XXXVII SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS - SBrT2019, 29/09/2019–02/10/2019, PETRÓPOLIS, RJ

Sobre a matriz de quadricovariância: definição, propriedades e aplicação

Flávio R. M. Pavan e Maria D. Miranda

***Resumo*— Em problemas multivariados, estatísticas de ordem dois são naturalmente representadas por uma matriz de covariância. Decomposições dessa matriz são úteis, por exemplo, nos contextos de estimação espectral e análise de componentes principais. Representações matriciais de estatísticas de ordem superior, por outro lado, não possuem uma definição direta. A matriz de quadricovariância, proposta inicialmente no contexto de análise de componentes independentes, é uma possível representação de estatísticas de ordem quatro. Embora possua aplicações relevantes, seu entendimento não é tão evidente. Neste artigo, revisita-se a definição de matriz de quadricovariância para variáveis aleatórias reais e abordam-se suas propriedades e uma aplicação. Vislumbra-se que essa abordagem possibilite melhor compreensão e aproveitamento de técnicas baseadas em estatísticas de ordem superior.**

***Palavras-Chave*— Matriz de quadricovariância, estatísticas de ordem superior, análise de componentes independentes.**

***Abstract*— In multivariate problems, second-order statistics are naturally represented by a covariance matrix. Factorizations of this matrix are useful, for instance, in the contexts of spectral estimation and principal component analysis. Matrix representations of higher-order statistics, however, are not straightforwardly defined. The quadricovariance matrix, initially proposed in the context of independent component analysis, is a possible representation of fourth-order statistics. In spite of relevant applications, its understanding is rather intricate. In this paper, the quadricovariance matrix definition for real random variables is revisited, and its properties and an application are presented. This approach is expected to provide a better comprehension and use of techniques based on higher-order statistics.**

***Keywords*— Quadricovariance matrix, higher-order statistics, independent component analysis.**

I. INTRODUÇÃO

Estatísticas de ordem dois são amplamente utilizadas na descrição e solução de problemas em que se deseja garantir a não correlação de variáveis aleatórias. Alguns exemplos são regressão linear, filtragem de Wiener, filtragem adaptativa e estimação espectral [1]. Estatísticas de ordem superior a dois, por outro lado, são relevantes na solução de problemas em que é necessário impor a não correlação de variáveis aleatórias em ordens superiores e, eventualmente, sua independência estatística. Em geral, problemas de estimação cega como desconvolução multiusuário e separação de fontes (BSS, do inglês *blind source separation*) podem ser abordados com estatísticas de ordem superior [2], [3]. Essas estatísticas também têm sido extensivamente aplicadas, implícita ou explicitamente, no contexto de aprendizado de máquina para extração de características e pré-processamento de dados [4], [5], [6]. Por outro lado, o entendimento teórico das estatísticas de ordem superior não é tão evidente quanto para o caso de ordem dois.

Flávio R. M. Pavan, e-mail: frmp@lcs.poli.usp.br; Maria D. Miranda, e-mail: maria@lcs.poli.usp.br. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Distribuições estatísticas podem ser descritas por parâmetros conhecidos como momentos e cumulantes [7], [8]. Em problemas multivariados, cumulantes de ordem dois são comumente representados por meio de uma matriz de covariância. É amplamente conhecido que a decomposição dessa matriz em autovalores e autovetores permite encontrar uma transformação linear que descorrelaciona variáveis aleatórias [7]. Dependendo do modelo estatístico considerado para os sinais em cada problema, diferentes aplicações podem decorrer da diagonalização da matriz de covariância. Alguns exemplos são o método de Pisarenko para identificação de senoides em ruído, a transformada de Karhunen-Loève para compressão de sinais e a análise de componentes principais (PCA, do inglês *principal component analysis*) para redução de dimensionalidade [9], [10].

Se a diagonalização da matriz de covariância permite descorrelacionar variáveis aleatórias, então é razoável imaginar que a aplicação do mesmo princípio a estatísticas de ordem superior possa tornar as variáveis aleatórias mais próximas da independência. No contexto de análise de componentes independentes (ICA, do inglês *independent component analysis*) aplicada ao problema de BSS, uma das primeiras soluções baseou-se na matriz de quadricovariância [11], [12], [13]. Trata-se de uma representação matricial consistente para cumulantes de ordem quatro, cuja notação usual é tensorial [13].

A representação de estatísticas de ordem superior por meio de matrizes é relevante, em especial na solução de problemas de estimação cega. Porém, sua definição e entendimento não são evidentes. Neste artigo, apresenta-se a definição de matriz de quadricovariância para variáveis aleatórias reais e abordam-se suas propriedades e uma aplicação. Na Seção II, recordam-se brevemente a definição de matriz de covariância, suas principais propriedades e aplicações. Na Seção III, apresenta-se a definição usual de matriz de quadricovariância e, na Seção IV, apresentam-se algumas de suas propriedades. Na Seção V, aborda-se a aplicação da matriz de quadricovariância a um problema de BSS. Isso motiva, na Seção VI, uma discussão acerca da completude dessa representação matricial. Por fim, as conclusões são apresentadas na Seção VII.

NOTAÇÃO E CONCEITOS PRELIMINARES

Ao longo do texto, grandezas aleatórias são representadas sublinhadas e grandezas vetoriais são representadas em negrito. O transposto de um vetor ou matriz é denotado por $(\cdot)^{\mathsf{T}}$ e o elemento da linha i e coluna j de uma matriz é denotado por $[\cdot]_{i,j}$. Considera-se que todas as variáveis aleatórias abordadas na sequência são reais e todos os seus momentos e cumulantes convergem. A esperança matemática de uma variável aleatória $\underline{a}$ é denotada por $\mathrm{E}[\underline{a}]$, e seu cumulante por $\mathrm{cum}[\underline{a}]$. Definindo-se o vetor aleatório $\underline{\boldsymbol{a}} = [\,\underline{a}_1 \;\; \underline{a}_2 \;\; \cdots \;\; \underline{a}_N\,]^\mathsf{T}$, representam-se momentos conjuntos e cumulantes conjuntos dos elementos de $\underline{\boldsymbol{a}}$ como

$$\begin{aligned} \mathcal{E}_{\underline{\boldsymbol{a}}}^{i,j,\ldots,\ell} &= \mathrm{E}[\underline{a}_i \underline{a}_j \cdots \underline{a}_\ell], \\ \mathcal{C}_{\underline{\boldsymbol{a}}}^{i,j,\ldots,\ell} &= \mathrm{cum}[\underline{a}_i, \underline{a}_j, \ldots, \underline{a}_\ell], \end{aligned}$$

com $i, j, \ldots, \ell = 1, 2, \ldots, N$.

Para cumulantes de um conjunto de variáveis aleatórias, são válidas as propriedades de multilinearidade, invariância à permutação de índices e à adição de constantes [8]. Para partições independentes de variáveis aleatórias, o cumulante conjunto é nulo [8]. Além disso, para variáveis aleatórias conjuntamente gaussianas, todos os cumulantes de ordem maior do que dois são iguais a zero [14]. Supondo as variáveis aleatórias de $\underline{\boldsymbol{a}}$ com médias nulas, os cumulantes conjuntos de ordem dois e quatro dessas variáveis podem ser expressos em função de momentos conjuntos como

$$\mathcal{C}_{\underline{\boldsymbol{a}}}^{i,j} = \mathcal{E}_{\underline{\boldsymbol{a}}}^{i,j}, \tag{1}$$

$$\mathcal{C}_{\underline{\boldsymbol{a}}}^{i,j,k,\ell} = \mathcal{E}_{\underline{\boldsymbol{a}}}^{i,j,k,\ell} - \mathcal{E}_{\underline{\boldsymbol{a}}}^{i,j}\mathcal{E}_{\underline{\boldsymbol{a}}}^{k,\ell} - \mathcal{E}_{\underline{\boldsymbol{a}}}^{i,k}\mathcal{E}_{\underline{\boldsymbol{a}}}^{j,\ell} - \mathcal{E}_{\underline{\boldsymbol{a}}}^{i,\ell}\mathcal{E}_{\underline{\boldsymbol{a}}}^{j,k}, \tag{2}$$

para $i, j, k, \ell = 1, 2, \ldots, N$ [8].

II. Matriz de covariância

Os cumulantes conjuntos de ordem dois, usualmente denominados covariâncias, possuem uma notação matricial que provém diretamente da sua definição.

Definição 1: A matriz de covariância do vetor aleatório $\underline{\boldsymbol{a}}$, denotada por $\boldsymbol{C}_{\underline{\boldsymbol{a}}} \in \mathbb{R}^{N\times N}$, é definida elemento a elemento como $[\boldsymbol{C}_{\underline{\boldsymbol{a}}}]_{i,j} = \mathcal{C}_{\underline{\boldsymbol{a}}}^{i,j}$.

Devido à invariância à permutação de índices dos cumulantes, nota-se que $\boldsymbol{C}_{\underline{\boldsymbol{a}}}$ é uma matriz simétrica, isto é, $\boldsymbol{C}_{\underline{\boldsymbol{a}}} = \boldsymbol{C}_{\underline{\boldsymbol{a}}}^\mathsf{T}$. Além disso, a partir de (1), constata-se que essa matriz é positiva semidefinida [7]. De acordo com a definição apresentada e utilizando (1), pode-se exprimir a matriz de covariância de $\underline{\boldsymbol{a}}$ genericamente como

$$\boldsymbol{C}_{\underline{\boldsymbol{a}}} = \mathrm{E}[(\underline{\boldsymbol{a}} - \mathrm{E}[\underline{\boldsymbol{a}}])\,(\underline{\boldsymbol{a}} - \mathrm{E}[\underline{\boldsymbol{a}}])^\mathsf{T}]. \tag{3}$$

Propriedade 1: Seja $\underline{\boldsymbol{a}}$ um vetor aleatório de N variáveis, com matriz de covariância $\boldsymbol{C}_{\underline{\boldsymbol{a}}} \in \mathbb{R}^{N\times N}$, e seja o vetor aleatório $\underline{\boldsymbol{b}} = \boldsymbol{G}\underline{\boldsymbol{a}}$ com $\boldsymbol{G} \in \mathbb{R}^{N\times N}$. A matriz de covariância de $\underline{\boldsymbol{b}}$ pode ser expressa como

$$\boldsymbol{C}_{\underline{\boldsymbol{b}}} = \boldsymbol{G}\boldsymbol{C}_{\underline{\boldsymbol{a}}}\boldsymbol{G}^\mathsf{T}. \tag{4}$$

Nota-se que $\boldsymbol{C}_{\underline{\boldsymbol{b}}}$ é positiva semidefinida, podendo ser expressa como $\boldsymbol{C}_{\underline{\boldsymbol{b}}} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^\mathsf{T}$, com $\boldsymbol{P}^{-1} = \boldsymbol{P}^\mathsf{T}$ e $\boldsymbol{\Lambda}$ matriz diagonal de elementos não negativos [15]. Comparando essa fatoração com (4), nota-se que existe $\underline{\boldsymbol{a}}$ com $\boldsymbol{C}_{\underline{\boldsymbol{a}}} = \boldsymbol{\Lambda}$ tal que $\underline{\boldsymbol{b}} = \boldsymbol{P}\underline{\boldsymbol{a}}$. Portanto, a diagonalização de $\boldsymbol{C}_{\underline{\boldsymbol{b}}}$ permite expressar os elementos de $\underline{\boldsymbol{b}}$ como combinações lineares de variáveis não correlacionadas. Essencialmente, a diagonalização da matriz de covariância pode ser compreendida como uma decomposição de variáveis aleatórias em uma base ortogonal [7].

III. Matriz de quadricovariância

Ao contrário dos cumulantes de ordem dois, a obtenção de uma representação matricial consistente para cumulantes de ordem superior não é tão direta [13], [16]. Por simplicidade, aborda-se a seguir o caso de cumulantes de ordem quatro de um vetor aleatório $\underline{\boldsymbol{a}}$. Por possuírem quatro índices livres, esses cumulantes podem ser representados por meio de uma hipermatriz $\mathfrak{C}_{\underline{\boldsymbol{a}}} \in \mathbb{R}^{N\times N\times N\times N}$ cujos elementos são dados por $[\mathfrak{C}_{\underline{\boldsymbol{a}}}]_{i,j,k,\ell} = \mathcal{C}_{\underline{\boldsymbol{a}}}^{i,j,k,\ell}$. Embora a notação usual de cumulantes de ordem superior seja tensorial [8], [17], [18], pode-se abordar o caso de ordem quatro sem lançar mão de ferramentas matemáticas mais sofisticadas [13]. Para esses cumulantes, é possível adotar representações matriciais que preservem suas propriedades (como linearidade, simetria, etc.) e sejam convenientes na resolução de problemas práticos [16]. A seguir, apresenta-se uma das possíveis representações matriciais de cumulantes de ordem quatro [12].

Definição 2: A matriz de quadricovariância de $\underline{\boldsymbol{a}}$, denotada por $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}) \in \mathbb{R}^{N\times N}$, é definida como

$$[\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V})]_{i,j} = \sum_{k=1}^{N}\sum_{\ell=1}^{N} \mathcal{C}_{\underline{\boldsymbol{a}}}^{i,j,k,\ell}\,[\boldsymbol{V}]_{k,\ell}\,, \tag{5}$$

em que $\boldsymbol{V} \in \mathbb{R}^{N\times N}$ é uma matriz de pesos.

A matriz de quadricovariância $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V})$ consiste em uma combinação linear de fatias da hipermatriz de cumulantes $\mathfrak{C}_{\underline{\boldsymbol{a}}}$. Cada fatia corresponde a uma matriz de cumulantes de ordem quatro $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}^{k,\ell} \in \mathbb{R}^{N\times N}$ tal que $\left[\boldsymbol{Q}_{\underline{\boldsymbol{a}}}^{k,\ell}\right]_{i,j} = \mathcal{C}_{\underline{\boldsymbol{a}}}^{i,j,k,\ell}$ para $k, \ell = 1, 2, \ldots, N$. A matriz de quadricovariância pode ser expressa em termos das fatias conforme

$$\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}) = \sum_{k=1}^{N}\sum_{\ell=1}^{N} \boldsymbol{Q}_{\underline{\boldsymbol{a}}}^{k,\ell}\,[\boldsymbol{V}]_{k,\ell}\,. \tag{6}$$

Exemplo 1: Para um vetor aleatório $\underline{\boldsymbol{a}} = [\,\underline{a}_1 \;\; \underline{a}_2\,]^\mathsf{T}$ de $N = 2$ elementos, a hipermatriz de cumulantes de ordem quatro $\mathfrak{C}_{\underline{\boldsymbol{a}}} \in \mathbb{R}^{2\times 2\times 2\times 2}$ possui $2^4 = 16$ elementos. Fatiando-se a hipermatriz, obtêm-se as seguintes matrizes

$$\boldsymbol{Q}_{\underline{\boldsymbol{a}}}^{1,1} = \begin{bmatrix} \mathcal{C}_{\underline{\boldsymbol{a}}}^{1,1,1,1} & \mathcal{C}_{\underline{\boldsymbol{a}}}^{1,2,1,1} \\ \mathcal{C}_{\underline{\boldsymbol{a}}}^{2,1,1,1} & \mathcal{C}_{\underline{\boldsymbol{a}}}^{2,2,1,1} \end{bmatrix}, \quad \boldsymbol{Q}_{\underline{\boldsymbol{a}}}^{1,2} = \begin{bmatrix} \mathcal{C}_{\underline{\boldsymbol{a}}}^{1,1,1,2} & \mathcal{C}_{\underline{\boldsymbol{a}}}^{1,2,1,2} \\ \mathcal{C}_{\underline{\boldsymbol{a}}}^{2,1,1,2} & \mathcal{C}_{\underline{\boldsymbol{a}}}^{2,2,1,2} \end{bmatrix}$$

$$\boldsymbol{Q}_{\underline{\boldsymbol{a}}}^{2,1} = \begin{bmatrix} \mathcal{C}_{\underline{\boldsymbol{a}}}^{1,1,2,1} & \mathcal{C}_{\underline{\boldsymbol{a}}}^{1,2,2,1} \\ \mathcal{C}_{\underline{\boldsymbol{a}}}^{2,1,2,1} & \mathcal{C}_{\underline{\boldsymbol{a}}}^{2,2,2,1} \end{bmatrix}, \quad \boldsymbol{Q}_{\underline{\boldsymbol{a}}}^{2,2} = \begin{bmatrix} \mathcal{C}_{\underline{\boldsymbol{a}}}^{1,1,2,2} & \mathcal{C}_{\underline{\boldsymbol{a}}}^{1,2,2,2} \\ \mathcal{C}_{\underline{\boldsymbol{a}}}^{2,1,2,2} & \mathcal{C}_{\underline{\boldsymbol{a}}}^{2,2,2,2} \end{bmatrix}.$$

Nota-se que todos os cumulantes de ordem quatro de $\underline{\boldsymbol{a}}$ são contemplados pelos elementos dessas matrizes. A partir das matrizes de cumulantes obtidas, a matriz $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}) \in \mathbb{R}^{2\times 2}$ resulta de (6). Escolhendo matrizes de pesos $\boldsymbol{V} \in \mathbb{R}^{2\times 2}$ tais que apenas um elemento seja igual a um e os demais sejam nulos, pode-se recuperar cada uma das matrizes de cumulantes previamente obtidas.

Diferentemente da matriz de covariância, a matriz $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V})$ não é única pois sua obtenção depende da escolha de $\boldsymbol{V}$. Essa dependência, aparentemente estranha, decorre da representação de cumulantes de ordem quatro de N variáveis aleatórias como matrizes $N \times N$. A notação introduzida em (5) sugere que as matrizes $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V})$ são elementos do espaço imagem do operador de quadricovariância

$$\boldsymbol{Q}_{\underline{\boldsymbol{a}}} : \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N}$$

$$\boldsymbol{V} \mapsto \boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}).$$

Dessa forma, a matriz de pesos $\boldsymbol{V}$ é parte integrante da representação de cumulantes de ordem quatro por meio da matriz de quadricovariância. Dependendo da aplicação, a escolha de $\boldsymbol{V}$ pode ser crucial na aplicabilidade de $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V})$ à resolução de um problema. O papel da matriz de pesos na representação da Definição 2 é abordado em mais detalhes a partir da Seção V.

IV. Propriedades referentes à matriz de quadricovariância

Propriedade 2: O operador $\boldsymbol{Q}_{\underline{\boldsymbol{a}}} : \mathbb{R}^{N\times N} \to \mathbb{R}^{N\times N}$ é linear, isto é, $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}_1 + \boldsymbol{V}_2) = \boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}_1) + \boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}_2)\ \forall\, \boldsymbol{V}_1, \boldsymbol{V}_2 \in \mathbb{R}^{N\times N}$ e $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(c\,\boldsymbol{V}) = c\,\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V})\ \forall\, c \in \mathbb{R}$.

Propriedade 3: $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}) = \boldsymbol{Q}_{\underline{\boldsymbol{a}}}^{\mathsf{T}}(\boldsymbol{V})\ \forall\, \boldsymbol{V} \in \mathbb{R}^{N\times N}$.

Propriedade 4: $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}) = \boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}^{\mathsf{T}})\ \forall\, \boldsymbol{V} \in \mathbb{R}^{N\times N}$.

Propriedade 5: Se $\underline{\boldsymbol{a}}$ possuir média nula, então as matrizes de quadricovariância podem ser expressas em função de momentos como

$$\begin{aligned}\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}) = \mathrm{E}\left[\left(\underline{\boldsymbol{a}}^{\mathsf{T}}\boldsymbol{V}\underline{\boldsymbol{a}}\right)\left(\underline{\boldsymbol{a}}\underline{\boldsymbol{a}}^{\mathsf{T}}\right)\right] - \mathrm{tr}(\boldsymbol{C}_{\underline{\boldsymbol{a}}}\boldsymbol{V})\,\boldsymbol{C}_{\underline{\boldsymbol{a}}} \\ - \boldsymbol{C}_{\underline{\boldsymbol{a}}}\boldsymbol{V}\boldsymbol{C}_{\underline{\boldsymbol{a}}} - \boldsymbol{C}_{\underline{\boldsymbol{a}}}\boldsymbol{V}^{\mathsf{T}}\boldsymbol{C}_{\underline{\boldsymbol{a}}},\end{aligned} \tag{7}$$

para toda $\boldsymbol{V} \in \mathbb{R}^{N\times N}$, em que $\mathrm{tr}(\cdot)$ denota o traço de $(\cdot)$ [12].

Comparando-se (2) e (7), nota-se certa semelhança. Particularmente, (2) poderia ser interpretada como uma expressão para cumulantes escalares $\mathcal{C}_{\underline{\boldsymbol{a}}}^{i,j,k,\ell}$ em função de momentos também escalares, e (7) como uma expressão para matrizes de quadricovariância $\boldsymbol{Q}_{\underline{\boldsymbol{a}}}(\boldsymbol{V})$ em função de matrizes de momentos.

Propriedade 6: Seja $\underline{\boldsymbol{a}}$ um vetor aleatório de N elementos mutuamente independentes e seja $\underline{\boldsymbol{b}} = \boldsymbol{G}\underline{\boldsymbol{a}}$ com $\boldsymbol{G} = [\,\boldsymbol{g}_1\ \ \boldsymbol{g}_2\ \ \cdots\ \ \boldsymbol{g}_N\,] \in \mathbb{R}^{N\times N}$. A matriz de quadricovariância de $\underline{\boldsymbol{b}}$ pode ser expressa como

$$\boldsymbol{Q}_{\underline{\boldsymbol{b}}}(\boldsymbol{V}) = \boldsymbol{G}\boldsymbol{\Lambda}_{\underline{\boldsymbol{a}}}(\boldsymbol{V})\,\boldsymbol{G}^{\mathsf{T}} \tag{8}$$

para toda $\boldsymbol{V} \in \mathbb{R}^{N\times N}$, sendo

$$\boldsymbol{\Lambda}_{\underline{\boldsymbol{a}}}(\boldsymbol{V}) = \mathrm{diag}\left(\mathcal{K}_{\underline{a}_1}\,\boldsymbol{g}_1^{\mathsf{T}}\boldsymbol{V}\boldsymbol{g}_1, \mathcal{K}_{\underline{a}_2}\,\boldsymbol{g}_2^{\mathsf{T}}\boldsymbol{V}\boldsymbol{g}_2, \ldots, \mathcal{K}_{\underline{a}_N}\,\boldsymbol{g}_N^{\mathsf{T}}\boldsymbol{V}\boldsymbol{g}_N\right)$$

em que $\mathrm{diag}(\cdot)$ denota uma matriz diagonal com elementos $(\cdot)$ e $\mathcal{K}_{\underline{a}_i} = \mathcal{C}_{\underline{\boldsymbol{a}}}^{i,i,i,i}$ denota o autocumulante de ordem quatro da i-ésima variável aleatória de $\underline{\boldsymbol{a}}$ [12], [13].

Demonstração: Inicialmente, aplica-se a Definição 2 a $\boldsymbol{Q}_{\underline{\boldsymbol{b}}}(\boldsymbol{V})$ para expressar essa matriz em termos dos cumulantes $\mathcal{C}_{\underline{\boldsymbol{b}}}^{i,j,k,\ell}$. Utilizando as propriedades de multilinearidade e partições independentes de cumulantes, pode-se expressar $\mathcal{C}_{\underline{\boldsymbol{b}}}^{i,j,k,\ell}$ em função de $\mathcal{C}_{\underline{\boldsymbol{a}}}^{i,j,k,\ell}$ na expressão de $\boldsymbol{Q}_{\underline{\boldsymbol{b}}}(\boldsymbol{V})$. Em seguida, manipulações algébricas permitem chegar a

$$\boldsymbol{Q}_{\underline{\boldsymbol{b}}}(\boldsymbol{V}) = \sum_{i=1}^{N} \mathcal{K}_{\underline{a}_i}(\boldsymbol{g}_i^{\mathsf{T}}\boldsymbol{V}\boldsymbol{g}_i)(\boldsymbol{g}_i\boldsymbol{g}_i^{\mathsf{T}}),$$

que, por fim, pode ser representada na forma compacta (8). ■

Diferentemente da Propriedade 1 para a matriz de covariância, a Propriedade 6 só pode ser garantida para uma transformação linear sobre um vetor aleatório de elementos independentes. Mesmo assim, é interessante comparar as transformações para a matriz de covariância [Eq. (4)] e quadricovariância [Eq. (8)] nesse caso particular. Enquanto a matriz $\boldsymbol{C}_{\underline{\boldsymbol{a}}}$ será diagonal contendo as variâncias dos elementos de $\underline{\boldsymbol{a}}$, a matriz $\boldsymbol{\Lambda}_{\underline{\boldsymbol{a}}}(\boldsymbol{V})$ dependerá, além das estatísticas de ordem quatro de cada elemento de $\underline{\boldsymbol{a}}$, também das colunas de $\boldsymbol{G}$ e da matriz de pesos $\boldsymbol{V}$.

Por fim, em virtude da Propriedade 3, a matriz de quadricovariância é ortogonalmente diagonalizável e possui autovalores reais, assim como a matriz de covariância.

V. Aplicação a um caso de separação cega

Dependendo do problema ao qual a matriz de quadricovariância é aplicada, pode ser possível relacionar sua decomposição em autovalores e autovetores com a fatoração (8). Em certos problemas de BSS, por exemplo, essa relação permite expressar os elementos de um vetor aleatório como combinações lineares de variáveis aleatórias independentes. Nesse contexto, a diagonalização da matriz de quadricovariância consiste em uma maneira de realizar a análise de componentes independentes. Trata-se de uma extensão da diagonalização da matriz de covariância que, conforme abordado na Seção II, permite decompor variáveis aleatórias em componentes não correlacionadas.

Seja um problema particular de BSS em que são observadas N variáveis aleatórias dispostas em um vetor $\underline{\boldsymbol{x}}$ satisfazendo o modelo linear e instantâneo $\underline{\boldsymbol{x}} = \boldsymbol{U}\underline{\boldsymbol{s}}$ em que $\boldsymbol{U} = [\,\boldsymbol{u}_1\ \ \boldsymbol{u}_2\ \ \cdots\ \ \boldsymbol{u}_N\,] \in \mathbb{R}^{N\times N}$ é uma matriz de mistura ortogonal[1] e $\underline{\boldsymbol{s}}$ é um vetor de N fontes independentes não observadas [3], [12]. Deseja-se separar as fontes com base apenas em observações das misturas. Por simplicidade, supõe-se que as estatísticas das misturas são conhecidas com exatidão. Se for possível determinar a matriz de mistura $\boldsymbol{U}$, pode-se separar as fontes conforme $\underline{\boldsymbol{y}} = \boldsymbol{U}^{\mathsf{T}}\underline{\boldsymbol{x}}$.

Aplicando-se a Propriedade 6 ao modelo dado, tem-se que a matriz de quadricovariância de $\underline{\boldsymbol{x}}$ pode ser expressa como

$$\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V}) = \boldsymbol{U}\boldsymbol{\Lambda}_{\underline{\boldsymbol{s}}}(\boldsymbol{V})\,\boldsymbol{U}^{\mathsf{T}} \tag{9}$$

com

$$\boldsymbol{\Lambda}_{\underline{\boldsymbol{s}}}(\boldsymbol{V}) = \mathrm{diag}\left(\mathcal{K}_{\underline{s}_1}\,\boldsymbol{u}_1^{\mathsf{T}}\boldsymbol{V}\boldsymbol{u}_1, \mathcal{K}_{\underline{s}_2}\,\boldsymbol{u}_2^{\mathsf{T}}\boldsymbol{V}\boldsymbol{u}_2, \ldots, \mathcal{K}_{\underline{s}_N}\,\boldsymbol{u}_N^{\mathsf{T}}\boldsymbol{V}\boldsymbol{u}_N\right). \tag{10}$$

Nesse caso, como $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ é ortogonalmente diagonalizável, tem-se que (9) é uma decomposição em autovalores e autovetores válida de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ para todo $\boldsymbol{V} \in \mathbb{R}^{N\times N}$. Em outras palavras, a matriz ortogonal de mistura diagonaliza a matriz de quadricovariância.

Como as estatísticas das misturas são conhecidas, $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ pode ser calculada aplicando-se a Definição 2 ao vetor $\underline{\boldsymbol{x}}$. Se todos os autovalores de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ dispostos em $\boldsymbol{\Lambda}_{\underline{\boldsymbol{s}}}(\boldsymbol{V})$ forem distintos, a decomposição em autovalores e autovetores de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ é única a menos de permutações e trocas de sinal dos autovetores [15]. Nesse caso, diz-se que $\boldsymbol{U}$ é identificável a partir da diagonalização de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ e as fontes podem ser separadas a menos de ambiguidades de permutação e sinal [3].

[1] Casos mais gerais em que a matriz de mistura não é ortogonal podem ser reduzidos ao modelo apresentado, por exemplo, a partir de uma etapa prévia de descorrelação em ordem dois [12].

Essas ambiguidades são previstas pelo teorema de Darmois-Skitovich e decorrentes de limitações teóricas desse problema particular de separação cega [3], [19].

Observando-se (10), nota-se que se $\mathcal{K}_{\underline{s}_i} = 0$ para mais de uma fonte, então $\boldsymbol{U}$ nunca será identificável. Felizmente, exceto pelo caso gaussiano que pode ser desconsiderado por limitações teóricas [19], distribuições com autocumulante de ordem quatro nulo são pouco usuais [13]. Excluído esse caso, resta saber como garantir que $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ tenha autovalores distintos sempre, independentemente das estatísticas das fontes e das colunas de $\boldsymbol{U}$. Aparentemente, a única alternativa viável seria tentar satisfazer essa condição a partir da escolha adequada de $\boldsymbol{V}$.

Exemplo 2: A fim de ilustrar o papel da matriz de pesos $\boldsymbol{V}$, considera-se o problema de BSS apresentado no início dessa seção com $N = 2$. Na Fig. 1, o diagrama de dispersão i.(a) representa fontes $\underline{\boldsymbol{s}}$ de mesmas estatísticas, e o diagrama i.(b) representa suas misturas $\underline{\boldsymbol{x}}$. Em i.(c) e i.(d) estão os diagramas dos resultados da separação obtidos a partir da diagonalização de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ com $\boldsymbol{V} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ e $\boldsymbol{V} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$, respectivamente. Já o diagrama indicado em ii.(a) representa fontes de estatísticas distintas, e o diagrama ii.(b) representa suas misturas. Em ii.(c) e ii.(d) estão os diagramas dos resultados da separação obtidos a partir da diagonalização de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ com $\boldsymbol{V} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ e $\boldsymbol{V} = \left[\begin{smallmatrix} 1 & -0{,}4 \\ 0 & 0 \end{smallmatrix}\right]$, respectivamente. Nota-se que uma mesma $\boldsymbol{V}$ pode permitir a identificabilidade de $\boldsymbol{U}$ em ii.(c), e não permiti-la em i.(c). Além disso, observa-se em i.(d) que a identificabilidade pode ser obtida com a alteração de $\boldsymbol{V}$, ou perdida como acontece em ii.(d).

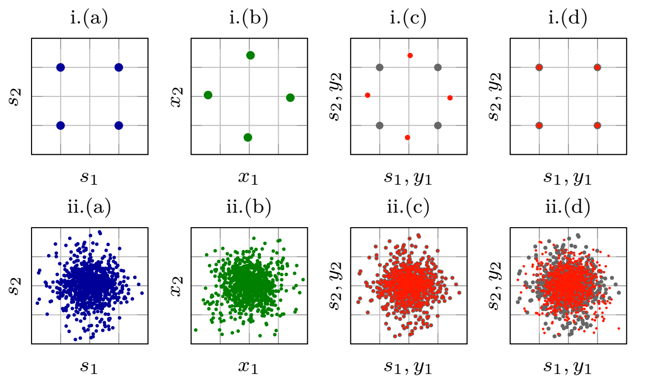

Fig. 1. Diagramas de dispersão; i. fontes discretas de mesma distribuição, ii. fontes contínuas de distribuições distintas com 10^3 realizações; (a) fontes, (b) misturas, (c) e (d) resultado da separação em vermelho, fontes em cinza.

Com o exemplo apresentado, tem-se pistas de como as estatísticas das fontes e as matrizes de pesos adotadas intervêm na obtenção da matriz ortogonal que leva à separação. Particularmente, considerando-se uma matriz de pesos identidade, *i.e.* $\boldsymbol{V} = \boldsymbol{I}_N$, a partir de (10) obtém-se $\boldsymbol{\Lambda}_{\underline{\boldsymbol{s}}}(\boldsymbol{I}_N) = \operatorname{diag}(\mathcal{K}_{\underline{s}_1}, \mathcal{K}_{\underline{s}_2}, \ldots, \mathcal{K}_{\underline{s}_N})$. Nesse caso, os autovalores de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ são distintos se, e somente se, as fontes possuírem autocumulantes de ordem quatro distintos [12]. Isso permite compreender o resultado dos diagramas i.(c) e ii.(c) da Fig. 1. Técnicas obtidas a partir dessa escolha de matriz de pesos começaram a ser estudadas no final dos anos 1980 e são denominadas *fourth-order blind identification* (FOBI) [11], [12]. Nota-se que a abordagem do FOBI funciona em um caso muito específico: quando as fontes possuem estatísticas de ordem quatro distintas. Uma solução aceitável para o problema de BSS, por outro lado, deve garantir a separação para o conjunto mais amplo possível de estatísticas das fontes e para quaisquer coeficientes do sistema misturador, observadas as condições teóricas de separabilidade do teorema de Darmois-Skitovich.

VI. Completude da representação matricial

Como consequência da Definição 2, nota-se que a descrição completa dos cumulantes de ordem quatro de $\underline{\boldsymbol{x}}$ está no operador $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ e não necessariamente em uma matriz específica $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ para um $\boldsymbol{V}$ dado. Isso motiva abordar o operador de quadricovariância em mais detalhes, a fim de verificar se é possível garantir a separação mesmo quando duas ou mais fontes possuem mesmas estatísticas [12], [13].

De modo a analisar $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ teoricamente, supõe-se inicialmente que as colunas $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_N$ da matriz de mistura $\boldsymbol{U}$ são conhecidas. Observando-se os autovalores de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ em (10), sugere-se considerar as matrizes de pesos $\boldsymbol{V}_i = \boldsymbol{u}_i \boldsymbol{u}_i^{\mathsf{T}} \in \mathbb{R}^{N \times N}$ para $i = 1, 2, \ldots, N$. Devido à ortogonalidade dos vetores $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_N$, obtém-se

$$\boldsymbol{\Lambda}_{\underline{\boldsymbol{s}}}(\boldsymbol{V}_i) = \operatorname{diag}(0, \ldots, 0, \mathcal{K}_{\underline{s}_i}, 0, \ldots, 0), \tag{11}$$

com elemento não necessariamente nulo apenas na i-ésima posição da diagonal principal. Isso ocorrerá para cada $i = 1, 2, \ldots, N$ considerado, de modo que a identificação da direção $\boldsymbol{u}_i$ será possível desde que $\mathcal{K}_{\underline{s}_i} \neq 0$. No caso em que a i-ésima fonte é a única que possui $\mathcal{K}_{\underline{s}_i} = 0$, pode-se desconsiderar a diagonalização da matriz $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V}_i)$. Nesse caso, após a determinação de todas as outras direções a partir da diagonalização das demais matrizes de quadricovariância, obtém-se $\boldsymbol{u}_i$ que complete uma base ortonormal com as demais direções. Sendo assim, a matriz $\boldsymbol{U}$ é identificável a partir da diagonalização simultânea de todas as matrizes $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V}_1), \boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V}_2), \ldots, \boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V}_N)$ se, e somente se, no máximo uma fonte possuir autocumulante de ordem quatro nulo.

Trata-se de uma condição mais abrangente do que a do FOBI, em cujo caso $\boldsymbol{U}$ não é identificável se duas ou mais fontes possuírem as mesmas estatísticas. Com base nessa escolha de matrizes de pesos e na diagonalização conjunta das respectivas matrizes de quadricovariância, propôs-se no início dos anos 1990 uma técnica amplamente utilizada e conhecida como *joint approximate diagonalization of eigenmatrices* (JADE) [12].

As matrizes de pesos $\boldsymbol{V}_i = \boldsymbol{u}_i \boldsymbol{u}_i^{\mathsf{T}}$ também são úteis para compreender melhor o operador de quadricovariância. Substituindo-se (11) em (9), tem-se

$$\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V}_i) = \boldsymbol{U} \boldsymbol{\Lambda}_{\underline{\boldsymbol{s}}}(\boldsymbol{V}_i) \boldsymbol{U}^{\mathsf{T}} = \mathcal{K}_{\underline{s}_i} \boldsymbol{u}_i \boldsymbol{u}_i^{\mathsf{T}} = \mathcal{K}_{\underline{s}_i} \boldsymbol{V}_i. \tag{12}$$

Do ponto de vista algébrico, é possível interpretar $\boldsymbol{V}_i$ como um autovetor do operador linear $\boldsymbol{Q}_{\underline{\boldsymbol{x}}} : \mathbb{R}^{N \times N} \to \mathbb{R}^{N \times N}$. Devido ao espaço matricial em que o operador atua, os autovetores de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ são denominados automatrizes [11], [12]. O autovalor ao qual a automatriz $\boldsymbol{V}_i$ está associada é, portanto, igual a $\mathcal{K}_{\underline{s}_i}$.

Além disso, expandindo-se (9), verifica-se que $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V})$ para uma matriz de pesos $\boldsymbol{V}$ qualquer pode sempre ser escrita como

$$\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V}) = \sum_{i=1}^{N} \left(\mathcal{K}_{\underline{s}_i} \boldsymbol{u}_i^{\mathsf{T}} \boldsymbol{V} \boldsymbol{u}_i\right) \boldsymbol{V}_i. \tag{13}$$

Observa-se que o espaço imagem de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ está contido no subespaço com base $\{\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_N\}$. Se $\mathcal{K}_{\underline{s}_i} \neq 0$ para $i = 1, 2, \ldots, N$, então $\{\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_N\}$ é uma base ortonormal do espaço imagem de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ sob o produto interno usual do espaço vetorial de matrizes [15]. Nesse caso, a dimensão do espaço imagem (*i.e.*, posto) de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ é igual a N. Se $\mathcal{K}_{\underline{s}_i} = 0$ apenas para a i-ésima fonte, então $\{\boldsymbol{V}_1, \boldsymbol{V}_2, \ldots, \boldsymbol{V}_N\} \setminus \{\boldsymbol{V}_i\}$ é uma base ortonormal do espaço imagem de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$. Nesse caso, o posto de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ passa a ser igual a $N - 1$. Conforme o teorema do núcleo e da imagem de álgebra linear [15], o núcleo de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ tem dimensão igual a $N^2 - \mathrm{posto}(\boldsymbol{Q}_{\underline{\boldsymbol{x}}})$.

Independentemente das estatísticas das fontes ou dos coeficientes do sistema misturador, nota-se que $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ não é inversível para $N > 1$, pois não é injetivo nem sobrejetivo. Porém, esse operador é diagonalizável, o que decorre de sua simetria [15]. De fato, pode-se mostrar que $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ é simétrico sob o produto interno usual, isto é, $\mathrm{tr}(\boldsymbol{Q}_{\underline{\boldsymbol{x}}}(\boldsymbol{V}_\mathrm{a})\, \boldsymbol{V}_\mathrm{b}^\mathsf{T}) = \mathrm{tr}(\boldsymbol{V}_\mathrm{a} \boldsymbol{Q}_{\underline{\boldsymbol{x}}}^\mathsf{T}(\boldsymbol{V}_\mathrm{b}))$ para todas $\boldsymbol{V}_\mathrm{a}, \boldsymbol{V}_\mathrm{b} \in \mathbb{R}^{N \times N}$. Logo, garante-se que $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$ possui uma base ortonormal de N^2 automatrizes, todas associadas a autovalores reais. Dessas, no máximo N automatrizes estão associadas a autovalores não nulos e geram a imagem do operador, e no mínimo $N(N-1)$ estão associadas ao autovalor nulo e geram o núcleo do operador.

Na prática, embora as direções $\boldsymbol{u}_i$ para $i = 1, 2, \ldots, N$ sejam desconhecidas, as automatrizes $\boldsymbol{V}_i$ podem ser determinadas tirando-se proveito do isomorfismo entre $\mathbb{R}^{N \times N}$ e $\mathbb{R}^{N^2}$ [12]. Após a proposição inicial do JADE [12], constatou-se que suas propriedades estatísticas eram prejudicadas devido à necessidade de se obter as automatrizes com base em estimativas dos cumulantes de ordem quatro [13], [20]. Em virtude disso, propôs-se substituir as automatrizes por um conjunto de matrizes independentes das direções $\boldsymbol{u}_i$, mas cujo subespaço vetorial gerado contivesse o autoespaço de $\boldsymbol{Q}_{\underline{\boldsymbol{x}}}$. Como as automatrizes são simétricas, propôs-se substituí-las por uma base canônica do subespaço de matrizes simétricas contido em $\mathbb{R}^{N \times N}$. Como a dimensão desse subespaço é igual a $N(N+1)/2$ e maior do que N para $N > 1$, torna-se necessária a diagonalização simultânea de um conjunto maior de matrizes de quadricovariância. Por um lado, perde-se em termos de custo computacional; por outro, garante-se segundo [13] melhor desempenho estatístico do estimador. Pode-se mostrar que a consideração dessas matrizes permite a identificação de $\boldsymbol{U}$ mesmo para duas ou mais fontes com autocumulantes de ordem quatro iguais [13], [20]. Na literatura, essas técnicas continuaram sendo denominadas da mesma forma ou apenas como AJD (do inglês *approximate joint diagonalization*) [20].

VII. Conclusões

Nesse artigo, abordou-se por meio de ferramentas matemáticas convencionais uma representação matricial de estatísticas de ordem quatro e sua aplicação a um problema de estimação cega. Por meio dessa aplicação, verificou-se que cuidado deve ser tomado ao utilizar representações matriciais de estatísticas de ordem superior a fim de considerar todo o conjunto de estatísticas necessário à resolução de um determinado problema. Representações tensoriais podem contribuir nessa interpretação, mas às custas de um ferramental matemático mais sofisticado [21], [8].

Quanto à aplicação ao problema de separação cega de fontes, abordou-se e interpretou-se o procedimento de [12], [13] baseado na consideração de um conjunto de matrizes de quadricovariância para se aproximar do limite teórico de separabilidade de Darmois-Skitovich [3], [19]. No entanto, essa abordagem não é única e alternativas vêm sendo propostas na literatura, como, por exemplo, em [22], [23]. Partindo da interpretação teórica aqui apresentada, deseja-se investigar em trabalhos futuros alternativas à escolha de matrizes de pesos $\boldsymbol{V}$ para garantir a identificabilidade da matriz de mistura $\boldsymbol{U}$ no contexto de separação de fontes.

Referências

[1] S. Haykin, *Adaptive Filter Theory*, 5th ed. Pearson, 2013.

[2] D. Godard, "Self-recovering equalization and carrier tracking in two-dimensional data communication systems," *IEEE Trans. Commun.*, vol. 28, no. 11, pp. 1867–1875, Nov. 1980.

[3] P. Comon, "Independent component analysis, a new concept?" *Signal Process.*, vol. 36, no. 3, pp. 287–314, 1994.

[4] L. J. Cao, K. S. Chua, W. K. Chong, H. P. Lee, and Q. M. Gu, "A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine," *Neurocomputing*, vol. 55, no. 1, pp. 321–336, 2003.

[5] A. Subasi and M. I. Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8659–8666, 2010.

[6] A. Hyvärinen, "Independent component analysis: recent advances," *Philos. Trans. R. Soc., A*, vol. 371, no. 1984, 2013.

[7] P. Z. Peebles, *Probability, Random Variables, and Random Signal Principles*, 4th ed. New York City: McGraw-Hill, 2000.

[8] P. McCullagh, *Tensor Methods in Statistics*. London: Chapman and Hall, 1987.

[9] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," *Philos. Trans. R. Soc., A*, vol. 374, no. 2065, 2016.

[10] X.-D. Zhang, *Matrix Analysis and Applications*. Cambdrige: Cambridge University Press, 2017.

[11] J.-F. Cardoso, "Eigen-structure of the fourth-order cumulant tensor with application to the blind source separation problem," in *Int. Conf. Acoust. Speech Signal Process.*, vol. 5, Apr. 1990, pp. 2655–2658.

[12] J.-F. Cardoso and A. Souloumiac, "Blind beamforming for non-Gaussian signals," *Proc. Inst. Elect. Eng., F: Radar and Signal Process.*, vol. 140, no. 6, pp. 362–370, Dec. 1993.

[13] J.-F. Cardoso, "High-order contrasts for independent component analysis," *Neural Comput.*, vol. 11, no. 1, pp. 157–192, Jan. 1999.

[14] B. Holmquist, "Moments and cumulants of the multivariate normal distribution," *Stochastic Anal. Appl.*, vol. 6, no. 3, pp. 273–278, 1988.

[15] G. H. Golub and C. F. V. Loan, *Matrix Computations*, 4th ed. Baltimore: The John Hopkins University Press, 2012.

[16] H. Ould-Baba, V. Robin, and J. Antoni, "Concise formulae for the cumulant matrices of a random vector," *Linear Algebra Appl.*, vol. 485, pp. 392–416, 2015.

[17] L. Qi, W. Sun, and Y. Wang, "Numerical multilinear algebra and its applications," *Front. Math. China*, vol. 2, no. 4, pp. 501–526, Oct. 2007.

[18] P. Comon, G. Golub, L.-H. Lim, and B. Mourrain, "Symmetric tensors and symmetric tensor rank," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1254–1279, 2008.

[19] F. R. M. Pavan and M. D. Miranda, "On the Darmois-Skitovich theorem and spatial independence in blind source separation," *J. Commun. Inf. Syst.*, vol. 33, no. 1, pp. 146–157, 2018.

[20] J. Miettinen, S. Taskinen, K. Nordhausen, and H. Oja, "Fourth moments and independent component analysis," *Stat. Sci.*, vol. 30, no. 3, pp. 372–390, Aug. 2015.

[21] L. De Lathauwer, "A link between the canonical decomposition in multilinear algebra and simultaneous matrix diagonalization," *SIAM J. Matrix Anal. Appl.*, vol. 28, no. 3, pp. 642–666, 2006.

[22] R. Martín-Clemente, J. I. Acha, and C. G. Puntonet, "Eigendecomposition of self-tuned cumulant-matrices for blind source separation," *Signal Process.*, vol. 84, no. 7, pp. 1201–1211, 2004.

[23] A. Dapena, H. J. Pérez-Iglesias, and V. Zarzoso, "Blind channel estimation based on maximizing the eigenvalue spread of cumulant matrices in (2 x 1) Alamouti's coding schemes," *Wireless Commun. Mobile Comput.*, vol. 12, no. 6, pp. 516–528, 2012.