

On The Positive Definiteness of Polarity Coincidence Correlation Coefficient Matrix

F. Haddadi, *Student Member, IEEE*, M. M. Nayebi, *Senior Member, IEEE*, and M. R. Aref

Abstract—Polarity coincidence correlator (PCC), when used to estimate the covariance matrix on an element-by-element basis, may not yield a positive semi-definite (PSD) estimate. Devlin et al. [1], claimed that element-wise PCC is not guaranteed to be PSD in dimensions $p > 3$ for real signals. However, no justification or proof was available on this issue. In this letter, it is proved that for real signals with $p \leq 3$ and for complex signals with $p \leq 2$, a PSD estimate is guaranteed. Counterexamples are presented for higher dimensions which yield invalid covariance estimates.

Index Terms—Polarity coincidence correlator, element-wise covariance estimate, positive semi-definite.

I. INTRODUCTION AND PRELIMINARIES

POLARITY coincidence correlator (PCC) is a robust and nonparametric estimator of bivariate correlation [1], [2]. It is also a fast and low-cost estimator for applications with extraordinary computational complexity. Radio astronomy is an instance in which PCC is by far the most favorable correlator [3].

Several researchers have investigated the statistical error of PCC as an estimate of bivariate correlation [4], [5]. In multivariate case, using PCC to estimate elements of the covariance matrix does not guarantee a PSD matrix estimator [1], [6]. Devlin et al. [1, Sec. 4.4], referring to a personal communication, claim that element-wise PCC (or "quadrant correlation"), may yield an invalid covariance estimate for $p > 3$ and real signals.

In this letter, we prove that for real signals with $p \leq 3$, and complex signals with $p \leq 2$, PCC estimate is PSD. For higher dimensions, counterexamples are presented which yield invalid covariance estimates.

Let x and y be two zero-mean real random variables with correlation coefficient r distributed with elliptical symmetry. It is well known that [6]:

$$r = \sin\left(\frac{\pi}{2}E\{\text{sgn}(x)\text{sgn}(y)\}\right) \quad (1)$$

This work was supported by Advanced Communication Research Institute (ACRI), Sharif University of Technology, Tehran, Iran. Authors are with the Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran (e-mails: farzanhadadi@yahoo.com, Nayebi@sharif.edu, and Aref@sharif.edu).

where

$$\text{sgn}(x) = \begin{cases} +1 & : x \geq 0 \\ -1 & : x < 0 \end{cases} \quad (2)$$

Using (1), an estimate of r from N iid observations $x_i, y_i, i = 1, \dots, N$ is given by

$$\hat{r} = \sin\left(\frac{\pi}{2} \frac{1}{N} \sum_{i=1}^N s_{xi}s_{yi}\right). \quad (3)$$

where $s_{xi} = \text{sgn}(x_i)$. In the complex case, we can define the complex sign function as $\text{sgn}_c(x) \triangleq \text{sgn}(\Re[x]) + j \text{sgn}(\Im[x])$, where $\Re[x]$ and $\Im[x]$ are real and imaginary parts of x , respectively. In the Appendix, it is shown that

$$\begin{aligned} \Re[r] &= \sin\left(\frac{\pi}{4}E\{\Re[\text{sgn}_c(x)\text{sgn}_c^*(y)]\}\right) \\ \Im[r] &= \sin\left(\frac{\pi}{4}E\{\Im[\text{sgn}_c(x)\text{sgn}_c^*(y)]\}\right) \end{aligned} \quad (4)$$

where $(\cdot)^*$ denotes complex conjugate. Similar to (3), an estimate for the complex case is obtained by replacing expectation with the average as

$$\begin{aligned} \hat{r}_R &= \sin\left(\frac{\pi}{4} \frac{1}{N} \sum_{i=1}^N [s_{xiR}s_{yiR} + s_{xiI}s_{yiI}]\right) \\ \hat{r}_I &= \sin\left(\frac{\pi}{4} \frac{1}{N} \sum_{i=1}^N [s_{xiI}s_{yiR} - s_{xiR}s_{yiI}]\right) \end{aligned} \quad (5)$$

where $(\cdot)_R$ and $(\cdot)_I$ denote real and imaginary parts, respectively.

II. MAIN RESULT

Let $\mathbf{R}_{p \times p}$ be the covariance matrix of p random signals with unit diagonal elements and off-diagonal elements $r_{ij} : i, j = 1, \dots, p$. For $p = 2$ case, a valid correlation estimate should satisfy $|\hat{r}| \leq 1$. For the real case of (3), $|\hat{r}| = |\sin(\cdot)| \leq 1$. For the complex case, regarding (5) define α and β such that $\hat{r}_R = \sin(\alpha)$ and $\hat{r}_I = \sin(\beta)$. Then

$$\alpha + \beta = \frac{\pi}{4N} \sum_{i=1}^N [s_{xiR}(s_{yiR} - s_{yiI}) + s_{xiI}(s_{yiR} + s_{yiI})] \quad (6)$$

and it can be easily checked that the argument of summation in (6) belongs to $\{\pm 2\}$. This yields $\alpha + \beta \leq \frac{\pi}{2}$. In

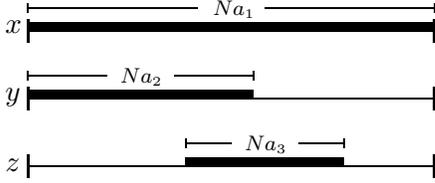


Fig. 1. Polarity coincidence diagram of x, y, z . Black strips denote the packed positions of polarity coincidences of each signal with signal x . The strips lengths Na_i are the number of polarity coincidences.

the same manner, we can show that $\pm \alpha \pm \beta \leq \frac{\pi}{2}$ which gives $|\alpha| + |\beta| \leq \frac{\pi}{2}$. Now it is straightforward to see that $|\hat{r}|^2 = \sin^2(\alpha) + \sin^2(\beta) \leq \sin^2(|\alpha|) + \sin^2(\frac{\pi}{2} - |\alpha|) = 1$.

For $p = 3$ and real signals, we calculate the valid range of the elements of a 3×3 covariance matrix. Then we show that PCC estimate lies in this range.

A. Valid Range of Covariance

Let $\mathbf{R} \in \mathbb{R}^{3 \times 3}$ be a covariance matrix with unit diagonal elements. Valid range of r_{23} should be calculated when $r_{12}, r_{13} \in [-1, +1]$ are fixed. It can be readily shown that $|\mathbf{R}| \geq 0$ implies that

$$|r_{23} - r_{12}r_{13}| \leq \sqrt{(1 - r_{12}^2)(1 - r_{13}^2)}. \quad (7)$$

B. PCC Covariance Estimate

Assume random sign sequences s_x, s_y, s_z with length N . Consider the positions of polarity coincidence with s_x as black positions or "+" and elsewhere as white or "-". Obviously all of the positions in s_x is "+" and (s_{yi}, s_{zi}) have four states of $\{++, +-, -+, --\}$. Since the permutation of the samples does not affect the estimate in (3), put the samples of s_x, s_y, s_z from left in the order of $\{++ + -, ++ ++, + - +, + - -\}$ as in Fig. 1. Then any random sign sequences of s_x, s_y and s_z can be replaced by the model in Fig. 1 with appropriate strip lengths Na_i (with $a_1 = 1$) and relative positions of strips.

Let \mathbf{R}_s be the covariance matrix of s_x, s_y, s_z with elements $r_{sik}, i, k = 1, 2, 3$. The maximum of $r_{s12} = +1$ occurs in $a_2 = 1$ and the minimum of $r_{s12} = -1$ in $a_2 = 0$. In fact, $r_{s12} = \frac{1}{N} \sum_{i=1}^N s_{xi}s_{yi} = \frac{1}{N}[Na_2 - (N - Na_2)] = 2a_2 - 1$, in other words

$$a_i = \frac{1 + r_{s1i}}{2}. \quad (8)$$

r_{s12} and r_{s13} are determined by the values of a_2 and a_3 , $r_{sii} = 1$, and the possible range of r_{s23} should be calculated. r_{s23} depends on the number of polarity coincidences of y and z which is maximum when the

TABLE I
COUNTEREXAMPLES FOR REAL AND COMPLEX DATA

Real Case		Complex Case	
s_x	+++ +	s_x	++ ++
s_y	++ --	s_y	++ -+
s_z	+++ -	s_z	++ --
s_w	++ - +		

strip of z is in the left corner, and minimum when it is in the right corner. After some calculations, the range of r_{s23} is found as

$$|r_{s12} + r_{s13}| - 1 \leq r_{s23} \leq 1 - |r_{s12} - r_{s13}|. \quad (9)$$

It should be noted that the effect of finite N is the quantization of the accessible values. Now, it can be readily verified that

$$\sin\left(\frac{\pi}{2}(1 - |r_{s12} - r_{s13}|)\right) = r_{12}r_{13} + \sqrt{(1 - r_{12}^2)(1 - r_{13}^2)} \quad (10)$$

and

$$\sin\left(\frac{\pi}{2}(|r_{s12} + r_{s13}| - 1)\right) = r_{12}r_{13} - \sqrt{(1 - r_{12}^2)(1 - r_{13}^2)}. \quad (11)$$

Therefore, $\hat{r}_{23} = \sin\left(\frac{\pi}{2}r_{s23}\right)$ satisfies (7). This, besides $|\hat{r}_{12}| < 1$ and $|\hat{r}_{13}| < 1$ can be used to show that $|\hat{\mathbf{R}}| \geq 0$ (as in (7)) and the assertion is proved that for $p = 3$ and real data, PCC estimate is a valid covariance matrix.

III. COUNTEREXAMPLES

In this section, some counterexamples are presented to show that PCC covariance estimate is not guaranteed to be PSD in dimensions $p > 3$ for real signals and $p > 2$ for complex signals. In real data case with $p = 4$ and number of observations $N = 4$, the real sign sequences in Table I results in an invalid covariance estimate. After simple computations, we will have $r_{s12} = r_{s34} = 0$ and $r_{s13} = r_{s14} = r_{s23} = r_{s24} = 0.5$. The covariance estimate will be

$$\hat{\mathbf{R}}_1 = \begin{bmatrix} 1 & 0 & 0.7 & 0.7 \\ 0 & 1 & 0.7 & 0.7 \\ 0.7 & 0.7 & 1 & 0 \\ 0.7 & 0.7 & 0 & 1 \end{bmatrix}$$

with eigenvalues $[-0.4, 1, 1, 2.4]$. Then $\hat{\mathbf{R}}_1$, with a negative eigenvalue, is not a valid covariance matrix.

We can augment this example to give a counterexample for dimension $p = 5$. Repeat each sign twice to have four signals with number of observations $2N$.

Note that the covariance matrix does not change. Now, add a new signal with alternating sign in each sample. The covariance estimate will be

$$\hat{\mathbf{R}}_{aug} = \begin{bmatrix} \hat{\mathbf{R}}_1 & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix}.$$

where $\mathbf{0}$ is the 4×1 vector of zeros. As a consequence of the structure of $\hat{\mathbf{R}}_{aug}$, eigenvalues of $\hat{\mathbf{R}}_1$ are also eigenvalues of $\hat{\mathbf{R}}_{aug}$. Therefore, $\hat{\mathbf{R}}_{aug}$ is an invalid covariance matrix. This procedure can continue to produce counterexamples for higher dimensions in real data case.

In case of complex signals, $p = 3$ and $N = 2$, a counterexample is given in Table I, where "−+" denotes $-1 + j$. The resulting estimate is

$$\hat{\mathbf{R}} = \begin{bmatrix} 1 & 0.7 - j0.7 & 0 \\ 0.7 + j0.7 & 1 & 0.7 - j0.7 \\ 0 & 0.7 + j0.7 & 1 \end{bmatrix}$$

with eigenvalues $[-0.4, 1, 2.4]$ which make $\hat{\mathbf{R}}$ an invalid covariance matrix. Augmentation of the complex signal set for higher dimensions is similar to the real case, except that the new added signal alternates between "++" and "--".

IV. APPLICATIONS OF THE RESULTS

In this section, we discuss the practical usefulness of the main results of this letter which focuses on low number of sensors. In the signal processing context, covariance estimation often arises in the multi-sensor applications where parameters of interest are functions of the true data covariance matrix. Although PCC estimate of the covariance matrix exhibits attractive features such as robustness and extremely low complexity, it cannot be guaranteed to be PSD in the applications with large number of sensors.

Selection of the number of sensors in an application depends on both nature of the problem and practical limitations. In theory, more sensors always results in a better estimate, as proved in many cases such as direction finding through examination of the Cramer-Rao bounds [7]. In practice, complexity issues usually limit the number of sensors. Large arrays are used whenever performance be of the main importance regardless of the cost. In such cases as DOA estimation in military environments (radar and sonar), thousands of sensors are not uncommon. Nevertheless, most low-cost civil applications use very few sensors. In the following, we consider some of these applications.

A. MIMO Communication Systems

Multiple antenna systems are an integral part of the most new wireless communication systems increasing user and data capacity (e.g. UMTS/W-CDMA, 802.11n WLAN, 60 GHz WPAN). Multiple antennas can provide diversity gain and/or better antenna gain through beamforming in base station and/or handset. Beamformers (e.g. conventional or Capon) usually utilize an estimate of the array covariance matrix [8], that may be obtained using PCC as a power-saving estimator. It is well known that performance improvement due to diversity gain reduces as the number of antennas increases. This, besides space limit on the handset and coupling phenomena have resulted in the prevalence of MIMO systems with very few (usually 2 to 4) antennas [9], [10].

B. Blind Source Separation (BSS)

BSS has found numerous potential applications in the field of audio signal processing [11]. An array of microphones is used to gather multiple signal mixtures and diverse methods are used to extract signals from these observations. A large class of BSS methods use real-valued inter-sensor covariances with different time lags to estimate the mixing matrix and desired signals (e.g. SOBI [12], JADE [13]). This also includes input signals whitening as a preprocessing that converts the convolutive source separation problem to a simpler independent component analysis (ICA) problem. This family of two-step algorithms is known as AMUSE (Algorithm for Multiple Source Extraction). PCC, as a fast correlator, can make real-time operation more feasible in these methods. For realistic situations where we have fewer sensors than sources, underdetermined methods are proposed [14]. Many methods are presented for the special case of 2 sensors and multiple sources (e.g. DUET [15], and [16]), and also quite few sensors are common to many realizations of the methods [12], [14].

APPENDIX COMPLEX PCC

Let x, y be two zero-mean, unit-variance, and circularly symmetric complex random variables with independent real and imaginary parts. To prove (4), we expand the expectation as

$$E\{\text{sgn}_c(x)\text{sgn}_c^*(y)\} = E\{[\text{sgn}(x_{RYR}) + \text{sgn}(x_{IYI})] + j[\text{sgn}(x_{IYR}) - \text{sgn}(x_{RYI})]\}. \quad (12)$$

Furthermore, $E\{xy^*\} = r$ implies that

$$\begin{aligned} E\{x_{RYR} + x_{IYI}\} &= r_R \\ E\{x_{IYR} - x_{RYI}\} &= r_I. \end{aligned} \quad (13)$$

Circular symmetry of x and y yields

$$\begin{aligned} E\{x_R y_R\} &= E\{x_I y_I\} = r_R/2 \\ E\{x_I y_R\} &= -E\{x_R y_I\} = r_I/2 \end{aligned} \quad (14)$$

and $E\{x_R^2\} = E\{x_I^2\} = E\{y_R^2\} = E\{y_I^2\} = \frac{1}{2}$. Then the correlation coefficients will be

$$\begin{aligned} \text{Cor}(x_R, y_R) &= \text{Cor}(x_I, y_I) = r_R \\ \text{Cor}(x_I, y_R) &= -\text{Cor}(x_R, y_I) = r_I. \end{aligned} \quad (15)$$

Substituting (15) and (1) into (12) gives

$$E\{\text{sgn}_c(x)\text{sgn}_c^*(y)\} = 4/\pi [\sin^{-1}(r_R) + j \sin^{-1}(r_I)] \quad (16)$$

which implies (4).

REFERENCES

- [1] S. J. Devlin, R. Gnanadesikan, and J. R. Kettenring, "Robust estimation and outlier detection with correlation coefficients", *Biometrika*, vol. 62, No. 3, pp. 531-545, March 1975.
- [2] S. S. Wolff, J. B. Thomas, and T. R. Williams, "The polarity-coincidence correlator: a nonparametric detection device", *IRE Trans. Information Theory*, vol. 8, No. 1, pp. 5-9, Jan. 1962.
- [3] P. C. Egau, "Correlation systems in radio astronomy and related fields", *IEE Proc.*, vol. 131, Pt. F, No. 1, pp. 32-39, Feb. 1984.
- [4] K. J. Gabriel, "Comparison of three correlation coefficient estimators for gaussian stationary processes", *IEEE Trans. Acoustic, Speech, Signal Proc.*, vol. ASSP-31, No. 4, pp. 1023-1025, Aug. 1983.
- [5] G. Jacovitti and R. Cusani, "Performance of normalized correlation estimators for complex processes", *IEEE Trans. Signal Proc.*, vol. 40, No. 1, pp. 114-128, Jan. 1992.
- [6] S. Visuri, V. Koivunen, and H. Oja, "Sign and rank covariance matrices", *J. Stat. Plan. Infer.*, vol. 91, No. 2, pp. 557-575, Feb. 2000.
- [7] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and Cramer-Rao bound", *IEEE Trans. Acoustic Speech Sig. Proc.*, vol 37, No. 5, pp. 720-741, May 1989.
- [8] B. D. Van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering", *IEEE ASSP Mag.*, pp. 4-24, Apr. 1988.
- [9] H. T. Hui, "Practical dual-helical antenna array for diversity-MIMO receiving antennas on mobile handsets", *IEE Proc. Microw. Antennas Propag.*, vol. 152, No. 5, pp. 367-372, Oct. 2005.
- [10] D. Browne, M. Manteghi, M. P. Fitz, and Y. Rahmat-samii, "Experiments with compact antenna arrays for MIMO radio communications", *IEEE Trans. Ant. Propag.*, vol. 54, No. 11, pp. 3239-3259, Nov. 2006.
- [11] K. Torkkola, "Blind separation for audio signals - are we there yet?", in *Proc. Workshop on Independent Component Analysis and Blind Signal Separation*, Aussois, France, Jan. 11-15, 1999.
- [12] A. Belouchrani, K. Abed-Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second-order statistics", *IEEE Trans. Signal Proc.*, vol. 45, No. 2, pp. 434-444, Feb. 1997.
- [13] J. F. Cardoso, "High-order contrast for independent component analysis", *Neural Computation*, vol. 11, No. 1, pp. 157-192, 1999.
- [14] S. Winter, W. Kellermann, H. Sawada, and S. Makino, "MAP-based underdetermined blind source separation of convolutive mixtures by hierarchical clustering and l1-norm minimization", *EURASIP J. Advances Signal Proc.*, pp. 1-12, 2007.
- [15] O. Yilmaz and S. Rickard, "Blind separation of speech mixtures via time-frequency masking", *IEEE Trans Signal Proc.*, vol. 52, No. 7, pp. 1830-1847, July 2004.
- [16] P. Comon, "Blind identification and source separation in 2×3 under-determined mixtures", *IEEE Trans. Signal Proc.*, vol. 52, No. 1, pp. 11-12, Jan. 2004.