Goodness-of-fit for regime-switching copula models with application to option pricing

Article *in* Canadian Journal of Statistics · October 2019

3 authors:

Bouchra R. Nasri
McGill University

Bruno N. Remillard
HEC Montréal - École des Hautes Études commerciales

Mamadou Yamar Thioub
HEC Montréal - École des Hautes Études commerciales

The Canadian Journal of Statistics
Vol. xx, No. yy, 2019, Pages 1–25
La revue canadienne de statistique

Goodness-of-fit for regime-switching copula models with application to option pricing

Bouchra R. Nasri [1*], Bruno N. Rémillard[2] and Mamadou Yamar Thioub[2]

[1] *McGill University, Department of Mathematics and Statistics, 805 Rue Sherbrooke O, Montréal (Québec), QC H3A 0B9*
[2] *HEC Montréal, 3000 chemin de la Côte Sainte-Catherine, Montréal (Québec), Canada H3T 2A7*

Key words and phrases: Goodness-of-fit; time series; copulas ; regime-switching models ; generalized error models.

MSC 2010: Primary 62M10; secondary 62P05

Abstract:
We consider several time series and for each of them, we fit an appropriate dynamic parametric model. This produces serially independent error terms for each time series. The dependence between these error terms is then modeled by a regime-switching copula. The EM algorithm is used for estimating the parameters and a sequential goodness-of-fit procedure based on Cramér-von Mises statistics is proposed to select the appropriate number of regimes. Numerical experiments are performed to assess the validity of the proposed methodology. As an example of application, we evaluate a European put-on-max option on the returns of two assets. In order to facilitate the use of our methodology, we have built a R package HMMcopula available on CRAN. *The Canadian Journal of Statistics* xx: 1–25; 2019 © 2019 Statistical Society of Canada

Résumé: Nous considérons plusieurs séries temporelles univariées, et pour chacune nous trouvons un modèle dynamique paramétrique approprié. Nous obtenons alors des termes d'erreur indépendants pour chaque série. La dépendance entre ces termes d'erreur est ensuite modélisée par une copule avec changement de régime. L'algorithme EM est utilisé pour estimer les paramètres et une procédure séquentielle de tests d'adéquation basés sur la statistique de Cramér-von Mises est proposée pour sélectionner le nombre approprié de régimes. Nous réalisons une série d'expériences numériques afin d'évaluer la validité et la performance de la méthodologie proposée. Comme exemple d'application, nous évaluons le prix d'une option de vente européenne sur le rendement maximal de deux titres en utilisant un modèle de copule à changement de régime. Finalement, afin de faciliter l'utilisation future de la méthodologie proposée, nous avons construit une librairie de fonctions basée sur le progiciel R, qui s'intitule HMMcopula, et qui est disponible gratuitement sur CRAN. *La revue canadienne de statistique* xx: 1–25; 2019 © 2019 Société statistique du Canada

1. INTRODUCTION

In finance, many instruments are based on several risky assets and their evaluation rest on the joint distribution of these assets. In fact, to determine this joint distribution, we must take into account the serial dependence in each asset, as well as the dependence between the assets. Underestimating the latter can have devastating financial and economic consequences, as exemplified by the 2008 financial crisis. We must also consider that the dependence may vary with time, potentially increasing in crisis periods. Some ways to take into account time-varying dependence have been proposed. Recently, Adams et al. (2017) fitted DCC-GARCH models (Engle, 2002)

* *Author to whom correspondence may be addressed.*
E-mail: bouchra.nasri@gmail.com

to multivariate time series, which is a bit restrictive in terms of dependence since it is based on the multivariate Gaussian distribution. To overcome this limitation, and because copulas are specially designed to model dependence, it is no wonder that many time-varying dependence models are based on copulas.

To our knowledge, the first papers involving time-dependent copulas were Patton (2004) and van den Goorbergh et al. (2005). In Patton (2004), the author fitted a Gaussian copula on monthly returns (assumed independent), where the correlation parameter was a function of covariates. In van den Goorbergh et al. (2005), the authors, in order to evaluate call-on-max options, fitted a copula family to the residuals of two GARCH time series, with a parameter expressed as a function of the volatilities. Note that both are special cases of what is now known as single-index copula (Fermanian and Lopez, 2018). One can also use the methodology proposed in Nasri and Rémillard (2019), where generalized error models are fitted to each time series, and the underlying copula has time-dependent parameters. In order to be able to take into account abrupt changes in the dependence, it can be appropriate to use regime-switching copulas.

This approach has been proposed recently for vines in Stöber and Czado (2014), Fink et al. (2017) and for hierarchical Archimedean models in Härdle et al. (2015). In all cases, the dynamic models for the marginal distributions were ARMA-GARCH, and there was no formal test of goodness-of-fit. The selection of the number of regimes was based on comparisons of likelihoods, using also rolling windows. There is not yet a theory supporting this method in our setting but results from Cappé et al. (2005)[Chapter 15] shows that the BIC selection criterion works for HMM with a discrete finite state space for the observations; unfortunately, here this hypothesis is not met.

In this article, we propose a formal goodness-of-fit test for regime-switching copulas, which was not done before. As a by-product, we obtain another way to select the number of regimes based on P-values. More precisely, in Section 2, we describe the model for the time series and we define regime-switching copulas. In Section 3, we detail the estimation procedure, the goodness-of-fit test, and the selection of the number of regimes. Numerical experiments are performed in Section 4 to assess the validity of the procedures to choose the number of regimes. In Section 5, we give an example of application for option pricing, along the same lines as van den Goorbergh et al. (2005) but with different data. Note that we have built a R package for regime switching copula models, HMMcopula available at CRAN (Thioub et al., 2018).

2. LINKING MULTIVARIATE TIME SERIES WITH REGIME-SWITCHING COPULAS

To introduce copula-based models, we proceed in two steps: first, for each univariate time series, we use a "generalized error model" (Du, 2016) to produce iid univariate series; second, regime-switching copulas are fitted to these series. To fix ideas, let $\mathbf{X}_t = (X_{1t}, \ldots, X_{dt})$, be a multivariate time series. For each $j \in \{1, \ldots, d\}$, let $\mathcal{F}_{j,t-1}$ contains information from the past of $X_{j1}, \ldots X_{j,t-1}$, and possibly information from exogenous variables as well. Further set $\mathcal{F}_t = \vee_{j=1}^{d} \mathcal{F}_{j,t}$. Assume that for each $j \in \{1, \ldots, d\}$, there exist continuous, increasing, and $\mathcal{F}_{j,t-1}$-measurable functions $G_{\boldsymbol{\alpha},jt}$ so that $\varepsilon_{jt} = G_{\boldsymbol{\alpha},jt}(X_{jt})$ are iid with continuous distribution function F_j and density f_j, for some unknown parameter $\boldsymbol{\alpha} \in \mathcal{A}$. Note that stochastic volatility models and Hidden Markov models (HMM) are particular cases of generalized error models. Next, to introduce the dependence between the time series, we choose a sequence of $\mathcal{F}_{t-1}$-measurable copulas C_t, so that the joint conditional distribution function K_t of $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, \ldots, \varepsilon_{dt})$ given $\mathcal{F}_{t-1}$ is $K_t(\mathbf{x}) = C_t\{\mathbf{F}(\mathbf{x})\}$, with $\mathbf{F}(\mathbf{x}) = (F_1(x_1), \ldots, F_d(x_d))^\top$, for any $\mathbf{x} = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$. In particular, $\mathbf{U}_t = \mathbf{F}(\boldsymbol{\varepsilon}_t) \sim C_t$, for every $t \in \{1, \ldots, n\}$.

This way of modeling dependence between several time series is usually applied to innovations of stochastic volatility models (van den Goorbergh et al., 2005; Chen and Fan, 2006; Patton, 2006; Rémillard, 2017). For example, suppose that $X_{1t} = \mu_{1t}(\boldsymbol{\alpha}) + \sigma_{1t}(\boldsymbol{\alpha})\varepsilon_{1t}$, $\varepsilon_{1t} \sim F_1$, where μ_{1t} and σ_{1t} are $\mathcal{F}_{1,t-1}$-measurable, and the innovations ε_{1t} are independent of $\mathcal{F}_{1,t-1}$. In this case, one could take $G_{\boldsymbol{\alpha},1t}(x_1) = \frac{x_1 - \mu_{1t}(\boldsymbol{\alpha})}{\sigma_{1t}(\boldsymbol{\alpha})}$, and then $\varepsilon_{1t} = G_{\boldsymbol{\alpha},1t}(X_{1t}) \sim F_1$. We can also consider Gaussian HMM models for some univariate time series. For example, suppose that there exists a Markov chain s_t on $\{1, \ldots, m\}$ with transition matrix Q so that given $s_1 = i_1, \ldots, s_n = i_n$, $X_{11}, \ldots, X_{1n}$ are independent, and $X_{1t} \sim N(\mu_{i_t}, \sigma^2_{it})$. If $\eta_{t-1}(k)$ is the probability of being in regime $k \in \{1, \ldots, m\}$ at time $t-1$ given the past observations $X_{11}, \ldots, X_{1,t-1}$, then the conditional distribution G_{1t} of X_{1t} given the past is $G_{1t}(x) = \sum_{k=1}^m W_{t-1}(k) F^{(k)}(x)$, where $W_{t-1}(k) = \sum_{j=1}^m \eta_{t-1}(j) Q_{jk}$ is the probability of being in regime k at time t given the past observations, and $F^{(k)}$ is the cdf of a Gaussian distribution with mean μ_k and variance σ_k^2. It then follows that the sequence $U_{1t} = G_{1t}(X_{1t})$ are iid uniform random variables.

After having chosen the generalized error models for each univariate time series, we need to choose the regime-switching copula model C_t for the multivariate series $\mathbf{U}_t$. This means that there exists a finite Markov chain τ_t on $\{1, \ldots, \ell\}$ with transition matrix P, so that given $\tau_1 = i_1, \ldots, \tau_n = i_n$, $\mathbf{U}_1, \ldots, \mathbf{U}_n$ are independent, and $\mathbf{U}_t \sim C_{\boldsymbol{\beta}_{i_t}}$, $t \in \{1, \ldots, n\}$, where $\{C_{\boldsymbol{\beta}}; \boldsymbol{\beta} \in \mathcal{B}\}$ is a given parametric copula family. Also we assume the usual smoothness conditions on the associated densities $c_{\boldsymbol{\beta}}$ so that the pseudo-maximum likelihood estimator exists. Note that for a given $j \in \{1, \ldots, d\}$, one needs that the values U_{jt}, $t \in \{1, \ldots, n\}$, are iid uniform. This is indeed true as proven in the following theorem.

Theorem 1. *Suppose that the multivariate time series $\mathbf{U}_t$ has distribution function C_t given $\mathcal{F}_{t-1}$. Then for any given $j \in \{1, \ldots, d\}$, the values U_{jt}, $t \in \{1, \ldots, n\}$, are iid uniform.*

Proof of Theorem 1. F. or simplicity, suppose that $j = 1$. By hypothesis, $\mathbb{P}(U_{1t} \leq u_1, \ldots, U_{dt} \leq u_d | \mathcal{F}_{t-1}) = C_t(u_1, \ldots, u_d)$. From the properties of copulas, one gets that $\mathbb{P}(U_{1t} \leq u_1 | \mathcal{F}_{t-1}) = C_t(u_1, 1, \ldots, 1) = u_1$. As a result, one may conclude that $U_{11}, \ldots, U_{1n}$ are iid uniform. ■

Since the generalized errors $\boldsymbol{\varepsilon}_t$ are not observable, $\boldsymbol{\alpha}$ being unknown, the latter must be estimated by a consistent estimator $\boldsymbol{\alpha}_n$. One can then compute the pseudo-observations $\mathbf{e}_{n,t} = (e_{n,1t}, \ldots, e_{n,dt})^\top = \mathbf{G}_{\boldsymbol{\alpha}_n,t}(\mathbf{X}_t)$, where $e_{n,jt} = G_{\boldsymbol{\alpha}_n,jt}(X_{jt})$, $j \in \{1, \ldots, d\}$ and $t \in \{1, \ldots, n\}$. Using these pseudo-observations might be a problem, but in Nasri and Rémillard (2019), it was shown that using the normalized ranks of these pseudo-observations, one can estimate the parameters $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_\ell$ and P, as if one was observing $\mathbf{U}_1, \ldots, \mathbf{U}_n$. The same applies to the goodness-of-fit test that will be defined in Section 3.2.

Based on Theorem 2, note that in order to simulate the multivariate time series, it suffices to generate $\mathbf{U}_t = (U_{1t}, \ldots, U_{dt})$ according to the regime-switching copula model, set $\varepsilon_{jt} = F_j^{-1}(U_{jt})$, and then compute $X_{jt} = G^{-1}_{\boldsymbol{\alpha},jt}(\varepsilon_{jt})$, $j \in \{1, \ldots, d\}$, and $t \in \{1, \ldots, n\}$.

3. ESTIMATION AND GOODNESS-OF-FIT TEST

We first present general regime-switching models which can be applied to univariate time series or copula. Then, we describe an estimation procedure and a goodness-of fit test for regime-switching copula models. Finally, we propose a sequential procedure for selecting the optimal number of regimes.

3.1. General regime-switching models

Let τ_t be a homogeneous discrete-time Markov chain on $S = \{1, \ldots, \ell\}$, with transition probability matrix P on $S \times S$. Given $\tau_1 = k_1, \ldots, \tau_n = k_n$, the observations $\mathbf{Y} = (Y_1, \ldots, Y_n)$ are independent with densities $g_{\boldsymbol{\beta}_{k_t}}$, $t \in \{1, \ldots, n\}$. Set $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_\ell, P)$. Then the joint density of $\tau = (\tau_1, \ldots, \tau_n)$ and $\mathbf{Y}$ is

$$f_{\boldsymbol{\theta}}(\tau, \mathbf{Y}) = \left(\prod_{t=1}^{n} P_{\tau_{t-1},\tau_t} \right) \times \prod_{t=1}^{n} g_{\boldsymbol{\beta}_{\tau_t}}(Y_t), \tag{1}$$

so one can write

$$\log f_{\boldsymbol{\theta}}(\tau, \mathbf{Y}) = \sum_{t=1}^{n} \log P_{\tau_{t-1},\tau_t} + \sum_{t=1}^{n} \log g_{\boldsymbol{\beta}_{\tau_t}}(Y_t). \tag{2}$$

Because the regimes τ_t are not observable, an easy way to estimate the parameters is to use the EM algorithm (Dempster et al., 1977), which proceeds in two steps: expectation (E step), where $Q_{\mathbf{y}}(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta}}\{\log f_{\tilde{\boldsymbol{\theta}}}(\tau, \mathbf{Y}) | \mathbf{Y} = \mathbf{y}\}$ is computed, and maximization (M step), where one computes

$$\boldsymbol{\theta}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} Q_y \left(\boldsymbol{\theta}, \boldsymbol{\theta}^{(k)} \right),$$

starting from an initial value $\boldsymbol{\theta}^{(0)}$. As $k \to \infty$, $\boldsymbol{\theta}^{(k)}$ converges to the maximum likelihood estimator of the density of $\mathbf{Y}$. The formulas for the EM steps are given in Appendix 6. As a particular case of regime-switching models, if $P_{ij} = \nu_j$, then one gets mixture models. In this case $\tau_1, \ldots, \tau_n$ are iid. The simplified formulas for the EM steps are given in Appendix . For application to copulas, the density $g_{\boldsymbol{\beta}}$ is the density of a parametric family of copulas $C_{\boldsymbol{\beta}}$, with $\boldsymbol{\beta} \in \mathcal{B}$. However $Y_1, \ldots, Y_n$ are not observable so they must be replaced by the normalized ranks of the pseudo-observations $\mathbf{e}_{n,t}$, i.e., $Y_{jt} = \mathrm{rank}(e_{n,jt})/(n+1)$.

3.2. Goodness-of-fit

In this section, we propose a methodology to perform a goodness-of-fit test on a multivariate time series, by using the Rosenblatt's transform. First, following Rémillard (2013), under the general regime-switching model described in Section 3.1, the conditional density f_t of Y_t given $Y_1, \ldots, Y_{t-1}$ can be expressed as a mixture viz.

$$f_t(y_t | y_1, \ldots, y_{t-1}) = \sum_{i=1}^{\ell} f^{(i)}(y_t) \sum_{j=1}^{\ell} \eta_{t-1}(j) P_{ji} = \sum_{i=1}^{\ell} f^{(i)}(y_t) W_{t-1}(i), \tag{3}$$

where $f^{(i)} = g_{\boldsymbol{\beta}_i}$ and

$$W_{t-1}(i) = \sum_{j=1}^{\ell} \eta_{t-1}(j) P_{ji}, \quad i \in \{1 \ldots \ell\}, \tag{4}$$

$$\eta_t(j) = \frac{f^{(j)}(y_t)}{Z_{t|t-1}} \sum_{i=1}^{\ell} \eta_{t-1}(i) P_{ij}, \quad j \in \{1, \ldots, \ell\}, \tag{5}$$

$$Z_{t|t-1} = \sum_{j=1}^{\ell} f^{(j)}(y_t) \sum_{i=1}^{\ell} \eta_{t-1}(i) P_{ij}. \tag{6}$$

Note that formulas (3)–(6) also hold for univariate Gaussian HMM; in this case, $f^{(j)}$ is the Gaussian density with mean μ_j and variance σ_j^2. Next, let $i \in \{1, \ldots, \ell\}$ be fixed and suppose that $Z = (Z_1, \ldots, Z_d)$ has density $f^{(i)}$. For any $q \in \{1, \ldots, d\}$, denote by $f_{1:q}^{(i)}$ the density of

$(Z_1, \ldots, Z_q)$. Also, let $f_q^{(i)}$ be the conditional density of Z_q given $Z_1, \ldots, Z_{q-1}$. Further denote by $F_q^{(i)}$ the distribution function corresponding to density $f_q^{(i)}$. The Rosenblatt's transform Ψ_t corresponding to the density (3) conditional on $y_1, \ldots, y_{t-1} \in \mathbb{R}^d$ is given by

$$\Psi_t^{(1)}(y_{1t}) = \sum_{i=1}^{\ell} W_{t-1}(i) F_1^{(i)}(y_{1t}), \tag{7}$$

and for $q \in \{2, \ldots, d\}$,

$$\Psi_t^{(q)}(y_{1t}, \ldots, y_{qt}) = \frac{\sum_{i=1}^{\ell} W_{t-1}(i) f_{1:q-1}^{(i)}(y_{1t}, \ldots, y_{q-1,t}) F_q^{(i)}(y_{tq})}{\sum_{i=1}^{\ell} W_{t-1}(i) f_{1:q-1}^{(i)}(y_{1t}, \ldots, y_{q-1,t})}. \tag{8}$$

Suppose now that $\mathbf{U}_1, \ldots, \mathbf{U}_n$ is a random sample of size n of d-dimensional vectors drawn from a joint continuous distribution $\mathbf{P}$ belonging to a parametric family of regime-switching copula models with ℓ regimes. Formally, the hypothesis to be tested is

$$\mathcal{H}_0 : \mathbf{P} \in \mathcal{P} = \{\mathbf{P}_{\boldsymbol{\theta}}; \boldsymbol{\theta} \in \mathcal{O}\} \quad vs \quad \mathcal{H}_1 : \mathbf{P} \notin \mathcal{P}.$$

Under $\mathcal{H}_0$, it follows that $\mathbf{V}_1 = \Psi_1(\mathbf{U}_1, \boldsymbol{\theta}), \mathbf{V}_2 = \Psi_2(\mathbf{U}_1, \mathbf{U}_2, \boldsymbol{\theta}), \ldots, \mathbf{V}_n = \Psi_n(\mathbf{U}_1, \ldots, \mathbf{U}_n, \boldsymbol{\theta})$ are iid uniform over $(0,1)^d$, where $\Psi_1(\cdot, \boldsymbol{\theta}), \ldots, \Psi_n(\cdot, \boldsymbol{\theta})$ are the Rosenblatt's transforms for the true parameters $\boldsymbol{\theta} \in \mathcal{O}$. However, $\boldsymbol{\theta}$ must be estimated, say by $\boldsymbol{\theta}_n$. Also, the random vectors $\mathbf{U}_1, \ldots, \mathbf{U}_n$ are not observable, so they must be replaced by the normalized ranks $\mathbf{u}_{n,t}$ of the pseudo-observations $\mathbf{e}_{n,t}$, $t \in \{1, \ldots, n\}$. Then, define the pseudo-observations $\mathbf{V}_{n,t} = \Psi_t(\mathbf{u}_{n,t}, \boldsymbol{\theta}_n)$, $t \in \{1, \ldots, n\}$, and for any $\mathbf{u} = (u_1, \ldots, u_d) \in [0,1]^d$, define the empirical process $D_n(\mathbf{u}) = \frac{1}{n}\sum_{t=1}^n \prod_{j=1}^d \mathbb{1}(V_{n,jt} \le u_j)$. To test $\mathcal{H}_0$ against $\mathcal{H}_1$, Genest et al. (2009) suggest to use the Cramér-von Mises type statistic S_n defined by

$$\begin{aligned} S_n = \mathcal{S}_n(\mathbf{V}_{n,1}, \ldots, \mathbf{V}_{n,n}) &= n \int_{[0,1]^d} \left\{ D_n(\mathbf{u}) - \prod_{j=1}^d u_j \right\}^2 d\mathbf{u} \\ &= \frac{1}{n}\sum_{t=1}^n \sum_{i=1}^n \prod_{q=1}^d \{1 - \max(V_{n,qt}, V_{n,qi})\} - \frac{1}{2^{d-1}} \sum_{t=1}^n \prod_{q=1}^d \left(1 - V_{n,qt}^2\right) + \frac{n}{3^d}. \end{aligned}$$

We can interpret S_n as the distance of our empirical distribution and the independence copula. Since $\mathbf{V}_{n,t}$, $t \in \{1, \ldots, n\}$ are almost uniformly distributed over $(0,1)^d$ under $\mathcal{H}_0$, large values of S_n lead to the rejection of the null hypothesis. Unfortunately, the limiting distribution of the test statistic will depend on the unknown parameter $\boldsymbol{\theta}$, but it does not depend on the estimated parameters of the univariate time series (Nasri and Rémillard, 2019). Therefore, we will use the parametric bootstrap described in Algorithm 1 to estimate P-values.

Algorithm 1 *For a given number of regimes ℓ, get estimator $\boldsymbol{\theta}_n$ of $\boldsymbol{\theta}$ using the EM algorithm described in Section 3.1, applied to the pseudo-observations $\mathbf{u}_{n,t}$, $t \in \{1, \ldots, n\}$. Then compute the statistic $S_n = \mathcal{S}_n(\mathbf{V}_{n,1}, \ldots, \mathbf{V}_{n,n})$, using the pseudo-observations $\mathbf{V}_{n,t} = \Psi_t(\mathbf{u}_{n,t}, \boldsymbol{\theta}_n)$, $t \in \{1, \ldots, n\}$. Then for $k = 1, \ldots, B$, B large enough, repeat the following steps:*

- *Generate a random sample $\mathbf{U}_1^*, \ldots, \mathbf{U}_n^*$ from distribution $\boldsymbol{P}_{\boldsymbol{\theta}_n}$, i.e., from a regime-switching copula model with parameter $\boldsymbol{\theta}_n$.*
- *Get the estimator $\boldsymbol{\theta}_n^*$ from $\mathbf{U}_1^*, \ldots, \mathbf{U}_n^*$.*

- *Compute the normalized ranks* $\mathbf{u}^*_{n,1}, \ldots, \mathbf{u}^*_{n,n}$ *from* $\mathbf{U}^*_1, \ldots, \mathbf{U}^*_n$.
- *Compute the pseudo-observations* $\mathbf{V}^*_{n,t} = \mathbf{\Psi}_t\left(\mathbf{u}^*_{n,t}, \boldsymbol{\theta}^*_n\right)$, $t \in \{1, \ldots, n\}$ *and calculate* $S_n^{(k)} = \mathcal{S}_n\left(\mathbf{V}^*_{n,1}, \ldots, \mathbf{V}^*_{n,n}\right)$.

Then, an approximate P*-value for the test based on the Cramér-von Mises statistic* S_n *is given by*

$$\frac{1}{B}\sum_{k=1}^{B} \mathbb{1}\left(S_n^{(k)} > S_n\right).$$

3.3. Selecting the number of regimes

There are many ways one could select the copula model and the number of regimes. In the literature on regime-switching copulas, see, e.g., Stöber and Czado (2014), Fink et al. (2017), it is often suggested to choose the model with the smallest AIC/BIC. However, there is no empirical study backing up this idea. Note also that model selection based on AIC or BIC does not guarantee that the model is correct. This is why one could also rely on the goodness-of-fit test described in the previous section. In the case of a Gaussian HMM, Rémillard (2013) suggested to choose the number of regimes ℓ^* as the first ℓ for which the P-value is larger than 5%. We will also use the same idea here. The consistency of all these procedures is investigated numerically in Section 4.

4. NUMERICAL EXPERIMENTS

In this section we consider Monte Carlo experiments for assessing the power of the proposed goodness-of-fit test and the validity of the procedures proposed in Section 3.3. To this end, we generated random samples of size $n \in \{250, 500, 1000\}$ from four regime-switching bivariate copula families: Clayton, Frank, Gaussian, and Gumbel with one, two, and three regimes. For the 1-regime model, all copulas have a Kendall's $\tau = .5$, while for the 2-regime copula, we took $\tau = .25$ for regime 1 and $\tau = 0.75$ for regime 2, with transition matrix $P = \begin{pmatrix} 0.25 & 0.75 \\ 0.50 & 0.50 \end{pmatrix}$. Finally, for the 3-regime copula, we took $\tau = .25$ for regime 1, $\tau = 0.5$ for regime 2, and $\tau = 0.75$ for regime 3, with transition matrix $P = \begin{pmatrix} 0.5 & 0.25 & 0.25 \\ 0.3 & 0.5 & 0.2 \\ 0.1 & 0.3 & 0.6 \end{pmatrix}$.

For each sample size and for each model with a given number of regimes, we performed 1000 replications and in each replication, when needed, $B = 100$ bootstrap samples were used to compute the P-value of the test statistic S_n.

In the first set of experiments, we assess the power of the proposed goodness-of-fit test for different copula families. To this end, we fix the number of regimes $\ell \in \{2, 3\}$ and vary the copula family. The results displayed in Table 1 show that for two regimes, the empirical levels are not significantly different from the target value of 5%. For three regimes, for all but the Frank copula with $n = 250$, the empirical levels are not significantly different from 5%. Next, for two and three regimes, the estimated power is quite good, and as expected, it increases with the sample size. From these results, we may conclude that the goodness-of-fit test can distinguish between copula families when the number of regimes is fixed.

In the second set of experiments, we assess the power of the goodness-of-fit test for different regimes. To this end, we fix the copula family and we vary the number of regimes $\ell \in \{1, 2, 3\}$. We observe from Table 2 that for all models, the empirical levels are not significantly different

TABLE 1: Percentage of rejection of $\mathcal{H}_0$ at the 5% level for copula models with $\ell \in \{2, 3\}$ regimes, with $N = 1000$ replications and $B = 100$ bootstrap samples.

Copula family under H_0

	$\ell = 2$				$\ell = 3$			
H_1	Clayton	Frank	Gaussian	Gumbel	Clayton	Frank	Gaussian	Gumbel
$n = 250$								
Clayton	**7.6**	45.8	54.9	89.8	**3.9**	65.1	51.9	92.5
Frank	24.4	**4.8**	14.5	36.1	21.7	**10.5**	10.2	36.2
Gaussian	20.1	8.2	**4.8**	18.6	23.2	19.9	**3.3**	16.1
Gumbel	41.3	16.7	8.5	**4.4**	36.0	29.2	14.7	**5.4**
$n = 500$								
Clayton	**6.5**	77.2	92.5	100	**5.20**	94.6	93.6	100
Frank	65.3	**5.5**	31.6	80.3	75.4	**8.0**	20.6	85.8
Gaussian	60.2	12.3	**5.5**	39.4	70.3	25.6	**4.6**	43.7
Gumbel	89.4	16.7	18.2	**6.3**	90.3	47.3	28.7	**4.7**
$n = 1000$								
Clayton	**5.3**	99.5	100	100	**5.4**	100.0	100.0	100.0
Frank	99.0	**5.3**	67.8	100	99.9	**7.0**	58.7	99.8
Gaussian	97.3	29.0	**5.8**	77.3	70.3	24.8	**3.6**	43.6
Gumbel	100	43.3	41.5	**5.4**	99.9	82.0	66.0	**6.1**

from the target value of 5%. Also, when the true number of regimes is one, the percentage of rejection of the null hypothesis of two or three regimes is also about 5%. Next, when the true number of regimes is two or three, then the null hypothesis of one regime is easily rejected, while generally, the percentage of rejection of the hypothesis of three (resp. two) regimes is about 5% when there are in fact two (resp. three) regimes. We can conclude that when there are more than one regime, the goodness-of-fit test rejects easily the null hypothesis of one regime. However, when the true number of regimes is $\ell > 2$, to get a good power for rejecting k regimes, with $1 < k < \ell$, one needs a large sample size. Furthermore, it seems that the goodness-of-fit test is likely to accept a model with more regimes than necessary. This justifies that we should select the least number of regimes with a P-value larger than 5%. This procedure is investigated next when there are one or two regimes. The results, displayed in Table 3, show that the proposed methodology works fine, especially when the sample size is large enough. We also tried this method of selection when there are three regimes but the results were not satisfactory, because most of the time, two regimes were selected, in agreement with the results in Table 2.

Finally, we repeated the second set of experiments using the AIC and BIC criteria instead of the goodness-of-fit test for a sample size of $n = 1000$ only. The results are given in Table 4 and they show that when the true model has one or two regimes, both criteria are quite good, the better one being the BIC. However, when the model has three regimes, the true number of regimes is almost never discovered with the BIC, while the percentage is a bit better with the AIC. These results are similar to those obtained in Table 2 by using our goodness-of-fit test.

From all these results, trying to distinguish between two and three regimes seems illusory when $n \leq 1000$. We checked the results of the estimation procedure when there are $\ell \geq 3$ regimes and the estimation errors are too large, because in this case, the number of parameters to estimate is at least ℓ^2. We believe that in increasing the sample size enough, the estimation errors would be smaller and one could then distinguish between two or three regimes. However,

TABLE 2: Percentage of rejection of $\mathcal{H}_0$ at the 5% level for the regime-switching Clayton, Frank, Gaussian, and Gumbel copula models with one, two, and three regimes, using $N = 1000$ replications and $B = 100$ bootstrap samples.

Copula family under H_0

	Clayton			Frank		
H_1	1 regime	2 regimes	3 regimes	1 regime	2 regimes	3 regimes
$n = 250$						
1 regime	**4.5**	3.8	4.0	**5.8**	5.3	6.4
2 regimes	99.6	**7.0**	4.5	75.4	**5.0**	10.2
3 regimes	79.7	5.6	**4.2**	36.3	7.0	**11.6**
$n = 500$						
1 regime	**4.2**	3.9	3.4	**5.4**	6.0	6.1
2 regimes	100.0	**7.0**	5.4	96.4	**5.2**	6.9
3 regimes	98.3	6.2	**5.6**	66.9	5.2	**7.9**
$n = 1000$						
1 regime	**4.4**	4.5	4.5	**5.5**	5.1	5.3
2 regimes	100.0	**7.3**	6.1	100.0	**5.7**	4.8
3 regimes	100.0	6.4	**5.3**	92.3	5.5	**6.1**
	Gaussian			**Gumbel**		
H_1	1 regime	2 regimes	3 regimes	1 regime	2 regimes	3 regimes
$n = 250$						
1 regime	**5.0**	6.2	6.7	**6.7**	6.0	4.9
2 regimes	94.4	**5.1**	4.6	59.2	**4.7**	4.5
3 regimes	57.6	4.8	**4.4**	26.1	4.8	**4.6**
$n = 500$						
1 regime	**5.4**	5.2	5.3	**5.0**	4.3	4.6
2 regimes	99.9	**5.4**	4.9	92.3	**5.2**	4.2
3 regimes	87.9	5.9	**4.5**	46.4	5.4	**4.9**
$n = 1000$						
1 regime	**4.9**	6.1	5.0	**5.6**	5.8	6.0
2 regimes	100.0	**4.4**	4.5	99.8	**7.1**	5.1
3 regimes	99.9	4.3	**4.4**	80.3	4.6	**4.2**

numerical experiments to prove this would require months of computations. In the end, we recommend to combine goodness-of-fit tests and the AIC/BIC criteria, to ensure at least that the chosen model is valid.

Note that one should also expect better results for the power of the goodness-of-fit test by taking a larger number of bootstrap samples. Here, in order to build the tables in a reasonable time, we restricted ourselves to bootstrap samples of size $B = 100$, which is quite small. In real life, we do not repeat the experiments $N = 1000$ times, so we may use at least $B = 1000$, especially when the P-value is around 5%. Furthermore, we did not consider the regime-switching Student copula since it has more parameters and, according to Table 6, the computation time is approximately 10 times longer for the Student family than for the Gumbel family, which has the longest computation time amongst the four other families.

TABLE 3: Estimation of the number of regimes ℓ^* for $N = 1000$ replications, using $B = 100$ bootstrap samples. Boldface values indicate the percentage of the correct choice of the number of regimes.

	Copula family							
	Clayton		Frank		Gaussian		Gumbel	
	Number of regimes		Number of regimes		Number of regimes		Number of regimes	
ℓ^*	1	2	1	2	1	2	1	2
$n = 250$								
1	**94.4**	0.8	**94.8**	25.1	**95.3**	5.4	**95.2**	37.5
2	2.3	**91.8**	1.5	**57.0**	1.6	**88.7**	1.5	**58.0**
3	0.5	2.9	0.7	2.3	0.4	2.3	0.4	1.2
≥ 4	2.8	4.5	3.0	15.6	2.7	3.6	3.6	3.3
$n = 500$								
1	**93.8**	0	**94.8**	2.4	**93.6**	0.1	**95.2**	8.6
2	2.4	**92.3**	1.6	**76.4**	1.9	**95.4**	1.2	**86.7**
3	0.4	1.8	0.7	2.9	0.5	1.6	0.6	1.0
≥ 4	3.4	5.9	2.9	18.3	4	2.9	3.0	3.7
$n = 1000$								
1	**95.0**	0.0	**94.1**	0.0	**95.6**	0.0	**94.9**	0.1
2	2.2	**94.3**	1.7	**79.5**	0.8	**95.2**	1.2	**93.7**
3	0.6	1.7	0.3	2.1	0.7	1.0	0.4	1.0
≥ 4	2.2	4.0	3.9	18.4	2.9	3.8	3.5	5.2

5. APPLICATION TO OPTION PRICING

In this application, we want to evaluate a European put-on-max option on Amazon (amzn) and Apple (aapl) stocks. The payoff of this option is given by $\Phi(s_1, s_2) = \max\{K - \max(s_1, s_2), 0\}$, where s_1 and s_2 are the values of the stocks at the maturity of the option, normalized to start at 1\$, and K is the strike price. An investor would be interested in such an option to protect the returns of his assets, since he will exercise the option if both returns are lower than $\log K$. Also this option is cheaper than a put-on-min. In order to evaluate this option, we need first to find the joint distribution of both assets. Next, we will choose an appropriate risk neutral probability measure.

5.1. Joint distribution

The first step is to fit dynamic models for the univariate time series. To this end, we used the adjusted prices of Amazon and Apple from January 1, 2015 to June 29, 2018. The sample size is 880 observations for each time series. The 879 daily log-returns of the stocks are shown in Figure 1. Since van den Goorbergh et al. (2005) used GARCH models for the log-returns of the assets they considered, we also tried to fit GARCH(p,q) models with Gaussian innovations, but we rejected the null hypothesis for $p, q \leq 3$. We next tried to fit Gaussian HMMs to the log-returns. Using the selection procedure described in Section 3.2, we obtained a Gaussian HMM with three regimes for the daily log-returns of Amazon as well as for the daily log-returns of Apple. Here, the P-values are 38.8% and 15.1% respectively, computed using $B = 1000$ bootstrap samples. The estimated parameters for both time series are given in Table 5. Note that the regimes are ordered by their stationary distribution (ν), meaning that the least frequent regime is 1, and the

TABLE 4: Percentage of selection of the number of regimes $\ell \in \{1, 2, 3\}$ for a sample size $n = 1000$ and $N = 1000$ replications, based on the AIC and BIC criteria. Boldface values indicate the percentage of the correct choice of the number of regimes.

Copula family under H_0

	Clayton					
		AIC			BIC	
H_1	1 regime	2 regimes	3 regimes	1 regime	2 regimes	3 regimes
1 regime	**97.4**	2.6	0	**100**	0	0
2 regimes	0	**96.4**	3.6	0	**99.9**	0.1
3 regimes	0	90.2	**9.8**	0	99.9	**0.1**
	Frank					
		AIC			BIC	
H_1	1 regime	2 regimes	3 regimes	1 regime	2 regimes	3 regimes
1 regime	**97.8**	2.2	0	**100**	0	0
2 regimes	0	**97.2**	2.8	0	**100**	0
3 regimes	0	98.2	**1.8**	4.4	95.6	**0**
	Gaussian					
		AIC			BIC	
H_1	1 regime	2 regimes	3 regimes	1 regime	2 regimes	3 regimes
1 regime	**98.4**	1.6	0	**100**	0	0
2 regimes	0	**98.1**	1.9	0	**100**	0
3 regimes	0	95.3	**4.7**	0	100	**0**
	Gumbel					
		AIC			BIC	
H_1	1 regime	2 regimes	3 regimes	1 regime	2 regimes	3 regimes
1 regime	**97.6**	2.4	0	**100**	0	0
2 regimes	0	**96.2**	3.8	0	**100**	0
3 regimes	0	96.4	**3.6**	0.9	99.1	**0**

most frequent regime is 3. As can be seen from Figure 1, for Amazon, regime 1 corresponds to large positive returns with a frequency of 2%, while for Apple, regime 1 consists in large negative values with a frequency of 10%. So in both cases, regime 1 is not a persistent state, while the two other regimes are much more persistent. For the two stocks, since regime 2 has always a negative mean, it can be interpreted as a bear market, while for regime 3, the mean μ_3 is positive, as well as $\mu_3 - \sigma_3^2/2$, so this regime can be interpreted as a bull market.

From now on, let X_{1t} denotes the log-returns of Amazon and let X_{2t} denotes the log-returns of Apple, and let F_{1t} and F_{2t} be the conditional distributions of X_{1t} and X_{2t} given the past observations, corresponding to the densities defined by Equation (3). Further set $U_{1t} = F_{1t}(X_{1t})$ and $U_{2t} = F_{2t}(X_{2t})$. As defined in Section 3, let $e_{n,jt} = F_{n,jt}(X_{jt})$, $j = 1, 2$, be the pseudo-observations, where $F_{n,jt}$ is the conditional distribution function computed with the parameters of Table 5. The graph of the normalized ranks of $\mathbf{u}_{n,t} = (u_{n,1t}, u_{n,2t})$ is displayed in Figure 2. Next, in order to select the appropriate regime-switching copula model, we performed goodness-of-fit tests using $B = 1000$ bootstrap samples to select the copula family and the number of regimes amongst the Clayton, Frank, Gaussian, Gumbel and Student families. Note that for the 1-

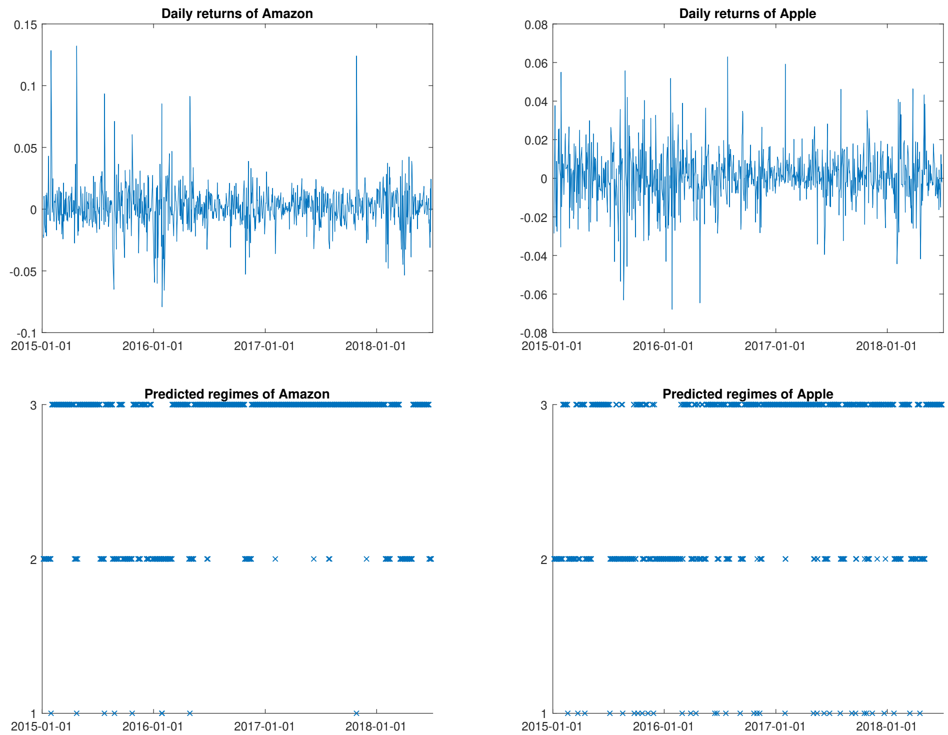

FIGURE 1: Daily log returns and predicted regimes for Amazon and Apple.

TABLE 5: Estimated parameters for the log-returns of Amazon and Apple, using Gaussian HMM. Here, ν is the stationary distribution of the regimes, and Q is the transition matrix.

	Amazon			Apple		
	Regime			Regime		
Parameter	1	2	3	1	2	3
$\mu \times 10^{-2}$	4.4122	-0.1179	0.1892	-0.2777	-0.0433	0.2215
$\sigma \times 10^{-3}$	57.2765	22.9170	10.1334	2.4846	20.6456	8.8176
ν	0.0199	0.2574	0.7227	0.1015	0.3894	0.5091
Q	$\begin{pmatrix} 0.1572 & 0.3978 & 0.4450 \\ 0.0545 & 0.8849 & 0.0606 \\ 0.0038 & 0.0300 & 0.9662 \end{pmatrix}$			$\begin{pmatrix} 0.0674 & 0.4154 & 0.5172 \\ 0.0000 & 0.8788 & 0.1212 \\ 0.1859 & 0.0098 & 0.8042 \end{pmatrix}$		

regime and 2-regime Student copulas, and for the 2-regime Gaussian copula, we took $B = 10000$ in order to get more precise results. The corresponding P-values are given in Table 6, together with the computation time in seconds for $B = 1000$ bootstrap samples,

and the BIC values. From this table, based on the P-values, we can see that the 2-regime Gaussian and the 2-regime Student copula models are valid, while the 1-regime Student copula model is almost acceptable. However, the estimated degrees of freedom of the 2-regime Student copula model are very large, indicating that it is indeed a 2-regime Gaussian copula. We then

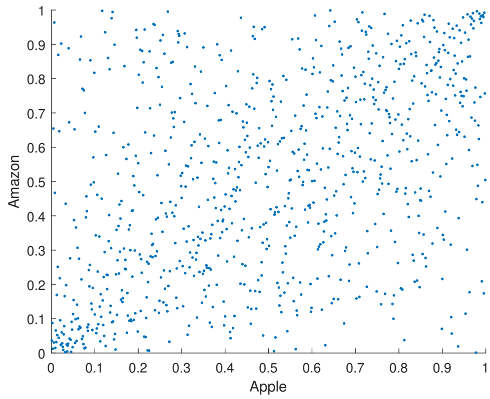

FIGURE 2: Scatter plot of the normalized ranks of the pseudo-observations $\mathbf{u}_{n,t}$ for Apple and Amazon.

restricted the degrees of freedom to be less than 25, and we obtained the upper bounded as their estimations. Looking at the BIC values, we see that the smallest one is for the 2-regime Gaussian copula. Based on these results, we choose the 2-regime Gaussian copula as the best model. Its estimated parameters appear in Table 7.

TABLE 6: P-values (in percentage) of the different regime-switching copula families, together with the computation time in seconds and the BIC criterion.

	Copula family									
	Clayton		Frank		Gaussian		Gumbel		Student	
	Number of regimes		Number of regimes		Number of regimes		Number of regimes		Number of regimes	
	1	2	1	2	1	2	1	2	1	2
P-value	0.0	1.0	0.6	0.8	0.0	**9.8**	0.0	0.0	**4.4**	**10.1**
Sec.	195	1272	128	490	175	712	198	1342	1715	15386
BIC	-174.0	-181.3	-188.8	-190.8	-179.6	-205.3	-168.7	-182.3	-201.8	-191.8

TABLE 7: Estimated parameters for the 2-regime Gaussian copula. Here, τ is Kendall's tau, $\rho = \sin(\pi\tau/2)$ is the correlation coefficient of the copula, ν is the stationary distribution of the regimes, and P is the transition matrix.

Parameter	Regime 1	Regime 2
τ	0.0859	0.5816
ρ	0.1346	0.7917
ν	0.5209	0.4791
P	$\begin{pmatrix} 0.7414 & 0.2586 \\ 0.2812 & 0.7188 \end{pmatrix}$	

5.2. Bivariate option pricing

In order to price an option with payoff $\Phi(S_{1n}, S_{2n})$ over n trading days, we perform a Monte Carlo simulation under a risk neutral measure. First, as in van den Goorbergh et al. (2005), we assume that the selected regime-switching copula model with parameters appearing in Table 7 is also valid under the risk neutral measure. Next, for the dynamic models of both time series, we assume that we still have Gaussian HMM, but with new parameters, namely $\tilde{\mu}_{jk} = r - \frac{\sigma_{jk}^2}{2}$, $\tilde{\sigma}_{jk} = \sigma_{jk}$, and $\tilde{Q}^{(j)} = Q^{(j)}$, where r is the risk free daily interest rate. This way, under the risk neutral measure, the discounted prices $e^{-rt}S_{jt} = e^{\sum_{i=1}^{t}(X_{ji}-r)}$ form a martingale, for each $j = 1, 2$.

The following steps illustrate the procedure to evaluate a European option with payoff Φ in the case of a general regime-switching copula with ℓ regimes, where each univariate time series is modeled by a Gaussian HMM with m_j regimes and parameters $\tilde{\mu}_{j1}, \ldots, \tilde{\mu}_{j\ell_j}$, $\tilde{\sigma}_{j1}, \ldots, \tilde{\sigma}_{j\ell_j}$, $\tilde{Q}^{(j)}$:

1. Generate $\mathbf{U}_t$, $t \in \{1, \ldots, n\}$, from the regime-switching copula model.
2. For $t \in \{1, \ldots, n\}$, and $j = 1, 2$, compute the conditional distribution function F_{jt} under the risk neutral measure, and set $X_{jt} = F_{jt}^{-1}(U_{jt})$.
3. For $j = 1, 2$, compute $S_{jn} = e^{\sum_{t=1}^{n} X_{ji}}$.
4. Repeat N times steps $1 - 3$, in order to get N independent values of (S_{1n}, S_{2n}).

The value of the option is then approximated by the average of the discounted values $e^{-rn}\Phi(S_{1n}, S_{2n})$.

To evaluate the put-on-max option, we used $N = 10000$ simulations, with a maturity of $n = 20$ trading days and a risk free rate $r = 4\%$.

Figure 3 displays the price of the option as a function of the strike K for the best model, i.e., the 2-regime Gaussian copula, versus the other four 2-regime copula families. As expected, the prices given by the 2-regime Gaussian copula and the 2-regime Student are almost identical. Note that the prices of the 2-regime Frank and Gumbel copulas are always lower than those of the 2-regime Gaussian copula, while those of the 2-regime Clayton copula are higher than those of the 2-regime Gaussian copula when the strike value is lower than 1.0005, and are lower when the strike is larger than 1.0005.

6. CONCLUSION

In this paper, for a regime-switching copula model, we proposed a methodology based on a goodness-of-fit test to select the copula family and the number of regimes. This methodology can also be used for mixtures of copula models, as well as for univariate HMM. We performed Monte Carlo simulations with a sample size $n \in \{250, 500, 1000\}$, and we showed that the level of the goodness-of-fit test is correct and that it is powerful enough to distinguish between regime-switching copula families and also to detect if there is more than one regime. The proposed procedure for selecting the number of regimes works when the sample size is large enough and there are less than three regimes. For three regimes or more, the sample size must be larger than 1000. As an example of application, we showed how to evaluate a European put-on-max option, but the proposed methodology can also be applied to a wide range of options on multivariate assets. The empirical results emphasize the importance of choosing the correct copula family.

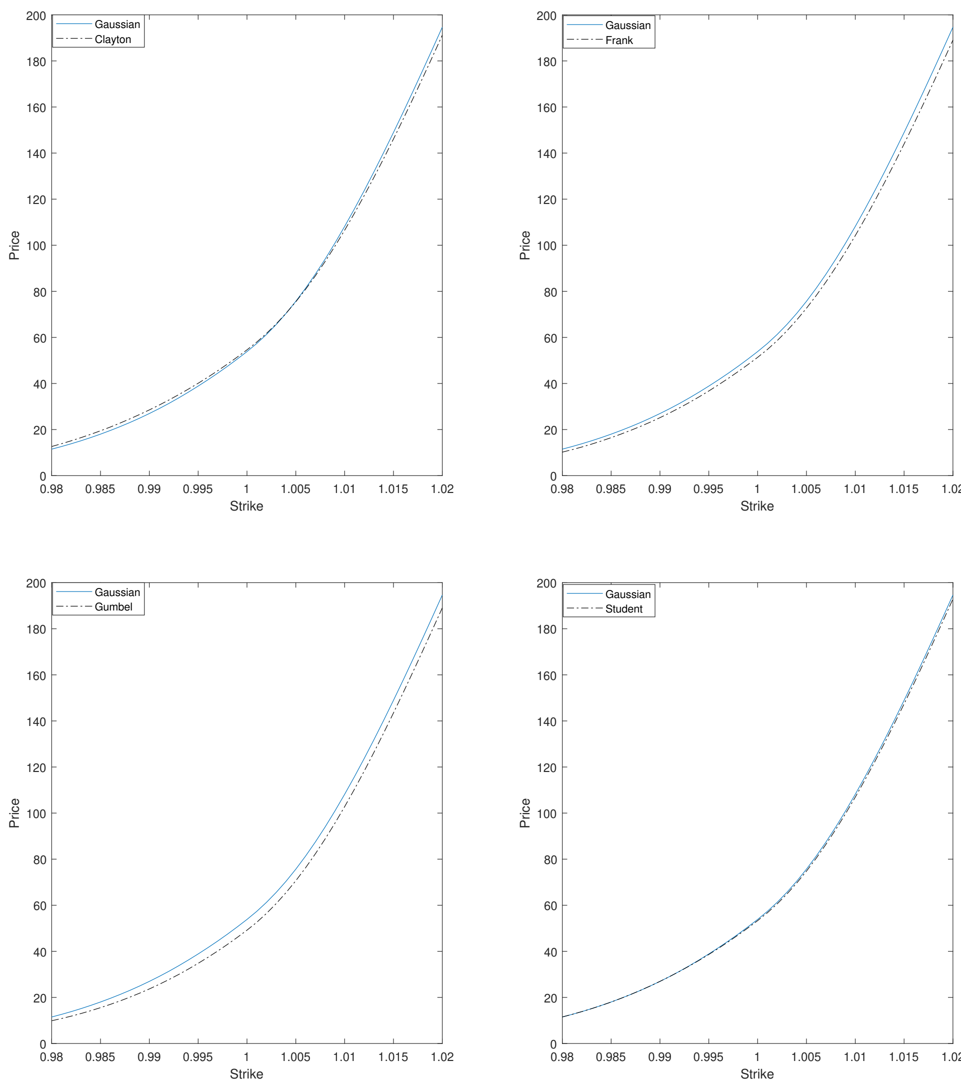

FIGURE 3: Comparison of put-on-max prices for $n = 20$ trading days maturity, as a function of the strike, between a 2-regime Gaussian copula and 2-regime Clayton, Frank, Gumbel and Student copula models.

Acknowledgments

The authors are grateful to the Guest Editor, Cody Hyndman and two anonymous referees for their comments and suggestions. Partial funding in support of this work was provided by the Natural Sciences and Engineering Research Council of Canada (Grant 04430–2014), the Canadian Statistical Sciences Institute (postdoctoral fellowship), the Fonds de recherche du Québec – Nature et technologies (2015–PR–183236 and postdoctoral fellowship 259667), and the Groupe d'études et de recherche en analyse des décisions (postdoctoral fellowship).

BIBLIOGRAPHY

Adams, Z., Füss, R., and Glück, T. (2017). Are correlations constant? Empirical and theoretical results on popular correlation models in finance. *Journal of Banking & Finance*, 84:9–24.

Cappé, O., Moulines, E., and Rydén, T. (2005). *Inference in Hidden Markov Models*. Springer Series in Statistics. Springer, New York.

Chen, X. and Fan, Y. (2006). Estimation of copula-based semiparametric time series models. *Journal of Econometrics*, 130:307–335.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B*, 39:1–38.

Du, Z. (2016). Nonparametric bootstrap tests for independence of generalized errors. *The Econometrics Journal*, 19(1):55–83.

Engle, R. (2002). Dynamic conditional correlation. *Journal of Business & Economic Statistics*, 20(3):339–350.

Fermanian, J.-D. and Lopez, O. (2018). Single-index copulas. *Journal of Multivariate Analysis*, 165:27–55.

Fink, H., Klimova, Y., Czado, C., and Stöber, J. (2017). Regime switching vine copula models for global equity and volatility indices. *Econometrics*, 5(1):1–38.

Genest, C., Rémillard, B., and Beaudoin, D. (2009). Omnibus goodness-of-fit tests for copulas: A review and a power study. *Insurance Math. Econom.*, 44:199–213.

Härdle, W. K., Okhrin, O., and Wang, W. (2015). Hidden Markov structures for dynamic copulae. *Econometric Theory*, 31(5):981—-1015.

Nasri, B. and Rémillard, B. (2019). Copula-based dynamic models for multivariate time series. *Journal of Multivariate Analysis*, 172:107–121.

Nasri, B. R., Rémillard, B. N., and Bouezmarni, T. (2019). Semi-parametric copula-based models under non-stationarity. *Journal of Multivariate Analysis*, 173:347–365.

Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics*, 2(1):130–168.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review*, 47(2):527–556.

Rémillard, B. (2013). *Statistical Methods for Financial Engineering*. Chapman and Hall/CRC Financial Mathematics Series. Taylor & Francis.

Rémillard, B. (2017). Goodness-of-fit tests for copulas of multivariate time series. *Econometrics*, 5(1):13.

Rémillard, B., Hocquard, A., Lamarre, H., and Papageorgiou, N. A. (2017). Option pricing and hedging for discrete time regime-switching model. *Modern Economy*, 8:1005–1032.

Stöber, J. and Czado, C. (2014). Regime switches in the dependence structure of multidimensional financial data. *Computational Statistics & Data Analysis*, 76:672–686. CFEnetwork: The Annals of Computational and Financial Econometrics.

Thioub M. Y., Nasri, B. R., Pieugueu, R., and Rémillard, B. N. (2019), *HMMcopula*, R package version 1.0.3.

van den Goorbergh, R. W. J., Genest, C., and Werker, B. J. M. (2005). Bivariate option pricing using dynamic copula models. *Insurance: Mathematics and Economics*, 37:101–114.

APPENDIX

Estimation for general regime-switching models

E-Step

Set $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\beta}}_1, \ldots, \tilde{\boldsymbol{\beta}}_\ell, \tilde{P})$. *Then, according to* (Rémillard, 2013, *Appendix 10.A*),

$$
\begin{aligned}
Q_y(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}) &= \mathbb{E}_{\boldsymbol{\theta}}\{\log f_{\tilde{\boldsymbol{\theta}}}(\tau, Y)|Y = y\} \\
&= \sum_{t=1}^{n} \sum_{j\in S} \sum_{k\in S} \mathbb{P}_{\boldsymbol{\theta}}(\tau_{t-1} = j, \tau_t = k|Y = y) \log \tilde{P}_{jk} + \sum_{t=1}^{n} \sum_{j\in S} \mathbb{P}_{\boldsymbol{\theta}}(\tau_t = j|Y = y) \log g_{\tilde{\boldsymbol{\beta}}_j}(y_t) \\
&= \sum_{t=1}^{n} \sum_{j\in S} \sum_{k\in S} \Lambda_{\boldsymbol{\theta},t}(j, k) \log \tilde{P}_{jk} + \sum_{t=1}^{n} \sum_{j\in S} \lambda_{\boldsymbol{\theta},t}(j) \log g_{\tilde{\boldsymbol{\beta}}_j}(y_t),
\end{aligned}
$$

where $\lambda_{\boldsymbol{\theta},t}(j) = P(\tau_t = j|Y = y)$ *and* $\Lambda_{\boldsymbol{\theta},t}(j, k) = P(\tau_{t-1} = j, \tau_t = k|Y = y)$, *for all* $t \in \{1, \ldots, n\}$ *and* $j, k \in S$. *Next, define for all* $j \in S$, $\bar{\eta}_{\boldsymbol{\theta},n}(j) = 1/\ell$, $\eta_{\boldsymbol{\theta},0}(j) = 1/\ell$,

$$
\begin{aligned}
\bar{\eta}_{\boldsymbol{\theta},t}(j) &= Prob(\tau_t = j|y_{t+1}, \ldots, y_n), \quad t = 1, \ldots, n-1, \\
\eta_{\boldsymbol{\theta},t}(j) &= Prob(\tau_t = j|y_1, \ldots, y_t), \quad t = 1, \ldots, n.
\end{aligned}
$$

It follows easily that for $t = 1, \ldots, n$, $\eta_t(j) = \frac{g_{\boldsymbol{\beta}_j}(y_t)}{Z_{t|t-1}} \sum_{i=1}^{\ell} \eta_{t-1}(i) P_{ij}$, *where*

$$
Z_{t|t-1} = \sum_{j=1}^{\ell} g_{\boldsymbol{\beta}_j}(y_t) \sum_{i=1}^{\ell} \eta_{t-1}(i) P_{ij}.
$$

Next, for all $i \in \{1, \ldots, l\}$, *and for all* $t = 0, \ldots, n-1$,

$$
\begin{aligned}
\bar{\eta}_{\boldsymbol{\theta},t}(i) &= \frac{\sum_{\beta=1}^{\ell} \bar{\eta}_{\boldsymbol{\theta},t+1}(\beta) P_{i\beta} g_{\boldsymbol{\beta}_\beta}(y_{t+1})}{\sum_{k=1}^{\ell} \sum_{\beta=1}^{\ell} \bar{\eta}_{\boldsymbol{\theta},t+1}(\beta) P_{k\beta} g_{\boldsymbol{\beta}_\beta}(y_{t+1})}, \\
\lambda_{\boldsymbol{\theta},t}(i) &= \frac{\eta_{\boldsymbol{\theta},t}(i) \bar{\eta}_{\boldsymbol{\theta},t}(i)}{\sum_{k=1}^{l} \eta_{\boldsymbol{\theta},t}(k) \bar{\eta}_{\boldsymbol{\theta},t}(k)}.
\end{aligned}
$$

Hence, for all $i, j \in \{1, \ldots, l\}$, *and for all* $t = 1, \ldots, n$,

$$
\Lambda_{\boldsymbol{\theta},t}(i, j) = \frac{P_{ij} \eta_{\boldsymbol{\theta},t-1}(i) \bar{\eta}_{\boldsymbol{\theta},t}(j) g_{\boldsymbol{\beta}_j}(y_t)}{\sum_{k=1}^{l} \sum_{\beta=1}^{l} P_{k\beta} \eta_{\boldsymbol{\theta},t-1}(k) \bar{\eta}_{\boldsymbol{\theta},t}(\beta) g_{\boldsymbol{\beta}_\beta}(y_t)}.
$$

As a result, for all $i \in \{1, \ldots, l\}$, *and for every* $t = 1, \ldots, n$, $\sum_{j=1}^{l} \Lambda_{\boldsymbol{\theta},t}(i, j) = \lambda_{\boldsymbol{\theta},t-1}(i)$.

M-Step

For this step, given $\boldsymbol{\theta}^{(k)}$, $\boldsymbol{\theta}^{(k+1)}$ *is defined as* $\boldsymbol{\theta}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} Q_y\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{(k)}\right)$. *Setting* $\lambda_t^{(k)}(i) = \lambda_{\boldsymbol{\theta}^{(k)},t}(i)$ *and* $\Lambda_t^{(k)}(i, j) = \Lambda_{\boldsymbol{\theta}^{(k)},t}(i, j)$, *it follows from Section that*

$$
\boldsymbol{\theta}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} \sum_{t=1}^{n} \sum_{i,j\in S} \Lambda_t^{(k)}(i, j) \log P_{ij} + \sum_{t=1}^{n} \sum_{i\in S} \lambda_t^{(k)}(i) \log g_{\boldsymbol{\beta}_i}(y_t).
$$

Using Lagrange multipliers, the function to maximize is $h(\boldsymbol{\theta}, \psi)$, where $\psi = (\psi_1, \ldots, \psi_\ell)$, and

$$h(\boldsymbol{\theta}, \psi) = \sum_{t=1}^{n} \sum_{i,j \in S} \Lambda_t^{(k)}(i,j) \log P_{ij} + \sum_{t=1}^{n} \sum_{i \in S} \lambda_t^{(k)}(i) \log g_{\boldsymbol{\beta}_i}(y_t) + \sum_{i=1}^{l} \psi_i \left(1 - \sum_{j=1}^{\ell} P_{ij}\right).$$

For $i, j \in S$ we have $\frac{\partial h}{\partial P_{i,j}} = \sum_{t=1}^{n} \Lambda_t^{(k)}(i,j) \frac{1}{P_{ij}} - \psi_i$. As a result, for any $i, j \in S$, the partial derivative of h with respect to P_{ij} is zero if and only if $\psi_i P_{ij} = \sum_{t=1}^{n} \Lambda_t^{(k)}(i,j)$. Summing over j yields that

$$\psi_i = \sum_{j=1}^{\ell} \psi_i P_{ij} = \sum_{j=1}^{\ell} \sum_{t=1}^{n} \Lambda_t^{(k)}(i,j) = \sum_{t=1}^{n} \lambda_{t-1}^{(k)}(i) = \sum_{t=1}^{n} \lambda_{\boldsymbol{\theta}^{(k)},t-1}(i).$$

Hence $P_{ij}^{(k+1)} = \sum_{t=1}^{n} \Lambda_t^{(k)}(i,j) \Big/ \sum_{t=1}^{n} \lambda_{t-1}^{(k)}(i)$. Also, maximizing h with respect to $\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_\ell$ amounts to maximize $\sum_{t=1}^{n} \sum_{i=1}^{\ell} \lambda_t^{(k)}(i) \log g_{\boldsymbol{\beta}_i}(y_t)$ with respect to $\boldsymbol{\beta}_i$, for all $i \in S$.

Estimation for general mixture models

This model is a particular case of regime-switching where $P_{ij} = \nu_j$, $j \in \{1, \ldots, \ell\}$. So, under this model, τ_t is a sequence of iid observations with distribution $\nu = (\nu_1, \ldots, \nu_\ell)$. The algorithm described previously can then be simplified. To this end, set $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_\ell, \nu)$. The joint density of $\tau = (\tau_1, \ldots, \tau_n)$ and Y is $f_{\boldsymbol{\theta}}(\tau, Y) = \left(\prod_{t=1}^{n} \nu_{\tau_t}\right) \times \prod_{t=1}^{n} g_{\boldsymbol{\beta}_{\tau_t}}(Y_t)$, yielding

$$\log f_{\boldsymbol{\theta}}(\tau, Y) = \sum_{t=1}^{n} \log P_{\tau_{t-1},\tau_t} + \sum_{t=1}^{n} \log g_{\boldsymbol{\beta}_{\tau_t}}(Y_t).$$

E-Step

Set $\tilde{\boldsymbol{\theta}} = (\tilde{\boldsymbol{\beta}}_1, \ldots, \tilde{\boldsymbol{\beta}}_\ell, \tilde{\nu})$. Then, according to the previous computations,

$$Q_y(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}) = \mathbb{E}_{\boldsymbol{\theta}}\{\log f_{\tilde{\boldsymbol{\theta}}}(\tau, Y) | Y = y\} = \sum_{t=1}^{n} \sum_{j \in S} \lambda_{\boldsymbol{\theta},t}(j) \left(\log \tilde{\nu}_j + \log g_{\tilde{\boldsymbol{\beta}}_j}(Y_t)\right), \qquad (1)$$

where $\lambda_{\boldsymbol{\theta},t}(j) = P_{\boldsymbol{\theta}}(\tau_t = j | Y = y) = \frac{\tilde{\nu_j} g_{\tilde{\boldsymbol{\beta}_{\tau_t}}}(y_t)}{\sum_{k=1}^{\ell} \tilde{\nu_k} g_{\tilde{\boldsymbol{\beta}_k}}(y_t)}$ for all $t \in \{1, \ldots, n\}$ and $j \in S$.

M-Step

For this step, given $\boldsymbol{\theta}^{(k)}$, $\boldsymbol{\theta}^{(k+1)}$ is defined as $\boldsymbol{\theta}^{(k+1)} = \arg\max_{\boldsymbol{\theta}} Q_y\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{(k)}\right)$. Setting $\lambda_t^{(k)}(i) = \lambda_{\boldsymbol{\theta}^{(k)},t}(i)$, one obtains

$$\begin{aligned} Q_y(\tilde{\boldsymbol{\theta}}, \boldsymbol{\theta}) &= \sum_{t=1}^{n} \sum_{j=1}^{\ell} \lambda_t^{(k)}(j) \left(\log \tilde{\nu}_j + \log g_{\tilde{\boldsymbol{\beta}}_j}(Y_t)\right) \\ &= \sum_{t=1}^{n} \sum_{j=1}^{\ell} \lambda_t^{(k)}(j) \log \tilde{\nu}_j + \sum_{t=1}^{n} \sum_{j=1}^{\ell} \lambda_t^{(k)}(j) \log g_{\tilde{\boldsymbol{\beta}}_j}(Y_t). \end{aligned}$$

For $j \in S$ *we have,* $\frac{\partial Q_y}{\partial \tilde{\nu}_j} = \frac{n}{\tilde{\nu}_j} \sum_{t=1}^{n} \lambda_t^{(k)}(j)$. *Hence* $\tilde{\nu_j}^{(k+1)} = \frac{\sum_{t=1}^{n} \lambda_t^{(k)}(j)}{n}$, $j \in \{1, \ldots, l\}$, *and*

$$\tilde{\boldsymbol{\beta}}_j^{(k+1)} = \arg\max_{\tilde{\boldsymbol{\beta}}_j} \sum_{t=1}^{n} \sum_{j=1}^{\ell} \lambda_t^{(k)}(j) \log g_{\tilde{\boldsymbol{\beta}}_j}(y_t).$$

Received 10 March 2019
Accepted 18 October 2019