

Article

A Generalized Equilibrium Transform with Application to Error Bounds in the Rényi Theorem with No Support Constraints

Irina Shevtsova ^{1,2,3,4,*}  and Mikhail Tselishchev ^{2,4*}¹ Department of Mathematics, School of Science, Hangzhou Dianzi University, Hangzhou 310018, China² Department of Mathematical Statistics, Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, GSP-1, 1-52 Leninskiye Gory, Moscow 119991, Russia³ Federal Research Center “Informatics and Control” of the Russian Academy of Sciences, Moscow 119333, Russia⁴ Moscow Center for Fundamental and Applied Mathematics, Moscow 119991, Russia

* Correspondence: ishevtsova@cs.msu.ru (I.S.); m.tselishchev@cs.msu.ru (M.T.)

Received: 18 March 2020; Accepted: 8 April 2020; Published: 13 April 2020



Abstract: We introduce a generalized stationary renewal distribution (also called the equilibrium transform) for arbitrary distributions with finite nonzero first moment and study its properties. In particular, we prove an optimal moment-type inequality for the Kantorovich distance between a distribution and its equilibrium transform. Using the introduced transform and Stein’s method, we investigate the rate of convergence in the Rényi theorem for the distributions of geometric sums of independent random variables with identical nonzero means and finite second moments without any constraints on their supports. We derive an upper bound for the Kantorovich distance between the normalized geometric random sum and the exponential distribution which has exact order of smallness as the expectation of the geometric number of summands tends to infinity. Moreover, we introduce the so-called asymptotically best constant and present its lower bound yielding the one for the Kantorovich distance under consideration. As a concluding remark, we provide an extension of the obtained estimates of the accuracy of the exponential approximation to non-geometric random sums of independent random variables with non-identical nonzero means.

Keywords: Rényi theorem; Kantorovich distance; zeta-metrics; Stein’s method; stationary renewal distribution; equilibrium transform; geometric random sum; characteristic function

1. Introduction

Let X_1, X_2, \dots be a sequence of independent and, for simplicity in this Introduction, identically distributed (i.i.d.) random variables (r.v.s) with $a := \mathbf{E}X_1 \neq 0$. Let N be a random variable independent of $\{X_1, X_2, \dots\}$ and having the geometric distribution $\text{Geom}(p)$ with parameter $p \in (0, 1)$, i.e., $\mathbf{P}(N = n) = p(1 - p)^{n-1}$ for $n \in \mathbb{N}$. Denote also $N_0 := N - 1$ the shifted geometric r.v. Let $S_n := \sum_{k=1}^n X_k$, $n \in \mathbb{N}$, $S_0 := 0$. The well-known Rényi theorem states that the distribution of a properly normalized geometric random sum S_N converges weakly to the exponential law as p tends to zero. More precisely,

$$W := \frac{S_N}{\mathbf{E}S_N} \xrightarrow{d} \mathcal{E} \text{ as } p \downarrow 0, \quad \text{where } \mathcal{E} \sim \text{Exp}(1) \text{ and } \mathbf{E}S_N = \mathbf{E}N \mathbf{E}X_1 = a/p. \quad (1)$$

Here, the notation $\text{Exp}(\lambda)$ stands for the exponential distribution with density $\lambda e^{-\lambda x} \mathbb{1}_{(0, \infty)}(x)$, $\lambda > 0$. Originally, Rényi proved Equation (1) under the additional assumption of nonnegativeness of

$\{X_n\}$. However, it can be made sure that Equation (1) holds also: (i) for alternating $\{X_n\}$ (by alternating r.v. we mean a r.v. that may take values of both signs); and (ii) for

$$W_0 := \frac{S_{N_0}}{\mathbf{E}S_{N_0}} = \frac{pS_{N_0}}{a(1-p)},$$

in place of W (still without any support assumptions on the distribution of $\{X_k\}$). This can be done, for example, by showing that the characteristic function (ch.f.) of W (and also of W_0) converges pointwisely to that of the exponential distribution.

The importance of every limit theorem only increases if it is accompanied by the corresponding estimates of the rate of convergence. There are several bounds on the accuracy of approximation in Equation (1), mainly w.r.t. the Kolmogorov (uniform) and ζ -metrics, which are cited below. All of them assume additional conditions on the distribution of random summands including the finiteness of higher-order moments.

Recall that both the Kolmogorov and ζ_s -metrics are defined as simple probability metrics with ζ -structure (see Section 2 of [1]) between probability distributions (d.f.s F, G) of r.v.s X, Y :

$$\zeta_{\mathcal{H}}(F, G) \equiv \zeta_{\mathcal{H}}(\mathcal{L}(X), \mathcal{L}(Y)) \equiv \zeta_{\mathcal{H}}(X, Y) := \sup_{h \in \mathcal{H}} \left| \int_{\mathbb{R}} h dF - \int_{\mathbb{R}} h dG \right| \tag{2}$$

for specific classes \mathcal{H} of real Borel functions on \mathbb{R} (to simplify the notation, here and in what follows, we use r.v.s as well as their distributions and d.f.s in the arguments of simple probability metrics interchangeably; this should not cause any misunderstanding). The Kolmogorov metric ρ is obtained with $\mathcal{H} = \{ \mathbb{1}_{(-\infty, a)}(x) \mid a \in \mathbb{R} \}$, the class of indicators of all open intervals with unbounded left endpoint:

$$\rho(F, G) := \sup_{x \in \mathbb{R}} |F(x) - G(x)|,$$

while ζ -metric of order $s > 0$, originally introduced by Zolotarev [2] (see also [3]) as an example of an ideal metric with ζ -structure, is defined as $\zeta_{\mathcal{H}}$ with $\mathcal{H} = \mathcal{F}_s^\infty$, where

$$\mathcal{F}_s^\infty := \{ h \in \mathcal{F}_s : h \text{ is bounded} \},$$

$$\mathcal{F}_s := \{ h: \mathbb{R} \rightarrow \mathbb{R} : |h^{(m)}(x) - h^{(m)}(y)| \leq |x - y|^{s-m} \quad \forall x, y \in \mathbb{R} \text{ with } m := \lceil s - 1 \rceil \in \mathbb{N}_0 \}, \quad s > 0,$$

that is,

$$\zeta_s(F, G) := \sup_{h \in \mathcal{F}_s^\infty} \left| \int_{\mathbb{R}} h dF - \int_{\mathbb{R}} h dG \right|. \tag{3}$$

Observe that $h \in \mathcal{F}_s$ iff $h' \in \mathcal{F}_{s-1}$, $s > 1$. If $\mathbf{E}|X|^s < \infty$ and $\mathbf{E}|Y|^s < \infty$, then $\zeta_s(F, G) < \infty$ and the least upper bound w.r.t. to $h \in \mathcal{F}_s^\infty$ in Equation (3) may be replaced with that over a wider class \mathcal{F}_s . For further properties of ζ_s -metrics, we refer to the works in [3,4] and Section 4 of [5].

In the present paper, we focus mostly on ζ_1 -metrics between distributions with finite first moments; under this assumption, the definition of ζ_1 -metric can be rewritten as

$$\zeta_1(F, G) = \sup_{h \in \text{Lip}_1} \left| \int_{\mathbb{R}} h dF - \int_{\mathbb{R}} h dG \right|, \tag{4}$$

where

$$\text{Lip}_c := \left\{ h: \mathbb{R} \rightarrow \mathbb{R} \mid |h(x) - h(y)| \leq c|x - y| \quad \forall x, y \in \mathbb{R} \right\}, \quad c > 0,$$

so that $\text{Lip}_1 = \mathcal{F}_1$. It is worth noting that ζ_1 has several alternative representations. The Kantorovich–Rubinstein theorem states that $\zeta_1(X, Y)$ is minimal with respect to the compound metric $\mathbf{E}|X - Y|$, while the results in [6] imply that the optimal coupling is attained at the comonotonic

pair (that is, with $(X, Y) = (F^{-1}(U), G^{-1}(U))$, U having the uniform distribution on $(0, 1)$, F^{-1}, G^{-1} being generalized inverse d.f.s):

$$\zeta_1(F, G) = \min_{\mathcal{L}(X', Y'): X' \stackrel{d}{=} X, Y' \stackrel{d}{=} Y} \mathbf{E}|X' - Y'| = \int_0^1 |F^{-1}(u) - G^{-1}(u)| du = \int_{-\infty}^{\infty} |F(x) - G(x)| dx. \quad (5)$$

The rightmost representation in Equation (5), as the mean metric between the d.f.s F and G , follows from the geometrical interpretation. The metric ζ_1 is also called the Kantorovich, or the Wasserstein distance.

Thus, coming back to the convergence rate estimates in Equation (1), we first mention the paper by Solov'yev [7], which gives the following uniform bound for nonnegative $\{X_k\}$, as pointed out in [8]:

$$\rho(W_0, \mathcal{E}) \leq 24p \frac{\gamma_r}{r-2}, \quad 2 < r \leq 3, \quad (6)$$

where $\gamma_r = (\mathbf{E}X_1^r/a^r)^{1/(r-1)}$.

Kalashnikov and Vsekhsvyatskii [9] proved a uniform upper bound for nonnegative summands in terms of their moments of order $s \in (1, 2]$:

$$\rho(W, \mathcal{E}) \leq Cp^{s-1} \frac{\mathbf{E}X_1^s}{a^s}, \quad (7)$$

where C is an absolute constant.

Kruglov and Korolev [10] gave the following nonuniform bound of the accuracy of the exponential approximation to the normalized geometric distribution (i.e., for degenerate $\{X_n\}$):

$$\begin{aligned} |\mathbf{P}(pN < x) - (1 - e^{-x})| &\leq x \mathbb{1}_{\{x < p\}} + (e^{-x} - e^{-Q(p)x}) \mathbb{1}_{\{x \geq p\}} \leq \\ &\leq x \left[\mathbb{1}_{\{x < p\}} + \frac{p}{2(1-p)} e^{-x} \mathbb{1}_{\{x \geq p\}} \right], \quad (8) \end{aligned}$$

where $Q(p) = (1 - p/2)/(1 - p)$.

Brown [8] proved an asymptotically exact (as $p \rightarrow 0$) upper bound for nonnegative summands, which does not require moments of order greater than two:

$$\rho(W_0, \mathcal{E}) \leq p \frac{\mathbf{E}X_1^2}{a^2} \max \left(1, \frac{1}{2(1-p)} \right). \quad (9)$$

Brown also showed that Equation (9) is tighter than Equation (6) for all $2 < r \leq 3$ and $p \in (0, 0.5]$. Moreover, Equation (9) can be treated as a specification of Equation (7) for $s = 2$ with a concrete value of C .

Sugakova [11] presented some bounds for the d.f. $F_{S_{N_0}}(t)$ for $t > 1$ using the characteristics of the renewal process built on top of independent and not necessary identically distributed alternating $\{X_n\}$ with identical means.

Kalashnikov [12] provided estimates of the rate of convergence in the Rényi theorem for i.i.d. alternating $\{X_n\}$ w.r.t. ζ_s -metrics of order $s \in [1, 2]$ and the uniform metric (the latter is done under the additional assumption of bounded density), in particular, for any $s \in (1, 2]$,

$$\zeta_s(W, \mathcal{E}) \leq p^{s-1} \zeta_s(X_1, \mathcal{E}), \quad (10)$$

$$\zeta_1(W, \mathcal{E}) \leq p \zeta_1(X_1, \mathcal{E}) + 2(1-p)p^{s-1} \zeta_s(X_1, \mathcal{E}), \quad (11)$$

provided that $\mathbf{E}X_1 = 1$.

Among other valuable things, Peköz and Röllin [13] exploited Stein’s method and equilibrium (stationary renewal) distributions (see Section 3) to estimate the Kantorovich distance between the exponential distribution and that of a normalized geometric random sum W of square integrable independent and not necessary identically distributed nonnegative random summands $\{X_n\}$ with identical positive means under the technical assumption $\mathbf{E}X_k = 1$:

$$\zeta_1(W, \mathcal{E}) \leq 2p \sum_{n=1}^{\infty} \mathbf{P}(N = n) \zeta_1(X_n, X_n^e), \tag{12}$$

where X_n^e has an equilibrium distribution w.r.t. $X_n, n \in \mathbb{N}$. Using the trivial bound $\zeta_1(X, Y) \leq \mathbf{E}|X| + \mathbf{E}|Y|$ that follows from representation (5) and holds true for arbitrary r.v.s X, Y with finite first moments, the inequality in Equation (12) can be naturally extended to

$$\zeta_1(W, \mathcal{E}) \leq 2p \sup_n \zeta_1(X_n, X_n^e) \leq p \sup_n (\mathbf{E}X_n^2 + 2), \tag{13}$$

as done in [14].

Equation (22) of Hung [15] gives the following bound for the Trotter distance between W and \mathcal{E} in the case of i.i.d. nonnegative summands $\{X_n\}$ with $\mathbf{E}X_1 = 1$:

$$d_T(W, \mathcal{E}; h) := \sup_{t \in \mathbb{R}} |\mathbf{E}h(W + t) - \mathbf{E}h(\mathcal{E} + t)| \leq p^{s-1} (\mathbf{E}X_1^2 + 3), \quad h \in \mathcal{F}_s^\infty, s \in (1, 2]. \tag{14}$$

Given that $\zeta_s(W, \mathcal{E}) = \sup_{h \in \mathcal{F}_s^\infty} d_T(W, \mathcal{E}; h)$, the estimate in Equation (14) may be rewritten as

$$\zeta_s(W, \mathcal{E}) \leq p^{s-1} (\mathbf{E}X_1^2 + 3) \quad \text{for } s \in (1, 2]. \tag{15}$$

To compare Equation (15) with Kalashnikov’s bound in Equation (10), observe that, by Theorem 1(i,c) below, the dual representation of $\zeta_s(X, Y)$ -metric as the minimal w.r.t. the compound metric $\mathbf{E}|X - Y|^s$ for $s \in (0, 1]$ (see, e.g., Corollary 5.2.2 of [4]), and, finally, Theorem 1(g) below, for $s \in (1, 2]$, we have

$$\begin{aligned} \zeta_s(X_1, \mathcal{E}) &= \zeta_{s-1}(X_1^e, \mathcal{E}^e) = \zeta_{s-1}(X_1^e, \mathcal{E}) = \inf_{\mathcal{L}(X, Y): X \stackrel{d}{=} X_1^e, Y \stackrel{d}{=} \mathcal{E}} \mathbf{E}|X - Y|^{s-1} \leq \\ &\leq \mathbf{E}|X_1^e - \mathcal{E}| + 1 \leq \mathbf{E}X_1^e + \mathbf{E}\mathcal{E} + 1 = \mathbf{E}X_1^2/2 + 2 < \mathbf{E}X_1^2 + 3, \end{aligned}$$

hence, Kalashnikov’s bound in Equation (10) is tighter than Equation (15).

Thus, most existing estimates of the rate of convergence in the Rényi theorem were obtained under the additional assumption of nonnegativeness of random summands $\{X_n\}$. However, there are many applications where geometric random sums appear with alternating random summands, for example, as profit-or-losses in financial mathematics, risk theory, queuing theory, etc. Hence, extensions of such sharp and natural estimates as Equations (9), (12), and (13), say, to the alternating random summands, would not only represent a theoretical interest, but can also be in great demand by various applications of probability theory.

In the present paper, we focus on ζ_1 -estimates, in particular, we extend bounds in Equations (12) and (13) to the alternating case. More precisely, in Theorem 4 below, we prove that, for square

integrable independent and not necessarily identically distributed random summands $\{X_n\}$ with identical nonzero means (for simplicity, equal to one), the following estimates hold:

$$\zeta_1(W, \mathcal{E}) \leq 2p \sum_{n=1}^{\infty} \mathbf{P}(N = n) \zeta_1(\mathcal{L}(X_n), \mathcal{L}^e(X_n)) \leq p \left(\mathbf{E}X_N^2 - 2 \mathbf{P}(X_N \leq 0) \right), \tag{16}$$

$$\zeta_1(W_0, \mathcal{E}) \leq \frac{2p}{1-p} \zeta_1(\delta_0, \mathcal{L}^e(X_N)) = \frac{p}{1-p} \mathbf{E}X_N^2, \tag{17}$$

where δ_0 is the Dirac measure concentrated in zero and $\mathcal{L}^e(X_n)$ is the equilibrium transform of $\mathcal{L}(X_n)$, which is a generalization of the equilibrium distribution introduced in Section 3 below and, generally speaking, is no more a probability measure (therefore, we write $\mathcal{L}^e(X_n)$ instead of $\mathcal{L}(X_n^e)$), but allows eliminating the support constraints on the distribution of X_n . The notion of the ζ_1 -metric between signed measures is introduced in Section 2 below and coincides with that of the ordinary ζ_1 -metric in case of probability measures. Thus, the intermediate estimate in Equation (16) coincides with estimate (12), but now also holds true for alternating random summands $\{X_n\}$. Furthermore, it can easily be seen that the right-hand side of Equation (16) does not exceed

$$p \sup_n \left(\mathbf{E}X_n^2 - 2 \mathbf{P}(X_n \leq 0) \right)$$

and, hence, is tighter than estimate (13) and does not require that $\{X_n\}$'s take only positive values. The comparison of estimates (16) and Kalashnikov's bound in Equation (11) with $s = 2$

$$\begin{aligned} \zeta_1(W, \mathcal{E}) &\leq p \zeta_1(X_1, \mathcal{E}) + 2p(1-p) \zeta_2(X_1, \mathcal{E}) = \\ &= p \zeta_1(X_1, \mathcal{E}) + 2p(1-p) \zeta_1(\mathcal{L}^e(X_1), \text{Exp}(1)) \end{aligned} \tag{18}$$

(for the equality here, see Theorem 1(i) below) is complicated in the general case, since, due to Theorem 3 below, the rightmost expression does not exceed

$$2p(2-p) \zeta_1(\mathcal{L}(X_1), \mathcal{L}^e(X_1)),$$

which is asymptotically twice greater than the intermediate expression in Equation (16), while the intermediate estimate in Equation (16), by the triangle inequality, yields the bound

$$\zeta_1(W, \mathcal{E}) \leq 2p \zeta_1(X_1, \mathcal{E}) + 2p \zeta_1(\mathcal{L}^e(X_1), \text{Exp}(1))$$

with the first term twice larger than that in Equation (18).

We use the same techniques and recipes as in [13]. First, we bound the left-hand side of Equation (16) from above with $\zeta_1(\mathcal{L}(W), \mathcal{L}^e(W))$ using Stein's method (see Theorem 3 in Section 4). Second, we estimate $\zeta_1(\mathcal{L}(W), \mathcal{L}^e(W))$ by the ζ_1 -distances between X_n and their equilibrium transforms $\mathcal{L}^e(X_n)$, $n \in \mathbb{N}$. Third, we construct an optimal upper bound for $\zeta_1(\mathcal{L}(X_n), \mathcal{L}^e(X_n))$ in terms of the second moments of X_n and $\mathbf{P}(X_n \leq 0)$, $n \in \mathbb{N}$ (see Theorem 2 in Section 3). The resulting upper bounds for $\zeta_1(W, \mathcal{E})$ and $\zeta_1(W_0, \mathcal{E})$ are given in Theorem 4 of Section 5. Furthermore, we provide asymptotic lower bounds for $\zeta_1(W, \mathcal{E})$ and $\zeta_1(W_0, \mathcal{E})$ (see Theorem 5 in Section 5) in terms of the so-called *asymptotically best constants* introduced in Section 5. The constructed lower bounds turn out to be asymptotically four times smaller than the upper ones. Finally, we extend the obtained estimates of the accuracy of the exponential approximation to non-geometric random sums of independent random variables with non-identical nonzero means of identical signs (see Theorem 6 in Section 5).

2. The Kantorovich Distance between Signed Measures

In the next sections, we need to calculate the Kantorovich (or ζ_1 -) distance between measures on $(\mathbb{R}, \mathcal{B})$ that are no longer probabilities, but still have unit mass on \mathbb{R} . Denote by \mathcal{M}^1 the linear space of signed measures on $(\mathbb{R}, \mathcal{B})$ with finite total variations and finite first moments, and by \mathcal{M}_0^1 the subspace of measures $\sigma \in \mathcal{M}^1$ with $\sigma(\mathbb{R}) = 0$.

The Kantorovich norm on \mathcal{M}_0^1 is defined as (see Section 3.2 of [16])

$$\|\sigma\|_K := \sup_{f \in \text{Lip}_1} \left| \int_{\mathbb{R}} f d\sigma \right|.$$

Now let $\mu, \nu \in \mathcal{M}^1$ and $\mu(\mathbb{R}) = \nu(\mathbb{R})$, so that $\mu - \nu \in \mathcal{M}_0^1$. The induced Kantorovich distance ζ_1 between μ and ν is

$$\zeta_1(\mu, \nu) := \|\mu - \nu\|_K = \sup_{f \in \text{Lip}_1} \left| \int_{\mathbb{R}} f d\mu - \int_{\mathbb{R}} f d\nu \right|. \tag{19}$$

It is easy to see that in the case of probability measures μ and ν Equation (19) coincides with the definition of ζ_1 -distance given in Equation (4).

Using the Jordan decompositions $\mu = \mu^+ - \mu^-$ and $\nu = \nu^+ - \nu^-$, as well as the alternative representation in Equation (5) of the ζ_1 -distance between nonnegative measures $\lambda = \mu^+ + \nu^-$ and $\pi = \nu^+ + \mu^-$ with $\lambda(\mathbb{R}) = \pi(\mathbb{R})$ in terms of their d.f.s, after a proper normalization, one can rewrite Equation (19) as

$$\zeta_1(\mu, \nu) = \zeta_1(\lambda, \pi) = \int_{\mathbb{R}} |F_\lambda(x) - F_\pi(x)| dx = \int_{\mathbb{R}} |F_\mu(x) - F_\nu(x)| dx, \tag{20}$$

where $F_\mu(x) = \mu((-\infty, x))$, $F_\nu(x) = \nu((-\infty, x))$, $x \in \mathbb{R}$, are the d.f.s of the signed measures μ and ν , respectively. In other words, the alternative representation of Zolotarev’s ζ_1 -distance in terms of d.f.s in Equation (5) is preserved for signed measures with identical masses of \mathbb{R} .

We also use the convolution of signed measures $\mu * \lambda$, which is defined word-for-word as that of probability distributions. The uniqueness and multiplication theorems (see, e.g., Chapter 6 of [17] or Section 3.8 of [18]) state that the characteristic function of μ (the Fourier–Stieltjes transform of F_μ)

$$\widehat{\mu}(t) := \int_{\mathbb{R}} e^{itx} \mu(dx) = \int_{\mathbb{R}} e^{itx} dF_\mu(x), \quad t \in \mathbb{R},$$

defines the signed measure μ as well as its d.f. F_μ uniquely and

$$\widehat{\mu * \nu} = \widehat{\mu} \cdot \widehat{\nu}.$$

The following lemma, which is a simple corollary to representation (20), shows that the well-known properties of homogeneity and regularity of the Kantorovich distance between probability distributions are preserved for signed measures, but with a slight correction.

Lemma 1. *The Kantorovich distance ζ_1 on the space \mathcal{M}_D^1 of finite signed Borel measures on the real line with the masses of \mathbb{R} equal to $D \in \mathbb{R}$ and finite first moments possesses the following properties:*

- (a) **Homogeneity of order 1.** *For every $\mu, \nu \in \mathcal{M}_D^1$ and $c \neq 0$, with $\mu_c(B) := \mu(cB)$, $\nu_c(B) := \nu(cB)$ and $cB := \{cx \mid x \in B\}$, $B \in \mathcal{B}$, we have*

$$\zeta_1(\mu_c, \nu_c) = \frac{1}{|c|} \zeta_1(\mu, \nu).$$

(b) Regularity. For all $\mu, \nu \in \mathcal{M}_D^1$ and $\lambda \in \mathcal{M}^1$, we have

$$\zeta_1(\mu * \lambda, \nu * \lambda) \leq |\lambda|(\mathbb{R}) \cdot \zeta_1(\mu, \nu),$$

where $|\lambda| := \lambda^+ + \lambda^-$ is the total variation of λ .

To avoid abusing the notation, in what follows, we also use $\zeta_1(F, G)$ for the Kantorovich distance between (signed) measures uniquely restored (Section 3.5, Theorem 3.29 of [19]) from distribution functions F and G .

3. The Equilibrium Transform of Probability Distributions

The notion of *equilibrium distribution* w.r.t. nonnegative r.v.s with finite positive means originally arises in the renewal theory as the distribution of the initial delay of a renewal process which makes its renewal rate constant (Chapter 11, §4 of [20]) and, more generally, the renewal process stationary (Chapter 5, §4 of [21]), which is why it is also called the *stationary renewal distribution*. Equilibrium distribution appears also as the limit distribution of the residual waiting times, or hitting probabilities (Chapter 11, §4 of [20]) and in the celebrated Pollaczek–Khinchin–Beekman formula which expresses the ruin probability in the classical risk process in terms of geometric random sum of i.i.d. r.v.s whose common distribution is the equilibrium transform of the distributions of claims. Due to the definition given in a more general form in Equation (21) below, equilibrium distribution is also called the *integrated tail* one ([12], p. 37, [22]). Concerning the equilibrium transform, we would also like to mention the work of Harkness and Shantaram [23] who considered the *iterated equilibrium transform* for d.f.s with nonnegative support and investigated limit theorems for normalized iterations, the description of limit laws being given in [24]. In particular, the authors of [23] calculated the ch.f. of the equilibrium transform that can be used as the definition of the equilibrium transform in the general case and hence, with the inverse formula, can give a hint to definition in Equation (21) of the equilibrium d.f. with no support constraints.

We introduce an extension of the equilibrium distribution that is applicable for alternating random variables with finite nonzero first moments, but leads out of the class of probability distributions.

Let P be a probability measure with the d.f. $F(x) = P((-\infty, x))$, $x \in \mathbb{R}$, ch.f. $f(t) = \int_{\mathbb{R}} e^{itx} P(dx) = \int_{\mathbb{R}} e^{itx} dF(x)$, $t \in \mathbb{R}$, and a finite first moment $a := \int_{\mathbb{R}} xP(dx) = \int_{\mathbb{R}} x dF(x)$. If a r.v. X (on some probability space $(\Omega, \Sigma, \mathbf{P})$) has the distribution P , we also write $P = \mathcal{L}(X)$, $f(t) = \mathbf{E}e^{itX} =: f_X(t)$, $F(x) = \mathbf{P}(X < x) =: F_X(x)$, $a = \mathbf{E}X$.

Definition 1. The equilibrium d.f. (distribution) w.r.t. the d.f. F (probability distribution P / law $\mathcal{L}(X)$) with $a \neq 0$ is a function of bounded variation (a (signed) measure $P^e / \mathcal{L}^e(X)$ on $\mathcal{B}(\mathbb{R})$ with the d.f.)

$$F^e(x) := \begin{cases} -\frac{1}{a} \int_{-\infty}^x F(y) dy, & \text{if } x \leq 0, \\ -\frac{\mathbf{E}X^-}{a} + \frac{1}{a} \int_0^x (1 - F(y)) dy, & \text{if } x > 0, \end{cases} \tag{21}$$

$$= \frac{1}{a} \left(x^+ - \int_{-\infty}^x F(y) dy \right), \quad x \in \mathbb{R}. \tag{22}$$

In Theorem 1(a) below, it is proved that F^e , indeed, has bounded variation and some useful properties of the equilibrium transform are stated as well.

We call $F^e / P^e / \mathcal{L}^e(X)$ the equilibrium transform (d.f./distribution) w.r.t. $F / P / \mathcal{L}(X) / X$ correspondingly, although it may not be a probability d.f./distribution at all. At the same time, it can be easily seen that $\mathcal{L}^e(X)$ is a probability measure if and only if X does not change sign (that is, if and only if P is concentrated either on $(-\infty, 0]$ or on $[0, \infty)$), in which case one might construct a random variable X^e with the distribution $\mathcal{L}(X^e) = \mathcal{L}^e(X)$ and such that X and X^e are either both nonnegative or both nonpositive.

In what follows, to indicate the r.v. whose equilibrium transform is considered, we use the corresponding lower index and write F_X^e and f_X^e for $(F_X)^e$ and $(f_X)^e$, respectively.

Theorem 1. Let X be a r.v. with the d.f. F and $a \neq 0$, and F^e be the equilibrium d.f. w.r.t. F defined in Equation (21). Then:

- (a) **Absolute continuity.** The function F^e has bounded variation on \mathbb{R} with

$$|\mathcal{L}^e(X)|(\mathbb{R}) = \mathbf{E}|X|/|\mathbf{E}X|, \quad F^e(-\infty) = 0, \quad F^e(+\infty) = 1,$$

and, hence, $\mathcal{L}^e(X)$ is a Borel measure with unit on \mathbb{R} ; moreover, F^e is a.c. with the Lebesgue derivative

$$p^e(x) = \begin{cases} -\frac{1}{a}F(x), & \text{if } x \leq 0, \\ \frac{1}{a}(1 - F(x)), & \text{if } x > 0, \end{cases} \tag{23}$$

and $\text{supp } \mathcal{L}^e(X)$ coincides with the convex hull of $\text{supp } \mathcal{L}(X)$.

- (b) **Characteristic function.** The ch.f. (Fourier–Stieltjes transform) of F^e has the form

$$f^e(t) := \int_{\mathbb{R}} e^{itx} dF^e(x) = \frac{f(t) - 1}{tf'(0)} = \frac{f(t) - 1}{ita}, \quad \text{if } t \neq 0, \quad \text{and } f^e(0) = 1. \tag{24}$$

- (c) **Fixed points.** $\mathcal{L}^e(X) = \mathcal{L}(X)$ iff $X \sim \text{Exp}(1/a)$, that is, if and only if $F(x) = (1 - e^{-x/a})\mathbb{1}_{(0,\infty)}(x)$ for some $a > 0$.

- (d) **Test functions.** F^e is the equilibrium d.f. w.r.t. X if and only if

$$\mathbf{E}g(X) - g(0) = \mathbf{E}X \cdot \int_{\mathbb{R}} g'(x) dF^e(x) \tag{25}$$

for all Lipschitz functions $g: \mathbb{R} \rightarrow \mathbb{R}$.

- (e) **Mixture preservation.** For arbitrary d.f.s F_1, F_2, \dots with identical nonzero expectations and a discrete probability distribution $p_n \geq 0, n \in \mathbb{N}, \sum_{n=1}^{\infty} p_n = 1$, we have

$$\left(\sum_{n=1}^{\infty} p_n F_n \right)^e = \sum_{n=1}^{\infty} p_n F_n^e. \tag{26}$$

- (f) **Homogeneity.** For all $c \in \mathbb{R} \setminus \{0\}$, we have

$$(F_{cX})^e(x) = F_X^e(x/c), \quad x \in \mathbb{R}, \tag{27}$$

or, in terms of (constant-sign) r.v.s, $(cX)^e \stackrel{d}{=} cX^e, c \in \mathbb{R} \setminus \{0\}$. In other words, equilibrium transform respects scaling.

- (g) **Moments.** If $\mathbf{E}|X|^{r+1} < \infty$ for some $r > 0$, then for all $k \in \mathbb{N} \cap [1, r]$ we have

$$\int_{\mathbb{R}} x^k dF^e(x) = \frac{\mathbf{E}X^{k+1}}{(k+1)\mathbf{E}X}, \quad \int_{\mathbb{R}} |x|^r dF^e(x) = \frac{\mathbf{E}X|X|^r}{(r+1)\mathbf{E}X}, \tag{28}$$

$$\int_{\mathbb{R}} x^k |dF^e|(x) = \frac{\mathbf{E}|X|X^k}{(k+1)|\mathbf{E}X|}, \quad \int_{\mathbb{R}} |x|^r |dF^e|(x) = \frac{\mathbf{E}|X|^{r+1}}{(k+1)|\mathbf{E}X|}. \tag{29}$$

- (h) **Single summand property.** Let N, X_1, X_2, \dots be independent r.v.s, such that $a_n := \mathbf{E}X_n \in (0, \infty)$, $n \in \mathbb{N}$, $\mathbf{P}(N \in \mathbb{N}_0) = 1$, $S_N := X_1 + \dots + X_N$, $S_0 := 0$, $A := \mathbf{E}S_N = \sum_{n=1}^{\infty} a_n \mathbf{P}(N \geq n)$ be finite, and M be a \mathbb{N} -valued r.v. with the distribution

$$\mathbf{P}(M = m) = \frac{a_m}{A} \mathbf{P}(N \geq m), \quad m \in \mathbb{N}.$$

Then,

$$\mathcal{L}^e(S_N) = \sum_{m=1}^{\infty} \mathbf{P}(M = m) \mathcal{L}(S_{m-1}) * \mathcal{L}^e(X_m), \tag{30}$$

where $*$ denotes the convolution of two Borel measures, or, in terms of (constant-sign) r.v.s,

$$S_N^e \stackrel{d}{=} S_{M-1} + X_M^e,$$

where all the r.v.s are independent. In particular, if $N \sim \text{Geom}(p)$ and all X_k 's have identical nonzero expectations, then $M \stackrel{d}{=} N$ and

$$\mathcal{L}^e(S_N) = \mathcal{L}^e(S_{N-1}) = \sum_{n=1}^{\infty} p(1-p)^{n-1} \mathcal{L}(S_{n-1}) * \mathcal{L}^e(X_n), \tag{31}$$

which can be also rewritten, in the case of i.i.d. $\{X_k\}$, in the form

$$\mathcal{L}^e(S_N) = \mathcal{L}^e(S_{N-1}) = \mathcal{L}(S_{N-1}) * \mathcal{L}^e(X_1).$$

- (i) **Relation between ζ -distances.** For arbitrary d.f.s F and G with finite moments of order $s > 1$ and identical expectations $a \neq 0$, we have

$$\zeta_s(F, G) = |a| \zeta_{s-1}(F^e, G^e). \tag{32}$$

Theorem 2 below provides also an optimal upper bound for $\zeta_1(F, F^e)$ given $F(0+)$ and the second-order moment of F .

Remark 1. Theorem 1(h) shows that the equilibrium transform of the geometric random sum of independent r.v.s with identical nonzero means does not depend on whether or not one takes the geometric distribution starting from zero.

Let us make several historical remarks. Some of the properties of the equilibrium distribution stated in Theorem 1 were known for a nonnegative r.v. X . Thus, the characteristic function of X^e given in Equation (24) was found in [23], Equation (25) was taken as the definition of (the distribution of) X^e in [13,14]. In Theorem 2.1 of [13], it was proved that the exponential distribution is the only fixed point of the equilibrium transform; this fact is proved directly also in Lemma 5.2 of [14]. In [14] (p. 268), it is observed that $(cX)^e \stackrel{d}{=} cX^e$ for $c > 0$. Some moment calculations were given in [22]. Single summand property for S_N was demonstrated in the proof of Theorem 3.1 of [13] for nonnegative, but not necessarily independent $\{X_k\}$. The fact that $\mathcal{L}^e(S_N) = \mathcal{L}^e(S_{N-1})$ for i.i.d. nonnegative $\{X_k\}$ was observed in [8] (p. 1394). The equality in Equation (32) for $F(0) = G(0) = 0$ and $s = 2$ was stated in [12] (p. 37).

To prove Theorem 1, we require the following auxiliary statement.

Lemma 2. For every $n \in \mathbb{N}$ and $z_1, \dots, z_n \in \mathbb{C}$, we have

$$\prod_{k=1}^n z_k - 1 = \sum_{k=1}^n (z_k - 1) \prod_{j=1}^{k-1} z_j = \sum_{k=1}^n (z_k - 1) \prod_{j=k+1}^n z_j, \tag{33}$$

where $\prod_{j=a}^b (\cdot) := 1$ for $b < a$.

Proof. We use the induction w.r.t. n . For $n = 1$ Equation (33) is trivial. Let Equation (33) hold for $n = 1, \dots, m - 1$; let us prove it for $n = m$. Using the inductive transition in the second equality below, we get

$$\prod_{k=1}^m z_k - 1 = (z_m - 1) \prod_{k=1}^{m-1} z_k + \prod_{k=1}^{m-1} z_k - 1 = (z_m - 1) \prod_{k=1}^{m-1} z_k + \sum_{k=1}^{m-1} (z_k - 1) \prod_{j=1}^{k-1} z_j = \sum_{k=1}^m (z_k - 1) \prod_{j=1}^{k-1} z_j.$$

The second equality in Equation (33) can be deduced from the first one just by the re-numeration of $\{z_k\}_{k=1}^n : z_k \leftarrow z_{n-k+1}, k = 1, \dots, n$. \square

Proof of Theorem 1. (a) It follows immediately from the definition in Equation (21) of F^e that F^e is a.c. with the density given in Equation (23). In turn, Equation (23) implies that $\text{supp } \mathcal{L}^e(X)$ is the convex hull of $\text{supp } \mathcal{L}(X)$ and, accounting for $|\mathcal{L}^e(X)|(\mathbb{R}) = \int |p^e(x)| dx = \mathbf{E}|X|/|\mathbf{E}X| < \infty$, also that F^e has bounded variation. The limiting values $F^e(\pm\infty)$ can be found directly using the definition of F^e .

(b) Using the density of F^e (see Equation (23)) and integrating by parts, we have

$$\begin{aligned} f^e(t) &= \frac{1}{a} \int_{\mathbb{R}} e^{itx} p^e(x) dx = \frac{1}{a} \int_{\mathbb{R}} e^{itx} (\mathbb{1}_{(0,\infty)}(x) - F(x)) dx = \frac{1}{ita} \int_{\mathbb{R}} (\mathbb{1}_{(0,\infty)}(x) - F(x)) de^{itx} = \\ &= \frac{1}{ita} \left[-e^{itx} F(x) \Big|_{-\infty}^0 + e^{itx} (1 - F(x)) \Big|_0^{\infty} + \int_{\mathbb{R}} e^{itx} dF(x) \right] = \frac{f(t) - 1}{ita}, \end{aligned}$$

which coincides with Equation (24).

(c) This statement follows immediately due to the uniqueness of the solution to the linear equation

$$f^e(t) \equiv \frac{f(t) - 1}{ita} = f(t) \iff f(t) = \frac{1}{1 - ita} \sim \text{Exp}(1/a).$$

(d)–(g) These statements follow from the definition and integration by parts for (d) and (g) or the linearity of the Lebesgue–Stieltjes integral for (e).

(h) Let us denote $f_0(t) \equiv 1, f_k(t) = \mathbf{E}e^{itX_k}, k \in \mathbb{N}, t \in \mathbb{R}$. Using the fact that

$$f_{S_N}(t) = \sum_{n=0}^{\infty} \mathbf{P}(N = n) \mathbf{E}e^{itS_n} = \sum_{n=0}^{\infty} \mathbf{P}(N = n) \prod_{k=0}^n f_k(t),$$

together with the equation for the equilibrium ch.f. in Equations (24) and (33), we get

$$\begin{aligned} f_{S_N}^e(t) &= \frac{f_{S_N}(t) - 1}{t f_{S_N}'(0)} = \frac{1}{itA} \sum_{n=1}^{\infty} \mathbf{P}(N = n) \left(\prod_{k=1}^n f_k(t) - 1 \right) = \\ &= \sum_{n=1}^{\infty} \mathbf{P}(N = n) \sum_{k=1}^n \frac{f_k(t) - 1}{itA} \prod_{j=1}^{k-1} f_j(t) = \\ &= \sum_{n=1}^{\infty} \mathbf{P}(N = n) \sum_{k=1}^n \frac{a_k}{A} f_k^e(t) f_{S_{k-1}}(t). \end{aligned}$$

Changing the order of summation, which is possible by virtue of the absolute convergence of the above series, and recalling the definition of $\mathcal{L}(M)$, we obtain

$$f_{S_N}^e(t) = \sum_{k=1}^{\infty} f_k^e(t) f_{S_{k-1}}(t) \cdot \frac{a_k}{A} \sum_{n=k}^{\infty} \mathbf{P}(N = n) = \sum_{k=1}^{\infty} f_k^e(t) f_{S_{k-1}}(t) \mathbf{P}(M = k),$$

which is equivalent to Equation (30) by virtue of the uniqueness theorem.

If now $N \sim \text{Geom}(p)$ and $a_1 = a_2 = \dots = a$, then $A = a\mathbf{E}N = a/p$, $\mathbf{P}(M = k) = p(1 - p)^{k-1} = \mathbf{P}(N = k)$, $k \in \mathbb{N}$. Denoting by M_0 a r.v. corresponding to $N_0 := N - 1$ with the distribution

$$\mathbf{P}(M_0 = k) := a_k \mathbf{P}(N_0 \geq k) / \sum_{k=1}^{\infty} a_k \mathbf{P}(N_0 \geq k) = \mathbf{P}(N_0 \geq k) / \mathbf{E}N_0 = p(1 - p)^{k-1}, \quad k \in \mathbb{N},$$

we observe that $M_0 \stackrel{d}{=} N \stackrel{d}{=} M$. This proves Equation (31).

(i) This statement follows from Theorem 4.2(a), Equation (4.20) of [5]. It can also be proved independently, namely, by virtue of (d) we have

$$\begin{aligned} \zeta_s(F, G) &= \sup_{h \in \mathcal{F}_s} \left| \int_{\mathbb{R}} h dF - \int_{\mathbb{R}} h dG \right| = |a| \sup_{h \in \mathcal{F}_s} \left| \int_{\mathbb{R}} h' dF^e - \int_{\mathbb{R}} h' dG^e \right| = \\ &= |a| \sup_{h \in \mathcal{F}_{s-1}} \left| \int_{\mathbb{R}} h dF^e - \int_{\mathbb{R}} h dG^e \right| = |a| \zeta_{s-1}(F, G). \quad \square \end{aligned}$$

To conclude this section, we construct an optimal upper bound for the Kantorovich distance between an arbitrary probability distribution with nonzero mean and its equilibrium transform given its second moment and the mass of nonpositive axis. Before formulating the corresponding result, we have to note that Cantelli's (one-sided Chebyshev's) inequality yields $\mathbf{P}(X \leq 0) \leq 1 - 1/\mathbf{E}X^2$ for an arbitrary r.v. X with $0 < \mathbf{E}X^2 < \infty$, and, hence,

$$\mathbf{E}X^2 \geq \frac{1}{1 - \mathbf{P}(X \leq 0)}.$$

This remark explains the choice of the domain of parameters q and b in the following Theorem 2.

Theorem 2. Take any $q \in [0, 1)$ and $b \geq \frac{1}{\sqrt{1-q}}$ and let X be a square integrable r.v. with $\mathbf{E}X = 1$, $\mathbf{E}X^2 = b^2$, and $\mathbf{P}(X \leq 0) = q$. Then,

$$\zeta_1(\mathcal{L}(X), \mathcal{L}^e(X)) \leq \frac{b^2}{2} - q, \tag{34}$$

where $\mathcal{L}^e(X)$ is the equilibrium transform of $\mathcal{L}(X)$. The equality in Equation (34) is attained for every $q \in (0, 1)$ and $b \geq \frac{1}{\sqrt{1-q}}$ on the two-point distribution $\mathcal{L}(X) = q\delta_u + (1 - q)\delta_v$ with

$$u = 1 - \sqrt{\frac{1-q}{q}(b^2 - 1)}, \quad v = 1 + \sqrt{\frac{q}{1-q}(b^2 - 1)}, \tag{35}$$

and for $q = 0$ and $b = 1$ on the degenerate distribution $\mathcal{L}(X) = \delta_1$.

Remark 2. With the account of Theorem 1(f) and Lemma 1(a), for arbitrary $\mathbf{E}X \neq 0$, Equation (34) takes the form

$$\zeta_1(\mathcal{L}(X), \mathcal{L}^e(X)) \leq \frac{1}{2} \cdot \frac{\mathbf{E}X^2}{|\mathbf{E}X|} - |\mathbf{E}X| \cdot \mathbf{P}(X \leq 0).$$

Proof of Theorem 2. Let F be the d.f. of X and F^e be its equilibrium transform. Consider the following functional on the space \mathbf{F} of probability d.f.s with unit mean and finite second moment:

$$J(F) = \zeta_1(F, F^e) - \frac{1}{2} \int_{\mathbb{R}} x^2 dF(x) + F(0+), \quad F \in \mathbf{F}. \tag{36}$$

Then, Equation (34) would follow from

$$\sup_{F \in \mathbf{F}} J(F) \leq 0. \tag{37}$$

Let us prove Equation (37).

Since $h \in \text{Lip}_1$ if and only if $(-h) \in \text{Lip}_1$, the modulus sign in the definition of $\zeta_1(F, F^e)$ (see Equation (19)) may be omitted. Hence, we can rewrite

$$J(F) = \sup_{h \in \text{Lip}_1} J_1(F, h), \quad \text{where} \quad J_1(F, h) = \int_{\mathbb{R}} h dF - \int_{\mathbb{R}} h dF^e - \frac{1}{2} \int_{\mathbb{R}} x^2 dF(x) + F(+0), \quad F \in \mathbf{F}.$$

Note that $J_1(F, h)$ is linear w.r.t. $F \in \mathbf{F}$ for every $h \in \text{Lip}_1$, by definition. According to Theorems 2 and 3 of [25], for any fixed $h \in \text{Lip}_1$, the least upper bound $\sup_{F \in \mathbf{F}} J_1(F, h)$ w.r.t. probability d.f. F satisfying two linear conditions (we can also fix the value $b^2 \geq 1$ of the second moment and then take the least upper bound w.r.t. all $b \geq 1$) coincides with that over the set of three-point distributions from \mathbf{F} . Since every three-point distribution has finite moments of all orders, the condition of finiteness of the second-order moments may be eliminated, so that

$$\sup_{F \in \mathbf{F}} J(F) = \sup_{h \in \text{Lip}_1} \sup_{F \in \mathbf{F}_3} J_1(F, h),$$

where \mathbf{F}_3 is the space of all discrete probability d.f.s with at most three jumps and unit first moment. Furthermore, according to Hoeffding [26], the least upper bound $\sup_{F \in \mathbf{F}_3} J_1(F, h)$ w.r.t. discrete probability d.f.s F with finite number of jumps and satisfying one moment condition is attained on two-point distributions, hence,

$$\sup_{F \in \mathbf{F}} J(F) = \sup_{h \in \text{Lip}_1} \sup_{F \in \mathbf{F}_2} J_1(F, h) = \sup_{F \in \mathbf{F}_2} J(F),$$

where \mathbf{F}_2 is the space of all discrete probability d.f.s with at most two jumps and unit first moment. Therefore, to prove Equation (37), it suffices to show that $J(F) \leq 0$ for every $F \in \mathbf{F}_2$

Let F correspond to a two-point distribution $p\delta_u + (1-p)\delta_v$ with $u < v$ and $p \in [0, 1)$. The condition $\int_{\mathbb{R}} x dF(x) = 1$ yields $u < 1 \leq v$ and $v = (1 - pu)/(1 - p)$, so that there are only three possibilities:

Case 1: $u \leq 0 < 1 \leq v$ and $p \in [0, 1)$. Then,

$$q = \mathbf{P}(X \leq 0) = p, \quad b^2 = \mathbf{E}X^2 = \frac{pu^2 - 2pu + 1}{1 - p}, \tag{38}$$

and, by definition of F^e given in Equation (21), we have

$$F^e(x) = \begin{cases} 0, & \text{for } x \leq u, \\ pu - px, & \text{for } u < x \leq 0, \\ pu + (1 - p)x, & \text{for } 0 < x \leq v, \\ 1, & \text{for } x > v. \end{cases}$$

Observing that the difference $F(x) - F^e(x)$ has exactly one sign change at $x = p(1 - u)/(1 - p) = v - 1 \in [0, v)$ and using Equation (20), after some elementary calculations, we get

$$\zeta_1(F, F^e) = \frac{1}{2}u^2p - up + \frac{1}{2}(1 - u)^2 \frac{p^2}{1 - p} + \frac{1}{2}(1 - p) \cdot 1,$$

and, hence,

$$J(F) = \zeta_1(F, F^e) - \frac{pu^2 - 2pu + 1}{2(1 - p)} + p = 0,$$

which means that $J(F) = 0$ for arbitrary two-point probability distribution with unit first moment and a nonpositive atom. Expressing u and v in terms of q and b^2 (see Equation (38)), we get Equation (35).

Case 2: $0 < u < 1 \leq v$ and $p \in [0, u]$. Then, $q = \mathbf{P}(X \leq 0) = 0$,

$$F^e(x) = \begin{cases} 0, & \text{for } x \leq 0, \\ x, & \text{for } 0 < x \leq u, \\ u + (1 - p)(x - u), & \text{for } u < x \leq v, \\ 1, & \text{for } x > v, \end{cases}$$

and by $F^e(x) - F(x) \geq 0$ for all $x \in \mathbb{R}$, we get $\zeta_1(F, F^e) = \frac{1}{2}u^2 + \frac{1}{2}(v - u)(u + 1 - 2p) = 1 - \frac{1}{2}\mathbf{E}X^2$. Hence,

$$J(F) = \zeta_1(F, F^e) - \frac{1}{2}\mathbf{E}X^2 + q = 1 - \mathbf{E}X^2 \leq 0,$$

since $\mathbf{E}X^2 \geq (\mathbf{E}X)^2 = 1$ by Jensen’s inequality. The equality here and, hence, in Equation (34) is attained in the case of degenerate distribution δ_1 .

Case 3: $0 < u < 1 < v$ and $p \in (u, 1)$. Then, $q = 0$ and F^e has the same form as in the previous case, but the function $F^e(x) - F(x)$ now has exactly one sign change at $x = p(1 - u)/(1 - p) = v - 1 \in (u, v)$, and, hence, $\zeta_1(F, F^e) = \frac{1}{2}u^2 + \frac{1}{2}(p - u)^2 \frac{1}{1 - p} + \frac{1}{2}(1 - p) \cdot 1$. Thus,

$$J(F) = \zeta_1(F, F^e) - \frac{1}{2}\mathbf{E}X^2 + q = u^2 - p < 0,$$

since $u^2 < u < p$ in this case, and the equality in Equation (37) (and, hence, in Equation (34)) is not attained. \square

Remark 3. Analyzing the proof, one can make sure that Equation (34) admits a slight improvement:

$$\zeta_1(\mathcal{L}(X), \mathcal{L}^e(X)) \leq \frac{\mathbf{E}X^2}{2} - \mathbf{P}(X \leq 0) - \mathbf{E}(1 - X)^2 \mathbb{1}_{(0,1]}(X)$$

for any r.v. X with $\mathbf{E}X = 1$ and finite second moment. The proof differs only by the appearance (subtraction) of an additional term $\int_{(0,1]} (1 - x)^2 dF(x)$ in definition in Equation (36) of $J(F)$, which is still linear w.r.t. F , and, hence, does not change the logic. One has only to check that the new $J(F)$ is nonpositive for two-point distributions. In Case 1, $J(F)$ is retained. In Cases 2 and 3, the additional term is of the form $p(1 - u)^2$ and it can be made sure that this term does not affect the sign of $J(F)$.

4. Stein’s Method

Stein’s method, first introduced in [27] for normal approximation, is a powerful technique that allows to estimate distances with ζ -structure (see Equation (2)) between probability distributions and a fixed target distribution (of a r.v.) Z . A complete survey on Stein’s method may be found, e.g., in [14]. Suppose that the distance $\zeta_{\mathcal{H}}$ is of the form given in Equation (2) for a specific class \mathcal{H} of real-valued

functions. As mentioned in the Introduction, this is the case for both uniform (Kolmogorov) and Kantorovich distances with $\mathcal{H} = \{ \mathbb{1}_{(-\infty, a)}(\cdot) \mid a \in \mathbb{R} \}$ and $\mathcal{H} = \text{Lip}_1$, respectively.

The first step of Stein’s method is to construct the so-called *Stein operator* \mathcal{A} in some space \mathcal{F} of real functions, such that

$$\mathbf{E}\mathcal{A}f(Z) = 0 \quad \forall f \in \mathcal{F}. \tag{39}$$

The second step is to find the solution f_h to the *Stein equation*

$$\mathcal{A}f_h(x) = h(x) - \mathbf{E}h(Z) \tag{40}$$

for every $h \in \mathcal{H}$. Once the solution is found, it becomes possible to estimate the distance between the distributions of X and Z as

$$\begin{aligned} \zeta_{\mathcal{H}}(X, Z) &= \sup_{h \in \mathcal{H}} \left| \int_{\mathbb{R}} h dF_X - \int_{\mathbb{R}} h dF_Z \right| = \sup_{h \in \mathcal{H}} \left| \int_{\mathbb{R}} h dF_X - \mathbf{E}h(Z) \right| = \\ &= \sup_{h \in \mathcal{H}} \left| \int_{\mathbb{R}} (h - \mathbf{E}h(Z)) dF_X \right| = \sup_{h \in \mathcal{H}} \left| \int_{\mathbb{R}} \mathcal{A}f_h dF_X \right| = \sup_{h \in \mathcal{H}} |\mathbf{E}\mathcal{A}f_h(X)|. \end{aligned} \tag{41}$$

The final estimate for $\zeta_{\mathcal{H}}(X, Z)$ is usually derived by bounding the latest expression in Equation (41) from above using the properties of the Stein operator \mathcal{A} and those of the solutions f_h to the Stein Equation (40).

It can be made sure that for $Z \stackrel{d}{=} \mathcal{E} \sim \text{Exp}(1)$ the following operator satisfies Equation (39) on the space \mathcal{F} of absolutely continuous functions with $\mathbf{E}|f'(\mathcal{E})| < +\infty$ and thus appears to be the Stein operator:

$$\mathcal{A}f(x) = f'(x) - f(x) + f(0). \tag{42}$$

Peköz and Röllin [13] found an explicit solution to Stein Equation (40) in this case:

$$f_h(x) = -e^x \int_x^{+\infty} \tilde{h}(t)e^{-t} dt, \quad \text{where } \tilde{h}(t) = h(t) - \mathbf{E}h(\mathcal{E}), \tag{43}$$

for every h with $\mathbf{E}|h(\mathcal{E})| < \infty$. Note that $f_h(0) = 0$.

The following theorem extends results of Peköz and Röllin [13] in Theorem 2.1 to distributions with no support constraints and provides estimates of the accuracy of the exponential approximation in terms of the Kantorovich distance characterizing the proximity to the equilibrium transform.

Theorem 3. *Let X be a square integrable r.v. with $\mathbf{E}X = 1$ and $\mathcal{E} \sim \text{Exp}(1)$. Then,*

$$\begin{aligned} \zeta_1(X, \mathcal{E}) &\leq 2 \zeta_1(\mathcal{L}(X), \mathcal{L}^e(X)), \\ \zeta_1(\mathcal{L}^e(X), \text{Exp}(1)) &\leq \zeta_1(\mathcal{L}(X), \mathcal{L}^e(X)), \end{aligned}$$

where $\mathcal{L}^e(X)$ is the equilibrium transform of $\mathcal{L}(X)$.

Proof. Let f_h be defined by Equation (43). Then, by Equations (41), (42), and (25), we have

$$\zeta_1(X, \mathcal{E}) = \sup_{h \in \text{Lip}_1} |\mathbf{E}\mathcal{A}f_h(X)| = \sup_{h \in \text{Lip}_1} |\mathbf{E}f'_h(X) - \mathbf{E}f_h(X)| = \sup_{h \in \text{Lip}_1} \left| \int_{\mathbb{R}} f'_h dF_X - \int_{\mathbb{R}} f'_h dF_X^e \right|$$

and

$$\begin{aligned} \zeta_1(\mathcal{L}^e(X), \text{Exp}(1)) &= \sup_{h \in \text{Lip}_1} \left| \int_{\mathbb{R}} h(x) dF_X^e(x) - \mathbf{E}h(\mathcal{E}) \right| = \sup_{h \in \text{Lip}_1} \left| \int_{\mathbb{R}} \tilde{h}(x) dF_X^e(x) \right| = \\ &= \sup_{h \in \text{Lip}_1} \left| \int_{\mathbb{R}} \mathcal{A}f_h(x) dF_X^e(x) \right| = \sup_{h \in \text{Lip}_1} \left| \int_{\mathbb{R}} f'_h(x) dF_X^e(x) - \int_{\mathbb{R}} f_h(x) dF_X^e(x) \right| = \\ &= \sup_{h \in \text{Lip}_1} \left| \int_{\mathbb{R}} f_h(x) dF_X(x) - \int_{\mathbb{R}} f_h(x) dF_X^e(x) \right|. \end{aligned}$$

In Lemma 4.1 of [13] (see also Lemma 5.3 of [14]), it is proved that $f_h \in \text{Lip}_1$ and $f'_h \in \text{Lip}_2$ for $h \in \text{Lip}_1$. This remark together with the observation that $\mathcal{L}(X)$ and $\mathcal{L}^e(X)$ have finite first moments immediately leads to the statement of the theorem. \square

Less formally, Theorem 3 states that, if $\mathcal{L}(X)$ and $\mathcal{L}^e(X)$ are close, then so are $\mathcal{L}(X)$ and $\text{Exp}(1)$, and, hence, may be regarded as the continuity theorem to the fixed-point property stated in Theorem 1(c).

5. Main Results

Theorem 4. Let X_1, X_2, \dots be a sequence of independent square integrable random variables with $\mathbf{E}X_n = a \neq 0$ and $S_n := \sum_{i=1}^n X_i$ for $n \in \mathbb{N}$, $S_0 := 0$. Let $p \in (0, 1)$, $N \sim \text{Geom}(p)$, be independent of all $\{X_n\}$, $N_0 := N - 1$, and $W := S_N/\mathbf{E}S_N = pS_N/a$, $W_0 := S_{N_0}/\mathbf{E}S_{N_0} = pS_{N_0}/(a(1-p))$ be normalized geometric random sums, $\mathcal{E} \sim \text{Exp}(1)$. Then,

$$\zeta_1(W, \mathcal{E}) \leq \frac{2p}{|a|} \sum_{n=1}^{\infty} \mathbf{P}(N = n) \zeta_1(\mathcal{L}(X_n), \mathcal{L}^e(X_n)) \leq p \left(\frac{\mathbf{E}X_N^2}{a^2} - 2\mathbf{P}(X_N \leq 0) \right), \tag{44}$$

$$\zeta_1(W_0, \mathcal{E}) \leq \frac{p}{1-p} \cdot \frac{\mathbf{E}X_N^2}{a^2}. \tag{45}$$

Before proceeding to the proof, we need the following auxiliary statement.

Lemma 3. Under the conditions of Theorem 4, we have

$$\begin{aligned} \zeta_1(\mathcal{L}(S_N), \mathcal{L}^e(S_N)) &\leq \sum_{n=1}^{\infty} p(1-p)^{n-1} \zeta_1(\mathcal{L}(X_n), \mathcal{L}^e(X_n)), \\ \zeta_1(\mathcal{L}(S_{N_0}), \mathcal{L}^e(S_{N_0})) &\leq \frac{\mathbf{E}X_N^2}{2|a|}. \end{aligned}$$

Proof. Let F_n be the d.f. of X_n , $n \in \mathbb{N}$. Then, according to Equation (20), Theorem 1(h), Tonelli's theorem, and an obvious fact that $\mathcal{L}(S_n) = \mathcal{L}(S_{n-1}) * \mathcal{L}(X_n)$, we have

$$\begin{aligned} \zeta_1(\mathcal{L}(S_N), \mathcal{L}^e(S_N)) &= \int_{\mathbb{R}} |F_{S_N}(x) - F_{S_N}^e(x)| dx \leq \\ &\leq \sum_{n=1}^{\infty} p(1-p)^{n-1} \int_{\mathbb{R}} \int_{\mathbb{R}} |F_n(x-s) - F_n^e(x-s)| dF_{S_{n-1}}(s) dx = \\ &= \sum_{n=1}^{\infty} p(1-p)^{n-1} \int_{\mathbb{R}} \int_{\mathbb{R}} |F_n(x-s) - F_n^e(x-s)| dx dF_{S_{n-1}}(s) = \\ &= \sum_{n=1}^{\infty} p(1-p)^{n-1} \zeta_1(\mathcal{L}(X_n), \mathcal{L}^e(X_n)), \end{aligned}$$

which proves the first claim of the lemma, and, similarly,

$$\begin{aligned} \zeta_1(\mathcal{L}(S_{N_0}), \mathcal{L}^e(S_{N_0})) &\leq \sum_{n=1}^{\infty} p(1-p)^{n-1} \int_{\mathbb{R}} \int_{\mathbb{R}} |\mathbb{1}_{(0,+\infty)}(x-s) - F_n^e(x-s)| dF_{S_{n-1}}(s) dx = \\ &= \sum_{n=1}^{\infty} p(1-p)^{n-1} \int_{\mathbb{R}} |\mathbb{1}_{(0,+\infty)}(x) - F_n^e(x)| dx = \sum_{n=1}^{\infty} p(1-p)^{n-1} \zeta_1(\delta_0, \mathcal{L}^e(X_n)), \end{aligned}$$

where δ_0 denotes the Dirac delta-measure concentrated in 0. As can easily be seen from the definition of the equilibrium transform given in Equation (21),

$$\begin{aligned} \text{if } a > 0, \quad \text{then } F^e(x) &\leq 0, \quad x \leq 0, \quad F(x) \leq 1, \quad x \geq 0, \\ \text{if } a < 0, \quad \text{then } F^e(x) &\geq 0, \quad x \leq 0, \quad F(x) \geq 1, \quad x \geq 0, \end{aligned}$$

hence, we write

$$|\mathbb{1}_{(0,+\infty)}(x) - F_n^e(x)| = \begin{cases} F_n^e(x) \operatorname{sign} a, & x \leq 0, \\ (1 - F_n^e(x)) \operatorname{sign} a, & x \geq 0, \end{cases}$$

and also using Equation (28), we obtain

$$\zeta_1(\delta_0, \mathcal{L}^e(X_n)) = \operatorname{sign} a \cdot \left(- \int_{-\infty}^0 F_n^e(x) dx + \int_0^{+\infty} (1 - F_n^e(x)) dx \right) = \operatorname{sign} a \cdot \int_{\mathbb{R}} x dF_n^e(x) = \frac{\mathbf{E}X_n^2}{2|a|}.$$

The second claim of the lemma follows now by the total probability formula and independence conditions. \square

Proof of Theorem 4. Due to the homogeneity of both the Kantorovich metric (Lemma 1(a)) and the equilibrium transform (Theorem 1(f)), without loss of generality, we can assume that $a = 1$. The second inequality in Equation (44) is the implication of Theorem 2, thus it remains only to prove the first inequality in Equation (44) and the inequality in Equation (45). Indeed, by Theorems 3 and 1(f) and Lemmas 1 and 3, we have

$$\begin{aligned} \zeta_1(W, \mathcal{E}) &\leq 2 \zeta_1(\mathcal{L}(W), \mathcal{L}^e(W)) = \\ &= 2p \zeta_1(\mathcal{L}(S_N), \mathcal{L}^e(S_N)) \leq 2p \sum_{n=1}^{\infty} \mathbf{P}(N = n) \zeta_1(\mathcal{L}(X_n), \mathcal{L}^e(X_n)), \end{aligned}$$

and

$$\begin{aligned} \zeta_1(W_0, \mathcal{E}) &\leq 2 \zeta_1(\mathcal{L}(W_0), \mathcal{L}^e(W_0)) = \\ &= \frac{2p}{1-p} \zeta_1(\mathcal{L}(S_{N_0}), \mathcal{L}^e(S_{N_0})) \leq \frac{p}{1-p} \mathbf{E}X_N^2. \quad \square \end{aligned}$$

Corollary 1. Under the conditions of Theorem 4 and $\sup_n \mathbf{E}X_n^2 < \infty$, we have

$$\zeta_1(W, \mathcal{E}) \leq \frac{2p}{|a|} \sup_n \zeta_1(\mathcal{L}(X_n), \mathcal{L}^e(X_n)) \leq p \sup_n \left(\frac{\mathbf{E}X_n^2}{a^2} - 2 \mathbf{P}(X_n \leq 0) \right), \tag{46}$$

$$\zeta_1(W_0, \mathcal{E}) \leq \frac{p}{(1-p)a^2} \sup_n \mathbf{E}X_n^2. \tag{47}$$

Remark 4. The right-hand side of Equation (47) is no less than that of Equation (46) because of the factor $\frac{1}{1-p} > 1$ and the absence of the nonpositive term $-2\mathbf{P}(X_n \leq 0)$. This result agrees with the intuition that W may be closer to \mathcal{E} than W_0 , because S_N contains a.s. one summand more than S_{N_0} .

Corollary 2. Under the conditions of Theorem 4, we have

$$\zeta_2(W, \mathcal{E}) \leq \frac{3p}{|a|} \sum_{n=1}^{\infty} \mathbf{P}(N = n) \zeta_1(\mathcal{L}(X_n), \mathcal{L}^e(X_n)) \leq \frac{3p}{2} \left(\frac{\mathbf{E}X_N^2}{a^2} - 2\mathbf{P}(X_N \leq 0) \right), \tag{48}$$

$$\zeta_2(W_0, \mathcal{E}) \leq \frac{p}{1-p} \cdot \frac{3\mathbf{E}X_N^2}{2a^2}. \tag{49}$$

Recently, Korolev and Zeifman [28] obtained a bound similar to Equation (49), but with the constant factor of 1/2 on the right-hand side instead of 3/2, i.e., three times smaller. The estimate in Equation (48) is also worse than Kalashnikov’s bound in Equation (10) obtained in the i.i.d. case and $\mathbf{E}X_1 = 1$, since Equation (10) with $s = 2$, by Theorem 3, yields

$$\zeta_2(W, \mathcal{E}) \leq p\zeta_1(X_1, \mathcal{E}) \leq 2p\zeta_1(\mathcal{L}(X_1), \mathcal{L}^e(X_1)),$$

while Equation (48) in the i.i.d. case with $\mathbf{E}X_1 = 1$ reduces to

$$\zeta_2(W, \mathcal{E}) \leq 3p\zeta_1(\mathcal{L}(X_1), \mathcal{L}^e(X_1)),$$

which is 1.5 times greater.

Proof. Using subsequently Theorem 1(i,c), the triangle inequality for the Kantorovich metric, Theorem 3, and Lemma 3 together with the homogeneity of the Kantorovich distance and the equilibrium transform, we obtain

$$\begin{aligned} \zeta_2(W, \mathcal{E}) &= \zeta_1(\mathcal{L}^e(W), \mathcal{L}^e(\mathcal{E})) = \zeta_1(\mathcal{L}^e(W), \mathcal{L}(\mathcal{E})) \leq \zeta_1(\mathcal{L}^e(W), \mathcal{L}(W)) + \zeta_1(W, \mathcal{E}) \leq \\ &\leq 3\zeta_1(\mathcal{L}^e(W), \mathcal{L}(W)) \leq \frac{3p}{|a|} \sum_{n=1}^{\infty} \mathbf{P}(N = n) \zeta_1(\mathcal{L}(X_n), \mathcal{L}^e(X_n)). \end{aligned}$$

Similarly,

$$\zeta_2(W_0, \mathcal{E}) \leq 3\zeta_1(\mathcal{L}^e(W_0), \mathcal{L}(W_0)) \leq \frac{3}{2} \cdot \frac{p}{1-p} \cdot \frac{\mathbf{E}X_N^2}{a^2}. \quad \square$$

To study the problem of the accuracy of the estimates obtained above in Equations (46) and (47), let us introduce the *asymptotically best constant* for the Kantorovich distance in the Rényi theorem for geometric random sums of i.i.d. r.v.s in a way similar to the definition of the asymptotically best constant [29] in the classical Berry–Esseen inequality (see also [3,30–35]):

$$C_{AB} := \sup_{\{X_n\} \sim \text{i.i.d.}: \mathbf{E}X_1 \neq 0, \mathbf{E}X_1^2 < \infty} \lim_{p \rightarrow +0} \zeta_1(W, \mathcal{E}) \frac{(\mathbf{E}X_1)^2}{p\mathbf{E}X_1^2}, \tag{50}$$

which serves as a lower bound to the constant C in the inequality

$$\zeta_1(W, \mathcal{E}) \leq Cp\mathbf{E}X_1^2 / (\mathbf{E}X_1)^2, \tag{51}$$

still if it is supposed to hold only for sufficiently small p . Similarly, define C_{AB}^0 for W_0 . The inequality in Equation (46) (similarly, Equation (47)) trivially yields the validity of Equation (51) with $C = 1$ for all $p \in (0, 1)$. Since

$$C \geq C_{AB},$$

it is easy to conclude that $C_{AB} \leq 1$.

Theorem 5. For the asymptotically best constants C_{AB}, C_{AB}^0 defined in Equation (50), for W and W_0 we have

$$C_{AB} \geq 1/4, \quad C_{AB}^0 \geq 1/4.$$

Proof. Taking all $X_n := 1$, we get $\mathbf{E}X_n = \mathbf{E}X_n^2 = 1$ and $W = pN, W_0 := pN_0/(1-p)$, where $N \sim \text{Geom}(p)$ and $N_0 := N - 1$. To estimate $\zeta_1(W, \mathcal{E})$, we use the definition of the Kantorovich distance in Equation (19) and take $h(x) = \frac{1}{t} \sin(tx) \in \text{Lip}_1$ as a test function, where $t \in \mathbb{R} \setminus \{0\}$ is the free parameter to be chosen later. Recalling the ch.f.s of the exponential and the geometric distributions, we obtain

$$\mathbf{E}h(\mathcal{E}) = \frac{1}{t} \mathfrak{S} \mathbf{E} e^{it\mathcal{E}} = \mathfrak{S} \frac{1}{t(1-it)} = \frac{1}{1+t^2},$$

$$\begin{aligned} \mathbf{E}h(W) &= \mathbf{E}h(pN) = \frac{1}{t} \mathfrak{S} \mathbf{E} e^{itpN} = \frac{1}{t} \mathfrak{S} \left[\frac{pe^{itp}}{1-(1-p)e^{itp}} \right] = \\ &= \frac{1}{t} \mathfrak{S} \left[\frac{pe^{itp}(1-(1-p)e^{-itp})}{1+(1-p)^2-2(1-p)\cos(tp)} \right] = \frac{p \sin(tp)}{tp^2 + 2t(1-p)(1-\cos(tp))}, \end{aligned}$$

$$\mathbf{E}h(W_0) = \mathbf{E}h\left(\frac{pN_0}{1-p}\right) = \frac{1}{t} \mathfrak{S} \mathbf{E} e^{itpN_0/(1-p)} = \frac{p(1-p) \sin\left(\frac{tp}{1-p}\right)}{tp^2 + 2t(1-p)\left(1-\cos\left(\frac{tp}{1-p}\right)\right)}.$$

Thus,

$$\begin{aligned} C_{AB} &\geq \limsup_{p \rightarrow +0} \sup_{t \neq 0} \frac{|\mathbf{E}h(W) - \mathbf{E}h(\mathcal{E})|}{p} \geq \sup_{t \neq 0} \lim_{p \rightarrow +0} \left| \frac{\mathbf{E}h(W) - \mathbf{E}h(\mathcal{E})}{p} \right| = \\ &= \sup_{t \neq 0} \lim_{p \rightarrow +0} \left| \frac{p^3 t^3 + o(p^3)}{p^3 t(t^2 + 1)^2 + o(p^3)} \right| = \sup_{t \neq 0} \frac{t^2}{(t^2 + 1)^2} = 1/4, \end{aligned}$$

and, similarly,

$$C_{AB}^0 \geq \sup_{t \neq 0} \lim_{p \rightarrow +0} \left| \frac{\mathbf{E}h(W_0) - \mathbf{E}h(\mathcal{E})}{p} \right| = \sup_{t \neq 0} \frac{t^2}{(t^2 + 1)^2} = 1/4. \quad \square$$

Theorem 1(h) allows extending Theorem 4 to non-geometric random sums of independent random variables with arbitrary means of identical signs. Namely, the following statement holds.

Theorem 6. Let X_1, X_2, \dots be a sequence of independent random variables, independent of all else, with

$$a_n := \mathbf{E}X_n > 0, \quad b_n := \mathbf{E}X_n^2 < \infty, \quad n \in \mathbb{N},$$

and $S_n := \sum_{i=1}^n X_i$ for $n \in \mathbb{N}, S_0 := 0$. Let N be a \mathbb{N}_0 -valued r.v.,

$$A := \mathbf{E}S_N = \sum_{n=1}^{\infty} a_n \mathbf{P}(N \geq n) < \infty,$$

and M be a \mathbb{N} -valued r.v. with the distribution

$$\mathbf{P}(M = m) = \frac{a_m}{A} \mathbf{P}(N \geq m), \quad m \in \mathbb{N}.$$

Assume also that $\mathbf{E}S_M < \infty$. Then, with $W := S_N/\mathbf{E}S_N = A^{-1}S_N$, for any joint distribution $\mathcal{L}(N, M)$, we have

$$\zeta_1(W, \mathcal{E}) \leq 2A^{-1} \left(\sup_n \mathbf{E}|X_n| \cdot \mathbf{E}|N - M| + \sum_{m \in \mathbb{N}} \mathbf{P}(M = m) \zeta_1(\mathcal{L}(X_m), \mathcal{L}^e(X_m)) \right) \leq \tag{52}$$

$$\leq 2A^{-1} \left(\sup_n \mathbf{E}|X_n| \cdot \mathbf{E}|N - M| + \mathbf{E} \left(\frac{b_M}{2a_M} - a_M \cdot \mathbf{P}(X_M \leq 0|M) \right) \right). \tag{53}$$

Remark 5. If both expectations $\mathbf{E}N$ and $\mathbf{E}M$ are finite, then $\mathbf{E}|N - M|$ in Equations (52) and (53) can be replaced with $\zeta_1(N, M)$.

Remark 6. Theorem 6 reduces to ([13], Theorem 3.1) in the case of nonnegative $\{X_n\}$ and to Theorem 4, Equation (44), in the case of $N \sim \text{Geom}(p)$ and identical $a := \mathbf{E}X_n \neq 0, n \in \mathbb{N}$. For shifted geometric N , i.e., $\mathbf{P}(N = n) = p(1 - p)^n, n \in \mathbb{N}_0$, under the assumptions of Theorem 4, Theorem 6 yields a bound

$$\begin{aligned} \zeta_1(W_0, \mathcal{E}) &\leq \frac{p}{1 - p} \left(2 \sup_n \frac{\mathbf{E}|X_n|}{|a|} + \sum_{n \in \mathbb{N}} \mathbf{P}(N = n - 1) \zeta_1(\mathcal{L}(X_n), \mathcal{L}^e(X_n)) \right) \leq \\ &\leq \frac{p}{1 - p} \left(\frac{\mathbf{E}X_{N+1}^2}{a^2} + 2 \sup_n \frac{\mathbf{E}|X_n|}{|a|} - \mathbf{P}(X_{N+1} \leq 0) \right), \end{aligned}$$

whose rightmost part is worse than the estimate in Equation (45), generally speaking (for example, in the i.i.d. case), since $\mathbf{E}|X_n| \geq |a|$ for all $n \in \mathbb{N}$ and $\mathbf{P}(X_{N+1} \leq 0) \leq 1$.

Proof of Theorem 6. By Theorem 3 and homogeneity of the Kantorovich distance and the equilibrium transform (see Lemma 1(a) and Theorem 1(f)), we have

$$\zeta_1(W, \mathcal{E}) \leq 2 \zeta_1(\mathcal{L}(W), \mathcal{L}(W^e)) = 2A^{-1} \zeta_1(\mathcal{L}(S_N), \mathcal{L}^e(S_N)). \tag{54}$$

Let us bound $\zeta_1(\mathcal{L}(S_N), \mathcal{L}^e(S_N))$ from above.

For a given joint distribution $\mathcal{L}(N, M)$, let $p_{nm} := \mathbf{P}(N = n, M = m), n \in \mathbb{N}_0, m \in \mathbb{N}$. Denoting $S_{j,k} := \sum_{i=j}^k X_i$ for $j \leq k$ and using the representation in Equation (20) and Theorem 1(h), we have

$$\begin{aligned} \zeta_1(\mathcal{L}(S_N), \mathcal{L}^e(S_N)) &= \int_{\mathbb{R}} |F_{S_N}(x) - F_{S_N}^e(x)| dx = \int_{\mathbb{R}} |F_{S_N}(x) - F_{S_{M-1}} * F_{X_M}^e(x)| dx = \\ &= \int_{\mathbb{R}} \left| \sum_{n \in \mathbb{N}_0, m \in \mathbb{N}} p_{nm} (F_{S_n}(x) - F_{S_{m-1}} * F_{X_m}^e(x)) \right| dx \leq \\ &\leq \sum_{n,m} p_{nm} \int_{\mathbb{R}} |F_{S_n}(x) - F_{S_{m-1}} * F_{X_m}^e(x)| dx \leq \\ &\leq \sum_{n < m} p_{nm} \int_{\mathbb{R}} |\mathbb{1}_{(0,+\infty)}(x) - F_{S_{n+1,m-1}} * F_{X_m}^e(x)| dx + \\ &\quad + \sum_{n \geq m} p_{nm} \int_{\mathbb{R}} |F_{S_{m,n}}(x) - F_{X_m}^e(x)| dx. \end{aligned}$$

Adding and subtracting $F_{S_{n+1,m}}(x)$ under the modulus sign in the integrands in the first sum (w.r.t. $n < m$) and $F_{X_m}(x)$ in the second one (w.r.t. $n \geq m$) and using further the triangle inequality and Lemma 1(b), we obtain

$$\begin{aligned}
 \zeta_1(\mathcal{L}(S_N), \mathcal{L}^e(S_N)) &\leq \sum_{n < m} p_{nm} \zeta_1(\delta_0, S_{n+1,m}) + \sum_{n \geq m} p_{nm} \zeta_1(S_{m+1,n}, \delta_0) + \\
 &+ \sum_{n,m} p_{nm} \zeta_1(\mathcal{L}(X_m), \mathcal{L}^e(X_m)) = \\
 &= \sum_{n,m} p_{nm} \mathbf{E} \left| \sum_{i=(n \wedge m)+1}^{n \vee m} X_i \right| + \sum_{m \in \mathbb{N}} \mathbf{P}(M = m) \zeta_1(\mathcal{L}(X_m), \mathcal{L}^e(X_m)) \leq \\
 &\leq \sup_i \mathbf{E}|X_i| \cdot \sum_{n,m} p_{nm} |n - m| + \sum_{m \in \mathbb{N}} \mathbf{P}(M = m) \zeta_1(\mathcal{L}(X_m), \mathcal{L}^e(X_m)) \\
 &= \sup_i \mathbf{E}|X_i| \cdot \mathbf{E}|N - M| + \sum_{m \in \mathbb{N}} \mathbf{P}(M = m) \zeta_1(\mathcal{L}(X_m), \mathcal{L}^e(X_m)).
 \end{aligned}$$

Substituting the latter bound into Equation (54) yields Equation (52). The bound in Equation (53) follows from Equation (52) by Theorem 2 (see also Remark 2). □

Author Contributions: Conceptualization, I.S.; methodology, I.S. and M.T.; formal analysis, I.S. and M.T.; investigation, I.S. and M.T.; writing—original draft preparation, I.S. and M.T.; writing—review and editing, I.S. and M.T.; supervision, I.S.; funding acquisition, and I.S. All authors have read and agreed to the published version of the manuscript.

Funding: The results of Sections 1–3 (including Theorem 1) were obtained under support by the Russian Science Foundation, project No. 18-11-00155. The rest of the study was funded by RFBR, project number 20-31-70054, and by the grant of the President of Russia No. MD–189.2019.1.

Acknowledgments: The authors would like to thank Professor Victor Korolev for the careful editing of the manuscript and to anonymous referee for a suggestion resulting in Theorem 6.

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

Abbreviations

The following abbreviations are used in this manuscript:

- r.v. random variable
- i.i.d. independent identically distributed
- d.f. distribution function
- ch.f. characteristic function
- a.s. almost sure
- a.c. absolute continuity, absolutely continuous
- w.r.t. with respect to

References

1. Zolotarev, V.M. Probability metrics. *Theory Probab. Appl.* **1984**, *28*, 278–302. [CrossRef]
2. Zolotarev, V.M. Ideal metrics in the problems of probability theory and mathematical statistics. *Austral. J. Statist* **1979**, *21*, 193–208. [CrossRef]
3. Zolotarev, V.M. *Modern Theory of Summation of Random Variables*; VSP: Utrecht, The Netherlands, 1997.
4. Rachev, S. *Probability Metrics and the Stability of Stochastic Models*; John Wiley and Sons: Chichester, UK, 1991.
5. Mattner, L.; Shevtsova, I.G. An optimal Berry–Esseen type theorem for integrals of smooth functions. *Lat. Am. J. Probab. Math. Stat.* **2019**, *16*, 487–530. [CrossRef]
6. Cambanis, S.; Simons, G.; Stout, W. Inequalities for $E_k(X, Y)$ when the marginals are fixed. *Z. Wahrsch. Verw. Geb.* **1976**, *36*, 285–294. [CrossRef]
7. Solov'yev, A.D. Asymptotic behaviour of the time of first occurrence of a rare event. *Engrgy Cybern.* **1971**, *9*, 1038–1048.
8. Brown, M. Error Bounds for Exponential Approximations of Geometric Convolutions. *Ann. Probab.* **1990**, *18*, 1388–1402. [CrossRef]

9. Kalashnikov, V.V.; Vsekhsvyatskii, S.Y. Metric estimates of the first occurrence time in regenerative processes. In *Stability Problems for Stochastic Models*; Lecture Notes in Mathematics; Springer: Berlin/Heidelberg, Germany, 1985; Volume 1155, pp. 102–130. [[CrossRef](#)]
10. Kruglov, V.M.; Korolev, V.Y. *Limit Theorems for Random Sums*; Moscow State University: Moscow, Russia, 1990. (In Russian)
11. Sugakova, E.V. Estimates in the Rényi theorem for differently distributed terms. *Ukr. Math. J.* **1995**, *47*, 1128–1134. [[CrossRef](#)]
12. Kalashnikov, V.V. *Geometric Sums: Bounds for Rare Events with Applications: Risk Analysis, Reliability, Queueing*; Mathematics and Its Applications; Springer: Dordrecht, The Netherlands, 1997.
13. Peköz, E.A.; Röllin, A. New rates for exponential approximation and the theorems of Rényi and Yaglom. *Ann. Probab.* **2011**, *39*, 587–608. [[CrossRef](#)]
14. Ross, N. Fundamentals of Stein’s method. *Probab. Surv.* **2011**, *8*, 210–293, doi:10.1214/11-PS182. [[CrossRef](#)]
15. Hung, T.L. On the rate of convergence in limit theorems for geometric sums. *Southeast Asian J. Sci.* **2013**, *2*, 117–130.
16. Bogachev, V.I. Weak Convergence of Measures. *Am. Math. Soc.* **2018**, *234*.
17. Glivenko, V.I. *Stieltjes Integral*, 2nd ed.; URSS: Moscow, Russia, 2007.
18. Bogachev, V.I. *Measure Theory. Vol. I, II.*; Springer: Berlin, Germany, 2007.
19. Folland, G.B. *Real Analysis*, 2nd ed.; John Wiley & Sons: New York, NY, USA, 1999.
20. Feller, W. *An Introduction to the Probability Theory and Its Applications. Vol. II.*, 2nd ed.; John Wiley: New York, NY, USA, 1971.
21. Asmussen, W. *Applied Probability and Queues*; Springer: New York, NY, USA, 2003.
22. Lin, X.S., Integrated Tail Distribution. Encyclopedia of Actuarial Science. *Am. Cancer Soc.* **2006**. [[CrossRef](#)]
23. Harkness, W.L.; Shantaram, R. Convergence of a sequence of transformations of distribution functions. *Pacific J. Math.* **1969**, *31*, 403–415. [[CrossRef](#)]
24. Shantaram, R.; Harkness, W.L. On a certain class of limit distributions. *Ann. Math. Stat.* **1972**, *43*, 2067–2071. [[CrossRef](#)]
25. Mulholland, H.P.; Rogers, C.A. Representation Theorems for Distribution Functions. *Proc. Lond. Math. Soc.* **1958**, *s3-8*, 177–223. [[CrossRef](#)]
26. Hoeffding, W. The extrema of the expected value of a function of independent random variables. *Ann. Math. Stat.* **1955**, *26*, 268–275. [[CrossRef](#)]
27. 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*; University of California Press: Berkeley, CA, USA, 1972; pp. 583–602.
28. Korolev, V.; Zeifman, A. Bounds for convergence rate in laws of large numbers for mixed Poisson random sums. *arXiv* **2020**, arXiv:2003.12495.
29. Esseen, C.G. A moment inequality with an application to the central limit theorem. *Skand. Aktuarietidskr.* **1956**, *39*, 160–170. [[CrossRef](#)]
30. Kolmogorov, A.N. Some recent works in the field of limit theorems of probability theory. *Bull. Mosc. Univ.* **1953**, *10*, 29–38. (In Russian)
31. Chistyakov, G.P. Asymptotically proper constants in the Lyapunov theorem. *J. Math. Sci.* **1999**, *93*, 480–483. [[CrossRef](#)]
32. Chistyakov, G.P. A new asymptotic expansion and asymptotically best constants in Lyapunov’s theorem. I. *Theory Probab. Appl.* **2002**, *46*, 226–242. [[CrossRef](#)]
33. Chistyakov, G.P. A new asymptotic expansion and asymptotically best constants in Lyapunov’s theorem. II. *Theory Probab. Appl.* **2002**, *46*, 516–522. [[CrossRef](#)]
34. Chistyakov, G.P. A new asymptotic expansion and asymptotically best constants in Lyapunov’s theorem. III. *Theory Probab. Appl.* **2003**, *47*, 395–414. [[CrossRef](#)]
35. Shevtsova, I.G. On the asymptotically exact constants in the Berry–Esseen–Katz inequality. *Theory Probab. Appl.* **2011**, *55*, 225–252. [[CrossRef](#)]

