

# A comparative analysis of Spearman's rho and Kendall's tau in normal and contaminated normal models

Weichao Xu <sup>a,\*</sup>, Yunhe Hou <sup>b</sup>, Y.S. Hung <sup>b</sup>, Yuexian Zou <sup>c</sup>

<sup>a</sup> Department of Automatic Control, School of Automation, Guangdong University of Technology, Guangzhou, Guangdong 510006, PR China

<sup>b</sup> Department of Electrical & Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong

<sup>c</sup> Advanced Digital Signal Processing Lab, Peking University Shenzhen Graduate School, Shenzhen, Guangdong 518055, PR China

## ARTICLE INFO

### Article history:

Received 13 September 2011

Received in revised form

20 July 2012

Accepted 6 August 2012

Available online 14 August 2012

### Keywords:

Bivariate normal

Correlation theory

Kendall's tau

Pearson's product moment correlation

coefficient

Spearman's rho

## ABSTRACT

This paper analyzes the performances of Spearman's rho (SR) and Kendall's tau (KT) with respect to samples drawn from bivariate normal and contaminated normal populations. Theoretical and simulation results suggest that, contrary to the opinion of equivalence between SR and KT in some literature, the behaviors of SR and KT are strikingly different in the aspects of bias effect, variance, mean square error (MSE), and asymptotic relative efficiency (ARE). The new findings revealed in this work provide not only deeper insights into the two most widely used rank-based correlation coefficients, but also a guidance for choosing which one to use under the circumstances where Pearson's product moment correlation coefficient (PPMCC) fails to apply.

© 2012 Elsevier B.V. All rights reserved.

## 1. Introduction

Correlation analysis is among the core research paradigms in nearly all branches of scientific and engineering fields, not to mention the area of signal processing [1–17]. Being interpreted as the strength of statistical relationship between two random variables [18], 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 [19]. A number of methods have been proposed and applied in the literature to assess the correlation between two random variables. Among these methods Pearson's product moment correlation coefficient (PPMCC) [20,21], Spearman's rho (SR) [22]

and Kendall's tau (KT) [22] are perhaps the most widely used [23].

The properties of PPMCC in bivariate normal samples (binormal model) is well known thanks to the fundamental work of Fisher [20]. It follows that, in the normal cases, (1) PPMCC is an asymptotically unbiased estimator of the population correlation  $\rho$  and (2) the variance of PPMCC approaches the Cramer–Rao lower bound (CRLB) as the sample size increases [18]. Due to its optimality, PPMCC has and will continue to play the dominant role when quantifying the intensity of correlation between bivariate random variables in the literature. However, sometimes the PPMCC might not be applicable when the following scenarios happen:

1. The data is incomplete, that is, only ordinal information (e.g. ranks) is available. This situation is not uncommon in the area of social sciences, such as psychology and education [22].
2. The underlying data is complete (cardinal) and follows a bivariate normal distribution, but is attenuated more

\* Corresponding author.

E-mail addresses: [wcxu@gdut.edu.cn](mailto:wcxu@gdut.edu.cn) (W. Xu), [yhhou@eee.hku.hk](mailto:yhhou@eee.hku.hk) (Y. Hou), [yshung@eee.hku.hk](mailto:yshung@eee.hku.hk) (Y.S. Hung), [zouyx@szpku.edu.cn](mailto:zouyx@szpku.edu.cn) (Y. Zou).

or less by some monotone nonlinearity in the transfer characteristics of sensors [24].

3. The data is complete and the majority follows a bivariate normal distribution, but there exists a tiny fraction of outliers with very large variance (impulsive noise) [25–27].

Under these circumstances, it would be more suitable to employ the two most popular nonparametric coefficients, SR and KT, which are (1) dependant only on ranks, (2) invariant under increasing monotone transformations [22], and (3) robust against impulsive noise [28,29]. Owing to these desired properties, SR and KT have drawn attentions of signal processing practitioners [30–33] in the late 1960s shortly after their being thoroughly studied in the literature of statistics [34–45]. However, despite the rich history of SR and KT, it is still unclear which one should we use in Scenarios 1–3 where the familiar PPMCC is inapplicable, even under the fundamental binormal model. Some researchers, such as Fieller et al. [44], preferred KT to SR based on empirical evidences; while others, such as Gilpin [46], asserted that SR and KT are equivalent in terms of hypothesis testing.

Aiming at resolving such inconsistency, in this work we investigate systematically the properties of SR and KT under the binormal model [20]. Moreover, to deal with Scenario 3 mentioned above, we also investigate their properties under the contaminated binormal model [28].

The rest part of this paper is structured as follows. Section 2 gives some basic definitions and summarizes the general properties of PPMCC, SR and KT. In Section 3, we lay the foundation of the theoretical framework in this study by outlining some critical results in the binormal model. Section 4 establishes the closed form formulae with respect to the expectations of SR and KT in the contaminated normal model. In Section 5, we focus on the performances of the three estimators of  $\rho$  constructed from SR and KT. Section 6 verifies the analytical results with Monte Carlo simulations. Finally, in Section 7 we provide our answers to the above raised question concerning the choice of Spearman's rho and Kendall's tau in practice when PPMCC fails to apply.

## 2. Basic definitions and general properties

### 2.1. Definitions

Let  $\{(X_i, Y_i)\}_{i=1}^n$  denote  $n$  independent and identically distributed (i.i.d.) data pairs drawn from a bivariate population with continuous joint distribution. Let  $P_j$  (resp.  $Q_j$ ) be the rank of  $X_j$  among  $X_1, \dots, X_n$  (resp.  $Y_j$  among  $Y_1, \dots, Y_n$ ) [22]. Let  $\bar{X}$  and  $\bar{Y}$  be the arithmetic mean values of  $X_i$  and  $Y_i$ , respectively. Let  $\text{sgn}(\cdot)$  stand for the sign of the argument. The sample versions of three well known classical correlation coefficient, PPMCC ( $r_P$ ), SR ( $r_S$ ), and KT ( $r_K$ ), are then defined as follows [19]:

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

$$r_S(X_i, Y_i) \triangleq 1 - \frac{6 \sum_{i=1}^n (P_i - Q_i)^2}{n(n^2 - 1)} \quad (2)$$

$$r_K(X_i, Y_i) \triangleq \frac{\sum_{i \neq j=1}^n \text{sgn}(X_i - X_j) \text{sgn}(Y_i - Y_j)}{n(n-1)} \quad (3)$$

To ease the following discussion, we will employ the symbol  $r_\eta(X_i, Y_i)$ ,  $\eta \in \{P, S, K\}$  as a compact notation for the three coefficients. For brevity, the arguments of  $r_\eta(X_i, Y_i)$  will be dropped in the sequel unless ambiguity occurs.

### 2.2. General properties

It follows that coefficients  $r_\eta$ ,  $\eta \in \{P, S, K\}$ , possess the following general properties:

1.  $r_\eta \in [-1, 1]$  for all  $(X, Y)$  (standardization);
2.  $r_\eta(X_i, Y_i) = r_\eta(Y_i, X_i)$  (symmetry);
3.  $r_\eta = \pm 1$  if  $Y$  is a positive (negative) linear transformation of  $X$  (shift and scale invariance);
4.  $r_S = r_K = \pm 1$  if  $Y$  is a monotone increasing (decreasing) function of  $X$  (monotone invariance);
5. the expectations of  $r_\eta$  equal zero if  $X$  and  $Y$  are independent (independence);
6.  $r_\eta(X_i, Y_i) = -r_\eta(-X_i, Y_i) = -r_\eta(X_i, -Y_i) = r_\eta(-X_i, -Y_i)$ ;
7.  $r_\eta$  converges to normal distribution when the sample size  $n$  is large.

Note that the first six properties are discussed in [19,23], and the last property follows from the asymptotic theory of  $U$ -statistics established by Hoeffding [38].

## 3. Auxiliary results in normal cases

In this section we provide some prerequisites concerning the orthant probabilities of normal distributions. These probabilities, contained in Lemma 1, are critical for the development of Lemmas 3 and 4 later on. Moreover, some well known results about the expectation and variance of PPMCC, SR and KT are collected in Lemma 2 for ease of exposition. For convenience, we use symbols  $\mathbb{E}(\cdot)$ ,  $\mathbb{V}(\cdot)$ ,  $\mathbb{C}(\cdot, \cdot)$ , and  $\text{corr}(\cdot, \cdot)$  in the sequel to denote the mean, variance, covariance, and correlation of (between) random variables, respectively. Symbols of *big oh* and *little oh* are utilized to compare the magnitudes of two functions  $u(t)$  and  $v(t)$  as the argument  $t$  tends to a limit  $L$  (might be infinite). The notation  $u(t) = O(v(t))$ ,  $t \rightarrow L$ , denotes that  $|u(t)/v(t)|$  remains bounded as  $t \rightarrow L$ ; whereas the notation  $u(t) = o(v(t))$ ,  $t \rightarrow L$ , denotes that  $u(t)/v(t) \rightarrow 0$  as  $t \rightarrow L$  [47]. Symbols of  $P_m^0(Z_1, \dots, Z_m) \triangleq \mathbb{P}(Z_1 > 0, \dots, Z_m > 0)$  are adopted to denote the positive orthant probabilities associated with multivariate normal random vectors  $[Z_1 \dots Z_m]$  of dimensions  $m = 1, \dots, 4$ , respectively. The notation  $R(\varrho_{rs})_{m \times m}$  stands for correlation matrix with each element  $\varrho_{rs} \triangleq \text{corr}(Z_r, Z_s)$ ,  $r, s = 1, \dots, m$ . Obviously the diagonal entries in  $R$  are all unities. For compactness, we will also use the symbol  $P_m^0(R)$  to denote  $P_m^0(Z_1, \dots, Z_m)$  in the sequel. Throughout the symbol  $\simeq$  stands for “approximately equals to”.

### 3.1. Orthant probabilities for normal distributions

**Lemma 1.** Assume that  $(Z_1, Z_2, Z_3, Z_4)$  follow a quadrivariate normal distribution with zero means and correlation matrix  $R = (\varrho_{rs})_{4 \times 4}$ . Define

$$H(t) \triangleq \begin{cases} 1 & (t > 0) \\ 0 & (t \leq 0) \end{cases} \quad (4)$$

Then the orthant probabilities

$$P_1^0(Z_1) = \mathbb{E}\{H(Z_1)\} = \frac{1}{2} \quad (5)$$

$$P_2^0(Z_1, Z_2) = \mathbb{E}\{H(Z_1)H(Z_2)\} = \frac{1}{4} \left( 1 + \frac{2}{\pi} \sin^{-1} \varrho_{12} \right) \quad (6)$$

$$\begin{aligned} P_3^0(Z_1, Z_2, Z_3) &= \mathbb{E}\{H(Z_1)H(Z_2)H(Z_3)\} \\ &= \frac{1}{8} \left( 1 + \frac{2}{\pi} \sum_{r=1}^2 \sum_{s=r+1}^3 \sin^{-1} \varrho_{rs} \right) \end{aligned} \quad (7)$$

$$\begin{aligned} P_4^0(Z_1, Z_2, Z_3, Z_4) &= \mathbb{E}\{H(Z_1)H(Z_2)H(Z_3)H(Z_4)\} \\ &= \frac{1}{16} \left( 1 + \frac{2}{\pi} \sum_{r=1}^3 \sum_{s=r+1}^4 \sin^{-1} \varrho_{rs} + W \right) \end{aligned} \quad (8)$$

where

$$W \triangleq \frac{1}{\pi^4} \iiint_{-\infty}^{+\infty} \frac{\exp\left(-\frac{1}{2} \mathbf{z} R \mathbf{z}^T\right)}{z_1 z_2 z_3 z_4} dz_1 dz_2 dz_3 dz_4 \quad (9)$$

$$= \sum_{\ell=2}^4 \frac{4}{\pi^2} \int_0^1 \frac{\varrho_{1\ell}}{[1 - \varrho_{1\ell}^2 u^2]^{1/2}} \sin^{-1} \left[ \frac{\alpha_{\ell}(u)}{[\beta_{\ell}(u) \gamma_{\ell}(u)]} \right] du \quad (10)$$

with

$$\alpha_2 = \varrho_{34} - \varrho_{23}\varrho_{24} - [\varrho_{13}\varrho_{14} + \varrho_{12}(\varrho_{12}\varrho_{34} - \varrho_{14}\varrho_{23} - \varrho_{13}\varrho_{24})]u^2$$

$$\alpha_3 = \varrho_{24} - \varrho_{23}\varrho_{34} - [\varrho_{12}\varrho_{14} + \varrho_{13}(\varrho_{13}\varrho_{24} - \varrho_{14}\varrho_{23} - \varrho_{12}\varrho_{34})]u^2$$

$$\alpha_4 = \varrho_{23} - \varrho_{24}\varrho_{34} - [\varrho_{12}\varrho_{13} + \varrho_{14}(\varrho_{14}\varrho_{23} - \varrho_{13}\varrho_{24} - \varrho_{12}\varrho_{34})]u^2$$

$$\beta_2 = \beta_3 = [1 - \varrho_{23}^2 - (\varrho_{12}^2 + \varrho_{13}^2 - 2\varrho_{12}\varrho_{13}\varrho_{23})u^2]^{1/2}$$

$$\gamma_2 = \beta_4 = [1 - \varrho_{24}^2 - (\varrho_{12}^2 + \varrho_{14}^2 - 2\varrho_{12}\varrho_{14}\varrho_{24})u^2]^{1/2}$$

$$\gamma_3 = \gamma_4 = [1 - \varrho_{34}^2 - (\varrho_{13}^2 + \varrho_{14}^2 - 2\varrho_{13}\varrho_{14}\varrho_{34})u^2]^{1/2}.$$

**Proof.** The first statement (5) is trivial. The second one (6) is usually called *Sheppard's theorem* in the literature, although it was proposed earlier by Stieltjes [48]. The third one (7) is a simple generalization of Sheppard's theorem based on the relationship [49]

$$P_3^0 = \frac{1}{2} \left[ 1 - \sum_{r=1}^3 P_1^0(Z_r) + \sum_{r=1}^2 \sum_{s=r+1}^3 P_2^0(Z_r, Z_s) \right].$$

The last one (8) is due to Childs [50] and is termed *Childs's reduction formula* throughout.  $\square$

### 3.2. Some well known results

**Lemma 2.** Let  $\{(X_i, Y_i)\}_{i=1}^n$  denote  $n$  i.i.d. bivariate normal data pairs with correlation coefficient  $\rho$ . Write  $S_1 \triangleq \sin^{-1} \rho$

and  $S_2 \triangleq \sin^{-1} \frac{1}{2} \rho$ . Then

$$\mathbb{E}(r_P) = \rho \left[ 1 - \frac{1 - \rho^2}{2n} + O(n^{-2}) \right] \rightarrow \rho \quad \text{as } n \rightarrow \infty \quad (11)$$

$$\mathbb{V}(r_P) = \frac{(1 - \rho^2)^2}{n - 1} + O(n^{-2}) \quad (12)$$

$$\mathbb{E}(r_S) = \frac{6}{\pi(n+1)} \left[ \sin^{-1} \rho + (n-2) \sin^{-1} \frac{\rho}{2} \right] \quad (13)$$

$$\rightarrow \frac{6}{\pi} \sin^{-1} \frac{\rho}{2} \quad \text{as } n \rightarrow \infty \quad (14)$$

$$\mathbb{E}(r_K) = \frac{2}{\pi} \sin^{-1} \rho \quad (15)$$

$$\mathbb{V}(r_K) = \frac{2}{n(n-1)} \left[ 1 - \frac{4S_1^2}{\pi^2} + 2(n-2) \left( \frac{1}{9} - \frac{4S_2^2}{\pi^2} \right) \right]. \quad (16)$$

**Proof.** The first three results, (11)–(13), were given by Hotelling [43], Fisher [21], and Moran [37], respectively; whereas the last two results, (15) and (16), were derived by Esscher [51].  $\square$

### 3.3. Variance of SR

**Lemma 3.** Let  $\{(X_i, Y_i)\}_{i=1}^n$ ,  $S_1$  and  $S_2$  be defined as in Lemma 2. Let  $W_c, W_d, \dots, W_q$  be defined as in (9) with respect to  $R_c, R_d, \dots, R_q$  that are tabulated in Table 1. Then the variance of  $r_S(X, Y)$  is

$$\begin{aligned} \mathbb{V}(r_S) &= \frac{6}{n(n+1)} + \frac{9(n-2)(n-3)}{n(n^2-1)(n+1)} [(n-4)\Omega_1(\rho) + \Omega_2(\rho)] \\ &\quad - \frac{36}{\pi^2 n(n^2-1)(n+1)} [3(n-2)(3n^2-15n+22)S_2^2 \\ &\quad + 12(n-2)^2 S_1 S_2 - 2(n-3)S_1^2] \end{aligned} \quad (17)$$

where

$$\Omega_1(\rho) = W_c + 8W_d + 2W_f \quad (18)$$

$$\Omega_2(\rho) = 6W_g + 8W_h + 6W_l + 2W_n + W_o + \frac{1}{3}. \quad (19)$$

Moreover, when  $n$  is sufficiently large,

$$\mathbb{V}(r_S) \simeq \frac{1}{n} \left[ 9\Omega_1(\rho) - \frac{324S_2^2}{\pi^2} \right]. \quad (20)$$

**Proof.** Using the technique developed by David and Mallows [45], it follows that the variance of  $r_S$  depends on a weighted summation of 24 quadrivariate normal orthant probabilities corresponding to  $R_c, R_d, \dots, R_q$  listed in Table 1. The lemma thus follows with the assistance of (10) in Lemma 1 after some tedious but straightforward derivations. See also the work of Croux and Dehon [29] and [52] for further discussion.  $\square$

**Remark 1.** Due to the complicated integrals involved in the expressions of  $W$ -terms in  $\Omega_1(\rho)$  and  $\Omega_2(\rho)$ , the variance of  $r_S$  cannot be expressed into elementary functions in general. Nevertheless, exact results can be obtained when  $\rho$  assumes some particular values. It can

**Table 1**Quantities for evaluation of  $\mathbb{E}(S^2)$  in Lemma 3.

Corr. matrix	No. of terms <sup>a</sup>	Representative term				Correlation coefficients						$P_4^0(Z_1, Z_2, Z_3, Z_4)^b$	$W(0)$	$W(1)$
		$Z_1$	$Z_2$	$Z_3$	$Z_4$	$\varrho_{12}$	$\varrho_{13}$	$\varrho_{14}$	$\varrho_{23}$	$\varrho_{24}$	$\varrho_{34}$			
$R_a$	$n^{[6]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_4 - X_5$	$Y_4 - Y_6$	$\frac{1}{2}\rho$	0	0	0	0	$\frac{1}{2}\rho$	$\frac{1}{16} + \frac{S_2}{4\pi} + \frac{S_2^2}{4\pi^2}$	–	–
$R_b$	$2n^{[5]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_4$	$Y_3 - Y_5$	$\rho$	0	0	0	0	$\frac{1}{2}\rho$	$\frac{1}{16} + \frac{S_1}{8\pi} + \frac{S_2}{8\pi} + \frac{S_1 S_2}{4\pi^2}$	–	–
$R_c$	$n^{[5]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_1 - X_4$	$Y_1 - Y_5$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{5}{48} + \frac{S_2}{2\pi} + \frac{W_c(\rho)}{16}$	$\frac{1}{9}$	$\frac{1}{5}$
$R_d$	$4n^{[5]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_2 - X_4$	$Y_2 - Y_5$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	0	0	$\frac{1}{2}\rho$	$\frac{1}{24} + \frac{S_2}{8\pi} + \frac{W_d(\rho)}{16}$	0	$\frac{1}{15}$
$R_e$	$2n^{[5]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_4 - X_2$	$Y_4 - Y_5$	$\frac{1}{2}\rho$	$\frac{1}{2}$	0	0	0	$\frac{1}{2}\rho$	$\frac{1}{12} + \frac{S_2}{4\pi} + \frac{W_e(\rho)}{16}$	–	–
$R_f$	$2n^{[5]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_4 - X_3$	$Y_4 - Y_5$	$\frac{1}{2}\rho$	0	0	$\frac{1}{2}\rho$	0	$\frac{1}{2}\rho$	$\frac{1}{16} + \frac{3S_2}{8\pi} + \frac{W_f(\rho)}{16}$	0	$\frac{2}{15}$
$R_g$	$4n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_1 - X_3$	$Y_1 - Y_4$	$\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{5}{48} + \frac{S_1}{8\pi} + \frac{3S_2}{8\pi} + \frac{W_g(\rho)}{16}$	$\frac{1}{9}$	$\frac{1}{3}$
$R_h$	$4n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_4$	$Y_3 - Y_1$	$\rho$	0	$\frac{1}{2}\rho$	0	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{24} + \frac{S_1}{8\pi} + \frac{W_h(\rho)}{16}$	0	$\frac{1}{3}$
$R_i$	$4n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_3 - X_4$	$Y_3 - Y_2$	$\frac{1}{2}\rho$	0	$\frac{1}{2}\rho$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{24} + \frac{S_2}{4\pi}$	–	–
$R_j$	$n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_4$	$Y_3 - Y_4$	$\rho$	0	0	0	0	$\rho$	$\frac{1}{16} + \frac{S_1}{4\pi} + \frac{S_1^2}{4\pi^2}$	–	–
$R_k$	$2n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_1 - X_2$	$Y_1 - Y_4$	$\frac{1}{2}\rho$	1	$\frac{1}{2}\rho$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{6} + \frac{S_2}{2\pi}$	–	–
$R_l$	$4n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_4 - X_1$	$Y_4 - Y_3$	$\frac{1}{2}\rho$	$\frac{1}{2}$	0	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{16} + \frac{S_2}{8\pi} + \frac{W_l(\rho)}{16}$	$\frac{1}{9}$	0
$R_m$	$n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_4 - X_2$	$Y_4 - Y_3$	$\frac{1}{2}\rho$	$\frac{1}{2}$	0	0	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{5}{48} + \frac{S_2}{4\pi} + \frac{W_m(\rho)}{16}$	–	–
$R_n$	$2n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_2 - X_4$	$Y_2 - Y_1$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$-\rho$	0	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{48} + \frac{S_1}{8\pi} + \frac{S_2}{4\pi} + \frac{W_n(\rho)}{16}$	$\frac{1}{9}$	0
$R_o$	$n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_4 - X_3$	$Y_4 - Y_2$	$\frac{1}{2}\rho$	0	$\frac{1}{2}\rho$	$\frac{1}{2}\rho$	0	$\frac{1}{2}\rho$	$\frac{1}{16} + \frac{S_2}{2\pi} + \frac{W_o(\rho)}{16}$	0	$\frac{1}{3}$
$R_p$	$2n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_2 - X_3$	$Y_2 - Y_4$	$\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{48} + \frac{S_1}{8\pi} + \frac{S_2}{8\pi} + \frac{W_p(\rho)}{16}$	–	–
$R_q$	$4n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_2$	$Y_3 - Y_4$	$\rho$	$\frac{1}{2}$	0	$\frac{1}{2}\rho$	0	$\frac{1}{2}\rho$	$\frac{1}{12} + \frac{S_1}{8\pi} + \frac{S_2}{4\pi} + \frac{W_q(\rho)}{16}$	–	–
$R_r$	$3n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_1 - X_3$	$Y_1 - Y_3$	$\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\rho$	$\frac{1}{9} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi} + \frac{S_1^2}{4\pi^2} - \frac{S_2^2}{4\pi^2}$	–	–
$R_s$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_1 - X_2$	$Y_1 - Y_3$	$\frac{1}{2}\rho$	1	$\frac{1}{2}\rho$	$\frac{1}{2}\rho$	1	$\frac{1}{2}\rho$	$\frac{1}{4} + \frac{S_2}{2\pi}$	–	–
$R_t$	$2n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_3$	$X_3 - X_1$	$Y_3 - Y_2$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$-\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{36} + \frac{S_1}{8\pi} + \frac{3S_2}{8\pi} - \frac{S_1^2}{8\pi^2} + \frac{S_2^2}{8\pi^2}$	–	–
$R_u$	$4n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_1 - X_2$	$Y_1 - Y_3$	$\rho$	1	$\frac{1}{2}\rho$	$\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{6} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$	–	–
$R_v$	$4n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_1$	$Y_3 - Y_2$	$\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{18} + \frac{S_1}{8\pi} + \frac{S_2}{8\pi} + \frac{S_1^2}{8\pi^2} - \frac{S_2^2}{8\pi^2}$	–	–
$R_w$	$2n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_1$	$Y_3 - Y_1$	$\rho$	$\frac{1}{2}$	$\frac{1}{2}\rho$	$\frac{1}{2}\rho$	$\frac{1}{2}$	$\rho$	$\frac{1}{36} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi} + \frac{S_1^2}{4\pi^2} - \frac{S_2^2}{4\pi^2}$	–	–
$R_x$	$n^{[2]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_1 - X_2$	$Y_1 - Y_2$	$\rho$	1	$\rho$	$\rho$	1	$\rho$	$\frac{1}{4} + \frac{S_1}{2\pi}$	–	–

<sup>a</sup> The symbol  $n^{[\kappa]}$  is a compact notation of  $n(n-1)\cdots(n-\kappa+1)$ , with  $\kappa$  being a positive integer.<sup>b</sup> The orthant probability  $P_4^0(Z_1, Z_2, Z_3, Z_4) \triangleq \mathbb{E}(H(Z_1)H(Z_2)H(Z_3)H(Z_4))$ . Notations  $S_1 \triangleq \sin^{-1} \rho$  and  $S_2 \triangleq \sin^{-1} \frac{1}{2} \rho$  are used for brevity.

be shown that

$$\Omega_1(0) = \frac{1}{9}, \quad \Omega_2(0) = \frac{5}{9} \quad (21)$$

$$\Omega_1(1) = 1, \quad \Omega_2(1) = \frac{16}{3} \quad (22)$$

Substituting  $\rho = 0$  and (21) into (17) leads directly to

$$\mathbb{V}(r_S)|_{\rho=0} = \frac{1}{n-1} \quad (23)$$

which is a well known result [22]. Substituting  $\rho = 1$  and (22) into (17) and (20) together with some simplifications yields the following obvious result:

$$\mathbb{V}(r_S)|_{\rho=1} = 0$$

And the same relationship also holds true for  $\rho = -1$  due to symmetry.

## 3.4. Covariance between SR and KT

**Lemma 4.** Let  $\{(X_i, Y_i)\}_{i=1}^n$ ,  $S_1$  and  $S_2$  be defined as in Lemma 2. Then the covariance between  $r_S(X, Y)$  and  $r_K(X, Y)$  is

$$\mathbb{C}(r_S, r_K) = \frac{12}{n(n^2-1)} \left[ \frac{7n-5}{18} + (n-4) \frac{S_1^2}{\pi^2} - 5(n-2) \frac{S_2^2}{\pi^2} - 6(n-2)^2 \frac{S_1 S_2}{\pi^2} + (n-2)(n-3) \Omega_3(\rho) \right] \quad (24)$$

$$\simeq \frac{12}{n} \left[ \Omega_3(\rho) - 6 \frac{S_1 S_2}{\pi^2} \right] \quad (\text{as } n \text{ large}) \quad (25)$$

where

$$\Omega_3(\rho) = \frac{1}{2} W_g + W_h. \quad (26)$$

**Proof.** See Appendix A.

**Corollary 1.** In Lemma 4, the covariance  $\mathbb{C}(r_S, r_K)$  can also be expressed as

$$\begin{aligned} \mathbb{C}(r_S, r_K) = & \frac{12}{n(n^2-1)} \left[ \frac{(n+1)^2}{18} + (n-4) \frac{S_1^2}{\pi^2} - 5(n-2) \frac{S_2^2}{\pi^2} \right. \\ & \left. - 6(n-2)^2 \frac{S_1 S_2}{\pi^2} + \frac{2}{\pi^2} (n-2)(n-3) \Omega_4(\rho) \right] \end{aligned} \quad (27)$$

$$\simeq \frac{12}{n} \left[ \frac{1}{18} + 2 \frac{\Omega_4(\rho)}{\pi^2} - 6 \frac{S_1 S_2}{\pi^2} \right] \quad (\text{as } n \text{ large}) \quad (28)$$

where

$$\begin{aligned} \Omega_4(\rho) = & \int_0^\rho \left[ \sin^{-1} \left( \frac{x}{3} \right) + 2 \sin^{-1} \left( \frac{x}{\sqrt{3}} \right) \right] \frac{dx}{\sqrt{1-x^2}} \\ & - 2 \int_0^\rho \sin^{-1} \left( \frac{x}{2} \sqrt{\frac{1-x^2}{9-3x^2}} \right) \frac{dx}{\sqrt{4-x^2}} \\ & + \int_0^\rho \sin^{-1} \left( \frac{x}{2} \sqrt{\frac{5-x^2}{3-x^2}} \right) \frac{dx}{\sqrt{4-x^2}} \\ & - 2 \int_0^\rho \sin^{-1} \left( x \sqrt{\frac{1-x^2}{12-6x^2}} \right) \frac{dx}{\sqrt{4-x^2}} \\ & + 2 \int_0^\rho \sin^{-1} \left( x \sqrt{\frac{3-x^2}{4-2x^2}} \right) \frac{dx}{\sqrt{4-x^2}} \end{aligned} \quad (29)$$

**Proof.** Inverting (8) yields

$$W = 16P_4^0 - 1 - \frac{2}{\pi} \sum_{r=1}^3 \sum_{s=r+1}^4 \sin^{-1} \varrho_{rs}. \quad (30)$$

Combining (26) and (30),  $\Omega_3(\rho)$  can be rewritten in terms of  $P_4^0$  and the correlation coefficients corresponding to  $R_g$  and  $R_h$  in Appendix 2 of [45]. This leads to

$$\Omega_3(\rho) = \frac{1}{18} + \frac{2}{\pi^2} \Omega_4(\rho). \quad (31)$$

The corollary thus follows directly by substituting (31) into (24) and (25), respectively.  $\square$

**Remark 2.** Both (24) and (27) are exact for any value of  $n \geq 4$  and  $|\rho| \leq 1$ . However, they are of different usefulness according to different numerical and analytical purposes. Formula (24) is more convenient for computation; whereas (27) is more convenient for analysis. For example, performing the Taylor expansion to (28) with the assistance of (29) gives

$$\begin{aligned} \mathbb{C}(r_S, r_K) \simeq & \frac{2}{3n} (1 - 1.24858961\rho^2 + 0.06830496\rho^4 \\ & + 0.07280482\rho^6 + 0.04025528\rho^8 + 0.02189277\rho^{10} + \dots) \end{aligned} \quad (32)$$

which agrees with the formula (51) obtained in [41], except for the coefficients of the last two terms, which we find to be 0.04025528 and 0.02189277, against their 0.04025526 and 0.01641362, respectively. Since  $\Omega_4(\rho)$  in

(27) is exact, we believe that (32) is more accurate than (51) in [41].

**Remark 3.** Due to the complicated integrals involved in (24) and (27),  $\mathbb{C}(r_S, r_K)$  cannot be expressed in elementary functions. However, exact results are attainable for  $\rho = 0$  and  $\rho = 1$  (−1). It follows that

$$\Omega_3(0) = \frac{1}{18} \quad (33)$$

$$\Omega_3(1) = \frac{1}{2} \quad (34)$$

Substituting (33) into (24) yields

$$\mathbb{C}(r_S, r_K)|_{\rho=0} = \frac{2}{3n} \frac{n+1}{n-1} \quad (35)$$

which is more readily to obtain on substitution of  $\rho = 0$  into (27). Regarding the case for  $\rho = 1$ , it is rather difficult by means of substituting  $\rho = 1$  into (27) and evaluating  $\Omega_4(1)$  based on (29) thereafter. Fortunately, with the help of (34), it follows readily from (24) and (25) that  $\mathbb{C}(r_S, r_K)|_{\rho=1} = 0$ . Due to symmetry, we also have  $\mathbb{C}(r_S, r_K)|_{\rho=-1} = 0$ .

#### 4. New results in contaminated normal model

The PPMCC is notoriously sensitive to the non-Gaussianity caused by impulsive contamination in the data. Even a single outlier can distort severely the value of PPMCC and hence result in misleading inference in practice. Assume that  $(X, Y)$  obeys the following distribution [28]:

$$(1-\epsilon)\mathcal{N}(\mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho) + \epsilon\mathcal{N}(\mu_X, \mu_Y, \lambda_X^2 \sigma_X^2, \lambda_Y^2 \sigma_Y^2, \rho') \quad (36)$$

where  $0 \leq \epsilon \leq 1$ ,  $\lambda_X \gg 1$ , and  $\lambda_Y \gg 1$ . Under this Gaussian contamination model<sup>1</sup> that frequently used in the literature of robustness analysis [25–27], it has been shown that, no matter how small  $\epsilon$  is, the expectation of the PPMCC  $\mathbb{E}(r_P) \rightarrow \rho'$  as  $\lambda_X \rightarrow \infty$  and  $\lambda_Y \rightarrow \infty$  [28]. On the other hand, as shown in the theorem below, SR and KT are more robust than PPMCC under the model (36).

**Theorem 1.** Let  $\{(X_i, Y_i)\}_{i=1}^n$  be i.i.d. samples generated from the model (36). Then for  $\epsilon$  around 0,

$$\lim_{\substack{\lambda_X \rightarrow \infty \\ \lambda_Y \rightarrow \infty}} \mathbb{E}(r_K) \simeq \frac{2}{\pi} \left[ (1-2\epsilon) \sin^{-1} \rho + 2\epsilon \sin^{-1} \rho' \right] \quad (37)$$

$$\lim_{\substack{n \rightarrow \infty \\ \lambda_X \rightarrow \infty \\ \lambda_Y \rightarrow \infty}} \mathbb{E}(r_S) \simeq \frac{6}{\pi} \left[ (1-3\epsilon) \sin^{-1} \frac{\rho}{2} + \epsilon \sin^{-1} \rho' \right]. \quad (38)$$

where the order of taking limits is arbitrary.

**Proof.** See Appendix B.

**Remark 4.** It was stated without substantial argument in [28] that, under the model (36), the first approximation of

<sup>1</sup> Note that here we only consider the centered and symmetric type of contamination that is frequently encountered in the literature of signal processing. For general situations, the readers are referred to [29].

$\mathbb{E}(r_S)$ , in terms of  $\epsilon$ , is of the following form:

$$\mathbb{E}(r_S) \simeq \frac{6}{\pi} \left[ (1-\epsilon) \sin^{-1} \frac{\rho}{2} + \epsilon \sin^{-1} \frac{\rho'}{2} \right] \quad (*)$$

as  $n \rightarrow \infty$ ,  $\lambda_X \rightarrow \infty$  and  $\lambda_Y \rightarrow \infty$ . This is quite inconsistent with our result (38) in Theorem 1. We will resolve the controversy between (38) and (\*) by Monte Carlo simulations in Section 6.

## 5. Estimators of the population correlation

In this section, we investigate the performance of several estimators of  $\rho$  based on SR and KT in terms of bias, MSE and ARE to PPMCC.

### 5.1. Asymptotic unbiased estimators

Inverting (11), (14) and (15), we have the following estimators of  $\rho$ :

$$\hat{\rho}_P \triangleq r_P \quad (39)$$

$$\hat{\rho}_S \triangleq 2 \sin\left(\frac{\pi}{6} r_S\right) \quad (40)$$

$$\hat{\rho}_K \triangleq \sin\left(\frac{\pi}{2} r_K\right). \quad (41)$$

Moreover, another estimator based on a mixture of  $r_S$  and  $r_P$  can be constructed as [22]

$$\hat{\rho}_M \triangleq 2 \sin\left(\frac{\pi}{6} r_S - \frac{\pi}{2} \frac{r_K - r_S}{n-2}\right) \quad (42)$$

based on the following relationship:

$$\mathbb{E}(r_S) = \frac{6}{\pi} \left( S_2 + \frac{S_1 - 3S_2}{n+1} \right). \quad (43)$$

In the sequel we will focus on the properties of the estimators defined in (39)–(42). Here the estimator  $\hat{\rho}_P$  in (39) is employed as a benchmark due to its optimality for normal samples, in the sense of approaching the CRLB [18] when the sample size is sufficiently large.

### 5.2. Bias effect for small samples

It is noteworthy that the four estimators in (39)–(42) are unbiased only for large samples. When the sample size is small, the bias effects, as shown in the following theorem, are not ignorable any more.

**Theorem 2.** Let  $\hat{\rho}_\zeta, \zeta \in \{P, S, K, M\}$ , be defined as in (39)–(42), respectively. Define  $\text{BIAS}_\zeta \triangleq \mathbb{E}(\hat{\rho}_\zeta - \rho)$ . Let  $S_1$  and  $S_2$  bear the same meanings as in Lemma 2. Write  $\sigma_S^2 \triangleq \mathbb{V}(r_S)$ ,  $\sigma_K^2 \triangleq \mathbb{V}(r_K)$  and  $\sigma_{S,K} \triangleq \mathbb{C}(r_S, r_K)$ . Then, under the same assumptions made as in Lemma 3,

$$\text{BIAS}_P \simeq -\frac{1}{2n} \rho(1-\rho^2) \quad (44)$$

$$\text{BIAS}_S \simeq \frac{\sqrt{4-\rho^2}}{n+1} (S_1 - 3S_2) - \frac{\pi^2 \rho}{72} \sigma_S^2 \quad (45)$$

$$\text{BIAS}_K \simeq -\frac{\pi^2 \rho}{8} \sigma_K^2 \quad (46)$$

$$\text{BIAS}_M \simeq -\frac{1}{72} \frac{\pi^2 \rho}{(n-2)^2} [(n+1)^2 \sigma_S^2 - 6(n+1) \sigma_{S,K} + 9\sigma_K^2] \quad (47)$$

**Proof.** The first statement (44) follows directly from (11) in Lemma 2. Now we proceed to evaluate  $\text{BIAS}_S$ ,  $\text{BIAS}_K$  and  $\text{BIAS}_M$ . For convenience, write  $\bar{r}_S \triangleq \mathbb{E}(r_S)$ ,  $\bar{r}_K \triangleq \mathbb{E}(r_K)$ ,  $\delta_S \triangleq r_S - \bar{r}_S$ , and  $\delta_K \triangleq r_K - \bar{r}_K$ . Expanding (40) around  $\bar{r}_S$  yields

$$\hat{\rho}_S = 2 \sin\left(\frac{\pi}{6} \bar{r}_S\right) + \frac{\pi}{3} \cos\left(\frac{\pi}{6} \bar{r}_S\right) \delta_S - \frac{\pi^2}{36} \sin\left(\frac{\pi}{6} \bar{r}_S\right) \delta_S^2 + \dots \quad (48)$$

Taking expectation of both sides in (48), applying  $\mathbb{E}(\delta_S) = 0$ ,  $\mathbb{E}(\delta_S^2) = \sigma_S^2$  and ignoring the high order infinitesimals, we have

$$\mathbb{E}(\hat{\rho}_S) \simeq 2 \sin\left(\frac{\pi}{6} \bar{r}_S\right) - \frac{\pi^2}{36} \sin\left(\frac{\pi}{6} \bar{r}_S\right) \sigma_S^2. \quad (49)$$

Substituting (43) into (49), expanding to the order of  $(n+1)^{-1}$ , and subtracting  $\rho$  thereafter, we obtain the result (45). In a similar way we have

$$\begin{aligned} \mathbb{E}(\hat{\rho}_K) &\simeq \rho - \frac{\pi^2 \rho}{8} \sigma_K^2 \\ \text{which leads directly to (46). Performing Taylor expansion of } \hat{\rho}_M(r_S, r_K) \text{ around } (\bar{r}_S, \bar{r}_K) \text{ till the second order, we have} \\ \hat{\rho}_M &= \hat{\rho}_M(\bar{r}_S, \bar{r}_K) + \frac{\partial(\hat{\rho}_M)}{\partial(r_S)} \delta_S + \frac{\partial(\hat{\rho}_M)}{\partial(r_K)} \delta_K \\ &\quad + \frac{1}{2} \left[ \frac{\partial^2(\hat{\rho}_M)}{\partial(r_S)^2} \delta_S^2 + \frac{\partial^2(\hat{\rho}_M)}{\partial(r_K)^2} \delta_K^2 + \frac{2\partial^2(\hat{\rho}_M)\delta_S\delta_K}{\partial(r_S)\partial(r_K)} \right] + \dots \end{aligned} \quad (50)$$

Taking expectation of both sides in (50), ignoring high order infinitesimals, applying the results  $\hat{\rho}_M(\bar{r}_S, \bar{r}_K) = \rho$ ,  $\mathbb{E}(\delta_S) = 0$ ,  $\mathbb{E}(\delta_K) = 0$ ,  $\mathbb{E}(\delta_S^2) = \sigma_S^2$ ,  $\mathbb{E}(\delta_K^2) = \sigma_K^2$ ,  $\mathbb{E}(\delta_S\delta_K) = \sigma_{S,K}$  along with the second order partial derivatives

$$\begin{aligned} \frac{\partial^2 \hat{\rho}_M(\bar{r}_S, \bar{r}_K)}{\partial(r_S)^2} &= -\frac{\pi^2 \rho}{36} \frac{(n+1)^2}{(n-2)^2} \\ \frac{\partial^2 \hat{\rho}_M(\bar{r}_S, \bar{r}_K)}{\partial(r_K)^2} &= -\frac{\pi^2 \rho}{4} \frac{1}{(n-2)^2} \\ \frac{\partial^2 \hat{\rho}_M(\bar{r}_S, \bar{r}_K)}{\partial(r_S)\partial(r_K)} &= \frac{\pi^2 \rho}{12} \frac{n+1}{(n-2)^2} \end{aligned}$$

and subtracting  $\rho$  thereafter, we arrive at the fourth theorem statement (47), thus completing the proof.  $\square$

**Remark 5.** From (44) to (47), it follows that, for all the four estimators,

- $\text{BIAS}_\zeta(\rho) = \text{BIAS}_\zeta(-\rho)$  (odd symmetry);
- $\rho \text{BIAS}_\zeta(\rho) \leq 0$  (negative bias);
- $\text{BIAS}_\zeta = 0$  for  $\rho \in \{-1, 0, 1\}$ ;
- $\text{BIAS}_\zeta \sim O(n^{-1})$  as  $n \rightarrow \infty$ .

Moreover, contrary to  $\text{BIAS}_P$  and  $\text{BIAS}_K$ ,  $\text{BIAS}_S$  and  $\text{BIAS}_M$  cannot be expressed into elementary functions due to the intractability involved in (17) and (24), the expressions of  $\mathbb{V}(r_S)$  and  $\mathbb{C}(r_S, r_K)$ , respectively.

**Table 2**  
Values of  $\Omega_1(\rho)$  and  $\Omega_2(\rho)$  in Lemma 3.

$\rho$	0.00+	0.10+	0.20+	0.30+	0.40+	0.50+	0.60+	0.70+	0.80+	0.90+
0.00	0.1111111111	0.1185038469	0.1408155407	0.1784569533	0.2321489296	0.3029841008	0.3925307865	0.5030051934	0.6375648509	0.8008401854
0.01	0.1111849294	0.1200591282	0.1438805317	0.1830890232	0.2384392872	0.3110659830	0.4025930108	0.5153148070	0.6525067154	0.8189917929
0.02	0.1114063976	0.1217636535	0.1470991369	0.1878822020	0.2449020129	0.3193363285	0.4128663989	0.5278679000	0.6677394272	0.8375074957
0.03	0.1117755552	0.1236177309	0.1504719575	0.1928374325	0.2515384762	0.3277970726	0.4233536847	0.5406684104	0.6832688919	0.8563967883
0.04	0.1122924682	0.1256216966	0.1539996256	0.1979556956	0.2583500955	0.3364502180	0.4340577002	0.5537204299	0.6991012791	0.8756696820
0.05	0.1129572289	0.1277759147	0.1576828058	0.2032380114	0.2653383401	0.3452978372	0.4449813799	0.5670282120	0.7152430387	0.8953367468
0.06	0.1137699564	0.1300807778	0.1615221950	0.2086854397	0.2725047311	0.3543420748	0.4561277650	0.5805961801	0.7317009189	0.9154091574
0.07	0.1147307963	0.1325367071	0.1655185233	0.2142990814	0.2798508436	0.3635851506	0.4675000084	0.5944289365	0.7484819855	0.9358987450
0.08	0.1158399207	0.1351441527	0.1696725544	0.2200800791	0.2873783079	0.3730293618	0.4791013791	0.6085312723	0.7655936432	0.9568180554
0.09	0.1170975289	0.1379035942	0.1739850867	0.2260296186	0.2950888115	0.3826770865	0.4909352680	0.6229081771	0.7830436585	0.9781804140
0.00	0.5555555556	0.5831779199	0.6669411548	0.8096548728	1.0164469967	1.2956050434	1.6602428005	2.1318398440	2.7490961353	3.5988067890
0.01	0.5558310525	0.5889973638	0.6784952270	0.8273423444	1.0409353198	1.3279500727	1.7021306947	2.1861334471	2.8213687297	3.7041428194
0.02	0.5566576312	0.5953785532	0.6906408759	0.8456746565	1.0661538717	1.3611599188	1.7451031281	2.2419045230	2.8959717933	3.8148964800
0.03	0.5580355547	0.6023235637	0.7033822725	0.8646587057	1.0921135054	1.3952518608	1.7891893159	2.2992092121	2.9730475022	3.9318774005
0.04	0.5599652624	0.6098346629	0.7167238250	0.8843017153	1.1188255760	1.4302440445	1.8344202199	2.3581081295	3.0527562596	4.0561698333
0.05	0.5624473698	0.6179143137	0.7306701849	0.9046112480	1.1463019652	1.4661555384	1.8808286962	2.4186669028	3.1352804112	4.1892949503
0.06	0.5654826698	0.6265651776	0.7452262544	0.9255952194	1.1745551090	1.5030063944	1.9284496603	2.4809568021	3.2208290418	4.3335266816
0.07	0.5690721334	0.6357901183	0.7603971935	0.9472619127	1.2035980265	1.5408177141	1.9773202724	2.5450554807	3.3096442738	4.4925913045
0.08	0.5732169115	0.6455922055	0.7761884285	0.9696199936	1.2334443518	1.5796117219	2.0274801460	2.6110478526	3.4020096999	4.6735571704
0.09	0.5779183355	0.6559747193	0.7926056602	0.9926785275	1.2641083679	1.6194118448	2.0789715836	2.6790271403	3.4982619270	4.8941554764

(a) In the upper panel are the values of  $\Omega_1(\rho)$ , and in the lower panel (shaded area) are the values of  $\Omega_2(\rho)$ .

(b)  $\Omega_1(0) = 1/9$ ,  $\Omega_1(1) = 1$ ,  $\Omega_1(-\rho) = \Omega_1(\rho)$ .

(c)  $\Omega_2(0) = 5/9$ ,  $\Omega_2(1) = 16/3$ ,  $\Omega_2(-\rho) = \Omega_2(\rho)$ .



### 5.3. Approximation of variances

Besides the bias effect just discussed, the variance is another important *figure of merit* when comparing the performance of the estimators  $\hat{\rho}_\zeta$ ,  $\zeta \in \{P, S, K, M\}$ . From (11), it follows that

$$\mathbb{V}(\hat{\rho}_P) \simeq \frac{(1-\rho^2)^2}{n-1} \quad (51)$$

By the delta method, it follows that [22]

$$\mathbb{V}(\hat{\rho}_S) \simeq \frac{\pi^2(4-\rho^2)}{36} \mathbb{V}(r_S) \quad (52)$$

$$\mathbb{V}(\hat{\rho}_K) \simeq \frac{\pi^2(1-\rho^2)}{4} \mathbb{V}(r_K) \quad (53)$$

Now we only need to deal with  $\mathbb{V}(\hat{\rho}_M)$ , which is stated below.

**Theorem 3.** Let  $\hat{\rho}_M$  be defined as in (42). Then, under the same assumptions made as in Lemma 3,

$$\mathbb{V}(\hat{\rho}_M) \simeq \frac{\pi^2(4-\rho^2)}{36(n-2)^2} [(n+1)^2\sigma_S^2 - 6(n+1)\sigma_{S,K} + 9\sigma_K^2] \quad (54)$$

**Proof.** Using the delta method [18], it follows that

$$\mathbb{V}(\hat{\rho}_M) \simeq \left[ \frac{\partial(\hat{\rho}_M)}{\partial(r_S)} \right]^2 \sigma_S^2 + \left[ \frac{\partial(\hat{\rho}_M)}{\partial(r_K)} \right]^2 \sigma_K^2 + 2 \frac{\partial(\hat{\rho}_M)}{\partial(r_S)} \frac{\partial(\hat{\rho}_M)}{\partial(r_K)} \sigma_{S,K} \quad (55)$$

The theorem thus follows with the substitutions of the partial derivatives

$$\frac{\partial \hat{\rho}_M(r_S, r_K)}{\partial(r_S)} = \frac{\pi n + 1}{6 n - 2} \sqrt{4 - \rho^2}$$

$$\frac{\partial \hat{\rho}_M(r_S, r_K)}{\partial(r_K)} = \frac{\pi - 1}{2 n - 2} \sqrt{4 - \rho^2}$$

into (55), respectively.  $\square$

### 5.4. Asymptotic relative efficiency

Thus far in this section we have established the analytical results with an emphasis on limited-sized bivariate normal samples. For a better understanding of the four estimators, we will shift our attention to the asymptotic properties of  $\hat{\rho}_\zeta$  in the sequel. Since  $\lim_{n \rightarrow \infty} \mathbb{E}(\hat{\rho}_\zeta) = \rho$ , we can compare their performances by means of the *asymptotic relative efficiency*, which is defined as [18]

$$\text{ARE}_\zeta(\rho) \triangleq \lim_{n \rightarrow \infty} \frac{\mathbb{V}(\hat{\rho}_P)}{\mathbb{V}(\hat{\rho}_\zeta)}, \quad \zeta \in \{P, S, K, M\} \quad (56)$$

As remarked before, we employ  $\hat{\rho}_P$  as a benchmark, since, for large-sized bivariate normal samples,  $\hat{\rho}_P$  approaches the Cramer–Rao lower bound (CRLB) [18]

$$\text{CRLB} = \frac{(1-\rho^2)^2}{n} \quad (57)$$

From (56) it is obvious that  $\text{ARE}_P = 1$ . Moreover, comparing (52) and (54), it is easily seen that  $\lim_{n \rightarrow \infty} \mathbb{V}(\hat{\rho}_S)/\mathbb{V}(\hat{\rho}_M) = 1$ , which leads readily to  $\text{ARE}_S = \text{ARE}_M$  by referring to (56). Then we only need to focus on  $\text{ARE}_S$  and  $\text{ARE}_K$  in the following discussion.

**Table 3**  
Values of  $\Omega_3(\rho)$  in Lemma 4.

$\rho$	0.00 +	0.10 +	0.20 +	0.30 +	0.40 +	0.50 +	0.60 +	0.70 +	0.80 +	0.90 +
0.00	0.0555555556	0.0579082728	0.0650488362	0.0772363477	0.0949464623	0.1189551518	0.1505074640	0.1916842312	0.2463473665	0.3235686523
0.01	0.0555790160	0.0584040661	0.0660345125	0.0787487694	0.0970477974	0.1217448819	0.1541471283	0.1964547749	0.2528155359	0.3333572518
0.02	0.0556494053	0.0589477686	0.0670708481	0.0803167902	0.0992127259	0.1246109820	0.1578844390	0.2013621185	0.2595091876	0.3437145088
0.03	0.0557667477	0.0595395710	0.0681582284	0.0819410523	0.1014422703	0.1275551071	0.1617222612	0.2064119567	0.2664434835	0.35477334972
0.04	0.0559310834	0.0601796821	0.0692970610	0.0836222291	0.1037375018	0.1305789987	0.1656636397	0.2116104637	0.2736356357	0.3665400166
0.05	0.0561424692	0.0608683287	0.0704877763	0.0853610263	0.1060995431	0.1336844907	0.1697118152	0.2169643536	0.2811053381	0.3793122556
0.06	0.0564009777	0.0616057558	0.0717308284	0.0871581833	0.1085295710	0.13688735157	0.1738702420	0.2224809499	0.2888753252	0.393319258
0.07	0.0567066983	0.0623292274	0.0730266953	0.0890144747	0.1110288192	0.1401481121	0.1781426083	0.2281682700	0.2969721099	0.4090062849
0.08	0.0570597366	0.0632280263	0.0743758804	0.0909307119	0.1135985821	0.1435104314	0.1825328587	0.2340351233	0.3054269741	0.4272272392
0.09	0.0574602150	0.0641134550	0.0757789125	0.0929077445	0.1162402177	0.1469627471	0.1870452202	0.2400912305	0.3142773322	0.4501481060

Note that (a)  $\Omega_3(0) = 1/18$ , (b)  $\Omega_3(1) = 1/2$ , and (c)  $\Omega_3(-\rho) = \Omega_3(\rho)$ .



**Theorem 4.** Let  $ARE_S$  and  $ARE_K$  be defined as in (56). Then

$$ARE_S(\rho) = \frac{36(1-\rho^2)^2}{(4-\rho^2)[9\pi^2\Omega_1(\rho)-324(\sin^{-1}\frac{1}{2}\rho)^2]} \quad (58)$$

$$ARE_K(\rho) = \frac{9(1-\rho^2)}{\pi^2-36(\sin^{-1}\frac{1}{2}\rho)^2} \quad (59)$$

**Proof.** Substituting (52) and (53) into (56) yields (58) and (59), respectively, and the proof completes. See also the proof in [29] based on influence functions.  $\square$

**Remark 6.** Due to the intractability of  $\Omega_1(\rho)$  in (58),  $ARE_S$  cannot be expressed into elementary functions in general. However, exact results are obtainable for  $\rho = 0, \pm 1$ . Substituting  $\rho = 0$  and  $\Omega_1(0) = 1/9$  into (58), it is easy to verify that

$$ARE_S(0) = \frac{9}{\pi^2} \simeq 0.9119$$

which is a well known result [22]. In our previous work [14] we also obtained that

$$ARE_S(\pm 1) = \frac{15+11\sqrt{5}}{57} \simeq 0.6947. \quad (60)$$

Now let us investigate  $ARE_K$ . It follows from (59) that,  $ARE_K$  is expressible as elementary functions of  $\rho$ , and is therefore more tractable than  $ARE_S$ . In other words, we can evaluate easily any value of  $ARE_K$  with respect to any value of  $\rho \neq \pm 1$ . For example, substituting  $\rho = 0$  into (59) yields

$$ARE_K(0) = \frac{9}{\pi^2}$$

which is identical to  $ARE_S(0)$  and also well known [22]. However, when  $\rho \rightarrow \pm 1$ , an extra effort is necessary, since both the numerator and denominator of (59) vanish in this

case. Apply L'Hopital's rule, we find the following result:

$$ARE_K|_{\rho \rightarrow \pm 1} = \frac{1}{4} \frac{\rho \sqrt{4-\rho^2}}{\sin^{-1}\frac{1}{2}\rho} \Big|_{\rho=\pm 1} = \frac{3\sqrt{3}}{2\pi} \simeq 0.8270 \quad (61)$$

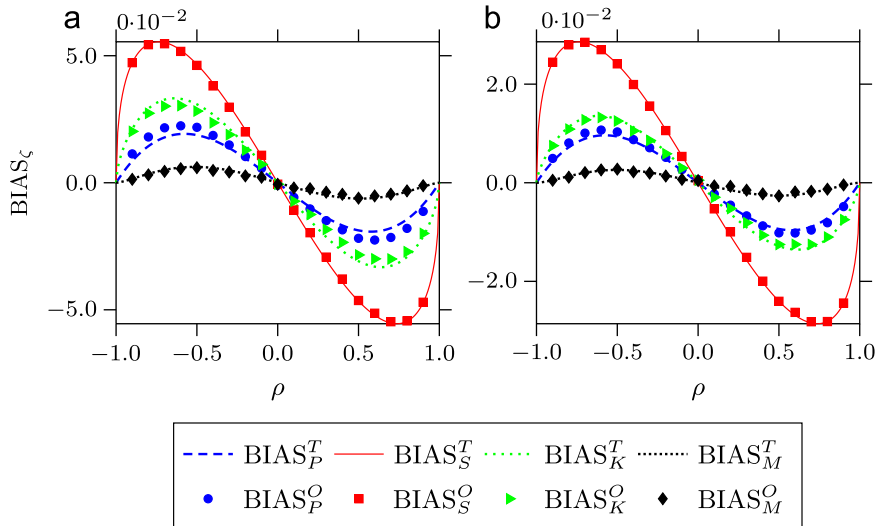
which is greater than  $ARE_S(\pm 1)$ . In fact, a comparison of  $ARE_S$  and  $ARE_K$  in Section 6 suggest that  $ARE_S \leq ARE_K$  for all  $\rho \in [-1, 1]$ .

## 6. Numerical results

In this section we aim at (1) tabulating the values of  $\Omega_1(\rho)$ ,  $\Omega_2(\rho)$  (in Lemma 3) and  $\Omega_3(\rho)$  (in Lemma 4) that are not expressible as elementary functions, (2) verifying the theoretical results established in Theorem 1, and (3) comparing the performances of the four estimators defined in (39)–(42) by means of *bias effect*, *mean square error* (MSE) and ARE under both the normal and contaminated normal models (Theorems 2–4). Throughout this section, Monte Carlo experiments are undertaken for  $10 \leq n \leq 100$ . A sample size of  $n=1000$  is considered large enough when we investigate the asymptotic behaviors. The number of trials is set to  $5 \times 10^5$  for reason of accuracy. All samples are generated by functions in the Matlab Statistics Toolbox™. Specifically, the normal samples are generated by `mvnrnd`, whereas the contaminated normal samples are generated by `gmdistribution` and `random`. The notation  $\rho = \rho_1(\Delta\rho)\rho_2$  represents a list of  $\rho$  starting from  $\rho_1$  to  $\rho_2$  with increment  $\Delta\rho$ .

### 6.1. Tables of $\Omega_1(\rho)$ , $\Omega_2(\rho)$ and $\Omega_3(\rho)$

Table 2 contains the values of  $\Omega_1(\rho)$  and  $\Omega_2(\rho)$  in (17), the first statement of Lemma 3 for  $\rho = 0(0.01)1$ . In the upper panel are the values of  $\Omega_1(\rho)$ ; whereas in the lower panel are the values of  $\Omega_2(\rho)$ . Due to the importance of  $\mathbb{V}(r_S)$  both in theory and in practice, the table is made as intensive and accurate as possible, with the increment  $\Delta\rho$  being 0.01, and the precision being up to 10 decimal



**Fig. 1.** Comparison of  $BIAS_\zeta$ ,  $\zeta \in \{P, S, K, M\}$  for (a)  $n=10$  and (b)  $n=20$ . Theoretical curves, denoted by  $BIAS_\zeta^T$  in the legend, are plotted over  $\rho = -1(0.01)1$  based on (44)–(47), respectively; whereas the simulation results, denoted by  $BIAS_\zeta^O$  in the legend, are plotted over  $\rho = -0.9(0.1)0.9$ .

places. In Table 3 are tabulated the values of  $\Omega_3(\rho)$  in (24) of Lemma 4 for  $\rho = 0(0.01)1$ . Because of the similar reasons, the increment  $\Delta\rho$  and precision are the same as those in Table 2. The values of  $\Omega_1(\rho)$ ,  $\Omega_2(\rho)$  and  $\Omega_3(\rho)$  with respect to  $\rho$  not included in Tables 2 and 3 can be easily obtained by interpolation. Given these tables, we can easily calculate the quantities that depend on  $\Omega_1(\rho)$ ,  $\Omega_2(\rho)$  and  $\Omega_3(\rho)$ , including  $\mathbb{V}(\hat{r}_S)$ ,  $\mathbb{V}(\hat{\rho}_S)$ ,  $\mathbb{V}(\hat{\rho}_M)$ ,  $\text{BIAS}(\hat{\rho}_S)$ ,  $\text{BIAS}(\hat{\rho}_M)$ ,  $\text{ARE}_S$ ,  $\text{ARE}_M$ , and so forth.

### 6.2. Verification of $\text{BIAS}_\zeta$ and $\mathbb{V}(\hat{\rho}_\zeta)$ in small samples

Fig. 1 shows the bias effects  $\text{BIAS}_\zeta$  corresponding to the four estimators  $\rho_\zeta$ ,  $\zeta \in \{P, S, K, M\}$  for  $n=10$  and  $n=20$ , respectively. It is clearly observed that the magnitudes of  $\text{BIAS}_\zeta$  can be ordered as  $|\text{BIAS}_M| < |\text{BIAS}_P| < |\text{BIAS}_K| < |\text{BIAS}_S|$ . That is, the performance of  $r_S$  is much worse than those of the other three in terms of bias effect in small samples. Moreover, it is also observed that (45) and (47) with respect to  $\text{BIAS}_S$  and  $\text{BIAS}_M$  are more accurate than (44) and (46) with respect to  $\text{BIAS}_P$  and  $\text{BIAS}_K$ . In other words, the former two formulae agree better than do the latter two formulae with the corresponding simulation results for a sample size  $n$  as small as 10. Nevertheless, the deviations from (44) and (46) to the corresponding simulation results are less noticeable when the sample size  $n$  is increased up to 20.

Table 4 lists, for each of the three samples sizes, 10, 20 and 30, (1) the theoretical results (51)–(54) with respect to  $\mathbb{V}(\hat{\rho}_\zeta)$  and (2) the corresponding observed variances from the Monte Carlo simulations. It can be seen that (52) and (54), with respect to  $\mathbb{V}(\hat{\rho}_S)$  and  $\mathbb{V}(\hat{\rho}_M)$ , are accurate enough even though the sample size is as small as  $n=10$ . On the other hand, unfortunately, the theoretical formula (51) for  $\mathbb{V}(\hat{\rho}_P)$  and (53) for  $\mathbb{V}(\hat{\rho}_K)$  deviate substantially from the corresponding observed simulation results for the same sample size  $n=10$ . However, it appears that these deviations become less noticeable for  $n=20$  and negligible for  $n=30$ . Therefore, it would be save to use (51)–(54) when approximating the variances of  $\hat{\rho}_\zeta$  for  $n \geq 20$  in practice.

### 6.3. Comparison of MSE in small samples

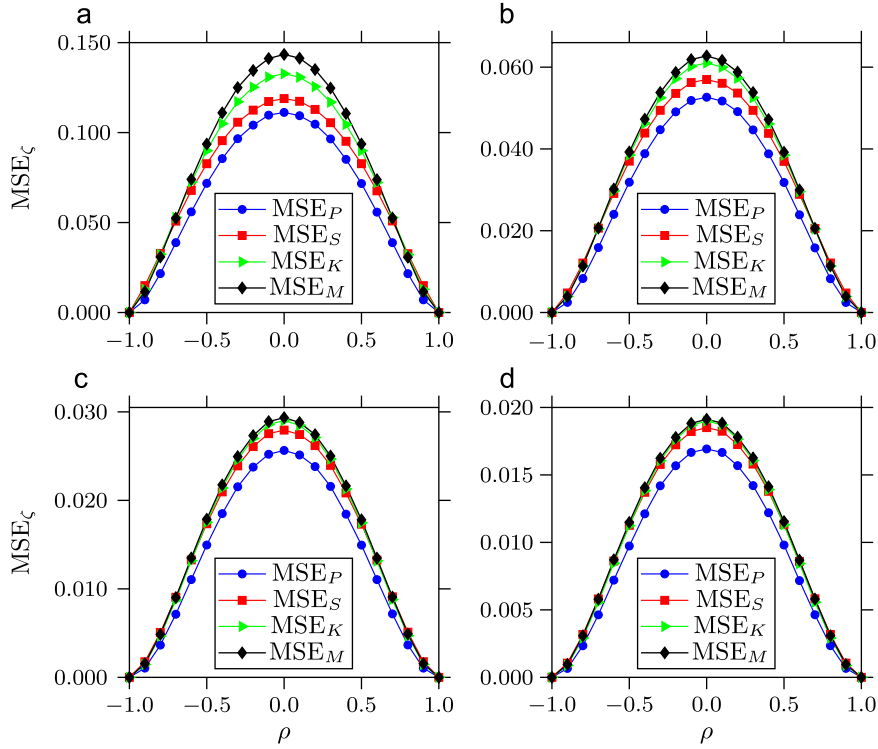
Contrary to  $\text{BIAS}_\zeta$  illustrated in Fig. 1, the magnitudes of the mean square errors

$$\text{MSE}_\zeta \triangleq \mathbb{E}[(\hat{\rho}_\zeta - \rho)^2], \quad \zeta \in \{P, S, K, M\}$$

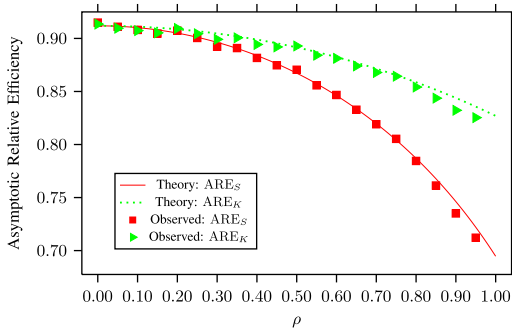
cannot be ordered in a consistent manner. It appears in Fig. 2 that (1)  $\text{MSE}_M > \text{MSE}_K > \text{MSE}_S > \text{MSE}_P$  when  $|\rho|$  is around 0, (2)  $\text{MSE}_S > \text{MSE}_K > \text{MSE}_P$  when  $|\rho|$  exceeds some threshold, which moves towards 0 with increase of  $n$ , and (3) the difference between  $\text{MSE}_K$  and  $\text{MSE}_S$  around  $\rho=0$  decreases steadily with increase of  $n$ . Furthermore, due to the asymptotic equivalence between  $\hat{\rho}_S$  and  $\hat{\rho}_M$ ,  $\text{MSE}_S$  and  $\text{MSE}_M$  becomes closer and closer as  $n$  increases.

**Table 4**  
Variances of  $\mathbb{V}(\hat{\rho}_\zeta)$ ,  $\zeta \in \{P, S, K, M\}$  for  $n = 10, 20, 30$ .

$\rho$	Sample size $n=10$						Sample size $n=20$						Sample size $n=30$					
	$\mathbb{V}(\hat{\rho}_P)$			$\mathbb{V}(\hat{\rho}_S)$			$\mathbb{V}(\hat{\rho}_K)$			$\mathbb{V}(\hat{\rho}_M)$			$\mathbb{V}(\hat{\rho}_P)$			$\mathbb{V}(\hat{\rho}_S)$		
	(53)	Obs.	(54)	(55)	Obs.	(56)	(55)	Obs.	(56)	(55)	Obs.	(56)	(53)	Obs.	(54)	(54)	Obs.	(56)
0.0	0.111	0.111	0.122	0.119	0.133	0.148	0.143	0.143	0.148	0.148	0.143	0.143	0.034	0.034	0.038	0.041	0.039	0.040
0.1	0.109	0.109	0.120	0.117	0.131	0.145	0.141	0.141	0.145	0.145	0.141	0.141	0.034	0.034	0.037	0.040	0.039	0.040
0.2	0.102	0.105	0.115	0.112	0.125	0.138	0.135	0.135	0.138	0.138	0.135	0.135	0.032	0.032	0.035	0.038	0.037	0.038
0.3	0.092	0.096	0.106	0.105	0.117	0.127	0.125	0.125	0.127	0.127	0.125	0.125	0.029	0.029	0.032	0.034	0.034	0.034
0.4	0.078	0.085	0.094	0.094	0.104	0.112	0.111	0.111	0.112	0.112	0.111	0.111	0.024	0.024	0.028	0.030	0.029	0.030
0.5	0.063	0.071	0.080	0.081	0.089	0.093	0.094	0.094	0.093	0.093	0.094	0.094	0.019	0.019	0.023	0.024	0.024	0.025
0.6	0.046	0.055	0.063	0.065	0.071	0.073	0.074	0.074	0.073	0.073	0.074	0.074	0.014	0.014	0.018	0.018	0.018	0.019
0.7	0.029	0.038	0.046	0.048	0.052	0.051	0.053	0.053	0.051	0.051	0.053	0.053	0.009	0.010	0.012	0.012	0.012	0.013
0.8	0.014	0.021	0.028	0.030	0.031	0.029	0.031	0.031	0.029	0.029	0.031	0.031	0.004	0.005	0.007	0.006	0.007	0.007
0.9	0.004	0.007	0.012	0.013	0.012	0.011	0.012	0.012	0.011	0.011	0.012	0.012	0.001	0.001	0.002	0.002	0.002	0.002



**Fig. 2.** Comparison of observed  $MSE_{\zeta}$ ,  $\zeta \in \{P, S, K, M\}$  for (a)  $n=10$ , (b)  $n=20$ , (c)  $n=40$ , and (d)  $n=60$  over  $\rho = -1(0.1)1$ , respectively.



**Fig. 3.** Verification and comparison of ARE<sub>S</sub> and ARE<sub>K</sub> for  $n=1000$  over  $\rho = 0(0.01)1$ , for theoretical curves, and  $\rho = 0(0.05)0.95$ , for simulation results. The results (60) and (61) are used to plot the two theoretical curves for  $\rho = 1$ , respectively.

#### 6.4. Verification and comparison of ARE<sub>S</sub> and ARE<sub>K</sub>

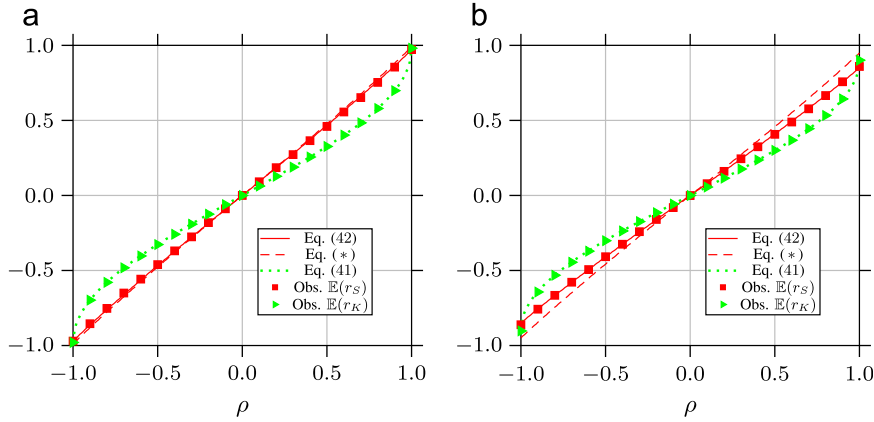
Fig. 3 verifies and compares the performance of  $\hat{\rho}_S$  and  $\hat{\rho}_K$  in terms of ARE. For purpose of numerical verification, simulation results for  $n=1000$  are superimposed upon the corresponding theoretical curves. Due to the asymptotic equivalence between  $\hat{\rho}_S$  and  $\hat{\rho}_M$ , the results with respect to ARE<sub>M</sub> are not included in Fig. 3. It can be observed that (1) the simulations agree well with our theoretical findings in (58) and (59), respectively, (2) ARE<sub>K</sub> lies consistently above ARE<sub>S</sub>, indicating the superiority of  $\hat{\rho}_K$  over  $\hat{\rho}_S$  for large samples, and (3) the performance of  $\hat{\rho}_S$  deteriorates severely as  $\rho$  approaching

unity, although it performs similarly as  $\hat{\rho}_K$  when  $\rho$  is small. Note that the remarks on ARE<sub>S</sub> also apply to ARE<sub>M</sub> due to the equivalence between  $\hat{\rho}_S$  and  $\hat{\rho}_M$  when the sample size  $n$  is large.

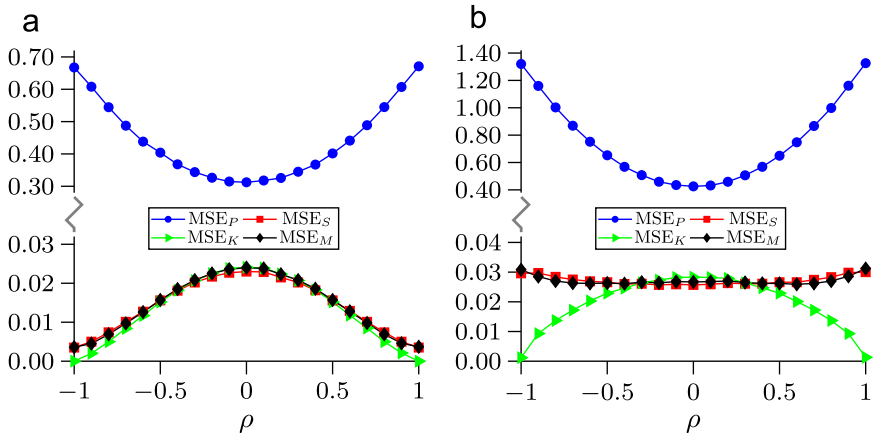
#### 6.5. Performance of $\hat{\rho}_{\zeta}$ in contaminated normal model

Fig. 4 supports to (1) verify the two statements concerning  $\mathbb{E}(r_K)$  and  $\mathbb{E}(r_S)$  in Theorem 1 under the contaminated Gaussian model (36) and (2) compare our formula (38) with the result of (\*) that asserted in [28]. Due to the lack of space, we only present the results for  $\epsilon = 0.01$  and  $\epsilon = 0.05$  under the sample size  $n=50$  here. For simplicity, the rest parameters of the model (36) are set to be  $\sigma_X = \sigma_Y = 1$ ,  $\lambda_X = \lambda_Y = 100$  and  $\rho' = 0$  throughout. It is seen that the observed values of  $\mathbb{E}(r_K)$  and  $\mathbb{E}(r_S)$  agree well with the corresponding theoretical results of (37) and (38) established in Theorem 1. On the other hand, however, the curves with respect to (\*), especially in Fig. 4(b), deviate obviously from the corresponding observed values.

Fig. 5 illustrates, in terms of MSE, the sensitivity of  $\hat{\rho}_p$  as well as the robustness of  $\hat{\rho}_S$ ,  $\hat{\rho}_K$  and  $\hat{\rho}_M$  to impulsive noise. It is shown in Fig. 5 that the MSE of  $\hat{\rho}_p$  is dramatically larger than those of the other three estimators, irrespective of how small the fraction  $\epsilon$  of impulsive component is. On the other hand, it is seen that, despite some minor negative (positive) differences for  $\rho$  around 0 ( $\pm 1$ ), MSE<sub>S</sub> and MSE<sub>M</sub> behave similarly with MSE<sub>K</sub> for  $\epsilon = 0.01$ . Nevertheless, MSE<sub>S</sub> and MSE<sub>M</sub> are much larger than MSE<sub>K</sub> for  $\epsilon = 0.05$  when  $\rho$  falls in the neighborhood of  $\pm 1$ . Combining Fig. 5(a) and (b), it would be reasonable to rank their performance as



**Fig. 4.** Verification of Theorem 1 for (a)  $\epsilon = 0.01$  and (b)  $\epsilon = 0.05$  under the sample size  $n = 50$  over  $\rho = (-1)0.1(1)$ , for simulations, and  $\rho = (-1)0.01(1)$ , for theoretical formulae (37) and (38). The rest parameters of the model (36) are set to be  $\sigma_X = \sigma_Y = 1$ ,  $\lambda_X = \lambda_Y = 100$  and  $\rho' = 0$ , respectively. The formula (\*) concerning  $\mathbb{E}(r_S)$  developed elsewhere [28] is also included for comparison.



**Fig. 5.** Performance comparison in terms of  $\text{MSE}_\zeta$ ,  $\zeta \in \{S, P, K, M\}$ , over  $\rho = -1(0.1)1$  in the contaminated normal model (36) for (a)  $\epsilon = 0.01$  and (b)  $\epsilon = 0.05$  under the sample size  $n = 50$ . The rest parameters of the model are set to be  $\sigma_X = \sigma_Y = 1$ ,  $\lambda_X = \lambda_Y = 100$  and  $\rho' = 0$ , respectively.

$\hat{\rho}_K \geq \hat{\rho}_S \sim \hat{\rho}_M \gg \hat{\rho}_P$  in terms of MSE under the contaminated normal model (36), where the symbol  $\sim$  stands for “is similar to”.

## 7. Concluding remarks

In this paper we have investigated systematically the properties of Spearman's rho and Kendall's tau for samples drawn from bivariate normal contained normal populations. Theoretical derivations along with Monte Carlo simulations reveal that, contrary to the opinion of equivalence between SR and KT in some literature, e.g. [46], they behave quite differently in terms of *mathematical tractability*, *bias effect*, *mean square error*, *asymptotic relative efficiency* in the normal cases and *robustness* properties in the contaminated normal model.

As shown in Lemma 3, SR is mathematically less tractable than KT, in the sense of the intractable terms  $\Omega_1(\rho)$  and  $\Omega_2(\rho)$  in the formula of its variance (17), in contrast with the closed form expression of  $\mathbb{V}(r_K)$  in (16). However, this mathematical inconvenience is, to some extent, offset by Table 2 provided in this work, especially

from the viewpoint of numerical accuracy. Moreover, as demonstrated in Fig. 1 and Table 4, the convergence speed of the asymptotic formulae (46) and (53) with respect to  $\text{BIAS}_K$  and  $\mathbb{V}(\hat{\rho}_K)$  are slower than those of  $\text{BIAS}_S$  and  $\mathbb{V}(\hat{\rho}_S)$  due to the high nonlinearity of the calibration (41). As a consequence, we do not attach too much importance to such mathematical advantage of KT over SR.

Now let us turn back to the question raised at the very beginning of this paper: Which one, SR or KT, should we use in practice when PPMCC is inapplicable? The answer to this question is different for different requirements of the task at hand. Specifically,

1. If *unbiasedness* is on the top priority list, then neither  $\hat{\rho}_S$  nor  $\hat{\rho}_K$  should be resorted to. The modified version  $\hat{\rho}_M$  that employs both SR and KT is definitely the best choice (cf. Fig. 1).
2. On the other hand, if minimal MSE is the critical feature and the sample size  $n$  is small, then  $\hat{\rho}_S$  ( $\hat{\rho}_K$ ) should be employed when the population correlation  $\rho$  is weak (strong) (cf. Fig. 2).

3. Since  $\hat{\rho}_K$  outperforms  $\hat{\rho}_S$  asymptotically in terms of ARE, then  $\hat{\rho}_K$  is the suitable statistic in large-sample cases (cf. Fig. 3).
4. If there is impulsive noise in the data, then it would be better to employ  $\hat{\rho}_K$ , in terms of MSE, although there is some minor advantage of  $\hat{\rho}_S$  when  $\rho$  is in the neighborhood of 0 (cf. Fig. 5).
5. Moreover, in terms of time complexity,  $\hat{\rho}_S$  appears to be superior to  $\hat{\rho}_K$ —the computational load of the former is  $O(n \log n)$ ; whereas the computational load of the latter is  $O(n^2)$  [12].

Possessing the desirable properties summarized in Section 2, Spearman's rho and Kendall's tau have found wide applications in the literature other than signal processing. With the new insights uncovered in this paper, these two rank based coefficients can play complementary roles under the circumstances where Pearson's product moment correlation coefficient is no longer effective.

## Acknowledgments

This work was supported in part by 100-Talents Scheme Funding from Guangdong University of Technology under Grant 112418006, in part by the Talent Introduction Special Funds from Guangdong Province under Grant 2050205, in part by a team project from Guangdong University of Technology under Grant GDUT2011-07, and in part by National Natural Science Foundation of China under Project 61271380.

## Appendix A

### Proof of Lemma 4. Let

$$\begin{aligned} S &\triangleq \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n H(X_i - X_j) H(Y_i - Y_k) \\ &= \sum_{i \neq j=1}^n H(X_i - X_j) H(Y_i - Y_j) \\ &\quad + \sum_{i \neq j \neq k=1}^n H(X_i - X_j) H(Y_i - Y_k) \end{aligned} \quad (\text{A.1})$$

and  $\mathcal{T}$  be the numerator of (3). Taking the expectation of both sides of (A.1) with the assistance of (6) in Lemma 2, it follows readily that

$$\mathbb{E}(S) = n^{[2]} \left( \frac{1}{4} + \frac{S_1}{2\pi} \right) + n^{[3]} \left( \frac{1}{4} + \frac{S_2}{2\pi} \right) \quad (\text{A.2})$$

where  $n^{[\kappa]} \triangleq n(n-1) \cdots (n-\kappa+1)$ , with  $\kappa$  being a positive integer. Define

$$I \triangleq \sum_{i \neq j} H(X_i - X_j) H(Y_i - Y_j) \quad (\text{A.3})$$

$$J \triangleq \sum_{i \neq j \neq k} H(X_i - X_j) H(Y_i - Y_k) \quad (\text{A.4})$$

$$K \triangleq \sum_{i \neq j} H(X_i - X_j) \quad (\text{A.5})$$

$$L \triangleq \sum_{i \neq k} H(Y_i - Y_k) \quad (\text{A.6})$$

Then, we have, from (3), (B.9) and (A.1) along with the relationship  $\text{sgn}(\cdot) = 2H(\cdot) - 1$ ,

$$S = I + J \quad (\text{A.7})$$

$$\mathcal{T} = 4I - 2K - 2L + n^{[2]} \quad (\text{A.8})$$

and hence

$$\begin{aligned} \mathbb{C}(r_S, r_K) &= \frac{12}{n^2(n-1)(n^2-1)} \mathbb{C}(S, \mathcal{T}) \\ &= \frac{12}{n^2(n-1)(n^2-1)} [\mathbb{E}(ST) - \mathbb{E}(S)\mathbb{E}(\mathcal{T})] \end{aligned} \quad (\text{A.9})$$

From (5) and (6), it follows that

$$\mathbb{E}(I) = n^{[2]} \left( \frac{1}{4} + \frac{S_1}{2\pi} \right) \quad \text{and} \quad \mathbb{E}(K) = \mathbb{E}(L) = \frac{n^{[2]}}{2}$$

Substituting these expectation terms into (A.8) gives

$$\mathbb{E}(\mathcal{T}) = 4n^{[2]} \left( \frac{1}{4} + \frac{S_1}{2\pi} \right) - n^{[2]} = \frac{2n^{[2]}}{\pi} S_1 \quad (\text{A.10})$$

Recall that we have obtained  $\mathbb{E}(S)$  in (A.2). Now the only difficulty lies in the evaluation of  $\mathbb{E}(ST)$  in (A.9). Multiplying (A.7) and (A.8), expanding and taking expectations term by term, we have

$$\begin{aligned} \mathbb{E}(ST) &= 4\mathbb{E}(IJ) - 2\mathbb{E}(KJ) - 2\mathbb{E}(LJ) \\ &\quad + 4\mathbb{E}(I^2) - 2\mathbb{E}(KI) - 2\mathbb{E}(LI) + n^{[2]}\mathbb{E}(S) \end{aligned} \quad (\text{A.11})$$

Now, resorting to Table 5, we are ready to evaluate the first six terms in (A.11). From (A.3) and (A.4), it follows that  $\mathbb{E}(IJ)$  is a summation of  $P_4^0$  terms of the form

$$\mathbb{E}\{ \underbrace{H(X_i - X_j)}_{Z_1} \underbrace{H(Y_i - Y_j)}_{Z_2} \underbrace{H(X_k - X_l)}_{Z_3} \underbrace{H(Y_k - Y_m)}_{Z_4} \}. \quad (\text{A.12})$$

Since, by (4),  $H(0) = 0$ , the term (A.12) vanishes for  $i=j$  or  $k=l$  or  $k=m$ . Then there are  $n^2(n-1)^2(n-2)$  nontrivial (A.12)-like terms left to be evaluated. It follows that the domain of the quintuple  $(i, j, k, l, m)$  can be partitioned into 13 disjoint and exhaustive subsets whose representative terms,  $Z_1, Z_2, Z_3, Z_4$ , are listed in the upper panel of Table 5. Summing up the corresponding  $P_4^0$ -terms in Table 5 leads directly to  $\mathbb{E}(IJ)$ . In a similar manner we can obtain  $\mathbb{E}(KJ)$  and  $\mathbb{E}(LJ)$ . With the assistance of the lower panel of Table 5, we also have the expressions of  $\mathbb{E}(I^2)$ ,  $\mathbb{E}(KI)$  and  $\mathbb{E}(LI)$ . Substituting these results and (A.2) into (A.11), subtracting the multiplication of (A.2) and (A.10) and substituting the resultant back into (A.9), we find that  $\mathbb{C}(r_S, r_K)$  is of the form (24) with

$$\Omega_3(\rho) = \frac{1}{4} W_g + \frac{1}{4} W_p + \frac{1}{2} W_h + \frac{1}{2} W_q$$

which simplifies to (26) by applying the identities

$$W_e = 2W_d, \quad W_g = W_p, \quad W_h = W_q, \quad \text{and} \quad W_m = 2W_l + \frac{1}{3}$$

obtained from an application of the relationship (8) to Appendix 2 of [45]. The lemma then follows.  $\square$

**Table 5**Quantities for evaluation of  $\mathbb{E}(ST)$  in Lemma 4.

Corr. matrix	No. of terms <sup>a</sup>	Representative term				$P_4^0(Z_1, Z_2, Z_3, Z_4)^b$	$P_3^0(Z_2, Z_3, Z_4)^b$	$P_3^0(Z_1, Z_3, Z_4)^a$
		$Z_1$	$Z_2$	$Z_3$	$Z_4$			
$R_b$	$n^{[5]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_4$	$Y_3 - Y_5$	$\frac{1}{16} + \frac{S_1}{8\pi} + \frac{S_2}{8\pi} + \frac{S_1 S_2}{4\pi^2}$	$\frac{1}{8} + \frac{S_2}{4\pi}$	$\frac{1}{8} + \frac{S_2}{4\pi}$
$R_g$	$n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_1 - X_3$	$Y_1 - Y_4$	$\frac{5}{48} + \frac{S_1}{8\pi} + \frac{3S_2}{8\pi} + \frac{W_g(\rho)}{16}$	$\frac{1}{6} + \frac{S_2}{2\pi}$	$\frac{1}{6} + \frac{S_2}{2\pi}$
$R_h$	$n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_4$	$Y_3 - Y_1$	$\frac{1}{24} + \frac{S_1}{8\pi} + \frac{W_h(\rho)}{16}$	$\frac{1}{12} + \frac{S_2}{4\pi}$	$\frac{1}{8}$
$R_h$	$n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_1$	$Y_3 - Y_4$	$\frac{1}{24} + \frac{S_1}{8\pi} + \frac{W_h(\rho)}{16}$	$\frac{1}{12} + \frac{S_2}{4\pi}$	$\frac{1}{8}$
$R_p$	$n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_2 - X_3$	$Y_2 - Y_4$	$\frac{1}{48} + \frac{S_1}{8\pi} + \frac{S_2}{8\pi} + \frac{W_p(\rho)}{16}$	$\frac{1}{12}$	$\frac{1}{12}$
$R_q$	$n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_2$	$Y_3 - Y_4$	$\frac{1}{12} + \frac{S_1}{8\pi} + \frac{S_2}{4\pi} + \frac{W_q(\rho)}{16}$	$\frac{1}{8} + \frac{S_2}{2\pi}$	$\frac{1}{6} + \frac{S_2}{4\pi}$
$R_q$	$n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_4$	$Y_3 - Y_2$	$\frac{1}{12} + \frac{S_1}{8\pi} + \frac{S_2}{4\pi} + \frac{W_q(\rho)}{16}$	$\frac{1}{8} + \frac{S_2}{2\pi}$	$\frac{1}{6} + \frac{S_2}{4\pi}$
$R_u$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_1 - X_2$	$Y_1 - Y_3$	$\frac{1}{6} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$	$\frac{1}{6} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$	$\frac{1}{4} + \frac{S_2}{2\pi}$
$R_u$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_1 - X_3$	$Y_1 - Y_2$	$\frac{1}{6} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$	$\frac{1}{6} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$	$\frac{1}{4} + \frac{S_2}{2\pi}$
$R_v$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_1$	$Y_3 - Y_2$	$\frac{1}{18} + \frac{S_1}{8\pi} + \frac{S_2}{8\pi} + \frac{S_1^2}{8\pi^2} - \frac{S_2^2}{8\pi^2}$	$\frac{1}{6}$	$\frac{1}{12} + \frac{S_2}{2\pi}$
$R_v$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_2$	$Y_3 - Y_1$	$\frac{1}{18} + \frac{S_1}{8\pi} + \frac{S_2}{8\pi} + \frac{S_1^2}{8\pi^2} - \frac{S_2^2}{8\pi^2}$	$\frac{1}{6}$	$\frac{1}{12} + \frac{S_2}{2\pi}$
$R_1$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_2 - X_1$	$Y_2 - Y_3$	0	$\frac{1}{12} - \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$	0
$R_2$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_2 - X_3$	$Y_2 - Y_1$	0	0	$\frac{1}{12} - \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$
$R_j$	$n^{[4]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_4$	$Y_3 - Y_4$	$\frac{1}{16} + \frac{S_1}{4\pi} + \frac{S_1^2}{4\pi^2}$	$\frac{1}{8} + \frac{S_1}{4\pi}$	$\frac{1}{8} + \frac{S_1}{4\pi}$
$R_r$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_1 - X_3$	$Y_1 - Y_3$	$\frac{1}{9} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi} + \frac{S_1^2}{4\pi^2} - \frac{S_2^2}{4\pi^2}$	$\frac{1}{6} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$	$\frac{1}{6} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$
$R_r$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_2$	$Y_3 - Y_2$	$\frac{1}{9} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi} + \frac{S_1^2}{4\pi^2} - \frac{S_2^2}{4\pi^2}$	$\frac{1}{6} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$	$\frac{1}{6} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$
$R_w$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_3 - X_1$	$Y_3 - Y_1$	$\frac{1}{36} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi} + \frac{S_1^2}{4\pi^2} - \frac{S_2^2}{4\pi^2}$	$\frac{1}{12} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$	$\frac{1}{12} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi}$
$R_w$	$n^{[3]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_2 - X_3$	$Y_2 - Y_3$	$\frac{1}{36} + \frac{S_1}{4\pi} + \frac{S_2}{4\pi} + \frac{S_1^2}{4\pi^2} - \frac{S_2^2}{4\pi^2}$	$\frac{1}{12} + \frac{S_1}{4\pi} - \frac{S_2}{4\pi}$	$\frac{1}{12} + \frac{S_1}{4\pi} - \frac{S_2}{4\pi}$
$R_x$	$n^{[2]}$	$X_1 - X_2$	$Y_1 - Y_2$	$X_1 - X_2$	$Y_1 - Y_2$	$\frac{1}{4} + \frac{S_1}{2\pi}$	$\frac{1}{4} + \frac{S_1}{2\pi}$	$\frac{1}{4} + \frac{S_1}{2\pi}$

<sup>a</sup> The symbol  $n^{[\kappa]}$  is a compact notation of  $n(n-1)\cdots(n-\kappa+1)$ , with  $\kappa$  being a positive integer.<sup>b</sup>  $P_4^0(Z_1, Z_2, Z_3, Z_4) \triangleq \mathbb{E}\{H(Z_1)H(Z_2)H(Z_3)H(Z_4)\}$ ,  $P_3^0(Z_2, Z_3, Z_4) \triangleq \mathbb{E}\{H(Z_2)H(Z_3)H(Z_4)\}$ , and  $P_3^0(Z_1, Z_3, Z_4) \triangleq \mathbb{E}\{H(Z_1)H(Z_3)H(Z_4)\}$ . Notations  $S_1 \triangleq \sin^{-1} \rho$  and  $S_2 \triangleq \sin^{-1} \frac{1}{2} \rho$  are used for brevity.

## Appendix B

**Proof of Theorem 1.** For the ease of the following discussion, we will use  $\phi(x, y)$  and  $\psi(x, y)$  to denote the pdfs of the two bivariate normal components in (36), respectively. From [37] and (A.8), it follows that the numerator  $\mathcal{T}$  of (3) can be simplified to

$$\mathcal{T} = 4 \sum_{i \neq j=1}^n \sum_{j=1}^n H(X_i - X_j) H(Y_i - Y_j) - n^{[2]} \quad (\text{B.1})$$

which yields

$$\mathbb{E}(\mathcal{T}) = 4n^{[2]} \underbrace{\mathbb{E}[H(X_1 - X_2)H(Y_1 - Y_2)]}_{E_1} - n^{[2]} \quad (\text{B.2})$$

by the i.i.d. assumption. To evaluate  $E_1$  in (B.2), we need the joint distribution of  $(X_1, Y_1, X_2, Y_2)$ , denoted by  $\varphi(x_1, y_1, x_2, y_2)$ , which is readily obtained as

$$\varphi = [(1-\epsilon)\phi_1 + \epsilon\psi_1][(1-\epsilon)\phi_2 + \epsilon\psi_2]$$

$$= \underbrace{(1-\epsilon)^2}_{\alpha_1} \underbrace{\phi_1 \phi_2}_{\phi_1} + \underbrace{\epsilon(1-\epsilon)}_{\alpha_2} \underbrace{\phi_1 \psi_2}_{\phi_2} + \underbrace{\epsilon(1-\epsilon)}_{\alpha_3} \underbrace{\phi_2 \psi_1}_{\phi_3} + \underbrace{\epsilon^2}_{\alpha_4} \underbrace{\psi_1 \psi_2}_{\phi_4} \quad (\text{B.3})$$

where  $\varphi$ ,  $\phi_i$ ,  $\psi_i$  are compact notations of  $\varphi(x_1, y_1, x_2, y_2)$ ,  $\phi(x_i, y_i)$  and  $\psi(x_i, y_i)$ ,  $i=1, 2$ , respectively. Write

$$U \triangleq \frac{X_1 - X_2}{\sqrt{\mathbb{V}(X_1 - X_2)}} \quad \text{and} \quad V \triangleq \frac{Y_1 - Y_2}{\sqrt{\mathbb{V}(Y_1 - Y_2)}}.$$

Then, with respect to  $\varphi_1$ ,  $\varphi_2$ ,  $\varphi_3$ , and  $\varphi_4$  in (B.3),  $(U, V)$  follows four standard bivariate normal distributions with correlations

$$\varrho_1 = \rho \quad (\text{B.4})$$

$$\varrho_2 = \frac{\rho + \lambda_X \lambda_Y \rho'}{\sqrt{1 + \lambda_X^2} \sqrt{1 + \lambda_Y^2}} \rightarrow \rho' \quad \text{as } \lambda_X, \lambda_Y \rightarrow \infty \quad (\text{B.5})$$

$$\varrho_3 = \frac{\rho + \lambda_X \lambda_Y \rho'}{\sqrt{1 + \lambda_X^2} \sqrt{1 + \lambda_Y^2}} \rightarrow \rho' \quad \text{as } \lambda_X, \lambda_Y \rightarrow \infty \quad (\text{B.6})$$



$$\varrho_4 = \rho' \quad (\text{B.7})$$

respectively. An application of Sheppard's theorem (6) to (B.2) along with (B.4)–(B.7) yields

$$\begin{aligned} \mathbb{E}(\mathcal{T}) &= 4n^{[2]} \sum_{i=1}^4 \alpha_i \left( \frac{1}{4} + \frac{1}{2\pi} \sin^{-1} \varrho_i \right) - n^{[2]} \\ &= \frac{2n^{[2]}}{\pi} [\alpha_1 \sin^{-1} \rho + 2\alpha_2 \sin^{-1} \varrho_2 + \alpha_4 \sin^{-1} \rho']. \end{aligned} \quad (\text{B.8})$$

Now it is not difficult to verify that the first statement (37) holds by (1) dividing both sides of (B.8) by  $n^{[2]}$ , (2) letting  $\lambda_X \rightarrow \infty$  and  $\lambda_Y \rightarrow \infty$ , and (3) ignoring the  $O(\epsilon^2)$  terms.

To prove the second statement (38), it suffices to evaluate  $\mathbb{E}(S)$  by the relationship [37]

$$r_S = \frac{S - \frac{1}{4}(n-1)^2}{n} \frac{12}{n^2 - 1}. \quad (\text{B.9})$$

Taking expectations of both sides in (A.1) along with the i.i.d. assumptions gives

$$\mathbb{E}(S) = n^{[2]} E_1 + n^{[3]} \underbrace{\mathbb{E}[H(X_1 - X_2)H(Y_1 - Y_3)]}_{E_2}. \quad (\text{B.10})$$

Since we have known  $E_1$  in the above development, now we only need to work out  $E_2$  in (B.10). Let  $\varpi(x_1, y_1, x_2, y_2, x_3, y_3)$ , abbreviated as  $\varpi$ , denote the pdf of the joint distribution of  $(X_1, Y_1, X_2, Y_2, X_3, Y_3)$ . Then, from (36) and the i.i.d. assumption,

$$\begin{aligned} \varpi &= [(1-\epsilon)\phi_1 + \epsilon\psi_1][(1-\epsilon)\phi_2 + \epsilon\psi_2][(1-\epsilon)\phi_3 + \epsilon\psi_3] \\ &= (1-\epsilon)^3 \underbrace{\phi_1\phi_2\phi_3}_{\varpi_1} + \epsilon(1-\epsilon)^2 \underbrace{(\phi_1\phi_2\psi_3 + \phi_1\psi_2\phi_3 + \psi_1\phi_2\phi_3)}_{\varpi_2} \\ &\quad + \epsilon^2(1-\epsilon) \underbrace{(\phi_1\phi_2\psi_3 + \psi_1\phi_2\psi_3 + \psi_1\psi_2\phi_3)}_{\varpi_3} + \epsilon^3 \underbrace{\psi_1\psi_2\psi_3}_{\varpi_4} \\ &\quad + \epsilon^2(1-\epsilon) \underbrace{(\phi_1\phi_2\psi_3 + \psi_1\phi_2\psi_3 + \psi_1\psi_2\phi_3)}_{\varpi_5} + \epsilon^3 \underbrace{\psi_1\psi_2\psi_3}_{\varpi_6} \\ &\quad + \epsilon^2(1-\epsilon) \underbrace{(\phi_1\phi_2\psi_3 + \psi_1\phi_2\psi_3 + \psi_1\psi_2\phi_3)}_{\varpi_7} + \epsilon^3 \underbrace{\psi_1\psi_2\psi_3}_{\varpi_8} \end{aligned} \quad (\text{B.11})$$

where  $\phi_i$  and  $\psi_i$  are compact notations of  $\phi(x_i, y_i)$  and  $\psi(x_i, y_i)$ ,  $i = 1, 2, 3$ , respectively. Define

$$V' = \frac{Y_1 - Y_3}{\sqrt{\mathbb{V}(Y_1 - Y_3)}}.$$

Then, with respect to  $\varpi_1$  to  $\varpi_8$  in (B.11),  $(U, V')$  follows eight standard bivariate normal distributions with correlations

$$\varrho_5 = \frac{\rho}{2} \quad (\text{B.12})$$

$$\varrho_6 = \frac{1}{\sqrt{2}} \frac{\rho}{\sqrt{1 + \lambda_Y^2}} \rightarrow 0 \quad \text{as } \lambda_Y \rightarrow \infty \quad (\text{B.13})$$

$$\varrho_7 = \frac{1}{\sqrt{2}} \frac{\rho}{\sqrt{1 + \lambda_X^2}} \rightarrow 0 \quad \text{as } \lambda_X \rightarrow \infty \quad (\text{B.14})$$

$$\varrho_8 = \frac{\lambda_X \lambda_Y \rho'}{\sqrt{1 + \lambda_X^2} \sqrt{1 + \lambda_Y^2}} \rightarrow \rho' \quad \text{as } \lambda_X, \lambda_Y \rightarrow \infty \quad (\text{B.15})$$

$$\varrho_9 = \frac{\rho}{\sqrt{1 + \lambda_X^2} \sqrt{1 + \lambda_Y^2}} \rightarrow 0 \quad \text{as } \lambda_X, \lambda_Y \rightarrow \infty \quad (\text{B.16})$$

$$\varrho_{10} = \frac{1}{\sqrt{2}} \frac{\lambda_X \rho'}{\sqrt{1 + \lambda_X^2}} \rightarrow \frac{\rho'}{\sqrt{2}} \quad \text{as } \lambda_X \rightarrow \infty \quad (\text{B.17})$$

$$\varrho_{11} = \frac{1}{\sqrt{2}} \frac{\lambda_Y \rho'}{\sqrt{1 + \lambda_Y^2}} \rightarrow \frac{\rho'}{\sqrt{2}} \quad \text{as } \lambda_Y \rightarrow \infty \quad (\text{B.18})$$

$$\varrho_{12} = \frac{\rho'}{2} \quad (\text{B.19})$$

Using Sheppard's theorem (6) again together with (B.11)–(B.19), we can obtain the expression of  $E_2$  and hence  $\mathbb{E}(S)$  in terms of  $n$ ,  $\epsilon$  and  $\varrho_1$  to  $\varrho_{12}$ . Substituting  $\mathbb{E}(S)$  into (B.9), letting  $n$ ,  $\lambda_X$ ,  $\lambda_Y \rightarrow \infty$  and ignoring the  $O(\epsilon^2)$  terms, we arrive at (38), the second theorem statement.  $\square$

## References

- [1] P. Schreier, A unifying discussion of correlation analysis for complex random vectors, *IEEE Transactions on Signal Processing* 56 (4) (2008) 1327–1336.
- [2] J. Via, I. Santamaria, Correlation matching approaches for blind OSTBC channel estimation, *IEEE Transactions on Signal Processing* 56 (12) (2008) 5950–5961.
- [3] X. Xu, Z. Ye, Y. Zhang, DOA estimation for mixed signals in the presence of mutual coupling, *IEEE Transactions on Signal Processing* 57 (9) (2009) 3523–3532.
- [4] E. Dougherty, S. Kim, Y. Chen, Coefficient of determination in nonlinear signal processing, *Signal Processing* 80 (10) (2000) 2219–2235.
- [5] M. Barkat, S. Dib, Cfar detection for two correlated targets, *Signal Processing* 61 (3) (1997) 289–295.
- [6] J. Lian, G. Garner, D. Muessig, V. Lang, A simple method to quantify the morphological similarity between signals, *Signal Processing* 90 (2) (2010) 684–688.
- [7] J. Girault, D. Kouamé, A. Ouahabi, Analytical formulation of the fractal dimension of filtered stochastic signals, *Signal Processing* 90 (9) (2010) 2690–2697.
- [8] J. Delmas, H. Abeida, Asymptotic distribution of circularity coefficients estimate of complex random variables, *Signal Processing* 89 (12) (2009) 2670–2675.
- [9] Y. He, T. Gan, W. Chen, H. Wang, Multi-stage image denoising based on correlation coefficient matching and sparse dictionary pruning, *Signal Processing* 92 (1) (2012) 139–149.
- [10] S. Zhou, H. Liu, Y. Zhao, L. Hu, Target spatial and frequency scattering diversity property for diversity mimo radar, *Signal Processing* 91 (2) (2011) 269–276.
- [11] W. Xu, C. Chang, Y.S. Hung, S.K. Kwan, P.C.W. Fung, Order statistic correlation coefficient and its application to association measurement of biosignals, in: *Proceedings of the International Conference on Acoustics, Speech, Signal Processing (ICASSP)*, vol. 2, 2006, pp. II-1068–II-1071.
- [12] W. Xu, C. Chang, Y.S. Hung, S.K. Kwan, P.C.W. Fung, Order statistics correlation coefficient as a novel association measurement with applications to biosignal analysis, *IEEE Transactions on Signal Processing* 55 (12) (2007) 5552–5563.
- [13] W. Xu, C. Chang, Y.S. Hung, P.C.W. Fung, Asymptotic properties of order statistics correlation coefficient in the normal cases, *IEEE Transactions on Signal Processing* 56 (6) (2008) 2239–2248.
- [14] W. Xu, Y.S. Hung, M. Niranjana, M. Shen, Asymptotic mean and variance of Gini correlation for bivariate normal samples, *IEEE Transactions on Signal Processing* 58 (2) (2010) 522–534.
- [15] M. Greco, F. Gini, A. Farina, Radar detection and classification of jamming signals belonging to a cone class, *IEEE Transactions on Signal Processing* 56 (5) (2008) 1984–1993.
- [16] F. Gini, A radar application of a modified Cramer–Rao bound: parameter estimation in non-Gaussian clutter, *IEEE Transactions on Signal Processing* 46 (7) (1998) 1945–1953.
- [17] F. Gini, J. Michels, Performance analysis of two covariance matrix estimators in compound-Gaussian clutter, *IEE Proceedings Part-F* 146 (3) (1999) 133–140.
- [18] A. Stuart, J.K. Ord, *Kendall's Advanced Theory of Statistics: Classical Inference and Relationship*, 5th edition, vol. 2, Edward Arnold, London, 1991.



- [19] J.D. Gibbons, S. Chakraborti, Nonparametric Statistical Inference, 3rd edition, Marcel Dekker, New York, 1992.
- [20] R.A. Fisher, On the 'probable error' of a coefficient of correlation deduced from a small sample, *Metron* 1 (1921) 3–32.
- [21] R.A. Fisher, Statistical Methods, Experimental Design, and Scientific Inference, Oxford University Press, New York, 1990.
- [22] M. Kendall, J.D. Gibbons, Rank Correlation Methods, 5th edition, Oxford University Press, New York, 1990.
- [23] D.D. Mari, S. Kotz, Correlation and Dependence, Imperial College Press, London, 2001.
- [24] S. Tumanski, Principles of Electrical Measurement, Taylor & Francis, New York, 2006.
- [25] D. Stein, Detection of random signals in Gaussian mixture noise, *IEEE Transactions on Information Theory* 41 (6) (1995) 1788–1801.
- [26] R. Chen, X. Wang, J. Liu, Adaptive joint detection and decoding in flat-fading channels via mixture Kalman filtering, *IEEE Transactions on Information Theory* 46 (6) (2000) 2079–2094.
- [27] Z. Reznic, R. Zamir, M. Feder, Joint source-channel coding of a Gaussian mixture source over the Gaussian broadcast channel, *IEEE Transactions on Information Theory* 48 (3) (2002) 776–781.
- [28] G.L. Shevlyakov, N.O. Vilchevski, Robustness in Data Analysis: Criteria and Methods, Modern Probability and Statistics, VSP, Utrecht, 2002.
- [29] C. Croux, C. Dehon, Influence functions of the Spearman and Kendall correlation measures, *Statistical Methods & Applications* 19 (4) (2010) 497–515.
- [30] D. Ruchkin, Error of correlation coefficient estimates from polarity coincidences (corresp.), *IEEE Transactions on Information Theory* 11 (2) (1965) 296–297.
- [31] M.C. Cheng, The clipping loss in correlation detectors for arbitrary input signal-to-noise ratios, *IEEE Transactions on Information Theory* 14 (3) (1968) 382–389.
- [32] H. Chadwick, L. Kurz, Rank permutation group codes based on Kendall's correlation statistic, *IEEE Transactions on Information Theory* 15 (2) (1969) 306–315.
- [33] V. Hansen, Detection performance of some nonparametric rank tests and an application to radar, *IEEE Transactions on Information Theory* 16 (3) (1970) 309–318.
- [34] F. Esscher, On a method of determining correlation from the ranks of variates, *Skandinavisk Aktuarietidskrift* 7 (2) (1924) 201–219.
- [35] M.G. Kendall, S.F.H. Kendall, B.B. Smith, The distribution of Spearman's coefficient of rank correlation in a universe in which all rankings occur an equal number of times, *Biometrika* 30 (3/4) (1939) 251–273.
- [36] H.E. Daniels, The relation between measures of correlation in the universe of sample permutations, *Biometrika* 33 (2) (1944) 129–135.
- [37] P.A.P. Moran, Rank correlation and product-moment correlation, *Biometrika* 35 (1/2) (1948) 203–206.
- [38] W. Hoeffding, A class of statistics with asymptotically normal distribution, *Annals of Mathematical Statistics* 19 (3) (1948) 293–325.
- [39] M.G. Kendall, Rank and product-moment correlation, *Biometrika* 36 (1/2) (1949) 177–193.
- [40] H.E. Daniels, Rank correlation and population models, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 12 (2) (1950) 171–191.
- [41] S.T. David, M.G. Kendall, A. Stuart, Some questions of distribution in the theory of rank correlation, *Biometrika* 38 (1/2) (1951) 131–140.
- [42] J. Durbin, A. Stuart, Inversions and rank correlation coefficients, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 13 (2) (1951) 303–309.
- [43] H. Hotelling, New light on the correlation coefficient and its transforms, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 15 (2) (1953) 193–232.
- [44] E.C. Fieller, H.O. Hartley, E.S. Pearson, Tests for rank correlation coefficients. i, *Biometrika* 44 (3/4) (1957) 470–481.
- [45] F.N. David, C.L. Mallows, The variance of Spearman's rho in normal samples, *Biometrika* 48 (1/2) (1961) 19–28.
- [46] A.R. Gilpin, Table for conversion of Kendall's tau to Spearman's rho within the context of measures of magnitude of effect for meta-analysis, *Educational and Psychological Measurement* 53 (1) (1993) 87–92.
- [47] R.J. Serfling, Approximation Theorems of Mathematical Statistics, Wiley Series in Probability and Mathematical Statistics, Wiley New York, 2002.
- [48] A. Stuart, J.K. Ord, S.F. Arnold, Kendall's Advanced Theory of Statistics: Distribution Theory, 6th edition, vol. 1, Edward Arnold, London, 1991.
- [49] S.S. Gupta, Probability integrals of multivariate normal and multivariate t, *Annals of Mathematical Statistics* 34 (3) (1963) 792–828.
- [50] D.R. Childs, Reduction of the multivariate normal integral to characteristic form, *Biometrika* 54 (1/2) (1967) 293–300.
- [51] F. Esscher, On a method of determining correlation from the ranks of the variates, *Skandinavisk Aktuarietidskrift* 7 (1924) 201–219.
- [52] C.B. Borkowf, Computing the nonnull asymptotic variance and the asymptotic relative efficiency of Spearman's rank correlation, *Computational Statistics & Data Analysis* 39 (3) (2002) 271–286.