Copulas: A personal view

Paul Embrechts

Department of Mathematics

ETH Zurich, Switzerland

First version: November 2007

This version: June 15, 2009

Abstract

Copula modeling has taken the world of finance and insurance, and well beyond, by storm. Why is this? In this paper I review the early start of this development, discuss some important current research, mainly from an applications point of view, and comment on potential future developments. An alternative title of the paper would be "Demystifying the copula craze". The paper also contains what I would like to call the **copula must-reads**.

Keywords: copula, extreme value theory, Fréchet–Hoeffding bounds, quantitative risk management, Value–at–Risk

1 Introduction

When I was asked by the organizers of the conference on "New Forms of Risk Sharing and Risk Engineering" to deliver the Keynote Speech, I decided to talk about dependence modeling and extremes.s Both themes are crucial to the Alternative Risk Transfer (ART) landscape. As a consequence I have written this review paper with a main emphasis on dependence, in particular on copulas. However, after all the papers, books, conferences, talks, discussions, software, applications, ... what is there more to say on the topic? Going back to my early days on the subject, I recall fondly my excitement upon learning some of the caveats out there in dependence (correlation) modeling. I would like to share these early experiences and discuss potential developments. From the outset, I insist that this is a very personal view with *in no way* a complete literature overview; others have done that much better. I have therefore left out numerous references/topics I could/should have included. Let me just mention two interesting references: Patton [37] reviews the use of copulas in econometric modeling, Genest et al. [19] gives a nice bibliometric overview. Also, there are no figures to be found in "my personal view on copulas"; a fact that can easily be compensated by googling "copula".

For me personally, the copula story started around 1995 when on two separate occasions I was contacted by practitioners from banking and insurance with risk management questions, the essence of which was: "Here we are given a multi-dimensional (i.e. several risk factors) problem for which we have marginal information together with an idea of dependence. When is this question well–posed?". One concrete actuarial question from that period was: given two marginal, one–period risks X_1, X_2 with lognormal distribution functions (dfs) $F_1 = LN(0,1)$, $F_2 = LN(0,16)$. How can one simulate from such a model if X_1 and X_2 have linear correlation $\rho = 0.5$, say. Well, first of all, the correlation information says something about the (or a) *bivariate* df of the random vector $(X_1, X_2)^\top$, i.e. about $F(x_1, x_2) = P(X_1 \leq x_1, X_2 \leq x_2)$. Note however that, in the above, that information is *not* given; we *only* know F_1, F_2 and ρ. What else would one need?

To start on this question here is an easy calculation which most unfortunately quickly floors those end–users of probabilistic and statistical techniques with an insufficient basic first course on probability (which our American colleagues would refer to as Probability 101 and some of my more mathematical colleagues would refer to as "Surely you had this in Kindergarten!").

The (easy) copula argument

– First note that for random variables (rvs) X_i with continuous dfs F_i, $i = 1, 2$, $U_1 = F_1(X_1)$, $U_2 = F_2(X_2)$ both are uniformly distributed rvs on $[0, 1]$ *(check!)*.

– Hence for some joint df F with marginal dfs F_1, F_2,

$$\begin{aligned} F(x_1, x_2) &= P(X_1 \le x_1, X_2 \le x_2) \\ &= P(F_1(X_1) \le F_1(x_1), F_2(X_2) \le F_2(x_2)) \\ &= P(U_1 \le F_1(x_1), U_2 \le F_2(x_2)) \\ &\overset{(N)}{=} C(F_1(x_1), F_2(x_2)) , \end{aligned} \tag{1}$$

where $\overset{(N)}{=}$ stands for "is denoted by".

– The C above is exactly *a* (careful here!) *copula*, a df on $[0, 1]^2$ with standard uniform marginals, C is the df of the random vector $(U_1, U_2)^\top$.

A remark is in order here: there is absolutely no real, compelling mathematical reason for transforming the marginal dfs of F to uniform dfs on $[0, 1]$, though it may be useful from a statistical point of view. In his 1940 papers, Hoeffding used the interval $[-1/2, 1/2]$. In multivariate Extreme Value Theory, it is standard to transform to unit Fréchet marginal dfs. In this context, Resnick [38] p. 265 writes "How one standardizes is somewhat arbitrary and depends on taste. Different specifications have led to (superficially) different representations in the literature." See also de Haan and Ferreira [9] p. 210. Finally, Klüppelberg and Resnick [29] present a problem from the realm of copulas where Pareto marginalization is useful; we will briefly come back to this in our concluding Section 5. For the moment it suffices to know that (i) transforming marginal dfs in multivariate statistical data to some standard df is often useful, (ii) in the copula case, one standardizes to a uniform df on $[0, 1]$, and (iii) many other transformations are possible, and relevant. The choice is often context/application dependent.

If we now return to our lognormal example, we see no immediate reason how the number ρ

should determine a/the function C, it is not even clear whether the problem has none, one or infinitely many solutions. In this case, it turns out that the problem posed has *no* solution.

Formula (1) *couples* the continuous marginal dfs F_1, F_2 to the joint df F via *a/the* copula C. Now it is time to concentrate on "a/the" issue and read/interpret (1) correctly. This is done through Sklar's Theorem, a result which has become rather well known in Quantitative Risk Management (QRM); see McNeil et al. [32], Theorem 5.3. Here is *a* version:

Theorem 1 (Sklar, 1959; the easy version) *Suppose $X_1, \ldots, X_d$ are rvs with continuous dfs $F_1, \ldots, F_d$ and joint df F, then there exists a* unique *copula C (a df on $[0,1]^d$ with uniform marginals) such that for all $\mathbf{x} = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$:*

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)) \ . \tag{2}$$

Conversely, given any *dfs $F_1, \ldots, F_d$ and copula C, F* defined *through (2) is a d–variate df with marginal dfs $F_1, \ldots, F_d$.* $\square$

One can also rewrite (2) for $\mathbf{u} = (u_1, \ldots, u_d)^\top \in [0,1]^d$ as

$$C(u_1, \ldots, u_d) = F\left(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\right) \ . \tag{3}$$

Here F_i^{-1} stands for the inverse of F_i properly defined, it is precisely at this point where the continuity of F_i comes in handy and problems start when F_i has jumps; see McNeil et al. [32], Appendix A.1.2, for details. Using formula (2) we can now construct in two stages a joint df F to our taste. Start with the marginal dfs $F_1, \ldots, F_d$ and "add" to these a copula C of your choice: 99.9% of the applications of copula technology use this completely straightforward formula (do not take the 0.1% too literally!). Stage 1 fixes the marginal dfs $F_1, \ldots, F_d$, whereas stage 2 allows the coupling of the marginals with a predescribed interdependence through the copula C. Reading formula (3) from right to left yields the construction of the copula C from any joint df with continuous marginal dfs $F_1, \ldots, F_d$, a recipe not unlike recipes you find in real cookbooks. One does not immediately encounter the lobster– or châteaubriand-copula, but Archimedean–, Gauss–, Maltesian–, t–, hyperbolic–, zebra– and elliptical copulas for instance do appear as recipes.

I realize that I did not answer several remaining questions:

- What about the "a/the" issue more in detail: surely discrete marginal dfs $F_1, \ldots, F_d$ *must* cause problems. Believe me, they do! The *full version* of Sklar's Theorem makes this clear. See Genest and Nešlehová [22] for an excellent primer on this issue and be prepared that everything that *can* go wrong, *will* go wrong. This paper is my **first must–read**, though Genest et al. [23] comes very close.
- Are there ways for constructing copulas (away from the $F, F_1, \ldots, F_d$ construction in (3))? Well, that's "easy". Write down a candidate function from $[0,1]^d$ to $[0,1]$ and "just" check that the marginals are standard uniform (usually a trivial task) and that C is indeed a d–variate df: this is often the hard bit and the reason for quotation marks around "easy" and "just". See the standard monographs Joe [25] and Nelsen [36] for all you want to learn on this, and much more! McNeil et al. [32] contains an introduction to the realm of copulas aimed at the quantitative risk manager. If you have mastered the basic theory above, you may venture out into the exciting land of copula–exotics; an interesting paper for instance giving you a guided walk through the copula country ruled by Archimedes is McNeil and Nešlehová [33]. This paper not only gives you ways to construct copulas with shapes you are likely to meet in your wildest dreams, the authors also show how beautiful and surprising some of the underlying mathematics can be.

An absolutely crucial next question now is:

- If we use the two–stage copula modeling approach (2) towards the construction of a multivariate model F, how would one fit such a model to data? One general idea is obvious. Start from a d–variate sample $\mathbf{X}_1, \ldots, \mathbf{X}_n$ from which d estimators $\widehat{F}_{1,n}, \ldots, \widehat{F}_{d,n}$ of the marginal dfs can be obtained. Use these $\widehat{F}_{i,n}$ to transform the $\mathbf{X}$–sample to $\mathbf{U}_1, \ldots, \mathbf{U}_n$ on $[0,1]^d$ and fit your favorite copula model. At any stage of this procedure you can use whatever technique you learned from statistics: parametric, non–parametric, semi–parametric or Bayesian. Of course, if relevant, one can use a full

Maximum Likelihood approach; see for instance Kim et al. [28]. This paper illustrates through simulations the clear advantages connected with the use of rank–based methods for the estimation of dependence parameters, and especially the superiority of the maximum pseudo–likelihood technique originally studied in Genest et al. [20].

However, fitting a (some) model(s) to data is one thing, showing that you have a good model reflecting the basic data characteristics is something different. Again, 99.9% of statistical estimation in copula applications to risk management concentrates on the former and largely neglects the latter (the goodness–of–fit step). The best advice I can give the reader at this point is to Google Christian Genest and find out what he and his co–workers have written on the subject. From this source you will find numerous papers with statistical theory and applications to several problems many of which nowadays stem from finance and insurance. For me, our **second must–read** is Genest and Favre [18].

Later on in this paper, I will continue with some further copula related questions you may want to look at. But now, it is time for

2 Another theorem

Recall that in Section 1, I already stated the result that there is no bivariate model with $LN(0,1)$, $LN(0,16)$ marginals and correlation $\rho = 0.5$. Why? Well, the (skewed) marginal dfs constrain the range of all possible correlations that a joint model with the given LN–marginals can have. The basic result starting the analysis goes back a very long time, to the 1940's.

Theorem 2 (Fréchet–Hoeffding bounds) *Suppose $F_1, \ldots, F_d$ are marginal dfs and F is any joint df with those given marginals, then for all $\mathbf{x} \in \mathbb{R}^d$,*

$$\left(\sum_{k=1}^{d} F_k\left(x_k\right) + 1 - d \right)^{+} \leq F(\mathbf{x}) \leq \min\left(F_1\left(x_1\right), \ldots, F_d\left(x_d\right)\right) . \tag{4}$$

$\square$

The right–hand side of (4) is always a df (the so-called *comonotonic* model), whereas the left–hand side is only a df for $d = 2$, in which case this model is referred to as *countermonotonicity*. As an important conclusion we learn that the marginal dfs constrain the possible joint dfs built on them (via (2)) as given in (4). This theorem can now be *translated* into the language of correlations and yields a result very useful for ART modeling:

Theorem 3 (Correlation bounds) *Suppose F_1, F_2 are non-degenerate dfs. Then for any bivariate model F with F_1, F_2 as marginal dfs, the corresponding linear correlation coefficient ρ_F satisfies:*

$$-1 \le \rho_{\min} \le \rho_F \le \rho_{\max} \le +1\,,$$

where all values in the closed interval $[\rho_{\min}, \rho_{\max}]$ can be achieved. One always has that $\rho_{\min} < 0$ and $\rho_{\max} > 0$ but it is possible that $\rho_{\min} > -1$ and/or $\rho_{\max} < +1$; $\rho_{\min}$ (respectively $\rho_{\max}$) corresponds to counter– (respectively co–)monotonicity. $\square$

With respect to the last statement, one can for instance show that for non-degenerate, positive rvs always $\rho_{\min} > -1$; see Property 5.2.7 of Denuit et al. [10]. Now we can return to the LN–example from Section 1 and compute for $F_1 = LN(0,1)$, $F_2 = LN\,(0,\sigma^2)$,

$$\rho_{\min} = \frac{e^{-\sigma} - 1}{\sqrt{(e-1)\,(e^{\sigma^2}-1)}}\,,$$

and

$$\rho_{\max} = \frac{e^{\sigma} - 1}{\sqrt{(e-1)\,(e^{\sigma^2}-1)}}\,.$$

When $\sigma = 4$, this becomes the interval $[-0.00025, 0.01372]$. The value $\rho = 0.5$ lies well outside this feasibility region; $\rho = 0.5$ can only be achieved for $\sigma \le 2.28$. I leave the obvious risk management consequences of the small absolute values of $\rho_{\min}$, $\rho_{\max}$ to the interested reader. The quantitative risk manager or ART modeler may wonder how to digest a (fairly standard LN) model with correlation $\rho = 0.01372$ and yet very strongly (even comonotonically) dependent marginals. As a consequence, be careful with the myth "low correlation (low β in the Capital Asset Pricing Model (CAPM) language), low dependence, diversification effect, ...". If necessary, pedagogical and technical support can be found in Embrechts et al. [13] or McNeil et al. [32], Figure 5.8. The former publication was one of the early papers that warned the QRM community for the imprudent use of linear correlation by discussing several correlation fallacies (it existed as an ETH Zurich RiskLab report by the end of 1998), the LN–case above being one of them. Though being somewhat biased, I would like to label this paper as the **third must–read**.

3 Some of the key copula models used

As we have seen so far (i.e. (2)), the notion of copula is both natural as well as easy for looking at multivariate dfs. But why do we witness such an incredible growth in papers published starting the end of the nineties (recall, the concept goes back to the fifties and even earlier, but not under that name). Here I can give three reasons: finance, finance, finance. In the eighties and nineties we experienced an explosive development of quantitative risk management methodology within finance and insurance, a lot of which was driven by either new regulatory guidelines or the development of new products; see for instance Chapter 1 in McNeil et al. [32] for the full story on the former. Two papers more than any others "put the fire to the fuse": the already mentioned 1998 RiskLab report Embrechts et al. [13] and at around the same time, the Li credit portfolio model Li [30]. Let me briefly touch upon the latter.

First recall the standard credit risk model for the pricing of corporate debt. In a simplified form, consider d companies with value processes $V_1, \ldots, V_d$ and respective debt $D_1, \ldots, D_d$. Company i defaults by the end of the accounting year if $V_i(1) < D_i(1)$, $i = 1, \ldots, d$. Hence the individual one–year default probabilities equal

$$P\left(V_1(1) \leq D_1(1)\right), \ldots, P\left(V_d(1) \leq D_d(1)\right) .$$

If one now assumes that, after a logarithmic transformation, $\left(V_1(t), \ldots, V_d(t)\right)_{t\leq 1}$ follows a d–

dimensional Brownian motion (hence look at V_i as log–value), the joint probability of default for instance corresponds to

$$P\left(V_1(1) \leq D_1(1), \ldots, V_d(1) \leq D_d(1)\right) , \tag{5}$$

which results in the calculation of the joint probability under a multivariate normal model. This model is however known for producing too low probabilities of joint defaults. One way to improve this defect would be to look at the d *marginal* normal (1–dimensional Brownian) models, but then put a copula on these which allows for more tail dependence, like a t–copula. The latter model is referred to as a meta–t model stressing the fact that it is mainly the copula that determines ultimate joint default rather than the marginal dfs. Li [30] takes a different approach by assuming that the time–to–default of a company can be modeled by an exponential rv. Hence our credit portfolio now consists of the random vector $(E_1, \ldots, E_d)^\top$ where $E_i \sim \mathrm{Exp}\left(\lambda_i\right)$, $i = 1, \ldots, d$, and equivalent to (5), one has to calculate

$$P\left(E_1 \leq 1, \ldots, E_d \leq 1\right) ,$$

or more generally, the joint distribution

$$P\left(E_1 \leq x_1, \ldots, E_d \leq x_d\right) ;$$

here the 1 corresponds to the one–year default horizon.

Now, in contrast to the multivariate normal or t, there does *not* exist a *standard* d–dimensional exponential df. So in order to make $E_1, \ldots, E_d$ dependent, while sticking to the exponential marginals, one can put any copula C on $(E_1, \ldots, E_d)$. Li [30] proposes for C a Gaussian copula which is relatively easy to calibrate to credit risk data. This meta–Gaussian model became enormously popular, in the end causing problems because the market too strongly believed in it; see Whitehouse [40]. One of my probability friends, at the height of the copula craze to credit risk pricing, told me that “The Gauss–copula is the worst invention ever for credit risk management.” Currently various more realistic (less unrealistic) versions have been worked out replacing for instance the Gauss–copula by a finite mixture of Gaussians or even an infinite mixture leading to elliptical copulas. Unfortunately, most of these models are inherently static and fail to incorporate the dynamics of markets, especially in periods of distress. The 2007 subprime crisis around the pricing of Collateralized Debt Obligations (CDOs) is a very clear proof of this. For an introduction to CDO–pricing, see McNeil et al.[32] or Bluhm and Overbeck [5]. As always, applied modelers should handle new technology with sufficient caution. In particular, finance seems to be prone to methodological herding.

Other examples where copulas enter naturally in the realm of QRM concern aggregation of risk measures within a given risk class, operational risk, say, or between risk classes (market, credit, operational, underwriting, ...). Some relevant papers to get started here are Embrechts and Puccetti [15], Degen et al. [8] for the former and Aas et al. [1] for the latter. The above papers discuss problems of the following type. Let $X_1 \dots, X_d$ be d one–period risks with marginal dfs $F_1, \dots, F_d$. For some function $\Psi : \mathbb{R}^d \to \mathbb{R}$, $\Psi(X_1, \dots, X_d)$ denotes a risk position. Given now some risk measure $\mathcal{R}$, one wants to calculate $\mathcal{R}(\Psi(X_1, \dots, X_d))$. For the latter one typically needs the joint df F of $(X_1, \dots, X_d)$. We are again in the situation of the Fréchet–Hoeffding bounds of Section 2. If no further information is available, one could try to calculate lower and upper bounds

$$\mathcal{R}_L \leq \mathcal{R}(\Psi(X_1, \dots, X_d)) \leq \mathcal{R}_U$$

which are conform with the available statistical model assumptions on $(X_1, \dots, X_d)$. For the fully unconstrained case we only know $F_1, \dots, F_d$. Copula theory may enter at the level of coding dependence information. For example, one could be interested in the sum–case $\Psi(x_1, \dots, x_d) = \sum_{i=1}^d x_i$, $\mathcal{R} = \mathrm{VaR}_\alpha$, for some $\alpha \in (0, 1)$, and require that the copula $C \geq C_I$, where $C_I(u_1, \dots, u_d) = u_1 \cdots u_d$ denotes the independence copula. Here VaR_α stands for the Value–at–Risk at confidence α and mathematically corresponds to the α–quantile of an underlying df; see McNeil et al. [32] or Jorion [26]. The resulting optimization problems, using (2), reduce to the calculation of

$$\mathrm{VaR}_{\alpha,L} = \inf\left\{\mathrm{VaR}_\alpha\left(\sum_{i=1}^d X_i\right), \quad X_i \sim F_i, \quad i = 1, \dots, d, \quad C \geq C_I\right\}$$

and

$$\mathrm{VaR}_{\alpha,U} = \sup\left\{\mathrm{VaR}_\alpha\left(\sum_{i=1}^d X_i\right), \quad X_i \sim F_i, \quad i = 1, \dots, d, \quad C \geq C_I\right\}.$$

Beyond the numerical calculation of such bounds, one is further interested in proving that these bounds are sharp. For $d > 2$, there are still several open problems; see for instance Embrechts and Puccetti [15]. I would like to add at this point that conditions like $C \geq C^*$, for some copula C^*, may not be most natural from an applied modeling point of view. More relevant notions of risk ordering are, for instance, to be found in Müller and Stoyan [35] and Denuit et al. [10].

So far we have looked at a random vector $\mathbf{X} = (X_1, \ldots, X_d)^\top$ with known one–dimensional marginal dfs $F_1, \ldots, F_d$. A natural extension could start with a multi–dimensional subvector decomposition of $\mathbf{X}$,

$$\mathbf{X} = \left(X_1, \ldots, X_{d_1}, X_{d_1+1}, \ldots, X_{d_1+d_2}, \ldots, X_{d_1+\cdots+d_{l-1}+1}, \ldots, X_{d_1+\cdots+d_l}\right)^\top$$

where $d = \sum_{i=1}^{l} d_i$ and the marginal subvector dfs of $\left(X_{d_1+\cdots+d_i+1}, \ldots, X_{d_1+\cdots+d_{i+1}}\right)$ are known. A typical example from the realm of finance corresponds to a global equity index for $\mathbf{X}$, say, and the subvectors correspond to more or less homogeneous subgroups such as banking, insurance, chemical, pharmaceutical, manufacturing, ... A further extension concerns looking at overlapping marginal subvector models. For some references on this, check for instance the notion of grouped t–copula in McNeil et al. [32], the QRM examples in Embrechts and Puccetti [16] and the early theory in Rüschendorf [39]. An application of this theory to operational risk modeling is given in Embrechts and Puccetti [17].

4 Which copula to use

One of the questions most commonly asked by practitioners or beginning young researchers to the field is "Which copula should we use for a certain problem?". This question quickly goes to the core of the criticism voiced by some on the copula craze. By definition, or indeed by Sklar's Theorem, a copula model for a multivariate df is two–stage,

$$F = C\left(F_1, \ldots, F_d\right) .$$

Certain multivariate models have a natural and legitimate reason for existence, so for instance the multivariate Gaussian $F = N_d\left(\boldsymbol{\mu}, \sum\right)$ which appears through the multivariate Central Limit Theorem. Likewise, elliptical dfs appear as scale–mixtures of multivariate Gaussians via

$$\mathbf{X} = \boldsymbol{\mu} + \sqrt{W}\mathbf{Z}\,,$$

where $\mathbf{Z} \sim N_d\left(\mathbf{0}, \sum\right)$, $\boldsymbol{\mu} \in \mathbb{R}^d$ and $W \geq 0$ is a scalar rv W, independent of $\mathbf{Z}$. For example, if $W^{-1} \sim \chi^2_\nu/\nu$, then $\mathbf{X} \sim t_d\left(\nu, \boldsymbol{\mu}, \sum\right)$, the d–dimensional t–distribution. One does not need copula theory to understand these models, though it helps to know that the t–copula always yields extremal clustering, whereas the Gauss–copula does not. I would like to stress that the above makes the multivariate normal– and t–distribution natural, not necessarily their corresponding copulas. Copulas also enter naturally in describing the dependence between multivariate extremes. Whereas univariate Extreme Value Theory (EVT) has a canonical definition of largest (smallest) value, the Multivariate case (MEVT) allows for several approaches. By far the most popular one is the coordinatewise one. Suppose $\mathbf{X}_1, \ldots, \mathbf{X}_n$ is a sample of iid d–dimensional random vectors. Denote the marginal maxima by

$$M_{in} = \max\left\{X_{ij}, \quad j = 1, \ldots, n\right\}\,, \quad i = 1, \ldots, d\,.$$

Suppose now that for some affine transformation $\alpha_{in}(x) = \frac{x - c_{in}}{d_{in}}$, $c_{in} \in \mathbb{R}$, $d_{in} > 0$, $\alpha_{in}\left(M_{in}\right)$ converges in distribution to some non–degenerate df H_i, then H_i is a generalized extreme value distribution (of the Fréchet, Gumbel or Weibull type). Hence if the vector of normalized maxima $\left(\alpha_{in}\left(M_{in}\right), \ldots, \alpha_{dn}\left(M_{dn}\right)\right)^\top$ converges in distribution, then the vector of marginal limit dfs equals $\left(H_1, \ldots, H_d\right)^\top$, where $H_1, \ldots, H_d$ are known through EVT; see Embrechts et al. [12] for details. Hence in order to fully characterize the joint df H of the limit random vector, one needs to know the copulas C which can appear in this way; they are referred to as *extreme value copulas*, C_{EV}. One can now readily characterize all C_{EV}s and work out the resulting MEVT based on them; see McNeil et al. [32] for details. A more geometric approach to MEVT is given in Balkema and Embrechts [3]; in that theory, copulas only appear marginally.

Besides the elliptical– and extreme value copulas, there are other more specialized copula models which appear in a natural way as the solution of some stochastic scheme (weak convergence limit, closure property, ...). One such example is the Clayton copula, which for $d = 2$ takes the form

$$C_\theta^{\mathrm{Cl}}(u_1, u_2) = \exp\left\{-\left((-\ln u_1)^\theta + (-\ln u_2)^\theta\right)^{1/\theta}\right\}, \quad 1 \leq \theta < \infty .$$

This copula appears as a natural dependence model for successive conditioning in upper semi–infinite rectangles. See Juri and Wüthrich [27] for the original paper on the subject, Alink [2] for several extensions and Balkema and Qi [4] for the ultimate uniqueness result as a closure property. Other interesting classes enter through special stochastic modeling, like the Marshall–Olkin copula, see Lindskog and McNeil [31], or copulas related to frailty models in (medical) survival analysis; see Hougaard [24] and Duchateau and Janssen [11] for the latter, where also the link to the class of Archimedean copulas is explained. For applications of frailty models to credit risk, see for instance McNeil et al. [32].

In summary, the question "which copula to use?" has no obvious answer. There definitely are many problems out there for which copula modeling is very useful. Within the world of copulas there are more and less natural ones, just like in the case of general joint dfs. Copula theory does not yield a magic trick to pull *the* model out of a hat. This situation led to some heated discussions between those in favor and those against copula thinking. Without taking sides in this discussion, I can recommend the interested critical reader to go through the arguments oneself; see Mikosch [34] for the paper, followed by a somewhat lively discussion and rejoinder. As so often, the "truth" lies somewhere in the middle: copulas form a most useful concept for a lot of applied modeling, they do not yield, however, a panacea for the construction of useful and well–understood multivariate dfs, and much less for multivariate stochastic processes. But none of the copula experts makes these claims. No doubt, copulas form an interesting class of probability measures which can be studied just from the point of view of Platonic beauty and interest and that is what overall (pure) mathematics is often about; see however Davies [7] on this more philosophical issue.

As a final remark: recent applications of copula techniques often lack historical perspective. In the spirit of the well–known quote "standing on the shoulders of giants" by Bernard de Chartres (12th Century) and made famous by Isaac Newton, researchers should always be aware where these "new" techniques originate, where their historical cradle stood. The edited volume Dall'Aglio et al. [6] (note the subtitle "Beyond the Copulas") contains several interesting papers on this. Notably, the contributions by Giorgio Dall'Aglio, Ludger Rüschendorf and Berthold Schweizer are still well worth reading.

5 What next

This brings us to the end of my somewhat personal and admittedly too brief copula journey. Besides consulting the basic texts Joe [25] and Nelsen [36], I urge you to go through the **must–reads** in full detail: in chronological order, **Embrechts et al. [13]**, **Genest and Favre [18]** and **Genest and Nešlehová [22]**. Make up your own mind on the critical Mikosch [34] including the discussions, and experiment/contribute yourself. The theory allows not only for good mathematics, but also for some fun. For the latter, see for instance Genest and MacKay [21], or you may want to listen to the DVD publication Embrechts and Nešlehová [14] on copulas and extreme value theory. And, as a final note: Thomas Mikosch compared the copula craze with Hans Christian Andersen's fairy tale "The Emperor's new clothes" where the child says "But he hasn't got anything on!". In a recent publication, Klüppelberg and Resnick [29], the authors end with "Religious Copularians have unshakable faith in the value of transforming a multivariate distribution to its copula. For the sceptics who believe the Emperor wears no clothes (Mikosch [34]), perhaps use of the Pareto copula convinces some of these the Emperor at least wears socks." It is my personal belief that over the years to come, research will be able to put further garments on the poor man so that eventually in Hans Christian Andersen's words we can truly say "Goodness! How well they suit your Majesty! What a wonderful fit! What a cut! What colors! What sumptuous robes!".

Acknowledgment: I would like to thank Christian Genest, Johanna Nešlehová and two anonymous referees for some useful comments on an earlier version of this paper.

References

[1] Aas, K., Dimakos, X.K. and Øksendal, A. (2007) Risk capital aggregation. *Risk Management* **9**, 82–107.

[2] Alink, S. (2007) *Copulas and Extreme Values*. PhD thesis, University of Nijmegen.

[3] Balkema, A.A. (Guus) and Embrechts, P. (2007) *High Risk Scenarios and Extremes. A Geometric Approach*. Zurich Lectures in Advanced Mathematics, European Mathematical Society Publishing House.

[4] Balkema, A.A. and Qi, Y.–C. (1998) Strictly stable laws for multivariate residual lifetimes. *Journal of Mathematical Sciences* **92**, 3873–3880.

[5] Bluhm, C. and Overbeck, L. (2007) *Structured Credit Portfolio Analysis, Baskets & CDOs*. Chapman & Hall, Boca Raton.

[6] Dall'Aglio, G. Kotz, S. and Salinetti, G. (Eds.) (1991) *Advances in Probability Distributions with Given Marginals. Beyond the Copulas.* Kluwer Academic Publishers, Dordrecht.

[7] Davies, E.B. (2007) Let Platonism die. *Newsletter of the European Mathematical Society* **64**, 24–25.

[8] Degen, M., Embrechts, P. and Lambrigger, D. (2007) The quantitative modelling of operational risk: between g–and–h and EVT. *ASTIN Bulletin* **37**, 265–291.

[9] de Haan, L. and Ferreira, A. (2006) *Extreme Value Theory. An Introduction.* Springer, Berlin.

[10] Denuit, M., Dhaene, J., Goovaerts, M. and Kaas, R. (2005) *Actuarial Theory for Dependent Risks. Measures, Orders and Models.* Wiley, Chichester.

[11] Duchateau, L. and Janssen, P. (2008) *The Frailty Model.* Springer, New York.

[12] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[13] Embrechts, P., McNeil, A.J. and Straumann, D. (2002) Correlation and dependency in risk management: properties and pitfalls. In *Risk Management: Value at Risk and Beyond* (ed. M. Dempster), pp. 176–223. Cambridge University Press, Cambridge.

[14] Embrechts, P. and Nešlehová, J. (2006) Copulas and Extreme Value Theory, in *Quantitative Financial Risk Management: Fundamentals, Models and Techniques.* Chapter IV: Applications to Credit Risk and Market Risk. DVD, Henry Stewart Talks.

[15] Embrechts, P. and Puccetti, G. (2006) Bounds for functions of dependent risks. *Finance and Stochastics* **10**, 341–352.

[16] Embrechts, P. and Puccetti, G. (2006) Bounds for functions of multivariate risks. *Journal of Multivariate Analysis* **27**, 526–547.

[17] Embrechts, P. and Puccetti, G. (2006) Aggregating risk capital, with an application to operational risk. *The Geneva Risk and Insurance Review* **31**, 71–90.

[18] Genest, C. and Favre, A.–C. (2007) Everything you always wanted to know about copula modeling but were afraid to ask. *Journal of Hydrologic Engineering* **12**, 347–368.

[19] Genest, C., Gendron, M. and Bourdeau–Brien, M. (2009) The advent of copulas in finance. *European J. Finance* **15**, to appear.

[20] Genest, C., Ghoudi, K. and Rivest, L.-P. (1995) A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* **82**, 543–552.

[21] Genest, C. and MacKay, R.J. (1986) The joy of copulas: bivariate distributions with uniform marginals. *The American Statistician* **40**, 280–283.

[22] Genest, C. and Nešlehová, J. (2007) A primer on copulas for count data. *ASTIN Bulletin* **37**, 475–515.

[23] Genest, C., Rémillard, B. and Beaudoin, D. (2009) Goodness–of–fit tests for copulas: a review and a power study. *Insurance: Mathematics and Economics* **44**, 199–213.

[24] Hougaard, P. (2001) *Analysis of Multivariate Survival Data.* Springer, New York.

[25] Joe, H. (1997) *Multivariate Models and Dependence Concepts.* Chapman & Hall, London.

[26] Jorion, P. (2007) *Value at Risk.* (Third Ed.) McGraw Hill, New York.

[27] Juri, A. and Wüthrich, M.V. (2002) Copula convergence theorems for tail events. *Insurance: Mathematics and Economics* **30**, 405–420.

[28] Kim, G., Silvapulle, M.J. and Silvapulle, P. (2007) Comparisons of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis* **51**, 2836–2850.

[29] Klüppelberg, C. and Resnick, S.I. (2009) The Pareto copula, aggregation of risk and the Emperor's socks. Preprint, Technical University of Munich.

[30] Li, D. (2001) On default correlation: a copula function approach. *Journal of Fixed Income* **9**, 43–54.

[31] Lindskog, F. and McNeil, A.J. (2003) Common Poisson shock models: applications to insurance and credit risk modelling. *ASTIN Bulletin* **33**, 209–238.

[32] McNeil, A.J., Frey, R. and Embrechts, P. (2005) *Quantitative Risk Management: Concepts, Techniques, Tools.* Princeton University Press, Princeton.

[33] McNeil, A.J. and Nešlehová, J. (2009) Multivariate Archimedean copulas, d–monotone functions and l_1–norm symmetric distributions. *Annals of Statistics* **37**, to appear.

[34] Mikosch, T. (2006) Copulas: tales and facts. *Extremes* **9**, 3–20. Discussion, 21–53. Rejoinder 55–62.

[35] Müller, A. and Stoyan, D. (2002) *Comparison Methods for Stochastic Models and Risks.* Wiley, Chichester.

[36] Nelsen, R.B. (2006) *An Introduction to Copulas.* Second Edition. Springer, New York.

[37] Patton, A. (2009) Copula–based models for financial time series. In Andersen, T.G., Davies, R.A., Kreiss, J.–P. and Mikosch, T. (Eds.) (2009) *Handbook of Financial Time Series*, pp. 767–785, Springer, Berlin.

[38] Resnick, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[39] Rüschendorf, L. (1985) Construction of multivariate distributions with given marginals. *Annals of the Institute of Statistical Mathematics* **37**, 225–233.

[40] Whitehouse, M. (2005) How a formula ignited market that burned some big investors. *The Wall Street Journal*, September 12, 2005.