UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

MIXED CONVEXITY AND OPTIMIZATION RESULTS FOR FUNCTIONS WITH INTEGER AND REAL VARIABLES WITH APPLICATIONS TO QUEUEING SYSTEMS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

EMRE TOKGÖZ
Norman, Oklahoma
2012

MIXED CONVEXITY AND OPTIMIZATION RESULTS FOR FUNCTIONS WITH INTEGER AND REAL VARIABLES WITH APPLICATIONS TO QUEUEING SYSTEMS

A DISSERTATION APPROVED FOR THE SCHOOL OF INDUSTRIAL ENGINEERING

BY

Dr. Hillel Kumin, Chair

Dr. Theodore B. Trafalis

Dr. Süleyman Karabük

Dr. K. Thulasiraman

Dr. Dean Hougen

ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Dr. Hillel Kumin, for encouraging me to start a doctoral degree in School of Industrial Engineering, for serving as a member in my mathematics doctoral degree, and for his tremendous support.

Secondly, I would like to thank Dr. Dean Hougen for serving in both of my doctoral committees. I would like to also thank Dr. Theodore B. Trafalis, Dr. K. Thulasiraman and Dr. Süleyman Karabük who supported me to achieve my goal, a doctoral degree in Industrial Engineering, by serving on my Industrial Engineering doctoral committee.

Third, I would like to thank my wife, Annina Tokgöz, who encouraged me to start and finish all the degrees I have received. Special thanks to my family members in Turkey, Beyhan Tokgöz, Kâmil Tokgöz, Bengü Çapkın, Gürol Çapkın and Dide Çapkın, and my family members in the United States, Annina Tokgöz, Elena Collier and Bob Collier, who supported me during the completion of all of my degrees. Words are not enough to express my appreciation for their priceless help, encouragement, and support. In particular, I would also like to thank my friends who provided moral support throughout my education.

Finally, I would like to thank the University of Oklahoma School of Industrial Engineering, Department of Mathematics, School of Computer Science, and the Graduate College for their financial support.

TABLE OF CONTENTS

List of Tables

LIST OF FIGURES

ABSTRACT

In this dissertation, a new method for obtaining convexity and optimization results for functions with integer and real (i.e. mixed) variables is introduced and this new method is applied to obtain mixed convexity and optimization results for some of the known mixed variable functions in the literature. These mixed variable functions include the Erlang delay and loss formulae in telecommunication systems, an $(S-1, S)$ inventory model (suggested by Das (1977)), and an $M/E_k/1$ queueing system model (suggested by Kumin (1973)). Local and global mixed convexity and optimization results for these mixed variable functions are obtained after introducing definitions for a condense discrete convex set, a condense discrete convex function, a discrete Hessian matrix, a mixed convex set, a mixed convex function, and a mixed Hessian matrix. Symbolic toolbox of MATLAB $R2009a$ is used to obtain symbolic results. Computational discrete and mixed convexity and optimization results are also obtained by using MATLAB $R2009a$. The results obtained in this work are important because prior to this work no joint convexity results for mixed functions for mixed functions have been defined. This dissertation obtains such joint results. In addition, for real variable functions that are strictly convex, it is well-known that any local minimum is also the global minimum. In this work, similar results are obtained for mixed strictly convex functions. A new Hessian matrix defined for mixed variable functions can be used to determine whether any local minimum is also the global minimum.

Chapter 1

INTRODUCTION

In the design of diverse systems, models obtained for optimization can have both integer and real (mixed) variables. These models include the Erlang delay, Erlang loss, $(S-1, S)$ inventory model (suggested by Das (1977)), and an $M/E_k/1$ queueing system model (suggested by Kumin (1973)). A practical way to obtain optimization results for functions with real variables is by determining the convexity of the proposed model since every local minimum is also a global minimum. The convexity definition of a real multivariate function is unique in real convex analysis; however, there are many convexity definitions of a discrete multivariate function. This variation of definitions in discrete convex analysis can result in different convexity results for the same model. Combinatorial and algorithmic techniques are often applied to solve discrete variable function optimization problems. Then the following natural question arises: Is there a practical way to obtain mixed convexity results for mixed variable functions that can be used to obtain optimization results for the corresponding models?

Convexity of mixed multivariate functions, functions with domain $\mathbb{Z}^n \times \mathbb{R}^m$, are obtained either for the integer variables when the real variables are assumed constant or the real variables when the integer variables are assumed constant. For example, Harel (2010) proves the real convexity of the Erlang delay formula with respect to the real variable traffic intensity (ρ) when the integer variable number of servers (s)

is assumed constant for more than two servers in the system. Let

s : Number of servers, λ : Arrival rate,

μ : Service Rate, ρ : Traffic intensity,

a : The offered load in the system,

B : The probability that all servers busy in an $M/M/s$ queue,

W_q : The average waiting time in the queue,

L_q : The number of customers waiting in the queue,

L : The number of customers waiting in the system,

where $\rho = \frac{\lambda}{s\mu}$. Table 1.1 summarizes some of the queueing system models with mixed variables, authors, and the corresponding obtained convexity results:

Table 1.1	Queueing system models & Convexity results	
Queueing System	*Model & Results*	*Author*
$M/D/s$	$W_q(s) = \sum_{n=1}^{\infty} e^{-na} \left[\sum_{j=ns}^{\infty} \frac{(na)^j}{j!} - \frac{s}{a} \sum_{j=(n+1)s}^{\infty} \frac{(na)^j}{j!} \right]$, W_q *convex in s.*	*Rolfe* (1971)
$M/M/s$	$W_q(s) = \frac{\rho^s}{s!\left(1-\frac{\rho}{s}\right)^2 s\mu} \left[\sum_{j=0}^{s-1} \frac{\rho^j}{j!} + \frac{\rho^s}{s!\left(1-\frac{\rho}{s}\right)} \right]$, W_q *convex in s.*	*Dyer & Proll* (1977)
$M/M/s$	$W_q(\rho) = \frac{B}{(1-\rho)s\mu}$, ($B$: *Erlang delay formula*), W_q *convex in* ρ.	*Lee & Cohen* (1983)
$M/M/s$	$L(\rho) = s\rho + B\frac{\rho}{1-\rho}$, L *is convex in* ρ.	*Grassmann* (1985)
$M/M/s$	$L_q(\rho) = \frac{\rho}{1-\rho}B$, L_q *is convex in* ρ.	*Lee & Cohen* (1985)
$M/M/s$	*Erlang delay & Erlang Loss formulae,* *convex in* ρ	*Harel* (2010)

All the convexity results given in Table 1.1 have the following common properties:

1. All the models are associated with a queueing system where μ is the service rate, the number of servers s is the integer variable, and the traffic intensity ρ is the real variable.

2. The convexity results are obtained when either ρ or s is assumed constant.

The results obtained in this research are important because prior to this work no joint convexity results for mixed functions for mixed functions have been defined. This dissertation obtains such joint results. In addition, for real variable functions that are strictly convex, it is well-known that any local minimum is also the global minimum. In this work, similar results are obtained for mixed strictly convex functions. A new Hessian matrix defined for mixed variable functions can be used to determine whether any local minimum is also the global minimum.

The convexity definition of real multivariate functions and the corresponding optimization results will be an integral part of the results obtained for condense discrete and mixed convexity and optimization of the corresponding functions; therefore, the convexity and optimization results obtained for discrete, real, and mixed multivariate functions will be introduced in the next chapter.

Chapter 2

CONVEXITY AND OPTIMIZATION

In this chapter, the definitions and results known in the literature relevant to this work are summarized for real (continuous), discrete, and mixed (integer and real) variable functions.

2.1 CONTINUOUS CONVEX ANALYSIS AND OPTIMIZATION

Real convex analysis is shaped by the lecture notes of Fenchel while lecturing at Princeton University in the early 1950's where concepts such as convex sets, cones, and functions are covered (Rockafellar (1970)). Fenchel ((1953), (1971), (1974), and (1987)), Bonnesen ((1971), (1974), and (1987)), and Rockafellar ((1970), (1974), and (1998)) are some of the leading researchers in real convex analysis of 1900's. The following two definitions are the basic definitions that we will be using throughout this work.

A set $D \subseteq \mathbb{R}^n$ is called convex if it satisfies the condition

$$x, y \in D,\ 0 \leq a \leq 1 \Rightarrow ax + (1-a)\,y \in D.$$

Definition 2.1 (Convex function): A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is called a convex function on a convex set $D \subseteq \mathbb{R}^n$ if and only if for $\forall\ x, y \in D$ and $0 \leq a \leq 1$, the inequality

$$f(ax + (1-a)y) \leq af(x) + (1-a)f(y) \qquad (2.1)$$

holds. f is called strictly convex if the inequality in (2.1) is a strict inequality when $0 < a < 1$.

Convexity results obtained for nonlinear real multivariable functions have important applications in the optimization of real variable functions (Borwein and Lewis (2000)). In particular for the case when a function is convex/concave local optimality guarantees global optimality and the global optimum can be found by descent algorithms. The (Legendre-Fenchel) duality and separation theorems hold which are well known results in real convex analysis (Rockafellar, (1970)). In addition, it is well known in real convex analysis that a C^2 function has a unique minimum point if it is strictly convex in a convex domain. The convexity of a C^2 function is equivalent to the positive definiteness of the corresponding Hessian matrix, the matrix that consists of all the second order partial derivatives of the function. Convexity of functions with constraints have also important applications in engineering where algorithmic approaches can also be taken (Boyd and Vandenberghe (2009)).

2.2 DISCRETE CONVEX ANALYSIS AND OPTIMIZATION

A real valued function f with a single discrete variable is defined to be discrete convex by Fox (1966) if its first forward differences are increasing or at least non-decreasing; That is,

$$f(x+2) - 2f(x+1) - f(x) \geq 0.$$

Discretely convex functions are introduced by Miller (1971) , and Integrally-convex functions are introduced by Favati and Tardella (1990). M, L, $M^{\natural}$, and $L^{\natural}$ convexity are introduced by Murota (1996) , Murota (1998), Murota and Shioura (1999), and Fujishige and Murota (2000), respectively. The Hessian matrices corresponding to M, L, $M^{\natural}$, and $L^{\natural}$ convexity are introduced by Hirai and Murota (2004) , and Moriguchi and Murota (2005). Yüceer (2002) introduced strong discrete convexity of functions with positive semi-definite matrix of second forward differences. D-convex function introduced by Ui (2006) has a unified form that includes discretely convex, integrally convex, M, $M^\natural$, L, and $L^\natural$ convex functions in local settings. Tokgöz, Nourazari and Kumin (2011) introduced the condense discrete convexity of multivariable functions by extending the discrete convexity definition of Fox (1966) which will be explained in chapter 4.

2.3 MIXED CONVEX ANALYSIS AND OPTIMIZATION

Mixed integer programming problems in operations research contain functions with integer and real variables where the constraints of these problems also have integer and real variables. Particular solutions to these problems can be obtained by using algorithmic approaches. A theoretical approach to the mixed convexity of mixed functions is suggested by Tokgöz, Maalouf, and Kumin (2009) where the mixed convexity of a mixed variable function associated with an $M/E_k/1$ queueing system is obtained with respect to the constraints. In this work we obtain local and global mixed convexity results after introducing definitions, and obtaining results for condense discrete convex sets and functions. In addition, optimization results for mixed convex functions are obtained. Tokgöz (2009) introduced the mixed convexity of mixed variable functions T, T^*, E, and E^* with the corresponding discrete convex counterparts M, L, $M^\natural$, and $L^\natural$ convexity definitions. In addition, Hessian matrices corresponding to the mixed T, T^*, E, and E^* functions are introduced to obtain mixed convexity results.

2.4 ORGANIZATION OF THE DISSERTATION

In this work, discrete and mixed convexity definitions of multivariable functions are introduced and the corresponding theoretical convexity and optimization results are obtained using these definitions. These convexity and optimization results are used to obtain numerical results for several mixed variable functions.

In chapter 2 we cite the literature on convexity and optimization of real, discrete, and mixed variable functions. In chapter 3, condense discrete convex function, condense discrete convex set, discrete Hessian matrix, local minimum, and global minimum definitions of a multivariate discrete function are introduced, and condense discrete convexity and optimization results are obtained with an application to an adaptation of Rosenbrock's function proposed by Yüceer (2002). In chapter 4, condense mixed convex function, condense mixed convex set, mixed Hessian matrix, local minimum, and global minimum definitions of a multivariate mixed function are introduced, and mixed convexity and optimization results are obtained. The mixed convexity and optimization definitions and results obtained in chapter 4 are used to obtain mixed convexity and optimization results corresponding to the $M/E_k/1$ queueing systems in chapter 5, to the Erlang Delay and Loss formulae in chapter 6, and to an $(S-1, S)$ inventory model proposed by Das in 1977 in chapter 7 where numerical and graphical illustration of the results are presented. In chapter 9, conclusion and possible future work are provided. In appendices A, B, and C the algorithms/programs and theoretical mixed convexity results will be obtained by using the symbolic toolbox of MATLAB $R2009a$ program and algebraic calculations.

CHAPTER 3

CONDENSE DISCRETE CONVEXITY AND OPTIMIZATION OF FUNCTIONS

In this section, condense discrete convexity of multivariate nonlinear discrete functions $\beta : \mathbb{Z}^n \to \mathbb{R}$ is introduced, which is a generalization of the integer convexity definition of Fox (1966) for a one variable discrete function to nonlinear multivariable discrete variable functions. Similar to the difference operator definition of Kiselman and Christer (2010), we define the first difference of an integer variable function $f : \mathbb{Z}^n \to \mathbb{R}$ by

$$\nabla_i f(x) = f(x + e_i) - f(x),$$

and the difference of the first difference, namely the second difference of f is defined by

$$\nabla_{ij}(f(x)) = f(x + e_i + e_j) - f(x + e_i) - f(x + e_j) + f(x),$$

where e_i represents the positive integer vectors of unit length at the i^{th} position of the function f. A discrete Hessian matrix H consisting of second differences $\nabla_{ij}\beta$ $(1 \leq i, j \leq n)$ corresponding to a condense discrete convex function $\beta : \mathbb{Z}^n \to \mathbb{R}$ in local settings is introduced, and convexity results are obtained for condense discrete functions similar to the convexity results obtained in real convex analysis. The discrete Hessian matrix H is shown to be symmetric, linear, and vanishes when the condense discrete function is affine.

Yüceer (2002) proves convexity results for a certain class of discrete convex functions and shows that the restriction of the adaptation of Rosenbrook's function from real variables to discrete variables does not yield a discretely convex function using Miller's (1971) discrete convexity definition. Here, it is shown that the adaptation of Rosenbrook's function considered by Yüceer (2002) is a condense discrete convex function where the set of local minimums is also the the set of global minimums.

3.1 CONDENSE DISCRETE CONVEXITY OF FUNCTIONS

In discrete convex analysis, Fox (1966) defined a single variable discrete function to be convex if the first forward differences of the given function are increasing or at least non-decreasing. A multivariable discrete L-convex function is defined to be the generalization of the Lovász extension of submodular set functions in (1988) by Murota. $L^{\#}-$convex functions are defined in (2000) by Fujishige and Murota. The concept of M-convex functions is introduced by Murota in (1996) and that of $M^{\#}$ convex functions by Murota-Shioura in (1999). The discrete analogue of Hessian matrices corresponding to multivariable discrete L, $L^{\#}$, M, and $M^{\#}$ functions are introduced by Hirai and Murota (2004), and Moriguchi and Murota (2005). Important applications of $L, L^{\#}, M$, and $M^{\#}$ discrete convex/concave functions appear in network flow problems (see Murota (2003) for details). The convexity properties of nonlinear integer variable, integer valued objective functions are investigated by Favati and Tardella (1990) where algorithmic approaches are also presented. Kiselman and Samieinia (2008) define the convex envelope, canonical extension and lateral convexity of multivariable discrete functions where the second difference of a function $f : \mathbb{Z}^n \to \mathbb{R}$ is introduced to define lateral convexity.

Let S be a subspace of a discrete n-dimensional space. A function $f : S \to R$ is defined to be discrete convex by Yüceer (2002) (using Miller's (1971) definition) if for all $x, y \in S$ and $\alpha > 0$ we have

$$\alpha f(x) + (1-\alpha) f(y) \geq \min_{u \in N(z)} f(u) \tag{3.1}$$

where $z = \alpha x + (1-\alpha)y$, $N(z) = \{u \in S : \|u - z\| < 1\}$, and $\|u\| = \max\{|u_i| : 1 \leq i \leq n\}$. This discrete convex function definition yields nonnegative second forward differences in each component, and a symmetric matrix of second forward cross differences. By imposing additional submodularity conditions on discrete convex functions, the concept of strong discrete convexity is introduced in [17]. A strong discrete convex function has a corresponding positive semi-definite matrix of second forward differences which has practical and computational implications. D-convex and semistrictly quasi D-convex functions are introduced in (2006) by Ui where D-convex functions have a unified form that includes discretely convex, integrally convex, M convex, $M^\natural$ convex, L convex, and $L^\natural$ convex functions in local settings.

We define a condense discrete convex set D to be the set of points that coincides with a real convex set on the integer lattice which is large enough to support the second difference of a given condense discrete function. We assume that the union of condense discrete convex sets are discrete convex sets as well. The following definition of an n-integer variable function holds for a certain class of discrete functions. The definition of a condense discrete convex function is based on its quadratic approximation in a condense discrete convex neighborhood $D \subset \mathbb{Z}^n$. The quadratic approximation of f is the quadratic discrete function that agrees with f in a local neighborhood. It is called "approximation" because the function under consideration may not agree with the quadratic function over the entire space. A local neighborhood needs to have the second difference points in it.

Definition 3.1: A discrete function $f : D \to \mathbb{R}$ on a condense discrete convex set $D \subset \mathbb{Z}^n$ is defined to be condense discrete convex if its quadratic approximation

$$f(x) = \frac{1}{2}x^T A x + b^T x + c \tag{3.2}$$

in the neighborhood D is strictly positive where A is the symmetric coefficient matrix of the quadratic approximation of f. f is called condense discrete concave if $-f$ is condense discrete convex. A is called the discrete coefficient matrix of f.

Proposition 3.1: Let $f : D \to \mathbb{R}$ be defined on a condense discrete convex set $D \subset \mathbb{Z}^n$ with its quadratic approximation defined by (3.2) .The coefficient matrix A corresponding to f is the symmetric matrix $[\nabla_{ij}(f)]_{n \times n}$.

Proof: We first prove the symmetry of the matrix $[\nabla_{ij}(f)]_{n \times n}$.

$$\begin{aligned}
\nabla_{ij} f(x) &= \nabla_i \left(f(x+e_j) - f(x)\right) \\
&= f(x+e_i+e_j) - f(x+e_i) - f(x+e_j) + f(x) \\
&= \nabla_j \left(f(x+e_i) - f(x)\right) \\
&= \nabla_j \left(\nabla_i f(x)\right) = \nabla_{ji} f(x).
\end{aligned}$$

Assuming that A is symmetric, for all i and j,

$$\begin{aligned}
\nabla_{ij}(f(x)) &= \frac{1}{2}\nabla_i \left((x+e_j)^T A (x+e_j) - x^T A x\right) \\
&= \frac{1}{2}\nabla_i \left(x^T A (x+e_j) + e_j^T A (x+e_j) - x^T A x\right) \\
&= \frac{1}{2}\nabla_i \left(x^T A x + x^T A e_j + e_j^T A x + e_j^T A e_j - x^T A x\right) \\
&= \frac{1}{2}\nabla_i \left(x^T A e_j + e_j^T A x + e_j^T A e_j\right) \\
&= \frac{1}{2}\left((x+e_i)^T A e_j - x^T A e_j + e_j^T A (x+e_i) - e_j^T A x\right) \\
&= \frac{1}{2}\left(x^T A e_j + e_i^T A e_j - x^T A e_j + e_j^T A x + e_j^T A e_i - e_j^T A x\right) \\
&= \frac{1}{2}\left(e_i^T A e_j + e_j^T A e_i\right) \\
&= a_{ij}.
\end{aligned}$$

Therefore

$$A_f = [a_{ij}]_{n \times n} = [\nabla_{ij} f]_{n \times n}.$$

which completes the proof.

Proposition 3.2: The coefficient matrix A_f of $f : D \to \mathbb{R}$ given above in proposition 3.1 satisfies the properties of the discrete Hessian matrix corresponding to real convex functions. That is, A_f is linear with respect to the condense discrete functions, symmetric, and vanishes when f is discrete affine.

Proof: Let $f_1 : S_1 \to \mathbb{R}$ and $f_2 : S_2 \to \mathbb{R}$ be condense discrete functions with the corresponding coefficient matrices A_{f_1} and A_{f_2}, respectively. Then

$$\begin{aligned}
\left[\nabla_{ij}(f_1 + f_2)\right]_{n\times n} &= A_{f_1+f_2} \\
&= A_{f_1} + A_{f_2} \\
&= \left[\nabla_{ij}(f_1)\right]_{n\times n} + \left[\nabla_{ij}(f_2)\right]_{n\times n}
\end{aligned}$$

which also proves the linearity of the second difference operator with respect to the condense discrete functions. The symmetry condition is proven in proposition 3.1.

Considering the condense discrete affine function f,

$$f(x) = \sum_{i=1}^{n} b_i x_i$$

the second difference operator vanishes since $\nabla_i(f) = b_i$ and $\nabla_{ij}(f) = 0$ for all i and j.

Theorem 3.1: A function $f : D \to \mathbb{R}$ is condense discrete convex if and only if the corresponding discrete Hessian matrix is positive definite in D.

Proof: Consider the discrete function

$$f(x) = x^T A_f x = \sum_{i,j=1}^{2} a_{ij} x_i x_j$$

where $a_{ij} \in \mathbb{R}$ for all $1 \le i, j \le 2$, and $x \in \mathbb{Z}^2$. We prove the case for 2×2 matrix and $n \times n$ matrix case follows similarly. Suppose A_f is positive definite.

Case 1: If we let $x = (1, 0)$, then

$$f(x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = a_{11} > 0.$$

Case 2: If we let $x = (0, 1)$, then

$$f(x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = a_{22} > 0.$$

To show $A_f > 0$ for any $x \neq 0$ consider the following cases.

Case 1: If we let $x = (x_1, 0)$ with $x_1 \neq 0$. Then,

$$f(x) = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = a_{11}x_1^2 > 0 \Leftrightarrow a_{11} > 0.$$

Case 2: If we let $x = (x_1, x_2)$ with $x_2 \neq 0$. Let $x_1 = tx_2$ for some $t \in \mathbb{R}$. Therefore we have

$$f(x) = \left(a_{11}t^2 + 2a_{12}t + a_{22}\right) x_2^2$$

where $f(x) > 0 \Leftrightarrow \varphi(t) = a_{11}t^2 + 2a_{12}t + a_{22} > 0$ since $x_2 \neq 0$. Note that

$$\begin{aligned} \varphi'(t) &= 2a_{11}t + 2a_{12} = 0 \\ &\Rightarrow t^* = -\frac{a_{12}}{a_{11}} \\ \varphi''(t) &= 2a_{11}. \end{aligned}$$

If $a_{11} > 0$ then

$$\begin{aligned} \varphi(t) &\geq \varphi(t^*) = \varphi\left(-\frac{a_{12}}{a_{11}}\right) = \frac{-a_{12}^2}{a_{11}} + a_{22} \\ &= \frac{1}{a_{11}} \det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}. \end{aligned}$$

Therefore if $a_{11} > 0$ and the determinant given above is positive then $\varphi(t) > 0$ for all $t \in \mathbb{R}$. Conversely, if $f(x) > 0$ for every $x \neq 0$ then $\varphi(t) > 0$ for some t, therefore

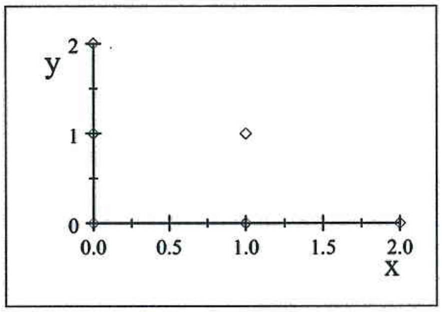
$$\begin{aligned} \varphi(t) &> 0 \Rightarrow a_{11} > 0, \ \textit{and} \ 4a_{12}^2 - 4a_{11}a_{22} = -4\det(A_f) < 0, \\ \varphi(t) &> 0 \Leftrightarrow a_{11} > 0 \ \textit{and} \ \det(A_f) > 0. \end{aligned}$$

which completes the proof.

3.1.1 An Example

Suppose we want to find the second difference of a function $f : \mathbb{Z}^2 \to \mathbb{R}$ defined by $f(x, y) = e^x + e^y$ on the the integer domain

$$S_1 = \{(0,0), (1,1), (0,1), (0,2), (1,0), (2,0)\}.$$



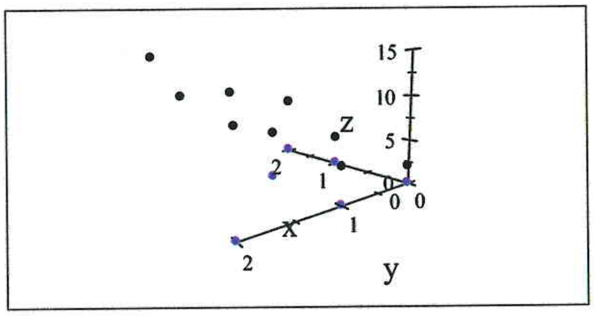
The first concern is whether we can find a quadratic discrete variable function in S_1 or not. As we want to find a quadratic function of the form

$$\begin{aligned} L(x) &= \frac{1}{2}\begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 & b_2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + c \\ &= \frac{1}{2}\left(a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_1x_2 + a_{22}x_2^2\right) + b_1x_1 + b_2x_2 + c \end{aligned}$$

we can find the coefficients corresponding to the approximation of f by using the values of f and determine the discrete quadratic function corresponding to f easily. This can be done by finding the values a_{ij}, b_i, and c by solving the system of equations

$$\begin{aligned} f(0,0) &= e^0 + e^0 = 2 = c \\ f(1,1) &= \frac{1}{2}\left(a_{11}(1)^2 + a_{12}(1)(1) + a_{21}(1)(1) + a_{22}1^2\right) + b_1(1) + b_2(1) + c \\ f(0,1) &= \frac{1}{2}a_{22}(1)^2 + b_2(1) + c \\ f(0,2) &= \frac{1}{2}a_{22}(2)^2 + b_2(2) + c \\ f(1,0) &= \frac{1}{2}a_{11}(1)^2 + b_1(1) + c \\ f(2,0) &= \frac{1}{2}a_{11}(2)^2 + b_1(2) + c \end{aligned}$$

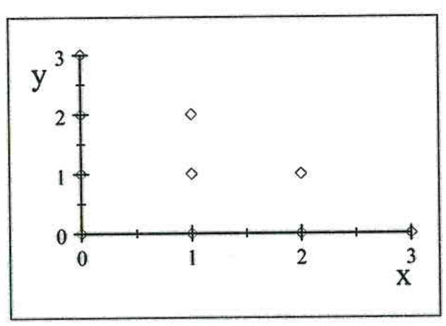
where there are six equations with six unknowns assuming that the matrix A is symmetric. As long as this system of equations has a solution, it is possible to find the quadratic approximation for any given function on a set where we can calculate the second differences.



Graph of $e^x + e^y$ on S_1.

The discussion above indicates that any well defined function on a set where the second difference elements exist in $\mathbb{Z}^2$ can have a quadratic approximation. In $\mathbb{Z}^n$, the number of elements in a set should be at least $\frac{1}{2}(n^2 + 3n + 2)$ to be able to find the quadratic approximation of the function in that neighborhood. For that reason, in chapter 4 it is assumed that the condense discrete convex set is large enough to support the second difference for the given function. We can also expand the set S_1 to another set

$$\begin{aligned} S_3 &= S_1 \cup S_2 \\ &= \{(0,0),(1,1),(0,1),(0,2),(1,0),(2,0)\} \cup \{(3,0),(2,1),(1,2),(0,3)\} \end{aligned}$$



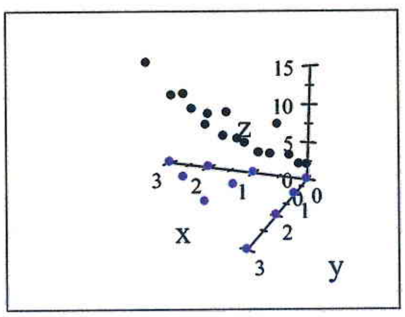
In this case, by using the subsets

$$D_1 = \{(1,0),(2,0),(3,0),(1,1),(1,2),(2,1)\},$$

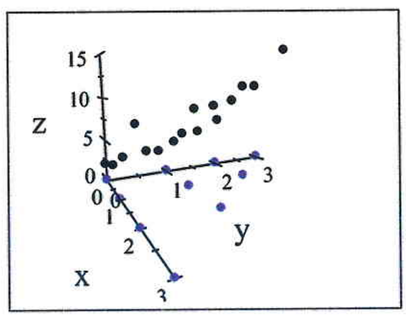
and

$$D_2 = \{(0,1),(0,2),(0,3),(1,1),(1,2),(2,1)\},$$

we find two other quadratic approximations for the function $f(x)$ in the neighborhood S_3. By finding the condense discrete convexity of these quadratic functions we find the convexity condition of the function in the neighborhood S_3.



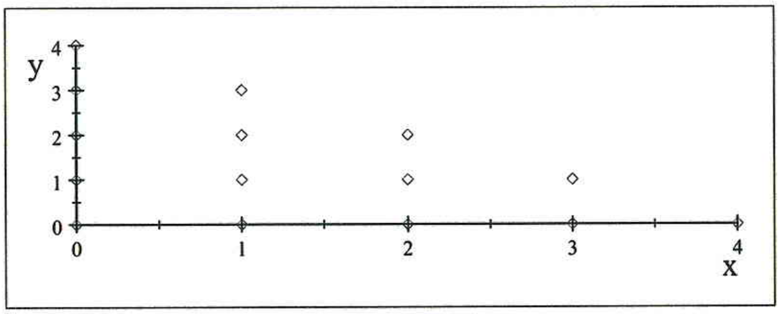
Graph of $e^x + e^y$ on S_3.



Graph of $e^x + e^y$ on S_3.

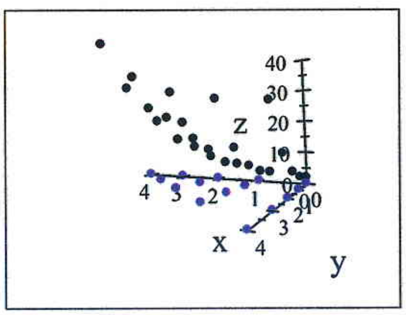
Neighborhood S_3 does not necessarily need extension further since the condense discrete convexity of the function for each point of the set S_3 is covered by using the method above. If we do want to extend the set S_3 to $\mathbb{Z}^2$ we can enlarge S_3 by adding elements as above where we can calculate the second differences. For example, we can expand S_3 to the set S_4 defined by

$$\begin{aligned} S_4 &= S_1 \cup S_2 \cup S_5 \\ &= \{(0,0),(1,1),(0,1),(0,2),(1,0),(2,0),(3,0),(2,1),(1,2),(0,3)\} \cup \\ &\quad \{(4,0),(3,1),(2,2),(1,3),(0,4)\} \end{aligned}$$

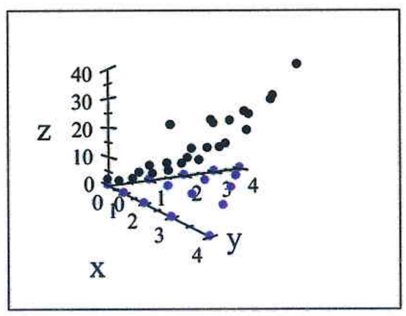


Set S_4.

The discrete function $f(x, y) = e^x + e^y$ with domain S_4 is illustrated below.



Graph of $e^x + e^y$ on S_4.



Graph of $e^x + e^y$ on S_4.

If we extend the domain of $f(x, y) = e^x + e^y$ from S_3 to $\mathbb{Z}^2$ then instead of checking the neighborhood convexity by finding what the coefficient matrix A is, we can check the second differences of the function itself. Therefore

$$H = \begin{bmatrix} e^x (e^2 - 2e + 1) & 0 \\ 0 & e^y (e^2 - 2e + 1) \end{bmatrix}$$

corresponds to the discrete Hessian matrix of $f(x, y) = e^x + e^y$. This Hessian matrix is clearly positive definite.

3.2 CONDENSE DISCRETE CONVEX FUNCTION OPTIMIZATION

To obtain minimization results for a given condense discrete convex function, we require the given condense discrete convex function to be C^1. By a C^1 condense discrete convex function $f : \mathbb{Z}^n \to \mathbb{R}$ we mean the extended function $f : \mathbb{R}^n \to \mathbb{R}$ to be a C^1 function. This extension depends on the given discrete function; however, the real extension of f has to agree with the discrete function f on the integer lattice. This extension might be possible by assuming the integer variable is a real variable. After defining the local and global minimum of condense discrete convex functions, we obtain convexity results for C^1 condense discrete convex functions. Condense discrete concave function maximization results follow similarly.

We let $\bigcup_{i=1}^{\infty} S_i = \mathbb{Z}^n$ where S_i is a non-empty sufficiently small condense discrete convex neighborhood to support quadratic approximation of f, $\bigcap_{i=1}^{\infty} S_i = \emptyset$, J is a finite index set, and $\{s_i\}$ is a singleton in $\mathbb{Z}^n$.

The partial derivative operator of a C^1 discrete function $f : \mathbb{Z}^n \to \mathbb{R}$ will be denoted by $\partial f(x) := \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}\right)$.

Definition 3.2: The local minimum of a condense discrete C^1 function $f : \mathbb{Z}^n \to \mathbb{R}$ is the minimal value of f in a local neighborhood $\bigcup_{i \in I} S_i$ which is also the smallest value in a neighborhood $N = \bigcup_{j \in J}\left(\bigcup_{i \in I_j} S_i\right)$ where J is a finite index set. The global minimum value of a condense discrete convex function $f : \mathbb{Z}^n \to \mathbb{R}$ is the minimum value of f in the entire integer space $\mathbb{Z}^n$.

We define the set of local minimums of a C^1 condense discrete convex function f by

$$\Psi = \{\rho = (\rho_1, ..., \rho_n) : \rho_i \in \{\lceil \gamma_i \rceil, \lfloor \gamma_i \rfloor\} \subset \mathbb{Z} \text{ for all } i\} \subset \mathbb{Z}^n.$$

where $\partial f(\gamma) = 0$ holds for $\gamma \in \mathbb{R}^n$. As the domain is $\mathbb{Z}^n$, we consider the solutions in Ψ where $\rho_i = \lceil \gamma_i \rceil$ or $\rho_i = \lfloor \gamma_i \rfloor$ is the solution for multivariable integer function f.

Example 3.1: We consider the following function $f : \mathbb{Z}^2 \to \mathbb{R}$ defined by

$$f(m, n) = \left(n - \frac{3}{100}m - \frac{1}{100}\right)^2 + \epsilon\left[(m-33)^2 + (n-1)^2\right]$$

where

$$\epsilon = \frac{1}{(33^2+1)\,100^2}.$$

Noting that the domain of f is $\mathbb{Z}^2$, the Hessian matrix corresponding to f is

$$H = \begin{bmatrix} \frac{18}{10^4} + 2\epsilon & \frac{-6}{100} \\ \frac{-6}{100} & 2 + 2\epsilon \end{bmatrix}$$

which is positive definite.

Supposing we want to find the local minimum corresponding to f starting from $(0,0)$, we need to visit the entire domain nodes and calculate all the corresponding function values which is computationaly costly. To be able to find the local minimum(s) we initially let the domain to be

$$S_1 = \{(0,0), (0,1), (1,0)\}.$$

Then the minimum in S_1 is

$$\begin{aligned} f(0,0) &= \left(0 - \frac{3}{100}0 - \frac{1}{100}\right)^2 + \frac{1}{(33^2+1)\,100^2}(0-33)^2 \\ &\quad + \frac{1}{(33^2+1)\,100^2}(0-1)^2 \\ &= 0.000\,2 \end{aligned}$$

since

$$\begin{aligned} f(0,1) &= \left(1 - \frac{3}{100}0 - \frac{1}{100}\right)^2 + \frac{1}{(33^2+1)\,100^2}(0-33)^2 \\ &\quad + \frac{1}{(33^2+1)\,100^2}(1-1)^2 \\ &= 0.98020 \\ f(1,0) &= \left(0 - \frac{3}{100}1 - \frac{1}{100}\right)^2 + \frac{1}{(33^2+1)\,100^2}(1-33)^2 \\ &\quad + \frac{1}{(33^2+1)\,100^2}(0-1)^2 \\ &= 0.001\,694 \end{aligned}$$

We extend the domain to $N_2 = \underset{1\leq i\leq 2}{\cup} S_i$ by defining

$$S_1 = \{(0,2), (1,0), (2,0)\}$$

The minimum in the domain S_2 can be found by using the following calculations:

$$\begin{aligned} f(0,2) &= \left(2 - \frac{3}{100}0 - \frac{1}{100}\right)^2 + \frac{1}{(33^2+1)\,100^2}(0-33)^2 + \frac{1}{(33^2+1)\,100^2}(2-1)^2 \\ &= 3.9602 \\ f(1,0) &= \left(0 - \frac{3}{100}1 - \frac{1}{100}\right)^2 + \frac{1}{(33^2+1)\,100^2}(1-33)^2 + \frac{1}{(33^2+1)\,100^2}(0-1)^2 \\ &= 1.694 \times 10^{-3} \\ f(2,0) &= \left(0 - \frac{3}{100}2 - \frac{1}{100}\right)^2 + \frac{1}{(33^2+1)\,100^2}(2-33)^2 + \frac{1}{(33^2+1)\,100^2}(0-1)^2 \\ &= 4.9883 \times 10^{-3} \end{aligned}$$

The minimum of f in S_2 occurs at $(1, 0, 1.694 \times 10^{-3})$ and in N_2 at $(0, 0, 0.0002)$. In its most general form, letting $k = i + j$ with $0 \leq i,\ j \leq k$, we can define

$$S_k = \{(i,j) : \text{for all } i \text{ and } j \text{ such that } k = i + j\}$$

and

$$N_k = \underset{1\leq l\leq k}{\cup} S_l.$$

Considering the neihgborhood N_{40}, it is easy to see that the local minimum occurs in N_{34} at $(33, 1, 0)$ which is the smallest value in $N = \underset{1\leq i\leq 40}{\cup} N_i \subset \mathbb{Z}^2$ (These calculations can be done by writing a simple computer program.) As it is noted above, it is costly to calculate the local minimum of f by visiting every node; therefore, noting that f is a C^1 function with domain $\mathbb{Z}^2$, the practical way to find the local minimum of f is by calculating $\partial f = 0$;

$$\begin{aligned} \frac{\partial f}{\partial m} &= -\frac{6}{100}\left(n - \frac{3}{100}m - \frac{1}{100}\right) + \frac{2}{(33^2+1)\,100^2}(m-33) = 0 \\ \frac{\partial f}{\partial n} &= 2\left(n - \frac{3}{100}m - \frac{1}{100}\right) + \frac{2}{(33^2+1)\,100^2}(n-1) = 0 \end{aligned}$$

which is satisfied when $(m, n) = (33, 1)$. It is easy to see that the local minimum is also the global minimum and vice versa in this case.

Lemma 3.1: Let $f : \mathbb{Z}^n \to \mathbb{R}$ be a C^1 condense discrete convex function in $N \subset \mathbb{Z}^n$. Then there exists a local minimum value in $N \subset \mathbb{Z}^n$ such that

$$f_0 = \min_{\beta \in \Psi} \{f(\beta)\}.$$

Proof: Let $f : N \to \mathbb{R}$ be a C^1 strict condense discrete convex function. Therefore f has a local minimum value $f(x_0)$ in some neighborhood $S = \underset{i \in I}{\cup} S_i$ by theorem 3.1. By definition of N, $\underset{i \in I}{\cup} S_i \subseteq N$ hence $f(x_0)$ is also the local minimum in the neighborhood N.

It is well known that the local minimum of a C^1 function f is obtained when the system of equations

$$\frac{\partial f(x)}{\partial x_i} = \lim_{t \to 0} \frac{f(x + te_i) - f(x)}{t} = 0$$

is solved simultaneously for all $i, 1 \leq i \leq n$. We first find $\partial f(x) = 0$ which implies the existence of a $\gamma_i \in \mathbb{R}$ for all i. Noting that the domain is $\mathbb{Z}^n$, we take the ceiling and floor of the components of γ_i to obtain the minimal point which consist of integer numbers $\lfloor \gamma_i \rfloor$ or $\lceil \gamma_i \rceil$ for all i. This gives a local minimum point $\beta \in \Psi$ and the corresponding value $f_0 = \min_{\beta \in \Psi} \{f(\beta)\}$.

Based on the statement of lemma 3.1 we try to find the minimum of f where β belongs to Ψ. Based on the set definition of Ψ we have to have $\partial f(\gamma) = 0$ meaning

$$\partial f(x) := \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right) = 0.$$

Here the basic assumption is simply there exists a β in Ψ so that the function value holds. If there does not exist such a β then clearly f_0 does not exist. $\min_{\beta \in \Psi}$ corresponds to existing at least one β satisfying the first partial derivative equal to zero since $\beta \in \Psi$. Considering the example $f(n, m) = e^n e^m$ in $\mathbb{Z}^2$, we cannot have the first partial derivatives of the function equal to zero for any point in $\mathbb{Z}^2$, therefore there does not exist a β to start checking the minimum for the function, therefore the lemma cannot be applied to this function.

The following result for condense discrete convex functions is a result similar to a result in real convex analysis.

Theorem 3.2: Let $f : \mathbb{Z}^n \to \mathbb{R}$ be a C^1 strict condense discrete convex function. Then the set of local minimums of f form a set of global minimums and vice versa.

Proof: Suppose $f : \mathbb{Z}^n \to \mathbb{R}$ be a C^1 condense discrete convex function. Let $\overset{\infty}{\underset{i=1}{\cup}} S_i = \mathbb{Z}^n$ where S_i are sufficiently small condense discrete neighborhoods that support quadratic approximation of f for all i, and $\overset{\infty}{\underset{i=1}{\cap}} S_i = \emptyset$. Let Ω_1 be the set of local minimum points of f in $\mathbb{Z}^n$, and Ω_2 be the set of global minimum points of f in $\mathbb{Z}^n$.

Let $f : \mathbb{Z}^n \to \mathbb{R}$ be a C^1 condense discrete convex function and suppose f has global minimum points in $\mathbb{Z}^n = \overset{\infty}{\underset{i=1}{\cup}} S_i$. Noting that f is nonlinear, there exists a finite collection of S_i, $\underset{i\in I_0}{\cup} S_i$, where the global minimum points are located. The solution set of $\frac{\partial f(x)}{\partial x_j} = 0$ for all $j, 1 \leq j \leq n$, gives the set of local minimums in S_i. Therefore for all $x \in \Omega_2$ there exists a set of integer vectors $y \in \Omega_1$ such that $\min_{x\in\Omega_1} f(x) = f(y)$ which indicates $\Omega_2 \subset \Omega_1$ since $\underset{i\in I_0}{\cup} S_i \subset N \subset \mathbb{Z}^n$.

Now suppose there exists a vector x_0 in a local neighborhood $S = \underset{i\in I_1}{\cup} S_i$ such that $x_0 \notin \Omega_2$ (Note that x_0 is not necessarily an element of Ω_1 since it is a local minimum in a local setting). x_0 is a local minimum which is not a global minimum in S, therefore there exist x_1 and y_1 such that $f(x_0) > f(x_1) > f(y_1)$ in $N = \underset{j\in J}{\cup} \left(\underset{i\in I_j}{\cup} S_i \right) \supset S$ where y_1 becomes the new local minimum of the local neighborhood N. Therefore y_0 is the new local minimum of N where x_0 is not a local minimum of N. Suppose y_0 is a local minimum that is not a global minimum otherwise it would be an element of Ω_2. Continuing to enlarge the local obtained neighborhoods in this way to the entire space $\mathbb{Z}^n$, we obtain a set of points in a local neighborhood D of

$\mathbb{Z}^n$ where local minimum points $x \in \Omega_1$ satisfy $f(x) < f(y)$ for all $y \in \mathbb{Z}^n - D$. Therefore $x \in \Omega_2$ and hence $\Omega_1 \subset \Omega_2$ which completes the proof.

Next we consider an adaptation of Rosenbrook's function suggested by Yüceer (2002) and show that this function is a condense discrete convex function.

3.2.1 AN EXAMPLE

Yüceer (2002) shows that the adaptation of Rosenbrook's function

$$g(k,\mu) = 25(2\mu - k)^2 + \frac{1}{4}(2-k)^2 \quad \text{where } k, \mu \in \mathbb{Z}. \tag{3.3}$$

is not a discretely convex function when continuous variables are restricted to the integer lattice. Here, we first prove the condense discrete convexity of the function given in (3.3) and then show that the set of local minimums is also the set of global minimums.

The diagonal elements of the discrete Hessian matrix that corresponds to $g(k,\mu)$ are

$$\begin{aligned}
\nabla_{11} g(k,\mu) &= 25(2\mu - k - 2)^2 + \frac{1}{4}k^2 - 50(2\mu - k - 1)^2 \\
&\quad -\frac{1}{2}(1-k)^2 + 25(2\mu - k)^2 + \frac{1}{4}(2-k)^2 \\
&= \frac{101}{2} > 0. \\
\nabla_{22} g(k,\mu) &= 25(2\mu + 4 - k)^2 - 50(2\mu + 2 - k)^2 + 25(2\mu - k)^2 \\
&= 200.
\end{aligned}$$

By the symmetry of the discrete Hessian matrix, the off diagonal elements of the discrete Hessian matrix are

$$\begin{aligned}
\nabla_{12} g &= \nabla_{21} g = 25(2\mu + 2 - k - 1)^2 - 25(2\mu - k - 1)^2 \\
&\quad -25(2\mu + 2 - k)^2 + 25(2\mu - k)^2 \\
&= -100.
\end{aligned}$$

Therefore

$$\begin{aligned}\det(H) &= 200.\frac{101}{2} - (100)^2 \\ &= 100.101 - (100)^2 \\ &= 100.\end{aligned}$$

This indicates that the discrete Hessian matrix is positive definite. Therefore, the adaptation of the Rosenbrook's function given in the equality (3.3).

It is important to note that the optimization results stated in this section assume the condense discrete convexity of a discrete function to find a local or global minimum if it exists. Based on Miller's (1971) definition of discrete convexity, the adaptation of Rosenbrock's function defined by (3.3) is not a discrete convex function; therefore, to obtain minimization results for (3.3), we have to check the function value at every point of the domain in every possible neighborhood to find out where the minimum value is since it is not a discrete convex function from Miller's discrete convexity point of view. However, by using the condense discrete convexity and the corresponding optimization results, we only need to first check whether it is a condense discrete convex function or not by finding the corresponding discrete Hessian matrix, and if it is a condense discrete convex then find the first partial derivatives to obtain the corresponding optimization results (clearly in the case when a solution for the equation $\partial f(x) = 0$ exists.) Therefore, calculating the minimum value of a condense discrete convex function is more practical compared to calculating the minimum value of a function by using Miller's discrete convexity definition. From an algorithmic solution point of view, even in a bounded set there is a huge computation difference between these two methods. Note that in lemma 3.1 the discrete set N is not necessarily a bounded set.

Next we show that the set of local minimums of the adaptation of the Rosenbrock's function is also the set of global minimums. Clearly, g is a C^1 function

therefore

$$\partial g\,(k,\mu)=0\Rightarrow\begin{cases}\frac{\partial g}{\partial k}=-50(2\mu-k)-\frac{1}{2}(2-k)=0\\ \frac{\partial g}{\partial \mu}=100(2\mu-k)=0\end{cases}$$

where simultaneous solution of this system of two equations indicate $k=2$ and $\mu=1$. Therefore the minimal value is $g(2,1)=0$. Since the adaptation of Rosenbrock's function is a C^1 condense discrete convex function, the local minimum point set which is the singleton $\{(2,1,0)\}$ is also the set of global minimum points.

CHAPTER 4

MIXED CONVEXITY AND OPTIMIZATION

In this section, mixed convexity and optimization of mixed (real and integer) variable set and function definitions will be introduced. In addition, mixed convexity and optimization results will be stated and proven. Examples of a mixed convex and a non-convex mixed function will be given in the last section. Many examples of optimization problems with mixed variable functions can be found in queueing systems and network designs where the results obtained in this work can be applied.

4.1 DEFINITIONS AND RESULTS

Let $V_1 \subseteq \mathbb{Z}^n$ be a condense discrete convex set and $V_2 \subseteq \mathbb{R}^m$ be a real convex set. A mixed convex set is the set of the form $V = V_1 \times V_2 \subseteq \mathbb{Z}^n \times \mathbb{R}^m$. Throughout this work g will be assumed to be a C^2 function with respect to its real variable unless stated otherwise, and the indices i,j and k, l will be used for the integer and real variables, respectively.

Definition 4.1: A mixed function $g : V \to \mathbb{R}$ on a mixed convex set $V \subseteq \mathbb{Z}^n \times \mathbb{R}^m$ is defined to be mixed convex if its quadratic approximation

$$g\left(x, y\right) = \frac{1}{2} x^T A x + x^T B^T y + \frac{1}{2} y^T C y + b^T x + c^T y + d \tag{4.1}$$

in the neighborhood V is strictly positive where A and C are the symmetric coefficient matrices of the quadratic approximation of g with respect to x and y, respectively. h is called mixed concave if $-h$ is mixed convex.

Proposition 4.1: Let $g : V \to \mathbb{R}$ be defined on a mixed convex set $V \subseteq \mathbb{Z}^n \times \mathbb{R}^m$ with its quadratic approximation given in (4.1). The coefficient matrix H_g corresponding to g is the symmetric matrix

$$H_g = \begin{bmatrix} [\nabla_{ij}(g)]_{n\times n} & \left[\frac{\partial}{\partial y_k}\nabla_j(g)\right]_{n\times m} \\ \left[\nabla_i\frac{\partial}{\partial y_l}(g)\right]_{m\times n} & \left[\frac{\partial^2 g}{\partial y_k \partial y_l}\right]_{m\times m} \end{bmatrix} \tag{4.2}$$

Proof: The symmetry of the matrix $[\nabla_{ij}(g)]_{n\times n}$ in the mixed Hessian matrix H_g follows from proposition 3.1. Clearly

$$\frac{\partial^2 g}{\partial y_k \partial y_l} = \frac{\partial^2 g}{\partial y_l \partial y_k}$$

yields to a symmetric matrix. The off diagonal block matrices of H_g satisfy the symmetry condition

$$\begin{aligned} \frac{\partial}{\partial y_k}(\nabla_j g(x,y)) &= \frac{\partial}{\partial y_k}(g(x+e_j,y) - g(x,y)) \\ &= \frac{\partial}{\partial y_k} g(x+e_j,y) - \frac{\partial}{\partial y_k} g(x,y) \\ &= \nabla_j \frac{\partial}{\partial y_k} g(x,y) \end{aligned}$$

for all j and k. Therefore H_g is a symmetric matrix.

Assuming A is symmetric, by proposition 3.1

$$\nabla_{ij}(g(x)) = a_{ij}$$

holds for all i and j. Therefore

$$A = [a_{ij}]_{n\times n} = [\nabla_{ij} g]_{n\times n}.$$

Straightforward calculations indicate

$$\left[\frac{\partial^2 g}{\partial y_k \partial y_l}\right]_{m\times m} = C.$$

The off diagonal elements satisfy

$$\begin{aligned}
\frac{\partial}{\partial y_k}\nabla_j\left(g\left(x,y\right)\right) &= \frac{\partial}{\partial y_k}\left(\left(x+e_j\right)^T B^T y - x^T B^T y\right) \\
&= \frac{\partial}{\partial y_k}\left(x^T B^T y + e_j^T B^T y - x^T B^T y\right) \\
&= \frac{\partial}{\partial y_k}\left(e_j^T B^T y\right) \\
&= b_{jk}.
\end{aligned}$$

Proposition 4.2: The coefficient matrix H_g of $g : V \to \mathbb{R}$ given in proposition 4.1 satisfies the properties of the mixed Hessian matrix corresponding to real convex functions. That is, H_g is linear with respect to the condense mixed functions, symmetric, and vanishes when g is mixed affine.

Proof: Let $g_t : W_t \to \mathbb{R}$ be condense mixed functions where W_t are mixed convex sets (for $t = 1, 2$) with the corresponding coefficient matrices

$$H_{g_t} = \begin{bmatrix} A_{g_t} & B_{g_t} \\ B_{g_t} & C_{g_t} \end{bmatrix},\ t = 1, 2.$$

Note that

$$\begin{aligned}
\frac{\partial}{\partial y_k}\nabla_j\left(g_1+g_2\right) &= \frac{\partial}{\partial y_k}\left\{g_1\left(x+e_j,y\right)+g_2\left(x+e_j,y\right)-\left[g_1\left(x,y\right)+g_2\left(x,y\right)\right]\right\} \\
&= \frac{\partial}{\partial y_k}\left(g_1\left(x+e_j,y\right)-g_1\left(x,y\right)\right)+\frac{\partial}{\partial y_k}\left(g_2\left(x+e_j,y\right)-g_2\left(x,y\right)\right)
\end{aligned}$$

indicating

$$\begin{aligned}
\left[\frac{\partial}{\partial y_k}\nabla_j\left(g_1+g_2\right)\right]_{n\times m} &= \left[\frac{\partial}{\partial y_k}\nabla_j\left(g_1\right)\right]_{n\times m}+\left[\frac{\partial}{\partial y_k}\nabla_j\left(g_2\right)\right]_{n\times m} \\
&= B_{g_1}+B_{g_2}
\end{aligned}$$

Therefore, by using the symmetry property obtained in proposition 4.1,

$$\begin{aligned}
H_{g_1+g_2} &= \begin{bmatrix} \left[\nabla_{ij}\left(g_1+g_2\right)\right]_{n\times n} & \left[\frac{\partial}{\partial y_k}\nabla_j\left(g_1+g_2\right)\right]_{n\times m} \\ \left[\frac{\partial}{\partial y_k}\nabla_j\left(g_1+g_2\right)\right]_{m\times n} & \left[\frac{\partial^2}{\partial y_k\partial y_l}\left(g_1+g_2\right)\right]_{m\times m} \end{bmatrix} \\
&= \begin{bmatrix} A_{g_1+g_2} & B_{g_1}+B_{g_2} \\ B_{g_1}+B_{g_2} & \frac{\partial^2 g_1}{\partial y_k\partial y_l}+\frac{\partial^2 g_2}{\partial y_k\partial y_l} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
&= \begin{bmatrix} A_{g_1} + A_{g_2} & B_{g_1} + B_{g_2} \\ B_{g_1} + B_{g_2} & C_{g_1} + C_{g_2} \end{bmatrix} \\
&= \begin{bmatrix} A_{g_1} & B_{g_1} \\ B_{g_1} & C_{g_1} \end{bmatrix} + \begin{bmatrix} A_{g_2} & B_{g_2} \\ B_{g_2} & C_{g_2} \end{bmatrix} \\
&= H_{g_1} + H_{g_2}
\end{aligned}$$

which also proves the linearity of the second difference operator with respect to the condense mixed functions. The symmetry condition is proven in proposition 4.1.

Considering the mixed affine function g,

$$g(x,y) = \sum_{i=1}^{n} d_i x_i + \sum_{k=1}^{m} w_k x_k,$$

the second difference operator vanishes since $\nabla_i(g) = b_i$ and $\nabla_{ij}(g) = 0$ for all i and j. Similarly $\nabla_i \frac{\partial}{\partial y_k} = 0$ and $\frac{\partial^2}{\partial y_k \partial y_l} = 0$ for all i, k, and l.

Theorem 4.1: A function $g : V \to \mathbb{R}$ is strict mixed convex if and only if the corresponding mixed Hessian matrix H_g is positive definite in V.

Proof: In the case when $m = 0$ the proof follows from theorem 3.1. In the case when $n = 0$ the result is well known from real convexity theory. Consider the mixed function

$$\begin{aligned}
g(x,y) &= \frac{1}{2}x^T Ax + x^T B^T y + \frac{1}{2}y^T Cy \\
&= ax^2 + 2bxy + cy^2
\end{aligned}$$

where $a, b, c \in \mathbb{R}$ and $(x,y) \in \mathbb{Z} \times \mathbb{R}$. We prove the case for 2×2 matrix and $(n+m) \times (n+m)$ matrix case follows similarly. Let $z = (x,y)$. Suppose H_g is positive definite.

Case 1: If we let $z = (1,0)$, then

$$g(z) = ax^2 + 2bxy + cy^2 = a > 0.$$

Case 2: If we let $z = (0, 1)$, then

$$g(z) = ax^2 + 2bxy + cy^2 = c > 0.$$

To show $H_g > 0$ for any $z \neq 0$ consider the following cases.

Case 1: If we let $x = (x, 0)$ with $x \neq 0$, then

$$g(x) = ax^2 + 2bxy + cy^2 = ax^2 > 0 \Leftrightarrow a > 0.$$

Case 2: Let $x = (x, y)$ with $y \neq 0$, and $x = ty$ for some $t \in \mathbb{R}$. Therefore we have

$$g(x) = \left(at^2 + 2bt + c\right) y^2$$

where $g(z) > 0 \Leftrightarrow \varphi(t) = at^2 + 2bt + c > 0$ since $y \neq 0$. Note that

$$\begin{aligned}
\varphi'(t) &= 2at + 2b = 0 \\
&\Rightarrow t^* = -\frac{b}{a} \\
\varphi''(t) &= 2a.
\end{aligned}$$

If $a > 0$ then

$$\begin{aligned}
\varphi(t) &\geq \varphi(t^*) = \varphi\left(-\frac{b}{a}\right) = \frac{-b^2}{a} + c \\
&= \frac{1}{a} \det \begin{bmatrix} a & b \\ b & c \end{bmatrix}.
\end{aligned}$$

Therefore if $a > 0$ and the determinant given above is positive then $\varphi(t) > 0$ for all $t \in \mathbb{R}$. Conversely, if $g(z) > 0$ for every $z \neq 0$ then $\varphi(t) > 0$ for some t, therefore

$$\begin{aligned}
\varphi(t) &> 0 \Rightarrow a > 0, \ \textit{and} \ 4b^2 - 4ac = -4 \det(H_g) < 0, \\
\varphi(t) &> 0 \Leftrightarrow a > 0 \ \ \textit{and} \ \det(H_g) > 0
\end{aligned}$$

which completes the proof.

To obtain minimization results for a given mixed convex function, the given mixed convex function will be required to be C^1 with respect to all of its variables. After defining the local and global minimum point concepts of mixed convex functions, we prove optimization results for C^1 mixed convex functions. Mixed concave function maximization results follow similarly.

Let $\mathbb{Z}^n \times \mathbb{R}^m = \bigcup_{i=1}^{\infty} S_i \times \bigcup_{j=1}^{\infty} R_j$ where $S_i \times R_j$ is a non-empty sufficiently small mixed convex neighborhood to support quadratic approximation of g, $\bigcap_{i=1}^{\infty} S_i = \emptyset$ for all S_i where S_i have at least one common element for all $i \in I$, I is a finite index set, and $\{(s_i, r_j)\}$ is a singleton in $\mathbb{Z}^n \times \mathbb{R}^m$.

The partial derivative operator of a C^1 mixed function $g : \mathbb{Z}^n \times \mathbb{R}^m \to \mathbb{R}$ will be denoted by

$$\partial g\left(x\right) := \left(\frac{\partial g}{\partial x_1}, \frac{\partial g}{\partial x_2}, ..., \frac{\partial g}{\partial x_n}, \frac{\partial g}{\partial y_1}, \frac{\partial g}{\partial y_2}, ..., \frac{\partial g}{\partial y_m}\right).$$

Definition 4.2: The local minimum of a mixed C^1 function $g : \mathbb{Z}^n \times \mathbb{R}^m \to \mathbb{R}$ is the minimal value of g in a local neighborhood $\bigcup_{i \in I} S_i \times \bigcup_{j \in I} R_j$ which is also the smallest value in a neighborhood $M = N \times R$ where I is a finite index set and $R = \bigcup_{j \in I} R_j$. The global minimum value of a mixed convex function $g : \mathbb{Z}^n \times \mathbb{R}^m \to \mathbb{R}$ is the minimum value of g in the entire mixed space $\mathbb{Z}^n \times \mathbb{R}^m$.

Define the set of local minimums of a C^1 mixed convex function g by

$$\Psi = \{\rho = (\rho_1, ..., \rho_n, \alpha_1, ..., \alpha_m) : \rho_i \in \{\lceil \gamma_i \rceil, \lfloor \gamma_i \rfloor\} \subset \mathbb{Z}\ \forall i, \alpha_j \in \mathbb{R}\ \forall j\} \subset \mathbb{Z}^n \times \mathbb{R}^m$$

where $\partial g\left(\gamma, \alpha\right) = 0$ holds for $(\gamma, \alpha) \in \mathbb{R}^{n+m}$. We consider the solutions in Ψ where $\rho_i = \lceil \gamma_i \rceil$ or $\rho_i = \lfloor \gamma_i \rfloor$ is the solution for multivariable mixed function g.

Lemma 4.1: Let $g : M \to \mathbb{R}$ be a C^1 mixed convex function in $M \subset \mathbb{Z}^n \times \mathbb{R}^m$. Then there exists a local minimum value in M such that

$$g_0 = \min_{\beta \in \Psi} \left\{g\left(\beta\right)\right\}.$$

Proof: Let $g : M \to \mathbb{R}$ be a C^1 strict mixed convex function. Suppose $\partial g(x, y) = 0$ holds in some neighborhood $S = \bigcup_{i \in I} S_i \times \bigcup_{j \in I} R_j \subseteq M$ for all $(x, y) \in M$. Therefore the local minimum of C^1 function g is obtained when the system of equations

$$\begin{aligned} \frac{\partial g(x)}{\partial x_i} &= \lim_{t \to 0} \frac{g(x + te_i) - g(x)}{t} = 0 \\ \frac{\partial g(x)}{\partial y_j} &= 0 \end{aligned}$$

are solved simultaneously for all i, $1 \le i \le n$, and for all j, $1 \le j \le m$. This indicates the existence of a $(\gamma, \alpha) \in \mathbb{R}^{n+m}$. For the integer variables in domain $\mathbb{Z}^n$, we take the ceiling and floor of the components of γ_i to obtain the minimal point which consist of integer numbers $\lfloor \gamma_i \rfloor$ or $\lceil \gamma_i \rceil$ for all i, $1 \le i \le n$. This gives a local minimum point $(\beta, \alpha) \in \Psi$ since there exists a unique minimum value of a real convex function and the corresponding value $g_0 = \min_{\beta \in \Psi} \{g(\beta)\}$.

Suppose $\partial g(x, y) \neq 0$ for some x and y. Then either $\partial g(x, y) > 0$ or $\partial g(x, y) < 0$ holds which in either case the minimum value is obtained for the boundary values of M for x and y satisfying $\partial g(x, y) \neq 0$.

The following theorem for mixed convex functions has a similar statement to the results obtained for real and condense discrete convex functions. It is evident that a condense mixed convex function can have more than one global minimum point since a condense discrete convex function can have more than one minimum point. A simple example of a condense mixed convex function $g : \mathbb{Z}^n \times \mathbb{R}^m \to \mathbb{R}$ with 2^n minimum points is

$$g(x) = \sum_{i=1}^{n} (x_i - 0.5)^2 + \sum_{j=1}^{m} (y_j - 0.5)^2.$$

Theorem 4.2: Let $g : \mathbb{Z}^n \times \mathbb{R}^m \to \mathbb{R}$ be a C^1 strict mixed convex function which has local and global minimum points. Then the set of local minimum points of g form a set of global minimum points and vice versa.

Proof: Suppose $g : \mathbb{Z}^n \times \mathbb{R}^m \to \mathbb{R}$ is a C^1 mixed convex function. Let $\bigcup_{i=1}^{\infty} S_i \times \bigcup_{j=1}^{\infty} R_j = \mathbb{Z}^n \times \mathbb{R}^m$ where $S_i \times R_j$ are sufficiently small condense mixed neighborhoods supporting quadratic approximation of g for all i and j, and $\bigcap_{i=1}^{\infty} S_i = \emptyset$. Let Φ_1 be the set of local minimum points of g in $\mathbb{Z}^n \times \mathbb{R}^m$, and Φ_2 be the set of global minimum points of g in $\mathbb{Z}^n \times \mathbb{R}^m$.

Let $g : \mathbb{Z}^n \times \mathbb{R}^m \to \mathbb{R}$ be a C^1 mixed convex function and suppose g has global minimum points in $\mathbb{Z}^n \times \mathbb{R}^m$. Noting that g is strict mixed convex, there exists a collection of $S_i \times R_j \subset \bigcup_{i \in I_0} S_i \times \bigcup_{j \in I_0} R_j$ where the global minimum points are located. Considering the quadratic approximation of g in the neighborhood $S_i \times R_j$, the solution set of $\partial g(z) = 0$ gives the set of local minimums in $S_i \times R_j$. Therefore for all $z_2 \in \Phi_2$ there exists a set of vectors $z_1 \in \Phi_1$ such that $\min_{z_1 \in \Phi_1} g(z_1) = g(z_2)$ which indicates $\Phi_2 \subset \Phi_1$ since $\bigcup_{i \in I_0} S_i \times \bigcup_{j \in I_0} R_j \subset \mathbb{Z}^n \times \mathbb{R}^m$.

Now suppose there exists a vector $z_0 = (x_0, y_0)$ in a local neighborhood $M_1 = \bigcup_{i \in I_1} S_i \times \bigcup_{j \in I_1} R_j$ such that $z_0 \notin \Phi_2$ (Note that z_0 is not necessarily an element of Φ_1 since it is a local minimum in a local setting). z_0 is a local minimum which is not a global minimum in M_1, therefore there exist z_1 and z_2 such that $g(z_0) > g(z_1) > g(z_2)$ in $M_2 = \bigcup_{j \in J} \left(\bigcup_{i \in I_j} S_i \right) \times \bigcup_{j \in J} \left(\bigcup_{i \in I_j} R_i \right) \supset M_1$ where z_2 becomes the new local minimum of the local neighborhood M_2. Therefore z_2 is the new local minimum of M_2 where z_0 is not a local minimum of M_2. Suppose z_2 is a local minimum that is not a global minimum otherwise it would be an element of Φ_2. Continuing to enlarge the local obtained neighborhoods in this way to the entire space $\mathbb{Z}^n \times \mathbb{R}^m$, a set of points in a local neighborhood V of $\mathbb{Z}^n \times \mathbb{R}^m$ is obtained where local minimum points $z \in \Phi_1$ satisfy $g(z) < g(\overline{z})$ for all $\overline{z} \in \mathbb{Z}^n - V$. Therefore $z \in \Phi_2$ and hence $\Phi_1 \subset \Phi_2$ which completes the proof.

Example 4.1: If a function is real convex with respect to its real variables when the integer variable is assumed constant, and if it is condense discrete convex when the real variables are assumed constant doesn't necessarily imply the mixed convexity of the proposed function. A simple counter example can be seen by choosing $\Psi : \mathbb{Z} \times \mathbb{R} \rightarrow \mathbb{R}$, $\Psi(\alpha, \beta) = (\alpha^2 + 0.5)(\beta^2 + 1)$. In this case,

$$H = \begin{bmatrix} 2(\beta^2 + 1) & 2\beta(2\alpha + 1) \\ 2\beta(2\alpha + 1) & 2(\alpha^2 + 0.5) \end{bmatrix}$$

and the choice of $\alpha = \beta = 1$ gives $\det(H) = -24$ which shows that Ψ is not a strict mixed convex function; however, it is easy to see that for each fixed α, $(\beta^2 + 1)$ is strict convex and for each fixed β,

$$\begin{aligned} \nabla_{11}\Psi &= \left[(\alpha + 2)^2 - 2(\alpha + 1)^2 + \alpha^2\right](\beta^2 + 1) \\ &= 2(\beta^2 + 1) \end{aligned}$$

strictly positive.

Example 4.2: An example which shows that the global minimum value is unique while the point that corresponds to the global minimum value is not necessarily unique is as follows:

Define a function $\Im : \mathbb{Z}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ by

$$\Im(\alpha, \beta) = \sum_{i=1}^{n} (\alpha_i - 1.5)^2 + \sum_{j=1}^{m} (\beta_j - j)^2$$

where $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n) \in \mathbb{Z}^n$ and $\beta = (\beta_1, \beta_2, ..., \beta_m) \in \mathbb{R}^m$. The global minimum points of $\Im$ are $(\alpha_i, \beta_j) \in \mathbb{Z}^n \times \mathbb{R}^m$ such that $\alpha_i = 1, 2$, and $\beta_j = j$ where the corresponding global minimum value is $0.25n$.

In the next section, it will be shown that the Rosenbrock's function with mixed variables is not a mixed convex function, and the adaptation of Rosenbrock's function (3.3) will be shown to be a mixed convex function.

4.2 EXAMPLES OF CONVEX AND NON-CONVEX MIXED FUNCTIONS

In this section examples of convex and non-convex mixed condense functions will be given.

4.2.1 AN EXAMPLE OF A MIXED CONVEX FUNCTION

The adaptation of Rosenbrock's function given in equation (3.3) failed to be a strong discrete convex function in domain $\mathbb{Z}^2$ as it was shown in Yüceer (2002). In chapter 3, equation (3.3) is shown to be a condense discrete convex function with the corresponding positive definite discrete Hessian matrix

$$\begin{aligned} H &= \begin{bmatrix} \nabla_{11} g(x,y) & \nabla_{12} g(x,y) \\ \nabla_{21} g(x,y) & \nabla_{22} g(x,y) \end{bmatrix} \\ &= \begin{bmatrix} 200 & -100 \\ -100 & \frac{101}{2} \end{bmatrix} \end{aligned}$$

In this section we consider the function introduced in equation (3.3) with domain $\mathbb{Z} \times \mathbb{R}$. The mixed Hessian matrix corresponding to $g : \mathbb{Z} \times \mathbb{R} \to \mathbb{R}$ will be shown to be a positive definite matrix after calculating the mixed Hessian matrix components. The diagonal elements of the mixed Hessian matrix that correspond to the function given in equation (3.3) are

$$\begin{aligned} \nabla_{11} g(x,y) &= 25(2y-x-2)^2 + \frac{1}{4}x^2 - 50(2y-x-1)^2 \\ &\quad -\frac{1}{2}(1-x)^2 + 25(2y-x)^2 + \frac{1}{4}(2-x)^2 \\ &= 25(2y-x)^2 - 100(2y-x) + 100 + \frac{1}{4}x^2 \\ &\quad -50(2y-x)^2 + 100(2y-x) - 50 \\ &\quad -\frac{1}{2}(1-x)^2 + 25(2y-x)^2 + \frac{1}{4}(2-x)^2 \end{aligned}$$

$$= 50 + \frac{1}{4}x^2 - \frac{1}{2}(1-x)^2 + \frac{1}{4}(2-x)^2$$
$$= 50 + \frac{x^2}{4} - \frac{1}{2} + x - \frac{x^2}{2} + 1 - x + \frac{x^2}{4}$$
$$= \frac{101}{2} > 0$$

and

$$\frac{d^2 g(x,y)}{dy^2} = 200$$

By the symmetry of the mixed Hessian matrix, the off diagonal elements of the mixed Hessian matrix are

$$\begin{aligned} \frac{d}{dy}(\nabla_1 g(x,y)) &= \nabla_1\left(\frac{dg(x,y)}{dy}\right) = 100(2y-x-1) - 100(2y-x) \\ &= -100 \end{aligned}$$

Therefore

$$\begin{aligned} \det(H) &= 200.\frac{101}{2} - (100)^2 \\ &= 100.101 - (100)^2 \\ &= 100 \end{aligned}$$

indicating positive definite mixed Hessian matrix. By theorem 4.1, the adaptation of Rosenbrock's function given in equation (3.3) is a strict mixed convex function. Clearly, g is a C^1 function hence

$$\partial g(x,y) = 0 \Rightarrow \begin{cases} \frac{\partial g}{\partial x} = -50(2y-x) - \frac{1}{2}(2-x) = 0 \\ \frac{\partial g}{\partial y} = 100(2y-x) = 0 \end{cases}$$

where simultaneous solution of this system of two equations indicate $x = 2$ and $y = 1$. Therefore the minimal value is $g(2,1) = 0$. Since (3.3) is a C^1 strict mixed convex function, the local minimum point set which is the singleton $\{(2,1,0)\}$ is also the set of global minimum points.

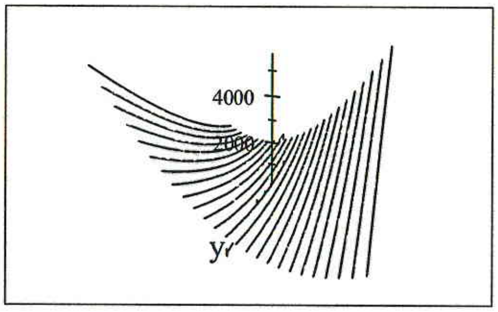


Fig.4.1: Rosenbrock's function $f : \mathbb{R}^2 \to \mathbb{R}$

4.2.2 AN EXAMPLE OF A FUNCTION THAT IS NOT MIXED CONVEX

It is well known and easy to check that Rosenbrock's function defined by

$$\begin{aligned} f &: \mathbb{R}^2 \to \mathbb{R} \\ (x,y) &\longmapsto (1-x)^2 + 100\left(y - x^2\right) \end{aligned}$$

is not a real convex function. The mixed Hessian matrix components corresponding to the Rosenbrock's function $f : \mathbb{Z} \times \mathbb{R} \to \mathbb{R}$ are

$$\begin{aligned} \nabla_{11} f(x,y) &= (1-(x+2))^2 + 100\left(y-(x+2)^2\right) \\ &\quad -2\left(1-(x+1)\right)^2 + 100\left(y-(x+1)^2\right) \\ &\quad +(1-x)^2 + 100\left(y-x^2\right) \\ &= -300x^2 - 600x - 498 + 300y \end{aligned}$$

$$\begin{aligned} \frac{\partial}{\partial y}\left(\nabla_1 f(x,y)\right) &= \frac{\partial}{\partial y}[(1-(x+1))^2 + 100\left(y-(x+1)^2\right) \\ &\quad -\left[(1-x)^2 + 100\,(y-x)\right]) \\ &= 100 - 100 \\ &= 0 \end{aligned}$$

and

$$\frac{\partial^2 f}{\partial y^2} = 0.$$

Therefore the mixed Hessian matrix corresponding to the Rosenbrock function is not a positive definite matrix indicating Rosenbrock's function is not a mixed convex function.

If the Rosenbrock's function is considered with the change in the variables, $f : \mathbb{R} \times \mathbb{Z} \to \mathbb{R}$, then the corresponding mixed Hessian matrix components change; however, the mixed convexity condition remains same for this function:

$$\begin{aligned}
\nabla_{11} f(x, y) &= (1 - x)^2 + 100\left((y + 2) - x^2\right) \\
&\quad -2\left[(1 - x)^2 + 100\left(y + 1 - x^2\right)\right] \\
&\quad + (1 - x)^2 + 100\left(y - x^2\right) \\
&= 0 \\
\frac{\partial^2 f}{\partial x^2} &= -198 \\
\nabla_1 \left(\frac{\partial}{\partial x} f(x, y)\right) &= \nabla_1 \left[-2(1 - x) + 100(-2x)\right] \\
&= 0
\end{aligned}$$

Chapter 5

MIXED CONVEXITY AND OPTIMIZATION APPLICATIONS IN QUEUEING SYSTEMS

In this section we will consider the mixed convexity of the functions associated to the $M/E_k/1$ queueing systems suggested by Kumin (1973).

5.1 MIXED CONVEXITY OF AN $M/E_k/1$ QUEUEING SYSTEM

In this section we consider an automated machine that can perform up to k operations in series. Each step takes the same mean time, with the times distributed exponentially. If the arrival process is Poisson, we have an $M/E_k/1$ queueing system. Kumin (1973) considers the following design problem: Assume that there is a linear cost C_1 associated with each operation, a linear cost C_2 associated with each unit of service, and a linear cost C_3 associated with the length of the queue, L_q. Then the design problem is to minimize

$$\begin{aligned} f \quad &: \quad \mathbb{Z} \times \mathbb{R} \rightarrow \mathbb{R} \\ (k, \mu) \quad &\longmapsto \quad C_1 k + C_2 \mu + C_3 E(L_q) \end{aligned}$$

subject to $\mu > \lambda > 0$, for $k = 1, 2...$ on the neighborhood $U = \{\mu \mid \mu > \lambda > 0\}$ where

$$E(L_q) = (\frac{\lambda}{\mu})^2 \frac{(k+1)}{2k(1 - \frac{\lambda}{\mu})}$$

and $\rho = \dfrac{\lambda}{\mu}$.

In this problem, we have

$$\begin{aligned}
\nabla_{11} f(k,\mu) &= (C_1(k+1) + C_2\mu + C_3(\frac{(k+2)}{2(k+1)}\frac{\rho^2}{(1-\rho)}) \\
&\quad -(C_1 k + C_2\mu + C_3(\frac{(k+1)}{2k}\frac{\rho^2}{(1-\rho)}) \\
&= \frac{C_3\lambda^2 + 2C_1 k(1+k)(\lambda-\mu)\mu}{2k(k+1)(\lambda-\mu)\mu} \\
\frac{\partial f(k,\mu)}{\partial\mu} &= \frac{C_3(k+1)\lambda^2(\lambda-2\mu) + 2C_2 k(\lambda-\mu)^2\mu^2}{2k(\lambda-\mu)^2\mu^2}
\end{aligned}$$

The mixed matrix elements would be

$$\begin{aligned}
\nabla_{11} f(k,\mu) &= -\frac{C_3\lambda^2}{(2k+3k^2+k^3)(\lambda-\mu)\mu} \\
\frac{\partial(\nabla_1 f(k,\mu))}{\partial\mu} &= -\frac{C_3\lambda^2(\lambda-2\mu)}{2k(k+1)(\lambda-\mu)^2\mu^2} \\
\nabla_1(\frac{\partial f(k,\mu)}{\partial\mu}) &= -\frac{C_3\lambda^2(\lambda-2\mu)}{2k(k+1)(\lambda-\mu)^2\mu^2} \\
\frac{\partial^2 f(k,\mu)}{\partial\mu^2} &= -\frac{C_3(k+1)\lambda^2(\lambda^2-3\lambda\mu+3\mu^2)}{k(\lambda-\mu)^3\mu^3}
\end{aligned}$$

Therefore, the mixed matrix is then

$$H_f = \begin{bmatrix} -\frac{C_3\lambda^2}{(2k+3k^2+k^3)(\lambda-\mu)\mu} & -\frac{C_3\lambda^2(\lambda-2\mu)}{2k(k+1)(\lambda-\mu)^2\mu^2} \\ -\frac{C_3\lambda^2(\lambda-2\mu)}{2k(k+1)(\lambda-\mu)^2\mu^2} & -\frac{C_3(k+1)\lambda^2(\lambda^2-3\lambda\mu+3\mu^2)}{k(\lambda-\mu)^3\mu^3} \end{bmatrix}$$

The results above show that the first principal minor

$$\nabla_{11} f(k,\mu) = -\frac{C_3\lambda^2}{(2k+3k^2+k^3)(\lambda-\mu)\mu}$$

is strictly positive and the determinant of this mixed matrix, which is found to be

$$\frac{(C_3)^2\lambda^4((2+k(7+4k))\lambda^2 - 4(1+k(5+3k))\lambda\mu + 4(1+k(5+3k))\mu^2)}{4k^2(k+1)^2(k+2)(\lambda-\mu)^4\mu^4}$$

is also strictly positive. Therefore, the mixed matrix is positive definite for all $C_3 > 0$, and all $\mu \in U$ which implies that the function $f(k, \mu)$ is convex in μ and integer convex in k. In addition, this function is a 2-smooth strictly mixed convex function; hence, the determinant of the mixed Hessian matrix is strictly positive and $f(k, \mu)$ has a unique global minimum value by corollary 5.4.

Chapter 6

MIXED CONVEXITY APPLICATIONS IN TELECOMMUNICATION SYSTEMS

In telecommunication engineering, the quality of the service and the grade of service are the two measures used to specify the quality of voice service. Grade of Service (GoS) is the probability of a call in a circuit group being blocked or delayed for more than a specified interval. This is always with reference to the busy hour when the traffic intensity is the greatest in the network. One way of viewing GoS is to consider it independently from the perspective of outgoing versus incoming calls, and is not necessarily equal in each direction or between different source-destination pairs.

A. K. Erlang used a set of assumptions that relied on the network losing calls when all circuits in a group were busy to calculate the Grade of Service of a specified group of circuits or routes. These assumptions are (Flood (1998)):

- All call arrivals and terminations are independent random events
- The average number of calls does not change
- Every outlet from a switch is accessible from every inlet
- Any call that encounters congestion is immediately lost.

By using these assumptions Erlang developed the Erlang-B (or Erlang loss) formula which describes the probability of congestion in a circuit group. The probability of congestion gives the Grade of Service experienced.

Convexity analysis of Erlang delay and loss formulae are useful in the study of multi-server queueing systems. This can be done by finding simple and sharp bounds (see for example Ko, Serfozo and Sivakumar (2004) , Fuhrmann, Kogan and Milito (1996) and Harel (1990)) These bounds also enable one to obtain related convexity results; however, these convexity results are obtained by holding constant either the number of servers or the traffic intensity. We provide a technique for determining convexity without fixing either the service rate or the number of servers. Convexity results obtained previously are special cases of the results provided in this work.

6.1 ERLANG DELAY AND LOSS FORMULAE

Suppose μ is the service rate of the circuit to the incoming calls, s is the number of circuits, λ is the arrival rate of new calls to the circuit, $c = \frac{\lambda}{\mu}$ is the offered load of the circuit and $\rho = \frac{c}{s}$ is the expected traffic intensity in Erlangs. To determine the Grade of Service of a network when the traffic load and number of circuits are known, telecommunications network operators make use the Erlang loss formula given by (6.1) which allows operators to determine whether each of their circuit groups meet the required Grade of Service, simply by monitoring the reference traffic intensity (Flood (1998)).

$$E_l\left(s,\mu\right) = \frac{1}{\sum\limits_{j=0}^{s} \left(\frac{\mu}{\lambda}\right)^{s-j} \frac{s!}{j!}} \tag{6.1}$$

For delay networks, the Erlang delay (or C) formula, E_d, allows network operators to determine the probability of delay depending on peak traffic and the number of circuits

$$E_d\left(s,\mu\right) = \frac{1}{(s-1)!\sum\limits_{j=0}^{s-1} \frac{(s-j)}{j!} \left(\frac{\mu}{\lambda}\right)^{s-j}} \tag{6.2}$$

when $0 < c \leq s$ (see for example Harel (1990)).

Jagers and van Doorn (1991) provided the real extensions of the Erlang loss formula E_l, E_{lx}, and the Erlang delay formula E_d, E_{dx}, by the formulas

$$E_{lx} = \frac{1}{\int\limits_0^\infty \frac{c}{e^{ct}} (1+t)^x \, dt} \quad \text{and} \quad E_{dx} = \frac{1}{\int\limits_0^\infty \frac{c}{e^{ct}} t (1+t)^{x-1} \, dt}, \forall x \in \mathbb{R}^+$$

For traffic calculations in most telecommunication queueing systems, the mathematics is based on the assumption that call arrivals are random and Poissonian. In telecommunication network queueing systems, the convexity of the Erlang delay and loss formulae play important role in the minimization of the circuits and the traffic intensity. We provide strong convexity results for the Erlang B and C formulae with boundary conditions.

In queueing theory, the Erlang B formula can be used to show whether the $M/G/s/s$ queueing system is full or not. The steady-state probability of delay in the $M/M/s$ queueing system can be obtained from the Erlang C formula. Harel and Zipkin (1998), Harel ((1981), (1988), and (1988)) obtains sharp bounds for the Erlang delay and loss formulae. Considering the large offered load case and the number of servers, Newell (1984) obtains asymptotic approximations for the Erlang loss formula. Following Halfin and Whitt (1981), asymptotic results are developed by Janssen, van Leeuwaarden and Zwart (2008). Sharp heavy traffic approximations are obtained by Halfin and Whitt (1981). Simple inequalities for some of the performance measures of multi-server queues are obtained by Sobel (1980). A family of upper bounds which are sufficiently close to the Erlang loss formula is obtained by Adelman (2008). Harel (1990) develops convexity results for the Erlang delay and loss formulae. Lee and Cohen (1982) outline the proof of the convexity of the Erlang delay formula in the traffic intensity when the number of servers is held constant and Harel (1990) provides a short proof for this result. [For a comprehensive approach to the Erlang loss function see Jagerman (1974) and see Whitt (2002) for applications of the Erlang delay and loss formulae.]

6.2 CONDENSE DISCRETE CONVEXITY RESULTS FOR THE ERLANG LOSS FORMULA

Erlang (1917) proposed the Erlang loss formula

$$l(s,\mu) = \frac{\left(\frac{\lambda}{\mu}\right)^s}{s!\sum_{j=0}^{s}\frac{1}{j!}\left(\frac{\lambda}{\mu}\right)^j} \tag{6.3}$$

for the number of busy servers under the following $M/M/s$ queue system assumptions:

- The customer arrival follows a Poisson Process with arrival rate λ; independent from each other,
- Service rate follows an exponential distribution with mean $\frac{1}{\mu}$,
- If some of the customers find all the servers busy, they leave the system and have no effect on the system.

Let

$$e^{\mu}(s) := \sum_{j=0}^{s}\frac{c^j}{j!}, \tag{6.4}$$

$$\epsilon^{\mu}(s) := \frac{1}{s!}\left(\frac{\lambda}{\mu}\right)^s. \tag{6.5}$$

e_s^{μ} and ϵ_s^{μ} will be used instead of $e^{\mu}(s)$ and $\epsilon^{\mu}(s)$ respectively for simplicity. Erlang loss formula given in (6.3) can be expressed as a function of the number of servers for fixed service rate μ by the formula

$$l^{\mu}(s) = \frac{\epsilon_s^{\mu}}{e_s^{\mu}}.$$

Lemma 6.1: The Erlang loss formula, $l^{\mu}(s)$, given in (6.3) is a condense discrete convex function with respect to the number of servers when the service rate is assumed constant for all $s \geq 1$ and $0 < \lambda \leq s\mu$.

Proof: The second difference of $l^{\mu}(s)$ is

$$\begin{aligned}\nabla_{11} l^{\mu}(s) &= \frac{\epsilon_{s+2}^{\mu}}{e_{s+2}^{\mu}} - 2\frac{\epsilon_{s+1}^{\mu}}{e_{s+1}^{\mu}} + \frac{\epsilon_{s}^{\mu}}{e_{s}^{\mu}} \\ &= \frac{\lambda^{s+2}}{(s+2)!\mu^{s+2}e_{s+2}^{\mu}} - 2\frac{\lambda^{s+1}}{(s+1)!\mu^{s+1}e_{s+1}^{\mu}} + \frac{\lambda^{s}}{s!\mu^{s}e_{s}^{\mu}}\end{aligned}$$

Letting

$$k^{\mu}(s) = \frac{\lambda^{s+2}}{(s+2)!\mu^{s+2}e_{s}^{\mu}e_{s+1}^{\mu}e_{s+2}^{\mu}}$$

we have

$$\nabla_{11} l^{\mu}(s) = k^{\mu}(s)\left[e_{s}^{\mu}e_{s+1}^{\mu} - 2\frac{\mu}{\lambda}(s+2)\,e_{s}^{\mu}e_{s+2}^{\mu} + \frac{\mu^2}{\lambda^2}(s+2)(s+1)\,e_{s+2}^{\mu}e_{s+1}^{\mu}\right]$$

By basic algebra and using the inequality $s+1 \geq 2$, $\frac{(s+1)\mu}{\lambda} \geq 2\frac{\mu}{\lambda}$ holds, and in addition

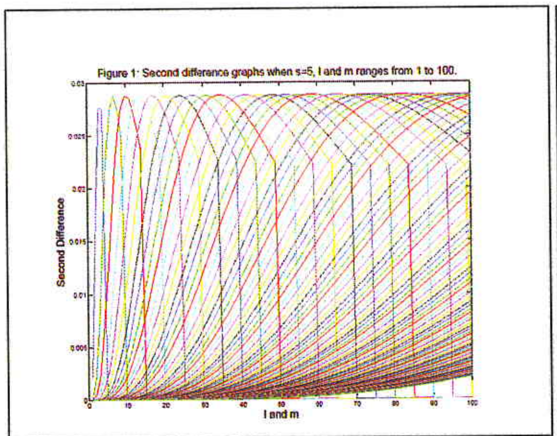
$$\frac{\mu}{\lambda}(s+1)\,e_{s+1}^{\mu} = \frac{\mu}{\lambda}(s+1)\left[e_{s}^{\mu} + \epsilon_{s+1}^{\mu}\right]$$

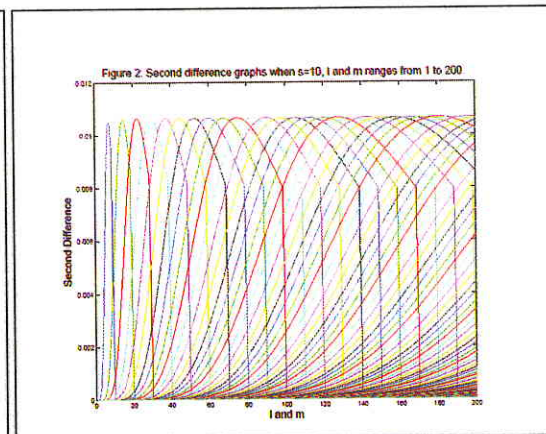
indicates

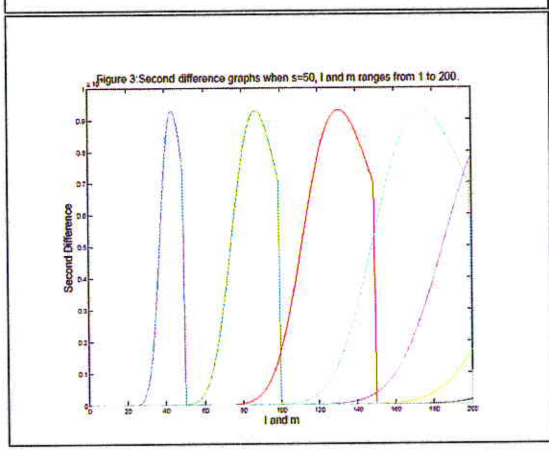
$$\begin{aligned}\frac{\mu}{\lambda}(s+1)\,e_{s}^{\mu} + \frac{\mu(s+1)\,\epsilon_{s+1}^{\mu}}{\lambda} &> 2e_{s}^{\mu} \\ \frac{\mu^2}{\lambda^2}(s+2)(s+1)\,e_{s+2}^{\mu}e_{s+1}^{\mu} &> 2\frac{\mu}{\lambda}(s+2)\,e_{s}^{\mu}e_{s+2}^{\mu}\end{aligned}$$

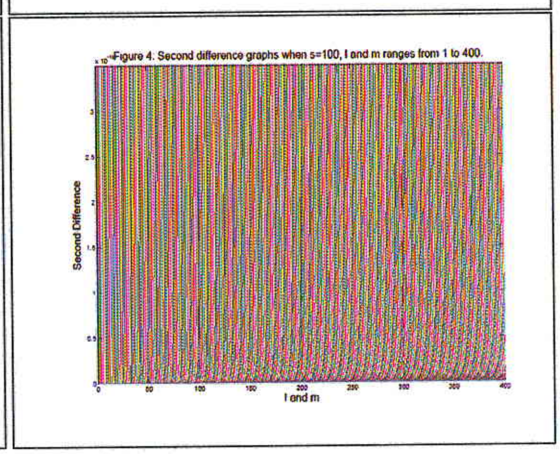
therefore $\nabla_{11} l^{\mu}(s) > 0$.

The following graphs illustrate special cases of lemma 6.1. We use the letters l to represent the arrival rate λ, and m to represent the service rate μ. Figure (1) illustrates the second difference graphs in the case when the number of servers is 5 and both the service and the arrival rates range from 1 to 100. Figures (2) illustrates the second difference graphs in the case when the number of servers is 10 and the service and the arrival rates range from 1 to 200. Figure (3) illustrates the second difference graphs in the case when the number of servers is 50 and the service and the arrival rates range from 1 to 200. Figure (4) illustrates the second difference graphs in the case when the number of servers is 100 and the service and the arrival rates range from 1 to 400.









6.3 CONDENSE DISCRETE CONVEXITY RESULTS FOR THE ERLANG DELAY FORMULA

In this section we consider an $M/M/s$ queue satisfying the conditions

- Unlimited capacity of waiting space in the queueing system,
- Constant arrival rate λ,
- Service rate is exponentially distributed,
- If a customer finds all the s servers busy then the customer joins the queue and waits until receiving service,

- No server can be idle if a customer is waiting.

The probability that all the servers busy in this $M/M/s$ queue, namely the Erlang delay formula, is defined to be the function

$$d(s,\mu) = \frac{1}{(s-1)!\sum_{i=0}^{s-1} \frac{s-i}{i!}\left(\frac{\mu}{\lambda}\right)^{s-i}} \tag{6.6}$$

when $s\mu > \lambda$ (Harel (1996)). By replacing the summation index i in the above equality with $i - s - 1$, the Erlang delay formula takes the form

$$d(s,\mu) = \frac{1}{(s-1)!\sum_{i=0}^{s-1} \frac{i+1}{(s-i)!}\left(\frac{\mu}{\lambda}\right)^{i}}.$$

In this section we will show that the Erlang delay formula is a condense discrete convex function with respect to the number of servers s when the service rate μ is assumed constant. This result will be used to show the mixed convexity of the Erlang delay formula.

Erlang delay formula given in (6.6) will be expressed as a function of the number of servers s for the constant service rate μ by either $d^{\mu}(s)$ or d_s^{μ} throughout this work.

Lemma 6.2: The Erlang delay function, $d^{\mu}(s)$, given in (6.6) is a condense discrete convex function of the number of servers, $s > 1$, for all $0 < \lambda \leq s\mu$ and fixed values of the service rate, μ.

Proof: Let

$$f^{\mu}(s) = (s-1)!\sum_{i=0}^{s-1} \frac{i+1}{(s-1-i)!}\left(\frac{\mu}{\lambda}\right)^{i+1} \tag{6.7}$$

$$h^{\mu}(s) = \frac{1}{f^{\mu}(s+2)\, f^{\mu}(s+1)\, f^{\mu}(s)} \tag{6.8}$$

By using the inequality $\lambda < s\mu$, we let $s\mu = \lambda + \delta$ for some $\delta > 0$.

Using (6.7) and (6.8), the second difference of the Erlang delay formula is

$$\nabla_{11} d^{\mu}(s) = d^{\mu}(s+2) - 2d^{\mu}(s+1) + d^{\mu}(s)$$
$$= h^{\mu}(s)\left[f^{\mu}(s+1) f^{\mu}(s) - 2f^{\mu}(s) f^{\mu}(s+2) + f^{\mu}(s+1) f^{\mu}(s+2)\right] \tag{6.9}$$

The last two terms on the left side of equation (6.9) are

$$-2f^{\mu}(s) f^{\mu}(s+2) = -2(s-1)!(s+1)!\left[\sum_{i=0}^{s-1} \frac{i+1}{(s-1-i)!}\left(\frac{\mu}{\lambda}\right)^{i+1}\right] \cdot \left[\sum_{i=0}^{s+1} \frac{i+1}{(s+1-i)!}\left(\frac{\mu}{\lambda}\right)^{i+1}\right],$$

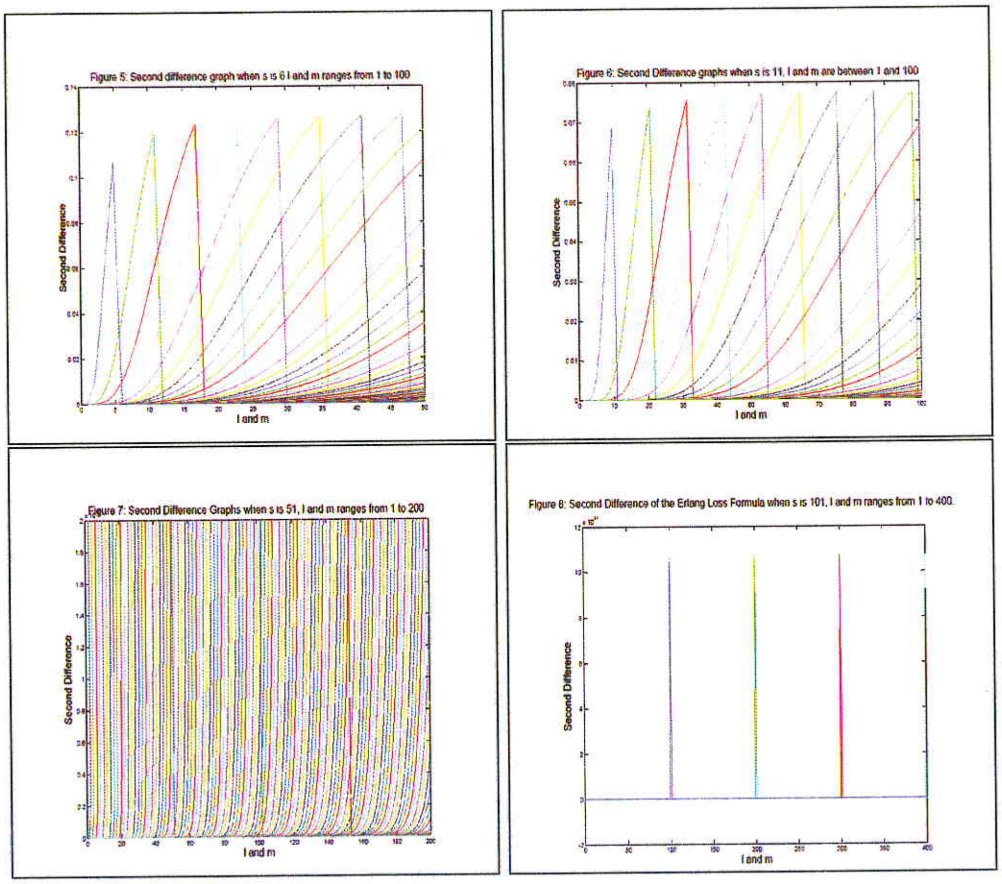
$$f^{\mu}(s+1) f^{\mu}(s+2) = (s+1)!(s+2)!\left[\sum_{i=0}^{s} \frac{i+1}{(s-i)!}\left(\frac{\mu}{\lambda}\right)^{i+1}\right]\left[\sum_{i=0}^{s+1} \frac{i+1}{(s+1-i)!}\left(\frac{\mu}{\lambda}\right)^{i+1}\right].$$

Therefore

$$\left[-2f^{\mu}(s) + f^{\mu}(s+1)\right] f^{\mu}(s+2) = \left[(s-1)!\right]^2 \left[s(s+1) \sum_{i=0}^{s+1} \frac{(i+1)}{(s+1-i)!}\left(\frac{\mu}{\lambda}\right)^{i+1}\right] \cdot \left[\sum_{i=0}^{s+1} (i+1)\left(\frac{\mu}{\lambda}\right)^{i+1}\left[\frac{-2}{(s-1-i)!} + \frac{(s+2)(s+1)s}{(s-i)!}\right]\right]$$

is positive since $(s+2)(s+1)s > 2(s-i)$ for all $0 \leq i \leq s+1$ which completes the proof.

The following graphs illustrate special cases of lemma 6.2. We use the letters l to represent the arrival rate λ, and m to represent the service rate μ. Figure (1) illustrates the second difference graphs in the case when the number of servers is 6 and both the service and the arrival rates range from 1 to 100. Figure (2) illustrates the second difference graphs in the case when the number of servers is 11 and the service and the arrival rates range from 1 to 100. Figure (3) illustrates the second difference graphs in the case when the number of servers is 51 and the service and the arrival rates range from 1 to 200. Figure (4) illustrates the second difference graphs in the case when the number of servers is 101 and the service and the arrival rates range from 1 to 400.



The illustrated graphs (1) – (4) agree with the general result obtained in lemma 6.2

6.4 COMPUTATIONAL MIXED CONVEXITY RESULTS FOR THE ERLANG LOSS FORMULA

In this section we will illustrate numerical mixed convexity results with several graphs in addition to the particular mixed convexity values. Let H_{E_l} be the mixed Hessian matrix corresponding to the Erlang Loss formula. Please see *Appendix A* for the determinant of the Hessian matrix corresponding to the Erlang Loss formula. Table 2 below displays several different values corresponding to the determinant of H_{E_l} for several different cases of s, μ, and λ when $s\mu = \lambda$. As it can be seen in the table all the particular cases indicate a positive definite Hessian matrix. In the case when the number of servers is 200, the service and the arrival rate change, Matlab $R2009b$ results in NaN (Not a number.)

Table 6.1 Determinant of H_{E_l} when $s\mu$ is equal to λ.

(μ, s)	$\lambda = 1$	(μ, s)	$\lambda = 10$	(μ, s)	$\lambda = 100$
$(0.1, 10)$	0.4955	$(1, 10)$	0.5×10^{-2}	$(10, 10)$	4.955×10^{-5}
$(0.02, 50)$	1.531	$(0.2, 50)$	1.53×10^{-2}	$(2, 50)$	1.5313×10^{-4}
$(0.01, 100)$	2.332	$(0.1, 100)$	2.33×10^{-2}	$(1, 100)$	2.332×10^{-4}
$(0.005, 200)$	NaN	$(0.05, 200)$	NaN	$(0.5, 200)$	NaN

In the case when $s\mu = \lambda$, the determinant of H_{E_l} for several different cases of s, μ, and λ are illustrated in Table 3. Similar to the case in Table 2 Matlab $R2009b$ program did not give a numerical result when the number of servers is 200, the service and the arrival rate change.

Table 6.2 Mixed Convexity Results for the Erlang Loss Formula when $s\mu$ is bigger than λ.

(μ, s)	$\lambda = 5$	(μ, s)	$\lambda = 50$	(μ, s)	$\lambda = 500$
$(0.6, 10)$	0.016	$(6, 10)$	1.598×10^{-4}	$(50.1, 10)$	1.982×10^{-6}
$(0.15, 50)$	3.1499×10^{-4}	$(1.1, 50)$	5.3323×10^{-4}	$(10.1, 50)$	6.1728×10^{-6}
$(0.06, 100)$	0.0171	$(0.6, 100)$	1.7077×10^{-4}	$(0.51, 100)$	9.467×10^{-4}
$(0.06, 150)$	1.2701×10^{-18}	$(0.6, 150)$	1.2701×10^{-20}	$(6, 150)$	1.2701×10^{-22}
$(0.03, 200)$	NaN	$(0.26, 200)$	NaN	$(2.6, 200)$	NaN

The following table (Table 4) illustrates the change in the determinant of the mixed Hessian matrix when the number of servers is 10, the service rate increments by $1 + \epsilon$ and the arrival rates considered are 1, 10, and 100.

Table 6.3 Mixed Convexity Results for the Erlang Loss Formula.

(μ, s)	$(1+\epsilon, 10)$ λ=1	$(1+\epsilon, 10)$ λ=10	$(10+\epsilon, 10)$ λ=100
$\epsilon = 100$	5.97×10^{-57}	4.906×10^{-37}	1.243×10^{-18}
$\epsilon = 10$	7.637×10^{-36}	1.243×10^{-16}	1.674×10^{-6}
$\epsilon = 1$	5.998×10^{-20}	1.674×10^{-4}	4.693×10^{-5}
$\epsilon = 0.1$	1.243×10^{-14}	0.0047	4.956×10^{-5}
$\epsilon = 0.01$	6.779×10^{-14}	0.005	4.9554×10^{-5}
$\epsilon = 0.001$	8.091×10^{-14}	0.005	4.955×10^{-5}

Table 5 indicates the change in the mixed convexity when the integer variable changes for the specified service and arrival rates.

Table 6.4 Mixed Convexity Results for the Erlang Loss Formula.

(μ, s)	$(0.1, 10+\epsilon)$ λ=1	$(1, 10+\epsilon)$ λ=10	$(10, 10+\epsilon)$ λ=100
$\epsilon = 1$	0.506	0.0051	5.0603×10^{-5}
$\epsilon = 2$	0.4524	0.0045	4.5240×10^{-5}
$\epsilon = 3$	0.3538	0.0035	3.5378×10^{-5}
$\epsilon = 4$	0.2411	0.0024	2.4107×10^{-5}
$\epsilon = 5$	0.1427	0.0014	1.4267×10^{-5}
$\epsilon = 6$	0.0733	7.3260×10^{-4}	7.3260×10^{-6}
$\epsilon = 7$	0.0327	3.2697×10^{-4}	3.2697×10^{-6}
$\epsilon = 8$	0.0127	1.2736×10^{-4}	1.2736×10^{-6}
$\epsilon = 9$	0.0044	4.3537×10^{-5}	4.3537×10^{-7}
$\epsilon = 10$	0.0013	1.3149×10^{-5}	1.3149×10^{-7}

The following graphs illustrate the determinant of the mixed Hessian matrix corresponding to the Erlang Loss formula. Figure (9) illustrates the determinant of the mixed Hessian matrix corresponding to the Erlang Loss formula when the number of the servers is 5, and the service and arrival rates range from 1 to 5 in the system. In this case it is easy to see that the mixed Hessian matrix is positive definite; therefore, Erlang Loss formula is a mixed convex function.

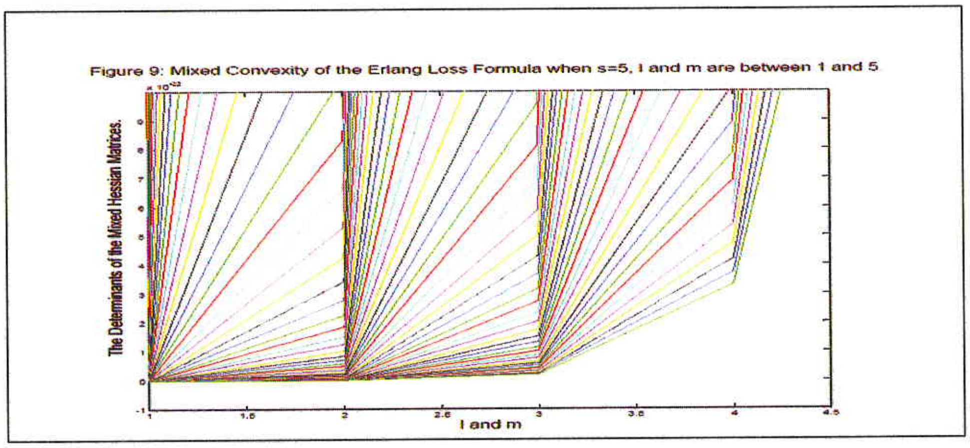
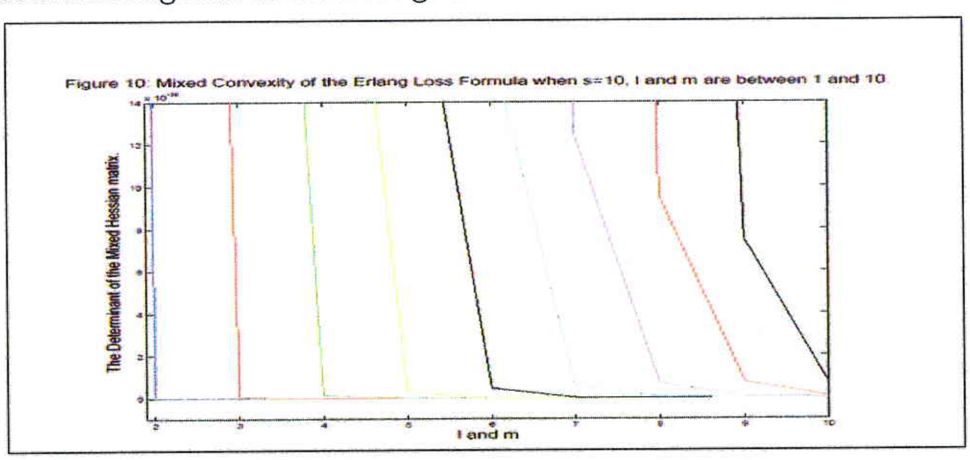
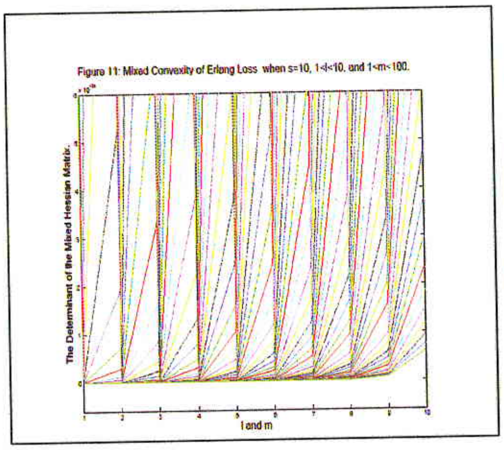


Figure (10) illustrates the determinant of the mixed Hessian matrix determinant corresponding to the Erlang Loss formula when the number of the servers in the systems is 10, and the service and arrival rates range from 1 to 10 in the system. In this case also it is easy to see that the mixed Hessian matrix is positive definite therefore Erlang Loss formula is again a mixed convex function.



Figures (11) and (12) illustrate the determinant of the mixed Hessian matrix for the Erlang Loss formula when the number of the servers and the arrival rate have the same conditions as in figure 10; however, the service rate ranges from 1 to 100 in figure (11) and ranges from 1 to 1,000 in fugure (12). In these two cases it is easy to see that the mixed Hessian matrix is positive definite therefore the Erlang Loss formula is again a mixed convex function.



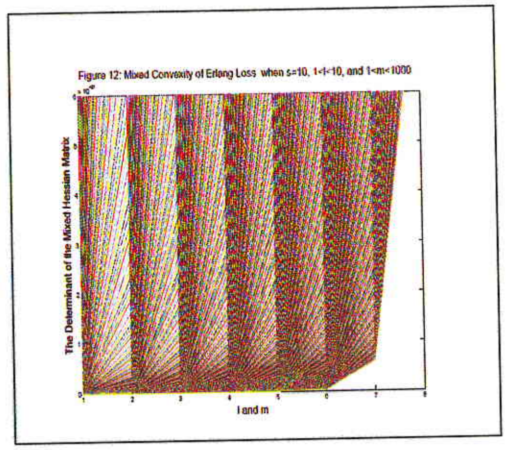


Figure (13) illustrates the determinant of the mixed Hessian matrix corresponding to the Erlang Loss formula when the number of the servers in the system is 50, the service rate increments from 0.1 to 1 by 0.1 and the arrival rates ranges from 1 to 10 in the system. The mixed convexity condition also holds in this case similar to the cases mentioned above.

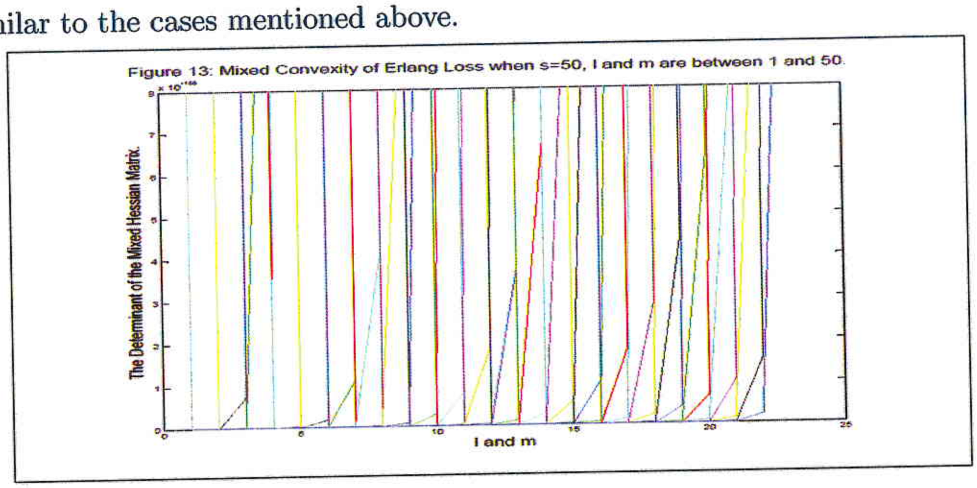
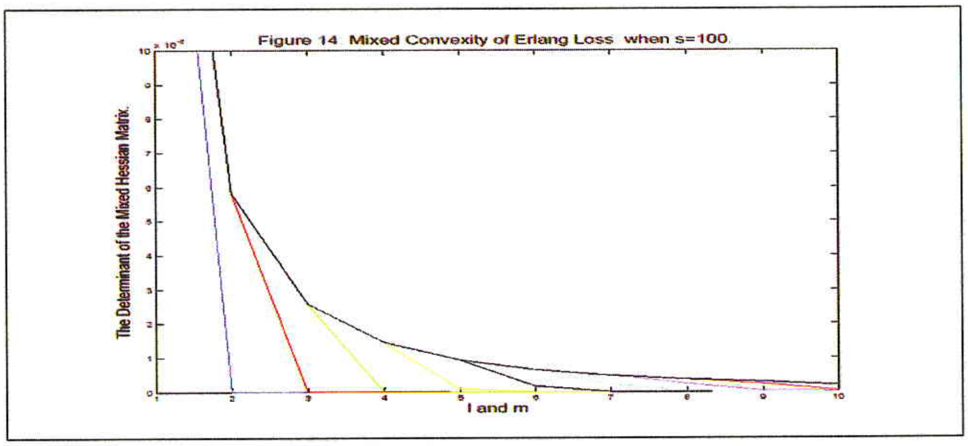


Figure (14) illustrates the determinant of the mixed Hessian matrix corresponding to the Erlang Loss formula when the number of the servers in the system is 100, the arrival rate λ ranges from 100 to 900, and the arrival rate μ ranges from 0.1λ to 10 with an increment of 1 in the system. The mixed convexity condition continues to hold in this case similar to the cases mentioned above.



Considering all the particular cases considered above, the mixed convexity holds for the Erlnag Loss formula.

6.5 COMPUTATIONAL MIXED CONVEXITY RESULTS FOR THE ERLANG DELAY FORMULA

In this section numerical mixed convexity results for several different special cases will be obtained. In addition some of the mixed convexity results corresponding to the Erlang Delay formula will be illustrated with graphs. Let H_{E_d} be the mixed Hessian matrix corresponding to the Erlang delay formula. Please see *Appendix B* for the determinant of the Hessian matrix corresponding to the Erlang Delay formula. Table 6 below displays several different values corresponding to the determinant of H_{E_d} for several different cases of s, μ, and λ when $s\mu = \lambda$. As it can be seen in the table all the particular cases indicate a positive definite Hessian matrix. In the case when the number of servers is 200, the service and the arrival rate change, Matlab $R2009b$ results in NaN (Not a number.)

Table 6.5 Determinant of H_{E_d} when $s\mu$ is equal to λ.

(μ, s)	$\lambda = 1$	(μ, s)	$\lambda = 10$	(μ, s)	$\lambda = 100$
$(\frac{1}{6}, 6)$	22.0148	$(\frac{10}{6}, 6)$	0.2201	$(\frac{100}{6}, 6)$	0.0022
$(\frac{1}{11}, 11)$	64.4792	$(\frac{10}{11}, 11)$	0.6448	$(\frac{100}{11}, 11)$	0.0064
$(\frac{1}{51}, 51)$	0.0196	$(\frac{10}{51}, 51)$	0.1961	$(\frac{100}{51}, 51)$	1.9608
$(\frac{1}{101}, 101)$	0.0099	$(\frac{10}{101}, 101)$	0.099	$(\frac{100}{101}, 101)$	0.9901

In the case when $s\mu = \lambda$, the determinant of H_{E_d} for several different cases of s, μ, and λ are illustrated in Table 7. Similar to the case in Table 6 Matlab 2009b program did not give a numerical result when the number of servers is 200, the service and the arrival rate change.

Table 6.6 Mixed Convexity Results for the Erlang Delay Formula.

(μ, s)	$\lambda = 6$	(μ, s)	$\lambda = 50$
$(\frac{7}{6}, 6)$	0.2576	$(\frac{51}{6}, 6)$	0.0085
$(\frac{7}{11}, 11)$	1.3159	$(\frac{51}{11}, 11)$	0.0227
$(\frac{7}{51}, 51)$	2.2282	$(1, 51)$	0.2624
$(\frac{7}{101}, 101)$	NaN	$(\frac{51}{101}, 101)$	NaN

(μ, s)	$\lambda = 500$	(μ, s)	$\lambda = 5000$
$(\frac{501}{6}, 6)$	8.7708×10^{-5}	$(\frac{5001}{6}, 6)$	8.8024×10^{-7}
$(\frac{501}{11}, 11)$	2.5463×10^{-4}	$(\frac{5001}{11}, 11)$	2.5759×10^{-6}
$(\frac{501}{51}, 51)$	0.0033	$(\frac{5001}{51}, 51)$	3.3190×10^{-5}
$(\frac{501}{101}, 101)$	NaN	$(\frac{5001}{101}, 101)$	NaN

The following table (Table 8) illustrates the change in the determinant of the mixed Hessian matrix when the number of servers is 10, the service rate increments by $1 + \epsilon$ and the arrival rates considered are $1,\ 10$, and 100.

Table 6.7 Mixed Convexity Results for the Erlang Delay Formula.

(μ, s)	$(\mu, 11)$ $\lambda=1$	$(\mu, 11)$ $\lambda=10$	$(\mu, 11)$ $\lambda=100$
$\mu = 101$	5.3314×10^{-63}	4.4675×10^{-41}	7.5981×10^{-20}
$\mu = 11$	5.8524×10^{-40}	1.1509×10^{-18}	0.0016
$\mu = 1.1$	1.1509×10^{-16}	0.1618	3.2407
$\mu = 1.01$	7.5981×10^{-16}	0.3133	3.8563
$\mu = 1.001$	9.2534×10^{-16}	0.3344	3.9271

Table 9 indicates the change in the mixed convexity when the integer variable changes for the specified service and arrival rates.

Table 6.8 Mixed Convexity Results for the Erlang Delay Formula.

(μ, s)	$(0.1, 10+\epsilon)$ $\lambda=1$	$(1, 10+\epsilon)$ $\lambda=10$	$(10, 10+\epsilon)$ $\lambda=100$
$\epsilon = 1$	33.681	0.3368	0.0034
$\epsilon = 2$	19.1326	0.1913	0.0019
$\epsilon = 3$	9.9449	0.0994	9.9449×10^{-4}
$\epsilon = 4$	4.7086	0.0471	4.7086×10^{-4}
$\epsilon = 5$	2.0349	0.0203	2.0349×10^{-4}
$\epsilon = 6$	0.8175	0.0082	8.1749×10^{-5}
$\epsilon = 7$	0.3196	0.0032	3.1961×10^{-5}
$\epsilon = 8$	0.1312	0.0013	1.3119×10^{-5}
$\epsilon = 9$	0.0617	6.1670×10^{-4}	6.1670×10^{-6}
$\epsilon = 10$	0.0352	3.5217×10^{-4}	3.5217×10^{-6}

The following graphs illustrate the determinant of the mixed Hessian matrix corresponding to the Erlang Delay formula. Figure (15) illustrates the determinant of the mixed Hessian matrix corresponding to the Erlang Delay formula when the number of the servers is 6, the service rate is between 1 and 10, and the arrival rate changes between 1 to 6 in the system. In this case it is easy to see that the mixed Hessian matrix is positive definite; therefore, Erlang Delay formula is a mixed convex function.

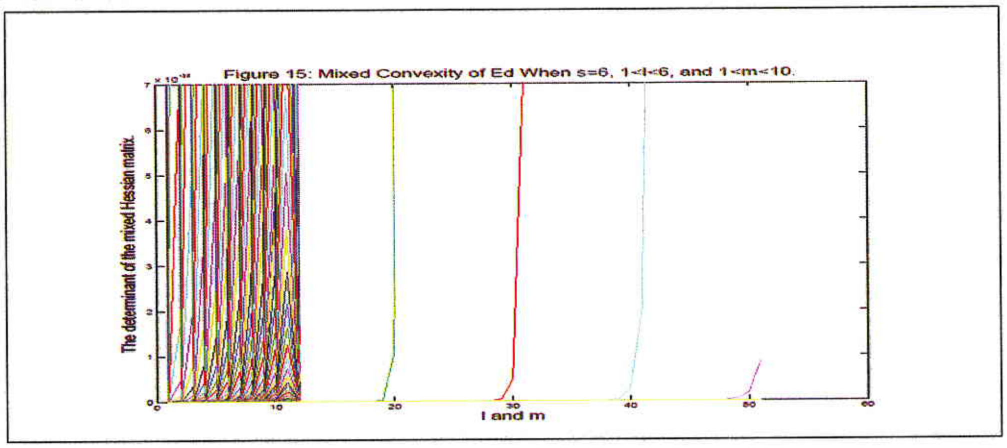
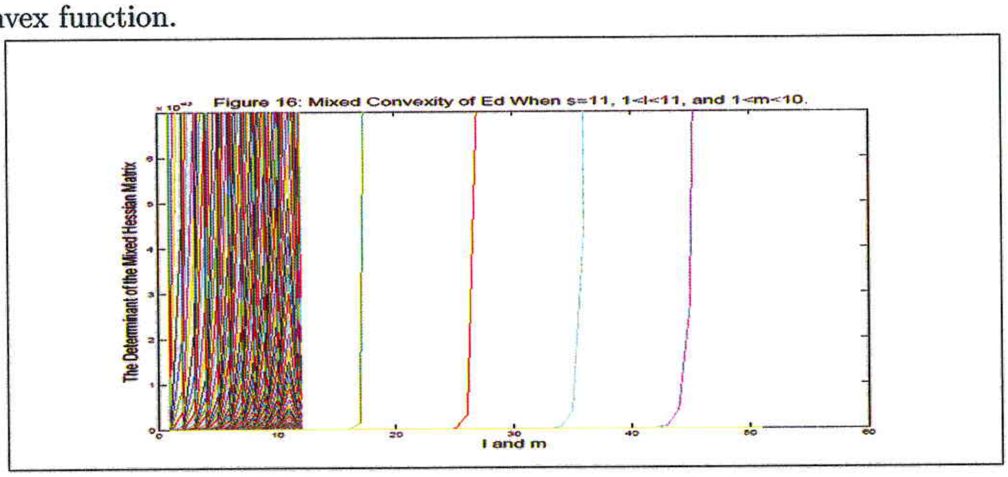
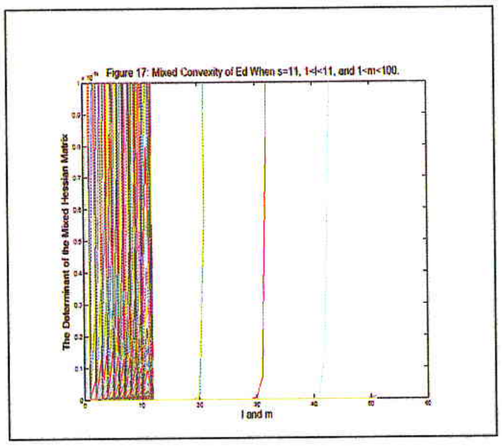


Figure (16) illustrates the determinant of the mixed Hessian matrix determinant corresponding to the Erlang Delay formula when the number of the servers in the systems is 11, and the service rate changes between 1 and 10, and the arrival rate ranges from 1 to 11 in the system. In this case also it is easy to see that the mixed Hessian matrix is positive definite therefore Erlang Delay formula is again a mixed convex function.



Figures (17) and (18) illustrate the determinant of the mixed Hessian matrix for the Erlang Delay formula when the number of the servers and the arrival rate have the same conditions as in figure 16; however, the service rate is between 1 and 100 in figure (17) and between 1 and 1,000 in fugure (18). In these two cases it is easy to see that the mixed Hessian matrix is positive definite therefore the Erlang Delay formula is again a mixed convex function.



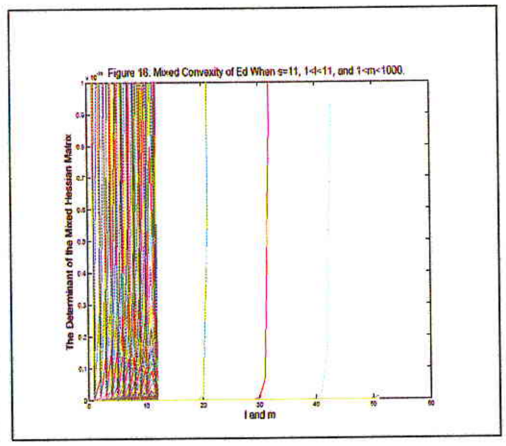


Figure (19) illustrates the determinant of the mixed Hessian matrix corresponding to the Erlang Delay formula when the number of the servers in the system is 50, the service rate increments from 0.1 to 1 by 0.1 and the arrival rates ranges from 1 to 10 in the system. The mixed convexity condition also holds in this case similar to the cases mentioned above.

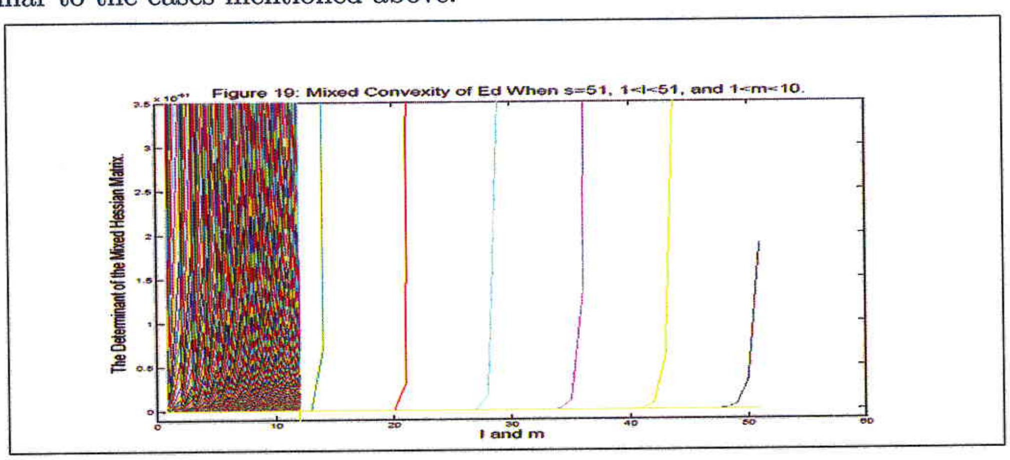
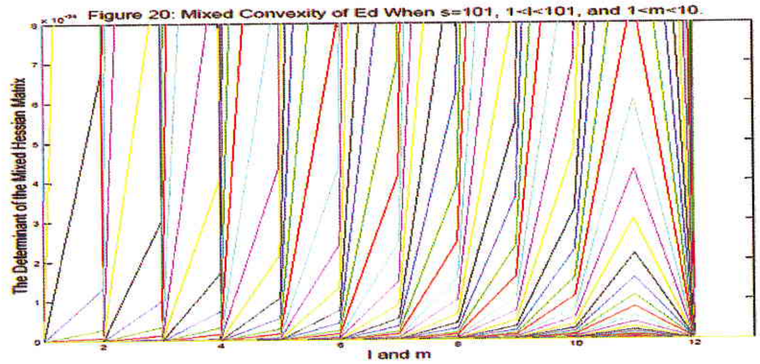


Figure (20) illustrates the determinant of the mixed Hessian matrix corresponding to the Erlang Delay formula when the number of the servers in the system is 101, the arrival rate λ ranges from 1 to 101, and the arrival rate μ ranges from 1 to 10 in the system. The mixed convexity condition continues to hold in this case similar to the cases mentioned above.

Figure 20: Mixed Convexity of Ed When s=101, 1<l<101, and 1<m<10.



Considering all the particular cases considered above, the mixed convexity holds for the Erlnag Delay formula.

Chapter 7

MIXED CONVEXITY AND OPTIMIZATION RESULTS FOR INVENTORY SCIENCE

The assumptions for the existing models in the literature were either completely backlogged or lost demand prior to Das's $(S-1, S)$ inventory model with partial backlogging (Das, 1977). For instance, Feeney-Sherbrooke (1966) analyzed the backlogging and lost-sale cases for any distribution of replenishment time where the demand is assumed to be compound Poisson. Gross-Harris (1973) allowed the replenishment time to depend on the level of unfilled demands when Poisson demand is considered to analyze backlogging. Hadley-Whitin (1963) considered any distribution of replenishment time and Poisson demand to analyze the lost-sale cases and backlogging. Galliher-Morse-Simond (1959) assumed the replenishment time to be either exponentially distributed or constant to analyze the backlogging case under the stuttering Poisson demand assumption. Among all these models that are suitable for recoverable items, the $(S-1, S)$ inventory model is appropriate for items when the demand is low but the item is expensive so that the cost of ordering in negligible compared with the costs of holding and shortage. Das's $(S-1, S)$ inventory model has a realistic demand condition in which the customers reaction to stockout is to wait a certain amount of time before cancelling their orders and it is also applicable to non-service systems with partial backlogging, such as production systems where the input material need not be processed immediately but processing must be done before a specific fixed amount of time (Das, 1977).

7.1 DAS'S INVENTORY MODEL

Following Das, let $N \in \mathbb{Z}^+$ be the initial number of items in stock, λ be the arrival rate of the demand per unit time (constant), ρ be the traffic intensity, $0 < \rho = \frac{\lambda}{\mu} < 1$, μ be the service rate, $K(N, \rho)$ be the expected total cost per unit time of operating the system, $L(\rho) = \frac{25}{\rho^2}$ be the arbitrarily chosen cost per unit time of receiving replenishment at the rate μ , C_h be the holding cost per unit per unit time, C_d be the cost per unit time of keeping an order waiting, C_l to be the penalty cost (in addition to TC_d) per lost order, $\alpha = \mu T$, and the tolerable delay for each customer be T units away from his arrival time. Then the total cost per unit time for the $(S-1, S)$ model has the closed function form

$$\begin{aligned} M(N,\rho) &= K(N,\rho) + L(\rho) \\ &= C_h \frac{1}{1 - e^{-\alpha(1-\rho)}\rho^{N+1}} \{N - \frac{\rho - \rho^{N+1}}{1-\rho}\} \\ &\quad + \frac{C_d}{1-\rho}\left[1 - e^{-\alpha(1-\rho)} + \alpha\rho(1-\rho)e^{-\alpha(1-\rho)}\right]\rho^{N+1} \\ &\quad + C_l\lambda(1-\rho)e^{-\alpha(1-\rho)}\rho^N + L(\rho). \end{aligned} \tag{7.1}$$

Analyzing the convexity and optimization of a function with unbounded domain $D \subseteq \mathbb{Z}^n \times \mathbb{R}^m$ for $n, m \geq 1$ can be difficult (Kumin 1973). Das (1977) emphasizes the difficulty of minimizing the cost function $M : D \to \mathbb{R}$, $D \subset \mathbb{Z}^+ \times \mathbb{R}^+$, with respect to the initial number of items in the stock and the service rate by applying the previously known methods as follows: "Sufficient conditions for the minimum of a function of mixed variables such as M have been discussed recently by Kumin in 1973. Unfortunately, these conditions are not easy to verify for the present case." Noting that the cost per unit time of receiving replenishment at the rate μ, $L(\rho) = \frac{25}{\rho^2}$, is a convex function of the real variable μ, he approximates the minimal value of the cost function by computing the values of the approximation function $M^*(\rho) = K^*(\rho) + L(\rho)$ for each ρ, $0 < \rho < 1$, where $K^*(\rho) = K(N^*(\rho), \rho)$ and $N^*(\rho)$ is the optimal initial stock for a given ρ. It is also shown that for all test problems $K^*(\rho)$ is found to be a strict convex function and so is $L(\rho)$, therefore $M^*(\rho)$ is a strict convex function. This is a useful technique to find the optimal ρ that minimizes $M^*(\rho)$ for $0 < \rho < 1$ for which any line-search can be applied; however a closed form solution to the suggested minimization problem cannot be obtained by this method as it requires calculations for all ρ, $0 < \rho < 1$. In this case finding an optimal solution to the model $M^*(N, \rho)$ implies finding a particular optimal solution for the original model $M(N, \rho)$.

The goal of in the following sections is to carry out a study to examine computational optimization results and obtain generalized mixed convexity results for the cost function M given in (7.1) by applying Theorem 4.1. These generalized mixed convexity results improve the results of Das and determine necessary and sufficient mixed convexity conditions for the cost function associated with the $(S-1, S)$ inventory model he suggested. The mixed convexity conditions obtained for the cost function M form a subset of the domain D, the convexity region containing possible minimal values of the cost function. The discrete convexity condition $(\nabla_{11}M > 0)$ of the cost function M which needs to be satisfied to carry out the mixed convexity results agrees with the discrete convexity results of Das. Some of the numerical results obtained in Section 7.2 are employed as counter examples for the mixed convexity of the cost function M. Considering the numerical data used by Das and the mixed convexity results obtained in Section 3, we find minimization results when the parameters and constant unknowns vary. In the last section we conclude that the cost function $M : D \to \mathbb{R}$ is not a mixed convex function over the entire space $(\mathbb{Z}^+)^5 \times (\mathbb{R}^+)^2$ with the real variables $(T, \mu) \in (\mathbb{R}^+)^2$ and integer variables λ, N, C_h, C_d, $C_l \in (\mathbb{Z}^+)^5$ under the constraint $\lambda < \mu$, but a region of mixed convexity can be determined by using Theorem 4.1. Note that even though naturally N is the only positive integer variable which represents the initial stock level, we consider the variables λ, C_h, C_d, and C_l to be integer as well since the numerical results presented in this work are for particular integer values of λ, C_h, C_d, and C_l. The mixed convexity conditions of the cost function M are not satisfied for $(N, \mu) \in \mathbb{Z}^+ \times \mathbb{R}^+$ with the particular integer values of λ, C_h, C_d, and C_l; therefore, the mixed convexity conditions cannot be satisfied for real values of λ, C_h, C_d, and C_l.

7.2 COMPUTATIONAL CONVEXITY RESULTS

A generalized closed-form mixed convexity result is obtained in chapter 5 to solve a mixed convexity problem raised by Kumin (1973). Das (1977) encountered a similar convexity problem to that of Kumin (1973) where the mixed convexity of the mixed variable cost function given in (7.1) was the concern. By using the equalities $\rho = \frac{\lambda}{\mu}$ and $\alpha = \mu T$, $M(N, \rho)$ given in (7.1) can be rewritten as

$$\begin{aligned} M(N,\mu) &= C_h \frac{1}{1-\left(e^{(-\mu+\lambda)T}\right)\left(\frac{\lambda}{\mu}\right)^{N+1}} \left(N - \frac{\frac{\lambda}{\mu} - \left(\frac{\lambda}{\mu}\right)^{N+1}}{1-\frac{\lambda}{\mu}}\right) \\ &\quad + \frac{C_d}{1-\frac{\lambda}{\mu}} \left[1 - e^{(-\mu+\lambda)T} + T\lambda\left(1-\frac{\lambda}{\mu}\right) e^{(-\mu+\lambda)T}\right] \left(\frac{\lambda}{\mu}\right)^{N+1} \\ &\quad + C_l\lambda\left(\left(\frac{\lambda}{\mu}\right)^{N} - \left(\frac{\lambda}{\mu}\right)^{N+1}\right) e^{(-\mu+\lambda)T} + \frac{25\mu^2}{\lambda^2}. \end{aligned}$$

In this section we obtain numerical mixed convexity results for $M(N, \mu)$ and provide relevant graphs based on the data used by Das. It is a difficult task to compute the elements and the determinant of the mixed Hessian matrix H corresponding to the cost function M; therefore, we provide the following (symbolic programming) algorithm.

Algorithm 7.1:

`Introduce symbols` $T, \lambda, \mu, N, C_l, C_h, C_d$

`Define` $M(N,\mu) = C_h \frac{1}{1-\left(e^{(-\mu+\lambda)T}\right)\left(\frac{\lambda}{\mu}\right)^{N+1}} \left(N - \frac{\frac{\lambda}{\mu}-\left(\frac{\lambda}{\mu}\right)^{N+1}}{1-\frac{\lambda}{\mu}}\right)$

$+\frac{C_d}{1-\frac{\lambda}{\mu}}\left[1-e^{(-\mu+\lambda)T}+T\lambda\left(1-\frac{\lambda}{\mu}\right)e^{(-\mu+\lambda)T}\right]\left(\frac{\lambda}{\mu}\right)^{N+1}+\frac{C_l\lambda\left(\left(\frac{\lambda}{\mu}\right)^N-\left(\frac{\lambda}{\mu}\right)^{N+1}\right)}{e^{(\mu-\lambda)T}}+\frac{25\mu^2}{\lambda^2}.$

Calculate $\nabla_1(M(N,\mu))=M(N+1,\mu)-M(N,\mu)$.

Calculate $\nabla_{11}(M(N,\mu))=M(N+2,\mu)-2M(N+1,\mu)+M(N,\mu)$.

Calculate $\frac{d}{d\mu}(\nabla_1(M(N,\mu)))$ //Differentiate $M(N+1,\mu)-M(N,\mu)$ with respect to μ.

Calculate $\det(H)=\nabla_{11}(M(N,\mu))\frac{d^2}{d\mu^2}(M(N,\mu))-\left(\frac{d}{d\mu}(\nabla_1 M(N,\mu))\right)^2$.

Considering $D\subset\mathbb{Z}^+\times\mathbb{R}^+$,

If $(\nabla_{11}(M)>0)$ {

M is strictly discrete convex with respect to its integer variable;

if $(\det(H)>0)$

M is strictly mixed convex with respect to its mixed variables;

end }

M is not strictly mixed convex.

Note that the differential of the first difference is the difference of the first differential by the symmetry of the mixed Hessian matrix. By applying this algorithm, closed-form mixed convexity conditions for $M(N,\mu)$ are obtained symbolically. In particular, numerical mixed convexity results for the cost function $M(N,\mu)$ associated with the $(S-1,S)$ inventory model under the time limit on backorders suggested by Das can be obtained by modifying this algorithm. These convexity results depend on the fixed variables T, λ, C_l, C_h and C_d where we choose $L(\rho)=\frac{25}{\rho^2}$. Considering the initial number of items in the stock and the service rate to be the only parameters of the cost function M, the mixed convexity conditions of the cost function consist of the integer convexity of M (i.e., $\nabla_{11}M(N,\mu)>0$) and the positive determinant of the mixed Hessian matrix corresponding to M (i.e., $\det(H)>0$). By using these two conditions, we specify the mixed convexity region to find the minimal values of the cost function in the domain and therefore generalize the results of Das. To specify the region for which we can find optimal values, we first check whether or not the cost function has a mixed convex structure in its domain.

In this section, we illustrate the graphs of the second difference of the cost function with respect to the initial number of items in the stock, the second derivative of the cost function with respect to the service rate and the determinant of the mixed Hessian matrix H that corresponds to the cost function $M(N, \mu)$, in addition to the related numerical results.

7.2.1 CHANGE IN THE ARRIVAL RATE (λ)

In this section we observe the effect of the change in arrival rate of the demands per unit time (λ) on the mixed convexity of the cost function $M(N, \mu)$ generated for the $(S-1, S)$ model. The following graphs assume that the arrival rate is either 1, 10 or 100, the tolerable delay for each customer is $T = 0.5$ units away from his arrival time, the holding cost per unit per unit time is 6, the cost per unit time of keeping an order waiting is 4, and the penalty cost per lost order is 35 when the initial number of items range between 1 and 40, and $\lambda + 1 \leq \mu = m \leq 40$. From this point on, use μ and m interchangeably.

Case 1: We first consider $T = 0.5$, $\lambda = 1$, $C_h = 6$, $C_d = 4$ and $C_l = 35$. In this case, using the basic assumption $0 < \rho = \frac{\lambda}{\mu} < 1$, figures 1-3 below are obtained for the second derivative of M, the second difference of M and the determinant for the mixed Hessian matrix H. These figures and the numerical data indicate the mixed convexity of the cost function for the $(S-1, S)$ model when the arrival rate is 1, the tolerable delay of the customer is 0.5 units away from his arrival time, holding cost per unit per unit time is 6, cost per unit time of keeping an order waiting is 4, the penalty cost per lost orders 35, and the initial number of items in the stock varies between 1 and 40 while the service rate ranges between 2 and 40.

Fig 1. Second difference of M when lambda is 1.

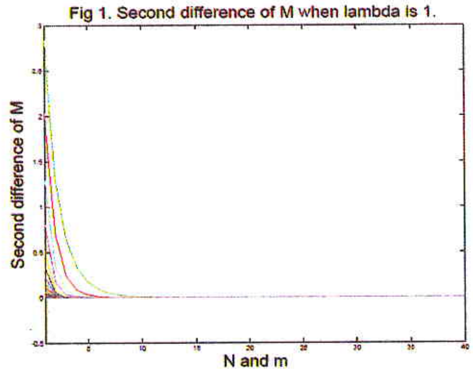


Fig 2. Second derivative of M when lambda is 1.

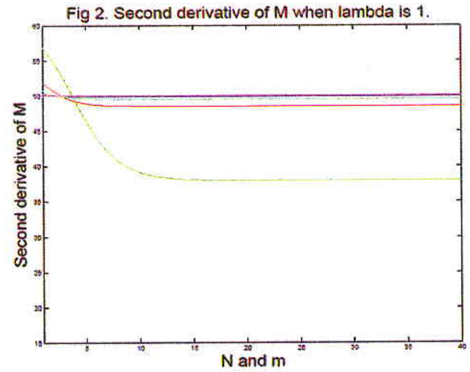
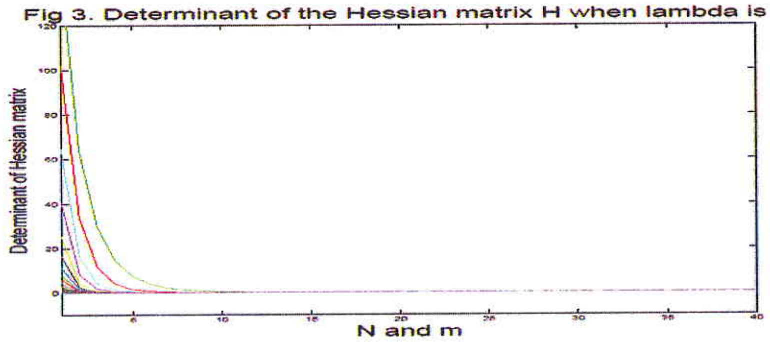


Fig 3. Determinant of the Hessian matrix H when lambda is 1



Case 2: Similar to the first case, we consider the case where $T = 0.5$, $\lambda = 10$, $C_h = 6$, $C_d = 4$, $C_l = 35$ and $\rho = \frac{\lambda}{\mu} < 1$; however, this time we consider the values $1 \leq N \leq 40$ and $11 \leq m \leq 40$. The graphs of the determinant of H, the second derivative of M and the second difference of M indicate that there exist negative values of the second derivative of M and the determinant of the mixed Hessian matrix H, therefore the mixed convexity condition of M is not satisfied on a bounded domain. Hence the cost function M does not have a mixed convex structure when the arrival rate is 10, the tolerable delay of the customer is 0.5 units away from his arrival time, holding cost per unit per unit time is 6, cost per unit time of keeping an order waiting is 4, the penalty cost per lost orders is 35, the initial number of

items in the stock varies between 1 and 40 while the service rate ranges between 11 and 40.

Fig 4. Second difference of M when lambda is 10.

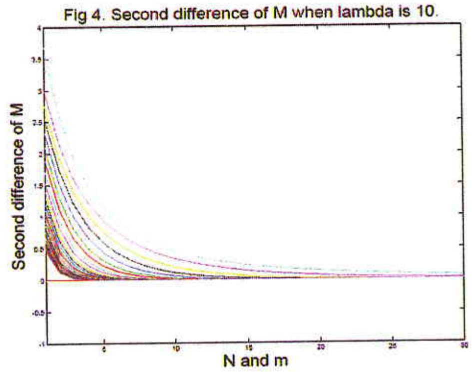


Fig 5. Second derivative of M when lambda is 10.

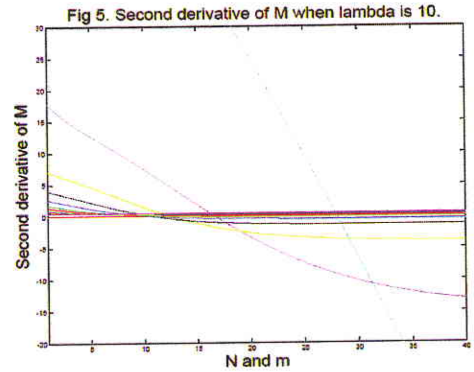
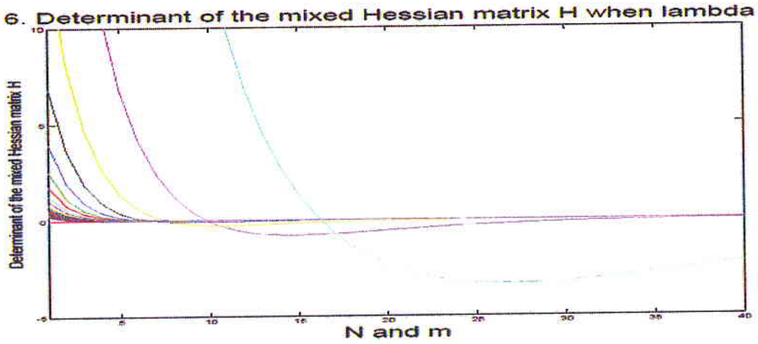
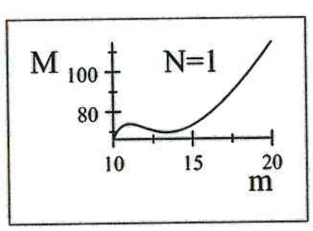
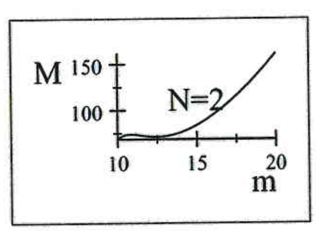


Fig 6. Determinant of the mixed Hessian matrix H when lambda is 10.



For the quantities considered in this case, the following graphs serve as counterexamples to the global mixed convexity (i.e., the mixed convexity in the entire domain) of the cost function M.





Case 3: Similar to the first two cases, we assume that the tolerable delay of the customers to be 0.5 units away from his arrival time, the arrival rate to be

100, the holding cost per unit per unit time to be 6, the cost per unit time of keeping an order waiting to be 4, the penalty cost per lost orders to be 35, the initial number of items in the stock to vary between 1 and 40 while the service rate changes between 101 and 200. Some of the values of the determinant of H and the second derivative of M are negative which violate the definition of mixed convexity. The integer convexity condition of the cost function for each fixed arrival rate holds based on the the second difference graph of the cost function M . This indicates that the cost function corresponding to the $(S-1, S)$ inventory model under the time limit on backorders suggested by Das (1977) is not necessarily a mixed convex function for the quantities $T = 0.5$, $\lambda = 100$, $C_h = 6$, $C_d = 4$ and $C_l = 35$, when $1 \leq N \leq 40$ and $101 \leq m \leq 200$.

Fig 7. Second difference of M when lambda is 100.

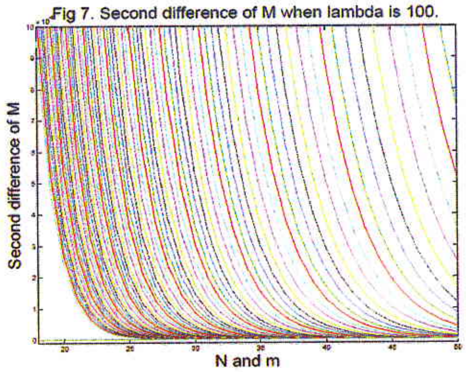


Fig 8. Second derivative of M when lambda is 100.

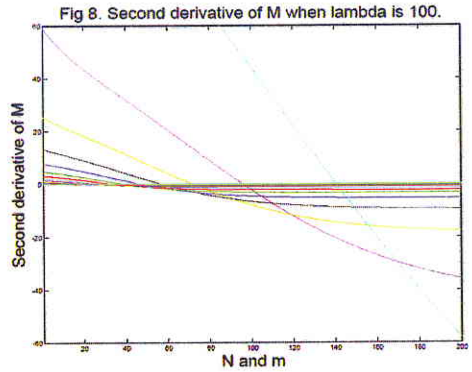
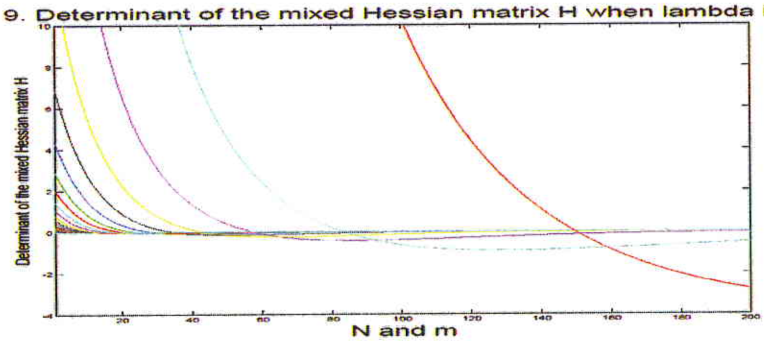


Fig 9. Determinant of the mixed Hessian matrix H when lambda is 100.



7.2.2 CHANGE IN THE HOLDING COST (C_h)

In this section, we observe the effect of the change of holding cost per unit per unit time (C_h) on the mixed convexity of the cost function M. The following figures are illustrated for the mixed convexity conditions of the cost function when the arrival rate is 5, the tolerable delay of the customer is 0.5 units away from his arrival time, the cost per unit time of keeping an order waiting is 4, the penalty cost per lost orders is 35 when the holding cost per unit per unit time ranges between 1 and 10, and the initial number of items in the stock ranges between 1 and 40.

Fig 10. Second difference of M when lambda is 5.

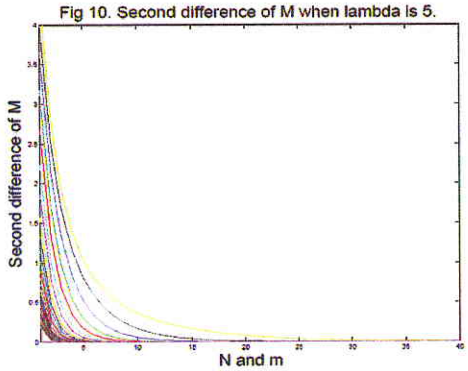


Fig 11. Second derivative of M when lambda is 5.

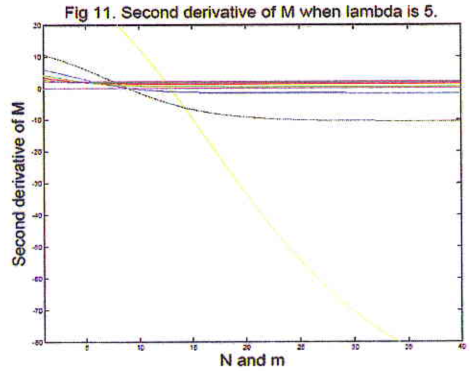
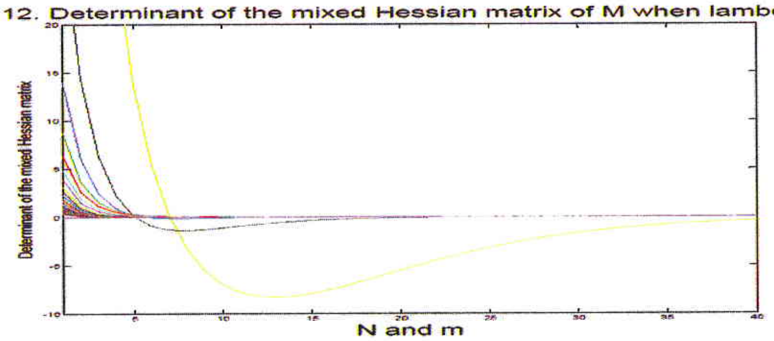


Fig 12. Determinant of the mixed Hessian matrix of M when lambda is 5.



The numerical data indicate that the second difference of M is strictly positive as was observed by Das; however the second derivative of M and the determinant of the mixed Hessian matrix H have negative values. Therefore, the cost function M

is not a strict mixed convex function when C_h ranges between 1 and 10 for the fixed parameter values considered in this case.

7.2.3 CHANGE IN THE COST OF KEEPING AN ORDER WAITING (C_d)

In this section, we observe the effect of C_d, the change in the cost of keeping an order waiting, on the mixed convexity conditions of the cost function.

The numerical data indicate that $\nabla_{11}M > 0$, as was observed by Das; however, $\frac{dM}{dm}$ and $\det(H)$ have negative values when the arrival rate of the demands is 5, the tolerable delay of the customer is 0.5 units away from his arrival time, the holding cost per unit per unit time is 6, and the penalty cost per lost orders is 35 when $1 \le C_d \le 10$, $1 \le N \le 40$ and $6 \le m \le 40$. In this case, we verified the integer convexity results of Das (1977) by obtaining $\nabla_{11}M > 0$ for various values of the cost of keeping an order waiting; however, the mixed convexity condition, $\det(H) > 0$, of the cost function does not hold.

Fig 13. Second difference of M when lambda is 5.

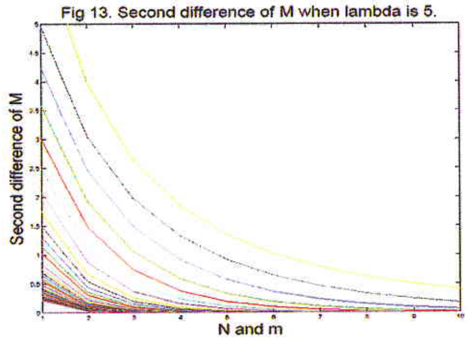


Fig 14. Second derivative of M when lambda is 5.

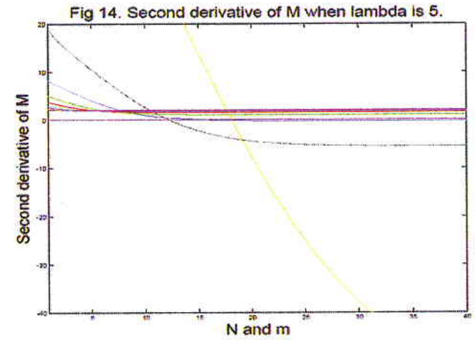
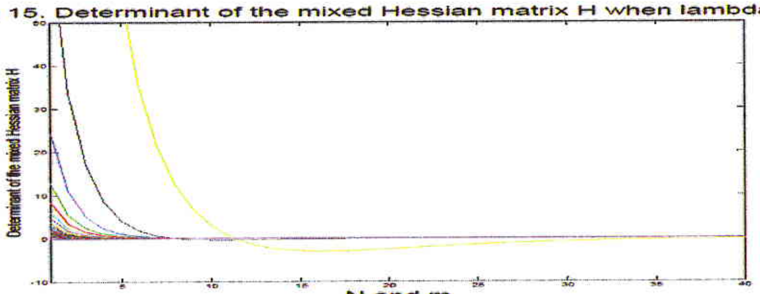


Fig 15. Determinant of the mixed Hessian matrix H when lambda is 5.



7.2.4 CHANGE IN THE PENALTY COST PER LOST ORDER (C_l)

The effect of the penalty cost per order (C_l) on the strict mixed convexity of the cost function under the parameter assumptions $\lambda = 5$, $T = 0.5$, $C_d = 4$ and $C_h = 6$ when $30 \leq C_l \leq 40$, $1 \leq N \leq 40$ and $6 \leq m \leq 40$ is similar to the effect of C_d and C_h on the mixed convexity of the cost function M. That is; The second difference of the cost function M is strictly positive, whereas the determinant of the mixed Hessian matrix is negative for some values of the parameters as illustrated in the following graphs.

Fig 16. Second difference of M when lambda is 5.

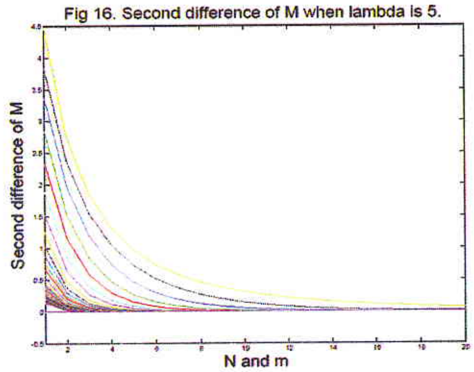


Fig 17. Second derivative graph of M when lambda is 5.

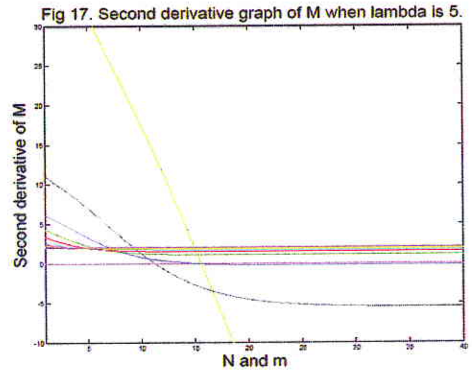
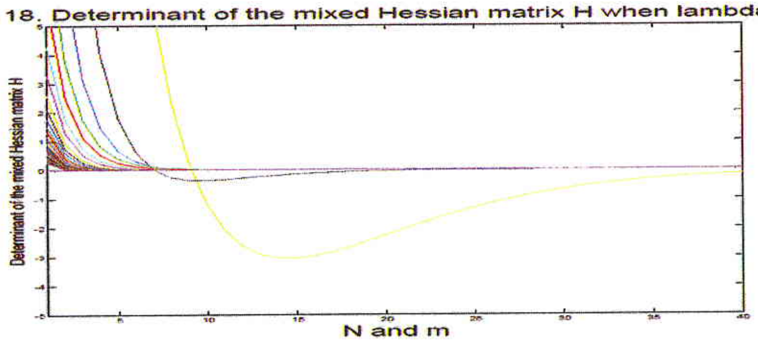
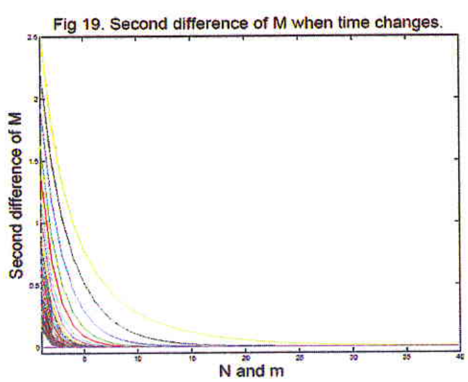


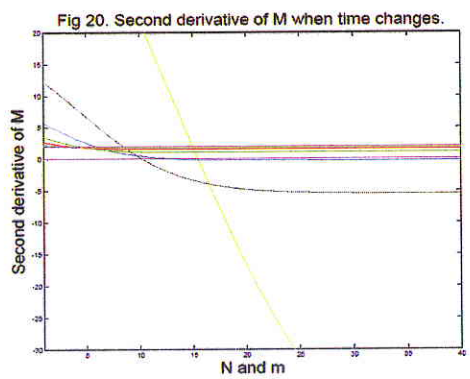
Fig 18. Determinant of the mixed Hessian matrix H when lambda is 5.

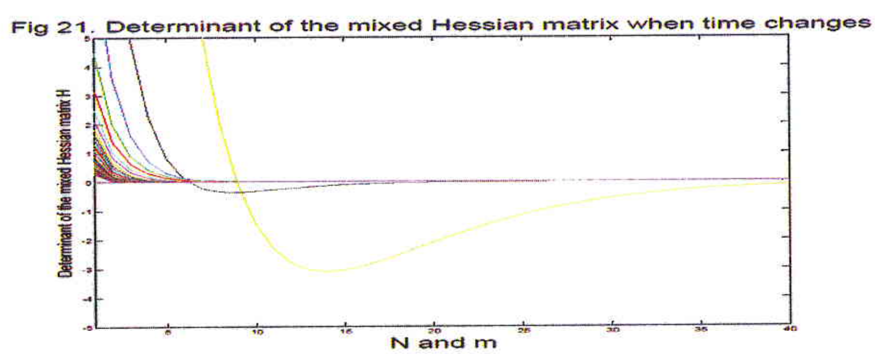


7.2.5 CHANGE IN THE TOLERABLE DELAY (T)

The change in the tolerable delay for each customer from his arrival time is an important factor in the mixed convexity of the cost function. Assume that the tolerable delay for each customer ranges between 0.1 and 0.9, the arrival rate is 5, the holding cost per unit per unit time is 6, the cost per unit time of keeping an order waiting is 4, the penalty cost per lost order is 35, the initial number of items range between 1 and 40, and the service rate ranges between 6 and 40. Under these assumptions, the numerical data indicate that the cost function is not a mixed convex function, but is an integer convex function. The following graphs illustrate the numerical data for this case.







In the cases considered above, the numerical data indicate that the cost function is not a strict mixed convex function for some T, λ, C_h, C_d and C_l where $L(\rho) = \frac{25}{\rho^2}$.

7.3 COMPUTATIONAL OPTIMIZATION RESULTS

Computational results for the optimization of mixed variable functions are necessary because of the complex nature of the problems encountered (see for example, Benders (2005) , Gümüş and Floudas (2005) , and Al-Yakoob, Sherali and Al-Jazzaf (2010) .) In the previous section, counter examples to the global mixed convexity of the cost function considered by Das are given for some μ and N. Considering the mixed convexity conditions of the cost function M obtained by applying Theorem 4.1, a subdomain of the domain of the cost function can be found to find the minimum values of the cost function. In this section, several computational optimization results are illustrated by using the mixed convexity results obtained for the cost function $M(N, \mu)$. In addition to the cases for which minimal values of the cost function M can be found, we consider the cases where the real convexity conditions of the cost function M hold but optimal values do not exist, and the real convexity conditions of the cost function M do not hold, and therefore minimal values do not exist.

Case 1: In this case we assume the arrival rate to be 1, the tolerable delay of the customers to be 0.5 units away from his arrival time, holding cost per unit per unit time is 6, cost per unit time of keeping an order waiting is 4, the penalty cost per lost orders is 35, and the initial number of items in the stock varies. Computations indicate the real convexity of the function $M(N, m)$ for each fixed N; however, minimum values of the cost function do not exist. That is, in the case in which the arrival rate of the customers is 1 and the tolerable delay for each customer is 0.5 units away from his arrival time, the minimal total cost per unit time does not exist for any of the fixed initial number of items in the stock when the service rate changes for $\lambda < \mu$. Figure 22 illustrates the graphs of M versus μ for various values of N and indicates the real convexity behavior of M with respect to μ.

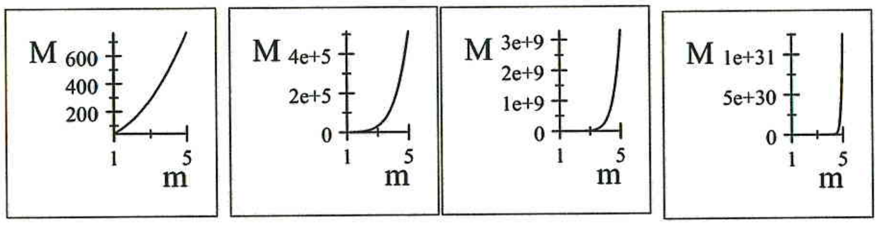


Fig 22 N=1, N=5, N=20, and N=40

Now suppose that the tolerable delay of the customers change from 0.5 to 0.9 and the other assumptions for this case remain the same. The change in the tolerable time delay does not result in a change where the minimal service required can be found for 1 item in the stock. Figure 23 illustrates the graphs of M versus μ for various values of N and indicates the real convexity behavior of M with respect to μ.

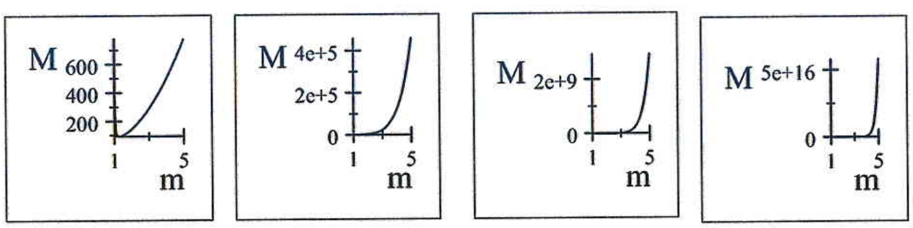


Fig 23 N=1, N=5, N=10, and N=20

Case 2: Assume the arrival rate to be 10, the tolerable delay of the customers to be 0.5 units away from his arrival time, holding cost per unit per unit time is 6, cost per unit time of keeping an order waiting is 4, the penalty cost per lost order is 35, and the initial number of items in the stock varies. In this case, figure 24, $N = 1$, is a counter example for the mixed convexity of the cost function. The numerical data indicate that M cannot be minimized for any number of initial number of items in stock.

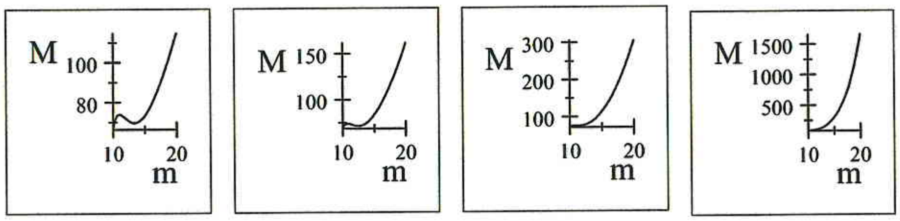


Fig 24 N=1, N=2, N=3, and N=5

Now suppose the tolerable delay of the customers change from 0.5 to 0.9 and the other assumptions in this case remain same. The increase in tolerable delay of the customers results in mixed convexity where the minimal total costs can be found. The numerical data indicate that the total costs are increasing as the initial number of items in the stock increases. In addition, the service rate decreases when the number of items in the initial stock increases. For various values of unknowns in this case, figures 25 – 28 illustrate the real convexity behavior of the cost function M versus μ for various values of N.

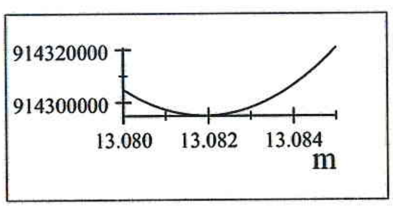


Fig 25. M versus m when N=1

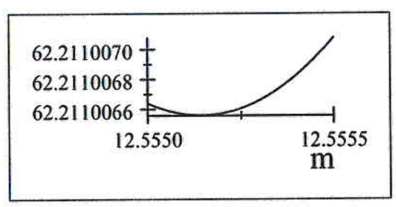


Fig 26. M versus m when N=2

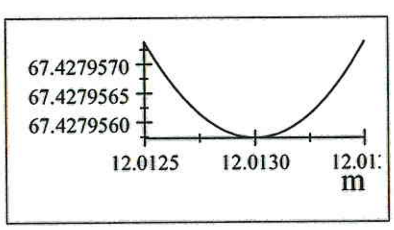


Fig 27. M versus m when N=3

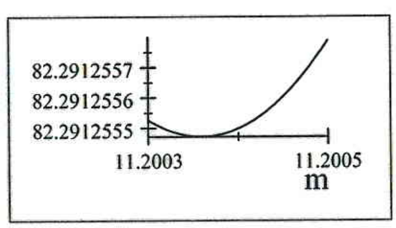


Fig 28. M versus m when N=5

Case 3: Suppose the arrival rate to be 100, the tolerable delay of the customer to be 0.5 units away from his arrival time, holding cost per unit per unit time is 6, cost per unit time of keeping an order waiting is 4, the penalty cost per lost orders is 35, and the initial number of items in the stock varies. The numerical data indicate that the mixed convexity holds where minimal total cost can be found. For various values of unknowns in this case, figures 29 – 32 illustrate the real convexity behavior of the cost function M versus μ for various values of N.

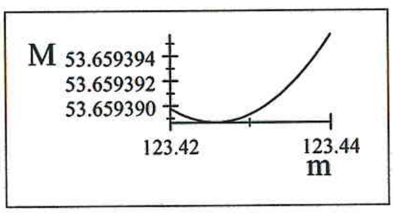


Fig 29. M versus m when

N=1

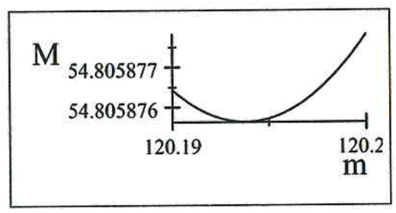


Fig 30. M versus m when

N=2

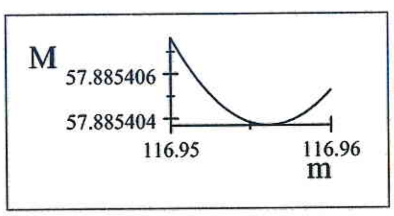


Fig 31. M versus m when

N=3

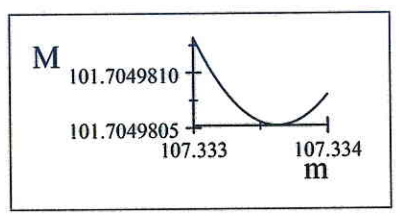


Fig 32. M versus m when

N=5

Under the assumptions of case 3, suppose we change the tolerable delay for each customer from 0.5 to 0.9 units away from his arrival time. This change results in an increase of the service rate and a decrease in the total cost. For various values of unknowns in this case, figures 33-36 illustrate the real convexity behavior of the cost function M versus μ for various values of N.

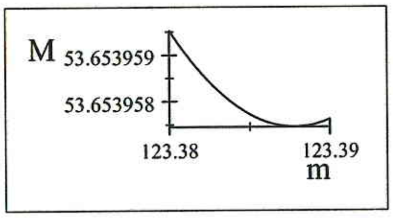


Fig 33. M versus m when

N=1

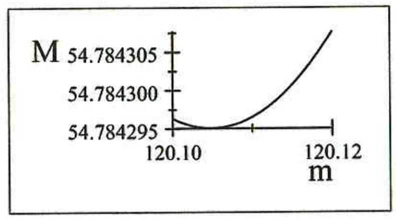


Fig 34. M versus m when

N=2

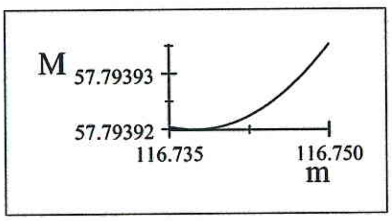


Fig 35. M versus m when

N=3

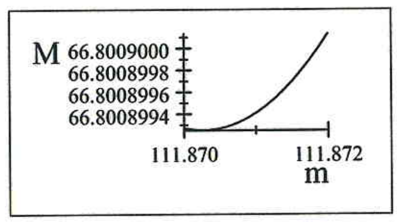


Fig 36. M versus m when

N=5

The following algorithm can be used to determine the minimal values of the cost function $M(N, \mu)$.

Algorithm 7.2:

`Define` $M(N, \mu)$.

`Apply algorithm 7.1 to find the mixed convexity region of` M `under the specified constraints and call it` D.

`Fix the integer variable` N `symbolically in region` D.

`Find the` μ `values by solving` $\frac{dM(N,\mu)}{d\mu} = 0$ `symbolically.`

The drawbacks of this algorithm are that the solution of the differential equation $\frac{dM(N,\mu)}{d\mu} = 0$ is not easy to obtain and the minimum value of the function depends on the strict mixed convex domain of the cost function.

The following table summarizes the convexity of the cost function M with respect to the initial number of items in the stock and the service rate.

Table 7.1 Convexity of $M(N, \mu)$ with varying parameters in λ

Convexity	λ 1	10	100
Integer convexity holds	$\mu \in [2, 40]$	$\mu \in [11, 40]$	$\mu \in [101, 200]$
Mixed convexity holds	*Locally Vary*	*Locally Vary*	*Locally Vary*

Table 7.2 Convexity of $M(N, \mu)$ for parameters C_h, C_d, C_l and T

Convexity	C_h	C_d	C_l	T
Integer convexity holds	$[1, 10]$	$[30, 40]$	$[1, 10]$	$[0.1, 0.9]$
Mixed convexity holds		*Locally Vary*		*Locally Vary*

The minimal values are summarized in the following table when M is mixed convex and $\frac{\partial M(N,\mu)}{\partial \mu} = 0$, the holding cost per unit per unit time is 6, the cost per unit time of keeping an order waiting is 4, and the penalty cost per lost order is 35 and $\lambda + 1 \leq \mu \leq 40$.

Table 7.3 Some of the Minimum Values of $M(N, \mu)$

Test Values		$1 \leq \mu \leq 40$	$11 \leq \mu \leq 40$	$101 \leq \mu \leq 400$
$T = 0.5$,	$N = 1$	D.N.E.	D.N.E.	53.6594
	$N = 2$			54.8059
	$N = 3$			57.8854
	$N = 5$			101.705
$T = 0.9$,	$N = 1$	D.N.E.	59.9143	53.654
	$N = 2$		62.211	54.7843
	$N = 3$		67.428	57.7939
	$N = 5$		82.2913	66.801

where $D.N.E.$ (does not exist) is the case when the condition $\frac{\partial M(N,\mu)}{\partial \mu} = 0$ is not satisfied. By using least squares we obtain the following best fitting models that fit the minimal value data of M with respect to N when T is fixed and μ changes:

$$f(N, 0.9, 11 \leq \mu \leq 40) = -0.180483N^3 + 2.54305N^2 - 4.069067N + 61.6208, \quad (7.2)$$

$$f(N, 0.5, 101 \leq \mu \leq 400) = 1.32757N^3 - 6.9989N^2 + 12.85023N + 46.4805, \quad (7.3)$$

$$f(N, 0.9, 101 \leq \mu \leq 400) = -0.110417N^3 + 1.602145N^2 - 2.90323N + 55.0655, \quad (7.4)$$

with the corresponding curves in figure 37. It can be seen that even for the case when $N \leq 2$, the approximation function given in (7.3) has a non-convex behavior. It is not always obvious that the second forward difference of the cost function M is positive for $N \geq 2$ which indicates the complexity of the cost function M.

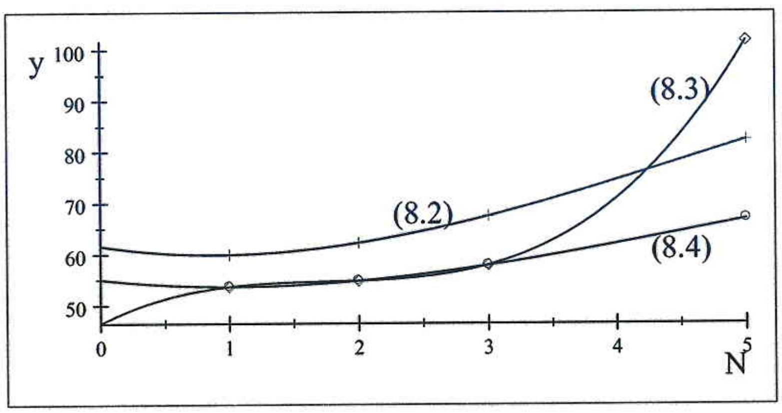


Fig 37 The graphs of the functions given in (8.2), (8.3), and (8.4).

7.4 CONCLUSION

In this work, computational mixed convexity and optimization results for the cost function corresponding to the $(S-1, S)$ inventory model under the time limit on backorders suggested by Das (1977) are given. A generalized convexity result for the cost function suggested by Das is also shown to find a region for which mixed convexity results hold. The integer convexity results obtained in this work support the integer convexity results obtained by Das for the cost function $M(N, \mu)$. However, the mixed convexity of the cost function with respect to the number of items initially in stock and the service rate parameters do not hold as shown by the numerical results in this work. Therefore, the same result holds true when a more general global mixed convexity result is considered for the cost function with respect to the integer variables λ (the arrival rate of the demand per unit time), N (the initial number of items in stock), C_h (the holding cost per unit per unit time), C_d (the cost per unit time of keeping an order waiting), and C_l (the penalty cost (in addition to TC_d) per lost order), and the real variables μ (the service rate) and T (the tolerable delay for each customer).

That is,

$$\begin{aligned} M &: \left(\mathbb{Z}^+\right)^5 \times \left(\mathbb{R}^+\right)^2 \to \mathbb{R}, \\ (\lambda, N, C_h, C_d, C_l, T, \mu) &\mapsto M(\lambda, N, C_h, C_d, C_l, T, \mu), \end{aligned}$$

does not satisfy the global strict mixed convexity properties with respect to the constraint $\lambda < \mu$. The violation of the global strict mixed convexity condition of $M(N, \mu)$ is based on the negative values of the second derivative of M in all the conditions other than when $T = 0.5$, $\lambda = 1$, $C_h = 6$, $C_d = 4$ and $C_l = 35$ when $1 \leq N \leq 40$ and $6 \leq m \leq 40$.

In this work, we obtained necessary and sufficient conditions to find a local mixed convexity result of the cost function M corresponding to the $(S-1, S)$ inventory model under the time limit on backorders suggested by Das, where minimal values of M can be obtained by using the mixed convexity region of the cost function.

Chapter 8

CONCLUSION AND FUTURE WORK

Convexity of multivariate mixed (integer and real variable) functions have important applications in many fields of study. Examples of such functions can be found in management science and telecommunication systems. In particular, some of these functions have number of servers to be the integer variable and the service rate to be the real variable. Special case convexity and optimization results for multivariate mixed functions are mainly obtained by either algorithmic or combinatorial approaches. In this work, a unified theoretical convexity method is introduced to obtain mixed convexity results of mixed variable functions. The main component of this mixed convexity method is a Hessian matrix which has a Hessian matrix for multivariate discrete functions as a special case. The mixed convexity of multivariate functions associated to

1. An $M/E_k/1$ queueing system suggested by Kumin (1973),
2. An $(S-1, S)$ inventory model suggested by Das (1977),
3. Erlang Delay and Loss formulae

are investigated using the method suggested in this work.

The optimization problems designed for queuing systems usually have mixed variable functions where the corresponding solutions obtained are particular solutions of the general problem. The future goal is to apply the method introduced in this work to obtain mixed convexity results for functions exist in queueing systems optimization problems. In addition, the results obtained for discrete convex functions can be improved for mixed variable functions.

REFERENCES

1.Al-Yakoob, S. M., Sherali, H. D. and Al-Jazzaf, M., (2010), A mixed-integer mathematical modeling approach to exam timetabling. Comput Manag Sci., Vol. 7, 19-46.

2.Adelman, D., (2008), A simple algebraic approximation to the Erlang loss system, Oper. Res. Let., Vol. 36(4), 484-491.

3.Andersson, J., and Melchiors, P., (2001), A two-echelon inventory model with lost sales, Int. J. Prod. Econ., Vol. 69, 307-315.

4.Benders, J. F., (2005), Partitioning procedures for solving mixed-variables programming problems, Comput. Manag. Sci., Vol. 2, 3-19.

5.Borwein, J. M., and Lewis, A. S., (2000), Convex analysis and nonlinear optimization: Theory and examples, Springer-Verlag, Berlin.

6.Boyd, S., and Vandenberghe, L., (2009), Convex optimization, Cambridge University Press, New York.

7.Cooper, R. B., (1981), Introduction to Queueing theory, 2nd edn. North-Holland, Amsterdam.

8.Das, C., (1977), The (S-1,S) Inventory model under time limit on back-orders, Oper. Res., Vol. 25(5), 835-850.

9.Denardo, E. V., (1982), Dynamic programming, Prentice-Hall, Englewood cli3s, NJ.

10.Dyer, M. E., and Proll, L. G., (1977), On the validity of Marginal Analysis for Allocating Servers in M/M/c Queues, Manag. Sci., Vol. 23, 1019-1022.

11.Erlang, A. K., (1917), Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges, Post Office Electrical Engineer's journal, Vol. 10, 189-197.

12.Favati, P., and Tardella, F., (1990), Convexity in Nonlinear Integer Programming, Ricerca Operativa, Vol. 53, 3-44.

13.Feeney, G., and Sherbrooke, C., (1966), The (S-1,S) Inventory policy under compound poisson demand, Manag. Sci., Vol. 12(5), 391-411.

14.Fenchel, W., (1953), Convex Cones, Sets, and Functions. Princeton, New Jersey: Princeton University Dept. of Mathematics. http://rasmusen.org /x/abros/fenchel.pdf.

15.Fenchel, W., and Bonnesen, T., (1971), Theorie der konvexen Körper, Bronx, New York, Chelsea Publishing Co.

16.Fenchel, W., and Bonnesen, T., (1974), Theorie der konvexen Körper, Berlin-New York, Springer-Verlag.

17.Fenchel, W., and Bonnesen, T., (1987), Theory of convex bodies, Moscow, Idaho, L. Boron, C. Christenson and B. Smith. BCS Associates.

18.Flood, J. E., (1998), Telecommunications Switching, Traffic and Networks, Chapter 4: Telecommunications Traffic, New York, Prentice-Hall.

19.Fox, B., (1966), Discrete optimization via marginal analysis, Manag. Sci., Vol. 13, 210–216.

20.Fuhrmann, S., Kogan, Y., and Milito, R. A., (1996), An adaptive autonomous network congestion controller, In Proceeding of the 35th IEEE Conference on Decision and Control, Kobe, Japan, 301-306.

21.Fujishige, S., and Murota, K., (2000), Notes on L-/M-convex functions and the separation theorems, Math. Program, Vol. 88, 129-146.

22.Galliher, H., Morse, P., and Simond, M., (1959), Dynamics of two classes of continuous-review inventory systems, Oper. Res., Vol. 7, 362-384.

23.Grassmann, W., (1983), The Convexity of the Mean Size of the M/M/c Queue with respect to the Traffic Intensity, J. Appl. Prob., Vol. 20, 916-919.

24.Gross, D., and Harris, C. M., (1973), Continuous-Review (s,S) inventory models with state-dependent lead times, Manag. Sci., Vol. 19, 567-574.

25.Gümüş, Z. H., and Floudas, C. A., (2005), Global optimization of mixed-integer bilevel programming problems, Comput. Manag. Sci., Vol. 2, 181-212.

26.Hadley, G. and Whitin, T., (1963), Analysis of inventory systems, Prentice-Hall, Englewood Cliffs, New Jersey.

27.Halfin, S., and Whitt, W., (1981), Heavy-trafic limits for for queues with many exponential servers, Oper. Res., Vol. 29, 567-588.

28.Harel, A., (1988), Sharp bounds and simple approximations for the Erlang delay and loss formulas, Manag. Sci., Vol. 34, 959-972.

29.Harel, A., (1988), Erratum to: sharp bounds and simple approximations for the Erlang delay and loss formulas, Manag. Sci., Vol. 34, 1277.

30.Harel, A., (1990), Convexity properties of the Erlang loss formula, Oper. Res., Vol. 38, 499-505.

31.Harel, A., (2010), Sharp and simple bounds for the Erlang delay and loss formulae, Queueing Syst., Vol. 64(2), 119-143.

32.Harel, A., and Zipkin, P., (1987), Strong convexity results for queueing systems, Oper. Res., Vol. 35(3), 405-418.

33.Hayya, J. C., Bagchi, U., and Ramasesh, R., (2011), Cost relationships in stochastic inventory systems: A simulation study of the (S, S-1, t=1) model, Int. J. Prod. Econ. Vol. 130, 196–202

34.Hirai, H. and Murota, K., (2004), M-convex functions and tree metrics, Japan J. Industrial Applied Math., Vol. 21, 391-403.

35.Jagerman, D. L., (1974), Some properties of the Erlang loss function, Bell Syst. Tech. J., Vol. 53, 525-551.

36.Jagers, A. A., and van Doorn, E. A., (1991), Convexity of functions which are generalizations of the Erlang loss function and the Erlang delay function. SIAM Rev., Vol. 33(2), 281-282.

37.Janssen, A. J. E. M., van Leeuwaarden, J. S. H., and Zwart, B., (2008), Corrected asymptotic for a multi-server queue in the Halfin-Whitt regime, Queueing Syst., Vol. 58(4), 261-301.

38.Kiselman, C.O., and Christer, O., (2010), Local minima, marginal functions, and separating hyperplanes in discrete optimization, In: Rajendra Bhatia (Ed.), Abstracts: Short communications; Posters. International Congress of Mathematicians, Hyderabad, August 19-27, 572-573.

39.Kiselman, C. O., and Samieinia S., (2010), Convexity of marginal functions in the discrete case, www2.math.uu.se/~kiselman/papersix.pdf, manuscript.

40.Ko, S. S., Serfozo, R., and Sivakumar, A. I., (2004), Reducing cycle times in manifacturing and supply chains by input and service rate smoothing. IIE Trans., Vol. 36(2), 145-153.

41.Kumin, H., (1973), On characterizing the extrema of a function of two variables, one of which is discrete, Manag. Sci., Vol. 20, 126-129.

42.Lee, H. L., and Cohen, M. A., (1983), A note on the Convexity of Performance Measures of M/M/c Queueing Systems, J. Appl. Prob., Vol. 20, 920-923.

43.Lee, H. L., and Cohen, M. A., (1985), Multi-Agent Customer Allocation in a Stochastic Service System, Manag. Sci., Vol. 31, 752-763.

44.Liu, L., and Cheung, K. L., (1997), Service constrained inventory models with random lifetimes and lead times, J. Oper. Res. Soc., Vol. 48(10), 1022-1028.

45.Miller, B. L., (1971), On minimizing nonseparable functions defined on the integers with an inventory application, SIAM J. Appl. Math., Vol. 21, 166-185.

46.Moinzadeh, K., (1989), Operating characteristics of the (S–1, S) inventory system with partial backorders and constant resupply times, Manag. Sci., Vol. 35(4), 472–477.

47.Moinzadeh, K., and Schmidt, C. P., (1991), An (S–1, S) inventory system with emergency orders, Oper. Res., Vol. 39(3), 308–321.

48.Moriguchi, S., and Murota, K., (2005), Discrete Hessian matrix for L-convex functions, IECE Trans. Fundamentals, E88-A.

49.Murota, K., (1996), Convexity and Steinitz's exchange property, Adv. Math., 272-311.

50.Murota, K., (1998), Discrete convex analysis, Math. Prog., Vol. 83, 313-371.

51.Murota, K., and Shioura, A., (2000), Extension of M-convexity and L-convexity to polyhedral convex functions, Advances in Applied Mathematics, Vol. 25, 352-427.

52.Newell, G. F., (1984), The $M/M/\infty$ Service system with ranked Servers in Heavy Traffic, Springer, New York.

53.Rockafellar, R. T., (1970), Convex analysis, Princten University Press, Princeton, NJ.

54.Rockafellar, R. T., (1974), Conjugate duality and optimization, SIAM Regional Conference Series in Applied Mathematics, Vol. 16, SIAM, Philadelphia.

55.Rockafellar, R. T., and Wets, R. J .-B., (1998), Variational analysis, Springer-Verlag, Berlin.

56.Rolfe, A. J., (1971), A Note on the Marginal Allocation in Multi-Server Facilities, Manag. Sci., Vol. 17, 656-658.

57.Smith, S. A., (1977), Optimal inventories for (S–1, S) system with no backorders, Manag. Sci., Vol. 23(5), 522–528.

58.Sobel, M. J., (1980), Simple inequality for multiserver queues, Manag. Sci., Vol. 26, 951-956.

59.Stoer, J., and Witzgall, C., (1970), Convexity and optimization in finite dimensions I, Springer-Verlag, Berlin.

60.Tokgöz, E., (2009), Algorithms for mixed convexity and optimization of 2-smooth mixed functions, Int. J. Pure Appl. Math., Vol. 57(1), 103-110.

61.Tokgöz, E., Maalouf, M., and Kumin, H., (2009), A Hessian matrix for functions with integer and continuous variables, Int. J. Pure Appl. Math., Vol. 57(2), 209-218.

62.Tokgöz, E., Nourazari, S., and Kumin, H., (2011), Convexity and optimization of condense discrete functions, P.M. Pardalos and S. Rebennack (Eds.), LNCS 6630, 33–42, Springer-Verlag Berlin Heidelberg.

63.Ui, T., (2006), A note on discrete convexity and local optimality, Japan J. Indust. Appl. Math., Vol. 23, 21-29.

64.Whitt, W., (2002), The Erlang B and C formulas: Problems and solutions. Unpublished Manuscript. http://www.columbia/~ww2040/ErlangBandCFormulas.pdf.

65.Yüceer, U., (2002), Discrete convexity: convexity for functions defined on discrete spaces., Discrete Appl. Math., Vol. 119, 297-304.

Appendix A

MIXED CONVEXITY OF THE ERLANG LOSS FORMULA

By using the following symbolic MATLAB R2009a program we obtain the determinant of the mixed Hessian matrix corresponding to the Erlang Loss formula.

```
syms l s m e_2 e_1 e eP1 eP2 eps_1 eps epsP1 epsP2
%----------------- All the epsilon's are written in terms of eps  ----------%
epsP2 = ((l^2)/((s+2)*(s+1)*m^2))*eps
epsP1 = (l/((s+1)*m))*eps
eps_1 = ((s*m)/l)*eps
%---------------- All the e terms written in terms of es ----------------%
e_2 = e - eps - eps_1
e_1 = e - eps
eP1 = e + epsP1
eP2 = e + epsP1 + epsP2
%------- Mixed convexity components of the Erlang Loss Fromula------------%
Nabla_11 = (epsP2*eP1*e - 2*epsP1*eP2*e + eps*eP2*eP1)/(eP2*eP1*e)
Diff_Nabla1 = (1/((m*eP1*e)^2))*(-eps*eP1*e^2 + epsP1*e^3 + eps_1*(eP1^2)*e
- eps*(eP1^2)*e_1)
Scnd_Diff = (eps/((m^4)*e^3))* ( s*(m^2)*(s+1)*e^2
- 2*l*m*(s+1)*e*e_1 + 2*(l^2)*(e_1)^2 - (l^2)*e*e_2)
Det = (Nabla_11)*(Scnd_Diff) - (Diff_Nabla1)^2
A = collect(Det,e,eps)
B = simplify(Det)
C = simple(Det)
R = pretty(C)
```

The output of the algorithm gives the following:

$$\begin{aligned} Det\,(H) \;=\; & -\frac{\lambda^2\epsilon^2}{m^4 e_s^4 e_{s+1}^4}\left[e_s^2 e_{s+1} + e_{s-1}e_{s+1}^2 - \frac{\lambda}{m\,(s+1)}e_s^3 - \frac{sm}{\lambda}\left(e_s e_{s+1}^2\right)\right]^2 \\ & +\frac{\epsilon^2}{m^4 e_s^4 e_{s+1} e_{s+2}}\left[e_{s+1}e_{s+2} - 2\frac{\lambda}{m\,(s+1)}\left(e_s e_{s+2}\right) + \frac{\lambda^2}{m^2\,(s+1)\,(s+2)}\left(e_s e_{s+1}\right)\right]\cdot \\ & \left[2\lambda^2 e_{s-1}^2 + \lambda^2 e_s\left(\frac{sm}{\lambda}\epsilon - e_{s-1}\right) + m^2 s\,(s+1)\,e_s^2 - 2\lambda m\,(s+1)\,e_s e_{s-1}\right] \end{aligned}$$

$$
\begin{aligned}
Det\,(H) &= -\left[\frac{\lambda\epsilon}{m^2}\left(\frac{1}{e_{s+1}}\right)+\frac{\lambda\epsilon}{m^2}\left(\frac{e_{s-1}}{e_s^2}\right)-\frac{\lambda^2\epsilon}{m^3\,(s+1)}\left(\frac{e_s}{e_{s+1}^2}\right)-\frac{s\epsilon}{m}\frac{1}{e_s}\right]^2 \\
&+\left[\frac{\epsilon^2}{m^4}\left(\frac{1}{e_s^4}\right)-\frac{2\lambda\epsilon^2}{m^5\,(s+1)}\left(\frac{1}{e_s^3e_{s+1}}\right)+\frac{\lambda^2\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{1}{e_s^3e_{s+2}}\right)\right]\cdot \\
&\left[2\lambda^2e_{s-1}^2+sm\lambda\epsilon e_s-\lambda^2e_{s-1}e_s+e_s^2m^2s\,(s+1)-2\lambda me_se_{s-1}\,(s+1)\right]
\end{aligned}
$$

After multiplying out the terms in the determinant we obtain

$$
\begin{aligned}
Det\,(H) &= \left[-\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{1}{e_{s+1}^2}\right)-\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}^2}{e_s^4}\right)-\frac{\lambda^4\epsilon^2}{m^6\,(s+1)^2}\left(\frac{e_s^2}{e_{s+1}^4}\right)-\frac{s^2\epsilon^2}{m^2}\frac{1}{e_s^2}\right] \\
&+\left[-2\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^2e_{s+1}}\right)+2\frac{\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_s}{e_{s+1}^3}\right)+2\frac{\lambda s\epsilon^2}{m^3}\left(\frac{1}{e_se_{s+1}}\right)\right] \\
&+\left[2\frac{\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}}{e_se_{s+1}^2}\right)+2\frac{\lambda s\epsilon^2}{m^3}\left(\frac{e_{s-1}}{e_s^3}\right)\right]-\left[2\frac{\lambda^2s\epsilon^2}{m^4\,(s+1)}\left(\frac{1}{e_{s+1}^2}\right)\right] \\
&+\frac{2\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}^2}{e_s^4}\right)+\frac{s\lambda\epsilon^3}{m^3}\left(\frac{1}{e_s^3}\right)-\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^3}\right)+\frac{s\,(s+1)\,\epsilon^2}{m^2}\left(\frac{1}{e_s^2}\right) \\
&-\frac{2\lambda\epsilon^2\,(s+1)}{m^3}\left(\frac{e_{s-1}}{e_s^3}\right)-\{\frac{4\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}^2}{e_s^3e_{s+1}}\right)+\frac{2\lambda^2\epsilon^3s}{m^4\,(s+1)}\left(\frac{1}{e_s^2e_{s+1}}\right) \\
&-\frac{2\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}}{e_s^2e_{s+1}}\right)+\frac{2\lambda s\epsilon^2}{m^3}\left(\frac{1}{e_se_{s+1}}\right)-\frac{4\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^2e_{s+1}}\right)\} \\
&+\frac{2\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}^2}{e_s^3e_{s+2}}\right)+\frac{\lambda^3\epsilon^3s}{m^5\,(s+1)\,(s+2)}\left(\frac{1}{e_s^2e_{s+2}}\right) \\
&-\frac{\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}}{e_s^2e_{s+2}}\right)+\frac{\lambda^2\epsilon^2s}{m^4\,(s+2)}\left(\frac{1}{e_se_{s+2}}\right)-\frac{2\lambda^3\epsilon^2}{m^5\,(s+2)}\left(\frac{e_{s-1}}{e_s^2e_{s+2}}\right)
\end{aligned}
$$

To be able to cancel out the common terms we number the components of the determinant as follows:

$$
\begin{aligned}
Det\,(H) &= -\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{1}{e_{s+1}^2}\right)-\underbrace{\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}^2}{e_s^4}\right)}_{\boxed{1.1}}-\frac{\lambda^4\epsilon^2}{m^6\,(s+1)^2}\left(\frac{e_s^2}{e_{s+1}^4}\right)-\underbrace{\frac{s^2\epsilon^2}{m^2}\frac{1}{e_s^2}}_{\boxed{2.1}} \\
&\underbrace{-2\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^2e_{s+1}}\right)}_{\boxed{5.1}}+2\frac{\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_s}{e_{s+1}^3}\right)+\underbrace{2\frac{\lambda s\epsilon^2}{m^3}\left(\frac{1}{e_se_{s+1}}\right)}_{\boxed{4.1}} \\
&+2\frac{\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}}{e_se_{s+1}^2}\right)+\underbrace{2\frac{\lambda s\epsilon^2}{m^3}\left(\frac{e_{s-1}}{e_s^3}\right)}_{\boxed{3.1}}-2\frac{\lambda^2s\epsilon^2}{m^4\,(s+1)}\left(\frac{1}{e_{s+1}^2}\right)
\end{aligned}
$$

$$+\underbrace{\frac{2\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}^2}{e_s^4}\right)}_{\boxed{1.2}}+\frac{s\lambda\epsilon^3}{m^3}\left(\frac{1}{e_s^3}\right)-\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^3}\right)+\underbrace{\frac{s\left(s+1\right)\epsilon^2}{m^2}\left(\frac{1}{e_s^2}\right)}_{\boxed{2.2}}$$

$$-\underbrace{\frac{2\lambda\epsilon^2\left(s+1\right)}{m^3}\left(\frac{e_{s-1}}{e_s^3}\right)}_{\boxed{3.2}}-\frac{4\lambda^3\epsilon^2}{m^5\left(s+1\right)}\left(\frac{e_{s-1}^2}{e_s^3e_{s+1}}\right)-\frac{2\lambda^2\epsilon^3 s}{m^4\left(s+1\right)}\left(\frac{1}{e_s^2e_{s+1}}\right)$$

$$+\frac{2\lambda^3\epsilon^2}{m^5\left(s+1\right)}\left(\frac{e_{s-1}}{e_s^2e_{s+1}}\right)-\underbrace{2\frac{\lambda\epsilon^2 s}{m^3}\left(\frac{1}{e_se_{s+1}}\right)}_{\boxed{4.2}}+\underbrace{\frac{4\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^2e_{s+1}}\right)}_{\boxed{5.2}}$$

$$+\frac{2\lambda^4\epsilon^2}{m^6\left(s+1\right)\left(s+2\right)}\left(\frac{e_{s-1}^2}{e_s^3e_{s+2}}\right)+\frac{\lambda^3\epsilon^3 s}{m^5\left(s+1\right)\left(s+2\right)}\left(\frac{1}{e_s^2e_{s+2}}\right)$$

$$-\frac{\lambda^4\epsilon^2}{m^6\left(s+1\right)\left(s+2\right)}\left(\frac{e_{s-1}}{e_s^2e_{s+2}}\right)+\frac{\lambda^2\epsilon^2 s}{m^4\left(s+2\right)}\left(\frac{1}{e_se_{s+2}}\right)-\frac{2\lambda^3\epsilon^2}{m^5\left(s+2\right)}\left(\frac{e_{s-1}}{e_s^2e_{s+2}}\right)$$

Simplifying the terms $(1.1)-(1.2)$ through $(5.1)-(5.2)$ we obtain

$$\begin{aligned}
(1.1)+(1.2) &= \frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}^2}{e_s^4}\right),\\
(2.1)+(2.2) &= \frac{s\epsilon^2}{m^2}\left(\frac{1}{e_s^2}\right),\\
(3.1)+(3.2) &= -\frac{2\lambda\epsilon^2}{m^3}\left(\frac{e_{s-1}}{e_s^3}\right),\\
(4.1)+(4.2) &= 0,\\
(5.1)+(5.2) &= 2\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^2e_{s+1}}\right).
\end{aligned}$$

Therefore after simplification we obtain the determinant of the mixed Hessian to be

$$Det\left(H\right) = -\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{1}{e_{s+1}^2}\right)+\underbrace{\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}^2}{e_s^4}\right)}_{\substack{\text{Follows from 1}\\ \boxed{7.1}}}$$

$$-\frac{\lambda^4\epsilon^2}{m^6\left(s+1\right)^2}\left(\frac{e_s^2}{e_{s+1}^4}\right)+\underbrace{\frac{s\epsilon^2}{m^2}\left(\frac{1}{e_s^2}\right)}_{\boxed{\text{Follows from 2}}}$$

$$+\underbrace{\frac{2\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^2 e_{s+1}}\right)}_{\boxed{\text{Follows from 5}}}+2\frac{\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_s}{e_{s+1}^3}\right)+2\frac{\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}}{e_s e_{s+1}^2}\right)$$

$$-2\frac{\lambda^2 s\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_{s+1}^2}\right)+\frac{\epsilon^3 s\lambda}{m^3}\left(\frac{1}{e_s^3}\right)-\underbrace{\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^3}\right)}_{\boxed{7.2}}-\underbrace{\frac{2\lambda\epsilon^2}{m^3}\left(\frac{e_{s-1}}{e_s^3}\right)}_{\boxed{\text{Follows from 3}}}$$

$$\underbrace{-\frac{4\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+1}}\right)}_{\boxed{6.1}}-\frac{2\lambda^2\epsilon^3 s}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}}\right)+\underbrace{\frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}}{e_s^2 e_{s+1}}\right)}_{\boxed{6.2}}$$

$$+\frac{2\lambda^4\epsilon^2}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right)+\frac{\lambda^3\epsilon^3 s}{m^5(s+1)(s+2)}\left(\frac{1}{e_s^2 e_{s+2}}\right)$$

$$-\frac{\lambda^4\epsilon^2}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}}{e_s^2 e_{s+2}}\right)+\frac{\lambda^2\epsilon^2 s}{m^4(s+2)}\left(\frac{1}{e_s e_{s+2}}\right)-\frac{2\lambda^3\epsilon^2}{m^5(s+2)}\left(\frac{e_{s-1}}{e_s^2 e_{s+2}}\right)$$

Simplifying the terms (6.1) with (6.2), and (7.1) with (7.2) we obtain

$$\begin{aligned}
(6.1)+(6.2) &= -\frac{4\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+1}}\right)+\frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}}{e_s^2 e_{s+1}}\right)\\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}}{e_s^3 e_{s+1}}\right)[-2e_{s-1}+e_s]\\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}}{e_s^3 e_{s+1}}\right)[-e_{s-1}+\epsilon_s]\\
&= -\frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+1}}\right)+\underbrace{\frac{2\lambda^3\epsilon^3}{m^5(s+1)}\left(\frac{e_{s-1}}{e_s^3 e_{s+1}}\right)}_{c_1}\\
&= -\frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+1}}\right)+c_1
\end{aligned}$$

$$\begin{aligned}
(7.1)+(7.2) &= \frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}^2}{e_s^4}\right)-\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^3}\right)\\
&= \frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}^2-e_s e_{s-1}}{e_s^4}\right)\\
&= \frac{\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}^2-(e_{s-1}+\epsilon_s)e_{s-1}}{e_s^4}\right)\\
&= \frac{\lambda^2\epsilon^2}{m^4}\left(\frac{-\epsilon_s e_{s-1}}{e_s^4}\right)=\frac{-\lambda^2\epsilon^3}{m^4}\left(\frac{e_{s-1}}{e_s^4}\right)
\end{aligned}$$

Therefore,

$$
\begin{aligned}
Det\,(H) \;=\; & -\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{1}{e_{s+1}^2}\right) - \frac{\lambda^4\epsilon^2}{m^6\,(s+1)^2}\left(\frac{e_s^2}{e_{s+1}^4}\right) + \frac{2\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_s}{e_{s+1}^3}\right) \\
& +2\frac{\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}}{e_s e_{s+1}^2}\right) - 2\frac{\lambda^2\epsilon^2 s}{m^4\,(s+1)}\left(\frac{1}{e_{s+1}^2}\right) \\
& -\frac{\lambda^2\boxed{\epsilon^3}}{m^4}\left(\frac{e_{s-1}}{e_s^4}\right) + \frac{s\lambda\boxed{\epsilon^3}}{m^3}\left(\frac{1}{e_s^3}\right) + \underbrace{\frac{s\epsilon^2}{m^2}\left(\frac{1}{e_s^2}\right)}_{\boxed{10.1}} - \underbrace{\frac{2\lambda\epsilon^2}{m^3}\left(\frac{e_{s-1}}{e_s^3}\right)}_{\boxed{10.2}} \\
& +\left[-\frac{2\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+1}}\right) + c_1\right] - \frac{2\lambda^2\boxed{\epsilon^3}s}{m^4\,(s+1)}\left(\frac{1}{e_s^2 e_{s+1}}\right) \\
& +\frac{2\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^2 e_{s+1}}\right) + \underbrace{\frac{2\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right)}_{\boxed{11.1}} \\
& +\frac{\lambda^3\boxed{\epsilon^3}s}{m^5\,(s+1)\,(s+2)}\left(\frac{1}{e_s^2 e_{s+2}}\right) - \underbrace{\frac{\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}}{e_s^2 e_{s+2}}\right)}_{\boxed{11.2}} \\
& +\frac{\lambda^2\epsilon^2 s}{m^4\,(s+2)}\left(\frac{1}{e_s e_{s+2}}\right) - \frac{2\lambda^3\epsilon^2}{m^5\,(s+2)}\left(\frac{e_{s-1}}{e_s^2 e_{s+2}}\right)
\end{aligned}
$$

We first simplify the terms given in (10.1) and (10.2);

$$
\begin{aligned}
(10.1) + (10.2) &= \frac{\epsilon^2}{m^2}\left(\frac{1}{e_s^3}\right)\left(se_s - 2\frac{\lambda}{m}e_{s-1}\right) \\
&= \frac{\epsilon^2}{m^3}\left(\frac{1}{e_s^3}\right)\left[sme_s - 2\lambda e_{s-1}\right],
\end{aligned}
$$

and second simplify the terms (11.1) and (11.2);

$$
\begin{aligned}
(11.1) + (11.2) &= \frac{\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}}{e_s^3 e_{s+2}}\right)(2e_{s-1} - e_s) \\
&= \frac{\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}}{e_s^3 e_{s+2}}\right)(e_{s-1} + e_{s-1} - e_s) \\
&= \frac{\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}}{e_s^3 e_{s+2}}\right)(e_{s-1} - \epsilon_s) \\
&= \frac{\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right) - \frac{\lambda^4\epsilon^3}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}}{e_s^3 e_{s+2}}\right).
\end{aligned}
$$

Therefore

$$
\begin{aligned}
Det\,(H) \;=\; & -\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{1}{e_{s+1}^2}\right) - \frac{\lambda^4\epsilon^2}{m^6\,(s+1)^2}\left(\frac{e_s^2}{e_{s+1}^4}\right) + \frac{2\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_s}{e_{s+1}^3}\right) \\
& +2\frac{\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}}{e_s e_{s+1}^2}\right) - 2\frac{\lambda^2\epsilon^2 s}{m^4\,(s+1)}\left(\frac{1}{e_{s+1}^2}\right) \\
& -\frac{\lambda^2\boxed{\epsilon^3}}{m^4}\left(\frac{e_{s-1}}{e_s^4}\right) + \frac{s\lambda\boxed{\epsilon^3}}{m^3}\left(\frac{1}{e_s^3}\right) + \underbrace{\frac{\epsilon^2}{m^3}\left(\frac{1}{e_s^3}\right)[sme_s - 2\lambda e_{s-1}]}_{\boxed{\text{Result of 10}}} \\
& -\frac{2\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+1}}\right) + c_1 - \frac{2\lambda^2\boxed{\epsilon^3}s}{m^4\,(s+1)}\left(\frac{1}{e_s^2 e_{s+1}}\right) + \frac{2\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^2 e_{s+1}}\right) \\
& +\underbrace{\frac{\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right) - \frac{\lambda^4\boxed{\epsilon^3}}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}}{e_s^3 e_{s+2}}\right)}_{\boxed{\text{Result of 11}}} \\
& +\frac{\lambda^3\boxed{\epsilon^3}s}{m^5\,(s+1)\,(s+2)}\left(\frac{1}{e_s^2 e_{s+2}}\right) + \frac{\lambda^2\epsilon^2 s}{m^4\,(s+2)}\left(\frac{1}{e_s e_{s+2}}\right) \\
& -\frac{2\lambda^3\epsilon^2}{m^5\,(s+2)}\left(\frac{e_{s-1}}{e_s^2 e_{s+2}}\right)
\end{aligned}
$$

In the above inequality, first we collect and write the negative terms that contains ϵ^2, second we collect and write the positive terms with ϵ^2, and last, we collect and write the terms that contain ϵ^3. Therefore we have

$$
\begin{aligned}
Det\,(H) \;=\; & -\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{1}{e_{s+1}^2}\right) \underbrace{-\frac{\lambda^4\epsilon^2}{m^6\,(s+1)^2}\left(\frac{e_s^2}{e_{s+1}^4}\right)}_{\boxed{8.1}} \underbrace{-\frac{2\lambda^2\epsilon^2 s}{m^4\,(s+1)}\left(\frac{1}{e_{s+1}^2}\right)}_{\boxed{12.1}} \\
& -\frac{2\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+1}}\right) - \frac{2\lambda^3\epsilon^2}{m^5\,(s+2)}\left(\frac{e_{s-1}}{e_s^2 e_{s+2}}\right) \\
& +\frac{2\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_s}{e_{s+1}^3}\right) + \frac{\epsilon^2}{m^3}\left(\frac{1}{e_s^3}\right)[sme_s - 2\lambda e_{s-1}] \\
& +c_1 + \frac{2\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}}{e_s e_{s+1}^2}\right) + \underbrace{\frac{2\lambda^2\epsilon^2}{m^4}\left(\frac{e_{s-1}}{e_s^2 e_{s+1}}\right)}_{\boxed{12.2}} \\
& +\underbrace{\frac{\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right)}_{\boxed{8.2}} + \frac{\lambda^2\epsilon^2 s}{m^4\,(s+2)}\left(\frac{1}{e_s e_{s+2}}\right)
\end{aligned}
$$

$$
-\frac{\lambda^2\boxed{\epsilon^3}}{m^4}\left(\frac{e_{s-1}}{e_s^4}\right)+\frac{s\lambda\boxed{\epsilon^3}}{m^3}\left(\frac{1}{e_s^3}\right)-\frac{2\lambda^2\boxed{\epsilon^3}s}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}}\right)
$$

$$
-\frac{\lambda^4\boxed{\epsilon^3}}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}}{e_s^3 e_{s+2}}\right)+\frac{\lambda^3\boxed{\epsilon^3}s}{m^5(s+1)(s+2)}\left(\frac{1}{e_s^2 e_{s+2}}\right)
$$

Simplifying (8.1) and (8.2) we have

$$
\begin{aligned}
(8.1)+(8.2) &= -\frac{\lambda^4\epsilon^2}{m^6(s+1)^2}\left(\frac{e_s^2}{e_{s+1}^4}\right)+\frac{\lambda^4\epsilon^2}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right)\\
&= \frac{\lambda^4\epsilon^2}{m^6(s+1)}\left[-\frac{1}{s+1}\left(\frac{e_s^2}{e_{s+1}^4}\right)+\frac{1}{s+2}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right)\right]\\
&= \frac{\lambda^4\epsilon^2}{m^6(s+1)^2(s+2)}\left(\frac{1}{e_s^3 e_{s+1}^4 e_{s+2}}\right)\left[-(s+2)e_s^5 e_{s+2}+(s+1)e_{s+1}^4 e_{s-1}^2\right]\\
&= \frac{\lambda^4\epsilon^2}{m^6(s+1)^2(s+2)}\left(\frac{1}{e_s^3 e_{s+1}^4 e_{s+2}}\right)\left[-e_s^5 e_{s+2}\right]+c_3\\
&= -\frac{\lambda^4\epsilon^2}{m^6(s+1)^2(s+2)}\left(\frac{e_s^2}{e_{s+1}^4}\right)+c_3
\end{aligned}
$$

Simplifying (12.1) and (12.2) we have

$$
\begin{aligned}
(12.1)+(12.2) &= 2\frac{\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}^2}\right)\left[(s+1)e_{s-1}e_{s+1}-se_s^2\right]\\
&= 2\frac{\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}^2}\right)\left[(s+1)(e_s-\epsilon)(e_s+\epsilon_{s+1})-se_s^2\right]\\
&= 2\frac{\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}^2}\right)\left[(s+1)\left(e_s^2-\epsilon e_s+\epsilon_{s+1}e_s-\epsilon\epsilon_{s+1}\right)-se_s^2\right]\\
&= 2\frac{\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}^2}\right)\left[se_s^2+e_s^2+(s+1)(-\epsilon e_s+\epsilon_{s+1}e_s-\epsilon\epsilon_{s+1})-se_s^2\right]\\
&= 2\frac{\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}^2}\right)\left[e_s^2+(s+1)(-\epsilon e_s+\epsilon_{s+1}e_s-\epsilon\epsilon_{s+1})\right]\\
&= 2\frac{\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}^2}\right)\left(e_s^2\right)\\
&\quad +2\frac{\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}^2}\right)(s+1)(-\epsilon e_s+\epsilon_{s+1}e_s-\epsilon\epsilon_{s+1})\\
&= 2\frac{\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_{s+1}^2}\right)+2\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{1}{e_s^2 e_{s+1}^2}\right)(-\epsilon_s e_s+\epsilon_{s+1}e_s-\epsilon_s\epsilon_{s+1})\\
&= 2\frac{\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_{s+1}^2}\right)+a_1
\end{aligned}
$$

Therefore we have

$$
\begin{aligned}
Det\,(H) \;=\; & \underbrace{-\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{1}{e_{s+1}^2}\right)}_{\boxed{13.1}} - \frac{2\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+1}}\right) \qquad (1) \\
& \underbrace{-\frac{\lambda^4\epsilon^2}{m^6\,(s+1)^2\,(s+2)}\left(\frac{e_s^2}{e_{s+1}^4}\right) + c_3}_{\boxed{\text{Remaining from 8}}} - \frac{2\lambda^3\epsilon^2}{m^5\,(s+2)}\left(\frac{e_{s-1}}{e_s^2 e_{s+2}}\right) \\
& + \frac{\epsilon^2}{m^3}\left(\frac{1}{e_s^3}\right)\left[sme_s - 2\lambda e_{s-1}\right] + \frac{2\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_s}{e_{s+1}^3}\right) \\
& + 2\frac{\lambda^3\epsilon^2}{m^5\,(s+1)}\left(\frac{e_{s-1}}{e_s e_{s+1}^2}\right) + \underbrace{2\frac{\lambda^2\epsilon^2}{m^4\,(s+1)}\left(\frac{1}{e_{s+1}^2}\right) + a_1}_{\boxed{13.3}} \\
& + \frac{\lambda^4\epsilon^2}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right) + \underbrace{\frac{\lambda^2\epsilon^2 s}{m^4\,(s+2)}\left(\frac{1}{e_s e_{s+2}}\right)}_{\boxed{13.2}} \\
& - \frac{\lambda^2\boxed{\epsilon^3}}{m^4}\left(\frac{e_{s-1}}{e_s^4}\right) + \frac{s\lambda\boxed{\epsilon^3}}{m^3}\left(\frac{1}{e_s^3}\right) + - \frac{2\lambda^2\boxed{\epsilon^3}s}{m^4\,(s+1)}\left(\frac{1}{e_s^2 e_{s+1}}\right) \\
& - \frac{\lambda^4\boxed{\epsilon^3}}{m^6\,(s+1)\,(s+2)}\left(\frac{e_{s-1}}{e_s^3 e_{s+2}}\right) + \frac{\lambda^3\boxed{\epsilon^3}s}{m^5\,(s+1)\,(s+2)}\left(\frac{1}{e_s^2 e_{s+2}}\right)
\end{aligned}
$$

First simplifying (13.1) with (13.2),

$$
\begin{aligned}
(13.1) + (13.2) \;=\; & -\frac{\lambda^2\epsilon^2}{m^4}\left(\frac{1}{e_{s+1}^2}\right) + \frac{\lambda^2\epsilon^2 s}{m^4\,(s+2)}\left(\frac{1}{e_s e_{s+2}}\right) \\
=\; & \frac{\lambda^2\epsilon^2}{m^4\,(s+2)\,e_s e_{s+2} e_{s+1}^2}\left[-\,(s+2)\,e_s e_{s+2} + s e_{s+1}^2\right] \\
=\; & \frac{\lambda^2\epsilon^2}{m^4\,(s+2)\,e_s e_{s+2} e_{s+1}^2}\left[-\,(s+2)\,e_s\left(e_s + \epsilon_{s+1} + \epsilon_{s+2}\right) + s\left(e_s + \epsilon_{s+1}\right)^2\right] \\
=\; & \frac{\lambda^2\epsilon^2}{m^4\,(s+2)\,e_s e_{s+2} e_{s+1}^2}\{-se_s^2 - 2e_s^2 \\
& -\,(s+2)\,e_s\left(\epsilon_{s+1} + \epsilon_{s+2}\right) + s\left(e_s^2 + 2e_s\epsilon_{s+1} + \epsilon_{s+1}^2\right)\} \\
=\; & \frac{\lambda^2\epsilon^2}{m^4\,(s+2)\,e_s e_{s+2} e_{s+1}^2}\{-se_s^2 - 2e_s^2 \\
& -\,(s+2)\,e_s\left(\epsilon_{s+1} + \epsilon_{s+2}\right) + se_s^2 + 2se_s\epsilon_{s+1} + s\epsilon_{s+1}^2\}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\lambda^2\epsilon^2}{m^4(s+2)\,e_s e_{s+2} e_{s+1}^2}\{-2e_s^2 - (s+2)\,e_s(\epsilon_{s+1}+\epsilon_{s+2}) + 2s e_s \epsilon_{s+1} + s\epsilon_{s+1}^2\} \\
&= \frac{\lambda^2\epsilon^2}{m^4(s+2)\,e_s e_{s+2} e_{s+1}^2}(-2e_s^2) + \frac{\lambda^2\epsilon^2}{m^4(s+2)}\left(\frac{1}{e_s e_{s+2} e_{s+1}^2}\right)\{-s e_s \epsilon_{s+2} \\
&\quad -2e_s(\epsilon_{s+1}+\epsilon_{s+2}) + s e_s \epsilon_{s+1} + s\epsilon_{s+1}^2\} \\
&= \frac{-2\lambda^2\epsilon^2}{m^4(s+2)}\left(\frac{e_s}{e_{s+2}e_{s+1}^2}\right) + c_3 \qquad (13.4)
\end{aligned}
$$

and then simplifying the outcome (13.4) with (13.3) we obtain

$$
\begin{aligned}
(13.4)+(13.3) &= \frac{-2\lambda^2\epsilon^2}{m^4(s+2)}\left(\frac{e_s}{e_{s+2}e_{s+1}^2}\right) + c_3 + 2\frac{\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{1}{e_{s+1}^2}\right) + a_1 \\
&= \frac{-2\lambda^2\epsilon^2}{m^4(s+2)}\left(\frac{e_s}{e_{s+2}e_{s+1}^2}\right) + \frac{2\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{e_{s+2}}{e_{s+2}e_{s+1}^2}\right) + \underbrace{a_1 + c_3}_{c_4} \\
&= \frac{-2\lambda^2\epsilon^2}{m^4(s+2)}\left(\frac{(s+1)\,e_s}{(s+1)\,e_{s+2}e_{s+1}^2}\right) + \frac{2\lambda^2\epsilon^2}{m^4(s+1)}\left(\frac{(s+2)\,e_{s+2}}{(s+2)\,e_{s+2}e_{s+1}^2}\right) + c_4 \\
&= \frac{2\lambda^2\epsilon^2}{m^4(s+1)(s+2)}\left(\frac{1}{e_{s+2}e_{s+1}^2}\right)[-(s+1)\,e_s + (s+2)\,e_{s+2}] + c_4 \\
&= \frac{2\lambda^2\epsilon^2}{m^4(s+1)(s+2)}\left(\frac{1}{e_{s+2}e_{s+1}^2}\right)[-(s+1)\,e_s + (s+1)\,e_{s+2} + e_{s+2}] + c_4 \\
&= \frac{2\lambda^2\epsilon^2}{m^4(s+1)(s+2)}\left(\frac{1}{e_{s+2}e_{s+1}^2}\right)[(s+1)(-e_s + e_{s+2}) + e_{s+2}] + c_4 \\
&= \frac{2\lambda^2\epsilon^2}{m^4(s+1)(s+2)}\left(\frac{1}{e_{s+2}e_{s+1}^2}\right)[(s+1)(\epsilon_{s+1} + \epsilon_{s+2}) + e_{s+2}] + c_4 \\
&= \frac{2\lambda^2\epsilon^2}{m^4(s+1)(s+2)}\left(\frac{1}{e_{s+1}^2}\right) + \frac{2\lambda^2\epsilon^2}{m^4(s+2)}\left(\frac{\epsilon_{s+1}+\epsilon_{s+2}}{e_{s+2}e_{s+1}^2}\right) + c_4 \\
&= \frac{2\lambda^2\epsilon^2}{m^4(s+1)(s+2)}\left(\frac{1}{e_{s+1}^2}\right) + c_5 \qquad (13.5)
\end{aligned}
$$

Therefore the determinant takes the form

$$
\begin{aligned}
Det\,(H) &= \underbrace{-\frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+1}}\right)}_{\boxed{14.1}} - \underbrace{\frac{2\lambda^3\epsilon^2}{m^5(s+2)}\left(\frac{e_{s-1}}{e_s^2 e_{s+2}}\right)}_{\boxed{15.1}} \\
&\quad - \frac{\lambda^4\epsilon^2}{m^6(s+1)^2(s+2)}\left(\frac{e_s^2}{e_{s+1}^4}\right) + c_3
\end{aligned}
$$

$$
\begin{aligned}
&+\underbrace{\frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_s}{e_{s+1}^3}\right)}_{\boxed{14.2}}+\underbrace{\frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}}{e_s e_{s+1}^2}\right)}_{\boxed{15.2}}\\
&+\underbrace{\frac{2\lambda^2\epsilon^2}{m^4(s+1)(s+2)}\left(\frac{1}{e_{s+1}^2}\right)+c_5}_{\boxed{13.5}}+\frac{\lambda^4\epsilon^2}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right)\\
&+\frac{\epsilon^2}{m^3}\left(\frac{1}{e_s^3}\right)[sme_s-2\lambda e_{s-1}]\\
&-\frac{\lambda^2\boxed{\epsilon^3}}{m^4}\left(\frac{e_{s-1}}{e_s^4}\right)+\frac{s\lambda\boxed{\epsilon^3}}{m^3}\left(\frac{1}{e_s^3}\right)-\frac{2\lambda^2\boxed{\epsilon^3}s}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}}\right)\\
&-\frac{\lambda^4\boxed{\epsilon^3}}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}}{e_s^3 e_{s+2}}\right)+\frac{\lambda^3\boxed{\epsilon^3}s}{m^5(s+1)(s+2)}\left(\frac{1}{e_s^2 e_{s+2}}\right)
\end{aligned}
$$

Adding (14.1) and (14.2) we obtain

$$
\begin{aligned}
(14.1)+(14.2) &= -\frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+1}}\right)+\frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_s}{e_{s+1}^3}\right)\\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}}\right)\left[-\frac{e_{s-1}^2}{e_s^3}+\frac{e_s}{e_{s+1}^2}\right]\\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)\left[-e_{s+1}^2 e_{s-1}^2+e_s^4\right]\\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)\left[-\left(e_s+\epsilon_{s+1}\right)^2\left(e_s-\epsilon_s\right)^2+e_s^4\right]\\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)\left[-\left(e_s^2+2e_s\epsilon_{s+1}+\epsilon_{s+1}^2\right)\left(e_s^2-2e_s\epsilon_s+\epsilon_s^2\right)+e_s^4\right]\\
&= \frac{2\lambda^3\epsilon_s^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)\{2e_s^3\epsilon_s-e_s^2\epsilon_s^2-2e_s^3\epsilon_{s+1}\\
&\quad +4e_s^2\epsilon_{s+1}\epsilon_s-\epsilon_{s+1}^2 e_s^2+2e_s\epsilon_{s+1}^2\epsilon_s-\epsilon_{s+1}^2\epsilon_s^2\}\\
&= \frac{2\lambda^3\epsilon_s^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)\{2e_s^2\left(e_s\epsilon_s-\frac{\epsilon_s^2}{2}-e_s\epsilon_{s+1}\right)\\
&\quad +4e_s^2\epsilon_{s+1}\epsilon_s-\epsilon_{s+1}^2 e_s^2+2e_s\epsilon_{s+1}^2\epsilon_s-\epsilon_{s+1}^2\epsilon_s^2\}\\
&= \frac{2\lambda^3\epsilon_s^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)\{2e_s^2\left(e_s\left(\frac{(s+1)m}{\lambda}\epsilon_{s+1}-\epsilon_{s+1}\right)-\frac{\epsilon_s^2}{2}\right)\\
&\quad +4e_s^2\epsilon_{s+1}\epsilon_s-\epsilon_{s+1}^2 e_s^2+2e_s\epsilon_{s+1}^2\epsilon_s-\epsilon_{s+1}^2\epsilon_s^2\}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{2\lambda^3\epsilon_s^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)\{2e_s^2\left(e_s\left(\frac{(s+1)m}{\lambda}-1\right)\epsilon_{s+1}-\frac{\epsilon_s^2}{2}\right) \\
&\quad +4e_s^2\epsilon_{s+1}\epsilon_s-\epsilon_{s+1}^2e_s^2+2e_s\epsilon_{s+1}^2\epsilon_s-\epsilon_{s+1}^2\epsilon_s^2\} \\
&= \frac{2\lambda^3\epsilon_s^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)\{2e_s^2\left(e_s\left(\frac{(s+1)m}{\lambda}-1\right)\epsilon_{s+1}-\frac{\epsilon_s^2}{2}\right) \\
&\quad +4e_s^2\epsilon_{s+1}\epsilon_s-\epsilon_{s+1}^2e_s^2+2e_s\epsilon_{s+1}^2\epsilon_s-\epsilon_{s+1}^2\epsilon_s^2\} \\
&= \frac{2\lambda^3\epsilon_s^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)k_1
\end{aligned}
$$

Note that k_1 is positive since

$$
\begin{aligned}
2e_s^2\left(e_s\left(\frac{sm+m}{\lambda}-1\right)\epsilon_{s+1}-\frac{\epsilon_s^2}{2}\right) &> 2e_s^2\left(e_s\left(\frac{m}{\lambda}\right)\epsilon_{s+1}-\frac{\epsilon_s^2}{2}\right) \text{ since } \frac{sm}{\lambda}>1, \\
&> 2e_s^2\left(e_s\left(\frac{m}{\lambda}\right)\epsilon_{s+1}-\frac{\epsilon_s^2}{2}\right) \\
&> 0
\end{aligned}
$$

since $\frac{2m}{\lambda}e_s\epsilon_{s+1} > \epsilon_s^2$ by direct computation. Clearly

$$
\begin{aligned}
4e_s^2\epsilon_{s+1}\epsilon_s-\epsilon_{s+1}^2e_s^2+2e_s\epsilon_{s+1}^2\epsilon_s-\epsilon_{s+1}^2\epsilon_s^2 &= e_s^2\epsilon_{s+1}\left(4\epsilon_s-\epsilon_{s+1}\right)+\epsilon_{s+1}^2\left(2e_s\epsilon_s-\epsilon_s^2\right) \\
&> 0.
\end{aligned}
$$

Adding (15.1) to (15.2) we obtain

$$
\begin{aligned}
(15.1)+(15.2) &= -\frac{2\lambda^3\epsilon^2}{m^5(s+2)}\left(\frac{e_{s-1}}{e_s^2e_{s+2}}\right)+\frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{e_{s-1}}{e_se_{s+1}^2}\right) \\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+1)(s+2)}\left(\frac{e_{s-1}}{e_s}\right)\left[-(s+1)\frac{1}{e_se_{s+2}}+(s+2)\frac{1}{e_{s+1}^2}\right] \\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+1)(s+2)}\left(\frac{e_{s-1}}{e_s}\right)\left[-(s+1)\frac{e_{s+1}^2}{e_{s+1}^2e_se_{s+2}}+(s+1)\frac{e_se_{s+2}}{e_se_{s+2}e_{s+1}^2}\right] \\
&\quad +\frac{2\lambda^3\epsilon^2}{m^5(s+1)(s+2)}\left(\frac{e_{s-1}}{e_s}\right)\left(\frac{1}{e_{s+1}^2}\right) \\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+2)}\left(\frac{e_{s-1}}{e_s}\right)\left[-\frac{e_{s+1}^2}{e_se_{s+2}e_{s+1}^2}+\frac{e_se_{s+2}}{e_se_{s+2}e_{s+1}^2}\right] \\
&\quad +\frac{2\lambda^3\epsilon^2}{m^5(s+1)(s+2)}\left(\frac{e_{s-1}}{e_se_{s+1}^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{2\lambda^3\epsilon^2}{m^5(s+2)}\left(\frac{e_{s-1}}{e_s}\right)\left(\frac{1}{e_se_{s+2}e_{s+1}^2}\right)\left[-e_{s+1}^2+e_se_{s+2}\right] \\
&\quad +\frac{2\lambda^3\epsilon^2}{m^5(s+1)(s+2)}\left(\frac{e_{s-1}}{e_se_{s+1}^2}\right) \\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+2)}\left(\frac{e_{s-1}}{e_s}\right)\left(\frac{1}{e_se_{s+2}e_{s+1}^2}\right)\left[-e_{s+1}^2+(e_{s+1}-\epsilon_{s+1})(e_{s+1}+\epsilon_{s+2})\right] \\
&\quad +\frac{2\lambda^3\epsilon^2}{m^5(s+1)(s+2)}\left(\frac{e_{s-1}}{e_se_{s+1}^2}\right) \\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+2)}\left(\frac{e_{s-1}}{e_s}\right)\left(\frac{1}{e_se_{s+2}e_{s+1}^2}\right)\left[e_{s+1}\epsilon_{s+2}-\epsilon_{s+1}e_{s+1}-\epsilon_{s+1}\epsilon_{s+2}\right] \\
&\quad +\frac{2\lambda^3\epsilon^2}{m^5(s+1)(s+2)}\left(\frac{e_{s-1}}{e_se_{s+1}^2}\right) \\
&= \frac{2\lambda^3\epsilon^2}{m^5(s+2)}\left(\frac{e_{s-1}}{e_s^2e_{s+2}e_{s+1}^2}\right)k_2+\frac{2\lambda^3\epsilon^2}{m^5(s+1)(s+2)}\left(\frac{e_{s-1}}{e_se_{s+1}^2}\right)
\end{aligned}
$$

where clearly $k_2 < 0$. Therefore

$$
\begin{aligned}
Det(H) &= \underbrace{\frac{2\lambda^3\epsilon_s^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3e_s^3}\right)k_1}_{\boxed{\text{Remaining from 14}}}\underbrace{-\frac{\lambda^4\epsilon^2}{m^6(s+1)^2(s+2)}\left(\frac{e_s^2}{e_{s+1}^4}\right)}_{\boxed{9.1}}+c_3 \\
&\quad \underbrace{+\frac{2\lambda^3\epsilon^2}{m^5(s+2)}\left(\frac{e_{s-1}}{e_s^2e_{s+2}e_{s+1}^2}\right)k_2+\frac{2\lambda^3\epsilon^2}{m^5(s+1)(s+2)}\left(\frac{e_{s-1}}{e_se_{s+1}^2}\right)}_{\boxed{\text{Remaining from 15.2}}} \\
&\quad +\frac{2\lambda^2\epsilon^2}{m^4(s+1)(s+2)}\left(\frac{1}{e_{s+1}^2}\right)+c_5+\underbrace{\frac{\lambda^4\epsilon^2}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}^2}{e_s^3e_{s+2}}\right)}_{\boxed{9.2}} \\
&\quad +\frac{\epsilon^2}{m^3}\left(\frac{1}{e_s^3}\right)\left[sme_s-2\lambda e_{s-1}\right] \\
&\quad -\frac{\lambda^2\boxed{\epsilon^3}}{m^4}\left(\frac{e_{s-1}}{e_s^4}\right)+\frac{s\lambda\boxed{\epsilon^3}}{m^3}\left(\frac{1}{e_s^3}\right)-\frac{2\lambda^2\boxed{\epsilon^3}s}{m^4(s+1)}\left(\frac{1}{e_s^2e_{s+1}}\right) \\
&\quad -\frac{\lambda^4\boxed{\epsilon^3}}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}}{e_s^3e_{s+2}}\right)+\frac{\lambda^3\boxed{\epsilon^3}s}{m^5(s+1)(s+2)}\left(\frac{1}{e_s^2e_{s+2}}\right)
\end{aligned}
$$

Adding (9.1) and (9.2) gives

$$
\begin{aligned}
(9.1)+(9.2) &= -\frac{\lambda^4\epsilon^2}{m^6(s+1)^2(s+2)}\left(\frac{e_s^2}{e_{s+1}^4}\right)+\frac{\lambda^4\epsilon^2}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right)\\
&= \frac{\lambda^4\epsilon^2}{m^6(s+1)^2(s+2)}\left(\frac{1}{e_s^3 e_{s+2} e_{s+1}^4}\right)\left(-e_s^5 e_{s+2}+(s+1)\,e_{s-1}^2 e_{s+1}^4\right)\\
&> 0
\end{aligned}
$$

because $e_{s+1}^4 > e_s^4$ and $(s+1)\,e_{s-1}^2 > e_s e_{s+2}$. Therefore

$$
\begin{aligned}
Det\,(H) &> \frac{2\lambda^3\epsilon^2}{m^5(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)k_1+\frac{2\lambda^3\epsilon^2}{m^5(s+2)}\left(\frac{e_{s-1}}{e_s^2 e_{s+2} e_{s+1}^2}\right)k_2\\
&\quad+\frac{2\lambda^3\epsilon^2}{m^5(s+1)(s+2)}\left(\frac{e_{s-1}}{e_s e_{s+1}^2}\right)+\frac{2\lambda^2\epsilon^2}{m^4(s+1)(s+2)}\left(\frac{1}{e_{s+1}^2}\right)\\
&\quad+c_5+\frac{\lambda^4\epsilon^2}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}^2}{e_s^3 e_{s+2}}\right)+\frac{\epsilon^2}{m^3}\left(\frac{1}{e_s^3}\right)\left[sme_s-2\lambda e_{s-1}\right]\\
&\quad\underbrace{-\frac{\lambda^2\boxed{\epsilon^3}}{m^4}\left(\frac{e_{s-1}}{e_s^4}\right)}_{\boxed{16.2}}+\underbrace{\frac{s\lambda\boxed{\epsilon^3}}{m^3}\left(\frac{1}{e_s^3}\right)}_{\boxed{16.1}}-\frac{2\lambda^2\boxed{\epsilon^3}s}{m^4(s+1)}\left(\frac{1}{e_s^2 e_{s+1}}\right)\\
&\quad\underbrace{-\frac{\lambda^4\boxed{\epsilon^3}}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}}{e_s^3 e_{s+2}}\right)}_{\boxed{17.1}}+\underbrace{\frac{\lambda^3\boxed{\epsilon^3}s}{m^5(s+1)(s+2)}\left(\frac{1}{e_s^2 e_{s+2}}\right)}_{\boxed{17.2}}.
\end{aligned}
$$

Note that

$$
\begin{aligned}
(16.1)+(16.2) &= \frac{s\lambda\boxed{\epsilon^3}}{m^3}\left(\frac{1}{e_s^3}\right)-\frac{\lambda^2\boxed{\epsilon^3}}{m^4}\left(\frac{e_{s-1}}{e_s^4}\right)\\
&= \frac{\lambda^2\boxed{\epsilon^3}}{m^4}\frac{1}{e_s^4}\left(\frac{sm}{\lambda}e_s-e_{s-1}\right)>0
\end{aligned}
$$

and

$$
\begin{aligned}
(17.1)+(17.2) &= -\frac{\lambda^4\boxed{\epsilon^3}}{m^6(s+1)(s+2)}\left(\frac{e_{s-1}}{e_s^3 e_{s+2}}\right)+\frac{\lambda^3\boxed{\epsilon^3}s}{m^5(s+1)(s+2)}\left(\frac{1}{e_s^2 e_{s+2}}\right)\\
&= \frac{\lambda^4\boxed{\epsilon^3}}{m^6(s+1)(s+2)}\left(\frac{1}{e_s^3 e_{s+2}}\right)\left[e_{s-1}+\frac{sm}{\lambda}e_s\right]\\
&= \frac{\lambda^4\boxed{\epsilon^3}}{m^6(s+1)(s+2)}\left(\frac{1}{e_s^3 e_{s+2}}\right)\left[-e_{s-1}+\frac{sm}{\lambda}e_{s-1}+\frac{sm}{\lambda}\epsilon_s\right]\\
&> \frac{\lambda^4\boxed{\epsilon^3}}{m^6(s+1)(s+2)}\left(\frac{1}{e_s^3 e_{s+2}}\right)\left(\frac{sm}{\lambda}\epsilon_s\right)>0 \text{ since } \frac{sm}{\lambda}>1.
\end{aligned}
$$

Therefore

$$
\begin{aligned}
Det\,(H) \quad > \quad & \frac{2\lambda^3\epsilon_s^2}{m^5\,(s+1)}\left(\frac{1}{e_{s+1}^3 e_s^3}\right)k_1 + \frac{2\lambda^3\epsilon^2}{m^5\,(s+2)}\left(\frac{e_{s-1}}{e_s^2 e_{s+2} e_{s+1}^2}\right)k_2 \\
& + \frac{2\lambda^3\epsilon^2}{m^5\,(s+1)\,(s+2)}\left(\frac{e_{s-1}}{e_s e_{s+1}^2}\right) + \frac{2\lambda^2\epsilon^2}{m^4\,(s+1)\,(s+2)}\left(\frac{1}{e_{s+1}^2}\right) \\
& + c_5 + \frac{\epsilon^2}{m^3}\left(\frac{1}{e_s^3}\right)\left[sme_s - 2\lambda e_{s-1}\right] - \frac{2\lambda^2\epsilon^3 s}{m^4\,(s+1)}\left(\frac{1}{e_s^2 e_{s+1}}\right)
\end{aligned}
$$

Because of the complicated nature of this inequality we obtain numerical results corresponding to the determinant of the mixed Hessian matrix.

Appendix B

MIXED CONVEXITY OF THE ERLANG DELAY FORMULA

In this section we determine the determinant of the mixed Hessian matrix corresponding to the Erlang delay formula

$$E_d(s,\mu) = \frac{1}{(s-1)!\sum_{i=0}^{s-1} \frac{(s-i)}{i!} \left(\frac{1}{s\rho}\right)^{s-i}}.$$

Next we rewrite the Erlang delay formula as a function of the service rate μ and the number of servers s by using the equality $\rho = \frac{\lambda}{s\mu}$.

$$\begin{aligned} E_d(s,\mu) &= \frac{1}{(s-1)!\sum_{i=0}^{s-1} \frac{(s-i)}{i!} \left(\frac{1}{s\frac{\lambda}{s\mu}}\right)^{s-i}} \\ &= \frac{1}{(s-1)!\sum_{i=0}^{s-1} \frac{(s-i)}{i!} \left(\frac{\mu}{\lambda}\right)^{s-i}} \\ &= \frac{1}{(s-1)!\sum_{i=0}^{s-1} \frac{(s-i)}{i!} \left(\frac{\mu}{\lambda}\right)^{s-i}} \end{aligned}$$

Replacing i with $s-i-1$ we have

$$\begin{aligned} E_d(s,\mu) &= \frac{1}{(s-1)!\sum_{i=0}^{s-1} \frac{(s-i)}{i!} \left(\frac{\mu}{\lambda}\right)^{s-i}} \\ &= \frac{1}{(s-1)!\sum_{s-i-1=0}^{s-i-1=s-1} \frac{(s-(s-i-1))}{(s-i-1)!} \left(\frac{\mu}{\lambda}\right)^{s-(s-i-1)}} \\ &= \frac{1}{(s-1)!\sum_{i=0}^{s-1} \frac{(i+1)}{(s-i-1)!} \left(\frac{\mu}{\lambda}\right)^{i+1}} \end{aligned}$$

Let $j = i + 1$ therefore

$$E_d(s,\mu) = \frac{1}{(s-1)!\sum_{j=1}^{s} \frac{j}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^j}$$

The first difference and the differential of the Erlang delay formula are

$$\begin{aligned}
\nabla_1 E_d &= \frac{1}{s!\sum_{j=1}^{s+1} \frac{j}{(s+1-j)!}\left(\frac{\mu}{\lambda}\right)^j} - \frac{1}{(s-1)!\sum_{j=1}^{s} \frac{j}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^j} \\
\frac{dE_d}{d\mu} &= \frac{-1}{(s-1)!}\left[\sum_{j=1}^{s} \frac{j}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^j\right]^{-2}\left[\sum_{j=1}^{s} \frac{j^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-1}\left(\frac{1}{\lambda}\right)\right] \\
&= \frac{-1}{\lambda(s-1)!}\left[\sum_{j=1}^{s} \frac{j}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^j\right]^{-2}\left[\sum_{j=1}^{s} \frac{j^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-1}\right]
\end{aligned}$$

and the Hessian matrix components are

$$\nabla_{11} E_d = \frac{1}{(s+1)!\sum_{j=1}^{s+2} \frac{j}{(s+2-j)!}\left(\frac{\mu}{\lambda}\right)^j} - 2\frac{1}{s!\sum_{j=1}^{s+1} \frac{j}{(s+1-j)!}\left(\frac{\mu}{\lambda}\right)^j} + \frac{1}{(s-1)!\sum_{j=1}^{s} \frac{j}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^j},$$

$$\begin{aligned}
\frac{d^2E_d}{d\mu^2} &= \frac{2}{\lambda^2(s-1)!}\left[\sum_{j=1}^{s} \frac{j}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^j\right]^{-3}\left[\sum_{j=1}^{s} \frac{j^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-1}\right]^2 \\
&\quad - \frac{1}{\lambda^2(s-1)!}\left[\sum_{j=1}^{s} \frac{j}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^j\right]^{-2}\left[\sum_{j=1}^{s} \frac{j(j-1)^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-2}\right] \\
&= \frac{1}{\lambda^2(s-1)!}\left[\sum_{j=1}^{s} \frac{j}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^j\right]^{-3}\{2\left[\sum_{j=1}^{s} \frac{j^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-1}\right]^2 \\
&\quad - \left[\sum_{j=1}^{s} \frac{j^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-1}\right]\left[\sum_{j=2}^{s} \frac{j(j-1)^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-2}\right]\},
\end{aligned}$$

and

$$
\begin{aligned}
\frac{d}{d\mu}\nabla_1 E_d &= \frac{d}{d\mu}\left(\frac{1}{s!\sum\limits_{j=1}^{s+1}\frac{j}{(s+1-j)!}\left(\frac{\mu}{\lambda}\right)^j} - \frac{1}{(s-1)!\sum\limits_{j=1}^{s}\frac{j}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^j}\right) \\
&= \frac{-1}{\lambda s!}\left[\sum_{j=1}^{s+1}\frac{j}{(s+1-j)!}\left(\frac{\mu}{\lambda}\right)^j\right]^{-2}\left[\sum_{j=1}^{s+1}\frac{j^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-1}\right] \\
&\quad + \frac{1}{\lambda(s-1)!}\left[\sum_{j=1}^{s}\frac{j}{(s+1-j)!}\left(\frac{\mu}{\lambda}\right)^j\right]^{-2}\left[\sum_{j=1}^{s}\frac{j^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-1}\right]
\end{aligned}
$$

Let

$$
\begin{aligned}
A &= (s+1)!\sum_{j=1}^{s+2}\frac{j}{(s+2-j)!}\left(\frac{\mu}{\lambda}\right)^j \\
B &= s!\sum_{j=1}^{s+1}\frac{j}{(s+1-j)!}\left(\frac{\mu}{\lambda}\right)^j \\
C &= (s-1)!\sum_{j=1}^{s}\frac{j}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^j \\
D &= \sum_{j=1}^{s}\frac{j^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-1} \\
E &= \sum_{j=2}^{s}\frac{j\,(j-1)^2}{(s-j)!}\left(\frac{\mu}{\lambda}\right)^{j-2} \\
F &= \sum_{j=1}^{s+1}\frac{j^2}{(s+1-j)!}\left(\frac{\mu}{\lambda}\right)^{j-1}
\end{aligned}
$$

By using this notation we have the Hessian matrix components as follows:

$$
\begin{aligned}
\frac{d^2 E_d}{d\mu^2} &= \frac{1}{\lambda^2(s-1)!}C^{-3}\left[2D^2 - CE\right] \\
\nabla_{11}E_d &= \frac{1}{(s+1)!}(ABC)^{-1}\left[BC - 2(s+1)AC + s(s+1)AB\right] \\
\frac{d}{d\mu}\nabla_1 E_d &= \frac{1}{\lambda s!}B^{-2}C^{-2}\left(-C^2F + sB^2D\right)
\end{aligned}
$$

Therefore the determinant of the mixed Hessian matrix corresponding to the Erlang delay formula is as follows:

$$\begin{aligned}
\det H &= \frac{1}{\lambda^2 (s-1)!} C^{-3} \left[2D^2 - CE\right] \frac{1}{(s+1)!} (ABC)^{-1} \left[BC - 2(s+1)AC + s(s+1)AB\right] \\
&\quad - \left[\frac{1}{\lambda s!} B^{-2} C^{-2} \left(-C^2 F + sB^2 D\right)\right]^2 \\
&= \frac{A^{-1}B^{-1}C^{-4}}{\lambda^2 (s+1)! (s-1)!} \left(2D^2 - CE\right) \left[BC - 2(s+1)AC + s(s+1)AB\right] \\
&\quad - \frac{B^{-4}C^{-4}}{\lambda^2 (s!)^2} \left[C^4 F^2 - 2sC^2 B^2 FD + s^2 B^4 D^2\right] \\
&= \frac{A^{-1}B^{-1}C^{-4}}{\lambda^2 s! (s+1)!} \{s \left(2D^2 B^3 - CEB^3\right) \left(BC - 2(s+1)AC + s(s+1)AB\right) \\
&\quad - (s+1) \left[AC^4 F^2 - 2sAC^2 B^2 FD + s^2 AB^4 D^2\right]\}
\end{aligned}$$

Letting

$$K = \frac{A^{-1}B^{-1}C^{-4}}{\lambda^2 s! (s+1)!}$$

we have

$$\begin{aligned}
\det H &= K\{2sB^4CD^2 - 4s(s+1)AB^3CD^2 + 2s^2(s+1)AB^4D^2 \\
&\quad -sB^4C^2E + 2s(s+1)AB^3C^2E - s^2(s+1)AB^4CE \\
&\quad - (s+1)AC^4F^2 + 2s(s+1)AB^2C^2DF - s^2(s+1)AB^4D^2\}
\end{aligned}$$

In this case the only two terms that we can cancel out are $2s^2(s+1)AB^4D^2$ and $s^2(s+1)AB^4D^2$. Therefore

$$\begin{aligned}
\det H &= K\{2sB^4CD^2 - 4s(s+1)AB^3CD^2 + s^2(s+1)AB^4D^2 \\
&\quad -sB^4C^2E + 2s(s+1)AB^3C^2E - s^2(s+1)AB^4CE \\
&\quad - (s+1)AC^4F^2 + 2s(s+1)AB^2C^2DF\}
\end{aligned}$$

The complex structure of the determinant is making it difficult to determine the determinant to be positive therefore we observed the special cases in chapter 8.