CSMA konferenciakötet külső borító

Contents

S. Baran, K. Sikolya, Parameter estimation in linear regression driven by a Wiener sheet . . . 3
V. E. Bening, N. K. Galieva, R. A. Korolev, On rate of convergence in distribution of asymptotically normal statistics based on samples of random size . . . 17
E. Csáki, M. Csörgő, A. Földes, P. Révész, Random walk on half-plane half-comb structure . . . 29
I. Fazekas, Zs. Karácsony, R. Vas, Joint asymptotic normality of the kernel type density estimator for spatial observations . . . 45
I. Fazekas, B. Porvázsnyik, A generalized allocation scheme . . . 57
I. Fazekas, T. Tómács, On weighted averages of double sequences . . . 71
J. Fritz, Compensated compactness and relaxation at the microscopic level . 83
R. Giuliano, C. Macci, Large deviations for some normalized sums of exponentially distributed random variables . . . 109
A. Gut, Strong limit theorems for random fields . . . 125
M. Khaziakhmetov, T. Zakharova, The using of wavelet analysis in climatic challenges . . . 159
S. V. Nagaev, Renewal theorems in the case of attraction to the stable law with characteristic exponent smaller than unity . . . 173
Cs. Noszály, Experiments on the distance of two-dimensional samples . . . 193
V. A. Reisen, F. F. Molinares, Robust estimation in time series with long and short memory properties . . . 207
R. Shevchuk, Pasting of two one-dimensional diffusion processes . . . 225
I. Shevtsova, Moment-type estimates with asymptotically optimal structure for the accuracy of the normal approximation . . . 241
B. Székely, A. Kőrösi, P. Vámos, J. Bíró, Discriminatory processor sharing with access rate limitations . . . 309
A. S. Tylyshchak, Simulation of vibrations of a rectangular membrane with random initial conditions . . . 325

ANNALES MATHEMATICAE ET INFORMATICAE 39. (2012)

ANNALES MATHEMATICAE ET INFORMATICAE

TOMUS 39. (2012)

HUNGARIA, EGER

CSMA konferenciakötet belső borító

ANNALES MATHEMATICAE ET INFORMATICAE

International journal for mathematics and computer science

Referred by
Zentralblatt für Mathematik
and
Mathematical Reviews

The journal of the Institute of Mathematics and Informatics of Eszterházy Károly College is open for scientific publications in mathematics and computer science, where the field of number theory, group theory, constructive and computer aided geometry as well as theoretical and practical aspects of programming languages receive particular emphasis. Methodological papers are also welcome. Papers submitted to the journal should be written in English. Only new and unpublished material can be accepted.

Authors are kindly asked to write the final form of their manuscript in LaTeX. If you have any problems or questions, please write an e-mail to the managing editor Miklós Hoffmann: hofi@ektf.hu

The volumes are available at http://ami.ektf.hu

Printed by Debrecen University Press
Technical editor: Tibor Tómács

ANNALES MATHEMATICAE ET INFORMATICAE

VOLUME 39. (2012)

INSTITUTE OF MATHEMATICS AND INFORMATICS
ESZTERHÁZY KÁROLY COLLEGE
HUNGARY, EGER

Proceedings of the

Conference on Stochastic Models and their Applications

Faculty of Informatics, University of Debrecen
Debrecen, Hungary, August 22–24, 2011

Edited by

István Fazekas
György Terdik

HU ISSN 1787-5021 (Print)
HU ISSN 1787-6117 (Online)

Annales Mathematicae et Informaticae
39 *(2012) pp. 3–15*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Parameter estimation in linear regression driven by a Wiener sheet*

Sándor Baran, Kinga Sikolya

Faculty of Informatics, University of Debrecen
Kassai út 26, H-4028 Debrecen, Hungary
baran.sandor@inf.unideb.hu, sikolya.kinga@inf.unideb.hu

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

The problem of estimating the parameters of linear regression $Z(s,t) = m_1 g_1(s,t) + \cdots + m_p g_p(s,t) + W(s,t)$ based on observations of Z on a spatial domain G of special shape is considered, where the driving process W is a standard Wiener sheet and $g_1, \ldots, g_p$ are known functions. We provide an expression for the maximum likelihood estimator of the unknown parameters based on the observation of the process Z on the set G. Simulation results are also presented, where the driving random sheets are simulated with the help of their Karhunen-Loève expansions.

Keywords: Wiener sheet, maximum likelihood estimation, Radon-Nikodym derivative.

MSC: 60G60; 62M10; 62M30.

1. Introduction

The Wiener sheet is one of the most important examples of Gaussian random fields. It has various applications in statistical modelling. Wiener sheet appears as limiting process of some random fields defined on the interface of the Ising model [12], it is used to model random polymers [9], to describe the dynamics of Heath–Jarrow–Morton type forward interest rate models [10] or to model random mortality surfaces [6]. Further, [7] considers the problem of estimation of the mean

*Research has been supported by the Hungarian Scientific Research Fund under Grant No. OTKA T079128/2009 and partially supported by TÁMOP 4.2.1./B-09/1/KONV-2010-0007/IK/IT project. The project is implemented through the New Hungary Development Plan co-financed by the European Social Fund, and the European Regional Development Fund.

in a nonparametric regression on a two-dimensional regular grid of design points and constructs a Wiener sheet process on the unit square with a drift that is almost the mean function in the nonparametric regression.

In this paper we consider a linear regression driven by a Wiener sheet, that is random field

$$Z(s,t) := m_1 g_1(s,t) + \cdots + m_p g_p(s,t) + W(s,t) \tag{1.1}$$

observed on a domain G, where $g_1, \ldots, g_p$ are known functions and W is a standard Wiener sheet, and we determine the maximum likelihood estimator (MLE) of the unknown parameters $m_1, \ldots, m_p$.

In principle, the Radon-Nikodym derivative of Gaussian measures might be derived from the general Feldman-Hajek theorem [11], but in most of the cases explicit calculations cannot be carried out. For example, if $p = 1$ and $g_1 \equiv 1$ (shifted Wiener sheet), the MLE of the unknown parameter is given in [13] and the estimator is expressed as a function of a usually unknown random variable satisfying some characterizing equation. In several cases the exact form of this random variable can be derived by a method proposed in [14] based on linear stochastic partial differential equations.

Special case 1: Baran et al. [3] studied the case, when $p = 1$ and the random field Z is observed on a rectangular domain $G := [a_1, a_2] \times [b_1, b_2]$, $[a_1, a_2], [b_1, b_2] \subset (0, \infty)$. Assuming that g_1 is absolutely continuous with respect to the Lebesgue measure and $\partial_1 \partial_2 g_1 \in L^2(G)$, they proved that the MLE of the shift parameter m_1 has the form $\widehat{m}_1 = A^{-1}\zeta$, where

$$A = \frac{g_1^2(a_1,b_1)}{a_1 b_1} + \int_{a_1}^{a_2} \frac{[\partial_1 g_1(u,b_1)]^2}{b_1} \mathrm{d}u + \int_{b_1}^{b_2} \frac{[\partial_2 g_1(a_1,v)]^2}{a_1} \mathrm{d}v \tag{1.2}$$
$$+ \iint_G [\partial_1 \partial_2 g_1(u,v)]^2 \mathrm{d}u \mathrm{d}v,$$

$$\zeta = \frac{g_1(a_1,b_1) Z(a_1,b_1)}{a_1 b_1} + \int_{a_1}^{a_2} \frac{\partial_1 g_1(u,b_1)}{b_1} Z(\mathrm{d}u, b_1) + \int_{b_1}^{b_2} \frac{\partial_2 g_1(a_1,v)}{a_1} Z(a_1, \mathrm{d}v) \tag{1.3}$$
$$+ \iint_G \partial_1 \partial_2 g_1(u,v) Z(\mathrm{d}u, \mathrm{d}v),$$

and it has normal distribution with mean m_1 and variance $1/A$. For $g_1 \equiv 1$ we have $\widehat{m}_1 = Z(a_1, b_1)$.

Special case 2: Arató N.M. [2] considered the case $p = 1$ and $g_1 \equiv 1$ and using Rozanov's method found the MLE of the shift parameter m_1 when the process is observed on a special domain

$$G \subset \widetilde{G} := \{(s,t) \in \mathbb{R}^2 \ : \ a \le s \le b,\, t \ge \gamma(s) \text{ or } s > b,\, t \ge \gamma(b)\}$$

containing an ε–strip of $\Gamma := \{(s, \gamma(s)) : s \in (a, b)\}$, i.e. for some $\varepsilon > 0$

$$\{(s,t) \in \mathbb{R}^2 : \ s \in [a, a+\varepsilon], t \in [\gamma(s), \gamma(a)] \text{ or } s \in [a+\varepsilon, b], t \in [\gamma(s), \gamma(s)+\varepsilon]\} \subset G,$$

where $\gamma : [a, b] \subset (0, \infty) \to \mathbb{R}$ is continuous and strictly decreasing with $\gamma(b) > 0$.

Baran et al. [4] considered the same model and under much weaker conditions $\gamma \in C^2(a, b)$, $\gamma'(a) := \lim_{s\downarrow a} \gamma'(s) \in [-\infty, 0]$ and $\gamma'(b) := \lim_{s\uparrow b} \gamma'(s) \in [-\infty, 0]$ exist, and

$$\int_a^b \frac{|\gamma'(s)\gamma''(s)|}{(1+\gamma'(s)^2)^2}\,\mathrm{d}s < \infty,$$

they proved the result of [2, Theorem 2]. They showed that the MLE of the shift parameter m_1 has the form $\widehat{m_1} = A^{-1}\zeta$, where

$$A = \frac{1}{b\gamma(b)} + \int_a^b \frac{\mathrm{d}s}{s^2\gamma(s)}, \qquad \zeta = c_1 Z(a, \gamma(a)) + c_2 Z(b, \gamma(b)) + \int_\Gamma y_1 Z + \int_\Gamma y_2 \partial_{\mathrm{n}} Z, \tag{1.4}$$

c_1, c_2 are constants depending on γ and γ' at a and b, y_1 and y_2 are functions of $\gamma, \gamma', \gamma''$, and $\partial_{\mathrm{n}} Z$ denotes the normal derivative of Z [4, Definition 4.1].

If $\gamma'(a) = -\infty$ we have

$$\zeta = \frac{Z(b, \gamma(b))}{b\gamma(b)} + \int_a^b \frac{Z(s, \gamma(s))}{s^2\gamma(s)}\mathrm{d}s - \int_a^b \frac{1}{s\gamma(s)} Z(\mathrm{d}s, \gamma(s)).$$

In the present paper we consider the same type of domain G as in [5] and give a natural extension of their result for the general model (1.1). We also present some simulation results to illustrate the theoretical ones where the Wiener sheet is simulated with the help of its Karhunen-Loève expansion (see e.g. [8]).

2. Model and estimator

Consider the model (1.1) with some given functions $g_1, \ldots, g_p : \mathbb{R}_+^2 \to \mathbb{R}$ and with unknown regression parameters $m_1, \ldots, m_p \in \mathbb{R}$. Let $[a, c] \subset (0, \infty)$ and $b_1, b_2 \in (a, c)$, let $\gamma_{1,2} : [a, b_1] \to \mathbb{R}$ and $\gamma_0 : [b_2, c] \to \mathbb{R}$ be continuous, strictly decreasing functions and let $\gamma_1 : [b_1, c] \to \mathbb{R}$ and $\gamma_2 : [a, b_2] \to \mathbb{R}$ be continuous, strictly increasing functions with $\gamma_{1,2}(b_1) = \gamma_1(b_1) > 0$, $\gamma_2(b_2) = \gamma_0(b_2)$, $\gamma_{1,2}(a) = \gamma_2(a)$ and $\gamma_1(c) = \gamma_0(c)$. Consider the curve $\Gamma := \Gamma_{1,2} \cup \Gamma_1 \cup \Gamma_2 \cup \Gamma_0$, where

$$\Gamma_{1,2} := \Big\{\big(s, \gamma_{1,2}(s)\big) : s \in [a, b_1]\Big\}, \qquad \Gamma_1 := \Big\{\big(s, \gamma_1(s)\big) : s \in [b_1, c]\Big\},$$
$$\Gamma_2 := \Big\{\big(s, \gamma_2(s)\big) : s \in [a, b_2]\Big\}, \qquad \Gamma_0 := \Big\{\big(s, \gamma_0(s)\big) : s \in [b_2, c]\Big\},$$

and for a given $\varepsilon > 0$ let $\Gamma^\varepsilon_{1,2}$, Γ^ε_1, Γ^ε_2 and Γ^ε_0 denote the inner ε-strip of $\Gamma_{1,2}$, Γ_1, Γ_2 and Γ_0, respectively, that is e.g.

$$\Gamma^\varepsilon_{1,2} := \big\{(s,t)\in\mathbb{R}^2 : s\in[a,a+\varepsilon],\ t\in[\gamma_{1,2}(s),\gamma_{1,2}(a)] \text{ or } s\in[a+\varepsilon,b_1],\ t\in[\gamma_{1,2}(s),\gamma_{1,2}(s)+\varepsilon]\big\}.$$

Suppose that there exists an $\varepsilon > 0$ such that

$$\Gamma^\varepsilon_1 \cap \Gamma^\varepsilon_2 = \emptyset \quad\text{and}\quad \Gamma^\varepsilon_{1,2}\cap\Gamma^\varepsilon_0 = \emptyset, \tag{2.1}$$

and consider the set $G := G_1 \cup G_2 \cup G_3$, where

$$\begin{aligned} G_1 &:= \big\{(s,t)\in\mathbb{R}^2 : s\in[a, b_1\wedge b_2],\, t\in[\gamma_{1,2}(s),\gamma_2(s)]\big\},\\ G_2 &:= \begin{cases} \big\{(s,t)\in\mathbb{R}^2 : s\in[b_1,b_2],\, t\in[\gamma_1(s),\gamma_2(s)]\big\}, & \text{if } b_1\le b_2,\\ \big\{(s,t)\in\mathbb{R}^2 : s\in[b_2,b_1],\, t\in[\gamma_{1,2}(s),\gamma_0(s)]\big\}, & \text{if } b_1 > b_2,\end{cases}\\ G_3 &:= \big\{(s,t)\in\mathbb{R}^2 : s\in[b_1\vee b_2, c],\, t\in[\gamma_1(s),\gamma_0(s)]\big\}.\end{aligned}$$

An example of such a set of observations can be seen of Figure 1.

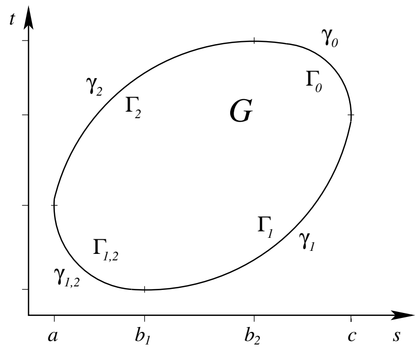


Figure 1: An example of a set of observations G

The following theorem is an extension of Theorem 2.1 of [5] and can be proved in a similar way. The proof is based on the discrete approximation method described in [3, 4, 5], which relies on the results of [1, Section 2.3.2].

Theorem 2.1. *If $g_1,\dots,g_p$ are twice continuously differentiable inside G and the partial derivatives $\partial_1 g_i, \partial_2 g_i$ and $\partial_1\partial_2 g_i$, $i=1,\dots,p$, can be continuously extended to G then the probability measures P_Z and P_W, generated on $C(G)$ by the sheets Z and W, respectively, are equivalent and the Radon-Nikodym derivative of P_Z with respect to P_W equals*

$$\frac{\mathrm{d}\mathsf{P}_Z}{\mathrm{d}\mathsf{P}_W}(Z) = \exp\left\{-\frac{1}{2}\big(\mathbf{m}^\top A\mathbf{m} - 2\zeta^\top\mathbf{m}\big)\right\},$$

where $A := \big(A_{k,\ell}\big)_{k,\ell=1}^{p}$, $\mathbf{m} := (m_1, \dots, m_p)^\top$ *and* $\zeta := \big(\zeta_1, \dots, \zeta_p\big)^\top$ *with*

$$
\begin{aligned}
A_{k,\ell} :=\ & \frac{g_k\big(b_1, \gamma_{1,2}(b_1)\big)\, g_\ell\big(b_1, \gamma_{1,2}(b_1)\big)}{b_1 \gamma_{1,2}(b_1)} \qquad (2.2)\\
&+ \int_a^{b_1} \frac{\big[g_k\big(s, \gamma_{1,2}(s)\big) - s\partial_1 g_k\big(s, \gamma_{1,2}(s)\big)\big]\big[g_\ell\big(s, \gamma_{1,2}(s)\big) - s\partial_1 g_\ell\big(s, \gamma_{1,2}(s)\big)\big]}{s^2 \gamma_{1,2}(s)}\, \mathrm{d}s\\
&+ \int_{b_1}^{c} \frac{\partial_1 g_k\big(s, \gamma_1(s)\big)\, \partial_1 g_\ell\big(s, \gamma_1(s)\big)}{\gamma_1(s)}\, \mathrm{d}s + \int_{\gamma_2(a)}^{\gamma_2(b_2)} \frac{\partial_2 g_k\big(\gamma_2^{-1}(t), t\big)\, \partial_2 g_\ell\big(\gamma_2^{-1}(t), t\big)}{\gamma_2^{-1}(t)}\, \mathrm{d}t\\
&+ \int_{\gamma_{1,2}(b_1)}^{\gamma_{1,2}(a)} \frac{\partial_2 g_k\big(\gamma_{1,2}^{-1}(t), t\big)\, \partial_2 g_\ell\big(\gamma_{1,2}^{-1}(t), t\big)}{\gamma_{1,2}^{-1}(t)}\, \mathrm{d}t + \iint_G \partial_1\partial_2 g_k(s,t)\, \partial_1\partial_2 g_\ell(s,t) \mathrm{d}s\, \mathrm{d}t,
\end{aligned}
$$

and

$$
\begin{aligned}
\zeta_k :=\ & \frac{g_k\big(b_1, \gamma_{1,2}(b_1)\big) Z\big(b_1, \gamma_{1,2}(b_1)\big)}{b_1 \gamma_{1,2}(b_1)} + \int_{b_1}^{c} \frac{\partial_1 g_k\big(s, \gamma_1(s)\big)}{\gamma_1(s)}\, Z\big(\mathrm{d}s, \gamma_1(s)\big) \qquad (2.3)\\
&+ \int_a^{b_1} \frac{\big[g_k\big(s, \gamma_{1,2}(s)\big) - s\partial_1 g_k\big(s, \gamma_{1,2}(s)\big)\big]}{s^2 \gamma_{1,2}(s)} \Big[Z\big(s, \gamma_{1,2}(s)\big)\, \mathrm{d}s - sZ\big(\mathrm{d}s, \gamma_{1,2}(s)\big)\Big]\\
&+ \int_{\gamma_2(a)}^{\gamma_2(b_2)} \frac{\partial_2 g_k\big(\gamma_2^{-1}(t), t\big)}{\gamma_2^{-1}(t)}\, Z\big(\gamma_2^{-1}(t), \mathrm{d}t\big) + \int_{\gamma_{1,2}(b_1)}^{\gamma_{1,2}(a)} \frac{\partial_2 g_k\big(\gamma_{1,2}^{-1}(t), t\big)}{\gamma_{1,2}^{-1}(t)}\, Z\big(\gamma_{1,2}^{-1}(t), \mathrm{d}t\big)\\
&+ \iint_G \partial_1\partial_2 g_k(s,t)\, Z(\mathrm{d}s, \mathrm{d}t).
\end{aligned}
$$

If $\det(A) \neq 0$ *then the maximum likelihood estimator of the parameter vector* $\mathbf{m}$ *based on the observations* $\{Z(s,t) : (s,t) \in G\}$ *has the form* $\widehat{\mathbf{m}} = A^{-1}\zeta$ *and has a* p*-dimensional normal distribution with mean* $\mathbf{m}$ *and covariance matrix* A^{-1}.

Remark 2.2. Observe that all six terms of matrix A are non-negative definite matrices, so A is non-negative definite, too. Hence, to ensure $\det(A) \neq 0$ it suffices to have at least one positive definite among the terms, which fulfils e.g. if $g_1, \dots, g_p$ are linearly independent.

Remark 2.3. We remark that the weighted L^2-Riemann integrals of partial derivatives of the Wiener sheet (and of other L^2-processes) along a curve are defined in the sense of [5, Definition 4.1]. This means that if Z is an L^2-process given along an ε-neighborhood of a curve $\Gamma := \big\{(s, \gamma(s)) : s \in [a,b]\big\}$, where $\gamma : [a,b] \to \mathbb{R}$ is strictly monotone and $y:[a,b]\to\mathbb{R}$ is a function, then

$$\int_a^b y(s)\,Z(\mathrm{d}s,\gamma(s)) := \operatorname*{l.i.m.}_{h\to 0}\frac{1}{h}\int_a^b y(s)\big[Z(s+h,\gamma(s))-Z(s,\gamma(s))\big]\,\mathrm{d}s,$$

$$\int_{\gamma(a)}^{\gamma(b)} y(\gamma^{-1}(t))\,Z(\gamma^{-1}(t),\mathrm{d}t) := \operatorname*{l.i.m.}_{h\to 0}\frac{1}{h}\int_{\gamma(a)}^{\gamma(b)} y(\gamma^{-1}(t))\big[Z(\gamma^{-1}(t),t+h)-Z(\gamma^{-1}(t),t)\big]\mathrm{d}t,$$

if the right hand sides exist.

Example 2.4. Consider the model

$$Z(s,t)=m_1(s^2+t^2)+m_2(s+t)+m_3(s\cdot t)+W(s,t),\qquad (s,t)\in G,$$

where $W(s,t),\ (s,t)\in[u-r,u+r]\times[v-r,v+r]$ is a standard Wiener sheet and G is a circle with center at (u,v) and radius r. Thus

$$\begin{aligned}
\gamma_{1,2}(s)&=v-\sqrt{r^2-(s-u)^2}, && s\in[u-r,u],\\
\gamma_1(s)&=v-\sqrt{r^2-(s-u)^2}, && s\in[u,u+r],\\
\gamma_2(s)&=v+\sqrt{r^2-(s-u)^2}, && s\in[u-r,u],\\
\gamma_0(s)&=v+\sqrt{r^2-(s-u)^2}, && s\in[u,u+r],\\
\gamma_{1,2}^{-1}(t)&=u-\sqrt{r^2-(t-v)^2}, && t\in[v-r,v],\\
\gamma_2^{-1}(t)&=u-\sqrt{r^2-(t-v)^2}, && t\in[v,v+r].
\end{aligned}$$

In this case the distinct elements of the symmetric matrix A defined by (2.2) are the following

$$\begin{aligned}
A_{1,1}&=\frac{(u^2+(v-r)^2)^2}{u(v-r)}+\int_{u-r}^{u}\frac{(\gamma_{1,2}^2(s)-s^2)^2}{s^2\gamma_{1,2}(s)}\mathrm{d}s+4\int_u^{u+r}\frac{s^2}{\gamma_1(s)}\mathrm{d}s\\
&\quad+4\int_{v-r}^{v}\frac{t^2}{\gamma_{1,2}^{-1}(t)}\mathrm{d}t+4\int_v^{v+r}\frac{t^2}{\gamma_2^{-1}(t)}\mathrm{d}t,\\
A_{1,2}&=\frac{(u^2+(v-r)^2)(u+v-r)}{u(v-r)}+\int_{u-r}^{u}\frac{\gamma_{1,2}^2(s)-s^2}{s^2}\mathrm{d}s\\
&\quad+2\int_u^{u+r}\frac{s}{\gamma_1(s)}\mathrm{d}s+2\int_{v-r}^{v}\frac{t}{\gamma_{1,2}^{-1}(t)}\mathrm{d}t+2\int_v^{v+r}\frac{t}{\gamma_2^{-1}(t)}\mathrm{d}t,\\
A_{2,2}&=\frac{(u+v-r)^2}{u(v-r)}+\int_{u-r}^{u}\frac{\gamma_{1,2}(s)}{s^2}\mathrm{d}s+\int_u^{u+r}\frac{1}{\gamma_1(s)}\mathrm{d}s
\end{aligned}$$

$$+\int_{v-r}^{v}\frac{1}{\gamma_{1,2}^{-1}(t)}\mathrm{d}t+\int_{v}^{v+r}\frac{1}{\gamma_{2}^{-1}(t)}\mathrm{d}t, \tag{2.4}$$

$$A_{1,3}=u^2+(v-r)^2+2r(u+2v)+r^2,$$

$$A_{2,3}=u+v+2r,$$

$$A_{3,3}=u(v-r)+r(v+2u)-\frac{3\pi r^2}{4},$$

while the components of $\zeta=\big(\zeta_1,\zeta_2,\zeta_3\big)^\top$ defined by (2.3) are

$$\zeta_1=\frac{\big(u^2+(v-r)^2\big)Z\big(b_1,\gamma_{1,2}(b_1)\big)}{u(v-r)}+\int_{u}^{u+r}\frac{2s}{\gamma_1(s)}\,Z\big(\mathrm{d}s,\gamma_1(s)\big)$$
$$+\int_{v-r}^{v}\frac{2t}{\gamma_2^{-1}(t)}\,Z\big(\gamma_2^{-1}(t),\mathrm{d}t\big)$$
$$+\int_{u-r}^{u}\frac{(\gamma_{1,2}^2(s)-s^2)}{s^2\gamma_{1,2}(s)}\Big[Z\big(s,\gamma_{1,2}(s)\big)\,\mathrm{d}s-sZ\big(\mathrm{d}s,\gamma_{1,2}(s)\big)\Big]$$
$$+\int_{v}^{v+r}\frac{2t}{\gamma_{1,2}^{-1}(t)}\,Z\big(\gamma_{1,2}^{-1}(t),\mathrm{d}t\big),$$

$$\zeta_2=\frac{\big(u+v-r\big)Z\big(b_1,\gamma_{1,2}(b_1)\big)}{u(v-r)}+\int_{u}^{u+r}\frac{1}{\gamma_1(s)}\,Z\big(\mathrm{d}s,\gamma_1(s)\big)$$
$$+\int_{v-r}^{v}\frac{1}{\gamma_2^{-1}(t)}\,Z\big(\gamma_2^{-1}(t),\mathrm{d}t\big) \tag{2.5}$$
$$+\int_{u-r}^{u}\frac{1}{s^2}\Big[Z\big(s,\gamma_{1,2}(s)\big)\,\mathrm{d}s-sZ\big(\mathrm{d}s,\gamma_{1,2}(s)\big)\Big]+\int_{v}^{v+r}\frac{1}{\gamma_{1,2}^{-1}(t)}\,Z\big(\gamma_{1,2}^{-1}(t),\mathrm{d}t\big),$$

$$\zeta_3=Z\big(b_1,\gamma_{1,2}(b_1)\big)+\int_{u}^{u+r}Z\big(\mathrm{d}s,\gamma_1(s)\big)+\int_{v-r}^{v}Z\big(\gamma_2^{-1}(t),\mathrm{d}t\big)$$
$$+\int_{v}^{v+r}Z\big(\gamma_{1,2}^{-1}(t),\mathrm{d}t\big)+\iint_{G}Z(\mathrm{d}s,\mathrm{d}t).$$

3. Simulation results

To illustrate the theoretical results of [2, 3, 4, 5] and of Theorem 2.1 we performed computer simulations using Matlab 2010a. In order to simulate a Wiener sheet $W(s,t),\ 0 \le s \le S,\ 0 \le t \le T$, we considered its Karhunen-Loève expansion, that is

$$W(s,t) \approx \sum_{j,k=1}^{n} \omega_{j,k} \frac{8\sqrt{ST}}{(\pi^2)(2k-1)(2j-1)} \sin\left(\frac{\pi(2j-1)t}{2T}\right) \sin\left(\frac{\pi(2k-1)s}{2S}\right), \quad (3.1)$$

where $\{\omega_{j,k} : 1 \le j,k \le n\}$ are independent standard normal random variables [8]. Figure 2 shows an approximation of the Wiener sheet with $n = 150$. Obviously, there are other methods of simulating a Wiener sheet e.g. with the help of discretization and using the independence of increments (see e.g. [15]). However, in order to calculate our estimators we need a method which provides us whole realizations of the sheet.

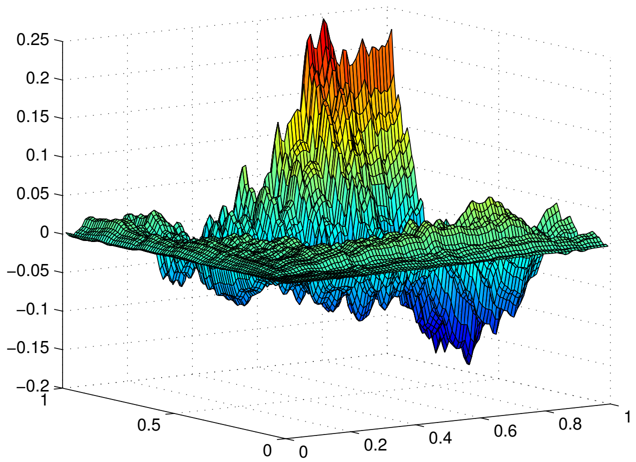


Figure 2: Simulation of Wiener sheet, $n = 150$

In each of the following examples 1000 independent samples of the driving Wiener sheet were simulated with n varying between 25 and 100 and the means of the estimates of the parameters and the empirical variances or covariance matrices of ζ defined by (1.3), (1.4) and (2.3), respectively, were calculated.

Example 3.1. Consider the model

$$Z(s,t) = W(s,t) + m(s^2 + t^2),$$

where $W(s,t)$, $(s,t) \in G = [1,3]^2$ is a standard Wiener sheet (see Special case 1). Components of A and of the approximation of ζ can be given in the following closed form:

$$A = \frac{4b_2^3 - b_1^3}{3a_1} + \frac{4a_2^3 - a_1^3}{3b_1} + 2a_1b_1,$$

$$\begin{aligned}
&\zeta - mA \approx \\
&\sum_{j,k=1}^{n} \omega_{j,k} \frac{8\sqrt{ST}}{(\pi^2)(2k-1)(2j-1)} \Bigg\{ \frac{a_1^2+b_1^2}{a_1b_1} \sin\left(\frac{\pi(2k-1)a_1}{2S}\right) \sin\left(\frac{\pi(2j-1)b_1}{2T}\right) \\
&+ \frac{2}{b_1} \sin\left(\frac{\pi(2j-1)b_1}{2T}\right) \left[\frac{2S}{\pi(2k-1)} \left(\cos\left(\frac{\pi(2k-1)a_2}{2S}\right) - \cos\left(\frac{\pi(2k-1)a_1}{2S}\right) \right) \right. \\
&\left. + a_2 \sin\left(\frac{\pi(2k-1)a_2}{2S}\right) - a_1 \sin\left(\frac{\pi(2k-1)a_1}{2S}\right) \right] + \frac{2}{a_1} \sin\left(\frac{\pi(2k-1)a_1}{2S}\right) \\
&\times \left[\frac{2T}{\pi(2j-1)} \left(\cos\left(\frac{\pi(2j-1)b_2}{2T}\right) - \cos\left(\frac{\pi(2j-1)b_1}{2T}\right) \right) \right. \\
&\left. + b_2 \sin\left(\frac{\pi(2j-1)b_2}{2T}\right) - b_1 \sin\left(\frac{\pi(2j-1)b_1}{2T}\right) \right] \Bigg\},
\end{aligned}$$

where

$$\zeta = \frac{(a_1^2+b_1^2)Z(a_1,b_1)}{a_1b_1} + \int_{a_1}^{a_2} \frac{2u}{b_1} Z(\mathrm{d}u, b_1) + \int_{b_1}^{b_2} \frac{2v}{a_1} Z(a_1, \mathrm{d}v).$$

The theoretical parameter value is $m = 5$, while $A = 33.3333$. On Figure 3 the means of the estimates of the parameter and the estimated variances of ζ are plotted versus the level n of the approximation (3.1). In case of $n = 100$ we have $\widehat{m} = 5.0007$ and $\widehat{A} = 33.7233$.

Example 3.2. Consider the model

$$Z(s,t) = W(s,t) + m,$$

where $W(s,t), (s,t) \in G$, is a standard Wiener sheet and G is a set satisfying conditions of Special case 2 and Γ is a part of a circle with center at the origin, that is $\gamma(s) = \sqrt{r^2 - s^2}$ with some $r > 0$ and with $[a,b] \subset (0,r)$ [4, Example 1.2] . Then

$$A = \frac{1}{r^2}\left(\frac{\sqrt{r^2-a^2}}{a} + \frac{b}{\sqrt{r^2-b^2}}\right), \qquad c_1 = \frac{\sqrt{r^2-a^2}}{r^2a}, \qquad c_2 = \frac{b}{r^2\sqrt{r^2-b^2}},$$

$$y_1(s, \sqrt{r^2-s^2}) \equiv 0, \qquad y_2(s, \sqrt{r^2-s^2}) \equiv -\frac{1}{r^2}$$

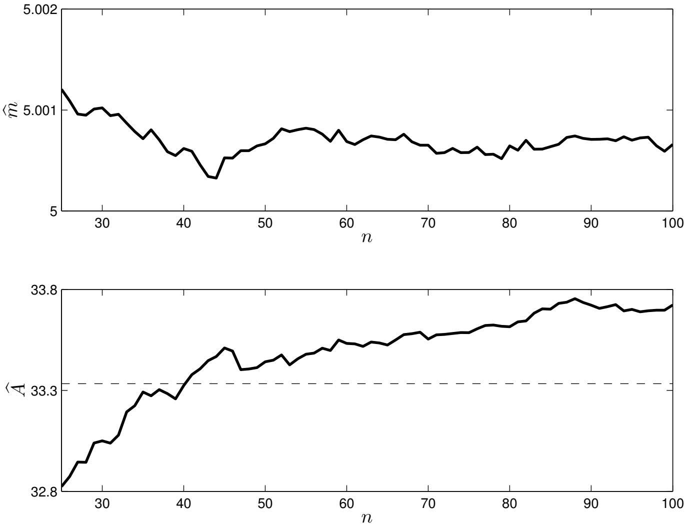


Figure 3: Means of the estimates of m and estimated variances of ζ in Example 3.1 for $25 \leq n \leq 100$

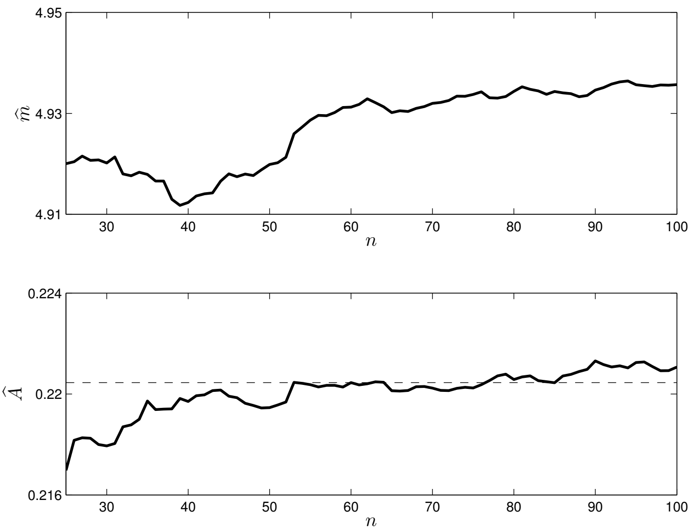


Figure 4: Means of the estimates of m and estimated variances of ζ in Example 3.2 for $25 \leq n \leq 100$

and

$$\zeta - mA \approx \sum_{j,k=1}^{n} \omega_{j,k} \frac{8\sqrt{ST}}{\pi^2 r^2 (2k-1)(2j-1)}$$

$$\times \left\{ \frac{\sqrt{r^2-a^2}}{a} \sin\left(\frac{\pi(2k-1)a}{2S}\right) \sin\left(\frac{\pi(2j-1)\sqrt{r^2-a^2}}{2T}\right) \right.$$
$$+ \frac{b}{\sqrt{r^2-b^2}} \sin\left(\frac{\pi(2k-1)b}{2S}\right) \sin\left(\frac{\pi(2j-1)\sqrt{r^2-b^2}}{2T}\right)$$
$$- \int_a^b \left\{ \frac{\pi(2k-1)s}{2S\sqrt{r^2-s^2}} \cos\left(\frac{\pi(2k-1)s}{2S}\right) \sin\left(\frac{\pi(2j-1)\sqrt{r^2-s^2}}{2T}\right) \right.$$
$$\left. \left. + \frac{\pi(2j-1)}{2T} \sin\left(\frac{\pi(2k-1)s}{2S}\right) \cos\left(\frac{\pi(2j-1)\sqrt{r^2-s^2}}{2T}\right) \right\} \mathrm{d}s \right\},$$

where ζ is defined by (1.4).

Let parameter value be $m = 5$ and choose $a = 1$, $b = 3$ and $r = 5$ yielding $A = 0.2205$. On Figure 4 the means of the estimates of the parameter and the estimated variances of ζ are plotted versus the level n of the approximation (3.1). In case of $n = 100$ we have $\widehat{m} = 4.9357$ and $\widehat{A} = 0.2213$.

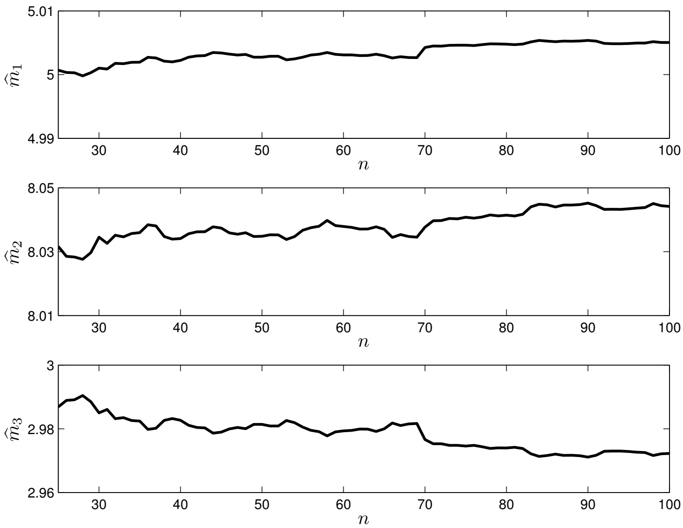


Figure 5: Means of the estimates of the components of $\mathbf{m}$ in Example 3.3 for $25 \leq n \leq 100$

Example 3.3. Consider the same model

$$Z(s,t) = m_1(s^2+t^2) + m_2(s+t) + m_3(s \cdot t) + W(s,t), \qquad (s,t) \in G,$$

as in Example 2.4, where $W(s,t)$, $(s,t) \in [0,8]^2$, is a standard Wiener sheet and G is a circle with center at $(6,6)$ and radius $r = 2$. In this case the entries of the matrix A defined by (2.4) and the approximation of the components of $\zeta = (\zeta_1, \zeta_2, \zeta_3)^\top$ defined by (2.5) can be calculated using numerical integration, where Matlab function `quad` is applied (recursive adaptive Simpson quadrature).

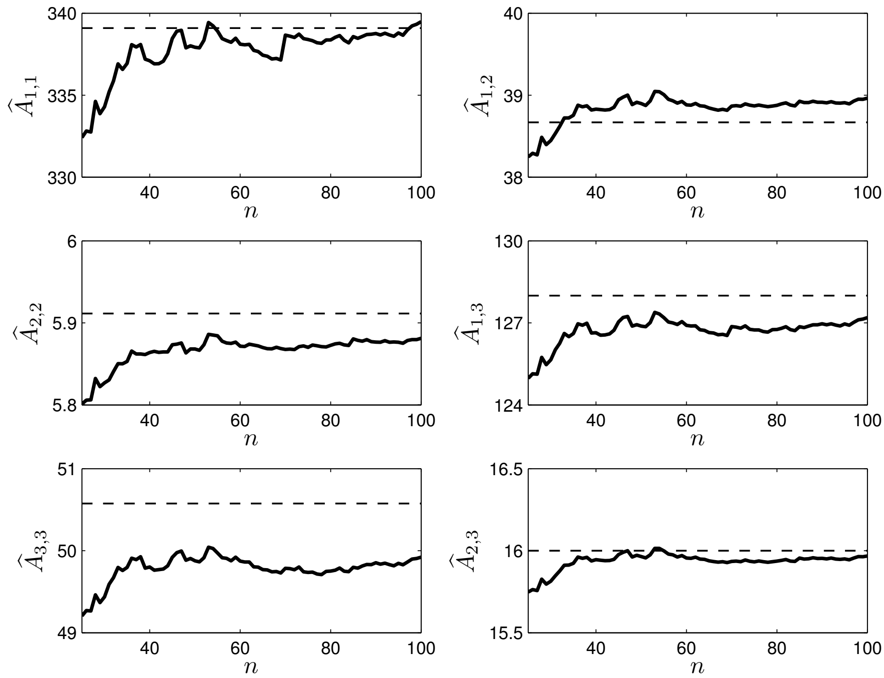


Figure 6: Estimated covariances of ζ in Example 3.3 for $25 \leq n \leq 100$

The theoretical parameter values are $m_1 = 5$, $m_2 = 8$ and $m_3 = 3$, while the theoretical covariance matrix of ζ equals

$$A = \begin{pmatrix} 339.0895 & 38.6688 & 128.0000 \\ 38.6688 & 5.9115 & 16.0000 \\ 128.0000 & 16.0000 & 50.5752 \end{pmatrix}.$$

On Figure 5 the means of the estimates of the three parameters, while on Figure 6 the estimated covariances of ζ are plotted versus the level n of the approximation (3.1). In case of $n = 100$ we have $(5.0050, 8.0442, 2.9723)$ for the mean and

$$\widehat{A} = \begin{pmatrix} 339.4824 & 38.9639 & 127.1914 \\ 38.9639 & 5.8811 & 15.9680 \\ 127.1914 & 15.9680 & 49.9207 \end{pmatrix}$$

for the covariance matrix.

References

[1] Arató, M. (1982) *Linear Stochastic Systems with Constant Coefficients. A Statistical Approach.* Springer-Verlag, Berlin; in Russian: Nauka, Moscow, 1989.

[2] Arató, N. M. (1997) Mean estimation of Brownian sheet. *Comput. Math. Appl.* **33**, no. 8, 13–25.

[3] Baran, S., Pap, G. and Zuijlen, M. v. (2003) Estimation of the mean of stationary and nonstationary Ornstein-Uhlenbeck processes and sheets. *Comput. Math. Appl.* **45**, no. 4-5, 563–579.

[4] Baran, S., Pap, G. and Zuijlen, M. v. (2004) Estimation of the mean of a Wiener sheet. *Stat. Inference Stoch. Process.* **7**, no. 3, 279–304.

[5] Baran, S., Pap, G., Zuijlen M. v. (2011) Parameter estimation of a shifted Wiener sheet. *Statistics* **45**, no. 4, 319–335.

[6] Biffis, E. and Millossovich, P. (2006) A bidimensional approach to mortality risk. *Decisions Econ. Finan.* **29**, no. 2, 71–94.

[7] Carter, A. V. (2006) A continuous Gaussian approximation to a nonparametric regression in two dimensions. *Bernoulli* **12**, no. 1, 143–156.

[8] Deheuvels, P., Peccati, G. and Yor, M. (2006) On quadratic functionals of the Brownian sheet and related processes. *Stochastic Process. Appl.* **116**, no. 3, 493–538.

[9] Douglas, J. F. (1996) Swelling and growth of polymers, membranes, and sponges. *Phys. Rev. E* **54**, no. 3, 2677–2689.

[10] Goldstein, R. (2000) The term structure of interest rates as a random field. *Rev. Financ. Stud.* **13**, no. 2, 365–384.

[11] Kuo, H. H. (1975) *Gaussian Measures in Banach Spaces.* Springer-Verlag, Berlin.

[12] Kuroda, K. and Manaka, H. (1987) The interface of the Ising model and the Brownian sheet. *J. Stat. Phys.* **47**, no. 5-6, 979–984.

[13] Rozanov, Yu. A. (1968) *Gaussian Infinite-dimensional Distributions.* (in Russian). Tr. Mat. Inst. Akad. Nauk SSSR, Nauka, Moscow.

[14] Rozanov, Yu. A. (1990) Some boundary problems for generalized random fields. *Theor. Probab. Appl.* **35**, no. 4, 707–724.

[15] Toth, D. (2004) Adding interior points to an existing Brownian sheet lattice. *Statist. Probab. Lett.* **66**, no. 3, 221–227.

Annales Mathematicae et Informaticae
***39** (2012) pp. 17–28*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

On rate of convergence in distribution of asymptotically normal statistics based on samples of random size

V. E. Bening[a], N. K. Galieva[b], R. A. Korolev[c]

[a]Lomonosov Moscow State University
Faculty of Computational Mathematics and Cybernetics
Institute of Informatics Problems of the Russian Academy of Sciences (IPI RAN)
bening@yandex.ru

[b]Kazakhstan Branch of Lomonosov Moscow State University

[c]Peoples' Friendship University of Russia

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

In the present paper we prove a general theorem which gives the rates of convergence in distribution of asymptotically normal statistics based on samples of random size. The proof of the theorem uses the rates of convergences in distribution for the random size and for the statistics based on samples of nonrandom size.

Keywords: sample of random size; asymptotically normal statistic; transfer theorem; rate of convergence; mixture of distributions; Laplace distribution; Student's distribution

1. Introduction

Asymptotic properties of distributions of sums of random number of random variables are subject of many papers (see e.g. Gnedenko&Fahim, 1969; Gnedenko, 1989; Kruglov&Korolev, 1990; Gnedenko&Korolev, 1996; Bening&Korolev, 2002; von Chossy&Rappl, 1983). This kind of sums are widely used in insurance, economics, biology, etc. (see Gnedenko, 1989; Gnedenko, 1998; Bening&Korolev, 2002). However, in mathematical statistics and its applications, there are common statistics that are not sums of observations. Examples are the rank statistics, U-statistics, linear combinations of order statistics, etc. In this case the statistics are often situations when the sample size is not predetermined and can be regarded as random. For example, in reliability testing the number of failed devices at a particular time is a random variable.

Generally, in most cases related to the analysis and processing of experimental data, we can assume that the number of random factors, influencing the observed values, is itself random and varies from observation to observation. Therefore, instead of different variants of the central limit theorem, proving the normality of the limiting distribution of classical statistics, in such situations we should rely on their analogues for samples of random size. This makes it natural to study the asymptotic behavior of distributions of statistics of general form, based on samples of random size. For example, Gnedenko (1989) examines the asymptotic properties of the distributions of sample quantiles constructed from samples of random size.

In this paper we estimate the rate of convergence of distribution functions of asymptotically normal statistics based on samples of random size. The estimations depend on the rates of convergences of distributions of the random size of sample and the statistic based on sample of nonrandom size. Such statements are usually called transfer theorems. In the present paper we prove transfer theorems concerning estimates of convergence rate.

In this paper we use the following notation and symbols: $\mathbb{R}$ as real numbers, $\mathbb{N}$ as positive integers, $\Phi(x)$, $\varphi(x)$ as standard normal distribution function and density.

In Section 2 we give a sketch of the proof of a general transfer theorem, Sections 3, 4 and 5 contain the main theorems, their proofs and examples.

Consider random variables $N_1, N_2, \ldots$ and $X_1, X_2, \ldots$ defined on a common measurable space $(\Omega, \mathcal{A}, \mathsf{P})$. The random variables $X_1, X_2, \ldots X_n$ denote observations, n is a nonrandom size of sample, the random variable N_n denotes a random size of sample and depends on a natural parameter $n \in \mathbb{N}$. Suppose that the random variables N_n take on positive integers for any $n \geq 1$, that is $N_n \in \mathbb{N}$, and do not depend on $X_1, X_2, \ldots$. Suppose that $X_1, X_2, \ldots$ are independent and identically distributed observations having a distribution function $F(x)$.

Let $T_n = T_n(X_1, \ldots, X_n)$ be some statistic, that is a real measurable function on observations $X_1, \ldots, X_n$. The statistic T_n is called asymptotically normal with parameters $(\mu, 1/\sigma^2)$, $\mu \in \mathbb{R}$, $\sigma > 0$, if

$$\mathsf{P}(\sigma\sqrt{n}(T_n - \mu) < x) \longrightarrow \Phi(x), \quad n \to \infty, \quad x \in \mathbb{R}, \tag{1.1}$$

where $\Phi(x)$ is the standard normal distribution function.

The asymptotically normal statistics are abundant. Recall some examples of asymptotically normal statistics: the sample mean (assuming nonzero variances), the maximum likelihood estimators (under weak regularity conditions), the central order statistics and many others.

For any $n \geq 1$ define the random variable T_{N_n} by

$$T_{N_n}(\omega) \equiv T_{N_n(\omega)}(X_1(\omega), \ldots, X_{N_n(\omega)}(\omega)), \quad \omega \in \Omega. \tag{1.2}$$

Therefore, T_{N_n} is a statistic constructed from the statistic T_n and from the sample of random size N_n.

In Gnedenko&Fahim (1969) and Gnedenko (1989), the first and second transfer theorems are proved for the case of sums of independent random variables and sample quantiles.

Theorem 1.1 (Gnedenko, 1989). *Let $X_1, X_2, \dots$ be independent and identically distributed random variables and $N_n \in \mathbb{N}$ denotes a sequence of random variables which are independent of $X_1, X_2, \dots$. If there exist real numbers $b_n > 0$, $a_n \in \mathbb{R}$ such that*

$$\textit{1.}\qquad \mathsf{P}\Big(\frac{1}{b_n}\sum_{i=1}^{n}(X_i - a_n) < x\Big) \longrightarrow \Psi(x), \quad n \to \infty$$

and

$$\textit{2.}\qquad \mathsf{P}\Big(\frac{N_n}{n} < x\Big) \longrightarrow H(x), \quad H(0+) = 0, \quad n \to \infty,$$

where $\Psi(x)$ and $H(x)$ are distribution functions, then, as $n \to \infty$,

$$\mathsf{P}\Big(\frac{1}{b_n}\sum_{i=1}^{N_n}(X_i - a_n) < x\Big) \longrightarrow G(x), \quad n \to \infty,$$

where the distribution function $G(x)$ is defined by its characteristic function

$$g(t) = \int_0^\infty (\psi(t))^z \,\mathrm{d}H(z)$$

and $\psi(t)$ is the characteristic function of $\Psi(x)$.

The proof of the theorem can be read in Gnedenko (1998).

Theorem 1.2 (Gnedenko, 1989). *Let $X_1, X_2, \dots$ be independent and identically distributed random variables and $N_n \in \mathbb{N}$ is a sequence of random variables which are independent of $X_1, X_2, \dots$, and let $X_{\gamma:n}$ be the sample quantile of order $\gamma \in (0,1)$ constructed from sample $X_1, \dots, X_n$. If there exist real numbers $b_n > 0$, $a_n \in \mathbb{R}$ such that*

$$\textit{1.}\qquad \mathsf{P}\Big(\frac{1}{b_n}(X_{\gamma:n} - a_n) < x\Big) \longrightarrow \Phi(x), \quad n \to \infty$$

and

$$\textit{2.}\qquad \mathsf{P}\Big(\frac{N_n}{n} < x\Big) \longrightarrow H(x), \quad H(0+) = 0, \quad n \to \infty,$$

where $H(x)$ is a distribution function, then, as $n \to \infty$,

$$\mathsf{P}\Big(\frac{1}{b_n}(X_{\gamma:N_n} - a_n) < x\Big) \longrightarrow G(x), \quad n \to \infty$$

where the distribution function $G(x)$ is a mixture of normal distribution with the mixing distribution H

$$G(x) = \int_0^\infty \Phi(x\sqrt{y})\, dH(y).$$

In Bening&Korolev (2005), the following general transfer theorem is proved for asymptotically normal statistics (1.1).

Theorem 1.3. *Let $\{d_n\}$ be an increasing and unbounded sequence of positive integers. Suppose that $N_n \to \infty$ in probability as $n \to \infty$. Let $T_n(X_1, \dots, X_n)$ be an asymptotically normal statistics, that is*

$$\mathsf{P}(\sigma\sqrt{n}(T_n - \mu) < x) \longrightarrow \Phi(x), \quad n \to \infty.$$

Then a necessary and sufficient condition for a distribution function $G(x)$ to satisfy

$$\mathsf{P}(\sigma\sqrt{d_n}(T_{N_n} - \mu) < x) \longrightarrow G(x), \quad n \to \infty,$$

is that there exists a distribution function $H(x)$ with $H(0+) = 0$ satisfying

$$\mathsf{P}(N_n < d_n x) \longrightarrow H(x), \quad n \to \infty, \quad x > 0,$$

and $G(x)$ has a form

$$G(x) = \int_0^\infty \Phi(x\sqrt{y})\, dH(y), \quad x \in \mathbb{R},$$

that is the distribution $G(x)$ is a mixture of the normal law with the mixing distribution H.

Now, we give a brief sketch of proof of Theorem 1.3 to make references later.

2. Sketch of proof of Theorem 1.3

The proof of Theorem 1.3 is closely related to the proof of Theorems 6.6.1 and 6.7.3 for random sums in Kruglov&Korolev (1990).

By the formula of total probability, we have

$$\mathsf{P}\Big(\sigma\sqrt{d_n}(T_{N_n} - \mu) < x\Big) - G(x)$$

$$= \sum_{k=1}^{\infty} \mathsf{P}(N_n = k)\mathsf{P}\Big(\sigma\sqrt{k}(T_k - \mu) < \sqrt{k/d_n}x\Big) - G(x)$$
$$= \sum_{k=1}^{\infty} \mathsf{P}(N_n = k)\Big(\Phi\big(\sqrt{k/d_n}x\big) - G(x)\Big)$$
$$+ \sum_{k=1}^{\infty} \mathsf{P}(N_n = k)\Big(\mathsf{P}\Big(\sigma\sqrt{k}(T_k - \mu) < \sqrt{k/d_n}x\Big) - \Phi\big(\sqrt{k/d_n}x\big)\Big)$$
$$\equiv J_{1n} + J_{2n}. \tag{2.1}$$

From definition of $G(x)$ the expression for J_{1n} can be written in the form

$$J_{1n} = \int_0^{\infty} \Phi(x\sqrt{y})\,\mathrm{d}\mathsf{P}(N_n < d_n y) - \int_0^{\infty} \Phi(x\sqrt{y})\,\mathrm{d}H(y)$$
$$= \int_0^{\infty} \Phi(x\sqrt{y})\,\mathrm{d}\big(\mathsf{P}(N_n < d_n y) - H(y)\big).$$

Using the formula of integration by parts for Lebesgue integral (see e.g. Theorem 2.6.11 in Shiryaev, 1995) yields

$$J_{1n} = -\int_0^{\infty} \big(\mathsf{P}(N_n < d_n y) - H(y)\big)\,\mathrm{d}\Phi(x\sqrt{y}). \tag{2.2}$$

By the condition of the present theorem,

$$\mathsf{P}(N_n < d_n y) - H(y) \longrightarrow 0, \quad n \to \infty$$

for any fixed $y \in \mathbb{R}$, therefore, by the dominated convergence theorem (see e.g. Theorem 2.6.3 in Shiryaev, 1995), we have

$$J_{1n} \longrightarrow 0, \quad n \to \infty.$$

Consider J_{2n}. For simplicity, instead of the condition for the statistic T_n to be asymptotically normal (see (1.1)), we suggest a stronger condition which describes the rate of convergence of distributions of T_n to the normal law. Suppose that the following condition is satisfied.

Condition 1. *There exist real numbers $\alpha > 0$ and $C_1 > 0$ such that*

$$\sup_x \Big|\mathsf{P}\Big(\sigma\sqrt{n}(T_n - \mu) < x\Big) - \Phi(x)\Big| \le \frac{C_1}{n^{\alpha}}, \quad n \in \mathbb{N}.$$

From the condition we obtain estimates for J_{2n}. We have

$$|J_{2n}| = \Big|\sum_{k=1}^{\infty} \mathsf{P}(N_n = k)\Big(\mathsf{P}\Big(\sigma\sqrt{k}(T_k - \mu) < \sqrt{k/d_n}x\Big) - \Phi\big(\sqrt{k/d_n}x\big)\Big)\Big|$$

$$\leq C_1 \sum_{k=1}^{\infty} \mathsf{P}(N_n = k)\frac{1}{k^\alpha} = C_1 \mathsf{E}(N_n)^{-\alpha} = \frac{C_1}{d_n^\alpha}\mathsf{E}(N_n/d_n)^{-\alpha}. \tag{2.3}$$

Since, by the condition of theorem, the random variables N_n/d_n have a weak limit, then the expectation $\mathsf{E}(N_n/d_n)^{-\alpha}$ is typically bounded. Because $d_n \to \infty$, from the last inequality it follows that

$$J_{2n} \longrightarrow 0, \quad n \to \infty.$$

3. The main results

Suppose that the limiting behavior of distribution functions of the normalized random size is described by the following condition.

Condition 2. *There exist real numbers $\beta > 0$, $C_2 > 0$ and a distribution $H(x)$ with $H(0+) = 0$ such that*

$$\sup_{x\geq 0}\Big|\mathsf{P}\Big(\frac{N_n}{n} < x\Big) - H(x)\Big| \leq \frac{C_2}{n^\beta}, \quad n \in \mathbb{N}.$$

Theorem 3.1. *If for the statistic $T_n(X_1, \dots, X_n)$ condition 1 is satisfied, for the random sample size N_n condition 2 is satisfied, then the following inequality holds*

$$\sup_x \Big|\mathsf{P}\Big(\sigma\sqrt{n}(T_{N_n} - \mu) < x\Big) - G(x)\Big| \leq C_1 \mathsf{E} N_n^{-\alpha} + \frac{C_2}{2n^\beta},$$

where the distribution $G(x)$ has the form

$$G(x) = \int_0^\infty \Phi(x\sqrt{y})\,\mathrm{d}H(y), \quad x \in \mathbb{R}.$$

Corollary 3.2. *The statement of the theorem remains valid if the normal law is replaced by any limiting distribution.*

Corollary 3.3. *If the moments $\mathsf{E}(N_n/n)^{-\alpha}$ are bounded uniformly in n, that is*

$$\mathsf{E}\Big(\frac{N_n}{n}\Big)^{-\alpha} \leq C_3, \quad C_3 > 0, \quad n \in \mathbb{N},$$

then the right side of the inequality in the statement of the theorem has the form

$$\frac{C_1 C_3}{n^\alpha} + \frac{C_2}{2n^\beta} = \mathcal{O}\big(n^{-\min(\alpha,\beta)}\big).$$

Corollary 3.4. *By Hölder's inequality for $0 < \alpha \le 1$, the following estimate holds*

$$\mathsf{E}N_n^{-\alpha} \le \Big(\mathsf{E}\frac{1}{N_n}\Big)^{\alpha},$$

which is useful from practical viewpoint. In this case, the right side of the inequality has the form

$$C_1\Big(\mathsf{E}\frac{1}{N_n}\Big)^{\alpha} + \frac{C_2}{2n^{\beta}}.$$

Corollary 3.5. *Note that, condition 2 means that the random variables N_n/n converge weakly to V which has the distribution $H(x)$. From the definition of weak convergence with function $x^{-\alpha}, \quad x \ge 1$, for $N_n \ge n, n \in \mathbb{N}$, it follows that*

$$\mathsf{E}\Big(\frac{N_n}{n}\Big)^{-\alpha} \longrightarrow \mathsf{E}\frac{1}{V^{\alpha}}, \quad n \to \infty,$$

that is the moments $\mathsf{E}(N_n/n)^{-\alpha}$ are bounded in n and, therefore, the estimate from Corollary 3.3 holds.

The case $N_n \ge n$ appears when the random variable N_n takes on values $n, 2n, \dots, kn$ with equal probabilities $1/k$ for any fixed $k \in \mathbb{N}$. In this case, the random variables N_n/n do not depend on n and, therefore, converge weakly to V which takes values $1, 2, \dots, k$ with equal probability $1/k$.

Corollary 3.6. *From the proof of the theorem it follows that skipping of conditions 1 and 2 yields the following statement*

$$\begin{aligned}
&\sup_x \Big|\mathsf{P}\Big(\sigma\sqrt{n}(T_{N_n} - \mu) < x\Big) - G(x)\Big| \\
&\le \sum_{k=1}^{\infty} \mathsf{P}\big(N_n = k\big) \sup_x \Big|\mathsf{P}\Big(\sigma\sqrt{k}(T_k - \mu) < x\Big) - \Phi(x)\Big| \\
&\quad + \frac{1}{2}\sup_{x\ge 0}\Big|\mathsf{P}\Big(\frac{N_n}{n} < x\Big) - H(x)\Big|.
\end{aligned}$$

Following the proof of Theorem 3.1 (see Section 2 and 4), we can formulate more general result.

Theorem 3.7. *Let a random element $\mathbf{X}_n$ in some measurable space and random variable N_n be defined on a common measurable space and independent for any $n \in \mathbb{N}$. Suppose that a real-valued statistic $T_n = T_n(\mathbf{X}_n)$ and the random variable N_n satisfy the following conditions.*

1. There exist real numbers $\alpha > 0$, $\sigma > 0$, $\mu \in \mathbb{R}$, $C_1 > 0$ and a sequence $0 < d_n \uparrow +\infty, n \to \infty$, such that

$$\sup_x \Big|\mathsf{P}\Big(\sigma\sqrt{d_n}(T_n - \mu) < x\Big) - \Phi(x)\Big| \le \frac{C_1}{n^{\alpha}}, \quad n \in \mathbb{N}.$$

2. There exist a number $C_2 > 0$, a sequence $0 < \delta_n \downarrow 0$, $n \to \infty$ and a distribution function $H(x)$ with $H(0+) = 0$ such that

$$\sup_{x \geq 0} \Big| \mathsf{P}\Big(\frac{N_n}{d_n} < x \Big) - H(x) \Big| \leq C_2 \delta_n, \quad n \in \mathbb{N}.$$

Then the following inequality holds

$$\sup_x \Big| \mathsf{P}\Big(\sigma \sqrt{d_n} (T_{N_n} - \mu) < x \Big) - G(x) \Big| \leq C_1 \mathsf{E} N_n^{-\alpha} + \frac{C_2}{2} \delta_n,$$

where the distribution function $G(x)$ has the form

$$G(x) = \int_0^\infty \Phi(x\sqrt{y}) \, \mathrm{d}H(y), \quad x \in \mathbb{R}.$$

4. Proof of Theorem 3.1

Suppose $x \geq 0$. Using formulas (2.1)–(2.3) with $d_n = n$ yields

$$\sup_{x \geq 0} \Big| \mathsf{P}\Big(\sigma \sqrt{n} (T_{N_n} - \mu) < x \Big) - G(x) \Big| \leq I_{1n} + I_{2n}, \tag{4.1}$$

where

$$I_{1n} = \sup_{x \geq 0} \int_0^\infty \big| \mathsf{P}(N_n < ny) - H(y) \big| \, \mathrm{d}\Phi(x\sqrt{y}), \tag{4.2}$$

$$I_{2n} = \sum_{k=1}^\infty \mathsf{P}(N_n = k) \sup_{x \geq 0} \Big| \mathsf{P}\Big(\sigma \sqrt{k} (T_k - \mu) < \sqrt{k/n} x \Big) - \Phi\big(\sqrt{k/n} x \big) \Big|. \tag{4.3}$$

To estimate the variable I_{1n} we use equality (4.2) and condition 2,

$$I_{1n} \leq \frac{C_2}{n^\beta} \sup_{x \geq 0} \int_0^\infty \mathrm{d}\Phi(x\sqrt{y}) = \frac{C_2}{2n^\beta}. \tag{4.4}$$

The series in I_{2n} (see (4.3)) is estimated by using condition 1.

$$I_{2n} \leq C_1 \sum_{k=1}^\infty \frac{1}{k^\alpha} \mathsf{P}(N_n = k) = C_1 \mathsf{E} N_n^{-\alpha}. \tag{4.5}$$

Note that the estimate (4.5) is valid for $x < 0$. For I_{1n} and negative x, we have (see (2.1) and (2.2))

$$I_{1n} = \sup_{x < 0} \Big| \int_0^\infty \big(\mathsf{P}(N_n < ny) - H(y) \big) \, \mathrm{d}\Phi(x\sqrt{y}) \Big|$$

$$= \sup_{x<0} \Big| \int_0^\infty \big(\mathsf{P}(N_n < ny) - H(y)\big)\, \mathrm{d}\Phi(|x|\sqrt{y}) \Big|$$

$$\le \sup_{x \ge 0} \int_0^\infty \big|\mathsf{P}(N_n < ny) - H(y)\big|\, \mathrm{d}\Phi(x\sqrt{y}),$$

and we can use (4.4) again. The statement of the theorem follows from (4.1), (4.4) and (4.5). The theorem is proved.

5. Examples

We consider two examples of use of Theorem 3.1 when the limiting distribution function $G(x)$ is known.

5.1. Student's distribution

Bening&Korolev (2005) shows that if the random sample size N_n has the negative binomial distribution with parameters $p = 1/n$ and $r > 0$, that is (in particular, for $r = 1$, it is the geometric distribution)

$$\mathsf{P}(N_n = k) = \frac{(k+r-2)\cdots r}{(k-1)!} \frac{1}{n^r} \Big(1 - \frac{1}{n}\Big)^{k-1}, \quad k \in \mathbb{N},$$

then, for an asymptotically normal statistic T_n the following limiting relationship holds (see Corollary 2.1 in Bening&Korolev, 2005)

$$\mathsf{P}(\sigma\sqrt{n}(T_{N_n} - \mu) < x) \longrightarrow G_{2r}(x\sqrt{r}), \quad n \to \infty, \tag{5.1}$$

where $G_{2r}(x)$ is Student's distribution with parameter $\gamma = 2r$, having density

$$p_\gamma(x) = \frac{\Gamma((\gamma+1)/2)}{\sqrt{\pi\gamma}\,\Gamma(\gamma/2)} \Big(1 + \frac{x^2}{\gamma}\Big)^{-(\gamma+1)/2}, \quad x \in \mathbb{R},$$

where $\Gamma(\cdot)$ is the gamma function, and $\gamma > 0$ is a shape parameter (if the parameter γ is a positive integer, then it is called the number of degrees of freedom). In our situation the parameter may be arbitrary small, and we have typical heavy-tailed distribution. If $\gamma = 2$, that is $r = 1$, then the distribution function $G_2(x)$ can be found explicitly

$$G_2(x) = \frac{1}{2}\Big(1 + \frac{x}{\sqrt{2 + x^2}}\Big), \quad x \in \mathbb{R}.$$

For $r = 1/2$, we obtain the Cauchy distribution.

Bening et al. (2004) gives an estimate of rate of convergence for random sample size, for $0 < r < 1$,

$$\sup_{x \ge 0} \Big| \mathsf{P}\Big(\frac{N_n}{\mathsf{E}N_n} < x\Big) - H_r(x) \Big| \le \frac{C_r}{n^{r/(r+1)}}, \quad C_r > 0, \quad n \in \mathbb{N}, \tag{5.2}$$

where

$$H_r(x) = \frac{r^r}{\Gamma(r)} \int\limits_0^x e^{-ry} y^{r-1}\,\mathrm{d}y, \quad x \geq 0,$$

for $r = 1$, the right side of the inequality can be replaced by $1/(n-1)$. So, $H_r(x)$ is a distribution with parameter $r \in (0, 1]$, and

$$\mathsf{E} N_n = r(n-1) + 1. \tag{5.3}$$

From

$$(1+x)^\gamma = \sum_{k=0}^{\infty} \frac{\gamma(\gamma-1)\cdots(\gamma-k+1)}{k!} x^k, \quad |x| < 1, \quad \gamma \in \mathbb{R},$$

we have

$$\mathsf{E} N_n^{-1} = \frac{1}{(n-1)(1-r)} \Big(\frac{1}{n^{r-1}} - 1 \Big) = \mathcal{O}(n^{-r}), \quad 0 < r < 1, \quad n \in \mathbb{N}. \tag{5.4}$$

If the Berry-Esseen estimate is valid for the rate of convergence of distribution of T_n, that is

$$\sup_x \Big| \mathsf{P}\Big(\sigma\sqrt{n}(T_n - \mu) < x\Big) - \Phi(x) \Big| = \mathcal{O}\Big(\frac{1}{\sqrt{n}}\Big), \quad n \in \mathbb{N}, \tag{5.5}$$

then from Theorem 3.1 with $\alpha = 1/2$, $\beta = r/(r+1)$, from relations (5.1)–(5.4) and Corollary 3.4, we have the following estimate

$$\sup_x \Big| \mathsf{P}\Big(\sigma\sqrt{n}(T_{N_n} - \mu) < x\Big) - G_{2r}(x\sqrt{r}) \Big|$$
$$= \mathcal{O}\Big(\frac{1}{n^{r/2}}\Big) + \mathcal{O}\Big(\frac{1}{n^{r/(r+1)}}\Big) = \mathcal{O}\Big(\frac{1}{n^{r/2}}\Big), \quad r \in (0,1),\ n \in \mathbb{N}. \tag{5.6}$$

5.2. Laplace distribution

Consider Laplace distribution with distribution function $\Lambda_\gamma(x)$ and density

$$\lambda_\gamma(x) \ = \ \frac{1}{\gamma\sqrt{2}} \exp\Big\{-\frac{\sqrt{2}|x|}{\gamma}\Big\}, \quad \gamma > 0, \quad x \in \mathbb{R}.$$

Bening&Korolev (2008) gives a sequence of random variables $N_n(m)$ which depends on the parameter $m \in \mathbb{N}$. Let $Y_1, Y_2, \ldots$ be independent and identically distributed random variables with some continuous distribution function. Let m be a positive integer and

$$N(m) = \min\{i \geq 1 : \max_{1 \leq j \leq m} Y_j < \max_{m+1 \leq k \leq m+i} Y_k\}.$$

It is well-known that such random variables have the discrete Pareto distribution

$$\mathsf{P}(N(m) \geq k) = \frac{m}{m+k-1}, \quad k \geq 1. \tag{5.7}$$

Now, let $N^{(1)}(m), N^{(2)}(m), \ldots$ be independent random variables with the same distribution (5.7). Define the random variable

$$N_n(m) = \max_{1 \le j \le n} N^{(j)}(m),$$

then Bening&Korolev (2008) shows that

$$\lim_{n\to\infty} \mathsf{P}\Big(\frac{N_n(m)}{n} < x\Big) = e^{-m/x}, \quad x > 0, \tag{5.8}$$

and, for an asymptotically normal statistic T_n, the following relationship holds

$$\mathsf{P}(\sigma\sqrt{n}(T_{N_n(m)} - \mu) < x) \longrightarrow \Lambda_{1/m}(x), \quad n \to \infty,$$

where $\Lambda_{1/m}(x)$ is the Laplace distribution function with parameter $\gamma = 1/m$.

Lyamin (2010) gives the estimate for the rate of convergence for (5.8),

$$\sup_{x\ge 0}\Big|\mathsf{P}\Big(\frac{N_n(m)}{n} < x\Big) - e^{-m/x}\Big| \le \frac{C_m}{n}, \quad C_m > 0, \quad n \in \mathbb{N}. \tag{5.9}$$

If the Berry-Esseen estimate is valid for the rate of convergence of distribution for the statistic (see (5.5)), then from Corollary 3.4 for $\alpha = 1/2$, $\beta = 1$ and from inequality (5.9), we have

$$\sup_x \Big|\mathsf{P}\Big(\sigma\sqrt{n}(T_{N_n(m)} - \mu) < x\Big) - \Lambda_{1/m}(x)\Big| = \mathcal{O}\Big((\mathsf{E}N_n^{-1}(m))^{1/2}\Big) + \mathcal{O}\big(n^{-1}\big). \tag{5.10}$$

Consider the variable $\mathsf{E}N_n^{-1}(m)$. From definition of $N_n(m)$ and inequality (5.7), we have

$$\mathsf{P}(N_n(m) = k) = \Big(\frac{k}{m+k}\Big)^n - \Big(\frac{k-1}{m+k-1}\Big)^n = mn \int\limits_{k-1}^{k} \frac{x^{n-1}}{(m+x)^{n+1}}\,\mathrm{d}x,$$

therefore,

$$\mathsf{E}N_n^{-1}(m) = \sum_{k=1}^{\infty} \frac{1}{k}\mathsf{P}(N_n(m) = k) = mn \sum_{k=1}^{\infty} \frac{1}{k} \int\limits_{k-1}^{k} \frac{x^{n-1}}{(m+x)^{n+1}}\,\mathrm{d}x$$

$$\le mn \sum_{k=1}^{\infty} \int\limits_{k-1}^{k} \frac{x^{n-2}}{(m+x)^{n+1}}\,\mathrm{d}x mn \int\limits_{0}^{\infty} \frac{x^{n-2}}{(m+x)^{n+1}}\,\mathrm{d}x.$$

To calculate the last integral we use the following formula (see formula 856.12 in Dwight, 1961)

$$\int\limits_0^\infty \frac{x^{m-1}}{(a+bx)^{m+n}}\,\mathrm{d}x = \frac{\Gamma(m)\Gamma(n)}{a^n b^m \Gamma(m+n)} \quad a, b, m, n > 0.$$

We have

$$\mathsf{E}N_n^{-1}(m) \leq mn\frac{\Gamma(n-1)\Gamma(2)}{m^2\Gamma(n+1)} = \frac{1}{m(n-1)} = O(n^{-1}).$$

Now, by this formula and (5.10), we obtain

$$\sup_x \Big|\mathsf{P}\Big(\sigma\sqrt{n}(T_{N_n(m)} - \mu) < x\Big) - \Lambda_{1/m}(x)\Big| = \mathcal{O}\Big(\frac{1}{\sqrt{n}}\Big).$$

References

[1] Bening, V. E. & Korolev V. Y. (2002). *Generalized Poisson Models and Their Applications in Insurance and Finance*, VSP Press, Netherlands.

[2] Bening, V. E. & Korolev, V. Y. (2005). On an application of the Student distribution in the theory of probability and mathematical statistics. *Theory Probab. Appl.*, **49**(3), 377–391.

[3] Bening, V. E. & Korolev, V. Y. (2008). Some statistical problems related to the Laplace distribution. *Informatics and its Applications, IPI RAN*, **2**(2), 19–34 [in Russian].

[4] Bening, V. E., Korolev, V. Y. & U Da (2004). Estimates of rates of convergences of distributions of some statistics to the Student distribution. *J. People's Friendship Univer. Russia*, **1**(12), 59–74.

[5] Von Chossy, R. & Rappl, G. (1983). Some approximation methods for the distribution of random sums. *Insurance: Mathematics and Economics*, **2**, 251–270.

[6] Dwight, H. B. (1961). *Tables of Integrals and Other Mathematical Data*, 4th edn. Macmillan, USA.

[7] Gnedenko, B. V. (1989). On estimation of the unknown parameters of distributions from a random number of independent observations. *Proceedings of Tbilisi Math. Inst., AN GSSR*, **92**, 146–150 [in Russian].

[8] Gnedenko, B. V. (1998). *Theory of Probability*, 6th edn. CRC Press, USA.

[9] Gnedenko, B. V. & Fahim, G. (1969). On a transfer theorem. *Soviet Math. Dokl.*, **10**, 769–772.

[10] Gnedenko, B. V. & Korolev V. Y. (1996). *Random Summation. Limit Theorems and Applications*, CRC Press, USA.

[11] Kruglov, V. M. & Korolev V. Y. (1990). *Limit Theorems for Random Sums*, Moscow University Press, Moscow [in Russian].

[12] Lyamin, O. O. (2011). On the rate of convergence of the distributions of certain statistics to the Laplace and Student distributions. *Moscow University Computational Mathematics and Cybernetics*, **35**(1), 37–46.

[13] Shiryaev, A. N. (1995). *Probability*, 2nd edn. Springer, USA.

Annales Mathematicae et Informaticae
39 *(2012) pp. 29–44*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Random walk on half-plane half-comb structure

Endre Csáki[a*], **Miklós Csörgő**[b†], **Antónia Földes**[c‡]

Pál Révész[d*]

[a]Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Budapest
`csaki.endre@renyi.mta.hu`

[b]School of Mathematics and Statistics, Carleton University, Ottawa, Ontario, Canada
`mcsorgo@math.carleton.ca`

[c]Department of Mathematics, College of Staten Island, CUNY, New York, U.S.A.
`Antonia.Foldes@csi.cuny.edu`

[d]Institut für Statistik und Wahrscheinlichkeitstheorie, Technische Universität Wien, Austria
`revesz.paul@renyi.mta.hu`

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

We study limiting properties of a random walk on the plane, when we have a square lattice on the upper half-plane and a comb structure on the lower half-plane, i.e., horizontal lines below the x-axis are removed. We give strong approximations for the components with random time changed Wiener processes. As consequences, limiting distributions and some laws of the iterated logarithm are presented. Finally, a formula is given for the probability that the random walk returns to the origin in $2N$ steps.

Keywords: Anisotropic random walk; Strong approximation; Wiener process; Local time; Laws of the iterated logarithm;

MSC: primary 60F17, 60G50, 60J65; secondary 60F15, 60J10

*Research supported by the Hungarian National Foundation for Scientific Research.
†Research supported by an NSERC Canada Discovery Grant at Carleton University.
‡Research supported by a PSC CUNY Grant, No. 64086-0042.

1. Introduction and main results

The properties of a simple symmetric random walk on the square lattice $\mathbb{Z}^2$ have been extensively investigated in the literature since Dvoretzky and Erdős (1951), and Erdős and Taylor (1960). For these and further results we refer to Révész (2005).

Subsequent investigations concern random walks on other structures of the plane. For example, a simple random walk on the 2-dimensional comb lattice that is obtained from $\mathbb{Z}^2$ by removing all horizontal lines off the x-axis was studied by Weiss and Havlin (1986), Bertacchi and Zucca (2003), Bertacchi (2006), Csáki et al. (2009, 2011).

These are particular cases of the so-called anisotropic random walk on the plane. The general case is given by the transition probabilities

$$\begin{aligned}&\mathbf{P}(\mathbf{C}(N+1)=(k+1,j)|\mathbf{C}(N)=(k,j))\\&=\mathbf{P}(\mathbf{C}(N+1)=(k-1,j)|\mathbf{C}(N)=(k,j))=\frac{1}{2}-p_j,\end{aligned}$$

$$\begin{aligned}&\mathbf{P}(\mathbf{C}(N+1)=(k,j+1)|\mathbf{C}(N)=(k,j))\\&=\mathbf{P}(\mathbf{C}(N+1)=(k,j-1)|\mathbf{C}(N)=(k,j))=p_j,\end{aligned}$$

for $(k,j)\in\mathbb{Z}^2$, $N=0,1,2,\ldots$ with $0<p_j\le 1/2$ and $\min_{j\in\mathbb{Z}} p_j<1/2$. See Seshadri et al. (1979), Silver et al. (1977), Heyde (1982) and Heyde et al. (1982). The simple symmetric random walk corresponds to the case $p_j=1/4$, $j=0,\pm1,\pm2,\ldots$, while $p_0=1/4$, $p_j=1/2$, $j=\pm1,\pm2,\ldots$ defines random walk on the comb.

In this paper we combine the simple symmetric random walk with random walk on a comb, when $p_j=1/4$, $j=0,1,2,\ldots$ and $p_j=1/2$, $j=-1,-2,\ldots$, i.e., we have a square lattice on the upper half-plane, and a comb structure on the lower half-plane. We call this model Half-Plane Half-Comb (HPHC) and denote the random walk on it by $\mathbf{C}(N)=(C_1(N),C_2(N))$, $N=0,1,2,\ldots$.

For the second component of the HPHC walk a theorem of Heyde et al. (1982) gives in this particular case, the following strong limit theorem.

Theorem A. *On an appropriate probability space one can construct a sequence* $C_2^{(N)}(\cdot)$ *and a process* $Y(\cdot)$ *such that*

$$\lim_{N\to\infty}\sup_{0\le t\le M}\left|\frac{C_2^{(N)}([Nt])}{\sqrt{N}}-Y(t)\right|=0\quad a.s.,$$

where $Y(\cdot)$ *is an oscillating Brownian motion (Wiener process) and* $M>0$ *is arbitrary.*

Our first result is a strong approximation of both components of the random walk $\mathbf{C}(\cdot)$ by certain time-changed Wiener processes (Brownian motions) with rates of convergence. Before stating it, we need some definitions. Assume that we have two independent standard Wiener processes $W_1(t), W_2(t),\ t \geq 0$, and consider

$$\alpha_2(t) := \int_0^t I\{W_2(s) \geq 0\}\, ds,$$

i.e., the time spent by W_2 on the non-negative side during the interval $[0,t]$. The process $\gamma_2(t) := \alpha_2(t) + t$ is strictly increasing, hence we can define its inverse: $\beta_2(t) := (\gamma_2(t))^{-1}$. Observe that the processes $\alpha_2(t)$, $\beta_2(t)$ and $\gamma_2(t)$ are defined in terms of $W_2(t)$ so they are independent from $W_1(t)$. Moreover, it can be seen that $0 \leq \alpha_2(t) \leq t$, and $t/2 \leq \beta_2(t) \leq t$.

Theorem 1.1. *On an appropriate probability space for the HPHC random walk* $\{\mathbf{C}(N) = (C_1(N), C_2(N)); N = 0,1,2,\ldots\}$ *with* $p_j = 1/4,\ j = 0,1,2,\ldots,\ p_j = 1/2,\ j = -1,-2,\ldots$ *one can construct two independent standard Wiener processes* $\{W_1(t);\ t \geq 0\}$, $\{W_2(t);\ t \geq 0\}$ *such that, as* $N \to \infty$, *we have with any* $\varepsilon > 0$

$$|C_1(N) - W_1(N - \beta_2(N))| + |C_2(N) - W_2(\beta_2(N))| = O(N^{3/8+\varepsilon}) \quad a.s.$$

We note that the process $W_2(\beta_2(t))$ is identical with $Y(t)$ of Theorem A, i.e., an oscillating Brownian motion. It is a diffusion with speed measure (see Heyde et al., 1982)

$$m(dy) = \begin{cases} 4\,dy & \text{for} \quad y \geq 0, \\ 2\,dy & \text{for} \quad y < 0. \end{cases}$$

For more details on oscillating Brownian motion we refer to Keilson and Wellner (1978).

2. Preliminaries

First we want to redefine our walk $\mathbf{C}(\cdot)$ as follows: On a suitable probability space consider two independent simple symmetric (one-dimensional) random walks $S_1(\cdot)$, and $S_2(\cdot)$. We may assume that on the same probability space we have a sequence of independent geometric random variables $\{G_i,\ i = 1,2,\ldots\}$, independent from $S_1(\cdot), S_2(\cdot)$, with distribution

$$\mathbf{P}(G_i = k) = \frac{1}{2^{k+1}},\ k = 0,1,2,\ldots$$

Now horizontal steps will be taken consecutively according to $S_1(\cdot)$, and vertical steps consecutively according to $S_2(\cdot)$ in the following way. Start from $(0,0)$, take G_1 horizontal steps (possibly $G_1 = 0$) according to $S_1(\cdot)$, then take 1 vertical step. If this arrives to the upper half-plane ($S_2(1) = 1$), then take G_2 horizontal steps. If, however, the first vertical step is on the negative direction ($S_2(1) = -1$), then continue with another vertical step, and so on. In general, if the random walk is on the upper half-plane ($y \geq 0$) after a vertical step, then take a random number of horizontal steps according to the next (so far) unused G_j, independently from the previous steps. On the other hand, if the random walk is on the lower half-plane ($y < 0$) then continue with vertical steps according to $S_2(\cdot)$ until it reaches the x-axis, and so on.

Now we define the local times of a random walk and a Wiener process. Let $\{S(n);\, n = 0, 1, \ldots\}$ be a simple symmetric random walk on the line, i.e., $S(0) = 0$, $S(n) = X_1 + \ldots + X_n$, where $\{X_1, X_2, \ldots\}$ are i.i.d. random variables with $\mathbf{P}(X_i = 1) = \mathbf{P}(X_i = -1) = 1/2$. The local time is defined by

$$\xi(x, n) := \sum_{i=0}^{n} I\{S(i) = x\}, \quad x \in \mathbb{Z},\ n = 0, 1, \ldots,$$

where $I\{\cdot\}$ is the indicator function. The local time $\eta(x, t)$ of a Wiener process $W(\cdot)$ is defined via

$$\int_A \eta(x, t)\, dx = \lambda\{s :\, 0 \leq s \leq t,\, W(s) \in A\}$$

for any $x \in \mathbb{R}$, $t \geq 0$, where $A \subset \mathbb{R}$ is any Borel set and λ is the Lebesgue measure.

Now we state some results needed to prove our Theorem 1.1. First we quote a result of Révész (1981), that amounts to the first simultaneous strong approximation of a simple symmetric random walk and that of its local time process on the integer lattice $\mathbb{Z}$.

Lemma A. *On an appropriate probability space for a simple symmetric random walk* $\{S(n);\, n = 0, 1, 2, \ldots\}$ *with local time* $\{\xi(x, n);\, x = 0, \pm 1, \pm 2, \ldots;\, n = 0, 1, 2, \ldots\}$ *one can construct a standard Wiener process* $\{W(t);\, t \geq 0\}$ *with local time process* $\{\eta(x, t);\, x \in \mathbb{R};\, t \geq 0\}$ *such that, as* $n \to \infty$, *we have for any* $\varepsilon > 0$

$$S(n) - W(n) = O(n^{1/4+\varepsilon}) \quad a.s.$$

and

$$\sup_{x \in \mathbb{Z}} |\xi(x, n) - \eta(x, n)| = O(n^{1/4+\varepsilon}) \quad a.s.,$$

simultaneously.

The following strong invariance principle is given in Horváth (1998).

Lemma B. *On the probability space of* Lemma A, *for any* $\varepsilon > 0$, *as* $n \to \infty$, *we have*

$$\left| \sum_{k=0}^{n} g(S(k)) - \int_0^n g(W(t))\, dt \right|$$

$$= \left| \sum_{j=-\infty}^{\infty} g(j)\xi(j,n) - \int_{-\infty}^{\infty} g(x)\eta(x,n)\,dx \right| = O(n^{a/2+3/4+\varepsilon}) \quad a.s.,$$

where $g(t) \geq 0$, $t \in \mathbb{R}$ *is a function such that for* $k \in \mathbb{Z}$ *we have* $g(t) = g(k)$, $k \leq t < k+1$ *and*

$$g(t) \leq C(|t|^a + 1)$$

for some $C > 0$ *and* $0 \leq a$.

For $n \geq 1$ let

$$A(n) := \sum_{i=0}^{n-1} I\{S(i) \geq 0\} = \sum_{j=0}^{\infty} \xi(j, n-1), \tag{2.1}$$

i.e., the time spent by the random walk $S(\cdot)$ on the non-negative side during the first $n-1$ steps. Let furthermore

$$\alpha(t) = \int_0^t I\{W(s) \geq 0\}\,ds = \int_0^{\infty} \eta(x,t)\,dx.$$

Applying Lemma B with $g(t) = I\{t \geq 0\}$, $a = 0$, and taking into account that $A(n+1) - A(n) \leq 1$, we have the following consequence.

Corollary A. *On the probability space of* Lemma A, *for any* $\varepsilon > 0$, *as* $n \to \infty$, *we have almost surely*

$$A(n) - \alpha(n) = O(n^{3/4+\varepsilon}).$$

Concerning the increments of the Wiener process we quote the following result from Csörgő and Révész (1981).

Lemma C. *Let* $0 < a_T \leq T$ *be a non-decreasing function of* T. *Then, as* $T \to \infty$, *we have almost surely*

$$\sup_{0 \leq t \leq T - a_T} \sup_{s \leq a_T} |W(t+s) - W(t)| = O(a_T^{1/2}(\log(T/a_T) + \log\log T)).$$

Put

$$f_v(z,y)\,dz\,dy := \mathbf{P}(W(v) \in dz,\ \alpha(v) \in dy),$$

the joint density function of $(W(v), \alpha(v))$. For $f_v(z,y)$ the following two formulas are known in the literature. The first one is due to Karatzas and Shreve (1984), (see also Borodin and Salminen, 1996), the second one is given in Nikitin and Orsingher (2000).

Lemma D. *For* $0 \leq y \leq v$ *we have*

$$f_v(z,y) = \begin{cases} \int_0^\infty \frac{s(s+z)}{\pi y^{3/2}(v-y)^{3/2}} \exp\left(-\frac{s^2}{2(v-y)} - \frac{(s+z)^2}{2y}\right) ds, & z \geq 0, \\ \\ \int_0^\infty \frac{s(s-z)}{\pi y^{3/2}(v-y)^{3/2}} \exp\left(-\frac{s^2}{2y} - \frac{(s-z)^2}{2(v-y)}\right) ds, & z < 0, \end{cases}$$

$$f_v(z,y)=\begin{cases}\int_{v-y}^{v}\frac{z\exp\left(-\frac{z^2}{2(v-s)}\right)}{2\pi s^{3/2}(v-s)^{3/2}}\,ds, & z\geq 0,\\ \\ \int_{y}^{v}\frac{|z|\exp\left(-\frac{z^2}{2(v-s)}\right)}{2\pi s^{3/2}(v-s)^{3/2}}\,ds, & z<0.\end{cases}$$

3. Proof of Theorem 1.1

Start with the construction of HPHC given in Section 2. Let H_N and V_N, the number of horizontal and vertical steps, respectively of the two-dimensional random walk $\mathbf{C}(\cdot)$ during the first N steps, i.e., $H_N+V_N=N$. Consider the two independent simple symmetric random walks $S_1(\cdot)$ and $S_2(\cdot)$ and the sequence of i.i.d. geometric random variables, which is indepedent from these two walks, as it was described in Section 2. Define $A_2(n)$ as in (2.1), in terms of $S_2(\cdot)$, i.e., $A_2(n)=\sum_{j=0}^{\infty}\xi_2(j,n-1)$, where $\xi_2(\cdot,\cdot)$ is the local time of $S_2(\cdot)$. Assume furthermore that on the same probability space we have strong approximations of (S_1,ξ_1) by (W_1,η_1) and that of (S_2,ξ_2) by (W_2,η_2) as described in Lemma A, where W_1 and W_2 are two independent Wiener processes on the line, and η_1 and η_2 are their respective local times.

Then, with $V_N=n$,

$$\sum_{j=1}^{A_2(n)}G_j\leq H_N\leq\sum_{j=1}^{A_2(n)+1}G_j$$

and since one term in the above sum is $O(\log N)$ a.s., and $EG_j=1$, with finite variance, we have

$$H_N=A_2(n)+O(A_2(n)^{1/2+\varepsilon})=A_2(n)+O(N^{1/2+\varepsilon})\quad a.s.,$$

as $N\to\infty$. Hence, using Corollary A, we have almost surely, as $N\to\infty$,

$$\alpha_2(n)+n=A_2(n)+O(N^{3/4+\varepsilon})+V_N=H_N+V_N+O(N^{3/4+\varepsilon})=N+O(N^{3/4+\varepsilon}).$$

Consequently,

$$V_N=n=\beta_2(\alpha_2(n)+n)=\beta_2(N+O(N^{3/4+\varepsilon}))=\beta_2(N)+O(N^{3/4+\varepsilon})$$

and

$$H_N=N-\beta_2(N)+O(N^{3/4+\varepsilon}).$$

Using Lemma C, this gives almost surely, as $N\to\infty$,

$$C_1(N)=S_1(H_N)=W_1(H_N)+O(H_N^{1/4+\varepsilon})=W_1(N-\beta_2(N))+O(N^{3/8+\varepsilon})$$

and

$$C_2(N)=S_2(V_N)=W_2(\beta_2(N))+O(N^{3/8+\varepsilon}),$$

proving Theorem 1.1. □

Remark 3.1. In the above argument we used the fact, that for $u, v > 0$, $\beta(u+v) - \beta(u) \le v$. To see this recall that $\beta(t)$ is the inverse of $\gamma(t) = \alpha(t) + t$. Hence

$$v = \gamma(\beta(u+v)) - \gamma(\beta(u)) = \alpha(\beta(u+v)) + \beta(u+v) - \alpha(\beta(u)) - \beta(u) \ge \beta(u+v) - \beta(u),$$

as $\alpha(t)$ is nondecreasing.

4. Limiting densities and consequences

First we give an integral expression for the joint density of the vector $(W_1(t - \beta(t)), W_2(\beta(t)))$, using Lemma D. Here, and throughout this section, $\beta(t)$ stands for $\beta_2(t)$, hence it is independent from W_1. The joint density of $W_1(t - \beta(t))$, $W_2(\beta(t))$, $\beta(t)$ is given by

$$\begin{aligned}&\mathbf{P}(W_1(t - \beta(t)) \in du,\, W_2(\beta(t)) \in dz,\, \beta(t) \in dv)\\ &= \frac{1}{\sqrt{2\pi(t-v)}} \exp\left(-\frac{u^2}{2(t-v)}\right) f_v(z, t-v)\, du\, dz\, dv.\end{aligned}$$

From this we get

Lemma 4.1.

$$\begin{aligned}&g_t(u,z)\, du\, dz := \mathbf{P}(W_1(t - \beta(t)) \in du,\, W_2(\beta(t)) \in dz)\\ &= \left(\int_{t/2}^{t} \frac{1}{\sqrt{2\pi(t-v)}} \exp\left(-\frac{u^2}{2(t-v)}\right) f_v(z, t-v)\, dv\right) du\, dz.\end{aligned}$$

The marginal density of $W_1(t - \beta(t))$ is given by

Lemma 4.2.

$$g_t^{(1)}(u)\, du := \mathbf{P}(W_1(t - \beta(t)) \in du) = \frac{1}{\pi\sqrt{2\pi t}} \exp\left(-\frac{u^2}{2t}\right) K_0\left(\frac{u^2}{2t}\right) du,$$

where $K_0(\cdot)$ is the modified Bessel function of the second kind.

Proof.

$$\mathbf{P}(W_1(t - \beta(t)) \in du) = \int_{t/2}^{t} \mathbf{P}(W_1(t - v) \in du,\, \beta(t) \in dv)$$

$$= \left(\int_{t/2}^{t} \frac{1}{\sqrt{2\pi(t-v)}} \exp\left(-\frac{u^2}{2(t-v)}\right) \frac{1}{\pi\sqrt{(t-v)(2v-t)}}\, dv\right) du$$

$$= \frac{1}{\pi\sqrt{2\pi}} \exp\left(-\frac{u^2}{t}\right) \int_0^\infty \frac{1}{\sqrt{y^2 t + y u^2}} e^{-y}\, dy\, du = \frac{1}{\pi\sqrt{2\pi t}} \exp\left(-\frac{u^2}{2t}\right) K_0\left(\frac{u^2}{2t}\right) du,$$

where the substitution

$$y = u^2\left(\frac{1}{2(t-v)} - \frac{1}{t}\right)$$

was made and the formula

$$\int_0^\infty \frac{e^{-px}\, dx}{\sqrt{x(x+a)}} = e^{ap/2} K_0\left(\frac{ap}{2}\right)$$

was used (see Gradsteyn and Ryzhik, 1994, 3.364.3). □

For the marginal density of $W_2(\beta(t))$ as follows, we refer to Heyde et al. (1982).

Lemma E.

$$g_t^{(2)}(z)\, dz = \mathbf{P}(W_2(\beta(t)) \in dz) = \begin{cases} 2\sqrt{\frac{2}{\pi t}}(\sqrt{2}-1)e^{-z^2/t}\, dz, & z \geq 0 \\ \sqrt{\frac{2}{\pi t}}(\sqrt{2}-1)e^{-z^2/2t}\, dz, & z < 0. \end{cases}$$

As a consequence of these Lemmas, we now obtain the joint and marginal limiting distributions of the HPHC random walk.

Corollary 4.3.

$$\lim_{N\to\infty} \mathbf{P}\left(\frac{C_1(N)}{\sqrt{N}} \leq x,\ \frac{C_2(N)}{\sqrt{N}} \leq y\right) = \int_{-\infty}^{x} \int_{-\infty}^{y} g_1(u,z)\, du\, dz,$$

$$\lim_{N\to\infty} \mathbf{P}\left(\frac{C_1(N)}{\sqrt{N}} \leq x\right) = \int_{-\infty}^{x} g_1^{(1)}(u)\, du,$$

$$\lim_{N\to\infty} \mathbf{P}\left(\frac{C_2(N)}{\sqrt{N}} \leq y\right) = \int_{-\infty}^{y} g_1^{(2)}(z)\, dz.$$

Corollary 4.4. *The following laws of the iterated logarithm hold.*

(i) $\limsup_{t\to\infty} \frac{W_1(t-\beta(t))}{\sqrt{t \log\log t}} = \limsup_{N\to\infty} \frac{C_1(N)}{\sqrt{N \log\log N}} = 1 \quad a.s.,$

(ii) $\liminf_{t\to\infty} \frac{W_1(t-\beta(t))}{\sqrt{t \log\log t}} = \liminf_{N\to\infty} \frac{C_1(N)}{\sqrt{N \log\log N}} = -1 \quad a.s.,$

(iii) $\limsup_{t\to\infty} \frac{W_2(\beta(t))}{\sqrt{t \log\log t}} = \limsup_{N\to\infty} \frac{C_2(N)}{\sqrt{N \log\log N}} = 1 \quad a.s.,$

(iv) $\liminf_{t\to\infty} \frac{W_2(\beta(t))}{\sqrt{t \log\log t}} = \liminf_{N\to\infty} \frac{C_2(N)}{\sqrt{N \log\log N}} = -\sqrt{2} \quad a.s.$

Proof. We give short proofs in the case of W_1 and W_2. The results for C_1 and C_2 then follow from Theorem 1.1. In the proof we repeatedly use the inequality

$$\frac{t}{2} \le \beta(t) \le t.$$

Proof of (i) and (ii). By the law of the iterated logarithm for W_1 we have for all large enough t

$$\begin{aligned} W_1(t-\beta(t)) &\le (1+\varepsilon)(2(t-\beta(t))\log\log(t-\beta(t)))^{1/2} \\ &\le (1+\varepsilon)(t\log\log t)^{1/2}, \end{aligned}$$

which gives an upper bound in (i).

To give a lower bound in (i), for any sufficiently small $\delta > 0$ define the events

$$A_n = \{W_1(u_n) \ge (1-\delta)(2u_n \log\log u_n)^{1/2}\}, \quad B_n = \{\alpha(u_n(1+\delta)) > u_n\},$$

$n = 1, 2, \ldots$. Then, with some sequence $\{u_n\}$ ($u_n = a^n$ with sufficiently large a will do), we have

$$\mathbf{P}(A_n \ i.o.) = 1, \qquad \mathbf{P}(B_n) > c > 0.$$

It follows from Klass (1976) that

$$\mathbf{P}(A_n B_n \ i.o.) \ge c > 0.$$

By the 0-1 law this probability is equal to 1. Let t_n be defined by

$$u_n = t_n - \beta(t_n) = \alpha(\beta(t_n)).$$

Since

$$B_n = \{\alpha(u_n(1+\delta)) > \alpha(\beta(t_n))\},$$

B_n implies

$$u_n \ge \frac{\beta(t_n)}{1+\delta} \ge \frac{t_n}{2(1+\delta)}.$$

Hence $A_n B_n$ implies

$$W_1(t-\beta(t_n)) \ge (1-\delta)\left(\frac{t_n \log\log t_n}{1+\delta}\right)^{1/2}.$$

Since $\delta > 0$ is arbitrary, this gives a lower bound in (i).

The proof of (ii) follows by symmetry.

Proof of (iii). We have infinitely often with probability 1

$$W_2(\beta(t)) \geq (1-\varepsilon)(2\beta(t)\log\log t)^{1/2} \geq (1-\varepsilon)(t\log\log t)^{1/2},$$

giving a lower bound in (iii).

To give an upper bound, we use the formula for the distribution of the supremum of $W_2(\beta(t))$ given in Corollary 2 of Keilson and Wellner (1978), which in our case is equivalent to

$$\mathbf{P}(\sup_{0\leq s\leq t} W_2(\beta(s)) > y)$$
$$= \frac{2\sqrt{2}}{1+\sqrt{2}}\sum_{k=0}^{\infty}\left(\frac{1-\sqrt{2}}{1+\sqrt{2}}\right)^k\left(1-\Phi\left(\frac{(2k+1)y\sqrt{2}}{\sqrt{t}}\right)\right).$$

From this it is easy to give the estimation

$$\mathbf{P}(\sup_{0\leq s\leq t} W_2(\beta(s)) > y) \leq c\exp\left(-\frac{y^2}{t}\right)$$

with some constant c, from which the upper estimation in (iii) follows by the usual procedure.

Proof of (iv). The lower estimation is easy. Namely we have

$$W_2(\beta(t)) \geq -(1+\varepsilon)(2\beta(t)\log\log\beta(t))^{1/2} \geq -(1+\varepsilon)(2t\log\log t)^{1/2}.$$

It remains to prove an upper estimation in (iv). By the law of the iterated logarithm for W_2

$$W_2(v) \leq -((2-\varepsilon)v\log\log v)^{1/2} \tag{4.1}$$

almost surely for infinitely many v tending to infinity. Let $\zeta(v)$ be the last zero of W_2 before v, i.e.,

$$\zeta(v) = \max\{u \leq v : W_2(u) = 0\}.$$

By Theorem 1 of Csáki and Grill (1988), for large v satisfying (4.1) we have $\zeta(v) \leq \varepsilon v$, and hence also $\alpha(v) \leq \zeta(v) \leq \varepsilon v$. Now put $v = \beta(t)$, i.e., $\alpha(v)+v = t \leq (1+\varepsilon)v$, from which $v = \beta(t) \geq t/(1+\varepsilon)$. Hence

$$W_2(v) = W_2(\beta(t)) \leq -\left(\frac{(2-\varepsilon)t\log\log t}{1+\varepsilon}\right)^{1/2}.$$

Since $\varepsilon > 0$ is arbitrary, this gives an upper bound in (iv).

This completes the proof of Corollary 4.4. □

Some related distributions can also be determined. For example, we can obtain the following result for the supremum of the first component.

Lemma 4.5.

$$\mathbf{P}(\sup_{0\le s\le t}|W_1(s-\beta(s))|\le u)$$
$$=\frac{4}{\pi}\sum_{j=0}^{\infty}\frac{(-1)^j}{2j+1}\exp\left(-\frac{(2j+1)^2\pi^2 t}{32u^2}\right)I_0\left(\frac{(2j+1)^2\pi^2 t}{32u^2}\right),$$

where I_0 is the modified Bessel function of the first kind given by

$$I_0(z)=\sum_{k=0}^{\infty}\frac{z^{2k}}{4^k(k!)^2}.$$

Proof.

$$\mathbf{P}(\sup_{0\le s\le t}|W_1(s-\beta(s))|\le u)=\int_{t/2}^{t}\mathbf{P}(\sup_{z\le t-v}|W_1(z)|\le u)\mathbf{P}(\beta(t)\in dv)$$

$$=\int_{t/2}^{t}\frac{4}{\pi}\sum_{j=0}^{\infty}\frac{(-1)^j}{2j+1}\exp\left(-\frac{(2j+1)^2\pi^2(t-v)}{8u^2}\right)\frac{1}{\pi\sqrt{(t-v)(2v-t)}}\,dv,$$

and using 3.384.2 and 9.235.1 of Gradsteyn and Ryzhik (1994), and some calculations, we obtain Lemma 4.5. □

Corollary 4.6.

$$\lim_{N\to\infty}\mathbf{P}\left(\frac{\sup_{0\le k\le N}|C_1(k)|}{\sqrt{N}}\le u\right)$$
$$=\frac{4}{\pi}\sum_{j=0}^{\infty}\frac{(-1)^j}{2j+1}\exp\left(-\frac{(2j+1)^2\pi^2}{32u^2}\right)I_0\left(\frac{(2j+1)^2\pi^2}{32u^2}\right),$$

5. Return probabilities

We give the probability that the random walk returns to the origin in 2N steps.

Theorem 5.1. *For $N\ge 1$*

$$\mathbf{P}(\mathbf{C}(2N)=(0,0))$$
$$=\frac{1}{2^{4N}}\left(\binom{2N}{N}+\sum_{n=1}^{N}\sum_{k=1}^{n}\sum_{j=1}^{k}\binom{2N-2n}{N-n}a_j a_{n+1-j}\left(b(n,2k)+b(n,2k-1)\right)\right),$$

where for $i=1,2,\dots,\quad n=1,2,\dots,N,\quad \ell=1,2,\dots,$

$$a_i=\frac{1}{2i-1}\binom{2i-1}{i},\qquad b(n,\ell)=\binom{2N-2n+\ell}{\ell}2^{2n-\ell}.$$

Proof. For $n \geq 1$ let

$$P(2n,r) = \mathbf{P}(S_2(2n) = 0, A_2(2n) = r), \quad Q(2n,r) = 2^{2n}P(2n,r).$$

Obviously $P(2n,r) = 0$ if $r > 2n$ or $r \leq 0$. Furthermore it is easy to see, that

$$P(2n,1) = \frac{1}{2n-1}\binom{2n-1}{n}\frac{1}{2^{2n}} = \frac{1}{2(2n-1)}\binom{2n}{n}\frac{1}{2^{2n}},$$
$$P(2n,2n) = \frac{1}{n+1}\binom{2n}{n}\frac{1}{2^{2n}}.$$

For $n = 1, 2\ldots, \quad r = 2, 3, \ldots 2n$, we have the following recursion for $P(2n,r)$.

$$\begin{aligned}
P(2n,r) &= \sum_{i=1}^{n}\mathbf{P}(S(1) < 0, \ldots, S(2i-1) < 0, S(2i) = 0)P(2n-2i, r-1)\\
&\quad + \sum_{i=1}^{n}\mathbf{P}(S(1) > 0, \ldots, S(2i-1) > 0, S(2i) = 0)P(2n-2i, r-2i)\\
&= \sum_{i=1}^{n}\frac{1}{2i-1}\binom{2i-1}{i}\frac{1}{2^{2i}}P(2n-2i, r-1)\\
&\quad + \sum_{i=1}^{n}\frac{1}{2i-1}\binom{2i-1}{i}\frac{1}{2^{2i}}P(2n-2i, r-2i),
\end{aligned}$$

where we define $P(0,0) = 1$. □

Now we need the following lemma.

Lemma 5.2. *For* $n = 1, 2, \ldots, \quad k = 1, 2, \ldots, n,$ *we have*

$$Q(2n, 2k-1) = Q(2n, 2k) \tag{5.1}$$

and

$$\begin{aligned}
Q(2n,2k) &= \sum_{j=1}^{k} a_j a_{n+1-j}\\
&= \sum_{j=1}^{k}\frac{1}{2j-1}\binom{2j-1}{j}\frac{1}{2n+1-2j}\binom{2n+1-2j}{n+1-j}.
\end{aligned} \tag{5.2}$$

Remark 5.3. It is obvious that

$$Q(2n+2, 1) = Q(2n, 2n).$$

Furthermore, we can conveniently reformulate the second statement as

$$Q(2n, 2k) = Q(2n, 2k-2) + a_k a_{n+1-k}.$$

In particular

$$Q(2n, 2n) = Q(2n+2, 2) = a_{n+1}.$$

Proof. We prove Lemma 5.2 with simultaneous induction. Clearly, for $n = 1$ and $k = 1$ both of our statements are correct. We suppose that (5.1) and (5.2) hold for all $m < n$ and $j \le 2k - 2$. First we prove (5.1). By our recursion formula and the induction hypothesis we have

$$Q(2n, 2k-1) = \sum_{j=1}^{n-k+1} a_j Q(2n-2j, 2k-2) + \sum_{j=1}^{k-1} a_j\, Q(2n-2j, 2k-2j-1)$$
$$= \sum_{j=1}^{n-k+1} a_j Q(2n-2j, 2k-2) + \sum_{j=1}^{k-1} a_j\, Q(2n-2j, 2k-2j).$$

Moreover,

$$Q(2n, 2k) = \sum_{j=1}^{n-k} a_j Q(2n-2j, 2k-1) + \sum_{j=1}^{k-1} a_j\, Q(2n-2j, 2k-2j)$$
$$= \sum_{j=1}^{n-k} a_j Q(2n-2j, 2k) + \sum_{j=1}^{k-1} a_j\, Q(2n-2j, 2k-2j).$$

Then

$$\begin{aligned}
&Q(2n, 2k) - Q(2n, 2k-1) \\
&= \sum_{j=1}^{n-k} a_j\, (Q(2n-2j, 2k) - Q(2n-2j, 2k-2)) - a_{n-k+1}\, Q(2k-2, 2k-2) \\
&= \sum_{j=1}^{n-k} a_j a_k a_{n+1-k-j} - a_{n-k+1}\, a_k = a_k \sum_{j=1}^{n-k} a_j a_{n+1-k-j} - a_{n-k+1}\, a_k \\
&= a_k Q(2n-2k, 2n-2k) - a_{n-k+1}\, a_k = a_k\, a_{n-k+1} - a_k\, a_{n-k+1} = 0,
\end{aligned}$$

which proves (5.1). To prove (5.2), consider

$$\begin{aligned}
Q(2n, 2k) - Q(2n, 2k-2) &= \sum_{j=1}^{n-k} a_j Q(2n-2j, 2k) + \sum_{j=1}^{k-1} a_j Q(2n-2j, 2k-2j) \\
&- \left(\sum_{j=1}^{n+1-k} a_j Q(2n-2j, 2k-2) + \sum_{j=1}^{k-2} a_j Q(2n-2j, 2k-2-2j) \right) \\
&= \sum_{j=1}^{n-k} a_j\, (Q(2n-2j, 2k) - Q(2n-2j, 2k-2)) - a_{n+1-k} Q(2k-2, 2k-2) \\
&+ \sum_{j=1}^{k-2} a_j\, (Q(2n-2j, 2k-2j) - Q(2n-2j, 2k-2-2j)) + a_{k-1} Q(2n-2k+2, 2)
\end{aligned}$$

$$= \sum_{j=1}^{n-k} a_j a_k a_{n-k+1-j} - a_{n+1-k} a_k + \sum_{j=1}^{k-2} a_j a_{k-j} a_{n+1-k} + a_{k-1}\, a_{n-k+1}$$
$$= a_k \sum_{j=1}^{n-k} a_j a_{n-k+1-j} - a_{n+1-k} a_k + a_{n+1-k} \sum_{j=1}^{k-2} a_j a_{k-j} + a_{k-1} a_{n-k+1}$$
$$= a_k Q(2n-2k, 2n-2k) - a_{n+1-k} a_k + a_{n+1-k} Q(2k-2, 2k-4) + a_{k-1} a_{n-k+1}$$
$$= a_k a_{n-k+1} - a_{n+1-k} a_k + a_{n+1-k} \left(Q(2k-2, 2k-2) - a_1 a_{k-1}\right) + a_{k-1} a_{n-k+1}$$
$$= a_{n+1-k} a_k - a_{n+1-k} a_{k-1} + a_{k-1} a_{n+1-k} = a_k a_{n+1-k},$$

proving (5.2). □

Returning to the proof of Theorem 5.1, let V_N and H_N be the number of vertical and horizontal steps, resp. as in the proof of Theorem 1.1. We have

$$\mathbf{P}(\mathbf{C}(2N) = (0,0)) = \mathbf{P}(H_{2N} = 2N, S_1(2N) = 0)$$
$$+ \sum_{n=1}^{N} \sum_{r=1}^{2n} \mathbf{P}(H_{2N} = 2N - 2n | S_2(2n) = 0, A_2(2n) = r)$$
$$\times P(2n, r) \mathbf{P}(S_1(2N - 2n) = 0).$$

For $n \geq 1$ we show that

$$\mathbf{P}(H_{2N} = 2N - 2n | S_2(2n) = 0, A_2(2n) = r) = \binom{2N - 2n + r}{r} \frac{1}{2^{2N-2n+r}}.$$

Under the condition $S_2(2n) = 0, A_2(2n) = r$, we have

$$H_{2N} = \sum_{i=1}^{r} G_i + G,$$

where G_i are i.i.d. geometric variables with

$$\mathbf{P}(G_i = k) = \frac{1}{2^{k+1}}, \quad k = 0, 1, \ldots$$

and G denotes the number of horizontal steps after the $2n$-th vertical step up to the total number of $2N$ steps. So

$$\mathbf{P}(H_{2N} = 2N - 2n | S_2(2n) = 0, A_2(2n) = r) = \sum_{k=0}^{2N-2n} \mathbf{P}\left(\sum_{i=1}^{r} G_i = k\right) \frac{1}{2^{2N-2n-k}}$$
$$= \sum_{k=0}^{2N-2n} \binom{k+r-1}{k} \frac{1}{2^{k+r}} \frac{1}{2^{2N-2n-k}} = \frac{1}{2^{2N-2n+r}} \sum_{k=0}^{2N-2n} \binom{k+r-1}{k}$$
$$= \binom{2N-2n+r}{r} \frac{1}{2^{2N-2n+r}}.$$

Hence we have

$$\begin{aligned}
&\mathbf{P}(\mathbf{C}(2N) = (0,0)) \\
&= \frac{1}{2^{4N}}\binom{2N}{N} + \sum_{n=1}^{N}\sum_{r=1}^{2n} P(2n,r)\binom{2N-2n}{N-n}\frac{1}{2^{2N-2n}}\binom{2N-2n+r}{r}\frac{1}{2^{2N-2n+r}} \\
&= \frac{1}{2^{4N}}\binom{2N}{N} + \sum_{n=1}^{N}\sum_{r=1}^{2n} Q(2n,r)\binom{2N-2n}{N-n}\frac{1}{2^{2N}}\binom{2N-2n+r}{r}\frac{1}{2^{2N-2n+r}}
\end{aligned}$$

and using Lemma 5.2 completes the proof of our Theorem 5.1.

References

[1] Bertacchi, D. (2006). Asymptotic behaviour of the simple random walk on the 2-dimensional comb. *Electron. J. Probab.* 11:1184–1203.

[2] Bertacchi, D. and Zucca, F. (2003). Uniform asymptotic estimates of transition probabilities on combs. *J. Aust. Math. Soc.* 75:325–353.

[3] Borodin, A.N. and Salminen, P. (1996). *Handbook of Brownian Motion – Facts and Formulae.* Birkhäuser, Boston.

[4] Csáki, E., Csörgő, M., Földes, A. and Révész, P. (2009). Strong limit theorems for a simple random walk on the 2-dimensional comb. *Electr. J. Probab.* 14:2371–2390.

[5] Csáki, E., Csörgő, M., Földes, A. and Révész, P. (2011). On the local time of random walk on the 2-dimensional comb. *Stoch. Process. Appl.* 121:1290–1314.

[6] Csáki, E. and Grill, K. (1988). On the large values of the Wiener process. *Stoch. Process. Appl.* 27:43–56.

[7] Csörgő, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics.* Academic Press, New York.

[8] Dvoretzky, A. and Erdős, P. (1951). Some problems on random walk in space. *Proc. Second Berkeley Symposium*, pp. 353–367.

[9] Erdős, P. and Taylor, S.J. (1960). Some problems concerning the structure of random walk paths. *Acta Math. Acad. Sci. Hungar.* 11:137–162.

[10] Gradshteyn, I.S. and Ryzhik, I.M. (1994). *Table of integrals, series, and products.* Academic Press, Boston, MA.

[11] Heyde, C.C. (1982). On the asymptotic behaviour of random walks on an anisotropic lattice. *J. Statist. Physics* 27:721–730.

[12] Heyde, C.C., Westcott, M. and Williams, R.J. (1982). The asymptotic behavior of a random walk on a dual-medium lattice. *J. Statist. Physics* 28:375–380.

[13] Horváth, L. (1998). Diffusion approximation for random walks on anisotropic lattices. *J. Appl. Probab.* 35:206–212.

[14] Karatzas, I. and Shreve, S.E. (1984). Trivariate density of Brownian motion, its local and occupation times, with application to stochastic control. *Ann. Probab.* 12:819–828.

[15] Keilson, J. and Wellner, J.A. (1978). Oscillating Brownian motion. *J. Appl. Probab.* 15:300–310.

[16] Klass, M. (1976). Toward a universal law of the iterated logarithm. I. *Z. Wahrsch. verw. Gebiete* 36:165–178.

[17] Nikitin, Y. and Orsingher, E. (2000). The intermediate arc-sine law. *Statist. Probab. Lett.* 49:119–125.

[18] Révész, P. (2005). *Random Walk in Random and Non-Random Environments*, 2nd ed. World Scientific, Singapore.

[19] Révész, P. (1981). Local time and invariance. *Lecture Notes in Math.* 861:128–145. Springer, New York.

[20] Seshadri, V., Lindenberg, K. and Shuler, K.E. (1979). Random Walks on Periodic and Random Lattices. II. Random Walk Properties via Generating Function Techniques. *J. Statist. Physics* 21:517–548.

[21] Silver, H., Shuler, K.E. and Lindenberg, K. (1977). Two-dimensional anisotropic random walks. In: *Statistical mechanics and statistical methods in theory and application (Proc. Sympos., Univ. Rochester, Rochester, N.Y., 1976)*, Plenum, New York, pp. 463–505.

[22] Weiss, G.H. and Havlin, S. (1986). Some properties of a random walk on a comb structure. *Physica A* 134:474–482.

Annales Mathematicae et Informaticae
39 *(2012) pp. 45–56*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Joint asymptotic normality of the kernel type density estimator for spatial observations

István Fazekas[a*], **Zsolt Karácsony**[b†], **Renáta Vas**[a‡]

[a]Faculty of Informatics, University of Debrecen, Hungary
fazekasi@inf.unideb.hu, vas.renata@inf.unideb.hu

[b]Department of Applied Mathematics, University of Miskolc, Hungary
matkzs@uni-miskolc.hu

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

The Central Limit Theorem is considered for m-dependent random fields. The random field is observed in a sequence of irregular domains. The sequence of domains is increasing and at the same time, the locations of the observations become more and more dense in the domains. The Central Limit Theorem is applied to obtain asymptotic normality of kernel type density estimators. It turns out that the covariance structure of the limiting normal distribution can be a combination of those of the continuous parametric and the discrete parametric results. Numerical evidence is presented.

Keywords: Asymptotic normality, central limit theorem, random field, kernel, infill-increasing setup

MSC: 60F05,62M30

*The work is supported by the Hungarian Scientific Research Fund under Grant No. OTKA T079128/2009.

†This author was supported by TAMOP-4.2.1.B-10/2/KONV-2010-0001 project with support by the European Union, co-financed by the European Social Fund.

‡The work is supported by the TÁMOP-4.2.2/B-10/1-2010-0024 project. The project is co-financed by the European Union and the European Social Fund.

1. Introduction

Consider a domain D in $\mathbb{R}^d$. We observe a random field $\xi(\cdot)$ in certain points of the domain D and we assume the following setup. Suppose that the random field $\xi(\cdot)$ is observed at finitely many locations i.e. at the elements $\mathbf{s}_{n1}, \dots, \mathbf{s}_{nn}$ lying in the sampling region $D_n \subset D$. Let $\mathcal{R}_n = \{\mathbf{s}_{n1}, \dots, \mathbf{s}_{nn}\}$ denote the n-th set of the locations of the observations. We shall use the notion of the mixed (or nearly infill or infill-increasing) domain sampling which means that the sampling region D_n increases and at the same time, the data sites $\{\mathbf{s}_{n1}, \dots, \mathbf{s}_{nn}\}$ fill in any given sub-region of D_n increasingly densely as $n \to \infty$. (Increasing domains means that $D_n \subseteq D_{n+1}$ and the size of D_n goes to infinity as $n \to \infty$.) This approach was studied e.g. by Lahiri [4], Lahiri, Kaiser, Cressie and Hsu [5], Fazekas and Chuprunov [2], Park, Kim, Park and Hwang [6] and Karácsony and Filzmoser [3]. It can be useful in geostatistics, environmental sciences etc.

To obtain asymptotic normality, we assume that the n-th set of observations is $\xi_n(\mathbf{s}_{n1}), \dots, \xi_n(\mathbf{s}_{nn})$, where $\xi_n(\cdot), n = 1, 2, \dots$ is a sequence of stationary random fields and $\xi_n(\cdot)$ is weakly dependent for any fixed n. For the sake of simplicity we suppose that $\xi_n(\cdot)$ is m-dependent. It is a restriction but it has an advantage namely that we can easily obtain a central limit theorem (CLT) for irregular domains. We mention that similar results can be obtained for mixing random fields as well (see e.g. Fazekas and Chuprunov [1], but there the domain is regular and the conditions are quite difficult to check). The main objective of Park, Kim, Park and Hwang [6] is to provide central limit theorems that could be applied easily in practice. In our paper we discuss some consequences of the results of Park, Kim, Park and Hwang [6].

The article is organized as follows. In Section 2, we introduce our notations and we recall the CLT for stationary random fields of Park, Kim, Park and Hwang [6]. In Section 3, we turn to the density estimator, we quote Theorem 3 of Park, Kim, Park and Hwang [6]. It states that under mild conditions the kernel type density estimator is asymptotic normal. In Section 4, we deal with the multidimensional extension of this theorem. Simulation evidence is presented here, too. The numerical examples show the unusual covariance structure of the limiting normal distribution. This covariance structure was first presented in Fazekas and Chuprunov [2]. That is, the asymptotic covariance of the kernel type density estimator for nearly infill sampling can be a combination of the covariances of the discrete and the continuous parameter models. Similar result is valid for the regression estimator (see Karácsony and Filzmoser [3]).

2. CLT for stationary random fields

Let us consider a zero mean strictly stationary random field $\{\xi(\mathbf{s}) : \mathbf{s} \in D\}$, $D \subseteq \mathbb{R}^d$. Here, the strict stationarity of the random field means that for any $\mathbf{s}_1, \dots, \mathbf{s}_k, \mathbf{t}$, the distribution of $(\xi(\mathbf{s}_1), \dots, \xi(\mathbf{s}_k))$ is the same as that of $(\xi(\mathbf{s}_1 + \mathbf{t}), \dots, \xi(\mathbf{s}_k + \mathbf{t}))$.

We assume that the random field $\xi(\cdot)$ is m-dependent. m-dependence means that m is the infimum of the numbers denoted by b such that if $\|\mathbf{s}_1 - \mathbf{s}_2\| > b$, then $\xi(\mathbf{s}_1)$ and $\xi(\mathbf{s}_2)$ are independent. Here, $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^d$.

For $\mathbf{u} \in \mathcal{R}_n$, let

$$I_{m,n}(\mathbf{u}) = \{\mathbf{s} \in \mathcal{R}_n : \|\mathbf{s} - \mathbf{u}\| \le m\}$$

and $\kappa_n = \max_{\mathbf{u}\in\mathcal{R}_n} \sharp\{I_{m,n}(\mathbf{u})\}$. So κ_n denotes the number of elements of the set $I_{m,n}(\mathbf{u})$ with maximal cardinality. Therefore κ_n is an indicator of the strength of dependence. To avoid the independent case, we assume that $\kappa_n > 0$ for each n. We suppose that the measure κ_n of density of locations satisfies

$$\kappa_n \sim n^a \quad \text{with a constant } 0 < a < 1. \tag{2.1}$$

Here for any two sequences $\{t_n\}$ and $\{v_n\}$ of positive numbers, the notation $t_n \sim v_n$ means that the relation

$$0 < c_1 \le \liminf_{n\to\infty}(t_n/v_n) \le \limsup_{n\to\infty}(t_n/v_n) \le c_2 < \infty$$

holds for positive constants c_1 and c_2.

For real valued sequences $\{a_n\}$ and $\{b_n\}$, the notation $a_n = \mathrm{o}(b_n)$ (resp. $a_n = \mathrm{O}(b_n)$) means that the sequence a_n/b_n converges to 0 (resp. is bounded). The sign $\mathbb{E}$ stands for expectation. Variance and covariance are denoted by var(.) and cov(.,.), respectively. The sign "$\Rightarrow$" denotes convergence in distribution. $\mathcal{N}(m, \Sigma)$ stands for the (vector) normal distribution with mean (vector) m and covariance (matrix) Σ.

First, recall the CLT for m-dependent random fields presented in Park, Kim, Park and Hwang [6].

Consider a series of strictly stationary m-dependent random fields $\{\xi_n(\mathbf{s}) : \mathbf{s} \in D\}$, $D \subseteq \mathbb{R}^d$, $n = 1, 2, \ldots$. For a fixed n, let us introduce the notation $S_n = \sum_{i=1}^n \xi_n(\mathbf{s}_{ni})$. Furthermore, let $\mathcal{T}_n = \{(i,j) : 0 < \|\mathbf{s}_{ni} - \mathbf{s}_{nj}\| \le m\}$, $\nu_n = \mathrm{var}(\xi_n(\mathbf{s}))$ and

$$\tau_n = \frac{1}{n\kappa_n} \sum_{(i,j)\in\mathcal{T}_n} \mathrm{cov}(\xi_n(\mathbf{s}_{ni}), \xi_n(\mathbf{s}_{nj})). \tag{2.2}$$

At this point we notice that $\mathrm{var}(S_n) = n\nu_n + n\kappa_n\tau_n$ and τ_n can be negative as well.

Theorem 2.1 (Theorem 2 of Park, Kim, Park and Hwang [6])**.** *Let $\{\xi_n\}$ be a sequence of strictly stationary random fields on $D \subset \mathbb{R}^d$ with $\mathbb{E}\xi_n(\mathbf{s}) = 0$. Assume that $\sup_{\mathbf{s}\in D} |\xi_n(\mathbf{s})|$ is bounded with probability one and $\mathbb{E}\left|\prod_{j=1}^l \xi_n(\mathbf{s}'_{nj})\right| = \mathrm{O}\left(\nu_n^l\right)$ holds uniformly for all the different points $\mathbf{s}'_{nj} \in \{\mathbf{s}_{n1}, \ldots, \mathbf{s}_{nn}\}$. If $\nu_n + \kappa_n\tau_n \ge \delta\kappa_n\nu_n^2$ for some $\delta > 0$, then we have*

$$\frac{S_n}{\sqrt{\mathrm{var}(S_n)}} \Rightarrow \mathcal{N}(0,1).$$

3. Application to density estimation

In Park, Kim, Park and Hwang [6], the CLT was applied to obtain asymptotic normality of the kernel type density estimator.

Let $\{Z(\mathbf{s}) : \mathbf{s} \in D\}$ be a strictly stationary m-dependent random field, $D \subseteq \mathbb{R}^d$. For each $z \in \mathbb{R}$, let $F(z) = P(Z(\mathbf{s}) \le z)$. We call the function F marginal distribution function. Assume that there exist the appropriate marginal density function f. Suppose that we observe the values of Z at the points $\mathbf{s}_{n1}, \dots, \mathbf{s}_{nn}$ in D. In this section we study the nonparametric estimation of the marginal density function. Consider the kernel type density estimator

$$\hat{f}_n(z) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{z - Z(\mathbf{s}_{ni})}{h_n}\right).$$

Here K is a kernel. We say that the function $K : \mathbb{R} \to [0, \infty)$ is a kernel if it is a bounded, continuous, symmetric density function (with respect to the Lebesgue measure) and

$$\lim_{|u| \to \infty} |u| K(u) = 0. \tag{3.1}$$

Let $f_{s_{ni}, s_{nj}}$ be the joint density function of $Z(\mathbf{s}_{ni})$ and $Z(\mathbf{s}_{nj})$. Let $z \in \mathbb{R}$ be fixed. Consider the following assumptions.

(1) (a) $f(z) > 0$,

(b) f is continuous at z,

(c) $f_{s_{ni}, s_{nj}}$ are equicontinuous at (z, z), i.e. if $(z_1, z_2) \to (z, z)$, then

$$\sup_{i,j} |f_{s_{ni}, s_{nj}}(z_1, z_2) - f_{s_{ni}, s_{nj}}(z, z)| \to 0,$$

(d) all finite dimensional densities of $Z(\mathbf{s}_{n1}), Z(\mathbf{s}_{n2}), \dots$ exist and are bounded and continuous,

(e) if $n \to \infty$, then

$$\frac{1}{n\kappa_n} \sum_{(i,j) \in \mathcal{T}_n} \{f_{s_{ni}, s_{nj}}(z, z) - f(z)^2\} \to \tau,$$

where τ is a nonnegative constant depending on z,

(f) $h^2 n^a, \ 0 < a < 1$ is bounded.

(2) The kernel K is bounded, nonnegative on $\mathbb{R}$ and satisfies $\int_{\mathbb{R}} K = 1$; $|z|K(z) \to 0$ as $|z| \to \infty$.

(3) $h_n > 0$ is a sequence satisfying $h_n \to 0$ and $nh_n \to \infty$ as $n \to \infty$.

(4) There exists a constant $\delta > 0$ such that

$$f(z)\int_{\mathbb{R}} K^2 + \tau\kappa_n h_n \geq \delta\kappa_n h_n.$$

Theorem 3.1 (Theorem 3 of Park, Kim, Park and Hwang [6]). *Let us suppose that the assumptions* (1)–(4) *hold.*

1. *Then*

$$\left\{n^{-1}h_n^{-1}f(z)\int_{\mathbb{R}} K^2 + n^{-1}\kappa_n\tau\right\}^{-\frac{1}{2}}\{\hat{f}_n(z) - \mathbb{E}\hat{f}_n(z)\} \Rightarrow \mathcal{N}(0,1).$$

2. *Suppose that f is twice differentiable in a neighbourhood of z and $\int uK(u)du=$ 0. Moreover, assume that f'' is continuous, bounded and $nh_n^5 \to 0$, $n\kappa_n^{-1}h_n^4 \to$ 0. Then*

$$\left\{n^{-1}h_n^{-1}f(z)\int_{\mathbb{R}} K^2 + n^{-1}\kappa_n\tau\right\}^{-\frac{1}{2}}\{\hat{f}_n(z) - f(z)\} \Rightarrow \mathcal{N}(0,1).$$

4. Joint asymptotic normality for the density estimator

In Park, Kim, Park and Hwang [6], the multivariate asymptotic normality was not considered.

Our aim is to study the multidimensional version of Theorem 3.1, i.e. the joint asymptotic normality of the kernel type density estimator.

Proposition 4.1. *Let $z_1, z_2, \ldots, z_q$ be given distinct real numbers. We assume that*

$$\frac{1}{n\kappa_n}\sum_{i,j\in\mathcal{T}_n}\left(f_{s_{ni},s_{nj}}(z_r,z_t) - f(z_r)f(z_t)\right) \to \tau_{rt} \text{ if } n\to\infty.$$

Let $W = \left(\frac{\tau_{ij}\kappa_n}{n}\right)_{1\le i,j\le q}$ and let V be a diagonal matrix with diagonal elements $\frac{1}{nh_n}f(z_i)\int_{-\infty}^{\infty}K^2(t)dt$, $i = 1,\ldots,q$. Let $\Sigma = V + W$.

Then under certain conditions, $(\hat{f}_n(z_i) - f(z_i), i = 1,\ldots,q)$ is asymptotically $\mathcal{N}(0,\Sigma)$. The structure of Σ is the following:

$$\Sigma = \frac{1}{nh_n}\begin{bmatrix} f(z_1)\int K^2(t)dt + \tau_{11}\kappa_n h_n & \tau_{12}\kappa_n h_n & \ldots & \tau_{1q}\kappa_n h_n \\ \tau_{21}\kappa_n h_n & f(z_2)\int K^2(t)dt + \tau_{22}\kappa_n h_n & \ldots & \tau_{2q}\kappa_n h_n \\ \vdots & & \ddots & \vdots \\ \tau_{q1}\kappa_n h_n & \ldots & \ldots & f(z_q)\int K^2(t)dt + \tau_{qq}\kappa_n h_n \end{bmatrix}.$$

To obtain this result one has to apply Theorem 2.1 and the Cramér-Wold device.

We can see that the asymptotic covariance matrix Σ has a special structure. In the diagonal, the expressions $f(z_i)\int K^2(t)dt$ come from the asymptotic covariance matrix of the discrete parameter model. On the other hand, the elements $\tau_{ij}\kappa_n h_n$ correspond to the asymptotic covariance matrix of the continuous parameter model. We mention that the asymptotic covariance matrices are well-known both for the discrete time and the continuous time models. The combination of the two covariance structures was first pointed out in Fazekas and Chuprunov [2] for the kernel type density estimator and then in Karácsony and Filzmoser [3] for the regression estimator. To underline the importance of the covariance structure, we mention the following. When calculating numerically the density estimator for a continuous time model, we approximate the estimator with a one corresponding to an infill-increasing model. However, the limiting covariance structures of those models can be distinct.

We present examples that give numerical evidence for the phenomena described in the above proposition. First we consider a one-dimensional regular domain D.

Example 1. Moving average on the real line.

We consider the process on the l-lattice points of the domain $D=[0,t]$ with $l=0.1$ and $t=200$. It means that the distance between two neighbours is $l=0.1$.

That is, the sample is $z_1=\xi(1/10),\dots,z_n=\xi(2000/10)$ with $n=2000$. The data generation for the simulation is easy. Let $y_1,\dots,y_{n+4}$ be i.i.d. standard normal random variables and choose

$$z_i = 0.05\cdot y_i + 0.2\cdot y_{i+1} + 0.5\cdot y_{i+2} + 0.2\cdot y_{i+3} + 0.05\cdot y_{i+4},\ i=1,\dots,n.$$

So $\xi(s)$ is a moving average process. We can see that the data is m-dependent with $m=5$. The marginal density is $f(x)=\frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{-x^2}{2\sigma^2}\right)$ where $\sigma=0.5788$.

Using these data, we calculated the estimation of the marginal density function of the random field at the points $x_1=-1.0$, $x_2=-0.5$, $x_3=0.0$, $x_4=0.5$ and $x_5=1.0$. We used two values of the bandwidth, $h_1=0.10$ and $h_2=0.01$, and applied the standard normal density function as kernel K.

The simulations were performed with MATLAB, 5000 repetitions of the procedure were made. The data sets for both bandwidths h_1 and h_2 were the same. The theoretical values of the density function and the average of their estimators are shown in Table 1. For both values of the bandwidths we can see a close similarity of the theoretical and the empirical values.

We calculated the empirical covariance matrices Σ_1 (corresponding to bandwidth h_1) and Σ_2 (corresponding to bandwidth h_2) for our estimators

$$(\widehat{f}_n(x_1),\dots,\widehat{f}_n(x_5)).$$

$$\Sigma_1=\begin{bmatrix} 0.3078 & 0.0516 & -0.1107 & -0.1475 & -0.0624\\ 0.0516 & 0.8053 & -0.1524 & -0.3343 & -0.1540\\ -0.1107 & -0.1524 & 0.9289 & -0.1485 & -0.1221\\ -0.1475 & -0.3343 & -0.1485 & 0.7853 & 0.0632\\ -0.0624 & -0.1540 & -0.1221 & 0.0632 & 0.3195 \end{bmatrix}\cdot 10^{-3};$$

$$\Sigma_2 = \begin{bmatrix} 2.2605 & 0.0244 & -0.1598 & -0.0875 & -0.0631 \\ 0.0244 & 6.7115 & -0.1994 & -0.3860 & -0.1832 \\ -0.1598 & -0.1994 & 9.8334 & -0.1701 & -0.2003 \\ -0.0875 & -0.3860 & -0.1701 & 6.8598 & 0.0881 \\ -0.0631 & -0.1832 & -0.2003 & 0.0881 & 2.2602 \end{bmatrix} \cdot 10^{-3}.$$

The difference in the diagonals of Σ_1 and Σ_2 is clearly visible. The off-diagonal elements are almost the same.

x	-1.0	-0.5	0.0	0.5	1.0
$f(x)$	0.1549	0.4746	0.6892	0.4746	0.1549
$\overline{\hat{f}_n(x)}$ with $h_1 = 0.10$	0.1590	0.4726	0.6794	0.4728	0.1599
$\overline{\hat{f}_n(x)}$ with $h_2 = 0.01$	0.1543	0.4747	0.6876	0.4763	0.1564

Table 1: Theoretical values of the density function and the average of their estimators for the data of Example 1.

Now calculate the additional terms in the diagonals of the covariance matrices described by Σ defined in Proposition 4.1. In our case the elements of the diagonal matrix V_k for the bandwidth h_k $(k = 1, 2)$ are

$$\frac{1}{n}\frac{1}{h_k} f(x_i) \int_{-\infty}^{\infty} K^2(u)du = \frac{1}{2000}\frac{1}{h_k} f(x_i) \frac{1}{2\sqrt{\pi}}.$$

Since in the infill-increasing case only the diagonals of the limit covariance matrices can be different for different values of the bandwidth, we show in Table 2 the ratio between the diagonals of the difference of the empirical covariance matrices, $diag(\Sigma_2 - \Sigma_1)$, and of the theoretical covariance matrices, $diag(V_2 - V_1)$.

x	-1.0	-0.5	0.0	0.5	1.0
$\frac{diag(\Sigma_2-\Sigma_1)}{diag(V_2-V_1)}$	0.9927	0.9803	1.0176	1.0082	0.9867

Table 2: Ratio between the diagonal of the difference of the empirical covariance matrices and that of the theoretical covariance matrices for the data of Example 1.

These are close to 1 as it is expected from the above proposition.

Finally, Figure 1 shows histograms of $\frac{1}{2}(\hat{f}_n(0.5) + \hat{f}_n(1.0))$ for the bandwidths $h_1 = 0.10$ (left picture) and $h_2 = 0.01$ (right picture). Figure 2 shows histograms of $\frac{1}{3}(\hat{f}_n(-1.0) + \hat{f}_n(-0.5) + \hat{f}_n(0.0))$ for the above bandwidths.

The histograms are presented together with the theoretical normal densities with means and variances estimated from the data used for the histograms. The approximate normality of the density estimator stated in the above proposition is reflected in these figures. Different bandwidths lead to different spreads of the normal distribution.

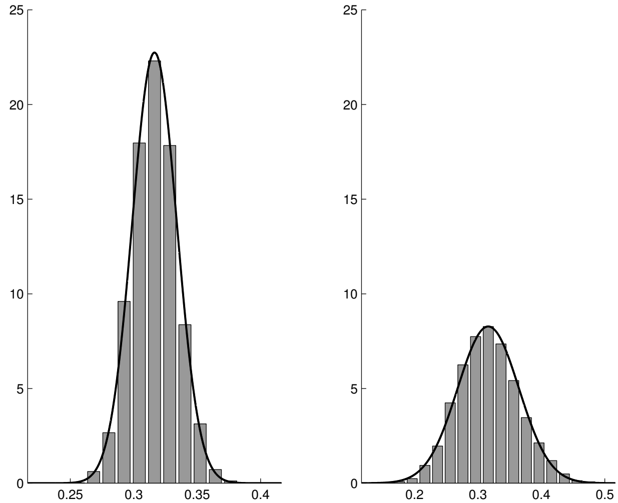


Figure 1: Histograms of $\frac{1}{2}(\hat{f}_n(0.5) + \hat{f}_n(1.0))$ for the bandwidths $h_1 = 0.10$ (left) and $h_2 = 0.01$ (right), together with the theoretical densities of the normal distribution for the data of Example 1.

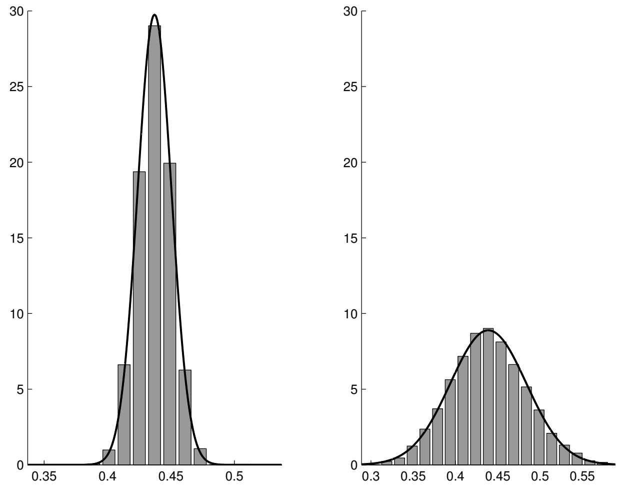


Figure 2: Histograms of $\frac{1}{3}(\hat{f}_n(-1.0) + \hat{f}_n(-0.5) + \hat{f}_n(0.0))$ for the bandwidths $h_1 = 0.10$ (left) and $h_2 = 0.01$ (right), together with the theoretical densities of the normal distribution for the data of Example 1.

Now we consider a two-dimensional domain with fractal-like shape.

Example 2. Two-dimensional moving average.

Now the locations will be the l-lattice points of the domain $D = [0, t]^2$ with $l = 0.1$ and $t = 10$. Thus the random field is $z_{(i,j)} = \xi_{(i/10, j/10)}, i, j = 1, \dots, 100$. Let $y_{k,l}$, $k, l = 1, \dots, 102$, be i.i.d. standard normal random variables, and let

$$z_{(i,j)} = \frac{1}{9} \sum_{k=i}^{i+2} \sum_{l=j}^{j+2} y_{k,l}, \quad i, j = 1, \dots, 100.$$

Therefore the random field is m-dependent with $m = 3$. The marginal density is $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-x^2}{2\sigma^2}\right)$ where $\sigma = 0.3333$.

Some points from the locations were omitted. In Figure 3, the small squares where the locations were deleted are marked with dark. We can see that in each white small square we have 16 sites of observations. Denote the set of the remaining locations by D. So the observations are $z_{(i,j)}$, $i, j \in D$. Therefore the actual sample size is 7056.

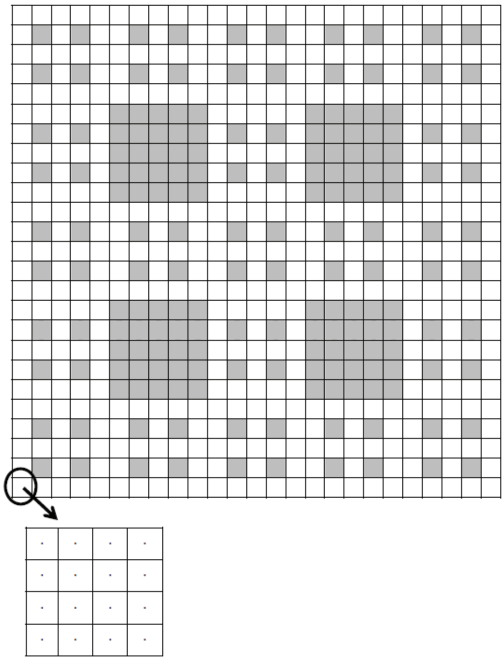


Figure 3: Sampling sites

It can be seen that the resulted domain is not convex. In the above proposition the asymptotic properties of the estimator remain true. It is clearly shown by the following numerical results.

As in the previous example, we calculated the density estimator $\hat{f}_n$ at the points $x_1 = -1.0$, $x_2 = -0.5$, $x_3 = 0.0$, $x_4 = 0.5$, $x_5 = 1.0$. We used the bandwidths

$h_1 = 0.10$ and $h_2 = 0.01$ and applied the standard normal density function as kernel K. The data sets for both bandwidths were the same, and 5000 repetitions were performed. Table 3 shows that the theoretical values of the density function and the average of their estimators are very similar.

x	-1.0	-0.5	0.0	0.5	1.0
$f(x)$	0.3886	0.9034	1.1968	0.9034	0.3886
$\overline{\hat{f}_n(x)}$ with $h = 0.10$	0.4087	0.8852	1.1460	0.8858	0.4085
$\overline{\hat{f}_n(x)}$ with $h = 0.01$	0.3907	0.9032	1.1965	0.9029	0.3895

Table 3: Theoretical values of the density function and the average of their estimators for the data of Example 2.

The empirical covariance matrices are

$$\Sigma_1 = \begin{bmatrix} 0.5124 & 0.3246 & -0.1801 & -0.4534 & -0.2921 \\ 0.3246 & 0.7406 & 0.0403 & -0.5479 & -0.4382 \\ -0.1801 & 0.0403 & 0.5769 & 0.0194 & -0.1941 \\ -0.4534 & -0.5479 & 0.0194 & 0.7785 & 0.3362 \\ -0.2921 & -0.4382 & -0.1941 & 0.3362 & 0.5089 \end{bmatrix} \cdot 10^{-3};$$

$$\Sigma_2 = \begin{bmatrix} 1.9357 & 0.2898 & -0.1783 & -0.5075 & -0.2852 \\ 0.2898 & 4.0989 & -0.0694 & -0.6534 & -0.5137 \\ -0.1783 & -0.0694 & 4.9750 & -0.1292 & -0.2899 \\ -0.5075 & -0.6534 & -0.1292 & 4.2037 & 0.3005 \\ -0.2852 & -0.5137 & -0.2899 & 0.3005 & 1.9322 \end{bmatrix} \cdot 10^{-3}$$

for the bandwidths h_1 and h_2, respectively. Again, the agreement of the off-diagonal elements and the difference in the diagonal becomes visible.

Similarly to the previous example, we show the ratios $\frac{diag(\Sigma_2 - \Sigma_1)}{diag(V_2 - V_1)}$ in Table 4. These are close to 1 as it was expected from our proposition.

x	-1.0	-0.5	0.0	0.5	1.0
$\frac{diag(\Sigma_2 - \Sigma_1)}{diag(V_2 - V_1)}$	1.0181	1.0331	1.0213	1.0537	1.0180

Table 4: Ratio between the diagonal of the difference of the empirical covariance matrices and that of the theoretical covariance matrices for the data of Example 2.

Finally, Figure 4 shows histograms of $\frac{1}{2}(\hat{f}_n(0.0) + \hat{f}_n(0.5))$ for the bandwidths $h_1 = 0.10$ (left picture) and $h_2 = 0.01$ (right picture). Figure 5 shows histograms of $\frac{1}{3}(\hat{f}_n(-1.0) + \hat{f}_n(-0.5) + \hat{f}_n(0.0))$ for the above bandwidths.

The histograms are presented together with the theoretical normal densities with means and variances estimated from the data used for the histograms. The approximate normality of the density estimator stated in the above proposition is reflected in these figures. Different bandwidths lead to different spreads of the normal distribution.

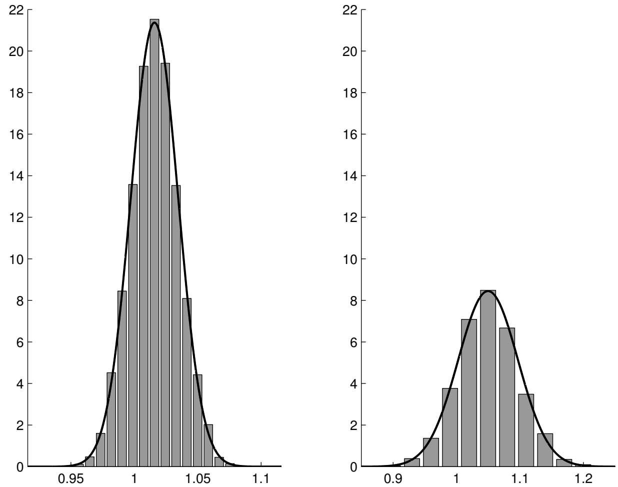


Figure 4: Histograms of $\frac{1}{2}(\hat{f}_n(0.0) + \hat{f}_n(0.5))$ for the bandwidths $h_1 = 0.10$ (left) and $h_2 = 0.01$ (right), together with the theoretical densities of the normal distribution for the data of Example 2.

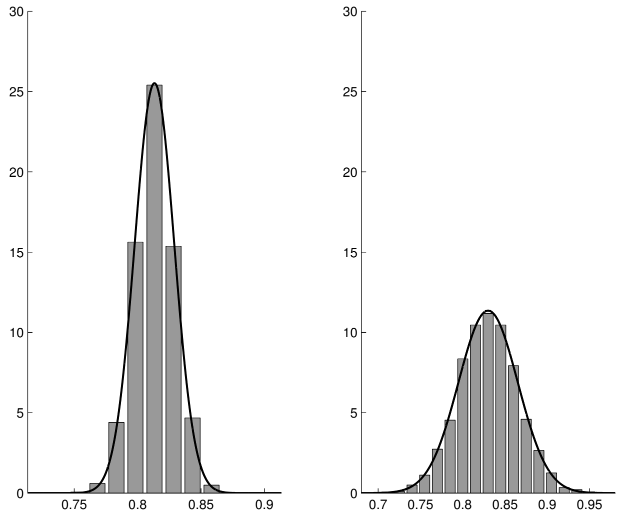


Figure 5: Histograms of $\frac{1}{3}(\hat{f}_n(-1.0) + \hat{f}_n(-0.5) + \hat{f}_n(0.0))$ for the bandwidths $h_1 = 0.10$ (left) and $h_2 = 0.01$ (right), together with the theoretical densities of the normal distribution for the data of Example 2.

5. Conclusions

In the paper, the kernel type density estimator $\hat{f}_n$ is considered. The underlying random field is m-dependent but the observation domain can be irregular. Nearly infill sampling scheme is supposed. Based on the CLT of Park, Kim, Park and Hwang [6] the joint asymptotic normality of $\hat{f}_1(x_1), \dots, \hat{f}_n(x_r)$ is obtained. The asymptotic covariance matrix is unusual in the sense that it is a combination of the covariance matrices in the continuous and the discrete parameter cases. Numerical evidence supports our results.

References

[1] Fazekas, I. and Chuprunov, A. (2004), A central limit theorem for random fields. *Acta Mathematica Academiae Paedagogicae Nyiregyhaziensis,* **20** (1), 93–104, www.emis.de/journals/AMAPN.

[2] Fazekas, I. and Chuprunov, A. (2006), Asymptotic normality of kernel type density estimators for random fields. *Stat. Inf. Stoch. Proc.* **9**, 161–178.

[3] Karácsony, Zs. and Filzmoser, P. (2010), Asymptotic normality of kernel type regression estimators for random fields. *Journal of Statistical Planning and Inference*, **140**, 872–886.

[4] Lahiri, S.N. (1999), Asymptotic distribution of the empirical spatial cumulative distribution function predictor and prediction bands based on a subsampling method. *Probab. Theory Related Fields*, **114** (1), 55–84.

[5] Lahiri, S.N., Kaiser, M.S., Cressie, N. and Hsu, N.J. (1999), Prediction of spatial cumulative distribution functions using subsampling. *J. Amer. Statist. Assoc.* **94** (445), 86–110.

[6] Park, B.U., Kim, T.Y., Park, T.-S., Hwang, S.Y. (2009), Practically Applicable Central Limit Theorem for Spatial Statistics. *Math. Geosci.* **41**, 555–569.

Annales Mathematicae et Informaticae
39 *(2012) pp. 57–70*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

A generalized allocation scheme*

István Fazekas[a], Bettina Porvázsnyik[b]

[a]University of Debrecen, Faculty of Informatics
Debrecen, Hungary
e-mail: fazekasi@inf.unideb.hu

[b]University of Debrecen, Faculty of Science and Technology
Debrecen, Hungary
e-mail: porv.bettina@gmail.com

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

The generalized allocation scheme was introduced by V.F. Kolchin [1]. Let $\xi_1, \xi_2, \dots, \xi_N$ be independent identically distributed non-negative integer valued non-degenerate random variables. Consider the random variables $\eta'_1, \dots, \eta'_N$ with joint distribution

$$\mathbb{P}\{\eta'_1 = k_1, \dots, \eta'_N = k_N\} = \mathbb{P}\Big\{\xi_1 = k_1, \dots, \xi_N = k_N \;\Big|\; \sum\nolimits_{i=1}^{N} \xi_i = n\Big\}.$$

Let ξ_i have Poisson distribution, then $(\eta'_1, \dots, \eta'_N)$ has polynomial distribution. Therefore $\{\eta'_1 = k_1, \dots, \eta'_N = k_N\}$ means that the contents of the boxes are $k_1, \dots, k_N$ after allocating n balls into N boxes during the usual allocation procedure.

Our aim is to study random variables $\eta_1, \dots, \eta_N$ with joint distribution

$$\mathbb{P}\{\eta_1 = k_1, \dots, \eta_N = k_N\} = \mathbb{P}\Big\{\xi_1 = k_1, \dots, \xi_N = k_N \;\Big|\; \sum\nolimits_{i=1}^{N} \xi_i \geq n\Big\}.$$

It can be considered as a general allocation scheme when we place at least n balls into N boxes. Let μ_{nN} denote the number of cases when $\{\eta_i = r\}$. That is μ_{nN} is the number of boxes containing r balls. We shall prove limit theorems for $\mathbb{P}\{\mu_{nN} = k\}$. Moreover, we shall consider the asymptotic behaviour of $\mathbb{P}\{\max_{1 \leq i \leq N} \eta_i \leq r\}$ and $\mathbb{P}\{\min_{1 \leq i \leq N} \eta_i \leq r\}$.

Keywords: generalized allocation scheme, conditional probability, law of large numbers, central limit theorem, Poisson distribution.

MSC: 60C05, 60F05

*Supported by the Hungarian Scientific Research Fund under Grant No. OTKA T079128/2009. Supported by the TÁMOP-4.2.2/B-10/1-2010-0024 project. The project is co-financed by the European Union and the European Social Fund.

1. Introduction

The usual allocation scheme is the following. Let n balls be placed successively and independently into N boxes. At any allocation the ball can fall into each box with probability $1/N$. This model was widely studied. See the early papers Weiss [13], Rényi [12], Békéssy [1] and the monograph Kolchin-Sevast'yanov-Chistyakov [8]. See also Chuprunov-Fazekas [2] for certain recent results.

A generalization of the usual allocation scheme was introduced by V.F. Kolchin (see the monographs of Kolchin [7] and Pavlov [10]). Let $\eta_1', \eta_2', \dots, \eta_N'$ be non-negative integer-valued random variables. In Kolchin's generalized allocation scheme the joint distribution of $\eta_1', \eta_2', \dots, \eta_N'$ can be represented as

$$\mathbb{P}\{\eta_1' = k_1, \dots, \eta_N' = k_N\} = \mathbb{P}\Big\{\xi_1 = k_1, \dots, \xi_N = k_N \ \Big| \ \sum\nolimits_{i=1}^N \xi_i = n\Big\}, \tag{1.1}$$

where $\xi_1, \xi_2, \dots, \xi_N$ are independent identically distributed non-negative integer valued non-degenerate random variables and $k_1, k_2, \dots, k_N$ are arbitrary non-negative integers, $k_1 + k_2 + \dots + k_N = n$. This scheme contains the usual allocation procedure, certain random forests, and several other models (see the monographs of Kolchin [7] and Pavlov [10]).

The usual allocation scheme is obtained as follows. Let ξ_i have Poisson distribution, i.e. $\mathbb{P}(\xi_i = k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, \dots$. Then

$$\mathbb{P}\{\xi_1 = k_1, \dots, \xi_N = k_N \ | \ \xi_1 + \dots + \xi_N = n\} = \frac{n!}{k_1! \dots k_N!}\left(\frac{1}{N}\right)^n$$

if $k_1 + \dots + k_N = n$. That is $(\eta_1', \dots, \eta_N')$ has polynomial distribution. Now $\{\eta_1' = k_1, \dots, \eta_N' = k_N\}$ means that the cell contents are $k_1, \dots, k_N$ after allocating n particles into N cells considering the usual allocation procedure.

The connection of the random forest and the generalized allocation scheme is the following. Let $\mathcal{T}_{n,N}$ denote the set of forests containing N labelled roots and n labelled non-root vertices. By Cayley's theorem, $\mathcal{T}_{n,N}$ has $N(n+N)^{n-1}$ elements. Consider uniform distribution on $\mathcal{T}_{n,N}$. Let η_i' denote the number of the non-root vertices of the ith tree. Then

$$\mathbb{P}\{\eta_1' = k_1, \dots, \eta_N' = k_N\} = \frac{n!}{k_1! \dots k_N!}\frac{(k_1+1)^{k_1-1} \dots (k_N+1)^{k_N-1}}{N(N+n)^{n-1}}.$$

Now let ξ_i have Borel distribution (see [5], [9]) $\mathbb{P}(\xi_i = k) = \frac{\lambda^k(1+k)^{k-1}}{k!}e^{-(k+1)\lambda}$, $k = 0, 1, \dots$, $\lambda > 0$. Then

$$\begin{aligned}&\mathbb{P}\{\xi_1 = k_1, \dots, \xi_N = k_N \ | \ \xi_1 + \dots + \xi_N = n\} \\ &= \frac{n!}{k_1! \dots k_N!}\frac{(k_1+1)^{k_1-1} \dots (k_N+1)^{k_N-1}}{N(N+n)^{n-1}}\end{aligned}$$

if $k_1 + \dots + k_N = n$. See [7, 2, 10]. Therefore $\eta_1', \dots, \eta_N'$ satisfy (1.1).

We can say that in the generalized allocation scheme we place n balls into N boxes. In the framework of the generalized allocation scheme several asymptotic results can be obtained. Let μ'_r be the number of the random variables $\eta'_1, \eta'_2, \dots, \eta'_N$ being equal to r $(r = 0, 1, \dots, n)$.

Observe that

$$\mu'_r = \mu'_{rnN} = \mu'_{nN} = \sum\nolimits_{i=1}^{N} \mathbb{I}_{\{\eta'_i = r\}} \tag{1.2}$$

can be considered as the number of boxes containing r balls. Here $\mathbb{I}_A$ is the indicator of the set A, i.e. $\mathbb{I}_A(x) = 1$ if $x \in A$ and $\mathbb{I}_A(x) = 0$ if $x \notin A$. (μ'_r, μ'_{rnN}, and μ'_{nN} are just different notations for the same quantity.)

Limit results for μ'_r can be obtained in the following way. Let ξ_0 be a random variable with the same distribution as ξ_1. Let $p_r = \mathbb{P}\{\xi_0 = r\}$ and $\mathbb{E}\xi_0 = a$. Introduce notation $S_N = \sum_{i=1}^{N} \xi_i$.

Denote by $\xi_0^{(r)}$ a random variable with distribution

$$\mathbb{P}\{\xi_0^{(r)} = k\} = \mathbb{P}\{\xi_0 = k \mid \xi_0 \neq r\}. \tag{1.3}$$

The expectation and the second moment of $\xi_0^{(r)}$ are the following $a_r = \mathbb{E}\xi_0^{(r)} = \dfrac{a - rp_r}{1 - p_r}$ and $\mathbb{E}\left(\xi_0^{(r)}\right)^2 = \dfrac{\mathbb{E}\xi_0^2 - r^2 p_r}{1 - p_r}$. Let $\xi_1^{(r)}, \dots, \xi_N^{(r)}$ be independent copies of $\xi_0^{(r)}$. Let $S_N^{(r)} = \sum_{i=1}^{N} \xi_i^{(r)}$. Denote by C_N^k the binomial coefficient $C_N^k = \binom{N}{k}$.

V.F. Kolchin proved in [7] the following lemma.

Lemma 1.1. *Let μ'_{nN} and $\xi_0^{(r)}$ be defined by* (1.2) *and* (1.3)*, respectively. Then*

$$\mathbb{P}\{\mu'{}_{nN} = k\} = C_N^k p_r^k (1 - p_r)^{N-k} \frac{\mathbb{P}\{S_{N-k}^{(r)} = n - kr\}}{\mathbb{P}\{S_N = n\}}. \tag{1.4}$$

Using this representation, normal and Poisson limit theorems were obtained (see [7], and [10]).

In [4] a modification of the generalized allocation scheme was studied, that is in (1.1) the condition was changed for $\sum_{i=1}^{N} \xi_i \leq n$.

In this paper we introduce another scheme, i.e. we use in (1.1) condition of the form $\sum_{i=1}^{N} \xi_i \geq n$. It can be considered as a general allocation scheme when we place at least n balls into N boxes. Let μ_{nN} denote the number of cases when $\{\eta_i = r\}$. That is μ_{nN} is the number of boxes containing r balls.

We shall prove limit theorems for $\mathbb{P}\{\mu_{nN} = k\}$. Moreover, we shall consider the asymptotic behaviour of $\mathbb{P}\{\max_{1 \leq i \leq N} \eta_i \leq r\}$ and $\mathbb{P}\{\min_{1 \leq i \leq N} \eta_i \leq r\}$.

In Section 2 $\mathbb{P}\{\mu_{nN} = k\}$ is studied. In sections 3 and 4 $\mathbb{P}\{\max_{1 \leq i \leq N} \eta_i \leq r\}$ and $\mathbb{P}\{\min_{1 \leq i \leq N} \eta_i \leq r\}$ are considered, respectively.

2. Another generalized allocation scheme

Let $\xi_1, \xi_2, \dots, \xi_N$ be independent identically distributed non-negative integer-valued non-degenerate random variables. Consider random variables $\eta_1, \eta_2, \dots, \eta_N$ with joint distribution

$$\mathbb{P}\{\eta_1 = k_1, \dots, \eta_N = k_N\} = \mathbb{P}\Big\{\xi_1 = k_1, \dots, \xi_N = k_N \;\Big|\; \sum\nolimits_{i=1}^{N} \xi_i \geq n\Big\}. \quad (2.1)$$

In this case, we place at least n balls into N boxes.

Example 2.1. Let ξ_i have Poisson distribution, i.e. $\mathbb{P}(\xi_i = k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, \dots$. Then

$$\mathbb{P}\{\xi_1 = k_1, \dots, \xi_N = k_N \mid \xi_1 + \dots + \xi_N \geq n\} = \frac{1}{k_1! \dots k_N!}\lambda^{k_0} \Big/ \sum_{k=n}^{\infty} \frac{(N\lambda)^k}{k!} \quad (2.2)$$

if $k_1 + \dots + k_N = k_0 \geq n$. Now, we place η (random number) balls into N boxes. Assume that $\eta \geq n$. Let η_i denote the number of balls in the ith box. Then

$$\begin{aligned}\mathbb{P}\{\eta_1 = k_1, \dots, \eta_N = k_N\} &= \sum_{i=n}^{\infty} \mathbb{P}(\eta_1 = k_1, \dots, \eta_N = k_N \mid \eta = i)\mathbb{P}(\eta = i) \\ &= \frac{k_0!}{k_1! \dots k_N!}\left(\frac{1}{N}\right)^{k_0} \mathbb{P}(\eta = k_0), \qquad (2.3)\end{aligned}$$

if $k_1 + \dots + k_N = k_0 \geq n$. If we choose the a'priori distribution of η as Poisson distribution truncated from below, i.e

$$\mathbb{P}(\eta = i) = \frac{(N\lambda)^i}{i!}e^{-N\lambda} \Big/ \sum_{k=n}^{\infty} \frac{(N\lambda)^k}{k!}e^{-N\lambda}, \quad i = n, n+1, \dots,$$

then we obtain (2.2). That is our scheme (2.1) with ξ_i having Poisson distribution describes the usual allocation when the number of balls is given by a truncated Poisson distribution.

Let

$$\mu_r = \mu_{rnN} = \mu_{nN} = \sum\nolimits_{i=1}^{N} \mathbb{I}_{\{\eta_i = r\}}$$

be the number of the boxes containing r balls. Then we have the following analogue of Kolchin's formula (1.4) for our model. Recall that $\xi_0^{(r)}$ is defined by (1.3).

Theorem 2.2. *For all* $k = 0, 1, 2, \dots, N$ *we have*

$$\mathbb{P}\{\mu_{nN} = k\} = C_N^k p_r^k (1 - p_r)^{N-k} \frac{\mathbb{P}\{S_{N-k}^{(r)} \geq n - kr\}}{\mathbb{P}\{S_N \geq n\}}. \quad (2.4)$$

Proof. (2.4) can be proved by a certain modification of the proof of Lemma 1.1.

Let $A_k^{(r)}$ be the event that exactly k of the random variables $\xi_1, \dots, \xi_N$ are equal to r. By (2.1), we have

$$\mathbb{P}\{\mu_{nN} = k\} = \mathbb{P}(A_k^{(r)} \mid S_N \geq n) = \frac{\mathbb{P}(A_k^{(r)}, S_N \geq n)}{\mathbb{P}(S_N \geq n)}.$$

Furthermore,

$$
\begin{aligned}
&\mathbb{P}(A_k^{(r)}, S_N \geq n) = \mathbb{P}(S_N \geq n |\, A_k^{(r)})\mathbb{P}(A_k^{(r)})\\
&= C_N^k p_r^k (1-p_r)^{N-k}\mathbb{P}(S_N \geq n |\, \xi_1 \neq r, \dots, \xi_{N-k} \neq r, \xi_{N-k+1} = r, \dots, \xi_N = r)\\
&= C_N^k p_r^k (1-p_r)^{N-k}\mathbb{P}(S_{N-k}^{(r)} \geq n - kr).
\end{aligned}
$$

Here we have used that $\xi_1, \dots, \xi_N$ are independent random variables and the event $A_k^{(r)}$ can occur C_N^k different ways, moreover

$$
\begin{aligned}
&\mathbb{P}(S_N \geq n |\, A_k^{(r)}) = \frac{\mathbb{P}(S_N \geq n \,, A_k^{(r)})}{\mathbb{P}(A_k^{(r)})}\\
&= \frac{\mathbb{P}(\xi_1 + \dots + \xi_N \geq n \,, \xi_1 \neq r, \dots, \xi_{N-k} \neq r, \xi_{N-k+1} = r, \dots, \xi_N = r)}{\mathbb{P}(\xi_1 \neq r, \dots, \xi_{N-k} \neq r, \xi_{N-k+1} = r, \dots, \xi_N = r)}\\
&= \frac{\mathbb{P}(\xi_1 + \dots + \xi_{N-k} \geq n - kr \,, \xi_1 \neq r, \dots, \xi_{N-k} \neq r, \xi_{N-k+1} = r, \dots, \xi_N = r)}{\mathbb{P}(\xi_1 \neq r, \dots, \xi_{N-k} \neq r, \xi_{N-k+1} = r, \dots, \xi_N = r)}\\
&= \frac{\mathbb{P}(\xi_1 + \dots + \xi_{N-k} \geq n - kr \,, \xi_1 \neq r, \dots, \xi_{N-k} \neq r)}{\mathbb{P}(\xi_1 \neq r, \dots, \xi_{N-k} \neq r)}. \qquad \square
\end{aligned}
$$

The proofs of our limit theorems are based on representation (2.4). First we consider two theorems with normal limiting distribution. Let $\alpha_{nN} = \frac{n}{N}$.

Theorem 2.3. *Let $\mathbb{E}\xi_0 = a$ be finite, $\mathbb{E}\xi_0^{(r)} = a_r$, $s_r^2 = p_r(1-p_r)$.*

(1) Let $d < a$. Then, uniformly for $\alpha_{nN} < d$, we have

$$
\mathbb{P}\{\mu_{nN} = k\} = \frac{1}{\sqrt{2\pi N} s_r} e^{-u^2/2}(1 + \mathrm{o}(1)), \tag{2.5}
$$

as $n, N \to \infty$ and $u = \frac{k - Np_r}{s_r N^{1/2}}$ belongs to an arbitrary bounded fixed interval.

(2) Suppose that $a_r < a$. Let $a_r < d_1 < d < a$. If k belongs to a bounded interval, then we have

$$
\lim_{n,N\to\infty,\, d_1 < \alpha_{nN} < d} \mathbb{P}\{\mu_{nN} = k\} = 0. \tag{2.6}
$$

Proof. (1) By the Moivre-Laplace Theorem we have

$$
C_N^k p_r^k (1-p_r)^{N-k} = \frac{1}{\sqrt{2\pi N} s_r} e^{-u^2/2}(1 + \mathrm{o}(1)), \tag{2.7}
$$

as $N \to \infty$ uniformly if $u = \frac{k - Np_r}{s_r N^{1/2}}$ belongs to a bounded fixed interval, where $s_r^2 = p_r(1-p_r)$.

As $\alpha_{nN} < d < a$, applying Kolmogorov's law of large numbers, we obtain

$$
\begin{aligned}
&\lim_{n,N\to\infty,\, \alpha_{nN} < d} \mathbb{P}\left\{\sum_{i=1}^{N-k} \xi_i^{(r)} \geq n - kr\right\} = 1,\\
&\lim_{n,N\to\infty,\, \alpha_{nN} < d} \mathbb{P}\left\{\sum_{i=1}^{N} \xi_i \geq n\right\} = 1.
\end{aligned} \tag{2.8}
$$

Now (2.4), (2.7) and (2.8) imply (2.5).

(2) Let $d_1 < \alpha_{nN} < d$. By Kolmogorov's law of large numbers, we have

$$\lim_{n,N\to\infty,\, d_1<\alpha_{nN}<d} \mathbb{P}\left\{\sum_{i=1}^{N-k} \xi_i^{(r)} \geq n - kr\right\} = 0. \tag{2.9}$$

We obtain (2.6) from (2.4), if we apply (2.7) and (2.9). □

Remark 2.4. It is easy to see that $a < a_r$, $a > a_r$ and $a = a_r$ if and only if $a > r$, $a < r$ and $a = r$, respectively.

Let Φ denote the standard normal distribution function. Recall that $a = \mathbb{E}\xi_0$, $a_r = \mathbb{E}\xi_0^{(r)}$ and $s_r^2 = p_r(1-p_r)$.

Theorem 2.5. *Suppose that $\mathbb{E}\xi_0^2 < \infty$. Denote by σ^2 the variance of ξ_0 and by σ_r^2 the variance of $\xi_0^{(r)}$. Assume $0 < \sigma^2, \sigma_r^2 < \infty$. Let $-\infty \leq C < \infty$. Then, as $n, N \to \infty$ such that $\sqrt{N}(\alpha_{nN} - a) \to C$, we have*

$$\mathbb{P}\{\mu_{nN} = k\} = \frac{1}{\sqrt{2\pi N} s_r} e^{-u^2/2} \left(\frac{1 - \Phi\left(\frac{C + us_r \frac{a-r}{1-p_r}}{\sqrt{1-p_r}\sigma_r}\right)}{1 - \Phi\left(\frac{C}{\sigma}\right)} + \mathrm{o}(1) \right), \tag{2.10}$$

for $u = \frac{k - Np_r}{s_r N^{1/2}}$ belonging to any bounded fixed interval.

Proof. As $\sigma^2 = \mathbb{D}^2(\xi_0) < \infty$ and $\sigma_r^2 = \mathbb{D}^2(\xi_0^{(r)}) < \infty$, by the central limit theorem, we obtain

$$\mathbb{P}\left\{\sum_{i=1}^{N} \xi_i \geq n\right\} = \mathbb{P}\left\{\frac{\sum_{i=1}^{N} \xi_i - Na}{\sqrt{N}\sigma} \geq \frac{\sqrt{N}(\alpha_{nN} - a)}{\sigma}\right\}$$
$$= 1 - \Phi\left(\frac{\sqrt{N}(\alpha_{nN} - a)}{\sigma}\right) + \mathrm{o}(1), \tag{2.11}$$

and similarly we obtain

$$\mathbb{P}\left\{\sum_{i=1}^{N-k} \xi_i^{(r)} \geq n - kr\right\} = 1 - \Phi\left(\frac{\sqrt{N}(\alpha_{nN} - a) + us_r \frac{a-r}{1-p_r}}{\sqrt{1-p_r}\sigma_r}\right) + \mathrm{o}(1). \tag{2.12}$$

Using (2.11), (2.12), and (2.7), relation (2.4) implies the desired result. □

Using large deviation theorems we can describe the relation between μ_{nN} and μ'_{nN}.

Let $X_1, X_2, \ldots$ be independent identically distributed non-negative non-degenerate random variables with lattice distribution (assume that the span of the distribution of X_1 is 1). Suppose that Cramér's condition is satisfied, that is $\mathbb{E}e^{\lambda_0 X_1} < \infty$ for some $\lambda_0 > 0$. Let $Z_N = X_1 + \cdots + X_N$. Introduce notation

$$M(h) = \mathbb{E}e^{hX_1}, \quad a(h) = (\ln(M(h)))', \quad v^2(h) = a'(h).$$

As X_1 is non-degenerate, therefore $a'(h) > 0$, so $a(\cdot)$ is strictly increasing.

We have the following lemma from [11].

Lemma 2.6. *Let x be an integer number and let $h = a^{-1}(\frac{x}{N})$. Then, as $N \to \infty$, we have*

$$\mathbb{P}(Z_N = x) = \frac{1}{v(h)\sqrt{2\pi N}} M^N(h) e^{-hx} \left(1 + \mathrm{O}\left(\frac{1}{N}\right)\right),$$
$$\mathbb{P}(Z_N \geq x) = \frac{1}{v(h)\sqrt{2\pi N}} M^N(h) e^{-hx} (1 - e^{-h})^{-1} \left(1 + \mathrm{O}\left(\frac{1}{N}\right)\right)$$

uniformly for x, with $Na(\varepsilon) \leq x \leq Na(\lambda_0 - \varepsilon)$, where ε is an arbitrary small positive number. In particular

$$\frac{\mathbb{P}(Z_N \geq x)}{\mathbb{P}(Z_N = x)} = (1 - e^{-h})^{-1}(1 + \mathrm{o}(1)). \tag{2.13}$$

Introduce notation

$$L(\lambda) = \mathbb{E}e^{\lambda \xi_0}, \quad L_r(\lambda) = \mathbb{E}e^{\lambda \xi_0^{(r)}}$$

where we assume that there exist positive constants $\lambda_0 > 0$ and $\lambda_0^{(r)} > 0$ such that $\mathbb{E}e^{\lambda_0 \xi_0} < \infty$ and $\mathbb{E}e^{\lambda_0^{(r)} \xi_0^{(r)}} < \infty$ (Cramér's condition). Let

$$m(\lambda) = (\ln(L(\lambda)))', \quad \sigma^2(\lambda) = m'(\lambda), \quad 0 \leq \lambda \leq \lambda_0\,,$$

$$m_r(\lambda) = (\ln(L_r(\lambda)))', \quad \sigma_r^2(\lambda) = m_r'(\lambda), \quad 0 \leq \lambda \leq \lambda_0^{(r)}.$$

As ξ_0 is non-degenerate, therefore $m(.)$ is strictly increasing. Assume that $0 < \mathbb{P}(\xi_0 = 0) < 1$. Moreover, if we additionally assume that $r \neq 0$ and $\mathbb{P}(\xi_0 = 0) + \mathbb{P}(\xi_0 = r) < 1$, then $\xi_0^{(r)}$ is non-degenerate, therefore similar property is valid for the function $m_r(.)$.

Let $h = m^{-1}(\alpha_{nN})$, $h_r = m_r^{-1}(\alpha_{nN})$, and $\beta(\alpha_{nN}) = \dfrac{1 - e^{-h}}{1 - e^{-h_r}}$.

Theorem 2.7. *Assume $r > 0$, $\mathbb{P}(\xi_0 = 0) > 0$, and $\mathbb{P}(\xi_0 = 0) + \mathbb{P}(\xi_0 = r) < 1$. Let $\max\{a, a_r\} < d_1 < d_2 < \min\{m(\lambda_0), m_r(\lambda_0^{(r)})\}$. Then, as $n, N \to \infty$, we have*

$$\mathbb{P}\{\mu_{nN} = k\} = \mathbb{P}\{\mu'_{nN} = k\}\beta(\alpha_{nN})(1 + \mathrm{o}(1)) \tag{2.14}$$

uniformly for $d_1 < \alpha_{nN} < d_2$.

Proof. We obtain Theorem 2.7 from (2.4) and from Lemma 1.1 , if we apply (2.13) both for ξ_i and for $\xi_i^{(r)}$. □

We shall use the so called power series distribution. Consider the random variable ξ_0 with the following distribution. Let $b_0,\ b_1, b_2, \dots$ be a sequence of non-negative numbers and let R denote the radius of convergence of the series

$$B(\theta) = \sum\nolimits_{k=0}^{\infty} \frac{b_k \theta^k}{k!}.$$

Assume that $R > 0$. Let $\xi_0 = \xi_0(\theta)$ have the following distribution

$$p_k = p_k(\theta) = \mathbb{P}\{\xi_0(\theta) = k\} = \frac{b_k \theta^k}{k! B(\theta)}, \quad k = 0, 1, 2, \dots . \tag{2.15}$$

Differentiating $B(\theta)$ for $0 \le \theta < R$, we obtain

$$\mathbb{E}\xi_0(\theta) = \frac{\theta B'(\theta)}{B(\theta)}, \quad \mathbb{D}^2\xi_0(\theta) = \frac{\theta^2 B''(\theta)}{B(\theta)} + \mathbb{E}\xi_0(\theta) - (\mathbb{E}\xi_0(\theta))^2$$

(see e.g. [7]).

We will assume that the distribution of the random variable $\xi_0(\theta)$ satisfies

$$b_0 > 0, \ b_1 > 0. \tag{2.16}$$

We emphasize that the distribution of $\xi_0 = \xi_0(\theta)$ is not fixed, it depends on θ.

We have the following Poisson limit theorem.

Theorem 2.8. *Suppose that the random variable $\xi_0 = \xi_0(\theta)$ has distribution* (2.15), *condition* (2.16) *is satisfied. Let $\theta \le K < R$. Let $r > 1$ and $\frac{n}{N^{1-\frac{1}{r}}} \to 0$. Let $N \to \infty$ such that $Np_r(\theta) \to \lambda$ for some $0 < \lambda < \infty$. Then for all $k \in \mathbb{N}$ we have*

$$\mathbb{P}\{\mu_{nN} = k\} = \frac{\lambda^k e^{-\lambda}}{k!}(1 + \mathrm{o}(1)). \tag{2.17}$$

Proof. Let $k \in \mathbb{N}$. By the Poisson limit theorem, one has

$$C_N^k p_r^k (1 - p_r)^{N-k} = \frac{\lambda^k e^{-\lambda}}{k!}(1 + \mathrm{o}(1)). \tag{2.18}$$

Relation $Np_r(\theta) \to \lambda$ implies that $\theta = \mathrm{o}(1)$, $B(\theta) = b_0 + \mathrm{o}(1)$, $B^{'}(\theta) = b_1 + \mathrm{o}(1)$ and $B^{''}(\theta) = b_2 + \mathrm{o}(1)$. Therefore $\theta = \left(\frac{r!(b_0\lambda + \mathrm{o}(1))}{Nb_r}\right)^{1/r}$.

We know that $\mathbb{E}\xi_0 = \frac{\theta B^{'}(\theta)}{B(\theta)}$. Therefore

$$\mathbb{E}\xi_0 = \frac{b_1}{b_0}\left(\frac{r!(b_0\lambda + \mathrm{o}(1))}{Nb_r}\right)^{1/r}(1 + \mathrm{o}(1)) = C\left(\frac{1}{N}\right)^{1/r}(1 + \mathrm{o}(1)). \tag{2.19}$$

Here and in what follows C denotes an appropriate constant (its value can be different in different formulae). Similarly

$$\mathbb{D}^2\xi_0 = C\left(\frac{1}{N}\right)^{1/r}(1 + \mathrm{o}(1)). \tag{2.20}$$

Now applying condition $\frac{n}{N^{1-\frac{1}{r}}} \to 0$, Chebishev's inequality and relations (2.19), (2.20), we obtain

$$\mathbb{P}\left\{S_N \geq n\right\} = (1 + \mathrm{o}(1)). \tag{2.21}$$

As $\mathbb{E}\xi_0^{(r)} = \dfrac{\mathbb{E}\xi_0 - rp_r}{1-p_r}$, so (2.19) and condition $\frac{n}{N^{1-\frac{1}{r}}} \to 0$ imply that

$$\mathbb{E}\xi_0^{(r)} = C\left(\frac{1}{N}\right)^{1/r}(1 + \mathrm{o}(1)). \tag{2.22}$$

We have

$$\mathbb{D}^2\xi_0^{(r)} = \frac{\mathbb{E}\xi_0^2}{1-p_r} - \frac{a^2}{(1-p_r)^2} + \frac{2arp_r}{(1-p_r)^2} - \frac{r^2p_r}{(1-p_r)^2}. \tag{2.23}$$

We obtain

$$\mathbb{D}^2\xi_0^{(r)} = C\left(\frac{1}{N}\right)^{1/r}(1 + \mathrm{o}(1)). \tag{2.24}$$

Now applying condition $\frac{n}{N^{1-\frac{1}{r}}} \to 0$, Chebishev's inequality and relations (2.22), (2.24), we obtain

$$\mathbb{P}\{S_{N-k}^{(r)} \geq n - kr\} = (1 + \mathrm{o}(1)). \tag{2.25}$$

Inserting (2.21), (2.25), and (2.18) into (2.4), we obtain (2.17). □

3. Limit theorems for $\max_{1\leq i\leq N}\eta_i$

Let $\eta_{(N)} = \max_{1\leq i\leq N}\eta_i$. $\eta_{(N)}$ is the maximal number of balls contained by any of the boxes.

Let $\xi_0^{(\leq r)}$ be a random variable with distribution

$$\mathbb{P}\{\xi_0^{(\leq r)} = k\} = \mathbb{P}\{\xi_0 = k \mid \xi_0 \leq r\}.$$

Let $\xi_i^{(\leq r)}$, $i = 1, \dots, N$, be independent copies of $\xi_0^{(\leq r)}$. Let $S_N^{(\leq r)} = \sum_{i=1}^N \xi_i^{(\leq r)}$ and $\mathbb{E}\xi_0^{(\leq r)} = a_{\leq r}$. We can see that $a_{\leq r} \leq a$. Moreover, $a_{\leq r} = a$ if and only if $\mathbb{P}(\xi_0 \leq r) = 1$, that is ξ_0 and $\xi_0^{(\leq r)}$ have the same distribution.

The following representation of $\eta_{(N)}$ is useful to obtain limit results.

Theorem 3.1. *We have*

$$\mathbb{P}\{\eta_{(N)} \leq r\} = (1 - P_r)^N \frac{\mathbb{P}\{S_N^{(\leq r)} \geq n\}}{\mathbb{P}\{S_N \geq n\}}, \tag{3.1}$$

for all $r \in \mathbb{N}$ where $P_r = \mathbb{P}\{\xi_0 > r\}$.

Proof.

$$
\begin{aligned}
\mathbb{P}\{\eta_{(N)} \le r\} &= \mathbb{P}\{\eta_1 \le r, \dots, \eta_N \le r\} \\
&= \mathbb{P}\Big\{\xi_1 \le r, \dots, \xi_N \le r \;\Big|\; \sum\nolimits_{i=1}^{N} \xi_i \ge n\Big\} \\
&= \frac{\mathbb{P}\Big\{\xi_1 \le r, \dots, \xi_N \le r \,, S_N \ge n\Big\}}{\mathbb{P}\Big\{S_N \ge n\Big\}} \\
&= \left(\mathbb{P}\{\xi_1 \le r\}\right)^N \frac{\mathbb{P}\Big\{S_N \ge n \;\Big|\; \xi_1 \le r, \dots, \xi_N \le r\Big\}}{\mathbb{P}\Big\{S_N \ge n\Big\}} \\
&= (1 - P_r)^N \frac{\mathbb{P}\{S_N^{(\le r)} \ge n\}}{\mathbb{P}\{S_N \ge n\}}. \qquad \square
\end{aligned}
$$

Theorem 3.2. *(1) Let $d < a_{\le r}$. Then for all fixed $r \in \mathbb{N}$, as $n, N \to \infty$, we have*

$$\mathbb{P}\{\eta_{(N)} \le r\} = (1 - P_r)^N (1 + \mathrm{o}(1)) \tag{3.2}$$

uniformly for $\alpha_{nN} < d$.

(2) Suppose that $a_{\le r} < a$ and $a_{\le r} < d_1 < d < a$. Then for all fixed $r \in \mathbb{N}$, as $n, N \to \infty$, we have

$$\mathbb{P}\{\eta_{(N)} \le r\} = (1 - P_r)^N \mathrm{o}(1) \tag{3.3}$$

uniformly for $d > \alpha_{nN} > d_1$.

Proof. (1) Apply Kolmogorov's law of large numbers for S_N and $S_N^{(\le r)}$ in (3.1). Then (3.2) follows.

(2) If $d_1 < \alpha_{nN} < d$ and we apply Kolmogorov's law of large numbers, then we obtain

$$\lim_{n,N\to\infty,\, d_1 < \alpha_{nN} < d} \mathbb{P}\left\{ \frac{S_N^{(\le r)}}{N} \ge \frac{n}{N} \right\} = 0. \tag{3.4}$$

$\square$

Theorem 3.3. *Suppose that $\mathbb{E}\xi_0^2 < \infty$ and let $\sigma_{\le r}^2$ be the variance of $\xi_0^{(\le r)}$. Let $-\infty \le C < \infty$. Then, for all $r \in \mathbb{N}$, as $n, N \to \infty$ such that $\sqrt{N}(\alpha_{nN} - a) \to C$, we have*

$$\mathbb{P}\{\eta_{(N)} \le r\} = (1 - P_r)^N \left(\frac{1 - \Phi\left(\frac{C}{\sigma_{\le r}}\right)}{1 - \Phi\left(\frac{C}{\sigma}\right)} + \mathrm{o}(1) \right), \quad \textit{for} \quad a_{\le r} = a, \tag{3.5}$$

$$\mathbb{P}\{\eta_{(N)} \le r\} = (1 - P_r)^N \cdot \mathrm{o}(1), \quad \textit{for} \quad a_{\le r} < a. \tag{3.6}$$

Proof. By the central limit theorem, we have

$$\mathbb{P}\left\{ \sum_{i=1}^{N} \xi_i^{(\le r)} \ge n \right\} = \mathbb{P}\left\{ \frac{\sum_{i=1}^{N} \xi_i^{(\le r)} - N a_{\le r}}{\sqrt{N}\sigma_{\le r}} \ge \frac{\sqrt{N}(\alpha_{nN} - a_{\le r})}{\sigma_{\le r}} \right\}$$

$$= 1 - \Phi\left(\frac{\sqrt{N}(\alpha_{nN} - a_{\le r})}{\sigma_{\le r}}\right) + \mathrm{o}(1). \tag{3.7}$$

In relation (3.1) apply (2.11) and (3.7) to obtain (3.5) and (3.6). □

Let $\eta'_{(N)} = \max_{1\le i\le N} \eta'_i$ be the maximum in the usual generalized allocation scheme (1.1). Using large deviation results, we can describe the relation of $\eta'_{(N)}$ and $\eta_{(N)}$.

Introduce notation

$$L_{\le r}(\lambda) = \mathbb{E}e^{\lambda\xi_0^{(\le r)}}$$

where we assume that there exist a positive constant $\lambda_0^{(\le r)} > 0$, such that

$$\mathbb{E}e^{\lambda_0^{(\le r)}\xi_0^{(\le r)}} < \infty \quad \text{(Cramér's condition)}.$$

Let

$$m_{\le r}(\lambda) = (\ln(L_{\le r}(\lambda)))', \quad \sigma^2_{\le r}(\lambda) = m'_{\le r}(\lambda), \quad 0 \le \lambda \le \lambda_0^{(\le r)}.$$

Let $r \ge 1$. If $\mathbb{P}(\xi_0 = 0) > 0$ and $\mathbb{P}(\xi_0 \le r) > \mathbb{P}(\xi_0 = 0)$, then $\xi_0^{(\le r)}$ is non-degenerate, therefore $m_{\le r}(.)$ is a strictly increasing function.

Let $h = m^{-1}(\alpha_{nN})$, $h_{\le r} = m^{-1}_{\le r}(\alpha_{nN})$, and $\beta_{\le r}(\alpha_{nN}) = \frac{1-e^{-h}}{1-e^{-h_{\le r}}}$.

Theorem 3.4. *Assume that* $r \ge 1$, $\mathbb{P}(\xi_0 = 0) > 0$, *and* $\mathbb{P}(\xi_0 \le r) > \mathbb{P}(\xi_0 = 0)$. *Let* $\max\{a, a_{\le r}\} < d_1 < d_2 < \min\{m(\lambda_0), m_{\le r}(\lambda_0^{(\le r)})\}$. *Then, for all* $r \in \mathbb{N}$ *as* $n, N \to \infty$, *we have*

$$\mathbb{P}\{\eta_{(N)} \le r\} = \mathbb{P}\{\eta'_{(N)} \le r\}\beta_{\le r}(\alpha_{nN})(1 + \mathrm{o}(1)) \tag{3.8}$$

uniformly for $d_1 < \alpha_{nN} < d_2$.

Proof. For the usual generalized allocation scheme, V.F. Kolchin in [7] obtained that

$$\mathbb{P}\{\eta'_{(N)} \le r\} = (1 - P_r)^N \frac{\mathbb{P}\{S_N^{(\le r)} = n\}}{\mathbb{P}\{S_N = n\}} \tag{3.9}$$

for all $r \in \mathbb{N}$ where $P_r = \mathbb{P}\{\xi_0 > r\}$.

Using (3.9) and (3.1) and applying (2.13) both for ξ_i and for $\xi_i^{(\le r)}$, the proof of Theorem 3.4 is complete. □

Theorem 3.5. *Suppose that the random variable* $\xi = \xi(\theta)$ *has distribution* (2.15), *condition* (2.16) *is satisfied and* $\theta \le K < R$. *Let* $r \in \mathbb{N}$. *Let* $\theta = \theta(N)$ *be such that* $Np_{r+1}(\theta) \to \lambda$ *where* $0 < \lambda < \infty$. *Then, as* $n, N \to \infty$ *such that* $\frac{n}{N^{r/(r+1)}} \to 0$, *we have*

$$\mathbb{P}\{\eta_{(N)} = r\} = e^{-\lambda} + \mathrm{o}(1), \tag{3.10}$$

$$\mathbb{P}\{\eta_{(N)} = r + 1\} = 1 - e^{-\lambda} + \mathrm{o}(1). \tag{3.11}$$

Proof. Relation $\frac{n}{N^{r/(r+1)}} \to 0$ implies that

$$B(\theta) = b_0 + \mathrm{o}(1) \quad \text{and} \quad \theta = \left(\frac{(r+1)!(b_0\lambda + \mathrm{o}(1))}{N b_{r+1}} \right)^{1/(r+1)}.$$

Using $r+1$ instead of r in the proof of Theorem 2.8, we obtain

$$\mathbb{P}\left\{S_N \geq n\right\} = (1 + \mathrm{o}(1)). \tag{3.12}$$

Let $r_1 \in \mathbb{N}$. Then

$$\begin{aligned}
\mathbb{E}\xi_0^{(\leq r_1)} &= \frac{\sum_{k=1}^{r_1} k \frac{b_k}{k! b_0} \left(\left(\frac{(r+1)!(b_0\lambda+\mathrm{o}(1))}{N b_{r+1}} \right)^{1/(r+1)} \right)^k}{\sum_{k=0}^{r_1} \frac{b_k}{k! b_0} \left(\left(\frac{(r+1)!(b_0\lambda+\mathrm{o}(1))}{N b_{r+1}} \right)^{1/(r+1)} \right)^k} (1 + \mathrm{o}(1)) \\
&= C \left(\frac{1}{N} \right)^{1/(r+1)} (1 + \mathrm{o}(1)).
\end{aligned} \tag{3.13}$$

Moreover,

$$\mathbb{D}^2 \xi_0^{(\leq r_1)} \leq C \left(\frac{1}{N} \right)^{1/(r+1)} (1 + \mathrm{o}(1)). \tag{3.14}$$

Using Chebishev's inequality, (3.13) and (3.14), we obtain

$$\mathbb{P}\{S_N^{(\leq r_1)} \geq n\} = (1 + \mathrm{o}(1)). \tag{3.15}$$

Using relations $\theta \to 0$ and $N p_{r+1}(\theta) \to \lambda$, we obtain

$$(1 - P_{r-1})^N = \mathrm{o}(1), \quad (1 - P_r)^N = e^{-\lambda} + \mathrm{o}(1), \quad (1 - P_{r+1})^N = 1 + \mathrm{o}(1). \tag{3.16}$$

Inserting (3.12), (3.15), and (3.16) into (3.1), we obtain

$$\mathbb{P}\{\eta_{(N)} \leq r-1\} = \mathrm{o}(1), \quad \mathbb{P}\{\eta_{(N)} \leq r\} = e^{-\lambda} + \mathrm{o}(1), \quad \mathbb{P}\{\eta_{(N)} \leq r+1\} = 1 + \mathrm{o}(1).$$

These relations imply (3.10) and (3.11). □

4. Limit theorems for $\min_{1 \leq i \leq N} \eta_i$

In this section we shall prove limit theorems for the minimal content of the boxes. Let $\eta_{(N-)} = \min_{1 \leq i \leq N} \eta_i$. Let $\xi_0^{(\geq r)}$ be a random variable with distribution $\mathbb{P}\{\xi_0^{(\geq r)} = k\} = \mathbb{P}\{\xi_0 = k \mid \xi_0 \geq r\}$. Let $\xi_i^{(\geq r)}$, $i = 1, \ldots, N$, be independent copies of $\xi_0^{(\geq r)}$. Let $S_N^{(\geq r)} = \sum_{i=1}^N \xi_i^{(\geq r)}$ and $\mathbb{E}\xi_0^{(\geq r)} = a_{\geq r}$. One can see that $\mathbb{E}\xi_0^{(\geq r)} \geq \mathbb{E}\xi_0$ and equality can happen if and only if $\xi_0^{(\geq r)} = \xi_0$.

We start with an appropriate representation of $\eta_{(N-)}$

Theorem 4.1. *We have*

$$\mathbb{P}\{\eta_{(N-)} \geq r\} = (1-Q_r)^N \frac{\mathbb{P}\{S_N^{(\geq r)} \geq n\}}{\mathbb{P}\{S_N \geq n\}}, \tag{4.1}$$

for all $r \in \mathbb{N}$ where $Q_r = \mathbb{P}\{\xi_0 < r\}$.

Proof.

$$\begin{aligned}
\mathbb{P}\{\eta_{(N-)} \geq r\} &= \mathbb{P}\{\eta_1 \geq r, \ldots, \eta_N \geq r\} \\
&= \mathbb{P}\Big\{\xi_1 \geq r, \ldots, \xi_N \geq r \;\Big|\; \sum\nolimits_{i=1}^N \xi_i \geq n\Big\} \\
&= \frac{\mathbb{P}\Big\{\xi_1 \geq r, \ldots, \xi_N \geq r\,, S_N \geq n\Big\}}{\mathbb{P}\Big\{S_N \geq n\Big\}} \\
&= (\mathbb{P}\{\xi_1 \geq r\})^N \frac{\mathbb{P}\Big\{S_N \geq n \;\Big|\; \xi_1 \geq r, \ldots, \xi_N \geq r\Big\}}{\mathbb{P}\Big\{S_N \geq n\Big\}} \\
&= (1-Q_r)^N \frac{\mathbb{P}\{S_N^{(\geq r)} \geq n\}}{\mathbb{P}\{S_N \geq n\}}. \qquad \square
\end{aligned}$$

Theorem 4.2. *Let $d < a$. Then for all $r \in \mathbb{N}$, as $n, N \to \infty$, we have*

$$\mathbb{P}\{\eta_{(N-)} \geq r\} = (1-Q_r)^N (1 + \mathrm{o}(1)) \tag{4.2}$$

uniformly for $\alpha_{nN} < d$.

Proof. We apply Kolmogorov's law of large numbers for S_N and $S_N^{(\geq r)}$ in (4.1). Then we obtain (4.2). □

Theorem 4.3. *Suppose that $\mathbb{E}\xi_0^2 < \infty$ and let $\sigma_{\geq r}^2$ be the variance of $\xi_0^{(\geq r)}$. Let $-\infty \leq C < \infty$. Then, for all $r \in \mathbb{N}$, as $n, N \to \infty$ such that $\sqrt{N}(\alpha_{nN} - a) \to C$, we have*

$$\mathbb{P}\{\eta_{(N-)} \geq r\} = (1-Q_r)^N \left(\frac{1 - \Phi\left(\frac{C}{\sigma_{\geq r}}\right)}{1 - \Phi\left(\frac{C}{\sigma}\right)} + \mathrm{o}(1) \right), \quad \textit{for} \quad a_{\geq r} = a, \tag{4.3}$$

$$\mathbb{P}\{\eta_{(N-)} \geq r\} = (1-Q_r)^N \left(\frac{1}{1 - \Phi\left(\frac{C}{\sigma}\right)} + \mathrm{o}(1) \right), \quad \textit{for} \quad a_{\geq r} > a. \tag{4.4}$$

Proof. By the central limit theorem, we have

$$\begin{aligned}
\mathbb{P}\left\{ \sum_{i=1}^N \xi_i^{(\geq r)} \geq n \right\} &= \mathbb{P}\left\{ \frac{\sum_{i=1}^N \xi_i^{(\geq r)} - N a_{\geq r}}{\sqrt{N}\sigma_{\geq r}} \geq \frac{\sqrt{N}(\alpha_{nN} - a_{\geq r})}{\sigma_{\geq r}} \right\} \\
&= 1 - \Phi\left(\frac{\sqrt{N}(\alpha_{nN} - a_{\geq r})}{\sigma_{\geq r}} \right) + \mathrm{o}(1).
\end{aligned} \tag{4.5}$$

In relation (4.1) apply (2.11) and (4.5). Then we obtain (4.3) and (4.4). □

References

[1] Békéssy, A. On classical occupancy problems. I. *Magy. Tud. Akad. Mat. Kutató Int. Közl.* **8** (1963), No.1-2, 59–71.

[2] Chuprunov, A.N. and Fazekas, I. Inequalities and strong laws of large numbers for random allocations. *Acta Math. Hungar.* **109** (2005), No.1-2, 163–182.

[3] Chuprunov, A.N. and Fazekas, I. An inequality for moments and its applications to the generalized allocation scheme. *Publ. Math. Debrecen,* **76** (2010), no. 3, 271–286.

[4] Chuprunov, A.N. and Fazekas, I. One some analogue of the generalized allocation scheme. (Russian) *Diskretnaya Matematika*, **21**(2012), No.1.

[5] Johnson, N. L., Kemp, A. W. and Kotz, S. *Univariate Discrete Distributions.* Third edition. Wiley Series in Probability and Statistics. Wiley-Interscience, Hoboken, NJ, 2005.

[6] Kolchin, A. V. and Kolchin, V. F. On transition of distributions of sums of independent identically distributed random variables from one lattice to another in the generalised allocation scheme. (Russian) *Diskret. Mat.* **18** (2006), no. 4, 113–127; translation in *Discrete Math. Appl.* **16** (2006), no. 6, 527–540.

[7] Kolchin, V.F. *Random Graphs.* Cambridge University Press, Cambridge, 1999.

[8] Kolchin, V.F. , Sevast'yanov, B.A. and Chistyakov, V.P. *Random allocations.* V.H. Winston & Sons, Washington D. C. 1978.

[9] Lerner, B., Lone, A. and Rao, M. On generalized Poisson distributions. *Probab. Math. Statist.* **17** (1997), no. 2, Acta Univ. Wratislav. No. 2029, 377–385.

[10] Pavlov, Yu. L. *Random Forests.* VSP, Utrecht, 2000.

[11] Rozovskii, L. V. Probabilities of large deviations for some classes of distributions satisfying the Cramér condition. (Russian) *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **298** (2003), *Veroyatn. i Stat.* 6, 161–185, 318–319; translation in *J. Math. Sci. (N. Y.)* **128** (2005), no. 1, 2585–2600.

[12] Rényi, A. Three new proofs and generalization of a theorem of Irving Weiss. *Magy. Tud. Akad. Mat. Kutató Int. Közl.* **7** (1962), No.1-2, 203–214.

[13] Weiss, I. Limiting distributions in some occupancy problems. *Ann. Math. Statist.* **29** (1958), No.3, 878–884.

Annales Mathematicae et Informaticae
39 *(2012) pp. 71–81*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

On weighted averages of double sequences

István Fazekas[a*], **Tibor Tómács**[b]

[a]University of Debrecen, Faculty of Informatics
Debrecen, Hungary
e-mail: fazekasi@inf.unideb.hu

[b]Eszterházy Károly College,
Institute of Mathematics and Computer Science
Eger, Hungary
e-mail: tomacs@ektf.hu

Dedicated to Mátyás Arató on his eightieth birthday

1. Introduction

The well known Kolmogorov strong law of large numbers states the following. If $X_1, X_2, \dots$ are independent identically distributed (i.i.d.) random variables with finite expectation and $\mathrm{E}\,X_1 = 0$, then the average $(X_1 + \dots + X_n)/n$ converges to 0 almost surely (a.s.). However, if we consider a double sequence, then we need another condition. Actually, if (X_{ij}) is a double sequence of i.i.d. random variables with $\mathrm{E}\,X_{11} = 0$, then $\mathrm{E}\,|X_{11}| \log^+ |X_{11}| < \infty$ implies that $\left(\sum_{i=1}^m \sum_{j=1}^n X_{ij}\right)/(mn)$ converges to 0 a.s., as n, m tend to infinity (see Smythe [6]).

For a double numerical sequence x_{ij} there are different notions of convergences. One can consider a strong version of convergence when x_{ij} converges as one of the indices i, j goes to infinity (this type of convergence was used in Fazekas [1]). Another version when x_{ij} converges as both indices i, j tend to infinity. However, in the second case convergence does not imply boundedness. To avoid unpleasant situations one can assume that the sequence is bounded. In this paper we shall study the so called bounded convergence of double sequences.

We shall prove two criteria for the bounded convergence of weighted averages of double sequences. Both criteria are based on subsequences. The subsequence is constructed by a well-known method: we proceed along a non-negative, increasing, unbounded sequence and pick up a member which is about the double

*Supported by the Hungarian Scientific Research Fund under Grant No. OTKA T079128/2009.

of the previous selected member of the sequence. (This method was applied e.g. in Fazekas–Klesov [2]). However, this method is not convenient for an arbitrary double sequence of weights. Therefore we apply weights of product type (it was considered e.g. in Noszály–Tómács [5]).

Our theorems can be considered as generalizations of some results in Fekete–Georgieva–Móricz [3], where harmonic averages of double sequences were considered. They obtained the following theorem.

$$\frac{1}{\ln m \ln n} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{x_{ij}}{ij} \xrightarrow{\text{b}} L, \quad \text{as} \quad m, n \to \infty \tag{1.1}$$

if and only if

$$\frac{1}{2^{m+n}} \max_{\substack{2^{2^{m-1}} < k \le 2^{2^m} \\ 2^{2^{n-1}} < l \le 2^{2^n}}} \left| \sum_{i=2^{2^{m-1}}+1}^{k} \sum_{j=2^{2^{n-1}}+1}^{l} \frac{x_{ij} - L}{ij} \right| \xrightarrow{\text{b}} 0, \text{ as } m, n \to \infty. \tag{1.2}$$

Here $\xrightarrow{\text{b}}$ means the bounded convergence. Our Theorem 2.4 is a generalization of this result for general weights.

Our results can also be considered as extensions of certain theorems of Móricz and Stadtmüller [4] where ordinary (that is not double) sequences were studied. In our proofs we apply ideas of [4].

2. Main results

Let $(x_{kl} : k, l = 1, 2, \dots)$ be a sequence of real numbers, and let $(b_k : k = 1, 2, \dots)$, $(c_l : l = 1, 2, \dots)$ be sequences of weights, that is, sequences of non-negative numbers for which

$$B_m := \sum_{k=1}^{m} b_k \to \infty, \quad \text{as} \quad m \to \infty, \tag{2.1}$$

$$C_n := \sum_{l=1}^{n} c_l \to \infty, \quad \text{as} \quad n \to \infty. \tag{2.2}$$

Let $a_{kl} := b_k c_l$, $A_{mn} := \sum_{k=1}^{m} \sum_{l=1}^{n} a_{kl}$ and $S_{mn} := \sum_{k=1}^{m} \sum_{l=1}^{n} a_{kl} x_{kl}$. The *weighted averages* Z_{mn} of the sequence (x_{kl}) with respect to the weights (a_{kl}) are defined by

$$Z_{mn} := \frac{1}{A_{mn}} S_{mn}$$

for n, m large enough so that $A_{mn} > 0$.

We define a sequence $m_0 = 0, m_1 = 1 < m_2 < m_3 < \dots$ of integers with the following property

$$B_{m_{i+1}-1} < 2B_{m_i} \le B_{m_{i+1}}, \quad i = 1, 2, \dots \tag{2.3}$$

Similarly, let $n_0 = 0, n_1 = 1 < n_2 < n_3 < \dots$ be a sequence of integers such that

$$C_{n_{j+1}-1} < 2C_{n_j} \le C_{n_{j+1}}, \quad j = 1, 2, \dots \tag{2.4}$$

In this paper we shall also use the following notation

$$\Delta_{st}^{mn} A := \sum_{k=s+1}^{m} \sum_{l=t+1}^{n} a_{kl}, \quad \Delta_{st}^{mn} S := \sum_{k=s+1}^{m} \sum_{l=t+1}^{n} a_{kl} x_{kl}.$$

Actually $\Delta_{st}^{mn} A$ is an increment on a rectangle (in other word two-dimensional difference) of the sequence A_{mn}. We note that

$$\frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \Delta_{m_i n_j}^{m_{i+1} n_{j+1}} S$$

is called the *moving average* of the sequence (x_{kl}) with respect to the weights (a_{kl}).

Definition 2.1. Let $(y_{kl} : k, l = 1, 2, \dots)$ be a sequence of real numbers, and let y be a real number. It is said that *bounded convergence*

$$y_{kl} \xrightarrow{\text{b}} y, \quad \text{as} \quad k, l \to \infty,$$

is satisfied if
(i) the sequence $(y_{kl} : k, l = 1, 2, \dots)$ is bounded; and
(ii) for every $\varepsilon > 0$ there exist positive integers k_0, l_0, such that

$$|y_{kl} - y| < \varepsilon \quad \text{for} \quad k \ge k_0, l \ge l_0. \tag{2.5}$$

Remark 2.2. Relation (2.5) does not imply that (y_{kl}) is bounded. For example if $y_{1l} = l$ for $l \ge 1$ and $y_{kl} = y$ for $k \ge 2, l \ge 1$, then (2.5) holds but (y_{kl}) is unbounded.

Theorem 2.3. *Suppose that conditions* (2.1) *and* (2.2) *are satisfied. Then for some constant* L, *we have*

$$Z_{m_i n_j} \xrightarrow{\text{b}} L, \quad \textit{as} \quad i, j \to \infty \tag{2.6}$$

if and only if

$$\frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \Delta_{m_i n_j}^{m_{i+1} n_{j+1}} S \xrightarrow{\text{b}} L, \quad \textit{as} \quad i, j \to \infty, \tag{2.7}$$

where the sequences (m_i) *and* (n_j) *are defined in* (2.3) *and* (2.4).

Theorem 2.4. *Assume that* $B_m/b_m \ge 1 + \delta$ *and* $C_m/c_m \ge 1 + \delta$ *for* m *being large enough where* $\delta > 0$. *Assume that conditions* (2.1), (2.2) *are satisfied. Then for some constant* L, *we have*

$$Z_{mn} \xrightarrow{\text{b}} L, \quad \textit{as} \quad m, n \to \infty \tag{2.8}$$

if and only if

$$\frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \max_{\substack{m_i < m \le m_{i+1} \\ n_j < n \le n_{j+1}}} \left| \sum_{k=m_i+1}^{m} \sum_{l=n_j+1}^{n} a_{kl}(x_{kl} - L) \right| \xrightarrow{\mathrm{b}} 0, \text{ as } i, j \to \infty, \tag{2.9}$$

where the sequences (m_i) *and* (n_j) *are defined in* (2.3) *and* (2.4).

The following two corollaries characterize the strong law of large numbers for weighted averages of a sequence of random variables with two-dimensional indices. These corollaries are consequences of Theorem 2.3 and 2.4.

Corollary 2.5. *Let* $(X_{kl} : k, l = 1, 2, \dots)$ *be a sequence of random variables. If conditions* (2.1) *and* (2.2) *are satisfied, then for some constant* L*, we have*

$$\frac{1}{A_{m_i n_j}} \sum_{k=1}^{m_i} \sum_{l=1}^{n_j} a_{kl} X_{kl} \xrightarrow{\mathrm{b}} L, \quad \text{as} \quad i, j \to \infty \quad \text{a.s.}$$

if and only if

$$\frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \sum_{k=m_i+1}^{m_{i+1}} \sum_{l=n_j+1}^{n_{j+1}} a_{kl} X_{kl} \xrightarrow{\mathrm{b}} L, \quad \text{as} \quad i, j \to \infty \quad \text{a.s.},$$

where the sequences (m_i) *and* (n_j) *are defined in* (2.3) *and* (2.4).

Corollary 2.6. *Let* $(X_{kl} : k, l = 1, 2, \dots)$ *be a sequence of random variables. Assume that* $B_m / b_m \ge 1 + \delta$ *and* $C_m / c_m \ge 1 + \delta$ *for* m *being large enough where* $\delta > 0$. *Assume that conditions* (2.1) *and* (2.2) *are satisfied. Then for some constant* L*, we have*

$$\frac{1}{A_{mn}} \sum_{k=1}^{m} \sum_{l=1}^{n} a_{kl} X_{kl} \xrightarrow{\mathrm{b}} L, \quad \text{as} \quad m, n \to \infty \quad \text{a.s.}$$

if and only if

$$\frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \max_{\substack{m_i < m \le m_{i+1} \\ n_j < n \le n_{j+1}}} \left| \sum_{k=m_i+1}^{m} \sum_{l=n_j+1}^{n} a_{kl}(X_{kl} - L) \right| \xrightarrow{\mathrm{b}} 0, \text{ as } i, j \to \infty \quad \text{a.s.},$$

where the sequences (m_i) *and* (n_j) *are defined in* (2.3) *and* (2.4).

Remark 2.7. In the above two corollaries L can be an a.s. finite random variable, as well.

Remark 2.8. The results of this section can be generalized for sequences with d-dimensional indices.

3. Proofs of Theorems 2.3 and 2.4

Proof of Theorem 2.3. Let ε be a fixed positive real number. First we prove the necessity. Assume that (2.6) is satisfied, that is, there exist integers i_0, j_0 such that

$$\left|Z_{m_i n_j} - L\right| < \varepsilon \quad \text{for all} \quad i \geq i_0, j \geq j_0,$$

furthermore $(Z_{m_i n_j})$ is a bounded sequence. So, if $i \geq i_0, j \geq j_0$, then we have

$$\begin{aligned}
&\left|\frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \Delta_{m_i n_j}^{m_{i+1} n_{j+1}} S - L\right| = \frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \left|\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} S - L \Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A\right| \\
&= \frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \big|(S_{m_{i+1} n_{j+1}} - L A_{m_{i+1} n_{j+1}}) - (S_{m_i n_{j+1}} - L A_{m_i n_{j+1}}) \\
&\qquad - (S_{m_{i+1} n_j} - L A_{m_{i+1} n_j}) + (S_{m_i n_j} - L A_{m_i n_j})\big| \\
&\leq \frac{A_{m_{i+1} n_{j+1}}}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \left(|Z_{m_{i+1} n_{j+1}} - L| + |Z_{m_i n_{j+1}} - L| + |Z_{m_{i+1} n_j} - L| + |Z_{m_i n_j} - L|\right) \\
&< 4\varepsilon \frac{A_{m_{i+1} n_{j+1}}}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} = 4\varepsilon \frac{B_{m_{i+1}}}{B_{m_{i+1}} - B_{m_i}} \frac{C_{n_{j+1}}}{C_{n_{j+1}} - C_{n_j}} \leq 16\varepsilon. \qquad (3.1)
\end{aligned}$$

Now, turn to the boundedness. Similarly as above

$$\begin{aligned}
\left|\frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \Delta_{m_i n_j}^{m_{i+1} n_{j+1}} S\right| &\leq \frac{B_{m_{i+1}}}{B_{m_{i+1}} - B_{m_i}} \frac{C_{n_{j+1}}}{C_{n_{j+1}} - C_{n_j}} \left(|Z_{m_{i+1} n_{j+1}}| + |Z_{m_i n_{j+1}}|\right. \\
&\qquad \left. + |Z_{m_{i+1} n_j}| + |Z_{m_i n_j}|\right) \leq \text{const.}, \qquad (3.2)
\end{aligned}$$

because $(Z_{m_i n_j})$ is bounded. Inequalities (3.1) and (3.2) imply (2.7).

Now, we turn to sufficiency. Assume that (2.7) is satisfied, that is, there exist integers i_0, j_0 such that

$$\left|\frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \Delta_{m_i n_j}^{m_{i+1} n_{j+1}} S - L\right| < \varepsilon \quad \text{for all} \quad i \geq i_0, j \geq j_0, \qquad (3.3)$$

furthermore $\left(\frac{1}{\Delta_{m_i n_j}^{m_{i+1} n_{j+1}} A} \Delta_{m_i n_j}^{m_{i+1} n_{j+1}} S\right)$ is a bounded sequence. If $i \geq i_0$ and $j \geq j_0$, then $m_{i+1} > m_{i_0}$ and $n_{j+1} > n_{j_0}$, so

$$\begin{aligned}
&Z_{m_{i+1} n_{j+1}} - L \\
&= \frac{1}{A_{m_{i+1} n_{j+1}}} (S_{m_{i+1} n_{j+1}} - L A_{m_{i+1} n_{j+1}}) = \frac{1}{A_{m_{i+1} n_{j+1}}} \sum_{k=1}^{m_{i+1}} \sum_{l=1}^{n_{j+1}} a_{kl}(x_{kl} - L) \\
&= \frac{1}{A_{m_{i+1} n_{j+1}}} \left(\sum_{k=1}^{m_{i_0}} \sum_{l=1}^{n_{j_0}} a_{kl}(x_{kl} - L) + \sum_{k=m_{i_0}+1}^{m_{i+1}} \sum_{l=n_{j_0}+1}^{n_{j+1}} a_{kl}(x_{kl} - L) \right.
\end{aligned}$$

$$+\sum_{k=1}^{m_{i_0}}\sum_{l=n_{j_0}+1}^{n_{j+1}} a_{kl}(x_{kl}-L)+\sum_{k=m_{i_0}+1}^{m_{i+1}}\sum_{l=1}^{n_{j_0}} a_{kl}(x_{kl}-L)\Bigg) \tag{3.4}$$

for all $i \geq i_0, j \geq j_0$.

Consider the first term in (3.4). Since $\frac{1}{A_{m_{i+1}n_{j+1}}} \to 0$, as $i \to \infty, j \to \infty$, then there exist integers $i_1 \geq i_0$ and $j_1 \geq j_0$, such that

$$\frac{1}{A_{m_{i+1}n_{j+1}}}\left|\sum_{k=1}^{m_{i_0}}\sum_{l=1}^{n_{j_0}} a_{kl}(x_{kl}-L)\right| < \varepsilon \quad \text{for all} \quad i \geq i_1, j \geq j_1. \tag{3.5}$$

Now, turn to the secont term in (3.4). If $i \geq k$, then

$$\frac{B_{m_{k+1}}-B_{m_k}}{B_{m_{i+1}}} = \frac{B_{m_{k+1}}-B_{m_k}}{B_{m_{k+1}}}\frac{B_{m_{k+1}}}{B_{m_{k+2}}}\frac{B_{m_{k+2}}}{B_{m_{k+3}}}\cdot\dots\cdot\frac{B_{m_i}}{B_{m_{i+1}}} \leq \left(\frac{1}{2}\right)^{i-k}.$$

Similarly, if $j \geq l$, then

$$\frac{C_{n_{l+1}}-C_{n_l}}{C_{n_{j+1}}} \leq \left(\frac{1}{2}\right)^{j-l}.$$

Hence we get from (3.3)

$$\begin{aligned}
&\frac{1}{A_{m_{i+1}n_{j+1}}}\left|\sum_{k=m_{i_0}+1}^{m_{i+1}}\sum_{l=n_{j_0}+1}^{n_{j+1}} a_{kl}(x_{kl}-L)\right| \\
&= \frac{1}{A_{m_{i+1}n_{j+1}}}\left|\sum_{k=i_0}^{i}\sum_{l=j_0}^{j}\sum_{s=m_k+1}^{m_{k+1}}\sum_{t=n_l+1}^{n_{l+1}} a_{st}(x_{st}-L)\right| \\
&= \left|\sum_{k=i_0}^{i}\sum_{l=j_0}^{j}\frac{B_{m_{k+1}}-B_{m_k}}{B_{m_{i+1}}}\frac{C_{n_{l+1}}-C_{n_l}}{C_{n_{j+1}}}\left(\frac{1}{\Delta_{m_k n_l}^{m_{k+1}n_{l+1}}A}\Delta_{m_k n_l}^{m_{k+1}n_{l+1}}S-L\right)\right| \\
&< \varepsilon\sum_{k=i_0}^{i}\left(\frac{1}{2}\right)^{i-k}\sum_{l=j_0}^{j}\left(\frac{1}{2}\right)^{j-l} < 4\varepsilon \quad \text{for all} \quad i \geq i_0, j \geq j_0.
\end{aligned} \tag{3.6}$$

For the third term in (3.4) we have

$$\begin{aligned}
&\frac{1}{A_{m_{i+1}n_{j+1}}}\left|\sum_{k=1}^{m_{i_0}}\sum_{l=n_{j_0}+1}^{n_{j+1}} a_{kl}(x_{kl}-L)\right| \\
&= \frac{1}{A_{m_{i+1}n_{j+1}}}\left|\sum_{k=0}^{i_0-1}\sum_{l=j_0}^{j}\sum_{s=m_k+1}^{m_{k+1}}\sum_{t=n_l+1}^{n_{l+1}} a_{st}(x_{st}-L)\right| \\
&= \left|\sum_{k=0}^{i_0-1}\sum_{l=j_0}^{j}\frac{B_{m_{k+1}}-B_{m_k}}{B_{m_{i+1}}}\frac{C_{n_{l+1}}-C_{n_l}}{C_{n_{j+1}}}\left(\frac{1}{\Delta_{m_k n_l}^{m_{k+1}n_{l+1}}A}\Delta_{m_k n_l}^{m_{k+1}n_{l+1}}S-L\right)\right|
\end{aligned}$$

$$\leq \frac{1}{B_{m_{i+1}}} \sum_{k=0}^{i_0-1} (B_{m_{k+1}} - B_{m_k}) \sum_{l=j_0}^{j} \left(\frac{1}{2}\right)^{j-l} \text{const.}$$

$$\leq \text{const.} \frac{1}{B_{m_{i+1}}} B_{m_{i_0}} \sum_{k=0}^{i_0-1} \frac{B_{m_{k+1}} - B_{m_k}}{B_{m_{i_0}}} \leq \text{const.} \frac{B_{m_{i_0}}}{B_{m_{i+1}}} \sum_{k=0}^{i_0-1} \left(\frac{1}{2}\right)^{i_0-1-k}$$

$$\leq \text{const.} \frac{B_{m_{i_0}}}{B_{m_{i+1}}} \to 0, \quad \text{as} \quad i \to \infty.$$

Hence, there exists $i_2 \geq i_1$ such that

$$\frac{1}{A_{m_{i+1}n_{j+1}}} \left| \sum_{k=1}^{m_{i_0}} \sum_{l=n_{j_0}+1}^{n_{j+1}} a_{kl}(x_{kl} - L) \right| < \varepsilon \quad \text{for all} \quad i \geq i_2, j \geq j_0. \tag{3.7}$$

Similarly, for the fourth term in (3.4) we obtain that there exists $j_2 \geq j_1$ such that

$$\frac{1}{A_{m_{i+1}n_{j+1}}} \left| \sum_{k=m_{i_0}+1}^{m_{i+1}} \sum_{l=1}^{n_{j_0}} a_{kl}(x_{kl} - L) \right| < \varepsilon \quad \text{for all} \quad i \geq i_0, j \geq j_2. \tag{3.8}$$

By (3.4)–(3.8), we have

$$|Z_{m_{i+1}n_{j+1}} - L| < 7\varepsilon \quad \text{for all} \quad i \geq i_2, j \geq j_2. \tag{3.9}$$

Finally, turn to the proof of boundedness.

$$|Z_{m_i n_j}| = \frac{1}{A_{m_i n_j}} \left| \sum_{k=1}^{m_i} \sum_{l=1}^{n_j} a_{kl} x_{kl} \right| = \frac{1}{A_{m_i n_j}} \left| \sum_{k=0}^{i-1} \sum_{l=0}^{j-1} \Delta_{m_k n_l}^{m_{k+1} n_{l+1}} S \right|$$

$$= \left| \sum_{k=0}^{i-1} \sum_{l=0}^{j-1} \frac{B_{m_{k+1}} - B_{m_k}}{B_{m_i}} \frac{C_{n_{l+1}} - C_{n_l}}{C_{n_j}} \frac{1}{\Delta_{m_k n_l}^{m_{k+1} n_{l+1}} A} \Delta_{m_k n_l}^{m_{k+1} n_{l+1}} S \right|$$

$$\leq \text{const.} \sum_{k=0}^{i-1} \frac{B_{m_{k+1}} - B_{m_k}}{B_{m_i}} \sum_{l=0}^{j-1} \frac{C_{n_{l+1}} - C_{n_l}}{C_{n_j}} \leq 4 \cdot \text{const.}$$

This inequality and (3.9) imply (2.6). Thus the theorem is proved. □

Proof of Theorem 2.4. Let ε be a fixed positive real number. First we prove the necessity. Assume that (2.8) is satisfied, that is, there exist integers M_0, N_0 such that

$$|Z_{mn} - L| < \varepsilon \quad \text{for all} \quad m \geq M_0, n \geq N_0, \tag{3.10}$$

furthermore (Z_{mn}) is a bounded sequence. Since we have

$$\sum_{k=m_i+1}^{m} \sum_{l=n_j+1}^{n} a_{kl}(x_{kl} - L) = A_{mn}(Z_{mn} - L) - A_{m_i n}(Z_{m_i n} - L)$$

$$- A_{mn_j}(Z_{mn_j} - L) + A_{m_in_j}(Z_{m_in_j} - L),$$

if $m > m_i$ and $n > n_j$, hence the ratio on the left-hand side in (2.9) is less than or equal to

$$\frac{A_{m_{i+1}n_{j+1}}}{\Delta_{m_in_j}^{m_{i+1}n_{j+1}}A}\Bigg(\max_{\substack{m_i<m\le m_{i+1}\\ n_j<n\le n_{j+1}}}|Z_{mn} - L| + \max_{n_j<n\le n_{j+1}}|Z_{m_in} - L| + \max_{m_i<m\le m_{i+1}}|Z_{mn_j} - L| + |Z_{m_in_j} - L|\Bigg). \tag{3.11}$$

There exist integers i_0, j_0 such that if $i \ge i_0$ and $j \ge j_0$, than $m_i \ge M_0$ and $n_j \ge N_0$. So (3.10) and (3.11) imply, that the ratio on the left-hand side in (2.9) is less than

$$\frac{A_{m_{i+1}n_{j+1}}}{\Delta_{m_in_j}^{m_{i+1}n_{j+1}}A}4\varepsilon \le 16\varepsilon \quad \text{for all} \quad i \ge i_0, j \ge j_0. \tag{3.12}$$

On the other hand, since (Z_{mn}) is a bounded sequence, so by (3.11), the ratio on the left-hand side in (2.9) is less than or equal to

$$\frac{A_{m_{i+1}n_{j+1}}}{\Delta_{m_in_j}^{m_{i+1}n_{j+1}}A}4\cdot\text{const.} \le 16\cdot\text{const.} \quad \text{for all} \quad i, j.$$

This fact and (3.12) imply (2.9).

Now we turn to sufficiency. Assume that (2.9) is satisfied. The ratio on the left-hand side in (2.9) is greater than or equal to

$$\frac{1}{\Delta_{m_in_j}^{m_{i+1}n_{j+1}}A}\left|\sum_{k=m_i+1}^{m_{i+1}}\sum_{l=n_j+1}^{n_{j+1}}a_{kl}(x_{kl} - L)\right| = \left|\frac{1}{\Delta_{m_in_j}^{m_{i+1}n_{j+1}}A}\Delta_{m_in_j}^{m_{i+1}n_{j+1}}S - L\right|,$$

so (2.7) is satisfied. Now, applying Theorem 2.3, we get that (2.6) is true. In the following parts of the proof, for fixed integers m, n let i, j be integers, such that

$$m_i < m \le m_{i+1} \quad \text{and} \quad n_j < n \le n_{j+1}.$$

We have

$$\begin{aligned}
Z_{mn} - L &= \frac{1}{A_{mn}}\sum_{k=1}^{m}\sum_{l=1}^{n}a_{kl}(x_{kl} - L)\\
&= \frac{1}{A_{mn}}\sum_{k=1}^{m_i}\sum_{l=1}^{n_j}a_{kl}(x_{kl} - L) + \frac{1}{A_{mn}}\sum_{k=m_i+1}^{m}\sum_{l=n_j+1}^{n}a_{kl}(x_{kl} - L)\\
&\quad + \frac{1}{A_{mn}}\sum_{k=m_i+1}^{m}\sum_{l=1}^{n_j}a_{kl}(x_{kl} - L) + \frac{1}{A_{mn}}\sum_{k=1}^{m_i}\sum_{l=n_j+1}^{n}a_{kl}(x_{kl} - L).
\end{aligned} \tag{3.13}$$

Consider the absolute values of all terms of this sum. For the first term, from (2.6) we get that

$$\frac{1}{A_{mn}}\left|\sum_{k=1}^{m_i}\sum_{l=1}^{n_j} a_{kl}(x_{kl}-L)\right|$$
$$=\frac{A_{m_in_j}}{A_{mn}}|Z_{m_in_j}-L| \le |Z_{m_in_j}-L| \xrightarrow{\text{b}} 0, \quad \text{as} \quad m,n\to\infty. \tag{3.14}$$

We shall use the following relations for the coefficients.

$$\frac{\Delta_{m_in_j}^{m_{i+1}n_{j+1}}A}{A_{mn}} = \frac{(B_{m_{i+1}}-B_{m_i})(C_{n_{j+1}}-C_{n_j})}{B_mC_n} \le \frac{B_{m_{i+1}}}{B_{m_i+1}}\frac{C_{n_{j+1}}}{C_{n_j+1}}$$
$$=\frac{B_{m_{i+1}-1}}{B_{m_i+1}}\left(1+\frac{b_{m_{i+1}}}{B_{m_{i+1}-1}}\right)\frac{C_{n_{j+1}-1}}{C_{n_j+1}}\left(1+\frac{c_{n_{j+1}}}{C_{n_{j+1}-1}}\right)$$
$$\le 4\left(1+\frac{b_{m_{i+1}}}{B_{m_{i+1}-1}}\right)\left(1+\frac{c_{n_{j+1}}}{C_{n_{j+1}-1}}\right)\le \text{const}. \tag{3.15}$$

To see the above relation, we mention that

$$\frac{B_{m-1}}{b_m}+1=\frac{B_{m-1}+b_m}{b_m}=\frac{B_m}{b_m}\ge 1+\delta,$$

because of the assumptions of the theorem. Therefore (b_m/B_{m-1}) is a bounded sequence. Similarly (c_n/C_{n-1}) is a bounded sequence, too.

Consider the second term in (3.13). From (3.15) and (2.9) we get that

$$\frac{1}{A_{mn}}\left|\sum_{k=m_i+1}^{m}\sum_{l=n_j+1}^{n} a_{kl}(x_{kl}-L)\right|$$
$$\le \frac{\Delta_{m_in_j}^{m_{i+1}n_{j+1}}A}{A_{mn}}\frac{1}{\Delta_{m_in_j}^{m_{i+1}n_{j+1}}A}\max_{\substack{m_i<t\le m_{i+1}\\ n_j<s\le n_{j+1}}}\left|\sum_{k=m_i+1}^{t}\sum_{l=n_j+1}^{s} a_{kl}(x_{kl}-L)\right| \xrightarrow{\text{b}} 0,$$
$$\text{as} \quad m,n\to\infty. \tag{3.16}$$

Now turn to the third and fourth terms on the left hand side of (3.13). With notation

$$\Phi_{it}:=\frac{1}{\Delta_{m_in_{t-1}}^{m_{i+1}n_t}A}\max_{m_i<s\le m_{i+1}}\left|\sum_{k=m_i+1}^{s}\sum_{l=n_{t-1}+1}^{n_t} a_{kl}(x_{kl}-L)\right|$$

we get that

$$\frac{1}{A_{mn}}\left|\sum_{k=m_i+1}^{m}\sum_{l=1}^{n_j} a_{kl}(x_{kl}-L)\right| \le \frac{1}{A_{mn}}\sum_{t=1}^{j}\left|\sum_{k=m_i+1}^{m}\sum_{l=n_{t-1}+1}^{n_t} a_{kl}(x_{kl}-L)\right|$$

$$\leq \frac{1}{A_{mn}} \sum_{t=1}^{j} \Delta_{m_i n_{t-1}}^{m_{i+1} n_t} A\,\Phi_{it} \leq \frac{B_{m_{i+1}} - B_{m_i}}{B_{m_i+1}} \sum_{t=1}^{j} \frac{C_{n_t} - C_{n_{t-1}}}{C_{n_j+1}} \Phi_{it}. \tag{3.17}$$

But

$$\frac{B_{m_{i+1}} - B_{m_i}}{B_{m_i+1}} < \frac{b_{m_{i+1}} + B_{m_i}}{B_{m_i+1}} < 1 + \frac{B_{m_{i+1}-1}}{B_{m_i}} \frac{b_{m_{i+1}}}{B_{m_{i+1}-1}} < 1 + 2\frac{b_{m_{i+1}}}{B_{m_{i+1}-1}},$$

which is bounded as we have already seen. Furthermore, for $t = 1, 2, \ldots, j$,

$$\frac{C_{n_t} - C_{n_{t-1}}}{C_{n_j+1}} = \frac{C_{n_t} - C_{n_{t-1}}}{C_{n_t}} \frac{C_{n_t}}{C_{n_{t+1}}} \frac{C_{n_{t+1}}}{C_{n_{t+2}}} \cdots \frac{C_{n_{j-1}}}{C_{n_j}} \frac{C_{n_j}}{C_{n_j+1}} \leq \left(\frac{1}{2}\right)^{j-t+1}.$$

Hence (3.17) implies that

$$\frac{1}{A_{mn}} \left| \sum_{k=m_i+1}^{m} \sum_{l=1}^{n_j} a_{kl}(x_{kl} - L) \right| \leq \text{const.} \sum_{t=1}^{j} \left(\frac{1}{2}\right)^{j-t} \Phi_{it}. \tag{3.18}$$

By (2.9), $\Phi_{it} \xrightarrow{\text{b}} 0$. This and (3.18) imply that the expression on the left-hand side in (3.18) is bounded. Moreover, there exist i_0, j_0 such that $\Phi_{it} < \varepsilon$ and at the same time $(1/2)^t < \varepsilon$ for all $i \geq i_0, t \geq j_0$. From these facts and applying that the sequence Φ_{it} is bounded, we get

$$\sum_{t=1}^{j} \left(\frac{1}{2}\right)^{j-t} \Phi_{it} = \sum_{t=1}^{j_0} \left(\frac{1}{2}\right)^{j-t} \Phi_{it} + \sum_{t=j_0+1}^{j} \left(\frac{1}{2}\right)^{j-t} \Phi_{it}$$
$$< \text{const.} \left(\frac{1}{2}\right)^{j/2} \sum_{t=1}^{j_0} \left(\frac{1}{2}\right)^{j/2-t} + 2\varepsilon < \text{const.}\varepsilon \quad \text{for all} \quad i \geq i_0, j \geq 2j_0.$$

So it follows from (3.18) that

$$\frac{1}{A_{mn}} \left| \sum_{k=m_i+1}^{m} \sum_{l=1}^{n_j} a_{kl}(x_{kl} - L) \right| \xrightarrow{\text{b}} 0, \quad \text{as} \quad m, n \to \infty. \tag{3.19}$$

By similar arguments, for the fourth term in (3.13), we have

$$\frac{1}{A_{mn}} \left| \sum_{k=1}^{m_i} \sum_{l=n_j+1}^{n} a_{kl}(x_{kl} - L) \right| \xrightarrow{\text{b}} 0, \quad \text{as} \quad m, n \to \infty. \tag{3.20}$$

Finally (3.13), (3.14), (3.16), (3.19) and (3.20) imply (2.8). Thus the theorem is proved. □

References

[1] Fazekas, I., Marcinkiewicz strong law of large numbers for B-valued random variables with multidimensional indices, Statistics and probability, Proc. 3rd Pannonian Symp., Visegrád/Hungary, 1982, Reidel, Dordrecht, 53–61, 1984.

[2] Fazekas, I., Klesov, O. I., A general approach to the strong law of large numbers, Theory of Probability Applications, 45/3, 568–583, 2000.

[3] Fekete, Á., Georgieva, I., Móricz, F., Characterizations of the convergence of harmonic averages of double numerical sequences, Mathematical Inequalities & Applications, Volume 14, no. 3, 555–573, 2011.

[4] Móricz, F., Stadtmüller, U., Characterization of the convergence of weighted averages of sequences and functions, Manuscript, 2011.

[5] Noszály, Cs., Tómács, T., A general approach to strong laws of large numbers for fields of random variables, Annales Univ. Sci. Budapest, 43, 61–78, 2000.

[6] Smythe, R. T., Strong laws of large numbers for r-dimensional arrays of random variables, Ann. Probability 1, no. 1, 164–170, 1973.

Annales Mathematicae et Informaticae
39 *(2012) pp. 83–108*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Compensated compactness and relaxation at the microscopic level*

József Fritz

Mathematical Institute, Budapest University of Technology and Economics
e-mail: `jofri@math.bme.hu`

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

This is a survey of some recent results on hyperbolic scaling limits. In contract to diffusive models, the resulting Euler equations of hydrodynamics develop shocks in a finite time. That is why the derivation of the macroscopic equations from a microscopic model requires a synthesis of probabilistic and PDE methods. In the case of two-component stochastic models with a hyperbolic scaling law the method of compensated compactness seems to be the only tool that we can apply. Since the associated *Lax entropies* are not preserved by the microscopic dynamics, a logarithmic Sobolev inequality is needed to evaluate *entropy production.* Extending the arguments of Shearer (1994) and Serre–Shearer (1994) to stochastic systems, the nonlinear wave equation of isentropic elastodynamics is derived as the hyperbolic scaling limit of the anharmonic chain with Ginzburg–Landau type random perturbations. The model of interacting exclusion of charged particles results in the Leroux system in a similar way. In the presence of an additional creation-annihilation mechanism the missing logarithmic Sobolev inequality is replaced by an associated relaxation scheme. In this case the uniqueness of the limit is also known.

Keywords: Anharmonic chain, Ginzburg–Landau model, interacting exclusions, creation and annihilation, hyperbolic scaling, vanishing viscosity limit, logarithmic Sobolev inequalities, Lax entropy pairs, compensated compactness, relaxation schemes.

MSC: Primary 60K31, secondary 82C22.

*Partially supported by Hungarian Science Foundation Grants K-60708 and K-100473.

1. Historical notes and references

The idea that the Euler equations of hydrodynamics ought to be derived from statistical mechanics goes back to Morrey (1955). He proposed a *scaling limit* to pass to the hyperbolic system of classical conservation laws when the number of particles goes to infinity. The natural scaling of mechanical and related asymmetric systems is *hyperbolic:* the microscopic time is speeded up at the same rate at which the size of the system goes to infinity. The theory of *diffusive scaling limits* seems to be more or less complete, see Kipnis–Landim (1989) for a comprehensive survey.[1] Here we concentrate on the hyperbolic scaling limit of stochastic systems. Various models are introduced, and the main ideas of several proofs are also outlined in the next sections. You shall see that progress in this direction is rather slow, there are many relevant open problems.

Basic principles: In theoretical physics it is commonly accepted that the *equilibrium states* of the microscopic system are specified by the *Boltzmann–Gibbs formalism,* and the evolved measure can be well approximated by means of such Gibbs states with space and time dependent parameters. This *principle of local equilibrium* is used then to determine the macroscopic flux of the conserved quantities of the underlying microscopic dynamics; this is the first crucial problem in the theory of *hydrodynamic limits* (HDL). However, a rigorous verification of any version of this principle is problematic because the standard argument is based on a strong form of the *ergodic hypothesis,* which amounts to a description of translation invariant stationary states of the microscopic system as superpositions of the equilibrium Gibbs random fields. This is certainly one of the hardest open problems of mathematics, it is much more difficult than the question of *metric transitivity* of the underlying stationary process, but it is much weaker than the claim of the principle of local equilibrium. A second principal difficulty in the theory of hyperbolic scaling limits comes from the complexity of the resulting macroscopic equations (conservation laws). The breakdown of the existence of global classical solutions is quite general, and the surviving weak solutions are usually not unique. The formation of the associated *shock waves* results in extremely strong fluctuations at the microscopic level, too. Concerning terminology and basic facts on HDL we refer to the textbooks by Spohn (1991) and Kipnis–Landim (1999), while to Hörmander (1997), Bressan (2000) and Dafermos (2005) on PDE theory.

Deterministic models: Of course, there exist some mechanical systems that admit explicit computations. However, the exactly solvable models of *one-dimensional hard rods* and *coupled harmonic oscillators* are not ergodic in the traditional sense. Besides the classical ones these systems admit a continuum of conservation laws, consequently the scaling limit of such models does not result in a closed system of a finite set of equations for the classical conservation laws, see the papers by Dobrushin and coworkers (1980, 1983, 1985). The treatment of more realistic

[1] More recent information can be found on the web site http://stokhos.shinshu-u.ac.jp/10thSALSIS/ of the 10-th Symposium on Stochastic Analysis of Large Scale Interacting System, Kochi (Japan) 2011.

mechanical systems is out of question, Sinai (1988) is the only scientist who dared to attack this issue. He claimed that the identification of the macroscopic flux does not require the strong ergodic hypothesis, the problem is still open.

Attractive systems: To avoid the hopeless issue of strong ergodicity of mechanical systems, stochastic models are only considered in the rest of the related literature on hydrodynamic limits. Appropriately chosen random effects regularize the dynamics, thus there is a good chance to identify the conservation laws and the associated stationary states of the microscopic system. The first result in this direction is due to Rost (1981), he managed to derive certain rarefaction wave solutions to the Burgers equation as HDL of the totally asymmetric simple exclusion process. Following some preliminary studies by various authors, a few years later Rezakhanlou (1991) extended his *coupling technique* for a large class of *attractive models.* Several more recent results in this direction are treated or mentioned by Kipnis–Landim (1989) and Bahadoran (2004). Although the appearance of shocks is not excluded, effective coupling in attractive models implies the *Kruzkov entropy condition* in a natural way, consequently the empirical process converges to the uniquely specified *weak entropy solution* of the associated single conservation law. We are mainly interested in the hydrodynamic limit of microscopic systems with two conservation laws, these are certainly not attractive.

Entropy and HDL in a smooth regime: Random effects might regularize even the classical dynamics in such a way that we have a description of stationary measures: translation invariant equilibrium states of finite specific entropy with respect to a given stationary measure are all superpositions of the classical equilibrium (Gibbs) states. As a next step, a fairly general theory of *asymptotic preservation of local equilibrium* has been initiated by Yau (1991). This means that if the initial distribution is close to local equilibrium in the sense of specific relative entropy, then this property remains in force as long as the macroscopic solution is smooth enough. His method has been extended to Hamiltonian dynamics[2] with conservative noise for continuous particle systems by Olla–Varadhan–Yau (1993). The hyperbolic (Euler) scaling limit yields the full set of the *compressible Euler equations.* The basic ideas of this approach are to be discussed in the next section.

The problem of shocks: In the case of a hyperbolic scaling limit the microscopic system simply does not have enough time to organize itself, even the asymptotic preservation of local equilibrium is a problematic issue in a regime of *shock waves.* Therefore the separation of the *slowly varying* conserved quantities from the other, *rapidly oscillating* ones is less transparent than in a smooth or diffusive regime.

The existence theory of parabolic equations or systems is based on the associated *energy inequalities,* and it is a quite natural idea of PDE theory to construct a parabolic approximation to the hyperbolic system of conservation laws by adding elliptic (viscid) terms to the right hand side of the equations under consideration. Since the related energy inequalities degenerate in this *small viscosity limit,* the standard compactness argument has to be replaced by a radically new technique

[2]The kinetic energy of the model is not the classical one because energy transport can not be controlled in that case.

called *compensated compactness,* see Hörmander (1997) or Dafermos (2005) with several further references.

The microscopic models of hydrodynamics imitate this approach, thus the situation is quite similar. The probabilistic a priori bounds we have in a diffusive scaling limit[3] do not work any more in case of a hyperbolic scaling limit. Therefore we have to extend the theory of compensated compactness to our microscopic systems, see Fritz (2001, 2004, 2011), Fritz–Tóth (2004), Fritz–Nagy (2006) and Bahadoran–Fritz–Nagy (2011). In this way we obtain convergence along subsequences to *weak solutions,* and the uniqueness of the limit ought to be the consequence of some additional information. The familiar *Lax entropy inequality* is only sufficient for weak uniqueness to a single conservation law. Unfortunately, in the case of systems the much deeper Oleinik type conditions of Bressan (2000) are required, and these strictly local bounds are not attainable by our present probabilistic techniques.

2. The anharmonic chain

It is perhaps the simplest mechanical system that exhibits a correct physical behavior, it is considered as a *microscopic model of one-dimensional elasticity.* The Hamiltonian of coupled oscillators of unit mass on $\mathbb{Z}$ reads as

$$H(\omega) := \sum_{k\in\mathbb{Z}} H_k(\omega), \quad H_k(\omega) := p_k^2/2 + V(q_{k+1} - q_k),$$

where $\omega = \{(p_k, q_k) : k \in \mathbb{Z}\}$ denotes a configuration of the infinite system, $p_k, q_k \in \mathbb{R}$ are the momentum (velocity) and position of the oscillator at site $k \in \mathbb{Z}$. In terms of the *deformation* variables $r_k := q_{k+1} - q_k$, the equations of motion read as

$$\dot{p_k} = V'(r_k) - V'(r_{k-1}) \quad \text{and} \quad \dot{r_k} = p_{k+1} - p_k \quad \text{for } k \in \mathbb{Z}; \tag{2.1}$$

in this formulation the interaction potential V needs not be symmetric. The existence of unique solutions in a space of configurations $\omega := \{(p_k, r_k) : k \in \mathbb{Z}\}$ with a sub-exponential growth is quite standard if V' is Lipschitz continuous, i.e. if V'' is bounded. The related iterative procedure shows also that the solutions of the infinite system can be well approximated by the solutions of its finite subsystems when the size of the finite system goes to infinity, see e.g. Fritz (2011) with further references.

Although (2.1) is a direct lattice approximation to the *p-system* $\partial_t u = \partial_x V'(v)$, $\partial_t v = \partial_x u$, its convergence is rather problematic. In PDE theory (2.1) is not considered as a stable numerical scheme for solving the p-system, thus we can not believe in its convergence. The right way of its regularization is suggested by the small viscosity approach, it is certainly not difficult to define stable approximation schemes in this way. However, the theory of hydrodynamic limits goes beyond numerical analysis as discussed below.

[3]See Fritz (1986), and Guo–Papanicolau–Varadhan (1988) for a more perfect treatment.

2.1. The compressible Euler equations

(2.1) reads as a *lattice system of conservation laws* for the total momentum $P := \sum p_k$, and for the total deformation $R := \sum r_k$, respectively: $\partial_t P = \partial_t R = 0$ are formal identities. Since $\partial_t H_k(\omega) = p_{k+1}V'(r_k) - p_k V'(r_{k-1})$ is a difference of currents, the total energy H is also preserved by the dynamics, therefore we expect to have three hydrodynamic equations: one for momentum, one for the deformation, and one for energy. In view of the principle of local equilibrium, the macroscopic fluxes of these conservative quantities are to be calculated by means of the stationary states of the dynamics.

Stationary states and thermodynamics: These are characterized by $\int \mathcal{L}_0\varphi \, d\lambda = 0$ for smooth *local functions* φ of a finite number of variables, where

$$\mathcal{L}_0\varphi := \sum_{k\in\mathbb{Z}} \left((V'(r_k) - V'(r_{k-1})) \frac{\partial\varphi}{\partial p_k} + (p_{k+1} - p_k) \frac{\partial\varphi}{\partial r_k} \right) \tag{2.2}$$

denotes the associated *Liouville operator.* Assuming $\lim V(x)/|x| = +\infty$ as $|x| \to +\infty$, it is easy to check that we have a three-parameter family $\lambda_{\beta,\pi,\gamma}$ of translation invariant product measures: $\beta > 0$ is the *inverse temperature,* $\pi \in \mathbb{R}$ denotes the *mean velocity,* and $\gamma \in \mathbb{R}$ is a *chemical potential.* Under $\lambda_{\beta,\pi,\gamma}$ the marginal Lebesgue density of any couple $(p_k, r_k) \sim (y, x)$ reads as $\exp(\gamma x - \beta I(y,x|\pi) - F(\beta,\gamma))$, where $I(y,x|\pi) := (y-\pi)^2/2 + V(x)$; the normalization

$$F(\beta,\gamma) := \log \iint_{\mathbb{R}^2} \exp\left(\gamma r - \beta I(y,x|\pi)\right) \, dy \, dx \tag{2.3}$$

is sometimes referred to as the *free energy.* Indeed, approximating the infinite system by its finite subsystems, it follows immediately that these product measures are really equilibrium states of (2.1). It is easy to see that $\mathcal{L}_0$ is antisymmetric with respect to any $\lambda_{\beta,\pi,\gamma}$.

Let us remark that there is a one-to-one correspondence between the parameters (β,π,γ) and the corresponding expected values (h,u,v) of the conservative quantities H_k, p_k and r_k with respect to $\lambda_{\beta,\pi,\gamma}$. It is plain that $u := \int p_k \, d\lambda_{\beta,\pi,\gamma} = \pi$ is the mean velocity. By a direct computation we see also that the equilibrium mean of the *internal energy* $I_k := I(p_k, r_k|\pi)$ at one site is given by $\chi := \int I_k \, d\lambda_{\beta,\pi,\gamma} = -F'_\beta(\beta,\gamma)$, thus the equilibrium mean of the total energy $H_k = p_k^2/2 + V(r_k)$ is just $h := \chi + \pi^2/2$, while $v = F'_\gamma(\beta,\gamma) = \int r_k \, d\lambda_{\beta,\pi,\gamma}$ is the mean deformation. Integrating by parts we obtain $\int V'(r_k) \, d\lambda_{\beta,\pi,\gamma} = \gamma/\beta$ for the equilibrium expectation of V'. The parameters β and γ can be expressed in terms of the *thermodynamical entropy*

$$S(\chi,v) := \sup\left\{\gamma v - \beta\chi - F(\beta,\gamma) : \beta > 0, \gamma \in \mathbb{R}\right\} \tag{2.4}$$

as follows. Since S is the convex conjugate of F, we have $\gamma = S'_v(\chi,v)$ and $\beta = -S'_\chi(\chi,v)$ if $v = F'_\gamma(\beta,\gamma)$ and $\chi = -F'_\beta(\beta,\gamma)$.

The hyperbolic scaling limit: We are interested in the asymptotic behavior of the *empirical processes* $u_\varepsilon(t,x) := p_k(t/\varepsilon)$, $v_\varepsilon(t,x) := r_k(t/\varepsilon)$ and $h_\varepsilon(t,x) :=$

$H_k(\omega(t/\varepsilon))$ if $|k\varepsilon - x| < \varepsilon/2$, as $0 < \varepsilon \to 0$. Of course it is assumed that at time zero these processes converge, at least in a weak sense to the corresponding initial values of the hydrodynamic equations.

In view of the physical principle of local equilibrium, the macroscopic currents of the conservative quantities should be calculated by means of a product measure of type $\lambda_{\beta,\pi,\gamma}$ with parameters depending on time and space. In this framework $\gamma/\beta = \int V'(r_k)\, d\lambda_{\beta,\pi,\gamma}$ is the mean current of momentum, and $\pi\gamma/\beta = \int p_k V'(r_{k-1})\, d\lambda_{\beta,\pi,\gamma}$ is the mean current of energy, consequently a formal calculation results in the triplet of compressible Euler equations:

$$\partial_t u = \partial_x J(\chi, v), \quad \partial_t v = \partial_x u \quad \text{and} \quad \partial_t h = \partial_x(uJ(\chi, v)), \tag{2.5}$$

where $J(\chi, v) := \gamma/\beta = -S'_v(\chi, v)/S'_\chi(\chi, v)$ and $\chi = h - u^2/2$, see Chen–Dafermos (1995) and Fritz (2001). Therefore $\partial_t \chi = J(\chi, v)\partial_x u$ and $\partial_t S(\chi, v) = 0$ along classical solutions, but we have to keep in mind that this system develops shock waves in a finite time.

2.2. Stochastic perturbations

As we have emphasized before, we are not able to materialize the heuristic derivation of the compressible Euler equations, the dynamics of the anharmonic chain should be regularized by a well chosen noise. There are several plausible tricks, we are going to consider Markov processes generated by an operator $\mathcal{L} = \mathcal{L}_0 + \sigma\, \mathcal{G}$, where $\mathcal{L}_0$ is the Liouville operator, while the Markov generator $\mathcal{G}$ is symmetric in equilibrium. Here $\sigma > 0$ may depend on the scaling parameter $\varepsilon > 0$, and $\varepsilon\sigma(\varepsilon)$ is interpreted as the coefficient of *macroscopic viscosity.* We are assuming that $\varepsilon\sigma(\varepsilon) \to 0$ as $\varepsilon \to 0$, then the effect of the symmetric component $\sigma\mathcal{G}$ diminishes in the limit. Our philosophy consists in adapting the *vanishing viscosity approach* of PDE theory to the microscopic theory of hydrodynamics. In a regime of shocks an additional technical condition: $\varepsilon\sigma^2(\varepsilon) \to +\infty$ is also needed.

Random exchange of velocities: As far as I understand, this is the weakest but still effective conservative noise. At the bonds of $\mathbb{Z}$ we have independently running clocks with exponential waiting times of parameter 1, and we exchange the velocities at the ends of the bond when the clock rings. The generator $\mathcal{G} = \mathcal{G}_{ep}$ of this exchange mechanism is acting on local functions as

$$\mathcal{G}_{ep}\varphi(\omega) = \sum_{k\in\mathbb{Z}} \left(\varphi(\omega^{k,k+1}) - \varphi(\omega)\right), \tag{2.6}$$

where $\omega^{k,k+1}$ denotes the configuration obtained from $\omega = \{(p_j, r_j)\}$ by exchanging p_k and p_{k+1}, the rest of ω remains unchanged. It is plain that $P = \sum p_k$, $R = \sum r_k$ and the total energy H are formally preserved by $\mathcal{G}_{ep}$, and the product measures $\lambda_{\beta,\pi,\gamma}$ are all stationary states of the Markov process generated by $\mathcal{L} := \mathcal{L}_0 + \sigma\, \mathcal{G}_{ep}$ if $\sigma > 0$.

This model was introduced by Fritz–Funaki–Lebowitz (1994), where the strong ergodic hypothesis is proven for lattice models with two conservation laws. The proof applies also in our case without any essential modification, see below. The *relative entropy* $\mathsf{S}[\mu|\lambda]$ of two probability measures on the same space is defined by $\mathsf{S} := \int \log f \, d\mu$, provided that $f = d\mu/d\lambda$ and the integral does exist; $\mathsf{S}[\mu|\lambda] = +\infty$ otherwise.[4] Let μ_n denote the joint distribution of the variables $\{(p_k, r_k) : |k| \le n\}$ with respect to μ, as a reference measure we choose $\lambda := \lambda_{1,0,0}$, and $f_n := d\mu_n/d\lambda$.

Theorem 2.1. *Suppose that μ is a translation invariant stationary measure of the process generated by $\mathcal{L} = \mathcal{L}_0 + \sigma\mathcal{G}_{ep}$. If the specific entropy of μ is finite, i.e. $\mathsf{S}[\mu_n|\lambda] = O(n)$, then μ is contained in the weak closure of the convex hull of our set $\{\lambda_{\beta,\pi,\gamma}\}$ of stationary product measures.*

On the ideas of the proof: The basic steps can be outlined as follows, for technical details see Theorems 2.4 and 3.1 of our paper cited above, or an improved version of the notes by Bernardin–Olla (2010). Since $\mathsf{S}[\mu_n|\lambda]$ is constant in a stationary regime, $\int \mathcal{L} \log f_n \, d\mu = 0$. The contribution of $\mathcal{L}_0$ consists of two boundary terms only because $\mathcal{L}_0$ is antisymmetric, while $-D_n[\mu|\lambda]$ is the essential part of the contribution of the symmetric $\mathcal{G}_{ep}$, where $D_n := -\int f_n \mathcal{G}_{ep} \log f_n \, d\lambda$. Due to the translation invariance of μ we see immediately that $(1/n)D_n[\mu|\lambda] \to 0$ as $n \to +\infty$. Moreover, $D_n \ge 0$ is a convex functional of μ, thus $D_{n+m} \ge D_n + D_m$, whence even $D_n[\mu|\lambda] = 0$ follows for all $n \in \mathbb{N}$. Therefore μ is symmetric with respect to any exchange of velocities, i.e. $\int \mathcal{G}_{ep}\varphi \, d\mu = 0$ is an identity, consequently the stationary Liouville equation $\int \mathcal{L}_0\varphi \, d\mu = 0$ also holds true.

Let $\phi(p)$ and $\psi(r)$ denote local functions depending only on the velocity and the deformation variables $p := \{p_j\}$, $r := \{r_j\}$, respectively. If φ_k and ψ_k are their translates by $k \in \mathbb{Z}$, then

$$\int \phi_k(p)\psi_k(r) \, d\mu = \int \phi_k(p)\psi_0(r) \, d\mu = \frac{1}{l}\sum_{j=0}^{l-1} \int \phi_{k+j}(p)\psi_0(r) \, d\mu$$

are identities, and the law of large numbers applies to the right hand side. For instance we see that given r, the conditional distribution of p is exchangeable, and it does not depend on the individual deformation variables r_j, thus the conditional expectation of any p_j is an invariant and tail measurable function $u \sim \pi$. Similarly, the conditional variance Q of velocities defines our first parameter, the inverse temperature β by $\beta := 1/Q$, it is an invariant function, too. Moreover, the entropy condition implies $\beta > 0$ almost surely.

On the other hand, for $\varphi = \psi(r)(p_k - u)$ the stationary Liouville equation yields

$$\int \psi(r)(V'(r_k) - V'(r_{k-1}) \, d\mu = \sum_{j\in\mathbb{Z}} \int \frac{\partial\psi(r)}{\partial r_j} \, (p_k - u)(p_{j+1} - p_j)) \, d\mu.$$

In view of the De Finetti–Hewitt–Savage theorem, the velocities are conditionally independent when r is given, consequently

$$\int \psi(r) \, (V'(r_k) - V'(r_{k-1})) \, d\mu = \int \frac{1}{\beta}\left(\frac{\partial\psi}{\partial r_k} - \frac{\partial\psi}{\partial r_{k-1}}\right) d\mu.$$

[4] The *entropy inequality* $\int \varphi \, d\mu \le \mathsf{S}(\mu|\lambda) + \log\int e^{\varphi} \, d\lambda$ is used in several probabilistic computations; $\varphi = \log f$ is the condition equality.

Now an obvious summation trick lets the law of large numbers work, whence

$$\int \psi(r)(V'(r_k) - \gamma)\, d\mu = \int \frac{1}{\beta}\,\frac{\partial \psi(r)}{\partial r_k}\, d\mu,$$

where the parameter γ is again invariant and tail measurable because it is the limit of the arithmetic averages of the $V'(r_j)$ variables. The stationary Liouville equation has been separated (localized) in this way, therefore the distribution of the deformation variables can be identified. Indeed, as β does not depend on r_k, the desired statement reduces to the differential characterization of the Lebesque measure by integrating by parts. In the case of velocities a similar argument results in

$$\int \phi(p)(p_k - \pi)\, d\mu = \int \frac{1}{\beta}\,\frac{\partial \phi(p)}{\partial p_k}\, d\mu,$$

consequently if the tail field is given, then the conditional distribution of $\omega = \{(p_k, r_k)\}$ under μ is just $\lambda_{\beta,\pi,\gamma}$.

It is interesting to note that Theorem 2.1 is not true for finite systems because the cited theorem on exchangeable variables applies to infinite sequences only.

Physical viscosity with thermal noise: Another popular model is obtained by adding a *Ginzburg-Landau type conservative noise* to the equations of velocities:

$$\begin{aligned} dp_k &= (V'(r_k) - V'(r_{k-1}))\, dt + \sigma\,(p_{k+1} + p_{k-1} - 2p_k)\, dt \\ &\quad + \sqrt{2\sigma}\,(dw_k - dw_{k-1}), \quad dr_k = (p_{k+1} - p_k)\, dt, \quad k \in \mathbb{Z}, \end{aligned} \tag{2.7}$$

where $\sigma > 0$ is a given constant, and $\{w_k : k \in \mathbb{Z}\}$ is a family of independent Wiener processes. Due to $V'' \in L^\infty$, the existence of unique strong solutions to this infinite system of stochastic differential equations is not a difficult issue, see e.g. Fritz (2001) with further references. The generator of the Markov process defined in this way can again be written as $\mathcal{L} := \mathcal{L}_0 + \sigma \mathcal{G}_p$, where $\mathcal{G}_p$ is now an elliptic operator. Total energy is not preserved any more, and a thermal equilibrium of unit temperature is maintained by the noise. It is easy to check that the product measures $\lambda_{\pi,\gamma} := \lambda_{1,\pi,\gamma}$ are all stationary, thus (2.5) reduces to the *p-system* (nonlinear sound equation) of elastodynamics:

$$\partial_t u = \partial_x S'(v) \quad \text{and} \quad \partial_t v = \partial_x u, \quad \text{that is} \quad \partial_t^2 v = \partial_x^2 S'(v) \tag{2.8}$$

because $\int V'(r_k)\, d\lambda_{\pi,\gamma} = \gamma = S'(v)$ if $\int r_k\, d\lambda_{\pi,\gamma} = v = F'(\gamma)$, where

$$S(v) := \sup_{\gamma}\{\gamma v - F(\gamma)\}; \quad F(\gamma) := \log \int_{-\infty}^{\infty} \exp(\gamma x - V(x))\, dx.$$

Let us remark that both F and S are infinitely differentiable, and $S''(v) = 1/F''(\gamma)$ is strictly positive and bounded.

The verification of the strong ergodic hypothesis is similar, but considerably simpler than in the previous case:

Theorem 2.2. *Translation invariant stationary measures of finite specific entropy are superpositions of our product measures* $\lambda_{\pi,\gamma}$.

For a complete proof see Theorem 13.1 in the notes by Fritz (2001). HDL of this model follows easily by the relative entropy argument of Yau. At a level $\varepsilon > 0$ of scaling $\mu_{t,\varepsilon,n}$ denotes the true distribution of the variables $\{(p_k(t), r_k(t)) : |k| \leq n\}$, and $\lambda_{t,\varepsilon} \sim \lambda_{\pi,\gamma}$ is a product measure with parameters $\pi = \pi_k(t,\varepsilon)$ and $\gamma = \gamma_k(t,\varepsilon)$ depending on space and time. We say that asymptotic local equilibrium holds true on the interval $[0,T]$ if we have a family $\{\lambda_{t,\varepsilon} : t \leq T/\varepsilon, \varepsilon \in (0,1]\}$ such that for all $\tau \leq T$

$$\lim_{\varepsilon \to 0} \sup_{n \geq 1/\varepsilon} \frac{\mathsf{S}[\mu_{\tau/\varepsilon,\varepsilon,n}|\lambda_{\tau/\varepsilon,\varepsilon}]}{2n+1} = 0. \tag{2.9}$$

Postulate this for $\tau = 0$, and suppose also that the prescribed initial values give rise to a continuously differentiable solution (u,v) to (2.8) on $[0,T], T > 0$. Then the approximate local equilibrium (2.9) remains in force for $\tau \leq T$, at least if the parameters π_k and γ_k of $\lambda_{t,\varepsilon}$ are chosen in a clever way, namely as they are predicted by the hydrodynamic equations (2.8). For example, we can put $\pi_k(t,\varepsilon) := u(\tau/\varepsilon, k/\varepsilon)$ and $\gamma_k(t,\varepsilon) := S'(v(\tau/\varepsilon, k/\varepsilon))$ if $t = \tau/\varepsilon$, but solutions to a discretized version of (2.8) can also be used. Therefore the empirical processes u_ε and v_ε converge in a weak sense to that smooth solution of (2.8). Indeed, the entropy inequality implies $-\log \lambda[A]\,\mu[A] \leq \mathsf{S}[\mu|\lambda] + \log 2$ for any event A, and in an exact local equilibrium $\lambda_{t,\varepsilon}$ the weak law of large numbers holds true with an exponential rate of convergence. Consequently (2.9) implies

Theorem 2.3. *Under the conditions listed above we have*

$$\lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} \varphi(x) u_\varepsilon(\tau,x)\,dx = \int_{-\infty}^{\infty} \varphi(x) u(\tau,x)\,dx$$

and

$$\lim_{\varepsilon \to 0} \int_{-\infty}^{\infty} \psi(x) v_\varepsilon(\tau,x)\,dx = \int_{-\infty}^{\infty} \psi(x) v(\tau,x)\,dx$$

in probability for all continuous φ, ψ *with compact support if* $\tau \leq T$, *where* (u,v) *is the preferred smooth solution to* (2.8).

The main ideas concerning the derivation of (2.9) are discussed in the next subsection, for a complete proof see that of Theorem 14.1 in Fritz (2001). In contrast to the result of Olla–Varadhan–Yau (1993) and other related papers, see also Theorem 2.4 below, the statement is not restricted to the periodic setting; the scaling limit here is considered on the infinite line. Such an extension of the original argument is based on the observation that the boundary terms of $\partial_t \mathsf{S}[\mu_{t,\varepsilon,n}|\lambda_{t,\varepsilon}]$ can be controlled by the associated *Dirichlet form* consisting of the volume terms of $\partial_t \mathsf{S}$. The first proof in this direction is due to Fritz (1990), see also Fritz–Nagy (2006), Bahadoran–Fritz–Nagy (2011) and Fritz (2011).

2.3. Derivation of the Euler equations in a smooth regime

Here we are going to outline Yau's method for the anharmonic chain with random exchange of velocities. The argument is similar but much more transparent than that of Olla–Varadhan–Yau (1993). The derivation of (2.8) is easier, its main steps are also included in the next coming calculations. Since the noise is not strong enough to control the flux of the relative entropy, we have to formulate the problem in a periodic setting: $p_k(0) = p_{k+n}(0)$ and $r_k(0) = r_{k+n}(0)$ for all k with some $n \in \mathbb{N}$. The evolved configuration remains periodic for all times, which means that the system can be considered on the discrete circle of length $n \to +\infty$. The coefficient $\sigma > 0$ can be kept fixed during the procedure of scaling because the only role of the exchange mechanism is to ensure the strong ergodic hypothesis. At a level $\varepsilon = 1/n$ of scaling let $\mu_{t,n}$ denote the evolved measure, and consider the *local equilibrium distributions* $\lambda_{t,n}$ of type $\lambda_{\beta,\pi,\gamma}$ with parameters depending on space and time: $\beta = \beta_k(t,n)$, $\pi = \pi_k(t,n)$ and $\gamma = \gamma_k(t,n)$.

Theorem 2.4. *Suppose that $(1/n)\mathsf{S}[\mu_{0,n}|\lambda_{0,n}] \to 0$ as $n \to +\infty$, and the related initial values determine a smooth solution (u,v,h) to (2.5) on the interval $[0,T]$ of time such that $\beta = -S'_\chi(\chi,v)$ remains strictly positive. Then*

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \psi(x) z_n(t,x)\,dx = \int_{-\infty}^{\infty} \psi(x) z(t,x)\,dx$$

in probability for all continuous ψ with compact support if $t \le T$, where (z_n, z) is any of the couples (u_n,u), (v_n,v), (h_n,h), and $u_n(t,x) := p_k(tn)$, $v_n(t,x) := r_k(tn)$, $h_n(t,x) := H_k(tn)$ if $|k - xn| < 1/2$.

In view of the argument we have sketched before Theorem 2.3, we have to show that if the parameters of $\lambda_{t,n}$ are defined by means of the smooth solution, then $(1/n)\mathsf{S}[\mu_{\tau n,n}|\lambda_{\tau n,n}] \to 0$ as $n \to +\infty$ for all $\tau \le T$, consequently the empirical processes converge in a weak sense to that solution of (2.5).

Calculation of entropy: Let $f_{t,n} := d\mu_{t,n}/d\lambda_{t,n}$ and consider the time evolution of $\mathsf{S}[\mu_{t,n}|\lambda_{t,n}] = \int \log f_{t,n}\,d\mu_{t,n}$. In the next coming calculations we are assuming that the evolved density $f_{t,n}(\omega) > 0$ is a continuously differentiable function. This hypothesis can be relaxed by means of a standard regularization procedure, see e.g. Fritz–Funaki–Lebowitz (1994). The required regularity of the parameters is a consequence of their construction via discretizing the macroscopic system (2.5). By a formal computation

$$\partial_t \mathsf{S}[\mu_{t,n}|\lambda_{t,n}] = \int (\partial_t + \mathcal{L}_0 + \sigma \mathcal{G}_{ep}) \log f_{t,n}\,d\mu_{t,n} \le \int (\partial_t + \mathcal{L}_0) f_{t,n}\,d\lambda_{t,n}$$

because $\int f_{t,n}\,d\lambda_{t,n} \equiv 1$, $\mathcal{L}_0 \log f_{t,n} = (1/f_{t,n})\mathcal{L}_0 f_{t,n}$, and the contribution of $\mathcal{G}_{ep}$ is certainly not positive. Moreover, as $\mathcal{L}_0$ is antisymmetric with respect to the Lebesgue measure, we have

$$\int (\partial_t f_{t,n} + f_{t,n}\partial_t \log g_{t,n})\,d\lambda_{t,n} = \int (\mathcal{L}_0 f_{t,n} + f_{t,n}\mathcal{L}_0 \log g_{t,n})\,d\lambda_{t,n} = 0,$$

where $g_{t,n}$ denotes the Lebesgue density of $\lambda_{t,n}$, consequently

$$\mathsf{S}[\mu_{t,n}|\lambda_{t,n}] \le \mathsf{S}[\mu_{0,n}|\lambda_{0,n}] - \int_0^t \int (\partial_s + \mathcal{L}_0) \log g_{s,n}\, d\mu_{s,n}\, ds. \tag{2.10}$$

On the other hand, as

$$\log g_{s,n} = \sum_{k=0}^{n-1} (\gamma_k r_k - \beta_k I_k - F(\beta_k, \gamma_k)),$$

where $I_k := I(p_k, r_k|\pi_k) = (p_k - \pi_k)^2/2 + V(r_k)$, by a direct calculation we obtain that

$$\partial_s \log g_{s,n} = \sum_{k=0}^{n-1} \Big(\dot\gamma_k (r_k - v_k) + \beta_k \dot\pi_k (p_k - \pi_k) - \dot\beta_k (I_k - \chi_k)\Big),$$

where "dot" indicates differentiation with respect to time.

There is a fundamental relation between the parameters β, π, γ of $\lambda_{n,t}$, namely

$$\sum_{k=0}^{n-1} ((\gamma_{k-1} - \gamma_k)\pi_k + (\beta_k \pi_k - \beta_{k+1}\pi_{k+1}) J_k + (\beta_{k+1} - \beta_k)\pi_{k+1} J_k) = 0.$$

As it is explained by Tóth–Valkó (2003), this identity is due to the conservation of the *thermodynamic entropy* in a smooth regime, which is a basic feature of all models with a proper physical motivation. On the other hand, it is a necessary requirement when we evaluate the rate of production of S in order to conclude (2.9). Indeed, we get

$$\begin{aligned}\mathcal{L}_0 \log g_{s,n} = &\sum_{k=0}^{n-1} (\gamma_{k-1} - \gamma_k)(p_k - \pi_k) \\ &+ \sum_{k=0}^{n-1} (\beta_k \pi_k - \beta_{k+1}\pi_{k+1})(V'(r_k) - J_k) \\ &+ \sum_{k=0}^{n-1} (\beta_{k+1} - \beta_k)(p_{k+1} V'(r_k) - \pi_{k+1} J_k),\end{aligned} \tag{2.11}$$

where $v_k := \int r_k\, d\lambda_{t,n}$, $\pi_k := u_k = \int p_k\, d\lambda_{t,n}$ and $\chi_k := \int I_k\, d\lambda_{t,n}$, finally $J_k = J(\chi_k, v_k) = \gamma_k/\beta_k := \int V'(r_k)\, d\lambda_{t,n}$. Notice that the local equilibrium mean of any of the last factors on the right hand sides of (2.11) above does vanish: for instance $\int (V'(r_k) - J_k)\, d\lambda_{t,n} = 0$.

The crucial step: The microscopic time t is as big as $t = n\tau$, thus there is a danger of explosion on the right hand side of (2.10) as $n \to +\infty$. However, due to the smoothness of the macroscopic solution, the nonlinear functions appearing in the sums above can be substituted by their block averages, and the celebrated

One-block Lemma, which is the main consequence of strong ergodicity, allows us to approximate the block averages by their *canonical equilibrium expectations*, see Lemma 3.1 in Guo–Papanicolau–Varadhan (1988) or Theorem 3.5 of Fritz (2001).

The wave equation: The case of (2.8) is quite simple because $\beta_k \equiv 1$ then, thus $V_k' = V'(r_k)$ is the only nonlinear function we are facing with. Block averages $\bar{\eta}_{l,k} := (1/l)(\eta_k + \eta_{k-1} + \cdots + \eta_{k-l+1})$ of size $l \in \mathbb{N}$ are also periodic functions of $k \in \mathbb{Z}$ with period n. Since $\int V_k'\,d\lambda_{1,\pi,\gamma} = S'(v_k) = J_k$ if v_k is the local equilibrium mean of r_k, $\bar{V}_{l,k}' \approx S'(\bar{r}_{l,k})$ is the desired substitution, which is valid as $l \to +\infty$ after $n \to \infty$. Presupposing $|\pi_{k+1} - \pi_k| = O(1/n)$ and $|v_{k+1} - v_k| = O(1/n)$ we write

$$\sum_{k=0}^{n-1}(\pi_k - \pi_{k+1})(V'(r_k) - S'(v_k)) \approx \sum_{k=0}^{n-1}(\pi_k - \pi_{k+1})(\bar{V}_{l,k}' - S'(v_k))$$
$$\approx \sum_{k=0}^{n-1}(\pi_k - \pi_{k+1})(S'(\bar{r}_{l,k}) - S'(v_k)) \approx \sum_{k=0}^{n-1}(\pi_k - \pi_{k+1})S''(v_k)(r_k - v_k).$$

The remainders including the squared differences coming from the expansion of $S'(\bar{r}_{l,k}) - S'(v_k)$ are estimated by means of the basic entropy inequality and the related large deviation bound; let us omit these technicalities. Comparing the leading terms we see that

$$\dot{\gamma} = S''(v_k)(\pi_{k+1} - \pi_k) \quad \text{and} \quad \dot{\pi}_k = \gamma_k - \gamma_{k-1}$$

is the right choice of the parameters because then there is a radical cancelation on the right hand side of (2.10). Since $\gamma_k = S'(v_k)$, this system is just a lattice approximation to (2.8), thus our conditions on the regularity of the parameters are also justified. Summarizing the calculations above, we get a bound

$$\mathsf{S}[\mu_{\tau,n}|\lambda_{t,n}] \le \mathsf{S}[\mu_{0,n}|\lambda_{0,n}] + \frac{K}{n}\int_0^t \mathsf{S}[\mu_{s,n}|\lambda_{s,n}]\,ds + R_n(T,l) \tag{2.12}$$

such that $R_n(T,l) \to 0$ as $n \to +\infty$ and then $l \to +\infty$, whence $\mathsf{S}[\mu_{\tau n,n}|\lambda_{\tau n,n}] = o(n)$ follows by the Grönwall inequality if $\tau \le T$.

The general case: It is a bit more complicated then the case of the p-system, the required substitutions read as

$$V'(r_k) \approx J(\bar{I}_{l,k}, \bar{r}_{l,k}) \approx J_k + J_\chi'(\chi_k, v_k)(\bar{I}_{l,k} - \chi_k) + J_v'(\chi_k, v_k)(\bar{r}_{l,k} - v_k),$$

and

$$\begin{aligned} p_{k+1}V'(r_k) \approx \bar{p}_{l,k+1}J(\bar{I}_{l,k}, \bar{r}_{l,k}) &\approx \pi_{k+1}J(\chi_k, v_k) \\ &+ J(\chi_k, v_k)(\bar{p}_{l,k+1} - \pi_{k+1}) \\ &+ \pi_{k+1}J_\chi'(\chi_k, v_k)(\bar{I}_{l,k} - \chi_k) + \pi_{k+1}J_v'(\chi_k, v_k)(\bar{r}_{l,k} - v_k). \end{aligned}$$

These steps are justified by the strong ergodicity of the dynamics (One-block Lemma), provided that $V'(r_k)$ and $\pi_{k+1}V'(r_k)$ can be replaced by their block averages. This second condition turns out to be a consequence of the smoothness of the macroscopic solution, see the construction below. The second order quadratic terms of the expansions above are estimated by means of the entropy inequality, we only need standard large deviation bounds.

To minimize $\mathsf{S}[\mu_{t,n}|\lambda_{t,n}]$, the parameters of $\lambda_{t,n}$ should be defined by means of a discretized version of the Euler equations. In fact we set

$$\pi_k = u_k, \quad \gamma_k = S'_v(\chi_k, v_k), \quad \beta_k = -S'_\chi(\chi_k, v_k),$$

where

$$\dot{v}_k = u_{k+1} - u_k, \quad \dot{u}_k = J(\chi_{k+1}, v_{k+1}) - J(\chi_k, v_k)$$

and $\dot{\chi}_k = J(\chi_k, v_k)(u_{k+1} - u_k)$, whence

$$\begin{aligned}
\beta_k \dot{\pi}_k &= (\gamma_k - \gamma_{k-1}) + (\beta_{k-1} - \beta_k)\, J_{k-1},\\
\dot{\gamma}_k &= (\beta_{k+1}\pi_{k+1} - \beta_k \pi_k) J'_v(\chi_k, v_k) + (\beta_k - \beta_{k+1}) J'_v(\chi_k, v_k),\\
\dot{\beta}_k &= (\beta_k \pi_k - \beta_{k+1}\pi_{k+1}) J'_\chi(\chi_k, v_k) + (\beta_{k+1} - \beta_k) J'_\chi(\chi_k, v_k)
\end{aligned}$$

follow by a direct computation.

As a consequence of these calculations, we see the expected cancelation of the sum of all leading terms on the right hand side of (2.10), while the remainders can be estimated by means of the entropy inequality. The summary of these computations results in (2.12), thus the proof can be terminated as it was outlined in the previous two paragraphs.

3. Compensated compactness via artificial viscosity

As we have already explained, randomness in the above modifications of the anharmonic chain implies convergence to a classical solution of the macroscopic system (2.5) or (2.8) by the strong ergodic hypothesis, but in a regime of shocks much more information is needed to pass to the hydrodynamic limit. Effective coupling techniques that we have for attractive models are not available in the case of two-component systems, compensated compactness seems to be the only tool we can use. The microscopic dynamics can not admit non-classical conservation laws because it should be ergodic in the strong sense, therefore a nontrivial Lax entropy is not conserved by the microscopic dynamics. In general, the flux of a Lax entropy exhibits a *non-gradient behavior*, but the standard spectral gap estimates of Varadhan (1994) are not sufficient for its control in this case, a *logarithmic Sobolev inequality* (LSI) is needed. The effective LSI is due to the *strong artificial viscosity* of our next model, we will consider a Ginzburg–Landau type stochastic system:

$$\begin{aligned}
dp_k &= (V'(r_k) - V'(r_{k-1}))\, dt + \sigma(\varepsilon)\,(p_{k+1} + p_{k-1} - 2p_k)\, dt\\
&\quad + \sqrt{2\sigma(\varepsilon)}\,(dw_k - dw_{k-1})
\end{aligned}$$

and

$$dr_k = (p_{k+1} - p_k)\,dt + \sigma(\varepsilon)\,(V'(r_{k+1}) + V'(r_{k-1}) - 2V'(r_k))\,dt + \sqrt{2\sigma(\varepsilon)}\,(d\tilde{w}_{k+1} - d\tilde{w}_k),$$

where $\{w_k : k \in \mathbb{Z}\}$ and $\{\tilde{w}_k : k \in \mathbb{Z}\}$ are independent families of independent Wiener processes. Of course, the macroscopic viscosity $\varepsilon\sigma(\varepsilon)$ vanishes as $\varepsilon \to 0$, but we also need $\varepsilon\sigma^2(\varepsilon) \to +\infty$ to suppress extreme fluctuations of Lax entropies. To have a standard existence and uniqueness theory for this infinite system of stochastic differential equations, we are assuming that V'' is bounded. The generator of the *Feller process* defined in this way reads as $\mathcal{L} = \mathcal{L}_0 + \sigma\mathcal{G}_p + \sigma\mathcal{G}_r$, where $\mathcal{G}_r$ is also elliptic. Additional conditions on the interaction potential V are listed below.

3.1. Conditions and main result

Just as in the case of (2.7), the same $\{\lambda_{\pi,\gamma} : \pi, \gamma \in \mathbb{R}\}$ is the family of stationary product measures, and the converse statement, i.e. the strong ergodic hypothesis can be proven in the same way. Therefore again (2.8) is expected to govern the macroscopic behavior of the system under hyperbolic scaling. The first crucial problem is the evaluation of $\mathcal{L}_0 h$ when h is a Lax entropy, we have to show that its dominant part is a difference of currents. These probabilistic calculations are based on a logarithmic Sobolev inequality. In view of the *Bakry–Emery criterion,* see Deuschel–Stroock (1989), we have to assume that V *is strictly convex*, i.e. V'' is bounded away from zero. On the other hand, the existence of weak solutions to (2.8) requires the condition of *genuine nonlinearity:* the third derivative S''' can not have more than one root, see DiPerna (1985), Shearer (1994) and Serre–Shearer (1994). In terms of V this is a consequence of one of the following assumptions.

(i) V' is strictly convex or concave on $\mathbb{R}$.

(ii) V is symmetric and $V'(r)$ is strictly convex or concave for $r > 0$.

The very same properties of the flux S' follow immediately by the theory of *total positivity.* Of course, small perturbations of such potentials also imply the required genuine nonlinearity of the macroscopic flux, $V(r) := r^2/2 - a\log\cosh(br)$ is an explicitly solvable example if $a > 0$ is small enough.

A technical condition: *asymptotic normality* requires the existence of positive constants α, V''_+, V''_- and R such that $|V''(r) - V''_+| \le e^{-\alpha r}$ if $r \ge R$, while $|V''(r) - V''_-| \le e^{\alpha r}$ if $r \le -R$.

Since we are not able to prove the uniqueness of the hydrodynamic limit, our only hypothesis on the initial distribution is an entropy bound: $\mathsf{S}[\mu_{0,\varepsilon,n}|\lambda_{0,0}] = O(n)$.

Let P_ε denote the distribution of the empirical process $(u_\varepsilon, v_\varepsilon)$, then the simplest version of our main result reads as:

Theorem 3.1. P_ε *is a tight family with respect to the weak local topology of the* L^2 *space of trajectories, and its limit distributions are all concentrated on a set of weak solutions to* (2.8).

The notion of weak convergence changes from step to step of the argument. We start with the *Young measure* of the block-averaged process, and at the end we get tightness in the strong local $L^p(\mathbb{R}_+^2)$ topology for $p < 2$; $\mathbb{R}_+^2 := \mathbb{R}_+ \times \mathbb{R}$. This strong form of our result is proven for a mollified version $(\hat{u}_\varepsilon, \hat{v}_\varepsilon)$ of the empirical process, it is defined a bit later, after (3.2). Compensated compactness is the most relevant keyword of the proofs.

3.2. On the ideas of the proof

We follow the argumentation of the vanishing viscosity approach. In a concise form (2.8) can be written as $\partial_t z + \partial_x \Phi(z) = 0$, where $z := (u, v)$, $\Phi(z) := -(S'(v), u)$, and its viscid approximation reads as $\partial_t z_\delta + \partial_x \Phi(z_\delta) = \delta\, \partial_x^2 z_\delta$. This parabolic system admits classical solutions if $\delta > 0$, and the original hyperbolic equation can be solved by sending $\delta \to 0$. The argument is not trivial at all, see e.g. Dafermos (2005). Our task is to extend this technology to microscopic systems.

Energy inequality: Observe first that the space integral of $W(z) := u^2/2 + S(v)$ is constant along classical solutions to the wave equation (2.8), moreover its viscid approximation satisfiess

$$\partial_t W(z_\delta) = \partial_x \left(u_\delta S'(v_\delta)\right) + \delta\, \partial_x \left(u_\delta \partial_x u_\delta + S'(v_\delta)\partial_x v_\delta\right) \\ - \delta \left((\partial_x u_\delta)^2 + S''(v_\delta)(\partial_x v_\delta)^2\right).$$

Since S is strictly convex, we have got a standard energy inequality: an L^2 bound for $\delta^{1/2}\, \partial_x z_\delta$. In a regime of shocks, however, this bound does not vanish as $\delta \to 0$, consequently a strong compactness argument is not available.

Young family: Nevertheless, a very weak form of compactness holds true at the level of the Young measure. The approximate solution z_δ can be represented by a measure Θ_δ on $\mathbb{R}_+^2 \times \mathbb{R}^2$ such that $d\Theta_\delta := dt\, dx\, \theta_{t,x}^\delta(dz)$, where $\theta_{t,x}^\delta$ is the *Dirac mass* sitting at the actual value $z_\delta(t, x)$ of z_δ. Since z_δ is locally bounded in $L^2(\mathbb{R}_+^2)$, we can select weakly convergent sequences from Θ_δ as $\delta \to 0$. Of course, the *Young family* $\{\theta_{t,x} : (t, x) \in \mathbb{R}_+^2\}$ of a limiting measure Θ of Θ_δ needs not be Dirac, thus we only have convergence to *measure valued solutions:* $\partial_t \theta_{t,x}(z) + \partial_x(\theta_{t,x}(\Phi(z))) = 0$ in the sense of distributions, where the abbreviation $\theta_{t,x}(\varphi(z)) := \int \varphi(z)\, \theta_{t,x}(dz)$ is used; we write $\theta_{t,x}(z)$ if $\varphi(z) \equiv z$. F The identification of measure valued solutions as weak solutions is the subject of the theory of compensated compactness, in fact the Dirac property of the limiting Young measure should be verified.

Compensated factorization: It is crucial that (2.8) admits a rich family of *Lax entropy pairs* (h, J), these are characterized by the conservation law: $\partial_t h(z) + \partial_x J(z) = 0$ along classical solutions. Let us now turn to the viscid approximation. We see that *entropy production*

$$X_\delta := \partial_t h(z_\delta) + \partial_x J(z_\delta) = \delta\, \partial_x(h'_u \partial_x u_\delta + h'_v \partial_x v_\delta) \\ - \delta \left(h''_{uu}(\partial_x u_\delta)^2 + 2h''_{uv}\partial_x u_\delta\, \partial_x v_\delta + h''_{vv}(\partial_x v_\delta)^2\right)$$

decomposes as $X_\delta = Y_\delta + Z_\delta$, where Y_δ vanishes in H^{-1}, while Z_δ is bounded in the space of measures. As a first consequence we get the *Lax entropy inequality*:

$X_\delta \le 0$ as a distribution if h is convex, but the famous *Div-Curl Lemma* is more relevant at this point. Let $\theta_{t,x}$ denote the Young family of a weak limit point Θ of the sequence of Young measures Θ_δ as $\delta \to 0$, then for couples (h_1, J_1) and (h_2, J_2) of Lax entropy pairs we have a compound factorization property:

$$\theta_{t,x}(h_1 J_2) - \theta_{t,x}(h_2 J_1) = \theta_{t,x}(h_1)\theta_{t,x}(J_2) - \theta_{t,x}(h_2)\theta_{t,x}(J_1) \tag{3.1}$$

almost everywhere on $\mathbb{R}^2$. In his pioneering papers Ronald DiPerna managed to show that (3.1) implies the Dirac property of the Young family, at least if the sequence of approximate solutions is uniformly bounded, see DiPerna (1985) with further references.

The microscopic evolution: The Ito lemma yields a parabolic energy inequality

$$\begin{aligned}\partial_t \mathsf{E} H_k\left(\omega(t)\right) = {} & \mathsf{E}\left(p_{k+1}V'(r_k) - p_k V'(r_{k-1})\right) \\ & + \sigma(\varepsilon)\mathsf{E}\left(p_k\left(p_{k+1} + p_{k-1} - 2p_k\right)\right) \\ & + \sigma(\varepsilon)\mathsf{E}\left(V'(r_k)(V'(r_{k+1}) + V'(r_{k-1}) - 2V'(r_k))\right)\end{aligned}$$

at the microscopic level. If $\varepsilon\sigma(\varepsilon)$ remains positive as $\varepsilon \to 0$, then the tightness in the local topology of $L^2(\mathbb{R})$ of the distributions of the time averaged process might follow from this bound in much the same way as it is done in PDE theory.[5] However, $\varepsilon\sigma(\varepsilon) \to 0$ as $\varepsilon \to 0$, thus the bound degenerates in the limit, consequently there is no hope to get tightness in L^2. That is why we say that a direct compactness argument does not work, the method of compensated compactness is needed.

In our case a difficult step of the usual non-gradient analysis can be avoided by considering the Lax entropy pairs (h, J) as functions of the block averaged empirical process $(\hat{u}_\varepsilon, \hat{v}_\varepsilon)$. Entropy production $X_\varepsilon := \partial_t h(\hat{u}_\varepsilon, \hat{v}_\varepsilon) + \partial_x J(\hat{u}_\varepsilon, \hat{v}_\varepsilon)$ is defined as a generalized function, without the condition $\varepsilon\sigma^2(\varepsilon) \to +\infty$ its fluctuations might explode in the limit even if we define the empirical processes in terms of block averages. The main difficulty is to identify the macroscopic flux in the microscopic expression of $\mathcal{L}_0 h$, and to show that the remainders do vanish in the limit. This is achieved by replacing block averages of the microscopic currents of momenta with their equilibrium expectations, a logarithmic Sobolev inequality plays a decisive role in the computations. This substitution transforms the evolution equation of h into a fairly transparent form: we can recover essentially the same structure which appears when the vanishing viscosity limit for (2.8) is performed. At this point can we launch the stochastic theory of compensated compactness, and the proof is terminated by referring to known results from PDE theory. Unfortunately we can not find bounded, *positively invariant regions* in stochastic situations as DiPerna (1985) did at the PDE level, but the results of Shearer (1994) and Serre–Shearer (1994) on an L^p theory of compensated compactness are applicable.

[5] In case of the diffusive models of Fritz (1986) and its continuations, an energy inequality implies this kind of tightness of the process in the space of trajectories. Guo–Papanicolau–Varadhan (1988) had raised the problem to the level of measures μ_t, and instead of energy and the H^{+1} norm of configurations, the relative entropy and its rate of production (Dirichlet form) are estimated to get the required a priori bounds including an energy inequality.

3.3. Stochastic theory of compensated compactness

Most computations involve *mesoscopic block averages* of size $l = l(\varepsilon)$ such that

$$\lim_{\varepsilon\to 0} \frac{l(\varepsilon)}{\sigma(\varepsilon)} = 0 \quad \text{and} \quad \lim_{\varepsilon\to 0} \frac{\varepsilon l^3(\varepsilon)}{\sigma(\varepsilon)} = +\infty.$$

For sequences ξ_k indexed by $\mathbb{Z}$ we define two kinds of block averages:

$$\bar{\xi}_{l,k} := \frac{1}{l}\sum_{j=0}^{l-1} \xi_{k-j} \quad \text{and} \quad \hat{\xi}_{l,k} := \frac{1}{l^2}\sum_{j=-l}^{l} (l-|j|)\,\xi_{k+j}. \tag{3.2}$$

For example, $\bar{V}'_{l,k}$ denotes the arithmetic mean of the sequence $\xi_j = V'(r_j)$. We start calculations with the "smooth" averages $\hat{\xi}_{l,k}$, the arithmetic means appear in canonical expectations. The corresponding empirical process $(\hat{u}_\varepsilon, \hat{v}_\varepsilon)$ and $(\bar{u}_\varepsilon\, \bar{v}_\varepsilon)$ are defined according to $\hat{u}_\varepsilon(t,x) := \hat{p}_{l,k}(t/\varepsilon)$ if $|\varepsilon k - x| < \varepsilon/2$, and so on. Since $\hat{u}_\varepsilon$ and $\hat{v}_\varepsilon$ are bounded in a mean sense in $L^2(dt,dx)$, the distributions $\hat{\mathsf{P}}_\varepsilon$ of the Young measure Θ form a tight family; these are now defined as $d\Theta_\varepsilon := dt\,dx\,\theta^\varepsilon_{t,x}(du)$, where $\theta^\varepsilon_{t,x}$ is the Dirac mass at the actual value of $(\hat{u}_\varepsilon, \hat{v}_\varepsilon)$. The Young family controls the asymptotic behavior of various functions of the empirical processes.

Given a Lax entropy pair (h,J), the associated entropy production is defined as

$$X_\varepsilon(\varphi,h) := -\int_0^\infty\int_{-\infty}^\infty \left(h(\hat{u}_\varepsilon,\hat{v}_\varepsilon)\varphi'_t(t,x) + J(\hat{u}_\varepsilon,\hat{v}_\varepsilon)\varphi'_x(t,x)\right)\,dx\,dt,$$

where the test function φ is compactly supported in the interior of $\mathbb{R}^2_+$. We call an entropy pair (h,J) *well controlled* if its entropy production decomposes as $X_\varepsilon(\varphi,h) = Y_\varepsilon(\varphi,h) + Z_\varepsilon(\varphi,h)$, and we have two random functionals $A_\varepsilon(\phi,h)$ and $B_\varepsilon(\phi,h)$ such that

$$|Y_\varepsilon(\psi\phi,h)| \le A_\varepsilon(\phi,h)\|\psi\|_+ \quad \text{and} \quad |Z_\varepsilon(\psi,h)| \le B_\varepsilon(\phi,h)\|\psi\|,$$

where $\|\cdot\|$ is the uniform norm, while $\|\cdot\|_+$ denotes the norm of the Sobolev space H^{+1}. Here the test function ϕ is compactly supported in the interior of $\mathbb{R}^2_+$, its role is to localize the problem. The factors A_ε and B_ε do not depend on ψ, moreover $\lim \mathsf{E}A_\varepsilon(\phi,h) = 0$ and $\limsup \mathsf{E}B_\varepsilon(\phi,h) < +\infty$ as $\varepsilon \to 0$.

Proposition 3.2. *If* (h_1,J_1) *and* (h_2,J_2) *are well controlled entropy pairs, then* (3.1) *holds true with probability one with respect to any limit distribution of* $\hat{\mathsf{P}}_\varepsilon$ *that we obtain as* $\varepsilon \to 0$.

This is the stochastic version of the Div-Curl Lemma above. The proof is not difficult, by means of the *Skorohod Embedding Theorem* it can be reduced to the original, deterministic version, see Fritz (2001), Fritz (2004) and Fritz–Tóth (2004). The main problem is the verification of its conditions, the logarithmic Sobolev inequality plays an essential role here.

3.4. The a priori bounds

Following Fritz (1990), our a priori bounds are all based on the next inequality that controls relative entropy and its rate of production. The initial condition implies that

$$\mathsf{S}[\mu_{t,\varepsilon,n}|\lambda_{0,0}] + \sigma(\varepsilon)\int_0^t D[\mu_{s,\varepsilon,n}|\lambda_{0,0}]\,ds \le C\left(t + \sqrt{n^2+\sigma(\varepsilon)t}\right)$$

for all t, n, ε with the same constant C, where D is the Dirichlet form, it is due to the elliptic perturbation of the anharmonic chain:

$$D[\mu_{t,\varepsilon,n}|\lambda_{0,0}] := \sum_{k=-n}^{n-1}\int (\nabla_1\partial_k\sqrt{f_n})^2\,d\lambda + \sum_{k=-n}^{n-1}\int (\nabla_1\tilde{\partial}_k\sqrt{f_n})^2\,d\lambda,$$

where $\nabla_l\xi_k := (1/l)(\xi_{k+l}-\xi_k)$, $f_n := d\mu_{t,\varepsilon,n}/d\lambda_{0,0}$, $\partial_k := \partial/\partial p_k$ and $\tilde{\partial}_k := \partial/\partial r_k$. This is the consequence of a system of differential inequalities:

$$\partial_t S_n + 2\sigma(\varepsilon)D_n \le K\left(S_{n+1} - S_n + \sigma(\varepsilon)\sqrt{S_{n+1}-S_n}\,\sqrt{D_{n+1}-D_n}\right),$$

where $S_n := \mathsf{S}[\mu_{t,\varepsilon,n}|\lambda_{0,0}]$ and $D_n := D[\mu_{t,\varepsilon,n}|\lambda_{0,0}]$ for brevity. For a proof of this local entropy bound see Fritz (2011) with further references.

LSI: The logarithmic Sobolev inequality we are going to use, can be stated as follows. Given $\bar{r}_{l,k} = v$, let $\mu_{l,k}^v$ and $\lambda_{l,k}^v$ denote the conditional distributions of the variables $r_k, r_{k+1}, ..., r_{k+l-1}$ with respect to μ and $\lambda_{0,0}$, and set $f_{l,k}^v := d\mu_{l,k}^v/d\lambda_{l,k}^v$, then

$$\int \log f_{l,k}^v\,d\mu_{l,k}^v \le l^2 C_{\mathrm{lsi}}\sum_{j=k}^{k+l-2}\int\left(\nabla_1\tilde{\partial}_k(f_{l,k}^v)^{1/2}\right)^2 d\lambda_{l,k}^v$$

for all μ, v, k, l with a universal constant C_{lsi} depending only on V. Of course, a similar inequality holds true for the conditional distributions of momenta. Combining this with the standard entropy inequality $\int\varphi\,d\mu \le \mathsf{S}[\mu|\lambda] + \log\int e^{\varphi}\,d\lambda$, the calculation of expectations reduces to large deviation bounds for the canonical distributions of the equilibrium measure $\lambda_{0,0}$. The most important consequence of the local entropy bound and this LSI is the evaluation of the microscopic current of momentum as follows:

$$\sum_{|k|<n}\int_0^t\int\left(\bar{V}_{l,k}' - S'(\bar{r}_{l,k})\right)^2 d\mu_{s,\varepsilon}\,ds \le C_1\left(\frac{nt}{l} + \frac{l^2\sqrt{n^2+\sigma(\varepsilon)t}}{\sigma(\varepsilon)}\right).$$

Similar bounds control the differences $\bar{r}_{l,k+l} - \bar{r}_{l,k}$ and $\hat{r}_{l,k} - \bar{r}_{l,k}$. Later on the validity of such a bound will be indicated as $\bar{V}_{l,k}' \approx S'(\bar{r}_{l,k})$, $\bar{r}_{l,k+l} \approx \bar{r}_{l,k}$, and so on.

Entropy flux: Finally, let us outline the crucial step of the evaluation of entropy production at a heuristic level. Consider a Lax entropy $h = h(u, v)$ with flux

$J = J(u,v)$ and expand J. The second order terms of the Lagrange expansion can be neglected, thus we have

$$\begin{aligned} X_{0,k} &:= \mathcal{L}_0 h(\hat{p}_{l,k},\hat{r}_{l,k}) + J(\hat{p}_{l,k+1},\hat{r}_{l,k+1}) - J(\hat{p}_{l,k},\hat{r}_{l,k}) \\ &\approx h'_u(\hat{p}_{l,k},\hat{r}_{l,k})(\hat{V}'_{l,k} - \hat{V}'_{l,k-1}) + h'_v(\hat{p}_{l,k},\hat{r}_{l,k})(\hat{p}_{l,k+1} - \hat{p}_{l,k}) \\ &\quad + J'_u(\hat{p}_{l,k},\hat{r}_{l,k})(\hat{p}_{l,k+1} - \hat{p}_{l,k}) + J'_v(\hat{p}_{l,k},\hat{r}_{l,k})(\hat{r}_{l,k+1} - \hat{r}_{l,k}). \end{aligned}$$

Since $h'_u(u,v)S''(v) + J'_v(u,v) = h'_v(u,v) + J'_u(u,v) = 0$,

$$X_{0,k} \approx h'_u(\hat{p}_{l,k},\hat{r}_{l,k})(\hat{V}'_{l,k} - \hat{V}'_{l,k-1}) - h'_u(\hat{p}_{l,k},\hat{r}_{l,k})S''(\hat{r}_{l,k})(\hat{r}_{l,k+1} - \hat{r}_{l,k}).$$

Observe now that $\hat{\xi}_{l,k+1} - \hat{\xi}_{l,k} = (1/l)(\bar{\xi}_{l,k+l} - \bar{\xi}_{l,k})$, thus the substitution $\bar{V}'_{l,k} \approx S'(\bar{r}_{l,k})$ results in $l\, X_{0,k} \approx 0$ as

$$l\, X_{0,k} \approx h'_u(\hat{p}_{l,k},\hat{r}_{l,k})\left(S'(\bar{r}_{l,k-1+l}) - S'(\bar{r}_{l,k-1}) - S''(\hat{r}_{l,k})(\bar{r}_{l,k+l} - \bar{r}_{l,k})\right).$$

Of course, the precise computation is much more complicated because in the formula X_ε of entropy production the terms $X_{0,k}$ have a factor $1/\varepsilon$. In fact, $(\varepsilon l(\varepsilon)\sigma(\varepsilon))^{-1}$ is the order of the replacement error; that is why we need $\varepsilon\sigma^2(\varepsilon) \to +\infty$ and the sharp explicit bounds provided by the logarithmic Sobolev inequality.

4. Relaxation of interacting exclusions

We consider ± 1 charges in an electric field, positive charges jump to the right on $\mathbb{Z}$, negative charges move to the left with unit jump rates in both cases such that two or more particles can not coexist at the same site. There is an interaction between these processes: if charges of opposite sign meet, then they jump over each other at rate 2. The configurations are doubly infinite sequences $\omega_k \in \{-1,0,1\}$ indexed by $\mathbb{Z}$, $\omega_k = 0$ indicates an empty site, and $\eta_k := \omega_k^2$ denotes the occupation number. The generator of the process is acting on local functions φ as

$$\mathcal{L}_0\varphi(\omega) = \frac{1}{2}\sum_{k\in\mathbb{Z}}(\eta_k + \eta_{k+1} + \omega_k - \omega_{k+1})(\varphi(\omega^{k,k+1}) - \varphi(\omega));$$

$\omega \to \omega^{k,k+1}$ indicates the exchange of ω_k and ω_{k+1}. This most interesting model had been introduced by Tóth–Valkó (2003), where its HDL in a smooth regime is demonstrated, too. The total charge $P = \sum \omega_k$ and particle number $R = \sum \eta_k$ are obviously preserved by the evolution, and the associated family of translation invariant stationary product measures $\{\lambda_{u,\rho}\}$ can be parametrized so that $\int \omega_k\, d\lambda_{u,\rho} = u$ and $\int \eta_k\, d\lambda_{u,\rho} = \rho$. Conservation of ω and η means that they are driven by currents, i.e. $\mathcal{L}_0\omega_k = \mathrm{j}^\omega_{k-1} - \mathrm{j}^\omega_k$ and $\mathcal{L}_0\eta = \mathrm{j}^\eta_{k-1} - \mathrm{j}^\eta_k$, where

$$\begin{aligned} \mathrm{j}^\omega_k &:= (1/2)\left(\eta_k + \eta_{k+1} - 2\omega_k\omega_{k+1} + \omega_k\eta_{k+1} - \eta_k\omega_{k+1} + \eta_k - \eta_{k+1}\right), \\ \mathrm{j}^\eta_k &:= (1/2)\left(\omega_k + \omega_{k+1} - \omega_k\eta_{k+1} - \eta_k\omega_{k+1} + \eta_k - \eta_{k+1}\right). \end{aligned}$$

Since $\int \mathrm{j}_k^\omega \, d\lambda_{u,\rho} = \rho - u^2$ and $\int \mathrm{j}_k^\eta \, d\lambda_{u,\rho} = u - u\rho$, the principle of local equilibrium suggests that under hyperbolic scaling a version of the Leroux system:

$$\partial_t u + \partial_x(\rho - u^2) = 0 \quad \text{and} \quad \partial_t \rho + \partial_x(u - u\rho) = 0 \tag{4.1}$$

governs the macroscopic evolution. The strong ergodic hypothesis can easily be proven by a standard entropy argument. In a regime of shock waves the method of compensated compactness is applied to derive the Leroux system; therefore an additional stirring mechanism:

$$\mathcal{G}_e \varphi(\omega) := \sum_{k \in \mathbb{Z}} \big(\varphi(\omega^{k,k+1}) - \varphi(\omega)\big)$$

is needed to regularize the process. The full generator reads as $\mathcal{L} := \mathcal{L}_0 + \sigma(\varepsilon)\, \mathcal{G}_e$, and our usual conditions $\varepsilon\sigma(\varepsilon) \to 0$ and $\varepsilon\sigma^2(\varepsilon) \to +\infty$ are assumed.

The statement is similar to the case of isentropic elastodynamics, the proof is based on the logarithmic Sobolev inequality what we have for the stirring generator $\mathcal{G}_e$, see Fritz–Tóth (2004), where HDL is proven in a periodic setting. The extension of this result to general initial values is explained by Fritz–Nagy (2006), the optimal version concerns the mollified empirical processes $\hat{u}_\varepsilon(t,x) := \hat{\omega}_{l,k}(t/\varepsilon)$ and $\hat{\rho}_\varepsilon(t,x) := \hat{\eta}_{l,k}(t/\varepsilon)$ if $|\varepsilon k - x| < \varepsilon/2$, where the block size $l = l(\varepsilon)$ is the same as in Section 3.

Theorem 4.1. *The distributions of our empirical processes form a tight family with respect to the strong local topology of* $L^1(\mathbb{R}_+^2)$*, and any limit distribution of* $(\hat{u}_\varepsilon, \hat{\rho}_\varepsilon)$ *is concentrated on a set of weak solutions to* (4.1)*. These weak solutions satisfy the Lax entropy condition, too.*

A uniqueness theorem for the Leroux system requires only a local bound on the total variation of the weak solution we have constructed, nevertheless we are not able to prove the uniqueness of the hydrodynamic limit.

4.1. Creation and annihilation of charges

In the paper Fritz–Nagy (2006) it was shown that an additional spin-flip mechanism relaxes the Leroux system to the Burgers equation $\partial_t \rho + \kappa\, \partial_x(\rho - \rho^2) = 0$ even in the case of shocks, where $\kappa = 0$ in the *completely symmetric* case. The replacement $u \approx \kappa\rho$ is due to a second logarithmic Sobolev inequality. The following modification of the above process of interacting exclusions is a caricature of *electrophoresis,* and it is interesting also from the point of view of mathematics because the PDE method of *relaxation schemes* is reformulated for the microscopic dynamics.

The model: Imagine that when two particles of opposite charge collide, then instead of jumping over each other, they may kill each other and disappear, while at two neighboring empty sites a couple $(+1, -1)$ can be created. The action $\omega \to \omega^{k+}$ of *creation* at the bond $b = (k, k+1)$ means that $(\omega_k, \omega_{k+1}) \to (1, -1)$ if $\omega_k = \omega_{k+1} = 0$, while *annihilation* $\omega \to \omega^{k-}$ is defined by $(\omega_k, \omega_{k+1}) \to (0, 0)$

if $\omega_k = 1$ and $\omega_{k+1} = -1$; at other sites the configuration is not altered. The generator of this process of *interacting exclusions with creation and annihilation* reads as $\mathcal{L}^* = \mathcal{L}_0 + \beta(\varepsilon)\,\mathcal{G}^*$, where

$$\mathcal{G}^*\varphi(\omega) := \sum_{k\in\mathbb{Z}}(1-\eta_k)(1-\eta_{k+1})(\varphi(\omega^{k+}) - \varphi(\omega)) + \frac{1}{4}\sum_{k\in\mathbb{Z}}(\eta_k+\omega_k)(\eta_{k+1}-\omega_{k+1})(\varphi(\omega^{k-}) - \varphi(\omega)).$$

Since we do not want to postulate smoothness of the macroscopic solution, the process should be regularized by stirring, thus the effective generator becomes $\mathcal{L} := \mathcal{L}^* + \sigma(\varepsilon)\,\mathcal{G}_e$. The factor $\sigma = \sigma(\varepsilon)$ is the same as above, and it is natural to assume that β is a positive constant because it is the parameter of the basic model.

Creation-annihilation violates the conservation of particle number, only total charge $\sum \omega_k$ is preserved by our stochastic dynamics. A product measure $\lambda_{u,\rho}$ will be stationary if $\lambda_{u,\rho}[\omega_k = 0, \omega_{k+1} = 0] = \lambda_{u,\rho}[\omega_k = 1, \omega_{k+1} = -1]$, that is $4(1-\rho)^2 = (\rho^2 - u^2)$, whence

$$\rho = F(u) := (1/3)(4 - \sqrt{4-3u^2}) \tag{4.2}$$

is the criterion of stationarity because the second root:

$$\tilde{F}(u) := (1/3)(4 + \sqrt{4-3u^2}) \geq 5/3 > 1.$$

Therefore our one-parameter family $\{\lambda_u^* : |u| < 1\}$ of stationary product measures is defined by $\lambda_u^* := \lambda_{u,F(u)}$. Of course, $\int \omega_k\,d\lambda_u^* = u$ and $\int \eta_k\,d\lambda_u^* = F(u)$, thus $\int \mathrm{j}_k^\omega\,d\lambda_u^* = F(u) - u^2$. On the other hand, $\mathcal{G}^*\omega_k = \mathrm{j}_{k-1}^{\omega*} - \mathrm{j}_k^{\omega*}$ is a difference of currents,

$$\mathrm{j}_k^{\omega*}(\omega) := (1/4)(\eta_k+\omega_k)(\eta_{k+1}-\omega_{k+1}) - (1-\eta_k)(1-\eta_{k+1}), \tag{4.3}$$

and $\int \mathrm{j}_k^{\omega*}\,d\lambda_{u,\rho} = C(u,\rho) := (3/4)(\rho - F(u))(\tilde{F}(u) - \rho)$, thus the equilibrium expectation of $\mathrm{j}_k^{\omega*}$ does vanish, consequently the principle of local equilibrium predicts

$$\partial_t u(t,x) + \partial_x(F(u) - u^2) = 0 \tag{4.4}$$

as the result of the hyperbolic scaling limit. Note that the flux is neither convex nor concave, thus the structure of shock waves may be rather complex.

It is not a surprise that the contribution of the creation-annihilation mechanism does not appear in the limit. The generator $\mathcal{G}^*$ is symmetric in $L^2(d\lambda_u^*)$, consequently a diffusive scaling would be needed to recover its action.

Main result. Assume that the initial distributions satisfy

$$\lim_{\varepsilon\to 0}\,\varepsilon\sum_{k\in\mathbb{Z}}\varphi(\varepsilon k)\omega_k(0) = \int_{-\infty}^{\infty}\psi(x)u_0(x)\,dx$$

in probability for all compactly supported $\varphi \in C(\mathbb{R})$. We say that a measurable and bounded $u = u(t,x)$ is a *weak entropy solution* to (4.4) with initial value u_0 if

$$\int_0^\infty \int_{-\infty}^\infty (\psi_t'(t,x)u(t,x) + \psi_x'(t,x)(F(u(t,x)) - u^2(t,x)))\,dx\,dt + \int_{-\infty}^\infty \psi(0,x)u_0(x)\,dx = 0,$$

and for all convex entropy pairs (h,J) we have the Lax inequality:

$$-X_\varepsilon(\psi,h) = \int_0^\infty \int_{-\infty}^\infty (\psi_t'(t,x)h(u) + \psi_x'(t,x)J(u))\;dx\,dt\; \geq 0 \tag{4.5}$$

whenever $0 \leq \psi \in C^1(\mathbb{R}^2)$ is compactly supported in the interior of $\mathbb{R}^2_+$. Entropy pairs of (4.4) are characterized by $J'(u) = (F'(u) - 2u)h'(u)$, that is $\partial_t h(u) + \partial_x J(u) = 0$ along classical solutions. Our effective empirical process $\hat{u}_\varepsilon$ is now defined as $\hat{u}_\varepsilon(t,x) := \hat{\omega}_{l,k}(t/\varepsilon)$ if $|\varepsilon k - x| < \varepsilon/2$; the mesoscopic block size $l = l(\varepsilon)$ is just as big as it was in the previous section.

In the paper by Bahadoran–Fritz–Nagy (2011) we prove:

Theorem 4.2. *The above conditions imply that*

$$\lim_{\varepsilon\to 0} \mathsf{E} \int_0^\tau \int_{-r}^r |u(t,x) - \hat{u}_\varepsilon(t,x)|\,dx\,dt = 0$$

for all $r, \tau > 0$*, where* u *is the uniquely specified weak entropy solution to* (4.4) *with initial value* u_0.

Let us remark that the coefficient $\beta > 0$ needs not be a constant, it is sufficient to assume that $\sigma(\varepsilon)\beta(\varepsilon) \to +\infty$ and $\varepsilon\sigma^2(\varepsilon)\beta^{-4}(\varepsilon) \to +\infty$ as $\varepsilon \to 0$.

4.2. Relaxation in action

The proof follows the standard technology of the stochastic theory of compensated compactness, the entropy production for entropy pairs of (4.4) has to be evaluated. Here the uniqueness of the hydrodynamic limit is a consequence of the Lax entropy inequality, see Chen–Rascle (2000), thus $\limsup X_\varepsilon(\psi,h) \leq 0$ is also needed for $\psi \geq 0$ and convex h. We are facing with the computation of four basic quantities, besides j_k^ω, j_k^η and $\mathrm{j}_k^{\omega *}$,

$$\begin{aligned}\mathcal{G}^*\eta_k = {} & (1-\eta_k)(1-\eta_{k+1}) - (1/4)(\eta_k + \omega_k)(\eta_{k+1} - \omega_{k+1}) \\ & + (1-\eta_{k-1})(1-\eta_k) - (1/4)(\eta_{k-1} + \omega_{k-1})(\eta_{k-1} - \omega_{k-1})\end{aligned}$$

should also be evaluated. Since $\mathcal{G}^*\eta_k = -\mathfrak{j}^{\omega*}_{k-1} - \mathfrak{j}^{\omega*}_k$, we have

$$\int \mathcal{G}^*\eta_k \, d\lambda_{u,\rho} = (3/2)(\rho - F(u))(\rho - \tilde{F}(u)) = -2C(u,\rho). \tag{4.6}$$

The macroscopic flux: The fundamental local bound on relative entropy and its rate of production holds true also in this case, see Lemma 3.1 of our paper, thus the logarithmic Sobolev inequality involving the Dirichlet form of $\mathcal{G}_e$ applies, too. In this way we can estimate canonical expectations given $\bar{\omega}_{l,k}$ and $\bar{\eta}_{l,k}$, see Lemmas 3.3–3.5 in Bahadoran–Fritz–Nagy (2011); the explicit upper bounds are the same as in Section 3.4. Therefore the replacements

$$\bar{\mathfrak{j}}^{\omega}_{l,k} \approx \bar{\eta}_{l,k} - (\bar{\omega}_{l,k})^2, \quad \bar{\mathfrak{j}}^{\eta}_{l,k} \approx \bar{\omega}_{l,k} - \bar{\omega}_{l,k}\bar{\eta}_{l,k}, \quad \bar{\mathfrak{j}}^{\omega*}_{l,k} \approx C(\bar{\omega}_{k,l}, \bar{\eta}_{l,k}) \tag{4.7}$$

and $\bar{\eta}^*_{l,k} \approx -2C(\bar{\omega}_{k,l}, \bar{\eta}_{l,k})$, where $\eta^*_j := \mathcal{G}^*\eta_j$ for convenience, are all allowed, moreover $\bar{\omega}_{l,k+l} \approx \bar{\omega}_{l,k} \approx \hat{\omega}_{l,k}$ and $\bar{\eta}_{l,k+l} \approx \bar{\eta}_{l,k}$.

Entropy production: Since $\mathcal{G}^*$ is reversible, the critical component of entropy production is induced by $\mathcal{L}_0$. Let us consider now an entropy pair (h, J) of (4.4), i.e. $J'(u) = (F'(u) - 2u)h'(u)$. In view of the asymptotic equivalence relations listed above, we obtain that

$$\begin{aligned} X^*_{0,k} &:= \mathcal{L}_0 h(\hat{\omega}_{l,k}) + J(\hat{\omega}_{l,k+1}) - J(\hat{\omega}_{l,k}) \approx h'(\hat{\omega}_{l,k})(\hat{\mathfrak{j}}^{\omega}_{l,k-1} - \hat{\mathfrak{j}}^{\omega}_{l,k}) \\ &\quad + (F'(\hat{\omega}_{l,k}) - 2\hat{\omega}_{l,k})\, h'(\hat{\omega}_{l,k})(\hat{\omega}_{l,k+1} - \hat{\omega}_{l,k}) \\ &\approx (1/l)h'(\hat{\omega}_{l,k})\left(\bar{\eta}_{l,k} - \bar{\eta}_{l,k+l} - (\bar{\omega}_{l,k})^2 + (\bar{\omega}_{l,k+l})^2\right) \\ &\quad + (1/l)h'(\hat{\omega}_{l,k})\left(F'(\bar{\omega}_{l,k}) - \bar{\omega}_{l,k} - \bar{\omega}_{l,k+l}\right)(\bar{\omega}_{l,k+l} - \bar{\omega}_{l,k}) \\ &\approx (1/l)h'(\hat{\omega}_{l,k})\left(\bar{\eta}_{l,k} - \bar{\eta}_{l,k+l} + F'(\bar{\omega}_{l,k})(\bar{\omega}_{l,k+l} - \bar{\omega}_{l,k})\right), \end{aligned}$$

whence the required $l\,X^*_{0,k} \approx 0$ would follow by the substitution $\bar{\eta}_{l,k} \approx F(\bar{\omega}_{l,k})$. Since we do not have the appropriate logarithmic Sobolev inequality, another tool must be found.

Relaxation schemes: Observe that $\bar{\eta}_{l,k}$ appears with a negative sign in the formula of $\mathcal{G}^*\bar{\eta}_{l,k}$, see also (4.6), thus there is a hope to experience *relaxation*, which results in $C(\bar{\omega}_{l,k}, \bar{\eta}_{l,k}) \to 0$ as $\varepsilon \to 0$. Although the evolution equations of $\bar{\omega}_{l,k}$ and $\bar{\eta}_{l,k}$ are rather complicated, the following couple of approximate identities reflects quite well the underlying structure. Applying the substitution relations (4.7) and neglecting obviously vanishing terms, we get

$$\begin{aligned} d\tilde{u}_\varepsilon + \partial_x(\tilde{\rho}_\varepsilon - \tilde{u}_\varepsilon^2)\,dt + \beta\,\partial_x C(\tilde{u}_\varepsilon, \tilde{\rho}_\varepsilon)\,dt \approx 0, \\ d\tilde{\rho}_\varepsilon + \partial_x(\tilde{u}_\varepsilon - \tilde{u}_\varepsilon\tilde{\rho}_\varepsilon) + (2\beta/\varepsilon)\,C(\tilde{u}_\varepsilon, \tilde{\rho}_\varepsilon)\,dt \approx 0, \end{aligned}$$

where $\tilde{u}_\varepsilon \sim \bar{\omega}_{l,k}$ and $\tilde{\rho}_\varepsilon \sim \bar{\eta}_{l,k}$ by mollification. Since

$$(\rho - F(u))C(u,\rho) \geq \Psi(u,\rho) := (1/2)\,(\rho - F(u))^2\,,$$

even the trivial Liapunov function Ψ can be applied to conclude that $\bar{\eta}_{l,k} \approx F(\bar{\omega}_{l,k})$. This trick works well if $\varepsilon\sigma^2(\varepsilon)\beta^2(\varepsilon) \to +\infty$ as $\varepsilon \to 0$, a slightly better result can be proven by replacing Ψ with a clever Lax entropy, see Bahadoran–Fritz–Nagy (2011).

5. Concluding remarks

In spite of some progress in the stochastic theory of compensated compactness, there are many relevant open problems whose solution seems to be hard or even hopeless at this time.

The Lax inequality: The dominant term of entropy production $X_\varepsilon(\psi, h)$ is generated by the elliptic components of $\mathcal{L} = \mathcal{L}_0 + \sigma(\varepsilon)\mathcal{G}_p + \sigma(\varepsilon)\mathcal{G}_r$. It is bounded in the space of measures, and the contribution of $\mathcal{G}_p$ is obviously not positive if $\psi \geq 0$ and h is convex. Our naive large deviation technique is not strong enough to exploit that V is convex. The Lax inequality restricts the set of limiting weak solutions, but in the case of systems it is not a known condition of uniqueness.

Uniqueness of HDL: This is a very hard problem in the case of a couple of conservation laws because any criterion of uniqueness presupposes a sharp local bound at fixed times. Unfortunately, in the case of stochastic models we are able to bound expectations of space-time integrals only. From the point of view of computations the microscopic systems of statistical physics are more complicated than the sophisticated numerical schemes of PDE theory.[6] For example, even the existence of positively invariant regions is a problematic issue.

Physical viscosity: It would be nice to materialize the argumentation of Serre–Shearer (1994) at the microscopic level, that is to consider hyperbolic scaling of the model $\mathcal{L} = \mathcal{L}_0 + \sigma\mathcal{G}_p$ in a regime of shocks. This is not easy because the Dirichlet form of $\mathcal{G}_p$ controls the distribution of velocities only, while the most crucial step consists of the substitution $\bar{V}'_{l,k} \approx S'(\bar{r}_{l,k})$. The less interesting case of $\mathcal{L} = \mathcal{L}_0 + \sigma\mathcal{G}_r$ seems to be simpler, but it not trivial at all.

The strength of artificial viscosity: The condition $\varepsilon\sigma^2(\varepsilon) \to +\infty$ is not necessary in the case of attractive models, but it is systematically applied in more general situations.

Euler equations with physical viscosity: HDL of the model $\mathcal{L} = \mathcal{L}_0 + (1/\varepsilon)\mathcal{G}_r$ results in the p-system of elastodynamics with artificial viscosity, see Theorem 3 in Fritz (1990). The derivation of the viscid Euler equations (1.5) of Chen–Dafermos (1995) is more complicated because then a momentum and energy preserving diffusive noise should be added to the equations of the anharmonic chain. To solve the resulting non-gradient problem, the spectral gap of the elliptic components of the generator ought to be determined.

References

[1] C. Bahadoran (2004): Blockage hydrodynamics of one-dimensional driven conservative systems. Ann. Probab. **32**:805–854.

[6] There are some rumors in the folklore of this science that the asymptotic behavior of stochastic systems is more stable than that of purely numerical procedures, but I have not heard about convincing experiments.

[2] C. Bahadoran, J. Fritz and K. Nagy (2011): Relaxation schemes for interacting exclusions. Electronic J. Probab. **16**:230–262.

[3] C. Bernardin and S. Olla (2010): Non-equilibrium macroscopic dynamics of chains of anharmonic oscillators. IHP Lecture Notes, in preparation. www.ceremade.dauphine.fr/∼olla.

[4] C. Boldrighini, R.L. Dobrushin and Yu.M. Suhov (1980): Hydrodynamics of hard rods. Uspechi Mat. Nauk. **35**:252–253.

[5] C. Boldrighini and R. L. Dobrushin (1983): One-dimensional hard rod caricature of hydrodynamics. J. Stat. Phys. **31**:577–616.

[6] A. Bressan: Hyperbolic Systems of Conservation Laws: The One Dimensional Cauchy Problem. Oxford Lecture Series in Math. Appl. **20,** Oxford 2000.

[7] G.-Q. Chen and C.M. Dafermos (1995): The vanishing viscosity method in one-dimensional thermoelasticity. Trans. AMS **347**:531–541.

[8] G.Q. Chen and M. Rascle (2000): Initial layers and uniqueness of weak entropy solutions to hyperbolic conservation laws. Arch. Rational Mech. Anal. **153**:205–220.

[9] C.M. Dafermos: Hyperbolic Conservation Laws in Continuum Physics. Grundlehren der Mathematishen Wissenschaften **325.** Springer 2005.

[10] J.-D. Deuschel and D.W. Stroock: Large Deviations. Acad. Press, Boston 1989.

[11] R.J. DiPerna (1985): Measure-valued solutions to conservation laws. Arch. Rational Mech. Anal. **88**:223–270.

[12] R.L. Dobrushin, A. Pellegrinotti, Yu.M. Suhov and L. Triolo (1985): One-dimensional harmonic lattice caricature of hydrodynamics. J. Stat. Phys. **38**:473–507.

[13] J. Fritz (1986): On the hydrodynamic limit of a scalar Ginzburg-Landau lattice model. The resolvent approach. In: Hydrodynamic Behaviour and Interacting Particle Systems. IMA Volumes in Mathematics and Its Applications. **9**:75–99. G. Papanicolau Ed. Springer 1987.

[14] J. Fritz (1990): On the diffusive nature of entropy flow in infinite systems: Remarks to a paper by Guo, Papanicolau and Varadhan. Commun. Math. Phys. **133**:331–352.

[15] J. Fritz, T. Funaki and J.L. Lebowitz (1994): Stationary states of random Hamiltonian systems. Probab. Theory Relat. Fields **99**:211–236.

[16] J. Fritz: An Introduction to the Theory of Hydrodynamic Limits. Lectures in Mathematical Sciences **18.** The University of Tokyo, Tokyo 2001.

[17] J. Fritz (2004): Entropy pairs and compensated compactness for weakly asymmetric systems. Advanced Studies in Pure Mathematics **39**:143–171.

[18] J. Fritz and B. Tóth (2004): Derivation of the Leroux system as the hydrodynamic limit of a two-component lattice gas. Commun. Math. Phys. **249**:1–27.

[19] J. Fritz and Katalin Nagy (2006): On uniqueness of the Euler limit of one-component lattice gas models. ALEA Latin American J. Probab. **1**:367–392.

[20] J. Fritz (2011): Microscopic theory of isothermal elastodynamics. Arch. Rational Mech. Anal. **201**:209–249.

[21] G.C. Guo, G. Papanicolau and S.R.S. Varadhan (1988): Nonlinear diffusion limit for a system with nearest neighbor interactions. Commun. Math. Phys. **8**:31–59.

[22] L. Hörmander Lectures on Nonlinear Hyperbolic Differential Equations. Springer 1997.

[23] C. Kipnis and C. Landim: Scaling Limit of Interacting Particle Systems. Grundlehren Math. Wissenschaften **320.** Springer 1999.

[24] C. Morrey (1955): On the derivation of hydrodynamics from statistical mechanics. Commun. Pure Appl. Math. **8**:279–326.

[25] S. Olla, S.R.S. Varadhan and H.-T. Yau (1993): Hydrodynamical limit for a Hamiltonian system with weak noise Commun. Math. Phys. **155**:523–560.

[26] F. Rezakhanlou (1991): Hydrodynamic limit for attractive particle systems on $\mathbf{Z}^d$. Commun. Math. Phys. **140**:417–448.

[27] H. Rost (1981): Nonequilibrium behaviour of a many particle process: density profile and local equilibrium. Z. Wahrsch. verw. Geb. **58**:41–53.

[28] D. Serre and J.W. Shearer (1994): Convergence with physical viscosity for nonlinear elasticity. Preprint.

[29] J.W. Shearer (1994): Global existence and compactness in L^p for the quasi-linear wave equation. Commun. Partial Differetial Equations. **19**:1829–1877.

[30] Ya.G. Sinai (1988): Dynamics of local equilibrium Gibbs distributions and Euler equations. Selecta Matematica Sovietica **7**:279–289.

[31] H. Spohn: Large Scale Dynamics of Interacting Particles. Texts and Monographs in Physics. Springer 1991.

[32] B. Tóth and B. Valkó (2003): Onsager relations and Eulerian hydrodynamic limit for systems with several conservation laws. J. Statist. Phys. **112**:497–521.

[33] S.R.S. Varadhan: Nonlinear diffusion limit for a system with nearest neighbor interactions II. In: Asymptotic Problems in Probability Theory: Stochastic Models and Diffusion on Fractals. Elworthy, K. D. and Ikeda, N. Eds. Pitman Research Notes in Mathematics. **283**:75–128. Wiley 1994.

[34] H.-T. Yau (1991): Relative entropy and hydrodynamics of Ginzburg–Landau models. Lett. Math. Phys. **22**:63–80.

Annales Mathematicae et Informaticae
39 *(2012) pp. 109–123*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Large deviations for some normalized sums of exponentially distributed random variables*

Rita Giuliano[a], **Claudio Macci**[b]

[a]Dipartimento di Matematica "L. Tonelli", Università di Pisa, Pisa, Italy
giuliano@dm.unipi.it

[b]Dipartimento di Matematica, Università di Roma Tor Vergata, Rome, Italy
macci@mat.uniroma2.it

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

We prove large deviation results for sequences of normalized sums which are defined in terms of triangular arrays of exponentially distributed random variables. We also present some examples: one of them might have applications in reliability theory because it concerns the spacings of i.i.d. exponentially distributed random variables; in another one we consider a sequence of logarithmically weighted means.

Keywords: large deviations, exponential distribution, Riemann-ζ function, triangular array, spacings, logarithmically weighted mean.

MSC: 60F05, 60F10, 60F15, 62G30, 11M06.

1. Introduction

Throughout the paper we use the symbol $Z \sim \mathcal{E}(\lambda)$ to mean that a random variable Z has exponential distribution with parameter λ, i.e. Z has continuous density $f_Z(t) = \lambda e^{-\lambda t} 1_{(0,\infty)}(t)$. The aim is to study the convergence and to present results on large deviations for the sequence $(R_n)_{n\geq 1}$ defined by

$$R_n := \frac{\sum_{j=1}^{n} T_j^{(n)}}{\gamma_n},$$

*The financial support of the Research Grant PRIN 2008 *Probability and Finance* is gratefully acknowledged.

where: $(T_j^{(n)})_{n\geq j\geq 1}$ is a triangular array of exponentially distributed random variables i.e., for every $n \geq 1$, $T_1^{(n)}, \ldots, T_n^{(n)}$ are independent and $T_j^{(n)} \sim \mathcal{E}(\lambda_j^{(n)})$ for some $(\lambda_j^{(n)})_{j\leq n}$; we put $\gamma_n := \sum_{j=1}^n s_{j,n}$ for $s_{j,n} := \frac{1}{\lambda_j^{(n)}}$, and we assume in the whole paper that $\lim_{n\to\infty} \gamma_n = +\infty$.

The theory of large deviations gives an asymptotic computation of small probabilities on exponential scale (we refer to [2] for this topic), and the basic concept of Large Deviation Principle (LDP from now on) consists of an upper bound for all closed sets and a lower bound for all open sets. Here we can prove the upper bound for all closed sets (Theorem 3.1) and the lower bound for a class of open sets (Theorem 3.2) which depends on a constant $c > 0$ appearing in the assumptions. It is worth noting that, if $c \geq 1$, this class of open sets coincides with all the open sets; therefore, as stated in Corollary 3.6 below, we have a full LDP if $c \geq 1$.

We remark that in our setting we obtain a linear rate function (see I in eq. (3.2) below). This situation is completely different from the classical one, in which all the random variables $(T_j^{(n)})_{n\geq j\geq 1}$ have the *same* exponential distribution, i.e. $\mathcal{E}(1)$ (see assumption (ii) in Theorem 3.1), and $\gamma_n = n$ (for all $n \geq 1$). In such a case $(R_n)_{n\geq 1}$ is a sequence of partial empirical means of i.i.d. random variables and, by the well-known Cramér Theorem (see e.g. Theorem 2.2.3 in [2]), the LDP holds with a strictly convex rate function.

We also give some illustrative examples. In Example 4.1 we have $\lambda_j^{(n)} = j$ for all $j = 1, \ldots, n$; in view of potential applications in reliability theory, we notice that (for every $n \geq 1$) the random variables $(T_j^{(n)})_{j\leq n}$ can be considered as the *spacings* of independent random variables with distribution $\mathcal{E}(1)$ (see Remark 4.2). Example 4.3 consists of a simple choice of $(\lambda_j^{(n)})_{n\geq j\geq 1}$ such that $\lim_{n\to\infty} \lambda_j^{(n)} = j$ for all $j \geq 1$. In some sense Example 4.4 comes up in natural way by considering a slight change of the values $(\lambda_j^{(n)})_{n\geq j\geq 1}$ in Example 4.3; an interesting feature is that the value $\zeta(2)$ (i.e. the Riemann-ζ function computed at 2) plays a crucial role in the computations; moreover we give a version of Example 4.4 which reveals a connection with the logarithmically weighted means as in the recent paper [3] (see Remark 4.5). The full LDP can be proved for Examples 4.1–4.3 only, since Corollary 3.6 can be applied only for those two examples.

The paper is organized as follows: in Section 2 we give some preliminary results and illustrate some facts about large deviations; in Section 3 we state our results; in Section 4 we present the examples; Section 5 contains the proofs.

2. Preliminaries on large deviations and first results

We start by giving some convergence results for the sequence $(R_n)_{n\geq 1}$.

Proposition 2.1. *Assume that*

$$\sup_{\substack{n\geq 1,\\ 1\leq j\leq n}} s_{j,n} = C < +\infty.$$

Then $R_n \longrightarrow 1$ *in probability as* $n \to \infty$.

Proof. Since $\mathbf{E}\big[T_j^{(n)}\big] = s_{j,n}$, we have

$$R_n - 1 = \frac{\sum_{j=1}^n \big(T_j^{(n)} - \mathbf{E}\big[T_j^{(n)}\big]\big)}{\gamma_n}$$

and, by Chebyshev inequality,

$$P\left(\left|\frac{\sum_{j=1}^n \big(T_j^{(n)} - \mathbf{E}\big[T_j^{(n)}\big]\big)}{\gamma_n}\right| > \epsilon\right)$$
$$\leq \frac{\mathbf{Var}\big(\sum_{j=1}^n T_j^{(n)}\big)}{\epsilon^2\gamma_n^2} = \frac{\sum_{j=1}^n s_{j,n}^2}{\epsilon^2\gamma_n^2} \leq C\left(\frac{\sum_{j=1}^n s_{j,n}}{\gamma_n}\right)\left(\frac{1}{\epsilon^2\gamma_n}\right) \to 0,$$

as $n \to \infty$. □

In some particular cases convergence in probability can be improved to almost sure convergence; this will be shown in the following

Proposition 2.2. *Let* $(X_j)_{j\geq 1}$ *be a sequence of i.i.d. random variables, with* $X_j \sim \mathcal{E}(1)$ *for every* j. *Assume that* $T_j^{(n)} := s_{j,n}X_j$. *If*

$$\frac{\sup_{1\leq j\leq n} s_{j,n}}{\gamma_n} = \mathrm{o}\Big(\frac{1}{\sqrt{n\log n}}\Big),$$

then $R_n \longrightarrow 1$ $P-$*a.s. as* $n \to \infty$.

Proof. Since $R_n - 1 = \sum_{j=1}^n a_{j,n}\big(X_j - 1\big)$ with $a_{j,n} = \frac{s_{j,n}}{\gamma_n}$, the result follows from Corollary 4 of [5]. □

The main asymptotic results in this paper concern large deviations. We start by recalling the definition of LDP, for which we refer to [2] (pages 4–5). Let $\mathcal{X}$ be a topological space equipped with its completed Borel σ-field. A sequence of $\mathcal{X}$-valued random variables $(Z_n)_{n\geq 1}$ satisfies the LDP with speed function v_n and rate function I if: $\lim_{n\to\infty} v_n = +\infty$; the function $I\colon \mathcal{X} \to [0;\infty]$ is lower semi-continuous;

$$\limsup_{n\to\infty} \frac{1}{v_n} \log P(Z_n \in F) \leq -\inf_{x\in F} I(x) \qquad \text{for all closed sets } F; \tag{2.1}$$

$$\liminf_{n\to\infty} \frac{1}{v_n} \log P(Z_n \in G) \geq -\inf_{x\in G} I(x) \qquad \text{for all open sets } G. \tag{2.2}$$

A rate function I is said to be *good* if its level sets $\{\{x \in \mathcal{X} : I(x) \leq \eta\} : \eta \geq 0\}$ are compact.

Throughout the paper we always have $\mathcal{X} = \mathbb{R}$ and we consider applications of Gärtner–Ellis Theorem (see e.g. Theorem 2.3.6 in [2]). The application of this theorem for the sequence $(Z_n)_{n\geq 1}$ consists in checking the existence of the function $\Lambda\colon \mathbb{R} \to (-\infty, \infty]$ defined by

$$\Lambda(\theta) := \lim_{n\to\infty} \frac{1}{v_n} \log \mathbf{E}[e^{\theta v_n Z_n}].$$

Then, if 0 belongs to the interior of $\{\theta \in \mathbb{R} : \Lambda(\theta) < \infty\}$ and if we set

$$I(x) := \sup_{\theta\in\mathbb{R}}\{\theta x - \Lambda(\theta)\}, \tag{2.3}$$

we have: (a) the upper bound (2.1); (b) the lower bound

$$\liminf_{n\to\infty} \frac{1}{v_n} \log P(Z_n \in G) \geq - \inf_{x\in G\cap\mathcal{F}} I(x) \qquad \text{for all open sets } G, \tag{2.4}$$

where $\mathcal{F}$ is the set of exposed points (see e.g. Definition 2.3.3 in [2]); (c) if Λ is essentially smooth (see e.g. Definition 2.3.5 in [2]) and lower semi-continuous, the LDP holds with a good rate function. Thus, if Λ is not essentially smooth, Gärtner–Ellis Theorem may provide a trivial non-sharp lower bound for open sets in terms of the exposed points of the rate function. It is exactly what happens in our case (see Theorem 3.1). Indeed Theorem 2.3.6 (b–c) in [2] would lead to the non-sharp lower bound (2.4) with $\mathcal{F} = \{1\}$, and this coincides with the sharp lower bound (2.2) if and only if $1 \in G$.

We point out that Corollary 3.6 here below provides an example in which the LDP holds, i.e. a case where the lower bound (2.4) (in terms of the exposed points) can be improved obtaining the lower bound for all open sets (2.2). Other examples are the one presented in Remark (d) after the statement of Theorem 2.3.6 in [2] where we have again a linear rate function (it is slightly different from the rate function I in eq. (3.2) below), and Exercise 2.3.24 in [2].

3. Statements of the main results

In order to apply Gärtner–Ellis Theorem, the first thing to do is to check the existence of the limit

$$\Lambda(\theta) := \lim_{n\to\infty} \frac{1}{v_n} \log \mathbf{E}[\exp(\theta v_n R_n)] = \lim_{n\to\infty} \frac{1}{v_n} \log \mathbf{E}\left[\exp\left(\theta \frac{v_n}{\gamma_n} \sum_{j=1}^{n} T_j^{(n)}\right)\right] \tag{3.1}$$

for all $\theta \in \mathbb{R}$, where v_n is the speed. We start with the following result where $v_n = \gamma_n$.

Theorem 3.1. *Let the following assumptions hold:*
(i) for each $n \geq 1$, the function $j \mapsto \lambda_j^{(n)}$ $(j = 1, \dots, n)$ is non-decreasing and $\lim_{n\geq j\to\infty} \lambda_j^{(n)} = +\infty$;

(ii) $n \mapsto \lambda_1^{(n)}$ *is ultimately monotone and* $\lim_{n\to\infty} \lambda_1^{(n)} = 1$.
Then the limit $\Lambda(\theta)$ *in* (3.1) *exists for every* $\theta \in \mathbb{R} \setminus \{1\}$ *with* $v_n = \gamma_n$*, and we have*

$$\Lambda(\theta) = \begin{cases} \theta & \textit{for } \theta < 1 \\ +\infty & \textit{for } \theta > 1. \end{cases}$$

It is easy to check that, if the limit $\Lambda(\theta)$ in (3.1) exists for $\theta = 1$, we have $\Lambda(1) \in [1, \infty]$ and the function I in (2.3) becomes

$$I(x) = \begin{cases} x - 1 & \text{for } x \geq 1 \\ +\infty & \text{for } x < 1. \end{cases} \tag{3.2}$$

Moreover, the function Λ is not essentially smooth; hence Gärtner–Ellis Theorem cannot give the sharp lower bound (2.2). In the next result we obtain a weak form of the lower bound by considering eq. (1.2.8) in [2].

Theorem 3.2. *Let the assumptions of Theorem 3.1 hold. Assume moreover that:*
(i) $\gamma_n \geq c \log n + \mathrm{o}(\log n)$ *ultimately* ($c > 0$ *constant);*
(ii) for $n \geq j \geq 1$, $\lambda_j^{(n)} - \lambda_1^{(n)} \geq j - 1$*;*
(iii) for each $n \geq 1$, $j \mapsto \frac{\lambda_j^{(n)} - \lambda_1^{(n)}}{j-1}$ ($j = 2, \ldots, n$) *is non-decreasing.*
Then, for $x \geq 1/c$ *and for all open sets* G *such that* $x \in G$*, we have*

$$\liminf_{n\to\infty} \frac{1}{\gamma_n} \log P(R_n \in G) \geq -I(x),$$

where I *is as in* (3.2).

Remark 3.3. Assumption (iii) of Theorem 3.2 holds for instance if, for each integer n, the (finite) sequence $j \mapsto \lambda_j^{(n)}$ is the restriction to $\mathbb{N} \cap [2, n]$ of a convex function $x \mapsto f_{(n)}(x)$ defined on $[1, n]$.

Remark 3.4. We notice for future reference that assumption (iii) of Theorem 3.2 implies that, for $i \neq j$,

$$\frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{|\lambda_j^{(n)} - \lambda_i^{(n)}|} \leq \frac{i-1}{|j-i|}.$$

In fact, for $j > i$, it gives

$$\frac{\lambda_j^{(n)} - \lambda_1^{(n)}}{\lambda_i^{(n)} - \lambda_1^{(n)}} \geq \frac{j-1}{i-1},$$

hence, by assumption (i) of Theorem 3.1,

$$\frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{|\lambda_j^{(n)} - \lambda_i^{(n)}|} = \frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{\lambda_j^{(n)} - \lambda_i^{(n)}} = \frac{1}{\frac{\lambda_j^{(n)} - \lambda_1^{(n)}}{\lambda_i^{(n)} - \lambda_1^{(n)}} - 1} \leq \frac{1}{\frac{j-1}{i-1} - 1} = \frac{i-1}{j-i} = \frac{i-1}{|j-i|}.$$

The proof for $i < j$ is similar.

Remark 3.5. A careful look at the proofs shows that assumption (ii) of Theorem 3.2 could be relaxed as follows:
$(ii)'$ There exists a sequence $(a_n)_{n\geq 1}$, with $\lim_{n\to\infty} a_n = 1$, such that, for every integer n and for each $j = 2, \ldots, n$

$$\lambda_j^{(n)} - \lambda_1^{(n)} \geq a_n(j-1).$$

It follows that, if $\big(\lambda_j^{(n)}\big)_{j\leq n}$ verifies $(ii)'$, the same happens for $\big(\widetilde{\lambda}_j^{(n)}\big)_{j\leq n}$ such that

$$\widetilde{\lambda}_j^{(n)} = d_n\big(\lambda_j^{(n)} + c_n\big),$$

where $(c_n)_{n\geq 1}$ is any sequence and $\lim_{n\to\infty} d_n = 1$.

It is obvious that the weaker form of the lower bound provided by Theorem 3.2 coincides with the lower bound (2.2) if $c \geq 1$. Thus, putting together the results of Theorems 3.1 and 3.2 and Gärtner Ellis Theorem, we get the following corollary.

Corollary 3.6. *Let the whole set of assumptions (i) and (ii) of Theorem 3.1 and (i), (ii) and (iii) of Theorem 3.2 hold. Moreover we assume that the limit $\Lambda(\theta)$ in* (3.1) *exists for $\theta = 1$ with $v_n = \gamma_n$. Then, if $c \geq 1$, $(R_n)_{n\geq 1}$ satisfies an LDP with speed $v_n = \gamma_n$ and rate function I as* (3.2).

4. Examples

In this section we present some examples checking for each of them that the assumptions of Theorems 3.1–3.2 hold. We remark that Corollary 3.6 is in force (and therefore the LDP holds) for Examples 4.1–4.3, where $c \geq 1$. Here is the first example.

Example 4.1. Let $(\lambda_j^{(n)})_{j\leq n}$ be defined by $\lambda_j^{(n)} := j$ for $j = 1, \ldots, n$ and $n \geq 1$.

Remark 4.2. Let $\{X_n : n \geq 1\}$ be independent random variables such that $X_n \sim \mathcal{E}(1)$ for all $n \geq 1$ and, for every $n \geq 1$, consider the order statistics $X_{n,n} \leq \cdots \leq X_{1,n}$ of $X_1, \ldots, X_n$; then the *spacings* $(T_j)_{j\leq n}$ defined by

$$T_j^{(n)} := X_{j,n} - X_{j+1,n}, \qquad j = 1, \ldots, n \qquad (where\ X_{n+1,n} = 0),$$

meet the framework of Example 4.1 (see for instance [1], Ex. 4.1.5, p. 185).

In this case the assumptions of Theorems 3.1–3.2 can be easily checked. Here we only notice that assumption (i) of Theorem 3.2 holds with $c = 1$ since $\gamma_n = \sum_{j=1}^n \frac{1}{j} \geq \log(n+1)$. Finally we can apply Corollary 3.6 because we have $\Lambda(1) = 1$ with $v_n = \gamma_n$ (this can be easily checked and we omit the details).

In the next Example 4.3 we consider a particular choice of the values $(\lambda_j^{(n)})_{j\leq n}$. It is worth noting that $\lim_{n\to\infty} \lambda_j^{(n)} = j$, which are the parameters in Example 4.1.

Example 4.3. Let $(\lambda_j^{(n)})_{j\leq n}$ be defined by $\lambda_j^{(n)} := \frac{1}{\frac{1}{j}-\frac{1}{n+1}} = \frac{(n+1)j}{n+1-j}$ for $j = 1,\dots,n$ and $n \geq 1$.

In this case the assumptions of Theorems 3.1–3.2 can be checked as follows. The assumptions (i) and (ii) of Theorem 3.1 are obvious. As to (i) of Theorem 3.2 (again with $c = 1$) we notice that

$$\gamma_n = \sum_{j=1}^{n} \frac{1}{j} - \sum_{j=1}^{n} \frac{1}{n+1} = \sum_{j=1}^{n} \frac{1}{j} - \frac{n}{n+1} \geq \log(n+1) - \frac{n}{n+1}.$$

Assumption (ii) of Theorem 3.2 holds since

$$\lambda_j^{(n)} - \lambda_1^{(n)} = \frac{n+1}{n+1-j} \cdot \frac{n+1}{n}(j-1) \geq j-1;$$

moreover, it is easily seen that the function $x \mapsto f_{(n)}(x) = \frac{(n+1)x}{n+1-x}$ is convex, and we deduce that also (iii) of Theorem 3.2 is verified, by Remark 3.3. Finally, as for Example 4.1, we can apply Corollary 3.6 because we have $\Lambda(1) = 1$ with $v_n = \gamma_n$ (this can be easily checked and we omit the details).

In the previous Example 4.3 we had

$$\frac{1}{\lambda_j^{(n)}} = \frac{1}{j} - \frac{1}{n+1} = \int_j^{n+1} \frac{1}{x^2}\, dx.$$

A natural idea is to investigate what happens if we substitute the integral with the sum over integers, i.e. if we consider $\sum_{k=j}^{n} \frac{1}{k^2}$ instead of $\int_j^{n+1} \frac{1}{x^2}\, dx$. Since in such a case $\lim_{n\to\infty} \frac{1}{\sum_{k=1}^{n} \frac{1}{k^2}} = \frac{1}{\zeta(2)} = \frac{6}{\pi^2} \simeq 0.608 \neq 1$, assumption (ii) of Theorem 3.1 is satisfied if we perform a "normalization"; this leads to the following

Example 4.4. Let $(\lambda_j^{(n)})_{j\leq n}$ be defined by $\lambda_j^{(n)} := \frac{\zeta(2)}{\sum_{k=j}^{n} \frac{1}{k^2}}$ for $j = 1,\dots,n$ and $n \geq 1$.

Remark 4.5. Let $(\lambda_j^{(n)})_{j\leq n}$ be as in Example 4.4 and let $(U_j)_{j\geq 1}$ be a sequence of independent random variables, and assume that they are uniformly distributed on $(0,1)$. Then we set

$$T_j^{(n)} := \frac{1}{\zeta(2)} \sum_{k=j}^{n} \frac{1}{k} F_k^{-1}(U_j) \qquad j = 1,\dots,n,$$

where $F_k^{-1}(u) = -\frac{1}{k}\log(1-u)$ (for $u \in (0,1)$) is the inverse of the distribution function of a random variable $Z \sim \mathcal{E}(k)$. This is a version of Example 4.4 because, for each fixed $n \geq 1$, $(T_1^{(n)},\dots,T_n^{(n)})$ are independent (obvious) and, for all $j = 1\dots,n$, $T_j^{(n)} = \frac{1}{\zeta(2)} \sum_{k=j}^{n} \frac{1}{k^2} F_1^{-1}(U_j) = (\lambda_j^{(n)})^{-1} F_1^{-1}(U_j)$ with $F_1^{-1}(U_j) \sim \mathcal{E}(1)$, and therefore $T_j^{(n)} \sim \mathcal{E}(\lambda_j^{(n)})$. Finally we remark that R_n is a logarithmically weighted mean as in [3] because, if we set $X_k := \sum_{j=1}^k F_k^{-1}(U_j)$, we have

$$R_n = \frac{\sum_{j=1}^n T_j^{(n)}}{\gamma_n} = \frac{\sum_{j=1}^n \frac{1}{\zeta(2)} \sum_{k=j}^n \frac{1}{k} F_k^{-1}(U_j)}{\sum_{j=1}^n \frac{1}{\zeta(2)} \sum_{k=j}^n \frac{1}{k^2}}$$
$$= \frac{\sum_{k=1}^n \frac{1}{k} \sum_{j=1}^k F_k^{-1}(U_j)}{\sum_{k=1}^n \sum_{j=1}^k \frac{1}{k^2}} = \frac{\sum_{k=1}^n \frac{1}{k} X_k}{\sum_{k=1}^n \frac{1}{k}}.$$

Now we have to check all the conditions of Theorems 3.1–3.2 for Example 4.4. The assumptions of Theorem 3.1 are obvious. Assumption (i) of Theorem 3.2 holds since

$$\gamma_n = \frac{1}{\zeta(2)} \sum_{j=1}^n \sum_{k=j}^n \frac{1}{k^2} = \frac{1}{\zeta(2)} \sum_{k=1}^n \sum_{j=1}^k \frac{1}{k^2} = \frac{1}{\zeta(2)} \sum_{k=1}^n \frac{1}{k} \geq \frac{1}{\zeta(2)} \log(n+1).$$

Note that in this case we have $c = \frac{1}{\zeta(2)} < 1$ and Corollary 3.6 cannot be applied; for completeness we check $\Lambda(1) = 1$ with $v_n = \gamma_n$.

Proof of $\Lambda(1) = 1$ with $v_n = \gamma_n$ for Example 4.4. We have to check that

$$\lim_{n\to\infty} \frac{-\sum_{j=1}^n \log(1 - s_{j,n})}{\sum_{j=1}^n s_{j,n}} = 1$$

because $\gamma_n = \sum_{j=1}^n s_{j,n}$ and

$$\log \mathbf{E}\left[\exp\left(\sum_{j=1}^n T_j^{(n)}\right)\right] = \sum_{j=1}^n \log \mathbf{E}\left[e^{T_j^{(n)}}\right]$$
$$= \sum_{j=1}^n \log \frac{\lambda_j^{(n)}}{\lambda_j^{(n)} - 1} = -\sum_{j=1}^n \log(1 - s_{j,n}).$$

Moreover, since $-\log(1 - s_{j,n}) \geq s_{j,n}$, it is enough to check

$$\limsup_{n\to\infty} \frac{-\sum_{j=1}^n \log(1 - s_{j,n})}{\sum_{j=1}^n s_{j,n}} \leq 1$$

and, noting that

$$\sum_{j=1}^n s_{j,n} = \frac{1}{\zeta(2)} \sum_{j=1}^n \sum_{k=j}^n \frac{1}{k^2} = \frac{1}{\zeta(2)} \sum_{k=1}^n \sum_{j=1}^k \frac{1}{k^2} = \frac{1}{\zeta(2)} \sum_{k=1}^n \frac{1}{k} \sim \frac{1}{\zeta(2)} \log n,$$

this is equivalent to

$$\limsup_{n\to\infty} \frac{-\sum_{j=1}^n \log(1 - s_{j,n})}{\log n} \leq \frac{1}{\zeta(2)}. \tag{4.1}$$

Now, since $s_{j,n} \leq s_{j,\infty} = 1 - s_{1,j-1}$ and $x \in [0,1) \mapsto -\log(1-x)$ is an increasing function, we get $\frac{-\sum_{j=1}^{n}\log(1-s_{j,n})}{\log n} \leq \frac{-\sum_{j=1}^{n}\log(s_{1,j-1})}{\log n}$; thus (4.1) is implied by

$$\lim_{n\to\infty} \frac{-\sum_{j=1}^{n}\log(s_{1,j-1})}{\log n} = \frac{1}{\zeta(2)}$$

or, equivalently (by Cesaro Theorem), $\lim_{n\to\infty} -n\log(s_{1,n-1}) = \frac{1}{\zeta(2)}$; in conclusion (4.1) is implied by

$$\frac{1}{\zeta(2)} = \lim_{n\to\infty} n(1 - s_{1,n-1}) = \lim_{n\to\infty} \frac{n}{\zeta(2)} \sum_{k=n}^{\infty} \frac{1}{k^2},$$

which can be be checked noting that

$$\frac{1}{\zeta(2)} = \frac{n}{\zeta(2)} \int_{n}^{\infty} \frac{1}{x^2} dx \leq \frac{n}{\zeta(2)} \sum_{k=n}^{\infty} \frac{1}{k^2} \leq \frac{n}{\zeta(2)} \int_{n-1}^{\infty} \frac{1}{x^2} dx = \frac{1}{\zeta(2)} \frac{n}{n-1}. \qquad \square$$

We conclude with the proof of assumptions (ii)–(iii) of Theorem 3.2 for Example 4.4.

Proof of assumption (ii) of Theorem 3.2 for Example 4.4. The condition is obvious for $j = 1$ and, from now on, we assume that $j = 2, \ldots, n$. Since

$$\begin{aligned}\lambda_j^{(n)} - \lambda_1^{(n)} &= \zeta(2) \frac{\sum_{k=1}^{j-1} \frac{1}{k^2}}{\left(\sum_{k=1}^{n} \frac{1}{k^2}\right)\left(\sum_{k=j}^{n} \frac{1}{k^2}\right)} \geq \frac{\sum_{k=1}^{j-1} \frac{1}{k^2}}{\left(\sum_{k=j}^{n} \frac{1}{k^2}\right)} \\ &= \frac{\sum_{k=1}^{j-1} \frac{1}{k^2}}{\sum_{k=1}^{n} \frac{1}{k^2} - \sum_{k=1}^{j-1} \frac{1}{k^2}} \geq \frac{\sum_{k=1}^{j-1} \frac{1}{k^2}}{\zeta(2) - \sum_{k=1}^{j-1} \frac{1}{k^2}} = \frac{1}{\zeta(2)\left(\sum_{k=1}^{j-1} \frac{1}{k^2}\right)^{-1} - 1},\end{aligned}$$

it suffices to show that the last quantity above is $\geq j - 1$ or, in equivalent form, that

$$\frac{\zeta(2)}{\sum_{k=1}^{j-1} \frac{1}{k^2}} \leq \frac{j}{j-1}.$$

With some algebra, the inequality to be proved can be transformed into the equivalent one

$$\zeta(2) - \sum_{k=j}^{\infty} \frac{1}{k^2} = \sum_{k=1}^{j-1} \frac{1}{k^2} \geq \zeta(2)\left(1 - \frac{1}{j}\right),$$

or, after simplification,

$$a_j := -\sum_{k=j}^{\infty} \frac{1}{k^2} + \frac{\zeta(2)}{j} \geq 0.$$

Since $\lim_{j\to\infty} a_j = 0$, it is enough to show that (a_j) is non-increasing, i.e. for every j

$$-\sum_{k=j+1}^{\infty}\frac{1}{k^2}+\frac{\zeta(2)}{j+1}\leq-\sum_{k=j}^{\infty}\frac{1}{k^2}+\frac{\zeta(2)}{j},$$

and therefore

$$0\geq\sum_{k=j}^{\infty}\frac{1}{k^2}-\sum_{k=j+1}^{\infty}\frac{1}{k^2}+\zeta(2)\Big(\frac{1}{j+1}-\frac{1}{j}\Big)=\frac{1}{j^2}-\frac{\zeta(2)}{j(j+1)}.$$

Multiplying by $j^2(j+1)$ we get the equivalent inequality

$$\big(\zeta(2)-1\big)j\geq 1,$$

which is true since

$$\big(\zeta(2)-1\big)j\geq 2\big(\zeta(2)-1\big)\simeq 1.28. \qquad \square$$

Proof of assumption (iii) of Theorem 3.2 for Example 4.4. For $k\geq 1$ we set $s_k := \sum_{h=1}^{k}\frac{1}{h^2}$ and, for $n\geq 2$ and $j=1,\ldots,n-1$, we set $d_j^{(n)} := \frac{s_j}{\left(s_n-s_j\right)j}$. Then we have

$$d_{j-1}^{(n)}=\frac{s_{j-1}}{(s_n-s_{j-1})(j-1)}=\frac{\frac{\lambda_j^{(n)}}{\lambda_1^{(n)}}-1}{j-1}=\frac{1}{\lambda_1^{(n)}}\cdot\frac{\lambda_j^{(n)}-\lambda_1^{(n)}}{j-1}\qquad(j=2,\ldots,n);$$

therefore we need to prove that the finite sequence $\left(d_j^{(n)}\right)_j$ is non-decreasing, i.e.

$$d_{j-1}^{(n)}\leq d_j^{(n)}\qquad(j=2,\ldots,n-1).$$

After rearranging we see that this is equivalent to

$$s_n\leq\frac{s_{j-1}s_j}{js_{j-1}-(j-1)s_j}\qquad(j=2,\ldots,n-1);\tag{4.2}$$

moreover $s_n\uparrow\zeta(2)$ as $n\uparrow\infty$ and the right hand side in (4.2) tends to $\zeta(2)$ as $j\to\infty$; hence it suffices to show that the right hand side in (4.2) is a non-increasing function of j, i.e.

$$\frac{s_{j-1}s_j}{js_{j-1}-(j-1)s_j}\geq\frac{s_js_{j+1}}{(j+1)s_j-js_{j+1}}\qquad(j\geq 2).$$

We check this inequality with some algebra and by taking into account that $s_{j-1}=s_j-\frac{1}{j^2}$ and $s_{j+1}=s_j+\frac{1}{(j+1)^2}$; indeed we get the inequality

$$s_j\leq\frac{2j}{j+1}=2\left(1-\frac{1}{j+1}\right),$$

which is obviously true, since

$$s_j=\sum_{h=1}^{j}\frac{1}{h^2}\leq\sum_{h=1}^{j}\frac{2}{h(h+1)}=2\sum_{h=1}^{j}\left(\frac{1}{h}-\frac{1}{h+1}\right)=2\left(1-\frac{1}{j+1}\right). \qquad \square$$

5. The proofs

Recall the notations $s_{j,n} := \big(\lambda_j^{(n)}\big)^{-1}$ and $\gamma_n = \sum_{j=1}^n s_{j,n}$, which will be systematically used in the sequel.

Proof of Theorem 3.1. We give several proofs according to different values of θ.
• Let us consider first the case $\theta < 1$ (excluding the case $\theta = 0$, which is trivial). Fix $\delta \in (0, \frac{1}{2})$. Assumption (i) assures that exists j_0 such that, for $j_0 \le j \le n$, we have

$$\big|s_{j,n}\theta\big| < \delta.$$

We write

$$\begin{aligned}
&\frac{1}{\gamma_n}\log \mathbf{E}\left[\exp\left(\theta\sum_{j=1}^n T_j^{(n)}\right)\right] \\
&= \frac{1}{\gamma_n}\sum_{j=1}^n \log \mathbf{E}\left[\exp\left(\theta T_j^{(n)}\right)\right] = -\frac{\sum_{j=1}^n \log\Big(1 - s_{j,n}\theta\Big)}{\gamma_n} \\
&= \left(-\frac{\sum_{j=1}^{j_0} \log\Big(1 - s_{j,n}\theta\Big)}{\gamma_n}\right) + \left(-\frac{\sum_{j=j_0+1}^{n} \log\Big(1 - s_{j,n}\theta\Big)}{\gamma_n}\right) = A_n + B_n.
\end{aligned}$$

We shall prove that

(a) $\lim_{n\to\infty} A_n = 0$; (b) $\theta \le \liminf_{n\to\infty} B_n \le \limsup_{n\to\infty} B_n \le \theta + |\theta|\delta$.

Proof of (a). We treat separately the two cases (a_1) $\theta > 0$ and (a_2) $\theta < 0$.

Proof of (a_1). Since $\theta < 1$, there exists $\epsilon > 0$ such that $\theta < 1 - \epsilon < 1$. By assumption (ii), $\lambda_1^{(n)} > 1 - \epsilon$ ultimately, so that (i) implies that, for every $j \le n$,

$$s_{j,n}\theta \le s_{1,n}\theta \le \frac{\theta}{1-\epsilon} < 1.$$

Hence ultimately we have

$$0 \le A_n = -\frac{\sum_{j=1}^{j_0} \log\Big(1 - s_{j,n}\theta\Big)}{\gamma_n} \le -\frac{\sum_{j=1}^{j_0} \log\Big(1 - \frac{\theta}{1-\epsilon}\Big)}{\gamma_n} \to 0, \qquad n \to \infty.$$

Proof of (a_2). In this case we have $s_{j,n}\theta \in (-\delta, 0]$, and therefore $0 \le \log(1 - s_{j,n}\theta) = \log(1 + s_{j,n}|\theta|)$; moreover the sequence $\big(s_{1,n}\big)_n$, being convergent (to 1), is bounded by some positive real number C; hence for every $j \le n$ we have $s_{j,n} \le s_{1,n} \le C$, which gives

$$|A_n| = \frac{\sum_{j=1}^{j_0} \log\Big(1 - s_{j,n}\theta\Big)}{\gamma_n} = \frac{\sum_{j=1}^{j_0} \log\Big(1 + s_{j,n}|\theta|\Big)}{\gamma_n}$$

$$\leq \frac{\sum_{j=1}^{j_0} \left| \log\left(1 + C|\theta|\right) \right|}{\gamma_n} \to 0, \quad n \to \infty.$$

Proof of (b). For $|x| < 1/2$ we have $x \leq -\log(1-x) \leq x + x^2$; hence

$$\theta \cdot \frac{\sum_{j=j_0+1}^{n} s_{j,n}}{\gamma_n} \leq B_n \leq \theta \cdot \frac{\sum_{j=j_0+1}^{n} s_{j,n}}{\gamma_n} + \theta^2 \cdot \frac{\sum_{j=j_0+1}^{n} s_{j,n}^2}{\gamma_n},$$

and it is enough to check
(b_1) $\lim_{n\to\infty} \frac{\sum_{j=j_0+1}^{n} s_{j,n}}{\gamma_n} = 1$ and (b_2) $\limsup_{n\to\infty} \frac{\sum_{j=j_0+1}^{n} s_{j,n}^2}{\gamma_n} \leq \frac{\delta}{|\theta|}$.
Proof of (b_1). We have

$$\frac{\sum_{j=j_0+1}^{n} s_{j,n}}{\gamma_n} = 1 - \frac{\sum_{j=1}^{j_0} s_{j,n}}{\gamma_n},$$

and (as we have seen before) $s_{j,n} \leq s_{1,n} \leq C$ for every $j \leq n$; we deduce that

$$0 \leq \frac{\sum_{j=1}^{j_0} s_{j,n}}{\gamma_n} \leq \frac{\sum_{j=1}^{j_0} C}{\gamma_n} \to 0, \qquad n \to \infty.$$

Proof of (b_2). By construction we have $s_{j,n}|\theta| < \delta$ for $n \geq j \geq j_0$; thus

$$0 \leq \frac{\sum_{j=j_0+1}^{n} s_{j,n}^2}{\gamma_n} \leq \frac{\delta}{|\theta|} \cdot \frac{\sum_{j=j_0+1}^{n} s_{j,n}}{\gamma_n} \leq \frac{\delta}{|\theta|} \cdot \frac{\sum_{j=1}^{n} s_{j,n}}{\gamma_n} = \frac{\delta}{|\theta|}.$$

• We pass to the case $\theta > 1$. Since $\lim_{n\to\infty} \lambda_1^{(n)} = 1$, there exists an integer n_0 such that, for every $n > n_0$, we have $\theta > \lambda_1^{(n)}$; hence

$$\frac{1}{\gamma_n} \log \mathbf{E}\left[\exp\left(\theta \sum_{j=1}^{n} T_j^{(n)}\right)\right] \geq \frac{1}{\gamma_n} \log \mathbf{E}\left[\exp\left(\theta T_1^{(n)}\right)\right] = +\infty. \qquad \square$$

Proof of Theorem 3.2. The inequality to be proved is trivial if $x < 1$ (because $I(x) = +\infty$) and if $x = 1$ it holds by Proposition 2.1 (because $I(x) = 0$); so, throughout this proof, we restrict our attention to the case $x > 1$. We choose $\epsilon > 0$ so small to have $(x - \epsilon, x + \epsilon) \subset G$; hence

$$P(R_n \in G) \geq P(x - \epsilon < R_n < x + \epsilon) \geq P(x < R_n < x + \epsilon).$$

The main proof consists in showing that we have

$$\liminf_{n\to\infty} \frac{1}{\gamma_n} \log P(x < R_n < x + \epsilon) \geq 1 - x - \epsilon;$$

(in fact we easily get

$$\liminf_{n\to\infty} \frac{1}{\gamma_n} \log P(R_n \in G) \geq 1 - x - \epsilon,$$

and let ϵ go to zero).

Let F and f be the distribution function and the density of $\sum_{j=1}^n T_j^{(n)}$ respectively. By Lagrange Theorem, there exists $\xi \in \big(x, x+\epsilon\big)$ such that

$$P(x < R_n < x+\epsilon) = F\big((x+\epsilon)\gamma_n\big) - F\big(x\gamma_n\big) = \epsilon \cdot \gamma_n \cdot f(\xi\gamma_n).$$

Passing to the logarithm and dividing by γ_n we get

$$\frac{1}{\gamma_n} \log P(x < R_n < x+\epsilon) = \frac{\log \epsilon}{\gamma_n} + \frac{\log \gamma_n}{\gamma_n} + \frac{\log\big(f(\xi\gamma_n)\big)}{\gamma_n},$$

and of course only the last summand has to be considered. According to a well known formula (see for instance [4], p. 308 and ff.), f has the form

$$\begin{aligned} f(t) &= (-1)^{n-1} \lambda_1^{(n)} \cdot \ldots \cdot \lambda_n^{(n)} \sum_{j=1}^n \frac{e^{-\lambda_j^{(n)} t}}{\prod_{i\neq j} (\lambda_j^{(n)} - \lambda_i^{(n)})} \\ &= \lambda_1^{(n)} \cdot \ldots \cdot \lambda_n^{(n)} \frac{e^{-\lambda_1^{(n)} t}}{\prod_{i\neq 1} (\lambda_i^{(n)} - \lambda_1^{(n)})} \cdot \left\{ 1 - \sum_{j=2}^n e^{-(\lambda_j^{(n)} - \lambda_1^{(n)})t} \cdot \prod_{i\neq 1,j} \frac{\lambda_1^{(n)} - \lambda_i^{(n)}}{\lambda_j^{(n)} - \lambda_i^{(n)}} \right\} \end{aligned}$$

(note that this formula is allowed because the values $\lambda_1^{(n)}, \ldots, \lambda_n^{(n)}$ are all different by the hypotheses). Then we take the logarithm and we get

$$\begin{aligned} \log f(t) = &\sum_{j=1}^n \log \lambda_j^{(n)} - \lambda_1^{(n)} t - \sum_{j=2}^n \log\big(\lambda_j^{(n)} - \lambda_1^{(n)}\big) \\ &+ \log \left\{ 1 - \sum_{j=2}^n e^{-(\lambda_j^{(n)} - \lambda_1^{(n)})t} \cdot \prod_{i\neq 1,j} \frac{\lambda_1^{(n)} - \lambda_i^{(n)}}{\lambda_j^{(n)} - \lambda_i^{(n)}} \right\}. \end{aligned}$$

Calculating in $t = \xi\gamma_n$ and dividing by γ_n we find

$$\begin{aligned} \frac{\log\big(f(\xi\gamma_n)\big)}{\gamma_n} = &\left(\frac{\log \lambda_1^{(n)}}{\gamma_n}\right) + \left(\frac{\sum_{j=2}^n \log \frac{\lambda_j^{(n)}}{\lambda_j^{(n)} - \lambda_1^{(n)}}}{\gamma_n}\right) + \left(-\lambda_1^{(n)}\xi\right) \\ &+ \left(\frac{1}{\gamma_n} \cdot \log \left\{ 1 - \sum_{j=2}^n e^{-(\lambda_j^{(n)} - \lambda_1^{(n)})\xi\gamma_n} \cdot \prod_{i\neq 1,j} \frac{\lambda_1^{(n)} - \lambda_i^{(n)}}{\lambda_j^{(n)} - \lambda_i^{(n)}} \right\}\right) \end{aligned}$$

$$=: A_n + B_n + C_n + D_n.$$

By the assumption (ii) of Theorem 3.1 we have $\lim_{n\to\infty} A_n = 0$ and $\lim_{n\to\infty} C_n = -\xi > -x - \epsilon$. So the proof will be complete if we show that (a) $\liminf_{n\to\infty} B_n \geq 1$ and (b) $\lim_{n\to\infty} D_n = 0$.

Proof of (a). For every pair x, y, with $0 < x < y$ the inequality

$$\log \frac{y}{y-x} \geq \frac{x}{y},$$

(which comes from $\log(1+t) \leq t$ putting $t = -\frac{x}{y}$), applied to $y = \lambda_j^{(n)}$ and $x = \lambda_1^{(n)}$ gives

$$B_n = \frac{\sum_{j=2}^n \log \frac{\lambda_j^{(n)}}{\lambda_j^{(n)} - \lambda_1^{(n)}}}{\gamma_n} \geq \frac{\sum_{j=2}^n \frac{\lambda_1^{(n)}}{\lambda_j^{(n)}}}{\gamma_n} = \lambda_1^{(n)} \frac{\sum_{j=2}^n \frac{1}{\lambda_j^{(n)}}}{\gamma_n} \to 1,$$

by assumption (ii) of Theorem 3.1.

Proof of (b). It suffices to show that $\lim_{n\to\infty} a_n = 0$, where

$$a_n := -\sum_{j=2}^n e^{-(\lambda_j^{(n)} - \lambda_1^{(n)})\xi\gamma_n} \cdot \prod_{i\neq 1,j} \frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{\lambda_i^{(n)} - \lambda_j^{(n)}}.$$

To begin with, we write

$$- \prod_{i\neq 1,j} \frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{\lambda_i^{(n)} - \lambda_j^{(n)}} = - \prod_{i=2}^{j-1} \frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{\lambda_i^{(n)} - \lambda_j^{(n)}} \cdot \prod_{i=j+1}^{n} \frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{\lambda_i^{(n)} - \lambda_j^{(n)}}$$
$$= (-1)^{j-1} \prod_{i=2}^{j-1} \frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{\lambda_j^{(n)} - \lambda_i^{(n)}} \cdot \prod_{i=j+1}^{n} \frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{\lambda_i^{(n)} - \lambda_j^{(n)}} = (-1)^{j-1} \prod_{i\neq 1,j} \frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{|\lambda_i^{(n)} - \lambda_j^{(n)}|};$$

by assumption (ii) of Theorem 3.2 and Remark 3.4, we have

$$|a_n| \leq \sum_{j=2}^n e^{-(\lambda_j^{(n)} - \lambda_1^{(n)})\xi\gamma_n} \cdot \prod_{i\neq 1,j} \frac{\lambda_i^{(n)} - \lambda_1^{(n)}}{|\lambda_i^{(n)} - \lambda_j^{(n)}|} \leq \sum_{j=2}^n e^{-(j-1)\xi\gamma_n} \cdot \prod_{i\neq 1,j} \left(\frac{i-1}{|i-j|}\right);$$

hence, by assumption (i) of Theorem 3.2,

$$|a_n| \leq \sum_{j=2}^n \left(\frac{1}{\mathrm{e}^{b_n}}\right)^{j-1} \cdot \left(\prod_{i\neq 1,j} \frac{i-1}{|i-j|}\right),$$

where $b_n := c\xi \log n + \mathrm{o}(\log n)$. Now

$$\prod_{i\neq 1,j} \frac{i-1}{|i-j|} = \frac{(n-1)!}{j-1} \frac{1}{\prod_{i=2}^{j-1}(j-i)} \frac{1}{\prod_{i=j+1}^{n}(i-j)} = \frac{(n-1)!}{(j-1)!(n-j)!} = \binom{n-1}{j-1};$$

thus

$$|a_n| \leq \sum_{j=2}^{n} \binom{n-1}{j-1} \left(\frac{1}{\mathrm{e}^{b_n}}\right)^{j-1} = \sum_{j=1}^{n-1} \binom{n-1}{j} \left(\frac{1}{\mathrm{e}^{b_n}}\right)^{j} = \left(1 + \frac{1}{\mathrm{e}^{b_n}}\right)^{n-1} - 1.$$

Now we show that

$$\lim_{n\to\infty} \left(1 + \frac{1}{\mathrm{e}^{b_n}}\right)^{n-1} = 1,$$

or equivalently

$$\lim_{n\to\infty} (n-1)\log\left(1 + \frac{1}{\mathrm{e}^{b_n}}\right) = 0.$$

In fact, since $c\xi > 1$ (because $\xi > x$ and $x \geq 1/c$), we have

$$b_n - \log(n-1) = c\xi \log n + \mathrm{o}(\log n) - \log(n-1) \to +\infty, \qquad n \to +\infty,$$

whence

$$\lim_{n\to\infty} (n-1)\log\left(1 + \frac{1}{\mathrm{e}^{b_n}}\right) = \lim_{n\to\infty} \frac{n-1}{\mathrm{e}^{b_n}} = 0. \qquad \square$$

Acknowledgements. We thank the referee for several hints and comments which led to an improvement of the paper. We also thank F. Amoroso for his help for number-theoretical results concerning Example 4.4 presented in the first version of the paper.

References

[1] P. Embrechts, C. Klüppelberg, T. Mikosch, Modelling Extremal Events, *Springer*, 1997.

[2] A. Dembo, O. Zeitouni, Large Deviations Techniques and Applications (Second Edition), *Springer-Verlag*, 1998.

[3] R. Giuliano, C. Macci, Large deviation principles for sequences of logarithmically weighted means. *J. Math. Anal. Appl.* **378**, no. 2 (2011) 555–570.

[4] S.M. Ross, Introduction to Probability Models (Tenth Edition). *Academic Press*, 2010.

[5] S.H. Sung, Almost sure convergence for weighted sums of i.i.d. random variables (II). *Bull. Korean Math. Soc.* **33**, no. 3 (1996) 419–425.

Annales Mathematicae et Informaticae
39 *(2012) pp. 125–157*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Strong limit theorems for random fields

Allan Gut

Department of Mathematics, Uppsala University, Uppsala, Sweden
e-mail: allan.gut@math.uu.se

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

The aim of the present paper is to review some joint work with Ulrich Stadtmüller concerning random field analogs of the classical strong laws.

In the first half we start, as background information, by quoting the law of large numbers and the law of the iterated logarithm for random sequences as well as for random fields, and the law of the single logarithm for sequences. We close with a one-dimensional LSL pertaining to windows, whose edges expand in an "almost linear fashion", viz., the length of the nth window equals, for example, $n/\log n$ or $n/\log\log n$. A sketch of the proof will also be given.

The second part contains some extensions of the LSL to random fields, after which we turn to convergence rates in the law of large numbers. Departing from the now legendary Baum–Katz theorem in 1965, we review a number of results in the multiindex setting. Throughout main emphasis is on the case of "non-equal expansion rates", viz., the case when the edges along the different directions expand at different rates. Some results when the power weights are replaced by almost exponential weights are also given.

We close with some remarks on martingales and the strong law.

Keywords: i.i.d. random variables, law of large numbers, law of the iterated logarithm, law of the single logarithm, random field, multiindex.

MSC: Primary 60F05, 60F15, 60G70, 60G60; Secondary 60G40.

1. Introduction

Let X, X_1, X_2, ... be independent, identically distributed (i.i.d.) random variables with partial sums S_n, $n \geq 1$, and set $S_0 = 0$. The two most famous strong laws are the Kolmogorov strong law and the Hartman–Wintner Law of the iterated logarithm:

Theorem 1.1 (The Kolmogorov strong law — LLN)**.** *Suppose that $X, X_1, X_2, \ldots$ are i.i.d. random variables with partial sums S_n, $n \geq 1$.*

(a) *If $E|X| < \infty$ and $E\,X = \mu$, then*

$$\frac{S_n}{n} \stackrel{a.s.}{\to} \mu \quad as \quad n \to \infty.$$

(b) *If $\frac{S_n}{n} \stackrel{a.s.}{\to} c$ for some constant c, as $n \to \infty$, then*

$$E|X| < \infty \quad and \quad c = E\,X.$$

(c) *If $E|X| = \infty$, then*

$$\limsup_{n\to\infty} \frac{|S_n|}{n} = +\infty.$$

Remark 1.2. Strictly speaking, we presuppose in (b) that the limit can only be a constant. That this is indeed the case follows from the Kolmogorov zero–one law. Considering this, (c) is somewhat more general than (b). For proofs and details, see e.g. Gut (2007), Chapter 6.

Theorem 1.3 (The Hartman–Wintner law of the iterated logarithm — LIL)**.** *Suppose that $X, X_1, X_2, \ldots$ are i.i.d. random variables with mean 0 and finite variance σ^2, and set $S_n = \sum_{k=1}^n X_k$, $n \geq 1$. Then*

$$\limsup_{n\to\infty} \, (\liminf_{n\to\infty}) \frac{S_n}{\sqrt{2\sigma^2 n \log\log n}} = +1 \quad (-1) \quad a.s. \tag{1.1}$$

Conversely, if

$$P\Big(\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{n \log\log n}} < \infty\Big) > 0,$$

then $E\,X^2 < \infty$, $E\,X = 0$, and (1.1) holds.

The sufficiency is due to Hartman and Wintner (1941). The necessity is due to Strassen (1966). For this and more, see e.g. Gut (2007), Chapter 8.

Remark 1.4. The Kolmogorov zero–one law tells us that the limsup is finite with probability zero or one, and, if finite, the limit equals a constant almost surely. Thus, assuming in the converse that the probability is positive is in reality assuming that it is equal to 1. This remark also applies to (e.g.) Theorem 1.8.

The Kolmogorov strong law, which relates to the first moment, was generalized by Marcinkiewicz and Zygmund (1937) into a result relating to moments of order between 0 and 2; cf. also Gut (2007), Section 6.7:

Theorem 1.5 (The Marcinkiewicz–Zygmund strong law)**.** *Let $0 < r < 2$. Suppose that $X, X_1, X_2, \ldots$ are i.i.d. random variables. If $E|X|^r < \infty$ and $E\,X = 0$ when $1 \leq r < 2$, then*

$$\frac{S_n}{n^{1/r}} \stackrel{a.s.}{\to} 0 \quad as \quad n \to \infty \quad \Longleftrightarrow \quad E|X|^r < \infty \quad and, \ if \quad 1 \leq r < 2, \ E\,X = 0.$$

The results so far pertain to partial sums, summing from X_1 and onwards. There exist, however, analogs pertaining to *delayed sums* or *windows* or *lag sums*, that have not yet reached the same level of attention, most likely because they are more recent.

In order to describe these results we define the concept of a *window*, say. Namely for any given sequence $X_1, X_2, \ldots$ we set

$$T_{n,n+k} = \sum_{j=n+1}^{n+k} X_j, \quad n \geq 0,\, k \geq 1.$$

The analogs of the strong law large numbers and the law of the iterated logarithm are due to Chow (1973) and Lai (1974), respectively.

Theorem 1.6 (Chow's strong law for delayed sums)**.** *Let* $0 < \alpha < 1$*, suppose that* $X, X_1, X_2, \ldots$ *are i.i.d. random variables, and set* $T_{n,n+n^\alpha} = \sum_{k=n+1}^{n+n^\alpha} X_k$*,* $n \geq 1$*. Then*

$$\frac{T_{n,n+n^\alpha}}{n^\alpha} \overset{a.s.}{\to} 0 \iff E|X|^{1/\alpha} < \infty \quad \text{and} \quad E\,X = 0.$$

This result has been extended in Bingham and Goldie (1988) by replacing the window width n^α by a self-neglecting function $\phi(n)$ which includes regularly varying functions $\phi(\cdot)$ of order $\alpha \in (0,1)$.

Remark 1.7. As pointed out in Chow (1973), the strong law remains valid for $\alpha = 1$, since

$$\frac{T_{n,2n}}{n} = 2 \cdot \frac{S_{2n}}{2n} - \frac{S_n}{n} \overset{a.s.}{\to} 0 \quad \text{as} \quad n \to \infty,$$

whenever the mean is finite and equals zero.

In analogy with the LIL, where an *iterated logarithm* appears in the normalisation, the following result, due to Lai (1974), is called the *law of the single logarithm* (LSL).

Theorem 1.8 (Lai's law of the single logarithm — LSL)**.** *Let* $0 < \alpha < 1$*. Suppose that* $X, X_1, X_2, \ldots$ *are i.i.d. random variables with mean* 0 *and variance* σ^2*, and set* $T_{n,n+n^\alpha} = \sum_{k=n+1}^{n+n^\alpha} X_k$*,* $n \geq 1$*. If*

$$E\,|X|^{2/\alpha}\big(\log^+|X|\big)^{-1/\alpha} < \infty,$$

then,

$$\limsup_{n\to\infty}\,(\liminf_{n\to\infty})\frac{T_{n,n+n^\alpha}}{\sqrt{2n^\alpha \log n}} = \sigma\sqrt{1-\alpha} \quad (-\sigma\sqrt{1-\alpha}) \quad a.s.$$

Conversely, if

$$P\Big(\limsup_{n\to\infty} \frac{|T_{n,n+n^\alpha}|}{\sqrt{n^\alpha \log n}} < \infty\Big) > 0,$$

then

$$E\,|X|^{2/\alpha}\big(\log^+|X|\big)^{-1/\alpha} < \infty \quad \text{and} \quad E\,X = 0.$$

We remark, in passing, that results of this kind may be useful for the evaluation of weighted sums of i.i.d. random variables for certain classes of weights, for example in connection with certain summability methods; see e.g., Bingham (1984), Bingham and Goldie (1983), Bingham and Maejima (1985), Chow (1973).

The aim of this paper is, in the first half, to present a survey of random field analogs, although with main focus on the LSL. We shall therefore content ourselves by simply providing appropriate references for the law of large numbers and the law of the iterated logarithm. However, our first result is an LSL for sequences, where the windows expand in an "almost linear fashion", viz., the length of the nth window equals, for example, $n/\log n$ or $n/\log\log n$. A skeleton of the proof will be given in Subsection 2.1, and a sketch in Subsection 2.2.

In the second part we first present some extensions of the LSL to random fields, that is, we consider a collection of i.i.d. random variables indexed by $\mathbf{Z}_+^d$, the positive integer d-dimensional lattice, and prove analogous results in that setting. Main emphasis is on the case when the expansion rates in the components are different.

Finally we turn to convergence rates in the law of large numbers. Departing from the legendary Baum–Katz (1965) theorem, more precisely, the Hsu–Robbins–Erdős–Spitzer–Baum–Katz theorem, relating the finiteness of sums such as $\sum_{n=1}^{\infty} n^{\text{power}} P(|S_n| > n^{\text{power}}\varepsilon)$ to moment conditions, we review a number of results in the multiindex setting. Once again, the non-equal expansion rates are the main point. Some results when the power weights are replaced by almost exponential weights are also presented.

A final section contains some remarks on martingale proofs of the law of large numbers and their relation to the classical proofs.

We close this introduction with some pieces of notation and conventions:

- For all results concerning the limsup of a sequence there exist "obvious" analogs for the liminf.
- In the following we shall, at times, for mutual convenience, abuse the notation "iff" to be interpreted as in, for example, Theorems 1.3 and 1.8 in LIL- and LSL-type results.
- C with or without indices denote(s) numerical constants of no importance that may differ between appearances.
- Any random variable without index denotes a generic random variable with respect to the sequence or field of i.i.d. random variables under investigation.
- $\log^+ x = \max\{\log x, 1\}$ for $x > 0$. We shall, however, occasionally be sloppy about the additional $+$-sign within computations.
- For simplicity, we shall permit ourselves, when convenient, to treat quantities such as n^α or $n/\log n$, and so on, as integers.
- Empty products, such as $\prod_{i=1}^{0} = 1$.

2. Between the LIL and LSL

There exist two boundary cases with respect to Theorem 1.8; the cases $\alpha = 0$ and $\alpha = 1$.

The case $\alpha = 0$ contains the trivial one; when the window reduces to a single random variable. More interesting are the windows $T_{n,n+\log n}$, $n \geq 1$, for which the so-called Erdős–Rényi law (cf. Erdős and Rényi (1970), Theorem 2, Csörgő and Révész (1981), Theorem 2.4.3) tells us that if $E\,X = 0$, and the moment generating function $\psi_X(t) = E\,\exp\{tX\}$ exists in a neighborhood of 0, then, for any $c > 0$,

$$\lim_{n\to\infty} \max_{0\leq k\leq n-k} \frac{T_{k,k+c\log k}}{c\log k} = \rho(c) \quad \text{a.s.},$$

where

$$\rho(c) = \sup\{x : \inf_t e^{-tx}\psi_X(t) \geq e^{-1/c}\},$$

where, in particular, we observe that the limit depends on the actual distribution of the summands.

For a generalization to more general window widths a_n, such that $a_n/\log n \to \infty$ as $n \to \infty$, but still assuming that the moment generating function exists, we refer, e.g., to Csörgő and Révész (1981), Theorem 3.1.1. Results where the moment condition is somewhat weaker than existence of a moment generating function were discussed in Lanzinger and Stadtmüller (2000).

For the boundary case at the other end, viz., $\alpha = 1$, one has $a_n = n$ and $T_{n,2n} \overset{d}{=} S_n$ and the correct norming is as in the LIL.

An interesting remaining case is when the window size is larger than any power less than one, and at the same time not quite linear. In order to present that one we need the concept of slow variation.

Definition 2.1. Let $a > 0$. A positive measurable function L on $[a, \infty)$ *varies slowly* at infinity, denoted $L \in \mathcal{SV}$, iff

$$\frac{L(tx)}{L(t)} \to 1 \quad \text{as} \quad t \to \infty \quad \text{for all} \quad x > 0.$$

The typical example one should have in mind is $L(x) = \log x$ (or possibly $L(x) = \log\log x$). Every positive function with a finite limit as $x \to \infty$ is slowly varying. An excellent source is Bingham, Goldie and Teugels (1987). Some basic facts can be found in Gut (2007), Section A.7.

With this definition in mind, our windows thus are of the form

$$T_{n,n+n/L(n)}, \tag{2.1}$$

where

$$L \in \mathcal{SV},\ L(\cdot) \nearrow \infty,\ L \text{ is differentiable, and } \frac{xL'(x)}{L(x)} \searrow \quad \text{as} \quad x \to \infty. \tag{2.2}$$

Here is now the corresponding LSL from Gut et al. (2010).

Theorem 2.2. *Suppose that X_1, X_2, … are i.i.d. random variables with mean* 0 *and finite variance σ^2. Set, for $n \geq 2$,*

$$d_n = \log \frac{n}{a_n} + \log\log n = \log L(n) + \log\log n,$$

and

$$f(n) = \min\{a_n \cdot d_n,\, n\},$$

where $f(\cdot)$ is an increasing interpolating function, i.e., $f(x) = f_{[x]}$ for $x > 0$. Then, with $f^{-1}(\cdot)$ being the corresponding (suitably defined) inverse function,

$$\limsup_{n\to\infty} \frac{T_{n,n+a_n}}{\sqrt{2a_n d_n}} = \sigma \quad \textit{a.s.} \quad \Longleftrightarrow \quad E\big(f^{-1}(X^2)\big) < \infty.$$

Remark 2.3. The "natural" necessary moment assumption is the given one with $f(n) = a_n d_n$. However, for very slowly increasing functions, such as $L(x) = \log\log\log\log x$, we have $f(n) = n$, that is the moment condition is equivalent to finite variance in such cases.

In order to get a flavor of the result, we begin by providing some examples. In the following two subsections we shall encounter a skeleton of the proof as well as a sketch of the same.

First, the two "obvious ones".

Example 2.4. If for some $p > 0$

$$E\,X^2 \frac{(\log^+ |X|)^p}{\log^+ \log^+ |X|} < \infty,$$

then

$$\limsup_{n\to\infty} \frac{T_{n,n+n/(\log n)^p}}{\sqrt{2(p+1)\frac{n}{(\log n)^p} \log\log n}} = \sigma \quad \text{a.s.}$$

Example 2.5. If $\sigma^2 = \operatorname{Var} X < \infty$, then

$$\limsup_{n\to\infty} \frac{T_{n,n+n/\log\log n}}{\sqrt{2n}} = \sigma \quad \text{a.s.}$$

And here are two more elaborate ones.

Example 2.6. Let, for $n \geq 9$, $a_n = n(\log\log n)^q/(\log n)^p$, p, $q > 0$. Then

$$d_n = \log\Big(\frac{n(\log\log n)^q}{n/(\log n)^p}\Big) + \log\log n \sim (p+1)\log\log n \quad \text{as} \quad n \to \infty,$$

so that, $f(n) = (p+1)n(\log\log n)^{q+1}/(\log n)^p$, and, hence,

$$f^{-1}(n) \sim Cn(\log n)^p/(\log\log n)^{q+1}$$

as $n \to \infty$, and the following result emerges.

If, for some p, $q > 0$,

$$E\,X^2 \frac{(\log^+ |X|)^p}{(\log^+ \log^+ |X|)^{q+1}} < \infty,$$

then

$$\limsup_{n\to\infty} \frac{T_{n,n+n(\log\log n)^q/(\log n)^p}}{\sqrt{2(p+1)\frac{n}{(\log n)^p}(\log\log n)^{q+1}}} = \sigma \quad \text{a.s.}$$

Example 2.7. Let $a_n = n/\exp\{\sqrt{\log n}\}$, $n \geq 1$, that is,

$$d_n = \log\exp\{\sqrt{\log n}\} + \log\log n = \sqrt{\log n} + \log\log n \sim \sqrt{\log n} \quad \text{as} \quad n \to \infty,$$

which yields $f(n) \sim n\sqrt{\log n}/\exp\{\sqrt{\log n}\}$ as $n \to \infty$, so that

$$f^{-1}(n) \sim n\exp\{\sqrt{\log n + 1/2}\}/\sqrt{\log n} \quad \text{as} \quad n \to \infty,$$

which tells us that if

$$E\,X^2 \frac{\exp\{\sqrt{2\log^+ |X|}\}}{\sqrt{\log^+ |X|}} < \infty,$$

then

$$\limsup_{n\to\infty} \frac{T_{n,n+n/\exp\{\sqrt{\log n}\}}}{\sqrt{2\frac{n}{\exp\{\sqrt{\log n}\}}\sqrt{\log n}}} = \sigma \quad \text{a.s.}$$

We refer to Gut et al. (2010) for details and further examples.

The proof of Theorem 2.2 has some common ingredients with that of the LIL, in the sense that one needs two truncations. One to match the Kolmogorov exponential bounds and one to match the moment requirement. Typically (and somewhat frustratingly) it is the thin central part that causes the main trouble in the proof. A weaker result is obtained if only the first truncation is made. The cost is that too much (although not much too much) integrability will be required. A proof in this weaker setting is hinted at in Remark 2.10. For more we refer to Gut et al. (2010), Section 6.

2.1. Skeleton of the proof of Theorem 2.2

As indicated a few lines ago, one begins by truncating at two levels—b_n and c_n, where the former is chosen to match the exponential inequalities, and the latter to match the moment assumption, after which one defines the truncated summands,

$$\begin{aligned} X_n' &= X_n I\{|X_n| \leq b_n\}, \\ X_n'' &= X_n I\{b_n < |X_n| < c_n\}, \\ X_n''' &= X_n I\{|X_n| \geq c_n\}, \end{aligned} \tag{2.3}$$

and, along with them, their expected values, partial sums, and windows: $E\,X'_n$, $E\,X''_n$, $E\,X'''_n$, S'_n, S''_n, S'''_n, and $T'_{n,n+n/L(n)}$, $T''_{n,n+n/L(n)}$, $T'''_{n,n+n/L(n)}$, respectively, where, in the following any object with a prime or a multiple prime refers to the respective truncated component.

Since truncation generally destroys centering one then shows that the truncated means are "small" and that $\mathrm{Var}(T'_{n,n+n/L(n)}) \approx n\sigma^2$.

With these quantities one now proceeds as follows:

The upper estimate:

1. Dispose of $T'''_{n_k,n_k+n_k/L_{n_k}}$;
2. Dispose of $T''_{n_k,n_k+n_k/L_{n_k}}$ (frequently the hard(est) part);
3. Upper exponential bounds for a suitable subsequence $T'_{n_k,n_k+n_k/L_{n_k}}$;
4. Borel–Cantelli 1 $\Longrightarrow$ $T'_{n_k,n_k+n_k/L_{n_k}}$ is OK;
5. $1+2+4$ $\Longrightarrow$ $\limsup T_{n_k,n_k+n_k/L_{n_k}} \leq \cdots$;
6. Filling gaps;
7. $5+6$ $\Longrightarrow$ $\limsup T_{n,n+n/L(n)} \leq \cdots$;

The lower estimate:

8. Lower exponential for a suitable subsequence $T'_{n_k,n_k+n_k/L_{n_k}}$;
9. Subsequence is sparse $\Longrightarrow$ independence;
10. Borel–Cantelli 2 $\Longrightarrow$ $T'_{n_k,n_k+n_k/L_{n_k}}$ is OK;
11. $1+2+10$ $\Longrightarrow$ $\limsup T_{n_k,n_k+n_k/L_{n_k}} \geq \cdots$;
12. $\limsup T_{n,n+n/L(n)} \geq \limsup T_{n_k,n_k+n_k/L_{n_k}} \geq \cdots$;
13. $7+12$ $\Longrightarrow$ $\limsup T_{n,n+n/L(n)} = \cdots$;
14. $\square$

Remark 2.8. This is the procedure in Gut et al. (2010). However, for some results one can even dispose of $T'''_{n,n+n/L(n)}$ and $T''_{n,n+n/L(n)}$ in Steps 1 and 2, respectively.

When it comes to choosing the appropriate subsequence it turns out that the choice should satisfy the relation

$$d_{n_k} \sim \log k \quad \text{as} \quad k \to \infty, \tag{2.4}$$

and for this to happen, the following lemma, which is due to Fredrik Jonsson, Uppsala, is crucial.

Lemma 2.9. *Suppose that $L \in \mathcal{SV}$ satisfies (2.2). Then*

$$\frac{\log(L(t)\,\log t)}{\log \varphi(t)} \to 1 \quad \textit{as} \quad t \to \infty.$$

Before presenting the proof we note that the lemma is more or less trivially true for slowly varying functions made up by logarithms or iterated ones.

Proof. Setting $\varphi^*(t) = L(t)\log t$ we have $\varphi(t) \le \varphi^*(t)$ since $L(\cdot) \nearrow$. For the opposite inequality an appeal to (2.2) shows that

$$\begin{aligned}
\varphi^*(t) &= \int_1^t \Big(L'(u)\log u + \frac{L(u)}{u}\Big)du = \int_1^t \frac{L'(u)\,u\,L(u)}{L(u)\,u}\Big(\int_1^u \frac{1}{v}\,dv\Big)du + \varphi(t) \\
&\le \int_1^t \frac{L(u)}{u}\Big(\int_1^u \frac{L'(v)}{L(v)}dv\Big)du + \varphi(t) \le \varphi(t)(1+\log(L(t))),
\end{aligned}$$

from which we conclude that

$$1 \ge \frac{\log\varphi(t)}{\log\varphi^*(t)} \ge 1 - \frac{\log(1+\log L(t))}{\log(L(t)\log t)} \to 1 \quad \text{as} \quad t \to \infty. \qquad \square$$

2.2. Sketch of the proof of Theorem 2.2

We introduce the parameters $\delta > 0$ and $\varepsilon > 0$ and truncate at

$$b_n = \frac{\sigma\delta}{\varepsilon}\sqrt{\frac{a_n}{d_n}} \quad \text{and} \quad c_n = \delta\sqrt{f(n)},$$

recalling that

$$a_n = n/L(n), \quad d_n = \log L(n) + \log\log n, \quad f(n) = \min\{a_n d_n, n\},$$

and set, in accordance with (2.3),

$$\begin{aligned}
X_n' &= X_n I\{|X_n| \le b_n\}, \\
X_n'' &= X_n I\{b_n < |X_n| < \delta\sqrt{f(n)}\}, \\
X_n''' &= X_n I\{|X_n| \ge \delta\sqrt{f(n)}\},
\end{aligned}$$

after which we check the appropriate smallness of the truncated means.

Next we choose a subsequence such that $d_{n_k} \sim \log k$.

In order to dispose of $T'''_{n_k,n_k+a_{n_k}}$ we observe that if $|T'''_{n_k,n_k+a_{n_k}}|$ surpasses the $\eta\sqrt{a_{n_k}d_{n_k}}$ then, necessarily, at least one of the corresponding X''':s is nonzero, which leads to

$$\sum_{k=1}^{\infty} P(|T'''_{n_k,n_k+a_{n_k}}| > \eta\sqrt{a_{n_k}d_{n_k}}) \le \sum_{k=1}^{\infty} a_{n_k} P(|X| > \frac{\eta}{2}\sqrt{f(n_k)}) < \infty, \qquad (2.5)$$

where the finiteness is a consequence of the moment assumption.

As for the second step, this is a technically pretty involved matter for which we refer to Gut et al. (2010).

For the analysis of $T'_{n_k,n_k+a_{n_k}}$ we use the Kolmogorov upper exponential bounds (see e.g., Gut (2007), Lemma 8.2.1) and obtain (after having taken care of the centering inflicted by the truncation),

$$\begin{aligned} P(|T'_{n,n+a_n}| > \varepsilon\sqrt{2a_nd_n}) &\le P(|T'_{n,n+a_n} - ET'_{n,n+a_n}| > \varepsilon(1-\delta)\sqrt{2a_nd_n}) \\ &\le 2\exp\Big\{-\frac{\varepsilon^2(1-\delta)^3}{\sigma^2}\cdot d_n\Big\} \quad \text{for } n \text{ large,} \end{aligned}$$

which, together with the previous estimates, shows that

$$\sum_{k=1}^{\infty} P(|T_{n_k,n_k+a_{n_k}}| > (\varepsilon+2\eta)\sqrt{2a_{n_k}d_{n_k}}) < \infty,$$

provided $\varepsilon > \sigma/(1-\delta)^{3/2}$, and thus, due to the arbitrariness of η and δ, and the first Borel–Cantelli lemma, that

$$\limsup_{k\to\infty} \frac{T_{n_k,a_{n_k}}}{\sqrt{2a_{n_k}d_{n_k}}} \le \sigma \quad \text{a.s.} \tag{2.6}$$

The next step (Step 6 in the above list) amounts to proving the same for the entire sequence, and this is achieved by showing that

$$\sum_k P\Big(\max_{n_k\le n\le n_{k+1}} \frac{S_{n+a_n}-S_n}{\sqrt{2a_n\,d_n}} > \sigma\Big) < \infty, \tag{2.7}$$

implying that

$$P(\max_{n_k\le n\le n_{k+1}} \frac{S_{n+a_n}-S_n}{\sqrt{2a_n\,d_n}} > \sigma \text{ i.o}) = 0,$$

which, together with (2.6), then will tell us that

$$\limsup_{n\to\infty} \frac{T_{n,n+a_n}}{\sqrt{2a_nd_n}} \le \sigma \quad \text{a.s.}$$

In order to prove (2.7) we first observe that, for any $\eta > 0$,

$$\begin{aligned} &P\Big(\max_{n_k\le n\le n_{k+1}} \frac{S_{n+a_n}-S_n}{\sqrt{2a_n\,d_n}} > (1+6\,\eta)\sigma\Big) \\ &\le P(\max_{n_k\le n\le n_{k+1}} (S_{n+a_n} - S_{n_k+a_{n_k}}) > 2\eta\,\sigma\sqrt{2a_{n_k}\,d_{n_k}}) \\ &\quad + P(\max_{n_k\le n\le n_{k+1}} (-S_n + S_{n_k}) > 2\eta\,\sigma\sqrt{2a_{n_k}\,d_{n_k}}) \\ &\quad + P(\max_{n_k\le n\le n_{k+1}} (S_{n_k+a_{n_k}} - S_{n_k}) > (1+2\eta)\,\sigma\sqrt{2a_{n_k}d_{n_k}}), \end{aligned}$$

after which (2.7), broadly speaking, follows by applying the Lévy inequality (cf. e.g. Gut (2007), Theorem 3.7.2) to each of the four terms.

This finishes the "proof" of the upper estimate, and it remains to take care of the lower one (Step 8 and onwards in the skeleton list).

After having checked that

$$\operatorname{Var} X'_k \geq \sigma^2 - 2E\,X^2 I\{|X_k| \geq b_k\} \geq \sigma^2(1-\delta),$$

for n large, so that

$$\operatorname{Var}(T'_{n,n+a_n}) \geq a_n \sigma^2 (1-\delta) \quad \text{for } n \text{ large},$$

we obtain, exploiting the lower exponential bound (see e.g. Gut (2007), Lemma 8.2.2), that, for any $\gamma > 0$,

$$\begin{aligned} &P(T'_{n,n+a_n} > \varepsilon\sqrt{2a_n d_n}) \\ &\geq P\big(T'_{n,n+a_n} - ET'_{n,n+a_n} > \frac{\varepsilon(1+\delta)}{\sigma\sqrt{(1-\delta)}}\sqrt{2\operatorname{Var}(T'_{n,n+a_n})d_n}\big) \\ &\geq \exp\big\{-\frac{\varepsilon^2(1+\delta)^2(1+\gamma)}{\sigma^2(1-\delta)}\cdot d_n\big\} \quad \text{for } n \text{ large.} \end{aligned}$$

Applying this lower bound to our subsequence and combining the outcome with (2.5) and the omitted analog for $T''_{n,n+n/L(n)}$ then yields

$$\limsup_{k\to\infty} \frac{T_{n_k,n_k+a_{n_k}}}{\sqrt{2a_{n_k}d_{n_k}}} \geq \sigma \quad \text{a.s.} \tag{2.8}$$

Finally, since the limsup for the entire sequence certainly is at least as large as that of the subsequence (Step 12 in the skeleton), we conclude that the lower bound (2.8) also holds for the entire sequence.

This completes (the sketch of) the proof (Step 14).

Remark 2.10. We close this section by recalling that a slightly weaker result may be obtained by truncation at $b_n = \sqrt{a_n/d_n}$ only, in which case $T''_{n,n+n/L(n)}$ and $T'''_{n,n+n/L(n)}$ are joined into one "outer" contribution. With the same argument as above, the previous computation then is replaced by

$$\sum_{n=1}^{\infty} P(|X| > \frac{\sigma\delta}{\varepsilon} b_n) < \infty,$$

where finiteness holds iff

$$E\,b^{-1}(|X|) < \infty.$$

If, for example, $L(n) = \log n$, then the moment condition $E\,X^2 \frac{\log^+|X|}{\log^+\log^+|X|} < \infty$ in Theorem 2.2 is replaced by the condition $E\,X^2 \log^+|X| \log^+\log^+|X| < \infty$; cf. Gut et al. (2010), Section 6.

3. The LLN and the LIL for random fields

We now turn our attention to random fields. But first, in order to formulate our results, we need to define the setup. Toward that end, let $\mathbf{Z}_+^d$, $d \geq 2$, denote the positive integer d-dimensional lattice with coordinate-wise partial ordering $\leq$, viz., for $\mathbf{m} = (m_1, m_2, \ldots, m_d)$ and $\mathbf{n} = (n_1, n_2, \ldots, n_d)$, $\mathbf{m} \leq \mathbf{n}$ means that $m_k \leq n_k$, for $k = 1, 2, \ldots, d$. The "size" of a point equals $|\mathbf{n}| = \prod_{k=1}^d n_k$, and $\mathbf{n} \to \infty$ means that $n_k \to \infty$, for all $k = 1, 2, \ldots, d$.

Next, let $\{X_{\mathbf{k}}, \, \mathbf{k} \in \mathbf{Z}_+^d\}$ be i.i.d. random variables with partial sums $S_{\mathbf{n}} = \sum_{\mathbf{k} \leq \mathbf{n}} X_{\mathbf{k}}$, $\mathbf{n} \in \mathbf{Z}_+^d$.

For random fields with i.i.d. random variables $\{X_{\mathbf{k}}, \, \mathbf{k} \in \mathbf{Z}_+^d\}$ the analog of Kolmogorov's strong law (see Smythe (1973)) reads as follows:

$$\frac{S_{\mathbf{n}}}{|\mathbf{n}|} = \frac{1}{|\mathbf{n}|} \sum_{\mathbf{k} \leq \mathbf{n}} X_{\mathbf{k}} \overset{a.s.}{\to} 0 \quad \Longleftrightarrow \quad E|X| \, (\log^+ |X|)^{d-1} < \infty \quad \text{and} \quad E\,X = 0. \tag{3.1}$$

For more general index sets, see Smythe (1974).

The analogous Marcinkiewicz–Zygmund law of large numbers was proved in Gut (1978):

$$\frac{1}{|\mathbf{n}|^{1/r}} S_{\mathbf{n}} \overset{a.s.}{\to} 0 \quad \Longleftrightarrow \quad E|X|^r (\log^+ |X|)^{d-1} < \infty \quad \text{and, if} \quad 1 \leq r < 2, \; E\,X = 0. \tag{3.2}$$

The Hartman–Wintner analog is due to Wichura (1973):

$$\begin{gathered} \limsup_{\mathbf{n} \to \infty} \frac{S_{\mathbf{n}}}{\sqrt{2|\mathbf{n}| \log\log |\mathbf{n}|}} = \sigma\sqrt{d} \quad \text{a.s.} \\ \Longleftrightarrow \\ E\,X^2 \frac{(\log^+ |X|)^{d-1}}{\log^+ \log^+ |X|} < \infty \quad \text{and} \quad E\,X = 0, \; E\,X^2 = \sigma^2. \end{gathered} \tag{3.3}$$

A variation on the theme concerns the same problems when one considers the index set $\mathbf{Z}_+^d$ restricted to a *sector*, which, for the case $d = 2$, equals

$$S_\theta^{(2)} = \{(x, y) \in \mathbb{Z}_+^2 : \theta x \leq y \leq \theta^{-1} x, \; 0 < \theta < 1\}. \tag{3.4}$$

In the limiting case $\theta = 1$, the sector degenerates into a diagonal ray, in which case the sums $S_{\mathbf{n}}$, $\mathbf{n} \in S_\theta^{(2)}$, are equivalent to the subsequence S_{n^2}, more generally, S_{n^d}, $n \geq 1$, of the sequence $\{S_n, \, n \geq 1\}$ when $d = 1$. In that case it is clear that the usual one-dimensional assumptions are sufficient for the LLN and the LIL. One may therefore wonder about the proper conditions for the sector—since extra logarithms are needed "at the other end" (as $\theta \to 0$).

Without going into any details we just mention that it has been shown in Gut (1983) that the law of large numbers as well as the law iterated logarithm hold under the same moment conditions as in the case $d = 1$, and that the limit points in the latter case are the same as in the Hartman–Wintner theorem (Theorem 1.3).

For some additional comments on this we refer to Section 10 toward the end of the paper.

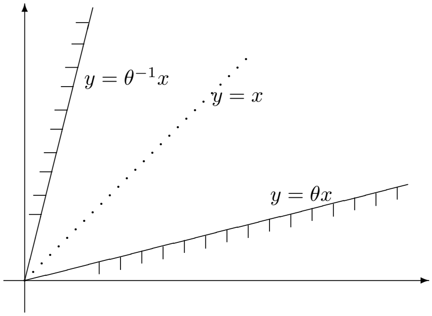


Figure 1: A sector ($d = 2$)

4. The LLN and LSL for windows

Having defined the general setup we also need the extension of the concept delayed sums or windows to this setting. A window here is an object $T_{\mathbf{n},\mathbf{n}+\mathbf{k}}$. For $d = 2$ we this is an incremental rectangle

$$T_{\mathbf{n},\mathbf{n}+\mathbf{k}} = S_{n_1+k_1,n_2+k_2} - S_{n_1+k_1,n_2} - S_{n_1,n_2+k_2} + S_{n_1,n_2} :$$

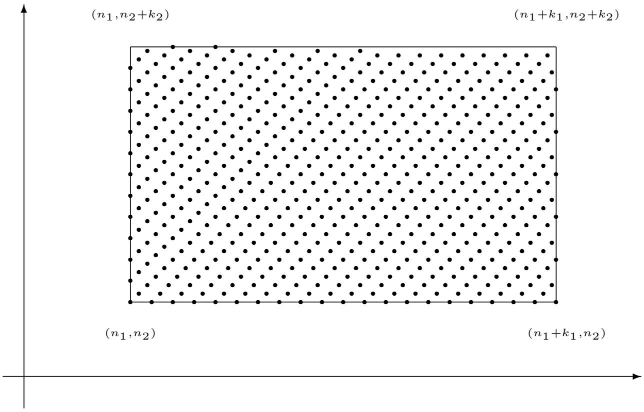


Figure 2: A typical window ($d = 2$)

In higher dimensions it is the analogous d-dimensional cube. A strong law for this setting can be found in Thalmaier (2009), Stadtmüller and Thalmaier (2009).

The extension of Theorem 1.8 to random fields runs as follows.

Theorem 4.1. *Let* $0 < \alpha < 1$, *and suppose that* $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d\}$ *are i.i.d. random variables with mean* 0 *and finite variance* σ^2. *If*

$$E\,X^{2/\alpha}(\log^+|X|)^{d-1-1/\alpha} < \infty,$$

then

$$\limsup_{\mathbf{n}\to\infty} \frac{T_{\mathbf{n},\mathbf{n}+\mathbf{n}^\alpha}}{\sqrt{2|\mathbf{n}|^\alpha \log|\mathbf{n}|}} = \sigma\sqrt{1-\alpha} \quad a.s.$$

Conversely, if

$$P\Big(\limsup_{\mathbf{n}\to\infty} \frac{|T_{\mathbf{n},\mathbf{n}+\mathbf{n}^\alpha}|}{\sqrt{|\mathbf{n}|^\alpha \log|\mathbf{n}|}} < \infty\Big) > 0,$$

then $E\,X^{2/\alpha}(\log^+|X|)^{d-1-1/\alpha} < \infty$ *and* $E\,X = 0$.

Some remarks on the proof will be given in Section 6.

4.1. An LSL for subsequences

The proof of the theorem is in the LIL-style, which, i.a., means that one begins by proving the sufficiency as well as the necessity along a suitable subsequence. Sticking to this fact one can, with very minor modifications of the proof of Theorem 4.1, prove the following *LSL for subsequences.* The inspiration for this result comes from the LIL-analog in Gut (1986).

Theorem 4.2. *Let* $0 < \alpha < 1$, *suppose that* $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d\}$ *are i.i.d. random variables with mean* 0 *and finite variance* σ^2, *and set*

$$\Lambda = \{\mathbf{n} \in \mathbf{Z}_+^d : \mathbf{n}_i = i^{\beta/(1-\alpha)},\ i \geq 1\}.$$

If

$$E\,X^{2/\alpha}(\log^+|X|)^{d-1-1/\alpha} < \infty,$$

then, for $\beta > 1$,

$$\limsup_{\substack{\mathbf{n}\to\infty \\ \{\mathbf{n}\in\mathbf{\Lambda}^*\}}} \frac{T_{\mathbf{n},\mathbf{n}+\mathbf{n}^\alpha}}{\sqrt{2|\mathbf{n}|^\alpha \log|\mathbf{n}|}} = \sigma\sqrt{\frac{1-\alpha}{\beta}} \quad a.s.$$

Conversely, if

$$P\Big(\limsup_{\substack{\mathbf{n}\to\infty \\ \{\mathbf{n}\in\mathbf{\Lambda}^*\}}} \frac{|T_{\mathbf{n},\mathbf{n}+\mathbf{n}^\alpha}|}{\sqrt{|\mathbf{n}|^\alpha \log|\mathbf{n}|}} < \infty\Big) > 0,$$

then $E\,X^{2/\alpha}(\log^+|X|)^{d-1-1/\alpha} < \infty$ *and* $E\,X = 0$.

For further details, see Gut and Stadtmüller (2008a), Section 6.

4.2. Different α:s

During a seminar in Uppsala on the previous material Fredrik Jonsson asked the question: "What happens if the α:s are different?"

In Theorem 4.1 the windows grow at the same rate in each coordinate; the edges of the windows are equal to n_k^α for all $k = 1, 2, \dots, d$. The focus now is to allow for different growth rates in different directions; viz., the edges of the windows will be $n_k^{\alpha_k}$, $k = 1, 2, \dots, d$, where, w.l.o.g., we assume that

$$0 < \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_d < 1.$$

Next, we define $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_d)$, and set, for ease of notation,

$$\mathbf{n}^{\boldsymbol{\alpha}} = (n_1^{\alpha_1}, n_2^{\alpha_2}, \dots, n_d^{\alpha_d}), \quad \text{and} \quad |\mathbf{n}^{\boldsymbol{\alpha}}| = \prod_{k=1}^{d} n_k^{\alpha_k}.$$

Furthermore, following Stadtmüller and Thalmaier (2009), we let p be equal to the number of α:s that are equal to the smallest one.

As for the strong law, the results in Thalmaier (2009), Stadtmüller and Thalmaier (2009), in fact, also cover the case of unequal α:s. For a Marcinkiewicz–Zygmund analog we refer to Gut and Stadtmüller (2009). For completeness we also mention Gut and Stadtmüller (2010), where some results concerning Cesàro summation are proved.

Here is now the generalization of Theorem 4.1. For a proof and further details we refer to Gut and Stadtmüller (2008b).

Theorem 4.3. *Suppose that* $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d\}$ *are i.i.d. random variables with mean* 0 *and finite variance* σ^2. *If*

$$E|X|^{2/\alpha_1} (\log^+ |X|)^{p-1-1/\alpha_1} < \infty,$$

then

$$\limsup_{\mathbf{n}\to\infty} \frac{T_{\mathbf{n},\mathbf{n}+\mathbf{n}^{\boldsymbol{\alpha}}}}{\sqrt{2|\mathbf{n}^{\boldsymbol{\alpha}}| \log |\mathbf{n}|}} = \sigma\sqrt{1-\alpha_1} \quad a.s.$$

Conversely, if

$$P\Big(\limsup_{\mathbf{n}\to\infty} \frac{|T_{\mathbf{n},\mathbf{n}+\mathbf{n}^{\boldsymbol{\alpha}}}|}{\sqrt{|\mathbf{n}^{\boldsymbol{\alpha}}| \log |\mathbf{n}|}} < \infty\Big) > 0,$$

then $E|X|^{2/\alpha_1} (\log^+ |X|)^{p-1-1/\alpha_1} < \infty$ *and* $E\,X = 0$.

Remark 4.4. If $\alpha_1 = \alpha_2 = \cdots = \alpha_d = \alpha$, then $p = d$ and $|\mathbf{n}^{\boldsymbol{\alpha}}| = |\mathbf{n}|^\alpha$, and the theorem reduces to Gut and Stadtmüller (2008a), Theorem 2.1 = Theorem 4.1 above.

Remark 4.5. For a result for subsequences analogous to Theorem 4.2; see Gut and Stadtmüller (2008b), Section 6.

We observe that the moment condition as well as the extreme limit points depend on the *smallest* α and its multiplicity. Heuristically this can be explained as follows. The longer the stretch of the window along a specific axis, the more cancellation may occur in that direction. Equivalently, the shorter the stretch, the wilder the fluctuations. This means that in order to "tame" the fluctuations it is (only) necessary to put conditions on the shortest edge(s).

4.3. Different α:s, $\log$, and $\log\log$

One can exaggerate the mixtures even further, namely, by combining edges that expand at different α-rates with edges that expand with different almost linear rates. Some results in this direction concerning the LLN can be found in Gut and Stadtmüller (2011b).

The paper Gut and Stadtmüller (2011a) is devoted to the LSL. First a result from that paper that extends Gut et al. (2010) to random fields for (iterated) logarithmic expansions and mixtures of them. For simplicity and illustrative purposes we stick to the case $d = 2$.

Theorem 4.6. *Let $\{X_{i,j},\, i,j \geq 1\}$ be i.i.d. random variables.*
(i) *If*

$$E\,X^2 \frac{(\log^+ |X|)^3}{\log^+ \log^+ |X|} < \infty \quad \text{and} \quad E\,X = 0,\ E\,X^2 = \sigma^2,$$

then

$$\limsup_{m,n\to\infty} \frac{T_{(m,n)\,,\,(m+m/\log m, n+n/\log n)}}{\sqrt{4mn \frac{\log\log m + \log\log n}{\log m \log n}}} = \sigma \quad a.s.$$

Conversely, if

$$P\Big(\limsup_{n\to\infty} \frac{|T_{(m,n)\,,\,(m+m/\log m, n+n/\log n)}|}{\sqrt{mn \frac{\log\log m + \log\log n}{\log m \log n}}} < \infty\Big) > 0,$$

then $E\,X^2 \frac{(\log^+ |X|)^3}{\log^+ \log^+ |X|} < \infty$ and $E\,X = 0$.
(ii) *If*

$$E\,X^2 \log^+ |X|\, \log^+ \log^+ |X| < \infty \quad \text{and} \quad E\,X = 0,\ E\,X^2 = \sigma^2,$$

then

$$\limsup_{m,n\to\infty} \frac{T_{(m,n)\,,\,(m+m/\log\log m, n+n/\log\log n)}}{\sqrt{2mn \frac{\log\log m + \log\log n}{\log\log m \log\log n}}} = \sigma \quad a.s.$$

Conversely, if

$$P\Big(\limsup_{n\to\infty} \frac{|T_{(m,n)\,,\,(m+m/\log\log m, n+n/\log\log n)}|}{\sqrt{mn \frac{\log\log m + \log\log n}{\log\log m \log\log n}}} < \infty\Big) > 0,$$

then $E\,X^2 \log^+|X| \log^+\log^+|X| < \infty$ *and* $E\,X = 0$.

(iii) *If*

$$E\,X^2(\log^+|X|)^2 < \infty \quad \text{and} \quad E\,X = 0,\ E\,X^2 = \sigma^2,$$

then

$$\limsup_{m,n\to\infty} \frac{T_{(m,n)\,,\,(m+m/\log m, n+n/\log\log n)}}{\sqrt{4mn\frac{\log\log m+\log\log n}{\log m \log\log n}}} = \sigma \quad a.s.$$

Conversely, if

$$P\Big(\limsup_{n\to\infty} \frac{|T_{(m,n)\,,\,(m+m/\log m, n+n/\log\log n)}|}{\sqrt{4mn\frac{\log\log m+\log\log n}{\log m \log\log n}}} < \infty\Big) > 0,$$

then $E\,X^2(\log^+|X|)^2 < \infty$ *and* $E\,X = 0$.

We conclude with an example where a logarithmic expansion is mixed with a power.

Theorem 4.7. *Let* $0 < \alpha < 1$, *and let* $\{X_{i,j},\ i,j \geq 1\}$ *be i.i.d. random variables. If*

$$E\,X^{2/\alpha}(\log^+|X|)^{-1/\alpha} < \infty \quad \text{and} \quad E\,X = 0,\ E\,X^2 = \sigma^2,$$

then

$$\limsup_{m,n\to\infty} \frac{T_{(m,n)\,,\,(m+m^\alpha,\, n+n/\log n)}}{\sqrt{2m^\alpha n\frac{(1-\alpha)\log(mn)}{\log n}}} = \sigma \quad a.s.$$

Conversely, if

$$P\Big(\limsup_{m,n\to\infty} \frac{|T_{(m,n)\,,\,(m+m^\alpha,\, n+n/\log n)}|}{\sqrt{m^\alpha n\frac{\log(mn)}{\log n}}} < \infty\Big) > 0,$$

then $E\,X^{2/\alpha}(\log^+|X|)^{-1/\alpha} < \infty$ *and* $E\,X = 0$.

5. Preliminaries

Proposition 5.1. *Let* $r > 0$ *and let* X *be a non-negative random variable. Then*

$$E\,X^r < \infty \quad \Longleftrightarrow \quad \sum_{n=1}^{\infty} n^{r-1}P(X \geq n) < \infty,$$

More precisely,

$$\sum_{n=1}^{\infty} n^{r-1}P(X \geq n) \leq E\,X^r \leq 1 + \sum_{n=1}^{\infty} n^{r-1}P(X \geq n).$$

As an example, consider the case $r = 1$, and suppose that $X_1, X_2, \ldots$ is an i.i.d. sequence. It then follows from the proposition that, for any $\varepsilon > 0$,

$$P(|X_n| > n\varepsilon \text{ i.o.}) = 0 \iff \sum_{n=1}^{\infty} P(|X_n| > n\varepsilon) < \infty \iff E\,|X| < \infty.$$

Suppose instead that we are facing an i.i.d. random field $\{X_{\mathbf{n}}, \mathbf{n} \in \mathbf{Z}_+^d\}$. What is then the relevant moment condition that ensures that

$$\sum_{n} P(|X_{\mathbf{n}}| > |\mathbf{n}|) < \infty\,? \quad \text{or, equivalently, that} \quad \sum_{\mathbf{n}} P(|X| > |\mathbf{n}|) < \infty\,? \tag{5.1}$$

In order to answer this question it turns out that we need the quantities

$$d(j) = \text{Card}\,\{\mathbf{k} : |\mathbf{k}| = j\} \quad \text{and} \quad M(j) = \text{Card}\,\{\mathbf{k} : |\mathbf{k}| \le j\},$$

which describe the "size" of the index set, and their asymptotics

$$\frac{M(j)}{j(\log j)^{d-1}} \to \frac{1}{(d-1)!} \quad \text{and} \quad d(j) = o(j^{\delta}) \text{ for any } \delta > 0 \quad \text{as} \quad j \to \infty; \tag{5.2}$$

cf. Hardy and Wright (1954), Chapter XVIII and Titchmarsh (1951), relation (12.1.1) (for the case $d = 2$). The quantity $d(j)$ itself has no pleasant asymptotics; $\liminf_{j\to\infty} d(j) = d$, and $\limsup_{j\to\infty} d(j) = +\infty$.

Now, exploiting the fact that all terms in expressions such as the second sum in (5.1) with equisized indices are equal, we conclude that

$$\sum_{\mathbf{n}} P(|X| > |\mathbf{n}|) = \sum_{j=1}^{\infty} \sum_{|\mathbf{n}|=j} d(j) P(|X| > j), \tag{5.3}$$

which, via partial summation yields the first half of following lemma. The second half follows via a change of variable.

Lemma 5.2. *Let $r > 0$, and suppose that X is a random variables. Then*

$$\sum_{\mathbf{n}} P(|X| > |\mathbf{n}|) < \infty \iff E\,M(|X|) < \infty \iff E|X|(\log^+ |X|)^{d-1} < \infty,$$

$$\sum_{\mathbf{n}} |\mathbf{n}|^{r-1} P(|X| > |\mathbf{n}|) < \infty \iff E\,M(|X|^r) < \infty \iff E|X|^r(\log^+ |X|)^{d-1} < \infty.$$

Reviewing the steps leading to the lemma one finds that if, instead, we consider the sector (recall (3.4)) one finds that

$$\sum_{\mathbf{n}\in S_\theta^d} P(|X| > |\mathbf{n}|) < \infty \iff E\,M(|X|) < \infty \iff E|X| < \infty. \tag{5.4}$$

Remark 5.3. Note that the first equivalence is the same as in Lemma 5.2, and that the second one is a consequence of the "size" of the index set.

For results such as Theorem 4.3, as well as for some of the results in Section 8 below, we shall need the more general index sets

$$M_{\boldsymbol{\alpha}}(j) = \mathrm{Card}\,\{\mathbf{k} : |\mathbf{k}^{\boldsymbol{\alpha}}| \le j^{\alpha_1}\} = \mathrm{Card}\,\{\mathbf{k} : \prod_{\nu=1}^{d} k_\nu^{\alpha_\nu/\alpha_1} \le j\}. \tag{5.5}$$

Generalizing Lemma 3 in Stadtmüller and Thalmaier (2009) in a straight forward manner yields the following analog of (5.2):

$$M_{\boldsymbol{\alpha}}(j) \sim c_{\boldsymbol{\alpha}}\, j\,(\log j)^{p-1} \quad \text{as} \quad j \to \infty \tag{5.6}$$

where $c_{\boldsymbol{\alpha}} > 0$, which, in turn, via partial summation, tells us that

$$\sum_{\mathbf{n}} P(|X| > |\mathbf{n}^{\boldsymbol{\alpha}}|) \asymp \sum_{j=1}^{\infty} (\log j)^{p-1} P(|X| > j^{\alpha_1}).$$

Using a slight modification of this, together with the fact that the inverse of the function $y = x^\alpha (\log x)^\kappa$ behaves asymptotically like $x = y^{1/\alpha}(\log y)^{-(\kappa/\alpha)}$, yields the next tool (Gut and Stadtmüller (2008a), Lemma 3.2, Gut and Stadtmüller (2008b), Lemma 3.1).

Lemma 5.4. *Let $\kappa \in \mathbb{R}$ and suppose that X is a random variable. Then,*

$$\sum_{\mathbf{n}} P\big(|X| > |\mathbf{n}^{\boldsymbol{\alpha}}|(\log|\mathbf{n}|)^\kappa\big) < \infty \quad \Longleftrightarrow \quad E|X|^{1/\alpha_1}(\log^+|X|)^{p-1-\kappa/\alpha_1} < \infty.$$

In particular, if $\alpha_1 = \alpha_2 = \cdots = \alpha_d = \kappa = 1/2$, then

$$\sum_{\mathbf{n}} P\big(|X| > \sqrt{|\mathbf{n}|\log|\mathbf{n}|}\big) < \infty \quad \Longleftrightarrow \quad E\,X^2(\log^+|X|)^{d-2} < \infty.$$

For illustrative reasons we also quote Gut and Stadtmüller (2008a), Lemma 3.3, as an example of the kind of technical aid that is required at times.

Lemma 5.5. *Let $\kappa \ge 1$, $\theta > 0$, and $\eta \in \mathbb{R}$.*

$$\sum_{i=2}^{\infty} \sum_{\{\mathbf{n}:|\mathbf{n}|=i^\kappa(\log i)^\eta\}} \frac{1}{|\mathbf{n}|^\theta} = \sum_{i=2}^{\infty} \frac{d(i^\kappa(\log i)^\eta)}{i^{\kappa\theta}(\log i)^{\eta\theta}} \begin{cases} < \infty, & \textit{when} \quad \theta > \frac{1}{\kappa}, \\ = \infty, & \textit{when} \quad \theta < \frac{1}{\kappa}. \end{cases}$$

6. Sketch of the proofs of Theorems 4.1 and 4.3

In this section we give som hints on the proofs of Theorems 4.1 and 4.3, in the sense that we shall point to differences and modifications compared to the proof of Theorem 2.2 in Section 2.2.

6.1. On the proof of Theorem 4.1

This time truncation is at

$$b_{\mathbf{n}} = b_{|\mathbf{n}|} = \frac{\sigma\delta}{\varepsilon}\frac{\sqrt{|\mathbf{n}|^{\alpha}}}{\log|\mathbf{n}|} \quad \text{and} \quad c_{\mathbf{n}} = \delta\sqrt{|\mathbf{n}|^{\alpha}\log|\mathbf{n}|},$$

for some (arbitrarily) small $\delta > 0$.

The first step differs slightly from the analog in the proof of Theorem 2.2, in that we now start by dispensing of the full double- and triple primed sequences (recall Remark 2.8).

As for the double primed contribution we argue that in order for the $|T'''_{\mathbf{n},\mathbf{n}+\mathbf{n}^{\alpha}}|$:s to surpass the level $\eta\sqrt{|\mathbf{n}^{\boldsymbol{\alpha}}|\log|\mathbf{n}|}$ infinitely often, for some $\eta > 0$ small, it is necessary that infinitely many of the X''':s are nonzero, and the latter event has probability zero by the first Borel–Cantelli lemma, since

$$\sum_{\mathbf{n}} P(|X_{\mathbf{n}}| > \eta\sqrt{|\mathbf{n}|^{\alpha}\log|\mathbf{n}|}) = \sum_{\mathbf{n}} P(|X| > \eta\sqrt{|\mathbf{n}|^{\alpha}\log|\mathbf{n}|}) < \infty,$$

where the finiteness is a consequence of the moment assumption and the second half of Lemma 5.4.

Taking care of $T''_{\mathbf{n},\mathbf{n}+\mathbf{n}^{\alpha}}$ is a bit easier this time, the argument being that in order for $|T''_{\mathbf{n},\mathbf{n}+\mathbf{n}^{\alpha}}|$ to surpass the level $\eta\sqrt{|\mathbf{n}|^{\alpha}\log|\mathbf{n}|}$ it is necessary that at least $N \geq \eta/\delta$ of the X'':s are nonzero, which, by stretching the truncation bounds to the extremes, some elementary combinatorics, and the moment assumption implies that

$$\begin{aligned}
&P(|T''_{\mathbf{n},\mathbf{n}+\mathbf{n}^{\alpha}}| > \eta\sqrt{|\mathbf{n}|^{\alpha}\log|\mathbf{n}|})\\
&\leq \binom{|\mathbf{n}|^{\alpha}}{N}\Big(P\big(b_{\mathbf{n}} < |X| \leq \delta\sqrt{(|\mathbf{n}| + |\mathbf{n}|^{\alpha})\log(|\mathbf{n}| + |\mathbf{n}|^{\alpha})}\,\big)\Big)^{N}\\
&\leq C\frac{(\log|\mathbf{n}|)^{N((3/\alpha)+1-d)}}{|\mathbf{n}|^{N(1-\alpha)}},
\end{aligned}$$

and, hence, that

$$\sum_{\mathbf{n}} P(|T''_{\mathbf{n},\mathbf{n}+\mathbf{n}/L(\mathbf{n})}| > \eta\sqrt{|\mathbf{n}|^{\alpha}\log|\mathbf{n}|}) < \infty \quad \text{for all} \quad \eta > \frac{\delta}{1-\alpha},$$

whenever $N(1-\alpha) > 1$ (and $N\delta \geq \eta$), after which another application of the first Borel–Cantelli lemma concludes that part of the proof.

As for $T'_{\mathbf{n},\mathbf{n}+\mathbf{n}^{\alpha}}$, the exponential bounds do the job as before;

$$P\big(T'_{\mathbf{n},\mathbf{n}+\mathbf{n}^{\alpha}} > \varepsilon\sqrt{2|\mathbf{n}|^{\alpha}\log|\mathbf{n}|}\big)\begin{cases}\leq \exp\Big\{-\dfrac{2\varepsilon^2(1-\delta)^2}{2\sigma^2}\log|\mathbf{n}|(1-\delta)\Big\},\\[2mm] \geq \exp\Big\{-\dfrac{2\varepsilon^2(1+\delta)^2}{2\sigma^2(1-\delta)}\log|\mathbf{n}|(1+\gamma)\Big\}.\end{cases}$$

Putting things together proves the theorem for suitably selected subsequences, and thus, in particular also the lower bound for the full field (remember Step 12 in the skeleton list).

It thus remains to verify the upper bound for the entire field.

Now, for the LIL and LSL one investigates the gaps between subsequence points with the aid of the Lévy inequalities, as we have seen in the proof of Theorem 2.2, Step 6. When $d \geq 2$, however, there are no gaps in the usual sense and one must argue somewhat differently.

Let us have a quick look at the situation when $d = 2$. First we must show that the selected subsequence (which we have not explicitly presented) is such that the subset of windows overlap, viz., that they cover all of $\mathbf{Z}_+^2$. Next, we select an arbitrary window

$$T_{((m,n),(m+m^\alpha,n+n^\alpha))}$$

and note that it is always contained in the union of (at most) four of the earlier selected ones:

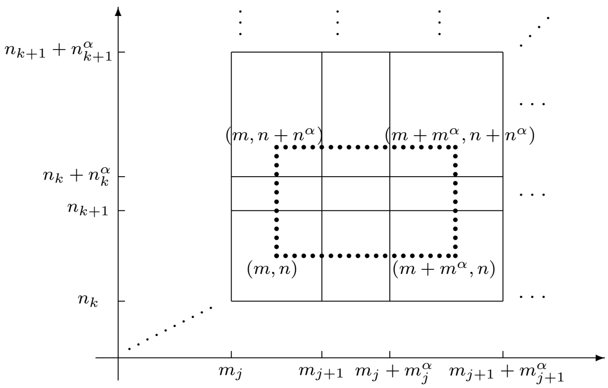


Figure 3: A dotted arbitrary window

One, finally, shows that the discrepancy between the arbitrary window and the selected ones is asymptotically negligible. This is a technical matter which we omit. Except for mentioning that one has to distinguish between the cases when the arbitrary window is located in "the center" of the index set or "close" to one of the coordinate axes (for a similar discussion cf. also Gut (1980), Section 4).

6.2. On the proof of Theorem 4.3

This proof runs along the same lines as the previous one with some additional technical complications, due to the non-equalness of the α:s. In order to illustrate this, consider the triple-primed windows.

Truncation now is at

$$b_{\mathbf{n}} = b_{|\mathbf{n}|} = \frac{\sigma\delta}{\varepsilon} \frac{\sqrt{|\mathbf{n}^{\boldsymbol{\alpha}}|}}{\log|\mathbf{n}|} \quad \text{and} \quad c_{\mathbf{n}} = \delta\sqrt{|\mathbf{n}^{\boldsymbol{\alpha}}| \log|\mathbf{n}|},$$

for $\delta > 0$ small; note $|\mathbf{n}^{\boldsymbol{\alpha}}|$ instead of $|\mathbf{n}|^{\alpha}$.

The argument for $T'''_{\mathbf{n},\mathbf{n}+\mathbf{n}^{\boldsymbol{\alpha}}}$:s is verbatim as before, and leads to the sum

$$\sum_{\mathbf{n}} P(|X_{\mathbf{n}}| > \eta\sqrt{|\mathbf{n}^{\boldsymbol{\alpha}}|\log|\mathbf{n}|}) = \sum_{\mathbf{n}} P(|X| > \eta\sqrt{|\mathbf{n}^{\alpha}|\log|\mathbf{n}|}) < \infty,$$

where the finiteness is a consequence of the moment assumption, which this time is a consequence of the first half of Lemma 5.4.

The remaining part of the proof amounts to analogous changes.

7. The Hsu–Robbins–Erdős–Spitzer–Baum–Katz theorem

One aspect of the seminal paper Hsu and Robbins (1947) is that it started an area of research related to convergence rates in the law of large numbers, which, in turn, culminated in the now classical paper Baum and Katz (1965), in which the equivalence of (7.1), (7.2), and (7.4) below was demonstrated. Namely, in Hsu and Robbins (1947) the authors introduced the concept of *complete convergence*, and proved that the sequence of arithmetic means of i.i.d. random variables converges completely to the expected value of the variables provided their variance is finite. The necessity was proved by Erdős (1949, 1950).

Theorem 7.1. *Let $r > 0$, $\alpha > 1/2$, and $\alpha r \geq 1$. Suppose that $X_1, X_2, \dots$ are i.i.d. random variables with partial sums $S_n = \sum_{k=1}^{n} X_k$, $n \geq 1$. If*

$$E|X|^r < \infty \quad \textit{and, if } r \geq 1, \quad E\,X = 0, \tag{7.1}$$

then

$$\sum_{n=1}^{\infty} n^{\alpha r-2} P(|S_n| > n^{\alpha}\varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 0; \tag{7.2}$$

$$\sum_{n=1}^{\infty} n^{\alpha r-2} P(\max_{1\leq k\leq n} |S_k| > n^{\alpha}\varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 0. \tag{7.3}$$

If $\alpha r > 1$ we also have

$$\sum_{n=1}^{\infty} n^{\alpha r-2} P(\sup_{k\geq n} |S_k/k^{\alpha}| > \varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 0. \tag{7.4}$$

Conversely, if one of the sums is finite for all $\varepsilon > 0$, then so are the others (for appropriate values of r and α), $E|X|^r < \infty$ and, if $r \geq 1$, $E\,X = 0$.

The Hsu–Robbins–Erdős part corresponds to the equivalence of (7.1) and (7.2) for the case $r = 2$ and $p = 1$. Spitzer (1956) verified the same for the case $r = p = 1$, and Katz (1963), followed by Baum and Katz (1965) took care of the equivalence between (7.1), (7.2), and (7.4) as formulated in the theorem. Chow (1973) proved that (7.3) holds iff (7.1) does, somewhat differently.

On the other hand, the equivalence of (7.2) and (7.3) is trivial one way and follows via the Lévy inequalities (more precisely via the standard Lévy inequalities as given in e.g. Gut (2007), Theorem 3.7.1 in conjunction with Proposition 3.6.1 there). The implication (7.4) $\Longrightarrow$ (7.2) is also trivial and the converse follows via a "slicing device" introduced in Baum and Katz (1965).

Remark 7.2. Strictly speaking, if one of the sums is finite for *some* $\varepsilon > 0$, then so are the others, and $E|X|^r < \infty$. However, we need convergence *for all* $\varepsilon > 0$ in order to infer that $E\,X = 0$ for the case $r \geq 1$. The same remark applies below.

Before continuing we pause for a moment and consider, for simplicity, the Hsu–Robbins–Erdős case $r = 2$ and $\alpha = 1$, for which the original proof of the implication (7.1) $\Longrightarrow$ (7.2) was technically very intricate.

The first and obvious attempt in order to find a simple proof of this implication fails, as is frequently the case, because of the divergence of the harmonic series. Namely, if $E\,X = 0$ and $\operatorname{Var} X = \sigma^2 < \infty$, then, by Chebyshev's inequality, we have

$$\sum_{n=1}^{\infty} P(|S_n| > n\varepsilon) \leq \sum_{n=1}^{\infty} \frac{\sigma^2}{n\varepsilon^2} = +\infty \quad \text{for any} \quad \varepsilon > 0.$$

However, a fascinating inequality, due to Kahane (1985) and Hoffmann-Jørgensen (1974), see also Gut (2007), Theorem 3.7.5, turns out to be an extremely efficient remedy.

Namely, the KHJ-inequality tells us that for independent *symmetric* random variables one has

$$P(|S_n| > 3n\varepsilon) \leq P(\max_{1\leq k\leq n} |X_k| > n\varepsilon) + 4\big(P(|S_n| > n\varepsilon)\big)^2, \tag{7.5}$$

which, since $P(\max_{1\leq k\leq n} |X_k| > n\varepsilon) \leq nP(|X| > n\varepsilon)$, yields

$$\begin{aligned}\sum_{n=1}^{\infty} P(|S_n| > 3n\varepsilon) &\leq \sum_{n=1}^{\infty} nP(|X| > n\varepsilon) + 4\big(P(|S_n| > n\varepsilon)\big)^2 \\ &\leq E(X/\varepsilon)^2 + 4\sum_{n=1}^{\infty}\Big(\frac{\sigma^2}{n\varepsilon^2}\Big)^2 = \frac{E\,X^2}{\varepsilon^2} + 4\frac{\sigma^4}{\varepsilon^4}\frac{\pi^2}{6},\end{aligned}$$

where, in the last inequality, we exploited Proposition 5.1.

Symmetrizing and desymmetrizing follow standard procedures. For a complete proof of the implication in the general case, one can iterate the KHJ-inequality and exploit the Marcinkiewicz–Zygmund (moment) inequalities in order to cover everything (except for the case $r = p$ for which truncation and a WLLN-type of argument is used). For details and a full proof we refer to Gut (2007), Section 6.11, the proof of which is based on Gut (1978), where the random field version, Theorem 8.1 below, was proved.

The beauty of this proof, thanks to KHJ, is the squaring of the Chebyshev estimate, in that $\sum_{n=1}^{\infty} n^{-1}$ (which is divergent) is replaced by $\sum_{n=1}^{\infty} n^{-2}$ (which is convergent).

We close by mentioning that for the limiting case $p = 2$ one is in the realm of the central limit theorem, and since the individual probabilities do not converge to zero in that case, there is of course no way of having their sums converge. However, by replacing, what would then be $\sqrt{n}$ by $\sqrt{n \log n}$ or even by $\sqrt{n \log\log n}$ there exist positive results; cf. Davis (1968a, 1968b), Lai (1974) for more.

8. The H-R-E-S-B-K theorem for random fields

The obvious question at this point is: What about random field versions?

Theorem 8.1. *Let $r > 0$ and $\alpha > 1/2$ with $\alpha r \geq 1$, suppose that $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d\}$ are i.i.d. random variables, and set $S_{\mathbf{n}} = \sum_{\mathbf{k} \leq \mathbf{n}} X_{\mathbf{k}}$, $\mathbf{n} \in \mathbf{Z}_+^d$. If*

$$E|X|^r (\log^+ |X|)^{d-1} < \infty \quad \text{and, if } r \geq 1, \quad E\,X = 0, \tag{8.1}$$

then

$$\sum_{\mathbf{n}} |\mathbf{n}|^{\alpha r - 2} P(|S_{\mathbf{n}}| > |\mathbf{n}|^{\alpha} \varepsilon) < \infty \quad \text{for all} \quad \varepsilon > 0\,; \tag{8.2}$$

$$\sum_{\mathbf{n}} |\mathbf{n}|^{\alpha r - 2} P(\max_{\mathbf{k} \leq \mathbf{n}} |S_{\mathbf{k}}| > |\mathbf{n}|^{\alpha} \varepsilon) < \infty \quad \text{for all} \quad \varepsilon > 0. \tag{8.3}$$

If $\alpha r > 1$ we also have

$$\sum_{j=1}^{\infty} j^{\alpha r - 2} P(\sup_{j \leq |\mathbf{k}|} |S_{\mathbf{k}}/|\mathbf{k}|^{\alpha}| > \varepsilon) < \infty \quad \text{for all} \quad \varepsilon > 0. \tag{8.4}$$

Conversely, if one of the sums is finite for all $\varepsilon > 0$, then $E|X|^r (\log^+ |X|)^{d-1} < \infty$ and, if $r \geq 1$, $E\,X = 0$.

This is Theorem 4.1 in Gut (1978). As for the proof we only mention that the KHJ- and the Marcinkiewicz–Zygmund inequalities concern sums and consequently remain valid also for random fields. The proof of (8.1) $\Longrightarrow$ (8.2) therefore follows along the same lines as above (with an application to Lemma 5.2 for the appropriate moment condition).

The same can be said about the equivalence (8.2) $\Longleftrightarrow$ (8.3) (with a $\mathbf{Z}_+^d$-version of the Lévy inequality replacing the standard one). The implication (8.4) $\Longrightarrow$ (8.2) is trivial again, and the converse follows via an elaboration of the slicing device of Baum and Katz (1965). We refer to Gut (1978) for details in the multiindex setting.

As the reader may have guessed by now, the next point on the agenda is the case of unequal α:s. Toward that end we first recall, from Subsection 4.2, that α is replaced by $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_d)$, where, as before,

$$p = \max\{k : \alpha_k = \alpha_1\},$$

although now,

$$\frac{1}{2} \le \alpha_1 \le \alpha_2 \le \cdots \le \alpha_d \le 1,$$

The reason for the lower bound 1/2 is, as was hinted at before, the central limit theorem. In fact, supposing that $\alpha_1 = 1/2$, then, for any $\varepsilon > 0$, we have

$$\sum_{\mathbf{n}} |\mathbf{n}|^{(r/2)-2} P(|S_{\mathbf{n}}| > |\mathbf{n}^{\boldsymbol{\alpha}}|\varepsilon) \ge \sum_{i=1}^{\infty} i^{(r/2)-2} P(|S_{i,1,1,\dots,1}| > \sqrt{i}\cdot 1\cdot 1 \cdots 1\cdot 1\cdot \varepsilon) = +\infty.$$

Our first result extends Theorem 8.1. The proof follows the basic lines of that of Theorem 8.1 with obvious changes, such as $|\mathbf{n}^{\boldsymbol{\alpha}}|$ instead of $|\mathbf{n}|^{\alpha}$, and the additional technicalities inflicted by the unequalness of the α:s. We refer to Gut and Stadtmüller (2012) for details.

Theorem 8.2. *Let $r > 0$, suppose that $\alpha_1 > 1/2$, that $\alpha_1 r \ge 1$, let $\{X_{\mathbf{k}},\, \mathbf{k} \in \mathbf{Z}_+^d\}$ be i.i.d. random variables, and set $S_{\mathbf{n}} = \sum_{\mathbf{k}\le\mathbf{n}} X_{\mathbf{k}}$, $\mathbf{n} \in \mathbf{Z}_+^d$. If*

$$E|X|^r(\log^+|X|)^{p-1} < \infty \quad \textit{and, if } r \ge 1, \quad E\,X = 0,$$

then

$$\sum_{\mathbf{n}} |\mathbf{n}|^{\alpha_1 r-2} P(|S_{\mathbf{n}}| > |\mathbf{n}^{\boldsymbol{\alpha}}|\varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 0\,;$$

$$\sum_{\mathbf{n}} |\mathbf{n}|^{\alpha_1 r-2} P(\max_{\mathbf{k}\le\mathbf{n}} |S_{\mathbf{k}}| > |\mathbf{n}^{\boldsymbol{\alpha}}|\varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 0.$$

If $\alpha_1 r > 1$ we also have

$$\sum_{j=1}^{\infty} j^{\alpha_1 r-2} P(\sup_{j\le|\mathbf{k}|} |S_{\mathbf{k}}/|\mathbf{k}^{\boldsymbol{\alpha}}|| > \varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 0.$$

Conversely, if one of the sums is finite for all $\varepsilon > 0$, then $E|X|^r(\log^+|X|)^{p-1} < \infty$ and, if $r \ge 1$, $E\,X = 0$.

In order to illustrate, once more, the efficiency of the KHJ-inequality we show how the proof for the first sum works in the special case when $\alpha_1 r = 2$ and the summands are symmetric. Following the procedure from the proof of Theorem 7.1 we obtain

$$\begin{aligned}\sum_{\mathbf{n}} P(|S_{\mathbf{n}}| > 3^j|\mathbf{n}^{\boldsymbol{\alpha}}|\varepsilon) &\le \sum_{\mathbf{n}} P(|X| > |\mathbf{n}^{\boldsymbol{\alpha}}|\varepsilon) + 4\sum_{\mathbf{n}} \Big(P(|S_{\mathbf{n}}| > |\mathbf{n}^{\boldsymbol{\alpha}}|\varepsilon)\Big)^2 \\ &\le \sum_{\mathbf{n}} P(|X| > |\mathbf{n}^{\boldsymbol{\alpha}}|\varepsilon) + \frac{4\sigma^4}{\varepsilon^4} \sum_{\mathbf{n}} \Big(\frac{|\mathbf{n}|\sigma^2}{|\mathbf{n}^{\boldsymbol{\alpha}}|^2\varepsilon^2}\Big)^2 \\ &= \sum_{\mathbf{n}} P(|X| > |\mathbf{n}^{\boldsymbol{\alpha}}|\varepsilon) + \frac{4\sigma^4}{\varepsilon^4} \prod_{i=1}^{d} \sum_{n_i=1}^{\infty} n_i^{-2(2\alpha_i-1)}.\end{aligned}$$

Now, the first sum is finite iff the desired moment condition is fulfilled (Lemma 5.2), and the second one is finite, since the last exponent $2(2\alpha_i - 1) > 1$ for all i.

Full details are given in Gut and Stadtmüller (2012), Section 3.

As mentioned some lines ago, there are no positive results when $\alpha_1 = 1/2$. However, by adding logarithms as in Lai (1974), Gut (1980), maybe ...?

In the following we first let "some" ($= p \leq d$) of the α:s be equal to $1/2$ with additional logarithms or iterated logarithms and some ($= d - p \geq 0$) of them be strictly larger than $1/2$, after which we consider the complete mixture with $q > p$ of the α:s being equal to $1/2$, the p first of them with additional logarithms, the $q-p$ next ones with additional iterated logarithms, and the $d-q$ largest ones being $> 1/2$. For proofs we refer to Gut and Stadtmüller (2012).

Theorem 8.3. *Let $r \geq 2$, suppose that $\alpha_1 = 1/2$ (and thus, in particular, that $\alpha_1 r \geq 1$), let $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d\}$ be i.i.d. random variables, and set $S_{\mathbf{n}} = \sum_{\mathbf{k}\leq\mathbf{n}} X_{\mathbf{k}}$, $\mathbf{n} \in \mathbf{Z}_+^d$. If*

$$E|X|^r(\log^+|X|)^{p-1-r/2} < \infty, \quad E\,X = 0, \quad \textit{and} \quad \operatorname{Var} X = \sigma^2 < \infty,$$

then

$$\sum_{\mathbf{n}} |\mathbf{n}|^{(r/2)-2} P\Big(|S_{\mathbf{n}}| > \sqrt{\prod_{i=1}^{p} n_i \cdot \log\big(\prod_{i=1}^{p} n_i\big) \prod_{i=p+1}^{d} n_i^{\alpha_i} \cdot \varepsilon}\Big) < \infty$$

for $\varepsilon > \sigma\sqrt{r-2}$;

$$\sum_{\mathbf{n}} |\mathbf{n}|^{(r/2)-2} P\Big(\max_{\mathbf{k}\leq\mathbf{n}} |S_{\mathbf{k}}| > \sqrt{\prod_{i=1}^{p} n_i \cdot \log\big(\prod_{i=1}^{p} n_i\big) \prod_{i=p+1}^{d} n_i^{\alpha_i} \cdot \varepsilon}\Big) < \infty$$

for $\varepsilon > \sigma\sqrt{r-2}$. If $\alpha_1 r > 1$, i.e. if $r > 2$, then we also have

$$\sum_{j=1}^{\infty} j^{(r/2)-2} P\Big(\sup_{j\leq|\mathbf{k}|} \Big|S_{\mathbf{k}} \Big/ \sqrt{\prod_{i=1}^{p} k_i \log k_i} \prod_{i=p+1}^{d} k_i^{\alpha_i}\Big| > \varepsilon\Big) < \infty \ \textit{for all}\ \varepsilon > \sigma\sqrt{r-2}.$$

Conversely, suppose that either $r = 2$ and $p \geq 2$, or that $r > 2$. If one of the sums is finite for some $\varepsilon > 0$, then $E|X|^r(\log^+|X|)^{p-1-r/2} < \infty$ and $E\,X = 0$.

Theorem 8.4. *Suppose that $\alpha_1 = 1/2$, that $p \geq 2$, let $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d\}$ be i.i.d. random variables, and set $S_{\mathbf{n}} = \sum_{\mathbf{k}\leq\mathbf{n}} X_{\mathbf{k}}$, $\mathbf{n} \in \mathbf{Z}_+^d$. If*

$$E\,X^2 \frac{(\log^+|X|)^{p-1}}{\log^+\log^+|X|} < \infty, \quad E\,X = 0, \quad \textit{and} \quad \operatorname{Var} X = \sigma^2,$$

then, for $\varepsilon > \sigma\sqrt{2p}$,

$$\sum_{\mathbf{n}} \frac{1}{|\mathbf{n}|} P\Big(|S_{\mathbf{n}}| > \sqrt{\prod_{i=1}^{p} n_i \cdot \log\log\big(\prod_{i=1}^{p} n_i\big) \prod_{i=p+1}^{d} n_i^{\alpha_i} \cdot \varepsilon}\Big) < \infty\,;$$

$$\sum_{\mathbf{n}} \frac{1}{|\mathbf{n}|} P\big(\max_{\mathbf{k}\le\mathbf{n}} |S_{\mathbf{k}}| > \sqrt{\prod_{i=1}^{p} n_i \cdot \log\log\big(\prod_{i=1}^{p} n_i\big) \prod_{i=p+1}^{d} n_i^{\alpha_i} \cdot \varepsilon}\big) < \infty.$$

Conversely, if one of the sums is finite for some $\varepsilon > 0$, *then* $E\,X^2 \frac{(\log^+|X|)^{p-1}}{\log^+\log^+|X|} < \infty$ *and* $E\,X = 0$.

Theorem 8.5. *Suppose that* $\alpha_1 = 1/2$, *that* $2 \le p < d$, *let* $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d\}$ *be i.i.d. random variables, and set* $S_{\mathbf{n}} = \sum_{\mathbf{k}\le\mathbf{n}} X_{\mathbf{k}}$, $\mathbf{n} \in \mathbf{Z}_+^d$. *If*

$$E\,X^2 \frac{(\log^+|X|)^{d-2}}{\log^+\log^+|X|} < \infty, \quad E\,X = 0, \quad \textit{and} \quad \operatorname{Var} X = \sigma^2,$$

then, for $\varepsilon > \sigma\sqrt{2p}$,

$$\sum_{\mathbf{n}} \frac{1}{|\mathbf{n}|} P\big(|S_{\mathbf{n}}| > \sqrt{\prod_{i=1}^{p} n_i \cdot \log\log\big(\prod_{i=1}^{p} n_i\big) \cdot \log\big(\prod_{i=p+1}^{d} n_i\big) \cdot \varepsilon}\big) < \infty\,;$$

$$\sum_{\mathbf{n}} \frac{1}{|\mathbf{n}|} P\big(\max_{\mathbf{k}\le\mathbf{n}} |S_{\mathbf{k}}| > \sqrt{\prod_{i=1}^{p} n_i \cdot \log\log\big(\prod_{i=1}^{p} n_i\big) \cdot \log\big(\prod_{i=p+1}^{d} n_i\big) \cdot \varepsilon}\big) < \infty.$$

Conversely, if one of the sums is finite for some $\varepsilon > 0$, *then* $E\,X^2 \frac{(\log^+|X|)^{d-2}}{\log^+\log^+|X|}$ *and* $E\,X = 0$.

Theorem 8.6. *Suppose that* $\alpha_1 = 1/2$, *that* $2 \le p < q < d$, *let* $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d\}$ *be i.i.d. random variables, and set* $S_{\mathbf{n}} = \sum_{\mathbf{k}\le\mathbf{n}} X_{\mathbf{k}}$, $\mathbf{n} \in \mathbf{Z}_+^d$. *If*

$$E|X|^2 \frac{(\log^+|X|)^{q-2}}{\log^+\log^+|X|} < \infty, \quad E\,X = 0, \quad \textit{and} \quad \operatorname{Var} X = \sigma^2,$$

then, for $\varepsilon > \sigma\sqrt{2p}$, *we have*

$$\sum_{\mathbf{n}} \frac{1}{|\mathbf{n}|} P\big(|S_{\mathbf{n}}| > \sqrt{\prod_{i=1}^{q} n_i \log\log\big(\prod_{i=1}^{p} n_i\big) \cdot \log\big(\prod_{i=p+1}^{q} n_i\big)} \cdot \prod_{i=q+1}^{d} n_i^{\alpha_i} \cdot \varepsilon\big) < \infty\,;$$

$$\sum_{\mathbf{n}} \frac{1}{|\mathbf{n}|} P\big(\max_{\mathbf{k}\le\mathbf{n}} |S_{\mathbf{k}}| > \sqrt{\prod_{i=1}^{q} n_i \log\log\big(\prod_{i=1}^{p} n_i\big) \cdot \log\big(\prod_{i=p+1}^{q} n_i\big)} \cdot \prod_{i=q+1}^{d} n_i^{\alpha_i} \cdot \varepsilon\big) < \infty.$$

Conversely, if one of the sums is finite for some $\varepsilon > 0$, *then* $E|X|^2 \frac{(\log^+|X|)^{q-2}}{\log^+\log^+|X|}$ *and* $E\,X = 0$.

Remark 8.7. When $p = d = 1$ Theorem 8.3 reduces to Lai (1974), Theorem 3, and for $p = d \ge 2$ to Gut (1980), Theorems 3.4 and 3.5. When $p = d$ in Theorem 8.4 one rediscovers Gut (1980), Theorem 6.2.

Remark 8.8. The reason for strict inequalities between p, q, and d in the last two results is that there is no "continuity" in the moment assumptions between those theorems and the earlier ones.

Remark 8.9. The first and necessary moment condition in Theorem 8.3 implies, in particular, that the variance is finite *except for the case when* $r = 2$ and $p = 1$. However, one can show (cf. Gut (1980), p. 301) that an intermediate condition is sufficient when $r = 2$ and $(p =)\, d = 1$ in the symmetric case. For the complicated precise condition and for more on this exceptional case we refer to Spătaru (2001). A similar remark applies to the case $p = 1$, since the variance is automatically finite unless $p = 1$.

A related problem occurs in the LIL where the proof of the necessity is "easy" when $d \geq 2$ and "hard" when $d = 1$.

9. Two additional problems

9.1. Other weights

In all results of the H-R-E-S-B-K kind the probabilities have had polynomial weights so far. So, what happens if the weights grow faster than polynomially? But not fast enough for the moment generating function to exist?

A first result in this direction is due to Lanzinger (1998), and corresponds to the equivalence of the moment condition and the convergence of the first sum for $d = 1$ (in a two-sided and, thus, stronger form) in the following result.

Theorem 9.1. *Let* $0 < \alpha < 1$*, and suppose that* $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d\}$ *are i.i.d. random variables with* $E\,X = 0$ *and partial sums* $S_{\mathbf{n}} = \sum_{\mathbf{k} \leq \mathbf{n}} X_{\mathbf{k}}$, $\mathbf{n} \in \mathbf{Z}_+^d$*. The following are equivalent:*

$$E \exp\{|X|^\alpha\}(\log^+ |X|)^{d-1} < \infty;$$

$$\sum_{\mathbf{n}} \exp\{|\mathbf{n}|^\alpha\} \cdot |\mathbf{n}|^{\alpha-2} P(|S_{\mathbf{n}}| > |\mathbf{n}|\varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 1;$$

$$\sum_{\mathbf{n}} \exp\{|\mathbf{n}|^\alpha\} \cdot |\mathbf{n}|^{\alpha-2} P(\max_{\mathbf{k} \leq \mathbf{n}} |S_{\mathbf{k}}| > |\mathbf{n}|\varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 1;$$

$$\sum_{j=1}^{\infty} \exp\{j^\alpha\} \cdot j^{\alpha-2} P(\sup_{j \leq |\mathbf{k}|} |S_{\mathbf{k}}/|\mathbf{k}|| > \varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 1.$$

There remains, in fact, an intermediate case, namely, when the weights are between polynomial and exponential in the following sense.

Theorem 9.2. *Let* $\alpha > 1$*, and suppose that* $\{X_{\mathbf{k}}, \mathbf{k} \in \mathbf{Z}_+^d\}$ *are i.i.d. random variables with* $E\,X = 0$ *and partial sums* $S_{\mathbf{n}} = \sum_{\mathbf{k} \leq \mathbf{n}} X_{\mathbf{k}}$, $\mathbf{n} \in \mathbf{Z}_+^d$*. The following are equivalent:*

$$E \exp\{(\log |X|)^\alpha\}(\log^+ |X|)^{d-1} < \infty;$$

$$\sum_{\mathbf{n}} \exp\{(\log|\mathbf{n}|)^\alpha\} \cdot \frac{(\log|\mathbf{n}|)^{\alpha-1}}{|\mathbf{n}|^2} P(|S_{\mathbf{n}}| > |\mathbf{n}|\varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 1;$$

$$\sum_{\mathbf{n}} \exp\{(\log|\mathbf{n}|)^\alpha\} \cdot \frac{(\log|\mathbf{n}|)^{\alpha-1}}{|\mathbf{n}|^2} P(\max_{\mathbf{k}\leq\mathbf{n}} |S_{\mathbf{k}}| > |\mathbf{n}|\varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 1;$$

$$\sum_{j=1}^{\infty} \exp\{(\log j)^\alpha\} \cdot \frac{(\log j)^{\alpha-1}}{j^2} P(\sup_{j\leq|\mathbf{k}|} |S_{\mathbf{k}}/|\mathbf{k}|| > \varepsilon) < \infty \quad \textit{for all} \quad \varepsilon > 1.$$

Once again, we refer to the original source Gut and Stadtmüller (2011c) for proofs and further details.

9.2. Last exit times

A strong limit theorem tells us, i.a., that the number of exceedances of some kind is a.s. finite. For the LLN (with obvious notation) this means that $P(|S_n| > n\varepsilon \text{ i.o.}) = 0$ for any $\varepsilon > 0$. Now, given this, one may ask for the number of them or the last time an exceedance occurs, which is called the *last exit time*, denoted $L(\varepsilon)$. The LLN is thus equivalent to the statement $P(L(\varepsilon) < \infty) = 1$. When $d = 1$ it is (maybe) more natural to put interest in $N(\varepsilon) =$ the number of exceedances, but, due to the partial order of $\mathbf{Z}_+^d$ we shall stick to last exit times here.

The point is that there is an obvious connection to the previous results. Namely, letting $a_{\mathbf{n}}$ denote $|\mathbf{n}|^\alpha$, $|\mathbf{n}^{\boldsymbol{\alpha}}|$, $\sqrt{|\mathbf{n}|\log|\mathbf{n}|}$, or $\sqrt{|\mathbf{n}|\log\log|\mathbf{n}|}$, then, for

$$L_d(\varepsilon) = \sup\{|\mathbf{n}| : |S_{\mathbf{n}}| > a_{\mathbf{n}}\varepsilon\},$$

we always have

$$\{L_d(\varepsilon) \geq j\} = \{\sup_{j\leq|\mathbf{k}|} |S_{\mathbf{k}}/a_{\mathbf{k}}| > \varepsilon\},$$

which implies, for example, that

$$E\big(L_d(\varepsilon)\big)^r \asymp \sum_{j=1}^{\infty} j^{r-1} P(\sup_{j\leq|\mathbf{k}|} |S_{\mathbf{k}}/a_{\mathbf{k}}| > \varepsilon),$$

after which the appropriate result above provides the relevant conditions for a moment of a given order to exist.

We confine ourselves with providing two examples, and leave it to the reader to invent the conclusion of his/her favorite choice.

Theorem 9.3. *Let $\alpha_1 > 1/2$, $\alpha_1 r > 1$, and set $L_d(\varepsilon) = \sup\{|\mathbf{n}| : |S_{\mathbf{n}}| > |\mathbf{n}^{\boldsymbol{\alpha}}|\varepsilon\}$. The following are equivalent:*

$$E|X|^r(\log^+|X|)^{p-1} < \infty \quad \textit{and, if} \quad r \geq 1,\ E\,X = 0,$$
$$E\big(L_d(\varepsilon)\big)^{\alpha_1 r-1} < \infty \quad \textit{for all} \quad \varepsilon > 0.$$

Theorem 9.3 is (of course) related to Theorem 8.2. When $p = d$ it reduces to Gut (1980), Theorem 8.1.

As a final remark we mention that, for $a_n = \sqrt{n \log\log n}$, Slivka (1969) showed that no finite moment exists for the corresponding counting variable, which immediately implies the same for the last exit times and, all the more, for $L_d(\varepsilon)$. However, it was shown in Gut (1980), Theorem 8.3, that logarithmic moments may exist. More precisely:

Theorem 9.4. *Let* $L_d(\varepsilon) = \sup\{|\mathbf{n}| : |S_{\mathbf{n}}| > \sqrt{|\mathbf{n}| \log\log |\mathbf{n}|\varepsilon}\}$*, and suppose that* $E\,X = 0$ *and* $\operatorname{Var} X = \sigma^2 < \infty$*. Then*
(a) $E\big(L_d(\varepsilon)\big)^r = +\infty$ *for all* $r > 0$ *and all* $\varepsilon > 0$.
(b) *If, in addition,* $EX^2 \frac{(\log^+ |X|)^d}{\log\log^+ |X|} < \infty$*, then* $E \log L_d(\varepsilon) < \infty$ *for* $\varepsilon > \sigma\sqrt{2(d+1)}$.

10. Martingales and the LLN for random fields

New problems appear in random field settings, because there exist *four* different definitions of martingales.

In the standard definition one defines a family of nested σ-algebras $\{\mathcal{F}_{\mathbf{n}}, \mathbf{n} \in \mathbf{Z}_+^d\}$ and an adapted family $\{X_{\mathbf{n}}, \mathbf{n} \in \mathbf{Z}_+^d\}$ of random variables, which together constitute a martingale iff

$$E(X_{\mathbf{n}} \mid \mathcal{F}_{\mathbf{m}}) = X_{\mathbf{m}} \quad \text{for} \quad \mathbf{m} \leq \mathbf{n}.$$

The martingale convergence theorem runs as follows.

Theorem 10.1. (a) *If* $\{X_{\mathbf{n}}, \mathbf{n} \in \mathbf{Z}_+^d\}$ *is a martingale, such that*

$$\sup_{\mathbf{n}} E|X_{\mathbf{n}}|(\log^+ |X_{\mathbf{n}}|)^{d-1} < \infty,$$

then $X_{\mathbf{n}}$ *converges almost surely as* $\mathbf{n} \to \infty$.
(b) *The same is true if the index set is a sector* S_θ^d *in* $\mathbf{Z}_+^d$.

Now, introducing a random field $\{Y_{\mathbf{n}}, \mathbf{n} \in \mathbf{Z}_+^d\}$ of i.i.d. random variables, it is known that the field $\{X_{\mathbf{n}} = \frac{1}{|\mathbf{n}|}\sum_{\mathbf{k}\leq\mathbf{n}} Y_{\mathbf{k}}\}$, where $\mathbf{n} \in \mathbf{Z}_+^d$ or $\mathbf{n} \in S_\theta^d$, of arithmetic means consitute reversed martingales to which Theorem 10.1 is applicable.

The LLN thus follows *immediately* from Theorem 10.1.

We may thus combine our knowledge about the law of large numbers and about martingales as follows:

- The LLN in $\mathbf{Z}_+^d$ holds iff $E\,M(|Y|) < \infty$ i.e., iff $E|Y|(\log^+ |Y|)^{d-1} < \infty$;
- The LLN in the sector S_θ^d holds iff $E\,M(|Y|) < \infty$ i.e., iff $E|Y| < \infty$;
- Martingale convergence holds in both cases iff $E|Y|(\log^+ |Y|)^{d-1} < \infty$.

The moral of the story is that for the *sector* the martingale proof yields a weaker result, since the LLN requires only finite mean. The explanation is that

- LLN: The decisive point concerning logarithms or not is the *size* of the index set.
- Martingales: Logarithms are present because of the *dimension* of the index set.

So, even though the martingale proof is an elegant so-called one-line proof it is inferior in cases such as the sector.

Acknowledgement

This paper reviews joint work with Professor Ulrich Stadtmüller that was presented at the *Conference on Stochastic Models and their Applications* in Debrecen in August, 2011. I wish to thank Professor István Fazekas for inviting me to the conference, for his generous hospitality, and for encouraging me to write this article. Finally, I would like to thank the University of Debrecen for financial support.

References

[1] Baum, L.E. and Katz, M. (1965). Convergence rates in the law of large numbers. *Trans. Amer. Math. Soc.* **120**, 108-123.

[2] Bingham, N.H. (1984). On Valiron and circle convergence. *Math. Z.* **186**, 273-286.

[3] Bingham, N.H. and Goldie, C.M. (1983). On one-sided Tauberian conditions. *Analysis* **3**, 150-188.

[4] Bingham, N.H. and Goldie, C.M. (1988). Riesz means and self-neglecting functions. *Math. Z.* **199**, 443-454.

[5] Bingham, N.H. and Maejima, M. (1985). Summability methods and almost sure convergence. *Z. Wahrscheinlickeitsth. verw. Gebiete* **68**, 383-392.

[6] Bingham, N.H., Goldie, C.M., and Teugels, J.L. (1987). *Regular Variation.* Cambridge University Press, Cambridge.

[7] Chow, Y.S. (1973). Delayed sums and Borel summability of independent, identically distributed random variables. *Bull. Inst. Math. Acad. Sinica* **1**, 207-220.

[8] Csörgö, M. and Révész, P. (1981). *Strong Approximations in Probability and Statistics*, Academic Press, New York.

[9] Davis, J.A. (1968a). Convergence rates for the law of the iterated logarithm. *Ann. Math. Statist.* **39**, 1479-1485.

[10] Davis, J.A. (1968b). Convergence rates for probabilities of moderate deviations. *Ann. Math. Statist.* **39**, 2016-2028.

[11] Erdős, P. (1949). On a theorem of Hsu and Robbins. *Ann. Math. Statist.* **20**, 286-291.

[12] Erdős, P. (1950). Remark on my paper "On a theorem of Hsu and Robbins". *Ann. Math. Statist.* **21**, 138.

[13] Erdős, P. and Rényi, A. (1970). On a new law of large numbers. *J. Analyse Math.* **23**, 103-111.

[14] Gut, A. (1978). Marcinkiewicz laws and convergence rates in the law of large numbers for random variables with multidimensional indices. *Ann. Probab.* **6**, 469-482.

[15] Gut, A. (1980). Convergence rates for probabilities of moderate deviations for sums of random variables with multidimensional indices. *Ann. Probab.* **8**, 298-313.

[16] Gut, A. (1983). Strong laws for independent and identically distributed random variables and convergence rates in indexed by a sector. *Ann. Probab.* **11**, 569-577.

[17] Gut, A. (1986). Law of the iterated logarithm for subsequences. *Probab. Math. Statist.* **7**, 27-58.

[18] Gut, A. (2007). *Probability: A Graduate Course,* Corr. 2nd printing. Springer-Verlag, New York.

[19] Gut, A. and Stadtmüller, U. (2008a). Laws of the single logarithm for delayed sums of random fields. *Bernoulli*, **14**, 249-276.

[20] Gut, A. and Stadtmüller, U. (2008b). Laws of the single logarithm for delayed sums of random fields II. *J. Math. Anal. Appl.* **346**, 403-413.

[21] Gut, A. and Stadtmüller, U. (2009). An asymmetric Marcinkiewicz-Zygmund LLN for random fields. *Statist. Probab. Lett.* **79**, 1016-1020.

[22] Gut, A. and Stadtmüller, U. (2010). Cesàro summation for random fields. *J. Theoret. Probab.* **23**, 715-728.

[23] Gut, A. and Stadtmüller, U. (2011a). On the LSL for random fields. *J. Theoret. Probab.* **24**, 422-449.

[24] Gut, A. and Stadtmüller, U. (2011b). On the strong law of large numbers for delayed sums and random fields. *Acta Math. Hungar.* **129**, 182-203.

[25] Gut, A. and Stadtmüller, U. (2011c). An intermediate Baum-Katz theorem. *Statist. Probab. Lett.* **81**, 1486-1492.

[26] Gut, A. and Stadtmüller, U. (2012). On the Hsu-Robbins-Erdős-Spitzer-Baum-Katz theorem for random fields. *J. Math. Anal. Appl.* **387**, 447-463.

[27] Gut, A., Jonsson F., and Stadtmüller, U. (2010). Between the LIL and the LSL. *Bernoulli* **16**, 1-22.

[28] Hardy, G.H. and Wright, E.M. (1954). *An Introduction to the Theory of Numbers,* 3rd ed. Oxford University Press.

[29] Hartman, P. and Wintner, A. (1941). On the law of the iterated logarithm. *Amer. J. Math.* **63**, 169-176.

[30] Hoffmann-Jørgensen, J. (1974). Sums of independent Banach space valued random variables. *Studia Math.* **LII**, 159-186.

[31] Hsu, P.L. and Robbins, H. (1947). Complete convergence and the law of large numbers. *Proc. Nat. Acad. Sci. USA* **33**, 25-31.

[32] Kahane, J.-P. (1985). *Some Random Series of Functions*, 2nd ed. Cambridge University Press, Cambridge.

[33] Katz, M. (1963). The probability in the tail of a distribution. *Ann. Math. Statist.* **34**, 312-318.

[34] Lai, T.L. (1974). Limit theorems for delayed sums. *Ann. Probab.* **2**, 432-440.

[35] Lanzinger, H. (1998). A Baum-Katz theorem for random variables under exponential moment conditions. *Statist. Probab. Lett.* **39**, 89-95.

[36] Lanzinger, H. and Stadtmüller, U. (2000). Strong laws for i.i.d. random variables with thin tails, *Bernoulli* **6**, 45-61.

[37] Marcinkiewicz, J. and Zygmund, A. (1937). Sur les fonctions indépendantes. *Fund. Math.* **29**, 60-90.

[38] Slivka, J. (1969). On the law of the iterated logarithm. *Proc. Nat. Acad. Sci. USA* **63**, 289-291.

[39] Smythe, R. (1973). Strong laws of large numbers for r-dimensional arrays of random variables. *Ann. Probab.* **1**, 164-170.

[40] Smythe, R. (1974). Sums of independent random variables on partially ordered sets. *Ann. Probab.* **2**, 906-917.

[41] Spătaru, A. (2001). On a series concerning moderate deviations. *Rev. Roumaine Math. Pures Appl.* **45** (2000), 883-896.

[42] Spitzer, F. (1956). A combinatorial lemma and its applications to probability theory. *Trans. Amer. Math. Soc.* **82**, 323-339.

[43] Spitzer, F. (1976). *Principles of Random Walk,* 2nd ed. Springer-Verlag, New York.

[44] Stadtmüller, U. and Thalmaier, M. (2009). Strong laws for delayed sums of random fields. *Acta Sci. Math. (Szeged)* **75**, 723-737.

[45] Strassen, V. (1966). A converse to the law of the iterated logarithm. *Z. Wahrsch. verw. Gebiete* **4**, 265-268.

[46] Thalmaier, M. (2009). *Grenzwertsätze für gewichtete Summen von Zufallsvariablen und Zufallsfeldern.* Dissertation, University of Ulm.

[47] Titchmarsh, E.C. (1951). *The Theory of the Riemann Zeta-function*, 2nd ed. Oxford University Press.

[48] Wichura, M.J. (1973). Some Strassen-type laws of the iterated logarithm for multiparameter stochastic processes with independent increments. *Ann. Probab.* **1**, 272-296.

Annales Mathematicae et Informaticae
39 *(2012) pp. 159–171*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

The using of wavelet analysis in climatic challenges

Maxim Khaziakhmetov, Tatiana Zakharova

Moscow State University
e-mail: `lsa@cs.msu.su`

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

Nowadays there are many methods for processing of digital signals. A classic example is Fourier analysis. Using this transformation we build a decomposition of a signal by frequencies, so the result is easy for interpretation. But this method works well only with stationary signals, when we can find periodic components. Also using Fourier transform we lose information about coordinates of events in the original signal, because the result of transformation is in terms of frequencies.

Of course, there are alternative ways of signal processing. Wavelet analysis is one of them. Wavelet transform works in a very simple manner. It divides the original signal into two parts – approximation and details. This dichotomization can be repeated many times and we'll have decomposition with multilevel detailization. There are 2 ways for further work: to analyze the result of transformation interpreting it as something or to execute some operations with the result and then use inverse Wavelet transform.

This article is about the using of wavelet analysis in climatic challenges. The work of the authors of this article was to analyze water vapor field of the Earth searching for numerical patterns.

1. Basic information about Wavelet transform

Information about wavelets required for further discussion is placed in this section of the article.

1.1. 1D wavelet transform

1.1.1. 1D discrete wavelet transform

Definition 1.1. The function $\varphi(x) \in L^2(\mathbb{R})$ is scaling function if it can be represented as:

$$\varphi(x) = \sqrt{2} \sum_{n \in \mathbb{Z}} h_n \varphi(2x - n), \tag{1.1}$$

where $h_n, n \in \mathbb{Z}$ satisfy the condition

$$\sum_{n \in \mathbb{Z}} |h_n|^2 < \infty.$$

The equation (1.1) is scaling equation, the set of coefficients $\{h_n\}_{n \in \mathbb{Z}}$ is scaling filter.

Widely known Haar function:

$$F(x) = \begin{cases} 1, x \in [0, 1), \\ 0, \text{ otherwise} \end{cases}$$

is scaling function, but centered Haar function

$$F(x) = \begin{cases} 1, x \in [-\frac{1}{2}, \frac{1}{2}), \\ 0, \text{ otherwise} \end{cases}$$

cannot be classified to scaling functions.

Definition 1.2. Orthogonal multiresolution decomposition (or multiresolution analysis, or MRA) of $L^2(\mathbb{R})$-space is a sequence of confined spaces:

$$\ldots \subset V_{-1} \subset V_0 \subset V_1 \subset \ldots,$$

with following properties:

1. $\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R})$,
2. $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$,
3. $f(x) \in V_j \iff f(2x) \in V_{j+1}$,
4. $\exists \varphi(x) \in V_0 : \{\varphi_{0,n}(x) = \varphi(x - n)\}_{n \in \mathbb{Z}}$ – an orthonormal basis of V_0-space.

Using properties 3 and 4 of Definition 1.2 we can conclude that

$$\{\varphi_{j,n}(x) = \sqrt{2^j} \varphi(2^j x - n)\}_{n \in \mathbb{Z}}$$

is an orthonormal basis of V_j-space.

$\forall j$ we have $V_j \subset V_{j+1}$. Let W_j be an orthogonal complement of V_j to V_{j+1}, i.e. $V_{j+1} = V_j \oplus W_j$. Then $V_{j+1} = V_{j-1} \oplus W_{j-1} \oplus W_j$, because $V_j = V_{j-1} \oplus W_{j-1}$. Repeating this procedure we'll have

$$V_{j+1} = \overline{\bigoplus_{k=-\infty}^{j} W_k}.$$

According to property 1 of Definition 1.2 we can conclude that $L^2(\mathbb{R})$-space has an orthogonal decomposition:

$$L^2(\mathbb{R}) = \overline{\bigoplus_{k=-\infty}^{+\infty} W_k}.$$

Definition 1.3. If a sequence of subspaces $\ldots \subset V_{-1} \subset V_0 \subset V_1 \subset \ldots$ is multiresolution analysis and $\forall j \in \mathbb{Z}$ W_j is the orthogonal complement of V_j to V_{j+1} then each such W_j is wavelet space and its elements are wavelets.

$\exists \psi(x) \in W_0$, called "mother wavelet", with properties

1. $\{\psi(x-n)\}_{n\in\mathbb{Z}}$ – an orthonormal basis of W_0-space,
2. $\{\psi_{j,n}(x) = \sqrt{2^j}\psi(2^j x - n)\}_{n\in\mathbb{Z}}$ – an orthonormal basis of W_j-space for every j.

Let $(f(x), g(x))$ be the scalar product in $L^2(\mathbb{R})$, i.e.

$$\langle f(x), g(x)\rangle = \int_{-\infty}^{+\infty} f(x)g(x)dx. \tag{1.2}$$

Let $f(x) \in V_{j+1}$ then we have the decomposition of $f(x)$:

$$f(x) = \sum_{k=-\infty}^{j} \sum_{n\in\mathbb{Z}} \langle f(x), \psi_{k,n}(x)\rangle \psi_{k,n}(x).$$

In practice the number of detailization levels is finite, so for $f(x)$ we have the following decomposition formula:

$$f(x) = \sum_{k=0}^{j} \sum_{n\in\mathbb{Z}} \langle f(x), \psi_{k,n}(x)\rangle \psi_{k,n}(x) + \sum_{n\in\mathbb{Z}} \langle f(x), \varphi(x-n)\rangle \varphi(x-n),$$

i.e. the signal from V_{j+1}-space is represented in terms of elements of spaces $\{V_0, W_0, \ldots, W_j\}$.

In case of discrete signal the formula (1.2) can be rewritten as the sum. If $x = \{x_n\}_{n\in\mathbb{Z}}$ is the digitization of the signal $x(t), t \in \mathbb{R}$, then wavelet coefficient

$a_{j,n}$ can be represented in terms of discrete convolution with the filter $h = \{h_n\}_{n\in\mathbb{Z}}$ corresponded to function $\psi_{j,n}$:

$$a_{j,n} = \sum_{k=-\infty}^{\infty} h_k x_{n-k}.$$

Usually the process of signal analysis starts from its representation in terms of a basis of V_{j+1}-space. Then we build a decomposition in bases of V_j č W_j constructing approximation and details. We can repeat decomposition or stop the process. So we have the following decomposition scheme:

$$\begin{array}{ccccccc} V_{j+1} & \longrightarrow & V_j & \longrightarrow \ldots \longrightarrow & V_1 & \longrightarrow & V_0 \\ \downarrow & & \downarrow & & \downarrow & & \\ W_j & & W_{j-1} & & W_0 & & \end{array}$$

We stopped at 0-index, so V_0-based approximation component consists of the most general trends of the original signal and W_0-based detailization includes the most spatially extended deviations from these trends.

As we noted in the abstract we can use decomposition coefficients as independent data and create a conclusion based on it or perform operations on them and then reconstruct the signal.

For the reconstruction process we have the following scheme:

$$\begin{array}{ccccccc} V_{j+1} & \longleftarrow & V_j & \longleftarrow \ldots \longleftarrow & V_1 & \longleftarrow & V_0 \\ \uparrow & & \uparrow & & \uparrow & & \\ W_j & & W_{j-1} & & W_0 & & \end{array}$$

1.1.2. 1D continuous wavelet transform

As we know, for 1D discrete wavelet transform the following formula is valid:

$$\psi_{j,n}(x) = \sqrt{2^j}\psi(2^j x - n), \text{ where } j,\ n \in \mathbb{Z}.$$

Continuous wavelet transform is constructed by allowing arbitrary real values of the parameters of scaling and shift (in discrete variant we should use powers of 2 for the scale parameter and integers for shift parameter). This generalization allows to select details of a signal with arbitrary size of their support.

Let $\psi(x)$ be wavelet. $\psi(x) \in L^2(\mathbb{R})$ and also

$$C_\psi = 2\pi \int_{-\infty}^{+\infty} |\omega|^{-1} |\hat{\psi}(\omega)|^2 d\omega < \infty. \tag{1.3}$$

(The relation (1.3) guarantees the existence of inverse continuous transform).

Let a be the scaling parameter and b is the shift parameter, then we have 2-parameter family of wavelets:

$$\psi_{ab}(x) = |a|^{-1/2}\psi\left(\frac{x-b}{a}\right), \text{ where } a,\ b \in \mathbb{R}.$$

1D continuous wavelet transform is defined by the formula:

$$W_\psi\,[f]\,(a,b) = (f, \psi_{a,b}) = |a|^{-1/2} \int\limits_{-\infty}^{+\infty} f(x)\overline{\psi\left(\frac{x-b}{a}\right)}dx. \tag{1.4}$$

It's obvious that the coefficients of 1D discrete wavelet transform can be computed as

$$c_{jk} = W_\psi[f]\left(\frac{1}{2^j}, \frac{k}{2^j}\right).$$

And as in discrete case we have inverse transform:

Theorem 1.4. *If $f(x) \in L^2(\mathbb{R})$ and (1.3) is satisfied, then we have inverse 1D continuous wavelet transform formula:*

$$f(x) = C_\psi^{-1} \iint W_\psi[f](a,b)\psi_{ab}(x)\frac{dadb}{a^2}.$$

1.2. 2D discrete wavelet transform

In this subsection we'll consider the case of functions from $L^2(\mathbb{R}^2)$-space.

The simplest way of generalization 1D wavelet transform to 2D wavelet transform is based on tensor product. We have the following representation for $L^2(\mathbb{R}^2)$-space:

$$L^2(\mathbb{R}^2) = L^2(\mathbb{R}) \otimes L^2(\mathbb{R}).$$

I.e. linear combinations of $f(x)g(y)$ construct dense set in $L^2(\mathbb{R}^2)$.

We'll define V_0^2 as tensor product of V_0:

$$V_0^2 = V_0 \otimes V_0.$$

According to this fact

$$\{\varphi_{0,k,n}(x,y) = \varphi(x-k)\varphi(y-n)\}_{k,n\in\mathbb{Z}}$$

is an orthonormal basis of V_0^2-space.

$V_j^2 = V_j \otimes V_j$ are scaled versions of V_0^2-space, for them we have the following relation

$$f(x,y) \in V_0^2 \iff f(2^jx, 2^jy) \in V_j^2.$$

So, as in 1D-case, there is a sequence of spaces $\ldots \subset V_{-1}^2 \subset V_0^2 \subset V_1^2 \subset \ldots$. Using equation $V_1 = V_0 \oplus W_0$ for $L^2(\mathbb{R}^2)$ we conclude that

$$
\begin{aligned}
V_1^2 = V_1 \otimes V_1 = (V_0 \oplus W_0) \otimes (V_0 \oplus W_0) = \\
= (V_0 \otimes V_0) \oplus (V_0 \otimes W_0) \oplus (W_0 \otimes V_0) \oplus (W_0 \otimes W_0) = \\
= V_0^2 \oplus (V_0 \otimes W_0) \oplus (W_0 \otimes V_0) \oplus (W_0 \otimes W_0).
\end{aligned}
$$

$V_0 \otimes W_0$, $W_0 \otimes V_0$, $W_0 \otimes W_0$ forms 2D-wavelet space W_0^2. The following facts exist:

space $V_0 \otimes W_0$ is constructed by shifts of function $\psi^H(x, y) = \varphi(x)\psi(y)$, we'll designate it as W_0^H – this is the space of horizontal details (ox-oriented homogeneous areas can be selected);

space $W_0 \otimes V_0$ is constructed by shifts of function $\psi^V(x, y) = \psi(x)\varphi(y)$, we'll designate it as W_0^V – this is the space of vertical details (oy-oriented homogeneous areas can be selected);

space $W_0 \otimes W_0$ is constructed by shifts of function $\psi^D(x, y) = \psi(x)\psi(y)$, we'll designate it as W_0^D – this is the space of diagonal details(diagonal inhomogeneous areas can be selected).

So we have the following decomposition:

$$V_{j+1}^2 = V_j^2 \oplus W_j^H \oplus W_j^V \oplus W_j^D \quad \forall j.$$

The following sets of functions are orthonormal bases of listed spaces:

$$
\begin{aligned}
&\{\varphi_{j,k,n}(x, y) = 2^j \varphi(2^j x - k)\varphi(2^j y - n)\}_{k,n \in \mathbb{Z}}; \\
&\{\psi_{j,k,n}^H(x, y) = 2^j \varphi(2^j x - k)\psi(2^j y - n)\}_{k,n \in \mathbb{Z}}; \\
&\{\psi_{j,k,n}^V(x, y) = 2^j \psi(2^j x - k)\varphi(2^j y - n)\}_{k,n \in \mathbb{Z}}; \\
&\{\psi_{j,k,n}^D(x, y) = 2^j \psi(2^j x - k)\psi(2^j y - n)\}_{k,n \in \mathbb{Z}}.
\end{aligned}
$$

I.e. there are 4 types of decomposition coefficients:

$$\langle f(x, y), \varphi_{j,k,n}(x, y)\rangle = 2^j \int\limits_{\mathbb{R}^2} f(x, y)\varphi(2^j x - k)\varphi(2^j y - n)dxdy;$$

$$\langle f(x, y), \psi_{j,k,n}^H(x, y)\rangle = 2^j \int\limits_{\mathbb{R}^2} f(x, y)\varphi(2^j x - k)\psi(2^j y - n)dxdy;$$

$$\langle f(x, y), \psi_{j,k,n}^V(x, y)\rangle = 2^j \int\limits_{\mathbb{R}^2} f(x, y)\psi(2^j x - k)\varphi(2^j y - n)dxdy;$$

$$\langle f(x, y), \psi_{j,k,n}^D(x, y)\rangle = 2^j \int\limits_{\mathbb{R}^2} f(x, y)\psi(2^j x - k)\psi(2^j y - n)dxdy;$$

Schemes of decomposition and reconstruction processes are

$$\begin{array}{ccccccc}
V_{j+1}^2 & \longrightarrow & V_j^2 & \longrightarrow \dots \longrightarrow & V_1^2 & \longrightarrow & V_0^2 \\
\downarrow & & \downarrow & & \downarrow & & \\
W_j^H, W_j^V, W_j^D & & W_{j-1}^H, W_{j-1}^V, W_{j-1}^D & & W_0^H, W_0^V, W_0^D & &
\end{array}$$

$$\begin{array}{ccccccc}
V_{j+1}^2 & \longleftarrow & V_j^2 & \longleftarrow \dots \longleftarrow & V_1^2 & \longleftarrow & V_0^2 \\
\uparrow & & \uparrow & & \uparrow & & \\
W_j^H, W_j^V, W_j^D & & W_{j-1}^H, W_{j-1}^V, W_{j-1}^D & & W_0^H, W_0^V, W_0^D & &
\end{array}$$

2. The task of analysis of water vapor field of the Earth

2.1. Introduction

The Earth atmosphere is very complex and unpredictable. Energy of the atmosphere is contained mostly in water vapor, because of its heat capacity. Study of processes in water vapor field can help us to explain and predict atmospheric phenomena, for example, cyclones and hurricanes' formation – they are the most interesting because of their consequences.

Many atmospheric phenomena have periodic nature, for example, it is easy to understand that water vapor field has an annual cycle of movements. But the most interesting phenomena for scientists are nonstationary. They should be localized in space and time and their parameters should be found so appropriate methods of research are required.

In our research of water vapor field we used wavelet analysis.

2.2. Essence of the task

For every day from 01.01.1999 to 31.12.2009 we had a digital map of the earth with the size 360×720 (grid is 0,5°). A value of every pixel on map is an average density (kg/m^2) of water vapor in spheric segment of the Earth with appropriate geographic coordinates and date (rounded to the nearest whole). So we had 3D-array of data from satellite images.

There are some problems with this array. The algorithm of value construction can be used only for water vapor field above the surface of the oceans, so by this reason the land surface pixels are filled with zeros. Also there are spaces on maps where in some days satellites didn't make images, these pixels are marked with "-20"-value.

It's obvious that the most convenient way to display this information is graphic – for every day we can draw a map of water vapor field as indexed image. The picture (1) demonstrates an example image (the matching of colors to values of density is on the right side): we can see all the problems listed above.

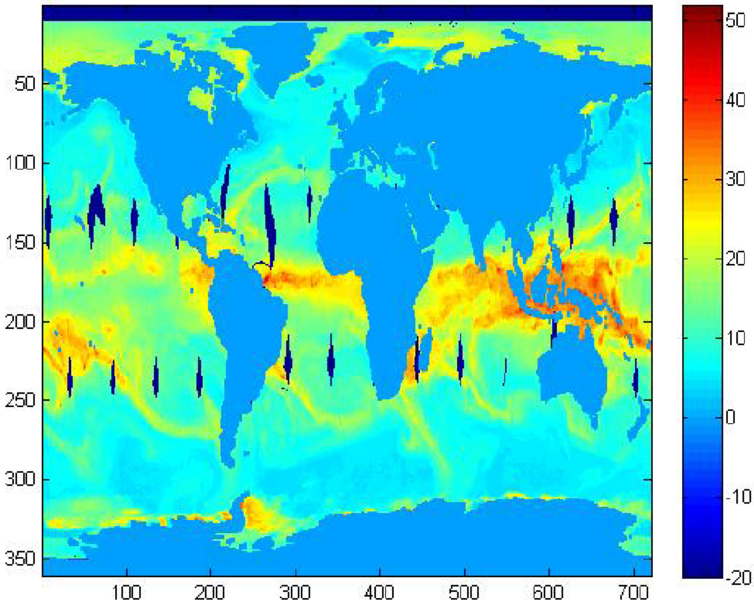


Figure 1: Water vapor field. 01.01.1999

So the essence of the task was to search for numeric patterns in this density array.

The preliminary analysis of data had given a very important fact. The variance of values for the first 7 years differs from the variance of values for the last 4 years (the algorithm of value construction had been changed by data provider), so we decided to use data from 01.01.1999 to 31.12.2005.

2.3. Research ways

We decided to divide research into 2 parts:

1. The research of time series in every discretization point;
2. Meridional analysis.

The main idea was to examine the results of wavelet decomposition on numeric patterns. All the algorithms were realized in Matlab using Wavelet Toolbox.

2.3.1. The research of time series

We had 3D-array where every 2D-layer is a density map. Fixing values of geographic coordinates we can extract time series for every point of discretization of the Earth's surface, so it had been done for data appropriate to surface of the oceans. An example of such time series is on the Figure 2.

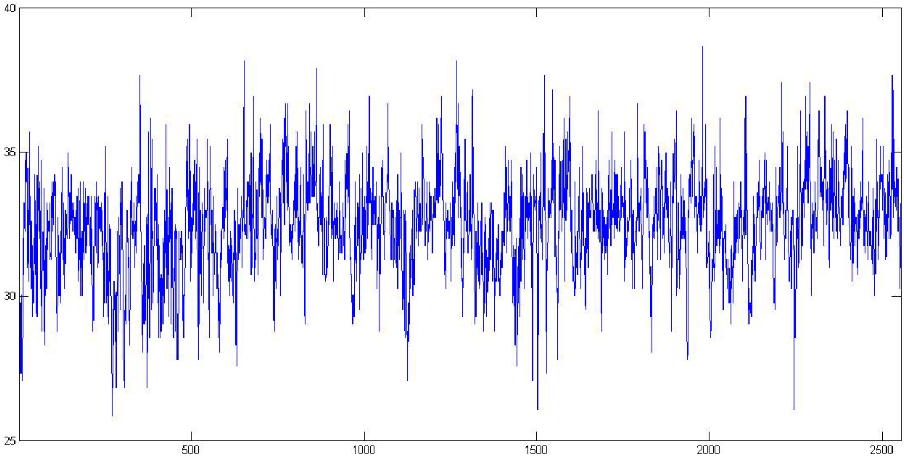


Figure 2: An example of time series (ox - time, oy - density)

In Matlab's Wavelet Toolbox we can analyze 1D arrays using both discrete and continuous wavelet transforms. But the second variant gives much more information about signal, because we can observe details of time series with arbitrary length instead of only the powers of 2. We had tested many wavelet families and for some of them found interesting patterns.

Results obtained with the use of wavelet Morlet (in terms of Matlab - "cmor1-1.5") we consider the most important. Wavelet Morlet is a complex function which can be written as

$$\psi(t) = e^{2\pi i t - \frac{t^2}{2}}.$$

Decomposition of every time series should be executed according to formula (1.4). On Figure 3 the result of 1D continuous complex wavelet transform for time series of point from Oceania is presented. This figure is built using Matlab's Wavelet Toolbox, so we can see specific Matlab notation for wavelet coefficients. For example, Ca, b means wavelet coefficient of continuous wavelet decomposition with scale parameter a and shift parameter b:

$$W_\psi[f](a, b) = Ca, b.$$

So we had such decomposition arrays for all available values of geographic coordinates. We had computed main frequencies for every pair (time series, values of scale parameter) using Fourier transform and for every value of scale parameter of decomposition generated a "frequency map". An example of such map is on the Figure 4. The area of 45° sl. — 45° nl. is presented because cyclonic activity is strong there.

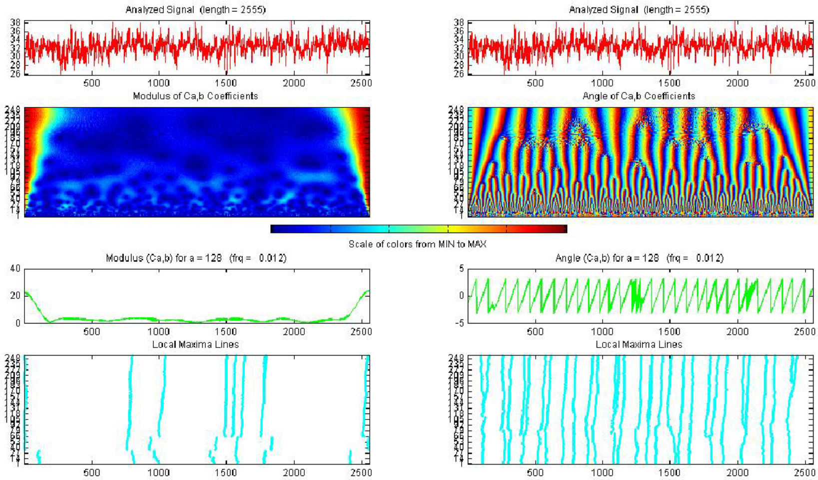


Figure 3: An example of continuous complex wavelet decomposition of time series which is presented on Figure 2

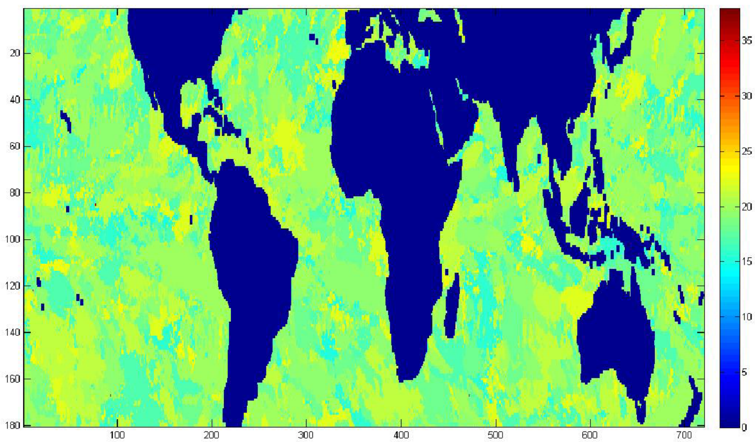


Figure 4: An example of "frequency map" (30-days activity)

Figure 4 demonstrates distribution of 30-days activity because the wavelet filter works with interval of coefficients of this length.

The staff of Institute of Space Researches had worked with this algorithm searching for physical interpretation of results represented on "frequency maps". They had found some new numeric patterns in the Earth's water vapor field such as subzones of variability of density of water vapor, known subzones had been localized better. Also all season effects and high day to day activity had been confirmed.

2.3.2. Meridional analysis

This part of researches is based on idea of joint analysis of data for every meridian. From original 3D array we extracted 2D array fixing the value of geographic longitude. So we had indexed images such as presented on Figure 5.

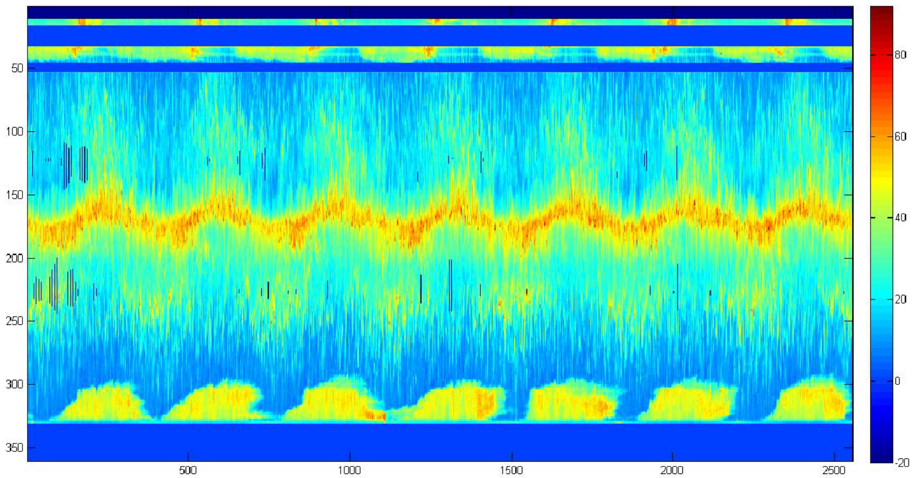


Figure 5: An example of distribution of density of water vapor during 1999-2005 for 21.5° el

According to strong variance of data from day to day we cannot use continuous wavelets. We will not reconstruct signal after decomposition, so we don't need any specific properties of wavelet basis in this context. Also it's comfortable when wavelet function is symmetric because transform results are unbiased. So we were using biorthogonal wavelets.

The physical interpretation of wavelet coefficients is rather simple. 2D discrete wavelet transform on every step of decomposition doubles the size of details it constructs. In our case it doubles them both by time and space. According to nature of selected basis we can conclude that the first level of decomposition consists of (day to day, 0,5° of latitude) details (for example the 5th level has details represented fluctuations during a month (approximately)). But we must take into account that there are 3 sorts of details: horizontal, vertical and diagonal. Horizontal details represent constant time nature of signal and difference by latitude. For vertical details we have an opposite situation. Diagonal details include space-and-time differences in signal.

There are many interesting facts we had found on decomposition images. In this article we are presented only some of them.

The first conclusion is based on the results of analysis of the first level of decompositions.We can say that day to day variability is stronger then interlatitudinal. Day to day activity is stronger in equatorial zone and mid-latitudes then in subtropics and circumpolar regions. Also it's easy to see that activity movings has seasonal component.

Another interesting fact we want to present in this article had been found on the 5th level of decomposition ("monthly details"). There is a strong difference between equatorial and subtropical monthly climatic activities. Similar result we see in circumpolar regions but we cannot be so sure in wavelet coefficients computed there because of edge effects.

2.4. Summary of current results

Distancing from facts listed above we can say that some relatively common results have already been achieved:

New mathematical methods for data processing has been proposed;

Using this methods the new concept "frequency map" has been introduced in the subject area;

Some numeric patterns for water vapor field have been discovered;

The hypothesis that "frequency maps" can help to predict atmospheric phenomena has been proposed.

2.5. Further plans

We have got many ideas about further work on this task and using of wavelet analysis in other research areas. We are working on the hypothesis of the relationship between water vapor field fluctuations and cyclonic activity. Many facts say that "frequency maps" can help us to make necessary mathematical tools for prediction of some atmospheric phenomena such as hurricanes. Some simple visual patterns were detected during comparison of sets of images such as Figures 6 and 7.

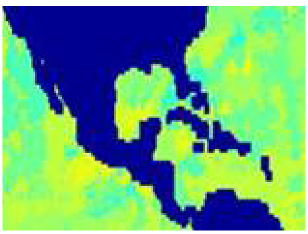


Figure 6: Element of frequency map

Figure 7: Trajectory of Catrina hurricane

So further researches are needed.

References

[1] Bogges A., Narkovich F. A. A First Course in Wavelets with Fourier Analysis. – Prentice Hall, 2001.

[2] Zakharova T. V., Shestakov O. V. Wavelet analysis. Study guide. – Max-Press, 2009 (in Russian).

[3] Mallat S. A wavelet tour of signal processing. – Academic Press, 1999.

[4] Smolentsev N. K. Basics of wavelet theory. Wavelets in Matlab. – DMK Press, 2005 (in Russian).

Annales Mathematicae et Informaticae
39 *(2012) pp. 173–191*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Renewal theorems in the case of attraction to the stable law with characteristic exponent smaller than unity*

S. V. Nagaev

Sobolev Institute of Mathematics, Novosibirsk
nagaev@math.nsc.ru, nagaevs@hotmail.com, nagaevs@academ.org

Dedicated to Mátyás Arató on his eightieth birthday

1. Introduction

Let X be a non-negative integer-valued random variable, $p_n = \mathbf{P}(X = n)$. Put $S_n = \sum_1^n X_j,\ n \geq 1$, where X_j are i.i.d. random variables which have the same distribution as X. In what follows we assume that $S_0 = 0$. Let $u_n = \sum_{k=0}^{\infty} \mathbf{P}(S_k = n)$ be the renewal probability at the instant n. Put $f(z) = \sum_{k=0}^{\infty} p_k z^k$. If $g(z)$ is an analytical function in some neighbourhood of zero, we denote the coefficient at z^n in Taylor series for $g(z)$ by $C_n(g(z))$.

In 1963 Garsia and Lamperti [1] proved that under the condition

$$\mathbf{P}(X > n) \sim L(n)n^{-\alpha}, \tag{1.1}$$

where $L(x)$ is a slowly-varying function, the asymptotic formula

$$u_n \sim \frac{\sin \pi\alpha}{\pi} L^{-1}(n)n^{\alpha-1}, \tag{1.2}$$

is valid, provided $1/2 < \alpha < 1$. The relation $a_n \sim b_n$ here and below indicates that $\lim\limits_{n\to\infty} a_n/b_n = 1$.

In 1968 Williamson [3] extended Garsia-Lamperti's result to the case that X belongs to the domain of attraction of a non-degenerate d - dimensional stable law with characteristic exponent α, $\ d/2 < \alpha < \min(d; 2)$.

To prove (1.2) Garsia and Lamperti used the purely analytical method based on analysis of behavior of the generating function $f(z)$ on the unit circle. On the

*This work was supported by RFBR 09-01-00048-a.

contrary, Williamson's approach is probabilistic with the local limit theorem by Rvacheva [4] as the starting point.

As to case $0 < \alpha \leq 1/2$, formula (1.2), generally speaking, is not true if we restrict our selves to condition (1.1). Corresponding counter-example is given in [3]. The point is that in the case $0 < \alpha \leq 1/2$ the existence of lacunas in the sequence p_n influences on the behavior of u_n. Therefore, additional constraints are necessary to provide the validity of (1.2). One such constraint was suggested by De Bruijn and Erdos [2] before [1] appeared, namely,

$$p_{n-1}p_{n+1} > p_n^2, \tag{1.3}$$

i.e. the sequence $\ln p_n$ is convex. Williamson [3] noticed that (1.2) remains true if the sequence p_n does not increase beginning with some number n. This condition is weaker than (1.3).

In the present work we use the condition

$$p_n \sim \frac{l(n)}{n^{1+\alpha}}, \ 0 < \alpha < 1, \tag{1.4}$$

where the function $l(x)$ is slowly varying. Notice that condition (1.1) with $L(n) = \alpha^{-1}l(n)$ follows from (1.4) (see Lemma 2.1 below). Condition (1.4) is discussed in our previous paper [5], namely, it is shown therein that if above-mentioned Williamson's condition is fulfilled, then (1.4) hold.

Theorem 1.1. *If condition (1.4) holds, then*

$$u_n \sim c(\alpha)\frac{\mathbf{P}(X = n)}{\mathbf{P}^2(X \geq n)} \sim \frac{\alpha^2 c(\alpha)}{l(n)n^{1-\alpha}}, \tag{1.5}$$

where $c(\alpha) = \sin \pi\alpha/\pi\alpha$.

The extreme case $p_n \sim n^{-1}l(n)$ is studied in [5]. It turns out that under this condition $u_n \sim \mathbf{P}(X = n)/\mathbf{P}^2(X \geq n)$. Since $c(\alpha) \to 1$ as $\alpha \to 0$, it implies that representation

$$u_n \sim c(\alpha)\frac{\mathbf{P}(X = n)}{\mathbf{P}^2(X \geq n)},$$

which is given in Theorem 1.1 is stable as $\alpha \to 0$. However, we can not say this about the relation $u_n \sim \alpha^2 c(\alpha)/l(n)n^{1-\alpha}$.

In proving Theorem 1.1 we apply the same approach as in [5]. However, to realize it was found more difficult in this case.

Remark. In [6] the renewal theorem is proved under condition that (1.1) holds and

$$p_n < c\mathbf{P}(X > n)n^{-1}$$

using Williamson's method. The proof is based on the following statement:
Assume that $F(0) = 0$ and (2.1) holds. Then for all $n \geq 1, z$ large enough and $x \geq z$

$$\mathbf{P}\{S_n \geq x, M_n \leq z\} \leq \{cz/x\}^{x/z},$$

where $M_n = \max\{X_1, X_2, ..., X_n\}$ *and* $S_n = \sum\limits_1^n X_i$ (see Lemma 2 in [6]).

The author of [6] asserts that this lemma is an immediate consequence of the inequality

$$\mathbf{P}(S_n \geq x) \leq \sum_{i=1}^{n} \mathbf{P}(X_i > y_i) + (eA_t^+/xy^{t-1})^{x/y},$$

where $S_n = \sum\limits_{j=1}^{n} X_j$, X_j are independent random variables, $y > \max\limits_i y_i$, $A_t^+ = \sum\limits_{j=1}^{n} \{X_j^t; X_j > 0\}$, $0 < t < 1$ (see Corollary 1.5 in [7]).

If X_j are i.i.d. equal to X by distribution, then

$$\mathbf{P}(S_n \geq x) \leq n\mathbf{P}(X > y) + \left(\frac{en\mathbf{E}\{X^t; X > 0\}}{xy^{t-1}} \right)^{x/y}.$$

If $X \leq y$, then

$$\mathbf{E}\{X^t; X > 0\} \leq y^t.$$

Consequently, in this case

$$\mathbf{P}(S_n \geq x) \leq \left(\frac{eny}{x} \right)^{x/y}.$$

This inequality differs from the inequality stated in [6] by the presence of n in the right-hand side. Thus, Lemma 2 of [6] does not follow from Corollary 1.5 of [7]), and, therefore, the former can not be considered as being proved.

Let $h_n = \sum\limits_{k=0}^{\infty} n^{-1}\mathbf{P}(S_k = n)$.

Theorem 1.2. *If condition (1.4) holds, then*

$$h_n \sim \frac{\alpha}{n}. \tag{1.6}$$

Notice that h_n is the derivative of the measure $\nu(A) := \sum\limits_{k\in A} h_k$ with respect to the counting measure. The measure $\nu(A)$ is a particular case of so called harmonic renewal measure. Recall that that the measure $\nu(\cdot) = \sum\limits_1^{\infty} n^{-1}F_n(\cdot)$, where F_n is n-th convolution of any distribution F on $\mathbf{R}^+$ is said to be harmonic renewal measure associated with F. In our case the distribution F is concentrated on the lattice of non-negative integers.The harmonic renewal function is defined by the equality $H(x) = \nu([0, x))$.

The next statement concerning the asymptotic behavior of $H(n)$ as $n \to \infty$ follows from Theorem 1.2.

Corollary 1.3. *If condition (1.4) holds, then*

$$H(n) \sim \alpha \ln n. \tag{1.7}$$

The asymptotic behavior of $H(x)$ for $x \to \infty$ is studied in [9, 10, 11, 12]. The case that F attracts to a stable law is considered in [9], namely , it is proved therein that under the condition $1 - F(x) \sim x^{-\alpha} L(x)$

$$\lim_{x\to\infty} (H(x) - \alpha \ln x + \ln L(x)) = \alpha \mathbf{C} - \ln \Gamma(1-\alpha),$$

where $\mathbf{C}$ is the Euler constant, $\Gamma(\cdot)$ is the gamma function. Of course, the last assertion is sharper than (1.7). Formula (1.7) is presented by reason of simplicity of proving.

2. Auxiliary results

Lemma 2.1. *For any* $\alpha > 0$

$$\sum_{k=n}^{\infty} \frac{l(k)}{k^{\alpha+1}} \sim \int_{n}^{\infty} \frac{l(y)}{y^{\alpha+1}} dy. \tag{2.1}$$

Proof. Put $p(x) = l(x)/x^{\alpha+1}$. Obviously,

$$\inf_{n\le y\le n+1} \frac{p(y)}{p(n)} \le \frac{1}{p(n)} \int_{n}^{n+1} p(y)dy \le \sup_{n\le y\le n+1} \frac{p(y)}{p(n)}. \tag{2.2}$$

It is easily seen that for every $n \le y \le n+1$

$$\left(\frac{n}{n+1}\right)^{\alpha+1} \inf_{n\le y\le n+1} \frac{l(y)}{l(n)} \le \frac{p(y)}{p(n)} \le \sup_{n\le y\le n+1} \frac{l(y)}{l(n)}. \tag{2.3}$$

In what follows we need Kamarata's representation

$$l(x) = a(x) \exp\left\{ \int_{1}^{x} \frac{\epsilon(u)}{u} du \right\},\ x \ge 1, \tag{2.4}$$

where $\lim_{n\to\infty} \varepsilon(u) = 0,\ \lim_{x\to\infty} a(x) = a,\ 0 < a < \infty$. Hence,

$$\frac{l(y)}{l(n)} = \frac{a(y)}{a(n)} \exp\left\{ \int_{n}^{y} \frac{\epsilon(u)}{u} du \right\}.$$

Obviously,

$$\lim_{n\to\infty} \sup_{n\le y\le n+1} \left| \int_{n}^{y} \frac{\varepsilon(u)}{u} du \right| = 0.$$

It follows from last two relations that

$$\lim_{n\to\infty} \sup_{n\le y\le n+1} \left| \frac{l(y)}{l(n)} - 1 \right| = 0. \tag{2.5}$$

Combining (2.2), (2.3) and (2.5), we have

$$\lim_{n\to\infty} \frac{1}{p(n)} \int_n^{n+1} p(y)dy = 1. \tag{2.6}$$

It is easily seen that

$$\inf_{k\ge n} \frac{1}{p(k)} \int_k^{k+1} p(y)dy \le \frac{\int_n^\infty p(y)dy}{\sum_{k=n}^\infty p(k)} \le \sup_{k\ge n} \frac{1}{p(k)} \int_k^{k+1} p(y)dy. \tag{2.7}$$

The conclusion of the Lemma follows from (2.6) and (2.7). □

Lemma 2.2. *For any* $\alpha > 0$

$$\int_x^\infty \frac{l(y)}{y^{\alpha+1}}dy \sim \frac{l(x)}{\alpha x^\alpha}. \tag{2.8}$$

Proof. By using (2.4), we have

$$\int_x^\infty \frac{l(y)}{y^{\alpha+1}}dy \sim \int_x^\infty \frac{l_0(y)}{y^{\alpha+1}}dy, \tag{2.9}$$

where

$$l_0(y) = \exp\left\{ \int_1^y \frac{\varepsilon(u)}{u}du \right\}. \tag{2.10}$$

Integrating by parts, we conclude that

$$\begin{aligned} \int_x^\infty \frac{l_0(y)}{y^{\alpha+1}}dy &= \frac{l_0(x)}{\alpha x^\alpha} + \frac{1}{\alpha}\int_x^\infty \frac{\varepsilon(u)l_0(y)}{y^{\alpha+1}}dy \\ &= \frac{l_0(x)}{\alpha x^\alpha}(1+o(1)) = \frac{l(x)}{\alpha x^\alpha}(1+o(1)). \end{aligned} \tag{2.11}$$

The desired result follows from (2.9) and (2.11). □

Note that (2.8) can be deduced from the asymptotic formula

$$\int_{\alpha}^{\infty} f(t)l(xt)dt \sim l(x)\int_{\alpha}^{\infty} f(t)dt,$$

where $\alpha > 0$, and $f(t)t^{\eta}$, $\eta > 0$, is integrable (see [8], Theorem 2.6), but not immediately. For this purpose one needs to make the change of variables $y = xt$ in the integral $\int_x^{\infty} y^{-\alpha-1}l(y)dy$. On the other hand, the method which is used in proving Lemma 2.2 allows to obtain very easily the statement the above mentioned Theorem 2.6 of [8].

Corollary 2.3. *Under condition (1.4)*

$$\mathbf{P}(X \geq n) \sim \frac{l(n)}{\alpha n^{\alpha}}. \tag{2.12}$$

Proof. Evidently,

$$\inf_{k\geq n} \frac{l(k)}{k^{\alpha+1}p_k} \leq \frac{\sum\limits_{k=n}^{\infty} l(k)k^{-\alpha-1}}{\sum\limits_{k=n}^{\infty} p_k} \leq \sup_{k\geq n} \frac{l(k)}{k^{\alpha+1}p_k}.$$

Hence, by (2.7)

$$\mathbf{P}(X \geq n) = \sum_{k\geq n} p_k \sim \sum_{k\geq n} \frac{l(k)}{k^{\alpha+1}} \sim \frac{l(n)}{\alpha n^{\alpha}},$$

which was to be proved. □

Lemma 2.4. *For any $\alpha < 1$*

$$\sum_{k=1}^{n} \frac{l(k)}{k^{\alpha}} \sim \frac{l(n)}{1-\alpha} n^{1-\alpha}. \tag{2.13}$$

Proof. Let $l_0(x)$ be defined by (2.10). Since $l_0(x) \sim l(x)$,

$$\sum_{k=1}^{n} \frac{l_0(k)}{k^{\alpha}} \sim \sum_{k=1}^{n} \frac{l(k)}{k^{\alpha}}. \tag{2.14}$$

Indeed,

$$1-\varepsilon < \frac{\sum\limits_{k=n(\varepsilon)}^{n} k^{-\alpha}l_0(k)}{\sum\limits_{k=n(\varepsilon)}^{n} k^{-\alpha}l(k)} < 1+\varepsilon$$

if $n(\varepsilon)$ is such that for $x > n(\varepsilon)$

$$1-\varepsilon < \frac{l_0(x)}{l(x)} < 1+\varepsilon.$$

It is easily seen that

$$\lim_{n\to\infty} \sum_{k=n(\epsilon)}^{n} k^{-\alpha} l(k) = \infty.$$

Therefore for sufficiently large n

$$1-2\varepsilon < \frac{\sum\limits_{k=n(\varepsilon)}^{n} k^{-\alpha} l_0(k)}{\sum\limits_{k=n(\varepsilon)}^{n} k^{-\alpha} l(k)} < 1+2\varepsilon.$$

Since ε can be made as small as we wish, hence the validity of (2.14) follows. By applying the Abel transform, we get

$$\sum_{k=1}^{n} \frac{l_0(k)}{k^\alpha} = l_0(n)\sum_{k=1}^{n} k^{-\alpha} + \sum_{k=1}^{n-1}(l_0(k)-l_0(k+1))\sum_{j=1}^{k} j^{-\alpha}. \tag{2.15}$$

It is easily seen that

$$l_0(k) - {}_0(k+1) = l_0(k)\left(1-\exp\left\{\int\limits_{k}^{k+1} \frac{\varepsilon(u)}{u} du\right\}\right).$$

Hence

$$|l_0(k) - {}_0(k+1)| < l_0(k)\left|\int\limits_{k}^{k+1} \frac{\varepsilon(u)}{u} du\right| = o(l_0(k)k^{-1}). \tag{2.16}$$

Further,

$$\sum_{k=1}^{n} k^{-\alpha} \sim \frac{n^{1-\alpha}}{1-\alpha}. \tag{2.17}$$

It follows from (2.16) and (2.17)

$$\sum_{k=1}^{n-1}(l_0(k)-l_0(k+1))\sum_{j=1}^{k} j^{-\alpha} = o\left(\sum_{k=1}^{n} \frac{l_0(k)}{k^\alpha}\right). \tag{2.18}$$

Combining (2.15)–(2.17), we conclude that

$$\sum_{k=1}^{n} l_0(k)k^{-\alpha} \sim \frac{l_0(n)}{1-\alpha} n^{1-\alpha}. \tag{2.19}$$

From (2.14) and (2.19) the result follows. □

Corollary 2.5. *Under conditions of Theorem 1.1*

$$\sum_{k=1}^{n}\mathbf{P}(X \geq k) \sim \frac{l(n)}{\alpha(1-\alpha)}n^{1-\alpha}. \tag{2.20}$$

Proof. According to Corollary 2.3 for any $\varepsilon > 0$ there exists $n(\varepsilon)$ such that for $n > n(\varepsilon)$

$$1-\varepsilon < \mathbf{P}(X \geq n)\Big/\frac{l(n)}{\alpha n^{\alpha}} < 1+\varepsilon.$$

Hence,

$$1-\varepsilon < \sum_{n(\varepsilon)<k\leq n}\mathbf{P}(X \geq k)/\alpha^{-1}\sum_{n(\varepsilon)<k\leq n}\frac{l(k)}{k^{\alpha}} < 1+\varepsilon.$$

On the other hand, since

$$\lim_{n\to\infty}\sum_{n(\varepsilon)<k\leq n}\frac{l(k)}{k^{\alpha}} = \infty$$

for every $\varepsilon > 0$,

$$\sum_{n(\varepsilon)<k\leq n}\frac{l(k)}{k^{\alpha}} \sim \sum_{k=1}^{n}\frac{l(k)}{k^{\alpha}}, \quad \sum_{n(\varepsilon)<k\leq n}\mathbf{P}(X \geq k) \sim \sum_{k=1}^{n}\mathbf{P}(X \geq k).$$

Therefore, for sufficiently large n

$$1-2\varepsilon < \alpha\sum_{k=1}^{n}\mathbf{P}(X \geq k)/\sum_{k=1}^{n}\frac{l(k)}{k^{\alpha}} < 1+2\varepsilon.$$

Hence, since ε is arbitrary, it follows that

$$\sum_{k=1}^{n}\mathbf{P}(X \geq k) \sim \alpha^{-1}\sum_{k=1}^{n}\frac{l(k)}{k^{\alpha}}.$$

To complete the proof it remains to apply Lemma 2.4. □

Lemma 2.6. *Under conditions of Theorem 1.1*

$$1-f(z) \sim (1-z)^{\alpha}L\left(\frac{1}{1-z}\right), \tag{2.21}$$

where

$$L(x) = \frac{\Gamma(1-\alpha)}{\alpha}l(x).$$

Proof. First of all,

$$\sum_{k=0}^{n} \mathbf{P}(X > k) z^k = \frac{1 - f(z)}{1 - z}.$$

It is easily seen that

$$\mathbf{P}(X > k) \sim \mathbf{P}(X \geq k).$$

Now, using Corollary 2.5 and the Abelian theorem (see, e.g. [13], Ch. XIII, section 5, Th. 5), we have

$$\begin{aligned} \frac{1 - f(z)}{1 - z} &\sim \frac{\Gamma(2-\alpha)}{\alpha(1-\alpha)}(1-z)^{\alpha-1} L(1-z) \\ &= \alpha^{-1}\Gamma(1-\alpha)(1-z)^{\alpha-1} l\left(\frac{1}{1-z}\right) = (1-z)^{\alpha-1} L\left(\frac{1}{1-z}\right), \end{aligned}$$

which is equivalent to the assertion of the Lemma. □

Lemma 2.7. *Under conditions of Theorem 1.1*

$$\sum_{k=0}^{n} u_k \sim \frac{n^\alpha}{\Gamma(\alpha+1)L(n)}, \tag{2.22}$$

where $L(x)$ is defined in Lemma 2.6.

Proof. Obviously,

$$u_k = C_k\left(\frac{1}{1 - f(z)}\right).$$

Applying Lemma 2.6 and the Tauberian theorem (see ref. in the proof of Lemma 2.6), we obtain the desired result. □

The next assertion is borrowed from [5].

Lemma 2.8. *The identity*

$$n u_n = \sum_{k=0}^{n-1} (n-k) p_{n-k} u_k^{(2)} \tag{2.23}$$

holds, where $u_n = \sum\limits_{k=0}^{\infty} \mathbf{P}(S_k = n)$, $u_n^{(2)} = \sum\limits_{k=0}^{n} u_{n-k} u_k$.

Lemma 2.9. *Under condition of Theorem 1.1 there exists the sequence θ_n such that $\lim\limits_{n\to\infty} \theta_n = 1$ and*

$$u_n^{(2)} \leq \frac{2^{1-\alpha}\theta_n n^\alpha}{\Gamma(\alpha+1)L(n)} \max_{n/2 \leq k \leq n} u_k. \tag{2.24}$$

Proof. It is easily seen that

$$u_n^{(2)} \le 2 \sum_{0\le k\le n/2} u_k u_{n-k} \le 2 \max_{n/2\le k\le n} u_k \sum_{0\le k\le n/2} u_k.$$

To complete the proof it is sufficient to apply Lemma 2.7. □

Lemma 2.10. *Under conditions of Theorem 1.1*

$$\sum_{k=1}^{n} u_k^{(2)} \sim \frac{n^{2\alpha}}{\Gamma(2\alpha+1)L^2(n)}, \tag{2.25}$$

where $L(x)$ is defined in Lemma 2.6.

Proof. It is easily seen that

$$u_k^{(2)} = C_k\left(\frac{1}{(1-f(z))^2}\right).$$

According to Lemma 2.6

$$(1-f(z))^{-2} \sim (1-z)^{-2\alpha} L^{-2}\left(\frac{1}{1-z}\right).$$

Applying the Tauberian theorem (see ref. in the proof of Lemma 2.6), we get the desired result. □

Lemma 2.11. *Under conditions of Theorem 1.1 for every fixed $0 < a < b < 1$*

$$\sum_{na\le k\le nb} l^{-2}(k)k^{2\alpha-1}(n-k)^{-\alpha} \sim l^{-2}(n)n^{\alpha}\int_a^b u^{2\alpha-1}(1-u)^{-\alpha}du. \tag{2.26}$$

Proof. First of all, notice that

$$\ln\frac{l_0(n)}{l_0(k)} = \int_k^n \frac{\varepsilon(u)}{u}du. \tag{2.27}$$

Consequently,

$$\lim_{n\to\infty} \sup_{na\le k\le nb}\left|\frac{l_0(n)}{l_0(k)} - 1\right| = 0. \tag{2.28}$$

This implies that

$$\sum_{na\le k\le nb} l_0^{-2}(k)k^{2\alpha-1}(n-k)^{-\alpha} \sim l_0^{-2}(n)\sum_{na\le k\le nb} k^{2\alpha-1}(n-k)^{-\alpha}.$$

Hence it follows that

$$\sum_{na\le k\le nb} l^{-2}(k)k^{2\alpha-1}(n-k)^{-\alpha} \sim l^{-2}(n) \sum_{na\le k\le nb} k^{2\alpha-1}(n-k)^{-\alpha}.$$

Further,

$$\sum_{na\le k\le nb} k^{2\alpha-1}(n-k)^{-\alpha} = n^{\alpha-1} \sum_{na\le k\le nb} \left(\frac{k}{n}\right)^{2\alpha-1}\left(1-\frac{k}{n}\right)^{-\alpha}$$
$$\sim n^{\alpha}\int\limits_a^b u^{2\alpha-1}(1-u)^{-\alpha}du.$$

The result follows from last two relations. □

3. The proof of Theorem 1.1

Let us write down formula (2.23) in the form

$$\begin{aligned} nu_n &= \Bigg(\sum_{0\le k<\sqrt{n}} + \sum_{\sqrt{n}\le k\le(1-\eta)n} + \sum_{(1-\eta)n<k\le n}\Bigg)(n-k)p_{n-k}u_k^{(2)} \\ &\equiv \sum\nolimits_1 + \sum\nolimits_2 + \sum\nolimits_3, \end{aligned} \tag{3.1}$$

where $0<\eta<1$. For any $\varepsilon>0$, sufficiently large n, and $k<\sqrt{n}$

$$p_{n-k} < (1+\varepsilon)\frac{l(n-k)}{(n-\sqrt{n})^{\alpha+1}}. \tag{3.2}$$

If $n-\sqrt{n}\le k\le n$, then

$$\frac{l_0(n)}{l_o(k)} = \exp\Bigg\{\int\limits_k^n \frac{\varepsilon(u)}{u}du\Bigg\} = 1+o(\ln n - \ln(n-\sqrt{n})) = 1+o\Bigg(\frac{1}{\sqrt{n}}\Bigg).$$

Consequently,

$$\max_{n-\sqrt{n}\le k\le n} l_0(k) \sim l_0(n). \tag{3.3}$$

It follows from (3.2) and (3.3) that

$$\sum\nolimits_1 = O\Bigg(\frac{l(n)}{n^{\alpha}}\sum_{k=1}^{[\sqrt{n}]} u_k^{(2)}\Bigg).$$

By Lemma 2.10

$$\sum_{k=1}^{[\sqrt{n}]} u_k^{(2)} = O\biggl(\frac{n^\alpha}{l^2(\sqrt{n})}\biggr). \tag{3.4}$$

Thus,

$$\sum\nolimits_1 = O\biggl(\frac{1}{l(\sqrt{n})}\biggr). \tag{3.5}$$

Let us turn to estimating $\sum_2$. It is easily seen that

$$\sum\nolimits_2 \sim \sum_{\sqrt{n}\le k\le(1-\eta)n} u_k^{(2)} \frac{l_0(n-k)}{(n-k)^\alpha} \equiv \sum\nolimits_4 . \tag{3.6}$$

Applying Abel's transformation, we have

$$\begin{aligned}\sum\nolimits_4 \sim \frac{l_0(n-\sqrt{n})}{(n-\sqrt{n})^\alpha} \sum_{\sqrt{n}\le k\le(1-\eta)n} u_k^{(2)} \\ - \sum_{\sqrt{n}\le k\le(1-\eta)n} \biggl(\frac{l_0(n-k-1)}{(n-k-1)^\alpha} - \frac{l_0(n-k)}{(n-k)^\alpha}\biggr) \sum_{j=[\sqrt{n}]}^{k} u_j^{(2)}.\end{aligned} \tag{3.7}$$

By Lemma 2.10

$$\sum_{\sqrt{n}\le k\le(1-\eta)n} u_k^{(2)} = \sum_{k\le(1-\eta)n} u_k^{(2)} - \sum_{k<\sqrt{n}} u_k^{(2)} \sim \frac{(1-\eta)^{2\alpha} n^{2\alpha}}{\Gamma(2\alpha+1)L^2(n)}. \tag{3.8}$$

Further,

$$\frac{l_0(k)}{k^\alpha} - \frac{l_0(k+1)}{(k+1)^\alpha} = l_0(k)\biggl(\frac{1}{k^\alpha} - \frac{1}{(k+1)^\alpha}\biggr) + \frac{l_0(k) - l_0(k+1)}{(k+1)^\alpha}. \tag{3.9}$$

Obviously,

$$\frac{1}{k^\alpha} - \frac{1}{(k+1)^\alpha} \sim \frac{\alpha}{k^{\alpha+1}}. \tag{3.10}$$

On the other hand,

$$\begin{aligned} l_0(k+1) - l_0(k) &= l_0(k)\biggl(\frac{l_0(k+1)}{l_0(k)} - 1\biggr) \\ &= l_0(k)\biggl(\exp\biggl\{\int_k^{k+1} \frac{\varepsilon(u)}{u} du\biggr\} - 1\biggr) = o\biggl(\frac{l_0(k)}{k}\biggr).\end{aligned} \tag{3.11}$$

It follows from (3.9)–(3.11) that

$$\frac{l_0(k)}{k^\alpha} - \frac{l_0(k+1)}{(k+1)^\alpha} \sim \frac{\alpha l_0(k)}{k^{\alpha+1}}. \tag{3.12}$$

Combining (3.6)–(3.8) and (3.12), we obtain

$$\begin{aligned}
\sum\nolimits_2 &\sim \frac{(1-\eta)^{2\alpha} l_0(n) n^\alpha}{\Gamma(2\alpha+1)L^2(n)} - \alpha \sum_{\sqrt{n}\le k\le(1-\eta)n} \frac{l_0(n-k)}{(n-k)^{\alpha+1}} \sum_{j=[\sqrt{n}]}^{k} u_j^{(2)} \\
&= \frac{(1-\eta)^{2\alpha}\alpha n^\alpha}{\Gamma(1-\alpha)\Gamma(2\alpha+1)a(n)L(n)} \\
&- \alpha \sum_{\sqrt{n}\le k\le(1-\eta)n} \frac{l_0(n-k)}{(n-k)^{\alpha+1}} \sum_{j=0}^{k} u_j^{(2)} + \alpha \sum_{j=0}^{\sqrt{[n]}-1} u_j^{(2)} \sum_{\sqrt{n}\le k\le(1-\eta)n} \frac{l_0(n-k)}{(n-k)^{\alpha+1}} \\
&= \frac{(1-\eta)^{2\alpha}\alpha n^\alpha}{\Gamma(1-\alpha)\Gamma(2\alpha+1)a(n)L(n)} - \alpha \sum\nolimits_5 + \alpha \sum\nolimits_6 .
\end{aligned} \tag{3.13}$$

Here $a(\cdot)$ is a factor in Karamata's representation (2.4) for $l(x)$. In view of (3.4)

$$\sum\nolimits_6 = O\biggl(\frac{l_0(n)}{l_0^2(\sqrt{n})}\biggr). \tag{3.14}$$

We now proceed to estimating $\sum_5$. By Lemma 2.10

$$\sum\nolimits_5 \sim c(\alpha) \sum_{\sqrt{n}\le k\le(1-\eta)n} L^{-2}(k) k^{2\alpha} \frac{l_0(n-k)}{(n-k)^{\alpha+1}} \equiv c(\alpha) \sum\nolimits_7, \tag{3.15}$$

where $c(\alpha) = 1/\Gamma(2\alpha+1)$. Applying the Abel transformation, we have

$$\begin{aligned}
\sum\nolimits_7 &\sim L^{-2}(n)(1-\eta)^{2\alpha} n^{2\alpha} \sum_{\sqrt{n}\le k\le(1-\eta)n} \frac{l_0(n-k)}{(n-k)^{\alpha+1}} \\
&- \sum_{\sqrt{n}\le k\le(1-\eta)n} (L^{-2}(k+1)(k+1)^{2\alpha} - L^{-2}(k)k^{2\alpha}) \sum_{j=[\sqrt{n}]}^{k} \frac{l_0(n-j)}{(n-j)^{\alpha+1}}.
\end{aligned} \tag{3.16}$$

In the same way as (3.12) we deduce that

$$L^{-2}(k+1)(k+1)^{2\alpha} - L^{-2}(k)k^{2\alpha} \sim 2\alpha L^{-2}(k)k^{2\alpha-1}.$$

Hence, denoting the second summand in (3.16) by $\sum_8$, we obtain

$$\sum\nolimits_8 \sim 2\alpha \sum_{\sqrt{n}\le k\le(1-\eta)n} L^{-2}(k)k^{2\alpha-1} \sum_{j=[\sqrt{n}]}^{k} \frac{l_0(n-j)}{(n-j)^{\alpha+1}}$$

$$\sim 2\alpha l_0(n) \sum_{\sqrt{n}\le k\le(1-\eta)n} L^{-2}(k)k^{2\alpha-1} \sum_{j=[\sqrt{n}]}^{k} (n-j)^{-\alpha-1}. \tag{3.17}$$

It is not difficult to check that for $\sqrt{n} \le k \le (1-\eta)n$

$$\alpha \sum_{j=\lfloor\sqrt{n}\rfloor} (n-j)^{-\alpha-1} = (n-k)^{-\alpha} - n^{-\alpha} + o(n^{-\alpha}).$$

Consequently,

$$\sum\nolimits_8 + 2n^{-\alpha} \sum_{\sqrt{n}\le k\le(1-\eta)n} L^{-2}(k)k^{2\alpha-1} \sim 2l_0(n) \sum_{\sqrt{n}\le k\le(1-\eta)n} L^{-2}(k)k^{2\alpha-1}(n-k)^{-\alpha}. \tag{3.18}$$

We need the identity

$$\begin{aligned}\sum_{\sqrt{n}\le k\le(1-\eta)n} &= \Bigg(\sum_{\sqrt{n}\le k<\varepsilon n} + \sum_{\varepsilon n\le k\le(1-\eta)n}\Bigg) L^{-2}(k)k^{2\alpha-1}(n-k)^{-\alpha} \\ &\equiv \sum\nolimits_9 + \sum\nolimits_{10}.\end{aligned} \tag{3.19}$$

It is easily seen that

$$\sum\nolimits_9 < (1-\varepsilon)^{-\alpha} n^{-\alpha} \sum_{\sqrt{n}\le k\le\varepsilon n} L^{-2}(k)k^{2\alpha-1}.$$

By using Lemma 2.4, we obtain that

$$\sum_{\sqrt{n}\le k\le\varepsilon n} L^{-2}(k)k^{2\alpha-1} \sim \frac{(\varepsilon n)^{2\alpha}}{2\alpha L^2(n)}.$$

Therefore, for sufficiently large n

$$\sum\nolimits_9 < (1-\varepsilon)^{-\alpha} \frac{\varepsilon^{2\alpha} n^{\alpha}}{2\alpha L^2(n)}. \tag{3.20}$$

On the other hand, by Lemma 2.11

$$\sum\nolimits_{10} \sim L^{-2}(n)n^{\alpha} \int_{\varepsilon}^{1-\eta} u^{2\alpha-1}(1-u)^{-\alpha}du. \tag{3.21}$$

It follows from (3.18) – (3.21) that

$$\sum\nolimits_8 + \frac{(1-\eta)^{2\alpha} n^{\alpha}}{\alpha L^2(n)} \sim \frac{2\alpha^2 n^{\alpha}}{\Gamma^2(1-\alpha)l(n)} \int_0^{1-\eta} u^{2\alpha-1}(1-u)^{-\alpha}du. \tag{3.22}$$

Combining (3.15), (3.16), (3.18) and (3.22) we obtain

$$\sum\nolimits_5 \sim \frac{(1-\eta)^{2\alpha}\alpha n^\alpha}{\Gamma(1-\alpha)\Gamma(2\alpha+1)L(n)} - \frac{2\alpha^2 n^\alpha}{\Gamma^2(1-\alpha)\Gamma(2\alpha+1)l(n)} I(\eta), \tag{3.23}$$

where $I(\eta) = \int\limits_0^{1-\eta} u^{2\alpha-1}(1-a)^{-\alpha}du$. Finally, it follows from (3.13), (3.14) and (3.23) that

$$\sum\nolimits_2 \sim \frac{2\alpha^3 n^\alpha}{\Gamma^2(1-\alpha)\Gamma(2\alpha+1)l(n)} I(\eta). \tag{3.24}$$

We now turn to estimating $\sum_3$. Evidently,

$$\sum\nolimits_3 < \max_{(1-\eta)n<k\le n} u_k^{(2)} \sum_{(1-\eta)n<k\le n} (n-k)p_{n-k}.$$

By Lemma 2.4

$$\sum_{(1-\eta)n<k\le n} (n-k)p_{n-k} \sim \sum_1^{[\eta n]} \frac{l(j)}{j^\alpha} \sim \frac{l(n)}{1-\alpha}(\eta n)^{1-\alpha}.$$

On the other hand, in view of (2.24)

$$\max_{(1-\eta)n<k\le n} u_k^{(2)} < \frac{2^{1-\alpha}n^\alpha}{\Gamma(\alpha+1)} \max_{(1-\eta)n<k\le n} \frac{\theta_k}{L(k)} \max_{(1-\eta)n/2\le j\le n} u_j.$$

As a result we obtain that

$$\sum\nolimits_3 = n\psi(n)(2\eta)^{1-\alpha} \max_{\delta n\le j\le n} u_j, \tag{3.25}$$

where

$$\psi(n) = \frac{\alpha b_n}{\Gamma(\alpha+1)\Gamma(1-\alpha)(1-\alpha)}, \quad 0 < \limsup_{n\to\infty} b_n \le 1, \quad \delta = \frac{1-\eta}{2}.$$

Notice that

$$\frac{\alpha}{\Gamma(\alpha+1)\Gamma(1-\alpha)} = \frac{1}{\Gamma(\alpha)\Gamma(1-\alpha)} = \frac{\sin\pi\alpha}{\pi}$$

(see [14], formula 8.334, 3). Consequently,

$$\psi(n) = \frac{\sin\pi\alpha}{(1-\alpha)\pi} b_n. \tag{3.26}$$

It follows from (3.1), (3.5), (3.24) and (3.25) that

$$u_n = \varphi(n) + (2\eta)^{1-\alpha}\psi(n) \max_{\delta n\le j\le n} u_j, \tag{3.27}$$

where

$$\varphi(n) = \frac{2\alpha^3 a_n n^{\alpha-1} I(\eta)}{\Gamma^2(1-\alpha)\Gamma(2\alpha+1)l(n)}, \ a_n \sim 1.$$

Let us fix $0 < \varepsilon < 1/2$. Let η be such that $(2\eta)^{1-\alpha} < \varepsilon$. Chose N so that $\psi(n) < 1$ for $n > N$. Let n_1 be the value of k for which $\max\limits_{\delta n \le k \le n} u_k$ is attained. In particular, it may be that $n_1 = n$. In this case $u_n < \varphi(n)/(1-\varepsilon)$. If $N < n_1 < n$, then

$$u_{n_1} < \varphi(n_1) + \varepsilon \max_{\delta n_1 \le j \le n_1} u_j$$

and consequently

$$u_n < \varphi(n) + \varepsilon\varphi(n_1) + \varepsilon^2 \max_{\delta n_1 \le j \le n_1} u_j. \tag{3.28}$$

If $\max\limits_{\delta n_1 \le j \le n_1} u_j = u_{n_1}$, then $u_{n_1} < \varphi(n_1)/(1-\varepsilon)$. Substituting this bound in (3.28), we have

$$u_n < \varphi(n) + \varepsilon\varphi(n_1) + \frac{\varepsilon^2}{1-\varepsilon}\varphi(n_1).$$

If $\max\limits_{\delta n_1 \le j \le n_1} u_j$ is attained for $N < j < n_1$, then, similarly, the following inequality is deduced

$$u_n < \varphi(n) + \varepsilon\varphi(n_1) + \varepsilon^2\varphi(n_2) + \frac{\varepsilon^3}{1-\varepsilon} \max_{\delta n_2 \le j \le n_2} u_j$$

and so forth.

There exist two possibilities: either for some $n_k > N$

$$\max_{\delta n_k \le j \le n_k} u_j = u_{n_k},$$

or for some $k = k_0$ the inequality $n_k < N$ is fulfilled. Consider the first case. First of all, notice that $n_k \ge \delta^k n$. Using Karamata's representation (2.4) for $l(n)$, we obtain

$$\frac{\varphi(n_j)}{\varphi(n)} = \frac{a_n a(n)}{a_{n_j} a(n_j)} \left(\frac{n}{n_j}\right)^{1-\alpha} \exp\left\{ - \int_{n_j}^{n} \frac{\varepsilon(u)}{u} \right\}.$$

Evidently,

$$\left| \int_{n_j}^{n} \frac{\varepsilon(u)}{u} du \right| < \sup_{n_j \le u \le n} |\varepsilon(u)| \ln \frac{n}{n_j} < -j\gamma \ln\delta, \ \gamma = \sup_{u > N} |\varepsilon(u)|.$$

Consequently, there exists ε_0 such that for $\varepsilon < \varepsilon_0$

$$\varepsilon^j \varphi(n_j) < \varepsilon^j \varphi(n) \exp\left\{ j\gamma \ln 2 \right\} < \varepsilon^{j/2}.$$

As a result we get that for $\varepsilon < \varepsilon_0$

$$u_n < \sum_{j=0}^{k-1} \varepsilon^j \varphi(n_j) + \frac{\varepsilon^k}{1-\varepsilon}\varphi(n_k) < \left(\sum_{j=0}^{k-1} \varepsilon^{j/2} + \frac{\varepsilon^{k/2}}{1-\varepsilon}\right)\varphi(n) < \frac{\varphi(n)}{1-\varepsilon^{1/2}}. \quad (3.29)$$

In the second case the recursion stops for $k = k_0 = \min\{k : n_k < N\}$. As a result we arrive at the bound

$$u_n < \frac{\varphi(n)}{1-\varepsilon^{1/2}} + \frac{\varepsilon^{k_0-1}}{1-\varepsilon} \max_{k\ge 0} u_k. \quad (3.30)$$

Since $n_k \ge \delta^k n$, $k_0 \ge \log_\delta \frac{N}{n}$. It implies that $\varepsilon^{k_0} \le \exp\{-2^{-1}\ln\varepsilon \log_\delta n\}$ for $n > N^2$. Consequently, for sufficiently small ε

$$\varepsilon^{k_0} = o(n^{-2}) = o(\varphi(n)). \quad (3.31)$$

It follows from (3.30) and (3.31) that $u_n < 2\varphi(n)$ for $n > N^2$ if ε sufficiently small. Returning to (3.27) we conclude that for sufficiently large n

$$0 < l(n)n^{1-\alpha}u_n - a_n c_1(\alpha)I(\eta) < 2\varepsilon n^{1-\alpha} l(n) \max_{\delta n\le k\le n} \varphi(k),$$

where $c_1(\alpha) = 2\alpha^3/\Gamma^2(1-\alpha)\Gamma(2\alpha+1)$. It is easily seen that

$$\limsup_{n\to\infty} n^{1-\alpha} l(n) \max_{\delta n\le k\le n} \varphi(k) \le \delta^{\alpha-1} c_1(\alpha)I(\eta).$$

It follows from two latter relations that

$$\lim_{n\to\infty} l(n)n^{1-\alpha}u_n = c_1(\alpha)I(0). \quad (3.32)$$

It remains to calculate $c_1(\alpha)I(0)$. Obviously,

$$I(0) = B(2\alpha, 1-\alpha) = \frac{\Gamma(2\alpha)\Gamma(1-\alpha)}{\Gamma(1+\alpha)}.$$

Consequently,

$$c_1(\alpha)I(0) = \frac{2\alpha^3\Gamma(2\alpha)}{\Gamma(1-\alpha)\Gamma(2\alpha+1)\Gamma(1+\alpha)} = \frac{\alpha}{\Gamma(1-\alpha)\Gamma(\alpha)} = \frac{\alpha\sin\pi\alpha}{\pi}. \quad (3.33)$$

It follows from (3.32) and (3.33) that

$$\lim_{n\to\infty} l(n)n^{1-\alpha}u_n = \frac{\alpha\sin\pi\alpha}{\pi}$$

On the other hand, by (2.12)

$$\frac{\mathbf{P}(X=n)}{\mathbf{P}^2(X\ge n)} \sim \frac{\alpha^2}{l(n)n^{1-\alpha}}.$$

Hence,

$$\frac{\sin\pi\alpha}{\pi\alpha}\frac{\mathbf{P}(X=n)}{\mathbf{P}^2(X\ge n)} \sim \frac{\alpha\sin\pi\alpha}{\pi l(n)n^{1-\alpha}} \sim u_n,$$

which was to be proved.

4. The proof of Theorem 1.2

According to definition

$$h_n = C_n(-\ln(1 - f(z))).$$

Hence,

$$nh_n = C_n\left(\frac{f'(z)}{1 - f(z)}\right).$$

Consequently,

$$h_n = \frac{1}{n}\sum_{k=0}^{n}(k+1)p_{k+1}u_{n-k}. \tag{4.1}$$

Applying Theorem 1.1, we have

$$\sum_{\varepsilon n \le k \le (1-\varepsilon)n}(k+1)p_{k+1}u_{n-k} \sim \frac{\alpha \sin \pi\alpha}{\pi}\sum_{\varepsilon n \le k \le (1-\varepsilon)n}(k+1)^{-\alpha}(n-k)^{\alpha-1}$$

$$\sim \frac{\alpha \sin \pi\alpha}{\pi}\int_{\varepsilon}^{1-\varepsilon} u^{-\alpha}(1-u)^{\alpha-1}du \equiv \frac{\alpha \sin \pi\alpha}{\pi}I(\varepsilon). \tag{4.2}$$

On the other hand, applying Lemmas 2.4 and 2.7, we have

$$\limsup_{n\to\infty}\sum_{0 \le k < \varepsilon n}(k+1)p_{k+1}u_{n-k} < \frac{\alpha}{\pi(1-\alpha)}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{1-\alpha} \tag{4.3}$$

and

$$\limsup_{n\to\infty}\sum_{(1-\varepsilon)n < k \le n}(k+1)p_{k+1}u_{n-k} < \frac{1}{\pi}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{\alpha}. \tag{4.4}$$

It follows from (4.2)–(4.4) that

$$\lim_{n\to\infty}\sum_{k=0}^{n}(k+1)p_{k+1}u_{n-k} = \alpha\frac{\sin \pi\alpha}{\pi}I(0). \tag{4.5}$$

Obviously,

$$I(0) = B(\alpha, 1-\alpha) = \Gamma(\alpha)\Gamma(1-\alpha) = \frac{\pi}{\sin \pi\alpha}. \tag{4.6}$$

Combining (4.1), (4.5), (4.6), we obtain that

$$h_n \sim \frac{\alpha}{n},$$

which was to be proved.

Acknowledgments. I thank the referee for helpful remarks.

References

[1] *Garsia A., Lamperti J.* A discrete renewal theorem with infinite mean. – Commentarii Mathematici Helvetici. 1963, **37**, 221–234.

[2] *De Bruijn N.G., Erdos P.* On a recursion formula and some Tauberian theorems. – J.Res.Nat.Bur. Stand., 1953, **50**, 161–164.

[3] *Williamson J.A.* Random walks and Riesz kernels. – Pac. J. Math., 1968, **25**, No 2, 393–415.

[4] *Rvacheva E.L.* On domains of attraction of multi - dimensional distributions. – Selected Translations in Mathematical Statistics and Probability, 1962, 183-203. L'vov Gos. Univ. Uch. Zap. 29, Ser. Meh.– Mat., 1954, **6**, No 29, p.5–44.

[5] *Nagaev S.V.* Renewal theorem in the absence of power moments. – Teor. Veroyatn. i Primen., 2011, **56**, No 1, 188–197.

[6] *Doney R.A.* One-sided local large deviation and renewal theorems in the case of infinite mean. Probab. Theory Related Fields. –Springer-Verlag, 1997, **107**, 1997, 451-465.

[7] *Nagaev S.V.* Large deviations of sums of independent random variables. – Ann. Prob., 1979, **7**, No 5, 745-789.

[8] *Seneta E.* Regularly varying functions. – Lecture Notes in Mathematics 508, Springer Verlag Berlin Heidelberg - New York, 1976.

[9] *Greenwood P., Omey I., J.L. Teugels J. L.* Harmonic renewal measures.– Z. Warsch. Verw. Gebiete, 1982, **59**, 391-409.

[10] *Grubel R.J.* On harmonic renewal measures. – Probab. Theory Rel. Fields, 1986, **71**, 393–403.

[11] *Grubel R.J.* Harmonic renewal sequences and the first positive sum. – J. London Math. Soc. 1988, **38**, No 2, 179–192.

[12] *Stam A.J.* Some theorems on harmonic renewal measures. – Stochastic processes and their Appl., 1991, **39**, 277–285

[13] *Feller W.* An Introduction to Probability Theory and Applications. – John Wiley and Sons, New York, London, Sydney, Toronto, 1971, **2**, 752p.

[14] *Gradshtein I., Ryzhik I.M.* Tables of integrals, sums and products.– Fiz.Mat.Giz., Moscow, 1962, 1100 p.

Annales Mathematicae et Informaticae
39 *(2012) pp. 193–206*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Experiments on the distance of two-dimensional samples

Csaba Noszály

University of Debrecen, Faculty of Informatics, Debrecen, Hungary
e-mail: noszaly.csaba@inf.unideb.hu

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

The distance of two-dimensional samples is studied. The distance is based on the optimal matching method. Simulation results are obtained when the samples are drawn from normal and uniform distributions.

Keywords: Optimal matching, simulation, Gaussian distribution, goodness of fit, general extreme value distribution.

MSC: 62E17, 62H10

1. Introduction

A well-known result in optimal matchings is the following (see Ajtai-Komlós-Tusnády [1]). Assume that both $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ are independent identically distributed (i.i.d.) random variables with uniform distribution on the two-dimensional unit square. Let $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ be independent of each other. Let

$$t_n = \min_{\pi} \sum_{i=1}^{n} ||X_{\pi(i)} - Y_i||, \tag{1.1}$$

where the minimum is taken over all permutations π of the first n positive integers. Then

$$C_1(n \log n)^{1/2} < t_n < C_2(n \log n)^{1/2} \tag{1.2}$$

with probability $1 - \mathrm{o}(1)$ (Theorem in [1]). t_n is the so-called transportation cost. Talagrand in [6] explains the specific feature of the two-dimensional case. In [7] it is explained that the transportation cost is closely related to the empirical process. So the following question arises. Can t_n serve as the basis of testing goodness of fit? Therefore to find the distribution of t_n is an interesting task. That problem was suggested by G. Tusnády.

Testing multidimensional normality is an important task in statistics (see e.g. [4]). In this paper we study a particular case of this problem. We study the fit to two-dimensional standard normality. The main idea is the following. Assume that we want to test if a random sample $X_1, \dots, X_n$ is drawn from a population with distribution F. We generate another sample $Y_1, \dots, Y_n$ from the distribution F. Then we try to find for any X_i a similar member of the sample $Y_1, \dots, Y_n$. We hope that the optimal matching of the two samples gives a reasonable statistic to test the goodness of fit.

In this paper we concentrate on three cases, that is when both $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ are standard normal, then both of them are uniform, finally when $X_1, \dots, X_n$ are normal and $Y_1, \dots, Y_n$ are uniform. We calculate the distances of the samples, then we find the statistical characteristics of the distances. The quantiles can serve as critical values of a goodness of fit test. Finally, we show some results on the distribution of our test statistic.

We use the classical notion of sample, i.e. $X_1, \dots, X_n$ is called a sample if $X_1, \dots, X_n$ are i.i.d. random variables.

For two given samples $X_i, Y_i \in \mathbb{R}^2$ $(i = 1, \dots, n)$ let us define the statistic T_n by

$$T_n = \min_{\pi \in S_n} \sum_{i=1}^{n} ||X_{\pi(i)} - Y_i||^2. \tag{1.3}$$

Here S_n denotes the set of permutations of $\{1, \dots, n\}$ and $||.||$ is the Euclidean norm. Formula (1.3) naturally expresses the 'distance' of two samples. We study certain properties of T_n for Gaussian and uniform samples. To this aim we made simulation studies for sample sizes $n = 2, \dots, 200$ with replication 1000 in each case. That is we generated two samples of sizes n, calculated T_n, then repeated this procedure 1000 times. Then we tried to fit the so called general extreme value (GEV) distribution (see [5], page 61) to the obtained data of size 1000. The distribution function of the general extreme value distribution is

$$F(x, \mu, \sigma, \xi) = \begin{cases} \exp\left(-\left[1 + \xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}\right), & \xi \neq 0; \\ \exp\left(-\exp\left(-\frac{(x-\mu)}{\sigma}\right)\right), & \xi = 0. \end{cases} \tag{1.4}$$

Here $\mu, \sigma > 0, \xi$ are real parameters. For further details see [5].

The values of T_n are obtained by Kuhn's Hungarian algorithm as described in [3]. We mention that a previous simulation study of T_n was performed in [2].

2. Simulation results for samples with common distribution

In this section we want to determine the distribution of T_n when the samples $X_1, \dots, X_n$ and $Y_1, \dots, Y_n$ have the same distribution. In terms of testing goodness of fit the task is the following.

Let $X_1, \ldots, X_n$ be a sample. We want to test the hypothesis

$$H_0 : \text{the distribution of } X_i \text{ is } F.$$

Generate another sample $Y_1, \ldots, Y_n$ from distribution F and calculate the test statistics T_n. If T_n is large, then we reject H_0. (In practice $X_1, \ldots, X_n$ are real life data, while $Y_1, \ldots, Y_n$ are random numbers.) To create a test we have to find some information on the distribution of T_n.

To obtain the distribution of T_n by simulation, we proceed as follows. For a fixed sample size n, $2n$ two-dimensional points are generated: $X_i = (X_{i1}, X_{i2})$, $Y_i = (Y_{i1}, Y_{i2})$, $i = 1, \ldots, n$, with independent coordinates. We restrict our attention to the simplest cases.

(a) Gaussian case when $X_{ij}, Y_{ij} \in \mathcal{N}(0,1)$, $i = 1, 2, \ldots, n$, $j = 1, 2$, i.e. they are standard normal.

(b) Uniform case when $X_{ij}, Y_{ij} \in \mathcal{U}(0,1)$, $i = 1, 2, \ldots, n$, $j = 1, 2$, i.e. they are uniformly distributed on $[0,1]$.

All the random variables involved are independent. Graphs of descriptives and tables of $5\%, 10\%, 90\%$ and 95% quantiles for selected sample sizes are presented in figures 1, 2 and tables 1, 4.

Figure 1(a) and Figure 2(a) show the sample mean and sample standard deviation of T_n, respectively, when both X_i and Y_i comes form two-dimensional standard normal. (They are calculated for each fixed n using 1000 replications.) Figure 1(b) and Figure 2(b) concern the case when both X_i and Y_i are uniform.

Table 1 shows the sample quantiles of T_n when both X_i and Y_i are two-dimensional standard normal. Each value is calculated for fixed n using 1000 replications. The upper quantile values (at 90% or 95%) can serve as critical values for the test

$$H_0 : X_i \text{ is two-dimensional standard normal.}$$

Table 2 contains the results when both samples are two-dimensional uniform (more precisely uniform on $[0,1] \times [0,1]$).

3. The mixed case

With the help of previous section's tables one can construct empirical confidence intervals for the distance T_n of two samples both in the Gaussian-Gaussian and uniform-uniform cases. In what follows we present some results on the distance T_n for the Gaussian-uniform case. For this aim we performed calculations for sample sizes $n = 2, \ldots, 200$ with 2000 replications in each cases. Note that here we used $\mathcal{U}(-\sqrt{3}, \sqrt{3})$ for the uniform variable because then we have $\mathbb{E}(Y_{ij}) = 0$ and $\mathbb{D}^2(Y_{ij}) = 1$.

Figure 3 and Table 3 concern the distribution of T_n when X_{ij} is standard normal and $Y_{ij} \in \mathcal{U}(-\sqrt{3}, \sqrt{3})$. That is the case when H_0 is not satisfied. If we compare the last columns (95% quantiles) of Table 3 and Table 1, then we see that our test is sensitive if the sample size is large ($n \geq 100$).

4. Fitting the *GEV*

To describe the distribution of T_n we fitted general extreme value distribution. For each fixed n we estimated the parameters of GEV from the 1000 replications. The maximum likelihood estimates of parameters ξ, μ, σ in (1.4) were obtained with MATLAB's *fitdist* procedure. Then we plotted the cumulative distribution function of the GEV. Figure 4(a), Figure 5(a) and Figure 6(a) show that the empirical distribution function of T_n fits well to the theoretical distribution function of the appropriate GEV when both X_i and Y_i are standard Gaussian. Figure 4(b), Figure 5(b) and Figure 6(b) show the same for uniformly distributed X_i and Y_i.

Figure 7 shows the empirical significance of Kolmogorov-Smirnov tests performed by *kstest*. The empirical p-values in Figure 7(a) and Figure 7(b) reveal that the fitting was succesful.

5. About the *GEV* parameters

To suspect something about the possible 'analytical form' of parameters ξ, σ, μ we made further simulations in the Gaussian case with 5000 replications for sample sizes $n = 2, \ldots, 500$. After several 'trial and error' attempts we got the following experimental results.

Figure 8 concern the functional form of the parameters. Here both X_i and Y_i were Gaussian. For each fixed n we fitted $GEV(\xi(n), \sigma(n), \mu(n))$. Then we approximated $\xi(n), \sigma(n)$ and $\mu(n)$ with certain functions. For example we obtained that $\xi(n)$ can be reasonably approximated with

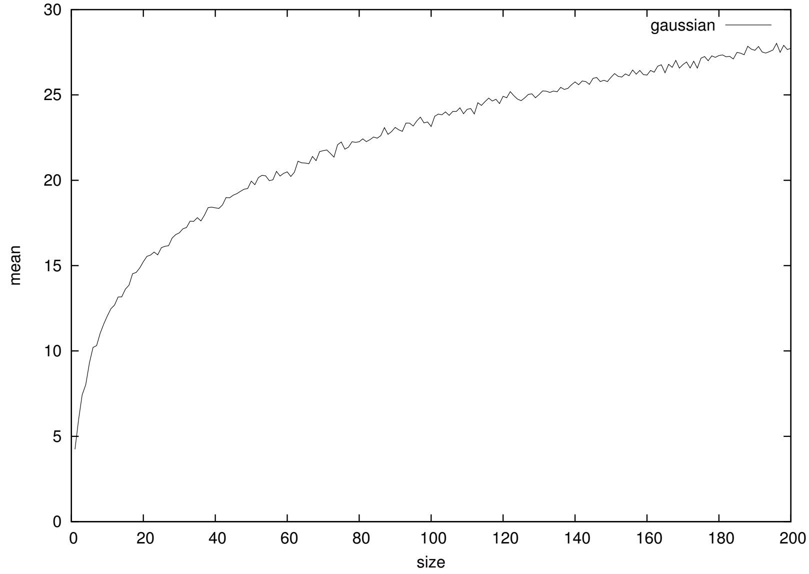
$$A/\sqrt{n} + B/\sqrt{\log(n)} + C$$

where A, B, C are given in Figure 8(a). Note that the classical goodness of fit measures (χ^2 and R^2) computed by *qtiplot* indicate tight fit.

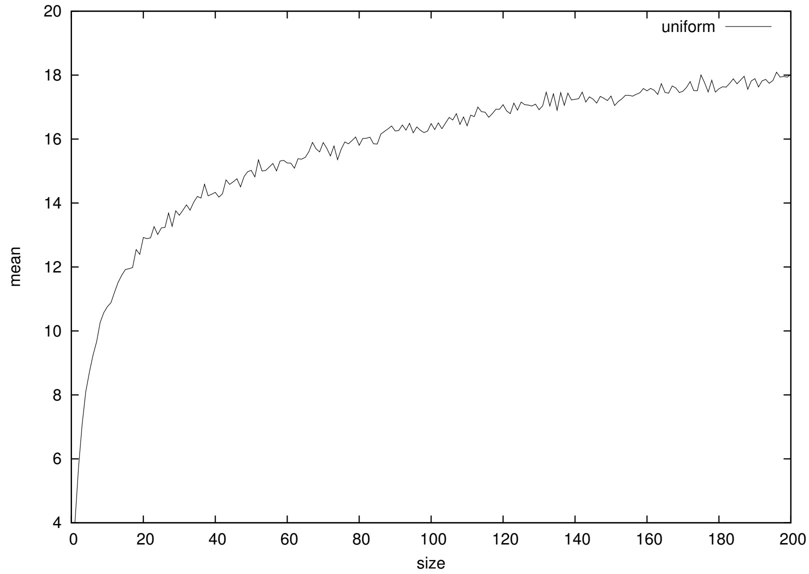
6. Tools

The Hungarian method was implemented in *C++* using the GNU g++ compiler. Most of the graphs were made with the help of the utility *gnuplot*. The fittings and the graphs of the last section were performed with *qtiplot*. MATLAB was used to compute the maximum likelihood estimators of the GEV.

7. Figures and tables

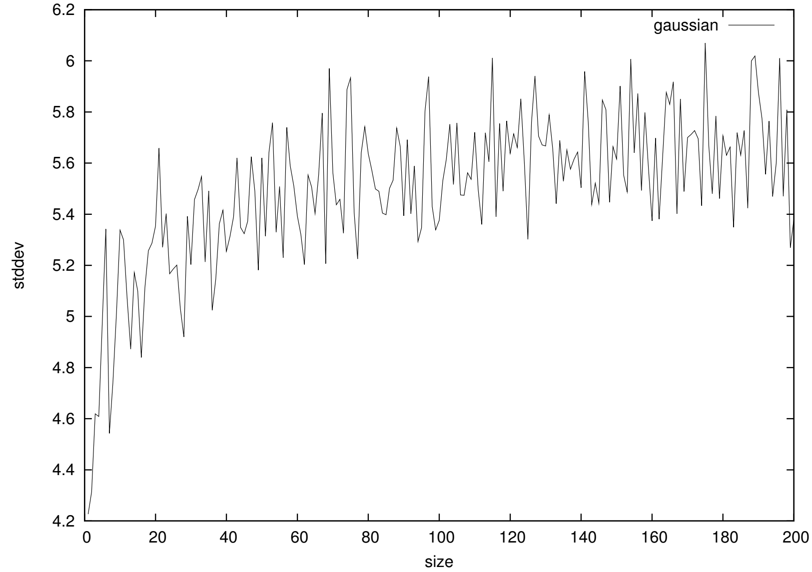


(a) Gaussian

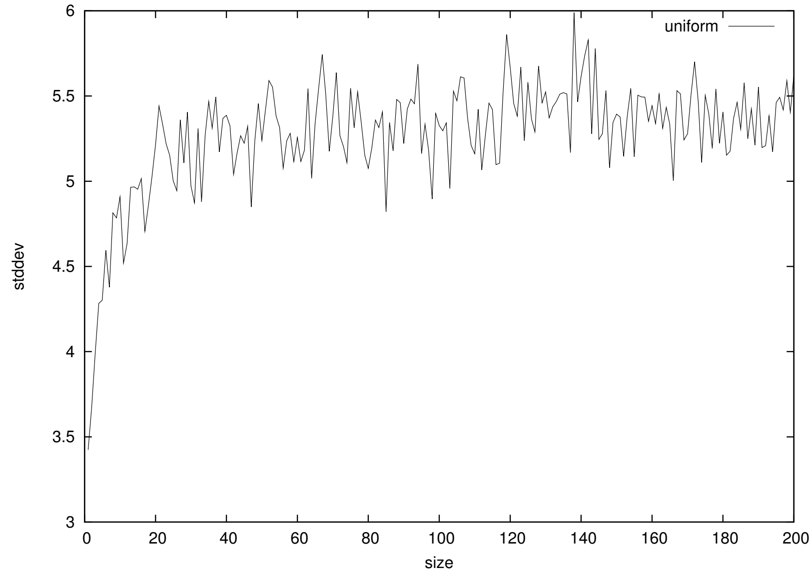


(b) uniform

Figure 1: Sample means



(a) Gaussian



(b) uniform

Figure 2: Sample standard deviations

size	mean	stddev	5%	10%	90%	95%
1	4.2412	4.2266	0.2381	0.4563	9.8020	12.2026
2	5.9673	4.3135	1.0836	1.7158	11.7292	14.2147
3	7.4222	4.6192	1.7995	2.5588	13.7402	16.0102
4	8.0206	4.6088	2.3981	3.2674	13.6933	16.0876
5	9.3045	4.9715	3.2585	4.1318	15.3954	18.3475
6	10.2078	5.3420	3.7299	4.5396	17.1818	19.4804
7	10.3123	4.5427	4.1995	5.2007	16.2449	19.0637
8	11.0368	4.7393	4.8979	5.7742	17.5610	19.8337
9	11.5844	5.0128	5.0169	5.7831	18.2823	20.9534
10	12.0570	5.3378	5.3833	6.3622	19.4668	21.8923
20	15.2328	5.3538	8.6359	9.5934	22.2431	24.9313
30	16.9197	5.2032	10.1785	11.0722	23.9962	27.0184
40	18.3806	5.2547	11.4259	12.6406	25.5429	28.0519
50	19.9502	5.6199	12.1530	13.5418	27.2146	30.6897
60	20.4902	5.3926	13.3070	14.3950	27.5160	31.5666
70	21.7366	5.5615	14.3281	15.5132	28.9196	32.1283
80	22.2543	5.6370	14.6116	15.8363	29.4256	32.3583
90	23.0996	5.3942	15.7942	16.9445	30.0369	33.0472
100	23.1510	5.3759	15.9969	17.0670	29.9957	33.3124
120	24.9210	5.6381	16.9766	18.4435	32.5778	34.7392
140	25.7610	5.5036	18.1150	19.5019	33.0206	35.3015
160	26.1585	5.3739	18.7503	20.0962	33.6252	36.3843
180	27.3072	5.7067	19.4513	20.7716	34.7677	37.3559
200	27.7257	5.3810	20.2195	21.4550	34.8416	37.1973

Table 1: Quantiles. (a) Gaussian case

size	mean	stddev	5%	10%	90%	95%
1	4.0262	3.4236	0.1970	0.4216	9.1473	10.7406
2	5.7465	3.6683	1.1099	1.7597	10.9533	12.8762
3	7.0817	3.9889	1.9709	2.5498	12.4090	14.4607
4	8.0923	4.2833	2.5078	3.3816	13.7593	16.6473
5	8.7164	4.3022	3.0438	4.0823	14.4455	17.0613
6	9.2447	4.5952	3.4694	4.3526	15.5264	18.0257
7	9.6608	4.3776	4.1441	5.0191	15.5752	17.7570
8	10.2707	4.8150	4.3921	5.3208	16.7853	19.6646
9	10.5731	4.7847	4.6570	5.4796	16.7946	19.8556
10	10.7589	4.9081	5.0283	5.7636	16.9497	19.6734
20	12.9231	5.2142	6.8068	7.5361	19.3753	22.0094
30	13.6183	4.9743	7.7449	8.5648	20.3468	23.5615
40	14.3316	5.3870	7.9332	9.0030	21.1094	24.8500
50	15.0187	5.2400	8.8329	9.6225	21.8464	25.6322
60	15.2523	5.2565	8.9552	9.8825	21.8732	25.0060
70	15.8911	5.3841	9.5833	10.5137	22.9240	26.1478
80	15.8035	5.0754	9.7065	10.5853	22.9256	26.3499
90	16.2536	5.2216	9.9975	10.7825	22.9050	26.0921
100	16.4830	5.3264	10.2617	10.8682	24.0967	27.2588
120	17.0734	5.6685	10.4419	11.3588	24.4638	28.9698
140	17.2442	5.6141	10.7627	11.5904	24.3573	27.6821
160	17.5099	5.4433	11.1689	12.0155	24.9285	27.9288
180	17.5731	5.4062	11.1337	12.0112	24.5775	27.8075
200	18.0244	5.6245	11.4052	12.3776	25.4038	28.4148

Table 2: Quantiles. (b) uniform case

size	mean	stddev	5%	10%	90%	95%
2	5.8818	4.1502	1.1652	1.6532	11.3164	13.6333
3	7.2554	4.1646	2.0590	2.8677	13.0166	15.0618
4	7.9939	4.4994	2.7365	3.3880	13.7425	16.6208
5	8.8754	4.4742	3.2754	4.1966	14.8754	17.3966
6	9.7783	4.5869	4.0448	4.7961	15.8904	18.3626
7	10.0745	4.5939	4.3386	5.2004	16.2399	18.8547
8	10.4734	4.4532	4.5207	5.4837	16.4532	18.8462
9	11.0886	4.6638	5.2157	6.1198	17.3704	19.7215
10	11.5361	4.8266	5.4715	6.4598	17.9202	20.9748
20	14.6339	5.1743	8.0674	9.0785	21.4606	24.9300
30	16.7707	5.1719	10.0346	11.0548	23.5244	26.3029
40	18.4228	5.4695	11.1791	12.2861	25.4337	28.2022
50	20.0629	5.6343	12.4183	13.9359	27.2744	30.4001
60	21.1895	5.4223	13.4741	14.7804	28.4744	31.1559
70	22.4896	5.6856	14.7198	16.0582	29.7835	33.4986
80	24.0810	5.8618	15.6983	17.2439	31.5593	34.0761
90	25.1524	6.0496	16.9018	18.1007	33.2026	36.5343
100	26.5301	6.1983	17.9469	19.3842	34.9823	37.9294
120	28.8502	6.3829	19.9032	21.3154	37.0409	40.2436
140	30.9987	6.5694	21.6081	23.2774	39.7266	42.9169
160	33.1575	6.7402	23.3054	25.1265	42.0649	45.1835
180	34.9612	6.7718	25.3528	26.8373	43.5126	47.4540
200	37.2655	7.0713	27.1785	28.8087	46.5339	49.9597

Table 3: Quantiles. Gaussian-uniform case

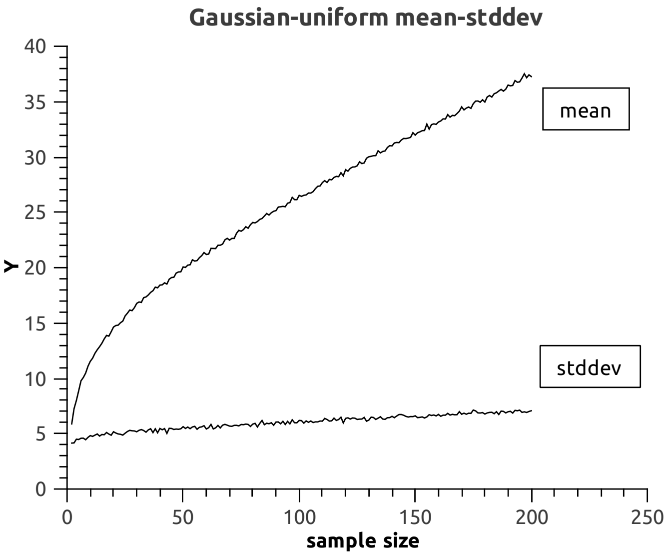
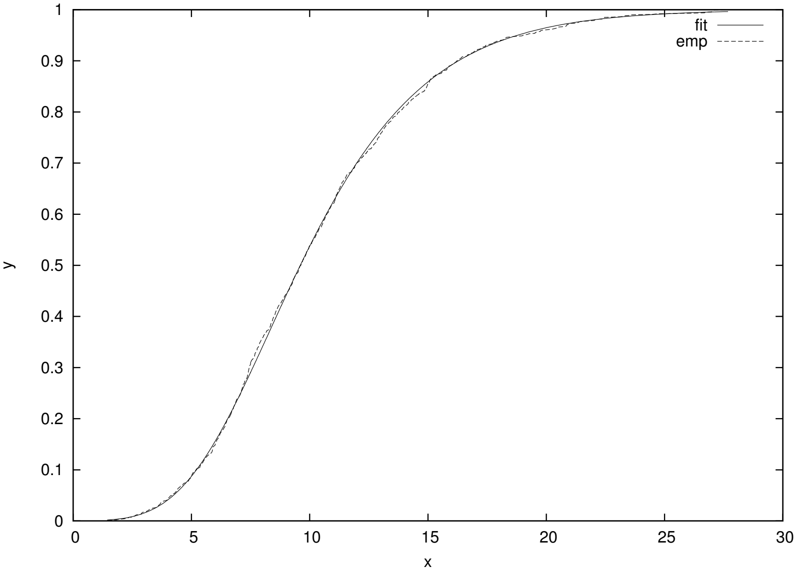
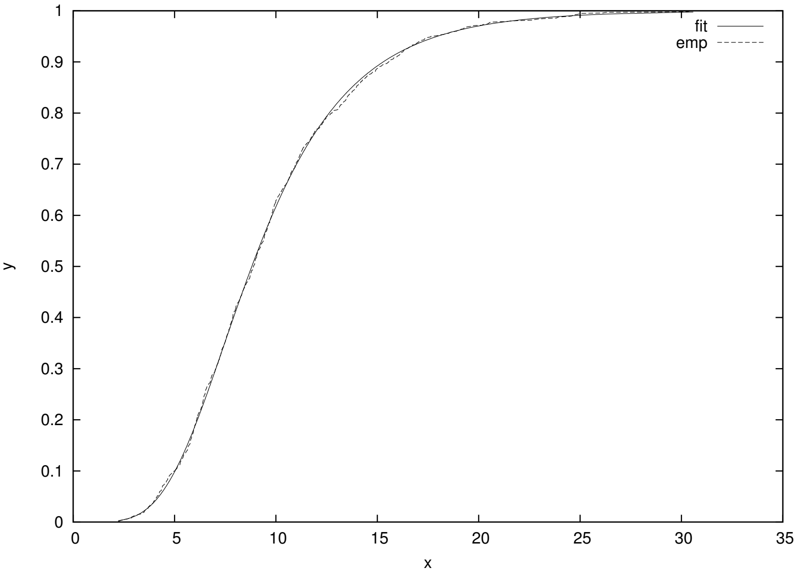


Figure 3: Sample means and standard deviations, Gaussian-uniform case

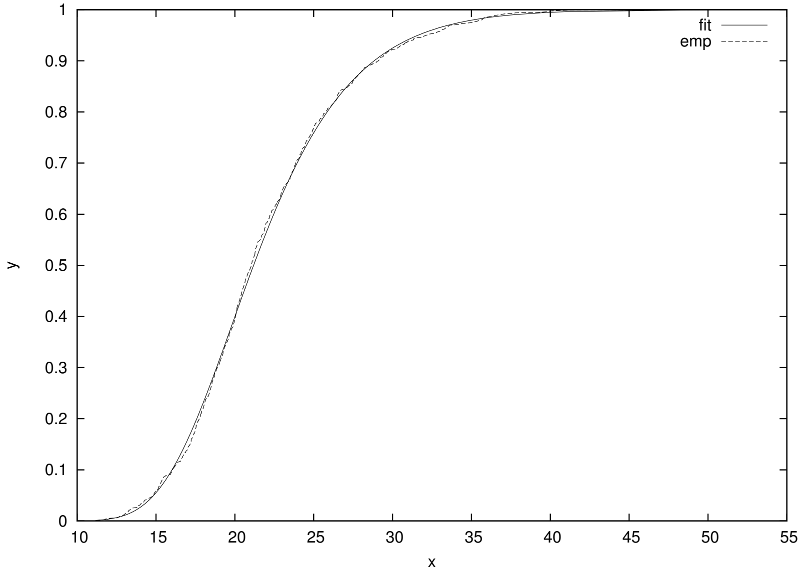


(a) Gaussian

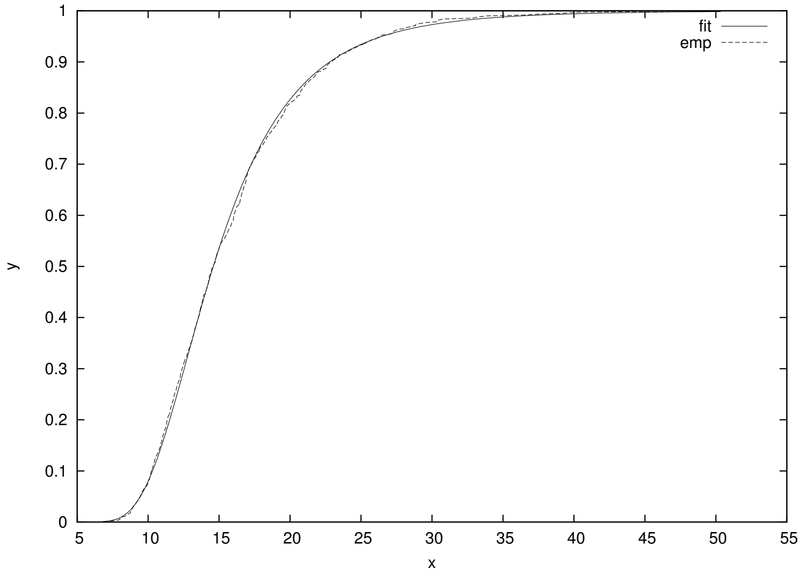


(b) uniform

Figure 4: Empirical and fitted cdf, $n = 7$

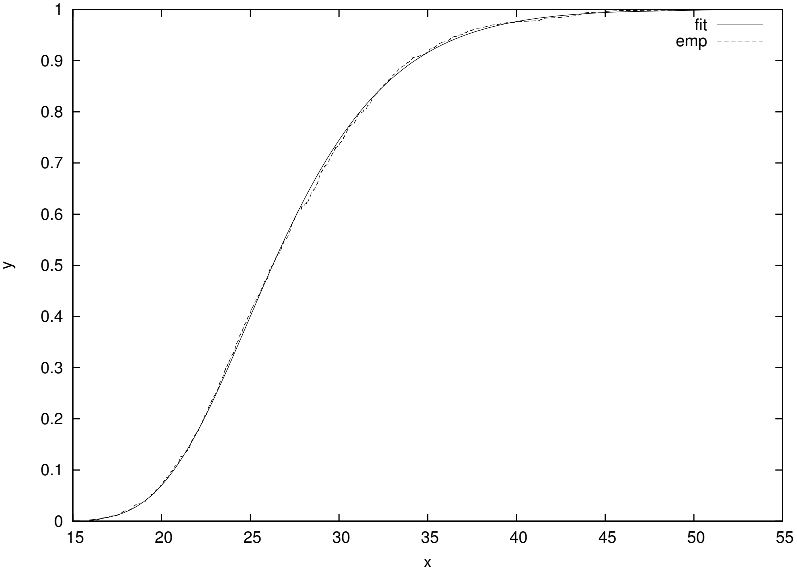


(a) Gaussian

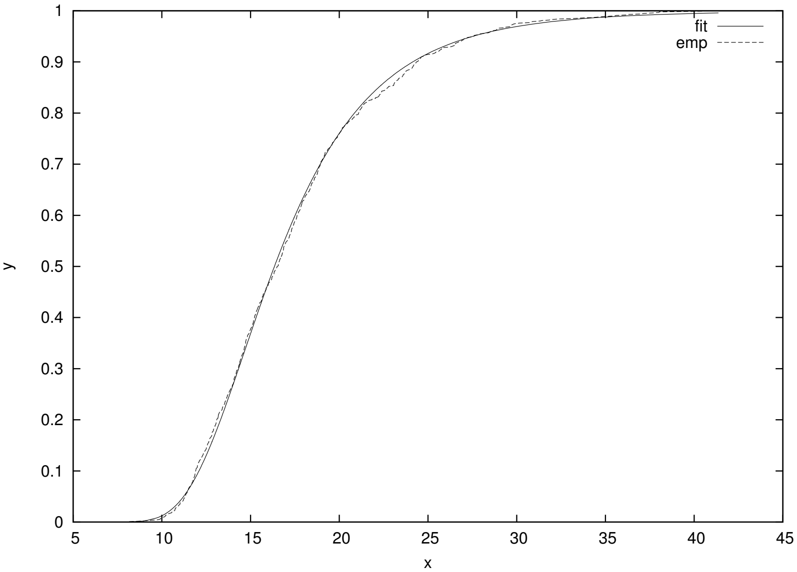


(b) uniform

Figure 5: Empirical and fitted cdf, $n = 77$

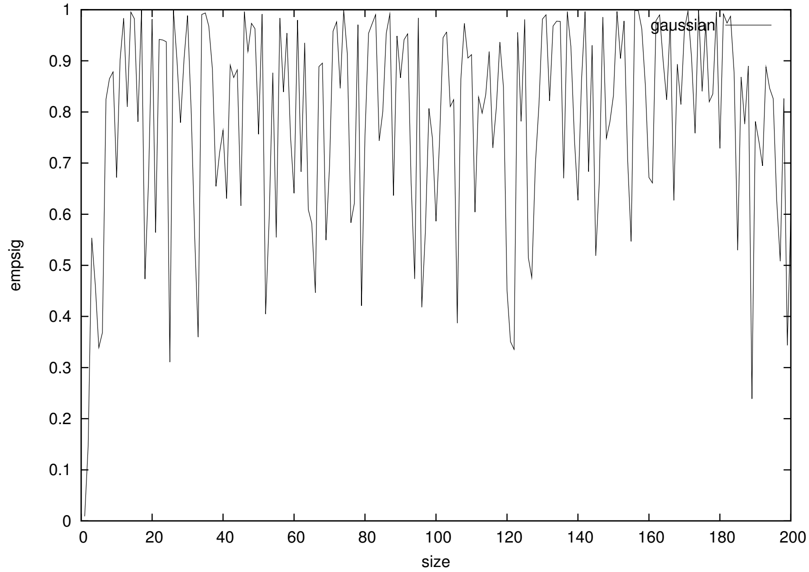


(a) Gaussian

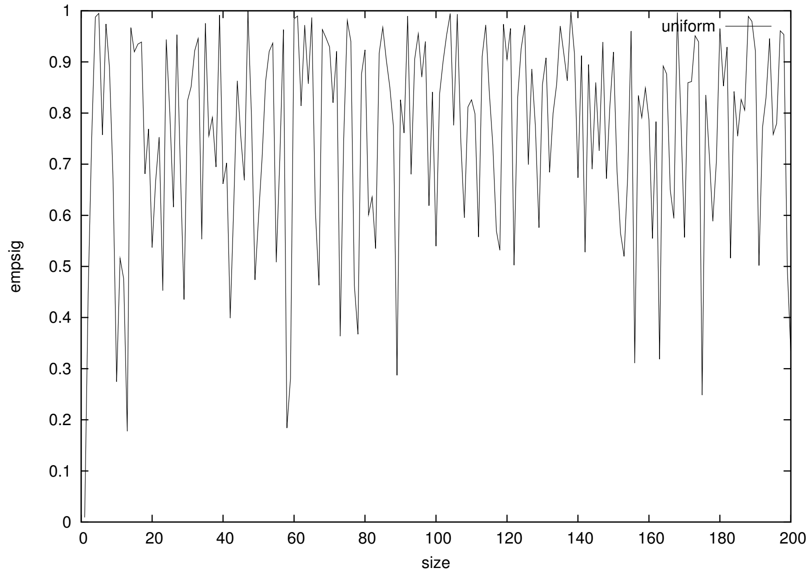


(b) uniform

Figure 6: Empirical and fitted cdf, $n = 177$

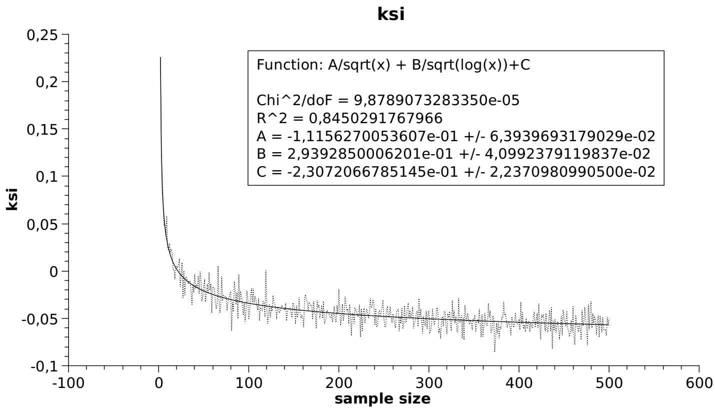


(a) Gaussian

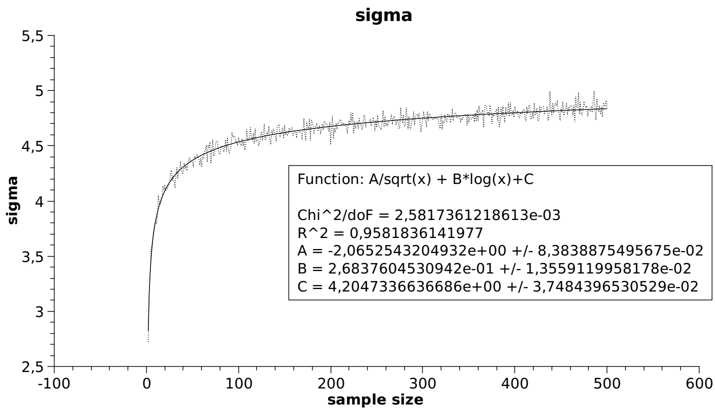


(b) uniform

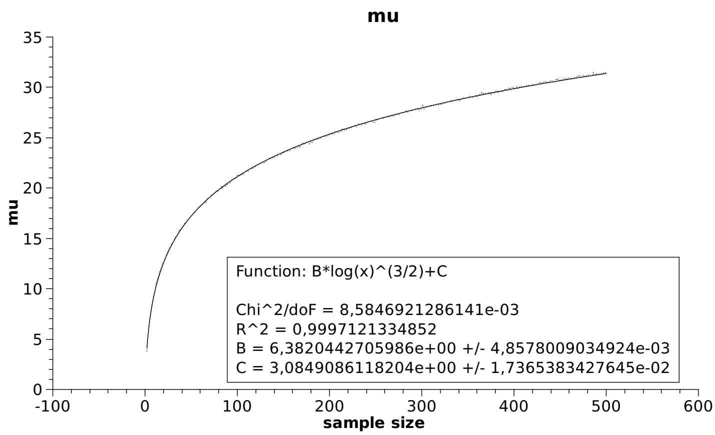
Figure 7: Empirical significance



(a) ξ



(b) σ



(c) μ

Figure 8: Estimating ξ, σ and μ

Acknowledgement. I would like to thank Prof. István Fazekas who suggested using GEV to this problem.

References

[1] M. Ajtai, J. Komlós, G. Tusnády, On optimal matchings, *Combinatorica*, 4(4):259–264, 1984.

[2] I. Fazekas, A. Roszjár, On test of two dimensional normality, *Technical Report, Univ. of Debrecen*, 153, 1995.

[3] E.G. Gol'shtein, D.B. Yudin, Problems of linear programming of transportation type, (Russian), *Nauka, Moscow*, 1969.

[4] N. Henze, Invariant tests for multivariate normality: A critical review, *Stat. Pap.*, 43(4):467–506, 2002.

[5] S. Kotz, S. Nadarajah, Extreme value distributions. Theory and applications, *Imperial College Press, London*, 2000.

[6] M. Talagrand, The generic chaining. Upper and lower bounds of stochastic processes, *Springer, Berlin*, 2005.

[7] J.E. Yukich, Optimal matching and empirical measures, *Proc. Am. Math. Soc.*, 107(4):1051–1059, 1989.

Annales Mathematicae et Informaticae
39 *(2012) pp. 207–224*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Robust estimation in time series with long and short memory properties

Valdério Anselmo Reisen, Fabio Fajardo Molinares

Departamento de Estatística, Universidade Federal do Espírito Santo, Vitória/ES, Brazil
valderioanselmoreisen@gmail.com, fabio.molinares@ufes.br.

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

This paper reviews recent developments of robust estimation in linear time series models, with short and long memory correlation structures, in the presence of additive outliers. Based on the manuscripts Fajardo, Reisen & Cribari-Neto 2009 [7] and Lévy-Leduc, Boistard, Moulines, Taqqu & Reisen 2011 [19], the emphasis in this paper is given in the following directions; the influence of additive outliers in the estimation of a time series, the asymptotic properties of a robust autocovariance function and a robust semiparametric estimation method of the fractional parameter d in ARFIMA(p, d, q) models. Some simulations are used to support the use of the robust method when a time series has additive outliers. The invariance property of the estimators for the first difference in ARFIMA model with outliers is also discussed. In general, the robust long-memory estimator leads to be outlier resistent and is invariant to first differencing.

Keywords: Additive outliers, ARFIMA model, long-memory, robustness.

1. Introduction

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stationary time series with spectral density that behaves like

$$f_X(\omega) \sim h(\omega) \mid \omega \mid^{-2d}, \text{ as } \omega \to 0 \tag{1.1}$$

where the spectral density $h(\omega)$ is a nonvanishing and continuously differentiable function with bounded derivative for $-\pi \leq \omega \leq \pi$, and $d < 0.5$.

A well-known stationary parametric model with the above spectral density is the ARFIMA(p, d, q) process, which is the solution of the equation

$$X_t - \mu = (1 - B)^{-d}\eta_t, \quad t \in \mathbb{Z}, \tag{1.2}$$

where $\eta_t = \frac{\Theta(B)}{\Phi(B)}\epsilon_t$ is an ARMA(p,q) process, μ is the mean (here it is assumed that $\mu = 0$), $\Phi(B) \equiv 1 - \sum_{j=1}^{p} \phi_j B^j$, $\Theta(B) \equiv 1 - \sum_{i=1}^{q} \theta_i B^i$ and p and q are positive integers (Hosking 1981 [11]). $\Phi(z)$ and $\Theta(z)$, with a scalar z, are the autoregressive and moving average polynomials with all roots outside the unit circle and share no common factors. d is the parameter that holds the memory of the process, that is, when $d \in (-0.5, 0.5)$ the ARFIMA(p,d,q) process is said to be invertible and stationary. Besides, for $d \neq 0$, its autocovariance decays at a hyperbolic rate ($\gamma(j) = O(j^{-1+2d})$). For $d = 0$, $d \in (-0.5, 0)$ or $d \in (0, 0.5)$, the process is said to be short-memory, intermediate-memory or long-memory, respectively. The long-memory property is related to the behavior of the autocovariances, which are not absolutely summable and the spectral density becomes unbounded at zero frequency. In the intermediate-memory region, the autocovariances are absolutely summable and, consequently, the spectral density is bounded.

The spectral density function of $\{X_t\}_{t\in\mathbb{Z}}$ is given by

$$f_X(\omega) = f_\eta(\omega)\left[2\sin\left(\frac{\omega}{2}\right)\right]^{-2d}, \omega \in [-\pi, \pi]. \tag{1.3}$$

$f_X(\omega)$ is continuous except for $\omega = 0$ where it has a pole when $d > 0$. A recent review of the ARFIMA model and its properties can be found in Palma 2007 [23] and Doukhan, Oppenheim & Taqqu 2003 [6].

Many estimators for the fractional parameter d in long-memory time series have already been proposed in the literature. Among them are the semiparametric procedures, a group which includes a wide variety of estimators based on the Ordinary Least Square (OLS) method. These procedures require the use of the spectral density parameterized within a neighborhood of zero frequency. Some references on this subject include the works of Geweke & Porter-Hudak 1983 [10], Reisen 1994 [26] and Robinson 1995 [27], among others. An overview of long-range dependence processes can be found in Beran 1994 [1] and Doukhan et al. 2003 [6].

Time series with outliers or atypical observations is quite common in any area of application. In the case where the data is time-dependent, several authors such as Ledolter 1989 [17], Chang, Tiao & Chen 1988 [4] and Chen & Liu 1993 [5] have studied the effect of outliers in a time series that follows ARIMA models. In general, they have concluded that the parameter estimates of ARMA models become more biased when the data contains outliers. Similar conclusion is also observed when estimating the fractional parameter in ARFIMA models. The outliers cause a substantial bias in the differencing parameter (Fajardo et al. 2009 [7]).

An autocovariance robust function was proposed by Ma & Genton 2000 [20]. The asymptotical properties of this function are studied by Lévy-Leduc et al. 2011 [19]. The results presented in Fajardo et al. 2009 [7], Lévy-Leduc et al. 2011 [19] and Lévy-Leduc, Boistard, Moulines, Taqqu & Reisen 2011 [18] are the motivations of this paper. The impact of outliers in the estimation of ARFIMA models under different context is here studied. The asymptotical properties of a robust autocovariance function is discussed and some empirical examples are used to illustrate the usefulness of a robust fractional parameter estimator. The invariance property of the estimator to the first difference is also empirically studied. The outline of this papers is as follows: Section 2 discusses the model and the impact of the outliers in time series. Section 3 summarizes the main results related to the robust auto-covariance estimator given in Lévy-Leduc et al. 2011 [19] and discusses the robust estimation of the fractional parameter in the ARFIMA model. Section 4 presents some empirical studies and an application is discussed in Section 5. Concluding remarks and future directions are given in Section 6.

2. The impact of outliers in time series

Suppose $x_1, ..., x_n$ is a partial realization of $\{X_t\}_{t\in\mathbb{Z}}$. Hence, the periodogram function is defined as $I_x(\omega) = (2\pi n)^{-1}|\sum_{t=1}^{n} x_t e^{i\omega t}|^2$. It follows that, when $d = 0$ in the ARFIMA model,

$$I_x(\omega) = 2\pi f_X(\omega)\frac{I_\epsilon(\omega)}{\sigma_\epsilon^2} + H(\omega) \tag{2.1}$$

where $\mathbb{E}[|H(\omega)|^2] = O(\frac{1}{n^{2\xi}})$ $(\xi > 0)$ is uniformly in $\omega \in [-\pi, \pi]$ (Theorem 6.2.2 in Priestley 1981 [25]) and $I_\epsilon(\cdot)$ is the periodogram of the residuals. From (1.2) and Theorem 6.1.1 in Priestley 1981 [25], asymptotic sample properties of $\frac{I_x(\omega)}{f_X(\omega)}$ are derived and they are summarized as follows. If $\{\epsilon_t\}_{t\in\mathbb{Z}}$ are normally distributed, for a fixed set of values of the Fourier frequencies $\omega_j = \frac{2\pi j}{n}$, $j = 1, ..., [n/2]$, where $[\cdot]$ means the integer part, asymptotically the set of variables $\frac{I_x(\omega_j)}{f_X(\omega_j)}$ is independently distributed, each distributed as $\frac{\chi_2^2}{2}$. At $\omega = 0$ and π, the distributions are χ_1^2 (for details see Priestley 1981 [25]). These asymptotic results for the periodogram lead to $\mathbb{E}\left[\frac{I_x(\omega_j)}{f_X(\omega_j)}\right] \to 1$ and $\text{var}\left[\frac{I_x(\omega_j)}{f_X(\omega_j)}\right] \to (1 + \delta(\omega_j))$ as $n \to \infty$, where

$$\delta(\omega_j) = 1 \text{ if } \omega_j = 0, \pi \text{ and } 0 \text{ otherwise.} \tag{2.2}$$

The above results establish the unbiasedness and inconsistency properties of $I_x(\omega_j)$.

Due to the singularity of $f_X(\omega)$ when $d > 0$, the standard results of the asymptotic distribution of the periodogram discussed previously can not be applied to $I_x(\omega_j)$ for small and fixed j. Hurvich & Beltrão 1993 [13] showed that $\lim_{n\to\infty} \mathbb{E}\left[\frac{I_x(\omega_j)}{f_X(\omega_j)}\right]$ depends on j and d, and exceeds unity for most $d \neq 0$ (Künsch 1986 [16]; Robinson 1995 [28]). For $j \neq k$, $\frac{I_x(\omega_j)}{f_X(\omega_j)}$ and $\frac{I_x(\omega_k)}{f_X(\omega_k)}$ are correlated, and for a fixed value j and Gaussian processes, the limiting distribution of $\frac{I_x(\omega_j)}{f_X(\omega_j)}$ is not exponential (Robinson 1995 [28]). That is, under the Gaussian assumption, Hurvich & Beltrão 1993 [13] show that the normalized periodogram $\frac{I(\omega)}{f_X(\omega)}$ is asymptotically distributed as the quadratic form

$$\frac{\alpha_1}{2}\chi_1 + \frac{\alpha_2}{2}\chi_2$$

where χ_1 and χ_2 are variables with Chi-squared distribution with one degree of freedom, $\alpha_1 = L_j(d) - 2L_j^*(d)$, $\alpha_2 = L_j(d) + 2L_j^*(d)$,

$$L_j(d) = \lim_{n\to\infty} \mathbb{E}\left\{\frac{I_x(\omega_j)}{f_X(\omega_j)}\right\} = \frac{2}{\pi}\int_{-\infty}^{\infty} \frac{\sin^2(\omega/2)}{(2\pi j-\omega)^2}\left|\frac{\omega}{2\pi j}\right|^{-2d} d\omega$$

and

$$L_j^*(d) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\sin^2(\omega/2)}{(2\pi j-\omega)(2\pi j+\omega)}\left|\frac{\omega}{2\pi j}\right|^{-2d} d\omega.$$

Let $\{Z_t\}_{t\in\mathbb{Z}}$ be a process contaminated by additive outliers, which is described by

$$Z_t = X_t + \sum_{j=1}^{m} \varpi_j Y_{j,t}, \tag{2.3}$$

where m is the maximum number of outliers; the unknown parameter ω_j represents the magnitude of the jth outlier, and $Y_{j,t}\ (\equiv Y_j)$ is a random variable ($r.v.$) with probability distribution $\Pr(Y_j=-1) = \Pr(Y_j=1) = \frac{p_j}{2}$ and $\Pr(Y_j=0) = 1-p_j$, where $\mathbb{E}[Y_j] = 0$ and $\mathbb{E}[Y_j^2] = \mathrm{var}(Y_j) = p_j$. Model 2.3 is based on the parametric models proposed by Fox 1972 [8]. Y_j is the product of *Bernoulli*(p_j) and *Rademacher* random variables; the latter equals 1 or -1, both with probability $\frac{1}{2}$. X_t and Y_j are independent random variables.

Some results related to the effects of outliers on the spectral density and on the autocorrelation functions of $\{Z_t\}_{t\in\mathbb{Z}}$ are presented as follows.

Proposition 2.1. *Suppose that $\{Z_t\}_{t\in\mathbb{Z}}$ follows Model 2.3.*

i. The autocovariance function (ACOVF) of $\{Z_t\}_{t\in\mathbb{Z}}$ is given by

$$\gamma_z(h) = \begin{cases} \gamma_X(0) + \sum_{j=1}^{m} \varpi_j^2 p_j, & \text{if } h=0,\\ \gamma_X(h), & \text{if } h\neq 0,\end{cases}$$

where $\gamma_X(h) = \mathbb{E}[X_tX_{t+h}] - \mathbb{E}[X_t]\mathbb{E}[X_{t+h}]$ with $h\in\mathbb{Z}$.

ii. The spectral density function of $\{Z_t\}$ is given by

$$f_Z(\omega) = f_X(\omega) + \frac{1}{2\pi}\sum_{j=1}^{m} \varpi_j^2 p_j, \quad \omega\in(-\pi,\pi],$$

where $f_X(\omega) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} \gamma_X(h)e^{-ih\omega}$.

Proposition 2.1 states that $\gamma_z(h)$, for $h=0$, depends on $\mathrm{var}(Y_j)$. $\gamma_Z(0)$ increases with $\mathrm{var}(Y_j)$ (see the proof in Fajardo et al. 2009 [7]). This relation between $R_Z(0)$

and $\text{var}(Y_j)$ will certainly affect the model parameter estimates because it reduces the magnitude of the autocorrelations and introduces loss of information on the pattern of serial correlation (see also Chan 1992, 1995 [2, 3]). The spectral form of $\{Z_t\}_{t\in\mathbb{Z}}$ (Model 2.3) when $\{X_t\}_{t\in\mathbb{Z}}$ follows an ARFIMA(p,d,q) model is given in the next lemma.

Lemma 2.2. *Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stationary and invertible ARFIMA(p,d,q) process. Also, let $\{Z_t\}_{t\in\mathbb{Z}}$ be such that $Z_t = X_t + \sum_{j=1}^{m} \varpi_j Y_j$, where m is the maximum number of outliers, the unknown parameter ϖ_j is the magnitude of the jth outlier and Y_j is a r.v. with probability distribution $\Pr(Y_j = -1) = \Pr(Y_j = 1) = \frac{p_j}{2}$ and $\Pr(Y_j = 0) = 1 - p_j$. The spectral density of $\{Z_t\}_{t\in\mathbb{Z}}$ is*

$$f_Z(\omega) = \frac{\sigma_\epsilon^2}{2\pi}\frac{|\Theta(e^{-i\omega})|^2}{|\Phi(e^{-i\omega})|^2}\left\{2\sin\left(\frac{\omega}{2}\right)\right\}^{-2d} + \frac{1}{2\pi}\sum_{j=1}^{m}\varpi_j^2 p_j.$$

The proof of Lemma 2.2 follows directly from Proposition 2.1.

The effects of an outlier on the sample autocovariance function and on the periodogram are given below.

Proposition 2.3. *Let $z_1, z_2, \ldots, z_n$ be generated from Model 2.3 with one outlier, and let the outlier occur at time $t = T$ with $h < T < n-h$. It follows that:*

i. The sample ACOVF is given by

$$\widehat{\gamma}_z(h) = \widehat{\gamma}_x(h) + \frac{\varpi}{n}(x_{T-h} + x_{T+h} - 2\bar{x}) + \frac{\omega^2}{n}\delta'(h) + o_p(n^{-1}), \tag{2.4}$$

where $\widehat{\gamma}_x(h) = \frac{1}{n}\sum_{t=1}^{n-h}(x_t - \bar{x})(x_{t+h} - \bar{x})$ *and* $\delta'(h) = \begin{cases} 1, & \text{when } h = 0,\\ 0, & \text{otherwise.}\end{cases}$

ii. The periodogram is given by

$$I_z(\omega) = I_x(\omega) + \Delta(\varpi), \quad \omega \in (-\pi, \pi],$$

where $I_x(\omega) = \frac{1}{2\pi}\sum_{h=-(n-1)}^{n-1}\widehat{\gamma}_x(h)e^{-ih\omega}$, *and*

$$\Delta(\varpi) = \frac{\varpi^2}{2\pi n} \pm \frac{\varpi}{\pi n}\left\{(x_T - \bar{x}) + \sum_{h=1}^{n-1}(x_{T-h} + x_{T+h} - 2\bar{x})\cos(h\omega)\right\} + o_p(n^{-1}).$$

These results show that outliers may substantially affect the inference performed on stationary models by revealing that there is information loss in the serial correlation dynamics of the process, which is translated into the parameter estimation process.

3. The autocovariance and spectral density robust functions

3.1. The autovariance function

Ma & Genton 2000 [20] proposed a scale covariance estimator which is based on $Q_n(\cdot)$, defined in the sequel, and on the following covariance identity

$$\text{cov}(X,Y) = \frac{1}{4ab}[\text{var}(aX+bY) - \text{var}(aX-bY)], \tag{3.1}$$

where X and Y are random variables, $a = \frac{1}{\sqrt{\text{var}(X)}}$ and $b = \frac{1}{\sqrt{\text{var}(Y)}}$ (Huber 2004 [12]).

Rousseeuw & Croux 1993 [29] proposed a robust scale estimator function $Q_n(\cdot)$ which is based on the τth order statistic of $\binom{n}{2}$ distances $\{|\eta_j - \eta_k|, j < k\}$, and can be written as

$$Q_n(\eta) = c \times \{|\eta_j - \eta_k|; j < k\}_{(\tau)}, \tag{3.2}$$

where $\eta = (\eta_1, \eta_2, \dots, \eta_n)'$, c is a constant used to guarantee consistency ($c = 2.2191$ for the normal distribution) and $\tau = \left\lfloor \frac{\binom{n}{2}+2}{4} \right\rfloor + 1$.

Based on identity (3.1) and on $Q_n(\cdot)$, Ma & Genton 2000 [20] proposed a highly robust estimator for the ACOVF:

$$\widehat{\gamma}_Q(h) = \frac{1}{4}\left[Q^2_{n-h}(\mathbf{u}+\mathbf{v}) - Q^2_{n-h}(\mathbf{u}-\mathbf{v})\right], \tag{3.3}$$

where $\mathbf{u}$ and $\mathbf{v}$ are vectors containing the initial $n-h$ and the final $n-h$ observations, respectively. The robust estimator for the autocorrelation function (ACF) is

$$\widehat{\rho}_Q(h) = \frac{Q^2_{n-h}(\mathbf{u}+\mathbf{v}) - Q^2_{n-h}(\mathbf{u}-\mathbf{v})}{Q^2_{n-h}(\mathbf{u}+\mathbf{v}) + Q^2_{n-h}(\mathbf{u}-\mathbf{v})}.$$

It can be shown that $|\widehat{\rho}_Q(h)| \leq 1$ for all h.

Influence Function and Breakdown Point

Influence Function (IF) is an important tool to understand the effect of the contamination of an outlier in any estimator. To define IF supposes that the empirical c.d.f. F_n of $x_1, ..x_n$, adequately normalized, converges. Following Huber 2004 [12], the influence function $x \to IF(x,T,F)$ is defined for a functional T at a distribution F and at point x as the limit

$$IF(x,T,F) = \lim_{\varepsilon \to 0+} \varepsilon^{-1}\{T(F + \varepsilon(\delta_x - F)) - T(F)\},$$

where δ_x is the Dirac distribution at x.

Breakdown Point (BP) indicates the largest proportion of outliers that the data may contain such that the estimator still gives some information about the distribution of the outlier-free data (Maronna, Martin & Yohai 2006 [21]). Rousseeuw & Croux 1993 [29] showed that the asymptotic BP of $Q_n(\cdot)$ is 50%, which means that the data can be contaminated by up to half of the observations with outliers and $Q_n(\cdot)$ will still yield sensible estimates.
The classical notion of sample BP of a scale estimator $S_n(\cdot)$ is given in Definition 3.1.

Definition 3.1. Let $\eta = (\eta_1, \eta_2, \ldots, \eta_n)'$ be a sample of size n. Let $\widetilde{\eta}$ be obtained by replacing any m observations of η by arbitrary values. The sample breakdown point of a scale estimator $S_n(\eta)$ is given by

$$\varepsilon_n^*(S_n(\eta)) = \max\left\{\frac{m}{n} : \sup_{\widetilde{\eta}} S_n(\widetilde{\eta}) < \infty \text{ and } \inf_{\widetilde{\eta}} S_n(\widetilde{\eta}) > 0\right\}.$$

The above BP definition holds for a scale estimator function of a time invariant random sample. As noted by Ma & Genton 2000 [20], in time series, the estimators are based on differences between observations apart by various time lag distances and usually have a BP with respect to these differences. Then, the time location of the outlier becomes important (see also, for example, Ledolter 1989 [17]). Therefore, the authors introduced the following definition of a temporal sample breakdown point of an autocovariance estimator $\hat{\gamma}_\eta(h)$ based on (3.1).

Definition 3.2. Let $\eta = (\eta_1, \eta_2, \ldots, \eta_n)'$ be a sample of size n and let $\widetilde{\eta}$ be obtained by replacing any m observations of η by arbitrary values. Denote by $\mathbb{I}_m$ a subset of size m of $\{1, 2, \ldots, n\}$. The temporal sample breakdown point of an autocovariance estimator $\hat{\gamma}_\eta(h)$ is given by

$$\begin{aligned}\varepsilon_n^{temp}(\widehat{\gamma}_\eta(h)) = \max\Bigg\{\frac{m}{n} &: \sup_{\mathbb{I}_m}\sup_{\widetilde{\eta}} S_{n-h}(\widetilde{\mathbf{u}} + \widetilde{\mathbf{v}}) < \infty, \inf_{\mathbb{I}_m}\inf_{\widetilde{\eta}} S_{n-h}(\widetilde{\mathbf{u}} + \widetilde{\mathbf{v}}) > 0, \\ &\sup_{\mathbb{I}_m}\sup_{\widetilde{\eta}} S_{n-h}(\widetilde{\mathbf{u}} - \widetilde{\mathbf{v}}) < \infty \text{ and } \inf_{\mathbb{I}_m}\inf_{\widetilde{\eta}} S_{n-h}(\widetilde{\mathbf{u}} - \widetilde{\mathbf{v}}) > 0\Bigg\},\end{aligned}$$

where $\widetilde{\mathbf{u}}$ and $\widetilde{\mathbf{v}}$ are derived from $\widetilde{\eta}$ as in (3.3).

Remark 3.3. The relation between the classical sample and the temporal sample breakdown points can be expressed by the following inequality (Ma & Genton 2000 [20]):

$$\frac{n-h}{2n}\varepsilon_n^*(\widehat{\gamma}_\eta(h)) \leq \varepsilon_n^{temp}(\widehat{\gamma}_\eta(h)) \leq \frac{1}{2}\varepsilon_n^*(\widehat{\gamma}_\eta(h)).$$

It then follows that since the sample breakdown point of the classical autocovariance estimator is zero, the temporal breakdown point of this estimator is also zero. This means that only one single outlier is enough to 'break' the estimator.

Ma & Genton 2000 [20] showed that the maximum temporal breakdown point of the highly robust autocovariance estimator is 25%, which is the highest possible breakdown point for an autocovariance estimator.

Results of the asymptotic properties of the robust aucovariance function for a Gaussian ARFIMA model are summarized as follows (see Lévy-Leduc et al. 2011 [19]).

Short-memory case

Let $\{X_t\}_{t\in\mathbb{Z}}$ be a stationary mean-zero Gaussian process given by Model 1.2 with $d = 0$, that is, the autocovariance function ($\gamma(h) = E(X_1 X_{h+1})$) of $\{X_t\}_{t\in\mathbb{Z}}$ satisfies

$$\sum_{h\geq 1} |\gamma(h)| < \infty.$$

The following theorems present the asymptotic behavior of the robust autocovariance estimator.

Theorem 3.4. *Let h be a non-negative integer. Under the assumption that the autocovariances are absolutely summable, the autocovariance estimator $\widehat{\gamma}_Q(h, X_{1:n}, \Phi)$ satisfies the following Central Limit Theorem:*

$$\sqrt{n}\,(\widehat{\gamma}_Q(h, X_{1:n}, \Phi) - \gamma(h)) \xrightarrow{d} \mathcal{N}(0, \check{\sigma}_h^2),$$

where

$$\check{\sigma}^2(h) = E[\psi^2(X_1, X_{1+h})] + 2\sum_{k\geq 1} E[\psi(X_1, X_{1+h})\psi(X_{k+1}, X_{k+1+h})] \tag{3.4}$$

where ψ is a function of $\gamma(h)$ and of IF (see, Theorem 4 in Lévy-Leduc et al. 2011 [19]).

Long-memory case

Now, let $d \neq 0$ in Model 1.2 and let $D = 1 - 2d$. The ACF behaves like

$$\gamma(h) = h^{-D} L(h),\ 0 < D < 1\ ,$$

where L is slowly varying at infinity and is positive for large h. Note that, for positive d, as previously stated, the ACF of the process is not absolutely summable.

Theorem 3.5. *Let h be a non negative integer. Then, $\widehat{\gamma}_Q(h, X_{1:n}, \Phi)$ satisfies the following limit theorems as n tends to infinity.*

- *If $D > 1/2$,*

$$\sqrt{n}\,(\widehat{\gamma}_Q(h, X_{1:n}, \Phi) - \gamma(h)) \xrightarrow{d} \mathcal{N}(0, \check{\sigma}^2(h))\ ,$$

where

$$\check{\sigma}^2(h) = \mathbb{E}[\psi^2(X_1, X_{1+h})] + 2\sum_{k\geq 1} \mathbb{E}[\psi(X_1, X_{1+h})\psi(X_{k+1}, X_{k+1+h})] \ ,$$

where ψ *is a function of* $\gamma(h)$ *and of IF (see, Theorems 4 and 5 in Lévy-Leduc et al. 2011 [19]).*

- *If* $D < 1/2$,

$$\beta(D)\frac{n^D}{\widetilde{L}(n)}\left(\widehat{\gamma}_Q(h, X_{1:n}, \Phi) - \gamma(h)\right) \xrightarrow{d} \frac{\gamma(0)+\gamma(h)}{2}(Z_{2,D}(1) - Z_{1,D}^2(1))$$

where $\beta(D) = \mathrm{B}((1-D)/2, D)$, B *denotes the Beta function, the processes* $Z_{1,D}(\cdot)$ *and* $Z_{2,D}(\cdot)$ *are defined by Equations 53 and 54, respectively, in Lévy-Leduc et al. 2011 [19], and*

$$\widetilde{L}(n) = 2L(n) + L(n+h)(1+h/n)^{-D} + L(n-h)(1-h/n)^{-D} \ . \tag{3.5}$$

Remark 3.6. For Model 1.2 with $1/4 < d < 1/2$, the robust autocovariance estimator $\widehat{\gamma}_Q(h, X_{1:n}, \Phi)$ has the same asymptotic behavior as the classical autocovariance estimator $\widehat{\gamma}_x(h)$.

Theories related to the use of the robust ACF function to obtain an spectral estimate are still opened questions. However, this was first empirically investigated by Fajardo et al. 2009 [7]. The authors considered a robust estimator of the spectral density based on the robust ACF function when the time series follows an ARFIMA Model. Their estimation method is discussed in the next sub-section.

3.2. The sample spectral function

The results discussed in the previous sections and the spectral representation of a stationary process justify the use of the robust ACF function in the calculus of an estimator of a spectral density.

As previously stated, for the stationary process $\{X_t\}_{t\in\mathbb{Z}}$, the spectral density is a real-valued function of the Fourier transform of the autocovariance function, that is,

$$f_X(\omega) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} \gamma_X(h) e^{-ih\omega} \tag{3.6}$$

where $\gamma_X(\cdot)$ is the autocovariance of the process.

Equation (3.6) suggests to replace $\gamma_X(\cdot)$ by its estimate to obtain an estimate of $f_X(\omega)$. The periodogram function is the classical tool to estimate the spectral function. Other variants of the periodogram are called smoothed window periodogram (see, for example, Priestley 1981 [25]). In the same direction, Fajardo et al. 2009 [7] suggested to use the robust autocovariance function as an estimator of the classical ACF to obtain a robust spectral function. Although the theoretical justification of this estimator is still an opened question, the authors have empirically shown that the robust spectral estimator can be an alternative method to estimate a time series with outliers. A robust spectral estimator is

$$I_Q(\omega) = \frac{1}{2\pi} \sum_{|h|<n} \kappa(h)\widehat{\gamma}_Q(h)\cos(h\omega), \tag{3.7}$$

where $\widehat{\gamma}_Q(h)$ is the sample autocovariance function given in (3.3) and $\kappa(h)$ is defined as

$$\kappa(h) = \begin{cases} 1, & |h| \leq M, \\ 0, & |h| > M. \end{cases}$$

$\kappa(h)$ is a particular case of the *lag window* functions used in classical spectral theory to obtain a consistent spectral estimator, and M is the truncation point which is a function of n, say $M = G(n)$, where $G(n)$ must satisfy $G(n) \to \infty$, $n \to \infty$, with $\frac{G(n)}{n} \to 0$. $G(n)$ is usually chosen to be $G(n) = n^{\beta}$, where $0 < \beta < 1$ (see, e.g. Priestley 1981 [25, pp. 433–437]). Note that, equivalently to the classical spectral estimation theories, other different *lag window* functions can be used to obtain a robust spectral estimator.

Since (3.7) does not have the same finite-sample properties as the periodogram, it is defined here as *robust truncated pseudo-periodogram*. For large h, the numbers of observations in the calculus of $\widehat{\gamma}_Q(h)$ are very small and, consequently, this function becomes very unstable. Then, to avoid these undesirable covariance estimates in the calculus of the estimator given in (3.7) justify the use of a truncation point M in the calculus of this sample function (see Fajardo et al. 2009 [7]). The authors suggested M that satisfies

$$M \leq h' = \min\left\{0 < h < n : \varepsilon_n^{temp}\left(\widehat{\gamma}_Q(h)\right) \leq \frac{m}{n}\right\} - 1.$$

4. Semiparametric estimation methods of d and empirical studies

The semiparametric estimation procedure based on the OLS estimator proposed by Geweke & Porter-Hudak 1983 [10](GPH) is considered. Since the GPH estimator is well-discussed in the literature, this method and its asymptotic statistical properties are briefly summarized as follows.

For a single realization $x_1, ..., x_n$ of $\{X_t\}_{t\in\mathbb{Z}}$, the GPH estimate of d is obtained from the regression equation

$$\log I_x(\omega_j) = a_0 - 2d\log\left[2\sin(\omega_j/2)\right] + \xi_j, j = 1, ..., m' \tag{4.1}$$

where ω_j is the Fourier frequency at j, m' is the bandwidth in the regression equation which has to satisfy $m' \to \infty$, $n \to \infty$, with $\frac{m'}{n} \to 0$ and $\frac{m'\log(m')}{n} \to 0$,

$a_0 = \log f_\eta(0) + \log \frac{f_\eta(\omega_j)}{f_\eta(0)} + \text{C}$, $\xi_j = \log \frac{I_x(\omega_j)}{f_X(\omega_j)} - C$ and $\text{C} = \varphi(1)$ ($\varphi(.)$ is the digamma function).

The GPH estimate of d is given by

$$d_{GPH} = (-0.5)\frac{\sum_{j=1}^{m'}(v_j - \bar{v})\log I_x(\omega_j)}{S_{vv}} \tag{4.2}$$

where $S_{vv} = \sum_{j=1}^{m'}(v_j - \overline{v})^2$, $v_j = \log\left\{4\sin^2(\omega_j/2)\right\}$.

Under some conditions, Hurvich, Deo & Brodsky 1998 [14] proved that the GPH-estimator is consistent for the memory parameter and asymptotically normal for Gaussian time series processes. The authors established that the optimal m' in (4.1) and (4.2) is of order $o(n^{4/5})$ and $(m')^{1/2}(d_{GPH} - d) \xrightarrow{d} N(0, \frac{\pi^2}{24})$.

To obtain a robust estimator of d, Fajardo et al. 2009 [7] proposed to replace in (4.1) the $\log I_x(\omega_j)$ by $\log I_Q(\omega_j)$ which gives the following OLS regression estimator

$$d_{GPHR} = -(0.5)\frac{\sum_{j=1}^{m'}(v_j - \bar{v})\log I_Q(\omega_j)}{S_{vv}}, \tag{4.3}$$

where S_{vv}, m' are defined as before and $I_Q(\omega)$ is the function given in (3.7). As previously mentioned, the asymptotical properties of d_{GPHR} still remains to be established. However, based on the following empirical investigation, the robust method seems to be a reasonable robust alternative method to estimate long-memory time series in the presence of additive outliers.

4.1. Numerical evaluation using the ARFIMA$(0, d, 0)$ model

The finite series were simulated from zero-mean ARFIMA models (Eq. 1.2) with $\{\epsilon_t\}_{t\in\mathbb{Z}}$, $t = 1, ..., n$, i.i.d. $N(0, 1)$. The models, parameters, sample sizes and empirical results are displayed in the following tables. The empirical mean, standard deviation (s.d.), bias and mean squared error (MSE) were obtained as a mean of 10.000 replications. The contaminated data were generated from Model 2.3 with $m = 1$, $p = 0.05$ for magnitude $\varpi = 10$ and bandwidth values for d_{GPH} and d_{GPHR} were computed for $\alpha = 0.7$ and truncation point $M = n^\beta$, $\beta = 0.7$. In the tables d_{GPH_c} and d_{GPHR_c} mean the estimates of d when the series has outliers. The simulations were carried out using the `Ox` matrix programming language (see `http://www.doornik.com`). The empirical study was divided into the following model properties: stationary and non-stationary processes.

Stationary model

Table 1 displays results for $d = 0.3, 0.45$ and $\alpha = \beta = 0.7$. From the table, it can be seen that when the series does not contain outliers, both estimators present similar behavior in the estimation of d, which is not a surprising result. However, the introduction of outliers in the series dramatically changes the performance of the classical estimator (GPH), in particular, it significantly underestimates the true parameter. On the other hand, in this scenario, the robust method (GPHR) seems to be not sensitive to outliers. Other cases were also simulated such as ARFIMA models with AR and MA parts and different values of p and ϖ. All cases indicated similar conclusions to the one given in Table 1. These are available upon request. Table 2 gives the estimates of d when different lag-windows are used to compute the robust periodogram estimator. The lag-windows are Parzen (P), Tukey-Hamming(TH) and Bartlett (B) and the fractional estimators were computed with the same bandwidths as in the previous case. The choice of the lag-window does not appear to be too important in the estimation of d since the estimates obtained from different lag-windows are, in general, numerically very close to each other. In other words, the estimates are not too sensitive to the choice of the lag-window. These lag-windows yield similarly accurate estimates compared to the one given in (3.7).

d	n		d_{GPH}	d_{GPH_c}	d_{GPHR}	d_{GPHR_c}
	100	mean	0.2988	0.1134	0.2584	0.2449
		s.d.	0.1735	0.1619	0.1558	0.1556
		bias	−0.0012	−0.1866	−0.0416	−0.0551
		MSE	0.0301	0.0610	0.0260	0.0272
	300	mean	0.3062	0.1007	0.2907	0.2837
0.30		s.d.	0.1005	0.0978	0.0926	0.0960
		bias	0.0062	−0.1993	−0.0093	−0.0163
		MSE	0.0101	0.0493	0.0087	0.0095
	800	mean	0.3003	0.1184	0.2949	0.2869
		s.d.	0.0679	0.0715	0.0573	0.0610
		bias	0.0003	−0.1816	−0.0051	−0.0131
		MSE	0.0046	0.0381	0.0033	0.0039
	100	mean	0.4561	0.1923	0.3975	0.3778
		s.d.	0.1722	0.1727	0.1506	0.1433
		bias	0.0061	−0.2577	−0.0525	−0.0722
		MSE	0.0297	0.0962	0.0254	0.0258
	300	mean	0.4594	0.2015	0.4329	0.4233
0.45		s.d.	0.0986	0.0976	0.1041	0.1013
		bias	0.0094	−0.2485	−0.0171	−0.0267
		MSE	0.0098	0.0713	0.0111	0.0110
	800	mean	0.4620	0.2306	0.4457	0.4349
		s.d.	0.0688	0.0809	0.0562	0.0576
		bias	0.0121	−0.2194	−0.0043	−0.0151
		MSE	0.0049	0.0547	0.0032	0.0035

Table 1: Simulation results; ARFIMA$(0, d, 0)$ model with $\alpha = \beta = 0.7$ and $\varpi = 0, 10$.

Non-stationary model

As is well-known, the GPH estimator has been widely used even for ARFIMA models with d in $(0.5, 1.0]$ (see, for example, Franco & Reisen 2007 [9], Hurvich & Ray 1995 [15],Olbermann, Lopes & Reisen 2006 [22], Phillips 2007 [24] among

uncontaminated series					
Parameter	n		d_P	d_{TH}	d_B
$d = 0.3$	100	mean	0.2699	0.2602	0.2459
		s.d.	0.1497	0.1575	0.1444
		bias	−0.0301	−0.0398	−0.0541
		MSE	0.0233	0.0264	0.0238
	300	mean	0.2880	0.2833	0.2857
		s.d.	0.1050	0.1037	0.0976
		bias	−0.0119	−0.0167	−0.0143
		MSE	0.0112	0.0110	0.0097
	800	mean	0.2985	0.2966	0.3001
		s.d.	0.0554	0.0584	0.0561
		bias	−0.0015	−0.0034	0.0001
		MSE	0.0031	0.0034	0.0031
contaminated series					
Parameter	n		d_P	d_{TH}	d_B
$d = 0.3$	100	mean	0.2504	0.2446	0.2419
		s.d.	0.1552	0.1482	0.1405
		bias	−0.0496	−0.0554	−0.0581
		MSE	0.0266	0.0250	0.0231
	300	mean	0.2806	0.2729	0.2796
		s.d.	0.1028	0.0925	0.0964
		bias	−0.0194	−0.0271	−0.0204
		MSE	0.0109	0.0093	0.0097
	800	mean	0.2934	0.2889	0.2928
		s.d.	0.0578	0.0606	0.0553
		bias	−0.0066	−0.0111	−0.0072
		MSE	0.0034	0.0038	0.0031

Table 2: Empirical results of d's estimators in ARFIMA$(0, d, 0)$ model using different lag-windows.

others).

Based on the theory discussed in the previous sections, the robust method can not be applied in a non-stationary time series. However, it may be interesting to verify if GPHR estimator is invariant to the first difference, i.e. estimative of the memory parameter based on the original data is equal to one plus the estimated d based on the differenced data.

Now, let Model 1.2 be defined with parameter $d^* = d + \kappa$, where $d \in (-0.5, 0.5)$, $\kappa > 0$, $\kappa \in \mathbb{Z}$. Then, Model 1.2, with zero-mean, becomes

$$X_t = (1 - B)^{-d^*} \eta_t, \quad t \in \mathbb{Z}. \tag{4.4}$$

Process given in (4.4) is non-stationary when $d^* \geq 0.5$; however, it is still persistent. For $d^* \in [0.5, 1.0)$ it is level-reverting in the sense that there is no long-run impact of an innovation on the value of the process. The level-reversion property no longer holds when $d^* \geq 1$. Note that when $d^* = 1$ the process is a random walk.

From Model 4.4 with $\kappa = 1$ and $p = q = 0$,

$$W_t = (1 - B) X_t, t \in \mathbb{Z},$$

is an $ARFIMA(0,d,0)$ process. Let $\hat{d}^*$ be the estimator of d^* and let $\hat{d}$ be the fractional estimator obtained from the differenced data. The main goal is to verify the equality $\hat{d}^* = \hat{d} + 1$ for uncontaminated and contaminated series. Based on the same simulation procedure previously described, series from Model 4.4 were generated and some cases are displayed in Table 3 (other cases are available upon request). Similar conclusions to the previous study are observed. Both estimators present equivalent performance when they are applied in the first difference of uncontaminated series. This suggests that both can be used in practical situations when dealing with non-stationary data. However, since the first difference does not eliminate the effect of an outlier, the estimates clearly indicate that caution has to be exercised when there is suspicion of outliers in the data. The GPH estimator presents poor performance in terms of bias (high positive bias) and MSE. In contrast to the GPH estimator, the GPHR method seems to be invariant to the first difference of non-stationary time series with outliers. This empirical study suggests that, in practical situations when dealing with non-stationary data with outliers, one solution is to apply the first difference in the series and then to estimate d with the robust estimator discussed in this paper.

Parameter	n		d_{GPH}	d_{GPH_c}	d_{GPHR}	d_{GPHR_c}
	300	mean	−0.2141	−0.5066	−0.1906	−0.2211
$d_X = 0.8, d_W = -0.2$		bias	0.0141	0.3066	−0.0094	0.0211
		s.d	0.1076	0.1469	0.1127	0.1421
		MSE	0.0118	0.1155	0.0128	0.0206
	800	mean	−0.1906	−0.4283	−0.2062	−0.2250
		bias	−0.0094	0.2283	0.0062	0.0251
		s.d	0.0630	0.0883	0.0851	0.1081
		MSE	0.0041	0.0599	0.0073	0.0123
	100	mean	−0.0048	−0.4166	−0.0449	−0.0871
		bias	0.0048	0.4166	0.0449	0.0871
		s.d	0.1763	0.2215	0.1620	0.1811
		MSE	0.0311	0.2226	0.0283	0.0404
	300	mean	−0.0122	−0.3230	−0.0273	−0.0426
$d_X = 1.0, d_W = 0.0$		bias	0.0122	0.3230	0.0273	0.0426
		s.d	0.1076	0.1296	0.1094	0.1277
		MSE	0.0117	0.1211	0.0127	0.0181
	800	mean	0.0059	−0.2181	−0.0107	−0.0222
		bias	−0.0059	0.2181	0.0107	0.0222
		s.d	0.0648	0.0823	0.0629	0.0909
		MSE	0.0042	0.0544	0.0041	0.0088

Table 3: Empirical results: ARFIMA$(0,d,0)$ model with differenced data and $\omega = 0, 10$.

5. Application

IGP-DI is the general price index with domestic availability and is calculated by Fundação Getúlio Vargas, Brazil. The series comprises monthly observations from

August 1994 to April 2011 (total of 201 observations). The series and its ACF are displayed in Figure 1. The observations of the months February 1999 (4.44%), October 2002 (4.21%) and November 2002 (5.84%) are possibly outliers. Looking at the plots in Figure 1, these suggest that the series is stationary and possess long-memory behavior. From the data and using the methodologies previously discussed, the parameter d is estimated and the results are displayed in Table 4. For this application, the estimates of d were computed from the original data (OD) and from the modified data (MD), where the observations of February 1999, October 2002 and November 2002 were replaced by the sample mean of the series. This analysis is a simple exercise to verify the robustness of the estimators in a real application and, also, to investigate whether the data contains outliers. The d' estimates of OD and MD series are given, respectively, on the left and right sides of Table 4. These estimates were calculated using different bandwidths in (4.2)($m' = n^{\alpha}$) and β was fixed as in the simulation study. In both series, for a fixed α, the robust methods present similar results. The estimates maintain the same empirical property across the bandwidth values. In contrast to the robust methods, the classical GPH estimator gives estimates that dramatically change from OD to MD data, showing that the observations replaced by the mean are possible atypical data.

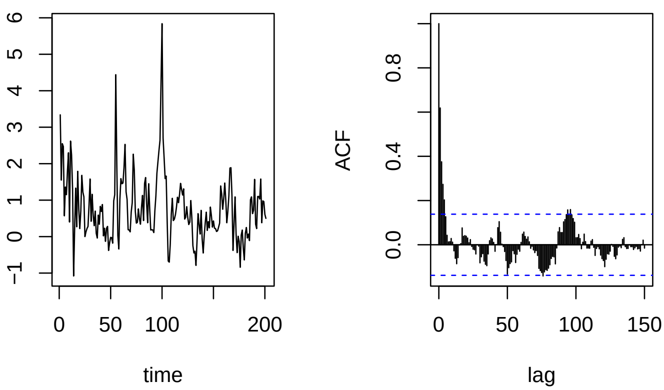


Figure 1: IGP-DI series and its sample autocorrelation function: period from Aug/94 to Apr/11.

6. Concluding remarks and future direction

This paper investigates the effect of outliers in the estimation of the fractional parameter d in the ARFIMA(p, d, q) model and, also, discusses the asymptotical and empirical properties of the robust autocovariance and spectral estimators, previously given in Fajardo et al. 2009 [7] and Lévy-Leduc et al. 2011 [19], for the case of time series with short and long-memory properties. These studies support the use

	Original time series				Modified time series			
Estimator	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$
d_{GPH}	0.0757	0.1205	0.3431	0.3759	0.3110	0.3116	0.3713	0.3875
	(0.3417)	(0.1869)	(0.1389)	(0.0888)	(0.1586)	(0.1077)	(0.0909)	(0.0683)
d_{GPHR_P}	0.1802	0.2335	0.2269	0.2397	0.1630	0.2077	0.2078	0.2230
	(0.0857)	(0.0745)	(0.0469)	(0.0331)	(0.0782)	(0.0603)	(0.0385)	(0.0251)
$d_{GPHR_{TH}}$	0.1718	0.1919	0.2125	0.2379	0.1545	0.1782	0.1968	0.2231
	(0.0742)	(0.0508)	(0.0303)	(0.0210)	(0.0673)	(0.0436)	(0.0259)	(0.0170)
d_{GPHR_B}	0.1522	0.1788	0.2047	0.2327	0.1379	0.1667	0.1896	0.2181
	(0.0641)	(0.0433)	(0.0262)	(0.0183)	(0.0586)	(0.0378)	(0.0227)	(0.0151)
d_{GPHR}	0.1662	0.2628	0.2454	0.2285	0.1500	0.2211	0.2215	0.2228
	(0.0862)	(0.0995)	(0.0671)	(0.0436)	(0.0794)	(0.0717)	(0.0511)	(0.0328)

Table 4: Estimates of d: IGP-DI data, period from Aug/94 to Apr/11.

of the robust estimators to estimate the long-memory parameter when Gaussian long-memory time series are contaminated with additive outliers. Non-stationary time series with outliers are also studied and the investigation reveals that the robust method can be used as an alternative estimation procedure in time series with fractional differences. As previously stated, the asymptotical properties of the robust estimator under the study still remain to be investigated. The robust ACF method discussed here has also been used in other contexts such as in the estimation of periodic process (Sarnaglia, Reisen & Lévy-Leduc 2010 [30]) and in seasonal ARFIMA processes (this is one of the current research of the authors).

Acknowledgements. The authors gratefully acknowledge partial financial supports from CNPq-Brazil and FAPES.

References

[1] J. Beran, On a class of M-estimators for gaussian long-memory models, *Biometrika*, 81:755–766, 1994.

[2] Wai-Sum Chan, A note on time series model specification in the presence outliers, *Journal of Applied Statistics*, 19:117–124, 1992.

[3] Wai-Sum Chan, Outliers and financial time series modelling: a cautionary note, *Mathematics and Computers in Simulation*, 39:425–430, 1995.

[4] I. Chang, G. C. Tiao and C. Chen, Estimation of time series parameters in presence of outliers, *Technometrics*, 30:1936–204, 1988.

[5] C. Chen and Lon-Mu Liu, Joint estimation of model parameters and outlier effects in time series, *Journal of the American Statistical Association*, 88:284–297, 1993.

[6] P. Doukhan, G. Oppenheim and M. Taqqu, *Theory and Applications of Long-Range Dependence*, Birkhäuser, 2003.

[7] F. Fajardo, V. A. Reisen and F. Cribari-Neto, Robust estimation in long-memory processes under additive outliers, *Journal of Statistical Planning and Inference*, 139:2511–2525, 2009.

[8] A. J. Fox, Outliers in time series, *Journal of the Royal Statistical Society*, 34(B):350–363, 1972.

[9] G. C. Franco and V. A. Reisen, Bootstrap approaches and confidence intervals for stationary and non-stationary long-range dependence processes, *Physica. A.*, 375:546–562, 2007.

[10] J. Geweke and S. Porter-Hudak, The estimation and application of long memory time series model, *Journal of Time Series Analysis*, 4:221–238, 1983.

[11] J. R. Hosking, Fractional differencing, *Biometrika*, 68:165–176, 1981.

[12] P. J. Huber, *Robust Statistics*, John Wiley & Sons, third edition, 2004.

[13] C. M. Hurvich and K. I. Beltrão, Asymptotics for low-frequency ordinates of the periodogram of a long-memory time series, *Journal of Time Series Analysis*, 14(5):455–472, 1993.

[14] C. M. Hurvich, R. Deo and J. Brodsky, The mean square error of geweke and porter-hudak's estimator of the memory parameter of a long-memory time series, *Journal of Time Series Analysis*, 19(1):19–46, 1998.

[15] C. M. Hurvich and B. K. Ray, Estimation of the memory parameter for nonstationary or noninvertible fractionally integrated processes, *Journal of Time Series Analysis*, 16(1):17–42, 1995.

[16] H. R. Künsch, Discrimination between monotonic trends and long-range dependence, *Journal of Applied Probability*, 23:1025–1030, 1986.

[17] J. Ledolter, The effect of additive outliers on the forecast from arma models, *International Journal of Forecasting*, 5:231–240, 1989.

[18] C. Lévy-Leduc, H. Boistard, E. Moulines, M. Taqqu and V. A. Reisen, Large sample behaviour of some well-known robust estimators under long-range dependence, *Statistics (Berlin)*, 45:59–71, 2011.

[19] C. Lévy-Leduc, H. Boistard, E. Moulines, M. Taqqu and V. A. Reisen, Robust estimation of the scale and the autocovariance function of gaussian short and long-range dependent processes, *Journal of Time Series Analysis*, 32:135–156, 2011.

[20] Y. Ma and M. Genton, Highly robust estimation of the autocovariance function, *Journal of Time Series Analysis*, 21:663–684, 2000.

[21] R. Maronna, R. D. Martin and V. Yohai, *Robust statistics*, John Wiley & Sons, 2006.

[22] B. P. Olbermann, S. R. Lopes and V. A. Reisen, Invariance in the first difference in arfima models, *Computational Statistics*, 21:445–461, 2006.

[23] W. Palma, *Long-Memory Time Series: Theory and Methods*, Wiley-Interscience, 2007.

[24] P. Phillips, Unit root log periodogram regression, *Journal of Econometrics*, 138:104–124, 2007.

[25] M. B. Priestley, *Spectral Analysis and Time Series*, Academic Press, 1981.

[26] V. A. Reisen, Estimation of the fractional difference parameter in the ARIMA(p, d, q) model using the smoothed periodogram, *Journal of Time Series Analysis*, 15:335–350, 1994.

[27] P. M. Robinson, Gaussian semiparametric estimation of long range dependence, *The Annals of Statistics*, 23:1630–1661, 1995.

[28] P. M. Robinson, Log-periodogram regression of time series with long range dependence, *The Annals of Statistics*, 23:1048–1072, 1995.

[29] P. J. Rousseeuw and C. Croux, Alternatives to the median absolute deviation, *Journal of the American Statistical Association*, 88:1273–1283, 1993.

[30] A. J. Q. Sarnaglia, V. A. Reisen and C. Lévy-Leduc, Robust estimation of periodic autoregressive processes in the presence of additives outliers, *Journal of Multivariate Analysis*, 2:2168–2183, 2010.

Annales Mathematicae et Informaticae
39 *(2012) pp. 225–239*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Pasting of two one-dimensional diffusion processes*

Roman Shevchuk

Ivan Franko National University of L'viv, Faculty of Mechanics and Mathematics,
Department of Higher Mathematics, Ukraine
r.v.shevchuk@gmail.com

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

By the method of classical potential theory we obtain an integral representation of the two-parameter semigroup of operators that describes an inhomogeneous Feller process on a line that is a result of pasting together two diffusion processes with the nonlocal boundary condition of non-transversal type.

Keywords: Feller semigroup, diffusion process, boundary condition of Feller-Wentzell

MSC: Primary 60J60

1. Introduction

Let $D_i = \{x \in \mathbb{R} : (-1)^i x > 0\}$, $i = 1, 2$, be the two domains on the line $\mathbb{R}$ with the common boundary $S = \{0\}$ and the closures $\overline{D}_i = D_i \cup \{0\}$, and let $T > 0$ be fixed. If Γ is $\overline{D}_i$ or $\mathbb{R}$, then we denote by $C_b(\Gamma)$ a Banach space of all functions $\varphi(x)$, real-valued, bounded and continuous on Γ with the norm

$$\|\varphi\| = \sup_{x \in \Gamma} |\varphi(x)|,$$

and by $C_2(\Gamma)$ the set of all functions φ, bounded and uniformly continuous on Γ together with their first- and second-order derivatives. Let φ_i be the restriction of any function $\varphi \in C_b(\mathbb{R})$ to D_i.

*Work partially supported by the State fund for fundamental researches of Ukraine and the Russian foundation for basic research, grant No. F40.1/023.

Assume that an inhomogeneous diffusion process is given in $D_i,\ i=1,2$, and it is generated by a second-order differential operator $A_s^{(i)},\ s\in[0,T]$, with the domain of definition $C_2(\overline{D}_i)$:

$$A_s^{(i)}\varphi_i(x)=\frac{1}{2}b_i(s,x)\frac{d^2\varphi_i(x)}{dx^2}+a_i(s,x)\frac{d\varphi_i(x)}{dx},\quad i=1,2, \tag{1.1}$$

where the diffusion coefficient $b_i(s,x)$ and the drift coefficient $a_i(s,x)$ satisfy the conditions:

1) there exist the constants b and B such that $0<b\le b_i(s,x)\le B$ for all $(s,x)\in[0,T]\times\overline{D}_i$;

2) the function $a_i(s,x)$ is bounded on $[0,T]\times\overline{D}_i$;

3) for all $s,s'\in[0,T],\ x,x'\in\overline{D}_i$ the next inequalities hold:

$$|b_i(s,x)-b_i(s',x')|\le c\left(|s-s'|^{\frac{\alpha}{2}}+|x-x'|^{\alpha}\right),$$

$$|a_i(s,x)-a_i(s',x')|\le c\left(|s-s'|^{\frac{\alpha}{2}}+|x-x'|^{\alpha}\right),$$

where c and α are the positive constants, $0<\alpha<1$.

Assume also that at the zero point the boundary operator L_s is defined by the formula

$$L_s\varphi(0)=\gamma(s)\varphi(0)+\int\limits_{D_1\cup D_2}[\varphi(0)-\varphi(y)]\mu(s,dy),\quad s\in[0,T], \tag{1.2}$$

where the function γ and the measure μ satisfy the following conditions:

a) the function $\gamma(s)$ is nonegative and Hölder continuous, with exponent $\frac{1+\alpha}{2}$, on [0, T];

b) $\mu(s,\cdot)$ is a nonnegative measure on $D_1\cup D_2$ such that $0<\mu(s,D_1\cup D_2)<\infty$, $s\in[0,T]$, and for all the functions f, bounded and measurable in $\mathbb{R}$, the integrals

$$G_f^{(i)}(s)=\int\limits_{D_i}f_i(y)\mu(s,dy),\quad i=1,2,$$

are Hölder continuous, with exponent $\frac{1+\alpha}{2}$, on $[0,T]$.

Note that the operator L_s is a particular case of Feller-Wentzell boundary operator ([1, 2]) which describes the behavior of a diffusion particle at the time when it reaches the origin. Its terms $\gamma(s)\varphi(0)$ and $\int\limits_{D_1\cup D_2}[\varphi(0)-\varphi(y)]\mu(s,dy)$ are supposed to correspond to the absorption phenomenon, and the inward jump phenomenon from the boundary, respectively.

The problem is to clarify the question of existence of the two-parameter semigroup of operators $T_{st},\ 0 \le s < t \le T$, describing the inhomogeneous Feller process in $\mathbb{R}$ such that in the domains D_1 and D_2 it coincides with the given diffusion processes generated by $A_s^{(1)}$ and $A_s^{(2)}$, respectively, and its behavior at the point zero is determined by the boundary condition

$$L_s\varphi(0) = 0. \tag{1.3}$$

This problem is also often called a problem of pasting together two diffusion processes on a line or a problem of the mathematical modeling of the diffusion phenomenon on a line with a membrane placed in a fixed point (see [3, 4]).

The investigation of the problem formulated above is performed by the analytical methods. Such an approach ([3]–[10]) permits to determine the required operator family by means of the solution of the corresponding problem of conjugation for a linear parabolic equation of the second order with variable coefficients, discontinuous at the zero point. This problem is to find a function $u(s,x,t) = T_{st}\varphi(x)$ satisfying the following conditions:

$$\frac{\partial u(s,x,t)}{\partial s} + A_s^{(i)}u(s,x,t) = 0, \quad 0 \le s < t \le T,\ x \in D_i,\ i = 1,2, \tag{1.4}$$

$$\lim_{s\uparrow t} u(s,x,t) = \varphi(x), \quad x \in D_1 \cup D_2, \tag{1.5}$$

$$u(s,0-,t) = u(s,0+,t), \quad 0 \le s < t \le T, \tag{1.6}$$

$$L_s u(s,0,t) = 0, \quad 0 \le s < t \le T, \tag{1.7}$$

where $\varphi \in C_b(\mathbb{R})$ is the given function. As we see, the condition (1.6) is the consequence of the Feller property of the required semigroup T_{st}, and the equality (1.7) corresponds to the non-transversal nonlocal boundary condition of Feller-Wentzell (1.2), (1.3). Note that in the transversal case (i.e., when the boundary condition of Feller-Wentzell includes the local terms of the orders higher than the order of the nonlocal one) the conjugation problem (1.4)–(1.7) was studied in [10] (cf. also [7, 8]).

A classical solvability of the problem (1.4)–(1.7) is established by the boundary integral equations method with the use of the ordinary parabolic simple-layer potentials that are constructed using the fundamental solutions of the uniformly parabolic operators. Application of this method permits us not only to prove the existence of the solution of the problem (1.4)–(1.7), but also to obtain its integral representation. It is necessary to note that we derived a generalization of the corresponding result obtained earlier in [6], where a similar problem was analyzed for the case of homogeneous diffusion processes. Furthermore, the boundary condition (1.3) considered there, had no term corresponding to the termination of the process at the zero point. The present paper can be also treated as a generalization of the work [9] devoted to construction of the two-parameter Feller semigroup that describes an inhomogeneous diffusion process on a half-line with the non-transversal nonlocal boundary condition of Feller-Wentzell.

We should also mention the works [11, 12, 13], where the problem of constructing of mathematical models of diffusion processes in mediums with membranes was studied by the methods of stochastic analysis.

2. Auxiliary propositions

Consider the Kolmogorov backward equations (1.4) ($i = 1, 2$). Assume that their coefficients $a_i(s, x)$ and $b_i(s, x)$ are defined on $[0, T] \times \mathbb{R}$ and in this domain they satisfy conditions 1)–3). These conditions imply the existence of the fundamental solutions of equations (1.4) in the domain $[0, T] \times \mathbb{R}$, i.e., the existence of the functions $G_i(s, x, t, y)$ defined for $0 \le s < t \le T,\ x, y \in \mathbb{R}$ such that:

- they are jointly continuous;
- for fixed $t \in (0, T],\ y \in \mathbb{R}$ they satisfy equations (1.4);
- for any function $\varphi \in C_b(\mathbb{R})$ and any $t \in (0, T],\ x \in \mathbb{R}$

$$\lim_{s \uparrow t} \int_{\mathbb{R}} G_i(s, x, t, y)\varphi(y)dy = \varphi(x).$$

Recall that (see [3, Ch. II, §2], [5, Addendum, §6], [14, Ch. IV, §§11–13]) the functions $G_i(s, x, t, y)$ are nonnegative, continuously differentiable with respect to s, twice continuously differentiable with respect to x and for $0 \le s < t \le T,\ x, y \in \mathbb{R}$ the following estimations hold:

$$|D_s^r D_x^p G_i(s, x, t, y)| \le c(t - s)^{-\frac{1+2r+p}{2}} \exp\left\{-h\frac{(y - x)^2}{t - s}\right\}, \tag{2.1}$$

where r and p are the nonnegative integers such that $2r + p \le 2$; D_s^r is the partial derivative with respect to s of order r; D_x^p is the partial derivative with respect to x of order p; c, h are positive constants[1]. Furthermore, $G_i(s, x, t, y)$ are represented as

$$G_i(s, x, t, y) = Z_{i0}(s, y - x, t, y) + Z_{i1}(s, x, t, y),$$

where

$$Z_{i0}(s, x, t, y) = [2\pi b_i(t, y)(t - s)]^{-\frac{1}{2}} \exp\left\{-\frac{(y - x)^2}{2b_i(t, y)(t - s)}\right\},$$

and the functions $Z_{i1}(s, x, t, y)$ satisfy the inequalities

$$|D_s^r D_x^p Z_{i1}(s, x, t, y)| \le c(t - s)^{-\frac{1+2r+p-\alpha}{2}} \exp\left\{-h\frac{(y - x)^2}{t - s}\right\}, \tag{2.2}$$

[1] We will subsequently denote various positive constants by the same symbol c (or h).

where $0 \le s < t \le T,\ x, y \in \mathbb{R},\ 2r + p \le 2,\ \alpha$ is the constant in 3).

Given the fundamental solutions $G_i,\ i = 1, 2$, we define the parabolic potentials that will be used to solve the problem (1.4)-(1.7), namely the Poisson potentials

$$u_{i0}(s, x, t) = \int_{\mathbb{R}} G_i(s, x, t, y)\varphi(y)dy, \quad 0 \le s < t \le T,\ x \in \mathbb{R},$$

and the simple-layer potentials

$$u_{i1}(s, x, t) = \int_s^t G_i(s, x, \tau, 0)V_i(\tau, t, \varphi)d\tau, \quad 0 \le s < t \le T,\ x \in \mathbb{R}, \tag{2.3}$$

where φ is the function in (1.5), and $V_i(s, t, \varphi),\ i = 1, 2$, are arbitrary functions, continuous in $0 \le s < t \le T$ for which the integrals on the right side of (2.3) exist. Note that (see [3, Ch. II, §3], [14, Ch. IV]) the functions $u_{i0},\ u_{i1},\ i = 1, 2$, are continuous in $0 \le s < t \le T,\ x \in \mathbb{R}$ and satisfy the equations (1.4) in the domains $(s, x) \in [0, t) \times \mathbb{R},\ (s, x) \in [0, t) \times (D_1 \cup D_2)$, respectively, and the initial conditions

$$\lim_{s \uparrow t} u_{i0}(s, x, t) = \varphi(x), \quad x \in \mathbb{R},$$
$$\lim_{s \uparrow t} u_{i1}(s, x, t) = 0, \quad x \in D_1 \cup D_2.$$

Furthermore, for the potentials $u_{i0},\ i = 1, 2$, the following estimations are valid:

$$|D_s^r D_x^p u_{i0}(s, x, t)| \le c\|\varphi\|(t - s)^{-\frac{2r+p}{2}}, \tag{2.4}$$

where $0 \le s < t \le T,\ x, y \in \mathbb{R},\ 2r + p \le 2$.

We will also use the next lemma.

Lemma 2.1. *Let $Q_f(s),\ s \in [0, T]$ be a family of linear functionals defined on $C_b(\mathbb{R})$ such that for all $f \in C_b(\mathbb{R})$ the functions $Q_f(s)$ are Hölder continuous with the same exponent $\beta \in (0, 1)$ on a closed interval $[0, T]$. Then for every $M > 0$ there exist a common constant $c > 0$ such that for all the functions $f \in C_b(\mathbb{R})$, bounded by M and for all $s,\ s' \in [0, T]$ the inequality*

$$|Q_f(s) - Q_f(s')| \le c|s - s'|^\beta$$

holds.

Proof. $f \mapsto |s - s'|^{-\beta} (Q_f(s) - Q_f(s'))$, for $s \ne s' \in [0, T]$ is a pointwise bounded family of linear functionals, hence it is uniformly bounded, which is the statement. □

3. Parabolic conjugation problem

In this section by the boundary integral equations method we establish the classical solvability of the problem (1.4)–(1.7).

Theorem 3.1. *Assume that the coefficients of the operators $A_s^{(i)}$, $i = 1, 2$, the function γ and the measure μ satisfy conditions 1)-3) and a), b). Then for any function $\varphi \in C_b(\mathbb{R})$ the problem (1.4)–(1.7) has a unique solution*

$$u(s, x, t) \in C^{1,2}([0, t) \times D_1 \cup D_2) \cap C([0, t) \times \mathbb{R}).$$

Furthermore,

$$|u(s, x, t)| \leq c\|\varphi\|, \quad 0 \leq s < t \leq T, \tag{3.1}$$

and this solution is represented as

$$u(s, x, t) = u_{i0}(s, x, t) + u_{i1}(s, x, t), \quad x \in \overline{D}_i,\ i = 1, 2, \quad 0 \leq s < t \leq T, \tag{3.2}$$

where a pair of functions (V_1, V_2) in (u_{11}, u_{21}) is a solution of some system of Volterra integral equations of the second kind.

Proof. We find a solution of the problem (1.4)-(1.7) of the form (3.2) with the unknown functions V_i to be determined. Without loss of generality we may assume that

$$\mu(s, D_1 \cup D_2) \equiv 1.$$

Therefore, the condition (1.7) reduces to

$$(\gamma(s) + 1)u(s, 0, t) - \int\limits_{D_1 \cup D_2} u(s, y, t)\mu(s, dy) = 0, \quad 0 \leq s < t \leq T. \tag{3.3}$$

If we substitute (3.2) into (3.3) then, upon using the relation (1.6), we get the following system of Volterra integral equations of the first kind for V_i:

$$\Phi_i(s, t, \varphi) = (\gamma(s) + 1) \int\limits_s^t G_i(s, 0, \tau, 0)V_i(\tau, t, \varphi)d\tau -$$

$$- \sum_{j=1}^{2} \int\limits_s^t \left(\int\limits_{D_j} G_j(s, y, \tau, 0)\mu(s, dy) \right) V_j(\tau, t, \varphi)d\tau, \quad i = 1, 2, \tag{3.4}$$

where

$$\Phi_i(s, t, \varphi) = \sum_{j=1}^{2} \int\limits_{D_j} u_{j0}(s, y, t)\mu(s, dy) - (\gamma(s) + 1)u_{i0}(s, 0, t), \quad i = 1, 2.$$

Now we have to reduce (3.4) to an equivalent system of Volterra integral equations of the second kind. For this purpose we consider the Holmgren's operator

$$\mathcal{E}(s,t)F = \sqrt{\frac{2}{\pi}}\frac{d}{ds}\int\limits_s^t (\rho - s)^{-\frac{1}{2}} F(s,t,\varphi)d\rho, \quad 0 \le s < t \le T$$

and apply it to the both sides of each equation in (3.4). After some straightforward simplifications, we get

$$\mathcal{E}(s,t)\Phi_i = -\frac{V_i(s,t,\varphi)}{\sqrt{b_i(s,0)}} + \sqrt{\frac{2}{\pi}}\frac{d}{ds}\int\limits_s^t \left(I_i^{(1)}(s,\tau) + \sqrt{\frac{\pi}{2b_i(\tau,0)}} \cdot \gamma(s) \right) V_i(\tau,t,\varphi)d\tau -$$

$$- \sqrt{\frac{2}{\pi}}\frac{d}{ds}\sum_{j=1}^{2}\int\limits_s^t I_j^{(2)}(s,\tau)V_j(\tau,t,\varphi)d\tau, \quad i = 1,2, \tag{3.5}$$

where

$$I_i^{(1)}(s,\tau) = \frac{1}{\sqrt{2\pi b_i(\tau,0)}}\int\limits_s^\tau (\rho - s)^{-\frac{1}{2}}(\tau - \rho)^{-\frac{1}{2}}(\gamma(\rho) - \gamma(s))d\rho +$$

$$+ \int\limits_s^\tau (\rho - s)^{-\frac{1}{2}}(\gamma(\rho) + 1)Z_{i1}(\rho,0,\tau,0)\Big] d\rho, \quad i = 1,2,$$

$$I_i^{(2)}(s,\tau) = \int\limits_s^\tau (\rho - s)^{-\frac{1}{2}}d\rho\int\limits_{D_i} G_i(s,y,\tau,0)\mu(\rho,dy), \quad i = 1,2.$$

In view of the properties a), b) of the function γ and the measure μ, respectively, as well as the inequalities (2.1), (2.2), it is easy to verify that

$$\lim_{s\uparrow\tau} I_i^{(1)}(s,\tau) = 0, \quad \lim_{s\uparrow\tau} I_i^{(2)}(s,\tau) = 0, \quad i = 1,2.$$

Hence (3.5) can be reduced to the following system of Volterra integral equations of the second kind:

$$V_i(s,t,\varphi) = \sum_{j=1}^{2}\int\limits_s^t K_{ij}(s,\tau)V_j(\tau,t,\varphi)d\tau + \Psi_i(s,t,\varphi), \quad i = 1,2, \tag{3.6}$$

where

$$K_{ii}(s,\tau) = \frac{r_i(s)}{2\sqrt{2\pi b_i(\tau,0)}}\int\limits_s^\tau (\rho - s)^{-\frac{3}{2}}(\tau - \rho)^{-\frac{1}{2}}(\gamma(\rho) - \gamma(s))d\rho +$$

$$+ r_i(s)\frac{d}{ds}\int\limits_s^\tau(\rho-s)^{-\frac{1}{2}}\left[(\gamma(\rho)+1)Z_{i1}(\rho,0,\tau,0)-\int\limits_{D_i}G_i(\rho,y,\tau,0)\mu(\rho,dy)\right]d\rho, i=1,2,$$

$$K_{ij}(s,\tau)=-r_i(s)\frac{d}{ds}\int\limits_s^\tau(\rho-s)^{-\frac{1}{2}}d\rho\int\limits_{D_j}G_j(\rho,y,\tau,0)\mu(\rho,dy),\quad i,j=1,2,\ i\neq j,$$

$$\Psi_i(s,t,\varphi)=-r_i(s)\sqrt{\frac{\pi}{2}}\mathcal{E}(s,t)\Phi_i,\quad r_i(s)=\frac{1}{\gamma(s)+1}\sqrt{\frac{2b_i(s,0)}{\pi}},\quad i=1,2.$$

Let us show that there exist a solution of the system of equations (3.6) which can be obtained by the method of successive approximations

$$V_i(s,t,\varphi)=\sum_{k=0}^{\infty}V_i^{(k)}(s,t,\varphi),\quad 0\leq s<t\leq T,\quad i=1,2, \tag{3.7}$$

where

$$V_i^{(0)}(s,t,\varphi)=\Psi_i(s,t,\varphi),$$
$$V_i^{(k)}(s,t,\varphi)=\sum_{j=1}^{2}\int\limits_s^t K_{ij}(s,\tau)V_j^{(k-1)}(\tau,t,\varphi)d\tau,\quad k=1,2,\ldots.$$

For this purpose, we have first to estimate the functions Ψ_i and the kernels K_{ij} in (3.6).

Consider the functions $\Psi_i(s,t,\varphi)$. Calculating the derivatives on the right side of their expressions, we obtain ($i=1,2$):

$$\Psi_i(s,t,\varphi)=r_i(s)\Phi_i(s,t,\varphi)(t-s)^{-\frac{1}{2}}$$
$$-\frac{r_i(s)}{2}\int\limits_s^t(\rho-s)^{-\frac{3}{2}}\left(\Phi_i(\rho,t,\varphi)-\Phi_i(s,t,\varphi)\right)d\rho. \tag{3.8}$$

Denote by Ψ_{i1} and Ψ_{i2} the first and second terms in (3.8), respectively. Using the estimation

$$|\Phi_i(s,t,\varphi)|\leq c\|\varphi\|, \tag{3.9}$$

that follows easily from the inequalities (2.4) (when $r=p=0$), we find that

$$|\Psi_{i1}(s,t,\varphi)|\leq c\|\varphi\|(t-s)^{-\frac{1}{2}}. \tag{3.10}$$

In order to estimate $\Psi_{i1}(s,t,\varphi)$ we consider first the increments $\Phi_i(\rho,t,\varphi)-\Phi_i(s,t,\varphi)$ and write them in the form

$$\Phi_i(\rho,t,\varphi)-\Phi_i(s,t,\varphi)=N_{i1}(s,\rho,t,\varphi)+N_2(s,\rho,t,\varphi),$$

where

$$N_{i1} = \sum_{j=1}^{2} \int_{D_j} [u_{j0}(\rho, y, t) - u_{j0}(s, y, t)]\mu(\rho, dy) - (\gamma(s) + 1)[u_{i0}(\rho, 0, t) - u_{i0}(s, 0, t)], \tag{3.11}$$

$$N_2 = \sum_{j=1}^{2} \int_{D_j} u_{j0}(s, y, t)(\mu(\rho, dy) - \mu(s, dy)).$$

Expressing by the Lagrange formula the increments $u_{j0}(\rho, y, t) - u_{j0}(s, y, t)$, $j = 1, 2$, and $u_{i0}(\rho, 0, t) - u_{i0}(s, 0, t)$ in (3.11) in terms of the values of their derivatives at the intermediate points and then using the inequalities (2.4), after some straightforward simplifications, we deduce that

$$|N_{i1}(s, \rho, t, \varphi)| \le c\|\varphi\|(t - \rho)^{-1}(\rho - s), \quad 0 \le s < \rho < t \le T. \tag{3.12}$$

Let us now estimate N_2. Note that $u_{j0}(s, y, t)$, $j = 1, 2$, as functions of y, belong to a class $C_b(\mathbb{R})$ and are bounded by $M = \|\varphi\|$. Hence, by Lemma 1,

$$\left| \int_{D_j} u_{j0}(s, y, t)(\mu(\rho, dy) - \mu(s, dy)) \right| \le c\|\varphi\|(\rho - s)^{\frac{1+\alpha}{2}},$$

and hence,

$$|N_2(s, \rho, t, \varphi)| \le c\|\varphi\|(\rho - s)^{\frac{1+\alpha}{2}}, \quad 0 \le s < \rho < t \le T. \tag{3.13}$$

Combining (3.12) and (3.13), we obtain

$$|\Phi_i(\rho, t, \varphi) - \Phi_i(s, t, \varphi)| \le c\|\varphi\| \left[(t - \rho)^{-1}(\rho - s) + (\rho - s)^{\frac{1+\alpha}{2}}\right]. \tag{3.14}$$

Further, using the inequalities (3.9) and (3.14), we get

$$|\Psi_{i2}(s, t, \varphi)| \le c\|\varphi\| \int_{s}^{\frac{s+t}{2}} \left[\left(t - \frac{s+t}{2}\right)^{-1} (\rho - s)^{-\frac{1}{2}} + (\rho - s)^{-1+\frac{\alpha}{2}} \right] d\rho$$
$$+ c\|\varphi\| \int_{\frac{s+t}{2}}^{t} (\rho - s)^{-\frac{3}{2}} d\rho \le c\|\varphi\|(t - s)^{-\frac{1}{2}}. \tag{3.15}$$

Combining (3.10) and (3.15), we conclude that

$$|\Psi_i(s, t, \varphi)| \le c_0\|\varphi\|(t - s)^{-\frac{1}{2}}, \quad 0 \le s < t \le T. \tag{3.16}$$

Proceeding by the same considerations[2] as ones leading to the estimation (3.16) we can also investigate the kernels $K_{ij}(s,\tau)$ in (3.6). We find the following result: the kernels $K_{ij}(s,\tau),\ i,j=1,2$, can be represented as

$$K_{ij}(s,\tau) = \widetilde{K}_{ij}(s,\tau) + \overline{K}_{ij}(s,\tau), \quad 0 \le s < \tau < t \le T, \tag{3.17}$$

where

$$\widetilde{K}_{ij}(s,\tau) = -r_i(s)\sqrt{\frac{\pi b_j(\tau,0)}{2}} \int\limits_{D_{j,\delta}} \frac{\partial Z_{j0}}{\partial y}(s,y,\tau,0)\mu(s,dy),$$

and $K_{ij}^{(2)}(s,\tau)$ satisfy the inequality

$$\left|\overline{K}_{ij}(s,\tau)\right| \le h(\delta)(\tau-s)^{-1+\frac{\alpha}{2}}. \tag{3.18}$$

Here $\delta,\ h(\delta)$ are any positive number and some constant depending on δ, respectively; $D_{j,\delta} = \{y \in D_j :\ |y| < \delta\}$. It is seen that K_{ij} have non-integrable singularity, which is caused by $\widetilde{K}_{ij}$, and therefore we do not know yet whether a solution of (3.6) exists, i.e., whether the series (3.7) converges. For this reason, using (3.16) and (3.17), we try to estimate each term $V_i^{(k)}$ of series (3.7) and then to prove the convergence of (3.7).

Consider first the functions $V_i^{(1)}$. We can write

$$\begin{aligned} V_i^{(1)}(s,t,\varphi) &= \sum_{j=1}^{2}\int\limits_s^t K_{ij}(s,\tau)V_i^{(0)}(\tau,t,\varphi)d\tau = \sum_{j=1}^{2}\int\limits_s^t \widetilde{K}_{ij}(s,\tau)\Psi_i(\tau,t,\varphi)d\tau \\ &+ \sum_{j=1}^{2}\int\limits_s^t \overline{K}_{ij}(s,\tau)\Psi_i(\tau,t,\varphi)d\tau = V_{i1}^{(1)} + V_{i2}^{(1)}. \end{aligned} \tag{3.19}$$

Using (3.16) and (3.18), we get

$$\left|V_{i2}^{(1)}(s,t,\varphi)\right| \le 2c_0 h(\delta)\|\varphi\| \frac{\Gamma\left(\frac{\alpha}{2}\right)\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1+\alpha}{2}\right)}(t-s)^{-\frac{1-\alpha}{2}}, \tag{3.20}$$

where c_0 and $h(\delta)$ are the constants in (3.16) and (3.18), respectively.

For the functions $V_{i1}^{(1)}$ we have

$$\begin{aligned} &\left|V_{i1}^{(1)}(s,t,\varphi)\right| \le \\ &\le c_0\|\varphi\| r_i(s)\sqrt{\frac{\pi}{2}}\sum_{j=1}^{2}\int\limits_s^t (t-\tau)^{-\frac{1}{2}}\sqrt{b_j(\tau,0)}d\tau \int\limits_{D_{j,\delta}} \left|\frac{\partial Z_{j0}}{\partial y}(s,y,\tau,0)\right|\mu(s,dy) \le \end{aligned}$$

[2]For further details cf. [9]

$$\leq c_0\|\varphi\|\frac{r_i(s)}{2b}\sum_{j=1}^{2}\int_s^t(t-\tau)^{-\frac{1}{2}}(\tau-s)^{-\frac{3}{2}}d\tau\int_{D_{j,\delta}}|y|e^{-\frac{y^2}{2B(\tau-s)}}\mu(s,dy)=$$

$$=c_0\|\varphi\|\frac{r_i(s)}{2b}\sum_{j=1}^{2}\int_{D_{j,\delta}}|y|e^{-\frac{y^2}{2B(t-s)}}\mu(s,dy)\int_s^t(t-\tau)^{-\frac{1}{2}}(\tau-s)^{-\frac{3}{2}}e^{-\frac{y^2}{2B(t-s)}\cdot\frac{t-\tau}{\tau-s}}d\tau. \tag{3.21}$$

The change of variables $z=\frac{t-\tau}{\tau-s}$ in the inner integral in the last relation in (3.21) leads to

$$\left|V_{i1}^{(1)}(s,t,\varphi)\right|\leq$$

$$\leq c_0\|\varphi\|\frac{r_i(s)}{2b}(t-s)^{-1}\sum_{j=1}^{2}\int_{D_{j,\delta}}|y|e^{-\frac{y^2}{2B(t-s)}}\mu(s,dy)\int_0^\infty z^{-\frac{1}{2}}e^{-\frac{y^2}{2B(t-s)}z}dz\leq$$

$$\leq c_0\|\varphi\|\frac{B}{b}(t-s)^{-\frac{1}{2}}\sum_{j=1}^{2}\int_{D_{j,\delta}}e^{-\frac{y^2}{2B(t-s)}}\mu(s,dy)\leq$$

$$\leq c_0\|\varphi\|\frac{B}{b}(t-s)^{-\frac{1}{2}}\max_{s\in[0,T]}\mu\left(s,D_{1,\delta}\cup D_{2,\delta}\right). \tag{3.22}$$

Combining (3.20) and (3.22), we arrive at the inequality

$$\left|V_{i}^{(1)}(s,t,\varphi)\right|\leq$$

$$\leq c_0\|\varphi\|(t-s)^{-\frac{1}{2}}\left(\frac{2h(\delta)T^{\frac{\alpha}{2}}\Gamma\left(\frac{\alpha}{2}\right)\cdot\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1+\alpha}{2}\right)}+\frac{B}{b}\max_{s\in[0,T]}\mu\left(s,D_{1,\delta}\cup D_{2,\delta}\right)\right).$$

Next, by mathematical induction method, we prove that the terms $V_i^{(k)}$ of series (3.7) satisfy the inequalities

$$\left|V_{i}^{(k)}(s,t,\varphi)\right|\leq c\|\varphi\|(t-s)^{-\frac{1}{2}}\sum_{n=0}^{k}C_k^n\cdot a^{(k-n)}m(\delta)^n,\quad k=0,1,2, \tag{3.23}$$

where

$$a^{(n)}=\frac{\left(2h(\delta)T^{\frac{\alpha}{2}}\Gamma\left(\frac{\alpha}{2}\right)\right)^n\cdot\Gamma\left(\frac{1}{2}\right)}{\Gamma\left(\frac{1+n\alpha}{2}\right)},\quad n=0,1,2,\ldots,k,$$

$$m(\delta)=\frac{B}{b}\max_{s\in[0,T]}\mu\left(s,D_{1,\delta}\cup D_{2,\delta}\right).$$

Let us fix $\delta=\delta_0$ such that, $m(\delta_0)<1$. Then in view of (3.23), we have

$$\sum_{k=0}^{\infty}\left|V_{i}^{(k)}(s,t,\varphi)\right|\leq c_0\|\varphi\|(t-s)^{-\frac{1}{2}}\sum_{k=0}^{\infty}\sum_{n=0}^{k}C_k^n a^{(k-n)}m(\delta_0)^n=$$

$$= c_0\|\varphi\|(t-s)^{-\frac{1}{2}}\sum_{k=0}^{\infty} a^{(k)}\sum_{n=0}^{\infty} C_{k+n}^{n} m(\delta_0)^n =$$
$$= c_0\|\varphi\|(t-s)^{-\frac{1}{2}}\sum_{k=0}^{\infty}\frac{a^{(k)}}{(1-m(\delta_0))^{k+1}} =$$
$$= c_0\|\varphi\|(t-s)^{-\frac{1}{2}}\sum_{k=0}^{\infty}\frac{\left(\frac{h(\delta_0)}{1-m(\delta_0)}T^{\frac{\alpha}{2}}\Gamma\left(\frac{\alpha}{2}\right)\right)^k}{\Gamma\left(\frac{1+k\alpha}{2}\right)}\cdot\frac{\Gamma(\frac{1}{2})}{1-m(\delta_0)}\ . \tag{3.24}$$

The estimation (3.24) ensures the absolute and uniform convergence of series (3.7). This means that the functions $V_i(s,t,\varphi)$, $i=1,2$, exist. Furthermore, they are continuous in $0 \le s < t \le T$ and satisfy the inequality

$$|V_i(s,t,\varphi)| \le c\|\varphi\|(t-s)^{-\frac{1}{2}}, \quad 0 \le s < t \le T. \tag{3.25}$$

Using estimations (2.1), (2.4) and (3.25) we derive the existence of a solution $u(s,x,t)$, $0 \le s < t \le T$ of conjugation problem (1.4)-(1.7) which is of the form (3.2), satisfies inequality (3.1) and belongs to $C^{1,2}([0,t)\times D_1\cup D_2)\cap C([0,t)\times\mathbb{R})$.

Thus, in order to complete the proof of the theorem it remains to prove the uniqueness of the solution of the conjugation problem (1.4)-(1.7). For this purpose, it suffices to note that the constructed function $u(s,x,t)$ in each of two domains $0 \le s < t \le T$, $x \in \overline{D}_1$ and $0 \le s < t \le T$, $x \in \overline{D}_2$ can be treated as a unique solution to the following first boundary-value parabolic problem:

$$\frac{\partial\omega(s,x,t)}{\partial s} + A_s^{(i)}\omega(s,x,t) = 0, \quad 0 \le s < t \le T,\ x \in D_i,\ i=1,2,$$
$$\lim_{s\uparrow t}\omega(s,x,t) = \varphi(x), \quad x \in D_i,\ i=1,2,$$
$$\omega(s,0,t) = \frac{1}{\gamma(s)+1}\int\limits_{D_1\cup D_2} u(s,y,t)\mu(s,dy), \quad 0 \le s < t \le T.$$

The proof of Theorem 1 is now complete. □

Remark 3.2. Let, in addition to the conditions of Theorem 1, the fitting condition

$$L_t\varphi(0) = 0,$$

holds, then the solution u of the problem (1.4)-(1.7) constructed in Theorem 1 belongs to

$$C^{1,2}([0,t)\times D_1\cup D_2)\cap C([0,t]\times\mathbb{R}).$$

4. Process with absorptions and jumps

Suppose that the conditions of Theorem 1 hold and consider the two-parameter family of linear operators T_{st}, $0 \le s < t \le T$, acting on the function $\varphi \in C_b(\mathbb{R})$ by the formula

$$T_{st}\varphi(x) = \int_{\mathbb{R}} G_i(s,x,t,y)\varphi(y)dy + \int_s^t G_i(s,x,\tau,0)V_i(\tau,t,\varphi)d\tau, \tag{4.1}$$

where the pair of functions $(V_1,\ V_2)$ is the solution of (3.6).

We introduce the subspace $C_L(\mathbb{R})$ of $C_b(\mathbb{R})$ which consists of all $\varphi \in C_b(\mathbb{R})$ with $L_t\varphi(0) = 0$. It is easily seen that the space $C_L(\mathbb{R})$ is closed in $C_b(\mathbb{R})$, and so it is a Banach space. Furthermore, it is invariant under the operators T_{st}, i.e.,

$$\varphi \in C_L(\mathbb{R}) \implies T_{st}\varphi \in C_L(\mathbb{R}).$$

Let us study properties of the operator family T_{st} in the space $C_L(\mathbb{R})$. First we note that

$$\lim_{n\to\infty} T_{st}\varphi_n(x) = T_{st}\varphi(x), \quad 0 \le s < t \le T, \quad x \in \mathbb{R},$$

for every sequence of functions $\varphi_n \in C_L(\mathbb{R})$ such that

$$\sup_n \|\varphi_n\| < \infty \quad \text{and} \quad \lim_{n\to\infty} \varphi_n(x) = \varphi(x), \quad x \in \mathbb{R}.$$

This property easily follows from Lebesgue bounded convergence theorem and it allows us to make all the following considerations, without loss of generality, under the condition that the function φ has compact support.

Now we prove that the cone of nonnegative functions remains invariant under the operators $T_{st},\ 0 \le s < t \le T$.

Lemma 4.1. *If $\varphi \in C_L(\mathbb{R})$ and $\varphi(x) \ge 0$ for all $x \in \mathbb{R}$, then $T_{st}\varphi(x) \ge 0$ for all $0 \le s < t \le T,\ x \in \mathbb{R}$.*

Proof. Let φ be any nonnegative function in $C_L(\mathbb{R})$ with compact support. If $\varphi \equiv 0$, then the assertion of the lemma is obvious. Consider now the case where the function φ not everywhere equals zero. Denote by m a minimum of the function $T_{st}\varphi(x)$ in the domain $(s,x) \in [0,t] \times \mathbb{R}$ and assume that $m < 0$. By the minimum principle ([15, Ch. II]), the value m can be attained only when $(s,x) \in [0,t] \times \{0\}$. Fix $s_0 \in [0,t]$ such that $T_{s_0t}\varphi(0) = m$. Then the following inequalities hold:

$$\gamma(s_0)T_{s_0t}\varphi(0) \le 0, \qquad \int_{D_1\cup D_2} [T_{s_0t}\varphi(0) - T_{s_0t}\varphi(y)]\mu(s,dy) < 0.$$

Consequently,

$$L_{s_0}T_{s_0t}\varphi(0) < 0.$$

Since, however, the condition (1.7) holds, we get a contradiction. This completes the proof of the lemma. □

By similar considerations to those in proof of Lemma 2, it can be easily verified that the operators T_{st} are contractive, i.e.,

$$\|T_{st}\| \leq 1, \quad 0 \leq s < t \leq T.$$

Finally, we show that the operator family T_{st} has a semigroup property

$$T_{st} = T_{s\tau}T_{\tau t}, \quad 0 \leq s < \tau < t \leq T.$$

This property is a consequence of the assertion of uniqueness of the solution of the problem (1.4)–(1.7) which we have already established above. Indeed, to find $u(s,x,t)$ when $\lim_{s\uparrow t} u(s,x,t) = \varphi(x)$, the problem (1.4)–(1.7) can be solved first in the time interval $[\tau, t]$, and then with the "initial" function $u(\tau, x, t) = T_{\tau t}\varphi(x)$, we derived, it can be solved in the time interval $[s, \tau]$. In other words, $T_{st}\varphi(x) = T_{s\tau}(T_{\tau t}\varphi)(x)$, $\varphi \in C_b(\mathbb{R})$, i.e., $T_{st} = T_{s\tau}T_{\tau t}$.

The properties of the operator family T_{st}, proved above, implies (see [5, Ch. II, §1]) the next theorem.

Theorem 4.2. *Let the conditions of Theorem 1 hold. Then the two-parameter semigroup of operators T_{st}, $0 \leq s < t \leq T$, defined by formula (4.1) describes the inhomogeneous Feller process in $\mathbb{R}$, such that in D_1 and D_2 it coincides with the diffusion processes generated by $A_s^{(1)}$ and $A_s^{(2)}$, respectively, and its behavior on $S = \{0\}$ is determined by the boundary condition (1.3).*

References

[1] W. Feller, *Diffusion processes in one dimension*, Trans. Amer. Math. Soc., **77** (1954), 1–31.

[2] A. D. Wentzell, *Semigroups of operators that correspond to a generalized differential operator of second order* (in Russian), Dokl. AN SSSR, **111** (1956), no. 2, 269–272.

[3] M. I. Portenko, *Diffusion Processes in Media with Membranes* (in Ukrainian), Institute of Mathematics of the NAS of Ukraine, Kyiv, 1995.

[4] B. I. Kopytko, *Sewing two nonhomogeneous diffusion processes on a straight line*, Ukrainian Math. J., **35** (1983), no. 2, 135–141.

[5] E. B. Dynkin, *Markov Processes* (in Russian), Fizmatgiz, Moscow, 1963.

[6] P. P. Kononchuk, *Pasting of two diffusion processes on a line with nonlocal boundary conditions*, Theory of Stochastic Processes, **14 (30)** (2008), no. 2, 52–59.

[7] P. P. Kononchuk, B. I. Kopytko, *Operator semigroups that describe Feller process on a line pasted from two diffusion processes* (in Ukrainian), Teor. Imovir. ta Matem. Statyst., **84** (2011), 84–93.

[8] P. P. Kononchuk, *One-dimensional model of the diffusion process with a membrane that is described by the Feller-Wentzell conjugation condition*, Theory of Stochastic Processes, **17 (33)** (2011), no. 1, 61–69.

[9] R. V. Shevchuk, *Inhomogeneous diffusion processes on a half-line with jumps on its boundary*, Theory of Stochastic Processes, **17 (33)** (2011), no. 1, 119–129.

[10] B. I. Kopytko, R. V. Shevchuk, *On pasting together two inhomogeneous diffusion processes on a line with the general Feller-Wentzell conjugation condition*, Theory of Stochastic Processes, **17 (33)** (2011), no. 2, 55–70.

[11] G. L. Kulinic, *On the limit behavior of the distribution of the solution of a stochastic diffusion equation*, Theory Probability Appl., **12** (1967), no. 3, 497–499.

[12] S. V. Anulova, *On stochastic differential equations with boundary conditions in a half-plane* (in Russian), Izv. AN SSSR Ser. Mat., **45** (1981), no. 3, 491–508.

[13] A. Yu. Pilipenko, *On the Skorokhod mapping for equations with reflection and possible jump-like exit from a boundary* (in Russian), Ukrain. Mat. Zh., **63** (2011), no. 9, 1241–1256.

[14] O. A. Ladyzhenskaya, V. A. Solonnikov, N. N. Ural'tseva, *Linear and Quasilinear Equations of Parabolic Type* (in Russian), Nauka, Moscow, 1967.

[15] A. Friedman, *Partial Differential Equations of Parabolic Type* (in Russian), Mir, Moscow, 1968.

Annales Mathematicae et Informaticae
39 *(2012) pp. 241–307*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Moment-type estimates with asymptotically optimal structure for the accuracy of the normal approximation*

Irina Shevtsova

Faculty of Computational Mathematics and Cybernetics
Lomonosov Moscow State University, Moscow, Russia
Institute for Informatics Problems of the Russian Academy of Sciences
e-mail: `ishevtsova@cs.msu.su`

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

For the uniform distance Δ_n between the distribution function of the standard normal law and the distribution function of the standardized sum of independent random variables $X_1,\dots,X_n$ with $\mathsf{E}X_j=0$, $\mathsf{E}|X_j|=\beta_{1,j}$, $\mathsf{E}X_j^2=\sigma_j^2$, $j=1,\dots,n$, for all $n\geqslant 1$ the bounds

$$\Delta_n\leqslant\frac{2\ell_n}{3\sqrt{2\pi}}+\frac{1}{2\sqrt{2\pi}B_n^3}\sum_{j=1}^n\beta_{1,j}\,\sigma_j^2+R(\ell_n),$$

$$\Delta_n\leqslant\inf_{c\geqslant 2/(3\sqrt{2\pi})}\left\{c\ell_n+\frac{K(c)}{B_n^3}\sum_{j=1}^n\sigma_j^3+R_c(\ell_n)\right\},$$

are proved, where $B_n^2=\sum_{j=1}^n\sigma_j^2$, $\ell_n=B_n^{-3}\sum_{j=1}^n\mathsf{E}|X_j|^3$, $R(\ell_n)\leqslant 6\ell_n^{5/3}$, $R_c(\ell_n)\leqslant\min\{3\ell_n^{7/6},A(c)\ell_n^{4/3}\}$ in the general case and $R(\ell_n)\leqslant 3\ell_n^2$, $R_c(\ell_n)\leqslant\min\{2\ell_n^{3/2},A(c)\ell_n^2\}$, if $X_1,\dots,X_n$ are identically distributed, $A(c)>0$ being a decreasing function of c such that $A(c)\to\infty$ as $c\to 2/(3\sqrt{2\pi})$. Moreover, the function $K(c)$ is optimal for each $c\geqslant 2/(3\sqrt{2\pi})$. In particular, $K\left((\sqrt{10}+3)/(6\sqrt{2\pi})\right)=0$, $K\left(2/(3\sqrt{2\pi})\right)=\sqrt{(2\sqrt{3}-3)/(6\pi)}=0.1569\dots$ It is shown that in the first inequality the coefficients $2/(3\sqrt{2\pi})$ and $\left(2\sqrt{2\pi}\right)^{-1}$

*Research supported by the Russian Foundation for Basic Research (projects 11-01-00515a, 11-07-00112a, 11-01-12026-ofi-m and 12-01-31125) and by the Ministry for Education and Science of Russia (grant MK–2256.2012.1 and State contract 16.740.11.0133).

are optimal and the lower bound $2/(3\sqrt{2\pi})$ for c in the second inequality is unimprovable. These results sharpen the well-known estimates due to H. Prawitz (1975), V. Bentkus (1991, 1994) and G. P. Chistyakov (1996, 2001). Also, an analog of the first inequality is proved for the case where the summands possess only the moments of order $2+\delta$ with some $0<\delta<1$. As a by-product, the von Mises inequality for lattice distributions is sharpened and generalized.

Keywords: central limit theorem, convergence rate estimate, normal approximation, Berry–Esseen inequality, asymptotically exact constant, characteristic function

MSC: 60F05, 60E10

1. Introduction

For $\delta\in[0,1]$ let $\mathcal{F}_{2+\delta}$ be the class of distribution functions (d.f.'s) $F(x)$ satisfying the conditions

$$\int\limits_{-\infty}^{+\infty} x\,dF(x)=0,\qquad \int\limits_{-\infty}^{+\infty}|x|^{2+\delta}\,dF(x)<\infty.$$

For $h>0$ let $\mathcal{F}^h_{2+\delta}$ denote the class of all lattice d.f.'s from $\mathcal{F}_{2+\delta}$ with span h. For $F\in\mathcal{F}_{2+\delta}$ set

$$\beta_r=\beta_r(F)=\int\limits_{-\infty}^{+\infty}|x|^r\,dF(x),\quad 0<r\leqslant 2+\delta,\quad \sigma^2=\beta_2.$$

For $\delta=0$ by $\mathcal{F}_2$ we mean the class of all d.f.'s with zero mean and finite second moment. It is easy to see that $\mathcal{F}_{2+\delta_1}\subset\mathcal{F}_{2+\delta_2}$ for any $0\leqslant\delta_1<\delta_2\leqslant 1$, and $\sigma^{2+\delta}\leqslant\beta_{2+\delta}$ for all $F\in\mathcal{F}_{2+\delta}$ and $\delta\in[0,1]$ by the Lyapounov inequality.

Let $X_1,\ldots,X_n$ be independent random variables (r.v.'s) defined on some probability space $(\Omega,\mathcal{A},\mathsf{P})$ with the corresponding d.f.'s $F_1,\ldots,F_n\in\mathcal{F}_{2+\delta}$. Denote

$$\sigma_j^2=\mathsf{E}X_j^2,\quad \beta_{r,j}=\mathsf{E}|X_j|^r,\quad 0<r\leqslant 2+\delta,\quad j=1,2,\ldots,n,$$

$$B_n^2=\sum_{j=1}^n\sigma_j^2,\quad \ell_n=\frac{1}{B_n^{2+\delta}}\sum_{j=1}^n\beta_{2+\delta,j},$$

$$\overline{F}_n(x)=\mathsf{P}(X_1+\ldots+X_n<xB_n)=(F_1*\ldots*F_n)(xB_n),$$

$$\Delta_n=\Delta_n(F_1,\ldots,F_n)=\sup_x|\overline{F}_n(x)-\Phi(x)|,\quad n=1,2,\ldots,$$

$\Phi(x)$ being the standard normal d.f. Assume, that $B_n>0$. It is easy to verify that under the above assumptions for any $n\geqslant 1$ we have

$$\ell_n\geqslant\frac{1}{B_n^{2+\delta}}\sum_{j=1}^n\sigma_j^{2+\delta}\geqslant n^{-\delta/2}.$$

If the r.v.'s $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.), then their common d.f. will be denoted by F $(= F_1 = \ldots = F_n)$. In this case we use the notation

$$\Delta_n(F) = \Delta_n(F_1, \ldots, F_n), \quad \sigma^2 = \mathsf{E}X_1^2 > 0, \quad \beta_{2+\delta} = \mathsf{E}|X_1|^{2+\delta}, \quad \beta_\delta = \mathsf{E}|X_1|^\delta.$$

Then

$$B_n = \sigma\sqrt{n}, \quad \ell_n = \frac{\beta_{2+\delta}}{\sigma^{2+\delta}n^{\delta/2}}.$$

In what follows, for a r.v. X the notation $X \in \mathcal{F}_{2+\delta}$ means that the d.f. $F(x) = \mathsf{P}(X < x)$, $x \in \mathbf{R}$, belongs to the class $\mathcal{F}_{2+\delta}$.

As is known, the rate of convergence in the central limit theorem of probability theory obeys the Berry–Esseen inequality

$$\Delta_n \leqslant C_{\text{BE}}(\delta) \cdot \ell_n, \quad n \geqslant 1, \ F_1, \ldots, F_n \in \mathcal{F}_{2+\delta}, \tag{1.1}$$

where $C_{\text{BE}}(\delta)$ depends only on δ [4, 8, 9]. Omitting the history of improvement of the constant $C_{\text{BE}}(1)$ the details of which can be found, for example, in the papers [19, 20], note that

$$0.4097\ldots = \frac{\sqrt{10}+3}{6\sqrt{2\pi}} \leqslant C_{\text{BE}}(1) \leqslant \begin{cases} 0.5600, & \text{in the general case,} \\ 0.4784, & \text{if } F_1 = \ldots = F_n, \end{cases}$$

see [10, 28, 20].[1] In 1966–1967 V. M. Zolotarev [37, 38, 39] suggested that $C_{\text{BE}}(1) = (\sqrt{10}+3)/(6\sqrt{2\pi})$. This hypothesis has been neither proved nor rejected yet.

For $0 < \delta < 1$ the best known upper estimates of the constants $C_{\text{BE}}(\delta)$ were obtained by W. Tysiak [30] for the general case (the second line in table 1) and by M. Grigorieva and I. Shevtsova [13] for the case of identically distributed summands (the third line in table 1). The first lower estimates were recently obtained by the author [29] (the fourth line in table 1).

In the case of identically distributed summands ($F_1 = \ldots = F_n = F$) and $\delta = 1$, inequality (1.1) takes the form

$$\Delta_n \leqslant C_{\text{BE}}(1) \cdot \frac{\beta_3}{\sigma^3\sqrt{n}}, \quad n \geqslant 1, \ F \in \mathcal{F}_3, \tag{1.2}$$

and along with the information concerning the two first moments also uses the value of the third absolute moment β_3.

[1] Recently, the presented upper bounds for $C_{\text{BE}}(1)$ were improved to $C_{\text{BE}}(1) \leqslant 0.5591$ in the general case by Ilya Tyurin (see "An improvement of the remainder in the Lyapounov theorem", Theory Probab. Appl., 2011, vol. 56, No. 4, p. 808-811 (in Russian)) and to $C_{\text{BE}}(1) \leqslant 0.4748$ in the i.i.d.-case by the author (see "On the absolute constants in the Berry–Esseen type inequalities for identically distributed summands", arXiv:1111.6554, 28 November 2011), the latest one — as a corollary to the estimate with an improved structure $\Delta_n \leqslant 0.33554(\beta_3/\sigma^3 + 0.415)/\sqrt{n}$, since $0.33554(\beta_3/\sigma^3 + 0.415) \leqslant 0.33554 \cdot 0.415\beta_3/\sigma^3 < 0.4748\beta_3/\sigma^3$ by virtue of the Lyapounov inequality. Independently, an estimate $C_{\text{BE}}(1) \leqslant 0.4774$ for the i.i.d.-case was obtained in the paper of I. Tyurin.

δ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$C_{\text{BE}}(\delta) \leqslant$	1.102	1.076	1.008	0.950	0.902	0.863	0.833	0.812	0.802
$C_{\text{BE}}(\delta) \leqslant$	0.6028	0.6094	0.6195	0.6342	0.6413	0.6276	0.6026	0.5723	0.5383
$C_{\text{BE}}(\delta) \geqslant$	0.4097	0.3603	0.3257	0.3000	0.2803	0.2651	0.2534	0.2446	0.2383

Table 1: Two-sided estimates of the constants $C_{\text{BE}}(\delta)$ from inequality (1.1) for some $\delta \in (0,1)$. The second line: the upper estimates in the general case [30]; the third line: improved estimates for the case of identically distributed summands [13]; the fourth line: the lower estimates [29].

On the other hand, as $n \to \infty$, if the summands are i.i.d. with arbitrary *fixed* (independent of n) d.f. $F \in \mathcal{F}_3$, then, as it was established in 1945 by Esseen [9], uniformly in x

$$\overline{F}_n(x) = \Phi(x) + \frac{\mathsf{E}X_1^3}{6\sigma^3} \cdot \frac{(1-x^2)e^{-x^2/2}}{\sqrt{2\pi n}} + \frac{h}{\sigma} \cdot \frac{H_n(x)e^{-x^2/2}}{\sqrt{2\pi n}} + o\left(\frac{1}{\sqrt{n}}\right), \tag{1.3}$$

where $h = hH_n(x) \equiv 0$, if F is non-lattice, and

$$H_n(x) = \frac{1}{2} - \left\{\left(x\sqrt{n} - \frac{an}{\sigma}\right)\frac{\sigma}{h}\right\}, \quad |H_n(x)| \leqslant \frac{1}{2},$$

if F is concentrated on the lattice $\{a + kh,\ k = 0, \pm 1, \pm 2, \ldots\}$ with span h, $\{x\}$ being the fractional part of $x \in \mathbf{R}$, whence Esseen deduced [10] that

$$\limsup_{n\to\infty} \Delta_n(F)\sqrt{n} = \frac{|\mathsf{E}X_1^3| + 3h\sigma^2}{6\sqrt{2\pi}\sigma^3}, \ \ F \in \mathcal{F}_3^h. \tag{1.4}$$

So, unlike (1.2), in the asymptotic relations (1.3) and (1.4) the third absolute moment $\mathsf{E}|X_1|^3$ does not take part at all whereas only the first three *original* moments are used as well as the parameter h, carrying the information on *the structure* of the basic distribution. The numerical characteristics mentioned above satisfy the relation [10, 40]

$$\sup_{h>0} \sup_{X \in \mathcal{F}_3^h} \frac{|\mathsf{E}X^3| + 3h\mathsf{E}X^2}{\mathsf{E}|X|^3} = \sqrt{10} + 3, \tag{1.5}$$

with supremum attained at the two-point distribution $\mathsf{P}(X = -h(4-\sqrt{10})/2) = (\sqrt{10}-2)/2$, $\mathsf{P}(X = h(\sqrt{10}-2)/2) = (4-\sqrt{10})/2$, called *the Esseen distribution.* From (1.4) and (1.5) it follows that for any $\mathcal{F} \in \mathcal{F}_3$

$$\limsup_{n\to\infty} \Delta_n(F)\sqrt{n} \leqslant \frac{\sqrt{10}+3}{6\sqrt{2\pi}} \cdot \frac{\beta_3}{\sigma^3}. \tag{1.6}$$

With the supremum attained at the Esseen distribution. This remark makes it possible to establish the lower estimate $C_{\text{BE}}(1) \geqslant (\sqrt{10}+3)/(6\sqrt{2\pi})$ as it was done by Esseen [10]. It is worth noticing for the sake of completeness that the normalized value of the third absolute moment of the Esseen distribution delivering the extremum in (1.5) and equality in (1.6) have the form

$$\beta_3/\sigma^3 = \sqrt{20(\sqrt{10}-3)/3} = 1.0401\ldots$$

So, if in (1.5) the supremum is sought not over all $X \in \mathcal{F}_3^h$, but under additional requirement that the ratio $\mathsf{E}|X|^3/(\mathsf{E}X^2)^{3/2}$ should be large enough, then the extremal value becomes smaller and hence, the lower estimate of the constant $C_{\text{BE}}(1)$ in (1.2) becomes more optimistic. This remark generates the hope (and explains) that the larger the value of the Lyapounov ratio β_3/σ^3, the smaller the upper estimate of the constant $C_{\text{BE}}(1)$ in (1.1) is.

Apparently, S. Zahl was the first to notice this [35, 36]. In 1963 he presented the structural improvement of inequality (1.1)

$$\Delta_n \leqslant \frac{0.651}{B_n^3}\sum_{j=1}^n \beta'_{3,j},$$

where

$$\beta'_{3,j} = \begin{cases} \beta_{3,j}, & \beta_{3,j} \geqslant 3\sigma_j^3/\sqrt{2}, \\ \sigma_j^3/\left(0.7804 - 0.1457\beta_{3,j}/\sigma_j^3\right), & \beta_{3,j} < 3\sigma_j^3/\sqrt{2}, \end{cases}$$

which more efficiently uses the information concerning the first three moments of random summands.

The next step in this direction was made in 1975 by H. Prawitz, from whose paper [25] one can deduce the estimate

$$\Delta_n \leqslant \ell_n \cdot A_1(\ell_n) + \frac{1}{2\sqrt{2\pi}B_n^3}\sum_{j=1}^n \sigma_j^3 + \frac{1}{4\pi B_n^4}\sum_{j=1}^n \sigma_j^4, \tag{1.7}$$

where $A_1(\ell)$ is a positive function of $\ell > 0$ with a complicated structure such that $A_1(\ell)$ does not increase for ℓ small enough and

$$\lim_{\ell\to 0} A_1(\ell) = \frac{1.0253}{6\sqrt{2\pi}} + \frac{1}{2\sqrt{2\pi}} = \frac{2}{3\sqrt{2\pi}} + \frac{0.0253}{6\sqrt{2\pi}} = 0.2676\ldots$$

Prawitz also described an algorithm for the computation of $A_1(\ell)$ for concrete values of ℓ. Since

$$\frac{1}{B_n^3}\sum_{j=1}^n \sigma_j^3 \leqslant \frac{1}{B_n^3}\sum_{j=1}^n \beta_{3,\,j} = \ell_n, \quad \frac{1}{B_n^4}\sum_{j=1}^n \sigma_j^4 \leqslant \ell_n^{4/3} = o(\ell_n),\ \ell_n \to 0,$$

from (1.7) it follows that

$$\Delta_n \leqslant \ell_n \cdot A_2(\ell_n), \tag{1.8}$$

where $A_2(\ell)$ is a positive function of $\ell > 0$ such that $A_2(\ell)$ does not increase for ℓ small enough and

$$\lim_{\ell\to 0} A_2(\ell) = \frac{1.0253}{6\sqrt{2\pi}} + \frac{1}{\sqrt{2\pi}} = \frac{7}{6\sqrt{2\pi}} + \frac{0.0253}{6\sqrt{2\pi}} = 0.4671\ldots\ .$$

Inequality (1.8) with concrete values of A_2 plays an important role in the problem of determination of upper estimates of the absolute constant $C_{\text{BE}}(1)$ in the Berry–Esseen inequality (1.1), since the algorithms which are traditionally used for these purposes cannot obtain the values of this constant which are less than A_2.

In the same paper [25], for identically distributed summands and $n \geqslant 2$, Prawitz announced the inequality

$$\Delta_n \leqslant \frac{2}{3\sqrt{2\pi}} \cdot \frac{\beta_3}{\sigma^3\sqrt{n-1}} + \frac{1}{2\sqrt{2\pi(n-1)}} + A_3 \cdot \ell_{n-1}^2, \tag{1.9}$$

where A_3 is an absolute positive constant and stated that the coefficient

$$\frac{2}{3\sqrt{2\pi}} = 0.2659\ldots$$

at the Lyapounov fraction in (1.9) cannot be made smaller. Unfortunately, the proof of this statement as well as that of inequality (1.9) were not published by Prawitz.

A strict proof of Prawitz' inequality (1.9), however, with a little worse remainder, follows from the papers of V. Bentkus [2, 3], in which for the case of arbitrary $F_1, \ldots, F_n \in \mathcal{F}_3$ and $n \geqslant 1$ the estimate

$$\Delta_n \leqslant \frac{2\ell_n}{3\sqrt{2\pi}} + \frac{1}{2\sqrt{2\pi}B_n^3}\sum_{j=1}^n \sigma_j^3 + A_4 \cdot \ell_n^{4/3} \leqslant \frac{7\ell_n}{6\sqrt{2\pi}} + A_4 \cdot \ell_n^{4/3} \tag{1.10}$$

was obtained, where A_4 is an absolute constant. The worse order of the remainder in (1.10) as compared with (1.9) is due to that the estimate (1.10) holds for arbitrary (not necessarily identical) $F_1, \ldots, F_n \in \mathcal{F}_3$.

So, even if the value of the constant A_4 in (1.10) were known, it would not be possible to obtain an estimate of the absolute constant $C_{\text{BE}}(1)$ in the Berry–Esseen inequality (1.1) lower than $7/(6\sqrt{2\pi}) = 0.4654\ldots$. For further progress in this problem, one has to improve the main term of asymptotic estimate (1.10).

In 1953 A. N. Kolmogorov [17] (also see the monographs of I. A. Ibragimov and Yu. V. Linnik [16] and V. M. Zolotarev [40]) formulated the problem of calculation of the so-called asymptotically exact constant

$$C_{\text{AE}} = \limsup_{\ell\to 0} \sup_{n\geqslant 1,\, F_1,\ldots,F_n :\, \ell_n=\ell} \frac{\Delta_n(F_1,\ldots,F_n)}{\ell},$$

for which from the papers of Esseen [10] and Bentkus [2, 3] it follows that

$$0.4097\ldots = \frac{\sqrt{10}+3}{6\sqrt{2\pi}} \leqslant C_{\text{AE}} \leqslant \frac{7}{6\sqrt{2\pi}} = 0.4654\ldots\ .$$

V. M. Zolotarev [38, 39, 40] held the opinion that C_{AE} coincides with its lower bound and together with A. N. Kolmogorov considered the problem of calculation of C_{AE} to be intermediate or auxiliary for the problem of calculation of the exact value of the absolute constant $C_{\text{BE}}(1)$ in (1.1). The gap of approximately 0.06 between the upper and lower bounds of C_{AE} presented above is due to the fact that the information on the *original* moments of summands is not taken into account in [25, 2, 3]. Since the summands are centered, the only informative original moment is the third one. S. V. Nagaev and V. I. Chebotarev [21] also noticed this and for the i.i.d. two-point summands proved the estimate $C_{\text{BE}}(1) \leqslant 0.4215$.

In 2001–2002 G. P. Chistyakov [7] obtained a new asymptotic expansion generalizing that due to Esseen (1.3) to the case of non-identically distributed summands. This new expansion allowed Chistyakov, as an intermediate step, to use the information concerning the original moments and other characteristics of the initial distributions and, as a result, to deduce the estimate

$$\Delta_n \leqslant \frac{\sqrt{10}+3}{6\sqrt{2\pi}} \cdot \ell_n + A_5 \cdot \ell_n^{40/39} |\ln \ell_n|^{7/6}, \tag{1.11}$$

where A_5 is an absolute constant. From (1.11) it follows that

$$C_{\text{AE}} = \frac{\sqrt{10}+3}{6\sqrt{2\pi}} = 0.4097\ldots,$$

thus Chistyakov proved the validity of Zolotarev's hypothesis concerning the exact value of the asymptotically exact constant C_{AE}.

Unfortunately, the particular value of the absolute constant A_5 in Chistyakov's inequality (1.11) was not given, so this fundamental result cannot be used for practical calculations, in particular, for the evaluation of the absolute constant $C_{\text{BE}}(1)$ in the Berry–Esseen inequality.

Nevertheless, the inequalities of Prawitz (1.9) and Bentkus (1.10) are interesting because in these inequalities the coefficient at the Lyapounov fraction is less than in Chistyakov's inequality (1.11):

$$0.2659\ldots = \frac{2}{3\sqrt{2\pi}} < \frac{\sqrt{10}+3}{6\sqrt{2\pi}} = 0.4097\ldots,$$

and hence, with large values of the ratio

$$\sum_{j=1}^{n} \beta_{3,j} \Big/ \sum_{j=1}^{n} \sigma_j^3$$

inequalities (1.9) and (1.10) are more precise than (1.11). This ratio may be arbitrarily large even in the case of identically distributed summands, for example, in the double array scheme where $\beta_3/\sigma^3 = \beta_3(n)/\sigma^3(n) \to \infty$, so that

$$\frac{1}{B_n^3} \sum_{j=1}^{n} \sigma_j^3 = \frac{1}{\sqrt{n}} = o(\ell_n) \quad \text{as } \ell_n = \frac{\beta_3(n)}{\sigma^3(n)\sqrt{n}} \to 0.$$

So, the unproved Prawitz' assertion that the coefficient $2/(3\sqrt{2\pi})$ at the Lyapounov fraction is unimprovable becomes exceptionally important. This assertion was proved only recently in [29] where the so-called *lower asymptotically exact constant*

$$\underline{C_{\text{AE}}} = \limsup_{\ell\to 0}\limsup_{n\to\infty}\sup_{F\colon\,\beta_3=\sigma^3\ell\sqrt{n}}\frac{\Delta_n(F)}{\ell}$$

was introduced (for the scheme of summation of identically distributed summands), which is an obvious lower bound for the coefficient under discussion, and it was demonstrated that $\underline{C_{\text{AE}}} = 2/(3\sqrt{2\pi})$.

The unimprovability of the first term in (1.9) naturally puts forward the question concerning the accuracy of the second term. No suggestions concerning the "exactness" of the coefficient at the second term in (1.9), (1.10) were stated by Prawitz or Bentkus. Actually, this question can be formulated in an even more general form: for any $c \geqslant \underline{C_{\text{AE}}}$ find the least possible value $K(c)$ providing the validity of the asymptotic estimate

$$\sup_{F\in\mathcal{F}_3\colon\,\beta_3=\rho\sigma^3}\Delta_n(F) \leqslant \frac{c\rho}{\sqrt{n}} + \frac{K(c)}{\sqrt{n}} + r_n(\rho)\cdot\frac{\rho}{\sqrt{n}},\quad n,\ \rho\geqslant 1,$$

in which the remainder $r_n(\rho)\geqslant 0$ satisfies the conditions

$$\limsup_{\ell\to 0}\limsup_{n\to\infty} r_n(\ell\sqrt{n}) = 0,\quad \sup_{\rho\geqslant 1}\limsup_{n\to\infty} r_n(\rho) = 0. \tag{1.12}$$

Apparently, for the first time this question was formulated in [29], where lower estimates of $K(c)$ were presented for $\underline{C_{\text{AE}}} \leqslant c \leqslant C_{\text{AE}}$. In particular, for $c = \underline{C_{\text{AE}}}$ in [29] it was shown that

$$K\left(\frac{2}{3\sqrt{2\pi}}\right) \geqslant \sqrt{\frac{2\sqrt{3}-3}{6\pi}} = 0.1569\ldots,$$

which is strictly less than the value of the coefficient $\left(2\sqrt{2\pi}\right)^{-1} = 0.1994\ldots$ at the second term in inequalities (1.9) and (1.10). Thus, the question of the "exactness" of the second term in (1.9) and (1.10) remained unanswered.

In the present paper we will prove that: for all $n\geqslant 1$ and $F_1,\ldots,F_n\in\mathcal{F}_3$

$$\Delta_n \leqslant \inf_{c\geqslant\underline{C_{\text{AE}}}}\left\{c\ell_n + \frac{K(c)}{B_n^3}\sum_{j=1}^n\sigma_j^3 + \min\left\{2.7176\ell_n^{7/6}, A(c)\ell_n^{4/3}\right\}\right\},$$

and for identically distributed summands

$$\Delta_n \leqslant \inf_{c\geqslant\underline{C_{\text{AE}}}}\left\{\frac{c\beta_3}{\sigma^3\sqrt{n}} + \frac{K(c)}{\sqrt{n}} + \min\left\{1.7002\ell_n^{3/2}, A(c)\ell_n^2\right\}\right\},$$

with the function $K(c)$ optimal for each $c\geqslant\underline{C_{\text{AE}}}$ (the optimality of this function is proved in remark 4.16), $A(c)>0$ being a decreasing function of c such that

$A(c)\to\infty$ as $c\to 2/(3\sqrt{2\pi})$. The function $K(c)$ decreases monotonically alternating its sign in a single point $c=(\sqrt{10}+3)/(6\sqrt{2\pi})$. So, the second term in the estimates presented above is negative for $c>(\sqrt{10}+3)/(6\sqrt{2\pi})$. The presence of a negative summand in the main term is rather unusual in estimates of the accuracy of the normal approximation, but makes it possible to obtain asymptotically exact estimates as simple corollaries of the results presented above even for symmetric Bernoulli distributions (see corollary 4.19) which distinguishes these results from previously known. In particular, for $c=C_{\text{AE}}$ we have

$$\Delta_n\leqslant\frac{\sqrt{10}+3}{6\sqrt{2\pi}}\cdot\ell_n+3.4314\cdot\ell_n^{4/3},\quad n\geqslant 1,\ F_1,\ldots,F_n\in\mathcal{F}_3,$$

$$\Delta_n\leqslant\frac{\sqrt{10}+3}{6\sqrt{2\pi}}\cdot\frac{\beta_3}{\sigma^3\sqrt{n}}+2.5786\cdot\ell_n^2,\quad n\geqslant 1,\ F_1=\ldots=F_n\in\mathcal{F}_3,$$

which improves Chistyakov's inequality (1.11) with respect to the remainder, whereas for $c=\underline{C_{\text{AE}}}$ we have

$$\Delta_n\leqslant\frac{2\ell_n}{3\sqrt{2\pi}}+\sqrt{\frac{2\sqrt{3}-3}{6\pi}\sum_{j=1}^n\frac{\sigma_j^3}{B_n^3}}+2.7176\cdot\ell_n^{7/6},\quad n\geqslant 1,\ F_1,\ldots,F_n\in\mathcal{F}_3,$$

$$\Delta_n\leqslant\frac{2}{3\sqrt{2\pi}}\cdot\frac{\beta_3}{\sigma^3\sqrt{n}}+\sqrt{\frac{2\sqrt{3}-3}{6\pi n}}+1.7002\cdot\ell_n^{3/2},\quad n\geqslant 1,\ F_1=\ldots=F_n\in\mathcal{F}_3,$$

which improves Prawitz' and Bentkus' inequalities (1.9), (1.10) with respect to the second term. Moreover, we will obtain the absolute improvements of Prawitz' and Bentkus' inequalities (1.9) and (1.10):

$$\Delta_n\leqslant\frac{2\ell_n}{3\sqrt{2\pi}}+\frac{1}{2\sqrt{2\pi}B_n^3}\sum_{j=1}^n\beta_{1,j}\,\sigma_j^2+5.4527\cdot\ell_n^{5/3},\quad n\geqslant 1,\ F_1,\ldots,F_n\in\mathcal{F}_3,$$

$$\Delta_n\leqslant\frac{2}{3\sqrt{2\pi}}\cdot\frac{\beta_3}{\sigma^3\sqrt{n}}+\frac{1}{2\sqrt{2\pi}}\cdot\frac{\beta_1}{\sigma\sqrt{n}}+2.4606\cdot\ell_n^2,\quad n\geqslant 1,\ F_1=\ldots=F_n\in\mathcal{F}_3,$$

in which the remainders have no worse order of decrease than in (1.9) and (1.10) but with specified constants and an improved function $\sum_{j=1}^n\beta_{1,j}\,\sigma_j^2\leqslant\sum_{j=1}^n\sigma_j^3$ of the two first moments in the second term with the same coefficient as in (1.9), (1.10). Below it will be shown that the value of the coefficient $\big(2\sqrt{2\pi}\big)^{-1}$ at this improved function of the two first moments yet cannot be made less (see remark 4.9). As well, similar estimates will be obtained for the case $0<\delta<1$, generalizing and sharpening the results of [11], where only the case of identically distributed summands was considered.

To prove the main results we use a combination of the method of characteristic functions (ch.f.'s) with the truncation method as well as some methods of convex analysis based on the works of W. Hoeffding [15] and V. M. Zolotarev [40].

It is worth noticing that in the preceding works dealing with the accuracy of the normal approximation, Prawitz' smoothing inequality was used, besides Prawitz himself, only by V. Bentkus [2, 3]. G. P. Chistyakov in [7] used Esseen's traditional smoothing inequality with the normal smoothing kernel, while in Prawitz' inequality, the smoothing function has a compact Fourier transform and does not have any probabilistic interpretation.

The paper is arranged as follows. In the second section we present new estimates for ch.f.'s implying, in particular, a generalization and improvement of the von Mises inequality for lattice distributions: for any $h>0$, $\delta\in(0,1]$ and $F\in\mathcal{F}^h_{2+\delta}$

$$\frac{h}{\sigma}\leqslant\frac{\beta_{2+\delta}}{\sigma^{2+\delta}}+\frac{\beta_\delta}{\sigma^\delta},$$

whereas in the original von Mises inequality $\delta=1$ and on the right-hand side there is $2\beta_3/\sigma^3$. In the third section a moment inequality is proved which improves (1.5) and plays the key role for the construction of the optimal function of moments in the resulting estimates. In the fourth section we formulate and prove new moment-type estimates of the accuracy of the normal approximation with optimal structure.

2. Estimates for characteristic functions

Denote

$$\varepsilon_n=B_n^{-(2+\delta)}\sum_{j=1}^n(\beta_{2+\delta,j}+\beta_{\delta,j}\,\sigma_j^2)=\ell_n+B_n^{-(2+\delta)}\sum_{j=1}^n\beta_{\delta,j}\,\sigma_j^2,$$

$$f_j(t)=\mathsf{E}e^{itX_j},\ j=1,2,\ldots,n,\quad \overline{f}_n(t)=\prod_{j=1}^n f_j\left(\frac{t}{B_n}\right),$$

$$r_n(t)=\left|\overline{f}_n(t)-e^{-t^2/2}\right|,\quad t\in\mathbf{R}.$$

As is well-known, if $X_1,\ldots,X_n$ are identically distributed, then

$$\overline{f}_n(t)=\left(f_1\left(\frac{t}{\sigma\sqrt{n}}\right)\right)^n,\quad t\in\mathbf{R}.$$

In this section new estimates for $|\overline{f}_n(t)|$ and $r_n(t)$ will be obtained.

Let $\theta_0(\delta)$ be the unique root of the equation

$$\delta\theta^2+2\theta\sin\theta+2(2+\delta)(\cos\theta-1)=0$$

within the interval $(0,2\pi)$. As this is so, $\pi<\theta_0(\delta)<2\pi$ for all $0<\delta\leqslant1$. Let

$$\varkappa_\delta\equiv\sup_{x>0}\frac{\left|\cos x-1+x^2/2\right|}{x^{2+\delta}}=\frac{\cos\theta_0(\delta)-1+\theta_0^2(\delta)/2}{\theta_0^{2+\delta}(\delta)}=\frac{\theta_0(\delta)-\sin\theta_0(\delta)}{(2+\delta)\theta_0^{1+\delta}(\delta)}.\quad(2.1)$$

Obviously,

$$\varkappa_\delta \leqslant \frac{1}{2\theta_0^\delta(\delta)} \leqslant \frac{1}{2\pi^\delta} \leqslant 1/2, \quad 0<\delta\leqslant 1. \tag{2.2}$$

For $\varepsilon>0$ let

$$\psi_\delta(t,\varepsilon)=\begin{cases} t^2/2-\varkappa_\delta\varepsilon|t|^{2+\delta}, & |t|<\theta_0(\delta)\varepsilon^{-1/\delta},\\ \dfrac{1-\cos\left(\varepsilon^{1/\delta}t\right)}{\varepsilon^{2/\delta}}, & \theta_0(\delta)\leqslant\varepsilon^{1/\delta}|t|\leqslant 2\pi,\\ 0, & |t|>2\pi\varepsilon^{-1/\delta}. \end{cases}$$

It is easy to see that the function $\psi_\delta(t,\varepsilon)$ decreases monotonically in ε for each fixed $t\in\mathbf{R}$ and all $0<\delta\leqslant 1$. Moreover, $\psi_\delta(t,\varepsilon)\geqslant 0$ for all $t\in\mathbf{R}$.

The following lemma plays the key role for the construction of estimates of the absolute value of a ch.f.

Lemma 2.1 (see [26]). *For any $x\in\mathbf{R}$ and $\theta_0(\delta)\leqslant\theta\leqslant 2\pi$*

$$\cos x\leqslant 1-a(\delta,\theta)x^2+b(\delta,\theta)|x|^{2+\delta},$$

where

$$a(\delta,\theta)=\frac{2+\delta}{\delta}\cdot\frac{1-\cos\theta}{\theta^2}-\frac{1}{\delta}\cdot\frac{\sin\theta}{\theta},$$
$$b(\delta,\theta)=\frac{2}{\delta}\cdot\frac{1-\cos\theta}{\theta^{2+\delta}}-\frac{1}{\delta}\cdot\frac{\sin\theta}{\theta^{1+\delta}}.$$

Theorem 2.2. *For any $F_1,\ldots,F_n\in\mathcal{F}_{2+\delta}$ and any $t\in\mathbf{R}$*

$$|\overline{f}_n(t)|\leqslant\left[1-\frac{2}{n}\,\psi_\delta(t,\varepsilon_n)\right]^{n/2}\leqslant\exp\{-\psi_\delta(t,\varepsilon_n)\}\leqslant\exp\left\{-t^2/2+\varkappa_\delta\varepsilon_n|t|^{2+\delta}\right\}.$$

Proof. Let X_j' be an independent copy of the r.v. X_j, $j=1,\ldots,n$. Then

$$\left|\overline{f}_n(t)\right|^2=\prod_{j=1}^n\left|f_j\left(\frac{t}{B_n}\right)\right|^2=\prod_{j=1}^n\mathsf{E}\cos\frac{t(X_j-X_j')}{B_n}.$$

Using lemma 2.1 and relations $\mathsf{E}(X_j-X_j')^2=2\sigma_j^2$, $\mathsf{E}|X_j-X_j'|^{2+\delta}\leqslant 2\big(\beta_{2+\delta,j}+\beta_{\delta,j}\,\sigma_j^2\big)$ (see, e. g., [34, p. 74, lemma 2.1.7]) we obtain

$$|\overline{f}_n(t)|^2\leqslant\prod_{j=1}^n\left(1-a(\delta,\theta)\frac{t^2\mathsf{E}(X_j-X_j')^2}{B_n^2}+b(\delta,\theta)\frac{|t|^{2+\delta}\mathsf{E}|X_j-X_j'|^{2+\delta}}{B_n^{2+\delta}}\right)$$
$$\leqslant\prod_{j=1}^n\left(1-2a(\delta,\theta)t^2\frac{\sigma_j^2}{B_n^2}+2b(\delta,\theta)|t|^{2+\delta}\frac{\beta_{2+\delta,j}+\beta_{\delta,j}\,\sigma_j^2}{B_n^{2+\delta}}\right).$$

The expression in brackets is an upper bound for the squared absolute value of the ch.f. $f_j(t)$ and, hence, is nonnegative. Since the geometric mean of nonnegative numbers is no greater than their arithmetic mean, for all $t \in \mathbf{R}$ and $\theta \in [\theta_0(\delta), 2\pi]$ we obtain

$$|\overline{f}_n(t)|^2 \leqslant \left[1 - \frac{2}{n}\sum_{j=1}^{n}\left(a(\delta,\theta)t^2\frac{\sigma_j^2}{B_n^2} - b(\delta,\theta)|t|^{2+\delta}\frac{\beta_{2+\delta,j}+\beta_{\delta,j}\,\sigma_j^2}{B_n^{2+\delta}}\right)\right]^n$$
$$= \left[1 - \frac{2}{n}\left(a(\delta,\theta)t^2 - b(\delta,\theta)\varepsilon_n|t|^{2+\delta}\right)\right]^n \equiv \left[1 - \frac{2}{n}\psi_\delta(t,\varepsilon_n,\theta)\right]^n,$$

where

$$\psi_\delta(t,\varepsilon,\theta) = a(\delta,\theta)t^2 - b(\delta,\theta)\varepsilon|t|^{2+\delta}, \quad t \in \mathbf{R},\ \varepsilon > 0,\ \theta_0(\delta) \leqslant \theta \leqslant 2\pi.$$

It can be made sure (see, e. g., [26]) that for any fixed $t \in \mathbf{R}$ the minimum of the right-hand side of the last estimate for $|\overline{f}_n(t)|^2$ is attained at

$$\theta = \min\left\{\max\left\{\theta_0(\delta), \varepsilon_n^{1/\delta}|t|\right\}, 2\pi\right\},$$

and

$$\psi_\delta(t,\varepsilon) = \max_{\theta_0(\delta)\leqslant\theta\leqslant 2\pi}\psi_\delta(t,\varepsilon,\theta) \geqslant \psi_\delta(t,\varepsilon,\theta_0(\delta)) = t^2/2 - \varkappa_\delta\varepsilon|t|^{2+\delta},$$

whence follows the statement of the lemma. □

For $n = 1$ from theorem 2.2 we obtain

Corollary 2.3. *For any r.v. $X \in \mathcal{F}_{2+\delta}$ for all $t \in \mathbf{R}$ there hold the estimates*

$$\left|\mathsf{E}e^{itX}\right|^2 \leqslant 1 - 2\psi_\delta\left(\sigma t, \beta_{2+\delta}/\sigma^{2+\delta} + \beta_\delta/\sigma^\delta\right) \leqslant 1 - \sigma^2t^2 + 2\varkappa_\delta\left(\beta_{2+\delta} + \beta_\delta\sigma^2\right)|t|^{2+\delta}.$$

Remark 2.4. For $\delta = 1$, in the paper of H. Prawitz [24] the first inequality of corollary 2.3 is proved as well as the second inequality of theorem 2.2. In the book of N. G. Ushakov [34] the second inequality of corollary 2.3 is proved for arbitrary $0 < \delta \leqslant 1$.

Remark 2.5. From corollary 2.3 it follows that $|f(t)|<1$ for $|t| < 2\pi(\beta_{2+\delta}/\sigma^2 + \beta_\delta)^{-1/\delta}$ for any d.f. $F \in \mathcal{F}_{2+\delta}$. A special role of the point $t = 2\pi(\beta_{2+\delta}/\sigma^2 + \beta_\delta)^{-1/\delta}$ is due to the fact that this is the least possible period of the ch.f. of a r.v. with fixed three absolute moments β_δ, σ^2 and $\beta_{2+\delta}$. Indeed, for the symmetric distribution $\mathsf{P}(X = \pm a) = 1/(2a^2)$, $\mathsf{P}(X = 0) = 1 - 1/a^2$ with $a = 1/\sqrt{2^\delta - 1}$ we have $\beta_\delta = a^{\delta-2}$, $\sigma^2 = 1$, $\beta_{2+\delta} = a^\delta$. It is easy to see that the ch.f. $f(t) = \mathsf{E}\cos(tX) = 1 - (1 - \cos(at))/a^2$ equals 1 for $t = \pi/a$, and with a specified above

$$\frac{\pi}{a} = \frac{2\pi}{a(1 + a^{-2})^{1/\delta}} = \frac{2\pi}{(\beta_{2+\delta} + \beta_\delta)^{1/\delta}}.$$

The fact mentioned in remark 2.5 can be used for the improvement of the von Mises inequality

$$\frac{h}{\sigma} \leqslant 2\frac{\beta_3}{\sigma^3},$$

relating the span of a lattice distribution with its moments. Namely, from corollary 2.3 it follows that

$$t_0 = \inf\{t > 0\colon |f(t)| = 1\} \geqslant 2\pi(\beta_{2+\delta}/\sigma^2 + \beta_\delta)^{-1/\delta}.$$

As is known, $t_0 < \infty$ if and only if $F \in \mathcal{F}^h_{2+\delta}$ with $h = 2\pi/t_0$. So, the following theorem holds.

Theorem 2.6. *For any $h > 0$ and $X \in \mathcal{F}^h_{2+\delta}$*

$$h \leqslant (\beta_{2+\delta}/\sigma^2 + \beta_\delta)^{1/\delta}. \tag{2.3}$$

For all $0 < \delta \leqslant 1$, this inequality is unimprovable in the sense that for any $h > 0$ we have

$$\sup\left\{h(\beta_{2+\delta}/\sigma^2 + \beta_\delta)^{-1/\delta}\colon X \in \mathcal{F}^h_{2+\delta}\right\} = 1, \quad 0 < \delta \leqslant 1,$$

moreover, the supremum is attained at the family of distributions of the form

$$\mathsf{P}\left(X = \frac{h}{1+u}\right) = \frac{u}{1+u} = 1 - \mathsf{P}\left(X = -\frac{uh}{1+u}\right), \quad u \to \infty.$$

For $\delta = 1$ the supremum is also attained at the extremal distribution $\mathsf{P}(X = h/2) = \mathsf{P}(X = -h/2) = 1/2$.

Theorem 2.2 and inequality (2.3) also improve the results of paper [26], in which $\sigma^\delta \geqslant \beta_\delta$ is used instead of β_δ.

Lemma 2.7. *For any $F_1, \ldots, F_n \in \mathcal{F}_{2+\delta}$ and $t \in \mathbf{R}$*

$$\begin{aligned} r_n(t) &\equiv \left|\overline{f}_n(t) - e^{-t^2/2}\right| \\ &\leqslant \sum_{j=1}^n \left| f_j\left(\frac{t}{B_n}\right) - \exp\left\{-\frac{\sigma_j^2 t^2}{2B_n^2}\right\}\right| \exp\left\{-\frac{t^2}{2}\left(1 - \frac{\sigma_j^2}{B_n^2}\right) + \varkappa_\delta \varepsilon_n |t|^{2+\delta}\right\}. \end{aligned}$$

Proof. In [25] it was proved that for any $A_j > 0$, $B_j \in \mathbb{C}$, $C_j \geqslant \max\{A_j, |B_j|\}$

$$\begin{aligned} \left|\prod_{j=1}^n B_j - \prod_{j=1}^n A_j\right| &\leqslant \frac{1}{2}\prod_{i=1}^n C_i \sum_{j=1}^n \frac{|B_j - A_j|}{C_j} + \frac{1}{2}\prod_{i=1}^n A_i \sum_{j=1}^n \frac{|B_j - A_j|}{A_j} \\ &\leqslant \sum_{j=1}^n \frac{|B_j - A_j|}{A_j}\prod_{i=1}^n C_i. \end{aligned}$$

Using this inequality with

$$B_j = f_j\left(\frac{t}{B_n}\right), \quad A_j = \exp\left\{-\frac{\sigma_j^2 t^2}{2B_n^2}\right\},$$

$$C_j = \exp\left\{-\frac{\sigma_j^2 t^2}{2B_n^2} + \varkappa_\delta(\beta_{2+\delta,j} + \beta_{\delta,j}\,\sigma_j^2)\frac{|t|^{2+\delta}}{B_n^{2+\delta}}\right\}$$

(the estimate $|B_j| \leqslant C_j$ follows from theorem 2.2), for $r_n(t)$ we obtain

$$r_n(t) = \left|\prod_{j=1}^n f_j\left(\frac{t}{B_n}\right) - \prod_{j=1}^n \exp\left\{-\frac{\sigma_j^2 t^2}{2B_n^2}\right\}\right| \leqslant$$
$$\leqslant \sum_{j=1}^n \left|f_j\left(\frac{t}{B_n}\right) - \exp\left\{-\frac{\sigma_j^2 t^2}{2B_n^2}\right\}\right| \exp\left\{-\frac{t^2}{2} + \varkappa_\delta \varepsilon_n |t|^{2+\delta} + \frac{\sigma_j^2 t^2}{2B_n^2}\right\}. \quad \square$$

The way we estimate $|f_j(t/B_n) - e^{-\sigma_j^2 t^2/(2B_n^2)}|$ in lemma 2.7 depends on whether $\delta = 1$ or not.

Lemma 2.8. *For any r.v. $X \in \mathcal{F}_{2+\delta}$ with the ch.f. $f(t)$ for all $t \in \mathbf{R}$ we have the estimates:*
if $\delta = 1$, then

$$\left|f(t) - e^{-\sigma^2 t^2/2}\right| \leqslant \frac{\beta_3 |t|^3}{6}, \tag{2.4}$$

$$\left|f(t) - e^{-\sigma^2 t^2/2}\right| \leqslant \frac{|t|^3}{6}\Big(\left|\mathsf{E}X^3\mathbf{1}(|X| \leqslant U)\right| + \mathsf{E}|X|^3\mathbf{1}(|X| > U)\Big) + $$
$$+ \frac{t^4}{24}\mathsf{E}|X|^4\mathbf{1}(|X| \leqslant U) + \frac{\sigma^4 t^4}{8} \tag{2.5}$$

for all $U \geqslant 0$;
if $0 < \delta \leqslant 1$, then

$$\left|f(t) - e^{-\sigma^2 t^2/2}\right| \leqslant \gamma_\delta \beta_{2+\delta}\,|t|^{2+\delta} + \sigma^4 t^4/8, \tag{2.6}$$

where

$$\gamma_\delta = \sup_{x>0}\left|e^{ix} - 1 - ix - (ix)^2/2\right|/x^{2+\delta}$$
$$= \sup_{x>0}\sqrt{\left(\frac{\cos x - 1 + x^2/2}{x^{2+\delta}}\right)^2 + \left(\frac{\sin x - x}{x^{2+\delta}}\right)^2}.$$

The values of γ_δ for some $0 < \delta \leqslant 1$ are presented in the second column of table 3. In particular, $\gamma_1 = 1/6$. The estimates given in lemma 2.8 were apparently first obtained for the case $0 < \delta < 1$ by W. Tysiak [30]. Nevertheless, for completeness we give their simple proof as well.

Proof. The first estimate follows from the works of I. Tyurin [31, 32], in which the inequality

$$\left|f(t) - e^{-\sigma^2 t^2/2}\right| \leqslant e^{-t^2/2}\int_0^{|t|}\frac{\beta_3 s^2}{2}\,e^{s^2/2}\,ds \leqslant \int_0^{|t|}\frac{\beta_3 s^2}{2}\,ds = \frac{\beta_3|t|^3}{6}, \quad t \in \mathbf{R},$$

was proved.

Further, using the inequality $|e^{-x}-1+x|\leqslant x^2/2$, $x\geqslant 0$, for all $t\in\mathbf{R}$ we obtain

$$|f(t)-e^{-\sigma^2t^2/2}|\leqslant\left|\mathsf{E}\left(e^{itX}-1-itX+\frac{t^2X^2}{2}\right)\right|+\left|e^{-\sigma^2t^2/2}-1+\frac{\sigma^2t^2}{2}\right|$$
$$\leqslant R(t)+\frac{\sigma^4t^4}{8},$$

where

$$R(t)=\left|\mathsf{E}\left(e^{itX}-1-itX-\frac{(itX)^2}{2}\right)\right|\leqslant R_1(t,U)+R_2(t,U),$$
$$R_1(t,U)=\left|\mathsf{E}\left(e^{itX}-1-itX-\frac{(itX)^2}{2}\right)\mathbf{1}(|X|\leqslant U)\right|,$$
$$R_2(t,U)=\mathsf{E}\left|e^{itX}-1-itX-\frac{(itX)^2}{2}\right|\mathbf{1}(|X|>U)$$

for any $U\geqslant 0$.

By the definition of γ_δ, $\left|e^{ix}-1-ix-(ix)^2/2\right|\leqslant\gamma_\delta|x|^{2+\delta}$, $x\in\mathbf{R}$, whence for $R_2(t,U)$ we obtain

$$R_2(t,U)\leqslant\gamma_\delta|t|^{2+\delta}\mathsf{E}|X|^{2+\delta}\mathbf{1}(|X|>U).$$

Adding and subtracting $(itX)^3/6\cdot\mathbf{1}(|X|\leqslant U)$ under the sign of expectation in $R_1(t,U)$, taking account of the inequality $\left|e^{ix}-1-ix-(ix)^2/2-(ix)^3/6\right|\leqslant x^4/24$, $x\in\mathbf{R}$, for $R_1(t,U)$ we obtain

$$R_1(t,U)\leqslant\left|\mathsf{E}\left(e^{itX}-1-itX-\frac{(itX)^2}{2}-\frac{(itX)^3}{6}\right)\mathbf{1}(|X|\leqslant U)\right|$$
$$+\frac{|t|^3}{6}\left|\mathsf{E}X^3\mathbf{1}(|X|\leqslant U)\right|\leqslant\frac{t^4}{24}\mathsf{E}X^4\mathbf{1}(|X|\leqslant U)+\frac{|t|^3}{6}\left|\mathsf{E}X^3\mathbf{1}(|X|\leqslant U)\right|.$$

So, for any $0<\delta\leqslant 1$ and $U\geqslant 0$ for all $t\in\mathbf{R}$ we have

$$|f(t)-e^{-\sigma^2t^2/2}|\leqslant\frac{\sigma^4t^4}{8}+\gamma_\delta|t|^{2+\delta}\mathsf{E}|X|^{2+\delta}\mathbf{1}(|X|>U)$$
$$+\frac{|t|^3}{6}\left|\mathsf{E}X^3\mathbf{1}(|X|\leqslant U)\right|+\frac{t^4}{24}\mathsf{E}X^4\mathbf{1}(|X|\leqslant U).$$

Setting $U=0$ in this inequality, we obtain the second estimate of the lemma, setting $\delta=1$ we obtain the third one. The lemma is completely proved. □

Remark 2.9. Note that using new optimal estimates for ζ-metrics obtained in [33], we can as well prove an analog of the first estimate of lemma 2.8 for the case of an arbitrary $0<\delta<1$ in the form

$$\left|f(t)-e^{-t^2/2}\right|\leqslant\frac{\beta_{2+\delta}|t|^{2+\delta}}{(1+\delta)(2+\delta)}\sup_{x>0}\frac{|e^{ix}-1|}{x^\delta},$$

however, it turns out that for all $0 < \delta < 1$

$$\frac{1}{(1+\delta)(2+\delta)}\sup_{x>0}\frac{|e^{ix}-1|}{x^\delta} > \sup_{x>0}\frac{\left|e^{ix}-1-ix-(ix)^2/2\right|}{x^{2+\delta}} = \gamma_\delta,$$

that is, the coefficient at $\beta_{2+\delta}|t|^{2+\delta}$ in this estimate will be greater than that in the third estimate of lemma 2.8. This circumstance is critical for the estimation of the remainder in the central limit theorem since it is this coefficient that determines the value of the constant at the main term. This is the reason why the third estimate of lemma 2.8 is more preferable, and will be used for our purposes.

3. The moment inequality

Theorem 3.1. *For any r.v. $X \in \mathcal{F}_3$, for all $\lambda \geqslant 1$ the inequality*

$$|\mathsf{E}X^3| + 3\mathsf{E}|X|\cdot\mathsf{E}X^2 \leqslant \lambda\mathsf{E}|X|^3 + M(p(\lambda),\lambda)(\mathsf{E}X^2)^{3/2}$$

holds, where

$$p(\lambda) = \frac{1}{2} - \sqrt{\frac{\lambda+1}{\lambda+3}}\sin\left(\frac{\pi}{6} - \frac{1}{3}\arctan\sqrt{\lambda^2 + 2\frac{\lambda-1}{\lambda+3}}\right),$$

$$M(p,\lambda) = \frac{1-\lambda+2(\lambda+2)p-2(\lambda+3)p^2}{\sqrt{p(1-p)}}, \quad 0 < p \leqslant \frac{1}{2},\ \lambda \geqslant 1,$$

with equality attained for each $\lambda \geqslant 1$ at the family of two-point distributions $\big\{\mathsf{P}\big(X = \sigma\sqrt{q/p}\big) = p = 1 - \mathsf{P}\big(X = -\sigma\sqrt{p/q}\big) : \sigma > 0\big\}$, where $p = p(\lambda)$, $q = 1 - p(\lambda)$.

The optimal values of the parameter $\lambda = \lambda(\beta_3)$, delivering the minimum to the right-hand side of the inequality in theorem 3.1 and the corresponding values $p = p(\lambda(\beta_3))$ are presented for some $\beta_3 = \mathsf{E}|X|^3/(\mathsf{E}X^2)^{3/2}$ in the fourth and seventh columns, respectively, of table 2 below.

Remark 3.2. It can be made sure that the function $p(\lambda)$ increases monotonically for $\lambda \geqslant 1$, varying within the limits

$$0.3169\ldots = \frac{1}{2}\left(1 - \sqrt{1-\sqrt{3}/2}\right) = p(1) \leqslant p(\lambda) \leqslant \lim_{\lambda\to\infty} p(\lambda) = \frac{1}{2}.$$

Moreover, as it will be seen from the proof, the function $M(p(\lambda),\lambda)$ can be represented as

$$M(p(\lambda),\lambda) = \sup_{0<p\leqslant 1/2}(\alpha_3(p) - \lambda\beta_3(p) + 3\beta_1(p)),$$

where $\alpha_3(p)$, $\beta_3(p)$, $\beta_1(p)$ are, respectively, the third original, third absolute and first absolute moments of the Bernoulli distribution assigning the probabilities p and $q = 1 - p$ to the points $\sqrt{q/p}$ and $-\sqrt{p/q}$:

$$M(p(\lambda),\lambda) = \sup\left\{\frac{q-p-\lambda(p^2+q^2)+6pq}{\sqrt{pq}} : 0 < p \leqslant \frac{1}{2},\ q = 1-p\right\}. \tag{3.1}$$

From this representation, first, it follows that the function $M(p,\lambda)$ decreases monotonically in $\lambda \geqslant 1$ for each $0 < p \leqslant 1/2$. The same property is inherent in $M(p(\lambda),\lambda)$, since for any $\lambda_1 \geqslant \lambda_2 \geqslant 1$ we have

$$M(p(\lambda_1),\lambda_1) \leqslant M(p(\lambda_1),\lambda_2) \leqslant \sup_{0<p<1/2} M(p,\lambda_2) = M(p(\lambda_2),\lambda_2).$$

Second, evidently,

$$M(p(\lambda),\lambda) \geqslant \frac{q-p-\lambda(p^2+q^2)+6pq}{\sqrt{pq}}\bigg|_{p=q=1/2} = 3-\lambda, \quad \lambda \geqslant 1,$$

with equality attained at $\lambda \to \infty$, so that

$$\inf_{\lambda \geqslant 1} \big(\lambda + M(p(\lambda),\lambda)\big) = \lim_{\lambda\to\infty} \big(\lambda + M(p(\lambda),\lambda)\big) = 3.$$

Thus, the function $M(p(\lambda),\lambda)$ decreases monotonically for all $\lambda \geqslant 1$, varying within the limits

$$2.3599\ldots = 2\sqrt{3\sqrt{3}(2-\sqrt{3})} = M(p(1),1) \geqslant M(p(\lambda),\lambda) > \lim_{\lambda\to\infty} M(p(\lambda),\lambda) = -\infty,$$

whence it follows that $M(p(\lambda),\lambda)$ alters its sign at the unique point $\lambda = \sqrt{10}$ corresponding to the value $p(\sqrt{10}) = 2 - \sqrt{10}/2 = 0.4188\ldots$, so that

$$M(p(\lambda),\lambda) < 0 \quad \Longleftrightarrow \quad \lambda > \sqrt{10}.$$

Since $p^2 + q^2 - \sqrt{pq} = -2pq - \sqrt{pq} + 1 > 0$ for all $p \in (0,1/2)$, $q = 1-p$, from (3.1) it also follows that the function

$$\lambda + M(p(\lambda),\lambda) = \sup\left\{\frac{q-p-\lambda(p^2+q^2-\sqrt{pq})+6pq}{\sqrt{pq}} : 0 < p \leqslant \frac{1}{2},\ q = 1-p\right\}$$

decreases monotonically, varying within the limits

$$3 < \lambda + M(p(\lambda),\lambda) \leqslant 1 + 2\sqrt{3\sqrt{3}(2-\sqrt{3})} = 3.3599\ldots, \quad \lambda \geqslant 1. \tag{3.2}$$

Using theorem 3.1 it is possible to improve a result due to C.-G. Esseen [10], according to which for a sequence of independent r.v.'s $X_1, X_2\ldots$ with the d.f. $F \in \mathcal{F}_3^h$ for some $h > 0$ such that $\mathsf{E}X_1^2 = 1$, $\mathsf{E}X_1^3 = \alpha_3$, $\mathsf{E}|X_1|^3 = \beta_3$, the relation

$$\psi(F) \equiv \limsup_{n\to\infty} \Delta_n\sqrt{n} = \frac{|\alpha_3| + 3h}{6\sqrt{2\pi}} \leqslant \frac{\sqrt{10}+3}{6\sqrt{2\pi}}\,\beta_3 \equiv \psi_1(\beta_3)$$

holds (see (1.4) and (1.6)).

On the other hand, according to (2.3) for h we have the estimate $h \leqslant \beta_3 + \beta_1$, whence it follows that in the case considered by Esseen

$$\psi(F) \leqslant \frac{|\alpha_3| + 3(\beta_3 + \beta_1)}{6\sqrt{2\pi}} \leqslant \inf_{\lambda \geqslant 1} \frac{(\lambda+3)\beta_3 + M(p(\lambda), \lambda)}{6\sqrt{2\pi}} =$$
$$= \inf_{c \geqslant 2/(3\sqrt{2\pi})} (c\beta_3 + K(c)) \equiv \psi_2(\beta_3), \quad (3.3)$$

where

$$K(c) = \frac{M(p(\lambda), \lambda)}{6\sqrt{2\pi}}\bigg|_{\lambda = 6\sqrt{2\pi}c - 3}.$$

Moreover, from theorem 3.1 it follows that c cannot be less than $2/(3\sqrt{2\pi}) = 0.2659\ldots$, and $K(c)$ in (3.3) can be made less for no $c \geqslant 2/(3\sqrt{2\pi})$. From (3.3) with $c = (\sqrt{10}+3)/(6\sqrt{2\pi}) = 0.4097\ldots$ (that corresponds to $\lambda = \sqrt{10}$, $K(c) = 0$) Esseen's bound follows, whereas (3.3) with $c = 2/(3\sqrt{2\pi})$ (that corresponds to $\lambda = 1$) implies the estimate

$$\psi(F) \leqslant \frac{2}{3\sqrt{2\pi}} \cdot \beta_3 + \sqrt{\frac{2\sqrt{3}-3}{6\pi}} < 0.2660\beta_3 + 0.1570, \quad (3.4)$$

which is more accurate than Esseen's bound $\psi(F) \leqslant \psi_1(\beta_3)$ for

$$\beta_3 \geqslant \frac{2\sqrt{3\sqrt{3}(2-\sqrt{3})}}{\sqrt{10}-1} = 1.0914\ldots,$$

although the value $c = 2/(3\sqrt{2\pi})$ (that is, $\lambda = 1$) is optimal in (3.3) only for $\beta_3 \geqslant 1.2185\ldots$

Comparing the functions $\psi_1(\beta_3)$ and $\psi_2(\beta_3)$, we conclude that their values coincide only at the unique point β_3 for which $c = (\sqrt{10}+3)/(6\sqrt{2\pi})$, $K(c) = 0$ (that corresponds to $\lambda = \sqrt{10}$, $p(\sqrt{10}) = 2 - \sqrt{10}/2$), that is, at the point

$$\beta_3 = \frac{p^2 + (1-p)^2}{\sqrt{p(1-p)}}\bigg|_{p = 2-\sqrt{10}/2} = \sqrt{20(\sqrt{10}-3)/3} = 1.0401\ldots,$$

and for all the rest of the values of $\beta_3 \geqslant 1$ the strict inequality $\psi_1(\beta_3) > \psi_2(\beta_3)$ holds. In particular, for $\beta_3 = 1$ (that is, for the symmetric Bernoulli distribution) $\psi_1(1) = (\sqrt{10}+3)/(6\sqrt{2\pi}) = 0.4097\ldots$, while

$$\psi_2(1) = \lim_{c\to\infty}(c + K(c)) = \lim_{\lambda\to\infty} \frac{\lambda + 3 + M(p(\lambda), \lambda)}{6\sqrt{2\pi}}$$
$$= \frac{1}{\sqrt{2\pi}} = 0.3989\ldots < \psi_1(1) - 0.0107.$$

The values of the functions $\psi_1(\beta_3)$ and $\psi_2(\beta_3)$ for some $\beta_3 \geqslant 1$ are presented in the second and third columns of table 2. The corresponding values of $c = c(\beta_3)$ and $K = K(c(\beta_3))$ delivering the minimum in (3.3) are presented in the fifth and sixth columns of table 2.

β_3	ψ_1	ψ_2	λ	c	K	p
1	0.4097	0.3989	+ inf	+ inf	− inf	1/2
1.01	0.4138	0.4111	7.2034	0.6784	-0.2741	0.4592
1.02	0.4179	0.4170	4.8305	0.5206	-0.1141	0.4424
1.03	0.4220	0.4218	3.7862	0.4512	-0.0430	0.4296
1.04	0.4261	0.4261	3.1682	0.4101	-0.0005	0.4189
1.05	0.4302	0.4300	2.7497	0.3823	0.0286	0.4095
1.06	0.4343	0.4337	2.4432	0.3619	0.0501	0.4011
1.07	0.4384	0.4373	2.2070	0.3462	0.0668	0.3934
1.08	0.4425	0.4407	2.0182	0.3336	0.0803	0.3863
1.09	0.4466	0.4440	1.8633	0.3233	0.0915	0.3796
1.10	0.4507	0.4471	1.7335	0.3147	0.1009	0.3733
1.12	0.4589	0.4533	1.5275	0.3010	0.1161	0.3618
1.14	0.4670	0.4592	1.3707	0.2906	0.1279	0.3513
1.16	0.4752	0.4649	1.2470	0.2823	0.1374	0.3416
1.18	0.4834	0.4705	1.1470	0.2757	0.1451	0.3326
1.20	0.4916	0.4760	1.0645	0.2702	0.1517	0.3243
1.21	0.4957	0.4787	1.0284	0.2678	0.1546	0.3203
1.22	0.4998	0.4813	1.0000	0.2659	0.1569	0.3169

Table 2: The values of the functions $\psi_1(\beta_3)$ and $\psi_2(\beta_3)$ for some β_3; optimal values of $c = (\lambda+3)/(6\sqrt{2\pi})$ delivering the minimum to $\psi_2(\beta_3)$ (see (3.3)); the corresponding values of $K(c)$ in (3.3); the parameter $p(\lambda)$ of the extremal distribution.

Proof of theorem 3.1. Since for $\sigma^2 \equiv \mathsf{E}X^2 = 0$ the statement of the theorem is obvious, in what follows we assume that $\sigma > 0$. Consider the functional

$$J_{\lambda,\sigma}(X) = \left(|\mathsf{E}X^3| + 3\mathsf{E}|X|\sigma^2 - \lambda\mathsf{E}|X|^3\right)/\sigma^3, \quad X \in \mathcal{F}_3.$$

Then the statement of the theorem is equivalent to

$$\sup_{\sigma>0}\ \sup_{X\in\mathcal{F}_3\colon \mathsf{E}X=0,\ \mathsf{E}X^2=\sigma^2} J_{\lambda,\sigma}(X) = M(p(\lambda),\lambda).$$

On the other hand, for any $\sigma > 0$

$$\sup_{X\in\mathcal{F}_3\colon \mathsf{E}X=0,\ \mathsf{E}X^2=\sigma^2} J_{\lambda,\sigma}(X) = \sup_{X\in\mathcal{F}_3\colon \mathsf{E}X=0,\ \mathsf{E}X^2=\sigma^2} J_{\lambda,\sigma}(-X)$$
$$= \sup_{X\in\mathcal{F}_3\colon \mathsf{E}X=0,\ \mathsf{E}X^2=\sigma^2} \widetilde{J}_{\lambda,\sigma}(X),$$

where

$$\widetilde{J}_{\lambda,\sigma}(X) = \left(\mathsf{E}X^3 + 3\mathsf{E}|X|\sigma^2 - \lambda\mathsf{E}|X|^3\right)/\sigma^3.$$

With the account of the results of W. Hoeffding [15] and V. M. Zolotarev [40] it is easy to see that for each $\sigma > 0$ the extremum of the moment-type functional

$\widetilde{J}_{\lambda,\sigma}(X)$ linear with respect to $F \in \mathcal{F}_3$ under two moment-type restrictions

$$\mathsf{E}X = 0, \quad \mathsf{E}X^2 = \sigma^2,$$

is attained on distributions concentrated in at most three points. Without loss of generality assume that the r.v. X takes the values $x < y \leqslant 0 < z$ with the probabilities

$$\mathsf{P}(X = x) = \frac{\sigma^2 + yz}{(z-x)(y-x)}, \quad \mathsf{P}(X = y) = -\frac{\sigma^2 + xz}{(z-y)(y-x)},$$

$$\mathsf{P}(X = z) = \frac{\sigma^2 + xy}{(z-x)(z-y)}, \quad -yz \leqslant \sigma^2 \leqslant -xz.$$

Then

$$\mathsf{E}|X| = \frac{2z(\sigma^2 + xy)}{(x-z)(y-z)}, \quad 3\mathsf{E}|X|\sigma^2 = \frac{6z\sigma^4 + 6xyz\sigma^2}{(x-z)(y-z)},$$

$$\mathsf{E}|X|^3 = \frac{(z^3 + a)\sigma^2 - xyz(xy - xz - yz - z^2)}{(z-x)(z-y)},$$

$$\mathsf{E}X^3 = (x+y+z)\sigma^2 + xyz = \frac{(z^3 - a)\sigma^2 + xyz(xy - xz - yz + z^2)}{(z-x)(z-y)},$$

$$a = a(x,y,z) = z(x^2 + y^2 + xy) - xy(x+y) > 0, \quad x < y \leqslant 0 < z,$$

$$\widetilde{J}_{\lambda,\sigma}(X) = \big(6z\sigma + (6xyz - (\lambda-1)z^3 - a(\lambda+1))\sigma^{-1} + \\ + xyz((\lambda+1)(xy - xz - yz) - (\lambda-1)z^2)\sigma^{-3}\big)/((z-x)(z-y))$$

and

$$\sup_{\sigma>0} \sup_{X \in \mathcal{F}_3 \colon \mathsf{E}X=0,\ \mathsf{E}X^2=\sigma^2} \widetilde{J}_{\lambda,\sigma}(X) = \sup_{X \in \mathcal{F}_3 \colon \mathsf{E}X=0} \sup_{\sigma>0} \frac{g(\sigma)}{(z-x)(z-y)},$$

where

$$g(\sigma) = g(\sigma, x, y, z, \lambda) = 6z\sigma + (6xyz - (\lambda-1)z^3 - a(\lambda+1))\sigma^{-1} +$$

$$+ xyz((\lambda+1)(xy - xz - yz) - (\lambda-1)z^2)\sigma^{-3}.$$

Show that the function $g(\sigma)$ is quasi-convex for $\sigma > 0$, namely, either $g(\sigma)$ increases monotonically for $\sigma > 0$ or there exists a point $\sigma_1 > 0$ such that $g(\sigma)$ decreases monotonically for $0 < \sigma < \sigma_1$ and increases monotonically for $\sigma > \sigma_1$. For this purpose differentiate $g(\sigma)$ and find the stationary points. We have

$$g'(\sigma) = 6z + (a(\lambda+1) + (\lambda-1)z^3 - 6xyz)\sigma^{-2} \\ - 3xyz((\lambda+1)(xy - xz - yz) - (\lambda-1)z^2)\sigma^{-4} \geqslant 0$$

if and only if

$$6\sigma^4+(a(\lambda+1)/z+(\lambda-1)z^2-6xy)\sigma^2+3xy((\lambda-1)z^2-(\lambda+1)(xy-xz-yz))\geqslant 0.$$

So, the equation $g'(\sigma)=0$ is equivalent to the quadratic equation with respect to σ^2. The latter either has no real roots and then $g'(\sigma)>0$ and $g(\sigma)$ increases, or has one real root which is the point of reflection of $g(\sigma)$ and then $g(\sigma)$ increases, or has two different real roots $\sigma_1<\sigma_2$ so that σ_1 is the point of maximum and σ_2 is the point of minimum. The desired property of the function g will be proved if we show that the smaller root σ_1 of the equation $g'(\sigma)=0$ is non-positive.

The smaller root s_1 of the quadratic equation $s^2+bs+c=0$ with two different roots has the form $s_1=-b-\sqrt{b^2-4c}$. It is obvious that $s_1\leqslant 0$ if and only if either $b>0$, or $b\leqslant 0$ and $c\leqslant 0$, that is, if the condition $b\leqslant 0$ implies $c\leqslant 0$. Apply this reasoning to $s=\sigma^2$,

$$b=(a(\lambda+1)/z+(\lambda-1)z^2-6xy)/6,\quad c=\frac{xy}{2z}((\lambda-1)z^3-(\lambda+1)z(xy-xz-yz)).$$

Indeed, the condition $b\leqslant 0$ implies $(\lambda-1)z^3\leqslant 6xyz-a(\lambda+1)$ and

$$c\cdot\frac{2z}{(\lambda+1)xy}\leqslant\frac{6xyz}{\lambda+1}-a-z(xy-xz-yz)\leqslant 3xyz-a-z(xy-xz-yz)=$$
$$=xz(y-x)-y^2z+(x+y)(xy+z^2)\leqslant 0$$

for all $\lambda\geqslant 1$ and $x<y\leqslant 0<z$. So, the maximum value of the function $g(\sigma)$ on the interval $-yz\leqslant\sigma^2\leqslant -xz$ is attained either at $\sigma^2=-yz$ and then $\mathsf{P}(X=x)=0$, or at $\sigma^2=-xz$ and then $\mathsf{P}(X=y)=0$, that is, the extremum of the functional $\widetilde{J}_{\lambda,\sigma}(X)$ is attained at two-point distributions of the r.v. X.

Now let $\mathsf{P}(X=\sigma\sqrt{q/p})=p$, $\mathsf{P}(X=-\sigma\sqrt{p/q})=q=1-p$, $0<p<1$. Then

$$\mathsf{E}X^3=\frac{q-p}{\sqrt{pq}}\,\sigma^3,\quad \mathsf{E}|X|^3=\frac{p^2+q^2}{\sqrt{pq}}\,\sigma^3=\frac{1-2pq}{\sqrt{pq}}\,\sigma^3,\quad \mathsf{E}|X|=2\sqrt{pq}\sigma.$$

Since $\mathsf{E}X^3<0$ for $p<1/2$, the range of the values of p under consideration can be restricted to the semi-interval $(0,1/2]$. Further, the functional

$$\widetilde{J}_{\lambda,\sigma}(X)=\frac{\mathsf{E}X^3-\lambda\mathsf{E}|X|^3+3\mathsf{E}|X|\sigma^2}{\sigma^3}=\frac{q-p-\lambda(1-2pq)+6pq}{\sqrt{pq}}=$$
$$=\frac{1-\lambda+2(\lambda+2)p-2(\lambda+3)p^2}{\sqrt{p(1-p)}}\equiv M(p,\lambda)$$

does not depend on σ and hence,

$$\sup_{\sigma>0}\ \sup_{X\in\mathcal{F}_3\colon\,\mathsf{E}X=0,\mathsf{E}X^2=\sigma^2}\widetilde{J}_{\lambda,\sigma}(X)=\sup_{0<p\leqslant 1/2}M(p,\lambda).$$

It remains to show that for each λ, $M(p,\lambda)$ attains its maximum value at the point $p=p(\lambda)$ specified in the formulation of Theorem 3.1.

Consider the zeroes of the derivative $M'_p(p,\lambda)$. We have

$$M'_p(p,\lambda)\cdot 2(p(1-p))^{3/2} = 4(\lambda+3)p^3-6(\lambda+3)p^2+6p+\lambda-1 \equiv h(p), \quad 0<p\leqslant 1/2.$$

Since

$$h''(p) = (12(\lambda+3)(p^2-p)+6)'_p = 12(\lambda+3)(2p-1)\leqslant 0$$

for $p\leqslant 1/2$, the function $h(p)$ is concave on the interval $(0,\,1/2]$. Moreover, $h(0+)=\lambda-1\geqslant 0$, $h(1/2)=-1<0$, that is, the function

$$h(p) = M'_p(p,\lambda)\cdot 2(p(1-p))^{3/2}$$

changes its sign at the unique point on the interval $(0,\,1/2]$, which delivers the maximum to the function $M(p,\lambda)$ for each $\lambda\geqslant 1$. It is easy to see that

$$p(\lambda) = \frac{1}{2}-\sqrt{\frac{\lambda+1}{\lambda+3}}\sin\left(\frac{\pi}{6}-\frac{1}{3}\arctan\sqrt{\lambda^2+2\frac{\lambda-1}{\lambda+3}}\right)\in(0,1/2],$$

for all $\lambda\geqslant 1$ and $h(p(\lambda))\equiv 0$, that is, $p(\lambda)$ is the point of the maximum of the function $M(p,\lambda)$. □

4. Estimates of the accuracy of the normal approximation to the distributions of sums of independent random variables

In addition to the notation introduced in section 1, let

$$\nu_n = 1+\sum_{j=1}^n \beta_{\delta,j}\,\sigma_j^2\Big/\sum_{j=1}^n\beta_{2+\delta,j}.$$

It is easy to see that the quantities ν_n, $\ell_n=\sum_{j=1}^n\beta_{2+\delta,j}$ are linked with the quantity ε_n introduced in section 2 by the relation $\varepsilon_n=\nu_n\ell_n$. Furthermore, by the Lyapounov inequality we have $1\leqslant\nu_n\leqslant 2$, and in the case of identically distributed summands we have

$$\nu_n = 1+\frac{\beta_\delta\sigma^2}{\beta_{2+\delta}}\leqslant 1+\frac{1}{n^{\delta/2}\ell_n}\leqslant 2. \tag{4.1}$$

We will also use the following inequality proved by H. Prawitz in [25]:

$$\sum_{j=1}^n\beta_{2+\delta,j}^r\leqslant\left(\sum_{j=1}^n\beta_{2+\delta,j}\right)^r = (B_n^{2+\delta}\ell_n)^r,\quad r\geqslant 1. \tag{4.2}$$

Before we proceed to the construction of new estimates of the accuracy of the normal approximation, note that

$$\kappa\equiv\sup_{F\in\mathcal{F}_2}\sup_x|F(x)-\Phi(x)| = \sup_{b>0}\left(\frac{1}{1+b^2}-\Phi(-b)\right)=0.54093\ldots \tag{4.3}$$

This relation is a consequence of lemma 12.3 from the monograph [5], establishing an upper bound for the uniform distance between F and Φ, and the paper [18] where the extremal two-point distribution was constructed. Relation (4.3) provides a universal estimate for all distributions with finite second moment. We will use this estimate for the purpose of bounding the range of the values of ℓ_n under consideration.

Recall that in section 2 by $f_j(t)$ we denoted the characteristic functions of the r.v.'s X_j, $j = 1, \ldots, n$, $\overline{f}_n(t) = \prod_{j=1}^n f_j(t/B_n)$, $r_n(t) = |\overline{f}_n(t) - e^{-t^2/2}|$.

The key role in the construction of estimates for Δ_n is played by Prawitz' smoothing inequality presented in the following lemma.

Lemma 4.1 (see [23]). *For all $n \geqslant 1$ and arbitrary d.f.'s $F_1, \ldots, F_n$ with zero expectations for any $0 < t_0 \leqslant 1$ and $T > 0$ there holds the inequality*

$$\Delta_n \leqslant 2\int_0^{t_0} |K(t)| r_n(Tt)\, dt + 2\int_{t_0}^{1} |K(t)| \cdot |\overline{f}_n(Tt)| dt +$$

$$+2\int_0^{t_0} \left|K(t) - \frac{i}{2\pi t}\right| e^{-T^2t^2/2} dt + \frac{1}{\pi}\int_{t_0}^{\infty} e^{-T^2t^2/2}\frac{dt}{t},$$

where

$$K(t) = \frac{1}{2}(1 - |t|) + \frac{i}{2}\left[(1 - |t|)\cot \pi t + \frac{\operatorname{sign} t}{\pi}\right], \quad -1 \leqslant t \leqslant 1,$$

furthermore, the function $K(t)$ satisfies the inequalities

$$|K(t)| \leqslant \frac{1.0253}{2\pi|t|}, \quad \left|K(t) - \frac{i}{2\pi t}\right| \leqslant \frac{1}{2}\left(1 - |t| + \frac{\pi^2 t^2}{18}\right), \quad -1 \leqslant t \leqslant 1.$$

The following lemma is important for the calculation of constants in the estimates of the normal approximation, to be constructed below. By $\mathcal{D}$ denote the class of real continuous nonnegative functions $J(z)$ defined for $z \geqslant 0$, which have a unique maximum and do not have a minimum for $z > 0$.

Lemma 4.2 (see [25, 11]). *Let $a < b$ and $k > 0$ be arbitrary constants, $g(s)$ and $G(s)$ be positive monotonically increasing differentiable functions on $a \leqslant s \leqslant b$. If the function*

$$\varphi(s) = \frac{G(s) - G(a)}{g^k(s)}, \quad a \leqslant s \leqslant b,$$

increases monotonically, then the function

$$J(z) = z^k \int_a^b e^{-zg(s)}\, dG(s), \quad z \geqslant 0,$$

belongs to the class $\mathcal{D}$.

If $G(a) = g(a) = 0$, then the condition that $\varphi(s)$ increases can be relaxed the requirement that the function

$$\psi(s) = \frac{G'(s)}{\left(g^k(s)\right)'}, \quad a \leqslant s \leqslant b,$$

increases.

Lemma 4.1 for all $F_1, \ldots, F_n \in \mathcal{F}_{2+\delta}$, $n \geqslant 1$, $0 < \delta \leqslant 1$, $0 < t_0 \leqslant t_1 \leqslant 1$ and $T > 0$ implies the estimate

$$\Delta_n \leqslant I_1 + I_2 + I_3 + I_4 + I_5,$$

where

$$I_1 = \frac{2}{T} \int_0^{t_0 T} \left| K\left(\frac{t}{T}\right) \right| r_n(t)\, dt,$$

$$I_2 = \frac{1.0253}{\pi} \int_{t_0}^{t_1} |\overline{f}_n(Tt)| \frac{dt}{t},$$

$$I_3 = 2 \int_{t_1}^{1} |K(t)| \cdot |\overline{f}_n(Tt)| dt,$$

$$I_4 = \int_0^{t_0} \left(1 - t + \frac{\pi^2 t^2}{18}\right) e^{-T^2 t^2/2} dt,$$

$$I_5 = \frac{1}{\pi} \int_{t_0}^{\infty} e^{-T^2 t^2/2} \frac{dt}{t}.$$

We will estimate the integrals I_2, I_3, I_4, I_5 in the same way as it was done in [25, 11]. We have

$$I_4 + I_5 = \int_0^{\infty} \left(1 - t + \frac{\pi^2 t^2}{18}\right) e^{-T^2 t^2/2}\, dt + \int_{t_0}^{\infty} \left(\frac{1}{\pi t} - 1 + t - \frac{\pi^2 t^2}{18}\right) e^{-T^2 t^2/2}\, dt$$
$$= \sqrt{\frac{\pi}{2}} \cdot \frac{1}{T} - \frac{1}{T^2} + \frac{\pi^{5/2}}{18\sqrt{2}} \cdot \frac{1}{T^3} + \frac{\widetilde{I}_4(T, t_0)}{T^2},$$

where

$$\widetilde{I}_4(T, t_0) = T^2 \int_{t_0}^{\infty} g(t) t e^{-T^2 t^2/2}\, dt, \quad g(t) = \frac{1}{t}\left(\frac{1}{\pi t} - 1 + t - \frac{\pi^2 t^2}{18}\right), \quad t > 0.$$

Since

$$\sup_{t>0} g'(t) = \sup_{t>0}\left(-\frac{2}{\pi t^3}+\frac{1}{t^2}-\frac{\pi^2}{18}\right) = \left(-\frac{2}{\pi t^3}+\frac{1}{t^2}-\frac{\pi^2}{18}\right)\Big|_{t=3/\pi} = -\frac{\pi^2}{54} < 0,$$

the function $g(t)$ decreases monotonically for $t > 0$ and hence,

$$\begin{aligned}\widetilde{I}_4(T,t_0) &\leqslant (g(t_0)\vee 0)T^2\int_{t_0}^{\infty} te^{-T^2t^2/2}\,dt\\ &= \frac{1}{t_0}\left(\frac{1}{\pi t_0}-1+t_0-\frac{\pi^2t_0^2}{18}\right)e^{-T^2t_0^2/2}\vee 0 \equiv J_4(T,t_0).\end{aligned}$$

The function $J_4(T,t_0)$ of $T>0$ is obviously in $\mathcal{D}$ for each fixed $t_0 \in (0,1]$.

Now choose the values of the parameters T and $t_1 \in (0,\,1]$. It is clear that for the efficient estimation of I_2 and I_3 we should use the upper bounds of $|\overline{f}_n(Tt)|$ which are almost everywhere strictly less than one. These upper bounds are given by theorem 2.2, but for their applicability we should assume that $T(\nu_n\ell_n)^{1/\delta} \leqslant 2\pi$. On the other hand, taking into account the term of the form $1/T$ in the estimate for I_4+I_5, we come to the conclusion that T should be taken as large as possible. Therefore finally we set

$$T = 2\pi\,(\nu_n\ell_n)^{-1/\delta}, \quad t_1 = t_1(\delta) = \frac{\theta_0(\delta)}{T(\nu_n\ell_n)^{1/\delta}} = \frac{\theta_0(\delta)}{2\pi}. \tag{4.4}$$

As it follows from the definition, $\theta_0(\delta) \in (\pi,\,2\pi)$, so that $t_1(\delta) \in (1/2,\,1)$ for all $0<\delta\leqslant 1$. Moreover, since $\nu_n \leqslant 2$, the quantities T and ℓ_n are linked by the inequalities

$$T \geqslant 2\pi(2\ell_n)^{-1/\delta}, \quad \ell_n \leqslant \left(\frac{2\pi}{T}\right)^{\delta}.$$

So, for the specified T and t_1 the estimates from theorem 2.2 and lemma 2.7 take the form

$$|\overline{f}_n(Tt)| \leqslant \exp\left\{-\frac{T^2t^2}{2}\Big(1-2\varkappa_\delta(2\pi|t|)^{\delta}\Big)\right\}, \quad t\in\mathbf{R}, \tag{4.5}$$

$$|\overline{f}_n(Tt)| \leqslant \exp\left\{-T^2\frac{1-\cos 2\pi t}{4\pi^2}\right\}, \quad t_1(\delta)\leqslant |t| \leqslant 1, \tag{4.6}$$

$$\begin{aligned} r_n(t) \leqslant \sum_{j=1}^{n}\left|f_j(t/B_n) - e^{-\sigma_j^2t^2/(2B_n^2)}\right| \times\\ \times\exp\left\{-\frac{t^2}{2}\left(1-\frac{\sigma_j^2}{B_n^2}-2\varkappa_\delta\Big(\frac{2\pi|t|}{T}\Big)^{\delta}\right)\right\}, \quad t\in\mathbf{R}. \end{aligned}\tag{4.7}$$

Using the estimate (4.5) in the integral I_2 and the estimate (4.6) in the integral I_3, for any $t_0 \leqslant t_1(\delta)$ we obtain

$$I_2 \leqslant \frac{1.0253}{\pi}\int_{t_0}^{t_1} \exp\left\{-\frac{T^2t^2}{2}\Big(1-2\varkappa_\delta(2\pi t)^\delta\Big)\right\}\,\frac{dt}{t},$$

$$\begin{aligned} I_3 &\leqslant 2\int_{t_1}^{1} |K(t)|\,\exp\left\{-T^2\frac{1-\cos 2\pi t}{4\pi^2}\right\}\,dt \\ &= 2\int_0^{1-t_1} |K(1-t)|\,\exp\left\{-T^2\frac{1-\cos 2\pi t}{4\pi^2}\right\}\,dt \\ &= \int_0^{1-t_1} t\sqrt{1+\Big(\frac{1}{\pi t}-\cot \pi t\Big)^2}\,\exp\left\{-T^2\frac{1-\cos 2\pi t}{4\pi^2}\right\}\,dt. \end{aligned}$$

As is known (see, e. g., [1, 4.3.91]), the cotangent can be expanded into simple fractions as follows:

$$f(x) \equiv \frac{1}{x}-\cot x = 2x\sum_{k=1}^{\infty}\frac{1}{\pi^2k^2-x^2}, \quad x \neq 0, \pm\pi, \pm 2\pi, \ldots,$$

whence it follows that the function $f(x)$ is nonnegative and increases monotonically for all $0 < x < \pi$ and hence, for any $0 < \epsilon < 1$ we have

$$I_3 \leqslant \int_0^{1-t_1} t\sqrt{1+\Big(\frac{1}{\pi(t\vee\epsilon)}-\cot \pi(t\vee\epsilon)\Big)^2}\,\exp\left\{-T^2\frac{1-\cos 2\pi t}{4\pi^2}\right\}\,dt \equiv J_3(T)/T^2.$$

Set

$$g(t) = \frac{1-\cos 2\pi t}{4\pi^2}, \quad dG(t) = t\sqrt{1+\Big(\frac{1}{\pi(t\vee\epsilon)}-\cot \pi(t\vee\epsilon)\Big)^2}\,dt, \quad 0 \leqslant t \leqslant 1-t_1.$$

Obviously, $g(0) = 0$ and $g(t)$ increases monotonically for $0 \leqslant t \leqslant 1/2 > 1-t_1$ (recall that $t_1 > 1/2$). Moreover, it can be made sure that the function $\sin t/t$ decreases for $0 < t \leqslant \pi$ and hence, on the interval $0 \leqslant t \leqslant 1-t_1 \leqslant 1/2$ the function

$$\frac{G'(t)}{g'(t)} = \frac{4\pi^2|K(1-t)|}{(1-\cos 2\pi t)'} = \frac{\pi t}{\sin 2\pi t}\sqrt{1+\Big(\frac{1}{\pi(t\vee\epsilon)}-\cot \pi(t\vee\epsilon)\Big)^2}$$

increases as the product of two monotonically increasing nonnegative functions. So, according to lemma 4.2, $J_3 \in \mathcal{D}$ for any $0 < \epsilon < 1$. Everywhere in what follows we use the value $\epsilon = 10^{-4}$.

Consider the upper bound for I_2 obtained above. It is easy to see that the function $t^2\left(1-2\varkappa_\delta(2\pi t)^\delta\right)$ is positive for $t\in(0,\,t_1]$, since, as it has been already mentioned, $t_1>1/2$, $\varkappa_\delta\leqslant\pi^{-\delta}/2$, and has a unique maximum at the point $t=t_{max}(\delta)=\left((2\pi)^\delta(2+\delta)\varkappa_\delta\right)^{-1/\delta}\in(0,t_1)$, and hence, there exists a unique root

$$t_2=t_2(\delta)\in(0,t_{max}(\delta))$$

of the equation

$$t^2\left(1-2\varkappa_\delta(2\pi t)^\delta\right)=t_1^2\left(1-2\varkappa_\delta(2\pi t_1)^\delta\right),\quad 0<t<t_1(\delta),$$

so that for all $t\in(t_2,t_1)$ we have

$$t^2\left(1-2\varkappa_\delta(2\pi t)^\delta\right)>t_1^2\left(1-2\varkappa_\delta(2\pi t_1)^\delta\right).$$

Splitting the integration domain in the upper bound for I_2 in two parts by the point t_2 we obtain the estimate

$$I_2\leqslant\left(J_{21}(T,t_0)+I_{22}(T,t_0)\right)/T^2,$$

where $J_{21}(T,t_0)=0$, if $t_0\geqslant t_2$, and

$$J_{21}(T,t_0)=\frac{1.0253}{\pi}T^2\int_{t_0}^{t_2}\exp\left\{-\frac{T^2t^2}{2}\left(1-2\varkappa_\delta(2\pi t)^\delta\right)\right\}\frac{dt}{t},\quad \text{if } t_0\leqslant t_2,$$

$$\begin{aligned}I_{22}(T,t_0)&=\frac{1.0253}{\pi}T^2\int_{t_0\vee t_2}^{t_1}\exp\left\{-\frac{T^2t^2}{2}\left(1-2\varkappa_\delta(2\pi t)^\delta\right)\right\}\frac{dt}{t}\\&\leqslant\frac{1.0253}{\pi}T^2\exp\left\{-\frac{T^2t_1^2}{2}\left(1-2\varkappa_\delta(2\pi t_1)^\delta\right)\right\}\int_{t_0\vee t_2}^{t_1}\frac{dt}{t}\\&=\frac{1.0253}{\pi}T^2\exp\left\{-\frac{T^2t_1^2}{2}\left(1-2\varkappa_\delta(2\pi t_1)^\delta\right)\right\}\ln\frac{t_1}{t_0\vee t_2}\equiv J_{22}(T,t_0).\end{aligned}$$

The function $J_{22}(T)$ obviously belongs to the class $\mathcal{D}$.

With a fixed $t_0\leqslant t_2(\delta)$, consider $J_{21}(T,t_0)$ as a function of $T>0$. As was mentioned above, on the interval $[t_0,t_2]$ the function $t^2\left(1-2\varkappa_\delta(2\pi t_1)^\delta\right)$ increases, therefore, according to lemma 4.2, for $J_{21}\in\mathcal{D}$ it suffices that the function

$$\frac{\ln t-\ln t_0}{t^2\left(1-K_\delta t^\delta\right)},\quad K_\delta=2\varkappa_\delta(2\pi)^\delta,$$

increases on $[t_0,\,t_2]$, which is equivalent to the inequality

$$\left(\frac{\ln t-\ln t_0}{t^2(1-K_\delta t^\delta)}\right)'=\frac{t(1-K_\delta t^\delta)-(\ln t-\ln t_0)(2t-(2+\delta)K_\delta t^{1+\delta})}{t^4(1-K_\delta t^\delta)^2}\geqslant 0,$$

$t_0 \leqslant t \leqslant t_2$. The last condition is satisfied, if t_0 satisfies the condition

$$\ln t_0 \geqslant \max_{t\in[t_0,\,t_2]} g(t), \quad g(t) = \ln t - \frac{1-K_\delta t^\delta}{2-(2+\delta)K_\delta t^\delta}.$$

Taking the derivative

$$g'(t) = \frac{(2+\delta)^2K_\delta^2t^{2\delta} - (4+(2+\delta)^2)K_\delta t^\delta + 4}{t(2-(2+\delta)K_\delta t^\delta)^2},$$

we find that $g'(t)$ changes its sign from positive to negative in the point

$$t^* = \left(\frac{4}{(2+\delta)^2K_\delta}\right)^{1/\delta} = \frac{1}{2\pi}\left(\frac{2}{(2+\delta)^2\varkappa_\delta}\right)^{1/\delta},$$

which maximizes the function $g(t)$ and

$$g(t^*) = \ln t^* - \frac{4+\delta}{2(2+\delta)},$$

and hence, for

$$\begin{aligned} t_0 &\geqslant \max_{t\in[t_0,\,t_2]} \exp\{g(t)\} = \exp\{g(t^*)\} \\ &= \frac{1}{2\pi}\left(\frac{2}{(2+\delta)^2\varkappa_\delta}\right)^{1/\delta} \exp\left\{-\frac{4+\delta}{2(2+\delta)}\right\} \equiv t_3(\delta) \end{aligned}$$

we have $J_{21} \in \mathcal{D}$. So,

$$I_2 + I_3 + I_4 + I_5 \leqslant \sqrt{\frac{\pi}{2}}\cdot\frac{1}{T} + \frac{J(T,t_0)}{T^2},$$

where

$$J(T,t_0) = 0 \vee \left(J_{21}(T,t_0) + J_{22}(T,t_0) + J_3(T) + J_4(T,t_0) - 1 + \frac{\pi^{5/2}}{18\sqrt{2}}\cdot\frac{1}{T}\right),$$

with the functions $J_{21}(T,t_0)$, $J_{22}(T,t_0)$, $J_3(T)$, $J_4(T,t_0)$ of $T>0$ belonging to $\mathcal{D}$ for each fixed t_0.

Finally, consider I_1. Estimating $r_n(t)$ by (4.7) with T defined in (4.4) we obtain

$$\begin{aligned} I_1 = \frac{2}{T}\int_0^{t_0T}\left|K\left(\frac{t}{T}\right)\right| r_n(t)dt \leqslant \frac{2}{T}\sum_{j=1}^n \int_0^{t_0T}\left|K\left(\frac{t}{T}\right)\right|\cdot\left|f_j(t/B_n) - e^{-\sigma_j^2t^2/(2B_n^2)}\right| \times \\ \times \exp\left\{-\frac{t^2}{2}\left(1 - \frac{\sigma_j^2}{B_n^2} - 2\varkappa_\delta\Big(\frac{2\pi t}{T}\Big)^\delta\right)\right\} dt \end{aligned}$$

$$
-\frac{2}{T}\sum_{j=1}^{n}\int\limits_{0}^{t_0T}\left|K\left(\frac{t}{T}\right)\right|\cdot\left|f_j(t/B_n)-e^{-\sigma_j^2t^2/(2B_n^2)}\right|e^{-t^2/2}\,dt
$$

$$
+\frac{2}{T}\sum_{j=1}^{n}\int\limits_{0}^{t_0T}\left|K\left(\frac{t}{T}\right)-\frac{iT}{2\pi t}\right|\cdot\left|f_j(t/B_n)-e^{-\sigma_j^2t^2/(2B_n^2)}\right|e^{-t^2/2}\,dt
$$

$$
+\frac{2}{T}\sum_{j=1}^{n}\int\limits_{0}^{t_0T}\left|\frac{iT}{2\pi t}\right|\cdot\left|f_j(t/B_n)-e^{-\sigma_j^2t^2/(2B_n^2)}\right|e^{-t^2/2}\,dt\leqslant I_{11}+I_{12}+I_{13},
$$

where

$$
I_{11}=\frac{2}{T}\sum_{j=1}^{n}\int\limits_{0}^{t_0T}\left|K\left(\frac{t}{T}\right)\right|\cdot\left|f_j(t/B_n)-e^{-\sigma_j^2t^2/(2B_n^2)}\right|e^{-t^2/2}\times
$$

$$
\times\left(\exp\left\{\frac{\sigma_j^2t^2}{2B_n^2}+\varkappa_\delta\left(2\pi/T\right)^\delta t^{2+\delta}\right\}-1\right)\,dt,
$$

$$
I_{12}=\frac{2}{T}\sum_{j=1}^{n}\int\limits_{0}^{t_0T}\left|K\left(\frac{t}{T}\right)-\frac{iT}{2\pi t}\right|\cdot\left|f_j(t/B_n)-e^{-\sigma_j^2t^2/(2B_n^2)}\right|e^{-t^2/2}\,dt,
$$

$$
I_{13}=\frac{1}{\pi}\sum_{j=1}^{n}\int\limits_{0}^{\infty}\frac{1}{t}\left|f_j(t/B_n)-e^{-\sigma_j^2t^2/(2B_n^2)}\right|e^{-t^2/2}\,dt.
$$

Note that for all $j=1,\dots,n$ and $t\leqslant t_0T$ with $T=2\pi\left(\nu_n\ell_n\right)^{-1/\delta}$

$$
\exp\left\{\frac{\sigma_j^2t^2}{2B_n^2}+\varkappa_\delta\left(2\pi/T\right)^\delta t^{2+\delta}\right\}-1
$$

$$
\leqslant\left(\frac{\sigma_j^2t^2}{2B_n^2}+\frac{\varkappa_\delta(2\pi)^\delta t^{2+\delta}}{T^\delta}\right)\exp\left\{\frac{\sigma_j^2t^2}{2B_n^2}+\frac{\varkappa_\delta(2\pi)^\delta t^{2+\delta}}{T^\delta}\right\}
$$

$$
\leqslant\left(\frac{\sigma_j^2t^2}{2B_n^2}+\varkappa_\delta\nu_n\ell_nt^{2+\delta}\right)\exp\left\{\frac{t^2}{2}\left(\frac{\sigma_j^2}{B_n^2}+2\varkappa_\delta(2\pi t_0)^\delta\right)\right\},
$$

Taking into account the estimates for $K(t)$ given by lemma 4.1 and the estimates (2.4), (2.6) for the modulus of the difference of the ch.f.'s from lemma 2.8 for the integral I_{11} we obtain

$$
I_{11}\leqslant\frac{1.0253}{\pi}\sum_{j=1}^{n}\int\limits_{0}^{t_0T}\left(\frac{\gamma_\delta\beta_{2+\delta,j}\,t^{2+\delta}}{B_n^{2+\delta}}+\frac{\sigma_j^4t^4}{8B_n^4}\,\mathbf{1}(\delta<1)\right)\left(\frac{\sigma_j^2t^2}{2B_n^2}+\varkappa_\delta\nu_n\ell_nt^{2+\delta}\right)\times
$$

$$
\times\exp\left\{-\frac{t^2}{2}\left(1-\frac{\sigma_j^2}{B_n^2}-2\varkappa_\delta(2\pi t_0)^\delta\right)\right\}dt
$$

$$= \frac{1.0253}{\pi}\sum_{j=1}^{n}\int\limits_0^\infty \Bigg[\gamma_\delta t^{4+\delta}\frac{\beta_{2+\delta,j}\sigma_j^2}{2B_n^{4+\delta}} + \gamma_\delta\varkappa_\delta\nu_n t^{4+2\delta}\ell_n\frac{\beta_{2+\delta,j}}{B_n^{2+\delta}}$$
$$+ \left(\frac{t^6\sigma_j^6}{16B_n^6} + \varkappa_\delta\nu_n t^{6+\delta}\ell_n\frac{\sigma_j^4}{8B_n^4}\right)\mathbf{1}(\delta<1)\Bigg]\times$$
$$\times \exp\left\{-\frac{t^2}{2}\left(1-\frac{\sigma_j^2}{B_n^2}-2\varkappa_\delta(2\pi t_0)^\delta\right)\right\}dt.$$

Estimate the exponent uniformly with respect to $j=1,\ldots,n$ by the inequality

$$\max_{1\leqslant j\leqslant n}\sigma_j^2 \leqslant \Big(\max_{1\leqslant j\leqslant n}\sigma_j\Big)^2 \leqslant \Big(\max_{1\leqslant j\leqslant n}\beta_{2+\delta,j}\Big)^{2/(2+\delta)}$$
$$\leqslant \Big(\sum_{j=1}^{n}\beta_{2+\delta,j}\Big)^{2/(2+\delta)} = B_n^2\ell_n^{2/(2+\delta)}.$$

Estimate the power-type multiplier by the Lyapounov inequality and relation (4.2) to obtain

$$I_{11} \leqslant \frac{1.0253}{16\pi}\int\limits_0^\infty\Bigg[(8\gamma_\delta\ell_n^{(4+\delta)/(2+\delta)}t^{4+\delta} + 16\gamma_\delta\varkappa_\delta\nu_n\ell_n^2 t^{4+2\delta}$$
$$+\left(\ell_n^{6/(2+\delta)}t^6 + 2\varkappa_\delta\nu_n\ell_n^{(6+\delta)/(2+\delta)}t^{6+\delta}\right)\mathbf{1}(\delta<1)\Bigg]\times$$
$$\times\exp\left\{-\frac{t^2}{2}\Big(1-\ell_n^{2/(2+\delta)}-2\varkappa_\delta(2\pi t_0)^\delta\Big)\right\}dt.$$

For the case of identically distributed summands, since $\sigma_j^2 = B_n^2/n \leqslant B_n^2\ell_n n^{-1+\delta/2}$ for all $j=1,\ldots,n$, we obtain the estimate

$$I_{11} \leqslant \frac{1.0253}{16\pi}\,\ell_n^2\int\limits_0^\infty\left(8\gamma_\delta t^{2+\delta}+\frac{t^4}{n^{1-\delta/2}}\,\mathbf{1}(\delta<1)\right)\left(\frac{t^2}{n^{1-\delta/2}}+2\varkappa_\delta\nu_n t^{2+\delta}\right)\times$$
$$\times\exp\left\{-\frac{t^2}{2}\Big(1-n^{-1}-2\varkappa_\delta(2\pi t_0)^\delta\Big)\right\}dt.$$

Assume that

$$t_0 < \frac{1}{2\pi}\left(\frac{1-\ell_n^{2/(2+\delta)}}{2\varkappa_\delta}\right)^{1/\delta} \equiv t_4(\delta,\ell_n).$$

The domain $t_3(\delta)\leqslant t_0 < t_4(\delta,\ell_n)$ is non-empty, if

$$\ell_n^{2/(2+\delta)} < 1-\frac{4}{(2+\delta)^2}\exp\left\{-\frac{\delta(4+\delta)}{2(2+\delta)}\right\} > 0,\quad 0<\delta\leqslant 1.$$

For t_0 and ℓ_n specified above introduce the function

$$Q(\ell_n, t_0, r) = \int_0^\infty t^r \exp\left\{-\frac{t^2}{2}\Big(1 - \ell_n^{2/(2+\delta)} - 2\varkappa_\delta (2\pi t_0)^\delta\Big)\right\}$$
$$= \frac{2^{(r-1)/2}\Gamma\left(\frac{r+1}{2}\right)}{\left(1 - \ell_n^{2/(2+\delta)} - 2\varkappa_\delta(2\pi t_0)^\delta\right)^{(r+1)/2}}, \quad r > 0.$$

It is obvious that $Q(\ell_n, t_0, r)$ increases monotonically in ℓ_n with fixed t_0 and r. So, for I_{11} for all $t_0 \leqslant t_4(\delta, \ell_n)$ we obtain

$$I_{11} \leqslant \frac{1.0253}{16\pi}\, \ell_n^{(4+\delta)/(2+\delta)}\Big(8\gamma_\delta Q\left(\ell_n, t_0, 4+\delta\right)$$
$$+ 16\nu_n \varkappa_\delta \gamma_\delta \ell_n^{\delta/(2+\delta)} Q\left(\ell_n, t_0, 4+2\delta\right) + \ell_n^{(2-\delta)/(2+\delta)} Q\left(\ell_n, t_0, 6\right)\mathbf{1}(\delta < 1)$$
$$+ 2\nu_n\varkappa_\delta \ell_n^{2/(2+\delta)} Q\left(\ell_n, t_0, 6+\delta\right)\mathbf{1}(\delta<1)\Big) \equiv \ell_n^{(4+\delta)/(2+\delta)} J_{11}(\ell_n, \nu_n, t_0)$$

in the general case, whereas for

$$\frac{1}{n} < 1 - \frac{4\exp\left\{-\frac{\delta(4+\delta)}{2(2+\delta)}\right\}}{(2+\delta)^2} \equiv (\bar{\ell}(\delta))^{2/(2+\delta)}, \ t_0 < \frac{1}{2\pi}\left(\frac{1-n^{-1}}{2\varkappa_\delta}\right)^{1/\delta} \equiv t_4\Big(\delta, n^{-1-\delta/2}\Big),$$

for identically distributed summands

$$I_{11} \leqslant \frac{1.0253}{16\pi}\, \ell_n^2\left[\frac{8\gamma_\delta}{n^{1-\delta/2}} Q\Big(\frac{1}{n^{1+\delta/2}}, t_0, 4+\delta\Big) + 16\nu_n\varkappa_\delta\gamma_\delta Q\Big(\frac{1}{n^{1+\delta/2}}, t_0, 4+2\delta\Big) + \right.$$
$$\left. + \Big(n^{-2+\delta} Q\Big(\frac{1}{n^{1+\delta/2}}, t_0, 6\Big) + \frac{2\nu_n\varkappa_\delta}{n^{1-\delta/2}}\, Q\Big(\frac{1}{n^{1+\delta/2}}, t_0, 6+\delta\Big)\Big)\mathbf{1}(\delta<1)\right] \equiv \ell_n^2 \widehat{J}_{11}(n, t_0).$$

Similarly, for I_{12} with the account of the definition of $T = 2\pi(\nu_n\ell_n)^{-1/\delta}$ we obtain

$$I_{12} \leqslant \frac{1}{T}\sum_{j=1}^n \int_0^{t_0 T} \left(1 - \frac{t}{T} + \frac{\pi^2 t^2}{18T^2}\right)\left(\frac{\gamma_\delta \beta_{2+\delta,j}\, t^{2+\delta}}{B_n^{2+\delta}} + \frac{\sigma_j^4 t^4}{8B_n^4}\,\mathbf{1}(\delta<1)\right) e^{-t^2/2}\, dt$$
$$\leqslant \frac{1}{T}\int_0^\infty \left(1 + \frac{\pi^2 t^2}{18T^2}\right)\left(\gamma_\delta \ell_n t^{2+\delta} + \frac{t^4}{8B_n^4}\sum_{j=1}^n \sigma_j^4 \mathbf{1}(\delta<1)\right) e^{-t^2/2}\, dt$$
$$= \frac{2^{(\delta-1)/2}\gamma_\delta}{\pi}\, \nu_n^{1/\delta} \ell_n^{(1+\delta)/\delta}\Gamma\Big(\frac{3+\delta}{2}\Big)\left(1 + \frac{3+\delta}{72}(\nu_n\ell_n)^{2/\delta}\right)$$
$$+ \frac{3(\nu_n\ell_n)^{1/\delta}}{16\sqrt{2\pi}B_n^4}\sum_{j=1}^n \sigma_j^4\left(1 + \frac{5}{72}(\nu_n\ell_n)^{2/\delta}\right)\mathbf{1}(\delta<1),$$

whence by the Lyapounov inequality and (4.2) it follows that in the general case

$$I_{12} \leqslant \ell_n^{(1+\delta)/\delta}\nu_n^{1/\delta}\left[\frac{2^{(\delta-1)/2}\gamma_\delta}{\pi}\,\Gamma\Big(\frac{3+\delta}{2}\Big)\left(1 + \frac{3+\delta}{72}(\nu_n\ell_n)^{2/\delta}\right)\right.$$

$$+\frac{3\ell_n^{(2-\delta)/(2+\delta)}}{16\sqrt{2\pi}}\left(1+\frac{5}{72}(\nu_n\ell_n)^{2/\delta}\right)\mathbf{1}(\delta<1)\Big]\equiv \ell_n^{(1+\delta)/\delta}J_{12}(\ell_n,\nu_n),$$

and in the case of identically distributed summands

$$I_{12}\leqslant \ell_n^{(1+\delta)/\delta}\nu_n^{1/\delta}\left[\frac{2^{(\delta-1)/2}\gamma_\delta}{\pi}\,\Gamma\Big(\frac{3+\delta}{2}\Big)\left(1+\frac{3+\delta}{72}(\nu_n\ell_n)^{2/\delta}\right)\right.$$
$$\left.+\frac{3n^{-1+\delta/2}}{16\sqrt{2\pi}}\left(1+\frac{5}{72}(\nu_n\ell_n)^{2/\delta}\right)\mathbf{1}(\delta<1)\right]\equiv \ell_n^{(1+\delta)/\delta}\widehat{J}_{12}(\ell_n,\nu_n,n).$$

Summarize the above reasoning as a lemma.

Lemma 4.3. *For $0<\delta\leqslant 1$ by $\theta_0(\delta)$ denote the unique root of the equation*

$$\delta\theta^2+2\theta\sin\theta+2(2+\delta)(\cos\theta-1)=0,\quad \pi<\theta<2\pi,$$

$$\varkappa_\delta=\sup_{x>0}\frac{\big|\cos x-1+x^2/2\big|}{x^{2+\delta}}=\frac{\cos\theta_0(\delta)-1+\theta_0^2(\delta)/2}{\theta_0^{2+\delta}(\delta)}=\frac{\theta_0(\delta)-\sin\theta_0(\delta)}{(2+\delta)\theta_0^{1+\delta}(\delta)},$$

$$\gamma_\delta=\sup_{x>0}\sqrt{\Big(\frac{\cos x-1+x^2/2}{x^{2+\delta}}\Big)^2+\Big(\frac{\sin x-x}{x^{2+\delta}}\Big)^2},$$

$t_1(\delta)=\theta_0(\delta)/(2\pi)$, let $t_2=t_2(\delta)$ be the unique root of the equation

$$t^2\left(1-2\varkappa_\delta(2\pi t)^\delta\right)=t_1^2(\delta)\left(1-2\varkappa_\delta(2\pi t_1(\delta))^\delta\right)$$

on the interval $(0,t_1(\delta))$. Let

$$\overline{\ell}(\delta)=\left(1-\frac{4}{(2+\delta)^2}\,\exp\left\{-\frac{\delta(4+\delta)}{2(2+\delta)}\right\}\right)^{1+\delta/2},$$

$$t_4(\delta,\ell)=\frac{1}{2\pi}\left(\frac{1-\ell^{2/(2+\delta)}}{2\varkappa_\delta}\right)^{1/\delta},\quad 0<\ell<\overline{\ell}(\delta).$$

$$t_3(\delta)=\frac{1}{2\pi}\left(\frac{2}{(2+\delta)^2\varkappa_\delta}\right)^{1/\delta}\exp\left\{-\frac{4+\delta}{2(2+\delta)}\right\}=t_4\left(\delta,\overline{\ell}(\delta)\right).$$

Then for all $n\geqslant 1$ and $F_1,\dots,F_n\in\mathcal{F}_{2+\delta}$ such that $\ell_n<\overline{\ell}(\delta)$ and for any t_0 from the interval

$$t_3(\delta)\leqslant t_0<\min\{t_1(\delta),t_4(\delta,\ell_n)\}$$

there holds the estimate

$$\Delta_n\leqslant\frac{1}{\pi}\sum_{j=1}^n\int_0^\infty\frac{1}{t}\left|f_j(t/B_n)-e^{-\sigma_j^2t^2/(2B_n^2)}\right|e^{-t^2/2}\,dt+\frac{(\nu_n\ell_n)^{1/\delta}}{2\sqrt{2\pi}}+$$
$$+\ell_n^{(4+\delta)/(2+\delta)}J_{11}(\ell_n,\nu_n,t_0)+\ell_n^{(1+\delta)/\delta}J_{12}(\ell_n,\nu_n)+\frac{(\nu_n\ell_n)^{2/\delta}}{4\pi^2}\,J\left(\frac{2\pi}{(\nu_n\ell_n)^{1/\delta}},t_0\right),$$

where

$$J_{11}(\ell,\nu,t_0)=\frac{1.0253}{16\pi}\Big(8\gamma_\delta Q\left(\ell,t_0,4+\delta\right)+16\nu\varkappa_\delta\gamma_\delta\ell^{\delta/(2+\delta)}Q\left(\ell,t_0,4+2\delta\right)$$
$$+\Big(\ell^{(2-\delta)/(2+\delta)}Q\left(\ell,t_0,6\right)+2\nu\varkappa_\delta\ell^{2/(2+\delta)}Q\left(\ell,t_0,6+\delta\right)\Big)\mathbf{1}(\delta<1)\Big),$$
$$1\leqslant\nu\leqslant 2,\ \ell>0,$$

$$Q(\ell,t_0,r)=\frac{2^{(r-1)/2}\Gamma\left(\frac{r+1}{2}\right)}{\left(1-\ell^{2/(2+\delta)}-2\varkappa_\delta(2\pi t_0)^\delta\right)^{(r+1)/2}},\quad 0<\ell<\overline{\ell}(\delta),\ r>0,$$

$$J_{12}(\ell,\nu)=\nu^{1/\delta}\left[\frac{2^{(\delta-1)/2}\gamma_\delta}{\pi}\Gamma\Big(\frac{3+\delta}{2}\Big)\left(1+\frac{3+\delta}{72}(\nu\ell)^{2/\delta}\right)\right.$$
$$\left.+\frac{3\ell^{(2-\delta)/(2+\delta)}}{16\sqrt{2\pi}}\left(1+\frac{5}{72}(\nu\ell)^{2/\delta}\right)\mathbf{1}(\delta<1)\right],\quad 1\leqslant\nu\leqslant 2,\ \ell>0,$$

$$J(T,t_0)=0\vee\left(J_{21}(T,t_0)+J_{22}(T,t_0)+J_3(T)+J_4(T,t_0)-1+\frac{\pi^{5/2}}{18\sqrt{2}}\cdot\frac{1}{T}\right),$$

$$J_{21}(T,t_0)=\frac{1.0253}{\pi}T^2\int\limits_{t_0\wedge t_2(\delta)}^{t_2(\delta)}\exp\left\{-\frac{T^2t^2}{2}\left(1-2\varkappa_\delta(2\pi t)^\delta\right)\right\}\frac{dt}{t},$$

$$J_{22}(T,t_0)=\frac{1.0253}{\pi}T^2\exp\left\{-\frac{T^2t_1^2(\delta)}{2}\left(1-2\varkappa_\delta(2\pi t_1(\delta))^\delta\right)\right\}\ln\frac{t_1(\delta)}{t_0\vee t_2(\delta)},$$

$$J_3(T)=T^2\int\limits_0^{1-t_1(\delta)}t\sqrt{1+\Big(\frac{1}{\pi(t\vee 10^{-4})}-\cot\pi(t\vee 10^{-4})\Big)^2}\times$$
$$\times\exp\left\{-T^2\frac{1-\cos 2\pi t}{4\pi^2}\right\}dt,$$

$$J_4(T,t_0)=0\vee\frac{1}{t_0}\left(\frac{1}{\pi t_0}-1+t_0-\frac{\pi^2t_0^2}{18}\right)e^{-T^2t_0^2/2},\quad T>0.$$

If $F_1=\ldots=F_n\in\mathcal{F}_{2+\delta}$*, then for all* $n\geqslant(\overline{\ell}(\delta))^{-2/(2+\delta)}$ *and* t_0 *such that*

$$t_3(\delta)\leqslant t_0<\min\left\{t_1(\delta),\ t_4\left(\delta,n^{-1-\delta/2}\right)\right\},$$

there holds the estimate

$$\Delta_n\leqslant\frac{1}{\pi}\sum_{j=1}^n\int\limits_0^\infty\frac{1}{t}\left|f_j(t/B_n)-e^{-\sigma_j^2t^2/(2B_n^2)}\right|e^{-t^2/2}\,dt+\frac{1}{2\sqrt{2\pi n}}\left(\frac{\beta_{2+\delta}}{\sigma^{2+\delta}}+\frac{\beta_\delta}{\sigma^\delta}\right)^{1/\delta}$$
$$+\ell_n^2\widehat{J}_{11}(n,\nu_n,t_0)+\ell_n^{(1+\delta)/\delta}\widehat{J}_{12}(\ell_n,\nu_n,n)+\frac{(\nu_n\ell_n)^{2/\delta}}{4\pi^2}J\left(\frac{2\pi}{(\nu_n\ell_n)^{1/\delta}},t_0\right),$$

where

$$\widehat{J}_{11}(n,\nu,t_0)=\frac{1.0253}{16\pi}\left[\frac{8\gamma_\delta}{n^{1-\delta/2}}Q\Big(\frac{1}{n^{1+\delta/2}},t_0,4+\delta\Big)\right.$$

$$
\begin{aligned}
&+16\nu\varkappa_\delta\gamma_\delta Q\Big(\frac{1}{n^{1+\delta/2}},t_0,4+2\delta\Big)+\Big(n^{-2+\delta}Q\Big(\frac{1}{n^{1+\delta/2}},t_0,6\Big)\\
&+\frac{2\nu\varkappa_\delta}{n^{1-\delta/2}}\,Q\Big(\frac{1}{n^{1+\delta/2}},t_0,6+\delta\Big)\Big)\mathbf{1}(\delta<1)\Big],\\
\widehat{J}_{12}(\ell,\nu,n)&=\nu^{1/\delta}\Big[\frac{2^{(\delta-1)/2}\gamma_\delta}{\pi}\,\Gamma\Big(\frac{3+\delta}{2}\Big)\Big(1+\frac{3+\delta}{72}(\nu\ell)^{2/\delta}\Big)\\
&+\frac{3n^{-1+\delta/2}}{16\sqrt{2\pi}}\Big(1+\frac{5}{72}(\nu\ell)^{2/\delta}\Big)\mathbf{1}(\delta<1)\Big],\ 1\leqslant\nu\leqslant 2,\ \ell>0,\ n\geqslant 1.
\end{aligned}
$$

With t_0 fixed, the functions $J_{21}(T,t_0)$, $J_{22}(T)$, $J_3(T)$, $J_4(T,t_0)$ of T for $T>0$ have at most one maximum and have no minima; $\widehat{J}_{11}(n,\nu,t_0)$, $\widehat{J}_{12}(\ell,\nu,n)$ decrease monotonically in $n\geqslant 1$ with ℓ and ν fixed; $t_4(\delta,\ell)$ decreases monotonically in ℓ; $J_{11}(\ell,\nu,t_0)$, $J_{12}(\ell,\nu)$, $\widehat{J}_{11}(\ell^{-2/\delta},\nu,t_0)$, $\widehat{J}_{12}(\ell,\nu,\ell^{-2/\delta})$ increase monotonically in ℓ; $J_{11}(\ell,\nu,t_0)$, $J_{12}(\ell,\nu)$, $\widehat{J}_{11}(n,\nu,t_0)$, $\widehat{J}_{12}(\ell,\nu,n)$ increase monotonically in $\nu\in[1,\,2]$, and

$$
\lim_{n\to\infty}\widehat{J}_{11}(n,\nu,t_0)=\frac{1.0253\cdot 2^{3/2+\delta}\nu\varkappa_\delta\gamma_\delta\Gamma(5/2+\delta)}{\pi(1-2\varkappa_\delta(2\pi t_0)^\delta)^{5/2+\delta}},\ 1\leqslant\nu\leqslant 2,\ t_3(\delta)\leqslant t_0\leqslant t_1(\delta),
$$

$$
\lim_{\ell\to 0}\sup_{n\geqslant\ell^{-2/\delta}}\widehat{J}_{12}(\ell,\nu,n)=\nu^{1/\delta}2^{(\delta-1)/2}\pi^{-1}\gamma_\delta\Gamma((3+\delta)/2),\quad 1\leqslant\nu\leqslant 2,
$$

$$
\lim_{T\to\infty}J(T,t_0)=0,\quad t_3(\delta)\leqslant t_0\leqslant t_1(\delta).
$$

The values of γ_δ, $\varkappa_\delta$, $t_1(\delta)$, $t_2(\delta)$, $t_3(\delta)$, $t_4(\delta,\ell)$, $\overline{\ell}(\delta)$ and

$$
N(\delta)=\inf\Big\{n\in\mathbf{N}\colon n>(\overline{\ell}(\delta))^{-2/(2+\delta)}\Big\}=1+\Big\lfloor(\overline{\ell}(\delta))^{-2/(2+\delta)}\Big\rfloor
$$

for some $0<\delta\leqslant 1$ and $\ell=0.1,\,0.01$ calculated with the accuracy to the fourth decimal digit are given in table 3.

Remark 4.4. On the right-hand sides of the inequalities in lemma 4.3 the "leading" terms are two first summands: the integral

$$
I_{13}=\frac{1}{\pi}\sum_{j=1}^{n}\int_0^\infty\frac{1}{t}\Big|f_j(t/B_n)-e^{-\sigma_j^2t^2/(2B_n^2)}\Big|\,e^{-t^2/2}\,dt
$$

and

$$
\frac{(\nu_n\ell_n)^{1/\delta}}{2\sqrt{2\pi}}=\sqrt{\frac{\pi}{2}}\cdot\frac{1}{T},
$$

appearing when the sum of the integrals I_4 and I_5 is estimated. It is interesting to clarify the nature of these summands and their contribution into the constants at the leading terms in the resulting estimates. For simplicity consider the case of identically distributed summands. As we will see below, the integral I_{13} contains the information concerning the "heavy-tailedness" of the distribution: the order of

δ	γ_δ	$\varkappa_\delta$	$t_1(\delta)$	$t_2(\delta)$	$t_3(\delta)$	$t_4(\delta, 0.1)$	$t_4(\delta, 0.01)$	$\overline{\ell}(\delta)$	$N(\delta)$
0.01	0.5225	0.4909	0.9950	0.0261	0.1356	0.0000	0.3566	0.0193	51
0.05	0.4885	0.4563	0.9761	0.0673	0.1370	0.1055	0.7887	0.0886	11
0.10	0.4498	0.4170	0.9539	0.1019	0.1386	0.2990	0.8613	0.1626	6
0.15	0.4149	0.3815	0.9331	0.1302	0.1401	0.4197	0.8798	0.2265	4
0.20	0.3833	0.3494	0.9132	0.1551	0.1416	0.4944	0.8841	0.2827	4
0.25	0.3548	0.3203	0.8941	0.1778	0.1431	0.5429	0.8826	0.3327	3
0.30	0.3290	0.2940	0.8756	0.1989	0.1444	0.5758	0.8784	0.3776	3
0.35	0.3058	0.2701	0.8576	0.2187	0.1457	0.5987	0.8725	0.4181	3
0.40	0.2847	0.2484	0.8399	0.2375	0.1469	0.6147	0.8658	0.4549	2
0.45	0.2657	0.2287	0.8226	0.2556	0.1480	0.6260	0.8584	0.4884	2
0.50	0.2486	0.2108	0.8054	0.2729	0.1490	0.6338	0.8507	0.5191	2
0.55	0.2331	0.1945	0.7884	0.2896	0.1500	0.6390	0.8427	0.5474	2
0.60	0.2193	0.1796	0.7716	0.3058	0.1509	0.6422	0.8345	0.5734	2
0.65	0.2070	0.1661	0.7548	0.3214	0.1517	0.6439	0.8262	0.5975	2
0.70	0.1960	0.1537	0.7380	0.3366	0.1524	0.6442	0.8177	0.6198	2
0.75	0.1865	0.1424	0.7212	0.3514	0.1530	0.6435	0.8091	0.6405	2
0.80	0.1783	0.1321	0.7044	0.3657	0.1536	0.6420	0.8005	0.6597	2
0.85	0.1715	0.1227	0.6875	0.3797	0.1540	0.6397	0.7918	0.6776	2
0.90	0.1665	0.1142	0.6705	0.3932	0.1544	0.6369	0.7830	0.6944	2
0.95	0.1637	0.1063	0.6533	0.4064	0.1547	0.6334	0.7741	0.7100	2
1.00	0.1666	0.0991	0.6359	0.4191	0.1550	0.6296	0.7652	0.7247	2

Table 3: The values of γ_δ, $\varkappa_\delta$, $t_1(\delta)$, $t_2(\delta)$, $t_3(\delta)$, $t_4(\delta,\ell)$, $\overline{\ell}(\delta)$ and $N(\delta) = 1 + \left\lfloor (\overline{\ell}(\delta))^{-2/(2+\delta)} \right\rfloor$ for some $0 < \delta \leqslant 1$ and $\ell = 0.1$, 0.01.

its decrease is completely determined by the maximum order of the finite moment of a summand (in our case $I_{13} = O(n^{-\delta/2})$) whereas the role of the corresponding characteristic of the distribution is played by the normalized moment of the maximum order $\beta_{2+\delta}/\sigma^{2+\delta}$. In other words, there exists such an *absolute* positive finite constant C that

$$I_{13} \leqslant C \cdot \frac{\beta_{2+\delta}}{\sigma^{2+\delta} n^{\delta/2}},$$

moreover, as is illustrated by the corresponding examples in [29], the order of this estimate is exact, if it is meant uniformly in $F \in \mathcal{F}_{2+\delta}$. The importance of the remark concerning the exactness of the order is conditioned by the fact that $\Delta_n(F) = o(n^{-\delta/2})$ for any fixed $F \in \mathcal{F}_{2+\delta}$ (see also [22]). But, on the other hand, if a distribution $F \in \mathcal{F}_{2+\delta}$ depends on n and the moment-type characteristic $\beta_{2+\delta}/\sigma^{2+\delta}$ is included in the estimate, then $\beta_{2+\delta}/\sigma^{2+\delta}n^{-\delta/2}$ is an exact characteristic of the rate of convergence.

Now consider the second term $\sqrt{\pi/2}/T$. Here the coefficient $\sqrt{\pi/2}$ is determined by the limit distribution which is normal in the case under consideration. The value of T chosen in the process of estimation of the integral I_3 is determined by the maximum length of a zero-left-ended interval on which it is possible to bound the absolute value of the ch.f. by a number less than one (see remark 2.5). So, the term under consideration contains the information concerning the smoothness of the pre-limit distribution. Moreover, since the sum of random variables is normalized by $\sqrt{n}$, the length of the interval on which the absolute value of the ch.f. is bounded by a number less than one is proportional to $\sqrt{n}$, that is, for $\delta < 1$ the effects due to the smoothness or discreteness of the original distribution disappear making no influence on the constant at the leading term of the estimate having the order $n^{-\delta/2}$. At the same time,for $\delta = 1$ the order of normalization of the sum of r.v.'s coincides with the order of the maximum length of the interval on which the absolute value of the ch.f. is bounded by a number less than one, therefore, the effects of "heavy-tailedness" revealing themselves in the integral I_{13} are added with the effects of "non-smoothness" which leads to abrupt increase (discontinuity) of the constant at the leading term of order $1/\sqrt{n}$ in the point $\delta = 1$.

Remark 4.5. Let $\nu \in [1, 2]$ and $\ell > 0$ be arbitrary numbers. For the purpose of construction of estimates of the function $J\left(2\pi(\nu_n\ell_n)^{-1/\delta}, t_0\right)$ with fixed t_0 uniform in $\ell_n \leqslant \ell$ and $\nu_n \in [1, \nu]$ consider the behavior of the functions $J_{21}(T, t_0)$, $J_{22}(T, t_0)$, $J_3(T)$, $J_4(T, t_0)$ of $T = 2\pi(\nu_n\ell_n)^{-1/\delta} \geqslant 2\pi(\nu\ell)^{-1/\delta} > 0$, which are components of $J(T, t_0)$. Obviously, the function $J_4(T, t_0)$ decreases monotonically in $T > 0$. Noticing that the function xe^{-ax} decreases monotonically for $x > 1/a > 0$ we conclude that $J_{22}(T, t_0)$ decreases monotonically for

$$T \geqslant \frac{\sqrt{2}}{t_1(\delta)\sqrt{1 - 2\varkappa_\delta(2\pi t_1(\delta))^\delta}} \equiv T_{22}(\delta).$$

If $t_3(\delta) \geqslant t_2(\delta)$, then $J_{21}(T, t_0) = 0$ for all $t_0 \geqslant t_3(\delta)$. And if $t_3(\delta) < t_2(\delta)$, then using the property of monotonic increase of the function $t^2(1 - 2\varkappa_\delta(2\pi t)^\delta)$ for $t \in (0, t_2(\delta))$ established in the proof of lemma 4.3 we similarly conclude that $J_{21}(T, t_0)$ decreases monotonically for

$$T \geqslant \frac{\sqrt{2}}{t_3(\delta)\sqrt{1 - 2\varkappa_\delta(2\pi t_3(\delta))^\delta}} \equiv T_{21}(\delta)$$

for each fixed $t_0 \geqslant t_3(\delta)$. Finally, for each fixed δ it is possible to find numerically the unique point $T_3(\delta)$ of the maximum of the function $J_3(T) \in \mathcal{D}$ such that $J_3(T)$ decreases monotonically for $T \geqslant T_3(\delta)$. So, if the numbers $\nu \in [1, 2]$ and $\ell > 0$ satisfy the inequality

$$\nu\ell \leqslant \left(\frac{2\pi}{\max\{T_{21}(\delta), T_{22}(\delta), T_3(\delta)\}}\right)^\delta \equiv \overline{\varepsilon}(\delta),$$

then

$$\max_{\ell_n \leqslant \ell,\, \nu_n \in [1,\nu]} J\left(\frac{2\pi}{(\nu_n\ell_n)^{1/\delta}}, t_0\right) \leqslant J\left(\frac{2\pi}{(\nu\ell)^{1/\delta}}, t_0\right).$$

The values of $T_{21}(\delta)$, $T_{22}(\delta)$, $T_3(\delta)$ and $\overline{\varepsilon}(\delta)$ are given in table 4. From this table it can be seen, in particular, that for $\ell_n \leqslant 0.3$ the monotonicity takes place for all $1 \leqslant \nu_n \leqslant 2$ and $0.01 \leqslant \delta \leqslant 1$ given in table 4.

Depending on whether $\delta = 1$ or not, to estimate the integral

$$I_{13} = \frac{1}{\pi}\sum_{j=1}^{n}\int_0^\infty \frac{1}{t}\left|f_j(t/B_n) - e^{-\sigma_j^2 t^2/(2B_n^2)}\right| e^{-t^2/2}\,dt$$

δ	$T_{21}(\delta) \leqslant$	$T_{22}(\delta) \leqslant$	$T_3(\delta) \leqslant$	$\overline{\varepsilon}(\delta) \geqslant$
0.01	74.1670	285.6369	1065.6543	0.9498
0.05	33.6579	59.2429	188.6696	0.8434
0.10	24.2258	30.8361	89.8283	0.7663
0.15	20.1242	21.3082	58.3999	0.7156
0.20	17.7237	16.5114	43.1128	0.6802
0.25	16.1158	13.6136	34.1103	0.6550
0.30	14.9517	11.6694	28.1896	0.6373
0.35	14.0650	10.2731	24.0043	0.6254
0.40	13.3653	9.2211	20.8912	0.6183
0.45	12.7987	8.4003	18.4862	0.6152
0.50	12.3308	7.7426	16.5734	0.6156
0.55	11.9386	7.2046	15.0164	0.6191
0.60	11.6060	6.7573	13.7250	0.6256
0.65	11.3215	6.3806	12.6372	0.6348
0.70	11.0764	6.0602	11.7091	0.6466
0.75	10.8643	5.7855	10.9083	0.6610
0.80	10.6802	5.5487	10.2111	0.6540
0.85	10.5202	5.3440	9.5992	0.6451
0.90	10.3813	5.1668	9.0585	0.6363
0.95	10.2609	5.0135	8.5779	0.6274
1.00	10.1571	4.8815	8.1488	0.6185

Table 4: The values of $T_{21}(\delta)$, $T_{22}(\delta)$, $T_3(\delta)$ and $\overline{\varepsilon}(\delta)$ for some δ.

we will use principally different techniques. The thing is that, as was mentioned above, for $\delta < 1$ the quantity

$$\frac{(\nu_n\ell_n)^{1/\delta}}{2\sqrt{2\pi}} = \frac{1}{2\sqrt{2\pi}}\Big(\ell_n + \frac{1}{B_n^{2+\delta}}\sum_{j=1}^n \beta_{\delta,j}\,\sigma_j^2\Big)^{1/\delta} \leqslant \frac{(2\ell_n)^{1/\delta}}{2\sqrt{2\pi}},$$

appearing in the estimate for Δ_n from lemma 4.3, is an infinitesimal of higher order of decrease than ℓ_n as $\ell_n \to 0$. Therefore, to estimate I_{13} it suffices to use traditional techniques. For $\delta = 1$ this quantity has the same order of decrease as the Lyapounov fraction ℓ_n and, as we will see below, makes the main contribution in the corresponding constant. The use of the same method as for $\delta < 1$ to estimate I_{13} makes it possible to obtain a new moment-type estimate whose structure is in some sense asymptotically optimal. But if this new estimate is used for the construction of the classical estimate with a single term, the Lyapounov fraction, then the coefficient $7/(6\sqrt{2\pi}) = 0.4654\ldots$ at the Lyapounov fraction in this classical estimate will be noticeably greater than its "exact" value $(\sqrt{10}+3)/6/\sqrt{2\pi} = 0.4097\ldots$. So, the new estimate with the asymptotically exact structure is too rough for the solution of the problem in the classical setting. Therefore, to estimate the integral I_{13} in the case $\delta = 1$ we will use another technique which is more delicate and is based on inequality (2.5) from lemma 2.8. This technique develops and sharpens the method used by G. P. Chistyakov in [7].

First consider the general case $\delta \leqslant 1$. With the account of estimates (2.4), (2.6) from lemma 2.8, for the integral I_{13} we obtain

$$I_{13} \leqslant \frac{1}{\pi}\sum_{j=1}^n \int_0^\infty \frac{1}{t}\left(\frac{\gamma_\delta \beta_{2+\delta,j}\, t^{2+\delta}}{B_n^{2+\delta}} + \frac{\sigma_j^4 t^4}{8B_n^4}\,\mathbf{1}(\delta<1)\right) e^{-t^2/2}\,dt$$

$$= C(\delta)\ell_n + \frac{1}{4\pi B_n^4}\sum_{j=1}^n \sigma_j^4 \mathbf{1}(\delta < 1),$$

where $C(\delta) = \gamma_\delta 2^{\delta/2}\Gamma(1+\delta/2)/\pi$. Further, by virtue of the Lyapounov inequality and (4.2) we conclude that

$$I_{13} \leqslant \begin{cases} C(\delta)\ell_n + \ell_n^{4/(2+\delta)}/(4\pi)\mathbf{1}(\delta<1), & \text{in the general case,} \\ C(\delta)\ell_n + (4\pi n)^{-1}\mathbf{1}(\delta<1), & \text{if } F_1 = \ldots = F_n. \end{cases}$$

So, from lemma 4.3 we obtain that for all $n \geqslant 1$ and $F_1, \ldots, F_n \in \mathcal{F}_{2+\delta}$ such that $\ell_n \leqslant \overline{\ell}(\delta)$ the estimate

$$\Delta_n \leqslant C(\delta)\cdot\ell_n + \frac{1}{2\sqrt{2\pi}}\left(\ell_n + \frac{1}{B_n^{2+\delta}}\sum_{j=1}^n \beta_{\delta,j}\,\sigma_j^2\right)^{1/\delta} + \begin{cases} \widetilde{C}_\delta(\ell_n)\cdot\ell_n^{4/(2+\delta)}, & 0<\delta<1, \\ \widetilde{C}_1(\ell_n)\cdot\ell_n^{5/3}, & \delta = 1, \end{cases} \tag{4.8}$$

holds, where

$$\widetilde{C}_\delta(\ell) = \frac{1}{4\pi} + \ell^{\frac{2-\delta(1-\delta)}{\delta(2+\delta)}} J_{12}(\ell,2) + \inf\Big\{\ell^{\delta/(2+\delta)} J_{11}(\ell,2,t_0) + \ell^{\frac{2(2-\delta)}{\delta(2+\delta)}}\cdot\frac{2^{2/\delta}}{4\pi^2}\sup_{0<\varepsilon\leqslant 2\ell} J\left(2\pi\varepsilon^{-1/\delta}, t_0\right) : t_3(\delta) \leqslant t_0 \leqslant t_1(\delta)\wedge t_4(\delta,\ell)\Big\},$$
$$\ell \in \big(0, \overline{\ell}(\delta)\big),\ \delta\in(0,1),$$

$$\widetilde{C}_1(\ell) = \ell^{1/3} J_{12}(\ell,2) + \inf\Big\{J_{11}(\ell,2,t_0) + \ell^{1/3}/\pi^2 \sup_{0<\varepsilon\leqslant 2\ell} J\left(2\pi\varepsilon^{-1/\delta}, t_0\right) : t_3(1) \leqslant t_0 \leqslant t_1(1)\wedge t_4(1,\ell)\Big\},$$
$$\ell \in \big(0, \overline{\ell}(1)\big),$$

and for all $n > \big(\overline{\ell}(\delta)\big)^{-2/(2+\delta)}$, $F_1 = \ldots = F_n \in \mathcal{F}_{2+\delta}$ and

$$t_0 \in \Big[t_3(\delta),\, t_1(\delta)\wedge t_4\left(\delta, n^{-1-\delta/2}\right)\Big)$$

we have

$$\Delta_n \leqslant C(\delta)\cdot\frac{\beta_{2+\delta}}{\sigma^{2+\delta}n^{\delta/2}} + \frac{1}{2\sqrt{2\pi n}}\left(\frac{\beta_{2+\delta}}{\sigma^{2+\delta}} + \frac{\beta_\delta}{\sigma^\delta}\right)^{1/\delta} + \frac{1}{4\pi n}\,\mathbf{1}(\delta<1) + \ell_n^2\bigg(\widehat{J}_{11}(n,\nu_n,t_0)$$

$$+\,\ell_n^{(1-\delta)/\delta}\widehat{J}_{12}(\ell_n,\nu_n,n)+\ell_n^{2(1-\delta)/\delta}\cdot\frac{\nu_n^{2/\delta}}{4\pi^2}\,J\left(\frac{2\pi}{(\nu_n\ell_n)^{1/\delta}},t_0\right)\Bigg). \qquad (4.9)$$

From (4.9) with the account of relations $n\geqslant\ell_n^{-2/\delta}$, $1\leqslant\nu_n\leqslant 2$ and the properties of the functions $\widehat{J}_{11}(n,\nu_n,t_0)$, $\widehat{J}_{12}(\ell_n,\nu_n,n)$, $t_4(\delta,n^{-1-\delta/2})$ described in lemma 4.3 it follows that, uniformly in n and ν_n,

$$\Delta_n\leqslant C(\delta)\cdot\frac{\beta_{2+\delta}}{\sigma^{2+\delta}n^{\delta/2}}+\frac{1}{2\sqrt{2\pi n}}\left(\frac{\beta_{2+\delta}}{\sigma^{2+\delta}}+\frac{\beta_\delta}{\sigma^\delta}\right)^{1/\delta}+\ell_n^2\cdot\widehat{C}_\delta(\ell_n),\quad \ell_n\leqslant\left(\overline{\ell}(\delta)\right)^{\delta/(2+\delta)},$$

where

$$\begin{aligned}\widehat{C}_\delta(\ell)&=\frac{1}{4\pi}\,\mathbf{1}(\delta<1)+\ell^{(1-\delta)/\delta}\widehat{J}_{12}(\ell,2,\ell^{-2/\delta})\\&\quad+\inf\Big\{\ell^{2(1-\delta)/\delta}\cdot\frac{2^{2/\delta}}{4\pi^2}\sup_{0<\varepsilon\leqslant 2\ell}J\left(2\pi\varepsilon^{-1/\delta},t_0\right)\\&\quad+\widehat{J}_{11}(\ell^{-2/\delta},2,t_0)\colon t_3(\delta)\leqslant t_0\leqslant t_1(\delta)\wedge t_4\left(\delta,\ell^{1+2/\delta}\right)\Big\},\\&\quad 0<\ell\leqslant\left(\overline{\ell}(\delta)\right)^{\delta/(2+\delta)}.\end{aligned}$$

For the calculation of the least upper bound of $J(2\pi\varepsilon^{-1/\delta},t_0)$ over $0<\varepsilon\leqslant 2\ell$ see remark 4.5.

Note that for each $0<\delta\leqslant 1$ the functions $\widetilde{C}_\delta(\ell)$ and $\widehat{C}_\delta(\ell)$ increase monotonically varying within the limits

$$\widetilde{C}_\delta(0)\equiv\lim_{\ell\to 0}\widetilde{C}_\delta(\ell)<\widetilde{C}_\delta(\ell)<\lim_{\ell\to\overline{\ell}(\delta)}\widetilde{C}_\delta(\ell)=+\infty,\quad 0<\ell<\overline{\ell}(\delta),$$

$$\widehat{C}_\delta(0)\equiv\lim_{\ell\to 0}\widehat{C}_\delta(\ell)<\widehat{C}_\delta(\ell)<\lim_{\ell\to(\overline{\ell}(\delta))^{\delta/(2+\delta)}}\widehat{C}_\delta(\ell)=+\infty,\quad 0<\ell<(\overline{\ell}(\delta))^{\delta/(2+\delta)},$$

$$\widetilde{C}_\delta(0)=\begin{cases}(4\pi)^{-1}=0.0795\ldots, & 0<\delta<1,\\ \dfrac{2\cdot 1.0253}{3\pi(1-4/9e^{-5/6})^3}=0.4142\ldots, & \delta=1,\end{cases}$$

$$\widehat{C}_\delta(0)=\begin{cases}\dfrac{1.0253\cdot 2^{5/2+\delta}\varkappa_\delta\gamma_\delta\Gamma(5/2+\delta)}{\pi\left(1-2\varkappa_\delta(2\pi t_3(\delta))^\delta\right)^{5/2+\delta}}+\dfrac{1}{4\pi}, & 0<\delta<1,\\ \dfrac{1.0253\cdot 5\varkappa_1}{\sqrt{2\pi}(1-4/9e^{-5/6})^{7/2}}+\dfrac{1}{3\pi}=0.5359\ldots, & \delta=1,\end{cases}$$

infinitely large values of the functions $\widetilde{C}_\delta(\ell)$ and $\widehat{C}_\delta(\ell)$ appear since $t_4(\delta,\ell)\to t_3(\delta)$ as $\ell\to\overline{\ell}(\delta)$, and for all $r>0$

$$\lim_{\ell\to\overline{\ell}(\delta)}\inf_{t_3(\delta)\leqslant t_0\leqslant t_1(\delta)\wedge t_4(\delta,\ell)}Q(\ell,t_0,r)=\lim_{\ell\to\overline{\ell}(\delta)}Q(\ell,t_3(\delta),r)=+\infty.$$

The values of $\widetilde{C}_\delta(\ell)$ for some $0<\delta\leqslant 1$ and ℓ are given in table 6. The values of $\widehat{C}_\delta(0)$ and $\widehat{C}_\delta(\ell)$ are given in table 7.

From inequality (4.9) one can also obtain improved estimates in a special scheme of a double array of row-wise i.i.d. summands:

$$F_j(x)=F_{j,n}(x)=F_{1,n}(x),\quad j=1,\ldots,n,$$

$$\beta_{2+\delta}=\beta_{2+\delta,n},\quad \sigma=\sigma_n,\quad \ell_n=\frac{\beta_{2+\delta,n}}{\sigma_n^{2+\delta}n^{\delta/2}},\quad n\geqslant 1.$$

The double array scheme admits such a dependence of the distributions $F_1,\ldots,F_n$ within each row on the number of the row n that whatever large n is, the Lyapounov fraction ℓ_n may remain fixed and, in particular, may be arbitrarily far from zero. Such a situation occurs, for example, in the construction of estimates of the rate of convergence of the distributions of Poisson random sums of i.i.d. summands with the use of the property of infinite divisibility of the compound Poisson distribution. The success in solving these problems directly depends on the quality of estimates of

$$\limsup_{n\to\infty}\ \sup_{F\in\mathcal{F}_{2+\delta}:\ |\ell_n(F)-\ell|\leqslant\theta_n}\Delta_n(F),$$

with $\ell>0$ and $\{\theta_n\}_{n\geqslant 1}$ being some infinitesimal sequence, to the construction of which we proceed. Recall that $\Delta_n(F)$ denotes the uniform distance between the d.f. of the standard normal law and the d.f. of the standardized sum of i.i.d. r.v.'s with the common d.f. $F\in\mathcal{F}_{2+\delta}$.

First note that for any $\ell>0$ and arbitrary infinitesimal sequence of nonnegative numbers $\{\theta_n\}_{n\geqslant 1}$ by virtue of (4.1) we have

$$\begin{aligned}1&\leqslant\limsup_{n\to\infty}\ \sup_{F_1=\ldots=F_n\in\mathcal{F}_{2+\delta}:\ |\ell_n-\ell|\leqslant\theta_n}\nu_n(F)\leqslant 1+\limsup_{n\to\infty}\ \sup_{\ell_n:\ |\ell_n-\ell|\leqslant\theta_n}\frac{1}{n^{\delta/2}\ell_n}\\&\leqslant 1+\lim_{n\to\infty}\frac{1}{n^{\delta/2}(\ell-\theta_n)}=1,\end{aligned}$$

and with account of the inequality $\varkappa_\delta\leqslant(2\theta_0(\delta))^{-1/\delta}$ (see (2.2))

$$\lim_{n\to\infty}t_4\left(\delta,n^{-1-\delta/2}\right)=\frac{(2\varkappa_\delta)^{-1/\delta}}{2\pi}\geqslant\frac{\theta_0(\delta)}{2\pi}=t_1(\delta).$$

Further, it is easy to make sure that for any $\ell>0$ and $t_0\in\big[t_3(\delta),\,t_1(\delta)\big)$ the relations

$$\begin{aligned}\limsup_{n\to\infty}\ \sup_{|\ell_n-\ell|\leqslant\theta_n}\widehat{J}_{11}(n,\nu_n,t_0)&=1.0253\pi^{-1}\varkappa_\delta\gamma_\delta Q(0,t_0,4+2\delta)\\&=1.0253\frac{2^{3/2+\delta}\varkappa_\delta\gamma_\delta\Gamma(5/2+\delta)}{\pi\left(1-2\varkappa_\delta(2\pi t_0)^\delta\right)^{5/2+\delta}},\end{aligned}$$

$$\limsup_{n\to\infty}\ \sup_{|\ell_n-\ell|\leqslant\theta_n}\widehat{J}_{12}(\ell_n,\nu_n,n)=2^{(\delta-1)/2}\pi^{-1}\gamma_\delta\times$$

$$\times\Gamma\Big(\frac{3+\delta}{2}\Big)\left(1+\frac{(3+\delta)\ell^{2/\delta}}{72}\right)\to\infty,\quad \ell\to\infty,$$

hold, where the least upper bounds are taken over all $F_1=\ldots=F_n\equiv F\in\mathcal{F}_{2+\delta}$ such that $|\ell_n(F)-\ell|\leqslant\theta_n$. So, from (4.9) for all $\ell>0$ follows the estimate

$$\limsup_{n\to\infty}\sup_{F\in\mathcal{F}_{2+\delta}\colon\,|\ell_n-\ell|\leqslant\theta_n}\Delta_n(F)\leqslant C(\delta)\cdot\ell+\frac{\ell^{1/\delta}}{2\sqrt{2\pi}}+C'_\delta(\ell)\cdot\ell^2,$$

where

$$C'_\delta(\ell)=\ell^{(1-\delta)/\delta}2^{(\delta-1)/2}\pi^{-1}\gamma_\delta\Gamma\Big(\frac{3+\delta}{2}\Big)\left(1+\frac{(3+\delta)\ell^{2/\delta}}{72}\right)$$
$$+\inf_{t_3(\delta)\leqslant t_0<t_1(\delta)}\left(1.0253\frac{2^{3/2+\delta}\varkappa_\delta\gamma_\delta\Gamma(5/2+\delta)}{\pi\left(1-2\varkappa_\delta(2\pi t_0)^\delta\right)^{5/2+\delta}}+\frac{\ell^{2(1-\delta)/\delta}}{4\pi^2}\sup_{0<\varepsilon\leqslant\ell}J\left(\frac{2\pi}{\varepsilon^{1/\delta}},t_0\right)\right).$$

For the calculation of the least upper bound of $J(2\pi\varepsilon^{-1/\delta},t_0)$ over $0<\varepsilon\leqslant\ell$ see remark 4.5. Note that for each $0<\delta\leqslant1$ the function $C'_\delta(\ell)$ increases monotonically varying within the limits

$$C'_\delta(0)\equiv\lim_{\ell\to0}C'_\delta(\ell)<C'_\delta(\ell)<\lim_{\ell\to\infty}C'_\delta(\ell)=+\infty,\quad 0<\delta\leqslant1,\ \ell>0,$$

$$C'_\delta(0)=\begin{cases}\frac{1.0253\cdot2^{3/2+\delta}\varkappa_\delta\gamma_\delta\Gamma(5/2+\delta)}{\pi(1-2\varkappa_\delta(2\pi t_3(\delta))^\delta)^{5/2+\delta}}=\frac{1.0253\cdot2^{3/2+\delta}\varkappa_\delta\gamma_\delta\Gamma(5/2+\delta)}{\pi\left(1-\frac{4}{(2+\delta)^2}\exp\left\{-\frac{\delta(4+\delta)}{2(2+\delta)}\right\}\right)^{5/2+\delta}}, & 0<\delta<1,\\ C'_{1-}(0)+\frac{1}{6\pi}=\frac{1.0253\cdot5\varkappa_1}{2\sqrt{2\pi}(1-4/9e^{-5/6})^{7/2}}+\frac{1}{6\pi}=0.2679\ldots, & \delta=1.\end{cases}$$

The values of $C'_\delta(0)$ and $C'_\delta(\ell)$ for some ℓ and $0<\delta\leqslant1$ are given in table 8.

To obtain estimates with constants $\widetilde{C}_\delta$, $\widehat{C}_\delta$, C'_δ at remainders bounded for all $\ell_n>0$, note that if $\ell_n\geqslant\ell$ for some $\ell>0$, then by virtue of (4.3) for any

$$A\geqslant\kappa-C(\delta)\cdot\ell-\frac{\ell^{1/\delta}}{2\sqrt{2\pi}}$$

the trivial estimate

$$C(\delta)\cdot\ell_n+\frac{1}{2\sqrt{2\pi}}\left(\ell_n+\frac{1}{B_n^{2+\delta}}\sum_{j=1}^n\beta_{\delta,j}\,\sigma_j^2\right)^{1/\delta}+A\geqslant C(\delta)\cdot\ell+\frac{\ell^{1/\delta}}{2\sqrt{2\pi}}+A\geqslant\kappa\geqslant\Delta_n,$$

holds so that the quantities $\widetilde{C}_\delta(\ell_n)\ell_n^{4/(2+\delta)}$ and $\widetilde{C}_1(\ell_n)\ell_n^{5/3}$ in (4.8) for $\ell_n\geqslant\ell$ can be respectively replaced by $\min\left\{\widetilde{C}_\delta(\ell_n)\ell_n^{4/(2+\delta)},\,\kappa-C(\delta)\cdot\ell-(2\sqrt{2\pi})^{-1}\ell^{1/\delta}\right\}$ and $\min\left\{\widetilde{C}_1(\ell_n)\ell_n^{5/3},\,\kappa-2/(3\sqrt{2\pi})\ell\right\}$ for any $\ell\in(0,\overline{\ell}(\delta))$. Note that

$$\kappa-C(\delta)\cdot\ell-\frac{\ell^{1/\delta}}{2\sqrt{2\pi}}\leqslant\kappa-\frac{\ell^{1/\delta}}{2\sqrt{2\pi}}\leqslant0\quad\text{for }\ell\geqslant(2\sqrt{2\pi}\kappa)^\delta.$$

Define $\widetilde{\ell}(\delta)$ as the unique root of the equation

$$\widetilde{C}_\delta(\ell)\cdot\ell^{4/(2+\delta)}=\kappa-C(\delta)\cdot\ell-\frac{\ell^{1/\delta}}{2\sqrt{2\pi}},\quad 0<\delta<1,$$

$$\widetilde{C}_1(\ell)\cdot\ell^{5/3}=\kappa-\frac{2\ell}{3\sqrt{2\pi}},\quad \delta=1,$$

on the interval $0<\ell<\overline{\ell}(\delta)\wedge(2\sqrt{2\pi}\kappa)^\delta=\overline{\ell}(\delta)$ (recall that, by definition, $\overline{\ell}(\delta)<1<(2\sqrt{2\pi}\kappa)^\delta$ for all $0<\delta\leqslant 1$, since $\kappa=0.54\ldots>1/2$, see (4.3)). The existence and uniqueness of $\widetilde{\ell}(\delta)$ follow from that on the interval under consideration the left-hand side of the equation is a continuous strictly monotonically increasing function taking *all* values from 0 to $+\infty$, and the right-hand side is a continuous strictly monotonically decreasing function taking positive values at small ℓ, that is, the graphs of these functions have a unique point of intersection on the interval $(0,\overline{\ell}(\delta))$. So, since the function $\widetilde{C}_\delta$ increases monotonically in ℓ, for any $\ell>0$ the estimate

$$\Delta_n\leqslant C(\delta)\cdot\ell_n+\frac{1}{2\sqrt{2\pi}}\Bigg(\ell_n+\frac{1}{B_n^{2+\delta}}\sum_{j=1}^n\beta_{\delta,j}\,\sigma_j^2\Bigg)^{1/\delta}+\begin{cases}\widetilde{C}_\delta\Big(\ell\wedge\widetilde{\ell}(\delta)\Big)\cdot\ell_n^{4/(2+\delta)}, & 0<\delta<1,\\ \widetilde{C}_1\Big(\ell\wedge\widetilde{\ell}(1)\Big)\cdot\ell_n^{5/3}, & \delta=1.\end{cases}$$

holds for all $n\geqslant 1$ and $F_1,\ldots,F_n\in\mathcal{F}_{2+\delta}$ such that $\ell_n\leqslant\ell$.

Similar reasoning also can be applied to the functions $\widehat{C}_\delta(\ell)$, $C'_\delta(\ell)$ with the only remark that for $C'_\delta(\ell)$ the root of the corresponding equation lies within the interval $\big(0,\,(2\sqrt{2\pi}\kappa)^\delta\big)$ which results in the following theorem.

Theorem 4.6. *For any $0<\delta\leqslant 1$ and $\ell>0$, for all $n\geqslant 1$ and $F_1,\ldots,F_n\in\mathcal{F}_{2+\delta}$ such that $\ell_n\leqslant\ell$, the following estimates hold: in the general case*

$$\Delta_n\leqslant C(\delta)\cdot\ell_n+\frac{1}{2\sqrt{2\pi}}\Bigg(\ell_n+\frac{1}{B_n^{2+\delta}}\sum_{j=1}^n\beta_{\delta,j}\,\sigma_j^2\Bigg)^{1/\delta}+\begin{cases}\widetilde{C}_\delta\Big(\ell\wedge\widetilde{\ell}(\delta)\Big)\cdot\ell_n^{4/(2+\delta)}, & 0<\delta<1,\\ \widetilde{C}_1\Big(\ell\wedge\widetilde{\ell}(1)\Big)\cdot\ell_n^{5/3}, & \delta=1,\end{cases}$$

in the case $F_1=\ldots=F_n$

$$\Delta_n\leqslant C(\delta)\cdot\frac{\beta_{2+\delta}}{\sigma^{2+\delta}n^{\delta/2}}+\frac{1}{2\sqrt{2\pi n}}\bigg(\frac{\beta_{2+\delta}}{\sigma^{2+\delta}}+\frac{\beta_\delta}{\sigma^\delta}\bigg)^{1/\delta}+\widehat{C}_\delta\Big(\ell\wedge\widehat{\ell}(\delta)\Big)\cdot\ell_n^2,$$

and also for any $\ell>0$ and arbitrary infinitesimal sequence of nonnegative numbers $\{\theta_n\}_{n\geqslant1}$

$$\limsup_{n\to\infty}\sup_{F\in\mathcal{F}_{2+\delta}:\,|\ell_n-\ell|\leqslant\theta_n}\Delta_n(F)\leqslant C(\delta)\cdot\ell+\frac{\ell^{1/\delta}}{2\sqrt{2\pi}}+C'_\delta\left(\ell\wedge\ell'(\delta)\right)\cdot\ell^2,$$

where

$$C(\delta)=\frac{\gamma_\delta 2^{\delta/2}}{\pi}\,\Gamma\left(\frac{2+\delta}{2}\right),$$

$$\begin{aligned}\widetilde{C}_\delta(\ell)=\frac{1}{4\pi}+\ell^{\frac{2-\delta(1-\delta)}{\delta(2+\delta)}}J_{12}(\ell,2)+\inf\Big\{\ell^{\delta/(2+\delta)}J_{11}(\ell,2,t_0)\\ +\ell^{\frac{2(2-\delta)}{\delta(2+\delta)}}\cdot\frac{2^{2/\delta}}{4\pi^2}\sup_{0<\varepsilon\leqslant 2\ell}J\left(2\pi\varepsilon^{-1/\delta},t_0\right):t_3(\delta)\leqslant t_0\leqslant t_1(\delta)\wedge t_4\left(\delta,\ell\right)\Big\},\\ 0<\delta<1,\end{aligned}$$

$$\begin{aligned}\widetilde{C}_1(\ell)=\inf\Big\{J_{11}(\ell,2,t_0)+\ell^{1/3}J_{12}(\ell,2)\\ +\ell^{1/3}\pi^{-2}\sup_{0<\varepsilon\leqslant 2\ell}J\left(2\pi\varepsilon^{-1/\delta},t_0\right):t_3(1)\leqslant t_0\leqslant t_1(1)\wedge t_4\left(1,\ell\right)\Big\},\end{aligned}$$

$$\begin{aligned}\widehat{C}_\delta(\ell)=\frac{\mathbf{1}(\delta<1)}{4\pi}+\ell^{(1-\delta)/\delta}\widehat{J}_{12}(\ell,2,\ell^{-2/\delta})\\ +\inf\Big\{\ell^{2(1-\delta)/\delta}\cdot\frac{2^{2/\delta}}{4\pi^2}\sup_{0<\varepsilon\leqslant 2\ell}J\left(2\pi\varepsilon^{-1/\delta},t_0\right)\\ +\widehat{J}_{11}(\ell^{-2/\delta},2,t_0)\colon t_3(\delta)\leqslant t_0\leqslant t_1(\delta)\wedge t_4\left(\delta,\ell^{1+2/\delta}\right)\Big\},\end{aligned}$$

$$\begin{aligned}C'_\delta(\ell)=\ell^{(1-\delta)/\delta}2^{(\delta-1)/2}\pi^{-1}\gamma_\delta\Gamma\Big(\frac{3+\delta}{2}\Big)\left(1+\frac{(3+\delta)\ell^{2/\delta}}{72}\right)\\ +\inf_{t_3(\delta)\leqslant t_0<t_1(\delta)}\left(1.0253\frac{2^{3/2+\delta}\varkappa_\delta\gamma_\delta\Gamma(5/2+\delta)}{\pi\left(1-2\varkappa_\delta(2\pi t_0)^\delta\right)^{5/2+\delta}}+\frac{\ell^{2(1-\delta)/\delta}}{4\pi^2}\sup_{0<\varepsilon\leqslant\ell}J\left(\frac{2\pi}{\varepsilon^{1/\delta}},t_0\right)\right),\end{aligned}$$

$\widetilde{\ell}(1)$ *is the unique root of the equation* $\widetilde{C}_1(\ell)\cdot\ell^{5/3}=\kappa-2\ell/(3\sqrt{2\pi})$ *on the interval* $0<\ell<\overline{\ell}(1)$, $\widetilde{\ell}(\delta)$, $\widehat{\ell}(\delta)$, $\ell'(\delta)$ *are respectively the unique roots of the equations*

$$\widetilde{C}_\delta(\ell)\cdot\ell^{4/(2+\delta)}=\kappa-C(\delta)\cdot\ell-\frac{\ell^{1/\delta}}{2\sqrt{2\pi}},\quad 0<\ell<\overline{\ell}(\delta),\ 0<\delta<1,$$

$$\widehat{C}_\delta(\ell)\cdot\ell^2=\kappa-C(\delta)\cdot\ell-\frac{\ell^{1/\delta}}{2\sqrt{2\pi}},\quad 0<\ell<\left(\overline{\ell}(\delta)\right)^{\delta/(2+\delta)},\ 0<\delta\leqslant 1,$$

$$C'_\delta(\ell)\cdot\ell^2=\kappa-C(\delta)\cdot\ell-\frac{\ell^{1/\delta}}{2\sqrt{2\pi}},\quad 0<\ell<(2\sqrt{2\pi}\kappa)^\delta,\ 0<\delta\leqslant 1,$$

on the intervals specified above; $\kappa=0.5409\ldots$ *is defined in* (4.3)*;* γ_δ, $\varkappa_\delta$, $t_1(\delta)$, $t_3(\delta)$, $t_4(\delta,\ell)$, $\overline{\ell}(\delta)$, $J_{11}(\ell,\nu,t_0)$, $\widehat{J}_{11}(n,\nu,t_0)$, $J_{12}(\ell,\nu)$, $\widehat{J}_{12}(\ell,\nu,n)$, $J(T,t_0)$, $T>0$, *are defined in lemma* 4.3.

$\delta =$	$C(\delta) \leqslant$	$\underline{C_{\rm AE}}(\delta) \geqslant$	$\delta =$	$C(\delta) \leqslant$	$\underline{C_{\rm AE}}(\delta) \geqslant$	$\delta =$	$C(\delta) \leqslant$	$\underline{C_{\rm AE}}(\delta) \geqslant$
0+	0.1693	0.0883	0.35	0.1017	0.0422	0.70	0.0709	0.0253
0.05	0.1561	0.0759	0.40	0.0956	0.0390	0.75	0.0685	0.0237
0.10	0.1444	0.0674	0.45	0.0902	0.0361	0.80	0.0665	0.0223
0.15	0.1339	0.0606	0.50	0.0854	0.0334	0.85	0.0650	0.0210
0.20	0.1245	0.0550	0.55	0.0810	0.0311	0.90	0.0642	0.0198
0.25	0.1161	0.0501	0.60	0.0772	0.0290	0.95	0.0642	0.0187
0.30	0.1085	0.0459	0.65	0.0738	0.0271	1−	0.0665	0.0177

Table 5: The values of $C(\delta)$ from theorem 4.6 which bounds above the asymptotically exact constant $C_{\rm AE}(\delta)$ (see theorem 4.12) rounded up to the fourth decimal digit and the corresponding values of the lower bound for the lower asymptotically exact constant $\underline{C_{\rm AE}}(\delta)$ (see (4.10)) for some $0 < \delta \leqslant 1$. By definition, $\underline{C_{\rm AE}}(\delta) \leqslant C_{\rm AE}(\delta) \leqslant C(\delta)$ for all $0 < \delta \leqslant 1$.

$\delta =$	$\widetilde{\ell}(\delta) \leqslant$	$\widetilde{C}_\delta(\widetilde{\ell}(\delta)) \leqslant$	$\widetilde{C}_\delta(0.1) \leqslant$	$\widetilde{C}_\delta(0.01) \leqslant$	$\widetilde{C}_\delta(10^{-3}) \leqslant$	$\widetilde{C}_\delta(10^{-4}) \leqslant$
0.05	0.0218	943.5902	943.5902	492.0103	290.6531	253.8418
0.10	0.0437	208.2037	208.2037	67.6270	43.7421	35.7650
0.15	0.0635	89.9006	89.9006	21.7830	13.7457	10.5124
0.20	0.0812	51.0184	51.0184	9.7720	5.8904	4.2460
0.25	0.0969	33.5946	33.5946	5.2585	3.0192	2.0712
0.30	0.1108	24.2825	20.0463	3.1846	1.7473	1.1531
0.35	0.1230	18.7024	12.7760	2.0993	1.1074	0.7110
0.40	0.1337	15.0778	8.8825	1.4770	0.7546	0.4767
0.45	0.1430	12.5785	6.5814	1.0951	0.5460	0.3431
0.50	0.1511	10.7742	5.1210	0.8479	0.4157	0.2623
0.55	0.1580	9.4240	4.1429	0.6812	0.3306	0.2111
0.60	0.1639	8.3842	3.4602	0.5648	0.2730	0.1773
0.65	0.1688	7.5649	2.9681	0.4814	0.2328	0.1543
0.70	0.1728	6.9071	2.6044	0.4203	0.2040	0.1381
0.75	0.1761	6.3715	2.3308	0.3748	0.1830	0.1266
0.80	0.1786	5.9306	2.1226	0.3405	0.1675	0.1182
0.85	0.1804	5.5650	1.9638	0.3146	0.1559	0.1119
0.90	0.1814	5.2610	1.8442	0.2953	0.1473	0.1073
0.95	0.1818	5.0102	1.7588	0.2816	0.1411	0.1040
1.00	0.2325	5.4527	1.6948	0.6317	0.4856	0.4427

Table 6: The values of $\widetilde{\ell}(\delta)$ and $\widetilde{C}_\delta(\ell)$ from theorem 4.6 for $\ell = \widetilde{\ell}(\delta)$, $0.1 \wedge \widetilde{\ell}(\delta)$, 0.01, 0.001, 0.0001 and some $0 < \delta \leqslant 1$; the fourth column contains the values of $\widetilde{C}_\delta\big(0.1 \wedge \widetilde{\ell}(\delta)\big)$. The optimal values of t_0 coincide with $t_3(\delta)$ (see table 3).

The values of $C(\delta)$, $\widetilde{\ell}(\delta)$, $\widehat{\ell}(\delta)$, $\ell'(\delta)$, $\widetilde{C}_\delta\big(\ell\big)$, $\widehat{C}_\delta\big(\ell\big)$, $C'_\delta(\ell)$ rounded above up to the fourth decimal digit are given in tables 5, 6, 7, 8 for some $0 < \delta \leqslant 1$ and $\ell \geqslant 0$. The computations were carried out in the Matlab R2011a environment.

Since $C(1) = 1/(6\sqrt{2\pi})$, from theorem 4.6 for $\delta = 1$ we obtain

Corollary 4.7. *For all $n \geqslant 1$ and $F_1, \ldots, F_n \in \mathcal{F}_3$*

$$\Delta_n \leqslant \frac{2\ell_n}{3\sqrt{2\pi}} + \frac{1}{2\sqrt{2\pi} B_n^3} \sum_{j=1}^{n} \beta_{1,j}\, \sigma_j^2 + 5.4527 \cdot \ell_n^{5/3}$$

$\delta =$	$\widehat{\ell}(\delta) \leqslant$	$t_0 =$	$\widehat{C}_\delta(\widehat{\ell}(\delta)) \leqslant$	$\widehat{C}_\delta(0.1) \leqslant$	$\widehat{C}_\delta(0.01) \leqslant$	$\widehat{C}_\delta\,(10^{-3}) \leqslant$	$\widehat{C}_\delta(0+) \leqslant$
0.05	0.0468	0.1370	243.6690	243.6690	243.6690	243.6690	243.6690
0.10	0.1050	0.1386	47.7282	47.7282	47.7282	47.7282	47.7282
0.15	0.1662	0.1401	18.7976	18.7973	18.7973	18.7973	18.7973
0.20	0.2283	0.1416	9.8319	9.8249	9.8246	9.8246	9.8246
0.25	0.2897	0.1431	6.0285	5.9929	5.9916	5.9916	5.9916
0.30	0.3407	0.1444	4.2951	4.0322	4.0288	4.0287	4.0287
0.35	0.3652	0.1457	3.6948	2.9060	2.8988	2.8987	2.8987
0.40	0.3795	0.1469	3.3818	2.2057	2.1932	2.1928	2.1928
0.45	0.3889	0.1480	3.1837	1.7448	1.7256	1.7246	1.7245
0.50	0.3950	0.1490	3.0525	1.4292	1.4018	1.3996	1.3994
0.55	0.3987	0.1525	2.9657	1.2069	1.1702	1.1661	1.1654
0.60	0.4005	0.1563	2.9104	1.0480	1.0007	0.9941	0.9923
0.65	0.4007	0.1588	2.8812	0.9338	0.8749	0.8652	0.8614
0.70	0.3996	0.1603	2.8742	0.8526	0.7811	0.7682	0.7608
0.75	0.3973	0.1613	2.8863	0.7968	0.7117	0.6957	0.6826
0.80	0.3940	0.1618	2.9157	0.7614	0.6618	0.6432	0.6216
0.85	0.3898	0.1622	2.9611	0.7435	0.6283	0.6081	0.5744
0.90	0.3847	0.1623	3.0228	0.7417	0.6097	0.5895	0.5389
0.95	0.3787	0.1623	3.1027	0.7571	0.6069	0.5886	0.5148
1.00	0.4180	0.1770	2.4606	0.6023	0.5403	0.5364	0.5360

Table 7: The values of $\widehat{\ell}(\delta)$ and $\widehat{C}_\delta(\ell)$ from theorem 4.6 for $\ell = \widehat{\ell}(\delta)$, $0.1 \wedge \widehat{\ell}(\delta)$, 0.01, 0.001 and $\ell \to 0+$ for some $0 < \delta \leqslant 1$. The third column contains the optimal values of t_0 delivering the infimum in $\widehat{C}_\delta(\widehat{\ell}(\delta))$, for other ℓ the optimal values of t_0 coincide with $t_3(\delta)$ (see table 3).

$\delta =$	$\ell'(\delta) \leqslant$	$t_0 =$	$C'_\delta(\ell'(\delta)) \leqslant$	$C'_\delta(0.5) \leqslant$	$C'_\delta(0.1) \leqslant$	$C'_\delta(0+) \leqslant$
0.05	0.0661	0.1370	121.7947	121.7947	121.7947	121.7947
0.10	0.1477	0.1386	23.8244	23.8244	23.8244	23.8244
0.15	0.2334	0.1401	9.3589	9.3589	9.3589	9.3589
0.20	0.3205	0.1416	4.8734	4.8734	4.8726	4.8726
0.25	0.4062	0.1431	2.9613	2.9613	2.9561	2.9561
0.30	0.4863	0.1444	1.9884	1.9884	1.9750	1.9746
0.35	0.5581	0.1457	1.4339	1.4291	1.4106	1.4096
0.40	0.6170	0.1469	1.1095	1.0805	1.0588	1.0566
0.45	0.6577	0.1480	0.9320	0.8508	0.8263	0.8225
0.50	0.6867	0.1490	0.8237	0.6935	0.6660	0.6599
0.55	0.7094	0.1500	0.7485	0.5833	0.5519	0.5429
0.60	0.7283	0.1513	0.6924	0.5051	0.4686	0.4564
0.65	0.7457	0.1621	0.6456	0.4497	0.4068	0.3909
0.70	0.7628	0.1717	0.6040	0.4110	0.3605	0.3406
0.75	0.7794	0.1801	0.5673	0.3847	0.3256	0.3015
0.80	0.7950	0.1874	0.5356	0.3680	0.2997	0.2710
0.85	0.8091	0.1937	0.5087	0.3565	0.2811	0.2474
0.90	0.8209	0.1988	0.4870	0.3492	0.2687	0.2297
0.95	0.8291	0.2026	0.4714	0.3469	0.2628	0.2176
1.00	0.8280	0.2044	0.4679	0.3559	0.2684	0.2680

Table 8: The values of $\ell'(\delta)$ and $C'_\delta(\ell)$ from theorem 4.6 for $\ell = \ell'(\delta)$, $0.5 \wedge \ell'(\delta)$, $0.1 \wedge \ell'(\delta)$ and $\ell \to 0+$ for some $0 < \delta \leqslant 1$. The third column contains the optimal values of t_0 delivering the infimum in $C'_\delta(\ell'(\delta))$, for other ℓ the optimal values of t_0 coincide with $t_3(\delta)$ (see table 3).

in the general case and

$$\Delta_n \leqslant \frac{2}{3\sqrt{2\pi}} \cdot \frac{\beta_3}{\sigma^3\sqrt{n}} + \frac{1}{2\sqrt{2\pi}} \cdot \frac{\beta_1}{\sigma\sqrt{n}} + 2.4606 \cdot \ell_n^2,$$

if $F_1 = \ldots = F_n$.

Remark 4.8. Corollary 4.7 improves the inequalities of Prawitz (1.9)

$$\Delta_n \leqslant \frac{2}{3\sqrt{2\pi}} \cdot \frac{\beta_3}{\sigma^3\sqrt{n-1}} + \frac{1}{2\sqrt{2\pi(n-1)}} + A_3 \cdot \ell_{n-1}^2, \quad n \geqslant 1,\ F_1 = \ldots = F_n \in \mathcal{F}_3,$$

and Bentkus (1.10)

$$\begin{aligned}\Delta_n &\leqslant \frac{2\ell_n}{3\sqrt{2\pi}} + \frac{1}{2\sqrt{2\pi}B_n^3}\sum_{j=1}^n \sigma_j^3 + A_4 \cdot \ell_n^{4/3} \\ &\leqslant \frac{7\ell_n}{6\sqrt{2\pi}} + A_4 \cdot \ell_n^{4/3}, \quad n \geqslant 1,\ F_1, \ldots, F_n \in \mathcal{F}_3,\end{aligned}$$

first, with respect to the second term, since $\beta_{1,j} \leqslant \sigma_j$, $j = 1, \ldots, n$, by the Lyapounov inequality, and second, with respect to the remainder, since it gives concrete values of the constants A_3 and A_4. And as regards the general case, corollary 4.7 also improves the order of decrease of the remainder to $\ell_n^{5/3}$ as compared with $\ell_n^{4/3}$ in Bentkus' inequality.

Remark 4.9. The values of the coefficients $2/(3\sqrt{2\pi})$ and $\left(2\sqrt{2\pi}\right)^{-1}$ in the estimates given in corollary 4.7 are optimal in the sense that whatever the coefficient at the second term is, the coefficient $2/(3\sqrt{2\pi})$ at the first term cannot be made less and for the given value $2/(3\sqrt{2\pi})$ of the coefficient at the first term, the coefficient at the second term cannot be made less than $\left(2\sqrt{2\pi}\right)^{-1}$. To make this sure it suffices to consider the estimates of the form

$$\Delta_n \leqslant C \cdot \frac{\beta_3}{\sigma^3\sqrt{n}} + K \cdot \frac{\beta_1}{\sigma\sqrt{n}} + A \cdot \ell_n^{1+\theta},$$

with some constants C, K, $A \in \mathbf{R}$ and $\theta > 0$ assuming that they hold for all (or at least for large enough) values of n and all $F_1 = \ldots = F_n \in \mathcal{F}_3$, and notice that by virtue of these estimates

$$\begin{aligned}\underline{C_{\text{AE}}} &= \limsup_{\ell\to 0}\limsup_{n\to\infty} \sup_{F\colon \beta_3=\sigma^3\ell\sqrt{n}} \frac{\Delta_n(F)}{\ell} \\ &\leqslant C + \limsup_{\ell\to 0}\limsup_{n\to\infty} \sup_{F\colon \beta_3=\sigma^3\ell\sqrt{n}} K \cdot \frac{\beta_1}{\sigma\sqrt{n}\ell} \leqslant C,\end{aligned}$$

since $K\beta_1/(\sigma\sqrt{n}\ell) \leqslant 0$ for $K \leqslant 0$, and for $K > 0$ by virtue of the Lyapounov inequality

$$K \cdot \frac{\beta_1}{\sigma\sqrt{n}\ell} \leqslant K \cdot \frac{1}{\ell\sqrt{n}} \to 0, \quad n \to \infty,$$

for any $\ell > 0$. So, with the account of the equality $\underline{C_{\text{AE}}} = 2/(3\sqrt{2\pi})$ [27] we conclude that for any $K \in \mathbf{R}$

$$C \geqslant \underline{C_{\text{AE}}} = \frac{2}{3\sqrt{2\pi}}.$$

Now let $C = 2/(3\sqrt{2\pi})$. Show that in this case K is no less than $\big(2\sqrt{2\pi}\big)^{-1}$. Indeed, by virtue of (1.4) we have

$$K \geqslant \sup_{X_1\in\mathcal{F}_3}\limsup_{n\to\infty}\frac{3\sqrt{2\pi n}\Delta_n(\mathsf{E}X_1^2)^{3/2}-2\mathsf{E}|X_1|^3}{3\sqrt{2\pi}\mathsf{E}|X_1|\mathsf{E}X_1^2}$$
$$=\sup_{h>0}\sup_{X\in\mathcal{F}_3^h}\frac{|\mathsf{E}X^3|+3h\mathsf{E}X^2-4\mathsf{E}|X|^3}{6\sqrt{2\pi}\mathsf{E}|X|\mathsf{E}X^2}.$$

Now letting $\mathsf{P}(X=-\sqrt{p/q})=q$, $\mathsf{P}(X=\sqrt{q/p})=p=1-q$, $0<p\leqslant 1/2$, we arrive at

$$\mathsf{E}X=0,\ \mathsf{E}X^2=1,\ \mathsf{E}X^3=\frac{q-p}{\sqrt{pq}},\ \mathsf{E}|X|=2\sqrt{pq},\ \mathsf{E}|X|^3=\frac{p^2+q^2}{\sqrt{pq}},\ h=\frac{1}{\sqrt{pq}},$$

and hence,

$$K\geqslant \sup_{0<p<1/2,\ q=1-p}\frac{q-p+3-4(p^2+q^2)}{12\sqrt{2\pi}pq}=\frac{1}{6\sqrt{2\pi}}\lim_{p\to 0+}\frac{3-4p}{1-p}=\frac{1}{2\sqrt{2\pi}}.$$

Remark 4.10. The estimate given in corollary 4.7, for summands with the common symmetric distribution $\mathsf{P}(X=\pm 1)=1/2$ with the moments $\beta_1=\sigma^2=\beta_3=1$, takes the form

$$\Delta_n\leqslant\frac{7}{6\sqrt{2\pi n}}+2.4606\ell_n^2=\frac{7\ell_n}{6\sqrt{2\pi}}+2.4606\ell_n^2.$$

On the other hand, for the distribution under consideration it follows from Esseen's asymptotic expansion (1.3) (see [9, 10]) that

$$\Delta_n=\frac{1}{\sqrt{2\pi n}}+o\left(\frac{1}{\sqrt{n}}\right)=\frac{\ell_n}{\sqrt{2\pi}}+o\left(\ell_n\right),\quad n\to\infty,$$

that is, the "exact" constant at the Lyapounov fraction ℓ_n is $7/6\approx 1.17$ times less than that given by the "optimal" estimate from corollary 4.7. Actually there is no paradox, since the obtained estimate is optimal in another sense, but the remark reveals the fact that to obtain estimates with "exact" coefficients at the Lyapounov fraction, the information concerning all first three *absolute* moments is not enough and it is required to use also the information concerning the *original* moments, the only informative of which is the third, since the summands are assumed centered.

Corollary 4.11. *For any $\ell>0$ and arbitrary infinitesimal sequence of nonnegative numbers $\{\theta_n\}_{n\geqslant 1}$*

$$\limsup_{n\to\infty}\sup_{F\in\mathcal{F}_3:\ |\ell_n-\ell|\leqslant\theta_n}\Delta_n(F)\leqslant\frac{2\ell}{3\sqrt{2\pi}}+C_1'\left(\ell\wedge 0.8280\right)\cdot\ell^2\leqslant 0.2660\cdot\ell+0.4679\cdot\ell^2,$$

where $C_1'(\ell)$ *is defined in theorem* 4.6. *In particular,* $C_1'(0.1)\leqslant 0.2684$, *and for all* $0<\ell\leqslant 0.1$

$$\limsup_{n\to\infty}\sup_{F\in\mathcal{F}_3\colon|\ell_n-\ell|\leqslant\theta_n}\Delta_n(F)\leqslant\frac{2\ell}{3\sqrt{2\pi}}+0.2684\cdot\ell^2<\begin{cases}0.2929\cdot\ell, & \ell\leqslant 0.1,\\ 0.2687\cdot\ell, & \ell\leqslant 10^{-2},\\ 0.2663\cdot\ell, & \ell\leqslant 10^{-3},\\ 0.2660\cdot\ell, & \ell\leqslant 10^{-4}.\end{cases}$$

Letting $\ell\to 0$, from theorem 4.6 one can obtain an upper bound for the asymptotically exact constant

$$C_{\text{AE}}(\delta)=\limsup_{\ell\to 0}\sup_{n\geqslant 1,\,F_1,\ldots,F_n\in\mathcal{F}_{2+\delta}\colon\,\ell_n=\ell}\Delta_n(F_1,\ldots,F_n)/\ell,\quad 0<\delta\leqslant 1.$$

Theorem 4.12. *For all* $0<\delta<1$ *the estimate* $C_{\text{AE}}(\delta)\leqslant C(\delta)$ *holds with* $C(\delta)$ *defined in theorem* 4.6. *In particular,*

$$\lim_{\delta\to 1-}C_{\text{AE}}(\delta)\leqslant\frac{1}{6\sqrt{2\pi}}<0.0665,\qquad \lim_{\delta\to 0+}C_{\text{AE}}(\delta)\leqslant\frac{\gamma_0}{\pi}<0.1693.$$

The values of $C(\delta)$ *for other* $0<\delta<1$ *are given in table* 5.

For the scheme of summation of identically distributed r.v.'s theorem 4.12 was proved in [11].

The lower bounds for the asymptotically exact constant $C_{\text{AE}}(\delta)$ for $0<\delta<1$ were obtained in [29] in terms of the so-called lower asymptotically exact constant

$$\underline{C_{\text{AE}}}(\delta)=\limsup_{\ell\to 0}\limsup_{n\to\infty}\sup_{F\in\mathcal{F}_{2+\delta}\colon\,\ell_n=\ell}\Delta_n(F)/\ell$$

and have the form

$$C_{\text{AE}}(\delta)\geqslant\underline{C_{\text{AE}}}(\delta)\geqslant\sup_{a\geqslant 0,\,b>0}\frac{\frac{4}{\sqrt{2+b^2}}\exp\left\{-\frac{a^2}{2(2+b^2)}\right\}+\frac{a^2+b^2}{\sqrt{2}}-2\sqrt{2}}{8M_{2+\delta}b^{2+\delta}e^{-a^2/(2b^2)}{}_1F_1\left(\frac{3+\delta}{2},\frac{1}{2},\frac{a^2}{2b^2}\right)},\tag{4.10}$$

where $\Gamma(\cdot)$ is the Euler's gamma-function, ${}_1F_1$ is the generalized hypergeometric function (the degenerate Meijer function), $M_{2+\delta}$ is the absolute moment of order $2+\delta$ of the standard normal law. The values of the lower bound mentioned above, as well as those of the corresponding upper bound, are given in table 5.

For $\delta=1$ from theorem 4.6 one can obtain only the estimate

$$C_{\text{AE}}(1)\leqslant\frac{7}{6\sqrt{2\pi}}=0.4654\ldots,$$

whereas Chistyakov [7] showed that actually

$$C_{\text{AE}}(1)=\frac{\sqrt{10}+3}{6\sqrt{2\pi}}=0.4097\ldots,$$

that is, the technique used above is too rough for the construction of asymptotically exact estimates in the classical setting in the case $\delta = 1$, and, as it has been noted in remark 4.10, the only way of sharpening of this technique is the use of the information concerning the third *original* moments. This information can be taken into account, if to estimate the absolute value of the difference of ch.f.'s in the integral

$$I_{13} = \frac{1}{\pi}\sum_{j=1}^{n}\int\limits_0^\infty \frac{1}{t}\left|f_j(t/B_n) - e^{-\sigma_j^2 t^2/(2B_n^2)}\right| e^{-t^2/2}\,dt$$

in lemma 4.3, inequality (2.5) given in lemma 2.8 is used. Taking into consideration that $\mathsf{E}|X_j|^4\mathbf{1}(|X_j| \leqslant U) \leqslant U\beta_{3,j}$, $j = 1,\ldots,n$, for any $U > 0$ we obtain

$$I_{13} \leqslant$$
$$\frac{1}{\pi}\sum_{j=1}^{n}\int\limits_0^\infty\!\left(\frac{t^2}{6B_n^3}\left(\left|\mathsf{E}X_j^3\mathbf{1}(|X_j|\leqslant U)\right| + \mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U)\right) + \frac{U\beta_{3,j}\,t^3}{24B_n^4} + \frac{\sigma_j^4 t^3}{8B_n^4}\right)e^{-t^2/2}dt$$
$$= \frac{1}{6\sqrt{2\pi}B_n^3}\sum_{j=1}^{n}\left(\left|\mathsf{E}X_j^3\mathbf{1}(|X_j|\leqslant U)\right| + \mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U)\right) + \frac{U\ell_n}{12\pi B_n} + \frac{1}{4\pi B_n^4}\sum_{j=1}^{n}\sigma_j^4,$$

so that

$$I_{13} + \frac{\nu_n\ell_n}{2\sqrt{2\pi}} = I_{13} + \frac{\ell_n}{2\sqrt{2\pi}} + \frac{1}{2\sqrt{2\pi}B_n^3}\sum_{j=1}^{n}\beta_{1,j}\,\sigma_j^2 \leqslant \frac{I_{14}}{\sqrt{2\pi}} + \frac{\ell_n}{2\sqrt{2\pi}},$$

where

$$I_{14} = \frac{1}{B_n^3}\sum_{j=1}^{n}\left(\frac{1}{6}\left|\mathsf{E}X_j^3\mathbf{1}(|X_j| \leqslant U)\right| + \frac{1}{2}\mathsf{E}|X_j|\mathsf{E}X_j^2\right) + \frac{U\ell_n}{6\sqrt{2\pi}B_n} + \frac{1}{2\sqrt{2\pi}B_n^4}\sum_{j=1}^{n}\sigma_j^4$$
$$+ \frac{1}{6B_n^3}\sum_{j=1}^{n}\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U), \quad U > 0.$$

The quantity I_{14} will be estimated in two steps.

1. Truncation. Denote $Y_j = X_j\mathbf{1}(|X_j| \leqslant U)$, $j = 1,\ldots,n$, $U > 0$. Then $X_j^k = Y_j^k + X_j^k\mathbf{1}(|X_j| \geqslant U)$ almost surely, $\mathsf{E}|Y_j|^k \leqslant \mathsf{E}|X_j|^k$, $k = 1, 2, 3$, and for all $j = 1,\ldots,n$

$$\mathsf{E}|X_j|\mathsf{E}X_j^2 \leqslant \mathsf{E}|Y_j|\mathsf{E}Y_j^2 + \mathsf{E}|Y_j|\mathsf{E}X_j^2\mathbf{1}(|X_j| > U) + \mathsf{E}|X_j|\mathbf{1}(|X_j| > U)\mathsf{E}X_j^2$$
$$\leqslant \mathsf{E}|Y_j|\mathsf{E}Y_j^2 + U^{-1}\left(\mathsf{E}|Y_j|^3\right)^{1/3}\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U)$$
$$+ U^{-2}\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U)\left(\mathsf{E}|X_j|^3\right)^{2/3} \leqslant \mathsf{E}|Y_j|\mathsf{E}Y_j^2 + \beta_{3,j}^{4/3}/U + \beta_{3,j}^{5/3}/U^2,$$

whence with the account of the relation

$$\sum_{j=1}^{n}\beta_{3,j}^r \leqslant \left(\sum_{j=1}^{n}\beta_{3,j}\right)^r = (B_n^3\ell_n)^r, \quad r \geqslant 1,$$

(see (4.2)) in the general case we obtain

$$\sum_{j=1}^{n} \mathsf{E}|X_j|\mathsf{E}X_j^2 \leqslant \sum_{j=1}^{n} \mathsf{E}|Y_j|\mathsf{E}Y_j^2 + \frac{B_n^4\ell_n^{4/3}}{U} + \frac{B_n^5\ell_n^{5/3}}{U^2}.$$

And if $X_1, \ldots, X_n$ are identically distributed, then $\beta_{3,j} = \beta_3 = B_n^3\ell_n/n$, $B_n = \sigma\sqrt{n}$ and hence,

$$\sum_{j=1}^{n} \mathsf{E}|X_j|\mathsf{E}X_j^2 \leqslant \sum_{j=1}^{n} \mathsf{E}|Y_j|\mathsf{E}Y_j^2 + \frac{B_n^4\ell_n^{4/3}}{Un^{1/3}} + \frac{B_n^5\ell_n^{5/3}}{U^2n^{2/3}}.$$

So, by the Lyapounov inequality and (4.2), for I_{14} we obtain

$$I_{14} \leqslant I_{15} + \frac{B_n\ell_n^{4/3}}{2U} + \frac{B_n^2\ell_n^{5/3}}{2U^2} + \frac{U\ell_n}{6\sqrt{2\pi}B_n} + \frac{\ell_n^{4/3}}{2\sqrt{2\pi}}, \quad U > 0,$$

in the general case and

$$I_{14} \leqslant I_{15} + \frac{B_n\ell_n^{4/3}}{2Un^{1/3}} + \frac{B_n^2\ell_n^{5/3}}{2U^2n^{2/3}} + \frac{U\ell_n}{6\sqrt{2\pi}B_n} + \frac{1}{2\sqrt{2\pi}n}, \quad U > 0,$$

in the case of identically distributed summands, where

$$I_{15} = \frac{1}{6B_n^3}\sum_{j=1}^{n} \left|\mathsf{E}Y_j^3\right| + \frac{1}{2B_n^3}\sum_{j=1}^{n} \mathsf{E}|Y_j|\mathsf{E}Y_j^2 + \frac{1}{6B_n^3}\sum_{j=1}^{n} \mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U).$$

Now choose the parameter U for the reason of equality of the orders of the "worst" terms in the obtained estimates for I_{14}, that is, so that for some free parameter $u > 0$ in the general case $U\ell_n/B_n = u^2B_n\ell_n^{4/3}/U$, and hence, $U = uB_n\ell_n^{1/6}$, and in the case of identically distributed summands $U\ell_n/B_n = u^2B_n\ell_n^{4/3}/(Un^{1/3})$, and hence, $U = uB_n(\ell_n/n)^{1/6}$, the parameter u being evaluated later. Then we obtain the estimates: in the general case

$$I_{14} \leqslant I_{15} + \ell_n^{7/6}\left(\frac{1}{2u} + \frac{u}{6\sqrt{2\pi}} + \frac{\ell_n^{1/6}}{2}\left(\frac{1}{u^2} + \frac{1}{\sqrt{2\pi}}\right)\right), \quad u > 0,$$

and, since

$$\frac{1}{n} \leqslant \frac{\ell_n^{4/3}}{n^{1/3}} = \frac{\ell_n^{7/6}}{n^{1/6}} \cdot \left(\frac{\ell_n}{n}\right)^{1/6},$$

in the case of identically distributed summands

$$I_{14} \leqslant I_{15} + \frac{\ell_n^{7/6}}{n^{1/6}}\left(\frac{1}{2u} + \frac{u}{6\sqrt{2\pi}} + \frac{1}{2}\Big(\frac{\ell_n}{n}\Big)^{1/6}\left(\frac{1}{u^2} + \frac{1}{\sqrt{2\pi}}\right)\right), \quad u > 0.$$

2. Centering. Since $\mathsf{E}X_j = 0$ for all $1 \leqslant j \leqslant n$, we have

$$|\mathsf{E}Y_j| = |\mathsf{E}X_j\mathbf{1}(|X_j| > U)| \leqslant \mathsf{E}|X_j|\mathbf{1}(|X_j| > U)$$

$$\leqslant U^{-2}\mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U)\leqslant U^{-2}\beta_{3,j}, \tag{4.11}$$

and hence,

$$|\mathsf{E}Y_j^3-\mathsf{E}(Y_j-\mathsf{E}Y_j)^3|=|3\mathsf{D}Y_j\mathsf{E}Y_j+(\mathsf{E}Y_j)^3|$$
$$\leqslant 3\mathsf{E}Y_j^2|\mathsf{E}Y_j|+|\mathsf{E}Y_j|^3\leqslant 3\beta_{3,j}^{5/3}/U^2+\beta_{3,j}^3/U^6,$$

whence for U chosen above, with the account of (4.2), we obtain

$$\sum_{j=1}^n\left|\mathsf{E}Y_j^3\right|\leqslant\sum_{j=1}^n\left|\mathsf{E}(Y_j-\mathsf{E}Y_j)^3\right|+\frac{3B_n^3\ell_n^{4/3}}{u^2}+\frac{B_n^3\ell_n^2}{u^6},\quad u>0,$$

in the general case and

$$\sum_{j=1}^n\left|\mathsf{E}Y_j^3\right|\leqslant\sum_{j=1}^n\left|\mathsf{E}(Y_j-\mathsf{E}Y_j)^3\right|+\frac{3B_n^3\ell_n^{4/3}}{u^2n^{1/3}}+\frac{B_n^3\ell_n^2}{u^6n},\quad u>0,$$

in the case of identically distributed summands. Similarly, for the terms of the second group in I_{15} we obtain

$$\mathsf{E}|Y_j|\mathsf{E}Y_j^2=\mathsf{E}|Y_j|\mathsf{D}Y_j+\mathsf{E}|Y_j|(\mathsf{E}Y_j)^2\leqslant\mathsf{E}|Y_j-\mathsf{E}Y_j|\mathsf{D}Y_j+|\mathsf{E}Y_j|\mathsf{D}Y_j+\mathsf{E}|Y_j|(\mathsf{E}Y_j)^2$$
$$\leqslant\mathsf{E}|Y_j-\mathsf{E}Y_j|\mathsf{D}Y_j+\left(U^{-2}(\mathsf{E}|Y_j|^3)^{2/3}+U^{-4}\beta_{3,j}(\mathsf{E}|Y_j|^3)^{1/3}\right)\mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U)$$
$$\leqslant\mathsf{E}|Y_j-\mathsf{E}Y_j|\mathsf{D}Y_j+\beta_{3,j}^{5/3}/U^2+\beta_{3,j}^{7/3}/U^4,$$

so that

$$\sum_{j=1}^n\mathsf{E}|Y_j|\mathsf{E}Y_j^2\leqslant\sum_{j=1}^n\mathsf{E}|Y_j-\mathsf{E}Y_j|\mathsf{D}Y_j+\frac{B_n^3\ell_n^{4/3}}{u^2}+\frac{B_n^3\ell_n^{5/3}}{u^4},\quad u>0,$$

in the general case and

$$\sum_{j=1}^n\mathsf{E}|Y_j|\mathsf{E}Y_j^2\leqslant\sum_{j=1}^n\mathsf{E}|Y_j-\mathsf{E}Y_j|\mathsf{D}Y_j+\frac{B_n^3\ell_n^{4/3}}{u^2n^{1/3}}+\frac{B_n^3\ell_n^{5/3}}{u^4n^{2/3}},\quad u>0,$$

in the case of identically distributed summands. With the parameter U specified above, denote

$$I_{16}=\frac{1}{6B_n^3}\sum_{j=1}^n\left|\mathsf{E}(Y_j-\mathsf{E}Y_j)^3\right|+\frac{1}{2B_n^3}\sum_{j=1}^n\mathsf{E}|Y_j-\mathsf{E}Y_j|\mathsf{D}Y_j+\frac{1}{6B_n^3}\sum_{j=1}^n\mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U).$$

Then we have: in the general case

$$I_{15}=\frac{1}{6B_n^3}\sum_{j=1}^n\left|\mathsf{E}Y_j^3\right|+\frac{1}{2B_n^3}\sum_{j=1}^n\mathsf{E}|Y_j|\mathsf{E}Y_j^2$$

$$\leqslant I_{16}+\frac{\ell_n^{4/3}}{u^2}+\frac{\ell_n^{5/3}}{2u^4}+\frac{\ell_n^2}{6u^6}=I_{16}+\ell_n^{7/6}\left(\frac{\ell_n^{1/6}}{u^2}+\frac{\ell_n^{1/2}}{2u^4}+\frac{\ell_n^{5/6}}{6u^6}\right),\quad u>0,$$

and in the case of identically distributed summands

$$I_{15}\leqslant I_{16}+\frac{\ell_n^{7/6}}{n^{1/6}}\left(\frac{1}{u^2}\Big(\frac{\ell_n}{n}\Big)^{1/6}+\frac{1}{2u^4}\Big(\frac{\ell_n}{n}\Big)^{1/2}+\frac{1}{6u^6}\Big(\frac{\ell_n}{n}\Big)^{5/6}\right),\quad u>0.$$

The application of the moment inequality of theorem 3.1 to the r.v.'s Y_j leads to the estimate

$$I_{16}\leqslant\frac{\lambda}{6B_n^3}\sum_{j=1}^n\mathsf{E}|Y_j-\mathsf{E}Y_j|^3+\frac{M(p(\lambda),\lambda)}{6B_n^3}\sum_{j=1}^n(\mathsf{D}Y_j)^{3/2}+\frac{1}{6B_n^3}\sum_{j=1}^n\mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U)$$

$$=\frac{\lambda\ell_n}{6}+\frac{M(p(\lambda),\lambda)}{6B_n^3}\sum_{j=1}^n\sigma_j^3+I_{17}-I_{18},\quad\lambda\geqslant1,$$

where

$$I_{17}=\frac{\lambda}{6B_n^3}\sum_{j=1}^n(\mathsf{E}|Y_j-\mathsf{E}Y_j|^3-\mathsf{E}|Y_j|^3)-\frac{M(p(\lambda),\lambda)}{6B_n^3}\sum_{j=1}^n\left(\sigma_j^3-(\mathsf{D}Y_j)^{3/2}\right),\quad(4.12)$$

$$I_{18}=\frac{\lambda-1}{6B_n^3}\sum_{j=1}^n\mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U),\quad(4.13)$$

with $p(\lambda)$ and $M(p,\lambda)$ defined in theorem 3.1. With the account of (4.11) we obtain

$$\mathsf{E}|Y_j-\mathsf{E}Y_j|^3-\mathsf{E}|Y_j|^3\leqslant3|\mathsf{E}Y_j|\mathsf{E}Y_j^2+|\mathsf{E}Y_j|^2\mathsf{E}|Y_j|$$
$$\leqslant\big(3\beta_{3,j}^{2/3}/U^2+\beta_{3,j}^{4/3}/U^4\big)\mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U).$$

By virtue of the inequality $(1-x)^\alpha\geqslant1-\alpha x$ which holds for all $0\leqslant x\leqslant1$, $\alpha\geqslant1$, we have

$$0\leqslant\sigma_j^3-(\mathsf{D}Y_j)^{3/2}=\sigma_j^3-\sigma_j^3\left(1-\frac{\mathsf{E}X_j^2\mathbf{1}(|X_j|>U)+(\mathsf{E}Y_j)^2}{\sigma_j^2}\right)^{3/2}$$

$$\leqslant\frac{3\sigma_j}{2}\left(\mathsf{E}X_j^2\mathbf{1}(|X_j|>U)+(\mathsf{E}Y_j)^2\right)\leqslant\frac{3}{2}\left(\beta_{3,j}^{1/3}/U+\beta_{3,j}^{4/3}/U^4\right)\mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U).$$

Noting that $M(p(\lambda),\lambda)\geqslant3-\lambda\geqslant1-\lambda$ for all $\lambda\geqslant1$ (see (3.2)) and using the estimates for the difference between the third moments and variances and denoting $b_j=\beta_{3,j}^{1/3}/U$ we obtain

$$I_{17}\leqslant\frac{\lambda}{6B_n^3}\sum_{j=1}^n\left(\mathsf{E}|Y_j-\mathsf{E}Y_j|^3-\mathsf{E}|Y_j|^3\right)+\frac{\lambda-1}{6B_n^3}\sum_{j=1}^n\left(\sigma_j^3-(\mathsf{D}Y_j)^{3/2}\right)$$

$$\leqslant \frac{1}{6B_n^3}\sum_{j=1}^n\Big(\lambda\left(3b_j^2+b_j^4\right)+\frac{3}{2}(\lambda-1)(b_j+b_j^4)\Big)\mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U)$$

$$=\frac{1}{6B_n^3}\sum_{j=1}^n\Big((\lambda-1)\Big(\frac{3}{2}\,b_j+3b_j^2+\frac{5}{2}\,b_j^4\Big)+3b_j^2+b_j^4\Big)\mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U),\quad(4.14)$$

so that

$$I_{17}-I_{18}\leqslant$$

$$\frac{\lambda-1}{6B_n^3}\sum_{j=1}^n\Big(\frac{3}{2}\,b_j+3b_j^2+\frac{5}{2}\,b_j^4-1\Big)\mathsf{E}|X_j|^3\mathbf{1}(|X_j|>U)+\frac{1}{6B_n^3}\sum_{j=1}^n\Big(\frac{3\beta_{3,j}^{5/3}}{U^2}+\frac{\beta_{3,j}^{7/3}}{U^4}\Big).$$

Let $b_0=0.36701\ldots$ be the unique root of the equation $1-\frac{3}{2}b-3b^2-\frac{5}{2}b^4=0$, $b>0$. Then we can guarantee that the first term in the estimate for $I_{17}-I_{18}$ is non-positive if $b_j\equiv\beta_{3,j}^{1/3}/U\leqslant b_0$, i.e. if $U\geqslant\beta_{3,j}^{1/3}/b_0$ for all $j=1,\ldots,n$. In the i.i.d. case, the condition $U\geqslant\beta_{3,j}^{1/3}/b_0$ is equivalent to $u\geqslant(\ell_n/n)^{1/6}/b_0$ and may be strengthened to $u\geqslant\ell_n^{1/2}/b_0$, since $n\geqslant 1/\ell_n^2$, while in the general case it follows from the condition $u\geqslant\ell_n^{1/6}/b_0$, since $\beta_{3,j}^{1/3}\leqslant B_n\ell_n^{1/3}$. Thus, we finally arrive at the estimate

$$I_{17}\leqslant\frac{\ell_n^{4/3}}{2u^2}+\frac{\ell_n^{5/3}}{6u^4}=\ell_n^{7/6}\Bigg(\frac{\ell_n^{1/6}}{2u^2}+\frac{\ell_n^{1/2}}{6u^4}\Bigg),\quad u\geqslant\ell_n^{1/6}/b_0=2.7246\ldots\ell_n^{1/6},$$

in the general case, and

$$I_{17}\leqslant\frac{\ell_n^{4/3}}{2u^2n^{1/3}}+\frac{\ell_n^{5/3}}{6u^4n^{2/3}}=\frac{\ell_n^{7/6}}{n^{1/6}}\bigg(\frac{1}{2u^2}\Big(\frac{\ell_n}{n}\Big)^{1/6}+\frac{1}{6u^4}\Big(\frac{\ell_n}{n}\Big)^{1/2},\quad u\geqslant\frac{1}{b_0}\Big(\frac{\ell_n}{n}\Big)^{1/6},$$

in the i.i.d. case.

Gathering the estimates for I_{14}, I_{15}, I_{16}, and I_{17}, in the general case for all $u\geqslant\ell_n^{1/6}/b_0$ we obtain

$$I_{14}\leqslant\frac{\lambda\ell_n}{6}+\frac{1}{6}M(p(\lambda),\lambda)\tau_n+\ell_n^{7/6}\Bigg(\frac{1}{2u}+\frac{u}{6\sqrt{2\pi}}+\frac{\ell_n^{1/6}}{2}\left(\frac{4}{u^2}+\frac{1}{\sqrt{2\pi}}\right)+\frac{2\ell_n^{1/2}}{3u^4}+\frac{\ell_n^{5/6}}{6u^6}\Bigg),$$

where

$$\tau_n=\frac{1}{B_n^3}\sum_{j=1}^n\sigma_j^3.$$

For $\ell>0$ denote

$$J_{13}(\ell)=\frac{1}{\sqrt{2\pi}}\inf\left\{\frac{1}{2u}+\frac{u}{6\sqrt{2\pi}}+\frac{\ell^{1/6}}{2}\left(\frac{4}{u^2}+\frac{1}{\sqrt{2\pi}}\right)+\frac{2\ell^{1/2}}{3u^4}+\frac{\ell^{5/6}}{6u^6}:u\geqslant\ell^{1/6}/b_0\right\}.$$

It is obvious that the function $J_{13}(\ell)$ monotonically and infinitely increases in $\ell > 0$ and

$$\lim_{\ell\to 0} J_{13}(\ell) = \frac{1}{\sqrt{2\pi}}\left(\frac{1}{2u} + \frac{u}{6\sqrt{2\pi}}\right)\bigg|_{u=\sqrt{3\sqrt{2\pi}}} = \frac{\sqrt{3}}{3(2\pi)^{3/4}} = 0.1454\ldots .$$

So, with the account of what has been said for arbitrary $\lambda \geqslant 1$ we obtain

$$I_{14} \leqslant \frac{\lambda\ell_n}{6} + \frac{1}{6}M(p(\lambda),\lambda)\tau_n + \begin{cases} \ell_n^{7/6}\cdot\sqrt{2\pi}J_{13}(\ell_n), & \text{in the general case,}\\ \ell_n^{7/6}n^{-1/6}\cdot\sqrt{2\pi}J_{13}(\ell_n/n), & \text{in the i.i.d. case,}\end{cases}$$

and, since $n \geqslant 1/\ell_n^2$,

$$I_{13}+\frac{\nu_n\ell_n}{2\sqrt{2\pi}} \leqslant \frac{I_{14}}{\sqrt{2\pi}}+\frac{\ell_n}{2\sqrt{2\pi}} \leqslant c\ell_n+K(c)\tau_n+\begin{cases} \ell_n^{7/6}\cdot J_{13}(\ell_n), & \text{in the general case,}\\ \ell_n^{3/2}\cdot J_{13}(\ell_n^3), & \text{in the i.i.d. case,}\end{cases}$$

where

$$c = \frac{\lambda+3}{6\sqrt{2\pi}} \geqslant \frac{2}{3\sqrt{2\pi}} = 0.2659\ldots, \quad K(c) = \frac{M(p(\lambda),\lambda)}{6\sqrt{2\pi}}\bigg|_{\lambda=6\sqrt{2\pi}c-3}.$$

So, from lemma 4.3 with the account of the estimates for $I_{13} + \nu_n\ell_n/(2\sqrt{2\pi})$ established above we finally obtain

$$\Delta_n \leqslant c\ell_n + K(c)\tau_n + R(\ell_n), \tag{4.15}$$

where $R(\ell) = \widetilde{C}(\ell)\cdot\ell^{7/6}$ in the general case and $R(\ell) = \widehat{C}(\ell)\cdot\ell^{3/2}$ in the i.i.d. case,

$$\widetilde{C}(\ell) = J_{13}(\ell)+\ell^{5/6}J_{12}(\ell,2)+\min_{t_3\leqslant t_0\leqslant t_1\wedge t_4}\Big\{\ell^{1/2}J_{11}(\ell,2,t_0)+\ell^{5/6}\pi^{-2}\max_{T\geqslant\pi/\ell}J(T,t_0)\Big\},$$

$$\widehat{C}(\ell) = \\ J_{13}\big(\ell^3\big) + \ell^{1/2}\Big(\widehat{J}_{12}(\ell,2,\ell^{-2/\delta})+\min_{t_3\leqslant t_0\leqslant t_1\wedge t_4}\Big\{\widehat{J}_{11}\big(\ell^{-2},2,t_0\big)+\pi^{-2}\max_{T\geqslant\pi/\ell}J(T,t_0)\Big\}\Big),$$

$t_3 = t_3(1) = 0.3566\ldots$, $t_1 = t_1(1) = 0.6359\ldots$, $t_4 = t_4(1,\ell) = (1-\ell^{2/3})/(4\pi\varkappa_1)$. Moreover, the functions $\widetilde{C}(\ell)$, $\widehat{C}(\ell)$ monotonically and infinitely increase on the intervals $0 < \ell < \overline{\ell}$ and $0 < \ell < (\overline{\ell})^{1/3}$ correspondingly, where $\overline{\ell} = \overline{\ell}(1) = \big(1-4/9e^{-5/6}\big)^{3/2} = 0.7247\ldots$.

Let us note that in (4.15) the "constants" $\widetilde{C}(\ell)$ and $\widehat{C}(\ell)$ in the remainder $R(\ell_n)$ do not depend on the choice of the coefficient c at the main term ℓ_n. But this "universality" contains a lack: the rate of decrease of the remainder $R(\ell_n)$ is too low than it could be for $c > 2/(3\sqrt{2\pi})$ if the remainder could depend on c. Indeed, in the final estimate for I_{17} we bounded $(1-\lambda)\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U)$ above by zero (recall that $c = (\lambda+3)/(6\sqrt{2\pi})$). The price for this operation is extremely high for $\lambda > 1$ (i.e. for $c > 2/(3\sqrt{2\pi})$), since the cubic tail $\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U)$ due to the truncation determines the rate of decrease of the remainder $R(\ell_n)$ in the final estimate (4.15), instead of being added to the main term and accurately estimated in a sum with $(1-\lambda)\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U)$. So, if this cubic tail is "transferred" to the main term, the remainder becomes better. Let us accomplish this transfer.

Estimating the integral I_{13} in the same way as above, for all $U \geqslant 0$ we obtain

$$I_{13} + \frac{\nu_n\ell_n}{2\sqrt{2\pi}} \leqslant \frac{U\ell_n}{12\pi B_n} + \frac{1}{4\pi B_n^4}\sum_{j=1}^n \sigma_j^4 + \frac{\ell_n}{2\sqrt{2\pi}} + \frac{I'_{14}}{\sqrt{2\pi}},$$

where

$$I'_{14} = \frac{1}{B_n^3}\sum_{j=1}^n\left(\frac{1}{6}\left|\mathsf{E}X_j^3\mathbf{1}(|X_j| \leqslant U)\right| + \frac{1}{2}\mathsf{E}|X_j|\mathsf{E}X_j^2 + \frac{1}{6}\,\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U)\right).$$

Truncating the moments $\mathsf{E}|X_j|$ and $\mathsf{E}X_j^2$ in the same way as in the integral I_{14} we obtain (recall that $Y_j = X_j\mathbf{1}(|X_j| \leqslant U)$)

$$\mathsf{E}|X_j|\mathsf{E}X_j^2 \leqslant \mathsf{E}|Y_j|\mathsf{E}Y_j^2 + \left(\beta_{3,j}^{1/3}/U + \beta_{3,j}^{2/3}/U^2\right)\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U), \quad j = 1,\ldots,n,$$

while the centering with the account of (4.11) leads to the estimates

$$\mathsf{E}|Y_j|\mathsf{E}Y_j^2 \leqslant \mathsf{E}|Y_j - \mathsf{E}Y_j|\mathsf{D}Y_j + \left(\beta_{3,j}^{2/3}/U^2 + \beta_{3,j}^{4/3}/U^4\right)\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U),$$

$$|\mathsf{E}Y_j^3| \leqslant |\mathsf{E}(Y_j - \mathsf{E}Y_j)^3| + \left(3\beta_{3,j}^{2/3}/U^2 + \beta_{3,j}^2/U^6\right)\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U), \quad j = 1,\ldots,n.$$

Summarizing the above estimates we obtain

$$I'_{14} \leqslant \frac{1}{B_n^3}\sum_{j=1}^n\left(\frac{1}{6}\left|\mathsf{E}(Y_j - \mathsf{E}Y_j)^3\right| + \frac{1}{2}\mathsf{E}|Y_j - \mathsf{E}Y_j|\mathsf{D}Y_j\right) + I_{19},$$

where

$$I_{19} = \frac{1}{6B_n^3}\sum_{j=1}^n\left(1 + 3b_j + 9b_j^2 + 3b_j^4 + b_j^6\right)\mathsf{E}|X_j|^3\mathbf{1}(|X_j| > U), \quad b_j = \beta_{3,j}^{1/3}/U.$$

Applying the moment inequality from theorem 3.1 to the r.v.'s Y_j we obtain

$$\begin{aligned} I'_{14} &\leqslant \frac{\lambda}{6B_n^3}\sum_{j=1}^n \mathsf{E}|Y_j - \mathsf{E}Y_j|^3 + \frac{M(p(\lambda),\lambda)}{6B_n^3}\sum_{j=1}^n(\mathsf{D}Y_j)^{3/2} + I_{19} \\ &= \frac{\lambda\ell_n}{6} + \frac{M(p(\lambda),\lambda)}{6B_n^3}\sum_{j=1}^n\sigma_j^3 + I_{17} - \frac{\lambda}{\lambda-1}I_{18} + I_{19}, \end{aligned}$$

with I_{17}, I_{18} defined in (4.12), (4.13) correspondingly. By virtue of the estimate (4.14) we have

$$I_{17} - \frac{\lambda}{\lambda-1} I_{18} + I_{19} \leqslant \frac{1}{6B_n^3} \sum_{j=1}^n \mathsf{E}|X_j|^3 \mathbf{1}(|X_j| > U) \times$$
$$\times \Big((1-\lambda)\Big(1 - \frac{3}{2}\, b_j - 3b_j^2 - \frac{5}{2}\, b_j^4\Big) + 3b_j + 12b_j^2 + 4b_j^4 + b_j^6 \Big).$$

As it was noticed above, $1-\frac{3}{2}b-3b^2-\frac{5}{2}b^4 > 0$ for $0 \leqslant b < b_0$, where $b_0 = 0.36701\ldots$ is the unique root of the equation $1-\frac{3}{2}b-3b^2-\frac{5}{2}b^4 = 0$, $b > 0$. Introduce the function

$$g(b) = \frac{3b + 12b^2 + 4b^4 + b^6}{1 - 3b/2 - 3b^2 - 5b^4/2}, \quad 0 \leqslant b < b_0.$$

Evidently, $g(b)$ increases monotonically varying within the limits

$$0 = \lim_{b\to 0} g(b) \leqslant g(b) < \lim_{b\to b_0} g(b) = +\infty, \quad 0 \leqslant b < b_0,$$

and therefore for each $\lambda > 1$ there exists a unique root of the equation $g(1/u) = \lambda - 1$ in the interval $u > 1/b_0 = 2.7246\ldots$. For $c = (\lambda+3)/(6\sqrt{2\pi}) \geqslant 2/(3\sqrt{2\pi})$ let u_c be the unique root of the equation $g(1/u) = 6\sqrt{2\pi}c - 4$, $u > 1/b_0$. It can easily be made sure that u_c decreases monotonically varying within the limits

$$2.7246\ldots = 1/b_0 = \lim_{c\to\infty} u_c \leqslant u_c \leqslant \lim_{c\to 2/(3\sqrt{2\pi})} u_c = +\infty, \quad c > 2/(3\sqrt{2\pi}).$$

If $b_j \equiv \beta_{3,j}^{1/3}/U \leqslant u_c^{-1}$, i.e. $U \geqslant u_c \beta_{3,j}^{1/3}$, for all $j = 1, \ldots, n$, then $I_{17} - \frac{\lambda}{\lambda-1} I_{18} + I_{19} \leqslant 0$, and thus we obtain the estimate

$$I_{13} + \frac{\nu_n \ell_n}{2\sqrt{2\pi}} \leqslant c\ell_n + K(c)\tau_n + R,$$

with c and $K(c)$ defined above, provided that $U \geqslant u_c \beta_{3,j}^{1/3}$ for all $j = 1, \ldots, n$, where

$$R = \frac{1}{4\pi B_n^4} \sum_{j=1}^n \sigma_j^4 + \frac{U\ell_n}{12\pi B_n} \leqslant \begin{cases} \dfrac{\ell_n}{12\pi}\big(3\ell_n^{1/3} + U/B_n\big), & \text{in the general case,} \\ \dfrac{\ell_n}{12\pi}\Big(3\Big(\dfrac{\ell_n}{n}\Big)^{1/3} + \dfrac{U}{B_n}\Big), & \text{in the i.i.d. case,} \end{cases}$$

Now choose the parameter U so that the orders of both terms in the above estimate for R coincide, i.e. let

$$U = \begin{cases} uB_n\ell_n^{1/3}, & \text{in the general case,} \\ uB_n(\ell_n/n)^{1/3}, & \text{in the i.i.d. case,} \end{cases}$$

$u \geqslant u_c$ being a free parameter. Then the condition $U \geqslant u_c \beta_{3,j}^{1/3}$ is satisfied for all $j = 1, \ldots, n$, since in the general case $\beta_{3,j}^{1/3}/U \leqslant B_n \ell_n^{1/3}/U = u^{-1} \leqslant u_c^{-1}$, as well as in the i.i.d. case $\beta_{3,j}^{1/3}/U = B_n(\ell_n/n)^{1/3}/U = u^{-1} \leqslant u_c^{-1}$ for all $j = 1, \ldots, n$. So, we have

$$I_{13} + \frac{\nu_n \ell_n}{2\sqrt{2\pi}} \leqslant c\ell_n + K(c)\tau_n + \begin{cases} \ell_n^{4/3}(3+u)/(12\pi), & \text{in the general case,} \\ \ell_n^{4/3} n^{-1/3}(3+u)/(12\pi), & \text{in the i.i.d. case,} \end{cases}$$

for arbitrary $u \geqslant u_c$. Evidently the value $u = u_c$ minimizes the right-hand side of the obtained estimate. Gathering the estimates from lemma 4.3 and theorem 4.6 we finally obtain

$$\Delta_n \leqslant c\ell_n + K(c)\tau_n + R(\ell_n, c),$$

with

$$R(\ell, c) = \begin{cases} \left(\frac{3+u_c}{12\pi} + \widetilde{C}_1(\ell)\ell^{1/3}\right)\ell^{4/3}, & \text{in the general case,} \\ \frac{(3+u_c)\ell^{4/3}}{12\pi n^{1/3}} + \widehat{C}_1(\ell)\ell^2 \leqslant \left(\frac{3+u_c}{12\pi} + \widehat{C}_1(\ell)\right)\ell^2, & \text{in the i.i.d. case,} \end{cases}$$

$\widetilde{C}_1(\ell)$, $\widehat{C}_1(\ell)$ defined in theorem 4.6.

As it follows from remark 3.2,

$$\inf_{c \geqslant 2/(3\sqrt{2\pi})} (c + K(c)) = \lim_{\lambda \to \infty} \frac{\lambda + 3 + M(p(\lambda), \lambda)}{6\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}},$$

and also $K(c) > 0$ if and only if $\lambda < \sqrt{10}$, that is, $c < (\sqrt{10} + 3)/(6\sqrt{2\pi}) = 0.4097\ldots$, whence it follows that for all c such that $2/(3\sqrt{2\pi}) \leqslant c \leqslant 1/\sqrt{2\pi} = 0.3989\ldots$ the estimates $K(c) > 0$ and

$$c\ell_n + K(c)\tau_n \geqslant c\ell_n \geqslant \frac{2\ell_n}{3\sqrt{2\pi}},$$

hold and for $c > 1/\sqrt{2\pi}$

$$\begin{aligned} c\ell_n + K(c)\tau_n &= \frac{\ell_n}{\sqrt{2\pi}} + \left(c - \frac{1}{\sqrt{2\pi}}\right)\ell_n + K(c)\tau_n \\ &\geqslant \frac{\ell_n}{\sqrt{2\pi}} + \left(c + K(c) - \frac{1}{\sqrt{2\pi}}\right)\tau_n \geqslant \frac{\ell_n}{\sqrt{2\pi}} \geqslant \frac{2\ell_n}{3\sqrt{2\pi}}, \end{aligned}$$

since $\ell_n \geqslant \tau_n$ by the Lyapounov inequality. So,

$$\inf_{c \geqslant 2/(3\sqrt{2\pi})} (c\ell_n + K(c)\tau_n) \geqslant \frac{2\ell_n}{3\sqrt{2\pi}}.$$

For the purpose of lowering the right bound of the interval of the values of ℓ under consideration and thus bound the range of the constants $\widetilde{C}(\ell)$, $\widehat{C}(\ell)$ above, note that if $\ell_n \geqslant \ell$ for some $\ell > 0$, then by virtue of (4.3) for any

$$A \geqslant \kappa - \frac{2\ell}{3\sqrt{2\pi}},$$

where $\kappa = 0.5409\ldots$ (see (4.3)), the trivial estimate

$$\inf_{c\geqslant 2/(3\sqrt{2\pi})}(c\ell_n + K(c)\tau_n) + A \geqslant \frac{2\ell}{3\sqrt{2\pi}} + A \geqslant \kappa \geqslant \Delta_n,$$

holds so that by virtue of the monotonicity of $R(\ell)$, in (4.15) for $\ell_n \geqslant \ell$ the quantity $R(\ell_n)$ can be replaced by

$$\min\left\{R(\ell),\, \kappa - \frac{2\ell}{3\sqrt{2\pi}}\right\} = R(\ell_R \wedge \ell),$$

where ℓ_R is the unique root of the equation

$$R(\ell) = \kappa - \frac{2\ell}{3\sqrt{2\pi}}$$

on the interval $(0,\,\overline{\ell}\,)$ in the general case and on the interval $\big(0, (\overline{\ell})^{1/3}\big)$ in the case of identically distributed summands. The existence of ℓ_R and its uniqueness follow from that on the interval under consideration the left-hand side of the equation is a continuous function which increases strictly monotonically and takes *all* values from 0 to $+\infty$, and the right-hand side is a continuous function which decreases strictly monotonically and takes positive values at small ℓ, that is, the graphs of these functions intersect in a single point. The same reasoning concerns $R(\ell_n, c)$ and is summarized in the following two theorems.

Theorem 4.13. *For any $\ell > 0$, for all $n \geqslant 1$ and $F_1, \ldots, F_n \in \mathcal{F}_3$ such that $\ell_n \leqslant \ell$ there hold the estimates:*

$$\Delta_n \leqslant \inf_{c\geqslant 2/(3\sqrt{2\pi})}\left\{c\ell_n + \frac{K(c)}{B_n^3}\sum_{j=1}^n \sigma_j^3\right\} + \widetilde{C}\big(\widetilde{\ell}\wedge\ell\big)\ell_n^{7/6}, \tag{4.16}$$

in the general case and

$$\Delta_n \leqslant \inf_{c\geqslant 2/(3\sqrt{2\pi})}\left\{c\ell_n + \frac{K(c)}{\sqrt{n}}\right\} + \widehat{C}\big(\widehat{\ell}\wedge\ell\big)\ell_n^{3/2}, \tag{4.17}$$

in the case of identically distributed summands, where

$$K(c) = \frac{M(p(\lambda),\lambda)}{6\sqrt{2\pi}}\bigg|_{\lambda = 6\sqrt{2\pi}c-3},$$

$$M(p,\lambda) = \frac{1-\lambda+2(\lambda+2)p-2(\lambda+3)p^2}{\sqrt{p(1-p)}}, \quad 0 < p \leqslant \frac{1}{2},\ \lambda \geqslant 1,$$

$$p(\lambda) = \frac{1}{2} - \sqrt{\frac{\lambda+1}{\lambda+3}}\sin\left(\frac{\pi}{6} - \frac{1}{3}\arctan\sqrt{\lambda^2 + 2\frac{\lambda-1}{\lambda+3}}\right), \quad \lambda \geqslant 1;$$

$$\widetilde{C}(\ell) = J_{13}(\ell) + \ell^{5/6}J_{12}(\ell,2) + \min_{t_3\leqslant t_0\leqslant t_1\wedge t_4(\ell)}\left\{\ell^{1/2}J_{11}(\ell,2,t_0) + \ell^{5/6}\pi^{-2}\max_{T\geqslant\pi/\ell}J(T,t_0)\right\};$$

$$\widehat{C}(\ell) = J_{13}(\ell^3)+\ell^{1/2}\Big(\widehat{J}_{12}(\ell,2,\ell^{-2})+\min_{t_3\leqslant t_0\leqslant t_1\wedge t_4(\ell^3)}\Big\{\widehat{J}_{11}(\ell^{-2},2,t_0)+\pi^{-2}\max_{T\geqslant\pi/\ell}J(T,t_0)\Big\}\Big);$$

$\widetilde{\ell} = 0.226547\ldots$, $\widehat{\ell} = 0.402361\ldots$ *are respectively the unique roots of the equations*

$$\widetilde{C}(\ell)\cdot\ell^{7/6} = \kappa - \frac{2\ell}{3\sqrt{2\pi}},\quad 0<\ell<\overline{\ell} = \Big(1-4/9e^{-5/6}\Big)^{3/2} = 0.7247\ldots,$$

$$\widehat{C}(\ell)\cdot\ell^{3/2} = \kappa - \frac{2\ell}{3\sqrt{2\pi}},\quad 0<\ell<(\overline{\ell}\,)^{1/3} = \sqrt{1-4/9e^{-5/6}} = 0.8982\ldots,$$

on the intervals specified above; $\kappa = 0.5409\ldots$ *is defined in* (4.3)*;*

$$J_{13}(\ell) = \frac{1}{\sqrt{2\pi}}\inf\Big\{\frac{1}{2u}+\frac{u}{6\sqrt{2\pi}}+\frac{\ell^{1/6}}{2}\Big(\frac{4}{u^2}+\frac{1}{\sqrt{2\pi}}\Big)+\frac{2\ell^{1/2}}{3u^4}+\frac{\ell^{5/6}}{6u^6} : u\geqslant u_0\ell^{1/6}\Big\},$$

$$t_3 = \frac{e^{-5/6}}{9\pi\varkappa_1} = 0.1550\ldots,\quad t_1 = \frac{\theta_0(1)}{2\pi} = 0.6359\ldots,\quad t_4(\ell) = \frac{1-\ell^{2/3}}{4\pi\varkappa_1},$$

$u_0 = 2.7246\ldots$ *is the unique root of the equation* $1-3/2u^{-1}-3u^{-2}-5/2u^{-4} = 0$, $u>0$; $J_{11}(\ell,\nu,t_0)$, $J_{12}(\ell,\nu)$, $\widehat{J}_{11}(n,\nu,t_0)$, $\widehat{J}_{12}(\ell,\nu,n)$, $J(T,t_0)$, $\theta_0(1)$, $\varkappa_1 = 0.0991\ldots$ *are defined in lemma* 4.3*. In particular,*

$$\widetilde{C}(0.226548)\leqslant 2.7176\ (\textit{with } t_0 = t_3 = 0.1550\ldots,\ u = 4.3173\ldots),$$

$$\widehat{C}(0.402362)\leqslant 1.7002\ (\textit{with } t_0 = 0.1802\ldots,\ u = 4.1157\ldots),$$

$$\widetilde{C}(0+) = \widehat{C}(0+) = \frac{\sqrt{3}}{3(2\pi)^{3/4}} = 0.1454\ldots\ (\textit{with } t_0 = t_3,\ u = \sqrt{3\sqrt{2\pi}} = 2.7422\ldots).$$

The values of $\widehat{C}(\ell)$, $\widetilde{C}(\ell)$ *for other* ℓ *are given in table* 9*, the functions* $\widetilde{C}(\ell)$, $\widehat{C}(\ell)$ *being monotonically increasing.*

ℓ	0.1	0.01	10^{-3}	10^{-4}	10^{-7}	10^{-20}
$\widetilde{u}(\ell) =$	4.1825	3.8521	3.5852	3.3724	2.9823	2.7440
$\widehat{u}(\ell) =$	3.5852	3.0782	2.8609	2.7813	2.7435	2.7422
$\widetilde{C}(\ell)\leqslant$	0.7802	0.2792	0.2110	0.1854	0.1577	0.1456
$\widehat{C}(\ell)\leqslant$	0.3861	0.2169	0.1682	0.1527	0.1458	0.1455

Table 9: The values of $\widetilde{C}(\ell)$, $\widehat{C}(\ell)$ from theorem 4.13 for some ℓ together with the optimal values of u from J_{13} which are denoted by $\widetilde{u}(\ell)$ for $\widetilde{C}(\ell)$ and $\widehat{u}(\ell)$ for $\widehat{C}(\ell)$. The optimal values of t_0 coincide with $t_3 = 0.1550\ldots$ in both cases.

Theorem 4.14. *For any $\ell > 0$, for all $n \geqslant 1$ and $F_1, \ldots, F_n \in \mathcal{F}_3$ such that $\ell_n \leqslant \ell$ there hold the estimates:*

$$\Delta_n \leqslant \inf_{c>2/(3\sqrt{2\pi})} \left\{ c\ell_n + \frac{K(c)}{B_n^3} \sum_{j=1}^n \sigma_j^3 + \widetilde{A}_c\big(\ell \wedge \widetilde{\ell}_c\big)\ell_n^{4/3} \right\}, \tag{4.18}$$

in the general case and

$$\Delta_n \leqslant \inf_{c>2/(3\sqrt{2\pi})} \left\{ c\ell_n + \frac{K(c)}{\sqrt{n}} + \widehat{A}_c\big(\ell \wedge \widehat{\ell}_c\big)\ell_n^2 \right\}, \tag{4.19}$$

in the case of identically distributed summands, where $K(c)$ is defined in theorem 4.13;

$$\widetilde{A}_c(\ell) = \frac{3+u_c}{12\pi} + \widetilde{C}_1(\ell) \cdot \ell^{1/3}, \quad \widehat{A}_c(\ell) = \frac{3+u_c}{12\pi} + \widehat{C}_1(\ell),$$

$\widetilde{C}_1(\ell)$, $\widehat{C}_1(\ell)$ *are defined in theorem* 4.6, u_c *is the unique root of the equation*

$$\frac{3u^{-1} + 12u^{-2} + 4u^{-4} + u^{-6}}{1 - 3/2u^{-1} - 3u^{-2} - 5/2u^{-4}} = 6\sqrt{2\pi}c - 4, \quad c > \frac{2}{3\sqrt{2\pi}},$$

in the interval $u > u_\infty$ with $u_\infty = 2.7246\ldots$ being the unique root of the equation $1 - 3/2u^{-1} - 3u^{-2} - 5/2u^{-4} = 0$, $u > 0$; $\widetilde{\ell}_c$, $\widehat{\ell}_c$ are respectively the unique roots of the equations

$$\widetilde{A}_c\big(\ell\big) \cdot \ell^{4/3} = \kappa - \frac{2\ell}{3\sqrt{2\pi}}, \quad 0 < \ell < \overline{\ell} = \Big(1 - 4/9e^{-5/6}\Big)^{3/2} = 0.7247\ldots,$$

$$\widehat{A}_c\big(\ell\big) \cdot \ell^2 = \kappa - \frac{2\ell}{3\sqrt{2\pi}}, \quad 0 < \ell < (\overline{\ell}\,)^{1/3} = \sqrt{1 - 4/9e^{-5/6}} = 0.8982\ldots,$$

on the intervals specified above; $\kappa = 0.5409\ldots$ is defined in (4.3). *The functions $\widetilde{A}_c\big(\ell\big)$, $\widehat{A}_c\big(\ell\big)$ increase monotonically in $\ell > 0$ and decrease monotonically in c, moreover*

$$\lim_{c\to 2/(3\sqrt{2\pi})} \inf_{\ell>0} \widetilde{A}_c(\ell) = \lim_{c\to 2/(3\sqrt{2\pi})} \inf_{\ell>0} \widehat{A}_c(\ell) = +\infty.$$

$$\lim_{c\to\infty} \widetilde{A}_c(\ell) - \widetilde{C}_1(\ell) \cdot \ell^{1/3} = \lim_{c\to\infty} \widehat{A}_c(\ell) - \widehat{C}_1(\ell) = \frac{3+u_\infty}{12\pi} = 0.1518\ldots, \quad \ell > 0,$$

$$\lim_{\ell\to 0} \widetilde{A}_c(\ell) = \frac{3+u_c}{12\pi},$$

$$\lim_{\ell\to 0} \widehat{A}_c(\ell) - \frac{3+u_c}{12\pi} = \widehat{C}_1(0) = \frac{1.0253 \cdot 5\varkappa_1}{\sqrt{2\pi}(1 - 4/9e^{-5/6})^{7/2}} + \frac{1}{3\pi} = 0.5359\ldots .$$

The values of $\widetilde{A}_c\big(\ell\big)$, $\widehat{A}_c\big(\ell\big)$ for some ℓ and c are given in table 10.

c	0.27	0.28	0.29	0.30	$\frac{\sqrt{10}+3}{6\sqrt{2\pi}}$	∞
$K(c)$	0.1521	0.1402	0.1287	0.1174	0.0000	$-\infty$
u_c	54.5687	18.8812	12.6629	10.0115	4.7345	2.7247
$\widetilde{\ell}_c$	0.2048	0.2220	0.2250	0.2263	0.2288	0.2298
$\widetilde{A}_c(\widetilde{\ell}_c)$	4.0313	3.5851	3.5160	3.4872	3.4314	3.4106
$\widetilde{A}_c(0.01)$	1.6632	0.7165	0.5516	0.4813	0.3413	0.2880
$\widetilde{A}_c(10^{-3})$	1.5757	0.6290	0.4641	0.3937	0.2538	0.2005
$\widetilde{A}_c(0+)$	1.5271	0.5805	0.4155	0.3452	0.2052	0.1519
$\widehat{\ell}_c$	0.3596	0.3942	0.4008	0.4036	0.4094	0.4116
$\widehat{A}_c(\widehat{\ell}_c)$	3.4449	2.8068	2.7046	2.6619	2.5786	2.5475
$\widehat{A}_c(0.07)$	2.1042	1.1576	0.9927	0.9223	0.7823	0.7290
$\widehat{A}_c(0.05)$	2.0902	1.1435	0.9786	0.9083	0.7683	0.7150
$\widehat{A}_c(0.03)$	2.0780	1.1314	0.9664	0.8961	0.7561	0.7028
$\widehat{A}_c(0.01)$	2.0674	1.1207	0.9558	0.8855	0.7455	0.6922
$\widehat{A}_c(0+)$	2.0631	1.1164	0.9515	0.8811	0.7412	0.6878

Table 10: Upper bounds of $K(c)$, u_c, $\widetilde{\ell}_c$, $\widetilde{A}_c(\ell)$, $\widehat{\ell}_c$, $\widehat{A}_c(\ell)$ from Theorem 4.14 for some ℓ and c.

Remark 4.15. Taking into account the properties of the functions $M(p(\lambda),\lambda)$ and $\lambda+M(p(\lambda),\lambda)$, $\lambda\geqslant 1$ described in remark 3.2 it can be made sure that the functions $K(c)$ and $c+K(c)$ decrease monotonically for all $c\geqslant 2/(3\sqrt{2\pi})$ varying within the limits

$$-\infty=\lim_{c\to\infty}K(c)<K(c)\leqslant K\left(\frac{2}{3\sqrt{2\pi}}\right)=\sqrt{\frac{2\sqrt{3}-3}{6\pi}}=0.1569\ldots,$$

$$0.3989\ldots=\frac{1}{\sqrt{2\pi}}=\lim_{c\to\infty}(c+K(c))<c+K(c)\leqslant\frac{2}{3\sqrt{2\pi}}+\sqrt{\frac{2\sqrt{3}-3}{6\pi}}=0.4228\ldots,$$

moreover the function $K(c)$ changes its sign at the unique point $c=(\sqrt{10}+3)/(6\sqrt{2\pi})=0.4097\ldots$.

Remark 4.16. As it was shown in [27, 29], the least possible value of the coefficient c at ℓ_n in estimates (4.16), (4.17) cannot be made less than $\underline{C_{\rm AE}}=2/(3\sqrt{2\pi})$. Furthermore, the estimates obtained in theorem 4.13 for each $c\geqslant\underline{C_{\rm AE}}$ are optimal in the sense that the value of the coefficient $K(c)$ cannot be made less. Indeed, even in the case of identically distributed summands, for all $c\geqslant\underline{C_{\rm AE}}$, obviously, $K(c)$ can be estimated as

$$K(c)\geqslant\sup_{X_1\in\mathcal{F}_3}\limsup_{n\to\infty}\frac{\sqrt{n}\Delta_n(\mathsf{E}X_1^2)^{3/2}-c\mathsf{E}|X_1|^3}{(\mathsf{E}X_1^2)^{3/2}}.$$

On the other hand, with the account of (1.4) we obtain

$$K(c) \geqslant \sup_{h>0} \sup_{X\in\mathcal{F}_3^h} \frac{|\mathsf{E}X^3| + 3h\mathsf{E}X^2 - 6\sqrt{2\pi}c\mathsf{E}|X|^3}{6\sqrt{2\pi}\,(\mathsf{E}X^2)^{3/2}}, \quad c \geqslant \underline{C_{\text{AE}}}.$$

Now letting $\mathsf{P}(X = -\sqrt{p/q}) = q$, $\mathsf{P}(X = \sqrt{q/p}) = p = 1-q$, $0 < p \leqslant 1/2$, we arrive at

$$\mathsf{E}X = 0, \quad \mathsf{E}X^2 = 1, \quad \mathsf{E}X^3 = \frac{q-p}{\sqrt{pq}}, \quad \mathsf{E}|X|^3 = \frac{p^2+q^2}{\sqrt{pq}}, \quad h = \frac{1}{\sqrt{pq}},$$

and hence, for all $c \geqslant \underline{C_{\text{AE}}}$

$$\begin{aligned} K(c) &\geqslant \frac{1}{6\sqrt{2\pi}} \sup\left\{\frac{q-p+3-6\sqrt{2\pi}c(p^2+q^2)}{\sqrt{pq}} : 0 < p \leqslant 1/2,\ q = 1-p\right\} \\ &= \frac{M(p(\lambda),\lambda)}{6\sqrt{2\pi}}\bigg|_{\lambda=6\sqrt{2\pi}c-3} \end{aligned}$$

by virtue of representation (3.1), which coincides with the definition of $K(c)$ (see theorem 4.13).

From theorems 4.13 and 4.14 with concrete c we can obtain some corollaries. For example, with $c = \underline{C_{\text{AE}}} = 2/(3\sqrt{2\pi})$ we have $K(c) = \sqrt{(2\sqrt{3}-3)/(6\pi)}$, and hence, theorem 4.13 implies

Corollary 4.17. *For all $n \geqslant 1$ and $F_1, \ldots, F_n \in \mathcal{F}_3$ there hold the estimates*

$$\Delta_n \leqslant \frac{2\ell_n}{3\sqrt{2\pi}} + \sqrt{\frac{2\sqrt{3}-3}{6\pi}\sum_{j=1}^n \sigma_j^3/B_n^3 + 2.7176\cdot\ell_n^{7/6}}$$

in the general case and

$$\Delta_n \leqslant \frac{2\ell_n}{3\sqrt{2\pi}} + \sqrt{\frac{2\sqrt{3}-3}{6\pi n} + 1.7002\cdot\ell_n^{3/2}}$$

in the case $F_1 = \ldots = F_n$, moreover, in each of the estimates the constant $\sqrt{(2\sqrt{3}-3)/(6\pi)} = 0.1569\ldots$ at the second term cannot be made less under the condition that the coefficient at the first term is fixed and equals $2/(3\sqrt{2\pi})$.

Corollary 4.17 sharpens the inequalities of Prawitz (1.9) and Bentkus (1.10) with respect to the second term by virtue of the smaller value of the constant

$$\sqrt{(2\sqrt{3}-3)/(6\pi)} = 0.1569\ldots$$

as compared to $\left(2\sqrt{2\pi}\right)^{-1} = 0.1994\ldots$ in (1.9), (1.10), but the "expense" of using the unimprovable constant at the second term is a worse order of decrease of the remainder, namely, $O\big(\ell_n^{3/2}\big)$ and $O\big(\ell_n^{7/6}\big)$ as compared with $O\big(\ell_n^{2}\big)$ in (1.9) and $O\big(\ell_n^{4/3}\big)$ in (1.10) respectively. However, here we specify concrete values of the constants.

With $c = C_{\text{AE}} = (\sqrt{10}+3)/(6\sqrt{2\pi})$ we have $K(c) = 0$, and hence, theorem 4.14 implies

Corollary 4.18. *For all $n \geqslant 1$ and $F_1, \ldots, F_n \in \mathcal{F}_3$ there hold the estimates*

$$\begin{aligned}
&\Delta_n \leqslant C_{\text{AE}} \cdot \ell_n + 3.4314 \cdot \ell_n^{4/3}, && \textit{for any } \ell_n,\\
&\Delta_n \leqslant C_{\text{AE}} \cdot \ell_n + 0.3413 \cdot \ell_n^{4/3} < 0.4833 \cdot \ell_n, && \ell_n \leqslant 0.01,\\
&\Delta_n \leqslant C_{\text{AE}} \cdot \ell_n + 0.2538 \cdot \ell_n^{4/3} < 0.4352 \cdot \ell_n, && \ell_n \leqslant 10^{-3},\\
&\Delta_n \leqslant C_{\text{AE}} \cdot \ell_n + 0.2053 \cdot \ell_n^{4/3} < 0.4098 \cdot \ell_n, && \ell_n \leqslant 10^{-11},
\end{aligned}$$

in the general case, and

$$\begin{aligned}
&\Delta_n \leqslant C_{\text{AE}} \cdot \ell_n + 2.5786 \cdot \ell_n^{2}, && \textit{for any } \ell_n,\\
&\Delta_n \leqslant C_{\text{AE}} \cdot \ell_n + 0.7683 \cdot \ell_n^{2} < 0.4482 \cdot \ell_n, && \ell_n \leqslant 0.05,\\
&\Delta_n \leqslant C_{\text{AE}} \cdot \ell_n + 0.7455 \cdot \ell_n^{2} < 0.4172 \cdot \ell_n, && \ell_n \leqslant 0.01,\\
&\Delta_n \leqslant C_{\text{AE}} \cdot \ell_n + 0.7412 \cdot \ell_n^{2} < 0.4098 \cdot \ell_n, && \ell_n \leqslant 10^{-5},
\end{aligned}$$

in the case $F_1 = \ldots = F_n$, where

$$C_{\text{AE}} = \frac{\sqrt{10}+3}{6\sqrt{2\pi}} = 0.4097\ldots\ .$$

This corollary improves Chistyakov's inequality (1.11)

$$\Delta_n \leqslant \frac{\sqrt{10}+3}{6\sqrt{2\pi}} \cdot \ell_n + A_5 \cdot \ell_n^{40/39} |\ln \ell_n|^{7/6},$$

with respect to the remainder: the order is improved, the value of the constant is explicitly specified. Moreover, comparing the leading term of Chistyakov's estimate (1.11)

$$\psi_1(F_1, \ldots, F_n) = \frac{\sqrt{10}+3}{6\sqrt{2\pi} B_n^3} \sum_{j=1}^{n} \beta_{3,\,j}$$

with those in theorems 4.13 and 4.14

$$\psi_2(F_1, \ldots, F_n) = \inf_{c \geqslant 2/(3\sqrt{2\pi})} \left(\frac{c}{B_n^3} \sum_{j=1}^{n} \beta_{3,\,j} + \frac{K(c)}{B_n^3} \sum_{j=1}^{n} \sigma_j^3 \right),$$

we notice that their values coincide if and only if

$$\sum_{j=1}^{n} \beta_{3,j} \Big/ \sum_{j=1}^{n} \sigma_j^3 = \sqrt{20(\sqrt{10}-3)/3} = 1.0401\ldots,$$

whereas in all the rest of the cases the strict inequality $\psi_1 > \psi_2$ holds, that is, the estimates in theorems 4.13 and 4.14 are more accurate. The optimal values of c delivering the infimum in ψ_2 can be found in the fifth column of table 2 for some values of the ratio $\ell_n/\tau_n = \sum_{j=1}^n \beta_{3,j}/\sum_{j=1}^n \sigma_j^3$, which is specified in the first column named β_3.

If the value of the Lyapounov fraction $\ell_n = B_n^{-3}\sum_{j=1}^n \beta_{3,j}$ coincides with that of $B_n^{-3}\sum_{j=1}^n \sigma_j^3$ (it is easy to see that this can be if and only if $\beta_{3,j} = \sigma_j^3$ for all $j = 1,\ldots,n$, that is, when the random summands have symmetric Bernoulli distributions $\mathsf{P}(X_j = \sigma_j) = \mathsf{P}(X_j = -\sigma_j) = 1/2$), then, as it follows from remark 3.2, the greatest lower bound in the estimates of theorem 4.14 is delivered as $c \to \infty$. So, one more corollary is valid.

Corollary 4.19. *For any $n \geqslant 1$ and $F_1,\ldots,F_n \in \mathcal{F}_3$ such that $\beta_{3,j} = \sigma_j^3$ for all $j = 1,\ldots,n$, the estimate*

$$\Delta_n \leqslant \frac{\ell_n}{\sqrt{2\pi}} + 3.4106 \cdot \ell_n^{4/3}$$

holds. If $F_1 = \ldots = F_n \in \mathcal{F}_3$ and $\mathsf{E}|X_1|^3 = (\mathsf{E}X_1^2)^{3/2}$, then for all $n \geqslant 1$

$$\Delta_n \leqslant \frac{1}{\sqrt{2\pi n}} + \frac{2.5475}{n} = \frac{\ell_n}{\sqrt{2\pi}} + 2.5475 \cdot \ell_n^2.$$

Corollary 4.19 completely agrees with the results of V. Bentkus [2, 3], G. P. Chistyakov [6, 7] and Ch. Hipp and L. Mattner [14] obtained for symmetric distributions. For the case of symmetric summands, in papers [2, 3] the estimate

$$\Delta_n \leqslant \frac{\ell_n}{\sqrt{2\pi}} + A_6 \cdot \ell_n^{4/3}, \tag{4.20}$$

was announced with the same rate of decrease of the remainder, but unknown constant A_6. In [7], Chistyakov proved an analog of (4.20) with a slightly heavier remainder of the order $O\big(\ell_n^{40/39}|\ln \ell_n|^{7/6}\big)$. Corollary 4.19 improves these results of Bentkus and Chistyakov for symmetric Bernoulli distributions. The unimprovability of the constant $1/\sqrt{2\pi}$ at the Lyapounov fraction in estimates of type (4.20) for symmetric distributions was proved in 1945 by C.-G. Esseen [9] (see also [12]).

Ch. Hipp and L. Mattner in [14] considered the case where the random summands have identical symmetric Bernoulli distribution and established that

$$\Delta_n = \begin{cases} \Phi\left(\dfrac{1}{\sqrt{n}}\right) - \dfrac{1}{2}, & n \text{ odd}, \\[2ex] \dfrac{n!}{2^{n+1}\big((n/2)!\big)^2}, & n \text{ even}, \end{cases}$$

whence it follows that $\Delta_n < 1/\sqrt{2\pi n}$ for all $n \geqslant 1$. Unlike [14], corollary 4.19 gives computable estimate with the asymptotically exact constant $1/\sqrt{2\pi}$ not only for *identically* distributed summands, but for the case of *arbitrary* symmetric Bernoulli distributions as well.

However, it should be noted that for the symmetric case, actually, by the methods originally adjusted for that case, one can considerably improve all the results obtained above. These improvements will be published elsewhere.

In conclusion the author expresses her sincere gratitude to V. Yu. Korolev for fruitful discussions and permanent attention to the work and to G. P. Chistyakov for useful discussions.

References

[1] M. Abramowitz, I. Stegun. Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables. — Courier Dover Publications, 1964, 1046 p.

[2] V. Bentkus. On the asymptotical behaviour of the constant in the Berry–Esseen inequality. — Preprint 91 – 078, Universität Bielefeld, 1991.

[3] V. Bentkus. On the asymptotical behaviour of the constant in the Berry–Esseen inequality. — J. Theoret. Probab., 1994, vol. 7, No. 2, p. 211–224.

[4] A. C. Berry. The accuracy of the Gaussian approximation to the sum of independent variates. — Trans. Amer. Math. Soc., 1941, vol. 49, p. 122–139.

[5] R. N. Bhattacharya, R. Ranga Rao. Normal Approximation and Asymptotic Expansions. — New York: Wiley, 1976, 316 p.

[6] G. P. Chistyakov. Asymptotically proper constants in the Lyapunov theorem. — Journal of Mathematical Sciences, 1999, vol. 93, No. 3, p. 480–483.

[7] G. P. Chistyakov. A new asymptotic expansion and asymptotically best constants in Lyapunov's theorem. I, II, III. — Theory Probab. Appl., 2002, vol. 46, No. 2, p. 226–242; vol. 46, No. 3, p. 516–522; 2003, vol. 47, No. 3, p. 395–414.

[8] C.–G. Esseen. On the Liapunoff limit of error in the theory of probability. — Ark. Mat. Astron. Fys., 1942, vol. A28, No. 9, p. 1–19.

[9] C.–G. Esseen. Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law. — Acta Math., 1945, vol. 77, No. 1, p. 1–125.

[10] C.–G. Esseen. A moment inequality with an application to the central limit theorem. — Skand. Aktuarietidskr., 1956, vol. 39, p. 160–170.

[11] M. O. Gaponova, I. G. Shevtsova. Asymptotic estimates of the absolute constant in the Berry–Esseen inequality for distributions with unbounded third moment. — Informatics and its applications, 2009, vol. 3, No. 4, p. 41–56 (in Russian).

[12] B. V. Gnedenko, A. N. Kolmogorov. Limit distributions for sums of independent random variables. — Addison–Wesley, 1954, 293 p.

[13] M. E. Grigorieva, I. G. Shevtsova. An improvement of the Katz–Berry–Esseen inequality. — Informatics and its applications, 2010, vol. 4, No. 2, p. 78–85 (in Russian).

[14] Ch. Hipp, L. Mattner. On the normal approximation to symmetric binomial distributions. — Theory Probab. Appl., 2008, vol. 52, No. 3, p. 516–523.

[15] W. Hoeffding. The extrema of the expected value of a function of independent random variables. — Ann. Math. Statist., 1955, vol. 26, No. 2, p. 268–275.

[16] I. A. Ibragimov, Yu. V. Linnik. Independent and Stationary Sequences of Random Variables. — Wolters–Noordhoff Publishing Groningen, the Netherlands, 1971, 438 p.

[17] A. N. Kolmogorov. Some recent works in the field of limit theorems of probability theory. — Bulletin of Moscow University, 1953, vol. 7, No. 10, p. 29–38 (in Russian).

[18] A. S. Kondrik, K. V. Mikhailov, V. I. Chebotarev. On the uniform estimate of the difference of distribution functions. — Abstracts of Communications of XXXI Far-East Academician E. V. Zolotov School-Seminar, Vladivostok, Russia, 2006, p. 16–17 (in Russian).

[19] V. Yu. Korolev, I. G. Shevtsova. On the upper bound for the absolute constant in the Berry–Esseen inequality. — Theory Probab. Appl., 2010, vol. 54, No. 4, p. 638–658.

[20] V. Korolev, I. Shevtsova. An improvement of the Berry–Esseen inequality with applications to Poisson and mixed Poisson random sums. — Scand. Actuar. J., 04 June 2010. DOI:10.1080/03461238.2010.485370, vol. 2012, issue 2, p. 81–105.

[21] S. V. Nagaev, V. I. Chebotarev. On the bound of proximity of the binomial distribution to the normal one. — Theory Probab. Appl., 2011, vol. 56, No. 2, p. 248–278 (in Russian).

[22] L. V. Osipov, V. V. Petrov. On an estimate of the remainder term in the central limit theorem. — Theory Probab. Appl., 1967, vol. 12, No. 2, p. 281–286.

[23] H. Prawitz. Limits for a distribution, if the characteristic function is given in a finite domain. — Skand. Aktuarietidskr., 1972, vol. 55, p. 138–154.

[24] H. Prawitz. Weitere Ungleichungen für den absoluten Betrag einer charakteristischen Funktion. — Scand. Actuar. J., 1975, p. 21–28.

[25] H. Prawitz. On the remainder in the central limit theorem. I. Onedimensional Independent Variables with Finite Absolute Moments of Third Order. — Scand. Actuar. J., 1975, No. 3, p. 145–156.

[26] I. G. Shevtsova. Some estimates for the characteristic functions with application to sharpening of the Mises inequality. — Informatics and its applications, 2009, vol. 3, No. 3, p. 69–78 (in Russian).

[27] I. G. Shevtsova. The lower asymptotically exact constant in the central limit theorem. — Dokl. Math., 2010, vol. 81, No. 1, p. 83–86.

[28] I. G. Shevtsova. An improvement of convergence rate estimates in Lyapunov's theorem. — Dokl. Math., 2010, vol. 82, No. 3, p. 862–864.

[29] I. G. Shevtsova. On the asymptotically exact constants in the Berry–Esseen–Katz inequality. — Theory Probab. Appl., 2011, vol. 55, No. 2, p. 225–252.

[30] W. Tysiak. Gleichmäßige und nicht-gleichmäßige Berry–Esseen–Abschätzungen. Dissertation, Wuppertal, 1983.

[31] I. S. Tyurin. On the accuracy of the Gaussian approximation. — Dokl. Math., 2009, vol. 80, No. 3, p. 840–843.

[32] I. S. Tyurin. On the convergence rate in Lyapunov's theorem. — Theory Probab. Appl., 2011, vol. 55, No. 2, p. 253–270.

[33] I. S. Tyurin. Some optimal bounds in CLT using zero biasing. — Stat. Prob. Letters, 2012, vol. 82, No. 3, p. 514-518.

[34] N. G. Ushakov. Selected Topics in Characteristic Functions. — Utrecht: VSP, 1999.

[35] S. Zahl. Bounds for the Central Limit Theorem Error. Ph.D. Thesis, Harward University, Cambridge, MA, 1963.

[36] S. Zahl. Bounds for the central limit theorem error. — SIAM Journal of Applied Mathematics, 1966, vol. 14, No. 6, p. 1225–1245.

[37] V. M. Zolotarev. An Absolute Estimate of the Remainder Term in the Central Limit Theorem. — Theory Probab. Appl., 1966, vol. 11, No. 1, p. 95–105.

[38] V. M. Zolotarev. Some inequalities in the theory of probability and their application to the sharpening of a theorem of A. M. Lyapunov. — Soviet Math. Dokl., 1967, vol. 8, p. 1427–1430.

[39] V. M. Zolotarev. A sharpening of the inequality of Berry–Esseen. — Z. Wahrsch. verw. Geb., 1967, Bd. 8, P. 332–342.

[40] V. M. Zolotarev. Modern Theory of Summation of Random Variables. — Utrecht: VSP, 1997.

Annales Mathematicae et Informaticae
39 *(2012) pp. 309–323*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Discriminatory processor sharing with access rate limitations

B. Székely, A. Kőrösi, P. Vámos, J. Bíró

Budapest University of Technology and Economics
biro@tmit.bme.hu

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

In the access part of communication networks user access rates are usually limited by technology and are much lower than the bottleneck link transmission capacity carrying the traffic flows aggregated. A possible model for bandwidth sharing of the bottleneck link is the Discriminatory Processor Sharing (DPS) models, in which the server capacity (link bandwidth) is distributed among different classes of users in an unequal manner. Recently, some DPS variants incorporating the access rate limits of users have been analyzed. These models are not bandwidth sparing in a sense, that the capacity share of a class may simply be cut at its access rate limit, and the incidentally residual bandwidth is not reused in other classes. In this paper we introduce and analyze a novel variant of DPS in which the original processor sharing effect and the access rate limit constraints are combined in a bandwidth economical way resulting a truly capacity-conserving operation. Besides the state space characterization of this model, two asymptotic behaviors are also presented. We also argue in the favor of practical significance of these asymptotics, that is it could greatly help in finding high quality approximate solutions of this DPS system, i.e.q in terms of the average waiting times of flows.

1. Introduction

The original discriminatory processor sharing (DPS) model has been presented and analyzed first in [7] and [11] for modeling purposes of time-sharing computer operation. In this model there are K number of classes of users, and the state of the system can be attributed by n_i denoting the number of class-i ($i = 1, \dots, K$) users in the C capacity processor sharing system. There is also a set of weights ϕ_i, $i = 1, \dots, K$ which can be used to control the sharing of the processor capacity among the classes of customers. More formally the (instantaneous) service rate of a class$-i$ customer is

$$c_i = \frac{\phi_i}{\sum_{j=1}^{K} \phi_j n_j} C. \qquad (1.1)$$

In [6] Fayolle et al. proved the results for DPS with respect to the steady-state average response times. In [12] Rege and Sengupta showed how to obtain the moments of the queue length distributions as the solutions to linear equations in case of exponential service time requirements, and they also presented a heavy-traffic limit theorem for the joint queue length distribution. These results were extended to phase type distributions by van Kessel et al. [14]. A further remarkable milestone in DPS analysis is [1] in which the authors showed that the mean queue lengths of all classes are finite under reasonable stability conditions, regardless of the higher moments of the service requirements.

Introducing capacity limits for the customers is mainly motivated by involving access rate limitations of users (e.g. in DSL-type access systems) into the modeling framework. In [10] Lindberger analyzed the M/G/R-PS system, which is a single-class processor sharing model with access rate limit b on the users ($R := C/p$ is the "number of servers" in this system). Several improvements of this model were studied for dimensioning purposes of IP access networks, e.g. in [13] and [5] still remaining at the single-class models.

In case of multi-class discriminative processor sharing with limited access rates the question of bandwidth re-distribution is an important issue, which was not addressed in the literature. This means that if users in a class can not fully utilize their service capacity share (bandwidth share) due to their access rate limit, the problem is how this unused bandwidth is re-distributed among the other classes. In one of the extreme cases, there is no re-distribution at all meaning that the possible remaining unused bandwidth due to rate limits is wasted. One can also interpret this as the server capacity may not be fully utilized, even in those cases when there is "enough" customers in the system. This approach is followed for example in the papers [8], [2].

In this paper we present and analyze the capacity conserving case of access rate limited discriminatory processor sharing, in which all the unused bandwidth left by rate limited customers are fully utilized by the other (non-limited) customers. This is referred to as bandwidth economical discriminatory processor sharing with access rate limitations. We characterize the state space of this model, with identifying those traffic classes which are compressed (whose users are not able to utilize its access rates) and those which are not compressed (which can receive service with their access rates) and with feasible computations for their respective service rates. Two asymptotic regimes of this bandwidth economical DPS are shown and their equivalence is proven. We present that the asymptotic equilibrium point of the bandwidth efficient system is always in the non-compressed region and can simply be formulated (for every class of users), as opposed to the more complicated asymptotic equilibrium of the previously analyzed model [2].

The significance of the fluid limits lies in the following. There is still no solution in the literature for the multi-class access rate limited DPS system (in case Poisson arrivals and exponential service time requirements the equilibrium of the underlying Markov chain, consequently, the expected response times are not known). Therefore, achieving high quality approximations of system parameters have an utmost importance, e.g. from viewpoint of dimensioning tasks of communication channels for elastic flows in aggregation part of access networks or of processing capacity in highly loaded computer systems like data centers [3]. One "extreme" type of access rate limited multi-class DPS is the limitless case (no compression imposed on the classes), for which Fayolle et al. have already given the solution [6] in terms of the steady-state average response times (by integro-differential equations), and also showed that in the special case of exponential distribution of the service time requirements, the steady-state average response times of classes can be obtained by solving a system of linear equations. The fluid limit is the other extreme case of this DPS system in the sense that some of (or all) classes are "infinitely" compressed (due to infinitely speed up the system), whilst the scaled down performance parameters remain (tend to) finite values. Operational systems to be modeled or dimensioned based on DPS models stand between these two extremes, surprisingly sometimes very close to the fluid limit.

2. DPS extended by the limits of service rates

In DPS for every pair of classes i, j the ratio of the service rates allocated to class-i and class-j users is equal to the ratio of the class weights (see formula (1.1)), that is

$$\frac{c_i}{c_j} = \frac{\phi_i}{\phi_j}, \ \forall i, j \in 1, \ldots, K. \tag{2.1}$$

The total amount of capacity (in a non-empty system) used by the users of classes is evidently C, i.e.

$$\sum_{i=1}^{K} n_i c_i = C. \tag{2.2}$$

Regarding the incorporation of access rate (customer service capacity share) limits into the DPS model, in [8] and [2] a very simple approach is followed. Namely, first computing the bandwidth shares of class-i users according to (1.1) and then cutting at the access rate limits p_i, i.e.

$$c_i = \min\left(\frac{\phi_i}{\sum_{j=1}^{K} \phi_j n_j} C, p_i\right). \tag{2.3}$$

The benefit of this bandwidth share calculation is its simplicity. Nevertheless the price for simplicity is that this approach is not a bandwidth saving one, because it may happen that the total amount of capacity used by the customers is smaller then the server capacity (the server capacity is not completely shared among the users), i.e.

$$\sum_{i=1}^{K} n_i c_i < C \tag{2.4}$$

even in those cases when there are “enough” users in the system, that is

$$\sum_{i=1}^{K} n_i p_i > C. \tag{2.5}$$

In this paper we follow the other “extreme” approach, in which all the unused parts of capacity shares due to access rate limits are redistributed among users which are not imposed by these limits on. Because redistribution and sharing the whole capacity C is possible when $\sum_{i=1}^{K} n_i p_i > C$, hereafter we assume the system is in this regime. Otherwise, when $\sum_{i=1}^{K} n_i p_i \leq C$, the bandwidth shares are trivially $c_i = p_i$. In what follows we define our bandwidth economical DPS.

Definition 2.1. The bandwidth economical DPS is such a discriminatory processor sharing system in which the bandwidth shares c_i of the users of K classes at a given state $\mathbf{n} = \{n_1, \ldots, n_K\}$ are determined by the following equations:

$$c_i = \min\left\{p_i, \frac{\phi_i}{\phi_j} c_j\right\} \ \forall i, j \in \{1, \ldots, K\},\ c_j < p_j \tag{2.6}$$

and

$$\sum_{i=1}^{K} n_i c_i = C \tag{2.7}$$

where p_i is the service rate limit of class$-i$ users, $0 < p_i \leq C$.

For the next lemma without loss of generality let us assume that

$$\frac{\phi_K}{p_K} \leq \frac{\phi_i}{p_i}, \forall i = 1, \ldots, K. \tag{2.8}$$

Lemma 2.2. *For class$-K$ users $c_K < p_K$ always holds.*

Proof. The proof is based on contradiction. Assume that $c_K = p_K$. Due to (2.6) and the assumption (2.8) above it follows that

$$c_i = \min\left\{p_i, \frac{\phi_i}{\phi_K} c_K\right\} = p_i,\ \forall i = 1, \ldots, K. \tag{2.9}$$

But in this case $\sum_{i=1}^{K} n_i c_i = \sum_{i=1}^{K} n_i p_i > C$ which contradicts to equation (2.7). □

In the next corollary we show the following statement:

Corollary 2.3. *There is a unique solution of equations (2.6) and (2.7) with respect to c_i, $i = 1, \ldots, K$.*

Proof. Because of Lemma 2.2 and (2.7) and the monotone increasing property of $\min\{p_i, \frac{\phi_i}{\phi_K} x\}$ w.r.t. x, a class$-K$ user bandwidth share is a unique solution of the equation

$$\sum_{i=1}^{K} n_i \min\left\{p_i, \frac{\phi_i}{\phi_K} x\right\} = C \tag{2.10}$$

with respect to x. Therefore, every other bandwidth share is also unique and can be calculated by using c_K and the equality

$$c_i = \min\{p_i, \frac{\phi_i}{\phi_K} c_K\}. \qquad \square$$

Let a numerical example be presented for this calculation. Let $C = 100$ [Mbit/s] and five classes (with index 1 to 5 in sequence) are set up with the following paramaters: $\mathbf{n} = (8, 15, 20, 10, 30)$, $\mathbf{p} = (2, 2, 1.5, 2, 10)$ [Mbit/s], $\phi = (10, 9, 5, 4, 1)$. The following table shows the ϕ_i/p_i ratios, the access rate limits p_i, the bandwidth shares in case of original DPS (without access rate limit), of DPS with access rate limit with simple cutting at the limits using formula (2.3), and the new bandwidth economical DPS according to equations (2.6) and (2.7).

class index	1	2	3	4	5
ϕ_i/p_i	5	4.5	3.33	2	1
p_i	2	2	1.5	2	10
orig. DPS	2.5974	2.3377	1.2987	1.0389	0.2597
equ (2.3)	2	2	1.2987	1.0389	0.2597
bw eco. DPS	2	2	1.5	1.3714	0.343

Table 1: Example of bandwidth shares of different DPS systems

The fifth line of the table clearly shows that in case of simple cutting DPS (using equation (2.3), or simple comparing the third and fourth lines of the table), the class-1 and class-2 users can utilize their access rates (they are uncompressed), while classes 3, 4 and 5 are compressed (they can not reach their access rates). It can also be observed that $\sum_{i=1}^{5} n_i c_i = 90.16$ Mbit/s, that is from the total capacity 100 Mbit/s almost ten percent is wasted.

On the contrary, the last row presenting the bandwidth share of the new DPS system shows, that not only class-1 and class-2 can achieve their access rate limits, but also class-3 became uncompressed, thanks to the redistribution[1] of the unused

[1]The term 'redistribution' is used because it can be shown that the following process results exactly the same solution: start with the original DPS bandwidth share, cut at the access rate limits, and redistribute the residual banwidths among the still compressed classes, which may result some classes become uncompressed. Repeat this until the bandwidth shares no longer change.

bandwidth left by class-1 and class-2 customers. Furthermore, class-4 and class-5 bandwidth shares are also higher than in the previous case, because they can also gain from bandwidth reuse. In this case, of course $\sum_{i=1}^{5} n_i c_i = 100$ Mbit/s, hence this is attributed as bandwidth economical.

Although the computational approach above is straightforward, it is worth exploring further the structure of the system. For this, let us assume again without restriction that

$$\frac{\phi_1}{p_1} \geq \frac{\phi_2}{p_2} \geq \ldots \geq \frac{\phi_K}{p_K} \; . \tag{2.11}$$

Lemma 2.4. *If $\sum_{i=1}^{K} n_i p_i > C$ there exists an i^*, $1 \leq i^* \leq K-1$ such that*

$$\sum_{k=1}^{i^*-1} n_k p_k + \sum_{k=i^*}^{K} n_k \phi_k \frac{p_{i^*}}{\phi_{i^*}} \leq C \text{ and} \tag{2.12}$$

$$\sum_{k=1}^{i^*} n_k p_k + \sum_{k=i^*+1}^{K} n_k \phi_k \frac{p_{i^*+1}}{\phi_{i^*+1}} > C \; . \tag{2.13}$$

Proof. Note that the function

$$f(i) = \sum_{k=1}^{i-1} n_k p_k + \sum_{k=i}^{K} n_k \phi_k \frac{p_i}{\phi_i}$$

is increasing w.r.t. i due to (2.11) and exceeds C for some $i^*+1 \leq K$, otherwise $f(K) = \sum_{i=1}^{K} n_i p_i \leq C$ would hold which is not true. □

As an important consequence of this lemma it is also worth noting that

$$\{1,\ldots,i\} \subset \mathcal{U}(\underline{n}) \text{ iff } \sum_{k=1}^{i-1} n_k p_k + \sum_{k=i}^{K} n_k \phi_k \frac{p_i}{\phi_i} \leq C \tag{2.14}$$

where $\mathcal{U}(\underline{n}) := \{1,\ldots,i^*\}$ is the set of uncompressed classes in the state $\underline{n}$.

Now the main theorem of this section is the following:

Theorem 2.5. *The unique solution of (2.6) and (2.7) can be expressed through i^* in the following way:*

$$c_k = p_k, \text{ if } k \leq i^* \text{ and} \tag{2.15}$$

$$c_k = \frac{\phi_k}{\sum_{i=i^*+1}^{K} \phi_i n_i} \left(C - \sum_{j=1}^{i^*} n_j p_j \right), \text{ if } i^* < k. \tag{2.16}$$

Proof. The validity of (2.7) can easily be checked. Next we show that (2.6) is fulfilled by c_k, c_l for which $k, l \in \mathcal{Z}(\underline{n}) := \{1,\ldots,K\} \setminus \mathcal{U}(\underline{n})$. In this case due to (2.13) and (2.11) $c_k < p_k$ and $c_l < p_l$. Moreover $c_k/p_k = c_l/p_l$ holds, therefore (2.6) is satisfied, that is $c_k = \min\{p_k, \frac{\phi_k}{\phi_l} c_l\}$.

Now assume that $l \in \mathcal{U}(\underline{n})$ and $k \in \mathcal{Z}(\underline{n})$. In this case $c_k < p_k$, therefore

$$\frac{\phi_l}{\phi_k} c_k = \frac{\phi_l}{\sum_{i=i^*+1}^{K} \phi_i n_i} \left(C - \sum_{j=1}^{i^*} n_j p_j \right) \tag{2.17}$$

which is not less than p_l due to (2.12) and (2.11). Hence,

$$c_l = \min\{p_l, \frac{\phi_l}{\phi_k} c_k\} = p_l$$

that is (2.6) is again fulfilled. □

3. Asymptotic behaviors of the bandwidth economical DPS

In this section we first show that the so-called fluid limit of the processor sharing model investigated in this paper exists. Then we find the equilibrium of the fluid limit. The stability of this equilibrium has been also proved, however, not presented in this paper. Assume that the service times are exponentially distributed and the arrival processes follow Poisson processes. Then in this case the number of jobs (of customers) in the system can be modeled by a Markov chain. The equilibrium of the Markov chain, consequently, the expected response times are not known. Fluid scaling is a possible asymptotic regime in which one may expect computing the equilibrium at least for the limiting structure. In fluid limit the arrival processes are accelerated by a common factor and the capacity of the server is speed up by the same factor. If the accelerating factor goes to infinity then in limit one gets the fluid limit of the number of waiting jobs. The limiting process of the number of waiting jobs is deterministic, it is a solution of a differential equation. The equilibrium of this differential equation can be found using analytical considerations. We remark that the fluid limit of many processor sharing model, as well as the one investigated in this paper, can be determined by using classical results presented in e.g. [4, Chapter 11].

For finding the fluid limit of our model first the transition rates are to be determined $q(\underline{n}, \underline{n} + \underline{l})$ from state $\underline{n}$ to $\underline{n} + \underline{l}$. Let $\underline{e}_k$ be a vector such that in $\underline{e}_j$ 1 stands at coordinate j and except this coordinate each coordinate is 0. For any $j = 1, \ldots, K$

$$\begin{aligned} q(\underline{n}, \underline{n} + \underline{e}_j) &= \lambda_j \\ q(\underline{n}, \underline{n} - \underline{e}_j) &= \mu_j n_j p_j && \text{if } j \in \mathcal{U}(\underline{n}) \\ q(\underline{n}, \underline{n} - \underline{e}_j) &= \mu_j n_j \phi_j \frac{C - \sum_{i \in \mathcal{U}(\underline{n})} p_i n_i}{\sum_{i \in \mathcal{Z}(\underline{n})} \phi_i n_i} && \text{if } j \in \mathcal{Z}(\underline{n}) \\ q(\underline{n}, \underline{n} + \underline{l}) &= 0 && \text{if } \underline{l} \neq \pm \underline{e}_k \\ & && \text{for some } k = 1, \ldots, K. \end{aligned} \tag{3.1}$$

Let $c_j(\underline{n})$ denote the bandwidth that a stream of class j obtains. We have

$$c_j(\underline{n}) = p_j \mathbf{I}\{j \in \mathcal{U}(\underline{n})\} + \phi_j \frac{C - \sum_{i\in\mathcal{U}(\underline{n})} p_i n_i}{\sum_{i\in\mathcal{Z}(\underline{n})} \phi_i n_i} \mathbf{I}\{j \in \mathcal{Z}(\underline{n})\}. \tag{3.2}$$

We remark that using (2.14), $c_j(\underline{n})$ can be given as an explicit function of $\underline{n}$ as follows:

$$\begin{aligned} c_j(\underline{n}) = p_j \mathbf{I}&\left\{\sum_{k=1}^{j-1} n_k p_k + \sum_{k=j}^{K} n_k \phi_k \frac{p_j}{\phi_j} \le C\right\} \\ &+ \phi_j \frac{C - \sum_{i\in\mathcal{U}(\underline{n})} p_i n_i}{\sum_{i\in\mathcal{Z}(\underline{n})} \phi_i n_i} \mathbf{I}\left\{\sum_{k=1}^{j-1} n_k p_k + \sum_{k=j}^{K} n_k \phi_k \frac{p_j}{\phi_j} > C\right\}. \end{aligned} \tag{3.3}$$

Of course, this definition makes sense for $\underline{n} \in \mathbb{R}_+^K$.

Let $\Pi_j^a(t), t \ge 0$ and $\Pi_j^d(t), t \ge 0$ for $j = 1, \dots, K$ be $2K$ independent Poisson processes with rate 1. Let $N_j(t)$ be the number of flows from class j in the system at time t. Then by the rates in (3.1) we have

$$N_j(t) = N_j(0) + \Pi_j^a\big(\lambda_j t\big) - \Pi_j^d\left(\int_0^t \mu_j N_j(s) c_j\left(\underline{N}(s)\right) \mathrm{d}s\right). \tag{3.4}$$

Let $\lambda_j^L = \lambda_j L$, $j = 1, \dots, K$, $C^L = CL$. Let $N_j^L(t)$ be the number of flows from class j in the system at time t if the arrival intensities to the classes are $\lambda_1^L, \dots, \lambda_K^L$ respectively and the capacity is C^L. Simply rewriting the equation (3.4) for $N^L(t), t \ge 0$ and dividing by L we get

$$\begin{aligned} \frac{N_j^L(t)}{L} = \frac{N_j^L(0)}{L} &+ \frac{1}{L}\Pi_j^a\big(L\lambda_j t\big) \\ &- \frac{1}{L}\Pi_j^d\left(L\int_0^t \mu_j \frac{N_j^L(s)}{L} c_j\left(\frac{\underline{N}^L(s)}{L}\right) \mathrm{d}s\right) \qquad j = 1, \dots, K. \end{aligned}$$

For the ease of notations we rewrite this equation. Introducing $n_j^L(t) = \frac{N_j^L(t)}{L}$ $j = 1, \dots, K$ we have

$$\begin{aligned} n_j^L(t) = n_j^L(0) &+ \frac{1}{L}\Pi_j^a\big(\lambda_j L t\big) \\ &- \frac{1}{L}\Pi_j^d\left(L\int_0^t \mu_j n_j^L(s) c_j\left(\underline{n}^L(s)\right) \mathrm{d}s\right) \qquad j = 1, \dots, K \end{aligned} \tag{3.5}$$

The theory presented in [4, Ch 6.4 and Ch 11.2] can be applied to the process $\underline{n}^L(t), t \geq 0$ for obtaining convergence to $\underline{n}(t), t \geq 0$ the solution of the system of equations

$$n_j(t) = n_j(0) + \lambda_j t - \int_0^t \mu_j n_j(s) c_j\left(\underline{n}(s)\right) \mathrm{d}s, \quad j = 1, \dots, K \tag{3.6}$$

as it is stated in the following theorem.

Theorem 3.1. *Assume that* $\lim_{L\to\infty} n_j^L(0) = n(0) \in [0, \infty)$ *for any* $j = 1, \dots, K$. *Then for every* $t \geq 0$,

$$\lim_{L\to\infty} \sup_{s\leq t} \left|\underline{n}^L(s) - \underline{n}(s)\right| = 0 \quad a.s. \tag{3.7}$$

Proof. We will apply Theorem 2.1 of [4, p 456]. We have to check three conditions. First, for any compact set $B \subset [0, \infty)^K$ the following bound holds

$$\sup_{\underline{n}\in B} n_j c_j(\underline{n}) < \infty \quad j = 1, \dots, K, \tag{3.8}$$

second, there exist M_B such that for any $j = 1, \dots, K$

$$|n_j c_j(\underline{n}) - m_j c_j(\underline{m})| \leq M_B \left|\underline{n} - \underline{m}\right| \quad \underline{n}, \underline{m} \in B. \tag{3.9}$$

Third,

$$\lim_{L\to\infty} n_j^L(0) = n(0) \in [0, \infty) \quad j = 1, \dots, K. \tag{3.10}$$

Using (3.3) Simple calculations show that (3.8) and (3.9) hold. The condition (3.10) is the same as the assumption of Theorem 3.1. Therefore, the convergence (3.7) holds. □

The main results of this section is the following.

Theorem 3.2. *If the function* $\underline{n}(t), t \geq 0$ *satisfies the equations (3.6) then in the stationary state* n_j^*, $j = 1, \dots, K$ *each class is uncompressed and the the following holds:*

$$n_j^* = \frac{\lambda_j}{\mu_j p_j} \quad j = 1, \dots, K.$$

Proof. For finding the stationary state $\underline{n}^*$ of the fluid limit differentiate $n_j(t)$, $j = 1, \dots, K$ with respect to t and find the solution of the system $n_j'(t) = 0$, $j = 1, \dots, K$. Using (3.6) and (3.2) one gets

$$\begin{aligned} 0 = n_j'(t) = \lambda_j - \mu_j n_j(t) \\ \cdot \left(p_j \mathbf{I}\{j \in \mathcal{U}(\underline{n}(t))\} + \phi_j \frac{C - \sum_{i\in\mathcal{U}(\underline{n})} p_i n_i(t)}{\sum_{i\in\mathcal{Z}(\underline{n})} \phi_i n_i(t)} \mathbf{I}\{j \in \mathcal{Z}(\underline{n}(t))\} \right) \end{aligned}$$

This means that in the stable state we have

$$\lambda_j = \mu_j p_j n_j^* \text{ if } j \in \mathcal{U}(\underline{n}^*),$$
$$\lambda_j = \mu_j n_j^* \frac{\phi_j}{\sum_{i\in\mathcal{Z}} \phi_i n_i^*}\left(C - \sum_{i\in\mathcal{U}} p_i n_i^*\right) \text{ if } j \in \mathcal{Z}(\underline{n}^*).$$

If there is at least one compressed class, that is, $\mathcal{Z}(\underline{n}^*) \neq \emptyset$ then we have for $j \in \mathcal{Z}(\underline{n}^*)$

$$\begin{aligned}\lambda_j &= \mu_j n_j^* \frac{\phi_j}{\sum_{i\in\mathcal{Z}(\underline{n}^*)} \phi_i n_i^*}\left(C - \sum_{i\in\mathcal{U}(\underline{n}^*)} p_i n_i^*\right)\\ &= \mu_j n_j^* \frac{\phi_j}{\sum_{i\in\mathcal{Z}(\underline{n}^*)} \phi_i n_i^*}\left(C - \sum_{i\in\mathcal{U}(\underline{n}^*)} \frac{\lambda_i}{\mu_i}\right)\\ &= \mu_j n_j^* \frac{\phi_j}{\sum_{i\in\mathcal{Z}(\underline{n}^*)} \phi_i n_i^*}\left(C - \sum_{i\in\mathcal{U}(\underline{n}^*)} C\varrho_i\right)\end{aligned}$$

since the definition $\varrho_i = \frac{\lambda_i}{\mu_i C}$. Dividing by $\mu_j C$ and using $\varrho_j = \frac{\lambda_j}{\mu_j C}$ one gets

$$\varrho_j = \frac{\phi_j n_j^*}{\sum_{i\in\mathcal{Z}(\underline{n}^*)} \phi_i n_i^*}\left(1 - \sum_{i\in\mathcal{U}(\underline{n}^*)} \varrho_i\right),$$

rearranging the terms on the right we have

$$\frac{\varrho_j}{1 - \sum_{i\in\mathcal{U}(\underline{n}^*)} \varrho_i} = \frac{\phi_j n_j^*}{\sum_{i\in\mathcal{Z}(\underline{n}^*)} \phi_i n_i^*},$$

then summing both sides over $j \in \mathcal{Z}(\underline{n}^*)$ one has

$$\frac{\sum_{j\in\mathcal{Z}(\underline{n}^*)} \varrho_j}{1 - \sum_{i\in\mathcal{U}(\underline{n}^*)} \varrho_i} = 1,$$

this is equivalent to $\sum_{j=1}^{K} \varrho_j = 1$ which is contradiction. Consequently, $\mathcal{Z}(\underline{n}^*) = \emptyset$ and for any $j = 1, \ldots, K$ $n_j^* = \frac{\lambda_j}{\mu_j p_j}$. □

It can also be shown that the equilibrium $\underline{n}^*$ is stable, nevertheless, due to the lack of space it is not performed here. It has been elaborated following the argumentation in [2, pp 48–49] and also using [9, Lemma 3].

Here we note that in this bandwidth economical DPS the fluid limit lies completely in the uncompressed region (every classes in the limit are uncompressed), and the closed form expression of the fluid limit of a class depends only on the class parameters (λ_j, μ_j, p_j), and is quite simple.

On the contrary, in case of the previously analyzed DPS [2] (based on equation (2.3)) the fluid limit has no closed form solution, an algorithm is needed to determine the compressed and uncompressed classes and the corresponding limits in the asymptotics. Furthermore, the limit of a class may depend on the parameters of other classes (see Proposition 1.3. in [2]).

4. Fluid limit as the number of servers goes to infinity

In the concept of fluid limit the intensity of the arrival processes and the capacity of the server increase in the same pace by a multiplier L. Consequently, the number of packets under service increases and the number of served packets in unit time increases as well. The first property can be rephrased as the number of servers ($\frac{C}{p_j}$) increases. It is natural to ask whether one can take an asymptotic regime in which the number of servers increases but the intensities of the arrivals and the capacity are fixed. If so, then what can be said about the limit process. A possible way of considering such an asymptotic is that we decrease the access rates by L and take $p_j^L = p_j/L$. This is not enough to obtain fluid scaling like set up because the number of served packets per unit time does not increase. One can get over this problem and obtain limit of similar kind as the fluid limit if the time of the system is accelerated too. This regime will be described in this section.

Let us fix C and λ_j and decrease the access rate limits p_j, such that $p_j^L = \frac{p_j}{L}$, $j = 1, \ldots, K$ for $L > 0$. Let $M_j^L(t)$ be the number of flows from class j in the system at time t if the access rate limits are p_j^L. It can be proved that the rescaled and time accelerated process has fluid limit.

Theorem 4.1. *Assume that* $\lim_{L\to\infty} \frac{M_j^L(Lt)}{L} = m(0) \in [0, \infty)$ *for any* $j = 1, \ldots, K$. *For the processes* $M_j^L(t)$, $j = 1, \ldots, K$ *defined above we have the following fluid limit*

$$\lim_{L\to\infty} \sup_{s\le t} \left| \frac{M_j^L(Lt)}{L} - \underline{n}(s) \right| = 0 \quad a.s. \tag{4.1}$$

where $\underline{n}(s)$ *is the solution of the differential equation (3.6). Consequently,*

$$\lim_{t\to\infty} \lim_{L\to\infty} \frac{M_j^L(Lt)}{L} = n_j^* \quad j = 1, \ldots, K, \tag{4.2}$$

where n_j^* *is defined in Theorem 3.2.*

Proof. We will prove that the process $\frac{\underline{M}^L(Lt)}{L}, t \ge 0$ satisfies equation (3.5). Consequently, Theorem 3.1 can be applied for $\frac{\underline{M}^L(Lt)}{L}, t \ge 0$ yielding the same convergence (4.1). Then one can conclude that Theorem 3.2 holds for the limit process (4.2) without any further modification.

Proving the process $\frac{M^L(Lt)}{L}, t \geq 0$ satisfies equation (3.5), we first rewrite the equation (3.4) for $M^L(t), t \geq 0$. We have

$$M_j^L(t) = M_j^L(0) + \Pi_j^a\big(\lambda_j t\big) - \Pi_j^d\left(\int\limits_0^t \mu_j M_j^L(s) c_j^{L*}\left(\underline{M}^L(s)\right)\,\mathrm{d}s\right),$$

where for any $\underline{m} \in [0,\infty)^K$ we define

$$c_j^{L*}(\underline{m}) = p_j \mathbf{I}\left\{\sum_{k=1}^{j} m_k \frac{p_k}{L} + \sum_{k=j+1}^{K} m_k \phi_k \frac{p_j}{\phi_j L} \leq C\right\}$$
$$+ \mu_j \phi_j \frac{C - \sum_{i\in\mathcal{U}(\underline{m})} p_i m_i}{\sum_{i\in\mathcal{Z}(\underline{m})} \phi_i n_i} \mathbf{I}\left\{\sum_{k=1}^{j} m_k \frac{p_k}{L} + \sum_{k=j+1}^{K} m_k \phi_k \frac{p_j L}{\phi_j} > C\right\}.$$

As previously we devide by L and for having fluid limit we speed up the time by L:

$$\frac{M_j^L(Lt)}{L} = \frac{M_j^L(0)}{L} + \frac{1}{L}\Pi_j^a\big(\lambda_j Lt\big)$$
$$- \frac{1}{L}\Pi_j^d\left(\int\limits_0^{Lt} \mu_j M_j^L(s) c_j^{L*}\left(\underline{M}^L(s)\right)\,\mathrm{d}s\right).$$

Using the fact that $\int_0^{Lt} f(s)\,\mathrm{d}s = \int_0^t Lf(Ls)\,\mathrm{d}s$ we have

$$\frac{M_j^L(Lt)}{L} = \frac{M_j^L(0)}{L} + \frac{1}{L}\Pi_j^a\big(\lambda_j Lt\big)$$
$$- \frac{1}{L}\Pi_j^d\left(L\int\limits_0^{t} \mu_j M_j^L(Ls) c_j^{L*}\left(\underline{M}^L(Ls)\right)\frac{1}{L}\,\mathrm{d}s\right). \tag{4.3}$$

From the definition of c_j^{L*} and c_j it follows that

$$c_j^{L*}\left(M_j^L(Lt)\right) = \frac{1}{L} c_j\left(\frac{M_j^L(Lt)}{L}\right).$$

This equation and (4.3) implies that

$$\frac{M_j^L(Lt)}{L} = \frac{M_j^L(0)}{L} + \frac{1}{L}\Pi_j^a\big(\lambda_j Lt\big)$$
$$- \frac{1}{L}\Pi_j^d\left(\int\limits_0^{t} \mu_j M_j^L(Ls) c_j\left(\frac{\underline{M}^L(Ls)}{M}\right)\,\mathrm{d}s\right). \tag{4.4}$$

Introducing $m_j^L(t) = \frac{M_j^L(Lt)}{L}$, (4.4) can be written as

$$m_j^L(t) = m_j^L(0) + \frac{1}{L}\Pi_j^a(\lambda_j Lt) - \frac{1}{L}\Pi_j^d\left(L\int_0^t \mu_j m_j^L(s) c_j\left(\underline{m}^L(s)\right)\,\mathrm{d}s\right).$$

which is the same as equation (3.5) and for the processes $m_j^L(t), t \geq 0$ we have fluid limit. □

4.1. The one-dimensional case

Let us consider the M/G/1-PS system as a special one-dimensional case of the multiclass Processor Sharing. The average number of customers in the sationary state of the system is $EN = \frac{\rho}{1-\rho}$ where $\rho = \frac{\lambda}{\mu C}$. It can easily be shown (also based on the previous discussion) that M/G/1-PS has a stable fluid limit, which is $n^* = \lim_{L\to\infty} \frac{EN^L}{L} = \rho$ where N^L is the average number of customers in the L times speed up M/G/1-PS system ($\lambda^L = L\lambda, C^L = LC$). Similarly to the multiclass case above, here it is also true that the very same fluid limit results if the number of servers goes to infinity (with C and λ fixed), that is $L := \frac{C}{p}$ tends to infinity (with p tending to zero) where p is the *access rate limit.* This observation is very important, because for every finite L the system is equivalent to the M/G/L-PS (in the literature often referred to as M/G/R-PS [10]) system whose solution is known. It means that in this single class case not only the two ‘extreme’ systems (the access rate limitless M/G/1-PS case when $L = 1$ and the fluid limit when $L = \infty$) can be characterized, but every system between them can be solved, thus the convergence to the limit can fully be described. Based on the formula for the average number of customers presented in [10] for M/G/L-PS, one can obtain

$$\frac{EN^L}{L} = \rho\left(1 + \frac{E_2(L, L\rho)}{L(1-\rho)}\right) \tag{4.5}$$

where E_2 is Erlang's second formula. It can easily be checked that the formula above gives $\frac{\rho}{1-\rho}$ for $L = 1$, and ρ for $L = \infty$. Of course, the M/G/R-PS itself has also fluid limit, which is $R\rho = \frac{\lambda}{\mu C/R} = \frac{\lambda}{\mu p}$ (see the similarities to the multiclass case in Theorem 3.2) and the convergence to the fluid limit can be characterized by using similar formula as in (4.5).

We strongly believe that this well characterizable convergence to the fluid limit of single class DPS can be utilized for solving the bandwidth economical multiclass access rate limited DPS, because the solution of the original model [6], and the fluid limit of the access rate limited multiclass DPS presented in this paper are already in our hands.

5. Conclusion

In this paper we have analyzed a bandwidth economical discriminatory processor sharing system with access rate limitations, as a possible and realistic model for bandwidth sharing of (elastic) network traffic flows subject to flow control and access rate limits. We have characterized the state-space and determined the unique state-dependent bandwidth shares of such a capacity conserving system, in which the unused capacity of users due to the effect of their access rate limits is fully re-distributed among other users. We have also presented two asymptotic regimes of the system which may help in the further research to obtain computationally tractable methods for evaluating the performance.

References

[1] K. Avrachenkov, U. Ayesta, P. Brown, and R. Núñez-Queija. Discriminatory processor sharing revisited. In *In: Proc. IEEE Infocom 2005, Miami FL*, pages 784–795, 2005.

[2] U. Ayesta and M. Mandjes. Bandwidth-sharing networks under a diffusion scaling. *Annals Operation Research*, 170(1):41–58, 2009.

[3] Niklas Carlsson and Martin Arlitt. Towards more effective utilization of computer systems. In *ICPE '11 Proceeding of the second joint WOSP/SIPEW International Conference on Performance Engineering*, 2011.

[4] S.N. Ethier and T.G. Kurtz. *Markov processes: Characterization and convergence.* Wiley New York, 1986.

[5] Z. Fan. Dimensioning Bandwidth for Elastic Traffic. *NETWORKING 2002: Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications*, pages 826–837, 2006.

[6] G. Fayolle, I. Mitrani, and R. Iasnogorodski. Sharing a processor among many job classes. *J. ACM*, 27(3):519–532, 1980.

[7] L. Kleinrock. Time-shared systems: A theoretical treatment. *J. of ACM*, 14(2):242–261, 1967.

[8] A. Lakshmikantha, R. Srikant, and CL Beck. Differential equation models of flow-size based priorities in internet routers. *International Journal of Systems, Control and Communications*, 2(1):170–196, 2010.

[9] L. Leskela. Stabilization of an overloaded queueing network using measurement-based admission control. *Journal of Applied Probability*, 43(1):231–244, 2006.

[10] K. Lindberger. Balancing quality of service, pricing and utilisation in multiservice networks with stream and elastic traffic. In *ITC-16: International Teletraffic Congress*, pages 1127–1136, 1999.

[11] T. M. O'Donovan. Direct solutions of m/g/1 processor-sharing models. *Operation Research*, 22:1232–1235, 1974.

[12] K.M. Rege and B. Sengupta. Queue length distribution for the discriminatory processor-sharing queue. *Operation Research*, 44:653–657, 1996.

[13] A. Riedl, T. Bauschert, and J. Frings. A framework for multi-service IP network planning. In *International Telecommunication Network Strategy and Planning Symposium (Networks)*, pages 183–190, 2002.

[14] G. van Kessel, R. Núñez-Queija, and S. Borst. Asymptotic regimes and approximations for discriminatory processor sharing. *ACM SIGMETRICS Performance Evaluation Review*, 32(2):44–46, 2004.

Annales Mathematicae et Informaticae
39 *(2012) pp. 325–338*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Simulation of vibrations of a rectangular membrane with random initial conditions

Anna Slyvka Tylyshchak

Kyiv National T. Shevchenko University, Ukraine

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

A new method is proposed in this paper to construct models for solutions of boundary-value problems for hyperbolic equations with random initial conditions. We assume that the initial conditions are strictly sub-Gaussian random fields (in particular, Gaussian random fields with zero mean). The models approximate solutions with a given accuracy and reliability in the uniform metric.

Keywords: Rectangular Membrane's Vibrations, Stochastic processes, Model of solution, Accuracy and Reliability.

MSC: 60G60; Secondary 60G15.

1. Introduction

We construct a model that approximates a solution of the boundary-value problem (2.1)–(2.3) for the hyperbolic equation with random initial conditions. The model is convenient to use when developing a software for computers. It approximates a solution with a given reliability and accuracy in the uniform metric.

We consider a strictly sub-Gaussian random field to model initial conditions in problem (2.1)–(2.3). Note that Gaussian fields are particular cases of strictly sub-Gaussian random fields.

It is known that a solution of the boundary-value problem can be represented, under certain conditions in the form of an infinite series, namely

$$u(x,y,t) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} V_{nm}(x,y)\left[a_{nm}\cos\sqrt{\lambda_{nm}}t + b_{nm}\sin\sqrt{\lambda_{nm}}t\right],$$

where $V_{nm}(x, y)$ are known functions and a_{nm} and b_{nm} are random variables whose joint distributions are known.

One can consider the following model for a solution of the boundary-value problem:

$$u(x, y, t, N) = \sum_{n=1}^{N} \sum_{m=1}^{N} V_{nm}(x, y) \left[a_{nm} \cos \sqrt{\lambda_{nm}} t + b_{nm} \sin \sqrt{\lambda_{nm}} t\right],$$

One can find the values of N for which $u(x,\ y,\ t,\ N)$ approximates the field $u(x;\ y,\ t)$ with a given reliability and accuracy.

The main disadvantage of this method is that the random variables a_{nm} and b_{nm} are independent only for very special initial conditions. Therefore it is practically impossible to apply this method for large N.

A new method is proposed in this paper to construct a model for a solution of the boundary problem (2.1)–(2.3). The idea of the method is, first, to model the initial conditions with a given accuracy and, second, to compute approximate values $\widetilde{a}_{nm}$ and $\widetilde{b}_{nm}$ of the coefficients a_{nm} and b_{nm}, respectively, by using the model for the initial conditions. The finite sum

$$\widetilde{u}(x, y, t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} V_{nm}(x, y) \left[\widetilde{a}_{nm} \cos \sqrt{\lambda_{nm}} t + \widetilde{b}_{nm} \sin \sqrt{\lambda_{nm}} t\right],$$

is considered as a model for the solution. We find values of N and an accuracy of the approximation of a_{nm} and b_{nm} by $\widetilde{a}_{nm}$ and $\widetilde{b}_{nm}$ for which this model approximates the solution of the boundary-value problem with a given reliability and accuracy in the uniform metric.

Note the paper consists of five section. The main result, Theorem 2.2 is stated in Section 2. The proof of the theorem is given in Section 3, and some examples are considered in Section 4. The model of a solution of a hyperbolic type equation with random initial conditions was investigated in the paper [7].

Note that all the results of the paper hold for the case where the initial conditions are zero mean Gaussian random fields. Some methods to model Gaussian and sub-Gaussian random processes and random fields can be found in the articles [4], [5] and the book [3].

2. Main result

Consider the problem of vibrations of a rectangular membrane [8] $0 < x < p$, $0 < y < q$:

$$u_{xx} + u_{yy} = u_{tt}, \tag{2.1}$$

$$u|_{t=0} = \xi(x, y), \frac{\partial u}{\partial t}|_{t=0} = \eta(x, y), \tag{2.2}$$

$$u|_{s} = 0, \tag{2.3}$$

where u is the deviation of the membrane from its equilibrium position, which coincides with the plane x, y, S is boundary of a rectangle $0 < x < p$, $0 < y < q$.

Let the initial conditions $\{\xi(x,y), x \in [0,p], y \in [0,q]\}$, $\{\eta(x,y), x \in [0,p], y \in [0,q]\}$ be an independent strictly sub-Gaussian stochastic processes (see [1]).

When solving problems similar (2.1)–(2.3) by Fourier's method, regardless of whether initial conditions are random or nonrandom, we look for a solution of the form

$$u(x,y,t) = \sum_{n=1}^{\infty}\sum_{m=1}^{\infty} V_{nm}(x,y)\left[a_{nm}\cos\sqrt{\lambda_{nm}}t + b_{nm}\sin\sqrt{\lambda_{nm}}t\right], \tag{2.4}$$

where

$$a_{nm} = \int_0^p\int_0^q \xi(x,y)V_{nm}(x,y)dxdy,$$

$$b_{nm} = \frac{1}{\sqrt{\lambda_{nm}}}\int_0^p\int_0^q \eta(x,y)V_{nm}(x,y)dxdy,$$

λ_{nm} and $V_{nm}(x,y)$ are eigenvalues and eigenfunctions of the Sturm-Liouville problem [8]:

$$V_{xx} + V_{yy} + \lambda V = 0,$$

$$V|_s = 0.$$

where λ_{nm} and $V_{nm}(x,y)$ have the following forms

$$\lambda_{nm} = \pi^2\left(\frac{n^2}{p^2} + \frac{m^2}{q^2}\right),$$

$$V_{nm}(x,y) = \frac{2}{\sqrt{pq}}\sin\frac{n\pi}{p}x\sin\frac{m\pi}{q}y, \tag{2.5}$$

where $n, m = 1, 2, \ldots$

In the papers by [6] (see also [2]) the theorems are formulated according to the conditions of which series (2.4) is the solution of problem (2.1)–(2.3).

Let's construct a model for a solution of problem (2.1)–(2.3) approximating the solution with a given reliability and accuracy in the uniform metric.

Let $\{\widehat{\xi}(x,y), x \in [0,p], y \in [0,q]\}$ and $\{\widehat{\eta}(x,y), x \in [0,p], y \in [0,q]\}$ be models of processes $\{\xi(x,y), x \in [0,p], y \in [0,q]\}$ and $\{\eta(x,y), x \in [0,p], y \in [0,q]\}$, respectively. Note that the models $\widehat{\xi}(x,y)$ and $\widehat{\eta}(x,y)$ are independent stochastic processes.

Put

$$\widehat{a}_{nm} = \int_0^p\int_0^q \widehat{\xi}(x,y)V_{nm}(x,y)dxdy,$$

$$\widehat{b}_{nm} = \frac{1}{\sqrt{\lambda_{nm}}} \int_0^p \int_0^q \widehat{\eta}(x,y) V_{nm}(x,y) dxdy.$$

The sum

$$u^N(x,y,t) = \sum_{n=1}^{N} \sum_{m=1}^{N} V_{nm}(x,y) \left[\widehat{a}_{nm} \cos\sqrt{\lambda_{nm}}t + \widehat{b}_{nm} \sin\sqrt{\lambda_{nm}}t\right] \tag{2.6}$$

is called a model of the process $u(x,y,t)$.

Definition 2.1. Let a solution $u(x,y,t)$ of problem (2.1)–(2.3) be represented in the form of series (2.4). We say that a model $u^N(x,y,t)$ approximates $u(x,y,t)$ with a given reliability $1-\gamma$ and accuracy δ in the uniform metric in the domain $D = [0,p] \times [0,q] \times [0,T]$ if

$$P\left\{ \sup_{(x,y,t)\in D} |u^N(x,y,t) - u(x,y,t)| > \delta \right\} \le \gamma.$$

Put

$$\Delta_N(x,y,t,N) = u(x,y,t) - u^N(x,y,t) = u_N(x,y,t) + V_N(x,y,t),$$

where

$$u_N(x,y,t) = \sum_{n=N+1}^{\infty} \sum_{m=N+1}^{\infty} V_{nm}(x,y) \left[a_{nm} \cos\sqrt{\lambda_{nm}}t + b_{nm} \sin\sqrt{\lambda_{nm}}t\right],$$

$$V_N(x,y,t) = \sum_{n=1}^{N} \sum_{m=1}^{N} V_{nm}(x,y) \left[(\widehat{a}_{nm} - a_{nm}) \cos\sqrt{\lambda_{nm}}t + (\widehat{b}_{nm} - b_{nm}) \sin\sqrt{\lambda_{nm}}t\right].$$

Below is the main result of the paper.

Theorem 2.2. *Let $\{\xi(x,y), x \in [0,p], y \in [0,q]\}$ and $\{\eta(x,y), x \in [0,p], y \in [0,q]\}$ be independent $SSub(\Omega)$ processes. Let the models $\{\widehat{\xi}(x,y), x \in [0,p], y \in [0,q]\}$ and $\{\widehat{\eta}(x,y), x \in [0,p], y \in [0,q]\}$ be such that*

$$\frac{1}{\sqrt{pq}} \int_0^p \int_0^q \sqrt{E\left(\widehat{\xi}(x,y) - \xi(x,y)\right)^2} dxdy \le \Lambda,$$

$$\frac{1}{\sqrt{pq}} \int_0^p \int_0^q \sqrt{E\left(\widehat{\eta}(x,y) - \eta(x,y)\right)^2} dxdy \le \Lambda.$$

Then the stochastic process $u^N(x,y,t)$ defined by (2.6), is a model of the stochastic process $u(x,y,t)$ that approximates it with reliability $1-\gamma$ and accuracy δ in the uniform metric in the domain $D = [0,p] \times [0,q] \times [0,T]$ if γ and N a such that

$$\left(T^{1/2} + p^{1/2} + q^{1/2}\right) A_N^2 \epsilon_0^2(N) < \delta,$$

$$\frac{1}{2}\left(\frac{\delta^{1/3}\left(\delta^{2/3}-3\left(T^{1/2}+p^{1/2}+q^{1/2}\right)^{\frac{2}{3}}A_N^{1/3}\epsilon_0^{1/3}(N)\right)}{\epsilon_0(N)}\right)^2 \geq \ln\frac{1}{\gamma},$$

where

$$A_N = \frac{2\pi}{p^{3/2}q^{3/2}}\left\{\left(\left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Ea_{nm}^2}(nq+mp)\right)^2 + \left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Eb_{nm}^2}(nq+mp)\right)^2\right)^{1/2} + 2\Lambda\left(\left(\sum_{n=1}^{N}\sum_{m=1}^{N}(nq+mp)\right)^2 + \left(\sum_{n=1}^{N}\sum_{m=1}^{N}\frac{pq}{\pi\sqrt{n^2q^2+m^2p^2}}(nq+mp)\right)^2\right)^{1/2}\right\},$$

$$\epsilon_0(N) = \frac{4}{\sqrt{pq}}\left\{\left(\left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Ea_{nm}^2}\right)^2 + \left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Eb_{nm}^2}\right)^2\right)^{1/2} + \Lambda\left(N^4 + \left(\frac{1}{\pi^2}\sum_{n=1}^{N}\sum_{m=1}^{N}\frac{pq}{\sqrt{n^2p^2+m^2q^2}}\right)^2\right)^{1/2}\right\}.$$

Remark 2.3. If the conditions of Theorem 4.3 in the paper [6] are hold true the series in definitions A_N and $\varepsilon_0(N)$ will converge.

3. Proof of Theorem 2.2

Since $\Delta_N(x,y,t,N)$ is a strictly sub-Gaussian stochastic process, we apply the result of the paper [6], and conclude that

$$P\left\{\sup_{(x,y,t)\in D}|\Delta_N(x,y,t,N)| > \delta\right\} \leq 2\widetilde{A}(\delta,\theta), \tag{3.1}$$

for all $0 < \theta < 1$, where

$$\widetilde{A}(\delta,0) = \exp\left\{-\frac{\left(\delta(1-\theta)-\frac{2}{\theta}I(\theta\epsilon_0)\right)^2}{2\epsilon_0^2}\right\}, \tag{3.2}$$

ϵ_0 is an arbitrary number such that

$$\epsilon_0 \geq \sup_{(x,y,t)\in D}\left(E|\Delta_N(x,y,t,N)|^2\right)^{1/2},$$

$$I(\theta\epsilon_0) = \frac{1}{\sqrt{2}} \int_0^{\theta\epsilon_0} \left(\ln\left(\frac{p}{2\sigma^{-1}(x)} + 1\right) + \ln\left(\frac{q}{2\sigma^{-1}(x)} + 1\right) + \ln\left(\frac{T}{2\sigma^{-1}(x)} + 1\right)\right)^{1/2} dx, \tag{3.3}$$

and where $\sigma(h)$ is a continuous increasing function such that $\sigma(0) = 0$ and

$$\sup_{\substack{|x-x_1|\le h\\|y-y_1|\le h\\|t-t_1|\le h}} \left(E|\Delta_N(x,y,t,N) - \Delta_N(x_1,y_1,t_1,N)|^2\right)^{1/2} \le \sigma(h),$$

$$\sup_{\substack{|x-x_1|\le h\\|y-y_1|\le h\\|t-t_1|\le h}} \left(E|u_N(x,y,t) + V_N(x,y,t) - u_N(x_1,y_1,t_1) - V_N(x_1,y_1,t_1)|^2\right)^{1/2}$$

$$\le \sup_{\substack{|x-x_1|\le h\\|y-y_1|\le h\\|t-t_1|\le h}} \left[\left(E|u_N(x,y,t) - u_N(x_1,y_1,t_1)|^2\right)^{1/2} + \left(E|V_N(x,y,t) - V_N(x_1,y_1,t_1)|^2\right)^{1/2}\right],$$

$$\sup_{(x,y,t)\in D} \left(E|\Delta_N(x,y,t,N)|^2\right)^{1/2} = \sup_{(x,y,t)\in D} \left(E|u_N(x,y,t) + V_N(x,y,t)|^2\right)^{1/2}$$

$$\le \sup_{(x,y,t)\in D} \left[\left(E|u_N(x,y,t)|^2\right)^{1/2} + \left(E|V_N(x,y,t)|^2\right)^{1/2}\right].$$

Since the stochastic processes $\xi(x,y)$ and $\eta(x,y)$ are independent, that is, a_{nm} and b_{nm} are independent, we obtain

$$E|u_N(x,y,t) - u_N(x_1,y_1,t_1)|^2$$

$$= E\left|\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty} V_{nm}(x,y)\left[a_{nm}\cos\sqrt{\lambda_{nm}}t + b_{nm}\sin\sqrt{\lambda_{nm}}t\right] - \sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty} V_{nm}(x_1,y_1)\left[a_{nm}\cos\sqrt{\lambda_{nm}}t_1 + b_{nm}\sin\sqrt{\lambda_{nm}}t_1\right]\right|^2$$

$$= E\left|\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty} a_{nm}\frac{2}{\sqrt{pq}}\left[\sin\frac{n\pi}{p}x\sin\frac{m\pi}{q}y\cos\sqrt{\lambda_{nm}}t - \sin\frac{n\pi}{p}x_1\sin\frac{m\pi}{q}y_1\cos\sqrt{\lambda_{nm}}t_1\right]\right.$$
$$\left. + \sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty} b_{nm}\frac{2}{\sqrt{pq}}\left[\sin\frac{n\pi}{p}x\sin\frac{m\pi}{q}y\sin\sqrt{\lambda_{nm}}t - \sin\frac{n\pi}{p}x_1\sin\frac{m\pi}{q}y_1\sin\sqrt{\lambda_{nm}}t_1\right]\right|^2$$

$$\leq \frac{4}{pq} \sum_{n=N+1}^{\infty} \sum_{m=N+1}^{\infty} \sum_{k=N+1}^{\infty} \sum_{l=N+1}^{\infty} |Ea_{nm}a_{kl}| \cdot \left|\sin \frac{n\pi}{p} x \sin \frac{m\pi}{q} y \cos \sqrt{\lambda_{nm}} t\right.$$

$$\left. - \sin \frac{n\pi}{p} x_1 \sin \frac{m\pi}{q} y_1 \cos \sqrt{\lambda_{nm}} t_1 \right| \cdot \left| \sin \frac{k\pi}{p} x \sin \frac{l\pi}{q} y \cos \sqrt{\lambda_{kl}} t \right.$$

$$\left. - \sin \frac{k\pi}{p} x_1 \sin \frac{l\pi}{q} y_1 \cos \sqrt{\lambda_{kl}} t_1 \right|$$

$$+ \frac{4}{pq} \sum_{n=N+1}^{\infty} \sum_{m=N+1}^{\infty} \sum_{k=N+1}^{\infty} \sum_{l=N+1}^{\infty} |Eb_{nm}b_{kl}| \cdot \left|\sin \frac{n\pi}{p} x \sin \frac{m\pi}{q} y \sin \sqrt{\lambda_{nm}} t\right.$$

$$\left. - \sin \frac{n\pi}{p} x_1 \sin \frac{m\pi}{q} y_1 \sin \sqrt{\lambda_{nm}} t_1 \right| \cdot \left| \sin \frac{k\pi}{p} x \sin \frac{l\pi}{q} y \sin \sqrt{\lambda_{kl}} t \right.$$

$$\left. - \sin \frac{k\pi}{p} x_1 \sin \frac{l\pi}{q} y_1 \sin \sqrt{\lambda_{kl}} t_1 \right|$$

$$= \frac{4}{pq} \left(\sum_{n=N+1}^{\infty} \sum_{m=N+1}^{\infty} \sqrt{Ea_{nm}^2} \left| \sin \frac{n\pi}{p} x \sin \frac{m\pi}{q} y \cos \sqrt{\lambda_{nm}} t \right. \right.$$

$$\left. \left. - \sin \frac{n\pi}{p} x_1 \sin \frac{m\pi}{q} y_1 \cos \sqrt{\lambda_{nm}} t_1 \right| \right)^2$$

$$+ \frac{4}{pq} \left(\sum_{n=N+1}^{\infty} \sum_{m=N+1}^{\infty} \sqrt{Eb_{nm}^2} \left| \sin \frac{n\pi}{p} x \sin \frac{m\pi}{q} y \sin \sqrt{\lambda_{nm}} t \right. \right.$$

$$\left. \left. - \sin \frac{n\pi}{p} x_1 \sin \frac{m\pi}{q} y_1 \sin \sqrt{\lambda_{nm}} t_1 \right| \right)^2 .$$

It is easy to check that

$$\left| \sin \frac{n\pi}{p} x \sin \frac{m\pi}{q} y \cos \sqrt{\lambda_{nm}} t - \sin \frac{n\pi}{p} x_1 \sin \frac{m\pi}{q} y_1 \cos \sqrt{\lambda_{nm}} t_1 \right|$$

$$\leq \left| \sin \frac{n\pi}{p} x - \sin \frac{n\pi}{p} x_1 \right| + \left| \sin \frac{m\pi}{q} y - \sin \frac{m\pi}{q} y_1 \right| + \left| \cos \sqrt{\lambda_{nm}} t - \cos \sqrt{\lambda_{nm}} t_1 \right|$$

$$\leq 2 \left| \sin \frac{\frac{n\pi}{p}(x - x_1)}{2} \right| + 2 \left| \sin \frac{\frac{m\pi}{q}(y - y_1)}{2} \right| + 2 \left| \sin \frac{\sqrt{\lambda_{nm}}(t - t_1)}{2} \right|$$

$$\leq \frac{n\pi}{p} h + \frac{m\pi}{q} h + \sqrt{\lambda_{nm}} h = \pi h \left(\frac{n}{p} + \frac{m}{q} + \sqrt{\frac{n^2}{p^2} + \frac{m^2}{q^2}} \right)$$

$$\leq 2\pi h \left(\frac{n}{p} + \frac{m}{q} \right) = 2\pi h \left(\frac{nq + pm}{pq} \right).$$

Similarly

$$\left| \sin \frac{n\pi}{p} x \sin \frac{m\pi}{q} y \sin \sqrt{\lambda_{nm}} t - \sin \frac{n\pi}{p} x_1 \sin \frac{m\pi}{q} y_1 \sin \sqrt{\lambda_{nm}} t_1 \right|$$

$$\leq 2\pi h\left(\frac{n}{p}+\frac{m}{q}\right)=\frac{2\pi h}{pq}(nq+mp).$$

Then

$$\begin{aligned}
&\left(E|u_N(x,y,t)-u_N(x_1,y_1,t_1)|^2\right)^{1/2}\\
&\leq\left(\frac{4}{pq}\left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Ea_{nm}^2}\left(\frac{2\pi h}{pq}(nq+mp)\right)\right)^2\right.\\
&\quad\left.+\frac{4}{pq}\left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Eb_{nm}^2}\left(\frac{2\pi h}{pq}(nq+mp)\right)\right)^2\right)^{1/2}\\
&=\frac{4\pi h}{p^{\frac{3}{2}}q^{\frac{3}{2}}}\left(\left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Ea_{nm}^2}(nq+mp)\right)^2\right.\\
&\quad\left.+\left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Eb_{nm}^2}(nq+mp)\right)^2\right)^{1/2}.
\end{aligned}\tag{3.4}$$

We also have

$$\begin{aligned}
&\left(E|V_N(x,y,t)-V_N(x_1,y_1,t_1)|^2\right)^{1/2}\\
&\leq\frac{4\pi h}{p^{\frac{3}{2}}q^{\frac{3}{2}}}\left(\left(\sum_{n=1}^{N}\sum_{m=1}^{N}\sqrt{E(\widehat{a}_{nm}-a_{nm})^2}(nq+mp)\right)^2\right.\\
&\quad\left.+\left(\sum_{n=1}^{N}\sum_{m=1}^{N}\sqrt{E(\widehat{b}_{nm}-b_{nm})^2}(nq+mp)\right)^2\right)^{1/2}.
\end{aligned}$$

One can easily obtain that

$$\begin{aligned}
E(\widehat{a}_{nm}-a_{nm})^2&=E\left(\int_0^p\int_0^q\left(\widehat{\xi}(x,y)-\xi(x,y)\right)V_{nm}(x,y)dxdy\right)^2\\
&=E\left(\frac{2}{\sqrt{pq}}\int_0^p\int_0^q\left(\widehat{\xi}(x,y)-\xi(x,y)\right)\sin\frac{n\pi}{p}x\sin\frac{m\pi}{q}y\right)^2\\
&\leq\left(\frac{2}{\sqrt{pq}}\int_0^p\int_0^q\sqrt{E\left(\widehat{\xi}(x,y)-\xi(x,y)\right)^2}dxdy\right)^2\leq 4\Lambda^2.
\end{aligned}\tag{3.5}$$

Similarly

$$E(\widehat{b}_{nm}-b_{nm})^2=4\Lambda^2\frac{p^2q^2}{\pi^2(n^2q^2+m^2p^2)}.\tag{3.6}$$

Then

$$\left[|V_N(x,y,t)-V_N(x_1,y_1,t_1)|^2\right]^{1/2} \le \frac{4\pi h}{p^{3/2}q^{3/2}}\left(\left(\sum_{n=1}^{N}\sum_{m=1}^{N} 2\Lambda(nq+mp)\right)^2 + \left(\sum_{n=1}^{N}\sum_{m=1}^{N} 2\Lambda\frac{pq}{\pi\sqrt{n^2q^2+m^2p^2}}(nq+mp)\right)^2\right)^{1/2}. \tag{3.7}$$

Thus we obtain from (3.4), (3.5), (3.6) and (3.7) that $\sigma(h)=hA_N$, where

$$A_N = \frac{2\pi}{p^{3/2}q^{3/2}}\left\{\left(\left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Ea_{nm}^2}(nq+mp)\right)^2 + \left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Eb_{nm}^2}(nq+mp)\right)^2\right)^{1/2} + 2\Lambda\left(\left(\sum_{n=1}^{N}\sum_{m=1}^{N}(nq+mp)\right)^2 + \left(\sum_{n=1}^{N}\sum_{m=1}^{N}\frac{pq}{\pi\sqrt{n^2q^2+m^2p^2}}(nq+mp)\right)^2\right)^{1/2}\right\}.$$

It is easy to see that

$$\begin{aligned}
&E|u_N(x,y,t)|^2\\
&\le E\left|\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}V_{nm}(x,y)\left[a_{nm}\cos\sqrt{\lambda_{nm}}t+b_{nm}\sin\sqrt{\lambda_{nm}}t\right]\right|^2\\
&=\left|\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sum_{k=N+1}^{\infty}\sum_{l=N+1}^{\infty}V_{nm}(x,y)V_{kl}(x,y)\left[Ea_{nm}a_{kl}\cos\sqrt{\lambda_{nm}}\cos\sqrt{\lambda_{kl}}t\right.\right.\\
&\quad\left.\left.+\ Eb_{nm}b_{kl}\sin\sqrt{\lambda_{nm}}\sin\sqrt{\lambda_{kl}}t\right]\right|\\
&\le\frac{4}{pq}\left(\left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Ea_{nm}^2}\right)^2+\left(\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{Eb_{nm}^2}\right)\right)
\end{aligned} \tag{3.8}$$

and

$$\begin{aligned}
&E|V_N(x,y,t)|^2\\
&\le\frac{4}{pq}\left(\left(\sum_{n=1}^{N}\sum_{m=1}^{N}\sqrt{E\left(\widehat{a}_{nm}-a_{nm}\right)^2}\right)^2+\left(\sum_{n=1}^{N}\sum_{m=1}^{N}\sqrt{E\left(\widehat{b}_{nm}-b_{nm}\right)^2}\right)^2\right).
\end{aligned} \tag{3.9}$$

Thus we obtain from (3.8) and (3.9) that

$$\epsilon_0(N) = \frac{4}{\sqrt{pq}} \left\{ \left(\left(\sum_{n=N+1}^{\infty} \sum_{m=N+1}^{\infty} \sqrt{a_{nm}^2} \right)^2 + \left(\sum_{n=N+1}^{\infty} \sum_{m=N+1}^{\infty} \sqrt{b_{nm}^2} \right)^2 \right)^{1/2} \right.$$
$$\left. + \Lambda \left(N^4 + \left(\frac{1}{\pi^2} \sum_{n=1}^{N} \sum_{m=1}^{N} \frac{pq}{\sqrt{n^2p^2 + m^2q^2}} \right)^2 \right)^{1/2} \right\}.$$

Substituting these values of $\sigma(h)$ and $\epsilon_0(N)$ in equality (3.3), we get for $z = \theta\epsilon_(N)$ that

$$I(z) = \frac{1}{\sqrt{2}} \int_0^z \left(\ln\left(\frac{pA_N}{2x} + 1\right) + \ln\left(\frac{qA_N}{2x} + 1\right) + \ln\left(\frac{TA_N}{2x} + 1\right) \right)^{1/2} dx$$
$$\leq \frac{1}{\sqrt{2}} \int_0^z \left(\ln\left(\frac{pA_N}{2x} + 1\right) \right)^{1/2} dx + \frac{1}{\sqrt{2}} \int_0^z \left(\ln\left(\frac{qA_N}{2x} + 1\right) \right)^{1/2} dx$$
$$+ \frac{1}{\sqrt{2}} \int_0^z \left(\ln\left(\frac{TA_N}{2x} + 1\right) \right)^{1/2} dx$$
$$\leq \left[\int_0^z \left(\frac{pA_N}{2x}\right)^{1/2} dx + \int_0^z \left(\frac{qA_N}{2x}\right)^{1/2} dx + \int_0^z \left(\frac{TA_N}{2x}\right)^{1/2} dx \right]$$
$$= \left(T^{1/2} + p^{1/2} + q^{1/2}\right) A_N^{1/2} z_0^{1/2}(N),$$

Then equality (3.2) can be rewritten as

$$\widetilde{A}(\delta,\theta) \leq exp\left\{ -\frac{1}{2} \left(\frac{\delta(1-\theta) - \frac{2}{\theta^{1/2}} \left(T^{1/2} + p^{1/2} + q^{1/2}\right) A_N^{1/2} \epsilon_0^{\frac{1}{2}}(N)}{\epsilon_0(N)} \right)^2 \right\}.$$

If

$$\left(T^{1/2} + p^{1/2} + q^{1/2}\right) A_N^2 \epsilon_0^2(N) < \delta,$$

then $\widetilde{A}(\delta,\theta)$ attains its minimum at

$$\theta = \frac{\left(T^{1/2} + p^{1/2} + q^{1/2}\right)^{2/3} A_N^{1/3} \epsilon_0^{1/3}(N)}{\delta^{2/3}}.$$

Namely

$$\min_\theta \widetilde{A}(\delta,\theta)$$

$$= \exp\left\{ -\frac{1}{2}\left(\frac{\delta^{1/3}\left(\delta^{2/3} - 3\left(T^{1/2}+p^{1/2}+q^{1/2}\right)^{\frac{2}{3}} A_N^{1/3}\epsilon_0^{1/2}(N)\right)}{\epsilon_0(N)}\right)^2\right\} \geq \ln\frac{1}{\gamma},$$

Therefore, given an accuracy δ, one can construct a model with reliability $1-\gamma$ if

$$\left(T^{1/2}+p^{1/2}+q^{1/2}\right) A_N^2\epsilon_0^2(N) < \delta,$$

$$\frac{1}{2}\left(\frac{\delta^{1/3}\left(\delta^{2/3} - 3\left(T^{1/2}+p^{1/2}+q^{1/2}\right) A_N^{1/3}\epsilon_0^N\right)^2}{\epsilon_0(N)}\right)^2 \geq \ln\frac{1}{\gamma}.$$

4. Example

Let $\eta(x,y)=0$, $p=q=\pi$, $T=\pi$, then the solution of problem (2.1)–(2.3) may be represented as:

$$u(x,y,t) = \frac{2}{\pi}\sum_{n=1}^{\infty}\sum_{m=1}^{N} a_{nm}\sin nx \sin my \cos\left(\sqrt{n^2+m^2}t\right).$$

Let's construct a model of the solution in the form:

$$\hat{u}^N(x,y,t) = \frac{2}{\pi}\sum_{n=1}^{N}\sum_{m=1}^{N} \hat{a}_{nm}\sin nx \sin my \cos\left(\sqrt{n^2+m^2}t\right).$$

Let the assumptions of Theorem 2.2 hold and let $\xi(x,y)$ be a Gaussian stochastic process such that

$$\xi(x,y) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty} \xi_{ij}\sin(i\,(x))\sin(j\,(y)),$$

where ξ_{ij} are independent Gaussian random variable with $E\xi_{ij}=0$, $E\xi_{ij}^2 = d^{ij}$. Here d^{ij} is a number such that $0<d^{ij}<1$. Let

$$\hat{\xi}(x,y) = \hat{\xi}_M(x,y) = \sum_{i=1}^{M}\sum_{i=1}^{M} \xi_{ij}\sin(i\,(x))\sin(j\,(y)).$$

Then

$$E\left(\xi(x,y)-\hat{\xi}_M(x,y)\right)^2 = \sum_{i=M+1}^{\infty}\sum_{j=M+1}^{\infty} d^{ij}\sin^2(i\,(x))\sin^2(j\,(y))$$

$$\leq \sum_{i=M+1}^{\infty}\sum_{j=M+1}^{\infty} d^{ij} = \sum_{i=M+1}^{\infty}\frac{d^{i(M+1)}}{1-d^i} \leq \frac{1}{1-d}\sum_{i=M+1}^{\infty} d^{i(M+1)}$$

$$= \frac{1}{1-d} \cdot \frac{d^{(M+1)^2}}{1-d^{M+1}} \leq \frac{d^{(M+1)^2}}{(1-d)^2}.$$

Note that given Λ, we chose M such that

$$\frac{1}{\pi}\int_0^\pi\int_0^\pi \sqrt{E\left(\xi(x,y)-\hat{\xi}_M(x,y)\right)^2}dxdy \leq \pi\sqrt{\frac{d^{(M+1)^2}}{(1-d)^2}} < \Lambda,$$

$$\frac{d^{(M+1)^2}}{(1-d)^2} < \frac{\Lambda^2}{\pi^2}$$

therefore,

$$M \geq \sqrt{\frac{\ln\left(\frac{\lambda^2}{\pi^2}(1-d)^2\right)}{\ln d}}.$$

In this case $b_{nm} = 0$,

$$a_{nm} = \int_0^\pi\int_0^\pi \xi(x,y)V_{nm}(x,y)dxdy = \frac{2}{\pi}\int_0^\pi\int_0^\pi \xi(x,y)\sin nx \sin my dxdy = 2\pi\xi_{nm},$$

that $Ea_{nm}^2 = 4\pi^2 d^{nm}$.

$$\hat{u}^N(x,y,t) = \frac{2}{\pi}\sum_{n=1}^{N}\sum_{m=1}^{N}\hat{a}_{nm}\sin nx \sin my \cos\left(\sqrt{n^2+m^2}t\right).$$

Thus

$$A_N = \frac{2}{\pi}\left\{2\pi\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{d^{nm}}\,(n+m) + \Lambda\sum_{n=1}^{N}\sum_{m=1}^{N}(n+m)\right\}$$

$$\leq \frac{2}{\pi}\left\{\frac{4\pi d^{\frac{(N+1)^2}{2}}\left(1+N+Nd^{\frac{N+1}{2}}\right)}{(1-d)(d^{\frac{N+1}{2}})^2} + \Lambda(1+N)N^2\right\},$$

$$\epsilon_0 = \frac{4}{\pi^2}\left\{2\pi\sum_{n=N+1}^{\infty}\sum_{m=N+1}^{\infty}\sqrt{d^{nm}} + \Lambda N^2\right\} \leq \frac{4}{\pi}\left\{\frac{2\pi d^{\frac{(N+1)2}{2}}}{(1-d)(1-d^{\frac{N+1}{2}})} + \Lambda N^2\right\}.$$

So, we have received the model, where N and Λ satisfy the following inequality

$$A_N^2\epsilon_0^2(N) < \frac{\delta}{3\sqrt{\pi}},$$

$$\left(\frac{\delta^{1/3}\left(\delta^{2/3} - (243\pi A_N\epsilon_0(N))^{1/3}\right)}{\epsilon_0(N)}\right)^2 \geq 2\ln\left(\frac{1}{\gamma}\right).$$

When some $\Lambda = 0.005$ and $N = 36$ the model $\hat{u}^N(x,y,t)$ approaches the random process $u(x,y,t)$ to reliability 0.99 and accuracy 0.01 in the uniform metric.

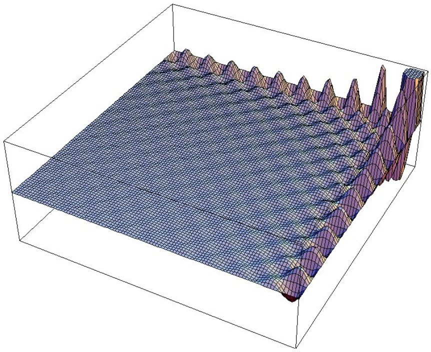


Figure 1: The model of membrane's vibration at the moment of time $t = 0$

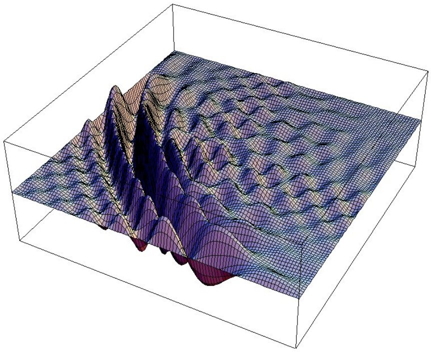


Figure 2: The model of membrane's vibration at the moment of time $t = 1$

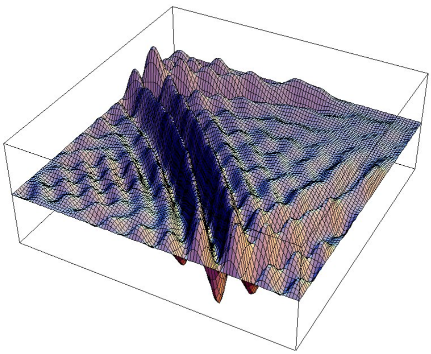


Figure 3: The model of membrane's vibration at the moment of time $t = 2$

References

[1] V. V. Buldygin and Yu. V. Kozachenko (2000), Metric Characterization of Random Variables and Randomprocesses. American Mathematical Society, Providence, Rhode.

[2] B. V. Dovgay, Yu. V. Kozachenko, and G. I. Slyvka-Tylyshchak (2008), The boundary-value problems of mathematical physics with random factors, "Kyiv university", Kyiv. (Unkrainian)

[3] Yu. V. Kozachenko, A. O. Pashko, and I.V. Rozora (2007), Modelling Stochastic Processes and fields, "Kyiv university", Kyiv. (Ukrainian)

[4] Yu. V. Kozachenko and I.V. Rozora (2003), Simulation of Gaussian stochastic proceses, Random Oper. and Stochastic equation,Vol. 11, No. 3.

[5] Yu. V. Kozachenko and I.V. Rozora (2004), Simulation of Gaussian stochastic fields, Theory of Stochastic processes, Vol. 10 (26), No. 1-2.

[6] Yu. V. Kozachenko and G. I. Slyvka (2003), Justification of the Fourier method for hyperbplic equations with random initial conditions, Teor. Probab. and Matem. Statist. 69.

[7] Yu. V. Kozachenko and G. I. Slyvka (2006), Modelling a Solution of a Hyperbolic Equation With Random Initial Conditions, Teor. Probab. and Matem. Statist. 74.

[8] G. N. Polozhiy (1964), Equations of Mathematical Physics, "Vyshaya shkola", Moscow. (Russian)

[9] G. I. Slyvka and A. M. Tegza (2005), Modelling a Solution of Problem of Homogeneous String Vibrations with Random Initial Conditions, Naukovij Visnik Uzhgorod University, No. 10-11. (Ukrainian)