MULTIPLE QUEUES WITH SIMULTANEOUS ARRIVALS

by
Toygar Emre
B.S., Industrial Engineering, Boğaziçi University, 2017

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Industrial Engineering
Boğaziçi University
2022

ACKNOWLEDGEMENTS

Initially, I would like to thank my supervisor Prof. Refik Güllü for his support, guidance and patience during my entire study. I deem that I could not complete my thesis without his confidence and encouragement. I am also very pleased that Assoc. Prof. Aybek Korugan and Assist. Prof. Zeynep Gökçe İşlier take place in my thesis committee suggesting distinct perspectives towards my thesis.

I am definitely indebted to Assoc. Prof. Aybek Korugan who actually instilled the inspiration into me to proceed in my thesis. Thanks to his queuing theory course I managed to take significant steps in the crucial part of my study.

Finally, I am also grateful to my family for their support. I strived for completing my master thesis in tough conditions and they pulled me through my mental and physical problems, thus made me focus on my study.

ABSTRACT

MULTIPLE QUEUES WITH SIMULTANEOUS ARRIVALS

Queuing theory problems have been the topic of deep research owing to the fact that so many difficulties are in existence and their significance in real life cases can not be ignored. Those problems can be observed in numerous sectors such as telecommunications, airlines, logistics, hospitals, computing, production and inventory. Besides, speed is the key word in today's world because population is almost at the peak, thus demands or requests must be met as much as possible. However, our world has limited sources that is why there has to be some delays and queues. Additionally, game theory is one of the most important topics and it comes into prominence due to increasing competition in the world. There are lots of organizations which dwell in aforementioned sectors and they need to compete with each other to maximize their benefits. Just as in queuing theory, application of game theory spans the huge part of real life problems involving so much burden. So, there are abundance of works which dive into the distinct branches of game theory. In this study, both queueing theory and game theory are taken into consideration. We include the concept of game analysis, server rate optimization, multiple queues, loss systems and simultaneous arrivals at the same time whereas the studies in literature just focus on some of them. In our first case, we apply a game theoretic approach to two loss queuing systems under specific assumptions. With the deployment of server rate optimization we reach Nash equilibrium points. We also provide some analytical derivations and validate them using simulations. In our second case, we deal with one loss system with an uncapacitated queue involving quasi birth death process. We find the steady state probabilities employing two different computation techniques and calculate the expected profit for each queue in the system.

ÖZET

AYNI ANDA VARIŞLARLA BİRDEN FAZLA KUYRUK

Kuyruk teorisi problemleri, sahip oldukları zorluklar sebebiyle ve gerçek hayat durumlarındaki öneminin göz ardı edilememesi nedeniyle derin araştırmalara konu olmuştur. Bu sorunlar telekomünikasyon, havayolları, lojistik, hastaneler, bilgisayar, üretim ve envanter gibi çok sayıda sektörde gözlemlenebilir. Ayrıca günümüz dünyasında hız anahtar kelimedir çünkü nüfus neredeyse zirvededir, bu nedenle talepler mümkün olduğunca hızla ve kayıpsız şekilde karşılanmalıdır. Ancak dünyamızın sınırlı kaynakları var, bu yüzden bazı gecikmeler ve kuyruklar olması kaçınılmaz olmaktadır. Ek olarak oyun teorisi en önemli konulardan biridir ve dünyada artan rekabet nedeniyle ön plana çıkmaktadır. Bahsedilen sektörlerde yaşayan ve faydalarını en üst düzeye çıkarmak için birbirleriyle rekabet etmeleri gereken birçok kuruluş var. Kuyruk teorisinde olduğu gibi, oyun teorisinin uygulanması, çok fazla yük içeren gerçek hayat problemlerinin büyük bir bölümünü kapsar. Dolayısıyla oyun teorisinin farklı dallarına odaklanan çok sayıda çalışma var. Bu çalışmada hem kuyruk teorisi hem de oyun teorisi ele alınmıştır. Çalışmamızda oyun analizi, sunucu oranı optimizasyonu, çoklu kuyruklar, kayıp sistemleri ve eşzamanlı varış kavramlarını aynı anda dahil ediyoruz ancak literatürdeki çalışmalar şu ana kadar sadece bazılarına odaklanmış vaziyettedir. İlk durumumuzda, belirli varsayımlar altında iki kayıp sıra sistemine oyun teorik bir yaklaşım uyguluyoruz. Sunucu hızı optimizasyonunun konuşlandırılmasıyla Nash denge noktalarına ulaşıyoruz. Ayrıca bazı analitik türevler sağlıyor ve simülasyonları kullanarak bunları doğruluyoruz. İkinci vakamızda, yarı doğum ölüm sürecini içeren, kapasitesi olmayan bir kuyruğa sahip bir kayıp sistemi ile ilgileniyoruz. İki farklı hesaplama tekniği kullanarak kararlı durum olasılıklarını buluyoruz ve sistemdeki her kuyruk için beklenen karı hesaplıyoruz.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

c_i	Cost of the i^{th} queue per unit time
P_{ij}	Stationary probability of $(i,\, j)$ state in two loss system
Q_1	Generator matrix in two loss systems
Q_2	Generator matrix in one loss systems
R	A fixed matrix for the solution of one loss system
r_i	Revenue of the i^{th} queue per unit time
X_1	Number of customers in the first queue of one loss system
X_2	Number of customers in the second queue of one loss system
Y_i	Number of customers in the i^{th} queue of two loss system
θ	Lagrange Multiplier
λ	Arrival rate with Poisson process
μ_i	Server rate of the i^{th} queue
ξ	Stationary probability vector in one loss system
ξ_i	Stationary probability vector for the i^{th} level
ξ_{ij}	Stationary probability of the i^{th} level with the j^{th} sublevel
π_i	Profit of the i^{th} queue

1. INTRODUCTION

As a general description, queuing theory refers to the mathematical calculations that belong to waiting lines and queues. In this environment there are always requests to try to get services as well as servers that provide the relevant services. Unfortunately, our world has limited sources and capacity, therefore it is almost impossible to serve the needs at the same time and somebody or something has to wait while others do not. Employing queuing theory, we are capable of investigating the system of lines and finding the essential values of that system such as the expected number of customers, optimal number of servers, optimal number of capacity, expected waiting time of customers. Also, different queuing disciplines can be included such as first-in first-out, prioritized, preemptive or non-preemptive. Owing to the fact that queuing systems can be applied to numerous areas like telecommunications, airlines, logistics, hospitals, computing, production, inventory there is an abundance of work handling the subject. Because of the wide range of queuing systems they are demonstrated in terms of Kendall notation which was first proposed by D.G.Kendall [1]. Kendall notation is written as A/S/c/K/N/D where A is arrival process, S is the distribution of the service time, c is number of servers, K is the capacity of the queue, N is the number of customers, and D is the queue discipline. For some queuing systems like M/M/c, where there is a poisson arrival process and exponential distributed service times with c servers, precise results may be achieved analytically. However, we have to employ approximation techniques or upper/lower bounds for the systems like G/G/c queues. Additionally, it is possible to obtain a queuing network where there are multiple waiting lines that customers, jobs or items can depart from a queue and arrive at another queue. Undoubtedly, queuing networks are harder, therefore pose more analytical and computational difficulties. Due to the fact that analytical solutions are extremely hard for those networks, computational methods which involve algorithms for product form or non-product form solutions are favoured over analytical methods.

Although queuing theory is a powerful study to derive significant results in real life problems, it may be insufficient to explain some interdependencies in systems in which competitors are in existence. Therefore, it may be better to merge queuing theory and game theory together in such circumstances. Game theory is capable of ensuring a quantitative approach for strategic decision making in a competitive environment where one player's payoff is dependent on other players' payoffs. As queuing theory, game theory also permeates lots of fields such as energy, operations research, economics, finance, wireless networks, psychology and business. Because of the fact that numerous types of competitions take place in real life cases, game theory also has a wide range of variety. Mostly, cooperative and non-cooperative games are examined where there is a competition between groups of players and no allowance for agreement respectively. Certainly, application of game theory to problems may require huge effort as in queue theory due to the fact that some problems may not have an explicit cost function for players, therefore appropriate cost functions need to be determined. Equilibrium point in a game is a candidate to be an optimal point with respect to payoffs, but those points may not exist, so employing randomized strategies those points may be acquired. However, the nature of the original game may change significantly. Furthermore, there may be multiple equilibrium points and in this case players may not be sure about which one to focus, but there are some methods to approach promising points. In short, both queuing theory and game theory are great tools to solve problems in real life. When there are some decision makers or there is some kind of competition in queuing systems/networks it is beneficial to combine those theories so as to attain satisfactory results.

In our thesis, we divide our investigation into two parts where the former one is mostly related to game theoretic view of multiple queues including two loss systems consisting of two M/M/1/1 queues whereas the latter one is considerably pertinent to quasi birth death processes involving one loss system which is comprised of M/M/1/1 and M/M/1 queues. In the first case, we have simultaneous arrivals with the Poisson process while servers' rates are exponentially distributed. Arrivals come towards the two loss systems as a couple and only join those systems if both queues are not at full capacity, otherwise both arrivals are lost. In addition, there is a competition between two M/M/1/1 queues because each of the queues is determined to maximize their profits and one of those queues tries to set its optimal server rate against the other queue's server rate. In the second case, we have again simultaneous arrivals with Poisson process whereas exponentially distributed servers' rates are in existence. Due to the fact that one of the queues is not capacitated, stability analysis needs to be taken into account in order to prevent the system explosion. Additionally, in order to find the expected profits for each queues stationary probabilities are needed to be analysed. Thus, quasi birth death processes are taken into consideration to reach those probabilities and two different solution techniques are used to derive them. Those models and assumptions actually suit real life problems such as in health care management or assemble to order systems. During the maintenance and production of big vehicles which require a large space such as ships or planes one loss system with M/M/1 queue model can be employed where big vehicles are located in M/M/1/1 queue while orders of their specific component create a M/M/1 queue. Also, during surgeries patients are under some operations in surgery rooms which can be deemed as an M/M/1/1 queue while patients' relatives or acquaintances are in M/M/1 queue for payments. In fact, it is possible to give numerous real case problems and deployment of those models presents a great potential to yield promising results.

As a remaining part of our thesis, we start out presenting our literature review in Chapter 2 in which there are plenty of significant works with respect to simultaneous/correlated demands-arrivals, service rate optimization and loss systems. Afterwards in Chapter 3 and Chapter 4, we initially introduce our two loss systems consisting of two M/M/1/1 queues and one loss system which is comprised of M/M/1/1 and M/M/1 queues respectively. Subsequently, we provide propositions with their proofs and validate those analytical demonstrations with some numerical results. Finally, we end up with our conclusion part summarizing our study in Chapter 5.

2. LITERATURE REVIEW

In queuing theory problems there are either queuing systems or queuing networks and they are fundamental ways of handling the calculations regarding flow of customers, jobs, items, objects while the scarcity of resources for providing services takes place. Nowadays, queues almost infiltrate our lives entirely due to the fact that we live in a rather complex world where the needs of people or customers are at the peak and those should be satisfied within a short time as much as possible. That is why there are a significant number of studies encompassing distinct aspects of queues and the number of those works are increasing gradually. Additionally, competitive structures are rising in a dramatic way owing to the fact that the number of organizations which strive for providing similar services to their customers are increasing and they always pursue to maximize their own benefits as much as possible with or without coalitions. Therefore, an organization needs to take action according to its rivals' position and game theoretic approach to such cases tends to result in sensible solutions.

In our study we examine two distinct cases. The first one consists of two M/M/1/1 queues each of which has arrivals according to Poisson process and server rate with exponential distribution. However, they are not totally independent queues because there are simultaneous arrivals, that means if there is an arrival for a M/M/1/1 queue, then there is also an arrival for the other M/M/1/1 queue. Also, those arrivals can join the queues only if the both queues are not at the full size. Each of those queues aims to maximize their own profits and always tend to obtain optimal server rate value against the other queue's server rate. The second one has one loss system with M/M/1 queue where the simultaneous arrivals take place as in the first case while servers' rates are exponentially distributed. Calculations related to quasi birth death processes are applied to the second case in order to acquire the stationary probabilities deploying two different methods, and to find the expected profits of each of the queues subsequently. Therefore, in this chapter we split our discussion into three parts where the first one includes simultaneous/correlated demands-arrivals involving queuing-production-inventory systems, quasi birth death processes, the second one is related to service rate optimisation of those systems and the final one involves the loss systems.

2.1. Simultaneous/Correlated Demands-Arrivals

The first stream of literature presents the studies related to simultaneous/correlated demands-arrivals that have arisen to point out the problems such as in assemble-to-order systems, multi-item inventory systems, make-to-order systems and other similar systems. Song, Liu and Xu [2] acquired performance measures for ATO systems where capacity, Poisson arrivals and exponential lead times exist. Matrix geometric method is employed in order to attain those measures. Also, they express that when ATO systems are large this method becomes computationally expensive, so the need for an efficient heuristic method is better to approximate those measures. Swaminathan and Tayur [3] investigated the case where final products are assembled using semi finished products according to arrival of orders, also all products are stored as semi finished base. They take the single-period model into consideration where production times are distributed generally and demands occur at the beginning of the period. They managed to solve the problem on a small scale but the real case which is actually a motivation for their study is hard to analyse. Cheung and Hausmann [4] analysed the multi-item spare inventory problem in which combination of items that need to be processed is linked with job type. Poisson process assumption is employed for jobs arrivals whereas exponentially or deterministically distributed infinite servers assumption to process relevant jobs takes place. For the process of jobs, a set of items according to job type need to be assembled as in assemble to order systems. They use approximations so as to calculate the expected number of jobs in the system. As an assumption, approximations involve independence across items for the number of items being processed. Additionally, they realise that the more items increase, the harder calculations become to compute. The work of Glasserman and Wang [5] includes trade-offs between lead-time and inventory in assemble to order systems. Poisson arrivals of demand and diverse distribution in servers are included in their work. They demonstrate a linear relationship between delivery lead time and inventory for a fixed service level. They provide us with examples involving a small number of items where the relationship strongly holds when the fill rate is so high. Zhang [6] investigates the expected time delay for systems which includes continuous review base-stock inventory policy according to assumption of independency for item delays. In those systems customers may entail several items. Additionally, Zhang shows that total time delay may be overestimated in consequence of correlated items if the assumption of independency among items is made. Also, inclusion of correlations for large systems makes problem hard to solve, therefore Zhang offers an algorithm which works well in fast service level case. Improvement of algorithm can be clearly seen if there is a decrease in demand entailing multiple items and utilisation of individual facilities. Xu and Li [7] dive into the correlated stochastic queues in order to investigate the problem of correlation between performance measures of items. In their work, the notion of "majorization" is introduced as a new tool to analyse the dependency in multivariate stochastic systems where correlated arrivals exist. They manage to apply their tool to assemble to order systems and demonstrate that adjusting input processes may boost system performance under pertinent conditions. With respect to assemble to order systems with partial order servers employing Markovian distributions, the number of performance bounds are analysed in the work of Dayanık, Song and Xu [8]. If the arriving order lacks some components, then the order just takes existing components according to partial order servers, which means orders are met partially. Also, their work includes an algorithm that can deal with non Markovian component production lead times and customer switching. Iravani *et al.* [9] examine the queuing system with correlated arrivals where parallel queue with bulk service exist. In order to obtain performance measures matrix geometric method is employed. In addition, extension of the principles of decomposition algorithm to examine various parallel queueing systems with correlated arrivals is presented. Li and Xu [10] examines the parallel queuing systems with correlated arrival processes to distinct queues where dependence structure and bounds are considered significantly. Aim of work is to obtain better knowledge about dependency and to derive various upper and lower bounds for the statistics of joint performance measures. Application of their results to synchronised queues shows how performance of measures may be changed according to different types of dependencies. Elhafsi *et al.* [11] investigates a single product assemble to order system which serves the demands of the final product which are assumed to have independent Poisson process with distinct rates and individual items based on relevant demand. If demand related to end product and items are not met, then this situation incurs cost for the system due to the lost demands and items. Additionally, production lead times are non identical and have an independent exponential distribution where produced items are held in stock. In this paper, they aim to determine optimal policy for inventory allocation and item production. Employing lots of numerical experiments they also propose three heuristic methods performing well due to the fact that numerically optimal policy is burdensome to acquire stemming from the curse of dimensionality of dynamic programming. Kushner [12] investigates a heavy traffic case of optimal control for assemble to order systems where demands consist of multiple end products each of which may require various components. Components are produced according to random production times whereas demand intervals are random and may occur in batches. As long as demands can be satisfied, assembly takes place otherwise demands are lost. Using numerical methods they achieve to present a reasonable optimal control policy. Lu *et al.* [13] analyse the optimal budget allocation among inventories in order to minimise weighted average of backorders for product types where each product order has a batch Poisson process and lead time for replenishment of components' inventory is stochastic including assemble to order systems. So as to construct surrogate optimization problems they come up with bound and approximations related to expected number of backorders, thus numerical examples are provided to show the effectiveness of bounds and approximations. Van Houdt [14] studies a broad class of semi Markovian queues introduced by Sengupta. The class on which his work is based includes many queues such as G/M/1, SM/MAP/1 as well as queues involving correlated inter-arrival and service times. In order to have attainable results in his work, the matrix geometric method pertains to quasi birth death process is deployed considerably. Deploying a stochastic programming formulation, Jaarsveld and Wolf [15] create an algorithm for inventory control in unequivocally efficient and scalable assemble to order systems. Considering a continuous time model they aim to find stock levels for components so as to minimise holding costs as well as backorder costs for products. Nevertheless, they analyse the effectiveness of underestimating stock out cases and derive an computable upper bound for the ease of optimal allocation.

Gao *et al.* [16] investigate a multi item, multiple classes of demand, assemble to order system where each item's inventory is based on item level and controlled by the base stock policy including finite capacity. Requiring a subset of items, arrivals of demands follow the Poisson process whereas replenishment of each item has exponential distribution. If the existence of unsatisfied demands takes place, then total order service and partial order service occur as stockout cases. Modelling the system as a queuing network, they actually obtain a quasi birth death process in which a matrix geometric solution method is employed to reach joint steady state distribution. They also display numerical examples to draw attention to how system performance changes with varying system parameters emphasizing the significance of involving machine failures. DeCroix *et al.* [17] dive into analysis of multi product assemble to order systems where inventory is based on component level and the finished products are obtained according to stochastic customer demands. Additionally, the system is exposed to stochastic returns of subsets of components as well as stochastic demand. They accomplish to identify various ways where returns complicate the behaviour of the system showing how to undermine or wipe out these complexities during the calculation of important performance metrics such as the immediate fill rate, the fill rate within time window and average backorders. Nevertheless, they display a method so as to compute near optimal base stock policy.

2.2. Service Rate Optimization

The second stream of literature encompasses the articles which are aimed at analyzing the service rate optimisation both in single queues and multiple queues. We also divide the part of multiple queues into two subparts one of which is a game theoretic approach which examines the relationships among people or organizations that have conflicted/different goals or partially compete and the other one is centralised decision making. Cooperative and non cooperative structures are taken into account revealing solution procedures in those articles that involve queuing, inventory and production systems.

2.2.1. Single Queue Server Rate Optimization

Yang *et al.* [18] studies M/M/2 queue with heterogeneous servers with multiple vacations and working breakdowns as a steady state analysis. They achieve to convert the problem into an quasi birth death process and employ matrix geometric method so as to calculate joint steady state probabilities. Additionally, they use heuristic method for the minimisation of cost approximating optimal service rates for two servers. As a result, they display some numerical results demonstrating the effects of distinct approximated optimal service rates and presents a practical example for an application of model. In the study of Yang *et al.* [19], they examine M/M/1 queue with secondary optional service that means when customer arrives, first he is served by a necessary server but some of customers may also be served by a second optional server as their wish. During the breakdown of servers, service is not ceased entirely, instead server rate is just reduced. In order to obtain steady state probabilities and compute some performance measures matrix geometric method is deployed. In addition, application of genetic algorithm is aimed at optimising the cost of problem heuristically approximating the optimal service rates. Lax and Indira [20] examine a queue with finite buffer multiple working vacations including balking, reneging and Bernoulli schedule vacation interruption under N policy. In this queue customer may prefer to join or balk, afterwards he may renege whereas server leaves the system for vacation. But the server may return to the system interrupting his vacation if there are at least N customers waiting in the queue. Under those circumstances, length of the system in steady state conditions is obtained employing recursive analysis. Also, using ant colony optimisation some performance measures and cost objective that is in terms of service rate and reneging rate is obtained heuristically. Hamasha *et al.* [21] suggest a Markov based service delivery model involving single server single queue and multiple server single queue structures. They aim to maximise the profit of the system acquiring the pertinent optimal parameters and also demonstrate that profit is so vulnerable to optimal service and arrival rates among other parameters. Ke *et al.* [22] focus on M/M/c balking retrial queue with vacation where single and multiple vacations are examined. They turn the problem into a quasi birth death process in which a matrix geometric method is used to get the steady state probabilities and performance measures. They also aim to minimise the cost objective in stationary conditions obtaining the optimal number of servers and optimal rates for servers employing Quasi Newton method, Nelder Mead method and heuristic methods such as simulated annealing. As a conclusion, they present numerical results derived from optimisation processes and provide us with an example for an application. Elijah *et al.* [23] study the part of the mining operations involving loading of material from the pit and hauling them to the processing plants which constitutes the half of the total operation costs of mine. Optimisation of productivity based on the application of shovel truck haulage system in limestone open pit mine is ensured using queuing theory methods. Their study includes multichannel queueing approach and so as to compute optimal interarrival rate and service rate for different numbers of trucks in system they construct a model. Liou *et al.* [24] examine a Markovian queue with a single unreliable server and infinite capacity where customers may not enter the system and leave the queue after entrance. Using the matrix geometric method stationary probabilities and stability conditions are calculated. Nevertheless, as a heuristic optimisation tool particle swarm optimisation is favoured so as to obtain optimal parameters that belong to the queue system according to cost minimization. Chiang *et al.* [25] display a systematic approach of nonlinear optimisation of queuing systems in an economic way. Demonstrating some convex structures of various queuing systems they apply convex optimisation methods to both single queues and networks. They also accomplish to show results that are pertinent to blocking probability minimisation with suitable service rate and arrival rate. Rapoport *et al.* [26] analyse a class of queuing problems in which endogenous arrival times based on non cooperative n-person games in normal form and single server exist. Employing multiple equilibria in pure strategies, some problems related to tacit coordination may arise. Their results acquired by deploying a Markov Chain algorithm to calculate symmetric mixed strategy equilibrium point presents that consistent and replicable patterns of arrival which leads to a concrete structure for mixed strategy equilibrium exists on aggregate level instead of individual level. Harrison and George [27] aims to minimise average cost per unit time over an infinite horizon including single server queue in which holding costs are a nondecreasing function in terms of queue length that is changed according to birth death process. They find the optimal service rates' equations that are non-decreasing as a function of queue length using the optimality equation of average cost dynamic programming as well as optimal service rates are bounded when holding cost function is bounded. Providing numerical examples they compare minimum average cost assuming state dependent service rates with that minimum average cost assuming fixed service rate. Moura *et al.* [28] suggest a model so as to study the interactions between hospitals and Original Equipment Manfacturer which ensures maintenance services for the advanced technology equipment used in healthcare institutions. In their work, maintenance services are provided for two distinct classes of hospitals one of which prefers to hire extended warranty whereas the other class prefers to pay for each maintenance intervention on demand regardless of priority. They assume that failures and repairs take place in a two class G/M/1 priority queuing system and Original Equipment Manfacturer tries to maximise its profits. Thus, they adopt a Stackelberg game where the Original Equipment Manfacturer is the leader and the customer is the follower. Xia *et al.* [29] examine the single server queue with Markovian arrival process and exponential service rate which depends on state of system such as queue length and phase of Markovian arrival process. Main goal is to minimise average total cost in the long run optimising service rates deploying matrix analytic methods and sensitivity based optimisation theory. In conclusion, they provide numerical examples to show the main results and try to observe the effect of the phase of the Markovian arrival process in the MAP/M/1 queuing system. Acharya and Rodriguez [30] investigate a M/M/1/m queuing system estimating maximum likelihood of the change point of arrival rate where the change occurs after a constant unknown number of finished services. They show that the estimators of maximum likelihood function are weakly consistent as well as compute the estimators of the traffic intensities after and before the change point. Dave and Shah [31] aim to obtain the unknown parameters of a stationary M/M/2 queuing system with heterogeneous servers employing maximum likelihood function in their paper. Observing the relevant queuing system for a fixed amount of time, they construct their log likelihood function using observations and optimise the log likelihood function under stability conditions. Wang *et al.* [32] analyse the maximum likelihood estimates of M/M/c queue with heterogeneous servers under stability conditions. After they give an example for M/M/3 case, they carry over the results to an M/M/c queue structure. Additionally, they provide confidence intervals of estimates and the expected number of customers in the system and the probability of empty system of M/M/c queue. Benes [33] studies a telephone exchange model that has an infinite number of trunks and traffic over the phone hinges on calling rate and mean holding time. Observations that are pertinent to the calling time and hang up time are noted during a finite interval to estimate the parameters employing likelihood function. As a result, he compares various estimators for relevant parameters of the model displaying their means, variances and distributions.

2.3. Multi Queue Server Rate Optimization

2.3.1. Game Theoretic Model

Due to the fact that according to some empirical works in health care and call centres pooling queues may induce operational inefficiencies compared to dedicated queues, Armony *et al.* [34] discuss that this situation may occurs when servers are strategic and display customer ownership which can be splitted into two parts whereas the first one is a point of service which is the case that servers internalise the holding cost of only their customers in service and the second one is the whole system which is the case that they internalise the holding costs of their customers in service and those in queue. They reveal that servers in a pooled queueing system prefer a lower capacity in equilibrium than in a dedicated queueing system utilising the model in which servers' choice of capacity is a non cooperative game. Sivaselvan [35] presents a game theoretic perspective pertaining to stream control mechanisms in multi class networks and provides us with a suitable framework for analysis of control mechanisms. Involving min max routing problem, control mechanism is aimed at deciding over which queue should be selected for the recently arrived customers ensuring the best performance under worst service conditions where service rate in each queue changes with the state of system and unknown to control mechanism Thus, zero sum Markov game is employed. Using value iterations technique, properties related to value of the game can be found. Timmer and Scheinhardt [36] analyze the queuing networks where there are different operators belonging to each queue who may cooperate so as to decrease the amount of waiting time. To get some insight into the case, they start out focusing on the Jackson tandem network where total capacity can be distributed over all queues. They pursue the answer to the question of whether or not the operators of queues will cooperate and how the costs will be shared while cooperation exists. Thus, they also examine networks consisting of two or three nodes. As we mentioned above, Timmer and Scheinhardt [36] incorporate tandem network with two or three nodes and deduce that cooperation is useful, but due to the fact that previous work is devoid of larger tandem networks and general Jackson networks analysis Timmer and Scheinhardt [37] mainly aims to give explicit cost allocation which is suitable for all operators involved in any Jackson network. Wee and Iyer [38] have a game theoretic approach to queuing models with holding costs where consolidation structure for queues exists or not. They deploy two server queuing systems where the choice of service rate, which is based on demand and holding cost allocation scheme arranged by the demand generating entity, is at the discretion of servers. In order to ensure maximum service rate for each of a pooled system where arriving demands are sent to the servers based on the current state of the system and a split system where arriving demands are assigned to one of the two servers based on their capacities, they obtain an optimal holding cost allocation scheme. Gonzalez and Herrero [39] dive into the cost sharing problem in which queues exist while sharing a medical service. They achieve to show that sharing the operating-theatre to treat patients who belong to distinct medical disciplines may result in cost reduction. Therefore, after they manage to calculate optimal fee per procedure related to usage of operating-theatre, they build the conditions under which cooperation among treatments decrease the post-operative costs.

2.3.2. Centralised Decision Making

Caryle *et al.* [40] deal with developing models which ease the burden of how to design systems modelled as parallel M/M/1 queues and how to assign service capacity optimally among the queues to increase the endurance of whole system for the worst case disruptions. They construct the problem as a three level sequential game of information between defender(that is a designated centralised planner to design parallel queues with the knowledge) and attacker(that is worst case disruptions). Cachon and Zhang [41] investigate a queueing model where capacities, processing rates are chosen by servers that are strategic and faster service has more cost. There is a buyer who is at discretion of demand allocation according to servers' performance and more demand is assigned to faster server. Thus, the buyer aims to minimise the average lead time received from the servers. As a result of study, performance-based allocation may be an efficient supply strategy for buyer if the buyer takes the servers' strategic behaviour into consideration. Gilbert and Weng [42] investigate a service network where an agency is just in charge of adjusting constraints on the expected waiting and service time of customers that means agency does not provide customers with actual service but it manages independently operated facilities. Agency should obtain a optimal strategy for the allocation of customers to the self interested operators minimising its own costs. In their paper two different customer allocations is compared, where there is a common queue in the former one and separate queue in the latter one. Lodree *et al.* [43] investigate a queuing system involving heterogeneous teams that work together to serve queues including three distinct prioritisation levels during a mass casualty event. In the paper they consider that the health condition of patients or casualties worsens as time goes on so minimisation of total loss is crucial in the system. Also there is a controller who is in charge of assigning doctors and nurses to relevant queues. Therefore, they aim to optimise the overall service rate of patient queues by assigning an optimal number of assignments of doctors and nurses to the queues. In order to optimise the model they apply heuristic methods and find an efficient method in their simulation studies. Weber and Stidham [44] study the optimal service rate control in some queuing networks where holding cost for a queue is a function of number of customers in that queue, service rate for a queue are function of queue length and arrival rate may or may not be controlled. They obtain a policy which is aimed at minimising the expected total discounted cost and extending their results to an average cost measure they present an example that the optimal policy may not be monotone if choice of server rates at each queue does not include zero. Rosberg *et al.* [45] study tandem queues including two M/M/1 queues where server rate of first queue is a function of both number of customers in first queue and second queue. Main goal in their paper is to find an optimal policy so as to minimise the expected total or average cost in the system. In the paper of Stidham and Weber [46], a large number of models and their results that are pertinent to control of queuing networks are examined under categories consisting of service rate control, admission control, routing and scheduling. They also pay attention to usage of Markov decision models to analyse the structure of optimal control rules. Ahn *et al.* [47] invenstigate the optimal control of two flexible parallel servers located in a two stage tandem queuing network where new arrivals first enter into first station and leave the system after second station. They split up the system into two where the former is called collaborative work forcing servers to be busy with the same job and the latter is non collaborative work forcing servers to be busy with unique jobs. Main problem is how to assign the servers over time to jobs considering both server stations so as to minimise the cost objective. They present sufficient conditions to obtain optimal policy in the both collaborative and non collaborative case. Nevertheless, Markov decision process formulation related to the problem is provided under the average cost optimality criterion demonstrating the existence of solution as well. As a result, some examples are displayed to test the optimal policies. Azaron and Ghomi [48] study the Jackson network ensuring the optimal control of arrival and server rates developing a new model in which the expected value of shortest path and total operating costs of the network are minimised. The queue nodes involved in Jackson network are not generalised, instead only M/M/1 and M/M/infinite queueing systems are included as nodes of the network whereas all other specifications of Jackson network structure are preserved. Converting the problem into a bicriteria optimal control problem, they use goal programming methods to acquire the optimal values. Finally some numerical examples are included in order to test their optimal policies.

2.4. Loss Systems

The final stream of literature involves the articles which are pertinent to loss systems that are widespread particularly in healthcare, queuing, production, inventory, telecommunications, banking, wireless, circuit-switched systems. Blocking behaviour in similar kinds of those systems are examined in a detailed form by various studies. Kobayashi and Mark [49] aim to construct some relationships between queuing networks and loss networks where stationary distributions can be written in terms of product form in both networks demonstrating how some properties derived for queuing networks can be shifted to the loss networks' studies thoroughly. Vinarskiy [50] studies a queuing network structure where there are multi class customers and nodes have M/M/1 queuing system sharing a capacitated waiting space. If an arrival finds a node at its full capacity, then that arrival is lost. Assuming each class input to a node is a Poisson process they obtain approximate analysis based on solving a system of nonlinear equations iteratively. Nevertheless, the existence and uniqueness of the solution found by the iterative algorithm is examined and proven thoroughly. In the paper of Bronschtein and Gertsbakh [51] exponential open queuing network is examined where various exponential servers exist in each node whose waiting space capacity is limited. Thus, arrival to a node is lost when that node is at its full capacity. Assuming total inflow to a node is superposition of external and internal Poisson flow, iterative solution technique based on a system of nonlinear equations constructed in terms of unknown nodal flow rates. They display a Markov chain based technique to obtain an approximate value of average conditional sojourn times both for customers who are done with their service process in the network and for those who are lost ultimately. Fernandez *et al.* [52] studies the economic impact of filtering policies in two stationary loss queuing network location model assuming arrival of requests for service is independent homogeneous Poisson process and service centres consist of finite number of service units. In their model, queue is not allowed and if there is not an idle server upon the recent arrival of request then arrival is lost, also even there is an idle servers system may reject some proportion of requests. Thus, they aim to minimise the overall operating costs finding the coverage of request as well as the locations of service centers. Song and Wu [53] analyse the referral incentive policy in two level healthcare delivery system which is aimed at leading more patients to prefer community health centres during their first treatment. In their study, they answer the questions about that referral incentive policy such as “Does the referral incentive policy really work in the sense of guiding more patients to community health centres?” and “Does the blocking phenomenon in general hospital have an impact on the effectiveness of referral incentive policy and how?” Thus, introducing a utility loss function they construct a queuing network that includes blocking. As a result, they find out that the proportion of patients who are steered to community health centres in the long run reaches a steady level and blocking situation has a significant effect on the proportion. Naumov and Samuilov [54] study the queue network that consists of resource multi server queuing systems having some losses where the occupation of serving of the accepted customer for resource amount is random according to a given distribution function that hinges on customer class and type of necessary servicing. Nevertheless, an arrival of a signal whose distribution is exponential as of the start of servicing may cease the customer servicing in a node and it may be ceased more than once during the customer sojourn in the network. Analytical derivations and equations are obtained so as to compute the joint distribution functions of the number of customers in the nodes and the volumes of resources occupied. Alnowibet and Perros [55] study the blocking probability in a non stationary queuing network involving multi rate loss queues and numerical computation of time varied mean number of customers in that network constructing a scheme of fixed point approximation technique. After they achieve to present how their model can be employed so as to investigate a single or multi class with multi rate loss queues, they extend their method to the investigation of non stationary queuing networks of multi rate loss queue. In conclusion, they make comparisons between exact and simulation results to prove the consistency of their work. In the paper of Ku and Jordan [56], the access control in a target multi server loss queue whose arrival stream consists of both upstream parallel multi server loss queue and stream of new customers. These systems mostly occur in computer and telecommunication networks where continued service to internal customers is favourable to acceptance of new customers. They achieve to maximise total discounted revenue in the system as well as to prove some properties which include monotonicity regarding system parameters. Lam [57] studies the class of queuing networks with state dependent lost arrivals and triggered arrivals that means new customers may enter into the network instantaneously. That class of network is particularly beneficial to model systems with population size constraints. He also finds sufficient conditions to ensure the stationary probability distribution for the network having product form, thus he enlarges the known classes of queuing networks whose stationary probabilities have product form.

Even though we have some common points with the papers mentioned above in our study, some applications we include differ from those papers considerably:

- Simultaneous/Correlated Demands-Arrivals:
  - (i) We include simultaneous arrivals in two different queuing systems where the latter encompasses M/M/1/1 and M/M/1 queues, and the former involves

two M/M/1/1 queues.

(ii) Also, in the system where we have M/M/1/1 and M/M/1 queues, we employ the recursive operations for R matrix as well as the special case of R matrix during the calculation of matrix geometric method.

- Server Rate Optimisation:
  (i) We have two M/M/1/1 queues with simultaneous arrivals where each of those queues aim to maximize their profits according to a given server rate that belongs to another queue.
  (ii) In addition to analysis for the existence of Nash equilibrium, we also prove the concavity and slope of response functions under some assumptions.
- Loss System:
  (i) We study both one loss system consisting of M/M/1/1 and M/M/1 queues and two loss systems consisting of two M/M/1/1 queues.
  (ii) If any capacitated queue is not idle in the queuing system upon arrivals, then simultaneous arrivals are lost.

Table 2.1. Classification of Some Related Works with Chosen Features.

	[3]	**[9]**	**[31]**	**[36]**	**[46]**	**[52]**	**[54]**	**Ours**
Game Analysis	-	-	-	X	-	-	-	X
Server Rate Optimization	-	-	X	X	X	X	-	X
Multiple Queues	-	X	-	X	X	-	X	X
Loss Systems	-	X	-	-	-	X	X	X
Simultaneous Arrivals	X	X	-	-	-	-	-	X

3. TWO LOSS SYSTEMS WITH SIMULTANEOUS ARRIVALS

3.1. Summary of Chapter

In this chapter, we focus on two loss systems consisting of two M/M/1/1 queues where simultaneous arrivals exist with poisson process and servers are exponential distributed. If one of the capacitated queue is not idle in the queuing system upon an arrival to that system, then simultaneous arrivals are lost immediately. We examine the Nash equilibrium of that system where the maximization of profits are aimed for each queue in existence of the fixed arrival rate as well as the slope and concavity of response functions based on server rates under some assumptions.

3.2. Model and Analysis

Modelling the system with Markovian structure we obtain the Figure 3.1 where the state space is (Y_1, Y_2) while Y_1 and Y_2 can only have 0 or 1. We define that Y_1 is the number of customers in the first M/M/1/1 queue while Y_2 is the number of customers in the second M/M/1/1 queue. Thus, we have four different cases in total and we are going to dive into the details of those states in further.

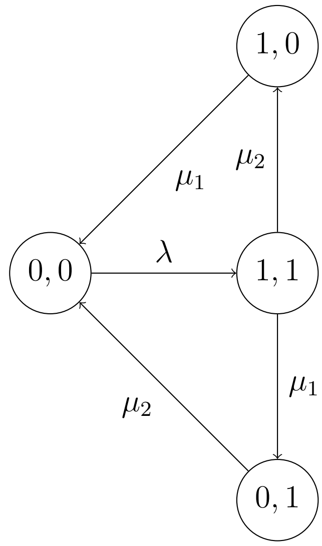

Figure 3.1. Model of Two Loss Systems.

In the model of two loss systems, arrival process is poisson with rate λ whereas server rates are exponential distributed with rate μ_1 and μ_2. Additionally, all states are communicating with each other and due to the fact that we have also a finite number of states stability is automatically ensured. According to Figure 3.1 we also deploy the

$$Q_1 = \begin{array}{c} \\ (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \end{array} \begin{array}{c} \begin{array}{cccc} (0,0) & (0,1) & (1,0) & (1,1) \end{array} \\ \begin{pmatrix} -\lambda & 0 & 0 & \lambda \\ \mu_2 & -\mu_2 & 0 & 0 \\ \mu_1 & 0 & -\mu_1 & 0 \\ 0 & \mu_1 & \mu_2 & -(\mu_1+\mu_2) \end{pmatrix} \end{array}$$

as a generator matrix which enables us to compute the stationary probabilities of states. In order to calculate the stationary probabilities we deduce the equations below according to generator matrix Q_1. Define that P_{00}, P_{01}, P_{10}, P_{11} are the stationary probabilities of states (0,0), (0,1), (1,0), (1,1) respectively. Let's write the balance equations:

$$\lambda P_{00} = \mu_2 P_{01} + \mu_1 P_{10}$$
$$\mu_2 P_{01} = \mu_1 P_{11}$$
$$\mu_1 P_{10} = \mu_2 P_{11}$$
$$(\mu_2 + \mu_1) P_{11} = \lambda P_{00}.$$

Using the equations above, we find the probabilities as:

$$P_{00} = \frac{\mu_1 \mu_2 (\mu_1 + \mu_2)}{\mu_1^2 \lambda + \mu_2^2 \lambda + \mu_1^2 \mu_2 + \mu_2^2 \mu_1 + \lambda \mu_1 \mu_2}$$
$$P_{01} = \frac{\lambda \mu_1^2}{\mu_1^2 \lambda + \mu_2^2 \lambda + \mu_1^2 \mu_2 + \mu_2^2 \mu_1 + \lambda \mu_1 \mu_2}$$
$$P_{10} = \frac{\lambda \mu_2^2}{\mu_1^2 \lambda + \mu_2^2 \lambda + \mu_1^2 \mu_2 + \mu_2^2 \mu_1 + \lambda \mu_1 \mu_2}$$
$$P_{11} = \frac{\lambda \mu_1 \mu_2}{\mu_1^2 \lambda + \mu_2^2 \lambda + \mu_1^2 \mu_2 + \mu_2^2 \mu_1 + \lambda \mu_1 \mu_2}.$$

Defining π_1 is the profit amount coming from first queue whereas π_2 is coming from second queue. Nevertheless while r_1 and r_2 are revenues per unit of time generated by first and second queue respectively, c_1 and c_2 are the costs incurred per unit of time, thus we can write profit functions as

$$\pi_1 = r_1 \mu_1 (P_{11} + P_{10}) - c_1 \mu_1$$
$$\pi_2 = r_2 \mu_2 (P_{11} + P_{01}) - c_2 \mu_2.$$

Here, we assume that $\frac{c_1}{r_1} < 1$ and employing some algebraic manipulation we write

$$\pi_1 = r_1 \lambda P_{00} - c_1 \mu_1$$

where

$$P_{00} = 1 - \frac{B}{B+C}$$

while

$$B = \lambda\mu_1^2 + \lambda\mu_2^2 + \lambda\mu_1\mu_2$$

and

$$C = \mu_1^2\mu_2 + \mu_1\mu_2^2.$$

Proposition 3.1. *Define $P_{block} = \frac{B}{B+C}$, then P_{block} is strictly convex in μ_1 and π_1 is strictly concave in μ_1 for each given $\mu_2 > 0$.*

Proof. Taking the derivative of P_{block} what we have is:

$$\frac{\partial P_{block}}{\partial \mu_1} = \frac{B_1C - BC_1}{(B+C)^2}$$

where

$$B_1 = \frac{\partial B}{\partial \mu_1}$$

and

$$C_1 = \frac{\partial C}{\partial \mu_1}.$$

Also define

$$D = B_1C - BC_1 = -2\lambda\mu_1\mu_2^3 - \lambda\mu_2^4.$$

Thus we have:

$$\frac{\partial^2 P_{block}}{\partial \mu_1^2} = \frac{D_1}{(B+C)^3}(-(B+C) + (B_1 + C_1)(2\mu_1 + \mu_2))$$

where

$$D_1 = -2\lambda\mu_2^3 = \frac{\partial D}{\partial \mu_1}.$$

So,

$$\frac{\partial^2 P_{block}}{\partial \mu_1^2} > 0$$

that means

$$\frac{\partial^2 \pi_1}{\partial \mu_1^2} < 0.$$

Therefore we can deduce that there is a unique solution to $\frac{\partial \pi_1}{\partial \mu_1} = 0$ for each given $\mu_2 > 0$. Additionally, using similar operations it can be shown that $\frac{\partial^2 \pi_2}{\partial \mu_2^2}$ is also strictly concave for each given $\mu_1 > 0$. □

Proposition 3.2. *Steady state probabilities of system consisting of two independent M/M/1/1 queues including same parameters may differ from the stationary probabilities we calculate for our current system.*

Proof. Let's examine the case of P_{01} where the first queue is idle while second queue is at full capacity. Remember for our current system

$$P_{01} = \frac{\lambda\mu_1^2}{\mu_1^2\lambda + \mu_2^2\lambda + \mu_1^2\mu_2 + \mu_2^2\mu_1 + \lambda\mu_1\mu_2}$$

and now recalling the formulas for stationary probabilities belong to M/M/c/K queues we can infer that

$$P_{01} = \frac{\lambda\mu_1}{(\lambda+\mu_1)(\lambda+\mu_2)}$$

for the system consisting of two independent M/M/1/1 queues. Now assume that those two P_{01} values are equal to each other having same parameters. Thus we obtain

$$\frac{\lambda\mu_1}{(\lambda+\mu_1)(\lambda+\mu_2)} = \frac{\lambda\mu_1^2}{\mu_1^2\lambda+\mu_2^2\lambda+\mu_1^2\mu_2+\mu_2^2\mu_1+\lambda\mu_1\mu_2}$$

that means

$$\lambda^2\mu_1 - \mu_2^2\lambda - \mu_2^2\mu_1 = 0$$

after some operations. We can observe that for a given μ_1 and μ_2 we have a second degree polynomial with respect to λ. Thus, using the root formulas belong to second order polynomial λ must have a positive root for given positive μ_1 and μ_2. This situation implies that just one λ value ensure the equivalence between two P_{01} probabilities mentioned above. In short, if we have an two independent M/M/1/1 queues in our system, then stationary probabilities may differ from the probabilities belong to our current system. $\square$

3.3. A Capacity Game

During the capacity game we consider that there is a fixed λ but servers' rates can change due to the fact that each queue aims to maximise their profits. Also, in our decentralised system parameters of a player such as cost, revenue and server rate are known by another player. Let's set

$$\lambda\mu_2 = a$$
$$\lambda\mu_2^2 = b$$
$$\frac{c_1}{r_1} = h_1$$
$$\frac{c_1}{r_1} = h_2$$

where h_1 and h_2 are strictly less than 1, in order to reduce some complex appearances. Now because of the maximization of profits we need to have $\frac{\partial\pi_1}{\partial\mu_1} = 0$ and $\frac{\partial\pi_2}{\partial\mu_2} = 0$. After some operations, what we obtain is

$$\begin{aligned}\frac{\partial\pi_1}{\partial\mu_1} = 0 \implies & (\lambda^2 + \mu_2^2 + 2\lambda\mu_2)h_1\mu_1^4 + (2\lambda\mu_2^2 + 2\mu_2^3 + 2\lambda a + 2a\mu_2)h_1\mu_1^3 + \\ & (\mu_2^4 + a^2 + 2\lambda b + 2b\mu_2 + 2a\mu_2^2)h_1\mu_1^2 + \\ & [(2b\mu_2^2 + 2ba)h - 2ba]\mu_1 - b^2 + h_1 b^2 = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial\pi_2}{\partial\mu_2} = 0 \implies & (\lambda^2 + \mu_1^2 + 2\lambda\mu_1)h_2\mu_2^4 + (2\lambda\mu_1^2 + 2\lambda a + 2\mu_1^3 + 2\mu_1 a)h_2\mu_2^3 + \\ & (\mu_1^4 + a^2 + 2a\lambda + 2a\mu_1 + 2a\mu_1^2)h_2\mu_2^2 + \\ & [(2b\mu_1^2 + 2ab)h_2 - 2ab]\mu_2 + b^2 h_2 - b^2 = 0.\end{aligned}$$

In short what we have is

$$\begin{aligned}& (\lambda^2 + \mu_2^2 + 2\lambda\mu_2)h_1\mu_1^4 + (2\lambda\mu_2^2 + 2\mu_2^3 + 2\lambda a + 2a\mu_2)h_1\mu_1^3 + \\ & (\mu_2^4 + a^2 + 2\lambda b + 2b\mu_2 + 2a\mu_2^2)h_1\mu_1^2 + \\ & [(2b\mu_2^2 + 2ba)h_1 - 2ba]\mu_1 - b^2 + h_1 b^2 = 0\end{aligned} \tag{3.1}$$

$$\begin{aligned}&(\lambda^2+\mu_1^2+2\lambda\mu_1)h_2\mu_2^4+(2\lambda\mu_1^2+2\lambda a+2\mu_1^3+2\mu_1 a)h_2\mu_2^3+\\&(\mu_1^4+a^2+2a\lambda+2a\mu_1+2a\mu_1^2)h_2\mu_2^2+\\&[(2b\mu_1^2+2ab)h_2-2ab]\mu_2+b^2h_2-b^2=0.\end{aligned} \tag{3.2}$$

We can observe that if $\mu_2 = 0$, then μ_1 must be 0 in the Equation (3.1) and if $\mu_2 > 0$, then zero value of μ_1 can not satisfy the Equation (3.1) but there is only one positive value root for μ_1 by Descartes rule of sign. Thus we actually have a function such as $\mu_1(\mu_2)$ which indicates an optimal value of μ_1 for a given μ_2 in our system. The same operations can also be applied for the Equation (3.2) and we can obtain $\mu_2(\mu_1)$ structure.

Now let's focus on π_1 case. We know that μ_1 can be written as a smooth function of μ_2 again by Descartes rule of sign where the smoothness comes from the fact that for each given positive μ_2 we derive μ_1 values from the unique positive root of polynomial in the Equation (3.1). While we have a fixed λ and are approaching to any positive μ_2 from both left and right, we are actually approaching to same value of the coefficients belong to the Equation (3.1) that means during the limitation with respect to μ_2 on both sides we reach same polynomial structure and we obtain same unique positive root for μ_1 satisfying the Equation (3.1). Thus we can infer that polynomial structure ensure the smoothness of $\mu_1(\mu_2)$ where $\mu_2 \geq 0$.

Proposition 3.3. *$\frac{d\mu_1}{d\mu_2}$ is always positive while $\mu_2 > 0$ if $\frac{d\mu_1}{d\mu_2}$ is positive in the vicinity of zero and at zero point.*

Proof. Let's assume that $\frac{d\mu_1}{d\mu_2}$ be positive in the vicinity of zero and at zero point, now we can express that if $\frac{d\mu_1}{d\mu_2}$ is negative at some point while $\mu_2 > 0$, then $\frac{d\mu_1}{d\mu_2}$ must be equal to zero at some point while $\mu_2 > 0$. Now examine whether or not $\frac{d\mu_1}{d\mu_2}$ is zero. So, first assume that $\frac{d\mu_1}{d\mu_2}$ is zero. Also employing implicit differentiation we can write

$$\frac{d\mu_1}{d\mu_2} = -\frac{\frac{\partial^2\pi_1}{\partial\mu_1\partial\mu_2}}{\frac{\partial^2\pi_1}{\partial\mu_1^2}}. \tag{3.3}$$

So, if $\frac{d\mu_1}{d\mu_2}$ is zero at some point of $\mu_2 > 0$, then

$$\frac{\partial^2 \pi_1}{\partial \mu_1 \partial \mu_2} = 0.$$

According to the fact that $\frac{\partial^2 \pi_1}{\partial \mu_1 \partial \mu_2}$ is zero what we have is

$$\begin{aligned}
&(2\mu_2 + 2\lambda)h_1\mu_1^4 + (4\lambda\mu_2 + 6\mu_2^2 + 2\lambda^2 + 4\lambda\mu_2)h_1\mu_1^3 + \\
&(4\mu_2^3 + 2\lambda^2\mu_2 + 4\lambda^2\mu_2 + 12\lambda\mu_2^2)h_1\mu_1^2 + \\
&[(8\lambda\mu_2^3 + 6\lambda^2\mu^2)h_1 - 6\lambda^2\mu_2^2]\mu_1 - 4\lambda^2\mu_2^3(1 - h_1) = 0.
\end{aligned} \tag{3.4}$$

Now multiply the Equation (3.1) with 3 and multiply the Equation (3.4) with μ_2 both on sides and subtraction of the Equation (3.1) from the Equation (3.4) yields

$$\begin{aligned}
&(-3\lambda^2 - \mu_2^2 - 4\lambda\mu_2)h_1\mu_1^4 + (-4\lambda\mu_2^2 - 4\lambda^2\mu_2)h_1\mu_1^3 + \\
&(\mu_2^4 - 3\lambda^2\mu_2^2)h_1\mu_1^2 + 2\lambda\mu_2^4 h_1\mu_1 - \lambda^2\mu_2^4(1 - h_1) = 0.
\end{aligned} \tag{3.5}$$

Recall that we only consider $\mu_2 > 0$, so coefficients of the Equation (3.5) can only be (-,-,+,+,-) or (-,-,-,+,-) or (-,-,0,+,-) that means total sign change is always 2 and according to Descartes sign rule for a fixed $\mu_2 > 0$ the Equation (3.5) either have no positive root or two positive roots. Let's examine those two different cases:

- Case 1 (no positive root): Recall that if $\mu_2 > 0$, then satisfying the Equation (3.1) μ_1 must be a unique positive root of it, but according to the Case 1 we should not have positive root, so there is a contradiction and Case 1 is wrong.
- Case 2 (two positive root): Recall that if $\mu_2 > 0$, then satisfying the the Equation (3.1) μ_1 must be a unique positive root of it because Descartes sign rule does not allow multiple positive roots, but according to the Case 2 we should have multiple positive roots, so there is again a contradiction and Case 2 is wrong.

To sum up, it is wrong to assume that $\frac{d\mu_1}{d\mu_2}$ is equal to zero at some point while $\mu_2 > 0$, therefore $\frac{d\mu_1}{d\mu_2}$ only have positive values for all $\mu_2 > 0$ if we assume that $\frac{d\mu_1}{d\mu_2}$ be positive in the vicinity of zero and at zero point. Now let's check that $\frac{d\mu_1}{d\mu_2}$ is positive in the vicinity of zero and at zero point. □

Proposition 3.4. *$\frac{d\mu_1}{d\mu_2}$ is positive when μ_2 is zero or μ_2 is in the vicinity of zero.*

Proof. Taking limit on both sides of the Equation (3.3) whereas $\mu_2 \to 0$, we have

$$\lim_{\mu_2 \to 0} -\frac{d\mu_1}{d\mu_2} = \frac{\lim_{\mu_2 \to 0} \frac{\partial^2 \pi_1}{\partial \mu_1 \partial \mu_2}}{\lim_{\mu_2 \to 0} \frac{\partial^2 \pi_1}{\partial \mu_1^2}}.$$

During the limitation we can consider the taylor series expansion of μ_1 and employ

$$\mu_1(\mu_2)\big|_{\mu_2=0^+} = \mu_1(\mu_2)\big|_{\mu_2=0} + \left.\frac{d\mu_1}{d\mu_2}\right|_{\mu_2=0} (\mu_2 - 0).$$

We know

$$\mu_1(\mu_2 = 0) = 0.$$

So, take μ_1 as $\left.\frac{d\mu_1}{d\mu_2}\right|_{\mu_2=0} \mu_2$ during the limitation. In order to reduce the complex visualisation, take

$$x = \left.\frac{d\mu_1}{d\mu_2}\right|_{\mu_2=0}.$$

Thus we have

$$\lim_{\mu_2 \to 0} -\frac{d\mu_1}{d\mu_2} = \frac{2\lambda^2 h_1 \mu_2^3 x^3 + 6\lambda^2 h_1 x^2 \mu_2^3 + (1-h_1)(-(4\lambda^2 + 6\lambda^2 x)\mu_2^3)}{6\lambda^2 h_1 x \mu_2^3 + 6\lambda^2 h_1 x^2 \mu_2^3 + 4\lambda^2 h_1 x^3 \mu_2^3 + (1-h_1)(-2\lambda^2 \mu_2^3)}.$$

Now observe that $-\frac{d\mu_1}{d\mu_2}$ is actually limiting to $-x$ while μ_2 is approaching to zero. If x equals to zero, then we have:

$$\lim_{\mu_2 \to 0} -\frac{d\mu_1}{d\mu_2} = 0 = \frac{(1-h_1)(-4\lambda^2 \mu_2^3)}{(1-h_1)(-2\lambda^2 \mu_2^3)} = 2$$

We can easily see that there is a contradiction because we have

$$0 = 2$$

assuming

$$\left.\frac{d\mu_1}{d\mu_2}\right|_{\mu_2=0} = x = 0.$$

So, this assumption is wrong. Also we can not have a negative value of $\left.\frac{d\mu_1}{d\mu_2}\right|_{\mu_2=0}$ because this would imply that when μ_2 increases starting at zero then μ_1 has a negative value, but we know $\mu_1(\mu_2 > 0)$ must be positive due to Descartes rule of sign satisfying the Equation (3.1). In short, $\left.\frac{d\mu_1}{d\mu_2}\right|_{\mu_2=0}$ can only be a positive number. Now let us examine the behaviour of $\frac{d^2\mu_1}{d\mu_2^2}$ as we have done for $\frac{d\mu_1}{d\mu_2}$ above. □

Proposition 3.5. *When we have $\mu_2 \to +\infty$, μ_1 has an asymptotic positive value.*

Proof. If $\mu_2 \to +\infty$, then the queue with second server behaves like a M/M/1/1 queue, so π_1 becomes $r_1\mu_1(\frac{\lambda}{\lambda+\mu_1}) - c_1\mu_1$ using M/M/c/K formulas for stationary probabilities. Let's take the first and second derivatives of π_1 as

$$\frac{d\pi_1}{d\mu_1} = \frac{\lambda r_1}{\lambda + \mu_1} - \frac{\lambda r_1 \mu_1}{(\lambda + \mu_1)^2} - c_1$$

$$\frac{d^2\pi_1}{d\mu_1^2} = \frac{-2\lambda r_1}{(\lambda + \mu_1)^2} + \frac{2\lambda r_1 \mu_1}{(\lambda + \mu_1)^3}.$$

We can observe that π_1 has an concave function with respect to μ_1 when μ_2 is at infinite, so we already know $\mu_1(\mu_2 > 0)$ must be positive by Descartes rule of sign regarding (3.1) and we have just found that μ_1 take a unique positive value when μ_2 is at infinite that means whereas μ_2 is going to infinite μ_1 must converge to a positive value. □

Proposition 3.6. *If* $\frac{d^2\mu_1}{d\mu_2^2}$ *is positive at some point while* $\mu_2 \geq 0$*, then* $\frac{d^2\mu_1}{d\mu_2^2}$ *must be positive for all* $\mu_2 \geq 0$*, or If* $\frac{d^2\mu_1}{d\mu_2^2}$ *is negative at some point while* $\mu_2 \geq 0$*, then* $\frac{d^2\mu_1}{d\mu_2^2}$ *must be negative for all* $\mu_2 \geq 0$*.*

Proof. Let's recall that μ_1 can be written as a smooth function of μ_2 again by Descartes rule of sign where the smoothness comes from the fact that for each given positive μ_2 we derive μ_1 values from the unique positive root of polynomial in the Equation (3.1). While we have a fixed λ and are approaching to any positive μ_2 from both left and right, we are actually approaching to same value of the coefficients belong to the Equation (3.1) that means during the limitation with respect to μ_2 on both sides we reach same polynomial structure and we obtain same unique positive root for mu_1 satisfying the Equation (3.1). Thus we can infer that polynomial structure ensure the smoothness of $\mu_1(\mu_2)$ where $\mu_2 \geq 0$, that is why we can include the first and second derivative of $\mu_1(\mu_2)$. Now let's glance at the Equation (3.3) and take ordinary derivative on both sides regarding μ_2, what we obtain is

$$\frac{\partial^2 \mu_1}{\partial \mu_2^2} = -\left(\frac{\frac{\partial^2 \pi_1}{\partial \mu_1 \partial \mu_2}' \frac{\partial^2 \pi_1}{\partial \mu_1^2}}{(\frac{\partial^2 \pi_1}{\partial \mu_1^2})^2} - \frac{\frac{\partial^2 \pi_1}{\partial \mu_1 \partial \mu_2} \frac{\partial^2 \pi_1}{\partial \mu_1^2}'}{(\frac{\partial^2 \pi_1}{\partial \mu_1^2})^2}\right).$$

If we assume that $\frac{d^2\mu_1}{d\mu_2^2}$ can be negative at some point for $\mu_2 \geq 0$ while $\frac{d^2\mu_1}{d\mu_2^2}$ can also be positive at other point for $\mu_2 \geq 0$, then $\frac{d^2\mu_1}{d\mu_2^2}$ must be equal to zero at some point $\mu_2 \geq 0$. Define

$$X = \frac{\partial^2 \pi_1}{\partial \mu_1 \partial \mu_2}$$

$$Y = \frac{\partial^2 \pi_1}{\partial \mu_1^2}$$

$$Z = \frac{d\mu_1}{d\mu_2}.$$

Thus, we have

$$-\frac{X}{Y} = Z$$

and due to our assumption

$$X'Y - XY' = 0$$

must be satisfied. So we can write

$$\frac{X}{Y} = \frac{X'}{Y'} = \frac{\frac{dX}{d\mu_2}}{\frac{dY}{d\mu_2}} = \frac{dX}{dY}$$

Using separation of variables technique (here we take the constant as zero as an example but whatever the constant is the division of X and Y is always positive) we find

$$\log_e X = \log_e Y \implies \log_e(\frac{X}{Y}) = 0$$

and this situation requires

$$X = Y$$

but this implies that

$$-\frac{X}{Y} = Z = -1$$

that is not meaningful because we have already proven that Z is always positive for $\mu_2 \geq 0$ in the Proposition 3.3 and the Proposition 3.4. In short we have a contradiction, and $\frac{d^2\mu_1}{d\mu_2^2}$ can not be zero for any $\mu_2 \geq 0$, thus we have proven our current proposition. $\square$

Proposition 3.7. $\frac{d^2\mu_1}{d\mu_2^2}$ *is negative for all* $\mu_2 \geq 0$.

Proof. Depending on the proof in the Proposition 3.6 we know $\frac{d^2\mu_1}{d\mu_2^2}$ can not be zero while $\mu_2 \geq 0$, so if we assume that $\frac{d^2\mu_1}{d\mu_2^2}$ is positive for $\mu_2 \geq 0$ then $\mu_1(\mu_2 \to +\inf)$ has an infinite value due to convexity which contradicts with the Proposition (3.5). Therefore $\frac{d^2\mu_1}{d\mu_2^2}$ must be negative for all $\mu_2 \geq 0$. □

In short, we have proven that $\frac{d\mu_1}{d\mu_2}$ is positive whereas $\frac{d^2\mu_1}{d\mu_2^2}$ is negative for all $\mu_2 \geq 0$. Now it is time to examine the Nash equilibrium for our system.

Proposition 3.8. *We have at least one Nash equilibrium point in our system.*

Proof. We already know that:
$\pi_1(\mu_1, \mu_2)$ and $\pi_2(\mu_1, \mu_2)$ are concave functions with respect to μ_1 and μ_2 respectively as proved in the Proposition 3.1. Also, $\mu_1 \in [0, \mu_1(+\infty)]$ and $\mu_2[0, \mu_2(+\infty)]$. Therefore we can conclude that there is at least one Nash equilibrium by standard theorem [58] in game theory. □

We can also see that $(0, 0)$ point satisfies the Nash equilibrium condition in our system due to the fact that

$$\pi_1(\mu_1, \mu_2 = 0) = -c_1\mu_1$$

and

$$\pi_2(\mu_1, \mu_2 = 0) = -c_2\mu_2$$

equations has optimal profits for $(\mu_1 = 0, \mu_2 = 0)$ point, so we have

$$\mu_1(\mu_2 = 0) = 0$$

and

$$\mu_2(\mu_1 = 0) = 0.$$

But we wonder if there may be multiple Nash equilibrium points?

Proposition 3.9. *In a symmetric game where $r_1 = r_2 = r$ and $c_1 = c_2 = c$, $(\mu_1(\mu_2 = 0), \mu_2(\mu_1 = 0)) = (0,0)$ is the only non-negative equilibrium if $r \leq 3c$. Otherwise, the only other non-negative equilibrium point that is given by $(\mu*, \mu*)$ where $\mu* = \frac{3\lambda}{2}(\sqrt{\frac{r}{3c}} - 1)$.*

Proof. We know that the two following equations must be satisfied which are

$$\frac{\partial \pi_1(\mu_1, \mu_2)}{\partial \mu_1} = 0$$

$$\frac{\partial \pi_2(\mu_1, \mu_2)}{\partial \mu_2} = 0.$$

Two equations above yield

$$r\lambda \frac{2\lambda\mu_1\mu_2^3 - \lambda\mu_2^4}{(\lambda\mu_1^2 + \lambda\mu_2^2 + \lambda\mu_1\mu_2 + \mu_1^2\mu_2 + \mu_2^2\mu_1)^2} = c$$

$$r\lambda \frac{2\lambda\mu_2\mu_1^3 - \lambda\mu_1^4}{(\lambda\mu_1^2 + \lambda\mu_2^2 + \lambda\mu_1\mu_2 + \mu_1^2\mu_2 + \mu_2^2\mu_1)^2} = c.$$

Due to symmetry at optimality we have

$$\mu_1 = \mu_2 = \mu$$

therefore we get

$$r\lambda \frac{2\lambda\mu^4 - \lambda\mu^4}{(\lambda\mu^2 + \lambda\mu^2 + \lambda\mu^2 + \mu^3 + \mu^3)^2} = c.$$

We can write the equation above as:

$$4c\mu^2 + 12\lambda c\mu + 3\lambda^2(3c - r) = 0.$$

According to second order polynomial with respect to μ, roots are: $\frac{-3\lambda}{2} + \lambda\sqrt{\frac{3r}{4c}}$ and $\frac{-3\lambda}{2} - \lambda\sqrt{\frac{3r}{4c}}$. For $\mu \geq 0$ we have

$$\mu = \frac{-3\lambda}{2} + \lambda\sqrt{\frac{3r}{4c}}.$$

But for $r \leq 3c$ we have $\mu \leq 0$ as a solution. $\square$

Additionally, we can encompass similar works in above propositions so as to analyse π_2, $\frac{d\mu_2}{d\mu_1}$ and $\frac{d^2\mu_2}{d\mu_1^2}$ whereas we have $\mu_2(\mu_1)$ in contrast to $\mu_1(\mu_2)$. In a nutshell, we can infer that we may have multiple or only one Nash equilibrium point for our system where there is always a strict concavity and positive slope for response functions demonstrating an asymptotic convergence at infinity.

3.4. The Centralised Model

In this section we aim to provide some explicit explanations and demonstrate why centralised model has a great potential to be better and more profitable than decentralised model by Nash equilibrium. Now let us give the model pertinent to centralised structure where $r_1 = r_2 = r$ and $c_1 = c_2 = c$. We know our initial model is

$$max(\pi_1 + \pi_2)$$

$$st : \mu_1, \mu_2 \geq 0.$$

However, using equal revenues and costs we can convert the model into

$$
\begin{aligned}
& max(2r\lambda P_{00} - \mu_1 c - \mu_2 c) \\
& st : \mu_1, \mu_2 \geq 0.
\end{aligned}
$$

According to model, there are some resemblances with symmetric game we have discussed in the Proposition 3.9 but in spite of the fact that both the centralised and decentralised models have same objectives, their feasible regions are different because in the centralised model $\mu_1 \geq 0$ and $\mu_2 \geq 0$ are independent on each other but in the decentralised model we need have some connection between $\mu_1 \geq 0$ and $\mu_2 \geq 0$ on a functional basis recalling that μ_1 is a unique positive root of the Equation (3.1) by Descartes rule while $\mu_2 > 0$. Therefore the feasible space of the centralised model is larger than the decentralised game and optimal point of the centralised model is larger than or equal to the decentralised model.

Now examine the hessian matrix of objective function in centralised model setting $2r\lambda P_{00} - \mu_1 c - \mu_2 c = M$ and what we have is

$$
\nabla^2 M = \begin{pmatrix} H_{1,1} & H_{1,2} \\ H_{2,1} & H_{2,2} \end{pmatrix}
$$

$$
H_{1,1} = \frac{2r\lambda 2\lambda\mu_2^3(-3\mu_1^2\mu_2 - 3\lambda\mu_1^2 - 3\mu_1\mu_2^2 - 3\lambda\mu_1\mu_2 - \mu_2^3)}{(\lambda\mu_1^2 + \lambda\mu_2^2 + \lambda\mu_1\mu_2 + \mu_1^2\mu_2 + \mu_2^2\mu_1)^3}
$$

$$
H_{1,2} = H_{2,1} = \frac{2r\lambda 2\lambda\mu_1^2\mu_2^2(\mu_1\mu_2 + 3\lambda\mu_1 + 3\lambda\mu_2)}{(\lambda\mu_1^2 + \lambda\mu_2^2 + \lambda\mu_1\mu_2 + \mu_1^2\mu_2 + \mu_2^2\mu_1)^3}
$$

$$
H_{2,2} = \frac{2r\lambda 2\lambda\mu_1^3(-\mu_1^3 - 3\mu_1^2\mu_2 - 3\mu_1\mu_2^2 - 3\lambda\mu_1\mu_2 - 3\lambda\mu_2^2)}{(\lambda\mu_1^2 + \lambda\mu_2^2 + \lambda\mu_1\mu_2 + \mu_1^2\mu_2 + \mu_2^2\mu_1)^3}.
$$

Proposition 3.10. *Hessian matrix of M is a negative semi definite matrix while $\mu_1 \geq 0$ and $\mu_2 \geq 0$.*

Proof. If hessian matrix of M is a negative semi definite matrix, then we have a concave objective function in the centralised model. Therefore we need to analyse $\vec{\mu}^T M \vec{\mu}$ where $\vec{\mu}$ is an arbitrary vector whose dimensions are μ_1 and μ_2.

$$\begin{aligned}\vec{\mu}^T M \vec{\mu} &= \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}\begin{pmatrix}H_{1,1} & H_{1,2}\\ H_{2,1} & H_{2,2}\end{pmatrix}\begin{pmatrix}\mu_1 & \mu_2\end{pmatrix}\\ &= \mu_1^2 H_{1,1} + 2\mu_1\mu_2 H_{1,2} + \mu_2^2 H_{2,2}\end{aligned}$$

where we have

$$\begin{aligned}\mu_1^2 H_{1,1} &= 2r\lambda(-6\lambda\mu_1^4\mu_2^4 - 6\lambda^2\mu_1^4\mu_2^3 - 6\lambda\mu_1^3\mu_2^5 - 6\lambda^2\mu_1^3\mu_2^4 - 2\lambda\mu_2^2\mu_2^6)\\ 2\mu_1\mu_2 H_{1,2} &= 2r\lambda(4\lambda\mu_1^4\mu_2^4 + 12\lambda^2\mu_1^4\mu_2^3 + 12\lambda^2\mu_1^3\mu_2^4)\\ \mu_2^2 H_{2,2} &= 2r\lambda(-2\lambda\mu_1^6\mu_2^2 - 6\lambda\mu_1^5\mu_2^3 - 6\lambda\mu_1^4\mu_2^4 - 6\lambda^2\mu_1^4\mu_2^3 - 6\lambda^2\mu_1^3\mu_2^4).\end{aligned}$$

As we can see that while $\mu_1 \geq 0$ and $\mu_2 \geq 0$, $\vec{\mu}^T M \vec{\mu}$ is lower than or equal to zero that means concavity. □

Proposition 3.11. *For the centralised model above, the optimal value may be strictly greater than the optimal value of the decentralised model with symmetric game in the Proposition 3.9 where $r > 3c$.*

Proof. We can write an equivalent model for our centralised model as

$$\begin{aligned}&min(-2r\lambda P_{00} + \mu_1 c + \mu_2 c)\\ &st : \mu_1, \mu_2 \geq 0.\end{aligned}$$

Initially, let's reduce the feasible region of our centralised model adding $\mu = \mu_1 = \mu_2 \geq 0$. Thus, by adding a little bit manipulation we can convert our centralised model into

$$\begin{aligned}&min(2c\mu - 2r\lambda P_{00})\\ &st : -\mu \leq 0.\end{aligned}$$

Also observe that the objective function of the centralised model above is convex in feasible region. In addition, our constraint is linear that means Slater condition is satisfied, so construct the Lagrange dual function with the Lagrange multiplier θ where

$$L(\theta, \mu) = 2c\mu - 2r\lambda(1 - \frac{3\lambda\mu^2}{3\lambda\mu^2 + 2\mu^3}) - \theta\mu.$$

We can express that the optimal value of the Lagrange dual function is equal to the optimal value of model above due to the fact that both the convex structure of model and Slater condition are satisfied. Now let's write the optimal conditions of model using the Lagrange dual function

$$\begin{aligned}\frac{\partial L}{\partial \mu} &= 2c - 2r\lambda\frac{3\lambda\mu^4}{(3\lambda\mu^2 + 2\mu^3)^2} - \theta \\ &= 0 \\ \theta\mu &= 0 \\ \theta &\geq 0 \\ \mu &\geq 0.\end{aligned}$$

Applying some operations we can have

$$8c\mu^2 + 24\lambda c\mu + 18\lambda^2 c - 6\lambda^2 r - 9\lambda^2\theta = 0.$$

Actually we have a second order polynomial with respect to μ as $a\mu^2 + b\mu + c = 0$, so in order to find the roots we need discriminant that is:

$$\begin{aligned}b^2 - 4ac =& 576\lambda^2c^2 - 4(144\lambda^2c^2 - 48\lambda^2 cr - 72\lambda\theta) \\ =& 576\lambda^2c^2 - 576\lambda^2c^2 + 192\lambda^2 cr + 288\lambda\theta\end{aligned}$$

Due to $r > 3c$ we have

$$-576\lambda^2c^2 + 192\lambda^2c^2 r \geq 0.$$

Thus, the magnitude of the discriminant is guaranteed to be larger than b in the polynomial, therefore we have only one positive root that is

$$\frac{-3\lambda}{2} + \sqrt{\frac{3\lambda^2 r}{4c} + \frac{72\lambda\theta}{196c}}.$$

As we can observe that if $\theta > 0$, then $\theta\mu$ is always strictly greater than zero, but if $\theta = 0$ then we satisfy $\theta \geq 0$, $\mu \geq 0$, $\theta\mu = 0$. So, we have actually reached the optimal point and value for the centralised model where $\mu_1 = \mu_2 = \mu$. The positive root we have just found is totally same with the positive root we have found in the Proposition 3.9 that includes symmetric game. In short, even if we reduce the feasible region of the centralised model we obtain the same optimality with the decentralised model that includes symmetric game. Thus, the optimality of the centralised model may be strictly larger than the decentralised model that involves Nash equilibrium solution with symmetric game. □

3.5. Numerical Results

In this section we aim to present our work we have done so far for capacity game chapter on some visual basis and we actually confirm our analytical results using some simulations.

In the Figure 3.2 we have two distinct Nash equilibrium points where there is a symmetric game involving $r_1 = r_2 = r$, $c_1 = c_2 = c$ and $r > 3c$ as we have proved in the Proposition 3.9. However, in the Figure 3.3 we have only one Nash equilibrium points where there is a symmetric game involving $r_1 = r_2 = r$, $c_1 = c_2 = c$ and $r \leq 3c$ as we have again proved in the Proposition 3.9. Nevertheless, in the Figure 3.4 we can observe that we have two distinct Nash equilibrium points where there is a unsymmetric game involving $r_1 \neq r_2$, $c_1 \neq c_2$.

In addition, we should pay attention to the fact that $\mu_1(\mu_2)$ and $\mu_2(\mu_1)$ have concave structures where their slopes are always positive and asymptotically converge to a positive values as we have proven in the the Proposition 3.3, 3.4, 3.5, 3.6 and 3.7.

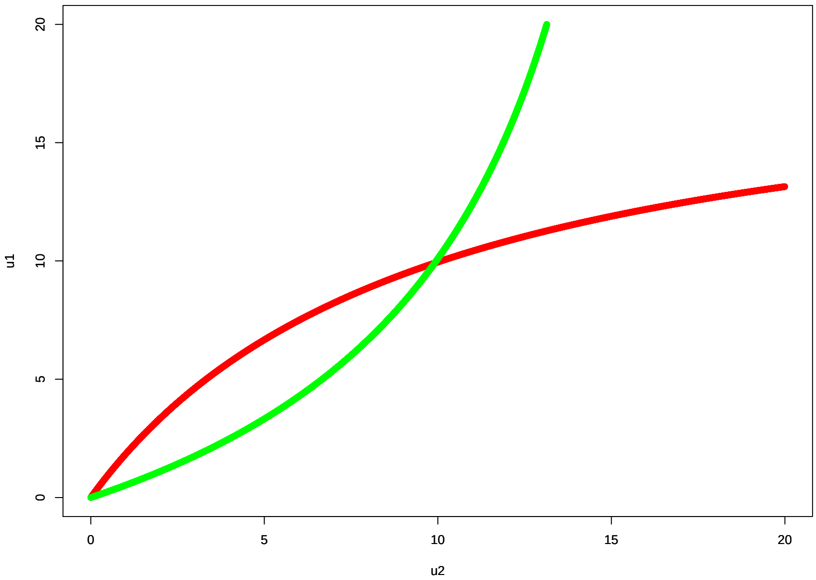

Figure 3.2. Symmetric Game with Two Nash Equilibriums.

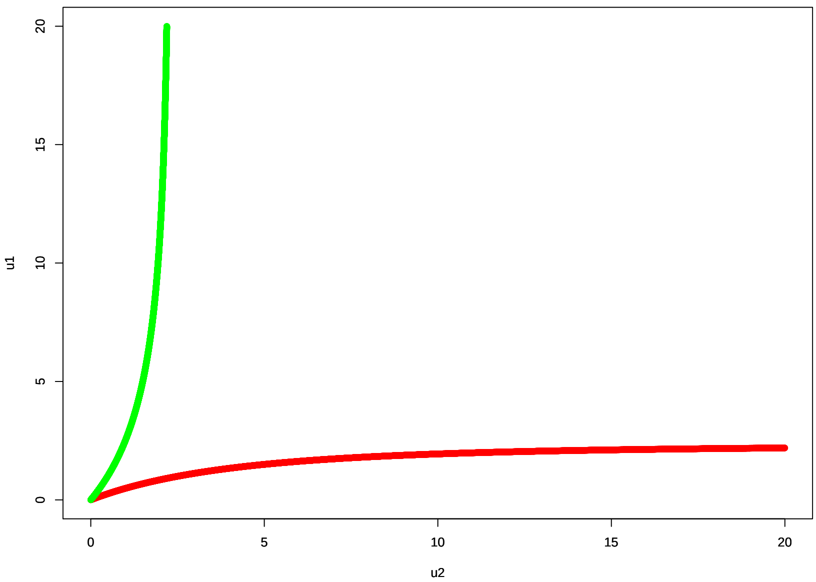

Figure 3.3. Symmetric Game with One Nash Equilibrium.

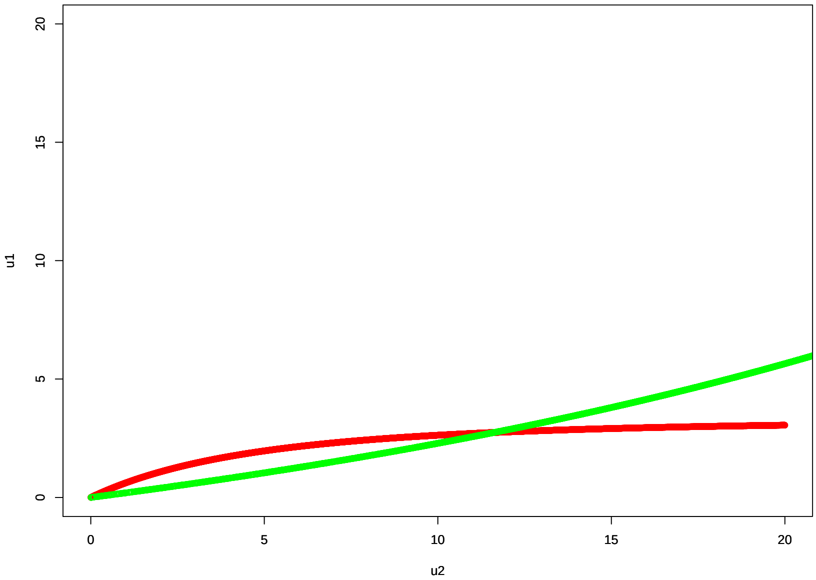

Figure 3.4. Unsymmetric Game with Two Nash Equilibriums.

4. ONE LOSS SYSTEM AND M/M/1 QUEUE WITH SIMULTANEOUS ARRIVALS: PERFORMANCE EVALUATION

4.1. Summary of Chapter

In this chapter, we focus on some notions and concepts which are not significantly included in the third chapter. In the previous chapter we mainly dealt with optimization processes based on service rates, but now we highly focus on some performance measurements of our new system such as steady state probabilities and expected profits. We analyse the system which consists of M/M/1/1 and M/M/1 queues where simultaneous arrivals exist with poisson process and servers are exponential distributed. If the capacitated queue is not idle in the queuing system upon an arrival to that system, then simultaneous arrivals are lost. We examine the system employing matrix geometric method that is aimed at deriving the steady state probabilities of the system. Additionally, we find the conditions that make the system be both irreducible and positive recurrent so as to ensure stability that is necessary to reach the steady state probabilities. Undoubtedly, the R matrix is so crucial to matrix geometric method and it needs to be found to obtain the stationary probabilities. Thus, we use the recursive technique to reach the R matrix under stability conditions. Nevertheless, we also apply a special case of R matrix that is capable of reducing large amount of computation during recursive operations and we compare the results of two distinct approaches during the computation of rate matrix.

4.2. Model and Analysis

Modelling the system with Markovian structure we obtain the following figure where the state space is (X_1, X_2) while X_1 can have all non-negative integers and X_2 can only have 0 or 1 setting that X_1 is the number of customers in the first M/M/1 queue, X_2 is the number of customers in the second M/M/1/1 queue. Thus, we have an infinite number of states. Additionally, all states are communicating with each other and due to the fact that we have also an infinite number of states stability can only be ensured under some conditions. Before jumping into details we need to observe the following Figure 4.1 to imagine the process.

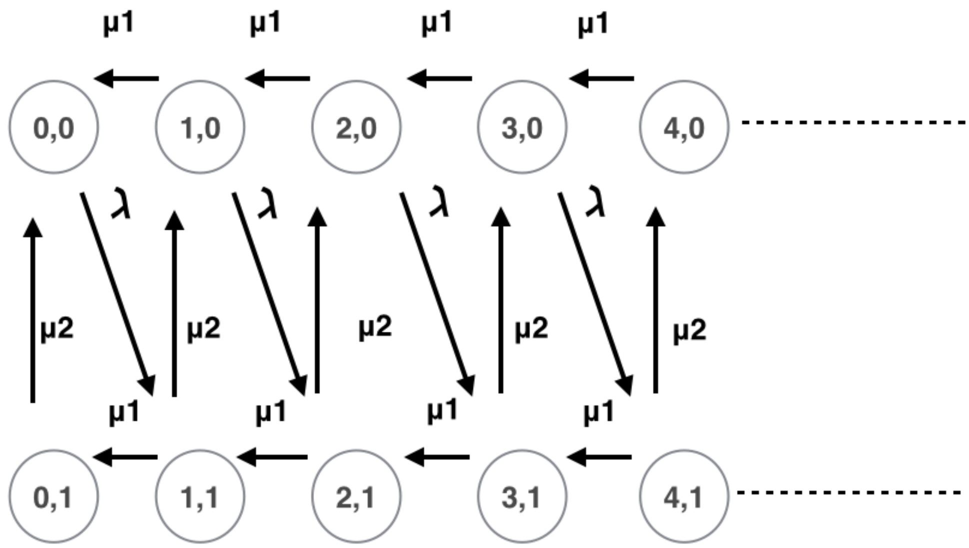

Figure 4.1. Model of One Loss System and M/M/1 Queue

In the Figure 4.1 we can observe that there is a quasi birth death process and each pair of states are communicating because of irreducibility. However, we have infinite number of states therefore, positive recurrent states must be provided satisfying stability condition which involves λ, μ_1, μ_2. Now let's take a glance at the generator matrix in the Figure 4.2.

$$
Q_2 = \begin{array}{cc}
 & \begin{array}{ccccc} (0,0)(0,1) & (1,0)(1,1) & (2,0)(2,1) & (3,0)(3,1) & \cdots \end{array} \\
\begin{array}{c} (0,0) \\ (0,1) \\ (1,0) \\ (1,1) \\ (2,0) \\ (2,1) \\ (3,0) \\ (3,1) \\ \vdots \end{array} &
\left(\begin{array}{c|c|c|c|c}
B_{00} & A_0 & & & \\
\hline
A_2 & A_1 & A_0 & & \\
\hline
 & A_2 & A_1 & A_0 & \\
\hline
 & & A_2 & A_1 & A_0 \\
\hline
 & & & \ddots & \ddots \quad \ddots
\end{array}\right)
\end{array}
$$

Figure 4.2. Generator Matrix

In the Figure 4.2, the generator matrix consists of only B_{00}, A_0, A_1, A_2 that are matrices with two rows and two columns. Now let's give the entities of those matrices as

$$B_{00} = \begin{pmatrix} -\lambda & 0 \\ \mu_2 & -\mu_2 \end{pmatrix}$$

$$A_2 = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_1 \end{pmatrix}$$

$$A_1 = \begin{pmatrix} -(\lambda + \mu_1) & 0 \\ \mu_2 & -(\mu_1 + \mu_2) \end{pmatrix}$$

$$A_0 = \begin{pmatrix} 0 & \lambda \\ 0 & 0 \end{pmatrix}.$$

We also consider ξ vector as an steady state probability vector in the following as

$$\xi^T = \begin{pmatrix} \xi_{00} & \xi_{01} & \xi_{10} & \xi_{11} & \xi_{20} & \xi_{21} & \cdots \end{pmatrix}$$

$$= \begin{pmatrix} \xi_0^T & \xi_1^T & \xi_2^T \cdots \end{pmatrix}$$

actually where

$$\xi_i^T = \begin{pmatrix} \xi_{i0} & \xi_{i1} \end{pmatrix}$$

indicating that ξ_i is the stationary probability vector of i^{th} level of the system and ξ_{i0} or ξ_{i1} are the stationary probabilities of i^{th} level with 0^{th} sub-level or i^{th} level with 1^{th} sub-level respectively. The level indicates the number of customers who belong to M/M/1 queue and the sub-level shows whether or not the M/M/1/1 system is idle. The solution form of quasi birth death process can be deemed as

$$\xi_i^T = \xi_{i-1}^T R \tag{4.1}$$

by G.Latouche and V.Ramaswami[58] where $i \geq 1$ and R is a matrix with two rows and two columns which indicates the expected time of visits to level i between two visits to level $i-1$ whereas the states of the irreducible system are positive recurrent. Let's give the balance equations in the following:

$$\xi^T Q_2 = 0$$

that implies

$$\begin{aligned} \xi_0^T B_{00} + \xi_1^T A_2 &= 0 \\ \xi_0^T A_0 + \xi_1^T A_1 + \xi_2^T A_2 &= 0 \\ \xi_1^T A_0 + \xi_2^T A_1 + \xi_3^T A_2 &= 0 \\ &\vdots \\ \xi_{i-1}^T A_0 + \xi_i^T A_1 + \xi_{i+1}^T A_2 &= 0. \end{aligned}$$

Merging the balance equations above and the solution form we mentioned before together what we obtain is

$$\begin{aligned} \xi_0^T B_{00} + \xi_1^T A_2 &= 0 \\ \xi_0^T A_0 + \xi_1^T (A_1 + RA_2) &= 0. \end{aligned}$$

Nevertheless, we need to include the normalisation equation, thus defining

$$S = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

what we have is

$$\begin{aligned} \xi_0^T B_{00} + \xi_1^T A_2 &= 0 \\ \xi_0^T A_0 + \xi_1^T (A_1 + RA_2) &= 0 \\ \sum_{n=0}^{\infty} \xi_n^T S &= 1. \end{aligned}$$

We can also convert the last equation above into

$$\xi_0^T S + (\xi_0^T R + \xi_0^T R^2 + \cdots) S = 1$$

which means

$$\xi_0^T S + \xi_1^T (\sum_{n=0}^{\infty} R^n) S = 1.$$

We need to express that the elements of R matrix has to be greater than or equal to zero by its nature [59], also recall that we have an irreducible and positive recurrent case. Therefore, we need to have $R^{n \to +\infty} = 0$ in order to ensure the Equation (4.1). Let's continue with the application of some algebric operations:

$$\begin{aligned}
&(\sum_{n=0}^{\infty} R^n) R = R + R^2 + R^3 + \cdots \\
&\implies \sum_{n=0}^{\infty} R^n - (\sum_{n=0}^{\infty} R^n) R = I \\
&\implies \sum_{n=0}^{\infty} R^n (I - R) = I \\
&\implies (I - R)^{-1} = \sum_{n=0}^{\infty} R^n.
\end{aligned}$$

Thus, using the equations above we can express that

$$\xi_0^T S + \xi_1^T (I - R)^{-1} = 1.$$

In short, using the following system of linear equations we can reach the steady state probabilities:

$$\begin{aligned}
&\begin{pmatrix} \xi_0^T & \xi_1^T \end{pmatrix} \begin{pmatrix} B_{00} & A_0 \\ A_2 & (A_1 + RA_2) \end{pmatrix} = \begin{pmatrix} 0 & 0 \end{pmatrix} \\
&\begin{pmatrix} \xi_0^T & \xi_1^T \end{pmatrix} \begin{pmatrix} S \\ (I - R)^{-1} S \end{pmatrix} = 1.
\end{aligned}$$

As we can observe that recalling the sub-levels we actually have four unknowns with five linear equations, so we can discard one of those equations and find ξ_0 and ξ_1 vectors. Afterwards, using the Equation (4.1) that is introduced by G.Latouche and V.Ramaswami [59] we accomplish to find the other steady state probabilities belong to other levels or sub-levels.

4.2.1. Recursive Technique

It is obvious that we need to derive what R matrix is. Therefore, initially we apply a recursive technique to obtain that matrix using $R^2A_2 + RA_1 + A_0$. We need to have a irreducible and positive recurrent system as well as have

$$\xi_j^T A_0 + \xi_j^T RA_1 + \xi_j^T R^2 A_2 = 0$$

with $j \geq 0$ which results from the balance equations with the deployment of the Equation (4.1) for all $i \geq 1$. So, $R^2A_2 + RA_1 + A_0$ has to be zero. The recursive technique we deploy according to [60] is based on the following equations in the Figure 4.3.

$R(0) = 0$;
for $k = 1$ to K **do**
  $n = k$;
  $R(n) = -(A_0 + R^2(n-1)A_2)A_1^{-1}$;
  $n = n + 1$;
end for

Figure 4.3. Recursive Algorithm for R matrix.

Proposition 4.1. *Recursive algorithm in the Figure 4.3 always converges for $\lambda > 0$, $\mu_1 > 0$ and $\mu_2 > 0$.*

Proof. Let's write the recursive operations in the following:

$$
\begin{aligned}
&R(0) = 0\\
&R(1) = -V\\
&R(2) = -V - V^2W\\
&R(3) = -V - (V^2 + 2V^3W + V^4W^2)W\\
&R(4) = -V - (V + V^2W + 2V^3W^2 + V^4W^3)^2W\\
&\vdots
\end{aligned}
$$

where we have

$$
-A_1^{-1} = \begin{pmatrix} \frac{1}{(\lambda+\mu_1)} & 0 \\ \frac{\mu_2}{(\lambda+\mu_1)(\mu_1+\mu_2)} & \frac{1}{(\mu_1+\mu_2)} \end{pmatrix}
$$

$$
-V = -A_0A_1^{-1} = \begin{pmatrix} \frac{\lambda\mu_2}{(\lambda+\mu_1)(\mu_1+\mu_2)} & \frac{\lambda}{(\mu_1+\mu_2)} \\ 0 & 0 \end{pmatrix}
$$

$$
-W = -A_2A_1^{-1} = \begin{pmatrix} \frac{\mu_1}{(\lambda+\mu_1)} & 0 \\ \frac{\mu_1\mu_2}{(\lambda+\mu_1)(\mu_1+\mu_2)} & \frac{\mu_1}{(\mu_1+\mu_2)} \end{pmatrix}.
$$

It can be seen that during the calculation of $R(n)$ the power of the V and W gradually increase, but because of the fact that the eigenvalues all of which are at the diagonal of lower or upper triangular matrices are strictly less than 1 whereas $\lambda > 0$, $\mu_1 > 0$ and $\mu_2 > 0$ we can express that the infinite power of V and W converge to zero matrix. Thus $R(n \to +\infty)$ has to be a finite matrix and has to converge. $\square$

Proposition 4.2. *All $R(n)$ matrices has non-negative elements during the recursive algorithm in the Figure 4.3 while $\lambda > 0$, $\mu_1 > 0$ and $\mu_2 > 0$.*

Proof. We apply induction proof for our proposition. So, initially assume that $R(n+$

1) $\geq R(n)$. Thus we have

$$\begin{aligned} R(n+2) &= -(A_0 + R^2(n+1)A_2)A_1^{-1} \\ &\geq -(A_0 + R^2(n)A_2)A_1^{-1} \\ &= R(n+1). \end{aligned}$$

But we already know that

$$R(0) = 0$$

and

$$R(1) = \begin{pmatrix} \frac{\lambda\mu_2}{(\lambda+\mu_1)(\mu_1+\mu_2)} & \frac{\lambda}{(\mu_1+\mu_2)} \\ 0 & 0 \end{pmatrix}$$

implying that $R(1) \geq R(0)$, thus we can express that our induction is consistent. $\square$

4.2.2. An Approach for Our Type of A_0 Matrices

During the usage of the recursive technique to reach the R matrix under stability conditions we may require large amount of computation. In order to alleviate that problem we can also apply a special case of R matrix as well as recursive operations referring to Adan *et al.* [61]. Now, let's show how to construct a special case of A_0 in our system.

Assume that

$$A_0 = w\beta^T,$$

where w and β are non-negative two dimensional vectors whereas

$$\beta^T e = 1$$

and

$$e = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

We know that

$$A_0 = \begin{pmatrix} 0 & \lambda \\ 0 & 0 \end{pmatrix}.$$

So we can set

$$w = \begin{pmatrix} \lambda \\ 0 \end{pmatrix}$$

and

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

Recalling our recursive iterations we can write

$$R(n) = -A_0 A_1^{-1} - R(n-1)^2 A_2 A_1^{-1},$$

where

$$R(0) = 0$$

and all

$$R(n) \geq 0.$$

Thus,

$$R(1) = -A_0 A_1^{-1} = w a_1^T$$

where a_1 is a nonnegative vector. Because A_1 is a diagonally dominant matrix having negative diagonals $-A_1^{-1}$ is a nonnegative matrix as in the following:

$$-A_1^{-1} = \begin{pmatrix} \frac{1}{(\lambda+\mu_1)} & 0 \\ \frac{\mu_2}{(\lambda+\mu_1)(\mu_1+\mu_2)} & \frac{1}{(\mu_1+\mu_2)} \end{pmatrix}.$$

In addition, $-A_2 A_1^{-1}$ is also a nonnegative matrix because of the structure of $-A_1^{-1}$ and we have:

$$-A_2 A_1^{-1} = \begin{pmatrix} \frac{\mu_1}{(\lambda+\mu_1)} & 0 \\ \frac{\mu_1\mu_2}{(\lambda+\mu_1)(\mu_1+\mu_2)} & \frac{\mu_1}{(\mu_1+\mu_2)} \end{pmatrix}.$$

Therefore we can write

$$\begin{aligned} R(2) &= w a_1^T - R^2(1) A_2 A_1^{-1} \\ &= w a_1^T + w a_1^T w a_1^T - A_2 A_1^{-1} \\ &= w a_2^T, \end{aligned}$$

where a_2 is a nonnegative vector. It can be seen that all $R(n)$ can be written in the form like $R(2)$ and $R(1)$. Then we can express that the R matrix we are looking for can be written as $R = wa^T$ where a is a nonnegative vector. If $R = wa^T$ is employed, then we can write the following:

$$\begin{aligned}
R^i &= wa^Twa^Twa^Twa^T \cdots wa^T \\
&= w(a^Tw)(a^Tw)(a^Tw)(a^T \cdots w)a^T \\
&= w(a^Tw)^{i-1}a^T \\
&= (a^Tw)^{i-1}wa^T \\
&= (a^Tw)^{i-1}R.
\end{aligned}$$

Now, multiplying the both sides of the last equation above with the eigenvector of R what we have is:

$$\begin{aligned}
R^iv &= (a^Tw)^{i-1}Rv \\
&\implies q^i = (a^Tw)^{i-1}q \\
&\implies q^{i-1} = (a^Tw)^{i-1}.
\end{aligned}$$

So, we can write $a^Tw = q$ where q is one of the eigenvalue of R matrix and v is the eigenvector.

We need to state that the rank of R matrix is one because all the rows are linearly dependent due to the wa^T structure. If a matrix has one rank, then it can not have any complex eigenvalues because of the fact that complex eigenvalues exist in pair and this make matrix rank at least two therefore there is either only one non zero real eigenvalue or just zero eigenvalue in this case. However, if all eigenvalues are zero this situation makes R matrix illogical because any power of R matrix equals zero and some steady state probabilities become exactly zero according to the Equation (4.1) but this creates contradiction in the irreducible and positive recurrent system. Therefore we need to focus on non-zero eigenvalue. Now recalling the equation

$$R^2A_2 + RA_1 + A_0 = 0$$

and merging that with the equation

$$R^i = q^{i-1}R$$

we obtain

$$R = -A_0(qA_2 + A_1)^{-1}.$$

It is already known that

$$q = a^T w \geq 0$$

but we can observe that if $q < 1$, then $(qA_2 + A_1)$ becomes a diagonally dominant matrix which is an invertible matrix as all diagonally dominant matrices. Because q is an eigenvalue of R we have

$$\begin{aligned}
det(R - qI) &= 0 \\
&= det(-A_0(qA_2 + A_1)^{-1} - qI) \\
&= det(-A_0(qA_2 + A_1)^{-1} - q(qA_2 + A_1)^{-1}(qA_2 + A_1)) \\
&= det(A_0 + q^2A_2 + qA_1)det(-(qA_2 - A_1)^{-1}) \\
&\implies det(A_0 + q^2A_2 + qA_1) = 0.
\end{aligned}$$

Therefore, we need to examine whether or not the determinant of $A_0 + q^2A_2 + qA_1$ is zero whereas $0 \leq q < 1$.

During the operations in both the recursive method and the special case we assume that the system is irreducible and positive recurrent, we already know the system is irreducible but we need to find the condition of stability so as to have positive recurrent states. By G.Latouche and V.Ramaswami [59] we can deploy

$$\alpha^T A_0 S < \alpha^T A_2 S$$

where α is the steady state probability vector derived from

$$\alpha^T A = 0$$

and

$$\alpha^T S = 1$$

while

$$A = A_2 + A_1 + A_0 = \begin{pmatrix} -\lambda & \lambda \\ \mu_2 & -\mu_2 \end{pmatrix}$$

and

$$S = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Therefore we can set

$$\lambda\alpha^T \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \mu_2\alpha^T \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$
$$\alpha^T S = 1.$$

Thus, we have

$$\alpha = \begin{pmatrix} \frac{\lambda}{(\lambda+\mu_2)} \\ \frac{\mu_2}{(\lambda+\mu_2)} \end{pmatrix}.$$

If $\alpha^T A_0 S < \alpha^T A_2 S$, then we have

$$\begin{pmatrix} 0 & \frac{\lambda\mu_2}{(\lambda+\mu_2)} \end{pmatrix} S < \begin{pmatrix} \frac{\mu_1\mu_2}{(\lambda+\mu_2)} & \frac{\lambda\mu_1}{(\lambda+\mu_2)} \end{pmatrix} S,$$

which implies

$$\lambda\mu_2 < \mu_1\mu_2 + \lambda\mu_1.$$

Now let's find the expected profits according to the stationary probabilities we have found for our quasi birth death system. Beginning with M/M/1/1 queue define that $P(X_2 = 1)$ is the stationary probability of having one customer in M/M/1/1 queue:

$$\begin{aligned}
P(X_2 = 1) &= \xi_{01} + \xi_{11} + \xi_{21} + \cdots \\
&= \xi_0^T \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \xi_1^T \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \xi_2^T \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \cdots \\
&= \xi_0^T \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \xi_0^T R \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \xi_0^T R^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \cdots \\
&= \xi_0^T (I + R + R^2 + \cdots) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&= \xi_0^T (\sum_{n=0}^{\infty} R^n) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
&= \xi_0^T (I - R)^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix}.
\end{aligned}$$

Therefore we can write

$$\pi_2 = \mu_2 P(X_2 = 1) r_2 - \mu_2 c_2$$
$$= \mu_2 \xi_0^T (I - R)^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} r_2 - \mu_2 c_2.$$

Now, it is time to find the expected profits of M/M/1 queue defining $P(X_1 = n)$ is the stationary probability of having n customer in M/M/1 queue, and we have

$$\pi_1 = \mu_1 (1 - P(X_1 = 0)) r_1 - \mu_1 c_1$$
$$= \mu_1 (1 - \xi_0^T S) r_1 - \mu_1 c_1.$$

4.3. Numerical Results

In that part, we compare both the recursive technique and the special case of A_0 during the calculation of R matrix. We present some results showing that R matrix does not change according to those techniques whereas involvement of the stability condition exists. We also display some results pertain to how R matrix changes with or without the stability condition based on $\lambda, \mu 1, \mu_2$. Therefore, we created the Table 4.1 of results belong to the recursion technique in the Figure 4.3 for a given $\lambda, \mu 1, \mu_2$. In the Table 4.1, it can be seen that if our input values satisfy the stability condition that is $\lambda \mu_2 < \mu_1 \mu_2 + \lambda \mu_1$, then the diagonal elements of R matrix has nonnegative values that are strictly less than one. Because of the fact that R matrix has to be an upper triangular matrix for our system all eigenvalues of R matrix lie along the diagonal. Thus, the infinite power of R matrices resulted from our recursive method involving stability condition have to converge to zero matrix and this circumstance is totally consistent the Equation (4.1) for $i \geq 1$ whereas irreducibility and positive recurrent states exist. But the infinite power of other R matrices whose diagonals are not strictly less than one can not converge to zero.

Table 4.1. Results of Recursive Method.

	Inputs as $\lambda, \mu 1, \mu 2$	**Outputs as R Matrix**
Experiment 1	1, 5, 2	$\begin{pmatrix} 0.0517 & 0.1483 \\ 0 & 0 \end{pmatrix}$
Experiment 2	4, 5, 2	$\begin{pmatrix} 0.1566 & 0.6434 \\ 0 & 0 \end{pmatrix}$
Experiment 3	8, 6, 3	$\begin{pmatrix} 0.2590 & 1.0744 \\ 0 & 0 \end{pmatrix}$
Experiment 4	8, 1, 9	$\begin{pmatrix} 1 & 0.8889 \\ 0 & 0 \end{pmatrix}$
Experiment 5	12, 2, 5	$\begin{pmatrix} 1 & 2.4 \\ 0 & 0 \end{pmatrix}$
Experiment 6	21, 3, 17	$\begin{pmatrix} 1 & 1.2353 \\ 0 & 0 \end{pmatrix}$

Recall that for the special case of A_0 we need to find whether or not the determinant of $A_0 + q^2 A_2 + qA_1$ is zero whereas $0 \leq q < 1$. Here we can turn our determinant structure into a polynomial structure regarding $\lambda, \mu 1, \mu 2$. As a polynomial what we have is

$$\mu_1^2 q^4 + (-2\mu_1^2 - \mu_1\mu_2 - \lambda\mu_1)q^3 + (\lambda\mu_1 + \lambda\mu_2 + \mu_1^2 + \mu_1\mu_2)q^2 + (-\lambda\mu_2)q = 0.$$

We also created the Table 4.2 of results belong to the special case for A_0 in the Figure 4.3 for a given $\lambda, \mu 1, \mu_2$. As we can see in the Table 4.2 possession of parameters that ensures stability yields a root that is strictly less than one and greater than 0. Therefore, we can embed that root into the equation that is

$$R = -A_0(qA_2 + A_1)^{-1}$$

which we derive for the special case of A_0 matrix in our system and find the same R matrices as in the recursive method.

Table 4.2. Results of Special Case for A_0.

	Inputs as $\lambda, \mu1, \mu2$	**Outputs as Polynomial Roots**
Experiment 1	1, 5, 2	$\begin{pmatrix} 0.0517 & 0 & 1 & 1.5483 \end{pmatrix}$
Experiment 2	4, 5, 2	$\begin{pmatrix} 0.1566 & 0 & 1 & 2.0434 \end{pmatrix}$
Experiment 3	8, 6, 3	$\begin{pmatrix} 0.2590 & 0 & 1 & 2.5744 \end{pmatrix}$
Experiment 4	8, 1, 9	$\begin{pmatrix} 1 & 0 & 6 & 12 \end{pmatrix}$
Experiment 5	12, 2, 5	$\begin{pmatrix} 1 & 0 & 2 & 7.5 \end{pmatrix}$
Experiment 6	21, 3, 17	$\begin{pmatrix} 1 & 0 & 4.1823 & 9.4843 \end{pmatrix}$

5. CONCLUSION

In the scope of this thesis, we first focus on two loss systems consisting of two M/M/1/1 queues where simultaneous arrivals exist with a Poisson process and servers are exponentially distributed. Modelling and solving our case we show the potential difference between two stationary probability vectors that belong to the case with simultaneous and independent arrivals. We also examine the Nash equilibrium of our system where the maximisation of profits are aimed for each queue in existence of fixed arrival rate as well as the slope and concavity of response functions based on server rates under some assumptions. Additionally, giving an example it is analytically shown that the centralised model in which the independent nonnegative server rates exist may have an optimal point strictly greater than the decentralised model. Finally, we manage to validate our analytical results with the numerical results. We can demonstrate the behaviour of slopes, concavity and the asymptotic structure of the response functions as well as the fact that two Nash equilibrium may not be satisfied under some conditions.

Secondly, we examine the system which consists of M/M/1/1 and M/M/1 queues where simultaneous arrivals exist with a Poisson process and servers are exponentially distributed. Applying the matrix geometric method we find the steady state probability vector as well as the expected profits in our system. Due to the fact that the R matrix plays a crucial role during the calculations we present two distinct techniques so as to derive the R matrix. We accomplish to prove that those techniques enable us to acquire the R matrix as it should be under the stability condition. In addition, by employing some experiments where we compare the outputs of those techniques for a given input we achieve to validate that those techniques yield the same R matrix and stationary probabilities under the stability condition. Furthermore, we obtain the expected profit generated by each queue benefiting from those probabilities.

As a future work, one can focus on a more detailed and generalized version of the third chapter including loss systems more than two. Thus, more complex cooperative and non-cooperative games can be examined where there is a competition between groups of players and no allowance for agreement respectively. Also, it can be assumed that simultaneous arrivals do not always exist involving that consecutive arrivals are correlated or there is a probability that arrivals are coming to the system simultaneously. Additionally, one can prefer to enlarge our system that exists in the fourth chapter creating sublevels more than one which requires a high dimensional quasi birth death process. The assumption of simultaneous arrivals also can be changed and more advanced methods may be employed to reach steady state probabilities, expected profits and other performance measurements such as expected waiting times.

REFERENCES

1. Kendall, D. G., "Stochastic Processes Occurring in The Theory of Queues and Their Analysis by The Method of The Imbedded Markov Chain", *The Annals of Mathematical Statistics*, pp. 338–354, 1953.

2. Song, J.-S., S. H. Xu and B. Liu, "Order-Fulfillment Performance Measures in an Assemble-to-Order System with Stochastic Leadtimes", *Operations Research*, Vol. 47, No. 1, pp. 131–149, 1999.

3. Swaminathan, J. M. and S. R. Tayur, "Managing Broader Product Lines Through Delayed Differentiation Using Vanilla Boxes", *Management Science*, Vol. 44, No. 12-part-2, pp. S161–S172, 1998.

4. Cheung, K. L. and W. H. Hausman, "Multiple Failures in a Multi-Item Spares Inventory Model", *Institute of Industrial Engineers Transactions*, Vol. 27, No. 2, pp. 171–180, 1995.

5. Glasserman, P. and Y. Wang, "Leadtime-Inventory Trade-Offs in Assemble-to-Order Systems", *Operations Research*, Vol. 46, No. 6, pp. 858–871, 1998.

6. Zhang, R. Q., "Expected Time Delay in Multi-Item Inventory Systems with Correlated Demands", *Naval Research Logistics (NRL)*, Vol. 46, No. 6, pp. 671–688, 1999.

7. Xu, S. H. and H. Li, "Majorization of Weighted Trees: A New Tool to Study Correlated Stochastic Systems", *Mathematics of Operations Research*, Vol. 25, No. 2, pp. 298–323, 2000.

8. Dayanik, S., J.-S. Song and S. H. Xu, "The Effectiveness of Several Performance Bounds for Capacitated Production, Partial-Order-Service, Assemble-to-Order Systems", *Manufacturing & Service Operations Management*, Vol. 5, No. 3,

pp. 230–251, 2003.

9. Iravani, S. M., K. L. Luangkesorn and D. Simchi-Levi, "A General Decomposition Algorithm for Parallel Queues with Correlated Arrivals", *Queueing Systems*, Vol. 47, No. 4, pp. 313–344, 2004.

10. Li, H. and S. H. Xu, "On The Dependence Structure and Bounds of Correlated Parallel Queues and Their Applications to Synchronized Stochastic Systems", *Journal of Applied Probability*, Vol. 37, No. 4, pp. 1020–1043, 2000.

11. Elhafsi, M., L. Zhi, H. Camus and E. Craye, "An Assemble-to-Order System with Product and Components Demand with Lost Sales", *International Journal of Production Research*, Vol. 53, No. 3, pp. 718–735, 2015.

12. Kushner, H. J., "Control and Optimal Control of Assemble to Order Manufacturing Systems Under Heavy Traffic", *Stochastics: An International Journal of Probability and Stochastic Processes*, Vol. 66, No. 3-4, pp. 233–272, 1999.

13. Lu, Y., J.-S. Song and D. D. Yao, "Backorder Minimization in Multiproduct Assemble-to-Order Systems", *Institute of Industrial Engineers Transactions*, Vol. 37, No. 8, pp. 763–774, 2005.

14. Van Houdt, B., "A Matrix Geometric Representation for The Queue Length Distribution of Multitype Semi-Markovian Queues", *Performance Evaluation*, Vol. 69, No. 7-8, pp. 299–314, 2012.

15. van Jaarsveld, W. and A. Scheller-Wolf, "Optimization of Industrial-Scale Assemble-to-Order Systems", *The Institute for Operations Research and the Management Sciences Journal on Computing*, Vol. 27, No. 3, pp. 544–560, 2015.

16. Gao, C., H. Shen and T. Cheng, "Order-Fulfillment Performance Analysis of an Assemble-to-Order System with Unreliable Machines", *International Journal of Production Economics*, Vol. 126, No. 2, pp. 341–349, 2010.

17. DeCroix, G. A., J.-S. Song and P. H. Zipkin, "Managing an Assemble-to-Order System with Returns", *Manufacturing & Service Operations Management*, Vol. 11, No. 1, pp. 144–159, 2009.

18. Yang, D.-Y., Y.-H. Chen and C.-H. Wu, "Modelling and Optimisation of a Two-Server Queue with Multiple Vacations and Working Breakdowns", *International Journal of Production Research*, Vol. 58, No. 10, pp. 3036–3048, 2020.

19. Yang, D.-Y. and Y.-H. Chen, "Computation and Optimization of a Working Breakdown Queue with Second Optional Service", *Journal of Industrial and Production Engineering*, Vol. 35, No. 3, pp. 181–188, 2018.

20. Laxmi, P. V. and S. Indira, "Ant Colony Optimisation for Impatient Customer Queue Under N-Policy and Bernoulli Schedule Vacation Interruption", *International Journal of Mathematics in Operational Research*, Vol. 10, No. 2, pp. 167–189, 2017.

21. Hamasha, M. M., G. Rumbe and M. T. Khasawneh, "Optimisation of Single-Queue Service Delivery Systems Using a Markovian Approach", *International Journal of Industrial and Systems Engineering*, Vol. 13, No. 4, pp. 424–441, 2013.

22. Ke, J.-C., F.-M. Chang and T.-H. Liu, "M/M/c Balking Retrial Queue with Vacation", *Quality Technology & Quantitative Management*, Vol. 16, No. 1, pp. 54–66, 2019.

23. Elijah, K., G. Joseph, M. Samuel and D. Mauti, "Optimisation of Shovel-Truck Haulage System in an Open Pit Using Queuing Approach", *Arabian Journal of Geosciences*, Vol. 14, No. 11, pp. 1–10, 2021.

24. Liou, C.-D., "Markovian Queue Optimisation Analysis with an Unreliable Server Subject to Working Breakdowns and Impatient Customers", *International Journal of Systems Science*, Vol. 46, No. 12, pp. 2165–2182, 2015.

25. Chiang, M., A. Sutivong and S. Boyd, “Efficient Nonlinear Optimizations of Queuing Systems”, *Global Telecommunications Conference, 2002, GLOBECOM’02.*, Vol. 3, pp. 2425–2429, 2002.

26. Rapoport, A., W. E. Stein, J. E. Parco and D. A. Seale, “Equilibrium Play in Single-Server Queues with Endogenously Determined Arrival Times”, *Journal of Economic Behavior & Organization*, Vol. 55, No. 1, pp. 67–91, 2004.

27. George, J. M. and J. M. Harrison, “Dynamic Control of a Queue with Adjustable Service Rate”, *Operations Research*, Vol. 49, No. 5, pp. 720–731, 2001.

28. Chagas Moura, M., J. M. Santana, E. L. Droguett, I. D. Lins and B. N. Guedes, “Analysis of Extended Warranties for Medical Equipment: A Stackelberg Game Model Using Priority Queues”, *Reliability Engineering & System Safety*, Vol. 168, pp. 338–354, 2017.

29. Xia, L., Q.-M. He and A. S. Alfa, “Optimal Control of State-Dependent Service Rates in a MAP/M/1 Queue”, *Transactions on Automatic Control*, Vol. 62, No. 10, pp. 4965–4979, 2017.

30. Acharya, S. K. and C. E. Villarreal-Rodríguez, “Change Point Estimation of Service Rate in an M/M/1/m Queue”, *International journal of mathematics in operational research*, Vol. 5, No. 1, pp. 110–120, 2013.

31. Dave, U. and Y. Shah, “Maximum Likelihood Estimates in a M/M/2 Queue with Heterogeneous Servers”, *Journal of the Operational Research Society*, Vol. 31, No. 5, pp. 423–426, 1980.

32. Wang, T.-Y., J.-C. Ke, K.-H. Wang and S.-C. Ho, “Maximum Likelihood Estimates and Confidence Intervals of an M/M/R Queue with Heterogeneous Servers”, *Mathematical Methods of Operations Research*, Vol. 63, No. 2, p. 371, 2006.

33. Beneš, V. E., “A Sufficient Set of Statistics for a Simple Telephone Exchange

Model", *Bell System Technical Journal*, Vol. 36, No. 4, pp. 939–964, 1957.

34. Armony, M., G. Roels and H. Song, "Pooling Queues with Strategic Servers: The Effects of Customer Ownership", *Operations Research*, Vol. 69, No. 1, pp. 13–29, 2021.

35. Vijayalakshmi, C. and K. Sivaselvan, "Stochastic Routing Mechanism on Multiclass Switching Network Using Game Theory", *Applied Mathematical Sciences*, Vol. 7, No. 21, pp. 1023–1032, 2013.

36. Timmer, J. and W. Scheinhardt, "How to Share The Cost of Cooperating Queues in a Tandem Network?", *2010 22nd International Teletraffic Congress (lTC 22)*, pp. 1–7, 2010.

37. Timmer, J. and W. Scheinhardt, "Cost Sharing of Cooperating Queues in a Jackson Network", *Queueing Systems*, Vol. 75, No. 1, pp. 1–17, 2013.

38. Wee, K.-E. and A. Iyer, "Consolidating or Non-Consolidating Queues: A Game Theoretic Queueing Model with Holding Costs", *Operations Research Letters*, Vol. 39, No. 1, pp. 4–10, 2011.

39. González, P. and C. Herrero, "Optimal Sharing of Surgical Costs in The Presence of Queues", *Mathematical Methods of Operations Research*, Vol. 59, No. 3, pp. 435–446, 2004.

40. Carlyle, W. M., S. G. Henderson and R. Szechtman, "Allocating Capacity in Parallel Queues to Improve Their Resilience to Deliberate Attack", *Naval Research Logistics (NRL)*, Vol. 58, No. 8, pp. 731–742, 2011.

41. Cachon, G. P. and F. Zhang, "Obtaining Fast Service in a Queueing System via Performance-Based Allocation of Demand", *Management Science*, Vol. 53, No. 3, pp. 408–420, 2007.

42. Gilbert, S. M. and Z. K. Weng, “Incentive Effects Favor Nonconsolidating Queues in a Service System: The Principal–Agent Perspective”, *Management Science*, Vol. 44, No. 12-part-1, pp. 1662–1669, 1998.

43. Lodree, E. J., N. Altay and R. A. Cook, “Staff Assignment Policies for a Mass Casualty Event Queuing Network”, *Annals of Operations Research*, Vol. 283, No. 1, pp. 411–442, 2019.

44. Weber, R. R. and S. Stidham, “Optimal Control of Service Rates in Networks of Queues”, *Advances in Applied Probability*, Vol. 19, No. 1, pp. 202–218, 1987.

45. Rosberg, Z., P. Varaiya and J. Walrand, “Optimal Control of Service in Tandem Queues”, *Transactions on Automatic Control*, Vol. 27, No. 3, pp. 600–610, 1982.

46. Stidham, S. and R. Weber, “A Survey of Markov Decision Models for Control of Networks of Queues”, *Queueing systems*, Vol. 13, No. 1, pp. 291–314, 1993.

47. Ahn, H.-S., I. Duenyas and M. E. Lewis, “Optimal Control of a Two-Stage Tandem Queuing System with Flexible Servers”, *Probability in the Engineering and Informational Sciences*, Vol. 16, No. 4, pp. 453–469, 2002.

48. Azaron, A. and S. F. Ghomi, “Optimal Control of Service Rates and Arivals in Jackson Networks”, *European Journal of Operational Research*, Vol. 147, No. 1, pp. 17–31, 2003.

49. Kobayashi, H. and B. L. Mark, “On Queuing Networks and Loss Networks”, *Proc. 1994 Annual Conference on Information Sciences and Systems*, Princeton NJ, 1994.

50. Vinarskiy, M., “A Method of Approximate Analysis of an Open Exponential Queuing Network with Losses due to Finite Shared Buffers in Multi-Queue Nodes”, *European Journal of Operational Research*, Vol. 258, No. 1, pp. 207–215, 2017.

51. Bronshtein, O. and I. B. Gertsbakh, "An Open Exponential Queuing Network with Limited Waiting Spaces and Losses: A Method of Approximate Analysis", *Performance Evaluation*, Vol. 4, No. 1, pp. 31–43, 1984.

52. Fernández, E., Y. Hinojosa and J. Puerto, "Filtering Policies in Loss Queuing Network Location Problems", *Annals of Operations Research*, Vol. 136, No. 1, pp. 259–283, 2005.

53. Guo, M., B. Li, S. Wu and J. Song, "Effectiveness of Referral Incentive Policy: Exploring Using Queuing Network Model with Blocking", *International Conference on Services Systems and Services Management*, pp. 1–6, 2011.

54. Naumov, V. A. and K. Samuilov, "Analysis of Networks of The Resource Queuing Systems", *Automation and Remote Control*, Vol. 79, No. 5, pp. 822–829, 2018.

55. Alnowibet, K. A. and H. Perros, "Nonstationary Analysis of The Loss Queue and of Queueing Networks of Loss Queues", *European Journal of Operational Research*, Vol. 196, No. 3, pp. 1015–1030, 2009.

56. Ku, C.-Y. and S. Jordan, "Access Control of Parallel Multiserver Loss Queues", *Performance Evaluation*, Vol. 50, No. 4, pp. 219–231, 2002.

57. Lam, S. S., "Queuing Networks with Population Size Constraints", *IBM Journal of Research and Development*, Vol. 21, No. 4, pp. 370–378, 1977.

58. Debreu, G., "A Social Equilibrium Existence Theorem", *Proceedings of the National Academy of Sciences*, Vol. 38, No. 10, pp. 886–893, 1952.

59. Latouche, G. and V. Ramaswami, *Introduction to Matrix Analytic Methods in Stochastic Modeling*, Society for Industrial and Applied Mathematics, 1999.

60. Neuts, M. F., *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*, Courier Corporation, 1994.

61. Adan, I., J. van Leeuwaarden and J. Selen, "Analysis of Structured Markov Processes", *ArXiv Preprint ArXiv:1709.09060*, 2017.