Generalized Measures of Correlation

Shurong Zheng[†], Ning-Zhong Shi[†] and Zhengjun Zhang[‡]

[†]School of Mathematics and Statistics and KLAS, Northeast Normal University,

Changchun City 130024, P. R. China

[‡]Department of Statistics, University of Wisconsin,

Madison, WI 53706, USA

July 3, 2010

Abstract

Applicability of Pearson's correlation as a measure of explained variance is by now well understood. One of its limitations is that it does not account for asymmetry in explained variance. Aiming to develop broad applicable correlation measures, we propose a pair of generalized measures of correlation (GMC) which deal with asymmetries in explained variances, and linear or nonlinear relations between random variables. We present examples under which the paired measures are identical, and they become a symmetric correlation measure which is the same as the squared Pearson's correlation coefficient. As a result, Pearson's correlation is a special case of GMC. Theoretical properties of GMC show that GMC can be applicable in numerous applications. In statistical inferences, the joint asymptotics of the kernel based estimators for GMC are derived and are used to test whether or not two random variables are symmetric in explaining variances. The testing results give important guidance in practical model selection problems. The efficiency of the test statistics is illustrated in simulation examples. In real data analysis, we present an important application of GMC in explained variances and market movements among three important economic and financial monetary indicators.

Key words and phrases: Linear dependence, nonlinear dependence, asymmetric correlation, nonparametric estimation, economic study.

1 Introduction

In almost all statistical inference problems, dealing with how random variables depend on each other plays a fundamental role in model selections. In the literature, since its introduction, Pearson's correlation coefficient has been the most dominant dependence measure used in numerous applications. It mainly depicts a symmetric and linear relation between two variables. Its theoretical properties have been thoroughly studied. Rodgers and Nicewander (1988) presented thirteen ways to look at the correlation coefficients. For many applications, it leads to meaningful and interesting interpretations of variables under study. However, it may also give misleading results in many applications. This phenomenon has been witnessed in many published papers, for example, O'Grady (1982), Ozer (1985), Drouet-Mari and Kotz (2001), and Zhang (2008), amongst many others. Recently, Zhang, Qi, and Ma (2010) show that the sample based Pearson's correlation coefficient is asymptotically independent of the quotient correlation coefficient, which is a very important property as it shows that these two correlation coefficients measure completely different dependencies between two random variables. Certainly, Pearson's correlation coefficient has its limitations in measuring variable dependencies. To overcome its limitations, various dependence measures have been proposed in the literature. We shall not detail them in the present work. We refer readers to Joe (1996) and Drouet-Mari and Kotz (2001) which are excellent books summarizing various dependence measures.

For Pearson's correlation coefficient, one of its limitations is that it does not account for asymmetry in explained variances which are often innate among nonlinearly dependent random variables. As a result, measures dealing with asymmetries are needed. In fact, studying the asymmetric dependent characteristics of random variables has drawn more and more attentions, especially in the studies of stock returns such as Hong, Tu and Zhou (2006), Zhang and Shinki (2006), and references therein. However, theoretical foundations of asymmetry in explained variances do not exist and are yet to be developed.

This paper is intended to introduce effective and broadly applicable statistical tools for dealing with asymmetry and nonlinear correlations between random variables. For simplicity of illustration, we regard 'linear' or 'symmetric' as a special case of 'nonlinear' or 'asymmetric'. In the case of 'linear and symmetric', Pearson's correlation coefficient is an extremely important and widely used analytical tool in statistical data analysis. New dependence measures that comprise Pearson's correlation coefficient as a special case should be of the greatest interest to practitioners. We aim to develop such dependence measures. In Section 2.1, we use a well known variance decomposition formula to introduce our proposed new measures of correlation: the generalized measures of correlation (GMC). Theoretical properties of GMC are illustrated in Section 2.2. One can see that our proposed GMC has various connections to Pearson's correlation coefficient, and especially they are identical to the squared Pearson's correlation coefficient when two random variables are related in a linear equation. A special case is that two random variables follow a bivariate normal distribution. More importantly, GMCs are nonzero while Pearson's correlation coefficient may have a zero value when two random variables are nonlinearly dependent. In addition, GMCs also have monotonic dependence properties in explained variances. One can also see that our proposed GMC may be used as an alternative statistical tool in Granger causality inference. For this purpose, we introduce two new measures: Auto generalized measures of correlation (AGMC) and Granger causality generalized measures of correlation (GcGMC).

The rest of the paper is structured as follows. In Section 3, we present a nonparametric method in computing our proposed GMC. The joint asymptotics of two GMC estimators are derived and they are used to test whether two explained variances are identical or not. Starting from theoretical foundations, we will analyze examples covering a wide range of dependency between random variables in Section 4. Particularly, we study three types of bivariate t random variables and derive their corresponding GMCs. We also calculate GMCs for a three sectional extreme value copula which shows an asymmetric dependence and extreme dependence between two underlying random variables. Numerical illustrations of GMC are presented in Section 5. One can see that GMC can be very useful dependence measures, especially when explained variance is concerned. In Section 6, we present real data analysis through three important economic variables: the exchange rate of Japanese Yen against US dollar, US federal funds rate, and Japan deposit rate. They are indicators of both countries' economy status: for example whether they are healthy or not. People have hoped that the comparisons may help reveal similarities and find answers (even solutions) to an economic recovery from the current international financial crisis. From a market perspective, plotting these variables shows no similarity, linear relationship, or co-monotone relationship. However, our peculiar GMC shall display economic changes between these two countries. Section 7 discusses potential extensions of the present paper and limitations of our proposed measures. Technical derivations are presented in Section 8.

2 Generalized measures of correlation: Definitions and Properties

2.1 Generalized measures of correlation

In computing coefficient of determination in a linear regression model, the total variation in response variable is partitioned into two component sums of squares, i.e. explained variation due to regression and unexplained variation. Here, we shall introduce our generalized measures of correlation based on a well known variance decomposition formula:

$$Var(X) = Var\big(E(X|Y)\big) + E\big(Var(X|Y)\big) \tag{1}$$

whenever $E(Y^2) < \infty$ and $E(X^2) < \infty$. Note that $E\big(Var(X|Y)\big)$ is the expected conditional variance of X given Y, and hence $E\big(Var(X|Y)\big)/Var(X)$ can certainly be interpreted as the explained variance of X by Y. We have

$$\frac{E\big(Var(X|Y)\big)}{Var(X)} = 1 - \frac{Var\big(E(X|Y)\big)}{Var(X)} = 1 - \frac{E[\{X - E(X|Y)\}^2]}{Var(X)}.$$

Similarly we can define the explained variance of Y given X. Therefore, it is natural to introduce a pair of generalized measures of correlation (GMC) as

$$\big\{GMC(Y|X),\ GMC(X|Y)\big\} = \Big\{1 - \frac{E[\{Y - E(Y|X)\}^2]}{Var(Y)},\ 1 - \frac{E[\{X - E(X|Y)\}^2]}{Var(X)}\Big\}. \tag{2}$$

This pair of GMC posseses many good properties which will be illustrated in detail in the following section. One of them is that the two measures are identical when (X, Y) is a bivariate normal random vector. The GMC can depict the nonlinear or asymmetric relation between two variables. They are true measures for explained variances.

2.2 Properties of generalized measures of correlation

We have the following proposition which shows that the GMC can measure the nonlinear and asymmetric relation between two variables. Proofs of propositions are postponed to Appendix section.

Proposition 2.1 *Suppose both X and Y have finite second moments. Then*

(i) *GMC is an indicator lying between zero and one, that is,*

$$0 \leq GMC(Y|X),\ \ GMC(X|Y) \leq 1,$$

and if X and Y are independent, then $GMC(Y|X) = 0,\ GMC(X|Y) = 0$.

(ii) *The relation of GMC and Pearson's correlation coefficient ρ_{XY} satisfies:*

- *If $\rho_{XY} = \pm 1$, then $GMC(Y|X) = 1$ and $GMC(X|Y) = 1$.*
- *If $\rho_{XY} \neq 0$, then $GMC(X|Y) \neq 0$ and $GMC(Y|X) \neq 0$.*
- *If $GMC(Y|X) = 0$ and/or $GMC(X|Y) = 0$, then $\rho_{XY} = 0$.*

(iii) *Suppose $Y = g(X) + \epsilon$, X and ϵ are independent, and both $g(X)$ and ϵ have finite second moments, where $g(\cdot)$ is a linear or nonlinear measurable function. Then*

$$GMC(Y|X) = \frac{Var\big(g(X)\big)}{Var\big(g(X)\big) + Var(\epsilon)}.$$

Particularly, if $g(x) = ax + b$ for $a \neq 0$ and b being constants, we have

$$GMC(Y|X) = \rho_{XY}^2.$$

For the extreme values of GMC, we have

$$GMC(Y|X) = 1 \Longleftrightarrow Y = g(X) \quad a.s.$$

Furthermore, If g is a one to one measurable function, then $GMC(Y|X) = GMC(X|Y) = 1$; If g is not one to one, then $GMC(Y|X) = 1 > GMC(X|Y) \geq 0$.

(iv) *Suppose $Y_1 = g_1(X) + \epsilon_1$ and $Y_2 = g_2(X) + \epsilon_2$, where ϵ_1 and ϵ_2 are independent of X, $g_1(\cdot)$ and $g_2(\cdot)$ are linear or nonlinear measurable functions. If either 1) $Var\big(g_1(X)\big) = Var\big(g_2(X)\big),\ Var(\epsilon_1) < Var(\epsilon_2)$; or 2) $Var\big(g_1(X)\big) > Var\big(g_2(X)\big),\ Var(\epsilon_1) = Var(\epsilon_2)$, we have*

$$GMC(Y_1|X) > GMC(Y_2|X).$$

(v) *If $Var(Y_1) = Var(Y_2)$ and $\inf_f E[\{Y_1 - f(X)\}^2] = \inf_g E[\{Y_2 - g(X)\}^2]$, where f, g are measurable functions, then*

$$GMC(Y_1|X) = GMC(Y_2|X).$$

If $Var(Y_1) = Var(Y_2)$ and $\inf_f E[\{Y_1 - f(X)\}^2] < \inf_g E[\{Y_2 - g(X)\}^2]$, then we have

$$GMC(Y_1|X) > GMC(Y_2|X).$$

Example 2.1 *Considering a special example $Y = X^2, X \sim N(0,1)$ where Y is a nonlinear measurable function in Proposition 2.1 (iii), we have generalized correlation coefficients $GMC(Y|X) = 1, GMC(X|Y) = 0, GMC(Y|X) \neq GMC(X|Y)$, but Pearson's correlation coefficient $\rho_{XY} = 0$.*

Remark 1 *In Proposition 2.1, it is clear in (i), (ii) and (iii) that GMC characterizes nonlinear or asymmetric relation between two variables, where 'linear' or 'symmetric' is considered as a special case of 'nonlinear' or 'asymmetric' respectively. If Y is perfectly nonlinearly dependent on X, $GMC(Y|X)$ is 1; and if X and Y are independent, the GMCs are 0. In (ii), when X and Y have a nonzero Pearson's linear correlation coefficient, the GMCs of X and Y are always greater than zero, but not reversely as shown in Example 2.1. On the other hand, as long as one of GMCs is zero, Pearson's correlation must be zero. These properties show that GMC has more applicabilities than Pearson's correlation coefficient has. They are strong indications of the generality of GMC as measure of dependence considering that Pearson's correlation only measures linear and symmetric relation. In (iv) if the linear or nonlinear relation of Y_1 on X is stronger than that of Y_2 on X, then $GMC(Y_1|X)$ is larger than $GMC(Y_2|X)$. It shows monotonicity of GMC, which is an important property in defining a correlation measure and in defining a prediction criterion for model/varaible selections. In (v) if X has the stronger ability to predict Y_1 than to predict Y_2, then GMC correlation $Y_1|X$ is larger than GMC correlation $Y_2|X$, that is, $GMC(Y_1|X) > GMC(Y_2|X)$.*

The following proposition shows that the squared Pearson's correlation coefficient is identical to the GMC under the bivariate normal distribution.

Proposition 2.2 *For the bivariate normal distribution, ρ_{XY} and $\big(GMC(Y|X), GMC(X|Y)\big)$ are equivalent in depicting the relation of X and Y, i.e., $GMC(Y|X) = GMC(X|Y) = \rho_{XY}^2$.*

The following proposition calculates GMCs when marginal distributions are uniform on [0,1]. The proof of the proposition is straightforward.

Proposition 2.3 *Suppose that X and Y have continuous distribution functions $F_X(x)$ and $F_Y(y)$ respectively. Then*

$$GMC\big(F_Y(Y)|X\big) = 12E\big(\{E(F_Y(Y)|X)\}^2\big)-3,\ GMC\big(F_X(X)|Y\big) = 12E\big(\{E(F_X(X)|Y)\}^2\big)-3. \tag{3}$$

Formulas in (3) can be compared with Spearman's correlation

$$\rho_S(X,Y) = corr\big(F_X(X), Y_Y(Y)\big) = 12E\big(F_X(X)F_Y(Y)\big) - 3,$$

which does not account for asymmetry in explained variances. Examples with uniform marginals will be illustrated in Section 4.

We argue that the three propositions above clearly show that the GMC is a true measure for explained variances, and for linear or nonlinear relations between two random variables. We note that the calculation of our proposed GMC involves computing the variance of conditional expectation and the expectation of conditional variance, which may be a difficult task in deriving explicit GMC formulas. In Section 4, we shall derive explicit forms of GMC in several joint distributional models.

2.3 Generalized measures of correlation in time series

In time series study, auto-correlation function is an important concept. Our GMC can naturally be extended to time series models. Suppose that $\{X_t, Y_t\},\ t > 0$ is a bivariate time series. We define auto generalized measures of correlation (AGMC) as:

$$AGMC_k(X_t) = GMC(X_t|X_{t-k}),\ AGMC_k(Y_t|X_t) = GMC(Y_t|X_{t-k}),\ k > 0. \tag{4}$$

Granger causality (Granger 1969) has been widely used in economics since the 1960s. It is a powerful statistical concept of causality that is based on prediction. It is normally tested in a bivariate linear autoregressive model of two variables X_t and Y_t. For simplicity, we assume an order one bivariate linear autoregressive model. We say Y_t Granger-causes X_t if

$$E[\{X_t - E(X_t|X_{t-1})\}^2] > E[\{X_t - E(X_t|X_{t-1}, Y_{t-1})\}^2], \tag{5}$$

i.e X_t can be better predicted using the histories of both X_t and Y_t than using the history of X_t alone. Similarly we say X_t Granger-causes Y_t if

$$E[\{Y_t - E(Y_t|Y_{t-1})\}^2] > E[\{Y_t - E(Y_t|Y_{t-1}, X_{t-1})\}^2]. \tag{6}$$

Using the fact $E\big(Var(X_t|X_{t-1})\big) = E[\{X_t - E(X_t|X_{t-1})\}^2]$ and

$$E[\{E(X_t|X_{t-1})-E(X_t|X_{t-1}, Y_{t-1})\}^2] = E[\{X_t-E(X_t|X_{t-1})\}^2]-E[\{X_t-E(X_t|X_{t-1}, Y_{t-1})\}^2],$$

one can see that (5) is equivalent to

$$1 - \frac{E[\{X_t - E(X_t|X_{t-1}, Y_{t-1})\}^2]}{E\big(Var(X_t|X_{t-1})\big)} > 0. \tag{7}$$

Similarly, (6) is equivalent to

$$1 - \frac{E[\{Y_t - E(Y_t|Y_{t-1}, X_{t-1})\}^2]}{E\big(Var(Y_t|Y_{t-1})\big)} > 0. \tag{8}$$

When both (5) and (6) are true, we have a *feedback system*. (7) and (8) can be extended to a more general form, and we introduce our Granger causality GMC as follows.

Definition 2.4 *Suppose that $\{X_t, Y_t\},\ t > 0$ is a bivariate stationary time series. Define Granger causality generalized measures of correlation (GcGMC) as:*

$$GcGMC(X_t|\mathcal{F}_{t-1}) = 1 - \frac{E[\{X_t - E(X_t|X_{t-1}, X_{t-2}, \dots, Y_{t-1}, Y_{t-2}, \dots)\}^2]}{E\big(Var(X_t|X_{t-1}, X_{t-2}, \dots)\big)}, \tag{9}$$

$$GcGMC(Y_t|\mathcal{F}_{t-1}) = 1 - \frac{E[\{Y_t - E(Y_t|Y_{t-1}, Y_{t-2}, \dots, X_{t-1}, X_{t-2}, \dots)\}^2]}{E\big(Var(Y_t|Y_{t-1}, Y_{t-2}, \dots)\big)}, \tag{10}$$

where $\mathcal{F}_{t-1} = \sigma(X_{t-1}, X_{t-2}, \dots, Y_{t-1}, Y_{t-2}, \dots)$.

- *If $GcGMC(X_t|\mathcal{F}_{t-1}) > 0$, we say Y Granger causes X.*
- *If $GcGMC(Y_t|\mathcal{F}_{t-1}) > 0$, we say X Granger causes Y.*
- *If $GcGMC(X_t|\mathcal{F}_{t-1}) > 0$ and $GcGMC(Y_t|\mathcal{F}_{t-1}) > 0$, we say that we have a* feedback system.
- *If $GcGMC(X_t|\mathcal{F}_{t-1}) > GcGMC(Y_t|\mathcal{F}_{t-1})$, we say that X is more influential than Y.*
- *IF $GcGMC(Y_t|\mathcal{F}_{t-1}) > GcGMC(X_t|\mathcal{F}_{t-1})$, we say that Y is more influential than X.*

It is easy to show that $0 \le GcGMC(X_t|\mathcal{F}_{t-1}) \le 1$ and $0 \le GcGMC(Y_t|\mathcal{F}_{t-1}) \le 1$. Other theoretical properties of GcGMC can be derived along the line of GMC. This work will make a single paper too overloaded, and we shall put the study of GcGMC in a separate project. We present the following proposition which relates GcGMC to Grainger's causality .

Proposition 2.5 *If $E\big(Var(X_t|X_{t-1})\big) = E\big(Var(Y_t|Y_{t-1})\big)$ and the strength of that X_t Granger causes Y_t is stronger than the strength of that Y_t Granger causes X_t, i.e.*

$$E[\{E(Y_t|Y_{t-1}) - E(Y_t|Y_{t-1}, X_{t-1})\}^2] > E[\{E(X_t|X_{t-1}) - E(X_t|X_{t-1}, Y_{t-1})\}^2],$$

then

$$GcGMC(Y_t|Y_{t-1}, X_{t-1}) > GcGMC(X_t|X_{t-1}, Y_{t-1}).$$

3 Nonparametric estimators of GMC

The GMC in (2) involves evaluations of conditional means and variances, which is not as easy as computing Pearson's correlation coefficient in practice. In the literature, there have been quite a few developments in estimating conditional variances such as Fan and Yao (1998) and Hensen (2009), amongst others. We propose to use nonparametric kernel based methods to estimate GMC. The construction of our estimators for each conditional variance in GMC is similar to the existing methods in the literature. The main task here is to establish the joint asymptotics of the estimators.

Throughout Section 3 and proofs of Lemma 3.1, Theorems 3.2 and 3.3, we denote $f^X(x)$ and $f^Y(y)$ as the density functions of X and Y respectively. $\{(X_i, Y_i),\ i = 1, \ldots, n\}$ is a random sample of (X, Y). Denote $R^1 = (-\infty, +\infty)$, $s_x = \inf\{x:\ f^X(x) > 0,\ x \in R^1\}$, $S^x = \sup\{x:\ f^X(x) > 0,\ x \in R^1\}$, $s_y = \inf\{y:\ f^Y(y) > 0,\ y \in R^1\}$, $S^y = \sup\{y:\ f^Y(y) > 0,\ y \in R^1\}$. For notational convenience, we drop the limits in all integrals. The lower and upper limits are s_x and S^x, respectively, in all integrals with respect to dx, and the lower and upper limits are s_y and S^y, respectively, in all integrals with respect to dy.

3.1 The estimators

First, GMC can be expressed as

$$GMC(Y|X) = 1 - \frac{E[\{Y - E(Y|X)\}^2]}{Var(Y)} = \frac{\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)} dx - \mu_Y^2}{\sigma_Y^2}$$

and

$$GMC(X|Y) = \frac{\int \frac{(\phi^{X|Y}(y))^2}{f^Y(y)} dy - \mu_X^2}{\sigma_X^2}$$

where $\mu_X = EX$, $\mu_Y = EY$, $\sigma_X^2 = Var(X)$, $\sigma_Y^2 = Var(Y)$, $\phi^{Y|X}(x) = \int yf(x,y)dy$, $\phi^{X|Y}(y) = \int xf(x,y)dx$, $f(x,y)$ is the joint density of X and Y. Let the kernel densities of (X,Y) be

$$\widehat{f}(x,y) = \frac{1}{nh^2} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) K\left(\frac{y - Y_i}{h}\right),$$

$$f_n^X(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right) \quad \text{and} \quad f_n^Y(y) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{y - Y_i}{h}\right)$$

where $K(.)$ is a kernel function and h is the bandwidth. Then the Nadaraya-Watson estimator is

$$\widehat{E}(Y|X = x) = \frac{\frac{1}{nh} \sum\limits_{i=1}^{n} Y_i K\left((x - X_i)/h\right)}{\frac{1}{nh} \sum\limits_{i=1}^{n} K\left((x - X_i)/h\right)} = \frac{\phi_n^{Y|X}(x)}{f_n^X(x)}$$

where $\phi_n^{Y|X}(x) = \frac{1}{nh} \sum\limits_{i=1}^{n} Y_i K\left(\frac{x - X_i}{h}\right)$. Similarly, we obtain

$$\widehat{E}Y = \int y f_n^Y(y) dy = \bar{Y} + h \cdot E_K^1,$$

$$\widehat{E}Y^2 = \int y^2 f_n^Y(y) dy = \frac{1}{n} \sum_{i=1}^{n} Y_i^2 + h^2 E_K^2 + 2h\bar{Y}E_K^1, \quad \widehat{Var}(Y) = S_Y^2 + h^2 Var_K$$

and

$$\widehat{E}[\{\widehat{E}(Y|X)\}^2] = \int [\widehat{E}(Y|X = x)]^2 f_n^X(x) dx = \int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)} dx$$

where $E_K^i = \int z^i K(z) dz$, $\bar{Y}$ and S_Y^2 are the sample mean and sample variance of $Y_1, \cdots, Y_n$, and $Var_K = \int z^2 K(z) dz - (\int zK(z)dz)^2$. Because $\widehat{E}[\{Y - \widehat{E}(Y|X)\}^2] = \widehat{E}Y^2 - \widehat{E}[\{\widehat{E}(Y|X)\}^2]$, then we have

$$\begin{aligned}
\frac{\widehat{E}[\{Y - \widehat{E}(Y|X)\}^2]}{\widehat{Var}(Y)} &= \frac{\frac{1}{n} \sum\limits_{i=1}^{n} Y_i^2 + h^2 \cdot E_K^2 + 2h \cdot \bar{Y} \cdot E_K^1 - \widehat{E}[\{\widehat{E}(Y|X)\}^2]}{S_Y^2 + h^2 \cdot Var_K} \\
&= 1 + \frac{\bar{Y}^2 - h^2 \cdot Var_K + h^2 \cdot E_K^2 + 2h \cdot \bar{Y} \cdot E_K^1 - \widehat{E}[\{\widehat{E}(Y|X)\}^2]}{S_Y^2 + h^2 \cdot Var_K}
\end{aligned}$$

$$= 1 + \frac{\bar{Y}^2 + h^2 \cdot (E_K^1)^2 + 2h \cdot \bar{Y} \cdot E_K^1 - \widehat{E}[\{\widehat{E}(Y|X)\}^2]}{S_Y^2 + h^2 \cdot Var_K}$$

$$= 1 + \frac{(\bar{Y} + h \cdot E_K^1)^2 - \widehat{E}[\{\widehat{E}(Y|X)\}^2]}{S_Y^2 + h^2 \cdot Var_K} = 1 + \frac{(\bar{Y} + h \cdot E_K^1)^2 - \int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)} dx}{S_Y^2 + h^2 Var_K}.$$

Then the kernel estimators of GMC are

$$\widetilde{GMC}(Y|X) = \frac{\int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)} dx - (\bar{Y} + hE_K^1)^2}{S_Y^2 + h^2 Var_K}$$

and

$$\widetilde{GMC}(X|Y) = \frac{\int \frac{(\phi_n^{X|Y}(y))^2}{f_n^Y(y)} dy - (\bar{X} + hE_K^1)^2}{S_X^2 + h^2 Var_K}$$

where $\phi_n^{X|Y}(y)$ is defined similarly to $\phi_n^{Y|X}(x)$.

3.1.1 Choice of bandwidth h

We shall use cross-validation method to choose h, which is a widely adopted procedure in the literature. Because $E(Y|X) = \min_{l(X)}(Y - l(X))^2$ and $E(X|Y) = \min_{l(Y)}(X - l(Y))^2$, we choose the optimal h as

$$h_{optimal} = \text{argmin}_{h>0} \left[\frac{\omega_1}{n} \sum_{k=1}^{n} (Y_k - \widehat{E}_h^{-k}(Y|X = X_k))^2 + \frac{\omega_2}{n} \sum_{k=1}^{n} (X_k - \widehat{E}_h^{-k}(X|Y = Y_k))^2 \right]$$

where ω_1 and ω_2 are weights, and $(\widehat{E}_h^{-k}(Y|X = x), \widehat{E}_h^{-k}(X|Y = y))$ are Nadaraya-Watson estimators computed based on data (X_i, Y_i), $i = 1, \cdots, k-1, k+1, \cdots, n$. Here ω_1 and ω_2 can be chosen as the inverse of the square root of sample variances of $Y_1, \cdots, Y_n$ and $X_1, \cdots, X_n$, that is, $\omega_1 = 1/S_Y$ and $\omega_2 = 1/S_X$.

3.2 Asymptotics of GMC estimators

Lemma 3.1 *Under Assumptions (a)-(c) in Appendix, we have*

$$\int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)} dx - \int \frac{(\phi^{Y|X}(x))^2}{f^X(x)} dx = T_n^{Y|X} + o_p(1/\sqrt{n}) + O_p(h^2) + O_p(\frac{1}{nh}) \tag{11}$$

and

$$\int \frac{(\phi_n^{X|Y}(y))^2}{f_n^Y(y)} dy - \int \frac{(\phi^{X|Y}(y))^2}{f^Y(y)} dy = T_n^{X|Y} + o_p(1/\sqrt{n}) + O_p(h^2) + O_p(\frac{1}{nh}) \tag{12}$$

where

$$T_n^{Y|X} = \int \frac{2(\phi_n^{Y|X}(x) - E\phi_n^{Y|X}(x))\phi^{Y|X}(x)}{f^X(x)}dx - \int \frac{(f_n^X(x) - Ef_n^X(x)) \cdot (\phi^{Y|X}(x))^2}{(f^X(x))^2}dx$$

and

$$T_n^{X|Y} = \int \frac{2(\phi_n^{X|Y}(y) - E\phi_n^{X|Y}(y))\phi^{X|Y}(y)}{f^Y(y)}dy - \int \frac{(f_n^Y(y) - Ef_n^Y(y)) \cdot (\phi^{X|Y}(y))^2}{(f^Y(y))^2}dy.$$

Theorem 3.2 *Let Assumptions (a)-(c) in Appendix be fulfilled. If* $nh^2 \to \infty$ *and* $nh^4 \to 0$, *then we have*

$$\sqrt{n}\begin{pmatrix} \widetilde{GMC}(Y|X) - GMC(Y|X) - (\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2)\left(\frac{1}{\sigma_Y^2 + h^2 Var_K} - \frac{1}{\sigma_Y^2}\right) \\ \widetilde{GMC}(X|Y) - GMC(X|Y) - (\int \frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy - \mu_X^2)\left(\frac{1}{\sigma_X^2 + h^2 Var_K} - \frac{1}{\sigma_X^2}\right) \end{pmatrix}$$

$$\Longrightarrow N(\mathbf{0}_{2\times 1}, A^T \Sigma A)$$

where $Var_K = \int z^2 K(z)dz - (\int zK(z))^2$,
$$\Sigma = Cov\left(\frac{\int\left(2Y_i - \frac{\phi^{Y|X}(x)}{f^X(x)}\right)\frac{\phi^{Y|X}(x)}{f^X(x)}\frac{1}{h}K\left(\frac{x-X_i}{h}\right)dx}{\sigma_Y^2}, \frac{Y_i}{\sigma_Y}, Y_i^2, \frac{\int\left(2X_i - \frac{\phi^{X|Y}(y)}{f^Y(y)}\right)\frac{\phi^{X|Y}(y)}{f^Y(y)}\frac{1}{h}K\left(\frac{y-Y_i}{h}\right)dy}{\sigma_X^2}, \frac{X_i}{\sigma_X}, X_i^2\right)$$
and

$$A = \begin{pmatrix} 1 & 0 \\ -\frac{2\mu_Y}{\sigma_Y} + \frac{2\left(\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2\right)\sigma_Y \mu_Y}{(\sigma_Y^2 + h^2 Var_K)^2} & 0 \\ -\frac{\left(\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2\right)}{(\sigma_Y^2 + h^2 Var_K)^2} & 0 \\ 0 & 1 \\ 0 & -\frac{2\mu_X}{\sigma_X} + \frac{2\left(\int \frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy - \mu_X^2\right)\sigma_X \mu_X}{(\sigma_X^2 + h^2 Var_K)^2} \\ 0 & -\frac{\left(\int \frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy - \mu_X^2\right)}{(\sigma_X^2 + h^2 Var_K)^2} \end{pmatrix}.$$

With the established joint asymptotics, we can make large sample inferences on explained variances. In the literature, Hotelling (1953) stated that the best present-day usage in dealing with correlation coefficients is Fisher's Z-transformation test of linear (in)dependence of two random variables. In our context, we argue that testing the equality of the explained variances is fundamentally important in model selection, model building and statistical inferences. We have the following testing problem:

$$H_0 : GMC(Y|X) = GMC(X|Y) \quad v.s. \quad H_1 : GMC(Y|X) \neq GMC(X|Y). \tag{13}$$

Theorem 3.3 *Let Assumptions (a)-(c) in Appendix be fulfilled. If $nh^2 \to \infty$ and $nh^4 \to 0$, then the test statistic for the testing problem (13) has the following asymptotic distribution under H_0*

$$\sqrt{n}(\widetilde{GMC}(Y|X) - \widetilde{GMC}(X|Y) - C_0) \to N(0, (1,-1)A^T\Sigma A(1,-1)^T)$$

where

$$\begin{aligned} C_0 &= \left(\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2\right)\left(\frac{1}{\sigma_Y^2 + h^2 Var_K} - \frac{1}{\sigma_Y^2}\right) \\ &\quad - \left(\int \frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy - \mu_X^2\right)\left(\frac{1}{\sigma_X^2 + h^2 Var_K} - \frac{1}{\sigma_X^2}\right) \end{aligned}$$

and the matrices A and Σ are the same as those in Theorem 3.2.

Proof of Theorem 3.3 is easily obtained by Theorem 3.2 and the delta method.

We note that when Theorems 3.2-3.3 are used to make statistical inference for $GMC(X|Y)$ and $GMC(Y|X)$, the unknown $\mu_X, \mu_Y, \sigma_X, \sigma_Y$, $\phi^{X|Y}, \phi^{X|Y}$, f^X, f^Y and Σ can be replaced by their consistent estimators. In Section 5, we use examples to demonstrate sample performances of the established theoretical results.

4 Derivations of GMCs in several joint distributions

From the previous section, we see that the GMCs between two bivariate normal random variables are identical. In the literature, there exist many parametric families of bivariate distribution functions. Some of them posses the property of having identical GMCs, while some of them do not. It will be very useful if we can present GMCs for each known family as GMCs are important population characteristics. It may be too ambitious a task to include every bivariate distribution in a single paper. We choose four families of bivariate distributions to illustrate the derivations of GMC. Two families posses the property of having identical GMCs, while the other two families do not. The purpose of our mathematical derivation of GMCs is to provide some guidance in application and to show what we can get. One can see that for some families, we can get explicit formulas for GMCs, while for other different families, we can not. Also the derivations for cases of non-identical GMCs are much more complicated than the derivations for cases of identical GMCs in our chosen examples.

Beyond bivariate normal distributions, the bivariate t distributions are also widely used in various applications. In the following proposition, we illustrate three types of bivariate t distributions with two having identical GMCs and one having different GMCs.

Proposition 4.1 *Suppose that $\{(X_i,\ Y_i),\ i = 1, \dots, m\}$ is an independent sequence of bivariate normal vectors. Define*

$$\bar{X} = \frac{1}{m}\sum_{i=1}^{m} X_i, \ \ \bar{Y} = \frac{1}{m}\sum_{i=1}^{m} Y_i, \ \ S_1^2 = \frac{1}{m-1}\sum_{i=1}^{m}(x_i - \bar{X})^2, \ \ S_2^2 = \frac{1}{m-1}\sum_{i=1}^{m}(Y_i - \bar{Y})^2.$$

Case 1: *Suppose $Var(X) = \sigma_1^2$, $Var(Y) = \sigma_2^2$, $Cov(X, Y) = \sigma_1\sigma_2\rho$ with $|\rho| < 1$. Define*

$$T_1 = \frac{\sqrt{m}\cdot\bar{X}}{S_1}, \quad T_2 = \frac{\sqrt{m}\cdot\bar{Y}}{S_2}.$$

Then $(T_1,\ T_2)$ follows a bivariate t distribution with degrees of freedom $df = m - 1$ (Siddiqui (1967)), and we have $GMC(T_1|T_2) = GMC(T_2|T_1)$.

Case 2: *Suppose $\sigma_1 = \sigma_2$. Define*

$$T_1 = \frac{\sqrt{m}\cdot\bar{X}}{S}, \quad T_2 = \frac{\sqrt{m}\cdot\bar{Y}}{S}.$$

where $S^2 = \frac{(m-1)(S_1^2+S_2^2)}{2m-1}$. Then $(T_1,\ T_2)$ follows a bivariate t distribution with degrees of freedom $df = m - 1$ (Patil and Liao (1970)), and we have $GMC(T_1|T_2) = GMC(T_2|T_1)$.

Case 3: *Let $X_i, i = 1, \cdots, m_1 + m_2 \overset{i.i.d.}{\sim} N(0, \sigma^2)$. Let*

$$\bar{X}_1 = \frac{1}{m_1}\sum_{i=1}^{m_1} X_i, \ \ \bar{X}_* = \frac{1}{m_2}\sum_{i=m_1+1}^{m_1+m_2} X_i$$

$$S_1^2 = \frac{1}{m_1-1}\sum_{i=1}^{m_1}(X_i - \bar{X}_1)^2, \ \ S_*^2 = \frac{1}{m_2-1}\sum_{i=m_1+1}^{m_1+m_2}(X_i - \bar{X}_*)^2,$$

$$\bar{X}_2 = \frac{1}{m_1+m_2}(m_1\bar{X}_1 + m_2\bar{X}_*), \quad S_2^2 = \frac{(m_1-1)S_1^2 + (m_2-1)S_*^2}{m_1+m_2-2},$$

and

$$T_1 = \frac{\sqrt{m_1}\cdot\bar{X}_1}{S_1}, \quad T_2 = \frac{\sqrt{m_1+m_2}\cdot\bar{X}_2}{S_2}.$$

Then $(T_1,\ T_2)$ follows a bivariate t distribution with degrees of freedom ($df1 = m_1 - 1,\ df2 = m_1 + m_2 - 2$), respectively (Bulgren et al. (1974)). When $m_1 > 2,\ m_2 \geq 2$, $(T_1,\ T_2)$ is a bivariate t distributed random vector, and we have $GMC(T_1|T_2) \neq GMC(T_2|T_1)$.

Recently, the study of copula functions has been becoming a major phenomenon in constructing joint distribution functions and modeling real data. In the literature, the bivariate Gumbel-Hougaard copula is widely used in many applications, especially in finance and in insurance. It is easy to show that the GMCs of a pair of random variables following a bivariate Gumbel-Hougaard copula are identical. In an effort to model multivariate extremal dependence, Zhang (2009) introduced a three sectional copula which partitions the probability space into three parts. We give a brief summary of the three sectional copula here. Suppose that X and Y are two loss random variables. Among the three parts, one part is related to computing the probability that the loss of Y is a times smaller the loss of X, one part is related to computing the probability that the loss of Y is b times larger than the loss of X, and the third part is related to computing the probability of the ratio of the loss of Y and the loss of X is between a and b with $a < b$. Zhang (2009) demonstrated that the three sectional copula performs as good as the Gumbel-Hougaard copula in modeling bivariate extreme dependence. However, the three sectional copula is able to account for either symmetry and asymmetry in explained variances by varying parameter values. In this paper, we further extend the three sectional copula to a model which gives a larger difference between the two GMCs with a price of adding a new parameter.

Suppose that U_1 and U_2 are independent uniform random variables on [0,1]. Define

$$\xi_1 = (-1/\log(U_1))^{1/\beta}; \xi_2 = (-1/\log(U_2))^{1/\beta};$$

$$\eta_1 = \max((1-\alpha_1)\xi_1,\ \alpha_1\xi_2);\ \eta_2 = \max(\alpha_2\xi_1,\ (1-\alpha_2)\xi_2);$$

and

$$X = \exp\left(-\frac{\alpha_1^\beta + (1-\alpha_1)^\beta}{\eta_1^\beta}\right);\ Y = \exp\left(-\frac{\alpha_2^\beta + (1-\alpha_2)^\beta}{\eta_2^\beta}\right) \tag{14}$$

where $\beta \geq 1$, $0 \leq \alpha_1,\ \alpha_2 \leq 1$ and $\alpha_1 + \alpha_2 < 1$.

Let $f_1(x) = \frac{x^\beta}{x^\beta + (1-x)^\beta}$ and $f_2(x) = \frac{(1-x)^\beta}{x^\beta + (1-x)^\beta}$. Then $f_1(x)$ is monotonically increasing in [0,1] and $f_2(x)$ is monotonically decreasing in [0,1]. We have the following proposition.

Proposition 4.2 *Under model (14), we have*

$$GMC(Y|X) = 12E\big(\{E(Y|X)\}^2\big) - 3,\ GMC(X|Y) = 12E\big(\{E(X|Y)\}^2\big) - 3 \tag{15}$$

where

$$E\big(\{E(Y|X)\}^2\big) = \frac{(f_2(a)+f_2(b))^2}{(f_2(b)+1)^2}\frac{f_2(b)}{2f_2(a)+f_2(b)} + \frac{f_1^2(a)}{(f_2(b)+1)^2}\frac{f_1(b)}{4f_1(a)-f_1(b)} + \frac{2(f_2(a)+f_2(b))f_1(a)f_1(b)f_2(b)}{(f_2(b)+1)^2(2f_1(a)f_2(b)+f_2(a)f_1(b))}$$

and

$$E\big(\{E(X|Y)\}^2\big) = \frac{(f_1(b)+f_1(a))^2}{(f_1(a)+1)^2}\frac{f_1(a)}{2f_1(b)+f_1(a)} + \frac{f_2^2(b)}{(f_1(a)+1)^2}\frac{f_2(a)}{4f_2(b)-f_2(a)} + \frac{2(f_1(b)+f_1(a))f_2(b)f_2(a)f_1(a)}{(f_1(a)+1)^2(2f_2(b)f_1(a)+f_1(b)f_2(a))}$$

with $a = 1-\alpha_1$ *and* $b = \alpha_2$.

It is easy to see that $GMC(Y|X)$ and $GMC(X|Y)$ are functions of a, b, and β. They are not identical. They are equal, or the difference is negligible, i.e., $|GMC(Y|X) - GMC(X|Y)| < 10^{-10}$, for some chosen values of α_1, α_2 and β, for example in the following choices, $(\alpha_1, \alpha_2) = (0.1, 0.2)$ and $\beta \geq 16.4021$, $(\alpha_1, \alpha_2) = (0.1, 0.1)$ and $\beta \geq 1$. The following table is a numerical illustration of choices of $(\alpha_1, \alpha_2, \beta)$.

	$\beta \geq$	α_1 0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
	0.1	1	16.4021	26.8361	56.0792	$+\infty$	27.0396	12.9395	7.9043	1
	0.2	16.4021	1	26.8361	56.0792	$+\infty$	27.0396	12.9355	1	
	0.3	26.8361	26.8361	1	56.0792	$+\infty$	27.0395	1		
	0.4	56.0792	56.0792	56.0792	1	$+\infty$	1			
α_2	0.5	$+\infty$	$+\infty$	$+\infty$	$+\infty$	1				
	0.6	27.0396	27.0396	27.0396	1					
	0.7	12.9355	12.9355	1						
	0.8	7.9043	1							
	0.9	1								

5 Simulation examples

In this section, we use the following simulation procedure:

(1) Simulate a bivariate random sample from a pre-specified joint distribution using software package R v2.9.

(2) Empirical Type I errors and Empirical powers are calculated based on 1000 repeated samples.

Example 5.1 *In this example, we simulate bivariate normal samples with variances* $\sigma_1^2 = \sigma_2^2 = 1$*. Figure 1 displays Type I errors for the sample size* $n = 50$ *and correlation coefficient* ρ *ranging from* 0 *to* 0.8 *with sizes of test* 10% *and* 5% *in the left panel and in the right panel respectively. Figure 2 is Type I errors for correlation coefficient* $\rho = 0.40$ *and changing sample sizes* $n = 30, 40, 50, 75, 150$*.*

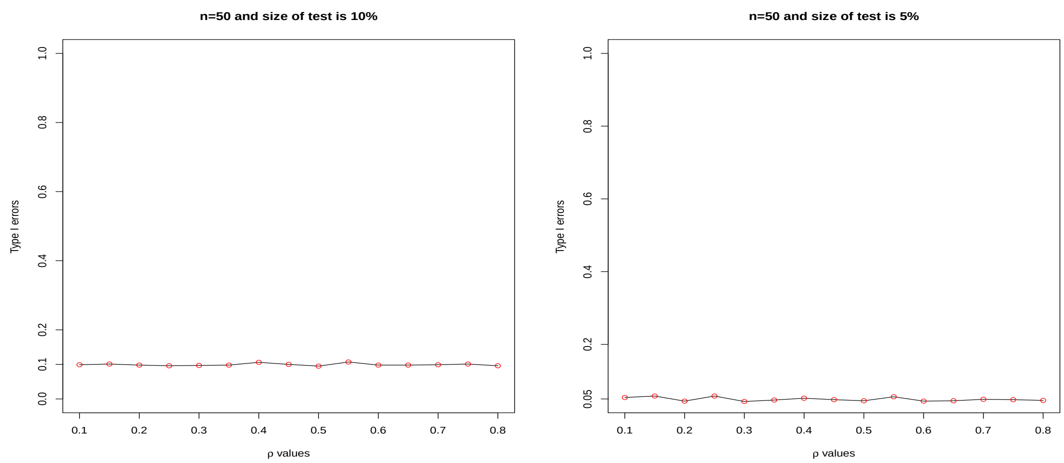

Figure 1: The demonstration of sample performance of bivariate normal distribution with sample size $n = 50$ and varying correlation coefficient values.

This example shows that for bivariate normal random variables, when the sample size is no less than 50, Type I error probabilities are well controlled within their nominal levels.

Example 5.2 *In this example, we simulate bivariate* t *samples (Case 2) with degrees of freedom* $df = 9$*. Figure 3 displays Type I errors for the sample size* $n = 50$ *and correlation coefficient* ρ *ranging from* 0 *to* 0.8 *with sizes of test* 10% *and* 5% *in the left panel and in the right panel respectively. Figure 4 is Type I errors for correlation coefficient* $\rho = 0.40$ *and changing sample sizes* $n = 30, 35, 40, 50, 75, 150$*.*

This example shows that for bivariate t random variables (Case 2), when the sample size is greater than 50, Type I error probabilities are controlled within their nominal levels.

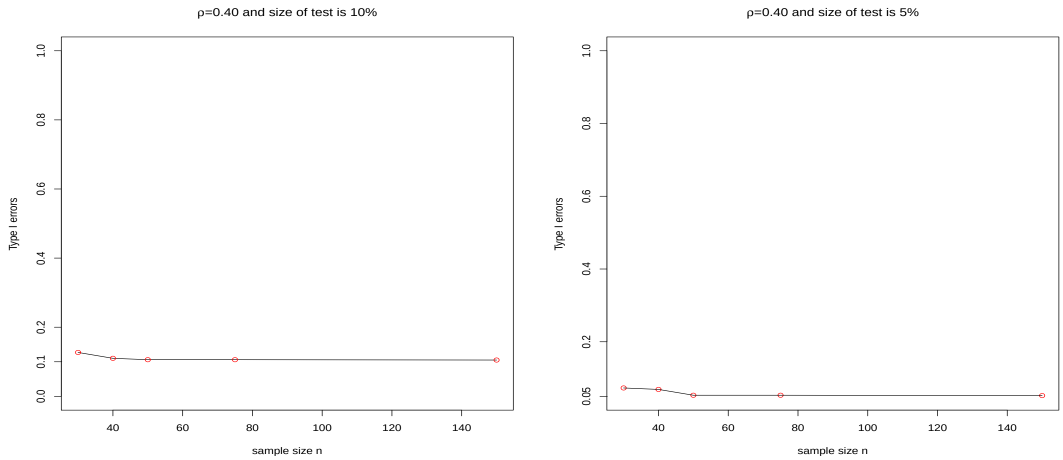

Figure 2: The demonstration of sample performance of bivariate normal distribution with correlation coefficient $\rho = 0.40$ and varying sample sizes.

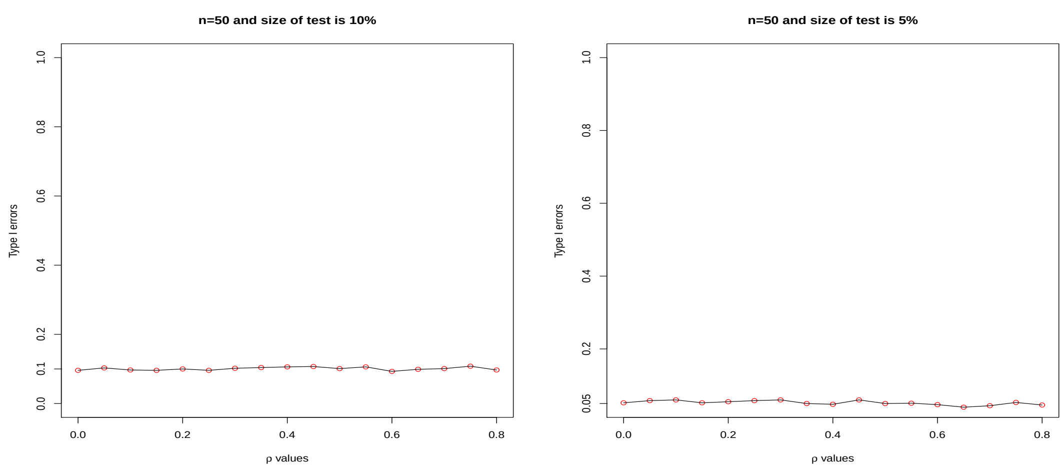

Figure 3: The demonstration of sample performance of bivariate t distribution (Case 2) with sample size $n = 50$ and varying correlation coefficient values.

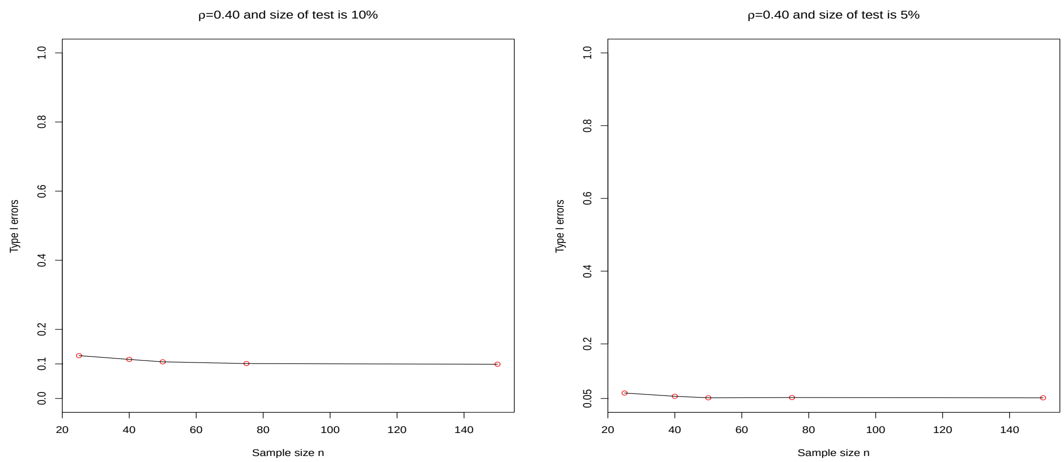

Figure 4: The demonstration of sample performance of bivariate t distribution (Case 2) with correlation coefficient $\rho = 0.40$ and varying sample sizes.

Example 5.3 *In this example, we simulate bivariate* t *samples (Case 3) with the degrees of freedom* $(df1 = 2, df2 = 21)$. *Figure 5 reveals empirical powers for the sample sizes* $n =$ 50, 75, 150, 200, 250, 500, 750. *The sizes of test are* $\alpha = 0.10$ *and* $\alpha = 0.05$, *respectively in the left panel and in the right panel.*

We can see that Example 5.3 clearly shows that with sufficiently large sample size, our proposed estimators and test statistics are able to tell whether explained variances are identical or not.

Example 5.4 *In this example, we simulate bivariate samples from Model (14). In Figure 6, we plot empirical powers for* $(\alpha_1, \alpha_2) = (0.4, 0.1))$ *for the sample sizes* $n = 25, 50, 75, 150, 200, 250, 500$, *and the sizes of test* $\alpha = 0.10$ *and* $\alpha = 0.05$ *respectively. In Figure 7, we plot powers for* $\alpha_2 = 0.1$ *and changing* α_1 *values while the sample sizes are fixed at* $n = 100, 300$.

Figure 6 demonstrates that when two GMCs are not identical, empirical powers are increasing along with increasing sample sizes. Figure 7 displays that Model (14) is a flexible copula model for a wide range of dependence and explained variances between random variables. In the figure, with the sample size being fixed, empirical powers give indications of how close of two GMCs.

Example 5.5 *In this example, we simulate bivariate sample from the following bivariate time series*

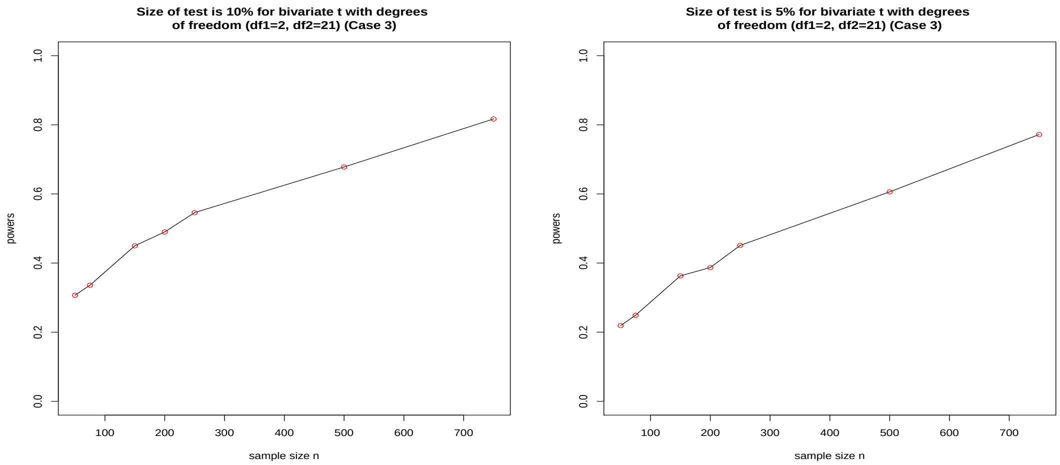

Figure 5: The demonstration of sample performance of bivariate t distribution (Case 3) with changing sample sizes.

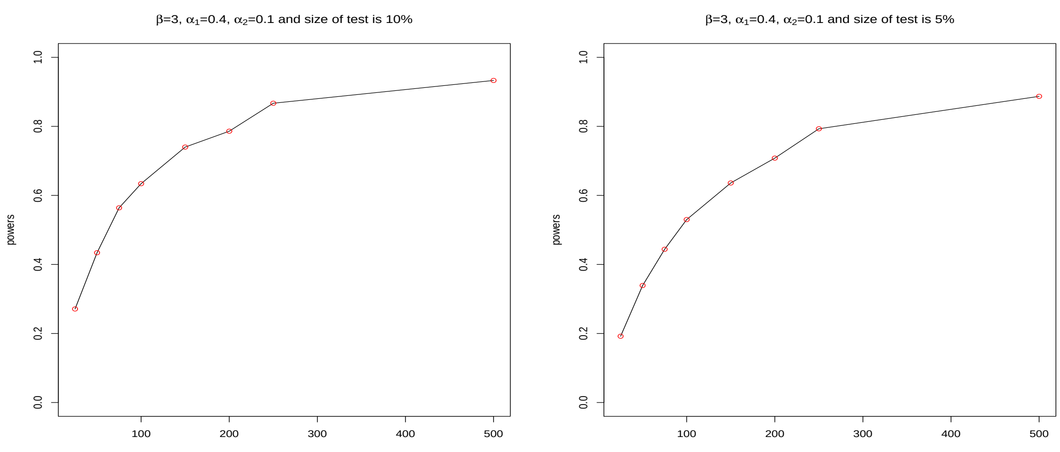

Figure 6: The demonstration of sample performance of Model (14) with changing sample sizes.

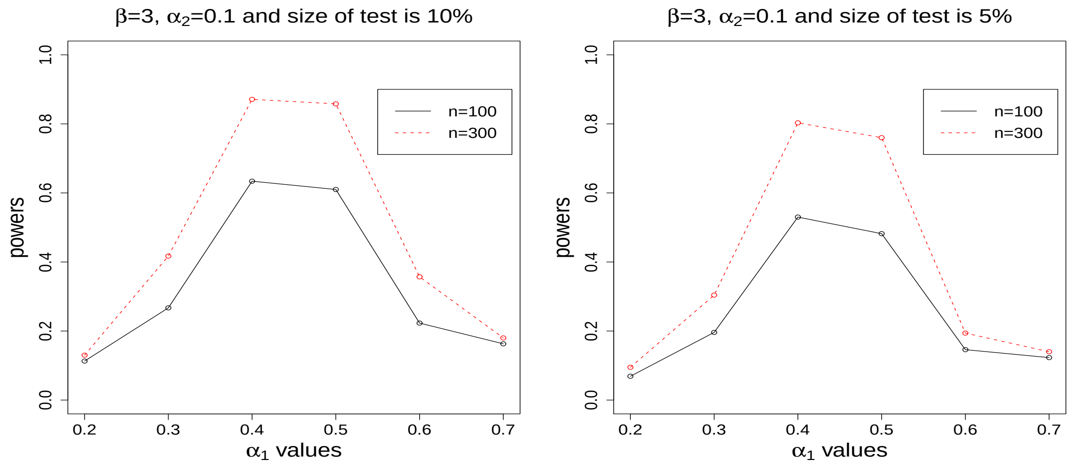

Figure 7: The demonstration of sample performance of Model (14) with sample sizes $n = 100, 300$ and changing a_1 values.

model:

$$X_i = 0.3X_{i-1} + 0.2X_{i-2} + \epsilon_i^X$$

$$Y_i = 0.5Y_{i-1} - 0.1Y_{i-2} + \epsilon_i^Y$$

where ϵ_i^X *and* ϵ_i^Y *follow a bivariate normal distribution with correlation coefficient* $\rho = 0.4$*. In Figure 8, we plot Type I errors for the sample sizes* $n = 25, 50, 75, 150, 200, 250, 500$*, and the sizes of test* $\alpha = 0.10$ *and* $\alpha = 0.05$ *respectively.*

We can see that with sufficiently large sample sizes, Type I errors are controlled within their corresponding nominal levels. This example suggests that GMCs are also suitable for time series data, which is a very important property in practice. In Section 6, we will apply GMC to economic time series data analysis.

One can see from these five examples that dealing with non-identical GMCs is a challenge task, which was also witnessed in Section 4 where mathematical derivations of GMCs for several cases of bivariate distributions were presented. Nevertheless, our simulation examples show that our proposed nonparametric estimators for GMCs are still able to efficiently estimate the values of GMCs and detect whether GMCs are identical or not with a sufficiently large sample size.

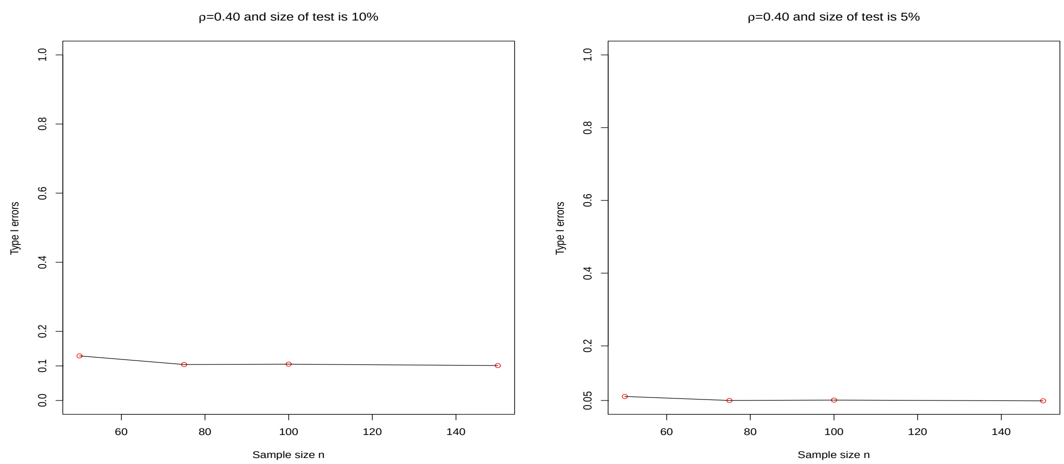

Figure 8: The demonstration of sample performance of bivariate time series with bivariate normal random errors ($\rho = 0.4$) and varying sample sizes.

6 Real data analysis

United States and Japan are two largest economies in the world. The relationship between these two economic powers is very strong and mutually advantageous. These two countries both suffered massive banking and financial crises as Japan started in 1989, which was followed by a long period of slow growth and deflation, and US started in 2008, which will remain a depressed economic growth. Comparing the US to Japan has drawn main attentions among politicians, economists, investors, and researchers. Among many comparisons, researchers have focused on illustrating national GDP, imports, exports, S&P500 index, Nikkei index, exchange rates, and other market variables. People have hoped that the comparisons may help reveal similarities and find answers (even solutions) to an economic recovery from the current international financial crisis.

This section aims to reveal an uncanny relationship via our proposed GMC using US and Japan economic variables. Particularly, we consider monthly average exchange rates from the Japanese Yen against the US dollar (JPY/USD), and monthly US federal funds rates, and monthly Japan deposit rates respectively. They are very important economic indicators. Our data source is International Monetary Fund (IMF), and the data is available at http://www.imf.org. The time range is from January 1957 to April 2009. They are plotted in Figure 9. From a market perspective, plotting these variables shows no similarity, linear relationship, or co-monotone relationship. However, our

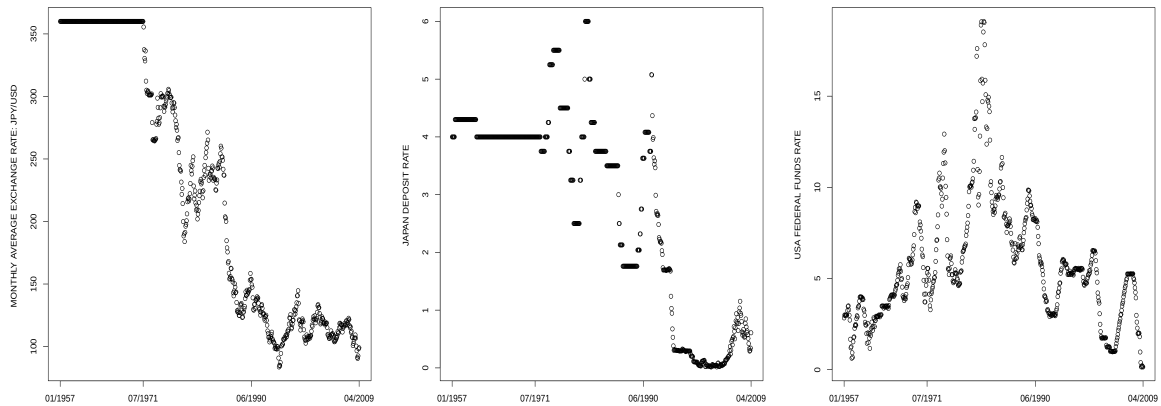

Figure 9: Plots of monthly average exchange rate of JPY/USD (left panel), monthly Japan deposit rates (middle panel), and US federal funds rates (right panel)

peculiar GMC shall display economic changes between these two countries.

When computing GMC using Japan deposit rates, we use data from August 1967 to April 2009 because Japan deposit rate did not show any changes before August 1967. When computing GMC using the exchange rates, we use data from July 1971 to April 2009 since the exchange rates before July 1971 did not show any changes. Considering that the Bureau of Economic Analysis (BEA) estimates of gross domestic product (GDP) are among the most widely scrutinized indicators of U.S. economic activity, and BEA releases a comprehensive revision about every 5 years, we calculate GMC using a five year window (60 months) and the following procedure:

- Suppose data are $\{(x_1, y_1), \ldots, (x_n, y_n), (x_{n+1}, y_{n+1}), \cdots, (x_{n+59}, y_{n+59})\}$.

  1) $i = 1$

  2) Use $\{(x_i, y_i), (x_{i+1}, y_{i+1}), \cdots, (x_{i+59}, y_{i+59})\}$ to compute $GMC(X_i|Y_i)$ and $GMC(Y_i|X_i)$

  3) $i = i + 1$ and repeat 2) until $i = n$.

One can see that this procedure will generate two dynamic GMC curves showing economic status changes over time. For notational convenience, we shall use brief letters 'E' to stand for the exchange rates of Japanese Yen against US dollar, 'J' to mean Japan deposit rates, and 'U' to indicate US federal funds rates. For example, GMC(E|U) stands for the proportion of variation of the exchange rates explained by US federal funds rates. The interpretations of the rest of notations are similar.

Figure 10 displays GMC(E|J), the absolute value of ordinary Pearson correlation $|cor(E, J)|$, and GMC(J|E) from July 1971 to April 2004. One can immediately see that the overall trends among the three curves are similar. The variations of $|cor(E, J)|$ look larger than the variations of GMC, which may indicate that the relationship between the two market variables is not linearly dependent. The curves of GMC reflect the history of Japan economy. The first valley of GMC curves occurred between 1972 and 1981. During this time period, the United States and Japan experienced three major economic conflicts: the first one led to the U.S. import surcharge of August 1971; the second one damaged Japanese confidence in its American connection and immediate impact on the political career of the then Prime Minister, Takeo Fukuda due to major U.S. pressure on Japan during 1977-78 to boost its domestic growth rate; the third one concerned the reemerging issue of security relations between the two countries; see Bergsten (1982) for more details. Notice that during this period, $GMC(E|J)$ is larger than $GMC(J|E)$, i.e. the strength of explained variance in exchange rates by the deposit rate is stronger than the strength of explained variance in deposit rates by the exchange rate. This phenomenon coincided with the damaged Japanese confidence in its American connection. It also clearly suggests that GMC can provide more information than Pearson's correlation can provide. The second valley occurred between 1986 and 1990 in which Japan experienced one of the great bubble economies in history. It began after the Japanese agreed, in the so-called Plaza Accord with the United States in 1985, to increase substantially the value of the Yen (which doubled by 1988), see Asher (1996) for more details. After the 1989 Japan economic crisis, the dynamic variations in computed empirical GMC are smaller than those in the proceeding time periods, which may indicate that other economic, social or political factors play a role in the variations of these two economic variables. We note that in the plot there are time points that GMC((E|J) and GMC(J|E) changed relative positions. These points tell that which economic variable is more influential in explained variances during a particular time period, which may also reflect the foreign relations between two countries. Based on the right panel of the figure, we can see that the deposit rate has more impacts in Japan economy growth than the exchange rate has.

Figure 11 compares GMC(E|U), the absolute value of ordinary Pearson correlation $|cor(E, U)|$, and GMC(U|E) from July 1971 to April 2004. The overall trends among the three curves are similar

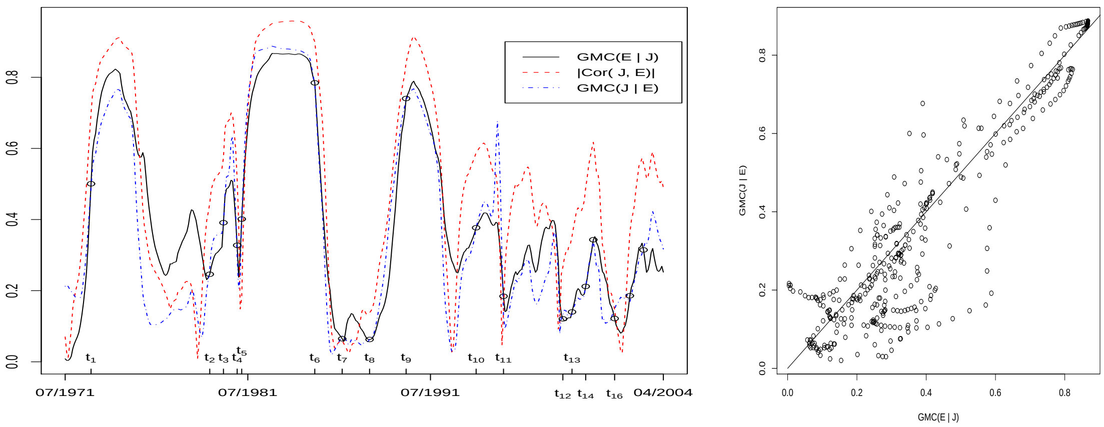

Figure 10: Plots of GMC and ordinary correlation coefficients between JPY/USD exchange rates and Japan deposit rates over time. Time points that GMC((E|J) and GMC(J|E) change relative position are $t_1 = 12/1972$, $t_2 = 06/1979$, $t_3 = 03/1980$, $t_4 = 12/1980$, $t_5 = 03/1983$, $t_6 = 03/1985$, $t_7 = 09/1986$, $t_8 = 03/1988$, $t_9 = 03/1990$, $t_{10} = 01/1994$, $t_{11} = 07/1995$, $t_{12} = 10/1998$, $t_{13} = 04/1999$, $t_{14} = 01/2000$, $t_{15} = 06/2000$, $t_{16} = 08/2001$, $t_{17} = 06/2002$, $t_{18} = 03/2003$. The right panel plots pairwise GMC with a straight line of 45 degrees.

to those in Figure 10. One can see that the influence of the federal funds rates on the exchange rates is more significant than the Japan deposit rates on the exchange rates. There are less number of points that GMC((E|U) and GMC(U|E) changed relative positions. We note that from October 1992 to August 1998, $GMC(U|E)$ is larger than $GMC(E|U)$, i.e. the strength of explained variance in the federal funds rates by the exchange rates is much stronger than the strength of explained variance in the exchange rates by the federal funds rates. This empirical finding is again coincided with the economic status during that time period as Griswold (1998) stated “From 1992 and 1997, the U.S. trade deficit almost tripled, while at the same time U.S. industrial production increased by 24 percent and manufacturing output by 27 percent. Trade deficits do not cost jobs. In fact rising trade deficits correlate with falling unemployment rates. Far from being a drag on economic growth, the U.S. economy has actually grown faster in years in which the trade deficit has been rising than in years in which the deficit has shrunk. Trade deficits may even be good news for the economy because they signal global investor confidence in the United States and rising purchasing power among domestic consumers.” This is another evidence that GMC is superior in explaining asymmetry of market

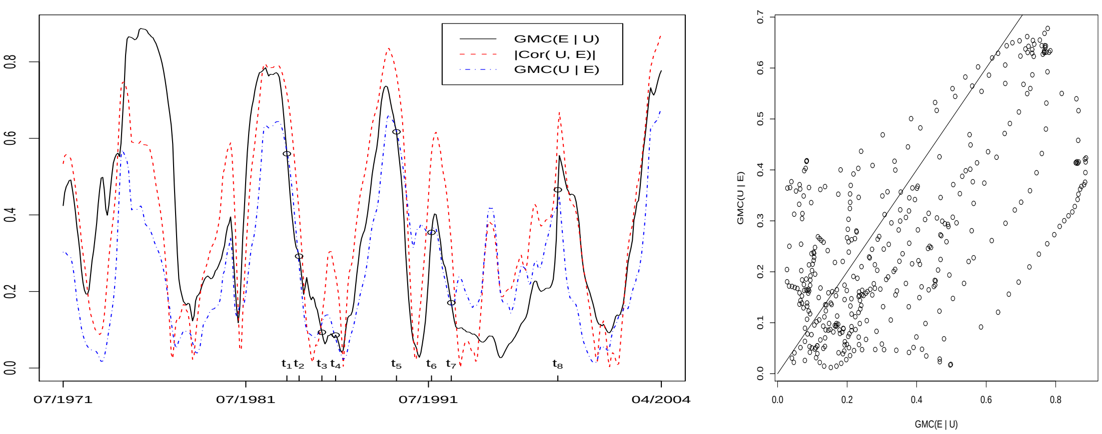

Figure 11: Plots of GMC and ordinary correlation coefficients between JPY/USD exchange rates and US federal funds rates over time. Time points that GMC((E|U) and GMC(U|E) change relative position are $t_1 = 10/1983$, $t_2 = 06/1984$, $t_3 = 09/1985$, $t_4 = 06/1986$, $t_5 = 10/1989$, $t_6 = 09/1991$, $t_7 = 10/1992$, $t_8 = 08/1998$. The right panel plots pairwise GMC with a straight line of 45 degrees.

movements while the ordinary correlation certainly can not achieve this purpose.

Figure 12 shows GMC(J|U), the absolute value of ordinary Pearson correlation $|cor(J, U)|$, and GMC(U|J) from July 1971 to April 2004. We see that the US federal funds rates and Japan deposit rates mutually influence each other. There are more number of points that GMC((J|U) and GMC(U|J) changed relative position. We note that between 1977 and 1981, $GMC(J|U)$ is much larger than $GMC(U|J)$, which tells that the US money regulation policy had a major impact in Japan deposit rates. During that time period, US President Jimmy Carter tried to combat economic weakness and unemployment by increasing government spending, and he established voluntary wage and price guidelines to control inflation. But the most important element in the war against inflation was the Federal Reserve Board, which clamped down hard on the money supply beginning in 1979. By refusing to supply all the money an inflation-ravaged economy wanted, the Fed caused interest rates to rise. As a result, consumer spending and business borrowing slowed abruptly. The economy soon fell into a deep recession (Source: US Economy in 1970s from U.S. Department of State). Before Japan 1989 economic crisis (time t_{11}), $GMC(J|U)$ is much smaller than $GMC(U|J)$ (t_{10} to t_{11}), while after the economic crisis, $GMC(J|U)$ is much larger than $GMC(U|J)$ (t_{11} to t_{12}), i.e. the

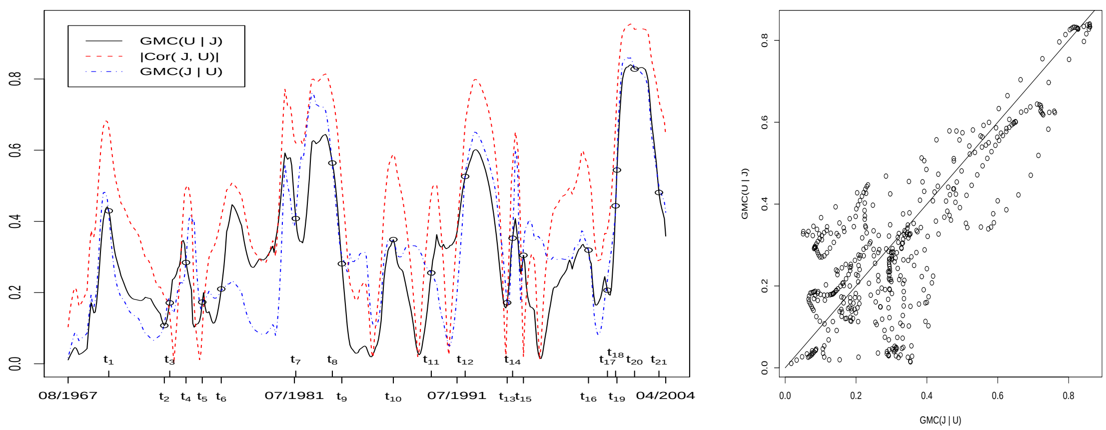

Figure 12: Plots of GMC and ordinary correlation coefficients between US federal funds rates and Japan deposit rates over time. Time points that GMC((J|U) and GMC(U|J) change relative position are $t_1 = 02/1970$, $t_2 = 07/1973$, $t_3 = 11/1973$, $t_4 = 11/1974$, $t_5 = 11/1975$, $t_6 = 01/1977$, $t_7 = 08/1981$, $t_8 = 11/1983$, $t_9 = 06/1984$, $t_{10} = 08/1987$, $t_{11} = 12/1989$, $t_{12} = 01/1992$, $t_{13} = 08/1994$, $t_{14} = 12/1994$, $t_{15} = 08/1995$, $t_{16} = 08/1999$, $t_{17} = 10/2000$, $t_{18} = 04/2001$, $t_{19} = 05/2001$, $t_{20} = 06/2002$, $t_{21} = 12/2003$. The right panel plots pairwise GMC with a straight line of 45 degrees.

explained variances in Japan economy had been delayed by the US economy. This is not surprising as at the beginning of an economic recession, the economy is weak, and it can hardly have an immediate economic bounce and recovery, i.e. the influence from other economic variable is weak.

Figure 13 compares auto generalized measures of correlation, i.e. AGMC. From the left panel and the middle panel, one can see that the lagged (one month) impact from the exchange rate on US federal funds rate and Japan deposit rate is higher than either the lagged-1 impact on the exchange rate from US federal funds rate or the lagged-1 impact from Japan deposit. This phenomenon may suggest that the exchange rates can be used to help researcher to build more reliable prediction models. The right panel in the figure shows that US economy is more influential than Japan.

This real data analysis clearly shows that our newly proposed GMC is more informative in explaining variations and movements in economic and financial monetary indicators. Our empirical findings show that there are some economic similarities between US and Japan, however the economic development dynamics between these two economic powers are asymmetric, and the uni-

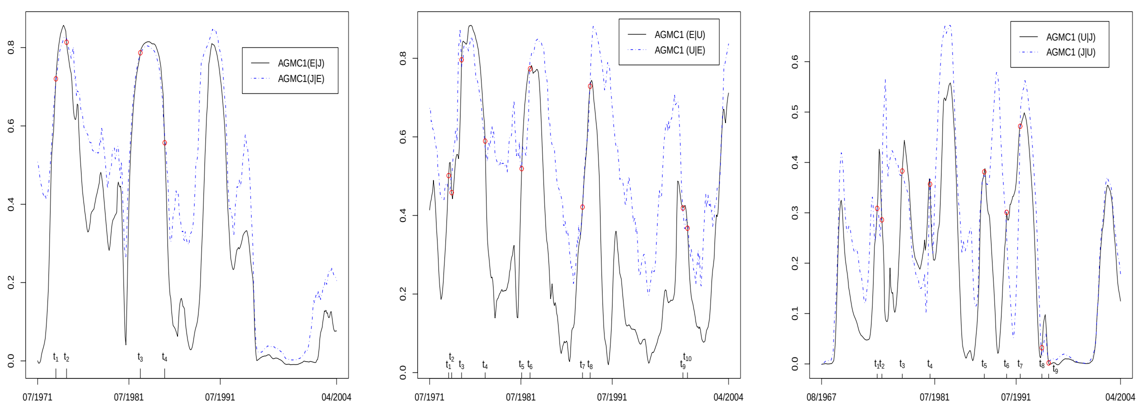

Figure 13: Plots of AGMC. The left panel is AGMC between JPY/USD exchange rates and Japan deposit rates over time. Time points that AGMC_1((E|J) and AGMC_1(J|E) change relative position are $t_1 = 07/1973$, $t_2 = 09/1974$, $t_3 = 10/1982$, $t_4 = 06/1985$. The middle panel is AGMC between JPY/USD exchange rates and US federal funds rates over time. Time points that AGMC_1((E|U) and AGMC_1(U|E) change relative position are $t_1 = 08/1973$, $t_2 = 12/1973$, $t_3 = 01/1975$, $t_4 = 08/1977$, $t_5 = 08/1981$, $t_6 = 07/1982$, $t_7 = 04/1988$, $t_8 = 02/1989$, $t_9 = 04/1999$, $t_{10} = 10/1999$. The right panel is AGMC between US federal funds rates and Japan deposit rates over time. Time points that AGMC_1((J|U) and AGMC_1(U|J)— change relative position are $t_1 = 06/1974$, $t_2 = 01/1975$, $t_3 = 07/1977$, $t_4 = 12/1980$, $t_5 = 08/1987$, $t_6 = 05/1990$, $t_7 = 01/1992$, $t_8 = 09/1994$, $t_9 = 07/1995$.

versal truth is still that US has more impacts in the world economy. As a result, our findings may be helpful in making monetary regulation policies.

7 Conclusions

We have demonstrated that our newly proposed GMC is superior in characterizing the asymmetry of explained variances. GMC contains the ordinary correlation coefficient as a special case when two random variables are related in a linear equation or they are bivariate normally distributed. Theoretical foundations of GMC show that when two random variables are correlated through a measurable function, at least one of GMC takes the extreme value one while the ordinary correlation coefficient can still be zero. GMC also obeys monotone properties. These properties are strong evidences that GMC is a true nonlinear dependence measure, especially in explained variances. It may be safe to say that GMC can be applied to many research areas where Pearson's correlation coefficient is either applicable or not applicable.

Our definitions of GMC are mainly for bivariate random variables. It is possible to extend the definitions to cases of multivariate random variables, i.e. we shall deal with the explained variance in X_1 by $X_2,\ X_3,\dots,\ X_k$. In an attempt to relate GMC to Granger causality between two time series, we introduced Granger causality generalized measures of correlation (GcGMC). We shall conduct a full study of properties of GcGMC in a separate project. GcGMC will be applied to bivariate time series study. We expect to obtain more meaningful results and discover things previously not revealed.

We note that the computation of Pearson's correlation coefficient is easy, while GMC involves conditional expectations, and hence it may be computationally challenge in practice. Based on our simulation examples, we found as long as sample size is 60 or larger, our nonparametric estimators give good approximated values of GMC. Our nonparametric estimators are kernel based estimators, which is a standard procedure in nonparametric statistics, and therefore they can easily be implemented in any existing software packages.

It is worth noting that Little and Rubin (1987) and Liu (1994) proposed a similar indicator to illustrate the fraction of the missing information in the data augmentation. To the best of our knowledge, the properties of their indicator as the correlation have not been discussed in detail.

Analog to their indicator, our generalized measures of correlation can be thought as an idea of increasing dimension to measure the asymmetry.

8 Appendix: Proofs of Properties

Proof of Proposition 2.1

(i). The proof is obvious.

(ii). If $\rho_{XY} = \pm 1$ or $Y = a + bX$ where $b \neq 0$, we have $GMC(Y|X) = 1$ and $GMC(X|Y) = 1$. When $GMC(Y|X) = 0$, we have $E(Y|X) = EY$ or $EXY = EXEY$, and then $\rho_{XY} = 0$. When the asymmetric correlation measure is zero, that is, $GMC(Y|X) = 0$ or $GMC(X|Y) = 0$, Pearson's correlation satisfies $\rho_{XY} = 0$. If $\rho_{XY} \neq 0$, we have $EXY \neq EX \cdot EY$. Then we obtain $GMC(X|Y) \neq 0 \quad \text{and} \quad GMC(Y|X) \neq 0$.

(iii). The proof of the first part is straightforward. We prove the second part here. Because

$$\begin{aligned} GMC(Y|X) = 1 \quad &\Longleftrightarrow E[\{Y - E(Y|X)\}^2] = 0 \Longleftrightarrow Y = E(Y|X) \ a.s. \\ &\Longleftrightarrow \text{Y is a measurable function of X} \end{aligned}$$

and

$$\begin{aligned} GMC(Y|X) = 0 \quad &\Longleftrightarrow E[\{Y - E(Y|X)\}^2] = Var(Y) \\ &\Longleftrightarrow E(Y|X) = EY \ a.s. \end{aligned}$$

then $GMC(Y|X) = 1 \Longleftrightarrow Y$ is a measurable function of X. Moreover, if X and Y are independent, then $GMC(Y|X) = 0$ and $GMC(X|Y) = 0$ because of $E(Y|X) = EY$ and $E(X|Y) = EX$.

(iv). The proof is straightforward.

(v). Because $E[\{Y - f(X)\}^2] = E[\{Y - E(Y|X)\}^2] + E[\{E(Y|X) - f(X)\}^2]$, we have $\inf_f E[\{Y - f(X)\}^2] = E[\{Y - E(Y|X)\}^2]$. Similarly, we have $\inf_g E[\{Z - g(X)\}^2] = E[\{Z - E(Z|X)\}^2]$. If $\inf_f E[\{Y - f(X)\}^2] = \inf_g E[\{Z - g(X)\}^2]$ and $Var(Y) = Var(Z)$, then we have $GMC(Y|X) = GMC(Z|X)$. If $\inf_f E[\{Y - f(X)\}^2] < \inf_g E[\{Z - g(X)\}^2]$ and $Var(Y) = Var(Z)$, then we have $GMC(Y|X) > GMC(Z|X)$. $\square$

Proof of Proposition 2.2. Suppose $(X, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Then we have

$$E(Y|X) = \mu_2 + \frac{\rho\sigma_1\sigma_2}{\sigma_2^2}(X - \mu_1), \qquad E(X|Y) = \mu_1 + \frac{\rho\sigma_1\sigma_2}{\sigma_1^2}(Y - \mu_2),$$

$$Var(X|Y) = E[\{X - E(X|Y)\}^2] = \sigma_1^2 - \frac{(\rho\sigma_1\sigma_2)^2}{\sigma_2^2},$$

and

$$Var(Y|X) = E[\{Y - E(Y|X)\}^2] = \sigma_2^2 - \frac{(\rho\sigma_1\sigma_2)^2}{\sigma_2^2}.$$

So we obtain $GMC(Y|X) = GMC(X|Y) = \rho^2$. $\square$

Proof of Proposition 2.5. We have

$$\begin{aligned} E[\{Y_t - E(Y_t|Y_{t-1}, X_{t-1})\}^2] &= E[\{Y_t - E(Y_t|Y_{t-1}) + E(Y_t|Y_{t-1}) - E(Y_t|Y_{t-1}, X_{t-1})\}^2] \\ &= E[\{Y_t - E(Y_t|Y_{t-1})\}^2] - E[\{E(Y_t|Y_{t-1}) - E(Y_t|Y_{t-1}, X_{t-1})\}^2] \end{aligned}$$

and

$$\begin{aligned} E[\{X_t - E(X_t|X_{t-1}, Y_{t-1})\}^2] &= E[\{X_t - E(X_t|X_{t-1}) + E(X_t|X_{t-1}) - E(X_t|X_{t-1}, Y_{t-1})\}^2] \\ &= E[\{X_t - E(X_t|X_{t-1})\}^2] - E[\{E(X_t|X_{t-1}) - E(X_t|X_{t-1}, Y_{t-1})\}^2]. \end{aligned}$$

Then

$$\frac{E[\{E(Y_t|Y_{t-1}) - E(Y_t|Y_{t-1}, X_{t-1})\}^2]}{E(Var(Y_t|Y_{t-1}))} = 1 - \frac{E[\{Y_t - E(Y_t|Y_{t-1}, X_{t-1})\}^2]}{E(Var(Y_t|Y_{t-1}))}$$

and

$$\frac{E[\{E(X_t|X_{t-1}) - E(X_t|X_{t-1}, Y_{t-1})\}^2]}{E(Var(X_t|X_{t-1}))} = 1 - \frac{E[\{X_t - E(X_t|X_{t-1}, Y_{t-1})\}^2]}{E(Var(X_t|X_{t-1}))}.$$

Thus

$$GcGMC(Y_t|X_{t-1}, Y_{t-1}) = \frac{E[\{E(Y_t|Y_{t-1}) - E(Y_t|Y_{t-1}, X_{t-1})\}^2]}{E(Var(Y_t|Y_{t-1}))}$$

and

$$GcGMC(X_t|Y_{t-1}, X_{t-1}) = \frac{E[\{E(X_t|X_{t-1}) - E(X_t|X_{t-1}, Y_{t-1})\}^2]}{E(Var(X_t|X_{t-1}))}.$$

If $E(Var(X_t|X_{t-1})) = E(Var(Y_t|Y_{t-1}))$ and the strength of that X_t Granger causes Y_t is stronger than the strength of that Y_t Granger causes X_t, i.e.

$$E[\{E(Y_t|Y_{t-1}) - E(Y_t|Y_{t-1}, X_{t-1})\}^2] > E[\{E(X_t|X_{t-1}) - E(X_t|X_{t-1}, Y_{t-1})\}^2],$$

we then have $GcGMC(Y_t|X_{t-1}, Y_{t-1}) > GcGMC(X_t|Y_{t-1}, X_{t-1})$.

The proof of Proposition 2.5 is completed. $\square$

The following assumptions are given for Lemma 3.1.

Assumption (a): $K(x)$ is a symmetric kernel about $x=0$ satisfying $\int K(x)=1$ and $x^2K(x)\in L_1(-\infty,\infty)$;

Assumption (b): The densities $f^X(x)$ and $f^Y(y)$ are bounded on the whole axis, and $I_x=\{x: f^X(x)=0, s_x<x<S^x\}$ and $I_y=\{y:\ f^Y(y)=0, s_y<y<S^y\}$ both have Lebesgue measure 0. Moreover, $\int\big(\int\int yf_x^{''}(x+\theta^*ht,y)\cdot t^2K(t)dtdy\big)^2dx$ and $\int\big(\int\int xf_y^{''}(x,y+\theta^*ht)\cdot t^2K(t)dtdx\big)^2dy$ are bounded for $0<\theta^*<1$.

Assumption (c): The fourth-order mixed moments of X and Y exist.

Proof of Lemma 3.1. We first establish the following equation.

$$
\begin{aligned}
&\int\frac{(\phi_n^{Y|X}(x))^2}{f^X(x)}dx-\int\frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx\\
&=\int\frac{(\phi_n^{Y|X}(x))^2}{f^X(x)}dx-\int\frac{(E\phi_n^{Y|X}(x))^2}{f^X(x)}dx+\int\frac{(E\phi_n^{Y|X}(x))^2}{f^X(x)}dx-\int\frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx\\
&=\int\frac{(\phi_n^{Y|X}(x))^2}{f^X(x)}dx-\int\frac{(E\phi_n^{Y|X}(x))^2}{f^X(x)}dx\\
&\quad+\int\frac{((E\phi_n^{Y|X}(x))-\phi^{Y|X}(x))((E\phi_n^{Y|X}(x))+\phi^{Y|X}(x))}{f^X(x)}dx\\
&=\int\frac{(\phi_n^{Y|X}(x))^2}{f^X(x)}dx-\int\frac{(E\phi_n^{Y|X}(x))^2}{f^X(x)}dx+O(h^2)\\
&=\int\frac{(\phi_n^{Y|X}(x)-E\phi_n^{Y|X}(x))^2}{f^X(x)}dx+\int\frac{2(\phi_n^{Y|X}(x)-E\phi_n^{Y|X}(x))E\phi_n^{Y|X}(x)}{f^X(x)}dx+O(h^2)\\
&=\int\frac{(\phi_n^{Y|X}(x)-E\phi_n^{Y|X}(x))^2}{f^X(x)}dx+\int\frac{2(\phi_n^{Y|X}(x)-E\phi_n^{Y|X}(x))\phi^{Y|X}(x)}{f^X(x)}dx+O(h^2) \qquad (16)
\end{aligned}
$$

where the third and last equality hold because under Assumption (a)-(c)

$$
\begin{aligned}
&E\phi_n^{Y|X}(x)-\phi^{Y|X}(x)\\
=\ &\int y\int K(t)\Bigg(f(x-ht,y)-f(x,y)\Bigg)dtdy\\
=\ &\int y\int K(t)\Bigg(-htf_x^{'}(x,y)+0.5h^2t^2f_x^{''}(x+\theta^*ht,y)\Bigg)dtdy\\
=\ &0.5h^2\int\int yf_x^{''}(x+\theta^*ht,y)\cdot t^2K(t)dtdy=O(h^2).
\end{aligned} \qquad (17)
$$

Similarly we have $E\phi_n^{X|Y}(y) - \phi^{X|Y}(y) = O(h^2)$. We now express the left-hand of (11) in a sum of five terms by

$$
\begin{aligned}
&\int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)}dx - \int \frac{(\phi_n^{Y|X}(x))^2}{f^X(x)}dx \\
= \;& -\int \frac{(f_n^X(x) - Ef_n^X(x))(\phi^{Y|X}(x))^2}{(f^X(x))^2}dx \\
&+ \int \frac{(f^X(x) - Ef_n^X(x))(\phi^{Y|X}(x))^2}{(f^X(x))^2}dx + \int \frac{(f^X(x) - f_n^X(x))^2(\phi^{Y|X}(x))^2}{(f^X(x))^2 f_n^X(x)}dx \\
&+ \int \frac{(f^X(x) - f_n^X(x))(\phi_n^{Y|X}(x) - \phi^{Y|X}(x))^2}{f^X(x) f_n^X(x)}dx \\
&+2 \int \frac{(f^X(x) - f_n^X(x))(\phi_n^{Y|X}(x) - \phi^{Y|X}(x))\phi^{Y|X}(x)}{f^X(x) f_n^X(x)}dx.
\end{aligned}
$$

Considering the second term in the sum above, we have

$$
\int \frac{|f^X(x) - Ef_n^X(x)| \cdot (\phi^{Y|X}(x))^2}{(f^X(x))^2}dx = O(h^2)
$$

because under Assumptions (a)-(c)

$$
\begin{aligned}
Ef_n^X(x) - f^X(x) &= \int K(t)\left[f^X(x - ht) - f^X(x)\right]dt \\
&= \int K(t)\left[-ht(f^X(x))' + 0.5h^2t^2(f^X(x + \theta^* ht))''\right]dt \\
&= 0.5h^2 \int \left(f^X(x + \theta^* ht)\right)'' \cdot t^2 K(t)dt = O(h^2).
\end{aligned} \tag{18}
$$

Similarly we have $Ef_n^Y(y) - f^Y(y) = O(h^2)$. For the third term in the sum, we have

$$
\begin{aligned}
&E\int \frac{(f^X(x) - f_n^X(x))^2(\phi^{Y|X}(x))^2}{(f^X(x))^2 f_n^X(x)}dx \\
&= \int E\frac{(f^X(x) - f_n^X(x))^2(\phi^{Y|X}(x))^2}{(f^X(x))^2 f_n^X(x)}dx \\
&= \int E\frac{(f^X(x) - f_n^X(x))^3(\phi^{Y|X}(x))^2}{(f^X(x))^2 f^X(x)}dx + \int E\frac{(f^X(x) - f_n^X(x))^2(\phi^{Y|X}(x))^2}{(f^X(x))^3}dx.
\end{aligned}
$$

Because

$$
\begin{aligned}
&\int E\frac{(f^X(x) - f_n^X(x))^2(\phi^{Y|X}(x))^2}{(f^X(x))^3}dx \\
&= \int \frac{(f^X(x) - Ef_n^X(x))^2 \cdot (\phi^{Y|X}(x))^2}{(f^X(x))^3}dx + \int \frac{E(f_n^X(x) - Ef_n^X(x))^2 \cdot (\phi^{Y|X}(x))^2}{(f^X(x))^3}dx \\
&+2\int \frac{(f^X(x) - Ef_n^X(x))E(f_n^X(x) - Ef_n^X(x)) \cdot (\phi^{Y|X}(x))^2}{(f^X(x))^3}dx
\end{aligned}
$$

$$= O(h^4) + O(\frac{1}{nh}) + O(h^2) = O(\frac{1}{nh}) + O(h^2)$$

by Assumptions (a)-(c) and (17)-(18), we have

$$E\int \frac{(f^X(x) - f_n^X(x))^2(\phi^{Y|X}(x))^2}{(f^X(x))^2 f_n^X(x)}dx = O(\frac{1}{nh}) + O(h^2).$$

For the fourth term in the sum, we have

$$\begin{aligned}
& E\int \frac{|f^X(x) - f_n^X(x)| \cdot (\phi_n^{Y|X}(x) - \phi^{Y|X}(x))^2}{f^X(x) f_n^X(x)}dx \\
\leq & \int E\frac{|f^X(x) - f_n^X(x)| \cdot (\phi_n^{Y|X}(x) - \phi^{Y|X}(x))^2}{f^X(x) f_n^X(x)}dx \\
\leq & \int E\frac{(f^X(x) + f_n^X(x)) \cdot (\phi_n^{Y|X}(x) - \phi^{Y|X}(x))^2}{f^X(x) f_n^X(x)}dx \\
= & \int E\frac{(\phi_n^{Y|X}(x) - \phi^{Y|X}(x))^2}{f_n^X(x)}dx + \int \frac{E(\phi_n^{Y|X}(x) - \phi^{Y|X}(x))^2}{f^X(x)}dx \\
= & O(\frac{1}{nh}) + O(h^2)
\end{aligned}$$

because the proof of

$$\int E\frac{(\phi_n^{Y|X}(x) - \phi^{Y|X}(x))^2}{f_n^X(x)}dx = O(\frac{1}{nh}) + O(h^2)$$

is similar to that of $E\int \frac{(f^X(x) - f_n^X(x))^2(\phi^{Y|X}(x))^2}{(f^X(x))^2 f_n^X(x)}dx$ and $\int \frac{E(\phi_n^{Y|X}(x) - \phi^{Y|X}(x))^2}{f^X(x)}dx = O(\frac{1}{nh}) + O(h^2)$ which is obtained through (17) and $Var(\phi_n^{Y|X}(x)) = O(\frac{1}{nh})$(Page 119 of Nadaraya (1989)).

Lastly for the fifth term in the sum, we have

$$\begin{aligned}
& 2E\int \frac{|f^X(x) - f_n^X(x)| \cdot |\phi_n^{Y|X}(x) - \phi^{Y|X}(x)| \cdot |\phi^{Y|X}(x)|}{f^X(x) f_n^X(x)}dx \\
& \leq 2\int \sqrt{E\frac{(f^X(x) - f_n^X(x))^2}{f_n^X(x)} \cdot E\frac{(\phi_n^{Y|X}(x) - \phi^{Y|X}(x)|)^2}{f_n^X(x)}} \cdot \frac{\phi^{Y|X}(x)}{f^X(x)}dx \\
& \leq 2\sqrt{\int E\frac{(f^X(x) - f_n^X(x))^2}{f_n^X(x)} \cdot \frac{(\phi^{Y|X}(x))^2}{(f^X(x))^2}dx \cdot \int E\frac{(\phi_n^{Y|X}(x) - \phi^{Y|X}(x)|)^2}{f_n^X(x)}dx} \\
& = O(\frac{1}{nh}) + O(h^2).
\end{aligned}$$

Combining the above expressions, we get

$$\begin{aligned}
\int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)}dx - \int \frac{(\phi_n^{Y|X}(x))^2}{f^X(x)}dx &= -\int \frac{(f_n^X(x) - Ef_n^X(x))(\phi^{Y|X}(x))^2}{(f^X(x))^2}dx \\
&\quad + O_p(\frac{1}{nh}) + O_p(h^2). \qquad (19)
\end{aligned}$$

Adding (16) and (19), we have

$$\begin{aligned}
&\int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)}dx - \int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx \\
= &\int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)}dx - \int \frac{(\phi_n^{Y|X}(x))^2}{f^X(x)}dx + \int \frac{(\phi_n^{Y|X}(x))^2}{f^X(x)}dx - \int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx \\
= &\int \frac{(\phi_n^{Y|X}(x) - E\phi_n^{Y|X}(x))^2}{f^X(x)}dx + \int \frac{2(\phi_n^{Y|X}(x) - E\phi_n^{Y|X}(x))\phi^{Y|X}(x)}{f^X(x)}dx \\
&- \int \frac{(f_n^X(x) - Ef_n^X(x)) \cdot (\phi^{Y|X}(x))^2}{(f^X(x))^2}dx + O_p(h^2) + O_p(\frac{1}{nh})
\end{aligned}$$

Similar to the proof of Theorem 4.1 of Nadaraya (1989), we have

$$\left(nh \int \frac{(\phi_n^{Y|X}(x) - E\phi_n^{Y|X}(x))^2}{f^X(x)}dx - C_1 \right) = O_p(\sqrt{h})$$

where C_1 is a constant depending on the kernel function $K(\cdot)$ and the joint density of (X, Y) under the conditions $nh^2 \to \infty$ and $nh^4 \to 0$. Then we obtain $\int \frac{(\phi_n^{Y|X}(x) - E\phi_n^{Y|X}(x))^2}{f^X(x)}dx = o_p\left(\frac{1}{\sqrt{n}}\right)$. So we have

$$\begin{aligned}
&\int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)}dx - \int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx \\
= &\underbrace{\int \frac{2(\phi_n^{Y|X}(x) - E\phi_n^{Y|X}(x))\phi^{Y|X}(x)}{f^X(x)}dx - \int \frac{(f_n^X(x) - Ef_n^X(x)) \cdot (\phi^{Y|X}(x))^2}{(f^X(x))^2}dx}_{T_n^{Y|X}} \\
&+ o_p(1/\sqrt{n}) + O_p(h^2) + O_p(\frac{1}{nh}).
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&\int \frac{(\phi_n^{X|Y}(y))^2}{f_n^Y(y)}dy - \int \frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy \\
= &\underbrace{\int \frac{2(\phi_n^{X|Y}(y) - E\phi_n^{X|Y}(y))\phi^{X|Y}(y)}{f^Y(y)}dy - \int \frac{(f_n^Y(y) - Ef_n^Y(y)) \cdot (\phi^{X|Y}(y))^2}{(f^Y(y))^2}dy}_{T_n^{X|Y}} \\
&+ o_p(1/\sqrt{n}) + O_p(h^2) + O_p(\frac{1}{nh}).
\end{aligned}$$

Proof of Theorem 3.2. By Lemma (3.1), we have

$$\begin{aligned}
&\left(\int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)}dx - \bar{Y}^2 \right) - \left(\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2 \right) \\
= &\ T_n^{Y|X} - (\bar{Y}^2 - \mu_Y^2) + o_p(1/\sqrt{n}) + O_p(h^2) + O_p(\frac{1}{nh}).
\end{aligned}$$

Then we obtain

$$\widetilde{GMC}(Y|X) - GMC(Y|X) = \frac{\int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)}dx - \bar{Y}^2}{S_Y^2 + h^2 Var_K} - \frac{\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2}{\sigma_Y^2}$$

$$= \frac{(\int \frac{(\phi_n^{Y|X}(x))^2}{f_n^X(x)}dx - \bar{Y}^2) - (\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2)}{S_Y^2 + h^2 Var_K} + \frac{\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2}{S_Y^2 + h^2 Var_K}$$

$$- \frac{\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2}{\sigma_Y^2 + h^2 Var_K} + \frac{\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2}{\sigma_Y^2 + h^2 Var_K} - \frac{\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2}{\sigma_Y^2}$$

$$= \frac{T_n^{Y|X}}{\sigma_Y^2} - \frac{\bar{Y}^2 - \mu_Y^2}{\sigma_Y^2} + \left(\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2\right)\left(\frac{1}{\sigma_Y^2 + h^2 Var_K} - \frac{1}{\sigma_Y^2}\right)$$

$$+ \left(\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2\right)\left(\frac{1}{S_Y^2 + h^2 Var_K} - \frac{1}{\sigma_Y^2 + h^2 Var_K}\right)$$

$$+ o_p(1/\sqrt{n}) + O_p(h^2) + O_p(\frac{1}{nh}).$$

Therefore,

$$\widetilde{GMC}(Y|X) - GMC(Y|X) - \left(\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2\right)\left(\frac{1}{\sigma_Y^2 + h^2 Var_K} - \frac{1}{\sigma_Y^2}\right)$$

$$= \frac{T_n^{Y|X}}{\sigma_Y^2} - \frac{\bar{Y}^2 - \mu_Y^2}{\sigma_Y^2} + \left(\int \frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx - \mu_Y^2\right)\left(\frac{1}{S_Y^2 + h^2 Var_K} - \frac{1}{\sigma_Y^2 + h^2 Var_K}\right)$$

$$+ o_p(1/\sqrt{n}) + O_p(h^2) + O_p(\frac{1}{nh}).$$

Similarly, we have

$$\widetilde{GMC}(X|Y) - GMC(X|Y) - \left(\int \frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy - \mu_X^2\right)\left(\frac{1}{\sigma_X^2 + h^2 Var_K} - \frac{1}{\sigma_X^2}\right)$$

$$= \frac{T_n^{X|Y}}{\sigma_X^2} - \frac{\bar{X}^2 - \mu_X^2}{\sigma_X^2} + \left(\int \frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy - \mu_X^2\right)\left(\frac{1}{S_X^2 + h^2 Var_K} - \frac{1}{\sigma_X^2 + h^2 Var_K}\right)$$

$$+ o_p(1/\sqrt{n}) + O_p(h^2) + O_p(\frac{1}{nh}).$$

By the central limit theorem, we have

$$\sqrt{n}\left(\frac{T_n^{Y|X}}{\sigma_Y^2}, \frac{\bar{Y} - \mu_Y}{\sigma_Y}, \frac{1}{n}\sum_{i=1}^n Y_i^2 - EY^2, \frac{T_n^{X|Y}}{\sigma_X^2}, \frac{\bar{X} - \mu_X}{\sigma_X}, \frac{1}{n}\sum_{i=1}^n X_i^2 - EX^2\right)^T \to N(\mathbf{0}_{6\times 1}, \Sigma)$$

where

$$\Sigma = Cov\left(\frac{\int \left(2Y_i - \frac{\phi^{Y|X}(x)}{f^X(x)}\right)\frac{\phi^{Y|X}(x)}{f^X(x)}\frac{1}{h}K\left(\frac{x - X_i}{h}\right)dx}{\sigma_Y^2}, \frac{Y_i}{\sigma_Y}, Y_i^2, \frac{\int \left(2X_i - \frac{\phi^{X|Y}(y)}{f^Y(y)}\right)\frac{\phi^{X|Y}(y)}{f^Y(y)}\frac{1}{h}K\left(\frac{y - Y_i}{h}\right)dy}{\sigma_X^2}, \frac{X_i}{\sigma_X}, X_i^2\right).$$

By the multivariate delta method, we have

$$\sqrt{n}\begin{pmatrix} \frac{T_n^{Y|X}}{\sigma_Y^2}-\frac{\bar{Y}^2-\mu_Y^2}{\sigma_Y^2}+(\int\frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx-\mu_Y^2)\left(\frac{1}{S_Y^2+h^2Var_K}-\frac{1}{\sigma_Y^2+h^2Var_K}\right) \\ \frac{T_n^{X|Y}}{\sigma_X^2}-\frac{\bar{X}^2-\mu_X^2}{\sigma_X^2}+(\int\frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy-\mu_X^2)\left(\frac{1}{S_X^2+h^2Var_K}-\frac{1}{\sigma_X^2+h^2Var_K}\right) \end{pmatrix}$$
$$\Longrightarrow N(\mathbf{0}_{2\times 1}, A^T\Sigma A)$$

where

$$A=\begin{pmatrix} 1 & 0 \\ -\frac{2\mu_Y}{\sigma_Y}+\frac{2\left(\int\frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx-\mu_Y^2\right)\sigma_Y\mu_Y}{(\sigma_Y^2+h^2Var_K)^2} & 0 \\ -\frac{\left(\int\frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx-\mu_Y^2\right)}{(\sigma_Y^2+h^2Var_K)^2} & 0 \\ 0 & 1 \\ 0 & -\frac{2\mu_X}{\sigma_X}+\frac{2\left(\int\frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy-\mu_X^2\right)\sigma_X\mu_X}{(\sigma_X^2+h^2Var_K)^2} \\ 0 & -\frac{\left(\int\frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy-\mu_X^2\right)}{(\sigma_X^2+h^2Var_K)^2} \end{pmatrix}.$$

Therefore,

$$\sqrt{n}\begin{pmatrix} \widetilde{GMC}(Y|X)-GMC(Y|X)-(\int\frac{(\phi^{Y|X}(x))^2}{f^X(x)}dx-\mu_Y^2)\left(\frac{1}{\sigma_Y^2+h^2Var_K}-\frac{1}{\sigma_Y^2}\right) \\ \widetilde{GMC}(X|Y)-GMC(X|Y)-(\int\frac{(\phi^{X|Y}(y))^2}{f^Y(y)}dy-\mu_X^2)\left(\frac{1}{\sigma_X^2+h^2Var_K}-\frac{1}{\sigma_X^2}\right) \end{pmatrix}$$
$$\Longrightarrow N(\mathbf{0}_{2\times 1}, A^T\Sigma A).$$

Proof of Proposition 4.1.

Case 1: The joint bivariate t density $f(t_1,t_2)$ of (T_1,T_2) is (see Siddiqui (1967))

$$\frac{\Gamma(m_1+2)(1-\rho^2)^{\frac{m_1+1}{2}}}{(2\pi)^{3/2}\Gamma(m_1+3/2)}\left[\left(1+\frac{t_1^2}{m_1}\right)\left(1+\frac{t_2^2}{m_1}\right)\right]^{-\frac{m_1+1}{2}}$$
$$\times\int(1-r^2)^{\frac{m_1-3}{2}}(1-b-cr)^{-m_1-\frac{1}{2}}F\left(\frac{1}{2},\frac{1}{2},m_1+\frac{3}{2},\frac{1+b+cr}{2}\right)dr$$

where $m_1=m-1$,

$$b=\frac{\rho t_1t_2}{m_1}\left[\left(1+\frac{t_1^2}{m_1}\right)\left(1+\frac{t_2^2}{m_1}\right)\right]^{-\frac{1}{2}},\quad c=\rho\left[\left(1+\frac{t_1^2}{m_1}\right)\left(1+\frac{t_2^2}{m_1}\right)\right]^{-\frac{1}{2}}$$

$$F(a,b;c;x)=1+\frac{ab}{c}\cdot\frac{x}{1!}+\frac{a(a+1)b(b+1)}{c(c+1)}\cdot\frac{x^2}{2!}+\cdots+\frac{\prod_{i=1}^{k}[(a+i-1)(b+i-1)]}{\prod_{i=1}^{k}(c+i-1)}\cdot\frac{x^k}{k!}+\cdots$$

Because in the joint density of (T_1, T_2), t_1 and t_2 is symmetric, we obtain that the GMCs satisfy $GMC(X|Y) = GMC(Y|X)$.

Case 2: The joint density of (T_1, T_2) for m odd and equal to $2k+3$ is

$$
\begin{aligned}
f(t_1,t_2) &= \frac{K_1}{2\pi(1-\rho^2)^{0.5}\cdot 4(k+1)} \sum_{i=0}^{k} C_k^i(-1)^{k-i}\left[\frac{1-\rho^2}{\rho}\right]^{2k-i+1}\Gamma(2k-i+1) \\
&\cdot\left\{\Gamma(i+2)\left[\frac{1}{2(1+\rho)}+\frac{t_1^2-2\rho t_1t_2+t_2^2}{8(k+1)(1-\rho^2)}\right]^{-(i+2)}\right. \\
&\left.-\sum_{j=0}^{2k-i}\left(\frac{\rho}{1-\rho^2}\right)^j\frac{\Gamma(i+j+2)}{j!}\left[\frac{1}{2(1-\rho)}+\frac{t_1^2-2\rho t_1t_2+t_2^2}{8(k+1)(1-\rho^2)}\right]^{-(i+j+2)}\right\}
\end{aligned}
$$

and for m even is

$$
f(t_1,t_2) = \sum_{i=0}^{+\infty}\frac{(m+i-1)(1-\rho^2)^{-0.5}q_i}{\pi 2^{m+i+1}(m-1)(1+\rho)^{m+i-1}}\left[\frac{1}{2(1+\rho)}+\frac{t_1^2-2\rho t_1t_2+t_2^2}{4(m-1)(1-\rho^2)}\right]^{-(m+i)}
$$

where $K_1 = \frac{1}{(\Gamma(k+1))^2(4(1-\rho^2))^{k+1}}$ and $q_i = C^i_{-0.5(m+1)}\left(\frac{2\rho}{1+\rho}\right)^i\left(\frac{1+\rho}{1-\rho}\right)^{\frac{m-1}{2}}$. Because in the joint density of (T_1, T_2), t_1 and t_2 is symmetric, we obtain that the generalized measures of correlation satisfy $GMC(X|Y) = GMC(Y|X)$.

Case 3: The bivariate t density of T_1 and T_2 is

$$
\begin{aligned}
f(t_1,t_2) &= \int_{0\le w_1\le w_2} f(t_1,t_2,w_1,w_2)dw_2dw_1 \\
&= \int_{0\le w_1\le w_2} A\sqrt{w_2}\cdot exp\left(-\frac{w_2}{2}-\frac{m_1+m_2}{2m_2}\left[\frac{t_1^2w_1}{m_1-1}-2Bt_1t_2\sqrt{w_1w_2}+\frac{t_2^2w_2}{m_1+m_2-2}\right]\right)dw_2dw_1
\end{aligned}
$$

where $A = \frac{K}{2\pi\left[\frac{m_2(m_1-1)(m_1+m_2-2)}{m_1+m_2}\right]^{1/2}}$, $B = \frac{\rho}{[(m_1-1)(m_1+m_2-2)]^{1/2}}$ and $K = \frac{1}{\Gamma\left(\frac{m_1-1}{2}\right)\Gamma\left(\frac{m_2-1}{2}\right)2^{(m_1+m_2-2)/2}}$ (see Bulgren et al. (1974)). Now, we compute

$$
E(T_1|T_2=t_2)f_{m_1+m_2-2}(t_2) = \int_{-\infty}^{+\infty} t_1 f(t_1,t_2)dt_1 = \int_{0\le w_1\le w_2}\int_{-\infty}^{+\infty} t_1 f(t_1,t_2,w_1,w_2)dt_1dw_2dw_1
$$

where

$$
\begin{aligned}
\int_0^{+\infty} t_1 f(t_1,t_2,w_1,w_2)dt_1 &= \int_0^{+\infty} t_1A\sqrt{w_2}exp\left(-\frac{w_2}{2}-\frac{m_1+m_2}{2m_2}\frac{t_2^2w_2}{m_1+m_2-2}\right) \\
&\cdot exp\left(-\frac{(m_1+m_2)w_1}{2m_2(m_1-1)}\left[t_1^2-2Bt_1t_2\frac{\sqrt{w_2}(m_1-1)}{\sqrt{w_1}}\right]\right)dt_1
\end{aligned}
$$

$$= A\sqrt{w_2}exp\left(-\frac{w_2}{2} - \frac{m_1+m_2}{2m_2}\frac{t_2^2 w_2}{m_1+m_2-2}\right)$$
$$\cdot exp\left(\frac{(m_1+m_2)B^2(m_1-1)t_2^2 w_2}{2m_2}\right) \cdot \frac{B(m_1-1)\sqrt{w_2}t_2\sqrt{2\pi}}{\sqrt{w_1}}\sqrt{\frac{m_2(m_1-1)}{(m_1+m_2)w_1}}$$
$$= exp\left(-\frac{w_2}{2} - \frac{m_1+m_2}{2m_2}\frac{t_2^2 w_2}{m_1+m_2-2} + \frac{(m_1+m_2)B^2(m_1-1)t_2^2 w_2}{2m_2}\right)$$
$$\cdot AB(m_1-1)\sqrt{2\pi}\sqrt{\frac{m_2(m_1-1)}{(m_1+m_2)}}t_2\frac{w_2}{w_1}.$$

Let $w_2 = w_2$ and $y = \frac{w_1}{w_2}$, we obtain that the Jacobian determinant is w_2, and then we get

$$\begin{aligned} h(t_2, y, w_2) &= \left.\int_{-\infty}^{+\infty} t_1 f(t_1, t_2) dt_1\right|_{y=\frac{w_1}{w_2}} \cdot w_2 \\ &= exp\left(-\frac{w_2}{2} - \frac{m_1+m_2}{2m_2}\frac{t_2^2 w_2}{m_1+m_2-2} + \frac{(m_1+m_2)B^2(m_1-1)t_2^2 w_2}{2m_2}\right) \\ &\quad \cdot AB(m_1-1)\sqrt{2\pi}\sqrt{\frac{m_2(m_1-1)}{(m_1+m_2)}}t_2 y \cdot w_2 \end{aligned}$$

Let $a = \frac{1}{2} + \frac{m_1+m_2}{2m_2}\frac{t_2^2}{m_1+m_2-2} - \frac{(m_1+m_2)B^2(m_1-1)t_2^2}{2m_2}$, we obtain

$$h(t_2, y, w_2) = AB(m_1-1)\sqrt{2\pi}\sqrt{\frac{m_2(m_1-1)}{(m_1+m_2)}}t_2 y \cdot w_2 exp\left(-a w_2\right).$$

We have

$$\begin{aligned} & E(T_1|T_2 = t_2) f_{m_1+m_2-2}(t_2) \int_0^1 \int_0^{+\infty} h(t_2, y, w_2) dw_2 dy \\ = \;& AB(m_1-1)\sqrt{2\pi}\sqrt{\frac{m_2(m_1-1)}{(m_1+m_2)}}t_2 \cdot \int_0^1 y dy \cdot \int_0^{+\infty} w_2 exp\left(-a w_2\right) dw_2 \\ = \;& AB(m_1-1)\sqrt{2\pi}\sqrt{\frac{m_2(m_1-1)}{(m_1+m_2)}}t_2 \cdot \frac{1}{2} \cdot \frac{1}{a^2} \\ = \;& AB(m_1-1)\frac{\sqrt{\pi}}{\sqrt{2}}\sqrt{\frac{m_2(m_1-1)}{(m_1+m_2)}} \cdot \frac{t_2}{\left(\frac{1}{2} + \frac{m_1+m_2}{2m_2}\frac{t_2^2}{m_1+m_2-2} - \frac{(m_1+m_2)B^2(m_1-1)t_2^2}{2m_2}\right)^2} \end{aligned}$$

where

$$f_{m_1+m_2-2} = \frac{\Gamma\left(\frac{(m_1+m_2-2)+1}{2}\right)}{\sqrt{(m_1+m_2-2)\pi}\Gamma\left(\frac{m_1+m_2-2}{2}\right)}\left(1 + \frac{t^2}{m_1+m_2-2}\right)^{-\frac{(m_1+m_2-2)+1}{2}};$$

$$
\begin{aligned}
& (E(T_1|T_2=t_2))^2 f_{m_1+m_2-2}(t_2) \\
= \;& A^2B^2(m_1-1)^2\frac{\pi}{2}\frac{m_2(m_1-1)}{(m_1+m_2)}\cdot\frac{t_2^2}{\left(\frac{1}{2}+\frac{m_1+m_2}{2m_2}\frac{t_2^2}{m_1+m_2-2}-\frac{(m_1+m_2)B^2(m_1-1)t_2^2}{2m_2}\right)^4} \\
& \cdot\frac{\sqrt{(m_1+m_2-2)\pi}\cdot\Gamma\left(\frac{m_1+m_2-2}{2}\right)}{\Gamma\left(\frac{(m_1+m_2-2)+1}{2}\right)}\left(1+\frac{t_2^2}{m_1+m_2-2}\right)^{\frac{(m_1+m_2-2)+1}{2}} \\
= \;& A^2B^2(m_1-1)^2\frac{\pi}{2}\frac{m_2(m_1-1)}{(m_1+m_2)}\frac{\sqrt{(m_1+m_2-2)\pi}\cdot\Gamma\left(\frac{m_1+m_2-2}{2}\right)}{\Gamma\left(\frac{(m_1+m_2-2)+1}{2}\right)} \\
& \cdot\frac{t_2^2\left(1+\frac{t_2^2}{m_1+m_2-2}\right)^2}{\left(\frac{1}{2}+\frac{m_1+m_2}{2m_2}\frac{t_2^2}{m_1+m_2-2}-\frac{(m_1+m_2)B^2(m_1-1)t_2^2}{2m_2}\right)^4} \\
= \;& A^2B^2(m_1-1)^2\frac{\pi}{2}\frac{m_2(m_1-1)}{(m_1+m_2)}\frac{\sqrt{(m_1+m_2-2)\pi}\cdot\Gamma\left(\frac{m_1+m_2-2}{2}\right)}{\Gamma\left(\frac{(m_1+m_2-2)+1}{2}\right)}\frac{a_2^2}{a_1^4}\cdot\frac{t_2^2\left(t_2^2+\frac{1}{a_2}\right)^2}{\left(t_2^2+\frac{1}{2a_1}\right)^4}
\end{aligned}
$$

$$
= A^2B^2(m_1-1)^2\frac{\pi}{2}\frac{m_2(m_1-1)}{(m_1+m_2)}\frac{\sqrt{(m_1+m_2-2)\pi}\cdot\Gamma\left(\frac{m_1+m_2-2}{2}\right)}{\Gamma\left(\frac{(m_1+m_2-2)+1}{2}\right)}\frac{a_2^2}{a_1^4}\cdot\frac{t_2^2\left(t_2+\frac{i}{\sqrt{a_2}}\right)^2\left(t_2-\frac{i}{\sqrt{a_2}}\right)^2}{\left(t_2+\frac{i}{\sqrt{2a_1}}\right)^4\left(t_2-\frac{i}{\sqrt{2a_1}}\right)^4}
$$

and

$$
a_1=\frac{m_1+m_2}{2m_2}\left(\frac{1}{m_1+m_2-2}-B^2(m_1-1)\right)=\frac{(m_1+m_2)(1-\rho)}{2m_2(m_1+m_2-2)},\quad a_2=\frac{1}{m_1+m_2-2}.
$$

Thus

$$
\begin{aligned}
E[\{E(T_1|T_2)\}^2] &= \int_{-\infty}^{+\infty}(E(T_1|T_2=t_2))^2 f_{m_1+m_2-2}(t_2)dt_2 \\
&= A^2B^2(m_1-1)^2\frac{\pi}{2}\frac{m_2(m_1-1)}{(m_1+m_2)}\frac{\sqrt{(m_1+m_2-2)\pi}\cdot\Gamma\left(\frac{m_1+m_2-2}{2}\right)}{\Gamma\left(\frac{(m_1+m_2-2)+1}{2}\right)}\frac{a_2^2}{a_1^4} \\
&\quad\cdot\int_{-\infty}^{+\infty}\frac{t_2^2\left(t_2+\frac{i}{\sqrt{a_2}}\right)^2\left(t_2-\frac{i}{\sqrt{a_2}}\right)^2}{\left(t_2+\frac{i}{\sqrt{2a_1}}\right)^4\left(t_2-\frac{i}{\sqrt{2a_1}}\right)^4}dt_2.
\end{aligned}
$$

Let $f(z)=\frac{\phi(z)}{\left(z-\frac{i}{\sqrt{2a_1}}\right)^4}$ where $\phi(z)=\frac{z^2\left(z+\frac{i}{\sqrt{a_2}}\right)^2\left(z-\frac{i}{\sqrt{a_2}}\right)^2}{\left(z+\frac{i}{\sqrt{2a_1}}\right)^4}$, we have

$$
\int_{-\infty}^{+\infty}\frac{t_2^2\left(t_2+\frac{i}{\sqrt{a_2}}\right)^2\left(t_2-\frac{i}{\sqrt{a_2}}\right)^2}{\left(t_2+\frac{i}{\sqrt{2a_1}}\right)^4\left(t_2-\frac{i}{\sqrt{2a_1}}\right)^4}dt_2=2\pi i\cdot Resf\left(\frac{i}{\sqrt{2a_1}}\right)=2\pi i\cdot\frac{\phi'''\left(\frac{i}{\sqrt{2a_1}}\right)}{3!}
$$

where

$$\phi'''(z) = \sum_{i=0}^{2} \sum_{j_1=0}^{min(2,3-i)} \sum_{j_2=0}^{min(2,3-i-j_1)} \frac{2!}{(2-i)!}\frac{2!}{(2-j_1)!}\frac{2!}{(2-j_2)!}\frac{(4+(3-i-j_1-j_2)-1)!}{(-1)^{(3-i-j_1-j_2)}\cdot 3!}$$
$$\cdot\frac{z^{2-i}\left(z+\frac{i}{\sqrt{a_2}}\right)^{2-j_1}\left(z-\frac{i}{\sqrt{a_2}}\right)^{2-j_2}}{\left(z+\frac{i}{\sqrt{2a_1}}\right)^{4+(3-i-j_1-j_2)}}.$$

$$\int_{-\infty}^{+\infty} \frac{t_2^2\left(t_2+\frac{i}{\sqrt{a_2}}\right)^2\left(t_2-\frac{i}{\sqrt{a_2}}\right)^2}{\left(t_2+\frac{i}{\sqrt{2a_1}}\right)^4\left(t_2-\frac{i}{\sqrt{2a_1}}\right)^4}dt_2 = \sum_{i=0}^{2} \sum_{j_1=0}^{min(2,3-i)} \sum_{j_2=0}^{min(2,3-i-j_1)} \frac{2\pi}{3!}\frac{2!}{(2-i)!}\frac{2!}{(2-j_1)!}\frac{2!}{(2-j_2)!}$$
$$\cdot\frac{(4+(3-i-j_1-j_2)-1)!}{(-1)^{(3-i-j_1-j_2)}\cdot 3!}\cdot\frac{\left(\frac{1}{\sqrt{2a_1}}\right)^{2-i}\left(\frac{1}{\sqrt{2a_1}}+\frac{1}{\sqrt{a_2}}\right)^{2-j_1}\left(\frac{1}{\sqrt{2a_1}}-\frac{1}{\sqrt{a_2}}\right)^{2-j_2}}{\left(\frac{\sqrt{2}}{\sqrt{a_1}}\right)^{4+(3-i-j_1-j_2)}}.$$

Then

$$E[\{E(T_1|T_2)\}^2] = A^2B^2(m_1-1)^2\frac{\pi}{2}\frac{m_2(m_1-1)}{(m_1+m_2)}\frac{\sqrt{(m_1+m_2-2)\pi}\cdot\Gamma\left(\frac{m_1+m_2-2}{2}\right)}{\Gamma\left(\frac{(m_1+m_2-2)+1}{2}\right)}\frac{a_2^2}{a_1^4}$$
$$\cdot\left[\sum_{i=0}^{2} \sum_{j_1=0}^{min(2,3-i)} \sum_{j_2=0}^{min(2,3-i-j_1)} \frac{2\pi}{3!}\frac{2!}{(2-i)!}\frac{2!}{(2-j_1)!}\frac{2!}{(2-j_2)!}\right.$$
$$\left.\cdot\frac{(4+(3-i-j_1-j_2)-1)!}{(-1)^{(3-i-j_1-j_2)}\cdot 3!}\cdot\frac{\left(\frac{1}{\sqrt{2a_1}}\right)^{2-i}\left(\frac{1}{\sqrt{2a_1}}+\frac{1}{\sqrt{a_2}}\right)^{2-j_1}\left(\frac{1}{\sqrt{2a_1}}-\frac{1}{\sqrt{a_2}}\right)^{2-j_2}}{\left(\frac{\sqrt{2}}{\sqrt{a_1}}\right)^{4+(3-i-j_1-j_2)}}\right].$$

Similarly, we have

$$E[\{E(T_2|T_1)\}^2] = A^2B^2(m_1+m_2-2)^2\frac{\pi}{2}\frac{m_2(m_1+m_2-2)}{(m_1+m_2)}\frac{\sqrt{(m_1-1)\pi}\cdot\Gamma\left(\frac{m_1-1}{2}\right)}{\Gamma\left(\frac{(m_1-1)+1}{2}\right)}\frac{(a_2^*)^2}{(a_1^*)^4}$$
$$\cdot\left[\sum_{i=0}^{2} \sum_{j_1=0}^{min(2,3-i)} \sum_{j_2=0}^{min(2,3-i-j_1)} \frac{2\pi}{3!}\frac{2!}{(2-i)!}\frac{2!}{(2-j_1)!}\frac{2!}{(2-j_2)!}\right.$$
$$\left.\cdot\frac{(4+(3-i-j_1-j_2)-1)!}{(-1)^{(3-i-j_1-j_2)}\cdot 3!}\cdot\frac{\left(\frac{1}{\sqrt{2a_1^*}}\right)^{2-i}\left(\frac{1}{\sqrt{2a_1^*}}+\frac{1}{\sqrt{a_2^*}}\right)^{2-j_1}\left(\frac{1}{\sqrt{2a_1^*}}-\frac{1}{\sqrt{a_2^*}}\right)^{2-j_2}}{\left(\frac{\sqrt{2}}{\sqrt{a_1^*}}\right)^{4+(3-i-j_1-j_2)}}\right]$$

where $a_1^* = (m_1+m_2)(1-\rho)/2m_2(m_1-1), \quad a_2^* = 1/(m_1-1)$.

Finally, we obtain

$$GMC(T_2|T_1) = 1 - \frac{E[\{T_2 - E(T_2|T_1)\}^2]}{Var(T_2)} = \frac{E[\{E(T_2|T_1)\}^2]}{Var(T_2)}$$

$$GMC(T_1|T_2) = 1 - \frac{E[\{T_1 - E(T_1|T_2)\}^2]}{Var(T_1)} = \frac{E[\{E(T_1|T_2)\}^2]}{Var(T_1)}$$

where $Var(T_1) = 2(m_1 - 1), Var(T_2) = 2(m_1 + m_2 - 2)$. Dividing the left-hands of the two equations above, we get

$$\frac{GMC(T_1|T_2)}{GMC(T_2|T_1)} = \frac{\Gamma(\frac{m_1+m_2-2}{2}) \cdot \Gamma(\frac{m_1}{2})}{\Gamma(\frac{m_1+m_2-1}{2}) \cdot \Gamma(\frac{m_1-1}{2})}$$

which shows that $GMC(T_1|T_2) \neq GMC(T_2|T_1)$ when $m_1 > 2, m_2 \geq 2$.

Proof of Proposition 4.2. Let $a = 1 - \alpha_1$ and $b = \alpha_2$. Then for $0 < x < 1$, $0 < y < 1$, we have

$$X \leq x, Y \leq y \Longleftrightarrow$$

$$U_1 \leq min\left\{exp\left(\frac{\log x}{1+\left(\frac{1-a}{a}\right)^\beta}\right), exp\left(\frac{\log y}{1+\left(\frac{1-b}{b}\right)^\beta}\right)\right\}$$

$$U_2 \leq min\left\{exp\left(\frac{\log x}{1+\left(\frac{a}{1-a}\right)^\beta}\right), exp\left(\frac{\log y}{1+\left(\frac{b}{1-b}\right)^\beta}\right)\right\}.$$

which gives

$$\begin{aligned} P(X \leq x, Y \leq y) &= \min\left\{exp\left(\frac{\log x}{1+\left(\frac{1-a}{a}\right)^\beta}\right), exp\left(\frac{\log y}{1+\left(\frac{1-b}{b}\right)^\beta}\right)\right\} \\ &\quad \cdot \min\left\{exp\left(\frac{\log x}{1+\left(\frac{a}{1-a}\right)^\beta}\right), exp\left(\frac{\log y}{1+\left(\frac{b}{1-b}\right)^\beta}\right)\right\} \end{aligned}$$

where $0 \leq x, y \leq 1$. Let $f_1(x) = \frac{x^\beta}{x^\beta+(1-x)^\beta}$ monotonically increasing and $f_2(x) = \frac{(1-x)^\beta}{x^\beta+(1-x)^\beta}$ monotonically decreasing. Because $\alpha_1 + \alpha_2 < 1$, we have $a > b$, and $x^{f_1(a)/f_1(b)} < x^{f_2(a)/f_2(b)}$. The joint distribution of (X, Y) is given by

$$P(X \leq x, Y \leq y) = \begin{cases} x^{f_1(a)} y^{f_2(b)}, & x^{f_1(a)/f_1(b)} < y < x^{f_2(a)/f_2(b)} \\ y, & y \leq x^{f_1(a)/f_1(b)} \\ x, & y \geq x^{f_2(a)/f_2(b)} \end{cases}$$

The conditional density $f)Y|X(y|x)$ has three cases.

Case 1: When $x^{f_1(a)/f_1(b)} < y < x^{f_2(a)/f_2(b)}$, we have

$$f_{Y|X}(y|x) = f_{X|Y}(x|y) = f_1(a) f_2(b) x^{f_1(a)-1} y^{f_2(b)-1}$$

Case 2: When $y \leq x^{f_1(a)/f_1(b)}$, we have

$$P(Y \leq x^{f_1(a)/f_1(b)} | x - \Delta \leq X \leq x) = \frac{x^{f_1(a)f_2(b)/f_1(b)} \left[x^{f_1(a)} - (x-\Delta)^{f_1(a)}\right]}{\Delta}$$

That is, $P(Y \leq x^{f_1(a)/f_1(b)}|X = x) = f_1(a)x^{\frac{f_1(a)}{f_1(b)}-1}$. When $y < x^{f_1(a)/f_1(b)}$, we have $P(Y \leq y|x - \Delta \leq X \leq x) = 0$, and $P(Y \leq y|X = x) = 0$. Then,

$$P(Y = x^{f_1(a)/f_1(b)}|X = x) = f_1(a)x^{\frac{f_1(a)}{f_1(b)}-1}.$$

Similarly, $P(X = y^{f_2(b)/f_2(a)}|Y = y) = f_2(b)y^{\frac{f_2(b)}{f_2(a)}-1}$.

Case 3. When $y \geq x^{f_2(a)/f_2(b)}$, we have

$$P(Y \leq y|x \leq X \leq x + \Delta) = 1 \to P(Y \leq y|X = x) = 1,$$
$$P(Y \leq x^{f_2(a)/f_2(b)}|x < X \leq x + \Delta) = \frac{x^{f_2(a)}(x + \Delta)^{f_1(a)} - x}{\Delta},$$
$$P(Y \leq x^{f_2(a)/f_2(b)}|X = x) = f_1(a).$$

Then $P(Y = x^{f_2(a)/f_2(b)}|X = x) = 1 - f_1(a) = f_2(a)$, $P(X = y^{f_1(b)/f_1(a)}|Y = y) = f_1(b)$. Now we compute

$$\begin{aligned} E(Y|X = x) &= \int_{x^{f_1(a)/f_1(b)}}^{x^{f_2(a)/f_2(b)}} y f_1(a) f_2(b) x^{f_1(a)-1} y^{f_2(b)-1} dy + f_1(a)x^{2f_1(a)/f_1(b)-1} + f_2(a)x^{f_2(a)/f_2(b)} \\ &= \tfrac{f_1(a)f_2(b)}{f_2(b)+1} x^{f_1(a)-1+f_2(a)(f_2(b)+1)/f_2(b)} - \tfrac{f_1(a)f_2(b)}{f_2(b)+1} x^{f_1(a)-1+f_1(a)(f_2(b)+1)/f_1(b)} \\ &\quad + f_1(a)x^{2f_1(a)/f_1(b)-1} + f_2(a)x^{f_2(a)/f_2(b)} \\ &= \tfrac{f_1(a)f_2(b)}{f_2(b)+1} x^{\frac{f_2(a)}{f_2(b)}} - \tfrac{f_1(a)f_2(b)}{f_2(b)+1} x^{\frac{2f_1(a)}{f_1(b)}-1} + f_1(a)x^{\frac{2f_1(a)}{f_1(b)}-1} + f_2(a)x^{\frac{f_2(a)}{f_2(b)}} \\ &= \tfrac{f_2(a)+f_2(b)}{f_2(b)+1} x^{\frac{f_2(a)+f_2(b)}{f_2(b)}-1} + \tfrac{f_1(a)}{f_2(b)+1} x^{\frac{2f_1(a)}{f_1(b)}-1} \\ &= \tfrac{(f_2(a)+f_2(b))}{f_2(b)+1} x^{\frac{2f_2(a)+f_2(b)}{2f_2(b)}-0.5} + \tfrac{f_1(a)}{f_2(b)+1} x^{\frac{4f_1(a)-f_1(b)}{2f_1(b)}-0.5}. \end{aligned}$$

and

$$\begin{aligned} E[\{E(Y|X)\}^2] &= \frac{(f_2(a) + f_2(b))^2}{(f_2(b) + 1)^2} \frac{f_2(b)}{2f_2(a) + f_2(b)} + \frac{f_1^2(a)}{(f_2(b) + 1)^2} \frac{f_1(b)}{4f_1(a) - f_1(b)} \\ &\quad + \frac{2(f_2(a) + f_2(b))f_1(a)}{(f_2(b) + 1)^2} \frac{1}{\frac{2f_2(a)+f_2(b)}{2f_2(b)} + \frac{4f_1(a)-f_1(b)}{2f_1(b)}} \\ &= \frac{(f_2(a) + f_2(b))^2}{(f_2(b) + 1)^2} \frac{f_2(b)}{2f_2(a) + f_2(b)} + \frac{f_1^2(a)}{(f_2(b) + 1)^2} \frac{f_1(b)}{4f_1(a) - f_1(b)} \\ &\quad + \frac{2(f_2(a) + f_2(b))f_1(a)f_1(b)f_2(b)}{(f_2(b) + 1)^2(2f_1(a)f_2(b) + f_2(a)f_1(b))}. \end{aligned}$$

From (3), we get $GMC(Y|X) = 12 \cdot E[\{E(Y|X)\}^2] - 3$. Similarly we have $GMC(X|Y) = 12 \cdot E[\{E(X|Y)\}^2] - 3$ where

$$E[\{E(X|Y)\}^2] = \frac{(f_1(b) + f_1(a))^2}{(f_1(a) + 1)^2} \frac{f_1(a)}{2f_1(b) + f_1(a)} + \frac{f_2^2(b)}{(f_1(a) + 1)^2} \frac{f_2(a)}{4f_2(b) - f_2(a)}$$

$$+\frac{2(f_1(b)+f_1(a))f_2(b)f_2(a)f_1(a)}{(f_1(a)+1)^2(2f_2(b)f_1(a)+f_1(b)f_2(a))}.$$

References

[1] Asher, D. L. (1996). What became of the Japanese "miracle." - economic development in Japan - economic myths explained *ORBIS*, Spring, 1996.

[2] Bergsten, C. F. (1982). What to do about the US -Japan economic conflict? *Foreign Affairs*, **60**, Suumer 6.

[3] Bulgren, W. G., Dykstra, R. L. and Hewett, J. E. (1974). A bivariate t-distribution with applications. *Journal of the American Statistical Association*, **69**(346), 525-532.

[4] Drouet-Mari, D. and Kotz, S. (2001). *Correaltion and dependence*. Imperial College Press.

[5] Fan, J. and Yao, Q. (1998). Efficient estimation of conditional variance functions in stochastic regression. *Biometrika* **85**, 645-660.

[6] Granger, C. (1969). Investigating causal relations by econometric methods and cross-spectral methods. *Econometrica* **34**, 424-438.

[7] Griswold, D. T. (1998). America's maligned and misunderstood trade deficit. USA Today, November 1998.

[8] Hansen, B. E. (2004). Nonparametric conditional density estimation. Unpublished manuscript,

[9] Hong, Y., Tu, J. and Zhou, G. (2007). Asymmetric correlation of stock returns: Statistical tests and economic evaluation. *The Review of Financial Studies*, **20**, 1547-1581.

[10] Joe, H. (1997) *M*ultivariate models and dependence concepts, Chapman & Hall.

[11] Nadaraya, E. A. (1989). *Nonparametric estimation of probability densities and regression curves*, Kluwer Academic Publishers.

[12] Little, R. J. A. and Rubin, D. B. (1987). *Statistical Analysis with Missing Data*, Wiley series in probability and mathematical statistics. Applied probability and statistics, Wiley, New York.

[13] Liu, J. S. (1994). Fraction of missing information and convergence rate of data augmentation, in J.Small and A.Lehman(eds), *Computationally Intensive Statistical Methods: Proceedings of the 26th symposium on the Interface*, Vol. 26 of Computing Science and Statistics, Interface Foundation of North America, North Carolina, PP. 490-497.

[14] O'Grady, K. E. (1982). Measures of Explained Variance: Cautions and Limitations. *Psyclogical Bulletin*, **92**, 766-777.

[15] Ozer, D. J. (1985). Correlation and Coefficient of Determination. *Psyclogical Bulletin*, **97**, 307-315.

[16] Patil, S. A. and Liao, S. H. (1970). The distribution of the ratios of means to the square root of the sum of variances of a bivariate normal sample. *The Annals of Mathematical Statistics*, **41**(2), 723-728.

[17] Rodgers, J. L. and Nicewander, W. A. (1988). Thirteen ways to look at the correlation coefficient. *American Statistician*, **42**, 59-66.

[18] Siddiqui, M. M. (1967). A bivariate t distribution. *The Annals of Mathematical Statistics*, **38**(1), 162-166.

[19] Zhang, Z. (2008). Quotient correlation: A sample based alternative to Pearson's correlation. *Annals of Statistics*. **36**, 1007–1030.

[20] Zhang, Z. (2009), On approximating max-stable processes and constructing extremal copula functions. *Statistical Inference for Stochastic Processes*, **12**, 89-114.

[21] Zhang, Z., Qi, Y., and Ma, X. (2010), Asymptotic independence of correlation coefficients. Manuscript.

[22] Zhang, Z. and Shinki, K. (2006), Extreme co-movements and extreme impacts in high frequency data in finance, *Journal of Banking and Finance*, 31, 1399-1415.