Semiparametric Bayesian modeling of nonstationary joint extremes: How do big tech's extreme losses behave?

Miguel de Carvalho[1,2] and Karla Vianey Palacios Ramirez[3]

[1] *School of Mathematics, University of Edinburgh, Edinburgh, UK*

[2] *Department of Mathematics–CIDMA, Universidade de Aveiro, Aveiro, Portugal*

[3] *School of Mathematics, Statistics and Physics, Newcastle University, Newcastle, UK*

Summary. Motivated by the hype surrounding AI and big tech stocks, we develop a model for tracking the dynamics of their combined extreme losses over time. Specifically, we propose a novel Bayesian model for inferring about the intensity of observations in the joint tail over time, and for assessing if two stochastic processes are asymptotically dependent. To model the intensity of observations exceeding a high threshold, we develop a Bayesian nonparametric approach that defines a prior on the space of what we define as EDI (Extremal Dependence Intensity) functions. In addition, a parametric prior is set on the coefficient of tail dependence. An extensive battery of experiments on simulated data show that the proposed method are able to recover the true targets in a variety of scenarios. An application of the proposed methodology to a set of big tech stocks—known as FAANG—sheds light on some interesting features on the dynamics of their combined losses over time.

Keywords: FAANG stocks, Mixture of finite Polya trees, Statistics of extremes, Multivariate extreme values, Nonparametric prior, Nonstationary extremal dependence.

1. Introduction

1.1. Data, financial rationale, and applied motivation

The rapid evolution of AI prompts serious concerns about its role in the next financial crisis (Financial Times—Editorial Board, 2023). While new developments offer benefits, some investors fear trading algorithms could cause the next market crash, while others worry an AI bubble—with everything AI-related getting inflated—could lead to a global meltdown. The substantial investments by major corporations in AI offer new opportunities, yet they also increase integration hence raising systemic risk. While many of these companies have been presenting in recent years steady financial results, the risk of another tech bubble like the 2 000 dot-com bubble cannot be ignored.

Motivated by this financial landscape, this paper will shed light on how the combined losses of a set of major AI tech stocks—known as FAANG (Meta's **F**acebook, **A**pple, **A**mazon, **N**etflix and Alphabet's **G**oogle)—has been evolving in recent years. Fig. 1 depicts the raw FAANG data over the period under analysis. Given the importance of FAANG stocks in the financial landscape—attracting everyone from retail investors to professional stakeholders—our empirical findings on how FAANG comove during periods of financial stress may be of broader interest in themselves.

Adding to the risks mentioned earlier, there is also the traditional risk of herd behavior according to which investors might irrationally follow market trends or the actions of peers, potentially leading to inflated asset prices and subsequent market bubbles (Avery and Zemsky, 1998; Chiang and Zheng, 2010; Cipriani and Guarino, 2014). Beyond FAANG,

Address for correspondence: M. de Carvalho, James Clerk Maxwell Building, The King's Buildings Peter Guthrie Tait Road, Edinburgh, EH9 3FD UK.
E-mail: Miguel.deCarvalho@ed.ac.uk

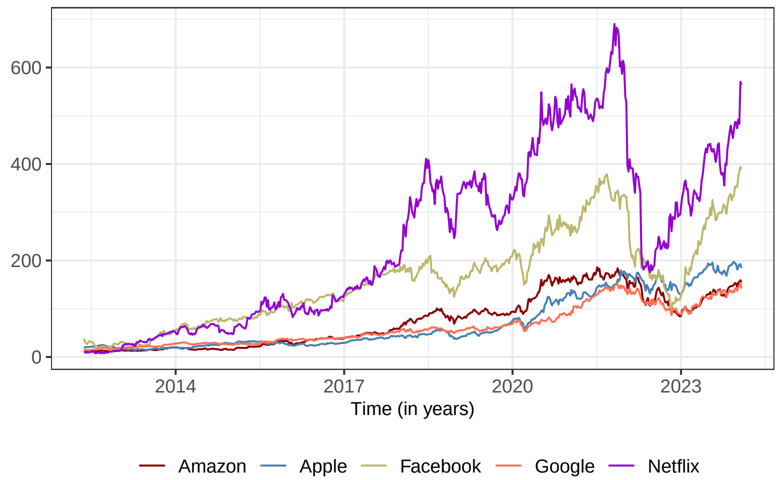

Fig. 1. FAANG prices at close over 2012–2024.

the stock market has embraced catchy names acronyms for top AI tech stocks, like *The Magnificent Seven* (Microsoft, Tesla, Nvidia, and FAANG minus Netflix), and *BAT* (Baidu, Alibaba, Tencent). While these labels streamline discussions, they also oversimplify the tech sector's complexity, prompting 'thinking fast'—akin to Kahneman's System 1 concept (Montier, 2010; Kahneman, 2011)—which is inadequate for complex decisions in settings like financial markets.

Stock market comovements have been widely studied in the financial literature (e.g., Forbes and Rigobon, 2002; Brooks and Del Negro, 2004; Morana and Beltratti, 2008; Rua and Nunes, 2009; Albuquerque and Vega, 2009; Jach, 2017; Ehrmann and Jansen, 2020). Despite this, and the fact that methodologies for modeling conditional copulas are well-known (e.g., Patton, 2006), few attempts (e.g., Poon et al., 2003; Castro et al., 2018; Lee et al., 2024) have been made however to examine the dynamics governing the comovement of *extreme losses* on stock markets. The current paper naturally allows for the latter inquiry as will be put forward by our FAANG data application.

1.2. Background and main contributions

Extreme value theory offers a sound probabilistic and statistical setup for dealing with record-breaking extreme events—such as stock market crashes, widespread flooding, wildfires and heatwaves—given its ability to extrapolate into the tails of a distribution (e.g., Embrechts et al., 1997; Coles, 2001; Beirlant et al., 2004). In a multivariate context, the degree of association between the extreme observations of a random vector with common margins, (X, Y), is often evaluated by,

$$\chi = \lim_{z \to \infty} P(X > z \mid Y > z). \tag{1}$$

The measure χ in (1) quantifies the probability of X being extreme, given that Y is extreme. If $0 < \chi \leq 1$ the variables are asymptotically dependent (AD), whereas if $\chi = 0$ they are said to be asymptotically independent (AI).

In this paper we develop a flexible Bayesian model for learning about the intensity of extreme observations of a random vector over time, as well as for assessing if two stochastic processes are asymptotically dependent. As it will be shown below, our methods have some links with a time-varying version of (1). This paper contributes to the recent literature on nonstationary multivariate extremes that has been developed for tracking how the extremal dependence within a random vector evolves over time or another covariate (e.g., de Carvalho, 2016; Mhalla et al., 2019; Castro et al., 2018; Gong and Huser, 2022). Our focus differs from that of the latter papers in a number of important ways. A key difference is that here the focus will be placed on the intensity of joint extreme observations—via a novel object which we refer to as the EDI (Extremal Dependence Intensity) function—whereas the approaches in the latter papers are mainly based on indexing the parameter of a multivariate extreme value distribution over time. As it will be shown later, the EDI can be interpreted as a measure of intensity for joint extremes, thereby learning about their frequency from the data. Contrarily to the latter approaches on nonstationary multivariate extremes, the proposed method do not require a componentwise block maxima framework and hence are not restricted to asymptotic dependence. Finally, the proposed approach does not require estimating an angular measure or a Pickands dependence function—which need to obey constraints that may be complicated to include in the inferences.

In addition, we develop Bayesian estimators for the two parameters in our framework, that is, the EDI and the coefficient of tail dependence. The proposed approach is suitable for modeling nonstationary joint extremes, as it has been conceived for tracking the dynamics of the intensity of extreme observations in the joint tail, and for assessing if two stochastic processes are asymptotically dependent. The EDI measures the intensity of extreme observations in the joint tail over time, and it is thus a measure of the degree of association of the extremes over time.

To learn the EDI from data, we define a prior on the space of all EDIs using a mixture of finite Polya trees (Müller et al., 2015, Section 3). The proposed inferences for the EDI are fully supported on the unit interval, and do not suffer from boundary bias. In the context of EDIs, our Polya tree-based approach relies on a parametric approach as a baseline model for the EDI, while allowing for deviations from it whenever the data provide evidence for that. As a byproduct, this paper also contributes to the relatively novel literature that interfaces non- and semiparametric Bayesian modeling with extreme value theory (e.g., Kottas et al., 2012; Fuentes et al., 2013; Marcon et al., 2016; Hanson et al., 2017; Padoan and Rizzelli, 2022).

1.3. Structure and organization

Section 2 introduces the proposed modeling framework. Bayesian inference for the proposed approach is developed in Section 3. Section 4 reports the main findings of a Monte Carlo simulation study, and Section 5 showcases an application of the proposed method to FAANG stocks. Finally, Section 6 discusses the main results and concludes the paper.

2. Modeling Time-Changing Joint Extremes

2.1. Framework

To streamline the presentation of the novel concepts to be introduced below, we start with a bivariate framework. Comments on how to proceed when the analysis involves more than two processes are given in Section 2.3. Let $\{(X_t, Y_t) : t \in [0, 1]\}$ be a collection of independent random vectors, and following standard practice in multivariate extreme value theory suppose that $\{X_t\}$ and $\{Y_t\}$ are unit Fréchet distributed; that is, X_s (Y_s) is independent of X_t (Y_t) for any $s < t$, and $P(X_t < z) = P(Y_t < z) = \exp(-1/z)$, with $z > 0$ for all t. Let $\chi(t) = \lim_{z\to\infty} P(X_t > z \mid Y_t > z)$ be a time-varying version of (1). The setup below can be used for modeling 'partially' as well as 'fully' AD processes ($\chi(t) \in (0, 1]$, for some t; or for all t, respectively), and AI processes ($\chi(t) = 0$, for all t).

To embed AI in our framework, we consider the following time-varying version of the assumption of Ledford and Tawn (1996),

$$P(X_t > z, Y_t > z) = \frac{L_t(z)}{z^{1/\gamma}}, \tag{2}$$

where $\gamma \in (0,1]$ is the coefficient of tail dependence and $L_t(z)$ is a time-varying slowly varying function ($L_t(zu)/L_t(z) \to 1$ as $z \to \infty$, for any $u > 0$). The processes $\{X_t\}$ and $\{Y_t\}$ are said to be positively associated at the extremes if $\gamma \in (1/2, 1)$, negatively associated if $\gamma \in (0, 1/2)$, and independent if $\gamma = 1/2$. For the sake of parsimony, we assume that γ is constant over time, and evidence presented in the Supporting Information (Section 2) suggests that this assumption seems reasonable for our case study.

For AD processes, the degree of dependence can be characterized by what we will refer to as the EDI (Extremal Dependence Intensity) function,

$$f(t) = \frac{\chi(t)}{\int_0^1 \chi(\tau)\,\mathrm{d}\tau} = \frac{\lim_{z\to\infty} P(X_t > z, Y_t > z)}{\int_0^1 \lim_{z\to\infty} P(X_\tau > z, Y_\tau > z)\,\mathrm{d}\tau}. \tag{3}$$

The EDI carries information on the intensity of observations in the joint tail over time, defined as $A = [u, \infty)^2 \times [0, t]$. This follows from the fact that for a sufficiently large u,

$$f(t) \propto \lim_{z\to\infty} P(X_t > z, Y_t > z) \approx P(X_t > u, Y_t > u) = \frac{\mathrm{d}\Lambda}{\mathrm{d}t}(A), \tag{4}$$

since the intensity measure

$$\Lambda(A) = E\left(\int_0^t J_\tau\,\mathrm{d}\tau\right) = \int_0^t P(X_\tau > u, Y_\tau > u)\,\mathrm{d}\tau,$$

as $J_\tau = \mathbb{1}_{\{X_\tau > u, Y_\tau > u\}} \sim \mathrm{Bern}\{P(X_\tau > u, Y_\tau > u)\}$, where $\mathbb{1}$ is the indicator function. A flat EDI, $f(t) \propto 1$, indicates a constant intensity of joint extreme observations over time, whereas if the EDI peaks at some period, it provides an indication of an higher intensity of joint extreme observations during that period. In other words, a flat EDI implies a pattern of stationary bivariate extremes, whereas a peak in the EDI implies an increase in the level of extremal dependence during that period. In the context of the current application, the point process of interest is that of joint extreme losses between pairs of stocks. Eq. (4) shows that the EDI, as defined in (3), can be regarded as the standardized intensity function for a given sufficiently large u.

2.2. *Examples of EDIs based on time-varying extreme value copulas*

Next, we illustrate some instances of EDIs based on time-varying extreme value copulas (de Carvalho, 2016; Castro et al., 2018). Recall that the time-varying bivariate extreme value copula is defined as

$$C_t(u, v) = G_t\left(\frac{-1}{\log(1-u)}, \frac{-1}{\log(1-v)}\right), \quad (u, v) \in [0,1]^2.$$

Here, $G_t(x, y)$ is a time-varying bivariate extreme value distribution. That is,

$$G_t(x, y) = \exp\left\{-\ell_t\left(\frac{1}{x}, \frac{1}{y}\right)\right\},$$

for $x, y > 0$, where the time-varying tail dependence function is

$$\ell_t(x, y) = \int_0^1 2\max\{wx, (1-w)y\}\, H_t(\mathrm{d}w), \tag{5}$$

and H_t is a time-varying angular measure that obeys the following moment constraint for all t,

$$\int_0^1 w H_t(\mathrm{d}w) = \frac{1}{2}.$$

Some parametric instances will be considered below to shed light on the interpretation of the EDI function. The coefficient of tail dependence of all models below is $\gamma = 1$; see Heffernan (2000).

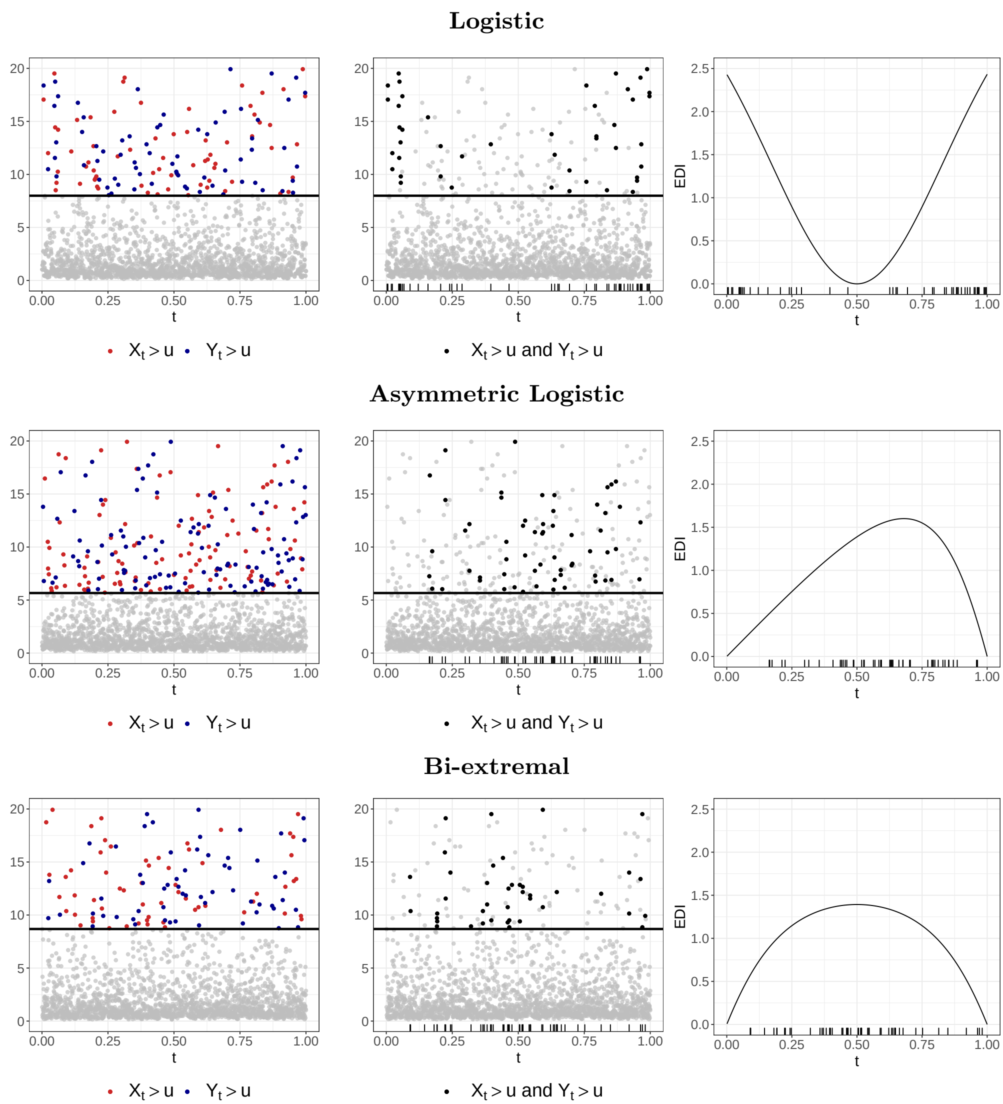

Fig. 2. EDI (Extremal Dependence Intensity) function for time-varying logistic extreme value copulas from Examples 1–2 along with simulated data ($T = 1\,000$). Left: Data above threshold. Middle: Rug of times of exceedances of Z_t above threshold and corresponding exceedances. Right: EDI function.

Example 1 (Logistic). The tail dependence function for the time-varying logistic extreme value copula is $\ell_t(x, y) = (x^{1/\alpha_t} + y^{1/\alpha_t})^{\alpha_t}$, for $x, y > 0$, where $0 < \alpha_t \leq 1$ and $t \in [0, 1]$.

The EDI for this model is

$$f(t) = \frac{2 - 2^{\alpha_t}}{2 - \int_0^1 2^{\alpha_\tau}\,\mathrm{d}\tau}. \tag{6}$$

A more general framework is provided by the following setup.

EXAMPLE 2 (BI-EXTREMAL). The tail dependence function for the time-varying bi-extremal extreme value copula is $\ell_t(x, y) = (1 - \psi_t)x + \{(\psi_t x)^{1/\alpha_t} + y^{1/\alpha_t}\}^{\alpha_t}$, for $x, y > 0$, where $0 < \alpha_t \leq 1$ and $0 \leq \psi_t \leq 1$ with $t \in [0, 1]$. The EDI for this model is

$$f(t) = \frac{1 + \psi_t - \{(\psi_t)^{1/\alpha_t} + 1\}^{\alpha_t}}{1 + \int_0^1 \psi_\tau - \{(\psi_\tau)^{1/\alpha_\tau} + 1\}^{\alpha_\tau}\,\mathrm{d}\tau}. \tag{7}$$

Examples 1 and 2 can be nested in the more general framework of a time-varying asymmetric logistic extreme value copula.

EXAMPLE 3 (ASYMMETRIC LOGISTIC). The tail dependence function for the time-varying asymmetric logistic extreme value copula is $\ell_t(x, y) = (1-\psi_{1,t})x+(1-\psi_{2,t})y+\{(\psi_{1,t}x)^{1/\alpha_t} + (\psi_{2,t}y)^{1/\alpha_t}\}^{\alpha_t}$, for $x, y > 0$, where $0 < \alpha_t \leq 1$ and $0 \leq \psi_{j,t} \leq 1$, for $j = 1, 2$ and $t \in [0, 1]$. The EDI for this model is

$$f(t) = \frac{2 + \psi_{1,t} + \psi_{2,t} - \{(\psi_{1,t})^{1/\alpha_t} + (\psi_{2,t})^{1/\alpha_t}\}^{\alpha_t}}{2 + \int_0^1 \psi_{1,\tau} + \psi_{2,\tau} - \{(\psi_{1,\tau})^{1/\alpha_\tau} + (\psi_{2,\tau})^{1/\alpha_\tau}\}^{\alpha_\tau}\,\mathrm{d}\tau}. \tag{8}$$

Fig. 2 shows the EDI underlying the three examples above along with $T = 1\,000$ simulated data points over $\{t_j \equiv j/T\}_{j=1}^{T}$. We set $\alpha_t = \sin(\pi t)$ for Example 1; also, we set $\alpha_t = 0.5$ and $\psi_t = \sin(\pi t)$ for Example 2; finally, we take $\alpha_t = 0.5$, $\psi_{1,t} = t$ and $\psi_{2,t} = \sin(\pi t)$ for Example 3. As it can be seen from Fig. 2, the more mass the EDI allocates to a period, the higher the degree of extremal dependence between X_t and Y_t.

2.3. Pairwise and multiwise analyses

Section 2.1 covered the case of two stochastic processes. When more than two processes are available, there are two options that complement themselves—the pairwise and the multiwise analyses. Let $\mathbf{Y}_t = (Y_{1,t}, \ldots, Y_{d,t})$, where each $\{Y_{1,t}\}, \ldots, \{Y_{d,t}\}$ is a collection of independent random variables with unit Fréchet margins. The *pairwise analysis* entails applying the principles from Section 2.1 to all $\binom{d}{2}$ pairs. The pairwise structure characterizes the so-called tail-dependence matrix of a d-dimensional vector $\mathbf{Y}_t = (Y_{1,t}, \ldots, Y_{d,t})$ (Embrechts et al., 2016, Definition 3.2) for all t, and it follows from Berman (1961) that $\mathbf{Y}_t$ is asymptotically independent at time t if all pairs $(Y_{i,t}, Y_{j,t})$ are asymptotically independent, with $i \neq j$. Another appealing aspect of the pairwise analysis is that it is rather convenient for visualizations. The pairwise structure is however insufficient to determine the higher order structure (e.g., not much of $P(Y_{1,t} > u, \ldots, Y_{d,t} > u)$ can be learned from from the pairs). Hence, we suggest complementing it with the following *multiwise approach.*

To embed AI in the multiwise setup, we set $Z_t = \min\{Y_{1,t}, \ldots, Y_{d,t}\}$ and additionally consider the following extension of (2),

$$P(Z_t > z) = \frac{L_t(z)}{z^{1/\gamma}}. \tag{9}$$

Independence in the multiwise approach corresponds to the case $\gamma = 1/d$.

For AD processes, the EDI function for that case naturally extends (3) as follows

$$f(t) = \frac{\lim_{z\to\infty} P(Y_{1,t} > z, \ldots, Y_{d,t} > z)}{\int_0^1 \lim_{z\to\infty} P(Y_{1,\tau} > z, \ldots, Y_{d,\tau} > z)\,\mathrm{d}\tau} = \frac{\lim_{z\to\infty} P(Z_t > z)}{\int_0^1 \lim_{z\to\infty} P(Z_t > z)\,\mathrm{d}\tau}. \tag{10}$$

Similarly to Section 2.1, $f(t)$ carries information on the intensity of observations in the joint tail over time, as the argument in (4) can be easily extended for $A = [u, \infty)^d \times [0, t]$, with u large.

3. Bayesian semiparametric inference for time-changing joint extremes

Our Bayesian approach is semiparametric, as it entails setting a prior on the coefficient of tail dependence and on the EDI. The proposed model can be completely characterized by the parameters $(F, \gamma) \in \mathscr{F} \times (0, 1]$, where $F(t) \equiv \int_0^t f(u)\,\mathrm{d}u$ is the cumulative EDI, and $\mathscr{F}$ is the space of all continuous distribution functions supported over the unit interval. Similarly to Poon et al. (2003), we suggest to first estimate γ and only if there is evidence in favor of asymptotic dependence do we infer about f.

3.1. Bayesian inference for the coefficient of tail dependence

Consider the standardized exceedances $\{E_1, \ldots, E_k\} = \{Z_t/u : Z_t > u\}$, for a sufficiently large u. Then, it follows from (9) that $P(Z_t/u > z \mid Z_t > u) \approx z^{-1/\gamma}$; that is, for a sufficiently large u the likelihood of the standardized exceedances $E = (E_1, \ldots, E_k)^{\mathrm{T}}$ is approximately that of a standard Pareto distribution. To conduct Bayesian inference for the coefficient of tail dependence γ, we need to set a prior on $(0, 1]$. Our prior consists of the mixture of a distribution supported on $(0, 1)$ along with a point mass at $\{1\}$ to induce shrinkage if there is evidence in favor of asymptotic dependence. This motivates the following hierarchical model,

$$\begin{cases} \mathsf{p}(E \mid \gamma) = \gamma^{-k} \prod_{j=1}^{k} E_j^{-(1+1/\gamma)} & (\textbf{Likelihood}) \\ \mathsf{p}(\gamma \mid \pi) = \pi \mathbb{1}_{\{1\}}(\gamma) + (1-\pi)\beta(\gamma; a_\gamma, b_\gamma), \quad \mathsf{p}(\pi) = \beta(\pi; a_\pi, b_\pi) & (\textbf{Prior}) \end{cases} \tag{11}$$

where $\beta(\cdot; a, b)$ is the density of a Beta distribution with parameters $a, b > 0$. For details on the posterior sampling algorithm see the Appendix.

3.2. Polya tree-based inference for the EDI function

Preparations

Bayesian inference for the EDI function involves defining a prior over $\mathscr{F}$. Our prior consists of a mixture of finite Polya trees (Hanson, 2006), and thus we start with some preparations on Polya trees (Lavine et al., 1992, 1994; Hanson and Johnson, 2002; Hanson, 2006; Christensen et al., 2008). It is well known that Polya trees can be regarded as random histograms (e.g., Rodriguez and Müller, 2013, Chapter 4), and like histograms they also involve bins. EDIs are supported on the unit interval after standardizing time, hence here we focus on distributions supported on $[0, 1]$. Since the EDI can be understood as a standardized intensity function, in our context bins correspond to subperiods of time.

To illustrate how Polya trees provide a natural extension of parametric models, below we consider the process of generalizing a family of distributions on the unit interval via a Polya tree; other statistical models can be generalized in a similar fashion (e.g., Christensen et al., 2008). A Polya tree entails a number of stages (J), a centering distribution function ($F_{0,\theta}$), and each stage involves a certain number of parameters. In the first stage, the unit interval is partitioned into two bins, $B_{1,1} = (0, m]$ and $B_{1,2} = (m, 1)$, where m is the median of the centering distribution. See Fig. 3 for the case of a Beta(5, 2) centering distribution, which will be used as a running example.

Let T_1 follow the first stage distribution; the parameters of the first stage quantify the amount of mass of T_1 that lies below and above the median of the centering distribution, that is

$$p_{1,1} \equiv P(T_1 \in B_{1,1}), \quad p_{1,2} \equiv P(T_1 \in B_{1,2}) = 1 - p_{1,1}.$$

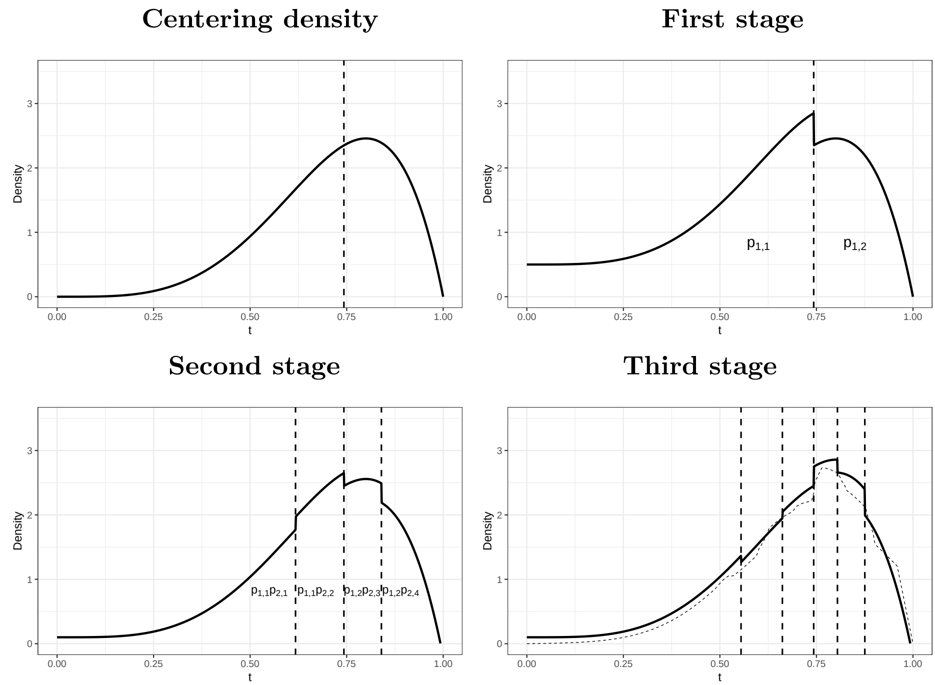

Fig. 3. Example of Polya tree densities centred at a Beta$(5, 2)$ density over stages 1–3; the third stage also shows a mixture of Polya trees mixing over $a \sim$ LN$(\log 2, .05)$ and $b \sim$ LN$(\log 5, .05)$. The dashed line represents the quantiles defining the bins.

Let's now move to the second stage; let T_2 follow the second stage distribution. We proceed as in the first stage but break the unit interval into four pieces with equal mass $B_{2,1} = (0, q_1]$, $B_{2,2} = (q_1, m]$, $B_{2,3} = (m, q_3]$, $B_{2,4} = (q_3, 1)$, where q_1 and q_3 are respectively the first and third quartiles of the centering distribution. The parameters of the second stage are conditional probabilities given the bins of the first stage, that is,

$$\begin{aligned} p_{2,1} &\equiv P(T_2 \in B_{2,1} \mid T_2 \in B_{1,1}), \quad p_{2,2} \equiv P(T_2 \in B_{2,2} \mid T_2 \in B_{1,1}), \\ p_{2,3} &\equiv P(T_2 \in B_{2,3} \mid T_2 \in B_{1,2}), \quad p_{2,4} \equiv P(T_2 \in B_{2,4} \mid T_2 \in B_{1,2}). \end{aligned}$$

The probability of each bin on the second stage is then

$$\begin{aligned} P(T_2 \in B_{2,1}) &= p_{1,1}p_{2,1}, \quad P(T_2 \in B_{2,2}) = p_{1,1}p_{2,2}, \\ P(T_2 \in B_{2,3}) &= p_{1,2}p_{2,3}, \quad P(T_2 \in B_{2,4}) = p_{1,2}p_{2,4}. \end{aligned}$$

See Fig. 3 for a blueprint of these stages; the subsequent stages extend analogously, and in general we would consider a sequence of bins (nested partitions), $\Pi = \{\Pi_j; j = 1, \ldots, J\}$, such that the jth level, $\Pi_j = \{B_{j,l} : l = 1, \ldots, 2^j\}$, partitions the unit interval using the quantiles of the centering as above.

Extending the principles and ideas discussed above, in a Polya tree the conditional distribution of the bins in the jth stage, given the bins of the previous stage, is

$$p_{j,2l-1} \equiv P(T_j \in B_{j,2l-1} \mid T_j \in B_{j-1,l}), \quad p_{j,2l} \equiv P(T_j \in B_{j,2l} \mid T_j \in B_{j-1,l}),$$

for every j and l. These conditional probabilities verify $p_{j,2l-1} + p_{j,2l} = 1$. Since the parameters of a Polya tree are all probabilities it is common to use independent Beta priors,

$p_{j,2l-1} \sim \mathrm{Beta}(\alpha_{j,2l-1}, \alpha_{j,2l})$. Below, we set $\alpha_{j,l} = \alpha j^2$ as this guarantees an absolutely continuous P with probability one in an infinite tree (Kraft, 1964).

The description above is for a Polya tree prior. A mixture of finite Polya trees completes the model by adding a prior on θ. This has the effect of introducing randomness on the starting and endpoints of the bins, which then smooths out the jumps noticeable in stages 1–3 of Fig. 3.

Bayesian inference for the EDI function via mixture of finite Polya trees
Inference is conducted by assuming that a realization of the process $\{Z_t\}$ is observed over a grid on the unit interval, $\mathbb{T} = \{t_1, \dots, t_T\}$. Let $I = \{t \in \mathbb{T} : Z_t > u\} = \{\tau_1, \dots, \tau_k\}$ be the times of k joint observations exceeding a high threshold. The hierarchical representation of our model for the EDI is as follows

$$I \mid F \sim F, \quad F \mid \theta, \Pi \sim \mathrm{PT}_J(\alpha, F_{0,\theta}), \quad \theta \sim \mathsf{p}(\theta). \tag{12}$$

Here, $\mathrm{PT}_J(\alpha, F_{0,\theta})$ is a Polya tree with two parameters: A centering cumulative EDI $(F_{0,\theta}(t))$; a precision parameter $(\alpha > 0)$. The parameter α controls how much deviations from the centering will be allowed, in the sense that the smaller the α the more we allow for deviations from the centering distribution. Following Hanson (2006), the Polya tree EDI density is

$$f(t \mid \Pi, \theta) = 2^J F(B_{J,l(t)} \mid \Pi, \theta) f_{0,\theta}(t), \tag{13}$$

where $f_{0,\theta} = \mathrm{d}F_{0,\theta}/\mathrm{d}t$, and where $l(t) \in \{1, \dots, 2^J\}$ identifies the bin at level J containing $t \in (0,1)$. Note that (13) implies that the Polya tree EDI, f, is a suitably 'tilted' version of the centering EDI, $f_{0,\theta}$. Some final comments on posterior sampling are in order. MCMC can be used to sample all parameters. For the conditional probabilities of the bins, a full conditional is available, allowing Gibbs sampling to be used. That is, the parameters α and θ can be updated by Metropolis–Hastings, whereas the conditional probabilities can be updated through,

$$p_{j,2l-1} \mid \{\tau_1, \dots, \tau_k\}, \alpha, \theta \sim \mathrm{Beta}(\alpha j^2 + k_{j,2l-1}, \alpha j^2 + k_{j,2l}),$$

where $k_{j,l}$ is the number of observations from I that lie on the jth bin, $B_{j,l}$, for $j = 1, \dots, J$ and $l = 1, \dots, 2^{j-1}$.

3.3. Preprocessing margins

The starting point of Section 2 has been that the random vector $\mathbf{Y}_t$ has unit Fréchet marginal distributions; in practice, this is achieved by using a suitable transformation of the raw process $\{\mathbf{R}_t = (R_{1,t}, \dots, R_{d,t})\}$ that sets

$$\mathbf{Y}_t = -(1/\log\{\mathbb{F}_{1,t}(R_{1,t})\}, \dots, 1/\log\{\mathbb{F}_{d,t}(R_{d,t})\}), \tag{14}$$

where $\mathbb{F}_{i,t}(r) = P(R_{i,t} \leq r)$, for all i. Below, we mainly focus on the case where marginals of $\mathbf{R}_t$ are time invariant (i.e. $\mathbb{F}_i \equiv \mathbb{F}_{i,t}$, for all i) as time invariance on the margins is sensible for applied settings such as the one examined in Section 5, and thus the practical implementation of mapping data into unit Fréchet distribution via (14) can be easily achieved by either using the empirical distribution function or a mixture of Polya trees for learning about $\mathbb{F}_i$. Still, for situations where there is evidence of nonstationary margins, one may always implement (14), converting the raw data to unit Fréchet margins using a time-varying distribution function estimator (e.g., Harvey and Oryshchenko, 2012; Nieto-Barajas et al., 2012) to learn about $\mathbb{F}_{i,t}$.

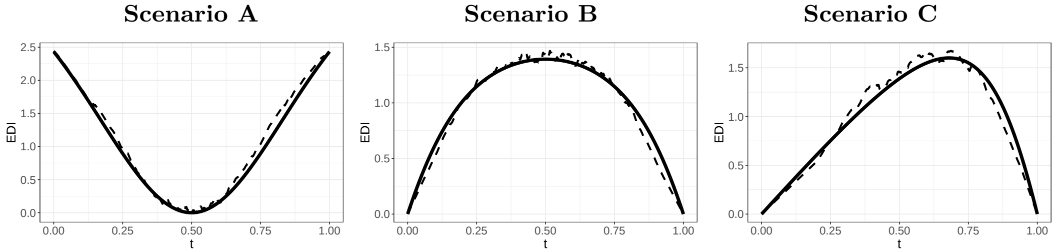

Fig. 4. Single sample experiment ($T = 1\,000$): Posterior median EDI obtained via the mixture of finite Polya trees over a single sample experiment (dashed) plotted against true (solid).

4. Numerical experiments on simulated data

4.1. Data generating processes and preliminary analysis

We now assess the performance of the proposed method using simulated data. First, we illustrate the methods on a single sample experiment and describe the simulation scenarios; a Monte Carlo simulation study will be presented in Section 4.2. The scenarios under which data are generated stem from Examples 1–3. Similarly to Fig. 2, $T = 1\,000$ observations over $\{t_j \equiv j/T\}_{j=1}^{T}$ are simulated from:

- **Scenario A**: Logistic extreme value copula from Example 1, with $\alpha_t = \sin(\pi t)$.
- **Scenario B**: Bi-extremal copula from Example 2, with $\alpha_t = 0.5$, and $\psi_t = \sin(\pi t)$.
- **Scenario C**: Asymmetric logistic extreme value copula from Example 3, with $\alpha = 0.5$, $\psi_{1,t} = t$, and $\psi_{2,t} = \sin(\pi t)$.

We have transformed the simulated data to unit Fréchet margins using the empirical distribution function. Some comments on learning about the EDI via a mixture of Polya trees are in order. We use a Beta distribution as the baseline distribution, that is $F_{0,\theta} = \text{Beta}(a, b)$, and we set $a \sim \text{Log-normal}(m_0, s_0)$ and $b \sim \text{Log-normal}(\tau_1, \tau_2)$; finally, we set $\alpha \sim \text{Gamma}(a_0, b_0)$. In terms of J, while earlier literature suggested rules such as $J = \lceil \log_2 k \rceil$ (Hanson and Johnson, 2002), it is by now well known that setting J around 5–8 provides identical inferences as larger values of J, regardless of k (e.g., Cipolli and Hanson, 2017). Keeping this in mind, we set $J = 8$. In terms of hyperparameters, we set $a_0 = 0.1$, $b_0 = .1$, $J = 8$, $m_0 = 0.5$, $s_0 = 1$, $\tau_1 = .01$, and $\tau_2 = .01$ for the EDI and $a_\gamma = 1$, $b_\gamma = 1$, $a_\pi = 1$ and $b_\pi = 1$ for the coefficient of tail dependence. For the threshold, u in Eq. (4), we consider the 0.95 quantile of $\min(X_t, Y_t)$ and run a burn-in period of $5\,000$ iterates, after which we saved $5\,000$ posterior iterates.

The outcome of a single-run experiment conducted according to the settings above is presented in Fig. 4. Such experiment allows us to anticipate some strengths and limitations of the proposed method. As it can be seen from this figure, our estimator is overall close to the true EDI and it thus captures the intensity of joint extreme observations. Also, the proposed Polya tree-based method does not suffer from boundary bias neither at $\{0\}$ nor at $\{1\}$.

A final comment is in order. For Scenarios A–C the true coefficient of tail dependence is $\gamma = 1$. We include in the Supporting Information (Section 1.2) Scenarios D and E for which $\gamma < 1$. As can be seen from the Supporting Information, the performance of the proposed method is also satisfactory for that setting, especially for higher T.

4.2. Monte Carlo simulation study

Here we report the main numerical findings from a Monte Carlo simulation study. We simulated $1\,000$ time series of length $T = 500, 1\,000, 5\,000$, and $10\,000$ from Scenarios A–

Table 1. Monte Carlo MISE (Mean Integrated Squared Error) for Scenarios A–C for the Polya tree-based EDI.

	Sample size (T)			
Scenario	500	1 000	5 000	10 000
A	0.0207	0.0182	0.0120	0.0103
B	0.0152	0.0106	0.0070	0.0057
C	0.0265	0.0126	0.0089	0.0080

C introduced in Section 4.1. Fig. 5 shows the EDI estimates obtained with the proposed method for Scenarios A–C over this Monte Carlo simulation study; we used the same prior as in Section 4.1, and have also transformed once more the simulated data to unit Fréchet margins using the empirical distribution function.

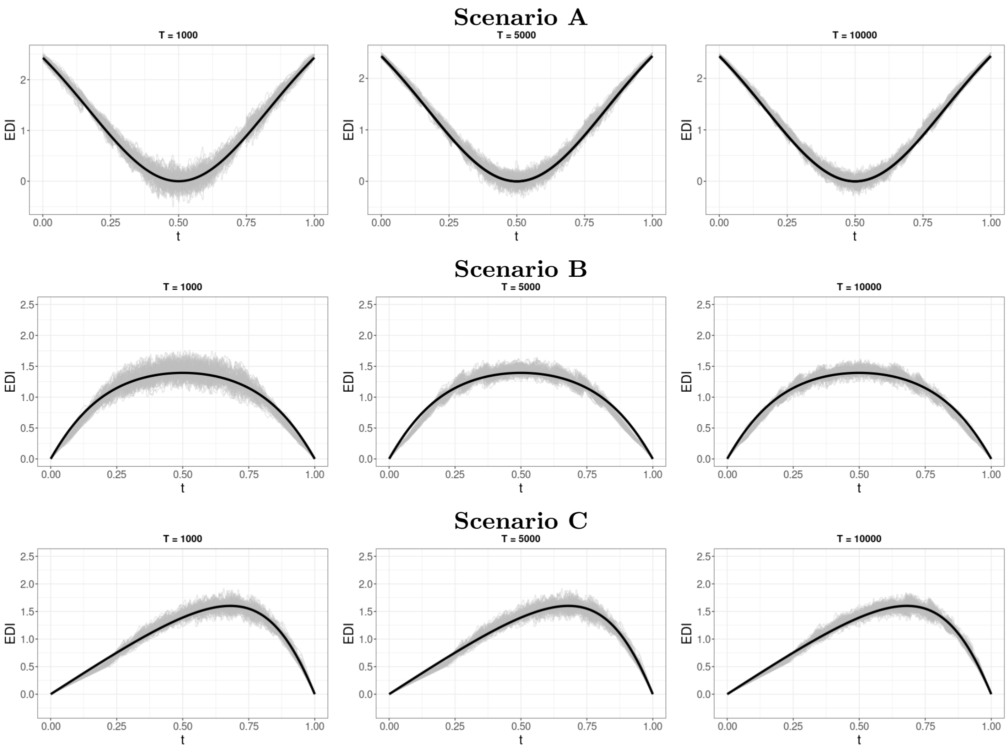

Fig. 5. Monte Carlo simulation study ($T = 1\,000$): 150 randomly selected posterior median EDI density estimates, resulting from the the Monte Carlo simulation study, obtained via a mixture of finite Polya trees (gray), compared with the true EDI (black).

We start with the EDI function. Fig. 5 suggests that the proposed Polya tree-based estimator for the EDI function performs well over Scenarios A–C, in line with the preliminary experiments from Section 4.1. Next, we move to the Bayesian estimator of the coefficient of tail dependence from Section 3. The Monte Carlo posterior median estimates reported in Table 2 suggest an overall good accuracy of the proposed Bayesian approach for learning about the coefficient of tail dependence. In an attempt to examine the frequentist properties of our Bayesian method from a numerical viewpoint, we also report in Table 2 the coverage probabilities, i.e. the number of times the true value γ was contained in the credible interval. As it can be seen from Table 2 the coverage probabilities reasonably follow the significance levels, especially for higher T, thus suggesting a good frequentist behavior of the proposed

Table 2. Monte Carlo posterior median coefficient of tail dependence and coverage probabilities I_x at levels $x = 0.90, 0.95$. For all scenarios γ = 1.

Scenario	Sample size (T)	Monte Carlo posterior median ($\hat{\gamma}$)	$I_{.90}$	$I_{.95}$
A	500	0.985	0.998	1.000
B	500	0.985	0.996	0.998
C	500	0.967	0.996	0.996
A	1 000	0.994	1.000	0.999
B	1 000	0.989	0.999	1.000
C	1 000	0.969	0.998	0.996
A	5 000	0.999	1.000	1.000
B	5 000	0.998	1.000	1.000
C	5 000	0.983	0.990	0.999
A	10 000	1.000	1.000	1.000
B	10 000	0.999	0.989	1.000
C	10 000	0.995	0.993	1.000

Bayesian approach. The appealing frequentist performance of the proposed method is reinforced by the Monte Carlo evidence from Table 1 that shows that the MISE (Mean Integrated Squared Error) decreases as the sample size increases across all simulation scenarios. In the Supporting Information (Section 1.2), we report an additional Monte Carlo experiment that assesses how performance varies when the dimension of the multivariate vector increases. As expected, for a fixed sample size, the accuracy of the fits is higher in the bivariate case.

5. Tracking extreme joint losses of FAANG stocks

5.1. Financial context and preprocessing

In this section we apply the proposed method to track the dynamics governing extreme joint losses of FAANG stocks. These stocks trade on the NASDAQ stock market and have attracted retail investors, money managers, and other professional stakeholders. For example, Warren Buffett was a key financial player investing on these stocks (Apple) in 2019; see the 2019 Hathaway's 13-F filing available from the Securities and Exchange Commission (SEC) webpage (`www.sec.gov`). The data were gathered from Yahoo Finance and consist of weekly closing prices from 1 Jan. 2012 to 1 Feb. 2024; this is mostly a period of sustained growth, and with a few sharp sell-offs over the COVID-19 era. The period under analysis also includes the beginning of the COVID-19 era, which some speculate will get these big tech companies to become even bigger (Wigglesworth, 2020).

Fig. 1 depicts the raw data. Since the focus of the analysis is on extreme losses, we use weekly negative returns as a unit of analysis. Some comments on preprocessing are in order. We transform the bivariate returns (R_t^X, R_t^Y) to unit Fréchet marginals (X_t, Y_t) using the transformation:

$$(X_t, Y_t) = -(1/\log G(R_t^X), 1/\log H(R_t^Y)),$$

where G and H are the respective marginal distribution functions for R_t^X and R_t^Y. We then work with exceedances of $Z_t = \min(X_t, Y_t)$ above the 0.95 quantile. We estimate G and H using a suitably rescaled empirical distribution function, that is $\widehat{G}(x) = 1/(T+1) \sum_{t=1}^{T} \mathbb{1}(R_t^X \leq x)$ with $\widehat{G}$ being analogously defined. Ljung–Box tests (Tsay, 2002, Chapter 2) were applied to the Z_t and no evidence in favor of Z_t being seriously correlated was found.

5.2. Learning about the dynamics of pairwise extreme losses

As noted in Section 3, similarly to Poon et al. (2003), we first estimate γ and only if there is evidence in favor of asymptotic dependence we estimate f. We recall that our prior for the coefficient of tail dependence includes a point mass at 1 to induce shrinkage if there is evidence in favor of asymptotic dependence. The obtained posterior summaries for the coefficient of tail dependence for all pairs of stocks are presented in Table 5.2 and suggest evidence in favor of asymptotic dependence for each pair of stocks. Given this, we next proceed to learn about the EDI. Bayesian inference for the EDI for all pairs of FAANG stocks is presented in Fig. 6. Some comments on implementation are in order. As in Section 4 we assume the number of levels to be $J = 8$. In terms of prior information, for the parameter $\alpha \sim \text{Gamma}(a_0, b_0)$ we use a non-informative prior $(a_0, b_0) = (0.1, 0.1)$, whereas for the parameter of the centring Beta distribution we set $a \sim \text{Log-normal}(\hat{\mu}_z, \hat{\sigma}_z)$ and $b \sim \text{Log-normal}(0.1, 0.1)$, where $\hat{\mu}_z$ and $\hat{\sigma}_z$ are respectively the sample mean and standard deviation. In terms of MCMC, we run a burn-in period of 5 000 iterates, after which we saved 5 000 posterior iterates. As can be seen from Fig. 6, most EDIs tend to peak around 2018–22 thus indicating that extreme joint losses have occurred mostly around that time. This aligns with several noteworthy financial episodes that took place over this turbulent period. Firstly, notable sell-offs occurred in 2018, attributed in part to regulatory concerns, with the tech sector already perceived at that time as vulnerable to herd investing (Bullock et al., 2018; Bullock and Williams, 2018). Secondly, the stock markets faced considerable turbulence in 2020–21 due to COVID-19, marked by events such as Black Monday (I–II) and Black Thursday. Thirdly, the onset of the pandemic era triggered various supply chain bottlenecks paving the way for a rising inflation rise and a decreased demand. The analysis presented above is for the raw returns. Yet, it is well known that returns can display dependence through higher moments, for example, via the volatility clustering phenomenon (e.g., Mandelbrot and Mandelbrot, 1997; Jondeau et al., 2007) Motivated by this, we performed the same analysis after prewhitening the returns using a GARCH (Generalized Autoregressive Conditional Heteroskedasticity) model; the results are presented in the Supporting Information (Section 2.2). The main empirical findings on the key dynamics of extremal dependence are largely consistent with those reported above, though as expected the resulting EDIs show slight variations.

In the Supporting Information we additionally comment on some links between the EDI and the subperiod estimator of Poon et al. (2003, Section 3.3.2) for $\chi(t)$. While the subperiod estimator Poon et al. is not a fair comparison to the EDI (by definition, $f(t) = \chi(t)/\int_0^1 \chi(\tau)\,\mathrm{d}\tau$) there are some links between the two that can easily established through an histogram estimator of the EDI. This is clarified through conceptual considerations and Monte Carlo evidence provided in the Supporting Information.

Finally, while the analysis in Fig. 6 offers an insightful post-mortem outlook on the frequency of joint extreme losses, an important practical question arises regarding predictive insights for the future. The future evolution of tail-dependence structure can be predicted by treating each MCMC trajectory of the EDI as a stochastic process to be forecasted. The forecasted EDI can be found in Supporting Information (Section 2.5), along with further technical details.

5.3. Multiwise analysis

Section 5.2 offered a pairwise analysis but in practice the interest often lies in more than two stocks. Here we focus on $d = 5$ FAANG stocks and conduct a multiwise analysis following the principles from Section 2.3. The reported analysis used the same prior and MCMC setup as in Section 5.2. Fig. 7 depicts the EDI associated with such time-varying minimum computed for all FAANG stocks. The EDI of this Z_t again showcases that the frequency of extreme joint losses was actually higher around late 2018. This is perhaps not surprising give that as

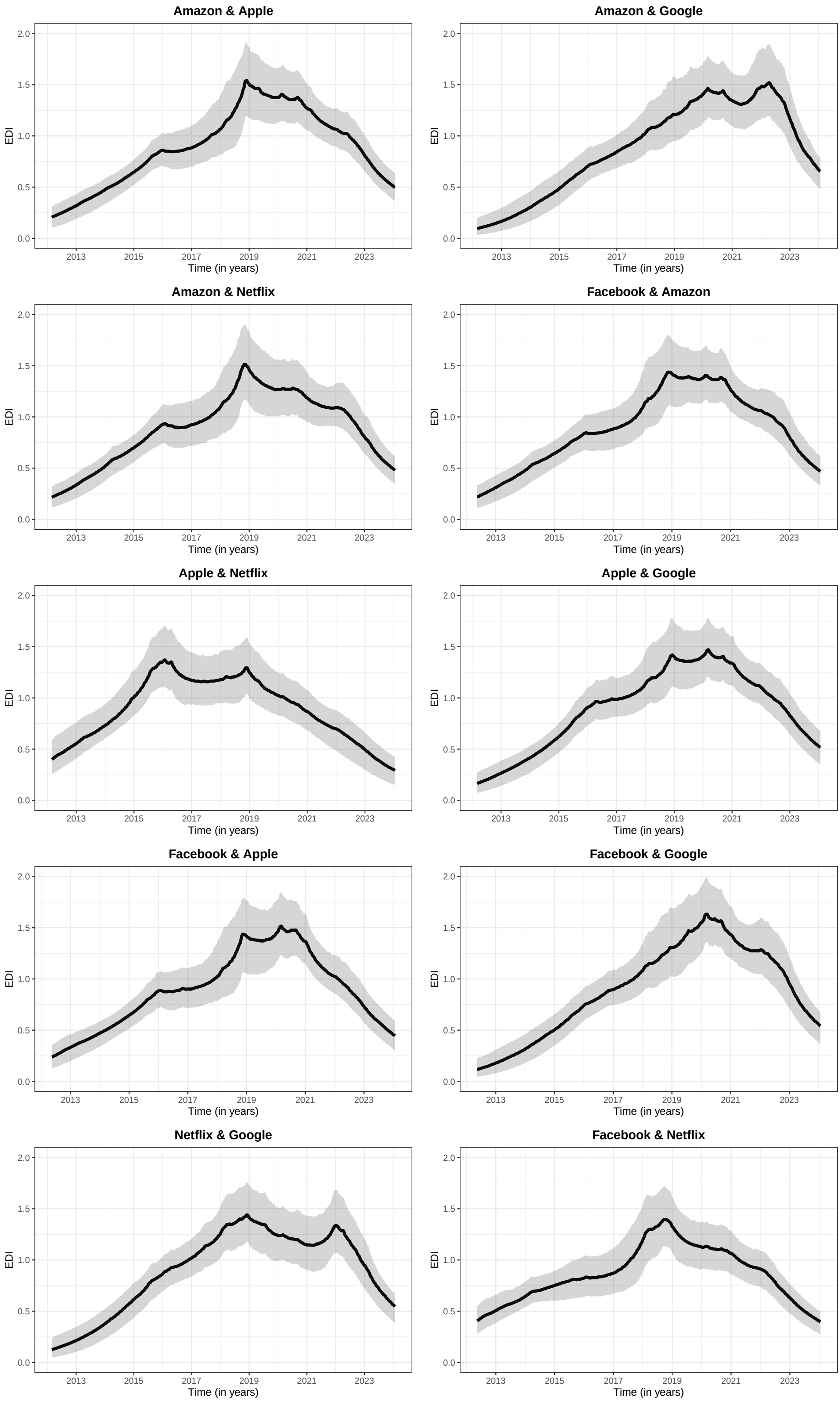

Fig. 6. Pairwise EDI for FAANG stocks: Posterior median of EDI based on a mixture of finite Polya trees along with pointwise credible bands.

Table 3. Coefficient of tail dependence for FAANG stocks: Posterior median and 95% credible intervals for pairwise analysis.

Pair of FAANG stocks	Lower limit	Posterior median	Upper limit
Facebook–Amazon	0.983	1.000	1.000
Facebook–Apple	0.999	1.000	1.000
Facebook–Netflix	1.000	1.000	1.000
Facebook–Google	1.000	1.000	1.000
Amazon–Apple	1.000	1.000	1.000
Amazon–Netflix	0.976	1.000	1.000
Amazon–Google	0.990	1.000	1.000
Apple–Netflix	0.970	1.000	1.000
Apple–Google	1.000	1.000	1.000
Netflix–Google	0.984	1.000	1.000

noted earlier a variety of joint sell-offs occurred in 2018 (Bullock et al., 2018; Bullock and Williams, 2018).

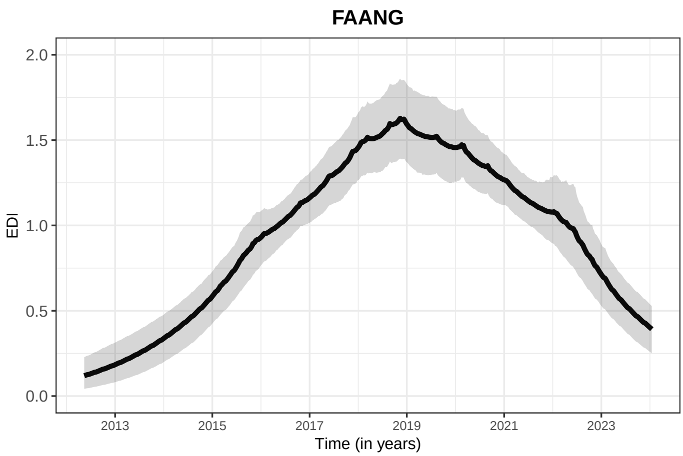

Fig. 7. Multiwise EDI for FAANG stocks: Posterior median EDI based on a mixture of finite Polya trees along with pointwise credible bands.

Also, many geopolitical issues, such as the US-China trade war (Liu and Woo, 2018; Li et al., 2018), along with US policy issues, including the impeachment of former President Trump (Jackman, 2017), may have been drivers behind some of these joint sell-offs. The posterior median coefficient of tail dependence for this multivariate analysis is 0.70 (CI = $(0.60, 0.83)$), thus suggesting that despite the sturdy growth of FAANG stocks over time, the comovements of their extreme losses is substantial; this matches the intuition from Fig. 1 where it can be seen that sharp dips for these stocks tend to be synchronized.

6. Discussion and closing remarks

This paper introduces a flexible Bayesian approach for modeling the time-changing nature of extreme observations of a random vector. The proposed framework is suitable for modeling time-varying extremal dependence, as it has been designed for tracking the dynamics governing the intensity of extreme observations in the joint tail, as well as for assessing if two stochastic processes are asymptotically dependent. In addition, we develop Bayesian estimators for the two targets of interest—the EDI and the coefficient of tail dependence. The EDI describes the intensity of the extreme observations in the joint tail over time, and it is thus a measure of the degree of association of the extremes over time. For learning about the EDI from data, we define a prior on the space of all EDIs using mixture of finite Polya trees. Our Polya tree-based approach relies on a parametric approach as a baseline model (say, a Beta centering distribution), while allowing for deviations from it whenever the data provide evidence for that. The application of our model to the so-called FAANG stocks revealed some interesting dynamics on the frequency of joint extremes over time, especially the fact the relative frequency of extreme joint losses has been higher over 2016–2019, than over the 2020 pandemic outbreak. Time-varying parameters have a long tradition in Econometrics and Statistics (e.g., Cooley and Prescott, 1976). In line with that tradition, the approach in this paper recognizes the time-changing nature of joint extreme events over time, in terms of their frequency, magnitude, and dependence.

Another natural avenue for future research entails modeling changepoints or structural changes in the intensity of extreme observations in the joint tail. While it is known that breaks in tail behavior are key in applications (e.g., Quintos et al., 2001; Lin and Kao, 2008, and references therein), most attention has been focused on modeling structural changes in the magnitude of the extreme observations rather than on their intensity function, and with the exception of de Carvalho et al. (2020) all previous developments are for the univariate setting. Modeling changepoints in the intensity of extreme observations in the joint tail would require setting a prior on the space of discontinuous EDI functions, and with the times of the breaks being themselves treated as a parameter . The assumptions in Section 2 accommodate both continuous and discontinuous EDIs; yet, instead of the mixtures of Polya trees that we have considered in this paper, perhaps Polya trees themselves may have a higher potential for learning about structural breaks in the intensity as they allow for jumps (see stages 1–3 in Fig. 3). Finally, the assumption of a constant γ was made for parsimony but could in principle be extended using the tail index regression framework of Wang and Tsai (2009), with the time index as a covariate and $Z_t = \min\{Y_{1,t}, \dots, Y_{d,t}\}$ as a response. Beyond the EDI, such extension involves specifying a second prior on a function space (as γ_t would become a function) and a rescaled Gaussian process could be a natural prior for that framework (Ghosal and Van der Vaart, 2015, Chapter 2). We leave such open problems for future analysis.

Acknowledgments

We thank the Editor, the Associate Editor, and two reviewers for insightful comments and constructive feedback. We thank Alejandro Jara, Tim Hanson, Vanda de Carvalho, and Wesley O. Johnson for insightful discussions on Polya trees.

Data Availability

Data are publicly available from Yahoo Finance.

Conflict of Interest

None declared.

Funding

The research was partially supported by The Royal Society of Edinburgh, CONACYT, and FCT Portugal through the CIDMA projects https://doi.org/10.54499/UIDB/04106/2020 and https://doi.org/10.54499/U
Funding from Abrdn, through the Centre for Investing Innovation, is also gratefully acknowledged.

Appendix A: Posterior sampling for the coefficient of tail dependence

This section derives the posterior inference algorithm for learning about the coefficient of tail dependence. The hierarchical model from Section 3.1, implies that, given γ and π, it follows that $E_1, \dots, E_k$ are independent with a distribution that depends only on γ, but not on π; that is, $\mathsf{p}(E \mid \gamma, \pi) = \gamma^{-k} \prod_{j=1}^{k} E_j^{-(1+1/\gamma)} = \mathsf{p}(E \mid \gamma)$, where $E = (E_1, \dots, E_k)^{\mathrm{T}}$. Hence, the full conditional density of $\gamma \in (0, 1]$ is

$$
\begin{aligned}
\mathsf{p}(\gamma \mid \pi, E) &\propto \mathsf{p}(E, \gamma, \pi) \\
&= \mathsf{p}(E \mid \gamma, \pi)\, \mathsf{p}(\pi, \gamma) \\
&= \mathsf{p}(E \mid \gamma)\, \mathsf{p}(\gamma \mid \pi)\, \mathsf{p}(\pi) \\
&\propto \mathsf{p}(E \mid \gamma)\, \mathsf{p}(\gamma \mid \pi) \\
&= \gamma^{-k} \prod_{j=1}^{k} E_j^{-(1+1/\gamma)} \{\pi \mathbb{1}_{\{1\}}(\gamma) + (1-\pi)\beta(\gamma; a_\gamma, b_\gamma)\}.
\end{aligned}
\tag{15}
$$

To update γ we use a Metropolis–Hastings step based on (15) where the proposal distribution is a mixture, $\mathsf{q}(\gamma^* \mid \gamma) = \omega \mathbb{1}_{\{1\}}(\gamma^*) + (1-\omega)\,\mathrm{TN}(\gamma^* \mid \gamma, 1)$; here, $\mathrm{TN}(\,\cdot \mid \gamma, 1)$ is the density of the truncated Normal distribution on $(0, 1)$, with mean γ and variance 1.

It follows from (11) that $\gamma \mid \pi$ is a mixture of a continuous and a discrete part ($\{1\}$), hence implying that

$$
\gamma = \begin{cases} 1, & \text{w.p. } \pi, \\ \gamma', & \text{w.p. } 1-\pi, \end{cases}
$$

where $\gamma' \sim \mathrm{Beta}(a_\gamma, b_\gamma)$. Let $\delta = \mathbb{1}_{\{1\}}(\gamma)$ be a binary latent indicator so that $\delta \sim \mathrm{Bern}(\pi)$. Beta–Bernoulli conjugacy implies that

$$
\mathsf{p}(\pi \mid \delta) \propto \mathsf{p}(\delta \mid \pi)\, \mathsf{p}(\pi) = \beta(\pi, a_\pi + \mathbb{1}_{\{1\}}(\gamma), b_\pi + \mathbb{1}_{(0,1)}(\gamma)),
$$

which motivates approximating the full conditional of $(\pi \mid \delta^{(1)}, \dots, \delta^{(i)})$ via a $\mathrm{Beta}(a_\pi + |r_i|, b_\pi + i - |r_i|)$ distribution, where $r_i = \{1 \leq j \leq i : \gamma^{(j)} = 1\}$. The quality of the latter approximation can be improved via thinning so to attenuate the dependence of the chain of δ's; in all results in the paper we used a thinning of 5. We found this approximation to a Gibbs step to work well in the Monte Carlo simulations of Section 3 and in the Supporting Information (Section 1.4).

Algorithm 1 summarizes the computational procedure based on the above derivations.

Algorithm 1 MONTE CARLO POSTERIOR SAMPLING FOR COEFFICIENT OF TAIL DEPENDENCE

1. Initialize $(\gamma^{(1)}, \pi^{(1)})$.
2. Sample $\gamma^* \sim \pi^{(i)} \mathbb{1}_{\{1\}} + (1 - \pi^{(i)})\mathrm{TN}(\gamma^{(i)}, 1)$ and compute the ratio

$$
R = \frac{\mathsf{p}(\gamma^* \mid \pi, E)\, \mathsf{q}(\gamma \mid \gamma^*)}{\mathsf{p}(\gamma \mid \pi, E)\, \mathsf{q}(\gamma^* \mid \gamma)}.
$$

   Accept $\gamma^{(i+1)} \equiv \gamma^*$ with probability $\min\{R, 1\}$; else, set $\gamma^{(i+1)} \equiv \gamma^{(i)}$.
3. Sample $\pi^{(i+1)}$ from $\mathrm{Beta}(a_\pi + |r_i|, b_\pi + i - |r_i|)$, where $r_i = \{j \leq i + 1 : \gamma^{(j)} = 1\}$.
4. Repeat Steps 2 and 3 until reaching stationarity.

References

Albuquerque, R. and Vega, C. (2009), Economic news and international stock market co-movement, *Review of Finance*, 13, 401–465.

Avery, C. and Zemsky, P. (1998), Multidimensional uncertainty and herd behavior in financial markets, *American economic review*, 724–748.

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004), *Statistics of extremes: Theory and applications*, Wiley, Hoboken, NJ.

Berman, S. M. (1961), Convergence to bivariate limiting extreme value distributions, *Annals of the Institute of Statistical Mathematics*, 13, 211–223.

Brooks, R. and Del Negro, M. (2004), The rise in comovement across national stock markets: Market integration or IT bubble? *Journal of Empirical Finance*, 11, 659–680.

Bullock, N., Rennison, J., and Authers, J. (2018), Shares in FAANGs suffer their worst one-day loss, *Financial Times.*

Bullock, N. and Williams, A. (2018), While confidence in tech trade has fallen, few are prepared to call the top for the sector, *Financial Times.*

Castro, D., de Carvalho, M., and Wadsworth, J. L. (2018), Time-varying extreme value dependence with application to leading European stock markets, *Annals of Applied Statistics*, 12, 283–309.

Chiang, T. C. and Zheng, D. (2010), An empirical analysis of herd behavior in global stock markets, *Journal of Banking & Finance*, 34, 1911–1921.

Christensen, R., Hanson, T., and Jara, A. (2008), Parametric nonparametric statistics: An introduction to mixtures of finite Polya trees, *The American Statistician*, 62, 296–306.

Cipolli, W. and Hanson, T. (2017), Computationally tractable approximate and smoothed Polya trees, *Statistics and Computing*, 27, 39–51.

Cipriani, M. and Guarino, A. (2014), Estimating a structural model of herd behavior in financial markets, *American Economic Review*, 104, 224–251.

Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values*, London: Springer.

Cooley, T. F. and Prescott, E. C. (1976), Estimation in the presence of stochastic parameter variation, *Econometrica*, 167–184.

de Carvalho, M. (2016), Statistics of extremes: Challenges and opportunities, in *Extreme Events in Finance: A Handbook of Extreme Value Theory and its Applications*, ed. Longin, F., Hoboken: Wiley.

de Carvalho, M., Leonelli, M., and Rossi, A. (2020), Tracking change-points in multivariate extremes, *arXiv:2011.05067.*

Ehrmann, M. and Jansen, D.-J. (2020), Stock Return comovement when investors are distracted: More, and more homogeneous, *European Central Bank Working Paper*, No 2412.

Embrechts, P., Hofert, M., and Wang, R. (2016), Bernoulli and tail-dependence compatibility, *The Annals of Applied Probability*, 26, 1636–1658.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997), *Modelling Extremal Events for Insurance and Finance*, New York: Springer.

Financial Times—Editorial Board (2023), How to prevent AI from provoking the next financial crisis, .

Forbes, K. J. and Rigobon, R. (2002), No contagion, only interdependence: Measuring stock market comovements, *The Journal of Finance*, 57, 2223–2261.

Fuentes, M., Henry, J., and Reich, B. (2013), Nonparametric spatial models for extremes: Application to extreme temperature data, *Extremes*, 16, 75–101.

Ghosal, S. and Van der Vaart, A. W. (2015), *Fundamentals of Nonparametric Bayesian Inference*, CUP, Cambridge.

Gong, Y. and Huser, R. (2022), Asymmetric tail dependence modeling, with application to cryptocurrency market data, *The Annals of Applied Statistics*, 16, 1822–1847.

Hanson, T. and Johnson, W. O. (2002), Modeling regression error with a mixture of Polya trees, *Journal of the American Statistical Association*, 97, 1020–1033.

Hanson, T. E. (2006), Inference for mixtures of finite Polya tree models, *Journal of the American Statistical Association*, 101, 1548–1565.

Hanson, T. E., de Carvalho, M., and Chen, Y. (2017), Bernstein polynomial angular densities of multivariate extreme value distributions, *Statistics and Probability Letters*, 128, 60–66.

Harvey, A. and Oryshchenko, V. (2012), Kernel density estimation for time series data, *International Journal of Forecasting*, 28, 3–14.

Heffernan, J. E. (2000), A directory of coefficients of tail dependence, *Extremes*, 3, 279–290.

Jach, A. (2017), International stock market comovement in time and scale outlined with a thick pen, *Journal of Empirical Finance*, 43, 115–129.

Jackman, S. (2017), Impeachment 101: The history, process and prospect of a Trump impeachment, *The United States Studies Centre at the University of Sydney.*

Jondeau, E., Poon, S.-H., and Rockinger, M. (2007), *Financial modeling under non-Gaussian distributions*, Springer Science & Business Media.

Kahneman, D. (2011), *Thinking Fast and Slow*, New York: Macmillan.

Kottas, A., Wang, Z., and Rodríguez, A. (2012), Spatial modeling for risk assessment of extreme values from environmental time series: A Bayesian nonparametric approach, *Environmetrics*, 23, 649–662.

Kraft, C. H. (1964), A class of distribution function processes which have derivatives, *J. Appl. Prob.*, 1, 385–388.

Lavine, M. et al. (1992), Some aspects of Polya tree distributions for statistical modelling, *The Annals of Statistics*, 20, 1222–1235.

— (1994), More aspects of Polya tree distributions for statistical modelling, *The Annals of Statistics*, 22, 1161–1176.

Ledford, A. W. and Tawn, J. A. (1996), Statistics for near independence in multivariate extreme values, *Biometrika*, 83, 169–187.

Lee, J., de Carvalho, M., Rua, A., and Avila, J. (2024), Bayesian smoothing for time-varying extremal dependence, *Journal of the Royal Statistical Society Series C: Applied Statistics*, 73, 581–597.

Li, C., He, C., and Lin, C. (2018), Economic impacts of the possible China–US trade war, *Emerging Markets Finance and Trade*, 54, 1557–1577.

Lin, C.-H. and Kao, T.-C. (2008), Multiple structural changes in the tail behavior: evidence from stock index futures returns, *Nonlinear Anal.: Real World Appl.*, 9, 1702–1713.

Liu, T. and Woo, W. T. (2018), Understanding the US-China trade war, *China Econ. J.*, 11, 319–340.

Mandelbrot, B. B. and Mandelbrot, B. B. (1997), *The variation of certain speculative prices*, Springer.

Marcon, G., Padoan, S. A., and Antoniano-Villalobos, I. (2016), Bayesian inference for the extremal dependence, *Electronic Journal of Statistics*, 10, 3310–3337.

Mhalla, L., de Carvalho, M., and Chavez-Demoulin, V. (2019), Regression Type Models for Extremal Dependence, *Scandinavian J. Statist.*, 49, 1141–1167.

Montier, J. (2010), *The Little Book of Behavioral Investing: How Not to be Your Own Worst Enemy*, vol. 35, New York: Wiley.

Morana, C. and Beltratti, A. (2008), Comovements in international stock markets, *Journal of International Financial Markets, Institutions and Money*, 18, 31–45.

Müller, P., Quintana, F. A., Jara, A., and Hanson, T. (2015), *Bayesian Nonparametric Data Analysis*, New York: Springer.

Nieto-Barajas, L. E., Müller, P., Ji, Y., Lu, Y., and Mills, G. B. (2012), A time-Series DDP for functional proteomics profiles, *Biometrics*, 68, 859–868.

Padoan, S. A. and Rizzelli, S. (2022), Consistency of Bayesian inference for multivariate max-stable distributions, *The Annals of Statistics*, 50, 1490–1518.

Patton, A. J. (2006), Modelling Asymmetric Exchange Rate Dependence, *International Economic Review*, 47, 527–556.

Poon, S.-H., Rockinger, M., and Tawn, J. (2003), Modelling extreme-value dependence in international stock markets, *Statistica Sinica*, 13, 929–953.

Quintos, C., Fan, Z., and Phillips, P. C. (2001), Structural change tests in tail behaviour and the Asian crisis, *Review of Economic Studies*, 68, 633–663.

Rodriguez, A. and Müller, P. (2013), Nonparametric Bayesian Inference, *NSF-CBMS Regional Conference Series in Probability and Statistics*, 9, i–110.

Rua, A. and Nunes, L. C. (2009), International comovement of stock market returns: A wavelet analysis, *Journal of Empirical Finance*, 16, 632–639.

Tsay, R. S. (2002), *Analysis of Financial Time Series*, New York: Wiley.

Wang, H. and Tsai, C.-L. (2009), Tail index regression, *Journal of the American Statistical Association*, 104, 1233–1240.

Wigglesworth, R. (2020), How Big Tech got even bigger in the COVID-19 era, *Financial Times*.