Robust numerical scheme for stochastic copula models

A. A. Novokhatskiy, M.E. Semenov

February 28, 2026

Contents

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 Markov property and known transition density of the Ornstein-Uhlenbeck process allow the high-dimensional likelihood integral to be factored into a sequence of two-dimensional integrals computed via the transfer matrix method on a discretized grid. This approach yields a numerically exact evaluation of the likelihood function, avoiding the variance and bias inherent in Monte Carlo estimators. We discuss implementation details including adaptive grid refinement and the use of sparse transfer matrices for long time series. The proposed model is validated on cryptocurrency market data, demonstrating a significant improvement in both log-likelihood and goodness-of-fit compared to static copula models.

Keywords Stochastic copula models, Transfer matrix method, Ornstein-Uhlenbeck process, Adaptive grid, Sparse matrices

1 Introduction

Modeling the joint behavior of financial asset returns requires capturing dependencies that may evolve over time. Copula models provide a flexible framework for constructing multivariate distributions by separating the modeling of marginal distributions from the dependence structure [Skl59, Nel07]. However, standard copula models with fixed parameters cannot account for the time-varying nature of dependencies observed in financial markets.

In our previous article [SS19], we presented the results of modeling risk measures based on the copula approach for an optimal portfolio on short financial time series comprising 253 observations in total. Our findings showed 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. It is necessary to take into account the dynamic changes in the relationships observed in the original data.

In the review [MR12] the authors distinguished between parametric, semi-parametric and non-parametric copula models that account for 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 [KHKK23]. 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 [CKL13], where the dependence of the copula on time is described by a parametric function with an autoregressive term. Stochastic autoregressive copula models (SCAR) [HM12] extend this concept by including a 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 [ACM16, KMM22] 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 [Mba24] 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 the class of nonparametric models, we refer to the paper [Har10] where the authors use a quantile function (estimating time-varying quantiles) to decide on a change in the model. The review by [MR12] presents several models where parameter responsible for the change in the model over time depends on the marginal distribution through the conditional variance [vdGGW05]. The article [NV23] 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).

A common limitation of Monte Carlo approaches used in existing SCAR models [HM12, LR03] is the bias inherent in the log-mean-exp estimator: by Jensen's inequality, $\mathbb{E}[\log \hat{L}] \leq \log L$, with the bias decaying only as $O(1/N)$ in the number of trajectories. For long time series this bias can be substantial, requiring an impractically large number of Monte Carlo paths.

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 a continuous-time stochastic process, the Ornstein-Uhlenbeck process, to capture the time-varying dependencies between assets. The key methodological contribution is the use of the *transfer matrix* approach, which exploits the Markov structure and known Gaussian transition density of the OU process to reduce the high-dimensional likelihood integral to a sequence of numerically tractable two-dimensional integrals evaluated on a grid via quadrature.

The primary objectives of this research are twofold: (1) to introduce time dynamics as a copula parameter through a stochastic differential equation in continuous time with a numerically exact likelihood evaluation via the transfer matrix method, and (2) to evaluate the effectiveness of the novel copula model using real datasets of different lengths from the cryptocurrency 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 develop a transfer matrix scheme that yields a numerically exact evaluation of the likelihood function, avoiding the variance and bias inherent in Monte Carlo estimators used in earlier SCAR models. We discuss implementation aspects of the transfer matrix method, including adaptive grid refinement and the use of sparse matrices to handle long time series efficiently.

The outcomes of this study can contribute to creation of more effective methods for modeling time-varying dependence structures in financial data.

The implementation of the proposed model is available as an open-source Python package at

This paper is structured as follows. This paper is structured as follows. Section 2 reviews various types of copula models, including Archimedean copulas, sampling techniques for them, and stochastic autoregressive copulas (SCAR). The continuous SCAR model and the transfer matrix method for exact likelihood computation are developed in Section 3, including implementation details for the numerical scheme. Section 4 focuses on goodness-of-fit tests for proposed copula models. Numerical experiments with real datasets from the cryptocurrency market are presented in Section 5. Finally, Section 6 concludes the paper and discusses the future work.

2 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 [Skl59].

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 [Nel07]. 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 of 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

$$f(x_1, \ldots, x_d) = c(F_1(x_1), \ldots, F_d(x_d)) \prod_{i=1}^{d} f_i(x_i), (x_1, \ldots, x_d) \in \mathbb{R}^d. \tag{2.1}$$

General problem of finding parameters of joint distribution implies solving the optimization problem for whole function (2.1), but it is convenient to solve two separate problems: for copula and for marginal distributions. For copula one can use 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*:

$$u_{ij} = \frac{rank(x_{ij})}{n+1}, \quad i = 1, ..., d, \quad j = 1, ..., n, \tag{2.2}$$

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)

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

2.1 Archimedian copulas

In order to construct a multivariate distribution based on copulas one should use a some class. In the study we used Archimedian copulas.

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

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

The following theorem [Kim74] 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 Archimedian copulas in explicit form are given in Table 1 (see [Nel07] for details). For given copulas the conditions of Theorem 1 holds.

Table 1: Archimedian 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:

$$C_{90}(u_1, u_2) = u_2 - C(1 - u_1, u_2),$$
$$C_{180}(u_1, u_2) = u_1 + u_2 - C(1 - u_1, 1 - u_2),$$
$$C_{270}(u_1, u_2) = u_1 - C(u_1, 1 - u_2).$$

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

2.2 Sampling Archimedian copulas

An efficient generation algorithm from Archimedian copula is closely related to the Laplace-Stieltjes transform of the inverse generator function (see Hofert [Hof08] 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*

$$\psi(t) = \int\limits_0^\infty e^{-tx} dF(x), \quad t \in [0,\infty).$$

The following theorem (Bernstein, see [Fel91]) 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 [MO88]).

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 $U_i = \psi\left(-\log(X_i)/V\right), \quad i = 1, ..., d$.

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 [Joe97]. 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 distribtuion 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 Archimedian 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	$1-(1-\exp\left(-t\right))^{1/\theta}$	$y_k = (-1)^{k+1}\binom{1/\theta}{k}, k \in \mathbb{N}$	$[1,\infty)$
Clayton	$(1+t\theta)^{-1/\theta}$	$\Gamma(1/\theta, 1)$	$(0,\infty)$

2.3 Stochastic autoregressive copulas (SCAR)

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

$$
\begin{aligned}
(u_{1t}, u_{2t}) &\sim c(u_{1t}, u_{2t}; \theta_t),\\
\theta_t &= \Psi(\lambda_t),\\
\lambda_t &= \alpha_1 + \alpha_2 \lambda_{t-1} + \alpha_3 \eta_t,
\end{aligned}
\tag{2.3}
$$

where $\Psi : \mathbb{R} \to \Theta$ is a suitable transformation that guarantees that the copula parameter remains in its domain of definition. It is assumed that the underlying process $\{\lambda_t\}_{t=1}^T$ is latent (i.e. unobservable) and is described by a 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 [LR03], discussed among other time-varying models in Manner and Reznikova [MR12] and was developed in Almeida, Czado, Manner [ACM16] for vine copulas.

Since the process $\{\lambda_t\}_{t=1}^T$ is unobservable, estimation for $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ cannot be obtained explicitly. The likelihood function has the form

$$
L(\alpha; U_1, U_2) = \int \prod_{t=1}^T c\left(u_{1t}, u_{2t}; \Psi(\lambda_t)\right) p\left(\lambda_t \mid \lambda_{t-1}; \alpha\right) d\Lambda,
\tag{2.4}
$$

where $U_{1,t} = \{u_{1i}\}_{i=1}^t$, $U_{2,t} = \{u_{2i}\}_{i=1}^t$ is a set of pseudo-observations known up to time t, and $p(\lambda_t|\lambda_{t-1}; \alpha)$ is the transition density. A straightforward Monte Carlo estimate of (2.4) is obtained by sampling N trajectories from the prior density p:

$$
\hat{L}_N(\alpha; U_1, U_2) = \frac{1}{N}\sum_{i=1}^N \prod_{t=1}^T c\left(u_{1t}, u_{2t}; \Psi\left(\lambda_t^{(i)}\right)\right),
\tag{2.5}
$$

where $\lambda_t^{(i)}$ is the value of the i-th trajectory at time t. This estimator, referred to as the *p-sampler*, is unbiased for L but is very inefficient for large T: the variance of the product grows exponentially, and the log-likelihood estimate $\log \hat{L}_N$ is biased downward by Jensen's inequality.

The Monte Carlo variance can be reduced by *efficient importance sampling* (EIS) [LR03]: one introduces an auxiliary density $m(\lambda_t|\lambda_{t-1}; a_t)$ with parameters a_t chosen to minimize the variance of the importance weights. The resulting estimator, known as the *m-sampler*, takes the form

$$
\hat{L}_N(\alpha; U_1, U_2) = \frac{1}{N}\sum_{i=1}^N \prod_{t=1}^T \frac{c\left(u_{1t}, u_{2t}; \Psi\left(\lambda_t^{(i)}\right)\right) p\left(\lambda_t^{(i)}|\lambda_{t-1}^{(i)}; \alpha\right)}{m\left(\lambda_t^{(i)}|\lambda_{t-1}^{(i)}; a_t\right)},
$$

where the trajectories $\lambda_t^{(i)}$ are now sampled from the density m. While the m-sampler substantially reduces the required number of trajectories (typically $100 < N < 200$), both Monte Carlo approaches remain subject to the Jensen bias in log-likelihood and to practical difficulties in the continuous-time setting (see Section 3.3 for a deterministic alternative).

3 Continuous SCAR model

3.1 Model formulation

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 (2.3) is the Ornstein-Uhlenbeck process

$$dx_t = (\alpha_1 + \alpha_2 x_t)\, dt + \alpha_3 dW_t, \quad t \geq 0, \tag{3.1}$$

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

$$dx_t = \theta\,(\mu - x_t)\, dt + \nu dW_t, \quad t \geq 0, \tag{3.2}$$

where the parameters of the processes (3.1) and (3.2) are related as follows:

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

Let the probability space $(\Omega, \mathcal{F}, P)$ be given. Consider on it the random process $\{x_t\}$, $t \in [0,1]$. Let us divide the interval $[0,1]$ by the points $0 \leq t_1 \leq t_2 \leq \ldots \leq t_T \leq 1$. Let us introduce the following probability measure:

$$\nu_{t_1,\ldots,t_T}(F_1 \times \ldots \times F_T) = \int_{F_1 \times \ldots \times F_T} \prod_{k=1}^{T} p(x_k, t_k | x_{k-1}, t_{k-1})\, dx_1 dx_2 \ldots dx_T,$$

where $p(x, t | x', t')$ – probability density function of transition from the state (x', t') to (x, t). We denote as $\alpha = (\theta, \mu, \nu)$ the set of constant parameters of stochastic equation (3.2) in canonical form. By analogy with the expression (2.4), define the likelihood function as follows:

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

$$= \int_{F_1 \times \ldots \times F_T} \prod_{t=1}^{T} c(u_{1t}, u_{2t}, \Psi(x_t) | U_{1,t-1}, U_{2,t-1}, x_{t-1}; \alpha)\, d\nu_{t_1,\ldots,t_T}(F_1 \times \ldots \times F_T), \quad (3.3)$$

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

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

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.

3.2 Transition density of the OU process

For a continuous stochastic Ito process of the form

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

the transition density $p(x, t | x', t')$ satisfies the Fokker-Planck equation

$$\frac{\partial p(x, t)}{\partial t} = -\frac{\partial}{\partial x}\left(A(x, t)p(x, t)\right) + \frac{\partial^2}{\partial x^2}\left(\frac{B^2(x, t)}{2}p(x, t)\right). \tag{3.4}$$

For the OU process (3.2), the solution of (3.4) is known explicitly. Multiplying both sides of (3.2) by the integrating factor $e^{\theta t}$ and integrating from t' to t yields

$$x_t = \mu + (x_{t'} - \mu)e^{-\theta(t-t')} + \nu \int_{t'}^{t} e^{-\theta(t-s)}\, dW_s. \tag{3.5}$$

The stochastic integral in (3.5) is a Gaussian random variable with zero mean and variance $\frac{\nu^2}{2\theta}(1 - e^{-2\theta(t-t')})$. Hence the conditional distribution is

$$x_t \,|\, x_{t'} \,\sim\, \mathcal{N}\!\left(\bar{x},\ \sigma_c^2\right),$$

where

$$\bar{x} = \mu + (x_{t'} - \mu)\, e^{-\theta \Delta t},$$
$$\sigma_c^2 = \frac{\nu^2}{2\theta}\left(1 - e^{-2\theta \Delta t}\right),$$

and $\Delta t = t - t'$ is the time step. It is convenient to introduce the notation

$$\rho = e^{-\theta \Delta t}, \qquad \sigma^2 = \frac{\nu^2}{2\theta},$$

so that $\sigma_c^2 = \sigma^2(1 - \rho^2)$ and $\bar{x} = \mu + \rho(x_{t'} - \mu)$. For a stationary OU process ($\theta > 0$), the initial distribution is given by the stationary (invariant) measure

$$x_0 \sim \mathcal{N}\!\left(\mu, \sigma^2\right).$$

Introducing centered coordinates $z = x - \mu$, the transition density and the stationary density take the form

$$p(z_t | z_{t'}) = \frac{1}{\sigma_c \sqrt{2\pi}} \exp\!\left(-\frac{(z_t - \rho\, z_{t'})^2}{2\sigma_c^2}\right), \tag{3.6}$$

$$p_0(z) = \frac{1}{\sigma \sqrt{2\pi}} \exp\!\left(-\frac{z^2}{2\sigma^2}\right). \tag{3.7}$$

3.3   Transfer matrix method for likelihood evaluation

Substituting the factorized joint density of the stationary OU process into (3.3), the likelihood function takes the form

$$L = \int \left[\prod_{t=0}^{T} f_t(x_t)\right] \cdot p_0(x_0) \prod_{t=1}^{T} p(x_t | x_{t-1})\, dx_0 \cdots dx_T,$$

where $f_t(x_t) = c(u_{1t}, u_{2t}; \Psi(x_t))$ is the copula density evaluated at the t-th observation for a given value of the latent process. The Markov property of the OU process ensures that this integral has a chain structure.

Backward messages.   Define the backward messages recursively. Starting from the last observation, set

$$m_T(x_{T-1}) = \int f_T(x_T)\, p(x_T | x_{T-1})\, dx_T. \tag{3.8}$$

For $t = T - 1, \ldots, 1$:

$$m_t(x_{t-1}) = \int f_t(x_t)\, p(x_t | x_{t-1})\, m_{t+1}(x_t)\, dx_t. \tag{3.9}$$

The likelihood is then

$$L = \int f_0(x_0)\, p_0(x_0)\, m_1(x_0)\, dx_0. \tag{3.10}$$

Each message m_t is a function of a single variable, so the problem is reduced from a $(T+1)$-dimensional integral to a sequence of T one-dimensional integrals.

Discretization on a grid.   In centered coordinates $z = x - \mu$, introduce a uniform grid of K points

$$z_j = -R\sigma + (j-1)\Delta z, \quad j = 1, \ldots, K, \qquad \Delta z = \frac{2R\sigma}{K-1},$$

where $R > 0$ is a parameter controlling the grid extent in units of the stationary standard deviation σ. Each function is represented as a vector of length K:

$$\mathbf{f}_t = (f_t(z_1 + \mu), \ldots, f_t(z_K + \mu)), \qquad \mathbf{m}_t = (m_t(z_1), \ldots, m_t(z_K)).$$

Transfer matrix. The transition density (3.6) is discretized as the matrix $\mathbf{T} \in \mathbb{R}^{K \times K}$:

$$T_{ji} = p(z_t = z_i \,|\, z_{t-1} = z_j) = \frac{1}{\sigma_c \sqrt{2\pi}} \exp\left(-\frac{(z_i - \rho \, z_j)^2}{2\sigma_c^2}\right),$$

where the row index j corresponds to the previous state $z_{t-1} = z_j$ and the column index i to the next state $z_t = z_i$.

The integrals in (3.8)–(3.9) are approximated by the trapezoidal quadrature. Let $\mathbf{w} \in \mathbb{R}^K$ be the vector of quadrature weights:

$$w_i = \Delta z \ \text{ for } i = 2, \ldots, K-1, \qquad w_1 = w_K = \frac{\Delta z}{2}.$$

Define the weighted transfer matrix $\widetilde{\mathbf{T}}$ with entries $\widetilde{T}_{ji} = T_{ji} \cdot w_i$. Then the backward recursion (3.9) becomes the matrix-vector operation

$$\mathbf{m}_t = \widetilde{\mathbf{T}} \left(\mathbf{f}_t \odot \mathbf{m}_{t+1}\right), \qquad t = T-1, \ldots, 1, \tag{3.11}$$

where $\odot$ denotes the Hadamard (element-wise) product, and the initial step is $\mathbf{m}_T = \widetilde{\mathbf{T}} \, \mathbf{f}_T$.

Final convolution. The stationary density (3.7) is discretized as the vector $\mathbf{p}_0 = (p_0(z_1), \ldots, p_0(z_K))$. The likelihood (3.10) is then

$$L \approx \sum_{i=1}^{K} f_0(z_i + \mu) \cdot p_0(z_i) \cdot m_1(z_i) \cdot w_i = \mathbf{w}^\top \left(\mathbf{f}_0 \odot \mathbf{p}_0 \odot \mathbf{m}_1\right).$$

Numerical stabilization. During the backward pass, the entries of $\mathbf{m}_t$ may decay (or grow) exponentially, leading to numerical underflow or overflow. To avoid this, after each step we normalize

$$\mathbf{m}_t \leftarrow \frac{\mathbf{m}_t}{c_t}, \qquad c_t = \max_j |m_t(z_j)|,$$

and accumulate the log-scale $S = \sum_t \log c_t$. The log-likelihood is then

$$\log L = \log\left(\sum_{i=1}^{K} f_0(z_i + \mu) \cdot p_0(z_i) \cdot m_1(z_i) \cdot w_i\right) + S. \tag{3.12}$$

Computational complexity. Each of the T backward steps requires one matrix-vector multiplication of cost $O(K^2)$ and one element-wise product of cost $O(K)$. The total complexity is $O(TK^2)$, which is efficient for typical values $T \sim 10^2$–10^3 and $K \sim 10^2$–10^3.

Comparison with Monte Carlo methods. The standard Monte Carlo estimator of the likelihood (the p-sampler (2.5)) computes $\hat{L} = \frac{1}{N} \sum_{i=1}^{N} \prod_t c(u_{1t}, u_{2t}; \Psi(x_t^{(i)}))$, then takes $\log \hat{L}$. By Jensen's inequality, $\mathbb{E}[\log \hat{L}] \leq \log L$, so this estimator is biased downward, with bias of order $O(1/N)$. Moreover, the variance of the product $\prod_t c(\cdot)$ grows exponentially with T, requiring an exponentially large number of trajectories for convergence. The transfer matrix method avoids both issues by computing the integral deterministically on a grid.

Finding copula parameters. To estimate the parameters $\alpha = (\theta, \mu, \nu)$ of the latent OU process, we follow the maximum likelihood approach.

Problem 1

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

where the likelihood function (3.12) is considered.

3.4 Implementation details

We now discuss several implementation aspects of the transfer matrix method that are important for achieving both accuracy and computational efficiency, particularly for long time series.

Adaptive grid refinement. The accuracy of the trapezoidal quadrature in (3.11) depends on how well the grid resolves the transition kernel. The conditional standard deviation $\sigma_c = \sigma\sqrt{1-\rho^2}$ decreases as $\theta\Delta t$ grows (i.e., when the mean-reversion is strong relative to the time step). If the grid spacing Δz is too coarse compared to σ_c, the Gaussian kernel $p(z_t|z_{t-1})$ is under-resolved and the quadrature error becomes significant.

To address this, we employ an adaptive rule for the grid size: given the user-specified minimum K and grid range R, we compute

$$K_{\text{eff}} = \max\left(K, \left\lceil \frac{2R\sigma}{\sigma_c/4} \right\rceil + 1\right), \tag{3.13}$$

which guarantees at least 4 grid points per conditional standard deviation σ_c. This ensures adequate resolution of the transition kernel regardless of the OU parameters. For moderate mean-reversion ($\theta\Delta t \ll 1$), the kernel is wide and $K_{\text{eff}} \approx K$; for strong mean-reversion ($\theta\Delta t \gg 1$), the kernel becomes narrow and the grid is automatically refined.

Dense and sparse transfer matrices. When the transition kernel is wide (i.e., the bandwidth of the matrix $\widetilde{\mathbf{T}}$ covers a significant fraction of the grid), the transfer matrix is essentially full and the dense matrix-vector multiplication $\widetilde{\mathbf{T}}\mathbf{v}$ with cost $O(K^2)$ is optimal. However, for long time series with many observations, Δt is small, the correlation $\rho = e^{-\theta\Delta t}$ is close to 1, and the conditional variance $\sigma_c^2 = \sigma^2(1-\rho^2)$ becomes small. In this regime the Gaussian kernel is sharply peaked: the matrix $\widetilde{\mathbf{T}}$ has only $O(b)$ significant entries per row, where $b = \lceil 5\sigma_c/\Delta z \rceil$ is the half-bandwidth.

When $b < K/4$, we store $\widetilde{\mathbf{T}}$ as a sparse matrix in CSR (Compressed Sparse Row) format. The matrix-vector product then costs $O(K \cdot b)$ instead of $O(K^2)$, which can provide a substantial speedup for large K. The selection between dense and sparse representations is made automatically based on the ratio b/K:

$$\text{storage} = \begin{cases} \text{dense}, & \text{if } b \geq K/4, \\ \text{sparse (CSR)}, & \text{if } b < K/4. \end{cases}$$

This strategy is particularly beneficial for high-frequency data, where T can be on the order of 10^4 and the adaptive grid (3.13) may increase K to several thousand.

Forward pass. In addition to the backward pass used for likelihood evaluation, several quantities of interest—the smoothed copula parameter $\bar{\theta}_k = \mathbb{E}[\Psi(x_k)|U_{1,k-1}, U_{2,k-1}]$, the mixture Rosenblatt transform (4.1), and the predictive distribution of x_T—require a *forward pass*. In the forward pass, we propagate the density $\alpha_k(z) \propto p(z_k|U_{1,k-1}, U_{2,k-1})$ from $k=0$ to $k=T$ by alternating updates (absorption of observation u_k) and predictions (propagation through the transition kernel):

$$\alpha_k(z) \leftarrow f_k(z) \cdot \alpha_k(z),$$
$$\alpha_{k+1}(z') = \int p(z'|z)\, \alpha_k(z)\, dz \approx \widetilde{\mathbf{T}}^\top(\boldsymbol{\alpha}_k \odot \mathbf{w}),$$

with normalization $\boldsymbol{\alpha}_k \leftarrow \boldsymbol{\alpha}_k / \max_j |\alpha_k(z_j)|$ to prevent under/overflow. Note that the forward pass uses the transpose $\widetilde{\mathbf{T}}^\top$, which for the sparse CSR format is computed via the CSC representation at no additional cost.

All forward-pass quantities share the same grid, transfer operator, and propagation loop; they differ only in the per-step computation performed on the predictive weights. This allows a unified implementation with a callback mechanism, reducing code duplication and ensuring that the adaptive grid and sparse/dense selection are applied consistently across all transfer matrix operations.

4 Goodness of fit for dynamic copula models

To assess the quality of copula models, we 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 the conditional distribution functions

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

If the distribution parameters are chosen correctly [HH15], 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, i.e., 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 [HH15].

For the bivariate case $(d = 2)$, the Rosenblatt transform reduces to

$$u_1' = u_1, \qquad u_2' = h(u_2, u_1; \theta) := C(u_2|u_1; \theta) = \frac{\partial C(u_1, u_2; \theta)}{\partial u_1},$$

where h is the h-function of the copula.

In the case of a constant copula parameter (MLE model), θ is a fixed scalar. For the stochastic copula model, the copula parameter $\theta_t = \Psi(x_t)$ is time-dependent and latent. Rather than plugging in a point estimate of x_t, which would ignore the uncertainty of the latent process, we compute the *mixture Rosenblatt transform* by integrating the h-function over the predictive distribution of the latent state:

$$u_{2,t}' = \mathbb{E}\big[h(u_{2t}, u_{1t}; \Psi(x_t)) \,\big|\, U_{1,t-1}, U_{2,t-1}\big] = \int h(u_{2t}, u_{1t}; \Psi(x_t)) \, p(x_t \,|\, U_{1,t-1}, U_{2,t-1}) \, dx_t. \quad (4.1)$$

The predictive density $p(x_t \,|\, U_{1,t-1}, U_{2,t-1})$ is obtained from the transfer matrix forward pass described in Section 3.3. On the grid, the computation reduces to a weighted sum:

$$u_{2,t}' \approx \sum_{j=1}^{K} h(u_{2t}, u_{1t}; \Psi(z_j + \mu)) \cdot \hat{p}_t(z_j) \cdot w_j,$$

where $\hat{p}_t(z_j)$ is the normalized forward message at time t and w_j are the quadrature weights.

In our case, when the copula parameter is time-dependent, we consider ordered sets $U_t = (u_{1t}, ..., u_{dt})$ – the ordering being determined by the simple ordering of the data in the original set.

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

$$Y = F_{\chi_d^2}\left(\sum_{i=1}^{d} \Phi^{-1}(u_i')^2\right)$$

is uniformly distributed on $[0, 1]$, where $F_{\chi_d^2}$ 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 Cramér-von Mises criterion.

5 Numerical results

In this section we apply the proposed continuous SCAR model with the transfer matrix likelihood evaluation to real cryptocurrency market data. We consider two datasets of different lengths to demonstrate the model's performance across different regimes and to illustrate the advantages of the transfer matrix approach over Monte Carlo estimation.

5.1 Data description

We use two datasets of BTC-USD and ETH-USD returns.

Dataset 1 (daily, $T = 1460$). Daily log-returns of BTC-USD and ETH-USD from January 2, 2020 to December 31, 2023, computed as $r_t = \log(P_t/P_{t-1})$. This dataset represents a moderate-length time series typical in financial applications.

Dataset 2 (30-minute, $T = 12{,}000$). Intraday 30-minute log-returns of BTC and ETH from January 1, 2020 to September 7, 2020. This high-frequency dataset tests the scalability of the transfer matrix method and highlights the regime where sparse matrices and adaptive grid refinement become essential.

In both cases, log-returns are transformed to pseudo-observations via the rank transformation (2.2).

5.2 Model comparison: MLE vs. Stochastic model

We fit the Gumbel copula (rotated by 180°) using two approaches: the classical MLE with a constant parameter and the proposed stochastic model with the Ornstein-Uhlenbeck latent process estimated via Transfer matrix approach. The quality of each model is assessed via the log-likelihood $\log L$ and the Cramér–von Mises goodness-of-fit test described in Section 4.

Dataset 1: daily returns ($T = 1460$). The results are presented in Table 3. The static MLE model yields a copula parameter $\hat{\theta} = 2.832$ and a GoF p-value of 0.010, indicating that the constant-parameter model is rejected at the 5% significance level. The SCAR-TM-OU model achieves a substantially higher log-likelihood (1045.5 vs. 955.6) and passes the GoF test with a p-value of 0.703, confirming that the time-varying dependence structure is a significantly better fit.

Table 3: Model comparison on daily BTC–ETH returns ($T = 1460$, Gumbel-180 copula)

Model	$\log L$	Parameters	GoF p-value
MLE (constant θ)	955.6	$\hat{\theta} = 2.832$	0.010
SCAR-TM-OU	1045.5	$(\hat{\theta}_{\mathrm{OU}}, \hat{\mu}, \hat{\nu}) = (58.96,\ 1.488,\ 4.531)$	0.703

Dataset 2: 30-minute returns ($T = 12{,}000$). The results for the high-frequency dataset are shown in Table 4. Here the discrepancy between the static and dynamic models is even more pronounced. The constant-parameter model yields a GoF p-value that is essentially zero, while the SCAR-TM-OU model achieves a p-value of 0.280. The large value of the estimated mean-reversion rate $\hat{\theta}_{\mathrm{OU}} = 93.74$ reflects the rapid dynamics of intraday dependence.

Table 4: Model comparison on 30-minute BTC–ETH returns ($T = 12{,}000$, Gumbel-180 copula)

Model	$\log L$	Parameters	GoF p-value
MLE (constant θ)	4525.5	$\hat{\theta} = 2.041$	< 0.001
SCAR-TM-OU	5075.0	$(\hat{\theta}_{\mathrm{OU}}, \hat{\mu}, \hat{\nu}) = (93.74,\ 1.099,\ 4.272)$	0.280

For $T = 12{,}000$ and $\Delta t \approx 8.3 \times 10^{-5}$, the conditional standard deviation σ_c is very small relative to the stationary σ, so the adaptive grid rule (3.13) increases K well above the default value and the sparse transfer matrix representation is selected automatically. This makes the computation tractable even for this large dataset.

5.3 Jensen bias in Monte Carlo estimation

To demonstrate the bias inherent in the Monte Carlo p-sampler (2.5), we evaluate the minus log-likelihood at the optimal parameters found by SCAR-TM-OU for Dataset 1, using $N_{\mathrm{tr}} \in \{100, 500, 1000, 10000, 100000\}$ trajectories, with 50 independent replications for each N_{tr}. The results are summarized in Table 5.

Table 5: Monte Carlo p-sampler: Jensen bias and variance ($T = 1460$, true $-\log L = 1045.5$)

N_{tr}	$-\log L$ (mean over 50 runs)	std
100	909.4	14.0
500	930.3	12.3
1,000	935.7	11.0
10,000	954.8	10.2
100,000	968.4	7.8

The Monte Carlo estimator significantly underestimates the log-likelihood: even with $N_{\mathrm{tr}} = 100000$, the mean estimate is -968.4 compared to the transfer matrix value of -1045.5, a bias of approximately 77 nats. The bias decreases slowly with N_{tr} (as $O(1/N)$ per Jensen's inequality), and the standard deviation remains substantial. This confirms that for time series of length $T \sim 10^3$, Monte Carlo-based SCAR estimation is unreliable without an impractically large number of trajectories, and the transfer matrix method provides a decisive advantage.

6   Conclusion

In this paper, we propose a stochastic copula model in continuous time for modeling time-varying dependence structures in financial data. Unlike existing SCAR models that rely on Monte Carlo integration, we develop a transfer matrix method that exploits the Markov structure and known Gaussian transition density of the Ornstein-Uhlenbeck process to evaluate the likelihood function deterministically on a grid. This approach reduces the original high-dimensional integral to a sequence of one-dimensional integrals, each computed via matrix-vector multiplication with trapezoidal quadrature. The resulting scheme has computational complexity $O(TK^2)$ and yields a numerically exact likelihood evaluation, free from the variance and downward bias inherent in the log-mean-exp Monte Carlo estimator.

We discuss implementation details that are essential for practical applicability: adaptive grid refinement that guarantees adequate resolution of the transition kernel regardless of the OU parameters, and the use of sparse transfer matrices that reduce the per-step cost from $O(K^2)$ to $O(Kb)$ when the kernel bandwidth b is small relative to K. These techniques make the method scalable to high-frequency datasets with T on the order of 10^4.

The proposed model was validated on cryptocurrency market data (BTC–ETH pairs) at two temporal resolutions: daily ($T = 1,460$) and 30-minute ($T = 12,000$). In both cases, the stochastic copula model showed a significant improvement in terms of the log-likelihood $\log L$ and p-values of the corresponding goodness-of-fit test compared to the classical model with constant parameter. The Monte Carlo p-sampler was shown to exhibit substantial downward bias even with 10^5 trajectories, confirming the practical necessity of the deterministic transfer matrix approach.

As possible future research directions, we note investigating alternative stochastic processes (fractional Brownian motion or Lévy processes), extending the model to higher-dimensional copulas via vine constructions, and applying the proposed framework to portfolio risk assessment.

References

[ACM16]   Carlos Almeida, Claudia Czado, and Hans Manner. Modeling high-dimensional time-varying dependence using dynamic d-vine models. *Applied Stochastic Models in Business and Industry*, 32(5):621–638, 2016.

[CKL13]   D. Creal, S. Koopman, and A. Lucas. Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777–795, 2013.

[Fel91]   William Feller. *An introduction to probability theory and its applications, Volume 2*, volume 81. John Wiley & Sons, 1991.

[Har10]   A. Harvey. Tracking a changing copula. *Journal of Empirical Finance*, 17:485–500, 2010.

[HH15]   Christian Hering and Marius Hofert. Goodness-of-fit tests for archimedean copulas in high dimensions. In Kathrin Glau, Matthias Scherer, and Rudi Zagst, editors, *Innovations in Quantitative Risk Management*, pages 357–373, Cham, 2015. Springer International Publishing.

[HM12]   C. M. Hafner and H. Manner. Dynamic stochastic copula models: estimation, inference and applications. *Journal of Applied Econometrics*, 27(2):269–295, 2012.

[Hof08]   Marius Hofert. Sampling archimedean copulas. *Computational Statistics and Data Analysis*, 52(12):5163–5174, 2008.

[Joe97]   H. Joe. *Multivariate Models and Multivariate Dependence Concepts*. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis, 1997.

[KHKK23]   Yunsun Kim, Sun-Young Hwang, Jong-Min Kim, and Sahm Kim. 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.

[Kim74]   Clark H Kimberling. A probabilistic interpretation of complete monotonicity. *Aequationes mathematicae*, 10:152–164, 1974.

[KMM22]   Janis Kielmann, Hans Manner, and Aleksey Min. Stock market returns and oil price shocks: A covar analysis based on dynamic vine copula models. *Empirical Economics*, 62:1543–1574, 2022.

[LR03]   Roman Liesenfeld and Jean-Francois Richard. Univariate and multivariate stochastic volatility models: Estimation and diagnostics. *Journal of Empirical Finance*, 10:505–531, 02 2003.

[Mba24]   J.C. Mba. Assessing portfolio vulnerability to systemic risk: a vine copula and APARCH-DCC approach. *Financ Innov*, 10:20, 2024.

[MO88]   Albert W Marshall and Ingram Olkin. Families of multivariate distributions. *Journal of the American statistical association*, 83(403):834–841, 1988.

[MR12]   Hans Manner and Olga Reznikova. A survey on time-varying copulas: Specification, simulations, and application. *Econometric Reviews*, 31(6):654–687, 2012.

[Nel07]   R.B. Nelsen. *An Introduction to Copulas*. Springer Series in Statistics. Springer New York, 2007.

[NV23]   Hoang Nguyen and Audronė Virbickaitė. Modeling stock-oil co-dependence with dynamic stochastic midas copula models. *Energy Economics*, 124:106738, 2023.

[Skl59]   M Sklar. Fonctions de repartition an dimensions et leurs marges. *Publ. inst. statist. univ. Paris*, 8:229–231, 1959.

[SS19]   M. Semenov and D. Smagulov. Copula Models Comparison for Portfolio Risk Assessment. In M. Kaz, T. Ilina, and G. Medvedev, editors, *Global Economics and Management: Transition to Economy 4.0*, Springer Proceedings in Business and Economics, pages 91–102. Springer, Cham, 2019.

[vdGGW05] R. W. J. van den Goorbergh, C. Genest, and B. J. M. Werker. Bivariate option pricing using dynamic copula models. *Insurance: Mathematics and Economics*, 37:101–114, 2005.