GENERALIZED KURTOSIS AND APPLICATIONS IN BLIND EQUALIZATION OF MIMO CHANNELS

Andre Tkacenko and P. P. Vaidyanathan

Dept. of Electrical Engineering
California Institute of Technology, Pasadena, CA 91125, USA
E-mail: andre@systems.caltech.edu, ppvnath@systems.caltech.edu

ABSTRACT

Recently, the problem of blind equalization of multiple-input multiple-output (MIMO) channels has received much attention from the communications and signal processing communities. Several methods proposed to solve this problem involve computing the kurtosis of the components of the observed output vector sequence. In this paper, we generalize the notion of kurtosis to vector sequences by proposing a definition for a kurtosis matrix of a random vector. It is shown that there are many analogies between our kurtosis matrix and the classical scalar kurtosis. We then use this matrix to develop a necessary condition for blind equalization of MIMO channels. Future work will include the development of a sufficient condition for equalization. [1]

1. INTRODUCTION

One problem which has recently received much attention from the communications and signal processing communities has been that of blind equalization of multiple-input multiple-output (MIMO) linear time-invariant (LTI) channels [2]. This arises, for example, in a multiuser wireless communications system, in which the goal is to remove inter and intrasymbol interference introduced by the channel. Typically, processing time is at a premium and as such, equalization must be done both blindly as well as adaptively.

Several methods have been proposed to solve this problem. Some involve filter bank precoders [4, 5], while others involve higher order statistical moments of the observed output sequences. In particular, the kurtosis, a fourth-order moment, has been found to be useful for blind equalization. The use of kurtosis for blind equalization was first pioneered by Shalvi and Weinstein [6] for the single-input single-output (SISO) case. In that paper, they found a globally convergent criterion under which equalization would occur and showed the usefulness of their criterion through examples.

Many attempts have been made to generalize the results of Shalvi and Weinstein to the MIMO case [1, 3]. In this paper, we consider generalizing their results along a different direction. We introduce the notion of a kurtosis *matrix* for a random vector. It will be seen that many analogies exist between our kurtosis matrix and the classical scalar kurtosis. We then use this matrix to obtain a *necessary* condition for blind equalization of MIMO channels. Development of a *sufficient* condition not yet been done and is the subject of future research.

[1]Work supported in part by the ONR grant N00014-99-1-1002, USA.

2. KURTOSIS MATRIX: DEFINITION AND PROPERTIES

The kurtosis of a scalar complex random variable x, which we denote here by K_x, is classically defined as follows [6].

$$K_x \triangleq E\left[|x|^4\right] - 2\left(E\left[|x|^2\right]\right)^2 - \left|E\left[x^2\right]\right|^2 \qquad (1)$$

Now let $\mathbf{x}$ be any $N \times 1$ complex random vector. We define the *kurtosis* of $\mathbf{x}$, denoted by $\mathbf{K_x}$, to be the following $N \times N$ matrix.

$$\begin{aligned}\mathbf{K_x} \quad &\triangleq \quad E\left[\left(\mathbf{x}\mathbf{x}^\dagger\right)^2\right] - \mathrm{Tr}\left(E\left[\mathbf{x}\mathbf{x}^\dagger\right]\right) E\left[\mathbf{x}\mathbf{x}^\dagger\right] \\ &\quad - \left(E\left[\mathbf{x}\mathbf{x}^\dagger\right]\right)^2 - E\left[\mathbf{x}\mathbf{x}^T\right]\left(E\left[\mathbf{x}\mathbf{x}^T\right]\right)^\dagger \qquad (2)\end{aligned}$$

When $N = 1$, the kurtosis matrix of (2) becomes the scalar kurtosis of (1). To justify the definition in (2), we will show that many analogies exist between the scalar kurtosis and our kurtosis matrix.

One famous property of the scalar kurtosis is that if x is any complex zero-mean Gaussian random variable, then $K_x = 0$ [6]. Analogously, we will show that if $\mathbf{x}$ is any complex zero-mean Gaussian random vector, then $\mathbf{K_x} = \mathbf{0}$. First we show that if $\mathbf{v}$ is any $M \times 1$ real zero-mean Gaussian random vector, then $\mathbf{K_v} = \mathbf{0}$.

2.1. Kurtosis of a real zero-mean Gaussian random vector

Theorem 1 *Let* $\mathbf{v}$ *be any* $M \times 1$ *real zero-mean Gaussian random vector with covariance matrix* $\mathbf{R_v}$ *(i.e.* $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{R_v})$*). Then,*

$$\mathbf{K_v} = \mathbf{0}$$

Proof: We have $\mathbf{R_v} = E\left[\mathbf{v}\mathbf{v}^T\right]$, and since $\mathbf{v}$ is real, we get,

$$\begin{aligned}\mathbf{K_v} \quad &= \quad E\left[\left(\mathbf{v}\mathbf{v}^T\right)^2\right] - \mathrm{Tr}\left(E\left[\mathbf{v}\mathbf{v}^T\right]\right) E\left[\mathbf{v}\mathbf{v}^T\right] \\ &\quad - \left(E\left[\mathbf{v}\mathbf{v}^T\right]\right)^2 - \left(E\left[\mathbf{v}\mathbf{v}^T\right]\right)^2 \\ &= \quad E\left[\left(\mathbf{v}\mathbf{v}^T\right)^2\right] - \left(\mathrm{Tr}\left(\mathbf{R_v}\right)\mathbf{I} + 2\mathbf{R_v}\right)\mathbf{R_v} \qquad (3)\end{aligned}$$

For $E[(\mathbf{v}\mathbf{v}^T)^2]$, we apply the *Mahalanobis* transformation [7],

$$\mathbf{u} \triangleq \mathbf{R_v}^{-\frac{1}{2}}\mathbf{v}$$

Note that $\mathbf{u} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Then, as $\mathbf{R_v}$ is a real symmetric matrix,

$$E\left[\left(\mathbf{v}\mathbf{v}^T\right)^2\right] \quad = \quad E\left[\left(\mathbf{R}_\mathbf{v}^{\frac{1}{2}}\mathbf{u}\mathbf{u}^T\mathbf{R}_\mathbf{v}^{\frac{1}{2}}\right)^2\right]$$

$$= \quad \mathbf{R}_{\mathbf{v}}^{\frac{1}{2}} \underbrace{E\left[\left(\mathbf{u}^T\mathbf{R}_{\mathbf{v}}\mathbf{u}\right)\mathbf{u}\mathbf{u}^T\right]}_{\mathbf{A}} \mathbf{R}_{\mathbf{v}}^{\frac{1}{2}} \quad (4)$$

Now, if we define u_k to be the k-th component of $\mathbf{u}$ for $0 \leq k \leq M-1$, then for the (p,q)-th element of $\mathbf{A}$, we have the following.

$$\begin{aligned} [\mathbf{A}]_{p,q} &= E\left[\left(\sum_{k,l} u_k [\mathbf{R}_{\mathbf{v}}]_{k,l} u_l\right) u_p u_q\right] \\ &= \sum_{k,l} [\mathbf{R}_{\mathbf{v}}]_{k,l} E[u_k u_l u_p u_q] \end{aligned} \quad (5)$$

for $0 \leq p,q \leq M-1$. As $\mathbf{u} \sim \mathcal{N}(\mathbf{0},\mathbf{I})$, the components of $\mathbf{u}$ are uncorrelated and hence independent [7]. Thus, for $p \neq q$, only the terms where k and l are exclusively p or q in (5) are nonzero. So,

$$[\mathbf{A}]_{p,q} = [\mathbf{R}_{\mathbf{v}}]_{p,q} + [\mathbf{R}_{\mathbf{v}}]_{q,p} = 2[\mathbf{R}_{\mathbf{v}}]_{p,q}\ , \ \ p \neq q \quad (6)$$

For $p = q$, we have the following.

$$\begin{aligned} [\mathbf{A}]_{p,p} &= \sum_{k,l\neq p} [\mathbf{R}_{\mathbf{v}}]_{k,l} E\left[u_k u_l u_p^2\right] + \sum_{k} [\mathbf{R}_{\mathbf{v}}]_{k,p} E\left[u_k u_p^3\right] \\ &= \sum_{l\neq p} [\mathbf{R}_{\mathbf{v}}]_{l,l} E\left[u_l^2 u_p^2\right] + [\mathbf{R}_{\mathbf{v}}]_{p,p} E\left[u_p^4\right] \\ &= \sum_{l\neq p} [\mathbf{R}_{\mathbf{v}}]_{l,l} E\left[u_l^2\right] E\left[u_p^2\right] + [\mathbf{R}_{\mathbf{v}}]_{p,p} E\left[u_p^4\right] \\ &= \sum_{l\neq p} [\mathbf{R}_{\mathbf{v}}]_{l,l} + [\mathbf{R}_{\mathbf{v}}]_{p,p} E\left[u_p^4\right] \\ &= \mathrm{Tr}(\mathbf{R}_{\mathbf{v}}) - [\mathbf{R}_{\mathbf{v}}]_{p,p} + [\mathbf{R}_{\mathbf{v}}]_{p,p} E\left[u_p^4\right] \end{aligned} \quad (7)$$

To calculate $E\left[u_p^4\right]$, recall that $u_p \sim \mathcal{N}(0,1)$ for all p since $\mathbf{u} \sim \mathcal{N}(\mathbf{0},\mathbf{I})$. So, using integration by parts, we can show that we have,

$$E\left[u_p^4\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^4 e^{-\frac{u^2}{2}}\, du = 3 \quad (8)$$

Using (8) in (7), we have,

$$[\mathbf{A}]_{p,p} = \mathrm{Tr}(\mathbf{R}_{\mathbf{v}}) + 2[\mathbf{R}_{\mathbf{v}}]_{p,p} \quad (9)$$

Combining (6) and (9), we can conclude that we have,

$$\mathbf{A} = \mathrm{Tr}(\mathbf{R}_{\mathbf{v}})\mathbf{I} + 2\mathbf{R}_{\mathbf{v}}$$

Substituting this into (4) yields the following.

$$E\left[\left(\mathbf{v}\mathbf{v}^T\right)^2\right] = \mathbf{R}_{\mathbf{v}}^{\frac{1}{2}}\left(\mathrm{Tr}(\mathbf{R}_{\mathbf{v}})\mathbf{I} + 2\mathbf{R}_{\mathbf{v}}\right)\mathbf{R}_{\mathbf{v}}^{\frac{1}{2}} = \left(\mathrm{Tr}(\mathbf{R}_{\mathbf{v}})\mathbf{I} + 2\mathbf{R}_{\mathbf{v}}\right)\mathbf{R}_{\mathbf{v}}$$

Finally substituting the above into (3), we find that $\mathbf{K}_{\mathbf{v}} = \mathbf{0}$ and this completes the proof. $\bigtriangledown\bigtriangledown\bigtriangledown$

To prove that the kurtosis of any *complex* zero-mean Gaussian random vector is likewise the zero matrix, we must relate the kurtosis of any complex random vector to that of the real vector formed by concatenating the real and imaginary parts.

2.2. Relating the kurtosis of a complex random vector to its real and imaginary parts

Any $N \times 1$ complex random vector $\mathbf{x}$ can be decomposed as,

$$\mathbf{x} = \mathbf{x}_r + j\mathbf{x}_i \quad (10)$$

where $\mathbf{x}_r$ and $\mathbf{x}_i$ are the real and imaginary parts of $\mathbf{x}$, respectively. Define the *real* $2N \times 1$ vector $\mathbf{v}$ as follows.

$$\mathbf{v} \triangleq \begin{bmatrix} \mathbf{x}_r \\ \mathbf{x}_i \end{bmatrix} \quad (11)$$

Clearly, $\mathbf{x}$ and $\mathbf{v}$ are related to each other via the following $N \times 2N$ linear transformation.

$$\mathbf{x} = \mathbf{T}\mathbf{v} \text{ where } \mathbf{T} \triangleq \begin{bmatrix} \mathbf{I}_N & j\mathbf{I}_N \end{bmatrix} \quad (12)$$

Then we have the following lemma.

Lemma 1 *Let $\mathbf{x}$ by any $N \times 1$ complex random vector with a decomposition into real and imaginary parts as in (10). Also, let $\mathbf{v}$ by the vector formed by concatenating the real and imaginary parts as in (11). Then, we have,*

$$\mathbf{K}_{\mathbf{x}} = \mathbf{T}\mathbf{K}_{\mathbf{v}}\mathbf{T}^\dagger$$

where $\mathbf{T}$ is the matrix given in (12).

Proof: For simplicity in (2) here, we will exploit the fact that $\mathrm{Tr}\left(E\left[\mathbf{y}\mathbf{y}^T\right]\right) = E\left[\mathbf{y}^T\mathbf{y}\right]$ and $\mathrm{Tr}\left(E\left[\mathbf{y}\mathbf{y}^\dagger\right]\right) = E\left[\mathbf{y}^\dagger\mathbf{y}\right]$ for any vector $\mathbf{y}$. Note that we have the following.

$$\mathbf{x}\mathbf{x}^\dagger = \mathbf{T}\mathbf{v}\mathbf{v}^T\mathbf{T}^\dagger\ ,\ \mathbf{x}^\dagger\mathbf{x} = \mathbf{v}^T\mathbf{v}$$

Here, the last equality comes from the fact that $\mathbf{x}$ and $\mathbf{v}$ have the same magnitude. Thus, we have,

$$\begin{aligned} E\left[\left(\mathbf{x}\mathbf{x}^\dagger\right)^2\right] &= E\left[\mathbf{x}\mathbf{x}^\dagger\mathbf{x}\mathbf{x}^\dagger\right] = E\left[\mathbf{x}\mathbf{v}^T\mathbf{v}\mathbf{x}^\dagger\right] \\ &= E\left[\mathbf{T}\mathbf{v}\mathbf{v}^T\mathbf{v}\mathbf{v}^T\mathbf{T}^\dagger\right] = \mathbf{T}E\left[\mathbf{v}\mathbf{v}^T\mathbf{v}\mathbf{v}^T\right]\mathbf{T}^\dagger \\ &= \mathbf{T}\left(E\left[\left(\mathbf{v}\mathbf{v}^T\right)^2\right]\right)\mathbf{T}^\dagger \end{aligned} \quad (13)$$

Also, we have,

$$\begin{aligned} E\left[\mathbf{x}\mathbf{x}^\dagger\right] &= E\left[\mathbf{T}\mathbf{v}\mathbf{v}^T\mathbf{T}^\dagger\right] = \mathbf{T}E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{T}^\dagger \\ E\left[\mathbf{x}\mathbf{x}^T\right] &= E\left[\mathbf{T}\mathbf{v}\mathbf{v}^T\mathbf{T}^T\right] = \mathbf{T}E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{T}^T \end{aligned}$$

Hence, we have the following.

$$\begin{aligned} E\left[\mathbf{x}^\dagger\mathbf{x}\right]E\left[\mathbf{x}\mathbf{x}^\dagger\right] &= E\left[\mathbf{v}^T\mathbf{v}\right]\cdot\mathbf{T}E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{T}^\dagger \\ &= \mathbf{T}\left(E\left[\mathbf{v}^T\mathbf{v}\right]E\left[\mathbf{v}\mathbf{v}^T\right]\right)\mathbf{T}^\dagger \end{aligned} \quad (14)$$

$$\left(E\left[\mathbf{x}\mathbf{x}^\dagger\right]\right)^2 = \mathbf{T}E\left[\mathbf{v}\mathbf{v}^T\right]\left(\mathbf{T}^\dagger\mathbf{T}\right)E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{T}^\dagger \quad (15)$$

$$\begin{aligned} E\left[\mathbf{x}\mathbf{x}^T\right]\left(E\left[\mathbf{x}\mathbf{x}^T\right]\right)^\dagger &= \mathbf{T}E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{T}^T\mathbf{T}^*E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{T}^\dagger \\ &= \mathbf{T}E\left[\mathbf{v}\mathbf{v}^T\right]\left(\mathbf{T}^\dagger\mathbf{T}\right)^*E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{T}^\dagger \end{aligned} \quad (16)$$

But note now that we have,

$$\mathbf{T}^\dagger\mathbf{T} = \begin{bmatrix} \mathbf{I}_N \\ -j\mathbf{I}_N \end{bmatrix} \begin{bmatrix} \mathbf{I}_N & j\mathbf{I}_N \end{bmatrix} = \begin{bmatrix} \mathbf{I}_N & j\mathbf{I}_N \\ -j\mathbf{I}_N & \mathbf{I}_N \end{bmatrix}$$

This can be written as,

$$\mathbf{T}^\dagger\mathbf{T} = \mathbf{I}_{2N} + \mathbf{C} \text{ where } \mathbf{C} \triangleq \begin{bmatrix} \mathbf{0} & j\mathbf{I}_N \\ -j\mathbf{I}_N & \mathbf{0} \end{bmatrix}$$

Note that $\mathbf{C}^* = -\mathbf{C}$. Using the above decomposition of $\mathbf{T}^\dagger\mathbf{T}$ in (15) and (16), we have the following.

$$\left(E\left[\mathbf{x}\mathbf{x}^\dagger\right]\right)^2 = \mathbf{T}\left(E\left[\mathbf{v}\mathbf{v}^T\right]\right)^2\mathbf{T}^\dagger + \mathbf{T}E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{C}E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{T}^\dagger \quad (17)$$

$$E\left[\mathbf{x}\mathbf{x}^T\right]\left(E\left[\mathbf{x}\mathbf{x}^T\right]\right)^\dagger = \mathbf{T}\left(E\left[\mathbf{v}\mathbf{v}^T\right]\right)^2\mathbf{T}^\dagger - \mathbf{T}E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{C}E\left[\mathbf{v}\mathbf{v}^T\right]\mathbf{T}^\dagger \quad (18)$$

Clearly, the terms in (17) and (18) with the matrix $\mathbf{C}$ will cancel each other out when we compute $\mathbf{K}_\mathbf{x}$. Combining (13), (14), (17), and (18), we have the following.

$$\mathbf{K}_\mathbf{x} = \mathbf{T}\left[E\left[\left(\mathbf{v}\mathbf{v}^T\right)^2\right] - E\left[\mathbf{v}^T\mathbf{v}\right]E\left[\mathbf{v}\mathbf{v}^T\right] - 2\left(E\left[\mathbf{v}\mathbf{v}^T\right]\right)^2\right]\mathbf{T}^\dagger \quad (19)$$

But since $\mathbf{v}$ is real, we have,

$$\mathbf{K}_\mathbf{v} = E\left[\left(\mathbf{v}\mathbf{v}^T\right)^2\right] - E\left[\mathbf{v}^T\mathbf{v}\right]E\left[\mathbf{v}\mathbf{v}^T\right] - 2\left(E\left[\mathbf{v}\mathbf{v}^T\right]\right)^2$$

Using the above equation in (19), we find that $\mathbf{K}_\mathbf{x} = \mathbf{T}\mathbf{K}_\mathbf{v}\mathbf{T}^\dagger$ which completes the proof. $\triangledown\triangledown\triangledown$

Combining Theorem 1 and Lemma 1, we can prove the following theorem.

2.3. Kurtosis of a complex zero-mean Gaussian random vector

Theorem 2 *Let $\mathbf{x}$ be any $N \times 1$ complex zero-mean Gaussian random vector. Then, we have,*

$$\mathbf{K}_\mathbf{x} = \mathbf{0}$$

Proof: Recall that a complex Gaussian random vector is one for which the real and imaginary parts are jointly Gaussian [7]. In other words, a complex vector $\mathbf{x}$ with decomposition into real and imaginary parts as in (10) is a complex Gaussian random vector if the real vector $\mathbf{v}$ formed by concatenating the real and imaginary parts as in (11) is a real Gaussian random vector. If $\mathbf{x}$ is zero-mean, then certainly $\mathbf{v}$ is also. Hence, from Theorem 2, we have $\mathbf{K}_\mathbf{v} = \mathbf{0}$. But from Lemma 1, we know that $\mathbf{K}_\mathbf{x} = \mathbf{T}\mathbf{K}_\mathbf{v}\mathbf{T}^\dagger$. Thus, $\mathbf{K}_\mathbf{x} = \mathbf{0}$ and this completes the proof. $\triangledown\triangledown\triangledown$

In the next section, we focus on the kurtosis of certain random processes and will continue to see many analogies between our kurtosis matrix and the classical scalar kurtosis.

3. KURTOSIS OF RANDOM VECTOR PROCESSES

In general, the kurtosis of a random vector process $\mathbf{x}(n)$, namely $\mathbf{K}_{\mathbf{x}(n)}$, will depend on n. If it does not, then for brevity, we will denote it by $\mathbf{K}_\mathbf{x}$. For the purpose of blind equalization, we focus on the following types of zero-mean random vector processes.

- Independent, identically distributed (i.i.d.) processes.
- i.i.d. processes passed through MIMO LTI systems.

3.1. Kurtosis of zero-mean i.i.d. processes

We say that a random vector processes $\mathbf{x}(n)$ is i.i.d. if all components of $\mathbf{x}(n)$, both spatial and temporal, are independent and identically distributed. Then, we have the following theorem.

Theorem 3 *Let $\mathbf{x}(n)$ be some $N \times 1$ zero-mean i.i.d. process and let $x_i(n)$ denote the i-th component of $\mathbf{x}(n)$ for $0 \le i \le N-1$. Then, we have,*

$$\mathbf{K}_{\mathbf{x}(n)} = \mathbf{K}_\mathbf{x} = K_x\mathbf{I}$$

where K_x is the kurtosis of any component of $\mathbf{x}(n)$ given by,

$$K_x = \mu_{x,4} - 2\sigma_x^4 - \lambda_{x,4} \quad (20)$$

where we have,

$$\mu_{x,4} \triangleq E\left[|x_i(n)|^4\right],\ \sigma_x^2 \triangleq E\left[|x_i(n)|^2\right],\ \lambda_{x,4} \triangleq \left|E\left[x_i^2(n)\right]\right|^2$$

for all $0 \le i \le N-1$ and $n \in \mathbb{Z}$.

Proof: As $\mathbf{x}(n)$ is an i.i.d. process, we have,

$$E\left[x_{i_0}(n_0)x_{i_1}^*(n_1)x_{i_2}(n_2)x_{i_3}^*(n_3)\right] = K_x\delta_{i_0,i_1,i_2,i_3}\delta_{n_0,n_1,n_2,n_3} + \sigma_x^4\delta_{i_0,i_1}\delta_{i_2,i_3}\delta_{n_0,n_1}\delta_{n_2,n_3} + \sigma_x^4\delta_{i_0,i_3}\delta_{i_1,i_2}\delta_{n_0,n_3}\delta_{n_1,n_2} + \lambda_{x,4}\delta_{i_0,i_2}\delta_{i_1,i_3}\delta_{n_0,n_2}\delta_{n_1,n_3} \quad (21)$$

where $\delta_{i_0,\ldots,i_{L-1}}$ denotes the Kronecker delta function which is unity for $i_0 = \cdots = i_{L-1}$ and zero otherwise. To justify this, note that the fourth-order moment from (21) can only take on three possible nonzero values, namely $\mu_{x,4}$, σ_x^4, and $\lambda_{x,4}$. When all component and temporal indices are respectively equal, then we should have $\mu_{x,4}$, which is the case here. The only other cases in which a nonzero value occurs are when sets of two component and temporal indices are exclusively equal, in which case we obtain either σ_x^4 or $\lambda_{x,4}$, depending on how the pairing is done. Thus,

$$\left[E\left[\left(\mathbf{x}(n)\mathbf{x}^\dagger(n)\right)^2\right]\right]_{p,q} = \sum_{i=0}^{N-1} E\left[x_p(n)\left|x_i(n)\right|^2 x_q^*(n)\right] = \begin{cases} 0\,, & p \neq q \\ \mu_{x,4} + (N-1)\sigma_x^4\,, & p = q \end{cases}$$

Hence, we have,

$$E\left[\left(\mathbf{x}(n)\mathbf{x}^\dagger(n)\right)^2\right] = \left(\mu_{x,4} + (N-1)\sigma_x^4\right)\mathbf{I} \quad (22)$$

Also, we have,

$$E\left[\mathbf{x}(n)\mathbf{x}^\dagger(n)\right] = \sigma_x^2\mathbf{I} \quad (23)$$

$$E\left[\mathbf{x}(n)\mathbf{x}^T(n)\right]\left(E\left[\mathbf{x}(n)\mathbf{x}^T(n)\right]\right)^\dagger = \lambda_{x,4}\mathbf{I} \quad (24)$$

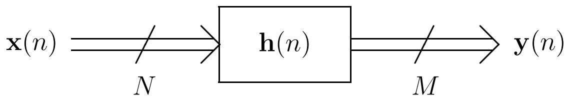

Fig. 1. MIMO LTI system.

Combining (22), (23), and (24), we obtain the following from (2).

$$\begin{aligned}\mathbf{K}_{\mathbf{x}(n)} &= \left(\mu_{x,4} + (N-1)\sigma_x^4\right)\mathbf{I} - N\sigma_x^4\mathbf{I} - \sigma_x^4\mathbf{I} - \lambda_{x,4}\mathbf{I} \\ &= \left(\mu_{x,4} - 2\sigma_x^4 - \lambda_{x,4}\right)\mathbf{I}\end{aligned}$$

As $\mathbf{K}_{\mathbf{x}(n)}$ does not depend on n, we denote it by $\mathbf{K}_{\mathbf{x}}$. Using the definition of K_x as given in (20), this completes the proof. $\bigtriangledown\bigtriangledown\bigtriangledown$

Note that if $\mathbf{x}(n)$ is a zero-mean i.i.d. process, then its autocorrelation matrix $\mathbf{R}_{\mathbf{x}} = E\left[\mathbf{x}(n)\mathbf{x}^\dagger(n)\right]$ is given by $\mathbf{R}_{\mathbf{x}} = \sigma_x^2\mathbf{I}$, which is analogous to the result proven in Theorem 3. This further justifies the definition of the kurtosis matrix proposed in (2).

3.2. Kurtosis of a filtered zero-mean i.i.d. process

Theorem 4 *Suppose that the $N \times 1$ zero-mean i.i.d. process $\mathbf{x}(n)$ is input to the $M \times N$ MIMO LTI system of Figure 1 with impulse response $\mathbf{h}(n)$. Then the $M \times 1$ output process $\mathbf{y}(n)$ has kurtosis,*

$$\mathbf{K}_{\mathbf{y}(n)} = \mathbf{K}_{\mathbf{y}} = K_x\left(\sum_m \mathbf{h}(m)\mathbf{\Lambda}(m)\mathbf{h}^\dagger(m)\right)$$

where K_x is as in (20) and $\mathbf{\Lambda}(m)$ is a diagonal $N \times N$ matrix such that,

$$\left[\mathbf{\Lambda}(m)\right]_{k,k} = \left[\mathbf{h}^\dagger(m)\mathbf{h}(m)\right]_{k,k} \tag{25}$$

Proof: By convolution, we have,

$$\mathbf{y}(n) = \sum_m \mathbf{h}(m)\mathbf{x}(n-m)$$

If we define $h_{p,q}(n) \triangleq \left[\mathbf{h}(n)\right]_{p,q}$ and denote the k-th component of $\mathbf{y}(n)$ by $y_k(n)$, then we have,

$$y_k(n) = \sum_m \sum_{l=0}^{N-1} h_{k,l}(m)x_l(n-m)\,,\ 0 \le k \le M-1$$

Using this along with (21), we can show that we have,

$$\begin{aligned}&\left[E\left[\left(\mathbf{y}(n)\mathbf{y}^\dagger(n)\right)^2\right]\right]_{p,q} = \\ &K_x\left(\sum_m \sum_{l=0}^{N-1} h_{p,l}(m)\left(\sum_{i=0}^{M-1}\left|h_{i,l}(m)\right|^2\right)h_{q,l}^*(m)\right) \\ &+ \sigma_x^4\left(\sum_m \sum_{i=0}^{M-1}\sum_{l=0}^{N-1}\left|h_{i,l}(m)\right|^2\right)\left(\sum_m \sum_{l=0}^{N-1} h_{p,l}(m)h_{q,l}^*(m)\right) \\ &+ \sigma_x^4\left(\sum_{m_0,m_1}\sum_{i,l_0,l_1} h_{p,l_0}(m_0)h_{i,l_0}^*(m_0)h_{i,l_1}(m_1)h_{q,l_1}^*(m_1)\right) \\ &+ \lambda_{x,4}\left(\sum_{m_0,m_1}\sum_{i,l_0,l_1} h_{p,l_0}(m_0)h_{i,l_0}(m_0)h_{i,l_1}^*(m_1)h_{q,l_1}^*(m_1)\right)\end{aligned}$$

From this, we can see that we have the following.

$$\begin{aligned}E\left[\left(\mathbf{y}(n)\mathbf{y}^\dagger(n)\right)^2\right] &= K_x\left(\sum_m \mathbf{h}(m)\mathbf{\Lambda}(m)\mathbf{h}^\dagger(m)\right) \\ &+ \sigma_x^4\mathrm{Tr}\left(\sum_m \mathbf{h}(m)\mathbf{h}^\dagger(m)\right)\left(\sum_m \mathbf{h}(m)\mathbf{h}^\dagger(m)\right) \\ &+ \sigma_x^4\left(\sum_m \mathbf{h}(m)\mathbf{h}^\dagger(m)\right)^2 \\ &+ \lambda_{x,4}\left(\sum_m \mathbf{h}(m)\mathbf{h}^T(m)\right)\left(\sum_m \mathbf{h}(m)\mathbf{h}^T(m)\right)^\dagger \end{aligned}\tag{26}$$

with $\mathbf{\Lambda}(m)$ as in (25). Also, we have,

$$\mathbf{R}_{\mathbf{y}} = E\left[\mathbf{y}(n)\mathbf{y}^\dagger(n)\right] = \sigma_x^2\sum_m \mathbf{h}(m)\mathbf{h}^\dagger(m) \tag{27}$$

$$\begin{aligned}&E\left[\mathbf{y}(n)\mathbf{y}^T(n)\right]\left(E\left[\mathbf{y}(n)\mathbf{y}^T(n)\right]\right)^\dagger = \\ &\lambda_{x,4}\left(\sum_m \mathbf{h}(m)\mathbf{h}^T(m)\right)\left(\sum_m \mathbf{h}(m)\mathbf{h}^T(m)\right)^\dagger\end{aligned}\tag{28}$$

Combining (26), (27), and (28) with (2), we obtain,

$$\mathbf{K}_{\mathbf{y}(n)} = \mathbf{K}_{\mathbf{y}} = K_x\left(\sum_m \mathbf{h}(m)\mathbf{\Lambda}(m)\mathbf{h}^\dagger(m)\right)$$

which completes the proof. $\bigtriangledown\bigtriangledown\bigtriangledown$

We now proceed to use the results of Theorem 4 to establish a criterion analogous to the one proposed by Shalvi and Weinstein [6] for blind equalization.

4. APPLICATION TO BLIND EQUALIZATION

Regarding the blind equalization problem, we typically assume that we have an $N \times 1$ zero-mean i.i.d. input, say $\mathbf{x}(n)$, only whose statistics are known. This input is applied to an $L \times N$ unknown channel $\mathbf{C}(z)$ with $L \ge N$, and the goal is to cascade this channel with an $N \times L$ equalizer $\mathbf{E}(z)$ such that the output $\mathbf{y}(n)$ is as close to the input $\mathbf{x}(n)$ as possible upto a permutation as well as a scaling and delaying of the components. If $\mathbf{H}(z)$ denotes the $N \times N$ cascade of the channel and the equalizer, namely $\mathbf{H}(z) = \mathbf{E}(z)\mathbf{C}(z)$, then we say that we have achieved equalization if we have the following [1].

$$\mathbf{H}(z) = \mathbf{P}\mathbf{D}(z) \tag{29}$$

where $\mathbf{P}$ is a permutation matrix and $\mathbf{D}(z)$ is a diagonal matrix of scales and delays given by,

$$\mathbf{D}(z) = \mathrm{diag}\left(c_0 z^{-n_0}, c_1 z^{-n_1}, \ldots, c_{N-1}z^{-n_{N-1}}\right)$$

For the single-input single-output (SISO) case, the criterion for blind equalization to occur is equivalent to saying that we have,

$$h(n) = c_0\delta(n-n_0) \tag{30}$$

In [6], only the SISO case was analyzed. A necessary and sufficient condition for equalization was derived using the inequality,

$$\sum_m |h(m)|^4 \le \left(\sum_m |h(m)|^2\right)^2 \tag{31}$$

for any scalar sequence $h(n)$ with equality iff $h(n)$ is as in (30). The term on the left hand side of (31) is proportional to the output kurtosis, whereas the right hand side is proportional to the square of the output variance. By globally maximizing the magnitude of the output kurtosis subject to keeping the output variance fixed, we obtain the equalization criterion of (30). Shalvi and Weinstein showed that this maximization problem had no local extrema and thus showed that an adaptive algorithm used for the maximization would converge to the desired criterion of (30) as best as possible.

Here, we consider developing a criterion similar to (31) for the MIMO case. In our inequality, the left hand side will be related to the output kurtosis matrix whereas the right hand side will be related to the output autocorrelation matrix. However, our criterion will only be a *necessary* condition for blind equalization.

Theorem 5 *Let $\mathbf{h}(n)$ denote some $M \times N$ matrix sequence. Then,*

$$\mathrm{Tr}\left(\sum_m \mathbf{h}(m)\mathbf{\Lambda}(m)\mathbf{h}^\dagger(m)\right) \le \mathrm{Tr}\left(\left(\sum_m \mathbf{h}(m)\mathbf{h}^\dagger(m)\right)^2\right)$$

where $\mathbf{\Lambda}(m)$ is as in (25). We have equality iff,

$$\mathbf{h}^\dagger(m)\mathbf{h}(n) = \mathbf{\Lambda}(m)\delta_{m,n}$$

Proof: For brevity, we define the matrices $\mathbf{A}$ and $\mathbf{B}$ as follows.

$$\mathbf{A} \triangleq \sum_m \mathbf{h}(m)\mathbf{\Lambda}(m)\mathbf{h}^\dagger(m)\,,\ \mathbf{B} \triangleq \left(\sum_m \mathbf{h}(m)\mathbf{h}^\dagger(m)\right)^2$$

Then, if $h_{p,q}(n)$ denotes the (p,q)-th element of $\mathbf{h}(n)$, we have,

$$\mathrm{Tr}\,(\mathbf{A}) = \sum_m \sum_l \left(\sum_k |h_{k,l}(m)|^2\right)^2$$

$$\begin{aligned}\mathrm{Tr}\,(\mathbf{B}) &= \sum_m \sum_l \left(\sum_k |h_{k,l}(m)|^2\right)^2 \\ &+ \sum_m \sum_{j,k\neq j} \left|\sum_l h_{l,j}(m)h^*_{l,k}(m)\right|^2 \\ &+ \sum_{m,n\neq m} \sum_{j,k} \left|\sum_l h_{l,j}(n)h^*_{l,k}(m)\right|^2\end{aligned}$$

Clearly we have,

$$\begin{aligned}\mathrm{Tr}\,(\mathbf{B}) - \mathrm{Tr}\,(\mathbf{A}) &= \sum_m \sum_{j,k\neq j} \left|\sum_l h^*_{l,k}(m)h_{l,j}(m)\right|^2 \\ &+ \sum_{m,n\neq m} \sum_{j,k} \left|\sum_l h^*_{l,k}(m)h_{l,j}(n)\right|^2\end{aligned}$$

and so $\mathrm{Tr}\,(\mathbf{B}) - \mathrm{Tr}\,(\mathbf{A}) \ge 0$ with equality iff we have,

$$\sum_l h^*_{l,k}(m)h_{l,j}(m) = 0 \ \ \forall\, m, j, k \neq j \qquad (32)$$

$$\text{and} \quad \sum_l h^*_{l,k}(m)h_{l,j}(n) = 0 \ \ \forall\, m, n \neq m, j, k \qquad (33)$$

In matrix form, the first condition in (32) is equivalent to saying,

$$\mathbf{h}^\dagger(m)\mathbf{h}(m) = \mathcal{D}(m)$$

where $\mathcal{D}(m)$ is a diagonal matrix for all m. Obviously, we must have $\mathcal{D}(m) = \mathbf{\Lambda}(m)$ in light of the definition of $\mathbf{\Lambda}(m)$ in (25). The second condition for equality in (33) is equivalent to saying,

$$\mathbf{h}^\dagger(m)\mathbf{h}(n) = \mathbf{0} \ \ \forall\, n \neq m$$

Combining these conditions, we conclude that,

$$\mathrm{Tr}\,(\mathbf{A}) \le \mathrm{Tr}\,(\mathbf{B}) \ \text{ with equality iff } \ \mathbf{h}^\dagger(m)\mathbf{h}(n) = \mathbf{\Lambda}(m)\delta_{m,n}$$

which completes the proof. ▽ ▽ ▽

When $M = N = 1$, we obtain the inequality (31) used in [6], as well as the condition for equality of (30). Despite this analogy, when $M = N > 1$, the condition for equality in Theorem 5 does not guarantee equalization. For example, when $M = N = 2$,

$$\mathbf{h}(n) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\delta(n) \Longleftrightarrow \mathbf{H}(z) = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

satisfies the condition for equality in Theorem 5 but $\mathbf{H}(z)$ is not of the desired form of (29). Conversely though, if $\mathbf{H}(z)$ is as in (29), then it can be shown that $\mathbf{h}(n)$ will satisfy the condition for equality in Theorem 5. Thus, for $M = N > 1$, Theorem 5 establishes a *necessary* condition for equalization and not a sufficient one.

5. CONCLUDING REMARKS

We have proposed a definition for a kurtosis matrix and have shown several analogies between this matrix and the classical scalar kurtosis. These analogies promote the validity of our definition. We have also shown that using the kurtosis matrix we can obtain a necessary condition for blind equalization of MIMO LTI systems. Using this matrix to establish a *sufficient* condition is the subject of future research. We can then apply the kurtosis matrix to the problem of blind equalization of MIMO channels in a practical setting.

6. REFERENCES

[1] Y. Inouye, “Criteria for blind deconvolution of multichannel linear time-invariant systems,” *IEEE Trans. Signal Processing*, 46(12):3432-3436, Dec. 1998.

[2] G. B. Giannakis, Y. Hua, P. Stoica, and L. Tong, editors, *Signal Processing Advances in Wireless and Mobile Communications, Volume 1: Trends in Channel Estimation and Equalization*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 2000.

[3] C. B. Papadias, “Globally convergent blind source separation based on a multiuser kurtosis maximization criterion,” *IEEE Trans. Signal Processing*, 48(12):3508-3519, Dec. 2000.

[4] A. Scaglione, G. B. Giannakis, and S. Barbarossa, “Redundant filter bank precoders and equalizers Part I: unification and optimal designs,” *IEEE Trans. Signal Processing*, 47(7):1988-2006, Jul. 1999.

[5] A. Scaglione, G. B. Giannakis, and S. Barbarossa, “Redundant filter bank precoders and equalizers Part II: synchronization and direct equalization,” *IEEE Trans. Signal Processing*, 47(7):2007-2022, Jul. 1999.

[6] O. Shalvi and E. Weinstein, “New criteria for blind deconvolution of nonminimum phase systems (channels),” *IEEE Trans. Inform. Theory*, 36(2):312-321, Mar. 1990.

[7] C. W. Therrien, *Discrete Random Signals and Statistical Signal Processing*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1992.