Article

Large Deviations for a Class of Multivariate Heavy-Tailed Risk Processes Used in Insurance and Finance

Miriam Hägele [†] and Jaakko Lehtomaa *,[†]

Department of Mathematics and Statistics, University of Helsinki, P.O. Box 64 (Gustaf Hällströmin katu 2), FI-00014 Helsinki, Finland; miriam.hagele@helsinki.fi
* Correspondence: jaakko.lehtomaa@helsinki.fi
† These authors contributed equally to this work.

Abstract: Modern risk modelling approaches deal with vectors of multiple components. The components could be, for example, returns of financial instruments or losses within an insurance portfolio concerning different lines of business. One of the main problems is to decide if there is any type of dependence between the components of the vector and, if so, what type of dependence structure should be used for accurate modelling. We study a class of heavy-tailed multivariate random vectors under a non-parametric shape constraint on the tail decay rate. This class contains, for instance, elliptical distributions whose tail is in the intermediate heavy-tailed regime, which includes Weibull and lognormal type tails. The study derives asymptotic approximations for tail events of random walks. Consequently, a full large deviations principle is obtained under, essentially, minimal assumptions. As an application, an optimisation method for a large class of Quota Share (QS) risk sharing schemes used in insurance and finance is obtained.

1. Introduction and Assumptions

1.1. Introduction

Applications in finance and insurance require multivariate models with heavy-tailed distributions to accurately describe multivariate risks. This includes understanding the possible dependence types of large observations. The case where such observations are restricted to a subset, say orthant, of the d-dimensional space $\mathbb{R}^d$ is studied in the setting of multivariate regular variation in Lehtomaa and Resnick (2020). Many studies on multivariate heavy-tailed distributions are built on the assumption of extremely heavy tails assuming, e.g., regular variation Hult and Lindskog (2006); Hult et al. (2005); Mikosch and Wintenberger (2016); Nyrhinen (2009). In this paper, we concentrate on the less studied situation where the large observations can be found from any direction and where the tails are not as heavy as regularly varying tails. Such situations appear naturally in the case of financial returns of portfolios since the tails are often observed to have a lognormal type distribution Hardy (2001); Jensen and Maheu (2018); Tegnér and Poulsen (2018) and the observations can be present in all orthants Lehtomaa and Resnick (2020).

We study asymptotic approximations of random walks, i.e., multivariate processes $(\mathbf{S}_n) := (\mathbf{S}_n)_{n=1}^{\infty}$ in $\mathbb{R}^d$ where

$$\mathbf{S}_n = \mathbf{X}_1 + \cdots + \mathbf{X}_n$$

and the increments $\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \ldots$ are independent and identically distributed (i.i.d.) random vectors. The class of studied increments is closely related to the class of multivariate subexponential vectors. Our class concerns lighter than polynomial tails where the variables have finite moments of all orders. There exist at least three different approaches in the literature

to define multivariate subexponentiality. The definitions in Cline and Resnick (1992); Omey (2006) require, in addition to subexponentiality of the marginal distributions, a multivariate version of long-tailedness. The approach in Samorodnitsky and Sun (2016) uses an alternative definition via fixed ruin sets in order to define a one-dimensional distribution function with respect to each set. The distribution class considered in this paper is consistent with the definition of Samorodnitsky and Sun (2016). For the one-dimensional case, Denisov et al. (2008) provides an overview of large deviation results for subexponential distributions.

We write **X** in product form as

$$\mathbf{X} = R\mathbf{U}.$$

The one-dimensional radius variable R controls the heaviness of the increments, and **U** indicates which directions (defined by unit vectors) are possible. Variable R can have, for example, Weibull or Lognormal type distribution. This definition can be extended to include the class of elliptical distributions, which frequently appear in the literature in applications in finance, see, for instance, Hult and Lindskog (2002); Klüppelberg et al. (2007). Notably, the tail decay speed of R is not restricted to a narrowly defined parametric class.

The proof methods are based on earlier results concerning one-dimensional random walks such as the ones presented in Lehtomaa (2017). A full large deviations principle with non-trivial rate function under, essentially, minimal assumptions on the distribution is also derived. This study complements Mikosch and Rodionov (2020), which considers lognormal distributions, and the result presented in Bazhba et al. (2020), which focuses on Weibull distributions in the one-dimensional setting. As an application, we get an optimisation method for Quota Share (QS) risk sharing schemes, which are widely used in the field of reinsurance. In a QS-contract, there are two participants called the ceding company and the reinsurance company. They agree to share a random risk Y so that one pays qY and the other pays $(1-q)Y$. Our aim is to optimise the portions q when a company buys reinsurance for all lines of business, i.e., each component of **X** is shared with a reinsurance company. The optimisation is obtained from the viewpoint of both the ceding and the reinsurance company.

1.2. Notation

We denote vectors by bold symbols and their components by upper indices, e.g., for $\mathbf{x} \in \mathbb{R}^d$ we write $\mathbf{x} = (x^1, \ldots, x^d)^T$. The inner product of the vectors **x** and **y** is denoted by $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{j=1}^d x^j y^j$ and $\|\mathbf{x}\|_2$ is the L_2-norm. Here, $\|\mathbf{x}\|_2$ is called the length of **x** and $\mathbf{x}/\|\mathbf{x}\|_2$ the direction of **x**. $\mathbb{S}^{d-1}$ is the d-dimensional unit sphere, the subset of $\mathbb{R}^d$ including all vectors with L_2-norm equal to one. The notation B° stands for the interior of the set B, $\overline{B}$ for its closure, and B^c for its complement. For $r \in \mathbb{R}_+$ and $S \subset \mathbb{S}^{d-1}$, we set

$$V_{r,S} := \left\{ \mathbf{x} \in \mathbb{R}^d : \|\mathbf{x}\|_2 > r, \frac{\mathbf{x}}{\|\mathbf{x}\|_2} \in S \right\}, \tag{1}$$

where the expression $A := B$ means A is defined by B. $B(\mathbf{x}, a)$ defines a ball centred at **x** with radius a, and we denote the inverse function of f by f^{-1}.

The asymptotic relation $f(x) \sim g(x)$, as $x \to \infty$ means $\lim_{x \to \infty} f(x)/g(x) = 1$ and the little-o notation $g(x) = o(f(x))$ means $\lim_{x \to \infty} g(x)/f(x) = 0$. We take the limit $x \to \infty$ or $n \to \infty$, where x denotes real and n natural numbers. The symbol $\mathbb{1}(C)$ denotes the indicator function of the event C, $\mathbb{P}(C)$ its probability, and $\mathbb{E}(X)$ stands for the expectation of X. By A, we denote a symmetric $d \times d$ matrix, and $\Omega \subset \mathbb{R}^d$ is the ellipsoid generated by the linear transformation $\Lambda : \mathbb{R}^d \to \mathbb{R}^d, \Lambda(\mathbf{x}) = A\mathbf{x}$ of the unit sphere, $\Omega = \Lambda(\mathbb{S}^{d-1})$.

1.3. Model Assumptions

The aim is to derive a large deviations principle for elliptical multivariate distributions with moderate heavy tails. Therefore, we study the asymptotic behaviour of the random walk $(\mathbf{S}_n)$, where

$$\mathbf{S}_n = \mathbf{X}_1 + \cdots + \mathbf{X}_n,$$

and $\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \ldots$ are i.i.d. increments. Here, $\mathbf{X}$ is the product of a heavy-tailed random variable R and a random vector $\mathbf{U}$ or $\mathbf{\Theta}$, similarly to the setting in Hägele (2020). We assume that $\mathbf{U}$ is distributed on the d-dimensional unit sphere $\mathbb{S}^{d-1}$, and that $\mathbf{\Theta}$ is distributed on a d-dimensional ellipsoid Ω.

We make the following, essentially minimal, assumptions on $R, \mathbf{U}$, and $\mathbf{\Theta}$. The assumptions are essentially minimal in the sense that omitting one or more of them typically causes the main results to not be true in the given set of non-parametric distributions. However, we assume certain level of smoothness of the tail of R. This assumption is made because it admits presenting the results in a less technical way.

(A1)   The tail function of the random variable R satisfies

$$- \log(\mathbb{P}(R > x)) \sim h(x), \tag{2}$$

as $x \to \infty$, where $h(x)$ is an increasing and concave function such that

(i) $h(x) = o(x)$ and

(ii) $\log(x) = o(h(x))$, as $x \to \infty$.

(A2)   The random vector $\mathbf{U} \in \mathbb{S}^{d-1}$ has a distribution on the d-dimensional unit sphere $\mathbb{S}^{d-1}$. Let $S \subset \mathbb{S}^{d-1}$ be a subset with positive Lebesgue measure. We assume that $\mathbb{P}(\mathbf{U} \in S) > 0$. In addition, $\mathbf{U}$ is assumed to be asymptotically independent of the random variable R in the sense that

$$\lim_{x \to \infty} \mathbb{P}(\mathbf{U} \in S | R > x) = \mathbb{P}(\mathbf{U} \in S),$$

and $\mathbb{E}(R\mathbf{U}) = \mathbf{0}$.

Remark 1. *The random vector $\mathbf{U}$ does not have to be uniformly distributed on the unit sphere. It is enough that for any non-empty open set $S \subset \mathbb{S}^{d-1}$, it holds that*

$$\mathbb{P}(R > 0, \mathbf{U} \in S) > 0.$$

For instance, in this model, it is possible to study a multivariate distribution which assigns more probability mass to some directions than others given that Assumption (A2), $\mathbb{E}(\mathbf{X}) = \mathbf{0}$, is still fulfilled.

We can then define the random vector $\mathbf{\Theta}$ by a linear transformation from the unit sphere $\mathbb{S}^{d-1}$ to the d-dimensional ellipsoid Ω centred at the origin. The linear transformation $\Lambda : \mathbb{R}^d \to \mathbb{R}^d$ with $\Lambda(\mathbf{x}) = A\mathbf{x}$, where A is a symmetric, positive definite matrix, generates the ellipsoid $\Omega = \Lambda(\mathbb{S}^{d-1})$. The random vector $\mathbf{\Theta}$ can then be written as a transformed vector, $\mathbf{\Theta} = \Lambda(\mathbf{U})$. If A is a diagonal matrix, the ellipsoid is orientated along the axes.

Instead of defining $\mathbf{\Theta}$ through the linear transformation Λ, we can write its definition in a similar way as for the random vector $\mathbf{U}$.

(A2')   The d-dimensional random vector $\mathbf{\Theta}$ is distributed on an ellipse or ellipsoid Ω centred at the origin with $\mathbb{E}(R\mathbf{\Theta}) = \mathbf{0}$. It holds for every set $S \subset \Omega$ with positive Lebesgue measure that $\mathbb{P}(\mathbf{\Theta} \in S) > 0$, and $\mathbf{\Theta}$ is asymptotically independent of the random variable R in the sense that $\lim_{x \to \infty} \mathbb{P}(\mathbf{\Theta} \in S | R > x) = \mathbb{P}(\mathbf{\Theta} \in S)$.

Remark 2. *Assumption (A1) implies that the random variable R is heavy-tailed in the sense that* $\mathbb{E}(e^{sR}) = \infty$ *for all* $s > 0$. *The fact that* $\log(x) = o(h(x))$ *implies* $\mathbb{E}(R^s) < \infty$ *for all* $s > 0$ *so the random variable R has finite moments of all orders. Furthermore, it follows from assumptions (A2) and (A2') that the support of the random vector* **U** *or* **Θ** *is the entire set* $\mathbb{S}^{d-1}$ *or* Ω.

Remark 3. *There are visual methods and statistical tests to analyse whether the data fulfils the assumptions. For the heaviness of the tail, one can test the one-dimensional empirical tail distribution of the norm of the observations, for example, with the method introduced in Asmussen and Lehtomaa (2017). The case where probability mass is concentrated on a subset of possible directions is studied in, e.g., Lehtomaa and Resnick (2020).*

Assumption (A1) is closely related to the class of subexponential distributions that is introduced, for instance, in Embrechts et al. (1997); Foss et al. (2011).

Lemma 1. *If Assumption (A1) holds with* (2) *as an equality for large enough arguments, the distribution of R belongs to the class of subexponential distributions.*

Proof. The statement follows directly from Theorem 2 of Teugels (1975), which gives a sufficient condition for subexponentiality of a distribution based on tail functions. The condition has three requirements, two of which are immediately true by our definition. To check the remaining condition, we can define an auxiliary function $g(x) := (x/h(x))^{1/2}$. Then $g(x) \to \infty$ and $x - g(x) \to \infty$, as $x \to \infty$. Without loss of generality, we can assume $h(0) \geq 0$, see Remark 5. Due to the concavity, it holds that

$$
\lim_{x \to \infty} \frac{\mathbb{P}(R > x - g(x))}{\mathbb{P}(R > x)} = \lim_{x \to \infty} \exp\left(-h\left(\left(1 - \frac{g(x)}{x} \right) x \right) + h(x) \right)
$$

$$
\geq \lim_{x \to \infty} \exp\left(-\left(1 - \frac{g(x)}{x} \right) h(x) + h(x) \right) = \lim_{x \to \infty} \exp\left(\frac{g(x)}{x} h(x) \right)
$$

$$
= \lim_{x \to \infty} \exp(g(x)^{-1}) = 1
$$

The corresponding upper bound of the limit is immediately valid by definition. □

Example 1. *Simple examples of distributions of R that fulfil Assumption (A1) include Weibull distributions with parameter* $\beta \in (0, 1)$ *and lognormal type distributions which are defined by the relation* $\mathbb{P}(R > x) \sim e^{-(\log(x))^p}$ *for* $x > x_0$, *where* $p > 1$.

Assumption (A1) can be used to obtain bounds even if it does not hold immediately for a given tail function $\mathbb{P}(R > x)$. For example, if R can be stochastically bounded by, say, R' and R'' in the sense that

$$
\mathbb{P}(R' > x) \leq \mathbb{P}(R > x) \leq \mathbb{P}(R'' > x)
$$

and the variables R' and R'' satisfy (A1) (possibly with different concave functions), a result can be obtained if the asymptotic behaviour concerning the upper and lower bounds coincides in a suitable sense. For a concrete example of this, recall that a random variable R belongs to the class of stretched exponential distributions if, for large enough x, inequalities

$$
l_1(x)e^{-l(x)x^\beta} \leq \mathbb{P}(R > x) \leq l_2(x)e^{-l(x)x^\beta}
$$

hold, where $\beta \in (0, 1)$ and l, l_1, l_2 are slowly varying functions. This class is studied in particular in Gantert (1998); Gantert et al. (2014). Here, Assumption (A1) is valid if $l(x)x^\beta$ is a concave function for large enough x. If it is not, we can still find, based on

Theorem 1 of Lehtomaa (2017), a function $\underline{h}(x)$ which satisfies Assumption (A1) and inequality $\mathbb{P}(R > x) \leq e^{-\underline{h}(x)}$ for large enough x and

$$\liminf_{x \to \infty} \frac{-\log \mathbb{P}(R > x)}{\underline{h}(x)} = 1.$$

This fact can be used in the proofs by replacing $\mathbb{P}(R > x)$ by $e^{-\underline{h}(x)}$ in suitable places in order to obtain results also for the stretched exponential class.

2. Asymptotics of Spherical Distributions

2.1. Large Deviations Principle

Throughout this section, we study the random walk $(\mathbf{S}_n)$ generated by random vectors of the form $\mathbf{X} = R\mathbf{U}$, where the random variable R fulfils Assumption (A1), and the random vector $\mathbf{U}$ fulfils Assumption (A2). We examine the probability of the asymptotic event that the random walk exceeds a threshold in a selected norm in order to prove a large deviations theorem. In this study, we choose to use the L_2-norm because it is, in our view, a natural choice when dealing with ellipses.

We start by considering a spherical distribution. The result is later extended to the setting of asymptotically elliptical heavy-tailed distributions. The proofs of the theorems stated below can be found in Section 2.3. The first result concerns logarithmic asymptotics of the norm of the random walk.

Theorem 1. *Let $a > 0$ be a fixed number. Suppose the increment of the random walk $(\mathbf{S}_n)$ is of the form $\mathbf{X} = R\mathbf{U}$, where assumptions (A1) and (A2) hold. Then,*

$$\lim_{n \to \infty} \frac{\log(\mathbb{P}(\|\mathbf{S}_n\|_2 > na))}{h(na)} = -1.$$

The asymptotic relation derived in Theorem 1 yields a full large deviations principle with non-trivial rate function for asymptotically spherical heavy-tailed distributions under an additional technical assumption.

Theorem 2. *Let $\mathbf{X} = R\mathbf{U}$, where R and $\mathbf{U}$ fulfil assumptions (A1) and (A2). Additionally, assume that, for $a > 0$, the limit*

$$\lim_{x \to \infty} \frac{h(ax)}{h(x)} \tag{3}$$

exists. Then, the process $\{\mathbf{S}_n/n\}$ satisfies the large deviations principle with rate function

$$I(\mathbf{x}) = \begin{cases} \lim_{n \to \infty} \frac{h(n\|\mathbf{x}\|_2)}{h(n)}, & \text{if } \mathbf{x} \neq \mathbf{0} \\ 0, & \text{if } \mathbf{x} = \mathbf{0} \end{cases}$$

and scale h, i.e.,

$$-\inf_{\mathbf{y} \in B^\circ} I(\mathbf{y}) \leq \liminf_{n \to \infty} \frac{\log(\mathbb{P}(\mathbf{S}_n/n \in B))}{h(n)}$$

$$\leq \limsup_{n \to \infty} \frac{\log(\mathbb{P}(\mathbf{S}_n/n \in B))}{h(n)} \leq -\inf_{\mathbf{y} \in \overline{B}} I(\mathbf{y})$$

for all Borel sets $B \subset \mathbb{R}^d$.

The large deviations principle in Theorem 2 is a multivariate equivalent of the large deviation principle in Lehtomaa (2017) for d-dimensional spherical random vectors.

Corollary 1. *If, in addition to the assumptions of Theorem 2, $I(x)$ is continuous for all $x \in \mathbb{R}^d$, and the Borel set B fulfils $\overline{B^\circ} = \overline{B}$, it holds that*

$$\lim_{n\to\infty} \frac{\log(\mathbb{P}(S_n \in nB))}{h(n)} = -\inf_{x\in B} I(x).$$

Proof. The claim follows directly from the assumed continuity of the rate function. □

Remark 4. *The existence of Limit (3) implies*

$$\lim_{x\to\infty} \frac{h(ax)}{h(x)} = a^\alpha$$

for some $\alpha \geq 0$ due to Theorem 1.4.1 in Bingham et al. (1989) so $h(x)$ is, in fact, a regularly varying function.

The rate function is symmetric with respect to the origin. Furthermore, the one-dimensional rate function along any line segment with an endpoint in the origin is concave like the rate function in the one-dimensional case examined in Lehtomaa (2017).

The following example examines the rate function for typical distributions that fulfil Assumption (A1).

Example 2.

(i) *Let R be Weibull distributed with parameter $\beta \in (0,1)$. Then, $h(x) = cx^\beta$ with some constant $c > 0$ so the index α from Remark 4 is equal to β. Furthermore,*

$$I(x) = \|x\|_2^\beta$$

for all $x \in \mathbb{R}^d$ and I is a good rate function. The rate function $I(x)$ is continuous so Corollary 1 holds.

(ii) *If R has a lognormal type distribution of the form $\mathbb{P}(R > x) \sim e^{-(\log(x))^p}$ for $x > x_0$ with some parameter $p > 1$, the index α from Remark 4 is 0, and*

$$I(x) = 1$$

for all $x \in \mathbb{R}^d \setminus \{0\}$ so the rate function I jumps at the origin. The rate function is non-trivial, but not good, according to the terminology used in the context of large deviations.

2.2. Auxiliary Results

In order to prove Theorem 1 and Theorem 2, we need some auxiliary lemmas. The auxiliary results study the projection of the random walk to a one-dimensional setting and its asymptotics.

The orthogonal projection $P_v(\cdot)$ on the subspace spanned by the vector $v \in \mathbb{S}^{d-1}$ is defined as

$$P_v(x) := \langle v, x \rangle v,$$

where the inner product defined as $p_v(x) := \langle v, x \rangle$ indicates the length of the projected vector in the subspace according to the L_2-norm.

The next result shows that the projection of the random vector X has the same asymptotic behaviour as $\|X\|_2$.

Lemma 2. *Suppose $X = RU$ where assumptions (A1) and (A2) hold. Let $a > 0$ and $v \in \mathbb{S}^{d-1}$. Then,*

$$\lim_{n\to\infty} \frac{\log(\mathbb{P}(p_v(X) > na))}{h(na)} = -1.$$

Proof. The asymptotic upper bound is due to

$$\log(\mathbb{P}(p_{\mathbf{v}}(\mathbf{X}) > na)) \leq \log(\mathbb{P}(R > na)) \sim -h(na)$$

since $\langle \mathbf{v}, \mathbf{U} \rangle \leq 1$.

To prove the asymptotic lower bound, we fix $\delta > 0$ and set $S(\mathbf{v}, \delta) \subset \mathbb{S}^{d-1}$ to be the δ-environment of the vector $\mathbf{v}$ in $\mathbb{S}^{d-1}$. Defining

$$c_{\mathbf{v}_\delta} := \min_{\mathbf{y} \in S(\mathbf{v}, \delta)} \langle \mathbf{v}, \mathbf{y} \rangle$$

which is positive choosing δ small enough, it holds that

$$
\begin{aligned}
&\log(\mathbb{P}(p_{\mathbf{v}}(\mathbf{X}) > na)) \\
\geq\ & \log(\mathbb{P}(\langle \mathbf{v}, \mathbf{U} \rangle R > na, \mathbf{U} \in S(\mathbf{v}, \delta))) \\
\geq\ & \log\left(\mathbb{P}\left(R > \frac{na}{c_{\mathbf{v}_\delta}}, \mathbf{U} \in S(\mathbf{v}, \delta) \right) \right) \\
=\ & \log\left(\mathbb{P}\left(R > \frac{na}{c_{\mathbf{v}_\delta}} \right) \right) + \log\left(\mathbb{P}\left(\mathbf{U} \in S(\mathbf{v}, \delta) \,\Big|\, R > \frac{na}{c_{\mathbf{v}_\delta}} \right) \right).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\liminf_{n \to \infty} \frac{\log(\mathbb{P}(p_{\mathbf{v}}(\mathbf{X}) > na))}{h(na)} \ &\geq\ \liminf_{n \to \infty} \frac{\log(\mathbb{P}(\mathbf{U} \in S(\mathbf{v}, \delta))) - h\left(\frac{na}{c_{\mathbf{v}_\delta}} \right)}{h(na)} \\
&=\ \liminf_{n \to \infty} \frac{-h\left(\frac{na}{c_{\mathbf{v}_\delta}} \right)}{h(na)}
\end{aligned}
$$

which holds for every $\delta > 0$ small enough and thus proves the claim since $c_{\mathbf{v}_\delta} \to 1$, as $\delta \to 0$. □

In the proof of Lemma 4, we divide the probability into the term caused by a single big jump and its complement. The following lemma provides an upper bound for the remaining term not caused by a single big jump.

Lemma 3. *Let $Y, Y_1, Y_2, \ldots$ be i.i.d. real-valued random variables with $\mathbb{E}(Y) = 0$ and finite moments of all order, and suppose $a > 0$. Furthermore, let*

$$\liminf_{x \to \infty} \frac{-(\log(\mathbb{P}(Y > x)))}{h(x)} \geq 1,$$

where $h(x)$ fulfils the assumptions on the function $h(x)$ stated in Assumption (A1), and additionally, $h(0) \geq 0$. Then,

$$\limsup_{n \to \infty} \frac{\log\left(\mathbb{P}\left(\sum_{i=1}^n Y_i > na, \max_{i=1}^n Y_i \leq na \right) \right)}{h(na)} \leq -1.$$

The proof uses similar ideas as the proof of Theorem 2 in Lehtomaa (2017).

Proof of Lemma 3. Due to the inequality

$$\mathbb{E}\left(e^{b_n \sum_{i=1}^n Y_i} \mathbb{1}\left(\sum_{i=1}^n Y_i > na, \max_{i=1}^n Y_i \leq na \right) \right)$$

$$\geq\ e^{b_n na} \mathbb{P}\left(\sum_{i=1}^n Y_i > na, \max_{i=1}^n Y_i \leq na \right)$$

and the independence of the random variables, it holds that

$$\mathbb{P}\left(\sum_{i=1}^{n} Y_i > na, \max_{i=1}^{n} Y_i \le na\right) \le \exp(-b_n na)\left(\mathbb{E}\left(e^{b_n Y}\mathbb{1}(Y \le na)\right)\right)^n.$$

To bound the expectation from above, we split it into two parts

$$
\begin{aligned}
\mathbb{E}\left(e^{b_n Y}\mathbb{1}(Y \le na)\right) &= \mathbb{E}\left(e^{b_n Y}\mathbb{1}(Y \le c_n)\right) + \mathbb{E}\left(e^{b_n Y}\mathbb{1}(c_n < Y \le na)\right) \\
&= E_1 + E_2.
\end{aligned}
$$

Taking $\delta \in (0,1)$ and setting $b_n := (1-\delta)h(na)/na \to 0$, one can choose $\varepsilon(n) \to 0$ such that $c_n := b_n^{-1}\varepsilon(n) \to \infty$, as $n \to \infty$. Then, it holds $b_n c_n \to 0$ and therefore one can apply Taylor series to the first term E_1,

$$
\begin{aligned}
E_1 &= \mathbb{E}((1 + b_n Y + o(b_n Y))\mathbb{1}(Y \le c_n)) \\
&= \mathbb{P}(Y \le c_n) + b_n \mathbb{E}(Y\mathbb{1}(Y \le c_n))(1 + o(1)).
\end{aligned}
$$

Integrating E_2 by parts and rewriting it in terms of the tail distribution of Y, one gets

$$
\begin{aligned}
E_2 &= e^{b_n na}\mathbb{P}(Y \le na) - e^{b_n c_n}\mathbb{P}(Y \le c_n) - \int_{c_n}^{na} b_n e^{b_n x}\mathbb{P}(Y \le x)\mathrm{d}x \\
&= e^{b_n c_n}\mathbb{P}(Y > c_n) - e^{b_n na}\mathbb{P}(Y > na) + b_n \int_{c_n}^{na} e^{b_n y}\mathbb{P}(Y > y)\mathrm{d}y.
\end{aligned}
$$

For every $\delta > 0$, it holds $\mathbb{P}(Y > y) \le \exp(-(1-\delta/2)h(y))$ for all $y \ge c_n$ choosing n large enough. Applying additionally Taylor series to the first term of the latter equation results in the upper bound

$$E_2 \le (1 + b_n c_n + o(b_n c_n))\mathbb{P}(Y > c_n) + b_n \int_{c_n}^{na} \exp\left(b_n y - \left(1 - \frac{\delta}{2}\right)h(y)\right)\mathrm{d}y.$$

Due to the concavity of the function $h(x)$ and the fact that $y \le na$, it holds that

$$h(y) \ge \frac{y}{na}h(na)$$

and therefore

$$
\begin{aligned}
& b_n \int_{c_n}^{na} \exp\left(b_n y - \left(1 - \frac{\delta}{2}\right)h(y)\right)\mathrm{d}y \\
={}& b_n \int_{c_n}^{na} \exp\left((1-\delta)\frac{y}{na}h(na) - (1-\delta)h(y) - \frac{\delta}{2}h(y)\right)\mathrm{d}y \\
\le{}& b_n \int_{c_n}^{na} \exp\left(-\frac{\delta}{2}h(y)\right)\mathrm{d}y \le b_n na \exp\left(-\frac{\delta}{2}h(c_n)\right).
\end{aligned}
$$

Applying the inequality $\log(x) \le x - 1$ to the term $n\log(E_1 + E_2)/h(na)$ yields

$$
\begin{aligned}
& \frac{n}{h(na)}\log\left(\mathbb{E}\left(e^{b_n Y}\mathbb{1}(Y \le na)\right)\right) \\
\le{}& \frac{n}{h(na)}\log\left(\mathbb{P}(Y \le c_n) + b_n \mathbb{E}(Y\mathbb{1}(Y \le c_n))(1 + o(1)) + \mathbb{P}(Y > c_n)\right. \\
& \left. {} + b_n c_n \mathbb{P}(Y > c_n)(1 + o(1)) + b_n na \exp\left(-\frac{\delta}{2}h(c_n)\right)\right) \\
\le{}& \frac{1 - \delta + o(1)}{a}(\mathbb{E}(Y\mathbb{1}(Y \le c_n)) + c_n \mathbb{P}(Y > c_n)) + (1-\delta)n\exp\left(-\frac{\delta}{2}h(c_n)\right).
\end{aligned}
$$

The first terms converge to zero, because $\mathbb{E}(Y) = 0$ and all moments are finite. Choosing $\varepsilon(n)$ such that $\log(n) = o(h(c_n))$, also the last term converges to zero. Finally,

$$
\limsup_{n\to\infty} \frac{\log\big(\mathbb{P}\big(\sum_{i=1}^{n} Y_i > na, \max_{i=1}^{n} Y_i \leq na\big)\big)}{h(na)}
$$
$$
\leq \limsup_{n\to\infty} \frac{-b_n na}{h(na)} + \frac{n}{h(na)} \log\Big(\mathbb{E}\big(e^{b_n Y} \mathbb{1}(Y \leq na)\big)\Big) = -(1 - \delta).
$$

The last inequality holds for any $\delta > 0$, which implies the claim. □

Remark 5. *If h is an increasing concave function with $h(x) = o(x)$ and $\log(x) = o(h(x))$, one can always construct an increasing concave function h' with $h'(0) \geq 0$ and $h'(x) \sim h(x)$, as $x \to \infty$ by changing h to a suitable linear function near the origin. This means that h' is a sub-additive function.*

We can now state the principle of a single big jump for projections of multivariate random walks.

Lemma 4. *Let $X = RU$, assume (A1) and (A2), and let $v \in \mathbb{S}^{d-1}$ and $a > 0$ be fixed. Then, it holds that*

$$
\lim_{n\to\infty} \frac{\log(\mathbb{P}(p_v(S_n) > na))}{\log(\mathbb{P}(p_v(X) > na))} = 1.
$$

Proof. At first, we show the asymptotic lower bound

$$
\liminf_{n\to\infty} -\frac{\log(\mathbb{P}(p_v(S_n) > na))}{\log(\mathbb{P}(p_v(X) > na))} \geq -1.
$$

Using the principle of a single big jump and the weak law of large numbers, it follows that

$$
\liminf_{n\to\infty} -\frac{\log(\mathbb{P}(p_v(S_n) > na))}{\log(\mathbb{P}(p_v(X) > na))}
$$
$$
\geq \liminf_{n\to\infty} -\frac{\log(\mathbb{P}(p_v(S_{n-1}) \leq \varepsilon na, p_v(X_n) > (1+\varepsilon)na))}{\log(\mathbb{P}(p_v(X) > na))}
$$
$$
\geq \liminf_{n\to\infty} -\frac{\log(\mathbb{P}(p_v(S_{n-1}) \leq \varepsilon na))}{\log(\mathbb{P}(p_v(X) > na))} - \frac{\log(\mathbb{P}(p_v(X) > (1+\varepsilon)na))}{\log(\mathbb{P}(p_v(X) > na))}
$$
$$
= \liminf_{n\to\infty} -\frac{h((1+\varepsilon)na)}{h(na)}.
$$

The last expression applies, additionally to the weak law of large numbers, Lemma 2. The fact that this holds for every $\varepsilon > 0$ implies the asymptotic lower bound.

It remains to show the asymptotic upper bound

$$
\limsup_{n\to\infty} -\frac{\log(\mathbb{P}(p_v(S_n) > na))}{\log(\mathbb{P}(p_v(X) > na))} \leq -1.
$$

Dividing the probability into the case where at least the projection of one random vector exceeds the threshold na and its complement implies

$$
\mathbb{P}(p_v(S_n) > na)
$$
$$
\leq n\mathbb{P}(p_v(X) > na) + \mathbb{P}(p_v(S_n) > na, p_v(X_i) \leq na \text{ for all } i = 1, \ldots n).
$$

Applying Lemma 1.2.15 in Dembo and Zeitouni (1993),

$$\limsup_{\varepsilon \to 0} \varepsilon \log(a^1_\varepsilon + a^2_\varepsilon) = \max\left(\limsup_{\varepsilon \to 0} \varepsilon \log(a^1_\varepsilon), \limsup_{\varepsilon \to 0} \varepsilon \log(a^2_\varepsilon)\right),$$

we can examine the terms separately since $(\log(\mathbb{P}(p_\mathbf{v}(\mathbf{X}) > na)))^{-1} \to 0$, as $n \to \infty$. We get

$$\limsup_{n \to \infty} -\frac{\log(\mathbb{P}(p_\mathbf{v}(\mathbf{S}_n) > na))}{\log(\mathbb{P}(p_\mathbf{v}(\mathbf{X}) > na))}$$

$$\leq \max\left(\limsup_{n \to \infty} -\frac{\log(n\mathbb{P}(p_\mathbf{v}(\mathbf{X}) > na))}{\log(\mathbb{P}(p_\mathbf{v}(\mathbf{X}) > na))},\right.$$

$$\left.\limsup_{n \to \infty} -\frac{\log(\mathbb{P}(p_\mathbf{v}(\mathbf{S}_n) > na, p_\mathbf{v}(\mathbf{X}_i) \leq na \text{ for all } i = 1, \ldots n))}{\log(\mathbb{P}(p_\mathbf{v}(\mathbf{X}) > na))}\right).$$

Clearly, the first term is -1. For the second term, we set $Y = p_\mathbf{v}(\mathbf{X})$. Since

$$\log(\mathbb{P}(p_\mathbf{v}(\mathbf{X}) > na)) = \log(\mathbb{P}(\langle \mathbf{v}, \mathbf{U} \rangle R > na)) \leq \log(\mathbb{P}(R > na)) \sim -h(na),$$

we can apply Lemma 3 with the help of Remark 5. Finally, applying Lemma 2 we get the upper bound -1 also for the second term, which completes the proof. $\square$

2.3. Proof of Theorems 1 and 2

We can now state the proofs of the main results of Section 2.

Proof of Theorem 1. First, we approximate the event $\{\|\mathbf{S}_n\|_2 > na\}$ by projections in different directions. The fact that the principle of a single big jump holds for any orthogonal projection yields the desired asymptotic behaviour. Since $\mathbb{P}(\|\mathbf{S}_n\|_2 > na) \geq \mathbb{P}(p_\mathbf{v}(\mathbf{S}_n) > na)$ for every $\mathbf{v} \in \mathbb{S}^{d-1}$, the asymptotic lower bound

$$\liminf_{n \to \infty} \frac{\log(\mathbb{P}(\|\mathbf{S}_n\|_2 > na))}{h(na)} \geq -1$$

is an immediate consequence of Lemmas 2 and 4.

To prove the corresponding upper bound, we cover the set $\{\mathbf{x} : \|\mathbf{x}\|_2 > na\}$ by a finite union of m sets defined by orthogonal projections and study the limit, as $m \to \infty$. To this end, let $m \geq 2d$. We aim to choose vectors $\mathbf{v}_k \in \mathbb{S}^{d-1}, k = 1, \ldots, m$, such that unions of form $\cup_{k=1}^m \{\mathbf{x} : p_{\mathbf{v}_k}(\mathbf{x}) > \varepsilon\}$, where $\varepsilon > 0$, can be used to cover the whole space except some neighbourhood of the origin. For example, in $\mathbb{R}^2$, an easy way to define the vectors $\mathbf{v}_k \in \mathbb{S}^1$ is to take $\mathbf{v}_k = (\cos(2k\pi/m), \sin(2k\pi/m))^T$.

Choosing $\mathbf{v}_k$ appropriately, for instance, such that they are uniformly spaced on the unit sphere, there exists some $\varepsilon(m) > 0$ such that

$$D := \cup_{k=1}^m \{\mathbf{x} : p_{\mathbf{v}_k}(\mathbf{x}) > (1 - \varepsilon(m))na\} \supset \{\mathbf{x} : \|\mathbf{x}\|_2 > na\}$$

and, more specifically, we can choose the numbers $\varepsilon(m)$ so that $\varepsilon(m) \to 0$, as $m \to \infty$. Hence,

$$\mathbb{P}(\|\mathbf{S}_n\|_2 > na) \leq \mathbb{P}(\mathbf{S}_n \in D) \leq \sum_{k=1}^m \mathbb{P}(p_{\mathbf{v}_k}(\mathbf{S}_n) > (1 - \varepsilon(m))na).$$

Lemma 1.2.15 in Dembo and Zeitouni (1993) yields

$$\limsup_{n\to\infty} \frac{\log(\mathbb{P}(\|\mathbf{S}_n\|_2 > na))}{h(na)}$$

$$\leq \limsup_{n\to\infty} \frac{\log\left(\sum_{k=1}^m \mathbb{P}(p_{\mathbf{v}_k}(\mathbf{S}_n) > (1-\varepsilon(m))na)\right)}{h(na)}$$

$$\leq \max_{k=1}^m \left(\limsup_{n\to\infty} \frac{\log\left(\mathbb{P}(p_{\mathbf{v}_k}(\mathbf{S}_n) > (1-\varepsilon(m))na)\right)}{h(na)}\right).$$

Applying Lemma 4 and Lemma 2, we get

$$\limsup_{n\to\infty} \frac{\log(\mathbb{P}(\|\mathbf{S}_n\|_2 > na))}{h(na)} \leq \limsup_{n\to\infty} \frac{-h((1-\varepsilon(m))na)}{h(na)}.$$

Letting $m \to \infty$, it holds that $\varepsilon(m) \to 0$ and the term above converges to -1. $\square$

Proof of Theorem 2. The proof of the large deviations principle is based on the fact that

$$\lim_{n\to\infty} \frac{\log(\mathbb{P}(\mathbf{S}_n \in V_{na,S}))}{h(na)} = -1, \tag{4}$$

since open sets of the form $V_{a,S}$ defined in (1), where $a > 0$ and S is an open subset of $\mathbb{S}^{d-1}$, generate $\mathbb{R}^d$.

The limit superior of (4) follows directly from Theorem 1 due to the inequality

$$\mathbb{P}(\mathbf{S}_n \in V_{na,S}) \leq \mathbb{P}(\|\mathbf{S}_n\|_2 > na).$$

Rewriting

$$\mathbb{P}(\mathbf{S}_n \in V_{na,S}) = \mathbb{P}(\|\mathbf{S}_n\|_2 > na)\mathbb{P}\left(\frac{\mathbf{S}_n}{\|\mathbf{S}_n\|_2} \in S \,\Big|\, \|\mathbf{S}_n\|_2 > na\right)$$

yields the limit inferior of (4) since the last probability is positive.

The weak law of large numbers implies $\lim_{n\to\infty} \mathbb{P}(\mathbf{S}_n/n \in B(\mathbf{0}, \varepsilon)) = 1$ for all $\varepsilon > 0$ and thus if $\mathbf{0} \in B \subset \mathbb{R}^d$

$$\lim_{n\to\infty} \frac{\log(\mathbb{P}(\mathbf{S}_n/n \in B))}{h(n)} = 0.$$

Finally, by (4), it is easy to obtain the inequalities

$$\limsup_{n\to\infty} \frac{\log(\mathbb{P}(\mathbf{S}_n/n \in F))}{h(n)} \leq -\inf_{y\in F} I(y) \text{ for all closed sets } F \subset \mathbb{R}^d,$$

$$\liminf_{n\to\infty} \frac{\log(\mathbb{P}(\mathbf{S}_n/n \in G))}{h(n)} \geq -\inf_{y\in G} I(y) \text{ for all open sets } G \subset \mathbb{R}^d$$

which result in the full large deviations principle. $\square$

3. Asymptotics in the Elliptical Case

3.1. Contraction Principle

To extend the result to asymptotically elliptically distributed random vectors, we apply the contraction principle, Theorem 4.2.1 in Dembo and Zeitouni (1993), to the large deviations result in Theorem 2.

In this section, we study the asymptotics of the random walk $(\mathbf{S}_n)$ with increments $\mathbf{X} = R\boldsymbol{\Theta}$, where $\boldsymbol{\Theta}$ is distributed on some d-dimensional ellipsoid Ω centred at the origin. For asymptotically elliptical distributions, a suitable linear map in the contraction principle is the bijective function $\Lambda : \mathbb{R}^d \to \mathbb{R}^d$ defined by $\Lambda(\mathbf{x}) = A\mathbf{x}$ mapping vectors from $\mathbb{S}^{d-1}$

to Ω. Here, A is a symmetric, positive definite and thus invertible $d \times d$ matrix, such that $\Lambda(\mathbb{S}^{d-1}) = \Omega$. Since Λ is linear,

$$\Lambda(\mathbf{S}_n) = \sum_{i=1}^{n} \Lambda(\mathbf{X}_i)$$

holds.

Theorem 3. *Let* $\mathbf{X} = R\mathbf{\Theta}$*, where* R *and* $\mathbf{\Theta}$ *fulfil assumptions (A1) and (A2') and Limit (3) exists. Let* A *be a symmetric, positive definite* $d \times d$ *matrix and define* $\Lambda : \mathbb{R}^d \to \mathbb{R}^d$ *by* $\Lambda(\mathbf{x}) = A\mathbf{x}$ *such that* Λ *maps* $\mathbb{S}^{d-1}$ *to* Ω*. Then, the process* $\{\mathbf{S}_n/n\}$ *satisfies the large deviations principle with scale* h*, so for all Borel sets* $B \subset \mathbb{R}^d$

$$-\inf_{\mathbf{y} \in \Lambda^{-1}(B^\circ)} I(\mathbf{y}) \leq \liminf_{n \to \infty} \frac{\log(\mathbb{P}(\mathbf{S}_n/n \in B))}{h(n)}$$

$$\leq \limsup_{n \to \infty} \frac{\log(\mathbb{P}(\mathbf{S}_n/n \in B))}{h(n)} \leq -\inf_{\mathbf{y} \in \Lambda^{-1}(\overline{B})} I(\mathbf{y})$$

where

$$I(\mathbf{x}) = \begin{cases} \lim_{n \to \infty} \frac{h(n\|\mathbf{x}\|_2)}{h(n)}, & \text{if } \mathbf{x} \neq \mathbf{0} \\ 0, & \text{if } \mathbf{x} = \mathbf{0} \end{cases}$$

and $\Lambda^{-1} : \mathbb{R}^d \to \mathbb{R}^d$ *with* $\Lambda^{-1}(\mathbf{x}) = A^{-1}\mathbf{x}$.

Example 3. *Similarly to Example 2, if* $\mathbf{X}$ *fulfils the assumptions of Theorem 3, where* R *follows a Weibull distribution with parameter* $\beta \in (0, 1)$,

$$I(\mathbf{x}) = \|\mathbf{x}\|_2^\beta$$

is a good rate function. If R *has a lognormal type distribution, the rate function is a constant everywhere except at the origin and hence not good.*

3.2. Proof of Theorem 3

The proof of the main result in this section relies on the linearity of the function Λ and the contraction principle for large deviation principles.

Proof of Theorem 3. The linear transformation

$$\Lambda(\mathbf{x}) = A\mathbf{x}$$

where A is a symmetric, positive definite $d \times d$ matrix is a bijective function $\Lambda : \mathbb{R}^d \to \mathbb{R}^d$. Its inverse function $\Lambda^{-1} : \mathbb{R}^d \to \mathbb{R}^d$ exists due to the invertibility of the matrix A and is $\Lambda^{-1}(\mathbf{x}) = A^{-1}\mathbf{x}$. The linear transformation Λ maps $\mathbb{S}^{d-1}$ to Ω. By the linearity of Λ, it holds $\Lambda(\mathbf{S}_n) = \sum_{i=1}^{n} \Lambda(\mathbf{X}_i)$, so the mapping of the random walk is equal to the sum of mappings of the increments.

The claim follows then from the contraction principle, see, for instance, Theorem 4.2.1 in Dembo and Zeitouni (1993). Applying the contraction principle with the continuous function Λ^{-1}, we get

$$\inf_{\mathbf{y} \in B} J(\mathbf{y}) = \inf_{\mathbf{y} \in B} \inf_{\mathbf{x} \in \mathbb{R}^d} \{I(\mathbf{x}) : \Lambda(\mathbf{x}) = \mathbf{y}\}$$

$$= \inf_{\mathbf{x} \in \mathbb{R}^d} \{I(\mathbf{x}) : \mathbf{x} \in \Lambda^{-1}(B)\} = \inf_{\mathbf{x} \in \Lambda^{-1}(B)} I(\mathbf{x}),$$

which completes the proof. $\quad\square$

4. Applications to Reinsurance

4.1. Introduction and Setting

An insurance company with d lines of business might optimise the asymptotic behaviour of their ruin probability by sharing its risks for some lines of business using quota share reinsurance contracts. Quota share reinsurance is a proportional reinsurance Embrechts et al. (1997), where the insurer (the ceding company) pays only a fixed ratio for each claim, while the reinsurance company pays the rest. In general contracts, the insurance company shares both the losses and the profits with the reinsurer.

A diagonal $d \times d$ matrix Q can be used to represent a quota share reinsurance strategy for an insurance company with d lines of business. The element $q_{k,k}$ refers then to the quota share ratio of the kth line of business, i.e., the ceding company pays $q_{k,k}X^k$ of the kth line of business, and the reinsurance company pays the remaining part $(1 - q_{k,k})X^k$. It is natural to assume that $q_{k,k} \in (0,1]$ for all indices $k = 1, \ldots, d$ since typically the insurance company keeps some share for every line of business. Under this assumption, Matrix Q is invertible.

For the ceding company, the aim could be to find a quota share reinsurance strategy defined by a matrix Q such that

$$\lim_{n \to \infty} \frac{\log(\mathbb{P}(\|\mathbf{S}_n/n\|_2 > a))}{h(n)} > \lim_{n \to \infty} \frac{\log(\mathbb{P}(\|Q\mathbf{S}_n/n\|_2 > a))}{h(n)} \tag{5}$$

because it reduces the asymptotic size of the ruin probability $\mathbb{P}(\|\mathbf{S}_n/n\|_2 > a)$.

We look at the quota share reinsurance from two different perspectives. In Section 4.2, we model the optimal quota share reinsurance strategy from the point of view of the ceding company, which gets reinsurance only for the lines of business with highest risks. Section 4.3 compares different quota share risk sharing strategies from the viewpoint of a reinsurance company that wants to offer quota share reinsurance while minimising their own risks.

We model the risk process of an insurance company with d lines of business as a d-dimensional random walk $(\mathbf{S}_n)$. The component of the increment $\mathbf{X}$ represents the difference between the claim size and the associated net premium in the corresponding line of business. We assume that $\mathbf{X}$ is asymptotically elliptically distributed: Let $\mathbf{X} = R\mathbf{\Theta}$ where R and $\mathbf{\Theta}$ fulfil assumptions (A1) and (A2'), and Ω is an ellipse or ellipsoid defined by the bijection $\Lambda : \mathbb{R}^d \to \mathbb{R}^d$ with $\Lambda(\mathbf{x}) = A\mathbf{x}$ where A is a symmetric, positive definite $d \times d$ matrix.

If there are several reinsurance strategies that yield the same right-hand side of Inequality (5), the insurance company chooses the one with the smallest premium. In this setting, we define the premium of the reinsurance strategy Q as

$$p(Q) = \mathbf{1}^T(I - Q)\mathbf{p} = \sum_{j=1}^{d}(1 - q_{j,j})p_j \tag{6}$$

where $\mathbf{p}$ is a positive premium vector, $\mathbf{1}$ denotes the d-dimensional vector with ones and I the $d \times d$ identity matrix. The positive premium vector $\mathbf{p}$ contains the premium rates of the reinsurances for each line of business when the entire component is reinsured. For example, p_j could be connected to the expected loss of the reinsurance company added with a safety loading. In (6), the constants p_j are multiplied by the factors $(1 - q_{j,j})$, where the q-coefficients can be selected by the ceding company. The premium vector is considered as given, and the ceding company cannot affect the values of this vector. Therefore, (6) is the premium for the entire reinsurance strategy. If the insurance company does not take reinsurance for the jth line of business, it selects $q_{j,j} = 1$. Hence, the premium for the reinsurance of the jth line of business is zero in this case.

4.2. Quota Share Reinsurance Strategy of the Ceding Company

The aim of the insurance company is to identify the riskiest lines of business and choose its quota share reinsurance strategy reducing these risks. The ceding company typically only wants to insure the highest risks. This is why it is natural to assume $q_{k,k} = 1$ for at least one index k, which represents the line of business with the lowest asymptotic risk.

Since quota share reinsurance is defined component-by-component, one can find an optimal reinsurance strategy for distributions on ellipsoids orientated along the axes. The general case can be mathematically reduced into this setting by rotating the original data. However, if the data is transformed using a rotation, suitable QS contracts might not be immediately available on the market because the new axes would not correspond to the lines of business.

Remark 6. *If Ω is a d-dimensional ellipsoid orientated along the axes, there exists a positive, diagonal $d \times d$ matrix A such that the mapping $\Lambda : \mathbb{R}^d \to \mathbb{R}^d, \Lambda(x) = Ax$ of the unit sphere $\mathbb{S}^{d-1}$ generates Ω. Furthermore, if A is a positive, diagonal $d \times d$ matrix, the ellipsoid generated by $\Lambda(\mathbb{S}^{d-1})$ is orientated along the axes.*

Theorem 4. *Let $X = R\Theta$, where Θ fulfils assumption (A2') and $\Omega = \Lambda(\mathbb{S}^{d-1})$ is defined by the linear transformation $\Lambda(x) = Ax$, where A is a diagonal $d \times d$ matrix with $a_{k,k} > 0$ for all $k = 1, \ldots, d$. Furthermore, we assume that R follows a Weibull distribution with parameter $\beta \in (0, 1)$. Then, the quota share reinsurance strategy defined by the matrix $Q = \min_{j=1}^{d} a_{j,j} A^{-1}$ yields the inequality*

$$\lim_{n \to \infty} \frac{\log(\mathbb{P}(\|S_n/n\|_2 > a))}{h(n)} \geq \lim_{n \to \infty} \frac{\log(\mathbb{P}(\|QS_n/n\|_2 > a))}{h(n)} \qquad (7)$$

for any $a > 0$, and Q minimises the right-hand side of Inequality (7) over all strategies. The minimum is unique under the additional condition that $p(Q)$ is also minimised.

Proof. The assumptions on R imply the large deviations principle

$$\lim_{n \to \infty} \frac{\log(\mathbb{P}(\|QS_n/n\|_2 > a))}{h(n)} = - \inf_{y \in B(0,a)^c} \|A^{-1}Q^{-1}y\|_2^{\beta}.$$

Therefore, it is sufficient to show that the matrix Q is the matrix that maximises

$$\max_{\tilde{Q} \in \mathcal{Q}} \inf_{x \in B(0,a)^c} \|A^{-1}\tilde{Q}^{-1}x\|_2 \qquad (8)$$

where $\mathcal{Q}$ is the set of $d \times d$ diagonal matrices with $q_{j,j} \in (0, 1]$ for all $j = 1, \ldots, d$ and $q_{k,k} = 1$ for at least one index $k \in \{1, \ldots, d\}$. Without loss of generality, we assume $\min_{j=1}^{d} a_{j,j} = a_{1,1}$. By the property $\|cx\|_2 = |c|\|x\|_2$, the infimum is always achieved at the boundary, so

$$\inf_{x \in B(0,a)^c} \|A^{-1}\tilde{Q}^{-1}x\|_2 = \inf_{\|x\|_2 = a} \|A^{-1}\tilde{Q}^{-1}x\|_2.$$

The fact that $\inf_{\|x\|_2=a} \|A^{-1}Q^{-1}x\|_2 = a/a_{1,1}$ follows directly from the definition of Q, since

$$A^{-1}Q^{-1}x = A^{-1}(\min_{j=1}^{d} a_{j,j} A^{-1})^{-1}x = \frac{1}{a_{1,1}} A^{-1}Ax = \frac{1}{a_{1,1}}x.$$

It remains to show that a matrix $\tilde{Q}$ that maximises (8) is of the form $\tilde{q}_{k,k} \leq a_{1,1}/a_{k,k}$ for all $k = 1, \ldots, d$. In order for $\tilde{Q}$ to maximise (8), it has to hold that

$$\|A^{-1}\tilde{Q}^{-1}x\|_2 = \left(\sum_{i=1}^{d} \frac{x_i^2}{a_{i,i}^2 \tilde{q}_{i,i}^2} \right)^{\frac{1}{2}} \geq \frac{a}{a_{1,1}}$$

for all $\mathbf{x}$ with $\|\mathbf{x}\|_2 = a$. Checking the inequality for a times the unit vectors, we get the condition $\tilde{q}_{k,k} \le a_{1,1}/a_{k,k}$ for all $k = 1, \ldots, d$. Due to the additional condition that $q_{k,k} = 1$ for at least one index, we need to set $q_{1,1} = 1$. Now, taking $\tilde{q}_{1,1} = 1$ and $\tilde{q}_{k,k} < a_{1,1}/a_{k,k}$ for all $k = 2, \ldots, d$ yields

$$\inf_{\|\mathbf{x}\|_2 = a} \|A^{-1}\tilde{Q}^{-1}\mathbf{x}\|_2 = \left(\sum_{i=1}^d \frac{x_i^2}{a_{i,i}^2 \tilde{q}_{i,i}^2} \right)^{\frac{1}{2}} \le \left(\frac{a^2}{a_{1,1}^2 \tilde{q}_{1,1}^2} \right)^{\frac{1}{2}} = \frac{a}{a_{1,1}}$$

so

$$\inf_{\|\mathbf{x}\|_2 = a} \|A^{-1}\tilde{Q}^{-1}\mathbf{x}\|_2 = \inf_{\|\mathbf{x}\|_2 = a} \|A^{-1}Q^{-1}\mathbf{x}\|_2.$$

Comparing the premium of the reinsurance strategies Q and $\tilde{Q}$, it is easy to obtain $p(Q) < p(\tilde{Q})$ due to the positivity of the vector $\mathbf{p}$. Thus, Q is the quota share reinsurance strategy that maximises (8) with the lowest premium. $\square$

Remark 7. *Theorem 4 can be extended to distributions for which Assumption (A1) hold and, in addition, $\lim_{n \to \infty} h(n\|\mathbf{x}\|_2)/h(n) = a^\alpha$ for $\alpha > 0$. If $\lim_{n \to \infty} h(n\|\mathbf{x}\|_2)/h(n)$ is a constant, a quota share reinsurance does not improve the asymptotic behaviour of the logarithmic ruin probability, and hence Theorem 4 does not hold for lognormal type distributions.*

The optimal matrix Q transforms the set Ω to a d-dimensional ball. Hence, the probability that the reinsured risk process exceeds a threshold in a selected norm has the same asymptotics in all directions. Taking reinsurance defined by the matrix Q reduces the risks of the riskier lines of business to the same level of the less risky line of business.

Remark 8. *If $a_{k,k} = a_{1,1}$ for all $k \in \{1, \ldots, d\}$, Ω describes a d-dimensional ball, and quota share reinsurance does not improve the asymptotic behaviour of the logarithmic ruin probability.*

If the ellipsoid is not orientated along the axes, it is not possible to find an optimal quota share reinsurance, which is defined component-by-component and results in equal asymptotic behaviour of all directions. Depending on the orientation of the ellipsoid, it might be possible to find a quota share reinsurance strategy defined component-by-component that still reduces the risks in the most risky directions and thus reduces the ruin probability.

Example 4. *Consider an insurance company with 3 lines of business that models their risks, the difference between the claim size and the associate net premium, by a random vector $\mathbf{X}$. Assume $\mathbf{X}$ fulfils the assumptions in Theorem 4 with*

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

The optimal QS strategy for the ceding company is then given by the matrix

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}.$$

The fact that $\min_{j=1}^3 a_{j,j} = a_{1,1}$ indicates that the first line of business is the less risky one and thus the insurance company decides to take reinsurance only for the second and third line of business. To minimise its own risks according to (7), the insurance company needs to take reinsurance of at least $1/2$ of the second line of business and $1/3$ of the third line of business. Taking more than the needed share of reinsurance does not reduce the risk, but only increases the premium of the reinsurance strategy.

4.3. Quota Share Risk Sharing of the Reinsurance Company

A reinsurance company is interested in optimising the risk sharing portfolio such that

$$\lim_{n\to\infty} \frac{\log(\mathbb{P}(\|\mathbf{S}_n/n\|_2 > a))}{h(n)} > \lim_{n\to\infty} \frac{\log(\mathbb{P}(\|(I-Q)\mathbf{S}_n/n\|_2 > a))}{h(n)},$$

where Q is a positive, diagonal $d \times d$ matrix with $0 < q_{j,j} < 1$ for all $j = 1,\ldots,d$. The condition $q_{j,j} < 1$ for all $j = 1,\ldots,d$ is due to the natural assumption that the reinsurer offers reinsurance for all lines of business. Comparing the situation with the quota share reinsurance strategy of the ceding company, the same situation leads to a square matrix of smaller dimension since the insurance company wants to get an offer for reinsurance only in the most risky lines of business, and covers the risks of the line of business with the lowest risk itself. The reinsurance company collects the premium

$$p(Q) = \mathbf{1}^T(I-Q)\mathbf{p} = \sum_{j=1}^{d}(1-q_{j,j})p_j$$

and covers the amount $(1-q_{j,j})X^j$ of the jth line of business. The following theorem states the optimising quota share strategy.

Theorem 5. *Let* $\mathbf{X} = R\boldsymbol{\Theta}$, *where* $\boldsymbol{\Theta}$ *fulfils assumption* (A2') *and* $\Omega = \Lambda(\mathbb{S}^{d-1})$ *is defined by the linear transformation* $\Lambda(\mathbf{x}) = A\mathbf{x}$, *where* A *is a diagonal* $d \times d$ *matrix with* $a_{k,k} > 0$ *for all* $k = 1,\ldots,d$. *Additionally, we assume that* R *has a Weibull distribution with parameter* $\beta \in (0,1)$. *Then, the quota share reinsurance defined by the matrix* $Q = I - A^{-1}/c$ *for some* $c > \max_{j=1}^{d} 1/a_{j,j}$ *yields the inequality*

$$\lim_{n\to\infty} \frac{\log(\mathbb{P}(\|\mathbf{S}_n/n\|_2 > a))}{h(n)} > \lim_{n\to\infty} \frac{\log(\mathbb{P}(\|(I-Q)\mathbf{S}_n/n\|_2 > a))}{h(n)}. \qquad (9)$$

Proof. The large deviations principle implies

$$\lim_{n\to\infty} \frac{\log(\mathbb{P}(\|\mathbf{S}_n/n\|_2 > a))}{h(n)} = -\inf_{\mathbf{y}\in B(\mathbf{0},a)^c} \|A^{-1}\mathbf{y}\|_2^{\beta} = -\inf_{\|\mathbf{y}\|_2=a} \|A^{-1}\mathbf{y}\|_2^{\beta}$$

and equivalently

$$\lim_{n\to\infty} \frac{\log(\mathbb{P}(\|(I-Q)\mathbf{S}_n/n\|_2 > a))}{h(n)} = -\inf_{\|\mathbf{y}\|_2=a} \|A^{-1}(I-Q)^{-1}\mathbf{y}\|_2^{\beta}.$$

For the jth unit vector $\mathbf{e}_j$ it holds $\|A^{-1}(a\mathbf{e}_j)\|_2 = a/a_{j,j}$, which results in

$$\inf_{\|\mathbf{y}\|_2=a} \|A^{-1}\mathbf{y}\|_2 \leq \min\left(\frac{a}{a_{1,1}},\ldots,\frac{a}{a_{d,d}}\right) \leq \max\left(\frac{a}{a_{1,1}},\ldots,\frac{a}{a_{d,d}}\right)$$
$$< ca = \inf_{\|\mathbf{y}\|_2=a} \|A^{-1}(I-Q)^{-1}\mathbf{y}\|_2.$$

This proves Inequality (9). □

As in Theorem 4, A is a diagonal matrix which implies that the ellipsoid that defines the distribution of $\mathbf{X}$ is orientated along the axes. The constant c defines the risk share of the reinsurance company and hence the amount that they reassure. The quota share ratio of the reinsurer of the jth line of business is $1 - q_{j,j} = 1/(a_{j,j}c)$ so increasing c reduces the risks for the reinsurer. As in Section 4.2 also the optimal risk sharing strategy of the reinsurance company includes bigger ratio for the lines of business with smaller risks.

The ceding company as well as the reinsurance company optimise their risks by taking or offering reinsurance that maps their share of the initial ellipsoid to a ball. The optimising strategy of the ceding company yields

$$AQ = \min_{i=1}^{d} a_{i,i} I$$

whereas the optimising strategy of the reinsurance company results in

$$A(I - \tilde{Q}) = \frac{1}{c} I,$$

where $c > \max_{i=1}^{d} 1/a_{i,i}$ or $1/c < \min_{i=1}^{d} a_{i,i}$. The radius of the ball defined by AQ is $\sqrt{1/\min_{i=1}^{d} a_{i,i}}$ and therefore smaller than the radius of the ball defined by $A(I - \tilde{Q})$ which is $\sqrt{c}$. Increasing c increases the radius of the ball generated by $A(I - \tilde{Q})$. In general, a bigger radius implies smaller risks for the insurance or reinsurance company. Increasing c reduces the share of the reinsurance company, and therefore also the risks of the reinsurance company. The minimum $\min_{i=1}^{d} a_{i,i}$ indicates the line of business with the lowest risk.

5. Conclusions

The assumptions of the main result require that the support of the random vectors is, asymptotically, the whole space. In particular, the components of increments are asymptotically dependent. However, the studied model admits more flexibility than many typical models in the sense that it does not require uniformly distributed random vectors on ellipses. This makes it possible to derive asymptotics for a wide class of zero-mean random walks. The general case with the non-zero expectation can be studied by centring the increments.

In economics and finance, quantification of tail risk plays a central role, because rare events can have a large impact on the operation of a financial entity. The derived large deviations principle quantifies, asymptotically, the probabilities of rare events of random walks. This makes it possible to derive further results in applications such as the one presented in the QS optimisation method. For financial data that fulfils the assumptions, the large deviations principle can be used to establish optimal bounds on the probabilities of rare events where the decay rate to zero of probabilities is not exponential. The derived results give theoretical estimates for probabilities that might be difficult to analyse numerically.

Further research could investigate the convergence rate of the tail probabilities in order to give a suitable threshold after which the large deviations principle gives a sufficiently accurate estimate. Direct approximation of probabilities by Monte Carlo simulation suggests that, to calculate the ratio in Theorem 2.1, at least one of n or a must be quite large. With a finite sample, the challenge is that the interesting events are tail events of which there are only few observations. It seems that a smaller n can be used for the heavier distributions in the studied class than for the ones with lighter tails.

In addition, it would be natural to ask how the set Ω can be deduced from observed data and if the large deviations principle holds even if Ω is, say, not the boundary of a convex set.

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

References

Asmussen, Søren, and Jaakko Lehtomaa. 2017. Distinguishing log-concavity from heavy tails. *Risks* 5: 10. doi:10.3390/risks5010010.

Bazhba, Mihail, Jose Blanchet, Chang-Han Rhee, and Bert Zwart. 2020. Sample path large deviations for lévy processes and random walks with weibull increments. *Annals of Applied Probability* 30: 2695–739.

Bingham, Nicholas, Charles Goldie, and Jozef Teugels. 1989. *Regular Variation*. Cambridge: Cambridge University Press.

Cline, Daren, and Sidney Resnick. 1992. Multivariate subexponential distributions. *Stochastic Processes and Their Applications* 42: 49–72. doi:10.1016/0304-4149(92)90026-M.

Dembo, Amir, and Ofer Zeitouni. 1993. *Large Deviations Techniques and Applications*. Boston: Jones and Bartlett Publishers.

Denisov, Denis, Antonius Dieker, and Vsevolod Shneer. 2008. Large deviations for random walks under subexponentiality: The big-jump domain. *The Annals of Probability* 36: 1946–91. doi:10.1214/07-AOP382.

Embrechts, Paul, Claudia Klüppelberg, and Thomas Mikosch. 1997. *Modelling Extremal Events*. Volume 33 of Applications of Mathematics (New York). Berlin: Springer. doi:10.1007/978-3-642-33483-2.

Foss, Sergey, Dmitry Korshunov, and Stan Zachary. 2011. *An Introduction to Heavy-Tailed and Subexponential Distributions*. Springer Series in Operations Research and Financial Engineering. New York: Springer. doi:10.1007/978-1-4419-9473-8.

Gantert, Nina. 1998, 07. Functional Erdős-Renyi laws for semiexponential random variables. *Annals of Probability* 26: 1356–69. doi:10.1214/aop/1022855755.

Gantert, Nina, Kavita Ramanan, and Franz Rembart. 2014. Large deviations for weighted sums of stretched exponential random variables. *Electronic Communications in Probability* 19: 1–14. doi:10.1214/ECP.v19-3266.

Hägele, Miriam. 2020. Precise asymptotics of ruin probabilities for a class of multivariate heavy-tailed distributions. *Statistics & Probability Letters* 166: 108871. doi:10.1016/j.spl.2020.108871.

Hardy, Mary. 2001. A regime-switching model of long-term stock returns. *North American Actuarial Journal* 5: 41–53.

Hult, Henrik and Filip Lindskog. 2002. Multivariate extremes, aggregation and dependence in elliptical distributions. *Advances in Applied Probability* 34: 587–608.

Hult, Henrik and Filip Lindskog. 2006. On regular variation for infinitely divisible random vectors and additive processes. *Advances in Applied Probability* 38: 134–48. doi:10.1239/aap/1143936144.

Hult, Henrik, Filip Lindskog, Thomas Mikosch, and Gennady Samorodnitsky. 2005. Functional large deviations for multivariate regularly varying random walks. *The Annals of Applied Probability* 15: 2651–80. doi:10.1214/105051605000000502.

Jensen, Mark and John Maheu. 2018. Risk, return and volatility feedback: A bayesian nonparametric analysis. *Journal of Risk and Financial Management* 11: 52. doi:10.3390/jrfm11030052.

Klüppelberg, Claudia, Gabriel Kuhn, and Liang Peng. 2007. Estimating the tail dependence function of an elliptical distribution. *Bernoulli* 13: 229–51. doi:10.3150/07-BEJ6047.

Lehtomaa, Jaakko. 2017. Large deviations of means of heavy-tailed random variables with finite moments of all orders. *Journal of Applied Probability* 54: 66–81. doi:10.1017/jpr.2016.87.

Lehtomaa, Jaakko, and Sidney Resnick. 2020. Asymptotic independence and support detection techniques for heavy-tailed multivariate data. *Insurance: Mathematics and Economics* 93: 262–77. doi:10.1016/j.insmatheco.2020.05.002.

Mikosch, Thomas, and Igor Rodionov. 2021. Precise Large Deviations for Dependent Subexponential Variables. *Bernoulli* 27: 1319–47.

Mikosch, Thomas, and Olivier Wintenberger. 2016. A large deviations approach to limit theory for heavy-tailed time series. *Probability Theory and Related Fields* 166: 233–69. doi:10.1007/s00440-015-0654-4.

Nyrhinen, Harri. 2009. On large deviations of multivariate heavy-tailed random walks. *Journal of Theoretical Probability* 22: 1–17. doi:10.1007/s10959-008-0194-2.

Omey, Edward 2006. Subexponential distribution functions in $\mathbf{R}^d$. *Journal of the Australian Mathematical Society* 138: 5434–49. doi:10.1007/s10958-006-0310-8.

Samorodnitsky, Gennady, and Julian Sun. 2016. Multivariate subexponential distributions and their applications. *Extremes* 19: 171–96. doi:10.1007/s10687-016-0242-8.

Tegnér, Martin, and Rolf Poulsen. 2018. Volatility is log-normal—But not for the reason you think. *Risks* 6: 46. doi:10.3390/risks6020046.

Teugels, Jozef. 1975. The class of subexponential distributions. *The Annals of Probability* 3: 1000–11.