Stochastic Copula Models in Continuous Time for Portfolio Risk Assessment

Aleksey Novokhatskiy and Mikhail Semenov†

Scientific Center for Information Technologies and Artificial Intelligence, Sirius University of Science and Technology, Olympic Ave., 1, Sirius, 354340, Krasnodar region, Russia.

*Corresponding author(s). E-mail(s): semenov.me@talantiuspeh.ru;

Contributing authors: aanovokhatskiy@gmail.com;

†These authors contributed equally to this work.

Abstract

This study aims to develop a copula model for constructing multivariate distributions considering the dynamics of financial assets. We propose a stochastic copula model in continuous time using an Ornstein-Uhlenbeck equation as a time-dependent copula parameter. The results demonstrate that the transition to continuous time allows one to use a Fokker-Planck equation to construct a stochastic processes with predetermined transition density, thereby reducing Monte Carlo variance. In particular, by introducing auxiliary coefficients the transition density for Ornstein-Uhlenbeck equation was modified and the explicit form of the corresponding stochastic differential equation for this density was provided. The stochastic model was applied to calculate the risk indicators Value at Risk (VaR) and Conditional value at risk (CVaR) for the cryptocurrency market.

Keywords: Stochastic copula models, Portfolio optimization, Conditional Value-at-Risk

1 Introduction

Banks, investment funds and other financial institutions constantly demonstrate significant interest in methods assessing the risks assioated with investment portfolios. In order to protect their investments, modern financial institutions often rely on indicators such as Value at Risk (VaR) and its more advanced counterpart, Conditional Value at Risk (CVaR), also known in the literature as Expected Shortfall (ES) [1].

In our previous article [2], we presented the results of modeling VaR and CVaR risk measures based on the copula approach for an optimal portfolio on short financial time series comprising 253 observations in total. Out finding showned that the R-vine copula model has a higher predictive ability compared to elliptical copula models. However, in the previous study we did not account for the possibility of dynamic changes in the relationship between variables. While this oversight may not be critical in short time series, it becomes increasingly important when modeling long time series. An urgent problem arises when modeling on a long time series. It is necessary to take into account the dynamic changes in the relationships observed in the original data.

In the review [3] the authors distinguished between parametric, semi-parametric and non-parametric copula models that accounting dynamic changes in the variable relationships. There is a brief outline of approaches.

Among the parametric models, we highlighted dynamic conditional correlation models and their further development into asymmetric dynamic conditional correlation models [4]. These models take into account the dynamics of conditional correlations through their dependence on previous shocks values (errors), adjusted for volatility. Autoregressive models and generalized autoregressive score (GAS) are based on observation-driven copula models [5], where the dependence of the copula on time is described by a parametric function with an autoregressive term. Stochastic autoregressive copula models (SCAR) [6] extend this concept by including parameter-driven model responsible for the change in the model over time, which is introduced using an independent stochastic process, rather than data as in previous models. Further development of the SCAR and GAS models have been proposed in [7, 8] using D-vine-based dynamic copula models. In these models, time-varying dependence parameters of copula model is achieved by introducing time series for copula parameters based on autoregressive moving average and generalized autoregressive conditional heteroskedasticity models. In [9] the author propose using a vine-copula in conjunction with the asymmetric power autoregressive conditional heteroscedasticity (APARCH) model to more accurately model the dependency between assets and account for complex dependencies and tail dynamics in the portfolio. In [10] authors propose copula-based trading strategy for cryptocurrency pairs and demonstrate that strategy exhibits strong risk-adjusted returns and outperforms conventional methods in terms of profitability and efficiency.

In the class of nonparametric models, we refer to the paper [11] where the authors use a quantile function (estimating time-varying quantiles) to decide on a change in the model. The review [3] presents several models where parameter responsible for the change in the model over time depends on the marginal distribution through

the conditional variance [12]. The article [13] introduces a two-dimensional dynamic stochastic copula, built on time series of different periodicities: daily for asset prices (stock index, oil price), and monthly for factors (inflation, interest rate).

The purpose of the paper is to develop a copula model for constructing multivariate distributions that account for dynamics of financial assets. The proposed model incorporates stochastic processes, such as the Ornstein-Uhlenbeck process to capture the time-varying dependencies between assets, enabling more precise risk assessments. Additionally, the proposed model includes efficiency enhancement techniques, such as importance sampling, to significantly reduce the computational complexity of the problem.

The primary objectives of this research are twofold: (1) to introduce time dynamics as a copula parameter through by a stochastic differential equation in continuous time, and (2) to evaluate the effectiveness of the novel copula model using real dataset from different portfolios in a market.

The key contributions of this paper are as follows.

Firstly, we propose a novel stochastic copula model for continuous cases based on the Ornstein-Uhlenbeck process. Secondly, we adapt Monte Carlo variance reduction techniques through the use of importance sampling and polynomial approximations for continuous case. Thirdly, we demonstrate how the proposed model can be effectively applied to assess risks on equal-weighted and CVaR-weighted portfolios.

The outcomes of this study can contribute to creation of more effective investment portfolio management strategies that account for complex interdependencies and temporal variations among assets.

This paper is structured as follows. Section 2, briefly describes methods for estimating marginal distributions. Section 3 reviews various types of copula models, including Archimedean copulas, sampling techniques for them, and stochastic autoregressive copulas (SCAR). The continuous SCAR model is developed in Section 4, where existence and uniqueness of solutions are discussed. Section 5 focuses on goodness-of-fit tests for proposed copula models. Risk metrics VaR and CVaR are introduced in Section 6, followed by testing their accuracy. Numerical experiments of VaR and CVaR calculation with real dataset from the cryptocurrency market are presented in Section 7. Finally, Section 8, concludes the paper and discusses the future work.

2 Marginal distribution estimation

Let S_t is asset price at time t. It is convenient to transform initial prices to logarithmic returns (log-return) r_t:

$$r_t = \log \frac{S_t}{S_{t-1}},$$

It is useful due to $r_t \in [-\infty, \infty]$, while $S_t \in [0, \infty]$. We would perform a further estimations for r_t variables.

It is known from numerous studies that the normal distribution, which is convenient for analysis, is not always suitable for financial assets risk assessment tasks due to insufficiently accurate prediction of the behavior of tails of distributions. Here we often talk about the presence of heavy tails in real distributions. In this paper, the following distributions are considered as marginal distributions of log-returns:

Normal

$$f(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{1}$$

where $x \in (-\infty,\infty)$, $\mu \in (-\infty,\infty)$, $\sigma > 0$.

Hyperbolic [14]

$$f(x,\lambda,\alpha,\beta,\delta,\mu) = \frac{(\gamma/\delta)^\lambda}{\sqrt{2\pi}K_\lambda(\delta\gamma)} e^{\beta(x-\mu)} \cdot \frac{K_{\lambda-1/2}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\left(\sqrt{\delta^2+(x-\mu)^2}/\alpha\right)^{(1/2-\lambda)}}, \tag{2}$$

where $x \in (-\infty,\infty)$, $\gamma = \sqrt{\alpha^2-\beta^2}$; $\lambda,\mu \in (-\infty,\infty)$; $\delta \geq 0, |\beta| < \alpha$, if $\lambda \geq 0$; $\delta > 0$ and $|\beta| \leq \alpha$, if $\lambda < 0$; K_p – modified Bessel function of second type of order p.

Stable [15]

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(t) e^{-ixt} dt, \tag{3}$$

where characteristic function ϕ has the form:

$$\phi(t,\alpha,\beta,c,\mu) = e^{it\mu - |ct|^\alpha (1 - i\beta sign(t)\Phi(\alpha,t))}.$$

The following parameterizations of the function $\Phi(a,t)$ are known from the literature.

Parameterization S_1

$$\Phi(\alpha,t) = \begin{cases} -\tan\frac{\pi\alpha}{2}, & \alpha \neq 1 \\ -\frac{2}{\pi}\log|t|, & \alpha = 1, \end{cases}$$

and parameterization S_0

$$\Phi(\alpha,t) = \begin{cases} -\tan\frac{\pi\alpha}{2}\left(|ct|^{1-\alpha}-1\right), & \alpha \neq 1 \\ -\frac{2}{\pi}\log|ct|, & \alpha = 1. \end{cases}$$

Meixner [16]

$$f(x,\alpha,\beta,\mu,\delta) = \frac{\left(2\cos\beta/2\right)^{2\delta}}{2\alpha\pi\Gamma(2\delta)} e^{\frac{\beta}{\alpha}(x-\mu)} \left|\Gamma\left(\delta + \frac{i(x-\mu)}{\alpha}\right)\right|^2, \tag{4}$$

where $\alpha > 0, -\pi < \beta < \pi, \delta > 0, \mu \in (-\infty,\infty)$.

Distribution parameters could be found using maximum likelihood estimation as follows. Let $f(x, \theta)$ – probability distribution function of random variable with unknown parameter set θ and $x_1, ..., x_n$ – is sample of values. Estimation for θ is:

$$\theta = \arg\max_{\theta} \log L(x_1, ..., x_n; \theta), \tag{5}$$

where logarithm of likelihood function $\log L(x_1, ..., x_n; \theta)$ has the form:

$$\log L(x_1, ..., x_n; \theta) = \sum_{i=1}^{n} \log f(x_i; \theta).$$

Among the statistical methods that allow us to test the empirical data for compliance with a given distribution, we can emphasize the following criteria: Kolmogorov-Smirnov, Cramer-von-Mises and Anderson-Darling criteria.

3 Copula models review

Consider a set of univariate random variables $X_1, ..., X_d$. In the case of financial datasets this variables could have a complex dependencies. To take them into account one could construct a multivariate distribution. One way to do this based on Sklar's theorem and copula approach [17].

Let $\boldsymbol{X} = (X_1, ..., X_d)$ – d-dimensional random variable. Distribution function of $\boldsymbol{X}$ could be represented as

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_d(x_d)), \quad (x_1, \ldots, x_d) \in \mathbb{R}^d,$$

where function C is d-dimensional function of one or more parameters called *copula*, which strict definition and properties could be found at [18]. It is useful to introduce vector $\boldsymbol{U} = (U_1, ..., U_d) = (F_1(x_1), ..., F_d(x_d))$ whose components are uniform at $[0, 1]$. Hence we can write the expression for copula density

$$c(u_1, ..., u_d) = \frac{\partial^d C(u_1, ..., u_d)}{\partial u_1 \cdots \partial u_d}, \quad u_1, ..., u_d \in [0, 1]$$

and density of joint distribution

$$f(x_1, \ldots, x_d) = c(F_1(x_1), \ldots, F_d(x_d)) \prod_{i=1}^{d} f_i(x_i), (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{6}$$

General problem of finding parameters of joint distribution implies solving the problem (5) for whole function (6), but it is convenient to solve two separate problems: for copula and for marginal distributions. For marginals we can use estimation (5) and for copula the following procedure. Let us define $U_i = F_i(X_i)$ – uniformly disturbed marginals and corresponding variables $u_i = F_i(x_i)$, $i = 1, ..., d$. Then copula density

will take the form $c(u_1, ..., u_d; \theta)$. Calculation of variables u_i implies the knowledge of the distribution function, but we can estimate the function $F_i(x_i)$ by non-parametric way using it's empirical version, which is equivalent to introducing a new variables called *pseudo observations*:

$$u_{ij} = \frac{rank(x_{ij})}{n+1}, \quad i = 1, ..., d, \quad j = 1, ..., n,$$

where $rank(x_{ij})$ – element number of x_{ij} in ranked series $x_{i1}, ..., x_{in}$. Denote as $U_i = \{u_{ij}\}_{j=1}^n$ – a set of observations, $i = 1, ..., d$. Then the maximum likelihood estimation of copula parameter θ_c is (such an approach could be found in literature as Omnibus or OM estimation)

$$\theta_c = \arg\max_{\theta} \log L(U_1, ..., U_d; \theta) = \sum_{j=1}^{n} \log c(u_{1j}, ..., u_{dj}; \theta).$$

3.1 Archimedean copulas

In order to construct a multivariate distribution based on copulas one should use a some class. In the study we used Archimedean copulas [7, 10, 18, 19].

Consider a strictly decreasing convex function $\phi(t; \theta)$ such that $\phi(1; \theta) = 0$, θ is fixed parameter. Function ϕ is called *generator* of copula. Formally, the Archimedean copula is a function satisfying the following parametrization

$$C(u_1, ..., u_d) = \phi^{-1}\left(\phi(u_1) + ... + \phi(u_d)\right).$$

The following theorem [20] establish necessary and sufficient conditions on generator ϕ to generate a multivariate copula.

Theorem 1. *Let ϕ be a continuous strictly decreasing function from $[0, 1]$ to $[0, \infty)$ such that $\phi(0) = \infty$ and $\phi(1) = 0$, and let ϕ^{-1} denote the inverse of ϕ. Then $C(u_1, ..., u_d)$ is an d-copula for all $n \geq 2$ if and only if ϕ^{-1} is completely monotone on $[0, \infty)$, i. e. $(-1)^k \frac{d^k \phi^{-1}(t)}{dt^k} \geq 0$ for any $k \in \mathbb{N}$.*

The generators of some Archimedean copulas in explicit form are given in Table 1 (see [10, 18] for details). For given copulas the conditions of Theorem 1 holds.

Table 1 Archimedean copulas

Copula	$\phi(t; \theta)$	$\phi^{-1}(t; \theta)$	$\theta \in$
Gumbel	$(-\log t)^\theta$	$\exp\left(-t^{1/\theta}\right)$	$[1, \infty)$
Frank	$-\log\left(\frac{\exp(-\theta t)-1}{\exp(-\theta)-1}\right)$	$-\frac{1}{\theta}\log\left(1+\exp(-t)(\exp(-\theta)-1)\right)$	$(0, \infty)$
Joe	$-\log\left(1-(1-t)^\theta\right)$	$1-(1-\exp(-t))^{(1/\theta)}$	$[1, \infty)$
Clayton	$\frac{1}{\theta}\left(t^{-\theta}-1\right)$	$(1+t\theta)^{-1/\theta}$	$(0, \infty)$

In the bivariate case it is useful to introduce a rotated (or survival) copulas. The copulas rotated on 90, 180, 270 degrees respectively have the form:

$$\begin{aligned} C_{90}(u_1, u_2) &= u_2 - C(1 - u_1, u_2), \\ C_{180}(u_1, u_2) &= u_1 + u_2 + C(1 - u_1, 1 - u_2) - 1, \\ C_{270}(u_1, u_2) &= u_1 - C(u_1, 1 - u_2). \end{aligned}$$

The form of the density does not change accurate to variable substitution. For example, $c_{180}(u_1, u_2) = c(1 - u_1, 1 - u_2)$.

3.2 Sampling Archimedean copulas

An efficient generation algorithm from Archimedean copula is closely related to the Laplace-Stieltjes transform of the inverse generator function (see Hofert [19] for the discussion and numerical studies). Hereinafter we will denote $\psi(t) = \phi^{-1}(t)$.

Definition 1. *Laplace-Stieltjes ($\mathcal{LS}$) transform of function $F(x)$ is given by*

$$\psi(t) = \int_0^\infty e^{-tx} dF(x), \quad t \in [0, \infty).$$

The following theorem (Bernstein, see [21]) establish a necessary and sufficient conditions when $\psi(t)$ is $\mathcal{LS}$-transform of some distribution function $F(x)$.

Theorem 2. *A function ψ on $[0, \infty)$ is the Laplace-Stieltjes transform of a distribution function if and only if ψ is completely monotone, i. e. $(-1)^k \frac{d^k \psi(t)}{dt^k} \geq 0$ for any $k \in \mathbb{N}$, and $\psi(0) = 1$.*

For copulas with generators satisfying Theorem 2 there is a simple and powerful sampling algorithm (Marshall, Olkin [22]).

Algorithm 1 Marshall and Olkin sampling algorithm

1: Sample $V \sim F = \mathcal{LS}^{-1}(\psi)$.
2: Sample i.i.d $X_i \sim U[0, 1], \quad i = 1, ..., d$.
3: Return $(U_1, ..., U_d)$, where $U_i = \psi\left(-\log\left(X_i\right)/V\right), \quad i = 1, ..., d$.

The crucial difficulty of Algorithm 1 is to find an inverse $\mathcal{LS}$-transformation $F(x) = \mathcal{LS}^{-1}(\psi)$, however, for a certain generators this transformation is known [23]. The corresponding transformations are listed in Table 2, where $\Gamma(\alpha, \beta)$ denotes a gamma distribution with shape parameter α and scale parameter β; $S(\alpha, \beta, c, \mu)$ denotes a stable distribution with exponent parameter α, skewness parameter β and location and scale parameters c and μ; y_k denotes a probability mass function of discrete distribution.

Table 2 The inverse Laplace-Stieltjes transform for Archimedean copulas

Copula	$\phi^{-1}(t;\theta)$	F	$\theta \in$
Gumbel	$\exp\left(-t^{1/\theta}\right)$	$S\left(1/\theta, 1, 0, \left(\cos\left(\frac{\pi}{2\theta}\right)\right)^{\theta}\right)$	$[1,\infty)$
Frank	$-\frac{1}{\theta}\log\left(1+\exp\left(-t\right)(\exp\left(-\theta\right)-1)\right)$	$y_k = \frac{\left(1-e^{-\theta}\right)^k}{k\theta}, k \in \mathbb{N}$	$(0,\infty)$
Joe	$1-\left(1-\exp\left(-t\right)\right)^{1/\theta}$	$y_k = (-1)^{k+1}\binom{1/\theta}{k}, k \in \mathbb{N}$	$[1,\infty)$
Clayton	$(1+t\theta)^{-1/\theta}$	$\Gamma(1/\theta, 1)$	$(0,\infty)$

3.3 Stochastic autoregressive copulas (SCAR)

Consider single-parameter model, for example Archimedean copula or bivariate elliptical copula. Without limiting the generality consider bivariate case and suppose that copula parameter is not fixed and follows the law

$$\begin{aligned}(u_{1t}, u_{2t}) &\sim c(u_{1t}, u_{2t}; \theta_t),\\ \theta_t &= \Psi(\lambda_t),\\ \lambda_t &= \alpha_1 + \alpha_2\lambda_{t-1} + \alpha_3\eta_t.\end{aligned} \tag{7}$$

where $\Psi : \mathbb{R} \to \Theta$ – a suitable transformation that guarantees that the copula parameter remains in its domain of definition. Note that for Archimedean copulas this model admits an extension to n dimensions, in which case expression $c(u_{1t}, u_{2t}; \theta_t)$ is simply replaced by expression $c(u_{1t}, u_{2t}, ..., u_{nt}; \theta_t)$, since the copula is uniparametric. It is assumed that the underlying process $\{\lambda_t\}_{t=1}^T$ is latent (i.e. unobservable) and describes by Gaussian autoregressive process of the first order $AR(1)$, where $\eta_t \overset{iid}{\sim} \mathcal{N}(0,1)$, $|\alpha_2| < 1$ to ensure stationarity of λ_t and $\alpha_3 > 0$. This approach was suggested in Liesenfeld and Richard [24], discussed among other time-varying models in Manner and Reznikova [3] and was developed in Almeida, Czado, Manner [7] for vine copulas.

Process $\{\lambda_t\}_{t=1}^T$ is unobservable, so estimation for $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ cannot be obtained explicitly. The paper [24] proposed expression for likelihood function in the general form

$$L(\alpha; U_1, U_2) = \int f(U_1, U_2, \Lambda; \alpha) d\Lambda,$$

here $U_1 = \{u_{1t}\}_{t=1}^T$, $U_2 = \{u_{2t}\}_{t=1}^T$, $t = 1, 2, \ldots, T$ and $f(\cdot)$ – joint density of observable variables (U_1, U_2) and latent process λ_t, and integration is performed over all possible states of the process λ_t. Using expression for copula density likelihood function can be represented in the form

$$L(\alpha; U_1, U_2) = \int \prod_{t=1}^{T} c\left(u_{1t}, u_{2t}, \lambda_t \,|\, U_{1,t-1}, U_{2,t-1}, \Lambda_{t-1}; \alpha\right) p\left(\lambda_t \,|\, \Lambda_{t-1}; \alpha\right) d\Lambda, \tag{8}$$

where $U_{1,t} = \{u_{1i}\}_{i=1}^t$, $U_{2,t} = \{u_{2i}\}_{i=1}^t$ – a set of observations (pseudo-observations) known up to the moment of t, $p(\lambda_t|\Lambda_{t-1}; \alpha)$ – probability density of transition to the state λ_t at the moment t from the set of previous states $\Lambda_{t-1} = \{\lambda_i\}_{i=1}^{t-1}$. It is not possible to calculate the integral (8) analytically, however, it is possible to obtain its

estimation by applying the Monte Carlo method:

$$\hat{L}_N(\alpha; U_1, U_2) = \frac{1}{N} \sum_{i=1}^{N} \left(\prod_{t=1}^{T} c\left(u_{1t}, u_{2t}, \lambda_t^{(i)} \,|\, U_{t-1}, \Lambda_{t-1}; \alpha\right) \right), \tag{9}$$

where N – number of random paths, $\lambda_t^{(i)}$ – value of the random process on the i trajectory with density p at time t (note that common random numbers should be used for generating random trajectories. Otherwise estimation would not be smooth). Hereafter, we will refer to the (9) estimation as *natural* or *p-sampler*. It should be noted that this estimation is very inefficient for large sample sizes.

The estimation (9) can be significantly optimized in terms of reducing Monte Carlo variance by introducing a new density of distribution $m\left(\lambda_t|\Lambda_{t-1}, a_t\right)$, where $a_t = (a_{1t}, a_{2t})$ – auxiliary parameters. Then the estimation will take the form

$$\hat{L}_N(\alpha; U_1, U_2) = \frac{1}{N} \sum_{i=1}^{N} \left(\prod_{t=1}^{T} \frac{f\left(u_{1t}, u_{2t}, \lambda_t^{(i)} \,|\, U_{t-1}, \Lambda_{t-1}(a_{t-1}); \alpha\right)}{m\left(\lambda_t^{(i)} \,|\, \Lambda_{t-1}^{(i)}(a_{t-1}); a_t\right)} \right), \tag{10}$$

here $\lambda_t^{(i)}(a_t)$ – value of the random process on the i trajectory at time t, but with density m. This technique is called *efficient importance sampling* (EIS) and the estimate will be called as m-sampler.

The function m can be chosen in the following form

$$m\left(\lambda_t|\Lambda_{t-1}; a_t\right) = \frac{k(\Lambda_t; a_t)}{\int k(\Lambda_t; a_t) d\lambda_t},$$

where $k(\Lambda_t; a_t)$ is *density kernel*. It can be shown that the problem of estimating the auxiliary coefficients reduces to the problem

$$\hat{a}_t(\alpha) = \arg\min_{a_t} \sum_{i=1}^{N} \Bigg[\log f\left(u_{1t}, u_{2t}, \lambda_t^{(i)} \,|\, U_{t-1}, \Lambda_{t-1}^{(i)}; \alpha\right) \times \\ \times \int k\left(\Lambda_{t-1}^{(i)}; \hat{a}_{t-1}\right) d\lambda_t - c_t - \log k\left(\Lambda_t^{(i)}; a_t\right) \Bigg]^2 \tag{11}$$

for $t = T, T-1, \ldots, 1$, under the condition $\int k\left(\Lambda_T^{(i)}(\alpha); \hat{a}_T(\alpha)\right) d\lambda_T \equiv 1$. The search for constants c_t is done in conjunction with a_t. If the function $k(\lambda_t; a_t)$ is chosen in the following form

$$k(\lambda_t, a_t) = p\left(\lambda_t|\Lambda_{t-1}\right) \exp\left(a_{1t}\lambda_t + a_{2t}\lambda_t^2\right), \tag{12}$$

then the problem (11) can be reduced to linear and the method of least squares can be applied, reducing the number of trajectories to the value $100 < N < 200$. In the case

of choosing $k(\lambda_t; a_t)$ as (12), the normalization multiplier can be computed explicitly

$$\xi_t = \int\limits_{-\infty}^{+\infty} k(\Lambda_t; a_t) d\lambda_t = \sqrt{\frac{v}{\alpha_3^2}} \exp\left[\left(\frac{\mu^2}{2v} - \frac{1}{2\alpha_3^2}\right)(\alpha_1 + \alpha_2\lambda_{t-1})^2\right],$$

where

$$v = \frac{\alpha_3^2}{1 - 2\alpha_3^2 a_2}, \quad \mu = v\left(\frac{\alpha_1 + \alpha_2\lambda_{t-1}}{\alpha_3^2} + a_1\right).$$

It is also can be shown that regression with density m has the form:

$$\lambda_t = p_1 + p_2\lambda_{t-1} + p_3\eta_t,$$

where

$$p_1 = v\left(\frac{\alpha_1}{\alpha_3^2} + a_1\right), \quad p_2 = v\frac{\alpha_2}{\alpha_3^2}, \quad p_3 = \sqrt{v}.$$

4 Continuous SCAR model

Suppose now that the process λ_t is a process in continuous time on the interval $t \in [0, 1]$. The continuous version of the discrete process (7) is the Ornstein-Uhlenbeck process

$$dx_t = (\alpha_1 + \alpha_2 x_t)\, dt + \alpha_3 dW_t, \quad t \geq 0, \tag{13}$$

where W_t – Wiener process. The same in canonical variables:

$$dx_t = \theta\,(\mu - x_t)\, dt + \nu dW_t, \quad t \geq 0, \tag{14}$$

where the parameters of the processes (13) and (14) are related as follows:

$$\mu = -\frac{\alpha_1}{\alpha_2}, \quad \theta = -\alpha_2, \quad \nu = \alpha_3.$$

Let us formalize the studied quantities in order to write down the model for the continuous case. Let the probability space $(\Omega, \mathcal{F}, P)$ be given. Consider on it some random process $\{x_t\}$, $t \in [0, 1]$. Let us divide the interval $[0, 1]$ by the points $0 \leq t_1 \leq t_2 \leq \ldots \leq t_k \leq 1$. In this case, the σ-algebra $\mathcal{F}$ will contain the σ-algebra $\mathcal{B}$ generated by a set of the form

$$\{\omega \in \Omega, \omega_1(t_1) \in F_1, \ldots, \omega_k(t_k) \in F_k\}, \quad F_i \subset \mathbb{R} \text{ – Borel set.}$$

Let us introduce the following probability measure on this set:

$$\nu_{t_1,\ldots,t_k}(F_1 \times \ldots \times F_k) = \int\limits_{F_1 \times \ldots \times F_k} \prod_{i=1}^{k} p(x_i, t_i | x_{i-1}, t_{i-1})\, dx_1 dx_2 \ldots dx_k,$$

where $p(x,t|x',t')$ – probability density function of transition from the state (x',t') to (x,t). To match the notations introduced for the discrete case we denote as $\alpha = (\mu,\theta,\nu)$ set of constant parameters of stochastic equation (14) in canonical form and by analogy with the expression (8), define the likelihood function as follows:

$$L(\alpha;U_1,U_2) = \mathbb{E}\left(\prod_{t=1}^{T} c(u_{1t},u_{2t},x_t|U_{1,t-1},U_{2,t-1},x_{t-1};\alpha)\right) =$$

$$= \int\limits_{F_1\times\ldots\times F_T} \prod_{t=1}^{T} c(u_{1t},u_{2t},x_t|U_{1,t-1},U_{2,t-1},x_{t-1};\alpha) d\nu_{t_1,\ldots,t_T}\left(F_1\times\ldots\times F_T\right), \quad (15)$$

where the partitioning of the interval $[0,1]$ is made by T points – according to a given set of observations. Now define the stochastic copula model in the continuous case

$$\begin{aligned}(u_{1t},u_{2t}) &\sim c(u_{1t},u_{2t};\theta_t),\\ \theta_t &= \Psi(x_t),\\ dx_t &= \theta\left(\mu - x_t\right)dt + \nu dW_t, \quad t\in[0,1],\end{aligned}$$

where $\Psi(x_t)$ is a real function that translates the values of the process x_t into the interval of parameter values allowed for a given copula. In further calculations we will use $\Psi(x) = x\tanh(x)$ for Clayton and Frank copulas, and $\Psi(x) = 1 + x\tanh(x)$ for Gumbel and Joe.

In the direct calculations we will use Monte Carlo estimates of (15). Let N be the number of trajectories of the random process $\{x_t\}$, then Monte Carlo estimation (p-sampler) of the integral (15) has the following form

$$\hat{L}(\alpha,U_1,U_2) = \frac{1}{N}\sum_{i=1}^{N}\left(\prod_{t=1}^{T} c(u_{1t},u_{2t},x_t^{(i)}|U_{1,t-1},U_{2,t-1},x_{t-1}^{(i)};\alpha)\right).$$

And again as in discrete case common random numbers should be used for generating random trajectories to ensure a smooth estimation. In order to construct an improved estimation (m-sampler) we need to analyze a transition density function $p(x,t|x',t')$. For a continuous stochastic Ito process, the function $p(x,t|x',t')$ can be found using the Fokker-Planck equation. In general, for a process of the form

$$dx_t = A(x,t)dt + B(x,t)dW_t, \quad (16)$$

Fokker-Planck equation has the form

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left(A(x,t)p(x,t)\right) + \frac{\partial^2}{\partial x^2}\left(\frac{B^2(x,t)}{2}p(x,t)\right). \quad (17)$$

The solution of (17) for process (14) is known

$$p(x,t|x',t') = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\overline{x})^2}{2\sigma^2}\right), \tag{18}$$

where

$$\overline{x} = \mu + (x' - \mu)\exp\left(-\theta(t-t')\right),$$
$$\sigma^2 = \frac{\nu^2}{2\theta}\left(1 - \exp\left(-2\theta\left(t-t'\right)\right)\right).$$

Let us introduce a new density

$$m(x,t|x',t') = p(x,t|x',t')g_1(x,t,\boldsymbol{a}|x',t')g_2(t,\boldsymbol{a}|x',t'), \tag{19}$$

where $g_1(x,t,\boldsymbol{a}|x',t')$ is the desired auxiliary function, $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ is the set of parameters which, in general, can depend on time t, $g_2(t,\boldsymbol{a}|x',t')$ is the normalization factor calculated from the condition

$$g_2(t,\boldsymbol{a}|x',t') = \left[\int_{-\infty}^{+\infty} g_1(x,t,\boldsymbol{a}|x',t')p(x,t|x',t')dx\right]^{-1}.$$

By analogy with the discrete representation (12), consider the function

$$g_1(x,t,\boldsymbol{a}) = \exp\left(a_1 x + a_2 x^2\right), \quad \boldsymbol{a} = (a_1, a_2) \in \mathbb{R}^2. \tag{20}$$

In this case, it is possible to calculate the normalizing factor

$$g_2(t,\boldsymbol{a}|x',t') = \left[\frac{1}{\sqrt{1-2a_2\sigma^2}} \exp\left(\frac{a_1^2\sigma^2 + 2\overline{x}(a_1 + a_2\overline{x})}{2 - 4a_2\sigma^2}\right)\right]^{-1},$$

under condition

$$1 - 2a_2\sigma^2 > 0. \tag{21}$$

It can be observed that density (19) is a normal distribution with the following parameters:

$$m(x,t|x',t') = \frac{1}{\sqrt{2\pi\tilde{\sigma}^2}} \exp\left(-\frac{(x-\tilde{x})^2}{2\tilde{\sigma}^2}\right), \tag{22}$$

where

$$\tilde{x} = \frac{\overline{x} + a_1\sigma^2}{1 - 2a_2\sigma^2}, \quad \tilde{\sigma}^2 = \frac{\sigma^2}{1 - 2a_2\sigma^2}.$$

The problem is to find a random process $\{x_t\}$ in the form (16) that satisfies the density (22). Therefore we need to determine the unknown functions $A(x,t)$ and $B(x,t)$. Since the equation is only one and there are two unknown functions, we need to determine or choose the function $B(x,t)$. For the Ornstein-Uhlenbeck process it could be done on the basis of normalizing conditions.

The equation (17) can be solved with respect to $A(x,t)$ in a rather general form by ordinary integration over the variable x. We have the following expression

$$A(x,t) = \frac{1}{p(x,t)}\left(C_1 + \frac{\partial}{\partial x}\left(\frac{1}{2}B^2(x,t)p(x,t)\right) - \frac{\partial}{\partial t}\int p(x,t)dx\right),$$

where C_1 – constant of integration. Using an explicit expression for density (22) and assuming $B(x,t) = B(t)$ we get

$$A(x,t) = \frac{\partial\tilde{x}}{\partial t} - \frac{(x-\tilde{x})}{2\tilde{\sigma}^2}\left(B^2(t) - \frac{\partial\tilde{\sigma}^2}{\partial t}\right) + C_1\exp\left(\frac{x(x-2\tilde{x})}{2\tilde{\sigma}^2}\right).$$

From the condition $A(x,t)|_{a_1,a_2=0} = \theta(\mu - x)$ we obtain that the integration constant $C_1 = 0$. Moreover, this condition obliges us to choose the function $B(t)$ such that $B(t)|_{a_1,a_2=0} = \nu$. In sum, we obtain that

$$A(x,t) = \frac{\partial\tilde{x}}{\partial t} - \frac{(x-\tilde{x})}{2\tilde{\sigma}^2}\left(B^2(t) - \frac{\partial\tilde{\sigma}^2}{\partial t}\right), \tag{23}$$

where

$$\frac{\partial\tilde{x}}{\partial t} = \frac{1}{1-2a_2\sigma^2}\left(\frac{\partial\overline{x}}{\partial t} + a_1\frac{\partial\sigma^2}{\partial t} + \frac{\partial a_1}{\partial t}\sigma^2\right) + \frac{2\tilde{x}}{1-2a_2\sigma^2}\left(\frac{\partial a_2}{\partial t}\sigma^2 + a_2\frac{\partial\sigma^2}{\partial t}\right), \tag{24}$$

$$\frac{\partial\tilde{\sigma}^2}{\partial t} = \frac{1}{(1-2a_2\sigma^2)^2}\left(\frac{\partial\sigma^2}{\partial t} + 2\sigma^4\frac{\partial a_2}{\partial t}\right), \tag{25}$$

$$\frac{\partial\overline{x}}{\partial t} = -\theta(x'-\mu)\exp\left(-\theta(t-t')\right), \tag{26}$$

$$\frac{\partial\sigma^2}{\partial t} = \nu^2\exp\left(-2\theta\left(t-t'\right)\right). \tag{27}$$

Despite the bulkiness of expressions (23)-(27) the solution of (16) could be obtained in a simple closed form

$$x_t = \tilde{x} + \frac{\nu e^{-\theta t}}{\sqrt{1-2a_2\sigma^2}}\int_0^t e^{\theta p}dW_p, \quad t \geq 0, \tag{28}$$

if we choose $B(x,t)$ as follows

$$B(t) = \frac{\nu}{\sqrt{1-2a_2\sigma^2}}, \tag{29}$$

where auxiliary parameters considered as time dependent: $a_1 = a_1(t)$, $a_2 = a_2(t)$. The proof of the solution (28) is given in Appendix A.

4.1 Existence and uniqueness of the solution

Let us now discuss the existence and uniqueness of the solution of the equation (16) for the obtained functions $A(x,t)$ and $B(x,t)$. There is the following theorem about existence and uniqueness of the solution [25].

Theorem 3. *Let $T > 0$, $A(x,t)$, $B(x,t)$ – measurable functions, satisfying the conditions:*

1. $|A(x,t)| + |B(x,t)| \leq C(1 + |x|), \quad x \in \mathbb{R}, \quad t \in [0,T],$
2. $|A(x,t) - A(y,t)| + |B(x,t) - B(y,t)| \leq D|x - y|, \quad x, y \in \mathbb{R}, \quad t \in [0,T],$

where C and D – certain constants. Let Z – random variable independent of the σ-algebra $\mathcal{F}_\infty$, generated by functions W_s, $s \geq 0$, so that

3. $\mathbb{E}(|Z|^2) < \infty$.

Then stochastic differential equation

$$dx_t = A(x,t)dt + B(x,t)dW_t, \quad 0 \leq t \leq T, \quad x_0 = Z,$$

has a unique t-continuous solution $x_t(\omega)$, such that $x_t(\omega)$ is consistent with the flow $\mathcal{F}_t^Z$ generated by Z and W_s, $s \leq t$ and

$$\mathbb{E}\left(\int_0^T |x_t|^2 \, dt\right) < \infty.$$

The proof of the theorem is given in the Oksendal's monograph [25].

Let us investigate the question of existence and uniqueness of the solution with respect to the functions (23) and (29). Based on the explicit form of the functions we can formally write down

$$A(x,t) = \phi_1(t) + \phi_2(t)x,$$

where $\phi_1(t)$ and $\phi_2(t)$ are continuous functions on $[0,1]$. Thus, the function $A(x,t)$ – is linear on x, and $B(x,t) = B(t)$ do not depend on x. Therefore, the first condition of Theorem 3 will be satisfied if we choose

$$C = \max\left(\sup_{t\in[0,1]} |\phi_1(t)|, \sup_{t\in[0,1]} |\phi_2(t)|, \sup_{t\in[0,1]} |B(t)|\right).$$

The second condition of the Theorem 3 will be satisfied if we put

$$D = \sup_{t\in[0,1]} |\phi_2(t)|.$$

Now, if we put $x_0 = \mu$, the third condition of Theorem 3 will be satisfied automatically. Hence, according to Theorem 3, the solution of the equation (16) for the chosen functions $A(x,t)$ and $B(x,t)$ exists and is unique if the coefficients a_1 and a_2 are chosen such that the normalization of the distribution $m(x,t)$ is not violated, i.e. the condition (21) must be satisfied for all $t \in [0,1]$.

Now we can conduct reasoning in the opposite direction. Consider the equation (16). As shown above, for the functions (23) and (29) the solution exists and is unique. Consequently, the probability density of transition from (x',t') to (x,t) also exists, which can be found from the Fokker-Plank equation (17). It is easy to see that for the given functions $A(x,t)$ and $B(x,t)$ the solution is the density (22).

4.2 Finding an explicit form of auxiliary parameters

It remains to consider the question of finding the auxiliary parameters. The main scheme of optimization problem follows the approach suggested in [24], but some numerical aspects differs for the continuous case. Hereinafter we denote as variables with hats $\hat{a}_{1,t_i}$, $\hat{a}_{2,t_i}$ a variables that are obtained from the least squares problem solving, where $t_i \in [0,1]$. In computer terms $\hat{a}_1 = \{\hat{a}_{1,t_1}, ..., \hat{a}_{1,t_T}\}$ and $\hat{a}_2 = \{\hat{a}_{2,t_1}, ..., \hat{a}_{2,t_T}\}$ represents an arrays of length T. For further purposes we would also need a smooth estimations of $\hat{a}_1$ and $\hat{a}_2$. We would denote them as $a_1(t)$, $a_2(t)$, where $t \in [0,1]$. Now let us write down an expression for the likelihood function $m(x,t|x',t')$:

$$L(\alpha; U_1, U_2) = \int\limits_{F_1\times...\times F_T} \prod_{t=1}^{T} c(u_{1t}, u_{2t}, x_t | U_{1,t-1}, U_{2,t-1}, x_{t-1}; \alpha) \cdot$$

$$\cdot\, d\nu_{t_1,...,t_T}(F_1 \times ... \times F_T) = \int\limits_{F_1\times...\times F_T} \prod_{t=1}^{T} c(u_{1t}, u_{2t}, x_t | U_{1,t-1}, U_{2,t-1}, x_{t-1}; \alpha) \cdot$$

$$\cdot \prod_{i=1}^{T} p(x_i, t_i | x_{i-1}, t_{i-1})\, dx_1 dx_2 \dots dx_T =$$

$$= \int\limits_{F_1\times...\times F_T} \prod_{t=1}^{T} c(u_{1t}, u_{2t}, x_t | U_{1,t-1}, U_{2,t-1}, x_{t-1}; \alpha) \prod_{i=1}^{T} \frac{p(x_i, t_i | x_{i-1}, t_{i-1})}{m(x_i, t_i | x_{i-1}, t_{i-1})} \cdot$$

$$\cdot \prod_{i=1}^{T} m(x_i, t_i | x_{i-1}, t_{i-1})\, dx_1 dx_2 \dots dx_T.$$

Let us generate N trajectories of a random process x_t with density $m(x,t)$ in accordance with the equation (16) and the functions (23), (29). Then the EIS Monte Carlo estimation (m-sampler) of this integral is

$$\hat{L}(\alpha, U_1, U_2) = \frac{1}{N}\sum_{k=1}^{N}\left(\prod_{i=1}^{T}\frac{c(u_{1i}, u_{2i}, x_i^{(k)}|U_{1,i-1}, U_{2,i-1}, x_{i-1}^{(k)};\alpha)}{m(x_i^{(k)}, t_i|x_{i-1}^{(k)}, t_{i-1})}\cdot\right.$$
$$\left.\cdot\, p(x_i^{(k)}, t_i|x_{i-1}^{(k)}, t_{i-1})\right), \quad (30)$$

The crucial condition for parameter selection, as in the algorithm for the discrete case, will be the most exact match between the numerator and denominator at each time $i = 1, ..., T$ for each of the trajectories $1 \le k \le N$

$$\frac{c(u_{1i}, u_{2i}, x_i^{(k)}|U_{1,i-1}, U_{2,i-1}, x_{i-1}^{(k)};\alpha)p(x^{(k)}, t_i|x_{i-1}^{(k)}, t_{i-1})}{m(x^{(k)}, t_i|x_{i-1}^{(k)}, t_{i-1})} \to 1.$$

Note that this expression is a function only of the parameters a_1, a_2. Passing from the functions themselves to their logarithms and considering all available trajectories we get the least squares problem

$$(\hat{a}_{1,t_i}, \hat{a}_{2,t_i}) = \underset{(a_1,a_2)}{\arg\min}\sum_{k=1}^{N}\left(\log\left[c(u_{1i}, u_{2i}, x_t^{(k)}|U_{1,i-1}, U_{2,i-1}, x_{i-1}^{(k)};\alpha\right]-\right.$$
$$\left.-\log\left[g_1(t_i, a_1, a_2|x^{(k)}, t_{i-1})\right] - \log\left[g_2(t_i, a_1, a_2|x^{(k)}, t_{i-1})\right]\right)^2, \quad (31)$$

where representation of the function $m(x,t)$ in the form (19) was used. Note that according to (20), the function $\log g_1(\cdot)$ is linear with respect to a_1, a_2 – this is the main advantage of representing the function $m(x,t)$ as (19). But $\log g_2(\cdot)$ is not linear, so solving the nonlinear problem (31) is not very computationally efficient. In order to overcome this difficulty, we modify the original problem to the following form:

$$(\hat{a}_{1,t_i}, \hat{a}_{2,t_i}) = \underset{(a_1,a_2)}{\arg\min}\sum_{k=1}^{N}\left(\log\left[c(u_{1i}, u_{2i}, x_i^{(k)}|U_{1,i-1}, U_{2,i-1}, x_{i-1}^{(k)};\alpha\right]-\right.$$
$$\left.-\log\left[g_1(t_i, a_1, a_2|x^{(k)}, t_{i-1})\right] - \log\left[g_2(t_i, \hat{a}_{1,t_{i-1}}, \hat{a}_{2,t_{i-1}}|x^{(k)}, t_{i-1})\right]\right)^2, \quad (32)$$

and start solving it by moving backwards in time from the moment $t = T$ to $t = 0$. As an initial condition we can set $\hat{a}_{1,T} = 0$, $\hat{a}_{2,T} = 0$ (but it is more convenient to shift the initial condition, which will be discussed below). Therefore, we substitute into the function g_2 the values $\hat{a}_{1,t_{i-1}}$, $\hat{a}_{2,t_{i-1}}$ already known from the previous step, which ensures linearity of the problem with respect to a_1, a_2. If we now execute the above algorithm several times (often it is enough, for example, $M = 5$ times), then by

successive approximations we will ensure convergence to the solution of the original problem.

At the end of this section we will formulate an algorithm for finding the model parameters in the continuous case and provide some remarks.

Algorithm 2 Auxiliary parameters in continuous case

Require: $M \geq 1$
Ensure: Estimation $\hat{L}(\alpha, U_1, U_2)$ (30)
1: **while** $j \leq M$ **do**
2: **if** $j = 1$ **then**
3: Generate N trajectories based on the stochastic equation (13)
4: Set $a_1(T) = 0$, $a_2(T) = 0$
5: **else if** $j > 1$ **then**
6: Generate N trajectories based on equation (28)
7: Set $\hat{a}_{i,T}^{(j)} = \frac{1}{T} \sum_{t=0}^{T} \hat{a}_{i,t}^{(j-1)}$, $i = 1, 2$
8: **end if**
9: Approximate $\hat{a}_{1,t}$, $\hat{a}_{2,t}$ by polynomials of degree j
10: $j = j + 1$
11: **end while**
12: Perform Monte Carlo estimation of the likelihood function by the formula (30)

First, in the Algorithm 2 in the line 4 we moving sequentially in the direction from $t = T$ to $t = 0$, we solve the optimization problem (32) applying threshold constraints on the parameter values (21).

Second, in the line 9 for each of the functions $a_1(t)$, $a_2(t)$ we additionally fix the rightmost point equal to the average of the found set $\hat{a}_{2,t}$. In case the approximation $a_2(t)$ exceeds the boundary value, we repeat sequentially the approximation with the regularization parameter α equal to 0.01, 0.05, 0.1, 0.5, 1.0, 5.0, 10.0. Finally, we approximate $a_1(t)$ with the same parameter α.

One computational aspect is important to note here. Despite the fact that in the expression (28) the quantities x' and t' are included explicitly (via $\tilde{\sigma}^2$ and $\tilde{x}$), here they have the sense of fixed parameters, namely

$$t' = 0, \quad x' = x_{t'=0},$$

whereas in the expression (19), when calculating the normalization coefficients, it is always

$$t - t' = dt, \quad x' = x_{t-dt}.$$

Moreover, in the expression (28)

$$t = \frac{i}{T},$$

where i is the number of current step in the discretization scheme, T is the total number of steps. In addition, we note that the proposed method allows us to construct processes with other functions than the given function $g_1(x, t, \boldsymbol{a})$. The only restriction is the solvability of the equation (17) with respect to $A(x, t)$ in elementary functions.

An example of a possible approximation and it is impact on random process trajectories shape is shown in Fig. 1 and Fig. 2. We see that the auxiliary coefficients have a significant effect on the shape of the trajectories with characteristic regions where the process dynamics is similar for all trajectories.

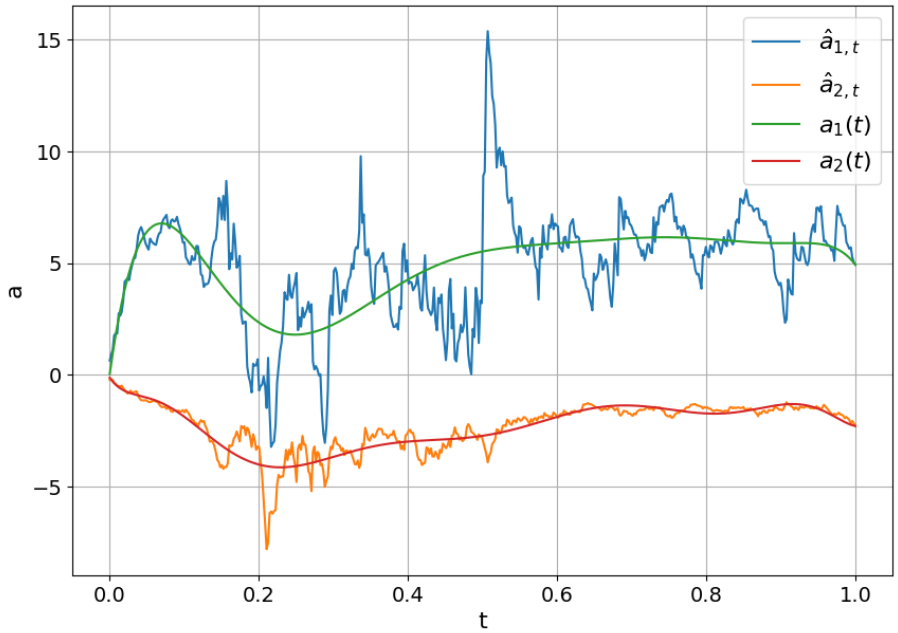

Fig. 1 Dependence of parameters $\hat{a}_{1,t}$ (blue line), $\hat{a}_{2,t}$ (orange line) and their approximation by polynomials of the tenth order $a_1(t)$ (green line), $a_2(t)$ (red line), $\alpha = 0$. Approximation parameters for $a_1(t)$: $R^2 = 0.36$, $MSE = 4.1$, for $a_2(t)$: $R^2 = 0.77$, $MSE = 0.3$. Rotated (180) Gumbel copula. Process parameters $\{x_t\}$: $\theta = 8.43$, $\mu = 1.64$, $\nu = 2.02$.

Remark 1. *The task* (32) *has a number of difficulties:*

1. *The initial problem is nonlinear. The method of successive approximations is used for linearization. The linear problem is ill-conditioned – a solution of least squares problem is calculated using a pseudo-inverse matrix.*
2. *The estimation is performed with a backward movement in time from T to 0. It is necessary to correctly set the initial condition. Numerical experiments show that the auxiliary parameters $a_1(t)$ and $a_2(t)$ vary in rather wide ranges and fixing the zero value at the rightmost point (as suggested for discrete case) leads to numerical instability, which has negative impact on the accuracy of the likelihood function. To*

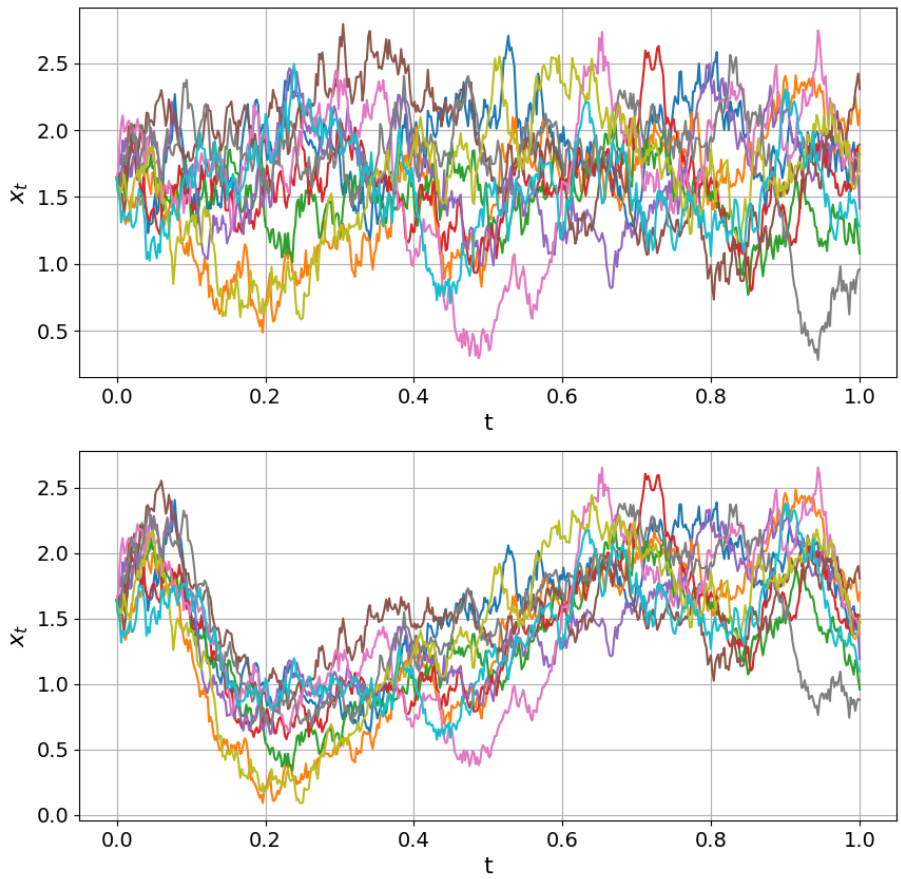

Fig. 2 A set of 10 random trajectories of the Ornstein-Uhlenbeck process (14) (upper figure) and the modified process (lower figure) (23), (29) with parameters $a_1(t)$ and $a_2(t)$ corresponded to Fig. 1. The same realizations of the Wiener process dW_t were used to construct the trajectories.

avoid this problem, we fix the rightmost point not equal to zero, but equal to the average of the found set from the previous iteration.

3. *The coefficients $a_1(t)$ and $a_2(t)$ are required to be continuous differentiable functions. Therefore, we have to approximate the obtained parameters $\hat{a}_{1,t}$, $\hat{a}_{2,t}$ by some smooth curve. For this problem ordinary polynomials are well suited, firstly, due to the simplicity of the function itself, and secondly, the problem of finding suitable coefficients of the polynomial is a linear least squares problem. We choose the degree of the polynomial in the calculations as follows: let $j \in 1, 2, ..., M$ be the number of the current iteration of the successive approximation algorithm. We set the degree of the polynomial equal to j, i. e. during the first iteration we approximate a straight line, then a parabola, and then up to a polynomial of degree M.*
4. *It is required to take into account the limit values for $a_2(t)$. There is no guarantee that the resulting smooth curve $a_2(t)$ everywhere on $t \in [0, 1]$ will be below the threshold. We solve this problem as follows: we approximate $a_2(t)$ by a polynomial*

using L_2-regularization with parameter $\alpha = 0$, then check whether the obtained approximation satisfies the threshold value (21), if not, we choose $\alpha = 0.01$, solve the problem again and check if the limit value is exceeded. If the limit value is not exceeded, we stop, otherwise we similarly solve the problem with coefficient α equal to 0.05, 0.1, 0.5, 1.0, 5.0, 10.0. In order to preserve the similarity of the parameter dynamics, we use the found value of α to approximate the parameter $a_1(t)$. If the regularization was not successful we set $a(t) = 0$.

4.3 Finding a copula parameters

To estimate the parameters α of the random process $\{x_t\}$, we could follow the logic of the maximum likelihood method.

Problem 1.

$$\alpha = \arg\max_{\alpha'} \log L(\alpha'; U_1, U_2).$$

However, for stochastic models more numerically stable optimization could be obtained using smoothed estimation for random process. Following the approach outlined in [24], a smoothed estimates of the random process have the following form. For discrete case

$$\mathbb{E}\left[\Psi(\lambda_t)|U_{1,t-1}, U_{2,t-1}\right] = \frac{\int \Psi(\lambda_t) p\left(\lambda_t|\Lambda_{t-1}, \alpha\right) f\left(U_{1,t-1}, U_{2,t-1}, \Lambda_{t-1}; \alpha\right) d\Lambda_t}{\int f\left(U_{1,t-1}, U_{2,t-1}, \Lambda_{t-1}; \alpha\right) d\Lambda_{t-1}}$$

and by analogy for continuous case

$$\mathbb{E}\left[\Psi(x_k)|U_{1,k-1}, U_{2,k-1}\right] = \\ = \frac{\int\limits_{F_1 \times \ldots \times F_{t_k}} \Psi(x_k) \prod\limits_{i=1}^{t_{k-1}} c(u_{1i}, u_{2i}, x_i|U_{1,i-1}, U_{2,i-1}, x_{i-1}) \cdot d\nu_{t_1,\ldots,t_k}}{\int\limits_{F_1 \times \ldots \times F_{t_{k-1}}} \prod\limits_{i=1}^{t_{k-1}} c(u_{1i}, u_{2i}, x_i|U_{1,i-1}, U_{2,i-1}, x_{i-1}) d\nu_{t_1,\ldots,t_{k-1}}}.$$

Then the smoothed copula parameter for discrete case is

$$\theta_t = \mathbb{E}\left[\Psi(\lambda_t)|U_{1,t-1}, U_{2,t-1}\right]$$

or

$$\theta_{t_k} = \mathbb{E}\left[\Psi(x_k)|U_{1,k-1}, U_{2,k-1}\right]$$

for continuous case. The expectation should be evaluated for each $t = 1, ..., T$ (for each t_k, $k = 1, ..., T$). For calculation here is used only first $t-1$ (t_{k-1}) observations. Note that calculation of numerator and denominator could be optimized using EIS technique as likelihood function (10). Then we can write down an expression for the

smoothed likelihood function

$$\tilde{L}(\alpha, U_1, U_2) = \prod_{i=1}^{T} c(u_{1i}, u_{2i}, \theta_i(\alpha)). \tag{33}$$

And the optimization problem would have the form
Problem 2.

$$\alpha = \arg\max_{\alpha'} \log \tilde{L}(\alpha'; U_1, U_2),$$

where the likelihood function (33) *is considered.*

We also note that it is important to choose a good initial point for optimization problem (we consider a local optimization). In our consideration we choose $(\theta_0, \mu_0, \nu_0) = (1.0, \mu_{\text{mle}}, 1.0)$, where μ_{mle} – the parameter obtained for the classical model with constant parameter. For the numerical stability we also constrain a latent process parameter $\theta > 0$. Starting point for the latent process is $x(t = 0) = \mu$.

5 Goodness of fit for dynamic copula models

To solve the problem of estimating the quality of copula models, we decided to use an approach based on the Rosenblatt transform. Consider a d-dimensional copula $C(u_1, ..., u_d)$, a vector $U = (u_1, ..., u_d)$ and an expression of the following form

$$C(u_i|u_1, ..., u_{i-1}) = \frac{\partial^{i-1} C(u_1, ..., u_i, 1, ..., 1)/\partial u_1 ... \partial u_{i-1}}{\partial^{i-1} C(u_1, ..., u_{i-1}, 1, ..., 1)/\partial u_1 ... \partial u_{i-1}}, \quad i = 2, ..., d.$$

where $C(u_i|u_1, ..., u_{i-1})$ is the conditional distribution function of the random variable u_i for given variables $u_1, ..., u_{i-1}$. It is argued that if the distribution parameters are chosen correctly [**?**], the quantities:

$$\begin{aligned}
u'_1 &= u_1, \\
u'_2 &= C(u_2|u_1), \\
u'_3 &= C(u_3|u_1, u_2), \\
&\vdots \\
u'_d &= C(u_d|u_1, ..., u_{d-1})
\end{aligned}$$

and the corresponding vector $U' = (u'_1, ..., u'_d)$ belong to an independent copula, or what is the same, the random variables $u'_1, ..., u'_d$ are pairwise independent and uniformly distributed on $[0, 1]$. The transformation $R : U \to U'$ is called the Rosenblatt transformation. The implementation details of this method are given in [**?**]. In our case, when the copula parameter is time-dependent, we will consider ordered sets $U_t = (u_{1t}, ..., u_{dt})$ – the ordering being determined by the simple ordering of the data in the original set. And for the copula parameter we consider the smoothed estimation θ_t introduced in Section 4.3.

From the independence and uniform distribution of the vector U' it follows that the univariate random variable Y:

$$Y = F_{\chi^2_d}\left(\sum_{i=1}^{d} \Phi^{-1}(u'_i)^2\right)$$

is uniformly distributed on $[0, 1]$, where $F_{\chi^2_d}$ is the distribution function of the χ^2-distribution with d degrees of freedom, Φ^{-1} is the inverse function to the distribution function of the standard normal distribution. The uniformity of Y can be checked, for example, by using the Cramer-von-Mises criterion.

6 Risk metrics VaR and CVaR

Let now pose the question of estimating the level of possible portfolio losses in a given probability interval. In this section we use the basic concepts and theorems given in the famous paper by Rockafellar and Uryasev [1] and adapt them to our model.

Let denote as $r = (r_1, ..., r_d)$ – the vector of logarithmic returns of portfolio W. Let us introduce the concept of return ρ_k:

$$\rho_k = \frac{s_k^i}{s_k^{i-1}} - 1 \tag{34}$$

where s_k^i is the value of asset number k at time i. Let us recall the concept of logarithmic return r_k introduced earlier:

$$r_k = \log\left(\frac{s_k^i}{s_k^{i-1}}\right).$$

For portfolios of d instruments we will consider the corresponding vectors $\rho = (\rho_1, ..., \rho_d)$ and $r = (r_1, ..., r_d)$. We denote the weights of assets in the portfolio as $w = (w_1, ..., w_d)$. Let us further define the type of financial portfolio whose risks we are going to consider.

Definition 2 (Investment portfolio)**.** *A portfolio W with instrument weights $w = (w_1, ..., w_d)$ will be called an investment portfolio if*

1. $\sum_{i=1}^{d} w_i = 1$,
2. $0 \leq w_i \leq 1$, $i = 1, ..., d$.

Let S_0 be the current value of the entire portfolio, S_1 be the value at the next point in time, s_i^0, s_i^1 be the corresponding instrument values. Next, we need to introduce a portfolio value loss function. Usually, in the literature, implying working with returns (34),

the loss function $f(w,\rho)$ is defined in the following form:

$$f(w,\rho) = 1 - \frac{S_1}{S_0} = 1 - \frac{\sum\limits_{i=1}^{d} s_i^1}{S_0} = 1 - \frac{\sum\limits_{i=1}^{d} \frac{s_i^1}{s_i^0} s_i^0}{S_0} = 1 - \sum_{i=1}^{d} \frac{s_i^1}{s_i^0} w_i = \sum_{i=1}^{d} \left(1 - \frac{s_i^1}{s_i^0}\right) w_i =$$

$$= -\sum_{i=1}^{d} \rho_i w_i = -(\rho, w).$$

The proposed form for loss function $f(w,\rho)$ is true only for the investment portfolio, and it is also used that $w_i = \frac{s_i^0}{S_0}$. However, this definition is not too convenient for our purposes, since we work with logarithmic returns. Therefore, we will use a more general representation of the loss function in our calculations:

$$f(w.r) = 1 - \frac{S_1}{S_0} = 1 - \frac{\sum\limits_{i=1}^{d} s_i^1}{S_0} = 1 - \frac{\sum\limits_{i=1}^{d} \frac{s_i^1}{s_i^0} s_i^0}{S_0} = 1 - \sum_{i=1}^{d} e^{r_i} w_i.$$

Let $p(r)$ be the joint distribution density of log-returns r. The probability that the value of the loss does not exceed the level of q is:

$$\Psi(w,q) = \int\limits_{f(w,r)\le q} p(r)\, dr.$$

Set some significance level γ (often chosen, for example, 0.95, 0.99, 0.999, etc.). Let us define the metric $\mathrm{VaR}_\gamma(w)$ of portfolio W as follows.
Definition 3 (Value-at-Risk, VaR)**.**

$$\mathrm{VaR}_\gamma(w) = q_\gamma(w) = \min\{q \in \mathbb{R} : \Psi(w,q) \ge \gamma\}.$$

In fact, $\mathrm{VaR}_\gamma(w)$ such a value of q that $\Psi(w,q) = \gamma$. However, it is possible that the function $p(r)$ plateaus at some point, then q is not the only value. Taking this fact into account in the definition of $\mathrm{VaR}_\gamma(w)$, the problem of minimization over the set of values $q \in \mathbb{R}$ arises. By calculating the average value-at-risk we arrive at the definition of the CVaR_γ metric.
Definition 4 (Conditional Value-at-Risk, CVaR)**.**

$$\mathrm{CVaR}_\gamma(w) = \phi_\gamma(w) = \frac{1}{1-\gamma} \int\limits_{f(r,w)\ge \mathrm{VaR}_\gamma(w)} f(w,r)p(r)\, dr.$$

Therefore CVaR is the average of the loss function over all events when losses exceed VaR. The multiplier $\frac{1}{1-\gamma}$ arises as a normalization factor. Indeed, from the definition

of 3 it follows that

$$\int\limits_{f(w,r)\geq \mathrm{VaR}_\gamma} p(r)dr = 1-\gamma.$$

The computation of CVaR_γ can be significantly optimized. For further calculations introduce auxiliary function $F_\gamma(w,q)$

$$F_\gamma(w,q) = q + \frac{1}{1-\gamma}\int\limits_{\mathbb{R}^d} (f(w,r)-q)_+ \, p(r)\, dr, \tag{35}$$

Where $x_+ = x$, if $x \geq 0$, 0 else. Let us recall two theorems, for the proof see the article [1].

Theorem 4. *The function $F_\gamma(w,q)$ is convex and continuously differentiable with respect to q. For arbitrary fixed w it is true that*

1. $\mathrm{CVaR}_\gamma(w) = \min\limits_{q\in\mathbb{R}} F_\gamma(w,q)$,
2. $\mathrm{VaR}_\gamma(w) =$ *left boundary of the set* $\arg\min\limits_{q\in\mathbb{R}} F_\gamma(w,q)$.

Theorem 5. *Let X be the set of admissible portfolio weights w. The problem of minimizing $\mathrm{CVaR}_\gamma(w)$ over the entire set X is equivalent to the problem of minimizing $F_\gamma(w,q)$ with respect to $(w,q)\in X\times\mathbb{R}$. In particular,*

$$\min_{w\in X} \mathrm{CVaR}_\gamma(w) = \min_{(w,q)\in X\times\mathbb{R}} F_\gamma(w,q).$$

Moreover, if $f(w,r)$ is convex with respect to w, then $F_\gamma(w,q)$ is convex with respect to (w,q) and $\mathrm{CVaR}_\gamma(w)$ is convex with respect to w. If X is a convex set, then the joint minimization problem is a convex programming problem.

Consider Monte Carlo estimates of the integral (35). Let the set of log-returns $(r^{(1)},...,r^{(N)})$ from the distribution $p(r)$ be known, then

$$\hat{F}_\gamma(w,q) = q + \frac{1}{N(1-\gamma)}\sum_{k=1}^{N}\left(f(r^{(k)},w)-q\right)_+.$$

The corresponding optimization problem will take the form.

Problem 3.

$$\begin{aligned}&\min_{(w,q)} \hat{F}_\gamma(w,q)\\ &\textit{subject to}\\ &w\in X.\end{aligned}$$

If we consider only investment portfolios and choose the loss function linear on w, then $\hat{F}_\gamma(w,q)$ will be a convex piecewise linear function. The problem of minimization

of such a function can be reduced to a linear programming problem if we introduce a set of auxiliary variables $v^{(j)} \geq 0$, $j = 1, ..., N$. Proceeding in this way, we obtain the following problem.

Problem 4.

$$\min_{(w,q,v^{(1)},...,v^{(N)})} q + \frac{1}{N(1-\gamma)} \sum_{j=1}^{N} v^{(j)}$$

subject to

$$v^{(j)} \geq f(w, r^{(j)}) - q, \quad j = 1, ..., N,$$

$$v^{(j)} \geq 0, \qquad j = 1, ..., N,$$

$$0 \leq w_i \leq 1, \qquad i = 1, ..., d,$$

$$\sum_{i=1}^{d} w_i = 1.$$

Therefore, to solve the Problem 4, it is sufficient to have a multivariate set of observations for a certain period, without using the law of their joint distribution. Such an approach can lead us to unexpected results. Often, zero weights can be obtained as a solution to the Problem 4, which is undesirable, especially when the set of investment instruments under consideration is quite small. The problem may lie in the fact that despite acceptable VaR and CVaR indicators, risks of *non-market* nature may come to the fore, which cannot be estimated within the framework of this model. However, it seems obvious that they multiply with decreasing number of instruments in the portfolio. To solve this problem, restrictions on the minimum level of weights are introduced. The linearity of the problem is not violated.

Let us now apply the information about joint distribution. Let an explicit form of the joint distribution $p(r)$ be obtained as a result of estimations:

$$p(r) = c(u_1, ..., u_d; \alpha)\, f_1(r_1) \cdot ... \cdot f_d(r_d),$$

where $c(u_1, ..., u_d; \alpha)$ – copula density, $u_i = F_i(r_i)$ – marginal distribution functions, $f_i(r_i)$ – marginal densities. The expression for the auxiliary function $F^c_\gamma(w, q)$, where the index c denotes the presence of copula in the formula, will take the following form

$$F^c_\gamma(w, q) = q + \frac{1}{1-\gamma} \int_{\mathbb{R}^d} (f(w, r) - q)_+ \, c(u_1, ..., u_d; \alpha) \cdot f_1(r_1)...f_d(r_d)\, dr_1...dr_d. \quad (36)$$

Using the algorithm 1 from Section 3.2 we can generate independently N vectors of pseudo observations $\left(u^{(1)}, ..., u^{(N)}\right)$ from copula $c(u_1, ..., u_d)$, where $u^{(k)} = \left(u_1^{(k)}, ..., u_d^{(k)}\right)$, thus obtaining a matrix of size $N \times d$. Now we need to apply an inverse transformation to logarithmic returns. Technically speaking, this transformation has the form $r_i^{(k)} = F_i^{-1}(u_i^{(k)})$, however, the honest computation of such an expression

could be a tedious problem, especially, for complex distributions like stable or generalized hyperbolic. We suggest to do the inversion non-parametrically as follows. Consider for simplicity a single distribution $F(x;\theta)$ of known form. Let the parameters θ be known (or estimated). To compute $r = F^{-1}(u)$, where $u \in (0,1)$ do the Algorithm 3.

Algorithm 3 Inverse cumulative distribution function

Require: $N \geq 10^5$
Ensure: $\hat{F}^{-1}(u)$
1: Sample $(x_1, ..., x_N)$ from distribution $F(x;\theta)$.
2: For given $u \in (0,1)$ compute inverse function $x = \hat{F}^{-1}(u)$ as quintile of set $(x_1, ..., x_N)$ of level u.

Such an approach allows us to not compute a complex integrals for distribution function numerically, which is possible, but an extremely slow operation, while sampling from known distribution, even with complex structure, could be considered as fast operation. For simpler distributions it is better to use exact formulas if they are available. Now we can apply Algorithm 3 to whole matrix $\left(u^{(1)}, ..., u^{(N)}\right)$ using distributions $F_1(r_1), ..., F_d(r_d)$. And for computing inverse function in multiple points $\left(u_i^{(1)}, ..., u_i^{(N)}\right)$ we generate a single sample $\left(r_i^{(1)}, ..., r_i^{(N)}\right)$ for any $i = 1, ..., d$. Thus we transform an original matrix of pseudo observations to matrix of log-returns $\left(r^{(1)}, ..., r^{(N)}\right)$ of size $N \times d$. Therefore, the Monte Carlo estimate of the integral (36) will take the form

$$\hat{F}_\gamma^c(w,q) = q + \frac{1}{N(1-\gamma)} \sum_{k=1}^{N} \left(f(r^{(k)}, w) - q \right)_+$$

and minimization problem is completely corresponds to Problem 3. Similarly, we could write the corresponding linear programming problem. With the information about joint distribution we can generate any large number of vectors $r^{(k)}$, but already at $N = 10^4$ the time of solving the problem will be too long to be used in real calculations (about several minutes when using python library scipy and optimize.linprog module on AMD Ryzen 5800h, 4.4 GHz processor). Therefore, we will not consider such a task in detail.

Let us make a remark concerning the stochastic copula model. Let as before the copula parameter be $\theta = \Psi(x_t)$, where x_t is a continuous stochastic Ito process on $t \in [0,T]$, where $T = 1$ as chosen in our consideration. Let $p(x_T) = p(x, T|x' = \mu, t' = 0)$ – the probability density of states of latent process $\{x_t\}$ at time $t = T$. The probability of such an event that the magnitude of portfolio losses at the next time instant does not exceed the level of q:

$$\Psi^{sc}(w,q) = \int\limits_{f(w,r) \leq q \times F_T} c(u_1, ..., u_d; \Lambda(x_T))\, f_1(r_1), ..., f_d(r_d) \cdot p(x_T)\, dr_1 ... dr_d\, dx_T,$$

where the index sc means the presence of stochastic copula in the formula. Here the integration is performed over logarithmic returns, as well as over the set of states of the random process $\{x_t\}$ at time $t = T$. The auxiliary function $F^{sc}_{\gamma}(w, q)$ will take the following form:

$$F^{sc}_{\gamma}(w,q) = q + \frac{1}{1-\gamma} \int\limits_{\mathbb{R}^d \times F_T} (f(w,r) - q)_+ \, c(u_1, ..., u_d; \Psi(x_T)) \cdot$$
$$\cdot f_1(r_1)...f_d(r_d) \cdot p(x_T)\, dr_1...dr_d\, dx_T. \quad (37)$$

Since the function (37) on the set of variables (w, q) has the same properties as the function $F^c(w, q)$ (35), the Theorems 4 and 5 also are applicable to it.

To build a Monte Carlo estimation we generate N trajectories of the random process $\{x_t\}$ on $t \in [0, T]$ and take the set of values of the random process at time $t = T$. This set could be easily obtained by generating from (18) since copula parameters here are known. Denote by $\theta^{(k)} = \Psi\left(x_T^{(k)}\right)$, $k = 1, ..., N$. Then generate independently N vectors $\left(u^{(1)}, ..., u^{(N)}\right)$ from copula $c(u_1, ..., u_d; \theta^{(k)})$, where the parameter $\theta^{(k)}$ differs for each iteration. Then apply Algorithm 3 to every column of matrix of pseudo observations distributions $\left(u^{(1)}, ..., u^{(N)}\right)$ using the distributions $F_1(r_1), ..., F_d(r_d)$, where $r^{(k)} = \left(r_1^{(k)}, ..., r_d^{(k)}\right)$ and thus obtaining matrix of log-returns $\left(r^{(1)}, ..., r^{(N)}\right)$ of size $N \times d$. The Monte Carlo evaluation of the integral (37) takes the form:

$$\hat{F}^{sc}_{\gamma}(w,q) = q + \frac{1}{N(1-\gamma)} \sum_{k=1}^{N} \left(f(r^{(k)}, w) - q\right)_+,$$

and again the minimization problem is completely corresponds to Problem 3. So we can use as a objective function in Problem 3 as well as functions $F_{\gamma}(w, q)$, $F^c_{\gamma}(w, q)$ and $F^{sc}_{\gamma}(w, q)$ in depending on the model under consideration. Hereafter, we are going to use last two functions.

7 Numerical results

As a dataset we will use a set of daily closing prices of the cryptocurrencies pair: ETH, XRP. Observation period from 2020-09-07 to 2022-01-19. The number of observations is 500 days. Statistical characteristics of logarithmic returns of this dataset are given in the Tables 3 and 4. The tables show that the time series under consideration have high volatility, as well as significant asymmetry and heavy tails of the corresponding distributions. In addition, there is a significant correlation between asset returns.

7.1 Estimation of marginal parameters

We investigate the applicability of the distributions (1)–(4) to the estimation of the parameters of marginal distributions. For this purpose, we used the Cramer-von-Mises

Table 3 Statistical characteristics of logarithmic returns: μ – mean, σ – standard deviation, S – skewness coefficient, K – kurtosis coefficient, V – average daily trading volume in billions U.S. dollars

Ticker	μ	σ	S	K	V
ETH	0.0043	0.052	−0.439	4.40	18.6
XRP	0.0022	0.078	0.039	10.72	0.5

Table 4 Correlation matrix of logarithmic returns

Ticker	ETH	XRP
ETH	1.00	0.56
XRP	0.56	1.00

statistical test and calculated the corresponding p-values. The results of the calculations are given in Table 5. The calculations show that a high value of p-value (> 0.05) is observed for hyperbolic and stable. In further calculations we will use the hyperbolic distribution as it provides the best computational performance among the complex distributions. Normal and Meixner distributions turned out to be inapplicable for this dataset.

Table 5 The p-values of the Cramer-von-Mises statistical test for estimating parameters of marginal distributions. The values lower than 0.05 denoted as 0.0.

Distribution	ETH	XRP
Normal	0.06	0.0
Hyperbolic	0.99	0.99
Stable	0.97	0.78
Meixner	0.55	0.0

7.2 Estimation of copula parameters

We investigate the accuracy of likelihood function calculations in the framework of stochastic copula models depending on the number of trajectories of a random process. Table 6 shows the results of calculations of the mean value of the log-likelihood function $\log L$ and the standard deviation σ over the 50 iterations of calculations. Based on the above data, we observe that the model with the Ornstein-Uhlenbeck process without modifications *scar-p-ou* (14), when the number of trajectories $N_{tr} = 10000$, provides similar accuracy compared to the model using the importance sampling technique *scar-m-ou* (28), when only $N_{tr} = 200$ trajectories are used. It is should be noted also that as the number of trajectories for the *scar-p-ou* model increases, not only does

σ decrease, but also the accuracy of the $\log L$ calculations increases, while the *scar-m-ou* model almost immediately gives a more stable prediction, however, too small a number of trajectories, in particular $N = 50$, may not be sufficient.

Furthermore, the calculations show that increasing the number of iterations of the successive approximation method from $M = 5$ to $M = 15$ has a positive effect on the Monte Carlo error of the calculations.

Table 6 Dependence of the standard deviation σ and the mean value of log-likelihood $\log L$ on the number of trajectories of the random process N_{tr} based on the results of 50 iterations of calculations. Rotated (180) Gumbel copula. The process parameters $\{x_t\}$: $\theta = 8.43$, $\mu = 1.64$, $\nu = 2.02$.

	scar-p-ou		*scar-m-ou*, $M = 5$		*scar-m-ou*, $M = 15$	
N_{tr}	$\log L$	σ	$\log L$	σ	$\log L$	σ
50	175.1	8.2	190.8	11.2	190.7	3.1
100	179.5	6.8	190.7	10.8	191.0	3.5
200	183.8	5.2	191.9	1.8	191.1	0.9
500	185.8	4.0	192.1	1.1	191.5	0.9
1000	187.1	3.6	192.4	0.7	191.9	0.8
5000	190.9	2.4	192.5	0.6	192.2	0.8
10^4	192.2	2.0	192.8	0.9	192.2	0.5
$5 \cdot 10^4$	193.9	1.6	192.9	0.7	192.5	0.5

7.3 Copula goodness of fit

Let us now consider the applicability of copulas of different types on the given dataset. Table 7 summarizes the results of estimating the parameters of Gumbel, Clayton, Frank, and Joe copulas for different computational methods: *MLE* – classical maximum likelihood estimation, as well as the above described *scar-p-ou* and *scar-m-ou*. The given data shows that the rotated (180) Gumbel copula performed the best on this dataset.

Table 7 Results of approximation of the Gumbel, Clayton, Frank, and Joe copulas for one iteration of the calculations and the corresponding value of $\log L$. The number of trajectories for *scar-m-ou* – 500, for *scar-p-ou* – 10000.

Copula	*MLE*	*scar-p-ou*	*scar-m-ou*
Gumbel	115	160	157
Gumbel (180)	158	187	193
Clayton	145	165	173
Frank	142	157	170
Joe (180)	147	172	175

Following the approach described at Section 5 we also performed a p-value calculation for the Rosenblatt transform method using Cramer-von-Mises test (Table 8). As one

can see the stochastic models show a significant improvement in p-value compared to the classical MLE.

Table 8 Quality estimation of copula models and corresponding p-values. The values lower than 0.02 denoted as 0.00.

Copula	*MLE*	*scar-p-ou*	*scar-m-ou*
Gumbel	0.00	0.76	0.31
Gumbel (180)	0.02	0.46	0.59
Clayton	0.00	0.37	0.22
Frank	0.02	0.77	0.71
Joe (180)	0.00	0.08	0.21

Since each of the models showed satisfactory results in terms of p-value, we chose the copula based on the value of $\log L$, as well as the overall stability of the calculations. For the dataset we consider, the rotated (180) Gumbel copula satisfies these criteria. We also note that stochastic models exhibit tangible variance in their $\log L$ and p-value estimates. Consequently, the results of a single estimate may vary from run to run.

7.4 VaR and CVaR results

Let us proceed to the calculation of VaR and CVaR in a time window. For this purpose, we have built a model with rotated (180) Gumbel copula and generalized hyperbolic marginals. We performed calculations of Problem 2 and 3 in a time window of 120 days. The following parameters were used: $N_{MC} = 10^5$ Monte Carlo iterations, $N_{tr} = 500$ latent process trajectories, $M = 10$. Note that, while working with stochastic copula models it is useful to pass the found on i-th iteration copula parameters as an initial point to the next optimization problem.

The results of CVaR metric calculations for different copula models are shown in Fig. 3. The behavior of two different copula models were compared here for equal-weighted portfolio. As one can see the models demonstrate similar results, which is very close to the empirical curve.

The difference between the classical model and stochastic one could be observed on the CVaR-optimized portfolio. The results of CVaR calculations are shown on Fig. 4. As one can see the optimal portfolios built by different models demonstrate similar estimations for CVaR. These models indeed shows more acceptable risk level in comparison to equal-weighted portfolio. However, the significant difference arises on the example of portfolio weights. For the bivariate investment portfolio it could be clearly demonstrated on a plot due to the fact, that if the weight of first instrument is w then for the second one is $1 - w$. This comparison is shown on Fig. 5. As it can be seen from the figure, the classical model actually resets the weight of one of the instruments to zero, which cannot be considered as an acceptable result. Despite the fact that the risk level predictions for both models are almost the same, the stochastic model shows more correct results.

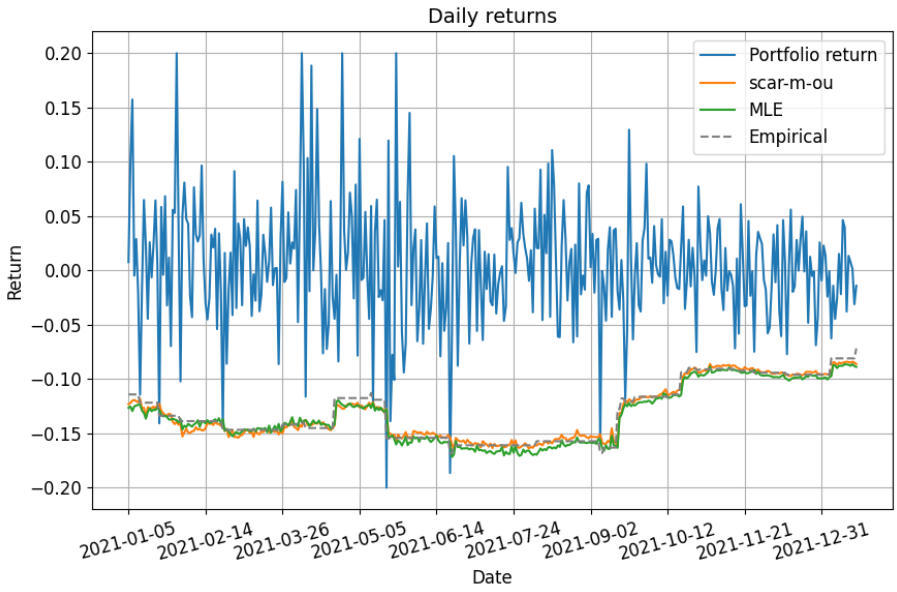

Fig. 3 Dynamics of the CVaR (95%) indicator of the equal-weighted portfolio. Returns on the figure were clipped to ±0.2.

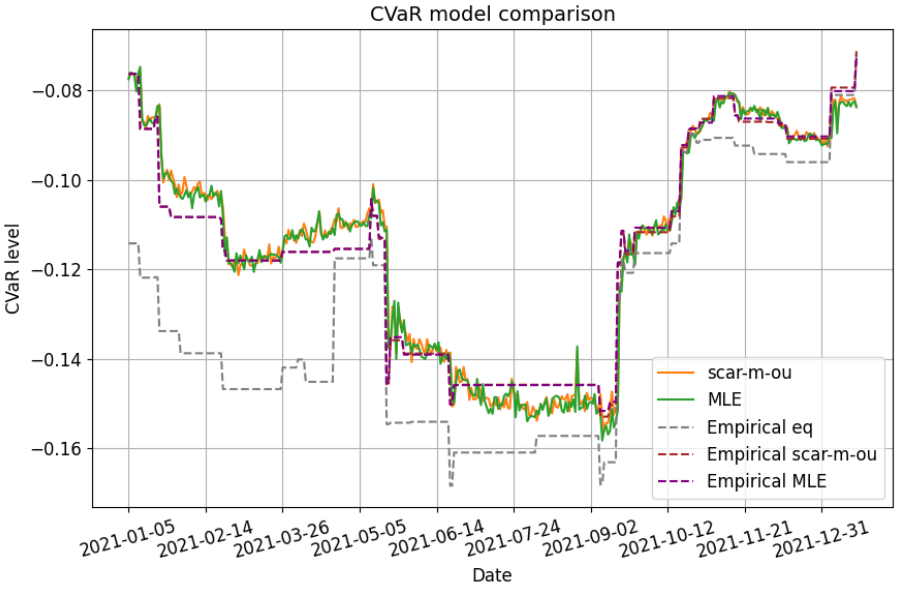

Fig. 4 Model comparison for dynamics of the CVaR (95%) indicator of the optimized portfolio.

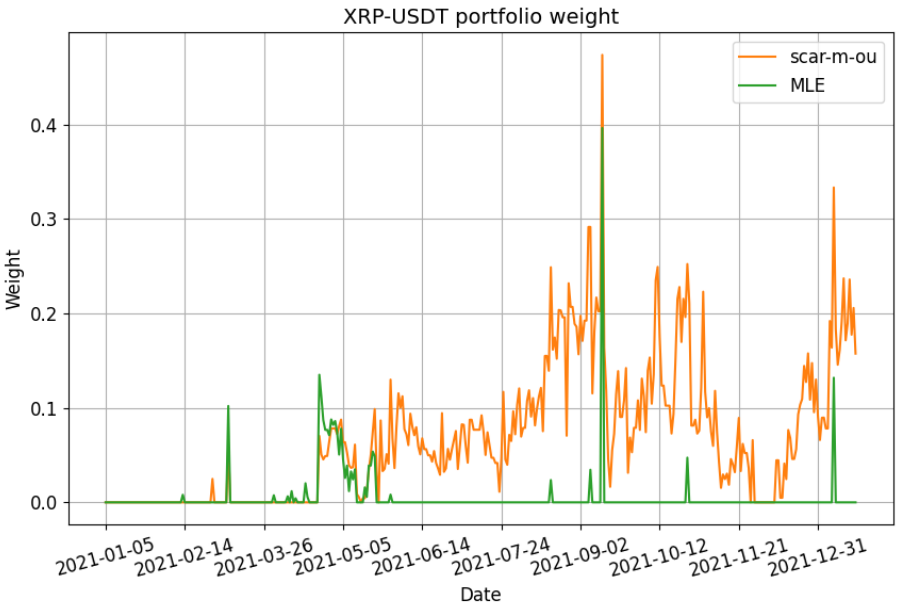

Fig. 5 Model comparison for dynamics of XRP-USDT weight in optimal bivariate portfolio.

8 Conclusion

In this paper, we propose the stochastic copula model in continuous time to solve the problem of risk assessment for financial portfolios. Unlike the discrete models considered in the literature earlier, continuous models are more complicated in the case of Monte Carlo variance reduction techniques due to the necessity to consider the drift and diffusion coefficients of the corresponding stochastic differential equations as a differentiable functions. As it was shown in the paper, this circumstance was possible to take into account by approximation of auxiliary parameters by polynomials of varying order. Furthermore, it has been shown that in the continuous case it is possible to construct a processes with a predetermined density. However, it is might be difficult to solve the problem of Monte Carlo variance reduction by the method of importance sampling due to the necessity of obtaining expressions integrable in elementary functions. Such a task may become a subject of further research.

The copula models proposed in this paper showed a significant improvements in terms of the likelihood function $\log L$ and p-values of corresponding goodness of fit test compared to the classic model with constant parameter. Proposed model have been used for calculation of portfolio Value-at-Risk and Conditional Value-at-Risk metrics. We have computed VaR and CVaR metrics on equal-weighted portfolio and CVaR-weighted portfolio.

Also as possible future research direction can be investigating alternative stochastic processes (fractional Brownian motion or Levy processes) and extending the model to higher-dimensional copulas.

Availability of data and materials

The dataset used and analyzed in the present study was acquired through the Binance API and is available as electronic supplementary material.

Abbreviations

APARCH: Asymmetric power autoregressive conditional heteroscedasticity
API: Application programming interface
AR(1): Autoregressive process of the first order
CVaR: Conditional Value at Risk
EIS: Efficient importance sampling
ES: Expected Shortfall
ETH: Ethereum
GAS: Generalized autoregressive score
LS: Laplace-Stieltjes
MLE: Maximum likelihood estimation
SCAR: Stochastic autoregressive copula models
VaR: Value at Risk
XRP: Ripple

Funding

This work was supported by the grant of the state program of the "Sirius" Federal Territory "Scientific and technological development of the "Sirius" Federal Territory" (Agreement No 18-03 date 10.09.2024).

Author information

Authors and Affiliations

Scientific Center for Information Technologies and Artificial Intelligence, Sirius University of Science and Technology, Sirius, Krasnodar region, Russia

Aleksey Novokhatskiy & Mikhail Semenov

Contributions

These authors contributed equally to this work.

Corresponding author

Correspondence to Mikhail Semenov.

Ethics declarations

Competing interests
The authors declare that they have no Conflict of interest.

References

[1] Rockafellar, R.T., Uryasev, S.: Optimization of conditional value-at-risk. Journal of Risk **2**(3), 21–41 (2000) https://doi.org/10.21314/JOR.2000.038

[2] Semenov, M., Smagulov, D.: Copula Models Comparison for Portfolio Risk Assessment. In: Kaz, M., Ilina, T., Medvedev, G. (eds.) Global Economics and Management: Transition to Economy 4.0. Springer Proceedings in Business and Economics, pp. 91–102. Springer, ??? (2019). https://doi.org/10.1007/978-3-030-26284-6_9

[3] Manner, H., Reznikova, O.: A survey on time-varying copulas: Specification, simulations, and application. Econometric Reviews **31**(6), 654–687 (2011) https://doi.org/10.1080/07474938.2011.608042

[4] Kim, Y., Hwang, S.-Y., Kim, J.-M., Kim, S.: Linear time-varying regression with copula–DCC–asymmetric–GARCH models for volatility: the co-movement between industrial electricity demand and financial factors. Applied Economics **55**(3), 255–272 (2023) https://doi.org/10.1080/00036846.2022.2086684

[5] Creal, D., Koopman, S., Lucas, A.: Generalized autoregressive score models with applications. Journal of Applied Econometrics **28**(5), 777–795 (2013) https://doi.org/10.1002/jae.1279

[6] Hafner, C.M., Manner, H.: Dynamic stochastic copula models: estimation, inference and applications. Journal of Applied Econometrics **27**(2), 269–295 (2012) https://doi.org/10.1002/jae.1197

[7] Almeida, C., Czado, C., Manner, H.: Modeling high-dimensional time-varying dependence using dynamic d-vine models. Applied Stochastic Models in Business and Industry **32**(5), 621–638 (2016) https://doi.org/10.1002/asmb.2182 https://onlinelibrary.wiley.com/doi/pdf/10.1002/asmb.2182

[8] Kielmann, J., Manner, H., Min, A.: Stock market returns and oil price shocks: A covar analysis based on dynamic vine copula models. Empirical Economics **62**, 1543–1574 (2022) https://doi.org/10.1007/s00181-021-02073-9

[9] Mba, J.C.: Assessing portfolio vulnerability to systemic risk: a vine copula and APARCH-DCC approach. Financial Innovation **10**, 20 (2024) https://doi.org/10.1186/s40854-023-00559-2

[10] Tadi, M., Witzany, J.: Copula-based trading of cointegrated cryptocurrency pairs.

Financial Innovation **11**, 40 (2025) https://doi.org/10.1186/s40854-024-00702-7

[11] Harvey, A.: Tracking a changing copula. Journal of Empirical Finance **17**, 485–500 (2010) https://doi.org/10.1016/j.jempfin.2009.10.004

[12] Goorbergh, R.W.J., Genest, C., Werker, B.J.M.: Bivariate option pricing using dynamic copula models. Insurance: Mathematics and Economics **37**, 101–114 (2005) https://doi.org/10.1016/j.insmatheco.2005.01.008

[13] Nguyen, H., Virbickaitė, A.: Modeling stock-oil co-dependence with dynamic stochastic midas copula models. Energy Economics **124**, 106738 (2023) https://doi.org/10.1016/j.eneco.2023.106738

[14] Barndorff-Nielsen, O.: Hyperbolic distributions and distributions on hyperbolae. Scandinavian Journal of statistics, 151–157 (1978)

[15] Nolan, J.: Stable Distributions: Models for Heavy-tailed Data. Birkhauser, ??? (2003)

[16] Schoutens, W., *et al.*: The Meixner Process: Theory and Applications in Finance. Eurandom Eindhoven, ??? (2002)

[17] Sklar, M.: Fonctions de repartition an dimensions et leurs marges. Publ. inst. statist. univ. Paris **8**, 229–231 (1959)

[18] Nelsen, R.B.: An Introduction to Copulas. Springer Series in Statistics. Springer, New York (2007). https://books.google.ru/books?id=B3ONT5rBv0wC

[19] Hofert, M.: Sampling archimedean copulas. Computational Statistics and Data Analysis **52**(12), 5163–5174 (2008) https://doi.org/10.1016/j.csda.2008.05.019

[20] Kimberling, C.H.: A probabilistic interpretation of complete monotonicity. Aequationes mathematicae **10**, 152–164 (1974)

[21] Feller, W.: An Introduction to Probability Theory and Its Applications, Volume 2 vol. 81. John Wiley & Sons, ??? (1991)

[22] Marshall, A.W., Olkin, I.: Families of multivariate distributions. Journal of the American statistical association **83**(403), 834–841 (1988) https://doi.org/10.2307/2289314

[23] Joe, H.: Multivariate Models and Multivariate Dependence Concepts. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, ??? (1997). https://doi.org/10.1201/9780367803896

[24] Liesenfeld, R., Richard, J.-F.: Univariate and multivariate stochastic volatility models: Estimation and diagnostics. Journal of Empirical Finance **10**, 505–531 (2003) https://doi.org/10.1016/S0927-5398(02)00072-5

[25] Øksendal, B.: Stochastic Differential Equations: An Introduction with Applications vol. 82. Journal of the American Statistical Association, ??? (2000). https://doi.org/10.1007/978-3-662-03185-8

Appendix A Solution for modified Ornstein–Uhlenbeck equation

Consider a liner stochastic equation in a general form

$$dX_t = (c_1(t)X_t + c_2(t))\,dt + (\sigma_1(t)X_t + \sigma_2(t))\,dW_t, \tag{A.38}$$

where $c_1(t), c_2(t), \sigma_1(t), \sigma_2(t)$ are some functions satisfying the Theorem 3. To build a general solution consider first a homogeneous equation

$$dS_t = c_1(t)S_t dt + \sigma_1(t)S_t dW_t. \tag{A.39}$$

To obtain the solution of (A.39) introduce a new variable $F = \log S$. Then using Ito's formula we get

$$d(\log S) = dF = \frac{\partial F}{\partial S} + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}dS^2 = \frac{1}{S}dS - \frac{1}{2S^2}dS^2 = \left(c_1 - \frac{1}{2}\sigma_1^2\right)dt + \sigma_1 dW_t.$$

By integration of the left and right parts from 0 to t we obtain

$$S_t = S_0 \exp\left(\int_0^t \left(c_1 - \frac{1}{2}\sigma_1^2\right)dp + \int_0^t \sigma_1 dW_p\right). \tag{A.40}$$

Now we can use equation (A.40) to solve a general problem (A.38). Let $S_0 = 1$ and consider $F = X_t/S_t$.

$$d\left(\frac{X_t}{S_t}\right) = d(X_t)\frac{1}{S_t} + X_t d\left(\frac{1}{S_t}\right) + d(X_t)d\left(\frac{1}{S_t}\right),$$

where

$$d\left(\frac{1}{S_t}\right) = -\frac{dS_t}{S_t^2} + \frac{(dS_t)^2}{S_t^3} = -\frac{c_1 dt + \sigma_1 dW_t}{S_t} + \frac{\sigma_1^2}{S_t}dt.$$

By collecting all terms we obtain

$$d\left(\frac{X_t}{S_t}\right) = \frac{c_2 - \sigma_1\sigma_2}{S_t}dt + \frac{\sigma_2}{S_t}dW_t.$$

So by integrating we obtain the solution of (A.38)

$$X_t = S_t \cdot \left(X_0 + \int_0^t \frac{c_2 - \sigma_1\sigma_2}{S_p} dp + \int_0^t \frac{\sigma_2}{S_p} dW_p \right), \tag{A.41}$$

where S_t satisfies (A.40) and $S_0 = 1$.

Consider now a modified Ornstein-Uhlenbeck equation

$$dx_t = A(x,t)dt + B(x,t)dW_t, \tag{A.42}$$

$$A(x,t) = \frac{\partial \tilde{x}}{\partial t} - \frac{(x-\tilde{x})}{2\tilde{\sigma}^2}\left(B^2(t) - \frac{\partial \tilde{\sigma}^2}{\partial t}\right),$$
$$B(t) = \frac{\nu}{\sqrt{1-2a_2\sigma^2}},$$

where auxiliary parameters are $a_1 = a_1(t)$, $a_2 = a_2(t)$ and $\tilde{x}$ and $\tilde{\sigma}^2$ satisfies (22). In our case we have the following equations for parameters

$$c_1(t) = -\frac{1}{2\tilde{\sigma}^2}\left(B^2(t) - \frac{\partial \tilde{\sigma}^2}{\partial t}\right),$$

$$c_2(t) = \frac{\partial \tilde{x}}{\partial t} + \frac{\tilde{x}}{2\tilde{\sigma}^2}\left(B^2(t) - \frac{\partial \tilde{\sigma}^2}{\partial t}\right),$$
$$\sigma_1(t) = 0,$$
$$\sigma_2(t) = \frac{\nu}{\sqrt{1-2a_2\sigma^2}}.$$

First of all, we calculate S_t. Since $\sigma_1 = 0$ the expression is greatly simplified and becomes an ordinary Riemann integral

$$S_t = \exp\left(\int_0^t c_1(p)dp\right).$$

Integral can be calculated explicitly

$$\int_0^t c_1(p)dp = -\int_0^t \frac{1}{2\tilde{\sigma}^2}\left(B^2(p) - \frac{\partial \tilde{\sigma}^2}{\partial p}\right)dp = -\int_0^t \frac{\theta - a_2\nu^2 - \sigma^2\frac{\partial a_2}{\partial p}}{1-2a_2\sigma^2}dp =$$
$$= -\left(\theta t + \frac{1}{2}\log\left(1-2a_2(t)\sigma^2(t)\right)\right),$$

where equations (25) and (27) were used. Therefore,

$$S_t = \frac{e^{-\theta t}}{\sqrt{1-2a_2\sigma^2}}.$$

Let us proceed to the calculation of the integrals in (A.41).

$$\int_0^t \frac{c_2 - \sigma_1\sigma_2}{S_p} dp = \int_0^t \frac{c_2}{S_p} dp = \int_0^t \frac{1}{S_p}\left(\frac{\partial \tilde{x}}{\partial p} + \frac{\tilde{x}}{2\tilde{\sigma}^2}\left(B^2 - \frac{\partial \tilde{\sigma}^2}{\partial p}\right)\right) dp =$$

$$= \int_0^t d\left(\frac{\tilde{x}}{S_t}\right) + \left(\frac{\tilde{x}}{S_p^2}\frac{\partial S_p}{\partial p} - \frac{\tilde{x}}{S_p}c_1(p)\right) dp = \int_0^t d\left(\frac{\tilde{x}}{S_p}\right) = \frac{\tilde{x}(t)}{S_t} - X_0,$$

where the following equalities were used

$$\frac{1}{S_t}\frac{\partial \tilde{x}}{\partial t} = d\left(\frac{\tilde{x}}{S_t}\right) + \frac{\tilde{x}}{S_t^2}\frac{\partial S}{\partial t},$$

$$\frac{\partial S}{\partial t} = S_t \cdot c_1.$$

The last integral in (A.41) has the form

$$\int_0^t \frac{\sigma_2}{S_p} dW_p = \int_0^t \nu e^{\theta p} dW_p.$$

Collecting all calculated terms, we can write down the solution of (A.42)

$$x_t = \tilde{x} + \frac{\nu e^{-\theta t}}{\sqrt{1 - 2a_2\sigma^2}} \int_0^t e^{\theta p} dW_p, \quad t \geq 0.$$