Dependence structure of risk factors and diversification effects

Chen Zhou[1]

Economics and Research Division
De Nederlandsche Bank
P.O. Box 98, 1000AB Amsterdam
The Netherlands

Abstract

In this paper, we study the aggregated risk from dependent risk factors under the multivariate Extreme Value Theory (EVT) framework. We consider the heavy-tailness of the risk factors as well a non-parametric tail dependence structure. This allows a large scope of models on the dependency. We assess the Value-at-Risk of a diversified portfolio constructed from dependent risk factors. Moreover, we examine the diversification effects under this setup.

Key words:
Aggregated risk, diversification effect, multivariate Extreme Value Theory
JEL: G11, C14

---

1. Introduction

Diversification is considered to be a major instrument in risk management. At a micro level, financial institutions use portfolios to diversify away the risk of individual risk factors. Diversification is applied to the downside risk of financial assets as well as other risk factors such as credit risk, operational risk, etc. At an industrial level, a large scope of recent financial innovations are essentially equivalent to a risk sound package with diversified risk factors. At a macro level, financial authorities such as central banks

Email address: `C.Zhou@DNB.nl` (Chen Zhou)

[1]Views expressed are that of the author and do not necessarily reflect official positions of De Nederlandsche Bank.

Preprint submitted to The Australasian Finance & Banking Conference August 18, 2009

monitor and supervise the diversified financial system in order to maintain financial stability. Therefore, it is necessary to have a general framework for evaluating aggregated risk and then assessing the diversification effect.

Having an accurate evaluation on aggregated risks is never an easy job. This requires caution on at least two stylized facts: firstly, each individual risk factor may have its own risk profiles such as heavy-tailness; secondly, among the risk factors, the dependence structure can be rather complex and thus difficult to model. Therefore, modeling the individual risk factors and investigating the dependence structure among them are the two major tasks in risk analysis and management. In this study, we try to approach both these two issues under the Extreme Value Theory (EVT) setup.

Initiated by Markowitz (1952), the classical mean-variance approach considers Gaussian (normal) distributed risks and takes the variance as a risk measure. The dependence structure of risk factors is modeled by the multivariate normal distribution. Under such a setup, the first generation of risk analysis and diversification effect study concentrates on the variance analysis. However, studies on economics and finance in the last 40 years observed quite some extreme phenomena which can not be explained by the classical Gaussian approach, see, e.g. Fama (1965).

The extreme phenomena motivate extensive studies on heavy-tailed behavior of the risk factors. A first taste on heavy-tailed model was the stable distribution which is a parametric model consisting of four parameters: the characteristic component, the scale parameter, the location parameter and the skewness parameter. As summarized in Fama and Miller (1972, Chap 6, Sec V), the symmetric stable distribution (with skewness 0) was applied to model the asset returns. In Fama and Miller's symmetric stable framework, by considering the the characteristic exponent α to be constant among all risk factors, the scale parameter is considered as an analog of the variance in the Gaussian framework which measures the risk. They further study the diversification effect with assuming independency among risk factors. By analyzing the scale parameter, it was concluded that the diversification effects can only be observed for $\alpha > 1$. When $\alpha = 1$, diversification has no effect. When $\alpha < 1$, increased diversification causes an increasing risk. This is recognized as a negative diversification effect. Note that the symmetric stable framework is still a full parametric approach similar to the Gaussian approach. Since in risk assessment, the major interest is on the tail of a distribution, full parametric models are not necessary, and also may not be preferable.

Starting from Roy (1952) and Arzac and Bawa (1977), safety-first criterion gets more and more attention in risk management. Value-at-Risk (VaR) which considers only the downside risk becomes a well accepted risk measure, since it does not rely on specific distribution assumption. For example, within the Gaussian framework, Gouriéroux et al. (2000) investigated the sensitivity of VaR. In fact, when fixing mean as zero, in Gaussian framework, VaR is directly connected with variance. The variance analysis is in fact a VaR approach. Similarly, in Fama and Miller's symmetric stable framework, when the characteristic component α is fixed, VaR is directly associated to the scale parameter. Hence, the study on the scale parameter in symmetric stable framework is also a VaR approach.

Recent developments in Extreme Value Theory (EVT) creates the possibility to model the tail of non-Gaussian returns without imposing assumptions on moderate level of the distribution. The EVT model focuses on the tail of a distribution function instead of the entire structure of the distribution. It is a semi-parametric model in the following sense. On the one hand, the tail of a heavy-tailed distribution is asymptotically Pareto distributed which is a parametric model with a shape parameter named the *tail index* and a *scale function*. On the other hand, the EVT model does not impose any parametric assumption on the moderate level of a distribution function. Such a semi-parametric feature is less restrictive than a specific parametric approach such as the normal distribution or the stable distribution. Notice that the stable distribution used by Fama and Miller belongs to the EVT model. Moreover, the characteristic exponent α of a symmetric stable distribution is equal to its tail index. Hence, the EVT model can be viewed as a generalization of the stable model in Fama and Miller's study. For detailed discussion on tail distribution modeling, see e.g. Embrechts et al. (1997). For applying the EVT model in finance, see e.g. Jansen and de Vries (1991).

Based on the EVT setup, the VaR can be approximately calculated thanks to the explicit Pareto distribution. As application in evaluating aggregated risks, Jansen et al. (2000) started the empirical exercises on portfolio selection under the VaR approach, followed by Jansen (2001) and Susmel (2001). In these studies, independency across risk factors are assumed, because datasets employed in these papers are the returns on market indices which are either markets across different countries or different markets for different assets such as equities and bonds. From the the diversification effect in symmetric stable framework, one may conjecture that the diversification effect for independent heavy-tailed risk factors depends on the tail index. This conjecture is confirmed in Hyung and de Vries (2002).

The independency assumption is not always valid, for example, when evaluating aggregated risks from assets in a specific market, or assessing the systemic risk of a financial system, it is necessary to study the dependence structure among risk factors.

To incorporate the dependency, specific models in modeling the dependence structure were considered in literature. In a series of papers by Hyung and de Vries (2002, 2005, 2007), the portfolio diversification problem is studied by assuming specific dependence structure such as the single-index factor market model, Capital Asset Pricing Model (CAPM). Hyung and de Vries (2002) confirm Fama and Miller's conclusion in a dependent setup with the specific dependence structure, CAPM. Since the CAPM model only creates a relatively simple dependence structure for modeling the dependence, it is still an open question for the diversification effect in a general dependence structure setup. We also refer to Ibragimov and Walden (2008) for more discussion.

The developments in multivariate EVT offers a large scope of models in tail dependency modeling. Again, the models are only on the tail part, which do not impose assumptions in the moderate dependence. Moreover, the tail dependence structure modeled by multivariate EVT can be even non-parametric, see, e.g. Huang (1992). For empirical applications, Hartmann et al. (2004) applied multivariate EVT to test the tail dependency among the stock markets of the G-5 countries as well as the bond markets.

The tail dependence structures in the multivariate EVT setup contains most of the parametric dependence models that have been used in literature. For example, the CAPM model used in Hyung and de Vries (2002), the archimedean model used in Embrechts et al. (2009) all belongs to the multivariate EVT framework. For application, Poon et al. (2004) considers the multivariate EVT setup for a portfolio selection problem.

In this paper, we study the aggregated risk from dependent risk factors under the multivariate EVT setup. The individual risk factors are modeled by heavy-tailed distributions while their tail dependence structure comes from multivariate EVT. We do not impose any parametric assumption ex-ante. We first link the VaR estimation to the scale function and the tail index of a heavy tailed distribution. In case the tail indices for all risk indicators are at different levels, we prove that within a constructed portfolio, the component involving the risk factor with the minimal tail index dominates the others in terms of tail risks, regardless the dependence structure. Then we look at the case in which the tail indices are at a constant level across all risk factors. In this case, we provide an explicit formula on how to calculate the aggregated risk. From the calculation, we confirm similar results as Fama and Miller, however, with taking into account dependency: the diversification effect is observed if and only if the tail indices are higher than 1 and the risk factors are not completely tail dependent. On contrast to Fama and Miller's result, in case of completely tail dependence, the diversification effect does not exist even if the tail indices are higher than 1. We demonstrate different examples on calculating aggregated risk under different dependence models as well as their corresponding diversification effects.

In the case when the tail indices are all equal and lower than 1, diversification dose not lead to risk reduction. Therefore, for risk control, one should compare the marginal VaRs and searching for the minimal risk. In case the marginal risks are at the same level, for risk managing purpose, we further investigate the dependence structure, and introduce the probability of domination to evaluate the linkage between each marginal risk factor and the systemic risk.

We remark that the sub- and superadditivity of VaR was studied in Embrechts et al. (2008) under the multivariate EVT setup. They considers homogeneous individual risk factors in the sense that the marginals share the same tail risk. Our approach is thus a generalization by allowing different marginal tail indices or scale functions. Moreover, as an extra contribution, our probability of dominance provides risk management suggestions to deal with the case when the tail indices are lower than 1. Such a situation is applied to (re)insurance when modeling catastrophic risks, see Ibragimov et al. (2008).

The paper is organized as follows. Section 2 reviews the Extreme Value Theory. Section 3 shows how to calculate the aggregated risk under the multivariate EVT setup. Section 4 discusses the diversification effects. Section 5 discusses the risk management implication from our theoretical result, in particular, when the marginal tail indices are at the same level and lower than 1, the probability of domination is introduced as a risk measurement. Section 6 concludes the paper.

2. Extreme Value Theory (EVT)

In this section, we review the Extreme Value Theory (EVT) framework in both univariate and multivariate setups. The univariate EVT provides a general approach in calculating Value-at-Risk, while the multivariate EVT provides a large scope of models on the dependence of extreme events.

2.1. Univariate EVT and Value-at-Risk (VaR)

Univariate EVT makes assumption on the tail of a distribution function. We only consider the heavy-tail case. Let X denote the loss generated from a certain risk factor, for instance, if R is the return of a certain asset, we could take $X = -R$. Denote F as the distribution function of X. Suppose X follows a heavy-tailed distribution, i.e. we have that

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-\alpha}, \tag{1}$$

where $\alpha > 0$ is called the *tail index*. It implies that $1-F(t) = t^{-\alpha}l(t)$, where $l(t)$ is a slowly varying function which is defined as

$$\lim_{t\to\infty} \frac{l(tx)}{l(t)} = 1.$$

In a narrow case where $l(t)$ is almost a constant, i.e. $l(t) \to A$ as $t \to \infty$, the tail of the distribution function of X has the following representation

$$P(X \geq x) = Ax^{-\alpha}[1 + o(1)],$$

as $x \to \infty$. In other words, the tail distribution of X is approximately Pareto distribution. This is the same setup as in Hyung and de Vries (2002, 2005).

Denote $VaR(\delta)$ as the Value-at-Risk of X at tail probability level δ, i.e. $P(X > VaR(\delta)) = \delta$. From the EVT setup, we have that

$$\delta = (VaR(\delta))^{-\alpha} l(VaR(\delta)),$$

which implies that

$$VaR(\delta) = \left(\frac{a(\delta)}{\delta}\right)^{1/\alpha}, \tag{2}$$

where $a(\delta) = l(VaR(\delta))$ is called the *scale function*. It can be verified that $a(\delta)$ is a slowly varying function as $\delta \to 0$. Thus, for small δ, $a(\delta)$ can be regarded as a constant function. In case $l(t) \sim A$, we get $a(\delta) \sim A$, as $\delta \to 0$.

In order to statistically estimate the VaR, it is necessary to estimate the tail index α as well as the scale function a. The estimation of the tail index is a major issue in extreme value statistics. Suppose we have a sample of observations $X_1, X_2, \cdots, X_n$. By ranking them, we get the order statistics $X_{n,1} \leq X_{n,2} \leq \cdots \leq X_{n,n}$. Hill (1975) proposed the so-called *Hill estimator* in estimating the tail index α as

$$\hat{\alpha}_H = \left(\frac{1}{k} \sum_{i=1}^{k} \log X_{n,n-i+1} - \log X_{n,n-k} \right)^{-1},$$

where $k = k(n)$ is a suitable intermediate sequence such that $k(n) \to \infty$ and $k(n)/n \to 0$ as $n \to \infty$. From the estimator, we observe that only k high order statistics from the top are used in the estimation.

Applying (2) with $\delta = k/n$, we get that

$$VaR(k/n) = \left(\frac{a(k/n)}{k/n} \right)^{1/\alpha}.$$

Since a remains at a constant level when δ approximates zero, for small δ, $a(\delta)$ can be well approximated by $a(k/n)$, together with (2), we have

$$\frac{VaR(\delta)}{VaR(k/n)} \approx \left(\frac{k/n}{\delta} \right)^{1/\alpha}.$$

Notice that the non-parametric estimation of $VaR(k/n)$ is $X_{n,n-k}$. We finally get a reasonable estimator of $VaR(\delta)$ as

$$\widehat{VaR}(\delta) = \left(\frac{k/n}{\delta} \right)^{1/\hat{\alpha}} X_{n,n-k}.$$

This estimator was proposed as a quantile estimator in Weissman (1978).[2]

In the case $a(\delta) \sim A$, instead of using a scale function, we could call A, the *scale*. Then we have the estimator of A as

$$\hat{A} = \hat{a}(k/n) = \frac{k}{n} \left(\widehat{VaR}(k/n) \right)^{\hat{\alpha}} = \frac{k}{n} \left(X_{n,n-k} \right)^{\hat{\alpha}}.$$

We could link the estimator of VaR to the estimator of the scale A as

$$\widehat{VaR}(\delta) = \left(\frac{\hat{A}}{\delta} \right)^{1/\hat{\alpha}}.$$

[2]Notice that the definition of VaR is exactly the same as the quantile of a certain distribution function.

Hence the estimation of VaR is determined by the estimations on the tail index α and the scale function $a(k/n)$. This can be viewed as a solution from the Pareto approximation:

$$\delta = P(X > VaR(\delta)) \approx A\,(VaR(\delta))^{-\alpha}\,.$$

Within the univariate EVT setup, although both the tail index and the scale function (or scale parameter) play a role in VaR evaluation, the tail index plays a more dominated role for extreme risks. Suppose we have two risk factors X and Y with tail indices α_1 and α_2, scale functions $a_1(\delta)$ and $a_2(\delta)$ respectively. If $\alpha_1 > \alpha_2$, then $1/\alpha_2 - 1/\alpha_1 > 0$. Hence, we have that

$$\lim_{\delta\to 0} \frac{VaR_X(\delta)}{VaR_Y(\delta)} = \lim_{\delta\to 0} \delta^{1/\alpha_2 - 1/\alpha_1} \frac{a_1(\delta)^{1/\alpha_1}}{a_2(\delta)^{1/\alpha_2}} = 0.$$

Here we use the fact that $\frac{a_1(\delta)^{1/\alpha_1}}{a_2(\delta)^{1/\alpha_2}}$ is a slowly varying function as $\delta \to 0$. This implies that X is less risky than Y. Hence, the risk factor with higher tail index exhibits less risk in extremes.

In the case that the tail indices are equal, we have $\alpha_1 = \alpha_2 = \alpha$, which implies that

$$\lim_{\delta\to 0} \frac{VaR_X(\delta)}{VaR_Y(\delta)} = \lim_{\delta\to 0} \left(\frac{a_1(\delta)}{a_2(\delta)}\right)^{1/\alpha}.$$

Thus, comparing the scale functions is important for the comparison of the VaRs. Here we present two properties of the scale function under equalized tail indices. Suppose the tail indices of X and Y) are both α, as $\delta \to 0$:
1) $a_{cX}(\delta) \sim c^\alpha a_X(\delta)$, for all $c > 0$;
2) $a_{X+Y}(\delta) \sim a_X(\delta) + a_Y(\delta)$, if X and Y are independent.
The second property follows from the Feller theorem, see Feller (1971, section VIII, 8). Paralleled to this, when we have the scale parameters A_X and A_Y, we have:
1) $A_{cX} = c^\alpha A_X$, for all $c > 0$;
2) $A_{X+Y} = A_X + A_Y$, if X and Y are independent.

2.2. Multivariate EVT and tail dependence

Multivariate EVT considers not only the tail behavior of each individual risk factor, but also the extreme co-movements among them. The setup splits marginal risks from the dependence structure as follows. Let $X = (X_1, \cdots, X_d)$ denote the losses of d individual risk factors. Each risk factor

X_i follows the univariate EVT setup with its own tail index α_i and scale function $a_i(t)$. Multivariate EVT models the dependence of the risk factors in extremal situation, i.e. the joint probability of extreme co-movements. For any $x_1, x_2, \cdots, x_d > 0$, as $\delta \to 0$, we assume that

$$\frac{P\left(X_1 > VaR_1(x_1\delta) \text{ or } \cdots \text{ or } X_d > VaR_d(x_d\delta)\right)}{\delta} \to L(x_1, x_2, \cdots, x_d). \quad (3)$$

where VaR_i denotes the Value-at-Risk of X_i, and L is a finite positive function. It is clear that L function characterize the co-movement of extreme events. We refer to de Haan and Ferreira (2006) for the properties of the L function.

L function is well connected with the modern instrument on dependence modeling–copula. Denote the joint distribution function of X as $F(x_1, \cdots, x_d)$ while the marginal distributions are denoted as $F_i(x_i)$ for $i = 1, \cdots, d$. Then there exists a unique distribution function $C(x_1, \cdots, x_d)$ on $[0, 1]^d$ such that

$$F(x_1, \cdots, x_d) = C(F_1(x_1), \cdots, F_d(x_d)),$$

where all marginal distribution of C are uniform distribution. C is called the *copula*. By decomposing F into marginal distributions and copula, we separate the marginal information and the dependence structure. Condition (3) is equivalent to the following relation. For any $x_1, x_2, \cdots, x_d > 0$, as $\delta \to 0$,

$$\frac{1 - C(1 - \delta x_1, \cdots, 1 - \delta x_d)}{\delta} \to L(x_1, x_2, \cdots, x_d). \quad (4)$$

Hence L function characterize a limit behavior of the copula C, i.e. it captures the tail behavior of the copula. This implies that L is independent from marginal information, thus marginal risks. Moreover, L-function does not contain dependence information at a moderate level. We call such a structure the *tail dependence structure*.

There are several ways to further characterize L function into an explicit expressions. Here we use one which serves our purpose, see de Haan and Ferreira (2006). Let H be any probability measure on

$$W = \{w = (w_1, \cdots, w_d) : w_1 + \cdots + w_d = 1, w_i \geq 0, i = 1, 2, \cdots, d\},$$

such that

$$\int_W w_i H(dw) = \frac{1}{d}, \quad \text{for } 1 \leq i \leq d.$$

Any qualified H leads to a L function as follows,

$$L(x_1, \cdots, x_d) = d \int_W \max_{1 \le i \le d} x_i w_i H(dw).$$

Conversely, any L function has the above representation with a suitable H. H is called the *spectral measure on* W.

Since the spectral measure is a rather abstract concept, we exhibit a few examples to show how it is connected to the tail dependence structure. We consider the simplest case $d = 2$.

Example 1) Suppose X_1 and X_2 are independent. Then the copula $C(x_1, x_2) = x_1 x_2$. From (4), we get that $L(x_1, x_2) = x_1 + x_2$. It is not difficult to verify that such a L function corresponds to a spectral measure H that concentrates its measure on two points $\{(1, 0), (0, 1)\}$ with probability $1/2$ each. Conversely, we observe that a H measure concentrates on the corners of W corresponds to the case of two idiosyncratic risks.

Example 2) Suppose U, V and W are independent and identically distributed (i.i.d.) with the standard Fréchet distribution function $\exp\{-1/x\}$ for $x > 0$. We define $X_1 = aU \vee (1-a)V$ and $X_2 = bU \vee (1-b)W$ for some $0 \le a, b \le 1$. The choice of the standard Fréchet distribution is simply for calculation convenience: both X_1 and X_2 follow the standard Fréchet distribution. Notice that the marginal distribution is independent from the L function and the spectral measure H. Our choice on the marginal distribution is not essential. With such a setup, the two risk factors shares a common risk U while each has its own idiosyncratic risk as V and W respectively. It is not difficult to calculate from (3) that $L(x_1, x_2) = ax_1 \vee bx_2 + (1-a)x_1 + (1-b)x_2$. One can then verify that such a L function corresponds to a spectral measure H that assigns positive measure on only three points $\{(a/(a+b), b/(a+b)), (1, 0), (0, 1)\}$ with probability $(a+b)/2$, $(1-a)/2$ and $(1-b)/2$, respectively. The measures on the two corner points $(1, 0)$ and $(0, 1)$ can be viewed as correspondences to the idiosyncratic risks while the measure on the inner point $(a/(a+b), b/(a+b))$ corresponds to the common risk. We remark that when $a = b = 0$, it turns to example 1), where there is no measure on any inner point. Correspondingly there is no common risk. This is called the *tail independence case*. Moreover, when $a = b = 1$, there is no measure on the corner points, while all the measures concentrate on the inner point $(1/2, 1/2)$. Hence there is no idiosyncratic risk. Back to our model setup, in this case, we get that $X_1 = X_2 = U$. Indeed, the two risk factors are completely dependent as we expected from the H measure. This is called the *completely tail dependent case.*

From the above two examples, we observe that for a H measure on W, the points on the corner points of W correspond to idiosyncratic risks, while the inner points correspond to common risks. When $d > 2$, this still holds, however, with more complicated structure: for example, the H measure may assign measures to boundary points of W but not the corners. We take $d = 3$ to demonstrate this situation. In this case, $W = \{w = (w_1, w_2, w_3) : w_1 + w_2 + w_3 = 1, w_i \geq 0, i = 1, 2, 3\}$, H measure could assign positive measures to points with $w_3 = 0$ while $w_1, w_2 > 0$. Such a point can be viewed as correspondence to a certain risk which influences X_1 and X_2 but not X_3.

All in all, we observe that H contains the information of the tail dependence structure. Hence, a careful study on the spectral measure H helps to understand the tail dependence structure of the risk factors.

In the above examples, H is always a discrete measure on finite points. In fact, as the number of potential independent risk components goes to infinite, H can be a discrete measure on infinite points, or even a continuous measure on W. In statistical inference, due to the finite dataset, our estimated H measure is always a discrete measure on finite points. For the statistical estimation on H, we refer to Huang (1992).

3. Aggregated risk with dependent risk factors

In Section 2, we discussed the univariate and multivariate EVT on modeling individual risk and tail dependence structure. We combine them to study aggregated risk with dependent risk factors.

Consider d dependent risk factors $X = (X_1, X_2, \cdots, X_d)$. We construct a portfolio

$$P = \sum_{i=1}^{d} c_i X_i$$

for some $c_1, \cdots, c_d > 0$. Since we study the tail behavior of the aggregated risk of P, it is only necessary to model the tail behavior of the risk factors. Following EVT, the risk profile of X can be decomposed into three parts: marginal tail indices, marginal scale functions and the L function or the spectral measure H which captures the tail dependence structure. We show that the tail property of P is determined by those three parts.

Suppose the marginal tail indices are $\alpha_1, \alpha_2, \cdots, \alpha_d > 0$, and the marginal scale functions are $a_1(\delta), a_2(\delta), \cdots, a_d(\delta)$. Without loss of generality, we assume that $\alpha_1 = \min_{1\leq i\leq d} \alpha_i$.

Firstly, we assume that α_1 is the unique minimum tail index. Then we have the following proposition. The proof is postponed to the Appendix.

Proposition 3.1. *Under the multivariate EVT setup, if α_1 is the unique minimum tail index, P has a regularly varying tail with tail index α_1 and a scale function $a_P(\delta)$ satisfying $a_P(\delta) \sim c_1^{\alpha_1} a_1(\delta)$ as $\delta \to 0$.*

From Proposition 3.1, we observe that, when constructing a portfolio based on heavy-tailed risk factors, the tail index of the portfolio is dominated by the minimal marginal tail index. Moreover, the scale function is also dominated by the componnet containing the risk factor with the minimal marginal tail index. We remark that the dependence structure H does not play a role.

Proposition 3.1 implies that in order to diversify away the risk, the optimal solution is to consider those risk factors which share the maximal marginal tail index. By doing that, it ensures the highest possible value of the portfolio tail index. If the one with the maximal marginal index is unique, it is the one with the minimal tail risk. However, when the maximal tail index is not unique, i.e. it is shared by a few risk factors, the diversification problem remains open: we need to evaluate a portfolio based on those risks which share the same tail indices as the maximum. They still follows the multivariate EVT setup, only the dimension might be reduced. Without loss of generality, we assume $\alpha_1 = \cdots = \alpha_d = \alpha$ in the rest of this section.

The following theorem deals with the aggregated risk in case all marginal indices are equal. The proof is in Appendix.

Theorem 3.1. *Suppose d risk factors $X = (X_1, \cdots, X_d)$ follow a d-dimensional EVT setup with equal marginal tail indices, α, marginal scale functions $a_i(\delta)$ for $1 \leq i \leq d$ and spectral measure H on W. Suppose all the marginal functions are comparable, i.e. there exists a function $a(\delta)$ such that $\frac{a_i(\delta)}{a(\delta)} \to m_i > 0$ as $\delta \to 0$. Without loss of generality, we can take $a(\delta) = \sum_{i=1}^d a_i(\delta)$. In that case $\sum_{i=1}^d m_i = 1$.*

For any positive constants $c_1, \cdots, c_d$, for a portfolio $P = \sum_{i=1}^d c_i X_i$, it *must have tail index α, and its scale function $a_P(\delta)$ satisfies*

$$\lim_{\delta\to 0}\frac{a_P(\delta)}{a(\delta)} = d\int_W \left(\sum_{i=1}^d c_i(m_i w_i)^{1/\alpha}\right)^\alpha H(dw).$$

Remark 3.1. *When $\alpha = 1$, by considering the property of H, the result of the theorem is simplified as*

$$a_P(\delta) \sim a(\delta)\sum_{i=1}^d c_i m_i. \sim \sum_{i=1}^d c_i a_i(\delta)$$

Hence, the scale function of the portfolio does not depend on the dependence structure. It is at the same level as the sum of the weighted marginal scale functions.

With the multivariate EVT setup, i.e. a general dependence structure assumption, Theorem 3.1 shows how to calculate the portfolio scale function, when the marginal tail indices are all equal across the risk factors. Since the tail index is at a constant level, the risk measured by VaR is now determined by the scale functions. Precisely, VaR can be calculated from (2) when the scale function is given by Theorem 3.1.

Here we give a few examples in which the dependence structure is parametrically specified. We calculate the scale function of the aggregated portfolio.

Example 3.1. *The individual risk factors are tail independent.*[3]

In this case, H measure only assigns positive measures to the corner points of W. Precisely, H is a discrete measure which concentrates its measure on d points $(1,0,\cdots,0),(0,1,\cdots,0),\cdots,(0,0,\cdots,1)$ with probability $1/d$ each. Thus, from Theorem 3.1, we get that as $\delta\to 0$,

$$a_P(\delta) \sim a(\delta)\sum_{i=1}^d \left(c_i m_i^{1/\alpha}\right)^\alpha = a(\delta)\sum_{i=1}^d c_i^\alpha m_i.$$

From the first property of the scale function in Section 2, we have that

$$a_{c_iX_i}(t) \sim c_i^\alpha a_i(\delta) \sim c_i^\alpha m_i a(\delta).$$

[3]Notice that the risk factors are independent implies that they are tail independent, not vice versa.

Hence we conclude that in the tail independence case, the scale function of the aggregated risk is at the same level as the sum of the scale functions of its each components. This is a generalization of the second property of the scale function in Section 2.

Example 3.2. *The individual risk factors are completely tail dependent.*[4]

In this case, H measure concentrates all its measure on a single inner point $(1/d, 1/d, \cdots, 1/d)$. Thus, from Theorem 3.1, we get that as $\delta \to 0$,

$$a_P(\delta) \sim a(\delta) \left(\sum_{i=1}^{d} c_i m_i^{1/\alpha} \right)^{\alpha}.$$

Example 3.3. *The risk factors follow the CAPM model.*

Following Hyung and de Vries (2002), we consider the single factor model as follows. Suppose $X_i = \beta_i R + R_i$ for $1 \leq i \leq d$, where R is a common risk factor with tail index α and scale function $a_R(\delta)$ and R_is are the idiosyncratic risk factors with the same tail index α' and scale functions $b_i(\delta)$, for $1 \leq i \leq d$.[5]

For simplicity, Hyung and de Vries (2002) considered the case that as $\delta \to 0$, $a_R(\delta)$ and $b_i(\delta)$ are at constant level. We follow such a setup, and denote the constant levels as a_R and b_i, respectively.

When $\alpha < \alpha'$, it is clear that all risk factors are dominated by the common risk R, thus it is the completely tail dependent case as in Example 3.2. When $\alpha > \alpha'$, each risk factor is dominated by its own idiosyncratic risk, thus it is the completely tail independence case as in Example 3.1.

We consider the case $\alpha = \alpha'$. We assume symmetry as $\beta_i = \beta$ and $b_i = b$, for $1 \leq i \leq d$. Following the definition of L-function, it can be verified that

$$L(x_1, \cdots, x_d) = \max_{1 \leq i \leq d} \frac{\beta^\alpha a_R}{\beta^\alpha a_R + b} x_i + \sum_{i=1}^{d} \frac{b}{d(\beta^\alpha a_R + b)} x_i.$$

Hence, $X = (X_1, \cdots, X_d)$ follows multivariate EVT setup, with corresponding H measure assigning measures to $d+1$ points: $(1, 0, \cdots, 0), (0, 1, \cdots, 0), \cdots, (0, 0, \cdots, 1)$

[4]The risk factors are completely dependent implies that they are completely tail dependent, not vice versa.

[5]When evaluating the risk of equities, R is normally considered as the market risk.

and $(1/d, \cdots, 1/d)$. The measure assigning to $(1/d, , \cdot, 1/d)$ is $\beta^\alpha a_R/(\beta^\alpha a_R + b)$, while the measures assigning to the rest d points are all equal to $b/d(\beta^\alpha a_R + b)$.

The setup of the H measure reflects the fact that there exist one common risk for all the risk factors, and idiosyncratic risk for each risk factor.

From the properties of the scale function in Section 2, we have that for all $1 \leq i \leq d$, as $\delta \to 0$, the marginal scale function satisfies

$$a_i(\delta) \sim \beta^\alpha a_R + b := a.$$

From Theorem 3.1, we can get the scale function of portfolio P as

$$a_P(\delta) \sim d \sum_{i=1}^{d} c_i^\alpha \cdot a \cdot \frac{b}{da} + \left(\sum_{i=1}^{d} c_i \right)^\alpha \cdot a \cdot \frac{\beta^\alpha a_R}{a} = b \sum_{i=1}^{d} c_i^\alpha + a_R \beta^\alpha \left(\sum_{i=1}^{d} c_i \right)^\alpha,$$

as $\delta \to 0$. This result agrees with the calculation in Hyung and de Vries (2002).

4. Diversification effects

As we shown in Proposition 3.1, diversification is beneficial only among the risk factors that share the same maximal marginal indices. Otherwise, the risk factor with the maximal marginal index will lead to the minimal tail risk. When considering a constant level of marginal indices, evaluating the VaR of a portfolio is equivalent to evaluating its scale function. Theorem 3.1 provides an explicit way to evaluate the scale function of an aggregated portfolio. In this section, we study whether diversification is indeed beneficial in such a case.

As in Fama and Miller (1972), in order to study the diversification effects, we assume the symmetric properties as that the individual risk factors has the same characteristics, i.e. $\alpha_i = \alpha$ and $a_i(\delta) = a(\delta)$ for $1 \leq i \leq d$. Thus, we could consider $m_i = 1$ for $1 \leq i \leq d$ in Theorem 3.1. Meanwhile, we consider the average portfolio P constructed by assigning equal weights to the considered risk factors as

$$P = \bar{X} = \sum_{i=1}^{d} \frac{1}{d} X_i.$$

We study whether the VaR of P is decreasing when the number of risk factors d is increasing.

From Theorem 3.1 we get that, as $\delta \to 0$,

$$a_{\bar{X}}(\delta) \sim a(\delta) d^{1-\alpha} \int_W \left(\sum_{i=1}^{d} w_i^{1/\alpha} \right)^{\alpha} H(dw) \tag{5}$$

In case $\alpha = 1$, from Remark 3.1, the scale function of a portfolio does not depend on the dependence structure. Hence, we get that $a_P(\delta) \sim a(\delta)$. Therefore, the scale function of the average portfolio is at the same level as the individual risk factors. This confirms Fama and Miller's conclusion that the diversification has no effect when $\alpha = 1$. Notice that this result always holds regardless the dependence structure. Hence we extend the Fama and Miller's result.

When $\alpha > 1$, we study the lower and upper bounds of the scale function of the average portfolio. From Jensen inequality,

$$\left(\sum_{i=1}^{d} \frac{1}{d} w_i^{1/\alpha} \right)^{\alpha} \leq \sum_{i=1}^{d} \frac{1}{d} w_i = \frac{1}{d},$$

for any $(w_1, \cdots, w_d) \in W$. Thus,

$$d^{1-\alpha} \int_W \left(\sum_{i=1}^{d} w_i^{1/\alpha} \right)^{\alpha} H(dw) \leq d^{1-\alpha} d^{\alpha-1} = 1.$$

We conclude that the scale function of the average portfolio is always bounded by the scale function of the individual risk factors. Hence, in terms of VaR, the average portfolio is less risky. On the other hand, the Jensen inequality turns to be equality if and only if all w_is are equal. Therefore, if and only if in the completely tail dependent case as shown in Example 3.2, the risk of the average portfolio is at the same level as the individual risk factors. In that case, diversification has not effect.

Besides investigating the existence of the diversification effect, we also study what is the potential speed of risk reducing by diversification, i.e. at which speed the risk is diversified away when imposing more risk factors. We call it the *maximum speed of diversification*. Notice that

$$\left(\sum_{i=1}^{d} w_i^{1/\alpha} \right)^{\alpha} \geq \sum_{i=1}^{d} w_i = 1, \tag{6}$$

for all $(w_1, \cdots, w_d)$ on W. From (5) we get that

$$\lim_{\delta \to 0} \frac{a_{\bar{X}}(\delta)}{a(\delta)} \geq d^{1-\alpha}.$$

Hence, in terms of the reduction of the scale function, the maximum speed of diversification is $d^{1-\alpha}$. Together with (2), the maximum speed of diversification at VaR level is $d^{1/\alpha - 1}$. This speed can be achieved if and only if the inequality (6) turns to be equality for all $(w_1, \cdots, w_d)$ on W. That is equivalent to the condition that there is only one dimension in $(w_1, \cdots, w_d)$ equals to 1 while all the rest equals to 0. This is the tail independence case as shown in Example 3.1. Thus the maximum speed of diversification can be achieved in the tail independence case. We summarize the above discussion in the following theorem.

Theorem 4.1. *Suppose d risk factors $X = (X_1, \cdots, X_d)$ follows a d-dimensional EVT setup with equal marginal tail indices, α, equal marginal scale functions $a(\delta)$ and spectral measure H on W. Denote the marginal VaR of each X_i at tail probability level δ as $VaR(\delta)$ and the VaR of portfolio $P = \bar{X}$ as $VaR_P(\delta)$. When $\alpha > 1$ we have that*

$$1 \geq \lim_{\delta \to 0} \frac{VaR_P(\delta)}{VaR(\delta)} \geq d^{1/\alpha - 1}.$$

The first inequality turns to be equality if and only if the risk factors are completely tail dependent. The second inequality turns to be equality if and only if the risk factors are tail independent.

From Theorem 4.1, we confirmed the existence of diversification effect under $\alpha > 1$ only except the completely tail dependent case. Moreover, the result on the maximum speed of diversification shows that assuming tail independence yields the lowest estimate on the aggregated risk. Thus, when the real dependence structure deviates from tail independence, ignoring the dependence structure may cause severe problem in terms of underestimating the aggregated risk. Hence it is necessary to model the tail dependence structure properly.

In case the dependence structure lies in between completely tail dependence and tail independence, the VaR of the average portfolio lies in between those of the two cases. We demonstrate one example: the CAPM model as in Example 3.3.

From the calculated scale function in Example 3.3, we get that, for the average portfolio $P = \bar{X}$,

$$a_P(\delta) \sim bd^{\alpha-1} + a_R\beta^\alpha.$$

Hence we observe that from the right hand side that, the diversification can not diversify away the market risk as shown in the second item, while the individual risks are diversified away at the maximum speed of diversification as shown in the first item. The mixture of the two effects gives a hybrid effect of diversification. Similar result has been obtained in Hyung and de Vries (2002).

The case $\alpha < 1$ is important in catastrophic risk analysis. In this case, similar to the $\alpha > 1$ case, one can obtain the following relation from the Jensen inequality.

$$d^{1-\alpha} \int_W \left(\sum_{i=1}^d w_i^{1/\alpha} \right)^\alpha H(dw) \geq d^{1-\alpha} d^{\alpha-1} = 1.$$

Hence, the diversified portfolio has higher risk than the individual risk factors. This confirms Fama and Miller's conclusion. Again, we generalized their result to the dependent case.

Summarizing the above discussion, we have the following three statements:
1) when $\alpha = 1$, the diversification has no effect regardless the dependence structure;
2) when $\alpha < 1$, the diversification has negative effects, i.e. leading to a higher risk;
3) when $\alpha > 1$, the diversification is in general leading to a lower risk. However, in the completely tail dependent case, the diversification has no effect. The maximum speed of diversification is achieved in the tail independent case. In other intermediate cases, the diversification effect depends on the dependence structure.

5. Risk management implications

Our results on the scale function calculation and diversification effects have direct implication for risk management. In this section we first discuss practical guideline in managing the tail risk of a diversified portfolio. Since this can only be achieved for the case $\alpha > 1$, we discuss an alternative solution for the $\alpha \leq 1$ case.

5.1. Managing tail risks in case $\alpha > 1$

In practice, when observing historical data on a set of heavy-tailed risk factors, we can estimate their marginal tail indices and scale functions from univariate EVT. For example, the Hill estimator provides reasonable estimates. By comparing the estimated marginal tail indices, when they significantly differ from each other, from Proposition 3.1, the risk of a diversified portfolio is dominated by the components involving the most heavy-tailed risk factors. Hence, to minimize the tail risk, it is necessary to choose the risk factors with the maximum estimated tail indices.

For a diversified portfolio constructed from the risk factors sharing the same estimated marginal indices, we apply Theorem 3.1 to calculate its the scale function. In order to do that, it is necessary to estimate the spectral measure H on W. Multivariate EVT provides non-parametric estimates of H, see de Haan and Ferreira (2006). After estimating H, the calculation on the scale of the portfolio is then straightforward.

From the discussion of diversification effect, in the case $\alpha > 1$, it is possible to achieve a smaller tail risk by constructing a diversified portfolio. In fact there exists a unique portfolio with the minimal tail risks. This can be achieved by minimizing the scale function of portfolio $P = \sum_{i=1}^{d} c_i X_i$ with the constrains that $\sum_{i=1}^{d} c_i = 1$ and $0 < c_i < 1$. Notice that

$$d \int_W \left(\sum_{i=1}^{d} c_i (m_i w_i)^{1/\alpha} \right)^{\alpha} H(dw)$$

is a strict convex function on $(c_1, \cdots, c_d) \in W$ for $\alpha > 1$. According to convex optimization theory there exists a unique optimal solution. Empirically, this can be solved by the Newton method, see, e.g. Dennis and Schnabel (1996) and Boyd and Vandenberghe (2004).

5.2. Measuring the linkage to systemic risk in case $\alpha \leq 1$

Diversification has no effect or negative effect in the case $\alpha \leq 1$. Therefore, it is not possible to reduce the tail risk by constructing a diversified portfolio. In this case, in order to minimize the tail risk, the optimal choice is not to diversify but compare the individual risks.

Suppose the individual risks are in the same level, say, all the marginal tail indices are equal, and all the marginal scale functions are at the same level, i.e. $\alpha_i = \alpha \leq 1$ and $a_i(\delta) \sim a(\delta)$, for all $1 \leq i \leq d$. In this case, the marginal VaRs are at the same level. Hence the individual risk factors do not differ in terms of tail risks. However, considering the dependence structure H, the risk factors may still differ in terms of the linkages to systemic risk.

In order to compare how the individual risk factors link to the systemic risk, we define an "extreme" situation for the system. We consider a case that at least one of the risk factors is extreme, i.e. $\bigvee_{j=1}^{d} X_j > t$ where t is a high threshold that goes to infinity. To measure the systemic linkage for a specific risk factor i to such an extreme situation, we consider the conditional probability that the $i-$th factor generates the highest loss, given the fact that the "extreme" situation occurs., i.e.

$$p_i := \lim_{t\to\infty} P(X_i = \bigvee_{j=1}^{d} X_j | \bigvee_{j=1}^{d} X_j > t).$$

p_i is called the *probability of dominance (POD)* of the $i-$th risk factor, where $1 \le i \le d$. For each risk factor, the POD indicates the probability to be the "most extreme" risk when the system is in "extreme". Therefore, the one with higher linkage to systemic risk, will exhibit a higher POD. The following theorem shows how to calculate the POD. For the proof, see Appendix.

Theorem 5.1. *Suppose $X = (X_1, \cdots, X_d)^T$ follows d-dimension EVT setup with equal marginal tail indices α, equal marginal scale functions $a(\delta)$ and spectral measure H on W. For any $1 \le i \le d$, the POD of the $i-$th risk factor is calculated as*

$$p_i = \frac{\int_{w_i=\max_{1\le j\le d} w_j} w_i H(dw)}{\int_W \max_{1\le j\le d} w_j H(dw)}.$$

Remark 5.1. *Similar results as in Theorem 5.1 can be found in de Haan (1984) and Resnick and Roy (1990). However, in these papers, the results are not expressed in terms of the H measure.*

We remark that the POD is only determined by the spectral measure H. This reflects the intuition that the POD is a measure of systemic linkage without taking the marginal risks into consideration. The POD can be statistically estimated thanks to the estimation of the spectral measure H. Therefore, in risk management, the risk factor with the minimum POD will be an optimal choice in the following sense: it has the lowest probability to be most risky one in an "extreme" situation. It has the minimum connection with the systemic risk.

Although we study the POD in order to provide an alternative solution for risk management in the case $\alpha \leq 1$, Theorem 5.1 does not assume the range of α. Hence, it can be applied to other values of α in order to evaluate the systemic linkages of risk factors. A potential application is to assess the systemic linkages in banking system, in order to manage the risk of a systemic banking crises.

6. Conclusion

This paper studies the tail risk of a portfolio constructed from dependent risk factors. We consider a multivariate EVT framework which models individual risk factors by heavy-tailed distributions, while allowing tail dependency among them characterized by a spectral measure H.

By employing VaR as the risk criterion, we first link the VaR calculation to the tail index and the scale function of a heavy tailed distribution. For a portfolio constructed from diversified risk factors, when the tail indices of the risk factors are not equal, both the tail index and the scale function of the portfolio is dominated by the component with the lowest tail index regardless the dependence structure. When the tail indices are at the same level, the portfolio shares the same tail index while we provide an explicit formula to calculate its scale function. The calculation takes the dependence structure into consideration.

Our study leads to a similar discussion on the diversification effect as in Fama and Miller (1972) under more general tail dependence structure. For $\alpha = 1$, there is no diversification effect. For $\alpha < 1$, the diversification causes a negative effect that is the increase of the tail risk. For $\alpha > 1$, our conclusion depends on the tail dependence structure. Usually, the diversification leads to a decrease of the tail risk. However, if the individual risks are completely tail dependent, the diversification has no effect. We also obtain the maximum speed of diversification when the diversification effect exists.

The theoretical result can be applied to risk management issues thanks to the statistical estimates on the marginal tail risk information and the dependence structure. Empirically, one could assess the risk of a specified portfolio. In the case $\alpha > 1$, it is possible to construct a portfolio with the minimum tail risk according to convex minimization. For the case $\alpha \leq 1$, since it is not possible to achieve a lower tail risk by diversification, for risk control, it is necessary to compare the individual risks. When the marginals bear the same individual risk, evaluating the systemic linkage between individual risk and the systemic risk is important. We propose the probability of dominance as a measure on such a linkage. Risk factors with a lower POD are less connected to the systemic risk. This result contributes to the literature in dealing with the extremely heavy-tailed case, $\alpha \leq 1$.

Our study shows that modeling the tail dependence structure of risk factors is important for assessing the tail risk of a diversified portfolio and obtaining diversification effect. Overlooking the potential dependency can lead to severe problems in evaluating the aggregated risk. For instance, assuming independency may lead to an underestimation of the aggregated risks. Moreover, in case diversification is not a valid instrument for risk control, investigating the dependence structure provides an alternative view by comparing the linkages to the systemic risk.

Acknowledgement The author thanks Laurens de Haan and conference participants at 13th International Congress on Insurance: Mathematics and Economics for useful discussion and comments.

Appendix

Proof of Proposition 3.1

Basrak et al. (2002) defines regularly variation for a random vector. It is not difficult to verify that when X belongs to $d-$dimensional EVT setup with positive marginal tail indices α_i, $1 \leq i \leq d$, X is a regularly varying random vector with tail index $\min_{1\leq i\leq d} \alpha_i$. From Theorem 1.1 in Basrak et al. (2002), the tail index of a linear combination of a regularly varying random vector is dominated by the minimum marginal tail index. Hence, $\sum_{i=1}^d c_i X_i$ must have tail index α_1. Now, we only need to prove the scale part. For simplicity we only consider the case where $a_1(\delta) \sim A_1$ as $\delta \to 0$.

From Theorem 1.1 in Basrak et al. (2002), we also get that $\sum_{i=2}^d c_i X_i$ must have the tail index $\min_{2\leq i\leq d} \alpha_i > \alpha_1$. Hence,

$$\lim_{t\to\infty} t^{\alpha_1} P(\sum_{i=2}^d c_i X_i > t) = 0.$$

Because for any given $0 < \varepsilon < 1$,

$$P(\sum_{i=1}^d c_i X_i > t) \leq P(c_1 X_1 > (1-\varepsilon)t) + P(\sum_{i=2}^d c_i X_i > \varepsilon t),$$

we get that

$$\limsup_{t\to\infty} t^{\alpha_1} P(\sum_{i=1}^d c_i X_i > t) \leq \frac{c_1^{\alpha_1} A_1}{(1-\varepsilon)^{\alpha_1}}.$$

By taking $\varepsilon \to 0$, we have that

$$\limsup_{t\to\infty} t^{\alpha_1} P(\sum_{i=1}^d c_i X_i > t) \leq c_1^{\alpha_1} A_1.$$

On the other hand, we have that

$$\liminf_{t\to\infty} t^{\alpha_1} P(\sum_{i=1}^d c_i X_i > t) \geq \liminf_{t\to\infty} t^{\alpha_1} P(c_1 X_1 > t) = c_1^{\alpha_1} A_1.$$

Combining the lower and upper boundary, the proposition is proved. $\square$

Proof of Theorem 3.1

In Barbe et al. (2006), this theorem is proved in a simpler case assuming all the marginal scale functions are equal to the same constant. Allowing different marginal scale functions generalizes their result by a similar proof without imposing extra difficulty. Nevertheless, we still give the proof here, since our proof is based on the concept "exponent measure" which is also used in proving Theorem 5.1. For details on the exponent measure, see e.g. de Haan and Ferreira (2006).

Denote $\mathbb{R}_+ = [0, \infty)$. Suppose ν is a measure defined on all Borel sets $A \subset \mathbb{R}_+^d$ with

$$\inf_{(x_1,\cdots,x_d)\in A} \max(x_1, \cdots, x_d) > 0 \tag{7}$$

such that

1) Homogeneity: for any Borel set A satisfying (7) and $\nu(\partial A) = 0$, and any $a > 0$,

$$\nu(aA) = a^{-1}\nu(A).$$

2) Marginal condition: for any $1 \leq i \leq d$ and any $x > 0$,

$$\nu\left\{(x_1, \cdots, x_d) \in \mathbb{R}_+^d \ : \ x_i > x\right\} = 1/x.$$

By denoting

$$A_{x_1,\cdots,x_d} = \{(s_1, \cdots, s_d) \ : \ \exists 1 \leq i \leq d \ \text{ s.t. } \ s_i > x_i\},$$

we can get a L-function from $L(x_1, \cdots, x_d) = \nu(A_{1/x_1,\cdots,1/x_d})$. Conversely, any L-function has such a representation with a suitable ν. ν is called the *exponent measure*.

These two measures H and ν can be transformed from one to the other in the following way. A point $x = (x_1, \cdots, x_d) \in \mathbb{R}_+^d / \{(0, \cdots, 0)\}$ can be mapped to $(r, w) \in (0, \infty) \times W$ by $r = \sum_{1\leq i\leq d} x_i$ and $w = x/(\sum_{i=1}^d x_i)$. It is a one-to-one mapping. Denote this mapping as π. For any Borel set A satisfying (7), $\pi(A)$ must be a Borel set in $(0, \infty) \times W$. In particular, for any $(x_1, \cdots, x_d)^T \in \mathbb{R}_+^d / \left\{(0, \cdots, 0)^T\right\}$,

$$\begin{aligned}\pi(A_{x_1,\cdots,x_d}) &= \{(r, w) \ : \ \exists 1 \leq i \leq d \ \text{ s.t. } rw_i > x_i\} \\ &= \left\{(r, w) \ : \ r > \frac{x_1}{w_1} \wedge \cdots \wedge \frac{x_d}{w_d}\right\}\end{aligned}$$

Hence, we have

$$\begin{aligned}
\nu(A_{x_1,\cdots,x_d}) &= L\left(\frac{1}{x_1},\cdots,\frac{1}{x_d}\right)\\
&= d\int_W \frac{w_1}{x_1}\vee\cdots\vee\frac{w_d}{x_d}H(dw)\\
&= d\int_W\left(\int_{\frac{x_1}{w_1}\wedge\cdots\wedge\frac{x_d}{w_d}}^{\infty}\frac{1}{r^2}dr\right)H(dw)\\
&= d\int_{(r,w)\in\pi(A_{x_1,\cdots,x_d})}\frac{1}{r^2}drH(dw)
\end{aligned}$$

This relation holds for any set $A_{x_1,\cdots,x_d}$, therefore, it also holds for any other Borel set A satisfying (7), i.e.

$$\nu(A) = d\int_{(r,w)\in\pi(A)}\frac{1}{r^2}drH(dw). \tag{8}$$

Now, we back to the proof Theorem 3.1 using the exponent measure ν.

Suppose X follows the $d-$dimensional EVT setup with all marginal tail indices α. Define $Y=(Y_1,\cdots,Y_d)^T$ by $Y_i=X_i^\alpha/m_i$. Then for any tail probability δ,

$$VaR_{X_i}(\delta) = \left(m_i VaR_{Y_i}(\delta)\right)^{1/\alpha}. \tag{9}$$

Moreover, since

$$P(Y_i>t) = P(X_i>(m_it)^{1/\alpha}) = \frac{1}{t}\frac{l_i((m_it)^{1/\alpha})}{m_i},$$

we get that Y_i has tail index 1 with slowly varying function $l_{Y_i}(t)=\frac{l_i((m_it)^{1/\alpha})}{m_i}$, where l_i is the corresponding slowly varying function for X_i. Combining with (9), we get that

$$\begin{aligned}
a_{Y_i}(\delta) &= l_{Y_i}\left(VaR_{Y_i}(\delta)\right)\\
&= \frac{1}{m_i}l_i\left((m_iVaR_{Y_i}(\delta))^{1/\alpha}\right)\\
&= \frac{1}{m_i}l_i(VaR_{X_i}(\delta))\\
&= \frac{1}{m_i}a_i(\delta)\sim a(\delta).
\end{aligned}$$

Hence, Y_i has tail index 1 and scale function $a(\delta)$. Thus $VaR(\delta)_{Y_i} \sim a(\delta)/\delta$ as $\delta \to 0$ Relation (3) implies that as $\delta \to 0$,

$$\frac{P\left(Y_1 > \frac{a(x_1\delta)}{x_1\delta} \text{ or } \cdots \text{ or } Y_d > \frac{a(x_d\delta)}{x_d\delta}\right)}{\delta} \to L(x_1, x_2, \cdots, x_d).$$

Together with the fact that a is a slowly varying function, we get that,

$$\begin{aligned}\frac{P\left(\frac{\delta}{a(\delta)}Y \in A_{\frac{1}{x_1},\cdots,\frac{1}{x_d}}\right)}{\delta} &= \frac{P\left(\frac{\delta}{a(\delta)}Y_1 > \frac{1}{x_1} \text{ or } \cdots \text{ or } \frac{\delta}{a(\delta)}Y_d > \frac{1}{x_d}\right)}{\delta} \\ &\to L(x_1, x_2, \cdots, x_d) = \nu(A_{\frac{1}{x_1},\cdots,\frac{1}{x_d}}).\end{aligned}$$

Thus, for any Borel set A such that $\nu(A)$ is defined, we have, as $\delta \to 0$,

$$\frac{P\left(\frac{\delta}{a(\delta)}Y \in A\right)}{\delta} \to \nu(A). \tag{10}$$

Particularly, we have

$$\begin{aligned}\frac{1}{\delta}Prob\left(\left(\frac{\delta}{a(\delta)}\right)^{1/\alpha} P > 1\right) &= \frac{1}{\delta}Prob\left(\left(\frac{\delta}{a(\delta)}\right)^{1/\alpha} \sum_{i=1}^{d} c_i X_i > 1\right) \\ &= \frac{1}{\delta}Prob\left(\sum_{i=1}^{d} c_i(m_i\frac{\delta}{a(\delta)}Y_i)^{1/\alpha} > 1\right) \\ &= \frac{1}{\delta}Prob\left(\frac{\delta}{a(\delta)}Y \in A^*\right) \\ &\to \nu(A^*),\end{aligned}$$

where $A^* = \left\{(x_1, \cdots, x_d) | \sum_{i=1}^{d} c_i(m_i x_i)^{1/\alpha} > 1\right\}$ is a Borel set. A calculation later will show that $\nu(A^*)$ is finite. Then we could calculate the VaR of the portfolio P as

$$VaR_P(\delta) = \left(\frac{a(\delta/\nu(A^*))}{\delta/\nu(A^*)}\right)^{1/\alpha} \sim \left(\frac{a(\delta)\nu(A^*)}{\delta}\right)^{1/\alpha}.$$

Hence we conclude that P has a tail index α with a scale function $a_P(\delta) \sim a(\delta)\nu(A^*)$.

As the last step of the proof, we calculate $\nu(A^*)$. This is straightforward from the relation between the exponent measure ν and the spectral measure H on W.

$$\begin{aligned}
\nu(A^*) &= d\int_{(r,w)\in\pi(A^*)} \frac{1}{r^2} dr H(dw) \\
&= d\int_{\sum_{i=1}^d c_i(m_i r w_i)^{1/\alpha}>1} \frac{1}{r^2} dr H(dw) \\
&= d\int_{r>\left(\sum_{i=1}^d c_i(m_i w_i)^{1/\alpha}\right)^{-\alpha}} \frac{1}{r^2} dr H(dw) \\
&= d\int_W \left(\sum_{i=1}^d c_i(m_i w_i)^{1/\alpha}\right)^{\alpha} H(dw).
\end{aligned}$$

The proof of Theorem 3.1 is complete. $\square$.

Proof of Theorem 5.1
To calculate the POD, we start with the transformation $Y_i = X_i^\alpha$ similar to the proof of Theorem 3.1. Hence, we have that

$$\begin{aligned}
p_i &= \lim_{t\to\infty} P\left(X_i = \bigvee_{i=1}^d X_i | \bigvee_{i=1}^d X_i > t\right) \\
&= \lim_{t\to\infty} P\left(Y_i = \bigvee_{i=1}^d Y_i | \bigvee_{i=1}^d Y_i > t\right) \\
&= \lim_{t\to\infty} \frac{P(Y\in B_i \bigcap C_t)}{P(Y\in C_t)},
\end{aligned}$$

where

$$B_i =: \left\{(x_1,\cdots,x_d)\in\mathbb{R}_+^d \ : \ x_i = \bigvee_{s=1}^d x_s\right\}$$

and

$$C_t := \left\{(x_1,\cdots,x_d)\in\mathbb{R}_+^d \ : \ \bigvee_{s=1}^d x_s > t\right\}.$$

Taking $t = a(\delta)/\delta$ and $\delta\to 0$, since $a(\delta)$ is slowly varying, we get $t\to\infty$.

Thus

$$
\begin{aligned}
p_i &= \lim_{t\to\infty} \frac{P(Y \in B_i \bigcap C_t)}{P(Y \in C_t)} \\
&= \lim_{\delta\to 0} \frac{P\left(Y \in B_i \bigcap \frac{a(\delta)}{\delta} C_1\right)}{P\left(Y \in \frac{a(\delta)}{\delta} C_1\right)} \\
&= \lim_{\delta\to 0} \frac{P\left(\frac{\delta}{a(\delta)} Y \in B_i \bigcap C_1\right)}{P\left(\frac{\delta}{a(\delta)} Y \in C_1\right)} \\
&\to \frac{\nu(B_i \bigcap C_1)}{\nu(C_1)},
\end{aligned}
$$

where ν is the corresponding exponent measure. The last step comes from relation (10). To continue with the calculation, we use (8) to transform the ν measure to the H measure.

$$
\begin{aligned}
\nu(B_i \bigcap C_1) &= d \int_{\pi(B_i \bigcap C_1)} \frac{1}{r^2} dr H(dw) \\
&= d \int_{rw_i > 1, w_i = \max_{1\le i\le d} w_i} \frac{1}{r^2} dr H(dw) \\
&= d \int_{w_i = \max_{1\le i\le d} w_i} w_i H(dw).
\end{aligned}
$$

Similarly,

$$
\nu(B_1) = d \int_W \max_{1\le i\le d} w_i H(dw).
$$

Combining these two, the proof is complete. $\square$

References

Arzac, E., Bawa, V., 1977. Portfolio choice and equilibrium in capital markets with safety-first investors. Journal of Financial Economics 4, 277–288.

Barbe, P., Fougères, A., Genest, C., 2006. On the tail behavior of sums of dependent risks. Astin Bulletin 36, 361–373.

Basrak, B., Davis, R. A., Mikosch, T., 2002. A characterization of multivariate regular variation. Annuals of Applied Probability 12, 908–920.

Boyd, S., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press.

de Haan, L., 1984. A spectral representation for max-stable processes. Annals of Probability 12, 1194–1204.

de Haan, L., Ferreira, A., 2006. Extreme Value Theory: An Introduction. Springer.

Dennis, J. E., Schnabel, R. S., 1996. Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Society for Industrial and Applied Mathematics.

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling extremal events: for insurance and finance. Springer.

Embrechts, P., Lambrigger, D., Wüthrich, M., 2008. Multivariate extremes and the aggregation of dependent risks: Examples and counter-examples. Extremes 12, 107–127.

Embrechts, P., Nešlehová, J., Wüthrich, M., 2009. Additivity properties for value-at-risk under archimedean dependence and heavy-tailedness. Insurance: Mathematics and Economics 44, 164–169.

Fama, E., 1965. The behavior of stock market prices. Journal of Business 38, 34–105.

Fama, E., Miller, M., 1972. The Theory of Finance. Holt, Rinehart and Winston, New York.

Feller, W., 1971. An Introduction to Probability Theory and Its Applications. Vol. 2. Wiley & Sons, New York.

Gouriéroux, C., Laurent, J., Scaillet, O., 2000. Sensitivity analysis of values at risk. Journal of Empirical Finance 7, 225–246.

Hartmann, P., Straetmans, S., de Vries, C., 2004. Asset market linkages in crisis periods. Review of Economics and Statistics 86, 313–326.

Hill, B., 1975. A simple general approach to inference about the tail of a distribution. Annals of Statistics 3(5), 1163–1173.

Huang, X., 1992. Statistics of bivariate extreme values. Ph.D. thesis, Tinbergen Institute.

Hyung, N., de Vries, C., 2002. Portfolio diversification effects and regular variation in financial data. Allgemeines Statistiches Archiv 86, 69–82.

Hyung, N., de Vries, C., 2005. Portfolio diversification effects of downside risk. Journal of Financial Econometrics 3, 107–125.

Hyung, N., de Vries, C., 2007. Portfolio selection with heavy tails. Journal of Empirical Finance 14, 383–400.

Ibragimov, R., Jaffee, D., Walden, J., 2008. Nondiversification traps in catastrophe insurance markets. Review of Financial Studies 22, 959.

Ibragimov, R., Walden, J., 2008. Portfolio diversification under local and moderate deviations from power laws. Insurance: Mathematics and Economics 42, 594–599.

Jansen, D., 2001. Limited downside risk in portfolio selection among u.s. and pacific basin equities. International Economic Journal 15, 1–22.

Jansen, D., de Vries, C., 1991. On the frequency of large stock returns: Putting booms and busts into perspective. Review of Economics and Statistics 73, 18–24.

Jansen, D., Koedijk, K., de Vries, C., 2000. Portfolio selection with limited downside risk. Journal of Empirical Finance 7, 247–269.

Markowitz, H., 1952. Portfolio selection. Journal of Finance 7, 77–91.

Poon, S., Rockinger, M., Tawn, J., 2004. Extreme value dependence in financial markets: Diagnostics, models, and financial implications. The Review of Financial Studies 17, No.2, 581–610.

Resnick, S., Roy, R., 1990. Multivariate extremal processes, leader processes and dynamic choice models. Advances in Applied Probability 22, 309–331.

Roy, A., 1952. Safety first and the holding of assets. Econometrica 20, 431–449.

Susmel, R., 2001. Extreme observations and diversification in latin american emerging equity markets. Journal of International Money and Finance 20, 971–986.

Weissman, I., 1978. Estimators of parameters and large quantiles based on the k largest observations. Journal of the American Statistical Association 73, 812–815.