NONPARAMETRIC ESTIMATION OF COPULAS FOR TIME SERIES

Jean-David FERMANIAN[a] and Olivier SCAILLET[b] [1]

[a] CDC Ixis Capital Markets, 56 rue de Lille, 75007 Paris, and CREST, 15 bd Gabriel Péri, 92245 Malakoff cedex, France. jdfermanian@cdcixis-cm.com

[b] HEC Genève and FAME, Université de Genève, 102 Bd Carl Vogt, CH - 1211 Genève 4, Suisse. scaillet@hec.unige.ch

revised February 2003 (November 2002)

Abstract

We consider a nonparametric method to estimate copulas, i.e. functions linking joint distributions to their univariate margins. We derive the asymptotic properties of kernel estimators of copulas and their derivatives in the context of a multivariate stationary process satisfactory strong mixing conditions. Monte Carlo results are reported for a stationary vector autoregressive process of order one with Gaussian innovations. An empirical illustration containing a comparison with the independent, comotonic and Gaussian copulas is given for European and US stock index returns.

Key words: Nonparametric, Kernel, Time Series, Copulas, Dependence Measures, Risk Management, Loss Severity Distribution.

JEL Classification: C14, D81, G10, G21, G22.

---

[1] We would like to thank Marc Henry for many stimulating discussions and proof checking, as well as the editor and a referee for constructive criticism. We have received fruitful comments from P. Doukhan, H. Hong and O. Renault. The second author receives support by the Swiss National Science Foundation through the National Center of Competence: Financial Valuation and Risk Management. Part of this research was done when he was visiting THEMA and IRES.

1 Introduction

Knowledge of the dependence structure between financial assets or claims is crucial to achieve performant risk management in finance and insurance. Measuring dependence by standard correlation is adequate in the context of multivariate normally distributed risks or for assessing linear dependence. Contemporary financial risk management however calls for other tools due to the presence of an increasing proportion of nonlinear risks (derivative assets) in trading books and the nonnormal behaviour of most financial time series (skewness and leptokurticity). Using estimates of risk dependence via conventional correlation coefficients neglects nonlinearities and leads in most cases to underestimation of the global risk of a portfolio. Furthermore it is now well admitted that the choice of the dependence structure, or similarly of the copula, is also often a key issue for numerous pricing models in finance and insurance. This is especially true concerning the pricing and hedging of credit sensitive instruments such as collateralised debt obligations (CDO) or basket credit derivatives [2].

The copula of a multivariate distribution can be considered as the part describing its dependence structure as opposed to the behaviour of each of its margins. One attractive property of the copula is its invariance under strictly increasing transformation of the margins [3]. In fact, the use of copulas allows solving a difficult problem, namely finding the whole multivariate distribution, by performing two easier tasks. The first step starts by modeling each marginal distribution. The second step consists of estimating a copula, which summarizes all the dependence structure. However this second task is still in its infancy for most of multivariate financial series partly because of the presence of temporal dependencies (serial autocorrelation, time varying heteroskedasticity,...) in returns of stock indices, credit spreads between obligors, interest rates of various maturities...

[2] see e.g. Frey and McNeil (2001), Li (2000).

[3] Note also that scale invariant measures of dependence such as Kendall's tau and Spearman's rho can be expressed by means of copulas. These quantities are often more informative and less misleading than the usual coefficient of correlation. See Embrechts et al. (2002) for a discussion.

Estimation of copulas has essentially been developped in the context of i.i.d. samples. If the true copula is assumed to belong to a parametric family $\mathcal{C} = \{C_\theta, \theta \in \Theta\}$, consistent and asymptotically normally distributed estimates of the parameter of interest can be obtained through maximum likelihood methods. There are mainly two ways to achieve this : a fully parametric method and a semiparametric method. The first method relies on the assumption of parametric marginal distributions. Each parametric margin is then plugged in the full likelihood and this full likelihood is maximized with respects to θ. Alternatively and without parametric assumptions for margins, the marginal empirical cumulative distribution functions can be plugged in the likelihood. These two commonly used methods are detailed in Genest et al. (1993) and Shi and Louis (1995) [4].

Beside these two methods, it is also possible to estimate a copula by some nonparametric methods based on empirical distributions following Deheuvels (1978), (1981a,b) [5]. The so-called empirical copulas resemble usual multivariate empirical cumulative distribution functions. They are highly discontinuous (constant on some data-dependent pavements) and cannot be exploited as graphical device. In fact they are useless to help finding a convenient parametric family of copulas by simple visual comparison on available data sets.

To our best knowledge, nonparametric estimation of copulas in the context of time dependence has not yet been studied theoretically in the literature. Clearly, this is an important omission since most financial series exhibit temporal dependence and copulas are becoming more and more popular among practitioners [6].

[4] see Cebrian, Denuit and Scaillet (2002) for inference under misspecified copulas.

[5] see some asymptotic properties and extensions in Fermanian et al. (2002).

[6] see Bouyé et al. (2000) for a survey of financial applications, Frees and Valdez (1998) for use in actuarial practice, and Patton (2001a,b), Rockinger and Jondeau (2001) for use in modelling conditional dependencies.

In this paper we propose a nonparametric estimation method for copulas for time series, and use a kernel based approach. Such an approach has the advantage to provide a smooth (differentiable) reconstitution of the copula function without putting any particular parametric a priori on the dependence structure between margins and without losing the usual parametric rate of convergence. The approach is developped in the context of multivariate stationary processes satisfying strong mixing conditions [7]. Once estimates of copulas and their derivatives are available (need of differentiability explains our choice of a kernel approach), concepts such as positive quadrant dependence and left tail decreasing behaviour may be empirically analysed. These estimates are also useful to draw simulated data satisfying the dependence structure infered from observations [8]. They are further needed to build asymptotic confidence intervals for our copula estimators. Nonparametric estimators of copulas may also lead to statistics aimed to assess independence between margins. These statistics are in the same spirit as kernel based tools used to test for serial dependence for a univariate stationary time series [9].

The paper is organized as follows. In Section 2 we outline our framework and recall the definition of copulas and some of their properties. In Section 3 we present the kernel estimators of copulas and their derivatives, and characterize their asymptotic behaviour. Their use in estimation of measures of dependence between margins is also briefly described. Section 4 contains some Monte Carlo results for a stationary vector autoregressive process of order one with Gaussian innovations. An empirical illustration on European and US stock index returns is provided in Section 5. Section 6 concludes. All proofs are gathered in an appendix.

[7] Intuitively a process is strong mixing or α-mixing when observations at different dates tend to behave more and more independently when the time interval between dates gets larger and larger. See Doukhan (1994) for relevant definitions and examples in ARMA and GARCH modelling with Gaussian errors.

[8] see Embrechts et al. (2002) for the description of an algorithm.

[9] see the survey of Tjostheim (1996).

2 Framework

We consider a strictly stationary process $\{\mathbf{Y}_t, t \in \mathbb{Z}\}$ taking values in $\mathbb{R}^n$ and assume that our data consist in a realization of $\{\mathbf{Y}_t, t = 1, ..., T\}$. These data may correspond to observed returns of n financial assets, say stock indices, at several dates. They may also correspond to simulated values drawn from a parametric model (VARMA, multivariate GARCH or diffusion processes), possibly fitted on another set of data. Simulations are often required when the structure of financial assets is too complex, as for some derivative products. This, in turn, implies that the sample length T can sometimes be controlled, and asked to be sufficiently large to get satisfying estimation results.

We denote by $f(\mathbf{y})$, $F(\mathbf{y})$, the p.d.f. and c.d.f. of $\mathbf{Y}_t = (Y_{1t}, ..., Y_{nt})'$ at point $\mathbf{y} = (y_1, ..., y_n)'$. The joint distribution F provides complete information concerning the behaviour of $\mathbf{Y}_t$. The idea behind copulas is to separate dependence and marginal behaviour of the elements constituting $\mathbf{Y}_t$. The marginal p.d.f. and c.d.f. of each element Y_{jt} at point y_j, $j = 1, ..., n$, will be written $f_j(y_j)$, and $F_j(y_j)$, respectively. A copula describes how the joint distribution F is "coupled" to its univariate margins F_j, hence its name. Before defining formally a copula and reviewing various useful dependence concepts, we would like to refer the reader to Nelsen (1999) and Joe (1997) for more extensive treatments.

Definition. (Copula)

A n-dimensional copula is a function C with the following properties:

1. $\mathrm{d}om\, C = [0,1]^n$.

2. *C is grounded, i.e. for every $\mathbf{u}$ in $[0,1]^n$, $C(\mathbf{u}) = 0$ if at least one coordinate $u_j = 0$, $j = 1, ..., n$.*

3. *C is n-increasing, i.e. for every $\mathbf{a}$ and $\mathbf{b}$ in $[0,1]^n$ such that $\mathbf{a} \leq \mathbf{b}$, the C-volume $V_C([\mathbf{a}, \mathbf{b}])$ of the box $[\mathbf{a}, \mathbf{b}]$ is positive.*

4. *If all coordinates of $\mathbf{u}$ are 1 except for some u_j, $j = 1, ..., n$, then $C(\mathbf{u}) = u_j$.*

The reason why a copula is useful in revealing the link between the joint distribution and its margins transpires from the following theorem.

Theorem 1. (Sklar's Theorem)
Let F be an n-dimensional distribution function with margins $F_1, ..., F_n$. Then there exists an n-copula C such that for all $\mathbf{y}$ in $\mathbb{R}^n$,

$$F(\mathbf{y}) = C(F_1(y_1), ..., F_n(y_n)). \tag{1}$$

If $F_1, ..., F_n$ are all continuous, then C is uniquely defined. Otherwise, C is uniquely determined on range $F_1 \times ... \times$ range F_n. Conversely, if C is an n-copula and $F_1, ..., F_n$ are distribution functions, then the function F defined by (1) is an n-dimensional distribution function with margins $F_1, ..., F_n$.

As an immediate corollary of Sklar's Theorem, we have

$$C(u_1, ..., u_n) = F(F_1^{-1}(u_1), ..., F_n^{-1}(u_n)), \tag{2}$$

where $F_1^{-1}, ..., F_n^{-1}$ are quasi inverses of $F_1, ..., F_n$, namely

$$F_j^{-1}(u_j) = \inf\{y | F_j(y) \geq u_j\}.$$

Note that, if F_j is strictly increasing, the quasi inverse is the ordinary inverse. Copulas are thus multivariate uniform distributions which describe the dependence structure of random variables. Besides as already mentioned, strictly increasing transformations of the underlying random variables result in the transformed variables having the same copula.

From expression (2), we may observe that the dependence structure embodied by the copula can be recovered from the knowledge of the joint distribution F and its margins F_j. These are the distributional objects that we propose to estimate nonparametrically by a kernel approach in the next section. Before turning our attention to this problem, let us review some relevant uses of copulas.

First copulas characterize independence and comonotonicity between random variables. Indeed, n random variables are independent if and only if $C(\mathbf{u}) = \prod_{j=1}^{n} u_j$, for all $\mathbf{u}$, and each random variable is almost surely a strictly increasing function of any of the others (comonotonicity) if and only if $C(\mathbf{u}) = \min(u_1, ..., u_n)$, for all $\mathbf{u}$. In fact copulas are intimately related to standard measures of dependence between two real valued random variables Y_{1t} and Y_{2t}, whose copula is C. Indeed, the population versions of Kendall's tau, Spearman's rho, Gini's gamma and Blomqvist's beta can be expressed as:

$$
\begin{aligned}
\tau_{Y_1,Y_2} &= 1 - 4\int_0^1 \int_0^1 \frac{\partial C(u_1,u_2)}{\partial u_1}\frac{\partial C(u_1,u_2)}{\partial u_2} du_1 du_2, && (3)\\
\rho_{Y_1,Y_2} &= 12\int_0^1 \int_0^1 C(u_1,u_2) du_1 du_2 - 3, && (4)\\
\gamma_{Y_1,Y_2} &= 4\int_0^1 [C(u_1, 1-u_1) + C(u_1,u_1)] du_1 - 2, && (5)\\
\beta_{Y_1,Y_2} &= 4C(1/2,1/2) - 1. && (6)
\end{aligned}
$$

Second copulas can be used to analyse how two random variables behave together when they are simultaneously small (or large). This will be useful in examining the joint behaviour of small returns, especially the large negative ones (big losses), which are of particular interest in risk management. This type of behaviour is best described by the concept known as positive quadrant dependence after LEHMANN (1966). Two random variables Y_{1t} and Y_{2t} are said to be positively quadrant dependent (PQD) if, for all (y_1, y_2) in $\mathbb{R}^2$,

$$P[Y_{1t} \leq y_1, Y_{2t} \leq y_2] \geq P[Y_{1t} \leq y_1]P[Y_{2t} \leq y_2]. \tag{7}$$

This states that two random variables are PQD if the probability that they are simultaneously small is at least as great as it would be if they were independent. Inequality (7) can be rewritten in terms of the copula C of the two random variables, since (7) is equivalent to the condition $C(u_1, u_2) \geq u_1 u_2$, for all (u_1, u_2) in $[0, 1]^2$.

Finally inequality (7) can be rewritten $P[Y_{1t} \leq y_1 | Y_{2t} \leq y_2] \geq P[Y_{1t} \leq y_1]$ by application of Bayes' rule. The PQD condition may be strengthened by requiring the conditional probability being a non increasing function of y_2. This implies that the probability that the return Y_{1t} takes a small value does not increase as the value taken by the other return increases. It corresponds to particular monotonicities in the tails. We say that a random variable Y_{1t} is left tail decreasing in Y_{2t}, denoted $LTD(Y_1|Y_2)$, if $P[Y_{1t} \leq y_1 | Y_{2t} \leq y_2]$ is a non increasing function of y_2 for all y_1. This in turn is equivalent to the condition that, for all u_1 in $[0,1]$, $C(u_1,u_2)/u_2$ is non increasing in u_2, or $\partial C(u_1,u_2)/\partial u_2 \leq C(u_1,u_2)/u_2$ for almost all u_2.

In short, concepts such as independence, PQD or LTD, may be characterized in terms of copulas, and thus may be checked (see for example the testing procedures developped in Denuit and Scaillet (2002) and Cebrian, Denuit and Scaillet (2002)), once copulas are empirically known. In the next section we develop estimation tools for that purpose.

3 Kernel estimators

We start with the definition of kernel estimators before moving to their asymptotic distributions.

3.1 Definitions

For given $u_j \in (0,1)$, $j=1,\ldots,n$, we assume that the c.d.f. F_j of Y_{jt}, is such that the equation $F_j(y)=u_j$ admits a unique solution denoted $\zeta_j(u_j)$, or more compactly ζ_j (if there is no ambiguity).

To build our estimators we need to introduce kernels, i.e. real bounded symmetric [10] functions k_j on $\mathbb{R}$ such that

$$\int k_j(x)\,dx = 1, \quad j=1,\ldots,n.$$

Let the n-dimensional kernel

$$k(\mathbf{x}) = \prod_{j=1}^{n} k_j(x_j),$$

and its primitive function

$$K(\mathbf{x}) = \prod_{j=1}^{n} \int_{-\infty}^{x_j} k_j = \prod_{j=1}^{n} K_j(x_j).$$

[10] The symmetry of the kernel may induce the so-called boundary bias for data with finite support. Boundary bias is due to weight allocation by the fixed symmetric kernel outside the density support when smoothing is carried out near the boundary. This may happen, for example, when considering smoothing of insurance loss data near the zero boundary.

For the sake of simplicity, we choose to work here with products of univariate kernels. We could however extend easily our results to more general k and K. Let us denote further

$$k(\mathbf{x};h)=\prod_{j=1}^{n}k_j\left(\frac{x_j}{h_j}\right),\quad K(\mathbf{x};h)=\prod_{j=1}^{n}K_j\left(\frac{x_j}{h_j}\right),$$

where the bandwidth h is a diagonal matrix with elements $(h_j)_{j=1}^n$ and determinant $|h|$ (for a scalar x, $|x|$ will denote its absolute value), while the individual bandwidths h_j are positive functions of T such that $h_j \to 0$ when $T \to \infty$. Moreover, we denote by h_* the largest bandwidth among $h_1,\ldots,h_n$. The p.d.f. of Y_{jt} at y_j, i.e. $f_j(y_j)$, will be estimated as usually by

$$\hat{f}_j(y_j)=(Th_j)^{-1}\sum_{t=1}^{T}k_j\left(\frac{y_j-Y_{jt}}{h_j}\right),$$

while the p.d.f. of $\mathbf{Y}_t$ at $\mathbf{y}=(y_1,...,y_n)'$, i.e. $f(\mathbf{y})$, will be estimated by

$$\hat{f}(\mathbf{y})=(T|h|)^{-1}\sum_{t=1}^{T}k(\mathbf{y}-\mathbf{Y}_t;h).$$

Hence, an estimator of the cumulative distribution of Y_{jt} at some point y_j is obtained as

$$\hat{F}_j(y_j)=\int_{-\infty}^{y_j}\hat{f}_j(x)dx, \tag{8}$$

while an estimator of the cumulative distribution of $\mathbf{Y}_t$ at $\mathbf{y}$ is obtained as

$$\hat{F}(\mathbf{y})=\int_{-\infty}^{y_1}...\int_{-\infty}^{y_n}\hat{f}(\mathbf{x})\,d\mathbf{x}. \tag{9}$$

If a single Gaussian kernel $k_j(x)=\varphi(x)$ is adopted, we get

$$\hat{F}_j(y_j)=T^{-1}\sum_{t=1}^{T}\Phi\left((y_j-Y_{jt})/h_j\right),$$

and

$$\hat{F}(\mathbf{y})=T^{-1}\sum_{t=1}^{T}\prod_{j=1}^{n}\Phi\left((y_j-Y_{jt})/h_j\right),$$

where φ and Φ denote the p.d.f. and c.d.f. of a standard Gaussian variable, respectively.

In order to estimate the copula at some point $\mathbf{u}$, we use a simple plug-in method, and exploits directly expression (2):

$$\hat{C}(\mathbf{u}) = \hat{F}(\hat{\boldsymbol{\zeta}}), \tag{10}$$

where $\hat{\boldsymbol{\zeta}} = (\hat{\zeta}_1, ..., \hat{\zeta}_n)'$ and $\hat{\zeta}_j = \inf_{y\in\mathbb{R}}\{y : \hat{F}_j(y) \geq u_j\}$. In fact $\hat{\zeta}_j$ corresponds to a kernel estimate of the quantile of Y_{jt} with probability level u_j [11].

3.2 Asymptotic distributions

The asymptotic normality of kernel estimators for copulas can be established under suitable conditions on the kernel, the asymptotic behaviour of the bandwidth, the regularity of the densities, and some mixing properties of the process.

Assumption 1. (kernel and bandwidth)
(a) Bandwidths satisfy $Th_*^2 \to 0$, *or*
(a') Bandwidths satisfy $Th_*^4 \to 0$ *and the kernel* k *is even,*
(b) The kernel k *has a compact support.*

Assumption 1 (b) could in fact be weakened, by controlling the tails of k, for instance by assuming $\sup_j |k_j(x)| \leq (1+|x|)^{-\alpha}$ for every x and some $\alpha > 0$, as in ROBINSON (1983). This type of assumption is satisfied by most kernels, in particular by the Gaussian kernel.

Assumption 2. (process)
(a) The process $(\mathbf{Y}_t)$ *is strong mixing with coefficients* α_t *such that* $\alpha_T = o(T^{-a})$ *for some* $a > 1$, *as* $T \to \infty$.
(b) The marginal c.d.f. F_j, $j = 1, \ldots, n$ *are continuously differentiable on the intervals* $[F_j^{-1}(a) - \varepsilon, F_j^{-1}(b) + \varepsilon]$ *for every* $0 < a < b < 1$ *and some* $\varepsilon > 0$, *with positive derivatives* f_j. *Moreover, the first partial derivatives of* F *exist and are Lipschitz continuous on the product of these intervals.*

[11] see e.g. GOURIÉROUX et al. (2000) and SCAILLET (2000) for use of this type of smooth quantile estimates in risk management and portfolio selection as well as AZZALINI (1981) for the asymptotic properties in the i.i.d. case.

The asymptotic behaviour of $\hat{C}$ is related to the limit in distribution of $T^{1/2}(\hat{F} - F)$, the smoothed empirical process associated with the sequence of $\mathbb{R}^n$-valued vectors $(\mathbf{Y}_t)_{t \geq 1}$. We first state the limiting behaviour of this smoothed empirical process before giving the limiting behaviour of the smoothed copula process.

Theorem 2. (Smoothed empirical process)
Under Assumptions 1 and 2, the smoothed empirical process $T^{1/2}(\hat{F} - F)$ tends weakly to a centered Gaussian process $\mathbb{G}$ in $l^\infty(\mathbb{R}^n)$ (the space of a.s. bounded functions on $\mathbb{R}^n$), endowed with the sup-norm. The covariance function of $\mathbb{G}$ is

$$Cov(\mathbb{G}(\mathbf{x}), \mathbb{G}(\mathbf{y})) = \sum_{t \in \mathbb{Z}} Cov(\mathbf{1}\{\mathbf{Y}_0 \leq \mathbf{x}\}, \mathbf{1}\{\mathbf{Y}_t \leq \mathbf{y}\}). \tag{11}$$

Moreover, $T^{1/2} \sup_{\mathbf{x}} |(\hat{F} - F)(\mathbf{x})| = o_P(1)$.

Theorem 3. (Smoothed copula process)
Under Assumptions 1 and 2, the process $T^{1/2}(\hat{C} - C)$ tends weakly to a centered Gaussian process $\phi'(\mathbb{G})$ in $l^\infty([0,1]^n)$ endowed with the sup-norm, where the limiting process is given by

$$\begin{aligned} \phi'(\mathbb{G})(u_1, \ldots, u_n) &= \mathbb{G}(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)) \\ &\quad - \sum_{j=1}^n \frac{\partial C}{\partial u_j}(u_1, \ldots, u_n) \mathbb{G}(+\infty, \ldots, F_j^{-1}(u_j), \ldots, +\infty). \end{aligned}$$

A direct consequence of Theorem 3 is the following asymptotic normality result.

Corollary 1. (Joint normality of copula estimators)
Under Assumptions 1 and 2, for any $(\mathbf{v}_1, \ldots, \mathbf{v}_d)$ in $]0,1]^{nd}$, the d-dimensional random vector

$$\mathbb{S} \equiv T^{1/2} \left((\hat{C} - C)(\mathbf{v}_1), \ldots, (\hat{C} - C)(\mathbf{v}_d) \right)$$

tends weakly to a centered Gaussian vector.

It is possible to derive an explicit form of the asymptotic covariance matrix of the vector $\mathbb{S}$ after some tedious computations (see Equation (16) at the end of the appendix). This covariance matrix will be used in the empirical section of this paper to build confidence intervals around copula estimates.

In the bivariate case $\mathbf{Y}_t = (Y_{1t}, Y_{2t})'$, we have seen that positive quadrant dependence is characterized by $C(u_1, u_2) - u_1 u_2 \geq 0$, while left tail decreasing behaviour of Y_{1t} (resp. Y_{2t}) in Y_{2t} (resp. Y_{1t}) is characterized by $C(u_1, u_2)/u_2 - \partial C(u_1, u_2)/\partial u_2 \geq 0$ (resp. $C(u_1, u_2)/u_1 - \partial C(u_1, u_2)/\partial u_1 \geq 0$). We have just developed a kernel estimator for C. It is thus natural to suggest an estimator for $\partial_p C(\mathbf{u}) = \partial C(\mathbf{u})/\partial u_p$ based on the differentiation of $\hat{C}(\mathbf{u})$ w.r.t. u_p:

$$\widehat{\partial_p C(\mathbf{u})} = \partial_p \left(\hat{C}(\mathbf{u}) \right) = \frac{\partial_p \hat{F}(\hat{\boldsymbol{\zeta}})}{\hat{f}_p(\hat{\zeta}_p)},$$

with

$$\hat{\boldsymbol{\zeta}} = (\hat{\zeta}_1, \ldots, \hat{\zeta}_n), \qquad \hat{\zeta}_j = \hat{F}_j^{-1}(u_j),\ j = 1, \ldots, n,$$

and

$$\partial_p \hat{F}(\hat{\boldsymbol{\zeta}}) = \frac{1}{T h_p} \sum_{t=1}^{T} k_p \left(\frac{\hat{\zeta}_p - Y_{pt}}{h_p} \right) \prod_{l \neq p} K_l \left(\frac{\hat{\zeta}_l - Y_{lt}}{h_l} \right).$$

The estimators $\hat{C}(\mathbf{u})$ and $\partial_p \hat{C}(\mathbf{u})$ will help to detect positive quadrant dependence and left tail decreasing behaviour through the empirical counterparts of the aforementioned inequalities.

Assumption 3. (kernel and bandwidth)

(a) Bandwidths satisfy $Th_p \to +\infty$, $Th_*^3 \to 0$ and k is even,

(b) Each kernel k_j has a compact support A_j, $j = 1, \ldots, n$, and the kernel k_p is twice continuously differentiable.

Assumption 4. (process)

(a) The process $(\mathbf{Y}_t)$ is strong mixing with coefficients α_t such that $\alpha_T = O(T^{-2})$, as $T \to \infty$.

(b) The first derivative $\partial_p F$ and the marginal density f_p are Lipschitz continuous.

(c) Let $f_{t,t'}$ be the density function of $(\mathbf{Y}_t, \mathbf{Y}_{t'})$ with respect to the Lebesgue measure. There exists an integrable function $\psi : \mathbb{R}^{2n-2} \longrightarrow \mathbb{R}$ such that, for every vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^n$, and

for some open subset $\mathcal{O}$ of $\mathbb{R}^2$ containing $A_p \times A_p$, we have

$$\sup_{t \neq t'} \sup_{(u_p, v_p) \in \mathcal{O}} f_{t,t'}(\mathbf{u}, \mathbf{v}) \leq \psi(\mathbf{u}_{-p}, \mathbf{v}_{-p}).$$

Then we get:

Theorem 4. (Joint normality of derivative estimators)
Under Assumptions 3 and 4, for any $\mathbf{u}_1, \ldots, \mathbf{u}_d \in]0,1[^{nd}$, the random vector

$$(Th_p)^{1/2} \left((\partial_p \hat{C} - \partial_p C)(\mathbf{u}_1), \ldots, (\partial_p \hat{C} - \partial_p C)(\mathbf{u}_d) \right)$$

tends weakly to a centered Gaussian vector.

Again the asymptotic covariance matrix $\Sigma^* = [\sigma^*_{ij}]_{1 \leq i,j \leq d}$ of this random vector admits a rather complex explicit form (see Equation (14) in the appendix).

Note that the extension of the previous propositions to higher derivatives $\partial^k \hat{C}(\mathbf{u})/(\partial u_1 \partial u_2 ... \partial u_k)$, $k \leq n$, is straightforward. Such estimates are for example required for the implementation of an empirical counterpart of the simulation algorithm described in EMBRECHTS et al. (2002).

As clear from Theorem 4, a random vector of derivatives of $T^{1/2}(\hat{C} - C)$ at some points $\mathbf{u}_1, \ldots, \mathbf{u}_d$ does not in general tend weakly to a vector of independent Gaussian variables. This is however the case when the components corresponding to the indices of the derivatives are all different. A similar result is also true for successive derivatives. Indeed, under some technical assumptions, the random vector

$$(Th_1^{m_1} \cdot \ldots \cdot h_n^{m_n})^{1/2} \left(\partial_1^{m_1} \ldots \partial_n^{m_n} (\hat{C} - C)(\mathbf{u}_1), \ldots, \partial_1^{m_1} \ldots \partial_n^{m_n} (\hat{C} - C)(\mathbf{u}_d) \right)$$

tends weakly to a centered Gaussian vector whose covariance matrix is diagonal if

1. $m_l \geq 1$ for every $l = 1, \ldots, n$, or

2. for the indices l such that $m_l \neq 0$, $u_{li} \neq u_{lj}$ for every pair (i, j).

We end up this section with the description of how to build the sample analogues of the dependence measures (3)-(6). First we may substitute a kernel estimator $\hat{C}$ for the unknown C in (4)-(6) to estimate ρ_{Y_1,Y_2}, γ_{Y_1,Y_2}, and β_{Y_1,Y_2}. According to Theorem 3, these estimators are consistent and asymptotically normally distributed. Second we may replace the unkown derivatives by $\partial\hat{C}(u_1,u_2)/\partial u_1$ and $\partial\hat{C}(u_1,u_2)/\partial u_2$ in order to estimate τ_{Y_1,Y_2}. This estimator will be consistent by Theorem 4.

4 Monte Carlo experiments

In this section we wish to investigate the finite sample properties of our kernel estimators. The experiments are based on a stationary vector autoregressive process of order one:

$$Y_t = A + BY_{t-1} + \nu_t,$$

where $\nu_t \sim N(0,\Sigma)$ and all the eigenvalues of B are less than one in absolute value. The marginal distribution of the process Y is Gaussian with mean $\mu = (Id - B)^{-1}A$ and covariance matrix satisfying vec $= (Id - B \quad B)^{-1}\mathrm{vec}\,\Sigma$ ($\mathrm{vec}\,M$ is the stack of the columns of the matrix M).

We start with a bivariate example where the components Y_{1t} and Y_{2t} are independent, and thus $C(u_1,u_2) = u_1u_2$. The parameters are $A = (1,1)'$, $\mathrm{vec}\,B = (0.25,0,0,0.75)'$ and $\mathrm{vec}\,\Sigma = (.75,0,0,1.25)'$. The parameters of the Gaussian stationary distribution are then $\mu = (1.33,4)'$ and vec $= (0.8,0,0,2.86)'$. The number of Monte Carlo replications is 5,000, while the data length is $T = 4^5 = 1024$ (roughly four trading years of quotes).

The nonparametric estimators make use of the product of two Gaussian kernels. Bandwidth values are based on the rule of thumb $h_i = \hat{\sigma}_i T^{-1/5}$, which uses the empirical standard deviation of each series. Table 1 gives bias and mean squared error (MSE) [12] of the kernel estimates $\hat{C}(u_1,u_2)$ for several pairs (u_1,u_2) in the tails and center of the distribution.

[12] Recall that the bias is defined as $E\hat{C} - C$ and the MSE as $E[(\hat{C} - C)^2]$.

Table 1: Bias and MSE of kernel estimates of copulas

Bias and MSE have been computed on 5000 random samples of length 1024 from a bivariate autoregressive process of order one with independent components.
All figures (true value of the Gaussian copula, bias and MSE) are expressed as percents of percents (10^{-4}).

$\times 10^{-4}$	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.5,.5)	C(.75,.75)	C(.95,.95)	C(.99,.99)
True	1.00	25.00	625.00	2500.00	5625.00	9025.00	9801.00
Bias	-0.09	-0.08	0.40	1.12	-0.90	-0.04	4.66
MSE	0.00	0.01	0.25	0.48	0.25	0.01	0.05

We may observe that the results are satisfactory both in terms of bias and MSE. The pairs located in the tails of the distribution are well estimated. The pairs $C(.01, .01)$, $C(.05, .05)$, in the left tail are of particular interest in risk management since they measure the dependence of joint extreme losses.

We continue with a bivariate example where the components Y_{1t} and Y_{2t} are dependent. The parameters are $A = (1, 1)'$, $\mathrm{vec}\, B = (0.25, 0.2, 0.2, 0.75)'$ and $\mathrm{vec}\, \Sigma = (.75, .5, .5, 1.25)'$. The parameters of the Gaussian stationary distribution are then $\mu = (3.05, 6.44)'$ and vec $\quad = (1.13, 1.49, 1.49, 3.98)'$ (correlation of 0.70). Since Y_{1t} and Y_{2t} are positively dependent, we get $C(u_1, u_2) > u_1 u_2$. Results are reported in Table 2. Bias and MSE are higher when compared with the previous independent case. They are however still satisfactory.

Table 2: Bias and MSE of kernel estimates of copulas

Bias and MSE have been computed on 5000 random samples of length 1024 from a bivariate autoregressive process of order one with dependent components.
All figures (true value of the Gaussian copula, bias and MSE) are expressed as percents of percents (10^{-4}).

$\times 10^{-4}$	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.5,.5)	C(.75,.75)	C(.95,.95)	C(.99,.99)
True	27.08	197.95	1511.74	3747.68	6511.74	9197.95	9827.08
Bias	-7.47	-34.88	-130.32	-172.28	-130.53	-35.25	-7.65
MSE	0.01	0.18	1.98	3.36	1.99	0.18	0.01

5 Empirical illustrations

This section illustrates the implementation of the estimation procedure described in Section 3. We analyse return data on two pairs of major stock indices: CAC40-DAX35, and S&P500-DJI. The data are one day returns recorded daily from 03/01/1994 to 07/07/2000, i.e. 1700 observations. We also report results for returns computed on a daily basis (rolling returns) with a holding period of ten days instead of one day. This holding period is favoured by the Basle Committee on Banking Supervision when quantifying trading risks. Note that our asymptotic results derived in a dependent data framework cover the high degree of autocorrelation that such a computation induces. We have selected bandwidth values according to the usual rule of thumb (empirical standard deviation over the sample length at the power one fifth), and used a Gaussian kernel [13].

We start with the pair of European indices CAC40-DAX35. The linear correlation coefficient between the two indices is 67.03%. Figure 1 reports plots of observed returns and rank statistics divided by the number of observations. The left column corresponds to one day returns while the right column corresponds to ten day returns. These plots show that the dependence is more pronounced for the ten day returns. Indeed we have more points situated on the diagonal for the ten day returns and their associated rank statistics.

[13] Other choices of kernels have not been found, as usually in nonparametric estimation, to affect the empirical results substantially.

– Please insert Figure 1 –

Figure 1 : Returns and rank statistics for CAC40-DAX35

One day and ten day returns of the pair CAC40-DAX35 between 03/01/1994 and 07/07/2000 are plotted on the first line. The second line shows the associated rank statistics divided by the number of observations.

Figure 2 gives 3D and contour plots of estimated copulas. The contour plot for the ten day returns is closer to the shape given by the comonotonic copula, namely successive straight lines at right angles on the diagonal, which also indicates higher dependence for the ten day returns.

– Please insert Figure 2 –

Figure 2 : Copula estimates and contour plots for CAC40-DAX35

Estimated copulas of one day and ten day returns of the pair CAC40-DAX35 between 03/01/1994 and 07/07/2000 are plotted on the first line. The second line shows the associated contour plots.

In Figure 3 we use estimates of the copulas to analyse positive quadrant dependence (PQD). The first line of graphs shows that $C(u_1, u_2) - u_1 u_2$ is greater than zero, which means that one day and ten day returns exhibit PQD. The difference is larger in the center of the distribution and decreases when we move to the extremes. Just below we find comparison w.r.t. the comonotonic copula, i.e. $\min(u_1, u_2) - C(u_1, u_2)$, and the Gaussian copula, i.e. $C(u_1, u_2) - C_{Gau}(u_1, u_2; \rho^*)$. The estimate $\hat{\rho}^*$ of the parameter ρ^* of the Gaussian copula is obtained using the equation $\rho^* = 2\sin(\pi\rho/6)$ linking ρ^* with the rank correlation ρ [14]. The hat of the ten day returns is lower than the hat of the one day

[14] Estimates obtained from the empirical rank correlation or its smoothed counterpart lead to virtually identical results.

returns for comotonicity. This again indicates a higher dependence for the former than for the latter. Interestingly the last line illustrates how smoothed copula estimates can be used as graphical device to detect adequacy of parametric copula models. Indeed we may observe that the Gaussian copula exhibits too low levels for small u_1, u_2 and large u_1, u_2. In the center of the distribution this is the reverse.

– Please insert Figure 3 –

Figure 3 : Comparison with independent, comonotonic and Gaussian copulas for CAC40-DAX35

The graphs successively compare nonparametric copula estimates with the independent copula, the comonotonic copula and the Gaussian copula for the one day and ten day returns of the pair CAC40-DAX35 between 03/01/1994 and 07/07/2000.

In Figure 4 copula derivatives are computed to study left tail decreasing behaviour (LTD). The two lines carry graphs of $C(u_1, u_2)/u_2 - \partial C(u_1, u_2)/\partial u_2$, and $C(u_1, u_2)/u_1 - \partial C(u_1, u_2)/\partial u_1$, respectively. Again we get positiveness, and LTD is thus present for both stock indices and both holding periods. Besides, LTD is heavier for the ten day holding period.

– Please insert Figure 4 –

Figure 4 : Left tail decreasing behaviour for CAC40-DAX35

The graphs show the left tail decreasing behaviour of the one day and ten day returns of the pair CAC40-DAX35 between 03/01/1994 and 07/07/2000.

For the pair of US indices S&P500-DJI, we get Figures 5-8. The linear correlation coefficient is equal to 93.25%. Comments made for European indices carry over. However the dependence is higher for US indices. Indeed we get further clustering around the diagonal in Figure 5, even closer contour plots to the comonotonic plot in Figure 6, and higher hats for positive quadrant dependence and comonotonicity in Figure 7.

– Please insert Figure 5 –

Figure 5 : Returns and rank statistics for S&P500-DJI

One day and ten day returns of the pair S&P500-DJI between 03/01/1994 and 07/07/2000 are plotted on the first line. The second line shows the associated order statistics divided by the number of observations.

– Please insert Figure 6 –

Figure 6 : Copula estimates and contour plots for S&P500-DJI

Estimated copulas of one day and ten day returns of the pair S&P500-DJI between 03/01/1994 and 07/07/2000 are plotted on the first line. The second line shows the associated contour plots.

– Please insert Figure 7 –

Figure 7 : Comparison with independent, comonotonic and Gaussian copulas for S&P500-DJI

The graphs successively compare nonparametric copula estimates with the independent copula, the comonotonic copula and the Gaussian copula for the one day and ten day returns of the pair S&P500-DJI between 03/01/1994 and 07/07/2000.

– Please insert Figure 8 –

Figure 8 : Left tail decreasing behaviour for S&P500-DJI

The graphs show the left tail decreasing behaviour of the one day and ten day returns of the pair S&P500-DJI between 03/01/1994 and 07/07/2000.

As already mentioned in the comments of Figures 3 and 7 the Gaussian copula tends to underestimate risk dependencies. One may then wonder what could be the impact of using a Gaussian copula in computing a risk measure like the Value at Risk (VaR). Recall that VaR for a two asset portfolio is implicitly defined through the equation:

$$P[-a_1 Y_{1t} - a_2 Y_{2t} > VaR(a_1, a_2; p)] = p, \tag{12}$$

where (a_1, a_2) is the portfolio allocation in percentage and p is a small probability level, say 1%. An empirical counterpart of Equation (12) under the assumption of a Gaussian copula and margins estimated by the corresponding individual empirical cumulative distribution functions is simply:

$$\sum_{t=2}^{T} \sum_{t'=2}^{T} \mathbf{1}\{-a_1 Y_{1(t)} - a_2 Y_{2(t')} > VaR(a_1, a_2; p)\} \Delta H_{Gau}(Y_{1(t)}, Y_{2(t')}; \hat{\rho}^*) = p,$$

where $Y_{1(t)}$, $Y_{2(t')}$ denote order statistics and

$$\begin{aligned} \Delta H_{Gau}(Y_{1(t)}, Y_{2(t')}; \hat{\rho}^*) &= C_{Gau}(\frac{t}{T}, \frac{t'}{T}; \hat{\rho}^*) - C_{Gau}(\frac{t-1}{T}, \frac{t'}{T}; \hat{\rho}^*) \\ &\quad - C_{Gau}(\frac{t}{T}, \frac{t'-1}{T}; \hat{\rho}^*) + C_{Gau}(\frac{t-1}{T}, \frac{t'-1}{T}; \hat{\rho}^*). \end{aligned}$$

Tables 3 and 4 compare empirical VaR, i.e. empirical quantiles [15] of the distribution of the portfolio losses $-a_1 Y_{1t} - a_2 Y_{2t}$, and VaR obtained under a Gaussian copula specification. We have considered an equally weighted portfolio, i.e. $a_1 = a_2 = 50\%$, and $p = 1\%$. Clearly the underestimation of risk dependencies by the Gaussian copula yields an underestimation of the VaR portfolio.

[15] Since smoothed quantiles are close to empirical ones we prefer to use here the empirical quantiles for our point estimates. They are easier to implement and faster to compute.

Table 3: Comparison of VaR estimates for CAC40-DAX35

VaR are computed for an equally weighted portfolio and a 99% loss probability level. VaR estimates correspond to empirical VaR and VaR obtained under a Gaussian copula specification. They are computed for the one day and ten day returns of the pair CAC40-DAX35 between 03/01/1994 and 07/07/2000.

	Emp (1 day)	Gau (1 day)	Emp (10 day)	Gau (10 day)
VaR	3.10%	2.75%	9.39%	8.53%

Table 4: Comparison of VaR estimates for S&P500-DJI

VaR are computed for an equally weighted portfolio and a 99% loss probability level. VaR estimates correspond to empirical VaR and VaR obtained under a Gaussian copula specification. They are computed for the one day and ten day returns of the pair S&P500-DJI between 03/01/1994 and 07/07/2000.

	Emp (1 day)	Gau (1 day)	Emp (10 day)	Gau (10 day)
VaR	2.53%	2.42%	7.04%	6.94%

Finally the two following tables deliver 90% confidence intervals based on the asymptotic normality result of Corollary 1 for copula estimators. Since the asymptotic variance involves an infinite number of lags (see Equation (16)) it is necessary to truncate after some point. We have chosen to keep 36 positive and 36 negative lags after having checked stability of variance estimates. Copula derivatives appearing in the asymptotic variance have been estimated with the estimators of Theorem 4, and covariances between indicator functions with their empirical average counterparts.

Table 5: Confidence intervals of copula estimates for CAC40-DAX35

Confidence intervals are based on a 90% level. All figures (copula estimate, upper bound and lower bound) are expressed as percents of percents (10^{-4}) and are given for the one day and ten day returns of the pair CAC40-DAX35 between 03/01/1994 and 07/07/2000.

$\times 10^{-4}$	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.5,.5)	C(.75,.75)	C(.95,.95)	C(.99,.99)
$\hat{C}$ (1 day)	25.71	202.51	1321.46	3427.40	6455.62	9248.57	9856.22
up (1 day)	40.45	259.78	1422.07	3556.42	6569.22	9310.66	9874.62
low (1 day)	10.97	145.23	1220.85	3298.38	6342.01	9186.48	9837.81
$\hat{C}$ (10 day)	43.99	186.88	1595.76	3843.69	6609.97	9338.08	9937.30
up (10 day)	70.28	247.10	1730.37	3992.00	6783.05	9425.27	9942.20
low (10 day)	17.71	126.66	1461.14	3695.37	6436.90	9250.90	9932.40

Table 6: Confidence intervals of copula estimates for S&P500-DJI

Confidence intervals are based on a 90% level. All figures (copula estimate, upper bound and lower bound) are expressed as percents of percents (10^{-4}) and are given for the one day and ten day returns of the pair S&P500-DJI between 03/01/1994 and 07/07/2000.

$\times 10^{-4}$	C(.01,.01)	C(.05,.05)	C(.25,.25)	C(.5,.5)	C(.75,.75)	C(.95,.95)	C(.99,.99)
$\hat{C}$ (1 day)	55.47	341.86	1859.67	4026.27	6839.94	9364.49	9879.73
up (1 day)	72.73	372.15	1946.85	4109.70	6904.53	9399.24	9893.02
low (1 day)	38.19	311.57	1772.49	3942.85	6775.35	9329.74	9866.45
$\hat{C}$ (10 day)	70.22	292.26	1781.33	4163.94	6910.56	9405.54	9952.28
up (10 day)	104.58	334.56	1894.72	4277.41	7014.44	9454.77	9958.42
low (10 day)	35.87	249.96	1667.95	4050.48	6806.68	9356.31	9946.13

6 Concluding remarks

In this paper we have proposed simple nonparametric estimation methods of copulas and their derivatives. The procedure relies on a kernel approach in the context of general stationary strong mixing multivariate processes, which provides smooth differentiable estimators. These estimators have proven to be empirically relevant to the analysis of dependencies among stock index returns. In particular they reveal the different types of dependence structures present in these data. Hence they complement ideally the existing battery of nonparametric tools by providing specific instruments dedicated to dependence measurement and joint risk analysis. They should also help to design goodness-of-fit tests for copulas. This is under current research.

APPENDIX

Proof of Theorem 2

Let us denote by F_T the empirical c.d.f. associated with $(\mathbf{Y}_t)_{t\geq 1}$, say

$$F_T(\mathbf{y}) = T^{-1}\sum_{t=1}^{T} \mathbf{1}\{\mathbf{Y}_t \leq \mathbf{y}\}.$$

We have for every $\mathbf{y}$ in $\mathbb{R}^n$

$$\hat{F}(\mathbf{y}) = \int K(\mathbf{y}-\mathbf{v};h)\, F_T(d\mathbf{v}) = \int F_T(\mathbf{y}-h\cdot\mathbf{v})\, k(\mathbf{v})\, d\mathbf{v},$$

by a n-dimensional integration by parts. The "dot" in $h\cdot\mathbf{v}$ denotes the componentwise product, i.e. $\mathbf{y}-h\cdot\mathbf{v}$ corresponds to the vector $(y_1-h_1v_1,\ldots,y_n-h_nv_n)$. Moreover,

$$T^{1/2}(\hat{F}-F)(\mathbf{y}) - T^{1/2}(F_T-F)(\mathbf{y}) = T^{1/2}\int [(F_T-F)(\mathbf{y}-h\cdot\mathbf{v})$$
$$- \quad (F_T-F)(\mathbf{y})]\; k(\mathbf{v})\, d\mathbf{v} + T^{1/2}\int [F(\mathbf{y}-h\cdot\mathbf{v}) - F(\mathbf{y})]\; k(\mathbf{v})\, d\mathbf{v} \equiv A_1 + A_2.$$

First, the equicontinuity of the process $T^{1/2}(F_T-F)$ is a consequence of its weak convergence (see Rio (2000)) :

Theorem 5. *Let $(\mathbf{Y}_t)_{t\geq 1}$ a stationary sequence in $\mathbb{R}^n$. If each marginal c.d.f. F_j, $j=1,\ldots,n$ is continuous, if the process is α-mixing and $\alpha_T = O(T^{-a})$ for some $a>1$, then there exists a Gaussian process $\mathbb{G}$, whose trajectories are uniformly continuous a.e. on $\mathbb{R}^n$ endowed with the pseudo-metric*

$$d(\mathbf{x},\mathbf{y}) = \sup_{i=1,\ldots,n} |F_i(x_i) - F_i(y_i)|,$$

and such that the empirical process $T^{1/2}(F_T-F)$ tends weakly to $\mathbb{G}$ in $l^\infty(\mathbb{R}^n)$. The covariance structure of the limiting process is given by Equation (11).

Thus, under such assumptions, and since k has a compact support, $\sup_{\mathbf{y}} |A_1| = o_P(1)$. Second, since F is Lipschitz continuous, the second term A_2 is $O(T^{1/2}h_*)$, or further $O(T^{1/2}h_*^2)$ if k is even. Thus, we have proved that

$$\sup_{\mathbf{y}\in\mathbb{R}^n} |T^{1/2}(\hat{F}-F)(\mathbf{y}) - T^{1/2}(F_T-F)(\mathbf{y})| = o_P(1), \tag{13}$$

hence the stated result. Note that we have proved Theorem 2 using the fact that F is Lipschitz continuous which is weaker than Assumption 2. $\square$

Proof of Theorem 3

By the functional Delta-Method, we deduce the weak convergence of $T^{1/2}(\hat{C} - C)$ in $l^\infty([a,b]^n)$, for every a, b, $0 < a < b < 1$, exactly as in Van der Vaart and Wellner (1996, p. 389), and we obtain the convergence of the finite dimensional distributions. Note that $T^{1/2}(\hat{C} - C)(\mathbf{u})$ is zero when one component of $\mathbf{u}$ is zero. Moreover, when one component of $\mathbf{u}$ is 1, say the j-th component, then $T^{1/2}(\hat{C} - C)(\mathbf{u})$ tends to a Gaussian random variable that would be obtained if we had forgotten all the j-th components (in other words, as if the observed random variables had been $(Y_{kt})_{k \neq j}$). Let us finally remark that the weak convergence of $T^{1/2}(\hat{C} - C)$ can be proved in $l^\infty([0,1]^n)$. This can be done by exploiting the proximity between $\hat{F}$ and F_T provided by Theorem 2, and by mimicking the proof of Theorem 10 in Fermanian et al. (2002). $\square$

Proof of Theorem 4

Let us remark that the Delta-Method provides us the limiting behaviour of the smoothed empirical quantiles.

Theorem 6. *Under Assumptions 1 and 2, for every $u_j \in]0,1[$ and every $j \in \{1, \ldots, n\}$, $\hat{\zeta}_j$ is a consistent estimator of $\zeta_j \equiv F_j^{-1}(u_j)$. Moreover,*

$$T^{1/2}(\hat{\zeta}_j - \zeta_j) \stackrel{law}{\longrightarrow} \mathcal{N}(0, \sigma^2(u_j)),$$

where

$$\sigma^2(u_j) = \frac{\sum_{t \in \mathbb{Z}} Cov(\mathbf{1}\{Y_{j0} \leq u_j\}, \mathbf{1}\{Y_{jt} \leq u_j\})}{f_j^2(\zeta_j)}.$$

In particular, each $\hat{\zeta}_j$ tends to ζ_j at the parametric rate $T^{-1/2}$. Note that Assumptions 3 and 4 imply Assumptions 1 and 2. We prove in the next theorem that the quantities $\partial_p \hat{F}$ and $\hat{f}_p$ converge at the slower rate $(Th_p)^{-1/2}$. Thus, it will be convenient to replace each random quantity $\hat{\zeta}_j$ by its limit ζ_j in $\partial_p \hat{F}(\hat{\boldsymbol{\zeta}})$ and $\hat{f}_p(\hat{\zeta}_p)$. The proof of the following asymptotic result will be given later in the appendix.

Theorem 7. *Under Assumptions 3 and 4, the random vector*

$$(Th_p)^{1/2}\left((\partial_p \hat{F} - \partial_p F)(\mathbf{y}_1), \ldots, (\partial_p \hat{F} - \partial_p F)(\mathbf{y}_d), (\hat{f}_p - f_p)(y_{p1}), \ldots, (\hat{f}_p - f_p)(y_{pd})\right)$$

tends weakly to a centered Gaussian vector whose covariance matrix $\Sigma = (\sigma_{i,j})_{1\leq i,j\leq 2d}$, *is characterized by*

$$\begin{aligned}
\sigma_{i,j} = \sigma_{j,i} = \partial_p F(\mathbf{y}_i \wedge \mathbf{y}_j) \int k_p^2, &\qquad \text{if } y_{pi} = y_{pj}, \\
\sigma_{i+d,j+d} = \sigma_{j+d,i+d} = f_p(y_{pi}) \int k_p^2, &\qquad \text{if } y_{pi} = y_{pj}, \\
\sigma_{i,j+d} = \sigma_{j+d,i} = \partial_p F(\mathbf{y}_i) \int k_p^2, &\qquad \text{if } y_{pi} = y_{pj}, \\
\sigma_{i,j} = \sigma_{i+d,j} = \sigma_{i,j+d} = \sigma_{i+d,j+d} = 0, &\qquad \text{otherwise},
\end{aligned}$$

for every i, j *in* $\{1, \ldots, d\}$.

We have denoted $\mathbf{y}_i \wedge \mathbf{y}_j$ the minimum of $\mathbf{y}_i$ and $\mathbf{y}_j$ componentwise, say $(\min(y_{1i}, y_{1j}), \ldots, \min(y_{ni}, y_{nj}))$. Let us now turn to the initial problem, namely the limit in law of

$$(Th_p)^{1/2}\left((\partial_p \hat{C} - \partial_p C)(\mathbf{u}_1), \ldots, (\partial_p \hat{C} - \partial_p C)(\mathbf{u}_d)\right),$$

where $(\mathbf{u}_1, \ldots, \mathbf{u}_d)$ is some point in $]0, 1[^{nd}$. Since every $\hat{\zeta}_j$ tends to $\zeta_j \equiv F_j^{-1}(u_j)$ at the rate $T^{-1/2}$, and since $\hat{f}_p(\zeta_p)$ tends to $f_p(\zeta_p)$ at the slower rate $(Th_p)^{-1/2}$, the asymptotic behaviour of $\hat{f}_p(\hat{\zeta}_p)$ will be the same as the one of $\hat{f}_p(\zeta_p)$. We however still need a continuity argument. If k_p is twice continuously differentiable, k_p'' being bounded,

$$\begin{aligned}
\hat{f}_p(\hat{\zeta}_p) &= \hat{f}_p(\zeta_p) + \hat{f}_p'(\zeta_p).(\hat{\zeta}_p - \zeta_p) + \frac{(\hat{\zeta}_p - \zeta_p)^2}{2Th_p^3} \sum_{t=1}^{T} k_p''\left(\frac{\zeta_p^* - Y_{pt}}{h_p}\right) \\
&= \hat{f}_p(\zeta_p) + \hat{f}_p'(\zeta_p).(\hat{\zeta}_p - \zeta_p) + \frac{(\hat{\zeta}_p - \zeta_p)^2}{2Th_p^3} O(T),
\end{aligned}$$

where $|\zeta_p^* - \zeta_p| \leq |\hat{\zeta}_p - \zeta_p|$. It can be proved easily that $\hat{f}_p'(\zeta_p)$ is $O_P(1)$ (for instance by use of an exponential inequality, or even the Markov inequality). Thus,

$$\begin{aligned}
&(Th_p)^{1/2}(\hat{f}_p(\hat{\zeta}_p) - f_p(\zeta_p)) = (Th_p)^{1/2}(\hat{f}_p - f_p)(\zeta_p) + O_P((Th_p)^{1/2} T^{-1/2}) \\
&+ \; O_P((Th_p)^{1/2} T^{-1} h_p^{-3}) = (Th_p)^{1/2}(\hat{f}_p - f_p)(\zeta_p) + O_P(h_p^{1/2}) + O_P(T^{-1/2} h_p^{-5/2}).
\end{aligned}$$

This quantity tends to zero when Th_p^5 tends to zero. Exactly in the same way, we get under the same assumptions:

$$(Th_p)^{1/2}(\partial_p \hat{F}(\hat{\boldsymbol{\zeta}}) - \partial_p F(\boldsymbol{\zeta})) = (Th_p)^{1/2}(\partial_p \hat{F} - \partial_p F)(\boldsymbol{\zeta}) + o_P(1).$$

Therefore, for any numbers $(\lambda_j)_{j=1,\dots,d}$, we get

$$\sum_{j=1}^{d} \lambda_j \left\{ \partial_p \hat{C}(\mathbf{u}_j) - \partial_p C(\mathbf{u}_j) \right\} = \sum_{j=1}^{d} \lambda_j \left\{ \frac{\partial_p \hat{F}(\hat{\boldsymbol{\zeta}}_j)}{\hat{f}_p(\hat{\zeta}_{pj})} - \frac{\partial_p F(\boldsymbol{\zeta}_j)}{f_p(\zeta_{pj})} \right\}$$
$$= \sum_{j=1}^{d} \lambda_j \left\{ \frac{\partial_p \hat{F}(\hat{\boldsymbol{\zeta}}_j)}{\hat{f}_p(\hat{\zeta}_{pj}) f_p(\zeta_{pj})} [f_p(\zeta_{pj}) - \hat{f}_p(\hat{\zeta}_{pj})] + \frac{\partial_p \hat{F}(\hat{\boldsymbol{\zeta}}_j) - \partial_p F(\boldsymbol{\zeta}_j)}{f_p(\zeta_{pj})} \right\}$$
$$= \sum_{j=1}^{d} \lambda_j \left\{ \frac{\partial_p F(\boldsymbol{\zeta}_j)}{f_p^2(\zeta_{pj})} [f_p(\zeta_{pj}) - \hat{f}_p(\zeta_{pj})] + \frac{\partial_p \hat{F}(\boldsymbol{\zeta}_j) - \partial_p F(\boldsymbol{\zeta}_j)}{f_p(\zeta_{pj})} + o_P((Th_p)^{-1/2}) \right\}.$$

Then we may apply Theorem 7 which delivers asymptotic normality. The asymptotic covariance matrix is also easily deduced from the covariance expressions of Theorem 7:

$$\sigma^*_{i,j} = \frac{1}{f_p(\zeta_{pi}) f_p(\zeta_{pj})} \sigma_{ij} + \frac{\partial_p F(\boldsymbol{\zeta}_i) \partial_p F(\boldsymbol{\zeta}_j)}{f_p^2(\zeta_{pi}) f_p^2(\zeta_{pj})} \sigma_{i+d,j+d}$$
$$- \frac{\partial_p F(\boldsymbol{\zeta}_j)}{f_p^2(\zeta_{pj}) f_p(\zeta_{pi}))} \sigma_{i,j+d} - \frac{\partial_p F(\boldsymbol{\zeta}_i)}{f_p^2(\zeta_{pi}) f_p(\zeta_{pj}))} \sigma_{i+d,j}. \square \qquad (14)$$

Proof of Theorem 7

It is sufficient to prove that

$$(Th_p)^{1/2} \sum_{i=1}^{d} \lambda_i (\partial_p \hat{F} - \partial_p F)(\mathbf{y}_i) + (Th_p)^{1/2} \sum_{i=1}^{d} \mu_i (\hat{f}_p - f_p)(y_{pi})$$

tends weakly to a centered Gaussian random variable whose variance takes the form $\sum_{i,j} \lambda_i \mu_j \overline{\sigma}^2_{i,j}$, for any real numbers λ_i, μ_j, $i, j = 1, ..., d$. Let us first deal with the bias term. Consider $(Th_p)^{1/2} \sum_{i=1}^{d} \lambda_i (E[\partial_p \hat{F}] - \partial_p F)(\mathbf{y}_i)$. As usually,

$$E[\partial_p \hat{F}(\mathbf{y}_i)] - \partial_p F(\mathbf{y}_i) = h_p^{-1} \int \partial_p K\left(\mathbf{y}_i - \mathbf{u}; h\right) F(d\mathbf{u}) - \partial_p F(\mathbf{y}_i)$$
$$= h_p^{-1} \int F(\mathbf{y}_i - h \cdot \mathbf{u})\, k_p'(u_p) \left(\prod_{l \neq p} k_l(u_l)\, du_l \right) du_p - \partial_p F(\mathbf{y}_i)$$
$$= \int \partial_p F(\mathbf{y}_i - h \cdot \mathbf{u}) k(\mathbf{u})\, d\mathbf{u} - \partial_p F(\mathbf{y}_i)$$
$$= O(h_*),$$

since $\partial_p F$ is Lipschitz continuous. Then, this bias term is negligible under our assumptions. Similarly, $(Th_p)^{1/2}\sum_{i=1}^d \mu_i(E[\hat{f}_p]-f)(y_{pi})$ is $o(1)$ under the same assumptions.

Second, to obtain the asymptotic normality, the simplest way to proceed is to apply Lemma 7.1 in ROBINSON (1983) [16]. In his notations, set $p=2d$, $a_i=h_p$, as well as

$$V_{itT}=\lambda_i\left\{(\partial_p K)(\mathbf{y}_i-\mathbf{Y}_t;h)-E\left[(\partial_p K)(\mathbf{y}_i-\mathbf{Y}_t;h)\right]\right\},$$

and

$$V_{itT}^*=\mu_i\left\{k_p(y_{pi}-Y_{pt};h_p)-E\left[k_p(y_{pi}-Y_{pt};h_p)\right]\right\}.$$

We now verify the conditions of validity of Lemma 7.1 in ROBINSON (1983).

Note that, by successive integration by parts, we get

$$\begin{aligned}
E[V_{itT}^2] &= \lambda_i^2\left\{\int(\partial_p K)^2(\mathbf{y}_i-\mathbf{u};h)\ F(d\mathbf{u})-\left(\int(\partial_p K)(\mathbf{y}_i-\mathbf{u};h)\ F(d\mathbf{u})\right)^2\right\}\\
&= \lambda_i^2\left\{\int F(\mathbf{y}_i-h\cdot\mathbf{u})2^n(kk')_p(u_p)\prod_{l\neq p}(kK)_l(u_l)\,d\mathbf{u}-\left(\int F(\mathbf{y}_i-h\cdot\mathbf{u})k_p'(u_p)\prod_{l\neq p}k_l(u_l)\,d\mathbf{u}\right)^2\right\}\\
&= h_p\lambda_i^2\left\{\int\partial_p F(\mathbf{y}_i-h\cdot\mathbf{u})2^{n-1}k_p^2(u_p)\prod_{l\neq p}(kK)_l(u_l)\,d\mathbf{u}\right.\\
&- \left.h_p\left(\int\partial_p F(\mathbf{y}_i-h\cdot\mathbf{u})k_p(u_p)\prod_{l\neq p}k_l(u_l)\,d\mathbf{u}\right)^2\right\}\\
&= 2^{n-1}h_p\lambda_i^2\partial_p F(\mathbf{y}_i)\int k_p^2(u_p)\prod_{l\neq p}(kK)_l(u_l)\,d\mathbf{u}+O(h_*^2)=\lambda_i^2 h_p\sigma_{ii}^2+O(h_*^2).
\end{aligned}$$

Moreover, by similar computations, if $i\neq j$,

$$\begin{aligned}
E[V_{itT}V_{jtT}] &= \lambda_i\lambda_j\left\{\int(\partial_p K)(\mathbf{y}_i-\mathbf{u};h)(\partial_p K)(\mathbf{y}_j-\mathbf{u};h)\ F(d\mathbf{u})\right.\\
&- \left.\left(\int(\partial_p K)(\mathbf{y}_i-\mathbf{u};h)\ F(d\mathbf{u})\right).\left(\int(\partial_p K)(\mathbf{y}_j-\mathbf{u};h)\ F(d\mathbf{u})\right)\right\}\\
&= \lambda_i\lambda_j\left\{\int F(\mathbf{y}_i-h\cdot\mathbf{u})\left[k_p'(u_p)k_p((y_{pj}-y_{pi})/h_p+u_p)+k_p(u_p)k_p'((y_{pj}-y_{pi})/h_p+u_p)\right]\right.\\
&\cdot \left.\prod_{l\neq p}\left[k_l(u_l)K_l((y_{lj}-y_{li})/h_l+u_l)+K_l(u_l)k_l((y_{lj}-y_{li})/h_l+u_l)\right]\,d\mathbf{u}+O(h_*^2).\right\}
\end{aligned}$$

[16] see e.g. BIERENS (1985) or BOSQ (1998) for some alternative sets of assumptions

If $y_{ip} \neq y_{jp}$, then $k_p((y_{pi}-y_{pj})/h_p+u_p)$ and $k'_p((y_{pi}-y_{pj})/h_p+u_p)$ is zero for every $u_p \in A_p$, for h_p sufficiently small. We get that $E[V_{itT}V_{jtT}] = 0$ in this case, for T sufficiently large. Thus, let us assume $y_{pi} = y_{pj}$. For any index $l \neq p$, notice that $K_l((y_{lj} - y_{li})/h_l + u_l)$ is zero if $y_{lj} < y_{li}$ and is one if $y_{lj} > y_{li}$, for T sufficiently large. In the first case, set the change of variable $y_{li} - h_l.u_l = y_{lj} - h_l.v_l$. The l-th factor in brackets becomes $k_l(v_l)K_l((y_{li} - y_{lj})/h_l + v_l)$, which is $k_l(v_l)$ for T sufficiently large. In the second case, the latter factor is $k_l(u_l)$. Thus, $E[V_{itT}V_{jtT}])$ is nonzero only if $y_{pi} = y_{pj}$ and, for T sufficiently large,

$$E[V_{itT}V_{jtT}] = \lambda_i\lambda_j\left\{\int F(\mathbf{y}_i \wedge \mathbf{y}_j - h\cdot\mathbf{u})2k'_p(u_p)k_p(u_p)\cdot\prod_{l\neq p}k_l(u_l)\,d\mathbf{u} + O(h_*^2).\right\} \quad (15)$$

By an integration by parts with respect to u_p (similarly as for $E[V_{itT}^2]$), we get easily

$$E[V_{itT}V_{jtT}] = \lambda_i\lambda_j h_p\sigma_{ij}^2 + o(h_*).$$

Similar computations yield

$$E[V_{itT}^*V_{jtT}^*] = h_p\mu_i\mu_j f_p(y_{pi})\int k_p^2 + o_P(h_p),$$

if $y_{pi} = y_{pi}$, and $E[V_{itT}^*V_{jtT}^*] = o_P(h_p)$ otherwise. Moreover,

$$\begin{aligned}
E[V_{itT}^*V_{jtT}] &= \mu_i\lambda_j\int k_p\left(\frac{y_{pi}-u_p}{h_p}\right)(\partial_p K)(\mathbf{y}_j - \mathbf{u}; h)\ F(d\mathbf{u}) + O_P(h_p^2) \\
= \quad & \mu_i\lambda_j\int F(\mathbf{y}_j - h\cdot\mathbf{u})\prod_{l\neq p}k_l(u_l)\left\{k_p(u_p)k'_p\left(\frac{y_{pi}-y_{pj}}{h_p} + u_p\right)\right. \\
& \left. + \ k'_p(u_p)k_p\left(\frac{y_{pi}-y_{pj}}{h_p} + u_p\right)\right\}\ d\mathbf{u} + O_P(h_p^2).
\end{aligned}$$

If $y_{pi} \neq y_{pj}$, the latter quantity is zero for T sufficiently large. Otherwise, it is equivalent to $\mu_i\lambda_j h_p\partial_p F(\mathbf{y}_j)\int k_p^2$.

Thus, condition $A.7.3$ of ROBINSON (1983) is satisfied. It remains to verify condition $A.7.4$. For every $i \neq j$ and $t \neq t'$,

$$
\begin{aligned}
E[|V_{itT}V_{jt'T}|] = \lambda_i\lambda_j \int & \left|(\partial_p K)(\mathbf{y}_i - \mathbf{u}; h) - \left(\int (\partial_p K)(\mathbf{y}_i - \mathbf{u}; h)\ F(d\mathbf{u})\right)\right| \\
\cdot \quad & \left|(\partial_p K)(\mathbf{y}_j - \mathbf{v}; h) - \left(\int (\partial_p K)(\mathbf{y}_j - \mathbf{v}; h)\ F(d\mathbf{v})\right)\right| F_{(\mathbf{Y}_t, \mathbf{Y}_{t'})}(d\mathbf{u}, d\mathbf{v}) \\
\leq \quad & Cst.h_p^2\lambda_i\lambda_j \int \left(\left|k_p(u_p)\prod_{l\neq p} K_l\left(\frac{y_{li}-u_l}{h}\right)\right| + 1\right)\left(\left|k_p(v_p)\prod_{l\neq p} K_l\left(\frac{y_{lj}-v_l}{h}\right)\right| + 1\right) \\
& f_{t,t'}(\mathbf{u}_{-p}, y_{pi} - h_p u_p, \mathbf{v}_{-p}, y_{pj} - h_p v_p)\, d\mathbf{u}\, d\mathbf{v}.
\end{aligned}
$$

Under our assumptions, the latter quantity if $O(h_*^2)$. Using similar boundings the same property is satisfied for $E[|V_{itT}V^*_{jt'T}|]$ and $E[|V^*_{itT}V^*_{jt'T}|]$.

The other conditions of Lemma 7.1 in ROBINSON (1983) are clearly satisfied. Note that our condition 4 (a) implies Robinson's one on the mixing coefficients, i.e. $\sum_{t=T}^{+\infty}\alpha_t = O(T^{-1})$. Hence we have indeed proved the stated result. $\square$

Asymptotic covariance matrix in Corollary 1

The limiting distribution of $\mathbb{S}$ is a centered Gaussian vector whose covariance matrix has the following (i,j)-th term:

$$E[\phi'(\mathbb{G})(\mathbf{v}_i)\phi'(\mathbb{G})(\mathbf{v}_j)] = E[\mathbb{G}(\boldsymbol{\zeta}_i)\mathbb{G}(\boldsymbol{\zeta}_j)] - \sum_{k=1}^n \partial_k C(\mathbf{v}_i) E[\mathbb{G}(+\infty,\dots,\zeta_{ki},\dots,+\infty)\mathbb{G}(\boldsymbol{\zeta}_j)]$$
$$- \sum_{k=1}^n \partial_k C(\mathbf{v}_j) E[\mathbb{G}(+\infty,\dots,\zeta_{kj},\dots,+\infty)\mathbb{G}(\boldsymbol{\zeta}_i)]$$
$$+ \sum_{k,l=1}^n \partial_k C(\mathbf{v}_i)\partial_l C(\mathbf{v}_j) E[\mathbb{G}(+\infty,\dots,\zeta_{ki},\dots,+\infty)\mathbb{G}(+\infty,\dots,\zeta_{lj},\dots,+\infty)]$$

As previously, for every $i = 1,\dots,d$, we have denoted by $\boldsymbol{\zeta}_i$ the n-dimensional vector

$$(\zeta_{1i},\dots,\zeta_{ni}) = (F_1^{-1}(v_{1i}),\dots,F_n^{-1}(v_{ni})).$$

Recalling Equation (11), we get

$$\begin{aligned} E[\phi'(\mathbb{G})(\mathbf{v}_i)\phi'(\mathbb{G})(\mathbf{v}_j)] &= \sum_{t\in\mathbb{Z}} Cov(\mathbf{1}\{\mathbf{Y}_0 \le \boldsymbol{\zeta}_i\}, \mathbf{1}\{\mathbf{Y}_t \le \boldsymbol{\zeta}_j\}) \\ &- \sum_{k=1}^n \partial_k C(\mathbf{v}_i) \sum_{t\in\mathbb{Z}} Cov(\mathbf{1}\{Y_{0k} \le \zeta_{ki}\}, \mathbf{1}\{\mathbf{Y}_t \le \boldsymbol{\zeta}_j\}) \\ &- \sum_{k=1}^n \partial_k C(\mathbf{v}_j) \sum_{t\in\mathbb{Z}} Cov(\mathbf{1}\{\mathbf{Y}_0 \le \boldsymbol{\zeta}_i\}, \mathbf{1}\{Y_{tk} \le \zeta_{kj}\}) \\ &+ \sum_{k,l=1}^n \partial_k C(\mathbf{v}_i)\partial_l C(\mathbf{v}_j) \sum_{t\in\mathbb{Z}} Cov(\mathbf{1}\{Y_{0k} \le \zeta_{ki}\}, \mathbf{1}\{Y_{tl} \le \zeta_{lj}\}). \end{aligned} \tag{16}$$

References

[1] Azzalini, A. (1981). A Note on the Estimation of a Distribution Function and Quantiles by a Kernel Method. *Biometrika*, **68**, 326-328.

[2] Bierens, H. (1985). Kernel Estimators of Regression Functions. *Advances in Econometrics*, Cambridge University Press, Cambridge, 99-144.

[3] Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction.* Lecture Notes in Statistics, Springer-Verlag, New-York.

[4] Bouyé, E., V. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli (2000). Copulas for finance. A reading guide and some applications. *GRO, Crédit Lyonnais*. Available on the web.

[5] Cebrian, A., M. Denuit, and O. Scaillet (2002). Testing for Concordance Ordering. DP FAME 41.

[6] Denuit, M., and O. Scaillet (2002). Nonparametric Tests for Positive Quadrant Dependence. DP FAME 44.

[7] Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Acad. R. Belg., Bull. Cl. Sci., 5 Série*, **65**, 274-292.

[8] Deheuvels, P. (1981a). A Kolmogorov-Smirnov type test for independence and multivariate samples. *Rev. Roum. Math. Pures et Appl.*, Tome XXVI, **2**, 213-226.

[9] Deheuvels, P. (1981b). A Nonparametric Test of Independence. *Publications de l'ISUP*, **26**, 29-50.

[10] Doukhan, P. (1994). *Mixing: Properties and Examples.* Lecture Notes in Statistics, Springer-Verlag, New-York.

[11] Embrechts, P., A. McNeil, and D. Straumann (2002). Correlation and Dependence in Risk Management: Properties and Pitfalls, in *Risk Management: Value at Risk and Beyond* edited by Dempster M., Cambridge University Press, Cambridge.

[12] Fermanian, J.D., D. Radulovic, and M. Wegkamp (2002). Weak convergence of empirical copula process. Working Paper CREST $2002-06$.

[13] Frees, E., and E. Valdez (1998). Understanding Relationships Using Copulas. *North American Actuarial J.*, **2**, $1-25$.

[14] Frey, R., and A. McNeil (2001). Modelling dependent defaults. Working paper.

[15] Genest, C., Ghoudi, K. and L.P. Rivest (1993). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, **82**, 543-552.

[16] Gouriéroux, C., J.P. Laurent and O. Scaillet (2000). Sensitivity Analysis of Values at Risk. *Journal of Empirical Finance*, **7**, 225-245.

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

[18] Lehmann, E. (1966). Some Concepts of Dependence. *Annals of Mathematical Statistics*, **37**, 1137-1153.

[19] Li, D. (2000). On default correlation : A copula function approach. *Journal of Fixed Income*, **9**, 43-54.

[20] Nelsen, R. (1999). *An Introduction to Copulas*. Lecture Notes in Statistics, Springer-Verlag, New-York.

[21] Patton, A. (2001a). Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula. UCSD WP 2001-09.

[22] Patton, A. (2001b). Estimation of Copula Models for Time Series of Possibly Different Lengths. UCSD WP 2001-17.

[23] Rio, E. (2000). *Théorie Asymptotique des Processus Aléatoires Faiblement Dépendants*. SMAI, Springer.

[24] Robinson, P. (1983). Nonparametric Estimators for Time Series. *Journal of Time Series Analysis*, **4**, 185-207.

[25] Rockinger, M., and E. Jondeau (2001). Conditional Dependency of Financial Series: An Application of Copulas. HEC Paris DP 723.

[26] Scaillet, O. (2000). Nonparametric Estimation and Sensitivity Analysis of Expected Shortfall, forthcoming in *Mathematical Finance*.

[27] Shi, J., and T. Louis (1995). Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, **51**, 1384-1399.

[28] Tjostheim, D. (1996). Measures of Dependence and Tests of Independence. *Statistics*, **28**, 249-284.

[29] Van der Vaart, A., and J. Wellner (1996). *Weak Convergence and Empirical Processes*. Springer-Verlag, New York.

Figures 1-8

Figure 1 : Returns and rank statistics for CAC40-DAX35

Figure 2 : Copula estimates and contour plots for CAC40-DAX35

Figure 3 : Comparison with independent, comonotonic and Gaussian copulas for CAC40-DAX35

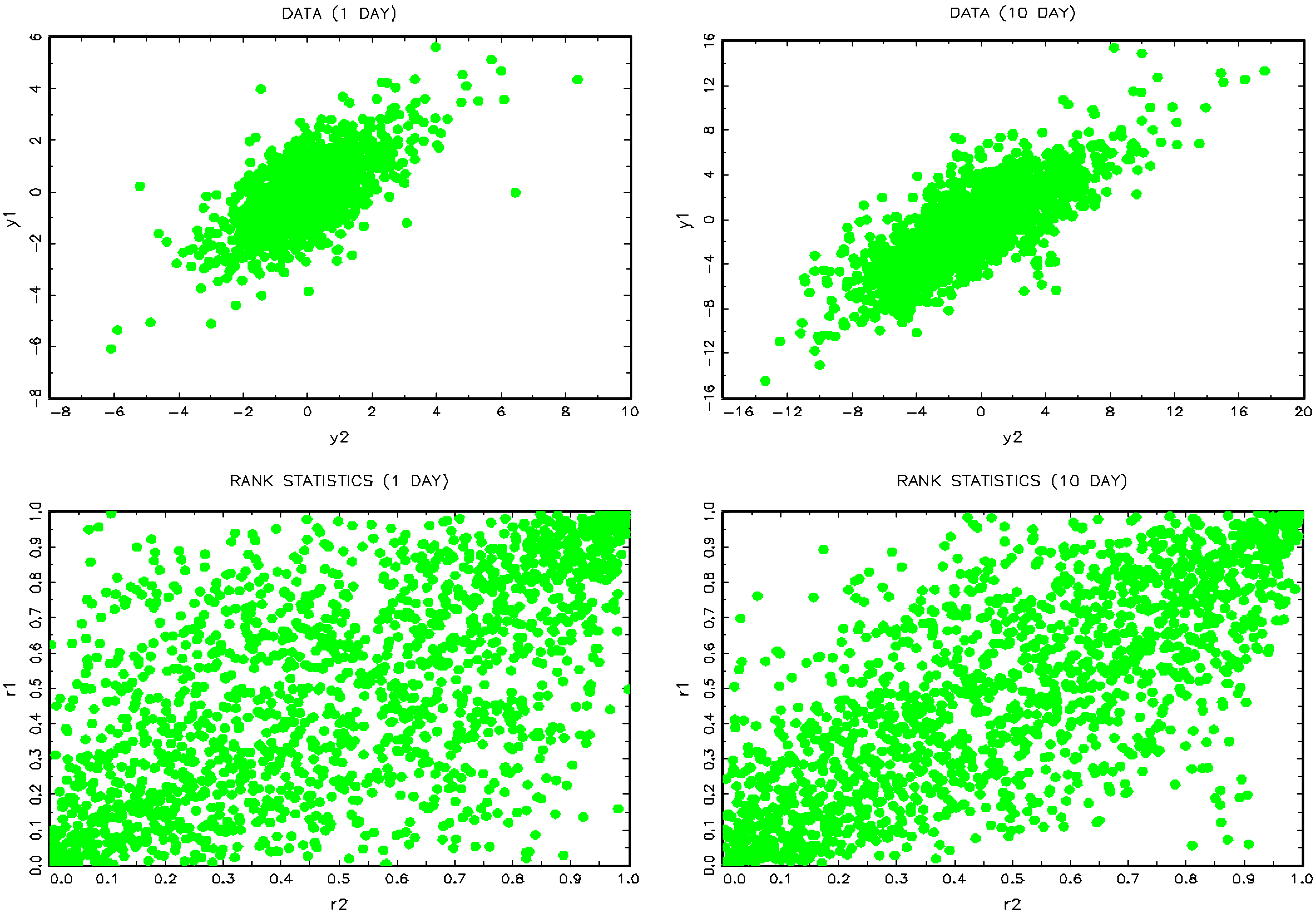
Figure 4 : Left tail decreasing behaviour for CAC40-DAX35

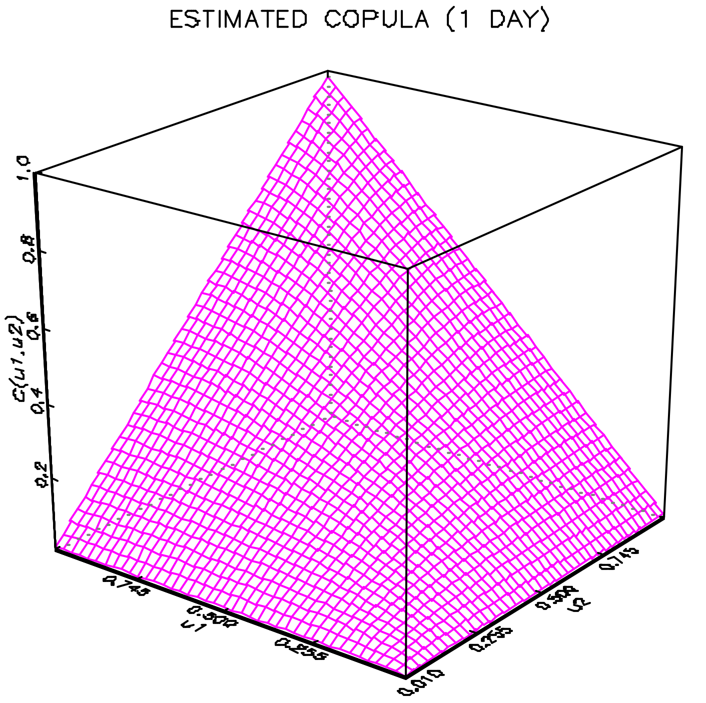
Figure 5 : Returns and rank statistics for S&P500-DJI

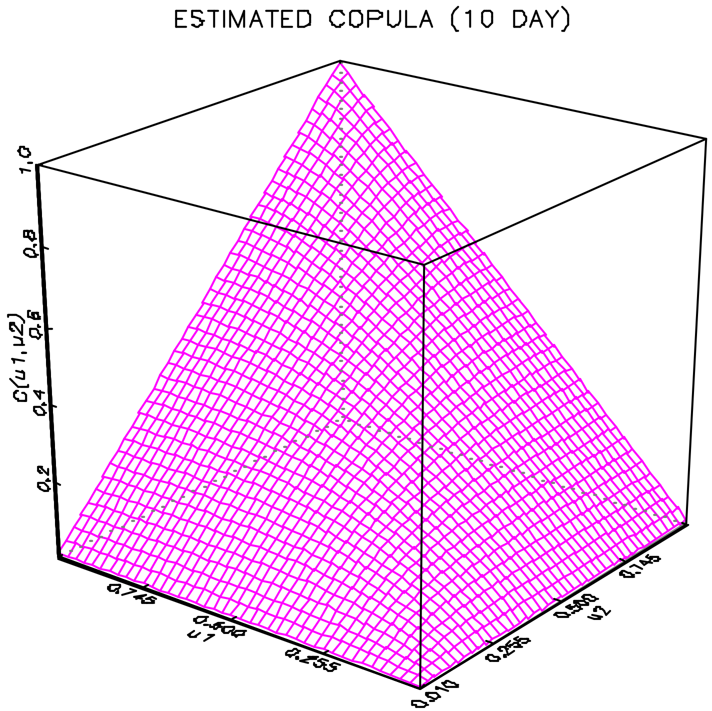
Figure 6 : Copula estimates and contour plots for S&P500-DJI

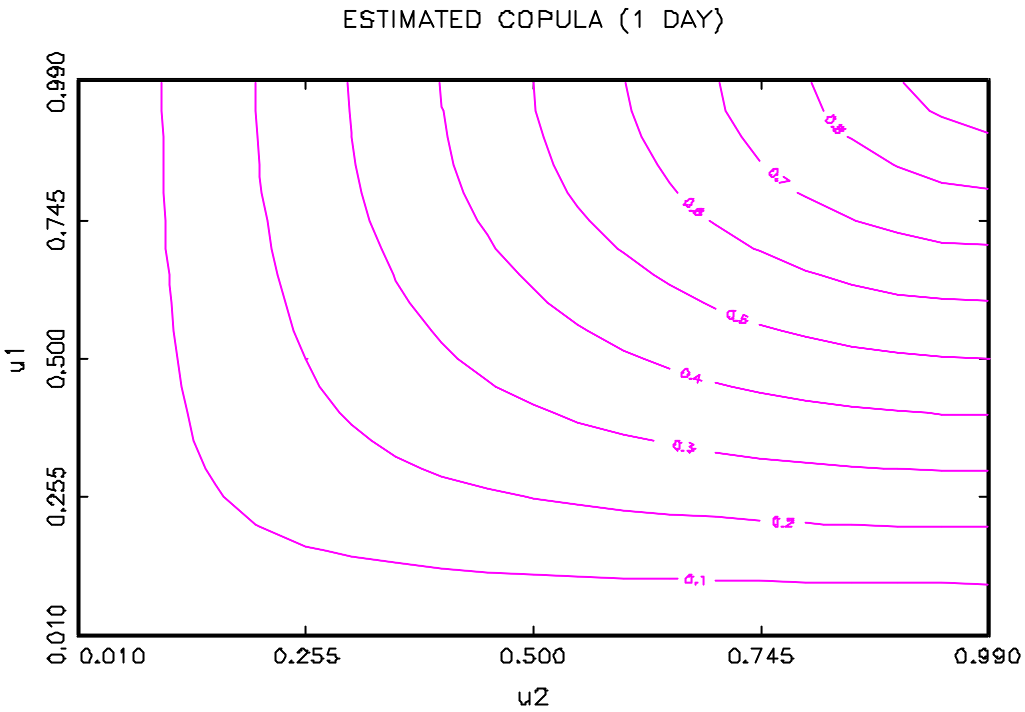
Figure 7 : Comparison with independent, comonotonic and Gaussian copulas for S&P500-DJI

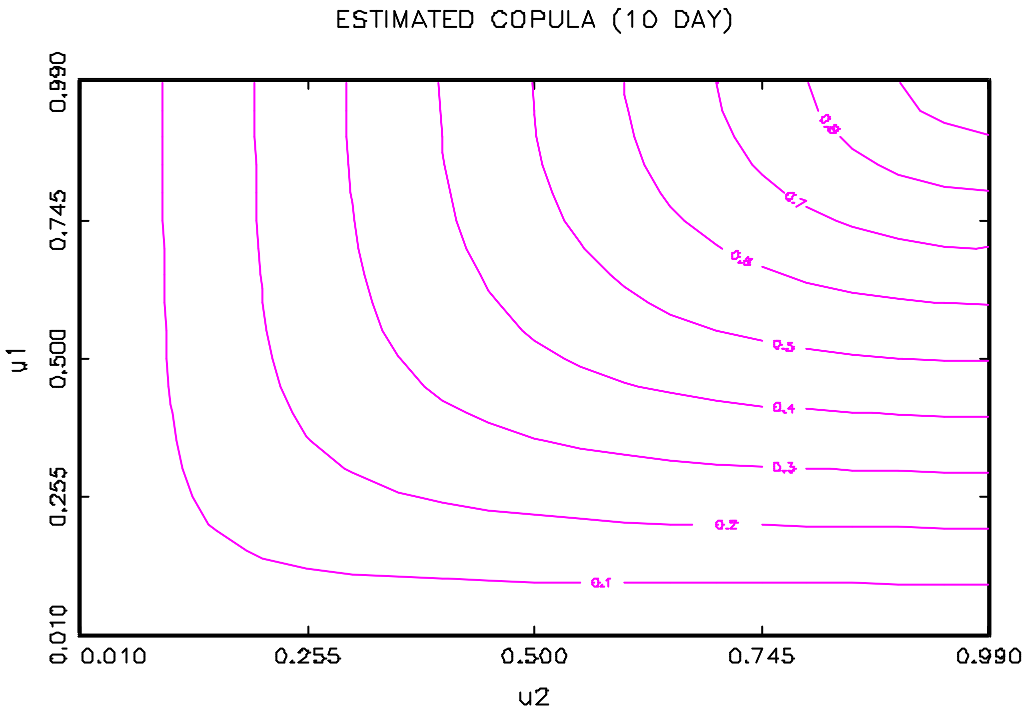
Figure 8 : Left tail decreasing behaviour for S&P500-DJI


DATA (1 DAY)
y1
y2
DATA (10 DAY)
y1
y2
RANK STATISTICS (1 DAY)
r1
r2
RANK STATISTICS (10 DAY)
r1
r2

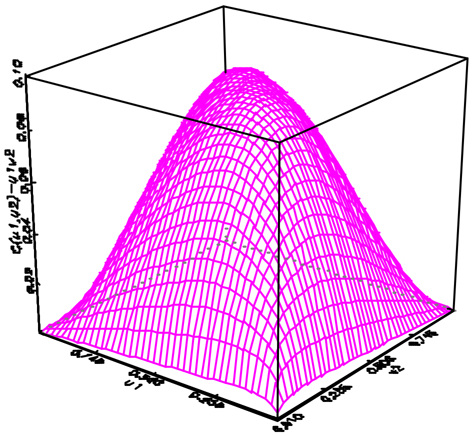

ESTIMATED COPULA (1 DAY)
C(u1,u2)
1.0
0.8
0.6
0.4
0.2
u1
0.745
0.500
0.255
u2
0.010
0.255
0.500
0.745


ESTIMATED COPULA (10 DAY)
C(u1,u2)
1.0
0.8
0.6
0.4
0.2
u1
0.745
0.500
0.255
u2
0.010
0.255
0.500
0.745

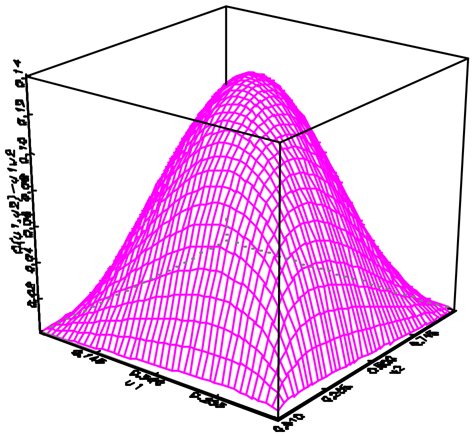

ESTIMATED COPULA (1 DAY)
u1
0.990
0.745
0.500
0.255
0.010
u2
0.010
0.255
0.500
0.745
0.990
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8


ESTIMATED COPULA (10 DAY)
u1
0.990
0.745
0.500
0.255
0.010
u2
0.010
0.255
0.500
0.745
0.990
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8

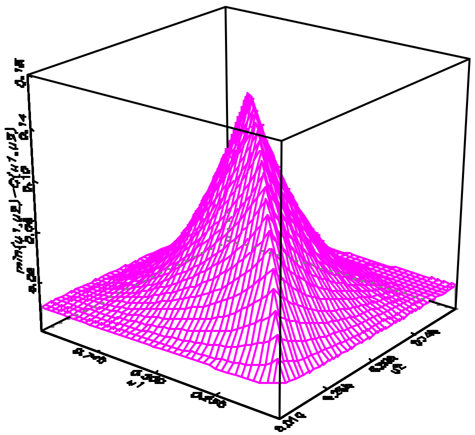
POSITIVE QUADRANT DEPENDENCE (1 DAY)



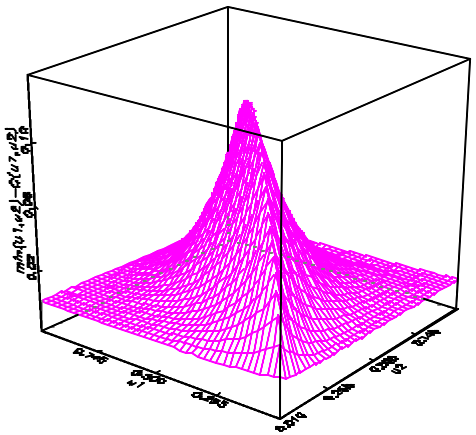
POSITIVE QUADRANT DEPENDENCE (10 DAY)



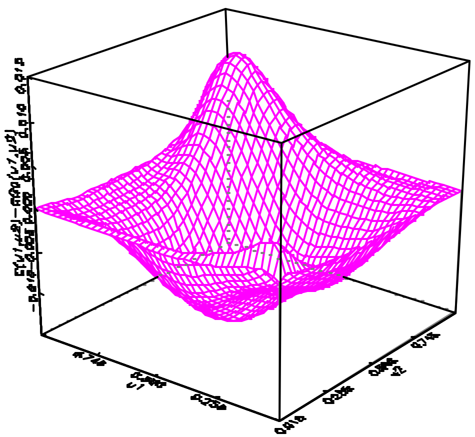
COMONOTONICITY (1 DAY)



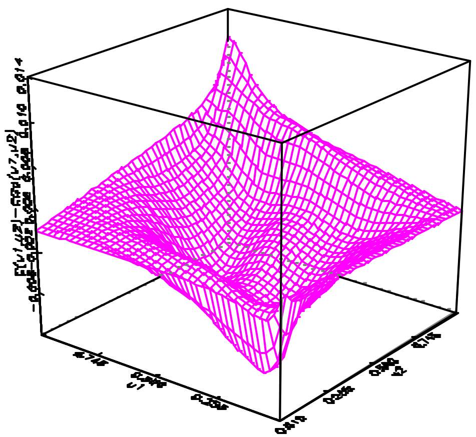
COMONOTONICITY (10 DAY)



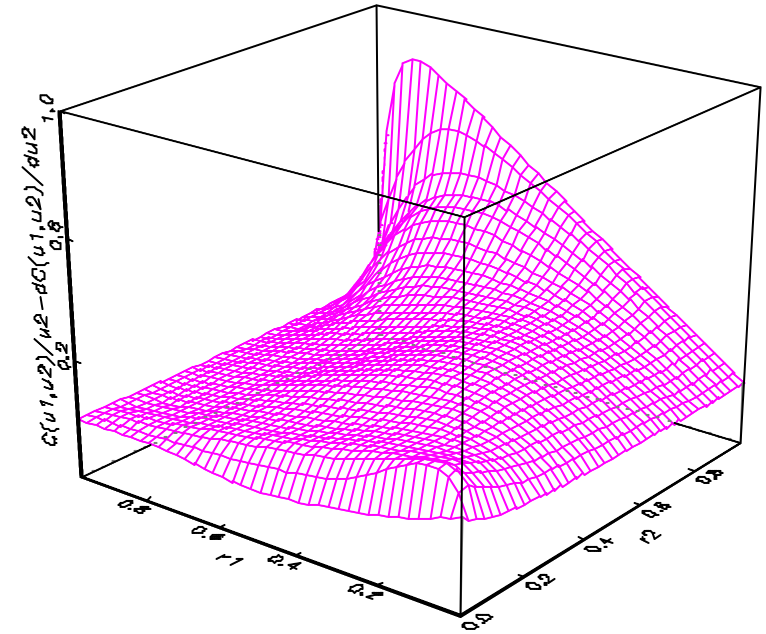
GAUSSIAN (1 DAY)



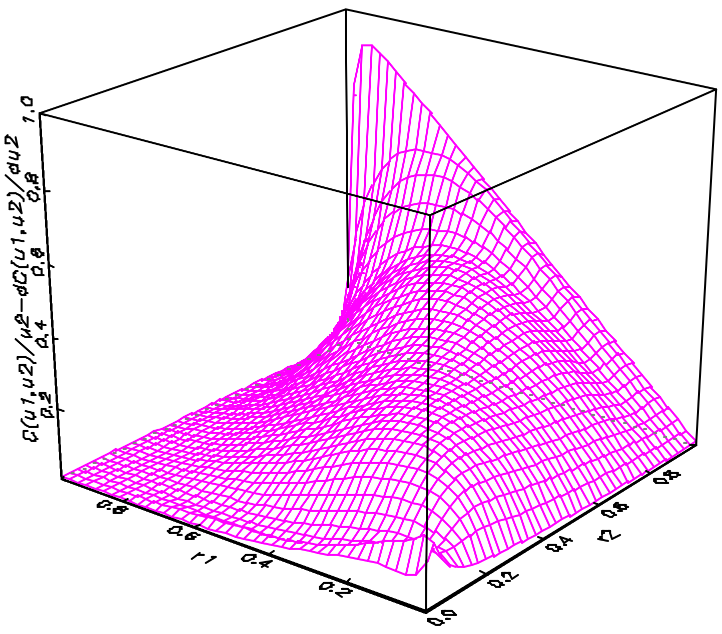
GAUSSIAN (10 DAY)



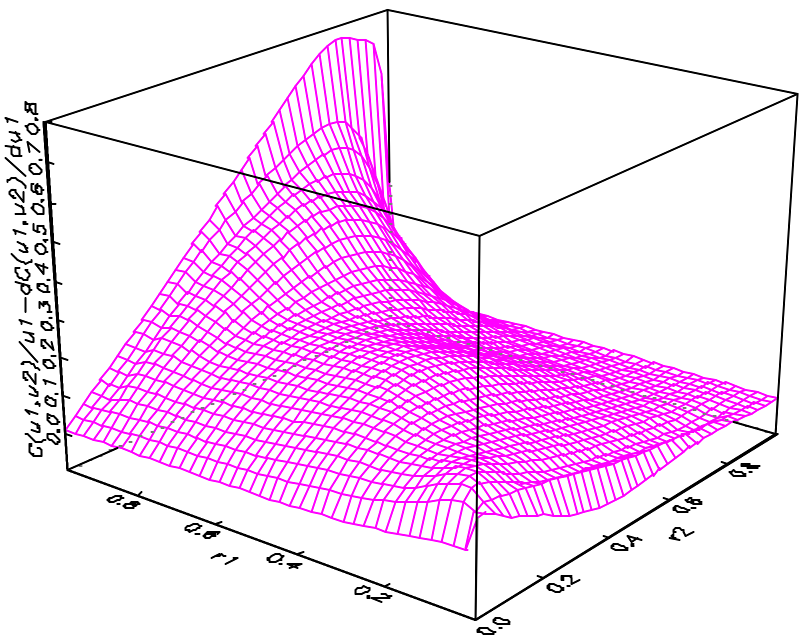
LEFT TAIL DEPENDENCE OF Y1 IN Y2 (1 DAY)



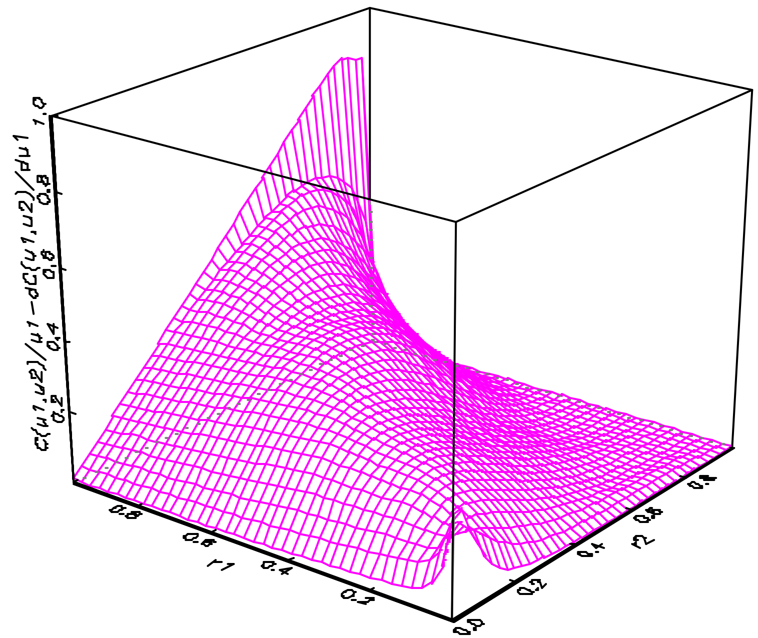
LEFT TAIL DEPENDENCE OF Y1 IN Y2 (10 DAY)

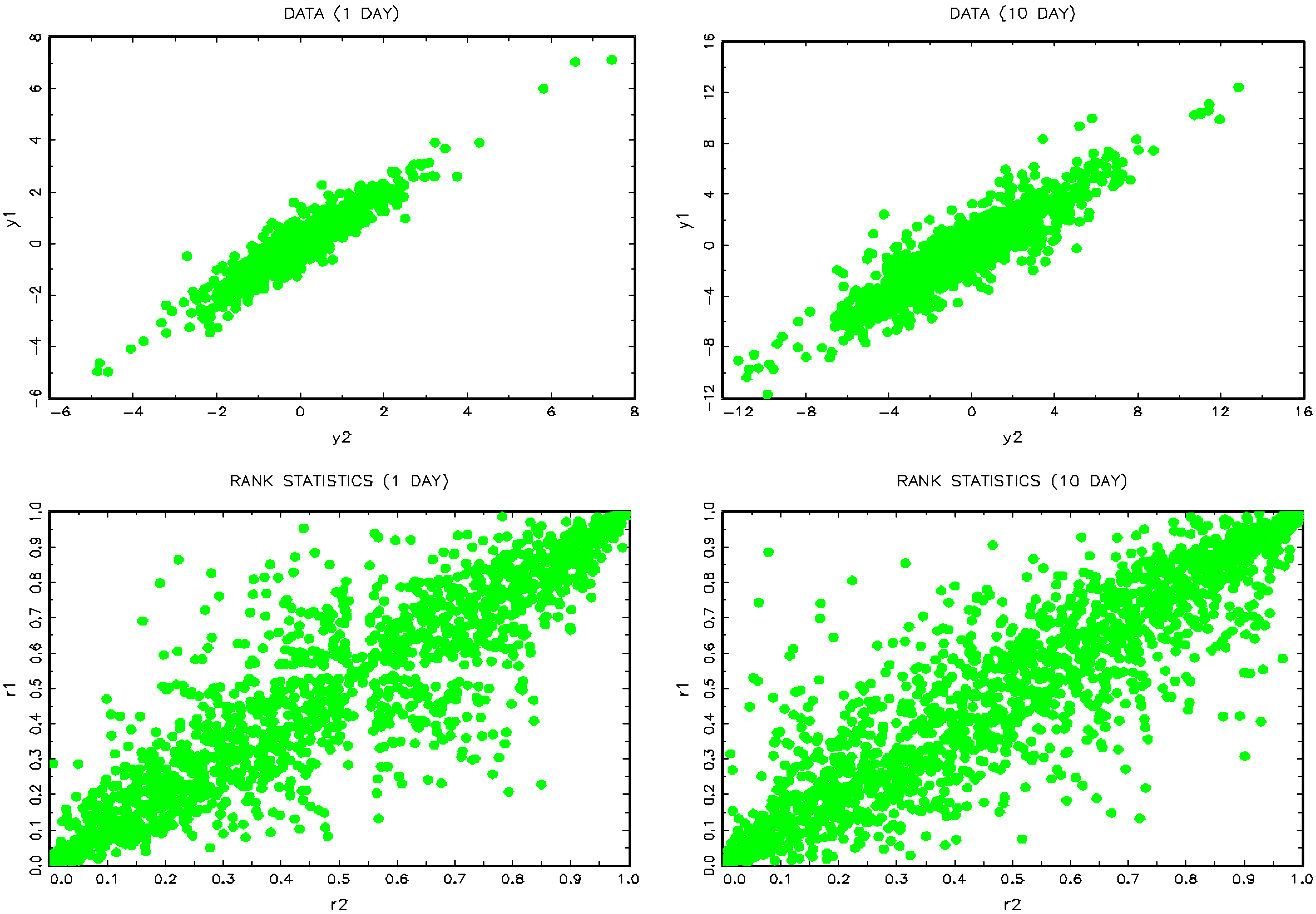


LEFT TAIL DEPENDENCE OF Y2 IN Y1 (1 DAY)

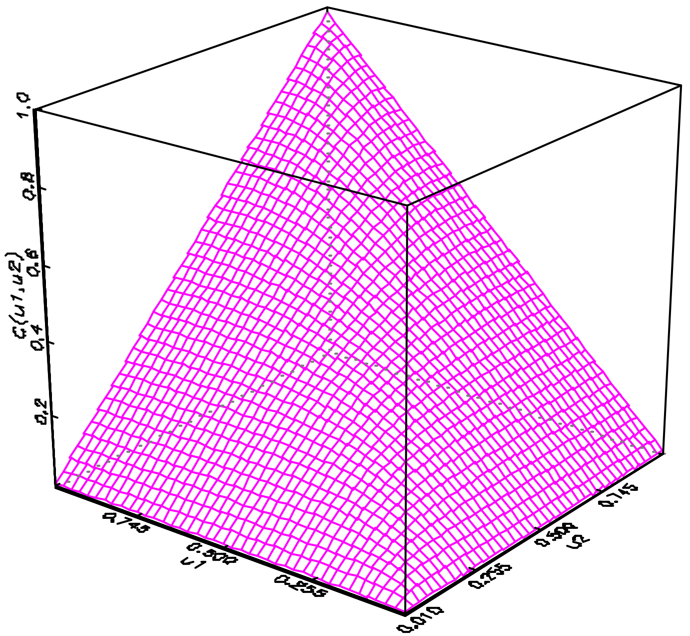


LEFT TAIL DEPENDENCE OF Y2 IN Y1 (10 DAY)

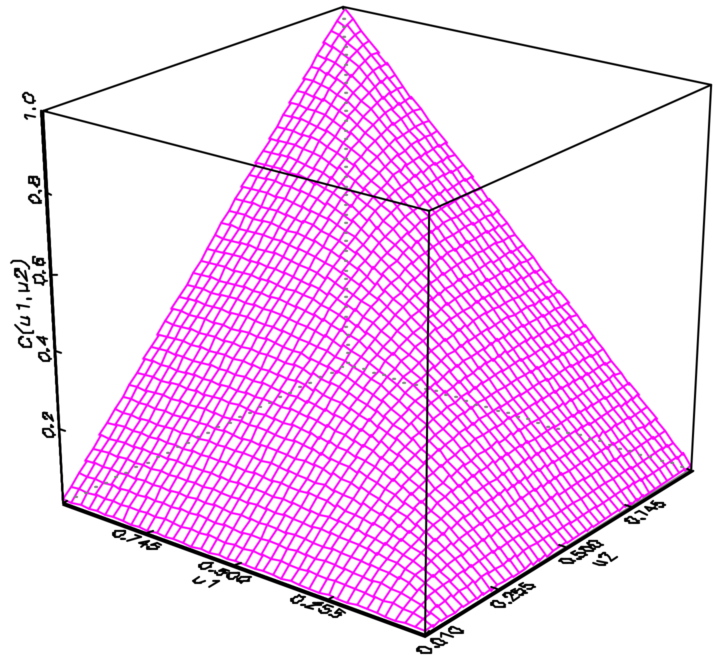



DATA (1 DAY)
DATA (10 DAY)
RANK STATISTICS (1 DAY)
RANK STATISTICS (10 DAY)
y1
y2
r1
r2

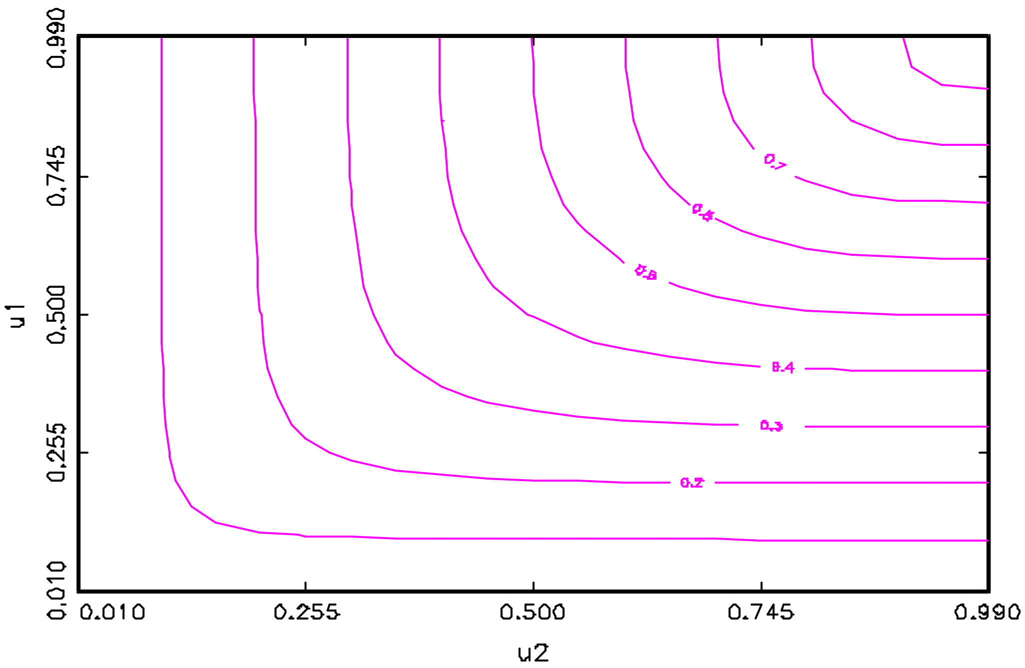
ESTIMATED COPULA (1 DAY)



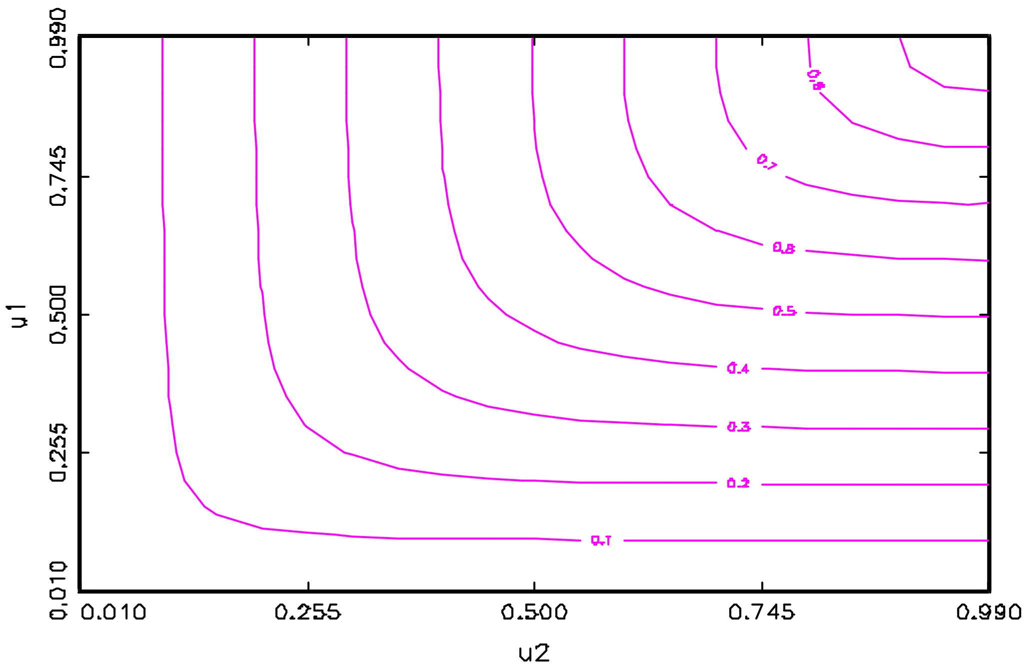
ESTIMATED COPULA (10 DAY)



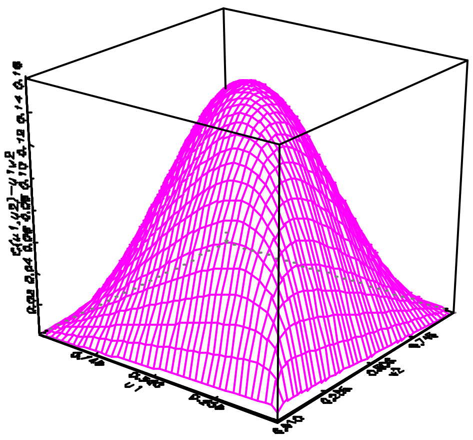
ESTIMATED COPULA (1 DAY)



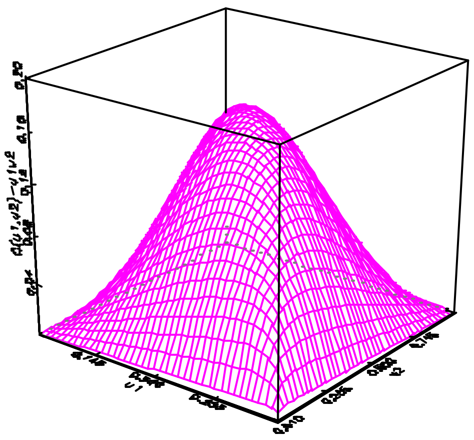
ESTIMATED COPULA (10 DAY)



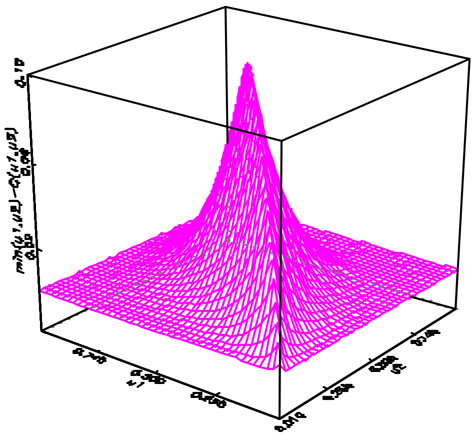
POSITIVE QUADRANT DEPENDENCE (1 DAY)



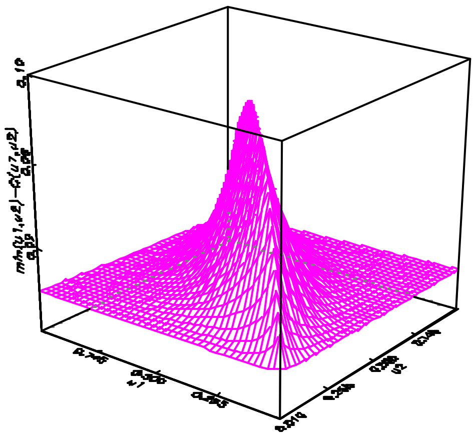
POSITIVE QUADRANT DEPENDENCE (10 DAY)



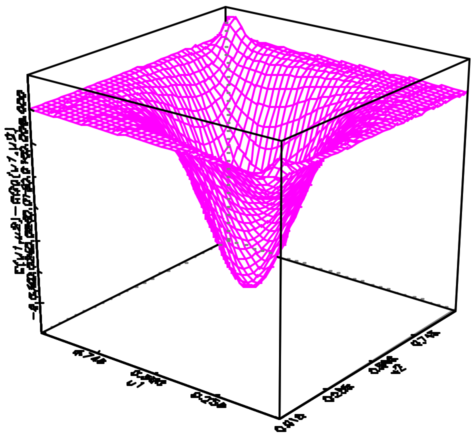
COMONOTONICITY (1 DAY)



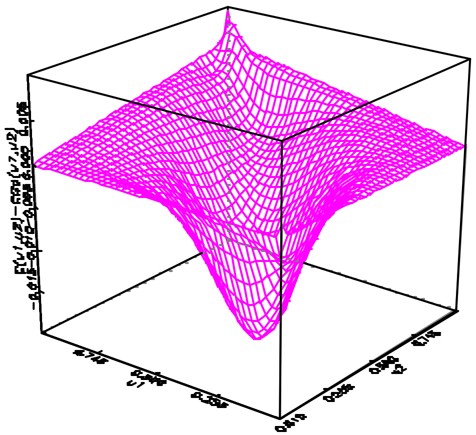
COMONOTONICITY (10 DAY)



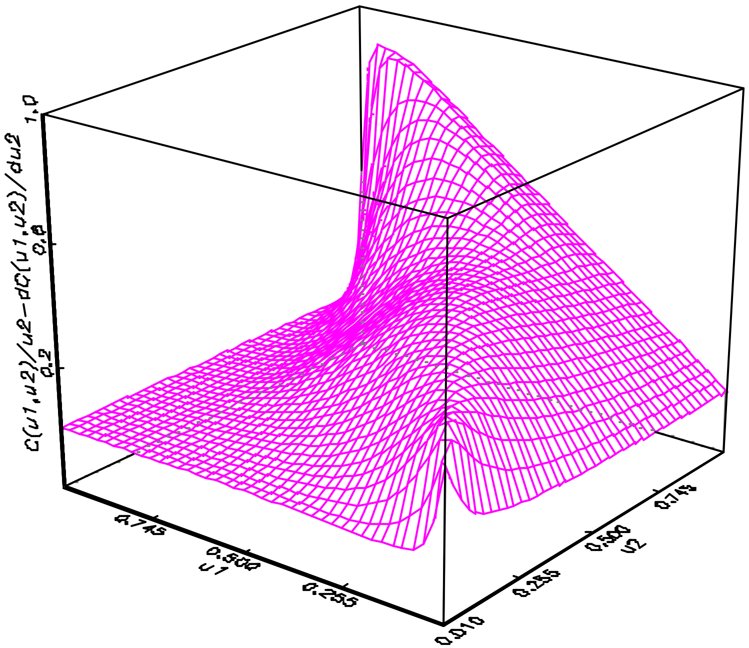
GAUSSIAN (1 DAY)



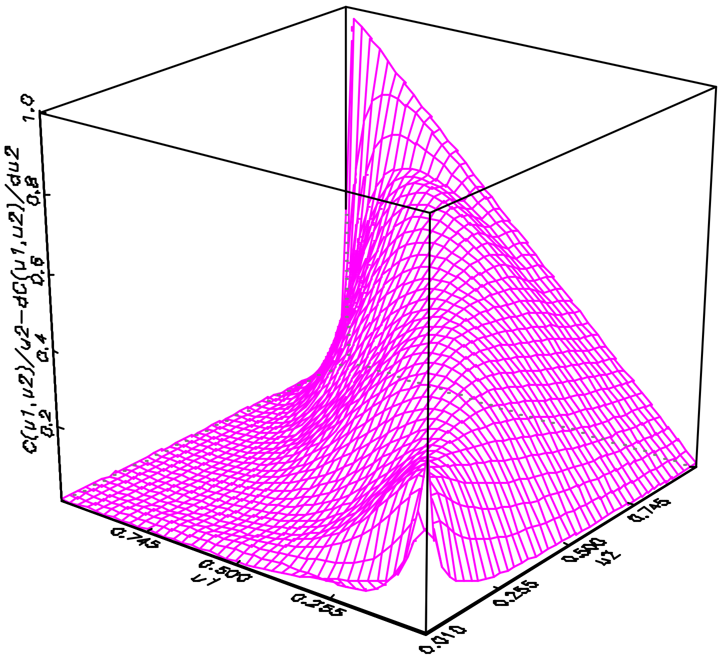
GAUSSIAN (10 DAY)



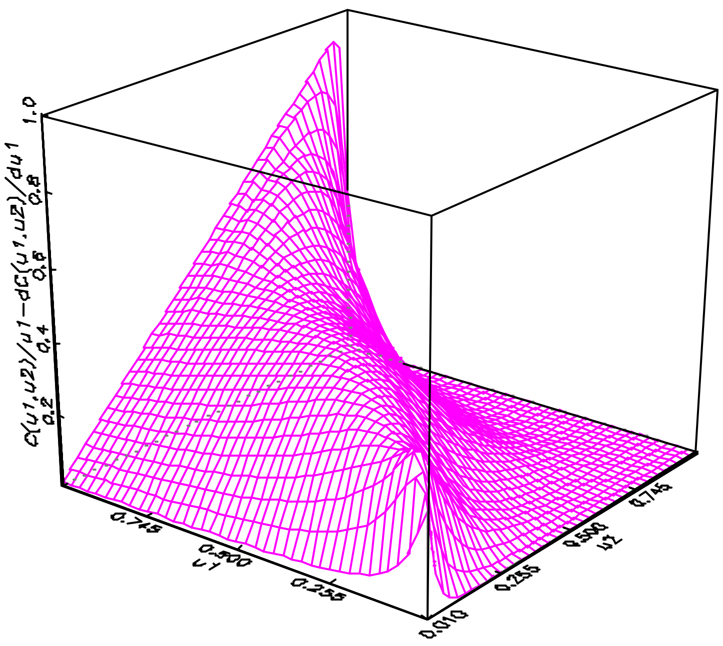
LEFT TAIL DEPENDENCE OF Y1 IN Y2 (1 DAY)



LEFT TAIL DEPENDENCE OF Y1 IN Y2 (10 DAY)



LEFT TAIL DEPENDENCE OF Y2 IN Y1 (1 DAY)



LEFT TAIL DEPENDENCE OF Y2 IN Y1 (10 DAY)