

Double t Copula Pricing of Structured Credit Products

Practical aspects of a trustworthy implementation

Frédéric D. Vrins*

This Version : February 2009

Abstract

In spite of its simplicity, the popular One Factor Gaussian Copula model remains the market standard for the valuation of CDO tranches and n -th to default. It suffers however from well-know weaknesses, mainly due to the tail behavior of the Normal distribution (namely : the tails are too light, and there is no tail dependence, whatever is the copula correlation). Alternative models have been proposed, among those is the double t copula, which does not share the Gaussian copula drawbacks while not being much more complex. In spite of its nice features, this framework suffers from some technical problems related to its implementation. Without a lot of care, this technique could easily lead to inconsistent results. In our opinion, these difficulties have prevent practitioners to really turn that theoretically sounding model into a workable pricing tool. This paper aims at filling this gap by giving routes toward a reliable implementation of the double t copula framework, throwing away the drawbacks of this framework compared to the Gaussian one.

The first purpose of this letter is to show that the implementation issues related to the double t model actually reduce to the estimation of integrals with respect to some student-t distributions. The second part of this note presents an efficient numerical method to perform this tedious task.

1 Introduction

Standard model for the pricing of multiple-names structured credit products is still the so-called *One Factor Gaussian Copula*. This model assumes that all the dependency between the constituents can be encapsulated into a single linear model involving Gaussian variables.

The popularity of this model results from its simplicity and tractability. However, some theoretical analysis have pointed out some key drawbacks of this model. First it strongly fails to replicate the market when one single correlation is used across the whole structure. Second, the choice of Gaussian variables, although having no consequence on the default times marginal distributions, do not have a fat-tail behavior, which is a desired property when assessing products related to extreme events, like credit events. In large dimensions, i.e. when the number of constituents becomes large, this problem becomes even more critical. Finally, in addition to the *too light tail* problem, we have a dependency problem. Whatever is the correlation plugged in the latent variable model, tail (that is extreme) events occur as if they were independent. Correlating such extreme events is at the limit impossible to do with a Gaussian mixing scheme. This implies in practice to push the model correlation very high in order to introduce dependency between these rare events. We believe this insensitivity of tail dependence to model correlation is the reason why the Gaussian copula model has problems to produce index CDO base correlations for the most senior tranches. Deeper discussion regarding this topic can be found in [1, 2, 3].

Researchers looked at other coupling schemes than the Gaussian copula, not suffering from the above drawbacks. Among the wide variety of proposed coupling functions, the double t copula received a lot of

*Dr F. Vrins is with the Financial Markets/Credit Derivatives Modeling Team of ING SWE. Contact: +32 2 557 13 13, frederic.vrins@ing.be, Aue Marnix 24, M2+4, 1050 Brussels, Belgium. The author is grateful to John Hull for discussions about existing implementations of the double t model and regarding the feedback provided about the present paper.

attention. Instead of combining Gaussian variables, it is based on a Student's t variables mixture. We refer to [1, 3, 4, 5] for more details on this model and its advantages. Although the purpose of this paper is not to discuss on the advantages of the double t from a model point of view, we illustrate the tail-dependency problem with simple charts for the reader who may not be familiar with this theoretical concept (see Figure 1).

Although this model fixes the light-tail and tail-dependence problems, it raises other issues, making it's practical use difficult. For instance, it requires the estimation of a double t quantile function (which is dependent upon the correlation level) which has indeed no known analytical form. Furthermore, it also requires the finding of an efficient technique for performing integration with respect to a Student's t distribution, which is not standard. Surprisingly, this is not discussed in the literature. This note aims at filling this gap.

Given the advantages of the double t copula, this paper provides the reader with a methodology to perform an efficient implementation of the copula. For instance, it is shown that tabulation of quantiles can be avoided. The first part of this paper is to reduce the technical difficulties of the double t implementation to the unique problem of integral computation of one specific form. Dealing with this specific integration is the topic of the second part of the paper. This last issue is shown to be indeed very critical, justifying the use of a dedicated quadrature methods.

2 Copula pricing framework

Pricing a structured credit product with N underlyings requires the computation of its expected present value, $\mathbb{E}[\text{Val}]$. By the law of iterated expectations, the conditional expectation is an unbiased estimator so, given the N -dimensional set of random default times $\vec{\tau} = [\tau_1, \dots, \tau_N]$, one gets

$$\mathbb{E}[\text{Val}] = \mathbb{E}[\mathbb{E}[\text{Val}|\vec{\tau}]] = \int_{t_1} \cdot \int_{t_N} \mathbb{E}[\text{Val}|\vec{\tau}] dF_{\vec{\tau}}(t_1, \cdot, t_N) ,$$

where $F_{\vec{\tau}}(t_1, \cdot, t_N)$ is the N -variate cumulative probability distribution of the default time vector $\vec{\tau}$. Defining $\sigma(\vec{\tau}) \doteq \bigvee_{i=1}^N \sigma(\tau_i)$, it is worth noting that Val is $\sigma(\vec{\tau})$ -measurable¹, so that we can define the deterministic function $\text{Val}(\vec{\tau}) \doteq \mathbb{E}[\text{Val}|\vec{\tau}]$.

In a "one factor" copula framework, default times are made dependent through a latent variable model,

$$Z_i = \rho Z + \sqrt{1 - \rho^2} \tilde{Z}_i ,$$

where Z is the common variable (describing the state of the economy) and the \tilde{Z}_i are i.i.d. random variables independent of Z . In this context, the i -th obligor has cumulative default probability

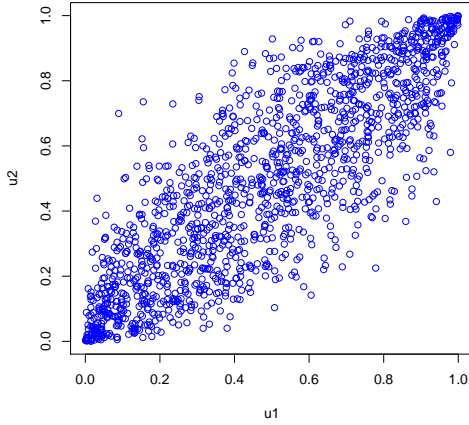
$$F_i(t) \doteq \mathbb{P}[\tau_i \leq t] = \mathbb{P} \left[Z_i \leq F_{Z_i}^{[-1]}(F_i(t)) \right] ,$$

where τ_i denotes its default time and $F_{Z_i}^{[-1]}$ is the quantile function of Z_i , the generalized inverse of its cumulative distribution, noted F_{Z_i} . The coefficient ρ is for that reason called the *loading factor*, and ρ^2 is the correlation between any $Z_i, Z_j, j \neq i$, sometimes called the *model* or *copula* correlation. When the variables Z and \tilde{Z}_i are zero-mean and unit-variance, so are the Z_i .

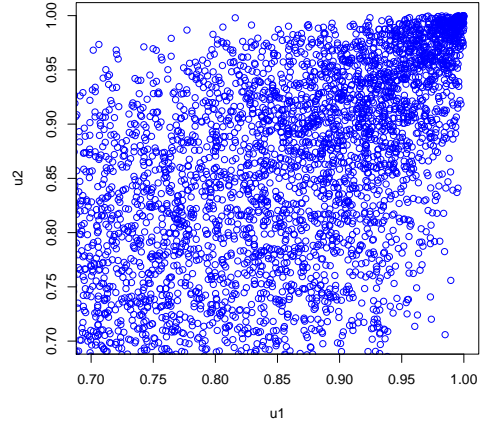
Conditionally on the common variable Z , the default times are independent, and

$$\mathbb{P}[\tau_1 \leq t_1, \dots, \tau_N \leq t_N] = F_{\vec{\tau}}(t_1, \cdot, t_N) = \mathbb{E} \left[\prod_{i=1}^N F_{i|Z}(t_i|Z) \right] .$$

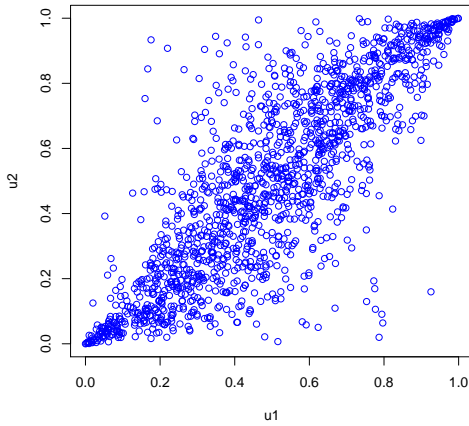
¹Indeed, once the default times are deterministically known, so is the size of the tranche at any time, and consequently so is the premium leg. Further, we also know how and when any default will impact the tranche, which gives the present value of any loss on the tranche resulting from a default. Actually, if T denotes the maturity date of the contract, Val is even \mathcal{F}_T -measurable, where $\mathcal{F}_t \doteq \bigvee_{i=1}^N \sigma(X_{\{\tau_i \leq s\}}; s \leq t)$ and $\mathcal{F}_T \subseteq \sigma(\vec{\tau})$.



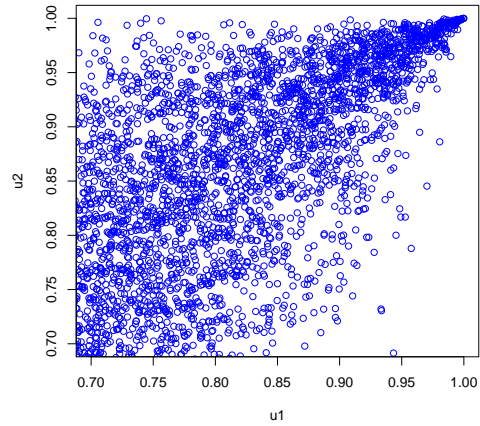
(a) Gaussian copula



(b) Gaussian copula (zoom in the top-right corner)



(c) Double t copula ($\nu = \bar{\nu} = 4$)



(d) Double t copula ($\nu = \bar{\nu} = 4$, zoom in the top-right corner)

Figure 1: Scatter plot of samples drawn from the Gaussian and the double t copula with correlation equal to 80%. One can see by looking at the global pictures that the samples are indeed correlated, but that the extreme samples appear to be almost independent in the case of the Gaussian (and they are, at the extreme, whatever is the copula correlation provided that it is not exactly 1) while this is not the case for the double t.

So,

$$\begin{aligned}\mathbb{E}[\text{Val}] &= \int_z \int_{t_1 \times \dots \times t_N} \text{Val}(\vec{\tau}) \prod_{i=1}^N dF_{i|Z}(\tau_i|z) dF_Z(z) \\ &= \int_z \left(\int_{t_1} \left(\cdot \left(\int_{t_N} \text{Val}(\vec{\tau}) dF_{N|Z}(t_N|z) \right) \cdot \right) dF_{1|Z}(t_1|z) \right) dF_Z(z) .\end{aligned}$$

Although the above expression does not correspond to the one which is used in practice (the latter involves the computation of the loss distribution using e.g. a recursion algorithm), this view is informative in the sense that it shows in a concise way that proper valuation of such a product in a one-factor copula framework only requires accurate:

- estimation of conditional distributions

$$F_{i|Z}(t|z) = \mathbb{P} \left[\rho z + \sqrt{1 - \rho^2} \tilde{Z}_i \leq F_{Z_i}^{[-1]}(F_i(t)) \right] = F_{\tilde{Z}_i} \left(\frac{F_{Z_i}^{[-1]}(F_i(t)) - \rho z}{\sqrt{1 - \rho^2}} \right) ,$$

- integration with respect to conditional distributions $F_{i|Z}$,
- integration with respect to the common state variable distribution F_Z .

Because all the \tilde{Z}_i are i.i.d., the distribution F_{Z_i} is not name-specific but, generally speaking, correlation-dependent. We can then drop the index i and use the simpler notations :

$$\begin{aligned}F(\cdot; \rho) &\doteq F_{Z_i}(\cdot) , \\ F^{[-1]}(\cdot; \rho) &\doteq F_{Z_i}^{[-1]}(\cdot) .\end{aligned}$$

The Gaussian copula framework is a specific case since we have $F(x; \rho) = F_{\tilde{Z}_i}(x) = \Phi(x)$, with $\Phi(x)$ the Normal standard cumulative distribution, and is independent of the loading factor. This is because the Gaussian function is stable under convolution : the sum of two Normal variables is again a Normal variable. When the weights are well-chosen, the resulting variable Z_i is again standard Normal. This is the case in the Gaussian Copula latent-variable model.

In the double t copula however, this is no longer the case. It is indeed the distribution of a weighted sum of two unit-variance Student's t variables ($Z \doteq Z(\nu) \sqrt{\frac{\nu-2}{\nu}}$ and $\tilde{Z}_i \doteq \tilde{Z}_i(\tilde{\nu}) \sqrt{\frac{\tilde{\nu}-2}{\tilde{\nu}}}$, with $Z(\nu)$, and $\tilde{Z}_i(\tilde{\nu})$ being independent Student's t random variables with $\nu > 0$ and $\tilde{\nu} > 0$ degrees of freedom, respectively, and $\tilde{Z}_i(\tilde{\nu}) \perp \tilde{Z}_j(\tilde{\nu})$ for $i \neq j$), which is a distribution not stable under convolution. Even worse, this distribution has no known analytical form. By contrast, noting $T_\nu(x) \doteq \mathbb{P}[Z(\nu) \leq x]$, we have $F_Z(z) = T_\nu \left(z / \sqrt{\frac{\nu-2}{\nu}} \right)$ and $F_{\tilde{Z}_i}(v) = T_{\tilde{\nu}} \left(z / \sqrt{\frac{\tilde{\nu}-2}{\tilde{\nu}}} \right)$.

Consequently, when proceeding to the valuation of a structured credit product under a copula approach, the only thing we need is i) the estimation of the function $F^{[-1]}(\cdot; \rho)$ together with ii) an efficient method to perform the numerical integration with respect to a Student's t distribution. The next sections are dedicated to these points.

3 Estimation of $F^{[-1]}(\cdot; \rho)$ function

The estimation of the quantile function of the double t random variable Z_i , $F^{[-1]}(\cdot; \rho)$, first require the estimation of the cumulative distribution $F(\cdot; \rho)$. From this, we can numerically obtain its inverse.

3.1 Estimation of $F(\cdot; \rho)$

We present two possible approaches for that purposes: the simulation-based approach and the semi-analytic technique.

3.1.1 Simulation-based approach

This technique aims at empirically estimating the function $F(\cdot; \rho)$ for some given value of the loading factor by combining Student's t samples according to the linear mixture model given by the copula. This can be done efficiently by taking advantage of known information regarding Z_i . For instance, it is a symmetric random variable, so that $F(-x; \rho) = 1 - F(x; \rho)$. In other words, we simply need to estimate $F(\cdot; \rho)$ on $[0, \infty[$. Further, note that if $\nu = \tilde{\nu}$, $F(\cdot; \rho) = F(\cdot; \sqrt{1 - \rho^2})$.

Here are the big lines of the process to estimate $F(x; \rho)$.

1. A two-dimensional grid needs to be built approach, meaning we will discretize our 2D (correlation, quantile) space. We build a correlation grid ranging in $[0, \sqrt{2}/2]$ with resolution defined by choosing a value for the step. Similarly we build a quantile grid, starting at 0, ending at some very large value parametrized by some given step. The probability grid ranges in $[0.5, 1]$.
2. The chosen method consists in generating n Student's t samples both for Z and \tilde{Z}_i , and combining them according to the mixture model.
3. From these samples, we estimate the CDF $F(x; \rho)$ as follows. Starting from zero and counting the number of samples lower than each of the (positive) quantile value stored in the grid, we come up with a frequency analysis. We have built a stepwise function, *starting from the bottom*. We can do the same *starting from the top*. We come up with two stepwise functions.
4. The steps in the estimation of $F(x; \rho)$ are then removed: every the constant pieces are replaced by linearly increasing pieces obtained through interpolation. The CDF curves are now strictly increasing on the quantile grid, as it should.
5. These two tables can then merged to come up with an average function².

Evaluating the function $F(x; \rho)$ would then require to look up for values in a grid, with proper interpolation techniques. This is very fast once the table is loaded. However, this may induce some inaccuracies. In addition, if the model needs to be ran for several couples of degrees of freedom $(\nu, \tilde{\nu})$, one table has to be performed for each such a pair (up to some redundancies).

Another method can be thought of, which is more flexible in the sense that no interpolation is needed. The function is estimated directly at the exact point (x, ρ) for any pair $(\nu, \tilde{\nu})$. This technique is presented below.

3.1.2 Semi-analytic approach

The CDF of Z_i , $F(x; \rho)$, can be computed by convolving the rescaled Student's t distributions. Another viewpoint which ends up with the same technique is to first conditioning in the latent variable model on the common state variable Z and then integrating out the latter, using the law of iterated expectation :

$$F(x; \rho | Z = z) = T_{\tilde{\nu}} \left(\frac{x - \rho \sqrt{\frac{\nu-2}{\nu}} z}{\sqrt{1 - \rho^2} \sqrt{\frac{\tilde{\nu}-2}{\tilde{\nu}}}} \right),$$

²Actually, we can build several "bottom" and "top" tables and compute their averages. This extension aims at artificially increasing the number of samples while avoiding computer memory issues.

and therefore,

$$F(x; \rho) = \mathbb{E}[F(x; \rho)|Z] = \int_{z=-\infty}^{\infty} T_{\tilde{\nu}} \left(\frac{x - \rho \sqrt{\frac{\nu-2}{\nu}} z}{\sqrt{1 - \rho^2} \sqrt{\frac{\tilde{\nu}-2}{\tilde{\nu}}}} \right) dT_{\nu}(z) .$$

It is interesting to note that one ends up again with the problem of integrating a function with respect to a Student's t distribution. Consequently, our three items listing in Section 2 the tasks to achieve for a proper implementation of the model all reduce to the accurate computation of such an integral.

3.2 Estimation of $F^{[-1]}(\cdot; \rho)$

Once we have the double t CDF $F(\cdot; \rho)$, we need to compute its inverse, the corresponding quantile function. Again, this can be done using a table or using a more flexible method.

It is usually advised in the literature to tabulate the values of the double-t quantile function. Given the double t CDF, this table can be directly computed. So, evaluating the quantile function $F^{[-1]}(p; \rho)$ at some point (p^*, ρ^*) only requires to loop up for the quantile value yielding the desired probability for the inputted value for the loading factor. Given the monotonic behavior of those functions, this can be managed very efficiently when efficient "look up" algorithms (like dichotomic search) are used. However, it suffers from several major drawbacks. First it requires to define the grid settings, that will impact precision. Increasing its size will i) increase the size of the file to load, and hence limit the portability of the code, and ii) will slow down the process because of the search procedure. Further, because there is no reason that the point (p^*, ρ^*) actually matches one of the table entries this approach requires to set up two-dimensional interpolations, together with the construction of several tables, if the model has to be ran with several pairs of $(\nu, \tilde{\nu})$ (because these parameters impact the double t CDF and consequently, on its generalized inverse).

For all these reasons, it might be interesting to avoid tabulation. This can be done very efficiently thanks to *root-finding* algorithms. We can look for the quantile x^* leading to a given probability p simply by searching for the (unique) root $x = x^*$ of the function $F(x; \rho) - p$. Practically, we are iterating on x such that $|F(x; \rho) - p| < \epsilon$ for some given precision threshold $\epsilon > 0$. For this purpose, popular algorithms like the Newton, the Secant or Brent's method can be used. Hence, the problem basically reduces to the double t CDF estimation.

This approach, combined to the semi-analytical procedure for the CDF computation is interesting. Although a bit less speed-efficient than the table-based approach (after the loading of the table, which could be time consuming but only needs to be performed once per run), it avoids the use of bi-dimensional interpolation. The value is estimated directly for the correct pair (p^*, ρ^*) . Further, it can be done *on-the-fly* for any pair $(\nu, \tilde{\nu})$ without having to rebuilt big tables.

Consequently, adopting the above technique, the only remaining point to properly implement the double t copula pricing framework is an efficient numerical integration method.³ This is a very critical point, which is the purpose of the next section.

4 Integrating with respect to the Student's t

Accurate pricing of structured credit product is extremely important, for several reasons.

1. First, because they are highly leveraged products : probabilities will be multiplied by loss levels, which can typically grow up to several billions.

³Actually, it can be shown that more specifically, the only remaining issue from a numerical perspective is to integrate *cumulative distributions* with respect to a Student's t distribution, namely either a shifted-rescaled Student's t cumulative distribution or a loss distribution. This is interesting because of the nice regularity properties this class of functions benefits from.

2. Second, the pricing of CDO tranches in the base correlation framework is highly sensitive to numerical errors since it involves the difficult task of subtracting a series of large numbers (expected losses at some evaluation times) computed on two distinct equity pieces priced with different loading factors $\rho_{attach}, \rho_{detach}$. This requires a very good accuracy, in order to avoid, for example, to have a contradiction as big as negative expected losses.
3. Third, we will often need to look up for *tail values*, and the probability related to extreme events need to be accurate. We cannot simply focus on methods giving satisfactory results for “likely” events according to the distribution, we need high accuracy when evaluating probabilities of extreme events. This is very misleading since even if two CDFs could look, at visual inspection, to be very close from each other, they could lead to very different quantile values when evaluated at probabilities close to 0, 1 because of the horizontal asymptotes.

Let us illustrate using a simple example why a high accuracy is so crucial. For instance, assume that for numerical reasons, the probability to have a loss (relative to the pool notional) $L(t)$ at some evaluation time t being larger than the maximum one, say K (defined as the sum of underlyings’ losses given default) is not exactly 1 (as it should) but rather $\Pr[L(t) \leq K] = 0.9999$. Although this value may seem very satisfactory at first sight, one could easily obtain the inconsistent result of a non-zero (and actually, several ten thousands) protection leg of a tranche attaching at a point higher than K . From a risk management perspective, this is most often unacceptable. As an example, assume $K = 60\%$. The expected loss on a 60% – 100% belongs to $[0.6 \times (1 - 0.9999), (1 - 0.9999)]$ per unit of pool’s notional. In the case of a 125-names pool, each underlying’s notional being $10M$, the expected loss on the above tranche is *bounded below* by $75k$ in the ad hoc currency unit. This is far from being a satisfactory result for such a fully protected tranche.⁴

Knowing how important is the accuracy of probabilities, we have to pay attention to each numerical computation. In the Gaussian copula case, one needs to perform a numerical integration with respect to the Gaussian function. This is not a big deal since specific quadrature techniques are dedicated to that problem (see below). This is no longer the case in the double t framework. In the sequel we show how critical is the choice of an appropriate integration technique, and present a specific method for achieving this task in our double t case.

A Gaussian quadrature technique is an efficient method for performing the numerical integration of a function $h(x)w(x)$ with respect to the variable x over a specific interval $[a, b]$. Depending on the weight function $w(x)$ and the interval $[a, b]$ a specific kind of quadrature to be used has to be determined. The integral is simply approximated by a weighted sum of the function $f(x)$ evaluated at some specific points (nodes):

$$\int_{x=a}^b h(x)w(x)dx \approx \sum_{i=1}^K \omega_i h(x_i) . \quad (1)$$

The *nodes* and the *weights* are given by the quadrature technique once the *number of integration points* K is known. This technique is very efficient in the sense that the approximation error is strictly zero for every function $h(\cdot)$ being a polynomial of order less than $2K$. With $K = 32$, the above approximation is then exact whatever is $h(\cdot)$ provided that it is a polynomial of order less than 64. Further, the approximation error decreases with K under some conditions. More information regarding this technique can be found in e.g. [6].

When valuing a CDO tranche or n -th to default in the Gaussian copula framework, we have $w(x) = \phi(x)$ where $\phi(x)$ is the standard Normal density and $h(\cdot)$ is finite and strictly positive everywhere.⁵ In this particular case, the *Gauss-Hermite* quadrature yields a very efficient approximation, even for limited K

⁴One could argue that because the recoveries are not *a priori* known, this tranche is not fully protected. However, the model *assumes* that recoveries are a given, and is not aware about the fact that they could fluctuate. So, from a model perspective, this tranche is indeed 100% safe, and the returned price should be consistent with that.

⁵Indeed, it is a conditional cumulative (loss or default) probability with support being the whole real line.

(typically, $K = 32$ is enough).⁶

Pricing this class of product in the double t copula framework requires, by contrast, to integrate a function with respect to a Student's t distribution with ν degrees of freedom ($w(x) = t_\nu(x)$) which is, as shown above, the only critical point of this model. It is tempting to simply apply any arbitrary quadrature with integration domain being the real line to the function $h(x) = f(x)t_\nu(x)/w(x)$, since in this case

$$\int_{\mathbb{R}} f(x)t_\nu(x)dx = \int_{\mathbb{R}} f(x)\frac{t_\nu(x)}{w(x)}w(x)dx = \int_{\mathbb{R}} h(x)w(x)dx \approx \sum_{i=1}^K \omega_i h(x_i) .$$

For example, one could simply approximate the integral $\int_{\mathbb{R}} f(x)t_\nu(x)dx$ by the Gauss-Hermite method applied to the function $h(x) = f(x)t_\nu(x)/\phi(x)$. Actually, although the above rescaling is perfectly legitimate, the question reduces to find the set of functions for which the approximation makes sense. It is intuitively obvious that the approximation is not equally good whatever is $f(x)$. We give below two good reasons showing that such a technique is not suitable.

1. Intuitively, this is not a good idea. Because the Gaussian density tends faster to zero than any (finite) power of x as $x \rightarrow \infty$, the integrand $h(x)$ will diverge to infinity as x tends to ∞ . Because of this problematic tail-behavior, it is likely that the approximation will have some problems to hold in such a case. As an illustration, we give a plot of the function $\log(h(u))$ where $h(u) = f(u)t_\nu(\tan(\pi/2u))/\phi(\tan(\pi/2u))$ with $f(u) = 1$ (Fig. 2(a)) and $f(u) = T_\nu(\tan(\pi/2u))$ (Fig. 2(b)) in the case $\nu = 4$. It can be shown that the function is continuous, but diverges to zero as $x \rightarrow |\infty|$ (see the appendix).
2. Theoretically, there is also an argument justifying to not use the above methodology. Although an approximation method for which no finite error bound is available is not necessarily worst than another for which a finite bound can be found, having a finite error bound is always a nice feature to have for an approximation technique. We show in the appendix that the error bound obtained by applying the Gauss-Hermite quadrature to the function $h(x) = f(x)t_\nu(x)/\phi(x)$ is not finite.

As a consequence, the above technique does not seem to be appropriate because of the huge (unbounded) values that the integrand can take; we could not simply re-define the function $h(x)$ to be integrated with respect to the kernel $w(x)$ to enter the framework of any other quadrature technique (the consequences of doing this in our application will be emphasized in a simulation below). Henceforth, we propose to circumvent this issue by selecting a specific quadrature, the *Gauss-Legendre* quadrature, for which the error bound has the same form as the *Gauss-Hermite* one. Specifically, in the case of the K -points Gauss-Legendre quadrature and for an integrand $f(x)$ with continuous $2K$ derivative $f^{(2K)}(x)$, the error made by approximating the integral $\int_{x=a}^b f(x)dx$ is bounded above by (see e.g. [7])

$$\frac{(b-a)^{2K+1}(K!)^4}{(2K+1)(2K!)^3} f^{(2K)}(\xi) ,$$

where ξ is some point in (a, b) . The framework of the K -points Gauss-Legendre integration provides the set of weights and nodes $\cup_{i=1}^K (x_i^{\mathcal{L}}, \omega_i^{\mathcal{L}})$ to approximate

$$\int_{x=-1}^1 f(x)dx \approx \sum_{i=1}^K f(x_i^{\mathcal{L}})\omega_i^{\mathcal{L}} . \tag{2}$$

Observe that the integration domain is now the interval $[-1, 1]$. In order to take profit of the Legendre integration technique, we hence first have to map our Student's t integration domain from \mathbb{R} to $[-1, 1]$. This

⁶Actually, these weights and abscis are not given for $\phi(x)$ but for $w(x) = e^{-x^2}$. However, both can be seen as equivalent since the functions are equivalent up to a change of variable and rescaling.

can be used through a change of variable. For example, using the $\arctan(\cdot)$ mapping :

$$\int_{-\infty}^{\infty} f(x)dT_{\nu}(x) = \int_{-\infty}^{\infty} f(x)t_{\nu}(x)dx = \frac{\pi}{2} \int_{u=-1}^1 \underbrace{\frac{f(\tan(\pi/2u))t_{\nu}(\tan(\pi/2u))}{\cos^2(\pi/2u)}}_{\doteq h(u)} du \approx \frac{\pi}{2} \sum_{i=1}^K \omega_i^{\mathcal{L}} h(x_i^{\mathcal{L}}) .$$

The above technique, named in the sequel *mapped Legendre* may seem inadequate because the integrand could, at first sight, take infinite values around $u = |1|$. We would then end up with a similar situation than the one faced with the Gauss-Hermite rule. The method for solving this integral is relevant provided that the function h does not diverge as $u \rightarrow |1|$. The integrand will be finite everywhere provided that $\lim_{u \rightarrow 1} h(u) = a$ and $\lim_{u \rightarrow -1} h(u) = b$ for some $-\infty < a, b < \infty$, where it is enough to study only one of the limits by symmetry.

Actually, it is easy to see that $a = b = 0$. Indeed, noting that $dT_{\nu}(\tan(u))/du = t_{\nu}(\tan(u))/\cos^2(u)$, one gets

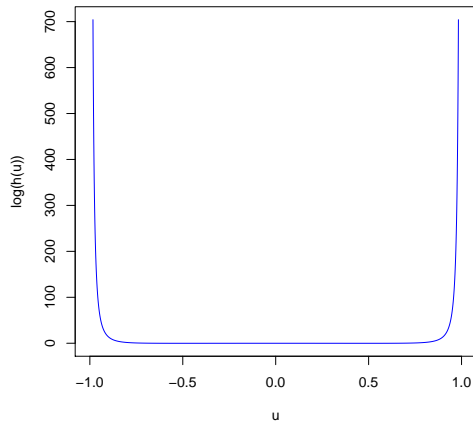
$$\lim_{u \uparrow \pi/2} \frac{t_{\nu}(\tan(u))}{\cos^2(u)} = \lim_{u \uparrow \pi/2} \frac{dT_{\nu}(\tan(u))}{du} = \lim_{x \rightarrow \infty} \frac{dT_{\nu}(x)}{dx} = \lim_{x \rightarrow \infty} t_{\nu}(x) = 0 ,$$

where the second equality results from the continuity of the $\tan(x)$ function in $] - \pi/2, \pi/2[$. Because the function $f(\cdot)$ we are dealing with have bounded limits (it is either a cumulative distribution defined on \mathbb{R}), this proves that $h(\cdot)$ is finite on its definition domain. As an illustration, we give a plot of the function $h(u)$ where $h(u) = f(u)t_{\nu}(\tan(\pi/2u))/\cos^2(\pi/2u)$ with $f(u) = 1$ (Fig. 2(c)) and $f(u) = T_{\nu}(\tan(\pi/2u))$ (Fig. 2(d)) in the $\nu = 4$ case. It can be shown that the function is continuous, bounded everywhere and converges smoothly to zero as $u \rightarrow |1|$.

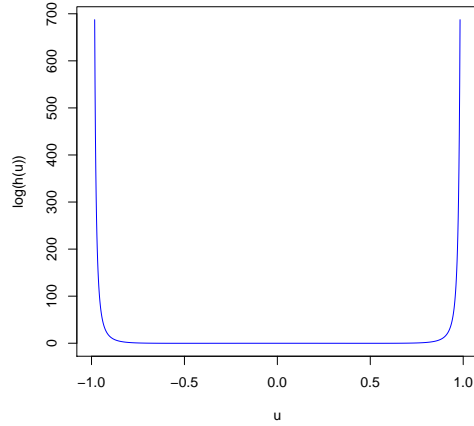
This shows that the *Gauss-Legendre* quadrature combined to the $\arctan(\cdot)$ mapping does not suffer from the drawback faced in the *Gauss-Hermite* one. In addition, it has the nice theoretical feature to have a finite error bound. This can be shown as follows. First, the integrand is bounded on the compact set $[-1, 1]$ because i) it is on any compact set belonging to $] - 1, 1[$ and ii) the limits at the boundaries exist (and equal 0). We can then call the *Boundedness Theorem*, which says that any function being continuous on a compact set is bounded. Since the integrand admits a continuous n -th order derivative for all n , any of these derivatives will also be bounded. This concludes the proof of the approximation error boundedness for any number of integration points K . Knowing how crucial it is to have high accuracies on our probabilities, it is interesting to conclude this paper by comparing the error made when the double t CDF $F(x, 0)$ is computed with the help of the two quadratures approximating $\int_z T_{\nu} \left(x / \sqrt{\frac{\nu-2}{\nu}} \right) dT_{\nu}(z)$. The $\rho = 0$ case is indeed interesting in the sense that the solution is analytically known :we should obviously have $F(x, 0) = T_{\nu} \left(x / \sqrt{\frac{\nu-2}{\nu}} \right)$. Figure 3 illustrates the approximation errors for both methods. It is seen that increasing the number of integration points does not improve drastically the approximation obtained through the Gauss-Hermite procedure, while the result is convincing when the mapped Legendre technique is used. Noting that this error can occur in any integration (i.e. when evaluating the double t CDF or when integrating out the common variable), one understands that it is possible to come up with unexpected inconsistent results if no specific attention is devoted to the integration (see the previous example, illustrating the consequences of having a probability of 0.9999 instead of 1 when pricing a *fully protected* CDO tranche).

5 Conclusion

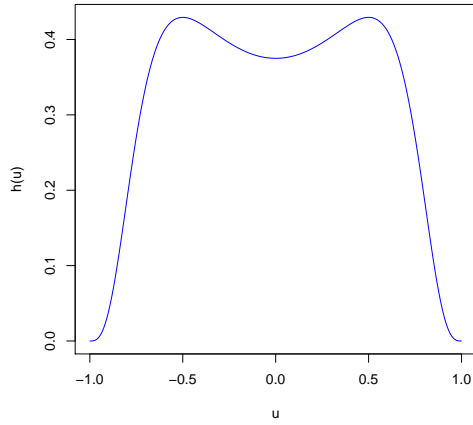
It is well-known from both analytical and practical studies that compared to the Gaussian copula, other “fatter-tails” coupling schemes are much more interesting when pricing large-dimensional products related to extreme events, like CDOs and n -th to default. Among those is the *double t* copula, which involves Student’s t distributions. One could then question why this last model didn’t become more widely accepted



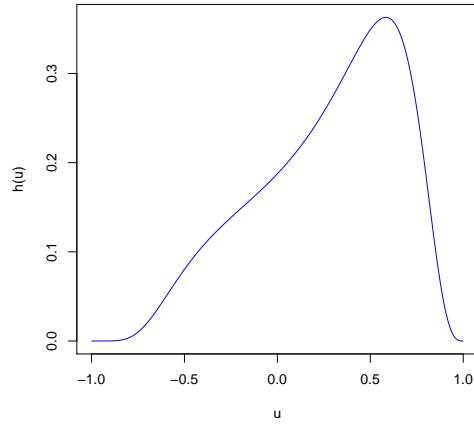
(a) $\log(h(u)), h(u) = \frac{t_4(\tan(\pi/2u))}{\phi(\tan(\pi/2u))}$



(b) $\log(h(u)), h(u) = T_4(\tan(\pi/2u)) \frac{t_4(\tan(\pi/2u))}{\phi(\tan(\pi/2u))}$



(c) $h(u), h(u) = \frac{t_4(\tan(\pi/2u))}{\cos^2(\pi/2u)}$



(d) $h(u), h(u) = T_4(\tan(\pi/2u)) \frac{t_4(\tan(\pi/2u))}{\cos^2(\pi/2u)}$

Figure 2: Top row : function to be integrated with respect to the standard Normal density ($w(x) = \phi(x)$). Bottom row, function to be integrated with respect to the function ($w(x) = 1$).

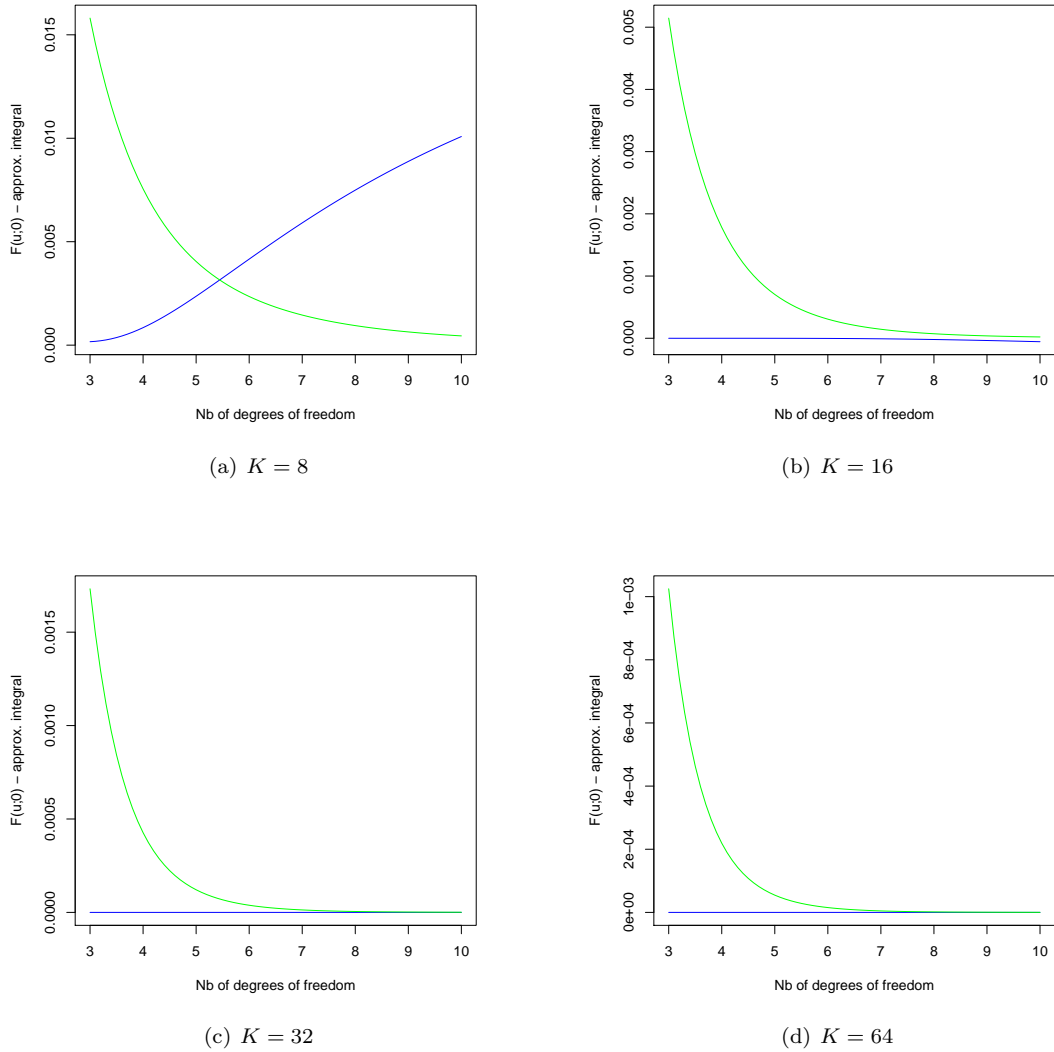


Figure 3: Approximation errors of the integral $\int_z T_\nu \left(x / \sqrt{\frac{\nu-2}{\nu}} \right) dT_\nu(z)$ with the *Gauss-Hermite* (green) or the *mapped Legendre* (blue) vs the number of degrees of freedom ν for several number of integration points K .

in the related community. Indeed, although it does not solve all the problems faced by the Gaussian copula, it already fixes quite a few major drawbacks. Actually, we think that practical implementation issues explain this situation. The double t copula model may seem to be at first sight quite unreliable and unstable when no specific attention is devoted to its practical implementation. This paper has dealt with this aspect of the model. It has been shown that i) the only difficulty actually reduces to the proper computation of a specific kind of integrals which have no known analytical solutions, but in parallel ii) even if we cannot blindly apply the technique used in the Gaussian copula framework, there exists an efficient method for performing this task. The presented results reduce the drawbacks of the double-t copula compared to the Gaussian one to the simple need of an iterative procedure for estimating the quantile function of the common variable.

Appendix : Gauss-Hermite quadrature applied to $f(x)t\nu(x)/\phi(x)$

Tail behavior of $f(x)t\nu(x)/\phi(x)$ ⁷

Let us show that $\lim_{x \rightarrow \infty} t\nu(x)/\phi(x) = \infty$. Noting that

$$t\nu(x) \doteq \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{2}\right)^{-\frac{\nu+1}{2}},$$

it is enough to prove that $\lim_{x \rightarrow \infty} x^k \phi(x) = 0$ for all $k \neq 0$. Indeed, this would imply that the result also apply to $k = \nu + 1$ for any $\nu > 2$ (this is needed for the variance of the Student's t variable to exist). Since the exponential growth overtakes any polynomial growth (i.e. increases faster than any power of x , this can be checked by looking at the Taylor expansion of the exponential function), we have

$$\lim_{x \rightarrow \infty} x^k \phi(x) = \lim_{x \rightarrow \infty} \frac{x^k}{e^{x^2/2}} = 0. \quad (3)$$

Let us define the set \mathcal{P} of functions $q(x)$ of the form

$$q(x) = \sum_{j=0}^J a_j x^j \left(\sum_{k=0}^{K_j} b_{j,k} p(x)^{-\beta_{j,k}} \right),$$

where $J < \infty$, $K_j < \infty$, $0 < \beta_{j,k} < \infty$ and $p(x)$ is a K -th order polynomial ($K < \infty$) with positive coefficients.

We call any member of \mathcal{P} a “ \mathcal{P} -polynomial”, and define $d(q(x))$ the *dominator* of the \mathcal{P} -polynomial $q(x)$:

$$d(q(x)) \doteq \text{sign}(a_{j^*} b_{j^*,k^*}), \quad (j^*, k^*) \doteq \text{argmax}_{j,k} \{j - \beta_{j,k} | a_j b_{j,k} \neq 0\}.$$

The dominator of $q(x)$ is the sign of the coefficient of the largest power of x involved in $q(x)$ (because the coefficients of $p(x)$ are all positive).

For all \mathcal{P} -polynomial $q_1(x)$, the growth property of the exponential function yields that

$$\lim_{x \rightarrow \infty} \frac{q_1(x)}{\phi(x)} \stackrel{(3)}{=} d(q_1(x)) \cdot \infty. \quad (4)$$

This shows that the limit of any ratio of a \mathcal{P} -polynomial by the Normal density is $\pm\infty$.

Because $t\nu(x) \in \mathcal{P}$ (with $J = 0$, $K_0 = 0$, $a_0 = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)}$, $b_{0,0} = 1$, $p(x) = (1 + x^2/2)$, $\beta_{0,0} = (\nu + 1)/2$), this proves that

$$\lim_{x \rightarrow \infty} \frac{t\nu(x)}{\phi(x)} = \infty.$$

⁷The proof presented here is made more complicated than needed because the general result will be useful in the next Section.

Same applies to $\lim_{x \rightarrow \infty} f(x) \frac{t\nu(x)}{\phi(x)}$ when, among other, $f(x)$ has a non-zero limit as $x \rightarrow \infty$, which is obviously the case when $f(x)$ is a cumulative distribution defined on the real line.

This can be visually checked on two specific examples in Fig. 2(a) and Fig. 2(b).

Bound of the Gauss-Hermite approximation of $\int_x (f(x)t\nu(x)/\phi(x)) \phi(x) dx$ ⁸

We shall first derive interesting properties of the set \mathcal{P} : (i) it is closed under addition, (ii) if $q(x) \in \mathcal{P}$ then $p(x)q(x) \in \mathcal{P}$ for all finite-order polynomial $p(x)$ and (iii) $dq(x)/dx \in \mathcal{P}$.⁹ The first two properties (i) and (ii) are fairly straightforward to check. Regarding the last one, the derivative of $q(x)$ is the sum of the derivatives of each term $a_j x^j \sum_{k=0}^{K_j} b_{j,k} p(x)^{-\beta_{j,k}}$, which are also particular \mathcal{P} -polynomials. Each of those derivatives can be seen to be themselves \mathcal{P} -polynomials (we drop the index $j \geq 0$ for convenience when not needed) :

$$d/dx \left(a_j x^j \sum_{k=0}^K b_k p(x)^{-\beta_k} \right) = j a_j x^{j-1} \sum_{k=0}^K b_k p(x)^{-\beta_k} + \frac{dp(x)}{dx} \overbrace{a_j x^j \sum_{k=0}^K -\beta_k b_k p(x)^{-(\beta_k+1)}}^{\in \mathcal{P}} .$$

Because $p(x)$ is a polynomial, so is $dp(x)/dx$ and we have using (ii) that both terms in the RHS of the above equation are \mathcal{P} -polynomials. Using (i), the whole RHS belongs to \mathcal{P} , and hence so is the derivative of each term involved in \mathcal{P} . Applying (i) once again, we have that the sum of all these derivatives is itself a \mathcal{P} -polynomial, which concludes the proof of (iii).

We now turn to the proof of the error bound, and consider the general case of $\nu > 2$, $\nu \in \mathbb{R}$, and will first show that the limit of any n -th order derivative of a ratio of the form $\frac{q(x)}{\phi(x)}$, $q(x) \in \mathcal{P}$ will be unbounded as $x \rightarrow \infty$. To that end, we first establish that the set

$$\mathcal{S} \doteq \{q(x)/\phi(x) : q(x) \in \mathcal{P}\}$$

is closed under differentiation. For instance, we have for any $q_1(x) \in \mathcal{P}$

$$\frac{d^n}{dx^n} \left(\frac{q_1(x)}{\phi(x)} \right) = \frac{d^{n-1}}{dx^{n-1}} \left(\frac{dq_1(x)/dx + xq_1(x)}{\phi(x)} \right) .$$

It turns out from the properties of \mathcal{P} that $q_2(x) \doteq xq_1(x) + dq_1(x)/dx$ belongs to \mathcal{P} and hence $q_2(x)/\phi(x) \in \mathcal{S}$.

It is obvious that $d(x^r q_1(x)) = d(q_1(x))$ for all $r \in \mathbb{N}$ (because $\operatorname{argmax}_j \{r + j - \alpha\} = \operatorname{argmax}_j \{j - \alpha\}$), and it can be checked that $d(q_2(x)) = d(xq_1(x))$ since the largest power of x in $xq_1(x) + dq_1(x)/dx$ will appear in the first of the two terms. Consequently, the n -th order derivative of any function $r_1(x) = q_1(x)/\phi(x) \in \mathcal{S}$ is the $(n-1)$ -th derivative of another specific function member of \mathcal{S} , say $q_2(x) = p_2(x)/\phi(x)$:

$$\frac{d^n}{dx^n} \left(\frac{q_1(x)}{\phi(x)} \right) = \frac{d^{n-1}}{dx^{n-1}} \left(\frac{q_2(x)}{\phi(x)} \right), \quad d(q_1(x)) = d(q_2(x)) .$$

By recursion, we have that for any $q_1(x) \in \mathcal{P}$ and for all $n \in \mathbb{N}$

$$\frac{d^n}{dx^n} \left(\frac{q_1(x)}{\phi(x)} \right) = \frac{q_{n+1}(x)}{\phi(x)}, \quad q_{n+1}(x) \in \mathcal{P}, \quad d(q_1(x)) = d(q_{n+1}(x)) .$$

Consequently, because we know from eq. (4) that the limit of any element of \mathcal{S} as $x \rightarrow \infty$ only depends on the dominator of the \mathcal{P} -polynomial involved, one gets that the limit of any n -th order derivative of any

⁸This proof may seem unnecessary involved. However, it requires the finding of a set which has very specific properties (like *closure* under some specific operations) in order to be able to apply a recursion. Each simpler investigated attempt has revealed nasty problems: quite difficult to detect but still making the proof not completely correct. We propose here the simplest “bug-free” version found.

⁹Note that $q(x)$ differs from $p(x)$ in the sense that the powers of x involved in $q(x)$ can be negative and fractional, while these exponents are restricted to be positive whole numbers for $p(x)$.

$r(x) = q(x)/\phi(x) \in \mathcal{S}$ as $x \rightarrow \infty$ is equal to the limit of $r(x)$, which is given by $d(q(x))$ and is either $-\infty$ or ∞ .

Finally, because $t_\nu(x) \in \mathcal{P}$ and so $t_\nu(x)/\phi(x) \in \mathcal{S}$ and $\lim_{x \rightarrow \infty} \frac{d^n}{dx^n} (t_\nu(x)/\phi(x)) = \lim_{x \rightarrow \infty} t_\nu(x)/\phi(x)$ (i.e. $\pm\infty$). Therefore, because the bound on the error is determined by the absolute value of the $2K$ -th derivative of $t_\nu(x)/\phi(x)$ evaluated at some (undetermined) real point (see for example [6, 8]), this point could be arbitrarily large in absolute value, given no finite bound for the error.

It can be checked that the same extends to the error bound of $f(x)t_\nu(x)\phi(x)$ up to some mild regularity conditions on the function $f(\cdot)$. For instance, it is enough to require that the limit of each of the n -th derivative of f (for all $n \in \mathbb{N}$ such that $n \leq 2K$) as $x \rightarrow |\infty|$ are bounded, which is trivially met in the case where $f(\cdot)$ is a cumulative distribution.

Disclaimer

The views expressed here are those of the author, and do not necessarily reflect the position of his employer.

References

- [1] X. Burtschell, J. Gregory and J.-P. Laurent, *A comparative analysis of CDO pricing models*, working paper. Available for download on <http://laurent.jeanpaul.free.fr>
- [2] P. Embrechts, A. McNeil and D. Straumann, *Correlation and Dependence in Risk Management : Properties and Pitfalls*. In : Risk Management : Value at Risk and Beyond, Cambridge University Press, 2002.
- [3] U. Cherubini, E. Luciano and W. Vecchiato, *Copula Methods in Finance*. Wiley Finance, 2004.
- [4] P. Wanitjirattikal and S. Kiatsupaibul, *Tail dependence of Student's t copula and Double t copula and their effects on Pricing Credit Derivatives*, Thammasat Int. J. Sc. Tech. 12(2), 2007. Available for download on http://www.tijsat.tu.ac.th/issues/2007/no2/2007_V12_No2_1.PDF
- [5] J. Hull and A. White, *Valuation of a CDO and n-th to Default Without Monte Carlo Simulation*. J. Derivatives, Winter 12(2), pp. 8-23, 2004. Available for download on <http://www.rotman.utoronto.ca/hull/downloadablepublications/HullWhiteCDOPaper.pdf>
- [6] F.B. Hildebrand, *Introduction to Numerical Analysis*. McGraw-Hill, 1956.
- [7] D. Kahaner, C. Moler and S. Nash, *Numerical Methods and Software*. Prentice-Hall, 1989.
- [8] J. Stoer and R. Bulirsh, *Introduction to Numerical Analysis*. Springer, 2002.