Limit theorems leading to Bose-Einstein, Maxwell-Boltzmann Statistics and Zipf-Mandelbrot Law

by

Tomasz Michal Lapinski

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Department of Statistics

October 2013

THE UNIVERSITY OF WARWICK

Contents

Declarations

I hereby declare that this thesis is the result of my own work and research, except where otherwise indicated. This thesis has not been submitted for examination to any institution other than the University of Warwick.

Signed:

Tomasz Lapinski

28th October 2013

Acknowledgments

I would like to dedicate special thanks to my supervisor Prof. Vassili Kolokoltsov for the ideas, patience and supervision during the time of my PhD. I would also like to thank my personal tutor John Warren and other staff at the Department for their support and the enjoyable times I have had.

For their support and faith in me I send special thanks to my parents and friends. During the time of this PhD I have had a very nice and supportive companionship and for that I thank my office mates and colleagues. I very much appreciate tea time with Duy, lunch with Helen and parties and pleasant times with other Warwick friends.

Tomasz Lapinski

28th October 2013

Abstract

In this thesis we develop the ideas introduced by V.P. Maslov in [9], [10] and [11], the new limit theorem which leads to Bose-Einstein, Maxwell-Boltzmann Statistics and Zipf-Mandelbrot Law. We independently constructed the proof for the theorem, based on Statistical Mechanics methodology, but with precise and rigorous estimates and rate of convergence. The proof involves approximation of the considered entropy, the partition function and specific Laplace type integral approximation which we had to develop specifically for this result. The proof also involved several minor estimates and approximations that are included in the work and the mathematical preliminaries which we used are attached in the appendix. In addition, we provide a step by step introduction to the underlying mathematical setting. Within the theorem we separated two cases of resulting distribution, this separation was mentioned in [11] however it was not developed further in that paper. The first case gives known distributions which are in the thesis title. Additionally, we construct two new fluctuation theorems with proof based on the proof of the main theorem. In terms of the application, we found that developed theory can be applied in the field of Econophysics. Based on the paper by F.Kusmartsev [16], we inferred that presented three distribution may correspond to the state of the economy of particular countries. Unified underlying framework might reflect the fact that these economies have one common structure.

List of Figures

1 Introduction

The main purpose of this thesis is to present new mathematical results related to Physics and indicate some possible applications within various scientific disciplines. As the title states, these new results are the limit theorems. The branch of Mathematics which deals with limit theorems is Probability Theory, hence these are probabilistic results. Further, as stated in the title, the outcomes of limit theorems, Bose-Einstein and Maxwell-Boltzmann statistics, are common distributions in the major field of Physics, Statistical Mechanics. The Zipf-Madlebort Law, which is the third outcome of the theorems, is a power law widely occurring in the Science of Complex Systems. Hence, to be more precise, this thesis is about a new result of Probability Theory related to Statistical Mechanics and Complexity Science.

In this introduction we provide an extensive background for the theorems. We include a broad literature review of the existing ideas of where the theorem originated from. We provide a short historical background of the fields and branches in which the theorems have fundaments. What is more, we describe in detail the particular results which are common and occur in this thesis. We also include a depiction of their development on the historical timeline. The last part of the introduction is an outline of the structure of the whole thesis.

The introduction chapter is structured into four sections. The first section is about the origin of the idea of the theorems. It is mostly a review of several papers by Prof. V.P. Maslov which seeded this idea and a short introduction of the author. The next section is about Statistical Physics. We provide historical outline of this field, we underline the significance of Thermodynamics in its development and other important historical facts and scientific achievements. We include a brief history of Bose-Einstain and Maxwell-Boltzmann statistics, together with their derivations, and which are common to physicists. The third section concerns Complexity Science. We begin with a little history of how this discipline evolved over time. Then we explain the emergence of power laws, in particular the Zipf-Mandelbrot Law. We will give the vast examples of power law systems to underline its significance in the real world. In the last section, the full outline of the thesis will be provided. We shortly describe the chapters which the thesis consist of and include some interplay between these chapters too.

1.1 Inspiration from the work of Prof. V.P. Maslov, Literature review

Viktor Pavlovich Maslov is a Professor at Lomonosov Moscow State University. He is a specialist in the field of mathematical physics but his research spreads over various branches of mathematical and natural sciences, particularly quantum theory, asymptotic analysis, operator theory and nanotechnology.
He has gained recognition as a scientist who has a grasp in uncovering mathematics behind various phenomena from physics and other natural sciences.
An example here can be his development of the first formal mathematical description of a nanostructure, which resulted in the introducing of an object called Lagrangian submanifold. V.P. Maslov is also known for the introduction of a Maslov index.
A peer-review journal Mathematical Notes, which is a translation of Matematicheskie Zametki, is the main mathematical journal of the Russian Academy of Science. Prof. V.P. Maslov is its editor-in-chief and there he publishes some of his findings. Among many branches of mathematics, one can find works published in number theory, functional analysis, topology, probability, operator and group theory, asymptotic and approximation methods spectral theory and other fields. Most of the publications which are fundamental for our work were released in this journal. For more information about V.P.Maslov see [6].

Here we will review four of his papers. The first paper 'Nonlinear averaging axioms in financial mathematics and stock price dynamics', provided some background to the nonlinear averages introduced by Maslov in economics and their connection to Statistical Mechanics. Then, in 'Nonlinear averaging in Economics' an extension of this nonlinear average to a more general context than economics is provided and a more explicit connection with statistical physics is given. The convergence of nonlinear average to Bose-Einstein statistics is also introduced. These findings are placed in the form of the limit theorem with drafts of the proof.

Finally in the third paper 'On a General Theorem of Set theory leading to the Gibbs, Bose-Einstain and Pareto Distributions as well as to the Zipf-Mandelbrot Law for the Stock Market' the nonlinear average is further generalised. This generalisation is of a mathematical nature. Instead of convergence to one statistics, i.e. Bose-Einstein, we have convergence to three, two others are Gibbs type and Pareto distribution. Obtaining one of three averages is determined by choice of some parameter. Our work is an extension and development of the findings of this paper. In the last paper of V.P.Maslov that we review, 'On Zipf's Law and Rank Distributions in Linguistics and Semiotics', he first underlines the significance of Zipf Law, recalls its origins and then introduces a new framework for how to model various systems with Zipf related laws. For us this paper was significant as it showed the generality of Zipf Law and related distributions in nature. Given that in previous paper the mathematical derivation of Zipf Law was given, exploring those various system modelled by Zipf Law was even more inspiring.

Review of Nonlinear averaging aximos in financial mathematics and stock price dynamics

First we consider the paper 'Nonlinear averaging axioms in financial mathematics and stock price dynamics' [10]. The author begins with an introduction to the certain type of nonlinear average and supports the fact of nonlinearity with two examples. In calculating the individual 'natural' capital, one has to consider many factors. One common way is to consider a credit which can be given to particular individuals and this depends on many factors. These factors can be regular income, employment status, age, number of dependencies , credit history and others. Obviously, the person's capital is not a linear dependence of possessed money and income.
Another example of nonlinear averaging occuring naturally is the stockholder's ability to influence the company, i.e. 51 percent of stock gives the right to decide 50 not. We see that the percentage of stocks possessed is not a linear dependence with ability to influence the company.
Further, the axioms of nonlinear averaging are introduced. He considers the

avarage of the form

$$y = f^{-1}\left(\sum_i \alpha_i f(x_i)\right),$$
$$x_i = \sum_{j=1}^{G} \lambda_j N_j,$$

where f is come convex function, α_i are weight factors and y is a nonlinear average of the incomes x_i.
Additionally the income x_i is composed of the incomes from G assets, each corresponding to outcomes λ_j and quantity of money N_j. We also have that $\sum_{j=1}^{G} N_j = N$ and N is the total amount of money invested.
Furthermore the 'degenerations' are included, i.e. there are G_1 same outcome λ_1 over which capital is redistributed and also G_2 of λ_2, and so on. Hence x_i are equal

$$x_i = \lambda_1 \sum_{j=1}^{G_1} N_j + \lambda_2 \sum_{j=G_1+1}^{G} N_j,$$

for two different outcomes λ_1, λ_2 only.
The axioms from the paper are the following

- Axiom 1 states that when there is only one income x_i then average simply becomes this income.
- Axiom 2 restricts that the coefficients α_i are independent of λ_j.
- Axiom 3 defines that two notes of money are indistinguishable.
- Axiom 4 states if we add some value ω to all λ_j then income x_i will increase by the same value $N\omega$.

The author applies these axioms to calculate the function f and weights α_i, this leads to the 'financial averaging formula' for two outcomes λ_1 and λ_2

$$y = \frac{1}{\beta}\log\left(\frac{(G-1)!N!}{(N+G-1)!}\sum_{N_1=0}^{N}\frac{(G_1+N_1-1)!}{(G_1-1)!N_1!}\frac{(G_2+N_2-1)!}{(G_2-1)!N_2!}\exp\Big(\beta(\lambda_1 N_1+\lambda_2 N_2)\Big)\right), \tag{1.1}$$

where

$$\alpha_i = \alpha_{N1} = \frac{(G-1)!N!}{(N+G-1)!}\frac{(G_1+N_1-1)!}{(G_1-1)!N_1!}\frac{(G_2+N_2-1)!}{(G_2-1)!N_2!}$$

and $N_2 = N - N_1$, $G_2 = G - G_1$.
It turns out that Axiom 3 about the indistinguishability of notes corresponds to assumptions about bosons in the Bose-Einstein statistics and the coefficients α_i correspond to a number of possible redistributions of N_1 boson particles over energy level with G_1 degenerations and N_2 bosons over G_2 degenerations.
What is more, the exponent function f and constant β correspond to the Gibbs factor.
As an example of such averaging, Prof. Maslov considers two groups of financial institutions. The first group gives return λ_1 and there are G_1 institutions in this group. The second provides outcome λ_2 and there are G_2 of them. Additionally, money deposited in the first group is subject to taxation proportional to the square of money deposited, while depositors of the second group get a subsidy which is also proportional to the square of money put in the second group institutions. Hence the income x_i is equal

$$x_i = \lambda_1 N_1 + \lambda_2 N_2 - \frac{V_1 N_1^2}{2N} + \frac{V_2 N_2^2}{2N},$$

where V_1, V_2 are constants corresponding to taxation and subsidy, and N_1 is money put in the first group and N_2 into second. The value of N_2 can be expressed via N_1, i.e. $N_2 = N - N_1$, then the 'financial averaging formula' (1.1) can be expressed as

$$y = \frac{1}{\beta} \ln \left(\sum_{N_1=0}^{N} \exp(F(N_1)) \right) \tag{1.2}$$

where $F(N_1)$ has from

$$\begin{aligned} F(N_1) =& \beta(\lambda_1 N_1 + \lambda_2 (N - N_1) - \frac{V_1 N_1^2}{2N} + \frac{V_2 (N - N_1)^2}{2N}) - \ln \frac{(n-1)!N!}{(N+n-1)!} + \ln \frac{(G_1 + N_1 - 1)!}{(G_1 - 1)!N_1!} + \\ & + \ln \frac{(G - G_1 + N - N_1 - 1)!}{(G - G_1 - 1)!(N - N_1)!}. \end{aligned}$$

Further, the author approximates $F(N_1) \approx N f(x)$ where $x = \frac{N_1}{N}$ as $N \to \infty$ with assumptions

$$\lim_{N \to \infty} \frac{G_1}{N} = g_1 > 0,$$
$$\lim_{N \to \infty} \frac{G_2}{N} = g_2 > 0,$$

and obtains

$$y = \frac{1}{\beta} \ln \left(\sum_{Nx=0}^{N} \exp(Nf(x)) \right),$$

Next the author uses method similar to Laplace approximation to find values of x which is a biggest weight in the average, i.e. maximum of $f(x)$ for large values of N.

The main conclusion of the paper is that finding several points of such maximum depend on the values of the parameter β.

Review of Nonlinear averaging in Economics

The second paper of V.P.Maslov we review is 'Nonlinear averaging in Economics', [9]. Here the author recalls the four Kolmogorov nonlinear averaging axioms. The class of functions which are obtained as a result of those axioms contains the function which was specified in nonlinear financial averaging from the previous paper. Then the fifth axiom is added and as a consequence, the class function is restricted to a function exactly the same as the one which comes from the Axioms of averaging in economy.

Further, the nonlinear average is introduced for the general case. There are n different prices and to each one corresponds number of financial instrument G_i having the price λ_i. The number of different possibilities the buyer can spend N_i amount of money in G_i number of instruments is given by the formula

$$\gamma_i(N_i) = \frac{(N_i + G_i - 1)!}{N_i!(G_i - 1)!}$$

Then $\mathcal{N} = (N_1, N_2, \dots, N_m)$ is a set corresponding to a particular allocation of money N , where $\sum_{i=1}^{n} N_i = N$. The number of different possibilities how such allocation can be done is equal

$$\gamma(\mathcal{N}) = \prod_{i=1}^{n} \gamma_i(N_i) = \prod_{i=1}^{n} \frac{(N_i + G_i - 1)!}{N_i!(G_i - 1)!}.$$

The expenditure for some particular allocation $\mathcal{N}$ is given by

$$x(\mathcal{N}) = \sum_{i}^{n} \lambda_i N_i,$$

and finally the nonlinear averaging for the general case is specified as

$$y = -\frac{1}{\beta}\ln\left(\frac{N!(G-1)!}{(N+G-1)!}\sum_{\{\mathcal{N}\}}\gamma(\mathcal{N})\exp(-\beta x(\mathcal{N}))\right).$$

where the sum is over all possible sets $\mathcal{N}$ denoted as $\{\mathcal{N}\}$ such that $\sum_{i=1}^{n} N_i = N$ and also $\sum_{i=1}^{n} G_i = G$.

Bought assets are additionally put into m groups with the index α, where the particular group has assets starting from the index i_α and ending on j_α, hence

$$i_\alpha \leq j_\alpha,\ i_{\alpha+1} = j_\alpha + 1,\ \alpha = 1, \ldots, m,\ i_1 = 1, j_m = n,$$

then we have also following

$$G_\alpha = \sum_{i=i_\alpha}^{j_\alpha} g_i, \qquad N_\alpha = \sum_{i=i_\alpha}^{j_\alpha} k_i.$$

As the author is interested in the behaviour of the average in the limit as $N \to \infty$ he makes assumptions on how the number of instruments increase as available money increases, i.e.

$$\lim_{N\to\infty}\frac{G}{N} = \widetilde{g},$$
$$\lim_{N\to\infty}\frac{G_\alpha}{N} = \widetilde{g_\alpha} > 0, \qquad \sum_{\alpha=1}^{m}\widetilde{g_\alpha} = \widetilde{g},$$
$$\lim_{N\to\infty}\frac{N_\alpha}{N} = \overline{n_\alpha} > 0, \qquad \sum \overline{n_\alpha} = 1.$$

Further, he claims that the average number of money put in certain groups is equal to

$$\overline{N}_\alpha(\beta, N) = \sum_{i=i_\alpha}^{j_\alpha}\frac{g_i}{\exp(\beta(\lambda_i + \nu)) - 1}$$

which corresponds to the number of particles on energy levels with energies λ_i and the number of level degenerations g_i in Bose-Einstein statistics.

The parameter ν is specified by the equation

$$N = \sum_{i=1}^{n}\frac{G_i}{\exp(\beta(\lambda_i + \nu)) - 1}.$$

and he introduces function $\Gamma(\beta, N)$

$$\Gamma(\beta, N) = \sum_{\{\mathcal{N}\}} \gamma(\mathcal{N}) \exp(-\beta x(\mathcal{N})) \tag{1.3}$$

Now we recall the main result of the paper, the limit theorem

Theorem 1. Let $\Delta = aN^{3/4+\delta}$ where a and $\delta < 1/3$ are some positive constants. Then for any $\epsilon > 0$, the following relation holds as $N \to \infty$

$$\begin{aligned}\frac{1}{\Gamma(\beta, N)} \sum_{\sum_{\alpha=1}^{m}(N_\alpha(\mathcal{N})-\overline{N}_\alpha(\beta,N))^2 \geq \Delta} &\gamma(\mathcal{N}) \exp(-\beta x(\mathcal{N})) = \\ =&O\Bigg(\exp\left(\frac{(1-\epsilon)a^2 N^{1/2+2\delta}}{2\widetilde{g}d}\right)\Bigg),\end{aligned}$$

where the summation is over the collection of the sets $\{\mathcal{N}\}$ such that condition $\sum_{\alpha=1}^{m}(N_\alpha(\mathcal{N}) - \overline{N}_\alpha(\beta, N))^2 \geq \Delta$ is satisfied and d is defined as

$$d = \frac{\exp(-\beta(\lambda_1 + \nu))}{(\exp(-\beta(\lambda_1 + \nu) - 1)^2}, \text{ for } \beta < 0,$$
$$d = \frac{\exp(-\beta(\lambda_n + \nu))}{(\exp(-\beta(\lambda_n + \nu) - 1)^2}, \text{ for } \beta > 0.$$

This can be put in the context of finance as the contribution to average expenditure which is the square difference from $\overline{N}_\alpha(\beta, N)$ by more than value $O(N^{3/4+\delta})$ is of exponentially small value for a sufficient large N.

The author provides draft of the proof of that theorem. It is a mixture of some methods from statistical physics and asymptotic analysis.

Review of On a General Theorem of Set theory leading to the Gibbs, Bose-Einstain and Pareto Distributions as well as to the Zipf-Mandelbrot Law for the Stock Market

Next, we review the paper 'On a General Theorem of Set theory leading to the Gibbs, Bose-Einstein and Pareto Distributions as well as to the Zipf-Mandelbrot Law for the Stock Market', [11]. The nonlinear average here is put in the broader context of sets. This time, instead of average expenditure we have a set of integers $\{\overline{N}_1, \overline{N}_2, \ldots, \overline{N}_n\}$, which are nonlinear average integers in the collection of the set

$\{\mathcal{N}\}$ such that $\mathcal{N} = (N_1, N_2, \ldots, N_m)$ and $\sum_{i=1}^{n} N_i = N$ for some integrer N. Now, let us consider the parameter s given as a limit

$$\lim_{N\to\infty} \frac{\ln N}{\ln n} = s,$$

which is a quotient of the sum of the integers $\sum_{i=1}^{n} N_i = N$ and the number of this integres itself. Depending on this parameter the author claims that the average integers $\overline{N}_i, i = 1, \ldots, m$ in the limit $N \to \infty$ are different depending on the parameter s

$$\begin{aligned}
&1)\ \overline{N}_i = e^{-\beta\lambda_i-\alpha}, && \text{for } s > 1,\\
&2)\ \overline{N}_i = \frac{1}{e^{\beta\lambda_i+\alpha} - 1}, && \text{for } s = 1,\\
&3)\ \overline{N}_i = \frac{1}{\beta\lambda_i + \alpha}, && \text{for } 0 < s < 1,
\end{aligned}$$

for $i = 1, \ldots, m$ where the parameters α and β are related to N and some parameter E by the conditions

$$\sum_{i=1}^{n} \overline{N}_i = N, \qquad \sum_{i=1}^{n} \lambda_i \overline{N}_i = E.$$

Then he considers the collection of all sets $\{\overline{N}\}$ and denote it by $\mathcal{M}$. Further he considers the subset $\mathcal{A} \subset \mathcal{M}$ such that

$$\mathcal{A} = \left\{ \{\mathcal{N}\}, \sum_{i=1}^{n} \left(N_i - \overline{N}_i \right) \leq \Delta \right\},$$

where

$$\begin{aligned}
&\Delta = \sqrt{N} \ln^{1/2+\epsilon} N && \text{for } s > 1,\\
&\Delta = \sqrt{n} \ln^{1/2+\epsilon} n && \text{for } s = 1,\\
&\Delta = \frac{N}{\sqrt{n}} \ln^{1/2+\epsilon} n && \text{for } 0 < s < 1,
\end{aligned}$$

is called a resolving power and we have the following theorem

Theorem 2. As $N \to \infty$ the following inequality holds

$$\frac{\mathcal{N}m(\mathcal{M}\backslash\mathcal{A})}{\mathcal{N}m(\mathcal{M})} \leq \frac{C}{n} + \frac{C}{N}, \tag{1.4}$$

where C is some constant and $\mathcal{N}m$ denotes the number of elements in the sets $\mathcal{M}\backslash\mathcal{A}$ and $\mathcal{M}$.

In other words, the theorem states that the contribution of the sets that differs by more than delta from the given average set is decreasing as $N \to \infty$ as $1/N$ and $1/n$.
Note, that th eabove theorem is similar to one from the previous paper, but there the author considered only the second case of the average and the contribution was exponentially small instead of $1/N$ and $1/n$.
Prof. Malsov includes the draft of the proof of that theorem in this paper. He uses there methods form analysis, asymptotic theory and statistical mechanics.
Throughout the paper the author connects the averages of the theorem with known distributions. The first one relates to Gibbs type distribution, second is Bose-Einstein statistics and the last one Pareto Distribution or Zipf-Mandlebrot Law.

On Zipf's Law and Rank Distributions in Linguistics and Semiotics

The last paper which we review is 'On Zipf's Law and Rank Distributions in Linguistics and Semiotics' [12]. In the beginning of this paper, the author introduces Zipf's Law. Taking a particular book, if one counts the occurring words in the text, takes the frequencies for occurrence of each one and orders them in descending order then one will get the relation which will be close to Zipf Law. The distance from the exact Zipf Law will vary from text to text, but for some it will be exactly Zipf.
Some mathematicians and linguists saw, through computing and the development of 'frequency dictionaries' the possibility of creating an algorithm to distinguish authorship.
However, V.P.Maslov is of the opinion that this situation with the frequency of words is not that simple. He claims that the factual frequency of particular words is actually higher than what one can count. One of the reasons is writing style, some words are omitted, some replaced, some are substituted as certain styles by default require that. Sometimes it may be because of shortcuts used in the style or

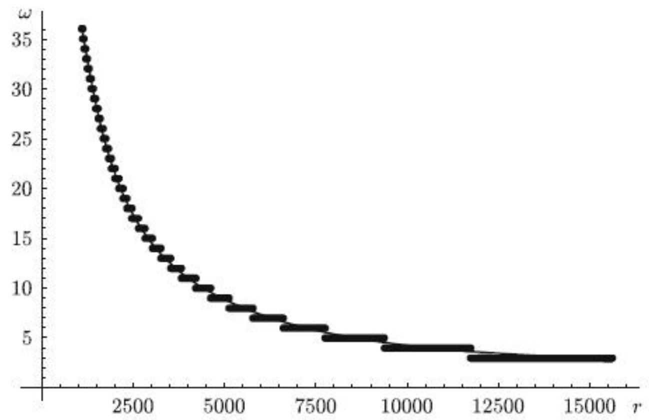


Figure 1.1: Zipf Law for the first volume of Leo Tolstoy's 'War and Peace'

meaning behind certain phrases, which might be much bigger than crude words. For a word Prof. Maslov defines this virtual frequency as

$$\widetilde{\omega}_i = \omega_i(1 + \alpha\omega_i^{\gamma}),$$

where α and γ are some parameters supposedly common to one text.
The main concept introduced in this paper is to extend the use of frequency dictionaries from just text to a more general context, which would be a sign system. Signs are of the interest of the discipline known as semiotics. In this general context, distinct words occurring in the book is a sign, its frequency is this signs cardinality and corresponding virtual frequency, virtual cardinality. The author gives various examples of the sign system, a book in the library with a given title, where the book database is a sign dictionary. Then the book requested in the database is a real cardinality, but if one adds book usages by colleagues, relatives this would be a virtual cardinality. Another example of sign could be a city. The number of people living in the city, the number from the census is a real cardinality but the number of people currently staying in the city, tourists, visitors of family, business visitors etc. is a virtual cardinality.
The example explored in more detail by the author is the prices of car brands, where the car brand is a sign and the car price is a cardinality. Then the virtual cardinality includes, with the exception of the original car price, many other expenses like insurance, gas, services and taxes.

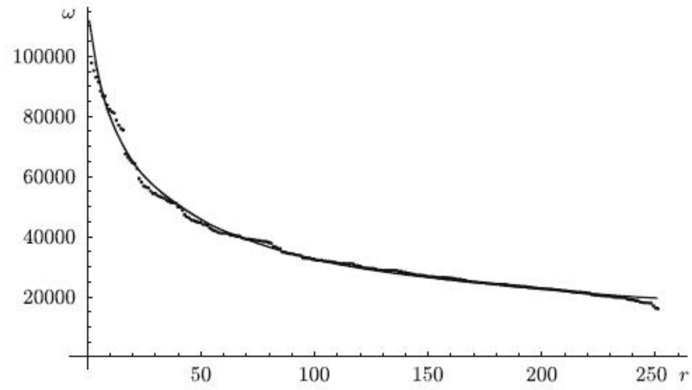


Figure 1.2: Model fit of car brands prices on American market

The next example is Japanese candles, i.e. a day to day changes of the asset prices, in the given example of some stock. The signs are the Japanese candles of particular size and the real frequency their amount. The virtual cardinality can correspond to the deal outside the stock exchange, by brokers themselves or other networks.

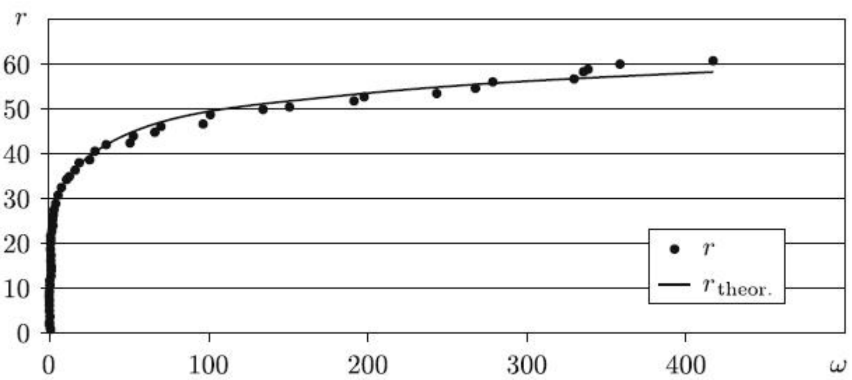


Figure 1.3: Japanese candles of the stock

1.2 Statistical Physics: Bose-Einstein and Maxwell-Boltzman statistics

This introduction is a compilation of relevant information from two books [5] and [17].

Statistical mechanics, in general, is about the systems of very many particles. Such system can be studied from two points of view: microscopic, "small scale", which is roughly the size of single atom or molecule, usually of the order 10Å and macroscopic, "large scale", where system is visible in the ordinary sense and it is of a size greater than 1 micron.

In the beginning, the "physical" homogeneous systems, such as liquids, gases or solids, have been investigated only from the macroscopic point of view, as the atomic nature of matter has not yet been well understood. Such description is based on the quantities which describe system as whole, macroscopic quantities. These quantities are related by the number of laws and together form a physics branch called 'thermodynamics". This theory was developed in consistent form by Clausius and Lord Kelvin in around 1850, and further extended by J.W. Gibbs in around 1877.

Significant progress in the understanding of matter on the microscopic level in the first half of the last century resulted in the development of quantum mechanics. Such development gave us the possibility to fully describe particles and the interaction between them on the microscopic scale.

Two theories, thermodynamics and quantum mechanics opened the way to form the theory which relates micro with macroscopic level. Statistical mechanics has emerged from their unification. It yields all the laws of thermodynamics plus a large number of relations connecting macroscopic quantities with the microscopic parameters.

1.2.1 Physical system under consideration, ideal mono-atomic gas in the equilibrium

Statistical mechanics is a broad field of physics, in the sense that it yields the results for the distinct systems consisting of variety of molecules in various states. However, one of the most common studied "class" of systems are the so-called ideal gases. Note that for ideal gases the type of molecules can vary, this can be mono or multi-atomic particles, bosons, fermions, helium, etc. Obviously this implies differences in the obtained results. Considered systems can "change significantly" in time, be time-dependent, i.e. non-equilibrium, or remain "stable", be in equilibrium. There is a different approach to obtaining results for systems which are in equilibrium and those which are non-equilibrium. Moreover, system state can vary, which can be measured by the macroscopic and microscopic quantities. Depending on the state of the system we might use classical or quantum mechanics to perform calculations. We will focus our attention on the model of ideal gas of single atom in equilibrium. Next, we will explain the above assumptions in details.

The state of the system can be described by the macro and microscopic quantities. For the gas the quantities which describe it as whole , i.e. macroscopically, are:

- Volume V,
- Energy E,
- Number of particles N,
- Entropy $S = k \log \Omega$ where Ω are systems accessible states and k is called Boltzmann constant,
- Temperature T, which represents the relative change energy when we change the entropy of system,
$$T = \frac{\partial E}{\partial S}.$$

The microscopic quantities are those which specify the states of single particles. In case of the ideal gas those are the vectors of the position r_i and momentum p_i, for i-th particle. Note that several other microscopic quantities can be derived from those basic ones, for example speed. The number of particles which have certain speed in system is also of microscopic quantity. To obtain such a quantity one would require the information about all particles momentum.

Generally, gases behave as an ideal gas only under certain conditions. Physically, this situation occurs when the concentration of the molecules is sufficiently small. However, speaking more rigorously, the potential energy, interactions, between particles have to be of negligible size. The following example illustrates this situation.

Let us consider the gas of the N molecules confined in the container of the volume V. The total energy of this system can be written as:

$$E = K + U + E_{int},$$

where K denotes the kinetic energy of the molecules. If the momentum of i-th molecule is given by the vector p_i then K is equal to:

$$K(p_1, p_2, \ldots, p_N) = \frac{1}{2m} \sum_{i=1}^{N} p_i^2,$$

where m is the mass of the single molecule. The quantity $U = U(r_1, r_2, \ldots, r_N)$ represents the potential energy of the mutual interaction of the particles and depends on the centre-of-mass positions of the molecules r_i. The term E_{int} is energy of the intermolecular interaction, which for the mono atomic gases $E_{int} = 0$. We call a gas an ideal if the potential energy of the interactions is negligibly small, i.e. $U \approx 0$. This usually can be achieved by increasing average distance between the molecules so that the collisions are relatively rare. Obviously, this can be achieved by, for example, decreasing the concentration of molecules N/V.

When we consider all possible configurations, i.e. microscopic states, as separate systems, those instances form so-called statistical ensemble. The aim of defining ensemble is to represent the probability of some systems while their macroscopic parameters have certain values. This implies that only the fraction of systems in ensemble will have this parameter of that specified value. We assume that, while system is in equilibrium, all configurations occur with the same probability, i.e. there is nothing special in any configuration to distinguish its occurrence. This is known in statistical mechanics as basic statistical postulate. We consider only the case of system in equilibrium, therefore this postulate will be valid.

It is important to mention that, on the microscopic level the system is governed by the laws of quantum mechanics. However, under some special circumstances for the large number of cases it obeys the laws of classical mechanics. The simplification to classical description has some significant consequences in computations. The essence of difference between classical and quantum description lies in the particles distinguishability. In other words, the number of systems in ensemble can be altered due to the distinguishability of particles. For example, if we interchange two particles and such change will result in obtaining two different states, then we deal with distinguishable particles. This occurs only in classical mechanics approximation. In quantum case the particles are considered as identical and therefore indistinguishable. Moreover, due to quantum mechanical results, we have two types of the identical particles: bosons and fermions. The difference between them is in the restriction of number of the particles which can have single energy value, i.e. occur on a single energy level. For fermions, only one particle in a single state is allowed, while there is no restriction for bosons. The difference is of significance while counting the number of constrained systems accessible states. Technically we can check whether we should use the classical approach. This can be done by measuring some macroscopic quantities and checking if the following condition holds

$$\left(\frac{V}{N}\right)^{\frac{1}{3}} \gg \frac{h}{\sqrt{3mkT}}.$$

We can infer if we can use classical mechanics when the concentration of molecules N/V is relatively small, temperature T and mass of molecule m are sufficiently high.

1.2.2 Derivation of Bose-Einstein and Maxwell-Boltzmann statistics

Both, Maxwell-Boltzmann and Bose-Einstein statistics give answers to the following question: how particles are distributed over the spectrum of available energies in the system, on average. Corresponding to our general overview of statistical mechanics, having the values of some macroscopic quantities, measurement, in our case fixed energy and the number of particles, we draw conclusions about system on the microscopic level. The difference between two statistics is in the distinguishability of particles, i.e. if we interchange two particles on two levels for Maxwell-Boltzmann this will count as two micro states but for Bose-Einstein this will be the same state, as essentially the number of particles on the energy levels didn't change.

In the literature there are two ways of deriving those statistics, the method of averages which is based on the grand canonical ensemble and the second, based on the entropy maximization, the method of most probable values.

Method of averages

The name of this method is related to the final result which is the average number of particles in strict sense. The general formula for the average number of particles $\overline{N}_i$ having the energy ε_i

$$\overline{N}_i = \frac{\sum\limits_r N_{i,r}\Omega(E_r, N_r)}{\sum\limits_r \Omega(E_r, N_r)} = \sum_r N_i P_r(E_r, N_r),$$

where summation is over all distinguishable states r for all energy E_r and number of particles N_r

$$\sum_{i=1}^{m} N_i = N,$$
$$\sum_{i=1}^{m} \varepsilon_i N_i = E,$$

and $\Omega(E_r, N_r)$ is the number of accessible states for given E_r and N_r. Each state r corresponds to the particular vector of number of particles $(N_1, N_2, \ldots, N_m)$.

Next step is to approximate the probability $P_r(E_r, N_r)$. However, to do that we first have to introduce the concept of reservoir.
Let us consider the situation when our system, denoted by A, is in contact with hypothetical heat and particle reservoir A'. Both systems $A + A' = A^{(o)}$ are isolated, i.e. do not exchange heat or particles with outside, only between each other. The total energy and particles are given by:

$$E + E' = E^{(o)} = constant,$$
$$N + N' = N^{(o)} = constant,$$

where N, E are the particles and energy of our considered system, while E' and N' are of the reservoir. The system A' is called reservoir because we assume that its energy and the total number of particles change insignificantly after contact with the considered system A, i.e. A' is much bigger than A. Which also means that $E^{(o)} \gg E_r$ and $N^{(o)} \gg N_r$ and we have the approximation

$$\frac{\partial^2 \Omega'}{\partial E'^2} E_r \ll \frac{\partial \Omega'}{\partial E'} \tag{1.5}$$

Then the probability $P_r(E_r, N_r)$ can be well approximated by considering the combined system $A^{(o)}$ instead of just A. Let $\Omega'(E', N')$ denote the number of states accessible, i.e. entropy, of the reservoir A' while A is in one of the definite states r with N_r particles and energy E_r. For that case the number of accessible states for the combined system $A^{(o)}$ is just a number of states accessible for the reservoir $S'(E^{(o)} - E_r, N^{(o)} - N_r)$ while the total amount of states accessible for $A^{(o)}$ is $\Omega^{(o)}(E^{(o)}, N^{(o)})$. The "starting"idea of the method is the equivalence of probability between finding A in state r and finding $A^{(o)}$ in fraction of states that A is in r and A' is one of the $S'(E^{(o)} - E_r, N^{(o)} - N_r)$. It can be expressed in formula:

$$P_r(E_r, N_r) = \frac{\Omega'(E^{(o)} - E_r, N^{(o)} - N_r)}{\Omega^{(o)}(E^{(o)}, N^{(o)})}.$$

We simplify the above formula by "extracting "the dependency of E_r and N_r from S'. The procedure is following. We first represent S' in terms of Taylor expansion

$$\ln \Omega'(E^{(o)} - E_r, N^{(o)} - N_r) = \ln \Omega'(E^{(o)}, N^{(o)}) - \left[\frac{\partial \ln \Omega'}{\partial E'}\right]_{E'=E^{(o)}} E_r - \\ - \left[\frac{\partial \ln \Omega'}{\partial N'}\right]_{N'=N^{(o)}} N_r + \dots,$$

where higher terms are neglected due to conditions (1.5).
The appearing derivatives are denoted

$$\lambda = \left[\frac{\partial \ln S'}{\partial E'}\right]_{E'=E^{(o)}}, \; \nu = \left[\frac{\partial \ln S'}{\partial N'}\right]_{N'=N^{(o)}}.$$

Then we exponentiate both sides

$$\Omega'(E^0 - E_r, N^0 - N_r) \approx \Omega'(E^0, N^0) e^{-\lambda E_r - \nu N_r},$$

and our probability is given by

$$P_r(E_r, N_r) = \frac{\Omega'(E^{(o)}, N^{(o)})}{\Omega^{(o)}(E^{(o)}, N^{(o)})} e^{-\lambda E_r - \nu N_r}.$$

Such class of probability distributions is called “grand canonical” distribution. The first part of the expression on the right side is independent of the particular state r and can be calculated from the normalization condition $\sum_r P_r = 1$ and eventually we have

$$P_r(E_r, N_r) = \frac{e^{-\lambda E_r - \nu N_r}}{\sum\limits_r e^{-\lambda E_r - \nu N_r}}.$$

Now, going back to our microscopic quantity - the average number of particles on some energy level is given by the expression

$$\overline{N}_i = \frac{\sum\limits_r N_i e^{-\lambda E_r - \nu N_r}}{\sum\limits_r e^{-\lambda E_r - \nu N_r}} = -\frac{1}{\lambda \mathcal{Z}} \frac{\partial \mathcal{Z}}{\partial \varepsilon_i}, \tag{1.6}$$

where two constant λ and ν are unknown and $\mathcal{Z}$ is called grand partition function and is given by

$$\mathcal{Z} = \sum_r e^{-\lambda E_r - \nu N_r}.$$

The grand partition function is altered for Maxwell-Boltzmann statistics and we have to add Gibbs correction factor to the exponent. Further we consider the “degeneration” of the energy level. For each N_i particles on the level with energy ε_i we can additionally redistribute them on the over the G_i sub-levels. Physically, this corresponds to the fact that some energy levels in the system are very close to one another and that is why they can be grouped as one. Due to this degenerations, the grand partition function is altered for both statistics and we provide the details separately for each statistics.

1. Maxwell-Boltzmann statistics
   Here the additional factor for the grand partition factor is due to the distinguishability of particles and level degeneration and is given by

   $$w_{M.B.} = \prod_{i=1}^{m} \frac{G_i^{N_i}}{N_i!},$$

   Hence the grand partition function is

   $$\mathcal{Z}_{M.B.} = \sum_r \prod_{i=1}^{m} \frac{G_i^{N_i}}{N_i!} e^{-\lambda E_r - \nu N_r}.$$

Next we transform it

$$\sum_r \prod_{i=1}^m \frac{G_i^{N_i}}{N_i!} e^{-\lambda E_r - \nu N_r} = \sum_{N=0}^{\infty} \sum_{N_1+N_2+\ldots+N_m=N} \frac{1}{N!} \frac{N!}{N_1!N_2!\ldots N_m!} \prod_{i=1}^m \left(G_i e^{-\lambda \varepsilon_i - \nu} \right)^{N_i},$$

and using the multinomial theorem

$$(x_1 + x_2 + \ldots + x_m)^N = \sum_{N_1+N_2+\ldots+N_m=N} N! \prod_{i=1}^m \frac{x_i^{N_i}}{N_i!}$$

we get

$$\sum_{N=0}^{\infty} \sum_{N_1+N_2+\ldots+N_m=N} \frac{1}{N!} \frac{N!}{N_1!N_2!\ldots N_m!} \prod_{i=1}^m \left(G_i e^{-\lambda \varepsilon_i - \nu} \right)^{N_i} = \sum_{N=0}^{\infty} \frac{1}{N!} \left(\sum_{i=1}^m G_i e^{-\lambda \varepsilon_i - \nu} \right)^N,$$

where the outcome expression is a series representation of the exponent, hence the grand partiton function for Maxwell-Boltzmann statistics is

$$\mathcal{Z}_{M.B.} = \exp \left(\sum_{i=1}^m G_i e^{-\lambda \varepsilon_i - \nu} \right),$$

and we calculate the statistics itself from formula (1.6)

$$\overline{N}_i = G_i e^{-\lambda \varepsilon_i - \nu}.$$

Regarding the parameters λ and ν, for system with sufficiently large numbers of particles, the number of accessible states $1/\sum_r e^{-\lambda E_r - \nu N_r}$ is rapidly increasing function of E' and N', on the other hand $e^{-\lambda E_r - \nu N_r}$ is rapidly decreasing. In that situation the function $e^{-\lambda E_r - \nu N_r}/\sum_r e^{-\lambda E_r - \nu N_r}$ which is our probability (1.2.2), experience very sharp maximum for some unknown values λ and ν. This sharp peak occurs for the related values $E_r = \widetilde{E}$ and $N_r = \widetilde{N}$. For other values E_r and N_r the probabilities $P_r(E_r, N_r) \approx 0$. We determine λ and ν by fixing quantities $\overline{E}$ and $\overline{N}$ obtained, possibly, by some macroscopic measurement. The probability of finding system in $\widetilde{E}$ and $\widetilde{N}$ is incomparably higher as for those values there was a sharp peak. Hence we can assume that $\widetilde{E}$ and $\widetilde{N}$ are also average values, namely $\overline{E} = \widetilde{E}$ and

$\overline{N} = \widetilde{N}$. Then we can calculate the parameters from the two equations

$$\overline{E} = \sum_{i=1}^{m} \varepsilon_i G_i e^{-\lambda\epsilon_i - \nu},$$
$$\overline{N} = \sum_{i=1}^{m} G_i e^{-\lambda\varepsilon_i - \nu}.$$

2. Bose-Einstein statistics

   In this case alteration of grand partition function is due to the sublevels G_i. We can express it in terms of combinatoric formula. A number of possibilities of redistributing N_1 indistinguishable particles over G_i sublevels is given

$$w_{B.E.} = \prod_{i=1}^{m} \frac{(N_i + G_i - 1)!}{N_i!(G_i - 1)!}, \tag{1.7}$$

   and then

$$\mathcal{Z}_{B.E.} = \prod_{i=1}^{m} \frac{(N_i + G_i - 1)!}{N_i!(G_i - 1)!} e^{-\lambda E_r - \nu N_r},$$

   which can be simplified by using generalised geometric series formula

$$\frac{1}{(1-x)^s} = \sum_{n=0}^{\infty} \binom{n+s-1}{n} x^n.$$

   hence we have

$$\mathcal{Z}_{B.E.} = \left(\sum_{N_1=0}^{\infty} \frac{(N_1 + G_1 - 1)!}{N_1!(G_1 - 1)!} e^{-(\lambda\varepsilon_1 + \nu)N_1} \right) \left(\sum_{N_2=0}^{\infty} \frac{(N_2 + G_2 - 1)!}{N_2!(G_2 - 1)!} e^{-(\lambda\epsilon_2 + \nu)N_2} \right) \dots =$$
$$= \left(\frac{1}{1 - e^{-\lambda\epsilon_1 - \nu}} \right)^{G_1} \left(\frac{1}{1 - e^{-\lambda\epsilon_2 - \nu}} \right)^{G_2} \dots = \prod_{i=1}^{m} \left(\frac{1}{1 - e^{-\lambda\epsilon_s - \nu}} \right)^{G_i},$$

   and from (1.6) we get the Bose-Einstein statistics

$$\overline{N}_i = \frac{G_i}{e^{\lambda\varepsilon_i + \nu} - 1}.$$

   Here, similar to the situation for the Maxwell-Boltzmann statistics, we calculate the statistics from the two equations

$$\overline{E} = \sum_{i=1}^{m} \varepsilon_i \frac{G_i}{e^{\lambda \epsilon_i + \nu} - 1},$$
$$\overline{N} = \sum_{i=1}^{m} \frac{G_i}{e^{\lambda \varepsilon_i + \nu} - 1}.$$

Method of most probable values

In this method we consider explicit function of number of accessible micro states in the system, the entropy, and we find the most probable micro state and assume this is the average state. We are given constraint for the number of particles and energy

$$\sum_{i=1}^{m} N_i = N, \qquad (1.8)$$
$$\sum_{i=1}^{m} \varepsilon_i N_i = E.$$

The number of accessible states to the system is

$$\Omega(N, E) = \sum_{\{N_i\}} W(\{N_i\}),$$

where sum $\{N_i\}$ is sum over all possible vectors $(N_1, N_2, \ldots, N_m)$ that conform to the conditions (1.8) and $W(\{N_i\})$ is a number of possible distribution corresponding to given vector of N_i's. Here we also consider the case when each $i - th$ level has G_i sublevels.

1. Maxwell-Boltzmann statistics
   In the case, due to distinguishability of particles, the $W(\{N_i\})$ is given by

$$W_{M.B.}(\{N_i\}) = \prod_{i=1}^{m} \frac{(G_i)^{N_i}}{N_i!}.$$

Then the systems entorpy is equal

$$S(N,E) = k\ln\Omega(N,E) = k\ln\Bigg(\sum_{\{N_i\}}\prod_{i=1}^{m}\frac{(G_i)^{N_i}}{N_i!}\Bigg)$$

Now, we find the vector $(N_1^*, N_2^*, \ldots, N_m^*)$ in the sum for which the entropy is largest and value for this vector will be much larger than for others in the thermodynamical limit, hence we have approximation

$$S(N,E) \approx k\ln\Bigg(\prod_{i=1}^{m}\frac{(G_i)^{N_i^*}}{N_i^*!}\Bigg), \; N \to \infty.$$

We find this vector by finding the maximum of the entropy using the Lagrange multipliers method. However, first we approximate the logarithm with known Stirling formula for factorials $\ln N! = N\ln N - N$, then the entropy is

$$S(N,E) \approx \sum_{i=1}^{m} N_i^* \ln\frac{G(i)}{N_i^*}.$$

Then the equation for maximum value is given by

$$\frac{\partial}{\partial N_i}\sum_{i=1}^{m}\Bigg[N_i\ln\frac{G(i)}{N_i} - \lambda\Bigg(\sum_{i=1}^{m}\varepsilon_i N_i - E\Bigg) - \nu\Bigg(\sum_{i=1}^{m} N_i - N\Bigg)\Bigg]_{N_i=N_i^*} = 0,$$

where λ and ν are Lagrange multipliers. The solution is

$$\sum_{i=1}^{m}\Bigg[\ln\frac{G(i)}{N_i} - \lambda\varepsilon_i - \nu\Bigg]_{N_i=N_i^*} = 0,$$

Hence we have

$$N_i^* = G_i e^{-\lambda\varepsilon_i - \nu},$$

which is Maxwell-Boltzmann statistics.

2. Bose-Einstein statistics
For this case we perform calculations analogical to the Maxwell-Boltzmann case. The entropy, due to indistinguishability of particles is given by

$$S(N,E) = k\ln\Omega(N,E) = k\ln\Bigg(\sum_{\{N_i\}}\prod_{i=1}^{m}\frac{(N_i+G_i-1)}{N_i!(G_i-1)!}\Bigg).$$

Then again, we apply Stirling approximation and obtain

$$S(N,E) \approx N_i \ln\left(\frac{G(i)}{N_i}+1\right) - G_i \ln\left(1+\frac{N_i}{G_i}\right), N \to \infty,$$

and the equation for the maximum values is

$$\frac{\partial}{\partial N_i} \sum_{i=1}^{m} \Bigg[N_i \ln\left(\frac{G(i)}{N_i}+1\right) - G_i \ln\left(1+\frac{N_i}{G_i}\right) -$$
$$\lambda\Bigg(\sum_{i=1}^{m} \varepsilon_i N_i - E\Bigg) - \nu\Bigg(\sum_{i=1}^{m} N_i - N\Bigg)\Bigg]_{N_i=N_i^*} = 0,$$

and outcome is

$$\left[\ln\left(\frac{G(i)}{N_i}+1\right) - \lambda\varepsilon_i - \nu\right]_{N_i=N_i^*} = 0,$$

and finally

$$N_i^* = \frac{G_i}{e^{\lambda\varepsilon_i-\nu}-1},$$

which is Bose-Einstein statistics.

1.3 Complexity Science: Zipf Law and other Power Laws

Across all disciplines of science the complex systems are common entities, basically the whole world is built of many complex systems. However, their structure and composition can be very different and the context in which they exists is various. It can be the virtual world inside a computer's memory, for example citation network or the computers or other electronic devices itself, like the Internet. It can be networks such as electricity or road networks. It can be social networks, the network of humans or other animals and their corresponding links, family, friendships or business relations. In nature, the crown of the sun and related to it solar flares, river systems or the coastal line can all be considered as complex systems, too. Fundamentally, any collection of similar entities that are in some way linked and interact with each other can be considered as complex systems. The analysis, categorizing and predicting of such systems, sooner or later was inevitable, therefore the need for an interdisciplinary field that would face this task has been obvious. That is how complexity science has emerged. Its emergence and evolution, however, was gradual and not linear. The picture 1.4 depicts a diagram of the evolution of Complexity Science, taken from [14].

There are a few disciplines close to and of similar origin as complexity science, this can be for example dynamic systems or agent based modeling. However, 'mainstream' complexity science emerged somewhere between 60's and 70's of the last century. It originated from few subfields, Cybernetics developed by Norbert Weiner, Systems Theory founded by Ludwig von Bertalanaffy and Dynamic systems theory. Over time several concepts were developed within complexity science, such as self-organisation and adaptation in the late 70's. Then in the 80's, Per Bak self organised criticality, related to emergence and dynamics in the systems and in the 90's and later there was a focus on the complex networks.

Most of the examples which we will discuss hare are taken from [7].
Among many complex systems particular class are those who manifests so-called power law behaviour. As an example we can recall the findings of Professor George Zipf, published in 1949 in the book 'Human Behavior and the Principle of Least effort'. He makes several interesting observations in the system of cities. He plotted the major cities of the world, starting from the biggest and ending with the smallest on the logarithmic plot. As a result he achieved a roughly straight line, see figure 1.5.
This fact was named after him Zipf Law and can be written by formula as

$$N(s) = \frac{1}{s}$$

where $N(s)$ is the number of cities with more than s inhabitants. In more general frameworks, it is a power law with exponent 1.

The next example of power systems are the earthquakes which occurred for a certain amount of time in a particular place. Figure 1.6 presents a power law for the earthquakes in New Madrid in USA earthquakes zone in the period 1974-1983 and the picture next to it shows the locations of these earthquakes. On one y-axis you have earthquake magnitude and on the y-axis the rank of particular earthquake. The plot is again a straight line which means it is a power law.

The other example is earthquakes which have occurred world wide since 1940. The data is from the USGS National Earthquake Information Center and its predecessors, the Coast and Geodetic Survey. Figure 1.7 presents the data on

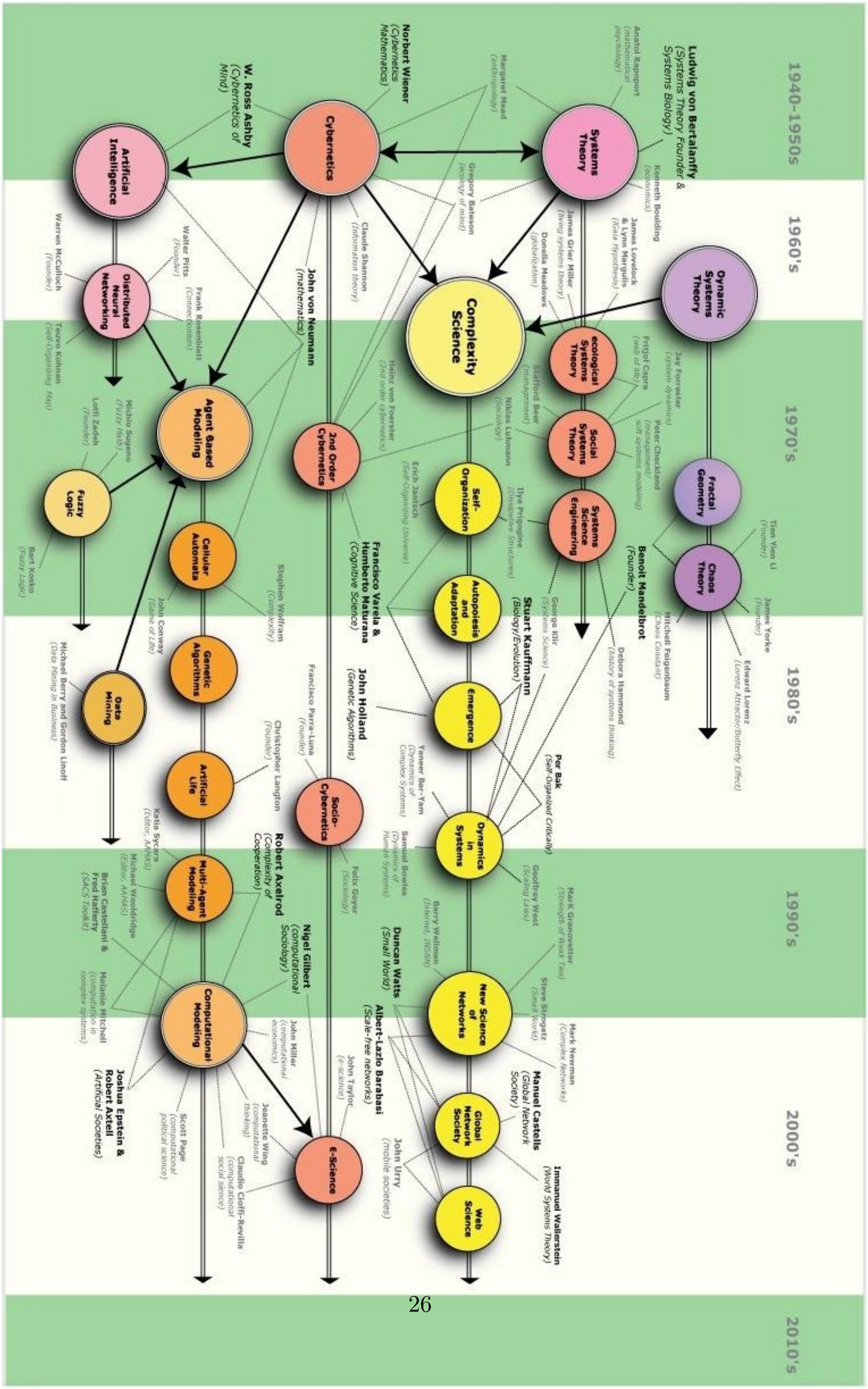


Figure 1.4: Diagram of evolution of Complexity Science

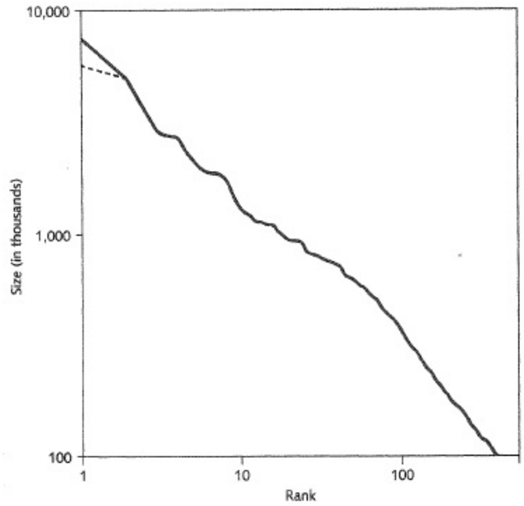


Figure 1.5: Ranking of city sizes around the year 1920

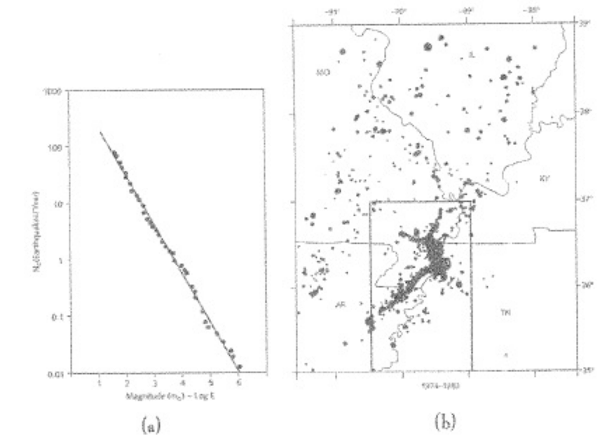


Figure 1.6: Illustration of Gutenberg-Richter law a) logarithmic plot of occured earthquakes , b) corresponding places of occurrence

a standard plot. On y-axis we have a earthquake magnitude and on x-axis the ranks of particular earthquakes. Here also we see the curve is a power law shape.

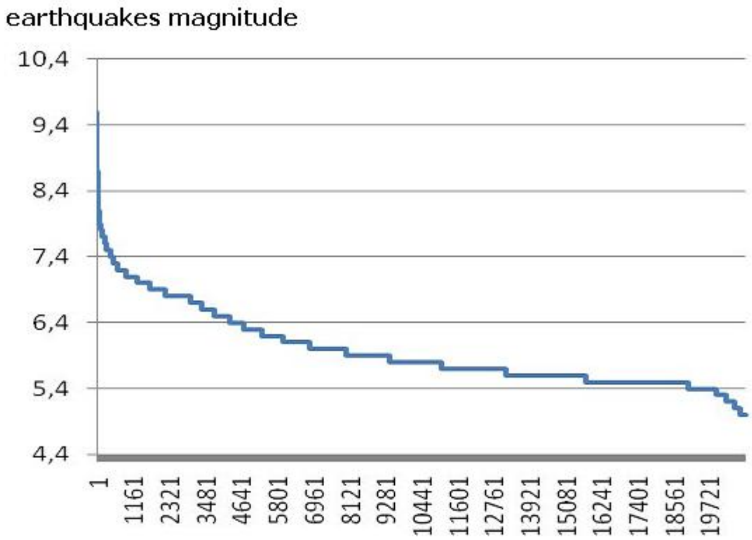


Figure 1.7: Illustration of earthquakes worldwide since 1940

We consider earth species extinctions as a system, take the mass extinctions through out the recorded history and plot them. On x-axis is the percent of organism extinct during geological stage and on y-axis number of such stages that occurred in the earth history. Here also manifests a power law shape.

The last example is a power spectrum of a traffic jam (Figure 1.9) on the logarithmic plot. Research by the Kai Nagel and Maya Paczuszki in 1995.

As we see from those examples, there are many various systems which experience power law behaviour and in some cases it is a Zipf Law. Many other examples could be recalled here. However, what is the important, is that power laws are purely experimental law. They have been obtained by mere observation and joining the plot, there is no mathematical framework, theory that fully explains those phenomenas.

1.4 Thesis outline

The thesis consist of five chapters and the appendix chapter, where the first was an introduction.

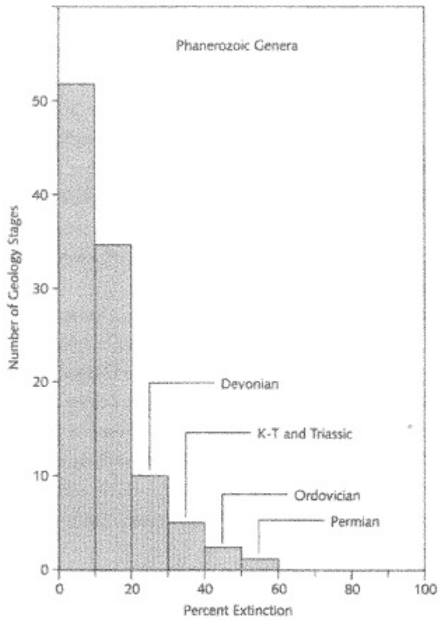


Figure 1.8: Plot of extinctions throughout the history of earth

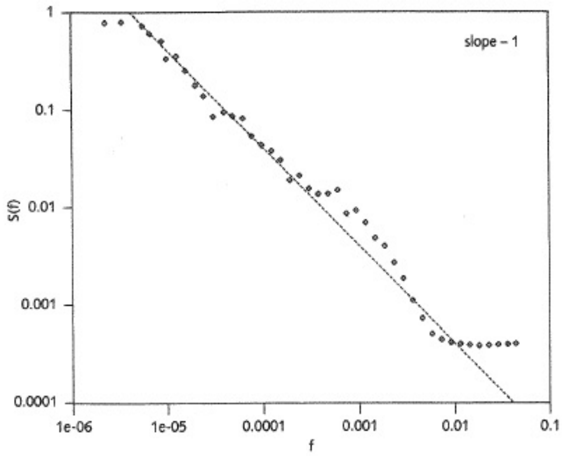


Figure 1.9: Power spectrum of traffic jam

The second chapter include the main results of the thesis. In several sections we provide mathematically rigorous limit theorem and corresponding fluctuation theorems with full proofs. The first theorem is built on already existing result developed by Prof V.P. Maslov. We introduced it in the Section 1 of the

Introduction as Theorem 2. The 'new version' we developed is a more rigorous and precise extension of that result, with a mathematically rigorous proof. The fluctuation theorems are the new results constructed on the fundaments of the first theorem. The results of that chapter are in the phase of preparation for publishing.

In the third chapter we present several results which we developed specially for the proofs of the Chapter 2. It consists of solutions of some optimization problems, some approximations and estimates. The results of this chapter are mathematically rigorous with proofs provided. One result is given without proof and is left for the future research.

The fourth chapter also is devoted for the results developed specifically for the proofs of Chapter 2. It includes some extension of Laplace approximation put in the few sections. These are new results however of minor relevancy. They were constructed based on the Laplace approximation in the book [13].

Last chapter is devoted to conclusions, applications and future research. We underline the contribution of our work to the field of Statistical Physics and Complexity Science. A short section on possible application is included. Finally, we emphasize possible future directions related to our work which can be conducted, some ideas which came across during our research and possible extensions of work already done.

In the Appendix we put all the well known results we used throughout the thesis, some minor results are proved and some basic definitions are also recalled. It consists of the Analysis, Probability, Asymptotics and Optimization.

2 Limit theorems

In this chapter we introduce and prove the main results of the thesis. The limit theorem and the corresponding fluctuation theorems.

We introduce a mathematical setting and all the assumptions on which the results are based in the section one.

The content of the Section 2 is the limit theorem about the convergence of the considered random variable to constant mean value. Corresponding estimate of the speed of convergence are also included. This result is an extension and more precise version of the Malsov Theorem , Theorem 2 in Introduction.

Next two sections are devoted to the fluctuation theorems. They provide information on the distribution of deviation of considered random variable from the maximum. As it turned out from the previous section there are two types of means, depending on the initial assumptions. As a result there are two fluctuation theorems. In Section 3 we have one case, when the mean is in the interior of the sample space and in the Section 4 the mean is on the boundary.

In the last Section we provide some additional results, estimates, used in proof the fluctuation theorems. For the transparency of the proofs we moved it to a separate section.

2.1 Introduction

This section consists of a step by step introduction of the mathematical setting which forms a background for the results of this thesis. Several assumption are made on the way in order to simplify the setting and make construction of the proofs possible.

For given integers $G, N > 0$, real number $E > 0$ and mapping $\varepsilon : \{1, 2, \dots G\} \to \mathbb{R}$ we introduce a probability space. The elementary events are uniformly distributed G-dimensional vectors of nonnegative integers $n_i, i =$

$1, \ldots, G$ satisfying constraints:

$$N = n_1 + n_2 + \ldots + n_G, \tag{2.1}$$

$$EN \geq \varepsilon(1)n_1 + \varepsilon(2)n_2 + \ldots + \varepsilon(G)n_G. \tag{2.2}$$

In physics we call such system micro-canonical ensemble.

Arbitrary elementary event can be illustrated as the random distribution of N balls in G boxes. Moreover, each box has 'weight' coefficient $\varepsilon(i)$ and the total 'weight' must be less or equal EN.

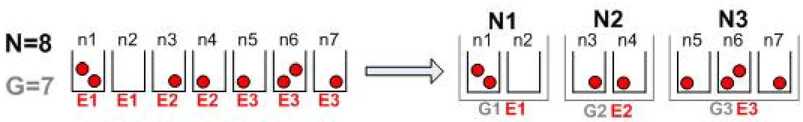
Furthermore, let us denote the image of the function ε as the set $\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m\}$ and without loss of generality it can be ordered $\varepsilon_1 < \varepsilon_2 < \ldots < \varepsilon_m$. To each element in the set corresponds a positive integer G_i, $i = 1, 2, \ldots, m$ representing the number of points in the domain of ε having the values ε_i, so that $G = \sum_{i=1}^{m} G_i$.

We can use this setting to define probability space in an alternative way. We consider the values G_i and $\varepsilon_i, i = 1, \ldots, m$ instead of the mapping ε. Respectively, the conditions (2.1) and (2.2) are reformulated

$$N = N_1 + N_2 + \ldots + N_m, \tag{2.3}$$

$$EN \geq \varepsilon_1 N_1 + \varepsilon_2 N_2 + \ldots + \varepsilon_m N_m, \tag{2.4}$$

where $N_i = n_{G_1+\ldots+G_{i-1}+1} + \ldots + n_{G_1+\ldots+G_{i-1}+2} + n_{G_1+\ldots+G_{i-1}+G_i}$ for $i = 1, \ldots, m$. Vectors satisfying above conditions form a sample space which will be denoted by $\Omega_{N,E}$. This situation, can be illustrated as distributing N balls over m 'bigger' boxes, where to each corresponds unique value ε_i. Then in each i-th 'bigger' box balls are distributed over G_i boxes.



For given vectors $\mathcal{N} = (N_1, \ldots, N_m)$ and $\mathcal{G} = (G_1, \ldots, G_m)$ the number of different combinations which can occur in such redistribution, exactly the logarithm of that number is denoted by $S(\mathcal{N})$ and called Entropy.

We count those combinations using formula from Combinatorics for the possible number of unordered arrangements of size r obtained by drawing from n

objects,

$$S(\mathcal{N}) = \ln \prod_{i=1}^{m} \frac{(N_i + G_i - 1)!}{N_i!(G_i - 1)!}. \tag{2.5}$$

Let us consider the discrete random vector denoted by $X_N = (X_1, X_2, \ldots, X_m)$ where $X_i = N_i/N,\ i = 1, \ldots, m$ and respectively sample space given by transformed conditions (2.3) and (2.4) is given by

$$\begin{aligned} &1 = x_1 + x_2 + \ldots + x_m, \\ &E \geq \varepsilon_1 x_1 + \varepsilon_2 x_2 + \ldots + \varepsilon_m x_m, \qquad x_i \in \left\{ \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N-1}{N}, 1 \right\}, \end{aligned}$$

and denoted by Ω_E and respectively entropy function

$$S(x, N) = \ln \prod_{i=1}^{m} \frac{(x_i N + G_i - 1)!}{(x_i N)!(G_i - 1)!}.$$

The probability mass function (pmf) of random variable X is given by

$$Pr(X = x) = \frac{1}{Z(N, E)} \prod_{i=1}^{m} \frac{(x_i N + G_i - 1)!}{(x_i N)!(G_i - 1)!}, \tag{2.6}$$

where $Z(N, E)$ is a normalization constant specified by

$$Z(N, E) = \sum_{\Omega_E} \prod_{i=1}^{m} \frac{(x_i N + G_i - 1)!}{(x_i N)!(G_i - 1)!}, \tag{2.7}$$

which is a total number of elementary events in the sample space Ω_E. Sometimes $Z(N, E)$ is called partition function.

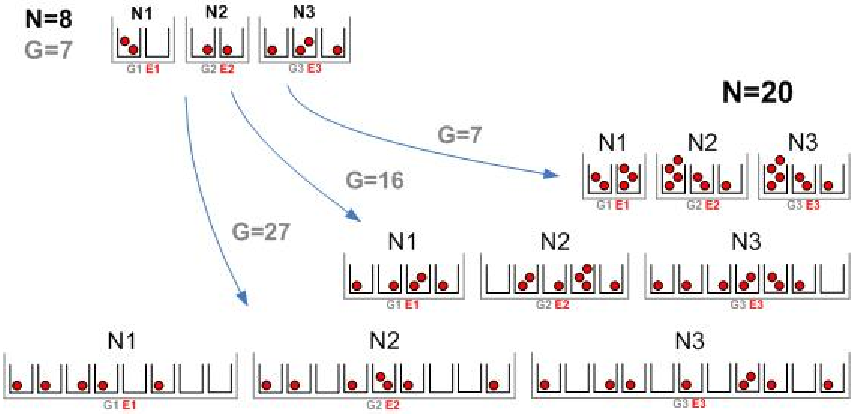
We are interested in the behaviour of random vector X as $N \to \infty$. We consider a particular case when $G = G(N)$ is an increasing function of N. Moreover, for each N the components G_i are equally weighted and their number m remains constant. Which means that for all N, $G_i = g_i G(N)$ for $i = 1, \ldots, m$ and some constants g_i such that $\sum_{i=1}^{m} g_i = 1$.

We distinguish three cases of function $G(N)$, depending on its asymptotic behaviour in $N \to \infty$

$$\begin{aligned} &1)\ \frac{G(N)}{N} \to \infty, \\ &2)\ \frac{G(N)}{N} \to c, \\ &3)\ \frac{G(N)}{N} \to 0, \end{aligned} \tag{2.8}$$

where c is some positive constant. The idea of three asymptotic cases is adopted from the paper of Maslov [11].

The picture below briefly illustrates the three cases.



2.2 Limit Theorem

The content of this section is our main result, the limit theorem which provides the mean values to which introduced in the previous section random variable converges. The two types of means are possible, depending on some sample space parameter. The proof is based on the convergence of corresponding moment generating function of the random variable. Additionally, the estimate for the speed of convergence of the moment generating function of considered random variable to mgf of mean is included.

Theorem 3 (Weak Law of large numbers)**.** Let X_N be the m-dimensional discrete random vector on the sample space Ω_E with pmf specified by (2.6). As $N \to \infty$ the random vector X_N converges in distribution to the constant vector $x^* = (x_1^*, x_2^*, \ldots, x_m^*)$. The exact values of the components of x^* depend on the sample space parameter E.

Let $\overline{g\varepsilon} = \frac{1}{m}\sum_{i=1}^{m} g_i\varepsilon_i$ and $g_{im}\varepsilon_{im} = \min_i g_i\varepsilon_i$, then

I) When $g_{im}\varepsilon_{im} < E < \overline{g\varepsilon}$ the components of x^* are

$$
\begin{aligned}
&1)\ x_i^* = \frac{g_i}{\lambda\varepsilon_i + \nu}, \quad && \text{if} \quad \frac{G(N)}{N} \to \infty,\\
&2)\ x_i^* = \frac{g_i}{e^{\lambda\varepsilon_i+\nu} - 1}, && \text{if} \quad \frac{G(N)}{N} \to c,\\
&3)\ x_i^* = \frac{g_i}{e^{\lambda\varepsilon_i+\nu}}, && \text{if} \quad \frac{G(N)}{N} \to 0,
\end{aligned}
$$

for $i = 1, \ldots, m$ and the parameters λ and ν are the solution of the system of equations

$$
\begin{aligned}
1 &= \sum_{i=1}^{m} x_i^*,\\
E &= \sum_{i=1}^{m} \varepsilon_i x_i^*.
\end{aligned}
$$

II) When $E \geq \overline{g\varepsilon}$ the components of x^* are

$$x_i^* = g_i,\ i = 1, \ldots, m.$$

Further, we have following estimates, distinct for the maximum in the interior and on the boundary

I) Maximum on the boundary

$$
\begin{aligned}
1)\ M_{X_N}(\xi) &= e^{\xi^T x^*} + O\left(\frac{1}{N^{1-\delta}}\right), \qquad && \text{when } \frac{1}{N} \gg \frac{N}{G(N)},\\
M_{X_N}(\xi) &= e^{\xi^T x^*} + O\left(\left(\frac{N}{G(N)}\right)^{1-\delta}\right), && \text{when } \frac{1}{N} \ll \frac{N}{G(N)},
\end{aligned}
$$

$$2)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{N^{1-\delta}}\right),$$

$$3)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{G(N)^{1-\delta}}\right), \qquad \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$
$$M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\left(\frac{G(N)}{N}\right)^{1-\delta}\right), \qquad \text{when } \frac{1}{G(N)} \ll \frac{G(N)}{N}$$

as $N \to \infty$.

II) Maximum in the interior

$$1)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{\sqrt{N}}\right), \qquad \text{when } \frac{1}{\sqrt{N}} \gg \frac{N}{G(N)},$$
$$M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\left(\frac{G(N)}{N}\right)^{1-\delta}\right)\Bigg), \qquad \text{when } \frac{1}{\sqrt{N}} \ll \frac{N}{G(N)},$$

$$2)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{\sqrt{N}}\right),$$

$$3)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{\sqrt{G(N)}}\right), \qquad \text{when } \frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N},$$
$$M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\left(\frac{G(N)}{N}\right)^{1-\delta}\right), \qquad \text{when } \frac{1}{\sqrt{G(N)}} \ll \frac{G(N)}{N},$$

as $N \to \infty$, valid for some arbitrary small constant δ, where $M_{X_N}(\xi)$ is moment generating function of the random vecto X_N.

Proof. We prove the theorem by showing convergence of the moment generating function of the random vector X_N to a constant vector x^* as $N \to \infty$.

The mgf of r.v. X_N is equal

$$M_X(\xi) = E[e^{\xi^T X}].$$

Evaluating the probability mass function we obtain following expression for $M_X(\xi)$

$$M_X(\xi) = \frac{1}{Z(N,E)} \sum_{\Omega_E} e^{\xi^T x} \prod_{i=1}^{m} \frac{(x_i N + G_i - 1)!}{(x_i N)!(G_i - 1)!}. \tag{2.9}$$

We start with approximating the first part of $M_{X_N}(\xi)$, i.e the normalization constant $Z(N,E)$, given by

$$Z(N,E) = \sum_{\Omega_E} e^{S(x,N)},$$

Let us consider only the first case of $G(N)$, (2.8). We use Lemma 1 from Section 1 in the Chapter III

$$e^{S(x,N)} = (2\pi)^{-\frac{m}{2}} e^{Nf_1(x,N)+R_1(N)}\left(1 + O\left(\frac{1}{N}\right)\right),\ N \to \infty$$

and then performing the summation over Ω_E and applying Triangle inequality on the LHS we get the following inequalities

$$Z(N,E) = (2\pi)^{-\frac{m}{2}} \sum_{\Omega_E} e^{Nf_1(x,N)+R_1(N)}\left(1 + O\left(\frac{1}{N}\right)\right),\ N \to \infty. \tag{2.10}$$

In the next step we approximate above sums using Lemma 7 from the Section 4 Ch.III

$$\sum_{\Omega_E} e^{Nf_1(x,N)+R_1(N)} = \int_{\Omega_E} e^{Nf_1(x,N)+R_1(N)} dx\left(1 + O\left(\frac{1}{N}\right)\right),\ N \to \infty,$$

and together with (2.10) we obtain

$$Z(N,E) = (2\pi)^{-\frac{m}{2}} \int_{\Omega_E} e^{Nf_1(x,N)+R_1(N)} dx\left(1 + O\left(\frac{1}{N}\right)\right),\ N \to \infty. \tag{2.11}$$

Then from Lemma 2 Section 2.1 Ch.III we have that functions $f_l(x),\ l = 1,2,3$ has two types of maximum depending on the sample space parameters E and $\varepsilon_i,\ i = 1,\ldots,m$. It can be on the boundary of the domain of optimization or in the interior of the domain. From the Lemma 3 in Section 2.2 of Chapter III, the function $f_l(x,N)$ has a unique maximum, and as $f_l(x,N) \to f_l(x), N \to \infty$ hence its maximum also is on the boundary of the domain or in the interior. For those two cases separately we apply Extended Laplace approximation from the Chapter IV.

I) When the maximum of $f_l(x,N)$ is attained on the boundary of the domain, we use Theorem 8 from Section 3 Chapter IV and for the first case we have

$$\int_{\Omega_E} e^{Nf_1(x,N)+R_1(N)}dx =$$
$$= e^{Nf_1(x^*(N),N)+R_1(N)}\frac{1}{N}\left(\frac{2\pi}{N}\right)^{\frac{m-1}{2}} \frac{|f_1'(x^*(N),N)|^{-1}}{\sqrt{|\det D^2 f_1(x^*(N),N)|}}\left(1+O\left(\frac{1}{N}\right)\right),$$

as $N\to\infty$, where $x^*(N)$ is a maximal point of $f_l(x,N),\ l=1,2,3$.
Then we combine above approximations with (2.11) and obtain for all three cases

1) $Z(N,E) =$
$$= e^{Nf_1(x^*(N),N)+R_1(N)}\frac{1}{2\pi}\frac{1}{N}\left(\frac{1}{N}\right)^{\frac{m-1}{2}} \frac{|f_1'(x^*(N),N)|^{-1}}{\sqrt{|\det D^2 f_1(x^*(N),N)|}}\left(1+O\left(\frac{1}{N}\right)\right),$$

2) $Z(N,E) =$
$$= e^{Nf_2(x^*(N),N)+R_2(N)}\frac{1}{2\pi}\frac{1}{N}\left(\frac{1}{N}\right)^{\frac{m-1}{2}} \frac{|f_2'(x^*(N),N)|^{-1}}{\sqrt{|\det D^2 f_2(x^*(N),N)|}}\left(1+O\left(\frac{1}{N}\right)\right),$$

3) $Z(N,E) = e^{G(N)f_3(x^*(N),N)+R_3(N)}\frac{1}{2\pi}\frac{1}{G(N)}\left(\frac{1}{G(N)}\right)^{\frac{m-1}{2}}\times$
$$\times\frac{|f_3'(x^*(N),N)|^{-1}}{\sqrt{|\det D^2 f_3(x^*(N),N)|}}\left(1+O\left(\frac{1}{G(N)}\right)\right),$$

as $N\to\infty$, where in the second case the alteration from the first case is only by the index of the function f_1. For the third case the alteration is in the index of f_1 and function $G(N)$ insted of N in th appriopriate places.

II) When the maximum of $f_l(x,N)$ is in the interior of the domain, we have the Extended Laplace approximation for the first case

$$\int_{\Omega_E} e^{Nf_1(x,N)+R_1(N)}dx =$$
$$= e^{Nf_1(x^*(N),N)+R_1(N)}\left(\frac{2\pi}{N}\right)^{\frac{m}{2}} \frac{1}{\sqrt{\det D^2 f_1(x^*(N),N)}}\left(1+O\left(\frac{1}{N}\right)\right),\ N\to\infty$$

where $x^*(N)$ is a maximal point.

Then we combine above approximations with (2.11) and obtain

$$1)\ Z(N,E) = e^{Nf_1(x^*(N),N)+R_1(N)}\left(\frac{1}{N}\right)^{\frac{m}{2}}\frac{1}{\sqrt{\det D^2 f_1(x^*(N),N)}}\left(1+O\left(\frac{1}{\sqrt{N}}\right)\right),$$

$$2)\ Z(N,E) = e^{Nf_2(x^*(N),N)+R_2(N)}\left(\frac{1}{N}\right)^{\frac{m}{2}}\frac{1}{\sqrt{\det D^2 f_2(x^*(N),N)}}\left(1+O\left(\frac{1}{\sqrt{N}}\right)\right),$$

$$3)\ Z(N,E) = e^{G(N)f_3(x^*(N),N)+R_3(N)}\left(\frac{1}{G(N)}\right)^{\frac{m}{2}}\times$$

$$\times\frac{1}{\sqrt{\det D^2 f_3(x^*(N),N)}}\left(1+O\left(\frac{1}{\sqrt{G(N)}}\right)\right),$$

as $N \to \infty$.

Analogically we approximate the other part of the mgf (2.9). The additional function under the sum does not affect the approximation of entropy nor the approximation of sum with the integral. In the Extended Laplace approximation this factor becomes function g in the Theorem. Hence we have

I) When the maximum $x^*(N)$ is on the boundary of the domain we have

$$1)\ \sum_{\Omega_E} e^{\xi^T x+S(x,N)} = e^{\xi^T x^*(N)+Nf_1(x^*(N),N)+R_1(N)}\frac{1}{2\pi}\left(\frac{1}{N}\right)^{\frac{m-1}{2}}\times$$

$$\times\frac{|f_1'(x^*(N),N)|^{-1}}{\sqrt{|\det D^2 f_1(x^*(N),N)|}}\left(1+O\left(\frac{1}{N}\right)\right),$$

$$2)\ \sum_{\Omega_E} e^{\xi^T x+S(x,N)} = e^{\xi^T x^*(N)+Nf_2(x^*(N),N)+R_2(N)}\frac{1}{2\pi}\left(\frac{1}{N}\right)^{\frac{m-1}{2}}\times$$

$$\times\frac{|f_2'(x^*(N),N)|^{-1}}{\sqrt{|\det D^2 f_2(x^*(N),N)|}}\left(1+O\left(\frac{1}{N}\right)\right),$$

$$3)\ \sum_{\Omega_E} e^{\xi^T x+S(x,N)} = e^{\xi^T x^*(N)+G(N)f_3(x^*(N),N)+R_3(N)}\frac{1}{2\pi}\left(\frac{1}{G(N)}\right)^{\frac{m-1}{2}}\times$$

$$\times\frac{|f_3'(x^*(N),N)|^{-1}}{\sqrt{|\det D^2 f_3(x^*(N),N)|}}\left(1+O\left(\frac{1}{N}\right)\right),$$

as $N \to \infty$.

II) When the maximum is inside the domain than we have

$$
\begin{aligned}
1)\ \sum_{\Omega_E} e^{\xi^T x + S(x,N)} &= e^{\xi^T x^*(N) + N f_1(x^*(N),N) + R_1(N)} \left(\frac{1}{N}\right)^{\frac{m}{2}} \times \\
&\times \frac{1}{\sqrt{\det D^2 f_1(x^*(N),N)}} \left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right), \\
2)\ \sum_{\Omega_E} e^{\xi^T x + S(x,N)} &= e^{\xi^T x^*(N) + N f_2(x^*(N),N) + R_2(N)} \left(\frac{1}{N}\right)^{\frac{m}{2}} \times \\
&\times \frac{1}{\sqrt{\det D^2 f_2(x^*(N),N)}} \left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right), \\
3)\ \sum_{\Omega_E} e^{\xi^T x + S(x,N)} &= e^{\xi^T x^*(N) + G(N) f_3(x^*(N),N) + R_3(N)} \left(\frac{1}{G(N)}\right)^{\frac{m}{2}} \times \\
&\times \frac{1}{\sqrt{\det D^2 f_3(x^*(N),N)}} \left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right),
\end{aligned}
$$

as $N \to \infty$.

Finally, we put together the approximations of the first and second part of mgf using Lemma 16 from the Appendix A.1 and cancel the identical terms. For two types of maximum we have separately

I) Maximum is on the boundary of the domain

$$
\begin{aligned}
1)\ M_{X_N}(\xi) &= e^{\xi^T x^*(N)} \left(1 + O\left(\frac{1}{N}\right)\right), \\
2)\ M_{X_N}(\xi) &= e^{\xi^T x^*(N)} \left(1 + O\left(\frac{1}{N}\right)\right), \qquad (2.12) \\
3)\ M_{X_N}(\xi) &= e^{\xi^T x^*(N)} \left(1 + O\left(\frac{1}{G(N)}\right)\right),
\end{aligned}
$$

as $N \to \infty$.

II) Maximum in the interior of the domain.

Here the situation is identical as for the boundary case but instead of N in the RHS we have $\sqrt{N}$ or $\sqrt{G(N)}$

$$
\begin{aligned}
&1)\ M_{X_N}(\xi) = e^{\xi^T x^*(N)}\left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right),\\
&2)\ M_{X_N}(\xi) = e^{\xi^T x^*(N)}\left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right),\\
&3)\ M_{X_N}(\xi) = e^{\xi^T x^*(N)}\left(1 + O\left(\frac{1}{\sqrt{G(N)}}\right)\right),
\end{aligned}
\tag{2.13}
$$

as $N \to \infty$.

Next we use following Taylor expansion

$$
e^{\xi^T x^*(N)} = e^{\xi^T x^*} + \xi e^{\xi^T x_\theta}(x^*(N) - x^*), \tag{2.14}
$$

where x^* is a maximum of limit functions of $f_l(x,N)$, $l = 1,2,3$ denoted by $f_l(x)$, given by Lemma 2 of Section 2.1 of Chapter III.
Further we substitute approximation for $(x^*(N) - x^*)$ given by Lemma 4 of Secion 3 in Chapter II and obtain

$$
\begin{aligned}
&1)\ e^{\xi^T x^*(N)} = e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\frac{1}{N^{1-\delta}}\right), && \text{when } \frac{1}{N} \gg \frac{N}{G(N)},\\
&\quad\ e^{\xi^T x^*(N)} = e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\left(\frac{N}{G(N)}\right)^{1-\delta}\right), && \text{when } \frac{1}{N} \ll \frac{N}{G(N)},\\
&2)\ e^{\xi^T x^*(N)} = e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\frac{1}{N^{1-\delta}}\right),\\
&3)\ e^{\xi^T x^*(N)} = e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\frac{1}{G(N)^{1-\delta}}\right), && \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},\\
&\quad\ e^{\xi^T x^*(N)} = e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\left(\frac{G(N)}{N}\right)^{1-\delta}\right), && \text{when } \frac{1}{G(N)} \ll \frac{G(N)}{N},
\end{aligned}
$$

as $N \to \infty$.
Now we combine it with approximations (2.12) and (2.13) for two cases of maximum

I) Maximum on the boundary

$$1)\ M_{X_N}(\xi) = \left[e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\frac{1}{N^{1-\delta}}\right)\right]\left(1 + O\left(\frac{1}{N}\right)\right), \qquad \text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$

$$M_{X_N}(\xi) = \left[e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\left(\frac{N}{G(N)}\right)^{1-\delta}\right)\right]\left(1 + O\left(\frac{1}{N}\right)\right), \qquad \text{when } \frac{1}{N} \ll \frac{N}{G(N)},$$

$$2)\ M_{X_N}(\xi) = \left[e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\frac{1}{N^{1-\delta}}\right)\right]\left(1 + O\left(\frac{1}{N}\right)\right),$$

$$3)\ M_{X_N}(\xi) = \left[e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\frac{1}{G(N)^{1-\delta}}\right)\right]\left(1 + O\left(\frac{1}{G(N)}\right)\right), \qquad \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$

$$M_{X_N}(\xi) = \left[e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\left(\frac{G(N)}{N}\right)^{1-\delta}\right)\right]\left(1 + O\left(\frac{1}{G(N)}\right)\right), \quad \text{when } \frac{1}{G(N)} \ll \frac{G(N)}{N},$$

as $N \to \infty$. Then we simplify above asymptotic equations and get

$$1)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{N^{1-\delta}}\right), \qquad \text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$

$$M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\left(\frac{N}{G(N)}\right)^{1-\delta}\right), \qquad \text{when } \frac{1}{N} \ll \frac{N}{G(N)},$$

$$2)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{N^{1-\delta}}\right),$$

$$3)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{G(N)^{1-\delta}}\right) \qquad \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$

$$M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\left(\frac{G(N)}{N}\right)^{1-\delta}\right), \qquad \text{when } \frac{1}{G(N)} \ll \frac{G(N)}{N},$$

as $N \to \infty$, where δ is some arbitrary small positive constant. Therefore, we get the final result for that case.

II) Maximum in the interior.

$$
\begin{aligned}
&1)\ M_{X_N}(\xi) = \left[e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\frac{1}{N^{1-\delta}}\right)\right]\left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right),\\
&\quad \text{when } \frac{1}{\sqrt{N}} \gg \frac{N}{G(N)},\\
&\quad M_{X_N}(\xi) = \left[e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\left(\frac{N}{G(N)}\right)^{1-\delta}\right)\right]\left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right),\\
&\quad \text{when } \frac{1}{\sqrt{N}} \ll \frac{\sqrt{N}}{G(N)},\\
&2)\ M_{X_N}(\xi) = \left[e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\frac{1}{N^{1-\delta}}\right)\right]\left(1 + O\left(\frac{1}{\sqrt{N}}\right)\right),\\
&3)\ M_{X_N}(\xi) = \left[e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\frac{1}{G(N)^{1-\delta}}\right)\right]\left(1 + O\left(\frac{1}{\sqrt{G(N)}}\right)\right),\\
&\quad \text{when } \frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N},\\
&\quad M_{X_N}(\xi) = \left[e^{\xi^T x^*} + \xi e^{\xi^T x_\theta} O\left(\left(\frac{G(N)}{N}\right)^{1-\delta}\right)\right]\left(1 + O\left(\frac{1}{\sqrt{G(N)}}\right)\right),\\
&\quad \text{when } \frac{1}{\sqrt{G(N)}} \ll \frac{G(N)}{N},
\end{aligned}
$$

as $N \to \infty$ and after simplification of above equation we get

$$
\begin{aligned}
&1)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{\sqrt{N}}\right), && \text{when } \frac{1}{\sqrt{N}} \gg \frac{N}{G(N)},\\
&\quad M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\left(\frac{N}{G(N)}\right)^{1-\delta}\right), && \text{when } \frac{1}{\sqrt{N}} \ll \frac{N}{G(N)},\\
&2)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{\sqrt{N}}\right), &&\\
&3)\ M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\frac{1}{\sqrt{G(N)}}\right) && \text{when } \frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N},\\
&\quad M_{X_N}(\xi) = e^{\xi^T x^*} + O\left(\left(\frac{G(N)}{N}\right)^{1-\delta}\right), && \text{when } \frac{1}{\sqrt{G(N)}} \ll \frac{G(N)}{N},
\end{aligned}
$$

as $N \to \infty$, which is our final result.

□

2.3 Fluctuation theorem, maximum in the interior of the domain

The fluctuation of the random variable from the mean value is introduced and proved in this section, the case when the maximum/mean is in the interior of the domain. As in the previous limit theorem the proof is based on the convergence of the moment generating functions. The speed of convergence of corresponding moment generating functions are included.

Theorem 4. For each case $G(N)$ given by (2.8) we have a m-dimensional random vector Y_N such that

$$\begin{aligned}
&1)\ Y_N = \sqrt{N}(X_N - x^*),\\
&2)\ Y_N = \sqrt{N}(X_N - x^*),\\
&3)\ Y_N = \sqrt{G(N)}(X_N - x^*),
\end{aligned}$$

defined on the discrete sample space Ω_E with pmf specified by (2.6).
Then for the sample space parameter $E \geq \overline{g\varepsilon}$, as $N \to \infty$ the distribution of the random vector Y converges to the multivariate normal $\mathcal{N}(0, -D^2 f_l(x^*)^{-1})$, where $l = 1, 2, 3$ indicates the case of $G(N)$.
Furthermore, we have estimates

$$\begin{aligned}
1)\ M_{Y_N}(\xi) &= e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi} + O\left(\frac{1}{N^{1/2-\delta}}\right), && \text{when } \frac{1}{N} \gg \frac{N}{G(N)},\\
M_{Y_N}(\xi) &= e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi} + O\left(\frac{N^{3/2-\delta}}{G(N)^{1-\delta}}\right), && \text{when } \frac{1}{\sqrt{N}} \gg \frac{N}{G(N)},\\
2)\ M_{Y_N}(\xi) &= e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi} + O\left(\frac{1}{N^{1/2-\delta}}\right), &&\\
3)\ M_{Y_N}(\xi) &= e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi} + O\left(\frac{1}{G(N)^{1/2-\delta}}\right), && \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},\\
M_{Y_N}(\xi) &= e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi} + O\left(\frac{G(N)^{3/2-\delta}}{N^{1-\delta}}\right), && \text{when } \frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N},
\end{aligned}$$

as $N \to \infty$, where δ is some arbitrary small constant.

Proof. The approach is analogical to the the proof in the previous section.
We first approximate the numerator and denominator of the mgf of Y_N using

Lemma 1 from Section 1, Ch. III, then Lemma 7 from the Section 4, Ch. III. Since the $E \geq \overline{g\varepsilon}$ from Lemma 2 Section 3, Ch.III we deduce that maximum of the approximated function is in the interior of the domain. Therefore we use appropriate Laplace approximation, Theorem 8, Section 2, Chapter IV. Finally we combine both approximations, for the numerator and denominator using Lemma 16 from the Appendix A.1. As a result we obtain following estimates

$$
\begin{aligned}
1)\ &\left| M_{Y_N}(\xi) - e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N) - f_1(x^*(N),N))} \frac{\sqrt{\det D^2 f_1(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_1(\widetilde{x}^*(N),N)}} \right| \leq \\
&\leq K_1^i \frac{1}{\sqrt{N}} e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N) - f_1(x^*(N),N))} \frac{\sqrt{\det D^2 f_1(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_1(\widetilde{x}^*(N),N)}},
\end{aligned}
$$

$$
\begin{aligned}
2)\ &\left| M_{Y_N}(\xi) - e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N) - f_2(x^*(N),N))} \frac{\sqrt{\det D^2 f_2(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_2(\widetilde{x}^*(N),N)}} \right| \leq \\
&\leq K_2^i \frac{1}{\sqrt{N}} e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N) - f_2(x^*(N),N))} \frac{\sqrt{\det D^2 f_2(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_2(\widetilde{x}^*(N),N)}},
\end{aligned}
$$

$$
\begin{aligned}
3)\ &\left| M_{Y_N}(\xi) - e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N) - f_3(x^*(N),N))} \frac{\sqrt{\det D^2 f_3(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_3(\widetilde{x}^*(N),N)}} \right| \leq \\
&\leq K_3^i \frac{1}{\sqrt{G(N)}} e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N) - f_3(x^*(N),N))} \frac{\sqrt{\det D^2 f_3(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_3(\widetilde{x}^*(N),N)}},
\end{aligned}
\tag{2.15}
$$

where $\widetilde{f}_l(x,N) = f(x,N) + \frac{1}{\sqrt{N}}\xi^T(x-x^*)$ for $l = 1,2$ and $\widetilde{f}_l(x,N) = f(x,N) + \frac{1}{\sqrt{G(N)}}\xi^T(x-x^*)$ for $l = 3$ and $\widetilde{x}^*(N)$ is a maximum of $\widetilde{f}(x,N)$. Next, we use the result of the Proposition 2 from Section 5 of this Chapter and multiply it by $\exp(N(\widetilde{f}_1(\widetilde{x}^*(N),N) - f_1(x^*(N),N)))$ to get

$$
\begin{aligned}
1)\ &\left| e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N) - f_1(x^*(N),N))} \frac{\sqrt{\det D^2 f_1(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_1(\widetilde{x}^*(N),N)}} - e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N) - f_1(x^*(N),N))} \right| \leq \\
&\leq K_1^{ii} \frac{1}{\sqrt{N}} e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N) - f_1(x^*(N),N))},
\end{aligned}
$$

$$2)\ \left| e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))} \frac{\sqrt{\det D^2 f_2(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_2(\widetilde{x}^*(N),N)}} - e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))} \right| \leq$$
$$\leq K_2^{ii} \frac{1}{\sqrt{N}} e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))}$$

$$3)\ \left| e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))} \frac{\sqrt{\det D^2 f_3(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_3(\widetilde{x}^*(N),N)}} - e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))} \right| \leq$$
$$\leq K_3^{ii} \frac{1}{\sqrt{G(N)}} e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))},$$

and combine above result with (2.15) and get

$$1)\ \left| M_{Y_N}(\xi) - e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N))} \right| \leq K_1^{iii} \frac{1}{\sqrt{N}} e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N))},$$
$$2)\ \left| M_{Y_N}(\xi) - e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))} \right| \leq K_2^{iii} \frac{1}{\sqrt{N}} e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))},$$
$$3)\ \left| M_{Y_N}(\xi) - e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))} \right| \leq\ K_3^{iii} \frac{1}{\sqrt{G(N)}} e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))},$$

where

$$1)\ K_1^{iii} = K_1^{ii} + K_1^{i}\left(1 + \frac{K_1^{ii}}{\sqrt{N}}\right),$$
$$2)\ K_2^{iii} = K_2^{ii} + K_2^{i}\left(1 + \frac{K_2^{ii}}{\sqrt{N}}\right),$$
$$3)\ K_3^{iii} = K_3^{ii} + K_3^{i}\left(1 + \frac{K_3^{ii}}{\sqrt{G(N)}}\right).$$

Now we use Proposition 1 from the Section 5 to approximate the expression in the exponent

$$1)\ N(\widetilde{f}_1(\widetilde{x}^*(N),N) - f_1(x^*(N),N)) = \sqrt{N}(x^*(N) - x^*)^T \left(\xi + \frac{\xi D^{\otimes 3} f_1(x_\theta(N),N)\xi}{\sqrt{N}} \right) +$$
$$+ \frac{1}{2} \xi^T D^2 f_1(x^*)^{-1} \xi + \frac{1}{\sqrt{N}} \left(\phi + \frac{\xi^T \kappa' \xi}{2\sqrt{N}} \right).$$

2) $N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))=\sqrt{N}(x^*(N)-x^*)^T\left(\xi+\dfrac{\xi D^{\otimes 3}f_2(x_\theta(N),N)\xi}{\sqrt{N}}\right)+$
$$+\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi+\frac{1}{\sqrt{N}}\left(\phi+\frac{\xi^T\kappa'\xi}{2\sqrt{N}}\right).$$

3) $G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))=\sqrt{G(N)}(x^*(N)-x^*)^T\left(\xi+\dfrac{\xi D^{\otimes 3}f_3(x_\theta(N),N)\xi}{\sqrt{G(N)}}\right)+$
$$+\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi+\frac{1}{\sqrt{G(N)}}\left(\phi+\frac{\xi^T\kappa'\xi}{2\sqrt{G(N)}}\right). \qquad (2.16)$$

Then we approximate the expression $(x^*(N)-x^*)$. From the Proposition 3 Section 4 Ch.III we have for all three cases

1. $\frac{G(N)}{N}\to\infty$
$$x^*(N)-x^*=\mathcal{K}_1 N^{-1+\delta}, \qquad \text{when } \frac{1}{N}\gg\frac{N}{G(N)},$$
$$x^*(N)-x^*=\mathcal{K}_1\left(\frac{N}{G(N)}\right)^{1-\delta}, \qquad \text{when } \frac{1}{N}\ll\frac{N}{G(N)}.$$

2. $\frac{G(N)}{N}\to c$
$$x^*(N)-x^*=\mathcal{K}_2 N^{-1+\delta}.$$

3. $\frac{G(N)}{N}\to 0$
$$x^*(N)-x^*=\mathcal{K}_3 G(N)^{-1+\delta}, \qquad \text{when } \frac{1}{G(N)}\gg\frac{G(N)}{N},$$
$$x^*(N)-x^*=\mathcal{K}_3\left(\frac{G(N)}{N}\right)^{1-\delta}, \qquad \text{when } \frac{1}{G(N)}\ll\frac{G(N)}{N},$$

where the form of approximations is different and $|\mathcal{K}_l|\le 1,\ l=1,2,3$. Then we substitute above estimates into (2.16) and after some rearrangements we have

$$1)\ N(\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N))=\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi+$$
$$+\frac{1}{N^{1/2-\delta}}\left(\mathcal{K}_1\xi+\mathcal{K}_1\frac{\xi D^{\otimes 3}f_1(x_\theta(N),N)\xi}{\sqrt{N}}+\frac{1}{N^\delta}\phi+\frac{1}{N^\delta}\frac{\xi^T\kappa'\xi}{2\sqrt{N}}\right),\text{ when }\frac{1}{N}\gg\frac{N}{G(N)},$$
$$N(\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N))=\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi+$$
$$+\frac{N^{3/2-\delta}}{G(N)^{1-\delta}}\left(\mathcal{K}_1\xi+\mathcal{K}_1\frac{\xi D^{\otimes 3}f_1(x_\theta(N),N)\xi}{\sqrt{N}}+\frac{G(N)^{1-\delta}}{N^{2-\delta}}\left(\phi+\frac{\xi^T\kappa'\xi}{2\sqrt{N}}\right)\right),\text{ when }\frac{1}{\sqrt{N}}\gg\frac{N}{G(N)},$$

$$2)\ N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))=\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi+$$
$$+\frac{1}{N^{1/2-\delta}}\left(\xi+\frac{\xi D^{\otimes 3}f_2(x_\theta(N),N)\xi}{\sqrt{N}}+\frac{1}{N^\delta}\phi+\frac{1}{N^\delta}\frac{\xi^T\kappa'\xi}{2\sqrt{N}}\right).$$

$$3)\ G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))=\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi+$$
$$+\frac{1}{G(N)^{1/2-\delta}}\left(\mathcal{K}_3\xi+\mathcal{K}_3\frac{\xi D^{\otimes 3}f_1(x_\theta(N),N)\xi}{\sqrt{G(N)}}+\frac{1}{G(N)^\delta}\phi+\frac{1}{G(N)^\delta}\frac{\xi^T\kappa'\xi}{2\sqrt{G(N)}}\right),$$
$$\text{when }\frac{1}{G(N)}\gg\frac{G(N)}{N},$$
$$G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))=\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi+$$
$$+\frac{G(N)^{3/2-\delta}}{N^{1-\delta}}\left(\mathcal{K}_3\xi+\mathcal{K}_3\frac{\xi D^{\otimes 3}f_3(x_\theta(N),N)\xi}{\sqrt{G(N)}}+\frac{N^{1-\delta}}{G(N)^{2-\delta}}\left(\phi+\frac{\xi^T\kappa'\xi}{2\sqrt{G(N)}}\right)\right),$$
$$\text{when }\frac{1}{\sqrt{G(N)}}\gg\frac{G(N)}{N},$$

where we omitted the case when $\frac{1}{\sqrt{G(N)}}\ll\frac{G(N)}{N}$ as the remainder would overcome the term $\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi$ which would contradict the theorem.

Then we introduce constants $K_1^{iv}, K_2^{iv}, K_3^{iv}$ and exponentiate the expressions

$$1)\ e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N))} = e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi + N^{-1/2+\delta}K_1^{iv}},$$

$$\text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$

$$e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N))} = e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi + \frac{N^{3/2-\delta}}{G(N)^{1-\delta}}K_1^{iv}},$$

$$\text{when } \frac{1}{\sqrt{N}} \gg \frac{N}{G(N)},$$

$$2)\ e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))} = e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi + N^{-1/2+\delta}K_2^{iv}}.$$

$$3)\ e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))} = e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi + G(N)^{-1/2+\delta}K_3^{iv}},$$

$$\text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$

$$e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))} = e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi + \frac{G(N)^{3/2-\delta}}{N^{1-\delta}}K_3^{iv}},$$

$$\text{when } \frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N}.$$

Further, we take Taylor approximation of the RHS at the point $e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi}$ and consequently obtain

$$1)\ \left|e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N))} - e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi}\right| =$$

$$= \frac{|K_1^{iv}|}{N^{1/2-\delta}} e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi + \theta_N N^{-1/2+\delta}K_1^{iv}},$$

$$\text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$

$$\left|e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N))} - e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi}\right| =$$

$$= \frac{N^{3/2-\delta}}{G(N)^{1-\delta}}|K_1^{iv}| e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi + \theta_N \frac{N^{3/2-\delta}}{G(N)^{1-\delta}}K_1^{iv}},$$

$$\text{when } \frac{1}{\sqrt{N}} \gg \frac{N}{G(N)},$$

$$2)\ \left|e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))} - e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi}\right| =$$
$$= \frac{K_2^{iv}}{N^{1/2-\delta}} e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi + \theta_N N^{-1/2+\delta} K_1^{iv}},$$

$$3)\ \left|e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))} - e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi}\right| =$$
$$= \frac{K_3^{iv}}{G(N)^{1/2-\delta}} e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi + \theta_N G(N)^{-1/2+\delta} K_3^{iv}},$$
when $\dfrac{1}{G(N)} \gg \dfrac{G(N)}{N}$,
$$\left|e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))} - e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi}\right| =$$
$$= \frac{G(N)^{3/2-\delta}}{N^{1-\delta}} K_3^{iv} e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi + \theta_N \frac{G(N)^{3/2-\delta}}{N^{1-\delta}} K_3^{iv}},$$
when $\dfrac{1}{\sqrt{G(N)}} \gg \dfrac{G(N)}{N}$,

and we introduce constants K_1^v, K_2^v, K_3^v such that

$$1)\ \left|e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N))} - e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi}\right| = \frac{K_1^v}{N^{1/2-\delta}},$$
when $\dfrac{1}{N} \gg \dfrac{N}{G(N)}$,
$$\left|e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N))} - e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi}\right| = \frac{N^{3/2-\delta}}{G(N)^{1-\delta}} K_1^v,$$
when $\dfrac{1}{\sqrt{N}} \gg \dfrac{N}{G(N)}$,

$$2)\ \left|e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))} - e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi}\right| = \frac{K_2^v}{N^{1/2-\delta}},$$

$$3)\ \left|e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))} - e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi}\right| = \frac{K_3^v}{G(N)^{1/2-\delta}},$$
$$\text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$
$$\left|e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))} - e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi}\right| = \frac{G(N)^{3/2-\delta}}{N^{1-\delta}}K_3^v,$$
$$\text{when } \frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N}.$$

Now combine above estimates with (2.15) and as a result we get

$$1)\ \left|M_{Y_N}(\xi) - e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi}\right| = \frac{K_1}{N^{1/2-\delta}},\ \text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$
$$\left|M_{Y_N}(\xi) - e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi}\right| = \frac{N^{3/2-\delta}}{G(N)^{1-\delta}}K_1, \text{when } \frac{1}{\sqrt{N}} \gg \frac{N}{G(N)},$$

$$2)\ \left|M_{Y_N}(\xi) - e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi}\right| = \frac{K_2}{N^{1/2-\delta}},$$

$$3)\ \left|M_{Y_N}(\xi) - e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi}\right| = \frac{K_3}{G(N)^{1/2-\delta}},\ \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$
$$\left|M_{Y_N}(\xi) - e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi}\right| = \frac{G(N)^{3/2-\delta}}{N^{1-\delta}}K_3, \text{when } \frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N},$$

where the constants are equal

$$1)\ K_1 = K_1^v + \frac{1}{N^\delta}K_1^{iii}\left(e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi} + \frac{K_1^v}{N^{1/2-\delta}}\right), \text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$
$$K_1 = K_1^v + \frac{G(N)^{1-\delta}}{N^{2-\delta}}K_1^{iii}\left(e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi} + \frac{N^{3/2-\delta}}{G(N)^{1-\delta}}K_1^v\right),$$
$$\text{when } \frac{1}{\sqrt{N}} \gg \frac{N}{G(N)},$$

$$2)\ K_2 = K_2^v + \frac{1}{N^\delta}K_2^{iii}\left(e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi} + \frac{K_2^v}{N^{1/2-\delta}}\right),$$

3) $K_3 = K_3^v + \frac{1}{G(N)^\delta} K_3^{iii}\left(e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi} + \frac{K_3^v}{G(N)^{1/2-\delta}}\right)$, when $\frac{1}{G(N)} \gg \frac{G(N)}{N}$,

$K_3 = K_3^v + \frac{N^{1-\delta}}{G(N)^{2-\delta}} K_3^{iii}\left(e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi} + \frac{G(N)^{3/2-\delta}}{N^{1-\delta}} K_3^v\right)$,

when $\frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N}$.

□

2.4 Fluctuation theorem, maximum on the boundary of the domain

The case of fluctuation theorem, when mean is on the boundary of the sample space is a content of this section. As previously, the proofs are done through convergence of the mgfs. Also, the estimates for the speed of convergence to the limiting distribution is included.

Theorem 5. For each case $G(N)$ given by (2.8) we have a m-dimensional random vector Y_N such that

$$1)\ Y_N = N(X_1 - x_1^*) + \sqrt{N}(\hat{X}_N - \hat{x}^*),$$
$$2)\ Y_N = N(X_1 - x_1^*) + \sqrt{N}(\hat{X}_N - \hat{x}^*),$$
$$3)\ Y_N = G(N)(X_1 - x_1^*) + \sqrt{G(N)}(\hat{X}_N - \hat{x}^*),$$

defined on the discrete sample space Ω_E with pmf specified by (2.6), where $x = (x_1, \hat{x})$.
Then for sample space parameter $g_{im}\varepsilon_{im} < E < \overline{g\varepsilon}$, as $N \to \infty$ the distribution of the random vector Y converges to the mixture of the multivariate normal along $\hat{x}$ and exponential distribution along x_1.

Furthermore, we have estimates

$$1)\ M_{Y_N}(\xi) = \frac{|f_1'(x^*)|}{|f_1'(x^*)+\xi|} e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi} + O\bigg(\frac{1}{N^{1/2-\delta}}\bigg),$$
$$\text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$
$$M_{Y_N}(\xi) = \frac{|f_1'(x^*)|}{|f_1'(x^*)+\xi|} e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi} + O\bigg(\frac{N^{3/2-\delta}}{G(N)^{1-\delta}}\bigg),$$
$$\text{when } \frac{1}{\sqrt{N}} \gg \frac{N}{G(N)},$$
$$2)\ M_{Y_N}(\xi) = \frac{|f_2'(x^*)|}{|f_2'(x^*)+\xi|} e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi} + O\bigg(\frac{1}{N^{1/2-\delta}}\bigg),$$
$$3)\ M_{Y_N}(\xi) = \frac{|f_3'(x^*)|}{|f_3'(x^*)+\xi|} e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi} + O\bigg(\frac{1}{G(N)^{1/2-\delta}}\bigg),$$
$$\text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$
$$M_{Y_N}(\xi) = \frac{|f_3'(x^*)|}{|f_3'(x^*)+\xi|} e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi} + O\bigg(\frac{G(N)^{3/2-\delta}}{N^{1-\delta}}\bigg),$$
$$\text{when } \frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N},$$

as $N \to \infty$, where $f_l'(x^*)$ is first derivative w.r.t. x_1 and $D^2 f_l(x^*)$ is $m-1$-dimensional matrix of second order derivatives w.r.t. $\hat{x}$.

Proof. The approach of proving is similar as in the proof of the previous limit theorems.

We approximate the numerator and the denominator of the mgf with appropriate theorems and lemmas and eventually obtain

$$1)\ \left| M_{Y_N}(\xi) - e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N) - f_1(x^*(N),N))} \frac{|f_1'(x^*(N),N)|}{|f_1'(\widetilde{x}^*(N),N)+\xi|} \frac{\sqrt{\det D^2 f_1(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_1(\widetilde{x}^*(N),N)}} \right| \leq$$
$$\leq K_1^i \frac{1}{\sqrt{N}} e^{N(\widetilde{f}_1(\widetilde{x}^*(N),N) - f_1(x^*(N),N))} \frac{|f_1'(x^*(N),N)|}{|f_1'(\widetilde{x}^*(N),N)+\xi|} \frac{\sqrt{\det D^2 f_1(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_1(\widetilde{x}^*(N),N)}},$$

$$2)\ \left|M_{Y_N}(\xi) - e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))}\frac{|f_2'(x^*(N),N)|}{|f_2'(\widetilde{x}^*(N),N)+\xi|}\frac{\sqrt{\det D^2 f_2(x^*(N),N)}}{\sqrt{\det D^2\widetilde{f}_2(\widetilde{x}^*(N),N)}}\right| \leq$$

$$\leq K_2^i\frac{1}{\sqrt{N}}e^{N(\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N))}\frac{|f_2'(x^*(N),N)|}{|f_2'(\widetilde{x}^*(N),N)+\xi|}\frac{\sqrt{\det D^2 f_2(x^*(N),N)}}{\sqrt{\det D^2\widetilde{f}_2(\widetilde{x}^*(N),N)}},$$

$$3)\ \left|M_{Y_N}(\xi) - e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))}\frac{|f_3'(x^*(N),N)|}{|f_3'(\widetilde{x}^*(N),N)+\xi|}\frac{\sqrt{\det D^2 f_3(x^*(N),N)}}{\sqrt{\det D^2\widetilde{f}_3(\widetilde{x}^*(N),N)}}\right| \leq$$

$$\leq K_3^i\frac{1}{\sqrt{G(N)}}\frac{|f_3'(x^*(N),N)|}{|f_3'(\widetilde{x}^*(N),N)+\xi|}e^{G(N)(\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N))}\frac{\sqrt{\det D^2 f_3(x^*(N),N)}}{\sqrt{\det D^2\widetilde{f}_3(\widetilde{x}^*(N),N)}}. \tag{2.17}$$

where $\widetilde{f}_l(x,N) = f_l(x,N) + N(X_1 - x_1^*) + \sqrt{N}(\hat{X}_N - \hat{x}^*)$ and with $G(N)$ for the third case. Then the determinants are approximated analogically to previous limit theorem, using Proposition 2 from Section 5.

Since the maximum along x_1 has no dependence on N, component $N(X_1 - x^*)$ in the function $\widetilde{f}_l(x,N)$ vanish and we can approximate the expression in the exponents similarly as for the previous fluctuation theorem, i.e. by Proposition 1 from Section 5. Hence, we get

$$1)\ \left|M_{Y_N}(\xi) - \frac{|f_1'(x^*(N),N)|}{|\widetilde{f}_1'(\widetilde{x}^*(N),N)+\xi|}e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi}\right| \leq K_1^{ii}\frac{1}{N^{1/2-\delta}}\frac{|f_1'(x^*(N),N)|}{|\widetilde{f}_1'(\widetilde{x}^*(N),N)+\xi|},$$

when $\dfrac{1}{N} \gg \dfrac{N}{G(N)}$,

$$\left|M_{Y_N}(\xi) - \frac{|f_1'(x^*(N),N)|}{|\widetilde{f}_1'(\widetilde{x}^*(N),N)+\xi|}e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi}\right| \leq K_1^{ii}\frac{N^{3/2-\delta}}{G(N)^{1-\delta}}\frac{|f_1'(x^*(N),N)|}{|\widetilde{f}_1'(\widetilde{x}^*(N),N)+\xi|},$$

when $\dfrac{1}{\sqrt{N}} \gg \dfrac{N}{G(N)}$,

$$2)\ \left|M_{Y_N}(\xi) - \frac{|f_2'(x^*(N),N)|}{|\widetilde{f}_2'(\widetilde{x}^*(N),N)+\xi|}e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi}\right| \leq K_2^{ii}\frac{1}{N^{1/2-\delta}}\frac{|f_2'(x^*(N),N)|}{|\widetilde{f}_2'(\widetilde{x}^*(N),N)+\xi|},$$

3) $\left|M_{Y_N}(\xi) - \frac{|f_3'(x^*(N),N)|}{|\widetilde{f}_3'(\widetilde{x}^*(N),N)+\xi|}e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi}\right| \le K_3^{ii}\frac{1}{G(N)^{1/2-\delta}}\frac{|f_3'(x^*(N),N)|}{|\widetilde{f}_3'(\widetilde{x}^*(N),N)+\xi|}$,

when $\frac{1}{G(N)} \gg \frac{G(N)}{N}$,

$\left|M_{Y_N}(\xi) - \frac{|f_3'(x^*(N),N)|}{|\widetilde{f}_3'(\widetilde{x}^*(N),N)+\xi|}e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi}\right| \le K_3^{ii}\frac{G(N)^{3/2-\delta}}{N^{1-\delta}}\frac{|f_3'(x^*(N),N)|}{|\widetilde{f}_3'(\widetilde{x}^*(N),N)+\xi|}$,

when $\frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N}$.

Since the first derivative is along x_1 and function f_l is of the form $f_l(x,N) = \sum_{i=1}^m f(x_i,N)$, and maximum along x_1 is independent of N we have $f_3'(\widetilde{x}^*(N),N) = f_3'(x^*,N)$ and $\widetilde{f}_3'(x^*(N),N) = f_3'(\widetilde{x}^*,N)$. Further, using the Proposition 4 from the Section 3 of Chapter III we have estimate

$$\frac{|f_l'(x^*(N),N)|}{|\widetilde{f}_l'(\widetilde{x}^*(N),N)+\xi|} = \frac{|f_l'(x^*)+\frac{K_l^{iii}}{N}|}{|f_l'(x^*)+\frac{K_l^{iii}}{N}+\xi|},$$

which after some manipulations is equal

$$\frac{|f_l'(x^*(N),N)|}{|\widetilde{f}_l'(\widetilde{x}^*(N),N)+\xi|} = \frac{|f_l'(x^*)|}{|f_l'(x^*)+\xi|} + K_l^{iv}\frac{1}{N},$$

valid for first two cases and for the third we have

$$\frac{|f_3'(x^*(N),N)|}{|\widetilde{f}_3'(\widetilde{x}^*(N),N)+\xi|} = \frac{|f_3'(x^*)|}{|f_3'(x^*)+\xi|} + K_l^{iv}\frac{1}{G(N)}.$$

Then we substitute above estimates into main estimate and get

1) $\left|M_{Y_N}(\xi) - \frac{|f_1'(x^*)|}{|f_1'(x^*)+\xi|}e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi}\right| \le K_1\frac{1}{N^{1/2-\delta}}$,

when $\frac{1}{N} \gg \frac{N}{G(N)}$,

$\left|M_{Y_N}(\xi) - \frac{|f_1'(x^*)|}{|f_1'(x^*)+\xi|}e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi}\right| \le K_1\frac{N^{3/2-\delta}}{G(N)^{1-\delta}}$,

when $\frac{1}{\sqrt{N}} \gg \frac{N}{G(N)}$,

$$2)\ \left| M_{Y_N}(\xi) - \frac{|f_2'(x^*)|}{|f_2'(x^*)+\xi|} e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi} \right| \leq K_2 \frac{1}{N^{1/2-\delta}},$$

$$3)\ \left| M_{Y_N}(\xi) - \frac{|f_3'(x^*)|}{|f_3'(x^*)+\xi|} e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi} \right| \leq K_3 \frac{1}{G(N)^{1/2-\delta}},$$

$$\text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$

$$\left| M_{Y_N}(\xi) - \frac{|f_3'(x^*)|}{|f_3'(x^*)+\xi|} e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi} \right| \leq K_3 \frac{G(N)^{3/2-\delta}}{N^{1-\delta}},$$

$$\text{when } \frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N},$$

where constants K_1, K_2, K_3 are

$$1)\ K_1 = K_1^{ii} \frac{|f_1'(x^*)|}{|f_1'(x^*)+\xi|} + K_1^{iv} \frac{1}{N^{1/2+\delta}} e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi} + \frac{K_1^{ii} K_1^{iv}}{N},$$

$$\text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$

$$K_1 = K_1^{ii} \frac{|f_1'(x^*)|}{|f_1'(x^*)+\xi|} + K_1^{iv} \frac{G(N)^{1-\delta}}{N^{5/2-\delta}} e^{\frac{1}{2}\xi^T D^2 f_1(x^*)^{-1}\xi} + \frac{K_1^{ii} K_1^{iv}}{N},$$

$$\text{when } \frac{1}{\sqrt{N}} \gg \frac{N}{G(N)},$$

$$2)\ K_2 = K_2^{ii} \frac{|f_2'(x^*)|}{|f_2'(x^*)+\xi|} + K_2^{iv} \frac{1}{N^{1/2+\delta}} e^{\frac{1}{2}\xi^T D^2 f_2(x^*)^{-1}\xi} + \frac{K_2^{ii} K_2^{iv}}{N},$$

$$3)\ K_3 = K_3^{ii} \frac{|f_3'(x^*)|}{|f_3'(x^*)+\xi|} + K_3^{iv} \frac{1}{G(N)^{1/2+\delta}} e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi} + \frac{K_3^{ii} K_3^{iv}}{G(N)},$$

$$\text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$

$$K_3 = K_3^{ii} \frac{|f_3'(x^*)|}{|f_3'(x^*)+\xi|} + K_3^{iv} \frac{N^{1-\delta}}{G(N)^{5/2-\delta}} e^{\frac{1}{2}\xi^T D^2 f_3(x^*)^{-1}\xi} + \frac{K_3^{ii} K_3^{iv}}{G(N)},$$

$$\text{when } \frac{1}{\sqrt{G(N)}} \gg \frac{G(N)}{N},$$

hence we get the final result. $\square$

2.5 Related results

This section is devoted to some related estimated, used in the previous two section. It is separated from the rest for the transparency of the proof.

Proposition 1. For the function $f_l(x,N) : \Omega \times \mathbb{N} \to \mathbb{R}$, $l = 1,2,3$, which has a unique maximum at the critical point $x^*(N)$ over the domain $\Omega \in \mathbb{R}^m$, we have another function defined for each case of (2.8)

$$\begin{aligned}
&1)\ \widetilde{f}_1(x,N) = f(x,N) + \frac{1}{\sqrt{N}}(x-x^*)^T\xi,\\
&2)\ \widetilde{f}_2(x,N) = f(x,N) + \frac{1}{\sqrt{N}}(x-x^*)^T\xi,\\
&3)\ \widetilde{f}_3(x,N) = f(x,N) + \frac{1}{\sqrt{G(N)}}(x-x^*)^T\xi,
\end{aligned}$$

where ξ has a value from the neighborhood of 0 and $x^* \in \Omega$. Then we have following estimates

$$\begin{aligned}
&1)\ \widetilde{x}^*(N) - x^*(N) = D^2 f_1(x^*(N),N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N},\\
&|\epsilon_N| \le \frac{F^{(3)}_{x_\theta}}{\left(F'^{(2)}_{x^*(N)}\right)^3}\left(|\xi| + \frac{F^3_{x_\theta}}{N^{1/6}}\right),
\end{aligned}$$

$$\begin{aligned}
&2)\ \widetilde{x}^*(N) - x^*(N) = D^2 f_2(x^*(N),N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N},\\
&|\epsilon_N| \le \frac{F^{(3)}_{x_\theta}}{\left(F'^{(2)}_{x^*(N)}\right)^3}\left(|\xi| + \frac{F^3_{x_\theta}}{N^{1/6}}\right),
\end{aligned}$$

$$\begin{aligned}
&3)\ \widetilde{x}^*(N) - x^*(N) = D^2 f_3(x^*(N),N)^{-1}\frac{\xi}{\sqrt{G(N)}} + \frac{\epsilon_N}{G(N)},\\
&|\epsilon_N| \le \frac{F^{(3)}_{x_\theta}}{\left(F'^{(2)}_{x^*(N)}\right)^3}\left(|\xi| + \frac{F^3_{x_\theta}}{G(N)^{1/6}}\right),
\end{aligned}$$

where $\widetilde{x}^*(N)$ is maximum of the function $\widetilde{f}(x,N)$.

Further, we have estimates

1) $\widetilde{f}_1(\widetilde{x}^*(N),N)-f_1(x^*(N),N)=\dfrac{1}{\sqrt{N}}(x^*(N)-x^*)^T\left(\xi+\dfrac{\xi D^{\otimes 3}f_1(x_\theta(N),N)\xi}{\sqrt{N}}\right)+$
$+\dfrac{1}{2N}\xi^T D^2 f_1(x^*)^{-1}\xi+\dfrac{1}{N^{3/2}}\left(\phi+\dfrac{\xi^T\kappa'\xi}{2\sqrt{N}}\right).$

2) $\widetilde{f}_2(\widetilde{x}^*(N),N)-f_2(x^*(N),N)=\dfrac{1}{\sqrt{N}}(x^*(N)-x^*)^T\left(\xi+\dfrac{\xi D^{\otimes 3}f_2(x_\theta(N),N)\xi}{\sqrt{N}}\right)+$
$+\dfrac{1}{2N}\xi^T D^2 f_2(x^*)^{-1}\xi+\dfrac{1}{N^{3/2}}\left(\phi+\dfrac{\xi^T\kappa'\xi}{2\sqrt{N}}\right).$

3) $\widetilde{f}_3(\widetilde{x}^*(N),N)-f_3(x^*(N),N)=\dfrac{1}{\sqrt{G(N)}}(x^*(N)-x^*)^T\left(\xi+\dfrac{\xi D^{\otimes 3}f_3(x_\theta(N),N)\xi}{\sqrt{G(N)}}\right)+$
$+\dfrac{1}{2G(N)}\xi^T D^2 f_3(x^*)^{-1}\xi+\dfrac{1}{G(N)^{3/2}}\left(\phi+\dfrac{\xi^T\kappa'\xi}{2\sqrt{G(N)}}\right).$

where

1) $|\phi|\le\dfrac{F^{(3)}_{x_\theta}}{6}\left|D^2 f_1(x^*(N),N)^{-1}\xi+\dfrac{\epsilon_N}{\sqrt{N}}\right|^3+\left|2\epsilon_N^T\xi+\dfrac{1}{2N^{1/2}}\epsilon_N^T D^2 f_1(x^*(N),N)\epsilon_N\right|,$
2) $|\phi|\le\dfrac{F^{(3)}_{x_\theta}}{6}\left|D^2 f_2(x^*(N),N)^{-1}\xi+\dfrac{\epsilon_N}{\sqrt{N}}\right|^3+\left|2\epsilon_N^T\xi+\dfrac{1}{2N^{1/2}}\epsilon_N^T D^2 f_2(x^*(N),N)\epsilon_N\right|,$
3) $|\phi|\le\dfrac{F^{(3)}_{x_\theta}}{6}\left|D^2 f_3(x^*(N),N)^{-1}\xi+\dfrac{\epsilon_N}{\sqrt{G(N)}}\right|^3+\left|2\epsilon_N^T\xi+\dfrac{1}{2G(N)^{1/2}}\epsilon_N^T D^2 f_3(x^*(N),N)\epsilon_N\right|$

Proof. First we prove for the case 1) and 2) of $G(N)$ and we drop the index of case in f temporally. Let us take some x_B such that $|x_B-x^*(N)|=N^{-1/3}$. Then we approximate $Df(x,N)$ with first order Taylor expansion at $x^*(N)$

$$Df(x,N)=Df(x^*(N),N)+D^2f(x_\theta(N))^T(x-x^*(N)).$$

Since $x^*(N)$ is a critical point, $Df(x^*(N),N)=o$. Taking the upper bound of the above expansion, together with the absolute value applied on both sided gives

$$|Df(x,N)|\ge F'^{(2)}_{x_\theta(N)}|x-x^*(N)|,$$

and for the point x_B we have

$$|Df(x_B, N)| \geq \frac{F'^{(2)}_{x_\theta(N)}}{N^{1/3}}. \tag{2.18}$$

Further, for the maximum $\widetilde{x}^*(N)$ we have

$$Df(\widetilde{x}^*(N), N) + \frac{\xi}{\sqrt{N}} = o,$$

hence, after changing the side of second term and applying absolute value on the both sides we obtain

$$|Df(\widetilde{x}^*(N), N)| = \frac{|\xi|}{\sqrt{N}}, \tag{2.19}$$

Since $f \in C^2$ and $\det D^2 f(x^*(N, N)) \neq 0$, by the inverse function theorem we have that the mapping $\mathcal{M} : x \to Df(x, N)$ is invertible in the neighborhood of $x^*(N)$ and the inverse function is in the class C^1. Hence, from the estimates (2.18) and (2.19) together with $Df(x^*(N), N) = o$ and knowing that $|x_B - x^*(N)| = N^{-1/3}$ we can infer following estimate

$$|\widetilde{x}^*(N) - x^*(N)| \leq N^{-1/3}. \tag{2.20}$$

Further we use second order Taylor expansion of $D\widetilde{f}(\widetilde{x}^*(N), N)$ at $x^*(N)$

$$\begin{aligned} D\widetilde{f}(\widetilde{x}^*(N), N) =& Df(x^*(N), N) - \frac{\xi}{\sqrt{N}} + D^2 f(x^*(N), N)^T (\widetilde{x}^*(N) - x^*(N)) + \\ & + D^3 f(x_\theta(N))(\widetilde{x}^*(N) - x^*(N))^{\otimes 2}, \end{aligned} \tag{2.21}$$

where $x_\theta(N)$ is some point between $\widetilde{x}^*(N)$ and $x^*(N)$. Notice, that as $D\widetilde{f}(\widetilde{x}^*(N), N) = o$ and $Df(x^*(N), N) = o$.

Now we get the upper bound of expansion (2.21) and apply absolute value on both sides

$$\left| D^2 f(x^*(N), N)^T (\widetilde{x}^*(N) - x^*(N)) - \frac{\xi}{\sqrt{N}} \right| \leq F^{(3)}_{x_\theta(N)} |\widetilde{x}^*(N) - x^*(N)|^2, \tag{2.22}$$

and use estimate (2.20)

$$\left|D^2 f(x^*(N),N)^T(\widetilde{x}^*(N)-x^*(N))-\frac{\xi}{\sqrt{N}}\right| \le F^{(3)}_{x_\theta(N)}N^{-2/3}. \tag{2.23}$$

Next we get the lower bound for (2.21)

$$\left|D^2 f(x^*(N),N)^T(\widetilde{x}^*(N)-x^*(N))-\frac{\xi}{\sqrt{N}}\right| \ge F'^{(2)}_{x^*(N)}|\widetilde{x}^*(N)-x^*(N)|-\left|\frac{\xi}{\sqrt{N}}\right|. \tag{2.24}$$

Next, we combine upper (2.23) and lower bound (2.24) into one inequality

$$F'^{(2)}_{x^*(N)}|\widetilde{x}^*(N)-x^*(N)|-\left|\frac{\xi}{\sqrt{N}}\right| \le F^{(3)}_{x_\theta(N)}N^{-2/3},$$

and after some manipulations one get estimate

$$|\widetilde{x}^*(N)-x^*(N)| \le \frac{\left(F'^{(2)}_{x^*(N)}\right)^{-1}}{\sqrt{N}}\left(|\xi|+F^{(3)}_{x_\theta(N)}N^{-1/6}\right),$$

Next we substitute above estimate into (2.22)

$$\left|D^2 f(x^*(N),N)^T(\widetilde{x}^*(N)-x^*(N))-\frac{\xi}{\sqrt{N}}\right| \le \frac{F^{(3)}_{x_\theta(N)}\left(F'^{(2)}_{x^*(N)}\right)^{-2}}{N}\left(|\xi|+F^{(3)}_{x_\theta(N)}N^{-1/6}\right)^2. \tag{2.25}$$

Further, we transform the LHS of the inequality

$$\begin{aligned}
&\left|D^2 f(x^*(N),N)^T(\widetilde{x}^*(N)-x^*(N))-\frac{\xi}{\sqrt{N}}\right| = \\
&= \left|D^2 f(x^*(N),N)^T(\widetilde{x}^*(N)-x^*(N))-D^2 f(x^*(N),N)^T\left(D^2 f(x^*(N),N)^T\right)^{-1}\frac{\xi}{\sqrt{N}}\right| = \\
&= \left|D^2 f(x^*(N),N)^T\left((\widetilde{x}^*(N)-x^*(N))-\left(D^2 f(x^*(N),N)^T\right)^{-1}\frac{\xi}{\sqrt{N}}\right)\right|,
\end{aligned}$$

calculate the lower bound of it

$$\begin{aligned}
&\left|D^2 f(x^*(N),N)^T\left((\widetilde{x}^*(N)-x^*(N))-\left(D^2 f(x^*(N),N)^T\right)^{-1}\frac{\xi}{\sqrt{N}}\right)\right| \ge \\
&\ge F'^{(2)}_{x^*(N)}\left|(\widetilde{x}^*(N)-x^*(N))-\left(D^2 f(x^*(N),N)^T\right)^{-1}\frac{\xi}{\sqrt{N}}\right|
\end{aligned}$$

and substitute back into (2.25) and dividing by $F'^{(2)}_{x^*(N)}$

$$\begin{aligned}
&\left| \widetilde{x}^*(N) - x^*(N) - \left(D^2 f(x^*(N),N)^T\right)^{-1} \frac{\xi}{\sqrt{N}} \right| \le \\
&\le \frac{1}{N} F^{(3)}_{x_\theta(N)} \left(F'^{(2)}_{x^*(N)}\right)^{-3} \left(|\xi| + F^{(3)}_{x_\theta(N)} N^{-1/6} \right)^2 .
\end{aligned}$$

Since $D^2 f(x^*(N),N)$ is a matrix of second derivatives of the continuous function $D^2 f(x^*(N),N)^T = D^2 f(x^*(N),N)$, hence we get the first result of the theorem, i.e

$$\begin{aligned}
\widetilde{x}^*(N) =& x^*(N) + D^2 f(x^*(N),N)^{-1} \frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}, \\
|\epsilon_N| \le & \frac{F^{(3)}_{x_\theta(N)}}{\left(F'^{(2)}_{x^*(N)}\right)^3} \left(|\xi| + \frac{F^{(3)}_{x_\theta(n)}}{N^{1/6}} \right),
\end{aligned}$$

Now, we derive the second result of the theorem. We substitute the above estimate into $f(\widetilde{x}^*(N),N)$, expand the function using 3-rd order Taylor expansion at $x^*(N)$ and apply absolute value

$$\begin{aligned}
&f(\widetilde{x}^*(N),N) = f\left(x^*(N) + D^2 f(x^*(N),N)^{-1} \frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}, N \right) = \\
&= f(x^*(N),N) + Df(x^*(N),N) \left(D^2 f(x^*(N),N)^{-1} \frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N} \right) + \\
&+ \frac{1}{2} \left(D^2 f(x^*(N),N)^{-1} \frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N} \right)^T D^2 f(x^*(N),N) \left(D^2 f(x^*(N),N)^{-1} \frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N} \right) + \\
&+ \frac{1}{6} \left\langle D^{\otimes 3} f(x_\theta(N),N), \left(D^2 f(x^*(N),N)^{-1} \frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N} \right)^{\otimes 3} \right\rangle,
\end{aligned}$$

where $x_\theta(N)$ is somwhere between $\tilde{x}^*(N)$ and $x^*(N)$.

Then we change side of some terms and apply absolute value

$$\Bigg| f(\widetilde{x}^*(N), N) = f\Bigg(x^*(N) + D^2 f(x^*(N), N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}, N\Bigg) = -$$

$$- f(x^*(N), N) - Df(x^*(N), N)\Bigg(D^2 f(x^*(N), N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}\Bigg) -$$

$$- \frac{1}{2}\Bigg(D^2 f(x^*(N), N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}\Bigg)^T D^2 f(x^*(N), N)\Bigg(D^2 f(x^*(N), N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}\Bigg) -$$

$$- \frac{1}{6}\Bigg\langle D^{\otimes 3} f(_\theta(N), N), \Bigg(D^2 f(x^*(N), N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}\Bigg)^{\otimes 3}\Bigg\rangle\Bigg| = 0,$$

and we calculate the upper bound and as $Df(x^*(N), N) = o$ we have

$$\Bigg| f(\widetilde{x}^*(N), N) - f(x^*(N), N) -$$

$$- \frac{1}{2}\Bigg(D^2 f(x^*(N), N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}\Bigg)^T D^2 f(x^*(N), N)\Bigg(D^2 f(x^*(N), N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}\Bigg)\Bigg| \leq$$

$$\leq \frac{F^{(3)}_{x_\theta}}{6}\Bigg|D^2 f(x^*(N), N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}\Bigg|^3,$$

consequently after some manipulations we get

$$\Bigg| f(\widetilde{x}^*(N), N) - f(x^*(N), N) - \frac{1}{2N}\xi^T D^2 f(x^*(N), N)^{-1}\xi - \frac{1}{N^{3/2}}\epsilon_N^T \xi -$$

$$- \frac{1}{2N^2}\epsilon_N^T D^2 f(x^*(N), N)\epsilon_N\Bigg| \leq \frac{F^{(3)}_{x_\theta}}{6}\Bigg|D^2 f(x^*(N), N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N}\Bigg|^3. \quad (2.26)$$

Now substitute into explicit expression for $\widetilde{f}(\widetilde{x}^*(N), N)$ into the first result of the theorem

$$\widetilde{f}(\widetilde{x}^*(N), N) = f(\widetilde{x}^*(N), N) + \frac{1}{\sqrt{N}}\xi^T\Bigg(x^*(N) + D^2 f(x^*(N), N)^{-1}\frac{\xi}{\sqrt{N}} + \frac{\epsilon_N}{N} - x^*\Bigg).$$

Then after some manipulations and applying absolute value we get

$$\Bigg|\widetilde{f}(\widetilde{x}^*(N), N) - f(\widetilde{x}^*(N), N) - \frac{1}{\sqrt{N}}(x^*(N) - x^*)^T\xi - \frac{1}{N}\xi^T D^2 f(x^*(N), N)^{-1}\xi - \frac{\xi^T\epsilon_N}{N^{3/2}}\Bigg| = 0,$$

and combining it with inequality (2.26) and applying triangle inequality we get

$$\Big|\widetilde{f}(\widetilde{x}^*(N),N) - f(x^*(N),N) - \frac{1}{2N}\xi^T D^2 f(x^*(N),N)^{-1}\xi - \frac{1}{N^{3/2}}\epsilon_N^T\xi -$$
$$- \frac{1}{2N^2}\epsilon_N^T D^2 f(x^*(N),N)\epsilon_N - \frac{1}{\sqrt{N}}(x^*(N)-x^*)^T\xi - \frac{1}{N}\xi^T D^2 f(x^*(N),N)^{-1}\xi - \frac{\xi^T\epsilon_N}{N^{3/2}}\Big| \leq$$
$$\leq \frac{F^{(3)}_{x_\theta}}{6}\left|D^2 f(x^*(N),N)^{-1}\frac{\xi}{\sqrt{N}} - \frac{\epsilon_N}{N}\right|^3,$$

which after some simplifications is

$$\left|\widetilde{f}(\widetilde{x}^*(N),N) - f(x^*(N),N) - \frac{1}{\sqrt{N}}(x^*(N)-x^*)^T\xi - \frac{3}{2N}\xi^T D^2 f(x^*(N),N)^{-1}\xi\right| \leq$$
$$\leq \frac{1}{N^{3/2}}\left(\frac{F^{(3)}_{x_\theta}}{6}\left|D^2 f(x^*(N),N)^{-1}\xi + \frac{\epsilon_N}{\sqrt{N}}\right|^3 + \left|2\epsilon_N^T\xi + \frac{1}{2N^{1/2}}\epsilon_N^T D^2 f(x^*(N),N)\epsilon_N\right|\right),$$

and finally we obtain for first two cases

$$\widetilde{f}(\widetilde{x}^*(N),N) - f(x^*(N),N) = \frac{1}{\sqrt{N}}(x^*(N)-x^*)^T\xi +$$
$$+ \frac{1}{2N}\xi^T D^2 f(x^*(N),N)^{-1}\xi + \frac{\phi}{N^{3/2}}, \tag{2.27}$$

where

$$|\phi_N| \leq \frac{F^{(3)}_{x_\theta(N)}}{6}\left|D^2 f(x^*(N),N)^{-1}\xi + \frac{\epsilon_N}{\sqrt{N}}\right|^3 + \left|2\epsilon_N^T\xi + \frac{1}{2N^{1/2}}\epsilon_N^T D^2 f(x^*(N),N)\epsilon_N\right|$$

Now we take first order Taylor expansion of the $D^2 f(x^*(N),N)^{-1}$ at x^*

$$D^2 f(x^*(N),N)^{-1} = D^2 f(x^*,N)^{-1} + (x^*(N)-x^*)D^{\otimes 3} f(x_\theta(N),N),$$

hence for $\xi^T D^2 f(x^*(N),N)^{-1}\xi$

$$D^2 f(x^*(N),N)^{-1} = D^2 f(x^*,N)^{-1} + (x^*(N)-x^*)D^{\otimes 3} f(x_\theta(N),N),$$

we have

$$\xi^T D^2 f(x^*(N),N)^{-1}\xi = \xi^T D^2 f(x^*,N)^{-1}\xi + \xi^T(x^*(N)-x^*)D^{\otimes 3} f(x_\theta(N),N)\xi,$$

Then we put it back into (2.27) and afetr some rearrangements we get

$$\widetilde{f}(\widetilde{x}^*(N), N) - f(x^*(N), N) = \frac{1}{\sqrt{N}}(x^*(N) - x^*)^T \left(\xi + \frac{\xi D^{\otimes 3} f(x_\theta(N), N)\xi}{\sqrt{N}} \right) + \\ + \frac{1}{2N}\xi^T D^2 f(x^*, N)^{-1}\xi + \frac{\phi}{N^{3/2}}.$$

Now we use Lemma 14 form the Section 4, Chapter III to approximate $D^2 f(x^*, N)^{-1}$ and get

$$\widetilde{f}(\widetilde{x}^*(N), N) - f(x^*(N), N) = \frac{1}{\sqrt{N}}(x^*(N) - x^*)^T \left(\xi + \frac{\xi D^{\otimes 3} f(x_\theta(N), N)\xi}{\sqrt{N}} \right) + \\ + \frac{1}{2N}\xi^T D^2 f(x^*)^{-1}\xi + \frac{1}{N^{3/2}}\left(\phi + \frac{\xi^T \kappa' \xi}{2\sqrt{N}} \right),$$

which is the final result for first two cases. For the third case we simple replace N with $G(N)$ everywhere in the estimate

$$\widetilde{f_3}(\widetilde{x}^*(N), N) - f_3(x^*(N), N) = \frac{1}{\sqrt{G(N)}}(x^*(N) - x^*)^T \left(\xi + \frac{\xi D^{\otimes 3} f_3(x_\theta(N), N)\xi}{\sqrt{G(N)}} \right) + \\ + \frac{1}{2G(N)}\xi^T D^2 f_3(x^*)^{-1}\xi + \frac{1}{G(N)^{3/2}}\left(\phi + \frac{\xi^T \kappa' \xi}{2\sqrt{G(N)}} \right),$$

where

$$|\phi_N| \leq \frac{F^{(3)}_{x_\theta}}{6}\left| D^2 f_3(x^*(N), N)^{-1}\xi + \frac{\epsilon_N}{\sqrt{G(N)}} \right|^3 + \left| 2\epsilon_N^T \xi + \frac{1}{2G(N)^{1/2}}\epsilon_N^T D^2 f_3(x^*(N), N)\epsilon_N \right|,$$

hence we get the final result. □

Proposition 2. For the function $f(x, N) : \Omega \times \mathbb{N} \to \mathbb{R}$, which has unique maximum at the critical point $x^*(N)$ over the domain $\Omega \in \mathbb{R}^m$, we have another function defined $\widetilde{f}(x, N) = f(x, N) + \frac{1}{\sqrt{N}}(x - x^*)^T \xi$, where ξ has a value from the neighborhood of 0. Then we have estimate

$$\left| \frac{\sqrt{\det D^2 f(x^*(N), N)}}{\sqrt{\det D^2 \widetilde{f}_\xi(\widetilde{x}^*(N), N)}} - 1 \right| \leq \frac{K}{\sqrt{N}} \tag{2.28}$$

where $K > 0$ is some constant and the matrices in the inequality are diagonal.

Proof. First we take first order Taylor approximation of $D^2 f(x^*(N), N)$ at $\widetilde{x}^*(N)$

$$D^2 f(x^*(N), N) = D^2 \widetilde{f}(\widetilde{x}^*(N), N) + < D^{\otimes 3} f(x_\theta(N), N), (\widetilde{x}^*(N) - x^*(N), N) >,$$

where $D^2 \widetilde{f}(\widetilde{x}^*(N), N) = D^2 f(\widetilde{x}^*(N), N)$ as the term $\frac{1}{\sqrt{N}}(x - x^*)^T \xi$ vanish in the second derivatives of x. Then we bound it and apply absolute vales on the both sides

$$\left| D^2 f(x^*(N), N) - D^2 \widetilde{f}(\widetilde{x}^*(N), N) \right| \leq F_\theta^{(3)} |\widetilde{x}^*(N) - x^*(N)|.$$

Next we substitute the estimate for $|\widetilde{x}^*(N) - x^*(N)|$ which estimated by the previous lemma

$$\left| D^2 f(x^*(N), N) - D^2 \widetilde{f}(\widetilde{x}^*(N), N) \right| \leq \frac{F_\theta^{(3)}}{\sqrt{N}} \left| D^2 f(x^*(N), N)^{-1} \xi + \frac{\epsilon N}{\sqrt{N}} \right|. \tag{2.29}$$

Since the function f hence $\widetilde{f}$ is decomposable $f(x, N) = \sum_{i=1}^m f(x_i, N)$ the matrices of the second derivatives are diagonal matrices. The diagonal elements of the matrices $D^2 f(x^*(N), N)$ and $D^2 \widetilde{f}(\widetilde{x}^*(N), N)$ we will denote respectively λ_i and μ_i, $i = 1, \ldots, m$. Then the elements of matrix on the LHS of (2.29) are given by

$$\left[\left| D^2 f(x^*(N), N) - D^2 \widetilde{f}(\widetilde{x}^*(N), N) \right| \right]_{i,j} = |\lambda_i - \mu_i|.$$

Since for the finite dimensional linear operator all norms are equivalent, the norm of the above matrix is equal

$$\left| D^2 f(x^*(N), N) - D^2 \widetilde{f}(\widetilde{x}^*(N), N) \right| = \max_j \sum_{i=1}^m |\lambda_i - \mu_i| = \sum_{i=1}^m |\lambda_i - \mu_i|,$$

hence by (2.29) we have a bound

$$\sum_{i=1}^m |\lambda_i - \mu_i| \leq \frac{F_\theta^{(3)}}{\sqrt{N}} |D^2 f(x^*(N), N)^{-1} \xi + \frac{\epsilon N}{\sqrt{N}}|. \tag{2.30}$$

Now we take LHS of inequality (2.28), write explicitly the matrices determinants

$$\left| \frac{\sqrt{\det D^2 f(x^*(N), N)}}{\sqrt{\det D^2 \widetilde{f}(\widetilde{x}^*(N), N)}} - 1 \right| = \left| \frac{\prod_{i=1}^m \sqrt{\lambda_i}}{\prod_{i=1}^m \sqrt{\mu_i}} - 1 \right|, \tag{2.31}$$

and further
$$\left|\frac{\prod_{i=1}^{m}\sqrt{\lambda_i}}{\prod_{i=1}^{m}\sqrt{\mu_i}}-1\right|=\left|\frac{\prod_{i=1}^{m}\sqrt{\lambda_i}-\prod_{i=1}^{m}\sqrt{\mu_i}}{\prod_{i=1}^{m}\sqrt{\mu_i}}\right|.$$
Then we transform the numerator by adding and deducing $\sqrt{\mu_1}\prod_{i=2}^{m}\sqrt{\lambda_i}$

$$\prod_{i=1}^{m}\sqrt{\lambda_i}-\sqrt{\mu_1}\prod_{i=2}^{m}\sqrt{\lambda_i}+\sqrt{\mu_1}\prod_{i=2}^{m}\sqrt{\lambda_i}-\prod_{i=1}^{m}\sqrt{\mu_i}=$$
$$(\sqrt{\lambda_1}-\sqrt{\mu_1})\prod_{i=2}^{m}\sqrt{\lambda_i}+\sqrt{\mu_1}\prod_{i=2}^{m}\sqrt{\lambda_i}-\prod_{i=1}^{m}\sqrt{\mu_i},$$

and again add and deduce $\sqrt{\mu_1\mu_2}\prod_{i=1}^{m}\sqrt{\lambda_i}$

$$(\sqrt{\lambda_1}-\sqrt{\mu_1})\prod_{i=2}^{m}\sqrt{\lambda_i}+\sqrt{\mu_1}\prod_{i=2}^{m}\sqrt{\lambda_i}-\sqrt{\mu_1\mu_2}\prod_{i=1}^{m}\sqrt{\lambda_i}+\sqrt{\mu_1\mu_2}\prod_{i=1}^{m}\sqrt{\lambda_i}-\prod_{i=1}^{m}\sqrt{\mu_i}=$$
$$=(\sqrt{\lambda_1}-\sqrt{\mu_1})\prod_{i=2}^{m}\sqrt{\lambda_i}+\sqrt{\mu_1}(\sqrt{\lambda_2}-\sqrt{\mu_2})\prod_{i=2}^{m}\sqrt{\lambda_i}+\sqrt{\mu_1\mu_2}\prod_{i=1}^{m}\sqrt{\lambda_i}-\prod_{i=1}^{m}\sqrt{\mu_i}.$$

We repeat this step until we get

$$=\sum_{i=1}^{m}(\sqrt{\lambda_i}-\sqrt{\mu_i})\prod_{j=1}^{i-1}\prod_{k=i}^{m}\sqrt{\mu_j\lambda_k}.$$

hence we have

$$\left|\prod_{i=1}^{m}\sqrt{\lambda_i}-\prod_{i=1}^{m}\sqrt{\mu_i}\right|=\left|\sum_{i=1}^{m}(\sqrt{\lambda_i}-\sqrt{\mu_i})\prod_{j=1}^{i-1}\prod_{k=i}^{m}\sqrt{\mu_j\lambda_k}\right|,$$

and by the triangle inequality and multiplicity of absolute value

$$\left|\prod_{i=1}^{m}\sqrt{\lambda_i}-\prod_{i=1}^{m}\sqrt{\mu_i}\right|\le\sum_{i=1}^{m}\left|\sqrt{\lambda_i}-\sqrt{\mu_i}\right|\prod_{j=1}^{i-1}\prod_{k=i}^{m}|\sqrt{\mu_j\lambda_k}| \tag{2.32}$$

Now, as
$$\lambda_i-\mu_i=\big(\sqrt{\lambda_i}-\sqrt{\mu_i}\big)\big(\sqrt{\lambda_i}+\sqrt{\mu_i}\big),$$
and therefore
$$\left|\sqrt{\lambda_i}-\sqrt{\mu_i}\right|=\frac{|\lambda_i-\mu_i|}{|\sqrt{\lambda_i}+\sqrt{\mu_i}|}.$$

Now combining (2.31) and (2.32) with above expression we get

$$\left|\frac{\sqrt{\det D^2 f(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}(\widetilde{x}^*(N),N)}}-1\right| \le \sum_{i=1}^{m} \frac{|\lambda_i-\mu_i|}{|\sqrt{\lambda_i}+\sqrt{\mu_i}|}\frac{\prod_{j=1}^{i-1}\prod_{k=i}^{m}|\sqrt{\mu_j\lambda_k}|}{\prod_{i=1}^{m}\sqrt{\mu_i}},$$

and after simplifications of last factor we get

$$\left|\frac{\sqrt{\det D^2 f(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}(\widetilde{x}^*(N),N)}}-1\right| \le \sum_{i=1}^{m} \frac{|\lambda_i-\mu_i|}{|\sqrt{\lambda_i}+\sqrt{\mu_i}|}\prod_{k=i}^{m}\sqrt{\frac{\lambda_k}{\mu_k}}.$$

Then we substitute into above inequality (2.30) and eventually obtain

$$\left|\frac{\sqrt{\det D^2 f(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}(\widetilde{x}^*(N),N)}}-1\right| \le \frac{F_\theta^{(3)}}{\sqrt{N}}\left|D^2 f(x^*(N),N)^{-1}\xi+\frac{\epsilon N}{\sqrt{N}}\right|\sum_{i=1}^{m}\frac{1}{|\sqrt{\lambda_i}+\sqrt{\mu_i}|}\prod_{k=i}^{m}\sqrt{\frac{\lambda_k}{\mu_k}}.$$

Since the function $f \in C^2$ the diagonal elements λ_i and μ_i, i.e. the second derivative for any N are bounded from below and above, hence we have a bounding constant

$$K \ge F_\theta^{(3)}\left|D^2 f(x^*(N),N)^{-1}\xi+\frac{\epsilon N}{\sqrt{N}}\right|\sum_{i=1}^{m}\frac{1}{|\sqrt{\lambda_i}+\sqrt{\mu_i}|}\prod_{k=i}^{m}\sqrt{\frac{\lambda_k}{\mu_k}},$$

for some fixed N. The finally we can write

$$\left|\frac{\sqrt{\det D^2 f(x^*(N),N)}}{\sqrt{\det D^2 \widetilde{f}_\xi(\widetilde{x}^*(N),N)}}-1\right| \le \frac{K}{\sqrt{N}}.$$

which is the result of the lemma. $\square$

3 Entropy related results

The entropy plays a key role in the main result of the thesis. By definition it is the number of possible states that can take for system for particular value of the parameters. Hence, it is essential in calculating the probability distribution in the statistical mechanics. This chapter contains the results related to the entropy of the considered system which are developed specifically for the proofs of the limit theorems in the second Chapter.

In the first section we provide an approximation of entropy. The method is based on the approximation of gamma function, which essentially is factorial but for the real numbers. It is much more complicated than for example if we use Stirling approximation but we are able to separate the error term form the approximation.

In the Section 2 we maximize the entropy using the methods of the convex optimization. We find explicit formula for the points of maximum. It turns out that there might be two type of maximum, on the boundary of the domain or in the interior. System parameters related to the domain over which the optimization is performed determines the type of maximum.

Section 4 consists of some related to the approximated entropy estimates. They are crucial for proving the limit theorems of the Chapter II.

The last section is approximation of so-called - partition function. It is the sum of entropies of all possible systems configurations. For the proof of the limit theorem we need this function. We need that it will be in the form of integral instead of sum. Although several attempts was taken to prove given result but the proof turned out to be more complicated than it seems and non of it was fully successful. As the result seems intuitively correct we include it but without proof.

For an easy referencing, we recall some related concepts from the Chapter II, like entropy function or related partition function.
For $N \in \mathbb{N}$, increasing discrete function $G(N) : \mathbb{N} \to \mathbb{N}$ and $x_i \in [0, 1]$, $g_i \in (0, 1)$,

$i = 1, \ldots, m$ we define the entropy $S : x \times N \to \mathbb{R}$ as

$$S(x, N) = \ln \prod_{i=1}^{m} \frac{(x_i N + g_i G(N) - 1)!}{(x_i N)!(g_i G(N) - 1)!} \tag{3.1}$$

where ! defines usual factorial and we have constraints on x_i and g_i, $\sum_{i=1}^{m} g_i = 1$ and $\sum_{i=1}^{m} x_i = 1$.
Moreover, we consider three cases of $G(N)$, distinguished by the behaviour as $N \to \infty$

$$\begin{aligned} &1)\ \frac{G(N)}{N} \to \infty, \\ &2)\ \frac{G(N)}{N} \to\ c, \\ &3)\ \frac{G(N)}{N} \to 0, \end{aligned} \tag{3.2}$$

where c is some positive constant. The corresponding partition function is defined

$$Z(N, E) = \sum_{\Omega_E} e^{S(x,N)}, \tag{3.3}$$

where Ω_E is some set in $\mathbb{R}_+^m$ such that following constraints are valid

$$\begin{aligned} &\sum_{i=1}^{m} x_i = 1, \\ &\sum_{i=1}^{m} \varepsilon_i x_i \leq E, \end{aligned}$$

where $E > 0$ and $\varepsilon_i,\ i = 1, \ldots, m$ are some constants such that $0 < \varepsilon_1 < \varepsilon_2 < \ldots < \varepsilon_m$.

3.1 Entropy approximation

This section contains the approximation of entropy (3.1). It is in the form of the Lemma. The rigorous proofs of it is also provided. The form of the approximated function is tailored such that it is convenient for the application of Laplace approximations performed in the proof of the limit theorem in the Chapter II.

Lemma 1. Following asymptotic equations are valid

$$
\begin{aligned}
&1)\ e^{S(x,N)} = (2\pi)^{-\frac{m}{2}} e^{Nf_1(x,N)+R_1(N)}\left(1+O\left(\frac{1}{N}\right)\right),\\
&2)\ e^{S(x,N)} = (2\pi)^{-\frac{m}{2}} e^{Nf_2(x,N)+R_2(N)}\left(1+O\left(\frac{1}{N}\right)\right),\\
&3)\ e^{S(x,N)} = (2\pi)^{-\frac{m}{2}} e^{G(N)f_3(x,N)+R_3(N)}\left(1+O\left(\frac{1}{G(N)}\right)\right),
\end{aligned}
$$

as $N \to \infty$ where

$$
\begin{aligned}
1)\ f_1(x,N) =& \sum_{i=1}^{m}\left[x_i \ln\frac{g_i}{x_i} + \left(x_i + g_i\frac{G(N)}{N}\right)\ln\left(1+\frac{x_iN}{g_iG(N)}\right) - \right.\\
&\left. - \frac{1}{2N}\ln(x_iN + g_iG(N)) - x_i\ln\left(1+\frac{1}{x_iN}\right) - \frac{1}{2N}\ln(x_iN+1)\right],\\
R_1(N) =& -\sum_{i=1}^{m}\frac{1}{2}\ln g_iG(N) + m + N\ln\frac{G(N)}{N},
\end{aligned}
$$

$$
\begin{aligned}
2)\ f_2(x,N) =& \sum_{i=1}^{m}\left[\left(x_i + g_i\frac{G(N)}{N}\right)\ln\left(x_i + g_i\frac{G(N)}{N}\right) - x_i\ln x_i - \right.\\
&\left. - \frac{1}{2N}\ln(x_iN + g_iG(N)) - x_i\ln\left(1+\frac{1}{x_iN}\right) - \frac{1}{2N}\ln(x_iN+1)\right],\\
R_2(N) =& -\sum_{i=1}^{m}\left[\frac{1}{2}\ln g_iG(N) - g_i\frac{G(N)}{N}\ln g_i\frac{G(N)}{N}\right] + m,
\end{aligned}
$$

$$
\begin{aligned}
3)\ f_3(x,N) =& \sum_{i=1}^{m}\left[g_i\ln x_i + \left(x_i\frac{N}{G(N)} + g_i\right)\ln\left(1+\frac{g_iG(N)}{x_iN}\right)\right] - \\
& - \frac{1}{2G(N)}\ln(x_iN + g_iG(N)) - \frac{x_iN}{G(N)}\ln\left(1+\frac{1}{x_iN}\right) - \frac{1}{2G(N)}\ln(x_iN+1)\Big],\\
R_3(N) =& \sum_{i=1}^{m}\left[\frac{1}{2}\ln g_iG(N) - g_iG(N)\ln g_iG(N)\right] + m + G(N)\ln N.
\end{aligned}
$$

Furthermore we have that

$$
f_l(x,N) \to f_l(x),\ \text{as } N \to \infty, \tag{3.4}
$$

for $l = 1, 2, 3$, where

$$1)\ f_1(x) = \sum_{i=1}^{m} \left[x_i \ln \frac{g_i}{x_i} + x_i \right],$$
$$2)\ f_2(x) = \sum_{i=1}^{m} \left[(x_i + g_i c) \ln(x_i + g_i c) - x_i \ln x_i \right],$$
$$3)\ f_3(x) = \sum_{i=1}^{m} \left[g_i \ln x_i + g_i \right].$$

Proof. Since $\Gamma(N) = (N-1)!$ we can write

$$\frac{(x_i N + g_i G(N) - 1)!}{(x_i N)!(g_i G(N) - 1)!} = \frac{\Gamma(x_i N + g_i G(N))}{\Gamma(x_i N + 1)\Gamma(g_i G(N))},$$

and if we introduce notation

$$\begin{aligned} x_i N + g_i G(N) &= \phi_i(N), \\ x_i N + 1 &= \psi_i(N), \\ g_i G(N) &= \theta_i(N), \end{aligned} \tag{3.5}$$

then

$$\frac{\Gamma(x_i N + g_i G(N))}{\Gamma(x_i N + 1)\Gamma(g_i G(N)} = \frac{\Gamma(\phi_i(N))}{\Gamma(\psi_i(N))\Gamma(\theta_i(N))}.$$

Further, let us introduce

$$\begin{aligned} \Phi_i(N) &= -\phi_i(N) + \left(\phi_i(N) - \frac{1}{2} \right) \ln \phi_i(N), \\ \Psi_i(N) &= -\psi_i(N) + \left(\psi_i(N) - \frac{1}{2} \right) \ln \psi_i(N), \\ \Theta_i(N) &= -\theta_i(N) + \left(\theta_i(N) - \frac{1}{2} \right) \ln \theta_i(N). \end{aligned} \tag{3.6}$$

First order approximation of gamma function by the Theorem 10 in the Appendix A.2, has asymptotic expansion

$$\Gamma(\lambda) \sim e^{-\lambda + (\lambda + \frac{1}{2}) \ln \lambda} \left[1 + \frac{1}{12\lambda} + \frac{1}{288\lambda^2} + \dots \right], \lambda \to \infty.$$

Since $\phi_i(N), \psi_i(N), \theta_i(N)$ are positive and increasing functions of N, they can be approximated using the above asymptotic expansion

$$\Gamma(\varphi_i(N)) \sim e^{\Psi_i(N)}\left[1+\frac{1}{12\varphi_i(N)}+\frac{1}{288\varphi_i(N)^2}+\dots\right],$$
$$\Gamma(\theta_i(N)) \sim e^{\Theta_i(N)}\left[1+\frac{1}{12\theta_i(N)}+\frac{1}{288\theta_i(N)^2}+\dots\right],$$
$$\Gamma(\phi_i(N)) \sim e^{\Phi_i(N)}\left[1+\frac{1}{12\phi_i(N)}+\frac{1}{288\phi_i(N)^2}+\dots\right],$$

where $\Phi_i(N), \Psi_i(N), \Theta_i(N)$ are given by (3.6).
From the definition of asymptotic expansion (Definition 4 in the Appendix A.2) we obtain first order approximation, valid for large N

$$\left|\Gamma(\theta_i(N)) - \sqrt{2\pi}e^{\Psi_i(N)}\right| \le K_{i,\psi}\left|\frac{1}{12\psi_i(N)}\sqrt{2\pi}e^{\Psi_i(N)}\right|,$$
$$\left|\Gamma(\phi_i(N)) - \sqrt{2\pi}e^{\Phi_i(N)}\right| \le K_{i,\phi}\left|\frac{1}{12\phi_i(N)}\sqrt{2\pi}e^{\Phi_i(N)}\right|,$$
$$\left|\Gamma(\theta_i(N)) - \sqrt{2\pi}e^{\Theta_i(N)}\right| \le K_{i,\theta}\left|\frac{1}{12\theta_i(N)}\sqrt{2\pi}e^{\Theta_i(N)}\right|,$$

where $K_{i,\psi}, K_{i,\phi}, K_{i,\theta}$ are some positive constants. If we consider constant $K_i = \max\{K_{i,\psi}, K_{i,\phi}, K_{i,\theta}\}$ then we can represent the above inequalities as

$$\left|\Gamma(\theta_i(N)) - \sqrt{2\pi}e^{\Psi_i(N)}\right| \le K_i\left|\frac{1}{12\psi_i(N)}\sqrt{2\pi}e^{\Psi_i(N)}\right|, \tag{3.7}$$
$$\left|\Gamma(\phi_i(N)) - \sqrt{2\pi}e^{\Phi_i(N)}\right| \le K_i\left|\frac{1}{12\phi_i(N)}\sqrt{2\pi}e^{\Phi_i(N)}\right|, \tag{3.8}$$
$$\left|\Gamma(\theta_i(N)) - \sqrt{2\pi}e^{\Theta_i(N)}\right| \le K_i\left|\frac{1}{12\theta_i(N)}\sqrt{2\pi}e^{\Theta_i(N)}\right|. \tag{3.9}$$

Next we combine approximations (3.9) and (3.8) using Lemma 14 from the Appendix A.2

$$|\Gamma(\psi_i(N))\Gamma(\theta_i(N)) - 2\pi e^{\Psi_i(N)+\Theta_i(N)}| \le \frac{K_i^2}{12\psi_i(N)12\theta_i(N)}2\pi e^{\Psi_i(N)+\Theta_i(N)}+$$
$$+\frac{K_i}{12\psi_i(N)}\sqrt{2\pi}e^{\Psi_i(N)+\Theta_i(N)}+\frac{K_i}{12\theta_i(N)}\sqrt{2\pi}e^{\Psi_i(N)+\Theta_i(N)}, \tag{3.10}$$

which holds for sufficient large N.

We set $K_{i,\psi,\theta}(N)$ to be

$$K_{i,\psi\theta}(N) = \frac{K_i^2}{12\psi_i(N)12\theta_i(N)} + \frac{K_i}{\sqrt{2\pi}12\psi_i(N)} + \frac{K_i}{\sqrt{2\pi}12\theta_i(N)},$$

and then we can write (3.10) as

$$\left|\Gamma(\psi_i(N))\Gamma(\theta_i(N)) - 2\pi e^{\Psi_i(N)+\Theta_i(N)}\right| \leq K_{i,\psi\theta}(N)2\pi e^{\Psi_i(N)+\Theta_i(N)}. \tag{3.11}$$

Now we use Lemma 13 from the Appendix A.2 to get the lower bound for (3.8) and (3.9)

$$\begin{aligned} K'_{i,\phi}\left|\frac{1}{12\phi_i(N)}\sqrt{2\pi}e^{\Phi_i(N)}\right| &\leq \left|\Gamma(\phi_i(N)) - \sqrt{2\pi}e^{\Phi_i(N)}\right|, \\ K'_{i,\theta}\left|\frac{1}{12\theta_i(N)}\sqrt{2\pi}e^{\Theta_i(N)}\right| &\leq \left|\Gamma(\theta_i(N)) - \sqrt{2\pi}e^{\Theta_i(N)}\right|, \end{aligned}$$

which holds for sufficient large N and $K_{i,\psi}, K_{i,\phi}, K_{i,\theta}$ are some positive constants. Further, if we introduce constant $K'_i = \max\{K_{i,\psi}, K_{i,\phi}, K_{i,\theta}\}$ we can represent the above inequalities as

$$K'_i\left|\frac{1}{12\phi_i(N)}\sqrt{2\pi}e^{\Phi_i(N)}\right| \leq \left|\Gamma(\phi_i(N)) - \sqrt{2\pi}e^{\Phi_i(N)}\right|, \tag{3.12}$$

$$K'_i\left|\frac{1}{12\theta_i(N)}\sqrt{2\pi}e^{\Theta_i(N)}\right| \leq \left|\Gamma(\theta_i(N)) - \sqrt{2\pi}e^{\Theta_i(N)}\right|, \tag{3.13}$$

and then we combine last two using Lemma 14 from A.2

$$K'_{i,\psi\theta}(N)2\pi e^{\Psi_i(N)+\Theta_i(N)} \leq |\Gamma(\psi_i(N))\Gamma(\theta_i(N)) - 2\pi e^{\Psi_i(N)+\Theta_i(N)}|, \tag{3.14}$$

where

$$K'_{i,\psi\theta}(N) = \frac{K'_i}{\sqrt{2\pi}12\psi_i(N)} + \frac{K'_i}{12\theta_i(N)} - \frac{K'^2_i}{12\psi_i(N)12\theta_i(N)}.$$

Now use Lemma 15 from A.2 for the inequalities (3.7), (3.11) and (3.14)

$$\left|\frac{\Gamma(\phi_i(N))}{\Gamma(\psi_i(N))\Gamma(\theta_i(N))} - \frac{\sqrt{2\pi}e^{\Phi_i(N)}}{2\pi e^{\Psi_i(N)+\Theta_i(N)}}\right| \leq \tag{3.15}$$
$$\leq \left(\frac{\frac{K_i}{12\phi_i(N)}\sqrt{2\pi}e^{\Phi_i(N)}}{K_{i,\psi\theta}(N)e^{\Psi_i(N)+\Theta_i(N)}} + \frac{|\sqrt{2\pi}e^{\Phi_i(N)}|}{|2\pi e^{\Psi_i(N)+\Theta_i(N)}|}\right)\frac{K'_{i,\psi\theta}(N)2\pi e^{\Psi_i(N)+\Theta_i(N)}}{|2\pi e^{\Psi_i(N)+\Theta_i(N)}| - K'_{i,\psi\theta}(N)2\pi e^{\Psi_i(N)+\Theta_i(N)}}.$$

Since exponential function is a positive function we can simplify the RHS of above inequality and introduce a new variable $K_{i,\phi\psi\theta}(N)$

$$\left(\frac{\frac{K_i}{12\phi_i(N)}\sqrt{2\pi}e^{\Phi_i(N)}}{K'_{i,\psi\theta}(N)e^{\Psi_i(N)+\Theta_i(N)}} + \frac{\sqrt{2\pi}e^{\Phi_i(N)}}{2\pi e^{\Psi_i(N)+\Theta_i(N)}}\right)\frac{K_{i,\psi\theta}(N)2\pi e^{\Psi_i(N)\Theta_i(N)}}{2\pi e^{\Psi_i(N)\Theta_i(N)} - K_{i,\psi\theta}(N)2\pi e^{\Psi_i(N)\Theta_i(N)}} =$$
$$= \frac{\sqrt{2\pi}e^{\Phi_i(N)}}{2\pi e^{\Psi_i(N)+\Theta_i(N)}}\left(\frac{2\pi K_i}{12\phi_i(N)K'_{i,\psi\theta}(N)} + 1\right)\frac{K_{i,\psi\theta}(N)}{1 - K_{i,\psi\theta}(N)} =$$
$$= \frac{1}{\sqrt{2\pi}}e^{\Phi_i(N)-\Psi_i(N)-\Theta_i(N)}\left(\frac{2\pi K_i + 12\phi_i(N)K'_{i,\psi\theta}(N)}{12\phi_i(N)K'_{i,\psi\theta}(N)}\right)\frac{K_{i,\psi\theta}(N)}{1 - K_{i,\psi\theta}(N)} =$$
$$= K_{i,\phi\psi\theta}(N)\frac{1}{\sqrt{2\pi}}e^{\Phi_i(N)-\Psi_i(N)-\Theta_i(N)}.$$

Hence (3.15) can be written as

$$\left|\frac{\Gamma(\phi_i(N))}{\Gamma(\psi_i(N))\Gamma(\theta_i(N))} - \frac{1}{\sqrt{2\pi}}e^{\Phi_i(N)-\Psi_i(N)-\Theta_i(N)}\right| \leq \tag{3.16}$$
$$\leq K_{i,\phi\psi\theta}(N)\frac{1}{\sqrt{2\pi}}e^{\Psi_i(N)-\Psi_i(N)-\Theta_i(N)}.$$

Now we use Lemma 14 from the Appendix A.2 to get expression for the m factors

$$\left|\prod_{i=1}^{m}\frac{\Gamma(\phi_i(N))}{\Gamma(\psi_i(N))\Gamma(\theta_i(N))} - \prod_{i=1}^{m}\frac{1}{\sqrt{2\pi}}e^{\Phi_i(N)-\Psi_i(N)-\Theta_i(N)}\right| \leq$$
$$\leq \prod_{i=1}^{m}K_{i,\phi\psi\theta}(N)\frac{1}{\sqrt{2\pi}}e^{\Phi_i(N)-\Psi_i(N)-\Theta_i(N)}+$$
$$+\sum_{j=1}^{m-1}\sum_{C_m^{m-j,j}}\prod_{k=1}^{m-j}\prod_{l=m-j+1}^{m}K_{k,\phi\psi\theta}(N)\frac{1}{\sqrt{2\pi}}e^{\Phi_i(N)-\Psi_i(N)-\Theta_i(N)},$$

where N must be sufficiently large and if we set

$$K_{\phi\psi\theta}(N) = \prod_{i=1}^{m} K_{i,\phi\psi\theta}(N) + \sum_{j=1}^{m-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{m-j} \prod_{l=m-j+1}^{m} K_{i_k,\phi\psi\theta}(N),$$

then we can write

$$\left| \prod_{i=1}^{m} \frac{\Gamma(\phi_i(N))}{\Gamma(\psi_i(N))\Gamma(\theta_i(N))} - \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi}} e^{\Phi_i(N)-\Psi_i(N)-\Theta_i(N)} \right| \leq \tag{3.17}$$
$$\leq K_{\phi\psi\theta}(N) \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi}} e^{\Phi_i(N)-\Psi_i(N)-\Theta_i(N)},$$

which is almost asymptotic equation from the Lemma we prove.

Explicit expressions of approximated function

Now we consider $\sum_{i=1}^{m}[\Psi_i(N) - \Phi_i(N) - \Theta_i(N)]$, we substitute its explicit expressions given by (3.6)

$$\sum_{i=1}^{m}[\Psi_i(N) - \Phi_i(N) - \Theta_i(N)] = \sum_{i=1}^{m} \Bigg[-\psi_i(N) + \left(\psi_i(N) - \frac{1}{2}\right) \ln \psi_i(N) +$$
$$+ \phi_i(N) - \left(\phi_i(N) - \frac{1}{2}\right) \ln \phi_i(N) + \theta_i(N) - \left(\theta_i(N) - \frac{1}{2}\right) \ln \theta_i(N) \Bigg],$$

and then substitute explicit expressions for $\psi_i(N), \phi_i(N), \theta_i(N)$ given by (3.5) and after some manipulations obtain

$$\sum_{i=1}^{m}[\Psi_i(N) - \Phi_i(N) - \Theta_i(N)] = \sum_{i=1}^{m} \Bigg[\left(x_i(N) + g_i G(N) - \frac{1}{2}\right) \ln(x_i N + g_i G(N)) -$$
$$\left(x_i N + \frac{1}{2}\right) \ln(x_i N + 1) - \left(g_i G(N) - \frac{1}{2}\right) \ln g_i G(N) + 1 \Bigg],$$

and after further manipulations

$$= \sum_{i=1}^{m} \Big[\big(x_i(N) + g_i G(N) \big) \ln(x_i N + g_i G(N)) - x_i N \ln x_i N - g_i G(N) \ln g_i G(N) -$$

$$- \frac{1}{2} \ln(x_i N + g_i G(N)) - x_i N \ln \left(1 + \frac{1}{x_i N} \right) - \frac{1}{2} \ln(x_i N + 1) + \frac{1}{2} \ln g_i G(N) + 1 \Big]. \tag{3.18}$$

First three terms are approximated analogically to the proof of Stiriling approximation in previous subsection, i.e.

$$1) \sum_{i=1}^{m} \Big[\big(x_i N + g_i G(N) \big) \ln \big(x_i N + g_i G(N) \big) - x_i N \ln x_i N - g_i G(N) \ln g_i G(N) \Big] =$$
$$= N \sum_{i=1}^{m} \left[x_i \ln \frac{g_i}{x_i} - \left(x_i + g_i \frac{G(N)}{N} \right) \ln \left(1 + \frac{x_i N}{g_i G(N)} \right) \right] + N \ln \frac{G(N)}{N},$$

For the second case we have

$$2) \sum_{i=1}^{m} \Big[\big(x_i N + g_i G(N) \big) \ln \big(x_i N + g_i G(N) \big) - x_i N \ln x_i N - g_i G(N) \ln g_i G(N) \Big] =$$
$$= N \sum_{i=1}^{m} \left[\left(x_i + g_i \frac{G(N)}{N} \right) \ln \left(x_i + g_i \frac{G(N)}{N} \right) - x_i \ln x_i - g_i \frac{G(N)}{N} \ln g_i \frac{G(N)}{N} \right],$$

and for the third

$$3) \sum_{i=1}^{m} \Big[\big(x_i N + g_i G(N) \big) \ln \big(x_i N + g_i G(N) \big) - x_i N \ln x_i N - g_i G(N) \ln g_i G(N) \Big] =$$
$$= G(N) \sum_{i=1}^{m} \left[g_i \ln x_i + \left(x_i \frac{N}{G(N)} + g_i \right) \ln \left(1 + \frac{g_i G(N)}{x_i N} \right) - g_i \ln g_i G(N) \right] + G(N) \ln N.$$

Hence we substitute it into (3.18) and take out N or $G(N)$ depending on case and obtain following

$$
\begin{aligned}
1)\ S(x,N) = N\sum_{i=1}^{m} &\Bigg[x_i \ln\frac{g_i}{x_i} + \left(x_i + g_i\frac{G(N)}{N}\right)\ln\left(1+\frac{x_iN}{g_iG(N)}\right) - \\
&- \frac{1}{2N}\ln(x_iN + g_iG(N)) - x_i\ln\left(1+\frac{1}{x_iN}\right) - \frac{1}{2N}\ln(x_iN+1) + \frac{1}{2N}\ln g_iG(N)\Bigg] + \\
&+ m + N\ln\frac{G(N)}{N} = Nf_1(x,N) + R_1(N).
\end{aligned}
$$

For the second case we have

$$
\begin{aligned}
2)\ S(x,N) = N\sum_{i=1}^{m} &\Bigg[\left(x_i + g_i\frac{G(N)}{N}\right)\ln\left(x_i + g_i\frac{G(N)}{N}\right) - x_i\ln x_i - g_i\frac{G(N)}{N}\ln\frac{g_iG(N)}{N} - \\
&- \frac{1}{2N}\ln(x_iN + g_iG(N)) - x_i\ln\left(1+\frac{1}{x_iN}\right) - \frac{1}{2N}\ln(x_iN+1) + \frac{1}{2N}\ln g_iG(N)\Bigg] + m = \\
&= Nf_2(x,N) + R_2(N),
\end{aligned}
$$

and for the third

$$
\begin{aligned}
3)\ S(x,N) = G(N)\sum_{i=1}^{m} &\Bigg[g_i\ln x_i + \left(x_i\frac{N}{G(N)} + g_i\right)\ln\left(1+\frac{g_iG(N)}{x_iN}\right) - g_i\ln g_iG(N) - \\
&- \frac{1}{2G(N)}\ln(x_iN + g_iG(N)) - \frac{x_iN}{G(N)}\ln\left(1+\frac{1}{x_iN}\right) - \frac{1}{2G(N)}\ln(x_iN+1) + \frac{1}{2G(N)}\ln g_iG(N)\Bigg] \\
&+ m + G(N)\ln N = G(N)f_3(x,N) + R_3(N).
\end{aligned}
$$

where here again $R_l(N)$, $l = 1,2,3$ include those terms which does not depend on x. We get the functions $f_l(x)$ by simply calculating the limits. Hence we get the expression from the theorem.

Constants estimates

We have following constants which occur in the upper bound (3.17)

$$\text{a) } K_{i,\psi\theta}(N) = \frac{K_i}{12\psi_i(N)} + \frac{K_i}{12\theta_i(N)} + \frac{K_i^2}{12\psi_i(N)12\theta_i(N)},$$
$$\text{b) } K'_{i,\psi\theta}(N) = \frac{K'_i}{12\psi_i(N)} + \frac{K'_i}{12\theta_i(N)} - \frac{K_i'^2}{12\psi_i(N)12\theta_i(N)},$$
$$\text{c) } K_{i,\phi\psi\theta}(N) = \left(\frac{K_i}{12\phi_i(N)K'_{i,\psi\theta}(N)}2\pi - 1\right)\frac{K_{i,\psi\theta}(N)}{1-K_{i,\psi\theta}(N)},$$
$$\text{d) } K_{\phi\psi\theta}(N) = \prod_{i=1}^{m} K_{i,\phi\psi\theta}(N) + \sum_{j=1}^{m-1}\sum_{C_m^{m-j,j}}\prod_{k=1}^{m-j} K_{i_k,\phi\psi\theta}(N)1_{i_l}.$$

For the first two constants we find upper and lower bound and for every constant we consider three asymptotic cases of $G(N)$ given by (3.2).

a) $K_{i,\psi\theta}(N)$ We will factorize $K_{i,\phi\theta}(N)$ into function of N and some function independent of N or at least bounded by function independent of N.

We start by substituting expressions for $\phi_i(N)$ and $\theta_i(N)$ given in (3.5) into $K_{i,\phi\theta}(N)$

$$K_{i,\psi\theta}(N) = \frac{K_i}{\sqrt{2\pi}12(x_iN+1)} + \frac{K_i}{\sqrt{2\pi}12g_iG(N)} + \frac{K_i'^2}{12(x_iN+1)g_iG(N)}. \tag{3.19}$$

Now we consider three cases described in (3.2) separately. As $N \to \infty$ we have

1) $\frac{G(N)}{N} \to \infty$,

We take out of the bracket $\frac{1}{N}$ of $K_{i,\phi\theta}(N)$

$$K_{i,\phi\theta}(N) = \frac{1}{N}\left[\frac{K_i}{\sqrt{2\pi}12(x_i+\frac{1}{N})} + \frac{K_iN}{\sqrt{2\pi}12g_iG(N)} + \frac{K_i'^2}{12(x_i+\frac{1}{N})12g_iG(N)}\right],$$

hence we can write

$$K_{i,\phi\theta}(N) = \frac{1}{N}{}_1K_{i,N,\phi\theta}.$$

In the limit $N \to \infty$ this function converges to a constant

$$\lim_{N\to\infty} {}_1K_{i,N,\phi\theta} = \frac{K_i}{\sqrt{2\pi}12x_i}.$$

2) $\frac{G(N)}{N} \to c$,

Here we do the same manipulations and

$$K_{i,\phi\theta}(N) = \frac{1}{N}\left[\frac{K_i}{\sqrt{2\pi}12(x_i + \frac{1}{N})} + \frac{K_i N}{\sqrt{2\pi}12 g_i G(N)} + \frac{K_i'^2}{12(x_i + \frac{1}{N})12 g_i G(N)}\right],$$

and shortly we can write

$$K_{i,\phi\theta}(N) = \frac{1}{N}{}_2K_{i,N,\phi\theta},$$

but the limit is

$$\lim_{N\to\infty} {}_2K_{i,N,\phi\theta} = \frac{K_i}{\sqrt{2\pi}12x_i} + \frac{K_i}{\sqrt{2\pi}12 g_i c}.$$

3) $\frac{G(N)}{N} \to 0$,

In this case we take out $G(N)$ out of the bracket

$$K_{i,\phi\theta}(N) = \frac{1}{G(N)}\left[\frac{K_i}{\sqrt{2\pi}12\big(x_i\frac{N}{G(N)} + \frac{1}{G(N)}\big)} + \frac{K_i}{\sqrt{2\pi}12 g_i} + \frac{K_i'^2}{12(x_i N + 1)12 g_i}\right],$$

Therefore we have

$$K_{i,\phi,\theta}(N) = \frac{1}{G(N)}{}_3K_{i,\phi\theta}(N),$$

and in the limit

$$\lim_{N\to\infty} {}_3K_{i,N,\phi\theta} = \frac{K_i}{\sqrt{2\pi}12 g_i}.$$

b)$K'_{i,\psi\theta}(N)$ We do the same manipulations with this constant as previously, but with minus in front of the last term.

c) $K_{i,\psi\phi\theta}(N)$ We bound it and factorize into the function of N and constant or some bounded function for each as case as $N \to \infty$

1. $\frac{G(N)}{N} \to \infty$,

   We have

$$K_{i,\psi\phi\theta}(N) = \left(\frac{2\pi K_i}{12\phi_i(N)K'_{i,\psi\theta}(N)} + 1\right)\frac{K_{i,\psi\theta}(N)}{1 - K_{i,\psi\theta}(N)},$$

and we substitute expression for for $K'_{i,\phi\theta}(N)$, $K_{i,\phi\theta}(N)$ and $\phi_i(N)$.

$$K_{i,\psi\phi\theta}(N) = \left(\frac{2\pi K_i N}{12(x_i N + g_i G(N))_1 K'_{i,\psi\theta}} + 1\right)\frac{\frac{{}_1K_{i,\psi\theta}}{N}}{1 - \frac{{}_1K_{i,\psi\theta}}{N}} =$$

and after some manipulations we get

$$= \left(\frac{2\pi K_i}{12(x_i + g_i\frac{G(N)}{N})_1 K'_{i,\psi\theta}} + 1\right)\frac{1}{N}\frac{{}_1K_{i,\psi\theta}}{1 - \frac{{}_1K_{i,\psi\theta}}{N}},$$

then

$$K_{i,\psi\phi\theta}(N) = \frac{1}{N}\left(\frac{2\pi K_i}{12(x_i + g_i\frac{G(N)}{N})_1 K'_{i,\psi\theta}} + 1\right)\frac{{}_1K_{i,\psi\theta}}{1 - \frac{{}_1K_{i,\psi\theta}}{N}}.$$

Hence, we can write

$$K_{i,\psi\phi\theta}(N) = \frac{1}{N}{}_1K_{i,N,\psi\phi\theta}.$$

2. $\frac{G(N)}{N} \to c$,

   Here situation is the same, again we put lower bounds for the $K'_{i,\Phi\Theta}(N)$, upper for the $K_{i,\Phi\Theta}(N)$ and $\psi_i(N)$

$$K_{i,\psi\phi\theta}(N) = \frac{1}{N}\left(\frac{2\pi K_i}{12(x_i + g_i\frac{G(N)}{N})_2 K'_{i,\psi\theta}} + 1\right)\frac{{}_2K_{i,\psi\theta}}{1 - \frac{{}_2K_{i,\psi\theta}}{N}}.$$

   Hence, we can shortly write

$$K_{i,\psi\phi\theta}(N) = \frac{1}{N}{}_2K_{i,N,\psi\phi\theta}.$$

3. $\frac{G(N)}{N} \to 0$,

   For third we take out $G(N)$ out of the bracket, rest is the same as for the second case

$$K_{i,\psi\phi\theta}(N) = \left(\frac{2\pi K_i}{12\phi_i(N)K'_{i,\psi\theta}(N)} + 1\right)\frac{K_{i,\psi\theta}(N)}{1 - K_{i,\psi\theta}(N)},$$

   and we substitute expression for for $K'_{i,\phi\theta}(N)$, $K_{i,\phi\theta}(N)$ and $\phi_i(N)$.

$$K_{i,\psi\phi\theta}(N) = \left(\frac{2\pi K_i G(N)}{12(x_i N + g_i G(N))_3 K'_{i,\psi\theta}} + 1\right)\frac{\frac{{}_3K_{i,\psi\theta}}{G(N)}}{1 - \frac{{}_3K_{i,\psi\theta}}{G(N)}} =$$

and after some manipulations we get

$$= \left(\frac{2\pi K_i}{12(x_i \frac{N}{G(N)} + g_i) {}_3K'_{i,\psi\theta}} + 1 \right) \frac{1}{G(N)} \frac{{}_3K_{i,\psi\theta}}{1 - \frac{{}_3K_{i,\psi\theta}}{G(N)}}.$$

Shortly we can write

$$K_{i,\psi\phi\theta}(N) = \frac{1}{G(N)} {}_3K_{i,N,\psi\phi\theta}. \tag{3.20}$$

d) $K_{\psi\phi\theta}(N)$

1. $\frac{G(N)}{N} \to \infty$,
   We substitute expression for $K_{i,\psi\phi\theta}(N)$ into $K_{\phi\psi\theta}$

$$K_{\phi\psi\theta}(N) = \prod_{i=1}^{m} \frac{1}{N} {}_1K_{i,N,\psi\phi\theta} + \sum_{j=1}^{m-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{m-j} \prod_{l=m-j+1}^{m} \frac{1}{N} {}_1K_{i_k,N,\psi\phi\theta} 1_{i_l},$$

factorize term $\frac{1}{N}$ and get

$$= \frac{1}{N} \left[\left(\frac{1}{N}\right)^{m-1} \prod_{i=1}^{m} {}_1K_{i,N,\psi\phi\theta} + \sum_{j=1}^{m-1} \left(\frac{1}{N}\right)^{j-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{m-j} \prod_{l=m-j+1}^{m} {}_1K_{i_k,N,\psi\phi\theta} 1_{i_l} \right],$$

and after some manipulations

$$= \frac{1}{N} \Bigg[\sum_{C_m^{m-j,j}} {}_1K_{i_1,N,\psi\phi\theta} + \left(\frac{1}{N}\right)^{m-1} \prod_{i=1}^{m} {}_1K_{i,N,\psi\phi\theta} +$$
$$+ \sum_{j=1}^{m-2} \left(\frac{1}{N}\right)^{m-j-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{m-j} \prod_{l=m-j+1}^{m} {}_1K_{i_k,N,\psi\phi\theta} 1_{i_l} \Bigg] = \frac{1}{N} K_{1,N},$$

hence we can write

$$K_{\phi\psi\theta}(N) = \frac{1}{N} K_{1,N},$$

where in the limit $N \to \infty$

$$K_{1,N} \to \sum_{i=1}^{m} {}_1K_{i,\psi\theta} = K_1.$$

2. $\frac{G(N)}{N} \to c$,
Here the situation is the same as in the previous case

$$K_{\phi\psi\theta}(N) = \frac{1}{N}\Bigg[\sum_{C_m^{m-j,j}} {}_1K_{i_2,N,\psi\phi\theta} + \left(\frac{1}{N}\right)^{m-1}\prod_{i=1}^{m} {}_2K_{i,N,\psi\phi\theta} +$$
$$+\sum_{j=1}^{m-2}\left(\frac{1}{N}\right)^{m-j-1}\sum_{C_m^{m-j,j}}\prod_{k=1}^{m-j}\prod_{l=m-j+1}^{m} {}_2K_{i_k,N,\psi\phi\theta}1_{i_l}\Bigg] = \frac{1}{N}K_{2,N},$$

hence we can write

$$K_{\phi\psi\theta}(N) = \frac{1}{N}{}_2K_{N,\phi\psi\theta},$$

where in the limit

$$K_{1,N} \to \sum_{i=1}^{m} {}_2K_{i,\psi\theta} = K_2.$$

3. $\frac{G(N)}{N} \to 0$,
For third case the only difference is $G(N)$ instead of N

$$K_{\phi\psi\theta}(N) = \frac{1}{G(N)}\Bigg[\sum_{C_m^{m-j,j}} {}_1K_{i_2,N,\psi\phi\theta} + \left(\frac{1}{G(N)}\right)^{m-1}\prod_{i=1}^{m} {}_2K_{i,N,\psi\phi\theta} +$$
$$+\sum_{j=1}^{m-2}\left(\frac{1}{G(N)}\right)^{m-j-1}\sum_{C_m^{m-j,j}}\prod_{k=1}^{m-j}\prod_{l=m-j+1}^{m} {}_2K_{i_k,N,\psi\phi\theta}1_{i_l}\Bigg] = \frac{1}{G(N)}K_{2,N},$$

$$K_{\phi\psi\theta}(N) = \frac{1}{G(N)}{}_3K_{N,\phi\psi\theta},$$

and in the limit

$$K_{3,N} \to \sum_{i=1}^{m} {}_3K_{i,\psi\theta} = K_3.$$

Now we substitute the estimates for the constant $K_{\psi\phi\theta}(N)$ into the inequalitie (3.17), together with substituting explicit expressions for $\prod_{i=1}^{m}[\Psi_i(N) - \Phi_i(N) - \Theta_i(N)]$ and obtain the final result. □

3.2 Optimization

The content of this section is devoted to the optimization problems related to the approximated entropy $f_l(x)$ and $f_l(x, N)$ which are outcome of the Lemma of the previous Section. The optimization problems are solved in the first two subsections. Some related results developed specifically for the optimization are included in the third subsection.

3.2.1 Optimization of the limit of the approximated entropy

Let $\mathbb{R}^m_+$ be a nonnegative orthant of $\mathbb{R}^m$. Then, the functions from the Theorem 6 in the Section 1, i.e. $f_l : \mathbb{R}^m_{++} \to \mathbb{R}$, $l = 1, 2, 3$ defined

$$f_1(x) = \sum_{i=1}^{m} x_i \ln \frac{g_i}{x_i} + x_i,$$
$$f_2(x) = \sum_{i=1}^{m} \Big[(x_i + g_i c) \ln(x_i + g_i c) - x_i \ln x_i \Big]$$
$$f_3(x) = \sum_{i=1}^{m} g_i \ln x_i + g_i,$$

where $c > 0$ and $g_i > 0,\ i = 1, \ldots, m$ are some constants and x_i is i-th component of x and $\sum_{i=1}^{m} g_i = 1$.

For each f_l we have optimization problem over the domain Ω_E recalled in the beginning of this chapter, i.e.

$$\begin{aligned} \text{maximize} \quad & f_l, \\ \text{subject to} \quad & \sum_{i=1}^{m} x_i = 1, \\ & \sum_{i=1}^{m} \varepsilon_i x_i \le E, \end{aligned} \tag{3.21}$$

where $E > 0$ and $\varepsilon_i,\ i = 1, \ldots, m$ are some constants such that $0 < \varepsilon_1 < \varepsilon_2 < \ldots < \varepsilon_m$.

Lemma 2. The solution of the optimization problem (3.21) exists only if $E > \varepsilon_1$ and this solution (optimal vector) is unique.

For $\overline{g\varepsilon} = \frac{1}{m} \sum_{i=1}^{m} g_i \varepsilon_i$, then if $E \ge \overline{g\varepsilon}$, then the optimal vector $x^* =$

$(x_1^*, x_2^*, \ldots, x_m^*)$ has components

$$x_i^* = g_i, i = 1, \ldots, m,$$

and if $g_{im}\varepsilon_{im} < E < \overline{\varepsilon}$, where $g_{im}\varepsilon_{im} = \min_i g_i\varepsilon_i$ respectively for each instance of f_l, the optimal vector has components

$$x_i^* = \frac{g_i}{e^{\lambda^*\varepsilon_i + \nu^*}},$$
$$x_i^* = \frac{g_i c}{e^{\lambda^*\varepsilon_i + \nu^*} - 1},$$
$$x_i^* = \frac{g_i}{\lambda^*\varepsilon_i + \nu^*},$$

for $i = 1, \ldots, m$, where the parameters λ^*, ν^* are uniquely determined by the system of equations

$$\sum_{i=1}^{m} x_i^* = 1,$$
$$\sum_{i=1}^{m} \varepsilon_i x_i^* = E.$$

Proof. We start by showing uniqueness of the solution, assuming it exists.

Let us denote by Ω the domain of optimization, i.e. the set of vectors satisfying the constraints of the problem (3.21) and implicit constraint from the definition of f, i.e. $x \in \mathbb{R}^m_{++}$. Equivalently, this set is an intersection of two m-dimensional simplexes, first is determined by origin and standard basis vectors of $\mathbb{R}^m$, i.e., $o, e_1, \ldots, e_m$ and second by vectors $0, \varepsilon_1 e_1/E, \ldots, \varepsilon_m e_m/E$. Since simplexes are convex sets, so their intersection and therefore the domain Ω is convex.

Now, lets show concavity of each instance of f_l. For all three functions, since they are twice differentiable, the matrices of second derivatives exists and are equal

$$
\begin{aligned}
[D^2 f(x)]_{i,j} &= -\delta_{ij}\frac{1}{x_i},\\
[D^2 f(x)]_{i,j} &= -\delta_{ij}\frac{g_i c}{x_i(x_i+g_i c)},\\
[D^2 f(x)]_{i,j} &= -\delta_{ij}\frac{g_i}{x_i^2},
\end{aligned}
$$

where $i, j = 1, \ldots, m$ and $\delta_{ij} = 1$ if $i = j$, otherwise $\delta_{ij} = 0$.
Since $c > 0, x_i > 0$ and $g_i > 0$ for $i = 1, \ldots, m$, the elements of above matrix are negative on the diagonal and zero elsewhere. Hence for all $x \in \mathbb{R}^m$, $x \neq o$ we have

$$x^T D^2 f_l(x) x < 0,$$

i.e. $D^2 f_l(x)$ is negative definite, which implies strict concavity of f_l.

Since problem (3.21) has convex domain, affine constraints and concave objective it is a convex optimization problem, for definition and terminology see Appendix A.4. Moreover, since objective is strictly concave the optimal vector is unique, if exists.

Now, we find the explicit form of the optimal vector. We start by representing (3.21) in the standard form

$$
\begin{aligned}
\text{minimize} \quad & -f_l, \\
\text{subject to} \quad & \varepsilon^T x - E \leq 0, \\
& \mathbf{1}^T x - 1 = 0,
\end{aligned}
\tag{3.22}
$$

where $\mathbf{1}$ is the unit vector and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)$.
Further, for the above problem we define Lagrange function $L : \mathbb{R}^m \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ for the problem

$$L(x, \lambda, \nu) = -f_l(z) + \lambda\Big(\varepsilon^T x - E\Big) + \nu\Big(\mathbf{1}^T z - 1\Big),$$

and corresponding Lagrange dual function (dual function) $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$

$$g(\lambda, \nu) = \inf_{x \in \Omega} L(x, \lambda, \nu) = \inf_{x \in \Omega} \Big\{ -f_l(x) + \lambda\Big(\varepsilon^T x - E\Big) + \nu\Big(\mathbf{1}^T x - 1\Big)\Big\},$$

where $\lambda, \nu \in \mathbb{R}$ are Lagrange multipliers.

Let p^* be the optimal value and x^* corresponding optimal vector of the problem (3.22), then for $\lambda \geq 0$ and any ν we have

$$\begin{aligned} g(\lambda, \nu) &= \inf_{x \in \Omega} L(x, \lambda, \nu) \leq L(x^*, \lambda, \nu) = \\ &= -f_l(x^*) + \lambda^* \Big(\varepsilon^T x^* - E\Big) + \nu(\mathbf{1}^T x^* - 1) \leq -f(x^*) = p^*, \end{aligned} \tag{3.23}$$

where the last inequality is valid since x^* is in the domain Ω.

Hence, from (3.23), for $\lambda \geq 0$ and any ν, function $g(\lambda, \nu)$ yields a lower bound for the optimal value, i.e.

$$g(\lambda, \nu) \leq p^*.$$

We find the biggest such lower bound by solving an optimization problem

$$\begin{aligned} \text{maximize} \quad & g(\lambda, \nu), \\ \text{subject to} \quad & \lambda \geq 0, \end{aligned}$$

which is Lagrange dual problem (dual problem) associated with the (primal) problem (3.22)
Note, it is a convex problem, irrespective of underlying problem, as $g(\lambda, \nu)$ is a point-wise infimum of a family of affine functions of (λ, ν). Hence, the maximum of $g(\lambda, \nu)$, if exists, is a global maximum.

Since the primal problem is convex and there exists $x \in \mathrm{Relint}(\Omega)$, (for the definition of Relint see Appendix A.4) with

$$\begin{aligned} &\mathbf{1}^T x - 1 = 0, \\ &\varepsilon^T x - E < 0, \end{aligned}$$

the Slater's conditions holds (see Appendix on Theory of Optimization A.4) therefore strong duality occurs and optimal point exists. Hence we have that

$$\begin{aligned} g(\lambda^*, \nu^*) &= \inf_{x \in \Omega} L(x, \lambda^*, \nu^*) = L(x^*, \lambda^*, \nu^*) \\ &= -f_l(x^*) + \lambda^* \Big(\varepsilon^T x^* - E\Big) + \nu^*(\mathbf{1}^T x^* - 1) = p^*. \end{aligned}$$

From the last equality it follows that the strong duality implies complementary slackness

$$\lambda^*\Big(\varepsilon^T x^* - E\Big) = 0, \tag{3.24}$$

where x^* and λ^* are optimal values.

Now, as $L(x, \lambda^*, \nu^*)$ is a sum of the convex and affine functions, it is a convex function of x. Further, since the function $L(x, \lambda^*, \nu^*)$ is differentiable with respect to x, $\inf_{x\in\Omega} L(x, \lambda^*, \nu^*)$ exists and is finite only if $\nabla_x L(x, \lambda^*, \nu^*) = o$ for some critical vector x^*. Hence, we get the gradient condition

$$-\nabla f_l(x^*) + \lambda^*\varepsilon + \nu^*\mathbf{1} = o, \tag{3.25}$$

where o is zero vector and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)$.

Now, if we put together the constraints of the primal and the dual problem, complementary slacknes (3.24) and gradient conditions (3.25) we arrive with Karush-Kuhn-Tucker (KKT) conditions, for details see Appendix A.4 ,

$$\varepsilon^T x^* - E \le 0, \tag{3.26a}$$

$$\mathbf{1}^T x^* - 1 = 0, \tag{3.26b}$$

$$\lambda^* \ge 0, \tag{3.26c}$$

$$\lambda^*\Big(\varepsilon^T x^* - E\Big) = 0, \tag{3.26d}$$

$$-\nabla f_l(x^*) + \lambda^*\varepsilon + \nu^*\mathbf{1} = o. \tag{3.26e}$$

For the convex optimization problem with the strong duality, these are necessary and sufficient conditions for the vectors x^* and (λ^*, ν^*) to be primal and dual optimal.

Now, we solve (3.26).

For the first function, $f_1(x) = \sum_{i=1}^m x_i \ln \frac{g_i}{x_i} + x_i$ from the gradient condition (3.26e) we obtain

$$x_i^* = \frac{g_i}{e^{\lambda^* E_i + \nu^*}} \tag{3.27}$$

Next, first four equations we represent as two cases of possible ranges of values of

λ^*. The first case is

$$\begin{aligned} \varepsilon^T x^* - E &\leq 0, \\ \mathbf{1}^T x^* - 1 &= 0, \\ \lambda^* &= 0. \end{aligned}$$

From Lemma 15 attach at the end of this section, the solution for the above system exists and is unique only if $E \geq \overline{g\varepsilon}$ and $(\lambda^*, \nu^*) = (0, 0)$. Hence (3.27) becomes

$$x_i = g_i, i = 1, \dots, m.$$

The other case is

$$\begin{aligned} \varepsilon^T x^* - E &= 0, \\ \mathbf{1}^T x^* - 1 &= 0, \\ \lambda^* &> 0, \end{aligned} \tag{3.28}$$

where we have equality in the first condition because of $\lambda^* \Big(\varepsilon^T x^* - E \Big) = 0$. By Lemma 4 from Subsection 3 of this section, solution (λ, ν) exists and is unique only if $g_{im}\varepsilon_{im} < E < \overline{g\varepsilon}$. Further, substituting (3.27) into two first conditions of (3.28) and we get the system of equations from which we can calculate parameters λ and ν explicitly. For the second and third case situation is analogical but we use respectively Lemma 5 and 6 from the Subsection 3 of this Section. Only the outcome of the gradient gives different result. □

3.2.2 Optimization of the approximated entropy

Let the functions $f_l : \mathbb{R}^m_{++} \times \mathbb{N} \to \mathbb{R}$, $l = 1, 2, 3$ are

$$
\begin{aligned}
f_1(x, N) = &\sum_{i=1}^{m} \left[x_i \ln \frac{g_i}{x_i} + \left(x_i + g_i \frac{G(N)}{N} \right) \ln \left(1 + \frac{x_i N}{g_i G(N)} \right) - \right.\\
&\left. - \frac{1}{2N} \ln(x_i N + g_i G(N)) - x_i \ln \left(1 + \frac{1}{x_i N} \right) - \frac{1}{2N} \ln(x_i N + 1) \right],\\
f_2(x, N) = &\sum_{i=1}^{m} \left[\left(x_i + g_i \frac{G(N)}{N} \right) \ln \left(x_i + g_i \frac{G(N)}{N} \right) - x_i \ln x_i - \right.\\
&\left. - \frac{1}{2N} \ln(x_i N + g_i G(N)) - x_i \ln \left(1 + \frac{1}{x_i N} \right) - \frac{1}{2N} \ln(x_i N + 1) \right],\\
f_3(x, N) = &\sum_{i=1}^{m} \left[g_i \ln x_i + \left(x_i \frac{N}{G(N)} + g_i \right) \ln \left(1 + \frac{g_i G(N)}{x_i N} \right) \right] - \\
&- \frac{1}{2G(N)} \ln(x_i N + g_i G(N)) - \frac{x_i N}{G(N)} \ln \left(1 + \frac{1}{x_i N} \right) - \frac{1}{2G(N)} \ln(x_i N + 1) \Big],
\end{aligned}
$$

where $c > 0$ and $g_i > 0,\ i = 1, \ldots, m$ are some constants and $\sum_{i=1}^{m} g_i = 1$.

For each $f_l(x, N)$ we have optimization problem

$$
\begin{aligned}
\text{maximize} \quad & f_l(x, N), \\
\text{subject to} \quad & \sum_{i=1}^{m} x_i = 1, \\
& \sum_{i=1}^{m} \varepsilon_i x_i \le E,
\end{aligned}
\tag{3.29}
$$

where $E > 0$ and $\varepsilon_i,\ i = 1, \ldots, m$ are some constants such that $0 < \varepsilon_1 < \varepsilon_2 < \ldots < \varepsilon_m$.

Lemma 3. For each instance of $f_l(x, N)$, for large enough N the solution of the above optimization problem (optimal vector) exists and is unique.

Proof. We start by showing uniqueness of the solution, assuming it exists.

Let us denote by Ω the domain of optimization, i.e. the set of vectors satisfying the constraints of the problem (3.29) and implicit constraint from the definition of f, i.e. $x \in \mathbb{R}^m_{++}$. Equivalently, this set is an intersection of two m-dimensional simplexes, first is determined by origin and standard basis vectors of $\mathbb{R}^m$, i.e., $o, e_1, \ldots, e_m$ and second by vectors $0, \varepsilon_1 e_1/E, \ldots, \varepsilon_m e_m/E$. Since simplexes are convex sets, so their intersection and therefore the domain Ω is convex.

Now, lets show concavity of each instance of $f_l(x, N)$. For all three functions, since they are twice differentiable, the matrices of second derivatives exists and are equal

$$
\begin{aligned}
[D^2 f_1(x,N)]_{i,j} =& -\delta_{ij}\bigg(\frac{1}{x_i} + \frac{1}{2(x_iN + g_iG(N))(x_i + g_iG(N)/N)} + \frac{1}{x_i^2 + x_i} - \\
& - \frac{1}{2(x_iN+1)(x_i+1/N)}\bigg), \\
[D^2 f_2(x,N)]_{i,j} =& -\delta_{ij}\bigg(\frac{g_i c}{x_i(x_i + g_i c)} + \frac{1}{2(x_iN + g_iG(N))(x_i + g_iG(N)/N)} + \\
& + \frac{1}{x_i^2 + x_i} - \frac{1}{2(x_iN+1)(x_i+1/N)}\bigg), \\
[D^2 f_2(x,N)]_{i,j} =& -\delta_{ij}\bigg(\frac{g_i}{x_i^2} + \frac{1}{2(x_iN + g_iG(N))(x_i + g_iG(N)/N)} + \frac{1}{x_i^2 + x_i} - \\
& - \frac{1}{2(x_iN+1)(x_i+1/N)}\bigg),
\end{aligned}
$$

where $i, j = 1, \ldots, m$ and $\delta_{ij} = 1$ if $i = j$, otherwise $\delta_{ij} = 0$.
Since $c > 0, x_i > 0$ and $g_i > 0$ for $i = 1, \ldots, m$, for large enough N the elements of above matrix are negative on the diagonal and zero elsewhere. Hence for all $x \in \mathbb{R}^m$, $x \neq o$ we have

$$x^T D^2 f_l(x, N) x < 0,$$

i.e. $D^2 f_l(x, N)$ is negative definite, which implies strict concavity of $f_l(x, N)$.

Since problem (3.29) has convex domain, affine constraints and concave objective it is a convex optimization problem, for definition and terminology see Appendix A.4. Moreover, since objective is strictly concave the optimal vector is unique, if exists.

Since the considered problem is convex and there exists $x \in \mathrm{Relint}(\Omega)$, (for the definition of Relint see Appendix A.4) with

$$
\begin{aligned}
\mathbf{1}^T x - 1 = 0, \\
\varepsilon^T x - E < 0,
\end{aligned}
$$

the Slater's conditions holds (for details see the Appendix A.4) therefore strong duality occurs and optimal point exists. $\square$

3.2.3 Related results

For the positive numbers $g_i, \varepsilon_i, i = 1, \ldots, m$ such that $\varepsilon_1 < \varepsilon_2 < \ldots < \varepsilon_m$, $\sum_{i=1}^{m} g_i = 1$ and some $E > 0$ we have the system of equations

$$\begin{aligned} \sum_{i=1}^{m} x_i &= 1, \\ \sum_{i=1}^{m} \varepsilon_i x_i &= E. \end{aligned} \tag{3.30}$$

where

$$x_i > 0,\ i = 1, \ldots, m, \tag{3.31}$$

Let $\overline{g\varepsilon} = \frac{1}{m}\sum_{i=1}^{m} g_i\varepsilon_i$ and $g_{im}\varepsilon_{im} = \min_i g_i\varepsilon_i$, $g_{iM}\varepsilon_{iM} = \max_i g_i\varepsilon_i$ then we following have lemmas

Lemma 4. For the system of equations (3.30) let

$$x_i = \frac{g_i}{e^{\lambda\varepsilon_i + \nu}},\ i = 1, \ldots, m. \tag{3.32}$$

where λ and ν are some parameters.

Then

i)	if $E = \overline{g\varepsilon}$,	then the solution is $(\lambda, \nu) = (0, 0)$ and is unique,
ii)	if $g_{im}\varepsilon_{im} < E < \overline{g\varepsilon}$,	then for $\lambda > 0$ the solution exists and is unique,
iii)	if $\overline{g\varepsilon} < E < g_{iM}\varepsilon_{iM}$,	then for $\lambda < 0$ the solution exists and is unique,
iv)	if $E \notin (g_{im}\varepsilon_{im}, g_{iM}\varepsilon_{iM})$,	then the solution does not exists.

Proof. We start with proof of uniqueness of the solution (λ, ν). First let us assume the solution exists. Then we substitute (3.32) into (3.30) and get

$$1 = \sum_{i=1}^{m} \frac{g_i}{e^{\lambda \varepsilon_i + \nu}}, \tag{3.33}$$

$$E = \sum_{i=1}^{m} \frac{g_i \varepsilon_i}{e^{\lambda \varepsilon_i + \nu}}, \tag{3.34}$$

From the first equation we have

$$\nu = \log\left(\sum_{i=1}^{m} g_i e^{-\lambda \varepsilon_i} \right). \tag{3.35}$$

Note that the function $\nu = \nu(\lambda)$ is strictly decreasing, hence one-to-one.
Next we substitute (3.35) into second equation of (3.33) and obtain

$$E = \frac{\sum_{i=1}^{m} g_i \varepsilon_i e^{-\lambda \varepsilon_i}}{\sum_{i=1}^{m} g_i e^{-\lambda \varepsilon_i}}.$$

Let us show that $E = E(\lambda)$ is ono-to-one function.
We calculate its derivative,

$$E'(\lambda) = \frac{-\sum_{i=1}^{m} g_i \varepsilon_i^2 e^{-\lambda \varepsilon_i} \left(\sum_{i=1}^{m} g_i e^{-\lambda \varepsilon_i} \right) + \left(\sum_{i=1}^{m} \varepsilon_i g_i e^{-\lambda \varepsilon_i} \right)^2}{\left(\sum_{i=1}^{m} g_i e^{-\lambda \varepsilon_i}\right)^2},$$

and represent as the difference of the expected values,

$$E'(\lambda) = E[\mathcal{E}]^2 - E[\mathcal{E}^2],$$

where $\mathcal{E}$ is a random variable with range $\Omega_{\mathcal{E}} = \{\varepsilon_1, \ldots, \varepsilon_m\}$ and pdf $f(\varepsilon_i) = g_i e^{-\lambda \varepsilon_i} / \sum_{j=1}^{m} g_i e^{-\lambda \varepsilon_j}$. By the Cauchy-Schwartz inequality $E'(\lambda)$ is strictly negative, therefore $E = E(\lambda)$ is strictly decreasing function, hence one-to-one.
Assuming the solution exists, since $E = E(\lambda)$ and $\nu = \nu(\lambda)$ are one-to-one, the solution (λ, ν) is unique.

Now lets prove the existence of the solution (λ, ν).

The function $E(\lambda)$ as $\lambda \to \pm\infty$ and $\lambda = 0$ takes values

$$\begin{aligned}
E(\lambda) =& \frac{\sum_{i=1}^{m} \varepsilon_i g_i e^{-\lambda\varepsilon_i}}{\sum_{i=1}^{m} g_i e^{-\lambda\varepsilon_i}} = \frac{g_{im}\varepsilon_{im} + \sum_{i=1, i\neq im}^{m} \varepsilon_i g_i e^{-\lambda(\varepsilon_i - \varepsilon_{im})}}{1 + \sum_{i=2}^{m} g_i e^{-\lambda(\varepsilon_i - \varepsilon_{im})}} \to g_{im}\varepsilon_{im}, \\
&\text{as } \lambda \to \infty, \\
E(\lambda) =& \frac{1}{m} \sum_{i=1}^{m} g_i \varepsilon_i, \quad \text{for } \lambda = 0, \\
E(\lambda) =& \frac{\varepsilon_{iM} + \sum_{i=1, i\neq iM}^{m} g_i \varepsilon_i e^{-\lambda(\varepsilon_i - \varepsilon_{iM})}}{1 + \sum_{i=1}^{m-1} e^{-\lambda(\varepsilon_i - \varepsilon_{iM})}} \to g_{iM}\varepsilon_{iM}, \quad \text{as } \lambda \to -\infty.
\end{aligned}$$

Since $E = E(\lambda)$ is strictly decreasing, points $g_{im}\varepsilon_{im}$ and $g_{iM}\varepsilon_{iM}$ are boundaries of its range, hence λ exists only if $E \in (g_{im}\varepsilon_{im}, g_{iM}\varepsilon_{iM})$ and further

$$\begin{aligned}
&\text{if } g_{im}\varepsilon_{im} < E < \overline{g\varepsilon}, && \text{then } \lambda > 0, \\
&\text{if } E = \overline{g\varepsilon}, && \text{then } \lambda = 0, \\
&\text{if } \overline{g\varepsilon} < E < g_{iM}\varepsilon_{iM}, && \text{then } \lambda < 0,
\end{aligned}$$

Now, from (3.35) we have

$$\begin{aligned}
&\nu(\lambda) \to \infty && \text{as } \lambda \to \infty, \\
&\nu(\lambda) = 0 && \text{for } \lambda = 0, \\
&\nu(\lambda) \to -\infty && \text{as } \lambda \to -\infty.
\end{aligned}$$

hence for any λ parameter ν exists.
Putting the results together we get the lemma. □

Lemma 5. For the system of equations (3.30) let

$$x_i = \frac{g_i}{\lambda\varepsilon_i + \nu}, \ i = 1, \ldots, m. \tag{3.36}$$

where λ and ν are some parameter.

Then,

$i)$ if $E = \overline{g\varepsilon}$, then the solution is $(\lambda, \nu) = (0, 1)$ and is unique,

$ii)$ if $g_{im}\varepsilon_{im} < E < \overline{g\varepsilon}$, then the solution exists and is unique for $\lambda > 0, \lambda < -\nu\varepsilon_1$,

$iii)$ if $\overline{g\varepsilon} < E < g_{iM}\varepsilon_{iM}$, then the solution exists and is unique for $\lambda < 0, \lambda > -\nu\varepsilon_m$,

$iv)$ if $E \neq (g_{im}\varepsilon_{im}, g_{iM}\varepsilon_{iM})$, then the solution does not exists.

Proof. We start with proof of uniqueness of the solution. First we assume it exists. Then we substitute (3.36) into (3.30), and get

$$1 = \sum_{i=1}^{m} \frac{g_i}{\lambda\varepsilon_i + \nu}, \tag{3.37}$$

$$E = \sum_{i=1}^{m} \frac{\varepsilon_i}{\lambda\varepsilon_i + \nu}. \tag{3.38}$$

and then we perform a substitution $\nu = \lambda\alpha$ and for $\lambda \neq 0$ we have

$$1 = \sum_{i=1}^{m} \frac{g_i}{\lambda(\varepsilon_i + \alpha)}, \tag{3.39a}$$

$$E = \sum_{i=1}^{m} \frac{g_i\varepsilon_i}{\lambda(\varepsilon_i + \alpha)}. \tag{3.39b}$$

From (3.39a) we obtain

$$\lambda = \sum_{i=1}^{m} \frac{g_i}{\varepsilon_i + \alpha}, \tag{3.40}$$

Except the singularities at the points $\alpha = -\varepsilon_i, i = 1, \ldots, m$, the function $\lambda = \lambda(\alpha)$ is strictly decreasing.
Note that by (3.31)

$$\frac{1}{\lambda(\varepsilon_i + \alpha)} > 0,\ i = 1, \ldots, m, \tag{3.41}$$

hence λ and α can take values

$$\lambda > 0,\ \alpha > -\varepsilon_1, \tag{3.42}$$

or

$$\lambda < 0,\ \alpha < -\varepsilon_m, \tag{3.43}$$

and if we define $\lambda = \lambda(\alpha)$ separately for the domains (3.42) and (3.43) it is also one-to-one.
Next we substitute substituting (3.40) into (3.39b) and get

$$E = \frac{\sum_{i=1}^{m} \frac{g_i \varepsilon_i}{\varepsilon_i+\alpha}}{\sum_{i=1}^{m} \frac{g_i}{\varepsilon_i+\alpha}}.$$

Let us show that function $E = E(\alpha)$ is one-to-one.
We calculate its derivative,

$$E'(\alpha) = \frac{-\sum_{i=1}^{m} \frac{g_i\varepsilon_i}{(\varepsilon_i+\alpha)^2}\left(\sum_{i=1}^{m}\frac{g_i}{\varepsilon_i+\alpha}\right) + \sum_{i=1}^{m}\frac{g_i\varepsilon_i}{\varepsilon_i+\alpha}\left(\sum_{i=1}^{m}\frac{g_i}{(\varepsilon_i+\alpha)^2}\right)}{\left(\sum_{i=1}^{m}\frac{g_i}{\varepsilon_i+\alpha}\right)^2}$$

and represent it in terms of the expectations

$$E'(\alpha) = E[\mathcal{E}]E\left[\frac{1}{\mathcal{E}+\alpha}\right] - E\left[\frac{\mathcal{E}}{\mathcal{E}+\alpha}\right], \tag{3.44}$$

where $\mathcal{E}$ and $(\mathcal{E}+\alpha)^{-1}$ are random variables with ranges

$$\Omega_{\mathcal{E}} = \{\varepsilon_1, \ldots, \varepsilon_m\},$$
$$\Omega_{(\mathcal{E}+\alpha)^{-1}} = \left\{\frac{1}{\varepsilon_1+\alpha}, \ldots, \frac{1}{\varepsilon_m+\alpha}\right\},$$

both with pdf $f_i = \frac{\frac{g_i}{(\varepsilon_i+\alpha)}}{\sum_{j=1}^{m}\frac{g_i}{\varepsilon_j+\alpha}}$.
Now, setting $g(\mathcal{E}) = \mathcal{E}, h(\mathcal{E}) = \frac{1}{\mathcal{E}+\alpha}$ and use special case of FKG inequality, see Appendix on Probability A.3 for the details,

$$E\left[\frac{\mathcal{E}}{\mathcal{E}+\alpha}\right] < E[\mathcal{E}]E\left[\frac{1}{\mathcal{E}+\alpha}\right],$$

which implies $E'(\alpha)$ is strictly positive, therefore $E = E(\alpha)$ is strictly increasing function. If we define $E(\alpha)$ for the domains (3.42) and (3.43) separately, it is also one-to-one. Therefore the parameters λ and α are unique.

Next we prove the existence of λ and α.
Let us start with showing the existence of α. The function $E(\alpha)$ as $\alpha \to -\varepsilon_1$,

$\alpha \to -\varepsilon_m$ and as $\alpha \to \pm\infty$ takes values

$$E(\alpha) = \frac{\sum_{i=1}^m \frac{g_i \varepsilon_i}{\varepsilon_i + \alpha}}{\sum_{i=1}^m \frac{g_i}{\varepsilon_i + \alpha}} = \frac{\frac{1}{\alpha}\sum_{i=1}^m \frac{g_i \varepsilon_i}{\frac{\varepsilon_i}{\alpha} + 1}}{\frac{1}{\alpha}\sum_{i=1}^m \frac{g_i}{\frac{\varepsilon_i}{\alpha} + 1}} = \frac{1}{m}\sum_{i=1}^m g_i \varepsilon_i, \qquad \text{as} \quad \alpha \to \pm\infty, \tag{3.45}$$

$$E(\alpha) = \frac{\sum_{i=1}^m \frac{g_i \varepsilon_i (\varepsilon_{im} + \alpha)}{\varepsilon_i + \alpha}}{\sum_{i=1}^m \frac{g_i (\varepsilon_{im} + \alpha)}{\varepsilon_i + \alpha}} = \frac{\varepsilon_{im} + \sum_{i=1, i \neq im}^m \frac{g_i \varepsilon_i (\varepsilon_{im} + \alpha)}{\varepsilon_i + \alpha}}{1 + \sum_{i=1, i \neq im}^m \frac{g_i (\varepsilon_{im} + \alpha)}{\varepsilon_i + \alpha}} = g_{im}\varepsilon_{im}, \qquad \text{as} \quad \alpha \to -\varepsilon_1, \tag{3.46}$$

$$E(\alpha) = \frac{\sum_{i=1}^m \frac{g_i \varepsilon_i (\varepsilon_{iM} + \alpha)}{\varepsilon_i + \alpha}}{\sum_{i=1}^m \frac{g_i (\varepsilon_{iM} + \alpha)}{\varepsilon_i + \alpha}} = \frac{\varepsilon_{iM} + \sum_{i=1, i \neq iM}^m \frac{g_i \varepsilon_i (\varepsilon_{iM} + \alpha)}{\varepsilon_i + \alpha}}{1 + \sum_{i=1, i \neq iM}^m \frac{g_i (\varepsilon_{iM} + \alpha)}{\varepsilon_i + \alpha}} = g_{iM}\varepsilon_{iM}, \quad \text{as} \quad \alpha \to -\varepsilon_m. \tag{3.47}$$

The equation (3.45) implies that $E(\alpha) \neq \frac{1}{m}\sum_{i=1}^m g_i \varepsilon_i$, however if we take original system of equations, i.e. (3.37), then for $(\lambda, \nu) = (0, 1)$ we have $E = \frac{1}{m}\sum_{i=1}^m g_i \varepsilon_i$. Taking that into account and equations (3.46), (3.47) we get that α exists if $E \in (g_{im}\varepsilon_{im}, g_{iM}\varepsilon_{iM})$.
Further, since function $E = E(\alpha)$ is strictly decreasing,

$$\begin{aligned}
&\text{if } g_{im}\varepsilon_{im} < E < \overline{g\varepsilon}, && \text{then } \alpha > -\varepsilon_1,\\
&\text{if } E = \overline{g\varepsilon}, && \text{then } \alpha \to \pm\infty,\\
&\text{if } \overline{g\varepsilon} < E < g_{iM}\varepsilon_{iM}, && \text{then } \alpha < -\varepsilon_m.
\end{aligned}$$

Now from (3.40)

$$\begin{aligned}
&\lambda(\alpha) \to \infty && \text{as } \alpha \to -\varepsilon_1\\
&\lambda(\alpha) \to 0 && \text{as } \alpha \to \pm\infty,\\
&\lambda(\alpha) \to -\infty && \text{as } \alpha \to -\varepsilon_m,
\end{aligned}$$

hence if α exists the parameter λ also exists.
Since $\nu = \frac{\alpha}{\lambda}$, λ and α exists and are unique, then the parameter ν also exists and is unique.
Putting the results together we get the outcome of the lemma. □

Lemma 6. For the system of equations (3.30) let

$$x_i = \frac{g_i}{e^{\lambda\varepsilon_i+\nu}-1}, \ i = 1, \ldots, m. \tag{3.48}$$

where λ and ν are some parameters.

Then

i)	if $E = \overline{g\varepsilon}$,	then the solution is $(\lambda, \nu) = (0, \log 2)$,
ii)	if $g_{im}\varepsilon_{im} < E < \overline{g\varepsilon}$,	then $\lambda > 0, \lambda < -\nu\varepsilon_1$,
iii)	if $\overline{g\varepsilon} < E < g_{iM}\varepsilon iM$,	then $\lambda < 0, \lambda > -\nu\varepsilon_m$,
iv)	if $E \notin (g_{im}\varepsilon_{im}, g_{iM}\varepsilon_{iM})$,	then the solution does not exists.

Proof. As in this case the parameters λ and ν in (3.48) cannot be factorized w provide the proof without full rigor regarding existence and uniqueness of parameters. We start with proof of uniqueness of the solution. First we assume it exists. Then we substitute (3.48) into (3.30), and get

$$1 = \sum_{i=1}^{m} \frac{g_i}{e^{\lambda\varepsilon_i+\nu}-1}, \tag{3.49}$$

$$E = \sum_{i=1}^{m} \frac{\varepsilon_i}{e^{\lambda\varepsilon_i+\nu}-1}. \tag{3.50}$$

and then we perform a substitution $\nu = \lambda\alpha$ and for $\lambda \neq 0$ we have

$$1 = \sum_{i=1}^{m} \frac{g_i}{e^{\lambda(\varepsilon_i+\alpha)}-1},$$

$$E = \sum_{i=1}^{m} \frac{g_i\varepsilon_i}{e^{\lambda(\varepsilon_i+\alpha)}-1},$$

Note that by (3.31)

$$\frac{g_i}{e^{\lambda(\varepsilon_i+\alpha)}-1} > 0, \ i = 1, \ldots, m,$$

hence λ and α can take values

$$\lambda > 0, \ \alpha > -\varepsilon_1, \tag{3.52}$$

or

$$\lambda < 0, \ \alpha < -\varepsilon_m, \tag{3.53}$$

For $\lambda = 0$ from the system of equation (3.49) we get that $\nu = \log 2$ and $E = \overline{g\varepsilon}$. Then we get that the solution exists if $E \in (g_{im}\varepsilon_{im}, g_{iM}\varepsilon_{iM})$ as from the second equation of (3.49) we have

$$g_{im}\varepsilon_{im} = g_{im}\varepsilon_{im} \sum_{i=1}^{m} \frac{1}{e^{\lambda\varepsilon_i+\nu} - 1} < \sum_{i=1}^{m} \frac{g_i\varepsilon_i}{e^{\lambda(\varepsilon_i+\alpha)} - 1},$$
$$g_{iM}\varepsilon_{iM} = g_{iM}\varepsilon_{iM} \sum_{i=1}^{m} \frac{1}{e^{\lambda\varepsilon_i+\nu} - 1} > \sum_{i=1}^{m} \frac{g_i\varepsilon_i}{e^{\lambda(\varepsilon_i+\alpha)} - 1},$$

i.e. the weighted sum cannot exceed its highest element or be smaller than lowest.

Since $E = E(\lambda, \nu)$ is a strictly decreasing function w.r.t variable λ we have that for $g_{im}\varepsilon_{im} < E < \overline{g\varepsilon}$ the corresponding parameters λ and ν are in the regime given by (3.52). For the values of E in $g_{iM}\varepsilon_{iM} > E > \overline{g\varepsilon}$ we have the other regime (3.53). Putting together the outcomes we get the final result. □

3.3 Related estimates

In this section we provide a various estimates related to the approximated entropy $f_l(x, N)$ and $f_l(x)$, for all three cases of function $G(N)$ given by (3.2), i.e. $l = 1, 2, 3$. The expressions for those functions are given by Lemma 1 in the first section of this Chapter. The first subsection is devoted to the estimate of the speed of convergence of the maximum of the function $f_l(x, N)$ to the maximum of $f_l(x)$, $l = 1, 2, 3$. The content of the second one are the estimates for the speed of convergence of first derivatives of the functions $f_l(x, N)$ to $f_l(x)$, $l = 1, 2, 3$. In the third subsection we estimate the speed of convergence of the inverse of second derivative matrices of functions $f_l(x, N)$ to $f_l(x)$, $l = 1, 2, 3$.

3.3.1 Estimates for the maximums

Note, that the maximum x^* of the function $f_l(x)$ exists and is unique for some range of parameter E and the domain Ω_E, which was shown in the Lemma 2 of the Section 2. Similar situation is with $f_l(x, N)$ shown in Lemma 3 of Section 2. For some $N > N_0$, $f_l(x, N)$ has a unique maximum $x^*(N)$ for $E > \varepsilon_1$ as then the domain is nonempty. Then we have following result

Proposition 3. For the maximum points $x^*(N)$ and x^* of the functions, respectively, $f_l(x, N)$ and $f_l(x)$ with $l = 1, 2, 3$ over the domain Ω we have following estimates for each case l

1. $\frac{G(N)}{N} \to \infty$

$$|x^* - x^*(N)| \le N^{-1+\delta}, \qquad \text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$
$$|x^* - x^*(N)| \le \left(\frac{N}{G(N)}\right)^{-1+\delta}, \qquad \text{when } \frac{1}{N} \ll \frac{N}{G(N)}.$$

2. $\frac{G(N)}{N} \to c$

$$|x^* - x^*(N)| \le N^{-1+\delta}.$$

3. $\frac{G(N)}{N} \to 0$

$$|x^* - x^*(N)| \le G(N)^{-1+\delta}, \qquad \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$
$$|x^* - x^*(N)| \le \left(\frac{G(N)}{N}\right)^{-1+\delta}, \qquad \text{when } \frac{1}{G(N)} \ll \frac{G(N)}{N}.$$

valid for sufficiently large N, where δ is some arbitrary small positive constant and the symbol $\gg$ is defined

$$f(x) \gg g(x), \text{ as } x \to \infty \iff \lim_{x\to\infty} \frac{f(x)}{g(x)} > 1.$$

Proof. We proof the theorem separately for each case.

1. $\frac{G(N)}{N} \to \infty$

From the proof of the previous section we have Lagrangian

$$L(x, \lambda^*, \nu^*) = -f_1(x) + \lambda^*\big(\varepsilon^T x - E\big) + \nu^*\big(\mathbf{1}^T x - 1\big),$$

where λ^* are ν^* are some parameters. Then one of the conditions for the maximum to exists and be unique , i.e. KKT conditions, is that that $\nabla_x L(x, \lambda^*, \nu^*) = 0$.
The similar situation will be for the function $f_1(x, N)$, in the Lagrangian we will have $f_1(x, N)$ instead of $f_1(x)$ and the corresponding Lagrangian will be $L(x, N, \lambda^*, \nu^*)$. Now we calculate explicit derivative of the $L(x, N, \lambda^*, \nu^*)$ and $L(x, \lambda^*, \nu^*)$.
For $i = 1, \ldots, m$ the partial derivatives are

$$\begin{aligned}
&\frac{\partial}{\partial x_i} L(x,\lambda^*, \nu^*) = \ln \frac{g_i}{x_i} + \lambda^*\varepsilon + \nu^*,\\
&\frac{\partial}{\partial x_i} L(x,N, \lambda^*, \nu^*) = \ln \frac{g_i}{x_i} + \ln\left(1 + \frac{x_i N}{g_i G(N)}\right) - \frac{1}{2}\frac{1}{x_i N + g_i G(N)} -\\
&\qquad - \ln\left(1 + \frac{1}{x_i N}\right) + \frac{1}{2}\frac{1}{x_i N + 1} + \lambda^*\varepsilon + \nu^*.
\end{aligned}$$

Hence we can also write that for the maximum point x^*

$$\begin{aligned}
&|DL(x^*, \lambda^*, \nu^*)| = 0,\\
&|DL(x^*(N), \lambda^*, \nu^*)| = \left| \ln\left(1 + \frac{x^*(N)N}{gG(N)}\right) - \frac{1}{2}\frac{1}{x^*(N)N + gG(N)} - \right.\\
&\qquad \left. - \ln\left(1 + \frac{1}{x^*(N)N}\right) + \frac{1}{2}\frac{1}{x^*(N)N + 1}\right|,
\end{aligned}$$

where in the second equation we substituted the first one.
In the next step we approximate the logarithms of the second equation using the approximation $\ln(1 + z) = z + O(z^2), z \to 0$ for any $|z| < 1$

$$\begin{aligned}
|DL(x^*(N),\lambda^*, \nu^*)| = &\left| \frac{xN}{gG(N)} + O\bigg(\frac{xN}{gG(N)}\bigg)^2 - \frac{1}{2}\frac{1}{xN + gG(N)} - \frac{1}{xN} + \right.\\
&\left. + O\bigg(\frac{1}{xN}\bigg)^2 + \frac{1}{2}\frac{1}{xN + 1}\right|.
\end{aligned}$$

Now transform RHS of above expression depending on which term is asymptotically stronger

$$
\begin{aligned}
|DL(x^*(N),\lambda^*,\nu^*)| &= \frac{1}{N}\left|\frac{x^*(N)N^2}{gG(N)} + O\left(\frac{x^*(N)N}{gG(N)}\right)^2 N - \frac{1}{2}\frac{1}{x^*(N)+g\frac{G(N)}{N}} - \frac{1}{x^*(N)} + \right. \\
&\quad \left. + O\left(\frac{1}{x^*(N)N}\right)^2 + \frac{1}{2}\frac{1}{x^*(N)+\frac{1}{N}}\right|, \qquad \text{when } \frac{1}{N} \gg \frac{N}{G(N)}, \\
|DL(x^*(N),\lambda^*,\nu^*)| &= \frac{N}{G(N)}\left|\frac{x^*(N)(N)}{g} + O\left(\frac{x^*(N)^2 N}{g^2 G(N)}\right) - \frac{1}{2}\frac{1}{x^*(N)G(N)+g\frac{G(N)^2}{N}} - \frac{1}{x^*(N)}\right. \\
&\quad \left. + O\left(\frac{1}{x^*(N)^2 G(N) N}\right) + \frac{1}{2}\frac{1}{x^*(N)+\frac{G(N)}{N}}\right|, \qquad \text{when } \frac{1}{N} \ll \frac{N}{G(N)}.
\end{aligned}
$$

Hence we can write

$$
\begin{aligned}
|DL(x^*(N),\lambda^*,\nu^*)| &\le \frac{1}{N}K_{1,1/N}, &\text{when } \frac{1}{N} \gg \frac{N}{G(N)}, \\
|DL(x^*(N),\lambda^*,\nu^*)| &\le \frac{N}{G(N)}K_{1,N/G}, &\text{when } \frac{1}{N} \ll \frac{N}{G(N)}.
\end{aligned}
\tag{3.54}
$$

where $K_{1,1/N}$ and $K_{1,N/G}$ are some constants depending on $x^*(N)$.
Then we approximate $|DL(x,\lambda^*,\nu^*)|$ with the first order Taylor expansion at the point x^*

$$|DL(x,\lambda^*,\nu^*)| = |DL(x^*,\lambda^*,\nu^*) + D^2L(x_\theta,\lambda^*,\nu^*)(x-x^*)|,$$

where x_θ is some between points x and x^*.
The first term on the RHS of above expansion is equal to 0 since it is condition for the maximum and $D^2L(x_\theta,\lambda^*,\nu^*)$ simply becomes $D^2f_1(x_\theta)$ since the expressions with λ^* and ν^* are equal to 0 when we make differentiation w.r.t x_i. Hence we have

$$|DL(x,\lambda^*,\nu^*)| = |D^2f_1(x_\theta)(x-x^*)|.$$

Further we bound from above the RHS to extract the vector $|x-x^*|$ and get

$$|DL(x,\lambda^*,\nu^*)| \ge F_{1,\theta}^{'(2)}|x-x^*|,$$

and then as x we take point on the ball x_B separated by $N^{-1+\delta}$ and $\big(N/G(N)\big)^{1-\delta}$ from the maximum, i.e. $|x_B - x^*| = N^{-1+\delta}$ or $|x_B - x^*| = (N/G(N))^{1-\delta}$, where δ is arbitrary small positive constant. Hence we can write

$$|DL(x_B, \lambda^*, \nu^*)| \geq F_{1,\theta}^{'(2)} N^{-1+\delta}, \qquad \text{when } \frac{1}{N} \gg \frac{N}{G(N)}, \tag{3.55}$$

$$|DL(x_B, \lambda^*, \nu^*)| \geq F_{1,\theta}^{'(2)} \left(\frac{N}{G(N)}\right)^{1-\delta}, \qquad \text{when } \frac{1}{N} \ll \frac{N}{G(N)}.$$

Since $f \in C^2$ and $\det D^2 f_1(x^*) \neq 0$, then by the inverse function theorem we have that the mapping $\mathcal{M} : x \to DL(x, \lambda^*, \nu^*)$ is invertible in the neighborhood of x^* and the inverse function is in the class C^1. Hence, from the estimates (3.54) and (3.55) together with $DL(x^*, \lambda^*, \nu^*) = 0$ and knowing that $|x_B - x^*| = N^{-1+\delta}$ or $|x_B - x^*| = (N/G(N))^{-1+\delta}$ we can infer following estimates

$$|x^* - x^*(N)| \leq N^{-1+\delta}, \qquad \text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$

$$|x^* - x^*(N)| \leq \left(\frac{N}{G(N)}\right)^{-1+\delta}, \qquad \text{when } \frac{1}{N} \ll \frac{N}{G(N)},$$

valid for sufficiently large N.

2. $\frac{G(N)}{N} \to c$
Here the approach is analogical. We first find the upper bound of $|DL(x^*(N), \lambda^*, \nu^*)|$, then appropriate lower bound for x_B and corresponding $|DL(x_B, \lambda^*, \nu^*)|$ and then infer the estimate $|x^* - x^*(N)|$.
In this case we have

$$|DL(x^*, \lambda^*, \nu^*)| = 0,$$

$$|DL(x^*(N), \lambda^*, \nu^*)| = \left| -\frac{1}{2}\frac{1}{x^*(N)N + gG(N)} - \ln\left(1 + \frac{1}{x^*(N)N}\right) + \frac{1}{2}\frac{1}{x^*(N)N + 1}\right|,$$

and approximating logarithm for the second equation we have

$$\begin{aligned}|DL(x^*(N), \lambda^*, \nu^*)| = \Bigg| &-\frac{1}{2}\frac{1}{x^*(N)N + gG(N)} - \frac{1}{x^*(N)N} - \\ &- O\left(\frac{1}{x^*(N)N}\right)^2 + \frac{1}{2}\frac{1}{x^*(N)N + 1}\Bigg| = \\ &= \frac{1}{N}\left| -\frac{1}{2}\frac{1}{x^*(N) + g\frac{G(N)}{N}} - \frac{1}{x^*(N)} - O\left(\frac{1}{x^*(N)^2 N}\right) + \frac{1}{2}\frac{1}{x^*(N) + \frac{1}{N}}\right|,\end{aligned}$$

hence

$$|DL(x^*(N), \lambda^*, \nu^*)| \le \frac{1}{N} K_2,$$

valid for large enough N, where K_2 is some positive constant.
We choose $|x_B - x^*| = N^{-1+\delta}$ where $\delta > 0$ is arbitrary small constant. Then analogically to previous case

$$|DL(x_B, \lambda^*, \nu^*)| \ge F_{2,\theta}^{'(2)} N^{-1+\delta},$$

Since by the inverse function theorem the mapping $\mathcal{M} : x \to DL(x, \lambda^*, \nu^*)$ is invertible and the inverse is continuous we can infer that

$$|x^* - x^*(N)| \le N^{-1+\delta}.$$

3. $\frac{G(N)}{N} \to 0$ Similarly we perform estimates in the last case.
   Firstly, we have

$$|DL(x^*, \lambda^*, \nu^*)| = 0,$$
$$|DL(x^*(N), \lambda^*, \nu^*)| = \left| O\left(\frac{g^2 G(N)}{x^*(N)^2 N}\right) - \frac{1}{2}\frac{1}{x^*(N)N + gG(N)} - \frac{1}{x^*(N)N} + \right.$$
$$\left. - O\left(\frac{1}{x^*(N)N}\right)^2 + \frac{1}{2}\frac{1}{x^*(N)N+1}\right|,$$

and we have to types of transformation for the second equation

$$|DL(x^*(N), \lambda^*, \nu^*)| = \frac{1}{G(N)}\left| O\left(\frac{g^2 G(N)^2}{x^*(N)^2 N}\right) - \frac{1}{2}\frac{1}{x^*(N)\frac{N}{G(N)} + g} - \frac{G(N)}{x^*(N)N} + \right.$$
$$\left. - O\left(\frac{1}{x^*(N)N}\right)^2 G(N) + \frac{1}{2}\frac{G(N)}{x^*(N)N+1}\right|, \qquad \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$
$$|DL(x^*(N), \lambda^*, \nu^*)| = \frac{G(N)}{N}\left| O\left(\frac{g^2}{x^*(N)^2}\right) - \frac{1}{2}\frac{1}{x^*(N)\frac{N^2}{G(N)} + gN} - \frac{1}{x^*(N)G(N)} + \right.$$
$$\left. - O\left(\frac{1}{x^*(N)^2 G(N) N}\right) + \frac{1}{2}\frac{1}{x^*(N)G(N) + \frac{G(N)}{N}}\right|, \text{ when } \frac{1}{G(N)} \ll \frac{G(N)}{N},$$

hence

$$|DL(x^*(N),\lambda^*,\nu^*)| \leq \frac{1}{G(N)}K_{3,1/G}, \qquad \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$
$$|DL(x^*(N),\lambda^*,\nu^*)| \leq \frac{G(N)}{N}K_{3,G/N}, \qquad \text{when } \frac{1}{G(N)} \ll \frac{G(N)}{N},$$

holding for some sufficiently large N, where $K_{3,1/G} > 0$ and $K_{3,G/N}$ are some constants. Now, analogically to other cases, we chose x_B such that $|x_B - x^*| = G(N)^{-1+\delta}$ or $|x_B - x^*| = (G(N)/N)^{1-\delta}$, where δ is arbitrary small positive constant. Then we have following estimates

$$|DL(x_B,\lambda^*,\nu^*)| \geq F^{'(2)}_{3,\theta}G(N)^{-1+\delta}, \qquad \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$
$$|DL(x_B,\lambda^*,\nu^*)| \geq F^{'(2)}_{3,\theta}\left(\frac{G(N)}{N}\right)^{1-\delta}, \qquad \text{when } \frac{1}{G(N)} \ll \frac{G(N)}{N}.$$

Since the mapping $\mathcal{M} : x \to DL(x,\lambda^*,\nu^*)$ is invertible and the invers is continuous we can infer that

$$|x^* - x^*(N)| \leq G(N)^{-1+\delta}, \qquad \text{when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$
$$|x^* - x^*(N)| \leq \left(\frac{G(N)}{N}\right)^{-1+\delta}, \qquad \text{when } \frac{1}{G(N)} \ll \frac{G(N)}{N}.$$

Putting together the outcomes for each case we get the result of the lemma. $\square$

3.3.2 Estimates for the first derivatives

Proposition 4. Given the functions $f_l(x,N)$ and $f_l(x)$, $l=1,2,3$, for the first derivative w.r.t. x_1 we have following estimates

$$
\begin{aligned}
&1)\ f_1'(x,N)=f_1'(x)+K_1\frac{1}{N}, && \text{when } \frac{1}{N}\gg\frac{N}{G(N)},\\
&\quad\ f_1'(x,N)=f_1'(x)+K_1\frac{N}{G(N)}, && \text{when } \frac{1}{N}\ll\frac{N}{G(N)},\\
&2)\ f_2'(x,N)=f_2'(x)+K_2\frac{1}{N}, && \\
&3)\ f_3'(x,N)=f_3'(x)+K_3\frac{1}{G(N)}, && \text{when } \frac{1}{G(N)}\gg\frac{G(N)}{N},\\
&\quad\ f_3'(x,N)=f_3'(x)+K_3\frac{G(N)}{N}, && \text{when } \frac{1}{G(N)}\ll\frac{G(N)}{N},
\end{aligned}
$$

valid for sufficiently large N, where K_1,K_2 and K_3 are some positive constants.

Proof. We start by taking the explicit expression for the function $f_l(x,N)$ and $f_l(x)$, we consider all thre cases of $G(N)$ separately.

1. $\frac{G(N)}{N}\to\infty$

$$
\begin{aligned}
f_1(x,N)=&\sum_{i=1}^{m}\Bigg[x_i\ln\frac{g_i}{x_i}+\left(x_i+g_i\frac{G(N)}{N}\right)\ln\left(1+\frac{x_iN}{g_iG(N)}\right)-\\
&-\frac{1}{2N}\ln(x_iN+g_iG(N))-x_i\ln\left(1+\frac{1}{x_iN}\right)-\frac{1}{2N}\ln(x_iN+1)\Bigg],
\end{aligned}
$$

and

$$
f_1(x)=\sum_{i=1}^{m}\left[x_i\ln\frac{g_i}{x_i}+x_i\right],
$$

Next we differentiate function $f_1(x,N)$ w.r.t. x_1 and obtain following

$$
\begin{aligned}
f_1'(x,N)=&\sum_{i=1}^{m}\Bigg[\ln\frac{g_i}{x_i}+\ln\left(1+\frac{x_iN}{g_iG(N)}\right)-\frac{1}{2}\frac{1}{x_iN+g_iG(N)}-\\
&-\ln\left(1+\frac{1}{x_iN}\right)+\frac{1}{2x_iN+2}\Bigg],
\end{aligned}
$$

and also differentiate $f_1(x)$

$$f_1'(x) = \sum_{i=1}^{m} \ln \frac{g_i}{x_i},$$

Then we apply approximation of logarithm $\ln(1+x) = x + O(x^2),\ x \in (0,1)$ for some terms of $f_1(x, N)$ and substitute the function $f_1(x)$

$$f_1'(x,N) = f_1'(x) + \sum_{i=1}^{m} \left[\frac{x_i N}{g_i G(N)} + O\left(\frac{x_i N}{g_i G(N)}\right)^2 - \frac{1}{2}\frac{1}{x_i N + g_i G(N)} - \right.$$
$$\left. - \frac{1}{x_i N} - O\left(\frac{1}{x_i N}\right)^2 + \frac{1}{2x_i N + 2}\right],$$

and perform some convenient manipulations

$$f_1'(x,N) = f_1'(x) + \sum_{i=1}^{m} \left[\frac{N}{G(N)}\left(\frac{x_i}{g_i} + O\left(\frac{x_i^2 N}{g_i^2 G(N)}\right)\right) + \frac{1}{N}\left(- \frac{1}{2}\frac{1}{x_i + g_i G(N)/N} - \right.\right.$$
$$\left.\left. - \frac{1}{x_i} - O\left(\frac{1}{x_i^2 N}\right) + \frac{1}{2x_i + 2/N}\right)\right].$$

Now, depending which term is asymptotically 'stronger' N or $G(N)/N$ we distinguish two cases

$$f_1'(x,N) = f_1'(x) + \frac{1}{N}\sum_{i=1}^{m} \left[\frac{N^2}{G(N)}\left(\frac{x_i}{g_i} + O\left(\frac{x_i^2 N}{g_i^2 G(N)}\right)\right) + \left(- \frac{1}{2}\frac{1}{x_i + g_i G(N)/N} - \right.\right.$$
$$\left.\left. - \frac{1}{x_i} - O\left(\frac{1}{x_i^2 N}\right) + \frac{1}{2x_i + 2/N}\right)\right],\ \text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$
$$f_1'(x,N) = f_1'(x) + \frac{N}{G(N)}\sum_{i=1}^{m} \left[\left(\frac{x_i}{g_i} + O\left(\frac{x_i^2 N}{g_i^2 G(N)}\right)\right) + \frac{G(N)}{N^2}\left(- \frac{1}{2}\frac{1}{x_i + g_i G(N)/N} - \right.\right.$$
$$\left.\left. - \frac{1}{x_i} - O\left(\frac{1}{x_i^2 N}\right) + \frac{1}{2x_i + 2/N}\right)\right],\ \text{when } \frac{1}{N} \ll \frac{N}{G(N)},$$

hence we get the final result for the first case,

$$f_1'(x,N) = f_1'(x) + K_1 \frac{1}{N},\ \text{when } \frac{1}{N} \gg \frac{N}{G(N)},$$
$$f_1'(x,N) = f_1'(x) + K_1 \frac{N}{G(N)},\ \text{when } \frac{1}{N} \ll \frac{N}{G(N)},$$

where the constant K_1 for two cases are the bounds

$$K_1 \geq \sum_{i=1}^{m}\Bigg[\frac{N^2}{G(N)}\bigg(\frac{x_i}{g_i}+O\bigg(\frac{x_i^2 N}{g_i^2 G(N)}\bigg)\bigg)+\bigg(-\frac{1}{2}\frac{1}{x_i+g_i G(N)/N}- \\ -\frac{1}{x_i}-O\bigg(\frac{1}{x_i^2 N}\bigg)+\frac{1}{2x_i+2/N}\bigg)\Bigg], \text{ when } \frac{1}{N}\gg\frac{N}{G(N)},$$

$$K_1 \geq \sum_{i=1}^{m}\Bigg[\bigg(\frac{x_i}{g_i}+O\bigg(\frac{x_i^2 N}{g_i^2 G(N)}\bigg)\bigg)+\frac{G(N)}{N^2}\bigg(-\frac{1}{2}\frac{1}{x_i+g_i G(N)/N}- \\ -\frac{1}{x_i}-O\bigg(\frac{1}{x_i^2 N}\bigg)+\frac{1}{2x_i+2/N}\bigg)\Bigg], \text{ when } \frac{1}{N}\ll\frac{N}{G(N)}.$$

2. $\frac{G(N)}{N} \to c$
For the second case we repeat the step. Differentiate $f_2(x,N)$ and $f_2(x)$ w.r.t. x_1, then approximate logarithms and substitute $f_2'(x)$ into $f_2'(x,N)$, eventually we get expression

$$f_2'(x,N) = f_2'(x)+\frac{1}{N}\sum_{i=1}^{m}\Bigg[-\frac{1}{2}\frac{1}{x_i+g_i G(N)/N}-\frac{1}{x_i}-O\bigg(\frac{1}{x_i^2 N}\bigg)+ \\ +\frac{1}{2x_i+2/N}\Bigg],$$

hence we get the final result with constant $K_2 > 0$ is defined

$$K_2 \geq \sum_{i=1}^{m}\Bigg[-\frac{1}{2}\frac{1}{x_i+g_i G(N)/N}-\frac{1}{x_i}-O\bigg(\frac{1}{x_i^2 N}\bigg)+\frac{1}{2x_i+2/N}\Bigg],$$

3. $\frac{G(N)}{N} \to 0$ Here perform again analogical steps as in the first case. This time, however, the cases are distinguished depending whether $G(N)$ or $\frac{N}{G(N)}$

is 'stronger'

$$f_3'(x,N) = f_3'(x) + \frac{1}{G(N)}\sum_{i=1}^{m}\Bigg[O\bigg(\frac{g_i^2 G(N)^2}{x_i^2 N}\bigg) + \bigg(-\frac{1}{2}\frac{1}{x_i + g_i G(N)/N} - \\ -\frac{1}{x_i} - O\bigg(\frac{1}{x_i^2 N}\bigg) + \frac{1}{2x_i + 2/N}\bigg)\Bigg], \text{ when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$

$$f_3'(x,N) = f_3'(x) + \frac{G(N)}{N}\sum_{i=1}^{m}\Bigg[O\bigg(\frac{g_i^2}{x_i^2}\bigg) + \frac{G(N)}{N^2}\bigg(-\frac{1}{2}\frac{1}{x_i + g_i G(N)/N} - \\ -\frac{1}{x_i} - O\bigg(\frac{1}{x_i^2 N}\bigg) + \frac{1}{2x_i + 2/N}\bigg)\Bigg], \text{ when } \frac{1}{G(N)} \ll \frac{G(N)}{N},$$

and K_3 for each case is defined

$$K_3 \geq \sum_{i=1}^{m}\Bigg[O\bigg(\frac{g_i^2 G(N)^2}{x_i^2 N}\bigg) + \bigg(-\frac{1}{2}\frac{1}{x_i + g_i G(N)/N} - \frac{1}{x_i} - O\bigg(\frac{1}{x_i^2 N}\bigg) + \\ + \frac{1}{2x_i + 2/N}\bigg)\Bigg], \text{ when } \frac{1}{G(N)} \gg \frac{G(N)}{N},$$

$$K_3 \geq \sum_{i=1}^{m}\Bigg[O\bigg(\frac{g_i^2}{x_i^2}\bigg) + \frac{G(N)}{N^2}\bigg(-\frac{1}{2}\frac{1}{x_i + g_i G(N)/N} - \frac{1}{x_i} - O\bigg(\frac{1}{x_i^2 N}\bigg) + \\ + \frac{1}{2x_i + 2/N}\bigg)\Bigg], \text{ when } \frac{1}{G(N)} \ll \frac{G(N)}{N},$$

hence we get the result of the lemma for the third case.

□

3.3.3 Estimates for the second derivatives

Proposition 5. Given the functions $f_l(x,N)$ and $f_l(x)$, $l = 1,2,3$, for the second derivative matrices we have following estimates

$$1)\ D^2 f_1(x,N)^{-1} = D^2 f_1(x)^{-1} + \frac{\kappa_1}{N},$$
$$2)\ D^2 f_2(x,N)^{-1} = D^2 f_2(x)^{-1} + \frac{\kappa_2}{N},$$
$$3)\ D^2 f_3(x,N)^{-1} = D^2 f_3(x)^{-1} + \frac{\kappa_3}{G(N)},$$

which holds for large enough N and K_1, K_2, K_3 are some positive constants.

Proof. We prove each case of $G(N)$ separately starting form the first one

1. $\frac{G(N)}{N} \to \infty$

We take the explicit expression for the functions $f_1(x,N)$ and $f_1(x)$ and differentiate twice w.r.t. x_i, $i = 1, \ldots, m$,

$$\begin{aligned}\frac{\partial^2}{\partial x_i^2} f_1(x,N) = &-\frac{1}{x_i} + \frac{1}{g_i G(N)/N + x_i} + \frac{1}{2(x_i N + g_i G(N))(x_i + g_i G(N)/N)} + \\ &+ \frac{1}{x_i^2 N + x_i} - \frac{1}{2(x_i N + 1)(x_i + 1/N)},\end{aligned}$$

and

$$\frac{\partial^2}{\partial x_i^2} f_1(x) = -\frac{1}{x_i},$$

where the mixed derivatives are equal to zero as the considered functions can be decomposed $f_1(x,N) = \sum_{i=1}^m f_1(x_i,N)$ and $f_1(x) = \sum_{i=1}^m f_1(x_i)$. Then we substitute derivatives of $f_1(x)$ into $f_1(x,N)$ and after some manipulations we get

$$\begin{aligned}\frac{\partial^2}{\partial x_i^2} f_1(x,N) =& \frac{\partial^2}{\partial x_i^2} f_1(x) + \frac{1}{N}\Bigg(\frac{1}{g_i G(N)/N^2 + x_i/N} + \frac{1}{x_i^2 + x_i/N} + \\ &+ \frac{1}{2(x_i + g_i G(N)/N)(x_i + g_i G(N)/N)} - \frac{1}{2(x_i + 1/N)(x_i + 1/N)}\Bigg),\end{aligned}$$

and we introduce the constant K_1 such that

$$\begin{aligned}K_1 \geq& \frac{1}{g_i G(N)/N^2 + x_i/N} + \frac{1}{x_i^2 + x_i/N} - \frac{1}{2(x_i + 1/N)(x_i + 1/N)} + \\ &+ \frac{1}{2(x_i + g_i G(N)/N)(x_i + g_i G(N)/N)}.\end{aligned}$$

Hence we have

$$\frac{\partial^2}{\partial x_i^2} f_1(x,N) = \frac{\partial^2}{\partial x_i^2} f_1(x) + \frac{K_{1,i}}{N}.$$

As the mixed derivatives vanish, the matrix $D^2 f_1(x,N)$ is a diagonal matrix, hence its inverse is simple the inverse its elements.

Therefore, the inverse of the second derivative of $f_1(x,N)$ is equal to

$$\frac{1}{\frac{\partial^2}{\partial x_i^2} f_1(x,N)} = \frac{1}{\frac{\partial^2}{\partial x_i^2} f_1(x) + \frac{K_{1,i}}{N}},$$

and after some manipulations we have that

$$\frac{1}{\frac{\partial^2}{\partial x_i^2} f_1(x,N)} = \frac{1}{\frac{\partial^2}{\partial x_i^2} f_1(x)} + \frac{K'_{1,i}}{N},$$

where $K'_{1,i}$ is some positive constant. Putting above expression into matrix form we get

$$D^2 f_1(x,N)^{-1} = D^2 f_1(x)^{-1} + \frac{\kappa_1}{N},$$

where κ_1 is a diagonal matrix with elements $K'_{1,i}$, $i = 1, \ldots, m$. Hence we get the result of the lemma for the first case.

2. $\frac{G(N)}{N} \to c$
   Here analogically to previous case, we differentiate twice functions $f_2(x,N)$ and $f_2(x)$, then substitute second derivative of $f_2(x)$ into the derivative of $f_2(x,N)$ and perform some rearrangements. We obtain following

$$\begin{aligned}\frac{\partial^2}{\partial x_i^2} f_2(x,N) =& \frac{\partial^2}{\partial x_i^2} f_2(x) + \frac{1}{N}\Bigg(\frac{1}{2(x_i + g_i G(N)/N)(x_i + g_i G(N)/N)} + \\ &+ \frac{1}{x_i^2 + x_i/N} - \frac{1}{2(x_i + 1/N)(x_i + 1/N)}\Bigg).\end{aligned}$$

   Then we introduce constant $K_{2,i}$

$$\begin{aligned}K_2 \geq & \frac{1}{x_i^2 + x_i/N} - \frac{1}{2(x_i + 1/N)(x_i + 1/N)} + \\ &+ \frac{1}{2(x_i + g_i G(N)/N)(x_i + g_i G(N)/N)},\end{aligned}$$

   and finally, in the matrix form, we get

$$\frac{\partial^2}{\partial x_i^2} f_2(x,N) = \frac{\partial^2}{\partial x_i^2} f_2(x) + \frac{K_{2,i}}{N}.$$

   As in the previous case we perform the inversion of that expression in order to get estimate for inverted matrices

$$\frac{1}{\frac{\partial^2}{\partial x_i^2} f_2(x,N)} = \frac{1}{\frac{\partial^2}{\partial x_i^2} f_2(x)} + \frac{K'_{2,i}}{N},$$

and eventually we get the final estimate

$$D^2 f_2(x, N)^{-1} = D^2 f_2(x)^{-1} + \frac{\kappa_2}{N},$$

where κ_2 is matrix with the diagonal elements $K'_{2,i}$, $i = 1, \ldots, m$.

3. $\frac{G(N)}{N} \to 0$
   In this case similarly we get the estimate for the second derivatives

$$\begin{aligned}\frac{\partial^2}{\partial x_i^2} f_3(x, N) =& \frac{\partial^2}{\partial x_i^2} f_3(x) + \frac{1}{G(N)}\biggl(\frac{1}{2(x_i N/G(N) + g_i)(x_i N/G(N) + g_i)} + \\ &+ \frac{1}{x_i^2 N/G(N) + x_i/G(N)} - \frac{1}{2(x_i N/G(N) + 1/G(N))(x_i + 1/N)}\biggr),\end{aligned}$$

and we define constant K_3

$$\begin{aligned}K_3 \geq& \biggl(\frac{1}{2(x_i N/G(N) + g_i)(x_i N/G(N) + g_i)} + \\ &+ \frac{1}{x_i^2 N/G(N) + x_i/G(N)} - \frac{1}{2(x_i N/G(N) + 1/G(N))(x_i + 1/N)}\biggr).\end{aligned}$$

Then we obtain the estimate for the inverse of the diagonal element

$$\frac{1}{\frac{\partial^2}{\partial x_i^2} f_3(x, N)} = \frac{1}{\frac{\partial^2}{\partial x_i^2} f_3(x)} + \frac{K'_{3,i}}{G(N)},$$

where $K'_{3,i}$ is some positive constant. Finally after putting above into matrix form for $i = 1, \ldots, m$, we get the result of the lemma, i.e.

$$D^2 f_3(x, N)^{-1} = D^2 f_3(x)^{-1} + \frac{\kappa_3}{G(N)},$$

where κ_3 is a diagonal matrix with elements $K'_{3,i}$, $i = 1, \ldots, m$.

□

3.4 Partition function approximation

In this section we introduce a Lemma for approximation of the sum of the partition function given by (3.3) but with the approximated entropy in the exponent instead of entropy with the appropriate integral.

Lemma 7. For defined above Ω_E, functions $f_l(x,N)$ and $R_l(N)$, $l = 1,2,3$ by Lemma 1 Section 1 of this Chapter we have approximation for each case of G(N)

$$1)\ \sum_{\Omega_E} e^{Nf_1(x,N)+R_1(N)} = \int_{\Omega_E} e^{Nf_1(x,N)+R_1(N)} dx \left(1 + O\left(\frac{1}{N}\right)\right),$$

$$2)\ \sum_{\Omega_E} e^{Nf_2(x,N)+R_2(N)} = \int_{\Omega_E} e^{Nf_2(x,N)+R_2(N)} dx \left(1 + O\left(\frac{1}{N}\right)\right),$$

$$3)\ \sum_{\Omega_E} e^{G(N)f_3(x,N)+R_3(N)} = \int_{\Omega_E} e^{G(N)f_3(x,N)+R_3(N)} dx \left(1 + O\left(\frac{1}{G(N)}\right)\right),$$

as $N \to \infty$.

Proof. Several approaches were taken in order to prove the theorem, however the proof is more complicated than anticipated and has been only partially completed, although the result seems intuitively correct. $\square$

4 Extended Laplace approximation

We consider a more general integral type than original Laplace one. The major difference lies in the dependence of the function in the exponent and its maximum on the limiting parameter, hence we call it 'Extended'. Those results are potentially publishable and the general method was based on the results in the [13], however literature review would have to be done for the confirmation.

In the beginning of the chapter we introduce the Extended integral itself. We consider two type of maximums, on the boundary of the domain and in the interior of the domain. Both types are used in the main result of the thesis.

In the first section we introduce a theorem when the maximum is on the boundary of the domain, where the point of maximum is not a critical point. The space on which integration is performed is one-dimensional.

In the second section we provide the approximation when the maximum is in the interior of the domain. The dimension of the space is finite.

The last section is for finite space when the maximum is on the boundary of the domain, and it is not a critical point.

We consider the integral

$$I(N) = \int_{\Omega} g(x) e^{Nf(x,N)} dx, \tag{4.1}$$

where $N > 0$ and

I) Set Ω is a convex subset of $\mathbb{R}^m$ with nonempty interior.

II) Function $f : \mathbb{R}^m \times \mathbb{N} \to \mathbb{R}$ is such that

i) for all $N \geq N_0$

$$\max_{x \in \Omega} f(x, N) = f(x^*(N), N)$$

and $x^*(N)$ is unique,

ii) $f \in C^n$, where $n \in \mathbb{N}$, N is fixed and $N \geq N_0$.

III) Function $g : \mathbb{R}^m \to \mathbb{R}$, $g \in C^k$ exists.

Further, we consider two types of maximum of f, the point $x^*(N)$ can be

- in the interior of Ω,
- on the boundary Ω the point of maximum is not a critical point.

4.1 One-dimensional function with the maximum on the boundary of the domain

The approximation of the integral (4.1), i.e.

$$I(N) = \int_\Omega g(x) e^{Nf(x,N)} dx. \tag{4.2}$$

with $\Omega = (x^*, a)$ and $f \in C^2$ and $g \in C^1$, is given by the following theorem

Theorem 6. For the above integral there exists $K > 0$ such that for sufficiently large N we have approximation

$$\begin{aligned}\left| \int_\Omega g(x) e^{Nf(x,N)} dx - g(x^*) e^{Nf(x^*,N)} \frac{1}{N} \frac{1}{|f'(x^*,N)|} \right| &\leq \\ \leq \frac{K}{N} |g(x^*)| e^{Nf(x^*,N)} \frac{1}{N} \frac{1}{|f'(x^*,N)|}&.\end{aligned}$$

Proof. Since

$$\frac{1}{N} \frac{1}{|f'(x^*,N)|} = \int_{x^*}^{\infty} e^{-N|f'(x^*,N)|(x-x^*)} dx,$$

we define

$$I_G(N) = g(x^*) e^{Nf(x^*,N)} \int_{x^*}^{\infty} e^{-N|f'(x^*,N)|(x-x^*)} dx. \tag{4.3}$$

Now, we introduce the set

$$U_N = \{x : |x - x^*| \leq \frac{1}{N^{1/2}},\ N \geq N_0\},$$

and decompose the integral (4.3) into two integrals, one over $U_N(x^*)$ and second over $\mathbb{R}\backslash U_N$

$$
\begin{aligned}
I_G(N) =& I_{G1}(N) + I_{G2}(N) = g(x^*)e^{Nf(x^*,N)} \int_{U_N} e^{-N|f'(x^*,N)|(x-x^*)}dx+ \\
&+ g(x^*)e^{Nf(x^*,N)} \int_{\mathbb{R}\backslash U_N} e^{-N|f'(x^*,N)|(x-x^*)}dx. \qquad (4.4)
\end{aligned}
$$

Let us use Taylor's Theorem to get 1-st order expansion of the function $g(x)$ at the point x^*,

$$g(x) = g(x^*) + g'(x_\theta)(x - x^*),$$

where x_θ is some point between x^* and x, and can be formally represented $x_\theta = x^* + \theta(x - x^*)$, $0 \leq \theta \leq 1$.
Then we substitute it into $I(N)$ and separate integrals, one with $g(x^*)$ and second with the other term of expansion

$$
\begin{aligned}
I(N) =& I_1(N) + I_2(N) = g(x^*) \int_{x^*}^{a} e^{Nf(x,N)}dx+ \\
&+ \int_{x^*}^{a} g'(x_\theta)(x - x^*)e^{Nf(x,N)}dx. \qquad (4.5)
\end{aligned}
$$

Next we decompose $I_1(N)$ into two, one over U_N and second over $(x^*, a)\backslash U_N$

$$I_1(N) = I_{11}(N) + I_{12}(N) = g(x^*) \int_{U_N} e^{Nf(x,N)}dx + g(x^*) \int_{x^*+N^{-1/2}}^{a} e^{Nf(x,N)}dx. \qquad (4.6)$$

Now we combine (4.5) and (4.6) and substitute it together with (4.4) into LHS of inequality given by this theorem and obtain

$$|I(N) - I_G(N)| = |I_{11}(N) + I_{12}(N) + I_2(N) - I_{G1}(N) - I_{G2}(N)|.$$

Then apply the triangle inequality four time on the RHS to separate the integrals except of $I_{11}(N)$ and $I_{G1}(N)$ and get

$$|I(N) - I_G(N)| \leq |I_{11}(N) - I_{G1}(N)| + |I_{12}(N)| + |I_2(N)| + |I_{G2}(N)|.$$

Each of the four terms we calculate separately

1) $|I_{11}(N) - I_{G1}(N)|$

Let us evaluate Taylor's theorem for $f(x,N)$ at x^* with $n=2$

$$f(x,N) = f(x^*,N) + f'(x^*,N)(x-x^*) + \frac{1}{2}f''(x_\theta,N)(x-x^*)^2,$$

where x_θ is a point between x and x^*. Then we substitute above expansion formula into $I_{11}(N)$ and evaluate expression for $I_{G1}(N)$

$$\begin{aligned}
&|I_{11}(N) - I_{G1}(N)| = \\
&= \bigg| g(x^*) \int_{U_N} e^{Nf(x^*,N)+Nf'(x^*,N)(x-x^*)+\frac{N}{2}f''(x_\theta,N)(x-x^*)^2} dx - \\
&\quad - g(x^*) \int_{U_N} e^{Nf(x^*,N)+Nf'(x^*,N)(x-x^*)} dx \bigg|,
\end{aligned}$$

and combine those two integrals

$$= \left| g(x^*) \int_{U_N} e^{Nf(x^*,N)+Nf'(x^*,N)(x-x^*)} \left(e^{\frac{N}{2}f''(x_\theta,N)(x-x^*(N))^2} dx - 1 \right) \right|.$$

We apply inequality from the Lemma 8 from the Appendix A.1 for $k=1$, i.e. $|e^t - 1| \leq |t| e^{|t|}$

$$\leq \left| g(x^*) \int_{U_N} e^{Nf(x^*,N)+Nf'(x^*,N)(x-x^*)} \frac{N}{2} f''(x_\theta,N)(x-x^*(N))^2 e^{\frac{N}{2}f''(x_\theta,N)(x-x^*(N))^2} dx \right|.$$

Since integration is over U_N, it is true that $|x-x^*| \leq \frac{1}{N^{1/2}}$ and $f'(x^*,N) < 0$, hence after appropriate substitution and basic manipulations we have

$$\begin{aligned}
|I_{11}(N) - I_{G1}(N)| &\leq \frac{N}{2} f''(x^*,N) e^{Nf(x^*,N)+\frac{1}{2}f''(x^*,N)} \times \\
&\times \left| g(x^*) \int_{U_N} |x-x^*|^2 e^{-N|f'(x^*,N)|(x-x^*)} dx \right|.
\end{aligned} \tag{4.7}$$

The integral above can be easily calculated using integration by parts

$$\begin{aligned}
\int_{U_N} |x-x^*|^2 e^{-N|f'(x^*,N)|(x-x^*)} dx &\leq \int_{x^*}^{\infty} |x-x^*|^2 e^{-N|f'(x^*,N)|(x-x^*)} dx \\
&\leq \frac{2}{(N|f'(x^*,N)|)^3}
\end{aligned}$$

Putting it together with all previous constants in (4.7) we finally get the approximation of the first term

$$|I_{11}(N) - I_{G1}(N)| \leq \frac{1}{N^2} \frac{f''(x^*, N)}{|f'(x^*, N)|^3} |g(x^*)| e^{Nf(x^*,N)+\frac{1}{2}f''(x^*,N)}. \qquad (4.8)$$

2) $|I_{12}(N)|$

Here we use 1-st order Taylor's expansion at $x^*(N)$

$$f(x, N) = f(x^*) + f'(x_\theta, N)(x - x^*),$$

where x_θ is some point between x and x^*. We insert it into $I_{12}(N)$ and with some basic manipulations get

$$|I_{12}(N)| = \left| g(x^*) e^{Nf(x^*,N)} \int_{x^*+N^{-1/2}}^{a} e^{-N|f'(x_\theta,N)|(x-x^*)} dx \right|,$$

and this is bounded by

$$\leq \left| g(x^*) e^{Nf(x^*,N)} \int_{x^*+N^{-1/2}}^{\infty} e^{-N|f'(x_\theta,N)|(x-x^*)} dx \right|,$$

which can be easily calculated

$$|I_{12}(N)| \leq \left| g(x^*) e^{Nf(x^*,N)}) \frac{1}{N|f'(x_\theta, N)|} e^{-N^{1/2}|f'(x_\theta,N)|} \right| \qquad (4.9)$$

3) $|I_2(N)|$

For this integral we again we substitute 1-nd order Taylor's expansion

$$|I_2(N)| = \left| \int_{x^*}^{a} g'(x_\theta)(x - x^*) e^{Nf(x^*,N)+Nf'(x_\theta,N)(x-x^*)} dx \right|,$$

and we perform some manipulations, bound it, then calculate the integral

$$|I_2(N)| \leq \left| g'(x_\theta) e^{Nf(x^*,N)} \int_{x^*}^{\infty} (x - x^*) e^{-N|f'(x_\theta,N)|(x-x^*)} dx \right| = \qquad (4.10)$$

$$= \left| g'(x_\theta) e^{Nf(x^*,N)} \frac{1}{(N|f'(x^*, N)|)^2} \right|.$$

4) $|I_{G2}(N)|$

The approximation procedure is the same as for $|I_{12}(N)|$ but instead of x_θ we have x^* in the first derivative,

$$|I_{12}(N)| \leq \left| g(x^*)e^{Nf(x^*,N)}) \frac{1}{N|f'(x^*,N)|} e^{-N^{1/2}|f'(x^*,N)|} \right|. \tag{4.11}$$

Now, we combine the approximation of four integrals (4.8), (4.9), (4.10) and (4.11).

$$\begin{aligned}
&|I_{11}(N) - I_{G1}(N)| + |I_{12}(N)| + |I_2(N)| + |I_{G2}(N)| \leq \\
&\leq \frac{1}{N^2} \frac{f''(x^*,N)}{|f'(x^*,N)|^3} |g(x^*)| e^{Nf(x^*,N) + \frac{1}{2}f''(x^*,N)} + \\
&+ \left| g(x^*)e^{Nf(x^*,N)}) \frac{1}{N|f'(x_\theta,N)|} e^{-N^{1/2}|f'(x_\theta,N)|} \right| + \\
&+ \left| g'(x_\theta)e^{Nf(x^*,N)} \frac{1}{(N|f'(x^*,N)|)^2} \right| + \left| g(x^*)e^{Nf(x^*,N)}) \frac{1}{N|f'(x^*,N)|} e^{-N^{1/2}|f'(x^*,N)|} \right|,
\end{aligned}$$

and we take out term $\frac{1}{N}|g(x^*(N))|e^{Nf(x^*,N)} \frac{1}{N|f'(x^*,N)|}$ out of the bracket

$$\begin{aligned}
&|I_{11}(N) - I_{G1}(N)| + |I_{12}(N)| + |I_2(N)| + |I_{G2}(N)| \leq \\
&\leq \frac{1}{N}|g(x^*(N))|e^{Nf(x^*,N)} \frac{1}{N|f'(x^*,N)|} \left| \frac{f''(x^*,N)}{|f'(x^*,N)|^2} e^{\frac{1}{2}f''(x^*,N)} + \frac{|f'(x^*,N)|}{|f'(x_\theta,N)|} N e^{-N^{1/2}|f'(x_\theta,N)|} + \right. \\
&\left. + \frac{|g'(x_\theta)|}{|g(x^*)f'(x^*,N)|} + N e^{-N^{1/2}|f'(x^*,N)|} \right|.
\end{aligned}$$

Since the function $f \in C^2$ in the domain of integration the derivatives are bounded for all N, the second and the last term of RHS are also bounded, hence we can fix $N = N_0$ to obtain the constant such that

$$\begin{aligned}
&\left| \frac{f''(x^*,N)}{|f'(x^*,N)|^2} e^{\frac{1}{2}f''(x^*,N)} + \frac{|f'(x^*,N)|}{|f'(x_\theta,N)|} N e^{-N^{1/2}|f'(x_\theta,N)|} + \frac{1}{|f'(x^*,N)|} + \right. \\
&\left. + N e^{-N^{1/2}|f'(x^*,N)|} \right| \leq K.
\end{aligned}$$

Therefore we have

$$|I(N) - I_G(N)| \leq \frac{K}{N}|g(x^*(N))|e^{Nf(x^*,N)} \frac{1}{N|f'(x^*,N)|},$$

which is our final result. □

4.2 m-dimensional function with the maximum in the interior of the domain

The approximation of the integral (4.1), i.e.

$$I(N) = \int_{\Omega} g(x) e^{Nf(x,N)} dx. \tag{4.12}$$

with $f \in C^3$ and $g \in C^1$, is given by the following theorem

Theorem 7. For the above integral there exists $K > 0$ such that for sufficiently large N we have approximation

$$\left| \int_{\Omega} g(x) e^{Nf(x,N)} dx - g(x^*(N)) e^{Nf(x^*(N),N)} \left(\frac{2\pi}{N}\right)^{\frac{m}{2}} \frac{1}{\sqrt{\det D^2 f(x^*(N),N)}} \right| \leq$$
$$\leq \frac{K}{\sqrt{N}} g(x^*(N)) e^{Nf(x^*(N),N)} \left(\frac{2\pi}{N}\right)^{\frac{m}{2}} \frac{1}{\sqrt{\det D^2 f(x^*(N),N)}}.$$

Proof. Since

$$\left(\frac{2\pi}{N}\right)^{\frac{m}{2}} \frac{1}{\sqrt{det(D^2 f(x^*(N),N))}} = \int_{\mathbb{R}^m} e^{\frac{N}{2}(x-x^*(N))^T D^2 f(x^*(N),N)(x-x^*(N))} dx,$$

we define

$$I_G(N) = g(x^*) e^{Nf(x^*(N),N)} \int_{\mathbb{R}^m} e^{\frac{N}{2}(x-x^*(N))^T D^2 f(x^*(N),N)(x-x^*(N))} dx. \tag{4.13}$$

Now, we introduce the set

$$U_N = \{x : |x - x^*(N)| \leq \frac{1}{N^{1/3}},\ N \geq N_0\}.$$

Next we decompose the integral (4.13) into two integrals, one over $U_N(x^*)$ and second over $\mathbb{R}^m \backslash U_N$

$$\begin{aligned} I_G(N) =& I_{G1}(N) + I_{G2}(N) = g(x^*(N)) e^{Nf(x^*(N),N)} \int_{U_N} e^{\frac{N}{2}(x-x^*(N))^T D^2 f(x^*(N),N)(x-x^*(N))} dx + \\ &+ g(x^*) e^{Nf(x^*(N),N)} \int_{\mathbb{R}^m \backslash U_N} e^{\frac{N}{2}(x-x^*(N))^T D^2 f(x^*(N),N)(x-x^*(N))} dx. \end{aligned} \tag{4.14}$$

Now let us use Taylor's Theorem to get 1-st order expansion of the function $g(x)$ at the point $x^*(N)$,

$$g(x) = g(x^*(N)) + Dg(x_\theta(N))^T(x - x^*(N)),$$

where $x_\theta(N)$ is some point between $x^*(N)$ and x, and can be formally represented $x_\theta(N) = x^*(N) + \theta(x - x^*(N))$, $0 \le \theta \le 1$.
Then we substitute it into $I(N)$ and separate integrals, one with $g(x^*(N))$ and second with the other term of expansion

$$\begin{aligned} I(N) =& I_1(N) + I_2(N) = g(x^*(N)) \int_\Omega e^{Nf(x,N)} dx + \\ &+ \int_\Omega Dg(x_\theta(N))^T (x - x^*(N)) e^{Nf(x,N)} dx. \end{aligned} \tag{4.15}$$

Next we decompose $I_1(N)$ into two, one over U_N and second over $\mathbb{R}\backslash U_N$

$$I_1(N) = I_{11}(N) + I_{12}(N) = g(x^*(N)) \int_{U_N} e^{Nf(x,N)} dx + g(x^*(N)) \int_{\Omega\backslash U_N} e^{Nf(x,N)} dx. \tag{4.16}$$

Now we combine (4.15) and (4.16) and substitute it together with (4.14) into LHS of inequality given by this theorem and obtain

$$|I(N) - I_G(N)| = |I_{11}(N) + I_{12}(N) + I_2(N) - I_{G1}(N) - I_{G2}(N)|.$$

Then apply the triangle inequality four time on the RHS to separate the integrals except of $I_{11}(N)$ and $I_{G1}(N)$ and get

$$|I(N) - I_G(N)| \le |I_{11}(N) - I_{G1}(N)| + |I_{12}(N)| + |I_2(N)| + |I_{G2}(N)|.$$

Each of the four terms we calculate separately

1) $|I_{11}(N) - I_{G1}(N)|$
   Let us evaluate Taylor's theorem for $f(x, N)$ at $x^*(N)$ with $n = 3$

$$\begin{aligned} f(x,N) =& f(x^*(N), N) + Df(x^*(N), N)^T (x - x^*(N)) + \\ &+ \frac{1}{2}(x - x^*(N))^T D^2 f(x^*(N), N)(x - x^*(N)) + F^{(3)}_{f(x_\theta(N))}(x - x^*(N))^{\otimes 3}, \end{aligned}$$

   where $x_\theta(N)$ is a point between x and $x^*(N)$ and for the definition of $F^{(3)}_{f(x_\theta(N))}$

see beginning of the Appendix on Analysis A.1. What more, as $x^*(N)$ is unique maximum, $Df(x^*(N),N)(x-x^*(N))=0$ for all x.
Then we substitute above expansion formula into $I_{11}(N)$ and evaluate expression for $I_{G1}(N)$

$$
\begin{aligned}
&|I_{11}(N)-I_{G1}(N)| = \\
&=\left|g(x^*(N))\int_{U_N} e^{Nf(x^*(N),N)+\frac{N}{2}(x-x^*(N))^T D^2 f(x^*(N),N)(x-x^*(N))+F^{(3)}_{Nf(x_\theta(N))}(x-x^*(N))^{\otimes 3}}dx-\right.\\
&\quad\left. -\,g(x^*(N))\int_{U_N} e^{Nf(x^*(N),N)+\frac{N}{2}(x-x^*(N))^T D^2 f(x^*(N),N)(x-x^*(N))}dx\right|,
\end{aligned}
$$

and combine these two integrals

$$
=\left|g(x^*(N))\int_{U_N} e^{Nf(x^*(N),N)+\frac{N}{2}(x-x^*(N))^T D^2 f(x^*(N),N)(x-x^*(N))}\left(e^{NF^{(3)}_{f(x_\theta(N))}(x-x^*(N))^{\otimes 3}}dx-1\right)\right|.
$$

Next we apply inequality from the Lemma 8 from the Appendix A.1 for $k=1$, i.e. $|e^t-1|\le |t|e^{|t|}$

$$
\begin{aligned}
\le&\left|g(x^*(N))\int_{U_N} e^{Nf(x^*(N),N)+\frac{N}{2}(x-x^*(N))^T D^2 f(x^*(N),N)(x-x^*(N))}\times\right.\\
&\left.\times\, NF^{(3)}_{f(x_\theta(N))}(x-x^*(N))^{\otimes 3}e^{NF^{(3)}_{f(x_\theta(N))}(x-x^*(N))^{\otimes 3}}dx\right|.
\end{aligned}
$$

Note that $F^{(3)}_{f(x_\theta(N))}(x-x^*(N))\le F^{(3)}(f(x^*(N)))|x-x^*(N)|^3$, for details see A.1 and since integration is over U_N it is true that $|x-x^*(N)|\le\frac{1}{N^{1/3}}$, hence after appropriate substitution and basic manipulations we have

$$
\begin{aligned}
|I_{11}(N)-I_{G1}(N)|\le NF^{(3)}(f(x_\theta(N)))e^{Nf(x^*(N),N)+F^{(3)}(f(x_\theta(N)))}\times \\
\times\left|g(x^*(N))\int_{U_N}|x-x^*(N)|^3 e^{\frac{N}{2}(x-x^*(N))^T D^2 f(x^*(N),N)(x-x^*(N))}dx\right|.
\end{aligned}
\tag{4.17}
$$

The integral above can be calculated applying Lemma 9 from the Appendix A.1

$$
\begin{aligned}
&\int_{U_N}|x-x^*(N)|^3 e^{\frac{N}{2}(x-x^*(N))^T D^2 f(x^*(N),N)(x-x^*(N))}dx\le\\
&\qquad\le\frac{||A_1^{-1}||^3}{\sqrt{|\det D^2 f(x^*(N),N)|}}\pi^{m/2}\left(\frac{N}{2}\right)^{-\frac{m+3}{2}}\frac{\Gamma(\frac{m+3}{2})}{\Gamma(\frac{m}{2})},
\end{aligned}
$$

where A_1 is such that $A_1^T A_1 = D^2 f(x^*(N), N)$ and $||A_1||^{-1}$ is norm of corresponding inverse transformation. Putting it together with all previous constants in (4.17) we finally get the approximation of the first term

$$|I_{11}(N) - I_{G1}(N)| \leq \frac{1}{\sqrt{N}} N^{-\frac{m}{2}} F^{(3)}(f(x_\theta(N))) \frac{||A_1^{-1}||^3}{\sqrt{|\det D^2 f(x^*(N), N)|}} \times \tag{4.18}$$

$$\times\, 2^{\frac{m+3}{2}} |g(x^*(N))| e^{N f(x^*(N),N) + F^{(3)}(f(x_\theta(N)))} \pi^{m/2} \frac{\Gamma(\frac{m+3}{2})}{\Gamma(\frac{m}{2})}.$$

2) $|I_{12}(N)|$
Here we use 2-nd order Taylor's expansion at $x^*(N)$

$$f(x, N) = f(x^*(N)) + Df(x^*(N), N)^T (x - x^*(N)) + \frac{1}{2}(x - x^*(N))^T D^2 f(x_\theta(N), N)(x - x^*(N)),$$

where $x_\theta(N)$ is some point between x and $x^*(N)$, and $Df(x^*(N), N) = 0$. We insert it into $I_{12}(N)$ and with some basic manipulations get

$$|I_{12}(N)| = \left| g(x^*(N)) e^{N f(x^*(N),N)} \int_{\mathbb{R}\setminus U_N} e^{\frac{N}{2}(x - x^*(N))^T D^2 f(x_\theta(N),N)(x - x^*(N))} dx \right|,$$

which can be calculated using Lemma 10 from the Appendix A.1, where the radius of sphere of the integration is $R = \frac{1}{N^{1/3}}$, $\alpha = N$, $k = 0$ and $A = D^2 f(x_\theta(N), N)$

$$\begin{aligned} |I_{12}(N)| \leq & \left| g(x^*(N)) e^{N f(x^*(N),N)}) \frac{||A_2^{-1}||^3}{\sqrt{|\det D^2 f(x_\theta(N), N)|}} \times \right. \\ & \left. \times\, e^{-\frac{N^{2/3}}{2}} \sum_{j=0}^{\frac{m}{2}-1} \binom{\frac{m}{2}-1}{j} N^{-\frac{2j}{3}} \pi^{\frac{m}{2}} \left(\frac{N}{2}\right)^{-\frac{m}{2}-1+j} \frac{\Gamma(\frac{m}{2} - j)}{\Gamma(\frac{m}{2})} \right|. \end{aligned}$$

where $A_2^T A_2 = D^2 f(x^*(N), N)$ For the simplicity of the bound we can take the term with highest power of N, i.e. the last term in the sum

$$\begin{aligned} |I_{12}(N)| \leq & \left| g(x^*(N)) e^{N f(x^*(N),N)}) \frac{||A_2^{-1}||^3}{\sqrt{|\det D^2 f(x_\theta(N), N)|}} \times \right. \\ & \left. \times\, e^{-\frac{N^{2/3}}{2}} m/2(m/2)! N^{-\frac{m-2}{3}} \pi^{\frac{m}{2}} \frac{\Gamma(1)}{\Gamma(\frac{m}{2})} \right|, \end{aligned}$$

and after some manipulations

$$|I_{12}(N)| \leq \Big| e^{-\frac{N^{2/3}}{2}} g(x^*(N)) e^{Nf(x^*(N),N)}) \times \tag{4.19}$$
$$\times \frac{||A_2^{-1}||^3}{\sqrt{|\det D^2 f(x_\theta(N),N)|}} m/2(m/2)! N^{-\frac{m-2}{3}} \pi^{\frac{m}{2}} \frac{\Gamma(1)}{\Gamma(\frac{m}{2})} \Big|.$$

3) $|I_2(N)|$

Again we substitute 2-nd order Taylor's expansion

$$|I_2(N)| = \left| \int_\Omega Dg(x^*(\theta))^T (x - x^*(\theta)) e^{Nf(x^*(N),N) + \frac{N}{2}(x-x^*(N))^T D^2 f(x_\theta(N),N)(x-x^*(N))} dx \right|.$$

Then we use Lemma 11 from A.1 to calculate the integral explicitly

$$|I_2(N)| \leq \frac{1}{\sqrt{N}} e^{Nf(x^*(N),N)} N^{-\frac{m}{2}} 2(2\pi)^{\frac{m-1}{2}} \left| \frac{Dg(x^*(N))^T D^2 f(x_\theta(N),N) Dg(x^*(N))}{\det(D^2 f(x_\theta(N),N))} \right|^{\frac{1}{2}}. \tag{4.20}$$

4) $|I_{G2}(N)|$

The approximation procedure is the same as for $|I_{12}(N)|$ but instead of $x_\theta(N)$ we have $x^*(N)$ in the determinant,

$$|I_{12}(N)| \leq \Big| e^{-\frac{N^{2/3}}{2}} g(x^*(N)) e^{Nf(x^*(N),N)}) \times \tag{4.21}$$
$$\times \frac{||A_2^{-1}||^3}{\sqrt{|\det D^2 f(x^*(N),N)|}} m/2(m/2)! N^{-\frac{m-2}{3}} \pi^{\frac{m}{2}} \frac{\Gamma(1)}{\Gamma(\frac{m}{2})} \Big|.$$

where $A_1^T A_1 = D^2 f(x^*(N), N)$.

Now combine the approximation of four integrals (4.18), (4.19), (4.20) and (4.21)

$$
\begin{aligned}
&|I_{11}(N) - I_{G1}(N)| + |I_{12}(N)| + |I_2(N)| + |I_{G1}(N)| \leq \\
&\leq \Bigg| \frac{1}{\sqrt{N}} N^{-\frac{m}{2}} F^{(3)}(f(x_\theta(N))) \frac{||A_1^{-1}||^3}{\sqrt{|\det D^2 f(x^*(N), N)|}} \times \\
&\times 2^{\frac{m+3}{2}} |g(x^*(N))| e^{Nf(x^*(N),N) + F^{(3)}(f(x_\theta(N)))} \pi^{m/2} \frac{\Gamma(\frac{m+3}{2})}{\Gamma(\frac{m}{2})} \Bigg| + \\
&+ \Bigg| e^{-\frac{N^{2/3}}{2}} g(x^*(N)) e^{Nf(x^*(N),N)}) \frac{||A_2^{-1}||^3}{\sqrt{|\det D^2 f(x_\theta(N), N)|}} m/2(m/2)! N^{-\frac{m-2}{3}} \pi^{\frac{m}{2}} \frac{\Gamma(1)}{\Gamma(\frac{m}{2})} \Bigg| + \\
&+ \Bigg| \frac{1}{\sqrt{N}} e^{Nf(x^*(N),N)} N^{-\frac{m}{2}} 2(2\pi)^{\frac{m-1}{2}} \left| \frac{Dg(x^*(N))^T D^2 f(x_\theta(N), N) Dg(x^*(N))}{\det(D^2 f(x_\theta(N), N))} \right|^{\frac{1}{2}} \Bigg| + \\
&+ \Bigg| e^{-\frac{N^{2/3}}{2}} g(x^*(N)) e^{Nf(x^*(N),N)}) \frac{||A_1^{-1}||^3}{\sqrt{|\det D^2 f(x^*(N), N)|}} m/2(m/2)! N^{-\frac{m-2}{3}} \pi^{\frac{m}{2}} \frac{\Gamma(1)}{\Gamma(\frac{m}{2})} \Bigg|,
\end{aligned}
$$

and we take out term $\frac{1}{\sqrt{N}} |g(x^*(N))| e^{Nf(x^*(N),N)} (\frac{2\pi}{N})^{\frac{m}{2}} |\det(D^2 f(x^*(N), N)|^{-1/2}$ out of the bracket and combine second with last expression

$$
\begin{aligned}
&|I_{11}(N) - I_{G1}(N)| + |I_{12}(N)| + |I_2(N)| + |I_{G1}(N)| \leq \\
&\leq \frac{1}{\sqrt{N}} |g(x^*(N))| e^{Nf(x^*(N),N)} \left(\frac{2\pi}{N} \right)^{\frac{m}{2}} \frac{1}{\sqrt{|\det(D^2 f(x^*(N), N))|}} \times \\
&\times \Bigg[\Bigg| F^{(3)}(f(x_\theta(N))) ||A_1^{-1}||^3 2^{\frac{m+3}{2}} e^{F^{(3)}(f(x_\theta(N)))} \pi^{m/2} \frac{\Gamma(\frac{m+3}{2})}{\Gamma(\frac{m}{2})} \Bigg| + \\
&+ \Bigg| \frac{\sqrt{|\det(D^2 f(x^*(N), N))|}}{|g(x^*(N))|} 2(2\pi)^{\frac{m-1}{2}} \left| \frac{Dg(x^*(N))^T D^2 f(x_\theta(N), N) Dg(x^*(N))}{\det(D^2 f(x_\theta(N), N))} \right|^{\frac{1}{2}} \Bigg| + \\
&+ \Bigg| e^{-\frac{N^{2/3}}{2}} ||A_1^{-1}||^3 m/2(m/2)! N^{-\frac{m-2}{3} + \frac{m}{2}} \pi^{\frac{m}{2}} \frac{\Gamma(1)}{\Gamma(\frac{m}{2})} \left(1 + \frac{||A_2^{-1}||^3 \sqrt{|\det D^2 f(x^*(N), N)|}}{||A_1^{-1}||^3 \sqrt{|\det D^2 f(x_\theta(N), N)|}} \right) \Bigg| \Bigg].
\end{aligned}
$$

Since the derivatives up to the third order are continuous, hence there are bounded and the last term in the bracket is bounded for fixed N and we can obtain the constant such that

$$\left|F^{(3)}(f(x_\theta(N)))||A_1^{-1}||^3 2^{\frac{m+3}{2}} e^{F^{(3)}(f(x_\theta(N)))}\pi^{m/2}\frac{\Gamma(\frac{m+3}{2})}{\Gamma(\frac{m}{2})}\right|+$$
$$+\left|\frac{\sqrt{|\det(D^2f(x^*(N),N))|}}{|g(x^*(N))|}2(2\pi)^{\frac{m-1}{2}}\left|\frac{Dg(x^*(N))^T D^2 f(x_\theta(N),N)Dg(x^*(N))}{\det(D^2f(x_\theta(N),N))}\right|^{\frac{1}{2}}\right|+$$
$$+\left|e^{-\frac{N^{2/3}}{2}}||A_1^{-1}||^3 m/2(m/2)!N^{-\frac{m-2}{3}+\frac{m}{2}}\pi^{\frac{m}{2}}\frac{\Gamma(1)}{\Gamma(\frac{m}{2})}\left(1+\frac{\sqrt{|\det D^2f(x^*(N),N)|}}{\sqrt{|\det D^2 f(x_\theta(N),N)|}}\right)\right|\leq K.$$

Therefore we have

$$|I(N)-I_G(N)|\leq\frac{K}{\sqrt{N}}|g(x^*(N))|e^{Nf(x^*(N),N)}\left(\frac{2\pi}{N}\right)^{\frac{m}{2}}\frac{1}{\sqrt{\det D^2f(x^*(N),N)}}.$$

Hence the theorem is proved. □

4.3 m-dimensional function with the maximum on the boundary of the domain

The approximation of the integral (4.1), i.e.

$$I(N)=\int_\Omega g(x)e^{Nf(x,N)}dx. \tag{4.22}$$

with $f\in C^4$ and $g\in C^1$, is given by the following theorem

Theorem 8. For the above integral there exists $K>0$ such that for sufficiently large N we have approximation

$$\left|\int_\Omega g(x)e^{Nf(x,N)}dx - g(x^*(N))e^{Nf(x^*(N),N)}\frac{1}{N}\left(\frac{2\pi}{N}\right)^{\frac{m-1}{2}}\frac{1}{|f'(x^*(N),N)|\sqrt{|\det D^2f(x^*(N),N)|}}\right|\leq$$
$$\leq\frac{K}{\sqrt{N}}|g(x^*(N))|e^{Nf(x^*(N),N)}\frac{1}{N}\left(\frac{2\pi}{N}\right)^{\frac{m-1}{2}}\frac{1}{|f'(x^*(N),N)|\sqrt{|\det D^2f(x^*(N),N)|}},$$

where the curve of the maximum is along the x_1 axis and

$$f'(x^*(N),N)=\frac{\partial}{\partial x_1}f(x,N)|_{x=x^*(N)}$$

and the matrix $D^2f(x^*(N),N)$ is defined

$$[D^2f(x^*(N),N)]_{i,j} = \frac{\partial^2}{\partial x_i \partial x_j} f(x,N)|_{x=x^*(N)}, \text{ for, } i,j = 2,\dots,m$$

In the case when curve of the maximum is in direction other than x_1 we perform rotation of coordinate system first. Further, if the maximum is along other variable then x_1 we simply swap the variables in the derivatives accordingly.

Proof. We start with introducing a new variables for the coordinate system in which the domain Ω is contained. The vector $x = (y, z_1, z_2, \dots, z_{m-1}) = (y, z)$. Since the maximum is along x_1 at the point of maximum of y, i.e. y^*, we define a curve of maximal values along y axis, i.e. $z = z^*(y,N)$. The domain corresponding to one cross-section with some $z^*(y,N)$ will be denoted by $\Omega(y)$. The corresponding domain along variable y will be denoted by the interval (y^*, Ω_y). Then the integral $I(N)$ can be decomposed in the following way

$$I(N) = \int_{y^*}^{\Omega_y} I(y,N)dy,$$

where

$$I(y,N) = \int_{\Omega(y)} g(y,z)e^{Nf((y,z),N)}dz. \tag{4.23}$$

Next, we apply Theorem 7 from the previous section to the integral $I(y,N)$

$$\begin{aligned}\Bigg| \int_{\Omega(y)} & g(y,z)e^{Nf((y,z),N)}dz - \\ & - g(y,z^*(y,N))e^{Nf((y,z^*(y,N)),N)}\left(\frac{2\pi}{N}\right)^{\frac{m}{2}} \frac{1}{\sqrt{\det D^2 f((y,z^*(y,N)),N)}} \Bigg| \leq \\ & \leq \frac{K(y)}{\sqrt{N}} g(y,z^*(y,N))e^{Nf((y,z^*(y,N)),N)}\left(\frac{2\pi}{N}\right)^{\frac{m}{2}} \frac{1}{\sqrt{\det D^2 f((y,z^*(y,N)),N)}}.\end{aligned}$$

where $K(y) > 0$ is a constant depending on y and inequality is valid for sufficiently large N.

Now we integrate over the variable y to obtain $I(N)$

$$\left| I(N) - \int_{y^*}^{\Omega_y} g(y, z^*(y,N)) e^{Nf((y,z^*(y,N)),N)} \left(\frac{2\pi}{N}\right)^{\frac{m}{2}} \frac{1}{\sqrt{\det D^2 f((y,z^*(y,N)),N)}} dy \right| \leq$$
$$\leq \int_{y^*}^{\Omega_y} \frac{K(y)}{\sqrt{N}} g(y, z^*(y,N)) e^{Nf((y,z^*(y,N)),N)} \left(\frac{2\pi}{N}\right)^{\frac{m}{2}} \frac{1}{\sqrt{\det D^2 f((y,z^*(y,N)),N)}} dy. \tag{4.24}$$

Then we apply Theorem 6 from the Section 1 for the integral in the LHS

$$\Bigg| \int_{y^*}^{\Omega_y} g(y, z^*(y,N)) e^{Nf(y,z^*(y,N),N)} \left(\frac{2\pi}{N}\right)^{\frac{m}{2}} \frac{1}{\sqrt{\det D^2 f((y,z^*(y,N)),N)}} dy -$$
$$- g(y^*, z^*(y^*,N)) e^{Nf((y^*,z^*(y^*,N)),N)} \frac{1}{N} \frac{1}{|f'((y^*,z^*(y^*,N)),N)|} \Bigg| \leq$$
$$\leq \frac{K}{N} |g(y^*, z^*(y^*,N))| e^{Nf((y^*,z^*(y^*,N)),N)} \frac{1}{N} \frac{1}{|f'((y^*,z^*(y^*,N)),N)|}, \tag{4.25}$$

and also for the integral on the RHS of (4.24). However, this time the function g in the Theorem 6 will be altered by the factor $K(y)$. Since from the Theorem 7 we have that $K = K(y)$, which is C^1 class function w.r.t. the variable y if $f \in C^4$ and $g \in C^1$, which is valid by the assumptions. Hence we can apply Theorem 6 for the considered integral

$$\Bigg| \int_{y^*}^{\Omega_y} \frac{K(y)}{\sqrt{N}} g(y, z^*(y,N)) e^{Nf((y,z^*(y,N)),N)} \left(\frac{2\pi}{N}\right)^{\frac{m}{2}} \frac{1}{\sqrt{\det D^2 f(y,z^*(y,N),N)}} dy -$$
$$- \frac{K(y^*)}{\sqrt{N}} g(y^*, z^*(y^*,N)) e^{Nf((y^*,z^*(y^*,N)),N)} \frac{1}{N} \frac{1}{|f'((y^*,z^*(y^*,N)),N)|} \Bigg| \leq$$
$$\leq \frac{K}{N} \frac{K(y^*)}{\sqrt{N}} |g(y^*, z^*(y^*,N))| e^{Nf((y^*,z^*(y^*,N)),N)} \frac{1}{N} \frac{1}{|f'((y^*,z^*(y^*,N)),N)|}. \tag{4.26}$$

Now, inserting inequalities (4.25) and (4.26) into (4.24) we obtain

$$\left| I(N) - g(y^*, z^*(y^*,N)) e^{Nf((y^*,z^*(y^*,N)),N)} \frac{1}{N} \frac{1}{|f'((y^*,z^*(y^*,N)),N)|} \right| \leq$$
$$\leq \left(\frac{K}{N} + \frac{K(y^*)}{\sqrt{N}} + \frac{K}{N} \frac{K(y)}{\sqrt{N}} \right) |g(y^*, z^*(y^*,N))| e^{Nf((y^*,z^*(y^*,N)),N)} \frac{1}{N} \frac{1}{|f'((y^*,z^*(y^*,N)),N)|}.$$

Hence, setting

$$\left(K(y^*) + \frac{K(y^*)}{\sqrt{N}} + \frac{K(y^*)K}{N}\right) \leq K',$$

for some fixed N and since $x^*(N) = (y^*, z^*(y^*, N))$, we get the result of the Theorem. $\square$

5 Conclusions, possible applications and future research

This last chapter is devoted to the conclusions regarding the content of Chapters 2 to 4. It emphasizes the results which are relevant to the corresponding fields and which are publishable. It also provides description of some relevant application of the thesis results and possible future research which can be conducted as a continuation of the work done.

The first section of this chapter presents the conclusions about the result of thesis, its relevancy, describes the contributions to the fields and the publishability of this results. The theorems contained in the Chapter 2 are mostly the subjects of the discussion.

The next section contains some possible application of our work. We put it into the context where the obtained theorems could have a valid contribution and extend the understanding of considered well-known phenomenas.

The last, third section, consist of directions, subjects and ideas which emerged during our work. We discuss a few topics which are relevant and where viable research could be done. This could be an extension of our work, relaxing the assumptions of underlying mathematical setting or independent result based on existing fundaments.

5.1 Conclusions

In this section we discuss the contribution of our work to the related fields of science. We describe the progress of developing the ideas in relation to already existing results introduced in the first Chapter.

The first subsection is a description of the obtained results with respect to work already done by V.P. Maslov. We put it in the context of Maslov's results introduced in the first Chapter and discuss the progress made towards comparing the developed theorems.

In the second subsection we present the contribution to Statistical Physics, the new ways of looking at previous, classical results. We underline new connections and new rigorous results and their publishability.

The third subsection consists of the presentation of the contribution to Complexity Science. We emphasize the significance of the developed theorems.

In the last subsection we discuss the connection of two fields, previously unrelated in this respect and on this level of the mathematical rigour.

5.1.1 Continuation of the work of Prof. V.P. Maslov

The inspiration for our work came from the ideas and concepts first introduced by Prof. V.P. Maslov. Initially we were mostly interested in the economical aspect of his work, contained in his two papers [9],[10]. However, we soon became aware that the proofs of relevant theorems are lacking full mathematical rigour and the connection to economics was vague. Those two papers are revised in the first two paragraphs of the first section of the Introduction Chapter.

Next, we found his result [11], which is constructed in more general framework and independently of the context of application. However the proof was lacking in rigour and the underlying mathematical fundaments were unclear. We found this result a good starting point for a PhD topic. It is reviewed in the third paragraph of the section on the work of Prof. V.P. Maslov. The main result of the thesis, Theorem 3 form the Section 2 of Chapter II is an extension and more precise version of the Prof. V.P. Maslovs theorem. The proof for our theorem was constructed with full rigour, from the beginning, independent of an already existing one. Further, the underlying mathematical setting on which the theorem is based is step by step, introduced in Section 1 of Chapter II. The theorem also precisely distinguishes two cases of solution of which was only mentioned in the original paper and the precise rate of convergence to the limit is also included. Further, based on this result we constructed two additional fluctuation theorems with proofs of similar structure. Those results are currently in the process of preparation for publication.

The boost for our work was a fact that one of outcomes of the V.P. Maslov theorem was a Zipf-Mandelbrot law. Its significance and the wide range of applications was emphasized in his paper [12].

5.1.2 Contribution to Statistical Physics

Our work brought a significant contribution to the field of Statistical Physics. Here we describe in details exactly which parts of the thesis are contributing.

Rigorous version of method of most probable values

In the literature on Statistical Physics, for example [5] and [17] the Bose-Einstein and Maxwell-Boltzmann statistics are classical results and there are two methods of deriving them. They are called 'method of averages' and 'method of the most probable values' in [17]. We introduce those methods in Section 2.2 of the Introduction. However, those methods usually lack full mathematical rigour, especially with estimates and the lack of speed of convergence of considered system to the limiting statistics. In our work we deliver a rigorous result built on the framework of the 'method of most probable values'. The precise approximations for the Entropy, Partition Function in Chapter III and the Laplace approximation in Chapter IV are constructed. Furthermore, the speed of convergence is also included in the final theorems. Speed of convergence is important as it allows for calculating the estimation error for the system of specific size and with some particular value of parameters, as in the nature thermodynamical limit is only an abstract simplification.

Two types of maximum for a standard problem

Our result distinguishes two types of solutions. When the maximum of the Entropy is in the critical point, usually inside the domain or on the boundary of the domain, those solutions are completely different and this fact was omitted in the reviewed literature [5],[17]. Only one of them represents the classical statistics of Maxwell-Boltzmann or Bose-Einstein. The optimization problem from which those two solutions emerge is contained in Section 2 of Chapter III.

Unification of Quantum Statistical Physics result with Classical one

The Bose-Einstein and Maxwell-Boltzman statistics, classically are derived from two different type of physical systems. The relation which indicates which statistic will be obtained is described in the Section 2.1 in Chapter I. Due to assumptions introduced by V.P. Maslov those two statistics are put under one framework. Statistic which we will obtain is controlled by one parameter of the system. Namely, the way how the degenerations of particular energy level increases as the number of particles of that level increase. If the number of degenerations increases faster than the number of particles, we get the Maxwell-Boltzmann, if the rate of increases is the same then Bose-Einstein emerges. The existence of such a framework is very interesting, as it unifies quantum with classical result and opens a way for a new interpretation of level degeneration and of those statistics itself.

Rigorous Fluctuation theorems

Last contribution to the Statistical Physics are the fluctuation theorems for the statistics contained in the Chapter II. For each type of solution we have a different fluctuation of statistics from the average. Surprisingly, when the limiting distributions are given by the Bose-Einstein and Maxwell-Boltzmann statistics the fluctuations are a mixture of Exponential and Gaussian distributions. For the other case the distribution of fluctuation are simply Gaussian. For those new results we provide rigorous proofs with the corresponding estimates for the speed of convergence.

5.1.3 Contribution to Complexity Science

In the section 3 of Chapter I we provided a brief introduction on Zipf Law and Power Laws which are common tools in the Complexity Science. We emphasized its wide range application in whole spectrum of systems occurring in the real world. We also mentioned that Power Laws as a description of systems behaviour emerged purely on an experimental basis, meaning that by mere observation of the systems evolution one concludes that its behaviour is governed by Power Law. In our work following V.P. Maslov, we have obtained the Power Law, i.e. Zipf-Mandelbrot Law in the theoretical manner. It is a natural result out of a mathematically defined probabilistic system, rather than observation of the effect only. Therefore this achievement establishes a fundaments for theory that mathematically formulate, describe and explains systems that manifests Power Law behaviour. Such a progress in explanation of the underlying mathematics for Complex Systems is a significant contribution to the discipline.

5.1.4 Interdisciplinary contribution - unification of Statistical Physics and Complexity Science

The Statistical Physics and Complexity Science deal with systems existing on a various scale and, what more, of a different nature. For Statistical Physics we consider systems rather than simple structures, solids liquids, gases etc. where the interactions are quite simple and evolution can be well described by mechanics. For Complexity, the systems are of a more complex nature, although they can be described quite well, the description of the evolution, interactions without crude approximations are basically impossible. The theorem we developed, Theorem 3 Chapter II, puts these two classes of systems under one framework, where one parameter determines what type of the system we deal with. Such a universal result is a valuable interdisciplinary contribution.

5.2 Possible application - Maxwell-Boltzmann, Bose-Einstein statistics and Zipf Law as a description of state of economy

Methods of Statistical Physics are widely used for the systems of other than physical nature. One of the fields where it become a major tool for analysis and prediction is Econophysics. Already in the past century scientists noticed that the market movements, such as distribution of money or debt in economy are somehow manifests similar behaviour as a thermodynamical system. The concepts of particles, energy levels, entropy and other have a corresponding analogy for the economy.

The paper of F.V. Kusmartsev [16] presents well the interplay between Statistical Physics and Economy. The market trading agents, their money, debt and wealth are put in the context of thermodynamics of grand canonical ensemble. The coefficients such as temperature, chemical potential, entropy and activity are also introduced in the context of economy, together with all the thermodynamical laws relating to them. The model is fitted into the real data of US economy between 1998 to 2008 and several conclusions are infered. First of all, the distribution of money across the indistinguishable trading agents has a Bose-Einstein distribution. The economy crisis has a reflection in the parameters of the system. The temperature and activity parameters are peaks in the time of the crisis. Furthermore, for particular range of the parameters values, such that classical limit is valid we going to have a Maxwell-Boltzmann distribution for the money spread over the trading agents.

The fact that the state of the economy, particularly distribution of money across market participants, can have Maxwell-Boltzmann and Bose-Einstein distribution, together with our theoretical result, which combines those two distribution and Zipf-Mandlebrot Law under one framework, indicates there might be Zipf-Mandelbrot type state too. The corresponding parameter which distinguishes three obtained distributions, rate of increase of level degenerations most likely has some interpretation in the economical context. Further, the fact that analysis in [16] is based on the wealthiest economy might also have an effect on obtaining statistics. We might get other statistics for countries of average wealth or poor ones. It would be interesting to conduct such a research, analysis of the several economies to recognize what statistics manifests in the certain types of economies. We might get an indication of what factor are the most influential in the transition from one statistics to another and compare it with the theoretical model. Further, the transitions between statistics could be analyzed in detail and the triggering factors in the economical context could be isolated, hence maybe some reliable prediction possible.

5.3 Future research

This section contains the directions, concepts and ideas that might be conducted as a future research, extension of our work. Some of them are based on the relaxation of the assumption or extending the underlying probabilistic system. We put them in several paragraphs.

Obtaining other power laws In our main theorem, Chapter 2 Theorem 3 we obtained a Zipf-Mandelbrot law, which is a particular case of Power Law. It might be possible and it would be very useful to have obtained mathematically other Power Laws, with different power coefficients. This might be obtained by altering the initial entropy approximation in Section 1 Chapter III by possibly increasing the precision of approximation. Then the maximal point of the resulting function could have other form, this might lead to altering the Zipf-Mandelbrot Law and obtaining some other Power Law. Hence, the research on how the form, precision of approximated entropy affects the obtained distribution would have to be done.

Infinite dimension of the system In our assumptions we considered a fixed number of energy levels. This assumption is stated in Section 1 of Second Chapter. It would be theoretically interesting and, what is more, closer to real system that the number of energy levels is infinite. Such a modification most likely would cause a chain of changes in other assumptions of the system and definitely in the proof of the theorems. Probably, the increasing dimension would have to be entangled with the condition for the rate of increase of energy level degeneration to obtain three distinct cases. Regarding the proof, Laplace approximation of Chapter IV would have to be extended to infinite dimensional space, also sum approximation of Section 4 Chapter III and other related estimates of Section 5 of Chapter II would have to be modified. Hence, there would be many modifications but rather of the technical nature.

Dynamics and correlations in the system The system under consideration, given by the assumptions in Section 1 of Chapter II is a static system, constructed from number of independent random variables, where each variable represents number of particles on particular energy level. A natural extension of that system is to make it time dependent. This definitely would alter the theorems themselves and more complex methods for proving would be necessary. The distributions to which system converges would have to be found. A side of time dependency, the correlations between occurrence of particles on energy levels might be another natural extension of assumptions. Here the outcome of theorem might change and the probability density function with correlation would definitely be altered. For those extensions the tools of advanced probability would have to be extensively used.

Large deviations theorems In addition to the fluctuations we could develop a theorem which would describe system behaviour for very rare events. This corresponds to the Large deviation theorems. The systems assumption would stay the same, however the construction of the proofs might be altered. It may be necessary to develop new technical results.

A Mathematical preliminaries

A.1 Analysis

Theorem 9 (Taylor's)**.** Suppose that f is a real function on the nontrivial convex closed set $A \in \mathbb{R}^m$, n is a positive integer, $f(n-1)$ is continuous on A, $f^{(n)}(t)$ exists for every $t \in A$. Then there exists a point x_θ between x^* and x, such that

$$f(x) = f(x^*) + \sum_{k=1}^{n-1} \frac{1}{k!} \sum_{i_1, i_2, \ldots, i_k = 1}^{m} \frac{\partial^k f(x^*)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}} (x_{i_1} - x^*_{i_1})(x_{i_2} - x^*_{i_2}) \ldots (x_{i_k} - x^*_{i_k}) + \tag{A.1}$$

$$+ \frac{1}{n!} \sum_{i_1, i_2, \ldots, i_n = 1}^{m} \frac{\partial^m f(x_\theta)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_m}} (x_{i_1} - x^*_{i_1})(x_{i_2} - x^*_{i_2}) \ldots (x_{i_m} - x^*_{i_m}),$$

where x_θ can be represented, $x_\theta = x^* + \theta(x - x^*)$, $0 < \theta < 1$.

The $n - th$ term in the Taylor's theorem can be represented

$$\sum_{i_1, i_2, \ldots, i_n = 1}^{m} \frac{f^m(x_\theta)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_m}} (x_{i_1} - x^*_{i_1}) \ldots (x_{i_m} - x^*_{i_m}) =< D^n f(x_\theta), (x - x^*)^{\otimes m} >, \tag{A.2}$$

where $\otimes$ is tensor product, $D : f \to \nabla f$ differentiation operator and $D^m = D \otimes D \otimes \ldots \otimes D = D^{\otimes m}$.

Basic functional analysis result, Riesz representation theorem states that we can represent every inner product as a functional. Hence

$$< D^m f(x_\theta), (x - x^*)^{\otimes m} >= F^{(m)}_{x_\theta} (x - x^*)^{\otimes m}, \tag{A.3}$$

where $F^{(m)}_{x_\theta} : \mathbb{R}^{mn} \to \mathbb{R}$. Hence, for $m \geq 4$ expansion (A.1) can also be represented as

$$f(x) = f(x^*) + Df(x^*)^T(x - x^*) + (x - x^*)^T D^2 f(x^*)(x - x^*) + \\ + \sum_{k=3}^{m-1} \frac{1}{k!} F_{x^*}^{(k)} (x - x^*)^{\otimes m} + \frac{1}{m!} F_{x_\theta}^{(m)} (x - x^*)^{\otimes m},$$

where $0 < \theta < 1$. Then for every $x \in A$ we define a constant

$$F_A^{(m)}(x_\theta) = \sup_{x \in A} \frac{|F_{x_\theta}^{(m)}(x - x^*)^{\otimes m}|}{|(x - x^*)^{\otimes m}|}.$$

If the set A is whole set on which function $F_A^m(x_\theta)$ is defined then it simply becomes norm of the functional $F_{x_\theta}^{(m)}$ and we denote it as $F^{(m)}(x_\theta)$. The existence of such constant is ensured, since functional $F_{x^*}^{(n)}$ is finite dimensional, hence it is bounded, i.e

$$|F_{x_\theta}^{(m)}(y)| \leq c|y|, \tag{A.4}$$

for every $y \in \mathbb{R}^{mn}$ and some $c > 0$. In our case $c = F_A^{(m)}(x_\theta)$.
Further, by the definition of the tensor product $|x^{\otimes k}| = |x|^k$, putting together (A.2), (A.3) and (A.4) we get that m-th term in Taylor's Theorem is bounded by

$$\sum_{i_1, i_2, \ldots, i_m}^{m} \frac{f^{(m)}(x_\theta)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_m}} (x_{i_1} - x_{i_1}^*)(x_{i_2} - x_{i_2}^*) \ldots (x_{i_m} - x_{i_m}^*) \leq F^{(m)}(x_\theta)|x - x^*|^m.$$

We also define a constant

$$0 \leq F_A'^{(m)}(x_\theta) = \inf_{x \in A} \frac{|F_{x_\theta}^{(m)}(x - x^*)^{\otimes m}|}{|(x - x^*)^{\otimes m}|},$$

which is a lower bound for m-th term in the Taylor expansion

$$F'^{(m)}(x_\theta)|x - x^*|^m \leq \sum_{i_1, i_2, \ldots, i_m}^{m} \frac{f^{(m)}(x_\theta)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_m}} (x_{i_1} - x_{i_1}^*)(x_{i_2} - x_{i_2}^*) \ldots (x_{i_m} - x_{i_m}^*).$$

Lemma 8. For any $n \in \mathbb{N}$ and $t \in \mathbb{R}$ we have the following inequalitiy

$$\left| e^t - \sum_{k=0}^{m} \frac{t^k}{k!} \right| \leq \frac{|t|^{m+1}}{(m+1)!} e^{|t|}$$

Proof. Using the power series representation of the exponential function we get

$$e^t - \sum_{k=0}^{m} \frac{t^k}{k!} = \sum_{k=m+1}^{\infty} \frac{t^k}{k!}.$$

Then we change the summation index in of RHS to $k' = k - (m+1)$

$$e^t - \sum_{k=0}^{m} \frac{t^k}{k!} = \sum_{k'=0}^{\infty} \frac{t^{k'+(m+1)}}{(k+(m+1))!}.$$

Now we take the absolute value of both sides and apply triangle inequality on RHS and get

$$\left| e^t - \sum_{k=0}^{m} \frac{t^k}{k!} \right| \leq \sum_{k'=0}^{\infty} \left| \frac{t^{k'+(m+1)}}{k'!(m+1)} \right|$$

Since, $(k+(n+1))! \geq (k)!(n+1)!$ and using multiplicative properties of absolute value we obtain

$$\left| e^t - \sum_{k=0}^{m} \frac{t^k}{k!} \right| \leq \frac{t^{m+1}}{(m+1)!} \sum_{k'=0}^{\infty} \frac{|t|^{k'}}{k'!}.$$

where the sum is series expansion of exponent and it is the desired result. □

Lemma 9. For any $k \in \mathbb{N}$ and $\alpha > 0$, the integral

$$\int_{\mathbb{R}^m} |x|^k e^{-\alpha|x|^2} dx = \pi^{n/2} \alpha^{-\frac{n+k}{2}} \frac{\Gamma(\frac{n+k}{2})}{\Gamma(\frac{n}{2})}.$$

Further, given symmetric negative definite m-dimensional matrix A, for any set $\Omega \in \mathbb{R}^m$, which includes origin we have bound

$$\int_{\Omega} |x|^k e^{\alpha x^T A x} dx \leq \frac{||Q^{-1}||^k}{\sqrt{|\det(A)|}} \pi^{m/2} \alpha^{-\frac{m+k}{2}} \frac{\Gamma(\frac{m+k}{2})}{\Gamma(\frac{m}{2})}.$$

where Q is orthogonal matrix such that $Q^T Q = A$ and $||Q^{-1}||$ norm of corresponding inverse transformation.

Proof. For the first result the proof is a standard result. We change the coordinates system to spherical and use alternative representation of gamma function, integral representation. The second integral is simply obtain by bounding it by the integral over whole space and then change of the variable $y = Qx$ where $Q^T Q = A$ and then applying the first result. □

Lemma 10. For some $k \in \mathbb{N}$, constant $\alpha > 0$ and $(m-1)$-dimensional sphere with radius R denoted by $S_{m-1}(R)$, with center in the origin, the integral

$$\int_{\mathbb{R}^m \setminus S_{m-1}(R)} |x|^k e^{-\alpha|x|^2} dx = e^{-\alpha R^2} \sum_{j=0}^{\frac{m+k}{2}-1} \binom{\frac{m+k}{2}-1}{j} R^{2j} \pi^{\frac{m}{2}} \alpha^{-\frac{m-k}{2}-1+j} \frac{\Gamma\left(\frac{m+k}{2}-j\right)}{\Gamma(\frac{m}{2})}.$$

Given symmetric negative definite m-dimensional matrix A and the set $\Omega \in \mathbb{R}^m$ such that $\Omega \setminus S_{m-1}(R)$ has non-empty interior and $R < 1$, the we have a upper bound for the above integral

$$\int_{\Omega \setminus S_{m-1}(R)} |x|^k e^{\alpha x^T A x} dx \leq \frac{||Q^{-1}||^k}{\sqrt{|\det(A)|}} e^{-\alpha R^2} \sum_{j=0}^{\frac{m+k}{2}-1} \binom{\frac{m+k}{2}-1}{j} R^{2j} \pi^{\frac{m}{2}} \alpha^{-\frac{m-k}{2}-1+j} \frac{\Gamma\left(\frac{m+k}{2}-j\right)}{\Gamma(\frac{m}{2})},$$

where Q is orthogonal matrix such that $Q^T Q = A$ and $||Q^{-1}||$ norm of corresponding inverse transformation.

Proof. For the first result we change the coordinate system into spherical one where $x = r s_r$, and radius $r = |x|$

$$\int_{\mathbb{R}^m \setminus S_{m-1}(R)} |x|^k \exp\{-|x|^2\} dy = \int_R^\infty \int_{S_{n-1}(r)} r^k e^{-\alpha r^2} ds_r dr$$

As the function under the integral does not depend on the surface coordinates we can independently integrate over the surface. The surface of sphere in m-dimensional space of radius r is given by

$$S_{m-1}(r) = 2\frac{\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} r^{m-1},$$

and then

$$\int_R^\infty \int_{S_{m-1}(r)} r^k e^{-\alpha r^2} ds_r dr = 2\frac{\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} \int_R^\infty r^{k+m-1} e^{-\alpha r^2} dr. \tag{A.5}$$

Now we make substitution $t = \alpha(r^2 - R^2)$ and get

$$\int_R^\infty r^{k+m-1} e^{-\alpha r^2} dr = e^{-\alpha R^2} \int_0^\infty e^{-t} \left(\frac{t}{\alpha} + R^2\right)^{\frac{m+k}{2}-1} \frac{dt}{2\alpha}. \tag{A.6}$$

For integer values of $\frac{m-k}{2} - 1$ the expression in the bracket can be represented as the sum

$$(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^{n-j} b^j.$$

We apply it to (A.6) and then do some manipulations we get

$$\begin{aligned}
&e^{-\alpha R^2} \int_0^\infty e^{-t} \left(\frac{t}{\alpha} + R^2 \right)^{\frac{m+k}{2}-1} \frac{dt}{2\alpha} = \\
&= e^{-\alpha R^2} \sum_{j=0}^{\frac{m+k}{2}-1} R^{2j} \binom{\frac{m+k}{2}-1}{j} \int_0^\infty e^{-t} \left(\frac{t}{\alpha} \right)^{\frac{m+k}{2}-1-j} \frac{dt}{2\alpha} = \\
&= \frac{1}{2\alpha} e^{-\alpha R^2} \sum_{j=0}^{\frac{m+k}{2}-1} R^{2j} \binom{\frac{m+k}{2}-1}{j} \alpha^{-\frac{m+k}{2}+1+j} \int_0^\infty e^{-t} t^{\frac{m+k}{2}-1-j} dt.
\end{aligned}$$

Since the gamma integral representation of gamma function is

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

with $z > 0$ and $m + k \geq 2$ we have

$$\begin{aligned}
&\frac{1}{2\alpha} e^{-\alpha R^2} \sum_{j=0}^{\frac{m+k}{2}-1} R^{2j} \binom{\frac{m+k}{2}-1}{j} \alpha^{-\frac{m-k}{2}+1+j} \int_0^\infty e^{-t} t^{(\frac{m+k}{2}-1-j)} dt = \\
&= \frac{1}{2\alpha} e^{-\alpha R^2} \sum_{j=0}^{\frac{m+k}{2}-1} R^{2j} \binom{\frac{m+k}{2}-1}{j} \alpha^{-\frac{m-k}{2}+1+j} \Gamma\left(\frac{m+k}{2} - j \right) dt.
\end{aligned}$$

Then we put it all together with (A.5) and (A.6) then performing some manipulation yields

$$\begin{aligned}
&2 \frac{\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} \int_R^\infty r^{k+m-1} e^{-\alpha r^2} dr = \\
&= 2 \frac{\pi^{\frac{m}{2}}}{\Gamma(\frac{m}{2})} \frac{1}{2\alpha} e^{-\alpha R^2} \sum_{j=0}^{\frac{m+k}{2}-1} R^{2j} \binom{\frac{m+k}{2}-1}{j} \alpha^{-\frac{m+k}{2}+1+j} \Gamma\left(\frac{m+k}{2} - j \right) = \\
&= e^{-\alpha R^2} \sum_{j=0}^{\frac{m+k}{2}-1} R^{2j} \binom{\frac{m+k}{2}-1}{j} \pi^{\frac{m}{2}} \alpha^{-\frac{m+k}{2}+j} \frac{\Gamma\left(\frac{m+k}{2} - j \right)}{\Gamma(\frac{m}{2})},
\end{aligned}$$

which the first result.
The second integral is again, obtain by simply bounding it by the integral over whole space $\mathbb{R}^m$ and then change of the variable $y = Qx$ where $Q^TQ = A$ and then applying the first result. $\square$

Lemma 11. Given symmetric, negative definite, m-dimensional matrix A and vector c we have

$$\int_{\mathbb{R}^m} |c^T x| e^{x^T A x} dx = 2(2\pi)^{(n-1)/2} \left| \frac{c^T A c}{\det(A)} \right|^{\frac{1}{2}}.$$

Proof. See, [2] , p. 30. $\square$

Lemma 12. Let β be a positive parameter. Then for all constants $\epsilon, k > 0$ always

$$\int_{|x|>\beta\epsilon} |x| \exp(-k|x|^2) dx \to 0, \beta \to \infty$$

Proof. By using spherical coordinates we have that

$$\int_{|x|>\beta\epsilon} |x| \exp(-k|x|^2) dx = \frac{2\pi^{n/2}}{\Gamma(n/2)} \int_{\rho>\beta\epsilon} \rho^n \exp(-k\rho^2) d\rho \to 0 \text{ as } \beta \to \infty.$$

$\square$

A.2 Asymptotic theory, approximations and related results

Let $f : A \to \mathbb{R}$ be a continuous function and $A = (a, \infty)$ for some a.

Definition 1 (Big O)**.** The function f is of order O of the function $g : A \to \mathbb{R}$ as $x \to \infty$ if there exists is a constant $K > 0$ and $x_K \in A$ such that for all $x > x_K$

$$|f(x)| \leq K|g(x)|,$$

and we write it symbolically

$$f(x) = O(g(x)),\ x \to \infty.$$

Definition 2 (Small o)**.** The function f is of order o of the function $g : A \to \mathbb{R}$ as $x \to \infty$ if for all $K > 0$ there exists $x_K \in A$ such that for all $x > x_K$

$$|f(x)| \leq K|g(x)|,$$

and we write it symbolically

$$f(x) = o(g(x)), \ x \to \infty.$$

Definition 3 (Asymptotic equivalence)**.** The functions f and $g : A \to \mathbb{R}$ are asymptotically equivalent as $x \to \infty$ if for all $K > 0$ there exists $x_K \in A$ such that for all $x > x_K$, $f(x) \neq 0$ and $g(x) \neq 0$ and

$$\left| \frac{f(x)}{g(x)} - 1 \right| \leq K,$$

and we write it symbolically

$$f(x) \sim g(x), \ x \to \infty.$$

Definition 4 (Asymptotic expansion)**.** The formal power series $\sum_{n=0}^{\infty} a_n x^{-n}$ is an asymptotic power series expansion of f, as $x \to \infty$ if for all $m \in \mathbb{N}$

$$f(x) = \sum_{n=0}^{m} a_n x^{-n} + O(x^{-(m+1)}), \ x \to \infty, \tag{A.7}$$

and we write it symbolically

$$f(x) \sim \sum_{n=0}^{\infty} a_n x^{-n}, \ x \to \infty.$$

If first few coefficients of power series are known then we write

$$f(x) \sim a_0 + \frac{a_1}{x} + \frac{a_2}{x^2} + \dots, \ x \to \infty.$$

Furthermore (A.7) can equivalently be represented as

$$f(x) = \sum_{n=0}^{m} a_n x^{-n} + \sigma(x), \tag{A.8}$$
$$\sigma(x) = O(x^{-(m+1)}), x \to \infty.$$

Lemma 13 (Lower bound from second order expansion)**.** For the asymptotic power series expansion of f given by above definition with $m = 1$ and $a_0 \neq 0$, $a_1 \neq 0$ there exists $K' > 0$ for the sufficiently large x such that

$$K' \frac{1}{|x|} < |f(x) - a_0|.$$

Proof. By Definition 8

$$f(x) = \sum_{n=0}^{m} a_n x^{-n} + O(x^{-(m+1)}),\ x \to \infty.$$

for all $m \in \mathbb{N}$.
Next, we evaluate the definition of big O for $m = 1$. Hence for large enough x there exists K

$$\left| f(x) - a_0 - \frac{a_1}{x} \right| \leq K \left| \frac{1}{x^2} \right|. \tag{A.9}$$

By the the symmetry of absolute value and triangle inequality we have

$$|b| - |a| \leq |a - b|, \tag{A.10}$$

for some vectors a, b.
Due to (A.10) the LHS of (A.9) has the lower bound

$$\left| \frac{a_1}{x} \right| - \left| f(x) - a_0 \right| \leq \left| f(x) - a_0 - \frac{a_1}{x} \right|,$$

then we combine it with RHS in (A.9) and after some manipulations we get

$$\left(|a_1| - \frac{K}{|x|} \right) \frac{1}{|x|} \leq \left| f(x) - a_0 \right|. \tag{A.11}$$

Now, we assume there exists a constant $K' > 0$ such that

$$K' \frac{1}{|x|} \leq \left(|a_1| - \frac{K}{|x|}\right) \frac{1}{|x|},$$

hence it has to fulfill the conditions

$$K' > 0, \tag{A.12a}$$

$$K' \leq |a_1| - \frac{K}{|x|}, \tag{A.12b}$$

$$x > x_K. \tag{A.12c}$$

where x_K in the last condition is from (A.9) as x must sufficiently large.
To find K' explicitly we first invert equation $(A.12b)$ and merge it with $(A.12a)$ and after some manipulations we get

$$|x| > K/|a_1|,$$

which together with (A.12c) implies

$$x > \max\{K/|a_1|, x_K\}.$$

Then we use above inequality to bound RHS of (A.12b)

$$|a_1| - \frac{K}{|\max\{K/|a_1|, x_K\}|} < |a_1| - \frac{K}{|x|},$$

hence we can set

$$0 < K' \leq |a_1| - \frac{K}{|\max\{K/|a_1|, x_K\}|},$$

and write formally

$$\exists_{K'>0} \exists_{x_{K'}} \forall_{x > x_{K'}}\ K' < |a_1| - \frac{K}{|x|},$$

where $x_{K'} = \max\{K/|a_1|, x_K\}$.
Then we combine it with (A.11) and obtain that there exists K', K for $x > x_{K'} = \max\{K/|a_1|, x_K\}$ such

$$K' \left|\frac{1}{x}\right| \leq \left|f(x) - a_0\right|,$$

and since $x_{K'} > x_K$ and K does not occur in the expression it can be simplified to

$$\exists_{K'>0}\exists_{x_{K'}}\forall_{x>x_{K'}}\ K'\left|\frac{1}{x}\right| < \left|f(x)-a_0\right|,$$

hence we transformed (A.9) into a lower bound which proves the result.

□

Let the function $\Gamma(\lambda) = (\lambda-1)!$ where ! is a usual factorial and $\lambda \in \mathbb{N}$. For $\lambda \in \mathbb{R}$ it is defined through its integral form

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

Theorem 10 (Gamma function approximation)**.** The Gamma function $\Gamma(\lambda)$ can be approximated

$$\Gamma(\lambda) \sim e^{-\lambda}\lambda_\lambda\left(\frac{2\pi}{\lambda}\right)^{1/2}\left[1+\frac{1}{12\lambda}+\frac{1}{288\lambda^2}+\dots\right].$$

Proof. See [1], p.60. □

Lemma 14. Given inequalities

$$|A_i - B_i| \le C_i,\ i=1,\dots,m, \tag{A.13}$$

$$|A_i - B_i| \ge C_i',\ i=1,\dots,m, \tag{A.14}$$

where $m \in \mathbb{N}$ and $C_i > 0, C_i' > 0,\ i=1,\dots,m$, following inequalities holds

$$\left|\prod_{i=1}^m A_i - \prod_{i=1}^m B_i\right| \le \prod_{i=1}^m C_i + \sum_{j=1}^{m-1}\sum_{C_m^{m-j,j}}\prod_{k=1}^{j}\prod_{l=j+1}^{m} C_{i_k}B_{i_l}, \tag{A.15}$$

$$\left|\prod_{i=1}^m A_i - \prod_{i=1}^m B_i\right| \ge \sum_{j=1}^{m-1}\sum_{C_m^{m-j,j}}\prod_{k=1}^{j}\prod_{l=j+1}^{m} C_{i_k}'B_{i_l} - \prod_{i=1}^m C_i. \tag{A.16}$$

Proof. We start by introducing equality

$$\prod_{i=1}^m A_i = (A_1 - B_1)\prod_{i=2}^m A_i + B_1\prod_{i=2}^m A_i,$$

which we obtained by adding and deducting $B_1\prod_{i=2}^m A_i$ to $\prod_{i=1}^m A_i$.
Then again, we add and deduce, but this time $B_1B_2\prod_{l=3}^m A_l$ and $(A_1-B_1)B_2\prod_{l=3}^m A_l$

and get

$$\prod_{i=1}^{m} A_i = (A_1-B_1)(A_2-B_2)\prod_{i=3}^{m} A_i+(A_1-B_1)B_2\prod_{i=3}^{m} A_i+B_1(A_2-B_2)\prod_{i=3}^{m} A_i+B_1B_2\prod_{i=3}^{m} A_i.$$

We repeat that step until all A_i's in the products are replaced by $(A_i - B_i)$, which eventually leads to the equation

$$\prod_{i=1}^{m} A_i - \prod_{i=1}^{m} B_i = \prod_{i=1}^{m}(A_i - B_i) + \sum_{j=1}^{m-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{j} \prod_{l=j+1}^{m} (A_{i_k} - B_{i_k})B_{i_l} \qquad \text{(A.17)}$$

where $\sum_{C_m^{m-j,j}}$ is a sum over possible arrangements of the elements of the set $\{1, 2, \ldots, m\}$ into two groups, where elements does not repeat and within the group the order does not matter. First group is of the size $m - j$ and second j and their elements correspond, respectively, to the indecies $i_k, k = 1, \ldots, m - j$ and $i_l, l = m - j + 1 \ldots, m$.
Next we take absolute value of both sides and apply triangle inequality on RHS of (A.17)

$$\left|\prod_{i=1}^{m} A_i - \prod_{i=1}^{m} B_i\right| \leq \left|\prod_{i=1}^{m}(A_i - B_i)\right| + \left|\sum_{j=1}^{m-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{j} \prod_{l=j+1}^{m} (A_{i_k} - B_{i_k})B_{i_l}\right|,$$

and then we again apply triangle inequality and use multiplicity of absolute value

$$\left|\prod_{i=1}^{m} A_i - \prod_{i=1}^{m} B_i\right| \leq \prod_{i=1}^{m}|A_i - B_i| + \sum_{j=1}^{m-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{j} \prod_{l=j+1}^{m} |A_{i_k} - B_{i_k}||B_{i_l}|.$$

Now, we bound first term by applying all m the product of inequalities given by (A.13) and

$$\prod_{i=1}^{m}|A_i - B_i| \leq \prod_{i=1}^{m} C_i,$$

and for the second term also by applying (A.13) and obtain

$$\left|\prod_{i=1}^{m} A_i - \prod_{i=1}^{m} B_i\right| \leq \prod_{i=1}^{m} C_i + \sum_{j=1}^{m-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{j} \prod_{l=j+1}^{m} C_{i_k}|B_{i_l}|,$$

which is our upper bound (A.15) .

For the lower bound proof is analogical. The triangle triangle inequality implies $|a+b| \geq |b| - |a|$ for any vectors a, b. We use that fact and multiplicity of absolute value on (A.17)

$$\left|\prod_{i=1}^{m} A_i - \prod_{i=1}^{m} B_i\right| \geq \left|\sum_{j=1}^{m-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{j} \prod_{l=j+1}^{m} (A_{i_k} - B_{i_k}) B_{i_l}\right| - \left|\prod_{i=1}^{m}(A_i - B_i)\right| \geq$$
$$\geq \sum_{j=1}^{m-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{j} \prod_{l=j+1}^{m} |A_{i_k} - B_{i_k}||B_{i_l}| - \prod_{i=1}^{m} |A_i - B_i|.$$

Then we apply inequalities (A.14) and obtain

$$\left|\prod_{i=1}^{m} A_i - \prod_{i=1}^{m} B_i\right| \geq \sum_{j=1}^{m-1} \sum_{C_m^{m-j,j}} \prod_{k=1}^{j} \prod_{l=j+1}^{m} C'_{i_k} |B_{i_l}| - \prod_{i=1}^{m} C_i.$$

which is our lower bound (A.16). □

Lemma 15. Given inequalities

$$|A_1 - B_1| \leq C_1, \tag{A.18}$$
$$|A_2 - B_2| \leq C_2, \tag{A.19}$$
$$|A_2 - B_2| \geq C'_2, \tag{A.20}$$

where $C_1 > 0, C_2 > 0, C'_1, C'_2$ are constants, following inequalities holds

$$\left|\frac{A_1}{A_2} - \frac{B_1}{B_2}\right| \leq \left(\frac{C_1}{C'_2} + \left|\frac{B_1}{B_2}\right|\right) \frac{C_2}{|B_2| - C_2}. \tag{A.21}$$

Proof. First we prove (A.21).

We start with dividing (A.18) by $|A_2 - B_2|$ and using (A.20) and as a result we get

$$\frac{|A_1 - B_1|}{|A_2 - B_2|} \leq \frac{C_1}{C'_2},$$

and since absolute value is multiplicative we can merge absolute values of numerator and denominator of LHS and obtain

$$\left|\frac{A_1 - B_1}{A_2 - B_2}\right| \leq \frac{C_1}{C_2'}. \tag{A.22}$$

Next, by some manipulations we transform the expression inside absolute value of LHS

$$\begin{aligned}\frac{A_1 - B_1}{A_2 - B_2} =& \frac{A_1}{A_2} - \frac{B_1}{B_2} + \frac{A_1 B_2^2 + A_2^2 B_1 - 2A_2 B_1 B_2}{A_2 B_2 (A_2 - B_2)} = \\ =& \frac{A_1}{A_2} - \frac{B_1}{B_2} + \frac{B_2}{A_2 - B_2}\left(\frac{A_1}{A_2} - \frac{B_1}{B_2}\right) + \frac{B_1}{A_2 - B_2}\left(\frac{A_2}{B_2} - 1\right) = \\ =& \left(\frac{A_1}{A_2} - \frac{B_1}{B_2}\right)\frac{A_2}{A_2 - B_2} + \frac{B_1}{B_2},\end{aligned}$$

and insert the result back to (A.22)

$$\left|\left(\frac{A_1}{A_2} - \frac{B_1}{B_2}\right)\frac{A_2}{A_2 - B_2} + \frac{B_1}{B_2}\right| \leq \frac{C_1}{C_2'}.$$

The triangle inequality implies that $|a| - |b| \leq |a + b|$ for any vectors a, b and we use that fact to obtain

$$\left|\left(\frac{A_1}{A_2} - \frac{B_1}{B_2}\right)\frac{A_2}{A_2 - B_2}\right| \leq \frac{C_1}{C_2'} + \left|\frac{B_1}{B_2}\right|, \tag{A.23}$$

and by the multiplicity of absolute value the LHS can be factorized

$$\left|\frac{A_1}{A_2} - \frac{B_1}{B_2}\right|\left|\frac{A_2}{A_2 - B_2}\right| \leq \frac{C_1}{C_2'} + \left|\frac{B_1}{B_2}\right|.$$

Then we divide both sides by $|A_2/(A_2 - B_2)|$ and apply inequality (A.19)

$$\left|\frac{A_1}{A_2} - \frac{B_1}{B_2}\right| \leq \left(\frac{C_1}{C_2'} + \left|\frac{B_1}{B_2}\right|\right)\frac{C_2}{|A_2|}. \tag{A.24}$$

As absolute value is symmetric, from triangle inequality we have

$$|a| \geq |b| - |a - b|,$$

valid for arbitrary vectors a, b. Hence, setting $a = A_2$ and $b = B_2$ yields

$$|A_2| \geq |B_2| - |A_2 - B_2|,$$

and after application of (A.19)

$$|A_2| \geq |B_2| - C_2,$$

which we apply in (A.24) and get inequality (A.21).

□

Lemma 16. Given equalities

$$A_1 - B_1 = \sigma_1, \tag{A.25}$$

$$A_2 - B_2 = \sigma_2, \tag{A.26}$$

where $\sigma_1, \sigma_2 \in \mathbb{R}$ are some constants, following equality holds

$$\left|\frac{A_1}{A_2} - \frac{B_1}{B_2}\right| \leq \left|\frac{B_1\sigma_2}{B_2(B_2\sigma_2)}\right| + \left|\frac{\sigma_1}{B_2 + \sigma_2}\right|.$$

Proof. First we divide (A.25) by (A.26) and get

$$\frac{A_1}{A_2} = \frac{B_1 + \sigma_1}{B_2 - \sigma_2}, \tag{A.27}$$

and then we transform LHS

$$\frac{B_1 + \sigma_1}{B_2 + \sigma_2} = \frac{B_1}{B_2 + \sigma_2} + \frac{\sigma_1}{B_2 + \sigma_2},$$

then add and deduce B_1/B_2 and perform some manipulations

$$\begin{aligned}
\frac{B_1}{B_2 + \sigma_2} + \frac{\sigma_1}{B_2 + \sigma_2} &= \frac{B_1}{B_2} + \frac{B_1}{B_2 + \sigma_2} - \frac{B_1}{B_2} + \frac{\sigma_1}{B_2 + \sigma_2} = \\
&= \frac{B_1}{B_2} + \frac{B_1 B_2}{B_2(B_2 + \sigma_2)} - \frac{B_1(B_2 + \sigma_2)}{B_2(B_2 + \sigma_2)} + \frac{\sigma_1}{B_2 + \sigma_2} = \\
&= \frac{B_1}{B_2} - \frac{B_1\sigma_2}{B_2(B_2 + \sigma_2)} + \frac{\sigma_1}{B_2 + \sigma_2},
\end{aligned}$$

and then we put it back to (A.27), change the side of B_1/B_2 and obtain

$$\frac{A_1}{A_2} - \frac{B_1}{B_2} = -\frac{B_1\sigma_2}{B_2(B_2+\sigma_2)} + \frac{\sigma_1}{B_2+\sigma_2}.$$

Then we take absolute value of both side, apply triangle inequality on the RHS and get

$$\left|\frac{A_1}{A_2} - \frac{B_1}{B_2}\right| \leq \left|\frac{B_1\sigma_2}{B_2(B_2+\sigma_2)}\right| + \left|\frac{\sigma_1}{B_2+\sigma_2}\right|,$$

which is the result from the lemma. □

A.3 Probability

Definition 5 (Moment generating function)**.** Let X be a random variable with cumulative distirbution function(cdf) F_X. The moment generating function (mgf) of X (or F_X), denoted by $M_X(t) = Ee^{tX}$, provided that expectation exists for t in some neighborhood of 0.

Definition 6 (Convergence in distribution)**.** A sequence of random variables, $X_1, X_2, \ldots$ converges in distribution to a random variable X if

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x),$$

at all points x where $F_X(x)$ is continuous.

Theorem 11 (Convergence of mgfs)**.** Suppose $\{X_n, n = 1, 2, \ldots\}$ is a sequence of random variables, each with mgf $M_{X_n}(t)$. Furthermore, suppose that

$$\lim_{n\to\infty} M_{X_n}(t) = M_X(t),$$

for all t in the neighborhood of 0 and $M_X(t)$ is a mgf.
Then there is a unique cdf F_X whose moments are determinant by $M_X(t)$ and, for all x where $F_X(x)$ is continuous, we have

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x).$$

Hence, convergence of mgfs in the neighborhood of 0 implies convergence of cdfs.

Proof. Check [15] p.66. □

Theorem 12. Let X be any random variable and $g(x)$ nondecreasing function and $h(x)$ non-increasing function, such that $Eg(X)$, $Eh(X)$, and $E(g(X)h(X))$ exist, then

$$E[g(X)h(X)] \leq E[g(X)]E[h(X)].$$

Proof. Inequality is a special case of FKG inequality, for more details, check [4]. $\square$

A.4 Theory of Optimization

Definition 7 (Optimization problem in standard form)**.** An optimization problem in standard form has the form

$$\begin{aligned} \text{minimize } & f_0(x), \\ \text{subject to } & f_i(x) \leq 0, i = 1, \ldots, m, \\ & h_i(x) = 0, i = 1, \ldots, p, \end{aligned}$$

The vector $x \in \mathbb{R}^m$ is the optimization variable of the problem. The function $f_0(x)$ the objective function. The inequalities $f_i(x) \leq 0$ are called inequality constraints and equalities $h_i(x) = 0$ are equality constraints.
The domain for which on which objective function and all constraints is defined $\mathcal{D}$ we call a domain of the optimization problem. Any point $x \in \mathcal{D}$ is feasible if it satisfies the all the constraints.
Furthermore, the optimal value p^* is defined as

$$p^* = \inf\{f_0(x) | f_i(x) \leq 0, i = 1, \ldots, m,\ h_i(x) = 0, i = 1, \ldots, p\}.$$

and x^* is an optimal point (vector) , if x^* is feasible and $f_0(x^*) = p^*$.

Definition 8 (Convex problem)**.** An optimization problem is convex problem if it is of the form

$$\begin{aligned} \text{minimize } & f_0(x). \\ \text{subject to } & f_i(x) \leq 0, i = 1, \ldots, m, \\ & h_i(x) = 0, i = 1, \ldots, p, \end{aligned}$$

where $f_0, f_1, \ldots, f_m$ are convex. Furthermore requirements must be met

- optimized function f_0 is convex,
- inequality constraint functions must be convex,
- equality constraints must be affine.

Definition 9 (Affine hull)**.** We define affine hull by set of all affine combinations of points in some set $A \subseteq \mathbb{R}^m$ is called the affine hull of A and denoted by aff(A):

$$\operatorname{aff}(A) = \{\theta_1 x_1 + \ldots + \theta_k x_k \mid x_1, \ldots, x_k \in A\}$$

Definition 10 (Relative interior)**.** We define relative interior of the set C as

$$\operatorname{relint}(C) = \{x \in C \mid \exists_{r>0} B(x, r) \cap \operatorname{aff} A \subseteq C\}$$

Theorem 13 (weak Slater's condition)**.** The Slater's condition hold if optimization problem is convex and there exists $x \in \operatorname{relint}(\mathcal{D})$ with

$$f_i(x) \leq 0,\ i = 1, \ldots, k,\ f_i(x) < 0,\ i = k+1, \ldots, m.$$

where f_i are inequality constraints and first k of them are affine and relint$(\mathcal{D})$ is relative interior of the domain. Moreover if Slater's conditions hold then optimal vector (λ^*, ν^*) exists and strong duality occurs.

Proof. See [3], p.227. □

Theorem 14 (KKT conditions for convex problem)**.** The Karush-Kuhn-Tucker (KKT) conditions are

$$\begin{aligned} f_i(x^*) \leq 0, \\ h_i(x^*) = 0, \\ \tilde{\lambda}_i \geq 0, \\ \tilde{\lambda}_i f_i(x^*) = 0, \\ \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^{m} \nu_i^* \nabla h_i(x^*) = o. \end{aligned}$$

For any convex problem with differentiable objective and constraint functions, any points that satisfy KKT conditions are primal and dual optimal and strong duality holds.

Furthermore, if Slater's conditions holds then KKT are necessary and sufficient conditions for the optimality.

Proof. See [3], p.244. □

Bibliography

[1] R. Wong.
Asymptotic Approximations for Integrals.
SIAM, (2001).

[2] K.W. Breitung.
Asymptotic Approximations for Probability Integral.
Springer, (1994).

[3] S.Boyd and L.Vandenberghe.
Convex Optimization.
Cambridge University Press, (2004).

[4] C. M. Fortuin, J. Ginibre, and P. W. Kasteleyn.
Correlation inequalities on some partially ordered sets.
Comm. Math. Phys,22 (1971), no. 2, 89-103.

[5] F. Reif.
Fundamentals of Statistical and Thermal Physics.
Waveland Press, Inc, 2009

[6] Dedication of Journal.
Fiftieth anniversary of research and teaching by Viktor Pavlovich Maslov.
Theoretical and Mathematical Physics, 155(2), 674-677, (2008).

[7] Per Bak.
How Nature Works.
Springer-Verlag, (1996).

[8] D.L. Cohn.
Measure Theory.
Birkhauser, (1980).

[9] V.P. Maslov.
Nonlinear Averages in Economics.
Mathematical Notes, 78 (2005), no. 3, 347-363.

[10] V.P. Maslov.
Nonlinear averaging axioms in financial mathematics and stock prices.
Theory Probability Applications, 48 (2004), no. 4, 723-733.

[11] V.P. Maslov.
On a General Theorem of Set Theory Leading to the Gibbs, Bose-Einstein, and Pareto Distributions as well as to the Zipf-Mandelbrot Law for the Stock Market.
Mathematical Notes, 78 (2005), no. 6, 807-813.

[12] Maslov V.P. Maslova T.V.
On Zipfs Law and Rank Distributions in Linguistics and Semiotics.
Mathematical Notes, 80 (2006), no. 5, 679-691.

[13] V.N. Kolokoltsov.
Semiclassical Analysis for Diffusions and Strochastic processes.
Springer, (2000).

[14] B. Castellani, Frederic W. Hafferty.
Sociology and Complexity Science, A New Field of Inquiry.
Springer, 2009.

[15] G.Casella and R.L.Berger.
Statistical Inference.
Wadsworth, (2002).

[16] F.V. Kusmartsev.
Statistical mechanics of economics I.
Physics Letters A, 375 (2011) 966-973.

[17] R.K.Pathria and Paul D.Beale.
Statistical Mechanics 3rd ed.
Elsevier, 2011