National Research University Higher School of Economics

Faculty of Mathematics

as a manuscript

Vladimir Panov

Non-classical methods of probabilistic and statistical analysis of the mixture models

Summary of the thesis
for the purpose of obtaining academic degree
Doctor of Science in Mathematics

Moscow - 2022

1 General description of the thesis

Relevance of the research topic

A mixture of probability distributions is a distribution of a random variable ξ, such that

$$\mathsf{P}\{\xi \in B\} = \int_A \mathsf{P}_{\vec{a}}(B) dG(\vec{a}), \qquad B \in \mathcal{B}(\mathbb{R}^d), \tag{1}$$

where $\mathsf{P}_{\vec{a}}$ is a parametric family of probability measures on $\left(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)\right)$, the mapping $\vec{a} \mapsto \mathsf{P}_{\vec{a}}(B)$ is measurable for any $B \in \mathcal{B}(\mathbb{R}^d)$, $A \subset \mathbb{R}^k$ is a set of possible values of the parameter $\vec{a}$, and G is the distribution function of this parameter. The family $\mathsf{P}_{\vec{a}}$ (and the distribution of ξ) may depend also on some other parameters, but we omit this dependence here for simplicity. If the set A consists of a finite number of elements, that is, $A = \{\vec{a}_1, ..., \vec{a}_m\}$, then the mixture is called finite. In this case, (1) can be written as

$$\mathsf{P}\{\xi \in B\} = \sum_{i=1}^{m} \pi_i \mathsf{P}_{\vec{a}_i}(B) \qquad B \in \mathcal{B}(\mathbb{R}^d), \tag{2}$$

where $\pi_i \geq 0 \; \forall i = 1..m, \sum_{i=1}^{m} \pi_i = 1$.

For finite models (2), the standard task of mathematical statistics is to estimate the elements of the support $\vec{a}_i, i = 1..m$, and the mixing distribution $(\pi_1, ..., \pi_m)$ based on the observations of the random variable ξ. The method of moments was applied to this problem as early as 1894 by Pearson[1] for the case when $\mathsf{P}_{\vec{a}}$ is a family of one-dimensional Gaussian distributions and $m = 2$. Classical statistical methods for finite mixtures (method of moments, method of maximum likelihood, Bayesian approaches), as well as the issues of model identifiability, were quite well studied in the 60 - 70th years of the last century.[2] Nowadays, the most popular approach for the parameter estimation in finite mixture models is the EM algorithm, which is a method for solving an optimisation problem that arises when calculating the maximum likelihood estimate.[3]

[1] Pearson, K. Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society of London.* Series A, 185: 71–110, 1894.
[2] Gupta, S. and Huang, W.-T. On mixtures of distributions: a survey and some new results on ranking and selection. *Sankhya: The Indian Journal of Statistics*, Series B, 245–290, 1981.
[3] McLachlan, G. and Krishnan, T. *The EM Algorithm and Extensions.* John Wiley & Sons, 2007.

The study of the probabilistic and statistical properties of the mixture model is a popular area of stochastic analysis. Let us mention that over the past 15 years, 6 PhD theses on the topics in this field have been defended at the Moscow State University.[4] The relevance of this topic is also confirmed by a large number of publications dealing with the applications of mixture models in finance, astronomy, image analysis, genomics, and in many other areas.[5] In fact, mixture models can be applied to the analysis of any data for which statistical clustering and classification tasks are relevant.[6]

At the same time, the existing methods of statistical analysis are not applicable to some new problems that have arisen in the literature recently. Let us list the main mathematical problems considered in this thesis.

1. Semiparametric estimation in mixture models. This direction is concentrated on the estimation of an unknown parameters of the family $\mathsf{P}_{\vec{a}}$ and an unknown absolutely continuous distribution G based on the observations from (1). This problem is well studied for the case when the distribution class G is given parametrically - for example, for the case of generalized hyperbolic distributions.[7] The model is of great interest for applications - in particular, in can be used for modeling the sizes of sand[8] and the sizes of diamonds in the deposits of south-west Africa.[9] More general semiparametric case was considered by Korsholm,[10] but its practical implementation meets serious computational difficulties, since one would need to solve rather

[4] Кокшаров, С.Н. *Асимптотические свойства смесей вероятностных распределений,* МГУ им. М.В. Ломоносова, 2007 (научный руководитель - Королёв В.Ю.).
Назаров, А.Л. *Приближенные методы разделения смесей вероятностных распределений,* МГУ им. М.В Ломоносова, 2013 (научный руководитель - Королёв В.Ю.).
Савинов Е.А. *Асимптотические свойства условных распределений непрерывных смесей,* МГУ им. М.В. Ломоносова, 2009 (научный руководитель - Шатских С.Я.).
Крылов В.А. *О некоторых свойствах смесей обобщенных гамма-распределений и их применениях,* МГУ им. М.В. Ломоносова, 2011 (научный руководитель - Матвеев В.Ф.).
Горшенин А.К. *Асимптотические свойства статистических процедур анализа смесей вероятностных распределений,* МГУ им. М.В. Ломоносова, 2011 (научный руководитель - Королёв В.Ю.).
Корчагин А.Ю. *Прогнозирование стохастических процессов с помощью сеточного метода разделения дисперсионно-сдвиговых смесей нормальных законов,* МГУ им. М.В. Ломоносова, 2015 (научный руководитель - Королёв В.Ю.).
[5] Frühwirth-Schnatter, S., Celeux, G., and Robert, C. *Handbook of Mixture Analysis.* CRC press, 2019.
[6] McNicholas, P. *Mixture Model-based Classification.* Chapman and Hall/CRC, 2016.
[7] Jørgensen B. *Statistical Properties of the Generalized Inverse Gaussian Distribution.* New York: Springer-Verlag, 1982.
[8] Barndorff-Nielsen O., Christensen C. Erosion, deposition, and size distributions of sand. *Proceedings of the Royal Society of London. Series A, Mathematical and physical sciences.* 417(1853): 335–352, 1988.
[9] Barndorff-Nielsen O. Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society. Series A.* 353(1674): 401–419, 1977.
[10] Korsholm L. The semiparametric normal variance-mean mixture model. *Scandinavian Journal of Statistics.* 27(2): 227–261, 2000

challenging optimisation problem. In this thesis, we propose an essentially new approach based on the properties of the superposition of Mellin and Laplace transforms. The method does not employ any parametric restrictions on G, and it can be applied for the estimation of any absolutely continuous mixing distribution.

2. Stochastic time-changed models. The classical task of financial mathematics is to construct the models that can realistically describe the jump-type dynamics of financial time series (for example, the dynamics of stocks returns). The majority of known approaches are based on the class of Lévy processes, which can be considered as the generalisation of the classical Black-Scholes model to the case when the trajectories are discontinuous. This class is often used for the construction of more complex models possessing the so-called stylised features of financial data.[11] The most popular constructions are the stochastic volatility models and stochastic time-changes in the Lévy processes.[12] Since the Lévy-based models are often constructed based on several stochastic processes, it would be natural to ask whether the important characteristics of these processes can be recovered from the observations of the model. For instance, one of these characteristics is the Blumenthal-Getoor index, which indicates the jump activity of the process.[13]

In the one-dimensional case, the concept of stochastic time change is that for some random process L_t, $t \geq 0$, the deterministic time t is replaced by a non-decreasing non-negative random process $\mathcal{T}(s)$, $s \geq 0$, which plays the role of random time. The considered models are closely related to mixtures: in fact, the distribution of the time-changed process is a mixture of probability distributions, and the Lévy measure characterising the behaviour of jumps is a mixture of (non-probabilistic) measures. A significant technical difficulty lies in the type of the data - while the results for the high-frequency data (that is, for the case when the time step between two consecutive observations tends to zero) were obtained in the papers by Aït-Sahalia and Jacod,[14] the case of low-frequency data (fixed time interval, but an infinite horizon), which is considered

[11] Cont, R. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance.* 1: 223-236, 2001.

[12] Barndorff-Nielsen, O. and Shiryaev, A. *Change of Time and Change of Measure.* World Scientific Publishing Company, 2015.

[13] Rosenbaum, M. and Tankov, P. Asymptotically optimal discretisation of hedging strategies with jumps. *The Annals of Applied Probability.* 24.3: 1002-1048, 2014.

[14] Aït-Sahalia, Y. and Jacod, J. Estimating the degree of activity of jumps in high frequency financial data. *The Annals of Statistics.* 37: 2202–2244, 2009.
Aït-Sahalia, Y. and Jacod, J. Identifying the succesive Blumenthal- Getoor indices of a discretely observed process. *The Annals of Statistics.* 40: 1430–1464, 2012.

in this thesis, was not studied earlier.[15]

3. Construction of nonparametric confidence sets for the density function. The most interesting problem in this field is to construct the $(1-\alpha)-$confidence sets that are *honest* with respect to a certain class of densities (for example, to a class of mixtures of absolutely continuous distributions) in the sense that the probability that a true function from this class belongs to a confidence set is larger than $(1-\alpha)$ uniformly on the class. Typically, the construction of confidence sets is based on the so-called SBR-type (Smirnov - Bickel - Rosenblatt) limit theorems, which yield the asymptotic behaviour of the maximal deviation of the considered density estimates.[16] Despite the long history of studying the question, theorems of this type are known only for kernel density estimates and projection estimates on some types of basis (the Haar and the Battle-Lemarie wavelets). This issue is confirmed by a number of recent articles on this topic published in the the Annals of Statistics.[17] In this thesis, we consider the construction of the honest confidence intervals based on the projection estimates, when using the basis of Legendre polynomials. The solution of this problem relies on some special asymptotic properties of non-stationary Gaussian processes, which are well studied for the special case of the nonstationarity - the so-called cyclostationarity,[18] but were not previously known in more general cases. Results of this kind are of particular interest for the analysis of mixtures of distributions, since classical methods for the construction of confidence sets in this case (for example, the method based on the kernel density estimates) lead to inadequate results.

4. Limit laws and phase transitions in mixture models. The general theory of the limit distributions for the sums and for the maxima of random variables is well-described in the brilliant books by Petrov[19] and Embrechts, Klüppelberg and Mikosch.[20] Nevertheless, the analysis of the limiting distribution in particular model can be rather tricky. Note, for example, that the limit

[15] Belomestny, D., and Reiss, M. Estimation and Calibration of Lévy models via Fourier methods. In: *Lévy Matters IV*. Springer, Cham, 1-76, 2015.

[16] Smirnov, N. V. On the construction of confidence regions for the density of distribution of random variables. *Doklady Akad. Nauk SSSR*. 74: 189–191, 1950.
Bickel, P, and Rosenblatt, M. On some global measures of the deviations of density function estimates. *The Annals of Statistics*. 1(6):1071–1095, 1973.

[17] Giné, E., Nickl, R. Confidence bands in density estimation. *The Annals of Statistics*. 38(2): 1122–1170, 2010.
Chernozhukov, V., Chetverikov, D., Kato, K. Anti-concentration and honest, adaptive confidence bands. *The Annals of Statistics*. 42(5): 1787–1818, 2014.

[18] Konstant, D., Piterbarg, V. Extreme values of the cyclostationary Gaussian random process. *Journal of Applied Probability*. 82–97, 1993.

[19] Petrov, V. *Sums of Independent Random Variables*, Vol. 82. Springer Science & Business Media, 2012.

[20] Embrechts, P., Klüppelberg, C., and Mikosch, T. *Modelling Extremal Events for Insurance and Finance*. Springer, 1997.

laws for the random energy model (REM) introduced by Derrida[21] in the beginning of 80-th, were fully described only 20 years later.[22] It would be a worth mentioning that the probabilistic analysis of the REM model is closely related to the parabolic Anderson problem, since it leads to the study of the same random exponential sum. Asymptotic behaviour of this sum, as well as the concepts of intermittency and localisation, were studied in the papers by Molchanov and his coauthors.[23] In this thesis, we aim to describe the limit laws for a new model of the REM type with a mixture distribution of energy levels. Models of this type are also motivated by the parabolic Anderson problem with the potential having a mixture distribution.

5. Analysis of the "purity" of a distribution. Due to the Lebesque theorem, any probability measure P can be represented as the sum of three measures

$$\mathsf{P}(\cdot) = a_1\mathsf{P}_d(\cdot) + a_2\mathsf{P}_{ac}(\cdot) + a_3\mathsf{P}_{sc}(\cdot),$$

where $a_1 + a_2 + a_3 = 1$ with $a_i \geq 0\ \forall i$, and the measures $\mathsf{P}_d, \mathsf{P}_{ac}, \mathsf{P}_{sc}$ are, respectively, the measures of discrete, absolute continuous and singular distributions. For the solution of statistical estimation problems, it is important to know whether some of parameters a_1, a_2, a_3 are equal to zero - for example, if $a_3 = 0$ (there is no singular component), then the estimation methods are significantly simplified. The Jessen - Wintner theorem says that the sum of a.s. convergent random sums has pure type, that is, two out of three numbers a_1, a_2, a_3 are equal to 0.[24] However, for specific models, the determination of the type can be a rather challenging task. A classical example is the Erdös problem dealing with the Bernoulli convolution, which is defined as a series $\mathcal{Z} = \sum_{n=0}^{\infty} \pm\rho^n$, where the signs are chosen randomly with probabilities $1/2$, and $\rho \in (0, 1)$.[25] The

[21] Derrida, B. Random - energy model: Limit of a family of disordered models. *Physical Review B.* 45(2): 79 - 82, 1980.
Derrida, B. Random - energy model: An exactly solvable model of disordered systems, *Physical Review Models.* 24(5): 2613 - 2626, 1981

[22] Bovier, A., Kurkova, I. and Löwe, M. Fluctuations of the free energy in the REM and the p-spin SK models. *The Annals of Probability.* 30(2): 605–651, 2002.
Ben Arous, G., Bogachev L. and Molchanov, S. Limit theorems for sums of random exponentials. *Probability Theory and Related Fields.* 132(4): 579–612, 2005.

[23] Gärtner, J., König, W. and Molchanov, S. Geometric characterization of intermittency in the parabolic Anderson model. *The Annals of Probability.* 35(2): 439–499, 2007.
Ben Arous, G., Molchanov, S., and Ramirez, A. Transition from the annealed to the quenched asymptotics for a random walk on random obstacles. *The Annals of Probability.* 33(6): 2149–2187, 2005.

[24] See Proposition 27.18 from Sato, K.-I. *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, 1999.

[25] Erdös, P. On a family of symmetric Bernoulli convolutions. *American Journal of Mathematics.* 61(4):974–976, 1939.

term "Bernoulli convolution" is motivated by the fact that $\mathcal{Z}$ is a convolution of infinite number of measures of the type $(\delta_{-\rho^n} + \delta_{\rho^n})/2$. The most well-known result in this field is the fact that $\mathcal{Z}$ has an absolutely continuous distribution for almost all $\rho \in (1/2, 1)$.[26]. In this thesis, we study the question of the type of the discrete Dickman-Goncharov distribution. This question is of great interest, since the number of applications of the Dickman–Goncharov distribution is growing: new applications have appeared in mathematics (random walks on solvable groups, random graph theory) and also in biology (models of growth and evolution of unicellular populations), finance (theory of extreme phenomena in finance and insurance), physics (the model of random energy levels), and other fields.[27]

The purpose and objectives of the study

The aim of the study is the development of new methods of statistical estimation in mixture models, as well as the solution of the relevant probabilistic problems that arise while establishing the statistical properties of these methods.

Let us list the particular tasks solved within the framework of this study.

(1) To develop an algorithm for semiparametric estimation of unknown parameters and unknown mixing distribution for the variance-mean Gaussian mixtures. To apply these statistical approaches to other related models, such as the continuous-time moving average Lévy processes. To derive the convergence rates of the proposed estimates. To prove the exponential mixing property for the processes from the considered class.

(2) For the time-changed Lévy processes, to propose an estimation scheme for the Blumenthal-Getoor index of unknown Lévy processes. To prove the optimality of the proposed estimates in the minimax sense for a certain subclass of Lévy processes. For a multidimensional model of time-changed stable processes, to develop a method for the representation of the processes from this class in the form of infinite series. To provide the empirical evidence that this model is more appropriate for describing stock prices than the classical time-changed Brownian motion, at least if the cumulative number of transactions is used for a stochastic time change.

[26]Solomyak, B. On the random series $\sum \pm\lambda_n$ (an Erdös problem). *The Annals of Mathematics.* 611-625, 1995.
[27]Molchanov, S. and Panov, V. The Dickman–Goncharov distribution. *Russian Mathematical Surveys,* 75(6):1089–1132, 2020.

(3) To develop an algorithm for the construction of confidence intervals for the probability density function, which are honest with respect to a certain subclass of absolutely continuous distributions. To apply similar ideas for the construction of confidence intervals for the density of the Lévy measure.

(4) To prove the limit laws and describe the phase transitions in a random energy model with the distribution of energy levels modelled by the Gaussian mixtures.

(5) To determine which components are equal to zero in the representation of the distribution function of the discrete Dickman-Goncharov law in the form of a mixture of three distributions of different types.

Key results and the structure of the thesis

This section lists the main results of the study and discusses the novelty of the obtained results.

In **Chapter 1**, we present an algorithm for semiparametric estimation of unknown parameters and an unknown mixing distribution for variance-mean Gaussian mixtures. We show that the convergence rate of the proposed estimates is determined by the properties of the Mellin transform of the density of the mixing distribution. The novelty of the presented method consists in the employing the properties of the superposition of the Mellin and Laplace transforms. The proposed approach is a significant contribution to this topic, because, unlike the previously known estimation methods, it is not based on the solution of difficult optimisation problems.[28]

The presented method is also adapted for continuous-time moving average Lévy processes. The algorithm is developed for the statistical estimation of the Lévy measure and other parameters of the Lévy process from the observations of the model, which is an integral over this process. The algorithm is new and can be applied to a wide class of models known as the ambit fields.[29] The properties of exponential mixing are proved for processes from the considered class, and upper bounds for the constructed estimates are derived.

[28]Korsholm L. The semiparametric normal variance-mean mixture model. *Scandinavian Journal of Statistics.* 27(2): 227–261, 2000.

[29]Podolskij, M. Ambit fields: survey and new challenges. In: *XI Symposium on Probability and Stochastic Processes. Birkhäuser, Cham, 2015.*

Chapter 2 deals with the statistical inference for the time-changed Lévy processes. We present a method for the estimation of the Blumenthal-Getoor indices of the Lévy processes used in the construction of this model. The novelty of the method lies in the use of the asymptotic properties of the characteristic function of the process for large values of the argument. The rates of convergence of the proposed estimates are obtained, and it is proved that these rates are optimal in the minimax sense.

We also present a new approach to the joint description of the returns of several stocks, based on the multidimensional time-changed Lévy processes. For the cases of subordinated Brownian motions and subordinated stable processes, we show a new approach for representing the processes from the considered class in the form of infinite series. It is shown that this method can be effectively used for the study of the returns of stock prices, and the constructed two-dimensional models reproduce well the correlations between various stock returns.

Chapter 3 focuses on the construction of honest confidence sets for an (unknown) distribution density based on projection estimates. It is shown that this problem is equivalent to the analysis of the asymptotic behaviour of non-stationary Gaussian processes of a certain type. From the technical side, the main difficulty is to describe this asymptotical behaviour up to the second term, and to show that the error terms are of polynomial order. The efficiency of the obtained results is illustrated by the examples of some mixture models.

We present also similar results for the density estimates of the Lévy measure. The main result of this part is the presentation of the sequences of accompanying laws that approximate the distribution of the maximum deviation of the estimates with errors of polynomial order. It is shown that the rates of convergence given in previous papers on this topic[30] are in fact of logarithmic order.

In **Chapter 4**, we explain the connection between the Anderson parabolic problem and the model of random energy levels. For the case, when the energy levels are modelled by the Gaussian mixtures, we study the asymptotic behaviour of the system depending on the parameters and describe the phase transitions. The results are a contribution to the theory of random energy models, since similar results were previously known only for the classical case, when the energy

[30] Figueroa-López, J.E. Sieve-based confidence intervals and bands for Lévy densities. *Bernoulli.* 17(2), 643–670, 2011.

levels are modelled by standard Gaussian random variables.[31]

Chapter 5 contains some new results on the type of the generalised Dickman-Goncharov distribution. For the case when the geometric distribution is used in the construction of this model, we describe the relation to the Erdös problem for the Bernoulli convolution. It is shown that in this case, the generalised Dickman - Goncharov distribution is absolutely continuous for almost all values of the parameter from a certain interval.

Theoretical and practical significance

The obtained theoretical results significantly expand the methodology of the statistical estimation for the mixture models. At the same time, Lévy-based models (in particular, time-changed models) are widely used to describe the dynamics of stock returns, and therefore the developed methods can be used to solve financial problems related to trading on the stock exchange. For example, the method of estimating the Blumenthal - Getoor index can be used to determine the degree of reliability of a financial asset, while methods of modelling multidimensional processes that well reproduce the dependence between financial assets, are used for testing the trading strategies on the stock exchange.

The presented results for the Lévy processes - in particular, the methods used for estimating the Blumenthal-Getoor index and the methods for modelling the multidimensional processes (see Section 2.2) - were included by the author in the course "Modelling of jump-type processes in economics" at the Higher School of Economics.

Approbation of the established results

The main results of this thesis were presented at the following international conferences and seminars.

1. *Bernoulli-IMS World Congress in Probability and Statistics* (online conference, organized by Seoul National University, July 2021), talk "The Dickman-Goncharov distribution".

[31]BenArous, G., Bogachev, L. and Molchanov, S. Limit theorems for sums of random exponentials. *Probability Theory and Related Fields*. 132(4): 579–612, 2005.

2. *Extreme Value Analysis Conference - EVA2019* (Zagreb, Croatia, July 2019), talk "Extremes of Gaussian non-stationary processes and improved confidence bands for densities".
3. *International Workshop on Applied Probability - IWAP2018* (Budapest, Hungary, Juny 2018), talk "Multivariate subordination of stable processes".
4. *German Probability and Statistics Days - GPSD2018* (Freiburg, Germany, February 2018), talk "Multivariate subordination of stable processes".
5. *Conference on Ambit Fields and Related Topics* (Aarhus, Denmark, August 2017), talk "Low-frequency estimation for moving-average Lévy processes".
6. *Probability Seminar Essen* (Essen, Germany, June 2017), talk "Low-frequency estimation for moving-average Lévy processes".
7. *Extreme Value Analysis Conference – EVA2017* (Delft, the Netherlands, June 2017), talk "Distribution of maximal deviation for Lévy density estimators".
8. *World Congress in Probability and Statistics* (Toronto, Canada, July 2016), talk "Low-frequency estimation of continuous-time moving average Lévy processes".
9. *German Probability and Statistics Days - GPSD2016* (Bochum, Germany, March 2016), talk "Statistical inference for fractional Lévy processes and related models".
10. *Conference on Stochastic Processes and their Applications - SPA2015* (Oxford, UK, July 2015), talk "Generalized Ornstein-Uhlenbeck process: Mellin transform of the invariant distribution and statistical inference".
11. *European Meeting of Statisticians - EMS2015* (Amsterdam, the Netherlands, July 2015), talk "Statistical inference for exponential functionals of Lévy processes".
12. *Statistical Inference for Lévy Processes* (Leiden, the Netherlands, September 2014), talk "Maximal deviation distribution for projection estimates of Lévy densities".

Moreover, the author has given 10 talks at the seminar "Stochastic analysis and its applications in economics" at the Faculty of Mathematics of the Higher School of Economics (seminar is coordinated by Prof. V. Konakov and Prof. A. Kolesnikov) and delivered 12 talks at the workshops organised by the International laboratory of stochastic analysis and its applications of the Higher School of Economics (https://lsa.hse.ru/).

List of author's published papers on the topic of the thesis

The list consists of 12 papers, among them 6 papers are published in journals with quartiles Q1/Q2 in Web of Science ([2], [3], [4], [6], [9], [12]), and 11 papers - in journals with quartiles Q1/Q2 in Scopus. Of the published papers, three are single-authored, and two are with co-authors - graduate students at HSE.

[1] Belomestny, D. and Panov, V. Semiparametric estimation in the normal variance-mean mixture model. *Statistics*, 52(3):571–589, 2018.

https://www.tandfonline.com/doi/abs/10.1080/02331888.2018.1425865?journalCode=gsta20

[2] Belomestny, D., Panov, V. and Woerner, J. Low-frequency estimation of continuous-time moving average Lévy processes. *Bernoulli*, 25(2):902–931, 2019.

https://projecteuclid.org/journals/bernoulli/volume-25/issue-2/Low-frequency-estimation
-of-continuous-time-moving-average-L%C3%A9vy-processes/10.3150/17-BEJ1008.short

[3] Belomestny D., Panov V. Abelian theorems for stochastic volatility models with application to the estimation of jump activity. *Stochastic Processes and their Applications*. 123 (1): 15-44, 2013.

https://www.sciencedirect.com/science/article/pii/S0304414912001925

[4] Belomestny, D. and Panov, V. Estimation of the activity of jumps in time-changed Lévy models. *Electronic Journal of Statistics,* 7:2970–3003, 2013.

https://projecteuclid.org/journals/electronic-journal-of-statistics/volume-7/issue-none...

[5] Panov, V. Series representations for multivariate time-changed Lévy models. *Methodology and Computing in Applied Probability,* 19(1):97–119, 2017

https://link.springer.com/article/10.1007/s11009-015-9461-8

[6] Panov, V. and Samarin, E. Multivariate asset-pricing model based on subordinated stable processes. *Applied Stochastic Models in Business and Industry,* 35(4):1060–1076, 2019.

https://onlinelibrary.wiley.com/doi/10.1002/asmb.2446

[7] Panov, V. Some properties of the one-dimensional subordinated stable model. *Statistics and Probability Letters,* 146:80–84, 2019.

https://www.sciencedirect.com/science/article/abs/pii/S0167715218303493

[8] Konakov, V., Panov, V. and Piterbarg, V. Extremes of a class of non-stationary Gaussian processes and maximal deviation of projection density estimates. *Extremes,* 24(3):617–651, 2021.

https://link.springer.com/article/10.1007/s10687-020-00402-2

[9] Konakov, V. and Panov, V. Sup-norm convergence rates for Lévy density estimation. *Extremes,* 19(3):371–403, 2016.

https://link.springer.com/article/10.1007/s10687-020-00402-2

[10] Molchanov, S. and Panov, V. Limit theorems for the alloy-type random energy model. *Stochastics,* 91(5):754–772, 2019.

https://www.tandfonline.com/doi/full/10.1080/17442508.2018.1545841

[11] Panov, V. Limit theorems for sums of random variables with mixture distribution. *Statistics and Probability Letters,* 129:379 – 386, 2017.

https://www.sciencedirect.com/science/article/abs/pii/S0167715217302213

[12] Molchanov, S. and Panov, V. The Dickman–Goncharov distribution. *Russian Mathematical Surveys,* 75(6):1089–1132, 2020.

https://iopscience.iop.org/article/10.1070/RM9976/meta?casa_token=2HoQHwUbDx4AAAAA:
bMJWR2yeYoRJoZaZXdjCXVmbRTA9SryGJxF7ISQS-TP6vK2mojbjSaXms3rx0fT5tLV9l5ZYMCQsjyA

2 Summary of the main results

2.1 Semiparametric estimation methods

2.1.1 Variance-mean Gaussian mixtures

We will begin the review of the results of this work by presenting a new semiparametric estimation method for the mixture model (1) in the case of a Gaussian family of measures $\mathsf{P}_{\vec{a}}$ with an unknown absolutely continuous mixing distribution G. More precisely, Section 1.1 of the thesis considers the model

$$p(x;\mu,G) = \int_{\mathbb{R}_+} \varphi_{(\mu s,s)}(x)\,dG(s) = \int_{\mathbb{R}_+} \frac{1}{\sqrt{2\pi s}} \exp\left\{-\frac{(x-\mu s)^2}{2s}\right\} dG(s), \qquad x\in\mathbb{R},$$

where $\varphi_{(\mu s,s)}$ is the density of the normal distribution with mean μs and variance s (normal variance-mean mixture). In this study, it is assumed that both parameter μ, and the density g of the mixing distribution are unknown. We aim to statistically estimate these objects from a sample $X_1,...,X_n$ having a distribution with the density $p(\cdot,\mu,G)$.

The estimation of the parameter μ is based on the representation of this parameter in the form

$$\mu = \frac{1}{2x}\log\left(\frac{p(x;\mu,G)}{p(-x;\mu,G)}\right).$$

It follows from this representation that μ is the only zero of the function

$$W(\rho) := \mathsf{E}\left[e^{-\rho X}w(X)\right], \quad \rho\in\mathbb{R},$$

where $w:\mathbb{R}\to\mathbb{R}$ is a Lipschitz function, possessing the property

$$w(x)\le 0 \;\text{ for }\; x\ge 0, \qquad w(-x) = -w(x), \qquad \operatorname{supp}(w)\subset[-A,A]$$

for some $A>0$. The estimate of the parameter μ is defined as follows

$$\widehat{\mu}_n := \inf\{\rho>0:\, \widehat{W}_n(\rho)=0\}\wedge M,$$

where M is a constant such that $\mu \in [0, M/2)$, and $\widehat{W}_n(\rho) := n^{-1}\sum_{i=1}^n e^{-\rho X_i} w(X_i)$ is a nonparametric estimate of $W(\rho)$. The constant M is a technical parameter, which is used for the proof of the theoretical properties (see Theorem 2 below).

For the estimation of the unknown distribution G, we employ the Mellin transform of a function f defined as $\mathcal{M}[f](z) := \int_{\mathbb{R}_+} x^{z-1} f(x)dx, z \in \mathbb{C}$. The key fact, yielding the statistical estimation method presented in Section 1, is that the superposition of the Mellin and Laplace transforms of the density function g can be represented as

$$\mathcal{M}\left[\mathscr{L}[g]\right](z) = \int_{\mathbb{R}_+} \phi_X(u) \left[\psi_\mu(u)\right]^{z-1} \psi'_\mu(u) du, \qquad z \in \mathbb{C}, \tag{3}$$

where $\psi_\mu(u) = \mathrm{i}\mu u - u^2/2,\ u \in \mathbb{R}$, is the characteristic exponent of the normal law with mean μ and unit variance. Moreover, the superposition of integral transforms in the left part of (3) is related to the Mellin transform of this density by the explicit relation

$$\mathcal{M}\left[g\right](z) = \frac{\mathcal{M}\left[\mathscr{L}[g]\right](1-z)}{\Gamma(1-z)} \tag{4}$$

for any z with $\mathrm{Re}(z) \in (0,1)$. The idea of sequential estimation of the functions $\psi_\mu(u)$, $\mathcal{M}\left[\mathscr{L}[g]\right](z)$, $\mathcal{M}\left[g\right](z)$ follows from relations (3) and (4). The last step (estimation of g) is based on the inverse Mellin transform. The resulting estimate is equal to

$$\begin{aligned}\widehat{g}_{n,\gamma}(x) = \frac{1}{2\pi n}\sum_{k=1}^n \int_0^{V_n} \left[\int_0^{U_n} e^{-\mathrm{i}uX_k} \left[\overline{\psi_{\hat{\mu}_n}(u)}\right]^{-\gamma-\mathrm{i}v} \overline{\psi'_{\hat{\mu}_n}(u)} du\right] \cdot \frac{x^{-\gamma-\mathrm{i}v}}{\Gamma(1-\gamma-\mathrm{i}v)} dv \\ + \frac{1}{2\pi n}\sum_{k=1}^n \int_{-V_n}^{0} \left[\int_0^{U_n} e^{\mathrm{i}uX_k} \left[\psi_{\hat{\mu}_n}(u)\right]^{-\gamma-\mathrm{i}v} \psi'_{\hat{\mu}_n}(u) du\right] \cdot \frac{x^{-\gamma-\mathrm{i}v}}{\Gamma(1-\gamma-\mathrm{i}v)} dv,\end{aligned}$$

where U_n, V_n are infinitely increasing sequences of positive numbers and $\gamma \in (0,1)$. For the theoretical study, we use the estimate $\widehat{g}^{\circ}_{n,\gamma}(x)$, which is derived from $\widehat{g}_{n,\gamma}(x)$ by replacing the estimate $\hat{\mu}_n$ by the true value μ. The following theorem shows that the convergence rate of the estimate $\widehat{g}^{\circ}_{n,\gamma}(x)$ is determined by the properties of the Mellin transform of the function g.

For any $r > 0$ and any random variable η with $\mathsf{E}[|\eta|^r] < \infty$, denote $\|\eta\|_r := (\mathsf{E}[|\eta|^r])^{1/r}$.

Theorem 1 (Theorem 1.2 in the thesis)**.** *Let* $U_n = n^{1/4}$ *and* $V_n = \varkappa \ln(n)$ *for some* $\varkappa > 0$.

1. *If the true density* g *belongs to the class of functions*

$$\mathcal{E}(\alpha, \gamma_\circ, \gamma^\circ, L) \quad := \quad \left\{ p : \sup_{\gamma \in (\gamma_\circ, \gamma^\circ)} \int_{\mathbb{R}} e^{\alpha|v|} \left| \mathcal{M}[p](\gamma + \mathrm{i}v) \right| dv \leq L \right\},$$

   where $\alpha \in \mathbb{R}_+, L > 0,\ 0 < \gamma_\circ < \gamma^\circ < 1/2$, *then for* $\varkappa = \gamma^\circ/(\pi + 2\alpha)$, *any* $x \in \mathbb{R}_+$, *and* n *large enough, it holds*

$$\sqrt{|x|^{2\gamma_\circ} \wedge 1} \cdot \|\widehat{g}^\circ_{n,\gamma}(x) - g(x)\|_2 \quad \leq \quad C_1 n^{-\alpha\varkappa},$$

   where $\gamma \in (\gamma_\circ, \gamma^\circ)$, *and* C_1 *depends only on the parameters of the class* $\mathcal{E}$.

2. *If the true density* g *belongs to the class*

$$\mathcal{P}(\beta, \gamma_\circ, \gamma^\circ, L) \quad := \quad \left\{ p : \sup_{\gamma \in (\gamma_\circ, \gamma^\circ)} \int_{\mathbb{R}} |v|^\beta \left| \mathcal{M}[p](\gamma + \mathrm{i}v) \right| dv \leq L \right\}$$

   for some $\beta \in \mathbb{R}_+, L > 0,\ 0 < \gamma_\circ < \gamma^\circ < 1/2$, *then for some* $\varkappa > 0$, *any* $x \in \mathbb{R}_+$ *and* n *large enough, it holds*

$$\sqrt{|x|^{2\gamma_\circ} \wedge 1} \cdot \|\widehat{g}^\circ_{n,\gamma}(x) - g(x)\|_2 \quad \leq \quad C_2 \log^{-\beta}(n),$$

   where $\gamma \in (\gamma_\circ, \gamma^\circ)$, *and* C_2 *depends only on the parameters of* $\mathcal{P}$.

The convergence rates of the estimates $\hat{\mu}_n$ and $\widehat{g}_{n,\gamma}$ are presented in the following statement.

Theorem 2 (Theorem 1.3 in the thesis)**.** *Let* $r \geq 2$ *and* $M > 0$, *and*

$$\Lambda(M, r) := \left\| \left(1 + e^{-MX}\right) X w(X) \right\|_r < \infty.$$

Then

$$\mathsf{E} \left\| \widehat{\mu}_n - \mu \right\|_r \quad \leq \quad C_3 n^{-1/2},$$

and for n large enough

$$\sqrt{|x|^{2\gamma_\circ} \wedge 1} \cdot \|\widehat{g}_{n,\gamma}(x) - \widehat{g}^{\circ}_{n,\gamma}(x)\|_2 \quad \leq \quad C_4 n^{-1/2},$$

where the constants C_3, C_4 depend on μ, M, r.

2.1.2 Continuous-time moving average Lévy processes

The idea of superposition of several integral transforms can be applied not only in the case of Gaussian mixtures, but also in more complex, but related models. This section presents the results of Section 1.2 for moving average models based on the Lévy processes. These processes are defined as

$$(5) \qquad Z_t = \int_{-\infty}^{\infty} \mathcal{K}(t-s)\, dL_s, \quad t \in \mathbb{R},$$

where $\mathcal{K} : \mathbb{R} \mapsto \mathbb{R}_+$ is a measurable deterministic function and $(L_t)_{t\in\mathbb{R}}$ is a two-sided Lévy process given by

$$L_t = \begin{cases} L_t^{(1)}, & \text{if } t \geq 0 \\ L_{-t-}^{(2)}, & \text{if } t < 0, \end{cases}$$

with $L^{(1)}, L^{(2)}$ - 2 independent copies of a Lévy process $\tilde{L}$ with Lévy triplet (γ, σ^2, ν).[32]

[32] Lévy processes are defined as the processes, which are equal to 0 at the initial time point, have independent and stationary increments, and possess the property of stochastic continuity. The Lévy measure is used for the description of jumps of a Lévy process $\vec{X}$. More precisely, the Levy measure of a set $B \subset \mathbb{R}^d/\{\vec{0}\}$ is equal to

$$\nu(B) = \mathsf{E}\left[\sharp \left\{ t \in [0,1] : \Delta\vec{X}_t \in B \right\}\right],$$

where $\Delta\vec{X}_t = \vec{X}_t - \vec{X}_{t-}$ is the size of the jump at time t. Characteristic function of the process $\vec{X}_t, t \geq 0$, is equal to

$$\phi_t(\vec{u}) = \mathsf{E}[e^{\mathrm{i}\langle \vec{u}, \vec{X}_t\rangle}] = e^{t\psi(u)},$$

where $\psi(u)$ is the characteristic exponent. It can be represented by the Levy-Khinchine formula as follows

$$\psi(u) = \mathrm{i}\langle\vec{\gamma}, \vec{u}\rangle - \frac{1}{2}\langle \vec{u}, \Sigma\vec{u}\rangle + \int_{\mathbb{R}} \left(e^{\mathrm{i}\langle \vec{u}, \vec{x}\rangle} - 1 - \mathrm{i}\langle \vec{u}, \vec{x}\rangle \mathbb{I}\{|\vec{x}| < 1\}\right) \nu(d\vec{x}),$$

with $\vec{\gamma} \in \mathbb{R}^d$ and Σ is a non-negatively defined matrix. Thus, the distribution of the process $\vec{X}$ is uniquely determined by the triplet $(\vec{\gamma}, \Sigma, \nu)$, known as the Lévy triplet.

The conditions $\mathcal{K} \in \mathscr{L}^1(\mathbb{R}) \cap \mathscr{L}^2(\mathbb{R})$ and $\int x^2 \nu(dx) < \infty$ guarantee the correctness of the definition[33] and the stationarity of Z. The characteristic function of the process Z can be represented as

$$\Phi(u) := \mathsf{E}\left[e^{\mathrm{i}uZ_t}\right] = \exp\left(\Psi(u)\right) \qquad \text{with} \quad \Psi(u) := \int_{\mathbb{R}} \psi(u\,\mathcal{K}(s))\,ds,$$

where $\psi(\cdot)$ is the characteristic exponent of $\tilde{L}$.

The Section 1.2 focuses on the problem of estimation the density of $\nu(x)$ of the process $\tilde{L}$ from the observations of the process Z_t at equidistant time points $t = \Delta, 2\Delta, ..., n\Delta$ for some $\Delta > 0$ (here and in the sequel we use the same notation for the Lévy measure and its density).

To simplify the presentation of the results, we will describe the case of the known parameter σ. We denote

$$\Psi_\sigma(u) := \Psi(u) + \frac{\sigma^2 u^2}{2} \int_{\mathbb{R}} \mathcal{K}^2(x)\,dx.$$

The Mellin transform of the second derivative of this function is related to the superposition of the Mellin and Fourier transforms of the function $\tilde{\nu}(x) := x^2\nu(x)$ as follows:

$$\mathcal{M}\left[\Psi_\sigma''\right](z) = -\mathcal{M}[\mathcal{F}[\tilde{\nu}]](z) \cdot \int_{\mathbb{R}} \left(\mathcal{K}(x)\right)^{2-z}\,dx$$

for all z such that $\int_{\mathbb{R}} (\mathcal{K}(x))^{2-\mathrm{Re}(z)}\,dx < \infty$ and $\int_{\mathbb{R}_+} \left|\mathcal{F}[\tilde{\nu}]\,(v)\right| \cdot v^{\mathrm{Re}(z)-1} dv < \infty$. Note that the superposition of integral transforms is equal to

$$\begin{aligned}
\mathcal{M}[\mathcal{F}[\tilde{\nu}]](z) &= \mathcal{M}\big[e^{\mathrm{i}\cdot}\big](z) \cdot \mathcal{M}\big[\tilde{\nu}_+\big](1-z) + \mathcal{M}\big[e^{-\mathrm{i}\cdot}\big](z) \cdot \mathcal{M}\big[\tilde{\nu}_-\big](1-z), \\
\mathcal{M}\Big[\overline{\mathcal{F}[\tilde{\nu}]}\Big](z) &= \mathcal{M}\big[e^{-\mathrm{i}\cdot}\big](z) \cdot \mathcal{M}\big[\tilde{\nu}_+\big](1-z) + \mathcal{M}\big[e^{\mathrm{i}\cdot}\big](z) \cdot \mathcal{M}\big[\tilde{\nu}_-\big](1-z).
\end{aligned}$$

where

$$\tilde{\nu}_+(x) := \tilde{\nu}(x) \cdot \mathbb{I}\{x \geq 0\}, \qquad \tilde{\nu}_-(x) := \tilde{\nu}(-x) \cdot \mathbb{I}\{x \geq 0\}.$$

The above formulas imply a method for the estimation the density of the Lévy measure ν. The method consists of two steps.

[33] The correctness of the definition is shown in Lemma 1.4 in the thesis. This lemma is based on the general theory of the integrals in form (5), given in the paper Rajput, B. and Rosiński, J. Spectral representations of infinitely divisible processes. *Probability Theory and Related Fields*, 82(3): 451–487, 1989.

1. At the first step, the Mellin transforms $\mathcal{M}\left[\Psi''_\sigma\right]$ and $\mathcal{M}\left[\overline{\Psi''_\sigma}\right]$ are estimated by

$$\widehat{\mathcal{M}_n}[\Psi''_\sigma](1-z) := \int_0^{U_n}\left[\frac{\widehat{\Phi}''_n(u)}{\widehat{\Phi}_n(u)}-\left(\frac{\widehat{\Phi}'_n(u)}{\widehat{\Phi}_n(u)}\right)^2+\sigma^2\int_{\mathbb{R}}\mathcal{K}^2(x)\,dx\right]u^{-z}\,du,$$

$$\widehat{\mathcal{M}_n}[\overline{\Psi''_\sigma}](1-z) := \int_0^{U_n}\left[\overline{\frac{\widehat{\Phi}''_n(u)}{\widehat{\Phi}_n(u)}-\left(\frac{\widehat{\Phi}'_n(u)}{\widehat{\Phi}_n(u)}\right)^2}+\sigma^2\int_{\mathbb{R}}\mathcal{K}^2(x)\,dx\right]u^{-z}\,du,$$

resp., where U_n is an unbounded increasing sequence of positive numbers and $\widehat{\Phi}_n(u) := n^{-1}\sum_{j=1}^n e^{\mathrm{i}uZ_{j\Delta}}$.

2. At the second step, the functions $\tilde{\nu}_+$ and $\tilde{\nu}_-$ are estimated for $x\in\mathbb{R}_+$ by

$$\tilde{\nu}_{n+}(x) := \frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}V_n}^{c+\mathrm{i}V_n}\left(\frac{\widehat{\mathcal{M}_n}\left[\Psi''_\sigma\right](z)}{Q_1(z)}-\frac{\widehat{\mathcal{M}_n}\left[\overline{\Psi''_\sigma}\right](z)}{Q_2(z)}\right)x^{-z}dz,$$

$$\tilde{\nu}_{n-}(x) := \frac{1}{2\pi\mathrm{i}}\int_{c-\mathrm{i}V_n}^{c+\mathrm{i}V_n}\left(\frac{\widehat{\mathcal{M}_n}\left[\overline{\Psi''_\sigma}\right](z)}{Q_1(z)}-\frac{\widehat{\mathcal{M}_n}\left[\Psi''_\sigma\right](z)}{Q_2(z)}\right)x^{-z}dz,$$

where V_n is an unbounded increasing sequence of positive numbers, and

$$Q_1(z) := -\frac{e^{\mathrm{i}\pi z}-e^{-\mathrm{i}\pi z}}{e^{\mathrm{i}\pi z/2}}\Gamma(z)\int_{\mathbb{R}}\left(\mathcal{K}(x)\right)^{2-z}dx,$$

$$Q_2(z) := -\frac{e^{\mathrm{i}\pi z}-e^{-\mathrm{i}\pi z}}{e^{-\mathrm{i}\pi z/2}}\Gamma(z)\int_{\mathbb{R}}\left(\mathcal{K}(x)\right)^{2-z}dx.$$

For any $x<0$, we set $\tilde{\nu}_{n+}(x)=\tilde{\nu}_{n-}(x)=0$. The estimate of the function $\tilde{\nu}(x)=x^2\nu(x)$ is defined as

$$\tilde{\nu}_n(x) := \tilde{\nu}_{n+}(x)+\tilde{\nu}_{n-}(-x), \qquad x\in\mathbb{R}. \tag{6}$$

To study the theoretical properties of $\tilde{\nu}_n$, the following assumption is introduced.

(A1) The density of the Lévy measure ν satisfies the condition $\int_{-1}^1|x|\nu(x)dx<\infty$, and for some $A>0$, $\alpha\in(0,1), \beta_+>0, \beta_->0, c\in(0,1)$ it holds

$$\begin{cases}\int_{\mathbb{R}}(1+|y|)^\alpha\left|\mathscr{F}[\tilde{\nu}](y)\right|dy & \le A,\\ \int_{\mathbb{R}}e^{\beta_\pm|u|}\left|\mathcal{M}[\tilde{\nu}_\pm](c+\mathrm{i}u)\right|du & \le A.\end{cases}$$

In Section 1.2.4, it is shown that this condition holds for all Lévy measures $\nu(x) = \nu_+(x)+\nu_-(-x)$ from the class

$$\nu_\pm(x) \quad = \quad \sum_{j=1}^{J^{(\pm)}} a_j^{(\pm)} x^{-\eta_j^{(\pm)}-1} e^{-\lambda_j^{(\pm)} x} \cdot \mathbb{I}\{x \geq 0\},$$

with $J^{(+)}, J^{(-)} \in \mathbb{N} \cup 0$, $a_j^{(+)}, a_j^{(-)} > 0$, $\eta_j^{(+)}, \eta_j^{(-)} < 1$, $\lambda_j^{(+)}, \lambda_j^{(-)} > 0$ for all j. Note that this class includes the tempered stable distributions.

From a technical point of view, the difficulty is that the property of alpha-mixing[34] is not proven for processes (5). To get around this difficulty, the upper bound for the supremum of the difference between $\tilde{\nu}_n(x)$ and $\tilde{\nu}(x)$ are proved on the complement of some specially selected set $\mathscr{A}_K$, the probability of which tends to zero at polynomial rate.

Theorem 3 (Theorem 1.7 in the thesis)**.** *Let us fix $K > 0$ and denote the set*

$$\mathcal{A}_K := \left\{ \max_{j=0,1,2} \left\| \frac{\Phi_n^{(j)}(u) - \Phi^{(j)}(u)}{\Phi(u)} \right\|_{U_n} \geq K\varepsilon_n \right\},$$

where for any $f : \mathbb{R} \mapsto \mathbb{R}$, $\|f\|_{U_n} := \sup_{u \in [-U_n, U_n]} |f(u)|$, ε_n - is a sequence of positive numbers such that $\varepsilon_n \to 0$ as $n \to \infty$, and

$$K\varepsilon_n \left(1 + \left\|\Psi'_\sigma\right\|_{U_n}\right) \leq 1/2.$$

Then on the set $\mathcal{A}_K^{\mathcal{C}}$ (compliment to $\mathcal{A}_K$), the estimate $\tilde{\nu}_n(x)$ possesses the property

$$(7) \quad \sup_{x \in \mathbb{R}} \{|x|^c |\tilde{\nu}_n(x) - \tilde{\nu}(x)|\} \leq \frac{1}{2\pi} \int_{\{|v| \leq V_n\}} \frac{\Omega_n}{\min\left(|Q_1(1-c-\mathrm{i}v)|, |Q_2(1-c-\mathrm{i}v)|\right)} dv + \frac{A}{2\pi} e^{-\beta V_n},$$

[34] The α-mixing coefficient between $\sigma-$algebras $\mathcal{B}$ and $\mathcal{C}$ is defined as

$$\alpha(\mathcal{B}, \mathcal{C}) = \sup_{B \in \mathcal{B}, C \in \mathcal{C}} |\mathsf{P}\{B \cap C\} - \mathsf{P}\{B\}\mathsf{P}\{C\}|.$$

The sequence ξ_k is said to have the alpha-mixing property if

$$\alpha_\xi(k) = \sup_{t \in \mathbb{Z}} \alpha\big(\sigma(\xi_s, s \leq t), \sigma(\xi_s, s \geq t+k)\big) \to 0 \qquad \text{as} \quad k \to \infty.$$

where

$$\Omega_n \;:=\; \frac{2KC}{1-c}\varepsilon_n\big(1+\|\Psi'_\sigma\|_{U_n}\big)U_n^{1-c} + \left(A + \frac{2^\alpha A}{1-c}\right)\int_{\mathbb{R}} \big[\mathcal{K}(x)\big]^{c+1}\big[1+U_n\mathcal{K}(x)\big]^{-\alpha}\,dx,$$

and $C > 0$ is a constant for all processes with Lévy measure satisfying the condition (A1).

Section 1.2.5 considers specific cases for which the right part of (7) can be given in a more explicit form. In addition, the section describes sufficient conditions for the convergence of the probability of the set $\mathscr{A}_K$ to zero with a polynomial rate for a certain choice of the sequence ε_n and the number K.

2.2 Time-changed models

In the one-dimensional case ($d = 1$), the concept of stochastic time change is that for some random process L_t, $t \geq 0$, the deterministic time t is replaced by a non-decreasing non-negative random process $\mathcal{T}(s), s \geq 0$, which plays the role of random time. That is,

$$X(s) = L_{\mathcal{T}(s)}, \quad s \geq 0, \tag{8}$$

where it is often assumed that the processes L and $\mathcal{T}$ are independent. This model is motivated by the Monroe theorem, yielding that the class time-changed Brownian motions essentially coincides with the class of semimartingales (note that in this theorem the processes L and $\mathcal{T}$ can be dependent).[35]

The economic interpretation of the time change is that for a specific financial instrument, "business time" $\mathcal{T}$ may run faster than physical time in some time periods. For example, such periods of time may be associated with an increased activity on the market, expressed in the number of transactions.[36] If the time change process $\mathcal{T}$ is itself a Lévy process (subordinator), then the final process X is also a Lévy process, and for any $s \geq 0$, the distribution of the process

[35] Monroe, I. Processes that can be embedded in Brownian motion. *The Annals of Probability.* 6:42-56, 1978.

[36] Clark P. A subordinated stochastic process model with fixed variance for speculative prices. *Econometrica.* 41:135-156, 1973.
Ané, T. and Geman, H. Order flow, transaction clock, and normality of asset returns. *The Journal of Finance.* 55(5): 2259-2284, 2000

$X(s)$ is a mixture of probability distributions:

$$\mathsf{P}\{X(s) \in B\} = \int_{\mathbb{R}+} \mathsf{P}_{L_u}(B) d\mathsf{P}_{\mathcal{T}(s)}(u), \qquad B \in \mathcal{B}(\mathbb{R}^d).$$

2.2.1 Estimation of the Blumenthal-Getoor indices

The Blumenthal-Getoor index of a Lévy process with the Lévy measure ν is defined as follows:

$$\mathtt{BG}(Z) = \inf \left\{ r > 0 : \int_{|x| \leq 1} |x|^r \nu(dx) < \infty \right\}.$$

It is known that this index lies in the interval $[0, 2]$ and characterizes the activity of small jumps of the process. Equivalently, the Blumenthal-Getoor index can be defined as a number α such that $\nu(\{x : |x| > \varepsilon\}) \asymp c\varepsilon^{-\alpha}$, $\varepsilon \to 0+$ for some constant $c \in (0, \infty)$.[37]

Section 2.1 presents a method for the statistical estimation of the Blumenthal-Getoor index of the processes L and $\mathcal{T}$. It is assumed that the observations of the process $X(s) = L_{\mathcal{T}(s)}$ are available at the time points $s = \Delta, 2\Delta, ..., n\Delta$ with a fixed $\Delta > 0$ (low-frequency data), and the processes L and $\mathcal{T}$ are not observed separately. The following restrictions are imposed on the model.

(A1) The process $L_t, t \geq 0$, defined on the space with filtration $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathsf{P})$, is a Lévy process with a Lévy triplet (μ, σ^2, ν), where $\sigma \neq 0$ and

$$\nu(\{x : |x| > \varepsilon\}) = \varepsilon^{-\gamma}(\beta_0 + \beta_1 \varepsilon^{\chi_1}(1 + O(\varepsilon))), \quad \text{as} \quad \varepsilon \to +0,$$

$\beta_0 > 0, \beta_1 \in \mathbb{R}, \chi_1 \in (0, \gamma)$, and γ is the Blumenthal-Getoor index of the process L.

(A2) Process $\mathcal{T} = \mathcal{T}(s), s \geq 0$, is an increasing process starting from 0 with almost all càdlàg trajectories, such that for any fixed s, the random variable $\mathcal{T}(s)$ is a stopping time with respect to the filtration $\mathcal{F}$. The processes $\mathcal{T}$ and L are independent. The Laplace transform

[37] Panov V. *Abelian Theorem for Stochastic Volatility Models and Semiparametric Estimation of the Signal Space.* PhD dissertation. Humboldt University (Berlin), 2012.

of a random variable $\mathcal{T}(\Delta)$ has the property

$$\mathcal{L}_\Delta(u) := \mathsf{E}[e^{-u\mathcal{T}(\Delta)}] \asymp Ae^{-\lambda u^\alpha \Psi(u)}, \qquad u \to +\infty, \tag{9}$$

with $\lambda > 0$, $A > 0$, $\alpha \in (0, 1]$, and the function $\Psi(u)$ for u large enough satisfies

$$|1 - \Psi(u)| \leq \beta_2 u^{-\chi_2}$$

with some $\chi_2, \beta_2 \geq 0$ (note that if L is a subordinator, then the coefficient α coincides with its Blumenthal-Getoor index).

The method of estimation of the parameters α and γ proposed in Section 2.1.5 is motivated by the fact (which is proven in Section 2.1.4) that the absolute value of the characteristic function of the process X_Δ can be represented as

$$\left|\phi^\Delta(u)\right| = \left|\mathsf{E}[e^{\mathrm{i}uX_\Delta}]\right| = A \exp\left\{-\tau_1 |u|^{2\alpha} \left(1 + \tau_2 |u|^{\gamma-2} + r(u)\right)\right\},$$

where $\tau_1, \tau_2 > 0$ and $r(u) = o(|u|^{\gamma-2})$. Denote the estimate of this function by

$$\hat{\phi}_n(u) := \frac{1}{n}\sum_{k=1}^{n} e^{\mathrm{i}u(X_{\Delta k} - X_{\Delta(k-1)})},$$

and define the estimates of α and γ as follows:

$$\begin{aligned}
\hat{\alpha}_n &:= \frac{1}{2}\int_0^\infty w^{U_n}(u) \log\left(-\log|\hat{\phi}_n(u)|\right) du, \\
\hat{\gamma}_n(\hat{\alpha}_n) &:= 2(1 - \hat{\alpha}_n) + \int_0^\infty w^{V_n}(u) \log\left(-\log\frac{|\hat{\phi}_n(u)|^{\theta^{2\hat{\alpha}_n}}}{|\hat{\phi}_n(\theta u)|}\right) du,
\end{aligned}$$

where the weight functions are defined as $w^{V_n}(u) = V_n^{-1} w^1(u/V_n)$, $w^{U_n}(u) = U_n^{-1} w^1(u/U_n)$, and U_n, V_n are two infinitely increasing sequences of positive numbers, w^1 is almost everywhere a smooth function supported on the interval $[\varepsilon, 1]$ with $\varepsilon \in (0, 1)$, and satisfying the properties

$$\int_\varepsilon^1 w^1(u)\, du = 0, \qquad \int_\varepsilon^1 w^1(u) \log u\, du = 1. \tag{10}$$

For the study of the theoretical properties of estimates, another restriction is introduced on the time change process $\mathcal{T}(s)$.

(A3) The sequence $T_k = \mathcal{T}(\Delta k) - \mathcal{T}(\Delta(k-1))$, $k \in \mathbb{N}$, is stationary and α-mixing, and moreover, the mixing coefficients $(\alpha_T(k))_{k\in\mathbb{N}}$ satisfy the property

$$\alpha_T(k) \leq \bar{\alpha}_0 \exp(-\bar{\alpha}_1 k), \quad j \in \mathbb{N}$$

with $\bar{\alpha}_0, \bar{\alpha}_1 > 0$.

In Theorem 2.11, a class of models $\mathscr{G}$ is introduced, consisting of time-changed models satisfying the conditions (A1)-(A3) with some restrictions on the parameters. The following statement is shown for this class.

Theorem 4 (Combination of theorems 2.11, 2.12, 2.17 in the thesis)**.**

1. *The estimate $\hat{\gamma}_n(\alpha)$ (that is, the estimate of of the parameter γ, provided that the parameter α is known) has a logarithmic rate of convergence to the true value of γ, but this order is optimal in the minimax sense. More precisely, there are positive constants $\varkappa, \delta, c, \Xi$ such that for any $\Xi_1 > \Xi$ and $\Xi_2 < \Xi$, it holds*

$$\begin{aligned} \sup_{\mathscr{G}} \mathsf{P}\left\{|\hat{\gamma}_n(\alpha) - \gamma| \geq \Xi_1 (\log n)^{-c}\right\} &< \varkappa\, n^{-1-\delta}, \\ \varliminf_{n\to\infty} \inf_{\hat{\gamma}_n^*} \sup_{\mathscr{G}} \mathsf{P}\left\{|\hat{\gamma}_n^* - \gamma| \geq \Xi_2 (\log n)^{-c}\right\} &> 0, \end{aligned}$$

where $\hat{\gamma}_n^*$ is any estimate of γ.

2. *In general, when the parameter α is not known, it is possible to choose the sequences U_n and V_n such that for some constant $\Xi_3 > 0$,*

$$\sup_{\mathscr{G}} \mathsf{P}\left\{|\hat{\gamma}_n(\hat{\alpha}_n) - \gamma| \geq \Xi_3 (\log n)^{-c}\right\} < \varkappa\, n^{-1-\delta},$$

that is, the convergence rate coincides with the convergence rate in the case of the known parameter α.

2.2.2 Series representation of the multivariate time-changed processes

Section 2.2 deals with the generalisation of the model (8) to the multidimensional case.

Consider the d-dimensional Lévy process $\vec{L}_t = (L_1(t), ..., L_d(t)),\ t \in \mathbb{R}_+,$ *with independent components* and d-dimensional subordinator $\vec{\mathcal{T}}(s) = (\mathcal{T}_1(s), ..., \mathcal{T}_d(s)),\ s \in \mathbb{R}_+$ (i.e., a Lévy process such that each component is a non-negative Lévy process), *with dependent components*, where $\mathcal{T}_i$ and L_i are independent for any $i = 1, ..., d$. Define the multivariate time change as

$$\vec{X}(s) = \Big(X_1(s), ..., X_d(s)\Big) := \Big(L_1(\mathcal{T}_1(s)), ..., L_d(\mathcal{T}_d(s))\Big), \qquad s \in \mathbb{R}_+. \tag{11}$$

To describe the dependence between the components of the time-changed process $\vec{\mathcal{T}}$ we use the concept of a Lévy copula.[38]

Definition. A d-dimensional Lévy copula $F : \bar{\mathbb{R}}^d \mapsto \bar{\mathbb{R}}$ is a function that satisfies the following properties

1. $F(\vec{u}) = 0$ if $u_i = 0$ for at least one $i = 1, ..., d$ (grounded function);
2. F is a d-increasing function;
3. $F^{(1)}(v) = ... = F^{(d)}(v) = v$, where

   $$F^{(j)}(v) = \lim_{u_1,...,u_{j-1},u_{j+1},u_d \to \infty} F(u_1, ..., u_{j-1}, v, u_{j+1}, ..., u_d),\ j = 1, ..., d,$$

   (uniform margins);
4. $F(u_1, ..., u_d) \neq \infty$ for any $(u_1, ..., u_d) \neq (\infty, ..., \infty)$.

Lévy copulas are closely related to the notion of a tail integral, which is defined for a Lévy measure ν as

$$U(x_1, ..., x_d) := (-1)^{h(x_1)+...+h(x_d)} \cdot \nu\Big(I(x_1) \times ... \times I(x_d)\Big),$$

[38] Cont, R. and Tankov, P. *Financial Modelling with Jump Processes.* Chapman & Hall. CRC Press, UK, 2004.

where

$$I(x) := \begin{cases} (x, +\infty), & \text{if } x > 0, \\ (-\infty, x), & \text{if } x < 0, \end{cases} \qquad \text{and} \qquad h(x) := \begin{cases} 2, & \text{if } x > 0, \\ 1, & \text{if } x < 0. \end{cases}$$

According to the analogue of Sklar's theorem, for a Lévy process with a tail integral U and marginal tail integrals $U_1, ..., U_d$, there is a Lévy copula F such that

$$U(x_1, ..., x_d) = F\left(U_1(x_1), ..., U_d(x_d)\right). \tag{12}$$

Conversely, for any Lévy copula F and any one-dimensional Lévy processes with residual integrals $U_1, ..., U_d$, there is a d-dimensional Lévy process with a residual integral U given by the formula (12), and marginal residual integrals $U_1, ..., U_d$. Note that Lévy copulas differ from the "ordinary" copulas only in the domain and the image of the function: "ordinary" copulas are defined on $[0, 1]^d$ and take values on $[0, 1]$.

In Section 2.2.6, the following theoretical result is given for the case when $\vec{L}$ is a multidimensional stable process.

Theorem 5 (Theorem 2.20 in the thesis)**.** *Consider the model (11), where $L_1, ...L_d$ are independent stable processes, and $\vec{\mathcal{T}} = (\mathcal{T}_1, ..., \mathcal{T}_d)$ is a d-dimensional subordinator (in general case, with dependent components) with the Laplace transform*

$$\mathscr{L}_{\mathcal{T}(s)}(\vec{u}) := \mathsf{E}\left[e^{\langle \vec{\mathcal{T}}(s), \vec{u}\rangle}\right] = \exp\Big\{ s \int_{\mathbb{R}} \left(e^{\langle \vec{u}, \vec{x}\rangle} - 1\right) \eta(d\vec{x}) \Big\}, \qquad \vec{u} \in \mathbb{R}^d, \quad s \geq 0,$$

where the Lévy measure η satisfies the condition

$$\int_{|\vec{x}| \leq 1} |\vec{x}|^{1/2} \eta(d\vec{x}) < \infty.$$

Denote by $F(u_1, ..., u_d)$ the Lévy copula between $\mathcal{T}_1, .., \mathcal{T}_d$. Assume that:

(1) the distribution function $\widetilde{F}(u_1, ..., u_{d-1}|v) = \partial F(u_1, .., u_{d-1}, v)/\partial v$ is a c.d.f. of an absolutely continuous distribution for any $v \geq 0$;

(2) there exist the functions $h_1, ..., h_{d-1} : \mathbb{R} \times \mathbb{R}_+ \mapsto \mathbb{R}$ and random variables $\xi_1, ..., \xi_{d-1}$ such

that

$$\mathsf{P}\left\{h_1(\xi_1, v) \le u_1, ..., h_{d-1}(\xi_{d-1}, v) \le u_{d-1}\right\} = \widetilde{F}(u_1, ..., u_{d-1}|v).$$

Then

$$\vec{X}(s) \stackrel{d}{=} \vec{Z}(s), \qquad \forall s \in [0,1],$$

with the d-dimensional stochastic process $\vec{Z}(s) = (Z_1(s), ..., Z_d(s))$ *defined as follows:*

$$\begin{aligned} Z_k(s) \quad &:= \quad \sum_{i=1}^{\infty} \left[\left(G_i^{(k)} - \mu_i \right) \left(U_k^{(-1)} \left(h_k(Q_i^{(k)}, \Gamma_i) \right) \right)^{1/\alpha_k} \right. \\ &\qquad \left. + \mu_i U_k^{(-1)} \left(h_k(Q_i^{(k)}, \Gamma_i) \right) \right] \mathbb{I}\left\{R_i \le s\right\} \end{aligned}$$

for $k = 1..(d-1)$ *and*

$$Z_d(s) \quad := \quad \sum_{i=1}^{\infty} \left[\left(G_i^{(d)} - \mu_i \right) \left(U_d^{(-1)}\left(\Gamma_i\right) \right)^{1/\alpha_k} + \mu_i U_d^{(-1)}\left(\Gamma_i\right) \right] \mathbb{I}\left\{R_i \le s\right\},$$

where

- $U_1, ..., U_d$ *are the tail integrals of the Lévy measures of subordinators* $T_1, ..., T_d$*, respectively, and* $U_1^{(-1)}, ..., U_d^{(-1)}$ *are their generalised inverse functions, that is*

$$U_i^{(-1)}(y) = \inf\left\{x > 0 : \quad U_i(x) < y\right\}, \quad i = 1...d, \quad y \in \mathbb{R}_+;$$

- $\Gamma_i, i = 1, 2, ...,$ *is the sequence of jumps of the standard Poisson process;*
- $R_i, i = 1, 2, ...,$ *is a sequence of independent identically distributed quantities with uniform distribution on* $[0,1]$*;*
- *for* $i = 1, 2, ...,$ $G_i^{(1)}, ...,$ $G_i^{(d)}$ *is a sequence of i.i.d. stable variables,* $G_i^{(j)} \sim S_{\alpha_j}(\sigma_j, \beta_j, 0)$*;*
- *for* $i = 1, 2, ...,$ $\vec{Q}_i := \left(Q_i^{(1)}, ..., Q_i^{(d-1)}\right)$ *is an i.i.d. sequence having the same distribution as* $(\xi_1, ..., \xi_{d-1})$*,*

and all sequences $\Gamma_i, R_i, G_i^{(1)}, ...,$ $G_i^{(d)}, \vec{Q}_i,$ $i = 1, 2, ...$ *are jointly independent.*

Note that in the time-changed models, the use of stable processes has a number of advantages in comparison with Brownian motion. Section 2.2.7 presents the results of the empirical analysis showing that the model of a two-dimensional subordinated stable process reproduces the dependence between the components well both in the case of highly correlated components (Section 2.2.7 considers the returns of Apple and Microsoft shares) and in the case of weak correlation (the returns of Apple and General Electric shares).

2.3 Construction of the honest confidence sets

2.3.1 Honest confidence sets for the density functions

In Section 3.1, we present a new approach for the construction of the confidence sets for the density and show the application of this method to some mixture models.

We say that $\mathcal{C}_n(x)$ is a $(1-\alpha)$-confidence set for p, honest with respect to a given class $\mathscr{F}$, if

$$\inf_{p\in\mathscr{F}} \mathsf{P}\Big\{p(x) \in \mathcal{C}_n(x), \quad \text{for all } x \in \mathbb{R}\Big\} \geq 1-\alpha+e_n, \tag{13}$$

where $e_n \to 0$ as $n \to \infty$. The standard approach to the construction of such confidence intervals consists in the application of theoretical facts known in the literature as the SBR-type theorems (Smirnov - Bickel - Rosenblatt), which yield that the distribution of the maximal deviation of the estimate $\hat{p}_n$, i.e.

$$\mathcal{D}[\hat{p}_n] = \sup_{u\in\mathbb{R}} \frac{|\hat{p}_n(u)-p(u)|}{\sqrt{p(u)}} \tag{14}$$

is asympotically close to the Gumbel distribution. Namely, as $n \to \infty$,

$$\sup_{p\in\mathscr{F}} \left| \mathsf{P}\left\{\mathcal{D}[\hat{p}_n] \leq \frac{x}{a_n} + b_n\right\} - e^{-e^{-x}} \right| \to 0, \qquad \forall\, x \in \mathbb{R} \tag{15}$$

for some deterministic sequences a_n and b_n. Note that the statements of the form (15) have been proved only for the kernel estimates and for projection estimates constructed using some wavelet bases (Haar basis and Battle-Lemarie basis).[39]

[39] Giné, E. and Nickl, R. *Mathematical Foundations of Infinite-dimensional Statistical Models*. Cambridge University Press, 2016.

Section 3.1 presents a method for the construction of honest confidence intervals with sequences e_n of polynomial order. The method is based on the projection estimates, which are defined for the density functions from the space $\mathscr{L}^2([A,B])$ ($[A,B]$ is a fixed interval) as follows. In the space $\mathscr{L}^2([A,B])$, we choose a basis $\Psi := \{\psi_0, \psi_1, \psi_2, ...\}$, divide $[A,B]$ on M intervals of the length $\delta = (B-A)/M$, and on each interval $I_m = [a_m, b_m] := [A+\delta(m-1), A+\delta m], m = 1..M$, we reproduce the basis Ψ

$$\psi_j^{(m)}(x) = \sqrt{M} \cdot \psi_j\Big(M(x-a_m)+A\Big), \quad m = 1..M, \;\; j = 0, 1, ...$$

Since $p \in \mathscr{L}^2([A,B])$, for any $M \in \mathbb{N}$,

$$p(x) = \sum_{m=1}^{M} \sum_{j=0}^{\infty} \left[\int \psi_j^{(m)}(u) p(u) du \right] \psi_j^{(m)}(x).$$

The projection estimate p is defined as

$$\hat{p}_n(x) \quad = \quad \sum_{m=1}^{M} \sum_{j=0}^{J} \left[\int \psi_j^{(m)}(u) d\mathsf{P}_n(u) \right] \psi_j^{(m)}(x),$$

where $J \in \mathbb{N}$ and $\mathsf{P}_n = n^{-1} \sum_{i=1}^{n} \delta_{X_i}$ is the empirical measure.

As shown in Section 3.1.4, the distribution of the maximum deviation of this estimate (in terms of (14)) is close to the distribution of the supremum of the absolute value of the Gaussian process

$$\Upsilon(x) = \int_I \Big(\sum_{j=0}^{J} \psi_j(x) \psi_j(u) \Big) dW(u), \qquad x \in [A,B],$$

where W is the Brownian motion. Due to Proposition 3.1,

$$\mathsf{P}\Big\{ \max_{t \in [A,B]} \big| \Upsilon(t) \big| \geq u \Big\} = G(u) + O\Big(e^{-u^2(1+\chi)/(2S)} \Big), \qquad u \to \infty,$$

for $S = \max_{t \in [A,B]} \{\mathsf{Var} \Upsilon_t\}$, some $\chi > 0$ and some decreasing function $G : \mathbb{R}_+ \mapsto \mathbb{R}$ such that $\lim_{x \to \infty} G(x) = 0$. Note that Proposition 3.1 is a new statement of the theory of extreme values for non-stationary Gaussian processes.

Below is the formulation of the statement about the sequence of accompanying laws.

Theorem 6 (Theorem 3.11 in the thesis)**.** *Assume that for some $J \in \mathbb{N}$, the functions $\psi_0, ..., \psi_J$ have the following properties.*

(A1) For any $j = 0..J,$, the function ψ_j is uniformly Hölder-continuous with some exponent $\alpha \in (0, 1]$, i.e., the Hölder coefficient

$$\left|\psi_j\right|_\alpha := \sup_{x \neq y,\, x,y \in [A,B]} \frac{|\psi_j(x) - \psi_j(y)|}{|x - y|^\alpha}$$

is finite.

(A2) The maximum of the sum $\sum_{j=0}^{J} \psi_j^2(x)$ is attained at a finite number of points.

Assume that the true density p belongs to the class

$$\mathcal{P}_{q,H,\beta} := \Big\{ p - \ p.d.f.\ , \quad p \in \mathscr{L}^2([A, B]), \quad \inf_{x \in [A,B]} p(x) \geq q, \quad \left|p\right|_\beta \leq H \Big\},$$

for some $q > 0, H > 0, \beta \in (0, 1]$. Denote the sequence of distribution functions

$$A_M(x) := \begin{cases} \exp\Big\{ -MG(x) \Big\}, & \text{if } x \geq c_M, \\ 0, & \text{if } x < c_M \end{cases}$$

with $c_M = (2S \log M)^{1/2} - S$. Then for $M = \lfloor n^\lambda \rfloor$ with $\lambda \in ((2\beta + 1)^{-1}, 1)$, and n large enough, it holds

$$\sup_{x \in \mathbb{R}} \left| \mathsf{P} \left\{ \sqrt{\frac{n}{M}} \mathcal{D}[\hat{p}_n] \leq x \right\} - A_M(x) \right| \ \leq \ \bar{c} n^{-\gamma}$$

with some $\bar{c}, \gamma > 0$.

The method of the construction of confidence intervals that are honest with respect to the class $\mathcal{P}_{q,H,\beta}$ is based on Theorem 6. As shown in Section 3.1.6, such intervals can be written as

$$\mathcal{C}_n(x) := \Big(\hat{p}_n(x) + (k^2_{\alpha,M}/2) - \big[\hat{p}_n(x)k^2_{\alpha,M} + (k^4_{\alpha,M}/4)\big]^{1/2},$$
$$\hat{p}_n(x) + (k^2_{\alpha,M}/2) + \big[\hat{p}_n(x)k^2_{\alpha,M} + (k^4_{\alpha,M}/4)\big]^{1/2}\Big),$$

where $k_{\alpha,M} := \sqrt{M/n} \cdot q_{\alpha,M}$ and $q_{\alpha,M}$ - $(1-\alpha)$ - quantile of the distribution A_M. Numerical experiments were provided for the density of a mixture of normal distributions

$$p(x) = \frac{1}{2}\phi_{(0,1)}(x) + \frac{1}{10}\sum_{j=0}^{4}\phi_{((j/2)-1,1/100)}(x),$$

known in the literature as the Bart Simpson density.[40]

2.3.2 Estimation of the Lévy measure

The techniques described in the previous section can also be applied in more complex models. Section 3.2 presents a method for the construction of confidence intervals for the density of the Lévy measure. Note that in the stochastic time-changed model (see Section 2.2 above), the Lévy measure of the process $X(s) = L_{\mathcal{T}(s)},\ s \geq 0$, is a mixture of probability distributions if the process $\mathcal{T}$ is a subordinator. Indeed, in this case, the Lévy measure of the process $X(s)$ is represented as

$$\nu(dx) = \int_{\mathbb{R}_+} \mu_t(dx)\nu_{\mathcal{T}}(dt),$$

where μ_t is the probability measure of the process L_t at time $t \geq 0$, and $\nu_{\mathcal{T}}$ is the Lévy measure of $\mathcal{T}$.

Assuming that observations of the process X_t are available at $t = \Delta, 2\Delta, ..., n\Delta$ with $\Delta \to 0$ as $n \to \infty$, we determine the projection estimate of the Lévy density

$$\hat{s}_n(x) := \frac{1}{n\Delta}\sum_{m=1}^{M}\sum_{j=0}^{J}\left[\sum_{k=1}^{n}\psi_j^{(m)}\left(X_\Delta^{(k)}\right)\right]\psi_j^{(m)}(x), \tag{16}$$

[40] Wasserman, L. *All of Nonparametric Statistics.* Springer Science and Business Media. 2006.

where $X_\Delta^{(k)} = X_{k\Delta} - X_{(k-1)\Delta},\ k = 1..n$. The analysis of the maximum deviation of this estimate is significantly more complex than in the case of projection estimates for the probability density function. For simplicity, we formulate the result for a special case of the trigonometric basis.

Theorem 7 (Theorem 3.22 in the thesis)**.** *Let $\hat{s}_n(x)$ be the estimate (16), constructed from a trigonometric basis, which is defined on the interval $[A, A+\delta)$ as*

$$\left\{\frac{1}{\sqrt{\delta}},\quad \sqrt{\frac{2}{\delta}}\cos\left(2j\pi(x-a)/\delta\right),\quad \sqrt{\frac{2}{\delta}}\sin\left(2j\pi(x-a)/\delta\right),\ \ j = 1, 2, ..\right\}.$$

Define the sequence of the distribution functions

$$A_M(x) := \begin{cases} \exp\left\{-2\exp\left\{-x - \frac{x^2}{4\log(h_1 M)}\right\} - 2M\left(1 - \Phi\left(u_M(x)\sqrt{2h_2}\right)\right)\right\}, \\ \hfill \text{if } x \geq -b_M^{3/2}, \\ 0, \hfill \text{if } x < -b_M^{3/2}, \end{cases}$$

where

$$u_M(x) := \frac{x}{a_M} + b_M$$

with

$$a_M = 2h_2 b_M, \quad b_M = \sqrt{\frac{\log(h_1 M)}{h_2}}, \quad h_1 = \sqrt{\frac{2\sum_{j=1}^{J/2} j^2}{J+1}}, \quad h_2 = \frac{B-A}{2(J+1)}.$$

Then for $T = n^{\varkappa}, M = o(n^{\varkappa/2}/\log n)$ with $\varkappa \in (0,1)$, it holds for n large enough

$$\sup_{x\in\mathbb{R}}\left|\mathsf{P}\left\{\sqrt{\frac{T}{M}}\mathcal{D}[\hat{s}_n] \leq u_M(x)\right\} - A_M(x)\right| \leq \bar{c}\, n^{-\gamma}.$$

with some $\bar{c}, \gamma > 0$.

2.4 Limit laws and phase transitions in mixture models

2.4.1 Parabolic Anderson problem with a potential having a mixture distribution

On the lattice $\mathbb{Z}^d$, consider the cube $Q_n = [-n, n]^d$ and random Anderson Hamiltonian

$$H_n = \Delta + \beta V_n(x, \omega),$$

where β is the reciprocal temperature, $V_n(x, \omega), x \in Q_n$, is the random i.i.d. potential, and

$$\Delta\psi(x) = \sum_{x' : |x' - x| = 1} \psi(x')$$

is the lattice Laplacian on Q_n with Dirichlet boundary condition $\psi(x) = 0,\ x \in \partial Q_n$. We assume a strong potential and choose $V_n(x, \omega) = \sqrt{n}\xi(x, \omega)$, where ξ is a Gaussian random variable.

Now consider the parabolic problem

$$\begin{aligned} \frac{\partial u}{\partial t} &= H_n u, \qquad t \geq 0,\ x \in Q_n, \\ u(t, x) &= 0, \qquad\qquad\quad x \in \partial Q_n, \\ u(0, x) &= \delta_y(x), \end{aligned}$$

where $y \in Q_n$ is considered as a parameter.

The fundamental solution of this problem is given by

$$u_n(t, x) = u_n(t, x, y) = \sum_{i=1}^{|Q_n|} e^{\lambda_{n,i} t} \psi_{n,i}(x) \psi_{n,i}(y), \tag{17}$$

where $\lambda_{n,i}, \psi_{n,i}$ are the eigenvalues and the normalised eigenfunctions of the operator H_n, that is, $H_n \psi_{n,i} = \lambda_{n,i} \psi_{n,i}$. Denote the random exponential sum

$$\operatorname{Tr} e^{tH_n} = \sum_{x \in Q_n} u_n(t, x, x) = \sum_{i=1}^{|Q_n|} e^{\lambda_{n,i} t}, \tag{18}$$

which for $t = \beta$ is close to $S_n(\beta) = \sum_{y \in Q_n} e^{\beta\sqrt{n}\xi(y, \omega)}$.

By analogy, the trace of the Hamiltonians with potentials

$$V_n(x,\omega) = \sqrt{n}\xi(x,\omega), \qquad \text{where} \quad \xi(x,\omega) = \begin{cases} \eta(x,\omega), & \text{with probability } 1/2, \\ \zeta(x,\omega), & \text{with probability } 1/2, \end{cases}$$

where η, ζ are independent Gaussian random variables, is close to the partition function of the REM model with a mixture distribution. The limit distributions and phase transitions of this model are considered in Section 4.1.

The mathematical object, which is studied in Section 4.1, can be presented as

$$\mathcal{S}_n(\beta) = \sum_{j=1}^{\lfloor e^n \rfloor} e^{\beta\sqrt{n}Z_j},$$

where $Z_1, ..., Z_{\lfloor e^n \rfloor}$ is a sequence of i.i.d. random variables having the distribution function

$$F_{a,\sigma}(x) = \frac{1}{2}\Phi\left(x\right) + \frac{1}{2}\Phi\left(\frac{x-\sqrt{n}a}{\sigma}\right)$$

with $a \in \mathbb{R}, \sigma > 0$.

Theorem 8 (Theorems 4.3, 4.4, 4.5 in the thesis)**.**

1. *The law of large numbers*

$$\frac{\mathcal{S}_n(\beta)}{\mathsf{E}[\mathcal{S}_n(\beta)]} \xrightarrow{p} 1, \qquad n \to \infty,$$

holds for $\beta < \beta^+$, where

$$\beta^+ = \begin{cases} \sqrt{2}/\sigma, & \text{if } a > \left(1-\sigma^2\right)/(\sqrt{2}\sigma), \\ \beta_\circ := \frac{2a}{1-\sigma^2}, & \text{if } \left(1-\sigma^2\right)/\sqrt{2} < a < \left(1-\sigma^2\right)/(\sqrt{2}\sigma), \\ \sqrt{2}, & \text{if } a < \left(1-\sigma^2\right)/\sqrt{2}. \end{cases}$$

2. *The central limit theorem*

$$\frac{\mathcal{S}_n(\beta) - \mathsf{E}[\mathcal{S}_n(\beta)]}{\sqrt{\mathsf{Var}(\mathcal{S}_n(\beta))}} \xrightarrow{d} \zeta_0, \qquad n \to \infty,$$

with $\zeta_0 \sim \mathcal{N}(0,1)$, *holds for* $\beta < \beta^+/2$.

3. *The convergence in distribution to stable laws holds in the following cases.*

   (i) If $a < \sqrt{2}(1-\sigma)$, *then there exists a deterministic sequence* $a_n^\sharp(\beta)$ *such that*

$$\frac{\mathcal{S}_n(\beta) - a_n^\sharp(\beta)}{\gamma_n(\beta)} \xrightarrow{d} \zeta_\beta, \qquad n \to \infty,$$

   for any $\beta > \beta^\sharp$, *where*

$$\beta^\sharp \;=\; \begin{cases} \beta_\diamond/2, & \text{if } \sigma < 1 \text{ and } a > (1-\sigma^2)/\sqrt{2}, \\ \sqrt{2}/2, & \text{elsewhere,} \end{cases}$$

   with $\beta_\diamond = \left((\sqrt{2}-a) - \sqrt{\left(\sqrt{2}-a\right)^2 - 2\sigma^2}\right)/\sigma^2$, *and* ζ_β *is a random variable having* $(\sqrt{2}/\beta)$*-stable distribution with Lévy triplet* $\left(0, 0, (2\pi)^{-1/2} x^{-\sqrt{2}/\beta} \mathbb{I}_{x>0} dx\right)$.

   (ii) If $a > \sqrt{2}(1-\sigma)$, *there exists a deterministic sequence* $\breve{a}_n(\beta)$ *such that*

$$\frac{\mathcal{S}_n(\beta) - \breve{a}_n(\beta)}{e^{\beta a n}\gamma_n(\beta\sigma)} \xrightarrow{d} \zeta_{\beta\sigma}, \qquad n \to \infty,$$

   for any $\beta > \breve{\beta}$, *where*

$$\beta^\sharp \;=\; \begin{cases} \beta_*/2, & \text{if } \sigma > 1 \text{ and } a < (1-\sigma^2)/(\sqrt{2}\sigma), \\ \sqrt{2}/(2\sigma), & \text{elsewhere} \end{cases}$$

   with $\beta_* = \left(\sigma\sqrt{2}+a\right) - \sqrt{\left(\sigma\sqrt{2}+a\right)^2 - 2}$.

2.5 Determining the type of the distribution

Section 5.2 presents new results on the type of the generalised Dickman-Goncharov distribution, defined as the distribution of a random variable $\mathcal{B}$ satisfying the distributional equation

$$\mathcal{B} \stackrel{d}{=} \mathcal{T} + \mathcal{B}\mathcal{X},$$

where $(\mathcal{T}, \mathcal{X})$ and $\mathcal{B}$ in the right-hand side are independent. In Section 5.2, we consider the case when $\mathcal{T} = \mathcal{X}$, and these variables take the values ρ^m with probabilities $qp^m,\ m = 0, 1, 2, ..,$ where $\rho \in (0, 1)$, $p \in (0, 1)$ and $q = 1 - p$ (discrete Dickman-Goncharov distribution). It is shown that in this case

$$\mathcal{B} \stackrel{d}{=} \sum_{n=0}^{\infty} Z_p^{(n)} \rho^n,$$

where $Z_p^{(1)}, Z_p^{(2)}, ...$ - i.i.d. r.v.'s taking the values $0, 1, 2, 3, ..$ with probabilities $p, pq, pq^2, pq^3, ...$ The following theorem holds.

Theorem 9 (Theorem 5.5 in the thesis)**.** *The distribution of the random variable $\mathcal{B}$ is absolutely continuous for almost all*

$$\rho \in \Big(\frac{q^2 + 1}{(q + 1)^2}, 1\Big).$$

3 Conclusion

The thesis presents new approaches to solving various probabilistic and statistical problems for mixture models. The obtained results are related to five different directions of study in this field (see p. 8). The results are published in twelve papers, among them eleven papers are published in the journals with quartiles Q1/Q2 (in Scopus). The main findings are presented in this summary in the form of nine theorems.