Optimization of Traffic Control in $\text{MMAP}[k]/\text{PH}[k]/S$ Catastrophic Queueing Model with PH Retrial Times and Preemptive Repeat Policy

Raina Raj$^1$, Vidyottama Jain$^1$

Central University of Rajasthan, Ajmer, India

Abstract

The presented study elaborates a multi-server catastrophic retrial queueing model considering preemptive repeat priority policy with phase-type (PH) distributed retrial times. For the sake of comprehension, the scenario of model operation prior and later to the occurrence of the disaster is referred to as the normal scenario and as the catastrophic scenario, respectively. In both scenarios, the arrival and service processes of all types of calls follow marked Markovian arrival process ($\text{MMAP}$) and PH distribution with distinct parameters, respectively. In the normal scenario, the incoming heterogeneous calls are categorized as handoff calls and new calls. An arriving new call will be blocked when all the channels are occupied, and consequently, will join the orbit (virtual space) of infinite capacity. From the orbit, the blocked new call can either retry for the service or exit the system following PH distribution. Whereas, an arriving handoff call is given preemptive repeat priority over a new call in service when all the channels are occupied and at least one of the channel is occupied with a new call otherwise the handoff call is dropped, and consequently, this preempted new call will join the orbit. In the catastrophic scenario, when a disaster causes the shut down of the entire system and failure of all functioning channels, a set of backup channels is quickly deployed to restore services. These failed channels are repaired following PH distribution. The bursty nature of the catastrophe arrival is represented by Markovian arrival process ($\text{MAP}$). Due to the emergency situation in the concerned area, calls to or from the emergency personnel are classified as emergency calls. Therefore, the incoming heterogeneous calls are divided into three categories: handoff calls, new calls, and emergency calls. Emergency calls are provided absolute preemptive priority over handoff calls and new calls due to the pressing need to save lives in such situations. Behaviour of the proposed system is modelled by the level-dependent quasi birth death ($\text{LDQBD}$) process. The Markov chain’s ergodicity criteria are established by demonstrating that it belongs to the class of asymptotically quasi-Toeplitz Markov chains ($\text{AQTM}$). For the approximate computation of the stationary distribution, a new approach is developed. The expressions of various performance measures have been derived for the numerical illustration. An optimization problem to
obtain optimal value of total number of backup channels has been formulated and dealt by employing non dominated sorting genetic algorithm-II (NSGA-II) approach.

**Keywords**— Catastrophe Phenomenon, marked Markovian Arrival Process, NSGA-II, Preemptive Repeat Priority Policy, Phase-Type Distribution, Retrial Queue.

1 Introduction

Over the past decade, there has been discovered a substantial attention in queuing models with disastrous events in communications and cellular networks. These disastrous events are either human-induced (e.g., fire, virus attack, power outages, etc.) or natural (e.g., flood, tsunami, cyclone, etc.), and are typically mentioned as catastrophe events. With the occurrence of a catastrophic event, all active and waiting customers/traffic are compelled to exit the system abruptly, rendering it inoperable. For a comprehensive survey of catastrophic events occurring in communication networks, refer to [9] and references cited therein. For these types of catastrophic queueing models, prioritization of traffic is also important in many cellular network applications e.g., in multiprocessor switching, voice traffic is provided priority over data traffic. The system’s various aspects, such as arrival discipline, service discipline, categories of services, and so on, ensure traffic prioritization. Many priority policies, such as guard channel policy, threshold policy, preemptive priority policy, non-preemptive priority policy, and others, have been proposed in the literature to this end (refer, [4, 8, 22, 25]). In this investigation, preemptive repeat priority policy has been implemented for the higher priority traffic. When all of the channels are occupied, but at least one of these channels is occupied by a lower priority traffic, the arriving higher priority traffic will preempt the service of the ongoing lower priority traffic. This preempted lower priority traffic will then join a virtual space, called orbit, and it will retry for its service from the scratch. The process of recurrent attempts to obtain the service is referred as retrial phenomenon and it’s been extensively explored by the researchers in the area of communications and cellular networks [20].

An overview of research works on the catastrophic queueing models without retrial phenomenon can be found in the articles (refer, [29, 3, 30]). Sudhesh et al. [27] considered a two heterogeneous servers queueing model with catastrophic event, server failure and repair. In their work, the waiting customers were presumed to be impatient and may exit the system at any time, when the system was down. However, the applicability of the above mentioned models has been diminished in the present scenario, since the incoming call arrival followed Poisson process and service times was considered to be exponentially distributed. In contrast to the memory-less property of stationary Poisson flow, the input stream of arrival contains burstiness and correlation properties. Thus, more generalized arrival and services processes are employed, such as Markovian arrival process (MAP), marked Markovian arrival process (MMAP) and phase-type (PH) distribution. Some of the pertinent studies with the consideration of more general processes are as follows. Chakravarthy [2] presented a MAP/PH/1 queueing model considering catastrophic event and delayed action. Recently, Kumar and Gupta [23] studied a discrete-time catastrophic model with population arrival following batch Bernoulli process, binomial catastrophe arrival occurred according to discrete-time renewal process according to which each individual either survived with probability $p$ or died with probability $1 - p$.

Reviews of some of the relevant literature for the catastrophic retrial queueing models are as follows. Wang et
al. [25] proposed a $M/G/1$ retrial queueing model with catastrophe phenomenon. In this study, the inter-retrial time was exponentially distributed and catastrophe arrival occurred according to Poisson process. On the similar track, Chakravarthy et al. [6] presented a $MAP/PH/1$ retrial queueing model with catastrophe phenomenon and repair process. The arrival of catastrophe followed Poisson process and failed channels were repaired following exponential distribution. Their model was a confined one due to the assumption of exponential distribution for retrial and repair processes. Recently, Ammar and Rajadurai [1] proposed a $M/G/1$ preemptive priority retrial queue with working breakdown services and disasters. They assumed when a failure occurred in the system, the major server was sent for the repair and the replacement server worked at a slow rate of service. Though the above mentioned studies had considered generalized arrival and/or service processes, yet retrial process was described through exponential distribution only. In wireless cellular networks, the inter-retrial times are notably brief in comparison to the service times. Since, the retrial attempt is just a matter of pushing one button, these retrial customers will make numerous attempts during any given service interval. Therefore, the consideration of exponential retrial times in place of non-exponential ones could lead to under or over estimating the system parameters as shown by various studies in the literature (refer, [7, 12, 18, 26]). Raj and Jain [19] presented a $MMAP[2]/PH[2]/S$ queueing model with $PH$ distributed retrial times and preemptive repeat priority policy. The proposed a traffic control optimization problem and employed heuristic approaches to obtain the optimal solution. Therefore, to obtain realistic performance measures for retrial phenomenon, a more generalized approach, $PH$ distribution has been applied in the provided model.

In this work, a multi-server catastrophic queueing model with $PH$ distribution for retrial process and preemptive repeat priority policy is introduced. To the best of authors’ knowledge, the proposed model is the first one that deals with such complex system. For the sake of clarity, the scenario of model functioning prior to the occurrence of a disaster (man-made/natural) is referred to as the normal scenario, and after the disaster, as the catastrophic scenario. In the normal scenario, the system will provide services to all incoming calls; however, in the catastrophic scenario, the system will collapse, flushing out all active and waiting calls. $MMAP$ and $PH$ distributions with varying parameters are used to describe the arrival and service processes of all types of incoming calls, respectively. In normal scenario, the incoming calls are classified as handoff call and new call. The new call, which finds all the channels busy upon its arrival will join the orbit (virtual space) of infinite capacity and will be referred as a retrial call ([17]). Following $PH$ distribution, the retrial call can either retry for service or quit the system without receiving it. When all the channels are occupied and at least one of the channels is occupied with the new call, an arriving handoff call is given preemptive priority over it; otherwise, the arriving handoff call is lost from the system. The preempted new call will join the orbit and retry for its service from the scratch.

In the catastrophic scenario, when a calamity strikes, the normally operating system is entirely shut down. Due to the sudden outbreak, all active and retrial calls are forced to terminate their processes and are removed from the system. Since a catastrophic event is considered bursty in nature, $MAP$ is an apt representation of the disaster’s arrival phase. Following $PH$ distribution, the failed channels are repaired promptly. When all of the channels fail in the system, a set of backup/standby channels is instantly installed in the affected area, and services are immediately restored. Calls to and from emergency services, such as hospitals, police, and fire departments, should be given precedence over other public calls in such tragic circumstances. Thus, the incoming heterogeneous calls are now classified as handoff calls, new calls and emergency calls. When all channels are occupied by either handoff calls or new calls, an arriving emergency call is given absolute preemptive priority over either handoff calls or new calls. If all the channels are occupied by both handoff and new calls, an arriving
emergency call will preempt the service of an ongoing new call. The arriving emergency call will start receiving service in place of the preempted handoff or new call. These preempted calls will be lost from the system as there is no concept of orbit in this scenario. The underlying process of the presented system is modelled by level dependent quasi-birth-death (LDQBD) process. The detailed study over LDQBD process can be found in [15] and [24]. Ergodicity conditions of the underlying Markov chain are obtained by proving that the Markov chain satisfies the properties of asymptotically quasi-Toeplitz Markov chains (AQTMC) [21]. A new algorithm is developed for efficient computation of the steady-state distribution. Further, the expressions of various performance measures have been derived for the numerical illustration. Due to the consideration of the preemptive priority policy, the blocking probability for emergency calls decreases and simultaneously the frequent termination of services for handoff and new calls increases the probability of preemption. Thus, an optimization problem to obtain optimal value of total number of backup channels has been formulated and dealt by employing non-dominated sorted genetic algorithm-II (NSGA-II) approach [11].

In this work, a multi-dimensional Markov chain has been constructed by appropriately choosing the components for PH distributed service and retrial times. There are two different ways to keep track of the phases for service and retrial processes. One of them, referred to as the TPFS (track-phase-for-server) in [16], keeps track of the current phase of service at each busy server. The second approach, known as the CSFP (count-server-for-phase) in [16], counts the number of servers having a certain phase of service. The TPFS approach yields a simpler and more transparent form of the blocks of the generator of the multi-dimensional Markov chain. However, if the PH distribution has N phases with n number of busy channels, then there are $N^n$ states of the service process. CSFP approach makes the multi-dimensional Markov chain less transparent. However, the number of states of the process is equal to $T_n = (n + N - 1)! / n!(N - 1)!$. Thus, for $N = 2$ and $n = 20$, the number of states are $1048576$ for TPFS approach and $21$ for the CSFP approach. This approach justifies to use the CSFP method in this work which is helpful in the efficient computation of steady-state distribution.

The layout of this work is arranged in seven sections. In Section 2 a $\text{M\text{M\text{A\text{P}}}[k]/PH[k]/S}$ model with PH distributed retrial times and catastrophe phenomenon is described. In Section 3 the infinitesimal generator matrix for the proposed LDQBD process has been derived. The ergodicity condition of the underlying process is obtained by proving that the Markov chain belongs to the class of AQTMC. An algorithm is proposed to compute steady-state probabilities. In Section 4 formulas of key performance measures to analyse the system efficiency are derived explicitly. Numerical illustrations to point out the impact of various intensities over the system performance are presented in Section 5. An optimization problem has been formulated to evaluate the behaviour of the system in Section 6. Finally, the underlying model is concluded with the insight for the future works in Section 7.

2 Model Description

2.1 Details and Assumptions

The presented study introduce with a $\text{M\text{M\text{A\text{P}}}[k]/PH[k]/S}$ catastrophic model with PH distributed retrial times. Here, S is defined as the total number of channels in the system. The proposed approach can be observed to work in two scenarios: normal and catastrophic. The model’s operation prior to the onset of a disaster can be described as normal, and the latter as catastrophic.
- **Normal Scenario:** In this scenario, the incoming heterogeneous calls are categorized in two classes \((k = 2)\) as handoff calls and new calls. Both types of calls have distinct arrival and service processes that follow the MMAP and PH distributions with different parameters, respectively. If all of the channels are occupied when a new call arrives, the new call will join the orbit (virtual space) of infinite capacity and be referred to as a retrial call \((\text{[17]}\)). The retrial call following PH distribution can either retry for service or exit the system without obtaining the service. It’s been considered that when the number of retrial calls \(l\) is between 0 and \(M\), the retrial rate is \(\theta_l\) and once the number exceed \(M\), the retrial rate is considered as \(\theta\). One of the following two scenarios may occur when a handoff call arrives to the system and all available channels are occupied. The arriving handoff call will be lost from the system if all of the channels are occupied by handoff calls in the first scenario. In the second scenario, among the occupied channels, if at least one of the channels is occupied with a new call, the arriving handoff call will be given preemptive priority over the new call in progress. The handoff call commenced its service in place of the preempted new call and this preempted new call joins the orbit.

- **Catastrophic Scenario:** The presented model is subject to catastrophic events such as power outages, virus assaults, natural disasters, fires, and so on. With the occurrence of such disaster events, all the calls in the system (the one in service and the one waiting for service) leave the system prematurely and the system becomes inactivated. Since a catastrophic event is bursty by nature, MAP is an appropriate representation of the disaster’s arrival phase. The failed channels are immediately repaired following PH distribution. In this study, it is considered that when the whole system is collapsed and all channels are failed, \(K\) backup/standby channels will start providing services at slow rate. These backup channels will stop working once one of the failed channels is fixed. Arriving calls are deemed lost from this point onward until all the channels are fixed. The occurrence of a disaster causes an emergency situation in the affected area. Thus, the incoming heterogeneous calls are now categorized in three classes \((k = 3)\) as handoff calls, new calls and emergency calls. The arrival and service processes of all types of calls follow MMAP and PH distributions with distinct parameters, respectively. To provide emergency services in catastrophic scenario, a general sense of priority should be attached to the emergency calls. Here, it is assumed that the emergency calls are provided preemptive priority over the handoff and new calls. When an emergency call arrives to the system, out of the following three cases, one might occur.

1. When an arriving emergency call finds all the channels occupied with the emergency calls only, the arriving emergency call will be lost form the system.

2. If at the arrival epoch of an emergency call, all the channels are occupied with handoff calls only, the service of a handoff call will be preempted and the emergency call will start receiving service in place of that preempted handoff call.

3. If at the arrival epoch of an emergency call, all the channels are occupied and at least one of the channels is occupied with a new call, the emergency call will preempt the service of that new call and commence service in its place.

It’s been considered that there is no orbit for the blocked or preempted calls when backup channels are working. As, the calls in the orbit need to wait some time before receiving service and it is unreasonable to make calls wait significant time before being admitted when there are urgent needs to save lives or properties.
3 Mathematical Analysis

The underlying process \( \{ \Xi(t), t \geq 0 \} \) for a cell is defined by the following state space:

\[
\Omega = \{ (l, \kappa_1, \kappa_2, j, i, v_1, v_2, s_1^{M_1}, s_2^{M_2}, \ldots, s_1^{M_N}, s_2^{M_N}, \ldots, s_1^{M_E}, s_2^{M_E}, \ldots, v^1, v^2, \ldots, v^R); \\
l \geq 0, \ 0 \leq \kappa_1 \leq S, \ 0 \leq \kappa_2 \leq S, \ 0 \leq j \leq S, \ 0 \leq i \leq K, \ 1 \leq v_1 \leq L_1, \ 1 \leq v_2 \leq L_2 \},
\]

where,

- \( l \) is the number of retrial calls,
- \( \kappa_1 \) is the number of handoff calls in the system receiving service,
- \( \kappa_2 \) is the number of new calls in the system receiving service,
- \( j \) is the number of channels under repair,
- \( i \) is the number of emergency calls in the system receiving service,
- \( v_1 \) is the current phase of MMAP for arrival of calls,
- \( v_2 \) is the current phase of MAP for arrival of catastrophe,
- \( s_1^{m_1} \) be the number of channels for handoff calls which are in phase \( m_1 \); \( s_1^{m_1} = 0, S, m_1 = 1, M_1 \),
- \( s_2^{m_2} \) be the number of channels for new calls which are in phase \( m_2 \); \( s_2^{m_2} = 0, S, m_2 = 1, M_2 \),
- \( s_2^{m_3} \) be the number of channels for emergency calls which are in phase \( m_3 \); \( s_2^{m_3} = 0, K, m_3 = 1, M_E \),
- \( \alpha, \beta, \gamma \) be the average service rates of handoff, new and emergency calls,
- \( \theta \) the average retrial rate of retrial call
- \( D_0, D_1 \) the square matrices of size \( L_2 \) that characterize the MAP
- \( (\alpha, \beta) \) representation of \( \phi \) distribution of new call with dimension \( M_N \)
- \( (\gamma, \Gamma) \) representation of \( \phi \) distribution of retrial call with dimension \( N \)
- \( \mathbb{R} \) row vector
- \( \mathbb{C} \) column vector

All the other assumptions are described in Table 3.

Table 1: Notation
where the infinitesimal generator matrix provided as follows:

\[
Q = \begin{pmatrix}
Q^0 & Q_{0,1} & 0 & 0 & 0 \\
Q^1 & Q_{1,1} & Q_{1,2} & 0 & 0 \\
Q_{2,0} & Q_{2,1} & Q_{2,2} & Q_{2,3} & 0 \\
Q_{3,0} & 0 & Q_{3,2} & Q_{3,3} & Q_{3,4} & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots \\
Q_{M,0} & 0 & 0 & Q_{M,M-1} & Q_{M,M} & \tilde{Q} \\
Q^+ & 0 & 0 & 0 & 0 & \tilde{Q}
\end{pmatrix}
\]

Upper Diagonal:

\[
D_{i,i+1} = \text{diag}(X_i(0), X_i(1), \ldots, X_i(S)) + \text{diag}^+ \{\tilde{X}_i(0), \tilde{X}_i(1), \ldots, \tilde{X}_i(S-1)\}; \ i \geq 0,
\]

\[
X_i(0) = \text{diag}\{X_i(k_1, k_2)\}; k_1 = 0, S, k_2 = 0, S - k_1,
\]

\[
X_i(k_1, k_2) = \begin{pmatrix}
\text{col}(X_i(0,0,0)); j = 0, S, \\
\text{col}(X_i(0,0,0)) + \text{col}(X_i(0,0,S)); \\
\forall S = K \text{ or } K < S \land k_1 + k_2 \leq K,
\end{pmatrix}
\]

\[
\tilde{X}_i(0) = \begin{pmatrix}
\tilde{X}_i(0,0,0); j = 0, S, \\
\tilde{X}_i(0,0,0) + \text{col}(X_i(0,0,S)); \\
\forall S = K \text{ or } K < S \land k_1 + k_2 \leq K,
\end{pmatrix}
\]

\[
\tilde{X}_i(k_1, k_2,0) = \begin{pmatrix}
\tilde{X}_i(k_1,0) \\
\tilde{X}_i(k_1,1) \\
\vdots \\
\tilde{X}_i(k_1,S-k_1)
\end{pmatrix}
\]

\[
\tilde{X}_i(k_1,0) = \begin{pmatrix}
\tilde{X}_i(0,0,0); j = 0, S - 1, \\
\tilde{X}_i(0,0,0) + \text{col}(X_i(0,0,S)); \\
\forall S = K \text{ or } K < S \land k_1 + k_2 > K,
\end{pmatrix}
\]

\[
\tilde{X}_i(k_1,0,0,0) = \begin{pmatrix}
\tilde{X}_i(0,0,0); j = 0, S - 1, \\
\tilde{X}_i(0,0,0) + \text{col}(X_i(0,0,S)); \\
\forall S = K \text{ or } K < S \land k_1 + k_2 > K,
\end{pmatrix}
\]

\[
X_i(k_1, k_2) = X_i(k_1, k_2,0) = X_i(k_1, k_2,0,0); \forall i \geq 1, \ X_i(k_1, k_2) = \tilde{X}_i(k_1, k_2,0) = \tilde{X}_i(k_1, k_2,0,0); \forall i \geq 1,
\]

\[
X_i(k_1, k_2, i) = C_N \otimes I_{T^{k_1} \times T^{k_2}} \otimes P_1(\gamma); i \geq 0, k_1 = 0, S, k_2 = S - k_1, j = i = 0,
\]

\[
\tilde{X}_i(k_1, k_2, i) = C_N \otimes I_{T^{k_1} \times T^{k_2}} \otimes P_{k_2}(\beta_N) \otimes I_{(T^{k_1} \times T^{k_2})}; i \geq 0, k_1 = 0, S - 1, k_2 = S - k_1, j = i = 0,
\]

\[
\tilde{Q} = D_{M,M+1}, \tilde{Q} = D_{2},
\]

\[
X_M(k_1, k_2, j) = C_N \otimes I_{T^{k_1} \times T^{k_2}} \otimes \tilde{P}_M(\gamma); k_1 = 0, S, k_2 = S - k_1, j = i = 0,
\]

\[
\tilde{X}_M(k_1, k_2, j) = C_N \otimes I_{T^{k_1} \times T^{k_2}} \otimes \tilde{P}_M(\gamma); k_1 = 0, S - 1, k_2 = S - k_1, j = i = 0,
\]

where \(P_M(\gamma)\) is \(T^N\) order square matrix.
Lower Diagonal:
\[ L_i = L_{i-1} + L_{i-1}' \quad \forall i \geq 2 \]
\[ Z_i(k_1) = \text{diag} \{ Z_i(k_1), Z_i(k_1, 1), \ldots, Z_i(k_1, S - k_1) \} + \text{diag}^+ \{ Z_i(k_1, 0), \hat{Z}_i(k_1, 1), \ldots, \hat{Z}_i(k_1, S - k_1 - 1) \} | k_1 = 0, S, \]
\[ Z_1(k_1, k_2) = \begin{cases} \text{row}(Z_1(0, 0)); \forall j = 0, S, \\ \text{row}(Z_1(k_1, k_2, 0), Z_1(k_1, k_2, S)); \forall S = K \text{ or } K < S \land k_1 + k_2 \leq K, \\ \forall K < S \land k_1 + k_2 > K, \end{cases} \]
\[ \hat{Z}_1(k_1, k_2) = \begin{cases} \text{row}(\hat{Z}_1(k_1, k_2, 0), \hat{Z}_1(k_1, k_2, S)); \forall S = K \text{ or } K < S \land k_1 + k_2 < K, \\ \forall K < S \land k_1 + k_2 > K, \end{cases} \]
\[ Z_i(k_1, k_2) = Z_i(k_1, k_2, 0) = Z_i(k_1, k_2, 0, 0); \forall i \geq 2, \quad \hat{Z}_i(k_1, k_2) = \hat{Z}_i(k_1, k_2, 0) = \hat{Z}_i(k_1, k_2, 0, 0); \forall i \geq 2, \]
\[ Z_i(k_1, k_2, i) = I_{L_1L_2U_k^N} \otimes L_i^{(1)}(i, \Gamma_1); \forall i \geq 1, k_1 = 0, S, k_2 = 0, S - 1, i = 0, \]
\[ \hat{Z}_i(k_1, k_2, i) = I_{L_1L_2U_k^N} \otimes P_{\beta_k}(\beta_1) \otimes L_i^{(2)}(i, \Gamma_2); \forall i \geq 1, k_1 = 0, S - 1, k_2 = 0, S - k_1 - 1, i = 0. \]
\[ D_2 = D_{M,M-1}, \quad \forall i \geq 2 \]
\[ D_2 = D_{M,M-1}, \quad \forall i \geq 2 \]
\[ \hat{Z}_M(k_1, k_2, i) = I_{L_1L_2U_k^N} \otimes P_{\beta_k}(\beta_1) \otimes L_M^{(2)}(M, \Gamma_2); \forall k_1 = 0, S - 1, k_2 = 0, S - k_1 - 1, i = 0. \]

where \( L_i^{(1)}(M, \Gamma_1), L_i^{(2)}(M, \Gamma_2) \) are \( T_{M,i} \) order square matrices.

Main Diagonal:
\[ M_i = \text{diag} \{ Y_i(0), Y_i(1), \ldots, Y_i(S) \} + \text{diag}^+ \{ \hat{Y}_i(0), \hat{Y}_i(1), \ldots, \hat{Y}_i(S - 1) \} + \text{diag}^+ \{ \tilde{Y}_i(1), \tilde{Y}_i(2), \ldots, \tilde{Y}_i(S) \}; \forall i \geq 0, \]
\[ Y_i(k_1) = \text{diag} \{ Y_i(k_1), Y_i(k_1, 1), \ldots, Y_i(k_1, S - k_1) \} + \text{diag}^+ \{ N_i(k_1, 0), N_i(k_1, 1), \ldots, N_i(k_1, S - k_1 - 1) \}
\]
\[ + \text{diag}^+ \{ S_i^N(k_1, 1), S_i^N(k_1, 2), \ldots, S_i^N(k_1, S - k_1) \}; k_1 = 0, S, \]
\[ Y_0(k_1, k_2) = \begin{cases} \text{diag} \{ Y_0(0, 0, 0), Y_0(0, 0, 1), \ldots, Y_0(0, 0, S) \} + \text{diag}^+ \{ R_0(0, 0, 1), R_0(0, 0, 2), \ldots, R_0(0, 0, S) \}, \\ \text{diag} \{ Y_0(k_1, k_2, 0), Y_0(k_1, k_2, S) \}; \forall S = K \text{ or } K < S \land k_1 + k_2 \leq K, \\ \forall K < S \land k_1 + k_2 > K, \end{cases} \]
\[ Y_0(k_1, k_2, i) = \begin{cases} Y_0(k_1, k_2, i); \forall j = 0, S - 1, \\ \text{diag} \{ Y_0(k_1, k_2, S), \ldots, Y_0(k_1, k_2, S - k_1, S - k_1 - 1) \} + \text{diag}^+ \{ S_0^N(k_1, k_2, S), \ldots, S_0^N(k_1, k_2, S - k_1 - 1) \}, \\ \text{diag} \{ E_0(k_1, k_2, S), \ldots, E_0(k_1, k_2, S - k_1 - 1) \}. \end{cases} \]
\[ N_0(0, 0) = \begin{pmatrix} N_0(0, 0, 0) & N_0(0, 0, 1) & \cdots & N_0(0, 0, S - 1) & 0 \\ 0 & 0 & \cdots & 0 & N_0(0, 0, S) \end{pmatrix}^T, \]
\[ N_0(k_1, k_2) = \begin{cases} \text{diag} \{ N_0(k_1, k_2, 0), N_0(k_1, k_2, S) \}; \forall S = K \text{ or } K < S \land k_1 + k_2 < K, \\ \text{col}(N_0(k_1, k_2, 0), 0); \forall K < S \land k_1 + k_2 = K, \\ N_0(k_1, k_2, 0); \forall K < S \land k_1 + k_2 > K, \end{cases} \]
\[ N_0(\kappa_1, \kappa_2, j; 0) = 0, S - 1, \]
\[ N_0(\kappa_1, \kappa_2, S, 0) \]
\[ \cdots \]
\[ N_0(\kappa_1, \kappa_2, S, K - \kappa_1 - \kappa_2 - 1) \]
\[ 0 \]
\[ S^N_0(\kappa_1, \kappa_2) = \text{diag}\{S^N_0(\kappa_1, \kappa_2, 0), S^N_0(\kappa_1, \kappa_2, S)\}; \forall S = K \text{ or } K < S \& \kappa_1 + \kappa_2 \leq K, \]
\[ \text{row}(S^N_0(\kappa_1, \kappa_2, 0)); \forall K < S \& \kappa_1 + \kappa_2 = K + 1, \]
\[ S^N_0(\kappa_1, \kappa_2, 0); \forall K < S \& \kappa_1 + \kappa_2 > K + 1, \]
\[ S^N_0(\kappa_1, \kappa_2, j; 0) \forall j = 0, S - 1, \]
\[ \cdots \]
\[ S^N_0(\kappa_1, \kappa_2, S, K - \kappa_1 - \kappa_2) \quad S^N_0(\kappa_1, \kappa_2, S, K - \kappa_1 - \kappa_2 + 1) \]
\[ Y_l(\kappa_1, \kappa_2) = Y_l(\kappa_1, \kappa_2, 0) = Y_l(\kappa_1, \kappa_2, 0, 0); l \geq 1, N_l(\kappa_1, \kappa_2) = N_l(\kappa_1, \kappa_2, 0) = N_l(\kappa_1, \kappa_2, 0, 0); l \geq 1, \]
\[ S^N_l(\kappa_1, \kappa_2) = S^N_l(\kappa_1, \kappa_2, 0) = S^N_l(\kappa_1, \kappa_2, 0, 0); l \geq 1, \]
\[ N_l(\kappa_1, \kappa_2, j; i) = \left\{ \begin{array}{l}
C_N \otimes I_{L_1 \times \sigma_1} \otimes P_{\nu}(\beta_N) \otimes I_{T_1}; \kappa_1 = 0, S - 1, \kappa_2 = 0, S - \kappa_1 - 1, j = i = 0, l \geq 0, \\
C_N \otimes I_{L_1 \times \sigma_1} \otimes P_{\nu}(\beta_N) \otimes I_{T_1 T_2}; \kappa_1 = 0, S - 1, \kappa_2 = 0, S - \kappa_1 - 1, j = S, i = 0, K - \kappa_1 - \kappa_2 - 1, l \geq 0, \\
I_{L_1 \times \sigma_1} \otimes L_{S - (\kappa_1 + \kappa_2)}(S - \kappa_1, A_N) \otimes I_{T_1}; \kappa_1 = 0, S - 1, \kappa_2 = 1, S - \kappa_1, j = i = 0, l \geq 0, \\
I_{L_1 \times \sigma_1} \otimes L_{K - (\kappa_1 + \kappa_2 + 1)}(K - \kappa_1 - 1, \hat{A}_N) \otimes I_{T_1 T_2}; \kappa_1 = 0, S - 1, \kappa_2 = 1, S - \kappa_1, j = S, i = 0, K - \kappa_1 - \kappa_2, \\
C \otimes I_{L_1 \times \sigma_1} \otimes P((\beta_\lambda) \otimes I_{T_2}; \kappa_1 = 0, S - 1, \kappa_2 = 1, S - \kappa_1, j = S, i = K - \kappa_1 - \kappa_2. \end{array} \right. \]
\[ E_0(\kappa_1, \kappa_2, S, i) = C_\epsilon \otimes I_{L_1 \times \sigma_1} \otimes P((\beta_\lambda) \otimes I_{T_2}; \kappa_1 = 0, S, \kappa_2 = 0, S - \kappa_1, i = 0, K - \kappa_1 - \kappa_2 - 1, \]
\[ S^F_0(\kappa_1, \kappa_2, S, i) = I_{L_1 \times \sigma_1} \otimes L_{K - (\kappa_1 + \kappa_2 + 1)}(K - \kappa_1 - 1, \hat{A}_N) \otimes I_{T_1}; \kappa_1 = 0, S, \kappa_2 = 0, S - \kappa_1, i = 1, K - \kappa_1 - \kappa_2, \]
\[ R_0(0, 0, j) = R_0(0, 0, j; 0) = I_{L_1 \times L_2} \otimes L_{S - (S, \hat{B})}; j = \overline{1, S}. \]
\[
Y_l(\kappa_1, \kappa_2, i, j) = \begin{cases}
C_0 \oplus D(1) + \Delta; \kappa_1 = \kappa_2 = j = i = \ell = 0, \\
(C_0 + C_N + C_\varepsilon) \oplus D(1) \oplus A_i(S, B) + \Delta; \kappa_1 = \kappa_2 = i = \ell = 0, j = 1, S - 1, \\
C_0 \oplus D(1) \oplus A_{\kappa_2}(S - \kappa_2, A_\varepsilon) \oplus A_{\kappa_2}(S - \kappa_1, A_N) \oplus A_i(K - \kappa_1 - \kappa_2, A_\varepsilon) \oplus A_i(S, B) + \Delta; \\
\kappa_1 = 0, S, \kappa_2 = S - \kappa_1, \ell = 0, j = S, i = 1, K - \kappa_1 - \kappa_2, \\
(C_0 + C_N + C_\varepsilon) \oplus D(1) \oplus A_i(K - \kappa_1 - \kappa_2, A_\varepsilon) \oplus A_i(S, B) + \Delta; \\
\kappa_1 = 0, \kappa_2 = 0, \ell = 0, j = S, i = K - \kappa_1 - \kappa_2, \\
(C_0 + C_N + C_\varepsilon) \oplus D(1) \oplus A_k(S - \kappa_2, A_\varepsilon) \oplus A_{\kappa_2}(S - \kappa_1, A_N) \oplus A_i(K - \kappa_1 - \kappa_2, A_\varepsilon) \\
\oplus A_i(S, B) + \Delta; \kappa_1 = 0, S, \kappa_2 = S - \kappa_1, \ell = 0, j = S, i = 1, K - \kappa_1 - \kappa_2, \\
C_0 \oplus D_0 \oplus A_{\kappa_1}(S - \kappa_2, A_\varepsilon) \oplus A_{\kappa_2}(S - \kappa_1, A_N) \oplus A_i(f, \Gamma) + \Delta; \\
\kappa_1 = 0, S, \kappa_2 = S - \kappa_1 - 1, \ell \geq 0, j = i = 0, \\
(C_0 + C_N) \oplus D_0 \oplus A_{\kappa_1}(S - \kappa_2, A_\varepsilon) \oplus A_{\kappa_2}(S - \kappa_1, A_N) \oplus A_i(f, \Gamma) \\
+ \Gamma_l l^2 T_{N_1} T_{N_2} \otimes L_2^{(2)}(f, \Gamma_l) P_l(\gamma) + \Delta'; \kappa_1 = 0, S, \kappa_2 = S - \kappa_1, \ell \geq 0, j = i = 0,
\end{cases}
\]

\[\Delta = \text{diag}\{I_{l_1 l_2} \otimes \Delta^{(\kappa_1, \kappa_2, \ell)}, i\}, \Delta' = \text{diag}\{I_{l_1 l_2} \otimes \Delta^{(\kappa_1, \kappa_2, \ell)}, i\}' = \Delta^{(\kappa_1, \kappa_2, \ell), i} - \text{diag}\{I_{\bar{m}_1 \bar{m}_2} \otimes T_{\bar{m}_1} T_{\bar{m}_2} \otimes L_2^{(2)}(f, \Gamma_l) P_l(\gamma)\}\]

\[\Delta^{(\kappa_1, \kappa_2, \ell)} = -\text{diag}\{A_{\kappa_1}(S - \kappa_2, A_\varepsilon) \oplus A_{\kappa_2}(S - \kappa_1, A_N) \oplus A_i(K - \kappa_1 - \kappa_2, A_\varepsilon) \oplus A_i(S, B) \oplus A_i(f, \Gamma)\}\]

\[
\hat{Y}_l(\kappa_1, 0) = \begin{pmatrix}
\hat{Y}_l(\kappa_1, 0) \\
\hat{Y}_l(\kappa_1, 1) \\
\vdots \\
\hat{Y}_l(\kappa_1, S - \kappa_1 - 1) \\
0
\end{pmatrix}, \\
\hat{Y}_0(\kappa_1, \kappa_2) = \begin{pmatrix}
0 & \cdots & 0 & 0 \\
0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0
\end{pmatrix}^T
\]

\[
\hat{Y}_0(\kappa_1, 0, 0, \cdots, \hat{Y}_0(0, 0, 0, S - 1, 0) \begin{pmatrix}
0 & \cdots & 0 & 0 \\
0 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0
\end{pmatrix}^T
\]

\[
\text{diag}\{\hat{Y}_0(\kappa_1, 0, 0), \hat{Y}_0(\kappa_1, 0, S)\}; \\
\forall S = K \text{ or } K < S \& \kappa_1 + \kappa_2 < K,
\]

\[
\text{col}\{\hat{Y}_0(\kappa_1, 0, 0), \hat{Y}_0(\kappa_1, 0, S)\}; \\
\forall S = K \text{ or } K < S \& \kappa_1 + \kappa_2 = K,
\]

\[
\hat{Y}_0(\kappa_1, \kappa_2, 0) = \begin{pmatrix}
\hat{Y}_0(\kappa_1, 0, 0) \\
\hat{Y}_0(\kappa_1, \kappa_2, S, 0) \\
\vdots \\
\hat{Y}_0(\kappa_1, \kappa_2, S, K - \kappa_1 - \kappa_2 - 1) \\
0
\end{pmatrix}, \\
\hat{Y}_l(\kappa_1, \kappa_2) = \hat{Y}_l(\kappa_1, \kappa_2, 0) = \hat{Y}_l(\kappa_1, \kappa_2, 0, 0); \ell \geq 1,
\]

\[
\hat{Y}_l(\kappa_1, \kappa_2, i, j) = \begin{cases}
C_\varepsilon \otimes I_{l_2} \otimes P_{\kappa_1}(\beta_{\kappa_1}) \otimes I_{\bar{m}_1 \bar{m}_2} \otimes T_{\bar{m}_1} T_{\bar{m}_2}; \kappa_1 = 0, S - 1, \kappa_2 = 0, S - \kappa_1 - 1, j = i = 0, \ell \geq 1, \\
C_\varepsilon \otimes I_{l_2} \otimes P_{\kappa_1}(\beta_{\kappa_1}) \otimes I_{\bar{m}_1 \bar{m}_2} \otimes T_{\bar{m}_1} T_{\bar{m}_2}; \kappa_1 = 0, S - 1, \kappa_2 = 0, S - \kappa_1 - 1, i = 0, \kappa_1 - \kappa_2, \ell = 0,
\end{cases}
\]
\[
\begin{align*}
\tilde{Y}_0(\kappa_1, \kappa_2) &= \begin{pmatrix} 
\tilde{Y}_0(1, 0, 0) & \cdots & \tilde{Y}_0(1, 0, S - 1) & 0 \\
0 & \cdots & 0 & \tilde{Y}_0(1, 0, S)
\end{pmatrix}, \\
\tilde{Y}_0(\kappa_1, \kappa_2) &= \begin{pmatrix} 
\tilde{Y}_0(\kappa_1, \kappa_2, 0), \tilde{Y}_0(\kappa_1, \kappa_2, S)\end{pmatrix}; \forall S = K \text{ or } K < S \& \kappa_1 + \kappa_2 \leq K, \\
\text{row}(\tilde{Y}_0(\kappa_1, \kappa_2, 0), \tilde{Y}_0(\kappa_1, \kappa_2, S)); \forall S = K \text{ or } K < S \& \kappa_1 + \kappa_2 = K + 1, \\
\tilde{Y}_0(\kappa_1, \kappa_2, 0); \forall K < S \& \kappa_1 + \kappa_2 > K + 1,
\end{align*}
\]
\[
\tilde{Y}_0(\kappa_1, \kappa_2, j) = \begin{pmatrix} 
\tilde{Y}_0(\kappa_1, \kappa_2, 0), \tilde{Y}_0(\kappa_1, \kappa_2, S) \\
\vdots \\
\tilde{Y}_0(\kappa_1, \kappa_2, S - 1), \tilde{Y}_0(\kappa_1, \kappa_2, S)
\end{pmatrix} : \kappa_1 \leq K, \kappa_2 = 0,
\]
\[
\tilde{Y}_i(\kappa_1, \kappa_2) = \tilde{Y}_i(\kappa_1, \kappa_2, 0) = \tilde{Y}_i(\kappa_1, \kappa_2, 0, 0); \forall i \geq 1,
\]
\[
\tilde{Y}_i(\kappa_1, \kappa_2, 0) = I_{L_0, L_T} \otimes L_{S-\kappa_1-1} (S - \kappa_2, \tilde{A}_N) \otimes I_{L_{N-T}L_T}; \forall \kappa_1 = 1, S, \kappa_2 = 0, S - \kappa_1, i \geq 0, \\
\tilde{Y}_i(\kappa_1, \kappa_2, S, i) = I_{L_0, L_T} \otimes L_{K-\kappa_1-1} (K - \kappa_2 - i, \tilde{A}_N) \otimes I_{L_{N-T}L_T}; \forall \kappa_1 = 1, S, \kappa_2 = 0, S - \kappa_1, i = 0, K - \kappa_1 - \kappa_2, \\
\tilde{Y}_i(\kappa_1, 0, S, K - \kappa_1 - \kappa_2) = C \otimes I_{N_1 \times N_2} \otimes P(\beta_E) \otimes I_{N_2}; \forall \kappa_1 = 1, S, i = K - \kappa_1.
\]

First column:
\[
L_{i, 0} = \begin{pmatrix} 
W_i(0) & 0 & \cdots & 0 \\
W_i(1) & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
W_i(S) & 0 & \cdots & 0
\end{pmatrix} ; i \geq 0, \quad W_i(\kappa_1) = \begin{pmatrix} 
W_i(\kappa_1, 0) & 0 & \cdots & 0 \\
W_i(\kappa_1, 1) & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
W_i(\kappa_1, S - \kappa_1) & 0 & \cdots & 0
\end{pmatrix} ; \kappa_1 = 0, S,
\]

where \(L_{i, 0}\) and \(W_i(\kappa_1)\) are square matrices of order \(S + 1\) and \(S - \kappa_1 + 1\), respectively and \(Q^+ = L_{M, 0}\),
\[
W_0(0, 0) = \text{diag}\{W_0(0, 0, 0), W_0(0, 0, 1), \ldots, W_0(0, 0, S)\},
\]
\[
W_0(\kappa_1, \kappa_2) = \begin{pmatrix} 
W_0(\kappa_1, \kappa_2, 0) & W_0(\kappa_1, \kappa_2, 1) & \cdots & W_0(\kappa_1, \kappa_2, S - 1) \\
0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0
\end{pmatrix} ; \forall S = K \text{ or } K < S \& \kappa_1 + \kappa_2 \leq K,
\]
\[
\text{row}(W_0(\kappa_1, \kappa_2, 0), W_0(\kappa_1, \kappa_2, 1), \ldots, W_0(\kappa_1, \kappa_2, S)); \forall K < S \& \kappa_1 + \kappa_2 > K,
\]
\[
W_0(\kappa_1, \kappa_2, 0); \forall j = 0, S - 1,
\]
\[
\text{diag}\{W_0(0, 0, S, 0), W_0(0, 0, S, 1), \ldots, W_0(0, 0, S, K)\}
\]
\[
W_0(\kappa_1, \kappa_2, j) = \begin{pmatrix} 
W_0(\kappa_1, \kappa_2, S, 0) \\
\vdots \\
W_0(\kappa_1, \kappa_2, S, K - \kappa_1 - \kappa_2 - 1)
\end{pmatrix} ; \forall \kappa_1 + \kappa_2 < K,
\]
\[
\text{row}(W_0(\kappa_1, \kappa_2, S, 0), W_0(\kappa_1, \kappa_2, S, 1), \ldots, W_0(\kappa_1, \kappa_2, S, K - \kappa_1 - \kappa_2)); \forall \kappa_1 + \kappa_2 \geq K,
\]
\[ W_t(\kappa_1, \kappa_2, j, 0) = I_{L_t} \otimes D_t \otimes \epsilon_{F}\mathbb{M}_{I_{T_M}}T_{T_M}\mathbb{M}_2 T_{T_M}\mathbb{M}_2 \otimes \Pi P(\alpha) \otimes \epsilon_{Z_N} ; \forall \kappa_1 = 0, \kappa_2 = 0, S - \kappa_1, j = 1, S \& j = \kappa_1 + \kappa_2, t \geq 0. \]

### 3.1 Ergodicity Condition

The modelled process \( \{\Xi(t), t \geq 0\} \) clearly has the traits of a level-dependent quasi-birth-death process, as evidenced by its structure. But, the existing definition of LQBD does not provide any results about the limiting/asymptotic behaviour of process when the countable element of Markov chain tends to infinity. Thus, the existing result of stability of constructive conditions of stability of LQBD process. Though, the behaviour of the proposed process satisfies the conditions imposed on the limiting behaviour of asymptotic quasi-toeplitz Markov chain (AQTM). In this order, we will compute matrices \( U_0, U_1 \) and \( U_2 \) defined as follows

\[
U_0 = \lim_{t \to \infty} T^{-1}_t D_{t, t-1}, \quad U_1 = \lim_{t \to \infty} T^{-1}_t D_{t, t}, \quad U_2 = \lim_{t \to \infty} T^{-1}_t D_{t, t+1},
\]

where \( T^{-1}_t \) is the diagonal matrix with diagonal entries defined as the modulus of the diagonal entries of the matrix \( D_{t, t}, t \geq 0 \). In the considered case, the matrices \( U_0 \), \( U_1 \) and \( U_2 \) have the following form.

\[
U_0 = \lim_{t \to \infty} T^{-1}_t D_{t, t-1} = T^{-1}_t D_0, \quad t > M,
\]

\[
U_1 = \lim_{t \to \infty} T^{-1}_t D_{t, t} + I = T^{-1}_t D_1, \quad t > M,
\]

\[
U_2 = \lim_{t \to \infty} T^{-1}_t D_{t, t+1} = T^{-1}_t D_2, \quad t > M,
\]

where \( T = \text{diag}\{T(0), T(1), \ldots, T(S)\}\); \( T(\kappa_1) = \text{diag}\{T(\kappa_1, 0), T(\kappa_1, 1), \ldots, T(\kappa_1, S - \kappa_1)\}\),

\[
T(\kappa_1, \kappa_2) = \begin{cases} 
A_0 \oplus \sum \oplus A_{\kappa_1} (S - \kappa_2, A_N) \oplus A_{\kappa_2} (S - \kappa_1, A_N) \oplus A_M (M, \gamma) ; & \kappa_1 = 0, S - \kappa_2 = 0, S - \kappa_1 - 1. \\
A_0 + C_N \oplus \sum \oplus A_{\kappa_1} (S - \kappa_2, A_N) \oplus A_{\kappa_2} (S - \kappa_1, A_N) \oplus A_M (M, \gamma) + I_{L_1 L_2 T_{\mathbb{M}_{I_{T_M}} T_{T_M}} \mathbb{M}_2} \otimes L^{(2)}_M (M, \gamma_2) \oplus P_M (\gamma) ; & \kappa_1 = 0, S, S - \kappa_2 = S - \kappa_1.
\end{cases}
\]

Here, \( A_0 = -C_0; \sum = -D_0 + D_1. \) Note that, in the following illustration \( \text{diag}, \text{diag}^+ \) and \( \text{diag}^- \) represents matrix with main diagonal, upper diagonal and lower diagonal, respectively.

**Theorem 3.1.** The necessary and sufficient condition for the ergodicity of the underlying process is the satisfaction of the inequality

\[
\lambda < \sum_{\kappa_1 = 0}^{S} \sum_{\kappa_2 = 0}^{S - \kappa_1} x^{(1)}_M \left( I^{(1)}_0 \right)^t (M, \gamma_1) e + \sum_{\kappa_1 = 0}^{S} \sum_{\kappa_2 = 0}^{S - \kappa_1} x^{(2)}_M \left( I^{(2)}_0 \right)^t (M, \gamma_2) e
\]

where \( x^{(1)}_M = x_M(\kappa_1, \kappa_2) (e_{L_2 T_{\mathbb{M}_{I_{T_M}} T_{T_M}} \mathbb{M}_2}) \) and \( x^{(2)}_M = x_M(\kappa_1, \kappa_2) (e_{L_2 T_{\mathbb{M}_{I_{T_M}} T_{T_M}} \mathbb{M}_2}) \) and \( x \) is the unique solution of \( x(D_0 + D_2 + D_1) = 0; x \in 1. \)

**Proof.** Following [21], a necessary and sufficient condition for ergodicity of the underlying process can be formulated in terms of the generator \( \mathcal{D} \) as follows

\[
x \mathcal{D} e < x \mathcal{D}_0 e
\]

(1)
where vector \( x \) is the unique solution to the system of linear equations

\[
x(\mathcal{D}_0 + \mathcal{D}_1 + \mathcal{D}_2) = 0; \quad xe = 1.
\] (2)

Let expression of \( x \) be of the form \( x = (\pi \otimes x_M(\kappa_1, 0), \pi \otimes x_M(\kappa_1, 1), \ldots, \pi \otimes x_M(\kappa_1, S - \kappa_1)) \), where \( x_M(\kappa_1, \kappa_2) \) vector is of size \( L_2^{T_{\kappa_1}^M} T_{\kappa_2}^N T_M^N \). Substituting the expression of \( x \) and block matrices in inequality \( \Pi \)

\[
\begin{align*}
    &\sum_{\kappa_1=0}^{S} (\pi \otimes x_M(\kappa_1, S - \kappa_1)) (C_N \otimes I_{L_2^{T_{\kappa_1}^M} T_{\kappa_2}^N} \otimes P_M(\gamma))e \\
    &+ \sum_{\kappa_1=1}^{S} (\pi \otimes x_M(\kappa_1, S - \kappa_1)) (C_N \otimes I_{L_2} \otimes P_{\kappa_1}(\beta_N) \otimes I_{T_{\kappa_2}^M} \otimes P_M(\gamma))e \\
    &< \sum_{\kappa_1=0}^{S} \sum_{k_2=0}^{S - \kappa_1} (\pi \otimes x_M(\kappa_1, \kappa_2)) (I_{L_1 L_2} T_{\kappa_2}^M T_{\kappa_2}^N \otimes L_0^{(1)}(M, \tilde{\Gamma}_1))e \\
    &+ \sum_{\kappa_1=0}^{S} \sum_{k_2=1}^{S - \kappa_1} (\pi \otimes x_M(\kappa_1, \kappa_2)) (I_{L_1 L_2} T_{\kappa_2}^M T_{\kappa_2}^N \otimes P_{\kappa_2}(\beta_N) \otimes L_0^{(2)}(M, \tilde{\Gamma}_2))e.
\end{align*}
\]

using the relations \( \pi C_N \lambda = \lambda \pi N, \lambda \pi N + \lambda \pi e = \lambda, \tilde{\Gamma} M(\gamma) e = e, P_{\kappa_1}(\beta_N) e = e, P_{\kappa_2}(\beta_N) e = e \), where \( e \) is a column vector with one of appropriate size, the inequality will be reduced to

\[
\lambda < \sum_{\kappa_1=0}^{S} \sum_{k_2=0}^{S - \kappa_1} x_M^{(1)}(1) L_0^{(1)}(M, \tilde{\Gamma}_1)e + \sum_{\kappa_1=0}^{S} \sum_{k_2=1}^{S - \kappa_1} x_M^{(2)}(2) L_0^{(2)}(M, \tilde{\Gamma}_2)e.
\]

Since the stability of the system cannot be defined in the catastrophic scenario, thus these conditions are derived for the normal scenario. System of equations \( \Pi \) has unique solution because the matrix of the system is an infinitesimal generator of the underlying process which defines joint distributions of the number of retrial calls, number of handoff calls receiving service and number of new calls receiving service. The left hand side of inequality \( \Pi \) is the total arrival rate of handoff calls and new calls in the system. In the right hand side, the first summand is the rate of departure from the system and second summand is the rate of starting the service for retrial calls when the retrial is successful. It is intuitively clear that the Markov chain describing queuing model under study is ergodic if and only if the total arrival rate is less than the maximum value of the total departure rate and successful retrial rate. When the number of retrial calls increases without bound and the retrial rate tends to infinity, the retrial queueing model approaches to the corresponding classical queueing model for which \( \lambda/S \mu \) becomes a necessary and sufficient condition for the stability.

\[\square\]

### 3.2 Stationary Distribution

Let \( z_\kappa = \{z_\kappa(0), z_\kappa(1), z_\kappa(2), \ldots, z_\kappa(M - 1), z_\kappa(M), \ldots\} \) be the steady-state probability vector of generator matrix \( \mathcal{D} \) satisfying \( z_\kappa \mathcal{D} = 0; \ q_\kappa e = 1 \). Here, element \( z_\kappa(0) \) contains

\[
1 \times \left( \sum_{\kappa_1=0}^{S} \sum_{\kappa_2=0}^{S - \kappa_1} L_1 L_2 T_{\kappa_1}^M T_{\kappa_2}^N \right) \left( \sum_{\kappa_1=0}^{S} \sum_{\kappa_2=0}^{S - \kappa_1} L_1 L_2 T_{\kappa_1}^M T_{\kappa_2}^N T_{\kappa_1}^R \right) \text{vector component}\)

and \( z_\kappa(\ell) \) contains

\[
1 \left( \sum_{\kappa_1=0}^{S} \sum_{\kappa_2=0}^{S - \kappa_1} L_1 L_2 T_{\kappa_1}^M T_{\kappa_2}^N \right) \text{elements}; \ \ell \geq 0, 0 \leq \kappa_1 \leq S, 0 \leq \kappa_2 \leq S. \text{ The derived structure of generator matrix lacks of the existing quasi-birth death structure and toeplitz like structure. Therefore, the existing approach for computing the stationary distribution for AQIMC can be employed because the Markov chain has a specific asymptotic behaviour. The originally proposed algorithm for AQIMC process was modified by Dudin and}
Dudina [14] later on. In their proposed approach, they tackled the challenges of larger order matrix computation and storage. Along with the substantial advantages, the proposed algorithm also has some following disadvantage.

- In their algorithm, the initial value for starting the computation process has been chosen randomly. They did not provide any reasoning on how to select the initial point \( i_0 \).
- After each unsuccessful iteration, to obtain a new searching interval, a randomly selected value \( s \) has been added in the existing interval. There is no reasoning behind the chosen value \( s \) also.
- While checking the termination criteria for matrix \( F \), they discarded some portion of the search interval at each failed iteration of Step 4.2 (case 3). They did not provided any clarification for the eliminated portion of the search interval.

These findings motivated for the development of a modified approach for computing the stationary distribution of the Markov chain under consideration.

### 3.2.1 Old Algorithm

**Step 1.** Set \( i_0, s \) randomly. Fix \( \epsilon_g \) and \( \epsilon_f \) as accuracy levels of matrices \( G_i \) and steady-state vector \( z_s(i) \), respectively.

**Step 2.1.** Set \( G^{(1)}_\kappa = O \) and \( G^{(2)}_\kappa = I \), \( \kappa = i_0 - 1 + s, \; i_f = -1 \).

**Step 2.2.** Compute the matrices as defined

\[
G^{(1)}_\kappa := -((\mathcal{D}_{\kappa+1,\kappa+1} + \mathcal{D}_{\kappa+1,\kappa+2}G^{(1)}_{\kappa+1}))^{-1}\mathcal{D}_{\kappa+1,\kappa},
\]

\[
G^{(2)}_\kappa := -((\mathcal{D}_{\kappa+1,\kappa+1} + \mathcal{D}_{\kappa+1,\kappa+2}G^{(2)}_{\kappa+1}))^{-1}\mathcal{D}_{\kappa+1,\kappa}.
\]

**Step 2.3.** Calculate \( \|G^{(1)}_\kappa - G^{(2)}_\kappa\| \). There can be three possible cases.

- **Case 1:** If \( \|G^{(1)}_\kappa - G^{(2)}_\kappa\| < \epsilon_g \) Go to step 3.
- **Case 2:** If \( \|G^{(1)}_\kappa - G^{(2)}_\kappa\| > \epsilon_g \) and \( \kappa \geq i_0 \), Set \( \kappa := \kappa - 1 \) and repeat step 2.2.
- **Case 3:** If \( \|G^{(1)}_\kappa - G^{(2)}_\kappa\| > \epsilon_g \) and \( \kappa = i_0 - 1 \), Set \( s := 2s \), \( \kappa = i_0 - 1 + 2s \), \( i_0 = \kappa \), and repeat step 2.1.

**Step 3.** Set \( G_\kappa = G^{(1)}_\kappa \). Compute \( G_i, i = (i_f + 1, \, \kappa - 1) \). Set \( B = Q_{\kappa,0} \) and compute \( B = Q_{0,0} + G_i B, \; i = (i_f + 1, \, \kappa - 1) \), and store.

**Step 4.1.** Set \( i = i_f + 1 \). If \( i = 0 \), find solution of \( z_s(0) B = 0; z_s(0) e = 1 \). Otherwise start \( i = 1 \) and compute \( z_s(i) = -z_s(i - 1)\mathcal{D}_{i-1,i}(\mathcal{D}_{i,i} + \mathcal{D}_{i,i+1}G_i)^{-1} \).

**Step 4.2.** Compute \( \|z_s(i)\| \). There can be three possible cases

- **Case 1:** If \( \|z_s(i)\| < \epsilon_f \), set \( i^* = i \) and go to Step 5.
- **Case 2:** If \( \|z_s(i)\| > \epsilon_f \), and \( i < \kappa \), increase \( i \) by one and go to Step 4.1.
- **Case 3:** If \( \|z_s(i)\| > \epsilon_f \), and \( i = \kappa \), set \( i_0 = \kappa + s, \; \kappa = i_0 - 1 + 2s \) and \( i_f = \kappa \) and go to Step 2.2.

**Step 5.** Calculate vectors \( z_s(\kappa), k = 0, \kappa^* \) as \( z_s(\kappa + 1) = c z_s(\kappa) \), where \( c = \frac{1}{z_s(0) + z_s(1) + \ldots + z_s(\kappa^*)} \) is a normalizing constant.

### 3.2.2 New Algorithm

**Step 1.** To fix the initial value \( i_0 \) (anticipated number of matrices \( G_i \), which is required for the computation of the stationary distribution), convert the original model into Poisson-exponentially distributed model, i.e., arrival of calls and catastrophe are defined by Poisson process and rest of the processes follow exponential distribution. For this model, fix a pre-defined small value \( \delta > 0 \) such that \( z_s(i_0) < \delta \). Here, \( z_s \) is the stationary distribution of
the system and \( z_s(i_0) \) is \( i_0 \)th component of stationary distribution.

**Explanation:** The termination criteria or the condition \( z_s(i_0) < \delta \) represents that for \( i \geq i_0 \), the steady-state vector is showing invariant behaviour for a pre-defined small value. This \( i_0 \) will work as an initial value for the original model. In the original model also, the value of \( i_0 \) represents the invariant behaviour of matrix \( G_i \) for \( i \geq i_0 \) which ultimately determines the invariant behaviour of steady-state vector.

**Step 2.** Fix \( s \) as some component of \( i_0 \), i.e., \( s = mi_0 \). (After many trials in the programming, it’s been observed that \( s = 2i_0 \) is appropriate choice for \( s \), as the solution converges fast.)

**Explanation:** If the Markov chain is ergodic, for any value of \( i_0 \), and for any fixed small arbitrary value \( \epsilon \) there exist a finite number \( s \), the probability that the Markov chain will transit from \( i \) to \( i + s \) without visiting state \( i - 1 \) is less than \( \epsilon \).

**Step 3.** Fix \( \epsilon_g \), and \( \epsilon_f \) as accuracy levels of matrices \( G_i \) and steady-state vector \( z_s(i) \), respectively.

**Step 4.1.** Set \( G^{(1)} = O \) and \( G^{(2)} = I \), \( \kappa = i_0 - 1 + s \), \( \epsilon_f = -1 \).

**Step 4.2.** Compute the matrices as defined

\[
G^{(1)}_\kappa := -(Q_{\kappa + 1, \kappa + 1} + Q_{\kappa + 1, \kappa + 2}G^{(1)}_{\kappa + 1})^{-1}Q_{\kappa + 1, \kappa},
\]

\[
G^{(2)}_\kappa := -(Q_{\kappa + 1, \kappa + 1} + Q_{\kappa + 1, \kappa + 2}G^{(2)}_{\kappa + 1})^{-1}Q_{\kappa + 1, \kappa}.
\]

**Step 4.3.** Calculate \( ||G^{(1)}_\kappa - G^{(2)}_\kappa|| \). There can be three possible cases.

- **case 1:** If \( ||G^{(1)}_\kappa - G^{(2)}_\kappa|| < \epsilon_g \), go to step 5.
- **case 2:** If \( ||G^{(1)}_\kappa - G^{(2)}_\kappa|| > \epsilon_g \) and \( \kappa \geq i_0 \). Set \( \kappa := \kappa - 1 \) and repeat step 4.2.
- **case 3:** If \( ||G^{(1)}_\kappa - G^{(2)}_\kappa|| > \epsilon_g \) and \( \kappa = i_0 - 1 \). Set \( s := 2s, \kappa = i_0 - 1 + 2s, i_0 = \kappa \), and repeat step 4.1.

**Step 5.** Set \( G_\kappa = G^{(1)}_\kappa \). Compute \( G_i, i = (\epsilon_f + 1, \kappa - 1) \). Set \( B = Q_{\kappa, 0} + G, B, i = (i_0 + 1, \kappa - 1) \), and store.

**Step 6.1.** Set \( i = \epsilon_f + 1 \). If \( i = 0 \), find solution of \( z_s(0)B = 0; z_s(0)e = 1 \). Otherwise start \( i = 1 \) and compute \( z_s(i) = -z_s(i-1)Q_{i-1, i} + Q_{i, i+1}G_i^{-1} \).

**Step 6.2.** Compute \( ||z_s(i)|| \). There can be three possible cases.

- **case 1:** If \( ||z_s(i)|| < \epsilon_f \), set \( i^* = i \) and go to step 7.
- **case 2:** If \( ||z_s(i)|| > \epsilon_f \), and \( i < \kappa \), increase \( i \) by one and go to step 6.1.
- **case 3:** If \( ||z_s(i)|| > \epsilon_f \), and \( i = \kappa \), set \( i_0 = \kappa + 1, \kappa = i_0 - 1 + 2s \) and \( i_f = \kappa \) and go to step 4.2.

**Step 7.** Calculate vectors \( z_s(\kappa), k = 0, i^* \) as \( z_s(\kappa + 1) = cz_s(\kappa) \), where \( c = \frac{1}{\sum_{j=0}^{i^*} z_s(j + 1) + \sum_{j=1}^{i_f} c_z(j + 1)} \) is a normalizing constant.

**Advantages:** The main advantage of the proposed algorithm is that the initial point will not be selected at random. Rather than starting the algorithm from the 0 or some random point, this algorithm starts the iteration with a logical initial point which saves less computation and time of matrices. The second point is that no calculation of \( G \) matrix is repeated. All the matrices calculated in previous iteration are stored and used in the next iteration. In this algorithm, a sequential search is done considering a search interval. No point has been discarded unlike the previous algorithm.

### 4 Performance Measures

The following relevant performance measures for the proposed system are calculated, after computing the steady-state distribution \( z_s \).
1. The probability that there are \( l \) number of retrial calls:

\[
P_{\text{orbit}}(l) = \sum_{\kappa_1=0}^{S} \sum_{\kappa_2=0}^{S-\kappa_1} \sum_{i=0}^{\infty} \left( \sum_{\kappa=0}^{S-\kappa_1} \right) P_{l,i}(\kappa_1, \kappa_2, 0, 0) e.
\]

2. Expected number of retrial calls:

\[
E_{\text{orbit}} = \sum_{l=0}^{\infty} l P_{\text{orbit}}(l) e.
\]

3. The probability that \( \kappa_1 \) number of handoff calls are receiving service:

\[
P_{\mathcal{E}}(\kappa_1) = \sum_{j=1}^{S-1} z_s(0, 0, 0, 0) e + \sum_{\kappa_1=0}^{S-\kappa_1} \sum_{\kappa_2=0}^{S-\kappa_1} z_s(0, \kappa_1, \kappa_2, 0, 0) e + \sum_{\kappa_1=0}^{S-\kappa_1} \sum_{\kappa_2=0}^{S-\kappa_1} \sum_{i=0}^{\infty} \left( \sum_{\kappa=0}^{S-\kappa_1} \right) P_{l,i}(\kappa_1, \kappa_2, 0, 0) e.
\]

4. The probability that \( \kappa_2 \) number of new calls are receiving service:

\[
P_{\mathcal{N}}(\kappa_2) = \sum_{j=1}^{S-1} z_s(0, 0, 0, 0) e + \sum_{\kappa_1=0}^{S-\kappa_1} \sum_{\kappa_2=0}^{S-\kappa_1} z_s(0, \kappa_1, \kappa_2, 0, 0) e + \sum_{\kappa_1=0}^{S-\kappa_1} \sum_{\kappa_2=0}^{S-\kappa_1} \sum_{i=0}^{\infty} \left( \sum_{\kappa=0}^{S-\kappa_1} \right) P_{l,i}(\kappa_1, \kappa_2, 0, 0) e.
\]

5. The probability that \( i \) number of emergency calls are receiving service:

\[
P_{\mathcal{I}}(i) = \sum_{j=1}^{S-1} z_s(0, 0, 0, 0) e + \sum_{\kappa_1=0}^{S-\kappa_1} \sum_{\kappa_2=0}^{S-\kappa_1} z_s(0, \kappa_1, \kappa_2, 0, 0) e + \sum_{\kappa_1=0}^{S-\kappa_1} \sum_{\kappa_2=0}^{S-\kappa_1} \sum_{i=0}^{\infty} \left( \sum_{\kappa=0}^{S-\kappa_1} \right) P_{l,i}(\kappa_1, \kappa_2, 0, 0) e.
\]

6. The probability that the system is under repair:

\[
P_{\mathcal{R}} = \sum_{i=0}^{\infty} z_s(0, 0, 0, 0) e + \sum_{\kappa_1=0}^{S-\kappa_1} \sum_{\kappa_2=0}^{S-\kappa_1} z_s(0, \kappa_1, \kappa_2, 0, 0) e + \sum_{\kappa_1=0}^{S-\kappa_1} \sum_{\kappa_2=0}^{S-\kappa_1} \sum_{i=0}^{\infty} \left( \sum_{\kappa=0}^{S-\kappa_1} \right) P_{l,i}(\kappa_1, \kappa_2, 0, 0) e.
\]

7. The dropping probability of a handoff call:

- in normal scenario:

\[
P_{\mathcal{D}} = \frac{1}{\lambda_e} \left( \sum_{i=0}^{\infty} z_s(0, 0, 0, 0) (C_{\mathcal{N}} \otimes I_{L_2 T_{\kappa_1} M_{\mathcal{T}} T_{\mathcal{N}}}) e \right).
\]

- in catastrophic scenario:

\[
P_{\mathcal{D}}^c = \frac{1}{\lambda_e} \left( \sum_{i=0}^{\infty} z_s(0, 0, 0, 0) (C_{\mathcal{N}} \otimes I_{L_2 T_{\kappa_1} M_{\mathcal{T}} T_{\mathcal{N}}}) e \right).
\]

8. The blocking probability of a new call in catastrophic scenario:

\[
P_{\mathcal{B}}^c = \frac{1}{\lambda_e} \left( \sum_{i=0}^{\infty} z_s(0, 0, 0, 0) (C_{\mathcal{N}} \otimes I_{L_2 T_{\kappa_1} M_{\mathcal{T}} T_{\mathcal{N}}}) e \right).
\]

9. The blocking probability of an emergency call:

\[
P_{\mathcal{E}} = \frac{1}{\lambda_e} (z_s(0, 0, 0, 0) (C_{\mathcal{E}} \otimes I_{L_2 T_{\kappa_1} M_{\mathcal{T}} T_{\mathcal{E}}}) e).
\]
10. Rate of losses due to the occurrence of catastrophe:
\[
\alpha_f = \delta \sum_{l=0}^{\infty} \sum_{\kappa_1=0}^{S} \sum_{\kappa_2=0}^{S-\kappa_1} z_s(l, \kappa_1, \kappa_2, 0, 0)(D_l \otimes I_{\xi_1} T_{N}^{M_{N}} T_{l-1}^{T_{N}})e.
\]

11. The probability that an arriving handoff call preempts the service of an ongoing new call in ordinary scenario:
\[
P_{\text{preempt}}^{\text{new}} = \frac{1}{\lambda_{\text{H}}^}\sum_{l=0}^{\infty} \sum_{\kappa_1=0}^{S-1} z_s(l, \kappa_1, S-\kappa_1, 0, 0)(C_{\xi} \otimes I_{l_{1}} T_{2}^{M_{N}} T_{l-1}^{T_{N}})e.
\]

12. The probability that an arriving emergency call preempts the service of an ongoing handoff/new call in catastrophic scenario:
\[
P_{\text{preempt}}^{\text{emr}} = \frac{1}{\lambda_{\text{H}}^}\sum_{k_{2}=1}^{K} z_s(0, 0, \kappa_2, S, K-\kappa_2)+ \sum_{k_{1}=1}^{K-\kappa_1} \sum_{\kappa_2=0}^{K-\kappa_1} z_s(0, \kappa_1, \kappa_2, S, K-\kappa_1-\kappa_2)(C_{\xi} \otimes I_{l_{1}} T_{2}^{M_{N}} T_{l-1}^{T_{N}})e.
\]

13. The intensity by which a retrial call is successfully connected to an available channel:
\[
\theta_{\text{r}}^{\text{succ}} = \sum_{l=1}^{\infty} \sum_{\kappa_1=0}^{S} \sum_{\kappa_2=0}^{S-\kappa_1} \theta z_s(l, \kappa_1, \kappa_2, 0, 0)(I_{l_{1}} T_{2}^{M_{N}} T_{l-1}^{T_{N}})e.
\]

5 Numerical Illustration

In this section, the qualitative behaviour of the proposed model is explored through a few experiments. All the numerical experiments have been conducted considering \(\delta = 10^{-12}, \epsilon_{g} = 10^{-10}\), and \(\epsilon_{f} = 10^{-10}\). For the numerical computation, the matrices for the \(MMAP\) are referred from [13] as follows in normal scenario,
\[
C_0 = \begin{pmatrix}
-0.8109843 & 0 \\
0 & -0.02632213
\end{pmatrix}, \quad C_{\xi} = \begin{pmatrix}
0.201398 & 0.0013479 \\
0.003665 & 0.0029153
\end{pmatrix}, \quad C_{N} = \begin{pmatrix}
0.5204914 & 0.004043935 \\
0.0013395 & 0.00874592
\end{pmatrix},
\]
and in catastrophic scenario as follows,
\[
C_0 = \begin{pmatrix}
-0.810 & 0 \\
0 & -0.026
\end{pmatrix}, \quad C_{\xi} = \begin{pmatrix}
0.20 & 0.003 \\
0.003 & 0.002
\end{pmatrix}, \quad C_{N} = \begin{pmatrix}
0.30 & 0.002 \\
0.005 & 0.004
\end{pmatrix}, \quad C_{\xi} = \begin{pmatrix}
0.30 & 0.002 \\
0.005 & 0.004
\end{pmatrix}.
\]
The correlation coefficients for both types of calls are \(C_{r}^{(1)} = C_{r}^{(2)} = 0.2\) and the variation coefficients for both types of calls are \(C_{r}^{(1)} = C_{r}^{(2)} = 12.34\). The arrival rates \(\lambda_{\xi} = 0.15, \lambda_{N} = 0.45\). Let \(PH\) distributions parameters for the service rates of a handoff and a new call be
\[
\beta_{\xi} = (0.05, \ 0.95), \quad A_{\xi} = \begin{pmatrix}
-0.031 & 0 \\
0 & -2.4
\end{pmatrix}, \quad \text{and} \quad \beta_{N} = (0.1, \ 0.9), \quad A_{N} = \begin{pmatrix}
-0.033 & 0 \\
0 & -2.52
\end{pmatrix},
\]
and \(\beta_{\xi} = (0,\ 1), \quad A_{\xi} = \begin{pmatrix}
-1 & 0 \\
0 & -1
\end{pmatrix}.
\]
The fundamental service rates are \(\mu_{\xi} = 0.5, \mu_{N} = 0.3\). The retrial rate of a retrial call, following \(PH\) distribution, is given by the parameters [2]
\[
\gamma = (1, \ 0), \quad \Gamma = \begin{pmatrix}
-2 & 2 \\
0 & -2
\end{pmatrix}, \quad \theta = 1.
\]
To demonstrate the feasibility of the developed model, some interesting observations of the proposed system are described through the following numerical experiments. These experiments will present the behaviour of performance measures with respect to arrival, service and retrial rates.

**Experiment 1:** The objective here is to analyze the impact of arrival rate ($\lambda_H$) and service rate ($\mu_H$) of handoff call over the loss probabilities, i.e., dropping probability of handoff call in normal scenario ($P^n_{nd}$) and preemption probability for new call in normal scenario ($P^n_{new\,preempt}$).

Figures 1(a) and 1(b) represent $P^n_{nd}$ and $P^n_{new\,preempt}$ as functions of total number of channels $S$ and $\lambda_H$. It can be observed from Figure 1(a) that $P^n_{nd}$ increases with respect to $\lambda_H$ for a fixed value of $S$. Moreover, under the same value of $\lambda_H$, $P^n_{nd}$ decreases with respect to $S$. When handoff calls arrive frequently in the system, all the channels are most likely to be occupied by handoff calls and consequently extra handoff calls will be dropped. Therefore, the value of $P^n_{nd}$ increases with increasing $\lambda_H$. In this scenario, if the total channels are increased in the system, more handoff call will be able to obtain the service, as a result, $P^n_{nd}$ decreases. Figure 1(b) exhibits the impact of $\lambda_H$ and $S$ over $P^n_{new\,preempt}$. It is seen from the graph that the values of $P^n_{new\,preempt}$ for different $S$, first increase, and then decrease. The cause for this behavior of $P^n_{new\,preempt}$ lies in the following explanation. When $\lambda_H$ is relatively small, an arriving handoff call often finds at least one channel available, and consequently the ongoing service of a new call is not preempted by the arriving handoff call. As $\lambda_H$ increases, the number of handoff calls also increase in the system. If an arriving handoff call finds all the channels occupied and at least one of them is serving a new call, the service of that new call will be preempted by the arriving handoff call. Hence, $P^n_{new\,preempt}$ increases and reaches maximum at some value of $\lambda_H$. Further, the decreasing behaviour of $P^n_{new\,preempt}$ is explained by the fact that, with the increment in $\lambda_H$, all the channels are occupied with handoff calls. Thus, the number of new calls in the service decreases and the probability that an arriving handoff call preempts the service of a new call decreases.

Further, it can also be observed from Figures 2(a) and 2(b) that, for the fixed value of $\mu_H$, $P^n_{nd}$ and $P^n_{new\,preempt}$ decrease as $S$ increases. An intuitive explanation for this finding can easily be given as follows. As $\mu_H$ increases, the calls are served with increasing rate, consequently the loss probabilities for both types of calls decrease. If $S$ is increased in the system, the chances for both types of calls to obtain the service also increase and consequently $P^n_{nd}$ and $P^n_{new\,preempt}$ decrease.

**Experiment 2:** The main motivation of this experiment is to show how the consideration of exponential retrial times might lead to under/over estimation of performance measures of the system.

Figures 3(a) and 3(b) exhibits the behaviour of the intensity by which a retrial call is successfully connected to an available channel $\theta^{succ}_r$ with respect to retrial rate $\theta$ and expected number of retrial calls $E_{orbit}$ with respect to new call arrival rate $\lambda_N$. It can be seen from the graphs that $\theta^{succ}_r$ increases as $\theta$ increases in the system because the probability that more retrial calls will get connection increases. A clear difference can be observed in the values of $\theta^{succ}_r$ for exponential retrial case and PH distributed retrial case. The value of $\theta^{succ}_r$ is high for PH distributed retrial case in comparison to exponentially distributed retrial case. This finding is useful in determining how taking exponentiality into account in the case of the retrial phenomenon can lead to under/over estimation of system performance. The retrial phenomena that occurs in communication networks is certainly not an exponential process, as is well acknowledged. Since the retrial of call operation completes in
many phases. Therefore, PH distribution is an apt representation. The obtained results justifies the above mentioned statement. The same increasing behaviour can be observed from the Figure 3(b) for $E_{orbit}$ with respect to $\lambda_N$. The value of $E_{orbit}$ is more in case of PH distributed retrial case than the exponentially distributed.

**Experiment 3:** The main purpose of this experiment is to observe the behaviour of loss probabilities in catastrophic scenario with respect to the arrival rate of emergency call ($\lambda_E$) and service rate of emergency call ($\mu_E$).

Figures 4(a) and 4(b) exhibit the behaviour of blocking probability for an emergency call $P_E$ with respect to $\lambda_E$ and $\mu_E$. $P_E$ increases with respect to $\lambda_E$ and decreases with respect to $\mu_E$. An intuitive explanation for this finding can easily be given as follows. As $\lambda_E$ increases, more emergency calls arrive to the system which leads to an increment in the blocking probability of emergency calls. The decreasing behaviour can be observed as $\mu_E$ increases in the system because the emergency calls are served with increasing rate, consequently $P_E$ decreases. If $K$ is increased in the system, the chances for emergency calls to obtain the service also increase and consequently $P_E$ decreases.

Figures 5(a) and 5(b) show the behaviour of blocking probability of new call $P_{cb}$ with respect to $\lambda_E$ and $\mu_E$. $P_{cb}$ increases as $\lambda_E$ increases in the system and decreases as $\mu_E$ increases. As the number of emergency calls increases in the system, less channels will be available for handoff and new calls. As a consequence, an arriving handoff/new call will be dropped from the system and $P_{cb}$ increases. The decreasing behaviour of $P_{cb}$ can be observed as $\mu_E$ increases in the system because the emergency calls will be served with an increasing rate, and more channels will be available for other calls, consequently $P_{cb}$ decreases. If $K$ increases in the system, the chances for handoff/new calls to obtain the service also increase and consequently $P_{cb}$ decreases.

Figures 6(a) and 6(b) represents the behaviour of preemption probability for emergency call $P_{preempt}^{emr}$ over $\lambda_E$ and $\mu_E$. Initially, for small values of $\lambda_E$, $P_{preempt}^{emr}$ increases which shows that due to the arrival of emergency calls, more handoff/new calls will be preempted from the system. After a certain value of $\lambda_E$, $P_{preempt}^{emr}$ starts decreasing. Since, after a certain time period, most of the channels will be occupied with the emergency calls only, consequently, there will be no preemption of handoff/new calls in the system. Therefore, $P_{preempt}^{emr}$ decreases as $\lambda_E$ increases. It can be observed from Figure 6(b) that $P_{preempt}^{emr}$ is a decreasing function of $\mu_E$ which is obvious. These observations are the main motivation for the formulation of the optimization problem illustrated in Section 6.

**6 Optimization Problem**

In the catastrophic scenario, the loss probabilities like blocking probability for emergency calls, blocking probability for new/handoff calls, and preemption probability for emergency calls are considered as performance determining factors for cellular networks. An increment in $P_E$, $P_{cb}$ as well as in $P_{preempt}^{emr}$ indicates unsatisfactory level of service for the customers. Thus, it is required to find the optimal values of parameters in such a way that the loss probabilities should not exceed some pre-defined values. In Section 5 a detailed analysis of loss probabilities with respect to several parameters has been provided. It can be observed from the results that $P_E$, $P_{cb}$ as well as in $P_{preempt}^{emr}$ are mostly affected by the arrival rate of emergency calls $\lambda_E$, service rate of emergency calls $\mu_E$ and the total number of backup channels $K$. The service provider certainly cannot determine $\lambda_E$, yet an approximated value of $\lambda_E$ can be estimated in order to keep sufficient backup channels to provide service. Thus,
Fig. 1: Dependence of the dropping probability $P_d^n$ and preemption probability $P_{\text{preempt}}^{\text{new}}$ over arrival rate of a handoff call $\lambda_H$.

Fig. 2: Dependence of the dropping probability $P_d^n$ and preemption probability $P_{\text{preempt}}^{\text{new}}$ over service rate of a handoff call $\mu_H$.

Fig. 3: Dependence of the intensity by which a retrial call is successfully connected to an available channel $\theta_{\text{succ}}$ over $\theta$ and expected number of retrial calls $E_{\text{orbit}}$ over arrival rate of a new call $\lambda_N$. 
Fig. 4: Dependence of the blocking probability for emergency call $P_e$ over arrival rate of an emergency call $\lambda_E$ and service rate of an emergency call $\mu_E$.

Fig. 5: Dependence of the blocking probability of new call $P_b^e$ over arrival rate of an emergency call $\lambda_E$ and service rate of an emergency call $\mu_E$.

Fig. 6: Dependence of the preemption probability for emergency call $P_{emr\_preempt}$ over arrival rate of an emergency call $\lambda_E$ and service rate of an emergency call $\mu_E$. 
Table 2: Optimal values of $\mu^*_c$, $\lambda^*_c$ and $K$ for different values of $S$ by applying NSGA-II method.
a non-trivial optimization problem is proposed with the decision variables $K$, $\mu_E$ and $\lambda_E$ given as:

$$\begin{align*}
\min & \quad K \\
\text{subject to} & \quad P_t(K, \lambda_E, \mu_E) \leq \epsilon_1, \\
& \quad P_E^c(K, \lambda_E, \mu_E) \leq \epsilon_2, \\
& \quad P_{\text{emr preemp}}(K, \lambda_E, \mu_E) \leq \epsilon_3, \\
& \quad K \leq S, \\
& \quad K, \lambda_E, \mu_E \geq 0.
\end{align*}$$

Here, $\epsilon_1$, $\epsilon_2$ and $\epsilon_3$ are pre-defined values depending on the tolerance of the system for $P_t$, $P_E^c$ and $P_{\text{emr preemp}}$, respectively. Consider $\epsilon_1 = \epsilon_2 = \epsilon_3 = 10^{-3}$ for the numerical computation of the proposed optimization problem.

The above defined constraints are non-linear and highly complex. Thus, a multi-objective evolutionary approach, non-dominated sorting genetic algorithm-II (NSGA-II) has been used to obtain the optimal solution. A modified version of non-dominated sorting genetic algorithm, was proposed by Deb et al. [11] as NSGA-II. The detailed analysis of NSGA-II algorithm can be found in [11]. The main steps of NSGA-II are provided as follows.

1. Initialize the population size $P$ based on the number of decision variables.

2. **Non-dominated sorting:** The initialized population is sorted on the basis of non-domination. Each solution is assigned a fitness or rank equal to its non-domination level. The steps of sort algorithm are as follows:
   - Initialize $S_p = \phi$. $S_p$ is set of all individuals that is being dominated by $p$.
   - Initialize $n_p = 0$. This would be the number of individuals that dominate $p$.
   - For each individual $q$ in $P$, if $p$ dominated $q$ then add $q$ to the set $S_p$ i.e., $S_p = S_p \cup \{q\}$ else if $q$ dominates $p$ then increment the domination counter for $p$ i.e., $n_p = n_p + 1$.
   - If $n_p = 0$ i.e., no individuals dominate $p$ then $p$ belongs to the first front; Set rank of individual $p$ to 1. Update the first front i.e., $F_1 = F_1 \cup \{p\}$. This is carried out for all the individuals in main population $P$.
   - Initialize the front counter to one, i.e., $i = 1$.
   - Define $Q = \phi$, the set for storing the individuals for $(i+1)^{th}$ front. Set $n_q = n_q - 1$, decrements the domination count for individual $q$ in $S_p$. Update the set $Q = Q \cup q$.
   - Increment the front counter by one and set the next front as $F_i = Q$. This step is carried out while the $i^{th}$ front is non-empty.

3. **Crowding Distance:** All the individuals after non-dominated sort are assigned a crowding distance value.

Crowding distance is assigned front wise and compared between two individuals.

4. **Selection:** Once the individuals are sorted based on non-domination and with crowding distance assigned, the selection is carried out using a crowded-comparison operator.

5. **Recombination and selection:** The offspring population is combined with the current generation population and selection is performed to set the individuals of the next generation.
Table 3: Optimal values of $\mu^*_E$, $\lambda^*_E$ and $K$ for different values of $S$ by applying NSGA-II method.

| $S$=5 | $\mu_N = 1$ | $\mu_K = 1$ | $\lambda_N = 1$ | $\lambda_K = 1$ | $\theta = 1$ |
|------|------------|-----------|-------------|-------------|------------|
| $K^*$ | $\lambda^*_E$ | $\mu^*_E$ | $P_E$ | $P_b$ | $P_{\text{preempt}}$ |
| 3 | 1.654 | 5.56205 | 0.000081836 | 0.000089307 | 0.00007471 |
| $S$=5 | $\mu_N = 1$ | $\mu_K = 1.5$ | $\lambda_N = 1$ | $\lambda_K = 1$ | $\theta = 1$ |
| $K^*$ | $\lambda^*_E$ | $\mu^*_E$ | $P_E$ | $P_b$ | $P_{\text{preempt}}$ |
| 3 | 1.654 | 5.5741 | 0.000080241 | 0.000088742 | 0.000073341 |
| $S$=5 | $\mu_N = 1$ | $\mu_K = 2$ | $\lambda_N = 1$ | $\lambda_K = 1$ | $\theta = 1$ |
| $K^*$ | $\lambda^*_E$ | $\mu^*_E$ | $P_E$ | $P_b$ | $P_{\text{preempt}}$ |
| 3 | 1.654 | 5.5787 | 0.000080240 | 0.000088701 | 0.000073339 |
| $S$=6 | $\mu_N = 1$ | $\mu_K = 1$ | $\lambda_N = 1$ | $\lambda_K = 1$ | $\theta = 1$ |
| $K^*$ | $\lambda^*_E$ | $\mu^*_E$ | $P_E$ | $P_b$ | $P_{\text{preempt}}$ |
| 4 | 1.5 | 5.74075 | 0.000047512 | 0.000057343 | 0.00009831 |
| $S$=6 | $\mu_N = 1$ | $\mu_K = 1.5$ | $\lambda_N = 1$ | $\lambda_K = 1$ | $\theta = 1$ |
| $K^*$ | $\lambda^*_E$ | $\mu^*_E$ | $P_E$ | $P_b$ | $P_{\text{preempt}}$ |
| 4 | 1.5 | 5.7221 | 0.000045442 | 0.000056134 | 0.000097745 |
| $S$=6 | $\mu_N = 1$ | $\mu_K = 2$ | $\lambda_N = 1$ | $\lambda_K = 1$ | $\theta = 1$ |
| $K^*$ | $\lambda^*_E$ | $\mu^*_E$ | $P_E$ | $P_b$ | $P_{\text{preempt}}$ |
| 4 | 1.5 | 5.7116 | 0.00004487 | 0.000055443 | 0.0000966412 |
| $S$=6 | $\mu_N = 1$ | $\mu_K = 2.5$ | $\lambda_N = 1$ | $\lambda_K = 1$ | $\theta = 1$ |
| $K^*$ | $\lambda^*_E$ | $\mu^*_E$ | $P_E$ | $P_b$ | $P_{\text{preempt}}$ |
| 4 | 1.5 | 5.70098 | 0.000042114 | 0.00005474 | 0.0000955414 |
| $S$=6 | $\mu_N = 1$ | $\mu_K = 3$ | $\lambda_N = 1$ | $\lambda_K = 1$ | $\theta = 1$ |
| $K^*$ | $\lambda^*_E$ | $\mu^*_E$ | $P_E$ | $P_b$ | $P_{\text{preempt}}$ |
| 4 | 1.5 | 5.70014 | 0.000042001 | 0.00005470 | 0.000095540 |

Tables 2 and 3 represent the optimal values of $k^*$, $\lambda^*_E$, and $\mu^*_E$ for different combinations of arrival rates and different values of $S$. All results were obtained by MATLAB software, which were run on a computer with Intel Core i7-6700 3.40GHz CPU and 8 GB of RAM. The obtained results provide the value of optimal backup channels for various combinations of total number of channels and arrival rates. The proposed optimization problem’s sensitivity analysis is useful in estimating the number of backup channels in an emergency scenario. These findings are extremely beneficial in communication systems and cellular networks.

7 Conclusions

Queuing models with catastrophic occurrences are a driving force in communications and cellular networks in cutting-edge wireless technology. In such types of catastrophic queueing models, traffic of different classes, e.g., video, voice, images, data, etc., are assigned different categories of importance, and consequently their services are effectuated in accordance with an appropriate priority policy. In cellular networks, the kinds of
systems where a higher priority traffic has an advantage in access to service compared to less important ones, are explored through priority policies. Therefore, in this study a MMAP[k]/PH[k]/S catastrophic queueing model with preemptive repeat priority policy and PH distributed retrial times is investigated. Due to the brief span of inter-retrial times in comparison to service times, a more generalized approach, PH distributed retrial times is used so that the performance of the system is not over or under estimated. The proposed model deals with three types of incoming calls, handoff call, new call and emergency call. Before the occurrence of a catastrophe, handoff call are provided preemptive priority over new call. Whereas, in the catastrophic scenario, emergency calls are provided preemptive priority over handoff and new calls. When a calamity strikes, backup channels are used to establish communication in the affected area. The underlying process of the presented system is modelled by AQTM.C. Ergodicity conditions of the underlying Markov chain are obtained by proving that the Markov chain belongs to the class of AQTM.C. A new algorithm is developed for approximate computation of the stationary distribution. Further, the expressions of various performance measures have been derived for the numerical illustration. Due to the consideration of the preemptive priority policy, the blocking probability for emergency calls decreases and simultaneously the frequent termination of services for handoff and new calls increases the probability of preemption. Thus, an optimization problem to obtain optimal value of total number of backup channels has been formulated and dealt by employing NSGA-II approach.

References

[1] Ammar, S.I. and Rajadurai, P., 2019. Performance analysis of preemptive priority retrial queueing system with disaster under working breakdown services. Symmetry, 11(3), p.419.

[2] Artalejo, J.R. and Gomez-Corral, A., 2007. Modelling communication systems with phase type service and retrial times. IEEE Communications Letters, 11(12), pp.955-957.

[3] Baumann, H. and Sandmann, W., 2012. Steady state analysis of level dependent quasi-birth-and-death processes with catastrophes. Computers & operations research, 39(2), pp.413-423.

[4] Brandwajn, A. and Begin, T., 2017. Multi-server preemptive priority queue with general arrivals and service times. Performance Evaluation, 115, pp.150-164.

[5] Chakravarthy, S.R., 2017. A catastrophic queueing model with delayed action. Applied Mathematical Modelling, 46, pp.631-649.

[6] Chakravarthy, S.R., Dudin, A.N. and Klimenok, V.I., 2010. A retrial queueing model with map arrivals, catastrophic failures with repairs, and customer impatience. Asia-Pacific Journal of Operational Research, 27(06), pp.727-752.

[7] Chakravarthy, S.R., 2020. A Retrial Queueing Model with Thresholds and Phase Type Retrial times. Journal of Applied Mathematics & Informatics, (3-4), p.351.

[8] Chang, W., 1965. Preemptive priority queues. Operations research, 13(5), pp.820-827.

[9] Dabrowski, C., 2015. Catastrophic event phenomena in communication networks: A survey. Computer Science Review, 18, pp.10-45.

[10] Dayar, T., 2012. Analyzing Markov chains using Kronecker products: theory and applications. Springer Science & Business Media.
[11] Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T.A.M.T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE transactions on evolutionary computation, 6(2), pp.182-197.

[12] Dharmaraja, S., Jindal, V. and Alfa, A.S., 2008. Phase-type models for cellular networks supporting voice, video and data traffic. Mathematical and computer modelling, 47(11-12), pp.1167-1180.

[13] Dudin, A., Kim, C., Dudin, S. and Dudina, O., 2016. Analysis and optimization of Guard Channel Policy with buffering in cellular mobile networks. Computer Networks, 107, pp.258-269.

[14] Dudin, S. and Dudina, O., 2019. Retrial multi-server queuing system with PHF service time distribution as a model of a channel with unreliable transmission of information. Applied Mathematical Modelling, 65, pp.676-695.

[15] He, Q.M., 2014. Fundamentals of matrix-analytic methods (Vol. 365). New York: Springer.

[16] He, Q.M. and Alfa, A.S., 2018. Space reduction for a class of multidimensional Markov chains: A summary and some applications. INFORMS Journal on Computing, 30(1), pp.1-10.

[17] Jain, V and Raj, R and Dharmaraja, S., 2021. Numerical optimization of loss system with retrial phenomenon in cellular networks. International Journal of Operational Research, In Press.

[18] Jain, V and Raj, R and Dharmaraja, S., 2021. Cost analysis of a $\text{MMA}[2]/\text{PH}[2]/S$ performability model with PH retrial times using simulated annealing method. Communication in Statistics- Theory and Methods, Communicated.

[19] Jain, V and Raj, R and Dharmaraja, S., 2021. Optimization of traffic control in $\text{MMA}[2]/\text{PH}[2]/S$ priority queueing model with PH retrial times and the preemptive repeat policy. Journal of Industrial and Management Optimization, Communicated.

[20] Kim, J. and Kim, B., 2016. A survey of retrial queueing systems. Annals of Operations Research, 247(1), pp.3-36.

[21] Klimenok, V. and Dudin, A., 2006. Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory. Queueing Systems, 54(4), pp.245-259.

[22] Krishnamoorthy, A., Babu, S. and Narayanan, V.C., 2008. $\text{MAP}/(\text{PH}/\text{PH})/c$ queue with self-generation of priorities and non-preemptive service. Stochastic Analysis and Applications, 26(6), pp.1250-1266.

[23] Kumar, N. and Gupta, U.C., 2020. Analysis of batch Bernoulli process subject to discrete-time renewal generated binomial catastrophes. Annals of Operations Research, 287(1), pp.257-283.

[24] Latouche, G. and Ramaswami, V., 1999. Introduction to matrix analytic methods in stochastic modeling. Society for Industrial and Applied Mathematics.

[25] Machihara, F., 1995. A bridge between preemptive and non-preemptive queueing models. Performance Evaluation, 23(2), pp.93-106.

[26] Shin, Y.W. and Moon, D.H., 2011. Approximation of $\text{M}/\text{M}/c$ retrial queue with PH-retrial times. European journal of operational research, 213(1), pp.205-209.

[27] Sudhesh, R., Savitha, P. and Dharmaraja, S., 2017. Transient analysis of a two-heterogeneous servers queue with system disaster, server repair and customers’ impatience. Top, 25(1), pp.179-205.

[28] Wang, J., Liu, B. and Li, J., 2008. Transient analysis of an $\text{M}/\text{G}/1$ retrial queue subject to disasters and server failures. European Journal of Operational Research, 189(3), pp.1118-1132.
[29] Yajima, M. and Phung-Duc, T., 2019. A central limit theorem for a Markov-modulated infinite-server queue with batch Poisson arrivals and binomial catastrophes. Performance Evaluation, 129, pp.2-14.

[30] Yechiali, U., 2007. Queues with system disasters and impatient customers when system is down. Queueing Systems, 56(3), pp.195-202.