
1	2.7034	2.6915	2.0026	2.6771	2.0142	1.0058
2	2.4588	2.4526	2.0305	2.4403	2.0396	1.0045
3	2.2387	2.2435	2.0244	2.2321	2.0324	1.0040

TABLE 3. Average values and RMSE's for the estimators of the zero- truncated Poisson.

Gamma prior with $\alpha = 1$ and $\beta = 1$, Sample Size $n=30$

Parameter						
True Value		Bayes Ave. RMSE		MLE Ave. RMSE		RMSE ratio MLE/Bayes
15.1380		15.1020	14.6255	15.1020	14.6256	1.0000
Reliability						
Time	Exact Value	Bayes Ave. RMSE		MLE Ave. RSME		RMSE ratio MLE/Bayes

1	15.5245	15.4889	14.6310	15.4882	14.6311	1.0000
2	15.4269	15.3892	14.6370	15.3890	14.6372	1.0000
3	15.3205	15.2819	14.6420	15.2821	14.6422	1.0000

TABLE 4. Average values and RMSE's for the estimators of the zero- truncated Poisson. Gamma prior with $\alpha = 2$ and $\beta = 5$, Sample Size n=30

Parameter						
True Value		Bayes		MLE		RMSE ratio
		Ave.	RMSE	Ave.	RMSE	MLE/Bayes
8.5200		8.4870	8.2306	8.4840	8.2311	1.0001
Reliability						
Time	Exact Value	Bayes		MLE		RMSE ratio
		Ave.	RMSE	Ave.	RSME	MLE/Bayes
1	9.0363	8.9992	8.2502	8.9961	8.2520	1.0002
2	8.8006	8.7641	8.2642	8.7617	8.2655	1.0002
3	8.5798	8.5466	8.2619	8.5438	8.2630	1.0001

In comparing the estimators, the Bayes ones have the smallest RMSE and are better. This is to be expected since the Bayes estimators take advantage of the known prior parameters α and β . By examining the RMSE ratios we can conclude that the estimates are sensitive to the choice of prior parameters and to sample size.

ACKNOWLEDGMENTS

The authors thank very much the anonymous referee and editor for his/her many helpful suggestions, which substantially simplified the paper.

REFERENCE

- [1] Poisson, S.D.: *Recherches sur la probabilité des Jugements en Matière Criminelle et en Matière Civile*, Precedees des Regles. Generales du calcul des Probabilities. Paris: Bachelier, Imprimeur Libraire pour les mathematiques, la physique, etc, 1837.
- [2] Newbold, E.M.: *Practical applications of the statistics of repeated events particularly to industrial accidents*, Jour. Royal Statist. Soc., 90 (1927), 518-535.
- [3] Jensen, A.: *A distribution model applicable to economics*, Munksguard, Copenhagen, 1954.
- [4] David, F.N.: *Games, Gods and Gambling*, Griffin, London 1962.
- [5] Moivre, A. de.: *The Doctrine of chances: or, A Method of Calculating the probability of events in Play*, London: Pearson, 1718.
- [6] Johnson, N.L., Kotz, S. and Kemp, A.W.: *Univariate discrete distributions* (second edition), John Wiley and Sons, New York, 1992.
- [7] Ahmad, M. and Roohi, A.: *Characterizations of the Poisson probability distribution*, Pakistan Journal of Statistics, 20(2) (2004), 301-304.
- [8] Roohi, A. and Ahmad, M.: *Estimation of Characterization of the parameter of Hyper-Poisson distribution using negative moments*, Pakistan Journal of Statistics, 19(1) (2003), 99-105.
- [9] Cohen, A.C.: *Estimating the parameter in a conditional Poisson distribution*, Biometrics, 16 (1960), 203-211.
- [10] David, F.N and Johnson, N.L.: *The truncated Poisson*, Biometrics, 8(1952), 275-285.

- [11] Plackett, R.L.: *The truncated Poisson distributions*, Biometrics, 9(1953), 485-488.
- [12] Murakami, M.: *Censored sample from truncated Poisson distribution*, Journal of the College of Arts and Sciences, Chiba University, 3(1961), 263-268.
- [13] Tate, R.F. and Goen, R.L.: *Minimum variance unbiased estimation for the truncated Poisson distribution*, Annals of Mathematical Statistics, 29 (1958), 755- 765.
- [14] Cohen A.C: *An extension of a truncated Poisson distribution*, Biometrics, 16(1960), 446-450.
- [15] Cohen, A.C.: *Estimation in a truncated Poisson distribution when zeroes and some ones are missing*, Journal of the American Statistical Association, 55(1960), 342- 348
- [16] Johnson, N.L. and Kotz, S.: *Discrete distributions* (First edition), John Wiley and Sons, New York 1969.
- [17] Consul, P.C.: *Generalized Poisson Distribution and its Applications*, Marcel Dekker, New York 1989.
- [18] Gelman, A., Carlin, J.B. Stern, H.S. and Rubin, D.B.: *Bayesian Data Analysis*, Chapman and Hall, New York 1995.
- [19] Kyriakoussis, A. and Papadopoulos, Alex S.: *On the Bayes estimators of the probability of success and reliability function of the zero-truncated binomial and negative binomial distributions*, Sankhya, 55(1993), Series B, Pt.2, 171-185.

ON ESTIMATION OF NEGATIVE POLYA-EGGENBERGER DISTRIBUTION AND ITS APPLICATIONS

ANWAR HASSAN^{1†} AND SHEIKH BILAL²

¹DEPARTMENT OF STATISTICS, UNIVERSITY OF KASHMIR, SRINAGAR INDIA

E-Mail address: anwar.hassan2007@gmail.com

²DEPARTMENT OF STATISTICS, DEGREE COLLEGE, BARAMULLA KASHMIR INDIA

E-Mail address: sbilal_sbilal@yahoo.com

ABSTRACT. In this paper, the negative Polya-Eggenberger distribution has been introduced by compounding negative binomial distribution with beta distribution of I-kind which generates a number of univariate contagious or compound (or mixture of) distributions as its particular cases. The distribution is unimode, over dispersed and all of its positive and negative integer moments exist. The difference equation of the proposed model shows that it is a member of the Ord's family of distribution. Some of its interesting properties have been explored besides different methods of estimation been discussed. Finally, the parameters of the proposed model have been estimated by using a computer programme in R-software. Application of the proposed model to some data, available in the literature, has been given and its goodness of fit demonstrated.

1. INTRODUCTION

The Polya-Eggenberger distribution and its inverse negative Polya-Eggenberger distribution was introduced by Polya and Eggenberger [1] through an urn model. They described these distributions as truly contagious distribution. The other type such as Neyman's contagious distribution [2] involves the "apparent contagion" as described by Feller [3]. The distribution in discussion has not been studied in detail so far. Though, its different parameterization is present in the literature. The first parameterization of the distribution is the beta-negative binomial distribution which is the generalization of hypergeometric distribution; see Johnson and Kotz's [10] for details. Another parameterization is the generalized Waring distribution introduced by Irwin [9] as an accident proneness-liability model which was subsequently applied by Irwin {[9], [14]} to data on accidents sustained by men in a soap factory. The negative Polya-Eggenberger distribution also belongs to Kemp and Kemp [6] generalized hyper-geometric distribution.

The negative Polya-Eggenberger distribution is related to Polya-Eggenberger

Received by the editors January 18, 2008.

2000 *Mathematics Subject Classification.* 62E15.

Key words and phrases. Negative binomial distribution, negative Polya-Eggenberger distribution, recurrence relations, probability generating function, chi-square fitting.

† Corresponding author.

distribution in the same way as negative binomial distribution is related to binomial distribution. It is well known fact that the negative binomial distribution has become increasingly popular as a more flexible alternative to the Poisson distribution especially when it is doubtful whether the strict requirements particularly independence for a Poisson distribution will be satisfied. There are various extensions/modifications of NBD in the literature including Engen's extended NBD {[13], [15]}, GNBD of Jain and Consul [12] and Weighted NBD; see Johnson et al. [16] for more details and explanations.

In this paper, the proposed model has been obtained by compounding the negative binomial distribution with beta distribution of first kind. In Gurland's [7] terminology, the proposed model represents a generalization of the negative binomial distribution. It has been shown that the proposed model exhibits more flexible alternative model to negative-binomial distribution and some of its generalizations. This has been demonstrated with the help of three data sets by the goodness of fit in the last section of this paper.

The organization of the paper is as follows: in section 2, we derive the proposed model. Section 3 deals with some interesting structural properties of the proposed model. In section 4, we discuss its relation with other distributions. Section 5 proposes different methods of estimation. Finally, in section 6, the parameters of the proposed model have been estimated by using a computer programme in R-software. Further, an application of the proposed model to three data sets has been given and its goodness of fit demonstrated.

2. THE PROPOSED MODEL

A certain mixture distribution arises when all (or some) parameters of a distribution vary according to some probability distribution called the mixing distribution. A well-known example of discrete-type mixture distribution is the negative-binomial distribution which can be obtained as a Poisson mixture with gamma distribution.

Let X has a conditional negative-binomial distribution with parameter p , that is, X has a conditional probability mass function (pmf)

$$P(X/p) = P(X = x/p) = \binom{n+x-1}{x} p^x (1-p)^n \quad \text{for } x=0,1,2,\dots \text{ and } 0 < p < 1, n > 0 \quad (2.1)$$

Now, suppose p is a continuous random variable with probability density function (pdf)

$$g(p) = \frac{1}{\beta(\gamma, \alpha)} p^{\gamma-1} (1-p)^{\alpha-1} \quad \text{for } 0 < p < 1, \quad (\alpha, \gamma) > 0 \quad (2.2)$$

Bhattacharya [8] showed that the conditional pmf of X is given by

$$f(x) = P(X = x) = \int_0^1 f(x/p) g(p) dp$$

The equation above together with (2.1) and (2.2) gives

$$P(X = x) = \binom{n+x-1}{x} \frac{\alpha(\alpha+1)\dots(\alpha+x-1)\gamma(\gamma+1)\dots(\gamma+n-1)}{(\alpha+\gamma)(\alpha+\gamma+1)\dots(\alpha+\gamma+n+x-1)}$$

Taking $\alpha = a/c$, $\gamma = b/c$, the equation above reduces to the negative Polya-Eggenberger distribution with pmf

$$P(X = x) = \binom{n+x-1}{x} \frac{a(a+c)\dots(a+x-1c)b(b+c)\dots(b+n-1c)}{(a+b)(a+b+c)\dots(a+b+n+x-1c)} \text{ for } x=0,1,2,\dots \quad (2.3)$$

The proposed model (2.3) can be put into different forms for the mathematical convenience and to study some of its properties. The model (2.3) in terms of ascending factorials can be put as

$$P(X = x) = \binom{n+x-1}{x} \frac{a^{[x,c]} b^{[n,c]}}{(a+b)^{[n+x,c]}} , \text{ for } x=0,1,2,\dots \quad (2.4)$$

Where $a^{[x,c]} = a(a+c)\dots(a+x-1c)$

Another form of (2.3) can be

$$P(X = x) = \binom{n+x-1}{x} \frac{\alpha^{[x]} \gamma^{[n]}}{(\alpha+\gamma)^{[n+x]}} , \text{ for } x=0,1,2,\dots \quad (2.5)$$

Where $\alpha^{[x]} = \alpha(\alpha+1)\dots(\alpha+x-1)$ and $\alpha = a/c$, $\gamma = b/c$. The model represented by (2.5) has been seen the most workable model, used through out this paper, for the mathematical computations.

Another form of (2.3) in terms of n , $p = \frac{a}{(a+b)}$, $Q = (1-p) = \frac{b}{(a+b)}$ and $\delta = \frac{c}{(a+b)}$ can be

$$P(X = x) = \binom{n+x-1}{x} \frac{\prod_{j=0}^{x-1} (p + j\delta) \prod_{j=0}^{n-1} (Q + j\delta)}{\prod_{j=0}^{n+x-1} (1 + j\delta)} , \text{ for } x=0,1,2,\dots \quad (2.6)$$

REMARKS: A number of special cases can be deduced from the proposed model (2.3) by assigning different set of values to its parameters. Some of the interesting cases deduced are negative-binomial distribution, beta-negative binomial distribution, negative hyper geometric distribution, geometric series distribution, Bernoulli-delta distribution (geometric) etc.

3. STRUCTURAL PROPERTIES

In this section, some of the interesting properties of the proposed model has been explored which are described as follows;

3.1 RECURRENCE RELATION BETWEEN PROBABILITIES

Expressing the pmf of the proposed model (2.5) as

$$P(X=x) = \frac{(n+x-1)!}{(n-1)!} \frac{\alpha^{x-1} \gamma^{n-1}}{x! (\alpha+\gamma)^{n+x}} \quad (3.1)$$

Taking $x = x+1$ in the equation above and dividing the resulting equation by (3.1), we get the recurrence relation of the proposed model as

$$P(X=x+1) = \left[\frac{(n+x)}{(x+1)} \frac{(\alpha+x)}{(\alpha+\gamma+n+x)} \right] P(X=x) \quad (3.2)$$

Which yields the difference equation of the proposed model as

$$\Delta P_{x-1} = \frac{\left\{ \frac{n\alpha-n-\alpha}{\gamma+1} - x \right\} P_x}{\left\{ \frac{n\alpha-n-\alpha}{\gamma+1} + \frac{1}{\gamma+1} \right\} + \left\{ \frac{n+\alpha+\gamma}{\gamma+1} - 1 \right\} x + \frac{x(x+1)}{\gamma+1}}$$

The difference equation above exhibits that the proposed model is a member of the Ord's family of distribution.

3.2. UNIMODALITY

The proposed model is a unimodal by the following result of Holgate [11]:

LEMMA. *If the mixing distribution is non-negative, continuous and unimodal then the resulting distribution is unimodal*

The proposed model is a unimodal since the mixing distribution is a beta distribution of I-kind which is unimodal. To show the unimodality of the distribution we have the following theorem.

Theorem 3.1. The proposed model is a unimodal for all values of (n, α, γ) and the mode is at $x=0$ if $n\alpha < 1$ and for $n\alpha > 1$ the mode is at some other point $x=M$ such that

$$\frac{n(\alpha-1) - (\alpha+\gamma)}{(\gamma+1)} < M < \frac{(n-1)(\alpha+1)}{(\gamma+1)} \quad (3.3)$$

Proof. The recurrence relation (3.2) gives the ratio

$$\frac{P(x+1)}{P(x)} = \frac{(n+x)}{(x+1)} \frac{(\alpha+x)}{(\alpha+\gamma+n+x)} \quad (3.4)$$

Which is less than one, that is,

$$\frac{P(x+1)}{P(x)} < 1 \text{ if } n\alpha < 1 \quad \forall (n, \alpha, \gamma) > 0$$

Hence, for $n\alpha < 1$, the ratio $\frac{P(x+1)}{P(x)}$ is a non-increasing function, therefore, the mode of the proposed model exists at $x=0$. Suppose for $n\alpha > 1$ the mode exists at $x=M$, then the ratio defined by (3.4) gives the two inequalities

$$\frac{P(M+1)}{P(M)} = \frac{(n+M)}{(M+1)} \frac{(\alpha+M)}{(\alpha+\gamma+n+M)} < 1 \quad (3.5)$$

and

$$\frac{P(M)}{P(M-1)} = \frac{n+M-1}{M} \frac{(\alpha+M-1)}{(\alpha+\gamma+n+M-1)} > 1 \tag{3.6}$$

By the inequality (3.5) we have

$$\frac{n(\alpha-1)-(\alpha+\gamma)}{(\gamma+1)} < M \tag{3.7}$$

and the inequality (3.6) gives

$$M < \frac{(n-1)(\alpha+1)}{(\gamma+1)} \tag{3.8}$$

On combining the inequalities (3.7) and (3.8), we get the result (3.3).

3.3. MEAN AND VARIANCE

Mean and variance of the proposed model can be easily obtained by using the properties of conditional mean and variance as follows;

MEAN: By the conditional mean we have

$$Mean = E(X) = E[E(X/p)] \tag{3.9}$$

Where $E(X/p)$ is the conditional expectation of X given p and for given p the random variable X follows (2.1) with mean and variance given by

$$\left. \begin{aligned} E(X/p) &= np^{-1}(1-p) \\ V(X/p) &= np^{-2}(1-p) \end{aligned} \right\} \tag{3.10}$$

The equation (3.9) together with (3.10) and (2.2) gives mean of the proposed model as

$$E(X) = \frac{n\alpha}{(\gamma-1)} \quad \gamma > 1$$

VARIANCE: Similarly, by the conditional variance we have

$$V(X) = E[V(X/p)] + V[E(X/p)] \tag{3.11}$$

Using (3.10) in the equation above, we get

$$V(X) = nE[p^{-2}(1-p)] + n^2E[p^{-2}(1-p)^2] - n^2\{E[p^{-1}(1-p)]\}^2$$

Since p is varying as (2.2), the equation above reduces to

$$V(X) = \frac{n}{\beta(\gamma, \alpha)} \int_0^1 p^{\gamma+1-1}(1-p)^{\alpha-3} dp + \frac{n^2}{\beta(\alpha, \gamma)} \int_0^1 p^{\gamma+2-1}(1-p)^{\alpha-3} dp - \{E(X)\}^2$$

By an application of beta integral, the equation above gives variance as

$$V(X) = \frac{n\alpha}{(\gamma-1)} + \frac{n(n+1)\alpha(\alpha+1)}{(\gamma-1)(\gamma-2)} - \left(\frac{n\alpha}{(\gamma-1)}\right)^2 \quad \text{for } \gamma > 2$$

3.4. RECURRENCE RELATION BETWEEN MOMENTS

The recurrence relation (3.2) gives

$$(1+x)^{(r+1)}P_{x+1}(n,\alpha,\gamma)=(1+x)^r\frac{(\alpha+x)(n+x)}{(\alpha+\gamma+n+x)}P_x(n,\alpha,\gamma)$$

Which subsequently reduces to

$$(1+x)^{(r+1)}P_{x+1}(n,\alpha,\gamma)=(1+x)^r\frac{n\alpha}{(\gamma-1)}P_x(n+1,\alpha+1,\gamma-1), \quad \gamma > 1 \quad (3.12)$$

Where $P_x(n+1,\alpha+1,\gamma-1)$ denotes the pmf of the proposed model with parameters $(n+1,\alpha+1,\gamma-1)$. Summing (3.12) over the values of x on both sides, we get the moment recurrence relation as

$$\mu'_{r+1}(n,\alpha,\gamma)=\frac{n\alpha}{(\gamma-1)}\sum_{j=0}^r\binom{r}{j}\mu'_j(n+1,\alpha+1,\gamma-1) \quad (3.13)$$

Where $\mu'_j(n+1,\alpha+1,\gamma-1)$ denotes the j th moment about origin of the proposed model with parameters $(n+1,\alpha+1,\gamma-1)$. The recurrence relation (3.13) gives the first four moments about origin as

$$\begin{aligned} \mu'_1 &= \frac{n\alpha}{(\gamma-1)}, & \gamma > 1 \\ \mu'_2 &= \frac{n\alpha}{(\gamma-1)} + \frac{n(n+1)\alpha(\alpha+1)}{(\gamma-1)(\gamma-2)}, & \gamma > 2 \\ \mu'_3 &= \frac{n\alpha}{(\gamma-1)} + \frac{3n(n+1)\alpha(\alpha+1)}{(\gamma-1)(\gamma-2)} + \frac{n(n+1)(n+2)\alpha(\alpha+1)(\alpha+2)}{(\gamma-1)(\gamma-2)(\gamma-3)}, & \gamma > 3 \\ \mu'_4 &= \frac{n\alpha}{(\gamma-1)} + \frac{7n(n+1)\alpha(\alpha+1)}{(\gamma-1)(\gamma-2)} + \frac{6n(n+1)(n+2)\alpha(\alpha+1)(\alpha+2)}{(\gamma-1)(\gamma-2)(\gamma-3)} \\ &+ \frac{n(n+1)(n+2)(n+3)\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{(\gamma-1)(\gamma-2)(\gamma-3)(\gamma-4)}, & \gamma > 4 \end{aligned}$$

Now, the central moments can be easily obtained from the moments about origin of the proposed model and are given by

$$\begin{aligned} \mu_2 &= \frac{n\alpha}{(\gamma-1)}\left[1 + \frac{(n+1)(\alpha+1)}{(\gamma-2)} - \frac{n\alpha}{(\gamma-1)}\right], & \gamma > 2 \\ \mu_3 &= \frac{n\alpha}{(\gamma-1)}\left[1 - \frac{3n\alpha}{(\gamma-1)} + 2\left(\frac{n\alpha}{(\gamma-1)}\right)^2\right] + \frac{3n(n+1)\alpha(\alpha+1)}{(\gamma-1)(\gamma-2)}\left[1 - \frac{n\alpha}{(\gamma-1)}\right] \\ &+ \frac{n(n+1)(n+2)\alpha(\alpha+1)(\alpha+2)}{(\gamma-1)(\gamma-2)(\gamma-3)}, & \gamma > 3 \\ \mu_4 &= \frac{n\alpha}{(\gamma-1)}\left[1 - \frac{4n\alpha}{(\gamma-1)} + 6\left(\frac{n\alpha}{(\gamma-1)}\right)^2 - 3\left(\frac{n\alpha}{(\gamma-1)}\right)^3\right] \\ &+ \frac{n(n+1)\alpha(\alpha+1)}{(\gamma-1)(\gamma-2)}\left[7 - \frac{12n\alpha}{(\gamma-1)} + 6\left(\frac{n\alpha}{(\gamma-1)}\right)^2\right] \end{aligned}$$

$$+ \frac{n(n+1)(n+2)\alpha(\alpha+1)(\alpha+2)}{(\gamma-1)(\gamma-2)(\gamma-3)} \left[6 - \frac{4n\alpha}{(\gamma-1)} \right] \\ + \frac{n(n+1)(n+2)(n+3)\alpha(\alpha+1)(\alpha+2)(\alpha+3)}{(\gamma-1)(\gamma-2)(\gamma-3)(\gamma-4)}, \quad \gamma > 4$$

The proposed model is over dispersed for $\gamma > 2$ and its coefficient of variation is given by

$$C.V. = 1 + \frac{(n+1)(\alpha+1)}{(\gamma-2)} - \frac{n\alpha}{(\gamma-1)}, \quad \gamma > 2$$

3.5. PROBABILITY GENERATING FUNCTION

Suppose $G_x(u)$ denotes the probability generating function of the proposed model then we have

$$G_x(u) = E(u^x) = \sum_{x=0}^{\infty} u^x \frac{(n+x-1)!}{(n-1)! x!} \frac{\alpha^{[x]} \gamma^{[n]}}{(\alpha+\gamma)^{[n+x]}}$$

Which yields the probability generating function of the proposed model as

$$G_x(u) = \frac{\gamma^{[n]}}{(\alpha+\gamma)^{[n]}} {}_2F_1[n, \alpha; \alpha+\gamma+n, u]$$

Where ${}_2F_1[n, \alpha; \alpha+\gamma+n, u]$ is a Gaussian hypergeometric function.

REMARKS: If we replace u by $(1+t)$ and $(1-t)$, the equation above yields the descending factorial moment and ascending factorial moment generating functions respectively.

3.6. FACTORIAL MOMENTS

The r th factorial moment about origin $\mu'_{(r)}$ of the proposed model is defined as

$$\mu'_{(r)} = E(X^{(r)}) = \sum_{x=0}^{\infty} x^{(r)} \frac{(n+x-1)!}{(n-1)! x!} \frac{\alpha^{[x]} \gamma^{[n]}}{(\alpha+\gamma)^{[n+x]}}$$

Which reduces to

$$\mu'_{(r)} = \sum_{x=r}^{\infty} \frac{(n+x-1)!}{(n-1)! (x-r)!} \frac{\alpha^{[x]} \gamma^{[n]}}{(\alpha+\gamma)^{[n+x]}}$$

Taking $x = x+r$, the equation above yields the r th factorial moment of the proposed model as

$$\mu'_{(r)} = \frac{n^{[r]} \alpha^{[r]} (\gamma-r-1)!}{(\gamma-1)!} \tag{3.14}$$

The expression (3.14) can be used to obtain the factorial moments of the proposed model.

3.7. NEGATIVE INTEGER FACTORIAL MOMENTS (NIFM)

The negative moments are useful in many problems of applied statistics, especially in life

testing and in survey sampling, where ratio estimates are used. In this section, we obtained the expression for the r th negative integer ascending factorial moment of the proposed model in terms of Gaussian hypergeometric function. Suppose $\phi'_{-[r]}$ denotes the r th negative integer ascending factorial moment of the proposed mode then we have

$$\phi'_{-[r]} = E[(x+1)^{[r]}]^{-1} = \sum_{x=0}^{\infty} \frac{1}{(x+1)^{[r]}} \frac{(n+x-1)!}{(n-1)! x!} \frac{\alpha^{[x]} \gamma^{[n]}}{(\alpha+\gamma)^{n+x}} \quad (3.15)$$

Where $\frac{1}{(x+1)^{[r]}} = \frac{x!}{(x+r)!} = \frac{1^{[x]}}{1^{[r]}(r+1)^{[x]}}$

Using the result above in (3.15), we get

$$\phi'_{-[r]} = \frac{\gamma^{[n]}}{r!(\alpha+\gamma)^{[n]}} {}_3F_2[n, \alpha, 1; r+1, \alpha+\gamma+n; 1]$$

Which gives the first four negative integer ascending factorial moment as

$$\phi'_{-[1]} = \frac{\gamma^{[n]}}{(\alpha+\gamma)^{[n]}} {}_3F_2[n, \alpha, 1; 2, \alpha+\gamma+n; 1]$$

$$\phi'_{-[2]} = \frac{\gamma^{[n]}}{2(\alpha+\gamma)^{[n]}} {}_3F_2[n, \alpha, 1; 3, \alpha+\gamma+n; 1]$$

$$\phi'_{-[3]} = \frac{\gamma^{[n]}}{6(\alpha+\gamma)^{[n]}} {}_3F_2[n, \alpha, 1; 4, \alpha+\gamma+n; 1]$$

$$\phi'_{-[4]} = \frac{\gamma^{[n]}}{24(\alpha+\gamma)^{[n]}} {}_3F_2[n, \alpha, 1; 5, \alpha+\gamma+n; 1]$$

4. RELATION WITH OTHER DISTRIBUTIONS

Theorem 4.1. Let X be a negative Polya-Eggenberger variate with parameters (n, α, γ) . If $\gamma \rightarrow \infty$ such that $\alpha\gamma^{-1} = \lambda$ and $\lambda = \theta n^{-1}$ as $n \rightarrow \infty$ then show that X tends to a Poisson distribution with parameter θ .

Proof: Expressing the pmf of the proposed model as

$$P(X=x) = \frac{(n+x-1)(n+x-2)\dots(n+1)n}{x!} \frac{\alpha(\alpha+1)\dots(\alpha+x-1)\gamma(\gamma+1)\dots(\gamma+n-1)}{(\alpha+\gamma)(\alpha+\gamma+1)\dots(\alpha+\gamma+n+x-1)} \quad (4.1)$$

Taking limit $\gamma \rightarrow \infty$ such that $\alpha\gamma^{-1} = \lambda$, the equation above gives

$$\lim_{\gamma \rightarrow \infty} P(X=x) = \left(1 + \frac{x-1}{n}\right) \left(1 + \frac{x-2}{n}\right) \dots \left(1 + \frac{1}{n}\right) \frac{(n\lambda)^x}{x!(1+\lambda)^{n+x}}$$

Substituting $\lambda = \frac{\theta}{n}$ and taking limit $n \rightarrow \infty$, the equation above reduces to the Poisson distribution with parameter θ .

Theorem 4.2. Let X be a negative Polya-Eggenberger variate with parameters (n, α, γ) . Show that zero-truncated negative Polya-Eggenberger distribution tends to logarithmic series distribution.

Proof: The pmf of the Zero-truncated negative Polya-Eggenberger distribution is

$$P(X = x) = \binom{n+x-1}{x} \frac{\alpha^{[x]} \gamma^{[n]}}{(\alpha + \gamma)^{[n+x]} (\alpha + \gamma)^{[n]} - \gamma^{[n]}} \quad (\alpha + \gamma)^{[n]} \quad , \quad x = 1, 2, \dots \quad (4.2)$$

Substituting $\alpha\gamma^{-1} = \lambda$ and proceeding to limit $\gamma \rightarrow \infty$, we get

$$\lim_{\gamma \rightarrow \infty} P(X = x) = \frac{n\Gamma(n+x)}{\Gamma(n+1)\Gamma(x+1)} \frac{\lambda^x}{(1+\lambda)^{n+x}} \frac{1}{1-(1+\lambda)^{-n}}$$

Taking $\frac{\lambda}{1+\lambda} = t$ in the equation above, we get

$$P(X = x) = \frac{n\Gamma(n+x)}{\Gamma(n+1)\Gamma(x+1)} \frac{t^x(1-t)^n}{1-(1-t)^n}$$

Proceeding to the limit $n \rightarrow 0$, the equation above reduces to the logarithmic series distribution.

5. ESTIMATION

5.1. MOMENT METHOD:

Let m'_1, m'_2, m'_3 be the sample moments about origin and μ'_1, μ'_2, μ'_3 be the population moments about origin of the proposed model. The method of moments consists in comparing the sample moments with the population moments of the proposed model, that is,

$$\bar{x} = \frac{n\alpha}{(\gamma-1)} \quad (5.1)$$

$$s^2 + \bar{x}^2 = \frac{n\alpha}{(\gamma-1)} + \frac{n(n+1)\alpha(\alpha+1)}{(\gamma-1)(\gamma-2)}, \quad \text{sample variance} = s^2 = m'_2 - \bar{x}^2 \quad (5.2)$$

Using (5.1) in (5.2), we get

$$\frac{l_1(\gamma-2) - \bar{x}^2(\gamma-1)}{\bar{x}} = n + \alpha + 1, \quad l_1 = s^2 + \bar{x}^2 - \bar{x} \quad (5.3)$$

By comparing the third sample moment with its corresponding population moment, we get

$$m'_3 = \frac{n\alpha}{(\gamma-1)} + \frac{3n(n+1)\alpha(\alpha+1)}{(\gamma-1)(\gamma-2)} + \frac{n(n+1)(n+2)\alpha(\alpha+1)(\alpha+2)}{(\gamma-1)(\gamma-2)(\gamma-3)} \quad (5.4)$$

The equation (5.4) together with (5.1) and (5.2) gives

$$\frac{(\gamma-3)(m'_3 - \bar{x} - 3l_1)}{l_1} - \frac{(\gamma-2)l_1}{\bar{x}} = n + \alpha + 3 \quad (5.5)$$

Eliminating n and α between (5.3) and (5.5), we get the estimate of γ as

$$\hat{\gamma} = \frac{\bar{x}(3m'_3 - 7l_1) - 4l_1^2 + \bar{x}^2(l_1 - 3)}{\bar{x}(m'_3 - 3l_1) - 2l_1^2 + \bar{x}^2(l_1 - 1)}$$

Substituting the value of n from (5.3) into (5.1), we get a quadratic equation in α as

$$\alpha^2 \bar{x} - \alpha[l_1(\hat{\gamma} - 2) - \bar{x}^2(\hat{\gamma} - 1) - \bar{x}] + \bar{x}^2(\hat{\gamma} - 1) = 0$$

Which can be solved for α . After estimating α , the value of n can be obtained either from (5.5) or (5.1).

5.2. USING MEAN AND FIRST THREE CELL FREQUENCIES:

Equating the first three probabilities of the proposed model with their corresponding relative frequencies $\frac{f_0}{N}, \frac{f_1}{N}, \frac{f_2}{N}$, we get

$$\frac{\gamma^{[n]}}{(\alpha + \gamma)^{[n]}} = \frac{f_0}{N} \quad (5.6)$$

$$\frac{n\alpha\gamma^{[n]}}{(\alpha + \gamma)^{[n+1]}} = \frac{f_1}{N} \quad (5.7)$$

$$\frac{n(n+1)\alpha(\alpha+1)\gamma^{[n]}}{2(\alpha + \gamma)^{[n+2]}} = \frac{f_2}{N} \quad (5.8)$$

Dividing (5.6) by (5.7) and then using (5.1) in the resulting equation, we get

$$\gamma = \frac{tf_1 + \bar{x}f_0}{\bar{x}f_0} \quad (5.9)$$

$$\text{Where } \alpha + \gamma + n = t \quad (5.10)$$

Using (5.1) and (5.10) in the equation obtained by dividing (5.8) with (5.7), we get

$$\gamma = \frac{2(t+1)f_2 - (t+1-\bar{x})f_1}{(x-1)f_1} \quad (5.11)$$

Eliminating γ between (5.9) and (5.11), we get the estimate of t as

$$\hat{t} = \frac{2\bar{x}f_0f_2}{f_1^2(\bar{x}-1) - 2\bar{x}f_0f_2 + \bar{x}f_0f_1}$$

The equation (5.9) together with the result obtained above gives the estimate of γ as

$$\hat{\gamma} = \frac{\hat{t}f_1 + 2\bar{x}f_0}{\bar{x}f_0}$$

Substituting the value of n from (5.10) into (5.1), we get a quadratic equation in α as

$$\alpha^2 - \alpha(\hat{t} - \hat{\gamma}) + \bar{x}(\hat{\gamma} - 1) = 0$$

Which can be used to estimate α . The estimate of n can be obtained from (5.10) or (5.1).

5.3. MAXIMUM LIKELIHOOD METHOD:

The log likelihood function of the proposed model is given by

$$\log L = N[\log(\gamma)^{[n]} - \log(\alpha + \gamma)^{[n]}] + \sum f_x \log(n)^{[x]} + \sum f_x \log(\alpha)^{[x]} - \sum f_x \log(\alpha + \gamma + n)^{[x]} - \sum f_x \log(x!)$$

Where f_x is the observed frequency for the variate value x and $N = \sum f_x$.

The method of maximum likelihood method of estimation gives the three likelihood equations for three unknown parameters as

$$\begin{aligned} 1) \quad \frac{\partial \log L}{\partial \alpha} = 0 &= \sum f_x \sum_{k=0}^{n+x-1} \frac{1}{\alpha + \gamma + k} - \sum f_x \sum_{k=0}^{x-1} \frac{1}{\alpha + k} \\ 2) \quad \frac{\partial \log L}{\partial \gamma} = 0 &= N \sum_{k=0}^{n-1} \frac{1}{\gamma + k} - \sum f_x \sum_{k=0}^{n+x-1} \frac{1}{\alpha + \gamma + k} \\ 3) \quad \frac{\partial \log L}{\partial n} = 0 &= N \frac{\partial \log(\gamma)^{[n]}}{\partial n} - N \frac{\partial \log(\alpha + \gamma)^{[n]}}{\partial n} \\ &+ \sum f_x \frac{\partial \log(n)^{[x]}}{\partial n} - \sum f_x \frac{\partial \log(\alpha + \gamma + n)^{[x]}}{\partial n} \end{aligned}$$

Converting the likelihood equation above in terms of gamma functions, we get

$$\begin{aligned} N \frac{\partial \log \Gamma(\gamma + n)}{\partial n} - N \frac{\partial \log \Gamma(\alpha + \gamma + n)}{\partial n} + \sum f_x \frac{\partial \log \Gamma(n + x)}{\partial n} - \sum f_x \frac{\partial \log \Gamma(n)}{\partial n} \\ + \sum f_x \frac{\partial \log \Gamma(\alpha + \gamma + n + x)}{\partial n} - \sum f_x \frac{\partial \log \Gamma(\alpha + \gamma + n)}{\partial n} = 0 \end{aligned} \tag{5.12}$$

The differentiation of the equation above is not straight forward and can be solved with the help of the following recurrence relation; (see pages 6-8 Johnson, Kotz and Kemp [16] for details)

$$\psi(x+n) = \psi(x) + \sum_{j=1}^n (x+j-1)^{-1}, \quad n=1,2,\dots \tag{5.13}$$

Where $\psi(x) = \frac{d}{dx} \{\log \Gamma(x)\} = \frac{\Gamma'(x)}{\Gamma(x)}$ is called digamma function. A good approximation for $\psi(x)$ is

$$\psi(x) = \log \Gamma(x) - \frac{1}{2x}$$

The relation (5.13) together with the above result gives

$$\psi(x+n) = \log \Gamma(x) - \frac{1}{2x} + \sum_{j=1}^n (x+j-1)^{-1}, \quad n=1,2,\dots \tag{5.14}$$

By an application of (5.14), the equation (5.12) gives the third likelihood equation as

$$\begin{aligned} N \left[\log \Gamma(\gamma + n) + \sum_{k=1}^{\gamma} (n+k-1)^{-1} \right] - 2N \left[\log \Gamma(\alpha + \gamma + n) + \sum_{k=1}^{\alpha+\gamma} (n+k-1)^{-1} \right] \\ + \sum f_x \left[\log \Gamma(n+x) + \sum_{k=1}^x (n+k-1)^{-1} \right] - N \log \Gamma(n) \end{aligned}$$

$$+ \sum f_x \left[\log \Gamma(n + \alpha + \gamma + x) + \sum_{k=1}^{\alpha + \gamma + x} (n + k - 1)^{-1} \right] = 0$$

The three likelihood equations are not simple to provide direct solution, however, different iterative procedures such as Fisher’s scoring method, Newton-Rampson method etc. can be employed to solve these equations. We may solve the following system of equations

$$(\hat{\theta} - \theta_0) \left[\frac{\partial^2 \log L}{\partial \theta^2} \right]_{\theta_0} = \left[\frac{\partial \log L}{\partial \theta} \right]_{\theta_0}$$

Where $\hat{\theta} = (n, \alpha, \gamma)$ is a parameter vector, is the ML estimate of θ and θ_0 is the trial value of θ which may be first obtained by equating the theoretical frequencies with the observed frequencies.

6. CHI-SQUARE FITTING

In this section, we present three data sets to examine the fitting of the proposed model and comparing that with the negative binomial distribution and generalized negative binomial distribution defined by Jain and Consul [12].

As mentioned in the previous section, the maximum likelihood equations are not straightforward to provide the maximum likelihood (ML) estimates of the parameters of the proposed model and thus need some iterative procedure such as Fisher’s scoring method, Newton-Rampson method etc. for their solution. The R-soft wear provides one among such solutions. The parameters have been estimated with the help of a computer program in R-soft wear and has been shown in the bottom of the table.

TABLE.1. Absenteeism among shift-workers in steel industry; data of Arbous and Sichel [5]

Count	Observed frequency	Expected frequencies		
		NBD	Jain and Consul’s[12] GNBD	Proposed model NPED
0	7	12.02	10.51	9.53
1	16	16.16	17.45	15.93
2	23	17.77	20.38	19.06
3	20	18.08	20.80	19.92
4	23	17.65	19.88	19.41
5	24	16.80	18.34	18.17
6	12	15.72	16.56	16.59
7	13	14.52	14.78	14.90
8	09	13.28	13.08	13.25
9	09	12.06	11.53	11.71
10	08	10.89	10.13	10.30
11	10	09.78	08.89	9.04
12	08	08.75	07.79	7.92
13	07	07.80	16.83	6.94

14	02	06.93	05.99	6.08
15	12	06.14	05.26	5.33
16	03	05.43	04.61	4.68
17	05	04.79	04.05	4.12
18	04	04.22	03.56	3.63
19	02	03.17	03.14	3.20
20	02	03.23	02.76	2.83
21	05	02.86	02.43	2.50
22	05	02.50	02.15	2.22
23	02	02.91	01.90	1.97
24	01	01.91	01.68	1.75
25-48	16	12.77	13.50	17.02
TOTAL	248	248	248	248
ML Estimate		$p=0.854$ $n=1.576$	$p=0.0001077$ $\beta=5978.528$ $n=29337.083$	$n=14.962954$ $\alpha=2.492821$ $\gamma=4.852530$
χ^2 d. f		14.92 17	27.79 16	10.20 16

TABLE 2. The data has been taken from Beall-Rescia [4], Table VII

Count	Observed frequency	Expected frequencies		
		NBD	Jain and Consul's[12] GNBD	Proposed model NPED
0	206	198.85	197.64	196.16
1	143	166.90	168.45	170.06
2	128	124.33	125.12	126.35
3	107	88.70	88.73	89.08
4	71	61.89	61.56	61.32
5	36	42.60	42.19	41.71
6	32	29.05	28.71	28.21
7	17	19.68	19.45	19.04
8	4	13.27	13.13	12.84
9	7	8.91	8.85	8.67
10	7	5.96	5.96	5.87
11	2	3.98	4.00	3.98
12	3	2.65	2.69	2.71
13	3	1.76	1.81	1.85
14	1	1.17	1.21	1.27
15	1	0.78	0.82	0.87
16	1	0.51	0.55	0.60
17	2	0.34	0.37	0.42
18	1	0.22	0.25	0.29

19	0	0.45	0.51	0.70
TOTAL	772	772	772	772
ML Estimate		n=1.2903314 p=0.6504948	n=1.4392633 β =1.0347577 p=0.6119716	n= 45.146373 α =1.429784 γ =27.880158
χ^2 d. f		18.27 9	18.18 8	17.68 8

TABLE 3. Accidents to 647 women working on H.E. Shells during 5 weeks

Count	Observed frequency	Expected frequencies		
		NBD	Jain and Consul's[12] GNBD	Proposed model NPED
0	447	445.89	445.17	445.58
1	132	134.90	136.76	135.88
2	42	44.59	43.09	43.41
3	21	14.09	14.27	14.40
4	3	4.96	4.93	4.94
5	2	2.57	2.78	2.79
TOTAL		647	647	647
ML Estimate		n= 0.86512 p=0.34969	n= 8.26367 β =4.84301 p=0.04424	n=17.78681 α = 0.98208 γ =38.51329
χ^2 d. f		4.46 2	4.32 1	4.15 1

From all the tables it is clear that the proposed model gives a very close fit as compared to other distributions. Thus, the proposed model provides a better alternative to explain the data than the compared distributions.

ACKNOWLEDGMENTS

The authors are thankful to the referees for useful suggestions which improved the quality and presentation of this paper.

REFERENCES

- [1] F. Eggenberger and G. Polya, *ZAMM: Z. Angew. Math. Mech*, Uber die Statistik verketteter vorgange, 1(1923), 279-289.
- [2] J. Neyman, *AMS: Ann. Math. Statist.*, On a new class of contagious distribution applicable in entomology and bacteriology, 10 (1939), 35-57.
- [3] W. Feller, *AMS: Ann. Math. Statist.*, On a general class of contagious distributions, 14 (1943), 389-400.
- [4] G. Beall and R. R. Rescia, *B: Biometrics*, A generalization of Neyman's contagious distributions, 9(1953), 354-57.

- [5] A. G. Arbous and H. S. Sichel, *B: Biometrika*, New techniques for the analysis of absenteeism data, 41(1954), 77.
- [6] C. D. Kemp and A. W. Kemp., JRSS: *Journal of the Royal Statistical Society, Series B*, Generalized hypergeometric distributions, 18, (1956 a), 202-211.
- [7] J. Gurland, *B: Biometrics*, A generalized class of contagious distribution, 14 (1957), 229-249.
- [8] S. K. Bhattacharya, *CSAB: Calcutta Statist. Assoc. Bull.*, Confluent hypergeometric distributions of discrete and continuous type with applications to accident proneness, 15(1966), 20-31.
- [9] J. O. Irwin, JRSS: *Journal of the Royal Statistical Society, Series A*, The generalized Waring distribution applied to accident theory, 131(1968), 205-225.
- [10] N. L. Johnson and S. Kotz, *Univariate Discrete Distribution*, New York: John Wiley and Sons, (1969).
- [11] P. Holgate, *B: Biometrika*, The modality of compound Poisson distribution, 57(1970), 665-667.
- [12] G. C. Jain and P. C. Consul, *SJAM: SIAM J. Appl. Math*, A Generalized Negative Binomial Distribution, 21(4) (1971), 501-513.
- [13] S. Engen, *B: Biometrika*, On species frequency models, 61(1974), 263-270.
- [14] J. O. Irwin, JRSS: *Journal of the Royal Statistical Society, Series A*, The generalized Waring distribution Part II, 138(1975 b), 204-227.
- [15] S. Engen, *Stochastic Abundance Model*. London: Chapman and Hall, (1978).
- [16] N.L. Johnson, S. Kotz, and A. W. Kemp, *Univariate Discrete distributions*, New York: John Wiley and Sons, (1993).

MULTIGRID CONVERGENCE THEORY FOR FINITE ELEMENT/FINITE VOLUME METHOD FOR ELLIPTIC PROBLEMS:A SURVEY

DO Y. KWAK¹

¹DEPARTMENT OF MATHEMATICAL SCIENCES, KAIST, DAEJEON, SOUTH KOREA
E-mail address: kdy@kaist.ac.kr

ABSTRACT. Multigrid methods finite element/finite volume methods and their convergence properties are reviewed in a general setting. Some early theoretical results in simple finite element methods in variational setting method are given and extension to nonnested-noninherited forms are presented. Finally, the parallel theory for nonconforming element[13] and for cell centered finite difference methods [15, 23] are discussed.

1. INTRODUCTION

In this survey article, we discuss various multigrid methods finite element/finite volume methods and their convergence properties. Multigrid methods have been very active area of research since it was introduced in 1960's[19]. It is one of the most efficient algorithm for solving system of linear equations; especially for elliptic problems. In this article, we review the methods in a more general setting. Some early theoretical results in simple finite difference method are given in [2, 18, 20] and for a finite element setting Bank [1] has given a nice proof and Brandt gave an extensive experiment including nonlinear problem, eigenvalue problem, non elliptic problems and/or applications such as image processing. For other cases, we refer to [11, 4, 16, 17, 21, 23, 24]. An excellent theory for conforming finite element method is arranged in the series of paper by Bramble et. al.[5, 6, 7, 9, 8, 3]. Finally, the parallel theory for nonconforming elements[13] and for cell centered finite difference methods [15, 23] are discussed.

2. MULTIGRID ALGORITHM

We consider the following problem:

$$\begin{aligned} Lu &= f & \text{in } \Omega \\ u &= 0 & \text{on } \partial\Omega, \end{aligned} \tag{1}$$

where Ω is a polygonal domain and L is a uniformly elliptic partial differential operator given by $Lu = -\nabla \cdot K\nabla u$ and $f \in L^2(\Omega)$. To discretize it we assume we have a sequence of finite

2000 *Mathematics Subject Classification.* 65N15, 65N30, 35J60.

Key words and phrases. Multigrid, finite element method, finite volume method.

The author's work was supported in part by KAIST.

dimensional space $M_k, k = 1, \dots, J$ with the equipped inner product $(\cdot, \cdot)_k$ which may or may not be a subspace of $H_0^1(\Omega)$. As a result, we obtain the following algebraic equation:

$$A_k x_k = f_k. \quad (2)$$

We introduce a discrete Galerkin form

$$(A_k x_k, v)_k = (f_k, v)_k, \quad v \in M_k. \quad (3)$$

To accomplish the philosophy of multigrid idea, we also assume prolongation operator $I_k : M_{k-1} \rightarrow M_k$ is given and its adjoint P_{k-1}^0 is defined by

$$(P_{k-1}^0 x_k, \phi)_{k-1} = (x_k, I_k \phi)_k, \quad x_k \in M_k, \phi \in M_{k-1}. \quad (4)$$

For each level, we need to relax the residual by a smoother $R_k : M_k \rightarrow M_k$, for $k = 2, \dots, J$. We take $R_1 = A_1^{-1}$ and we only consider the case R_k is symmetric; for nonsymmetric case, see [7].

MULTIGRID ALGORITHM.

Let y_k be an approximation to the solution of $A_k x_k = f_k$, we first define $B_1(y_1, f_1) = A_1^{-1} f_1$. For $k > 1$ define $B_k(y_k, f_k)$ recursively

- (1) Set $x_k^0 = 0, q^0 = 0$.
- (2) $x_k^\ell = x_k^{\ell-1} + R_k(f_k - A_k x_k^{\ell-1}), \quad \ell = 1, 2, \dots, m$.
- (3) $x_k^{m+1} = x_k^m + I_k q^p, \quad p = 1, 2$ where
 $q^i = q^{i-1} + B_{k-1}(0, P_{k-1}^0[f_k - A_k x_k^m - A_k q^{i-1}])$.
- (4) $x_k^\ell = x_k^{\ell-1} + R_k(f_k - A_k x_k^{\ell-1}), \quad \ell = m+2, \dots, 2m+1$
- (5) Set $B_k(y_k, f_k) = x_k^{2m+1}$.

When $p = 1$ it is a V -cycle, and when $p = 2$ it is a W -cycle. Then it is easy to see that

$$q^p = (I - I_k(I - B_{k-1}A_{k-1})^p)A_{k-1}^{-1}P_{k-1}^0A_k(x_k - x_k^m). \quad (5)$$

We shall derive a recurrence relation for $I - B_k A_k$. Let $K_k = I - R_k A_k$. First letting $P_{k-1} = A_{k-1}^{-1}P_{k-1}^0A_k$, we have

$$\begin{aligned} x_k - x_k^{m+1} &= x_k - x_k^m - q^p \\ &= (I - I_k B_{k-1} A_{k-1})^p A_{k-1}^{-1} P_{k-1}^0 A_k (x_k - x_k^m) \\ &= (I - I_k B_{k-1} A_{k-1})^p P_{k-1} K_k^m x_k \end{aligned}$$

on M_k . Then

$$\begin{aligned} x_k - x_k^{2m+1} &= K_k^m (x_k - x_k^{m+1}) = K_k^m (I - I_k B_{k-1} A_{k-1})^p P_{k-1} K_k^m x_k \\ &= K_k^m [I - I_k P_{k-1} + I_k (I - B_{k-1} A_{k-1})^p] P_{k-1} K_k^m x_k \end{aligned}$$

so that

$$I - B_k A_k = K_k^m [I - I_k P_{k-1} + I_k (I - B_{k-1} A_{k-1})^p] P_{k-1} K_k^m.$$

3. CONVERGENCE THEORY FOR THE CONFORMING FINITE ELEMENT METHOD

In this section, we assume the space M_k consists of continuous functions which are piecewise linear on each element (triangles or rectangles) where each element in τ_k is divided by connecting mid point of its side to produce elements in τ_{k+1} so that all the spaces M_k are nested. Hence the transfer operator I_k is the natural injection operator. Also, let

$$a(u, v) = \int_{\Omega} K \nabla u \cdot \nabla v \, dx = (f, v), \quad v \in H_0^1(\Omega)$$

be the usual variational formulation. Obviously the following is satisfied:

$$\mathbf{(A.0):} \quad (A_k I_k u, I_k u) = (A_{k-1} u, u), \quad u \in M_{k-1}.$$

The convergence proof in this case is based on the following assumptions.

$\mathbf{(A.1):}$ There exists some $0 < \alpha \leq 1$ such that

$$A((I - P_{k-1})u, u) \leq C_{\alpha}^2 \left(\frac{\|A_k u\|_k^2}{\lambda_k} \right)^{\alpha} A(u, u)^{1-\alpha} \text{ for all } u \in M_k,$$

$\mathbf{(A.2):}$ $\frac{\|u\|_k^2}{\lambda_k} \leq C_R (R_k u, u)_k$, for all $u \in M_k$,

where λ_k is the largest eigenvalue of A_k . The following result is in [5].

Theorem 3.1. *Let $p = 1$ and $m = \text{constant}$. Then*

$$A((I - B_k^s A_k)u, u) \leq \delta_k A(u, u),$$

$$\text{with } \delta_k = \frac{C_{\alpha, k}}{C_{\alpha, k} + m^{\alpha}}.$$

This result can be improved if we use a product form of multigrid algorithm: We first consider presmoothing only with $p = 1$. Then

$$(I - B_k^n A_k) = [(I - P_{k-1}) + (I - B_{k-1}^n A_{k-1})P_{k-1}]K_k^m \text{ on } M_k.$$

For the improvement of result, we derive a recursive relation: Let $T_k = (I - K_k^m)P_k$. Then we have

$$I - B_k^n A_k P_k = [I - B_{k-1}^n A_{k-1} P_{k-1}](I - T_k).$$

The corresponding relation for symmetric smoother (with postsmoothing also) is given by

$$(I - B_k^s A_k) = K_k^m [(I - P_{k-1}) + (I - B_{k-1}^n A_{k-1})^p P_{k-1}]K_k^m \text{ on } M_k$$

so that

$$(I - B_k^s A_k) = (I - B_k^n A_k)^*(I - B_k^n A_k). \quad (6)$$

Then we have the following result.

Theorem 3.2. [8] *The V-cycle converges with $\delta_J = 1 - \frac{1}{C_J}$, with no regularity assumptions other than H^1 .*

Proof. For the proof, we use

$$(I - B_J^n A_J) = (I - T_J) \cdots (I - T_0)$$

$$E_k = (I - T_k) E_{k-1}, \quad E_J = (I - B_J A_J)$$

and the following assumptions:

(A.3): There exists $Q_k : M_J \rightarrow M_k$ such that

$$\|(Q_k - Q_{k-1}u)\|_k^2 \leq C_1 \lambda_k^{-1} A((u, u)), \quad k = 1, \dots, J$$

(A.4): $A(Q_k u, Q_k u) \leq C_2 A(u, u)_k, \quad k = 1, \dots, J.$

□

With a more ingenious analysis shows the convergence can be further improved where the convergence rate is independent of the number of levels:

Theorem 3.3. [3] *Let $u \in H^{1+\alpha}(\Omega)$ with $\alpha > 0$. Then the V-cycle with one smoothing converges with $\delta_J = 1 - \frac{1}{C}$.*

4. NONNESTED SPACE OR NONINHERITED FORMS.

In the case of nonconforming finite element where $M_k \not\subset H_0^1(\Omega)$ or finite difference methods we usually do not have ‘variational equality’, i.e, the following

$$A_k(I_k u, I_k v) = A_{k-1}(u, v), \quad u, v \in M_{k-1}$$

do not hold. This also happens when treating curved boundaries with straight edged element. The recurrence relation becomes

$$I - B_k A_k = K_k^m [(I - I_k P_{k-1}) + I_k (I - B_{k-1} A_{k-1})^p P_{k-1}] K_k^m.$$

Now the condition (A.0), (A.1) have to be replaced by

(B.0): $A_k(I_k u, I_k u) \leq C_* A_{k-1}(u, u), \quad u \in M_{k-1}.$

(B.1): $A((I - I_k P_{k-1})u, u) \leq C_\alpha^2 \left(\frac{\|A_k u\|_k^2}{\lambda_k} \right)^\alpha A(u, u)^{1-\alpha}, \quad u \in M_k.$

The typical finite difference method applied to the Laplace operator with linear or bilinear element give rise to a bilinear form which satisfy (B.0) with $C_* = 1$. In this case, we have V-cycle result.

Theorem 4.1. *If $C_* \leq 1$, then the multigrid V-cycle algorithm converges with $\delta_k = \frac{C(\alpha, k)}{C(\alpha, k) + m^\alpha}$.*

W-cycle result

For general elliptic problem, however, FDM does not yields $C_* = 1$. In most cases C_* is strictly greater than 1. The same is true for nonconforming finite element. For these cases, we need to use $p = 2$ (W-cycle).

Theorem 4.2. *Assume (B.0) and (B.1) hold and $p = 2$. Then if m is sufficiently large we have $\delta_k = \frac{C}{C_{+m(k)^\alpha}}$. If in addition,*

$$A_k(I_k u, I_k u) \leq 2A_{k-1}(u, u)$$

holds then the W -cycle algorithm converge with the same δ with $m = 1$.

Nonconforming Elements. For various nonconforming elements even $C_* \leq 2$ do not holds generally. Thus, we investigate specific case in detail. Let $V_k, k = 1, \dots, J$ denote the nonconforming finite element spaces. The variational for usually satisfies

$$a_k(I_k v, I_k v) \leq C_* a_{k-1}(v, v), \quad \forall v \in V_{k-1}, \quad (7)$$

for some constant C . For triangular nonconforming element, it is known that $C_* > 2$. Hence W -cycle with large m only converges. However, for rectangular case, we have convergence for $m = 1$.

Theorem 4.3. [12] *The following estimate holds*

$$a_k(I_k v, I_k v) \leq 2a_{k-1}(v, v), \quad \forall v \in V_{k-1}$$

and thus the W -cycle converges with $m = 1$.

Cell Centered Method. Similar framework can be used to prove convergence of multigrid applied to the cell centered finite difference. First consider the following model problem:

$$-\nabla \cdot p \nabla \tilde{u} = f \text{ in } \Omega, \quad (8)$$

$$\tilde{u} = 0 \text{ on } \partial\Omega. \quad (9)$$

Integrating by parts on each cell, we obtain

$$-\int_{\partial E_{ij}^k} p \frac{\partial \tilde{u}}{\partial n} ds = \int_{E_{ij}^k} f dx \quad (10)$$

for $i, j = 1, \dots, n$. We approximate (10) by

$$p \frac{\partial \tilde{u}}{\partial n} \approx p_{i,j+1/2} \frac{u_{i,j+1} - u_{i,j}}{h}$$

to obtain a system of linear equation of the form

$$\bar{A}_k \bar{u} = \bar{f}, \quad (11)$$

whose multgrid method was first considered in [4]. However, they used a natural operator whose V -convergence is very slow. In fact, they show

$$A(I_k^n v, I_k^n v) = 2A(v, v)$$

and, hence, only W -cycle works. However, we can design a new prolongation operator and improve the energy norm. Define a weighted prolongation operator \mathcal{I}_{k-1}^k as follows:

$$\begin{aligned} (\mathcal{I}_{k-1}^k v)_{I,J} &= \frac{1}{4}(2v_{i,j} + v_{i,j+1} + v_{i+1,j}), \\ (\mathcal{I}_{k-1}^k v)_{I-1,J} &= \frac{1}{4}(2v_{i,j} + v_{i,j+1} + v_{i-1,j}), \\ (\mathcal{I}_{k-1}^k v)_{I,J-1} &= \frac{1}{4}(2v_{i,j} + v_{i+1,j} + v_{i,j-1}), \\ (\mathcal{I}_{k-1}^k v)_{I-1,J-1} &= \frac{1}{4}(2v_{i,j} + v_{i-1,j} + v_{i,j-1}). \end{aligned}$$

We can now prove the following crucial energy norm estimate.

Proposition 1. *For all $v \in V_{k-1}$, we have*

$$A_k(\mathcal{I}_{k-1}^k v, \mathcal{I}_{k-1}^k v) \leq C(p)A_{k-1}(v, v), \quad (12)$$

where $C(p) = 1$ if the coefficient p is constant and $C(p) = 1 + O(h_k)$ if p is a general Lipschitz continuous function.

Lemma 4.1. *Regularity and approximation property holds for $\alpha = \frac{1}{2}$.*

With these preparations, we prove the following: [15]

Theorem 4.4. *With $E_k = I - B_k A_k$, we have the following: If p is constant, then V -cycle converges and we have*

$$0 \leq A_k(E_k u, u) \leq \delta_k A_k(u, u) \quad \forall u \in V_k \quad (13)$$

where $\delta_k = \frac{Ck}{Ck + \sqrt{m}}$.

Triangular case. Similar idea can be used to design the prolongation operator for triangular element: Observe

$$A_{k-1}(v, v) = \theta \sum_{i \neq j} (v_i - v_j)^2, \quad (14)$$

where the sum is taken for all pairs of adjacent triangles i and j . Let $u = I_k^t v$. Then

$$A_k(u, u) = \theta \sum_{I \neq J} (u_I - u_J)^2. \quad (15)$$

If we define I_k to be certain average, it can be shown that

$$A_k(I_k^t v, I_k^t v) = 2\theta \sum_{i \neq j} (v_i - v_j)^2.$$

We find that the result does not change even if we change the weight as long as the sum of coefficients is 1. For details, see [25].

5. SOME V -CYCLE THEORY FOR THE NONCONFORMING CASE

The V -cycle convergence with the nonconforming element has been computationally observed but the proof seemed to be an open problem for a while. It was partially resolved by Chen and later Chen and Kwak[13] with no regularity assumption.

The obstacle of the convergence proof is that we have $C_*' > 1$ in (7). The idea is to judiciously introduce a new bilinear form so that (7) holds with $C_*' = 1$.

Given finite element spaces

$$V_0, V_1, \dots, V_J,$$

and coarse-to-fine grid operators $I_k : V_{k-1} \rightarrow V_k$ for $k = 1, \dots, J$.

Assume a quadratic forms $a_J(\cdot, \cdot)$ on $V_J \times V_J$ and define the iterates $H_k^J = I_J \cdots I_{k+1} : V_k \rightarrow V_J$, of I_k and their adjoint Λ_J^k by

$$H_k^J = I_J \cdots I_{k+1} : V_k \rightarrow V_J, \quad k = 0, \dots, J-1, \quad (16)$$

$$a_k(\Lambda_J^k v, w) = a_J(v, H_k^J w), \quad k = 0, \dots, J. \quad (17)$$

Using a_J form, we define new quadratic forms $b_k(\cdot, \cdot)$ on $V_k \times V_k$ by

$$b_k(v, w) = a_J(H_k^J v, H_k^J w), \quad \forall v, w \in V_k, k = 0, \dots, J.$$

If we let

$$(A_k v, w)_k = b_k(v, w), \quad \forall w \in V_k, k = 0, \dots, J,$$

then we have

$$\begin{aligned} (A_k I_k v, I_k v) &= b_k(I_k v, I_k v) \\ &= a_J(H_k^J I_k v, H_k^J I_k v) \\ &= a_J(H_{k-1}^J v, H_{k-1}^J v) \\ &= b_{k-1}(v, v) \\ &= (A_{k-1} v, v). \end{aligned}$$

Theorem 5.1 (Chen 96). *With certain regularity V -cycle MG for nonconforming element converges with $\delta_k = \frac{C_\alpha}{C_\alpha + m\alpha}$.*

Sketch of proof: We define the operators $P_{k-1} : V_k \rightarrow V_{k-1}$ and $P_{k-1}^0 : V_k \rightarrow V_{k-1}$ by

$$b_{k-1}(P_{k-1} v, w) = b_k(v, I_k w), \quad \forall w \in V_{k-1}, k = 1, \dots, J,$$

and

$$(P_{k-1}^0 v, w)_{k-1} = (v, I_k w)_k, \quad \forall w \in V_{k-1}, k = 1, \dots, J.$$

As usual, we need "approx. Regularity" for $b_k(\cdot, \cdot)$ form:

$$|(A_k(I - I_k P_{k-1})v, v)_k| \leq Ch_k \|A_k v\|_k \|v\|_{1,k} \quad (18)$$

For a proof, we need properties of $a(\bar{P}_{k-1} v, w) = b_k(v, I_k w)$ and the following lemma.

Lemma 5.1.

$$\begin{aligned} C_1 \|v\|_{\mathcal{E},k} &\leq \|v\|_{1,k} \leq C_2 \|v\|_{\mathcal{E},k} \\ \|\Lambda_J^k v\|_{\mathcal{E},k} &\leq C \|v\|_{\mathcal{E},k} \\ \|v - H_k^H v\| &\leq Ch_k \|v\|_{\mathcal{E},k}. \end{aligned}$$

This result works for smooth problem. The case of rough coefficients with only one smoothing is provided below.

Theorem 5.2 (Chen, Kwak 96). *(No regularity) For problems with large jump coefficients, V-cycle multigrid with one smoothing converges with $\delta = 1 - \frac{1}{CJ}$.*

For the proof, we have to resort to the product algorithm introduced in section 3. For that purpose we need several new operators. First define $\Pi_J^k : V_J \rightarrow V_k$ by

$$b_k(\Pi_J^k v, w) = b_J(v, H_k^J w), v \in V_J, w \in V_k.$$

Then the following properties hold.

Lemma 5.2. *It holds that*

$$P_{k-1}^0 A_k = A_{k-1} P_{k-1}, \quad (19)$$

$$P_{k-1} I_k = I \quad \text{on } V_{k-1}. \quad (20)$$

Lemma 5.3. *We have*

$$\begin{aligned} \Pi_J^{k-1} &= P_{k-1} \Pi_J^k. \\ \Pi_J^k &= P_k P_{k+1} \dots P_{J-1}; \\ \Pi_J^k H_k^J &= I \quad \text{on } V_k. \end{aligned}$$

With the iterated transfer operators H_k^J and Π_J^k let $S^k = H_k^J (I - J_k^{m(k)}) \Pi_J^k$. Then we have

$$I - H_k^J B_k^n A_k \Pi_J^k = (I - H_{k-1}^J B_{k-1}^n A_{k-1} \Pi_J^{k-1}) (I - S^k).$$

Therefore, by induction we obtain

$$I - B_J^n A_J = (I - S^0) \dots (I - S^J).$$

Denote the norm induced by the a_k form by $\|\cdot\|_{\mathcal{E},k}$. It has been shown [12] that the norms $\|\cdot\|_{\mathcal{E},k}$ and $\|\cdot\|_{1,k}$ are equivalent:

$$C_5 \|v\|_{\mathcal{E},k} \leq \|v\|_{1,k} \leq C_6 \|v\|_{\mathcal{E},k}, \quad \forall v \in V_k, k = 0, \dots, J, \quad (21)$$

with C_5 and C_6 independent of k . Assume the existence of $Q_J^k : V_J \rightarrow V_k$ such that $Q_J^k H_k^J v = v$ and

$$\|Q_J^k v\|_{1,k} \leq C \|v\|_{1,J}.$$

Later, Q_J^k will be given as a product of certain projection operators T_k .

Lemma 5.4. *There exist constants C independent of k such that*

$$\|Q_J^k v\|_{\mathcal{E},k} \leq C \|v\|_{\mathcal{E},J}, \quad \forall v \in V_J, \quad (22)$$

$$\|T_{k-1} v\|_{\mathcal{E},k-1} \leq C \|v\|_{\mathcal{E},k}, \quad \forall v \in V_k. \quad (23)$$

Now we apply product form of multigrid algorithm: For that purpose, we need to verify two conditions:

$$\|(Q_J^k - I_k Q_J^{k-1})v\|_k^2 \leq C_1 \lambda_k^{-1} b_J(v, v), \quad k = 1, \dots, J, \quad (24)$$

$$b_k(Q_J^k v, Q_J^k v) \leq C_2 b_J(v, v), \quad k = 0, \dots, J-1. \quad (25)$$

(25) follows from (22); To prove (24), observe that

$$Q_J^k - I_k Q_J^{k-1} = (I - I_k T_{k-1}) Q_J^k.$$

By approximation properties of I_k and T_{k-1}

$$\begin{aligned} \|(Q_J^k - I_k Q_J^{k-1})v\|_k &= \|(I - I_k T_{k-1})Q_J^k v\| \\ &\leq C (\|(I - T_{k-1})Q_J^k v\| + \|(I - I_k)T_{k-1}Q_J^k v\|) \\ &\leq Ch_k \|v\|_{\mathcal{E},J}, \quad v \in V_J. \end{aligned}$$

This, together with (21), implies the assumption (24). With these preparations, we can prove the result using similar frame in section 3.

Remark 5.1. *Similar result holds for nonsymmetric problem[14].*

Properties of transfer operator. First we consider triangular nonconforming element. We consider two sets of intergrid transfer operators $I_k : V_{k-1} \rightarrow V_k$ and $T_{k-1} : V_k \rightarrow V_{k-1}$ as follows.

$$(I_k v)(q) = \begin{cases} 0 & \text{if } q \in \Gamma, \\ v(q) & \text{if } q \notin \partial E \text{ for any } E \in \mathcal{E}_{k-1}, \\ \frac{1}{2} \{v|_{E_1}(q) + v|_{E_2}(q)\} & \text{if } q \in \partial E_1 \cap \partial E_2. \end{cases}$$

If $v \in V_k$ and q is the midpoint of an edge e of a triangle in \mathcal{E}_{k-1} , following [12], we define $T_{k-1} v \in V_{k-1}$ by

$$(T_{k-1} v)(q) = \frac{1}{2}(v(q_1) + v(q_2)),$$

where q_1 and q_2 are the respective midpoints of the edges e_1 and e_2 in \mathcal{E}_k , which form the edge e in \mathcal{E}_{k-1} . Note that the definition of T_{k-1} automatically preserves the zero nodal values on boundary edges. Also, it can be seen that

$$T_{k-1} I_k v = v, \quad v \in V_{k-1}, \quad k = 1, \dots, J. \quad (26)$$

As in [12], the iterated intergrid transfer operators

$$H_k^J = I_J \cdots I_{k+1} : V_k \rightarrow V_J, \quad k = 0, \dots, J-1, \quad (27)$$

$$Q_J^k = T_k \cdots T_{J-1} : V_J \rightarrow V_k, \quad k = 0, \dots, J. \quad (28)$$

satisfies

$$Q_J^k H_k^J v = v, \quad v \in V_k.$$

Similar result holds in the case of the Q_1 -nonconforming element.

Remark 5.2. *Finally, one can consider a perturbed problem. Perturbation comes from a variety of sources: Nonsymmetric term, numerical integration, finite difference methods, treating curved boundaries, nonconforming element, etc. These cases can be handled almost universally by introducing certain perturbed bilinear form and estimating the energy norm appropriately and W – cycle can be shown to converge if the unperturbed problem converge. See [22] for details.*

REFERENCES

- [1] R. Bank and T. Dupont, *An optimal order process for solving finite element equations*, Math. Comp., Vol. 36, pp. 35–51, (1981).
- [2] D. Braess and W. Hackbusch, *A new convergence proof for the multigrid method including the V –cycle*, SIAM J. Numer. Anal., Vol. 20, pp. 967–975, (1983).
- [3] J. H. Bramble, *Multigrid methods*, Pitman Research Notes in Mathematics Series 294, Longman Scientific, 1993.
- [4] J. H. Bramble, R. Ewing, J. E. Pasciak, and J. Shen, *The analysis of multigrid algorithms for cell centered finite difference methods*, Adv. Comput. Math., Vol. 5, no 1. pp. 15–29, (1996).
- [5] J. H. Bramble and J. E. Pasciak, *New convergence estimates for multigrid algorithms*, Math. Comp., Vol. 49, pp. 311–329, (1987).
- [6] J.H. Bramble and J.E. Pasciak, *The analysis of smoothers for multigrid algorithms*, Math. Comp. Vol 58, pp. 467-488(1992)
- [7] J.H. Bramble, D.Y.Kwak, and J.E. Pasciak *Uniform convergence of multigrid V -cycle iterations for indefinite and non-symmetric problems*, SIAM. J. Numer. Anal. Vol 31, 6, pp. 1746-1763(1994).
- [8] J.H. Bramble, J.E. Pasciak, J. Wang, and J. Xu, *Convergence estimates for multigrid algorithms without regularity assumptions*, Math. Comp. Vol 57, pp. 23–45 (1991).
- [9] J. H. Bramble, J. E. Pasciak, and J. Xu, *The analysis of multigrid algorithms with non-nested spaces or non-inherited quadratic forms*, Math. Comp., Vol. 56, pp. 1–34, (1991).
- [10] A. Brandt, *Guide to multigrid development*, in Multigrid Methods, Lecture Notes in Mathematics, W. Hackbusch and U. Trottenberg Ed., Vol. 960, pp. 220–312, (1981).
- [11] S. Brenner *An optimal-order multigrid method for P_1 nonconforming finite elements*, Math. Comp. Vol 52, pp. 1–15, (1989).
- [12] Z. Chen and P. Oswald, *Convergence of multigrid methods for nonconforming finite elements*, Technical Report, Southern Methodist University, 1996.
- [13] Z. Chen and Do Y. Kwak, *Convergence of Multigrid Methods for nonconforming finite Element Without Regularity Assumptions*, J. Comp. Appl. Math. (1998)
- [14] Z. Chen and Do Y. Kwak, *V -cycle Galerkin-multigrid methods for nonconforming methods for nonsymmetric and indefinite problems*, Applied Numerical Mathematics, 1998.
- [15] Do Y. Kwak, *V -cycle multigrid for cell-centered finite differences*, SIAM J. of Scientific Computing, V. 1999.
- [16] Z. Chen, Do Y. Kwak and Y. Yon, *The analysis of multigrid algorithms for nonconforming and mixed methods for second order elliptic problems*, SIAM J. of Scientific Computing. Vol 19 (2), 502-515.
- [17] R. E. Ewing and J. Shen, *A multigrid algorithm for the cell-centered finite difference scheme*, The Proceedings of 6-th Copper Mountain Conference on Multigrid Methods, April 1993, NASA Conference Publication 3224.
- [18] W. Hackbusch, *Multi-grid convergence theory*, in Multigrid Methods, Lecture Notes in Mathematics, W. Hackbusch and U. Trottenberg Ed., Vol. 960, pp. 177–219, (1981).

- [19] Fedorenko, R.P. *The speed of convergence of one iterative process*, USSR Comput. Math. and Math. Phys. pp. 1092–1096 (1961).
- [20] W. Hackbusch, *Multigrid Methods and Applications*, Springer Series in Computational Mathematics, Springer-Verlag, (1985).
- [21] M. Khalil and P. Wesseling, *Vertex-centered and cell centered multigrid for interface problems*, J. Comp. Physics, Vol. 98, pp. 1–10(1992).
- [22] Do. Y. Kwak, *A General Multigrid Framework for a Class of Perturbed Problems*, Contemporary Mathematics, V. 295, 2002.
- [23] Do Y. Kwak and J. S. Lee, *Multigrid Algorithm For The Cell Centered Finite Difference Method II: Discontinuous Coefficient Case*, Numerical Methods for PDE. V 20, No 5, 2004, pp742-764.
- [24] Do Y. Kwak and J. S. Lee, *Multigrid analysis For higher order Finite Difference Scheme*, Journal of Numerical Mathematics, V12, No4, pp. 285-296 (2004).
- [25] Do Y. Kwak, H. J. Kwon and S. Lee *Multigrid algorithm for cell centered finite difference on triangular meshes*, Applied Math. and Comp.(105) (1999) 77-85.
- [26] R. Rannacher and S. Turek *Simple nonconforming quadrilateral Stokes element*, Numer. Meth. Partial Diff. Equations Vol 8, pp. 97–111(1992).

A NEW VERSION OF FIRST RETURN TIME TEST OF PSEUDORANDOMNESS

DONG HAN KIM

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF SUWON, HWASEONG 445-743, KOREA.
E-mail address: kimdh@suwon.ac.kr.

ABSTRACT. We present a new version of the first return time test for pseudorandomness. Let R_n be the first return time of initial n -block with overlapping. An algorithm to calculate the probability distribution of the first return time R_n for each starting block is presented and used to test pseudorandom number generators. The standard Z -test for $\log R_n$ is applied to test the pseudorandom number generators.

1. INTRODUCTION

We introduce a new version of testing pseudorandom number generators (PRNGs) based on the first return time of the initial n -block for some fixed length n in a randomly generated binary sequence. The first return time is closely related to entropy, which plays a key role in the information theory. Entropy is defined as the limit of $-\frac{1}{n} \sum_i p_i \log p_i$ as n goes to infinity where the p_i 's are the possibility of each block of length n appears in the source. Entropy measures the amount of randomness and the maximum compression rate.

For each binary sequence $x = (x_1, x_2, \dots)$, define the first return time (recurrence time) by

$$R_n(x) = \min\{j \geq 1 : x_1 \dots x_n = x_{j+1} \dots x_{j+n}\},$$

To study the data compression algorithm like the Lempel-Ziv code[19], Wyner and Ziv showed that $\frac{1}{n} \log R_n(x)$ converges to entropy in measure[16]. Since then there have been many works on the relation between entropy and the first return time (e.g. [7], [8], [11], [18]).

Define the waiting time by

$$W_n(x, y) = W_n(x_1^n, y) = \min\{j \geq 1 : x_1 \dots x_n = y_j \dots y_{j+n-1}\}.$$

Wyner and Ziv showed that $\frac{1}{n} \log W_n(x, y)$ converges to entropy for Markov chains[16]. Generally $\frac{1}{n} \log W_n(x, y)$ converges to the entropy for weakly Bernoulli processes, but the convergence does not hold in general ergodic processes[15].

2000 *Mathematics Subject Classification.* 65C10.

Key words and phrases. The first return time, pseudorandom number generator, probability density, entropy, randomness.

This work was supported by the Korea Research Foundation Grant funded by the Korean Government (MOEHRD, Basic Research Promotion Fund) (KRF-2007-331-C00016).

Maurer [10] presented the nonoverlapping first return time algorithm in testing PRNGs, His algorithm corresponds to the nonoverlapping first return time:

$$R_{(n)}(x) = \min\{j \geq 1 : x_1 \dots x_n = x_{j+1} \dots x_{j+n}\}.$$

Put $v(r) = r \sum_{i=1}^{\infty} (1-r)^{i-1} \log_2 i$, and $w(r) = r \sum_{i=1}^{\infty} (1-r)^{i-1} (\log_2 i)^2$. Since $\Pr(B) = 2^{-n}$, $E[\log R_{(n)}] = v(2^{-n})$ and $E[(\log_2 R_{(n)})^2] = w(2^{-n})$. Then

$$\lim_{r \rightarrow 0^+} [v(r) + \log r] = -\gamma / \ln 2 = -0.832 \dots \equiv C,$$

where $\gamma = \lim_{n \rightarrow \infty} (\sum_{i=1}^n (1/i) - \ln n)$ is Euler's constant. Similarly,

$$\lim_{r \rightarrow 0^+} [w(r) - (\log r)^2 + 2C \log r] = 4.11 \dots \equiv D.$$

So the expectation of $\log R_{(n)}$ is $n + C$ and the variation is $D - C^2$. See [2] and [4] for related results.

An overlapping algorithm of the first return time test of randomness is considered by Choe and the author[3]. They studied the return time which does not allows the overlapping between the initial block and the first recurrent block:

$$R'_n(x) \equiv \min\{j \geq n : x_1 \dots x_n = x_{j+1} \dots x_{j+n}\}.$$

Though the nonoverlapping algorithm is relatively easier to analyze, but overlapping method is more natural and efficient since nonoverlapping algorithm requires n times more sample random digits to be applied. The first return time test using $R'_n(x)$ can be regards as the waiting time test of $W_n(x_1^n, x_{n+1}^\infty)$. In this article we consider the first return time $R_n(x)$ for testing pseudorandomness. Kac's lemma[5] which states that $E[R_n | x_1 \dots x_n = B] = 1 / \Pr(B)$ and the convergence of $\log R_n / n$ to entropy make R_n be more natural to consider than W_n for the randomness test though the methods in this article are not very different from that of using R'_n in [3]. For a related result of testing PRNGs using the first return time on the cyclic group, see [6].

In Section 2 we develop a formula for computing $\Pr(R_n = i | x_1 \dots x_n = B)$ exactly. First, we classify blocks with the same distribution of R_n . Next, we use two sequences $r_k(B)$ and $s_k(B)$ which have the information on the probability distribution of R_n for $x_1 \dots x_n = B$ and find the recurrence relation among them to compute $\Pr(R_n = i | x_1 \dots x_n = B)$ for each i . In Section 3 we apply the standard Z -tests for $\log R_n / n$. For each block of length 14 we compute the expectation and the standard deviation of $\log R_n / n$ and the deviation of the experimental data from the theoretical prediction is used to test PRNG's. Unlike the return time test in [3], we calculate the Z -values for each starting blocks, which enable us to have more sharp test result.

2. THE PROBABILITY DISTRIBUTION OF THE FIRST RETURN TIME

A *block* is a finite sequence of elements of alphabet $\mathcal{A} = \{0, 1\}$ and an *n-block* is a block of length n . For an n -block $B = b_1 b_2 \dots b_n \in \mathcal{A}^n$, we write $B_i^j = b_i b_{i+1} \dots b_j$, $1 \leq i \leq j \leq n$. Throughout the paper $\mathcal{A} = \{0, 1\}$.

TABLE 1. The Expectation of R_n and $\log R_n$

Block	$\bar{\Lambda}(B)$	$\Lambda(B)$	$E[R_n]$	$E[\log R_n]$	$\text{Var}[\log R_n]$
00000000	1,2,3,4,5,6,7	1	256	4.122127	18.37019
00000001	\emptyset	\emptyset	256	7.299403	2.441935
00000010	7	7	256	7.273498	2.589157
00000100	6,7	6,7	256	7.219351	2.905512
00001000	5,6,7	5,6,7	256	7.106875	3.576236
00010001	4	4	256	7.055111	3.986235
00100001	5	5	256	7.183896	3.147559
00100010	4,7	4,7	256	7.031221	4.110117
00100100	3,6,7	3,7	256	6.717126	6.102838
01000001	6	6	256	7.244771	2.763759
01000010	5,7	5,7	256	7.158986	3.283393
01001001	3,6	3	256	6.738698	6.005312
01010101	2,4,6	2	256	6.015615	10.32028

Since the distribution of return time is different from block to block. We classify the blocks to each set of blocks have the same return time distribution.

Definition 2.1. Let B be an n -block.

$$\bar{\Lambda}(B) = \{m : B_{m+1}^n = B_1^{n-m}, 1 \leq m < n\},$$

$$\Lambda(B) = \{m : m \in \bar{\Lambda}(B), i \nmid m \text{ for any } i \in \bar{\Lambda}(B)\}.$$

Table 2 shows $\bar{\Lambda}(B)$ and $\Lambda(B)$ for some 8-blocks. For more of the definition of $\Lambda(B)$, see [3].

Lemma 2.2 ([3], Lemma 3). *Let B be an n -block.*

- (i) *If $\lambda \in \bar{\Lambda}(B)$ and $\lambda < m < n$, then $\lambda \in \bar{\Lambda}(B_1^m)$.*
- (ii) *If $B = B_{m+1}^n B_1^m$ for some $1 \leq m < n$, then there is $\lambda \in \Lambda(B)$ such that λ divides n and m .*
- (iii) *If there is $\lambda \in \Lambda(B)$ with $\lambda \leq n/2$, then each $\lambda' \in \Lambda(B)$, $\lambda' \neq \lambda$, satisfies $\lambda' > n - \lambda$.*

Definition 2.3. Let B be an n -block. For each $k \geq n$ denote $\mathcal{F}(B, k)$ by the set of k -blocks of

$$\mathcal{F}(B, k) = \{C : C_1^n = B, C_{i+1}^{i+n} \neq B \text{ for any } i \geq 1\},$$

and for $k \geq 1$ let $\mathcal{S}(B, k)$ be the set of k -blocks defined by

$$\mathcal{S}(B, k) = \{C : (CB)_1^n = B, (CB)_{i+1}^{i+n} \neq B \text{ for any } i, 1 \leq i < k\}.$$

Clearly for $k \geq n$ we have $\mathcal{S}(B, k) \subset \mathcal{F}(B, k)$. Note that

$$x_1^k \in \mathcal{F}(B, k) \text{ if and only if } x_1^n = B, \text{ and } R_n(x) > k - n, \quad k \geq n \quad (1)$$

$$x_1^k \in \mathcal{S}(B, k) \text{ and } x_{k+1}^{k+n} = B \text{ if and only if } x_1^n = B \text{ and } R_n(x) = k, \quad k \geq 1. \quad (2)$$

The following shows the relation between $\mathcal{F}(B, k)$ and $\mathcal{S}(B, k)$.

Lemma 2.4. *Let B be an n -block. (i) For $k > n$ we have*

$$\mathcal{F}(B, k) \cup \{CB : C \in \mathcal{S}(B, k - n)\} = \{C \in \mathcal{A}^k : C_1^{k-1} \in \mathcal{F}(B, k - 1)\},$$

where the union is disjoint. (ii) For $k \geq n$ we have

$$\mathcal{F}(B, k) \setminus \mathcal{S}(B, k) = \bigcup_{\lambda \in \overline{\Lambda}(B)} \{CB_1^\lambda : C \in \mathcal{S}(B, k - \lambda)\},$$

where the unions are disjoint.

Proof. (i) By the definition of $\mathcal{F}(B, k)$ and $\mathcal{S}(B, k)$, it is clear that

$$\mathcal{F}(B, k) \cup \{CB : C \in \mathcal{S}(B, k - n)\} \subset \{C \in \mathcal{A}^k : C_1^{k-1} \in \mathcal{F}(B, k - 1)\}.$$

Let C be an k -block with $C_1^{k-1} \in \mathcal{F}(B, k - 1)$. Then either $C_{k-n+1}^k = B$ or $C_{k-n+1}^k \neq B$. If $C_{k-n+1}^k = B$, then $C_1^{k-n} \in \mathcal{S}(B, k)$ and $C = C_1^{k-n}B$. When $C_{k-n+1}^k \neq B$, $C \in \mathcal{F}(B, k)$.

(ii) Take a k -block $C \in \mathcal{F}(B, k) \setminus \mathcal{S}(B, k)$. Then for some s with $0 < s < n$ we have $(CB)_{k-s+1}^{k-s+n} = B$ i.e., $s \in \overline{\Lambda}(B)$ and $C_{k-s+1}^k = B_1^s$. If we put λ as the largest number of such s 's, then we have $C_1^{k-\lambda} \in \mathcal{S}(B, k - \lambda)$. Hence we have

$$\mathcal{F}(B, k) \setminus \mathcal{S}(B, k) \subset \bigcup_{\lambda \in \overline{\Lambda}(B)} \{CB_1^\lambda : C \in \mathcal{S}(B, k - \lambda)\}.$$

From the definition of $\mathcal{F}(B, k)$ and $\mathcal{S}(B, k)$, we have

$$\mathcal{F}(B, k) \setminus \mathcal{S}(B, k) = \bigcup_{\lambda \in \overline{\Lambda}(B)} \{CB_1^\lambda : C \in \mathcal{S}(B, k - \lambda)\}.$$

Now we prove the disjointness of the union: Suppose that there exist $\lambda, \lambda' \in \overline{\Lambda}(B)$, $\lambda < \lambda'$ such that $CB_1^\lambda = C'B_1^{\lambda'}$ for some $C \in \mathcal{S}(B, k - \lambda)$ and $C' \in \mathcal{S}(B, k - \lambda')$. Then we have $B_1^\lambda = B_{\lambda'-\lambda+1}^{\lambda'}$. Note that Lemma 2.2(i) implies $\lambda \in \overline{\Lambda}(B_1^{\lambda'})$ and $B_{\lambda+1}^{\lambda'} = B_1^{\lambda'-\lambda}$. Hence we have

$$B_1^{\lambda'} = B_1^\lambda B_{\lambda+1}^{\lambda'} = B_{\lambda'-\lambda+1}^{\lambda'} B_{\lambda+1}^{\lambda'} = B_{\lambda'-\lambda+1}^{\lambda'} B_1^{\lambda'-\lambda},$$

and from Lemma 2.2(ii) we have $\lambda_0 \in \Lambda(B)$ such that $\lambda = \ell\lambda_0$ and $\lambda' = \ell'\lambda_0$ for some positive integers ℓ and ℓ' with $\ell < \ell'$. Hence we have $C = C'B_1^{\lambda_0} \cdots B_1^{\lambda_0}$ and this contradicts $C \in \mathcal{S}(B, k - \lambda)$. \square

Definition 2.5. Define $r_k(B)$ and $s_k(B)$ by

$$\begin{aligned} r_k(B) &= \Pr(x_1 \cdots x_k \in \mathcal{F}(B, k)), & k \geq n, \\ s_k(B) &= \Pr(x_1 \cdots x_k \in \mathcal{S}(B, k)), & k \geq 1. \end{aligned}$$

Then we have (1) implies that

$$\Pr(x_1^n = B, R_n(x) > k - n) = r_k(B) \text{ for } k \geq n$$

and (2) yields

$$\Pr(x_1^n = B, R_n(x) = k) = \Pr(B)s_k(B) \text{ for } k \geq 1.$$

Now we can calculate the distribution of the first return time by the following theorem, which is directly obtained by Lemma 2.4.

Theorem 2.6. *For i.i.d. processes, if $k > n$, we have*

$$r_k(B) = r_{k-1}(B) - \Pr(B)s_{k-n}(B).$$

For $k \geq n$

$$s_k(B) = r_k(B) - \sum_{\lambda \in \overline{\Lambda}(B)} \Pr(B_1^\lambda)s_{k-\lambda}(B).$$

And for initial values we have

$$r_n(B) = \Pr(B) \text{ and } s_i(B) = \begin{cases} 0 & \text{if } i < n, i \notin \Lambda(B), \\ \Pr(B_1^i) & \text{if } i \in \Lambda(B). \end{cases}$$

Proof. The first recurrence relation is implied (i) and The second recurrence relation is directly obtained by Lemma 2.4 (ii).

Since $\mathcal{F}(B, n) = \{B\}$, we have $r_n(B) = \Pr(B)$. When $i < n$, each $C \in \mathcal{S}(B, n)$ should satisfies that $C = B_1^i$ and $(CB)_1^n = B_1^i B_1^{n-i} = B$, which implies that $i \notin \Lambda(B)$. Therefore, if $i < n$ and $i \notin \Lambda(B)$, then an $\mathcal{S}(B, i) = \emptyset$ or $s_i = 0$ and if $i \in \Lambda(B)$, then $\mathcal{S}(B, i) = \{B_1^i\}$ or $s_i = \Pr(B_1^i)$. \square

For random binary sequences we have the followings:

$$r_k(B) = r_{k-1}(B) - 2^{-n}s_{k-n}(B), \quad k > n.$$

$$s_k(B) = r_k(B) - \sum_{\lambda \in \overline{\Lambda}(B)} 2^{-\lambda}s_{k-\lambda}(B), \quad k \geq n$$

with initial values

$$r_n(B) = 2^{-n} \text{ and } s_i(B) = \begin{cases} 0 & \text{if } i < n, i \notin \Lambda(B), \\ 2^{-i} & \text{if } i \in \Lambda(B). \end{cases}$$

For $(1/2, 1/2)$ i.i.d. processes the sequence $\{r_k(B)\}$ is same for the blocks with the same $\Lambda(B)$. Thus we classify all the n -blocks using $\Lambda(B)$ and compute s_k for each block B from different classes. The computation of $s_k(B)$ for every n -block B is necessary for the application in later sections and it is done recursively on computers.

TABLE 2. The Z -test for $n = 14$ and sample size = 100,000

Generator	number of blocks such that				Mean	Variance
	$Z < -2.57$	$Z < -1.96$	$Z > 1.96$	$Z > 2.57$		
Randu	6667	6737	8923	8744	4.99	799.97
ANSI	2110	2639	7394	6024	1.09	12.16
MS	2163	2692	7364	6001	1.09	11.90
Fishman	23	206	201	28	-0.01	0.78
ICG	69	398	404	81	0.00	1.00
Ran0	25	229	204	30	0.00	0.77
Ran1	31	243	242	27	0.00	0.79
Ran2	95	434	401	79	-0.01	1.02
Ran3	73	417	415	79	0.01	1.00
F90	93	437	421	101	0.01	1.03

3. TEST FOR PSEUDORANDOM NUMBER GENERATORS

We apply Z -test for $\log R_n$ to test PRNGs given in Section 4. We calculate for each n -block B the expectations and the standard deviations for $\log R_n$ numerically by computer using the values $s_k(B)$ from Theorem 2.6.

First, we construct a long binary sequence $x = (x_1x_2\dots)$ by juxtaposing binary numbers of the random bits from pseudorandom number generators. Define $L(B)$ by

$$L(B) = \{j : x_j \dots x_{j+n-1} = B\}.$$

Put

$$L(B) = \{\ell_1(B), \ell_2(B), \ell_3(B), \dots\}$$

with increasing order, i.e., $\ell_k(B) < \ell_{k+1}(B)$ for all $k \geq 1$. We obtain the sample mean of $E[\log R_n | x_1^n = B]$ by

$$\frac{1}{M} \sum_{i=1}^M \log(\ell_{i+1}(B) - \ell_i(B)),$$

where M is the sample size.

For each 2^n blocks we compare the theoretical values and the sample mean values where the sample size for every generator is 100,000 in our experiments and $n = 14$. The standard Z -test for $\log R_n$ is applied for these all 2^n blocks. We obtain the sample averages of $\log R_n$ by observing the recurrence times of each block and we compare them with $E[\log R_n]$ and $Var[\log R_n]$.

Table 2 show the number of blocks such that $Z < -2.57$, $Z < -1.96$, $Z > 1.96$, and $Z > 2.57$. Note that the total number of block is $2^{14} = 16384$. If the absolute value of Z -value is larger than 1.96 and 2.57, then the corresponding generator fails the test for the corresponding n with statistical confidence of 95% and 99%, respectively.

TABLE 3. The variance test for Type I blocks for Z -values ($n = 14$)

Generator	Type I-1	Type I-2	Type I-3	Type I-4	Type I-5	Type I-6
Fishman	0.66	0.79	0.97	0.76	0.67	0.68
ICG	1.22	0.97	0.91	0.95	1.16	0.87
Ran0	0.87	0.72	0.72	0.64	0.71	0.86
Ran1	0.78	0.93	0.79	0.78	0.88	0.76
Ran2	1.30	0.90	1.13	0.96	1.17	0.84
Ran3	0.96	1.07	0.93	0.97	1.05	0.96
F90	0.92	1.14	0.77	0.99	1.02	0.90

TABLE 4. The variance test for Type II blocks for Z -values ($n = 14$)

Generator	type II-1	Type II-2	Type II-3	Type II-4	Type II-5
Fishman	0.87	0.75	0.73	0.88	0.76
ICG	1.04	1.06	1.00	0.95	0.84
Ran0	0.95	0.75	0.73	0.93	0.79
Ran1	0.93	0.83	0.91	0.95	0.78
Ran2	1.16	0.89	0.88	1.26	1.27
Ran3	0.91	0.90	1.02	0.96	1.16
F90	1.03	1.03	1.00	1.01	1.13

In Table 2, Mean and Variance denote the mean and variance of Z -values among the all 2^{14} blocks. For ideal generators have the mean and variance near 0 and 1 respectively. Apparently, Randu, ANSI, and MS seem to fail the test.

Since the Z -values among 2^n blocks are highly correlated, we need to reduce the correlation among the sample values. For example, the return time of 00000000000000 and 000000000000001 are highly related since there is a big chance that 000000000000001 appears right after 00000000000000. Therefore, we test the variance test on only special kind of blocks. A binary block will be regarded as an integer in binary expansion, i.e., for $B = 00000001000010$ we will say $B = (1000010)_{(2)} = 130$. We will take a set of blocks of the form $B \equiv a \pmod{b}$ or $B = bk + a$ for some integer $k \geq 0$, of which blocks are not easily overlapped with each other. If two blocks $B = bk + a$ and $B' = bk' + a$ are overlapped, then we have

$$B = CB_0, \quad B' = B_0C'$$

for some ℓ -blocks C and C' . In this case, $2^\ell(bk + a) + m' = 2^{14}m + bk' + a$ for some m and m' with $0 \leq m, m' \leq 2^\ell - 1$. Note that m and m' are representation of C and C' . Therefore we have

$$(2^\ell - 1)a + m \equiv dm' \pmod{b},$$

where $2^{14} \equiv d \pmod{b}$. Choose b as $2^{14} \equiv 1 \pmod{b}$. Then if $B = CB_0$ and $B' = B_0C'$ for some ℓ -blocks C and C' , then we have

$$(2^\ell - 1)a \equiv m \pmod{b}, \quad \text{for some } 0 \leq m \leq 2^\ell - 1. \quad (3)$$

Then possible value of b for $2^{14} \equiv 1 \pmod{b}$ with $1 < b < 2^{14} - 1$ is 3, 43, 127, 129, 381 and 5461.

Choose $b = 127$. Then for each $a = 64, 72, 84, 106, 118, 126$ there is no ℓ with $1 \leq \ell \leq 6$ which satisfies (3). When $b = 129$, if we pick $a = 65, 83, 108, 120, 128$, then (3) is not satisfied for any ℓ with $1 \leq \ell \leq 6$. We say the block B is of type I-1 (respectively I-2, I-3, I-4, I-5 and I-6), if the integer obtained by the binary block is $64 \pmod{127}$ (respectively $72 \pmod{127}$, $84 \pmod{127}$, $106 \pmod{127}$, $118 \pmod{127}$ and $126 \pmod{127}$). Similarly B is of type II-1 (respectively II-2, II-3, II-4 and II-5), if the integer obtained by the binary block is $65 \pmod{129}$ (respectively $83 \pmod{129}$, $108 \pmod{129}$, $120 \pmod{129}$ and $128 \pmod{129}$).

The variance test of the Z -values over the blocks of each type is applied. The test results are presented in Table 3 and 4. Test I-1,2,3,4,5,6 and II-1,2,3,4,5 denote the tests on the blocks of type I-1,2,3,4,5,6 and II-1,2,3,4,5 respectively. The number of each type I-1, I-2, I-3, I-4, I-5, I-6 blocks is 129 and degree of freedom of the corresponding chi-distribution is $n - 1 = 128$. For each type I blocks, if the sample variance is bigger than 1.26 or less than 0.77, than it fails the variance test with 5% significance level and if the sample variance is bigger than 1.35 or less than 0.71, than it fails the variance test with 1% significance level. The number of each type II-1, II-2, II-3, II-4, II-5 blocks is 127 and the degree of freedom of the chi-distribution is $n - 1 = 126$. For each type II blocks, if the sample variance is bigger than 1.26 or less than 0.77, than it fails the variance test with 5% significance level and if the sample variance is bigger than 1.35 or less than 0.71, than it fails the variance test with 1% significance level. Tests for Randu, ANSI, and MS are skipped because they fail the previous test. In this test all generator which is made up of one linear congruential generator fail the test.

4. GENERATORS

The following is a list of PRNGs tested in Section 3. We generate binary sequences using the algorithms listed in Table 5. A linear congruential generator $LCG(M, a, b)$ means the algorithm given by

$$X_{n+1} \equiv aX_n + b \pmod{M}.$$

Randu is an outdated generator developed by IBM in the sixties. ANSI and Microsoft are the generators used in C libraries by American National Standard Institute and Microsoft, respectively. For a prime p , the inversive congruential generator $ICG(p, a, b)$ is that

$$X_{n+1} \equiv a\overline{X_n} + b \pmod{p},$$

where \overline{X} is the multiplicative inverse of x modulo p . The generators Ran0, Ran1, Ran2 and Ran3 are from [13]. Ran0 is the linear congruential generator by Park and Miller[12]. Ran1 is

TABLE 5. The tested random number generators

Name	Generator	Period
Randu	$LCG(2^{31}, 65539, 0)$	2^{29}
ANSI	$LCG(2^{31}, 1103515245, 12345)$	2^{31}
Microsoft	$LCG(2^{31}, 214013, 2531011)$	2^{31}
Fishman	$LCG(2^{31} - 1, 950706376, 0)$	$2^{31} - 2$
ICG	$ICG(2^{31} - 1, 1, 1)$	$2^{31} - 1$
Ran0	$LCG(2^{31} - 1, 16807, 0)$	$2^{31} - 2$
Ran1	Ran0 with shuffle	$> 2^{31} - 2$
Ran2	L'Ecuyer's algorithm with shuffle	$> 2.3 \times 10^{18}$
Ran3	$X_n \equiv X_{n-55} - X_{n-24} \pmod{2^{31}}$	$\geq 2^{55} - 1$
F90	Ran0 combined with shift register	$\sim 3.1 \times 10^{18}$

Ran0 with Bays-Durham shuffle. Ran2 is L'Ecuyer's generator[9] made up of

$$LCG(2147483563, 40014, 0) \text{ and } LCG(2147483399, 40692, 0)$$

with Bays-Durham shuffle. Ran3 is a subtractive lagged Fibonacci sequences. F90[14] is Ran0 combined with a Marsaglia shift register generator, which is the form of $X_{n+1} = X_n(I \oplus L^{13})(I \oplus R^{17})(I \oplus L^5)$, where \oplus denotes the binary exclusive-or operation and L (resp. R) is the bitwise left-shift (resp. right-shift).

REFERENCES

- [1] G.H. Choe, C. Kim, and D.H. Kim, *Applications of ergodic theory to pseudorandom numbers*, Bull. Korean Math. Soc. **35** (1998), 173–187.
- [2] G.H. Choe and D.H. Kim, *Average convergence rate of the first return time*, Colloq. Math. **84/85** (2000), 159–171.
- [3] G.H. Choe and D.H. Kim, *The first return time test of pseudorandom numbers*, J. Comput. Appl. Math. **143** (2002), no. 2, 263–274.
- [4] J. Oliver, *A central limit theorem for non-overlapping return times*, J. Appl. Probab., **43** (2006), 32–47.
- [5] M. Kac, *On the notion of recurrence in discrete stochastic processes*, Bull. Amer. Math. Soc. **53** (1947), 1002–1010.
- [6] C. Kim, G.H. Choe and D.H. Kim, *Tests of randomness by the gambler's ruin algorithm*, Appl. Math. Comput. **199** (2008), 195–210.
- [7] D.H. Kim, *The recurrence of blocks for Bernoulli processes*, Osaka J. Math. **40** (2003), no. 1, 171–186.
- [8] I. Kontoyiannis, *Asymptotic recurrence and waiting times for stationary processes*, J. Theor. Probab. **11** (1998), 795–811.
- [9] P. L'Ecuyer, *Efficient and portable combined random number generators*, Comm. ACM. **31** (1988), no. 6, 742–749.
- [10] U. Maurer, *A universal statistical test for random bit generators*, J. Cryptology **5** (1992), 89–105.
- [11] D. Ornstein and B. Weiss, *Entropy and data compression schemes*, IEEE Trans. Inform. Theory **39** (1993), no. 1, 78–83.
- [12] S. Park and K. Miller, *Random number generators: good ones are hard to find*, Comm. ACM. **31** (1988), no. 10, 1192–1201.

- [13] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in C*, 2nd. ed., Cambridge Univ. Press, Cambridge, 1992.
- [14] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, *Numerical Recipes in Fortran 90*, Cambridge Univ. Press, Cambridge, 1996.
- [15] P. Shields, *Waiting times: positive and negative results on the wyner-ziv problem*, J. Theor. Probab., **6** (1993), 499–519.
- [16] A.D. Wyner and J. Ziv, *Some asymptotic properties of the entropy of stationary ergodic data source with applications to data compression*, IEEE Trans. Inform. Theory **35** (1989), no. 6, 1250–1258.
- [17] A.J. Wyner, *Strong Matching Theorems and Applications to Data Compression and Statistics*, Ph.D. thesis, Stanford University, Department of Statistics, 1993.
- [18] A.J. Wyner, *More on recurrence and waiting times*, Ann. Appl. Probab. **9** (1999), no. 3, 780–796.
- [19] J. Ziv and A. Lempel, *A universal algorithm for sequential data compression*, IEEE Trans. Inform. Theory **23** (1977), no. 3, 337–343.

A COMPUTATION METHOD IN PERFORMANCE EVALUATION IN CELLULAR COMMUNICATION NETWORK UNDER THE GENERAL DISTRIBUTION MODEL

KYUNGSUP KIM¹

¹COMPUTER SCIENCE AND ENGINEERING, CHUNGNAM NATIONAL UNIVERSITY, YOONSEONG-GU, DAEJEON, KOREA

E-mail address: kskim@cse.cnu.ac.kr

ABSTRACT. The paper considers the computation method in the performance evaluation of cellular network in the phase-type distribution assumptions that the channel holding times induced from mobility are modeled by well-fitted distributions to reflect an actual situation. When we consider a phase-type distribution model instead of exponential distribution, the complexity of the computation increases exponentially even though the accuracy is improved. We consider an efficient numerical algorithm to compute the performance evaluations in cellular networks such as a handoff call dropping probability, new call blocking probability, and handoff arrival rate. Numerical experiment shows that numerical analysis results are well approximated to the results of simulation.

1. INTRODUCTION

We consider performance evaluation problems of the call admission control in cellular networks. Various handoff priority-based call admission control (CAC) schemes have been proposed [1] [2]. A guard channel scheme proposed by Hong and Rappaport [3] is that a number of channels in each cell are reserved for exclusive use by handoff calls and remaining channels are shared by both new and handoff calls. In the queue priority scheme, when all channels are occupied, either new calls are queued while handoff calls are dropped [4], or new calls are dropped while handoff calls are queued [5] [6], or both calls are queued and rearranged [7]. In this paper, a handoff prioritization strategy with guard channel and queue is considered.

We note that there are some rough approximations in cellular system modeling in the past research literature. First, the channel holding times for new calls and handoff calls have been assumed to be independent, exponentially distributed, and have the same average values [3] [8]. It is known that the distribution of the cell dwell time induced by user's mobility should be modeled by a general distribution in order to evaluate exactly the performance metrics. Orlik and Rappaport proposed that the sum of exponential distributions (Hyper-exponential function) can

2000 *Mathematics Subject Classification.* 68M10,68M20.

Key words and phrases. Queueing, Performance Evaluation, Call admission control, cellular network .

The author's work was supported in part by Korea Research Foundation Grant funded by Korea Government(MOEHRD, Basic Research Promotion Fund) (KRF-2005-003- D00300).

be approximated to the dwell time distribution [9] [10]. Second, there are different classes of telephone traffics with different qualities of service (QoS) in the modern communication network. For examples, the new call and handoff call has different characteristics such as channel holding times and cell dwell times [3] [8], and the real time traffic and non-real time traffics have different QoSs. Therefore, we introduce multidimensional Markov model to handle the multi-class traffics. However, if we assume more general model and increase the dimension in order to get more exact solutions, then the complexity of computation increase exponentially. We handle the numerical computation problems in computing performance evaluations in the call admission control of cellular networks

In this paper, we develop a numerical algorithm computing a performance of a CAC using a channel reservation and handoff queueing, when there are different mobility patterns of users. We note that the mobility of wireless users impacts the performance, such as the blocking probability and the mean of delay. Thus, more realistic analytical model of the mobility and service rate is needed. General distribution model has computation complexity problems due to the exponential increase of dimension. Thus we propose a novel numerical method to compute the performance values of the CAC as the values for a handoff call dropping probability, new call blocking probability, and handoff arrival rate, when new call arrival rates are known.

The remainder of this paper is structured as follows. In section 2, we consider a model of mobility, and the computation method for the $M/Ph/C/K$ queueing model. The matrix solution of the analytical model for Channel guard scheme with handoff buffer is presented and the numerical algorithm is discussed in section 3. In Section 4, the numerical results of the analytical model are verified by some numerical examples. Finally, conclusions are drawn.

2. TRAFFIC MODEL IN MOBILE NETWORKS

2.1. Mobility Modeling. The probability that a new call is blocked is denoted by new call blocking probability (CBP) (P_{nb}) and the probability that a handoff call is dropped is denoted by Handoff Call Dropping Probability (CDP) (P_{hd}). These quantities are most significant QoS metrics in CAC scheme. When new calls and handoff calls are competing for the usage of a finite channel resource in a cell, their claims for QoS are different. From users' point of view, a call forced to terminate during service is more annoying than the new call blocked at its start. Therefore, handoff calls are commonly given a higher priority in accessing the wireless channel. This can be realized by handoff priority-based Call Admission Control (CAC) Schemes. We note that the mobility patterns of mobile user such as slow or fast speed influence to the QoS in wireless networks. The mobility plays an important role in the performance of a cellular networks.

Let λ_n be the arrival rate for new calls and λ_h be the arrival rate for handoff call. λ_h depends on λ_n . The cell dwell times are modeled as exponential distributions in some past literature, but real fast user's ones are not exponentially distributed [8]. We can reasonably assume that the holding channel times of the static user is an exponential distribution. The average cell dwell times for high speed users depend on the speed of the users. The users are moving in a random movement pattern in a cellular network. We need a general fitting model for approximating the

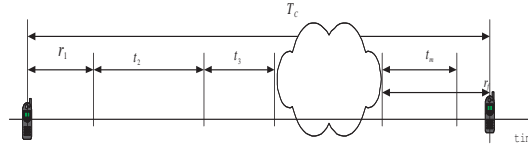


FIGURE 1. The time diagram for call holding time and cell dwell time

measured data for fitting. The distribution of a Hyper-Erlang random variable X is defined by

$$f(x) = \sum_{i=1}^p \alpha_i \frac{(\mu_i x)^{r_i-1}}{(r_i-1)!} \mu_i \exp(-\mu_i x) \tag{1}$$

where the mean rate μ is $\mu = \sum_{i=1}^p \alpha_i r_i \mu_i$ for some given p .

We discuss how the distributions of the cell dwell time and the call holding time influence the distributions of the new and handoff call channel holding times in Figure 1. Let T_c denote the lifetime of the call holding that is the length from the instant of admission by the base station to the instant when the connection is terminated in the cell or in another cell after several more handovers. If the call holding time has an exponential distribution, denoted by $f_c(x)$, then the residual call length of the handover also has the same distribution, due to the memoryless property. Figure 1 shows diagram for our study similar to that in [11]. Let t_i for $2 \leq i \leq m$ be the typical cell dwell time in a cell for a handoff user, r_1 be a cell dwell time in the first cell for a new call. r_f is a residual call holding time. We assume that cell dwell time are generally distributed as (1) depending on the mobility of a mobile user

The number of handoff times H that a mobile crosses different boundaries during a call holding time is a random variable depending on the cell size, call holding time and mobility parameter [12]. If the density function of the independent identical distribution calling holding times has a general distribution, then the handoff rate for a nonblocking call is given by

$$E[H] = \frac{(1 - P_{nb})}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} \frac{\mu[1 - f^*(s)]}{s^2[1 - (1 - P_{hd})f^*(s)]} f_c^*(-s) ds, \tag{2}$$

where $f_c(t)$ is defined by the distribution of call holding time and $f(t)$ is defined by the distribution of cell dwell time with mean rate μ (refer to [11]). Then the handoff rate λ_h is

$$\lambda_h = \lambda_n(1 - P_{nb})E[H]. \tag{3}$$

We can see that the handoff call arrival rate λ_h depends on the user mobility and the new call arrival rate (λ_n) (3).

2.2. M/Ph/C/K Modeling. We model call admission control of a single cell by using $M/Ph/C/K$ model. New call and handoff arrival processes are assumed Poisson processes with arrival rates λ_n and λ_h , respectively. Cell dwell times are modeled by the server with phase-type distribution. A phase-type distribution has finite states. For k processes, we require a $k + 2$ component vector $(n; n_q, n_k, n_{k-1}, \dots, n_1)$ to describe the state of a single cell behavior where n is the total number of customers in the system, the remaining i components n_i for $0 \leq i \leq k$ represent

the distribution of customers in various phases of service, and n_q is the number of customers in the queue. A lexicographic set \mathcal{M}_n is defined by

$$\mathcal{M}_n = \{(n; n_q, n_k, \dots, n_1)\} \quad (4)$$

where $n - n_q = \sum_{i=1}^k n_i$. Here, $n_q = 0$ if $n \leq C$. $n_q = n - C$ if $n > C$. Let us define a lexicographic ordering relation on the set \mathcal{M}_n . Let us define the ordering relation (\prec) of the lexicographic labeling by $(n; n_q, n_k, \dots, n_1) \prec (n; n'_q, n'_k, \dots, n'_1)$ if it satisfies one of the following statements:

- (1) there is a first j from the left side such that $n_j < n'_j$, and $n_i = n'_i$ for all $i > j$ if $0 \leq n \leq C$ ($n_q = n'_q = 0$)
- (2) there is a first j from the left side such that $n_j < n'_j$, $n_q = n'_q$ and $n_i = n'_i$ for all $i > j$ if $n \geq C$.

The cardinality of a lexicographic ordering set \mathcal{M}_n for n customers is defined by m_n ($|\mathcal{M}_n| = m_n$). Accordingly, the cardinality of \mathcal{M}_n can be computed by

$$m_n = \begin{cases} \frac{(n+k-1)!}{n!(k-1)!}, & \text{when } 1 \leq n \leq C; \\ \frac{(C+k-1)!}{C!(k-1)!}, & \text{when } K \geq n > C. \end{cases} \quad (5)$$

where $n!$ is the factorial of n and K is the maximal number of customers in the system. This gives the number of the total elements of the state space, $T = 1 + \sum_{n=1}^K m_n$. There is an one-to-one mapping such that $(n; n_q, n_k, n_{k-1}, \dots, n_1)$ corresponds to $[n, l]$ for $1 \leq l \leq m_n$ with the same order. The state probability $p_{n; n_q, n_k, n_{k-1}, \dots, n_1}$ represents the probability for which there are n customers in system and n_i customers in the phase i of service for $i = 1, \dots, k$, where $n_k + n_{k-1} + \dots + n_1 = \min\{n, C\}$. If there is no ambiguity, we use the simplified notation $p_{[n, l]} = p_{n; n_q, n_k, n_{k-1}, \dots, n_1}$. The vector-valued balanced equation can be written by

$$\begin{aligned} D_0 \vec{p}_0 &= V_0 \vec{p}_0 + W_0 \vec{p}_1 \\ D_n \vec{p}_n &= U_n \vec{p}_{n-1} + V_n \vec{p}_n + W_n \vec{p}_{n+1} \\ D_K \vec{p}_K &= U_K \vec{p}_{K-1} + V_K \vec{p}_K, \end{aligned} \quad (6)$$

where $\vec{p}_n = [p_{[n, 1]} \ p_{[n, 2]} \ \dots \ p_{[n, m_n]}]^t$ is a column vector with dimension m_n , m_n is the number of the possible states when there are n customers [13]. V_n is a matrix that presents a transition from an internal state of a server into another internal state without changing the numbers of customers in the queue and system. W_n is denoted a transition matrix induced by a customer departure from the system, U_n is a transition matrix induced by a new customer arrival. D_n is a diagonal matrix whose (i, i) entry equals the sum of all the entries in the i -th columns of the matrices U_{n+1} , V_n and W_{n-1} , i.e.

$$D_n(i, i) = \sum_{j=1}^{m_{n-1}} U_n(j, i) + \sum_{j=1}^{m_n} V_n(j, i) + \sum_{j=1}^{m_{n+1}} W_n(j, i).$$

Refer to the global state description of $M/Ph/C$ in Chapter 6 [14] to know the detail description. The normalization equation is satisfied as follows: $\sum_{n=0}^N \sum_{i=1}^{m_n} p_{[n,i]} = 1$. The steady state probabilities p_n can be computed by $\sum_{i=1}^{m_n} p_{[n,i]}$.

3. ANALYSIS OF CALL ADMISSION CONTROL SCHEME

3.1. CAC with Guard channel and Handoff Queueing. We will analyze the performance of a CAC scheme with handoff queueing and guard channels. We assume that the channel holding times for new calls and handoff calls are independent and have different distribution [3] [12] [9]. The one-dimensional Markov chain model for CAC schemes assuming that cell dwell times of new calls and handoff calls are identically distributed may not be appropriate. Therefore, the multi-dimensional Markov chain model is needed.

Let us consider the channel holding time. There are two kinds of channel holding times : a new call channel holding and handoff call channel holding time. Let t_{nh} and t_{hh} denote the new call channel holding time and the handoff call channel holding time, respectively. The new call channel holding time is $t_{nh} = \min\{T_c, r_1\}$ and the handoff call channel holding time is $t_{hh} = \min\{r_f, t_m\}$. We separated calls into new calls and handoff calls when considering the channel holding time. We need to consider the channel holding time for merged traffic of new calls with rate λ_n and handoff calls with rate λ_h . We assume that the distributions of channel holding times is approximated by hyper-Erlang distribution. We study the multi-dimensional Markov chain under the assumption that some random variable, such as dwell time may be modeled by the hyper-Erlang distribution. We study the multi-dimensional Markov chain under the assumption that some random variable, such as cell dwell time may be modeled by the Hyper-Erlang distribution [12] [9]. We develop an algorithm that computes the blocking probability of new calls and the dropping probability of handoff calls. We develop an algorithm that computes the blocking probability of new calls and the dropping probability of handoff calls.

Figure 2 shows an example for $C = 6$, $M = 2$, $N = 1$ and $B = 3$. Let p_{n_1, n_2} denote the steady-state probability that there are n_1 new calls and n_2 handoff calls in the cell. Let us consider the two-dimensional steady-state probability $p(n_1, n_2)$ occurring in the $M/Ph/C/N$ queueing system. The Poisson interarrival and Hyper-Erlang distributions in $M/Ph/C/N$ queueing system can model two-dimensional Markov chain as a generalized version of two-dimensional Markov chain introduced in [15]. We assume that each base station has a finite buffer size B . Let C be the total number of channels in a cell and M and N be the number of channels only assigned for new calls and handoff calls, respectively. There are $C - M - N$ shared channels that can be used by either type of call. All channels are employed in a first-come first-serve manner. The queue model can be described by a two-dimensional (i, j) Markov chain, where i and j denote the numbers of existing new calls and handoff calls in a cell, respectively. The two-dimensional state space is given by

$$\mathcal{S} = \{(i, j) | 0 \leq i < M, 0 \leq j \leq C - M + B \quad \text{or} \\ M \leq i \leq C - N, 0 \leq j \leq C - i + B\}.$$

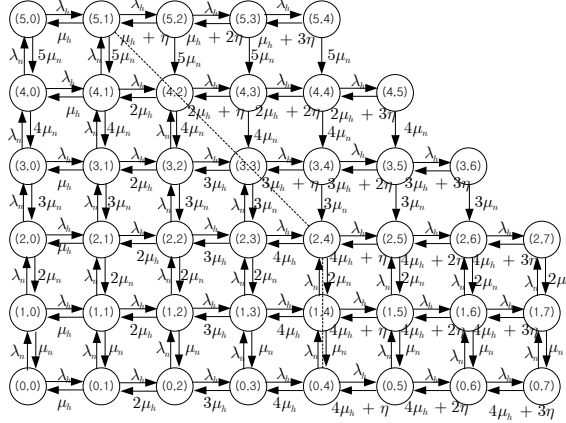


FIGURE 2. Two dimension Markov Chain

The state space \mathcal{S} of the two-dimensional Markov chain can be divided into three parts as follows:

$$\begin{aligned} \mathcal{S}_1 &= \{(i, j) | i + j < C, 0 < i < C - N, 0 < j < C - M\} \\ \mathcal{S}_2 &= \{(i, j) | i \leq M, C - M \leq j \leq C - M + B\} \\ \mathcal{S}_3 &= \{(i, j) | i + j \geq C, M \leq i \leq C - N\}. \end{aligned}$$

Let us define a vector $\vec{n}_i = (n_{i,q}, n_{i,k}, \dots, n_{i,1})$. For c servers, we require a $k+2$ component vector to describe the state of the system, which is defined by $(n_i; \vec{n}_i) = (n_i; n_{i,q}, n_{i,k}, \dots, n_{i,1})$ in (4) where $n_i = n_{i,q} + \sum_{j=1}^k n_{i,j}$ for $i = 1, 2$. The new call and handoff states can be re-ordered by lexicographical labeling. Let p_{n_1, n_2} denote the state probability that there are n_1 new calls and n_2 handoff calls. The probability p_{n_1, n_2} can be divided into the vector-valued probability \vec{p}_{n_1, n_2} by lexicographical ordering. The transition diagram for the new and handoff call bounding schemes with buffers for handoff is modeled by the two-dimensional Markov chain.

Define $p_{n_1, n_2}(l_1, l_2)$ by $p_{n_1, n_2}(l_1, l_2) = p_{n_1; \vec{n}_1, n_2; \vec{n}_2}$ where \vec{n}_i is l_i -th elements of \mathcal{M}_{n_i} for $i = 1, 2$. Then, the state vector \vec{p}_{n_1, n_2} can be defined by

$$\vec{p}_{n_1, n_2} = \begin{bmatrix} p_{n_1, n_2}(0, 0) \\ p_{n_1, n_2}(0, 1) \\ \vdots \\ p_{n_1, n_2}(0, m_{n_2}) \\ p_{n_1, n_2}(1, 0) \\ p_{n_1, n_2}(1, 1) \\ \vdots \\ p_{n_1, n_2}(m_{n_1}, m_{n_2}) \end{bmatrix}. \quad (7)$$

The state vector \vec{p}_{n_1, n_2} can be defined such that $s = m_{n_1}s_1 + s_2$ element is $p_{n_1, n_2}(s_1, s_2)$. Then state probability p_{n_1, n_2} is the sum of all the elements of the state vector probability \vec{p}_{n_1, n_2} since the probability p_{n_1, n_2} can be divided into the vector-valued probability \vec{p}_{n_1, n_2} by lexicographical ordering. In Figure 2, we can see that \mathcal{S} can be divided into \mathcal{S}_1 , \mathcal{S}_2 , and \mathcal{S}_3 . First, compute the transition matrix of state $(i, j) \in \mathcal{S}_1$. We define U_I, W_I, V, I_U, I_W for convenience) from $\mathcal{M}_{n_1} \otimes \mathcal{M}_{n_1}$ to $\mathcal{M}_{\vec{n}_1} \otimes \mathcal{M}_{\vec{n}_2}$ as follows:

$$\begin{aligned} U_I(n_1, n_2) &\triangleq U_{n_1} \otimes I_{m_{n_2}} \\ W_I(n_1, n_2) &\triangleq W_{n_1} \otimes I_{m_{n_2}} \\ V(n_1, n_2) &\triangleq V_{n_1} \otimes I_{m_{n_2}} + I_{m_{n_1}} \otimes V_{n_2} \\ I_U(n_1, n_2) &\triangleq I_{m_{n_1}} \otimes U_{n_2} \\ I_W(n_1, n_2) &\triangleq I_{m_{n_1}} \otimes W_{n_2} \end{aligned} \quad (8)$$

where U_{n_j}, V_{n_j} and W_{n_j} for $j = 1, 2$ are defined by (6), I_m is an $m \times m$ identity matrix, \otimes is a Kronecker product, and (n_1, n_2) is a feasible state in \mathcal{S} . Then, the global balanced equation can be written by

$$\begin{aligned} D(n_1, n_2)\vec{p}_{n_1, n_2} &= U_I(n_1, n_2) \cdot \vec{p}_{n_1-1, n_2} + \\ &V(n_1, n_2)\vec{p}_{n_1, n_2} + W_I(n_1, n_2) \cdot \vec{p}_{n_1+1, n_2} \\ &+ I_U(n_1, n_2) \cdot \vec{p}_{n_1, n_2-1} + \\ &I_W(n_1, n_2) \cdot \vec{p}_{n_1, n_2+1} \end{aligned} \quad (9)$$

where $D(n_1, n_2) = \{\lambda_n + \lambda_h + n_1\mu_n + n_2\mu_h\}I_{m_{n_1} \cdot m_{n_2}}$, the negative new call state probability \vec{p}_{-1, n_2} are zero m_{n_2} vectors, and the negative handoff call state probability $\vec{p}_{n_1, -1}$ are zero m_{n_1} vectors. Second, for (i, j) included inside the region of \mathcal{S}_2 (not a boundary value), we can derive U_I, W_I, V, I_U, I_W . $D(n_1, n_2)$ is computed by $D(n_1, n_2) = \{\lambda_n + \lambda_h + n_1\mu_n + (C - M)\mu_h + (n_2 - C + M)\eta\}I$. Third, for (i, j) included inside of the region \mathcal{S}_3 , we note from Fig. 2 that there are no arrows in the places of new call arrival rates because new calls are blocked. Thus, U_I 's are zero matrices. For W_I , after a new call is served, the remaining

capacity is used by handoff calls, because there are handoff calls waiting in the queue. Thus, the number of handoff calls instantly receiving service is added by 1 in the first service phase. W_I is computed by

$$W_I(n_1, n_2) \triangleq W_{n_1} \otimes \begin{bmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

$D(n_1, n_2)$ is computed by $D(n_1, n_2) = \{\lambda_h + n_1\mu_n + (C - n_1)\mu_h + (n_2 - C + n_1)\eta\}I$. Finally, we should carefully consider the boundary states, such as $\mathcal{S}_1 \cap \mathcal{S}_2$, $\mathcal{S}_1 \cap \mathcal{S}_3$, $\mathcal{S}_2 \cap \mathcal{S}_3$ and other boundaries of \mathcal{S}_i for $i = 1, 2, 3$. In order to handle the total state probability easily, we define a vector \vec{P}_n such that

$$\vec{P}_n = \begin{pmatrix} \vec{p}_{0,n} \\ \vec{p}_{1,n-1} \\ \vdots \\ \vec{p}_{n-1,1} \\ \vec{p}_{n,0} \end{pmatrix}$$

where the total number of new calls and handoff calls is n . Then we can obtain a global equation, as follows: $\mathcal{U}_n \vec{P}_{n-1} + (\mathcal{V}_n - \mathcal{D}_n) \vec{P}_n + \mathcal{W}_n \vec{P}_{n+1} = 0$, for $0 \leq n \leq K$, where \mathcal{U}_n , \mathcal{V}_n , \mathcal{W}_n and \mathcal{D}_n are defined by

$$\mathcal{U}_n = \begin{bmatrix} I_U(0, n) & 0 & \cdots & 0 \\ U_I(1, n-1) & I_U(1, n-1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & U_I(n, 0) \end{bmatrix}$$

$$\mathcal{W}_n = \begin{bmatrix} I_W(0, n) & W_I(0, n) & \cdots & 0 \\ 0 & I_W(1, n-1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & W_I(n, 0) \end{bmatrix}$$

$$\mathcal{V}_n = \text{diag}[V(0, n), V(1, n-1), \dots, V(n, 0)]$$

$$\mathcal{D}_n = \text{diag}[D(0, n), D(1, n-1), \dots, D(n, 0)].$$

\mathcal{D}_n is a diagonal matrix whose (i, i) entry equals the sum of all the entries in the i -th columns of the matrices \mathcal{W}_{n-1} , \mathcal{V}_n and \mathcal{U}_{n+1} . The above global equation can be written by the following

transition probability matrix

$$Q = \begin{bmatrix} \mathcal{A}_0 & \mathcal{W}_0 & \cdots & 0 & 0 \\ \mathcal{U}_1 & \mathcal{A}_1 & \cdots & 0 & 0 \\ 0 & \mathcal{U}_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathcal{A}_{K-1} & \mathcal{W}_{K-1} \\ 0 & 0 & \cdots & \mathcal{U}_K & \mathcal{A}_K \end{bmatrix},$$

in which we set $\mathcal{A}_i = (\mathcal{V}_i - \mathcal{D}_i)$.

In order to handle the total state probability easily, we define a vector \vec{P}_n such that $\vec{P}_n = [\vec{p}_{0,n} \ \vec{p}_{1,n-1} \ \cdots \ \vec{p}_{n-1,1} \ \vec{p}_{n,0}]$ where the total number of new calls and handoff calls is n . Then, we can obtain a global balance equation, as follows:

$$\mathcal{U}_n \vec{P}_{n-1} + (\mathcal{V}_n - \mathcal{D}_n) \vec{P}_n + \mathcal{W}_n \vec{P}_{n+1} = 0, \quad (10)$$

for $0 \leq n \leq K$. The detail derivation is omitted for the space saving. We solve the following equations

$$\begin{aligned} Q\vec{P} &= 0 \\ E\vec{P} &= 1 \end{aligned} \quad (11)$$

where $E = [1 \ 1 \ \cdots \ 1]$. Using the results, let us compute the new call dropping probability, the terminated handoff call probability, and queueing delay as QoS metrics. A new call arrival is blocked when it arrives at the state $(i, j) \in \mathcal{S}_3$. Therefore, the new call blocking probability P_{nb} is the sum of the conditional state probabilities when a new call arrives in the state $(i, j) \in \mathcal{S}_1$, such as $P_{nb} = \sum_{(i,j) \in \mathcal{S}_3} p_{i,j}$. The dropping probability of a handoff call can be calculated as the fraction of the incomplete handoff calls whose mobile leave the handoff area prior to their coming into the first queue position and getting a channel. The dropping probability $P_{hd|(i,j)}$ is defined by $P_{hd} = \sum_{(i,j) \in \mathcal{HD}} p_{i,j}$ where $\mathcal{HD} = \{(i, j) \in \mathcal{S} | j = C - M + B, \text{ or } i + j = C\}$.

We consider several methods for computing stationary probability distributions for large Markov Chain. There are two methods, direct and iterative, for solving linear systems. Iterative methods are the most commonly used methods for obtaining the stationary probability vector from the infinitesimal generator. In iterative method, the involved operations do not alter the form of the matrix and thus compact storage, which minimize the amount memory required to store the matrix, may be conveniently implemented, since the matrices involved are large and sparse. With direct methods, the elimination of one non-zero element of the matrix during the reduction phase often results in the creation of several non-zero elements in the position which contained zero. A successful direct method must incorporated a means of overcoming these difficulties.

When we use a good initial approximation, we should expect to compute the real solution in relatively few iterations. This is especially beneficial when a series of related experiments is being conducted and there is a little change. But iterative methods have the disadvantage

that they often require a long time to converge to the desired solution. Direct methods can be recommended if they obtain the solution in less time. However, we solve the linear system by direct method keeping sparse matrix.

3.2. Fixed point algorithm. We want to compute the CAC parameters as the values for a handoff call dropping probability P_{hd} , new call blocking probability P_{nb} , and handoff arrival rate λ_h , when new call arrival rates λ_n are known. These values can not be computed by using local information in a single cell, but need the global information. However, it is impossible to know the global information, because the total cellular system is very large and dynamics. The values for λ_h depends on the integration on each the drop or blocking probability of the total cellular system. So, we believe that a local value of λ_h measured in single cell is not a steady state value and is a dynamical value depending on instant state. Therefore, the values for P_{hd} , P_{nb} , λ_h , and λ_n should be predicted by an iterative method under the simplified model similar to [16] [3]. Beginning with an proper initial guess for the unknowns, the equations are solved numerically using an iterative method. This section shows how to use an iterative technique to compute P_{nb} , and P_{hd} using the equations derived in Sec. 3.1. The iterative algorithm is as follows:

Algorithm 1 (Fixed point algorithm). *Compute P_{nb} , and P_{hd} :*

- *Input parameters: the new call arrival rate λ_n , the number of channels C , and the mean and derivation of the cell dwell time.*
- *Output values: the handoff call arrival rate λ_h , the new call blocking probability P_{nb} , and the handoff dropping probability P_{hd} .*
 - (1) *Select initial values for P_{nb} and P_{hd} .*
 - (2) *Compute the handoff call rate λ_h as (3) (the instant value can directly be measured).*
 - (3) *Update old values:*

$$P_{nb,old} \leftarrow \alpha P_{nb,old} + (1 - \alpha) P_{nb}$$

$$P_{hd,old} \leftarrow \alpha P_{hd,old} + (1 - \alpha) P_{hd},$$

where $0 \leq \alpha < 1$.

- (4) *Compute the new-call-blocking and handoff-call-dropping probabilities (P_{nb} and P_{hd} , respectively) by using the results in Sec. 3.1.*
- (5) *If $|P_{nb,old} - P_{nb}|$ and $|P_{hd,old} - P_{hd}|$ are larger than the given thresholds, then go to step 2. Otherwise, go to the final step*
- (6) *The values for λ_h , P_{nb} and P_{hd} converge.*

In Step 3, α is an exponential moving average factor. The convergence rate depends on α .

4. NUMERICAL RESULTS

In this section, the numerical computation results obtained with our analytical model are discussed. We compare the performance QoS metrics of CAC with guard channel and finite

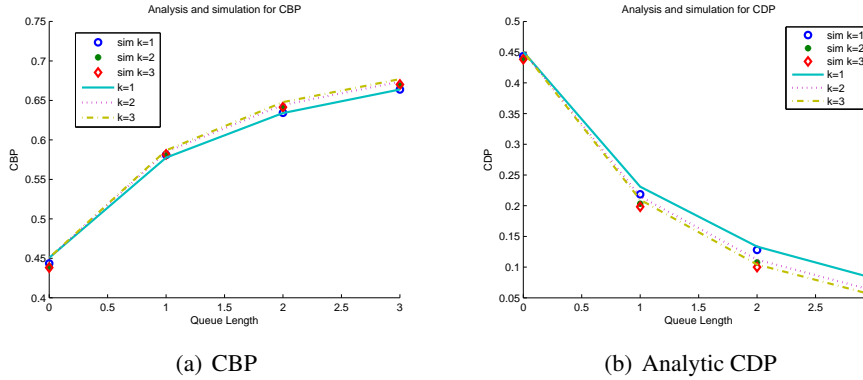


FIGURE 3. The loss probabilities(CBP and CDP) for Guard channel with respect to queue length

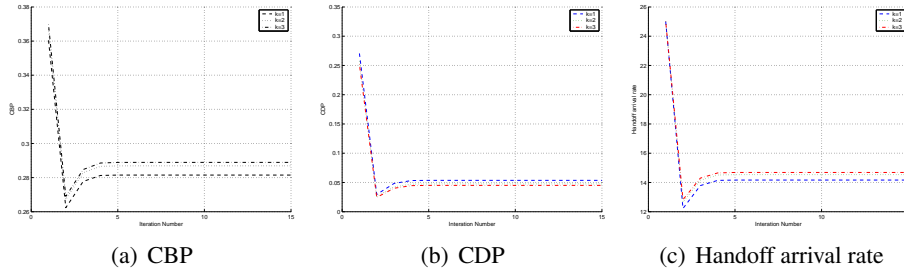


FIGURE 4. The CDP and CBP probabilities with respect to Erlang Index k

queueing for various parameter settings to find the critical parameters of the performance under the assumption that the cell dwell time distribution is an Erlang distribution as a special case.

Figure 3 illustrates the effects of the change of Erlang index on the new-call-blocking and handoff-call-dropping probabilities, depending on queue length with respect to each Erlang index. We set $\lambda_n = 10$, $\lambda_h = 10$, $C = 3$, $M = 0$, $N = 0$, $\mu_h = \mu_n = 5$, and $0 \leq B \leq 5$. We compare the different Erlang distributions with same mean $\frac{1}{\mu}$, but different variances $\frac{1}{(k\mu^2)}$, when guard channel schemes with queue is used. We can see that there is some differences in the blocking and dropping probabilities for different Erlang Indices. We can verify that the results of our analysis are almost equal to the results derived by the event-driven simulation in Fig 3. Here, arrival process in each cell is generated with identical independent distribution. However, it is known as Erlang loss that the steady state probabilities for an $M/G/C/C$ is the same as those of an $M/M/C/C$ with the same arrival process and the channel number [16]. Here, we can also see that there are some differences of the dropping and blocking probabilities for $M/E_k/C/K$ with respect to the Erlang index k .

We consider the effects of both the mobility and traffic types on the network performance for different Erlang Indexes. We set $\lambda_n = 20$, $C = 6$, $M = 2$, $N = 3$, $\mu_h = \mu_n = 5$, $\eta = 2.5$, and $B = 5$. In this example, the handoff arrival rate λ_h is computed by using (3). Figure 4 shows differences of the new-call-blocking probability and the handoff-dropping-probability, with respect to k . λ_h , P_{nb} , and P_{hd} are computed by Algorithm 1. In numerical experiment, we can see that α is closely related to the convergence of the algorithm. When $\alpha = 0$, algorithm does not converge but rather, it oscillates. Thus, in order to prevent the divergence of the algorithm, we use the exponential moving average filter. Figure 4 also shows the convergence of P_{hd} , P_{hb} , and λ_h . There is a trade-off between the convergence and the stability of the algorithm in choosing α .

5. CONCLUSION

We have developed an analytical model for a cellular system that utilizes CAC with guard channel and handoff queueing under the assumption that cell dwell time has a phase-type distribution. We have made CAC performance analysis reflected mobility effect. We have proposed a numerical algorithm to compute the QoS metrics, such as the new-call-blocking probability and the handoff-call forced-terminated probability when the distribution of channel holding times is an Erlang distribution. The complexity of the computation increases exponentially as the dimension of phase increase. By numerical experiment results, we have verified that there should be a tradeoff between the exactness of the performance model and the computational complexity. We have verified the analysis reliance by using event-driven simulation. In future works, the multidimensional mobility should be researched.

REFERENCES

- [1] M. H. Ahmed, "Call admission control in wireless networks : A comprehensive survey," *IEEE Communications Surveys*, vol. 7, no. 1, pp. 50–69, 2005.
- [2] I. Katzela and M. Naghshineh, "Channel assignment schemes for cellular mobile telecommunication systems," *IEEE Personal Communications*, vol. 3, pp. 77–92, 1996.
- [3] D. Hong and S. S. Rappaport, "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures," *IEEE transactions Vehicular Techonology*, vol. 35, no. 3, pp. 38–58, 1986.
- [4] R. A. Guerin, "Queueing-blocking system with two arrival streams and guard channels," *IEEE Transactions Vehicular Techonology*, vol. 36, no. 2, pp. 153–163, 1988.
- [5] E. D. Re, R. Fantacci, and G. Giambene, "Handover queueing strategies with dynamic and fixed channel allocation techniques in low earth orbit mobile satellite systems," *IEEE Transactions on Communications*, vol. 47, no. 1, pp. 89–102, 1999.
- [6] C. H. Yoon and C. K. Un, "Performance of personal portable radio telephone systems with and without guard channels," *IEEE Journal on Selected Areas in Communications*, vol. 11, no. 6, pp. 911–917, 1993.
- [7] K. Lau and S. Maric, "Mobility of queued call requests of a new call queueing technique for cellular systems," *IEEE Transactions on Vehicular Technology*, vol. 47, no. 2, pp. 480–489, 1998.
- [8] J. Hou and Y. Fang, "Mobility-based call admission control schemes for wireless mobile networks," *Wireless Communications and Mobile Computing*, vol. 1, pp. 269–282, 2001.

- [9] P. V. Orlik and S. S. Rappaport, "A model for teletraffic performance and channel holding time characterization in wireless cellular communication with general session and dwell time distributions," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 5, pp. 788–803, 1998.
- [10] Y. Fang, "Hyper-erlang distribution model and its application in wireless mobile networks," *Wireless Networks*, vol. 7, pp. 211–219, 2001.
- [11] Y. Fang, I. Chlamtac, and Y. Lin, "Channel occupancy times and handoff rate for mobile computing and PCS networks," *IEEE Transactions on Computers*, vol. 47, no. 6, pp. 679–692, 1998.
- [12] M. Zonoozi and P. Dassanayake, "User mobility modeling and characterization of mobility patterns," *IEEE Journal on Selected Areas in Communications*, vol. 15, no. 7, pp. 1239–1252, 1997.
- [13] L. P. Seelen, "An algorithm for Ph/Ph/c queues," *European Journal of Operational Research*, vol. 23, pp. 118–127, 1986.
- [14] L. Lipsky, *Queueing Theory: A linear algebraic approach*. Macmillan Publishing, 1992.
- [15] Y. Haung and J. Ho, "Distributed call admission control for a heterogeneous PCS network," *IEEE Transactions on Computers*, vol. 51, no. 12, pp. 1400–1409, 2002.
- [16] Y. Lin and I. Chlamtac, "A model with generalized holding and cell residence times for evaluating handoff rates and channel occupancy times in PCS networks," *International Journal of Wireless Information Networks*, vol. 4, no. 3, pp. 163–171, 1997.