Journal of Banking & Finance 69 (2016) 20–36

Contents lists available at ScienceDirect

Journal of Banking & Finance

journal homepage: www.elsevier.com/locate/jbf

Structure and estimation of Lévy subordinated hierarchical Archimedean copulas (LSHAC): Theory and empirical tests ☆

Wenjun Zhu [a], Chou-Wen Wang [b,c], Ken Seng Tan [d,*]

[a] School of Finance, Nankai University, 38 Tongyan Road, Jinnan District, Tianjin 300350, China
[b] Department of Finance, National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan
[c] Risk and Insurance Research Center, College of Commerce, National Chengchi University, Taiwan
[d] Department of Statistics and Actuarial Science, University of Waterloo, 200 University Avenue West, Waterloo N2L 3G1, ON, Canada

ARTICLE INFO

Article history:
Received 3 May 2015
Accepted 29 January 2016
Available online 22 March 2016

JEL classification:
C13
C16
G15
G17

Keywords:
High dimensional modeling
Hierarchical Archimedean copulas
Lévy subordinators
Downside risk

ABSTRACT

Lévy subordinated hierarchical Archimedean copulas (LSHAC) are flexible models in high dimensional modeling. However, there is limited literature discussing their applications, largely due to the challenges in estimating their structures and their parameters. In this paper, we propose a three-stage estimation procedure to determine the hierarchical structure and the parameters of a LSHAC. This is the first paper to empirically examine the modeling performances of LSHAC models using exchange traded funds. Simulation study demonstrates the reliability and robustness of the proposed estimation method in determining the optimal structure. Empirical analysis further shows that, compared to elliptical copulas, LSHACs have better fitting abilities as well as more accurate out-of-sample Value-at-Risk estimates with less parameters. In addition, from a financial risk management point of view, the LSHACs have the advantage of being very flexible in modeling the asymmetric tail dependence, providing more conservative estimations of the probabilities of extreme downward co-movements in the financial market.

1. Introduction

The construction and estimation of high dimensional copulas are challenging problems, yet they are critical and essential for financial risk management (Kole et al., 2007; Patton, 2009). Therefore, investigating the theoretical properties and empirical applications of high dimensional modeling with copulas have attracted much attentions in the literature (Bu et al., 2011; Dias and Embrechts, 2004; Embrechts et al., 2003; Hoesli and Peka, 2013; Kumar and Okimoto, 2011; Patton, 2006; Zhou and Gao, 2012). Elliptical copula models are inadequate to capture the nonlinear dependence in the financial returns (Embrechts et al., 2002), and additionally the number of parameters in the elliptical copulas grows quadratically with the dimension. Vine copulas, also known as pair copula constructions, facilitate extensions from bivariate copulas to higher dimensions.[1] Although extremely flexible, some outstanding issues need to be adequately addressed for vine copula models, including testing the "simplifying assumptions", selecting appropriate bivariate models from the huge number of potential candidates, designing spatial vines, and goodness-of-fit tests for high dimensional vine copulas, etc.

Archimedean copulas (AC), though have a very small number of parameters irrespective of dimensions, suffer from the exchangeable structures, which makes AC inadequate to model complex dependence structures (Weiß and Scheffer, 2015). In an attempt to overcome the exchangeability issue of ACs, the hierarchical Archimedean copula (HAC) has been proposed (Joe, 1997). This approach partially overcomes the exchangeability by "nesting" two or more ACs with appropriate grouping, thereby providing a more flexible framework. Despite their advantages, there are *compatible conditions* which the generators need to be satisfied to ensure that the resulting HAC yields a valid multivariate distribution. These conditions, however, can be difficult to verify and hence also restrict the practical applications of HACs.

☆ The authors acknowledge the research funding from the Natural Sciences and Engineering Research Council of Canada, Society of Actuaries CAE Research Grant, and Global Risk Institute (GRI) Research Grant. The second author was also supported in part by the MOST 101-2410-H-327-029.

* Corresponding author. Tel.: +1 (519) 888 4567x36688; fax: +1 (519) 746 1875.
E-mail addresses: zhu-wenjun@outlook.com (W. Zhu), chouwenwang@gmail.com (C.-W. Wang), kstan@uwaterloo.ca (K.S. Tan).

[1] A detailed introduction to vine copulas can be found in Aas et al. (2009), and the estimation of vine copulas is introduced in Kurowicka and Cooke (2006) for Gaussian vines and Aas et al. (2009) for non-Gaussian vines.

http://dx.doi.org/10.1016/j.jbankfin.2016.01.011
0378-4266/

In this paper, we advocate Lévy subordinated hierarchical Archimedean copulas (LSHAC), which is a family of HAC models constructed from Lévy subordinators. Hering et al. (2010) introduce the construction and simulation of LSHACs, while Mai and Scherer (2012) discuss LSHAC within a h-extendible copula framework. By inducing dependence within each group with Lévy subordinators, the hard-to-check *compatible conditions* are conveniently solved, leading to a huge potential application to financial modeling.

While LSHACs have substantially enlarged the family of HACs, they have never been employed in modeling the high-dimensional financial data, mostly because of the difficulties in determining the suitable hierarchical structures as well as estimating the parameters for LSHACs. The recursive multi-stage maximum likelihood (ML) estimation procedure proposed by Okhrin et al. (2013b) is efficient for HACs with the same generator functions such as Gumbel generator or Clayton generator (hereafter, we call these models as All-GM-HACs or All-CL-HACs), but will be computationally demanding for general HAC models with different generators. Moreover, their technique provides sub-optimal structure as well as ML estimators because of its recursive nature. In addition, LSHAC models are constructed in such a way that the parameters in the outer layers in the hierarchy should not be estimated later than the inner layer parameters, meaning the bottom-up recursive procedure by Okhrin et al. (2013b) is not applicable for LSHACs.

Motivated by these observations, this paper attempts to fill up these gaps by providing a comprehensive study of some of the outstanding issues in the construction and estimation of LSHACs. In doing so, we explicitly construct a multi-level LSHAC in a fully general setting by developing a notation system, an integral representation and the corresponding sampling algorithm. In addition, we propose to exploit the hierarchical clustering analysis to efficiently determine the structure of LSHACs with a new proposed dissimilarity metric, called τ-Euclidean metric. Using a simulation study, we demonstrate that the best grouping results provided by the τ-Euclidean metric can correctly identify the true structure more than 95% of the times. Once the optimal structure of a LSHAC is identified, an augmented inference for margin (AIFM) method is proposed to estimate the remaining LSHAC parameters.

We illustrate the economic significance of the LSHAC models by fitting the high dimensional financial data and forecasting the out-of-sample Value-at-Risk of an equally-weighted portfolio. To the best of our knowledge, this is the first paper to apply the LSHACs in financial modeling and risk management. Our results indicate that the LSHACs provide more accurate Value-at-Risk forecasts compared to the benchmark model of Student's t copula. In addition, we model the joint downward movements of the assets in the market, as appropriately modeling the downside risk is important for financial risk managers (Ang and Bekaert, 2002; Ang and Chen, 2002; Boubaker and Sghaier, 2013; Brooks et al., 2005; Harvey and Siddique, 1999; Jondeau and Rockinger, 2003; Das and Uppal, 2004; Longin and Solnik, 2001). With the flexibility of modeling asymmetric tail dependences, the LSHAC models will help financial risk managers to more accurately estimate the extreme downside co-movements of the market, which is critically important for risk management and portfolio optimization.

The remainder of this paper proceeds as follows. Section 2 introduces the LSHAC model and three stage estimation procedure. In particular, a new dissimilarity metric based on the hierarchical clustering analysis is proposed to determine the structure of LSHACs. Section 3 investigates the efficiency of the proposed method with a simulation study. In Section 4, we empirically test the LSHACs by forecasting the Value-at-Risk and modeling the downside co-movements. Section 5 concludes the paper.

2. Methodology

2.1. Hierarchical Archimedean copulas (HACs)

A function $C:[0,1]^d \to [0,1]$, $C(u_1,u_2,\ldots,u_d)=\psi(\psi^{-1}(u_1)+,\ldots,\psi^{-1}(u_d))$ defines a d-dimensional Archimedean copula (AC) if $\psi \in \mathcal{G}=\{\psi:[0,\infty)\to[0,1]|\psi_{\lim u\to\infty}(u)=0,\psi(0)=1,(-1)^k\frac{d^k}{du^k}\psi(u)\geqslant 0,k\in\mathbb{N}\}$ (Kimberling, 1974; Nelsen, 2006). Functions in the class of $\mathcal{G}$ is known as *completely monotonic* (c.m.). ψ is called the generator of the corresponding AC and ψ^{-1} is its general inverse, defined by $\psi^{-1}(u)=\inf\{t:\psi(t)\leqslant u\}$. The class of c.m. functions also coincides with the class of Laplace transforms on $[0,\infty)$ (Feller, 2008). Hence, copulas defined by the c.m. generators are also known as the Laplace transform AC (LT-AC).

The advantage of the AC family is that it simplifies the dependence modeling in high dimension with only one parameter. The drawback of such simplification is that the resulting distribution leads to the exchangeability phenomenon; i.e. the distribution of random variables $(u_1,u_2,\ldots,u_d)$ is invariant under permutation. To address this problem, the hierarchical Archimedean copula (HAC) model has been proposed by nesting the random variables into a hierarchy. HAC was first introduced by Joe (1997), and discussed within a more general framework by Savu and Trede (2010). Sampling algorithms are discussed by Whelan (2004), McNeil (2008), and Hofert (2012). The HAC model is best illustrated with an example. Assuming that a six-dimensional HAC is given by

$$C(u_1,\ldots,u_6)=C_{\psi_0}(C_{\psi_{1,1}}(C_{\psi_{2,1}}(u_1,u_2),u_3),C_{\psi_{1,2}}(C_{\psi_{2,2}}(u_4,u_5),u_6)). \tag{1}$$

Note that (1) is a three-level HAC with five generators. The copula C_{ψ_0} with generator ψ_0 is known as the outer copula while copulas $C_{\psi_{1,1}}$ and $C_{\psi_{1,2}}$ ($C_{\psi_{2,1}}$ and $C_{\psi_{2,2}}$), with generators $\psi_{1,1}$ and $\psi_{1,2}$ ($\psi_{2,1}$ and $\psi_{2,2}$), are the inner copulas at level 1 (level 2), respectively. Besides that ψ_0 and $\psi_{i,j}\,(i,j=1,2)$ should be c.m., to ensure (1) is a valid copula, the conditions ψ_0 and $\psi_{i,j}\in\mathcal{G}$ and $(\psi_0^{-1}\circ\psi_{i,j})'$ and $(\psi_{1,k}^{-1}\circ\psi_{2,k})'\in\mathcal{G}(i,j,k=1,2)$, called *compatible conditions*, need to be satisfied. Note that the notation "$\circ$" denotes function composition. The *compatible conditions* cause the construction of the HACs more challenging. If all the generators of a HAC are from the same AC family, these conditions are not too difficult to verify, since in most cases the copula parameters should be monotonic from top to deeper levels (Okhrin et al., 2013b; Embrechts et al., 2003). However, if a HAC is constructed from mixed generators involving different families, one has to verify the *compatible conditions* on a case-by-case basis (Savu and Trede, 2010; Hofert, 2012). For this reason most empirical studies on the HAC models have focused on either All-GM-HACs or All-CL-HACs (Savu and Trede, 2010; Okhrin et al., 2013b; Okhrin et al., 2013a; Choroś-Tomczyk et al., 2013). Hering et al. (2010) circumvent this hard-to-check *compatible conditions* by constructing two-level HACs via Lévy subordinators. Mai and Scherer (2012) consider a h-extendible framework in which LSHAC is one of the special cases. They provide a stochastic representation of three-level LSHACs and explained that the stochastic representation can be extended to higher levels in an iterative way.

2.2. General framework of LSHAC

In this subsection, we extend the model in Hering et al. (2010) by providing an integral representation of a general L-level LSHAC exhibited in Fig. 1. The notation system and its corresponding sampling algorithm are described in Appendix A. Specifically, let $\{S_t:0\leqslant t\leqslant T\}$ be a Lévy subordinator, i.e., a stochastically

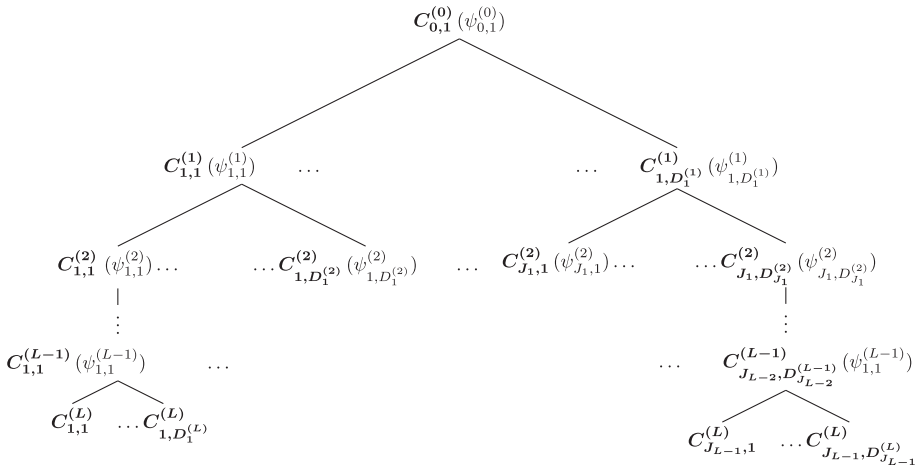


Fig. 1. General framework of a LSHAC model. The notation system of this LSHAC is introduced in Appendix A. The functions in the parentheses are the AC generators for the corresponding inner copulas.

continuous non-decreasing Lévy process, which has zero start, stationary and independent increments (see Tankov (2004); Proposition 3.10). The Laplace transform of S_t satisfies the following equation:

$$E(e^{-\omega S_t}) = \exp(-t\Psi(\omega)), \quad \forall \omega > 0, \tag{2}$$

where the non-decreasing function $\Psi : [0,\infty) \to [0,\infty)$ is called the *Laplace exponent* of S_t. Theorem 2.1 provides the integral representation of the L-level LSHAC, with the notations in Appendix A, in terms of the Laplace transform.

Theorem 2.1. *Given the structure of a LSHAC in Fig. 1, the copula function, $C(u_1,\ldots,u_d)$, can be constructed as*

$$\int_0^\infty \prod_{j_1=1}^{D_{j_0}^{(1)}} \int_0^\infty \prod_{j_2=1}^{D_{s_1}^{(2)}} \int_0^\infty \cdots \int_0^\infty \prod_{j_{L-1}=1}^{D_{s_{L-2}}^{(L-1)}} \int_0^\infty \prod_{j_L=1}^{D_{s_{L-1}}^{(L)}} \times \left(F_{s_{L-2}j_{L-1}}^{(L-1)}(C_{s_{L-1}j_L}^{(L)})\right)^{v_{s_{L-2}j_{L-1}}^{(L-1)}} (dG)_{j_{L-1}}^{(L-1)}, \tag{3}$$

where

$$(dG)_{j_0}^{(0)} = dG_{0,1}^{(0)}(v_{0,1}^{(0)}),$$

and

$$(dG)_{j_l}^{(l)} = d\widetilde{G}_{s_{l-1}j_l}^{(l)}(v_{s_{l-1}j_l}^{(l)}; v_{s_{l-2}j_{l-1}}^{(l-1)}) \ldots d\widetilde{G}_{s_0 j_1}^{(1)}(v_{s_0 j_1}^{(1)}; v_{0,1}^{(0)}) dG_{0,1}^{(0)}(v_{0,1}^{(0)}).$$

Proof. Proof of Theorem 2.1 is provided in Appendix B. □

It follows from the above theorem that the following corollary provides an expression of the inner generators.

Corollary 2.1. *At level l, where $1 \leqslant l \leqslant L$, the j_l-th copula generator in position s_{l-1} : $\psi_{s_{l-1}j_l}^{(l)}$, can be expressed as*

$$\psi_{s_{l-1}j_l}^{(l)}(u) = \psi_{0,1}^{(0)} \overset{l-1}{\underset{i=1}{\odot}} \widetilde{\Psi}_{s_{i-1}j_i}^{(i)}(u), \tag{4}$$

where $\odot_{k=1}^n f_k := f_1 \circ \ldots \circ f_n$, and $\psi_{s_{l-1}j_l}^{(l)}$ is c.m.

Corollary 2.1 states that at each level of a LSHAC, the generator can be constructed from composing an outer AC generator and a sequence of Laplace exponents of Lévy subordinators. Tables 1 and 2 list, respectively, examples of c.m. Archimedean generators, together with their tail dependences, and examples of Lévy subordinators, which are used in the simulation study in Section 3 and the empirical analysis in Section 4. In terms of the tail dependences, in contrast to the Gaussian copula with no tail dependence and Student's t copula with symmetric tail dependence, the LSHAC models can model upper tail dependence (GM family), lower tail dependence (CL family), tail independence (IG family), or asymmetric dependence at both tails (C_{12} and C_{14} family). For financial market, where the extreme events usually happen asymmetrically, LSHAC

Table 2
Lévy Subordinators. G: Gamma process; GM: Stable process; IG: the Inverse Gaussian process.

Subordinator	$\Psi(u)$	Parameters
G	$\Psi_G = a\log(1+\frac{u}{b})$	$a>0,\ b>0$
GM	$\Psi_{GM} = u^a$	$0<a<1$
IG	$\Psi_{IG} = a\sqrt{2u+b^2} - ab$	$a>0,\ b>0$

Table 1
Archimedean Copula (AC) generators. CL: Clayton family; GM: Gumbel family; IG: Inverse Gaussian family; C12: (12)th copula in Nelsen (2006, p. 116); C14: (14)th copula in Nelsen (2006, p. 116).

Family	$\psi(u)$	λ_l	λ_u	Parameter
GM	$\psi_{GM}(u) = \exp\left(-x^{\frac{1}{\theta}}\right)$	0	$2-2^{\frac{1}{\theta}}$	$\theta \geqslant 1$
CL	$\psi_{CL}(u) = (1+u)^{-\frac{1}{\theta}}$	$2^{-\frac{1}{\theta}}$	0	$\theta > 0$
IG	$\psi_{IG}(u) = \exp\left(\frac{1}{\theta}(1-\sqrt{1+2\theta^2 x})\right)$	0	0	$\theta > 0$
C_{12}	$\left(1+x^{\frac{1}{\theta}}\right)^{-1}$	$2^{-\frac{1}{\theta}}$	$2-2^{\frac{1}{\theta}}$	$\theta \geqslant 1$
C_{14}	$\psi_{C_{14}}(u) = \left(1+x^{\frac{1}{\theta}}\right)^{-\theta}$	$\frac{1}{2}$	$2-2^{\frac{1}{\theta}}$	$\theta \geqslant 1$

may have potential advantage of capturing the extreme co-movements more accurately.

It follows immediately from Corollary 2.1 that the All-GM-HAC model, which is currently the most commonly used HAC in the empirical analysis, is a special case of LSHACs. This property is expressed in Corollary 2.2 below, and it is also mentioned with a three level HAC example in Mai and Scherer (2012).

Corollary 2.2. *For an All-GM-HAC, the lth level copula generator* $\psi^{(l)}(u)$ *can be expressed as* ($l \geqslant 1$):

$$\psi^{(l)}(u) = \psi^{(0)} \overset{l-1}{\underset{k=1}{\odot}} \widetilde{\Psi}^{(k)}(u) = \exp\left(-u^{\prod_{k=1}^{l-1}\frac{1}{\theta_k}}\right).$$

From the parameterization in Tables 1 and 2, $\psi^{(0)}$ *represents a GM generator with* $\theta = \theta_0,\ \theta_0 \geqslant 1$; $\widetilde{\Psi}^{(k)}$ *denotes the kth GM subordinator with* $a = 1/\theta_k,\ \theta_k \geqslant 1$.

2.3. Structure and estimation of a LSHAC

In this subsection, we discuss an estimation procedure for LSHACs, focusing on the determination of the hierarchical structure. Given a d-dimensional sample data with T observations, $\boldsymbol{X}_T = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d)_{T\times d}$, the log-likelihood function of the sample is defined by

$$L(\boldsymbol{\theta}) = \sum_{j=1}^{d}\sum_{t=1}^{T}\log f_j\left(x_{t,j}|\mathcal{F}_{t-1};\boldsymbol{\theta}^M\right) + \sum_{t=1}^{T}\log\left(c\left(F_1(x_{t,1}),\ldots,F_d(x_{t,d})|\mathcal{F}_{t-1};\boldsymbol{\theta}\right)\right), \tag{5}$$

where $\boldsymbol{\theta} = (\boldsymbol{\theta}^M, \boldsymbol{\theta}^C, \mathcal{S})$ is the parameter vector to be estimated, including the marginal parameter set, $\boldsymbol{\theta}^M$, the copula parameter set, $\boldsymbol{\theta}^C$, and the hierarchy structure $\mathcal{S}$; $\mathcal{F}_t$ is the information available up to time t; c is the corresponding copula density; F_j is the marginal cumulative distribution function (c.d.f.) of $\boldsymbol{x}_j$ with density f_j, where $j = 1, \ldots, d$.

The classical inference for margin (IFM) estimation for copulas, in which $\boldsymbol{\theta}^M$ and $\boldsymbol{\theta}^C$ is calibrated in a two-step estimation procedure, is widely used and yields asymptotically efficient estimates (Joe, 1997; Patton, 2006). However, the ML estimation can only be employed to an HAC with a known hierarchical structure. Consequently, we propose an augmented IFM (AIFM) method with a three-stage procedure, determining the hierarchical structure of a LSHAC by using hierarchical clustering method.

In the first stage we obtain the ML estimator of each margin's parameter set, $\boldsymbol{\theta}_j^M,\ j = 1, \ldots, d$, from

$$\widehat{\boldsymbol{\theta}_{\boldsymbol{j}}^{\boldsymbol{M}}} = \underset{\boldsymbol{\theta}_j^M}{\operatorname{argmax}}\sum_{t=1}^{T}\log f_j\left(x_{t,j}|\mathcal{F}_{t-1};\boldsymbol{\theta}_j^M\right) \tag{6}$$

and produce the *pseudo-sample* $\boldsymbol{u} = (\boldsymbol{u_1}, \ldots, \boldsymbol{u_d})'$ by probability transformation with the estimated marginal distribution functions, namely

$$\boldsymbol{u} = (\boldsymbol{u_1}, \ldots, \boldsymbol{u_d})' = \left(\widehat{F}_1(\boldsymbol{x_1};\widehat{\boldsymbol{\theta}_{\boldsymbol{1}}^{\boldsymbol{M}}}), \ldots, \widehat{F}_d(\boldsymbol{x_d};\widehat{\boldsymbol{\theta}_{\boldsymbol{d}}^{\boldsymbol{M}}})\right)', \tag{7}$$

where $\widehat{F}_1(\boldsymbol{x_1};\widehat{\boldsymbol{\theta}_{\boldsymbol{1}}^{\boldsymbol{M}}}), \ldots, \widehat{F}_d(\boldsymbol{x_d};\widehat{\boldsymbol{\theta}_{\boldsymbol{d}}^{\boldsymbol{M}}})$ represent the estimates of the marginal probability transformations.

Let $\mathcal{S}$ be the true hierarchical structure that underlies the LSHAC. Given that $\mathcal{S}$ is unknown in practice, the objective of the second stage is to determine $\widehat{\mathcal{S}}$ that closely resemblances $\mathcal{S}$. As noted earlier, determining the optimal structure of a LSHAC is one of the key issues that has largely been ignored in the literature of LSHACs, despite its critical role on dependence modeling. Here we propose to determine the optimal structure of a LSHAC by resorting to the hierarchical clustering analysis (Ward, 1963; Székely and Rizzo, 2005; Zhang et al., 2013).

The hierarchical clustering procedure entails choosing an appropriate metric of dissimilarity (or equivalently similarity) between each pair of the pseudo sample, $(\boldsymbol{u_1}, \ldots, \boldsymbol{u_d})'$, where $\boldsymbol{u_j} = (u_{1j}, \ldots, u_{Tj})'$ and $j = 1, \ldots, d$, to construct a symmetric proximity matrix $\zeta = [d_{i,j}]$, where $d_{i,j}$ denotes a proximity index between the ith and the jth variables. Larger $d_{i,j}$ represents a higher level of dissimilarity. In hierarchical clustering, the Euclidean metric, defined by $d_{ij}^E = \sqrt{\sum_{t=1}^{T}(u_{t,i} - u_{tj})^2}$, is one of the most commonly used dissimilarity metrics. Another plausible family of dissimilarity metric is based on the association coefficients between variables. Examples of these association measures are Pearson's correlation coefficient, Spearman's rho and Kendall's τ.[2] The dissimilarity metric based exclusively on either of the above Euclidean distance or the association coefficient is unsatisfactory in that it takes into account of only one possible source of dissimilarity. More specifically, Euclidean metric fails to consider the association effect between variables while the association metric fails to consider the closeness of the data. To address this problem, we propose the following dissimilarity metric

$$d_{i,j}^{\tau-E} = \frac{\left(d_{i,j}^E\right)^{\alpha}}{\left(1+\tau_{i,j}\right)^{\beta}}, \tag{8}$$

where $\tau_{i,j}$ denotes the Kendall's τ between $\boldsymbol{u_i}$ and $\boldsymbol{u_j}$, α and β are appropriately chosen parameters. The above metric is new and we refer it as *τ-adjusted-Euclidean* (hereafter, τ-Euclidean) metric. Note that the Kendall's τ appears explicitly in Eq. (8). Other association measures such as that based on the Pearson's correlation coefficient or the Spearman's rho could have been used. Here we have chosen Kendall's τ as this is a popular metric for determining the structure of HAC models (see Okhrin et al., 2013b).

We emphasize that the metric based on τ-Euclidean offers a great deal of flexibility for measuring dissimilarity. In particular, by setting $(\alpha = 1, \beta = 0)$ and $(\alpha = 0, \beta = 1)$, the above τ-Euclidean metric reduces to the Euclidean metric and the Kendall's τ metric, respectively. More importantly, the relative magnitude of the parameters α and β provides an effective way of controlling the relative emphasis of the metric dictated by the Euclidean metric and the Kendall's τ metric. For example, a large α and a small β lead to a greater importance on the Euclidean distance and less emphasis on the association measure. Therefore, the proposed new metric has the capability of taking into consideration of both Euclidean distance and association effect jointly. This implies that a lower $d_{i,j}^E$ and a higher dependence (i.e., a larger $\tau_{i,j}$) lead to a smaller $d_{i,j}^{\tau-E}$, indicating lower dissimilarity.

Given the pseudo-sample, $\boldsymbol{u} = (\boldsymbol{u_1}, \ldots, \boldsymbol{u_d})'$, obtained from Eq. (7), and the estimated hierarchical structure, $\widehat{\mathcal{S}}$, from the second stage, the final stage is to determine the ML estimator of the copula parameter set $\boldsymbol{\theta}^C$ according to

$$\widehat{\boldsymbol{\theta}^C} = \underset{\boldsymbol{\theta}^C}{\operatorname{argmax}}\sum_{t=1}^{T}\log\left(c\left(F_1(x_{t,1}),\ldots,F_d(x_{t,d})|\mathcal{F}_{t-1};\widehat{\boldsymbol{\theta}^M},\boldsymbol{\theta}^C,\widehat{\mathcal{S}}\right)\right). \tag{9}$$

[2] We may also consider other distance measures, such as Tanimoto measure (see Goshtasby, 2012). According to the literature, Tanimoto measure usually produces similar results as the association coefficients, and hence we do not consider this measure in this paper.

The resulting AIFM estimator is denoted by $\widehat{\theta} = (\widehat{\theta^M}, \widehat{\theta^C}, \widehat{\mathcal{S}})$, with an optimal hierarchical structure, $\widehat{\mathcal{S}}$, obtained from hierarchical clustering analysis.

3. Simulation analysis

Resorting to a simulation study, this section provides an in-depth analysis to confirm the superiority of the τ-Euclidean metric in correctly identifying the structure of LSHACs. In our benchmark example, we assume a LSHAC model with the following given structure $\mathcal{S}$,

$$C(u_1,\ldots,u_6) = C_{0,1}^{(0)}(C_{1,1}^{(1)}(C_{2,1}^{(2)}(C_{1,1}^{(3)}(u_1,u_2),u_3),u_4),C_{1,2}^{(1)}(u_5,u_6)) \quad (10)$$

and generators:

$$\psi_{0,1}^{(0)}(u) = \psi_{GM}(u) = \exp(-u^{1/\theta}), \quad (11)$$

$$\psi_{1,1}^{(1)}(u) = \psi_{GM\circ G}(u) = \exp\left(-\left(a_{1,1}\log\left(1+\frac{u}{b_{1,1}}\right)\right)^{\frac{1}{\theta}}\right), \quad (12)$$

$$\psi_{1,2}^{(1)}(u) = \psi_{GM\circ IG}(u) = \exp\left(-\left(a_{1,2}\sqrt{2u+b_{1,2}^2}-a_{1,2}b_{1,2}\right)^{\frac{1}{\theta}}\right), \quad (13)$$

$$\psi_{2,1}^{(2)}(u) = \psi_{GM\circ G\circ IG}(u) = \exp\left(-\left(a_{1,1}\log\left(1+\frac{a_{2,1}}{b_{1,1}}\left(\sqrt{2u+b_{2,1}^2}-b_{2,1}\right)\right)\right)^{\frac{1}{\theta}}\right), \quad (14)$$

$$\psi_{1,1}^{(3)}(u) = \psi_{GM\circ G\circ IG\circ GM}(u) = \exp\left(-\left(a_{1,1}\log\left(1+\frac{a_{2,1}}{b_{1,1}}\left(\sqrt{2\exp(-u^{1/\theta_{3,1}})+b_{2,1}^2}-b_{2,1}\right)\right)\right)^{\frac{1}{\theta}}\right). \quad (15)$$

Here the subscripts denote the outer generator and the Lévy subordinators used to construct the corresponding inner generators. For example, $\psi_{GM\circ G}(u)$ is an inner generator constructed by a GM outer generator and a Lévy subordinator G. The parameter set of this LSHAC model is $\theta^C = (\theta, a_{11}, b_{11}, a_{12}, b_{12}, a_{21}, b_{21}, \theta_{31}) = (1.3, 1.3, 10, 0.3, 9, 0.08, 9, 0.5)$. The experiment procedure is given as follows:

Step 1: Sample N sets of copula parameters $\widetilde{\theta_n^C}$, $n = 1,\ldots,N$, from uniform distributions with range $[\theta^C(1-\pi), \theta^C(1+\pi)]$.

Step 2: For each nth set of copula parameters, where $n = 1,\ldots,N$, sample M independent batches each of sample size T from a LSHAC with parameters $\widetilde{\theta_n^C}$ and structure $\mathcal{S}$, using Algorithm Appendix A.1.

Step 3: For each mth simulated batch of sample size T, where $m = 1,\ldots,M$, estimate $\widehat{\mathcal{S}_{n,m}}$ using the hierarchical clustering analysis with the dissimilarity metrics discussed in Section 2.3.

Step 4: Calculate the reliability ratio, ρ_n, which measures the relative proportion of the estimated structures $\widehat{\mathcal{S}_{n,m}}$, $m = 1,\ldots,M$, that correctly identify the true structure $\mathcal{S}$; i.e.,

$$\rho_n = \sum_{m=1}^{M}\frac{\mathbb{I}_{\widehat{\mathcal{S}_{n,m}}=\mathcal{S}}}{M}, \quad n = 1,\ldots,N,$$

where $\mathbb{I}_{\widehat{\mathcal{S}_{n,m}}=\mathcal{S}}$ is an indicator variable with value equals to one if the estimated structure coincides with the true structure, zero otherwise.

Recall that one of the advantages of our proposed τ-Euclidean matric lies in its flexibility. To provide additional analysis on the impact of metrics on the reliability ratio, we consider different combinations of α and β, where $\alpha,\ \beta \in \{0,1,2,3\}$. While there are 16 possible pairs of $(\alpha,\ \beta)$, effectively we only need to consider 9 of such pairs for determining the hierarchical structures $\widehat{\mathcal{S}_{n,m}}$, where $n = 1,\ldots,N$ and $m = 1,\ldots,M$.[3] We let $N = 1000$, $M = 100$, and $T = 1000$. In addition, we use $\pi = 5\%,\ 10\%$, and 20% to reflect the parameter uncertainty in the LSHACs. The mean and variance of the reliability ratio (ρ_n) are summarized in Table 3.

- From Table 3, it is clear that the more general τ-Euclidean metric is more superior than that based on either the Euclidean distance or the Kendall's τ. This also justifies the importance of integrating both metrics. The reliability ratio could reach over 95.6% for $\alpha = 2$ and $\beta = 3$ and with $\pi = 5\%$ parameter uncertainly. Not only that this metric yields the highest reliability ratio, its variability (as measured by its sample variance) is also the smallest.
- As the degree of parameter uncertainty increases (i.e. by increasing π from 5% to 20%), the reliability ratio deteriorates slightly with increasing variability. This phenomenon is consistent for all proximity metrics. It is, however, worth pointing out that while the performance declines with increasing parameter uncertainty, the changes are quite small and hence this provides some indication of the robustness of the underlying proximity metric at identifying the true structure of the underlying LSHAC.
- While the Euclidean metric is the worst among the three metrics, it is comforting to know that it still has a success rate of at least 64%.

Fig. 2 plots the empirical cumulative distribution functions (eCDF) of the reliability ratios for three proximity metrics for $\pi = 5\%,\ 10\%$, and 20%, including the Euclidean metric $(\alpha = 1,\ \beta = 0)$, Kendall's τ metric $(\alpha = 0,\ \beta = 1)$, and the best performance τ-Euclidean metric with $(\alpha = 2,\ \beta = 3)$. It is also of interest to note that the eCDF of the reliability ratios based on τ-Euclidean metric lies under those of the other two metrics. According to the definition of stochastic ordering (see, for example, Hadar and Russell (1969)), the results of τ-Euclidean metric is first order stochastic dominance over the Euclidean and the Kendall's τ metric, indicating a superiority of τ-Euclidean metric.

4. Empirical analysis

In this section, the LSHAC models are applied to financial data using the three-stage estimation methodology proposed in the preceding sections. Gaussian copula and Student's t copula are used as the benchmark models. In term of model comparison, we resort to the commonly used criteria: Bayesian information criterion (BIC; Schwarz, 1978). In order to present economic benefit of applying LSHAC models, we perform out-of-sample Value-at-Risk (VaR) forecasting for an equally-weighted portfolio. Finally, the probabilities of the market downward co-movements are calculated based on the VaR and the conditional tail expectation (CTE) criteria.

4.1. Data

Special attention is needed when studying co-movement probabilities considering that markets have different opening and closing time, which will, to some extent, drive market dependence. In order to address this issue, we consider the exchange traded funds (ETF)

[3] There are 16 pairs of $(\alpha,\ \beta)$ when $\alpha,\ \beta \in \{0,1,2,3\}$. We exclude the pairs $(\alpha = 2, \beta = 0)$ and $(\alpha = 3, \beta = 0)$, which are the same as $(\alpha = 1, \beta = 0)$; the pairs $(\alpha = 0, \beta = 2)$ and $(\alpha = 0, \beta = 3)$, which are the same as $(\alpha = 0, \beta = 1)$; the pairs $(\alpha = 2, \beta = 2)$ and $(\alpha = 3, \beta = 3)$, which are the same as $(\alpha = 1, \beta = 1)$; as well as the pair $(\alpha = 0, \beta = 0)$, leading to a total of 9 effective pairs of (α, β).

Table 3
Mean and Variance of the Reliability Ratios (ρ_n) from different dissimilarity metrics. The Euclidean metric is a special case of the τ-Euclidean metric when $\alpha = 1$ and $\beta = 0$, with its results listed in the first two rows. The Kendall's τ metric is a special case of the τ-Euclidean metric when $\alpha = 0$ and $\beta = 1$, and the mean and variance of its reliability ratios are listed in the third and forth rows. The best structure estimating results are achieved when $\alpha = 2$ and $\beta = 3$, with the average reliability ratio being 95.62% when $\pi = 5\%$.

Dissimilarity metric	Statistics	$\pi = 5\%$	$\pi = 10\%$	$\pi = 20\%$
Euclidean	Mean	0.6590	0.6574	0.6392
$(\alpha = 1, \beta = 0)$	Variance	0.0018	0.0040	0.0134
Kendall's τ	Mean	0.8700	0.8616	0.8466
$(\alpha = 0, \beta = 1)$	Variance	0.0045	0.0113	0.0350
τ-Euclidean	Mean	0.8384	0.8323	0.8241
$(\alpha = 1, \beta = 1)$	Variance	0.0023	0.0054	0.0179
τ-Euclidean	Mean	0.9131	0.9057	0.8920
$(\alpha = 1, \beta = 2)$	Variance	0.0011	0.0029	0.0092
τ-Euclidean	Mean	0.9439	0.9394	0.9228
$(\alpha = 1, \beta = 3)$	Variance	0.0007	0.0016	0.0058
τ-Euclidean	Mean	0.9049	0.8969	0.8814
$(\alpha = 2, \beta = 1)$	Variance	0.0014	0.0032	0.0104
τ-Euclidean	Mean	0.9562	0.9490	0.9281
$(\alpha = 2, \beta = 3)$	Variance	0.0005	0.0010	0.0035
τ-Euclidean	Mean	0.9446	0.9412	0.9193
$(\alpha = 3, \beta = 1)$	Variance	0.0006	0.0013	0.0048
τ-Euclidean	Mean	0.9187	0.9121	0.8926
$(\alpha = 3, \beta = 2)$	Variance	0.0008	0.0013	0.0037

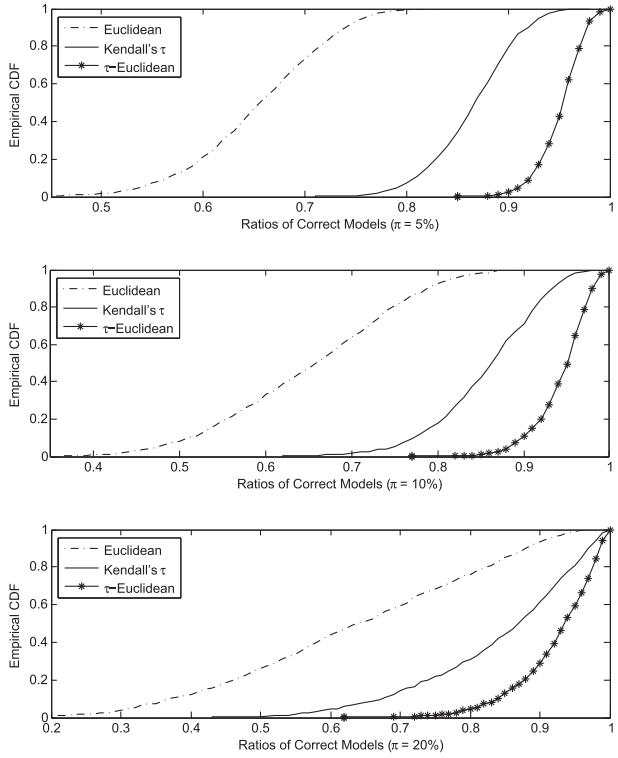


Fig. 2. Empirical cumulative distribution functions (eCDF) of the reliability ratios ($\boldsymbol{\rho_n}$). From the top to the bottom, $\pi = 5\%$, 10%, and 20% under the three proximity metrics. Dotted lines are for Euclidean metric; solid lines are for Kendall's τ metric; Lines with stars are for τ-Euclidean metric when $\alpha = 2$ and $\beta = 3$.

that are invested in the U.S. ETF data have the advantage of being traded at the same time, making it more convenient to study the joint market co-movement. Our sample portfolio consists of daily log-returns of six U.S. ETFs. Specifically, they are iShare MSCI Germany ETF (EWG), iShare MSCI France ETF (EWQ), iShare MSCI Netherlands ETF (EWN), iShare MSCI Dow Jones ETF (IYY), iShare MSCI Australia ETF (EWA), iShare MSCI Hong Kong ETF (EWH).[4]

To ensure geographical diversification, we ensure that the selected ETFs cover a diverse set of international markets, with three ETFs from the European market (EWG, EWQ, EWN), one from the American market (IYY), one from the Australia market (EWA), and one from the Asian market (EWH). The complete sample covers the period from January 1, 2011 to September 30, 2015, including 1193 trading days. The sample period covers a number of important events including the post European sovereign debt crisis, the sharp drops in the "August 2011 stock markets fall", and the "2015 stock market selloff". Log returns of the six ETFs are displayed in Fig. 3.

4.2. Estimation results

To consider the stylized facts of financial data, such as volatility clustering and positive excess kurtosis (McNeil et al., 2010), in the first stage, the generalized autoregressive conditional heteroskedasticity (GARCH $(1, 1)$[5]) models (Engle, 1982; Bollerslev, 1986) are used. The GARCH (1,1) model are specified as follows:

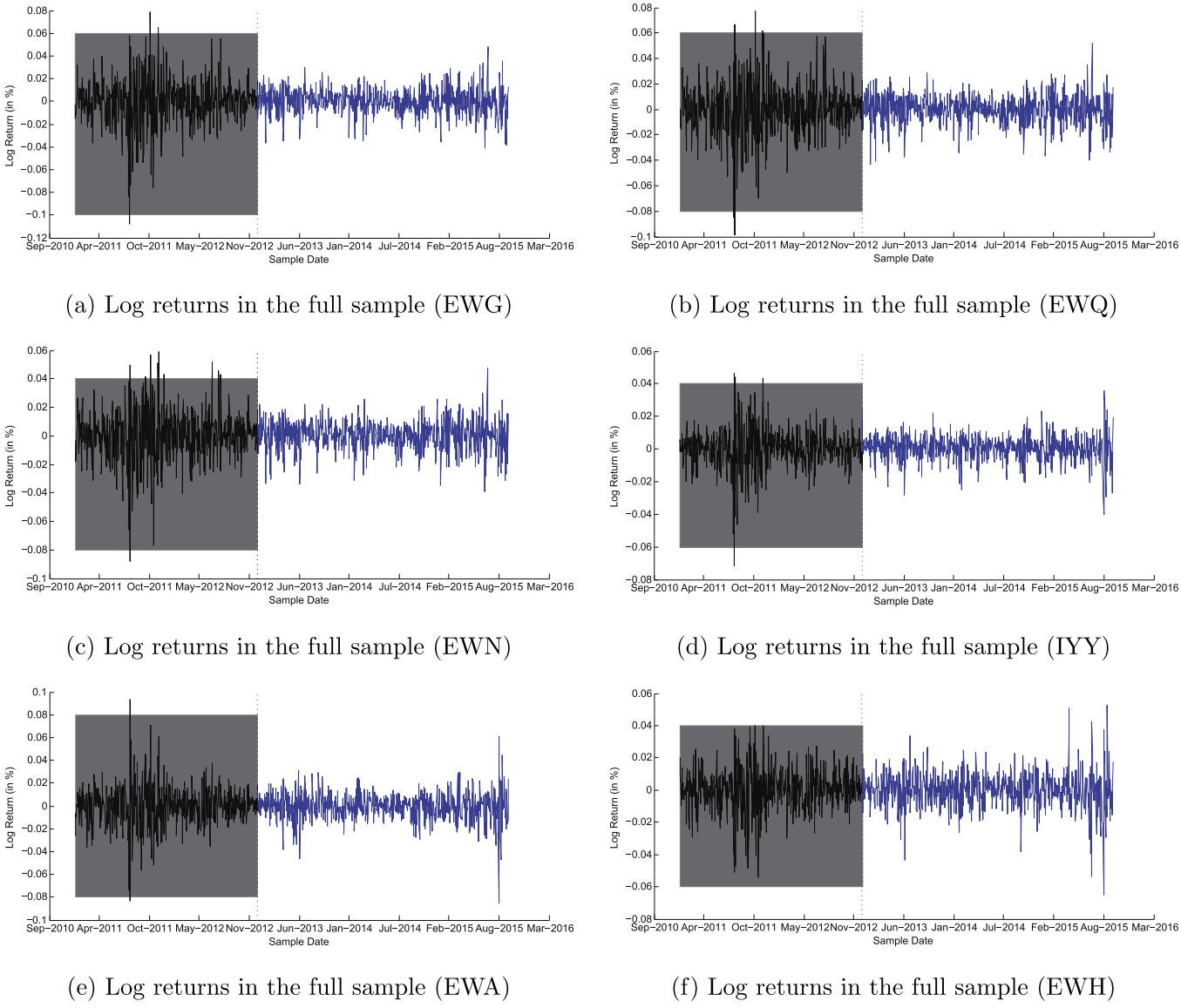
$$r_{i,t} = c_i + \epsilon_{i,t}, \quad \epsilon_{i,t} = \sqrt{h_{i,t}} z_{i,t}, \quad h_{i,t} = \omega_i + \eta_{i,1} h_{i,t-1} + \xi_{i,1} \epsilon^2_{i,t-1},$$

where $r_{i,t}$ is the log-return of the ith index at time t; $\epsilon_{i,j}$ and $z_{i,t}$ are the residual and standard residual of the ith index at time t, respectively; $h_{i,t}$ is the conditional variance of the ith index at time t based on the information at $t-1$; and c_i, ω_i, $\eta_{i,1}$, and $\xi_{i,1}$ are parameters. In addition, to capture the heavy tails of the data, the standard residuals $z_{i,t}$ are modeled as the generalized hyperbolic (GH) distribution (Barndorff-Nielsen, 1997). Hyperbolic, Normal-inverse Gaussian (NIG), student's t, and Variance Gamma (VG) are included within the GH family.

Table 4 exhibits the empirical results for the margins, together with the standard errors of the parameters given in the parenthesis. The best fitted residual models are Student's t for EWQ, EWA, as well as EWH, and VG for EWG, EWN and IYY. Pseudo sample data, obtained from probability transformations, are used to estimate LSHACs in the second and the third stages of the estimation procedure.

[4] The data were collected from finance.yahoo.com.

[5] In our analysis, the optimal order of the autoregressive moving average-generalized autoregressive conditional heteroskedasticity (ARMA (m,n)-GARCH (p, q)) model are selected based on BIC. GARCH (1,1) model is selected as the optimal model for all the six stock indices.



(a) Log returns in the full sample (EWG)

(b) Log returns in the full sample (EWQ)

(c) Log returns in the full sample (EWN)

(d) Log returns in the full sample (IYY)

(e) Log returns in the full sample (EWA)

(f) Log returns in the full sample (EWH)

Fig. 3. Time series plots of log returns in the full sample. The figure shows plots of the log returns on the six ETFs (EWG, EWQ, EWN, IYY, EWA, EWH). The complete sample is divided into the in-sample period from January 1, 2011 to December 31, 2012, including 501 trading days (shaded in gray), and the out-of-sample period from January 1, 2013 to September 30, 2015, including 692 trading days (highlighted in blue). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

In the second stage, the τ-Euclidean metric in Eq. (8) based on hierarchical clustering analysis is applied to the pseudo samples obtained from the first stage. In order to be consistent with the simulation study in Section 3, we use the 9 pairs of $(\alpha,\ \beta)$ listed in Table 3 to estimate the hierarchical structure of the ETF data. For this data set, the 9 pairs of $(\alpha,\ \beta)$ have identified five distinct hierarchical structures, as depicted in Fig. 4.

The estimated hierarchical structures based on the Euclidean metric $(\alpha = 1,\ \beta = 0)$, the Kendall 's τ metric $(\alpha = 0,\ \beta = 1)$ and the τ-Euclidean metric with $\alpha = \beta = 1$ are identical. This is a four-level structure and is denoted by $\widehat{\mathcal{S}_1}$ in Fig. 4a. This structure first isolates the Asian EWH from the rest of the ETFs. Among the remaining ETFs, the European ETFs EWG and EWQ are firstly grouped together, then joined by EWN. Meanwhile, IYY from the U.S. and EWA from Australia are in the other subgroup. This indicates that based on the three proximity metrics, i.e., $(\alpha = 0, \beta = 1)$, $(\alpha = 1, \beta = 0)$, and $(\alpha = 1, \beta = 1)$, the Hong Kong market is less similar to the ETFs from the other markets. In addition, $\widehat{\mathcal{S}_1}$ also indicates that within the European subgroup, Germany market and France market are more alike compared to the Netherlands market.

The second type of structure, dented by $\widehat{\mathcal{S}_2}$, is estimated by the pair $(\alpha = 1, \beta = 2)$. It is a four-level structure graphically represented in Fig. 4b. This structure collects the two ETFs from Germany (EWG) and Netherlands (EWN) into one group, and groups the remaining indices together. The four dimensional part on the left of the hierarchy is a nested structure in which EWQ from France and EWA from Australia are firstly grouped into a subgroup, then joined together by EWH from Hong Kong at the next level, with IYY from the U.S. joined finally. This structure separates the Germany and Netherlands from the other markets.

$\widehat{\mathcal{S}_3}$ and $\widehat{\mathcal{S}_4}$ have the same hierarchical structure as $\widehat{\mathcal{S}_1}$, but with different combinations of ETFs. $\widehat{\mathcal{S}_4}$, which is estimated based on $(\alpha = 2, \beta = 1), (\alpha = 3, \beta = 1)$, and $(\alpha = 3, \beta = 2)$, is almost the same as $\widehat{\mathcal{S}_1}$ except that in the three dimensional part on the left of the

Table 4
Marginal estimation results. The standard errors of estimated parameters are listed in the parenthesis. ν, α, and β are parameters of the residual distributions.

	Europe				America	
	EWG	EWQ	EWN	IYY	EWA	EWH
Distribution	VG	t	VG	VG	t	t
$c(\times 10^{-4})$	7.4800 (0.0156)	9.8347 (0.0094)	4.4019 (0.0248)	5.7783 (0.0736)	8.5840 (0.0688)	6.2272 (0.0918)
$\omega(\times 10^{-6})$	3.8456 (0.0192)	7.2770 (0.01328)	2.7497 (0.0193)	3.9969 (0.0144)	3.0814 (0.0131)	6.0600 (0.0371)
η_1	0.0888 (0.0038)	0.0865 (0.0121)	0.0702 (0.0055)	0.1262 (0.0005)	0.0966 (0.0019)	0.1131 (0.0003)
ξ_1	0.9046 (0.0000)	0.8984 (0.0000)	0.9217 (0.0100)	0.8450 (0.0000)	0.8957 (0.0000)	0.8543 (0.0001)
ν	–	7.2813 (0.0005)	–	–	7.9848 (0.0058)	7.9689 (0.0011)
α	2.1555 (0.0000)	–	2.3119 (0.0001)	2.2446 (0.0022)	–	–
β	−0.4055 (0.0008)	–	−0.3402 (0.0176)	−0.3068 (0.0154)	–	–

hierarchy, EWG and EWN are first grouped together then joined by EWQ. In contrast, $\widehat{\mathcal{S}_1}$ first collects EWG and EWQ and then groups EWN in the three dimensional subgroup. $\widehat{\mathcal{S}_3}$ $(\alpha = 1, \beta = 3)$, on the other hand, is drastically different. It firstly separates IYY from the U.S. and collects the remaining indices together. In the five dimensional part, EWG from Germany and EWN from Netherlands are in the same subgroup while EWQ from France, EWA from Australia and EWH from Hong Kong are in the other subgroup.

The hierarchical structure estimated with $(\alpha = 2, \beta = 3)$, denoted by $\widehat{\mathcal{S}_5}$, is graphically represented in Fig. 4e. The resulting structure first partitions the ETFs into two broad groups of European market (EWG, EWQ, and EWN) and non-European market (IYY, EWA, and EWH). In addition, within the European subgroup, EWG and EWQ are in the same small subgroup. Similarly, IYY and EWA are classified together within the non-European subgroup, implying that the U.S. market and Australia market are more similar when compared to the market between the U.S. and Hong Kong.

For the final and third stage of the estimation procedure, the fitting results of the LSHAC models with the five structures are compared, with the outer copulas and the Lévy subordinators chosen from Tables 1 and 2, respectively. The LSHACs are highly flexible with a large number of candidate models. To be more specific, to calibrate the six dimensional LSHAC with five generators in Fig. 4, we have $5 \times 3^4 = 405$ candidates for each structure. For the length restriction of this paper, only the best and the worst LSHAC models for each structure are displayed in Table 5. Elliptical copulas including Gaussian copula and t copula are also calibrated to serve as benchmarks.

From the fitting results in Table 5, the best LSHAC model for $\widehat{\mathcal{S}_4}$ is given by the C_{12} family, and is better than both Gaussian and Student's t copula, which is highlighted in bold with two stars in the parenthesis as the best LSHAC model. In addition, the best LSHAC model for $\widehat{\mathcal{S}_1}$, which is constructed from the C_{14} family, is better than Gaussian copula. Compared to the elliptical copulas, LSHAC models have the following two advantages. First, LSHACs generally have fewer parameters, increasing the efficiency of the models (see the last column in Table 5). Second, LSHAC models are very flexible in modeling the tail dependence. Gaussian copula has no tail dependence while Student's t copula have symmetric tail dependence. These tail dependence characteristics are not suitable for financial data. In contrast, the LSHACs is able to provide much greater flexibility in modeling the tail dependence. In particular, according to Table 1, C_{12} and C_{14} families are able to model asymmetric tail dependence, which according to the estimated results in Table 5, are more applicable to the financial data considered in this paper.

To further analyze the modeling abilities of different models, similar to Okhrin et al. (2013b), we also perform the moving window procedure, where each copula model are estimated with the sample from a moving window with the window width of 100 observations. The resulting series of BIC values are calculated and displayed in Fig. 5, including six copula models (Student's t copula, best LSHAC copula with structure $\widehat{\mathcal{S}}_j$, $j = 1, \ldots, 5$). The LSHAC models have the best performances for 89.55% ($\widehat{\mathcal{S}_4}$: 84.82%, $\widehat{\mathcal{S}_5}$: 4.73%) of the days, while the Student's t copula dominates the remaining 10.45% times.

Based on the estimating results in Table 5 and Fig. 5, the hierarchical structures of the LSHAC models have considerable effects in the estimation results. For our data set, $\widehat{\mathcal{S}_4}$ offers better fitting results compared to the other structures. Moreover, the moving window results show that the best LSHAC with $\widehat{\mathcal{S}_4}$ dominates 84.82% of the times while the LSHAC with $\widehat{\mathcal{S}_5}$ dominates only 4.73%.

Finally, it is worth emphasizing that all the other LSHAC models perform much better than the All-GM-HAC. Recall that the All-GM-HAC is the most popular HAC model currently used in the empirical analysis. By constructing HACs involving the Lévy subordinators, the LSHAC broadens the HAC family and hence provides greater flexibility for practical applications. The proximity metric proposed in this paper further facilitates the determination of the optimal structure of LSHACs, greatly enhancing the applicability of LSHAC models.

4.3. Backtesting

In the preceding sections, the LSHAC models have shown better fitting abilities based on information criteria. In this section, we illustrate the superiority of the LSHAC models by comparing their out-of-sample Value-of-Risk (VaR) forecasting with the benchmark Student's t copula. VaR is one of the most popular measures used by banks and insurance companies for quantifying risks. Loosely speaking, the VaR of a return random variable R at a given level $\alpha_0 \in (0, 1)$ corresponds to the worst return at probability α_0. For more information about VaR and backtesting, refer to McNeil et al. (2010, pp. 55–59).

The objective of our backtesting analysis is to calculate VaR at $\alpha_0 = 95\%$ and $\alpha_0 = 99\%$ for each trading day, t, in the out-of-sample period from January 1, 2013 to September 30, 2015. The copulas we tested in this study include Student's t copula, best LSHAC copula with structure $\widehat{\mathcal{S}}_j$, $j = 1, \ldots, 5$. We use the 501 in-sample historical data (from January 1, 2011 to December 31, 2012), $\boldsymbol{X}_s = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_d)_{s\times d}$, where $s = t - 500, t - 499, \ldots, t$, to make VaR forecast for day $t + 1$ with different copula models as follows:

Step 1: Starting from t, we generate N pseudo samples from a LSHAC with parameters $\widehat{\theta^C_{\text{LSHAC}}}$ and structure $\widehat{\mathcal{S}}_j, j = 1, \ldots, 5$, or from a t copula with parameters $\widehat{\theta^C_t}$.

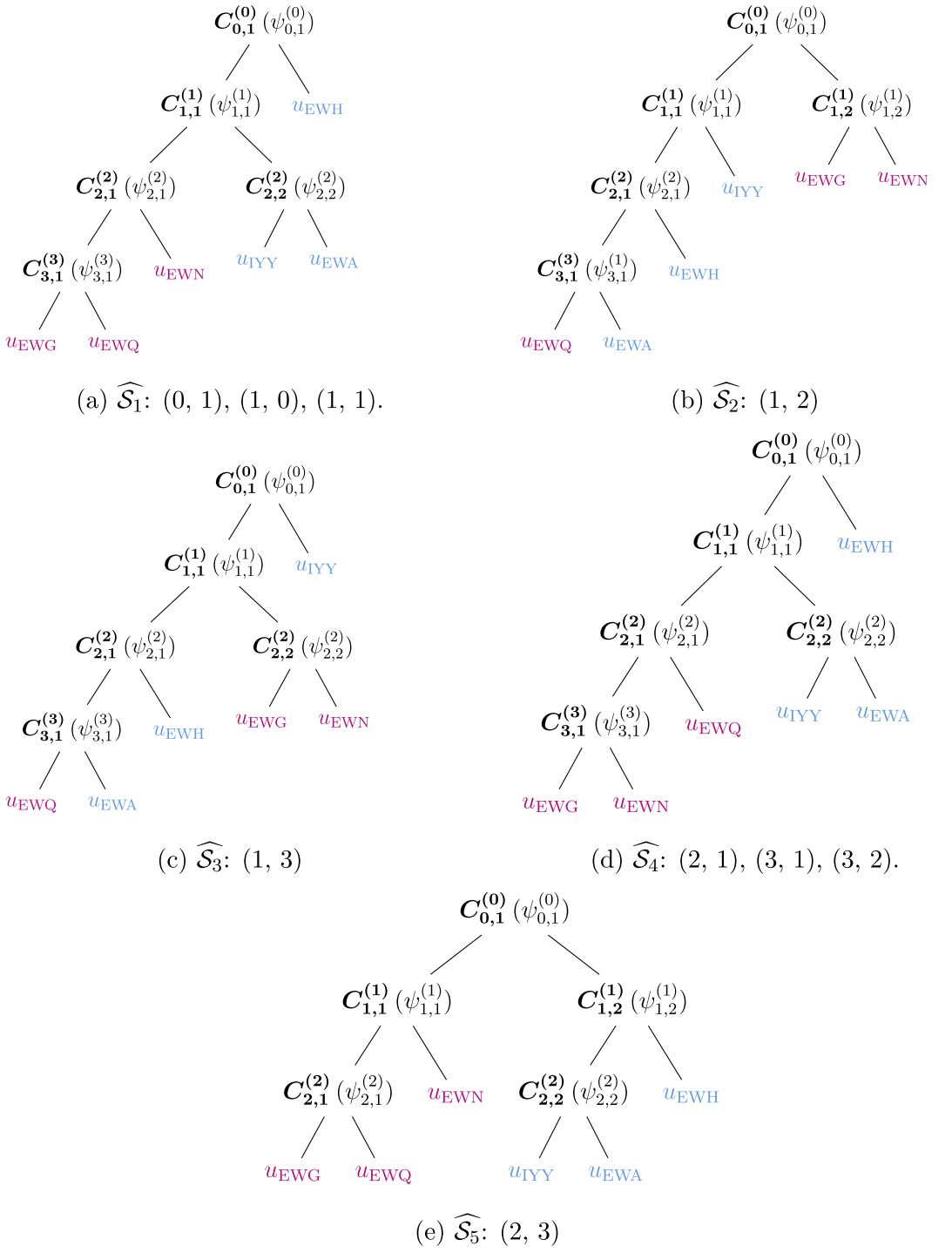


Fig. 4. Estimated hierarchical structures. The figure shows estimated tree structures. In consistent with the simulation analysis in Section 3, results are reported for $\alpha \in \{0,1,2,3\}$ and $\beta \in \{0,1,2,3\}$. Each sub-caption shows the pairs of (α,β) that estimate the corresponding hierarchical structures. Note that when $\alpha = 1,\ \beta = 0,\ \tau$-Euclidean degenerates to the Euclidean metric, and when $\alpha = 0,\ \beta = 1$, it degenerates to the Kendall's τ metric. Countries in the European markets are in red while the other countries from non-European markets are in green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Step 2: For each nth pseudo sample, $n = 1,\ldots,N$, we recover the margins with parameters $\widehat{\boldsymbol{\theta}^M}$. More specifically, we apply the inverse probability transform to convert them to the estimated GH distributions and hence the log-returns, $\widetilde{\boldsymbol{r}}^n_{t+1} = (\widetilde{r}^n_{1,t+1},\ldots,\widetilde{r}^n_{6,t+1})$, according to the estimated GARCH(1,1) model.

Step 3: Compute the simulated portfolio return as $r_{p,t+1} = \sum_{d=1}^{6} \frac{1}{6} r_{d,t+1}$, and estimate the sample α_0-level VaR for the day $t+1$.

Step 4: Update the information set with the actual portfolio return, $r_{p,t+1}$. Let $t = t+1$, reestimate all parameters and repeat the first three steps to forecast the portfolio VaR for the next day.

Table 5
Fitting results for the five structures in Fig. 4. This table shows copula estimating results, with the first two rows presenting results with Gaussian and Student's t copula as competing models. When estimating, we go through LSHAC models with the outer copulas chosen from Table 1 and the Lévy subordinators selected from Table 2. However, for the length restriction of this paper, only the best and the worst LSHAC models for each structure are displayed. The best fitting result, highlighted in bold with two stars in the parentheses, is given by a LSHAC with structure $\hat{\mathcal{S}}_4$ from the C_{12} family, which has asymmetric upper and lower tail dependence. The last two columns show the BIC values and the number of parameters of each model.

Model						BIC	No. para
Elliptical copulas							
Gaussian copula						−1782.0	15
Student's t copula						−1844.9	16
Structure	Generators in each level					BIC	No. para
LSHAC							
$\widehat{\mathcal{S}_1}$	$\mathbf{C}^{(0)}_{0,1}$	$\mathbf{C}^{(1)}_{1,1}$	$\mathbf{C}^{(2)}_{2,1}$	$\mathbf{C}^{(2)}_{2,2}$	$\mathbf{C}^{(3)}_{3,1}$		
	GM	GM	GM	GM	GM	−1543.71	5
	C_{14}	GM	IG	IG	GM	−1810.58	7
$\widehat{\mathcal{S}_2}$	$\mathbf{C}^{(0)}_{0,1}$	$\mathbf{C}^{(1)}_{1,1}$	$\mathbf{C}^{(2)}_{1,2}$	$\mathbf{C}^{(2)}_{2,1}$	$\mathbf{C}^{(3)}_{3,1}$		
	GM	GM	GM	GM	GM	−1472.75	5
	C_{12}	GM	GM	GM	GM	−1717.72	5
$\widehat{\mathcal{S}_3}$	$\mathbf{C}^{(0)}_{0,1}$	$\mathbf{C}^{(1)}_{1,1}$	$\mathbf{C}^{(2)}_{2,1}$	$\mathbf{C}^{(2)}_{2,2}$	$\mathbf{C}^{(3)}_{3,1}$		
	GM	GM	GM	GM	GM	−1472.87	5
	$\mathbf{C}_{12}$	GM	GM	GM	GM	−1717.71	5
$\widehat{\mathcal{S}_4}$	$\mathbf{C}^{(0)}_{0,1}$	$\mathbf{C}^{(1)}_{1,1}$	$\mathbf{C}^{(2)}_{2,1}$	$\mathbf{C}^{(2)}_{2,2}$	$\mathbf{C}^{(3)}_{3,1}$		
	GM	GM	GM	GM	GM	−1583.37	5
	$\mathbf{C}_{12}$	GM	GM	GM	IG	**−1851.84** (**)	6
$\widehat{\mathcal{S}_5}$	$\mathbf{C}^{(0)}_{0,1}$	$\mathbf{C}^{(1)}_{1,1}$	$\mathbf{C}^{(2)}_{1,2}$	$\mathbf{C}^{(2)}_{2,1}$	$\mathbf{C}^{(2)}_{2,2}$		
	GM	GM	GM	GM	GM	−1510.95	5
	$\mathbf{C}_{12}$	IG	IG	GM	GM	−1778.62	7

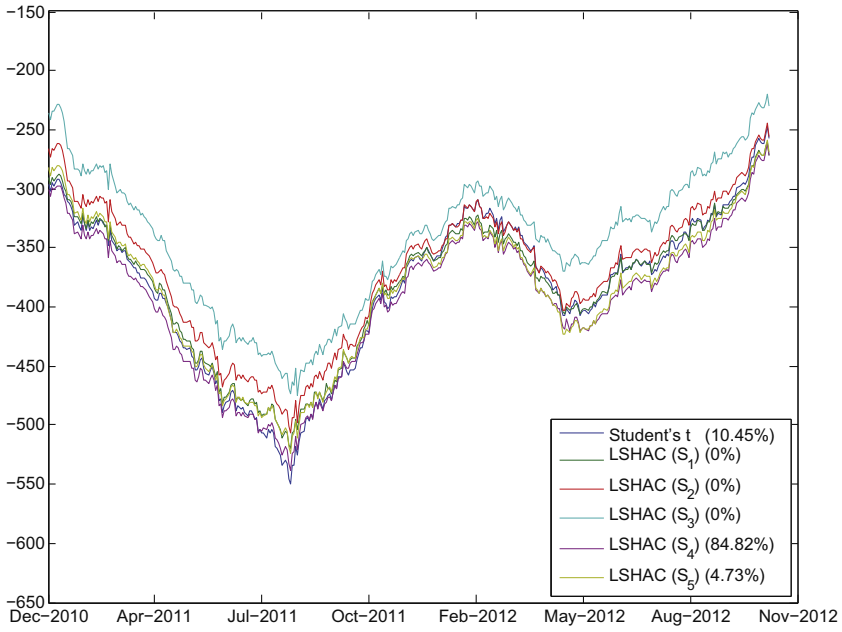


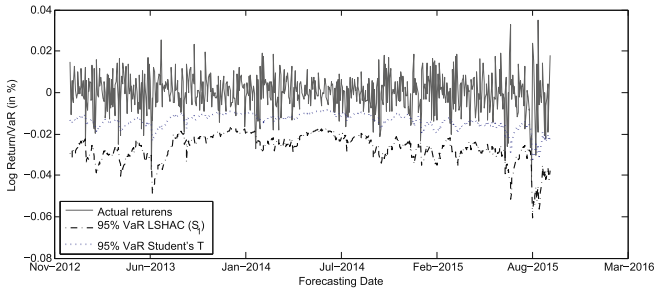
Fig. 5. BIC values from moving window estimation. This figure shows the moving window estimation results with window width of 100 observations. Copula models are shown in the legend with the percentage of the days, when the particular models is dominating, displayed in the corresponding parentheses.

In our simulation, we set $N = 100,000$. With the simulated VaR, we calculate the number of violations and collect the results in Table 6. Figs. 6 and 7 show the 95%- and 99%-level VaR estimates, respectively, as well as the actual portfolio returns. The out-of-sample period includes the whole year in 2013 (252 trading days), the whole year in 2014 (252 trading days) and January 1–September 30 in 2015 (188 trading days). In Table 6, the expected violations as well as the realized violations with different models are reported. The "Error" columns show for each dependence model the average discrepancy per year between realized and expected number of violations. The backtesting results with the LSHAC model with structures $\widehat{\mathcal{S}_2}$, $\widehat{\mathcal{S}_3}$, $\widehat{\mathcal{S}_4}$, and $\widehat{\mathcal{S}_5}$ are quite remarkable since the three-year average error of exceedances is −0.2 for 95%-VaR and −0.64 for 99%-VaR. In contrast, the three-year average error based on Student's t copula is 3.13 for 95% and 1.03 for 99%, respectively.

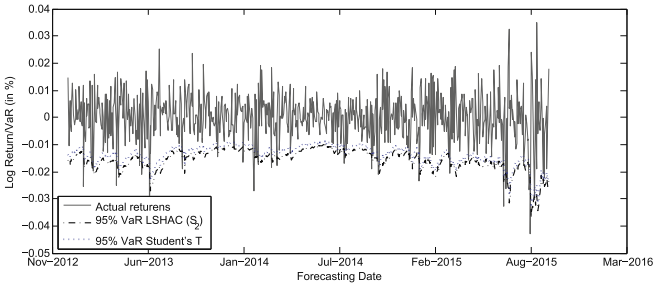
It is also noteworthy that although four out of five LSHACs perform better than the benchmark model, the LSHAC with $\widehat{\mathcal{S}_1}$ is overly conservative in estimating the portfolio VaR and seems to overestimate the portfolio risk. It indicates that although the LSHAC model

Table 6
Number of violations of the 95% and 99% VaR estimate calculated. The backtesting results for the 95% significant level are reported in the third to sixth column and the backtesting results for the 99% significant level are reported in the seventh to tenth column. Expected number of violations are also reported. The two columns named "Error" show for each dependence model the average absolute discrepancy per year between realized and expected number of violations. The equal-weight portfolio consist of the returns of the six ETFs in Section 4.1 including EWA, EWG, EWH, EWN, EWQ, and IYY. The full sample covers the period from January 1, 2011 to September 30, 2015 (1193 trading days). The backtests are performed on the out-of-sample data consists of 692 trading days including the whole year of 2013 (252 trading days), the whole year of 2014 (252 trading days) and January 1–September 30 in 2015 (188 trading days).

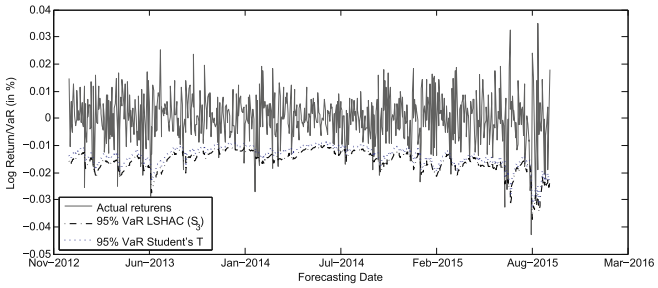
		Results for the 95% VaR estimate				Results for the 99% VaR estimate			
		2013	2014	2015	Error	2013	2014	2015	Error
Expected violations		12.6	12.6	9.4	–	2.52	2.52	1.88	–
Realized violations	LSHAC-$\widehat{\mathcal{S}_1}$	2	1	2	9.87	0	0	0	2.31
	LSHAC-$\widehat{\mathcal{S}_2}$	12	11	11	1.26	2	1	2	0.72
	LSHAC-$\widehat{\mathcal{S}_3}$	12	11	11	1.26	2	1	2	0.72
	LSHAC-$\widehat{\mathcal{S}_4}$	12	11	11	1.26	2	1	2	0.72
	LSHAC-$\widehat{\mathcal{S}_5}$	12	11	11	1.26	2	1	2	0.72
	Student's t	13	17	14	3.13	4	2	4	1.37



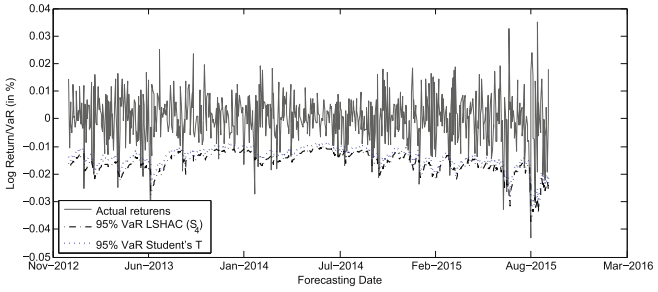
(a) Actual portfolio return, 95%-VaR forecasts: $\widehat{\mathcal{S}_1}$



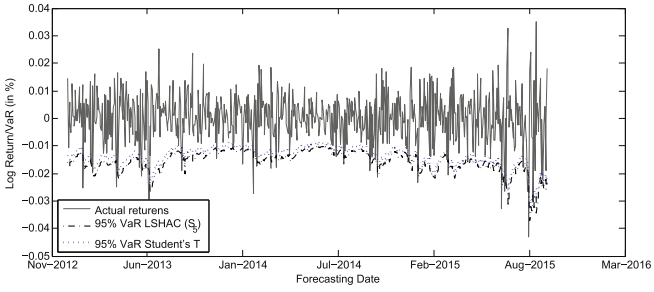
(b) Actual portfolio return, 95%-VaR forecasts: $\widehat{\mathcal{S}_2}$



(c) Actual portfolio return, 95%-VaR forecasts: $\widehat{\mathcal{S}_3}$

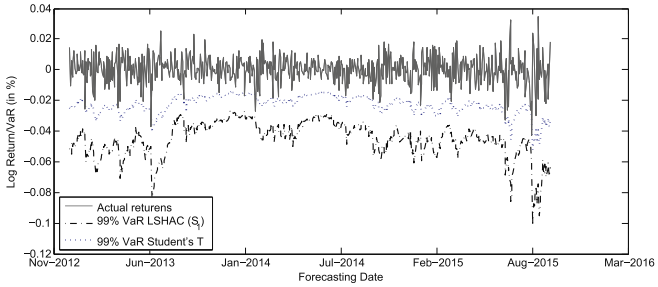


(d) Actual portfolio return, 95%-VaR forecasts: $\widehat{\mathcal{S}_4}$

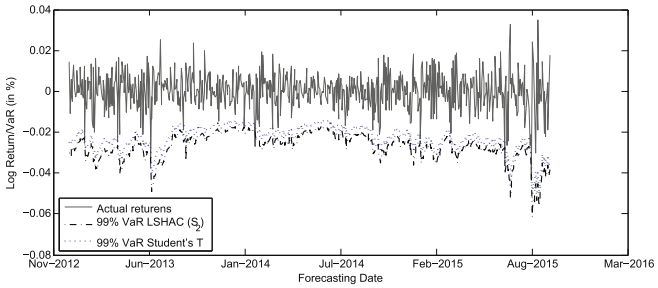


(e) Actual portfolio return, 95%-VaR forecasts: $\widehat{\mathcal{S}_5}$

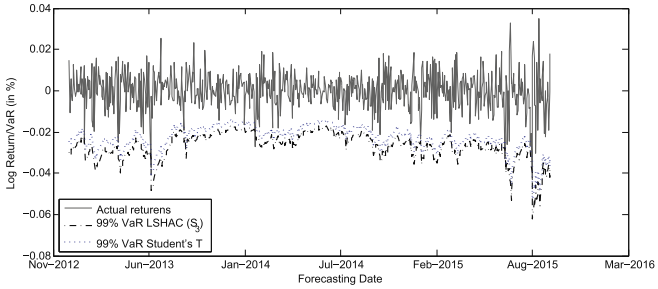
Fig. 6. Comparison of realized portfolio log returns and 95%-VaR forecasts. The figures show plots of the realized log returns on equally-weighted portfolio consisting of the six ETFs (EWG, EWQ, EWN, IYY, EWA, EWH) in solid dark gray lines. Each figure shows 95%-VaR forecasts using LSHAC models with one of the five estimated structures in dash-dot black lines. The VaR estimates based on Student's t copula is also shown as a benchmark in dashed blue lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



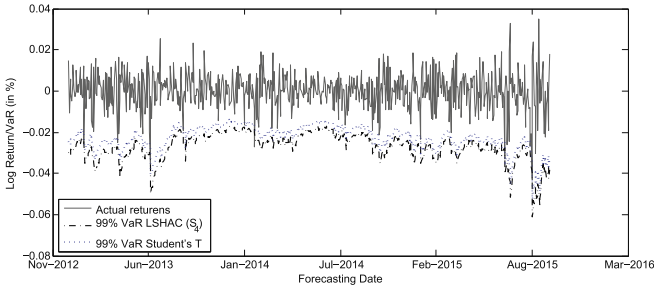
(a) Actual portfolio return, 99%-VaR forecasts: $\widehat{\mathcal{S}_1}$



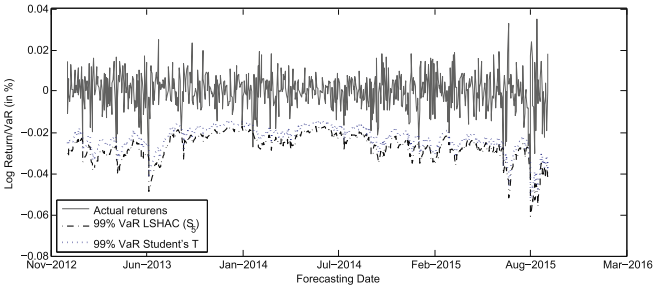
(b) Actual portfolio return, 99%-VaR forecasts: $\widehat{\mathcal{S}_2}$



(c) Actual portfolio return, 99%-VaR forecasts: $\widehat{\mathcal{S}_3}$



(d) Actual portfolio return, 99%-VaR forecasts: $\widehat{\mathcal{S}_4}$



(e) Actual portfolio return, 99%-VaR forecasts: $\widehat{\mathcal{S}_5}$

Fig. 7. Comparison of realized portfolio log returns and 99%-VaR forecasts. The figures show plots of the realized log returns on equally-weighted portfolio consisting of the six ETFs (EWG, EWQ, EWN, IYY, EWA, EWH) in solid dark gray lines. Each figure shows 99%-VaR forecasts using LSHAC models with one of the five estimated structures in dash-dot black lines. The VaR estimates based on Student's t copula is also shown as a benchmark in dashed blue lines. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

with $\widehat{\mathcal{S}_1}$ is competitive with the LSHAC with $\widehat{\mathcal{S}_4}$ from a statistical point of view (both models have very close BIC values), they may be very different economically, as shown in this VaR estimating example. Another possible explanation of this result is the tail dependence behavior of the two LSHAC models. To be more specific, the best LSHAC model of $\widehat{\mathcal{S}_1}$ is from C_{14} family, with a fixed lower tail dependence of 0.5. This restriction may lead to the unsatisfying VaR estimating ability of the C_{14}-LSHAC model. The LSHAC model of $\widehat{\mathcal{S}_4}$, on the other hand, is constructed with C_{12} family, which is more flexible in modeling both upper and lower tail dependence (see Table 1).

Based on results collected in Table 6, it is difficult to see the differences between the LSHAC models with structures $\widehat{\mathcal{S}_2}$, $\widehat{\mathcal{S}_3}$, $\widehat{\mathcal{S}_4}$, and $\widehat{\mathcal{S}_5}$. To further analysis the accuracy of the VaR estimates of different models, we perform the two-tailed conditional coverage backtest (CC Test) proposed in Christoffersen (1998) and the duration-based Weibull test of independence (WB Test) by Christoffersen (2004). The results are presented in Table 7. The results confirm our finding that the four LSHAC models with structures $\widehat{\mathcal{S}_2}$, $\widehat{\mathcal{S}_3}$, $\widehat{\mathcal{S}_4}$, and $\widehat{\mathcal{S}_5}$ are better than the benchmark Student's t copula most of the time and are also better than the LSHAC model with structure $\widehat{\mathcal{S}_1}$.

4.4. Extreme downward co-movement risk management

As cautioned in Kole et al. (2007) that the choice of copula models can have a profound effect on quantifying risk. It is therefore important to use a copula model that has the capability of reflecting key characteristics of the underlying problem. For quantifying financial risk involving multiple risk factors, the dependence, particularly the tail dependence, is one of the most important factors. In this subsection, the importance of choosing an appropriate copula model is further highlighted by analyzing the extreme downward co-movements. In our analysis, we use the same set of assets as in Section 4.1 to compute the following joint downward probability at time t:

$$\Pr(R_{1,t} \leqslant d_{1,t}, \ldots R_{6,t} \leqslant d_{6,t}) \tag{16}$$

Table 7
CC Test and WB Test results. This table presents the results of the two-tailed conditional coverage backtest (CC Test) proposed in Christoffersen (1998) and the duration-based Weibull test of independence (WB Test) by Christoffersen (2004). on the 95%- and 99%-level VaR estimates of the portfolio based on a variety of dependence models. The equal-weight portfolio consist of the returns of the six ETFs in Section 4.1 including EWA, EWG, EWH, EWN, EWQ, and IYY. The full sample covers the period from January 1, 2011 to September 30, 2015 (1193 trading days). The backtests are performed on the out-of-sample data consists of 692 trading days including the whole year of 2013 (252 trading days), the whole year of 2014 (252 trading days) and January 1–September 30 in 2015 (188 trading days).

Model	2013		2014		2015		2013–2015	
	CC test	WB test	CC test	WB test	CC test	WB test	CC test	WB test
LSHAC-$\widehat{\mathcal{S}_1}$	0.0008	0.2140	0.0001	0.0859	0.0113	0.7795	0.0000	0.0256
LSHAC-$\widehat{\mathcal{S}_2}$	0.1714	0.2081	0.2853	0.3507	0.4330	0.7518	0.1055	0.5163
LSHAC-$\widehat{\mathcal{S}_3}$	0.1797	0.9155	0.4045	0.5564	0.4325	0.7510	0.4844	0.9057
LSHAC-$\widehat{\mathcal{S}_4}$	0.1720	0.2081	0.4029	0.5573	0.4355	0.7457	0.1277	0.5968
LSHAC-$\widehat{\mathcal{S}_5}$	0.1712	0.2079	0.2861	0.3527	0.4338	0.7477	0.1052	0.5207
Student's t	0.1725	0.2075	0.1765	0.0811	0.4331	0.7518	0.0826	0.6129
LSHAC-$\widehat{\mathcal{S}_1}$	0.0000	0.0008	0.0000	0.0868	0.0000	0.0060	0.0000	0.0883
LSHAC-$\widehat{\mathcal{S}_2}$	0.9177	0.2146	0.543	0.9862	0.9654	0.7759	0.7090	0.3269
LSHAC-$\widehat{\mathcal{S}_3}$	0.9170	0.2151	0.5416	0.9851	0.9664	0.7806	0.7069	0.3261
LSHAC-$\widehat{\mathcal{S}_4}$	0.9187	0.2155	0.5423	0.9853	0.9684	0.7764	0.7117	0.3261
LSHAC-$\widehat{\mathcal{S}_5}$	0.9198	0.2143	0.5407	0.9943	0.9681	0.7738	0.7062	0.3253
Student's t	0.9161	0.2136	0.5386	0.9929	0.9612	0.7780	0.7099	0.3255

where $R_{i,t}$ and $d_{i,t}$ denote, respectively, the time-t return random variable and its time-t extreme downward return, where $i = 1, \ldots, 6$. In practice, the extreme downward returns $d_{i,t}$ are often quantified by some appropriately chosen risk measures. Besides VaR mentioned in the preceding sections, conditional tail expectation (CTE) is also a popular measure for risk management. CTE measures the average worst returns given that the return is worse than its VaR, therefore, CTE is a better measure of tail risk and is also commonly known as the Expected Shortfall (ES) or the Conditional Value at Risk (CVaR) (Föllmer and Schied, 2011).

By setting $d_{i,t},\ i = 1, \ldots, 6$, as $\mathrm{VaR}_{i,t}(\alpha_0)$, which is the time-$t$ VaR of the return of index i at confidence level α_0, the resulting downward co-movement probability (16) is termed as the probability of the "bad time" at time t (see also Kole et al., 2007). Similarly, when $d_{i,t},\ i = 1, \ldots, 6$, is the corresponding CTE, i.e. $\mathrm{CTE}_{i,t}(\alpha)$, then (16) becomes the probability of the "extreme bad time" at time t. The step-by-step estimating procedure for the probabilities of the "bad time" and the "extreme bad time" is summarized as follows:

Step 1: Calculate simulated log-returns, $\widetilde{\boldsymbol{r}}^n_{t+1} = (\widetilde{r}^n_{1,t+1}, \ldots, \widetilde{r}^n_{6,t+1})$, according to **Step 1** and **Step 2** in Section 4.3.

Step 2: For each ith asset, $i = 1, \ldots, 6$, we calculate $\widetilde{\mathrm{VaR}}_{i,t+1}(\alpha_0)$ as the α_0-quantile of the sample of $\widetilde{r}^1_{i,t+1}, \ldots, \widetilde{r}^N_{i,t+1}$, and compute the $\widetilde{\mathrm{CTE}}_{i,t+1}(\alpha_0)$ as the average of those samples that exceed $\widetilde{\mathrm{VaR}}_{i,t+1}(\alpha_0)$.

Step 3: Calculate the "bad time" and "extreme bad time" probabilities as follows,

$$\widetilde{P}^{\mathrm{bad}}_{t+1} = \sum_{n=1}^{N} \frac{\mathbb{I}_{\widetilde{\boldsymbol{r}}^n_{t+1} \leqslant \widetilde{\mathbf{VaR}}_{t+1}(\alpha)}}{N}, \quad \widetilde{P}^{\mathrm{ext}}_{t+1} = \sum_{n=1}^{N} \frac{\mathbb{I}_{\widetilde{\boldsymbol{r}}^n_{t+1} \leqslant \widetilde{\mathbf{CTE}}_{t+1}(\alpha)}}{N},$$

where $\widetilde{\boldsymbol{r}}^n_{t+1} = (\widetilde{r}^n_{1,t+1}, \ldots, \widetilde{r}^n_{6,t+1})$, $\widetilde{\mathbf{VaR}}_{t+1} = (\widetilde{\mathrm{VaR}}_{1,t+1}, \ldots, \widetilde{\mathrm{VaR}}_{6,t+1})$, $\widetilde{\mathbf{CTE}}_{t+1} = (\widetilde{\mathrm{CTE}}_{1,t+1}, \ldots, \widetilde{\mathrm{CTE}}_{6,t_0+1})$, $\mathbb{I}_{\widetilde{\boldsymbol{r}}^n_{t+1} \leqslant \widetilde{\mathbf{VaR}}_{t+1}(\alpha)}$ is an indicator representing the event of a "bad time", and $\mathbb{I}_{\widetilde{\boldsymbol{r}}^n_{t+1} \leqslant \widetilde{\mathbf{CTE}}_{t+1}(\alpha)}$ is an indicator of the event of an "extreme bad time".

Step 4: Update the information set with the actual portfolio return $r_{p,t+1}$. Let $t = t+1$, reestimate all parameters and repeat the first three steps to forecast the probabilities for the next day $t+1$. Then calculate the means and standard deviations of the estimated "bad time" and "extreme bad time" probabilities.

The downside comovement estimation results are displayed in Table 8. It is instructive to compare the probability of a "bad time" if all six indices were independent. By assuming that the probability of "bad time" for each asset is 0.05, then the joint probability for all six independent indices to experience a "bad time" is $(0.05)^6 = 1.56 \times 10^{-8}$. Hence, the results in Table 8 clearly highlights the importance of the dependence and hence the choice of copulas for modeling the financial market. It has been studied by the literature that elliptical copulas cannot provide sufficient estimates to the downside risk in the financial market due to their symmetric tail (in-)dependence (Weiß and Scheffer, 2015; Siburg et al., 2015). Therefore, given the financial turmoil that we have experienced and using Student's t copula as the benchmark model, we find that the Student's t copula produces probabilities of the "bad time" and the "extreme bad time" that seem too optimistic. On the other hand, the LSHAC models with asymmetric tail dependence appear to be more consistent with practice. It is also interesting to note that although the LSHAC model with structure $\widehat{\mathcal{S}_1}$ has substantially different VaR-forecast results, the downside comovement probabilities are similar to the other LSHAC models. Taking into consideration the number of parameters, the fitting abilities based on BIC presented in Section 4.2, the backtesting results in Section 4.3, as well as the flexibility in modeling the tail

Table 8
Estimation results of extreme downward co-movements. The first two columns of this table show the estimated mean probabilities of the extreme downward co-movements. Standard deviation of the estimations are displayed in the parentheses. The last two columns show the corresponding waiting time for the "bad time" and "extreme bad time" (denoted as "Ext. Time" in the table).

Model	Prob. of bad time	Prob. of ext. time	Wait of bad time	Wait of ext. time
$\widehat{\mathcal{S}_1}$	0.0040 (0.000195)	0.001396 (0.000119)	12.05	34.12
$\widehat{\mathcal{S}_2}$	0.0028 (0.000189)	0.001014 (0.000111)	16.76	46.98
$\widehat{\mathcal{S}_3}$	0.0029 (0.000138)	0.001022 (0.000076)	16.68	46.59
$\widehat{\mathcal{S}_4}$	0.0030 (0.000132)	0.002173 (0.000110)	15.83	21.91
$\widehat{\mathcal{S}_5}$	0.0029 (0.000134)	0.001044 (0.000075)	16.43	45.63
Student's t	0.0013 (0.000141)	0.000399 (0.000079)	35.89	119.35

dependence, these results lead to the justification of the LSHAC with structure $\widehat{\mathcal{S}_4}$.

5. Conclusions

High dimensional multivariate estimation has always been challenging, yet critical, in the financial and risk management modeling. Traditional elliptical copulas such as the Gaussian copulas and the Student's t copulas have limited flexibility, in addition to their undesirable quadratic growth in the number of parameters with dimension. These issues can be partially solved by using LSHACs to capture hierarchal dependence structure. The advantages of LSHAC models are their tractability to create new generators in terms of an outer generator and a series of Lévy subordinators, and more importantly, their flexibility to model the tail dependencies with the desirable hierarchical dependence structure that reflects the properties of the underlying data.

This paper constructs a generalized multi-level LSHAC model and designs an estimation procedure that focuses on a suitable grouping method to determine the hierarchical structure. We employ hierarchical clustering analysis with a newly proposed τ-Euclidean metric to determine the grouping of variables. The simulation study shows that the τ-Euclidean metric achieves a balance between Euclidean distance and association measures, providing a better performance in identifying the true structure. In the empirical analysis, the proposed estimation methodology is applied to estimating the out-of-sample Value-at-Risk and the extreme downside co-movement probabilities of a equally-weighted portfolio comprised of six ETFs from the European, American, Australian, and Asian markets. General LSHACs with the structure determined by τ-Euclidean metric produces better modeling performances and more accurate VaR forecasts than elliptical copulas and All-GM-HACs. In particular, with more flexibility to model the tail dependence by choosing appropriate outer generators, LSHAC models provide more accurate estimations to the joint downside movements and hence give better guidance to the financial risk managers. The empirical results also indicate that selecting a suitable structure of a LSHAC has a significant impact on high-dimensional estimation.

Appendix A. Notation system and sampling algorithm of a *L*-level LSHAC

It is indeed theoretically demanding to construct a LSHAC in a fully general setting. To this end, we provide a notation system and an integral representation (Marshall and Olkin, 1988; Joe, 1997; Whelan, 2004; McNeil, 2008; Hofert, 2008) of a general L-level LSHAC exhibited in Fig. 1 by introducing the following notation:

- For $l = 0, 1, \ldots, L$, let l denote the index level of LSHAC and J_l denote the number of copulas at level l.
- At level 0:
  - There is only one copula, denoted by $C_{0,1}^{(0)}$, and hence by construction $J_0 = 1$. This is also known as the outer copula.
  - There is a random time variable $V_{0,1}^{(0)}$ at which the Lévy subordinators for all subsequent groups are evaluated. We denote its corresponding c.d.f. as $G_{0,1}^{(0)}(v)$ and its LT-AC generator as $\psi_{0,1}^{(0)}$.
- At level l:
  - For $l = 1, \ldots, L$ and $j = 1, \ldots, J_{l-1}$, let $D_j^{(l)}$ be the number of copulas at level l "emanating" from the jth copula in the previous level $l-1$. Note that the following condition must hold:

$$\sum_{j=1}^{J_{l-1}} D_j^{(l)} = J_l, \quad l = 1, \ldots, L \tag{A.1}$$

and $J_L = d$.

  - Let $C_{j,k}^{(l)}$ with generator $\psi_{j,k}^{(l)}$ be the kth copula in the j-th cluster with size $D_j^{(l)}$. It is emanated from the jth copula at level $l-1$, for $l = 1, \ldots, L, j = 1, \ldots, J_{l-1}$, and $k = 1, \ldots, D_j^{(l)}$;
  - The mth adjacent copula emanated from $C_{j,k}^{(l)}$ at level $l+1$ is denoted as $C_{s,m}^{(l+1)}$, where $s = 1, \ldots, J_l$, is the position of $C_{s,m}^{(l+1)}$, satisfying

$$s = \left(\sum_{q=1}^{j-1} D_q^{(l)} \right) \mathbb{I}_{\{j>1\}} + k, \tag{A.2}$$

where $\mathbb{I}_{\{\cdot\}}$ is the indicator function.
- At level L:
  - We partition $(u_1, \ldots, u_d)$ into J_{L-1} groups and define

$$C_{j,k}^{(L)} = u_s, \quad s = 1, \ldots, J_L, \quad j = 1, \ldots, J_{L-1}, \quad k = 1, \ldots, D_j^{(L)}, \tag{A.3}$$

where s satisfies (A.2).
- Additional definitions:
  - Let $X(t)$ denote a Lévy subordinator evaluated at time t, with corresponding c.d.f $\widetilde{G}(x;t)$ and Laplace exponent $\widetilde{\Psi}$. Define function $F_{s_{l-1}j_l}^{(l)}(u)$ as

$$F_{s_{l-1}j_l}^{(l)}(u) = \exp(-\psi_{s_{l-1}j_l}^{(l)-1}(u)). \tag{A.4}$$

Given the above definition of $F_{s_{l-1}j_l}^{(l)}$, the following function

$$\left(F_{s_{l-1}j_l}^{(l)}(u)\right)^v = \exp\left(-v\psi_{s_{l-1}j_l}^{(l)-1}(u)\right) \tag{A.5}$$

is a valid c.d.f. for any positive v (Marshall and Olkin, 1988). Let $\Psi_{s_{l-1}j_l}^{(l)} = \psi_{s_{l-2}j_{l-1}}^{(l-1)-1} \circ \psi_{s_{l-1}j_l}^{(l)}$ be the Laplace exponent of a Lévy subordinator, $X_{s_{l-1}j_l}^{(l)}$, with c.d.f. $G_{s_{l-1}j_l}^{(l)}$, then the generator given by

$$\begin{aligned}\tilde{\psi}_{s_{l-1}j_l}^{(l)}(u;v) &= \left(F_{s_{l-2}j_{l-1}}^{(l-1)}\left(\psi_{s_{l-1}j_l}^{(l)}(u)\right)\right)^v \\ &= \exp\left(-v\psi_{s_{l-2}j_{l-1}}^{(l-1)-1} \circ \psi_{s_{l-1}j_l}^{(l)}(u)\right) \\ &= \exp\left(-v\Psi_{s_{l-1}j_l}^{(l)}(u)\right),\end{aligned} \tag{A.6}$$

is also a c.m. LT-AC generator, where $l = 2, \ldots, L-1$ (Feller, 2008; Nelsen, 2006).

Random samples from a LSHAC can be simulated relatively easily by recognizing that $\left(F_{s_{L-2}j_{L-1}}^{(L-1)}(C_{s_{L-1}j_L}^{(L)})\right)^{v_{s_{L-2}j_{L-1}}^{(L-1)}}$, where $j_L = 1, \ldots, D_{s_{L-1}}^{(L)}$, is a valid c.d.f. for any positive $v_{s_{L-2}j_{L-1}}^{(L-1)}$ (see (A.5), (A.6) and Theorem 2.1). More specifically, if $Y_{s_{L-1}j_L}$ is a uniform random variable on $(0,1)$, then given $V_{s_{L-2}j_{L-1}}^{(L-1)}$ with c.d.f. $\widetilde{G}_{s_{L-2}j_{L-1}}^{(L-1)}(x;t)$, a random sample of $C_{s_{L-1}j_L}^{(L)}$ can be obtained via inverse transform as

$$C_{s_{L-1}j_L}^{(L)} = \psi_{s_{L-2}j_{L-1}}^{(L-1)}\left(-\frac{\log(Y_{s_{L-1}j_L})}{V_{s_{L-2}j_{L-1}}^{(L-1)}}\right). \tag{A.7}$$

In summary, for a multi-level LSHAC with a general structure displayed in Fig. 1, the random samples can be simulated by a sequential procedure formally described in Algorithm Appendix A.1.

Algorithm Appendix A.1 (Sampling an L-level LSHAC).

Step 1: Generate a random variable $V_{0,1}^{(0)}$ with c.d.f. $G_{0,1}^{(0)}(x)$.

Step 2: For $l = 1,\ldots,L-1,\ s_{l-1} = 1,\ldots,J_{l-1},\ j_l = 1,\ldots,D_{s_{l-1}}^{(l)}$, generate a random variable $V_{s_{l-1}j_l}^{(l)}$ with c.d.f. $\widetilde{G}_{s_{l-1}j_l}^{(l)}(x; V_{s_{l-2}j_{l-1}}^{(l-1)})$.

Step 3: Generate a series of independent uniform random variables: $Y_{s_{L-1}j_L}, j_L = 1,\ldots,D_{s_{L-1}}^{(L)}$.

Step 4: Return $\overline{U}_{s_{L-1}j_L} = \psi_{s_{L-2}j_{L-1}}^{(L-1)}\left(-\frac{\log(Y_{s_{L-1}j_L})}{V_{s_{L-2}j_{L-1}}^{(L-1)}}\right) = \psi_{0,1}^{(0)} \odot_{i=1}^{L-1} \widetilde{\Psi}_{s_{i-1}j_i}^{(i)}\left(-\frac{\log(Y_{s_{L-1}j_L})}{V_{s_{L-2}j_{L-1}}^{(L-1)}}\right)$.

Then $(\overline{U}_{1,1},\ldots,\overline{U}_{s_{L-1}j_L}^{(L)},\ldots,\overline{U}_{s_{L-1}J_L}^{(L)})$ is a sample from copula $C(u_1,\ldots,u_d)$.

Note that when $L = 2$ and $d = J$, Algorithm Appendix A.1 reduces to the sampling algorithm of a two-level LSHAC proposed by Hering et al. (2010).

Appendix B. Proof of Theorem 2.1

We prove Theorem 2.1 by induction.

For level $y = 0$, since $\psi_{0,1}^{(0)}$ is a LT-AC generator with c.d.f $G(v_{0,1}^{(0)})$, we have

$$\psi_{0,1}^{(0)} = \int_0^\infty \exp\left\{-v_{0,1}^{(0)} \cdot u\right\} dG_{0,1}^{(0)}(v_{0,1}^{(0)}), \tag{B.1}$$

$$\left(F_{0,1}^{(0)}(u)\right)^v = \exp\left\{-v\psi_{0,1}^{(0)-1}(u)\right\}. \tag{B.2}$$

According to the definition of LSHAC, the copulas emanated from $C_{0,1}^{(0)}$ with generator $\psi_{0,1}^{(0)}$ are

$$\left\{C_{s_0j_1}^{(1)} | s_0 = 1,\quad j_1 = 1,\ldots,D_{s_0}^{(1)}\right\}. \tag{B.3}$$

Consequently, we have

$$C_{0,1}^{(0)} = C_{0,1}^{(0)}(C_{s_0,1}^{(1)},\ldots,C_{s_0,D_{s_0}^{(1)}}^{(1)}) = \psi_{0,1}^{(0)}\left(\sum_{j_1=1}^{D_{s_0}^{(1)}} \psi_{0,1}^{(0)-1}(C_{s_0j_1}^{(1)})\right) \tag{B.4}$$

Then, using (B.1) and (B.2) yields

$$\begin{aligned} C(u_1,u_2,\ldots,u_d) &= \int_0^\infty \prod_{j_1=1}^{D_{s_0}^{(1)}} \exp\left\{-v_{0,1}^{(0)}\psi_{0,1}^{(0)-1}(C_{s_0j_1}^{(1)})\right\} dG_{0,1}^{(0)}(v_{0,1}^{(0)}), \\ &= \int_0^\infty \prod_{j_1=1}^{D_{s_0}^{(1)}} \left(F_{0,1}^{(0)}(C_{s_0j_1}^{(1)})\right)^{v_{0,1}^{(0)}} dG_{0,1}^{(0)}(v_{0,1}^{(0)}), \\ &= \int_0^\infty \prod_{j_1=1}^{D_{s_0}^{(1)}} \left(F_{0,1}^{(0)}(C_{s_0j_1}^{(1)})\right)^{v_{0,1}^{(0)}} (dG)_{j_0}. \end{aligned} \tag{B.5}$$

Similarly, for level $y = 1$, the copulas emanated from $C_{s_0j_1}^{(1)}$ with generator $\psi_{s_0j_1}^{(1)}$ are

$$\left\{C_{s_1j_2}^{(2)} | s_1 = \left(\sum_{m=1}^{s_0-1} D_m^{(1)}\right)\mathbb{I}_{\{s_0>1\}} + j_1, j_2 = 1,\ldots,D_{s_1}^{(2)}\right\}. \tag{B.6}$$

Since $s_0 = 0$, it is equivalent to

$$\left\{C_{s_1j_2}^{(2)} | s_1 = j_1, j_2 = 1,\ldots,D_{s_1}^{(2)}\right\}. \tag{B.7}$$

Therefore, we have

$$C_{s_0j_1}^{(1)} = C_{s_0j_1}^{(1)}\left(C_{s_1,1}^{(2)},\ldots,C_{s_1,D_{s_1}^{(2)}}^{(2)}\right) = \psi_{s_0j_1}^{(1)}\left(\sum_{j_2=1}^{D_{s_1}^{(2)}} \psi_{s_0j_1}^{(1)-1}(C_{s_1j_2}^{(2)})\right). \tag{B.8}$$

Let

$$\tilde{\psi}_{s_0j_1}^{(1)}(u; v_{0,1}^{(0)}) = \left(F_{0,1}^{(0)}(\psi_{s_0j_1}^{(1)}(u))\right)^{v_{0,1}^{(0)}} = \exp\left\{-v_{0,1}^{(0)}\psi_{0,1}^{(0)-1} \circ \psi_{s_0j_1}^{(1)}(u)\right\} \tag{B.9}$$

and

$$\widetilde{\Psi}_{s_0j_1}^{(1)}(u) = \psi_{0,1}^{(0)-1} \circ \psi_{s_0j_1}^{(1)} \tag{B.10}$$

be the Laplace exponent of a Lévy subordinator, $X_{s_0j_1}^{(1)}$, with c.d.f. $\widetilde{G}_{s_0j_1}^{(1)}$. According to the property of Laplace exponent of a Lévy subordinator expressed in (2),

$$\begin{aligned} \tilde{\psi}_{s_0j_1}^{(1)}(u; v_{0,1}^{(0)}) &= \exp\left\{-v_{0,1}^{(0)}\psi_{0,1}^{(0)-1} \circ \psi_{s_0j_1}^{(1)}(u)\right\}, \\ &= \exp\left\{-v_{0,1}^{(0)}\widetilde{\Psi}_{s_0j_1}^{(1)}(u)\right\}, \\ &= E\left(\exp\left\{-uX_{J_0j_1}^{(1)}(v_{0,1}^{(0)})\right\}\right), \\ &= \int_0^\infty \exp\left\{-uv_{s_0j_1}^{(1)}(v_{0,1}^{(0)})\right\} d\widetilde{G}_{s_0j_1}^{(1)}(v_{s_0j_1}^{(1)}; v_{0,1}^{(0)}). \end{aligned} \tag{B.11}$$

As proved in Theorem 2.1 of Hering et al. (2010), derivative of $\psi_{0,1}^{(0)-1} \circ \psi_{s_0j_1}^{(1)}$ defined according to (B.10) is c.m. According to Joe (1997) and McNeil (2008), $\tilde{\psi}_{s_0j_1}^{(1)}(u)$ is a LT-AC generator. As a result, we can rewrite (B.5) according to (B.8) and (B.9),

$$\begin{aligned} C(u_1,\ldots,u_d) &= \int_0^\infty \prod_{j_1=1}^{D_{s_0}^{(1)}} \left(F_{0,1}^{(0)}(C_{s_0j_1}^{(1)})\right)^{v_{0,1}^{(0)}} dG_{0,1}^{(0)}(v_{0,1}^{(0)}), \\ &= \int_0^\infty \prod_{j_1=1}^{D_{s_0}^{(1)}} \tilde{\psi}_{s_0j_1}^{(1)}\left(\sum_{j_2=1}^{D_{s_1}^{(2)}} \psi_{s_0j_1}^{(1)-1}(C_{s_1j_2}^{(2)}); v_{0,1}^{(0)}\right) dG_{0,1}^{(0)}(v_{0,1}^{(0)}), \\ &= \int_0^\infty \prod_{j_1=1}^{D_{s_0}^{(1)}} \int_0^\infty \exp\left(-v_{s_0j_1}^{(1)}\left(\sum_{j_2=1}^{D_{s_1}^{(2)}} \psi_{s_0j_1}^{(1)-1}(C_{s_1j_2}^{(2)})\right)\right) \\ &\quad \times d\widetilde{G}_{s_0j_1}^{(1)}(v_{s_0j_1}^{(1)}; v_{0,1}^{(0)}) dG_{0,1}^{(0)}(v_{0,1}^{(0)}), \end{aligned} \tag{B.12}$$

Similarly, let $F_{s_0j_1}^{(1)}(u)$ satisfy

$$\left(F_{s_0j_1}^{(1)}(u)\right)^v = \exp\left(-v\psi_{s_0j_1}^{(1)-1}(u)\right), \tag{B.13}$$

then we have

$$\begin{aligned} &C(u_1,\ldots,u_d) \\ &= \int_0^\infty \prod_{j_1=1}^{D_{s_0}^{(1)}} \int_0^\infty \prod_{j_2=1}^{D_{s_1}^{(2)}} \left(F_{s_0j_1}^{(1)}(C_{s_1j_2}^{(2)})\right)^{v_{s_0j_1}^{(1)}} d\widetilde{G}_{s_0j_1}^{(1)}(v_{s_0j_1}^{(1)}; v_{0,1}^{(0)}) dG_{0,1}^{(0)}(v_{0,1}^{(0)}), \\ &= \int_0^\infty \prod_{j_1=1}^{D_{s_0}^{(1)}} \int_0^\infty \prod_{j_2=1}^{D_{s_1}^{(2)}} \left(F_{s_0j_1}^{(1)}(C_{s_1j_2}^{(2)})\right)^{v_{s_0j_1}^{(1)}} (dG)_{j_1}^{(1)}. \end{aligned} \tag{B.14}$$

Therefore, (3) is satisfied at level $y = 1$. Now let us assume at level $y : 0 \leqslant y \leqslant l-2 (l \geqslant 2)$, the following equation holds

$$C(u_1,\ldots,u_d) = \int_0^\infty \prod_{j_1=1}^{D_{s_0}^{(1)}} \int_0^\infty \prod_{j_2=1}^{D_{s_1}^{(2)}} \cdots \int_0^\infty \prod_{j_{y+1}=1}^{D_{s_y}^{(y+1)}} \left(F_{s_{y-1}j_y}^{(y)}(C_{s_yj_{y+1}}^{(l-1)})\right)^{v_{s_{y-1}j_y}^{(y)}} (dG)_{j_y}^{(y)}. \tag{B.15}$$

For notation consistency we let $s_{-1} = 0,\ j_0 = 1$. Then at level $y+1$, the copulas emanated from $C_{s_yj_{y+1}}^{(y+1)}$ with generator $\psi_{s_yj_{y+1}}^{(y+1)}$ are

$$\left\{C^{(y+2)}_{s_{y+1}j_{y+2}}|s_{y+1}=\left(\sum_{m=1}^{s_y-1}D^{(y+1)}_m\right)\mathbb{I}_{\{s_y-1>1\}}+j_{y+1},j_{y+2}=1,\ldots,D^{(y+2)}_{s_{y+1}}\right\}. \tag{B.16}$$

As a result, we have

$$C^{(y+1)}_{s_y j_{y+1}}=\psi^{(y+1)}_{s_y j_{y+1}}\left(\sum_{j_{y+2}=1}^{D^{(y+2)}_{s_{y+1}}}\psi^{(y+1)-1}_{s_y j_{y+1}}(C^{(y+2)}_{s_{y+1}j_{y+2}})\right). \tag{B.17}$$

Let

$$\tilde{\psi}^{(y+1)}_{s_y j_{y+1}}(u;v^{(y)}_{s_{y-1}j_y})=\left(F^{(y)}_{s_{y-1}j_y}(\psi^{(y+1)}_{s_y j_{y+1}}(u))\right)^{v^{(y)}_{s_{y-1}j_y}} \tag{B.18}$$

$$=\exp\left\{-v^{(y)}_{s_{y-1}j_y}\psi^{(y)-1}_{s_{y-1}j_y}\circ\psi^{(y+1)}_{s_y j_{y+1}}(u)\right\} \tag{B.19}$$

and

$$\widetilde{\Psi}^{(y+1)}_{s_y j_{y+1}}(u)=\psi^{(y)}_{s_{y-1}j_y}\circ\psi^{(y+1)}_{s_y j_{y+1}}(u) \tag{B.20}$$

be the Laplace exponent of a Lévy subordinator, denoted as $X^{(y+1)}_{s_y j_{y+1}}$, with c.d.f. $\widetilde{G}_{s_y,j^{(y+1)}_{y+1}}$. Consequently, substituting (B.20) into (B.19) yields

$$\begin{aligned}\tilde{\psi}^{(y+1)}_{s_y j_{y+1}}(u;v^{(y)}_{s_{y-1}j_y})&=\exp\left\{-v^{(y)}_{s_{y-1}j_y}\widetilde{\Psi}^{(y+1)}_{s_y j_{y+1}}(u)\right\}\\&=E\left(\exp\left\{-uX^{(y+1)}_{s_y j_{y+1}}(v^{(y)}_{s_{y-1}j_y})\right\}\right)\\&=\int_0^\infty\exp\left\{-uv^{(y+1)}_{s_y j_{y+1}}\right\}d\widetilde{G}^{(y+1)}_{s_y j_{y+1}}(v^{(y+1)}_{s_y j_{y+1}};v^{(y)}_{s_{y-1}j_y})\end{aligned} \tag{B.21}$$

Similarly, $\tilde{\psi}^{(y+1)}_{s_y j_{y+1}}$ is a LT-AC generator. Therefore, (B.15) can be rewritten as

$$\begin{aligned}C(u_1,\ldots,u_d)=&\int_0^\infty\prod_{j_1=1}^{D^{(1)}_{s_0}}\int_0^\infty\prod_{j_2=1}^{D^{(2)}_{s_1}}\cdots\int_0^\infty\prod_{j_{y+1}=1}^{D^{(y+1)}_{s_y}}\tilde{\psi}^{(y+1)}_{s_y j_{y+1}}\\&\times\left(\sum_{j_{y+1}=1}^{D^{(y+1)}_{s_y}}\psi^{(y+1)-1}_{s_y j_{y+1}}\left(C^{(y+2)}_{s_{y+1}j_{y+2}};v^{(y)}_{s_{y-1}j_y}\right)\right)(dG)^{(y)}_{j_y},\end{aligned} \tag{B.22}$$

which is equivalent to

$$\begin{aligned}&\int_0^\infty\prod_{j_1=1}^{D^{(1)}_{s_0}}\int_0^\infty\prod_{j_2=1}^{D^{(2)}_{s_1}}\cdots\int_0^\infty\prod_{j_{y+1}=1}^{D^{(y+1)}_{s_y}}\int_0^\infty\\&\quad\times\exp\left(-v^{(y+1)-1}_{s_y j_{y+1}}\sum_{j_{y+2}=1}^{D^{(y+2)}_{s_{y+1}}}\psi^{(y+1)-1}_{s_y j_{y+1}}(C^{(y+2)}_{s_{y+1}j_{y+2}})\right)\end{aligned} \tag{B.23}$$

$$d\widetilde{G}_{s_y j_{y+1}}(v^{(y+1)}_{s_y j_{y+1}};v^{(y)}_{s_{y-1}j_y})(dG)^{(y)}_{j_y}. \tag{B.24}$$

Let $F^{(y+1)}_{s_y j_{y+1}}$ satisfy

$$(F^{(y+1)}_{s_y j_{y+1}})^v=\exp\left(-v\psi^{(y+1)-1}_{s_y j_{y+1}}(u)\right). \tag{B.25}$$

Then we have

$$\begin{aligned}C(u_1,\ldots,u_d)=&\int_0^\infty\prod_{j_1=1}^{D^{(1)}_{s_0}}\int_0^\infty\prod_{j_2=1}^{D^{(2)}_{s_1}}\cdots\int_0^\infty\prod_{j_{y+1}=1}^{D^{(y+1)}_{s_y}}\int_0^\infty\prod_{j_{y+2}=1}^{D^{(y+2)}_{s_{y+1}}}F^{(y+1)}_{s_y j_{y+1}}\\&\times(C^{(y+2)}_{s_{y+1}j_{y+2}}))(dG)^{(y+1)}_{j_{y+1}},\end{aligned} \tag{B.26}$$

which means that (3) is satisfied at level $y+1$. Therefore, by mathematical induction, (3) holds for all $l=1,\ldots,L-1$, and this completes the proof of Theorem 2.1. □

References

Aas, K., Czado, C., Frigessi, A., Bakken, H., 2009. Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics 44 (2), 182–198.

Ang, A., Bekaert, G., 2002. International asset allocation with regime shifts. Review of Financial Studies 15 (4), 1137–1187.

Ang, A., Chen, J., 2002. Asymmetric correlations of equity portfolios. Journal of Financial Econometrics 63 (3), 443–494.

Barndorff-Nielsen, O.E., 1997. Normal inverse Gaussian distributions and stochastic volatility modelling. Scandinavian Journal of Statistics 24 (1), 1–13.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31 (3), 307–327.

Boubaker, H., Sghaier, N., 2013. Portfolio optimization in the presence of dependent financial returns with long memory: a copula based approach. Journal of Banking & Finance 37 (2), 361–377.

Brooks, C., Burke, P., Heravi, S., Persand, G., 2005. Autoregressive conditional kurtosis. Journal of Financial Econometrics 3 (3), 399–421.

Bu, R., Giet, L., Hadri, K., Lubrano, M., 2011. Modeling multivariate interest rates using time-varying copulas and reducible nonlinear stochastic differential equations. Journal of Financial Econometrics 9 (1), 198–236.

Choroś-Tomczyk, B., Härdle, W.K., Okhrin, O., 2013. Valuation of collateralized debt obligations with hierarchical Archimedean copulae. Journal of Empirical Finance 24, 42–62.

Christoffersen, P.F., 1998. Evaluating interval forecasts. International Economic Review 39 (4), 841–862.

Christoffersen, P.F., 2004. Backtesting value-at-risk: a duration-based approach. Journal of Financial Econometrics 2 (1), 84–108.

Das, S.R., Uppal, R., 2004. Systemic risk and international portfolio choice. The Journal of Finance 59 (6), 2809–2834.

Dias, A., Embrechts, P., 2004. Dynamic copula models for multivariate high-frequency data in finance. Manuscript, ETH Zurich.

Embrechts, P., Lindskog, F., McNeil, 2003. Modelling dependence with copulas and applications to risk management. Handbook of Heavy Tailed Distributions in Finance 8 (1), 329–384.

Embrechts, P., McNeil, A., Straumann, D., 2002. Correlation and dependence in risk management: properties and pitfalls. In: Dempster, M. (Ed.), Risk Management: Value at Risk and Beyond:. Cambridge University Press, pp. 176–223.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50 (4), 987–1007.

Feller, W., 2008. An Introduction to Probability Theory and its Applications. John Wiley & Sons Inc.

Föllmer, H., Schied, A., 2011. Stochastic Finance: An Introduction in Discrete Time, Second revised and extended edition. Walter de Gruyter, Berlin.

Goshtasby, A.A., 2012. Image Registration. Similarity and Dissimilarity Measures. Springer, London.

Hadar, J., Russell, W., 1969. Rules for ordering uncertain prospects. American Economic Review 59 (1), 25–34.

Harvey, C.R., Siddique, A., 1999. Autoregressive conditional skewness. Journal of Financial and Quantitative Analysis 34 (4), 465–487.

Hering, C., Hofert, M., Mai, J.-F., Scherer, M., 2010. Constructing hierarchical Archimedean copulas with Lévy subordinators. Journal of Multivariate Analysis 101 (6), 1428–1433.

Hoesli, M., Peka, K., 2013. Volatility spillovers, comovements and contagion in securitized real estate markets. The Journal of Real Estate Finance and Economics 47 (1), 1–35.

Hofert, M., 2008. Sampling Archimedean copulas. Computational Statistics and Data Analysis 52 (12), 5163–5174.

Hofert, M., 2012. A stochastic representation and sampling algorithm for nested Archimedean copulas. Journal of Statistical Computation and Simulation 82 (9), 1239–1255.

Joe, H., 1997. Multivariate Models and Multivariate Dependence Concepts. Taylor & Francis.

Jondeau, E., Rockinger, M., 2003. Conditional volatility, skewness and kurtosis: existence, persistence, and comovements. Journal of Economic Dynamics and Control 27 (10), 1699–1737.

Ward Jr., J.H., 1963. Hierarchical grouping to optimize an objective function. Journal of American Statistical Association 58 (301), 236–244.

Kimberling, C.H., 1974. A probabilistic interpretation of complete monotonicity. Aequationes Mathematicae 10 (2–3), 152–164.

Kole, E., Koedijk, K., Verbeek, M., 2007. Selecting copulas for risk management. Journal of Banking & Finance 31 (8), 2405–2423.

Kumar, M.S., Okimoto, T., 2011. Dynamics of international integration of government securities' markets. Journal of Banking & Finance 35 (1), 142–154.

Kurowicka, D., Cooke, R.M., 2006. Uncertainty Analysis with High Dimensional Dependence Modelling. John Wiley & Sons.

Longin, F.M., Solnik, B., 2001. Extreme correlation of international equity markets. The Journal of Finance 56 (2), 649–676.

Mai, J.-F., Scherer, M., 2012. H-extendible copulas. Journal of Multivariate Analysis 110 (0), 151–160, Special Issue on Copula Modeling and Dependence.

Marshall, A.W., Olkin, I., 1988. Families of multivariate distributions. Journal of the American Statistical Association 83 (403), 834–841.

McNeil, A.J., 2008. Sampling nested Archimedean copulas. Journal of Statistical Computation and Simulation 78 (6), 567–581.

McNeil, A.J., Frey, R., Embrechts, P., 2010. Quantitative Risk Management: Concepts, Techniques, and Tools. Princeton University Press.

Nelsen, R.B., 2006. An Introduction to Copulas, second ed. Springer.
Okhrin, O., Odening, M., Xu, W., 2013a. Systemic weather risk and crop insurance: the case of China. Journal of Risk and Insurance 80 (2), 351–372.
Okhrin, O., Okhrin, Y., Schmid, W., 2013b. On the structure and estimation of hierarchical Archimedean copulas. Journal of Econometrics 173 (2), 189–204.
Patton, A.J., 2006. Modelling asymmetric exchange rate dependence. International Economic Review 47 (2), 527–556.
Patton, A.J., 2009. Copula-based models for financial time series. In: Mikosch, T., Krei, J.-P., Davis, R.A., Andersen, T.G. (Eds.), Handbook of Financial Time Series. Springer, Berlin Heidelberg, pp. 767–785.
Savu, C., Trede, M., 2010. Hierarchies of Archimedean copulas. Quantitative Finance 10 (3), 295–304.
Schwarz, G.E., 1978. Estimating the dimension of a model. Annual of Statistics 6 (2), 461–464.
Siburg, K.F., Stoimenov, P., Weiß, G.N.F., 2015. Forecasting portfolio-value-at-risk with nonparametric lower tail dependence estimates. Journal of Banking & Finance 54, 129–140.
Székely, G.J., Rizzo, M.L., 2005. Hierarchical clustering via joint between-within distance: extending ward's minimum variance method. Journal of Classification 22 (2), 151–183.
Tankov, P., 2004. Financial Modelling with Jump Processes. CRC Press.
Weiß, G.N.F., Scheffer, M., 2015. Mixture pair-copula-constructions. Journal of Banking & Finance 54, 175–191.
Whelan, N., 2004. Sampling from Archimedean copulas. Quantitative Finance 4 (3), 339–352.
Zhang, W., Zhao, D., Wang, X., 2013. Agglomerative clustering via maximum incremental path integral. Pattern Recognition 46 (11), 3056–3065.
Zhou, J., Gao, Y., 2012. Tail dependence in international real estate securities markets. Journal of Real Estate Finance and Economics 45 (1), 128–151.