

Contrast Functions for Non-circular and Circular Sources Separation in Complex-Valued ICA

Y. S. Kim and S. S. Park

Abstract—In this paper, the complex-valued ICA problem is studied in the context of blind complex-source separation. We formulate the complex ICA problem in a general setting, and define the superadditive functional that may be used for constructing a contrast function for circular complex sources separation. We propose several contrast functions and study their properties. Finally, we also discuss relevant issues and present the convex analysis of a specific contrast function.

1. INTRODUCTION

In this paper, we study the complex-valued ICA problem in the context of blind complex-source separation. We formulate the complex ICA problem in a general setting, and define the superadditive functional that may be used for constructing a contrast function for circular complex sources separation. We propose several contrast functions and study their properties. Finally, we also discuss relevant issues and present the convex analysis of a specific contrast function.

2. PROBLEM FORMULATION

Let $\mathbf{X} = (X_1, \dots, X_k)$ be a random vector of k independent complex-valued sources, where X_k is a complex-valued random variable with probability density function H_k , $k = 1, \dots, k$. Let $\mathbf{A} = (A_1, \dots, A_k)$ be a matrix of k independent complex-valued sources, where A_k is a complex-valued random variable with probability density function A_k , $k = 1, \dots, k$.

3. CONCLUSION

z)

- **Linearity of expectation:**

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y].$$
- **Variance of sum of independent variables:**

$$\mathbb{E}[X^2] = \mathbb{E}[X]^2 + \text{var}[X].$$
- **Variance of sum of correlated variables:**

$$\text{var}[X + Y] = \text{var}[X] + \text{var}[Y] + 2\text{cov}[X, Y].$$

Covariance matrix:

$$C_{ij} = \mathbb{E}[(x_i - \mathbb{E}[x_i])(x_j^* - \mathbb{E}[x_j^*])] = \mathbb{E}[x_i x_j^*] - \mathbb{E}[x_i]\mathbb{E}[x_j^*].$$

uncorrelated $C_{ij} = 0$ for $i \neq j$.

Complex random vector:

$$\mathbf{z} = [z_1, \dots, z_n]^T, \quad \mathbf{z}^H = [z_1^*, \dots, z_n^*]^T \equiv (\mathbf{z}^*)^T \mathbf{H}$$

$\mathbb{E}[\mathbf{z}]$, $\text{cov}[\mathbf{z}] \equiv \mathbb{E}[(\mathbf{z} - \mathbb{E}[\mathbf{z}])(\mathbf{z} - \mathbb{E}[\mathbf{z}])^H]$, **pseudo-covariance matrix**
 $\text{pcov}[\mathbf{z}] \equiv \mathbb{E}[(\mathbf{z} - \mathbb{E}[\mathbf{z}])(\mathbf{z} - \mathbb{E}[\mathbf{z}])^T]$.

Definition 1: A $n \times n$ matrix \mathbf{A} is called **Hermitian** if $\mathbf{A} = \mathbf{A}^H$.
 A $n \times n$ matrix \mathbf{A} is called **circulant** if $A_{ij} = A_{(i-j) \bmod n, j}$.
 A $n \times n$ matrix \mathbf{A} is called **second-order circular** if $\mathbb{E}[z_i] = 0$ and $\mathbb{E}[z_i^2] = 0$.

Strongly uncorrelated: $\mathbb{E}[\mathbf{z}\mathbf{z}^H] = \mathbf{I}$.
Symmetric: $\mathbb{E}[\mathbf{z}\mathbf{z}^T] = \mathbf{A}$, $\mathbf{A} = \mathbf{A}^T$.
Zero-mean: $\mathbb{E}[\mathbf{z}\mathbf{z}^T] = \mathbf{0}$.

skewness $(\kappa) = \mathbb{E}[|z|^3] / (\mathbb{E}[|z|^2])^{3/2}$, (1)
kurtosis $(\kappa) = \mathbb{E}[|z|^4] - 2(\mathbb{E}[|z|^2])^2 - |\mathbb{E}[z^2]|^2$. (2)

Definition 2: The **order** of a random variable X is the number of non-zero eigenvalues of its covariance matrix. For a complex random vector \mathbf{z} , the order is the number of non-zero eigenvalues of $\text{cov}[\mathbf{z}]$.

Jacobian: $J(\mathbf{z}) = \frac{\partial J(\mathbf{z})}{\partial \mathbf{z}^*}$, $\mathbf{z} \in \mathbb{C}^N$.

Gradient of a scalar function:

$$\nabla J \equiv \frac{\partial J(\mathbf{z})}{\partial \mathbf{z}^*} = \frac{1}{2} \left(\frac{\partial J(\mathbf{z})}{\partial \mathbf{z}} + \frac{\partial J(\mathbf{z})}{\partial \mathbf{z}} \right) = 0.$$

$\frac{\partial J(\mathbf{z})}{\partial \mathbf{x}} = \frac{\partial J(\mathbf{z})}{\partial \mathbf{z}} = 0$.

Definition 3: A function $J(\mathbf{z})$ is called **convex** if $J(\lambda \mathbf{z}_1 + (1-\lambda)\mathbf{z}_2) \leq \lambda J(\mathbf{z}_1) + (1-\lambda)J(\mathbf{z}_2)$ for $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{C}^N$ and $0 \leq \lambda \leq 1$.

$J(\lambda \mathbf{z}_1 + (1-\lambda)\mathbf{z}_2) \leq \lambda J(\mathbf{z}_1) + (1-\lambda)J(\mathbf{z}_2)$

$\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{C}^N$, $0 \leq \lambda \leq 1$.
Hessian: $\mathbf{H} = \frac{\partial^2 J(\mathbf{z})}{\partial \mathbf{z} \partial \mathbf{z}^H}$.

Jacobian of a vector function: $J(\mathbf{z}) = \frac{\partial \mathbf{y}}{\partial \mathbf{z}^*}$.

Linear transformation: $\mathbf{y} = \mathbf{A}\mathbf{x}$, $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{x} \in \mathbb{C}^n$.
Orthogonal matrix: $\mathbf{A}^H = \mathbf{A}^{-1}$.
Block matrix: $\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$, $\mathbf{A}_{11} \in \mathbb{C}^{k \times k}$, $\mathbf{A}_{22} \in \mathbb{C}^{(n-k) \times (n-k)}$.

- **Block diagonal matrix:** $\mathbf{A} = \text{diag}(\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k)$.
- **Block upper triangular matrix:** $\mathbf{A}_{ij} = \mathbf{0}$ for $i > j$.
- **Block lower triangular matrix:** $\mathbf{A}_{ij} = \mathbf{0}$ for $i < j$.

Block matrix inversion: $\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}_{11}^{-1} & \dots \\ \dots & \dots \end{bmatrix}$.

Block matrix multiplication: $\mathbf{A} \mathbf{B} = \begin{bmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \dots \\ \dots & \dots \end{bmatrix}$.

Block matrix addition: $\mathbf{A} + \mathbf{B} = \begin{bmatrix} \mathbf{A}_{11} + \mathbf{B}_{11} & \dots \\ \dots & \dots \end{bmatrix}$.

Block matrix subtraction: $\mathbf{A} - \mathbf{B} = \begin{bmatrix} \mathbf{A}_{11} - \mathbf{B}_{11} & \dots \\ \dots & \dots \end{bmatrix}$.

Block matrix scalar multiplication: $c\mathbf{A} = \begin{bmatrix} c\mathbf{A}_{11} & \dots \\ \dots & \dots \end{bmatrix}$.

Block matrix transpose: $(\mathbf{A}^T)_{ij} = A_{ji}$.

Block matrix Hermitian transpose: $(\mathbf{A}^H)_{ij} = A_{ji}^*$.

Block matrix inverse: $(\mathbf{A}^{-1})_{ij} = (A_{ji})^{-1}$.

Block matrix determinant: $|\mathbf{A}| = |\mathbf{A}_{11}| |\mathbf{A}_{22}|$.

Block matrix trace: $\text{tr}(\mathbf{A}) = \text{tr}(\mathbf{A}_{11}) + \text{tr}(\mathbf{A}_{22})$.

Block matrix rank: $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}_{11}) + \text{rank}(\mathbf{A}_{22})$.

Block matrix eigenvalues: λ is an eigenvalue of \mathbf{A} if and only if λ is an eigenvalue of \mathbf{A}_{11} or \mathbf{A}_{22} .

Block matrix eigenvectors: \mathbf{v} is an eigenvector of \mathbf{A} if and only if \mathbf{v} is an eigenvector of \mathbf{A}_{11} or \mathbf{A}_{22} .

Block matrix singular values: σ is a singular value of \mathbf{A} if and only if σ is a singular value of \mathbf{A}_{11} or \mathbf{A}_{22} .

Block matrix singular vectors: \mathbf{u} and \mathbf{v} are singular vectors of \mathbf{A} if and only if \mathbf{u} and \mathbf{v} are singular vectors of \mathbf{A}_{11} or \mathbf{A}_{22} .

Block matrix singular value decomposition: $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H$.

Block matrix singular value decomposition: $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^H$.

$$(6), \quad \mathbf{W} = \mathbf{A}^{-1} \mathbf{y} \quad (7)$$

7, 13 :

$$\Delta \mathbf{W} = \eta(\mathbf{I} - \psi(\mathbf{y})\mathbf{y}^H)\mathbf{W}. \quad (9)$$

Liouville's theorem, boundedness, analyticity

$\psi(\cdot)$ (7-9, 13, 14);

$$I(1, \dots, n) = \mathbb{E}_{p(\mathbf{y})} \left[\log \frac{p(\mathbf{y})}{\prod_{i=1}^n p(i)} \right] \\ = \int p(\mathbf{y}) \log \frac{p(\mathbf{y})}{\prod_{i=1}^n p(i)} d\mathbf{y} \quad (4)$$

(4)

$$I(1, \dots, n) = -\frac{1}{2} \log \left(\frac{|\mathbf{C}_y|}{\prod_{i=1}^n C_{ii}} \right) = -\frac{1}{2} \sum_{i=1}^n \log(\lambda_i) \quad (5)$$

$$C_{ij} = \mathbb{E}[(y_i - \mathbb{E}[y_i])(y_j^* - \mathbb{E}[y_j^*])], \quad \mathbf{C}_y \mathbf{u} = \lambda \mathbf{\Lambda} \mathbf{u}, \\ \mathbf{C}_y = \text{cov}[\mathbf{y}], \quad \mathbf{\Lambda} = \{\mathbf{C}_{11}, \mathbf{C}_{22}, \dots, \mathbf{C}_{nn}\}.$$

$$I(1, \dots, n) = \sum_{i=1}^n H(i) - H(\mathbf{y}),$$

$$H(\mathbf{y}) = \sum_{i=1}^n H(y_i) - \log |\mathbf{W}|, \quad H(\mathbf{y}) = H(\mathbf{x}) + \log |\mathbf{W}| \quad (4)$$

$$J(\mathbf{W}) = \sum_{i=1}^n H(i) - \log |\mathbf{W}| - H(\mathbf{x}), \quad (6)$$

$$H(\mathbf{x}) = \sum_{i=1}^n H(x_i) - \log |\mathbf{W}| \quad (6)$$

\mathbf{W}^*

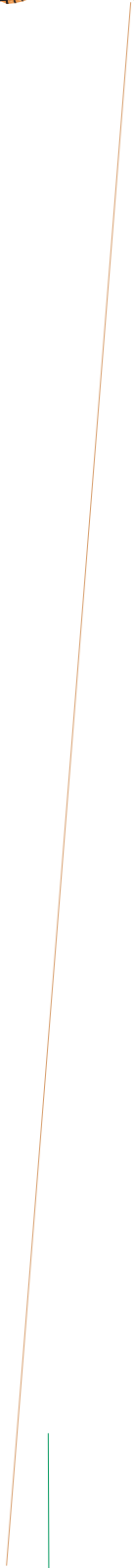
$$\nabla_{\mathbf{W}^*} J(\mathbf{W}) = (\mathbb{E}_{\mathbf{y}}[\psi(\mathbf{y})\mathbf{y}^H] - \mathbf{I})\mathbf{W}^{-H}, \quad (7)$$

$$\mathbf{W}^{-H} = [\psi(1), \dots, \psi(n)]^T, \quad \mathbf{W}^{-1},$$

$$\psi(i) = \frac{d \log p(i)}{d y_i^*} = -\frac{\frac{\partial p(y_i)}{\partial y_i} + \frac{\partial p(y_i)}{\partial y_i}}{p(i)} \\ = \frac{\partial \log p(i)}{\partial y_i} + \frac{\partial \log p(i)}{\partial y_i^*}, \quad (8)$$

$\psi(\cdot)$ complex score function. H

$$(i) \equiv (i, i) \quad (8)$$



$Q(\mathbf{z}) > 0$ for $\mathbf{z} \in \mathbb{R}^2$; $Q(\mathbf{z}) = 0$ if and only if $\mathbf{z} = \mathbf{0}$. \square

Corollary 1: $Q(\mathbf{z}_1 + \mathbf{z}_2) = Q(\mathbf{z}_1) + Q(\mathbf{z}_2)$ (19)

$$Q(\mathbf{z}_1 + \mathbf{z}_2) = Q(\mathbf{z}_1) + Q(\mathbf{z}_2) \quad (19)$$

$$Q(\mathbf{z}_1) = Q(\mathbf{z}_2) = 0.$$

Proof:

$$Q^2(\mathbf{z}_1 + \mathbf{z}_2) = Q^2(\mathbf{z}_1) + Q^2(\mathbf{z}_2) + 2Q(\mathbf{z}_1)Q(\mathbf{z}_2) \geq Q^2(\mathbf{z}_1) + Q^2(\mathbf{z}_2),$$

$$Q(\mathbf{z}_1) = Q(\mathbf{z}_2) = 0.$$

$Q(\mathbf{z}_1) = Q(\mathbf{z}_2) = 0$ is sufficient. \square

$\mathbf{z}_1 = |z_1|e^{j\theta_1}$, $\mathbf{z}_2 = |z_2|e^{j\theta_2}$ ($\theta_1, \theta_2 \in \mathbb{R}$)

$$|\mathbf{z}_1 + \mathbf{z}_2| = \sqrt{|z_1|^2 + |z_2|^2 + 2|z_1||z_2|\cos(\theta_1 - \theta_2)}e^{j\theta},$$

$$\theta = \arctan \frac{|z_1|\sin\theta_1 + |z_2|\sin\theta_2}{|z_1|\cos\theta_1 + |z_2|\cos\theta_2}.$$

$$|\mathbf{z}_1 + \mathbf{z}_2| \leq |z_1| + |z_2|.$$

Theorem 3: $Q(\mathbf{z}) = Q(|z|) = \sqrt{Q^2(|z|\cos\theta) + Q^2(|z|\sin\theta)}$

$$Q^2(|z|\cos\theta) \geq Q^2(|z_1|\cos\theta_1) + Q^2(|z_2|\cos\theta_2)$$

$$Q^2(|z|\sin\theta) \geq Q^2(|z_1|\sin\theta_1) + Q^2(|z_2|\sin\theta_2).$$

Proof:

$$Q^2(|z|\cos\theta) \geq Q^2(|z_1|\cos\theta_1) + Q^2(|z_2|\cos\theta_2)$$

$$Q^2(|z|\sin\theta) \geq Q^2(|z_1|\sin\theta_1) + Q^2(|z_2|\sin\theta_2).$$

$$Q^2(|z|\cos\theta) \geq Q^2(|z_1|\cos\theta_1) + Q^2(|z_2|\cos\theta_2)$$

$$Q^2(|z|\sin\theta) \geq Q^2(|z_1|\sin\theta_1) + Q^2(|z_2|\sin\theta_2).$$

$$\begin{aligned} \bar{Q}^2(\mathbf{z}) &= Q^2(|z|) = Q^2(|z|\cos\theta) + Q^2(|z|\sin\theta) \\ &\geq Q^2(|z_1|) + Q^2(|z_2|) = \bar{Q}^2(\mathbf{z}_1) + \bar{Q}^2(\mathbf{z}_2), \end{aligned} \quad (20)$$

Lemma 1: $Q(\mathbf{z}) = Q(|z|)$ ($\mathbf{z} \in \mathbb{C}$)

$\mathbf{z} = [z_1, z_2]^T$, $\mathbf{z} \in \mathbb{R}^2$; $Q(\alpha\mathbf{z}) = |\alpha|Q(\mathbf{z})$ ($\alpha \in \mathbb{R}$); $Q(\alpha\mathbf{z}) = |\alpha|Q(\mathbf{z})$ ($\alpha \in \mathbb{C}$).

Proof:

$Q(\alpha\mathbf{z}) = |\alpha|Q(\mathbf{z})$ ($\alpha \in \mathbb{R}$). \square

Theorem 4:

$Q(\mathbf{z}) \geq 2$ if and only if $Q(\mathbf{z}_1) \geq 2$ and $Q(\mathbf{z}_2) \geq 2$.

Proof:

$Q(\mathbf{z}) = \sqrt{Q^2(|z_1|\cos\theta) + Q^2(|z_2|\sin\theta)}$.

Remark:

$\mathbf{y} = \mathbf{E}[\mathbf{x}\mathbf{x}^H] = \mathbf{U}\Sigma\mathbf{U}^H$, $\mathbf{z} = \mathbf{D}^{-1/2}\mathbf{U}^H\mathbf{x}$, $\tilde{\mathbf{A}} = \mathbf{D}^{-1/2}\mathbf{U}^H\mathbf{A}$, $\tilde{\mathbf{A}}\mathbf{E}[\mathbf{z}\mathbf{z}^H]\tilde{\mathbf{A}}^H = \mathbf{I}$.

B. Examples of Contrast Functions

1) **Range Function:**

$$R(\cdot) = d, \quad \{ (i) \geq 0 \mid |i - 0| \leq d \} \quad (21)$$

$d \in \mathbb{R}^+$

Theorem 5:

$$R(\alpha + \beta) = R(\alpha) + R(\beta). \quad (22)$$

Proof:

$$\{ \alpha + \beta \} \subseteq \{ \alpha \} \cup \{ \beta \} \implies |\alpha + \beta| \leq |\alpha| + |\beta| \implies \square$$

Corollary 2:

$$R(\alpha + \beta) = |\alpha| R(\beta) + |\beta| R(\alpha), \quad \alpha, \beta \in \mathbb{C}, \quad (23)$$

$$Q(i) = R(i), \quad (17)$$

$$J(\mathbf{W}) = \sum_{i=1}^n \log \left(\sum_{j=1}^n |ij| R(j) \right) - \log |\mathbf{W}|. \quad (24)$$

Definition 6:

$$\int_0^d (|i|) d|i| \approx 1 \quad (0 < d < +\infty).$$

2) Shannon Entropy Function: A -k

$$H(i).$$

$$\bar{Q}(i) = Q(|i|) = \exp(H(|i|)), \quad (25)$$

$$\text{entropy power inequality, } k \quad (25)$$

3

$$(25) \quad (17) \quad (6)$$

$$\{ i \}, \quad \{ |i| \}$$

16, 17.

3) Rényi Entropy Function:

$(0 < \in \mathbb{R})$:

$$H_k(i) = \frac{1}{1-k} \log \left(\int (i)^k d i \right). \quad (26)$$

$\rightarrow 1$ k, = 2,

$$H_2(i) = -\log \left(\int (i)^2 d i \right).$$

Jensen inequality,

$$H_2(i) \leq H(i).$$

$$H_k(i) \geq H_r(i) >$$

Lemma 2: $\{ : (i) = g(|i|) \}$, $\exp(H_2(i))$,

$$\text{Proof: } H_2(i) = H_2(i + \alpha)$$

$$\begin{aligned} H_2(\alpha) &= -\log \left(\int g(|\alpha|)^2 d|\alpha| \right) \\ &= -\log \left(\int g(|i|)^2 d|i| \right) + \log |\alpha|. \end{aligned}$$

$$\exp(H_2(\alpha)) = |\alpha| \exp(H_2(i)).$$

$$\exp(H_2(i)) \quad \square$$

$$\exp(H_2(|i|)) = \frac{1}{\int p(|y_i|)^2 dy_i}, \quad \bar{Q}(i) = Q(|i|) =$$

$$26,27; \quad \bar{Q} \quad \{ |i| \}, \quad k \quad 18.$$

4) Fisher Information Function:

$$\mathbf{G} = \mathbb{E}[\psi(\cdot)^2] = \mathbb{E} \left[\left(\frac{d \log(\cdot)}{d} \right)^2 \right], \quad (27)$$

$$\psi(\cdot) = \frac{d \log p(\cdot)}{d} = \frac{p'(\cdot)}{p(\cdot)}, \quad Q(\cdot) = \mathbf{G}^{-1/2}, \quad Q$$

15.

$$\bar{Q}(\cdot) = \mathbb{E} \left[\left(\frac{p'(|i|)}{p(|i|)} \right)^2 \right]^{-1/2}$$

$$\psi_k(\cdot) =$$

$$Q(\omega) = (\Sigma)^{1/2} \equiv (\text{cov}[\tilde{\mathbf{z}}])^{1/2},$$

$$\begin{aligned} Q(\omega) &= (\text{cov}[\tilde{\mathbf{z}}'])^{1/2} = (\text{cov}[\Omega\tilde{\mathbf{z}}])^{1/2} \\ &= (\Omega\text{cov}[\tilde{\mathbf{z}}])^{1/2} = (\Omega)^{1/2} (\text{cov}[\tilde{\mathbf{z}}])^{1/2} \\ &= (\omega^2 + \omega^2)^{1/2} (\Sigma)^{1/2} = |\omega| (\Sigma)^{1/2}. \end{aligned}$$

$$\begin{aligned} Q(\omega) &= Q(\omega) \dots \dots \dots Q(\omega) \\ Q(\omega) &= \dots \dots \dots Q(\omega) \\ \dots \dots \dots Q(\omega) \dots \dots \dots \end{aligned}$$

... .. A

- A, (... ..) (... ..) 22 - 24, A.
- A 25 (... .. W) × $\mathcal{U}(\cdot)$, 2
- A A (... ..)

$$\mathcal{Y}^j \dots \mathcal{W}^A$$

... .. A 7, 13, 14, 5, 12. A 6 - 14. H A k k ;

A

... .. (... ..) $\mathcal{L}(\dots)$, $\mathcal{N}(\cdot) \left(1 + \frac{\kappa^3}{2} \mathcal{H}_3(\cdot) + \frac{\kappa^4}{2} \mathcal{H}_4(\cdot) \right)$