

Article

# Sharp Estimates for Proximity of Geometric and Related Sums Distributions to Limit Laws

Alexander Bulinski <sup>1,\*</sup>  and Nikolay Slepov <sup>2</sup> 

<sup>1</sup> Faculty of Mathematics and Mechanics, Lomonosov Moscow State University, Leninskie Gory 1, 119991 Moscow, Russia

<sup>2</sup> Department of Higher Mathematics, Moscow Institute of Physics and Technology, National Research University, 9 Institutskiy per., Dolgoprudny, 141701 Moscow, Russia

\* Correspondence: alexander.bulinski@math.msu.ru

**Abstract:** The convergence rate in the famous Rényi theorem is studied by means of the Stein method refinement. Namely, it is demonstrated that the new estimate of the convergence rate of the normalized geometric sums to exponential law involving the ideal probability metric of the second order is sharp. Some recent results concerning the convergence rates in Kolmogorov and Kantorovich metrics are extended as well. In contrast to many previous works, there are no assumptions that the summands of geometric sums are positive and have the same distribution. For the first time, an analogue of the Rényi theorem is established for the model of exchangeable random variables. Also within this model, a sharp estimate of convergence rate to a specified mixture of distributions is provided. The convergence rate of the appropriately normalized random sums of random summands to the generalized gamma distribution is estimated. Here, the number of summands follows the generalized negative binomial law. The sharp estimates of the proximity of random sums of random summands distributions to the limit law are established for independent summands and for the model of exchangeable ones. The inverse to the equilibrium transformation of the probability measures is introduced, and in this way a new approximation of the Pareto distributions by exponential laws is proposed. The integral probability metrics and the techniques of integration with respect to sign measures are essentially employed.

**Keywords:** probability metrics; Stein method; geometric sums; generalization of the Rényi theorem; generalized transformation of equilibrium for probability measures and its inverse; generalized gamma distribution

**MSC:** 60F99; 60E10; 60G50; 60G09



**Citation:** Bulinski, A.; Slepov, N. Sharp Estimates for Proximity of Geometric and Related Sums Distributions to Limit Laws. *Mathematics* **2022**, *10*, 4747. <https://doi.org/10.3390/math10244747>

Academic Editors: Alexander Tikhomirov and Vladimir Ulyanov

Received: 29 October 2022

Accepted: 7 December 2022

Published: 14 December 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The theory of sums of random variables belongs to the core of modern probability theory. The fundamental contribution to the formation of the classical core was made by A. de Moivre, J. Bernoulli, P.-S. Laplace, D. Poisson, P.L. Chebyshev, A.A. Markov, A.M. Lyapunov, E. Borel, S.N. Bernstein, P. Lévy, J. Lindeberg, H. Cramér, A.N. Kolmogorov, A.Ya. Khinchin, B.V. Gnedenko, J.L. Doob, W. Feller, Yu.V. Prokhorov, A.A. Borovkov, Yu.V. Linnik, I.A. Ibragimov, A. Rényi, P. Erdős, M. Csörgö, P. Révész, C. Stein, P. Hall, V.V. Petrov, V.M. Zolotarev, J. Jacod and A.N. Shiryaev among others. The first steps led to limit theorems for appropriately normalized partial sums of sequences of independent random variables. Besides the laws of large numbers, special attention was paid to emergence of Gaussian and Poisson limit laws. Note that despite many efforts to find necessary and sufficient conditions for the validity of the central limit theorem (the term was proposed by G. Pólya for a class of limit theorems describing weak convergence of distributions of normalized sums of random variables to the Gaussian law), this problem was completely resolved for independent summands only in the second part of the 20th century in the

works by V.M. Zolotarev and V.I. Rotar. Also in the last century, the beautiful theory of infinitely divisible and stable laws was constructed. New developments of infinite divisibility along with classical theory can be found in [1]. For exposition of the theory of stable distributions and their applications, we refer to [2], see also references therein.

Parallel to partial sums of a sequence of random variables (and vectors), other significant schemes have appeared, for instance, the arrays of random variables. Moreover, in physics, biology and other domains, researchers found that it was essential to study the sums of random variables when the number of summands was random. Thus, the random sums with random summands became an important object of investigation. One can mention the branching processes which stem from the 19th century population models by I.J. Bienaymé, F. Galton and H.W. Watson that are still intensively being developed, see, e.g., [3]. In the theory of risk, it is worth recalling the celebrated Cramér–Lundberg model for dynamics of the capital of an insurance company, see, e.g., Ch. 6 in [4]. Various examples of models described by random sums are considered in Ch. 1 of [5], including (see Example 1.2.1) the relationship between certain random sums analysis and the famous Pollaczek–Khinchin formula in queuing theory. A vast literature deals with the so-called geometric sums. There, one studies the sum of independent identically distributed random variables, and the summation index follows the geometric distribution, being independent with summands. Such random sums can model many real world phenomena, e.g., in queuing, insurance and reliability, see the Section “Origin of Geometric Sums” in the Introduction of [6]. Furthermore, a multitude of important stochastic models described by systems of dependent random variables occurred to meet diverse applications, see, e.g., [7]. In particular, the general theory of stochastic processes and random fields arose in the last century (for introduction to random fields, see, e.g., [8]).

An intriguing problem of estimating the convergence rate to a limit law was addressed by A.C. Berry and C.-G. Esseen. Their papers initiated the study of proximity for distribution functions of the normalized partial sums of independent random variables to the distribution function of a standard Gaussian law in the framework of the classical theory of random sums.

To assess the proximity of distributions, we will employ various integral probability metrics. Usually, for random variables  $Y, Z$  and a specified class  $\mathcal{H}$  of functions  $h : \mathbb{R} \rightarrow \mathbb{R}$ , one sets

$$d_{\mathcal{H}}(Y, Z) := \sup_{h \in \mathcal{H}} |\mathbb{E}[h(Y)] - \mathbb{E}[h(Z)]| \in [0, \infty]. \quad (1)$$

Clearly,  $d_{\mathcal{H}}(Y, Z)$  is a functional depending on  $law(Y)$  and  $law(Z)$ , i.e., distributions of  $Y$  and  $Z$ . A class  $\mathcal{H}$  should be rich enough to guarantee that  $d_{\mathcal{H}}$  possesses the properties of a metric (or semi-metric). The general theory of probability metrics is presented, e.g., in [9,10]. In terms of such metrics, one often compares the distribution of a random variable  $Y$  under consideration with that of a target random variable  $Z$ . In Section 2, we recall the definitions of the Kolmogorov and Kantorovich (alternatively called Wasserstein) distances and Zolotarev ideal metrics corresponding to the adequate choice of  $\mathcal{H}$ , denoted below as  $\mathcal{K}$ ,  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively.

It should be emphasized that for sums of random variables, deep results were established along with creation and development of different methods of analysis. One can mention the method of characteristic functions due to the works of J.Fourier, P.-S.Laplace and A.M.Lyapunov, the method of moments proposed by P.L.Chebyshev and developed by A.A.Markov, the Lindeberg method of employing auxiliary Gaussian random variables and the Bernstein techniques of large and small boxes. In 1972, C.Stein in [11] (see also [12]) introduced the new method to estimate the proximity of the distribution under consideration to a normal law. Furthermore, this powerful method was developed in the framework of classical limit theorems of the probability theory. We describe this method in Section 2. Applying the Stein method along with other tools, one can establish in certain cases the sharp estimates of closeness between a target distribution and other ones in specified metrics (see, e.g., [13,14]). We recommend the books [15,16] and the paper [17]

for basic ideas of the ingenious Stein method. The development of this techniques under mild moment restrictions for summands is treated in [18,19]. We mention in passing that there are deep generalizations of Stein techniques involving generators of certain Markov processes; a compact exposition is provided, e.g., on p. 2 of [20].

In the theory of random sums of random summands, the limit theorems with exponential law as a target distribution play a role similar to the central limit theorem for (nonrandom) sums of random variables. Here, one has to underline the principal role of the Rényi classical theorem for geometric sums published in [21]. Recall this famous result. Let  $X_1, X_2, \dots$  be a sequence of independent identically distributed (i.i.d.) random variables such that  $\mu := \mathbb{E}[X_1] \neq 0$ . Take a geometric random variable  $N_p$  with parameter  $p \in (0, 1)$ , defined as follows:

$$\mathbb{P}(N_p = k) = p(1 - p)^k, \quad k \in \mathbb{N} \cup \{0\}. \tag{2}$$

Assume that  $N_p$  and  $(X_n)_{n \in \mathbb{N}}$  are independent. Set  $S_0 := 0, S_n := X_1 + \dots + X_n, n \in \mathbb{N}$ . Then,

$$W_p := \frac{S_{N_p}}{\mathbb{E}[S_{N_p}]} \xrightarrow{\mathcal{D}} Z \sim \text{Exp}(1) \text{ as } p \rightarrow 0+, \tag{3}$$

where  $\xrightarrow{\mathcal{D}}$  stands for convergence in distribution, and  $Z$  follows the exponential law  $\text{Exp}(\lambda)$  with parameter  $\lambda = 1, \mathbb{E}[S_{N_p}] = \mu(1 - p)/p$ . In fact, instead of  $N_p$ , A.Rényi considered the shifted geometric random variable  $N(p)$  such that  $\mathbb{P}(N(p) = k) = p(1 - p)^{k-1}, k \in \mathbb{N}$ . Clearly,  $N_p$  has the same law as  $N(p) - 1$ . He supposed that i.i.d. random variables  $X_1, X_2, \dots$  are non-negative, and  $N(p)$  and  $(X_n)_{n \in \mathbb{N}}$  are independent. Then,  $S_{N(p)}/\mathbb{E}[S_{N(p)}]$  converges in distribution to  $Z \sim \text{Exp}(1)$  as  $p \rightarrow 0+$ , where  $\mathbb{E}[S_{N(p)}] = \mu/p$ . It was explained in [22] that both statements are equivalent and the assumption of nonnegativity of summands can be omitted.

Building on the previous investigations discussed below in this section, we study different instances of quantifying the approximation of random sums by limit laws and also extend the Stein method employment. The main goals of our paper are the following: (1) to find sharp estimates (i.e., optimal ones which cannot be diminished) of proximity of geometric sums of independent (in general non-identically distributed) random variables to exponential law using the probability metric  $d_{\mathcal{H}_2}$ ; (2) to prove the new version of the Rényi theorem when the summands are described by a model of exchangeable random variables, establishing the due non-exponential limit law together with an optimal bound of the convergence rate applying  $d_{\mathcal{H}_2}$ ; (3) to obtain the exact convergence rate of appropriately normalized random sums of random summands to the generalized gamma distribution when the number of summands follows the generalized negative binomial distribution employing  $d_{\mathcal{H}_2}$ ; (4) to introduce the inverse transformation to an “equilibrium distribution transformation”, give full description of its existence and demonstrate the advantage of applying the Stein method combined with that inverse transform; and (5) to use such approach in deriving the new approximation in the Kolmogorov metric  $d_{\mathcal{K}}$  of the Pareto distribution by an exponential one, which is important in signal processing.

The main idea is to apply the Stein method and deduce (Lemma 2) new estimates of the solution of Stein’s equation (corresponding to an exponential law  $\text{Exp}(\lambda)$  as a target distribution) when a function  $h$  appearing in its right-hand side belongs to a class  $\mathcal{H}_2$ . This entails the established sharp estimates. The integral probability metrics and the techniques of integration with respect to sign measures are essentially employed. It should be stressed that we consider random summands which take, in general, positive and negative values and in certain cases need not have the same law.

Now, we briefly comment on the relevance of the five groups of the paper results mentioned above. Some upper bounds for convergence rates in Equation (3) were obtained previously by different tools (the renewal techniques and the memoryless property of the geometric distribution), and the estimates were not sharp. We refer to the results by A.D. Soloviev, V.V. Kalashnikov and S.Y. Vsekhsvyatskii, M. Brown, V.M. Kruglov and V.Yu. Korolev, where the authors either used the Kolmogorov distance or proved specified

nonuniform estimates for differences of the corresponding distribution functions. For instance, in [23] the following estimate was proved

$$\sup_{x \in \mathbb{R}} |\mathbb{P}(W_p \leq x) - \mathbb{P}(Z \leq x)| \leq p \frac{\mathbb{E}[X_1^2]}{\mu^2} \max \left\{ 1, \frac{1}{2(1-p)} \right\},$$

where  $Z \sim \text{Exp}(1)$ . Moreover, this estimate is asymptotically exact when  $p \rightarrow 0+$ . Some improvements are in [24] under certain (hazard rate) assumptions. E.V. Sugakova obtained a version of the Rényi theorem for independent, in general, not identically distributed random variables. We also mention contributions by V.V. Kalashnikov, E.F. Peköz, A. Röllin, N. Ross and T.L. Hung which gave the estimates in terms of the Zolotarev ideal metrics. We do not reproduce all these results here since they can be viewed on pages 3 and 4 of [22] with references where they were published.

In Corollary 3.6 of [25] for nondegenerate i.i.d. positive random variables  $X_1, X_2, \dots$  with mean  $\mu$  and finite second moment, it was proved that

$$\zeta_2(pS(p), Z(1/\mu)) \leq p(\mathbb{E}[X_1^2] + 2\mu^2),$$

where  $S(p) := \sum_{j=1}^{N(p)} X_j$ ,  $\zeta_2$  is the Zolotarev ideal metric of order two,  $Z(\lambda) \sim \text{Exp}(\lambda)$ ,  $\lambda > 0$ . In [22], the estimates for proximity of geometric sums distributions to  $Z \sim \text{Exp}(1)$  were provided in the Kantorovich and  $\zeta_2$  metrics. A substantial contribution of the authors of [22] is the study of random summands  $X_1, X_2, \dots$  that need not be positive (see also [26]). The general estimate for deviation of  $W_p$  from  $Z \sim \text{Exp}(1)$  in the ideal metric of order  $s$  was proved in [27]. We do not assume that  $W_p$  is constructed by means of i.i.d. random variables and, moreover, demonstrate that our estimate (for summands taking real values) involving the metric  $d_{\mathcal{H}_2}$  is sharp.

The exchangeable random variables form an important class having various applications in statistics and combinatorics, see, e.g., [28]. As far as we know, the model of exchangeable random variables is studied in the context of random sums for the first time here. It is interesting that instead of the exponential limit law we indicate explicit expression of the new limit law. In addition, we establish the sharp estimate of proximity of random sums distributions to this law using  $d_{\mathcal{H}_2}$ .

A natural generalization of the Rényi theorem is to study the summation index following non-geometrical distribution. In this way, the upper bound of the convergence rate of random sums of random summands to generalized gamma distribution was proved in [29]. Theorem 3.1 in [30] contains the estimates in the Kolmogorov and Kantorovich distances for approximations of non-negative random variable law by specified (nongeneralized) gamma distribution. The proof relies on Stein’s identity for gamma distribution established in H.M.Luk’s PhD thesis (see the reference in [30]). New estimates of the solutions of the gamma Stein equation are given in [31]. We derive the sharp estimate for approximation of random sums by generalized gamma law using the Zolotarev metric of order two. In a quite recent paper [32] the author established deep results concerning further generalizations of the Rényi theorem. Namely, Theorem 1 of [32] demonstrates how one can provide the upper bounds of the convergence rate of specified random sums to a more general law than an exponential one using the estimates in the Rényi theorem. This approach is appealing since the author employs the ideal metric of order  $s > 0$ . However, the sharpness of these estimates was not examined.

Note that in [33] the important “equilibrium transformation of distributions” was proposed and employed along with the Stein techniques. We will consider this transformation  $X^e$  for a random variable  $X$  in Section 7 and also tackle other useful transformations. In the present paper, the inverse to the “equilibrium distribution transformation” is introduced. We completely describe the possibility to construct such transformation and provide an explicit formula for the corresponding density. The idea to apply such inverse transformation whenever it exists is based on the result [33] demonstrating that one can obtain a more

precise estimate for proximity in the Kantorovich metric between  $X^e$  and  $Z$  than between  $X$  and  $Z$ , where  $Z \sim \text{Exp}(1)$  and  $\mathbb{E}[X] = 1$ ,  $\mathbb{E}[X^2] < \infty$ . We extend this result. Moreover, we prove that in this way one can obtain a new estimate of approximation of the Pareto distribution by an exponential one. It is shown that our new estimate is advantageous for a wide range of parameters of the Pareto distribution. Let  $X^e \sim \text{Pareto}(\alpha, \beta)$ , i.e., the distribution function of  $X^e$  is

$$F^e(x) = 1 - \left( \frac{\beta}{x + \beta} \right)^\alpha, \quad x \geq 0, \quad \alpha > 0, \quad \beta > 0.$$

We show that the preimage  $X \sim \text{Pareto}(\alpha + 1, \beta)$ . Thus, for any  $\alpha > 2$ ,  $\beta > 0$ , one has  $d_{\mathcal{K}}(X^e, Z) \leq 1/(\alpha - 1)$ , where  $Z \sim \text{Exp}(\alpha/\beta)$  and  $d_{\mathcal{K}}$  stands for the Kolmogorov distance. This bound is more precise than the previous ones applied in signal processing, see, e.g., [34].

This paper is organized as follows. After the Introduction, the auxiliary results are provided in Section 2. Here we include the material important for understanding the main results. We recall the concept of probability metrics, consider the Kolmogorov and the Kantorovich distances and examine the Zolotarev ideal metrics. We describe the basic ideas of Stein's method, especially for the exponential target distribution. In this section, we formulate a simple but useful Lemma 1 concerning the essential supremum of the Lipschitz function, an important Lemma 2 giving the solution of the Stein equation for different functional classes. We explain the essential role of the generalized equilibrium transformation proposed in [22] which permits study of the summands taking both positive and negative values. We formulate Lemma 3 to be able to solve an integral equation involving the generalized equilibrium transformation when  $\mathbb{E}[X] \neq 0$  and  $\mathbb{E}[X^2] < \infty$ . The proofs of auxiliary lemmas are placed in Appendix A. Section 3 is devoted to an approximation of the normalized geometric sums  $W_p$  by an exponential law. Here, the sharp convergence rate is found (see Theorem 1) by means of the probability metric  $d_{\mathcal{H}_2}$ . The proof is based on the Lebesgue–Stieltjes integration techniques, the formula of integration by parts for functions of bounded variations, Lemma 2, various limit theorems for integrals and the important result of [22] concerning the estimates involving the Kantorovich distance. In Section 4, for the first time an analog of the Rényi theorem is proved for a model of exchangeable random variables proposed in [35]. We demonstrate (Theorem 2) that, in contrast to Rényi's theorem, the limit distribution for random sums under consideration is a specified mixture of two explicitly indicated laws. Moreover, the sharp convergence rate to this limit law is obtained (Theorem 3) by means of  $d_{\mathcal{H}_2}$ . In Section 5, the distance between the generalized gamma law and the suitably normalized sum of independent random variables is estimated when the number of summands has the generalized negative binomial distribution. Theorem 4 demonstrates that this estimate is sharp. For the proof, we employ various truncation techniques, the transformations of parameters of initial random variables, the monotone convergence theorem and explicit formula for the generalized gamma distribution moments of order  $\delta > 0$ , obtained in [27]. Section 6 provides the pioneering study of the same problem in the framework of exchangeable random variables and also gives the sharp estimate for the  $d_{\mathcal{H}_2}$  metric (Theorem 5). In Section 7, we introduce the inverse to the equilibrium transformation of the probability measures. Lemma 6 contains a full description of situations when a unique preimage  $X$  of a random variable  $X^e$  exists and gives an explicit formula for distribution of  $X$ . This approach permits us to obtain the new estimates of closeness of probability measures in the Kolmogorov and Kantorovich metrics (Theorem 6). In particular, due to Theorem 6 and Lemmas 2, 6, it becomes possible to find a useful estimate of proximity of the Pareto law to the exponential one (Example 2). Section 8 containing the conclusions and indications for further research work is followed by Appendix A and the list of references.

## 2. Auxiliary Results

Let  $\mathcal{K} := \{h : h_z(x) = \mathbb{I}\{x \leq z\}, x, z \in \mathbb{R}\}$ , where  $\mathbb{I}\{A\} := 1$  if  $A$  holds and zero otherwise. The choice  $\mathcal{H} = \mathcal{K}$  in Equation (1) corresponds to the Kolmogorov distance. Note that  $h$  above is a function in  $x$ , whereas  $z$  is the index parameterizing the class.

A function  $h : \mathbb{R} \rightarrow \mathbb{R}$  is called the Lipschitz one if

$$Lip(h) := \sup_{x,u \in \mathbb{R}; x \neq u} \frac{|h(x) - h(u)|}{|x - u|} < \infty. \tag{4}$$

Then,

$$|h(x) - h(u)| \leq C|x - u|, \quad x, u \in \mathbb{R}, \tag{5}$$

and in light of Equation (4),  $Lip(h)$  is the smallest possible constant  $C$  appearing in Equation (5). We write  $Lip(C)$ , where  $C \in [0, \infty)$  for a collection of the Lipschitz functions having  $Lip(h) \leq C$ . For  $s > 0$  set  $m = m(s) := \lceil s - 1 \rceil \in \mathbb{N} \cup \{0\}$  (where, for  $a \in \mathbb{R}$ ,  $\lceil a \rceil$  stands for the minimal integer number which is equal or greater than  $a$ ). Introduce a class of functions

$$\mathcal{H}_s := \{h : \mathbb{R} \rightarrow \mathbb{R}, |h^{(m)}(x) - h^{(m)}(u)| \leq |x - u|^{s-m}, x, u \in \mathbb{R}\}, \quad s > 0.$$

As usual,  $h^{(0)}(x) = h(x), x \in \mathbb{R}$ . We write  $d_{\mathcal{H}_s}$  for a metric defined according to Equation (1) with  $\mathcal{H} = \mathcal{H}_s$ . V.M. Zolotarev and many other researchers defined an ideal metric  $\zeta_s$  of order  $s > 0$  involving only bounded functions from  $\mathcal{H}_s$ . We will use collections  $\mathcal{H}_1$  and  $\mathcal{H}_2$  without assumption that functions  $h$  are bounded on  $\mathbb{R}$ . This is the reason why we write  $d_{\mathcal{H}_s}$  instead of  $\zeta_s$ . Thus, we employ

$$\mathcal{H}_1 := \{Lip(1)\}, \quad \mathcal{H}_2 := \{h : h' \in Lip(1)\}.$$

Note that in definitions of  $\mathcal{H}_2$  we deal with  $h \in C^{(1)}$ , where the space  $C^{(1)}(\mathbb{R})$  consists of functions  $h : \mathbb{R} \rightarrow \mathbb{R}$  such that  $h'(x)$  exists for all  $x \in \mathbb{R}$ , and  $h'$  is continuous on  $\mathbb{R}$  (evidently the Lipschitz function is continuous). One calls  $d_{\mathcal{H}_1}$  the Kantorovich metric (the term Wasserstein metric appears in the literature as well). One also uses the bounded Kantorovich metric when the class  $\mathcal{H}_1$  contains all the bounded functions from  $Lip(1)$ . The metric  $\zeta_s$  was introduced in [36] and called an ideal metric in light of its important properties. The properties of  $\zeta_s$  metrics, where  $s > 0$ , are collected in Sec. 2 of [32]. We mention in passing that various functionals are ubiquitous in assessing the proximity of distributions. In this regard, we refer, e.g., to [37,38].

To apply the Stein method, we begin with fixing the target random variable  $Z$  (or its distribution) and describe a class  $\mathcal{H}$  to estimate  $d_{\mathcal{H}}(Y, Z)$  for a random variable  $Y$  under consideration. Then, the problem is to indicate an operator  $T$  (with specified domain of definition) so that the Stein equation

$$Tf(x) = h(x) - \mathbb{E}[h(Z)] \tag{6}$$

has a solution  $f_h(x), x \in \mathbb{R}$ , for each function  $h \in \mathcal{H}$ . After that, one can substitute  $Y$  instead of  $x$  in Equation (6) and take the expectation of both sides, assuming that all these expectations are finite. As a result, one comes to the relation

$$\mathbb{E}[Tf_h(Y)] = \mathbb{E}[h(Y)] - \mathbb{E}[h(Z)]. \tag{7}$$

It is not a priori clear why the estimation of the left-hand side of Equation (7) is more adequate than the estimation of  $|\mathbb{E}[h(Y)] - \mathbb{E}[h(Z)]|$  for  $h \in \mathcal{H}$ . However, in many situations, justifying the method this occurs. The choice of  $T$  depends on the distribution of  $Z$ . Note that in certain cases (e.g., when  $Z$  follows the Poisson law) one considers functions  $f$  defined on a subset of  $\mathbb{R}$ . We emphasize that the construction of operator  $T$  is a nontrivial problem, see, e.g., [33,39–41].

The basic idea in this way is the following. For many probability distributions (Gaussian, Laplace, Exponential, etc.), one can find an operator  $T$  characterizing the law of a target variable  $Z$ . In other words, for a rather large class of functions  $f$ ,  $\mathbb{E}[Tf(Y)] = 0$  if and only if  $law(Y) = law(Z)$  (i.e., the laws of  $Y$  and  $Z$  coincide). Thus, if  $|\mathbb{E}[Tf_h(Y)]|$  is small enough for a suitable class of functions  $h$ , this leads to the assertion that the law of  $Y$  is close (in a sense) to the law of  $Z$ . One has to verify that this kind of “continuity” takes place. Clearly, if for any  $h \in \mathcal{H}$ , where  $\mathcal{H}$  defines the integral probability metric in Equation (1), one can find a solution  $f_h$  of Equation (6), then the relation  $\mathbb{E}[Tf_h(Y)] = 0$  for all  $f_h, h \in \mathcal{H}$ , yields  $d_{\mathcal{H}}(Y, Z) = 0$  and, consequently,  $law(Y) = law(Z)$ .

Further, we assume that  $Z \sim Exp(\lambda)$ , i.e.,  $Z$  has exponential distribution with parameter  $\lambda > 0$ . In this case (see, e.g., Sec. 5 in [17]), one uses the operator

$$Tf(x) := f'(x) - \lambda f(x) + \lambda f(0), \quad x \in \mathbb{R}, \quad \lambda > 0, \tag{8}$$

and writes the Stein Equation (6) as follows

$$f'(x) - \lambda f(x) + \lambda f(0) = h(x) - \mathbb{E}[h(Z)], \quad x \in \mathbb{R}. \tag{9}$$

It should be stipulated that  $\mathbb{E}[h(Z)] \in \mathbb{R}$  for a test function  $h \in \mathcal{H}$ , and there exists a differentiable solution  $f$  of Equation (9). Therefore, if one can find such solution  $f$ , then

$$\mathbb{E}[f'(Y)] - \lambda \mathbb{E}[f(Y)] + \lambda f(0) = \mathbb{E}[h(Y)] - \mathbb{E}[h(Z)] \tag{10}$$

under the hypothesis that all these expectations are finite. If  $f : \mathbb{R} \rightarrow \mathbb{R}$  is absolutely continuous, then (see, e.g., Theorem 13.18 of [42]) for almost all  $x \in \mathbb{R}$  with respect to the Lebesgue measure, there exists  $f'(x)$ . Moreover, one can find an integrable (on each interval) function  $g : \mathbb{R} \rightarrow \mathbb{R}, x \in \mathbb{R}$ , to guarantee, for each  $x, u \in \mathbb{R}$ , that

$$f(x) = f(u) + \int_u^x g(v)dv, \tag{11}$$

where  $g(v) = f'(v)$  for almost all  $v \in \mathbb{R}$ . Thus,  $(Tf)(x)$  is defined for such  $f$  according to Equation (8) for almost all  $x \in \mathbb{R}$ . In general, for an arbitrary random variable  $Y$ , one cannot write  $\mathbb{E}[(Tf)(Y)]$  since the value of expectation depends on the choice of a version of  $(Tf)(x), x \in \mathbb{R}$ . Really, let  $B \in \mathcal{B}(\mathbb{R})$  be such that  $m(B) = 0$ , where  $m$  stands for the Lebesgue measure. Assume that  $Y$  takes values in  $B$ . Then, it is clear that  $\mathbb{E}[(Tf)(Y)]$  depends on the choice of a function  $(Tf)(x)$  version defined on  $\mathbb{R}$ . However, if the distribution  $\mathbb{P}_Y$  of a random variable  $Y$  has a density with respect to  $m$ , then  $\mathbb{E}[(Tf)(Y)]$  will be the same for any version of  $Tf$  (with respect to the Lebesgue measure). In certain cases, the Stein operator is applied to smoothed functions (see, e.g., [33,43]). Otherwise, Equation (6) does not hold at each point of  $\mathbb{R}$  (see, e.g., Lemma 2.2 in [16]), and complementary efforts are needed. For our study, it is convenient to employ in Equation (8) for  $T$  in the capacity of  $f'(x), x \in \mathbb{R}$ , the right derivative. In many cases, for a real-valued function  $f$  defined on a fixed set  $D \subset \mathbb{R}$  one considers  $\sup_{x \in D} |f(x)|$  as “essential supremum”. Recall that a function  $\tilde{f}$  is a version of  $f$  (and vice versa) if the measure (here the Lebesgue measure) of points  $x$  such that  $\tilde{f}(x) \neq f(x)$  is zero. The notation  $\|f\|_\infty$  means that one takes  $\inf_{\tilde{f}} \sup_{x \in D} |\tilde{f}(x)|$ , where  $\tilde{f}$  belongs to the class of all versions of  $f$ . Clearly,  $\|f\|_\infty$  will be the same if we change  $f$  on a subset of  $D$  having a measure which is equal to zero. Thus, we write  $\|f'\|_\infty$  instead of  $\|g\|_\infty$  appearing in Equation (11). The following simple observation is useful. Its proof is provided in Appendix A.

**Lemma 1.** *A function  $h$  is the Lipschitz function on  $\mathbb{R}$  with  $Lip(h) = C < \infty$  if and only if  $h$  is absolutely continuous and (its essential supremum)  $\|h'\|_\infty = C < \infty$ .*

**Remark 1.** Note that  $0 \leq h(x) \leq 1, x \in \mathbb{R}$ , for any  $h \in \mathcal{K}$ . If, for some positive constant  $C, h \in \text{Lip}(C)$ , then Equation (5) yields that  $|h(x)| \leq C|x| + |h(0)|$ . If  $h'$  is a Lipschitz function (with  $\text{Lip}(h') = C$ ), then  $h''(x)$  exists for almost all  $x \in \mathbb{R}$  and an application of Lemma 1 gives

$$|h'(x) - h'(0)| = \left| \int_0^x h''(u)du \right| \leq C|x|, \quad x \in \mathbb{R}.$$

Consequently,  $|h'(x)| \leq A|x| + B$  for some positive  $A, B$  (one can take  $A = C, B = |h'(0)|$ ) and any  $x \in \mathbb{R}$ . As  $h'(x)$  is continuous on each interval, it follows that  $|h(x)| \leq ax^2 + b|x| + c$  for some positive  $a, b, c$  and all  $x \in \mathbb{R}$  ( $a = C/2, b = |h'(0)|, c = |h(0)|$ ). Therefore,  $|h(x)| \leq A_0x^2 + B_0$  for some positive  $A_0, B_0$  and each  $x \in \mathbb{R}$ .

**Lemma 2.** For any  $\lambda > 0$  and each  $h \in \mathcal{K} \cup \mathcal{H}_1 \cup \mathcal{H}_2$ , the equation

$$f'(x) - \lambda f(x) = h(x), \quad x \in \mathbb{R}, \tag{12}$$

has a solution

$$f_h(x) = -e^{\lambda x} \int_x^\infty h(u)e^{-\lambda u} du, \quad x \in \mathbb{R}, \tag{13}$$

where  $f_h(0) = -\mathbb{E}[h(Z)]/\lambda$ . If  $h \in \mathcal{K}$ , then for all  $x \in \mathbb{R}$  there exists  $f'_h(x)$  and  $\|f'_h\|_\infty \leq 1$ . If  $h \in \mathcal{H}_1 \cup \mathcal{H}_2$ , then  $f'_h$  is defined on  $\mathbb{R}$  and  $\|f'_h\|_\infty \leq \|h'\|_\infty/\lambda$ . For  $h \in \mathcal{H}_2$ , a function  $f''_h$  is defined on  $\mathbb{R}$  and  $\|f''_h\|_\infty \leq \min\{2\|h'\|_\infty, \|h''\|_\infty/\lambda\}$ .

The right-hand side of Equation (13) is well defined for each  $x \in \mathbb{R}$  in light of Remark 1. Lemma 4.1 of [33] contains for  $\lambda = 1$  some statements of Lemma 1. We will use the above estimates for any  $\lambda > 0$ . Estimates for  $h \in \mathcal{H}_2$  were not considered in [33]. The proof of Lemma 2 is given in Appendix A.

The following concept was introduced in [33].

**Definition 1 ([33]).** Let  $X$  be a non-negative random variable with finite  $\mathbb{E}[X] > 0$ . One says that a random variable  $X^e$  has distribution of equilibrium with respect to  $X$  if for any Lipschitz function  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,

$$\mathbb{E}[f(X)] - f(0) = \mathbb{E}[X]\mathbb{E}[f'(X^e)]. \tag{14}$$

Note that Definition 1 deals separately with distributions of  $X$  and  $X^e$ . One says that  $X^e$  is the result of the equilibrium transformation applied to  $X$ . The same terminology is used for transition from  $\text{law}(X)$  to  $\text{law}(X^e)$ . For the sake of completeness, we explain in Appendix A (Comments to Definition 1) why one can take the law of  $X^e$  having a density with respect to the Lebesgue measure

$$p^e(x) = \begin{cases} \frac{1}{\mathbb{E}[X]}\mathbb{P}(X > x), & x \geq 0, \\ 0, & x < 0, \end{cases} \tag{15}$$

to guarantee the validity of Equation (14).

**Remark 2.** For a non-negative random variable  $X$  with finite  $\mathbb{E}[X] > 0$ , one can construct a random variable  $X^e$  having a density (15). Accordingly, we then have a random vector  $(X, X^e)$  with specified marginal distributions. However, the joint law of  $X$  and  $X^e$  is not fixed and can be chosen in appropriate way. If  $X_1, X_2, \dots$  is a sequence of independent random variables, we will assume that a sequence  $(X_n, X_n^e)_{n \in \mathbb{N}}$  consists of independent vectors, and these vectors are independent with all considered random variables which are independent with  $(X_n)_{n \in \mathbb{N}}$ .

In the recent paper [22], a generalization of the equilibrium transformation of distributions was proposed without assuming that random variable  $X$  is non-negative.

**Definition 2** ([22]). Let  $X$  be a random variable having a distribution function  $F(x) := \mathbb{P}(X \leq x)$ ,  $x \in \mathbb{R}$ . Assume the existence of finite  $\mathbb{E}[X] \neq 0$ . An equilibrium distribution function corresponding to  $X$  (or  $F(x)$ ) is introduced by way of

$$F^e(x) := \begin{cases} -\frac{1}{\mathbb{E}[X]} \int_{-\infty}^x F(u)du, & x \leq 0, \\ -\frac{\mathbb{E}[X^-]}{\mathbb{E}[X]} + \frac{1}{\mathbb{E}[X]} \int_0^x (1 - F(u))du, & x > 0, \end{cases} \tag{16}$$

where  $X^- := X\mathbb{I}\{X < 0\}$ . This function can be written as  $F^e(x) = \int_{-\infty}^x p^e(u)du$ , where

$$p^e(x) = \begin{cases} -\frac{1}{\mathbb{E}[X]} F(x), & x \leq 0, \\ \frac{1}{\mathbb{E}[X]} (1 - F(x)), & x > 0, \end{cases} \tag{17}$$

thus,  $p^e$  is a density (with respect to the Lebesgue measure) of a signed measure  $Q^e$  corresponding to  $F^e$ . In other words, Equation (17) demonstrates the Jordan decomposition (see, e.g., Sec. 29 of [44]) of  $Q^e$ .

Clearly, for a non-negative random variable, the functions defined in Equation (15) and Equation (16) coincide. For a nonpositive random variable, the function  $F^e$  appearing in Equation (16) is a distribution function of a probability measure. In general, when  $X$  can take positive and negative values, the function introduced in Equation (16) is not a distribution function. We will call  $F^e$  the generalized equilibrium distribution function. Note that  $|p^e(x)| \leq \frac{1}{|\mathbb{E}[X]|}$ . Thus,  $F^e$  is the Lipschitz function and consequently continuous ( $F^e(x)$  is well defined for each  $x \in \mathbb{R}$  since  $\mathbb{E}[X]$  is finite and nonzero). Moreover,  $F^e$  is absolutely continuous being the Lipschitz function. Each absolutely continuous function has bounded variation. If  $G$  is a function of bounded variation, then  $G = G_1 - G_2$ , where  $G_1$  and  $G_2$  are nondecreasing functions (see, e.g., [42], Theorem 12.18). One can employ the canonical choice  $G_1(x) := Var_0^x(G)$ , where  $Var_a^b(G)$  means the variation of  $G$  on  $[a, b]$ ,  $-\infty < a \leq b < \infty$  (if  $a > b$  then  $Var_a^b(G) := -Var_b^a(G)$ ). If  $G$  is right-continuous (on  $\mathbb{R}$ ), then evidently  $G_1$  and  $G_2$  are also right-continuous. Thus, for a right-continuous  $G$  having bounded variation, a nondecreasing function  $G_i$  in its representation corresponds to a  $\sigma$ -finite measure  $Q_i$  on  $\mathcal{B}(\mathbb{R})$ ,  $i = 1, 2$ . More precisely, there exists a unique  $\sigma$ -finite measure  $Q_i$  on  $\mathcal{B}(\mathbb{R})$  such that, for each finite interval  $(a, b]$ ,  $Q_i((a, b]) = G_i(b) - G_i(a)$ ,  $i = 1, 2$ . Recall that one writes for the Lebesgue–Stieltjes integral with respect to a function  $G$

$$\int_{\mathbb{R}} f(u)dG(u) := \int_{\mathbb{R}} f(u)dG_1(u) - \int_{\mathbb{R}} f(u)dG_2(u), \tag{18}$$

whenever the integrals in the right-hand side exist (with values in  $[-\infty, \infty]$ ), and the cases  $\infty - \infty$  or  $-\infty + \infty$  are excluded. The integral  $\int_{\mathbb{R}} f(u)dG_i(u)$  means the integration with respect to measure  $Q_i$ ,  $i = 1, 2$ . The signed measure  $Q$  corresponding to  $G$  is  $Q_1 - Q_2$ . Thus,  $\int_{\mathbb{R}} f(u)dG(u)$  means the integration with respect to signed measure  $Q$ . Note that if  $G = U_1 - U_2$  where  $U_i$  is right-continuous and nondecreasing ( $i = 1, 2$ ), then

$$\int_{\mathbb{R}} f(u)dG_1(u) - \int_{\mathbb{R}} f(u)dG_2(u) = \int_{\mathbb{R}} f(u)dU_1(u) - \int_{\mathbb{R}} f(u)dU_2(u). \tag{19}$$

The left-hand side and the right-hand side of Equation (19) make sense simultaneously, and if so, are equal to each other. Indeed, for any finite interval  $(a, b]$  ( $a \leq b$ ), one has  $G_1(b) - G_1(a) - (G_2(b) - G_2(a)) = U_1(b) - U_1(a) - (U_2(b) - U_2(a))$ . Thus, the signed measures corresponding to  $G_1 - G_2$  and  $U_1 - U_2$  coincide on  $\mathcal{B}(\mathbb{R})$ . We mention in passing that one can also employ the Jordan decomposition of a signed measure.

For  $F^e$  introduced in Equation (16), the analog of Equation (15) has the form

$$\mathbb{E}[f(X)] - f(0) = \mathbb{E}[X] \int_{\mathbb{R}} f'(x)dF^e(x). \tag{20}$$

Taking into account Equation (17), one can rewrite Equation (20) equivalently as follows

$$\mathbb{E}[f(X)] - f(0) = \int_{(-\infty,0]} f'(x)(-F(x))dx + \int_{(0,\infty)} f'(x)(1 - F(x))dx. \tag{21}$$

The right-hand side of the latter relation does not depend on the choice of a version of  $f'$ . Due to Theorem 1(d) of [22], Equation (20) is valid for any Lipschitz function  $f$ . Evidently, an arbitrary function  $f \in \mathcal{H}_2$  need not be the Lipschitz one and vice versa.

**Lemma 3.** *Let  $X$  be a random variable such that  $\mathbb{E}[X^2] < \infty$  and  $\mathbb{E}[X] \neq 0$ . Then, Equation (20) is satisfied for all  $f \in \mathcal{H}_2$ .*

The proof is provided in Appendix A.

### 3. Limit Theorem for Geometric Sums of Independent Random Variables

Consider  $N_p \sim \text{Geom}(p)$ , see Equation (2). In other words,  $N_p$  has a geometric distribution with parameter  $p$ . Let  $X_1, X_2, \dots$  be a sequence of independent random variables such that  $\mathbb{E}[X_k] = \mu$ , where  $\mu \in \mathbb{R}$ ,  $\mu \neq 0$ ,  $k \in \mathbb{N}$ . Assume that  $N_p$  and  $(X_n)_{n \in \mathbb{N}}$  are independent. Consider a normalized geometric sum

$$W_p := \frac{p}{\mu(1-p)} \sum_{k=1}^{N_p} X_k, \tag{22}$$

introduced in Equation (3). Since  $N_p$  can take zero value, set, as usual,  $\sum_{k=1}^0 X_k := 0$ . One can see that  $W_p$  can be viewed as a random sum  $S_p := \sum_{k=1}^{N_p} X_k$  normalized by  $\mathbb{E}[X]\mathbb{E}[N_p]$ .

**Lemma 4.** *Let  $X_1, X_2, \dots$  and  $N_p$ , where  $p \in (0, 1)$ , be random variables described above in this Section. Then, the following relations hold:*

$$\begin{aligned} \mathbb{E}[W_p] &= 1, \quad \mathbb{E}|W_p| \leq \frac{\sup_{k \in \mathbb{N}} \mathbb{E}|X_k|}{|\mu|}, \\ \mathbb{E}[W_p^2] &= \frac{p}{\mu^2(1-p)} \mathbb{E}[X_{N_p+1}^2] + 2. \end{aligned} \tag{23}$$

**Proof.** Recall that

$$\mathbb{E}[N_p] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = \frac{1-p}{p}, \tag{24}$$

$$\mathbb{E}[N_p^2] = \sum_{k=1}^{\infty} k^2p(1-p)^{k-1} = \frac{(1-p)(2-p)}{p^2}. \tag{25}$$

Thus, one has

$$\mathbb{E}[W_p] = \frac{p}{\mu(1-p)} \sum_{k=1}^{\infty} k\mu\mathbb{P}(N_p = k) = \frac{p}{1-p} \mathbb{E}[N_p] = 1.$$

Clearly,  $\mathbb{E}|X_k| < \infty$  since  $\mathbb{E}[X_k]$  is finite ( $k \in \mathbb{N}$ ). Therefore

$$\mathbb{E}|W_p| \leq \frac{p}{|\mu|(1-p)} \sum_{k=1}^{\infty} k\mathbb{E}|X_k|\mathbb{P}(N_p = k) \leq \frac{\sup_{k \in \mathbb{N}} \mathbb{E}|X_k|}{|\mu|}.$$

Set  $v_k := \mathbb{E}[X_k^2]$ ,  $k \in \mathbb{N}$ . One has

$$\mathbb{E}[S_p^2] = \sum_{k=1}^{\infty} \mathbb{P}(N_p = k) \mathbb{E}\left(\sum_{i=1}^k X_i\right)^2 = \sum_{k=1}^{\infty} p(1-p)^k \left(\sum_{i=1}^k v_i + k(k-1)\mu^2\right). \tag{26}$$

According to Equations (24) and (25) one derives the formula

$$\sum_{k=1}^{\infty} p(1-p)^k(k(k-1)\mu^2) = \mu^2 \left( \frac{(1-p)(2-p)}{p^2} - \frac{1-p}{p} \right) = 2 \left( \frac{\mu(1-p)}{p} \right)^2. \tag{27}$$

Convergence of the series  $\sum_{k=1}^{\infty} p(1-p)^k \sum_{i=1}^k v_i$  having non-negative terms holds simultaneously with the validity of inequality  $\mathbb{E}[W_p^2] < \infty$ . Changing the order of summation, we obtain

$$\sum_{k=1}^{\infty} p(1-p)^k \sum_{i=1}^k v_i = \sum_{i=1}^{\infty} (1-p)^i v_i = \left( \frac{1-p}{p} \right) \mathbb{E}[X_{N_p+1}^2].$$

The latter formula and Equations (26), (27) yield

$$\begin{aligned} \mathbb{E}[W_p^2] &= \left( \frac{p}{\mu(1-p)} \right)^2 \mathbb{E}[S_p^2] = \left( \frac{p}{\mu(1-p)} \right)^2 \left( \left( \frac{1-p}{p} \right) \mathbb{E}[X_{N_p+1}^2] + 2 \left( \frac{\mu(1-p)}{p} \right)^2 \right) \\ &= \frac{p}{\mu^2(1-p)} \mathbb{E}[X_{N_p+1}^2] + 2. \end{aligned}$$

Equation (23) is established.  $\square$

The proof of Theorem 3.1 in [45] shows for non-negative i.i.d. random variables  $X_1, X_2, \dots$  (when  $\mu = 1$ , see Formula (3.15) in [45]) that the equilibrium transformation of  $W_p$  distribution has the following form:

$$W_p^e = \frac{p}{\mu(1-p)} \left( \sum_{k=1}^{N_p} X_k + X_{N_p+1}^e \right) = W_p + \frac{p}{\mu(1-p)} X_{N_p+1}^e, \tag{28}$$

where  $X_{N_p+1}^e$  means that we construct  $X_1^e, X_2^e, \dots$  and then take a random index  $N_p + 1$ . In other words,

$$X_{N_p+1}^e = \sum_{n=0}^{\infty} X_{n+1}^e \mathbb{I}\{N_p = n\}.$$

It was explained in Section 2 that a generalized equilibrium distribution function  $F_{W_p}^e(x)$  (see Definition 2) need not be a distribution function when the summands  $X_1, X_2, \dots$  can take values of different signs. However, employing this function, one can establish the following result.

**Theorem 1.** *Let  $X_1, X_2, \dots$  be a sequence of independent random variables having finite  $\mathbb{E}[X_k] = \mu$ , where  $\mu \neq 0, k \in \mathbb{N}$ . Assume that  $N_p$  and  $(X_n)_{n \in \mathbb{N}}$  are independent, where  $N_p \sim \text{Geom}(p)$ ,  $0 < p < 1$ . If  $Z \sim \text{Exp}(1)$ , then*

$$d_{\mathcal{H}_2}(W_p, Z) = \frac{\mathbb{E}[X_{N_p+1}^2]}{2\mu^2} \left( \frac{p}{1-p} \right) \tag{29}$$

where  $W_p$  was introduced in Equation (22).

**Proof.** If  $\mathbb{E}[W_p^2] = \infty$ , then  $d_{\mathcal{H}_2}(W_p, Z) = \infty$  since, for a function  $h(x) = x^2/2, x \in \mathbb{R}$ , belonging to  $\mathcal{H}_2$ , one has  $\mathbb{E}[h(W_p)] = \infty$ , whereas  $\mathbb{E}[h(Z)] < \infty$ . According to Equation (23),  $\mathbb{E}[W_p^2]$  and  $\mathbb{E}[X_{N_p+1}^2]$  are both finite or infinite simultaneously. Consequently, Equation (29) is true when  $\mathbb{E}[W_p^2] = \infty$ .

Let us turn to the case  $\mathbb{E}[W_p^2] < \infty$ . At first, we obtain an upper bound for  $d_{\mathcal{H}_2}(W_p, Z)$ . Take  $h \in \mathcal{H}_2$ . Applying Lemmas 1 and 2 and Remark 1, one can write due to Stein's Equation (10) that

$$|\mathbb{E}[h(W_p)] - \mathbb{E}[h(Z)]| = |\mathbb{E}[f'_h(W_p)] - \mathbb{E}[f_h(W_p)] + f(0)|. \tag{30}$$

Using the generalized equilibrium distribution transformation (20) one obtains:

$$|\mathbb{E}[f'_h(W_p)] - \mathbb{E}[f_h(W_p)] + f(0)| = \left| \int_{\mathbb{R}} f'_h(x) dF_{W_p}(x) - \int_{\mathbb{R}} f'_h(x) dF_{W_p}^e(x) \right|. \tag{31}$$

Due to Lemma 3 this is true, for  $h \in \mathcal{H}_2$ , because  $f_h \in \mathcal{H}_2$  according to Lemma 2 (with  $\lambda = 1$ ). Next, we employ the relation

$$\int_{\mathbb{R}} f'_h(x) dF_{W_p}(x) - \int_{\mathbb{R}} f'_h(x) dF_{W_p}^e(x) = \int_{\mathbb{R}} f'_h(x) d(F_{W_p} - F_{W_p}^e)(x). \tag{32}$$

Evidently, one can write  $\int_{\mathbb{R}} |f'_h(x)| dF_{W_p}(x) < \infty$ . The notation  $dF_{W_p}^e(x)$  in the integral refers to the Lebesgue–Stieltjes integral with respect to a function  $F_{W_p}^e(x)$  of bounded variation. In fact, the integral with integrator  $dF_{W_p}^e(x)$  means that integration employs a signed measure  $Q_p^+ - Q_p^-$ , where  $Q_p^+$  and  $Q_p^-$  have the following densities with respect to the Lebesgue measure:

$$q_p^+(x) := (1 - F_{W_p}(x))\mathbb{I}\{(0, \infty)\}, \quad q_p^-(x) := F_{W_p}(x)\mathbb{I}\{(-\infty, 0]\}, \quad x \in \mathbb{R},$$

we took into account that  $\mathbb{E}[W_p] = 1$  according to Lemma 4. Then, for any  $-\infty < a < b < \infty$ , one ascertains that variation of  $F_{W_p}^e$  on  $[a, b]$  is given by formula  $Var_a^b(F_{W_p}^e) = \int_a^b |p_{W_p}^e(u)| du$  (see, e.g., Theorem 4.4.7 [46]). Note that for any  $-\infty < a < b < \infty$ ,

$$\int_a^b |p_{W_p}^e(u)| du \leq \mathbb{E}|W_p| < \infty$$

according to Lemma 4. Thus,  $F_{W_p}^e$  is a function of bounded variation. In the right-hand side of Equation (32), we take the Lebesgue–Stieltjes integral with respect to the function of bounded variation  $(F_{W_p} - F_{W_p}^e)(x)$ ,  $x \in \mathbb{R}$ . Let  $F_{W_p}^e(x) = F_{p,1}^e(x) - F_{p,2}^e(x)$ ,  $x \in \mathbb{R}$ , where  $F_{p,i}^e$  are nondecreasing right-continuous functions (even continuous since  $F_{W_p}^e$  is continuous),  $i = 1, 2$ . Thus,

$$F_{W_p}(x) - F_{W_p}^e(x) = (F_{W_p}(x) + F_{p,2}^e(x)) - F_{p,1}^e(x), \quad x \in \mathbb{R}.$$

With the help of Equations (18) and (19) one makes sure that, for each  $n \in \mathbb{N}$ ,

$$\begin{aligned} \int_{(-n,n]} f'_h(x) d(F_{W_p} - F_{W_p}^e)(x) &= \int_{(-n,n]} f'_h(x) d(F_{W_p}(x) + F_{p,2}^e(x)) - \int_{(-n,n]} f'_h(x) d(F_{p,1}^e(x)) \\ &= \int_{(-n,n]} f'_h(x) dF_{W_p}(x) + \int_{(-n,n]} f'_h(x) dF_{p,2}^e(x) - \int_{(-n,n]} f'_h(x) dF_{p,1}^e(x) \\ &= \int_{(-n,n]} f'_h(x) dF_{W_p}(x) - \int_{(-n,n]} f'_h(x) d(F_{p,1}^e(x) - F_{p,2}^e(x)) \\ &= \int_{(-n,n]} f'_h(x) dF_{W_p}(x) - \int_{(-n,n]} f'_h(x) dF_{W_p}^e(x). \end{aligned}$$

All the integrals in the latter formulas are finite. According to Lemma 2 and Remark 1, one can write  $|f'_h(x)| \leq A_0|x| + B_0$ , where  $A_0, B_0$  are positive constants. Thus, the Lebesgue theorem on dominated convergence ensures that

$$\lim_{n \rightarrow \infty} \int_{(-n,n]} f'_h(x) dF_{W_p}(x) = \int_{\mathbb{R}} f'_h(x) dF_{W_p}(x),$$

where the latter integral is finite. Indeed,

$$\int_{\mathbb{R}} (A_0|x| + B_0) dF_{W_p}(x) = A_0\mathbb{E}|W_p| + B_0 < \infty \tag{33}$$

according to Lemma 4. By the same Lemma, one has  $\mathbb{E}[W_p] = 1$ . Therefore, on account of Equation (17), the following relation holds:

$$\int_{(-n,n]} f'_h(x) dF_{W_p^e}(x) = \int_{(-n,0]} f'_h(x)(-F_{W_p}(x))dx + \int_{(0,n]} f'_h(x)(1 - F_{W_p}(x))dx,$$

whereas Corollary 2, Sec. 6, Ch. II of [47] and Lemma 4 entail that

$$\int_{(-\infty,0]} (A_0|x| + B_0)F_{W_p}(x)dx + \int_{(0,\infty)} (A_0|x| + B_0)(1 - F_{W_p}(x))dx \leq A_0\mathbb{E}[W_p^2] + B_0\mathbb{E}|W_p| < \infty. \quad (34)$$

The Lebesgue theorem on dominated convergence for  $\sigma$ -finite measures and Equation (34) yield

$$\lim_{n \rightarrow \infty} \int_{(-n,n]} f'_h(x) dF_{W_p^e}(x) = \int_{\mathbb{R}} f'_h(x) dF_{W_p^e}(x),$$

where the latter integral is finite. Now, we show that

$$\lim_{n \rightarrow \infty} \int_{(-n,n]} f'_h(x) d(F_{W_p} - F_{W_p^e})(x) = \int_{\mathbb{R}} f'_h(x) d(F_{W_p} - F_{W_p^e})(x). \quad (35)$$

Note that  $f'_h(x)\mathbb{I}_{(-n,n]}(x) \rightarrow f'_h(x)$  at each  $x \in \mathbb{R}$  as  $n \rightarrow \infty$ . To apply the version of the Lebesgue theorem to integrals over a signed measure, it suffices (see, e.g., [48], p. 74) to verify that

$$\int_{\mathbb{R}} |f'_h(x)| |d(F_{W_p} - F_{W_p^e})(x)| < \infty,$$

where  $|dG|$  means that one evaluates an integral with respect to the measure corresponding to the total variation of a measure determined by a right-continuous function  $G$  of bounded variation. The extension of the Lebesgue theorem on dominated convergence for signed measures is an immediate corollary of the Jordan decomposition mentioned above. Using this decomposition, one obtains the inequality

$$\int_{\mathbb{R}} |f'_h(x)| |d(F_{W_p} - F_{W_p^e})(x)| \leq \int_{\mathbb{R}} |f'_h(x)| |dF_{W_p}(x)| + \int_{\mathbb{R}} |f'_h(x)| |dF_{W_p^e}(x)|.$$

Due to Remark 1 one has  $|f'_h(x)| \leq A_0|x| + B_0$  for all  $x \in \mathbb{R}$  and some positive constants  $A_0, B_0$ . Then, Equations (33) and (34) yield (as  $F_{W_p}$  generates probability measure)

$$\int_{\mathbb{R}} (A_0|x| + B_0) dF_{W_p}(x) + \int_{\mathbb{R}} (A_0|x| + B_0) |dF_{W_p^e}(x)| < \infty.$$

The functions  $f'_h$  and  $F_{W_p} - F_{W_p^e}$  are right-continuous and have bounded variation. Then each of them can be represented as the difference of right-continuous nondecreasing functions, and using for any  $n \in \mathbb{N}$  the integration by parts formula (see, e.g., Theorem 11, Sec. 6, Ch. 2, [47]), one has

$$\begin{aligned} \int_{(-n,n]} f'_h(x) d(F_{W_p} - F_{W_p^e})(x) &= f'_h(x)(F_{W_p}(x) - F_{W_p^e}(x))|_{-n}^n - \int_{(-n,n]} (F_{W_p}(x) - F_{W_p^e}(x)) df'_h(x). \end{aligned}$$

Since the integral in the right-hand side of Equation (35) is finite, it holds

$$f'_h(x)(F_{W_p}(x) - F_{W_p^e}(x)) \rightarrow 0, \quad x \rightarrow -\infty \text{ or } x \rightarrow \infty \quad (36)$$

(the proof is similar to the proof of Corollary 2, Sec. 6, Ch. 2 in [47]). Then,

$$\int_{\mathbb{R}} f'_h(x) d(F_{W_p} - F_{W_p}^e)(x) = - \lim_{n \rightarrow \infty} \int_{(-n,n]} (F_{W_p}(x) - F_{W_p}^e(x)) df'_h(x).$$

The function  $f'_h$  is absolutely continuous according to Lemma 2. Hence (see also Equations (36) and (A12) in Appendix A) we get

$$\begin{aligned} \left| \int_{\mathbb{R}} f'_h(x) d(F_{W_p}(x) - F_{W_p}^e(x)) \right| &= \lim_{n \rightarrow \infty} \left| \int_{(-n,n]} (F_{W_p}(x) - F_{W_p}^e(x)) f''_h(x) dx \right| \\ &\leq \|f''_h\|_{\infty} \int_{\mathbb{R}} |F_{W_p}(x) - F_{W_p}^e(x)| dx \leq \int_{\mathbb{R}} |F_{W_p}(x) - F_{W_p}^e(x)| dx, \end{aligned} \tag{37}$$

because  $\|f''_h\|_{\infty} \leq \|h'\|_{\infty} \leq 1$  due to Lemmas 1 and 2. Using the homogeneity of the Kantorovich metric for signed measures which is derived from formula (20) of [22] (see Lemma 1 (a) there) and applying Lemma 3 of that paper, we can write

$$\begin{aligned} \int_{\mathbb{R}} |F_{W_p}(x) - F_{W_p}^e(x)| dx &= \frac{p}{|\mu|(1-p)} \int_{\mathbb{R}} |F_{S_{N_p}}(x) - F_{S_{N_p}}^e(x)| dx \\ &\leq \frac{\mathbb{E}[X_{N_p+1}^2]}{2\mu^2} \left( \frac{p}{1-p} \right). \end{aligned} \tag{38}$$

Relations (30), (31), (32), (37), (38) and Lemmas 1 and 2 guarantee that  $d_{\mathcal{H}_2}(W_p, Z)$  does not exceed the right-hand side of Equation (29).

Now, we turn to the lower bounds for  $d_{\mathcal{H}_2}(W_p, Z)$ . Choose  $h(x) = x^2/2$  as the test function. Since  $h \in \mathcal{H}_2$ , we can write

$$d_{\mathcal{H}_2}(W_p, Z) \geq |\mathbb{E}[h(W_p)] - \mathbb{E}[h(Z)]| = \frac{1}{2} |\mathbb{E}[W_p^2] - \mathbb{E}[Z^2]|. \tag{39}$$

For a random variable  $Z$  following the exponential law  $Exp(1)$ , one has  $\mathbb{E}[Z^2] = 2$ . Formula (23) of Lemma 4 yields

$$d_{\mathcal{H}_2}(W_p, Z) \geq \frac{\mathbb{E}[X_{N_p+1}^2]}{2\mu^2} \left( \frac{p}{1-p} \right).$$

Taking into account formula (38), we come to the desired statement. The proof is complete.  $\square$

**Remark 3.** Evidently,

$$\mathbb{E}[X_{N_p+1}^2] = \sum_{n=0}^{\infty} \mathbb{E}[X_{n+1}^2] p(1-p)^n.$$

Thus, one obtains

$$\mathbb{E}[X_{N_p+1}^2] \leq \sup_{n \in \mathbb{N}} \mathbb{E}[X_n^2],$$

and the latter inequality becomes an equality when  $\mathbb{E}[X_n^2] = \mathbb{E}[X_1^2]$  for all  $n \in \mathbb{N}$ . Therefore, the statement of Theorem 1 can be written as follows

$$d_{\mathcal{H}_2}(W_p, Z) \leq \frac{\sup_{n \in \mathbb{N}} \mathbb{E}[X_n^2]}{2\mu^2} \left( \frac{p}{1-p} \right),$$

and this becomes an equality when  $\mathbb{E}[X_n^2] = \mathbb{E}[X_1^2]$  for all  $n \in \mathbb{N}$ .

**Remark 4.** In [22], the authors proved the following inequality

$$d_{\mathcal{H}_2}(W_p, Z) \leq \frac{3\mathbb{E}[X_{N_p+1}^2]}{2\mu^2} \left( \frac{p}{1-p} \right).$$

We established the sharp estimate with a factor  $1/2$  instead of  $3/2$  having employed Equation (20) for a class of functions comprising solutions of the Stein equation for  $h \in \mathcal{H}_2$ . The estimate with factor  $1/2$  was also obtained in the recent paper [49] but for i.i.d. summands. The lower bounds were not provided there. In our Theorem 1, the summands have the same expectations but need not have the same distribution.

**Remark 5.** If the summands of  $W_p$  are non-negative, we consider  $W_p^e$  appearing in Equation (28). Applying Theorem 1(i) [22] to relation (29), one obtains

$$d_{\mathcal{H}_1}(W_p^e, Z) = \frac{\mathbb{E}[X_{N_p+1}^2]}{2\mu^2} \frac{p}{1-p}.$$

For  $i \in \mathbb{N}$ , consider a random variable  $X_i$  having distribution  $\text{Exp}(1/\mu)$ . Then  $X_i^e \sim \text{Exp}(1/\mu)$ , and, consequently,  $X_{N_p+1}^e \sim \text{Exp}(1/\mu)$ . We can choose  $X_i^e, i \in \mathbb{N}$ , according to Remark 2. Then, the distribution of  $W_p^e$  will be the same if we change  $X_{N_p+1}^e$  to  $X_{N_p+1}$  in Equation (28). In such a way,  $W_p^e$  is a normalized sum of a random number of independent random variables. Using the homogeneity of the Kantorovich metric, one has

$$d_{\mathcal{H}_1}\left(\frac{p}{\mu} \sum_{k=1}^{N_p+1} X_k, (1-p)Z\right) = (1-p)d_{\mathcal{H}_1}\left(\frac{p}{\mu(1-p)} \sum_{k=1}^{N_p+1} X_k, Z\right) = \frac{\mathbb{E}[X_{N_p+1}^2]}{2\mu^2} p. \tag{40}$$

Therefore, for an arbitrary sequence  $(X_k)_{k \in \mathbb{N}}$  satisfying conditions of Theorem 1, the upper bound for the left-hand side of Equation (40) is not less than the right-hand side of Equation (40).

#### 4. Limit Theorem for Geometric Sums of Exchangeable Random Variables

Now, we consider exchangeable random variables  $X_1, X_2, \dots$  satisfying the dependence condition proposed in [35]. Namely, assume that for all  $n \in \mathbb{N}, t_j \in \mathbb{R} (j = 1, \dots, n)$  and some  $\rho \in [0, 1]$

$$\mathbb{E}\left[e^{i(t_1 X_1 + \dots + t_n X_n)}\right] = \rho \mathbb{E}\left[e^{i X_1(t_1 + \dots + t_n)}\right] + (1-\rho) \prod_{j=1}^n \mathbb{E}\left[e^{i t_j X_j}\right], \tag{41}$$

where  $i^2 = -1$ . The cases of  $\rho = 0$  and  $\rho = 1$  correspond, respectively, to independent random variables and those possessing the property of comonotonicity. The latter means that for  $\rho = 1$  the joint behavior of  $X_1, \dots, X_n$  is strongly correlated and coincides with one of a vector  $(X_1, \dots, X_1)$ .

**Theorem 2.** Let  $X_1, X_2, \dots$  be exchangeable random variables with  $\mathbb{E}[X_1] = \mu, \mu \neq 0$  satisfying condition (41) for some  $\rho \in (0, 1)$ . Suppose that  $(X_n)_{n \in \mathbb{N}}$  and  $N_p$  are independent, where  $N_p \sim \text{Geom}(p), p \in (0, 1)$ . In contrast to the Rényi theorem, one has

$$W_p \xrightarrow{D} Y, \quad p \rightarrow 0+,$$

where the law of  $Y$  is the following mixture

$$\mathbb{P}_Y = \rho \mathbb{P}_{V X_1/\mu} + (1-\rho) \mathbb{P}_Z, \tag{42}$$

random variables  $X_1, V$  are independent and  $V \sim \text{Exp}(1), Z \sim \text{Exp}(1)$ .

**Proof.** Let  $\tilde{X}_1, \tilde{X}_2, \dots$  be independent copies of  $X_1, X_2, \dots$ , respectively. Suppose that  $\tilde{X}_1, \tilde{X}_2, \dots$  are independent with  $N_p$ . Set  $S_0 := 0, \tilde{S}_0 := 0, \tilde{S}_n := \tilde{X}_1 + \dots + \tilde{X}_n, n \in \mathbb{N}$ . Denote the characteristic function of a random variable  $\zeta$  by  $f_\zeta(t), t \in \mathbb{R}$ . For each  $t \in \mathbb{R}$ , using Equation (41), one has

$$\begin{aligned}
 f_{S_{N_p}}(t) &= \sum_{n=0}^{\infty} \mathbb{E} \left[ e^{itS_n} \right] \mathbb{P}(N_p = n) \\
 &= \mathbb{P}(N_p = 0) + \sum_{n=1}^{\infty} \left( \rho \mathbb{E} \left[ e^{iX_1tn} \right] + (1 - \rho) \prod_{j=1}^n \mathbb{E} \left[ e^{iX_j} \right] \right) \mathbb{P}(N_p = n) \\
 &= p + \sum_{n=0}^{\infty} \left( \rho \mathbb{E} \left[ e^{iX_1tn} \right] + (1 - \rho) \prod_{j=1}^n \mathbb{E} \left[ e^{i\tilde{X}_j} \right] \right) \mathbb{P}(N_p = n) - \rho p - (1 - \rho)p \\
 &= \rho f_{X_1 N_p}(t) + (1 - \rho) \sum_{n=0}^{\infty} f_{\tilde{S}_n}(t) \mathbb{P}(N_p = n) = \rho f_{X_1 N_p}(t) + (1 - \rho) f_{\tilde{S}_{N_p}}(t).
 \end{aligned}$$

For each  $t \in \mathbb{R}$ , one has

$$f_{W_p}(t) = \rho f_{\frac{p}{\mu(1-p)} X_1 N_p}(t) + (1 - \rho) f_{\tilde{W}_p}(t), \tag{43}$$

where  $\tilde{W}_p = \frac{p}{\mu(1-p)} \sum_{j=1}^{N_p} \tilde{X}_j$ .

According to the classical Rényi theorem,  $\tilde{W}_p \xrightarrow{\mathcal{D}} Z$  as  $p \rightarrow 0+$ , where  $Z \sim \text{Exp}(1)$ . Note that  $T_p := \frac{p}{1-p} N_p \xrightarrow{\mathcal{D}} V$  as  $p \rightarrow 0+$ , where  $V \sim \text{Exp}(1)$ . In fact, one can apply Theorem 1 with  $X_j \equiv 1, j \in \mathbb{N}$  to check this. For each  $t \in \mathbb{R}$ , taking into account that  $T_p$  and  $X_1$  are independent and applying the Lebesgue theorem on dominated convergence, we see that

$$\mathbb{E} \left[ e^{itT_p X_1} \right] = \mathbb{E} \left[ \mathbb{E} \left[ e^{itT_p X_1} \mid X_1 \right] \right] = \int_{\mathbb{R}} e^{itT_p x} dF_{X_1}(x) \rightarrow \int_{\mathbb{R}} e^{itV x} dF_{X_1}(x) = \mathbb{E} \left[ e^{itV X_1} \right], \quad p \rightarrow 0+,$$

since  $X_1$  and  $V$  are independent. Hence,

$$\frac{p}{\mu(1-p)} X_1 N_p \xrightarrow{\mathcal{D}} \frac{V X_1}{\mu}, \quad p \rightarrow 0+$$

is true. In light of Equation (43),

$$W_p \xrightarrow{\mathcal{D}} Y, \quad p \rightarrow 0+,$$

here the law of  $Y$  is the mixture of distributions  $V X_1 / \mu$  and  $Z$  provided by Equation (42). The proof is complete.  $\square$

**Theorem 3.** Assume that  $N_p$  and  $(X_n)_{n \in \mathbb{N}}$  satisfy conditions of Theorem 2. Let  $\mu_2 = \mathbb{E}[X_1^2]$ . Then,

$$d_{\mathcal{H}_2}(W_p, Y) = \frac{\mu_2}{2\mu^2} \left( \frac{p}{1-p} \right). \tag{44}$$

**Proof.** Relation (43) for characteristic functions implies that the following equality of distributions holds

$$W_p \stackrel{\mathcal{D}}{=} \frac{p}{\mu(1-p)} \left( (1 - \mathbb{I}_\rho) N_p X_1 + \mathbb{I}_\rho \tilde{S}_{N_p} \right), \tag{45}$$

where indicator  $\mathbb{I}_\rho$  equals 1 and 0 with probabilities  $1 - \rho$  and  $\rho$ , respectively, and is independent of all the variables under consideration. Assume at first that  $\mu_2 < \infty$ . Then, for  $h \in \mathcal{H}_2$ ,

$$\mathbb{E}[h(W_p)] = \rho \mathbb{E} \left[ h \left( \frac{p}{\mu(1-p)} N_p X_1 \right) \right] + (1 - \rho) \mathbb{E}[h(\tilde{W}_p)].$$

In view of Equation (42) one has

$$\mathbb{E}[h(Y)] = \rho \mathbb{E}\left[h\left(\frac{VX_1}{\mu}\right)\right] + (1 - \rho)\mathbb{E}[h(Z)].$$

The latter two formulas and the triangle inequality yield

$$\begin{aligned} &|\mathbb{E}[h(W_p)] - \mathbb{E}[h(Y)]| \\ &\leq \rho \left| \mathbb{E}\left[h\left(\frac{p}{\mu(1-p)} N_p X_1\right)\right] - \mathbb{E}\left[h\left(\frac{VX_1}{\mu}\right)\right] \right| + (1 - \rho) |\mathbb{E}[h(\tilde{W}_p)] - \mathbb{E}[h(Z)]|. \end{aligned} \tag{46}$$

By means of Theorem 1 we have

$$\sup_{h \in \mathcal{H}_2} |\mathbb{E}[h(\tilde{W}_p)] - \mathbb{E}[h(Z)]| = \frac{\mu_2}{2\mu^2} \left(\frac{p}{1-p}\right). \tag{47}$$

For each  $h \in \mathcal{H}_2$ , taking into account the independence of  $X_1, N_p, V$ , one can write

$$\begin{aligned} &\left| \mathbb{E}\left[h\left(\frac{p}{\mu(1-p)} N_p X_1\right)\right] - \mathbb{E}\left[h\left(\frac{VX_1}{\mu}\right)\right] \right| \\ &= \left| \int_{\mathbb{R}} \left( \mathbb{E}\left[h\left(\frac{p}{\mu(1-p)} N_p X_1\right)\right] - \mathbb{E}\left[h\left(\frac{xV}{\mu}\right)\right] \right) dF_{X_1}(x) \right|. \end{aligned}$$

Due to homogeneity of  $d_{\mathcal{H}_2}$  we infer from Theorem 1 that

$$\begin{aligned} &\sup_{h \in \mathcal{H}_2} \left| \mathbb{E}\left[h\left(\frac{p}{\mu(1-p)} N_p X_1\right)\right] - \mathbb{E}\left[h\left(\frac{xV}{\mu}\right)\right] \right| = d_{\mathcal{H}_2}\left(\frac{px}{\mu(1-p)} N_p, \frac{xV}{\mu}\right) \\ &= \left(\frac{x}{\mu}\right)^2 d_{\mathcal{H}_2}\left(\frac{p}{(1-p)} \sum_{k=1}^{N_p} 1, V\right) = \frac{1}{2} \left(\frac{x}{\mu}\right)^2 \frac{p}{1-p}. \end{aligned}$$

Consequently, it holds

$$\begin{aligned} &\left| \mathbb{E}\left[h\left(\frac{p}{\mu(1-p)} N_p X_1\right)\right] - \mathbb{E}\left[h\left(\frac{VX_1}{\mu}\right)\right] \right| \\ &\leq \frac{p}{2(1-p)} \int_{\mathbb{R}} \left(\frac{x}{\mu}\right)^2 dF_{X_1}(x) = \frac{\mu_2}{2\mu^2} \left(\frac{p}{1-p}\right). \end{aligned} \tag{48}$$

Equations (46), (47) and (48) lead to the upper bound for  $d_{\mathcal{H}_2}(W_p, Y)$ .

Note that a function  $h(x) = x^2/2, x \in \mathbb{R}$ , belongs to  $\in \mathcal{H}_2$  and therefore

$$\sup_{\mathcal{H}_2} |\mathbb{E}[h(W_p)] - \mathbb{E}[h(Y)]| \geq \frac{1}{2} (\mathbb{E}[W_p^2] - \mathbb{E}[Y^2]). \tag{49}$$

Note that  $\mathbb{E}[Z^2] = \mathbb{E}[V^2] = 2$  because  $Z \sim Exp(1)$  and  $V \sim Exp(1)$ . The random variables  $X_1, V, Z$  are independent. Thus, in light of Equation (42), one has

$$\mathbb{E}[Y^2] = 2\rho \frac{\mu_2}{\mu^2} + 2(1 - \rho). \tag{50}$$

By means of Equations (45), (23) and (25) we obtain

$$\begin{aligned} \mathbb{E}[W_p^2] &= \left(\frac{p}{\mu(1-p)}\right)^2 \rho \mathbb{E}[N_p^2] \mathbb{E}[X_1^2] + (1 - \rho) \mathbb{E}[\tilde{W}_p^2] \\ &= \left(\frac{p}{\mu(1-p)}\right)^2 \rho \frac{(1-p)(2-p)}{p^2} \mu_2 + (1 - \rho) \left(\frac{p}{\mu^2(1-p)} \mu_2 + 2\right) = \\ &= \frac{\mu_2}{\mu^2} \left(\rho \frac{2-p}{1-p} + (1 - \rho) \frac{p}{1-p}\right) + 2(1 - \rho). \end{aligned} \tag{51}$$

Equations (50) and (51) permit to find  $\mathbb{E}[W_p^2] - \mathbb{E}[Y^2]$ . Hence Equation (49) leads to the inequality

$$\begin{aligned} \sup_{\mathcal{H}_2} |\mathbb{E}[h(W_p)] - \mathbb{E}[h(Y)]| \\ \geq \left(\frac{1}{2}\right) \frac{\mu_2}{\mu^2} \left( \rho \left( \frac{2-p}{1-p} - 2 \right) + (1-\rho) \frac{p}{1-p} \right) = \left(\frac{1}{2}\right) \frac{\mu_2}{\mu^2} \frac{p}{1-p}. \end{aligned} \tag{52}$$

Now, let  $\mu_2 = \infty$ . Then,  $d_{\mathcal{H}_2}(W_p, Y) = \infty$  according to Equation (52). The proof is complete.  $\square$

### 5. Convergence of Random Sums of Independent Summands to Generalized Gamma Distribution

Statements concerning weak convergence of geometric sums distributions to exponential law are often just particular cases of more general results concerning the convergence of random sums of random summands to generalized gamma law when the number of summands follows the generalized negative binomial distribution, see, e.g., [27,29,49]). The recent work [29] demonstrated how it is possible to study the mentioned general case employing the estimates of proximity of geometric sums distributions to exponential law. We introduce some notation to apply Theorem 1 for analysis of the distance between the distributions of random sums and the generalized gamma law.

Introduce a random variable  $G_{r,\lambda}$  such that  $G_{r,\lambda} \sim G(r, \lambda)$ , where  $G(r, \lambda)$  is the gamma law with positive parameters  $r$  and  $\lambda$ , i.e., its density with respect to the Lebesgue measure has the form

$$g(z; r, \lambda) = \frac{\lambda^r z^{r-1}}{\Gamma(r)} e^{-\lambda z} \mathbb{I}_{(0,\infty)}(z), \quad z \in \mathbb{R},$$

$\Gamma(r)$  being the gamma function. For  $r = 1$ , one has  $G(1, \lambda) = Exp(\lambda)$ . Clearly, for  $a > 0$ ,  $aG_{r,\lambda} \sim G(r, \lambda/a)$ . Set  $G_{r,\alpha,\lambda}^* := G_{r,\lambda}^{1/\alpha}$ , where  $\alpha > 0$ . One says that random variable  $G_{r,\alpha,\lambda}^*$  has the generalized gamma distribution  $G^*(r, \alpha, \lambda)$ . According to Equation (5) of [29], the density of  $G_{r,\alpha,\lambda}^*$  is given by formula

$$g^*(z; r, \alpha, \lambda) = \frac{|\alpha| \lambda^r z^{\alpha r - 1}}{\Gamma(r)} e^{-\lambda z^\alpha} \mathbb{I}_{(0,\infty)}(z), \quad z \in \mathbb{R}.$$

Also it is known (see Equation (6) in [29]) that, for  $r \in (0, 1)$ ,  $\alpha \in (0, 1]$  and  $\lambda > 0$ , the following relation holds

$$g^*(z; r, \alpha, \lambda) = \int_0^1 \frac{u}{1-u} e^{-\frac{u}{1-u} z} q(u; r, \alpha, \lambda) du, \quad z > 0, \tag{53}$$

where  $q$  is a density of a specified random variable  $Y_{r,\alpha,\lambda}$  such that support of its distribution belongs to  $(0, 1)$  (see Remark 3 [49]). We only note that for  $\alpha = 1$  the density  $q$  admits a representation

$$q\left(u; r, 1, \frac{b}{1-b}\right) = b^r \left(\frac{\sin \pi r}{\pi}\right) \frac{(1-u)^{r-1}}{u(u-b)^r} \mathbb{I}_{(b,1)}(u), \quad b \in (0, 1).$$

Consider a random variable  $N_{r,\alpha,p}^*$  having the generalized negative binomial distribution  $GNB(r, \alpha, p)$ , where  $r > 0$ ,  $\alpha \neq 0$  and  $p \in (0, 1)$ , i.e.,

$$\mathbb{P}(N_{r,\alpha,p}^* = k) = \int_0^\infty \frac{z^k}{k!} e^{-z} g^*\left(z; r, \alpha, \frac{p}{1-p}\right) dz, \quad k = 0, 1, \dots \tag{54}$$

Thus  $GNB(r, \alpha, p)$  has a mixed Poisson distribution. One can verify that  $GNB(r, 1, p)$  coincides with  $NB(r, p)$ , where  $NB(r, p)$  is the negative binomial law. Recall that  $N_{r,p} \sim NB(r, p)$  if

$$\mathbb{P}(N_{r,p} = k) = \frac{\Gamma(k+r)}{k!\Gamma(r)} p^r (1-p)^k, \quad k = 0, 1, \dots$$

Note also that  $N_{1,p} \sim \text{Geom}(p)$ .

Introduce the random variables

$$W_{r,\alpha,p}^* := \frac{1}{\mu} \left( \frac{p}{1-p} \right)^{1/\alpha} \sum_{k=1}^{N_{r,\alpha,p}^*} X_k, \quad S_{r,\alpha,p}^* := \sum_{k=1}^{N_{r,\alpha,p}^*} X_k, \tag{55}$$

where  $N_{r,\alpha,p}^* \sim \text{GNB}(r, \alpha, p)$ ,  $r > 0, \alpha \neq 0, p \in (0, 1)$ , and  $\mathbb{E}[X_k] = \mu, \mu \neq 0, k \in \mathbb{N}$ . We assume that  $(X_n)_{n \in \mathbb{N}}$  and  $N_{r,\alpha,p}^*$  are independent, where  $r > 0, \alpha \neq 0, p \in (0, 1)$ .

**Theorem 4.** Let  $(X_n)_{n \in \mathbb{N}}$  be a sequence of independent random variables having  $\mathbb{E}[X_n] = \mu, \mu \neq 0, n \in \mathbb{N}$ . Then, for  $W_{r,\alpha,p}^*$  introduced in Equation (55) with parameters  $r \in (0, 1), \alpha \in (0, 1], p \in (0, 1)$  and  $G_{r,1}$  having the gamma distribution  $G(r, 1)$ , the following relation holds

$$d_{\mathcal{H}_2}(W_{r,\alpha,p}^*, G_{r,1}^{1/\alpha}) = \frac{1}{2\mu^2} \left( \frac{p}{1-p} \right)^{2/\alpha} \int_0^1 \mathbb{E}[X_{N_u+1}^2] \left( \frac{1-u}{u} \right) q \left( u; r, \alpha, \frac{p}{1-p} \right) du, \tag{56}$$

whenever the right-hand side of Equation (56) is finite. Here,  $N_u := N_{1,1,u}^*, N_u \sim \text{Geom}(u), u \in (0, 1)$  and  $q$  appeared in Equation (53).

**Proof.** Without loss of generality, we can assume that  $\mu = 1$ ; otherwise, we consider  $\tilde{X}_n := \frac{X_n}{\mu}, n \in \mathbb{N}$ . For such sequence,  $\mathbb{E}[\tilde{X}_{N_u+1}^2] = \frac{1}{\mu^2} \mathbb{E}[X_{N_u+1}^2]$ . Note that  $\frac{1-p}{p} G_{r,1}$  has the same distribution as  $G_{r,p/(1-p)}$ . Applying the homogeneity property of the ideal probability metric of order two, one has

$$d_{\mathcal{H}_2}(W_{r,\alpha,p}^*, G_{r,1}^{1/\alpha}) = \left( \frac{p}{1-p} \right)^{2/\alpha} d_{\mathcal{H}_2}(S_{r,\alpha,p}^*, G_{r,p/(1-p)}^{1/\alpha}).$$

The proof of Theorem 1 [29] starts with establishing for any bounded Borel function  $h, r \in (0, 1), \alpha \in (0, 1]$  and  $p \in (0, 1)$ , that

$$\mathbb{E}[h(G_{r,p/(1-p)}^{1/\alpha})] = \int_0^1 \mathbb{E} \left[ h \left( \frac{1-u}{u} Z \right) \right] q \left( u; r, \alpha, \frac{p}{1-p} \right) du, \tag{57}$$

where  $Z \sim \text{Exp}(1)$ , and

$$\mathbb{E}[h(S_{r,\alpha,p}^*)] = \int_0^1 \mathbb{E}[h(S_{1,1,u}^*)] q \left( u; r, \alpha, \frac{p}{1-p} \right) du. \tag{58}$$

Let us examine these relations for each  $h \in \mathcal{H}_2$ . Recall that in light of Remark 1  $|h(x)| \leq A_0 x^2 + B_0$  for some positive constants  $A_0$  and  $B_0$  (which depend on  $h$ ), we write  $h = h^+ - h^-$ , where  $h^+(x) := h(x)\mathbb{I}\{h(x) \geq 0\}, h^-(x) := -h(x)\mathbb{I}\{h(x) \leq 0\}$ . Set  $h_n(x) := h^+(x)\mathbb{I}_{(-n,n]}(x), n \in \mathbb{N}$ . Then,  $h_n$  and  $n \in \mathbb{N}$  are bounded Borel functions such that for each  $x \in \mathbb{R}, 0 \leq h_n(x) \nearrow h^+(x)$  as  $n \rightarrow \infty$ . Hence, the monotone convergence theorem yields

$$\mathbb{E}[h^+(G_{r,p/(1-p)}^{1/\alpha})] = \lim_{n \rightarrow \infty} \mathbb{E}[h_n(G_{r,p/(1-p)}^{1/\alpha})].$$

Note that, for each  $u \in (0, 1), \mathbb{E} \left[ h_n \left( \frac{1-u}{u} Z \right) \right] \nearrow \mathbb{E} \left[ h^+ \left( \frac{1-u}{u} Z \right) \right]$ . Applying the monotone convergence theorem once again, we obtain

$$\int_0^1 \mathbb{E} \left[ h^+ \left( \frac{1-u}{u} Z \right) \right] q \left( u; r, \alpha, \frac{p}{1-p} \right) du = \lim_{n \rightarrow \infty} \int_0^1 \mathbb{E} \left[ h_n \left( \frac{1-u}{u} Z \right) \right] q \left( u; r, \alpha, \frac{p}{1-p} \right) du.$$

So, Equation (57) is valid if instead of  $h$  belonging to  $\mathcal{H}_2$  we write  $h^+$ . Obviously,  $0 \leq h^+(x) \leq |h(x)| \leq A_0x^2 + B_0, x \in \mathbb{R}, n \in \mathbb{R}$ . Thus,

$$\mathbb{E}\left[h^+\left(G_{r,p/(1-p)}^{1/\alpha}\right)^2\right] \leq A_0\mathbb{E}\left(G_{r,p/(1-p)}^{2/\alpha}\right) + B_0 < \infty.$$

According to [27] (page 8), for  $\delta > 0$ , one has

$$\mathbb{E}\left[(G_{r,\alpha,\lambda}^*)^\delta\right] = \frac{\Gamma(r + \frac{\delta}{\alpha})}{\lambda^{\delta/\alpha}\Gamma(r)}. \tag{59}$$

This permits us to write  $\mathbb{E}\left(G_{r,p/(1-p)}^{2/\alpha}\right) = \mathbb{E}\left[(G_{r,1,p/(1-p)}^*)^{2/\alpha}\right] < \infty$ .

In the same manner, we demonstrate that Equation (57) is valid if instead of  $h \in \mathcal{H}_2$  we take  $h^-$ . Moreover,  $\mathbb{E}\left[h^-\left(G_{r,p/(1-p)}^{1/\alpha}\right)\right]$  is finite. Therefore, Equation (57) holds for any  $h \in \mathcal{H}_2$ , and for such  $h, \mathbb{E}\left[h\left(G_{r,p/(1-p)}^{1/\alpha}\right)\right]$  is finite.

By the monotone convergence theorem  $\mathbb{E}\left[h^+(S_{r,\alpha,p}^*)\right] = \lim_{n \rightarrow \infty} \mathbb{E}\left[h_n(S_{r,\alpha,p}^*)\right]$ . In a similar way,  $\mathbb{E}\left[h_n(S_{1,1,u}^*)\right] \nearrow \mathbb{E}\left[h^+(S_{1,1,u}^*)\right]$  as  $n \rightarrow \infty$ , and applying this theorem once again, we obtain

$$\int_0^1 \mathbb{E}\left[h^+(S_{1,1,u}^*)\right]q\left(u; r, \alpha, \frac{p}{1-p}\right) du = \lim_{n \rightarrow \infty} \int_0^1 \mathbb{E}\left[h_n(S_{1,1,u}^*)\right]q\left(u; r, \alpha, \frac{p}{1-p}\right) du.$$

Taking into account that Equation (58) is valid for bounded Borel functions  $h_n$ , one ascertains that Equation (58) holds if we replace  $h$  by  $h^+$ . To show the latter integral is finite, we note that  $0 \leq h^+(x) \leq |h(x)| \leq A_0x^2 + B_0$ , for some positive  $A_0, B_0$  and all  $x \in \mathbb{R}$ . Formula (23) of Lemma 4 yields, for each  $u \in (0, 1)$ ,

$$\mathbb{E}\left[(S_{1,1,u}^*)^2\right] \leq \frac{1-u}{u} \mathbb{E}\left[X_{N_u+1}^2\right] + 2\frac{(1-u)^2}{u^2}.$$

It was assumed above that the right-hand side of Equation (56) is finite. So,

$$\int_0^1 \mathbb{E}\left(A_0\left(\frac{1-u}{u} \mathbb{E}\left[X_{N_u+1}^2\right] + 2\frac{(1-u)^2}{u^2}\right) + B_0\right)q\left(u; r, \alpha, \frac{p}{1-p}\right) du < \infty,$$

since in light of Equation (57), taking  $h(x) = 1$  and  $h(x) = \frac{x^2}{2}$  (these functions belong to  $\mathcal{H}_2$ ),  $x \in \mathbb{R}$ , we obtain, respectively,

$$\int_0^1 q\left(u; r, \alpha, \frac{p}{1-p}\right) du = 1, \tag{60}$$

$$\mathbb{E}\left[Z^2\right] \int_0^1 \frac{(1-u)^2}{u^2} q\left(u; r, \alpha, \frac{p}{1-p}\right) du = \mathbb{E}\left(G_{r,p/(1-p)}^{2/\alpha}\right) < \infty.$$

We demonstrate analogously that Equation (58) holds upon replacing  $h \in \mathcal{H}_2$  with  $h^-$  and if the right-hand side of Equation (56) is finite, it follows that

$$\int_0^1 \mathbb{E}\left[h^-(S_{1,1,u}^*)\right]q\left(u; r, \alpha, \frac{p}{1-p}\right) du$$

is finite as well. Consequently, Equation (58) is established for each  $h \in \mathcal{H}_2$  (whenever the right-hand side of Equation (56) is finite) and  $\mathbb{E}\left[h(S_{r,\alpha,p}^*)\right]$  is finite for such  $h$ . Therefore, for  $h \in \mathcal{H}_2$  and fixed  $\alpha, r, p$ , one has

$$\begin{aligned} & \mathbb{E}[h(S_{r,\alpha,p}^*)] - \mathbb{E}[h(G_{r,p/(1-p)}^{1/\alpha})] \\ &= \int_0^1 \left( \mathbb{E}[h(S_{1,1,u}^*)] - \mathbb{E}\left[h\left(\frac{1-u}{u}Z\right)\right] \right) q\left(u; r, \alpha, \frac{p}{1-p}\right) du =: J(h). \end{aligned}$$

By Theorem 1, for  $h \in \mathcal{H}_2$ , it holds

$$\begin{aligned} \left| \mathbb{E}[h(S_{1,1,u}^*)] - \mathbb{E}\left[h\left(\frac{1-u}{u}Z\right)\right] \right| &\leq d_{\mathcal{H}_2}\left(S_{1,1,u}^*, \frac{1-u}{u}Z\right) = \left(\frac{1-u}{u}\right)^2 d_{\mathcal{H}_2}\left(\frac{u}{1-u}S_{1,1,u}^*, Z\right) \\ &\leq \left(\frac{1-u}{u}\right)^2 \frac{u}{1-u} \left(\frac{1}{2}\right) \mathbb{E}[X_{N_u+1}^2] = \left(\frac{1}{2}\right) \frac{1-u}{u} \mathbb{E}[X_{N_u+1}^2], \end{aligned}$$

where we take into account that  $N_{1,u}^* \sim NB(1, u)$ , and  $NB(1, u)$  coincides with  $Geom(u)$ . Thus,  $\frac{u}{1-u}S_{1,1,u}^*$  can be written as

$$\frac{u}{1-u} \sum_{k=1}^{N_u} X_k,$$

where  $N_u \sim Geom(u)$ ,  $N_u$  and  $(X_k)_{k \in \mathbb{N}}$  are independent.

Therefore, for each  $h \in \mathcal{H}_2$ ,  $\left(\frac{p}{1-p}\right)^{2/\alpha} |J(h)|$  is bounded by the right-hand side of Equation (56), and so the desired upper bound is obtained (recall that  $\mu = 1$ ).

Now, we turn to the lower bound of  $d_{\mathcal{H}_2}(W_{r,\alpha,p}^*, G_{r,1}^{1/\alpha})$ . Take  $h(x) = x^2/2$  belonging to  $\mathcal{H}_2$ . Then, applying Equation (23) to evaluate  $\mathbb{E}[(S_{1,1,u}^*)^2]$ , one has

$$\begin{aligned} & d_{\mathcal{H}_2}(W_{r,\alpha,p}^*, G_{r,1}^{1/\alpha}) \\ & \geq \frac{1}{2} \left(\frac{p}{1-p}\right)^{2/\alpha} \left| \int_0^1 \left( \mathbb{E}[(S_{1,1,u}^*)^2] - \left(\frac{1-u}{u}\right)^2 \mathbb{E}[G_{1,1}^2] \right) q\left(u; r, \frac{p}{1-p}\right) du \right| \\ & = \frac{1}{2} \left(\frac{p}{1-p}\right)^{2/\alpha} \int_0^1 \left(\frac{1-u}{u}\right)^2 \mathbb{E}[X_{N_u+1}^2] q\left(u; r, \frac{p}{1-p}\right) du, \quad (61) \end{aligned}$$

where  $G_{1,1} = Z \sim Exp(1)$ . Thus, Equation (61) completes the proof.  $\square$

**Corollary 1.** Let conditions of Theorem 4 be satisfied and also  $\mu_2 = \sup_{n \in \mathbb{N}} \mathbb{E}[X_n^2] < \infty$ . Then, the right-hand side of Equation (56) is finite and

$$d_{\mathcal{H}_2}(W_{r,\alpha,p}^*, G_{r,1}^{1/\alpha}) \leq \frac{\mu_2}{2\mu^2} \left(\frac{p}{1-p}\right)^{1/\alpha} \frac{\Gamma(r + \frac{1}{\alpha})}{\Gamma(r)}.$$

The inequality becomes an equality if  $\mu_2 = \mathbb{E}[X_n^2]$  for all  $n \in \mathbb{N}$ . In particular, if  $\alpha = 1$  then  $\frac{\Gamma(r+1)}{\Gamma(r)} = r$ .

**Proof.** According to Equation (57), for  $h(x) = x$ ,  $x \in \mathbb{R}$ ,

$$\mathbb{E}[G_{r,p/(1-p)}^{1/\alpha}] = \mathbb{E}[Z] \int_0^1 \left(\frac{1-u}{u}\right) q\left(u; r, \alpha, \frac{p}{1-p}\right) du.$$

Thus, the following relation is valid.

$$\int_0^1 \left(\frac{1-u}{u}\right) q\left(u; r, \alpha, \frac{p}{1-p}\right) du = \mathbb{E}[G_{r,p/(1-p)}^{1/\alpha}]. \quad (62)$$

Due to [27] (see page 8 there), for  $\delta > 0$ , one has  $\mathbb{E}[G_{r,\alpha,\lambda}^*] = \frac{\Gamma(r+1/\alpha)}{\lambda^{1/\alpha}\Gamma(r)}$ . Therefore,

$$\mathbb{E}[G_{r,p/(1-p)}^{1/\alpha}] = \mathbb{E}[G_{r,\alpha,p/(1-p)}^*] = \left(\frac{1-p}{p}\right)^{\frac{1}{\alpha}} \frac{\Gamma(r + \frac{1}{\alpha})}{\Gamma(r)}.$$

For  $\alpha = 1$ , we obtain  $\mathbb{E}[G_{r,p/(1-p)}] = \frac{1-p}{p} \frac{\Gamma(r+1)}{\Gamma(r)} = r \frac{(1-p)}{p}$ .  $\square$

### 6. Convergence of Random Sums of Exchangeable Summands to Generalized Gamma Distribution

Consider the model of exchangeable random variables  $X_1, X_2, \dots$  described in Section 4. Introduce the distribution of a random variable  $U_{r,\alpha,\lambda}^*$  as the following mixture

$$\mathbb{P}_{U_{r,\alpha,\lambda}^*} = \rho \mathbb{P}_{\left(\frac{V_{r,\alpha,\lambda}^* X_1}{\mu}\right)} + (1-\rho) \mathbb{P}_{Z_{r,\alpha,\lambda}^*}, \tag{63}$$

where  $\rho \in [0, 1]$ ,  $\alpha > 0$ ,  $r > 0$ ,  $\mu := \mathbb{E}[X_1]$ ,  $\mu \neq 0$ , random variables  $X_1, V_{r,\alpha,\lambda}^*$  are independent,  $V_{r,\alpha,\lambda}^* \sim G^*(r, \alpha, \lambda)$ ,  $Z_{r,\alpha,\lambda}^* \sim G^*(r, \alpha, \lambda)$ . Since  $\mathbb{E}[G_{r,\lambda}^{2/\alpha}] = \frac{\Gamma(r+2/\alpha)}{\lambda^{2/\alpha}\Gamma(r)}$  (see, e.g., page 8 [27]), one has

$$\mathbb{E}[(U_{r,\alpha,\lambda}^*)^2] = \left(\rho \frac{\mathbb{E}[X_1^2]}{\mu^2} + (1-\rho)\right) \frac{\Gamma(r + 2/\alpha)}{\lambda^{2/\alpha}\Gamma(r)}. \tag{64}$$

Due to the properties of generalized gamma distributions, for any positive number  $c$ ,

$$\begin{aligned} \frac{1}{c^\alpha} U_{r,\alpha,\lambda}^* &= \frac{1}{c^\alpha} \left( (1 - \mathbb{I}_\rho) \frac{V_{r,\alpha,\lambda}^* X_1}{\mu} + \mathbb{I}_\rho Z_{r,\alpha,\lambda}^* \right) \\ &= \left( (1 - \mathbb{I}_\rho) \frac{V_{r,\alpha,\lambda}^* X_1}{\mu} + \mathbb{I}_\rho Z_{r,\alpha,c\lambda}^* \right) = U_{r,\alpha,c\lambda}^*, \end{aligned} \tag{65}$$

where indicator  $\mathbb{I}_\rho$  equals 1 and 0 with probabilities  $1 - \rho$  and  $\rho$ , respectively, and is independent with all the variables under consideration. Note that  $U_{1,1,1}^*$  has the same distribution as a random variable  $Y$ , having the law defined in Equation (42). Recall that the generalized negative binomial distribution  $GNB(r, \alpha, p)$  is the law of a random variable  $N_{r,\alpha,p}^*$ , see Equation (54). We will use the following result.

**Lemma 5.** *If  $r > 0$ ,  $\alpha \neq 0$ ,  $p \in (0, 1)$ , then for  $N_{r,\alpha,p}^* \sim GNB(r, \alpha, p)$  one has*

$$\mathbb{E}[N_{r,\alpha,p}^*] = \mathbb{E}[G_{r,\alpha,p/(1-p)}^*], \quad \mathbb{E}[N_{r,\alpha,p}^*(N_{r,\alpha,p}^* - 1)] = \mathbb{E}[(G_{r,\alpha,p/(1-p)}^*)^2]. \tag{66}$$

**Proof.** According to Equation (54), for each  $n \in \mathbb{N}$ ,

$$\begin{aligned} \sum_{k=1}^n k \mathbb{P}(N_{r,\alpha,p}^* = k) &= \int_0^\infty z \sum_{k=1}^n \frac{z^{k-1}}{(k-1)!} e^{-z} g^*(z; r, \alpha, \frac{p}{1-p}) dz, \\ \sum_{k=2}^n k(k-1) \mathbb{P}(N_{r,\alpha,p}^* = k) &= \int_0^\infty z^2 \sum_{k=2}^n \frac{z^{k-2}}{(k-2)!} e^{-z} g^*(z; r, \alpha, \frac{p}{1-p}) dz. \end{aligned}$$

The desired statement follows from the monotone convergence theorem for the Lebesgue integral by letting  $n \rightarrow \infty$ .  $\square$

**Theorem 5.** *Let  $X_1, X_2, \dots$  be exchangeable random variables, introduced in Section 4, such that  $\mathbb{E}[X_1] = \mu$ ,  $\mathbb{E}[X_1^2] = \mu_2 < \infty$ . Assume that for some  $\rho \in (0, 1)$  Equation (41) holds. Suppose that  $(X_n)_{n \in \mathbb{N}}$  and  $N_{r,\alpha,p}^*$  are independent, where  $N_{r,\alpha,p}^* \sim GNB(r, \alpha, p)$ . Then, for  $W_{r,\alpha,p}^*$  defined in Equation (55) with parameters  $r \in (0, 1)$ ,  $\alpha \in (0, 1]$ ,  $p \in (0, 1)$  and  $U_{r,\alpha,1}^*$  given in Equation (63), one has*

$$d\mathcal{H}_2(W_{r,\alpha,p}^*, U_{r,\alpha,1}^*) = \frac{\mu_2}{2\mu^2} \left(\frac{p}{1-p}\right)^{1/\alpha} \frac{\Gamma(1+\frac{1}{\alpha})}{\Gamma(r)}. \tag{67}$$

**Proof.** Without loss of generality, we can assume that  $\mu = 1$ ; otherwise, we consider  $\tilde{X}_n := X_n/\mu, n \in \mathbb{N}$ . For such sequence,  $\tilde{\mu}_2 = \mathbb{E}\tilde{X}_1^2 = \mu_2/\mu^2$ . Note that Equation (58) is true for dependent summands (see Theorem 1 [29]). Furthermore, for bounded  $h(t), t \in \mathbb{R}$ , function  $h_x(t) = h(xt)$  is also bounded for any  $x \in \mathbb{R}$ . Thus, an employment of Equation (63) gives

$$\mathbb{E}[h(U_{r,\alpha,\lambda}^*)] = \rho \int_{\mathbb{R}} \mathbb{E}[h_x(G_{r,\lambda}^{1/\alpha})] dF_{X_1}(x) + (1-\rho)\mathbb{E}[h(G_{r,\lambda}^{1/\alpha})]. \tag{68}$$

Now we apply Equation (57) with bounded  $h_x$  and by Fubini’s theorem obtain:

$$\begin{aligned} \int_{\mathbb{R}} \mathbb{E}[h_x(G_{r,\lambda}^{1/\alpha})] dF_{X_1}(x) &= \int_{\mathbb{R}} \int_0^1 \mathbb{E}\left[h_x\left(\frac{1-u}{u}V^*\right)\right] q(u;r,\alpha,\lambda) du dF_{X_1}(x) \\ &= \int_0^1 \mathbb{E}\left[h\left(\frac{1-u}{u}X_1V^*\right)\right] q(u;r,\alpha,\lambda) du, \end{aligned} \tag{69}$$

where  $X_1$  and  $V^*$  are independent and  $V^* \sim Exp(1)$ . Apply Equation (57) for the second summand of Equation (68). Then, Equation (69) yields

$$\begin{aligned} \mathbb{E}[h(U_{r,\alpha,\lambda}^*)] &= \rho \int_0^1 \mathbb{E}\left[h\left(\frac{1-u}{u}X_1V^*\right)\right] q(u;r,\alpha,\lambda) du + (1-\rho) \int_0^1 \mathbb{E}\left[h\left(\frac{1-u}{u}Z^*\right)\right] q(u;r,\alpha,\lambda) du \\ &= \int_0^1 \mathbb{E}\left[h\left(\frac{1-u}{u}U_{1,1,1}^*\right)\right] q(u;r,\alpha,\lambda) du, \end{aligned} \tag{70}$$

where  $Z^* \sim Exp(1)$  and  $U_{1,1,1}^*$  have the same distribution as  $Y$ , see Equation (42).

Recall that, for  $h \in \mathcal{H}_2$ , an inequality  $|h(x)| \leq A_0x^2 + B_0$  holds for all  $x \in \mathbb{R}$  and some positive constants  $A_0, B_0$  (see Remark 1). Moreover,  $\mathbb{E}[(U_{r,\alpha,\lambda}^*)^2] < \infty$  according to Equation (64). So, employing bounded  $h_n(x) = h(x)\mathbb{I}_{(-n,n]}(x)$  tending to  $h(x) \in \mathcal{H}_2$  as  $n \rightarrow \infty$ , one can invoke the Lebesgue dominated convergence theorem to claim that  $\lim_{n \rightarrow \infty} \mathbb{E}[h_n(U_{r,\alpha,\lambda}^*)] = \mathbb{E}[h(U_{r,\alpha,\lambda}^*)]$ . We take into account that

$$\int_0^1 \mathbb{E}\left|h_n\left(\frac{1-u}{u}U_{1,1,1}^*\right)\right| q(u;r,\alpha,\lambda) du \leq A_0\mathbb{E}[(U_{1,1,1}^*)^2] \int_0^1 \left(\frac{1-u}{u}\right)^2 q(u;r,\alpha,\lambda) du + B_0.$$

The integral in the right-hand side of the latter formula is finite by Equation (60) and  $\mathbb{E}[(U_{1,1,1}^*)^2] < \infty$  in accord with Equation (64). Thus, it is possible to apply the Lebesgue dominated convergence theorem to obtain

$$\lim_{n \rightarrow \infty} \int_0^1 \mathbb{E}\left[h_n\left(\frac{1-u}{u}U_{1,1,1}^*\right)\right] q(u;r,\alpha,\lambda) du = \int_0^1 \mathbb{E}\left[h\left(\frac{1-u}{u}U_{1,1,1}^*\right)\right] q(u;r,\alpha,\lambda) du$$

for any  $h \in \mathcal{H}_2$ . So, Equation (70) holds for all  $h \in \mathcal{H}_2$ .

In a similar way,  $\lim_{n \rightarrow \infty} \mathbb{E}[h_n(S_{r,\alpha,p}^*)] = \mathbb{E}[h(S_{r,\alpha,p}^*)]$  for  $h \in \mathcal{H}_2$ . According to the Cauchy–Bunyakovsky–Schwarz inequality for identically distributed variables  $X_1, X_2, \dots$  we have  $|\mathbb{E}[X_iX_j]| \leq \mu_2$  for  $i, j \in \mathbb{N}$  and consequently

$$\begin{aligned} \mathbb{E}[(S_{r,\alpha,p}^*)^2] &= \sum_{k=0}^{\infty} \mathbb{P}(N_{r,\alpha,p}^* = k) \mathbb{E}\left[\left(\sum_{j=1}^k X_j\right)^2\right] \\ &\leq \mu_2 \sum_{k=0}^{\infty} \mathbb{P}(N_{r,\alpha,p}^* = k) k^2 = \mu_2 \mathbb{E}[(N_{r,\alpha,p}^*)^2]. \end{aligned} \tag{71}$$

Equations (59) and (66) entail that  $\mathbb{E}[(N_{r,\alpha,p}^*)^2] < \infty$ . Thus, the dominated convergence theorem guarantees that  $\lim_{n \rightarrow \infty} \mathbb{E}[h_n(S_{r,\alpha,p}^*)] = \mathbb{E}[h(S_{r,\alpha,p}^*)]$ . Furthermore, one can demonstrate that, for each  $h \in \mathcal{H}_2$ ,

$$\lim_{n \rightarrow \infty} \int_0^1 \mathbb{E}[h_n(S_{1,1,u}^*)]q(u; r, \alpha, \lambda) du = \int_0^1 \mathbb{E}[h(S_{1,1,u}^*)]q(u; r, \alpha, \lambda) du. \tag{72}$$

For this purpose we note that Equation (71) implies

$$\int_0^1 \mathbb{E}|h_n(S_{1,1,u}^*)|q(u; r, \alpha, \lambda) du \leq C + A\mu_2 \int_0^1 \mathbb{E}[(N_{1,1,u}^*)^2]q(u; r, \alpha, \lambda) du.$$

According to Equation (66) one has

$$\int_0^1 \mathbb{E}[(N_{1,1,u}^*)^2]q(u; r, \alpha, \lambda) du = \int_0^1 (\mathbb{E}[(G_{1,1,u/(1-u)}^*)^2] + \mathbb{E}[G_{1,1,u/(1-u)}^*])q(u; r, \alpha, \lambda) du.$$

The latter integral is finite because one can take  $h(x) = x$  and  $h(x) = x^2/2$  in Equation (57) and invoke Equation (59). Then, it is possible to use the dominated convergence theorem once again to establish Equation (72).

Now, combining Equation (58) and Equation (70) leads for any  $h \in \mathcal{H}_2$  to the relation

$$\begin{aligned} &\mathbb{E}[h(S_{r,\alpha,p}^*)] - \mathbb{E}[h(U_{r,\alpha,p/(1-p)}^*)] \\ &= \int_0^1 \left( \mathbb{E}[h(S_{1,1,u}^*)] - \mathbb{E}\left[h\left(\frac{1-u}{u}U_{1,1,1}^*\right)\right] \right) q\left(u; r, \alpha, \frac{p}{1-p}\right) du. \end{aligned} \tag{73}$$

Note that a random variable  $N_{1,1,u}^*$  follows the geometric distribution  $Geom(u)$  with parameter  $u \in (0, 1)$ . For each  $h \in \mathcal{H}_2$  and any  $u \in (0, 1)$ , by Theorem 3 and in view of  $d_{\mathcal{H}_2}$  homogeneity, we obtain

$$\begin{aligned} &\left| \mathbb{E}[h(S_{1,1,u}^*)] - \mathbb{E}\left[h\left(\frac{1-u}{u}U_{1,1,1}^*\right)\right] \right| \leq d_{\mathcal{H}_2}\left(S_{1,1,u}^*, \frac{1-u}{u}U_{1,1,1}^*\right) \\ &= \left(\frac{1-u}{u}\right)^2 d_{\mathcal{H}_2}(W_u, Y) \leq \left(\frac{1-u}{u}\right)^2 \left(\frac{u}{1-u}\right) \frac{\mu_2}{2} = \left(\frac{1-u}{u}\right) \frac{\mu_2}{2}. \end{aligned} \tag{74}$$

Employing Equations (73), (74) and (62) one deduces

$$d_{\mathcal{H}_2}(S_{r,\alpha,p}^*, U_{r,\alpha,p/(1-p)}^*) \leq \frac{\mu_2}{2} \int_0^1 \left(\frac{1-u}{u}\right) q\left(u; r, \alpha, \frac{p}{1-p}\right) du = \frac{\mu_2}{2} \mathbb{E}[G_{r,p/(1-p)}^{1/\alpha}]. \tag{75}$$

Equation (65) implies by virtue of  $d_{\mathcal{H}_2}$  homogeneity that

$$d_{\mathcal{H}_2}(W_{r,\alpha,p}^*, U_{r,\alpha,1}^*) = \left(\frac{p}{1-p}\right)^{2/\alpha} d_{\mathcal{H}_2}(S_{r,\alpha,p}^*, U_{r,\alpha,p/(1-p)}^*). \tag{76}$$

Combining Equations (59), (75) and (76) we conclude that the right-hand side of Equation (67) is an upper bound for  $d_{\mathcal{H}_2}(W_{r,\alpha,p}^*, U_{r,\alpha,1}^*)$ .

Choosing  $h(x) = x^2/2$  in Equation (73), upon employing Equation (52) and Equation (62) one infers:

$$\begin{aligned} &d_{\mathcal{H}_2}(W_{r,\alpha,p}^*, G_{r,1}^{1/\alpha}) \geq \\ &\geq \frac{1}{2} \left(\frac{p}{1-p}\right)^{2/\alpha} \left| \int_0^1 \left( \mathbb{E}[(S_{1,1,u}^*)^2] - \left(\frac{1-u}{u}\right)^2 \mathbb{E}[(U_{1,1,1}^*)^2] \right) q\left(u; r, \alpha, \frac{p}{1-p}\right) du \right| = \\ &= \frac{\mu_2}{2} \left(\frac{p}{1-p}\right)^{2/\alpha} \int_0^1 \left(\frac{1-u}{u}\right) q\left(u; r, \alpha, \frac{p}{1-p}\right) du = \frac{\mu_2}{2} \left(\frac{p}{1-p}\right)^{2/\alpha} \mathbb{E}[G_{r,\alpha,p/(1-p)}^*]. \end{aligned}$$

Using Equation (59) once again, we see that the right-hand side of Equation (67) is a lower bound for  $d_{\mathcal{H}_2}(W_{r,\alpha,p}^*, U_{r,\alpha,1}^*)$ .  $\square$

### 7. Inverse to Equilibrium Transformation

The development of Stein’s method is closely connected with various transformations of distributions. Let a random variable  $W \geq 0$  and  $0 < \mu = \mathbb{E}[W] < \infty$ . Then, one says that a random variable  $W^s$  has the  $W$ -size biased distribution if for all  $f$  such that  $\mathbb{E}[Wf(W)]$  exists

$$\mathbb{E}[Wf(W)] = \mu\mathbb{E}[f(W^s)].$$

The connection of this transformation with Stein’s equation was considered in [50,51]. It was pointed out in [51] that this transformation works well for combinatorial problems, such as counting the number of vertices in a random graph having prespecified degrees, see also [52]. In [53], another transformation was introduced. Namely, if a random variable  $W$  has mean zero and variance  $\sigma^2 \in (0, \infty)$ , then the authors of [53] write (Definition 1.1) that a variable  $W^*$  has  $W$ -zero biased distribution whenever, for all differentiable  $f$  such that  $\mathbb{E}Wf(W)$  exists, the following relation holds

$$\mathbb{E}[Wf(W)] = \sigma^2\mathbb{E}[f'(W^*)].$$

This definition is inspired by an equation  $\mathbb{E}[Wf(W)] = \sigma^2\mathbb{E}[f'(W)]$  characterizing the normal law  $N(0, \sigma^2)$ . The authors of [53] explain that  $W^*$  always exists if  $\mathbb{E}[W] = 0$  and  $\text{var}W \in (0, \infty)$ . Zero-based coupling for products of normal random variables is treated in [54]. In Sec. 2 of [30], it is demonstrated that the gamma distribution is uniquely characterised by the property that its size-biased distribution is the same as its zero-biased distribution. Two generalizations of zero biasing were proposed in [55], see p. 104 of that paper for discussion of these transformations. We refer also to survey [56].

Now, we turn to the equilibrium distribution transformation introduced in [33] and concentrate on approximation of the law under consideration by means of an exponential law, see the corresponding Definition 1 in Section 2.

According to the second part of Theorem 2.1 of [33] (in our notation), for  $Z \sim \text{Exp}(1)$  and non-negative random variable  $X$  with  $\mathbb{E}[X] = 1$  and  $\mathbb{E}[X^2] < \infty$  the following estimate holds

$$d_{\mathcal{H}_1}(X, Z) \leq 2\mathbb{E}|X^e - X|,$$

and at the same time

$$d_{\mathcal{H}_1}(X^e, Z) \leq \mathbb{E}|X^e - X|. \tag{77}$$

The authors of [33] also proved that  $d_{\mathcal{K}}(X^e, Z) \leq \mathbb{E}|X^e - X|$ . Notice that the estimate for  $d_{\mathcal{H}_1}(X^e, Z)$  is more precise than that for  $d_{\mathcal{H}_1}(X, Z)$ .

Now we turn to Equation (77) and demonstrate how to find the distribution of  $X$  when we know the distribution of  $X^e$ . In other words, we concentrate on the inverse of an equilibrium distribution transformation.

Assume that  $\mathbb{E}[X] > 0$ . Recall that a random variable  $X^e$  exists if  $F^e(x)$  appearing in Equation (16) is a distribution function. The latter statement for  $\mathbb{E}[X] > 0$  is equivalent to nonnegativity of  $X$ . Indeed, for non-negative  $X$ ,  $F^e(x)$  coincides with a distribution function having a density (15). If  $F^e(x)$  is a distribution function and  $\mathbb{E}[X] > 0$  in Equation (16), then  $F^e(x) \geq 0$  for  $x < 0$  only if  $F(x) = 0$  for  $x < 0$ .

Thus a random variable  $X^e$  has a (version of) density  $p^e(x)$  introduced in Equation (15). Obviously, the function  $p^e(x)$  has the following properties. It is nonincreasing on  $[0, \infty)$  and  $p^e(x) = 0$  for  $x < 0$ . This density is right-continuous on  $[0, \infty)$  and consequently  $p^e(0) < \infty$ . Now, we are able to provide a full description of the class of densities for random variables  $X^e$  relevant to all non-negative  $X$  with positive mean.

**Lemma 6.** *Let a non-negative random variable  $X^e$  have a version of density (with respect to the Lebesgue measure)  $p^e(x)$ ,  $x \in \mathbb{R}$ , such that this function is nonincreasing on  $[0, \infty)$ ,  $p^e(x) = 0$  for*

$x < 0$ , and there is finite  $\lim_{x \rightarrow 0^+} p^e(x)$ . Then, there exists a unique preimage of  $X^e$  distribution having the distribution function  $F$  continuous at  $x = 0$ . Namely,

$$F(x) = \begin{cases} 1 - \frac{p^e(x)}{p^e(0)}, & x \geq 0, \\ 0, & x < 0. \end{cases} \tag{78}$$

**Proof.** First of all, note that  $p^e(0) > 0$  as otherwise  $p^e(x) = 0$  for all  $x \in \mathbb{R}$  ( $p^e$  is a nonincreasing function on  $[0, \infty)$ ). We also know that there exist a left-sided limit and a right-sided limit of  $p^e$  at each point  $x \in (0, \infty)$  as well as the right-sided limit of  $p^e$  at  $x = 0$ . The set of discontinuity points of  $p^e$  is at most countable, and we can take a version which is right continuous at each point of  $[0, \infty)$ . Then, Equation (78) introduces a distribution function. Consider a random variable  $X$  with distribution function  $F$  and check the validity of Equation (14).

The integration by a parts formula yields, for any  $b > 0$ ,

$$1 \geq \int_0^b p^e(x) dx = bp^e(b) + p^e(0) \int_0^b x dF(x). \tag{79}$$

Summands in the right-hand side of Equation (79) are non-negative. Therefore, for any  $b > 0$ ,  $\mathbb{E}[X\mathbb{I}(X \leq b)] \leq 1/p^e(0)$ . Hence, the monotone convergence theorem implies that  $\mathbb{E}[X]$  is finite. According to Equation (78)

$$bp^e(b)/p^e(0) = b(1 - F(b)) = b\mathbb{P}(X > b) \rightarrow 0, \quad b \rightarrow \infty, \tag{80}$$

since  $\mathbb{E}[X] < \infty$ . Taking in the Equation (79) limit as  $b \rightarrow \infty$ , one obtains  $1 = p^e(0)\mathbb{E}[X]$ . Now, we are ready to verify Equation (14). For any Lipschitz function  $f$ ,  $\mathbb{E}[f(X)]$  is finite and

$$\mathbb{E}[f(X)] = \int_0^\infty f(x)dF(x) = -\frac{1}{p^e(0)} \int_0^\infty f(x)dp^e(x).$$

Taking into account Equation (80), we infer that  $f(b)p^e(b) \rightarrow 0$  as  $b \rightarrow \infty$ . Consequently, applying integration by parts once again ( $f$  has bounded variation), we obtain

$$\begin{aligned} \mathbb{E}[X]\mathbb{E}[f'(X^e)] &= \frac{1}{p^e(0)} \int_0^\infty f'(x)p^e(x) dx = \frac{1}{p^e(0)} \int_0^\infty p^e(x)df(x) \\ &= \frac{1}{p^e(0)} \left[ -f(0)p^e(0) - \int_0^\infty f(x)dp^e(x) \right] = \mathbb{E}[f(X)] - f(0). \end{aligned}$$

Uniqueness of  $X$  distribution corresponding to  $X^e$  is a consequence of Equation (15) and continuity of  $F(x)$  at  $x = 0$ . Indeed, assume that for  $X_1$  and  $X_2$  one has  $X_1^e = X_2^e$ . Then, Equation (15) yields that for almost all  $x \geq 0$ ,

$$\frac{1}{\mathbb{E}[X_1]} \mathbb{P}(X_1 > x) = \frac{1}{\mathbb{E}[X_2]} \mathbb{P}(X_2 > x), \tag{81}$$

and therefore  $\mathbb{P}(X_1 > x) = c\mathbb{P}(X_2 > x)$ , where  $c$  is a positive constant (the equilibrium distribution in Definition 1 is introduced for random variables with positive expectation only). Since  $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_2 = 0) = 0$ , one has  $\mathbb{P}(X_1 > 0) = \mathbb{P}(X_2 > 0)$ . Let  $x_n \rightarrow 0^+$ ,  $n \rightarrow \infty$ , where the points  $x_n$  belong to the set considered in Equation (81) to ensure that  $c = 1$ . Thus, distributions of  $X_1$  and  $X_2$  coincide.  $\square$

**Remark 6.** Let  $X_p$  be the Bernoulli random variable taking values 1 and 0 with probabilities  $p$  and  $1 - p$ , respectively. Then, it is easily seen that the distribution of  $X_p^e$  is uniform on  $[0, 1]$ . Thus, in contrast to Lemma 6, without assumption of continuity of  $F$  at a point  $x = 0$  one can not guarantee, in general, the preimage uniqueness for the inverse transformation to the equilibrium one.

In the proof of Lemma 6, we find out that  $\mathbb{E}[X] = 1/p^e(0)$ . Set  $\lambda = p^e(0)$ ,  $Z \sim \text{Exp}(\lambda)$ . Then,  $\mathbb{E}[X] = \mathbb{E}[Z]$ . Further, we suppose that this choice of  $\lambda$  is made.

Recall that random variables  $U$  and  $V$  are stochastically ordered if either  $\mathbb{P}(U \leq x) \leq \mathbb{P}(V \leq x)$ , for every  $x \in \mathbb{R}$ , or the opposite inequality holds (for all  $x \in \mathbb{R}$ ). Now, we clarify one of the Theorem 2.1 of [33] statements (see also Theorem 3 [22], where the result similar to Theorem 2.1 of [33] is formulated employing the generalized distributions).

**Theorem 6.** *Let a random variable  $X^e$  satisfy conditions of Lemma 6, and  $\mathbb{E}[X^e] < \infty$  and  $X$  be a preimage of the equilibrium transformation. Then, Equation (77) holds. Moreover, the inequality becomes an equality when  $X$  and  $X^e$  are stochastically ordered.*

**Proof.** Apply the Stein Equation (10) along with equilibrium transformation (14). Then, in light of  $\mathbb{E}[X] = \frac{1}{\lambda}$  and  $\mathbb{E}f_h(X) - f_h(0) = \frac{1}{\lambda}\mathbb{E}f'_h(X^e)$ , we can write

$$\begin{aligned} |\mathbb{E}[h(X^e)] - \mathbb{E}[h(Z)]| &= |\mathbb{E}(f'_h(X^e) - \lambda f_h(X^e)) + \lambda f(0)| \\ &= \lambda |\mathbb{E}(f_h(X^e) - f_h(X))| \leq \lambda \|f'_h\|_\infty \mathbb{E}|X^e - X| \leq \|h'\|_\infty \mathbb{E}|X^e - X|. \end{aligned} \tag{82}$$

The last inequality in (82) is true due to Lemma 2. Now, we demonstrate that equality in (82) can be attained. Taking  $h(x) = x - \frac{1}{\lambda}$ , we have a solution  $f_h(x) = -\frac{1}{\lambda}x$  of Equation (12). Then,

$$|\mathbb{E}[h(X^e)] - \mathbb{E}[h(Z)]| = \lambda |\mathbb{E}(f_h(X^e) - f_h(X))| = |\mathbb{E}(X^e - X)|.$$

Employing the integration by parts formula, one can show that the expression in the right-hand side of the last equality is equal to the Kantorovich distance between  $X$  and  $X^e$  when these variables are stochastically ordered. Note that  $x(1 - F(x)) \rightarrow 0$ ,  $x(1 - F^e(x)) \rightarrow 0$  as  $x \rightarrow \infty$  and  $xF(x) \rightarrow 0$ ,  $xF^e(x) \rightarrow 0$  as  $x \rightarrow -\infty$  because  $\mathbb{E}[X]$  and  $\mathbb{E}[X^e]$  are finite. Thus,

$$\begin{aligned} |\mathbb{E}[X^e] - \mathbb{E}[X]| &= \left| \int_{\mathbb{R}} x(dF_{X^e}(x) - dF_X(x)) \right| \\ &= \left| - \int_{\mathbb{R}} (F_{X^e}(x) - F_X(x)) dx \right| = \int_{\mathbb{R}} |F_{X^e}(x) - F_X(x)| dx, \end{aligned}$$

since  $F_{X^e}(x) \geq F_X(x)$  (or  $\leq$ ) for all  $x \in \mathbb{R}$ . It is well-known that the Kantorovich distance is the minimal one for the metric  $\tau(U, V) = \mathbb{E}|U - V|$  (see, e.g., [9], Ch. 1, §1.3). Therefore,

$$\int_{\mathbb{R}} |F_{X^e}(x) - F_X(x)| dx = \inf \mathbb{E}|U - V|,$$

where the infimum has taken over all joint laws  $(U, V)$  such that  $\mathbb{P}_U = \mathbb{P}_{X^e}$  and  $\mathbb{P}_V = \mathbb{P}_X$  (see also Remark 2 and [10], Corollary 5.3.2). Consequently, in the framework of Theorem 6,  $|\mathbb{E}[X^e] - \mathbb{E}[X]| = \mathbb{E}|X^e - X|$ .  $\square$

**Remark 7.** *One can show that by means of Lemma 2 and Equation (82) it is possible to provide an estimate*

$$d_{\mathcal{K}}(X^e, Z) \leq \lambda \mathbb{E}|X^e - X|. \tag{83}$$

For each function  $h$  belonging to  $\mathcal{K}$ , in a similar way to Equation (82), one can apply Equation (10) together with equilibrium transformation. Now, it is sufficient to study the Stein equation with right derivative. Formula (13) gives a solution of the Stein equation according to Lemma 2. Note that for  $f_h$ , the right derivative coincides almost everywhere with the derivative, and the law of  $X^e$  is absolutely continuous according to Equation (15). Thus, for the Lipschitz function  $f_h$  (see Lemma 2), one can use an equilibrium transformation.

**Example 1.** Consider the distribution functions  $F_\varepsilon(x)$  of random variables  $X_\varepsilon$ , taking values  $\varepsilon$  and  $2 - \varepsilon$  with probabilities  $1/2, 0 < \varepsilon < 1$ . Formula (15) yields that  $X_\varepsilon^e$  has the following piece-line structure

$$F_\varepsilon^e(x) = \begin{cases} 0, & \text{if } x < 0, \\ x, & \text{if } 0 \leq x < \varepsilon, \\ x/2 + \varepsilon/2, & \text{if } \varepsilon \leq x < 2 - \varepsilon, \\ 1, & \text{if } 2 - \varepsilon \leq x. \end{cases}$$

If  $\varepsilon \geq 1/2$  then, for all  $x \in \mathbb{R}$ , the following inequality holds:  $F_\varepsilon^e(x) \geq F_\varepsilon(x)$ , i.e.,  $X_\varepsilon$  and  $X_\varepsilon^e$  are stochastically ordered. We see that for  $\varepsilon < 1/2$ , the inequality is violated in the right neighborhood of a point  $\varepsilon$ . Thus, there are beside the stochastically ordered pairs  $(X, X^e)$  also those of a different kind.

Now, we turn to another example of stochastically ordered  $X$  and  $X^e$ .

**Example 2.** Take  $X^e$  having the Pareto distribution. The notation  $X^e \sim \text{Pareto}(\alpha, \beta)$  means that  $X^e$  has a density  $f^e(x) = \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}}$  ( $x \geq 0$ ) and the corresponding distribution function  $F^e(x) = 1 - \left(\frac{\beta}{x+\beta}\right)^\alpha$ , where  $x \geq 0, \alpha > 0, \beta > 0$ .

Further, we consider only  $\alpha > 1$ , since in this case there exists finite  $\mathbb{E}[X^e] = \frac{\beta}{\alpha-1}$ . By means of Lemma 6, we obtain the distribution of the preimage of the equilibrium transformation

$$F(x) = 1 - \frac{f^e(x)}{f^e(0)} = 1 - \frac{\alpha\beta^\alpha}{(x+\beta)^{\alpha+1}} \frac{\beta^{\alpha+1}}{\alpha\beta^\alpha} = 1 - \left(\frac{\beta}{x+\beta}\right)^{\alpha+1}, \quad x \geq 0.$$

Thus one can state that  $X \sim \text{Pareto}(\alpha + 1, \beta)$ . It is not difficult to see that  $F^e(x) \leq F(x)$  for  $x \in \mathbb{R}$ , i.e., the random variables  $X^e$  and  $X$  are stochastically ordered. Due to Theorem 6, one has

$$d_{\mathcal{H}_1}(X^e, Z) = \mathbb{E}|X^e - X| = \mathbb{E}[X^e] - \mathbb{E}[X] = \frac{\beta}{\alpha - 1} - \frac{\beta}{\alpha} = \frac{\beta}{\alpha(\alpha - 1)}, \tag{84}$$

$$d_{\mathcal{K}}(X^e, Z) \leq \frac{\alpha}{\beta} \mathbb{E}|X^e - X| = \frac{1}{\alpha - 1}.$$

In such a way we find the bound for the Kolmogorov distance between the distributions  $\text{Pareto}(\alpha, \beta)$  and  $\text{Exp}(\alpha/\beta)$ . This relation demonstrates the convergence rate of  $d_1(X^e, Z)$  to zero as  $\alpha \rightarrow \infty$ . The estimate is nontrivial for  $\alpha > 2$ .

**Remark 8.** It is interesting that estimation of the proximity of the Pareto law to the Exponential one became important in signal processing, see [34] and references therein. Let  $X \sim \text{Pareto}(\alpha, \beta)$ , where  $\alpha > 0, \beta > 0$ , and  $Z \sim \text{Exp}(\lambda)$ . In [34], the author indicates that the Pinsker–Csiszár inequality was employed to derive

$$d_{\mathcal{K}}(X, Z) \leq \sqrt{2D_{KL}(X||Z)}, \tag{85}$$

where  $D_{KL}(X||Z)$  is the Kullback–Leibler divergence between laws of  $X$  and  $Z$ . More precisely, in the left-hand side of Equation (85) one can write the total variation distance  $d_{TV}(X, Z)$  between distributions of  $X$  and  $Z$ . Clearly,  $d_{\mathcal{K}}(X, Z) \leq d_{TV}(X, Z)$ . By evaluating  $D_{KL}(X||Z)$  and performing an optimal choice of parameter  $\lambda$ , it was demonstrated (formula (19) in [34]) that, for  $\alpha > 1$  and any  $\beta > 0$ ,

$$d_{\mathcal{K}}(X, Z) \leq \sqrt{\frac{2}{\alpha(\alpha - 1)}} \tag{86}$$

if  $\lambda = \frac{\alpha-1}{\beta}$ . The author of [34] on page 8 writes that in his previous work [57] the inequality

$$d_{\mathcal{K}}(X, Z) \leq \frac{3}{\alpha} \tag{87}$$

was established with the same choice of  $\lambda$ . Next, he also writes that “in the most cases  $\alpha > 2$ ” and notes that the estimate in Equation (86) involving the Kullback–Leibler divergence is more precise for  $\alpha > \frac{9}{7}$  than the estimate in Equation (87) obtained by the Stein method. Moreover, on page 4 of [34] we read: “The problem with the Stein approach is that the bounds do not suggest a suitable way in which, for a given Pareto model, an appropriate approximating Exponential distribution can be specified”. However, we have demonstrated that application of the inverse equilibrium transformation together with the Stein method permits indicating, whenever  $\alpha > 2$ , the corresponding Exponential distribution with proximity closer than the right-hand sides of Equation (86) and Equation (87) can provide.

### 8. Conclusions

Our principle goal was to find the sharp estimates of the proximity of random sums distributions to exponential and more general laws. This goal is achieved when we employ the probability metric  $d_{\mathcal{H}_2}$ . Thus, it would be valuable to find the best possible approximations of random sums distributions by means of specified laws using the metrics  $\zeta_s$  of order  $s > 0$ . The results of [32] provide the basis for this approach.

There are various complementary refinements of the Rényi theorem. One approach is related to the employment of Brownian motion. It is interesting that in [58] (p. 1071) the authors proposed an explanation of the Rényi theorem involving the embedding theorem. We provide a little bit different complete proof. Let  $X_1, X_2, \dots$  be i.i.d. random variables with mean  $\mu := \mathbb{E}X_1$  and  $\sigma^2 := \text{var}X_1 < \infty$ , whereas  $S_n, n \in \mathbb{N}$ , denote the corresponding partial sums. According to Theorem 12.6 of [59], which is due to A.V. Skorokhod and V. Strassen, there exists a standard Brownian motion  $B(t), t \geq 0$ , (perhaps it is defined on an extension of initial probability space) such that

$$\frac{1}{\sqrt{t}} \sup_{0 \leq u \leq t} |S_{[u]} - \mu u - \sigma B(u)| \xrightarrow{\mathbb{P}} 0, \quad t \rightarrow \infty, \tag{88}$$

and

$$\lim_{t \rightarrow \infty} \frac{S_{[t]} - \mu t - \sigma B(t)}{\sqrt{2t \log \log t}} = 0 \quad \text{a.s.}, \tag{89}$$

where  $\xrightarrow{\mathbb{P}}$  stands for convergence in probability, and a.s. means almost surely. Thus, in light of Equation (89), we can write, for  $t \geq 0$ ,

$$S_{[t]} = \mu t + \sigma B(t) + R(t), \tag{90}$$

where  $\sup_{0 \leq u \leq t} R(u) / \sqrt{t} \xrightarrow{\mathbb{P}} 0$  and  $R(t) / \sqrt{2t \log \log t} \rightarrow 0$  a.s. when  $t \rightarrow \infty$ . Substitute  $N_p$  (see Equation (2)) in Equation (90) instead of  $t$ . It is easily seen that  $N_p \xrightarrow{\mathbb{P}} \infty$  (i.e., for each  $t > 0$ , one has  $\mathbb{P}(N_p \leq t) \rightarrow 0$  as  $p \rightarrow 0+$ ) and by means of characteristic functions one can verify that  $pN_p \xrightarrow{\mathcal{D}} Z$  as  $p \rightarrow 0+$ , where  $Z \sim \text{Exp}(1)$ . Therefore,  $\mu pN_p \xrightarrow{\mathcal{D}} \mu Z, p \rightarrow 0+$ . In the proof of Lemma 4, we showed (Equation (24)) that  $\mathbb{E}[N_p] = (1 - p) / p$ . Consequently,

$$\begin{aligned} \text{var}[pB(N_p)] &= p^2 \mathbb{E}[B(N_p)^2] = p^2 \sum_{k=0}^{\infty} \mathbb{E}[B(k)^2] p(1 - p)^k \\ &= p^2 \sum_{k=0}^{\infty} kp(1 - p)^k = p^2 \mathbb{E}[N_p] = p^2 \frac{1 - p}{p} = p(1 - p) \rightarrow 0, \quad p \rightarrow 0+. \end{aligned}$$

Hence,  $p\sigma B(N_p) \xrightarrow{\mathbb{P}} 0$  as  $p \rightarrow 0+$ . Now, we demonstrate that  $pR(N_p) \xrightarrow{\mathbb{P}} 0$ ,  $p \rightarrow 0+$ . For any  $\varepsilon > 0$  and any  $t > 0$ ,

$$\begin{aligned} \mathbb{P}(p|R(N_p)| > \varepsilon) &\leq \mathbb{P}(p|R(N_p)| > \varepsilon, N_p \leq t) + \mathbb{P}(N_p > t) \\ &\leq \mathbb{P}(p \sup_{0 \leq u \leq t} |R(u)| > \varepsilon) + \mathbb{P}(N_p > t). \end{aligned}$$

In light of Equation (88), for arbitrary  $\gamma > 0$  and  $\varepsilon > 0$ , one can take  $t_0 = t_0(\gamma)$  such that  $\mathbb{P}(\sup_{0 \leq u \leq t_0} |R(u)| > \varepsilon\sqrt{t_0}) < \gamma/2$ . Then, for any  $0 < p \leq 1/\sqrt{t_0}$ , we obtain

$$\mathbb{P}(p \sup_{0 \leq u \leq t_0} |R(u)| > \varepsilon) < \gamma/2.$$

Since  $N_p \xrightarrow{\mathbb{P}} \infty$ , we can find  $p_0 > 0$  such that  $\mathbb{P}(N_p > t_0) < \gamma/2$  if  $0 < p \leq p_0$ . Therefore,  $R(N_p) \xrightarrow{\mathbb{P}} 0$  as  $p \rightarrow 0+$ . The Slutsky lemma yields the desired relation

$$pS_{N_p} \xrightarrow{\mathcal{D}} \mu Z, \quad p \rightarrow 0+,$$

which implies Equation (3). However, it seems that there is no clear intuitive reason why the law of the random sum converges to an exponential in the Rényi theorem. Moreover, in Ch. 3, Sec. 2 “The Rényi Limit Theorem” of [20] (see Sec. 2.1 “Motivation”), one can find examples demonstrating that intuition behind the Rényi theorem is poor.

Actually, relation (90) leads to refinements of Equation (3). In [58], it is proved that if  $X_1$  has finite exponential moments and other specified conditions are satisfied then there exists a more sophisticated approximation for distribution of  $W_p$ , and its accuracy is estimated. The results are applied to the study of  $M/G/1$  queue for both light-tailed and heavy-tailed service time distributions. Note that in [58], Section 5, the authors study the model where the distribution of  $X_1$  can depend on  $p$ . For future research, it would be desirable to establish analogues of our theorems for such a model.

The results concerning the accuracy of approximating a distribution under consideration by an exponential law are applicable to some queuing models. Let, for a queue  $M/G/1$ , the inter-arrival times follow  $Exp(\lambda)$  distribution and  $S$  stand for the general service time. Introduce the stationary waiting time  $W$  and define  $\rho := \lambda\mathbb{E}[S]$  to be its load. Due to [60], if  $\mathbb{E}[S^3] < \infty$  then  $(1 - \rho)W \xrightarrow{\mathcal{D}} Z$  as  $\rho \rightarrow 1$ , where  $Z \sim Exp(1)$ . Theorem 3.1 of [45] contains an upper bound of  $d_{\mathcal{H}_1}(W_p, Z)$ , where  $Z \sim Exp(1)$ . This estimate is used by the authors for analysis of queueing systems with a single server. It would be interesting to obtain the sharp approximations in the framework of queueing systems.

For the model of exchangeable random variables, Theorem 2 in Section 2 ensures the weak convergence of distributions under consideration to specified mixture of explicitly indicated laws. Theorem 3 proves the sharp convergence rate estimate to this limit law by means of the ideal probability metric of the second order. It would be worthwhile to establish such an estimate of the distributions proximity applying the Lévy–Prokhorov distance because convergence in this metric is equivalent to the weak convergence of distributions of random variables. All the more, at present there is no unified theory of probability metrics. In this regard, one can mention Proposition 1.2 of [17] stating that if a random variable  $Z$  has the Lebesgue density bounded by  $C$  then, for any random variable  $Y$ ,

$$d_{\mathcal{K}}(Y, Z) \leq \sqrt{Cd_{\mathcal{H}_1}(Y, Z)}.$$

However, this estimate only gives the sub-optimal convergence rates. We also highlight the important total variation distance  $d_{TV}$ . The authors of [61] study the sum  $W := \sum_{j \in J} X_j$ , where  $\{X_j, j \in J\}$  is a family of locally dependent non-negative integer-valued random variables. Using the perturbations of Stein’s operator, they establish the upper bounds of  $d_{TV}(W, M)$  where the law of  $M$  is a mixture of Poisson distribution and either binomial or

negative binomial distribution. It would be desirable to obtain the sharp estimates and, moreover, consider a more general model where the set of summation is random. In this connection, it seems helpful to employ the paper [62], where the authors proved results concerning the weak convergence of distributions of statistics constructed from samples of random size. In addition, it would be interesting to extend these results to stratified samples by invoking Lemma 1 of [63].

Special attention is paid to various generalizations of the geometric sums. In Theorem 3.3 of [64], the authors consider random sums with summation index  $T_n := Y_1 + \dots + Y_n$ , where  $Y_1, Y_2, \dots$  are i.i.d. random variables following the geometric law  $Geom(p)$ , see Equation (2). Then, they show that  $S_{T_n} / \mathbb{E}[S_{T_n}]$  converge in distribution to the gamma law with certain parameters as  $p \rightarrow 0+$ . In [62], it is demonstrated that the Linnik and the Mittag-Leffler laws arise naturally in the framework of limit theorems for random sums. Hopefully, in future the complete picture of limit laws involving general theory of distributions mixtures will appear. In addition, it is desirable to study various models of random sums of dependent random variables. On this track, it could be useful to consider the decompositions of exchangeable random sequences extending the fundamental de Finetti theorem, see, e.g., [65].

One can try to generalize the results of Section 7 for accumulative laws proposed in [66]. These laws are akin to both the Pareto distribution and the lognormal distribution. In addition, we refer to [43] where the “variance-gamma distributions” were studied. These distributions form a four-parameter family and comprise as special and limiting cases the normal, gamma and Laplace distributions. Employment of these distributions permits enlarging a range of applications in modeling and fitting real data.

To complete the indication of further research directions, we note that the next essential and nontrivial step is to establish the limit theorem in functional spaces for processes generated by a sequence of random sums of random variables. For such stochastic processes, one can obtain the analogues of the classical invariance principles.

**Author Contributions:** Conceptualization, A.B. and N.S.; methodology, A.B. and N.S.; formal analysis, A.B. and N.S.; investigation, A.B. and N.S.; writing—original draft preparation, A.B. and N.S.; writing—review and editing, A.B. and N.S.; supervision, A.B.; project administration, A.B.; funding acquisition, A.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** The work was supported by the Lomonosov Moscow State University project “Fundamental Mathematics and Mechanics”.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors are grateful to Alexander Tikhomirov for invitation to present manuscript for this issue. In addition, they would like to thank three anonymous Reviewers for the careful reading of the manuscript and valuable remarks.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Proof of Lemma 1.** If  $Lip(h) = C < \infty$ , then  $h$  is absolutely continuous (see, e.g., §13 in [42]), and consequently there exists  $h'(x)$  for almost all  $x \in \mathbb{R}$ . Thus,  $|h'(x)| \leq C$  for almost all  $x \in \mathbb{R}$  in light of Equation (4). Assume that essential supremum  $\|h'\|_\infty = C_0 < C$ . Then, for any  $\varepsilon > 0$ , one can find a version of  $h'$ , defined on  $\mathbb{R}$ , such that  $\sup_{x \in \mathbb{R}} |h'(x)| \leq C_0 + \varepsilon$ . (It was explained in Section 2 that one can consider a measurable extension of  $h'$  to  $\mathbb{R}$ ). Then, due to Equation (11) with  $h$  instead of  $f$  we obtain Equation (5) with  $C_0 + \varepsilon$  instead of  $C$ . Consequently,  $Lip(h) \leq C_0 + \varepsilon < C$ . We come to the contradiction.

On the other hand, let  $h$  be absolutely continuous. Then, for almost all  $x \in \mathbb{R}$ , there exists  $h'(x)$  and Equation (11) is valid for  $h$  instead of  $f$ . Assume that essential supremum

$\|h'\|_\infty = C < \infty$ . Then, for any  $\varepsilon > 0$  there is a version of  $h'$  such that  $\sup_{x \in \mathbb{R}} |h'(x)| \leq C + \varepsilon$ . According to Equation (11), the relation (5) holds with  $C + \varepsilon$  instead of  $C$ . Since  $\varepsilon > 0$  can be taken as an arbitrary small, one can claim that  $Lip(h) \leq C$ . Suppose that  $Lip(h) \leq C_0 < C$ . Then, for almost all  $x \in \mathbb{R}$ , there exists  $h'$  and  $|h'| \leq C_0$ . Thus, we found a version with  $\|h'\|_\infty \leq C_0$ . The contradiction shows that  $Lip(h) = C$ . Hence, the desired statement is proved.  $\square$

**Proof of Lemma 2.** Let  $x_0$  be a continuity point of a function  $h \in \mathcal{K} \cup \mathcal{H}_1 \cup \mathcal{H}_2$ . Then, the same is true for a function  $h(u)e^{-\lambda u}$ ,  $u \in \mathbb{R}$ . Hence, the function  $\int_x^\infty h(u)e^{-\lambda u} du$  has a derivative  $-h(x_0)e^{-\lambda x_0}$  at point  $x_0$  (in light of Remark 1 an integral  $\int_x^\infty h(u)e^{-\lambda u} du$  is well defined for any  $x \in \mathbb{R}$ ). Thus, for each point  $x$  of continuity  $h$  there exists

$$f'_h(x) = -\lambda e^{\lambda x} \int_x^\infty h(u)e^{-\lambda u} du - e^{\lambda x}(-h(x)e^{-\lambda x}) = \lambda f_h(x) + h(x). \tag{A1}$$

For each fixed  $z \in \mathbb{R}$  and a function  $h(x) = \mathbb{I}\{x \leq z\}$ , where  $x \in \mathbb{R}$ , Equation (12) is verified in a similar way for the right derivative  $f'_h$  at point  $z \in \mathbb{R}$ . Taking  $x = 0$  in Equation (12), we obtain  $-\mathbb{E}[h(Z)]/\lambda$ . Evidently,  $-e^{\lambda x} \int_x^\infty e^{-\lambda u} du = -1/\lambda$ . Therefore, Equation (A1) yields

$$f'_h(x) = -\lambda e^{\lambda x} \int_x^\infty (h(u) - h(x))e^{-\lambda u} du. \tag{A2}$$

If a function  $h$  belongs to  $\mathcal{K}$ , then, for any  $u, x \in \mathbb{R}$ , the following inequality holds  $|h(u) - h(x)| \leq 1$ . Consequently, for  $h \in \mathcal{K}$ , one has  $\|f'_h\|_\infty \leq 1$  (where  $f'_h$  means a right derivative of a version of  $f'_h$ , and we operate with essential supremum).

Taking into account Lemma 1, for a function  $h \in \mathcal{H}_1$  and any  $x \leq u$ , one can write  $|h(u) - h(x)| \leq Lip(h)(u - x) = \|h'\|_\infty(u - x)$ . For  $h \in \mathcal{H}_2$  and  $x \leq u$ , by the Lagrange finite-increments formula,  $|h(u) - h(x)| \leq |h'(v)|(u - x) \leq \|h'\|_\infty(u - x)$ , where  $x < v < u$ . Hence, for any  $x \in \mathbb{R}$  and  $h \in \mathcal{H}_1 \cup \mathcal{H}_2$ ,

$$|f'_h(x)| = \lambda e^{\lambda x} \int_x^\infty (h(u) - h(x))e^{-\lambda u} du \leq \lambda e^{\lambda x} \|h'\|_\infty \int_x^\infty (u - x)e^{-\lambda u} du = \frac{\|h'\|_\infty}{\lambda}$$

since

$$\lambda e^{\lambda x} \int_x^\infty (u - x)e^{-\lambda u} du = \int_0^\infty \lambda v e^{-\lambda v} dv = \frac{1}{\lambda}. \tag{A3}$$

Taking into account Equation (12), one can see that, for any  $h \in \mathcal{H}_2$ ,  $f'_h = \lambda f_h + h$ , where  $f_h$  and  $h$  have derivatives at each point  $x \in \mathbb{R}$ . Using Equation (A2) and Equation (A3), we obtain, for  $x \in \mathbb{R}$ ,

$$\begin{aligned} f''_h(x) &= \lambda f'_h(x) + h'(x) = -\lambda^2 e^{\lambda x} \int_x^\infty (h(u) - h(x))e^{-\lambda u} du + h'(x) \\ &= -\lambda^2 e^{\lambda x} \int_x^\infty (h(u) - h(x) - h'(x)(u - x))e^{-\lambda u} du. \end{aligned} \tag{A4}$$

By means of Equation (A3) and the Lagrange finite-increments formula we can write

$$|f''_h(x)| \leq 2\|h'\|_\infty \lambda^2 e^{\lambda x} \int_x^\infty (u - x)e^{-\lambda u} du = 2\|h'\|_\infty. \tag{A5}$$

Let us apply the Taylor formula with integral representation of the residual term:

$$h(u) = h(x) + h'(x)(u - x) + R(u, x), \quad R(u, x) = \int_x^u (u - t)h''(t)dt, \quad u, x \in \mathbb{R}. \tag{A6}$$

This representation known for the Riemann integral (see, e.g., [67], §9.17) holds in the framework of the Lebesgue integral if it is possible to use the recurrent integration by parts for  $R(u, x)$ , i.e.,

$$\int_x^u (u - t)h''(t)dt = -h'(x)(u - x) + \int_x^u h'(t)dt = -h'(x)(u - x) + h(u) - h(x). \tag{A7}$$

Integral in the left-hand side of Equation (A7) exists by virtue of Lemma 1 since  $h' \in \text{Lip}(1)$ . Therefore,  $h''(x)$  is defined for almost all  $x \in \mathbb{R}$  and (essential supremum)  $\|h''\| \leq 1$ . The latter equality in Equation (A7) is obvious since  $h'$  is continuous function on  $\mathbb{R}$ . The first equality in Equation (A7) is valid due to the integration by parts formula for the Lebesgue integral. Indeed, functions  $h'(t)$  and  $(u - t)$  are absolutely continuous for  $t$  belonging to  $[x, u]$ . Thus, we can apply, e.g., Theorem 13.29 of [42] to justify the first equality in Equation (A7). Consequently, due to Equation (A4) and Equation (A6) one can write

$$\begin{aligned} |f_h''(x)| &\leq \left| -\lambda^2 e^{\lambda x} \int_x^\infty \left( \int_x^u (u - t)h''(t)dt \right) e^{-\lambda u} du \right| \\ &\leq \frac{\|h''\|_\infty}{2} \left| \int_x^\infty \lambda^2 (u - x)^2 e^{-\lambda(u-x)} du \right| = \frac{\|h''\|_\infty \Gamma(3)}{2\lambda} = \frac{\|h''\|_\infty}{\lambda}, \end{aligned} \tag{A8}$$

where  $\Gamma(\alpha) := \int_0^\infty u^{\alpha-1} e^{-u} du, \alpha > 0$ . Relations Equation (A5) and Equation (A8) lead to the last statement of Lemma 2. The proof is complete.  $\square$

**Comments to Definition 1.** For each Lipschitz function  $f$ , one can claim that  $\mathbb{E}[f(X)]$  is finite since  $\mathbb{E}|X| < \infty$  and, in light of Remark 1, one has  $|f(x)| \leq C|x| + |f(0)|$ , where  $C = \text{Lip}(f), x \in \mathbb{R}$ . Clearly, it is sufficient to verify Equation (14) for any Lipschitz function  $f$  such that  $f(0) = 0$  (otherwise we take the Lipschitz function  $f(x) - f(0), x \in \mathbb{R}$ ). Evidently,  $p^e(x), x \in \mathbb{R}$ , introduced by Equation (15), is a probability density because for non-negative random variable  $X$  according to [47], Ch.2, formula (69)

$$\mathbb{E}[X] = \int_{[0,\infty)} \mathbb{P}(X > u) du. \tag{A9}$$

We will show that, for such  $f$  and a density  $p^e$  of  $X^e$ , one has

$$\int_{[0,\infty)} f(u) dF(u) = \int_{[0,\infty)} f'(u) \mathbb{P}(X > u) du, \tag{A10}$$

where  $F$  is a distribution function of  $X$  and  $\mathbb{E}[X] \neq 0$ . We take integrals over  $[0, \infty)$  as  $X \geq 0$  and  $p^e(x) = 0$  for  $x < 0$ .

We know that a function  $f$  has a derivative at almost all points  $x \in \mathbb{R}$ . Therefore, the right-hand side of Equation (A10) does not depend on the choice of a version  $f'$  ( $\mathbb{P}(X > u)$  is a measurable bounded function). The integral in the right-hand side of Equation (A10) is finite because  $\|f'\| \leq C$  in light of Lemma 1 and since the right-hand side of Equation (A9) is finite. One can take the integrals over  $(0, \infty)$  in Equation (A10) as  $f(0) = 0$  and  $m(\{0\}) = 0$ , where  $m$  stands for the Lebesgue measure.

Function  $f$  is a function of finite variation (as  $f$  is the Lipschitz function). Therefore,  $f = f_1 - f_2$  where  $f_1$  and  $f_2$  are nondecreasing functions. We can take the canonical representation with  $f_1(x) = \text{Var}_0^x(f)$  and  $f_2(x) = f(x) - f_1(x), x \in \mathbb{R}$ , where  $\text{Var}_a^b(f)$  is the variation of  $f$  on  $[a, b], a < b$  (see, e.g., [42], Theorem 12.18). If  $f \in \text{Lip}(C)$ , then  $\text{Var}_a^b(f) \leq C(b - a)$ . For  $a < c < b$ , one has (see, e.g., [42], Lemma 12.15)

$$\text{Var}_a^c(f) + \text{Var}_c^b(f) = \text{Var}_a^b(f).$$

We see that such  $f_1$  and  $f_2$  are the Lipschitz functions when  $f$  is the Lipschitz one. Hence, for almost all  $x \in \mathbb{R}$ , there exist  $f_1'(x), f_2'(x)$  and  $f'(x) = f_1'(x) - f_2'(x)$ . Thus, it is enough to demonstrate that

$$\int_{(0,\infty)} f_i(u) dF(u) = \int_{(0,\infty)} f_i'(u) \mathbb{P}(X > u) du, \quad i = 1, 2.$$

These integrals are finite since  $f_1$  and  $f_2$  are the Lipschitz functions. Note that

$$\int_{(0,\infty)} f_i(u)dF(u) = - \int_{(0,\infty)} f_i(u)d(1 - F(u)) = - \int_{(0,\infty)} f_i(u)d\mathbb{P}(X > u).$$

By applying Theorem 11 of Sec. 6, Ch. 2 [47], one obtains, for each  $b > 0$ , nondecreasing continuous function  $f_i$  and a nondecreasing right-continuous function  $(-\mathbb{P}(X > u))$ , the following formula:

$$\begin{aligned} \int_{(0,b]} f_i(u)d\mathbb{P}(X > u) &= f_i(b)\mathbb{P}(X > b) - f_i(0)\mathbb{P}(X > 0) - \int_{(0,b]} \mathbb{P}(X > u)df_i(u) \quad (A11) \\ &= f_i(b)\mathbb{P}(X > b) - \int_{(0,b]} \mathbb{P}(X > u)f_i'(u)du. \end{aligned}$$

We take into account that  $f_i(0) = 0$  and the  $\sigma$ -finite measure  $Q_i$  corresponding to  $f_i$  is absolutely continuous w.r.t.  $m$ , and the Radon–Nikodým derivative  $\frac{dQ_i}{dm}(x) = f_i'(x)$ ,  $x \in \mathbb{R}$ ,  $i = 1, 2$ . In addition, we can write  $\mathbb{P}(X > u)$  in Equation (A11) since for at almost all  $u \in \mathbb{R}$  the left-limit of this function coincides with  $\mathbb{P}(X > u)$  (there exist at most a countable set of jumps of  $\mathbb{P}(X > u)$ ,  $u \in \mathbb{R}$ ). Obviously,  $f_i(b)\mathbb{P}(X > b) \rightarrow 0$  as  $b \rightarrow \infty$  because  $|f_i(u)| \leq A_iu + B_i$  for some positive  $A_i, B_i$  and all  $u \in \mathbb{R}$ . Indeed, according to formula (73) of Sec. 6, Ch. 2 of [47] the condition  $\mathbb{E}|X| < \infty$  yields

$$b\mathbb{P}(|X| > b) \rightarrow 0, \quad b \rightarrow \infty.$$

By the Lebesgue dominated convergence theorem one infers that

$$\int_{(0,b]} f_i(u)d\mathbb{P}(X > u) \rightarrow \int_{(0,\infty)} f_i(u)d\mathbb{P}(X > u), \quad b \rightarrow \infty.$$

and

$$\lim_{b \rightarrow \infty} \int_{(0,b]} \mathbb{P}(X > u)f_i'(u)du = \int_{(0,\infty)} \mathbb{P}(X > u)f_i'(u)du.$$

This permits to claim the validity of Equation (A10) which entails the desired Equation (15).

**Proof of Lemma 3.** For  $f \in \mathcal{H}_2$ , in light of Remark 1 one can state that  $|f(x)| \leq A_0x^2 + B_0$  for some positive numbers  $A_0$  and  $B_0$ . Let  $F$  be a distribution function of  $X$ . Since  $\mathbb{E}[X^2] < \infty$ , due to Corollary 2, Sec. 6, Ch. 2, v.1, [47] one has

$$x^2F(x) \rightarrow 0, \quad x \rightarrow -\infty; \quad x^2(1 - F(x)) \rightarrow 0, \quad x \rightarrow \infty.$$

Hence, we obtain that  $f(x)F(x) \rightarrow 0$  as  $x \rightarrow -\infty$  and  $f(x)(1 - F(x)) \rightarrow 0$  as  $x \rightarrow \infty$ . Continuous function  $f$  has a bounded variation. Thus  $f = f_1 - f_2$  where  $f_1$  and  $f_2$  are nondecreasing continuous functions. Thus, for any  $a < 0$  and  $i = 1, 2$ , the integration by parts formula (see, e.g., Theorem 11, Sec. 6, Ch. 2, [47]) and Equation (18) give

$$\begin{aligned} \int_{(a,0]} (f_1(x) - f_2(x))dF(x) &= f(0)F(0) - f(a)F(a) - \left( \int_{(a,0]} F(x)df_1(x) - \int_{(a,0]} F(x)df_2(x) \right) \\ &= f(0)F(0) - f(a)F(a) - \int_{(a,0]} F(x)df(x). \end{aligned}$$

We take into account that the integrands are bounded measurable functions and the measures corresponding to  $F, f_1$  and  $f_2$  are finite on any interval  $(a, 0]$ . Therefore such integrals are finite. According to the Lebesgue theorem on dominated convergence (recall that  $\mathbb{E}[X^2] < \infty$ ) one has

$$\lim_{a \rightarrow -\infty} \int_{(a,0]} f(x)dF(x) = \int_{(-\infty,0]} f(x)dF(x),$$

and the limit is finite. The monotone convergence theorem for  $\sigma$ -finite measure yields

$$\lim_{a \rightarrow -\infty} \left( \int_{(a,0]} F(x)df_1(x) - \int_{(a,0]} F(x)df_2(x) \right) = \int_{(-\infty,0]} F(x)df_1(x) - \int_{(-\infty,0]} F(x)df_2(x).$$

We have seen that  $f(a)F(a) \rightarrow 0$  as  $a \rightarrow -\infty$ . Hence, in light of Equation (18)

$$\int_{(-\infty,0]} F(x)df_1(x) - \int_{(-\infty,0]} F(x)df_2(x) = \int_{(-\infty,0]} F(x)df(x).$$

Therefore, for  $i = 1, 2$ , each integral  $\int_{(-\infty,0]} F(x)df_i(x)$  is finite as  $\int_{(-\infty,0]} F(x)df(x)$  is finite. Thus,

$$\int_{(-\infty,0]} f(x)dF(x) = f(0)F(0) - \int_{(-\infty,0]} F(x)df(x) = f(0)F(0) + \int_{(-\infty,0]} (-F(x))f'(x)dx,$$

as  $f$  is absolutely continuous. Indeed, for any  $x \in \mathbb{R}$ ,

$$f(x) = f(0) + \int_{(0,x]} f'(u)du,$$

where (continuous)  $f' \in L^1[a, b]$  for any finite interval  $[a, b]$ . Thus,  $(f')^+ \in L^1[a, b]$  and  $(f')^- \in L^1[a, b]$ . Set

$$f_1(x) := f(0) + \int_{(0,x]} (f'(u))^+ du, \quad f_2(x) := \int_{(0,x]} (f'(u))^- du.$$

Then  $f_1$  and  $f_2$  are nondecreasing continuous functions on  $\mathbb{R}$ ,  $f = f_1 - f_2$  and

$$\int_{(a,0]} F(x)df(x) = \int_{(a,0]} F(x)df_1(x) - \int_{(a,0]} F(x)df_2(x),$$

where these three integrals are finite. For (non-negative)  $\sigma$ -finite measures corresponding to  $f_1$  and  $f_2$ , one can write

$$\int_{(a,0]} F(x)df_1(x) = \int_{(a,0]} F(x)(f'(x))^+ dx, \quad \int_{(a,0]} F(x)df_2(x) = \int_{(a,0]} F(x)(f'(x))^- dx.$$

Thus, one has

$$\begin{aligned} \int_{(a,0]} F(x)df(x) &= \int_{(a,0]} F(x)(f'(x))^+ dx - \int_{(a,0]} F(x)(f'(x))^- dx \\ &= \int_{(a,0]} F(x)((f'(x))^+ - (f'(x))^-) dx = \int_{(a,0]} F(x)f'(x) dx. \end{aligned} \tag{A12}$$

The bound  $\|f'\| \leq 1$  follows from Lemma 1. Therefore, the Lebesgue theorem on dominated convergence yields (as  $\mathbb{E}|X| < \infty$ )

$$\lim_{a \rightarrow -\infty} \int_{(a,0]} F(x)f'(x) dx = \int_{(-\infty,0]} F(x)f'(x) dx.$$

We have demonstrated that

$$\int_{(-\infty,0]} F(x)df(x) = \int_{(-\infty,0]} F(x)f'(x) dx.$$

In a similar way, we consider  $\int_{(0,b]} (1 - F(x))dx$  and letting  $b \rightarrow \infty$  come to relation

$$- \int_{(0,\infty)} f(x)d(1 - F(x)) = f(0)(1 - F(0)) + \int_{(0,\infty)} (1 - F(x))df(x)$$

$$= f(0)(1 - F(0)) + \int_{(0,\infty)} (1 - F(x))f'(x)dx.$$

This establishes Equation (21).  $\square$

## References

1. Steutel, F.W.; Van Harn, K. *Infinite Divisibility of Probability Distributions on the Real Line*; Marcel Dekker: New York, NY, USA, 2004.
2. Nolan, J.P. *Univariate Stable Distributions. Models for Heavy Tailed Data*; Springer: Cham, Switzerland, 2020.
3. Jagers, P. Branching processes: Personal historical perspective. In *Statistical Modeling for Biological Systems*; Almudevar, A., Oakes, D., Hall, J., Eds.; Springer: Cham, Switzerland, 2020; pp. 1–12. [\[CrossRef\]](#)
4. Schmidli, H. *Risk Theory*; Springer: Cham, Switzerland, 2017.
5. Gnedenko, B.V.; Korolev V.Y. *Random Summation. Limit Theorems and Applications*; CRC Press: Boca Raton, FL, USA, 1996.
6. Kalashnikov V.V. *Geometric Sums: Bounds for Rare Events with Applications*; Kluwer Academic: Dordrecht, The Netherlands, 1997.
7. Pinski, M.A.; Karlin, S. *An Introduction to Stochastic Modeling*, 4th ed.; Academic Press: Amsterdam, The Netherlands, 2011.
8. Bulinski, A.; Spodarev, E. Introduction to random fields. In *Stochastic Geometry, Spacial Statistics and Random Fields. Asymptotic Methods*; Spodarev, E., Ed.; Springer: Berlin, Germany, 2013; pp. 277–336. [\[CrossRef\]](#)
9. Zolotarev, V.M. *Modern Theory of Summation of Random Variables*; De Gruyter: Berlin, Germany, 1997.
10. Rachev, S.T.; Klebanov, L.B.; Stoyanov, S.V.; Fabozzi, F.J. *The Methods of Distances in the Theory of Probability and Statistics*; Springer: New York, NY, USA, 2013.
11. Stein, C. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*; Statistical Laboratory of the University of California: Berkeley, CA, USA, 1972; pp. 583–602.
12. Stein, C. *Approximate Computation of Expectations, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 7*; Institute of Mathematical Statistics: Hayward, CA, USA, 1986.
13. Slepov, N.A. Convergence rate of random geometric sum distributions to the Laplace law. *Theory Probab. Appl.* **2021**, *66*, 121–141. [\[CrossRef\]](#)
14. Tyurin, I.S. On the convergence rate in Lyapunov’s theorem. *Theory Probab. Appl.* **2011**, *55*, 253–270. [\[CrossRef\]](#)
15. Barbour, A.D.; Chen, L.H.Y. (Eds.) *An Introduction to Stein’s Method*; World Scientific: Singapore, 2005.
16. Chen, L.H.Y.; Goldstein, L.; Shao, Q.-M. *Normal Approximation by Stein’s Method*; Springer: Heidelberg, Germany, 2011.
17. Ross, N. Fundamentals of Stein’s method. *Probab. Surv.* **2011**, *8*, 210–293. [\[CrossRef\]](#)
18. Arras, B.; Breton, J.-C.; Deshayes, A.; Durieu, O.; Lachièze-Rey, R. Some recent advances for limit theorems. *ESAIM Proc. Surv.* **2020**, *68*, 73–96. [\[CrossRef\]](#)
19. Arras, B.; Houdré, C. *On Stein’s Method for Infinitely Divisible Laws with Finite First Moment*, 1st ed.; Springer: Cham, Switzerland, 2019.
20. Chen, P.; Nourdin, I.; Xu, L.; Yang, X.; Zhang, R. Non-integrable Stable Approximation by Stein’s Method. *J. Theor. Probab.* **2022**, *35*, 1137–1186. [\[CrossRef\]](#)
21. Rényi, A. (Hungarian) A characterization of Poisson processes. *Magyar Tud. Akad. Mat. Kutató. Int. Közl.* **1957**, *1*, 519–527.
22. Shevtsova, I.; Tselishchev, M. A generalized equilibrium transform with application to error bounds in the Rényi theorem with no support constraints. *Mathematics* **2020**, *8*, 577. [\[CrossRef\]](#)
23. Brown, M. Error bounds for exponential approximations of geometric convolutions. *Ann. Probab.* **1990**, *18*, 1388–1402. [\[CrossRef\]](#)
24. Brown, M. Sharp bounds for exponential approximations under a hazard rate upper bound. *J. Appl. Probab.* **2015**, *52*, 841–850. [\[CrossRef\]](#)
25. Hung, T.L.; Kein, P.T. On the rates of convergence in weak limit theorems for normalized geometric sums. *Bull. Korean Math. Soc.* **2020**, *57*, 1115–1126. [\[CrossRef\]](#)
26. Shevtsova, I.; Tselishchev, M. On the accuracy of the exponential approximation to random sums of alternating random variables. *Mathematics* **2020**, *8*, 1917. [\[CrossRef\]](#)
27. Korolev, V.; Zeifman, A. Bounds for convergence rate in laws of large numbers for mixed Poisson random sums. *Stat. Probab.* **2021**, *168*, 108918. [\[CrossRef\]](#)
28. Aldous, D.J. More Uses of Exchangeability: Representations of Complex Random Structures. In *Probability and Mathematical Genetics: Papers in Honour of Sir John Kingman*; Bingham, N.H., Goldie, C.M., Eds.; Cambridge University Press: Cambridge, UK, 2010.
29. Shevtsova, I.; Tselishchev, M. On the accuracy of the generalized gamma approximation to generalized negative binomial random sums. *Mathematics* **2021**, *9*, 1571. [\[CrossRef\]](#)
30. Liu, Q.; Xia, A. Geometric sums, size biasing and zero biasing. *Electron. Commun. Probab.* **2022**, *27*, 1–13. [\[CrossRef\]](#)
31. Döbler, C.; Peccati, G. The Gamma Stein equation and noncentral de Jong theorems. *Bernoulli* **2018**, *24*, 3384–3421. [\[CrossRef\]](#)
32. Korolev, V. Bounds for the rate of convergence in the generalized Rényi theorem. *Mathematics* **2022**, *10*, 4252. [\[CrossRef\]](#)
33. Peköz, E.A.; Röllin, A. New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **2011**, *39*, 587–608. [\[CrossRef\]](#)
34. Weinberg, G.V. Kulback-Leibler divergence and the Pareto-Exponential approximation. *SpringerPlus* **2016**, *5*, 604. [\[CrossRef\]](#)

35. Daly, F. Gamma, Gaussian and Poisson approximations for random sums using size-biased and generalized zero-biased couplings. *Scand. Actuar. J.* **2022**, *24*, 471–487. [[CrossRef](#)]
36. Zolotarev, V.M. Ideal metrics in the problem of approximating the distributions of sums of independent random variables. *Theory Probab. Appl.* **1977**, *22*, 433–449. [[CrossRef](#)]
37. Gibbs, A.L.; Su, F.E. On choosing and bounding probability metrics. *Int. Stat. Rev.* **2002**, *70*, 419–435. [[CrossRef](#)]
38. Janson, S. Probability Distances. 2020. Available online: [www2.math.uu.se/~svante](http://www2.math.uu.se/~svante) (accessed on 1 September 2022).
39. Peköz, E.A.; Röllin, A.; Ross, N. Total variation error bounds for geometric approximation. *Bernoulli* **2013**, *19*, 610–632. [[CrossRef](#)]
40. Slepov, N.A. Generalized Stein equation on extended class of functions. In Proceedings of the International Conference on Analytical and Computational Methods in Probability Theory and Its Applications, Moscow, Russia, 23–27 October 2017; pp. 75–79.
41. Ley, C.; Reinert, G.; Swan, Y. Stein’s method for comparison of inivariate distributions. *Probab. Surv.* **2017**, *14*, 1–52. [[CrossRef](#)]
42. Yeh, J. *Real Analysis. Theory of Measure and Integration*, 2nd ed.; World Scientific: Singapore, 2006.
43. Gaunt, R.E. Wasserstein and Kolmogorov error bounds for variance gamma approximation via Stein’s method I. *J. Theor. Probab.* **2020**, *33*, 465–505. [[CrossRef](#)]
44. Halmos, P.R. *Measure Theory*; Springer: New York, NY, USA, 1974.
45. Gaunt, R.E.; Walton, N. Stein’s method for the single server queue in heavy traffic. *Stat. Probab. Lett.* **2020**, *156*, 108566. [[CrossRef](#)]
46. Muthukumar, T. Measure Theory and Lebesgue Integration. 2018. Available online: [home.iitk.ac.in/~tmk](http://home.iitk.ac.in/~tmk) (accessed on 1 September 2022).
47. Shiryaev, A.N. *Probability-1*; Springer: New York, NY, USA, 2016.
48. Burkill, L.C. *The Lebesgue Integral*; Cambridge University Press: Cambridge, UK, 1963.
49. Korolev, V.; Zeifman, A. Generalized negative binomial distributions as mixed geometric laws and related limit theorems. *Lith. Math. J.* **2019**, *59*, 366–388. [[CrossRef](#)]
50. Baldi, P.; Rinott, Y.; Stein, C. A normal approximations for the number of local maxima of a random function on a graph. In *Probability, Statistics and Mathematics, Papers in Honor of Samuel Karlin*; Anderson, T.W., Athreya, K.B., Iglehart, D.L., Eds.; Academic Press: San-Diego, CA, USA, 1989; pp. 59–81. [[CrossRef](#)]
51. Goldstein, L.; Rinott, Y. Multivariate normal approximations by Stein’s method and size bias couplings. *J. Appl. Prob.* **1996**, *33*, 1–17. [[CrossRef](#)]
52. Goldstein, L. Berry-Esseen bounds for combinatorial central limit theorems and pattern occurrences, using zero and size biasing. *J. Appl. Probab.* **2005**, *42*, 661–683. [[CrossRef](#)]
53. Goldstein, L.; Reinert, G. Stein’s method and the zero bias transformation with application to simple random sampling. *Ann. Appl. Probab.* **1997**, *7*, 935–952. [[CrossRef](#)]
54. Gaunt, R.E. On Stein’s method for products of normal random variables and zero bias couplings. *Bernoulli* **2017**, *23*, 3311–3345. [[CrossRef](#)]
55. Döbler, C. Distributional transformations without orthogonality relations. *J. Theor. Probab.* **2017**, *30*, 85–116. [[CrossRef](#)]
56. Arratia, R.; Goldstein, L.; Kochman, F. Size bias for one and all. *Probab. Surv.* **2019**, *16*, 1–61. [[CrossRef](#)]
57. Weinberg, G.V. Validity of whitening-matched filter approximation to the Pareto coherent detector. *IET Signal Process* **2012**, *6*, 546–550. [[CrossRef](#)]
58. Blanchet, J.; Glinn, P. Uniform renewal theory with applications to expansions of random geometric sums. *Adv. Appl. Prob.* **2007**, *39*, 1070–1097. [[CrossRef](#)]
59. Kallenberg, O. *Foundations of Modern Probability*; Springer: New York, NY, USA, 1997.
60. Kingman, J.F.C. On queues in heavy traffic. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1962**, *24*, 383–392. [[CrossRef](#)]
61. Su, Z.; Wang, X. Approximation of sums of locally dependent random variables via perturbation of Stein operator. *arXiv* **2022**, arXiv:2209.09770.v2.
62. Korolev, V.Y.; Zeifman, A.I. Convergence of statistics constructed from samples with random sizes to the Linnik and Mittag-Leffler distributions and their generalizations. *J. Korean Stat. Soc.* **2017**, *46*, 161–181. [[CrossRef](#)]
63. Bulinski, A.; Kozhevin, A. New version of the MDR method for stratified samples. *Stat. Optim. Inf. Comput.* **2017**, *5*, 1–18. [[CrossRef](#)]
64. Ginag, L.T.; Hung, T.L. An extension of random summations of independent and identically distributed random variables. *Commun. Korean Math. Soc.* **2018**, *33*, 605–618. [[CrossRef](#)]
65. Farago, A. Decomposition of Random Sequences into Mixtures of Simpler Ones and Its Application in Network Analysis. *Algorithms* **2021**, *14*, 336. [[CrossRef](#)]
66. Feng, M.; Deng, L.-J.; Chen, F.; Perc, M.; Kurths, J. The accumulative law and its probability model: An extension of the Pareto distribution and the log-normal distribution. *Proc. R. Soc. A* **2020**, *476*, 20200019. [[CrossRef](#)] [[PubMed](#)]
67. Nikolsky, S.M. *A Course of Mathematical Analysis, v. 1*; Mir Publishers: Moscow, Russia, 1987.