

DYNAMIC
DEPENDENCE STRUCTURES
FOR MULTIVARIATE HIGH-FREQUENCY
DATA IN FINANCE

Paul Embrechts

Alexandra Dias

Department of Mathematics, ETH-Zürich
www.math.ethz.ch/~embrechts

September 2004

ACKNOWLEDGEMENT

We thank Olsen Data for having provided us with the high-frequency data used in this study

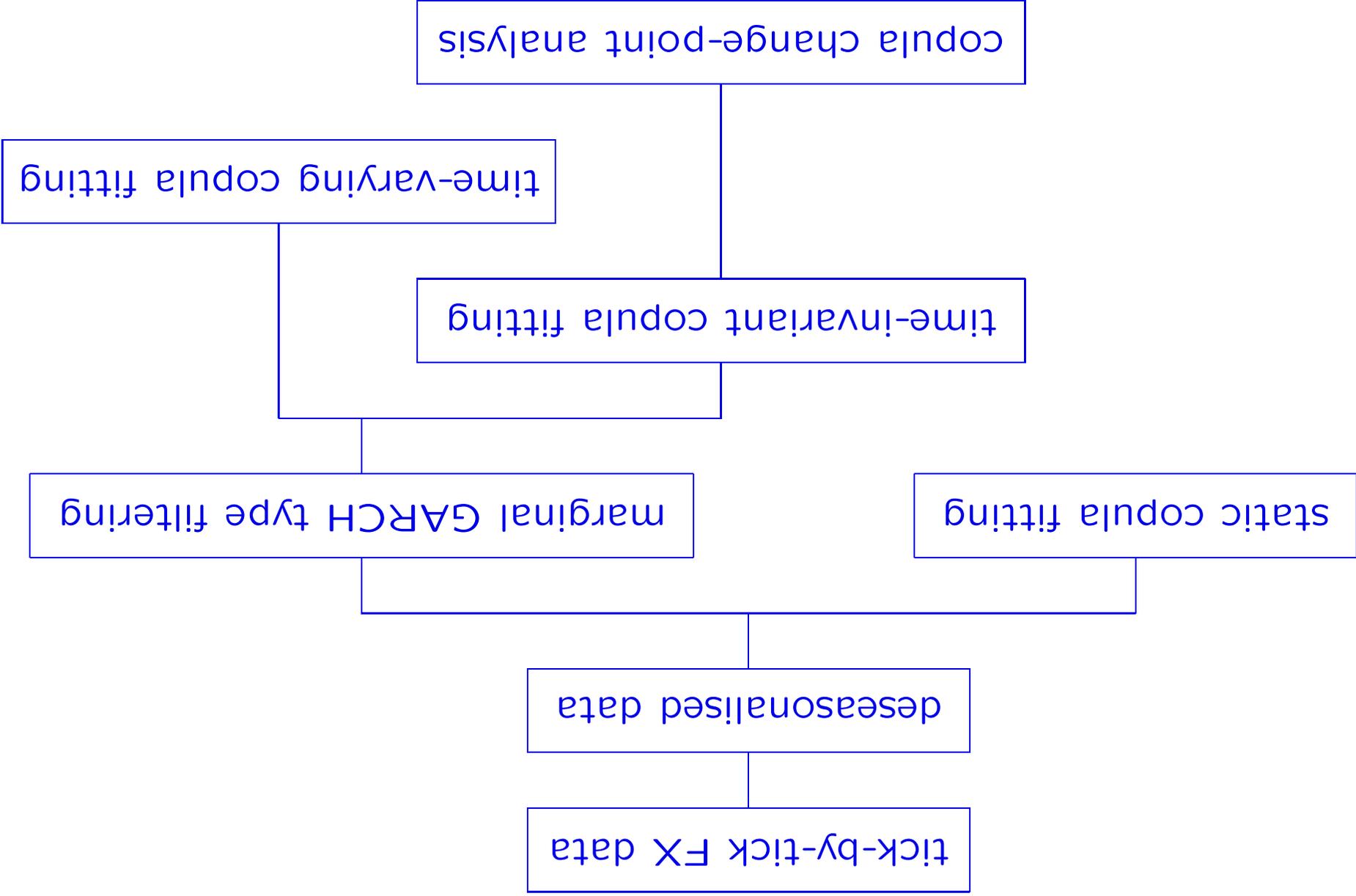
REFERENCES

- [1] Breymann, W., Dias, A., and Embrechts, P. (2003). Dependence structures for multivariate high-frequency data in finance. *Quant. Finance*, 3:1–14.
- [2] Dias, A., and Embrechts, P. (2003). Dynamic copula models for multivariate high-frequency data in finance. Research paper, ETH-Zürich.

OVERVIEW

- Motivation
- The FX data
- Deseasonalisation
- Static copula fitting
- Dynamic dependence structure modelling
 - Time-invariant copula model
 - Tail dependence analysis
 - Time-varying copula model
 - Copula change-point detection
- Conclusion

OVERVIEW OF THE FX COPULA ANALYSIS



MOTIVATION

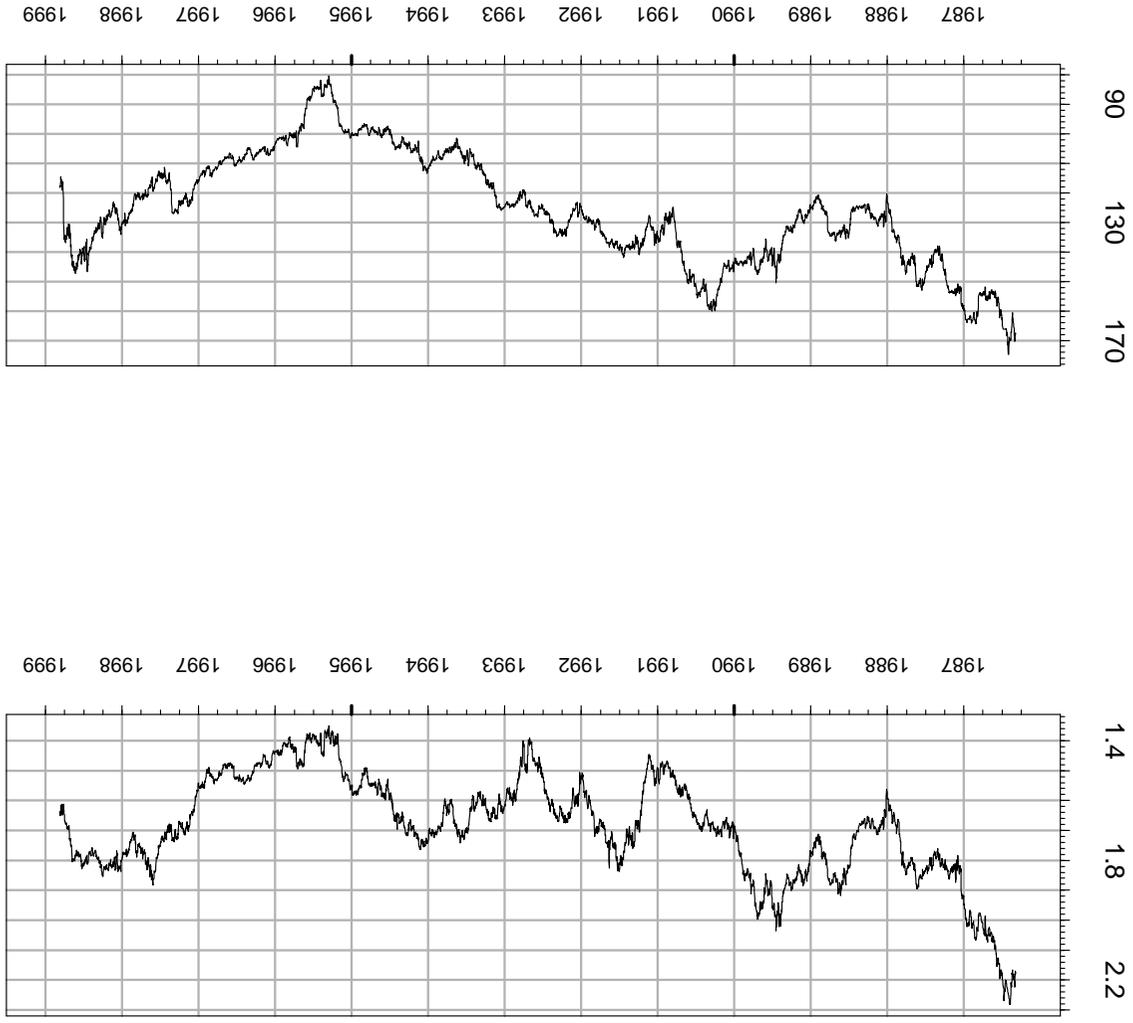
- The Goal:
 - Studying the **dynamic dependence structure across time scales**
 - Why?
 - The change of the behavior as a function of the time horizon may contain important information
 - Improves extrapolation from small to large time horizons
- Requires:
 - Characterising dependence for horizons from minutes to months
 - Here:
 - Restriction to high-frequency region (**1 hour – 1 day**)
- Peculiarities of high-frequency data are taken into account

THE DATA

- Tick-by-tick bid and ask quotes
- Period: Febr 1986–Dec 1998
- Collected and filtered by Olsen Data
- Irregularly spaced
- About 10 million data points for a single series
- Regularisation to 5 min. time series by linear interpolation
- Reduction to logarithmic middle prices:

$$\xi_{\alpha,t} = \frac{\log \left(p_{\alpha,t}^{Bid} \cdot p_{\alpha,t}^{Ask} \right)}{2}$$

FX PRICES FOR USD/DEM AND USD/JPY



DESEASONALISATION OF FINANCIAL DATA

- High frequency financial data present strong **seasonalities**
- Main periodicities: **daily** and **weekly**

• Seasonalities cover more subtle statistical properties

• Affected by Daylight Saving Time (DST)

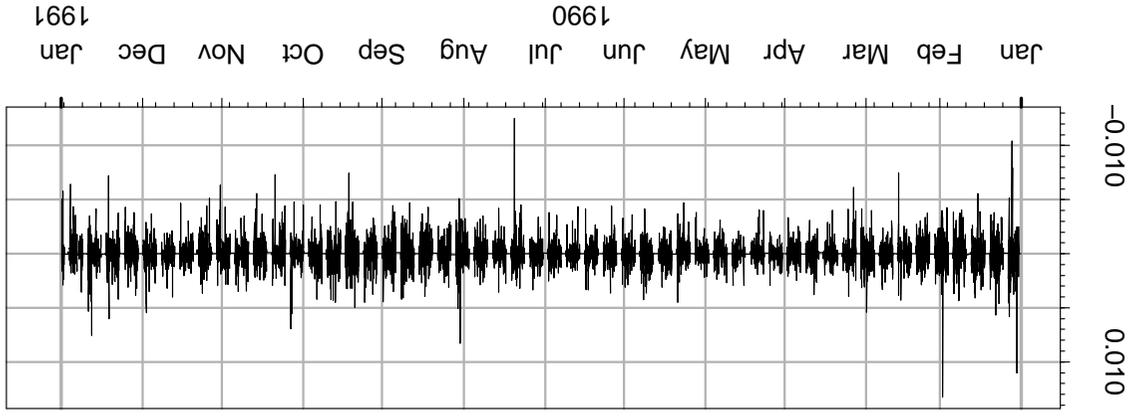
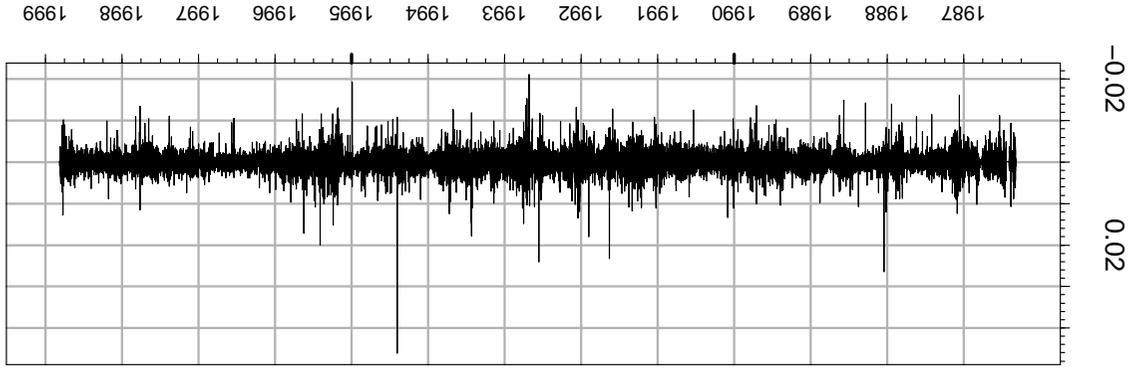
• Theory of stochastic processes favors time transformation to an activity-based time scale, but:

→ Loss of synchronicity in the multivariate case

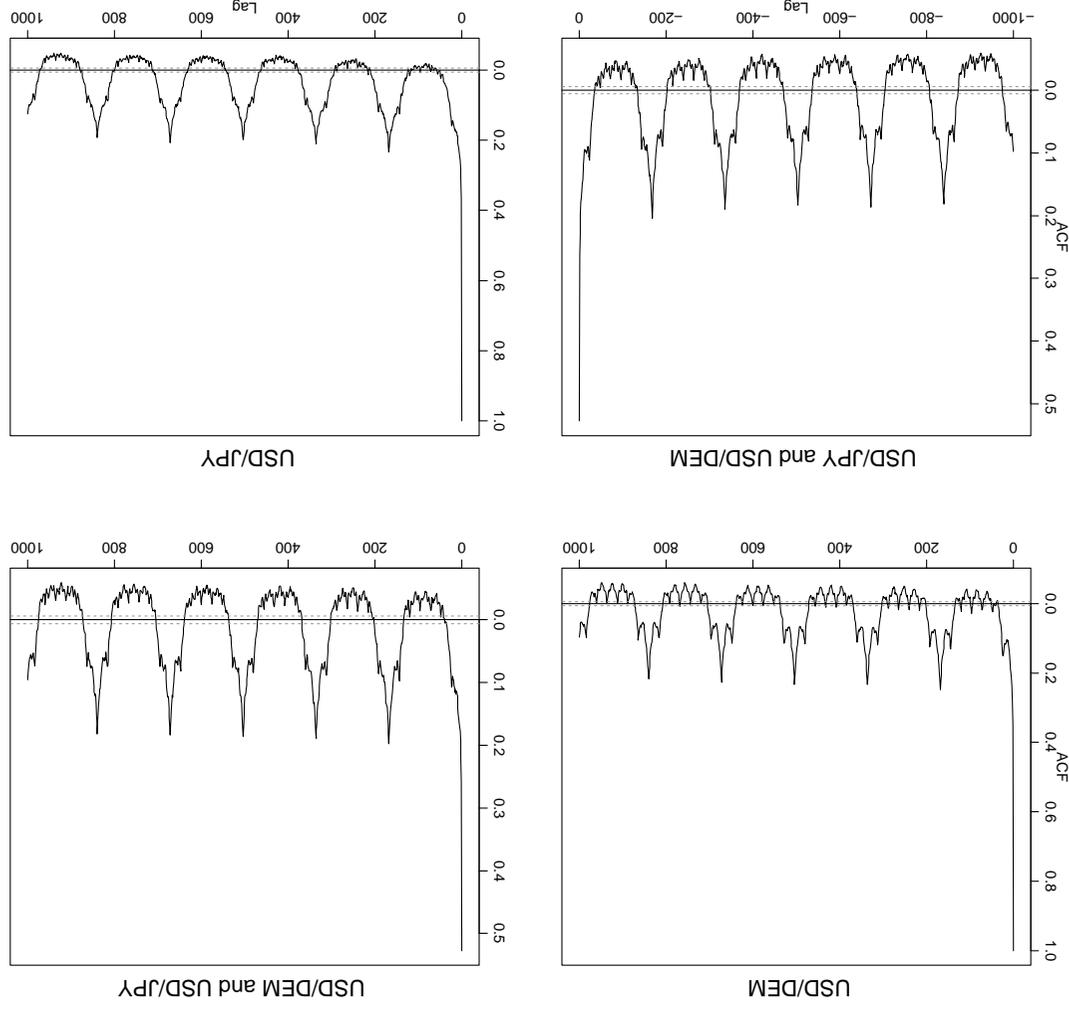
• Instead:

→ **Volatility weighting** based on weekly activity pattern

HOURLY RETURNS OF USD/DEM



AUTOCORRELATION FUNCTIONS OF ABSOLUTE RETURNS



MODELLING REQUIREMENTS

- The flexibility of modelling arbitrary patterns that display abrupt volatility changes (\rightarrow Japanese lunch break)
- Taking into account slow temporal changes in the habits of the market participants, institutional changes, etc
- Keeping track of Daylight Saving Time (DST) to take into account the one hour displacement between DST and non-DST periods
- The modelling of the geographical decomposition of market activity to take into account local public holidays and other irregularities

THE VOLATILITY PATTERN

- Integrated squared volatility: $V_t^2 = \sum_{t' \leq t} (v_{t'}[\delta])^2$

- Volatility wrt horizon ΔT :

$$\Delta V_2^t[\Delta T] \equiv V_2^t - V_2^{t-\Delta T} = \sum_{i=0}^{n-1} (v_{t-i\delta}[\delta])^2 = (v_t[\Delta T])^2$$

- Deseasonalised returns:

$$x_t[\Delta T] = \frac{\sqrt{\Delta V_2^t[\Delta T]}}{\xi_t - \xi_{t-\Delta T}}$$

- $\delta = 5$ minutes: elementary time step; $n = \Delta T/\delta$

- Aggregation property:

$$x_t[\Delta T] = \frac{\sqrt{\Delta V_2^t[\Delta T]}}{\sqrt{\Delta V_2^{t-\Delta T_2}[\Delta T_1] + x_{t-\Delta T_2}[\Delta T_1] \Delta V_2^{t-\Delta T_2}[\Delta T_2]}} = \sqrt{\Delta V_2^t[\Delta T]}$$

COMPUTING THE VOLATILITY PATTERN

- Decomposition of the volatility:

$$\sigma_{t}^2 = a_t \left(v_{t}^{(d)} \right)^2$$

– with relative market activity factor a_t and

– volatility averaged over DST period d conditional to the time in the week, $t = t \bmod (1 \text{ week})$:

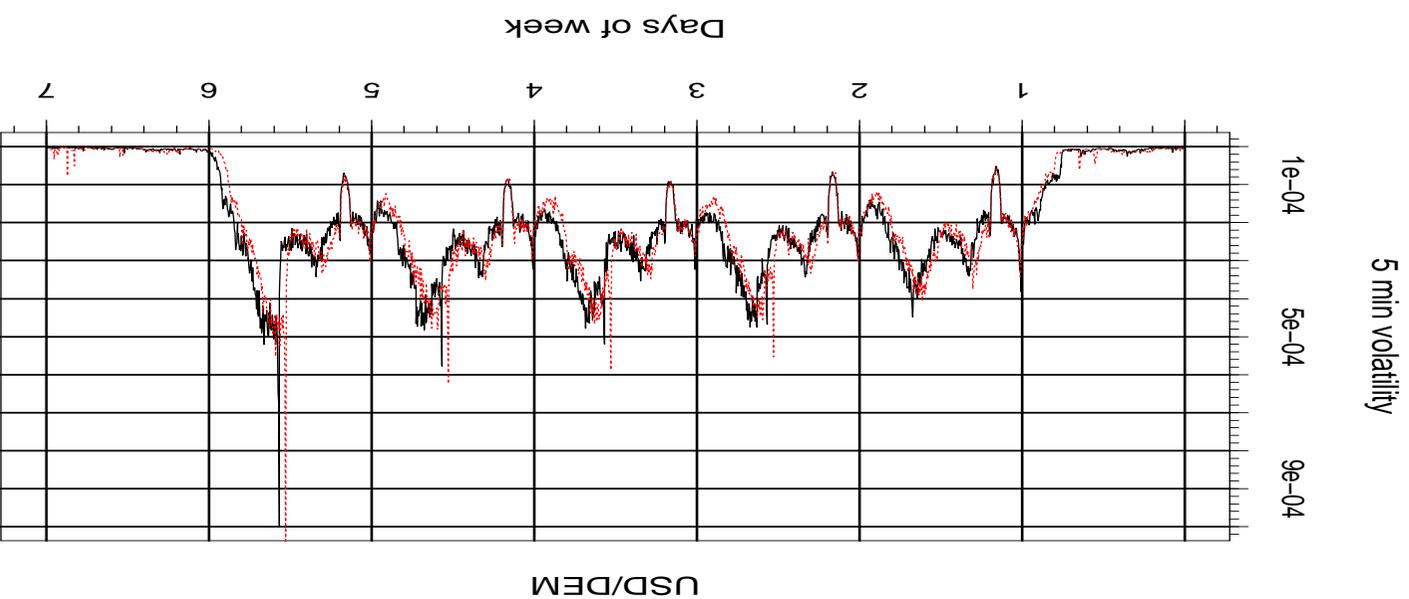
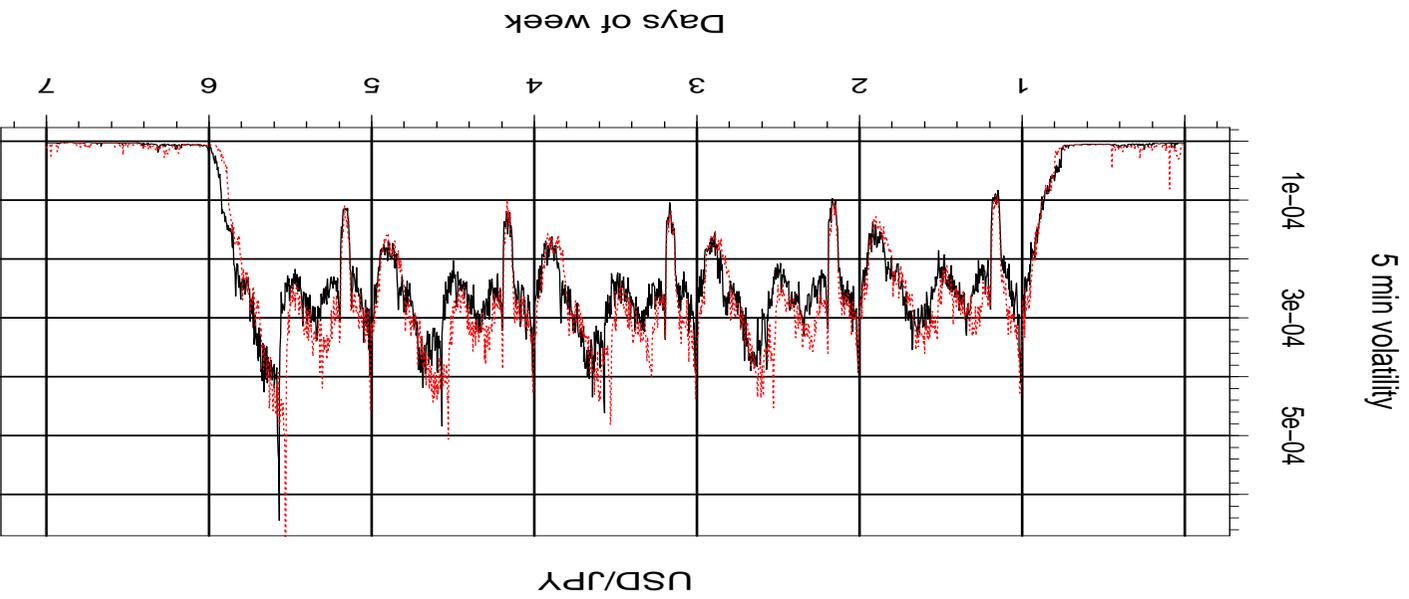
$$\frac{1}{N_d} \sum_{i=1}^N v_{t_i+t}^2 = \left(v_{t}^{(d)} \right)^2$$

- Weekend volatility:

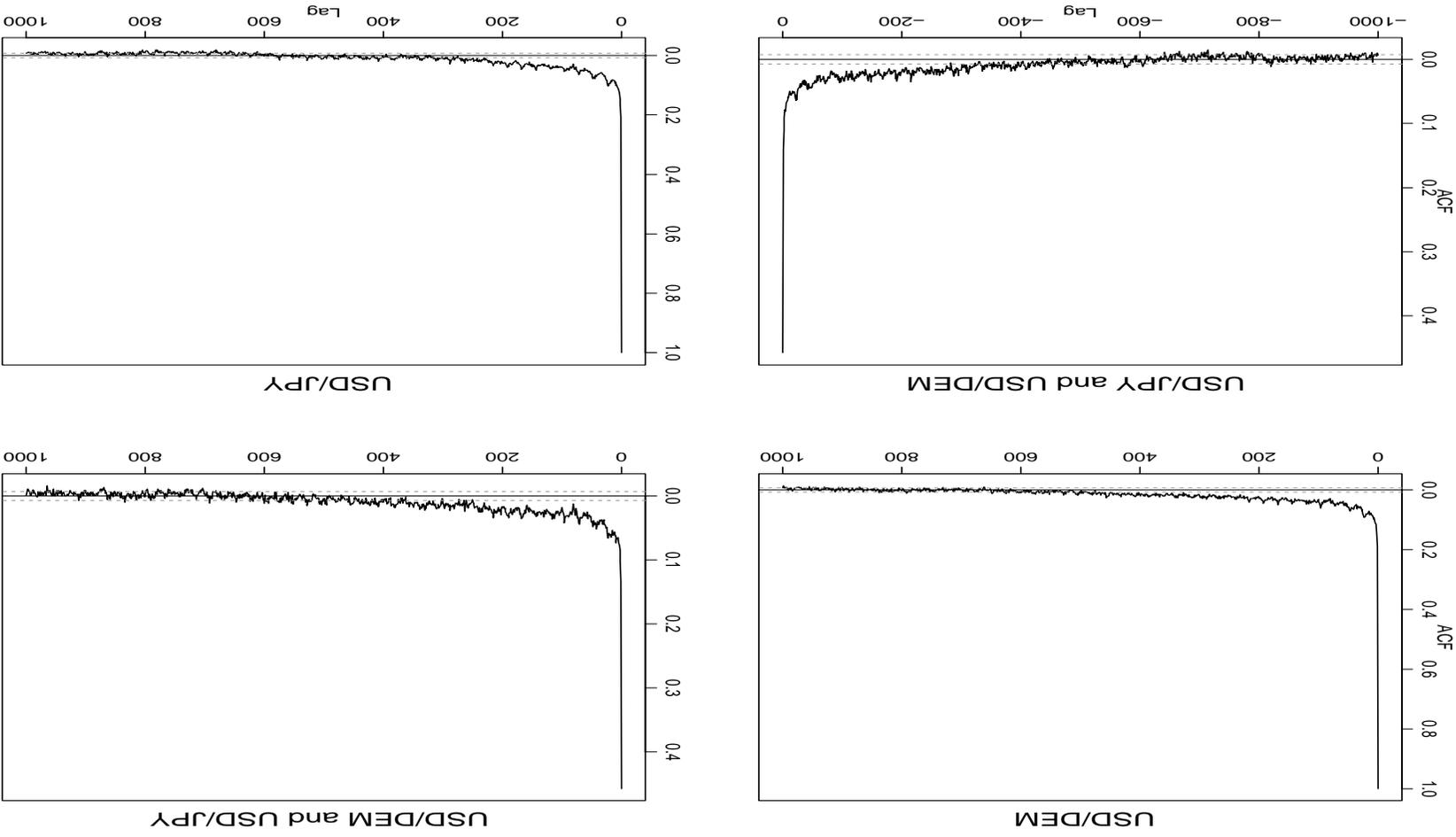
$$\sigma_{t}^2 = \left(v_{t}^{(m)} \right)^2 \left(\frac{\Delta L^m}{\delta} \right)^2$$

– with weekend length $\Delta L^m = t_{(end)}^m - t_{(start)}^m$

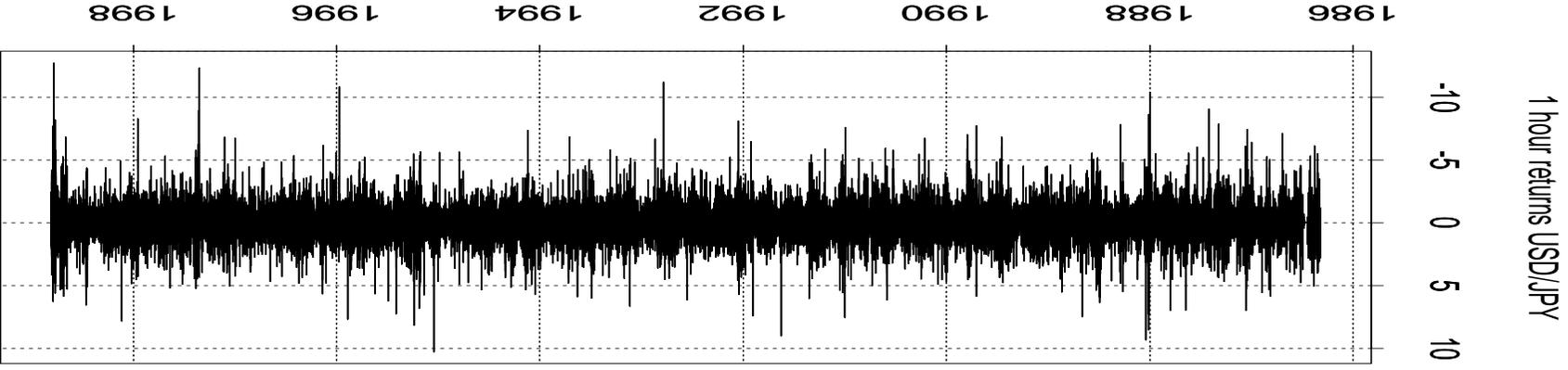
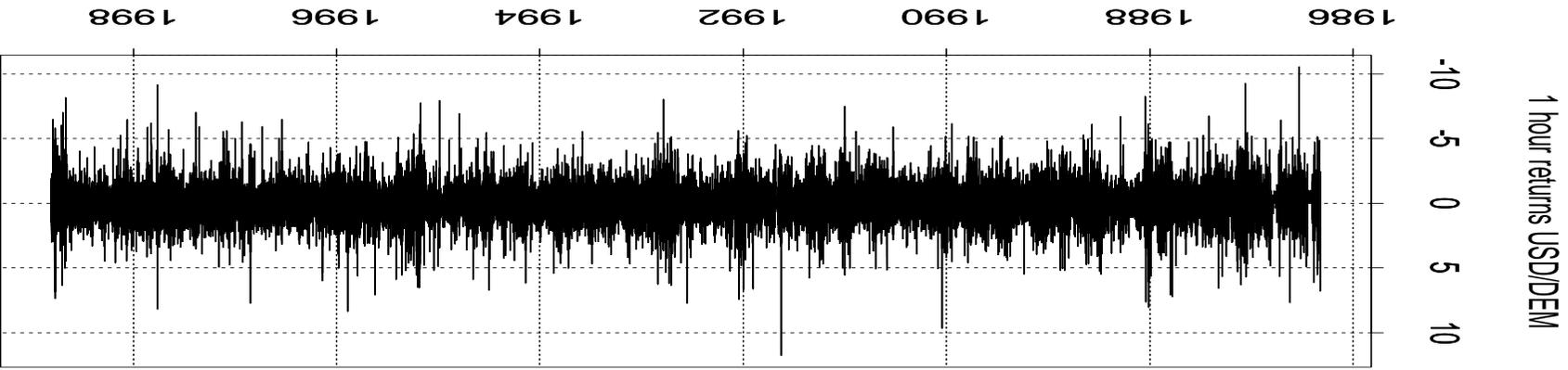
THE WEEKLY VOLATILITY PATTERN



ACF OF DESEASONALISED HOURLY ABSOLUTE RETURNS



DESEASONALISED RETURNS



DEPENDENCE STRUCTURE MODELLING

- Static copula fitting
- Marginal GARCH type filtering
- Time-invariant copulas across time scales
 - Copula fitting
 - Goodness of fit and ellipticity test
 - Tail coefficient
 - Spectral measure and bivariate excesses
- Multivariate (matrix-diagonal) GARCH
- Time-varying copulas across time scales
- Change-point detection

DEPENDENCE STRUCTURE

$\mathbf{X} = (X_1, \dots, X_d)$ d -dimensional vector of risk factors

- **Marginal** distributions (assumed continuous)

$$F_1(x_1) = P(X_1 \leq x_1), \dots, F_d(x_d) = P(X_d \leq x_d)$$

- **Joint** distribution

$$\begin{aligned} F_{\mathbf{X}}(\mathbf{x}) &= P(X_1 \leq x_1, \dots, X_d \leq x_d) \\ &= P(F_1(X_1) \leq F_1(x_1), \dots, F_d(X_d) \leq F_d(x_d)) \\ &= P(U_1 \leq F_1(x_1), \dots, U_d \leq F_d(x_d)) \end{aligned}$$

$(U_1, \dots, U_d) \stackrel{d}{\sim} C$ on $[0, 1]^d$: the **copula**

- **Copula** representation (Sklar):

$$F_{\mathbf{X}}(\mathbf{x}) = C(F_1(x_1), \dots, F_d(x_d))$$

$$C(u_1, u_2, \dots, u_d) = F_{\mathbf{X}}(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$$

FAMILIES OF COPULAS

- **Gaussian** copula for correlation ρ :

$$C_{Ga}^{\rho}(u, v) = \int_{\Phi^{-1}(u)}^{\infty} \int_{\Phi^{-1}(v)}^{\infty} \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{(s^2 - 2\rho st + t^2)}{2(1-\rho^2)}\right\} ds dt$$

- **t**-copula for ν degrees of freedom and correlation ρ :

$$C_{t,\nu}^{\rho}(u, v) = \int_{t_{-1}^{\nu}(u)}^{\infty} \int_{t_{-1}^{\nu}(v)}^{\infty} \frac{2^{\nu} \sqrt{1-\rho^2}}{(s^2 - 2\rho st + t^2)^{\nu+1/2}} ds dt$$

- **Clayton** copula:

$$C_{Cl}^{\beta}(u, v) = \max\left[-\left\{(-\log u)^{1/\beta} + (-\log v)^{1/\beta}\right\}^{\beta}, 0\right]$$

FAMILIES OF COPULAS

- Gumbel copula:

$$C_{G_n}^{\beta}(u, v) = \exp \left[- \left\{ (-\log u)^{1/\beta} + (-\log v)^{1/\beta} \right\}^{\beta} \right]$$

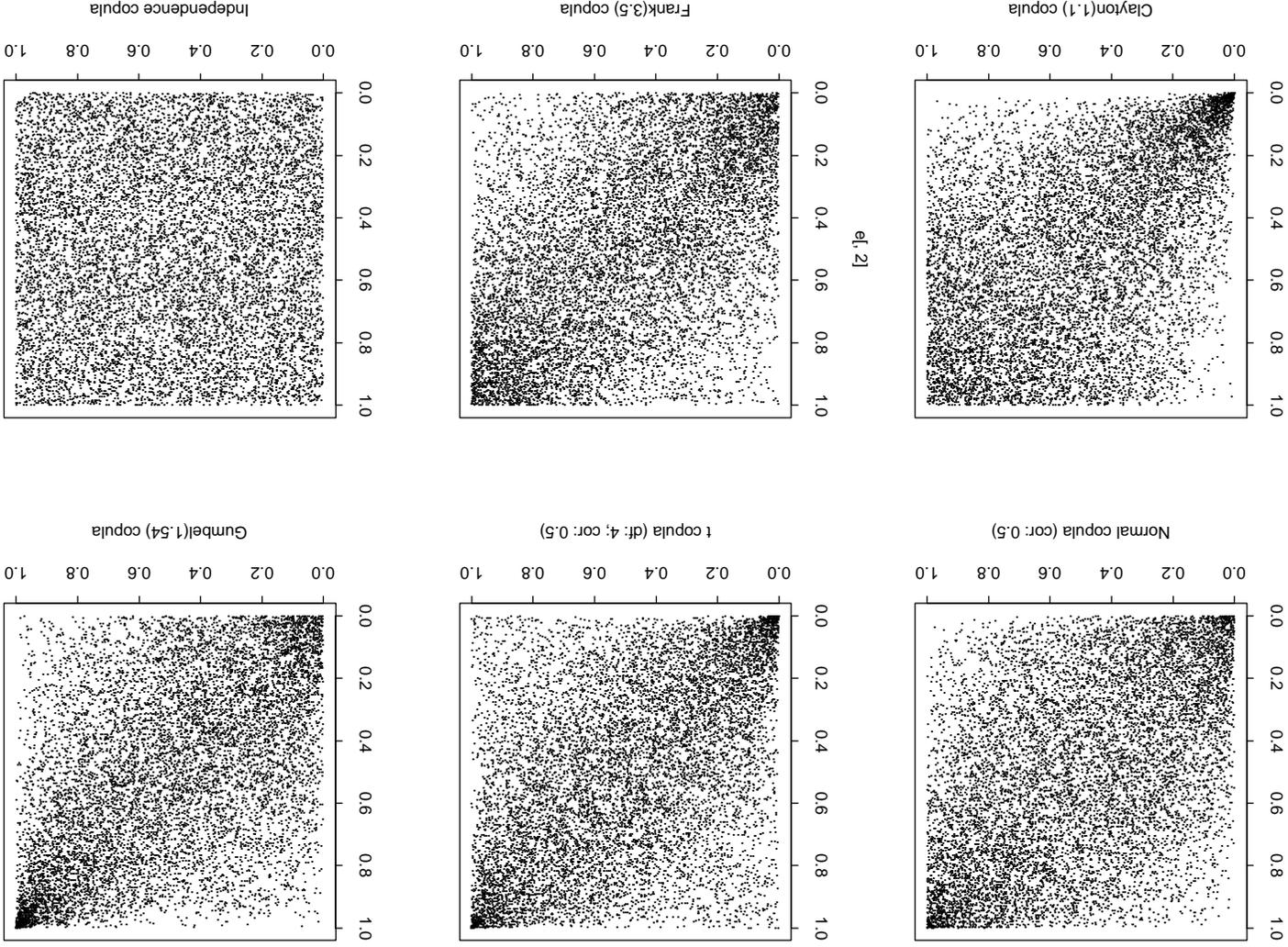
- Mixture of Gumbel and survival Gumbel copulas:

$$C(u_1, u_2; \theta) = \theta_3 C_{G_n}^{u_1, u_2; \theta_1} + (1 - \theta_3) \left(2u_1 - 1 + (1 - u_1)^{2^{1/\theta_2}} \right)$$

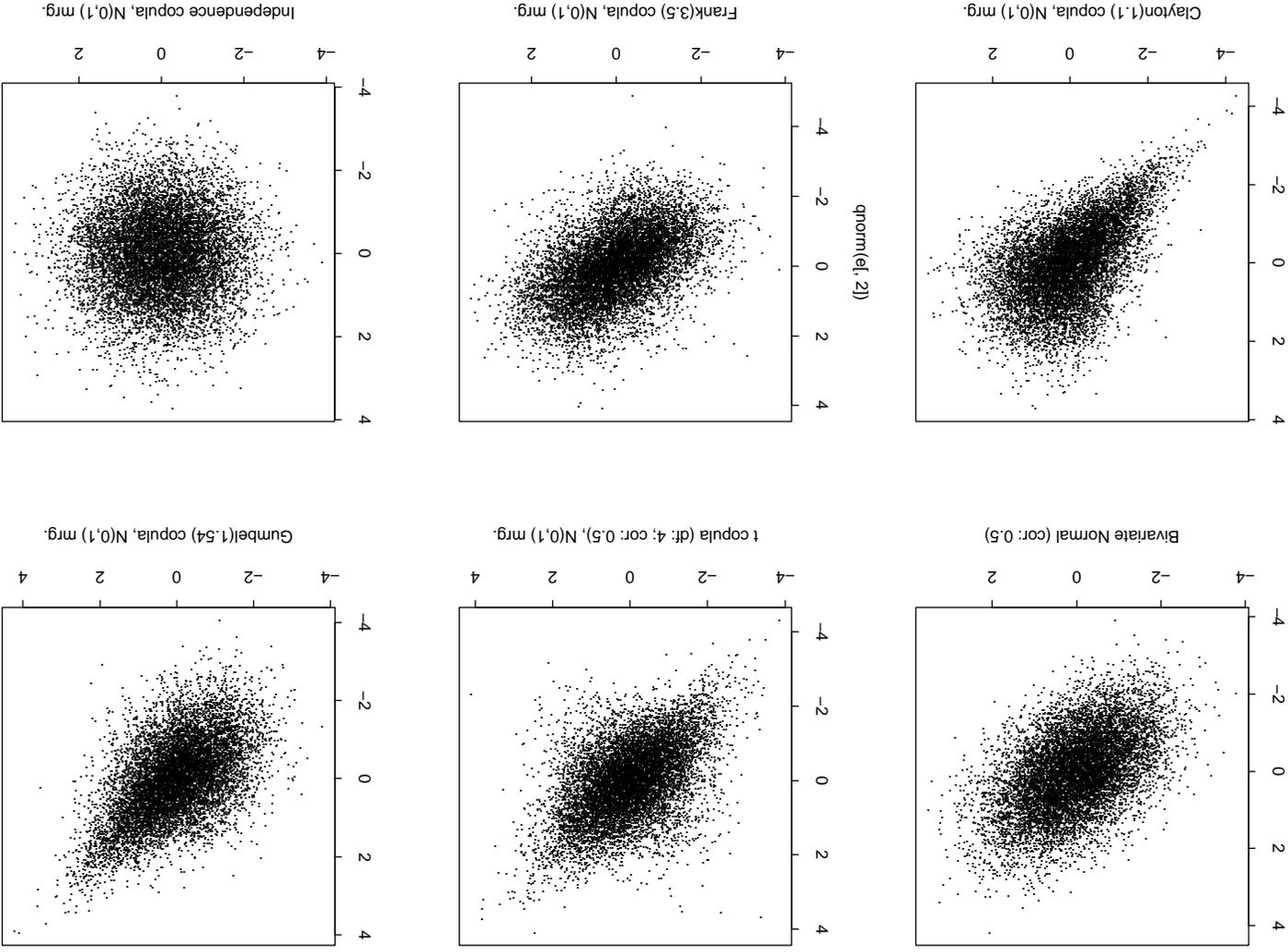
- Frank copula:

$$C_{Fr}^{\beta}(u, v) = -\frac{1}{\beta} \log \left[1 + \frac{(e^{-\beta u} - 1)(e^{-\beta v} - 1)}{e^{-\beta} - 1} \right]$$

COPULA DENSITIES FOR SELECTED COPULAS

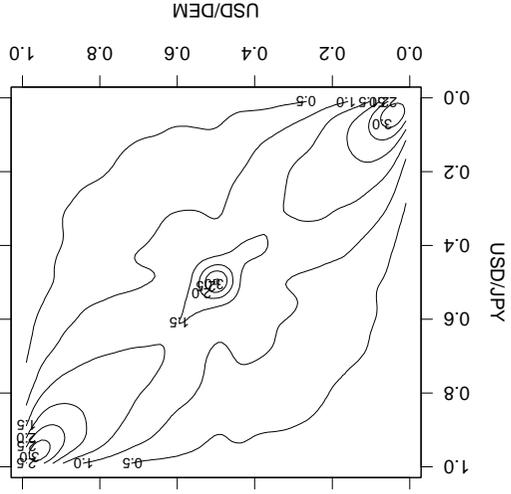


COPULA DENSITIES WITH NORMAL MARGINS

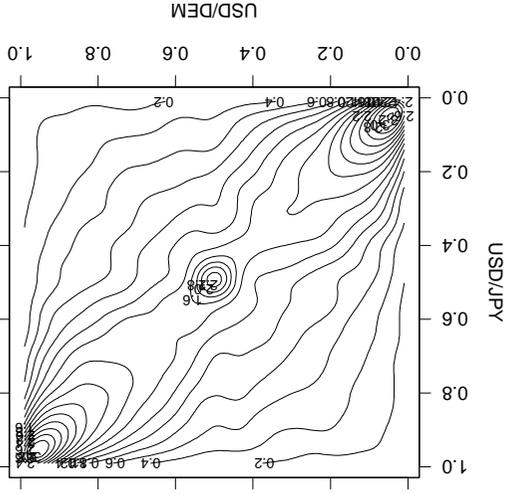


CONTOUR-PLOTS OF DESEASONALISED USD/DEM AND USD/JPY RETURNS

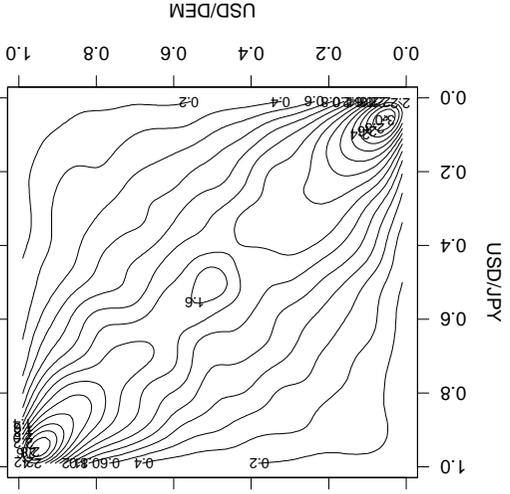
1 Hour returns



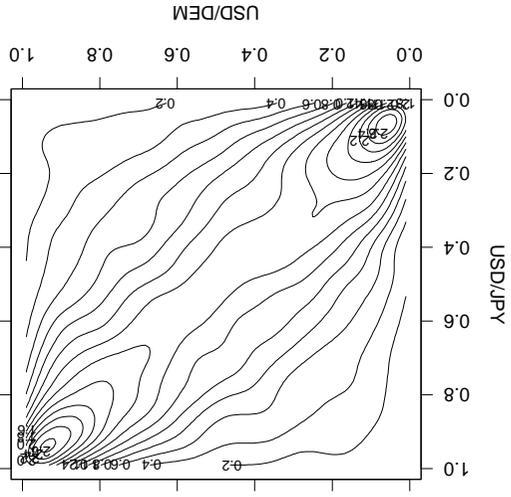
2 Hours returns



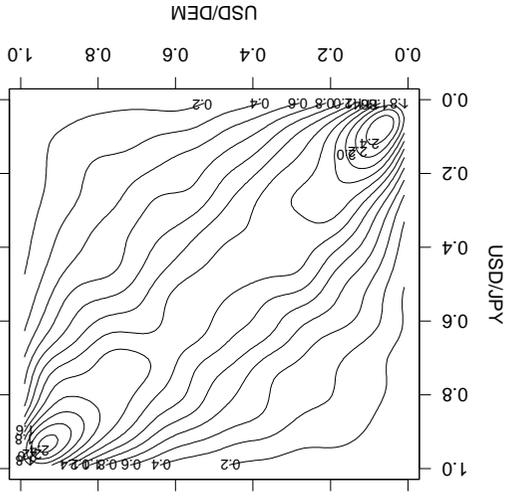
4 Hours returns



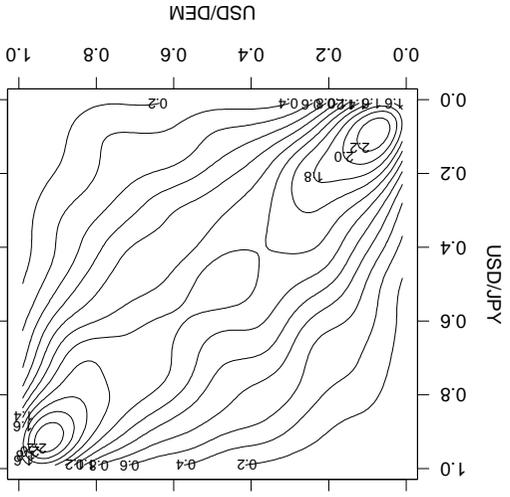
8 Hours returns



12 Hours returns



1 Day returns



FITTING STATIC COPULA MODELS

SUMMARY OF THE RESULTS

- Use pseudo-likelihood method
- t-copula fits best across all frequencies
- Persistence of tail dependence
- Data are non-elliptical for higher frequencies

Reference: [1]

MARGINAL GARCH TYPE FILTERING

Univariate ARMA-GARCH model:

$$X_t = \mu_t + \epsilon_t$$

$$\mu_t = \mu + \sum_{p_1=1}^i \phi_i (X_{t-i} - \mu) + \sum_{q_1=1}^j \theta_j \epsilon_{t-j}$$

$$\epsilon_t = \sigma_t Z_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{p_2=1}^i \alpha_i \epsilon_{t-i}^2 + \sum_{q_2=1}^j \beta_j \sigma_{t-j}^2$$

where $Z_t \sim t_v$

Marginal residuals: $\hat{z}_t = (x_t - \hat{\mu}_t) / \hat{\sigma}_t$

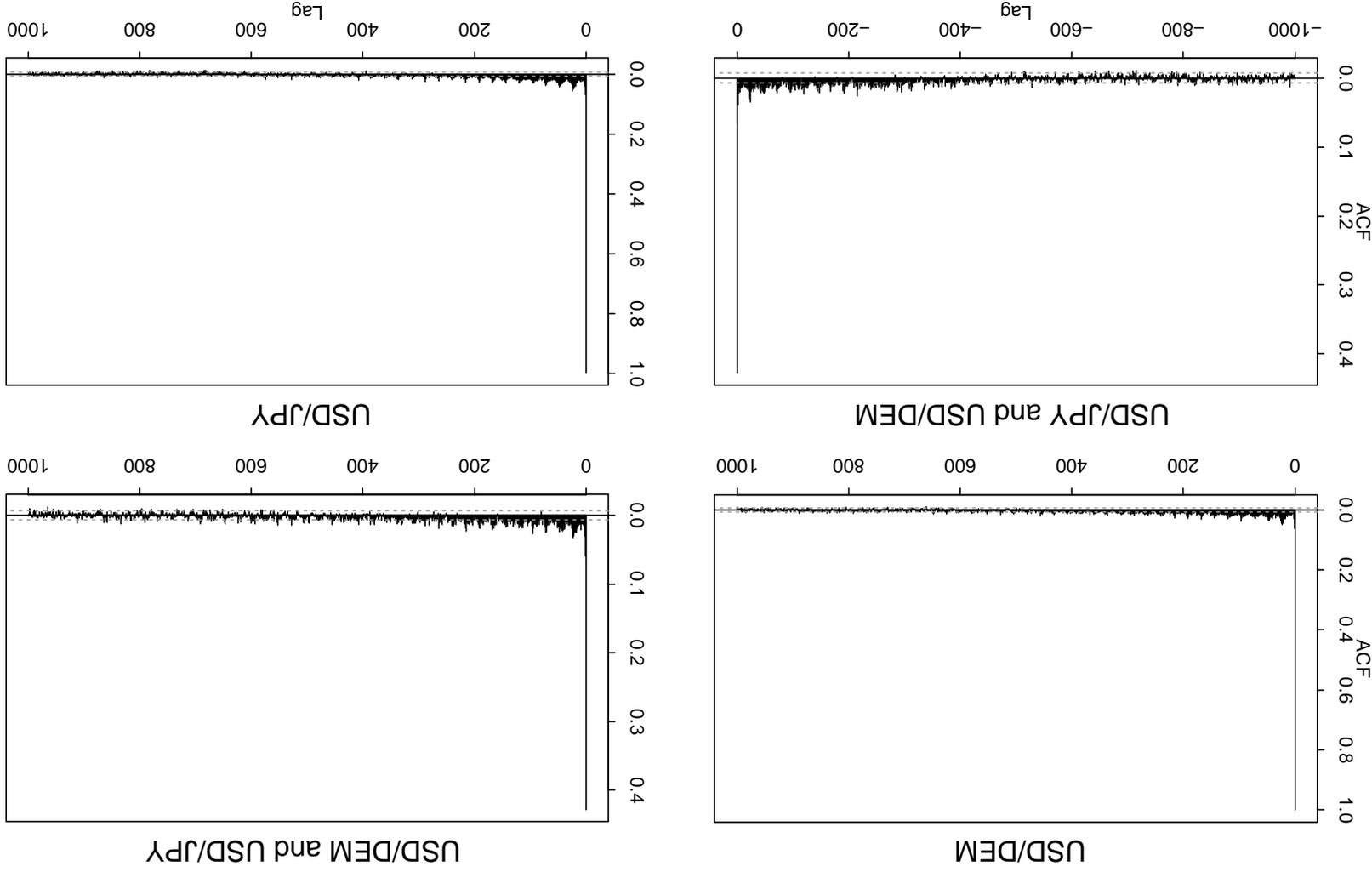
* leverage effect: $\sigma_t^2 = \alpha_0 + \sum_{p_2=1}^i \alpha_i |\epsilon_{t-i}| + \gamma_i \epsilon_{t-i} + \sum_{q_2=1}^j \beta_j \sigma_{t-j}^2$

Frequency	p_1	q_1	p_2	q_2	USD/DEM	\hat{v} (s.e.)
1 hour	—	—	1	1	3.693 (0.054)	
2 hours	2	2	1	1	3.708 (0.044)	
4 hours	—	5	1	1	3.975 (0.105)	
8 hours	2	4	1	1	4.679 (0.234)	
12 hours	1	—	1	1	5.385 (0.326)	
1 day	1	—	1	1	5.797 (0.556)	

Frequency	p_1	q_1	p_2	q_2	USD/JPY	\hat{v} (s.e.)
1 hour	—	—	1	1	3.654 (0.052)	
2 hours	1	—	2	1	3.759 (0.077)	
4 hours	4	4	2*	1	3.819 (0.109)	
8 hours	2	2	1*	1	4.357 (0.195)	
12 hours	1	—	1*	1	4.574 (0.251)	
1 day	10	—	1*	1	4.889 (0.412)	

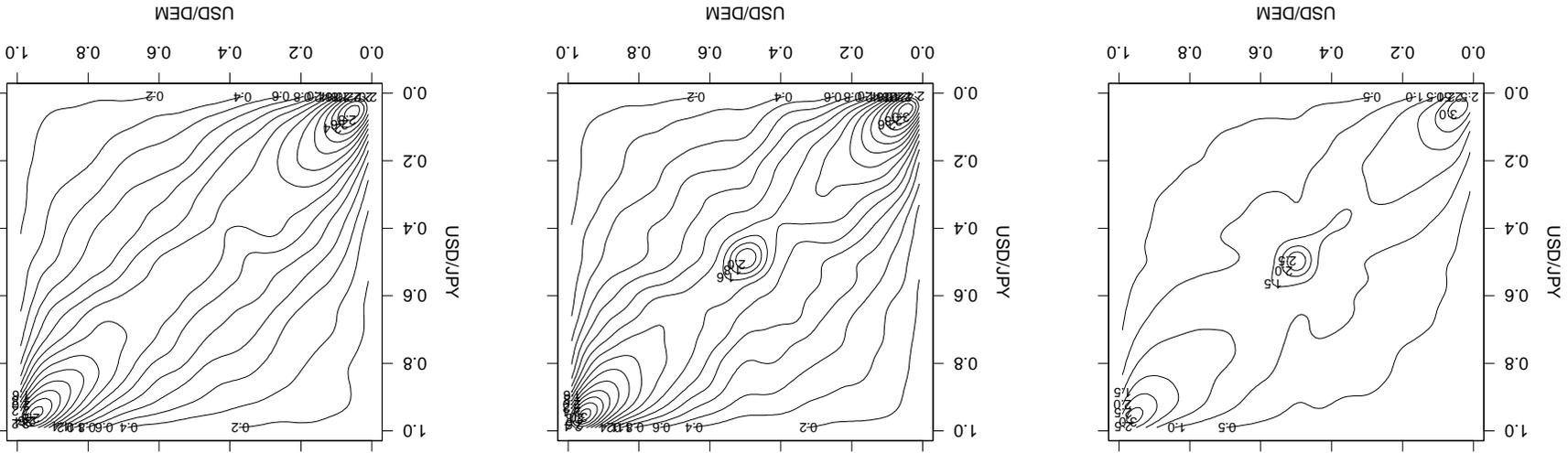
Sample autocorrelograms

Absolute values of the one hour
USD/DEM and USD/JPY residuals

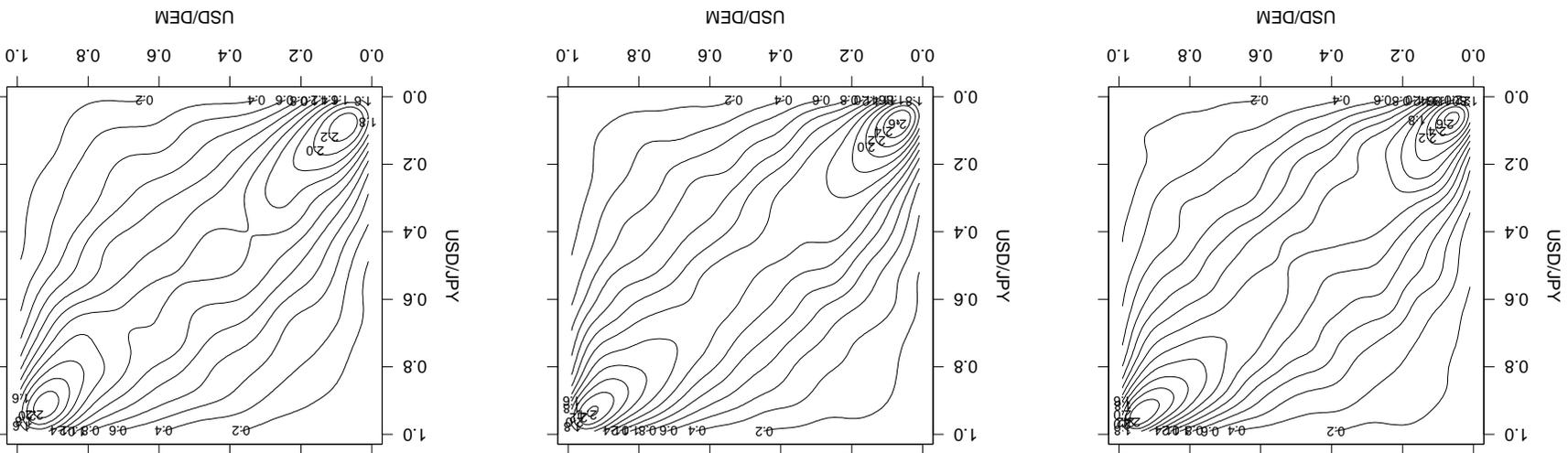


CONTOUR-PLOTS OF FILTERED USD/DEM AND USD/JPY RETURNS

1 Hour residuals 2 Hours residuals 4 Hours residuals

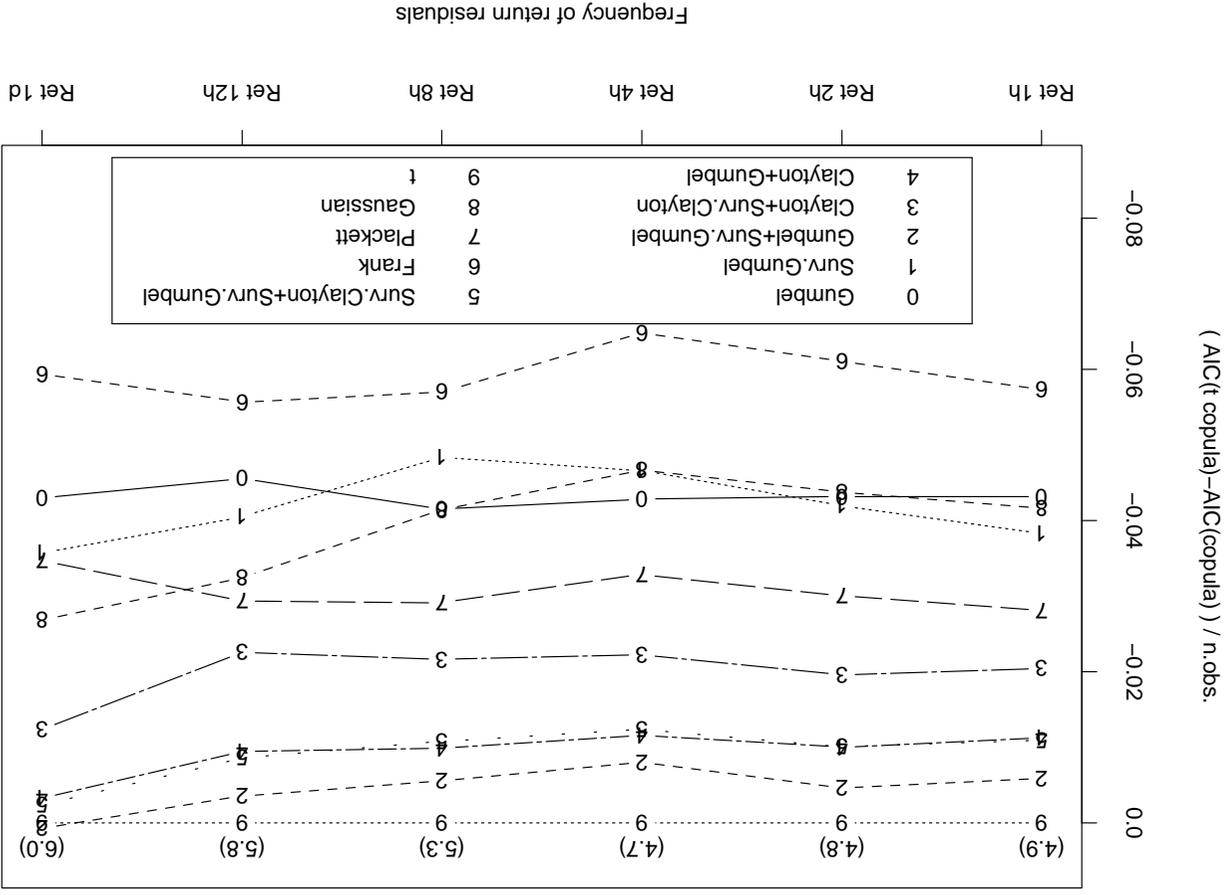


8 Hours residuals 12 Hours residuals 1 Day residuals



GOODNESS OF FIT

FOR DIFFERENT FREQUENCIES RELATIVE TO THE t-COPULA



GOODNESS OF FIT

of the t and Gumbel mixture models
fitted to the residuals

Frequency	sample size	t-model	Gumbel mixture
1 hour	78,239	0	0
2 hour	39,119	0	0
4 hour	19,559	0.0348	0.0006
8 hour	9,779	0.3808	0.1079
12 hour	6,519	0.2471	0.1949
1 day	3,259	0.7211	0.6775

ELLIPTICALITY TEST FOR THE BIVARIATE RESIDUALS

Frequency	Original margins	t margins
1 hour	0	0
2 hours	0	0
4 hours	0	0.348
8 hours	0.001	0.069
12 hours	0.145	0.501
1 day	0.389	0.451

TAIL-DEPENDENCE ANALYSIS

- Tail-dependence coefficient

Let X_1 and X_2 be random variables with distributions F_1 and F_2 , such that

$$\lim_{u \leftarrow 1^-} P \left(X_2 \geq F_2^{-1}(u) \mid X_1 \geq F_1^{-1}(u) \right) = \lambda^U$$

exists. If $\lambda^U \in (0, 1]$ then (X_1, X_2) has upper tail-dependence coefficient λ^U and (X_1, X_2) has no upper tail dependence if $\lambda^U = 0$

- Spectral analysis

Assumption: Multivariate regular variation of (X_1, X_2)

- Asymptotic clustering

Bivariate Archimedean excesses have a Clayton copula (asymptotically)

[3] Juri and Wüthrich (2002). Copula convergence theorems for tail events. Insurance Math. Econom., 30: 405–420

TAIL-DEPENDENCE COEFFICIENT ESTIMATES FOR THE RESIDUALS

t copula	λ	λ_T	λ_U
1 hour	0.242	0.209	0.250
2 hour	0.261	0.207	0.269
4 hour	0.273	0.265	0.225
8 hour	0.261	0.216	0.289
12 hour	0.247	0.288	0.224
1 day	0.240	0.286	0.226

SPECTRAL ANALYSIS

- Suppose that the d -dimensional random vector \mathbf{X} has a **regularly varying tail distribution**, i.e., the tail behaviour of \mathbf{X} is characterised by a tail index α and the limit

$$P(\|\mathbf{X}\| > tx, \mathbf{X}/\|\mathbf{X}\| \in \cdot) \xrightarrow{v} x^{-\alpha} P(\Theta \in \cdot),$$

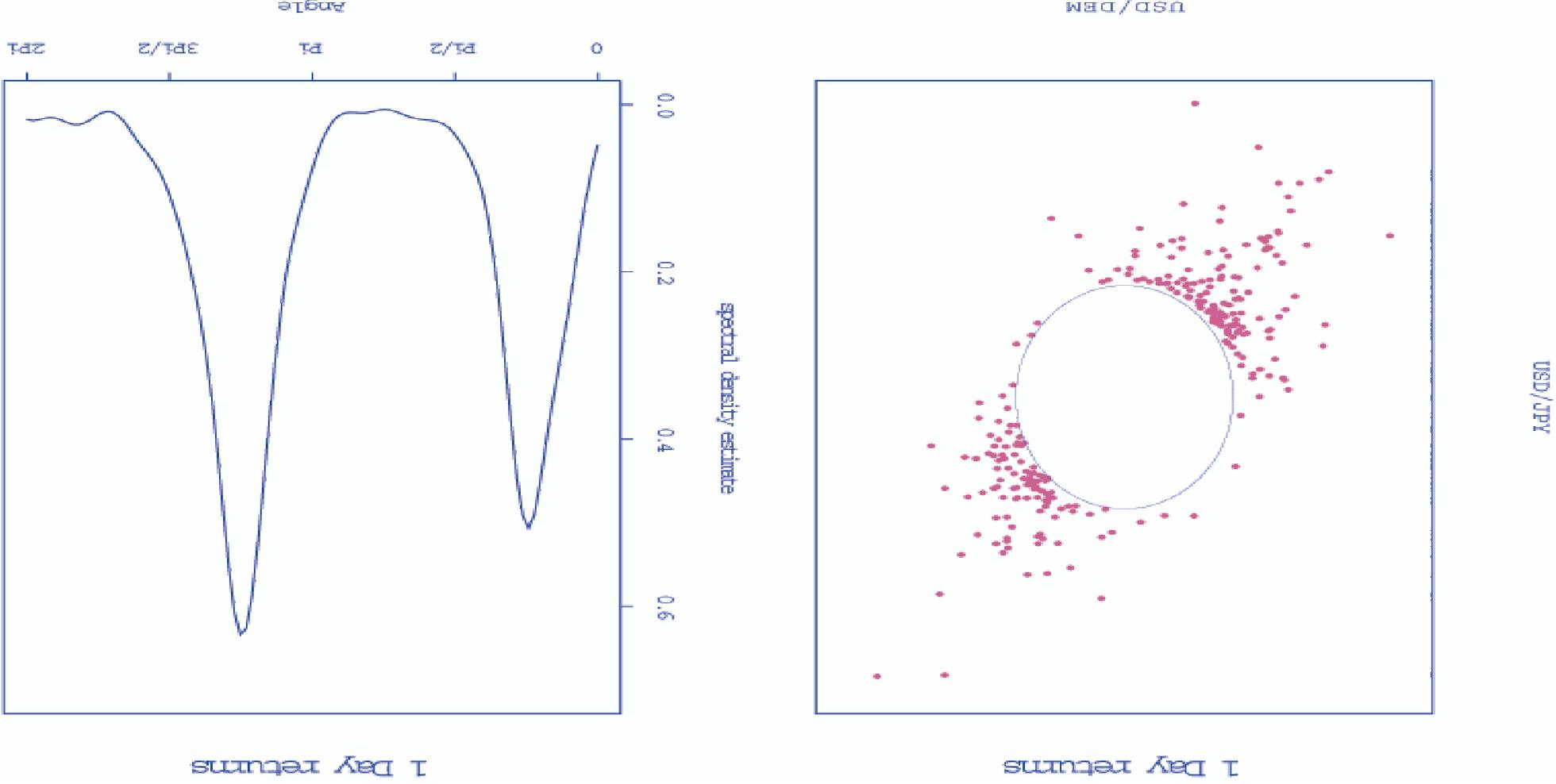
where $x > 0$, $t \rightarrow \infty$, exists. The distribution function of Θ is the **spectral distribution** of \mathbf{X}

- **Estimator:**

$$\hat{P}(\Theta \in S) = \frac{1}{n} \sum_{i=1}^{k_{n,n}} \epsilon_{\mathbf{X}_i / \|\mathbf{X}_i\|_{k_{n,n}}} V(S)$$

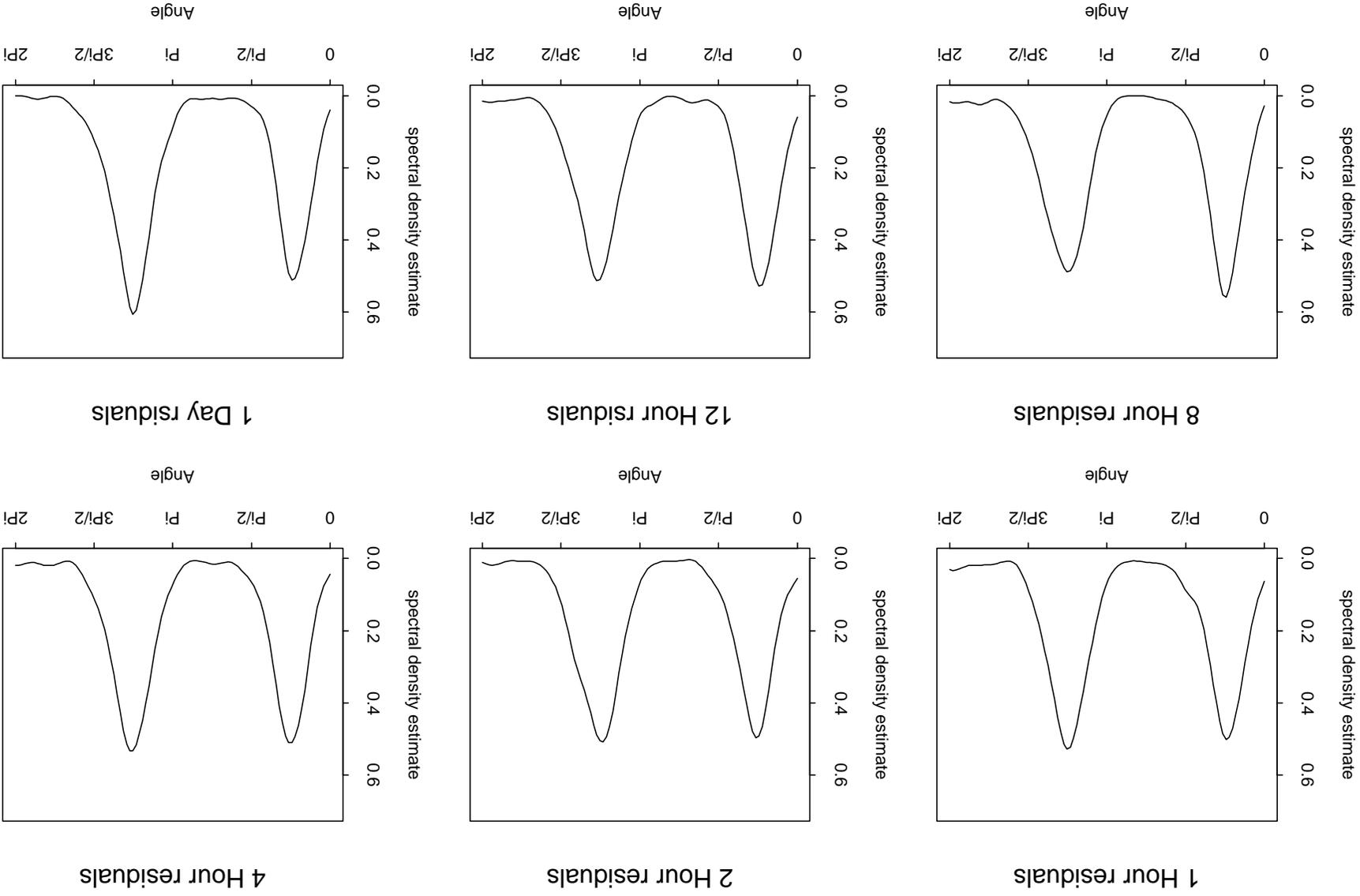
where $V(S) = \{\mathbf{x} \in S_{d-1}^+ : \mathbf{x} / \|\mathbf{x}\| \in S\}$

SPECTRAL MEASURE ESTIMATION



SPECTRAL ANALYSIS

FOR THE BIVARIATE RESIDUALS



ASYMPTOTIC CLUSTERING OF BIVARIATE EXCESSES

- **Extreme tail** dependence copula relative to a threshold t :

$$C_t(u, v) = P(U \leq F_t^{-1}(u), V \leq F_t^{-1}(v) | U \leq t, V \leq t)$$

with conditional distribution function

$$F_t(u) := P(U \leq u | U \leq t, V \leq t), 0 \leq u \leq 1$$

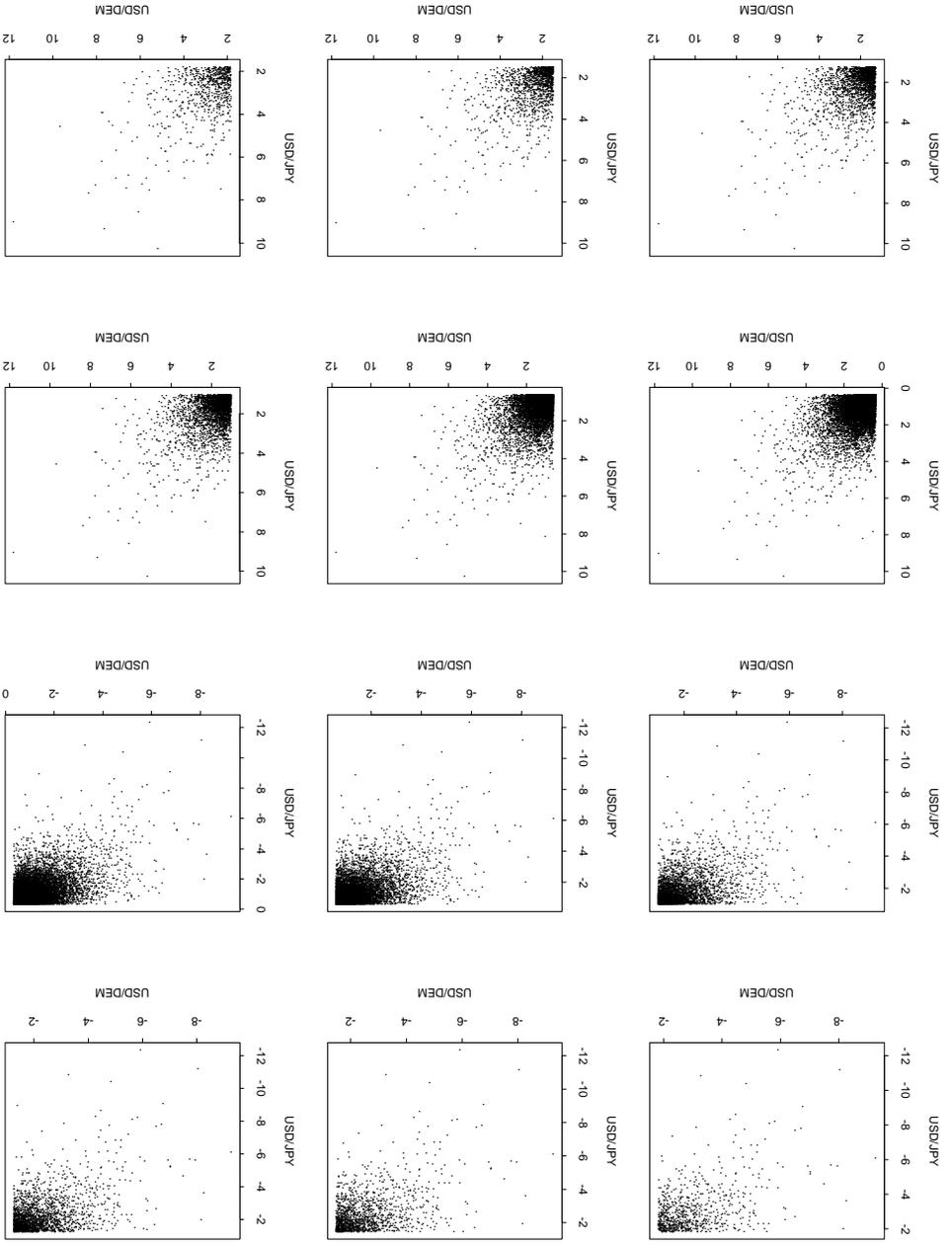
- **Archimedean** copulas: \exists cont., strictly decreasing function $\phi : [0, 1] \mapsto [0, \infty]$ with $\phi(1) = 0$, s.t.

$$C(u, v) = \phi_{[-1]}(\phi(u) + \phi(v))$$

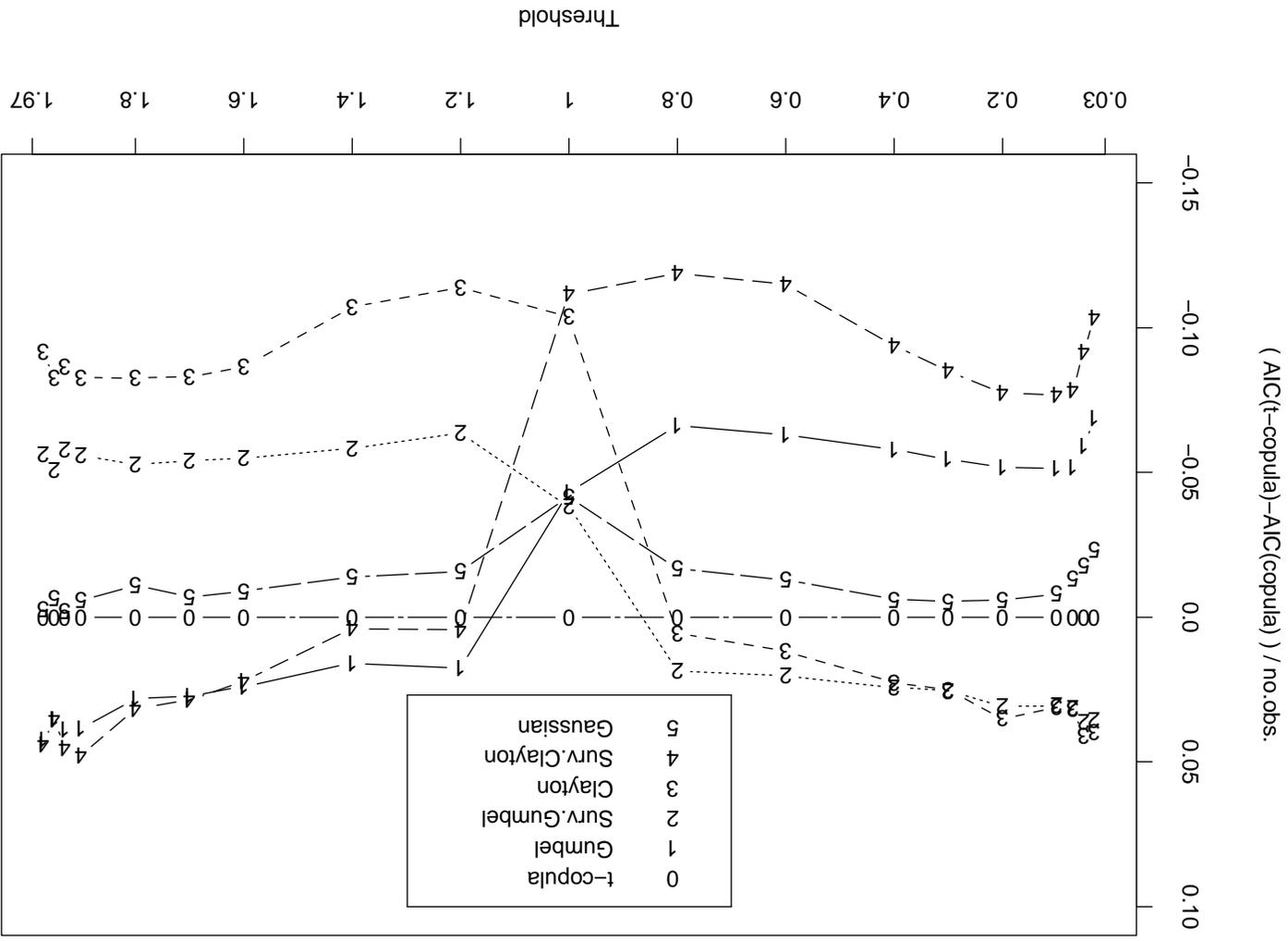
- For “sufficiently regular” **Archimedean** copulas (Juri and Wüthrich (2002)):

$$\lim_{t \rightarrow 0^+} C_t(u, v) = C_{\alpha}^{\text{Clayton}}(u, v)$$

BIVARIATE EXCESSES FOR DIFFERENT THRESHOLDS



COPULA FITTING OF MULTIVARIATE EXCESSES



MULTIVARIATE (MATRIX-DIAGONAL) GARCH

d -dimensional matrix-diagonal GARCH process with an AR component:

$$\begin{aligned} X_t &= \mu_t + \epsilon_t \\ \mu_t &= \mu + \sum_{i=1}^{p_1} M_i (X_{t-i} - \mu) \\ \epsilon_t &= \Sigma_t^{1/2} Z_t \end{aligned}$$

$$\Sigma_t = A_0 A_t^0 + \sum_{i=1}^{p_2} (A_i A_i^0) \otimes (\epsilon_{t-i} \epsilon_{t-i}^0) + \sum_{j=1}^q (B_j B_j^0) \otimes \Sigma_{t-j}$$

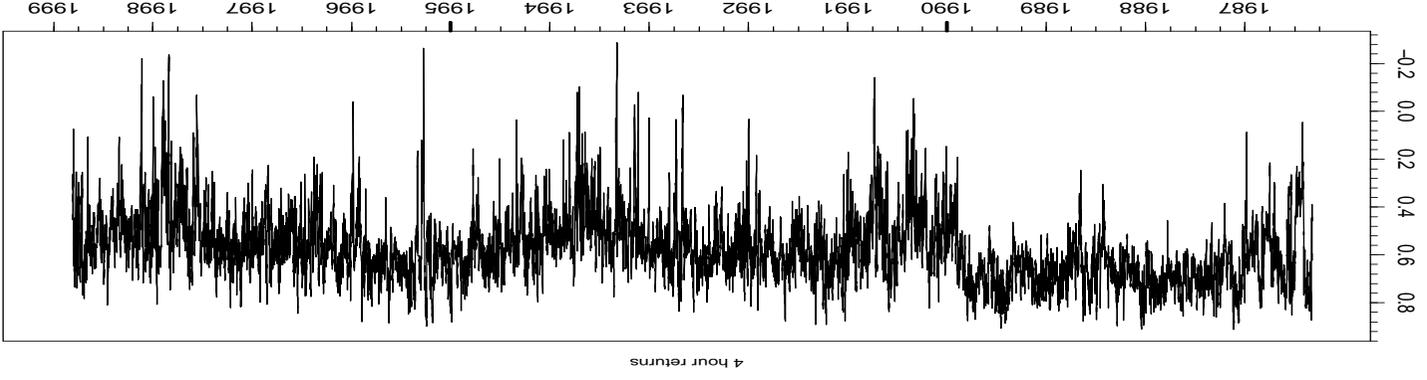
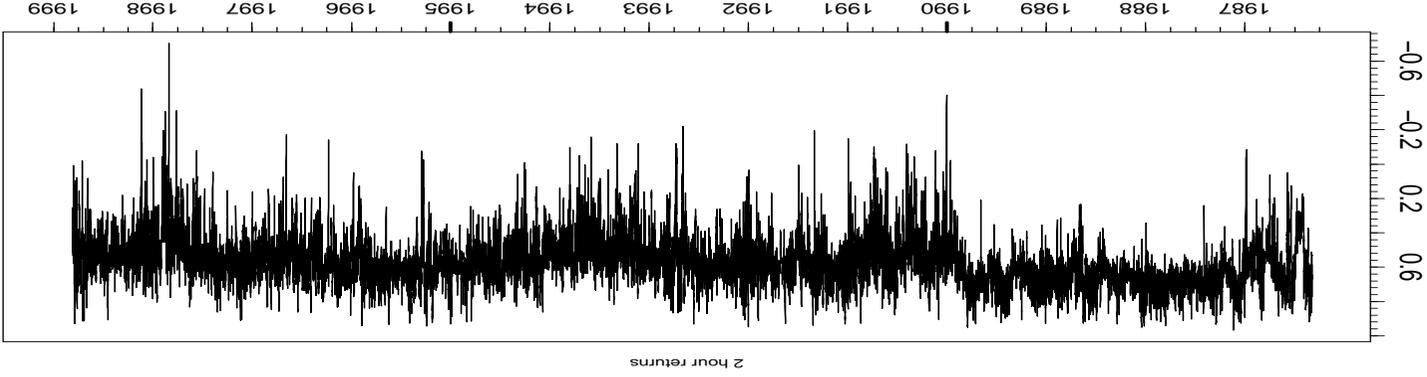
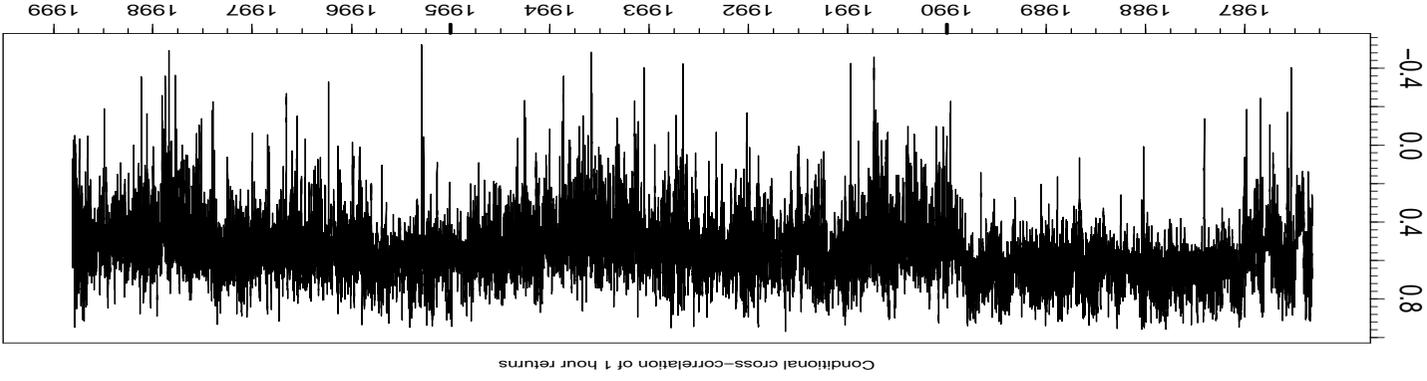
where

A_i and B_j are lower triangular $d \times d$ matrices
 M_i is a full matrix in $\mathbb{R}^{d \times d}$

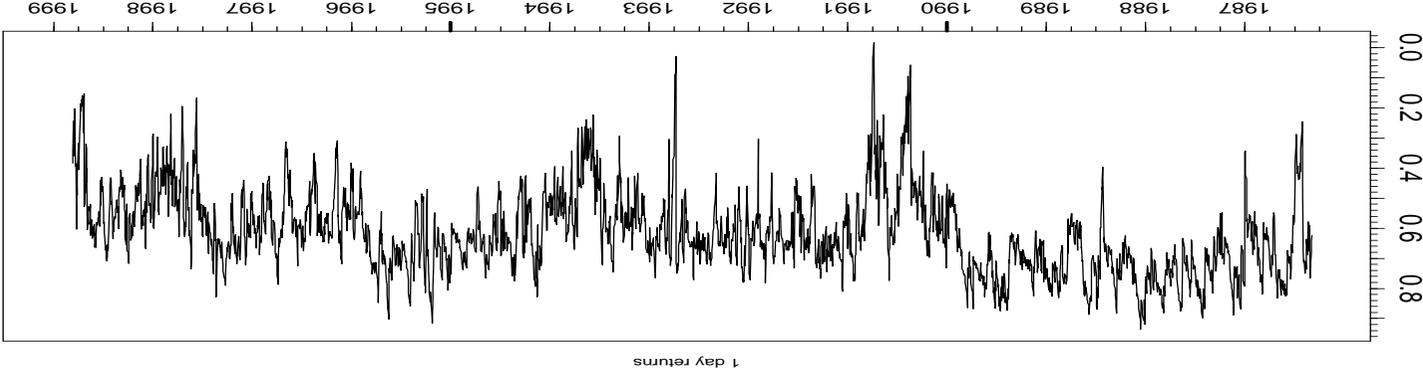
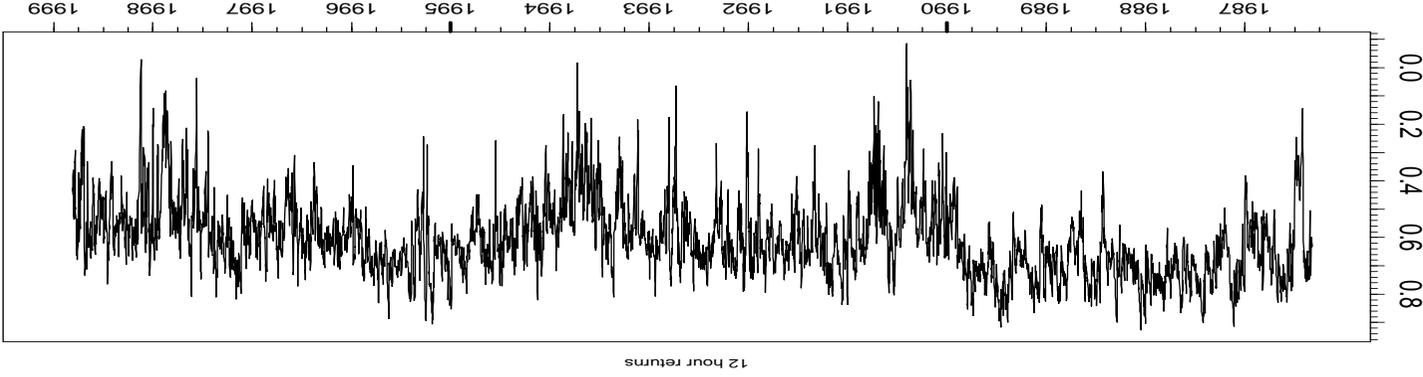
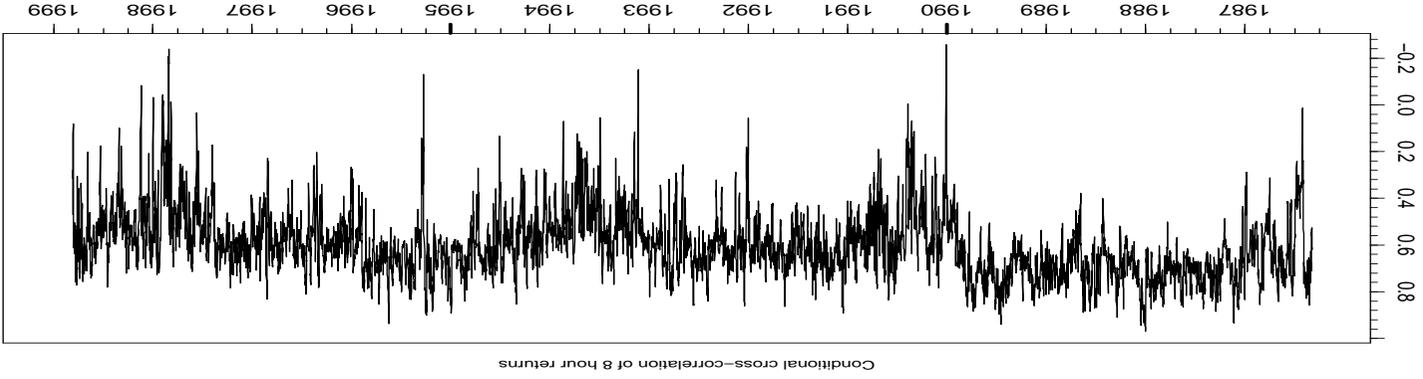
$(Z_t)_{t \in \mathbb{Z}}$ is an iid vector sequence with zero mean vector and unit variances
 Σ_t is the conditional covariance matrix of the vector ϵ_t

\otimes is the Hadamard product

CONDITIONAL CROSS-CORRELATION ESTIMATED BY MATRIX-DIAGONAL GARCH



CONDITIONAL CROSS-CORRELATION ESTIMATED BY MATRIX-DIAGONAL GARCH



TIME-VARYING COPULAS ACROSS TIME SCALES

- Multivariate GARCH model with time-varying copula:

$$X_t = c + \epsilon_t$$

$$\epsilon_t = \sigma_t Z_t$$

$$\sigma_t^2 = A_0 + \sum_{i=1}^p A_i (\epsilon_{t-i} \epsilon_{t-i}^T) + \sum_{j=1}^q B_j \otimes \sigma_{t-j}^2$$

where A_i and B_j are diagonal $d \times d$ matrices

- Assume that Z_t has a bivariate t-copula with time dependent parameters

$$v_t = \nu \quad \text{for all } t,$$

$$p_t = h_{-1}^{-1}(r_0 + r_1 z_{1,t-1} z_{2,t-1} + s_1 h(p_{t-1})),$$

where $h(\cdot)$ is Fisher's transformation for the correlation

$$h(d) = \log \left(\frac{1+d}{1-d} \right).$$

Dynamic copula modelling

- Standardised residual return series obtained from the univariate filtering:
$$\{(z_{1,t}, z_{2,t}) : t = 0, \dots, n\}$$

- Standardized residual returns mapped into $[0, 1]^2$:

$$\left\{ \left(t_{\hat{v}_1} \left(\sqrt{\frac{\hat{v}_1}{\hat{v}_1 - 2}} z_{1,t}, t_{\hat{v}_2} \left(\sqrt{\frac{\hat{v}_2}{\hat{v}_2 - 2}} z_{2,t} \right) \right) : t = 1, \dots, n \right\} \cdot$$

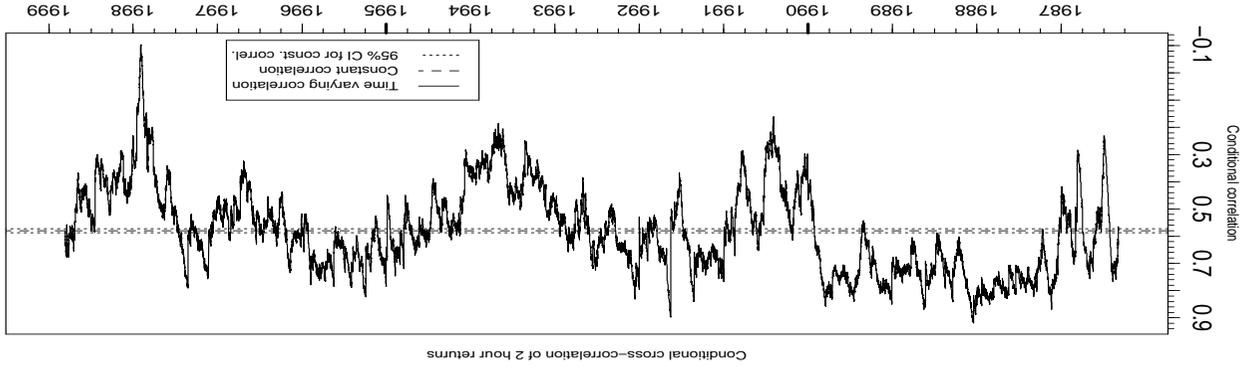
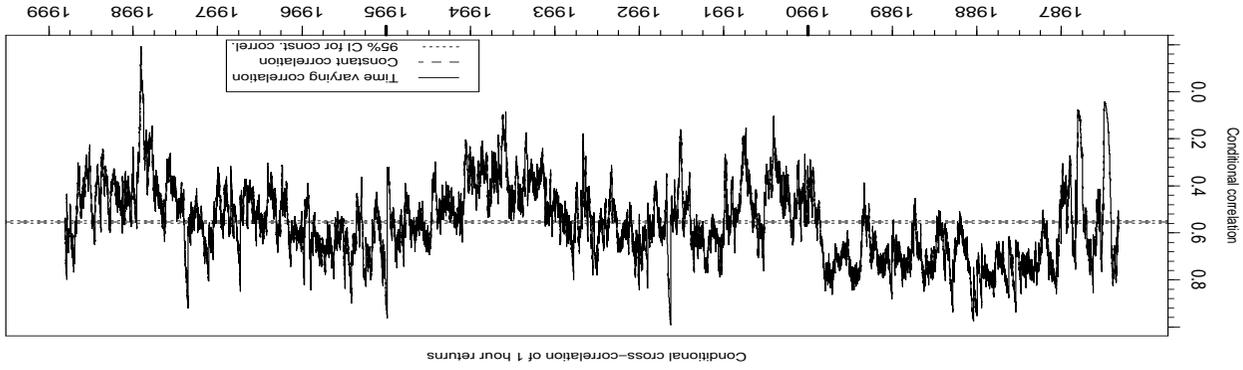
Dynamic copula results

Time frequency	non-dynamic		dynamic	
	\hat{v}	\hat{d}	\hat{v}	\hat{d}
1 hour	4.935 (0.108)	0.558 (0.002)	6.330 (0.167)	0.0005 (0.0002)
			\hat{r}_1	0.0193 (0.0010)
			\hat{s}_1	0.9921 (0.0005)
	AIC	-31517.70	AIC	-34488.72
2 hours	4.822 (0.147)	0.580 (0.003)	6.203 (0.230)	-0.0004 (0.0002)
			\hat{r}_1	0.0128 (0.0009)
			\hat{s}_1	0.9952 (0.0004)
	AIC	-17192.73	AIC	-19349.29
4 hours	4.669 (0.195)	0.592 (0.005)	6.072 (0.313)	-0.0008 (0.0002)
			\hat{r}_1	0.0147 (0.0011)
			\hat{s}_1	0.9947 (0.0004)
	AIC	-9085.848	AIC	-10262.23

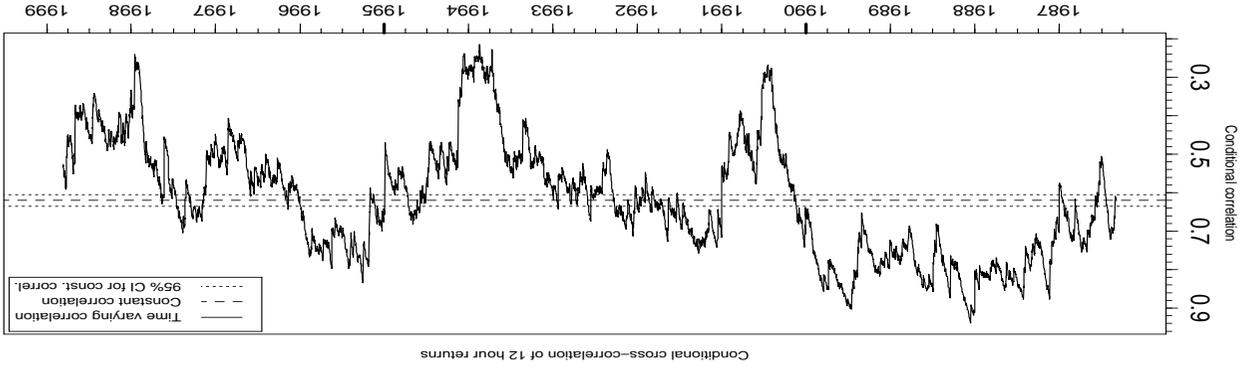
Dynamic copula results

Time frequency	non-dynamic	dynamic
8 hours	$\hat{\nu}$ 5.296 (0.339) $\hat{\rho}$ 0.612 (0.006) AIC -4813.6	$\hat{\nu}$ 7.206 (0.584) $\hat{\rho}_0$ 0.0005 (0.0005) \hat{r}_1 0.0173 (0.0014) \hat{s}_1 0.9927 (0.0006) AIC -5456.312
12 hours	$\hat{\nu}$ 5.830 (0.499) $\hat{\rho}$ 0.620 (0.008) AIC -3299.16	$\hat{\nu}$ 8.053 (0.884) $\hat{\rho}_0$ 0.0002 (0.0008) \hat{r}_1 -0.0249 (0.0023) \hat{s}_1 0.9901 (0.0010) AIC -3744.28
1 day	$\hat{\nu}$ 5.945 (0.758) $\hat{\rho}$ 0.619 (0.011) AIC -1644.549	$\hat{\nu}$ 8.573 (1.455) $\hat{\rho}_0$ -0.0023 (0.0017) \hat{r}_1 -0.0343 (0.0041) \hat{s}_1 0.9846 (0.0021) AIC -1881.760

Time-varying cross-correlations estimated by the time-varying copula-based model



Time-varying cross-correlations estimated by the time-varying copula-based model



THE GENERAL CHANGE-POINT PROBLEM

X_1, X_2, \dots, X_n independent, from

$$C(x; \theta_1, \eta_1), \dots, C(x; \theta_n, \eta_n)$$

$$H_0: \theta_1 = \theta_2 = \dots = \theta_n \quad \text{and} \quad \eta_1 = \eta_2 = \dots = \eta_n$$

$$H_A: \theta_1 = \dots = \theta_{k^*} \neq \theta_{k^*+1} = \dots = \theta_n \quad \text{and} \quad \eta_1 = \eta_2 = \dots = \eta_n$$

Likelihood ratio test

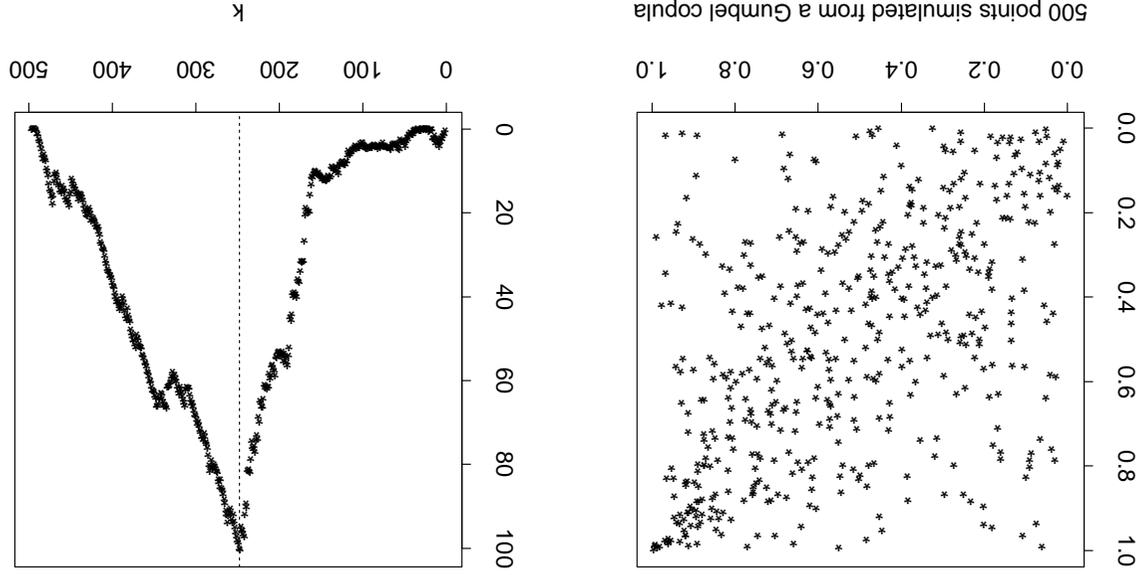
$$Z_n = \max_{1 \leq k \leq n} (-2 \log(\Lambda_k))$$

with

$$\Lambda_k = \frac{\prod_{1 \leq i \leq k} f(X_i; \theta, \eta) \prod_{k < i \leq n} f(X_i; \theta', \eta')}{\prod_{1 \leq i \leq n} f(X_i; \theta, \eta)}$$

CHANGE-POINT DETECTION

A SIMULATED EXAMPLE



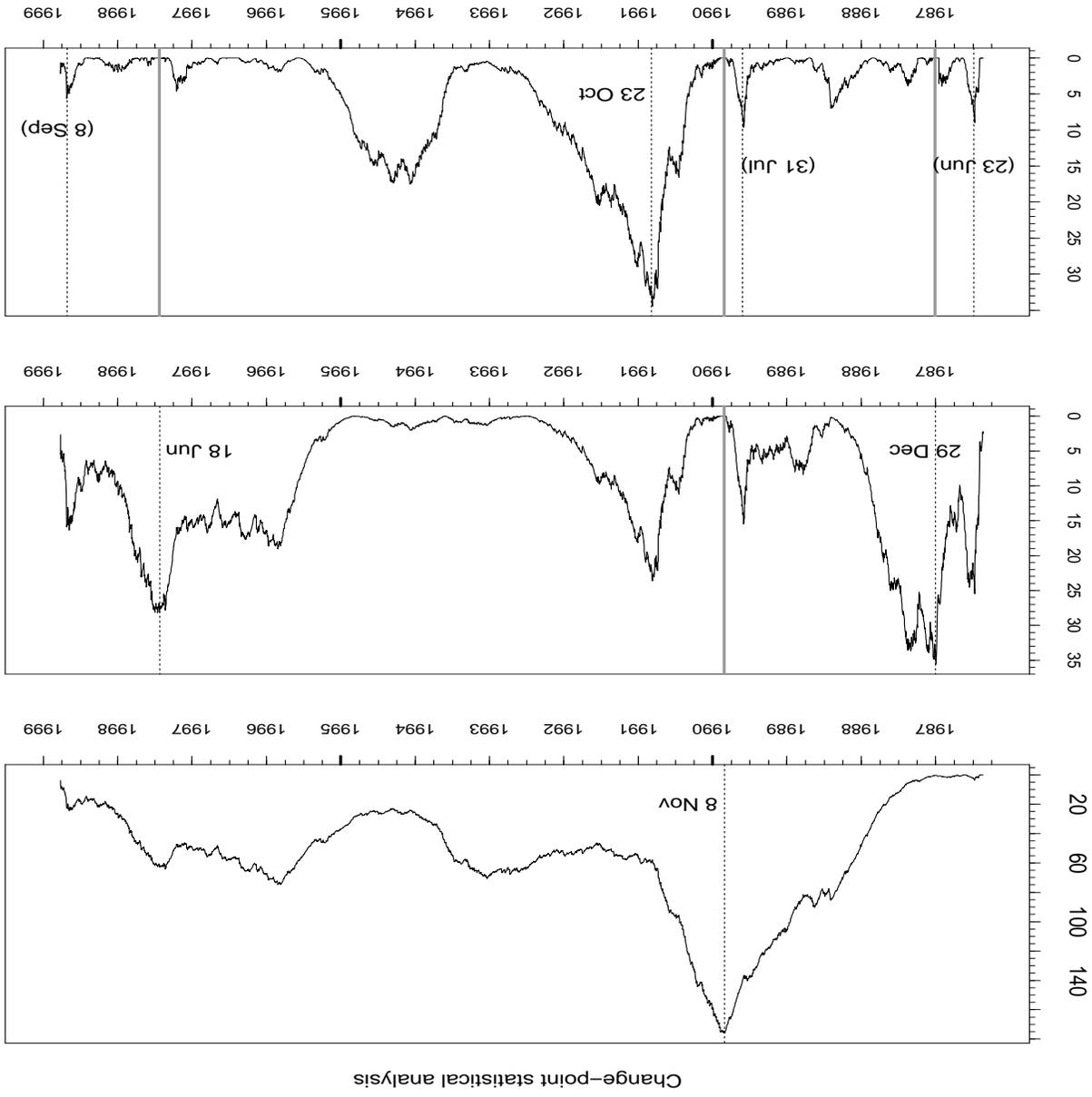
SEVERAL PROBLEMS

- The power of the test
- Boundary changes
- Size of the change
- Confidence intervals

Change-point analysis for USD/DEM and USD/JPY spot rate residuals

$z_{n\text{ obs}}^{1/2}$	n	$P\left(Z_n^{1/2} > z_{n\text{ obs}}^{1/2}\right)$	$H_0(0.95)$	Time of change
13.26	3 259	0	reject	8 Nov. 1989
5.96	923	0.0000004	"	29 Dec. 1986
5.31	2 336	0.0000143	"	18 June 1997
2.99	176	0.0689621	not rej.	(23 June 1986)
3.10	747	0.0709747	"	(31 July 1989)
5.86	1 985	0.0000007	reject	23 Oct. 1990
2.36	351	0.3380491	not rej.	(8 Sep. 1998)
2.78	1 736	0.1873493	"	(21 Oct. 1996)
2.86	249	0.1061709	"	(21 March 1990)

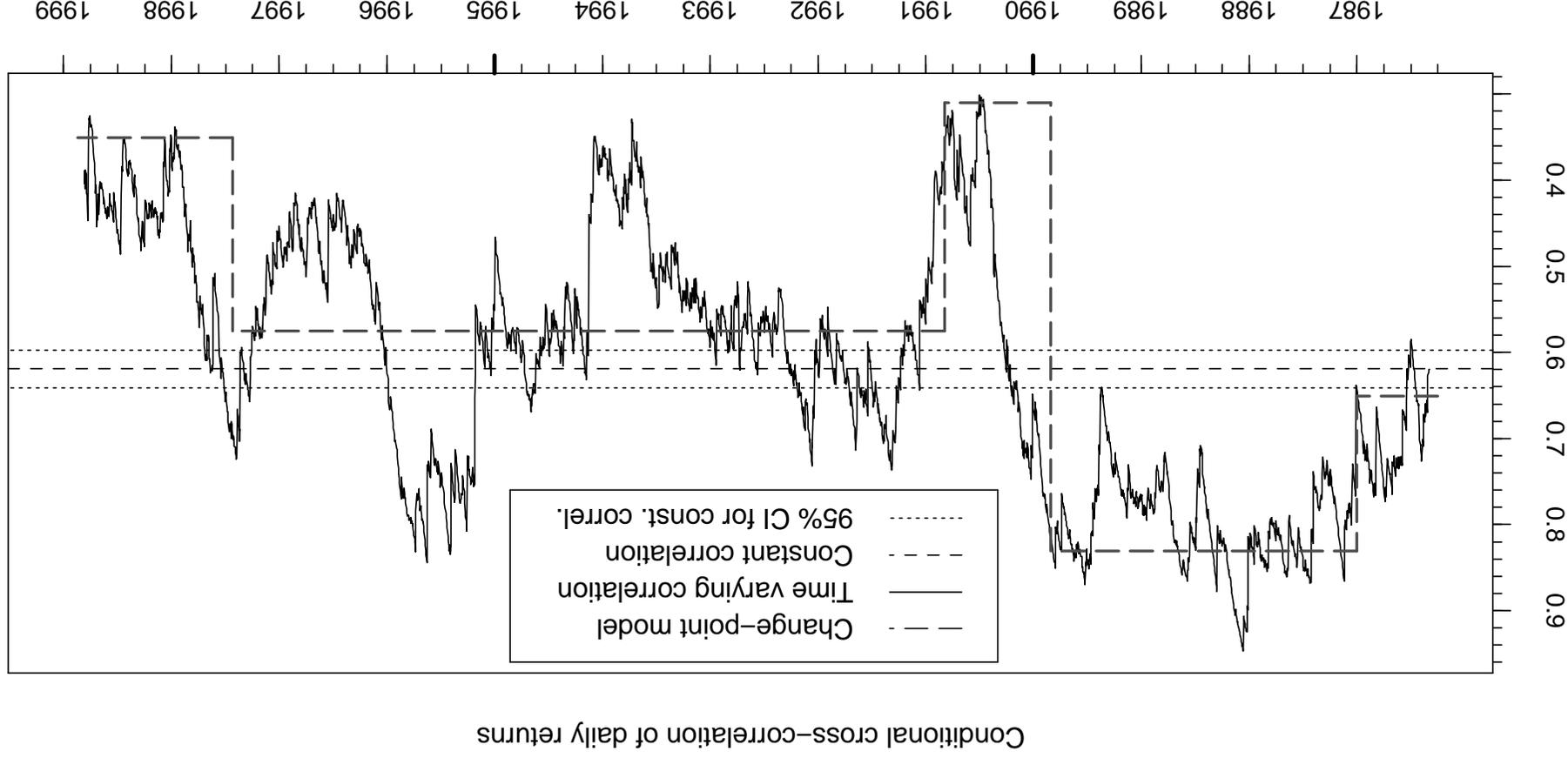
Change-point analysis for USD/DEM and USD/JPY spot rate residuals



Change-points detected on the USD/DEM and USD/JPY spot rate residuals

- December, 29 1986
- November, 8 1989: fall of the Berlin wall
- October, 23 1990: burst in the Japanese asset price bubble
- June, 18 1997: beginning of the Asia crisis

Estimated t-copula conditional correlation of daily returns on the FX USD/DEM and USD/JPY spot rates



Summary and further work

Summary:

- Deserialisation for multivariate tick-by-tick data
- Static versus dynamic copula fitting
- t-copula is overall fine
- Persistent tail-dependence
- Tested for ellipticity
- Change-point analysis

Further work:

- $d > 2$, other models
- Deserialisation, multivariate time change
- Tests for parameter constancy
- Link to multivariate EVT
- General Juri and Wüthrich (2002) result