The realized copula of volatility*

Kim Christensen[†‡] Wenjing Liu[§] Zhi Liu[¶] Yoann Potiron[‖]

April, 2026

Abstract

We study a new measure of codependency in the second moment of a continuous-time multivariate asset price process, which we name the realized copula of volatility. The statistic is based on local volatility estimates constructed from high-frequency asset returns and affords a nonparametric estimator of the empirical copula of the latent stochastic volatility. We show consistency of our estimator with in-fill asymptotic theory, either with a fixed or increasing time span. In the latter setting, we derive a functional central limit theorem for the empirical process associated with the measurement error of the time-invariant marginal copula of volatility. We also develop a goodness-of-fit test to evaluate hypotheses about the shape of the latter. In a simulation study, we demonstrate that our estimator is a good proxy of both the empirical and marginal copula of volatility, even with a moderate amount of high-frequency data recorded over a relatively short sample. The goodness-of-fit test is found to exhibit size control and excellent power. We implement our framework on high-frequency transaction data from futures contracts that track the U.S. equity and treasury bond market. A Gumbel copula is found to offer a near-perfect bind between the realized variance processes in these data.

JEL Classification: C14; C32; C58; C80.

Keywords: Copula; empirical process; functional central limit theorem; high-frequency data; nonparametric estimation; stochastic volatility; tail dependence.

*Christensen appreciates funding from the Independent Research Fund Denmark to support this work (DFF 1028–00030B). Zhi Liu's research is supported by MYRG-GRG2024-00190-FST-UMDF, MYRG-GRG2025-00093-FST, and APAEM Seed Grant in Financial Econometrics from the University of Macau. Potiron acknowledges financial support from the Japanese Society for the Promotion of Science (23H00807).

[†]Department of Economics and Business Economics, Aarhus University, Denmark. E-mail: kim@econ.au.dk

[‡]Research fellow at the Danish Finance Institute (DFI).

[§]Department of Mathematics, University of Macau, Macau SAR, China. E-mail: yc27481@um.edu.mo

[¶]Department of Mathematics, University of Macau, Macau SAR, China. E-mail: liuzhi@um.edu.mo

[‖]Faculty of Business and Commerce, Keio University, Japan. E-mail: potiron@fbc.keio.ac.jp

1 Introduction

Modeling the comovement of stochastic volatility across several assets is paramount to risk measurement and management, portfolio allocation, and multi-asset and derivative pricing. It has long been recognized that financial markets are interconnected, causing asset return volatilities—and correlations—to surge in tandem during financial distress, triggering significant volatility spillover effects, posing a potential systemic risk (see, e.g., Diebold and Yilmaz, 2009; Billio, Getmansky, Lo, and Pelizzon, 2012, and references therein). An understanding of volatility codependency is also important for pricing of volatility risk (i.e., the variance risk premium) across asset classes (Carr and Wu, 2009; Bollerslev and Todorov, 2011). The importance of the second moment structure of asset returns is further fueled by the so-called leverage effect that captures a negative association between return and volatility (Black, 1976; Christie, 1982). Thus, from a risk aggregation view the relevant target is often the dependence structure in second moment of returns (namely volatilities and correlations), rather than in the raw returns.

In practice, the assessment of volatility comovements is complicated by the fact that volatility is unobserved, so it requires an empirical route based on an observable proxy. It therefore remains a relatively unexplored area. In recent years, however, the availability of high-frequency data has made it possible to estimate volatility nonparametrically with the realized variance (e.g., Andersen and Bollerslev, 1998; Barndorff-Nielsen and Shephard, 2002). Most of this literature concentrates on estimation of integrated measures of variation, defined as an integral of a smooth function of the latent spot variance process (see Jacod and Rosenbaum, 2013). However, as Li, Todorov, and Tauchen (2013) point out, the mapping from the probability distribution of spot volatility to the distribution of integrated volatility is not one-to-one. Thus, the "integrated" approach may be partially informative about some distributional aspects of volatility and codependency, but it faces a limitation due to the inherent "identification failure." This may weaken, or outright eliminate, power for testing against relevant alternatives. They suggest to work directly with the pathwise distribution of spot volatility via a so-called occupation measure (e.g., Geman and Horowitz, 1980), which recovers the empirical distribution function of the point-in-time volatility process (see also Li, Todorov, and Tauchen, 2016a,b).[1] Regardless, the volatility occupation literature is entirely univariate and has concentrated on the marginal distribution, rather than looking at the interrelationship between various volatility processes.

The starting point of our analysis is the remarkable result from probability theory known as Sklar's theorem (Sklar, 1959). It states that for any continuous random vector $(X_1, \ldots, X_d)$ with cumulative distribution function $H(x_1, \ldots, x_d) = \mathbb{P}(X_1 \leq x_1, \ldots, X_d \leq x_d)$, there exists a unique function $C : [0,1]^d \to [0,1]$ called a *copula*, such that $H(x_1, \ldots, x_d) = C(H_1(x_1), \ldots, H_d(x_d))$, where $H_i(x_i) = \mathbb{P}(X_i \leq x_i)$ is the marginal distribution of X_i.[2] Conversely, given $C(u_1, \ldots, u_d)$ and

[1] In addition, Christensen, Thyrsgaard, and Veliyev (2019) and Zhang, Li, and Bollerslev (2022) made the univariate framework robust to microstructure noise.

[2] The concise mathematical definition is as follows: A d-dimensional copula $C : [0,1]^d \to [0,1]$ is the cumulative distribution function of a random vector $(U_1, \ldots, U_d)$, where the marginal distribution of U_i is uniform, i.e.

$H_1(x_1), \ldots, H_d(x_d)$, $H(x_1, \ldots, x_d) = C(H_1(x_1), \ldots, H_d(x_d))$ is a probability distribution. Copulas provide a marginal-free representation of dependence in a multivariate setting, and they enable the researcher to model the marginal features of the data separately.[3] Indeed, many measures of concordance (and discordance), such as Kendall's tau and Spearman's rho, can be expressed entirely in terms of the copula (see Section 6). The latter are often more informative than the Pearson's coefficient of correlation, which captures *linear* dependence. As emphasized by Embrects, McNeil, and Straumann (1999), correlation is insufficient for risk management. This is explained by the fact that financial markets typically exhibit *nonlinear* dependence and, further, may possess a high degree of tail concentration (or corner clustering).[4] Therefore, copulas—or descriptive statistics derived from them—are a better representation of dependence. Another advantage of rank-based statistics is their invariance to monotonic transformations (that do not alter the ordering of the data). However, a model-free technique for analyzing the multivariate dependence structure of volatility is still lacking. In particular, tools for estimating and conducting inference on the copula of volatility are not available.

We close that gap and propose to study the dependence structure of several stochastic volatility processes. We make a four-fold contribution to the literature. First, we extend Li, Todorov, and Tauchen (2013) to the multivariate setting. We operate within a general, albeit standard, continuous-time arbitrage-free Itô semimartingale framework for a (multi-dimensional) log-price process, where the random volatility can exhibit rather unrestricted dynamics. We replace the latent spot variance processes by localized, jump-robust, realized variance measures constructed from discretely observed high-frequency returns over short time intervals. The estimation errors are controllable by the sampling interval and bandwidth selection (e.g., Jacod and Protter, 2012). We allow for very active jumps in the price process. The latter can be regarded as "nuisance parameters" in our setting, and we handle them with a truncation device (e.g., Mancini, 2009). Second, we construct a nonparametric estimator of volatility codependency, called the *realized copula of volatility*, enabling a meaningful cross-asset comparison irrespective of the marginal distribution of the individual volatility processes. Third, we develop the necessary large-sample theory for our estimator aligned with its target at each horizon. In the in-fill asymptotic limit with fixed time span, we show the realized copula of volatility converges, uniformly, in probability to its empirical counterpart that captures the volatility dependency over the time interval observed so far. Importantly, we allow for nonstationary volatility. In a double asymptotic in-fill and long-span setting—with stationary and weakly dependent volatility—we establish a functional central limit theorem in the

$C(u_1, \ldots, u_d) = \mathbb{P}(U_1 \leq u_1, \ldots, U_d \leq u_d)$ with $\mathbb{P}(U_i \leq u_i) = u_i$, for $i = 1, \ldots, d$ and $0 \leq u_i \leq 1$.

[3]The term copula originates from the Latin word for "bond" or "tie".

[4]The application of copula theory had a revival in financial economics after Li (2000) proposed to model the default correlation in portfolios of bond loans with a Gaussian copula. This enabled tractable analysis of structured credit products (such as collateralized debt obligations). The model became a standard pricing tool on Wall Street. However, since the Gaussian copula has no tail dependence, it tends to understate the probability of concurrent defaults, as seen during the financial crisis (see, for instance, Salmon, 2009, for a critique of the approach). A comprehensive review of the application of copula theory to economic and financial time series is Patton (2012).

form of weak convergence of the measurement error of the realized volatility copula. We show how this can be exploited to construct pointwise or, with a bit more effort, uniform confidence bands. A goodness-of-fit test for assessing the shape of the volatility copula is also proposed. The latter procedure can, for instance, be used to gauge the adequacy of various parametric copulas; an approach we look further at in our application. This builds upon and extends earlier results on the empirical copula process (Deheuvels, 1979; Fermanian, Radulović, and Wegkamp, 2004). However, as explained above, our framework is distinctly more complicated than the discrete-time setup with independent and identically distributed data (see Fermanian and Scaillet (2002) for related work in the time series setting). Fourth, we deliver an implementable inference procedure by designing a Newey and West (1994)-type heteroscedasticity and autocorrelation consistent estimator of the long-run asymptotic variance of our realized copula of volatility.

The small sample properties of our estimator are investigated through a Monte Carlo experiment. We demonstrate its efficacy over the entire support of the copula. In addition, we evaluate the goodness-of-fit test in order to illuminate its size and power. In our empirical work, we consider an application to U.S. equity and treasury futures markets. This leads to forceful evidence that a Gumbel copula—exhibiting upper-tail dependence—offers a near-perfect description of the realized copula of volatility. Thus, our nonparametric procedure complements the existing literature and facilitates a marginal-free analysis of volatility codependency, offering a more refined assessment of systemic risk transmission.

The paper progresses as follows. In Section 2, we present the theoretical framework and define the empirical copula of the volatility process based on an occupation measure. In Section 3, we adopt a spot volatility estimator based on discrete high-frequency data, from which we can construct the realized copula of volatility. In Section 4, we state our assumptions and develop large-sample theory for our estimator. In particular, we show its consistency for the empirical copula of volatility in the in-fill limit with fixed time span. We also derive a functional central limit theorem for the empirical process associated with the estimation error of the time-invariant copula of volatility in a double asymptotic in-fill and long-span setting. We further propose an estimator of the long-run asymptotic variance and show how to exploit the functional analysis to construct uniform confidence bands. At last, we develop a goodness-of-fit test. Section 5 reports a detailed Monte Carlo study to evaluate the finite sample properties of our realized copula of volatility, along with a parametric version of our goodness-of-fit test. In Section 6, we present an empirical application. We conclude in Section 7. The proofs of our statistical analysis are postponed to the Appendix.

2 The setting

We suppose that a filtered probability space (or stochastic basis) $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P}\right)$ describes the evolution of a pair of log-price processes, X and Y, that evolve in continuous-time over an interval

$[0,T]$.[5] Here, $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $s \leq t \leq T$ is a filtration that represents past and current information available to market participants at any time.

As consistent with the "no free lunch with vanishing risk" principle from financial economics, we assume that the motion of (X, Y) is described by an Itô semimartingale (Delbaen and Schachermayer, 1994).[6] This means that it has the component-wise representation

$$\mathrm{d}Z_t = b_t^Z \mathrm{d}t + \sigma_t^Z \mathrm{d}W_t^Z + \int_{|\delta^Z(s,z)|\leq 1} \delta^Z(t,z)(\mu^Z - \nu^Z)(\mathrm{d}z, \mathrm{d}t) + \int_{|\delta^Z(s,z)|>1} \delta^Z(t,z)\mu^Z(\mathrm{d}z, \mathrm{d}t), \quad (2.1)$$

where the drift $(b_t^Z)_{t\geq 0}$ and the volatility $(\sigma_t^Z)_{t\geq 0}$ are adapted càdlàg processes, while $(W_t^Z)_{t\geq 0}$ is a standard Brownian motion. In addition, μ^Z is a Poisson random measure on $(\mathbb{R}_+, \mathbb{R})$ with compensator $\nu^Z(\mathrm{d}t, \mathrm{d}z) = dt \otimes \lambda^Z(\mathrm{d}z)$, λ^Z is a σ-finite measure on $\mathbb{R}$, and $\delta^Z : \Omega \times \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R}$ is predictable. Here, and in other places, we employ a generic process Z (typically with $Z = X, Y$) to avoid repeating definitions.

In general, a semimartingale can be decomposed into a finite variation component and a local martingale. The "Itô" classifier imposes an absolute continuity condition on these, so that we can express them as integrals of "spot" processes. This restriction is common, because the latter are more amenable to statistical analysis from high-frequency data. Apart from that, our framework is nonparametric, model-free, and suffices to capture the most dominant features of empirical asset price processes, such as time-varying expected returns, stochastic volatility, and leverage effect. Moreover, it encompasses the presence of large and small price jumps, where the latter can be infinitely active and of infinite variation. We also permit complex within- and cross-asset dependencies between the various model components. For example, the standard Brownian motions can be correlated (i.e. $\rho_t \mathrm{d}t = \mathrm{d}\langle W^X, W^Y\rangle_t$, where $\langle W^X, W^Y\rangle_t$ is the quadratic covariation process). Moreover, the jumps in volatility can be related to the occurrence of jumps in the log-price, and there can be common jumps in both (for instance Todorov and Tauchen, 2011; Bibinger and Winkelmann, 2018). We are, however, going to impose additional structure and smoothness conditions for our econometric procedure.

Our aim is to study the empirical copula of volatility, which begins with a notion of the empirical distribution function of volatility, defined as

$$H_T(x,y) = \frac{1}{T}\int_0^T \mathbb{1}_{\{V_t^X \leq x, V_t^Y \leq y\}} \mathrm{d}t, \quad (2.2)$$

where $0 < x, y < \infty$, while $V_t^X = (\sigma_t^X)^2$ and $V_t^Y = (\sigma_t^Y)^2$ are point-in-time variances of X and Y, and $\mathbb{1}_{\{\cdot\}}$ is the indicator function. $H_T(x,y)$ measures the relative amount of time in $[0,T]$ that the stochastic volatility processes spent at different levels over their support. Hence, (2.2) is the pathwise counterpart of the cumulative distribution function of volatility. Without the T

[5] Our theoretical framework extends trivially to d-dimensional processes, for any fixed $d \geq 2$. We concentrate on the bivariate setting, because it conveys the main idea and avoids the cost of extra notation.

[6] The Itô semimartingale assumption is commonly motivated in economics by the efficient markets hypothesis and rational expectations theory (see, e.g., Samuelson, 1965).

normalization, $H_T(x, y)$ is a multivariate extension of the volatility occupation measure studied in Li, Todorov, and Tauchen (2013), which corresponds to the marginal empirical distribution function $F_T(x) = H_T(x, \infty)$ and $G_T(y) = H_T(\infty, y)$.

The empirical copula of volatility is motivated by Sklar's theorem. In particular, since $H_T(x, y)$ in (2.2) is the pathwise analogue of the distribution function of the latent variance process (V_t^X, V_t^Y) over $[0, T]$, it is natural to separate its marginal features from its dependence structure. We therefore define the empirical copula of volatility as

$$C_T(u, v) = \frac{1}{T} \int_0^T \mathbb{1}_{\{F_T(V_t^X) \leq u, G_T(V_t^Y) \leq v\}} \mathrm{d}t, \tag{2.3}$$

where $0 < u, v < 1$.

If $F_T(x)$ and $G_T(y)$ are invertible (e.g., continuous and strictly increasing), (2.3) can be written as $C_T(u, v) = H_T(F_T^{-1}(u), G_T^{-1}(v))$, which leads to the "inversion formula" $H_T(x, y) = C_T(F_T(x), G_T(y))$. Thus, the empirical distribution function of volatility can be expressed as the marginal empirical distribution of the individual volatility processes and a copula that captures their dependence structure up to T. $C_T(u, v)$ delivers a margin-free characterization of volatility codependency that is invariant to strictly monotone transformations of the marginal volatility processes. This is useful in financial applications, where the marginal behavior of volatility can differ markedly across assets, while their dependence structure can be decided separately. In addition, and in contrast to Pearson's measure of linear correlation, the copula can accommodate nonlinear dependence and tail concentration.

Without prior knowledge of the volatility processes, however, we define the empirical quantile function of $F_T(x)$ and $G_T(y)$ as the generalized inverse:

$$Q_T^X(u) \equiv F_T^{-1}(u) = \inf_{x \geq 0}\{x : F_T(x) \geq u\} \quad \text{and} \quad Q_T^Y(v) \equiv G_T^{-1}(v) = \inf_{y \geq 0}\{y : G_T(y) \geq v\} \tag{2.4}$$

with the convention $\inf \emptyset = \infty$.[7] In the general setting, we still set $C_T(u, v) = H_T(Q_T^X(u), Q_T^Y(v))$, but it should now be given an infimum-based interpretation. Hence, $C_T(u, v)$ measures the proportion of time that the volatility processes spend below the smallest coordinate x and y for which $F_T(x)$ and $G_T(y)$ exceed u and v.

To the best of our knowledge, both the multivariate extension of the empirical distribution function of volatility, and the associated copula, are novel to the literature.

3 High-frequency estimation

In practice, V^X and V^Y are not directly observed, so neither is the empirical distribution function of volatility nor the copula transformation. Our empirical strategy is based on the premise that we can approximate the spot variance processes with a realized measure constructed from discretely

[7] The generalized inverse of a distribution function is non-decreasing, left-continuous, and admits a limit from the right. If the distribution function is discontinuous, the quantile function has flat spots, and vice versa.

observed high-frequency data of X and Y over small time intervals. In the asymptotic theory, we are going to both increase the number of blocks and shrink the timespan of each block, while padding it with an increasing number of observations, such that we attain a better and better assessment of the volatility's path. In doing so, we control for the price jump component. We also deal with the drift, but this is easier since it is asymptotically negligible in our setting.[8]

We assume that (X, Y) are sampled at equidistant time points $(i\Delta_n)_{i=0}^n$, where Δ_n is the time gap and $n = \lfloor T/\Delta_n \rfloor$ denotes the number of observations over $[0, T]$. We define an increment of the process Z between $(i-1)\Delta_n$ and $i\Delta_n$ as $\Delta_i^n Z = Z_{i\Delta_n} - Z_{(i-1)\Delta_n}$, for $i = 1, \ldots, n$. Then, as in Li, Todorov, and Tauchen (2013), we adopt the following estimator of V_t^Z (at time t):

$$\hat{V}_t^Z = \begin{cases} \frac{1}{k_n\Delta_n} \sum_{j=1}^{k_n} \left(\Delta^n_{k_n\lfloor \frac{t}{k_n\Delta_n}\rfloor + j} Z\right)^2 \mathbb{1}_{\{|\Delta^n_{k_n\lfloor \frac{t}{k_n\Delta_n}\rfloor + j} Z| \leq \alpha\Delta_n^{\varpi}\}}, & 0 \leq t < \lfloor \frac{T}{k_n\Delta_n}\rfloor k_n\Delta_n, \\ \frac{1}{k_n\Delta_n} \sum_{j=1}^{k_n} \left(\Delta^n_{n-j+1} Z\right)^2 \mathbb{1}_{\left\{|\Delta^n_{n-j+1} Z| \leq \alpha\Delta_n^{\varpi}\right\}}, & \lfloor \frac{T}{k_n\Delta_n}\rfloor k_n\Delta_n \leq t \leq T. \end{cases} \tag{3.1}$$

$\hat{V}_t^Z$ is the spot realized variance (e.g., Jacod and Protter, 2012). k_n is the number of log-price increments included in the estimation on each subinterval and tends to infinity in the asymptotic analysis, whereas $\alpha > 0$ and $\varpi \in (0, 1/2)$ relate to the jump truncation device of Mancini (2009). The threshold $\alpha\Delta_n^{\varpi}$ is designed to remove log-price increments perturbed by the price jump component and is asymptotically shrinking, slowly enough, to zero. The rate conditions of the tuning parameters are made explicit below.[9]

We define the realized distribution function of volatility, an estimator of the empirical distribution function of volatility, as follows:

$$\widehat{H}_{n,T}(x, y) = \frac{1}{T} \int_0^T \mathbb{1}_{\left\{\hat{V}_t^X \leq x, \hat{V}_t^Y \leq y\right\}} \mathrm{d}t. \tag{3.2}$$

The univariate version can be retrieved as $\widehat{F}_{n,T}(x) = \widehat{H}_{n,T}(x, \infty) = \frac{1}{T}\int_0^T \mathbb{1}_{\left\{\hat{V}_t^X \leq x\right\}} \mathrm{d}t$ and $\widehat{G}_{n,T}(y) = \widehat{H}_{n,T}(\infty, y) = \frac{1}{T}\int_0^T \mathbb{1}_{\left\{\hat{V}_t^Y \leq y\right\}} \mathrm{d}t$. The latter are again equivalent to those proposed in Li, Todorov, and Tauchen (2013).

Then, the realized copula of volatility is the statistic:

$$\widehat{C}_{n,T}(u, v) = \frac{1}{T} \int_0^T \mathbb{1}_{\left\{\widehat{F}_{n,T}(\hat{V}_t^X) \leq u, \widehat{G}_{n,T}(\hat{V}_t^Y) \leq v\right\}} \mathrm{d}t, \tag{3.3}$$

and the realized quantile functions of volatility are extended as follows:

$$\widehat{Q}^X_{n,T}(u) = \inf_{x \geq 0}\{x : \widehat{F}_{n,T}(x) \geq u\} \qquad \text{and} \qquad \widehat{Q}^Y_{n,T}(v) = \inf_{y \geq 0}\{y : \widehat{G}_{n,T}(y) \geq v\}, \tag{3.4}$$

such that $\widehat{C}_{n,T}(u, v) = \widehat{H}_{n,T}(\widehat{Q}^X_{n,T}(u), \widehat{Q}^Y_{n,T}(v))$.

[8] Christensen, Oomen, and Renò (2022) propose a drift burst model, in which the drift term is locally of larger order than the volatility component.

[9] Note that in much of the following, we suppress the explicit dependence of various statistics on n, when it can be avoided without causing confusion.

4 Asymptotic theory

To derive the asymptotic theory for the realized distribution function of volatility and realized copula of volatility, we prove a consistency result in distinct settings with i) $\Delta_n \to 0$ and T fixed, and ii) $\Delta_n \to 0$ and $T \to \infty$. In the first, the realized statistic is a natural estimator of the empirical counterpart. In the second, the realized statistic converges to the invariant distribution, assuming it exists. Under further smoothness conditions and weak dependence of the volatility process, we establish a central limit theorem. The latter is done in the functional sense of weak convergence of the probability measures associated with the empirical process of the normalized measurement error of volatility.

To begin with, we introduce a number of additional conditions. The first assumption concerns the regularity of the price jump component.

ASSUMPTION **1.** *The process Z follows (2.1) with b^Z locally bounded and σ^Z càdlàg. Moreover, for a constant $r \in [0,2]$, and a sequence of stopping times $(\tau_m)_{m=1}^{\infty}$, such that $\tau_m \to \infty$ as $m \to \infty$, there exists a sequence of functions $(\Gamma_m)_{m=1}^{\infty}$ such that for each $m : \min(|\delta^Z(\omega,t,z)|,1) \leq \Gamma_m(z)$ for (ω,t,z) with $t \leq \tau_m(\omega)$, and $\int_{\mathbb{R}} \Gamma_m(z)^r \lambda^Z(\mathrm{d}z) < \infty$.*

The exponent r in Assumption 1 is an upper bound on the generalized Blumenthal-Getoor index, an adaptation from Lévy processes to the semimartingale setting, see Jacod and Protter (2012, Lemma 3.2.1).[10] It is therefore related to the jump activity of the log-price process with lower values of r being more binding. The point $r = 1$ separates jump processes with sample paths of finite and infinite variation, i.e., it concerns their absolute summability, while $r = 0$ implies that the jumps are finitely active (on finite time intervals, almost surely). r plays a crucial role in determining rate conditions for the tuning parameters in our estimation procedure, and also in establishing the rate of convergence of our estimator.

ASSUMPTION **2.** *$H_T(x,y)$ is continuous, almost surely.*

The continuity of the empirical distribution function imposed by Assumption 2 is used to control the approximation error in the uniform metric.

THEOREM 1. *We suppose that Assumption 1 with $r = 2$ and Assumption 2 hold. As $\Delta_n \to 0$ and $k_n \to \infty$, such that $k_n \Delta_n \to 0$, and with T fixed, it follows that*

$$\sup_{(x,y)\in\mathbb{R}_+^2} |\widehat{H}_{n,T}(x,y) - H_T(x,y)| \xrightarrow{\mathbb{P}} 0. \tag{4.1}$$

Furthermore, let $\mathcal{Q} = \{(u,v) \in [0,1]^2 : Q_T^X(u)$ is continuous at u and $Q_T^Y(v)$ is continuous at v, almost surely$\}$. Then, for each $(u,v) \in \mathcal{Q}$, it further holds that

$$\widehat{C}_{n,T}(u,v) \xrightarrow{\mathbb{P}} C_T(u,v). \tag{4.2}$$

[10] The Blumenthal-Getoor index of a Lévy process is defined as $\beta = \inf\left\{r \geq 0 : \int_{|x|<1} |x|^r \nu(\mathrm{d}x) < \infty\right\}$, where ν is the Lévy measure. β can be interpreted as the smallest power at which the distribution of the small jumps (arbitrarily chosen to be of size $|x| < 1$) has finite rth moment.

Moreover, if $Q_T^X(u)$ and $Q_T^Y(v)$ are continuous, almost surely, on $[0,1]^2$:

$$\sup_{(u,v)\in[0,1]^2} |\widehat{C}_{n,T}(u,v) - C_T(u,v)| \xrightarrow{\mathbb{P}} 0. \tag{4.3}$$

Theorem 1 establishes uniform convergence in probability of $\widehat{H}_{n,T}(x,y)$ as an estimator of $H_T(x,y)$. As $\Delta_n \to 0$ and $k_n \to \infty$, such that $\Delta_n k_n \to 0$, we get an increasing number of error-free estimates of the spot variance, forming a dense subset on $[0,T]$, from which we can construct an arbitrarily accurate estimate of $H_T(x,y)$. The exact rate of k_n does not influence this result, so long as $\Delta_n k_n \to 0$. We can also show pointwise convergence in probability of $\hat{C}_{n,T}(x,y)$ on the continuity set $\mathcal{Q}$. However, in the discontinuous case at a point (u,v) where either $Q_T^X(u)$ or $Q_T^Y(v)$ are not continuous, even uniform convergence of $\widehat{H}_{n,T}(x,y)$ does not imply that $Q_{n,T}^X(u) \xrightarrow{\mathbb{P}} Q_T^X(u)$ (or $Q_{n,T}^Y(u) \xrightarrow{\mathbb{P}} Q_T^Y(v)$), so $\widehat{C}_{n,T}(u,v)$ is generally inconsistent for $\widehat{C}_T(u,v)$ at that point. We can strengthen the convergence of $\widehat{C}_{n,T}(x,y)$ to a uniform statement if the limiting empirical quantile function is continuous, almost surely.

The smoothness imposed on $H_T(x,y)$ does not require the volatility path itself to be continuous. Even if V_t^X or V_t^Y exhibit jumps, $H_T(x,y)$ can be continuous provided the sample path variation is rich enough to fill out the gaps. The intuition is that $H_T(x,y)$ aggregates the amount of time that the processes spend below a given threshold, and it can thus "iron out" local irregularities through temporal aggregation.

To establish the asymptotic theory for the combined in-fill and long-span setting, we need stricter control of the error embedded in the recovery of the volatility path. First, we need to restrict the càdlàg dynamic of the volatility processes.

Assumption **3.** *The volatility process σ^Z is of the form:*

$$\sigma_t^Z = \sigma_0^Z + \int_0^t \tilde{b}_s^Z \mathrm{d}s + \int_0^t \tilde{\sigma}_s^Z \mathrm{d}\tilde{W}_s^Z + \int_0^t \int_{\mathbb{R}} \tilde{\delta}^Z(s,z)(\tilde{\mu}^Z - \tilde{\nu}^Z)(\mathrm{d}s,\mathrm{d}z). \tag{4.4}$$

Here, $\tilde{b}^Z$ and $\tilde{\sigma}^Z$ are adapted and locally bounded processes, $\tilde{W}^Z$ is a standard Brownian motion (that may depend on W^Z), while $\tilde{\delta}^Z$ is a predictable function. Moreover, for a constant $\tilde{r} \in [0,2]$, and a sequence of stopping times $(\tilde{\tau}_m)_{m=1}^\infty$ such that $\tilde{\tau}_m \to \infty$ as $m \to \infty$, there exists a sequence of functions $(\tilde{\Gamma}_m)_{m=1}^\infty : \min\{|\tilde{\delta}^Z(\omega,t,z)|,1\} \leq \tilde{\Gamma}_m(z)$ for (ω,t,z) with $t \leq \tilde{\tau}_m(\omega)$, and $\int_{\mathbb{R}} \tilde{\Gamma}_m(z)^{\tilde{r}} \tilde{\lambda}^Z(\mathrm{d}z) < \infty$.

Assumption 3 states that the volatility is an Itô semimartingale. In contrast to the log-price process, this is not a consequence of the no-arbitrage principle. However, it is commonly assumed, because it restricts the local behavior of volatility that is necessary to apply standard estimates for semimartingales in the high-frequency paradigm (see Jacod and Protter, 2012). For example, if σ^Z is a diffusion, the assumption follows by an application of Itô's Lemma under certain smoothness conditions. Either way, it is fulfilled for most stochastic volatility models applied in practice, even when volatility has sources of randomness not captured by Z. The ability to correlate the Brownian motions of Z and σ^Z is particularly relevant for financial applications that involve equity data, in order to capture the so-called leverage effect (Black, 1976; Christie, 1982). Note that (4.4) does

rule out volatility models driven by a fractional Brownian motion, which feature prominently in some strands of research (e.g., Comte and Renault, 1998; Gatheral, Jaisson, and Rosenbaum, 2018). The main difficulty is that in the "rough" regime, the volatility behaves very erratically over short time intervals, and this makes it harder to control the discretization error (see Chong and Todorov, 2023). Christensen, Thyrsgaard, and Veliyev (2019) allow σ^Z to satisfy a Hölder-type condition, in the expectation sense. It may be possible to extend our theoretical framework in this direction as well, but we leave it for a future endeavor.

Second, we restrict the memory of volatility processes. To accomplish this, we introduce the additional notation for a generic process Z:

$$\alpha_t = \sup_{A \in \mathcal{F}_s, B \in \mathcal{F}^{s+t}} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|. \tag{4.5}$$

Here, $\mathcal{F}_s = \sigma(Z_u; u \leq s)$ and $\mathcal{F}^t = \sigma(Z_u; u \geq t)$ are the backward- and forward-looking σ-algebras generated by Z until time s and from time t, respectively. Moreover, α_t is the mixing coefficient function that measures the degree of stochastic (in)dependence between them. This is used in the definition of strong mixing introduced by Rosenblatt (1956a).

Definition **1.** *Z is strong mixing (or α-mixing) if $\alpha_t \to 0$ as $t \to \infty$.*

Assumption **4.** *The volatility process $(V_t^X, V_t^Y)_{t \geq 0}$ is stationary and α-mixing in the sense of Definition 1 with $\alpha_t = O(t^{-(1+\tau)})$ for some $\tau > 0$.*

Assumption 4 is essential in deriving the asymptotic distribution theory. It corresponds to Assumption 5 in Christensen, Thyrsgaard, and Veliyev (2019), but for the two-dimensional case, and is again a weak regularity condition that can be verified for most stochastic volatility models. The stationary condition makes it meaningful to speak of a time-invariant distribution function of volatility, $H(x, y) = \mathbb{P}(V_0^X \leq x, V_0^Y \leq y)$, from which $F(x) = \mathbb{P}(V_0^X \leq x)$ and $G(y) = \mathbb{P}(V_0^Y \leq y)$ can be extracted. Sklar's theorem then guarantees existence, and uniqueness under further conditions, of the copula $C(u, v) = \mathbb{P}(F(V_0^X) \leq u, G(V_0^Y) \leq v)$. The question is then to what extent $\widehat{H}_{n,T}(x, y)$, consistently estimating $H_T(x, y)$ as $\Delta_n \to 0$, also provides a good proxy of the stationary distribution $H(x, y)$, since there is only a single realization to work with. This, at a bare minimum, additionally requires $T \to \infty$, so sufficient conditions to apply the ergodic theorem on the volatility path are needed to prompt the law of large numbers for consistency. This is implied by the strong mixing condition, and the polynomial decay of α_t ensures that volatility persistence decays suitably fast to derive the limiting distribution of the statistic. In this regard, we should note that since our CLT is derived for a sequence of bounded random variables, all moments exist, so the usual moment condition "trade-off" in Davydov's inequality is irrelevant, and hence the requirement for the mixing coefficient function is the weakest possible and the autocorrelation function merely needs to be integrable.

A prominent counterexample (ruled out by Assumption 3), where Assumption 4 typically fails, is long-memory processes, such as the fractional Gaussian noise with Hurst exponent $H \in (0.5, 1.0)$, which is stationary and weak mixing, but not strong mixing.

The next assumption imposes a uniform boundedness on the coefficient processes.

ASSUMPTION **5.** *For all $p \geq 1$,*

$$\sup_{t\geq 0}\left\{\mathbb{E}\left[|b_t^Z|^p + |\sigma_t^Z|^p + |\tilde{b}_t^Z|^p + |\tilde{\sigma}_t^Z|^p\right] + \int_{\mathbb{R}}(1\wedge\Gamma(z))^p\lambda^Z(\mathrm{d}z)\right\} \leq C, \tag{4.6}$$

for a constant C.

This follows Assumption 1 in Christensen, Thyrsgaard, and Veliyev (2019) and Andersen, Thyrsgaard, and Todorov (2019). It imposes a light-tailedness condition on the distribution of the various coefficient processes driving the continuous part of X and Y, and the large price jumps, but not for the jumps of volatility. It is a common assumption in the long-span setting. The existence of moments of any order is more than we require, but it streamlines the exposition. The technical reason for this condition is that the standard estimates for the increments of semimartingales based on the localization procedure of Jacod and Protter (2012, Section 4.4.1) presumes T is fixed. The above assumption facilitates the derivation of such estimates over $[0,\infty)$.

ASSUMPTION **6.** *$H_T(x,y)$ is continuously differentiable on $\mathbb{R}^2_+$, almost surely. The partial derivatives (that exist, almost surely) are denoted by $\partial_x H_T(x,y)$ and $\partial_y H_T(x,y)$, respectively, and we assume that $\partial_\bullet H_T(x,y) \neq 0$ such that $\mathbb{E}\left[\partial_\bullet H_T(x,y)\right] \leq C$ for almost every $(x,y)\in\mathbb{R}^2_+$, where C is a bounding constant that does not depend on (x,y).*

This combines Assumption 2 of Christensen, Thyrsgaard, and Veliyev (2019) with Assumption B of Li, Todorov, and Tauchen (2013). The smoothness condition is used to derive a uniform upper bound for estimation error of the empirical copula function. The requirement $\partial_\bullet H_T(x,y) \neq 0$ rules out "flat spots" in the marginal distribution to ensure that $H_T(x,y) \mapsto C_T(u,v)$ is smooth and non-singular, so that we can apply the functional delta method without getting a degenerate limit process with zero variance. We note that the boundedness in expectation is weaker than asking $\mathbb{E}[\sup_{(x,y)\in\mathbb{R}^2_+}\partial_\bullet H_T(x,y)] \leq C$, since the latter requires almost sure control of the pathwise behavior of volatility.

At last, we state an extra condition that is required to establish the central limit theorem. To accomplish this, we introduce the notation $k_n = \Delta_n^{-\gamma}$ and for any $\iota > 0$

$$\begin{aligned} d_n &= \Delta_n^{\gamma-1+(2-r)\varpi} + \Delta_n^{\gamma/2-\iota} + \Delta_n^{(1-\gamma)/2-\iota} + \Delta_n^{\frac{1-\gamma}{1+\tilde{r}}-\iota} \text{ and} \\ d_n' &= \Delta_n^{\gamma/r-(1-\varpi)-\iota} + \Delta_n^{\gamma/2-\iota} + \Delta_n^{(1-\gamma)/2-\iota} + \Delta_n^{\frac{1-\gamma}{1+\tilde{r}}-\iota}. \end{aligned} \tag{4.7}$$

ASSUMPTION **7.** *We assume that $\frac{r-1}{r} < \varpi < \frac{1}{2}$ and $r(1-\varpi) < \gamma < 1$ for $r > 1$. In addition, $0 < \varpi < \frac{1}{2}$ and $1-(2-r)\varpi < \gamma < 1$ for $r \leq 1$. Moreover, $\frac{1-\gamma}{1+\tilde{r}} > 0$ for $\tilde{r} > 0$. Finally, we assume that either of the following conditions holds for any $\iota > 0$:*

$$1).\ r \leq 1 \text{ and } \sqrt{T}d_n \to 0 \quad \text{or} \quad 2).\ r > 1 \text{ and } \sqrt{T}d_n' \to 0. \tag{4.8}$$

This is identical to the conditions in Theorem 4 of Li, Todorov, and Tauchen (2013), except for the appearance of $\sqrt{T}$ since we operate in the long-span setting. It implies that the volatility error, $\hat{V}_t^Z - V_t^Z$, is dominated by T, so the high-frequency discretization part is asymptotically negligible. The four terms in d_n and d_n' represent the various sources of the estimation error in $H_{n,T}(x,y)$. Specifically, the first term captures the order of the truncation procedure, induced by replacing truncated increments by those of the continuous part. It depends on whether the jump component has finite or infinite variation. The second and third terms capture the sampling variability and the discretization bias in approximating the spot volatilities. The last order is for the collection of summands over the sampling intervals that contain jumps (with a size larger than some level) in the volatility. The choices of ϖ, r, and γ imply $d_n \to 0$ and $d_n' \to 0$. Thus, we can assume $T \to \infty$ such that Assumption 7 holds. In particular, for $r \in (0, \frac{1}{2})$ and $\tilde{r} = 0$, we can take $\varpi \in (\frac{3}{8-4r}, \frac{1}{2})$ and $\gamma = 1/2$, from which it follows that $d_n = O(\Delta_n^{\frac{1}{4}-\iota})$. In this case, $T = O(\Delta_n^{-1/4+\epsilon})$ with $\epsilon > 0$ verifies the requirement.

The next theorem delivers the functional CLT of the estimation procedure for the stationary distribution function of volatility and the associated copula. It is grounded in a double asymptotic setting with $\Delta_n \to 0$ and $T \to \infty$ and extends Theorem 1 in Li, Todorov, and Tauchen (2013) and Theorem 3.1 in Christensen, Thyrsgaard, and Veliyev (2019).

THEOREM 2. *Suppose that Assumptions 1 - 7 hold (with $r = 2$ in Assumption 1 and $\tilde{r} = 2$ in Assumption 3). As $\Delta_n \to 0$ and $T \to \infty$, it follows that*

$$\sqrt{T}\left(\widehat{H}_{n,T}(x,y) - H(x,y)\right) \Rightarrow \mathcal{G}. \tag{4.9}$$

Here, "$\Rightarrow$" denotes weak convergence in the space $\mathbb{D}(\mathbb{R}_+^2)$ of càdlàg functions equipped with the uniform topology. Also, $\mathcal{G}$ is a Brownian bridge on $\mathbb{R}_+^2$ with covariance function

$$\operatorname{cov}\big(\mathcal{G}(x,y), \mathcal{G}(x',y')\big) \equiv \operatorname{avar}_{\mathcal{G}}(x,y,x',y') = 2\int_0^\infty \Big(H_t(x,y,x',y') - H(x,y)H(x',y')\Big)\mathrm{d}t, \tag{4.10}$$

where $H_t(x,y,x',y') = P(V_0^X \le x, V_0^Y \le y, V_t^X \le x', V_t^Y \le y')$. Moreover, it follows that

$$\sqrt{T}\left(\widehat{C}_{n,T}(u,v) - C(u,v)\right) \Rightarrow \mathcal{C}, \tag{4.11}$$

where $\mathcal{C}$ is a stochastic process on $[0,1]^2$, derived from $\mathcal{G}$, that is defined as:

$$\mathcal{C}(u,v) = \mathbf{g}(u,v)^\top \mathbf{Z}(u,v). \tag{4.12}$$

Here,

$$\begin{aligned} \mathbf{Z}(u,v) &= \left[\mathcal{G}\big(Q^X(u), Q^Y(v)\big), \mathcal{G}\big(Q^X(u), \infty\big), \mathcal{G}\big(\infty, Q^Y(v)\big)\right]^\top, \\ \mathbf{g}(u,v) &= \left[1, -\partial_u C(u,v), -\partial_v C(u,v)\right]^\top, \end{aligned} \tag{4.13}$$

while $Q^X(u) = \lim_{T\to\infty} Q_T^X(u)$ and $Q^Y(v) = \lim_{T\to\infty} Q_T^Y(v)$.

The proof of Theorem 2 draws on the functional CLT for stationary mixing sequences of bounded random variables (see Theorem 18.5.4 in Ibragimov (1975) and Theorem 5.2 in Dehay (2005)) building on Rosenblatt (1956a). We reiterate that the rate of convergence and the asymptotic distribution of $\widehat{H}_{n,T}(x,y)$ and $\widehat{C}_{n,T}(x,y)$ are unaffected by the error from recovering volatility, $V_t^Z - \widehat{V}_t^Z$, which is asymptotically negligible under the rate conditions in (4.7).

The covariance function of $\mathcal{C}(u,v)$ follows from Theorem 2 and the functional delta rule, i.e.

$$\text{avar}_{\mathcal{C}}(u,v,u',v') \equiv \text{cov}\big(\mathcal{C}(u,v),\mathcal{C}(u',v')\big) = \mathbf{g}(u,v)^\top \Gamma(u,v,u',v')\mathbf{g}(u',v'), \tag{4.14}$$

where

$$\Gamma(u,v,u',v') = \text{cov}\big(\mathbf{Z}(u,v),\mathbf{Z}(u',v')\big). \tag{4.15}$$

As a consequence,

$$\begin{aligned} \sqrt{T}\left(\widehat{H}_{n,T}(x,y) - H(x,y)\right) &\xrightarrow{d} N\big(0, \text{avar}_{\mathcal{G}}(x,y,x,y)\big), \\ \sqrt{T}\left(\widehat{C}_{n,T}(u,v) - C(u,v)\right) &\xrightarrow{d} N\big(0, \text{avar}_{\mathcal{C}}(u,v,u,v)\big), \end{aligned} \tag{4.16}$$

where $\xrightarrow{d}$ is regular convergence in law. This result can be used to construct pointwise confidence intervals, or conduct hypothesis tests, once an estimator of the asymptotic covariance function has been designed.

To construct such an estimator, we set:

$$\widehat{\text{avar}}_{\mathcal{G}}(x,y,x',y') = 2\int_0^{T^\xi} \left\{\widehat{H}_{t,n,T}(x,y,x',y') - \widehat{H}_{n,T}(x,y)\widehat{H}_{n,T}(x',y')\right\} \mathrm{d}t, \tag{4.17}$$

with $\widehat{H}_{t,n,T}(x,y,x',y') = \frac{1}{T-T^\xi}\int_0^{T-T^\xi} \mathbb{1}_{\left\{\hat{V}_s^X \le x, \hat{V}_s^Y \le y, \hat{V}_{s+t}^X \le x', \hat{V}_{s+t}^Y \le y'\right\}} \mathrm{d}s$, where ξ is a small positive constant.

Corollary **1.** *We suppose that Assumptions 1 - 7 hold, and that $\xi \in (0, 1/3)$. As $\Delta_n \to 0$ and $T \to \infty$, we deduce that*

$$\widehat{\text{avar}}_{\mathcal{G}}(x,y,x',y') \xrightarrow{\mathbb{P}} \text{avar}_{\mathcal{G}}(x,y,x',y'). \tag{4.18}$$

In the above, T^ξ plays the role of the "lag length" in the estimation of the long-run variance of a stationary time series, which should not grow too fast for a consistent estimation. In fact, the optimal order of the bandwidth is often $T^{1/3}$ (see, e.g., Newey and West, 1994). The choice of ξ relates to this observation.

By Slutsky's theorem,

$$\frac{\sqrt{T}\left(\widehat{H}_{n,T}(x,y) - H(x,y)\right)}{\sqrt{\widehat{\text{avar}}_{\mathcal{G}}(x,y,x,y)}} \xrightarrow{d} N(0,1). \tag{4.19}$$

Conducting inference about $C(u,v)$ is a more complicated endeavor, since it also involves partial derivatives of $C(u,v)$ on top of the covariance function $\text{avar}_{\mathcal{G}}(x,y,x',y')$. To construct estimators

of the former, we propose a nonparametric kernel-based smoothing approach:[11]

$$\begin{aligned}\widehat{\partial_u C}(u,v) &= \frac{1}{hT}\int_0^T\int_0^v K\left(\frac{\widehat{F}_{n,T}(\hat{V}_t^X)-u}{h}, \frac{G_{n,T}(\hat{V}_t^Y)-w}{h}\right)\mathrm{d}w\mathrm{d}t,\\ \widehat{\partial_v C}(u,v) &= \frac{1}{hT}\int_0^T\int_0^u K\left(\frac{\widehat{F}_{n,T}(\hat{V}_t^X)-w}{h}, \frac{G_{n,T}(\hat{V}_t^Y)-v}{h}\right)\mathrm{d}w\mathrm{d}t.\end{aligned} \tag{4.20}$$

Here, K is defined on a compact subset of $\mathbb{R}^2$, Lipschitz continuous, and such that

$$\int_{\mathbb{R}^2} K(x,y)(x^2+y^2)\mathrm{d}x\mathrm{d}y \leq C \qquad \text{and} \qquad \int_{\mathbb{R}^2} K(x,y)\mathrm{d}x\mathrm{d}y = 1. \tag{4.21}$$

This is a common regularity condition for kernel density estimation. Note that K is not required to be symmetric.

To show consistency of $\widehat{\partial_u C}(u,v)$ and $\widehat{\partial_v C}(u,v)$, we take a slight detour by forming a preliminary nonparametric estimator of the copula density, i.e. $c(u,v) = \partial_u\partial_v C(u,v)$,

$$\hat{c}_{n,T}(u,v) = \frac{1}{Th^2}\int_0^T K\left(\frac{\widehat{F}_{n,T}(\hat{V}_t^X)-u}{h}, \frac{\widehat{G}_{n,T}(\hat{V}_t^Y)-v}{h}\right)\mathrm{d}t. \tag{4.22}$$

The following condition is sufficient to get consistency of $\hat{c}_{n,T}(u,v)$.

Assumption **8.** *$H_T(x,y)$ is three times continuously differentiable on $\mathbb{R}^2_+$.*

In other words, we implicitly assume that the empirical copula density $c_T(u,v)$ exists and is itself continuously differentiable on $[0,1]^2$.

Corollary **2.** *We suppose that Assumptions 1 - 8 hold. As $\Delta_n \to 0$, $T \to \infty$, $h \to 0$, such that*

$$\frac{d_n}{h^3} \vee \frac{1}{\sqrt{T}h^3} \to 0, \tag{4.23}$$

where d_n is defined in (4.7)*, for any $(u,v) \in [0,1]^2$ it holds that*

$$\hat{c}_{n,T}(u,v) \xrightarrow{\mathbb{P}} c(u,v). \tag{4.24}$$

The partial derivatives of $C(u,v)$ can be written as $\partial_u C(u,v) = \int_0^v c(u,w)\mathrm{d}w$ and $\partial_v C(u,v) = \int_0^u c(w,v)\mathrm{d}w$ by Assumption 8. Hence, $\widehat{\partial_u C}(u,v)$ and $\widehat{\partial_v C}(u,v)$ are Riemann approximations of these integrals based on the kernel density estimator $\hat{c}_{n,T}(u,v)$. Therefore, the consistency of the partial derivatives in (4.20) follows directly, namely

$$\widehat{\partial_u C}(u,v) \xrightarrow{\mathbb{P}} \partial_u C(u,v) \quad \text{and} \quad \widehat{\partial_v C}(u,v) \xrightarrow{\mathbb{P}} \partial_v C(u,v). \tag{4.25}$$

In view of (4.14) and (4.25), we define the following plug-in estimator for the covariance functional of the realized copula of volatility, $\mathrm{avar}_{\mathcal{C}}(u,v,u',v')$:

$$\widehat{\mathrm{avar}}_{\mathcal{C}}(u,v,u',v') = \hat{\mathbf{g}}(u,v)^\top \hat{\Gamma}(u,v,u',v')\hat{\mathbf{g}}(u',v'), \tag{4.26}$$

[11] This idea was introduced in Rosenblatt (1956b).

where

$$\hat{\mathbf{g}}(u,v) = \Big[1, -\widehat{\partial_u C}(u,v), -\widehat{\partial_v C}(u,v)\Big]^\top, \qquad \hat{\Gamma}(u,v,u',v') = \operatorname{cov}\big(\hat{\mathbf{Z}}(u,v), \hat{\mathbf{Z}}(u',v')\big), \tag{4.27}$$

and

$$\hat{\mathbf{Z}}(u,v) = \big[\mathcal{G}\big(Q^X_{n,T}(u), Q^Y_{n,T}(v)\big), \mathcal{G}\big(Q^X_{n,T}(u), \infty\big), \mathcal{G}\big(\infty, Q^Y_{n,T}(v)\big)\big]^\top \tag{4.28}$$

is a three-dimensional, mean zero Gaussian vector whose covariance is estimated via the long-run covariance functional in (4.17). Then, by dominance of convergence in probability,

$$\frac{\sqrt{T}\left(\widehat{C}_{n,T}(u,v) - C(u,v)\right)}{\sqrt{\widehat{\operatorname{avar}}_{\mathcal{C}}(u,v,u,v)}} \xrightarrow{d} N(0,1). \tag{4.29}$$

4.1 Uniform confidence bands

The pointwise inference derived above ensures asymptotically correct coverage and testing size for each fixed value of (u,v), or (x,y), but it conceals that the functional CLT established in Theorem 2 offers stronger control of the measurement errors. In this subsection, we exploit this observation to construct uniform confidence bands for the volatility copula.

Corollary **3.** *Under the assumptions of Theorem 2:*

$$\sup_{(u,v)\in[0,1]^2} \sqrt{T}\left(\widehat{C}_{n,T}(u,v) - C(u,v)\right) \xrightarrow{d} \sup_{(u,v)\in[0,1]^2} \mathcal{C}(u,v), \tag{4.30}$$

where $\mathcal{C}$ is the Gaussian process defined in Theorem 2.

Corollary 3 employs the continuous mapping theorem for functional convergence in law, since the supremum functional is continuous in the uniform metric. However, the limiting distribution of the supremum point is non-standard, so we propose a simulation procedure to construct valid confidence bands.

We define the $(1-\alpha)$-quantile of the right-hand side in (4.30) as follows:

$$q^{CB}_{1-\alpha} = \inf\left\{c > 0 : \mathbb{P}\left(\sup_{(u,v)\in[0,1]^2} |\mathcal{C}(u,v)| \le c\right) \ge 1-\alpha\right\}. \tag{4.31}$$

Thus, it follows that as $T \to \infty$,

$$\mathbb{P}\left(\sup_{(u,v)\in[0,1]^2} \left|\widehat{C}_{n,T}(u,v) - C(u,v)\right| \le \frac{q^{CB}_{1-\alpha}}{\sqrt{T}}\right) \to 1-\alpha, \tag{4.32}$$

so an asymptotic $(1-\alpha)$ uniform confidence band for $C(u,v)$ is given by

$$\left\{\left[\widehat{C}_{n,T}(u,v) \pm \frac{q^{CB}_{1-\alpha}}{\sqrt{T}}\right], \quad \text{for all } (u,v) \in [0,1]^2\right\}. \tag{4.33}$$

To compute the constant $q_{1-\alpha}^{CB}$, we follow a Monte Carlo approach, which builds on a consistent estimator of the covariance function of $\mathcal{C}(u,v)$. To implement it, we discretize $[0,1]^2$ into an $m \times m$ grid $\{(u_i, v_j)\}_{i,j=1}^m$ and form the corresponding $m^2 \times m^2$ covariance matrix, $\widehat{\text{avar}}_{\mathcal{C}}$. Based on the Cholesky decomposition of $\widehat{\text{avar}}_{\mathcal{C}} = LL^\top$, we repeat the following procedure for $b = 1, \ldots, B$:

1. Draw $z^{(b)} \overset{\text{i.i.d.}}{\sim} N(0, I_{m^2})$.
2. Set $\mathcal{Z}^{(b)} = Lz^{(b)}$, such that $\mathcal{Z}^{(b)} \sim N(0, \widehat{\text{avar}}_{\mathcal{C}})$.
3. Compute $S_{(b)} = \max_{1 \leq k \leq m^2} \left| \mathcal{Z}_k^{(b)} \right|$.
4. Retrieve the $(1-\alpha)$-quantile of $\left\{ S_{(1)}, \ldots, S_{(B)} \right\}$:

$$\widehat{q}_{1-\alpha}^{CB} = S_{(\lceil (1-\alpha)B \rceil)}. \tag{4.34}$$

5. Construct the feasible uniform confidence band:

$$\left\{ \left[\widehat{C}_{n,T}(u,v) \pm \frac{\widehat{q}_{1-\alpha}^{CB}}{\sqrt{T}} \right], \quad \text{for all } (u,v) \in [0,1]^2 \right\}. \tag{4.35}$$

The validity of this resampling procedure is formally established by the next proposition, which is a functional version of Slutsky's theorem.

PROPOSITION 1. *We assume the conditions of Theorem 2 and Corollary 1 - 2 hold, such that*

$$\sqrt{T} \left(\widehat{C}_{n,T}(u,v) - C(u,v) \right) \Rightarrow \mathcal{C}, \tag{4.36}$$

where $\mathcal{C}$ is a centered Gaussian process with continuous covariance function $\text{avar}_{\mathcal{C}}(u,v,u,v)$ bounded away from zero, i.e. $\inf_{(u,v) \in [0,1]^2} \text{avar}_{\mathcal{C}}(u,v,u,v) > 0$. Also, let $\widehat{\text{avar}}_{\mathcal{C}}$ be the plug-in estimator defined in (4.26), which consistently estimates the covariance function of $\mathcal{C}$ at the grid points. Conditionally on $\widehat{\text{avar}}_{\mathcal{C}}$, let $(\widehat{\mathcal{C}}_{n,T})$ be a centered Gaussian processes with covariance matrix $\widehat{\text{avar}}_{\mathcal{C}}$. Write

$$\widehat{q}_{1-\alpha}^{CB} = \inf \left\{ c > 0 : \mathbb{P} \left(\sup_{i,j} \left| \widehat{\mathcal{C}}_{n,T}(u_i, v_j) \right| \leq c \mid \widehat{\text{avar}}_{\mathcal{C}} \right) \geq 1 - \alpha \right\}, \tag{4.37}$$

Then, as the grid is made arbitrarily fine, i.e. $m \to \infty$, with $\Delta_n \to 0$ and $T \to \infty$, for any constant $c \geq 0$, we deduce that

$$\mathbb{P} \left(\sup_{(u,v) \in [0,1]^2} \left| \widehat{\mathcal{C}}_{n,T}(u,v) \right| \leq c \mid \widehat{\text{avar}}_{\mathcal{C}} \right) \xrightarrow{\mathbb{P}} \mathbb{P} \left(\sup_{(u,v) \in [0,1]^2} |\mathcal{C}(u,v)| \leq c \right), \tag{4.38}$$

It follows from the proposition that $\widehat{q}_{1-\alpha}^{CB} \xrightarrow{\mathbb{P}} q_{1-\alpha}^{CB}$. Therefore, the uniform confidence band constructed from $\widehat{q}_{1-\alpha}^{CB}$ has the correct asymptotic coverage:

$$\mathbb{P} \left(\sup_{(u,v) \in [0,1]^2} \left| \widehat{C}_{n,T}(u,v) - C(u,v) \right| \leq \frac{\widehat{q}_{1-\alpha}^{CB}}{\sqrt{T}} \right) \to 1 - \alpha. \tag{4.39}$$

4.2 Goodness-of-fit testing

The previous analysis can also be extended to construct a goodness-of-fit test of the volatility copula function, which is another prominent application of our functional convergence results. In particular, we introduce a null hypothesis and an alternative hypothesis:

$$\mathcal{H}_0 : C = C_0 \quad \text{and} \quad \mathcal{H}_1 : C \neq C_0. \tag{4.40}$$

Here, equality is almost everywhere on $[0,1]^2$.

To construct a testing procedure, we let

$$G_{n,T}(u,v) \equiv \sqrt{T}\left(\widehat{C}_{n,T}(u,v) - C_0(u,v)\right). \tag{4.41}$$

We now arrive at the following statement, which again follows from the continuous mapping theorem for functional convergence applied to Theorem 2.

COROLLARY **4.** *Under the assumptions of Theorem 2 and conditional on $\mathcal{H}_0$:*

$$\|G_{n,T}\|_{L_2}^2 \xrightarrow{d} \|\mathcal{C}\|_{L_2}^2. \tag{4.42}$$

The limiting distribution in Corollary 4 (a sort of weighted infinite mixture of χ^2 random variables) is again not standard, so we employ a Monte Carlo procedure, closely related to the one outlined above, to form a test of $\mathcal{H}_0$ versus $\mathcal{H}_1$. In particular, we let $q_{1-\alpha}^{GF}$ be a consistent estimator of the $(1-\alpha)$-quantile of the distribution of $\|\mathcal{C}\|_{L_2}^2$, and we define a rejection region:

$$\mathcal{R} = \left\{ \|G_{n,T}\|_{L_2}^2 \geq q_{1-\alpha}^{GF} \right\}. \tag{4.43}$$

Then, in view of Corollary 4,

$$\mathbb{P}(\mathcal{R} \mid \mathcal{H}_0) \to \alpha \quad \text{and} \quad \mathbb{P}(\mathcal{R} \mid \mathcal{H}_1) \to 1. \tag{4.44}$$

To obtain $q_{1-\alpha}^{GF}$, we approximate the distribution of the random variable $\|\mathcal{C}\|_{L_2}^2$ by a weighted Riemann sum on a, possibly unequally spaced, grid $\{(u_i, v_j)\}_{i,j=1}^m \subset [0,1]^2$, i.e. $\sum_{1\leq i,j\leq m} w_{ij}\mathcal{C}(u_i,v_j)^2$, where w_{ij} are weights induced by the grid and $\sum_{i,j} w_{ij} = 1$. Note that $\{\mathcal{C}(u_i,v_j)\}_{i,j=1}^m$ is a sequence of m^2 normal random variates with explicit covariance matrix given by (4.15) following Theorem 2, which we denote by Γ. Therefore,

$$\sum_{i,j} w_{ij}\mathcal{C}(u_i,v_j)^2 \overset{d}{=} \sum_{k=1}^{m^2} \pi_k \chi_k^2, \tag{4.45}$$

where χ_k^2 are i.i.d. $\chi^2(1)$ random variables and π_k are the eigenvalues of $W^{1/2}\Gamma W^{1/2}$ with $W = \operatorname{diag}\big(\{w_{ij}\}_{i,j=1}^m\big)$, for $k = 1, \ldots, m^2$.

We draw repeated observations of the right-hand side of (4.45) based on the eigenvalues from the estimator $W^{1/2}\widehat{\Gamma}W^{1/2}$ to simulate the distribution of $\|G_{n,T}\|_{L_2}^2$ under $\mathcal{H}_0$, from which an appropriate quantile can be extracted.[12]

[12]Since $\widehat{\Gamma}$ is an estimator of a covariance matrix, it can possess negative eigenvalues. We follow Andersen, Su, Todorov, and Zhang (2024) and retain only those terms in the sum associated with positive eigenvalues.

To estimate the covariance matrix Γ, we rearrange the pairs $\{(u_i, v_j)\}_{i,j=1}^m$ as $\{(u'_k, v'_k)\}_{k=1,\dots,m^2}$ and compute $\widehat{\Gamma}_{k\ell} = \hat{g}_k^\top \hat{A}_{k\ell} \hat{g}_\ell$, where

$$\begin{aligned} \hat{g}_k &= \Big[1, -\widehat{\partial_u C}(u_k, v_k), -\widehat{\partial_v C}(u_k, v_k)\Big]^\top, \\ \hat{A}_{kl} &= \operatorname{cov}\big(\hat{\mathbf{Z}}(u_k, v_k), \hat{\mathbf{Z}}(u_l, v_l)\big), \end{aligned} \tag{4.46}$$

and

$$\hat{\mathbf{Z}}_k = \big[\mathcal{G}\big(Q^X_{n,T}(u_k), Q^Y_{n,T}(v_k)\big), \mathcal{G}\big(Q^X_{n,T}(u_k), \infty\big), \mathcal{G}\big(\infty, Q^Y_{n,T}(v_k)\big)\big]^\top. \tag{4.47}$$

It is also possible to test a composite null hypothesis:

$$\mathcal{H}_0 : C \in \mathcal{S} \qquad \text{versus} \qquad \mathcal{H}_a : C \notin \mathcal{S},$$

where $\mathcal{S} = \{C_\theta : \theta \in \Theta \subseteq \mathbb{R}^p\}$ is a statistical model for the copula function, and θ is a p-dimensional parameter vector.

In practice, the true value of the copula parameter, θ_0, is unknown. A complication occurs if the parameter is estimated from the full sample, in which case the estimator, $\hat{\theta}$, is typically $\sqrt{T}$-consistent matching the convergence rate of $\widehat{C}_{n,T}(u, v)$. Then, $G_{n,T}(u, v) = \sqrt{T}(\widehat{C}_{n,T}(u, v) - C_{\hat{\theta}}(u, v)) = \sqrt{T}(\widehat{C}_{n,T}(u, v) - C_{\theta_0}(u, v)) - \sqrt{T}(\hat{\theta} - \theta_0)^\top \partial_\theta C_{\theta_0}(u, v) + o_p(1)$, where $\partial_\theta C_{\theta_0}(u, v) = \partial_\theta C_\theta|_{\theta=\theta_0}$, so $G_{n,T}$ converges in law to a Gaussian process *plus* a bias term that is a linear functional of $\sqrt{T}(\hat{\theta} - \theta_0) \xrightarrow{d} \mathcal{W} \sim N(0, \Sigma_\theta)$. Hence, the L^2-statistic $\|G_{n,T}\|^2_{L_2} \xrightarrow{d} \|\mathcal{C} - \langle \partial_\theta C_{\theta_0}, \mathcal{W}\rangle\|^2_{L_2}$ is a quadratic form of a Gaussian process, which depends on unknown quantities. In other words, the asymptotic distribution is non-pivotal, as it depends on the parameter vector θ_0, the gradient $\partial_\theta C_{\theta_0}(u, v)$ and the law of $(\mathcal{W}, \mathcal{C})$. However, we can easily resample the empirical process and the plug-in estimator of the covariance function using a parametric or multiplier bootstrap under $C_{\hat{\theta}}$ to approximate the distribution of $\|G_{n,T}\|^2_{L_2}$ and construct critical values without analytically having to remove the dependence on θ_0.

On the other hand, to avoid the effect of $\hat{\theta}$, we can estimate the copula parameter such that $\hat{\theta} - \theta_0$ achieves a convergence rate faster than $\widehat{C}_{n,T}(u, v) - C_{\theta_0}(u, v)$.[13] In this case, the test statistic is asymptotically pivotal and we can replace C_0 by $C_{\hat{\theta}}$ in the previous testing procedure with a single null hypothesis and attain a valid inference that does not involve nuisance parameters and is shape-only. Hence, it suffices to prepare a single table of critical values.

5 Simulation analysis

In this section, we conduct a Monte Carlo study to explore the finite sample properties of our realized estimator of the empirical copula of stochastic volatility. We restrict attention to the bivariate setting and assume that a pair of asset log-price processes follow the motion

$$\mathrm{d}X_{i,t} = \mu_i \mathrm{d}t + \sqrt{V_{i,t}} \mathrm{d}W_{i,t}, \tag{5.1}$$

[13] This can be accomplished, for example, by reserving data in a shorter time interval, say $[0, T^\eta]$ with $0 < \eta < 1$, to construct $\widehat{C}_{n,T}$, but still employing data over the whole time interval $[0, T]$ to get $\hat{\theta}$.

where

$$\mathrm{d}\log V_{i,t} = \kappa\left(\eta_i - \log V_{i,t}\right)\mathrm{d}t + \beta_i \mathrm{d}B_{i,t}, \tag{5.2}$$

for $i = 1$ and 2.

In the above, W and B are Brownian motions with leverage correlation $\mathrm{corr}[\mathrm{d}W_{i,t}\mathrm{d}B_{i,t}] = \rho = -\sqrt{0.5}$. We assume a common speed of mean reversion $\kappa = 10$. The other parameters are $\mu_1 = 0.10$, $\eta_1 = \log(0.05)$, $\beta_1 = 3$, $\mu_2 = 0.05$, $\eta_2 = \log(0.01)$, and $\beta_2 = 3$. This design is intended to represent a risky asset with an expected return and volatility of 10% and 20% and a near risk-free asset with an expected return and volatility of 5% and 10%. It emulates our empirical work in Section 6, where we look at high-frequency data from stock index and medium-duration treasury bond futures contracts.

A routine calculation for the stationary Gaussian Ornstein-Uhlenbeck process shows that the unconditional distribution of $\log V_{i,t}$ has the form

$$\log V_{i,t} \sim N\left(\eta_i, \frac{\beta_i^2}{2\kappa}\right). \tag{5.3}$$

We denote the distribution function as $F_{\log V_i}(v)$, for $-\infty < v < \infty$.

The univariate log-variance processes are tied together with a Gumbel copula:

$$C(u,v) = \exp\left(-\left[(-\log u)^\theta + (-\log v)^\theta\right]^{1/\theta}\right), \quad \text{for } 0 \le u, v \le 1, \tag{5.4}$$

which is parameterized by $\theta \in [1, \infty)$. A larger value indicates a stronger relationship. The independence copula—$C(u,v) = uv$—corresponds to $\theta = 1$, while the comonotonicity copula—$C(u,v) = \min(u,v)$—induced by the upper Fréchet-Hoeffding bound is the limit as $\theta \to \infty$. In our empirical application, we explore the properties of this particular copula, because our goodness-of-fit test suggests that it provides an excellent description of the dependence between the log-variance of the equity and treasury bond market. In line with our maximum likelihood estimate from that section, we set $\theta = 2$.

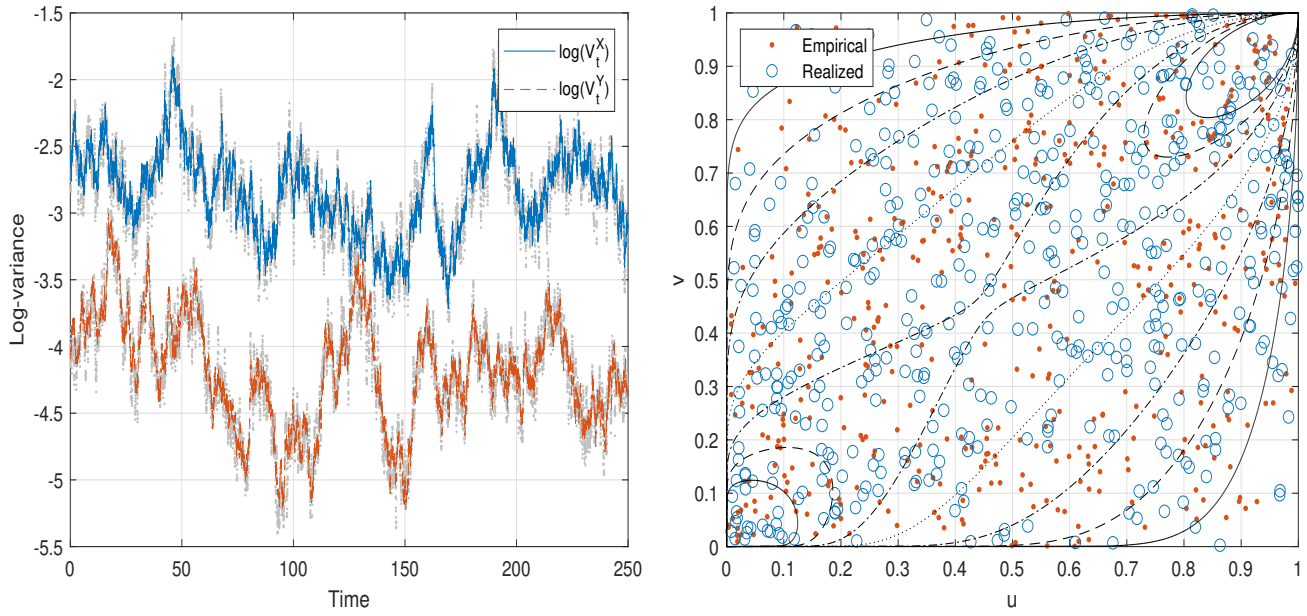
To construct a realization of this process, we begin by drawing from the Gumbel copula to generate a pair of dependent random variables $(U_{1,0}, U_{2,0})$, where $U_{i,0} \sim U(0,1)$. These serve to initialize the log-variance processes via $\log V_{i,0} = F^{-1}_{\log V_i}(U_{i,0})$. We then apply an Euler approximation with constant step size Δ to iteratively generate the sample path of (5.1)–(5.2) forward in time, using the Metropolis–Hastings algorithm to ensure that each proposal update is consistent with the target distribution.[14]

In the "continuous-time" version of the model, we set $\Delta = 1/N$ with $N = 23{,}400$. We assume data is captured over a time interval $[0, T]$ with $T = 50, 100, \ldots, 2000$, but the log-price process is only observed at a much coarser equidistant grid $t_i = i/n$ with $n = 39, 78, 390$, for $i = 0, 1, \ldots, nT$. This design is intended to imitate a security that is traded in a 6.5 hour window each day and whose price is recorded at the ten-, five-, and one-minute frequency over at most an eight-year

[14]The symbol Δ does double-duty in this context. In isolation it is the time increment $\Delta = t - s$, but when it is applied to a stochastic process Y, it represents the difference operator $\Delta Y_{i,t} = Y_{i,t} - Y_{i,s}$.

Figure 1: Illustration of bivariate SV model with Gumbel copula.

Panel A: Log-variance. Panel B: Copula.



Note. In Panel A, we show a realization of the log-variance processes of X and Y together with the associated log-realized variance series. The latter appear as a grey-shaded color in the background. In Panel B, we construct a scatter of 500 randomly selected pairs of the foregoing series, after transforming them through their empirical distribution function. We contrast this with selected contours of the probability density function of the Gumbel copula, which are increasing as we move toward the 45°-line.

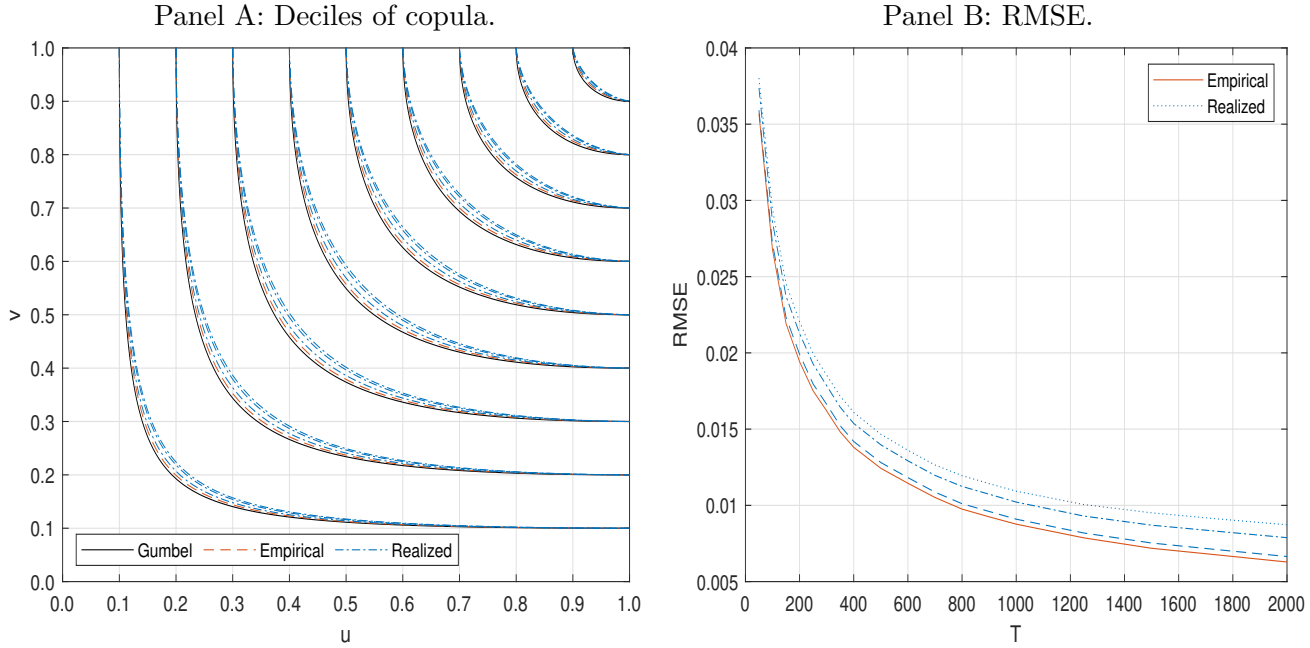
period. This resembles the dataset analyzed in our empirical application. To estimate the latent spot volatility, associated with each n we employ a bandwidth of $h_n = 36, 48, 120$. With the above interpretation, this amounts to a six-, four-, and two-hour window. Hence, as n increases we collect a larger amount of high-frequency data while reducing the time span of the estimation window, as stipulated by the rate conditions in the asymptotic theory. These choices follow Li, Todorov, and Tauchen (2013) and Christensen, Thyrsgaard, and Veliyev (2019).[15]

A sample path of each log-variance process, together with the log-realized variance, is illustrated in Panel A of Figure 1. In Panel B, we report a scatter of 500 randomly selected pairs of the foregoing series, after transforming them through their empirical distribution functions. We contrast this with selected contours of the probability density function of the Gumbel copula. The latter are increasing as we move toward the 45°-line and eventually disconnect at the lower-left and upper-right corners. The main takeaway is that while the log-variance observations are dispersed as expected, the estimation error in the log-realized variance shifts several observations into regions where the Gumbel copula has almost no concentration of mass.

The outcome of the full-blown simulation analysis is summarized in Figure 2. We perform

[15]In unreported results that are available at request, we conducted a more extensive set of numerical experiments with other bandwidth choices. This showed that the average squared estimation error of the spot volatility estimator was basically flat over an extended region, so long as h_n was not too small or too large.

Figure 2: Simulation result.



Note. In Panel A, we construct a decile plot of the distribution function of the Gumbel copula, which we compare to the deciles of the empirical copula of volatility (based on $T = 2{,}000$) and the realized version (with $n = 30$ to $n = 390$ as we approach the empirical copula). In Panel B, we plot the RMSE of the empirical and realized copula of volatility in terms of their estimation accuracy for the limiting Gumbel copula.

$M = 1{,}000$ repetitions in total. In Panel A, we construct a decile plot of the bivariate distribution function of the Gumbel copula, which we contrast against the empirical copula of volatility—based on $T = 2{,}000$—along with the estimated one recovered by the realized version. We observe that even with a conservative amount of high-frequency data, the empirical copula is close to its stationary counterpart with only some minor deviations. In contrast, the contours of the realized version are generally farther away and located to the northeast of the empirical copula. The explanation is that the sampling variation in the realized variances render them less correlated and shifts mass in the direction of the independence copula.[16] As n increases, however, the realized copula of volatility steadily trends toward the empirical counterpart, which is a validation of the asymptotic theory developed in Section 4.

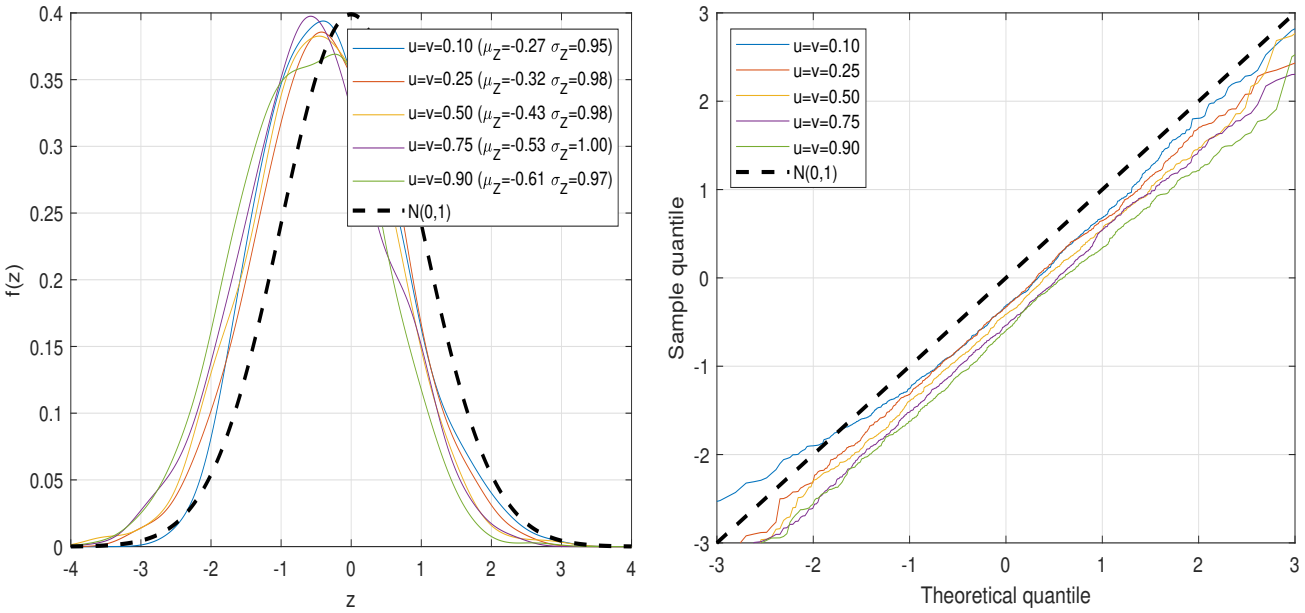
In Panel B of Figure 2, we report the root mean squared error (RMSE):

$$\text{RMSE}(x) = \frac{1}{M}\sum_{m=1}^{M}\sqrt{\int_0^1\int_0^1 [C_m^x(u,v) - C(u,v)]^2 \mathrm{d}u\mathrm{d}v}, \tag{5.5}$$

where $C(u,v)$ is the Gumbel copula given by the expression in (5.4) and $C_m^x(u,v)$ represents either the empirical ($x = \text{Empirical}$) or realized ($x = \text{Realized}$) estimate during the mth simulation. The figure reveals that, as expected, the RMSE of the empirical copula of volatility approaches zero, as $T \to \infty$, which is consistent with our theoretical exploration. It takes about a quadrupling of the

[16] The level curves of the independence copula are given by $u = c/v$ for $c \in (0,1)$.

Figure 3: Finite sample distribution of studentized realized copula of volatility.
Panel A: Kernel density estimate. Panel B: Q–Q plot.



Note. In Panel A, we report a kernel density estimate (based on a Gaussian kernel with bandwidth chosen by Silverman's rule of thumb) of the studentized statistic $Z_{\mathcal{C}} = \sqrt{T}\left(\widehat{C}_{n,T}(u,v) - C(u,v)\right)/\sqrt{\text{avar}_{\mathcal{C}}(u,v,u,v)}$, for $n = 390$, $T = 2{,}000$, and $u = v = (0.10, 0.25, 0.50, 0.75, 0.90)$. A standard normal curve (dashed line) is superimposed as a visual benchmark. In each setting, we also report the sample average and standard deviation of $Z_{\mathcal{C}}$. In Panel B, we show the corresponding normal Q–Q plot. The Monte Carlo experiment has 1,000 replications.

long-span dimension to reduce the RMSE in half, which reflects its convergence rate. Turning to the realized copula of volatility, we see that the RMSE is generally higher, and while it too drops as T increases, it tends to flatten out strictly above zero. This is because the estimation error does not vanish with a fixed n. Boosting n, however, lessens the discretization error, and by the time that $n = 390$ the realized version is within proximity of the empirical copula (its target in the in-fill limit as $n \to \infty$).

To gauge the functional CLT for the realized copula of volatility as an approximation to its finite sample behavior, we next explore the distribution of the asymptotic pivot:

$$Z_{\mathcal{C}} = \frac{\sqrt{T}\big(\widehat{C}_{n,T}(u,v) - C(u,v)\big)}{\sqrt{\text{avar}_{\mathcal{C}}(u,v,u,v)}} \xrightarrow{d} N(0,1), \tag{5.6}$$

for $u = v = 0.10, 0.25, 0.50, 0.75, 0.90$.

Figure 5 presents the outcome. We base the analysis on $n = 390$ and $T = 2{,}000$, which adheres to our empirical work in the next section, and the estimator of the asymptotic variance introduced in Section 4. In Panel A, we show kernel density estimates, while Panel B reports the associated quantile-quantile (Q–Q) plot. We include the sample average and sample standard deviation of $Z_{\mathcal{C}}$ for each setting. We observe that the kernel density estimates are well-aligned with the standard normal curve, and the Q–Q plots follow the 45^o-line reasonably. The dispersion is

almost as predicted, but the test statistic does exhibit a minuscule downward bias of about 0.2-0.5 standard deviation. Importantly, these numbers are more or less identical to those we compute for the empirical copula of volatility.[17]

At last, we turn our attention toward the goodness-of-fit test. We calculate rejection rates of the test statistic in (4.42) from Corollary 4 with $n = 390$ and $T = 250, 500, 1000, 2000$. To curb the computational cost, we implement the bootstrapping algorithm for producing critical values with 5,000 draws on a 5×5 grid $(u, v) \in \{0.10, 0.25, 0.50, 0.75, 0.90\}^2$, yielding 25 evaluation markers.[18]

The null hypothesis (size) is based on the Gumbel copula with $\theta = 2$, which has Kendall's tau given by $\tau = 1 - 1/\theta$, so that our choice implies $\tau = 0.5$. The alternative hypothesis (power), is given by three distinct parametric models. The first benchmark is the independence copula, $C(u, v) = uv$. This corresponds to the absence of volatility codependency with $\tau = 0$ and represents a stark violation of $\mathcal{H}_0$. It provides a natural baseline for assessing the ability of the test to detect tail concentration. The second choice is the Clayton copula, $C(u, v) = (u^{-\alpha} + v^{-\alpha} - 1)^{-1/\alpha}$ with parameter $\alpha \geq 0$. This is another example of an Archimedean copula used to capture positive tail dependence, but in contrast to the Gumbel copula, the dependence is located in the other extreme of the distribution. The Clayton copula has $\tau = \alpha/(\alpha + 2)$, and we set $\alpha = 2$—or $\tau = 0.5$—to be consistent with the Gumbel. This design intends to isolate departures from the null in the form of dependence in the opposite direction, while fixing the overall rank correlation summarized by Kendall's tau. At last, we complement the analysis with a power experiment based on an alternative member of the Gumbel class. Specifically, we examine a Gumbel copula with $\theta = 1.5$—or $\tau = 1/3$—that also has upper tail concentration, albeit slightly weaker compared to the null.

The results are reported in Table 1. We inspect the nominal significance levels 10%, 5%, and 1%. The size analysis is shown in Panel A. We observe that the test is to a small extent oversized for the lowest value of T, but it rapidly settles around the expected level. Moving to Panel B, we learn that the power versus the independence copula is almost unity across scenarios, even at the smallest level of significance. Moreover, the test has a good ability to distinguish between the Gumbel null and the Clayton alternative, especially with larger sample sizes. Lastly, while the rejection rates against the Gumbel(1.5) alternative are initially higher than those of the Clayton copula, but they increase less fast with T. The power function is possibly flatter in the Gumbel direction, where the functional form is identical and only the strength of the upper tail dependence is different.

Overall, our simulations suggest that the realized copula of volatility can reliably infer the latent volatility copula at a sampling frequency that is relevant for practical analysis. Moreover, the small sample properties of the standardized statistic are close to those implied by Theorem 2, which

[17]The kernel density estimates and Q–Q plots for the empirical copula of volatility are excluded from the presentation for brevity, but they can be shared at request.

[18]It is arguably preferable to use a more refined grid, or possibly even alter how the points are scattered over the domain. The intuition is that if the test statistic gets a better resolution of the realized copula of volatility, it has more scope to detect departures from the null, and this may improve rejection rates under the alternative. However, this also increases the dimensionality of the required covariance matrix, impacting memory usage and runtime speed. In line with this argument, we surmise that the power of the test reported here is conservative.

Table 1: Rejection rates of the goodness-of-fit test.

	$T = 250$			$T = 500$			$T = 1000$			$T = 2000$		
	10%	5%	1%	10%	5%	1%	10%	5%	1%	10%	5%	1%
Panel A: Size												
Gumbel(2)	0.135	0.088	0.035	0.116	0.072	0.024	0.111	0.068	0.023	0.094	0.056	0.017
Panel B: Power												
Independence	0.895	0.845	0.699	0.987	0.972	0.910	1.000	1.000	1.000	1.000	1.000	1.000
Clayton	0.230	0.167	0.085	0.330	0.194	0.087	0.550	0.312	0.092	0.936	0.814	0.371
Gumbel(1.5)	0.428	0.335	0.174	0.511	0.418	0.255	0.672	0.581	0.378	0.850	0.780	0.600

Note. We report Monte Carlo rejection rates based on 1,000 replica. The simulation design is described in the main text. Panel A with size is for the null hypothesis. Panel B reports power against the three alternatives listed in the first column.

is assuring, because it means that confidence intervals and hypothesis tests based on the normal distribution are accurate. We also uncover that the goodness-of-fit test has appropriate size control and excellent power against many relevant alternatives. A word of caution, though. Our goodness-of-fit test merely provides evidence against the null hypothesis as a whole. Thus, a rejection of a simple null in favor of the alternative should not be interpreted as identifying the exact source of misspecification. Indeed, it can arise either because the assumed functional form of the copula is wrong, or because the family is correctly specified but the concrete parameter value is not.

6 Empirical application

The forensic analysis presented here aims to determine the codependency between the volatility processes of a pair of leading financial indicators. We delve into high-frequency transaction data from futures contracts that track the aggregate U.S. equity index and treasury bond market, namely the E-mini S&P 500 (ES) and 10-year treasury note (TY). The chosen instruments are listed on the Globex platform at the Chicago Mercantile Exchange (CME). They trade around the clock five days per week, but we restrict attention to the most active hours from 9:30am to 4:00pm EST that overlap with the NYSE trading session. The data at our disposal were purchased from Tickdata (`https://www.tickdata.com/`). The sample covers the period from March 18, 2010 to October 14, 2021. We remove short days associated with a reduced trading schedule, which leaves $T = 2{,}891$ days for our empirical investigation.[19]

In Table 2, we report a few descriptive statistics of the retained data. The "N" column reports the average daily number of transactions (in 1000s). These futures contracts are very liquid. However, sampling the tick-by-tick data at the maximal resolution can be detrimental for volatility estimation, because of microstructure noise, such as price discreteness and bid-ask spread (e.g.,

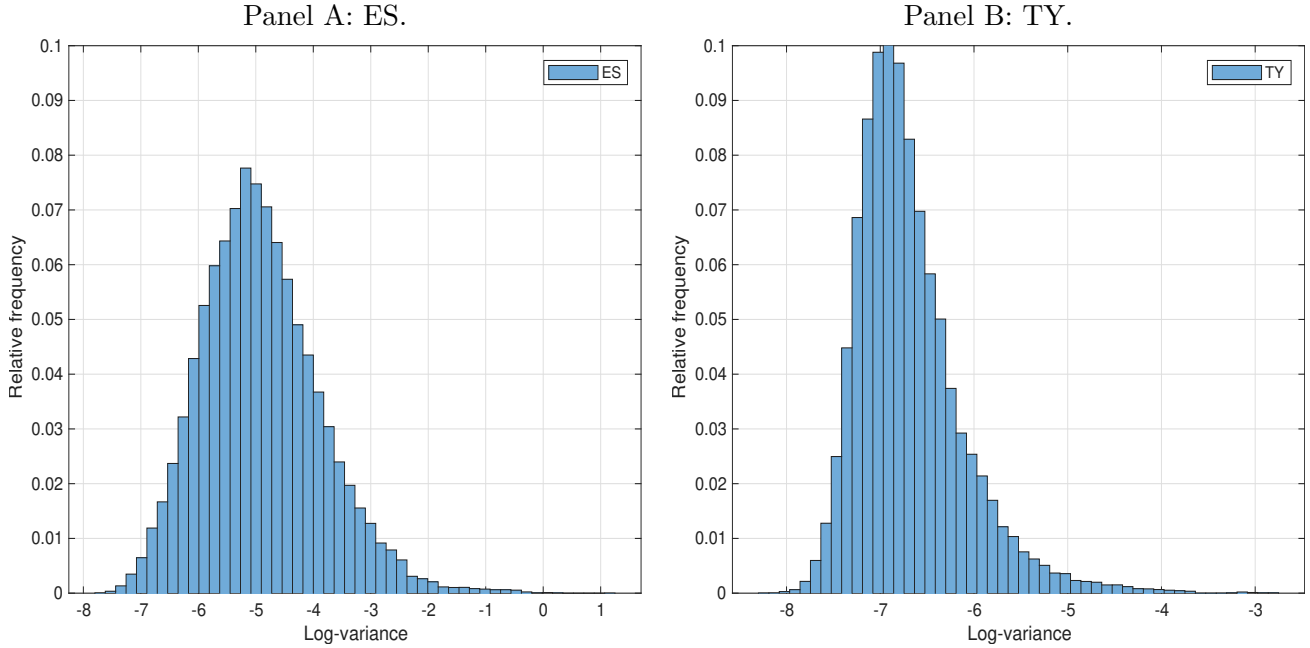
[19]To construct a consecutive price series, we employ a built-in procedure in the extraction software delivered by Tickdata. It rolls the front contract over into the back contract about a week before expiration.

Table 2: Descriptive statistics of CME futures contract data.

		Log-realized variance							
Ticker	N	Min.	1Q	Median	Mean	3Q	Max.	Skewness	Kurtosis
ES	338.78	-7.72	-5.63	-5.00	-4.91	-4.29	1.23	0.59	3.78
TY	90.95	-8.27	-7.08	-6.80	-6.69	-6.43	-2.80	1.39	6.42

Note. Ticker is the short name of the futures contract, N is the average daily number of transactions (in 1000s), whereas Min., 1Q, Median, Mean, 3Q, Max., Skewness, and Kurtosis are for the distribution of the log-realized variance over the whole sample.

Figure 4: Histogram of log-realized variance.



Note. In Panel A, we show a histogram (plotted in terms of relative frequency) of the log-realized variance of ES, whereas Panel B is the associated plot for the log-realized variance of TY.

Hansen and Lunde, 2006). To alleviate this concern, we downsample to a one-minute equidistant frequency using previous tick imputation, so there are $n = 390$ log-returns available per day to excavate the volatility process.[20]

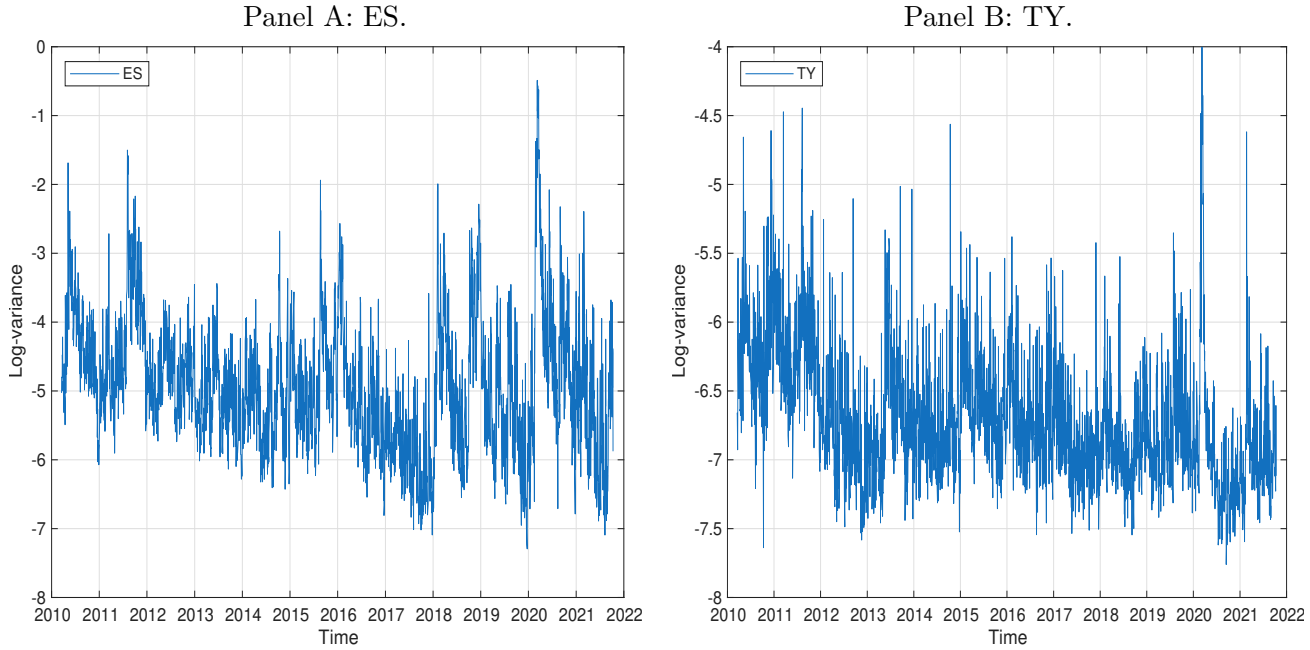
In what follows, we study the log-spot realized variance, which is less influenced by extreme outliers. We remark that by the invariance principle the copula of a random vector does not change under a strictly increasing transformation, so the realized copula of volatility is unaffected by studying the log-transform. In Figure 4, we show a relative frequency histogram for the log-realized variance of ES in Panel A, whereas Panel B is for TY. In agreement with the sample average and quantiles reported in Table 2, we observe that volatility in the stock market is, on average, higher than in the bond market. Furthermore, and consistent with Andersen, Bollerslev, Diebold,

[20]We follow the design of the simulation section in terms of choosing the various tuning parameters that are required to calculate the spot realized variance.

and Ebens (2001); Andersen, Bollerslev, Diebold, and Labys (2003), the shapes are close to the Gaussian bell curve, although both exhibit a mild positive skewness and excess kurtosis compared to the normal distribution. As evident from the right-hand side of Table 2, this effect is more pronounced in TY than ES.

To provide an initial assessment on the degree of volatility codependency in these time series, Figure 5 reports the average daily log-spot realized variance for ES in Panel A and TY in Panel B. The smoothing attenuates the measurement error in the spot volatility estimator a bit to better capture any comovement. There is forceful evidence of a positive relationship, most notably during periods of elevated market distress. To support this claim, in Panel A of Figure 6 we draw a scatter plot of the log-realized variance of ES (on the x-axis) against the log-realized variance of TY (on the y-axis). The graph reinforces the previous impression, which is further backed by Pearson's and Kendall's sample correlation coefficients.[21]

Figure 5: Time series of log-realized variance.



Note. We show the time series of the average daily log-realized variance of ES and TY. The sample period is from March 18, 2010 to October 14, 2021, or $T = 2{,}891$ days in total.

The realized copula of volatility is illustrated in Panel B of Figure 6. We plot 500 randomly drawn pairwise observations of $\Big(\widehat{F}_{n,T}\big(\log(\hat{V}_t^{\mathrm{ES}})\big), \widehat{G}_{n,T}\big(\log(\hat{V}_t^{\mathrm{TY}})\big)\Big)$ and add decile contours based on the whole sample. We contrast the latter with those implied by a parametric Gumbel copula, where the single parameter was estimated by maximum likelihood to be $\hat{\theta}_{\mathrm{MLE}} = 1.4786$. Overall, there is a close alignment between them.

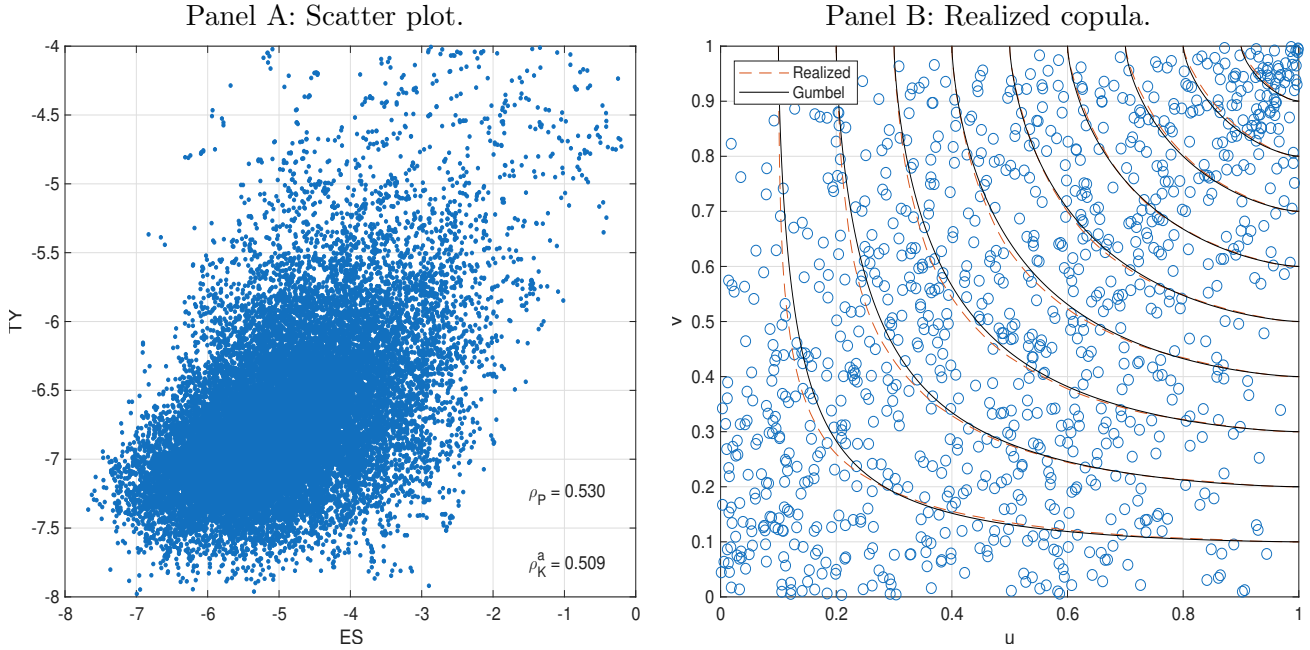
[21] Kendall's tau has been transformed as $\hat{\rho}_{\mathrm{K}}^{a} = \sin(\pi/2\hat{\rho}_{\mathrm{K}}^{u})$, where a (u) denotes the adjusted (unadjusted = raw) estimator. The correction—motivated by Greiner's equality—makes it unbiased for the correlation coefficient of a random sample from a bivariate normal distribution and, hence, its scale is comparable to the Pearson's measure.

To provide an alternative graphical representation of the realized copula of volatility, we construct a couple of descriptive statistics that are widely used in the literature. The first is an exceedance measure called tail concentration. It studies the amount of probability mass in the lower-left and upper-right quadrant of a bivariate distribution function, here in the form of a copula (see, e.g., Joe, 1993; Sibuya, 1960). In particular, the lower (L) and upper (U) tail concentration functions are defined as

$$L(z) = \mathbb{P}(U < z \mid V < z) = \frac{C(z,z)}{z} \quad \text{and} \quad U(z) = \mathbb{P}(U > z \mid V > z) = \frac{1 - 2z + C(z,z)}{1 - z}, \tag{6.1}$$

for any $z \in (0,1)$.[22]

Figure 6: Scatter plot of log-realized variance and realized copula of volatility.



Note. In Panel A, we show a scatter of log-realized variance of ES plotted against the log-realized variance of TY. In the lower right-hand corner, ρ_x is the sample correlation coefficient, where $x = \text{P}$ is Pearson's linear correlation and $x = \text{K}$ is Kendall's tau rank correlation. In Panel B, we draw 500 pairwise observations at random from our nonparametric realized copula of volatility. We superimpose deciles computed over the whole sample, which we contrast with those inferred from the cumulative distribution function of a Gumbel copula (with $\hat{\theta}_{\text{MLE}} = 1.4786$).

At one end of the domain, L is convergent toward $\lim_{z \to 1^-} L(z) = 1$, whereas R has the limit $\lim_{z \to 0^+} R(z) = 1$. This holds for every copula, and so it is not very informative. In the other end, however, $L(z)$ and $R(z)$ can converge to any number in the interval $[0,1]$. Provided the limits exist, the lower and upper tail concentration parameters are defined as $\lambda_L = \lim_{z \to 0^+} L(z)$ and $\lambda_U = \lim_{z \to 1^-} R(z)$. If either limit is nonzero, we say there is asymptotic dependence in that tail. Furthermore, noting that $L(0.5) = R(0.5)$, it is common to plot $T(z) = \min(L(z), R(z))$ as the aggregate tail concentration function.

[22] We note that the tail concentration function is symmetric, since its value does not change if we swap the position of U and V, because the latter are uniformly distributed.

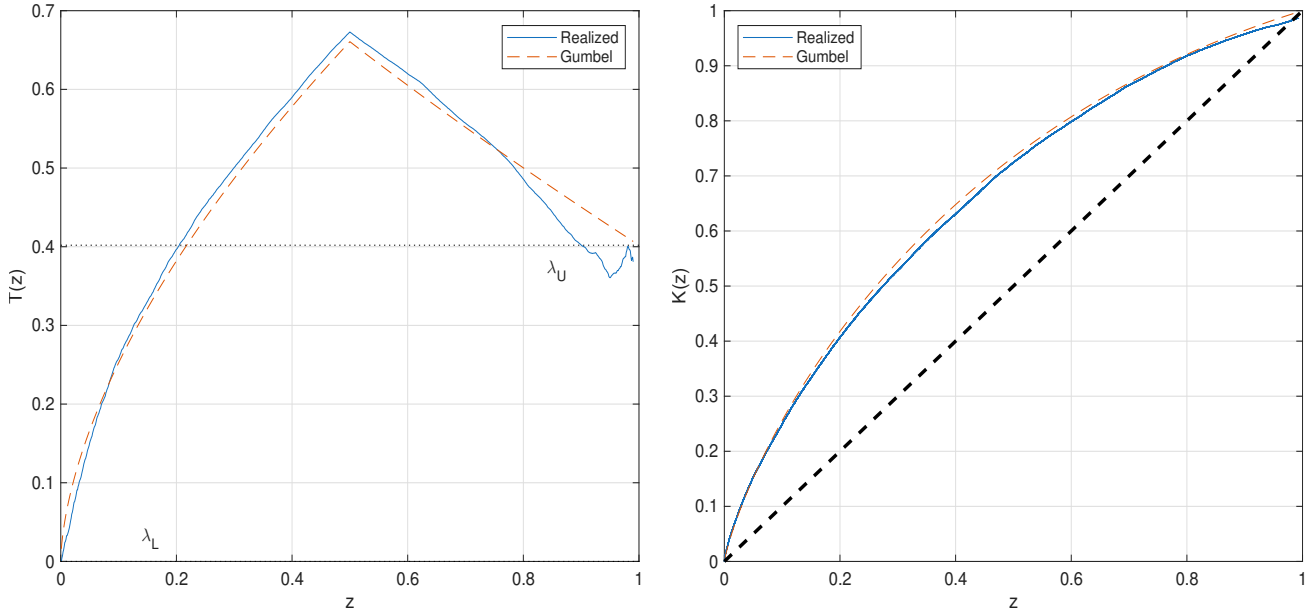
The second measure is Kendall's distribution function of a copula, which is defined as

$$K(z) = \mathbb{P}(C(U,V) \leq z). \tag{6.2}$$

Here, $C(U,V)$ is a univariate random variable, and (U,V) are uniform with bivariate distribution function $C(u,v)$. This can be interpreted as the multivariate version of the probability integral transform (see, e.g., Genest and Rivest, 1993).[23]

Figure 7: Graphical representation of the realized copula of volatility.

Panel A: Tail concentration. Panel B: Kendall's distribution function.



Note. In Panel A, we show the tail concentration function of the nonparametric realized copula of volatility, which we compare to the one of the parametric Gumbel copula, where the coefficient was estimated by maximum likelihood at $\hat{\theta}_{\text{MLE}} = 1.4786$. Also reported (as dashed horizontal lines) are the lower and upper tail concentration parameter, which are $\lambda_L = 0$ and $\lambda_U = 2 - 2^{1/\theta}$ for the Gumbel copula. In Panel B, we plot the Kendall's distribution function from (6.2). The 45°-line is a lower bound on $K(z)$ induced by the Fréchet-Hoeffding theorem.

We plot the estimated tail concentration function of the realized copula between the ES and TY (log-)variance processes in Panel A of Figure 7. In the graph, we also report the tail dependence of the Gumbel copula, again with $\hat{\theta}_{\text{MLE}} = 1.4786$. We recall that for the Gumbel copula, $\lambda_L = 0$ and $\lambda_U = 2 - 2^{1/\theta}$, which are superimposed as horizontal dotted lines. Overall, there is compelling evidence in favor of asymmetric tail dependence. While the lower tail concentration fizzles out and converges to zero, there is a nontrivial upper tail parameter, which suggests that elevated levels of volatility in the equity and bond markets are likely to occur in parallel. This feature is embedded in the Gumbel copula. Indeed, the tail concentration function of the realized copula of volatility and the associated one of the Gumbel copula are extremely close with a minimal distance, except for

[23] The connection between this probability measure and Kendall's tau is $\rho_{\text{K}} = 3 - 4\int_0^1 K(z)\mathrm{d}z$ (see, e.g., Schweizer and Wolff, 1981).

a small deviation in the far upper-right tail. We observe that the Gumbel copula has a Kendall's rank correlation of $\rho_{\mathrm{K}} = 1 - 1/\theta$. It thus describes variables that exhibit positive comovement, which is stronger for extremely large values of θ. We estimate an upper tail concentration of about $\hat{\lambda}_U = 0.40$ and a rank correlation of $\hat{\rho}^a_{\mathrm{K}} = 0.50$.

We corroborate this in Panel B of the figure by showing the estimated Kendall's distribution function. We overlay the population counterpart of the Gumbel copula. The latter is expressible in closed-form, i.e. $K(z) = z - \log(z^{1/\theta})$. We observe a near-perfect evolution with this measure over the entire support.

To assess goodness-of-fit, we examine a composite null hypothesis with the Gumbel, Clayton, and Independence copulas as contenders for describing the data. The Clayton and Independence copulas are both heavily rejected with a P-value close to zero, however. Conversely, the Gumbel copula is more appropriate with a P-value $= 0.1136$.

Taken together, the goodness-of-fit test coupled with the descriptive analysis from Figures 6 – 7 provide overwhelming evidence that the Gumbel copula offers a very good approximation of the codependency between the variance processes of the ES and TY futures contracts.

7 Conclusion

We propose the realized copula of volatility as a nonparametric tool for studying the codependency in the stochastic volatility component of a multivariate continuous-time asset price process. In analogy with the classical copula framework, the aim is to decouple the marginal behavior of the univariate volatility processes from their dependence structure. The distinctive challenge in this setting is that volatility is latent and has to be recovered with a realized measure constructed from discretely observed high-frequency data of the price process. Thus, inference on the realized copula of volatility requires control of measurement error.

We start by proposing a multivariate extension of the realized distribution function of volatility by Li, Todorov, and Tauchen (2013), from which the realized copula of volatility is derived. Under in-fill asymptotics with a fixed time span, we then show that it affords a consistent estimator of the empirical copula of the latent stochastic volatility (measuring the codependency of volatility on the part of the sample path observed so far). In a double-asymptotic framework with the long-span dimension also going to infinity, our estimator converges to the stationary marginal copula of volatility, provided it exists. We derive a functional central limit theorem for the realized-copula process, which supports uniform inference on the volatility copula. To render the limit theory operational, we design a feasible estimator of the asymptotic covariance function. Based on this, we develop pointwise confidence intervals, uniform confidence bands, and a goodness-of-fit test to evaluate a hypothesis about the marginal copula of volatility.

A simulation study sheds light on the finite sample properties of the realized copula of volatility and the associated inference procedures under realistic sampling schemes. We show that it accurately recovers the empirical copula of volatility for a fixed time span and converges at the

expected rate toward the marginal copula of volatility as the long-span dimension increases. The goodness-of-fit test is found to exhibit good size control and demonstrate good power.

In our empirical application, we construct the realized copula of volatility from transaction price data on futures contracts tracking the aggregate U.S. equity index and treasury bond market. The evidence presented points toward a Gumbel copula with asymptotic upper-tail dependence as being highly representative of the behavior of the realized variance processes in these data, in line with the notion that volatility across major asset classes tends to spike in tandem during financial market stress.

The asymptotic theory developed here for the latent volatility of semimartingale processes and the associated inference based on local realized proxies of spot variance can be extended in several directions in future research. First, the baseline analysis can be pursued in the presence of market microstructure noise by building on the univariate treatment of the realized distribution function of volatility in Christensen, Thyrsgaard, and Veliyev (2019). A second avenue is to allow the realized copula of volatility to vary with the state of the economy or with observable macro-financial indicators, so that dependence in volatility can differ across regimes or along the business cycle. A third extension is to embed the volatility copula in a higher-dimensional framework together with a model for time-varying correlations, thereby delivering a unified copula-based representation of the dynamics of the full covariance matrix. These developments require additional structure but can be built directly on the econometrics and inference procedures developed in this paper.

A Appendix

In this section, we prove the theoretical results presented in the main text. Throughout, we denote by C a generic positive constant. In the fixed T setting, without loss of generality from the standard localization procedure, as stated in Section 4.4.1 of Jacod and Protter (2012), Assumptions 1 and 3 can be strengthened into the following stronger statements.

ASSUMPTION B. *The processes b^Z and σ^Z are bounded. Moreover, for some bounded nonnegative function Γ on $\mathbb{R}$, $|\delta^Z(\omega,t,z)| \leq \Gamma(z)$ and $\int_{\mathbb{R}}(1 \wedge \Gamma(z))^r \lambda(\mathrm{d}z) < \infty$.*

ASSUMPTION $\tilde{\text{B}}$. *The processes $\tilde{b}^Z$ and $\tilde{\sigma}^Z$ are bounded. Moreover, for some bounded nonnegative function $\tilde{\Gamma}$ on $\mathbb{R}$, $|\tilde{\delta}^Z(\omega,t,z)| \leq \tilde{\Gamma}(z)$ and $\int_{\mathbb{R}}(1 \wedge \tilde{\Gamma}(z))^{\tilde{r}} \tilde{\lambda}(\mathrm{d}z) < \infty$.*

The following lemma, which extends Lemma 1 of Li, Todorov, and Tauchen (2013) to the multivariate setting, gives the consistency for the estimation of a general functional of the variance processes.

LEMMA 1. *Let $g : \mathbb{R}_+^2 \mapsto [0,1]$ be a measurable function and D_g be the collection of discontinuity point pairs of g. Suppose that*

(i) Assumption 1 holds for $r = 2$.

(ii) For Lebesgue almost everywhere $s \in [0,T]$, $\mathbb{P}\big((V_s^X, V_s^Y) \in D_g\big) = 0$.

Then, as $\Delta_n \to 0$,

$$\int_0^T g(\hat{V}_s^X, \hat{V}_s^Y)\mathrm{d}s \xrightarrow{\mathbb{P}} \int_0^T g(V_s^X, V_s^Y)\mathrm{d}s.$$

PROOF OF LEMMA 1: We denote $k_n\Delta_n = u_n$ and set $\hat{V}_s^{Z+} = \hat{V}^{Z+}_{\lceil \frac{s}{u_n} \rceil u_n}$, for $Z = X, Y$. By Lemma 1 of Li, Todorov, and Tauchen (2013), it follows that $\hat{V}_s^{Z+} \xrightarrow{\mathbb{P}} V_s$ for all $s \in \big(0, (\lfloor \frac{T}{u_n} \rfloor - 1)u_n\big)$, delivering that $(\hat{V}_s^{X+}, \hat{V}_s^{Y+}) \xrightarrow{\mathbb{P}} (V_s^X, V_s^Y)$. Thus, by condition (ii) and the Bounded Convergence Theorem (BCT), for Lebesgue almost everywhere $s \in [0,T]$, we get

$$\mathbb{E}\left[\left|g(\hat{V}_s^{X+}, \hat{V}_s^{Y+}) - g(V_s^X, V_s^Y)\right|\right] = \mathbb{E}\left[\left|(g(\hat{V}_s^{X+}, \hat{V}_s^{Y+}) - g(V_s^X, V_s^Y))\mathbb{1}_{\{(V_s^X, V_s^Y) \notin D_g\}}\right|\right] \to 0.$$

Furthermore,

$$\int_0^T g(\hat{V}_s^X, \hat{V}_s^Y)\mathrm{d}s = \int_0^{(\lfloor \frac{T}{u_n} \rfloor - 1)u_n} g(\hat{V}_s^{X+}, \hat{V}_s^{Y+})\mathrm{d}s + \int_0^{u_n} g(\hat{V}_s^X, \hat{V}_s^Y)\mathrm{d}s + \int_{\lfloor \frac{T}{u_n} \rfloor u_n}^T g(\hat{V}_s^X, \hat{V}_s^Y)\mathrm{d}s.$$

Since the function g is bounded on $[0,1]$, then

$$\mathbb{E}\left(\left|\int_0^T \big(g(\hat{V}_s^X, \hat{V}_s^Y) - g(V_s^X, V_s^Y)\big)\mathrm{d}s\right|\right) \leq Ku_n + \int_0^{(\lfloor \frac{T}{u_n} \rfloor - 1)u_n} \mathbb{E}\left(\left|g(\hat{V}_s^{X+}, \hat{V}_s^{Y+}) - g(V_s^X, V_s^Y)\right|\right)\mathrm{d}s.$$

To complete the proof, we follow up with a second application of the BCT. ■

We now turn our attention to the proof of Theorem 1, which is established with Lemma 1 as a precursor for the concrete test function $g(\cdot,\cdot\cdot) = \mathbb{1}_{\{\cdot\leq x,\cdot\cdot\leq y\}}$.

Proof of Theorem 1: This is a bivariate version of Theorem 1 of Li, Todorov, and Tauchen (2013) and Theorem 3.1 of Christensen, Thyrsgaard, and Veliyev (2019).

Part (1): We fix a point $(x,y) \in \mathbb{R}^2_+$. By Assumption 6, H_T is almost surely continuous, so

$$H_T(x,y) = H_T(x-,y) = H_T(x,y-) = H_T(x-,y-),$$

almost surely. Hence, we deduce that

$$\int_0^T \mathbb{P}(V_t^X = x, V_t^Y = y)\mathrm{d}t = T \cdot \mathbb{E}\left[H_T(x,y) - H_T(x,y-) - H_T(x-,y) + H_T(x-,y-)\right] = 0.$$

Therefore, we get $\mathbb{P}(V_t^X = x, V_t^Y = y) = 0$ for Lebesgue almost everywhere for $t \in [0,T]$, which establishes condition (ii) in Lemma 1. The pointwise consistency of $\widehat{H}_{n,T}(x,y)$, namely $\widehat{H}_{n,T}(x,y) \xrightarrow{\mathbb{P}} H_T(x,y)$, follows from Theorem 1 of Li, Todorov, and Tauchen (2013) and Lemma 1. Since $\widehat{H}_{n,T}$ and H_T are bounded and non-decreasing, pointwise convergence also implies locally uniform convergence.

Since the paths of the volatility processes σ_t^X and σ_t^Y are càdlàg, for any $\eta > 0$, we can find $0 < M < \infty$, such that

$$\mathbb{P}\left(\sup_{t\in[0,T]}\{V_t^X \vee V_t^Y\} > M\right) < \eta. \tag{A.1}$$

This implies that $\mathbb{P}(H_T(M,M) \neq 1) < \eta$, hence, $\mathbb{P}(F_T(M) \neq 1) < \eta$ and $\mathbb{P}(G_T(M) \neq 1) < \eta$. Then, for any $\epsilon > 0$, we get the decomposition

$$\begin{aligned}
\mathbb{P}\left(\sup_{(x,y)\in\mathbb{R}^2_+} |\widehat{H}_{n,T}(x,y) - H_T(x,y)| > \epsilon\right) &\leq \mathbb{P}\left(\sup_{(x,y)\in[0,M]^2} |\widehat{H}_{n,T}(x,y) - H_T(x,y)| > \epsilon\right) \\
&\quad + \mathbb{P}\left(\sup_{(x,y)\in(M,\infty)\times[0,M]} |\widehat{H}_{n,T}(x,y) - H_T(x,y)| > \epsilon\right) \\
&\quad + \mathbb{P}\left(\sup_{(x,y)\in[0,M]\times(M,\infty)} |\widehat{H}_{n,T}(x,y) - H_T(x,y)| > \epsilon\right) \\
&\quad + \mathbb{P}\left(\sup_{(x,y)\in(M,\infty)^2} |\widehat{H}_{n,T}(x,y) - H_T(x,y)| > \epsilon\right) \\
&\equiv \mathrm{I}_{n,T} + \mathrm{II}_{n,T} + \mathrm{III}_{n,T} + \mathrm{IV}_{n,T}.
\end{aligned}$$

The locally uniform convergence of $H_{n,T}(x,y)$ implies that $\limsup_{n\to\infty} \mathrm{I}_{n,T} = 0$, which is therefore asymptotically negligible. Second,

$$\mathbb{P}\left(\sup_{x>M,y\leq M} |\widehat{H}_{n,T}(x,y) - H_T(x,y)| > \epsilon\right) \leq \mathbb{P}\left(\sup_{x>M,y\leq M} |\widehat{H}_{n,T}(x,y) - \widehat{H}_{n,T}(M,y)| > \epsilon/3\right)$$

$$
\begin{aligned}
&+ \mathbb{P}\left(\sup_{y\le M} |\widehat{H}_{n,T}(M,y) - H_T(M,y)| > \epsilon/3\right) \\
&+ \mathbb{P}\left(\sup_{x>M, y\le M} |H_T(M,y) - H_T(x,y)| > \epsilon/3\right) \\
&\le \mathbb{P}\left(\sup_{y\le M} |\widehat{H}_{n,T}(\infty,y) - \widehat{H}_{n,T}(M,y)| > \epsilon/3\right) \\
&+ \mathbb{P}\left(\sup_{y\le M} |\widehat{H}_{n,T}(M,y) - H_T(M,y)| > \epsilon/3\right) \\
&+ \mathbb{P}\left(\sup_{y\le M} |H_T(M,y) - H_T(\infty,y)| > \epsilon/3\right).
\end{aligned}
$$

Now, by (A.1) and locally uniform convergence, we obtain

$$
\begin{aligned}
\limsup_{n\to\infty} \mathbb{P}\left(\sup_{y\in[0,M]} |\widehat{H}_{n,T}(\infty,y) - \widehat{H}_{n,T}(M,y)| > \epsilon/3\right) &\le \limsup_{n\to\infty} \mathbb{P}\left(|1 - \widehat{F}_{n,T}(M)| > \epsilon/3\right) \\
&\le \limsup_{n\to\infty} \mathbb{P}\left(|1 - F_T(M)| > \epsilon/6\right) \\
&+ \limsup_{n\to\infty} \mathbb{P}\left(|\widehat{F}_{n,T} - F_T(M)| > \epsilon/6\right) \\
&< \eta.
\end{aligned}
$$

We also get that

$$
\limsup_{n\to\infty} \mathbb{P}\left(\sup_{y\le M} |\widehat{H}_{n,T}(M,y) - H_T(M,y)| > \epsilon/3\right) = 0,
$$

and

$$
\mathbb{P}\left(\sup_{y\le M} |H_T(M,y) - H_T(\infty,y)| > \epsilon/3\right) \le \mathbb{P}(|1 - F_T(M)| > \epsilon/3\Big) < \eta.
$$

Taking η arbitrarily small, we conclude that $\limsup_{n\to\infty} \mathrm{II}_{n,T} = 0$ and $\limsup_{n\to\infty} \mathrm{III}_{n,T} = 0$.

Finally, following the above line of reasoning,

$$
\begin{aligned}
\limsup_{n\to\infty} \mathbb{P}\left(\sup_{(x,y)\in[M,\infty)^2} |\widehat{H}_{n,T}(x,y) - H_T(x,y)| > \epsilon\right) &\le \limsup_{n\to\infty} \mathbb{P}\left(|1 - \widehat{H}_{n,T}(M,M)| > \epsilon/3\right) \\
&+ \limsup_{n\to\infty} \mathbb{P}\left(|\widehat{H}_{n,T}(M,M) - H_T(M,M)| > \epsilon/3\right) \\
&+ \limsup_{n\to\infty} \mathbb{P}\left(|1 - H_T(M,M)| > \epsilon/3\right) \\
&\le 2\eta.
\end{aligned}
$$

Again, we can choose η as arbitrarily small as needed, so it follows that $\limsup_{n\to\infty} \mathrm{IV}_{n,T} = 0$. Hence, the proof of part (1) is done.

Part (2): By Corollary 1 in Li, Todorov, and Tauchen (2013), which builds on Lemma 21.2 in van der Vaart (1998), we deduce that

$$
\left(\widehat{Q}^X_{n,T}(u), \widehat{Q}^Y_{n,T}(v)\right) \xrightarrow{\mathbb{P}} \left(Q^X_T(u), Q^Y_T(v)\right).
$$

Note that

$$|\widehat{C}_{n,T}(u,v) - C_T(u,v)| \leq |\widehat{H}_{n,T}(Q^X_{n,T}(u), Q^Y_{n,T}(v)) - H_T(Q^X_{n,T}(u), Q^Y_{n,T}(v))| \\ + |H_T(Q^X_{n,T}(u), Q^Y_{n,T}(v)) - H_T(Q^X_T(u), Q^Y_T(v))|,$$

so pointwise convergence of $C_{n,T}(u,v)$ follows from the convergence of $\widehat{H}_{n,T}(x,y)$ and pointwise continuity of $Q^X_T(u)$ and $Q^Y_T(v)$. The uniform convergence follows from part (1) and the uniform continuity of $H_T(x,y)$. ■

The proof of Theorem 2 relies on the upcoming lemma, which shows that the estimation error of the realized distribution function of volatility and the realized copula of volatility, centered around their empirical counterpart and scaled with $\sqrt{T}$, converges to zero in probability as $\Delta_n \to 0$ and $T \to \infty$. It extends Lemma 2 and Theorem 2 of Li, Todorov, and Tauchen (2013) for T fixed.

LEMMA 2. *We suppose that Assumptions 1 - 6 hold and that $k_n = \Delta_n^{-\gamma}$. We choose $\frac{r-1}{r} < \varpi < \frac{1}{2}$ and $r(1-\varpi) < \gamma < 1$ for $r > 1$ or $0 < \varpi < \frac{1}{2}$ and $1 - (2-r)\varpi < \gamma < 1$ for $r \leq 1$. Then, under either of the following conditions:*

1. *V^X_t and V^Y_t are continuous on $\mathbb{R}_+$, and:*
   1). *$r \leq 1$: $\sqrt{T}\left(\Delta_n^{\gamma-1+(2-r)\varpi} + \Delta_n^{\gamma/2-\iota} + \Delta_n^{(1-\gamma)/2-\iota}\right) \to 0$,*
   2). *$r > 1$: $\sqrt{T}\left(\Delta_n^{\gamma/r-(1-\varpi)-\iota} + \Delta_n^{\gamma/2-\iota} + \Delta_n^{(1-\gamma)/2-\iota}\right) \to 0$,*
2. *V^X_t or V^Y_t are discontinuous on $\mathbb{R}_+$, and:*
   - *$r \leq 1$: $\sqrt{T}\left(\Delta_n^{\gamma-1+(2-r)\varpi} + \Delta_n^{\gamma/2-\iota} + \Delta_n^{(1-\gamma)/2-\iota} + \Delta_n^{\frac{1-\gamma}{1+r}-\iota}\right) \to 0$,*
   - *$r > 1$: $\sqrt{T}\left(\Delta_n^{\gamma/r-(1-\varpi)-\iota} + \Delta_n^{\gamma/2-\iota} + \Delta_n^{(1-\gamma)/2-\iota} + \Delta_n^{\frac{1-\gamma}{1+r}-\iota}\right) \to 0$.*

Then, it follows as $\Delta_n \to 0$ and $T \to \infty$ that

$$\sqrt{T}\left(\widehat{H}_{n,T}(x,y) - H_T(x,y)\right) \xrightarrow{\mathbb{P}} 0 \qquad \text{and} \qquad \sqrt{T}\left(\widehat{C}_{n,T}(u,v) - C_T(u,v)\right) \xrightarrow{\mathbb{P}} 0.$$

PROOF OF LEMMA 2: 1) If the volatility processes V^X and V^Y are continuous, we define

$$\eta^X_n = \sup_{t\in[0,\infty)} |\hat{V}^X_t - V^X_t|, \quad \eta^Y_n = \sup_{t\in[0,\infty)} |\hat{V}^Y_t - V^Y_t|,$$
$$\beta^X_n = \sup_{x\in[0,\infty)} |\widehat{F}_{n,T}(x) - F_T(x)|, \quad \beta^Y_n = \sup_{y\in[0,\infty)} |\widehat{G}_{n,T}(y) - G_T(y)|.$$

Following Lemma 2 and Theorem 2 of Li, Todorov, and Tauchen (2013), we know that η^X_n and β^X_n are both of order $O_P(d_n)$, with

$$d_n = \begin{cases} \Delta_n^{\gamma-1+(2-r)\varpi} + \Delta_n^{\gamma/2-\iota} + \Delta_n^{(1-\gamma)/2-\iota}, & r \leq 1, \\ \Delta_n^{\gamma/r-(1-\varpi)-\iota} + \Delta_n^{\gamma/2-\iota} + \Delta_n^{(1-\gamma)/2-\iota}, & r > 1. \end{cases}$$

Now, let $\alpha_n^X \equiv \sup_{t\in[0,\infty)} |\widehat{F}_{n,T}(\hat{V}_t^X) - F_T(V_t^X)|$ and $\alpha_n^Y \equiv \sup_{t\in[0,\infty)} |\widehat{G}_{n,T}(\hat{V}_t^Y) - G_T(V_t^Y)|$. Then,

$$\begin{aligned}
\left|\widehat{C}_{n,T}(u,v) - C_T(u,v)\right| &= \frac{1}{T}\int_0^T \left|\mathbb{1}_{\left\{\widehat{F}_{n,T}(\hat{V}_t^X)\le u, \widehat{G}_{n,T}(\hat{V}_t^Y)\le v\right\}} - \mathbb{1}_{\left\{F_T(V_t^X)\le u, G_T(V_t^Y)\le v\right\}}\right| \mathrm{d}t \\
&\le \frac{1}{T}\int_0^T \left|\mathbb{1}_{\{\widehat{F}_{n,T}(\hat{V}_t^X)\le u\}} - \mathbb{1}_{\{F_T(V_t^X)\le u\}}\right| \mathbb{1}_{\{\widehat{G}_{n,T}(\hat{V}_t^Y)\le v\}} \mathrm{d}t \\
&+ \frac{1}{T}\int_0^T \left|\mathbb{1}_{\{\widehat{G}_{n,T}(\hat{V}_t^Y)\le v\}} - \mathbb{1}_{\{G_T(V_t^Y)\le v\}}\right| \mathbb{1}_{\{\widehat{F}_{n,T}(V_t^X)\le u\}} \mathrm{d}t \\
&\le \frac{1}{T}\int_0^T \mathbb{1}_{\{u-\alpha_n^X \le F_T(V_t^X)\le u+\alpha_n^X\}} \mathrm{d}t + \frac{1}{T}\int_0^T \mathbb{1}_{\{v-\alpha_n^Y \le G_T(V_t^Y)\le v+\alpha_n^Y\}} \mathrm{d}t.
\end{aligned}$$

Notice that

$$\frac{1}{T}\int_0^T \mathbb{1}_{\{F_T(V_t^X)\le u\}} \mathrm{d}t = \frac{1}{T}\int_0^T \mathbb{1}_{\{V_t^X \le Q_T^X(u)\}} \mathrm{d}t = F_T(Q_T^X(u)) = u.$$

Hence,

$$\sqrt{T}\left|\widehat{C}_{n,T}(u,v) - C_T(u,v)\right| \le K\sqrt{T}(\alpha_n^X + \alpha_n^Y).$$

To calculate the order of α_n^X, we observe that

$$\begin{aligned}
\alpha_n^X &\le \sup_{t\in[0,\infty)} |\widehat{F}_{n,T}(\hat{V}_t^X) - F_T(\hat{V}_t^X)| + \sup_{t\in[0,\infty)} |F_T(\hat{V}_t^X) - F_T(V_t^X)| \\
&\le \beta_n^X + \sup_{t\in[0,\infty)} |F_T(\hat{V}_t^X) - F_T(V_t^X)|.
\end{aligned}$$

To handle the second term in the inequality, namely $\sup_{t\in[0,\infty)} |F_T(\hat{V}_t^X) - F_T(V_t^X)|$, we follow the proof of Lemma 2(b) in Li, Todorov, and Tauchen (2013) by observing that

$$\sup_{t\in[0,\infty)} |F_T(\hat{V}_t^X) - F_T(V_t^X)| \le \int_{\min\{V_t^X, \hat{V}_t^X\}}^{\max\{V_t^X, \hat{V}_t^X\}} f_T(z)\mathrm{d}z = O_P(d_n). \tag{A.2}$$

Similarly, we can get that

$$\alpha_n^Y = O_P(d_n).$$

Therefore, if the conditions in 1) hold, we obtain

$$\sqrt{T}(\alpha_n^X + \alpha_n^Y) = o_P(1).$$

2) To deal with the situation, where the volatility processes can jump, we introduce the set

$$\mathcal{I}_n(\epsilon) = \left\{0 \le i \le \lfloor T/u_n \rfloor - 1 : \mu\left(I_{n,i} \times \left\{z : \tilde{\Gamma}(z) > \epsilon\right\}\right) = 0\right\}.$$

where $I_{n,i} = [iu_n, (i+1)u_n)$. We also denote $B_n(\epsilon) = \cup_{i\in\mathcal{I}_n(\epsilon)} I_{n,i}$ and $B_n^c(\epsilon) = [0,T] \setminus B_n(\epsilon)$. Here, $\mathcal{I}_n(\epsilon)$ is a set of time indices that exclude the locations of the big jumps in volatility of size larger than ϵ. In this case,

$$\left|\widehat{C}_{n,T}(u,v) - C_T(u,v)\right| \le \frac{1}{T}\sum_{k=1}^{\lfloor T\rfloor} \int_{t_{k-1}}^{t_k} \left|\mathbb{1}_{\{\widehat{F}_{n,T}(\hat{V}_t^X)\le u\}} - \mathbb{1}_{\{F_T(V_t^X)\le u\}}\right| \mathrm{d}t + \frac{1}{T}.$$

where $t_k = \lfloor k/u_n \rfloor u_n$. For every k, it holds that

$$\int_{t_{k-1}}^{t_k} \left| \mathbb{1}_{\{\widehat{F}_{n,T}(\hat{V}_t^X) \leq u\}} - \mathbb{1}_{\{F_T(V_t^X) \leq u\}} \right| \mathrm{d}t = \int_{[t_{k-1},t_k] \cap B_n(\epsilon)} \left| \mathbb{1}_{\{\widehat{F}_{n,T}(\hat{V}_t^X) \leq u\}} - \mathbb{1}_{\{F_T(V_t^X) \leq u\}} \right| \mathrm{d}t + \int_{[t_{k-1},t_k] \cap B_n^c(\epsilon)} \left| \mathbb{1}_{\{\widehat{F}_{n,T}(\hat{V}_t^X) \leq u\}} - \mathbb{1}_{\{F_T(V_t^X) \leq u\}} \right| \mathrm{d}t.$$

Reusing Li, Todorov, and Tauchen (2013, Lemma 5), we deduce that

$$\int_{[t_{k-1},t_k] \cap B_n(\epsilon)} \left| \mathbb{1}_{\{\widehat{F}_{n,T}(\hat{V}_t^X) \leq u\}} - \mathbb{1}_{\{F_T(V_t^X) \leq u\}} \right| \mathrm{d}t = O_P(d_n),$$

and

$$\int_{[t_{k-1},t_k] \cap B_n^c(\epsilon)} \left| \mathbb{1}_{\{\widehat{F}_{n,T}(\hat{V}_t^X) \leq u\}} - \mathbb{1}_{\{F_T(V_t^X) \leq u\}} \right| \mathrm{d}t = O_P(d_n').$$

Here, $d_n' = \Delta_n^{(1-\gamma)/(1+\tilde{r})-\theta}$ and both equalities hold uniformly in k.

Therefore, if the conditions in 2) hold,

$$\sqrt{T}\left(\widehat{C}_{n,T}(u,v) - C_T(u,v)\right) \xrightarrow{\mathbb{P}} 0.$$

The proof that $\sqrt{T}(\widehat{H}_{n,T}(x,y) - H_T(x,y)) \xrightarrow{\mathbb{P}} 0$ follows this route too. ■

Now, we deliver the proof of Theorem 2. This is based on an application of the central limit theorem for mixing stationary sequences of bounded random variables (see Theorem 18.5.4 in Ibragimov (1975), Theorem 5.2 in Dehay (2005), and Lemma 2).

Proof of Theorem 2: We define the empirical process:

$$G_T \equiv \left\{ \sqrt{T}(H_T(x,y) - H(x,y)), (x,y) \in \mathbb{R}_+^2 \right\}.$$

To prove the weak convergence $G_T(x,y) \Rightarrow \mathcal{G}(x,y)$, we need to verify (1) convergence of finite-dimensional distributions and (2) tightness. To this end, we write

$$H_T(x,y) - H(x,y) = \frac{1}{T}\sum_{k=1}^{\lfloor T \rfloor} Z_k(x,y) + \frac{1}{T}\int_{\lfloor T \rfloor}^{T} \mathbb{1}_{\left\{V_t^X \leq x, V_t^Y \leq y\right\}} \mathrm{d}t,$$

where $Z_k(x,y) = \int_{k-1}^{k} \mathbb{1}_{\left\{V_t^X \leq x, V_t^Y \leq y\right\}} \mathrm{d}t - H(x,y)$. The second term on the right side of the equality is a remainder term, for which we immediately conclude that

$$\sqrt{T}\frac{1}{T}\int_{\lfloor T \rfloor}^{T} \mathbb{1}_{\left\{V_t^X \leq x, V_t^Y \leq y\right\}} \mathrm{d}t \leq \frac{1}{\sqrt{T}} \to 0,$$

almost surely. Thus, it suffices to show that

$$\zeta_m(x,y) \equiv \frac{1}{\sqrt{m}}\sum_{k=1}^{m} Z_k(x,y) \Rightarrow \mathcal{G},$$

where $m = \lfloor T \rfloor$.

(1) *Finite-dimensional convergence in law*: First, for a fixed coordinate $(x, y) \in \mathbb{R}^2_+$, we note that $Z_k(x, y)$ is a sequence of bounded random variables. So, letting $\alpha(n)$ be the strong mixing coefficient of the sequence $(Z_k)_{k=1}^{\infty}$, we deduce that $\alpha(n) \leq \alpha_{n-1}$. Hence, by Assumption 4

$$\sum_{n=1}^{\infty} \alpha(n) \leq \sum_{t=0}^{\infty} \alpha_t \leq C \sum_{t=0}^{\infty} t^{-(1+\gamma)} < \infty.$$

Then, we can deduce by Theorem 18.5.4 of Ibragimov (1975) that

$$\zeta_m(x, y) \xrightarrow{d} N\left(0, \sigma^2(x, y)\right).$$

Here, the long-run variance is given by

$$\sigma^2(x, y) = 2 \int_0^{\infty} (H_t(x, y) - H(x, y)^2) \mathrm{d}t,$$

with $H_t(x, y) = \mathbb{P}(V_0^X \leq x, V_0^Y \leq y, V_t^X \leq x, V_t^Y \leq y)$.

More generally, for any finite collection $\{(x_i, y_i)\}_{i=1}^r$ with $(x_i, y_i) \in \mathbb{R}^2_+$ and any $v = (v_1, \ldots, v_r)^\top \in \mathbb{R}^r$, we aim to explain the behavior of the linear combination

$$L_m = \sum_{i=1}^{r} v_i \zeta_m(x_i, y_i) = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} D_k$$

where

$$D_k = \sum_{i=1}^{r} v_i Z_k(x_i, y_i).$$

Next, $\{D_k\}$ is again a bounded, strong mixing sequence with mixing coefficient

$$\sum_{h=1}^{\infty} \alpha_D(h) \leq r \sum_{h=1}^{\infty} \alpha(h) < \infty,$$

so by Theorem 18.5.4 in Ibragimov (1975), we once again conclude that

$$L_m \xrightarrow{d} N(0, \sigma_L^2),$$

with

$$\sigma_L^2 = \mathrm{var}(D_0) + 2 \sum_{h=1}^{\infty} \mathrm{cov}(D_0, D_h) = v^\top \Sigma_r v,$$

in which $\Sigma_r = (\Sigma_{ij})_{1 \leq i,j \leq r}$, and

$$\Sigma_{ij} = \lim_{m \to \infty} \mathrm{cov}\left(\zeta_m(x_i, y_i), \zeta_m(x_j, y_j)\right) = 2 \int_0^{\infty} \left(H_t(x_i, y_i, x_j, y_j) - H(x_i, y_i) H(x_j, y_j)\right) \mathrm{d}t,$$

since the covariance between $\zeta_m(x_i, y_i)$ and $\zeta_m(x_j, y_j)$ is given by

$$\mathrm{cov}\left(\zeta_m(x_i, y_i), \zeta_m(x_j, y_j)\right) = \frac{1}{m} \int_0^m \int_0^m \Big(H_{s,t}(x_i, y_i, x_j, y_j) - H(x_i, y_j) H(x_j, y_j)\Big) \mathrm{d}t \mathrm{d}s$$

$$
\begin{aligned}
&= \frac{1}{m}\int_0^m \int_0^m \big(H_{|t-s|}(x_i, y_i, x_j, y_j) - H(x_i, y_i)H(x_j, y_j)\big)\mathrm{d}t\mathrm{d}s \\
&= \frac{2}{m}\int_0^m \int_s^m \big(H_{t-s}(x_i, y_i, x_j, y_j) - H(x_i, y_i)H(x_j, y_j)\big)\mathrm{d}t\mathrm{d}s \\
&= 2\int_0^m \left(1 - \frac{t}{m}\right)\big(H_t(x_i, y_i, x_j, y_j) - H(x_i, y_i)H(x_j, y_j)\big)\mathrm{d}t.
\end{aligned}
$$

Here, $H_{s,t}(x_i, y_i, x_j, y_j) = \mathbb{P}\big(V_s^X \leq x_i, V_s^Y \leq y_i, V_t^X \leq x_j, V_t^Y \leq y_j\big)$, and the stationary condition in Assumption 4 is applied for the second equality in the display.

As v was arbitrary, the Cramér-Wold device implies that

$$
\big(\zeta_m(x_1, y_1), \ldots, \zeta_m(x_r, y_r)\big) \xrightarrow{d} N_r\big(\mathbf{0}, \Sigma_r\big).
$$

(2) *Tightness*: To verify tightness of the sequence of processes $\{\zeta_m(x, y)\}$, we extend the arguments presented in the proof of Theorem 5.2 from Dehay (2005). In particular, we deduce that

$$
\limsup_{m\to\infty} \mathbb{P}^*\left(\sup_{\rho((x_i,y_i),(x_j,y_j))<\delta} |\zeta_m(x_i, y_i) - \zeta_m(x_j, y_j)| > \eta\right) < \epsilon,
$$

for all $\eta > 0$ and $\epsilon > 0$. Here, $\rho\left((x_i, y_i), (x_j, y_j)\right) = |H(x_i, y_i) - H(x_j, y_j)|$ and $\mathbb{P}^*$ is the outer probability measure associated with $\mathbb{P}$. Hence, $G_T(x, y) \Rightarrow \mathcal{G}(x, y)$.

Next, we show that

$$
\sqrt{T}\left(C_{n,T}(u, v) - C(u, v)\right) \Rightarrow \mathcal{C}(u, v).
$$

Consider the mapping $\phi(H)(u, v) = H(Q^X(u), Q^Y(v))$. By Fermanian, Radulović, and Wegkamp (2004, Lemma 2), the mapping ϕ is Hadamard differentiable at the point $H^*(\cdot, \cdot\cdot) = H(Q^X(\cdot), Q^Y(\cdot\cdot))$. Moreover, the Hadamard derivative of ϕ at this point is given by

$$
\begin{aligned}
\phi'_{H^*}(h)(u, v) = h\big(Q^X(u), Q^Y(v)\big) &- \frac{\partial H}{\partial x}(Q^X(u), Q^Y(v)) \cdot \frac{h(Q^X(u), \infty)}{f(Q^X(u))} \\
&- \frac{\partial H}{\partial y}(Q^X(u), Q^Y(v)) \cdot \frac{h(\infty, Q^Y(v))}{g(Q^Y(v))}.
\end{aligned}
$$

By the previous result, namely $G_T \Rightarrow \mathcal{G}$, and the functional delta method, we obtain

$$
\sqrt{T}\left(C_T(x, y) - C(x, y)\right) \Rightarrow \phi'_{H^*}(\mathcal{G}),
$$

where the limit process can be written as

$$
\phi'_{H^*}(\mathcal{G}) = \mathcal{G}(Q^X(u), Q^Y(v)) - \frac{\partial C}{\partial u}(u, v)\mathcal{G}(Q^X(u), \infty) - \frac{\partial C}{\partial v}(u, v)\mathcal{G}(\infty, Q^Y(v)),
$$

with

$$
\begin{aligned}
\frac{\partial C}{\partial u}(u, v) &= \frac{\partial H(Q^X(u), Q^Y(v))}{\partial x} \cdot \frac{1}{f(Q^X(u))}, \\
\frac{\partial C}{\partial v}(u, v) &= \frac{\partial H(Q^X(u), Q^Y(v))}{\partial y} \cdot \frac{1}{g(Q^Y(v))}.
\end{aligned}
$$

In view of Lemma 2, the proof is complete. ■

PROOF OF COROLLARY 1: The proof builds on Theorem 4.2 in Dehay (2005) and Lemma 2 above. It is enough to prove that

$$\int_0^{T^\xi} \left\{ \widehat{H}_{t,n,T}(x,y,x',y') - H_{t,T}(x,y,x',y') \right\} \mathrm{d}t \xrightarrow{\mathbb{P}} 0,$$
$$\int_0^{T^\xi} \left\{ \widehat{H}_{n,T}(x,y)\widehat{H}_{n,T}(x',y') - H_T(x,y)H_T(x',y') \right\} \mathrm{d}t \xrightarrow{\mathbb{P}} 0,$$
$$2\int_0^{T^\xi} \{H_{t,T}(x,y,x',y') - H_T(x,y)H_T(x',y')\} \mathrm{d}t - C_{\mathcal{G}}(x,y,x',y') \xrightarrow{\mathbb{P}} 0.$$

Here, $H_{t,T}(x,y,x',y') = \frac{1}{T-T^\xi}\int_0^{T-T^\xi} 1_{\left\{V_s^X \le x, V_s^Y \le y, V_{s+t}^X \le x', V_{s+t}^Y \le y'\right\}} \mathrm{d}s$.

The convergence of the third term follows directly from the proof of Proposition 4.2 in Dehay (2005).

To establish convergence of the first term, we observe that

$$\left|\widehat{H}_{t,n,T}(x,y,x',y') - H_{t,T}(x,y,x',y')\right| \le \frac{1}{T-T^\xi}\int_0^{T-T^\xi} \left|\mathbb{1}_{\left\{\hat{V}_s^X \le x, \hat{V}_s^Y \le y\right\}} - \mathbb{1}_{\{V_s^X \le x, V_s^Y \le y\}}\right| \mathrm{d}s$$
$$+ \frac{1}{T-T^\xi}\int_t^{T-T^\xi+t} \left|\mathbb{1}_{\left\{\hat{V}_s^X \le x', \hat{V}_s^Y \le y'\right\}} - \mathbb{1}_{\{V_s^X \le x', V_s^Y \le y'\}}\right| \mathrm{d}s.$$

Here, we apply the inequality $|ab - cd| \le |a-c| + |b-d|$, for $a, b, c, d \in [0,1]$, and the change of variable $r = s + t$.

Exploiting the arguments presented in the proof for $\widehat{H}_{n,T}(x,y) - H_T(x,y)$ in Lemma 2, we can show that

$$\sup_{t\in[0,T^\xi]} \left|\widehat{H}_{t,n,T}(x,y,x',y') - H_{t,T}(x,y,x',y')\right| = o_P\left(\frac{1}{\sqrt{T}}\right).$$

Hence,

$$\left|\int_0^{T^\xi} \left\{ \widehat{H}_{t,n,T}(x,y,x',y') - H_{t,T}(x,y,x',y') \right\} \mathrm{d}t\right| = o_P\left(T^{\xi-\frac{1}{2}}\right).$$

Finally, we deal with the asymptotic negligibility of the second term. By Lemma 2, we get

$$\left|\int_0^{T^\xi} \left\{ \widehat{H}_{n,T}(x,y)\widehat{H}_{n,T}(x',y') - H_T(x,y)H_T(x',y') \right\} \mathrm{d}t\right|$$
$$\le T^\xi \left(\left|\widehat{H}_{n,T}(x,y) - H_T(x,y)\right| + \left|\widehat{H}_{n,T}(x',y') - H_T(x',y')\right| \right) = o_P\left(T^{\xi-\frac{1}{2}}\right).$$

Since $\xi \in (0, 1/3)$, the proof is complete. ■

PROOF OF COROLLARY 2: By an algebraic manipulation, we note that

$$|c_{n,T}(u,v) - c(u,v)| \le \frac{1}{Th^2}\int_0^T \left| K\left(\frac{\widehat{F}_{n,T}(\hat{V}_t^X) - u}{h}, \frac{\widehat{G}_{n,T}(\hat{V}_t^Y) - v}{h}\right) - K\left(\frac{F_T(V_t^X) - u}{h}, \frac{G_T(V_t^Y) - v}{h}\right)\right| \mathrm{d}t$$
$$+ \frac{1}{Th^2}\int_0^T \left| K\left(\frac{F_T(V_t^X) - u}{h}, \frac{G_T(V_t^Y) - v}{h}\right) - K\left(\frac{F(V_t^X) - u}{h}, \frac{G(V_t^Y) - v}{h}\right)\right| \mathrm{d}t$$

$$+\left|\frac{1}{Th^2}\int_0^T K\left(\frac{F(V_t^X)-u}{h},\frac{G(V_t^Y)-v}{h}\right)\mathrm{d}t - c(u,v)\right|.$$

First, we get by (A.2) that

$$\begin{aligned}
&\sup_{t\in[0,\infty)}\left|K\left(\frac{\widehat{F}_{n,T}(\hat{V}_t^X)-u}{h},\frac{\widehat{G}_{n,T}(\hat{V}_t^Y)-v}{h}\right)-K\left(\frac{F_T(V_t^X)-u}{h},\frac{G_T(V_t^Y)-v}{h}\right)\right| \\
&\leq \frac{C}{h}\left(\sup_{t\in[0,\infty)}\left|\widehat{F}_{n,T}(\hat{V}_t^X)-F_T(V_t^X)\right|+\sup_{t\in[0,\infty)}\left|\widehat{G}_{n,T}(\hat{V}_t^Y)-G_T(V_t^Y)\right|\right) \\
&= O_P\left(\frac{d_n}{h}\right).
\end{aligned} \tag{A.3}$$

Since $\sqrt{T}(F_T(x)-F(x))$ converges to a Brownian bridge by Theorem 2, we can readily deduce that $\sup_{x\in[0,\infty)}\sqrt{T}\big(F_T(x)-F(x)\big)=O_P(1)$. This yields

$$\begin{aligned}
&\sup_{t\in[0,\infty)}\left|K\left(\frac{F_T(V_t^X)-u}{h},\frac{G_T(V_t^Y)-v}{h}\right)-K\left(\frac{F(V_t^X)-u}{h},\frac{G(V_t^Y)-v}{h}\right)\right| \\
&\leq \frac{C}{h}\left(\sup_{x\in[0,\infty)}|F_T(x)-F(x)|+\sup_{y\in[0,\infty)}|G_T(y)-G(y)|\right) \\
&= O_P\left(\frac{1}{\sqrt{T}h}\right).
\end{aligned} \tag{A.4}$$

Moreover, as $\big(F(V_t^X),G(V_t^Y)\big)\in[0,1]^2$ has a density $c(u,v)$:

$$\begin{aligned}
\mathbb{E}\left[\frac{1}{h^2}K\left(\frac{F(V_t^X)-u}{h},\frac{G(V_t^Y)-v}{h}\right)\right] &= \frac{1}{h^2}\int_0^1\int_0^1 K\left(\frac{w-u}{h},\frac{z-v}{h}\right)c(w,z)\mathrm{d}w\mathrm{d}z \\
&= \int_0^1\int_0^1 K(p,q)\,c(u+ph,v+qh)\mathrm{d}p\mathrm{d}q \\
&= c(u,v)+O_P(h).
\end{aligned}$$

Hence, in the above meaning $\frac{1}{Th^2}\int_0^T K\left(\frac{F(V_t^X)-u}{h},\frac{G(V_t^Y)-v}{h}\right)\mathrm{d}t$ is an asymptotically unbiased estimator of $c(u,v)$. Following this line of reasoning, we also conclude that

$$\mathbb{E}\left[\frac{1}{h^2}K\left(\frac{F(V_t^X)-u}{h},\frac{G(V_t^Y)-v}{h}\right)\cdot\frac{1}{h^2}K\left(\frac{F(V_s^X)-u}{h},\frac{G(V_s^Y)-v}{h}\right)\right]=c_{|t-s|}(u,v,u,v)+O_P(h).$$

Here, $c_t(\cdot,\cdot,\cdot,\cdot)$ is the joint density function of $(F(V_0^X),G(V_0^Y),F(V_t^X),G(V_t^X))$. Therefore, we obtain

$$\begin{aligned}
\mathrm{var}\left[\frac{1}{Th^2}\int_0^T K\left(\frac{F(V_t^X)-u}{h},\frac{G(V_t^Y)-v}{h}\right)\mathrm{d}t\right] &= \frac{1}{T}\int_0^T\big(c_t(u,v,u,v)-c(u,v)^2\big)\,\mathrm{d}t+O_P(h) \\
&= O_P\left(\frac{1}{T}+h\right).
\end{aligned} \tag{A.5}$$

The latter is obtained by Assumption 4, which implies a Castellana-Leadbetter-type condition (see, for instance, Equation (2.1) of Dehay, 2005). In turn, this establishes that

$$\int_0^T \big(c_t(u,v,u,v) - c(u,v)^2\big)\,\mathrm{d}t = O_P(1).$$

Combining (A.3) – (A.5), we get

$$|c_{n,T}(u,v) - c(u,v)| = O_P\left(\frac{d_n}{h^3} \vee \frac{1}{\sqrt{T}h^3} \vee \left(\frac{1}{\sqrt{T}} + \sqrt{h}\right)\right).$$

The proof is complete by the chosen order of h and T (relative to Δ_n). ■

Proof of Corollary 3: First, Theorem 2 shows the the functional convergence

$$\sqrt{T}\big(\widehat{C}_{n,T}(u,v) - C(u,v)\big) \Rightarrow \mathcal{C} \quad \text{in} \quad \big(\mathbb{D}([0,1]^2), \|\cdot\|_\infty\big),$$

where $\|f\|_\infty = \sup_{(u,v)\in[0,1]^2} |f(u,v)|$ and $\mathcal{C}$ is the centered Gaussian process defined in (4.12) with almost surely continuous sample paths on $[0,1]^2$.

Now, consider the mapping $\Psi\colon \mathbb{D}([0,1]^2) \longrightarrow \mathbb{R}$:

$$\Psi(f) = \sup_{(u,v)\in[0,1]^2} |f(u,v)|.$$

For any $f, g \in \mathbb{D}([0,1]^2)$,

$$\big|\Psi(f) - \Psi(g)\big| = \Big|\sup_{(u,v)} |f(u,v)| - \sup_{(u,v)} |g(u,v)|\Big| \leq \sup_{(u,v)} |f(u,v) - g(u,v)| = \|f-g\|_\infty.$$

Hence, Ψ is Lipschitz continuous in this metrized space.

Since $\sqrt{T}(\widehat{C}_{n,T}(u,v) - C(u,v)) \Rightarrow \mathcal{C}$ in $\big(\mathbb{D}([0,1]^2), \|\cdot\|_\infty\big)$ and Ψ is continuous on this space, the continuous mapping theorem yields

$$\Psi\big(\sqrt{T}(\widehat{C}_{n,T}(u,v) - C(u,v))\big) \Rightarrow \Psi(\mathcal{C}).$$

That is,

$$\sup_{(u,v)\in[0,1]^2} \sqrt{T}\big(\widehat{C}_{n,T}(u,v) - C(u,v)\big) \xrightarrow{d} \sup_{(u,v)\in[0,1]^2} \mathcal{C}(u,v),$$

as $\Delta_n \to 0$ and $T \to \infty$. This establishes (4.30). ■

Proof of Proposition 1: We divide the proof in three parts.

(1) *Finite-dimensional Gaussian approximation*: We start by fixing a finite grid $\{(u_i, v_j) : 1 \leq i,j \leq m\} \subset [0,1]^2$ and define

$$Z_{n,T}(u_i,v_j) = \sqrt{T}\big(\widehat{C}_{n,T}(u_i,v_j) - C(u_i,v_j)\big).$$

together with the corresponding m^2-dimensional random vector

$$Z_{n,T}^{(m)} = \big(Z_{n,T}(u_i,v_j)\big)_{1\leq i,j\leq m}.$$

By Theorem 2 and the continuous mapping theorem, as $\Delta_n \to 0$ and $T \to \infty$,

$$Z_{n,T}^{(m)} \Rightarrow Z^{(m)},$$

where $Z^{(m)} = \big(\mathcal{C}(u_i, v_j)\big)_{1\le i,j\le m}$ is a centered Gaussian vector with covariance matrix

$$\Sigma_m = \big\{\mathrm{avar}_{\mathcal{C}}(u_i, v_j, u_k, v_l)\big\}_{1\le i,j,k,l\le m}.$$

By Corollary 1,

$$\widehat{\Sigma}_m = \Big\{\widehat{\mathrm{avar}}_{\mathcal{C}}(u_i, v_j, u_k, v_l)\Big\}_{1\le i,j,k,l\le m} \xrightarrow{\mathbb{P}} \Sigma_m.$$

Conditional on $\widehat{C}_{\mathcal{C}}$, we generate

$$\widehat{Z}_{n,T}^{(m)} = \big(\widehat{\mathcal{C}}_{n,T}(u_i, v_j)\big)_{1\le i,j\le m} \sim N\big(0, \widehat{\Sigma}_m\big).$$

Since $\widehat{\Sigma}_m \xrightarrow{\mathbb{P}} \Sigma_m$, the characteristic functions of $N(0, \widehat{\Sigma}_m)$ converge pointwise to those of $N(0, \Sigma_m)$. Hence, by Lévy's continuity theorem,

$$\mathcal{L}\big(\widehat{Z}_{n,T}^{(m)} \mid \widehat{C}_{\mathcal{C}}\big) = \mathcal{L}\big(N(0, \widehat{\Sigma}_m)\big) \xrightarrow{\mathbb{P}} \mathcal{L}\big(N(0, \Sigma_m)\big) = \mathcal{L}\big(Z^{(m)}\big),$$

In words, the conditional law of $\widehat{Z}_{n,T}^{(m)}$ converges in probability to that of $Z^{(m)}$.

Let $\phi_m : \mathbb{R}^{m^2} \to \mathbb{R}$ be the functional $\phi_m(z) = \max_{1\le i,j\le m} |z_{ij}|$. By the continuous mapping theorem, for any continuity point $c > 0$ of the distribution function of $\phi_m(Z^{(m)})$,

$$\mathbb{P}\Big(\phi_m(\widehat{Z}_{n,T}^{(m)}) \le c \mid \widehat{C}_{\mathcal{C}}\Big) \xrightarrow{\mathbb{P}} \mathbb{P}\big(\phi_m(Z^{(m)}) \le c\big).$$

Or, equivalently,

$$F_{n,T}^{(m)}(c \mid \widehat{C}_{\mathcal{C}}) \xrightarrow{\mathbb{P}} F_m(c),$$

where

$$F_{n,T}^{(m)}(c \mid \widehat{C}_{\mathcal{C}}) = \mathbb{P}\left(\sup_{1\le i,j\le m} \big|\widehat{\mathcal{C}}_{n,T}(u_i, v_j)\big| \le c \mid \widehat{C}_{\mathcal{C}}\right) \quad \text{and} \quad F_m(c) = \mathbb{P}\left(\sup_{1\le i,j\le m} |\mathcal{C}(u_i, v_j)| \le c\right).$$

(2) *Quantile consistency*: We now let $m \to \infty$ such that the grid becomes dense. Then, since $\mathcal{C}$ has continuous sample paths on $[0,1]^2$, almost surely, it follows that

$$\sup_{1\le i,j\le m} |\mathcal{C}(u_i, v_j)| \to \sup_{(u,v)\in[0,1]^2} |\mathcal{C}(u, v)|,$$

almost surely, so the sequence of distribution functions $(F_m)_{m=1}^{\infty}$ converge pointwise to

$$\lim_{m\to\infty} F_m(c) = F(c) = \mathbb{P}\left(\sup_{(u,v)\in[0,1]^2} |\mathcal{C}(u, v)| \le c\right),$$

at every continuity point c of F. Hence, for any fixed compact interval $[0, M]$ and $\varepsilon > 0$, we can first choose m large enough so that

$$\sup_{c\in[0,M]} |F_m(c) - F(c)| < \varepsilon,$$

and then choose (n,T) sufficiently large so that for every $\varepsilon,\eta>0$,

$$\mathbb{P}\left(\sup_{c\in[0,M]}\left|F_{n,T}^{(m)}(c\mid\widehat{C}_{\mathcal{C}})-F_m(c)\right|>\varepsilon\right)<\eta.$$

Combining the two inequalities yields

$$\sup_{c\in[0,M]}\left|F_{n,T}^{(m)}(c\mid\widehat{C}_{\mathcal{C}})-F(c)\right|\xrightarrow{\mathbb{P}}0.$$

Hence, $F_{n,T}^{(m)}(\cdot\mid\widehat{C}_{\mathcal{C}})\xrightarrow{\mathbb{P}}F(\cdot)$ uniformly on compact subsets of $(0,\infty)$.

To complete part (b), let

$$q_{1-\alpha}^{CB}=\inf\{c>0:F(c)\geq1-\alpha\}\quad\text{and}\quad\widehat{q}_{1-\alpha}^{CB}=\inf\{c>0:F_{n,T}^{(m)}(c\mid\widehat{C}_{\mathcal{C}})\geq1-\alpha\}$$

denote the unique $(1-\alpha)$-quantile of $F(\cdot)$ and $F_{n,T}^{(m)}(\cdot\mid\widehat{C}_{\mathcal{C}})$, respectively. The uniform convergence of $F_{n,T}$ to F and continuity of F at $q_{1-\alpha}^{CB}$ imply that

$$\widehat{q}_{1-\alpha}^{CB}\xrightarrow{\mathbb{P}}q_{1-\alpha}^{CB},$$

see van der Vaart (1998, Lemma 21.2).

(3) *Asymptotic coverage of the confidence band*: In the final part, we set

$$\mathcal{X}_{n,T}=\sup_{(u,v)\in[0,1]^2}\sqrt{T}\big(\widehat{C}_{n,T}(u,v)-C(u,v)\big)\quad\text{and}\quad\mathcal{X}=\sup_{(u,v)\in[0,1]^2}\mathcal{C}(u,v).$$

By Corollary 3, $\mathcal{X}_{n,T}\Rightarrow\mathcal{X}$ as $\Delta_n\to0$ and $T\to\infty$, and furthermore $\widehat{q}_{1-\alpha}^{CB}\xrightarrow{\mathbb{P}}q_{1-\alpha}^{CB}$ with $\mathbb{P}(\mathcal{X}X\leq q_{1-\alpha}^{CB})=1-\alpha$. An application of Slutsky's lemma then delivers

$$\mathbb{P}\left(\mathcal{X}_{n,T}\leq\widehat{q}_{1-\alpha}^{CB}\right)\to\mathbb{P}\left(\mathcal{X}\leq q_{1-\alpha}^{CB}\right)=1-\alpha.$$

Equivalently,

$$\mathbb{P}\left(\sup_{(u,v)\in[0,1]^2}\left|\widehat{C}_{n,T}(u,v)-C(u,v)\right|\leq\widehat{q}_{1-\alpha}^{CB}/\sqrt{T}\right)\to1-\alpha,$$

which shows that the uniform confidence band

$$\left\{C:\sup_{(u,v)\in[0,1]^2}\left|\widehat{C}_{n,T}(u,v)-C(u,v)\right|\leq\widehat{q}_{1-\alpha}^{CB}/\sqrt{T}\right\}$$

has asymptotic coverage $1-\alpha$. ■

Proof of Corollary 4: Let

$$G_{n,T}(u,v)=\sqrt{T}\big(\widehat{C}_{n,T}(u,v)-C_0(u,v)\big)\qquad(u,v)\in[0,1]^2,$$

be the empirical process associated with $\mathcal{H}_0:C=C_0$. By Theorem 2,

$$G_{n,T}\Rightarrow\mathcal{C}\quad\text{in}\quad\big(\mathbb{D}([0,1]^2),\|\cdot\|_\infty\big),$$

under $\mathcal{H}_0$, where $\mathcal{C}$ is the centered Gaussian process from Corollary 3, which has almost surely continuous, and bounded, sample paths on $[0,1]^2$.

Now, consider the functional $\Phi : \mathbb{D}([0,1]^2) \to \mathbb{R}_+$ defined as

$$\Phi(f) = \int_0^1 \int_0^1 f^2(u,v)\mathrm{d}u\mathrm{d}v.$$

For any sequence $(f_n)_{n=1}^{\infty}$ such that $f_n \to f$ in $\big(\mathbb{D}([0,1]^2), \|\cdot\|_\infty\big)$:

$$\sup_{(u,v)\in[0,1]^2} \big|f_n(u,v)^2 - f(u,v)^2\big| \leq 2\bigg(\sup_{(u,v)} |f_n(u,v)| + \sup_{(u,v)} |f(u,v)|\bigg) \sup_{(u,v)} |f_n(u,v) - f(u,v)|.$$

Since $\|f_n\|_\infty \to \|f\|_\infty < \infty$, the right-hand side converges to zero, and so $f_n^2 \to f^2$ uniformly on $[0,1]^2$. It thus follows that

$$\lim_{n\to\infty} \Phi(f_n) = \lim_{n\to\infty} \int_0^1 \int_0^1 f_n^2(u,v)\mathrm{d}u\mathrm{d}v = \int_0^1 \int_0^1 f^2(u,v)\mathrm{d}u\mathrm{d}v = \Phi(f),$$

showing that Φ is continuous on $\big(\mathbb{D}([0,1]^2), \|\cdot\|_\infty\big)$.

So, by the continuous mapping theorem

$$\Phi\big(G_{n,T}\big) = \int_0^1 \int_0^1 G_{n,T}^2(u,v)\mathrm{d}u\mathrm{d}v \xrightarrow{d} \int_0^1 \int_0^1 \mathcal{C}^2(u,v)\mathrm{d}u\mathrm{d}v,$$

as $\Delta_n \to 0$ and $T \to \infty$. This establishes (4.42). ■

References

Andersen, T. G., and T. Bollerslev, 1998, "Answering the skeptics: Yes, standard volatility models do provide accurate forecasts," *International Economic Review*, 39, 885–905.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens, 2001, "The distribution of realized stock return volatility," *Journal of Financial Economics*, 61, 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys, 2003, "Modeling and forecasting realized volatility," *Econometrica*, 71, 579–625.

Andersen, T. G., T. Su, V. Todorov, and Z. Zhang, 2024, "Intraday periodic volatility curves," *Journal of the American Statistical Association*, 119, 1181–1191.

Andersen, T. G., M. Thyrsgaard, and V. Todorov, 2019, "Time-varying periodicity in intraday volatility," *Journal of the American Statistical Association*, 114, 1695–1707.

Barndorff-Nielsen, O. E., and N. Shephard, 2002, "Econometric analysis of realized volatility and its use in estimating stochastic volatility models," *Journal of the Royal Statistical Society: Series B*, 64, 253–280.

Bibinger, M., and L. Winkelmann, 2018, "Common price and volatility jumps in noisy high-frequency data," *Electronic Journal of Statistics*, 12, 2018–2073.

Billio, M., M. Getmansky, A. W. Lo, and L. Pelizzon, 2012, "Econometric measures of connectedness and systemic risk in the finance and insurance sectors," *Journal of Financial Economics*, 104, 535–559.

Black, F., 1976, "Studies of stock market volatility changes," in *Proceedings of the American Statistical Association*, pp. 177–181.

Bollerslev, T., and V. Todorov, 2011, "Tails, fears and risk premia," *Journal of Finance*, 66, 2165–2211.

Carr, P., and L. Wu, 2009, "Variance Risk Premiums," *Review of Financial Studies*, 22, 1311–1341.

Chong, C. H., and V. Todorov, 2023, "The fine structure of volatility dynamics," Working paper.

Christensen, K., R. C. A. Oomen, and R. Renò, 2022, "The drift burst hypothesis," *Journal of Econometrics*, 227, 461–497.

Christensen, K., M. Thyrsgaard, and B. Veliyev, 2019, "The realized empirical distribution function of stochastic variance with application to goodness-of-fit testing," *Journal of Econometrics*, 212, 556–583.

Christie, A. A., 1982, "The stochastic behavior of common stock variances: Value, leverage and interest rate effects," *Journal of Financial Economics*, 10, 407–432.

Comte, F., and E. Renault, 1998, "Long memory in continuous-time stochastic volatility models," *Mathematical Finance*, 8, 291–323.

Dehay, D., 2005, "On invariant distribution function estimation for continuous-time stationary processes," *Bernoulli*, 11, 933–948.

Deheuvels, P., 1979, "La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance," *Académie Royale de Belgique. Bulletin de la Classe des Sciences*, 65, 274–292.

Delbaen, F., and W. Schachermayer, 1994, "A general version of the fundamental theorem of asset pricing," *Mathematische Annalen*, 300, 463–520.

Diebold, F. X., and K. Yilmaz, 2009, "Measuring financial asset return and volatility spillovers, with application to global equity markets," *Economic Journal*, 119, 158–171.

Embrects, P., A. J. McNeil, and D. Straumann, 1999, "Correlation and dependence in risk management: Properties and pitfalls," Working paper.

Fermanian, J.-D., D. Radulović, and M. Wegkamp, 2004, "Weak convergence of empirical copula processes," *Bernoulli*, 10, 847–860.

Fermanian, J.-D., and O. Scaillet, 2002, "Nonparametric estimation of copulas for time series," *Journal of Risk*, 5, 25–54.

Gatheral, J., T. Jaisson, and M. Rosenbaum, 2018, "Volatility is rough," *Quantitative Finance*, 18, 933–949.

Geman, D., and J. Horowitz, 1980, "Occupation densities," *Annals of Probability*, 8, 1–67.

Genest, C., and L.-P. Rivest, 1993, "Statistical inference procedures for bivariate Archimedean copulas," *Journal of the American Statistical Association*, 88, 1034–1043.

Hansen, P. R., and A. Lunde, 2006, "Realized variance and market microstructure noise," *Journal of Business and Economic Statistics*, 24, 127–161.

Ibragimov, I., 1975, *Independent and stationary sequences of random variables*. Wolters-Noordhoff, Groningen, 1st edn.

Jacod, J., and P. E. Protter, 2012, *Discretization of Processes*. Springer, Berlin, 2nd edn.

Jacod, J., and M. Rosenbaum, 2013, "Quarticity and other functionals of volatility: Efficient estimation," *Annals of Statistics*, 41, 1462–1484.

Joe, H., 1993, "Parametric families of multivariate distributions with given margins," *Journal of Multivariate Analysis*, 46, 262–282.

Li, D. X., 2000, "On default correlation: A copula function approach," *Journal of Fixed Income*, 9, 43–54.

Li, J., V. Todorov, and G. Tauchen, 2013, "Volatility occupation times," *Annals of Statistics*, 41, 1865–1891.

——— , 2016a, "Estimating the volatility occupation time via regularized Laplace inversion," *Econometric Theory*, 32, 1253–1288.

——— , 2016b, "Inference theory for volatility functional dependencies," *Journal of Econometrics*, 193, 17–34.

Mancini, C., 2009, "Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps," *Scandinavian Journal of Statistics*, 36, 270–296.

Newey, W. K., and K. D. West, 1994, "Automatic lag selection in covariance matrix estimation," *Review of Economic Studies*, 61, 631–653.

Patton, A. J., 2012, "A review of copula models for economic time series," *Journal of Multivariate Analysis*, 110, 4–18.

Rosenblatt, M., 1956a, "A central limit theorem and a strong mixing condition," *Proceedings of the National Academy of Sciences of the United States of America*, 42, 43–47.

——— , 1956b, "Remarks on some nonparametric estimates of a density function," *Annals of Mathematical Statistics*, 27, 832–837.

Salmon, F., 2009, "Recipe for disaster: The formula that killed Wall Street," Accessed at: `https://www.wired.com/2009/02/wp-quant/`.

Samuelson, P. A., 1965, "Proof that properly anticipated prices fluctuate randomly," *Industrial Management Review*, 6, 41–49.

Schweizer, B., and E. F. Wolff, 1981, "On nonparametric measures of dependence for random variables," *Annals of Statistics*, 9, 879–885.

Sibuya, M., 1960, "Bivariate extreme statistics, I," *Annals of the Institute of Statistical Mathematics*, 11, 195–210.

Sklar, A., 1959, "Fonctions de répartition à n dimensions et leurs marges," *Publications de l'Institut de Statistique de l'Université de Paris*, 8, 229–231.

Todorov, V., and G. Tauchen, 2011, "Volatility jumps," *Journal of Business and Economic Statistics*, 29, 356–371.

van der Vaart, A. W., 1998, *Asymptotic Statistics*. Cambridge University Press, Cambridge, 1st edn.

Zhang, C., J. Li, and T. Bollerslev, 2022, "Occupation density estimation for noisy high-frequency data," *Journal of Econometrics*, 227, 189–211.