Asymptotic Mean and Variance of Gini Correlation for Bivariate Normal Samples

Weichao Xu, *Member, IEEE*, Y. S. Hung, *Senior Member, IEEE*, Mahesan Niranjan, *Member, IEEE*, and Minfen Shen, *Member, IEEE*

Abstract—This paper derives the asymptotic analytical forms of the mean and variance of the Gini correlation (GC) with respect to samples drawn from bivariate normal populations. The asymptotic relative efficiency (ARE) of the Gini correlation to Pearson's product moment correlation coefficient (PPMCC) is investigated under the normal assumptions. To gain further insight into GC, we also compare the Gini correlation to other two closely related correlation coefficients, namely, the order statistics correlation coefficient (OSCC) and Spearman's rho (SR). Theoretical and simulation results suggest that the performance of GC lies in between those of OSCC and SR when estimating the correlation coefficient of the bivariate normal population. The newly found theoretical results along with other desirable properties enable GC to be a useful alternative to the existing coefficients, especially when one wants to make a trade-off between the efficiency and robustness to monotone nonlinearity.

Index Terms—Asymptotic relative efficiency (ARE), bivariate normal, correlation theory, concomitant, delta method, Gini correlation (GC), order statistics correlation coefficient (OSCC), Pearson's product moment correlation coefficient (PPMCC), quadrivariate normal, ranks, Spearman's rho (SR), symmetric Gini correlation (SGC), trivariate normal.

I. INTRODUCTION

CORRELATION analysis is an important and active theme of research prevailing in many scientific and engineering areas, including nearly all branches of modern signal processing [1]–[11]. Correlation is often interpreted as the strength of statistical relationship between two random variables obeying a joint probability distribution [12]. As a measure of such strength, correlation should be large and positive if there is a high probability that large (small) values of one variable occur in conjunction with large (small) values of another. And it should be large and negative if the direction is reversed [13]. There is plenty of methods proposed and applied in the literature to quantify the correlation between two random

variables. Among these methods the Pearson's product moment correlation coefficient (PPMCC) [14], [15], Spearman's rho (SR) [12] and Kendall's tau [12] are perhaps the most widely used [16]. Besides these three classical correlation coefficients, various estimators based on M-estimation [17] and order statistics [18]–[20] among others, have been proposed in the literature.

There are many strengths and weaknesses of the correlation coefficients mentioned above. The PPMCC, which utilizes the whole information contained in the variates, is optimal when measuring the correlation between bivariate normal variables (binormal model) [21]. However, it might performs poorly when the data is attenuated by nonlinear transformations. On the other hand, the two rank correlation coefficients, Spearman's rho and Kendall's tau, are not as efficient as the PPMCC under the binormal model; nevertheless they are invariant under increasing monotone transformations, thus often considered as robust alternatives to the PPMCC when the data deviates from binormal model. Despite their robustness and stability in non-normal cases, the M-estimator-based correlation coefficients suffer great losses (up to 63%) of asymptotic relative efficiency (ARE) to the PPMCC for normal samples [17]. Such heavy loss of efficiency might not be compensated by their robustness in practice [17]. Recently, Xu *et al.* proposed a new estimator, called order statistics correlation coefficient (OSCC), which bridges the gap between the PPMCC and SR [18]–[20]. While possessing certain extent of robustness, the OSCC might perform worse than SR for data strongly distorted by increasing nonlinearity [19].

In addition to the limitations just mentioned, nearly all the above methods fail to handle satisfactorily the two special cases: 1) the data is only half complete, that is, one variable is ordinal while the other is cardinal and 2) the data is complete, but the distribution of one variable is highly skewed compared to the other. The conventional strategy with such kinds of data would be to rank the cardinal variable(s) and use SR, but while doing this we unavoidably lose information [22]. To deal with these two special cases, an interesting correlation coefficient called *Gini correlation* (GC) has been proposed [22]. Because of its desirable properties, GC has been widely used in the field of economic data analysis since late 1980s. However, the understanding of GC is far from complete due to the lack of knowledge on its distribution for data coming from bivariate normal distributions. Such knowledge is essential when analyzing the performance of GC under the binormal model, which is the benchmark model in the literature of signal processing [23]–[25]. Aiming at exploring the potential feasibility of GC

in the context of signal processing, we derive in this paper the asymptotic analytic expressions of the mean and variance of GC when samples are drawn from a bivariate normal population with correlation ρ. We will also compare GC to three closely related coefficients, namely, the classical PPMCC and SR, and the OSCC recently proposed [18]–[20].

Our theoretical contribution in this paper is threefold. First, we find the closed form expression of the mean and variance of GC with respect to ρ, the population correlation of the parent bivariate normal distribution. This makes GC more mathematical tractable than SR, whose variance contains elliptic integrals that cannot be expressed in elementary functions. Second, we obtain the ARE of GC to PPMCC, ranging from 95.5% to 88.7% with increase of $|\rho|$ from 0 to 1. Third, we derive the exact ARE of SR to PPMCC when $|\rho|$ tends to 1, being 69.5%. As indicated by these theoretical results, GC is not optimal when estimating the population correlation ρ. However, its minor loss of efficiency ($\leq 11.3\%$) is well offset by the advantageous properties depicted in the next section.

The remainder of this paper is organized as follows. Section II gives some basic definitions and the general properties concerning GC. In Section III, we derive the asymptotic expressions of the mean and variance of GC for samples drawn from bivariate Gaussian populations. In Section IV, we analyze the ARE of GC as well as those of OSCC and SR to PPMCC. Section V verifies the analytical results with Monte Carlo simulations. Finally, in Section VI we summarize our main findings and our conclusion on the Gini correlation.

II. GINI CORRELATION AND ITS GENERAL PROPERTIES

A. Definition and General Properties

Let $(X_i, Y_i), i = 1, \ldots, n$ denote n independent and identically distributed (i.i.d.) data pairs drawn from a bivariate population with continuous joint cumulative distribution function. Rearranging the data pairs in ascending order with respect to the magnitudes of X, we get a new sequence of data pairs $(X_{(i)}, Y_{[i]})$, where $X_{(1)} < \cdots < X_{(n)}$ are termed the order statistics of X and $Y_{[1]}, \ldots, Y_{[n]}$ the associated concomitants [26]–[28]. Reversing the roles of X and Y, we can also obtain $Y_{(1)}, \ldots, Y_{(n)}$ and the associated $X_{[1]}, \ldots, X_{[n]}$. As proposed in [22], two Gini correlations with respect to (X_i, Y_i) are defined as

$$r_G(Y, X) \triangleq \frac{\frac{1}{n(n-1)} \sum_{i=1}^{n} (2i - 1 - n) Y_{[i]}}{\frac{1}{n(n-1)} \sum_{i=1}^{n} (2i - 1 - n) Y_{(i)}} \quad (1)$$

and

$$r_G(X, Y) \triangleq \frac{\frac{1}{n(n-1)} \sum_{i=1}^{n} (2i - 1 - n) X_{[i]}}{\frac{1}{n(n-1)} \sum_{i=1}^{n} (2i - 1 - n) X_{(i)}}. \quad (2)$$

It can be shown that GC possesses the following general properties [22]:

1) $r_G(Y, X) \in [-1, 1]$ for all (X, Y);
2) $r_G(Y, X) = r_G(X, Y) = \pm 1$ if Y is a monotone increasing (decreasing) function of X;
3) the expectations of $r_G(Y, X)$ and $r_G(X, Y)$ equal zero if X and Y are independent;

4) $r_G(+, +) = -r_G(-, +) = -r_G(+, -) = r_G(-, -)$ for both $r_G(Y, X)$ and $r_G(X, Y)$;
5) $r_G(Y, X)$ is invariant under all strictly monotone transformations of X;
6) $r_G(Y, X)$ is scale and shift invariant with respect to both X and Y;
7) $r_G(Y, X)$ converges to normal distribution when the sample size n is sufficiently large.

Note that in general r_G is not symmetric, that is, $r_G(Y, X) \neq r_G(X, Y)$. Such asymmetry violates the axioms of correlation measurement [13], [16]. However, this problem can be easily solved with a revised version

$$r_G{}^*(X, Y) \triangleq \frac{1}{2} [r_G(Y, X) + r_G(X, Y)] \quad (3)$$

if symmetry is a critical feature in practice. Here and elsewhere, $r_G{}^*(X, Y)$ is termed as *symmetric Gini correlation* (SGC). For simplicity, the arguments of $r_G(Y, X)$ and $r_G{}^*(X, Y)$ will be omitted in the following discussion unless ambiguity occurs.

B. Relations With Other Correlation Coefficients

It can be shown that GC is closely related to two other coefficients, namely, OSCC and SR. As demonstrated below, these three coefficients can be linked together within a general framework.

Let (X_i, Y_i) be defined as before. Suppose that X_j is at the kth position in the sorted sequence $X_{(1)} < \cdots < X_{(n)}$, the number $1 \leq k \leq n$ is termed the rank of X_j and is denoted by P_j. Similarly we can get the rank of Y_j which is denoted by Q_j [12]. Write $P = \{P_1, \ldots, P_n\}$ and $Q = \{Q_1, \ldots, Q_n\}$. Denote $(\mathcal{X}_i, \mathcal{Y}_i)$ as suitable quantities associated with (X_i, Y_i). These values may be variates or ranks. A generalized coefficient Γ of $(\mathcal{X}, \mathcal{Y})$ can be defined by

$$\Gamma(\mathcal{X}, \mathcal{Y}) \triangleq \frac{\sum_{i=1}^{n} \left[\mathcal{X}_{(i)} - \mathcal{X}_{(n-i+1)} \right] \mathcal{Y}_{[i]}}{\sum_{i=1}^{n} \left[\mathcal{X}_{(i)} - \mathcal{X}_{(n-i+1)} \right] \mathcal{Y}_{(i)}}. \quad (4)$$

This general definition of (4) covers OSCC, GC and SR as particular cases when variates and/or ranks are substituted suitably.

1) OSCC as a Particular Case: Replacing $\mathcal{X}$ with X and $\mathcal{Y}$ with Y yields

$$\Gamma(X, Y) = \frac{\sum_{i=1}^{n} \left[X_{(i)} - X_{(n-i+1)} \right] Y_{[i]}}{\sum_{i=1}^{n} \left[X_{(i)} - X_{(n-i+1)} \right] Y_{(i)}} \quad (5)$$

which is the order statistics correlation coefficient $r_X(X, Y)$ defined in [18]–[20].

2) GC as a Particular Case: By the definition of order statistics and ranks, it is obvious that $P_{(i)} = i$. Replacing $\mathcal{X}$ with P and $\mathcal{Y}$ with Y in (4) yields

$$\Gamma(P, Y) = \frac{\sum_{i=1}^{n} (2i - 1 - n) Y_{[i]}}{\sum_{i=1}^{n} (2i - 1 - n) Y_{(i)}} \quad (6)$$

which is the Gini correlation $r_G(Y, X)$ defined in (1).

3) SR as a Particular Case: Replacing $\mathcal{X}$ and $\mathcal{Y}$ with P and Q respectively in (4) gives

$$\Gamma(P, Q) = \frac{\sum_{i=1}^{n} (2i - 1 - n) Q_{[i]}}{\sum_{i=1}^{n} (2i - 1 - n) Q_{(i)}} \quad (7)$$

which has been shown to be the Spearman's rho [20] that conventionally expressed as [12]

$$r_S(X,Y) \triangleq 1 - \frac{6\sum_{i=1}^n (P_i - Q_i)^2}{n(n^2-1)}. \tag{8}$$

The general coefficient Γ thus embraces r_X, r_G, and r_S as special cases under different combinations of variates and ranks in (4). From (5)–(8) we have

$$r_G(Y,X) = r_X(P,Y)$$
$$r_G(Q,X) = r_S(X,Y)$$
$$r_X(P,Q) = r_S(X,Y) \tag{9}$$

which means that GC can be interpreted as a particular case of OSCC and SR a particular case of GC. In other words, GC lies halfway from variate-variate to rank-rank arrangement in (4) and thus plays an intermediate role between OSCC and SR.

III. ASYMPTOTIC PROPERTIES OF GINI CORRELATION IN NORMAL SAMPLES

In this section we establish the analytical results concerning the expectation and variance of GC and SGC for samples drawn from binormal populations. For brevity we use symbols $\mathcal{E}(\cdot)$, $\mathcal{V}(\cdot)$, and $\mathcal{C}(\cdot,\cdot)$ in the sequel to denote the mean, variance and covariance of (between) random variables, respectively. Symbols of *big oh* and *little oh* are employed to compare the magnitudes of two functions $u(x)$ and $v(x)$ as the argument x tends to a limit L (not necessarily finite). The notation $u(x) = O(v(x))$, $x \to L$, denotes that $|u(x)/v(x)|$ remains bounded as $x \to L$; whereas the notation $u(x) = o(v(x))$, $x \to L$, denotes that $u(x)/v(x) \to 0$ as $x \to L$ [29].

A. Auxiliary Results

Lemma 1: Assume that W_1, W_2, W_3, W_4 follow a quadrivariate normal distribution with zero mean, unity variance and correlations $\varrho_{rs} = \mathcal{C}(W_r, W_s)$ for $r, s = 1, \ldots, 4$. Write $H(t) = 1$ for $t > 0$ and $H(t) = 0$ for $t \le 0$. Then

$$\begin{aligned}
\mathcal{I} &\triangleq \mathcal{E}\{H(W_1)H(W_2)W_3 W_4\} \\
&= \frac{1}{2\pi} \frac{\varrho_{13}\varrho_{24} + \varrho_{14}\varrho_{23}}{\sqrt{1-\varrho_{12}^2}} \\
&\quad - \frac{1}{2\pi} \frac{\varrho_{12}(\varrho_{13}\varrho_{14} + \varrho_{23}\varrho_{24})}{\sqrt{1-\varrho_{12}^2}} + \varrho_{34}\left(\frac{1}{4} + \frac{\sin^{-1}\varrho_{12}}{2\pi}\right).
\end{aligned} \tag{10}$$

Lemma 2: Assume that W_1, W_2, W_3 follow a trivariate normal distribution with zero mean, unity variance and correlations $\varrho_{rs} = \mathcal{C}(W_r, W_s)$ for $r, s = 1, \ldots, 3$. Let $H(t)$ be defined as in Lemma 1. Then

$$\mathcal{K} \triangleq \mathcal{E}\{H(W_1)W_2 W_3\} = \frac{\varrho_{23}}{2}. \tag{11}$$

Lemma 3: Let $(Z_i^1, Z_i^2, Z_i^3, Z_i^4)_{i=1}^n$ be n i.i.d. quadruples drawn from a quadrivariate normal population with zero mean,

unity variance and correlations $\eta_{rs} = \mathcal{C}(Z^r, Z^s)$ for $r, s = 1, \ldots, 4$. Let $\tilde{Z}_i^r$ denote respectively the ranks of Z_i^r for $r = 1, \ldots, 4$. Put

$$U = \frac{1}{n(n-1)} \sum_{i=1}^n \left(2\tilde{Z}_i^1 - 1 - n\right) Z_i^3$$

and

$$V = \frac{1}{n(n-1)} \sum_{i=1}^n \left(2\tilde{Z}_i^2 - 1 - n\right) Z_i^4.$$

Then

$$\mathcal{C}(U,V) = \frac{2}{\pi} \frac{\Delta_1 + (n-2)\Delta_2 - (2n-3)\eta_{13}\eta_{24}}{n(n-1)} \tag{12}$$

where

$$\Delta_1 = \frac{\eta_{13}\eta_{24} + \eta_{14}\eta_{23}}{\sqrt{1-\eta_{12}^2}} - \frac{\eta_{12}(\eta_{13}\eta_{14} + \eta_{23}\eta_{24})}{\sqrt{1-\eta_{12}^2}} + \eta_{34}\sin^{-1}\eta_{12} \tag{13}$$

and

$$\Delta_2 = \frac{4\eta_{13}\eta_{24} + \eta_{14}\eta_{23}}{\sqrt{4-\eta_{12}^2}} - \frac{\eta_{12}(\eta_{13}\eta_{14} + \eta_{23}\eta_{24})}{\sqrt{4-\eta_{12}^2}} + \eta_{34}\sin^{-1}\frac{\eta_{12}}{2}. \tag{14}$$

In order to avoid distraction, we relegate the proofs of these lemmas in Appendices I to III, respectively.

B. Properties of GC and SGC in Normal Samples

Theorem 1: Let $(X_i, Y_i), i = 1, \ldots, n$ be n i.i.d. sample pairs drawn from a bivariate normal population with correlation coefficient ρ. Then

$$\begin{aligned}
\mathcal{E}(r_G) = \mathcal{E}(r_G{}^*) &= \rho - \frac{2(n-2)}{n(n-1)} \\
&\quad \times \left(\sin^{-1}\frac{\rho}{2} + \rho\sqrt{4-\rho^2} - \sqrt{3}\rho\right) \\
&\quad + \frac{\pi}{3}\frac{(n+1)\rho}{n(n-1)} + \frac{2}{n(n-1)}(\sin^{-1}\rho + \rho\sqrt{1-\rho^2}) + o(n^{-1})
\end{aligned} \tag{15}$$

$$\begin{aligned}
\mathcal{V}(r_G) &= \frac{\pi}{3}\frac{n+1}{n(n-1)} + \left[\frac{n+1}{6} + \frac{2(n-2)\sqrt{3}}{\pi}\right]\frac{2\pi\rho^2}{n(n-1)} \\
&\quad - \frac{4(n-2)\left(\rho\sin^{-1}\frac{\rho}{2} + \rho^2\sqrt{4-\rho^2}\right)}{n(n-1)} \\
&\quad - \frac{4(\rho\sin^{-1}\rho + \rho^2\sqrt{1-\rho^2})}{n(n-1)} + o(n^{-1}) \tag{16}
\end{aligned}$$

and

$$\begin{aligned}
\mathcal{V}(r_G{}^*) &= \frac{1-\rho^2}{n(n-1)}\left[\sqrt{1-\rho^2} - \rho\sin^{-1}\rho + \frac{(n+1)\pi}{6}\right] \\
&\quad + \frac{(n-2)(1-\rho^2)}{n(n-1)}\left(\frac{1-\rho^2}{\sqrt{4-\rho^2}} - \rho\sin^{-1}\frac{\rho}{2}\right) \\
&\quad + o(n^{-1}). \tag{17}
\end{aligned}$$

Proof: From Property 6 and (3), both r_G and $r_G{}^*$ are scale and shift invariant. Therefore, without loss of any generality, we assume in the sequel that (X_i, Y_i) are i.i.d. sample pairs drawn from a standard bivariate normal population with zero mean, unity variance and correlation coefficient ρ.

The first identity in the statement (15) follows readily from (3) and the assumption just made. Now we focus on the expressions of $\mathcal{E}(r_G)$, $\mathcal{V}(r_G)$, and $\mathcal{V}(r_G{}^*)$. Let U_1, V_1 denote, respectively, the numerator and denominator of (1); and U_2, V_2, the numerator and denominator of (2). Then, by the *delta method* [30], it follows that

$$\mathcal{E}(r_G) = \frac{\mathcal{E}(U_1)}{\mathcal{E}(V_1)} + \frac{\mathcal{V}(V_1)\mathcal{E}(U_1)}{\mathcal{E}^3(V_1)} - \frac{\mathcal{C}(U_1, V_1)}{\mathcal{E}^2(V_1)} + \delta_1(n) \quad (18)$$

$$\mathcal{V}(r_G) = \frac{\mathcal{E}^2(U_1)}{\mathcal{E}^2(V_1)} \left[\frac{\mathcal{V}(U_1)}{\mathcal{E}^2(U_1)} \right.$$
$$\left. + \frac{\mathcal{V}(V_1)}{\mathcal{E}^2(V_1)} - \frac{2\mathcal{C}(U_1, V_1)}{\mathcal{E}(U_1)\mathcal{E}(V_1)} \right] + \delta_2(n) \quad (19)$$

$$\mathcal{V}(r_G{}^*) = \frac{1}{4} \left[\mathcal{V}\left(\frac{U_1}{V_1}\right) + \mathcal{V}\left(\frac{U_2}{V_2}\right) + 2\mathcal{C}\left(\frac{U_1}{V_1}, \frac{U_2}{V_2}\right) \right] \quad (20)$$

where

$$\mathcal{C}\left(\frac{U_1}{V_1}, \frac{U_2}{V_2}\right)$$
$$= \frac{\mathcal{C}(U_1, U_2)}{\mathcal{E}(V_1)\mathcal{E}(V_2)} + \frac{\mathcal{E}(U_1)\mathcal{E}(U_2)\mathcal{C}(V_1, V_2)}{\mathcal{E}^2(V_1)\mathcal{E}^2(V_2)}$$
$$- \frac{\mathcal{E}(U_2)\mathcal{C}(U_1, V_2)}{\mathcal{E}(V_1)\mathcal{E}^2(V_2)} - \frac{\mathcal{E}(U_1)\mathcal{C}(V_1, U_2)}{\mathcal{E}^2(V_1)\mathcal{E}(V_2)} + \delta_3(n). \quad (21)$$

The terms $\delta_1(n)$, $\delta_2(n)$, and $\delta_3(n)$ are higher order infinitesimals to be determined later on. To evaluate (18)–(20), it suffices to find $\mathcal{E}(V_\ell)$, $\mathcal{E}(U_\ell)$, $\mathcal{V}(V_\ell)$, $\mathcal{V}(U_\ell)$, and $\mathcal{C}(U_\ell, V_\kappa)$ for $\ell, \kappa = 1, 2$. From Nair's formulas [31], we have

$$\mathcal{E}(V_1) = \mathcal{E}(V_2) = \frac{1}{\sqrt{\pi}}. \quad (22)$$

It follows that order statistics $X_{(i)}$ and their concomitants $Y_{[i]}$ satisfy the following relationship [32]:

$$Y_{[i]} = \rho X_{(i)} + z_i \quad (23)$$

where $X_{(i)}$ and z_i are mutually independent, the latter being i.i.d. normal with zero mean and variance $1 - \rho^2$. A substitution of (23) into the expression of U_1 leads to

$$U_1 = \frac{1}{n(n-1)} \left[\rho \sum (2i - 1 - n)X_{(i)} + \sum (2i - 1 - n)z_i \right]$$

yielding

$$\mathcal{E}(U_1) = \frac{\rho}{\sqrt{\pi}}. \quad (24)$$

By an analogous argument we also find

$$\mathcal{E}(U_2) = \frac{\rho}{\sqrt{\pi}}. \quad (25)$$

Before evaluating their variances and covariances, we first rewrite U_ℓ and V_ℓ, $\ell = 1, 2$, in other forms to facilitate the following derivations. Let $\mathbf{X} \triangleq [X_1, \ldots, X_n]^T$, $\mathbf{Y} \triangleq [Y_1, \ldots, Y_n]^T$, $\mathbf{X}_O \triangleq [X_{(1)}, \ldots, X_{(n)}]^T$, $\mathbf{Y}_C \triangleq$

$[Y_{[1]}, \ldots, Y_{[n]}]^T$, $\mathbf{P} \triangleq [P_1, \ldots, P_n]^T$, $\mathbf{J} \triangleq [1, \ldots, n]^T$, $\mathbf{Y}_O \triangleq [Y_{(1)}, \ldots, Y_{(n)}]^T$. Denote $\mathbb{P}$ and $\mathbb{Q}$, both of size $n \times n$, as the permutation matrices mapping $\mathbf{X}$ to $\mathbf{X}_O$ and $\mathbf{Y}$ to $\mathbf{Y}_O$, respectively. That is, $\mathbf{X}_O = \mathbb{P}\mathbf{X}$ and $\mathbf{Y}_O = \mathbb{Q}\mathbf{Y}$. Then it follows directly that $\mathbf{Y}_C = \mathbb{P}\mathbf{Y}$, $\mathbf{Y}_O = \mathbb{Q}\mathbf{Y}$ and $\mathbf{J} = \mathbb{P}\mathbf{P} = \mathbb{Q}\mathbf{Q}$. As permutation matrices, $\mathbb{P}$ and $\mathbb{Q}$ must satisfy $\mathbb{P}^T\mathbb{P} = \mathbb{I}$ and $\mathbb{Q}^T\mathbb{Q} = \mathbb{I}$, where $\mathbb{I}$ is the identity matrix [33]. Thus,

$$\sum iY_{[i]} = \mathbf{J}^T\mathbf{Y}_C = \mathbf{P}^T\mathbb{P}^T\mathbb{P}\mathbf{Y} = \mathbf{P}^T\mathbf{Y} = \sum P_iY_i$$
$$\sum iY_{(i)} = \mathbf{J}^T\mathbf{Y}_O = \mathbf{Q}^T\mathbb{Q}^T\mathbb{Q}\mathbf{Y} = \mathbf{Q}^T\mathbf{Y} = \sum Q_iY_i.$$

Similarly we also have

$$\sum iX_{[i]} = \sum Q_iX_i \text{ and } \sum iX_{(i)} = \sum P_iX_i.$$

These identities allow us to rewrite U_ℓ and V_ℓ, $\ell = 1, 2$, as

$$U_1 = \frac{1}{n(n-1)} \sum_{i=1}^{n} (2P_i - 1 - n)Y_i \quad (26)$$

$$V_1 = \frac{1}{n(n-1)} \sum_{i=1}^{n} (2Q_i - 1 - n)Y_i \quad (27)$$

$$U_2 = \frac{1}{n(n-1)} \sum_{i=1}^{n} (2Q_i - 1 - n)X_i \quad (28)$$

$$V_2 = \frac{1}{n(n-1)} \sum_{i=1}^{n} (2P_i - 1 - n)X_i. \quad (29)$$

Having (26)–(29), we can carry out $\mathcal{V}(U_\ell)$, $\mathcal{V}(V_\ell)$ and $\mathcal{C}(U_\ell, V_\kappa)$ for $\ell, \kappa = 1, 2$, $\ell \neq \kappa$, with the assistance of Lemma 3. Write $Z^1 = Z^3 = X(Y)$, $Z^2 = Z^4 = Y(X)$. Then Z^1, Z^2, Z^3, Z^4 follow a quadrivariate normal distribution with zero mean, unity variance and correlations $\eta_{12} = \eta_{14} = \eta_{23} = \eta_{34} = \rho$, $\eta_{13} = \eta_{24} = 1$. Substituting these correlations into (12)–(14) along with the identity $\mathcal{V}(U_\ell) = \mathcal{C}(U_\ell, U_\ell)$, $\ell = 1, 2$, it follows that

$$\mathcal{V}(U_\ell) = \frac{1}{n(n-1)} \left\{ \frac{n+1}{3} + \frac{2\left[(n-2)\sqrt{3} - (2n-3)\right]\rho^2}{\pi} \right\}. \quad (30)$$

In a similar way we obtain, for $\ell, \kappa = 1, 2$, $\ell \neq \kappa$

$$\mathcal{V}(V_\ell)$$
$$= \frac{1}{n(n-1)} \left[\frac{n+1}{3} + \frac{2(n-2)\sqrt{3}}{\pi} - \frac{2(2n-3)}{\pi} \right] \quad (31)$$

$$\mathcal{C}(U_\ell, V_\ell)$$
$$= \frac{2}{\pi} \frac{1}{n(n-1)} \left[\sin^{-1}\rho + \rho\sqrt{1 - \rho^2} \right.$$
$$\left. + (n-2)\left(\sin^{-1}\frac{\rho}{2} + \rho\sqrt{4 - \rho^2}\right) - (2n-3)\rho \right] \quad (32)$$

$$\mathcal{C}(U_\ell, U_\kappa)$$
$$= \frac{2}{\pi} \frac{1}{n(n-1)} \left[\sqrt{1 - \rho^2} + \rho\sin^{-1}\rho \right.$$
$$\left. + (n-2)\left(\frac{1 + 2\rho^2}{\sqrt{4 - \rho^2}} + \rho\sin^{-1}\frac{\rho}{2}\right) - (2n-3)\rho^2 \right]. \quad (33)$$

$$\mathcal{C}(V_\ell, V_\kappa)$$
$$= \frac{2}{\pi} \frac{1}{n(n-1)} \left[\sqrt{1-\rho^2} + \rho \sin^{-1} \rho \right.$$
$$\left. + (n-2)\left(\sqrt{4-\rho^2} + \rho \sin^{-1} \frac{\rho}{2} \right) - (2n-3) \right] \quad (34)$$

and

$$\mathcal{C}(V_\ell, U_\kappa)$$
$$= \frac{\rho}{n(n-1)} \left[\frac{n+1}{3} + \frac{2(n-2)\sqrt{3}}{\pi} - \frac{2(2n-3)}{\pi} \right]. \quad (35)$$

From (30)–(35), we can see that the variance and covariance terms are all of order $O(n^{-1})$ as $n \to \infty$. It then follows that $\delta_1(n) \sim o(n^{-1})$, $\delta_2(n) \sim o(n^{-1})$ and $\delta_3(n) \sim o(n^{-1})$ as n large [30].

Now we are ready to write the complete forms of $\mathcal{E}(r_G)$, $\mathcal{V}(r_G)$ and $\mathcal{V}(r_G{}^*)$ with all the necessary terms being known. On substitution of (22), (24), (25), and (30)–(32) into (18) and (19), respectively, we find that 1) $\mathcal{E}(r_G)$ is in the form to the right side of the second equal sign in (15), 2) the second statement (16) holds true concerning the expression of $\mathcal{V}(r_G)$. Putting (22), (24), (25), and (30)–(35) into (20) and (21) and tidying up, we finally arrive at (17), the third theorem statement. ∎

It is easy to verify that when n is large, (15)–(17) simplify, respectively, to

$$\left. \begin{array}{c} \mathcal{E}(r_G) \\ \mathcal{E}(r_G{}^*) \end{array} \right\} \simeq \rho - \frac{1}{n} \left[2\sin^{-1} \frac{\rho}{2} + 2\rho\sqrt{4-\rho^2} - \left(\frac{\pi}{3} + 2\sqrt{3} \right) \rho \right] \quad (36)$$

$$\mathcal{V}(r_G) \simeq \frac{1}{n} \left[\frac{\pi}{3} + \left(\frac{\pi}{3} + 4\sqrt{3} \right) \rho^2 - 4\rho \sin^{-1} \frac{\rho}{2} - 4\rho^2 \sqrt{4-\rho^2} \right] \quad (37)$$

and

$$\mathcal{V}(r_G{}^*) \simeq \frac{1-\rho^2}{n} \left(\frac{\pi}{6} - \rho \sin^{-1} \frac{\rho}{2} + \frac{1-\rho^2}{\sqrt{4-\rho^2}} \right). \quad (38)$$

C. Asymptotic Distribution

It follows from Property 7 that $r_G(Y, X)$ and $r_G(X, Y)$ converges in distribution to normality as $n \to \infty$. Therefore, $r_G{}^*$, the linear combination of $r_G(Y, X)$ and $r_G(X, Y)$, also converges in distribution to normality as $n \to \infty$. Then, given the results established in Theorem 1, we are able to assert that, for samples drawn from normal populations and $n \to \infty$, both

$$\frac{\sqrt{n}\,(r_G - \rho)}{\left[\frac{\pi}{3} + \left(\frac{\pi}{3} + 4\sqrt{3} \right) \rho^2 - 4\rho\sin^{-1} \frac{\rho}{2} - 4\rho^2\sqrt{4-\rho^2} \right]^{\frac{1}{2}}}$$

and

$$\frac{\sqrt{n}\,(r_G{}^* - \rho)}{\left[(1-\rho^2)\left(\frac{\pi}{6} - \rho\sin^{-1} \frac{\rho}{2} + \frac{1-\rho^2}{\sqrt{4-\rho^2}} \right) \right]^{\frac{1}{2}}}$$

converge in distribution to standard normal distribution.

IV. ASYMPTOTIC RELATIVE EFFICIENCY

It follows from Theorem 1 that for normal samples

$$\lim_{n\to\infty} \mathcal{E}(r_G) = \rho \quad \text{and} \quad \lim_{n\to\infty} \mathcal{E}(r_G{}^*) = \rho.$$

This means that r_G and $r_G{}^*$ are unbiased estimators of ρ when the sample size n is large. On the other hand, it is well known that Pearson's product moment correlation coefficient [14]

$$r_P(X, Y) \triangleq \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (39)$$

is another asymptotic unbiased estimator of ρ. Moreover, it has been shown that OSCC and SR can also serve as estimators of ρ. Hence, we have five asymptotic unbiased estimators of ρ, as

$$\hat{\rho}_P = r_P \quad (40)$$
$$\hat{\rho}_G = r_G \quad (41)$$
$$\hat{\rho}_G^* = r_G{}^* \quad (42)$$
$$\hat{\rho}_X = r_X \quad (43)$$
$$\hat{\rho}_S = 2\sin\left(\frac{\pi r_S}{6} \right) \quad (44)$$

where $\hat{\rho}_X$ comes from [20] and $\hat{\rho}_S$ from [12].

Given (40)–(44), we can compare their performances by means of the *asymptotic relative efficiency*, which is defined by [21]

$$\text{ARE} \triangleq \lim_{n\to\infty} \frac{\mathcal{V}(\hat{\rho}_P)}{\mathcal{V}(\hat{\rho}_{\bar{P}})} \quad (45)$$

where $\hat{\rho}_{\bar{P}}$ denotes one of the four estimators defined in (41)–(44). Note that here we employ $\hat{\rho}_P$ as a benchmark, since, for binormal samples with n large, $\hat{\rho}_P$ approaches the Cramér–Rao lower bound (CRLB), that is [21]

$$\mathcal{V}(\hat{\rho}_P) = \mathcal{V}(r_P) = \underbrace{\frac{(1-\rho^2)^2}{n}}_{\text{CRLB}} + o(n^{-1}). \quad (46)$$

Substituting (46) and (16) into (45) and letting $n \to \infty$, we have

$$\text{ARE}_G = \frac{(1-\rho^2)^2}{\frac{\pi}{3} + \left(\frac{\pi}{3} + 4\sqrt{3} \right) \rho^2 - 4\rho\sin^{-1} \frac{\rho}{2} - 4\rho^2\sqrt{4-\rho^2}}. \quad (47)$$

It follows readily by substituting $\rho = 0$ to (47) that

$$\text{ARE}_G|_{\rho=0} = \frac{3}{\pi} \simeq 0.9549. \quad (48)$$

An extra effort should be made when $\rho \to \pm 1$, since (47) becomes an indeterminate form of "0/0" in this case. Applying L'Hôpital's rule twice along with some simplifications yields

$$\lim_{\rho\to\pm 1} \text{ARE}_G = \frac{12}{\pi + 6\sqrt{3}} \simeq 0.8867. \quad (49)$$

Combining (48) and (49), it follows that $\hat{\rho}_G$ looses an efficiency of

$$1 - \text{ARE}_G(0) \simeq 4.5\%$$

and

$$1 - \text{ARE}_G(\pm 1) \simeq 11.3\%$$

for $\rho = 0$ and $\rho = \pm 1$, respectively.

Through similar steps, we find that, for $\hat{\rho}_G^*(= r_G{}^*)$,

$$\text{ARE}_G^* = \frac{(1-\rho^2)\sqrt{4-\rho^2}}{1-\rho^2 + \left(\frac{\pi}{6} - \rho\sin^{-1}\frac{\rho}{2}\right)\sqrt{4-\rho^2}} \tag{50}$$

$$\text{ARE}_G^*|_{\rho=0} = \frac{6}{3+\pi} \simeq 0.9769 \tag{51}$$

and

$$\lim_{\rho\to\pm1}\text{ARE}_G^* = \frac{12}{\pi+6\sqrt{3}} \simeq 0.8867. \tag{52}$$

It has been shown [20] that for normal samples

$$\mathcal{V}(\hat{\rho}_X) = \mathcal{V}(r_X) = \frac{(1-\rho^2)^2}{n} + o(n^{-1})$$

which means

$$\text{ARE}_X = 1. \tag{53}$$

Now we proceed to evaluate ARE_S. It follows that asymptotically [12]

$$\mathcal{E}(r_S) = \frac{6}{\pi}\sin^{-1}\frac{\rho}{2} \tag{54}$$

and [34]

$$\mathcal{V}(r_S) = \frac{1}{n}\left[1 - \frac{324}{\pi^2}\left(\sin^{-1}\frac{\rho}{2}\right)^2 + \frac{144}{\pi^2}\sum_{i=1}^{4}\alpha_i(\rho)\right] \tag{55}$$

where

$$\alpha_1(\rho) = 2\int_0^\rho \sin^{-1}\left(\frac{x}{2}\sqrt{\frac{4-x^2}{3-x^2}}\right)\frac{dx}{\sqrt{4-x^2}} \tag{56}$$

$$\alpha_2(\rho) = \int_0^\rho \sin^{-1}\left(\frac{x}{6-2x^2}\right)\frac{dx}{\sqrt{4-x^2}} \tag{57}$$

$$\alpha_3(\rho) = \int_0^\rho \sin^{-1}\left(\frac{x}{2}\sqrt{\frac{4-x^2}{4-2x^2}}\right)\frac{dx}{\sqrt{4-x^2}} \tag{58}$$

$$\alpha_4(\rho) = \frac{1}{2}\int_0^\rho \sin^{-1}\left(\frac{x^3}{8-4x^2}\right)\frac{dx}{\sqrt{4-x^2}}. \tag{59}$$

Applying the delta method to (44) yields [12]

$$\mathcal{V}(\hat{\rho}_S) = \frac{\pi^2}{9}\left(1 - \frac{\rho^2}{4}\right)\mathcal{V}(r_S) \tag{60}$$

as n being sufficiently large. Substituting (46), (54) and (60) into (45) yields

$$\text{ARE}_S = \frac{36(1-\rho^2)^2}{(4-\rho^2)\left[\pi^2 - 324\left(\sin^{-1}\frac{\rho}{2}\right)^2 + 144\sum_{i=1}^4 \alpha_i(\rho)\right]}. \tag{61}$$

Due to the elliptic integrals involved in (56)–(59), ARE_S cannot be expressed into elementary functions. That is, in general we have to conduct numerical integrations so as to calculate $\alpha_i(\rho)$ and hence ARE_S. Nevertheless, exact results can be obtained for some special cases. Substituting $\rho = 0$ into (61) yields

$$\text{ARE}_S|_{\rho=0} = \frac{9}{\pi^2} \simeq 0.9119 \tag{62}$$

which is a well known result (see [12] and references therein). However, the property of ARE_S for $|\rho|$ approaching unity remains to be addressed. With some tedious calculations, we find the following new result (Appendix IV)

$$\text{ARE}_S|_{\rho\to\pm1} = \frac{15 + 11\sqrt{5}}{57} \simeq 0.6947. \tag{63}$$

By (62) and (63), it follows that $\hat{\rho}_S$ looses an efficiency of

$$1 - \text{ARE}_S(0) \simeq 8.8\%$$

and

$$1 - \text{ARE}_S(\pm1) \simeq 30.5\%$$

for $\rho = 0$ and $\rho = \pm1$, respectively.

V. NUMERICAL RESULTS

This section purports to 1) verify the theoretical results established in previous sections, 2) compare the performance of the five estimators defined in (40)–(44) by means of the root mean-square error (RMSE) [7], [8], and 3) reveal the advantage of Gini correlation in non-Gaussian case. Throughout this section, Monte Carlo experiments are undertaken for $20 \le n \le 100$ with respect to the means and variances of r_G. A sample size of $n = 1000$ is employed when we verify and compare ARE_G, ARE_G^*, ARE_X, and ARE_S. The number of trials is set to 10^5 for reason of accuracy. All the bivariate normal samples are generated by Matlab function `mvnrnd`.

A. Verification of $\mathcal{E}(r_G)$, $\mathcal{E}(r_G{}^*)$, $\mathcal{V}(r_G)$ and $\mathcal{V}(r_G{}^*)$

Table I lists, for each of the three sample sizes corresponding to 20, 50, and 100, 1) the theoretical expectation $\mathcal{E}(r_G)$ and $\mathcal{E}(r_G{}^*)$ from (15), 2) the asymptotic version (36), and 3) the observed mean from the Monte Carlo simulations. It is easily seen that the large sample approximation (15) is almost accurate to three figures for n as small as 20. Therefore, it would be safe to use (15) for $n \ge 20$ in practice.

On the other hand, as shown in Table II, the convergence speeds of the theoretical variance formulas (16) and (37) of r_G are rather slow. It appears that (37) gives values consistently smaller than those of (16) as well as the observed values. However, the difference between (16), (37) and the simulation results are less noticeable for $n = 50$ and ignorable for $n = 100$. These similar phenomena of $\mathcal{V}(r_G{}^*)$ with respect to (17) and (38) can also be observed in Table III. Therefore, it is accurate enough to approximate $\mathcal{V}(r_G)$ and $\mathcal{V}(r_G{}^*)$ respectively by (37) and (38) for $n \ge 50$. A sample size of $n \ge 100$ can be considered large in practice.

B. Verification of ARE

Fig. 1 depicts the theoretical curves of ARE_G, ARE_X and ARE_S superimposed with corresponding simulation results for $n = 1000$. Owing to the property of even symmetry, we only plot the results with respect to positive ρ. The simulation results are generated over $0 \le \rho \le 0.95$; whereas the theoretical curves are plotted over $0 \le \rho \le 1$ with an increment $\Delta\rho = 0.05$. Results of (49) and (63) are employed to plot the theoretical curves for $\rho = 1$. It can be observed that 1) the simulations agree well with our theoretical findings in (47), (53), and (61), respectively,

TABLE I
THEORETICAL AND OBSERVED $\mathcal{E}(r_G)$ AND $\mathcal{E}(r_G{}^*)$ FOR $n = 20, 50, 100$

ρ	$n = 20$				$n = 50$				$n = 100$			
	(15)	(36)	$O_1{}^a$	$O_2{}^b$	(15)	(36)	O_1	O_2	(15)	(36)	O_1	O_2
0.0	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.1	0.097	0.098	0.097	0.097	0.099	0.099	0.099	0.099	0.100	0.100	0.099	0.100
0.2	0.195	0.195	0.196	0.196	0.198	0.198	0.198	0.198	0.199	0.199	0.199	0.199
0.3	0.293	0.293	0.293	0.294	0.297	0.297	0.298	0.296	0.299	0.299	0.299	0.299
0.4	0.391	0.392	0.390	0.391	0.397	0.397	0.396	0.397	0.398	0.398	0.399	0.398
0.5	0.490	0.491	0.489	0.490	0.496	0.496	0.496	0.496	0.498	0.498	0.498	0.498
0.6	0.590	0.590	0.590	0.589	0.596	0.596	0.596	0.596	0.598	0.598	0.598	0.598
0.7	0.691	0.691	0.690	0.691	0.696	0.696	0.696	0.696	0.698	0.698	0.698	0.698
0.8	0.792	0.793	0.792	0.791	0.797	0.797	0.797	0.797	0.799	0.799	0.799	0.799
0.9	0.895	0.896	0.895	0.895	0.898	0.898	0.898	0.898	0.899	0.899	0.899	0.899

^{a} Observed mean values of r_G.
^{b} Observed mean values of r_G^*.

TABLE II
THEORETICAL AND OBSERVED $\mathcal{V}(r_G)$ FOR $n = 20, 50, 100$

ρ	$n = 20$			$n = 50$			$n = 100$		
	(16)	(37)	Obs.	(16)	(37)	Obs.	(16)	(37)	Obs.
0.0	0.0579	0.0524	0.0565	0.0218	0.0209	0.0217	0.0107	0.0105	0.0106
0.1	0.0568	0.0514	0.0558	0.0214	0.0205	0.0211	0.0105	0.0103	0.0105
0.2	0.0536	0.0484	0.0526	0.0202	0.0194	0.0201	0.0099	0.0097	0.0099
0.3	0.0485	0.0436	0.0481	0.0182	0.0174	0.0182	0.0089	0.0087	0.0089
0.4	0.0417	0.0373	0.0422	0.0156	0.0149	0.0157	0.0076	0.0075	0.0077
0.5	0.0337	0.0300	0.0350	0.0126	0.0120	0.0127	0.0061	0.0060	0.0062
0.6	0.0250	0.0220	0.0262	0.0093	0.0088	0.0095	0.0045	0.0044	0.0046
0.7	0.0163	0.0141	0.0177	0.0060	0.0056	0.0062	0.0029	0.0028	0.0029
0.8	0.0084	0.0071	0.0095	0.0030	0.0028	0.0032	0.0015	0.0014	0.0015
0.9	0.0025	0.0020	0.0030	0.0009	0.0008	0.0010	0.0004	0.0004	0.0004

TABLE III
THEORETICAL AND OBSERVED $\mathcal{V}(r_G{}^*)$ FOR $n = 20, 50, 100$

ρ	$n = 20$			$n = 50$			$n = 100$		
	(17)	(38)	Obs.	(17)	(38)	Obs.	(17)	(38)	Obs.
0.0	0.0553	0.0512	0.0546	0.0211	0.0205	0.0211	0.0104	0.0102	0.0103
0.1	0.0542	0.0502	0.0538	0.0207	0.0201	0.0206	0.0102	0.0100	0.0102
0.2	0.0512	0.0473	0.0509	0.0195	0.0189	0.0194	0.0096	0.0095	0.0096
0.3	0.0463	0.0427	0.0466	0.0176	0.0171	0.0179	0.0087	0.0085	0.0086
0.4	0.0398	0.0366	0.0404	0.0151	0.0146	0.0151	0.0074	0.0073	0.0075
0.5	0.0322	0.0294	0.0334	0.0122	0.0118	0.0124	0.0060	0.0059	0.0060
0.6	0.0238	0.0216	0.0255	0.0090	0.0087	0.0092	0.0044	0.0043	0.0044
0.7	0.0155	0.0139	0.0172	0.0058	0.0056	0.0060	0.0028	0.0028	0.0029
0.8	0.0080	0.0070	0.0092	0.0030	0.0028	0.0031	0.0014	0.0014	0.0015
0.9	0.0024	0.0020	0.0029	0.0009	0.0008	0.0009	0.0004	0.0004	0.0004

Fig. 1. Asymptotic Relative Efficiencies of ARE_G, ARE_G and ARE_S for $n = 1000$. Monte Carlo results are plotted over $0 \leq \rho \leq 0.95$ with an increment $\Delta\rho = 0.05$. The results with respect to $\rho < 0$ are not plotted due to the property of even symmetry. Results of (49) and (63) are employed to plot the theoretical curves for $\rho = 1$. It can be seen that the simulations agree well with our theoretical findings in (47), (53), and (61).
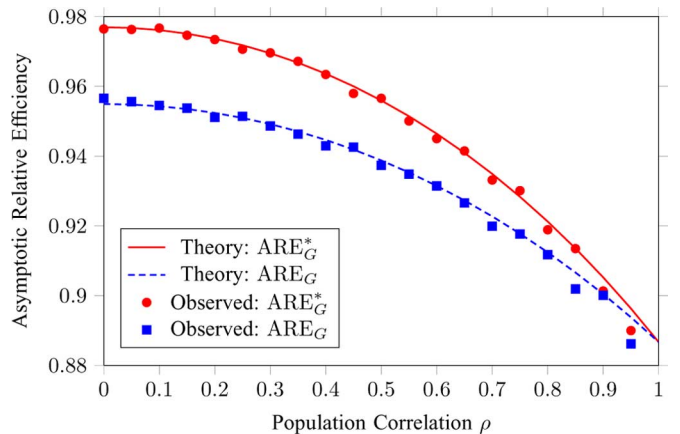
Fig. 2. Comparison of ARE_G^* and ARE_G for $n = 1000$. Simulation results are plotted over $0 \leq \rho \leq 0.95$ with an increment $\Delta\rho = 0.05$. The results with respect to $\rho < 0$ are not plotted due to the property of even symmetry. The results of (49) and (52) are utilized to plot the two theoretical curves for $\rho = 1$. Good agreements can be observed between the simulations and the theoretical results (47) and (50). The curve of ARE_G^* lies consistently above the curve of ARE_G, indicating the extra gain of efficiency resulting from the symmetrization in (3).

2) the efficiencies of $\hat{\rho}_G$ and $\hat{\rho}_S$ decrease monotonically with increase of ρ, 3) the performance of $\hat{\rho}_S$ deteriorates severely as ρ approaching unity, although it performs relatively well when ρ is small, and 4) ARE_G lies in between ARE_X and ARE_S, indicating its intermediate role as remarked before. Moreover, $\mathrm{ARE}_G \geq 90\%$ for $|\rho| \leq 0.9$. That means, equivalently, r_G looses an efficiency of less than 10% for a wide range of ρ.

Fig. 2 contrasts the difference between ARE_G and ARE_G^*. For purpose of numerical verification, simulation results are superimposed on corresponding theoretical curves. We can see that ARE_G^* lies consistently above ARE_G, indicating the extra gain of efficiency resulting from the symmetrization in (3). Such gain of efficiency decreases gradually with increase of $|\rho|$. When
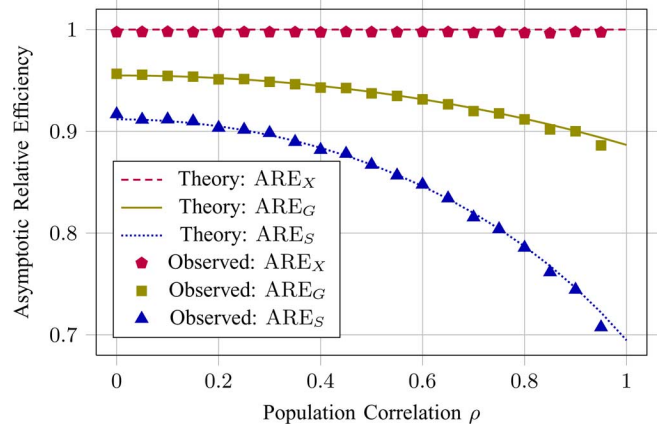
TABLE IV
RMSE FOR NORMAL SAMPLES WITH $n = 30$

ρ	CRLB	$\hat{\rho}_P$	$\hat{\rho}_X$	$\hat{\rho}_G^*$	$\hat{\rho}_G$	$\hat{\rho}_S$
0.0	0.1826	0.1848	0.1897	0.1879	0.1908	0.1931
0.1	0.1807	0.1829	0.1874	0.1861	0.1891	0.1913
0.2	0.1753	0.1804	0.1843	0.1833	0.1859	0.1885
0.3	0.1661	0.1688	0.1731	0.1722	0.1746	0.1781
0.4	0.1534	0.1583	0.1624	0.1620	0.1644	0.1684
0.5	0.1369	0.1423	0.1456	0.1458	0.1480	0.1529
0.6	0.1168	0.1239	0.1273	0.1280	0.1304	0.1365
0.7	0.0931	0.1009	0.1036	0.1046	0.1062	0.1138
0.8	0.0657	0.0728	0.0745	0.0761	0.0773	0.0858
0.9	0.0347	0.0380	0.0393	0.0410	0.0416	0.0511

TABLE V
RMSE FOR NON-NORMAL SAMPLES WITH $n = 30$

ρ	CRLB†	$\hat{\rho}_P$	$\hat{\rho}_X$	$\hat{\rho}_G^*$	$\hat{\rho}_G$	$\hat{\rho}_S$
0.0	0.1826	0.1848	0.2754	0.2745	0.1908	0.1931
0.1	0.1807	0.1912	0.2765	0.2725	0.1891	0.1913
0.2	0.1753	0.2044	0.2675	0.2611	0.1859	0.1885
0.3	0.1661	0.2270	0.2610	0.2496	0.1746	0.1781
0.4	0.1534	0.2539	0.2460	0.2308	0.1644	0.1684
0.5	0.1369	0.2848	0.2248	0.2040	0.1480	0.1529
0.6	0.1168	0.3188	0.2019	0.1727	0.1304	0.1365
0.7	0.0931	0.3521	0.1692	0.1343	0.1062	0.1138
0.8	0.0657	0.3914	0.1298	0.0959	0.0773	0.0858
0.9	0.0347	0.4303	0.0782	0.0508	0.0416	0.0511

† Not the real CRLB; reference only.

$|\rho| = 1$, both ARE_G and ARE_G^* reach their lower bound of 0.8867, as indicated by (49) and (52).

C. Comparison of RMSE for Normal Samples

From their definitions in (47), (50), (53) and (61), respectively, it is seen that the ARE of $\hat{\rho}_G$, $\hat{\rho}_G^*$, $\hat{\rho}_X$ and $\hat{\rho}_S$ to $\hat{\rho}_P$ are all independent of the sample size n. While this is convenient when measuring the performance of estimators, ARE does not take into account the bias effect that may not be ignorable for small sample size. Considering this shortcoming of ARE, we employ RMSE [7], [8], defined by

$$\mathrm{RMSE} \triangleq \sqrt{\mathcal{E}[(\hat{\rho} - \rho)^2]}$$

as a complementary performance measurement. As customary in statistical signal processing, the CRLB $(1 - \rho^2)/\sqrt{n}$ [21] is also included in the comparative study. Due to the length constraint of the paper, we only show the results with respect to $n = 30$ here.

From Table IV we can see that, when the sample size n is as small as 30, 1) for normal samples, as expected, the best is $\hat{\rho}_P$ by its minimal RMSE (closest to the CRLB), 2) the rank-based estimator $\hat{\rho}_S$, which has the maximal RMSE, performs the worst, 3) $\hat{\rho}_X$, $\hat{\rho}_G^*$ and $\hat{\rho}_G$ have comparatively moderate RMSE, indicating their intermediate behavior among all the five estimators, 4) the RMSE of $\hat{\rho}_X$ is only slightly smaller than that of $\hat{\rho}_G$, and 5) $\hat{\rho}_G^*$ even outperforms $\hat{\rho}_X$ when ρ is small (less than 0.5).

D. Comparison of RMSE for Non-Normal Samples

Thus far we have focused on the properties of GC for samples obeying bivariate normal distribution. Theoretical and simulation results suggest that, although not optimal, GC performs reasonably well not only for large sample size (in terms of ARE), but also for small sample size (in terms of RMSE). To gain further insight into the GC, now we investigate its performance for non-normal data, which is generated by performing an increasing nonlinear transformation on the X-variable of the standard binormal samples. Specifically, we first construct $X_i' = \exp\{(X_i + \mu)\sigma\}$, $i = 1, \ldots, n$, with $\mu = 10$ and $\sigma = 4$ in this study. By doing so, the joint distribution of X' and Y becomes non-Gaussian. We then compute the five estimators between X' and Y and summarize the results of RMSE in Table V. It is easily seen that, 1) the RMSE of $\hat{\rho}_P$ increases drastically

with the underlying ρ, 2) except for the null case ($\rho = 0$), the RMSE of $\hat{\rho}_G$ is the lowest, 3) like $\hat{\rho}_S$, $\hat{\rho}_G$ is invariant under the monotone transformation X' of X (Property 5), 4) the performances of $\hat{\rho}_X$ and $\hat{\rho}_G^*$ are worsened due to the effect of monotone transformation on the X-variable, and 5) the performance can be ordered as $\hat{\rho}_G > \hat{\rho}_S > \hat{\rho}_G^* > \hat{\rho}_X > \hat{\rho}_P$.

VI. CONCLUSION

In this paper we have investigated the asymptotic properties of the Gini correlation for samples drawn from bivariate normal populations. Monte Carlo simulations verified that the asymptotic results established in Theorem 1 are applicable for data of relative small sample sizes. For a better understanding of its behavior, we also compared GC with three other closely related correlation coefficients, namely, the classical Pearson's product moment correlation coefficient [14], Spearman's rho [12], and the order statistics correlation coefficient recently proposed [18]–[20]. Theoretical derivations and simulation results suggest that Gini correlation bridges the gap between the order statistics correlation coefficient and Spearman's rho. As an asymptotic unbiased estimator of the population correlation ρ, the variance of GC can be expressed in a closed form expression with respect to ρ. This makes GC more mathematically tractable than Spearman's rho, whose variance involves complex elliptic integrals that cannot be expressed in elementary functions. As indicated by the efficiency analysis, GC is not optimal when estimating the population correlation ρ. However, its loss of efficiency is confined within 4.5% to 11.3%, much lighter than that of SR, ranging from 8.8% to 30.5%.

Possessing the desirable properties summarized in Section II, Gini correlation has found wide applications in the literature of statistics and econometrics. With the mathematical tractability and high efficiency revealed in this paper, Gini correlation can play a complementary role as a measure of statistical relationship in many other research areas including statistical signal processing.

APPENDIX I
PROOF OF LEMMA 1

Proof: Let $\varphi(w, \varrho)$ be the probability density function (pdf) of the quadrivariate normal random variables

W_1, W_2, W_3, W_4. Define the associated characteristic function as [35]

$$\phi(t,\varrho) \triangleq \mathcal{E}\left\{\exp\left(\imath \sum_{j=1}^{4} t_j w_j\right)\right\}$$

$$= \exp\left\{-\frac{1}{2}\sum_{j=1}^{4} t_j^2 - \sum_{j<k}^{4} \varrho_{jk} t_j t_k\right\}$$

where $\imath$ denotes the imaginary unit $\sqrt{-1}$. Write

$$\psi(t,w,\varrho) = \frac{\phi(t,\varrho)}{16\pi^4}\exp\left\{-\imath\sum_{j=1}^{4} t_j w_j\right\}.$$

Then

$$\mathcal{I} = \iiiint_{-\infty}^{+\infty} H(w_1)H(w_2)w_3 w_4$$

$$\times \varphi(w,\varrho)dw_1 dw_2 dw_3 dw_4$$

$$= \iint_{0}^{+\infty} dw_1 dw_2 \iint_{-\infty}^{+\infty} w_3 w_4 dw_3 dw_4$$

$$\times \underbrace{\iiiint_{-\infty}^{+\infty} \psi(t,w,\varrho)dt_1 dt_2 dt_3 dt_4}_{\text{The pdf } \varphi(w,\varrho).}.$$

Write $\Omega = l + m + n + p + q + r$. An application of Taylor expansion yields [35]

$$\phi(t,\varrho) = \exp\left\{-\frac{1}{2}\sum_{j=1}^{4} t_j^2\right\}$$

$$\times \sum_{l,m,n,p,q,r=0}^{\infty} \frac{\varrho_{12}^l \varrho_{13}^m \varrho_{14}^n \varrho_{23}^p \varrho_{24}^q \varrho_{34}^r}{l!m!n!p!q!r!}$$

$$\times (-1)^{\Omega} t_1^{l+m+n} t_2^{l+p+q} t_3^{m+p+r} t_4^{n+q+r}.$$

We can then write

$$\mathcal{I} = \sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\sum_{p=0}^{\infty}\sum_{q=0}^{\infty}\sum_{r=0}^{\infty} \mathcal{A}_{lmnpqr}\varrho_{12}^l\varrho_{13}^m\varrho_{14}^n\varrho_{23}^p\varrho_{24}^q\varrho_{34}^r \tag{64}$$

where

$$\mathcal{A}_{lmnpqr} = (-1)^{\Omega}\frac{\mathcal{G}_{l+m+n}\mathcal{G}_{l+p+q}\mathcal{H}_{m+p+r}\mathcal{H}_{n+q+r}}{l!m!n!p!q!r!}$$

$$\mathcal{G}_s = \frac{1}{2\pi}\int_{0}^{+\infty} dw \int_{-\infty}^{+\infty} t^s \exp\left\{-\frac{1}{2}t^2 - \imath tw\right\} dt \tag{65}$$

and

$$\mathcal{H}_s = \frac{1}{2\pi}\int_{-\infty}^{+\infty} w\, dw \int_{-\infty}^{+\infty} t^s \exp\left\{-\frac{1}{2}t^2 - \imath tw\right\} dt. \tag{66}$$

It has been shown in [35] that

$$\mathcal{G}_s = \begin{cases} \frac{1}{2} & s = 0 \\ 0 & s = 2,4,6,\dots \\ \frac{(2k)!}{\imath\sqrt{2\pi}2^k k!} & s = 2k+1, k = 0,1,2,\dots. \end{cases} \tag{67}$$

Now we evaluate $\mathcal{H}_s$. From [35] it follows that

$$\frac{1}{2\pi}\int_{-\infty}^{+\infty} t^s \exp\left\{-\frac{t^2}{2} - \imath tw\right\} dt = \frac{1}{(-\imath)^s}\left(\frac{d}{dw}\right)^s \frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}}. \tag{68}$$

Substituting (68) into (66) and integrating by part yield

$$\mathcal{H}_s = \underbrace{\frac{1}{(-\imath)^s}\int_{-\infty}^{+\infty} w\left(\frac{d}{dw}\right)^s \frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}} dw}_{=0 \text{ for } s=0}$$

$$= \underbrace{\frac{1}{(-\imath)^s} w\left(\frac{d}{dw}\right)^{s-1}\frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\bigg|_{-\infty}^{+\infty}}_{=0 \text{ for } s\geq 1}$$

$$\underbrace{-\frac{1}{(-\imath)^s}\int_{-\infty}^{+\infty}\left(\frac{d}{dw}\right)^{s-1}\frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}} dw}_{=-\imath \text{ for } s=1}$$

$$= \underbrace{-\frac{1}{(-\imath)^s}\left(\frac{d}{dw}\right)^{s-2}\frac{e^{-\frac{w^2}{2}}}{\sqrt{2\pi}}\bigg|_{-\infty}^{+\infty}}_{=0 \text{ for } s\geq 2}$$

$$= \begin{cases} 0, & \text{for } s = 0; \\ -\imath, & \text{for } s = 1; \\ 0, & \text{for } s \geq 2. \end{cases} \tag{69}$$

It can be seen that $m + p + r$ and $n + q + r$ are suffixes of the $\mathcal{H}$-terms in (65). From (69) it follows that only $\mathcal{H}_1$ is non-null, hence $\mathcal{I}$ is non-null subject to the following conditions:

$$\begin{cases} m + p + r = 1 \\ n + q + r = 1 \\ m,n,p,q,r \geq 0. \end{cases} \tag{70}$$

It is easy to verify that there are only five solutions to system (70), as follows:
1) $m = q = 1, p = n = r = 0$;
2) $m = n = 1, p = q = r = 0$;
3) $p = q = 1, m = n = r = 0$;
4) $p = n = 1, m = q = r = 0$;
5) $r = 1, m = n = p = q = 0$.

Substituting these identities into (64) yields, respectively

$$\mathcal{I}_1 = \underbrace{\sum_{l=0}^{\infty}\frac{(-1)^{l+2}}{l!}\mathcal{G}_{l+1}^2\mathcal{H}_1^2\varrho_{12}^l}_{\mathcal{J}_1(\varrho_{12})}\varrho_{13}\varrho_{24} \tag{71}$$

$$\mathcal{I}_2 = \underbrace{\sum_{l=0}^{\infty}\frac{(-1)^{l+2}}{l!}\mathcal{G}_{l+2}\mathcal{G}_l\mathcal{H}_1^2\varrho_{12}^l}_{\mathcal{J}_2(\varrho_{12})}\varrho_{13}\varrho_{14} \tag{72}$$

$$\mathcal{I}_3 = \mathcal{J}_2(\varrho_{12})\varrho_{23}\varrho_{24} \tag{73}$$

$$\mathcal{I}_4 = \mathcal{J}_1(\varrho_{12})\varrho_{14}\varrho_{23} \tag{74}$$

and

$$\mathcal{I}_5 = \underbrace{\sum_{l=0}^{\infty} \frac{(-1)^{l+1}}{l!}\mathcal{G}_l^2 \mathcal{H}_1^2 \varrho_{12}^l}_{\mathcal{J}_3(\varrho_{12})} \varrho_{34}. \tag{75}$$

Now we evaluate $\mathcal{J}_1(\varrho_{12})$, $\mathcal{J}_2(\varrho_{12})$, and $\mathcal{J}_3(\varrho_{12})$, which are key terms to find $\mathcal{I} = \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4 + \mathcal{I}_5$. Substituting (67) and (69) into the expressions of $\mathcal{J}_1(\varrho_{12})$, $\mathcal{J}_2(\varrho_{12})$ and $\mathcal{J}_3(\varrho_{12})$ and by some simplifications, we find

$$\begin{aligned}
\mathcal{J}_1(\varrho_{12}) &= \frac{1}{2\pi}\sum_{k=0}^{\infty}\frac{(2k)!\varrho_{12}^{2k}}{2^{2k}(k!)^2} \\
&= \frac{1}{2\pi}\frac{1}{\sqrt{1-\varrho_{12}^2}}
\end{aligned} \tag{76}$$

$$\begin{aligned}
\mathcal{J}_2(\varrho_{12}) &= -\frac{1}{2\pi}\sum_{k=0}^{\infty}\frac{(2k)!\varrho_{12}^{2k+1}}{2^{2k}(k!)^2} \\
&= -\frac{1}{2\pi}\frac{\varrho_{12}}{\sqrt{1-\varrho_{12}^2}}
\end{aligned} \tag{77}$$

and

$$\begin{aligned}
\mathcal{J}_3(\varrho_{12}) &= \frac{1}{4} + \frac{1}{2\pi}\sum_{k=0}^{\infty}\frac{(2k)!\varrho_{12}^{2k+1}}{2^{2k}(k!)^2(2k+1)} \\
&= \frac{1}{4} + \frac{\sin^{-1}\varrho_{12}}{2\pi}.
\end{aligned} \tag{78}$$

A summation of (71)–(75) along with the results of (76)–(78) gives

$$\begin{aligned}
\mathcal{I} = \frac{1}{2\pi}\frac{\varrho_{13}\varrho_{24} + \varrho_{14}\varrho_{23}}{\sqrt{1-\varrho_{12}^2}} &- \frac{1}{2\pi}\frac{\varrho_{12}(\varrho_{13}\varrho_{14} + \varrho_{23}\varrho_{24})}{\sqrt{1-\varrho_{12}^2}} \\
&+ \varrho_{34}\left(\frac{1}{4} + \frac{\sin^{-1}\varrho_{12}}{2\pi}\right)
\end{aligned}$$

thus completing the proof. ∎

APPENDIX II
PROOF OF LEMMA 2

Proof: By similar procedures employed in the Proof of Lemma 1, we arrive at

$$\mathcal{K} = \sum_{l=0}^{\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\mathcal{B}_{lmn}\varrho_{12}^l\varrho_{13}^m\varrho_{23}^n$$

where

$$\mathcal{B}_{lmn} = (-1)^{l+m+n}\frac{\mathcal{G}_{l+m}\mathcal{H}_{l+n}\mathcal{H}_{m+n}}{l!m!n!}.$$

Recall that $\mathcal{G}_0 = 1/2$, $\mathcal{G}_2 = 0$ and $\mathcal{H}_1 = -\imath$. It follows that $\mathcal{K}$ is non-null only when $n = 1, l = m = 0$. Thus,

$$\mathcal{K} = \mathcal{B}_{001}\varrho_{12}^0\varrho_{13}^0\varrho_{23}^1 = \mathcal{G}_0\varrho_{23} = \frac{\varrho_{23}}{2}$$

hence the result. ∎

APPENDIX III
PROOF OF LEMMA 3

Proof: From [35] it follows that the ranks of Z_i^1 and Z_i^2 can be expanded as

$$\tilde{Z}_i^1 = \sum_{k=1}^{n} H\left(Z_i^1 - Z_k^1\right) + 1 \tag{79}$$

and

$$\tilde{Z}_i^2 = \sum_{k=1}^{n} H\left(Z_i^2 - Z_k^2\right) + 1. \tag{80}$$

Then

$$\begin{aligned}
\mathcal{C}(U,V) &= \mathcal{E}(UV) - \mathcal{E}(U)\mathcal{E}(V) \\
&= \frac{4A - 2(n-1)B + (n-1)^2C - D}{n^2(n-1)^2}
\end{aligned} \tag{81}$$

where

$$A = \sum_{i,j,k,l}^{n}\mathcal{E}\left\{H\left(Z_i^1 - Z_k^1\right)H\left(Z_j^2 - Z_l^2\right)Z_i^3 Z_j^4\right\} \tag{82}$$

$$\begin{aligned}
B &= \sum_{i,j,k}^{n}\mathcal{E}\left\{H\left(Z_i^1 - Z_k^1\right)Z_i^3 Z_j^4\right\} \\
&+ \sum_{i,j,l}^{n}\mathcal{E}\left\{H\left(Z_j^2 - Z_l^2\right)Z_i^3 Z_j^4\right\}
\end{aligned} \tag{83}$$

$$C = \sum_{i,j}^{n}\mathcal{E}\left(Z_i^3 Z_j^4\right) = n\eta_{34} \tag{84}$$

and

$$\begin{aligned}
D &= n^2(n-1)^2\mathcal{E}(U)\mathcal{E}(V) \\
&= \frac{n^2(n-1)^2\eta_{13}\eta_{24}}{\pi}.
\end{aligned} \tag{85}$$

Write $W_1 = (Z_i^1 - Z_k^1)/\sqrt{2}$, $W_2 = (Z_j^2 - Z_l^2)/\sqrt{2}$, $W_3 = Z_i^3$, and $W_4 = Z_j^4$. Then W_1, W_3, W_4 and W_2, W_3, W_4 obey respectively trivariate normal distribution with zero mean and unity variance when $i \neq k$ and $j \neq l$. According to the definition of $H(\cdot)$, it is obvious that $H(Z_i^1 - Z_k^1) = H(W_1)$ and $H(Z_j^2 - Z_l^2) = H(W_2)$. Then by Lemma 2 and the i.i.d. assumption, it follows that

$$\begin{aligned}
B &= \sum_{i\neq k,j}^{n}\mathcal{E}\{H(W_1)W_3 W_4\} \\
&+ \sum_{i,j\neq l}^{n}\mathcal{E}\{H(W_2)W_3 W_4\} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k\neq i}^{n}\frac{\mathcal{C}\left(Z_i^3, Z_j^4\right)}{2} \\
&+ \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{l\neq j}^{n}\frac{\mathcal{C}\left(Z_i^3, Z_j^4\right)}{2} \\
&= \sum_{i=1}^{n}\sum_{k\neq i}^{n}\left(\frac{\eta_{34}}{2}\right) + \sum_{j=1}^{n}\sum_{l\neq j}^{n}\left(\frac{\eta_{34}}{2}\right) \\
&= n(n-1)\eta_{34}.
\end{aligned} \tag{86}$$

TABLE VI
QUANTITIES FOR EVALUATION OF $A = \sum_{i,j,k,l}^{n} \mathcal{E}\{H(Z_i^1 - Z_k^1)H(Z_j^2 - Z_l^2)Z_i^3 Z_j^4\}$ IN LEMMA 3

Subset	No. of Terms*	Representative Term				Correlation Coeffcients between W_r and W_s					
		W_1	W_2	W_3	W_4	ϱ_{12}	ϱ_{13}	ϱ_{14}	ϱ_{23}	ϱ_{24}	ϱ_{34}
$\mathcal{D}_1 : i = j \neq k = l$	$n^{[2]}$	$(Z_1^1 - Z_2^1)/\sqrt{2}$	$(Z_1^2 - Z_2^2)/\sqrt{2}$	Z_1^3	Z_1^4	η_{12}	$\eta_{13}/\sqrt{2}$	$\eta_{14}/\sqrt{2}$	$\eta_{23}/\sqrt{2}$	$\eta_{24}/\sqrt{2}$	η_{34}
$\mathcal{D}_2 : l = i \neq j = k$	$n^{[2]}$	$(Z_1^1 - Z_2^1)/\sqrt{2}$	$(Z_2^2 - Z_1^2)/\sqrt{2}$	Z_1^3	Z_2^4	$-\eta_{12}$	$\eta_{13}/\sqrt{2}$	$-\eta_{14}/\sqrt{2}$	$-\eta_{23}/\sqrt{2}$	$\eta_{24}/\sqrt{2}$	0
$\mathcal{D}_3 : i = j \neq k \neq l$	$n^{[3]}$	$(Z_1^1 - Z_2^1)/\sqrt{2}$	$(Z_2^2 - Z_3^2)/\sqrt{2}$	Z_1^3	Z_1^4	$\eta_{12}/2$	$\eta_{13}/\sqrt{2}$	$\eta_{14}/\sqrt{2}$	$\eta_{23}/\sqrt{2}$	$\eta_{24}/\sqrt{2}$	η_{34}
$\mathcal{D}_4 : k = j \neq i \neq l$	$n^{[3]}$	$(Z_1^1 - Z_2^1)/\sqrt{2}$	$(Z_2^2 - Z_3^2)/\sqrt{2}$	Z_1^3	Z_2^4	$-\eta_{12}/2$	$\eta_{13}/\sqrt{2}$	$-\eta_{14}/\sqrt{2}$	0	$\eta_{24}/\sqrt{2}$	0
$\mathcal{D}_5 : l = i \neq j \neq k$	$n^{[3]}$	$(Z_1^1 - Z_3^1)/\sqrt{2}$	$(Z_2^2 - Z_1^2)/\sqrt{2}$	Z_1^3	Z_2^4	$-\eta_{12}/2$	$\eta_{13}/\sqrt{2}$	0	$-\eta_{23}/\sqrt{2}$	$\eta_{24}/\sqrt{2}$	0
$\mathcal{D}_6 : l = k \neq i \neq j$	$n^{[3]}$	$(Z_1^1 - Z_3^1)/\sqrt{2}$	$(Z_2^2 - Z_3^2)/\sqrt{2}$	Z_1^3	Z_2^4	$\eta_{12}/2$	$\eta_{13}/\sqrt{2}$	0	0	$\eta_{24}/\sqrt{2}$	0
$\mathcal{D}_7 : i \neq j \neq k \neq l$	$n^{[4]}$	$(Z_1^1 - Z_3^1)/\sqrt{2}$	$(Z_2^2 - Z_4^2)/\sqrt{2}$	Z_1^3	Z_2^4	0	$\eta_{13}/\sqrt{2}$	0	0	$\eta_{24}/\sqrt{2}$	0

* The symbol $n^{[m]}$ is a compact notation of $n(n-1)\cdots(n-m+1)$, with m being a positive integer.

Now we deal with A. Since $H(0) = 0$, the terms in (82) vanish when $i = k$ or $j = l$. Hence, there are $n^2(n-1)^2$ nontrivial terms left to be determined. It follows that the domain of the quadruple (i, j, k, l) can be partitioned into seven disjoint and exhaustive subsets, as shown in Table VI. For compactness, the symbol $n^{[m]}$ in Table VI and below denotes $n(n-1)\cdots(n-m+1)$, with m being a positive integer. Then (82) can be rewritten as

$$A = \sum_{j=1}^{7} A_j \qquad (87)$$

where

$$A_j = \sum_{\mathcal{D}_j} \mathcal{E}\left\{H(Z_i^1 - Z_k^1)H\left(Z_j^2 - Z_l^2\right)Z_i^3 Z_j^4\right\}$$

with $\mathcal{D}_j, j = 1, \ldots, 7$ being defined in Table VI. Write

$$\lambda_1 = \eta_{34}\left(\frac{1}{4} + \frac{\sin^{-1}\eta_{12}}{2\pi}\right)$$

and

$$\lambda_2 = \eta_{34}\left(\frac{1}{4} + \frac{\sin^{-1}\frac{\eta_{12}}{2}}{2\pi}\right).$$

Due to the i.i.d. assumption, it follows that

$$
\begin{aligned}
A_1 &= \sum_{i=j\neq k=l} \mathcal{E}\left\{H\left(Z_i^1 - Z_k^1\right)H\left(Z_j^2 - Z_l^2\right)Z_i^3 Z_j^4\right\} \\
&= n^{[2]}\mathcal{E}\left\{H\left(Z_1^1 - Z_2^1\right)H\left(Z_1^2 - Z_2^2\right)Z_1^3 Z_1^4\right\} \\
&= n^{[2]}\mathcal{E}\left\{H\left(\frac{Z_1^1 - Z_2^1}{\sqrt{2}}\right)H\left(\frac{Z_1^2 - Z_2^2}{\sqrt{2}}\right)Z_1^3 Z_1^4\right\}. \quad (88)
\end{aligned}
$$

Write $W_1 = (Z_1^1 - Z_2^1)/\sqrt{2}$, $W_2 = (Z_1^2 - Z_2^2)/\sqrt{2}$, $W_3 = Z_1^3$, and $W_4 = Z_1^4$. Then W_1, W_2, W_3, W_4 follow a quadrivariate normal distribution with zero mean, unity variance and correlation coefficients $\varrho_{12} = \eta_{12}$, $\varrho_{13} = \eta_{13}/\sqrt{2}$, $\varrho_{14} = \eta_{14}/\sqrt{2}$, $\varrho_{23} = \eta_{23}/\sqrt{2}$, $\varrho_{24} = \eta_{24}/\sqrt{2}$, and $\varrho_{34} = \eta_{34}$ (Table VI). Substituting (10), the result of Lemma 1, and the associated ϱ_{rs} into (88) yields

$$A_1 = n^{[2]}\left[\frac{\eta_{13}\eta_{24} + \eta_{14}\eta_{23}}{4\pi\sqrt{1-\eta_{12}^2}} - \frac{\eta_{12}(\eta_{13}\eta_{14} + \eta_{23}\eta_{24})}{4\pi\sqrt{1-\eta_{12}^2}} + \lambda_1\right]. \qquad (89)$$

Similarly we can obtain

$$A_2 = n^{[2]}\left[\frac{\eta_{13}\eta_{24} + \eta_{14}\eta_{23}}{4\pi\sqrt{1-\eta_{12}^2}} - \frac{\eta_{12}(\eta_{13}\eta_{14} + \eta_{23}\eta_{24})}{4\pi\sqrt{1-\eta_{12}^2}}\right] \qquad (90)$$

$$A_3 = n^{[3]}\left[\frac{\eta_{13}\eta_{24} + \eta_{14}\eta_{23}}{2\pi\sqrt{4-\eta_{12}^2}} - \frac{\eta_{12}(\eta_{13}\eta_{14} + \eta_{23}\eta_{24})}{4\pi\sqrt{4-\eta_{12}^2}} + \lambda_2\right] \qquad (91)$$

$$A_4 = n^{[3]}\left(\frac{\eta_{13}\eta_{24}}{2\pi\sqrt{4-\eta_{12}^2}} - \frac{\eta_{12}\eta_{13}\eta_{14}}{4\pi\sqrt{4-\eta_{12}^2}}\right) \qquad (92)$$

$$A_5 = n^{[3]}\left(\frac{\eta_{13}\eta_{24}}{2\pi\sqrt{4-\eta_{12}^2}} - \frac{\eta_{12}\eta_{23}\eta_{24}}{4\pi\sqrt{4-\eta_{12}^2}}\right) \qquad (93)$$

$$A_6 = \frac{n^{[3]}\eta_{13}\eta_{24}}{2\pi\sqrt{4-\eta_{12}^2}} \qquad (94)$$

$$A_7 = \frac{n^{[4]}\eta_{13}\eta_{24}}{4\pi}. \qquad (95)$$

Summing up (89)–(95) and tidying up, we arrive at

$$A = \frac{1}{2\pi}\left(n^{[2]}\Delta_1 + n^{[3]}\Delta_2 + \frac{n^{[4]}\eta_{13}\eta_{24}}{2}\right) + \frac{n(n-1)^2\eta_{34}}{4} \qquad (96)$$

where Δ_1 and Δ_2 are defined in (13) and (14), respectively. A substitution of (96), (84)–(86) into (81) along with some simplifications yields (12), thus completing the proof. ∎

APPENDIX IV
DERIVATION OF ARE_s FOR $\rho \to \pm 1$

Proof: Write

$$\beta(\rho) = \pi^2 - 324\sin^{-1}\left(\frac{\rho}{2}\right)^2 + 144\sum_{i=1}^{4}\alpha_i(\rho).$$

Let $U_3 = 36(1-\rho^2)^2$ and $V_3 = (4-\rho^2)\beta(\rho)$ be the numerator and denominator of (55), respectively. Because of the even symmetry of U_3 and V_3, it is sufficient for us to deal with the $\rho = 1$ case only in the following. From (8) we have $r_S = 1$ for sure when $\rho = 1$. That is, $\mathcal{V}(r_S)$ vanishes when ρ reaches unity. Then from (55) we have $\beta(1) = 0$. Therefore, (61) becomes an indeterminate form of 0/0 when $\rho \to 1$. It is easy to verify that

the first derivative of U_3 vanishes at $\rho = 1$. Differentiating V_3 with respect to ρ and substituting $\rho = 1$ thereafter produce

$$V_3'(1) = -2(1)\beta(1) + 3\beta'(1) = 3\beta'(1). \qquad (97)$$

It follows that

$$\beta'(\rho) = 144 \sum_{i=1}^{4} \alpha_i'(\rho) - \frac{648 \sin^{-1} \frac{\rho}{2}}{\sqrt{4 - \rho^2}}. \qquad (98)$$

From (56)–(59) we have

$$\alpha_1'(1) = \frac{2}{\sqrt{3}} \sin^{-1}\left(\frac{\sqrt{6}}{4}\right) \qquad (99)$$

$$\alpha_2'(1) = \frac{1}{\sqrt{3}} \sin^{-1}\left(\frac{1}{4}\right) \qquad (100)$$

$$\alpha_3'(1) = \frac{1}{\sqrt{3}} \sin^{-1}\left(\frac{\sqrt{6}}{4}\right) \qquad (101)$$

$$\alpha_4'(1) = \frac{1}{2\sqrt{3}} \sin^{-1}\left(\frac{1}{4}\right) \qquad (102)$$

and hence

$$\sum_{i=1}^{4} \alpha_i'(1) = \sqrt{3} \underbrace{\left[\sin^{-1}\left(\frac{\sqrt{6}}{4}\right) + \frac{1}{2} \sin^{-1}\left(\frac{1}{4}\right)\right]}_{\theta = \theta_1 + \frac{\theta_2}{2}}. \qquad (103)$$

Now we prove that $\theta = \pi/4$ by showing $\sin^2 \theta = 1/2$, as

$$\sin^2\left(\theta_1 + \frac{\theta_2}{2}\right)$$

$$= \left(\sin\theta_1 \cos\frac{\theta_2}{2} + \cos\theta_1 \sin\frac{\theta_2}{2}\right)^2$$

$$= \left(\frac{\sqrt{6}}{4}\sqrt{\frac{1}{2} + \frac{\sqrt{15}}{8}} + \frac{\sqrt{10}}{4}\sqrt{\frac{1}{2} - \frac{\sqrt{15}}{8}}\right)^2 = \frac{1}{2}.$$

Then, for $\rho = 1$, (98) turns out to be

$$\beta'(1) = 144 \times \frac{\sqrt{3}\pi}{4} - \frac{648\sqrt{3}\pi}{18} = 0. \qquad (104)$$

Because both U_3' and V_3' vanish at $\rho = 1$, we have to check their second derivatives. It follows that

$$U_3''(\rho) = 36(12\rho^2 - 4) \qquad (105)$$

and

$$V_3''(\rho) = \underbrace{-2\beta(\rho) - 4\rho\beta'(\rho)}_{=0 \text{ for } \rho = 1} + (4 - \rho^2)\beta''(\rho) \qquad (106)$$

where

$$\beta''(\rho) = 144 \sum_{i=1}^{4} \alpha_i''(\rho) - 648\left[\frac{1}{4 - \rho^2} + \frac{\rho \sin^{-1}\frac{\rho}{2}}{(4 - \rho^2)^{\frac{3}{2}}}\right]. \qquad (107)$$

A series of tedious but straightforward derivations leads to

$$\alpha_1''(1) = \frac{7\sqrt{5}}{15} + \frac{2\sqrt{3}}{9} \sin^{-1}\left(\frac{\sqrt{6}}{4}\right) \qquad (108)$$

$$\alpha_2''(1) = \frac{2\sqrt{5}}{15} + \frac{\sqrt{3}}{9} \sin^{-1}\left(\frac{1}{4}\right) \qquad (109)$$

$$\alpha_3''(1) = \frac{\sqrt{5}}{3} + \frac{\sqrt{3}}{9} \sin^{-1}\left(\frac{\sqrt{6}}{4}\right) \qquad (110)$$

$$\alpha_4''(1) = \frac{\sqrt{5}}{6} + \frac{\sqrt{3}}{18} \sin^{-1}\left(\frac{1}{4}\right). \qquad (111)$$

Summing up (108)–(111) gives

$$\sum_{i=1}^{4} \alpha_i''(1)$$

$$= \frac{11\sqrt{5}}{10} + \frac{\sqrt{3}}{3} \overbrace{\left[\sin^{-1}\left(\frac{\sqrt{6}}{4}\right) + \frac{1}{2} \sin^{-1}\left(\frac{1}{4}\right)\right]}^{\theta = \pi/4}$$

$$= \frac{11\sqrt{5}}{10} + \frac{\sqrt{3}\pi}{12}. \qquad (112)$$

Applying L'Hôpital's rule along with (105)–(107) and (112), we finally arrive at

$$\lim_{\rho \to 1} \text{ARE}_S = \frac{36(12 - 4)}{432\left(\frac{11\sqrt{5}}{10} + \frac{\sqrt{3}\pi}{12}\right) - 1944\left(\frac{1}{3} + \frac{\sqrt{3}\pi}{54}\right)}$$

$$= \frac{15 + 11\sqrt{5}}{57}$$

which also holds for $\rho \to -1$ due to the property of even symmetry. ∎

REFERENCES

[1] A. Chorti and D. T. Hristopulos, "Nonparametric identification of anisotropic (elliptic) correlations in spatially distributed data sets," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4738–4751, Oct. 2008.

[2] C. Chaux, L. Duval, A. Benazza-Benyahia, and J.-C. Pesquet, "A nonlinear stein-based estimator for multichannel image denoising," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3855–3870, Aug. 2008.

[3] R. Tao, F. Zhang, and Y. Wang, "Fractional power spectrum," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4199–4206, Sep. 2008.

[4] D. Axehill, F. Gunnarsson, and A. Hansson, "A low-complexity high-performance preprocessing algorithm for multiuser detection using gold sequences," *IEEE Trans. Signal Process.*, vol. 56, no. 9, pp. 4377–4385, Sep. 2008.

[5] M. Beko, J. Xavier, and V. Barroso, "Further results on the capacity and error probability analysis of noncoherent MIMO systems in the low SNR regime," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 2915–2930, Jul. 2008.

[6] M. Greco, F. Gini, and A. Farina, "Radar detection and classification of jamming signals belonging to a cone class," *IEEE Trans. Signal Process.*, vol. 56, no. 5, pp. 1984–1993, May 2008.

[7] F. Gini, "A radar application of a modified Cramer-Rao bound: Parameter estimation in non-Gaussian clutter," *IEEE Trans. Signal Process.*, vol. 46, no. 7, pp. 1945–1953, Jul. 1998.

[8] F. Gini and J. Michels, "Performance analysis of two covariance matrix estimators in compound-Gaussian clutter," *Proc. Inst. Electr. Eng.—F*, vol. 146, no. 3, pp. 133–140, Jun. 1999.

[9] L. Liu and M. Amin, "Performance analysis of GPS receivers in non-Gaussian noise incorporating precorrelation filter and sampling rate," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 990–1004, Mar. 2008.

[10] G. Giunta, "A note on the computational complexity of high correlation estimators," *IEEE Trans. Signal Process.*, vol. 39, no. 2, pp. 485–486, Feb. 1991.

[11] G. Jacovitti and R. Cusani, "Performance of normalized correlation estimators for complex processes," *IEEE Trans. Signal Process.*, vol. 40, no. 1, pp. 114–128, Jan. 1992.

[12] M. Kendall and J. D. Gibbons, *Rank Correlation Methods*, 5th ed. New York: Oxford Univ. Press, 1990.

[13] J. D. Gibbons and S. Chakraborti, *Nonparametric Statistical Inference*, 3rd ed. New York: Marcel Dekker, 1992.

[14] R. A. Fisher, "On the 'probable error' of a coefficient of correlation deduced from a small sample," *Metron*, vol. 1, pp. 3–32, 1921.

[15] R. A. Fisher, *Statistical Methods, Experimental Design, and Scientific Inference*. New York: Oxford Univ. Press, 1990.

[16] D. D. Mari and S. Kotz, *Correlation and Dependence*. London, U.K.: Imperial College Press, 2001.

[17] G. L. Shevlyakov and N. O. Vilchevski, *Robustness in Data Analysis: Criteria and Methods*, ser. Modern probability and statistics. Utrecht, The Netherlands: VSP, 2002.

[18] W. Xu, C. Chang, Y. S. Hung, S. K. Kwan, and P. C. W. Fung, "Order statistic correlation coefficient and its application to association measurement of biosignals," *Proc. Int. Conf. Acoustics, Speech, Signal Process. (ICASSP) 2006*, vol. 2, pp. II-1068–II-1071, 2006.

[19] W. Xu, C. Chang, Y. S. Hung, S. K. Kwan, and P. C. W. Fung, "Order statistics correlation coefficient as a novel association measurement with applications to biosignal analysis," *IEEE Trans. Signal Process.*, vol. 55, no. 12, pp. 5552–5563, Dec. 2007.

[20] W. Xu, C. Chang, Y. Hung, and P. Fung, "Asymptotic properties of order statistics correlation coefficient in the normal cases," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2239–2248, Jun. 2008.

[21] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics: Volume 2 Classical Inference and Relationship*, 5th ed. London, U.K.: Edward Arnold, 1991.

[22] E. Schechtman and S. Yitzhaki, "A measure of association base on Gini's mean difference," *Commun. Statist.-Theor. Meth.*, vol. 16, no. 1, pp. 207–231, 1987.

[23] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, ser. Prentice-Hall signal processing series. Englewood Cliffs, NJ: Prentice-Hall PTR, 1993.

[24] S. M. Kay, *Fundamentals of Statistical Signal Processing*, ser. Prentice-Hall Signal Processing Series. Englewood Cliffs, NJ: Prentice-Hall PTR, 1998.

[25] S.-C. Chan and Y.-X. Zou, "A recursive least m-estimate algorithm for robust adaptive filtering in impulsive noise: Fast algorithm and convergence performance analysis," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 975–991, Apr. 2004.

[26] H. David and H. Nagaraja, *Order Statistics*, 3rd ed. Hoboken, NJ: Wiley-Interscience, 2003.

[27] N. Balakrishnan and C. R. Rao, *Order Statistics: Applications*, ser. Handbook of statistics; v. 17. New York: Elsevier, 1998.

[28] N. Balakrishnan and C. R. Rao, *Order Statistics: Theory & Methods*, ser. Handbook of statistics; v. 16. New York: Elsevier, 1998.

[29] R. J. Serfling, *Approximation Theorems of Mathematical Statistics*, ser. Wiley Series in Probability and Mathematical Statistics. New York: Wiley, 2002.

[30] A. Stuart and J. K. Ord, *Kendall's Advanced Theory of Statistics: Volume 1 Distribution Theory*, 6th ed. London, U.K.: Edward Arnold, 1994.

[31] U. S. Nair, "The standard error of Gini's mean difference," *Biometrika*, vol. 28, no. 3/4, pp. 428–436, 1936.

[32] H. David and H. Nagaraja, *Order Statistics*, 3rd ed. Hoboken, NJ: Wiley-Interscience, 2003.

[33] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

[34] F. N. David and C. L. Mallows, "The variance of Spearman's rho in normal samples," *Biometrika*, vol. 48, pp. 19–28, 1961.

[35] P. A. P. Moran, "Rank correlation and product-moment correlation," *Biometrika*, vol. 35, no. 1/2, pp. 203–206, 1948.

Weichao Xu (M'06) received the B.Eng. and M.Eng. degrees in electrical engineering from the University of Science and Technology of China, in 1993 and 1996, respectively, and the Ph.D. degree in biomedical engineering from the University of Hong Kong in 2002.

Since 2003, he has been a Research Associate with the Department of Electrical and Electronic Engineering, the University of Hong Kong. His research interests are in the areas of mathematical statistics, machine learning, digital signal processing, and applications.

Y. S. Hung (M'88–SM'02) received the B.Sc. (Eng.) degree in electrical engineering and the B.Sc. degree in mathematics from the University of Hong Kong, Hong Kong, and the M.Phil. and Ph.D. degrees from the University of Cambridge, Cambridge, U.K.

He was a Research Associate with the University of Cambridge and a Lecturer with the University of Surrey, Surrey, U.K. In 1989, he joined the University of Hong Kong, where he is currently a Professor and the Head of the department. His research interests include robust control systems theory, robotics, computer vision, and biomedical engineering.

Prof. Hung was a recipient of the Best Teaching Award in 1991 from the Hong Kong University Students Union. He is a chartered engineer and a fellow of IET and HKIE.

Mahesan Niranjan (M'06) received the B.Sc. degree from the University of Peradeniya, Sri Lanka, in 1982, the M.E.E. degree from Eindhoven University of Technology, Eindhoven, The Netherlands, in 1985, both in electronics engineering, and the Ph.D. degree from the University of Cambridge, Cambridge, U.K., in 1990.

Previously, he held a lectureship in the University of Cambridge from 1990 to 1998, a professorship in the University of Sheffield from 1999 to 2008, where he served as Head of Computer Science from 2002 to 2004 and Dean of the Faculty of Engineering from 2006 to 2008. In February 2008, he was appointed a Professor of electronics and computer science at the University of Southampton, where he is currently head of the Information: Signals, Images and Systems (ISIS) research group. His research interests are in the algorithmic and applied aspects of machine learning, and he has authored or coauthored about 100 papers in peer reviewed journals and conferences. He has been Program Chair of several international workshops and has acted as a co-organizer of a six-month program on Neural Networks and Machine Learning at the Isaac Newton Institute for Mathematical Sciences, Cambridge, U.K.

Minfen Shen (M'09) received the M.Eng. degree in signal processing and Ph.D. degrees in communication and information systems from the University of Science and Technology of China, Hefei, China, respectively.

Since graduation, he has been with the Department of Electrical and Electronic Engineering, Shantou University, Shantou, China, where he is currently a Professor. He is also an Adjunct Professor of the Department of Scientific Technology, University of Science and Technology of China, Hefei, China. His research interests include random signal processing, computational intelligence, biomedical signal analysis, nonlinear signal and image processing, and chaos.