# A true measure of dependence

Li, Hui

26 February 2016

# A true measure of dependence

Hui Li[1]

The strength of dependence between random variables is an important property that is useful in a lot of areas. Various measures have been proposed which detect mostly divergence from independence. However, a true measure of dependence should also be able to characterize complete dependence where one variable is a function of the other. Previous measures are mostly symmetric which are shown to be insufficient to capture complete dependence. A new type of nonsymmetric dependence measure is presented that can unambiguously identify both independence and complete dependence. The original Rényi's axioms for symmetric measures are reviewed and modified for nonsymmetric measures.

## 1. Introduction

The need for a measure to detect degree of dependence is an important task in a lot of areas, including economics, finance, artificial intelligence, engineering, bioinformatics and neuroscience. Although numerous measures have been proposed and studied with a vast amount of literature, a true measure of dependence is still elusive in the popular research. By a true measure of dependence, we mean a measure that can detect the dependence spectrum between two opposite extremes: independence and complete dependence, such that it takes the minimum value zero exactly on independence and takes the maximum value one exactly on complete dependence. Besides, the measure should be invariant under reasonably smooth transformations on the underlying variables under different parametrization. This also requires that the measure can detect nonlinear dependence. It is well-known that some measures satisfy the latter condition, such as Shannon's mutual information (Shannon and Weaver, 1949) and Hellinger distance (Granger *et al.*, 2004) as well as their generalizations, Rényi's mutual information (Rényi, 1961) and Tsallis entropy (Tsallis, 1988). These measures are also able to detect independence. However, they are more like measure of divergence from independence as they cannot detect for sure the other extreme of complete dependence or functional relationship. Besides, all these measures are symmetric in the underlying variables, such that they may not be able to measure directed dependence such as causality. Mutual information has been extended to

[1]Email: lichtung2014@gmail.com. The views expressed in this article are the author's own and do not represent the views of his employer.

MSC2010 subject classifications: 62H20, 62B10, 62G99, 94A17

Keywords: Nonsymmetric dependence measure, complete dependence, $*$ product on copula, Data Processing Inequality (DPI)

conditional mutual information through transfer entropy and is used for directed information and causality detection in dynamic systems and time series analysis (Hlaváčková-Schindler *et al.*, 2007). But it still cannot predict for sure if there is true causality. Recent research indicates that, in order to be able to detect complete dependence, especially for continuous random variables, we may have to rely on a new class of nonsymmetric dependence measures discovered in the past few years. These new measures are based on the concept of copula (Nelson, 2006), which captures the dependence structure between continuous random variables by stripping off the marginal distributions from the joint distribution. Copula is already very popular in the study of dependence in various areas (see, for example, Jaworski *et al.*, 2009). Many dependence measures can be re-casted into copula form, which gives them the benefit of not depending on marginal distribution and parametrization. But the proper way to use copula to capture complete dependence is beginning to be understood only recently through the research on the relation between copula function and Markov processes (Darsow *et al.*, 1992). As an analog of the Chapman-Kolmogorov equation in Markov process, a $*$ product is defined on copulas such that the set of copulas form the algebraic structure of a semigroup with a null element and a unit element. With the unit element, one can define the left and right inverse of a copula if they exist. This leads to the definition of several different subsets of the copulas: the set of left invertible copulas, the set of right invertible copulas, the set of copulas that are both left and right invertible, the set of copulas that are the $*$ product of a left invertible copula and a right invertible copula. The third set is the smallest, which is included in all other three sets. The last set is the largest which includes the other three sets. The first two sets have non-empty intersection but are otherwise different. It turns out that left invertible or right invertible copulas imply functional relation between the underlying variables, albeit in opposite direction. The copulas in the third set thus imply mutual complete dependence or 1-1 functional relationship. But the last set may not imply any functional dependence in either direction. It has been shown (Li, 2015a) that the symmetric dependence measures mentioned above all take the maximum value, which could be infinity, on the last set, such that they do not characterize complete dependence in either direction. To capture the functional relationship, which is directional, we have to introduce nonsymmetric dependence measures which take maximum value on either of the first two sets. This kind of measures has been discovered recently in Dette *et al.* (2010) and Trutschnig (2011), and a new framework for nonsymmetric dependence measures has been proposed in (Li, 2015a). It is the purpose of the current paper to introduce them to mainstream research.

## 2. Data processing inequality and symmetric dependence measures

Let us first introduce the concept of copula (Nelson, 2006). A bivariate copula is a bivariate distribution function on the two-dimensional unit square $[0,1] \times [0,1]$ with uniform marginals on $[0,1]$. Copulas are of interest because they link one-dimensional marginal distributions to joint distributions. Sklar (1959) showed that, for any continuous random variables $X, Y$ with joint distribution $F_{XY}$, there is a copula function $C(u, v)$ on $[0,1] \times [0,1]$ such that $F_{XY}(x, y) = C(F_X(x), F_Y(y))$, where $F_X$ and $F_Y$ represent the cumulative distribution functions of $X$ and $Y$ respectively.

Next we introduce a key concept from information theory, the Data Processing Inequality (DPI) (see Cover and Thomas, 1991). Random variables $X, Y, Z$ form a Markov chain if $X$ and $Z$ are conditionally independent given $Y$. A dependence measure $D(X, Y)$, which measures the dependence between $X, Y$, satisfies DPI if $D(X, Z) \leq D(Y, Z)$ whenever $X, Y, Z$ form a Markov chain. It is well-known that Shannon's mutual information, defined as

$$I(X, Y) = \iint dx\, dy\, f(x, y) log_2 \frac{f(x,y)}{f(x)(y)} \tag{1}$$

where $f(x, y)$ is the density function of the joint distribution of $X, Y$ and $f(x)$, $f(y)$ are the density functions of the marginal distributions of $X, Y$, satisfies the DPI condition. It characterizes the general loss of information when transmitted through a noisy communication channel.

Interestingly, the operation on the transition matrices for the Markov chain in the discrete case can be naturally expressed with copulas through the operation of the $*$ product in the continuous case. For any two copula functions A and B, the $*$ product is defined as

$$(A * B)(u, v) = \int_0^1 \partial_2 A(u, t) \cdot \partial_1 B(t, v) dt \tag{2}$$

If continuous random variables $X, Y, Z$ form a Markov chain, then $C_{XZ} = C_{XY} * C_{YZ}$, which was proved in Darsow *et al.* (1992).

It is known that the independent copula $\Pi(u, v) = uv$ is the null element and the copula $M(u, v) = \min(u, v)$ is the unit element for the $*$ product, such that, for any copula $C$,

$$C * \Pi = \Pi * C = \Pi \tag{3}$$

$$C * M = M * C = C \tag{4}$$

With the unit element, we can naturally define if a copula is left invertible or right invertible. So copula $C$ is left invertible if there exists a copula A such that $A * C = M$. Similarly, $C$ is called right invertible if there exists a copula B such that $C * B = M$. $C$ is invertible if it is both left and right invertible. Obviously $M$ is invertible and $\Pi$ is not invertible. Note that left or right invertible copulas are all singular copulas with support on a zero-measure set.

It has been proved in Darsow *et al.* (1992) that

(a) The copula between continuous random variables $X$ and $Y$ is left invertible if and only if there is a measurable function $f$ such that $Y = f(X)$ almost surely.

(b) The copula between continuous random variables $X$ and $Y$ is right invertible if and only if there is a measurable function $g$ such that $X = g(Y)$ almost surely.

Thus invertible copulas imply mutual complete dependence between two random variables. We denote the set of left invertible copulas as $\mathcal{L}$ and the set of right invertible copulas as $\mathcal{R}$. Then $\mathcal{L} \cap \mathcal{R}$ will be the set of invertible copulas. We also define the set $\mathcal{L} * \mathcal{R} = \{L * R, L \in \mathcal{L}, R \in \mathcal{R}\}$. It is easy to see that the set $\mathcal{L} * \mathcal{R}$ contains both $\mathcal{L}$ and $\mathcal{R}$ as subsets and is larger than $\mathcal{L} \cup \mathcal{R}$. As shown later in an example, the copula for circular relationship $X^2 + Y^2 = 1$ belongs to the set $\mathcal{L} * \mathcal{R}$, which does not imply a functional relationship and is also singular.

Now we are ready to discuss the first important result of this paper. As copula captures all the dependence information between two random variables, we would expect a dependence measure to be a function of copula for continuous random variables or $D(X,Y) = D(C_{XY})$. So the DPI condition can be generalized as $D(C_{XZ}) \leq D(C_{YZ})$ if $C_{XZ} = C_{XY} * C_{YZ}$, which includes the case of a Markov chain. A dependence measure is symmetric if $D(X,Y) = D(Y,X)$. For symmetric measures, DPI condition also implies $D(C_{XZ}) = D(C_{ZX}) \leq D(C_{YX}) = D(C_{XY})$.

Proposition. If a symmetric dependence measure satisfies the generalized DPI condition, then it takes its maximum value on $\mathcal{L} * \mathcal{R}$.

Proof. It turns out that the inverse of a left or right invertible copula is its transpose $C^T(u,v) = C(v,u)$, see Darsow *et al.* (1992). For any element $L * R \in \mathcal{L} * \mathcal{R}$, we have $L^T * L * R * R^T = M$. So

$$D(M) = D(L^T * (L * R) * R^T) \leq D((L * R) * R^T) \leq D(L * R)$$

But, as $D(L * R) = D(L * R * M) \leq D(M)$, we must have $D(L * R) = D(M)$. Obviously, $D(M)$ is the maximum value of the dependence measure. As $D(\Pi) = D(\Pi * C) \leq D(C)$, $D(\Pi)$ is the minimum value of the dependence measure. □

Therefore a symmetric dependence measure will not be able to characterize complete dependence if it satisfies DPI condition. It is shown in Li (2015a) that Shannon's mutual information in copula form

$$I(C) = \int_0^1 \int_0^1 c(u,v) \cdot \log(c(u,v)) du dv \quad (5)$$

as a special case ($\alpha \to 1$) of both Rényi's mutual information

$$R_\alpha(C) = \frac{1}{\alpha - 1} log \left[\int_0^1 \int_0^1 c^\alpha(u,v) du dv\right], \qquad \alpha > 0 \quad (6)$$

and Tsallis entropy,

$$\Delta_\alpha(C) = \frac{1}{\alpha - 1} \left[\int_0^1 \int_0^1 c^\alpha(u,v) du dv - 1\right], \qquad \alpha > 0 \quad (7)$$

and Hellinger distance as a scaled special case ($\alpha = \frac{1}{2}$) of Tsallis entropy,

$$H(C) = \int_0^1 \int_0^1 \left[1 - c(u, v)^{\frac{1}{2}}\right] dudv \tag{8}$$

all satisfy the DPI condition through Jensen's inequality as the dependence measures are convex functions of the copula density $c(u, v) = \partial_u \partial_v C(u, v)$. The general form of DPI for symmetric bivariate dependence measures was also discussed in Kinney and Atwal (2014). So the symmetric measures cannot be true measures of complete dependence as they can take maximum value at least on $\mathcal{L} * \mathcal{R}$ even when there may be no functional relationship between the underlying variables.

## 3. Nonsymmetric dependence measures

However, recent research on copulas has already revealed that true measures of complete dependence exist. These measures take maximum value only on the set of left or right invertible copulas such that they capture the true functional relationship. The salient feature is that they are all nonsymmetric, which should actually be expected as functional relationship is indeed nonsymmetric. Unlike the symmetric dependence measures which use copula density or copula itself, the nonsymmetric dependence measures use partial derivative of copula, or conditional cumulative distribution function (Darsow *et al.*, 1992)

$$\partial_1 C(u, v) = \frac{\partial C(u,v)}{\partial u} = P(V \le v | U = u) \tag{9}$$

where uniform random variables $U = F_X(X), V = F_Y(Y)$ have the joint distribution $C(u, v)$. The new measures are defined as the average distance between the cumulative distribution of one variable conditional on the other variable and the unconditional cumulative distribution of the one variable. More specifically, the measure has the general form (Li, 2015a),

$$\begin{aligned} \tau_\alpha(C) &= \left(\frac{(\alpha+1)(\alpha+2)}{2} \int_0^1 \int_0^1 |P(V \le v | U = u) - P(V \le v)|^\alpha dudv\right)^{1/\alpha} \\ &= \left(\frac{(\alpha+1)(\alpha+2)}{2} \int_0^1 \int_0^1 |\partial_1 C(u, v) - v|^\alpha dudv\right)^{1/\alpha}, \qquad \alpha \ge 1 \end{aligned} \tag{10}$$

which measures the dependence of $Y$ on $X$. In the special case of $\alpha = 2$, the measure becomes

$$\tau_2^2(C) = 6 \int_0^1 \int_0^1 (\partial_1 C(u, v) - v)^2 dudv \tag{11}$$

which was first discussed in Dette *et al.* (2010). The $\alpha = 1$ case was discussed in Trutschnig (2011).

It is easy to prove that $0 \leq \tau_2(C) \leq 1$. $\tau_2(C) = 0$ if and only if $\partial_1 C(u, v) = v$ or $C(u, v) = uv$, thus $X, Y$ are independent.

$$\tau_2^2(C) = 6 \int_0^1 \int_0^1 \partial_1 C(u, v)^2 du dv - 2$$

$$\leq 6 \int_0^1 \int_0^1 \partial_1 C(u, v) du dv - 2 = 1 \quad (12)$$

where we have used the properties $0 \leq \partial_1 C(u, v) \leq 1$ and $C(1, v) = P(V \leq v) = v$, as can be seen from Eq. (9). Therefore $\tau_2(C) = 1$ if and only if $\partial_1 C(u, v) = P(V \leq v | U = u) = 0, 1$ almost surely. Intuitively, $P(V \leq v | U = u)$ is non-decreasing in $v$ and $P(V \leq 1 | U = u) = 1$, so it has a jump from 0 to 1 at certain value of $v$ for each $u$. Thus $V$ should be a function of $U$.

It has been proved in Darsow *et al.* (1992) that

(a) C is left invertible if and only if for each $v \in I$, $\partial_1 C(u, v) \in \{0,1\}$ for almost all $u \in I$;

(b) C is right invertible if and only if for each $u \in I$, $\partial_2 C(u, v) \in \{0,1\}$ for almost all $v \in I$.

Thus $\tau_2(C) = 1$ if and only if $C \in \mathcal{L}$ is left invertible or $Y$ is a function of $X$. So this defines a new class of true measure of dependence.

Besides the distance forms in Eq. (10), there are also entropy forms for the new measures, see Li (2015a). An example would be

$$R(C) = \int_0^1 \int_0^1 \frac{\partial_1 C(u,v)}{v} \cdot \log\left(\frac{\partial_1 C(u,v)}{v}\right) du dv \quad (13)$$

For the bivariate normal copula with correlation $\rho$, it is easy to calculate

$$\tau_2(C) = \sqrt{\frac{3}{\pi} \arcsin\left(\frac{1+\rho^2}{2}\right) - \frac{1}{2}} \quad (14)$$

Thus $\tau_2(C) = 0$ if $\rho = 0$, and $\tau_2(C) = 1$ if $\rho = \pm 1$. $\tau_2$ is an increasing function of $|\rho|$.

As an example, we consider three singular copulas with supports shown in Figure 1 where $C_1 \in \mathcal{L}$ is left invertible, $C_2 \in \mathcal{R}$ is right invertible, $C_3 = C_1 * C_2 \in \mathcal{L} * \mathcal{R}$ is not left or right invertible. $C_1$ corresponds to the relationship $Y = 1 - X^2$ for $X \in [-1,1]$, $C_2$ corresponds to the relationship $Y = Z^2$ for $Z \in [-1,1]$, and $C_3$ corresponds to the circular relationship $X^2 + Z^2 = 1$ where there is no direct functional relationship between random variables $X$ and $Z$. Straight calculation reveals that

$$\tau_2(C_1) = 1, \tau_2\left({C_1}^T\right) = \frac{1}{2}, I(C_1) = \infty, H(C_1) = 1$$

$$\tau_2(C_2) = \frac{1}{2}, \tau_2\left({C_2}^T\right) = 1, I(C_2) = \infty, H(C_2) = 1 \quad (15)$$

$$\tau_2(C_3) = \frac{1}{2}, \tau_2\left({C_3}^T\right) = \frac{1}{2}, I(C_3) = \infty, H(C_3) = 1$$

Shannon's mutual information produces infinity whenever a copula has singular component, while Hellinger distance gives one whenever a copula is singular. Thus neither of them can differentiate between the three dependence relationships, but the new nonsymmetric dependence measure gives reasonable results. For example, $\tau_2(C_1) = 1$ implies that $Y$ is a function of $X$, $\tau_2\left({C_1}^T\right) = \frac{1}{2}$ implies that $X$ is half dependent on $Y$ as, for each value of $Y$, $X$ has equal probability to be one of two values. Similarly $\tau_2(C_2) = \frac{1}{2}$ implies that $Z$ is half dependent on $Y$, $\tau_2\left({C_2}^T\right) = 1$ implies that $Y$ is a function of $Z$. $\tau_2(C_3) = \tau_2\left({C_3}^T\right) = \frac{1}{2}$ implies that $X$ is half dependent on $Z$ and $Z$ is half dependent on $X$ for a circular relationship.

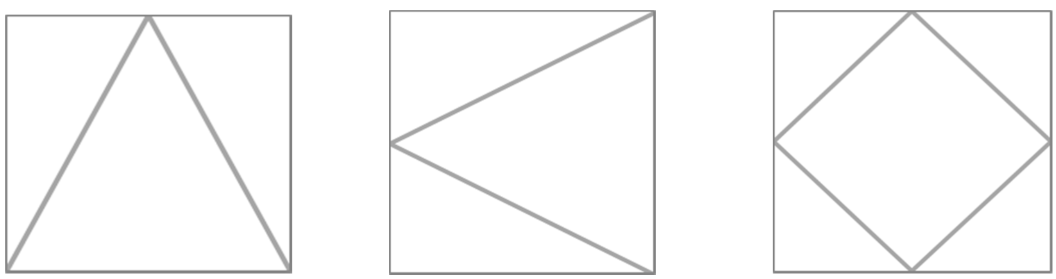

Figure 1. The support of singular copulas $C_1$, $C_2$, $C_3$ on $[0,1] \times [0,1]$.

The new measure can be generalized to multivariate case (Li, 2015b) and also applies to discrete random variables (Li, 2015c). A similar example for the discrete case was given in Li (2015c), also showing the insufficiency of Shannon's mutual information or Hellinger distance in detecting dependence for circular relationship.

One generalization to multivariate case (Li, 2015b) is to consider how much one random variable $Y$ depends on a group of random variables $X_1, X_2, \ldots, X_n$. For example, Equation (11) can be extended as

$$\tau^2(C) = 6\int_0^1 \cdots \int_0^1 \left(P(V \le v | U_1 = u_1, \cdots, U_n = u_n) - P(V \le v)\right)^2 dC(u_1, \cdots, u_n) dv$$

$$= 6\int_0^1 \cdots \int_0^1 \left( \frac{\frac{\partial^n}{\partial u_1 \cdots \partial u_n} C(u_1, \cdots, u_n, v)}{\frac{\partial^n}{\partial u_1 \cdots \partial u_n} C(u_1, \cdots, u_n)} - v \right)^2 \frac{\partial^n}{\partial u_1 \cdots \partial u_n} C(u_1, \cdots, u_n) du_1 \cdots du_n dv$$

$$ (16) $$

where the uniform random variables $U_1 = F_{X_1}(X_1), \cdots, U_n = F_{X_n}(X_n), V = F_Y(Y)$ have the joint distribution $C(u_1, \cdots, u_n, v)$. Again, $\tau(C) = 0$ if and only if $Y$ is independent of the group $X_1, X_2, \dots, X_n$, although $X_1, X_2, \dots, X_n$ may be dependent on each other. Besides, $\tau(C) = 1$ if and only if $Y$ is almost surely a function of $X_1, X_2, \dots, X_n$, which includes the case that $Y$ is a function of a subset of $X_1, X_2, \dots, X_n$.

## 4. Modified Rényi's axioms

Next we review the well-known Rényi's axioms for symmetric measures and show how to modify the conditions for nonsymmetric measures.

Rényi (1959) introduced a set of axioms as the criteria of a symmetric nonparametric measure of dependence $D(X, Y)$ for two random variables $X, Y$ on a common probability space:

a) $D(X, Y)$ is defined for all non-constant random variables $X, Y$;

b) $D(X, Y) = D(Y, X)$;

c) $0 \leq D(X, Y) \leq 1$;

d) $D(X, Y) = 0$ if and only if $X, Y$ are independent;

e) $D(X, Y) = 1$ if either $Y = f(X)$ or $X = g(Y)$ almost surely for some Borel-measurable functions $f, g$;

f) If $f$ and $g$ are Borel-measurable bijections on $\mathbb{R}$, then $D(f(X), g(Y)) = D(X, Y)$;

g) If $X, Y$ are jointly normal with correlation coefficient ρ, then $D(X, Y) = |\rho|$.

Rényi's condition b) specifies symmetry but is not absolutely necessary. Besides, in condition e), functional relationship is only sufficient but not necessary, which means it may take maximum value on non-functional relationship, as we have shown for Shannon's mutual information whose scaled version (Linfoot, 1957) $L(X, Y) = \sqrt{1 - e^{-2I(X,Y)}}$ satisfies all Rényi's axioms.

For nonsymmetric dependence measures, we may modify the conditions as follows (Li, 2015a):

a') $D(X, Y)$ is defined for all continuous random variables $X, Y$;

b') $D(X, Y)$ may not be equal to $D(Y, X)$;

c') $0 \leq D(X, Y) \leq 1$;

d') $D(X,Y) = 0$ if and only if $X, Y$ are independent;

e') $D(X,Y) = 1$ if and only if $Y = f(X)$ almost surely for a Borel-measurable function $f$;

f') If $g$ is a Borel-measurable bijection on $\mathbb{R}$, then $D(g(X),Y) = D(X,Y)$.

g') If $X, Y$ are jointly normal with correlation coefficient ρ, then $D(X,Y)$ is a strictly increasing function of $|\rho|$.

Condition a') restricts the random variables to continuous ones such that the copula between them is uniquely defined (Nelson, 2006). Condition b') specifies that the dependence measure can be nonsymmetric where $D(X,Y)$ measures dependence of $Y$ on $X$ and $D(Y,X)$ measures dependence of $X$ on $Y$. But if a copula is symmetric or $C(u,v) = C(v,u)$, then $D(X,Y) = D(Y,X)$. Conditions c') and d') are the same as Rényi's conditions c) and d) as independence is a symmetric property. Condition e') is more explicit about the nonsymmetric nature of dependence and is stronger as $D(X,Y) = 1$ happens if and only if $Y = f(X)$. As a nonsymmetric measure, condition f') only requires the measure to be invariant under bijective transformations on $X$. It is shown in Li (2015a) that distance-like dependence measures are also invariant under monotonic transformations on $Y$. We relax Rényi's condition g) to g'), since the measures are normally an increasing function of $|\rho|$ as shown in Eq. (14) for the special case $\tau_2$. Note that bivariate normal copula $C$ is symmetric, so $D(C) = D(C^T)$.

These new conditions can be shown to be satisfied by the new dependence measure in Eq. (10). Condition f') can be proved through the generalized DPI for nonsymmetric dependence measures Li (2015a). If $g$ is a bijection, then $X$, $g(X)$, $Y$ and $g(X)$, $X$, $Y$ both form Markov chain. Thus we have both $D(X,Y) \le D(g(X),Y)$ and $D(g(X),Y) \le R(X,Y)$, which leads to $D(X,Y) = D(g(X),Y)$. Thus if a measure satisfies DPI, then it satisfies condition f').

For the multivariate case, the conditions have been extended in Li (2015b).

## 5. Conclusion

To summarize, we have shown that well-known symmetric dependence measures have the issue of not capturing the complete dependence as they take maximum value on a larger set of dependence relationship which includes non-functional relationship. A new type of nonsymmetric dependence measure is introduced which takes minimum value zero exactly on independence and takes maximum value one exactly on complete dependence. This true measure of dependence will be better positioned for detecting hidden relationships including causality. It will be interesting to see how the new measure helps in areas where mutual information is normally used as dependence measure, especially in causal inference.

## References


1. Cover, T. M. and Thomas, J. A. (1991). *Elements of information theory*. New York: John Wiley & Sons.
2. Darsow, W. F., Nguyen, B. and Olsen, E. T. (1992). Copulas and Markov processes. *Illinois J. Math.* **36** 600-642.
3. Dette, H., Siburg, K. F. and Stoimenov, P. A. (2010). A copula-based nonparametric measure of regression dependence. *Scand. J. Stat.* **40** 21-41.
4. Granger, C. W., Maasoumi, E. and Racine, J. (2004). A dependence metric for possibly nonlinear processes. *J. Time Ser. Anal.* **25** 649-669.
5. Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M. and Bhattacharya, J. (2007). Causality detection based on information-theoretic approaches in time series analysis. *Phys. Rep.* **441** 1-46.
6. Jaworski, P., Durante, F., Härdle, W. and Rychlik, T. (Eds.) (2010). *Copula Theory and Its Applications*. Heidelberg: Springer.
7. Kinney, J. B. and Atwal, G. S. (2014). Equitability, mutual information and the maximal information coefficient. *Proc. Nat. Acad. Sci.* **111** 3354-3359.
8. Li, H. (2015a). On Nonsymmetric Nonparametric Measures of Dependence. *arXiv:1502.03850*.
9. Li, H. (2015b). A new class of nonsymmetric multivariate dependence measures. *arXiv:1511.02744*.
10. Li, H. (2015c). Nonsymmetric dependence measures: the discrete case. *arXiv:1512.07945*.
11. Linfoot, E. H. (1957). An informational measure of correlation. *Information and Control* **1** 85-89.
12. Nelson, R. B. (2006). *An introduction to copulas*. New York: Springer.
13. Rényi, A. (1959). On measures of dependence. *Acta Math. Acad. Sci. Hungar.* **10** 441-451.
14. Rényi, A. (1961). On measures of entropy and information. In: *Proceedings of the Fourth Berkley Symposium on Mathematical Statistics and Probability* **1** 547-561.
15. Shannon, C. E. and Weaver, W. (1949). *The mathematical theory of communication*. Urbana: University of Illinois Press.
16. Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8** 229-231.
17. Tsallis, C. (1988). Possible generalization of Boltzmann-Gibbs statistics. *J. Stat. Phys.* **52** 479-487.
18. Trutschnig, W. (2011). On a strong metric on the space of copulas and its induced dependence measure. *J. Math. Anal. Appl.* **384** 690-705.