QUANTIFYING EEG SYNCHRONY USING COPULAS

Satish G. Iyengar[a], Justin Dauwels[b], Pramod K. Varshney[a] and Andrzej Cichocki[c]

[a] Dept. of EECS, Syracuse University, NY
[b] Massachusetts Institute of Technology, Cambridge, MA
[c] RIKEN Brain Science Institute, Japan

ABSTRACT

In this paper, we consider the problem of quantifying synchrony between multiple simultaneously recorded electroencephalographic signals. These signals exhibit nonlinear dependencies and non-Gaussian statistics. A copula based approach is presented to model the joint statistics. We then consider the application of copula derived synchrony measures for early diagnosis of Alzheimer's disease. Results on real data are presented.

Index Terms— EEG, Copula theory, Kullback-Leibler divergence, Statistical dependence

1. INTRODUCTION

Quantifying synchrony between electroencephalographic (EEG) channels is helpful for detection of many neurophysiological diseases. For example, a decrease in synchrony among different channels is indicative of Alzheimer's disease (AD) while enhanced synchronization is known to be related to epileptic seizures. In this paper, we consider the problem of detecting the onset of AD. It has been reported that AD is the sixth leading cause of death in the United States [1]. Deaths due to AD have increased by as much as 47.1% between the years 2000-2006. Medication for AD is most effective when applied at an early stage.

Many studies have reported that AD causes a decreased inter-channel dependence when compared to their age-matched control subjects. It has been shown that EEG signals of AD patients exhibit reduced coherence when compared to the age-matched control patients (See [2, 3] for an in-depth review). Several measures have been proposed for EEG synchrony quantification ([3–8] provide recent reviews on EEG synchrony measures). However, reliable detection of AD remains to be a difficult problem especially for patients in the pre-symptomatic phase (also known as Mild Cognitive Impairment (MCI)). This poses severe limitations in the diagnosis of early stage AD. While some measures such as correlation coefficient, corr-entropy coefficient, coh-entropy and mutual information are bivariate measures and hence are applicable only to pairs of channels, measures such as Granger causality, even though multivariate, may fail to account for any nonlinear inter-channel dependence.

In this work, we propose a copula based approach to quantify synchrony between brain signals recorded from multiple locations on the scalp. The joint probability density function (PDF) $f(\mathbf{z_1}, \mathbf{z_2}, \ldots, \mathbf{z_N})$ defines the joint statistics of observations from N EEG channels and thus captures all of the synchrony information that may be present between the EEG channels. Measuring dependence in such a multivariate setting has been largely carried out assuming multivariate Gaussian distribution. This is because dependence quantification is straightforward in the Gaussian case. For example, a zero-correlation always implies statistical independence. However, this is not true with other distributions. The Gaussian model fails to characterize any nonlinear dependence and higher order correlations that may be present between the variables. Further, the multivariate Gaussian model constrains the marginals to also follow the Gaussian disribution. As we will show in the following sections, copula models help alleviate these limitations. Following the same approach as in [4], we then evaluate the classification (normal vs. subjects with MCI) performance using the leave-one-out cross validation procedure. We use the same data set as in [4] thus allowing comparison with other synchrony measures studied in [4].

The paper is structured as follows. We discuss copula theory and its use in modeling multivariate time series data in Section 2. Section 3 describes the EEG data set used in the present analysis and considers the application of copula theory for EEG synchrony quantification. Results are reported in Section 4 followed by some concluding remarks in Section 5.

2. COPULA THEORY

Copulas are functions that "stitch" the univariate marginal distributions to form a multivariate distribution function. Sklar (1959) was the first to define copula functions.

Theorem 1 *(Sklar's Theorem)*
Let $F(\bar{\mathbf{z}})$ be the joint cumulative distribution function (CDF) with continuous marginals $F_1(z_1)$, $F_2(z_2)$, $\cdots$, $F_N(z_N)$. Then there exists a copula function $C(\cdot)$ such that for all $z_1, z_2, \cdots, z_n$ in $[-\infty, \infty]$,

$$F(\bar{\mathbf{z}}) = C(F_1(z_1), F_2(z_2), \cdots, F_N(z_N)). \tag{1}$$

For continuous marginals, $C(\cdot)$ is unique; otherwise $C(\cdot)$ is uniquely determined on $RanF_{Z_1} \times RanF_{Z_2} \cdots \times RanF_{Z_N}$ where $RanX$ denotes the range of X. Conversely, if $C(\cdot)$ is a copula and $F_1(z_1)$, $F_2(z_2)$, $\cdots$, $F_N(z_N)$ are marginal CDFs then the function $F(\bar{\mathbf{z}})$ in (1) is a valid joint CDF with the given marginals.

The N^{th} order differentiation of (1) gives an expression for the copula based joint PDF.

$$f^c(\bar{\mathbf{z}}) \quad = \quad \underbrace{\left(\prod_{n=1}^{N} f_n(z_n)\right)}_{f^p(\bar{\mathbf{z}})} c(F_1(z_1), \cdots, F_N(z_N)), \tag{2}$$

where we use the *superscript c* to denote that $f^c(\bar{\mathbf{z}})$ is the copula representation of the joint PDF. The function $f^p(\bar{\mathbf{z}})$, in the RHS of (2), is the zero-dependence or the product distribution. It is interesting to note the form of (2). The copula density re-weights the product distribution to incorporate the dependence structure between the N

variables. Copula models thus factorize a joint distribution such that the dependence structure is separated from the product of marginals. One can thus construct joint distributions with arbitrary marginals and the desired dependence structure.

2.1. Copula based Synchrony Measure

Let $z_{1k}, z_{2k}, \ldots, z_{Nk}$ denote measurements from N different channels where k is the time index. In the context of time series modeling, copulas have been utilized to characterize both (a) the temporal dependence, *i.e.,* by considering the conditional PDF $f\left(z_{i(k)}|z_{i(k-1)}, \ldots, z_{i(k-m)}\right)$ (m is the model order), and (b) the multivariate dependence across multiple time series $\{z_{1k}, z_{2k}, \ldots z_{Nk}\}$ conditioned on some past information set $\mathcal{F}_{(k-1)}$ [9, 10]. They have been used intensively in modeling financial time series data (*e.g.*, [11, 12]).

In this paper, we consider the application of copulas to measure EEG synchrony. Specifically, we use copulas to estimate the $d-$dimensional joint PDF $f\left(\{z_1\}_{(k-m)}^{k}, \ldots, \{z_N\}_{(k-m)}^{k}\right)$ where $d = N \cdot (m+1)$. Now, although Sklar's theorem proves the existence of a copula density $c(\cdot)$ for any multivariate PDF with continuous marginals such that it can be expressed as in (2), identifying the true underlying copula density is difficult. A common approach then is to select a copula density (say $h(\cdot)$) *a priori* and fit the desired marginals. Several copula functions have been defined in the literature [13]. However, the popular ones especially in the multivariate setting are the Gaussian and the Student's t copula functions.

The d-dimensional Gaussian copula is given as

$$h_g(u_1, \ldots, u_d) = \frac{\phi_R(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d))}{\phi(\Phi^{-1}(u_1)) \cdots \phi(\Phi^{-1}(u_d))}, \tag{3}$$

where the index R refers to the correlation matrix and $u_i = F_i(z_i) \sim \mathcal{U}(0,1)$ (uniform distribution). Φ and ϕ are standard Gaussian CDF and PDF respectively. Like other copulas, it provides the flexibility of having disparate marginals.

The t-copula is given as,

$$h_t(u_1, \ldots, u_d) = \frac{f_{\nu,R}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d))}{\prod_{i=1}^{d} f_\nu(t_\nu^{-1}(u_i))}, \tag{4}$$

where $f_{\nu,R}(\cdot)$ and $t_\nu(\cdot)$ are the standard $t-$joint density and CDF with ν degrees of freedom respectively. Thus, in addition to the correlation matrix R, the t-copula is also parameterized by ν - the degrees of freedom (4), which is directly related to the so called tail dependence or the joint probability of extreme events [14]. For example, in the case of a bivariate random vector (X_1, X_2), the tail dependence measures the nature of $c(F_1(x_1), F_2(x_2))$ in the upper-right and lower-left quadrants of $\mathbb{I}^2$ - the unit square [13]. It is quantified by the upper and lower tail dependence coefficients λ_u and λ_l defined as

$$\lambda_u = \lim_{q \to 1} P(X_2 > F_2^{-1}(q) | X_1 > F_1^{-1}(q)) \tag{5}$$

$$\lambda_l = \lim_{q \to 0} P(X_2 \leq F_2^{-1}(q) | X_1 \leq F_1^{-1}(q)) \tag{6}$$

provided the limits exist.

For the t-copula density, the upper and lower tail dependence coefficients are equal and can be derived as

$$\begin{aligned} \lambda_u = \lambda_l &= 2t_{\nu+1}\left(-\sqrt{(\nu+1)}\frac{\sqrt{1-\rho}}{\sqrt{1+\rho}}\right) \\ &= \lambda. \end{aligned} \tag{7}$$

Thus, ν is inversely proportional to the tail dependence; the number of degrees of freedom ν decreases with increasing tail dependency λ. Further, as $\nu \to \infty$, the t-copula approaches the Gaussian copula.

Now, given the measurements, we use the canonical maximum likelihood (CML) approach to infer the parameters of the copula density (ν for the t-copula and R for both the t and Gaussian copula densities); the approach allows us to estimate the copula parameters without any assumptions on the parametric forms of the marginal distributions [15]. The entries of the correlation matrix R quantify the coupling between different signals and can be used to define synchrony measures. In the Gaussian copula, this coupling is linear, while it is nonlinear in the case of the t-copula. Moreover, we suggest the use of multiinformation [16] as a global synchrony measure which can summarize the whole dependence structure as a single number.

Multiinformation '$\mathcal{I}(\cdot)$' between d random variables $X_1, \ldots, X_d$ is given as

$$\begin{aligned} \mathcal{I}(X_1; \cdots; X_d) &= \int f(X_1, \ldots, X_d) \log \frac{f(X_1, \ldots, X_d)}{\prod_{i=1}^{d} f(X_i)} \\ &= \int f(X_1, \ldots, X_d) \log \frac{f^p(\cdot)c(\cdot)}{f^p(\cdot)} \qquad (8) \\ &= \mathbb{E}_f \log c(\cdot), \qquad (9) \end{aligned}$$

where $\mathcal{I}(\cdot) \geq 0$. It is zero when variables are statistically independent.

However, note that it is impossible to compute (9) since the true joint PDF $f(\cdot)$ and hence the copula $c(\cdot)$ is unknown. We approximate the integral over $f(\cdot)$ in (8) and (9) by a sample expectation, and replace the true copula $c(\cdot)$ by a trial copula $h(\cdot)$ chosen *a priori*, e.g., Gaussian or t-copula, leading to the following expression:

$$\hat{\mathcal{I}}_h(\cdot) = \frac{1}{L}\sum_{l} \log h(F_1(x_{1l}), \cdots, F_d(x_{dl})). \tag{10}$$

where L denotes the sample size[1].

Application of an arbitrary copula density $h(\cdot)$ ($\neq c(\cdot)$) introduces model mismatch errors and the performance of a copula based algorithm is thus highly dependent on the choice of $h(\cdot)$. We quantify this error due to copula misspecification using the Kullback Leibler (KL) divergence $\left(D(f||f^h)\right)$ between the true joint PDF and its copula based estimate given as,

$$\begin{aligned} D(f||f^h) &= \int f \log \frac{f}{f^h} = \int f \log \frac{f}{f^p} - \int f \log h(\cdot) \\ &= \int f \log c - \int f \log h \qquad (11) \\ &= \mathcal{I}(\cdot) - \underbrace{\mathbb{E}_f \log h(\cdot)}_{\mathcal{I}_h(\cdot)}. \qquad (12) \end{aligned}$$

where (10) can be used to approximate $\mathcal{I}_h(\cdot)$. Model mismatch error (quantified by the KL divergence D) is zero only when $h(\cdot) = c(\cdot)$, since then $f^h(\cdot) = f(\cdot)$. Note that $\mathcal{I}(\cdot)$, the first term in the RHS of (12), is fixed for given $f(\cdot)$. Thus, higher the value of $\mathcal{I}_h(\cdot)$, smaller is the mismatch error.

[1] For example, for $N = 2$ and model order m,

$$\hat{\mathcal{I}}_h(\cdot) = \frac{1}{L}\sum_{k} \log h(u_{1k}, \ldots, u_{1(k-m)}, u_{2k}, \ldots, u_{2(k-m)}),$$

where $u_{il} = F_{il}(z_{il})$.

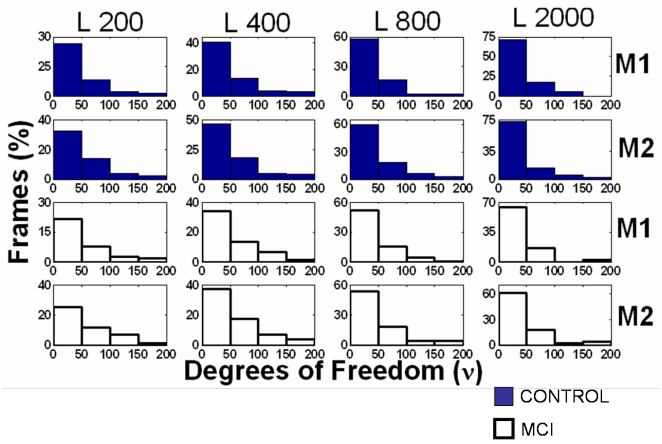

Fig. 1. Normalized histogram of ν estimates. Lower values for ν suggest non-zero tail dependence.

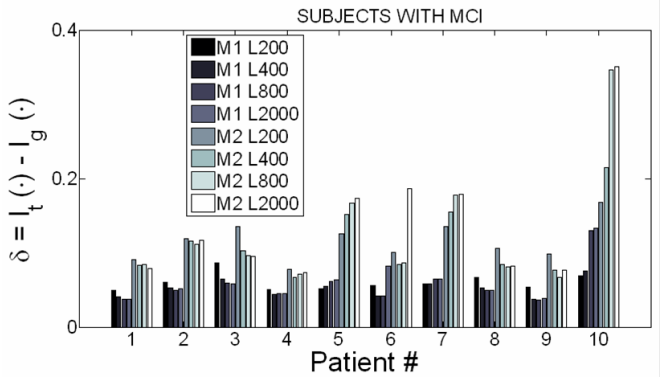

Fig. 2. Difference in multiinformation estimates based on t and Gaussian copula functions

The proposed copula based approach is applicable for modeling and analysis of generic time series. However, our focus in this study is EEG synchrony quantification which we discuss below.

3. EXPERIMENT DESCRIPTION AND MODEL FITTING

In this section, we describe the EEG data in brief and also address the issue of selecting $h(\cdot)$ that *best* fits the data.

3.1. EEG Data

Silver (Ag/AgCl) electrodes (disks of diameter 8mm) were placed on 21 sites according to 10-20 international system, with the reference electrode on the right ear-lobe. EEG were obtained with Biotop 6R12 (NEC San-ei, Tokyo, Japan) at the rate of 200 samples per second. The acquired data was then band pass filtered (4 Hz - 30 Hz) using a third-order Butterworth filter. A common reference (right ear-lobe) was used for data analysis.

The subjects comprised two study groups. The first was a group with 25 patients diagnosed as suffering from MCI when the EEG recordings were carried out. Later on, they all developed mild AD. The criteria for inclusion into the MCI group were a mini mental state exam (MMSE) score = 24, though the average score in the MCI group was 26 (standard deviation 1.8). The second group known as control set consisted of 56 age-matched, healthy subjects who had no memory or other cognitive impairments. The average MMSE of this control group was 28.5 (standard deviation 1.6). The ages of the two groups are 71.9 ± 10.2 and 71.7 ± 8.3, respectively.

All recording sessions were conducted with the subjects in an awake but resting state with eyes closed. A 5-minute long EEG was recorded. Only those subjects were retained in the analysis whose EEG recordings showed maximum artifact-free data (ascertained by an EEG expert by visual inspection). Based on this requirement, the number of subjects in the two groups described above was further reduced to 22 and 38, respectively. From each subject, one artifact-free EEG segment of 20s was analyzed (for each of the 21 channels).

Further, the 21 EEG channels were clustered into five groups and the signals in each group were averaged thus obtaining five spatially averaged EEG signals per subject. (The five groups correspond to frontal, left temporal, central, right temporal and occipital regions. See [4] for more details.). Further analysis and synchrony results are obtained for the spatially averaged EEG data (*i.e.,* $N = 5$). Since the EEG time series is highly non-stationary, each signal is divided into small segments or frames, each of length L samples. We consider several frame lengths in our analysis ($L = 200, 400, 800, 2000$). Synchrony measures are obtained by averaging over all the frames.

3.2. Copula Selection

We now justify our preference for the t-copula over the Gaussian copula density for EEG synchrony quantification.

Figure 1 shows the normalized histogram of ν estimated from the spatially averaged EEG data for model orders one and two (M1 and M2) and for different frame lengths (L). (See [17] for extension of the definition and analysis of ν to the multivariate t-copula). It is evident from Fig. 1 that an appreciable number of frames do exhibit tail dependence (i.e. have smaller values for ν). For example, for M2 L2000, about 75 % of the total frames have $\nu < 50$. The Gaussian copula with zero upper and lower tail dependence will be unable to *capture* this characteristic of the EEG data. The t-copula can also characterize the *Gaussian-like* behavior of certain frames with larger values of ν as it *nests* the Gaussian copula (i.e., it approaches the Gaussian copula as $\nu \to \infty$).

Next, we estimate multiinformation for each frame by approximating $\mathbb{E}_f \log h(\cdot)$ by (10). We then average over all frames to obtain estimates of $\mathcal{I}(\cdot)$ using both the Gaussian copula ($\hat{\mathcal{I}}_g(\cdot)$) and the t-copula ($\hat{\mathcal{I}}_t(\cdot)$) functions. We plot in Fig. 2, the difference δ between the two estimates $\hat{\mathcal{I}}_t(\cdot)$ and $\hat{\mathcal{I}}_g(\cdot)$ of multiinformation for ten patients. It is clear from Fig. 2, that δ is always positive, i.e., $\mathcal{I}_t(\cdot) \geq \mathcal{I}_g(\cdot)$ for any choice of the model order and frame length. Thus, the t-copula provides a better estimate (in the KL divergence sense) of the true joint PDF than the Gaussian copula density. Though results for only ten patients are presented here, we observed that the non-negativity of δ was true for all patients. In fact the t-copula was a better fit for every frame as well (*i.e.*, $\delta \geq 0$ for every frame). Similar results were obtained for all patients in the control set.

Thus, the above analysis indicates that the t-copula is more suitable for EEG synchrony quantification.

4. RESULTS

In the above sections, we discussed the copula approach to jointly model multiple EEG signals. Having learnt the copula parameters

Table 1. Classification rates using linear and quadratic discriminant analysis with leave-one-out cross validation

Features	Linear	Quadratic	Copula Features
ρ_{SES}	70.0%	70.0%	-
ffDTF	68.3%	75.0%	-
ρ_{SES}, ffDTF	83.3%	83.3%	-
ρ_{SES}, ffDTF, $h_g(\cdot)$	85.0%	80.0%	$\bar{\sigma}_\rho$
	85.0%	81.7%	$\bar{\mu}_\rho$
	83.3%	80.0%	$\hat{\mathcal{I}}_g(\cdot)$
ρ_{SES}, ffDTF, $h_t(\cdot)$	85.0%	81.67%	$\bar{\sigma}_\rho$
	85.0%	83.30%	$\bar{\mu}_\rho$
	83.3%	81.67%	$\hat{\mathcal{I}}_t(\cdot)$
	85.0%	85.0%	ν

from data, we now proceed to the task of distinguishing MCI patients from the age-matched control subjects. In order to compare our results with [4], we use the same classification method - both linear and quadratic discriminant analysis with leave-one-out cross validation [18] as used in [4].

For each frame of length L samples, we first compute μ_l, the mean of the absolute values of all entries of the correlation matrix R estimated using CML. Two features, $\bar{\mu}_\rho$ and $\bar{\sigma}_\rho$, are then obtained by computing the average and standard deviation of μ_l over all frames. Estimates for the multiinformation $\hat{\mathcal{I}}_h$ and the number of degrees of freedom ν are also obtained in the similar way (*i.e.*, by averaging over all the frames).

Table 1 shows the leave-one-out classification rate for different features obtained with optimal parameter settings (model order and frame length). Classification rates obtained using the Gaussian copula are also included. Results in rows 1, 2 and 3 of Table 1 have been reported in [4] and correspond to the stochastic event synchrony measure (ρ_{SES}), the full frequency directed transfer function (ffDTF) - a Granger causality measure [5], and the combination of both respectively. It can be seen that using copula measures (both Gaussian and Student's t) in conjunction with ρ_{SES} and ffDTF improves the classification rate to 85%. Several existing synchrony measures were considered in [4] and the authors observed no improvement beyond 83.3%.

Next, we compare the Gaussian copula and the t-copula models (*e.g*, classification rates obtained using $[\rho_{SES}, \text{ffDTF}, \hat{\mathcal{I}}_g(\cdot)]$ vs. that using $[\rho_{SES}, \text{ffDTF}, \hat{\mathcal{I}}_t(\cdot)]$). It can be seen that the performance of the t-copula feature set is never lower than that of the Gaussian copula density. This is probably because the t-copula nests the Gaussian copula as discussed in Section 3.2.

5. CONCLUDING REMARKS

We presented a copula based approach for the quantification of EEG synchrony. We analyzed the two commonly used copula models, the multivariate Gaussian and the Student's t copula for modeling multiple EEG channels. The important issue of selecting the copula that best fits the data was also discussed. We also considered the problem of distinguishing MCI patients from the age-matched control subjects and show that copula based features when used in conjunction with other synchrony measures help enhance the detection of AD onset. Our results prove the feasibility of the copula approach and more general multivariate copula functions that are not limited to correlation matrix for dependence characterization will be investigated in the future.

6. REFERENCES

[1] "2009 Alzheimer's Disease, Facts and Figures, Executive Summary." http://www.alz.org/alzheimers_disease_facts_figures.asp.

[2] J. Jeong, "EEG dynamics in patients with Alzheimer's disease," *Clinical Neurophysiology*, vol. 115, pp. 1490–1505, 2004.

[3] P. Uhlhaas and W. Singer, "Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology," *Neuron*, vol. 52, pp. 155–168, 2006.

[4] J. Dauwels, F. Vialatte, and A. Cichocki, "A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG," *NeuroImage (in press)*. http://www.dauwels.com/files/NeuroImage2009.pdf.

[5] M. Kaminski and H. Liang, "Causal influence: Advances in neurosignal analysis," *Critical review in biomedical engineering*, vol. 33, pp. 347–430, 2005.

[6] E. Pereda, R. Q. Quiroga, and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signals," *Progress in Neurobiology*, vol. 77, pp. 1–37, 2005.

[7] C. J. Stam, "Nonlinear dynamical analysis of EEG and MEG: review of an emerging field," *Clinical Neurophysiology*, vol. 116, pp. 2266–2301, 2005.

[8] M. Breakspear, "Dynamic connectivity in neural systems: theoretical and empirical considerations," *Neuroinformatics*, vol. 2, 2004.

[9] A. J. Patton, *Handbook of Financial Time Series*, ch. Copula-Based Models for Financial Time Series, pp. 767–785. Springer, 2009.

[10] X. Chen and Y. Fan, "Estimation of copula-based semiparametric time series models," *Journal of Econometrics*, vol. 130, pp. 307–335, 2006a.

[11] G. M. Rockinger and E. Jondeau, "Conditional dependency of financial series: An application of copulas," *SSRN eLibrary*, 2001. http://ssrn.com/paper=262642.

[12] C. Genest, M. Gendron, and M. Bourdeau-Brien, "The advent of copulas in finance," *The European Journal of Finance*, vol. 00, pp. 1–10, 2009.

[13] R. B. Nelsen, *An Introduction to Copulas*. New York: Springer-Verlag, 1999.

[14] S. Demarta and A. J. McNeil, "The t copula and related copulas," *International Statistical Review*, vol. 73, pp. 111–129, 2005.

[15] E. Bouyé, V. Durrleman, A. Nikeghbali, G. Riboulet, and T. Roncalli, "Copulas for finance - a reading guide and some applications," *SSRN eLibrary*, 2000. http://ssrn.com/paper=1032533.

[16] M. Studený and J. Vejnarová, *Learning in graphical models*, ch. The multiinformation function as a tool for measuring stachastic dependence, pp. 261–297. MIT Press, 1999.

[17] Y. Chan and H. Li, "Tail dependence for multivariate t-copulas and its monotonicity," *Insurance: Mathematics and Economics*, vol. 42, pp. 763–770, 2008.

[18] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Wiley-Interscience, 2 ed., 2000.