A Comparative Analysis of Spearman’s rho and Kendall’s tau in Normal and Contaminated Normal Models

Weichao Xu\textsuperscript{a,*}, Yunhe Hou\textsuperscript{b}, Y. S. Hung\textsuperscript{b}, Yuexian Zou\textsuperscript{c}

\textsuperscript{a}Department of Automatic Control, Faculty of Automation, Guangdong University of Technology, No. 100 Waihuan Xi Road, Guangzhou Higher Education Mega Center, Panyu District, Guangzhou, Guangdong, P. R. China 510006

\textsuperscript{b}Department of Electrical & Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong

\textsuperscript{c}Advanced Digital Signal Processing Lab, Peking University Shenzhen Graduate School, Shenzhen, Guangdong, P. R. China 518055

Abstract

This paper analyzes the performances of the Spearman’s rho (SR) and Kendall’s tau (KT) with respect to samples drawn from bivariate normal and contaminated normal populations. The exact analytical formulae of the variance of SR and covariance between SR and KT are obtained based on the Childs’s reduction formula for the quadrivariate normal positive orthant probabilities. Asymptotic closed-form expectations with respect to SR and KT are established under the bivariate contaminated normal models. The bias, mean square error (MSE) and asymptotic relative efficiency (ARE) of the three estimators based on SR and KT to the Pearson’s product moment correlation coefficient (PPMCC) are investigated in both the normal and contaminated normal models. 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, and asymptotic relative efficiency. 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 the PPMCC fails to apply.

Keywords: Bivariate normal, Correlation theory, Kendall’s tau, Pearson’s product moment correlation coefficient, Spearman’s rho

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 \cite{1–17}. Being interpreted as the strength of statistical relationship between two random variables \cite{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 the 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 asymptotic unbiased estimator of the population correlation \( \rho \), and 2) the variance of PPMCC approaches the Cramer-Rao lower bound (CRLB) with increase of the sample size [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 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]. Moreover, SR and KT are much more suitable for hardware implementation since the corresponding algorithms involve only whole numbers [cf. Eq. (2) and Eq. (3)]. Owing to these desired properties, SR and KT have drawn attentions of signal processing practitioners [29–32] in the late 1960s shortly after their being thoroughly studied in the literature of statistics [33–44]. However, despite the rich history of SR and KT, it is still unclear which one should we use in Scenarios 1 to 3 where the familiar PPMCC is inapplicable, even under the fundamental binormal model. Some researchers, such as Fieller et al [43], preferred KT to SR based on empirical evidences; while others, such as Gilpin [45], 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. Suppose that $X_j$ is at the $k$th position in the sorted sequence $X(1) < \cdots < X(n)$. The number $k$ 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 $Q_j$. 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 $\cdot$. The three well known classical correlation coefficient, PPMCC ($r_P$), SR ($r_S$), and KT ($r_K$), are then defined as follows [19]:

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

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

$$r_K(X,Y) \triangleq \frac{\sum_{i \neq j} \text{sgn}(X_i - X_j) \text{sgn}(Y_i - Y_j)}{n(n-1)}.$$  

To ease the following discussion, we will employ the symbol $r_\eta(X,Y), \eta \in \{P, S, K\}$ as a compact notation for the three coefficients. For brevity, the arguments of $r_\eta(X,Y)$ 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. $r_\eta(X,Y) \in [-1, 1]$ for all $(X,Y)$ (standardization);
2. $r_\eta(X,Y) = r_\eta(Y,X)$ (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. $r_\eta(+,+)=r_\eta(-,+)=r_\eta(+,-)=r_\eta(-,-)$;
7. \( r_n \) converges to normal distribution when the sample size \( n \) is large.

Note that the first six properties are discussed in [19] and [23], and the last property follows from the asymptotic theory of \( U \)-statistics established by Hoeffding [37].

2.3. Relationships Among PPMCC, SR and KT

From their expressions (1)–(3), it appears that the three coefficients PPMCC, SR and KT are quite different. However, as demonstrated below, these three coefficients are closely related with each other.

2.3.1. Daniel’s Generalized Correlation Coefficient

Consider the \( n \) data pairs \((X_i, Y_i), i = 1, \ldots, n\), at hand. To each pair of \( X \)’s, \((X_i, X_j)\), we can allot a score \( a_{ij} \) such that \( a_{ij} = -a_{ji} \) and \( a_{ii} = 0 \). In a similar manner, we can also allot a score \( b_{ij} \) to the ordered pair of \( Y \)’s, \((Y_i, Y_j)\). The Daniel’s generalized coefficient \( \Gamma \) is then defined by [35]

\[
\Gamma \triangleq \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} b_{ij}}{\left( \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 \frac{n}{2} \right)^{\frac{1}{2}}}.
\]

This general setup covers PPMCC, SR and KT as special cases with respect to different systems of scores [35]:

- Replacing \( a_{ij} \) by \( X_j - X_i \) and \( b_{ij} \) by \( Y_j - Y_i \) in (4) gives the PPMCC \( r_P \) defined in (1);
- Replacing \( a_{ij} \) by \( P_j - P_i \) and \( b_{ij} \) by \( Q_j - Q_i \) in (4) gives the SR \( r_S \) defined in (2);
- Replacing \( a_{ij} \) by sgn\((X_j - X_i)\) and \( b_{ij} \) by sgn\((Y_j - Y_i)\) in (4) gives the KT \( r_K \) defined in (3).

2.3.2. Inequalities between SR and KT

It is possible to state certain inequalities connecting the values of SR and KT based on a given set of \( n \) observations. The first one, ascribed to Daniel [39], is

\[
-1 \leq \frac{3(n + 2)}{n - 2} r_K - 2 \frac{n + 1}{n - 2} r_S \leq 1
\]

which, for large \( n \), becomes

\[
-1 \leq 3r_K - 2r_S \leq 1.
\]

The second one, due to Durbin and Stuat [41], states that

\[
r_S \leq 1 - \frac{1 - r_K}{2(n + 1)} \left[ (n - 1)(1 - r_K) + 4 \right].
\]
Combing (5) and (6) and letting \( n \to \infty \) yield the bounds of SR, in terms of KT, as
\[
\begin{align*}
\frac{3}{2} r_K - \frac{1}{2} & \leq r_S \leq \frac{1}{2} + r_K - \frac{1}{2} r_K^2, \quad r_K \geq 0 \\
\frac{3}{2} r_K + \frac{1}{2} & \geq r_S \geq \frac{1}{2} r_K^2 + r_K - \frac{1}{2}, \quad r_K \leq 0.
\end{align*}
\]

2.3.3. Relationship of SR to Other Coefficients

Besides the PPMCC and KT, SR is also closely related to other correlation coefficients, e.g., the order statistics correlation coefficient (OSCC) [11–13] and the Gini correlation (GC) [46]. In fact, SR can be reduced from the OSCC and GC by replacing the variates with corresponding ranks [14].

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 Lemma 3 and Lemma 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 \( E(\mathbf{\Lambda}) \), \( V(\mathbf{\Lambda}) \), \( C(\mathbf{\Lambda}, \mathbf{\Phi}) \), and \( \text{corr}(\mathbf{\Lambda}, \mathbf{\Phi}) \) 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 \( u(\mathbf{\Lambda}) \) and \( v(\mathbf{\Lambda}) \) as the argument \( \mathbf{\Lambda} \) tends to a limit \( L \) (might be infinite). The notation \( u(\mathbf{\Lambda}) = O(v(\mathbf{\Lambda})) \), \( \mathbf{\Lambda} \to L \), denotes that \( |u(\mathbf{\Lambda})|/v(\mathbf{\Lambda})| \) remains bounded as \( \mathbf{\Lambda} \to L \); whereas the notation \( u(\mathbf{\Lambda}) = o(v(\mathbf{\Lambda})) \), \( \mathbf{\Lambda} \to L \), denotes that \( u(\mathbf{\Lambda})/v(\mathbf{\Lambda}) \to 0 \) as \( \mathbf{\Lambda} \to L \) [47]. Symbols of \( P_0^m(Z_1, \ldots, Z_m) \) are adopted to denote the positive orthant probabilities associated with multivariate normal random vectors \( [Z_1 \cdots Z_m] \) of dimensions \( m = 1, \ldots, 4 \), respectively. The notation \( R(\varrho_{rs})_{m \times m} \) stands for correlation matrix with each element \( \varrho_{rs} \equiv \text{corr}(Z_r, Z_s) \), \( r, s = 1, \ldots, m \). Obviously the diagonal entries in \( R \) are all unities. For compactness, we will also use the symbol \( P_0^m(R) \) to denote \( P_0^m(Z_1, \ldots, Z_m) \) in the sequel.

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
\[
H(\mathbf{\Lambda}) \triangleq \begin{cases} 
1 & (\mathbf{\Lambda} > 0) \\
0 & (\mathbf{\Lambda} \leq 0).
\end{cases}
\]

5
Then the orthant probabilities

\[ P_1^0(Z_1) \triangleq \mathbb{E} \{ H(Z_1) \} \]
\[ = \frac{1}{2} \tag{8} \]
\[ P_2^0(Z_1, Z_2) \triangleq \mathbb{E} \{ H(Z_1)H(Z_2) \} \]
\[ = \frac{1}{4} \left( 1 + \frac{2}{\pi} \sin^{-1} \varphi_{12} \right) \tag{9} \]
\[ P_3^0(Z_1, Z_2, Z_3) \triangleq \mathbb{E} \{ H(Z_1)H(Z_2)H(Z_3) \} \]
\[ = \frac{1}{8} \left( 1 + \frac{2}{\pi} \sum_{r=1}^{2} \sum_{s=r+1}^{3} \sin^{-1} \varphi_{rs} \right) \tag{10} \]
\[ P_4^0(Z_1, Z_2, Z_3, Z_4) \triangleq \mathbb{E} \{ H(Z_1)H(Z_2)H(Z_3)H(Z_4) \} \]
\[ = \frac{1}{16} \left( 1 + \frac{2}{\pi} \sum_{r=1}^{3} \sum_{s=r+1}^{4} \sin^{-1} \varphi_{rs} + W \right) \tag{11} \]

where

\[ W \triangleq \frac{1}{\pi^2} \iiint_{-\infty}^{\infty} \frac{\exp \left( -\frac{1}{2} z^T R z \right)}{z_1 z_2 z_3 z_4} \, dz_1 \, dz_2 \, dz_3 \, dz_4 \tag{12} \]
\[ = \sum_{\ell=2}^{4} \frac{4}{\pi^2} \int_{0}^{1} \frac{\varphi_{1\ell}}{[1 - \varphi_{1\ell} u^2]^2} \sin^{-1} \left( \frac{\alpha_\ell(u)}{\beta_\ell(u) \gamma_\ell(u)} \right) \, du \tag{13} \]

with

\[ \alpha_2 = \varphi_{34} - \varphi_{23} \varphi_{24} - [\varphi_{13} \varphi_{14} + \varphi_{12} (\varphi_{12} \varphi_{34} - \varphi_{14} \varphi_{23} - \varphi_{13} \varphi_{24})] u^2 \]
\[ \alpha_3 = \varphi_{24} - \varphi_{23} \varphi_{34} - [\varphi_{12} \varphi_{14} + \varphi_{13} (\varphi_{13} \varphi_{24} - \varphi_{14} \varphi_{23} - \varphi_{12} \varphi_{34})] u^2 \]
\[ \alpha_4 = \varphi_{23} - \varphi_{24} \varphi_{34} - [\varphi_{12} \varphi_{13} + \varphi_{14} (\varphi_{14} \varphi_{23} - \varphi_{13} \varphi_{24} - \varphi_{12} \varphi_{34})] u^2 \]
\[ \beta_2 = \beta_3 = \left[ 1 - \varphi_{23}^2 - (\varphi_{12}^2 + \varphi_{13}^2 - 2 \varphi_{12} \varphi_{13} \varphi_{23}) u^2 \right]^{1/2} \]
\[ \gamma_2 = \gamma_4 = \left[ 1 - \varphi_{24}^2 - (\varphi_{12}^2 + \varphi_{14}^2 - 2 \varphi_{12} \varphi_{14} \varphi_{24}) u^2 \right]^{1/2} \]
\[ \gamma_3 = \gamma_4 = \left[ 1 - \varphi_{34}^2 - (\varphi_{13}^2 + \varphi_{14}^2 - 2 \varphi_{13} \varphi_{14} \varphi_{34}) u^2 \right]^{1/2} \]

Proof. The first statement (8) is trivial. The second one (9) is usually called Sheppard’s theorem in the literature, although it was proposed earlier by Stieltjes [48]. The third one (10) 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 (11) is due to Childs [50] and is termed the Childs’s reduction formula throughout. □

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 \( \rho \). Let \( r_p, r_s \) and \( r_k \) be the PPMCC, SR and KT that defined in (1)-(3), respectively. Write
$S_1 \triangleq \sin^{-1} \rho$ and $S_2 \triangleq \sin^{-1} \frac{1}{2} \rho$. Then
\[
\mathbb{E}(r_P) = \rho \left[ 1 - \frac{1 - \rho^2}{2n} + O \left( \frac{1}{n^2} \right) \right] \to \rho \text{ as } n \to \infty \tag{14}
\]
\[
\mathbb{V}(r_P) = \frac{(1 - \rho^2)^2}{n - 1} + O \left( \frac{1}{n^2} \right) \tag{15}
\]
\[
\mathbb{E}(r_S) = \frac{6}{\pi(n + 1)} \left[ \sin^{-1} \rho + (n - 2) \sin^{-1} \frac{\rho}{2} \right] \to \frac{6}{\pi} \sin^{-1} \frac{\rho}{2} \text{ as } n \to \infty \tag{16}
\]
\[
\mathbb{E}(r_K) = \frac{2}{\pi} \sin^{-1} \rho \tag{17}
\]
\[
\mathbb{V}(r_K) = \frac{2}{n(n - 1)} \left[ 1 - \frac{4S_1^2}{\pi^2} + 2(n - 2) \left( \frac{1}{9} - \frac{4S_2^2}{\pi^2} \right) \right]. \tag{18}
\]

Proof. The first three results, (14)–(16), were given by Hotelling [42], Fisher [21], and Moran [36], respectively; whereas the last two results, (18) and (19), were derived by Esscher [51]. \hfill \square

3.3. Variance of $r_S$

**Lemma 3.** Let $\{(X_i, Y_i)\}_{i=1}^n$, $S_1$ and $S_2$ be defined as in Lemma 2. Write $\xi \in \{c, d, e, f, g, h, l, m, n, o, p, q\}$.

Let $W_\xi$ be defined as in (12) with respect to $R_\xi$ that tabulated in Table 4. Then the variance of $r_S(X, Y)$ is
\[
\mathbb{V}(r_S) = \frac{6}{n(n+1)} + \frac{9(n-2)(n-3)}{n(n^2-1)(n+1)} \left[ (n-4)\Omega_1(\rho) + \Omega_2(\rho) \right] - \frac{36}{\pi^2 n(n^2-1)(n+1)} \left[ 3(n-2)(3n^2-15n+22)S_2^2 + 12(n-2)^2 S_1 S_2 - 2(n-3)S_1^2 \right] \tag{20}
\]
where
\[
\Omega_1(\rho) = W_c + 8W_d + 2W_f \tag{21}
\]
\[
\Omega_2(\rho) = 6W_g + 8W_h + 6W_i + 2W_n + W_o + \frac{1}{3}. \tag{22}
\]

Moreover, when $n$ is sufficiently large,
\[
\mathbb{V}(r_S) \simeq \frac{1}{n} \left[ 9\Omega_1(\rho) - \frac{324S_2^2}{\pi^2} \right]. \tag{23}
\]

Proof. See Appendix A. \hfill \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 be shown that (Appendix B)
\[
\Omega_1(0) = \frac{1}{9}, \quad \Omega_2(0) = \frac{5}{9}, \tag{24}
\]
\[
\Omega_1(1) = 1, \quad \Omega_2(1) = \frac{16}{3}. \tag{25}
\]
Substituting $\rho = 0$ and (24) into (20) leads directly to
\[ V(r_S)\big|_{\rho=0} = \frac{1}{n-1} \] (26)
which is a well known result [22]. Substituting $\rho = 1$ and (25) into (20) and (23) together with some simplifications yields the following obvious result
\[ V(r_S)\big|_{\rho=1} = 0. \]
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
\[
C(r_S, r_K) = \frac{12}{n(n^2-1)} \left[ \frac{7n-5}{18} + (n-4) \frac{S_1^2}{\pi^2} - 5(n-2) \frac{S_2^2}{\pi^2} \right. \\
-6(n-2)^2 \frac{S_1 S_2}{\pi^2} + (n-2)(n-3) \Omega_3(\rho) \left. \right] \]
(27)
\[
\approx \frac{12}{n} \left[ \Omega_3(\rho) - 6 \frac{S_1 S_2}{\pi^2} \right] \quad \text{(as } n \text{ large)} \] (28)
where
\[
\Omega_3(\rho) = \frac{1}{2} W_g + W_h. \] (29)

**Proof.** See Appendix C.

**Corollary 1.** In Lemma 4, the covariance $C(r_S, r_K)$ can also be expressed as
\[
C(r_S, r_K) = \frac{12}{n(n^2-1)} \left[ \frac{(n+1)^2}{18} + (n-4) \frac{S_1^2}{\pi^2} - 5(n-2) \frac{S_2^2}{\pi^2} \right. \\
-6(n-2)^2 \frac{S_1 S_2}{\pi^2} + \frac{2}{\pi^2} (n-2)(n-3) \Omega_4(\rho) \left. \right] \]
(30)
\[
\approx \frac{12}{n} \left[ \frac{1}{18} + 2 \frac{\Omega_4(\rho)}{\pi^2} - 6 \frac{S_1 S_2}{\pi^2} \right] \quad \text{(as } n \text{ large)} \] (31)
where
\[
\Omega_4(\rho) = \int_0^\rho \left[ \sin^{-1} \left( \frac{x}{3} \right) + 2 \sin^{-1} \left( \frac{x}{\sqrt{3}} \right) \right] \frac{dx}{\sqrt{1 - x^2}} \\
-2 \int_0^\rho \sin^{-1} \left( \frac{x}{2} \sqrt{\frac{1-x^2}{9-3x^2}} \right) \frac{dx}{\sqrt{4-x^2}} \\
+ \int_0^\rho \sin^{-1} \left( \frac{x}{2} \sqrt{\frac{5-x^2}{3-x^2}} \right) \frac{dx}{\sqrt{4-x^2}} \\
-2 \int_0^\rho \sin^{-1} \left( \frac{x}{\sqrt{12-6x^2}} \right) \frac{dx}{\sqrt{4-x^2}} \\
+2 \int_0^\rho \sin^{-1} \left( \frac{x}{\sqrt{\frac{3-x^2}{4-2x^2}}} \right) \frac{dx}{\sqrt{4-x^2}}. \] (32)
Proof. Inverting (11) yields

\[ W = 16P_4^0 - 1 - \frac{2}{\pi} \sum_{r=1}^{3} \sum_{s=r+1}^{4} \sin^{-1} \varrho_{rs}, \quad (33) \]

Combining (29) and (33), \( \Omega_3(\rho) \) can be rewritten in terms of \( P_4^0 \) and the correlation coefficients corresponding to \( R_g \) and \( R_h \) in Appendix 2 of [44]. This leads to

\[ \Omega_3(\rho) = \frac{1}{18} + \frac{2}{\pi^2} \Omega_4(\rho). \quad (34) \]

The corollary thus follows directly by substituting (34) to (27) and (28), respectively. \( \square \)

Remark 2. Both (27) and (30) 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 (27) is more convenient for computation; whereas (30) is more convenient for analysis. For example, performing the Taylor expansion to (31) with the assistance of (32) gives

\[ C(r_S, r_K) \simeq \frac{2}{3n} \left[ 1 - 1.24858961\rho^2 + 0.06830496\rho^4 
+ 0.07280482\rho^6 + 0.04025528\rho^8 + 0.02189277\rho^{10} + \cdots \right] \quad (35) \]

which agrees with the formula (51) obtained in [40], 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 (30) is exact, we believe that (35) is more accurate than (51) in [40].

Remark 3. Due to the complicated integrals involved in (27) and (30), \( C(r_S, r_K) \) cannot be expressed in elementary functions. However, exact results are attainable for \( \rho = 0 \) and \( \rho = 1 \) (−1). It follows that (Appendix 7)

\[ \Omega_3(0) = \frac{1}{18} \quad (36) \]
\[ \Omega_3(1) = \frac{1}{2} \quad (37) \]

Substituting (36) into (27) yields

\[ C(r_S, r_K)\big|_{\rho=0} = \frac{2}{3n} \frac{n + 1}{n - 1} \quad (38) \]

which is more readily to obtain on substitution of \( \rho = 0 \) into (30). Regarding the case for \( \rho = 1 \), it is rather difficult by means of substituting \( \rho = 1 \) into (30) and evaluating \( \Omega_4(1) \) based on (32) thereafter. Fortunately, with the help of (37), it follows readily from (27) and (28) that \( C(r_S, r_K)\big|_{\rho=1} = 0 \). Due to symmetry, we also have \( C(r_S, r_K)\big|_{\rho=-1} = 0 \).

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]

$$
\varepsilon N \left( \mu_X, \mu_Y, \sigma_X^2, \sigma_Y^2, \rho \right) + \varepsilon N \left( \mu_X, \mu_Y, \mu_X^2 \sigma_X^2, \lambda Y^2 \sigma_Y^2, \rho' \right)
$$

(39)

where $\varepsilon \equiv 1 - \epsilon$, $0 \leq \epsilon \leq 1$, $\lambda_X \gg 1$, and $\lambda_Y \gg 1$. Under this Gaussian contamination model 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 $E(r_P) \to \rho'$ as $\lambda_X \to \infty$ and $\lambda_Y \to \infty$ [28]. On the other hand, as shown in the theorem below, SR and KT are more robust than PPMCC under the model (39).

**Theorem 1.** Let $\{(X_i, Y_i)\}_{i=1}^n$ be i.i.d. samples generated from the model (39). Let $r_S$ and $r_K$ be the SR and KT defined in (2) and (3), respectively. Then

$$
\lim_{\epsilon \to 0, \lambda_X \to \infty, \lambda_Y \to \infty} E(r_K) = \frac{2}{\pi} \left[ (1 - 2\epsilon) \sin^{-1} \rho + 2\epsilon \sin^{-1} \rho' \right]
$$

(40)

$$
\lim_{n \to \infty, \lambda_X \to \infty, \lambda_Y \to \infty} E(r_S) = \frac{6}{\pi} \left[ (1 - 3\epsilon) \sin^{-1} \frac{\rho}{2} + \epsilon \sin^{-1} \frac{\rho'}{2} \right].
$$

(41)

**Proof.** See Appendix D.

**Remark 4.** It was stated without substantial argument in [28] that, under the model (39), $E(r_s)$ is of the following form

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

(41) as $\epsilon \to 0$, $\lambda_X \to \infty$ and $\lambda_Y \to \infty$. This is quite inconsistent with our result (41) in Theorem 1. We will resolve the controversy between (41) and (41) by Monte Carlo simulations in Section 6.

5. **Estimators of the Population Correlation**

In this section, we investigate the performance of the estimators of $\rho$ based on SR and KT in terms of bias, MSE and ARE to PPMCC. To gain further insight into their relationship, the correlation between the two estimators $\hat{\rho}_S$ and $\hat{\rho}_K$ (defined below) is also derived.

### 5.1. Asymptotic Unbiased Estimators

Inverting (14), (17) and (18), we have the following estimators of $\rho$

$$
\hat{\rho}_P \triangleq r_P
$$

(42)

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

(43)

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

(44)
Moreover, another estimator based on a mixture of $r_S$ and $r_P$ can be constructed as [22]

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

based on the following relationship

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

In the sequel we will focus on the properties of the estimators defined in (42)–(45). Here the estimator $\hat{\rho}_P$ in (42) 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 (42)–(45) 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 (42)–(45), 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 $\sigma^2_S \triangleq \mathbb{V}(r_S)$, $\sigma^2_K \triangleq \mathbb{V}(r_K)$ and $\sigma_{S,K} \triangleq \mathbb{C}(r_S, r_K)$. Then, under the same assumptions made as in Lemma 3,

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

(47)

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

(48)

$$\text{BIAS}_K \simeq -\frac{\pi^2 \rho}{8} \sigma^2_K$$

(49)

$$\text{BIAS}_M \simeq -\frac{1}{72 (n-2)^2} \left[ (n+1)^2 \sigma^2_S - 6(n+1)\sigma_{S,K} + 9 \sigma^2_K \right].$$

(50)

**Proof.** The first statement (47) follows directly from (14) in Lemma 2. Now we proceed to evaluate $\text{BIAS}_S$, $\text{BIAS}_K$ and $\text{BIAS}_M$. For convenience, write $\tau_S \triangleq \mathbb{E}(r_S)$, $\tau_K \triangleq \mathbb{E}(r_K)$, $\delta_S \triangleq r_S - \tau_S$, and $\delta_K \triangleq r_K - \tau_K$. Expanding (43) around $\tau_S$ yields

$$\hat{\rho}_S = 2 \sin \left( \frac{\pi}{6} \tau_S \right) + \frac{\pi}{3} \cos \left( \frac{\pi}{6} \tau_S \right) \delta_S - \frac{\pi^2}{36} \sin \left( \frac{\pi}{6} \tau_S \right) \delta_S^2 + \cdots .$$

(51)

Taking expectation of both sides in (51), applying $\mathbb{E}(\delta_S) = 0$, $\mathbb{E}(\delta_S^2) = \sigma^2_S$ and ignoring the high order infinitesimals, we have

$$\mathbb{E}(\hat{\rho}_S) \simeq 2 \sin \left( \frac{\pi}{6} \tau_S \right) - \frac{\pi^2}{36} \sin \left( \frac{\pi}{6} \tau_S \right) \sigma^2_S.$$  

(52)

Substituting (46) into (52), expanding to the order of $(n + 1)^{-1}$, and subtracting $\rho$ thereafter, we obtain the result (48). In a similar way we have

$$\mathbb{E}(\hat{\rho}_K) \simeq \rho - \frac{\pi^2 \rho}{8} \sigma^2_K.$$  

(53)
which leads directly to (49). Performing Taylor expansion of $\hat{\rho}_M(r_S, r_K)$ around $(\overline{r_S}, \overline{r_K})$ till the second order, we have

$$
\hat{\rho}_M = \hat{\rho}_M(\overline{r_S}, \overline{r_K}) + \frac{\partial (\hat{\rho}_M)}{\partial (r_S)} \delta_S + \frac{\partial (\hat{\rho}_M)}{\partial (r_K)} \delta_K + \frac{1}{2} \left[ \frac{\partial^2 (\hat{\rho}_M)}{\partial (r_S)^2} \delta_S^2 + \frac{\partial^2 (\hat{\rho}_M)}{\partial (r_K)^2} \delta_K^2 + 2 \frac{\partial^2 (\hat{\rho}_M)}{\partial (r_S) \partial (r_K)} \delta_S \delta_K \right] + \cdots .
$$

Taking expectation of both sides in (53), ignoring high order infinitesimals, applying the results

$$
\hat{\rho}_M(\overline{r_S}, \overline{r_K}) = \rho, \quad E(\delta_S) = 0, \quad E(\delta_K) = 0, \quad E(\delta_S^2) = \sigma_S^2, \quad E(\delta_K^2) = \sigma_K^2, \quad E(\delta_S, \delta_K) = \sigma_{S,K},
$$

along with the second order partial derivatives

$$
\frac{\partial^2 (\hat{\rho}_M)}{\partial (r_S)^2} = -\frac{\pi^2 \rho}{36} \frac{(n + 1)^2}{(n - 2)^2},
$$

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

$$
\frac{\partial^2 (\hat{\rho}_M)}{\partial (r_S) \partial (r_K)} = \frac{\pi^2 \rho}{12} \frac{n + 1}{(n - 2)^2},
$$

and subtracting $\rho$ thereafter, we arrive at the forth theorem statement (50), thus completing the proof.

Remark 5. From (47)–(50), 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 $n \to \infty$.

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 (20) and (27), the expressions of $V(r_S)$ and $C(r_S, r_K)$, respectively.

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 (14), it follows that

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

By the delta method, it follows that [22]

$$
V(\hat{\rho}_S) \simeq \frac{\pi^2 (4 - \rho^2)}{36} V(r_S) \quad (55)
$$

$$
V(\hat{\rho}_K) \simeq \frac{\pi^2 (1 - \rho^2)}{4} V(r_K). \quad (56)
$$

Now we only need to deal with $V(\hat{\rho}_M)$, which is stated below.
Theorem 3. Let \( \hat{\rho}_M \) be defined as in (45). Then, under the same assumptions made as in Lemma 3,
\[
\mathbb{V}(\hat{\rho}_M) \simeq \frac{\pi^2(4 - \rho^2)}{36(n-2)^2} \left[ (n+1)^2 \sigma_S^2 - 6(n+1) \sigma_{S,K} + 9 \sigma_K^2 \right].
\] (57)

Proof. Using the delta method [18], it follows that
\[
\mathbb{V}(\hat{\rho}_M) \simeq \left[ \frac{\partial(\hat{\rho}_M)}{\partial(r_S)} \right]^2 \sigma_S^2 + \left[ \frac{\partial(\hat{\rho}_M)}{\partial(r_K)} \right]^2 \sigma_K^2 + 2 \frac{\partial(\hat{\rho}_M)}{\partial(r_S)} \frac{\partial(\hat{\rho}_M)}{\partial(r_K)} \sigma_{S,K}.
\] (58)

The theorem thus follows with substitutions of the partial derivatives
\[
\frac{\partial \hat{\rho}_M(r_S, r_K)}{\partial(r_S)} = \frac{\pi n + 1}{6n - 2} \sqrt{4 - \rho^2}
\]
\[
\frac{\partial \hat{\rho}_M(r_S, r_K)}{\partial(r_K)} = \frac{\pi - 1}{2n - 2} \sqrt{4 - \rho^2}
\]
into (58), 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 fourt estimators, we will shift our attention to the asymptotic properties of \( \hat{\rho}_\zeta \) in the sequel. Since \( \lim_{n \to \infty} \mathbb{E}(\hat{\rho}_\zeta) = \rho \), we can compare their performances by means of the \textit{asymptotic relative efficiency}, which is defined as [18]
\[
\text{ARE}_\zeta \triangleq \lim_{n \to \infty} \frac{\mathbb{V}(\hat{\rho}_P)}{\mathbb{V}(\hat{\rho}_\zeta)}, \quad \zeta \in \{P, S, K, M\}.
\] (59)

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]
\[
\text{CRLB} = \frac{(1 - \rho^2)^2}{n}.
\] (60)

From (59) it is obvious that \( \text{ARE}_P = 1 \). Moreover, comparing (55) and (57), it is easily seen that 
\( \lim_{n \to \infty} \mathbb{V}(\hat{\rho}_S)/\mathbb{V}(\hat{\rho}_M) = 1 \), which leads readily to \( \text{ARE}_S = \text{ARE}_M \) by referring to (59). Then we only need to focus on \( \text{ARE}_S \) and \( \text{ARE}_K \) in the following discussion.

Theorem 4. Let \( \text{ARE}_S \) and \( \text{ARE}_K \) be defined as in (59). Then

\[
\text{ARE}_S = \frac{36(1 - \rho^2)^2}{(4 - \rho^2) \left[ 9\pi^2 \Omega_1(\rho) - 324 \left( \sin^{-1} \frac{1}{2} \rho \right)^2 \right]} \] (61)

\[
\text{ARE}_K = \frac{9(1 - \rho^2)}{\pi^2 - 36 \left( \sin^{-1} \frac{1}{2} \rho \right)^2}.
\] (62)

Proof. Substituting (55) and (56) into (59) yields (61) and (62), respectively, and the proof completes. \( \square \)
Remark 6. Due to the intractability of $\Omega_1(\rho)$ in (61), $\text{ARE}_S$ 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 (61), it is easy to verify that

$$\text{ARE}_S(0) = \frac{9}{\pi^2} \simeq 0.9119$$

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

$$\text{ARE}_S(\pm 1) = \frac{15 + 11\sqrt{5}}{57} \simeq 0.6947. \quad (63)$$

Now let us investigate $\text{ARE}_K$. It follows from (62) that, $\text{ARE}_K$ is expressible as elementary functions of $\rho$, and is therefore more tractable than $\text{ARE}_S$. In other words, we can evaluate easily any value of $\text{ARE}_K$ with respect to any value of $\rho \neq \pm 1$. For example, substituting $\rho = 0$ into (62) yields

$$\text{ARE}_K(0) = \frac{9}{\pi^2}$$

which is identical to $\text{ARE}_S(0)$ and also well known [22]. However, when $\rho \to \pm 1$, an extra effort is necessary, since both the numerator and denominator of (62) vanish in this case. Apply the L’Hopital’s rule, we find the following result

$$\text{ARE}_K\bigg|_{\rho \to \pm 1} = \frac{1}{4} \frac{\rho\sqrt{4 - \rho^2}}{\sin^{-1} \frac{1}{2}\rho} \bigg|_{\rho = \pm 1} = \frac{3\sqrt{3}}{2\pi} \simeq 0.8270 \quad (64)$$

which is greater than $\text{ARE}_S(\pm 1)$. In fact, a comparison of $\text{ARE}_S$ and $\text{ARE}_K$ in Section 6 suggest that $\text{ARE}_S \leq \text{ARE}_K$ 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 (42)–(45) 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 $\Delta \rho$.  

14
6.1. Tables of $\Omega_1(\rho)$, $\Omega_2(\rho)$ and $\Omega_3(\rho)$

Table 1 contains the values of $\Omega_1(\rho)$ and $\Omega_2(\rho)$ in (20), 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 $V(r_S)$ 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 ten decimal places. In Table 2 are tabulated the values of $\Omega_3(\rho)$ in (27) 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 1. The values of $\Omega_1(\rho)$, $\Omega_2(\rho)$ and $\Omega_3(\rho)$ with respect to $\rho$ not included in Tables 1 and 2 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 $V(r_S)$, $V(\hat{\rho}_S)$, $V(\hat{\rho}_M)$, BIAS($\hat{\rho}_S$), BIAS($\hat{\rho}_M$), ARE$_S$, ARE$_M$, and so forth.

6.2. Verification of BIAS$_\zeta$ and $V(\hat{\rho}_\zeta)$ in Small Samples

Fig. 1 shows the bias effects 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 BIAS$_\zeta$ can be ordered as BIAS$_M < BIAS_P < BIAS_K < BIAS_S$. 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 (48) and (50) with respect to BIAS$_S$ and BIAS$_M$ are more accurate than (47) and (49) with respect to BIAS$_P$ and 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 (47) and (49) to the corresponding simulation results are less noticeable when the sample size $n$ is increased up to 20.

Table 3 lists, for each of the three samples sizes, 10, 20 and 30, 1) the theoretical results (54)–(57) with respect to $V(\hat{\rho}_\zeta)$ and 2) the corresponding observed variances from the Monte Carlo simulations. It can be seen that (55) and (57), with respect to $V(\hat{\rho}_S)$ and $V(\hat{\rho}_M)$, are accurate enough even though the sample size is as small as $n = 10$. On the other hand, unfortunately, the theoretical formula (54) for $V(\hat{\rho}_P)$ and (56) for $V(\hat{\rho}_K)$ 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 (54)–(57) when approximating the variances of $\hat{\rho}_\zeta$ for $n \geq 20$ in practice.

6.3. Comparison of MSE in Small Samples

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

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

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

6.4. Verification and Comparison of \( \text{ARE}_S \) and \( \text{ARE}_K \)

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 \( \text{ARE}_M \) are not included in Fig. 3. It can be observed that 1) the simulations agree well with our theoretical findings in (61) and (62), respectively, 2) \( \text{ARE}_K \) lies consistently above \( \text{ARE}_S \), 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 \( \text{ARE}_S \) also apply to \( \text{ARE}_M \) due to the equivalence between \( \hat{\rho}_S \) and \( \hat{\rho}_M \) when the sample size \( n \) is large.

6.5. Performance of \( \hat{\rho}_\varsigma \) in Contaminated Normal Model

Fig. 4 puports to 1) verify the two statements concerning \( \mathbb{E}(r_K) \) and \( \mathbb{E}(r_S) \) in Theorem 1 under the contaminated Gaussian model (39), and 2) compare our formula (41) 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 (39) 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 (40) and (41) 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 (±1), MSE\(_S\) and MSE\(_M\) behave similarly with MSE\(_K\) for \( \epsilon = 0.01 \). Nevertheless, MSE\(_S\) and MSE\(_M\) are much larger than MSE\(_K\) for \( \epsilon = 0.05 \) when \( \rho \) falls in the neighborhood of ±1. Combing Fig. 5(a) and (b), it would be reasonable to rank their performance as \( \hat{\rho}_K \geq \hat{\rho}_S \sim \hat{\rho}_M \gg \hat{\rho}_P \) in terms of MSE under the contaminated normal model (39), where the symbol \( \sim \) stands for “is similar to”.

7. Concluding Remarks

In this paper we have investigated systematically the properties of the 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. [45], 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 (20), in contrast with the closed form expression of $\mathcal{V}(r_K)$ in (19). However, this mathematical inconvenience is, to some extent, offset by Table 1 provided in this work, especially from the viewpoint of numerical accuracy. Moreover, as demonstrated in Fig. 1 and Table 3, the convergence speed of the asymptotic formulae (49) and (56) with respect to $\text{BIAS}_K$ and $\mathcal{V}(\hat{\rho}_K)$ are slower than those of $\text{BIAS}_S$ and $\mathcal{V}(\hat{\rho}_S)$ due to the high nonlinearity of the calibration (44). 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 $\hat{\rho}_S$ or $\hat{\rho}_K$ 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. One 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 and 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.
Appendix A

Proof of Lemma 3. Using the technique developed by Moran [36] for finding $E(r_S)$, it follows that the ranks can be expressed as

$$P_i = \sum_{j=1}^{n} H(X_i - X_j) + 1 \quad (A.1)$$

$$Q_i = \sum_{k=1}^{n} H(Y_i - Y_k) + 1 \quad (A.2)$$

where $H(\bullet)$ is defined in (7). Substituting (A.1) and (A.2) into (2) yields

$$r_S = \frac{S - \frac{1}{4}(n - 1)^2}{\frac{12}{n^2 - 1}} \quad (A.3)$$

where

$$S = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} H(X_i - X_j)H(Y_i - Y_k)$$

$$= \sum_{i \neq j=1}^{n} H(X_i - X_j)H(Y_i - Y_j) \quad (A.4)$$

$$+ \sum_{i \neq j \neq k=1}^{n} H(X_i - X_j)H(Y_i - Y_k).$$

Then

$$\text{Var}(r_S) = \frac{144}{n^2(n^2 - 1)^2} \left[ E(S^2) - E^2(S) \right]. \quad (A.5)$$

Taking the expectation of both sides of (A.4) with the assistance of (9) in Lemma 2, it follows readily that

$$E(S) = n^{[2]} \left( \frac{1}{4} + \frac{S_1}{2\pi} \right) + n^{[3]} \left( \frac{1}{4} + \frac{S_2}{2\pi} \right) \quad (A.6)$$

where $n^{[\kappa]} \triangleq n(n-1)\cdots(n-\kappa+1)$, with $\kappa$ being a positive integer. Now the variance of $r_S$ depends on the evaluation of $E(S^2)$, which is a weighted summation of 24 quadrivariate normal orthant probabilities $P_0^0(R_\xi) = E(Z_1Z_2Z_3Z_4)$ corresponding to $R_\xi$ listed in Table 4 [44]. Collecting the terms of $P_0^0(R_\xi)$ in Table 4, subtracting the square of the right side of (A.6) and substituting the resultant into (A.5) along with some simplifications, we obtain the expression of (20) with

$$\Omega_1(\rho) = W_c + 4W_d + 2W_e + 2W_f \quad (A.7)$$

$$\Omega_2(\rho) = 4(W_g + W_h + W_l + W_q) + 2(W_n + W_p) + W_m + W_o. \quad (A.8)$$

An application of the relationship (11) to Appendix 2 of [44] yields

$$W_c = 2W_d, \quad W_g = W_p, \quad W_h = W_q, \quad \text{and} \quad W_m = 2W_l + \frac{1}{3}. \quad (A.9)$$

Substituting (A.9) into (A.7) and (A.8) yields (21) and (22), respectively. Hence the first theorem statement (20) follows. Ignoring the $o(n^{-1})$ terms in (20) yields the second statement (23), thus completing the proof. \[\square\]
Appendix B

Derivations of $\Omega_1(\rho)$, $\Omega_2(\rho)$ and $\Omega_3(\rho)$ for $\rho = 0, 1$. From (21), (22) and (29), it suffices to evaluate $W_{\xi}$, $\xi^\prime \in \{c, d, f, g, h, l, n, o\}$ for $\rho = 0, 1$; and with (33), it suffices to evaluate $P^0_4(R_{\xi'})$ for $\rho = 0, 1$. It follows readily from Appendix 2 of [44] that for $\rho = 0$, $P^0_4(R_c) = P^0_4(R_g) = 1/9$, $P^0_4(R_d) = P^0_4(R_h) = 1/24$, $P^0_4(R_f) = P^0_4(R_o) = 1/16$, $P^0_4(R_l) = 1/18$, $P^0_4(R_n) = 1/36$. Then, with the help of (33), we have the values $W_{\xi'}(0)$ as listed in the $W(0)$-column of Table 4. Using these $W_{\xi'}(0)$ values with the relationships (21), (22) and (29) yields $\Omega_1(0) = 1/9$, $\Omega_2(0) = 5/9$ and $\Omega_3(0) = 1/18$, respectively.

When $\rho$ approaches unity, it is rather tricky to evaluate the values $W_{\xi'}(1)$. Substituting $\rho = 1$ directly into the integrals in (13) or the integrals in Appendix 2 of [44] will not lead to any tractable argument. We have to investigate case by case. From Table 4, it is seen that the off-diagonal elements of $R_c$ are all $1/2$ when $\rho = 1$. Then we have $P^0_4(R_c)|_{\rho=1} = 1/5$ [52], and hence $W_c(1) = 1/5$ by (33). From [53] it follows that $P^0_4(R_f)|_{\rho=1} = 2/15$ and $P^0_4(R_m)|_{\rho=1} = 1/6$. Then we have, by (33), $W_f(1) = 2/15$ and $W_m(1) = 1/3$, the latter yielding $W_l(1) = 0$ from the identity $W_m = W_l + 1/3$ in (A.9). Substituting $R_f|_{\rho=1}$ into (12) and exchanging $z_1$ and $z_2$ gives $W_f(1) = W_c(1)$, which implies that $W_d(1) = 1/15$ by the identity $W_c = 2W_d$ in (A.9). Similarly, we also have $W_o(1) = W_m(1) = 1/3$ upon substitution of $R_m|_{\rho=1}$ into (12) and exchange of $z_3$ and $z_4$. It is easy to verify that $P^0_4(R_n)$ vanishes as $\rho \to 1$, since in this case $Z_1 = -Z_4$ and $H(Z_1)H(Z_2)H(Z_3)H(Z_4) = 0$ by the definition of $H(\bullet)$ in (7). Then $W_n(1) = 0$ by applying the relationship (33) once more. When $\rho$ approaches unity, it follows that $P^0_4(R_g)$ and $P^0_4(R_h)$ degenerate to two trivariate normal orthant probabilities that have closed form expressions (10). Specifically, it follows that $P_4(R_g)|_{\rho=1} = 1/4$ and $P_4(R_h)|_{\rho=1} = 1/8$, yielding $W_g(1) = 1/3$ and $W_h(1) = 1/3$, respectively. Having all the values of $W_{\xi'}(1)$, as listed in the $W(1)$-column of Table 4, and the three relationships (21), (22) and (29), we obtain $\Omega_1(1) = 1$, $\Omega_2(1) = 16/3$ and $\Omega_3(1) = 1/2$, respectively, and the evaluations complete.

Appendix C

Proof of Lemma 4. Let $S$ be the same as in (A.4) and $T$ be the numerator of (3). Define

$$I \equiv \sum_{i \neq j} H(X_i - X_j)H(Y_i - Y_j) \quad (C.1)$$

$$J \equiv \sum_{i \neq j, j \neq k} H(X_i - X_j)H(Y_i - Y_k) \quad (C.2)$$

$$K \equiv \sum_{i \neq j} H(X_i - X_j) \quad (C.3)$$

$$L \equiv \sum_{i \neq k} H(Y_i - Y_k). \quad (C.4)$$
Then, we have, from (3), (A.3) and (A.4) along with the relationship \( \text{sgn}(\mathbf{A}) = 2\mathbf{H}(\mathbf{A}) - 1 \),

\[
S = I + J 
\]

\[
T = 4I - 2K - 2L + n^{[2]} 
\]

and hence

\[
\mathbb{C}(r_S, r_K) = \frac{12}{n^2(n-1)(n^2 - 1)} \mathbb{C}(S, T) 
\]

\[
= \frac{12}{n^2(n-1)(n^2 - 1)} \left[ \mathbb{E}(ST) - \mathbb{E}(S)\mathbb{E}(T) \right]. 
\]

From (8) and (9), 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 (C.6) gives

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

Recall that we have obtained \( \mathbb{E}(S) \) in (A.6). Now the only difficulty lies in the evaluation of \( \mathbb{E}(ST) \) in (C.7). Multiplying (C.5) and (C.6), expanding and taking expectations term by term, we have

\[
\mathbb{E}(ST) = 4\mathbb{E}(IJ) - 2\mathbb{E}(KJ) - 2\mathbb{E}(LJ) 
\]

\[
+ 4\mathbb{E}(I^2) - 2\mathbb{E}(KI) - 2\mathbb{E}(LI) + n^{[2]}\mathbb{E}(S). 
\]

Now, resorting to Table 5, we are ready to evaluate the first six terms in (C.9). From (C.1) and (C.2), it follows that \( \mathbb{E}(IJ) \) is a summation of \( P_4^0 \) terms of the form

\[
\mathbb{E}\{ H(X_i - X_j)H(Y_i - Y_j)H(X_k - X_l)H(Y_k - Y_m) \}. 
\]

Since, by definition (7), \( H(0) = 0 \), the term (C.10) vanishes for \( i = j \) or \( k = l \) or \( k = m \). Then there are \( n^2(n-1)^2(n-2) \) nontrivial (C.10)-like terms left to be evaluated. It follows that the domain of the quintuple \( (i, j, k, l, m) \) can be partitioned into thirteen 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.6) into (C.9), subtracting the multiplication of (A.6) and (C.8) and substituting the resultant back into (C.7), we find that \( \mathbb{C}(r_S, r_K) \) is of the form (27) 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 (29) by applying the identities in (A.9). The theorem then follows. \( \square \)
Appendix D

Proof of Theorem 1. For 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 (39), respectively. From (A.1), (A.2) and (C.6), it follows that the numerator \( T \) of (3) can be simplified to

\[
T = 4 \sum_{i \neq j=1}^{n} H(X_i - X_j) H(Y_i - Y_j) - n^{[2]} \tag{D.1}
\]

which yields

\[
E(T) = 4n^{[2]} E[H(X_1 - X_2)H(Y_1 - Y_2)] - n^{[2]} \tag{D.2}
\]

by the i.i.d. assumption. To evaluate \( E_1 \) in (D.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

\[
\varphi = [(1 - \epsilon) \phi_1 + \epsilon \psi_1] [(1 - \epsilon) \phi_2 + \epsilon \psi_2] = (1-\epsilon)^2 \phi_1 \phi_2 + \epsilon(1-\epsilon) \phi_1 \psi_2 + \epsilon(1-\epsilon) \phi_2 \psi_1 + \epsilon^2 \psi_1 \psi_2 \tag{D.3}
\]

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

\[
U \triangleq \frac{X_1 - X_2}{\sqrt{V(X_1 - X_2)}} \text{ and } V \triangleq \frac{Y_1 - Y_2}{\sqrt{V(Y_1 - Y_2)}}.
\]

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

\[
\varphi_1 = \rho \tag{D.4}
\]

\[
\varphi_2 = \frac{\rho + \lambda_X \lambda_Y \rho'}{\sqrt{1 + \lambda_X^2 \lambda_Y^2}} \to \rho' \text{ as } \lambda_X, \lambda_Y \to \infty \tag{D.5}
\]

\[
\varphi_3 = \frac{\rho + \lambda_X \lambda_Y \rho'}{\sqrt{1 + \lambda_X^2 \lambda_Y^2}} \to \rho' \text{ as } \lambda_X, \lambda_Y \to \infty \tag{D.6}
\]

\[
\varphi_4 = \rho' \tag{D.7}
\]

respectively. An application of the Sheppard’s theorem (9) to (D.2) along with (D.4)–(D.7) yields

\[
E(T) = 4n^{[2]} \sum_{i=1}^{4} \alpha_i \left( \frac{1}{4} + \frac{1}{2\pi} \sin^{-1} \varphi_i \right) - n^{[2]} \tag{D.8}
\]

\[
= \frac{2n^{[2]} \alpha_1}{\pi} \left[ \alpha_1 \sin^{-1} \rho + 2\alpha_2 \sin^{-1} \varphi_2 + \alpha_4 \sin^{-1} \rho' \right].
\]

Now it is not difficult to verify that the first statement (40) holds by 1) dividing both sides of (D.8) by \( n^{[2]} \), 2) letting \( \lambda_X \to \infty \) and \( \lambda_Y \to \infty \), and 3) ignoring the \( O(\epsilon^2) \) terms.

To prove the second statement (41), it suffices to evaluate \( E(S) \) by the relationship (A.3). Taking expectations of both sides in (A.4) along with the i.i.d. assumptions gives

\[
E(S) = n^{[2]} E_1 + n^{[3]} E[H(X_1 - X_2)H(Y_1 - Y_3)]. \tag{D.9}
\]
Since we have known $E_1$ in the above development, now we only need to work out $E_2$ in (D.9). 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 (39) and the i.i.d. assumption,
\[
\varpi = \left[ (1-\epsilon)\phi_1 + \epsilon\psi_1 \right] \left[ (1-\epsilon)\phi_2 + \epsilon\psi_2 \right] \left[ (1-\epsilon)\phi_3 + \epsilon\psi_3 \right]
\]
\[
= (1-\epsilon)^3 \varpi_1 + \epsilon(1-\epsilon)^2 \left( \varpi_2 \phi_3 \phi_3 + \varpi_3 \phi_3 \phi_3 + \varpi_4 \phi_3 \phi_3 \right)
\]
\[
+ \epsilon^2 (1-\epsilon) \left( \varpi_5 \psi_3 \psi_3 + \varpi_6 \psi_3 \psi_3 + \varpi_7 \psi_3 \psi_3 \right) + \epsilon^3 \varpi_8 \psi_3 \psi_3 .
\]

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
\[
V' = \frac{Y_1 - Y_3}{\sqrt{V(Y_1 - Y_3)}}.
\]

Then, with respect to \( \varpi_1 \) to \( \varpi_8 \) in (D.10), \((U, V')\) follows 8 standard bivariate normal distributions with correlations
\[
\varrho_5 = \frac{\rho}{2} \quad \text{(D.11)}
\]
\[
\varrho_6 = \frac{1}{\sqrt{2}} \frac{\sqrt{1 + \lambda_Y^2}}{\sqrt{1 + \lambda_Y^2}} \to 0 \text{ as } \lambda_Y \to \infty \quad \text{(D.12)}
\]
\[
\varrho_7 = \frac{1}{\sqrt{2}} \frac{\sqrt{1 + \lambda_X^2}}{\sqrt{1 + \lambda_X^2}} \to 0 \text{ as } \lambda_X \to \infty \quad \text{(D.13)}
\]
\[
\varrho_8 = \frac{\lambda_X \lambda_Y \rho'}{\sqrt{1 + \lambda_X^2} \sqrt{1 + \lambda_Y^2}} \to \rho' \text{ as } \lambda_X, \lambda_Y \to \infty \quad \text{(D.14)}
\]
\[
\varrho_9 = \frac{\rho}{\sqrt{1 + \lambda_X^2} \sqrt{1 + \lambda_Y^2}} \to 0 \text{ as } \lambda_X, \lambda_Y \to \infty \quad \text{(D.15)}
\]
\[
\varrho_{10} = \frac{1}{\sqrt{2}} \frac{\lambda_X \rho'}{\sqrt{1 + \lambda_X^2}} \to \rho' \text{ as } \lambda_X \to \infty \quad \text{(D.16)}
\]
\[
\varrho_{11} = \frac{1}{\sqrt{2}} \frac{\lambda_Y \rho'}{\sqrt{1 + \lambda_Y^2}} \to \rho' \text{ as } \lambda_Y \to \infty \quad \text{(D.17)}
\]
\[
\varrho_{12} = \frac{\rho'}{2}. \quad \text{(D.18)}
\]

Using the Sheppard’s theorem (9) again together with (D.10)–(D.18), we can obtain the expression of \( E_2 \) and hence \( E(S) \) in terms of \( n, \epsilon \) and \( \varrho_{12} \).\( \varrho_{12} \). Substituting \( E(S) \) into (A.3), letting \( n, \lambda_X, \lambda_Y \to \infty \) and ignoring the \( O(\epsilon^2) \) terms, we arrive at (41), the second theorem statement.

References


[34] 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.


[50] D. R. Childs, Reduction of the multivariate normal integral to characteristic form, Biometrika 54 (1/2) (1967) 293–300.


Figure 1: Comparison of BIAS_ζ, ζ ∈ {P, S, K, M} for (a) n = 10 and (b) n = 20. Theoretical curves, denoted by BIAS_T^ζ in the legend, are plotted over ρ = −0.9(0.1)0.9 based on (47)–(50), respectively; whereas the simulation results, denoted by BIAS_O^ζ in the legend, are plotted over ρ = −1(0.01)1.
Figure 2: Comparison of observed $\text{MSE}_\zeta$, $\zeta \in \{P,S,K,M\}$ for (a) $n = 10$, (b) $n = 20$, (c) $n = 40$, and (d) $n = 60$ over $\rho = -1(0.1)1$, respectively.
Figure 3: Verification and Comparison of $\text{ARE}_S$ and $\text{ARE}_K$ for $n = 1000$ over $\rho = 0(0.01)1$, for theoretical curves, and $\rho = 0(0.05)0.95$, for simulation results. The results (63) and (64) are used to plot the two theoretical curves for $\rho = 1$, respectively.
Figure 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 (40) and (41). The rest parameters of the model (39) are set to be $\sigma_X = \sigma_Y = 1$, $\lambda_X = \lambda_Y = 100$ and $\rho' = 0$, respectively. The formula (\ast) concerning $\mathbb{E}(r_S)$ developed elsewhere [28] is also included for comparison.
Figure 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 (39) 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.
<table>
<thead>
<tr>
<th>( \rho )</th>
<th>0.00+</th>
<th>0.10+</th>
<th>0.20+</th>
<th>0.30+</th>
<th>0.40+</th>
<th>0.50+</th>
<th>0.60+</th>
<th>0.70+</th>
<th>0.80+</th>
<th>0.90+</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.1111111111</td>
<td>0.1185038469</td>
<td>0.1408155407</td>
<td>0.1784569533</td>
<td>0.2321489296</td>
<td>0.3029841008</td>
<td>0.3925307865</td>
<td>0.5030051934</td>
<td>0.6375648509</td>
<td>0.8008401854</td>
</tr>
<tr>
<td>0.01</td>
<td>0.1111849204</td>
<td>0.1200591282</td>
<td>0.1438805317</td>
<td>0.1830890232</td>
<td>0.2384392872</td>
<td>0.3110659830</td>
<td>0.4025930108</td>
<td>0.5151480701</td>
<td>0.6525067154</td>
<td>0.8189017929</td>
</tr>
<tr>
<td>0.02</td>
<td>0.1114063976</td>
<td>0.1217636535</td>
<td>0.1470991369</td>
<td>0.1878820200</td>
<td>0.2449021209</td>
<td>0.3193632585</td>
<td>0.4128663989</td>
<td>0.5278679000</td>
<td>0.6677394272</td>
<td>0.8370574957</td>
</tr>
<tr>
<td>0.03</td>
<td>0.1117755552</td>
<td>0.1236177309</td>
<td>0.1504719575</td>
<td>0.1923743255</td>
<td>0.2513847622</td>
<td>0.3277970726</td>
<td>0.423356847</td>
<td>0.5406684104</td>
<td>0.6832688919</td>
<td>0.8563978883</td>
</tr>
<tr>
<td>0.04</td>
<td>0.1122924682</td>
<td>0.1256216966</td>
<td>0.1539996256</td>
<td>0.1979556956</td>
<td>0.2583500955</td>
<td>0.3364502180</td>
<td>0.4340577002</td>
<td>0.5537204292</td>
<td>0.6991017291</td>
<td>0.8756698200</td>
</tr>
<tr>
<td>0.05</td>
<td>0.1129572289</td>
<td>0.1277751479</td>
<td>0.1576828058</td>
<td>0.2032380114</td>
<td>0.2653383401</td>
<td>0.3452987382</td>
<td>0.4498413799</td>
<td>0.5670822120</td>
<td>0.7152430387</td>
<td>0.8953367468</td>
</tr>
<tr>
<td>0.06</td>
<td>0.1137695654</td>
<td>0.1300607778</td>
<td>0.1615221950</td>
<td>0.2068654397</td>
<td>0.2725047311</td>
<td>0.3543420744</td>
<td>0.4561277650</td>
<td>0.5805961801</td>
<td>0.7317090189</td>
<td>0.9154091574</td>
</tr>
<tr>
<td>0.07</td>
<td>0.1147307963</td>
<td>0.1325367071</td>
<td>0.1655185233</td>
<td>0.2142990814</td>
<td>0.2795804836</td>
<td>0.3635815060</td>
<td>0.4675000084</td>
<td>0.5944289365</td>
<td>0.7484198555</td>
<td>0.9358987450</td>
</tr>
<tr>
<td>0.08</td>
<td>0.1158392070</td>
<td>0.1351441527</td>
<td>0.1696725544</td>
<td>0.2200800791</td>
<td>0.2873783079</td>
<td>0.3730293618</td>
<td>0.4791013791</td>
<td>0.5607822080</td>
<td>0.7152430387</td>
<td>0.9561805554</td>
</tr>
<tr>
<td>0.09</td>
<td>0.1170975289</td>
<td>0.1379035942</td>
<td>0.1739850867</td>
<td>0.2260296183</td>
<td>0.2950888115</td>
<td>0.3826770865</td>
<td>0.4903526080</td>
<td>0.6229081771</td>
<td>0.7830436585</td>
<td>0.9781804140</td>
</tr>
</tbody>
</table>

Table 1: Values of \( \Omega_1(\rho) \) and \( \Omega_2(\rho) \) in Lemma 3

(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) \( \Omega_2(0) = 5/9 \), \( \Omega_2(1) = 16/3 \), \( \Omega_2(-\rho) = \Omega_2(\rho) \).
Table 2: Values of $\Omega_3(\rho)$ in Lemma 4

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>0.00+</th>
<th>0.10+</th>
<th>0.20+</th>
<th>0.30+</th>
<th>0.40+</th>
<th>0.50+</th>
<th>0.60+</th>
<th>0.70+</th>
<th>0.80+</th>
<th>0.90+</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>0.0555555556</td>
<td>0.0579082728</td>
<td>0.0650488362</td>
<td>0.0772363477</td>
<td>0.0949464623</td>
<td>0.1189551518</td>
<td>0.1505074640</td>
<td>0.1916842312</td>
<td>0.2463473665</td>
<td>0.3235686523</td>
</tr>
<tr>
<td>0.01</td>
<td>0.0555790160</td>
<td>0.0584040661</td>
<td>0.0660345125</td>
<td>0.0787487694</td>
<td>0.0970477974</td>
<td>0.1217448819</td>
<td>0.1541471283</td>
<td>0.1964547749</td>
<td>0.2528155359</td>
<td>0.3333572518</td>
</tr>
<tr>
<td>0.02</td>
<td>0.0556494053</td>
<td>0.0589477686</td>
<td>0.0670708481</td>
<td>0.0803167902</td>
<td>0.0992127259</td>
<td>0.1246109820</td>
<td>0.1578844390</td>
<td>0.2013621185</td>
<td>0.2595091876</td>
<td>0.3437145088</td>
</tr>
<tr>
<td>0.03</td>
<td>0.0557666747</td>
<td>0.0595395710</td>
<td>0.0681582284</td>
<td>0.0819410523</td>
<td>0.1014227030</td>
<td>0.1275551071</td>
<td>0.1612226120</td>
<td>0.2064119567</td>
<td>0.2664434835</td>
<td>0.3547334972</td>
</tr>
<tr>
<td>0.04</td>
<td>0.0559310834</td>
<td>0.0601796821</td>
<td>0.0692970621</td>
<td>0.0836222291</td>
<td>0.1037375018</td>
<td>0.1305789987</td>
<td>0.1656636397</td>
<td>0.2116104637</td>
<td>0.2736356537</td>
<td>0.3665400166</td>
</tr>
<tr>
<td>0.05</td>
<td>0.0561424692</td>
<td>0.0608683287</td>
<td>0.0704877763</td>
<td>0.0853610263</td>
<td>0.1060995431</td>
<td>0.1336844907</td>
<td>0.1697118152</td>
<td>0.2169643536</td>
<td>0.2811053381</td>
<td>0.3793122556</td>
</tr>
<tr>
<td>0.06</td>
<td>0.0564009777</td>
<td>0.0616057558</td>
<td>0.0713082824</td>
<td>0.0871581833</td>
<td>0.1085295710</td>
<td>0.1368735157</td>
<td>0.1738702420</td>
<td>0.2224809499</td>
<td>0.2888732522</td>
<td>0.3933192258</td>
</tr>
<tr>
<td>0.07</td>
<td>0.0567066983</td>
<td>0.0623922724</td>
<td>0.0730266953</td>
<td>0.0890144747</td>
<td>0.1110288192</td>
<td>0.1401481121</td>
<td>0.1781426083</td>
<td>0.2281682700</td>
<td>0.2969721099</td>
<td>0.409062849</td>
</tr>
<tr>
<td>0.08</td>
<td>0.0570597366</td>
<td>0.0632280263</td>
<td>0.0743578804</td>
<td>0.0909307119</td>
<td>0.1135985821</td>
<td>0.1435104314</td>
<td>0.1825328587</td>
<td>0.2340351233</td>
<td>0.3054269741</td>
<td>0.4272272392</td>
</tr>
<tr>
<td>0.09</td>
<td>0.0574602150</td>
<td>0.0641134550</td>
<td>0.0757789125</td>
<td>0.0929077445</td>
<td>0.1162402177</td>
<td>0.1469627471</td>
<td>0.1870452202</td>
<td>0.2409132085</td>
<td>0.3142773322</td>
<td>0.4501481060</td>
</tr>
</tbody>
</table>

Note that (a) $\Omega_3(0) = 1/18$, (b) $\Omega_3(1) = 1/2$, and (c) $\Omega_3(-\rho) = \Omega_3(\rho)$. 
<table>
<thead>
<tr>
<th>Sample size $n = 10$</th>
<th>Sample size $n = 20$</th>
<th>Sample size $n = 30$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V(\hat{\rho}_P)$</td>
<td>$V(\hat{\rho}_S)$</td>
<td>$V(\hat{\rho}_K)$</td>
</tr>
<tr>
<td>$\rho$ (53) Obs. (54) Obs. (55) Obs. (56) Obs.</td>
<td>$\rho$ (53) Obs. (54) Obs. (55) Obs. (56) Obs.</td>
<td>$\rho$ (53) Obs. (54) Obs. (55) Obs. (56) Obs.</td>
</tr>
<tr>
<td>0.0 0.1111 0.1220 0.1190 0.1520 0.1330 0.1480 0.1430 0.0530 0.0580 0.0570 0.0650 0.0610 0.0640 0.0630 0.0340 0.0340 0.0370 0.0370 0.0410 0.0390 0.0400 0.0400</td>
<td>0.1 0.1090 0.1200 0.1170 0.1500 0.1310 0.1450 0.1410 0.0520 0.0520 0.0570 0.0560 0.0640 0.0600 0.0610 0.0620 0.0340 0.0340 0.0370 0.0370 0.0400 0.0390 0.0400 0.0390</td>
<td>0.2 0.1020 0.1050 0.1150 0.1120 0.1420 0.1250 0.1380 0.1350 0.0490 0.0490 0.0540 0.0540 0.0600 0.0570 0.0590 0.0590 0.0320 0.0320 0.0350 0.0350 0.0350 0.0350 0.0380 0.0370 0.0370</td>
</tr>
<tr>
<td>0.3 0.0920 0.0960 0.1060 0.1050 0.1290 0.1250 0.1270 0.1250 0.0440 0.0450 0.0490 0.0490 0.0550 0.0520 0.0540 0.0540 0.0290 0.0290 0.0320 0.0320 0.0340 0.0340 0.0340 0.0340</td>
<td>0.4 0.0780 0.0850 0.0940 0.0940 0.1130 0.1040 0.1120 0.1110 0.0370 0.0390 0.0430 0.0430 0.0470 0.0460 0.0470 0.0470 0.0240 0.0250 0.0280 0.0280 0.0300 0.0290 0.0300 0.0300</td>
<td>0.5 0.0630 0.0710 0.0800 0.0810 0.0950 0.0890 0.0930 0.0940 0.0300 0.0320 0.0360 0.0360 0.0390 0.0380 0.0390 0.0390 0.0190 0.0200 0.0230 0.0230 0.0240 0.0240 0.0240 0.0250</td>
</tr>
<tr>
<td>0.6 0.0460 0.0550 0.0630 0.0650 0.0700 0.0710 0.0730 0.0740 0.0220 0.0240 0.0280 0.0280 0.0290 0.0290 0.0300 0.0300 0.0140 0.0150 0.0180 0.0180 0.0180 0.0180 0.0180 0.0190 0.0190</td>
<td>0.7 0.0290 0.0380 0.0460 0.0480 0.0480 0.0520 0.0510 0.0530 0.0140 0.0160 0.0190 0.0200 0.0190 0.0200 0.0200 0.0210 0.0090 0.0100 0.0120 0.0120 0.0120 0.0120 0.0120 0.0130</td>
<td>0.8 0.0140 0.0210 0.0280 0.0300 0.0260 0.0260 0.0230 0.0230 0.0070 0.0080 0.0110 0.0110 0.0100 0.0110 0.0110 0.0110 0.0040 0.0050 0.0070 0.0070 0.0060 0.0070 0.0070 0.0070</td>
</tr>
<tr>
<td>0.9 0.0040 0.0070 0.0120 0.0130 0.0090 0.0120 0.0110 0.0110 0.0020 0.0020 0.0040 0.0040 0.0030 0.0040 0.0040 0.0040 0.0010 0.0010 0.0020 0.0020 0.0020 0.0020 0.0020 0.0020</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The orthant probability

\[ R \]

is a compact notation of

\[ n \]

\[ g \]

\[ q \]

\[ j \]

\[ h \]

\[ i \]

\[ f \]

\[ e \]

\[ d \]

\[ c \]

\[ b \]

\[ a \]

\[ X \]

\[ Y \]

\[ Z \]

Corr. Matrix No. of Terms

<table>
<thead>
<tr>
<th>Corr. Matrix</th>
<th>No. of Terms*</th>
<th>Representative Term</th>
<th>Correlation Coefficients</th>
<th>( P_4^0(Z_1, Z_2, Z_3, Z_4) )†</th>
<th>( W(0) )</th>
<th>( W(1) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_a )</td>
<td>( n^6 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_b )</td>
<td>( 2n^5 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_c )</td>
<td>( n^5 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_d )</td>
<td>( 4n^5 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_e )</td>
<td>( 2n^5 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_f )</td>
<td>( 2n^5 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_g )</td>
<td>( 4n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_h )</td>
<td>( 4n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_i )</td>
<td>( 4n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_j )</td>
<td>( n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_k )</td>
<td>( 2n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_l )</td>
<td>( 4n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_m )</td>
<td>( n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_n )</td>
<td>( 2n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_o )</td>
<td>( 2n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_p )</td>
<td>( 2n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_q )</td>
<td>( 4n^4 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_r )</td>
<td>( 3n^3 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_s )</td>
<td>( n^3 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_t )</td>
<td>( 2n^3 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_u )</td>
<td>( 2n^3 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_v )</td>
<td>( 4n^3 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_w )</td>
<td>( 2n^3 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>( R_x )</td>
<td>( n^2 )</td>
<td>( X_1 \cdots X_4 )</td>
<td>0 0 0 0 0 4 ( \rho )</td>
<td>( 1^{16} + \frac{S_1}{\pi} + \frac{S_2}{\pi^2} )</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

* The symbol \( n^k \) is a compact notation of \( n(n-1) \cdots (n-k+1) \), with \( k \) being a positive integer.
† The orthant probability \( P_4^0(Z_1, Z_2, Z_3, Z_4) \) \( = \mathbb{E} \{ H(Z_1)H(Z_2)H(Z_3)H(Z_4) \} \). Notations \( S_1 \triangleq \sin^{-1} \rho \) and \( S_2 \triangleq \sin^{-1} \frac{1}{2} \rho \) are used for brevity.
### Table 5: Quantities for evaluation of $E(ST)$ in Lemma 4

<table>
<thead>
<tr>
<th>Corr. Matrix</th>
<th>No. of Terms*</th>
<th>Representative Term</th>
<th>$P_0(z_1, z_2, z_3, z_4)$</th>
<th>$P_0'(z_2, z_3, z_4)$</th>
<th>$P_0'(z_1, z_2, z_3, z_4)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R_0$</td>
<td>$n^{10}$</td>
<td>$X_1 - X_2$ $Y_1 - Y_2$ $X_3 - X_4$ $Y_3 - Y_4$ $Z_1$</td>
<td>$1 + \frac{S_1}{4\pi} + \frac{S_2}{8\pi} + \frac{S_3}{16\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
</tr>
<tr>
<td>$R_1$</td>
<td>$n^{11}$</td>
<td>$X_1 - X_2$ $Y_1 - Y_2$ $X_3 - X_4$ $Y_3 - Y_4$ $Z_1$</td>
<td>$1 + \frac{S_1}{4\pi} + \frac{S_2}{8\pi} + \frac{S_3}{16\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
</tr>
<tr>
<td>$R_2$</td>
<td>$n^{12}$</td>
<td>$X_1 - X_2$ $Y_1 - Y_2$ $X_3 - X_4$ $Y_3 - Y_4$ $Z_1$</td>
<td>$1 + \frac{S_1}{4\pi} + \frac{S_2}{8\pi} + \frac{S_3}{16\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
</tr>
<tr>
<td>$R_3$</td>
<td>$n^{13}$</td>
<td>$X_1 - X_2$ $Y_1 - Y_2$ $X_3 - X_4$ $Y_3 - Y_4$ $Z_1$</td>
<td>$1 + \frac{S_1}{4\pi} + \frac{S_2}{8\pi} + \frac{S_3}{16\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
</tr>
<tr>
<td>$R_4$</td>
<td>$n^{14}$</td>
<td>$X_1 - X_2$ $Y_1 - Y_2$ $X_3 - X_4$ $Y_3 - Y_4$ $Z_1$</td>
<td>$1 + \frac{S_1}{4\pi} + \frac{S_2}{8\pi} + \frac{S_3}{16\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
</tr>
<tr>
<td>$R_5$</td>
<td>$n^{15}$</td>
<td>$X_1 - X_2$ $Y_1 - Y_2$ $X_3 - X_4$ $Y_3 - Y_4$ $Z_1$</td>
<td>$1 + \frac{S_1}{4\pi} + \frac{S_2}{8\pi} + \frac{S_3}{16\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
<td>$\frac{1}{8} + \frac{S_2}{8\pi}$</td>
</tr>
</tbody>
</table>

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

† $P_0(z_1, z_2, z_3, z_4) \triangleq E \{ H(z_1)H(z_2)H(z_3)H(z_4) \}$, $P_0'(z_2, z_3, z_4) \triangleq E \{ H(z_2)H(z_3)H(z_4) \}$, and $P_0'(z_1, z_2, z_3, z_4) \triangleq E \{ H(z_1)H(z_2)H(z_3)H(z_4) \}$. Notations $S_1 \triangleq \sin^{-1} \rho$ and $S_2 \triangleq \sin^{-1} \frac{1}{2} \rho$ are used for brevity.