ASYMMETRIC TAIL DEPENDENCE MODELING, WITH APPLICATION TO CRYPTOCURRENCY MARKET DATA

BY YAN GONG[a] AND RAPHAËL HUSER[b]

*Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division,
King Abdullah University of Science and Technology (KAUST),* [a]*yan.gong@kaust.edu.sa,* [b]*raphael.huser@kaust.edu.sa*

Since the inception of Bitcoin in 2008, cryptocurrencies have played an increasing role in the world of e-commerce, but the recent turbulence in the cryptocurrency market in 2018 has raised some concerns about their stability and associated risks. For investors it is crucial to uncover the dependence relationships between cryptocurrencies for a more resilient portfolio diversification. Moreover, the stochastic behavior in both tails is important, as long positions are sensitive to a decrease in prices (lower tail), while short positions are sensitive to an increase in prices (upper tail). In order to assess both risk types, we develop in this paper a flexible copula model which is able to distinctively capture asymptotic dependence or independence in its lower and upper tails simultaneously. Our proposed model is parsimonious and smoothly bridges (in each tail) both extremal dependence classes in the interior of the parameter space. Inference is performed using a full or censored likelihood approach, and we investigate by simulation the estimators' efficiency under three different censoring schemes which reduce the impact of nonextreme observations. We also develop a local likelihood approach to capture the temporal dynamics of extremal dependence among pairs of leading cryptocurrencies. We here apply our model to historical closing prices of five leading cryotocurrencies which share large cryptocurrency market capitalizations. The results show that our proposed copula model outperforms alternative copula models and that the lower-tail dependence level between most pairs of leading cryptocurrencies and, in particular, Bitcoin and Ethereum has become stronger over time, smoothly transitioning from an asymptotic independence regime to an asymptotic dependence regime in recent years, whilst the upper tail has been relatively more stable overall at a weaker dependence level.

1. Introduction. Because of the confidentiality, integrity, and speed of transactions of virtual operations, the use of cryptocurrencies among private users and businesses has increased at a fast rate since Bitcoin was initially created about a decade ago, and the transaction volume has grown considerably. However, cryptocurrencies have also been subject to major shocks and a highly volatile "roller coaster" in recent years which has triggered important concerns about the stability and the risks associated with cryptomarkets. One of the critical periods in the history of cryptomarkets is the 2017 boom, which was then followed by a sudden major crash in 2018, where the Bitcoin price dropped from about 18,000 U.S. dollars (USD) in December 2017 to only 3500 USD in December, 2018. Later, due to the worldwide COVID-19 pandemic and a fear of a global recession, the year 2020 challenged the traditional financial system, leading to a broader adoption of cryptocurrencies by various reputable financial institutions. These conditions combined with the major investment of 1.5 billion USD by Tesla in cryptocurrencies, boosted cryptomarkets with Bitcoin reaching an all-time high capitalization of more than 61,000 USD on March 12, 2021. But then in May, 2021, the Bitcoin price lost about half of its value, dropping quickly to about 30,000 USD,

following suggestions that Tesla has sold or will sell its Bitcoin holdings and new regulations from the Chinese government to support their cryptocurrency crackdown. Overall, cryptocurrency transactions are being increasingly scrutinized, are either adopted (even by countries, such as El Salvador and Cuba, who were the first ones to consider Bitcoin as a legal tender in June and August, 2021, respectively), and/or regulated, and in some cases banned (e.g., China declared all cryptocurrency transactions illegal in September 2021). As of October, 2021, the Bitcoin price has recovered to about 50,000 USD. Nevertheless, cryptcurrencies remain highly volatile and their systemic risks mostly unknown.

For decades, the statistical modeling of extreme events has played a fundamental role in a wide range of financial risk assessment studies (see, e.g., Castro-Camilo, de Carvalho and Wadsworth (2018), Embrechts, Klüppelberg and Mikosch (1997), Poon, Rockinger and Tawn (2003, 2004)), and Borri (2019) has recently shown that some leading cryptocurrencies are indeed highly exposed to tail risk within cryptomarkets; see also de Carvalho, Leonelli and Rossi (2020), Feng, Wang and Zhang (2018), Gkillas, Bekiros and Siriopoulos (2018), Nguyen, Chevapatrakul and Yao (2020) for other papers studying tail risk in cryptocurrencies. In particular, Feng, Wang and Zhang (2018) have shown that the lower and upper tail dependence structures among cryptocurrencies are asymmetric and have found that the dependence strength has increased after August 2016, suggesting high and growing systematic extreme risks. Apart from these recent contributions, the tail dependence relationships among the different cryptocurrencies, representing large simultaneous gains and losses, is still largely unexplored. For investors the behavior in both tails is important, as long positions are sensitive to a decrease in prices (lower tail), while short positions are sensitive to an increase in prices (upper tail). Huynh, Nguyen and Duong (2018) also pointed out that contagion risk among cryptocurrency returns exists and portfolio diversification is required for investors. Furthermore, asymmetric tail dependence structures are widely observed and studied in the modeling of financial assets (Alcock and Satchell (2018), Patton (2004, 2006)), yet the dependence structures commonly fitted to real data often tend to lack flexibility in the tails which is crucial with regard to risk management and mitigation.

In order to assess such risks, theoretically justified models that are resilient for extrapolating joint tail probabilities to the most extreme levels are needed, and extreme value theory (EVT) provides a natural theoretical framework; see Davison and Huser (2015) for a review on statistics of extremes. In the multivariate framework the two most prominent classes of asymptotic models in the extreme-value literature are max-stable distributions (Castruccio, Huser and Genton (2016), Padoan, Ribatet and Sisson (2010), Tawn (1988, 1990), Vettori, Huser and Genton (2018)) and multivariate Pareto distributions (Kiriliouk et al. (2019), Rootzén, Segers and Wadsworth (2018), Rootzén and Tajvidi (2006)). While the former are designed to model block maxima, the latter are used for high threshold exceedances. To use them in practice, we first need to choose a finite block size (or threshold), and we then keep only the block maxima (or observations exceeding the threshold) for fitting. While this modeling approach has solid theoretical foundations based on asymptotic arguments, it leads, in practice, to a large loss of information (by discarding all nonextreme data). Moreover, it adds the difficulty of choosing an appropriate block size (or threshold), which is especially tricky for nonstationary or heteroscedastic time series (Scarrott and MacDonald (2012)), and it does not provide any information about the bulk of the distribution. By contrast, in this paper we seek to develop a single flexible multivariate dependence model for the entire dataset that possesses high flexibility in both the lower and the upper tails while keeping a smooth transition between the two.

Essentially, two asymptotic regimes can prevail in each tail, namely, *asymptotic dependence* (AD) or *asymptotic independence* (AI). Mathematically, let $X = (X_1, X_2)^\top \sim F_X$ be a random vector with margins F_{X_1}, F_{X_2} assumed to be continuous for simplicity and define the

uniform random variables $U_1 = F_{X_1}(X_1), U_2 = F_{X_2}(X_2) \sim \text{Unif}(0, 1)$ such that the vector $U = (U_1, U_2)^\top$ follows the joint distribution

$$(1.1) \qquad C(u_1, u_2) = F_X\{F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2)\},$$

called the *copula* of X. It is unique when the marginal distributions F_{X_1}, F_{X_2} are continuous. Then, X is said to be AD in the upper tail if

$$(1.2) \qquad \chi_U = \lim_{t \to 1} \Pr(U_1 > t \mid U_2 > t) = \lim_{t \to 1} \frac{1 - 2t + C(t, t)}{1 - t} > 0,$$

whereas it is AI if $\chi_U = 0$, implicitly assuming that the limit exists. An analogous, symmetric definition holds for the lower tail; see Section 3.3. Loosely speaking, AI implies that the dependence strength weakens and eventually vanishes as events become more extreme, whereas AD means that it eventually stabilizes to some positive level. In practice, this distinction is key, as it determines the risk that future unprecedented extreme events might occur simultaneously. Under AD there is a positive probability that extreme events occur together, no matter how extreme they are, while under AI, this probability is zero for the most extreme events (i.e., in the limit). Max-stable distributions are always AD in the upper tail and AI in the lower tail (Ledford and Tawn (1996)). Therefore, they are unsuitable for the modeling of a wide range of processes with weakening upper-tail dependence or strong lower-tail dependence. Alternatively, various types of dependence structures may be used. The Gaussian copula is the most widely-used dependence model, but it is tail-symmetric, AI in both tails, and possesses a rigid-tail structure. The Student-t copula, which stems from a specific Gaussian scale mixture and generalizes the Gaussian copula, is also tail-symmetric and is AD in both tails. The tail properties of Gaussian scale mixtures and other types of elliptical models have been explored in depth among others by Hashorva (2010), Huser, Opitz and Thibaud (2017) and Engelke, Opitz and Wadsworth (2019) (see also the references therein). In particular, Huser, Opitz and Thibaud (2017) proposed a specific copula model that has a smooth transition between AD and AI on the boundary of the parameter space, but it remains tail-symmetric. In the same vein, exploiting various types of random scale constructions, Wadsworth et al. (2017) and Huser and Wadsworth (2019) proposed flexible bivariate and spatial copula models that can capture AI and AD in the upper tail only with the transition in the interior of the parameter space. Another related paper is Krupskii, Huser and Genton (2018) who studied the tail properties of a certain class of factor copula models; see also Zhang et al. (2021) for a related normal mean-variance mixture model capturing AI in both tails.

Building upon and extending the recent work of Huser and Wadsworth (2019) who proposed a spatial extreme model for the upper tail only, we here develop in this paper a new parsimonious copula model that is able to distinctively control the AD/AI regime in both the lower and upper tails. Our proposed model has a small number of parameters, and yet, it can capture a wide variety of dependence structures ranging from independence to complete dependence, while including nontrivial AD and AI cases characterized by slow and rapid joint-tail decay rates, respectively. Moreover, the transition between AD and AI takes place in the interior of the parameter space (for each tail) which greatly facilitates inference on the extremal dependence class. Unlike classical asymptotic extreme-value models, our model possesses high flexibility at subasymptotic levels, and so it can also be used to model the full dataset while still capturing the lower- and upper-tail behaviors accurately. We further extend the model to a skewed version which enjoys even more flexibility. To make inference, we propose and compare a full likelihood and various censored likelihood approaches, exploring three different censoring schemes that are specifically designed to prioritize model calibration in the lower and upper-tail regions while downweighting the contribution of nonextreme observations in the bulk. In particular, our full-likelihood approach is similar to other papers that

study extremes and assess the joint-tail risk using the complete dataset (Aulbach, Bayer and Falk (2012), Hazra, Reich and Staicu (2020), Leonelli and Gamerman (2020), Vrac, Naveau and Drobinski (2007)). Furthermore, we also develop a (weighted) local likelihood approach that can capture complex time-varying dependence behaviors, to uncover how the extremal dependence among any two leading cryptocurrencies has evolved over time.

The paper is organized as follows. In Section 2 we present the dataset, namely the historical closing prices of five leading cryptocurrencies, which currently share most of the cryptocurrency market capitalizations, and we discuss some basic statistical preprocessing. In Section 3 we detail the construction of our new copula model, and we formally derive its tail dependence properties. In Section 4 we describe (global and local) likelihood-based inference, using either the full likelihood or various censored likelihoods prioritizing calibration in the tails, and conduct an extensive simulation study. In Section 5 we apply our methodology to cryptocurrency data in order to uncover their complex time-varying extremal dependence structures in both tails. We finally conclude in Section 6 with some discussion and perspectives for future research.

2. Cryptocurrency market data and preprocessing.

Unlike traditional currencies, a cryptocurrency is a digital currency that is not emitted by a central bank, nor supported financially by the national currency. Being decentralized, a cryptocurrency is not as much affected by political decisions nor any other intermediates, and it uses cryptographic algorithms to secure financial activities and safeguard the confidentiality of transactions. For these attractive reasons the use of cryptocurrencies has grown considerably over the last decade. Bitcoin (BTC), which was initially created by Satoshi Nakamoto (Nakamoto (2008)) in 2008 and released in 2009, was the first cryptocurrency. Nowadays, there are thousands of different cryptocurrencies available in the market (see https://coinmarketcap.com/all/views/all/). Figure 1 shows the historical daily adjusted closing prices and log returns of five leading cryptocurrencies from December 31, 2015, to October 5, 2021. The values represent the relative prices with respect to USD downloaded from Yahoo Finance on October 6, 2021. The cryptocurrencies represented in Figure 1 are Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC), Monero (XMR) which occupied a large portion of the market capitalizations.

We split the entire dataset into two parts: a training set (December 31, 2015, to April 29, 2020), used for model fitting in our data analysis, and a validation set (April 30, 2020, to October 5, 2021), used to assess the out-of-sample performance of our proposed model. We chose the validation period to start on April 30, 2020, that is, just before the second booming period in the cryptocurrency history characterized by major upward and downward shocks. This allows us to assess whether our proposed model is resilient enough to accurately predict joint-tail risk in a challenging, highly volatile situation.

Although cryptocurrencies are generally believed to behave differently from traditional currencies, the log returns present similar characteristics, such as high volatility clusters and heavy tails. In order to extract stationary residuals (from which we will estimate joint-tail probabilities of simultaneous extremes), we then filter the log returns by fitting a time-varying ARMA(1, 1)–GARCH(1, 1) model to each time series separately with a moving window approach; see Brockwell and Davis (2002) and the Supplementary Material (Gong and Huser (2022a)) for details on time series models. The chosen marginal model was found to perform well and to be the best one after some experimentation and model selection procedure based on the Bayesian information criterion.

Figure 2 displays bivariate scatterplots of standardized residuals for all pairs of cryptocurrencies during the training period. In the Supplementary Material (Gong and Huser (2022a)) we further report the same plots on standard uniform margins, obtained after transforming the residuals using the empirical probability integral transform based on ranks. From a quick
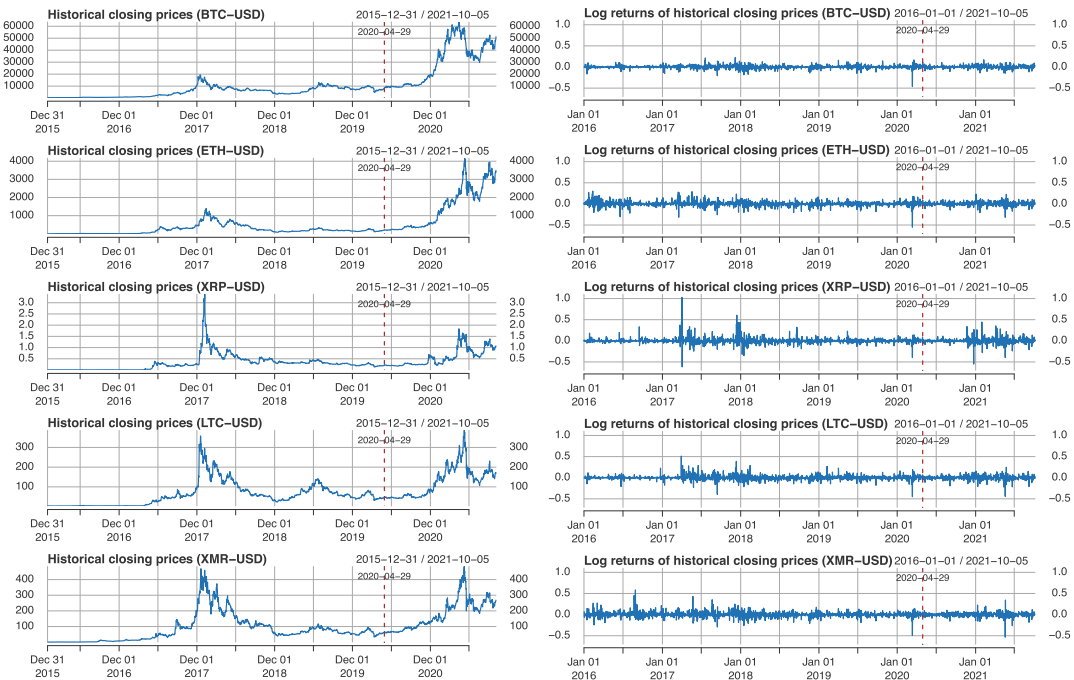
FIG. 1. *Historical daily adjusted closing prices* (*left*) *and log returns* (*right*) *of five leading cryptocurrencies* (*from top to bottom*: *BTC, ETH, XRP, LTC, and XMR*) *for the period from December* 31, 2015, *to October* 5, 2021. *The values represent the relative prices with respect to USD* (*downloaded on October* 6, 2021, *from Yahoo Finance*). *Training and validation sets are separated by the vertical dashed line.*

glimpse the overall correlations between pairs of cryptocurrencies appear to be rather weak in the bulk, while the dependence strength seems to be stronger in the lower tail than the upper tail. Moreover, from the color of the points (indicative of time), more recent observations also seem more strongly tail-dependent, as demonstrated by dark points being more concentrated around the diagonal line. To further investigate whether the lower-tail dependence structure between each pair of cryptocurrencies varies over time, we estimate the (symmetric version of) the tail-dependence coefficient (1.2) using a simple nonparametric moving-window estimator with window size chosen to provide a reasonable bias-variance trade-off. Figure 3
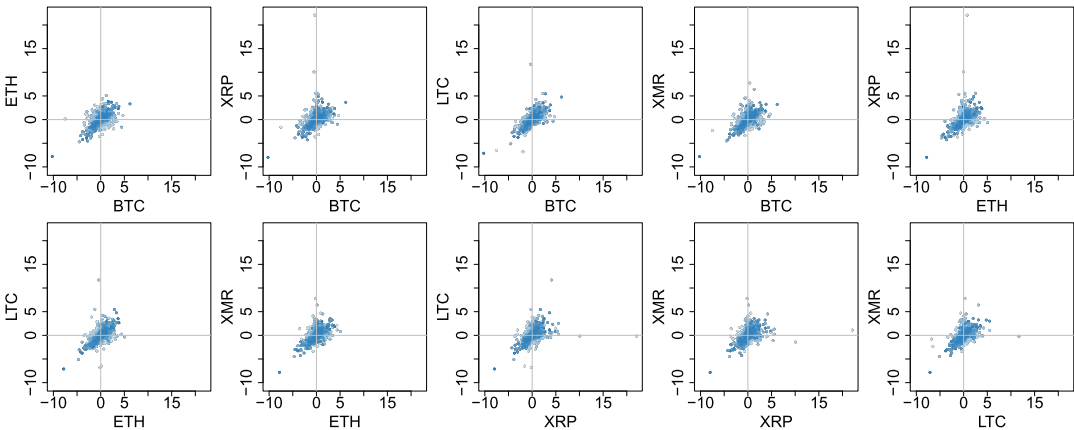

FIG. 2. *Bivariate scatterplots of standardized residuals extracted from fitting a time-varying* ARMA(1, 1)–GARCH(1, 1) *model to the log returns of all five leading cryptocurrencies during the training period. Darker points appear later in time.*
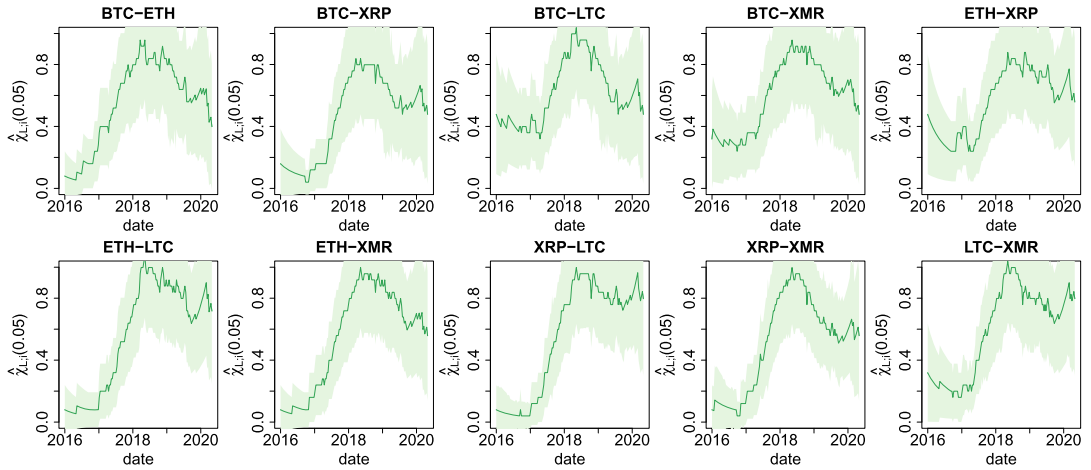
FIG. 3. *Time-varying lower-tail dependence coefficients* $\chi_L = \lim_{t \to 0} C(t,t)/t$, *estimated nonparametrically using the moving window estimator* $\widehat{\chi}_{L;i}(t) = \sum_{j \in J_i} \mathbb{I}(u_{j1} < t, u_{j2} < t)/(|J_i| t)$, *where* $J_i = \{\max(i - \tau, 1), \ldots, \min(i + \tau, n)\}$, $i = 1, \ldots, n$, *with threshold* $t = 0.05$ *and window size* $\tau = 500$, *based on uniformly distributed residuals* $(u_{11}, u_{12})^\top, \ldots, (u_{n1}, u_{n2})^\top$ *for all pairs of cryptocurrencies during the training period. Dark lines show point estimates, while light shaded areas display a 95% confidence envelopes obtained by the Delta method.*

reports the results. These nonparametric estimates clearly show that all pairs of cryptocurrencies under consideration have undergone a major shift in their lower-tail dependence structures over the study period, evolving from weak to very strong dependence in recent years, and potentially indicating a regime switch from AI to AD. This suggests increasing systemic risks. The upper-tail dependence coefficients, plotted in the Supplementary Material (Gong and Huser (2022a)), suggest that the upper-tail dependence structure is generally weaker and may also have evolved over time, though this is less clear. However, empirical estimates are naturally very variable and also cannot provide deep insights into whether a regime switch from AI to AD has truly occurred in each tail. This provides us with a strong motivation to perform a more in-depth model-based study of the time-varying patterns driving the co-occurrence of low or high values for the different pairs of cryptocurrencies under consideration. In Section 5 we analyze the data further by fitting various copula models, using different inference approaches, in order to assess both tail-dependence structures in a joint framework, and to accurately quantify the tail risk among these five leading cryptocurrencies.

3. Modeling.

3.1. *Model construction.*   We here describe the construction of our copula model used to assess the lower and upper dependence structures among cryptocurrency data.

In order to construct a parsimonious dependence model that possesses high tail flexibility, we mix an asymptotically independent random vector with a perfectly dependent random vector on a suitable marginal scale. Specifically, let $R \sim F_R$ be a random variable with asymmetric Laplace distribution, denoted $\mathrm{AL}(\delta_L, \delta_U)$,

$$(3.1) \qquad F_R(r) = \begin{cases} \dfrac{\delta_L}{\delta_L + \delta_U} \exp(r/\delta_L), & r \le 0, \\ 1 - \dfrac{\delta_U}{\delta_L + \delta_U} \exp(-r/\delta_U), & r > 0, \end{cases} \qquad r \in \mathbb{R},$$

where $\delta_L, \delta_U \in (0, 1)$ are scale parameters for the lower and upper tails, respectively. Furthermore, let $W_1, W_2 \sim F_W$ have the $\mathrm{AL}(1 - \delta_L, 1 - \delta_U)$ distribution, and assume that the
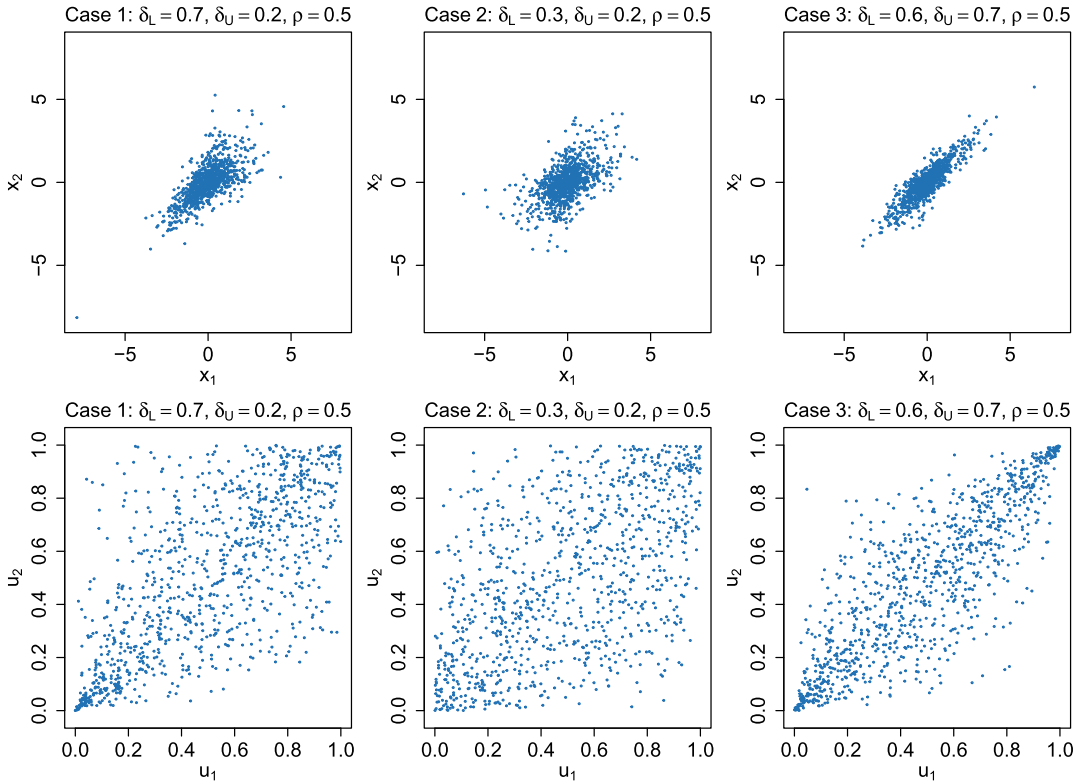
FIG. 4. *1000 independent samples from model* (3.3) *with correlation* $\rho = 0.5$ *and tail parameters* $\delta_L = 0.7, \delta_U = 0.2$ *(left),* $\delta_L = 0.3, \delta_U = 0.2$ *(middle) and* $\delta_L = 0.6, \delta_U = 0.7$ *(right). Simulated data are plotted on the original scale of* X *in* (3.3) *(top) or transformed into the standard uniform marginal scale (bottom).*

bivariate random vector $W = (W_1, W_2)^\top$ is driven by a Gaussian copula with correlation $\rho \in (-1, 1)$. In other words, the joint distribution of W satisfies

$$(3.2) \qquad \Pr(W_1 \leq w_1, W_2 \leq w_2) = \Phi_\rho[\Phi^{-1}\{F_W(w_1)\}, \Phi^{-1}\{F_W(w_2)\}],$$

where Φ and Φ_ρ denote the univariate standard Gaussian distribution and bivariate standard Gaussian distribution with correlation ρ, respectively. Our dependence model is now defined through the random vector $X = (X_1, X_2)^\top$ with components

$$(3.3) \qquad X_1 = R + W_1, \qquad X_2 = R + W_2.$$

As the random variable R is common to both X_1 and X_2, it can be interpreted through the perfectly dependent random vector $R = (R, R)^\top$, while the random vector W has a Gaussian dependence structure and is, therefore, asymptotically independent. Noting that the $\mathrm{AL}(\delta_L, \delta_U)$ distribution converges to a degenerate distribution with all its mass at zero as $\delta_L, \delta_U \to 0$, the dependence structure of X thus interpolates between that of W (Gaussian) as $\delta_L, \delta_U \to 0$ and that of R (perfect dependence) as $\delta_L, \delta_U \to 1$. Moreover, similarly to the model of Huser and Wadsworth (2019), which is designed for capturing the upper-tail behavior only, when $\delta_U > 0.5$, R intuitively "dominates" W in the upper-tail region, which induces strong upper-tail dependence, and the opposite is true when $\delta_U < 0.5$. The same holds for the lower tail controlled by the parameter δ_L. Hence, high flexibility can here be achieved in both the lower- and upper-joint tails using this parsimonious three-parameter $(\delta_L, \delta_U, \rho)$ model. To illustrate this, we plot in Figure 4 random samples from the model (3.3) with different parameter values, showing that a wide range of tail behaviors can be generated.

REMARK 1.   The construction (3.3) is only used to define a model with flexible lower- and upper-tail dependence structures. In practice, however, we first transform the data to the standard uniform scale and we then fit the copula associated with X. More details are given in Section 3.2 and Section 4.

REMARK 2.   The assumption (3.2) that the vector W has a Gaussian dependence structure is mainly made for computational convenience and to obtain the Gaussian copula model as a special case for X when $\delta_L, \delta_U \to 0$. However, evidence of both tail asymmetry and permutation asymmetry (i.e., asymmetry with respect to two diagonals of the unit square) has been found in some financial applications; see, for example, Krupskii (2017). In our model construction the Gaussian copula may be replaced by any other copula model that is asymptotically independent in both tails without affecting the tail dependence structures of X. Another interesting model for W is the skew-normal copula (Azzalini and Dalla Valle (1996)), which has additional "skewness" or "slant" parameters and can capture both permutation and tail asymmetry, thus increasing flexibility in the bulk and tail regions. This proposed model extension is illustrated in the Supplementary Material (Gong and Huser (2022a)) with simulated samples and fitted in our application in Section 5.

In the following section, Section 3.2, we derive the expressions related to the copula associated with our model (3.3), and in Section 3.3 we formally derive its tail dependence properties.

3.2. *Expressions for the associated copula.*   We first derive the marginal and joint distributions and densities of the vector $X = (X_1, X_2)^\top$, as defined in (3.3), from which the corresponding copula expressions can then be deduced. Let f_R denote the $\mathrm{AL}(\delta_L, \delta_U)$ density of R obtained by differentiating (3.1) with respect to the argument, r. The common marginal distribution F_X of X_i, $i = 1, 2$, is thus

$$
\begin{aligned}
F_X(x) &= \Pr(X_i \le x) = \Pr(R + W_i \le x) \\
&= \int_{\mathbb{R}} \Pr(W_i \le x - r) f_R(r) \, \mathrm{d}r \\
&= \frac{1}{\delta_L + \delta_U} \left\{ \int_{-\infty}^{0} \Pr(W_i \le x - r) \exp(r/\delta_L) \, \mathrm{d}r \right. \\
&\quad \left. + \int_{0}^{\infty} \Pr(W_i \le x - r) \exp(-r/\delta_U) \, \mathrm{d}r \right\}.
\end{aligned}
$$

(3.4)

By plugging the $\mathrm{AL}(1 - \delta_L, 1 - \delta_U)$ distribution of W_i, $i = 1, 2$, into (3.4), we can establish after some tedious but straightforward calculations that for $\delta_L, \delta_U \ne 1/2$, the marginal distribution of our model is equal to

$$
F_X(x) = \begin{cases} K_1(\delta_L, \delta_U) \exp\left(\dfrac{x}{\delta_L}\right) - K_2(\delta_L, \delta_U) \exp\left(\dfrac{x}{1 - \delta_L}\right), & x \le 0, \\[2mm] 1 + K_3(\delta_L, \delta_U) \exp\left(-\dfrac{x}{\delta_U}\right) - K_4(\delta_L, \delta_U) \exp\left(-\dfrac{x}{1 - \delta_U}\right), & x > 0, \end{cases}
$$

where the normalizing constants are $K_1(\delta_L, \delta_U) = \delta_L^3 \{(\delta_L + \delta_U)(2\delta_L - 1)(1 + \delta_L - \delta_U)\}^{-1}$, $K_2(\delta_L, \delta_U) = (\delta_L - 1)^3 \{(2\delta_L - 1)(\delta_L - \delta_U - 1)(2 - \delta_L - \delta_U)\}^{-1}$, $K_3(\delta_L, \delta_U) = \delta_U^3 \{(\delta_L + \delta_U)(2\delta_U - 1)(\delta_L - \delta_U - 1)\}^{-1}$, and $K_4(\delta_L, \delta_U) = (\delta_U - 1)^3 \{(2\delta_U - 1)(1 + \delta_L - \delta_U)(2 - \delta_U - \delta_U)\}^{-1}$. The intermediate cases when $\delta_L = 1/2$ and/or $\delta_U = 1/2$ can be established separately, and are reported in Appendix A for completeness. The marginal density f_X is easily derived from the above formula for F_X by differentiation.

Using (3.2), the joint distribution $F_X(x_1, x_2)$ of X_1 and X_2 may be expressed as

$$F_X(x_1, x_2) = \Pr(X_1 \leq x_1, X_2 \leq x_2) = \Pr(R + W_1 \leq x_1, R + W_1 \leq x_2)$$

(3.5)
$$= \int_{\mathbb{R}} \Pr(W_1 \leq x_1 - r, W_2 \leq x_2 - r) f_R(r) \, dr$$

$$= \int_{\mathbb{R}} \Phi_\rho [\Phi^{-1}\{F_W(x_1 - r)\}, \Phi^{-1}\{F_W(x_2 - r)\}] f_R(r) \, dr,$$

which involves the bivariate standard Gaussian distribution Φ_ρ, the standard Gaussian quantile function Φ^{-1}, the AL(δ_L, δ_U) density, f_R, and the AL$(1 - \delta_L, 1 - \delta_U)$ distribution, F_W. By differentiating under the integral sign, we obtain the joint density $f_X(x_1, x_2)$ as

$$f_X(x_1, x_2)$$

(3.6)
$$= \int_{\mathbb{R}} \frac{\partial^2}{\partial x_1 \partial x_2} \Pr(W_1 \leq x_1 - r, W_2 \leq x_2 - r) f_R(r) \, dr$$

$$= \int_{\mathbb{R}} \phi_\rho [\Phi^{-1}\{F_W(x_1 - r)\}, \Phi^{-1}\{F_W(x_2 - r)\}] \prod_{i=1}^{2} \frac{f_W(x_i - r)}{\phi\{F_W(x_i - r)\}} f_R(r) \, dr,$$

where ϕ and ϕ_ρ denote the univariate standard Gaussian density and the bivariate standard Gaussian density with correlation ρ, respectively. Similarly, we can derive the partial derivatives of the distribution $F_X(x_1, x_2)$ which are required for the censored likelihood inference approach described in Section 4. Writing ∂_1 and ∂_2 to denote differentiation with respect to the first and second arguments, respectively, we have

$$\partial_1 F_X(x_1, x_2) = \int_{\mathbb{R}} \Phi([\Phi^{-1}\{F_W(x_2 - r)\} - \rho \Phi^{-1}\{F_W(x_1 - r)\}]/\sqrt{1 - \rho^2})$$

(3.7)
$$\times \phi[\Phi^{-1}\{F_W(x_1 - r)\}] \frac{f_W(x_1 - r)}{\phi\{F_W(x_1 - r)\}} f_R(r) \, dr,$$

while $\partial_2 F_X(x_1, x_2)$ may be obtained by interchanging the labels.

REMARK 3. If the vector $W = (W_1, W_2)^\top$ is chosen to have a different dependence structure (e.g., with a skew-normal copula), the marginal distribution (3.4) and its density remain unchanged, while the joint distribution, density, and partial derivatives in (3.5), (3.6), and (3.7) are obtained in a similar form but with some slight modifications.

Now, define $U_i = F_X(X_i) \sim \text{Unif}(0, 1)$, $i = 1, 2$. The copula C, associated with $X = (X_1, X_2)^\top$, contains all the information about the dependence structure and is obtained as in (1.1), while its density and partial derivatives may be expressed as

$$c(u_1, u_2) = \frac{f_X\{F_X^{-1}(u_1), F_X^{-1}(u_2)\}}{f_X\{F_X^{-1}(u_1)\} f_X\{F_X^{-1}(u_2)\}},$$

(3.8)
$$\partial_i C(u_1, u_2) = \frac{\partial_i F_X\{F_X^{-1}(u_1), F_X^{-1}(u_2)\}}{f_X\{F_X^{-1}(u_i)\}},$$

for $i = 1, 2$. Notice that F_X and f_X are here available in closed form, which makes copula computations much more efficient than, for example, the models of Huser, Opitz and Thibaud (2017), where the marginal distribution and density are known only up to a unidimensional integral. The marginal quantile function F_X^{-1}, however, is not available in closed form but can be approximated efficiently using numerical root-finding algorithms. Similarly, it is impossible to obtain explicit expressions for F_X, f_X, and $\partial_i F_X$ in (3.5), (3.6), and (3.7), respectively,

but numerical integration routines may be used to accurately approximate them, and we have found that a simple finite integral computed from 10^4 subintervals works quite well for most parameter values. Overall, the computational burden, due to (3.5), (3.6), and (3.7), is roughly equivalent to that required for the model proposed by Huser and Wadsworth (2019).

3.3. *Tail dependence structures.* We now detail the lower- and upper-tail properties of our proposed model (3.3) and show that it can indeed capture a wide range of joint-tail decay rates in each tail.

We consider, for each threshold $t \in (0, 1)$, the tail coefficients

(3.9)
$$\chi_L(t) = \Pr(U_1 < t \mid U_2 < t) = \frac{C(t, t)}{t},$$
$$\chi_U(t) = \Pr(U_1 > t \mid U_2 > t) = \frac{1 - 2t + C(t, t)}{1 - t},$$

and their limits $\chi_L = \lim_{t \to 0} \chi_L(t)$ and $\chi_U = \lim_{t \to 1} \chi_U(t)$, expressed through the copula C of the random vector $\boldsymbol{U} = (U_1, U_2)^\top$. The coefficients χ_L and χ_U determine the asymptotic dependence class (AI/AD) in the lower and upper tails, respectively; recall the definition (1.2). In the asymptotically independent case the extremal dependence strength is more precisely described using the coefficient of tail dependence (Ledford and Tawn (1996)), sometimes also called the residual dependence coefficient, characterizing the rate of tail decay toward independence. Assume that the lower- and upper-tail coefficients admit the following expansions:

$$\chi_L(t) \sim \mathcal{L}_L(t^{-1})t^{1/\eta_L - 1}, \quad t \to 0, \qquad \chi_U(t) \sim \mathcal{L}_U\{(1 - t)^{-1}\}(1 - t)^{1/\eta_U - 1}, \quad t \to 1,$$

where the tail-specific functions $\mathcal{L}_{L/U}(\cdot)$ are slowly-varying at infinity, that is, they satisfy $\mathcal{L}_{L/U}(ax)/\mathcal{L}_{L/U}(x) \to 1$, as $x \to \infty$ for any real $a > 0$, and $0 < \eta_{L/U} \le 1$ are the coefficients of lower- and upper-tail dependence, respectively. If $\eta_{L/U} < 1$ or $\mathcal{L}_{L/U}(x) \to 0$ as $x \to \infty$, then $\chi_{L/U} = 0$, and we get asymptotic independence with $\eta_{L/U}$ controlling the tail-decay rate toward independence. In other cases, $\chi_{L/U} > 0$, and we get asymptotic dependence.

The following proposition details the lower- and upper-tail structures of Model (3.3) and establishes the corresponding extremal dependence classes. The proof relies on general results for random scale constructions (Engelke, Opitz and Wadsworth (2019)) and is postponed to Appendix B.

PROPOSITION 1 (Asymptotic dependence class and $\chi_{L/U}, \eta_{L/U}$ coefficients). *Consider a random vector $\boldsymbol{X}$, defined as in (3.3). Then, we have the following cases:*

(i) *Case 1: $\delta_{L/U} \le 1/2$. Then, $\boldsymbol{X}$ is asymptotically independent in its lower/upper tail with $\chi_{L/U} = 0$ and coefficient of lower/upper tail dependence obtained as*

$$\eta_{L/U} = \begin{cases} \delta_{L/U}/(1 - \delta_{L/U}), & \delta_{L/U} > (1 + \rho)/(3 + \rho), \\ (1 + \rho)/2, & \delta_{L/U} \le (1 + \rho)/(3 + \rho). \end{cases}$$

(ii) *Case 2: $\delta_{L/U} > 1/2$. Then, $\boldsymbol{X}$ is asymptotically dependent in its lower/upper tail with coefficient of tail dependence $\eta_{L/U} = 1$ and, writing $s_L = -1$ and $s_U = 1$,*

$$\chi_{L/U} = \mathrm{E}\left(\min\left[\frac{\exp(\frac{s_{L/U}}{\delta_{L/U}}W_1)}{\mathrm{E}\{\exp(\frac{s_{L/U}}{\delta_{L/U}}W_1)\}}, \frac{\exp(\frac{s_{L/U}}{\delta_{L/U}}W_2)}{\mathrm{E}\{\exp(\frac{s_{L/U}}{\delta_{L/U}}W_2)\}}\right]\right).$$

In order to visualize the various types of dependence structures that our model can produce, Figure 5 displays $\chi_L(t)$ and $\chi_U(t)$ for $t \in (0, 1)$. The next section discusses how to perform (full or censored and global or local) likelihood inference for our model.
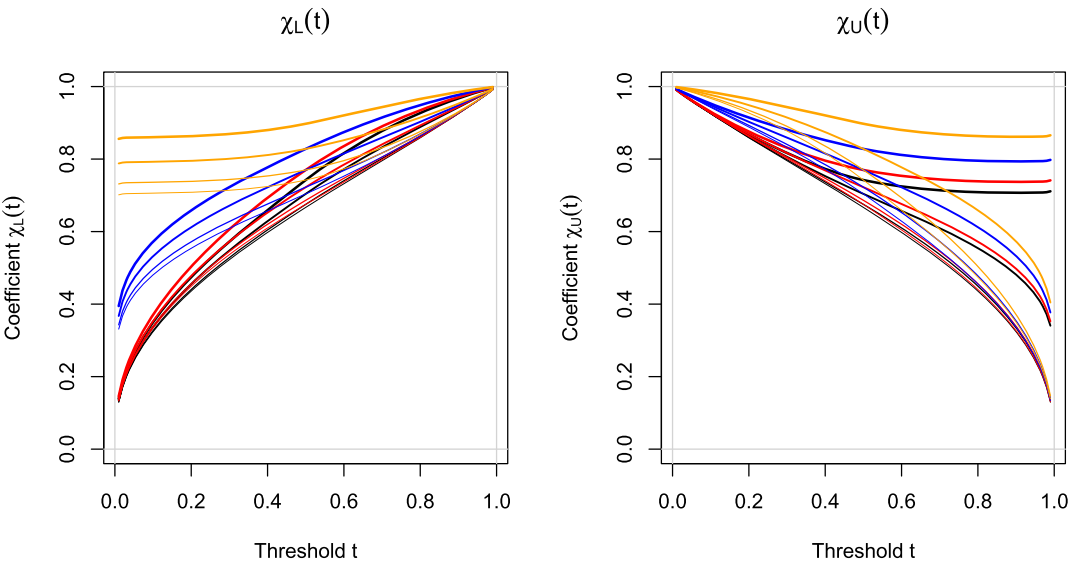
$$\chi_L(t) \qquad\qquad \chi_U(t)$$

FIG. 5. *Coefficients $\chi_L(t) = \mathrm{Pr}(U_1 < t \mid U_2 < t)$ (left) and $\chi_U(t) = \mathrm{Pr}(U_1 > t \mid U_2 > t)$ (right), with threshold $t \in [0.01, 0.99]$, for a random vector $U = (U_1, U_2)^\top$ on the uniform scale stemming from the model (3.3) with correlation $\rho = 0.5$ and tail parameters $\delta_L = 0, 0.2, 0.5, 0.8$ (black, red, blue, orange), $\delta_U = 0, 0.2, 0.5, 0.8$ (thin to thick curves).*

4. Inference.

4.1. *Full and censored likelihood approaches.* Let $Y_1, \ldots, Y_n$ denote n independent copies from a random vector $Y = (Y_1, Y_2)^\top$ that shares the same copula as the vector X in (3.3) but possesses potentially different marginal distributions $F_{Y,1}, F_{Y,2}$. In other words, the joint distribution of Y may be expressed as

$$F_Y(y_1, y_2) = \mathrm{Pr}(Y_1 \leq y_1, Y_2 \leq y_2) = C\{F_{Y,1}(y_1), F_{Y,2}(y_2)\},$$

where C is our copula model, defined in Section 3.2. In order to estimate the underlying copula C from an observed random sample $(y_{11}, y_{12})^\top, \ldots, (y_{n1}, y_{n2})^\top$, we adopt a two-step estimation approach. First, we estimate marginal distributions and transform the data to the standard uniform scale. To achieve this goal, we may either estimate $F_{Y,1}$ and $F_{Y,2}$ by fitting a parametric model to each margin or, more simply, by using the (nonparametric) empirical distribution functions $\widehat{F}_{Y,1}, \widehat{F}_{Y,2}$ based on ranks. The probability integral transform can then be used to get pseudo-uniform scores as $u_{j1} = \widehat{F}_{Y,1}(y_{j1})$ and $u_{j2} = \widehat{F}_{Y,2}(y_{j2})$, $j = 1, \ldots, n$. If the assumption of temporal stationarity is doubtful, it may also be possible to fit a dynamic model for the margins of Y by assuming that they vary over time according to some temporal covariate or, more flexibly, by adopting a local (parametric or nonparametric) estimation approach akin to the one discussed in Section 4.2 for the copula structure. Second, to estimate the dependence parameters, we adopt a likelihood-based approach. Under stationarity the full likelihood for our copula model (3.3) may be written as

$$(4.1) \qquad L(\theta) = \prod_{j=1}^{n} c(u_{j1}, u_{j2}), \qquad \theta = (\delta_L, \delta_U, \rho)^\top \in (0, 1) \times (0, 1) \times (-1, 1),$$

where the copula density c, defined in Section 3.2, depends on the model parameters $\theta = (\delta_L, \delta_U, \rho)^\top$. Maximizing (4.1) yields the full likelihood estimator $\widehat{\theta}_{\mathrm{Full}}$ which has well-known appealing large-sample properties.
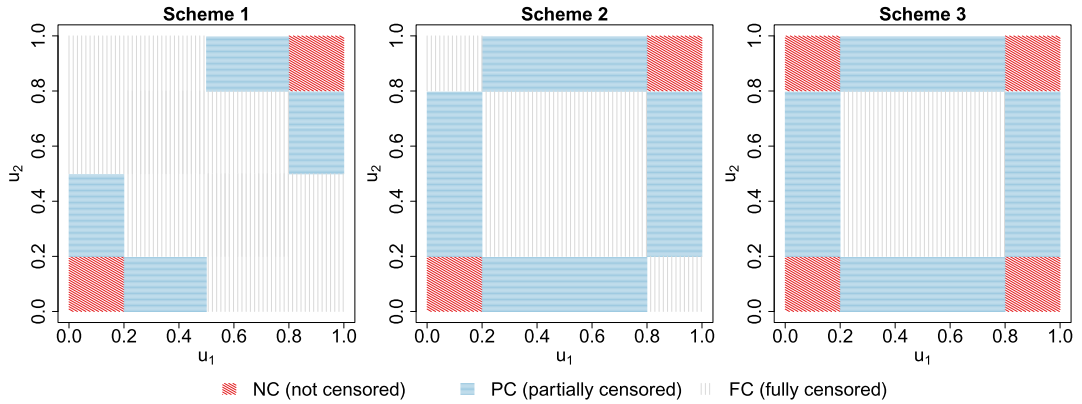
FIG. 6. *Three different censoring schemes putting the emphasis on the lower- and upper-joint tails which may be used in the censored likelihood approach.*

To prioritize calibration in the tails and reduce the influence of nonextreme observations (from the bulk), we can use instead various censored likelihoods of the form

$$(4.2) \quad L(\boldsymbol{\theta}) = \prod_{j \in A} L_{\mathrm{NC}}(u_{j1}, u_{j2}) \times \prod_{k=1}^{K_B} \prod_{j \in B_k} L_{\mathrm{PC}_1}^k(u_{j1}) \times \prod_{k=1}^{K_C} \prod_{j \in C_k} L_{\mathrm{PC}_2}^k(u_{j2}) \times \prod_{j \in D} L_{\mathrm{FC}},$$

where $L_{\mathrm{NC}}(u_{j1}, u_{j2}) = c(u_{j1}, u_{j2})$ are all noncensored likelihood contributions, while L_{FC} denotes fully censored likelihood contributions, involving the copula C, and $L_{\mathrm{PC}_1}^k(u_{j1})$ and $L_{\mathrm{PC}_2}^k(u_{j2})$ are (different types of) partially-censored likelihood contributions, the computation of which relies on the partial derivatives $\partial_1 C$ and $\partial_2 C$, respectively. Typically, the set A will correspond to noncensored points lying in the lower- and upper-joint tail regions (i.e., "genuine joint extremes" where both variables are small or large simultaneously); $\bigcup_{k=1}^{K_B} B_k$ and $\bigcup_{k=1}^{K_C} C_k$ will correspond to partially censored points located along the edges of the unique square (i.e., "partial extremes" with one component being extreme and the other not), and the set D will correspond to censored points lying in the bulk near the center of the unit square (i.e., "nonextremes" where none of the variables are small or large). Here, we propose three censoring schemes that are illustrated in Figure 6 which correspond to different definitions of what "joint extremes" actually means. Scheme 1 is perhaps the most natural one, as it only exploits information from extreme points near the joint upper and lower tails, while Scheme 2 uses extra information from points that are extremely large (respectively, small) in one variable and moderately small (respectively, large) in the other variable, and Scheme 3 also fully uses information from points that are extremely large (respectively, small) in one variable with extremely small (respectively, large) in the other. We stress here that, whatever the censoring scheme chosen, we always assume that our copula model is valid to describe the *whole* dataset (even in censored or partially censored subregions). The purpose of our proposed censoring schemes is thus simply to provide an inference method for estimating all three parameters δ_L, δ_U, and ρ jointly, while *prioritizing* calibration in the joint tails to make sure that the model fits joint extremes as best as possible subject to the model constraints. The choice of censoring scheme should be dictated by the application. The corresponding likelihood contributions are specific to each censoring scheme; see Appendix C for more details. Each of these censoring schemes depends on two quantiles $t_L, t_U \in (0, 1)$, defining the lower-tail and upper-tail censoring levels, respectively. In the sequel we take t_L to be a low quantile (such as, e.g., 0.01 or 0.1) and $t_U = 1 - t_L$. In Section 4.3 we perform an extensive simulation study to assess the performance of the censored likelihood estimators $\widehat{\boldsymbol{\theta}}_{\mathrm{Cens}}$, maximizing (4.2) under the censoring Schemes 1, 2 and 3 and various censoring levels.

4.2. *Local estimation approach for time-varying copula models.* As exemplified in Figure 1, financial market data are often nonstationary over time with volatility clusters appearing in periods of stress, and recent papers have proposed methods to estimate extremal (marginal) trends in heteroscedastic time series (de Haan and Zhou (2021), Einmahl, de Haan and Zhou (2016)). Beyond marginal distributions, Poon, Rockinger and Tawn (2003) and Castro-Camilo, de Carvalho and Wadsworth (2018) have realized and demonstrated that the dependence structure of such data may also vary over time. We also observe this phenomenon in Figure 3 for the cryptocurrencies under investigation with the lower-tail dependence becoming stronger in recent years. To estimate the temporal dynamics of extremal dependence, Castro-Camilo, de Carvalho and Wadsworth (2018) and Mhalla, de Carvalho and Chavez-Demoulin (2019) suggested using a (nonparametric) kernel estimator and (semiparametric) vector generalized additive models of the spectral density, respectively. We here instead address this issue by proposing a local copula-based likelihood estimation approach that can capture complex trends in a very flexible way.

Each full or censored likelihood in (4.1) and (4.2), respectively, can be rewritten as a product of likelihood contributions, namely, $L(\boldsymbol{\theta}) = \prod_{j=1}^{n} L_j(\boldsymbol{\theta})$. We now assume that the dependence structure smoothly evolves over time, and so we estimate a family of parameters $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$ (one for each time point). To do this, we replace the (full or censored) likelihood function by a family of weighted local likelihoods to be maximized, which have the form

$$(4.3) \qquad L(\boldsymbol{\theta}_i) = \prod_{j=1}^{n} \omega_\tau(|j - i|) L_j(\boldsymbol{\theta}_i), \quad i = 1, \ldots, n,$$

where $\omega_\tau(h) \geq 0$ is a nonnegative weight function (or "kernel") with bandwidth $\tau > 0$, down-weighting observations that are distant in time. For example, we can take the biweight function $\omega_\tau(h) = \{1 - (h/\tau)^2\}_+^2$ with compact support $[-\tau, \tau]$ which smoothly decays to zero at the endpoints $-\tau$ and τ. Other (symmetric or asymmetric) kernels may also be used. As always with local approaches, the choice of the kernel is not so important but the bandwidth is crucial, as it leads to a bias-variance trade-off which controls the smoothness of trends in estimated parameters. Small bandwidths lead to parameter estimates that are very variable but with a lot of local detail, while large bandwidths lead to smooth estimates with low variability. A good bandwidth usually lies in between these two extremes and is typically chosen pragmatically based on the results' interpretability.

4.3. *Simulation study.* To compare full and censored likelihood estimators, based on (4.1) and (4.2), respectively, we now conduct an extensive simulation study in well-specified and misspecified settings. Furthermore, we also demonstrate the performance of the local estimation approach based on (4.3) in a time-varying context.

4.3.1. *Well-specified stationary setting.* We start by simulating $n = 1000$ independent samples from the model (3.3) with correlation $\rho = 0.5$ and tail parameters $\delta_L = 0.7, \delta_U = 0.2$ (Case 1: strong lower-tail dependence, weak upper-tail dependence), $\delta_L = 0.3, \delta_U = 0.2$ (Case 2: weak lower- and upper-tail dependence), and $\delta_L = 0.6, \delta_U = 0.7$ (Case 3: strong lower- and upper-tail dependence). These cases, illustrated in Figure 4, cover various combinations of extremal dependence classes in each tail. We then estimate the model parameters $\boldsymbol{\theta} = (\delta_L, \delta_U, \rho)^\top$, using the full likelihood estimator $\widehat{\boldsymbol{\theta}}_{\text{Full}}$ in (4.1) and the three censored likelihood estimators $\widehat{\boldsymbol{\theta}}_{\text{Cens}}$ in (4.2), illustrated in Figure 6, using lower-tail censoring levels of $t_L = 0.01, 0.02, 0.05, 0.1, 0.2$ and upper-tail censoring levels equal to $t_U = 1 - t_L$. This yields 16 estimators in total (1 full likelihood, + 3 censoring schemes, times 5 censoring levels). We then repeat this experiment 300 times to produce boxplots of estimated parameters. The results for Case 1 are reported in Figure 7. Results for Cases 2 and 3 are similar and reported
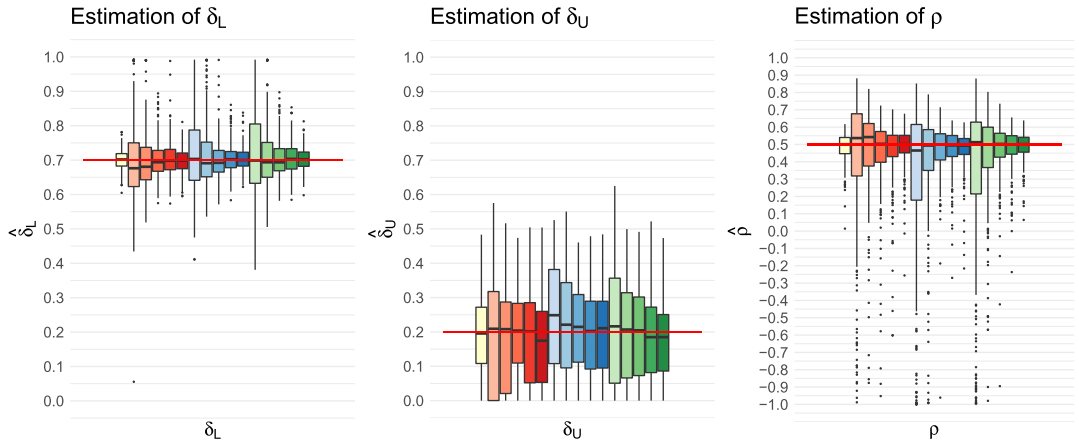
FIG. 7. *Results for Case 1 in the well-specified stationary setting with true values set to* $\delta_L = 0.7, \delta_U = 0.2, \rho = 0.5$. *The panels display boxplots of estimated values for* δ_L (*left*), δ_U (*middle*), *and* ρ (*right*), *based on the full likelihood estimator (yellow), and censored likelihood estimators, based on censoring scheme 1 (red), scheme 2 (blue), and scheme 3 (green). Lower-tail censoring levels of* $t_L = 0.01, 0.02, 0.05, 0.1, 0.2$ (*from lighter to darker red/blue/green colors*) *and upper-tail censoring levels equal to* $t_U = 1 - t_L$.

in the Supplementary Material (Gong and Huser (2022a)). Essentially, the results show that all estimation approaches work well, and the full likelihood estimator is the most efficient as expected. All three types of censored likelihood estimators perform similarly. Moreover, high censoring levels (such as $t_L = 0.01, t_U = 0.99$ or $t_L = 0.02, t_U = 0.98$), which put a strong emphasis on the tails and prioritize model calibration for joint extreme events, result in much higher uncertainty, owing to the largely reduced effective sample size. In contrast, with low censoring levels (such as $t_L = 0.2, t_U = 0.8$ or $t_L = 0.1, t_U = 0.9$), the variability of censored likelihood estimators is almost equivalent to the full likelihood case.

We then repeat the simulation study for Case 1 but consider increasing sample sizes $n = 500, 1000, 2000$. The results are reported in the Supplementary Material (Gong and Huser (2022a)). As expected, the variability of estimated parameters is reduced by increasing the sample size, and the boxplots' interquartile ranges roughly decrease at rate $n^{1/2}$ which corroborates asymptotic theory.

4.3.2. *Misspecified stationary setting.* To assess the flexibility of our parsimonious copula model and explore the effect of censoring nonextreme observations, we now investigate a misspecified setting, whereby the data are simulated from the bivariate Gumbel (also called "logistic") extreme-value copula, that is,

$$(4.4) \qquad C_{\text{Gum}}(u_1, u_2) = \exp(-[\{-\log(u_1)\}^{1/\alpha} + \{-\log(u_2)\}^{1/\alpha}]^\alpha),$$

where $\alpha \in (0, 1]$ is the dependence parameter, interpolating from independence ($\alpha = 1$) to perfect positive dependence ($\alpha \to 0$). This extreme-value copula is known to be asymptotically dependent in the upper tail with $\chi_U = 2 - 2^\alpha$ and $\eta_U = 1$ and asymptotically independent in the lower tail with $\chi_L = 0$ and $\eta_L = 2^{-\alpha}$; see Tawn (1988, 1990) and Ledford and Tawn (1996). We simulate $n = 1000$ independent samples from (4.4) with $\alpha = 0.2, 0.5, 0.8$ (from strong to weak dependence) and then fit our model (3.3) instead to assess its flexibility in capturing the lower and upper extremal dependence classes in this misspecified setting. We consider the full likelihood estimator and the three censored likelihood estimators presented above with censoring level $t_L = 0.01, 0.02, 0.05$ and $t_U = 1 - t_L$. As before, we repeat the experiment 300 times to compute performance metrics. Table 1 reports the results for the case

Results for misspecified stationary setting by simulating data from the Gumbel copula model (4.4) *with* $\alpha = 0.5$,
but fitting our copula model stemming from (3.3). *For each estimator* (*left column*), *we report* (*from left to right*)
the percentage of times that χ_L *is estimated to be zero* (*the true value*), *the median and median absolute deviation*
(*MAD*) *of the* η_L *and* χ_U *estimates, and the percentage of times that* η_U *is estimated to be one* (*the true value*)

True values	$\chi_L = 0$	$\eta_L = 0.71$	$\chi_U = 0.59$	$\eta_U = 1$
Estimators	$\%\{\widehat{\chi}_L = 0\}$	$\widehat{\eta}_L$, Median/MAD	$\widehat{\chi}_U$, Median/MAD	$\%\{\widehat{\eta}_U = 1\}$
Full likelihood	99%	0.75/0.07	0.67/0.12	100%
Cens., Scheme 1, $t_L = 0.01$	97%	0.77/0.13	0.68/0.17	97%
Cens., Scheme 1, $t_L = 0.02$	98%	0.76/0.10	0.69/0.17	99%
Cens., Scheme 1, $t_L = 0.05$	99%	0.75/0.08	0.69/0.16	100%
Cens., Scheme 2, $t_L = 0.01$	98%	0.74/0.09	0.69/0.16	100%
Cens., Scheme 2, $t_L = 0.02$	98%	0.73/0.08	0.69/0.16	99%
Cens., Scheme 2, $t_L = 0.05$	100%	0.74/0.06	0.68/0.14	100%
Cens., Scheme 3, $t_L = 0.01$	96%	0.79/0.14	0.71/0.22	97%
Cens., Scheme 3, $t_L = 0.02$	99%	0.76/0.09	0.69/0.16	100%
Cens., Scheme 3, $t_L = 0.05$	99%	0.75/0.07	0.68/0.15	100%

$\alpha = 0.5$. The cases $\alpha = 0.2$ and $\alpha = 0.8$ are reported in the Supplementary Material (Gong and Huser (2022a)). When the dependence strength is moderate, our model succeeds in estimating the tail dependence classes in most cases, and there is little difference between the estimators considered in this simulation experiment. The coefficients η_L and χ_U appear to be quite well estimated in most cases, albeit with a slight positive bias. This might be due to the shared correlation parameter ρ being common to both tails, hence restricting the possible tail structures that can be estimated.

We further perform another experiment by simulating data from the (misspecified) Coles–Tawn extreme-value copula model (Coles and Tawn (1991)), which captures permutation-asymmetry (i.e., nonexchangeability in both arguments) and has two dependence parameters controlling the overall dependence strength and the extent of asymmetry. The results reported in the Supplementary Material (Gong and Huser (2022a)) are very similar to Table 1 under both mild and strong asymmetry, showing an almost-perfect identification of the asymptotic dependence class but a slight positive bias for $\widehat{\eta}_L$ and $\widehat{\chi}_U$. Our proposed copula model is thus very flexible and our inference approach robust enough to provide accurate tail dependence estimates, even in highly misspecified settings.

4.3.3. *Dynamic time-varying setting.* Finally, we simulate data in a nonstationary setting, where both marginal parameters and the copula structure vary smoothly over time, similarly to our real data analysis in Section 5. The goal is to assess whether our proposed weighted local likelihood approach, based on (4.3), can accurately recover the underlying dynamic dependence structure. To mimic the real data application, we simulate two time series according to different ARMA(1, 1)–GARCH(1, 1) models with time-varying parameters, and we link them together through our copula model (3.3) with constant δ_U and ρ parameters but time-varying δ_L parameter, representing increasing lower tail dependence strength over time. Further details about the marginal model, including plots of true marginal parameters, are given in the Supplementary Material (Gong and Huser (2022a)), while the true copula parameters $\boldsymbol{\theta}_i = (\delta_{L;i}, \delta_{U;i}, \rho_i)^\top$ for the $n = 1500$ time points $i = 1, \ldots, n$ are shown in Figure 8. Specifically, the data are simulated with $\delta_{L;i} = 0.4\Phi(10i/n - 5) + 0.2$, where $\Phi(\cdot)$ is the standard Gaussian distribution, $\delta_{U;i} = 0.3$ and $\rho_i = 0.5$, $i = 1, \ldots, n$. Thus the true lower tail dependence structure transitions from asymptotic independence (with $\delta_{L;1} \approx 0.2$)
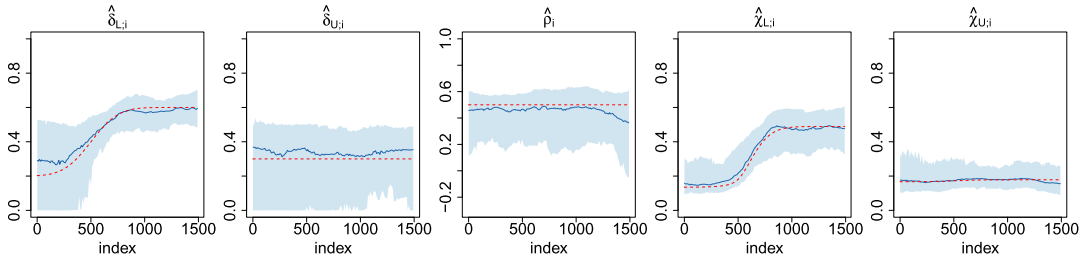
FIG. 8. *Time-varying copula parameter estimates* $\widehat{\boldsymbol{\theta}}_i = (\widehat{\delta}_{L;i}, \widehat{\delta}_{U;i}, \widehat{\rho}_i)^\top$ *(first three panels) and the corresponding lower- and upper-tail coefficient estimates* $\widehat{\chi}_{L;i}(0.05)$ *and* $\widehat{\chi}_{U;i}(0.95)$ *(last two panels),* $i = 1, \ldots, n$, *for the dynamic copula model simulation study (see Section 4.3.3 for more details). True parameters are plotted as dashed lines. Pointwise medians across the 100 experiments are plotted as dark solid lines. Light shaded areas are pointwise 95% confidence intervals calculated from the 100 experiments.*

to asymptotic dependence (with $\delta_{L;n} \approx 0.6$), while the upper tail dependence structure remains at a weak asymptotic independence level. We estimate marginal parameters in a first step by fitting a time-varying ARMA(1, 1)–GARCH(1, 1) using a moving window approach; then, after transforming the data to the uniform scale, we estimate the dependence parameters, based on (4.3), with full (i.e., noncensored) likelihood contributions and weight function $\omega_\tau(h) = \{1 - (h/\tau)^2\}_+^2$ with bandwidth $\tau = 500$. The bandwidth is chosen similarly for marginal and dependence parameter estimation in a way to provide a reasonable bias-variance trade-off. We repeat the experiment 100 times to assess the overall estimation uncertainty (representing both marginal and dependence estimation uncertainties). The results are presented in Figure 8. The pointwise median of parameter estimates across the 100 experiments follows the true parameters very closely, even in the middle of the time period when the parameter $\delta_{L;i}$ is evolving quite rapidly. There seems to be a slight positive bias in $\widehat{\delta}_{L;i}$ for the initial time points which might be due to edge effects that are characteristic of local estimation approaches. Nevertheless, the parameter estimate $\widehat{\delta}_{L;i}$ appears unbiased for later time points, and the estimates of the tail summary statistics $\chi_{L;i}(0.05)$ and $\chi_{U;i}(0.95)$ are very well estimated. Overall, our proposed approach works well for estimating dynamic dependence structures: it can clearly capture time-varying patterns and extract signal from the data, although the estimation uncertainty is relatively large, especially for low values of $\delta_{L;i}$ and $\delta_{U;i}$. This is due to the lower effective sample size of local estimation approaches. Variability of parameter estimates may be reduced by increasing the bandwidth τ (i.e., considering less local estimators), though at a cost in larger bias. Therefore, this emphasizes once more the importance of having a parsimonious (but flexible) copula model which makes our proposed model (3.3) especially appealing. Moreover, despite the variability of parameter estimates, the summaries $\chi_{L;i}(0.05)$ and $\chi_{U;i}(0.95)$ are very well estimated with low uncertainty.

5. Application: Assessment of tail risk among leading cryptocurrencies.

5.1. *Global estimation of extremal dependence.* We now come back to the analysis that we started in Section 2. To uncover the tail dependence structure among leading cryptocurrencies, we fit our proposed copula model (3.3) to the historical daily prices of pairs of cryptocurrencies (pretransformed to the uniform scale) during the training period. In this section we first assume that the dependence structure is stationary and use the various (global) full and censored likelihood estimators detailed in Section 4.1, while in Section 5.2 we further explore its time evolution using the local estimation approach from Section 4.3.3. We here focus on the pair BTC–ETH for illustration, as it corresponds to the two cryptocurrencies with the largest market capitalizations, while in Section 5.2 we also discuss results for others pairs
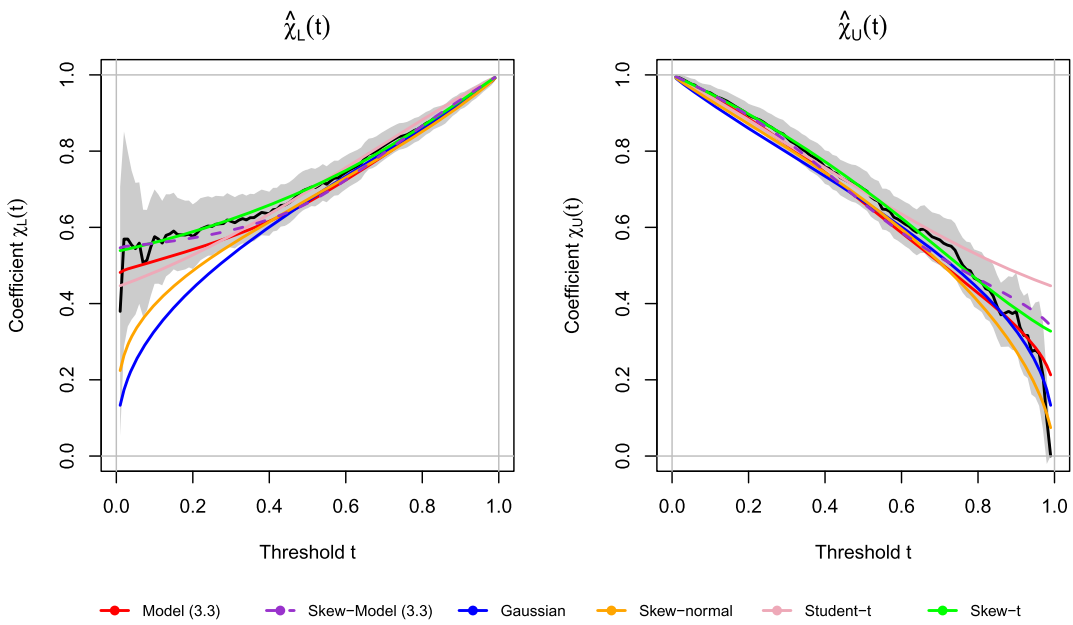
FIG. 9. *Coefficients* $\chi_L(t)$ *(left) and* $\chi_U(t)$ *(right), for* $t \in (0, 1)$, *estimated nonparametrically (black), using Model* (3.3) *(red), the skewed version of Model* (3.3) *(purple), the Gaussian copula (blue), the skew-normal copula (orange), the Student-t copula (pink) and the skew-t copula (green), from the BTC–ETH training data with corresponding 95%-bootstrap confidence envelope (grey). Results are based on the (global) full likelihood approach.*

of cryptocurrencies (reported in the Supplementary Material (Gong and Huser (2022a)) for completeness). For comparison purposes and to illustrate the performance of our proposed copula model, we also fit the skewed version of our model (based on the skew-normal copula for the vector $\boldsymbol{W}$ in (3.3)) as well as more traditional copula models, including the Gaussian copula, the skew-normal copula (Azzalini and Dalla Valle (1996)), the Student-t copula (Demarta and McNeil (2005)), and the skew-t copula (Arellano-Valle and Genton (2010), Demarta and McNeil (2005)).

Figure 9 plots the coefficients $\chi_L(t)$ and $\chi_U(t)$ in (3.9), estimated nonparametrically or from the fitted copula models using the full likelihood (4.1). While the upper tail (joint gains of BTC and ETH) appears to be asymptotically independent, with $\chi_U(t)$ decreasing to zero as $t \to 1$, the lower tail (joint losses of BTC and ETH) has much stronger dependence and appears to be asymptotically dependent. However, the uncertainty surrounding these empirical estimates is also quite high, and so the fit of our model bridging AD/AI classes provides more insight. As the data appear to be clearly tail asymmetric from Figure 9, the symmetric copula models (Gaussian and Student-t) provide a poor fit in one or both tails. Moreover, the skew-normal copula is AI in both tails and underestimates the lower tail probabilities. From Figure 9, our proposed model (3.3), its skewed version, and the skew-t copula seem to provide the best fits in the lower tail. Among these three models, our proposed copula model (3.3) performs best in the upper tail, at least visually. However, as the coefficients $\chi_L(t)$ and $\chi_U(t)$ plotted in Figure 9 provide only partial information about the dependence structure, we also consider more comprehensive information criteria to quantitatively determine which model provides the best overall fit.

Table 2 reports the estimated parameters for all models based on the (global) full likelihood approach and the censored likelihood based on censoring scheme 1 and censoring level $t_L = 0.1, t_U = 1 - t_L = 0.9$. To objectively compare the models, we also report the Akaike information criterion (AIC). Note that the censored likelihood is a valid likelihood function

TABLE 2

Estimated parameters $\widehat{\delta}_L$ (lower tail), $\widehat{\delta}_U$ (upper tail), $\widehat{\rho}$ (correlation), $\widehat{\alpha}_1$ (skewness for first margin), $\widehat{\alpha}_2$ (skewness for second margin), and $\widehat{v}$ (degrees of freedom) with 95% confidence intervals (CI), based on a parametric bootstrap procedure, and the Akaike information criteria (AIC), obtained by fitting the different copula models to the cryptocurrency training dataset (here, for the pair BTC and ETH). The estimators used are based on the (global) full likelihood and the censored likelihood estimator using censoring scheme 1 and censoring level $t_L = 0.1$, $t_U = 1 - t_L$; recall Section 4.1. For each inference approach, the best model (lowest AIC value) appears in bold

Copula	Cens. level	$\widehat{\delta}_L$ (95% CI)	$\widehat{\delta}_U$ (95% CI)	$\widehat{\rho}$ (95% CI)	$\widehat{\alpha}_1$ (95% CI)	$\widehat{\alpha}_2$ (95% CI)	$\widehat{v}$ (95% CI)	AIC
Model (3.3)	full lik.	0.65 (0.65, 0.75)	0.50 (0.50, 0.61)	−0.30 (−0.97, −0.30)	–	–	–	−775.3
	$t_L = 0.1$	0.77 (0.68, 0.84)	0.53 (0.45, 0.58)	−1.00 (−1.00, −0.35)	–	–	–	504.4
Skew-Model (3.3)	full lik.	0.71 (0.68, 0.75)	0.57 (0.55, 0.62)	−0.80 (−0.96, −0.63)	−0.16 (−0.15, 0.80)	1.25 (−0.25, 0.82)	–	−803.3
	$t_L = 0.1$	0.63 (0.43, 0.81)	0.45 (0.39, 0.56)	0.68 (−0.03, 0.93)	−0.26 (−2.26, 0.94)	0.53 (−2.11, 0.93)	–	**355.8**
Gaussian	full lik.	–	–	0.51 (0.46, 0.56)	–	–	–	−463.2
	$t_L = 0.1$	–	–	0.80 (0.74, 0.85)	–	–	–	595.2
Skew-normal	full lik.	–	–	0.70 (0.36, 0.77)	0.09 (−7.93, 1.42)	−4.74 (−7.90, 1.54)	–	−489.0
	$t_L = 0.1$	–	–	0.92 (0.89, 0.94)	−3.29 (−7.56, 0.85)	0.0006 (−8.72, 0.003)	–	541.7
Student-t	full lik.	–	–	0.57 (0.52, 0.61)	–	–	1.89 (1.66, 2.22)	−767.3
	$t_L = 0.1$	–	–	0.50 (0.44, 0.57)	–	–	1.85 (1.59, 2.42)	540.5
Skew-t	full lik.	–	–	0.66 (0.60, 0.72)	−0.38 (−0.74, −0.10)	−0.63 (−1.05, −0.30)	1.83 (1.62, 2.14)	**−811.7**
	$t_L = 0.1$	–	–	0.66 (0.55, 0.76)	−0.47 (−0.94, −0.13)	−0.83 (−2.50, −0.34)	1.85 (1.61, 2.50)	520.3

itself (based on a different dataset comprised of threshold exceedances together with exceedance indicators) so that the AIC can be used as a sensible model selection criterion in that setup also. The ranking of models is consistent with Figure 9. Without surprise, the Gaussian copula is by far the worst, followed by the skew-normal copula, both of which cannot capture AD. The Student-*t* copula can capture AD and thus has an increased performance with a lower AIC value but is worse than the skew-*t* copula and our proposed copula model (3.3) or its skewed version which have additional flexibility to capture tail asymmetry. Overall, these three best models have fairly equivalent goodness-of-fit performances, although the skew-*t* copula has here a slightly lower AIC value when considering the full likelihood approach. We note, however, that our proposed model (3.3) (with three parameters) is more parsimonious than the skew-*t* model (with four parameters), and this is a crucial aspect to take into consideration with low sample sizes or local likelihood approaches. Interestingly, when the censored likelihood approach is used, our copula model (3.3) and its skewed version have by far the best performances. This may be explained by the higher tail flexibility of our proposed model, as the censored likelihood approach precisely emphasizes calibration in the lower and upper tail regions.

From Table 2 the estimated lower tail parameter $\widehat{\delta}_L$ in our model (3.3), or its skewed version, is estimated to be larger than 0.5 (with a 95% confidence interval, excluding 0.5) which confirms that big losses of BTC and ETH are indeed asymptotically dependent. The
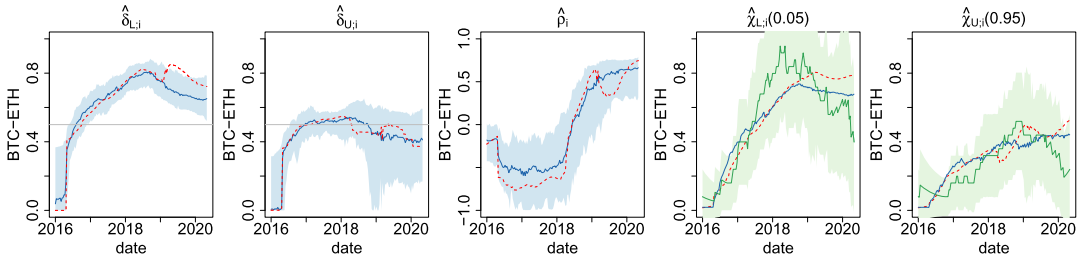
FIG. 10. *Time-varying copula parameter estimates* $\widehat{\boldsymbol{\theta}}_i = (\widehat{\delta}_{L;i}, \widehat{\delta}_{U;i}, \widehat{\rho}_i)^\top$ *(first three panels) and the corresponding lower and upper tail coefficient estimates* $\widehat{\chi}_{L;i}(0.05)$ *and* $\widehat{\chi}_{U;i}(0.95)$ *(last two panels),* $i = 1, \ldots, n$, *obtained by fitting the copula model* (3.3) *to cryptocurrency data (BTC and ETH) over the training period, using the local likelihood approach* (4.3) *with a biweight kernel* $\omega_\tau(h) = \{1 - (h/\tau)^2\}_+^2$ *and bandwidth* $\tau = 500$. *In each panel, dashed lines are model-based estimates based on full likelihood contributions. The horizontal grey lines at 0.5 in the first two panels correspond to the boundary between AI and AD regimes. Shaded areas (first three panels) are 95%-parametric bootstrap confidence envelopes for the model-based estimates, and dark solid lines are pointwise bootstrap medians. Dark solid lines (last two panels) show nonparametric estimates of the tail coefficients* $\chi_L(0.05)$ *and* $\chi_U(0.95)$ *based on a moving window approach (similar to Figure 3), while the corresponding shaded areas are 95% theoretical confidence envelopes.*

asymptotic dependence class in the upper tail controlled by the parameter δ_U is less clear, and so we cannot make any firm statements about the limiting joint behavior of BTC and ETH in the upper tail. A benefit of our proposed model is that it can account for the uncertainty of the asymptotic dependence class, and it can estimate it without making any prior assumptions.

5.2. *Time-varying estimation of extremal dependence.* While the models fitted in Section 5.1 already provide helpful insights into the asymmetric tail dependence structure of BTC and ETH, this analysis (based on a stationarity assumption) only reveals the "time-averaged" dependence behavior between these cryptocurrencies, thus lacking important information about any changes that might have occurred during the study period. Since ETH is a much more recent cryptocurrency than BTC and ETH was still very "immature" in early 2016, we might expect that their tail dependence structure has evolved over time. Similar considerations hold for the other pairs of cryptocurrencies. In our preliminary exploratory analysis reported in Figure 3 and the Supplementary Material (Gong and Huser (2022a)), empirical results indeed strongly suggest that cryptocurrencies have become more interdependent over the training period, especially in terms of their joint extremes (both gains or losses); see also the two right-most panels of Figure 10 for the pair BTC–ETH. However, whether a regime shift from AI to AD has truly occurred is not clear from purely nonparametric estimates of the tail coefficients $\chi_{L;i}(0.05)$ and $\chi_{U;i}(0.95)$, $i = 1, \ldots, n$. In order to assess this more precisely, we now fit our copula model (3.3) to the training BTC–ETH dataset, using the weighted local likelihood approach outlined in Section 4.2, with a biweight kernel $\omega_\tau(h) = \{1 - (h/\tau)^2\}_+^2$ and bandwidth $\tau = 500$ (as in the simulation study). Given that local estimation approaches always suffer from small effective sample sizes (thus increased variability), we here consider the local likelihood (4.3) with full (rather than censored) likelihood contributions, and we specifically choose our model (3.3) for its parsimony and tail flexibility. The bandwidth $\tau = 500$ was chosen to provide reasonably smooth estimates and performed well in our simulation study. Moreover, notice that, despite this quite large bandwidth, the estimates at a given point in time will be mostly influenced by observations in the relatively near past or future, since the biweight function $\omega_\tau(h)$ decays to zero as $h \to \pm\tau$ with $\omega_\tau(h) \approx 0.5$ when $h = 270$.

Figure 10 displays the time-varying parameter estimates $\widehat{\boldsymbol{\theta}}_i = (\widehat{\delta}_{L;i}, \widehat{\delta}_{U;i}, \rho_i)^\top$ (red dashed lines in the three leftmost panels) as well as the resulting time-varying tail coefficients
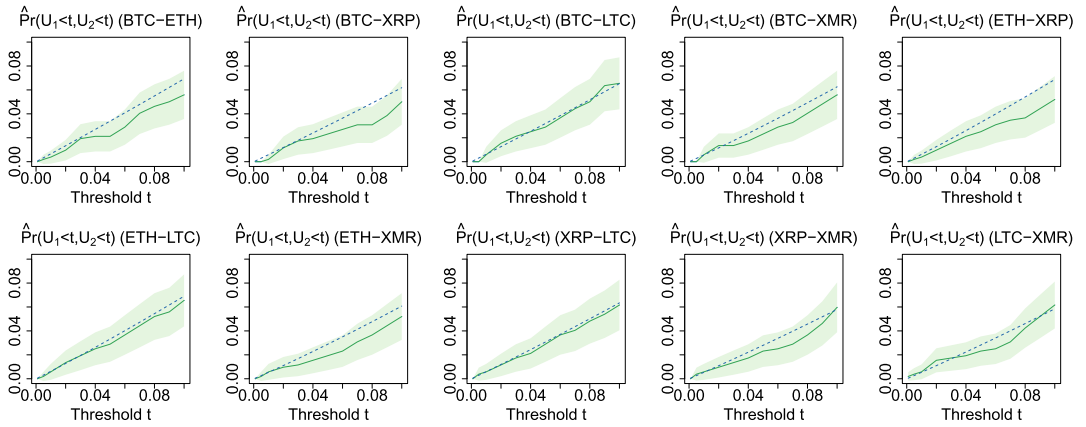
FIG. 11. *Joint lower-tail probabilities* $\Pr(U_1 < t, U_2 < t)$, $t \in (0, 0.1)$, *estimated for all pairs of cryptocurrencies (different panels). Dashed lines are model-based estimates obtained from the local fit of the April 29, 2020, considering the bootstrap median as estimated parameters. Dark solid lines are empirical estimates based on the validation dataset (from April 30, 2020, to October 5, 2021), while light shaded areas display a 95% confidence envelopes obtained by the Delta method.*

$\widehat{\chi}_{L;i}(0.05)$ and $\widehat{\chi}_{U;i}(0.95)$, $i = 1, \ldots, n$. The horizontal grey lines in the plots of $\widehat{\delta}_{L;i}$ and $\widehat{\delta}_{U;i}$ represent the critical threshold of 0.5, defining the boundary between AI and AD regimes.

While the upper tail parameter $\widehat{\delta}_{U;i}$ is fairly constant and often around 0.5 or below (implying AI), the lower tail parameter $\widehat{\delta}_{L;i}$ is quite low and remains below 0.5 until early 2017, before quickly rising around mid 2017 and reaching the level of $\widehat{\delta}_{L;i} \approx 0.8$ (implying AD) in 2018. The lower joint tail of ETH and BTC has thus transitioned from an AI regime to an AD regime. Similar patterns emerge in the tail coefficients $\widehat{\chi}_{L;i}(0.05)$ and $\widehat{\chi}_{U;i}(0.95)$. Interestingly, this fast regime switch coincides with the 2017 boom, while the strong dependence period coincides with the 2018 cryptocurrency crash and the period of high market stress.

Furthermore, the corresponding $\widehat{\chi}_{L;i}(0.05)$ and $\widehat{\chi}_{U;i}(0.95)$ (red dashed lines) in the last two plots follow quite well the nonparametric time-varying empirical estimates of tail coefficients, providing evidence that the model is able to capture the true tail dependence structures accurately. Overall, our results, therefore, agree with Feng, Wang and Zhang (2018) who found that systemic extreme risks in cryptomarkets have grown considerably in recent years. We expect that our analysis, when extended to other cryptocurrencies, might be helpful to investors who want to build a resilient portfolio through diversification. The full results of all 10 pairs of cryptocurrencies under study can be found in the Supplementary Material (Gong and Huser (2022a)).

5.3. *Model validation based on the most recent historical period.* In order to assess the out-of-sample performance of our proposed copula model (3.3) in a challenging, strongly volatile situation of high market stress and to see if it can accurately predict joint tail risk in the near future, we now estimate the joint tail probabilities $\Pr(U_1 < t, U_2 < t)$, $t \in (0, 0.1)$, and $\Pr(U_1 > t, U_2 > t)$, $t \in (0.9, 1)$, for all 10 pairs of cryptocurrencies considered, based on the estimated models obtained from the local likelihood fit of April 29, 2020, that is, on the last day of the training period. We then compare these model-based estimates with empirical probability estimates obtained from the validation dataset, that is, from April 30, 2020, to October 5, 2021, under the working assumption of stationarity. Figures 11 and 12 show the results for the lower and upper tails, respectively. Each panel in these figures corresponds to a different pair of cryptocurrencies. Model-based estimates (obtained from the training dataset) are displayed as dashed lines, while empirical estimates (obtained from the validation dataset)
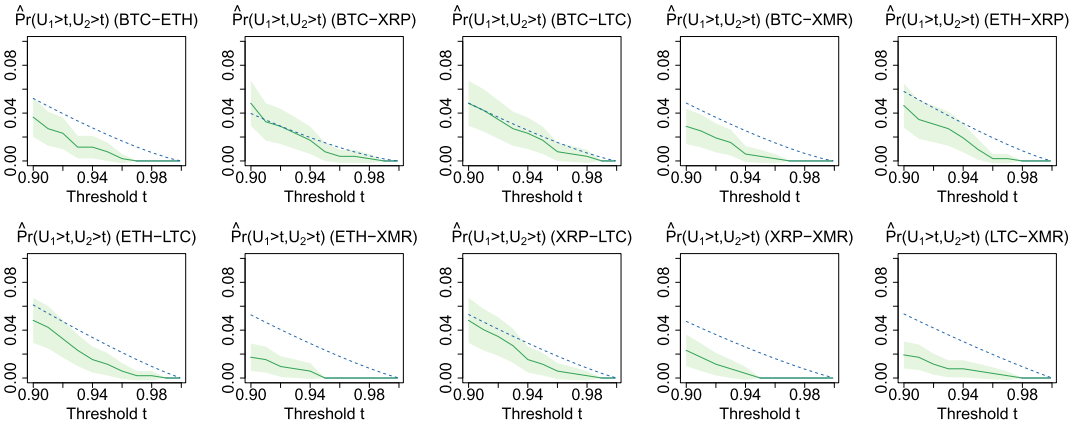
FIG. 12. *Joint upper-tail probabilities* $\Pr(U_1 > t, U_2 > t)$, $t \in (0.9, 1)$, *estimated for all pairs of cryptocurrencies* (*different panels*). *Dashed lines are model-based estimates obtained from the local fit of the April 29, 2020, considering the bootstrap median as estimated parameters. Dark solid lines are empirical estimates based on the validation dataset* (*from April 30, 2020, to October 5, 2021*), *while light shaded areas display a 95% confidence envelopes obtained by the Delta method.*

are displayed as dark solid lines. Light shaded areas correspond to 95% confidence envelopes computed by the Delta method.

For the lower tail (Figure 11), model-based estimates are strikingly similar to empirical estimates. This shows that our proposed model performs very well at capturing the lower tail dependence structure over the validation period and that our methodology is useful to accurately assess tail risk related to downward shocks in the relatively near future. For the upper tail (Figure 12), however, there are some discrepancies between model-based and empirical estimates, showing that our model has a tendency to overestimate the upper-tail dependence structure. This can be explained in two ways: first, the local model fitted at the end of the training period might suffer from mild edge effects (as often the case with local estimation approaches) which might prevent it from accurately capturing the weakening upper-tail dependence behavior around the end of the training period. Second, the upper-tail dependence structure might have significantly weakened over the validation period in a way that our local fit of April 29, 2020, cannot foresee nor predict.

Overall, the out-of-sample performance of our proposed model appears to be reasonable, especially in the left tail. In future research it would be interesting to further investigate whether the lower- and upper-tail dependence structures among cryptocurrencies have again been subject to significant changes in the most recent period.

6. Conclusion. In this paper we have proposed a new parsimonious copula model that possesses high flexibility in both the lower and upper tails. This model bridges asymptotic dependence and independence in the interior of the parameter space which simplifies inference on the extremal dependence class. Our model has similarities with the model proposed by Huser and Wadsworth (2019), but, unlike the latter, it is also very flexible in the lower tail. To the best of our knowledge, it is the first copula model that can capture and separately control both asymptotic dependence and independence in each joint tail with a smooth transition between dependence classes. Inference can be performed by maximum likelihood, using either full likelihood contributions or various types of censored likelihood contributions designed to prioritize calibration in the tails. Furthermore, we have also developed a local likelihood approach that can be used to uncover complex time trends driving the lower- and upper-tail dependence structures.

We have applied our new model to understand the tail dependence dynamics of cryptocurrency market price data, focusing on five leading cryptocurrencies. We have shown that our proposed model, despite its simplicity, outperforms other popular copula models, and we also note that the appealing parsimony of our proposed model becomes a crucial aspect to take into consideration in case of low sample sizes or when a local likelihood inference approach is used, as in our case study. Our analysis suggests that the upper-tail dependence strength has remained relatively stable at a moderate level for most pairs of cryptocurrencies under study, whereas the lower tail, representing the big joint losses, has become more and more dependent in recent years, transitioning from a weak asymptotic independence regime to a strong asymptotic dependence regime in some cases (e.g., Bitcoin–Ethereum). Interestingly, we have found that this regime switch coincides with the fast 2017 boom followed by the 2018 cryptocurrency crash. From a practical perspective our results could help to detect market risk contagion and be a useful source of information for investors who seek to diversify their portfolio. In this paper we analyzed the data until April 29, 2020, which just precedes a major rise in cryptocurrency prices, and we used the most recent market data until October 5, 2021, for model validation. Overall, we have found that our proposed model provides satisfactory performances both in-sample and out-of-sample, especially as far as the lower-tail dependence structure is concerned, and we thus conclude that our methodology can be helpful for assessing tail risk in the near future under high market stress.

We emphasize that our model is useful to analyze the extremal dependence of losses and gains jointly in a single statistical model. As our copula model describes the full range of the distribution (unlike most models for extremes which usually focus on one tail only), it may also be used as a building block for improving existing stochastic financial data simulators.

Although we focused in this paper on the bivariate setting, there is no conceptual problem for generalizing our model to the multivariate or spatial case (by taking a D-dimensional vector W in Section 3), but inference would be more challenging. This opens the door to the joint modeling of multiple cryptocurrencies, although it would be tricky to design a multivariate model with distinct asymptotic dependence regimes among different pairs of variables. Moreover, while we have here assumed that W has a Gaussian or skew-normal copula, it could be replaced by any other copula model that is asymptotically independent in both tails, without affecting the asymptotic tail results. Thus, the model construction is quite general and could be extended to a wide range of more complex and flexible copula models.

APPENDIX A: MARGINAL DISTRIBUTIONS OF OUR MODEL IN THE CASES WHERE $\delta_L = 1/2$ AND/OR $\delta_U = 1/2$

When $\delta_L, \delta_U \neq 1/2$, the marginal distributions of our model (3.3) are given in Section 3.2. The intermediate cases when $\delta_L = 1/2$ and/or $\delta_U = 1/2$ may be established separately or as the limits $\delta_L \to 1/2$ and/or $\delta_U \to 1/2$. When $\delta_L = 1/2$ and $\delta_U \neq 1/2$, we have

$$F_X(x) = \begin{cases} \dfrac{2x}{(1+2\delta_U)(2\delta_U-3)}\exp(2x) - \dfrac{-12\delta_U+12\delta_U^2-5}{(1+2\delta_U)^2(2\delta_U-3)^2}\exp(2x), \\ \quad x \le 0, \\ 1 - \dfrac{4\delta_U^3}{(1+2\delta_U)^2(2\delta_U-1)}\exp\left(-\dfrac{x}{\delta_U}\right) - \dfrac{4(\delta_U-1)^3}{(\delta_U-3)^2(2\delta_U-1)}\exp\left(-\dfrac{x}{1-\delta_U}\right), \\ \quad x > 0; \end{cases}$$

when $\delta_L \neq 1/2$ and $\delta_U = 1/2$, we have

$$
F_X(x) = \begin{cases}
\dfrac{4\delta_L^3}{(1+2\delta_L)^2(2\delta_L - 1)} \exp\left(\dfrac{x}{\delta_L}\right) + \dfrac{4(\delta_L - 1)^3}{(2\delta_L - 3)^2(2\delta_L - 1)} \exp\left(\dfrac{x}{1-\delta_L}\right), \\
\quad x \leq 0, \\
1 + \dfrac{2x}{(1+2\delta_L)(2\delta_L - 3)} \exp(-2x) + \dfrac{-12\delta_L + 12\delta_L^2 - 5}{(1+2\delta_L)^2(2\delta_L - 3)^2} \exp(-2x), \\
\quad x > 0;
\end{cases}
$$

finally, when $\delta_L = \delta_U = 1/2$, we have

$$
F_X(x) = \begin{cases}
\dfrac{1}{2}(1-x)\exp(2x), & x \leq 0, \\
1 - \dfrac{1}{2}(1+x)\exp(-2x), & x > 0.
\end{cases}
$$

APPENDIX B: PROOF OF PROPOSITION 1 ON TAIL DECAY RATES

To prove Proposition 1, we will exploit results on the extremal dependence of random scale constructions from Engelke, Opitz and Wadsworth (2019). In order to apply these results, we need first to put our model (3.3) in random scale form. By taking the exponential on both components of the random vector $X = (X_1, X_2)^\top$, we obtain the vector $\tilde{X} = (\tilde{X}_1, \tilde{X}_2)^\top$ with components $\tilde{X}_1 = \tilde{R}\tilde{W}_1$, $\tilde{X}_2 = \tilde{R}\tilde{W}_2$, where $\tilde{R} = \exp(R)$ and $\tilde{W}_i = \exp(W_i)$, $i = 1, 2$. Notice that, because the exponential is a monotone increasing function, the new random vector $\tilde{X}$ has the same dependence structure (i.e., copula) as X. Now, for $r > 1$, we obtain from (3.1) that $\Pr(\tilde{R} > r) = \Pr\{R > \log(r)\} = \frac{\delta_U}{\delta_L + \delta_U} r^{-1/\delta_U}$ which implies that $\tilde{R}$ is regularly varying at infinity with index $-1/\delta_U$. Similarly, for $w > 1$, $\Pr(\tilde{W}_i > w) = \Pr\{W_i > \log(w)\} = \frac{1-\delta_U}{2-\delta_L-\delta_U} w^{-1/(1-\delta_U)}$ which implies that $\tilde{W}_i$ is regularly varying at infinity with index $-1/(1-\delta_U)$. Moreover, clearly $\Pr(\tilde{W}_i > 0) = 1$, $i = 1, 2$. Furthermore, let $\varepsilon > 0$, and define $\tilde{\varepsilon} = \varepsilon\delta_U > 0$. We have

$$
\mathrm{E}(\tilde{W}_i^{1/\delta_U + \varepsilon}) = \int_0^\infty \Pr(\tilde{W}_i^{1/\delta_U + \varepsilon} > w)\,dw = \int_0^\infty \Pr(\tilde{W}_i > w^{\delta_U/(1+\tilde{\varepsilon})})\,dw
$$

$$
= \underbrace{\int_0^1 \Pr(\tilde{W}_i > w^{\delta_U/(1+\tilde{\varepsilon})})\,dw}_{:=I_1} + \frac{1-\delta_U}{2-\delta_L-\delta_U} \underbrace{\int_1^\infty w^{-\delta_U/\{(1+\tilde{\varepsilon})(1-\delta_U)\}}\,dw}_{:=I_2}.
$$

While the integral I_1 is bounded above by one, the integral I_2 is finite if and only if $\delta_U/\{(1+\tilde{\varepsilon})(1-\delta_U) > 1$, that is, $\delta_U > 1/(2+\tilde{\varepsilon})$. Letting $\varepsilon \to 0$, we conclude from Table 2 of Engelke, Opitz and Wadsworth (2019) that, when $\delta_U > 1/2$, the coefficient of tail dependence of X is $\eta_U = 1$ and

$$
\chi_U = \mathrm{E}\left[\min\left\{\frac{\tilde{W}_1^{1/\delta_U}}{\mathrm{E}(\tilde{W}_1^{1/\delta_U})}, \frac{\tilde{W}_2^{1/\delta_U}}{\mathrm{E}(\tilde{W}_2^{1/\delta_U})}\right\}\right].
$$

This coincides with the results of Proposition 1, Case 2, by plugging $\tilde{W}_i = \exp(W_i)$, $i = 1, 2$. The lower-tail coefficient χ_L can be derived by symmetry when flipping the sign of X in (3.3).

On the other hand, when $\delta_U < 1/2$, then $1/\delta_U > 1/(1-\delta_U)$. Therefore, because $W = (W_1, W_2)^\top$ is Gaussian with correlation ρ (and thus has $\chi_U = 0$ and $\eta_U = (1+\rho)/2$ according to Sibuya (1960) and Ledford and Tawn (1996)), we deduce from Proposition 5 of

Engelke, Opitz and Wadsworth (2019) that the vector X has $\chi_U = 0$ and that the coefficient of tail dependence is equal to

$$\eta_U = \begin{cases} \delta_U/(1 - \delta_U), & \delta_U > (1 + \rho)/(3 + \rho), \\ (1 + \rho)/2, & \delta_U \le (1 + \rho)/(3 + \rho), \end{cases}$$

as needed. The expressions for χ_L and η_L are obtained by symmetry.

The case $\delta_U = 1/2$ can be deduced by applying Proposition 6(3c) of Engelke, Opitz and Wadsworth (2019).

APPENDIX C: CENSORED LIKELIHOOD EXPRESSIONS

In Section 4, we describe censored likelihoods of the form (4.2) and consider three different censoring schemes, illustrated in Figure 6. For illustration, we here detail the censored likelihood contributions for Scheme 3. Assume that the censoring levels for the lower and upper tail are $0 < t_L < t_U < 1$ for both margins, and write the censored likelihood as $L(\boldsymbol{\theta}) = \prod_{j=1}^n L_j(\boldsymbol{\theta})$. Then, the censored likelihood contributions $L_j(\boldsymbol{\theta})$ are

$$L_j(\boldsymbol{\theta}) = \begin{cases} c(u_{j1}, u_{j2}), & j \in A; \\ \partial_1 C(u_{j1}, t_U) - \partial_1 C(u_{j1}, t_L), & j \in B_1; \\ \partial_2 C(t_U, u_{j2}) - \partial_2 C(t_L, u_{j1}), & j \in C_1; \\ C(t_L, t_L) + C(t_U, t_U) - C(t_L, t_U) - C(t_U, t_L), & j \in D, \end{cases}$$

where the sets are $A = \{j = 1, \ldots, n : \{u_{j1} < t_L \text{ or } u_{j1} > t_U\} \text{ and } \{u_{j2} < t_L \text{ or } u_{j2} > t_U\}\}$ (noncensored, NC), $B_1 = \{j = 1, \ldots, n : \{u_{j1} < t_L \text{ or } u_{j1} > t_U\} \text{ and } t_L \le u_{j2} \le t_U\}$ (partially censored, PC$_1$), $C_1 = \{j = 1, \ldots, n : t_L \le u_{j1} \le t_U \text{ and } \{u_{j2} < t_L \text{ or } u_{j2} > t_U\}\}$ (partially censored, PC$_2$), and $D = \{j = 1, \ldots, n : t_L \le u_{j1}, u_{j2} \le t_U\}$ (fully censored, FC). The expressions for the other censoring schemes are similar, although Scheme 1 has two different types of partial censoring likelihoods with $K_B = K_C = 2$ in (4.2) (rather than $K_B = K_C = 1$ for Schemes 2 and 3), and the formula is thus slightly more involved.

SUPPLEMENTARY MATERIAL

Supplement to "Asymmetric tail dependence modeling, with application to cryptocurrency market data" (DOI: 10.1214/21-AOAS1568SUPPA; .pdf). We provide in Gong and Huser (2022a) further details about the skewed version of our Model (3.3), extra simulation results in well-specified and misspecified settings, our treatment of marginal distributions, as well as more details on our cryptocurrency application (detailing in particular the results for all ten pairs of cryptocurrencies under study).

R code for "Asymmetric tail dependence modeling, with application to cryptocurrency market data" (DOI: 10.1214/21-AOAS1568SUPPB; .zip). We provide in Gong and Huser (2022b) the R code to fit our proposed copula model using a full likelihood approach, or any of the censored likelihood approaches proposed in the paper.

REFERENCES

ALCOCK, J. and SATCHELL, S. (2018). *Asymmetric Dependence in Finance*: *Diversification*, *Correlation and Portfolio Management in Market Downturns*. Wiley, New York.

ARELLANO-VALLE, R. B. and GENTON, M. G. (2010). Multivariate extended skew-*t* distributions and related families. *Metron* **68** 201–234. MR3041150 https://doi.org/10.1007/BF03263536

AULBACH, S., BAYER, V. and FALK, M. (2012). A multivariate piecing-together approach with an application to operational loss data. *Bernoulli* **18** 455–475. MR2922457 https://doi.org/10.3150/10-BEJ343

AZZALINI, A. and DALLA VALLE, A. (1996). The multivariate skew-normal distribution. *Biometrika* **83** 715–726. MR1440039 https://doi.org/10.1093/biomet/83.4.715

BORRI, N. (2019). Conditional tail-risk in cryptocurrency markets. *J. Empir. Finance* **50** 1–19.

BROCKWELL, P. J. and DAVIS, R. A. (2002). *Introductin to Time Series and Forecasting*, 2nd ed. Springer, New York.

CASTRO-CAMILO, D., DE CARVALHO, M. and WADSWORTH, J. (2018). Time-varying extreme value dependence with application to leading European stock markets. *Ann. Appl. Stat.* **12** 283–309. MR3773394 https://doi.org/10.1214/17-AOAS1089

CASTRUCCIO, S., HUSER, R. and GENTON, M. G. (2016). High-order composite likelihood inference for max-stable distributions and processes. *J. Comput. Graph. Statist.* **25** 1212–1229. MR3572037 https://doi.org/10.1080/10618600.2015.1086656

COLES, S. G. and TAWN, J. A. (1991). Modelling extreme multivariate events. *J. Roy. Statist. Soc. Ser. B* **53** 377–392. MR1108334

DAVISON, A. C. and HUSER, R. (2015). Statistics of extremes. *Annu. Rev. Stat. Appl.* **2** 203–235.

DE CARVALHO, M., LEONELLI, M. and ROSSI, A. (2020). Tracking change-points in multivariate extremes. ArXiv preprint. Available at arXiv:2011.05067.

DE HAAN, L. and ZHOU, C. (2021). Trends in extreme value indices. *J. Amer. Statist. Assoc.* **116** 1265–1279. MR4309271 https://doi.org/10.1080/01621459.2019.1705307

DEMARTA, S. and MCNEIL, A. J. (2005). The *t* copula and related copulas. *Int. Stat. Rev.* **73** 111–129.

EINMAHL, J. H. J., DE HAAN, L. and ZHOU, C. (2016). Statistics of heteroscedastic extremes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 31–51. MR3453645 https://doi.org/10.1111/rssb.12099

EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events*: *For Insurance and Finance*. *Applications of Mathematics* (*New York*) **33**. Springer, Berlin. MR1458613 https://doi.org/10.1007/978-3-642-33483-2

ENGELKE, S., OPITZ, T. and WADSWORTH, J. (2019). Extremal dependence of random scale constructions. *Extremes* **22** 623–666. MR4031852 https://doi.org/10.1007/s10687-019-00353-3

FENG, W., WANG, Y. and ZHANG, Z. (2018). Can cryptocurrencies be a safe haven: A tail risk perspective analysis. *Appl. Econ.* **50** 4745–4762.

GKILLAS, K., BEKIROS, S. and SIRIOPOULOS, C. (2018). Extreme correlation in cryptocurrency markets. Available at SSRN 3180934.

GONG, Y. and HUSER, R. (2022a). Supplement to "Asymmetric tail dependence modeling, with application to cryptocurrency market data." https://doi.org/10.1214/21-AOAS1568SUPPA

GONG, Y. and HUSER, R. (2022b). R code for "Asymmetric tail dependence modeling, with application to cryptocurrency market data." https://doi.org/10.1214/21-AOAS1568SUPPB

HASHORVA, E. (2010). On the residual dependence index of elliptical distributions. *Statist. Probab. Lett.* **80** 1070–1078. MR2651047 https://doi.org/10.1016/j.spl.2010.03.001

HAZRA, A., REICH, B. J. and STAICU, A.-M. (2020). A multivariate spatial skew-t process for joint modeling of extreme precipitation indexes. *Environmetrics* **31** e2602.

HUSER, R., OPITZ, T. and THIBAUD, E. (2017). Bridging asymptotic independence and dependence in spatial extremes using Gaussian scale mixtures. *Spat. Stat.* **21** 166–186. MR3692183 https://doi.org/10.1016/j.spasta.2017.06.004

HUSER, R. and WADSWORTH, J. L. (2019). Modeling spatial processes with unknown extremal dependence class. *J. Amer. Statist. Assoc.* **114** 434–444. MR3941266 https://doi.org/10.1080/01621459.2017.1411813

HUYNH, T. L. D., NGUYEN, S. P. and DUONG, D. (2018). Contagion risk measured by return among cryptocurrencies. In *International Econometric Conference of Vietnam* 987–998. Springer, Berlin.

KIRILIOUK, A., ROOTZÉN, H., SEGERS, J. and WADSWORTH, J. L. (2019). Peaks over thresholds modeling with multivariate generalized Pareto distributions. *Technometrics* **61** 123–135. MR3933664 https://doi.org/10.1080/00401706.2018.1462738

KRUPSKII, P. (2017). Copula-based measures of reflection and permutation asymmetry and statistical tests. *Statist. Papers* **58** 1165–1187. MR3720957 https://doi.org/10.1007/s00362-016-0743-1

KRUPSKII, P., HUSER, R. and GENTON, M. G. (2018). Factor copula models for replicated spatial data. *J. Amer. Statist. Assoc.* **113** 467–479. MR3803479 https://doi.org/10.1080/01621459.2016.1261712

LEDFORD, A. W. and TAWN, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83** 169–187. MR1399163 https://doi.org/10.1093/biomet/83.1.169

LEONELLI, M. and GAMERMAN, D. (2020). Semiparametric bivariate modelling with flexible extremal dependence. *Stat. Comput.* **30** 221–236. MR4064619 https://doi.org/10.1007/s11222-019-09878-w

MHALLA, L., DE CARVALHO, M. and CHAVEZ-DEMOULIN, V. (2019). Regression-type models for extremal dependence. *Scand. J. Stat.* **46** 1141–1167. MR4033807 https://doi.org/10.1111/sjos.12388

NAKAMOTO, S. (2008). Bitcoin: A peer-to-peer electronic cash system. Available at https://bitcoin.org/bitcoin.pdf.

NGUYEN, L. H., CHEVAPATRAKUL, T. and YAO, K. (2020). Investigating tail-risk dependence in the cryptocurrency markets: A LASSO quantile regression approach. *J. Empir. Finance* **58** 333–355.

PADOAN, S. A., RIBATET, M. and SISSON, S. A. (2010). Likelihood-based inference for max-stable processes. *J. Amer. Statist. Assoc.* **105** 263–277. MR2757202 https://doi.org/10.1198/jasa.2009.tm08577

PATTON, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *J. Financ. Econom.* **2** 130–168.

PATTON, A. J. (2006). Modelling asymmetric exchange rate dependence. *Internat. Econom. Rev.* **47** 527–556. MR2216591 https://doi.org/10.1111/j.1468-2354.2006.00387.x

POON, S.-H., ROCKINGER, M. and TAWN, J. (2003). Modelling extreme-value dependence in international stock markets. *Statist. Sinica* **13** 929–953. MR2026056

POON, S.-H., ROCKINGER, M. and TAWN, J. A. (2004). Extreme value dependence in financial markets: Diagnostics, models, financial implications. *Rev. Financ. Stud.* **17** 581–610.

ROOTZÉN, H., SEGERS, J. and WADSWORTH, J. L. (2018). Multivariate generalized Pareto distributions: Parametrizations, representations, and properties. *J. Multivariate Anal.* **165** 117–131. MR3768756 https://doi.org/10.1016/j.jmva.2017.12.003

ROOTZÉN, H. and TAJVIDI, N. (2006). Multivariate generalized Pareto distributions. *Bernoulli* **12** 917–930. MR2265668 https://doi.org/10.3150/bj/1161614952

SCARROTT, C. and MACDONALD, A. (2012). A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT* **10** 33–60. MR2912370

SIBUYA, M. (1960). Bivariate extreme statistics. I. *Ann. Inst. Statist. Math. Tokyo* **11** 195–210. MR0115241 https://doi.org/10.1007/bf01682329

TAWN, J. A. (1988). Bivariate extreme value theory: Models and estimation. *Biometrika* **75** 397–415. MR0967580 https://doi.org/10.1093/biomet/75.3.397

TAWN, J. A. (1990). Modelling multivariate extreme value distributions. *Biometrika* **77** 245–253.

VETTORI, S., HUSER, R. and GENTON, M. G. (2018). A comparison of dependence function estimators in multivariate extremes. *Stat. Comput.* **28** 525–538. MR3761339 https://doi.org/10.1007/s11222-017-9745-7

VRAC, M., NAVEAU, P. and DROBINSKI, P. (2007). Modeling pairwise dependencies in precipitation intensities. *Nonlinear Process. Geophys.* **14** 789–797.

WADSWORTH, J. L., TAWN, J. A., DAVISON, A. C. and ELTON, D. M. (2017). Modelling across extremal dependence classes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **79** 149–175. MR3597968 https://doi.org/10.1111/rssb.12157

ZHANG, Z., HUSER, R., OPITZ, T. and WADSWORTH, J. L. (2021). Modeling spatial extremes using normal mean-variance mixtures. *Extremes*. To appear.