PARAMETRIC ESTIMATION OF CUMULANTS

Yair Noam and Joseph Tabrikian

Dept. of ECE, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel
Email:{naimy,joseph}@ee.bgu.ac.il

ABSTRACT

The problem of higher-order cumulants estimation is addressed in this paper. Higher-order cumulants are necessary in many applications, such as blind source separation (BSS) and blind deconvolution. In these applications, the cumulants are usually estimated using sample estimation. In this paper, a parametric method for cumulants estimation using the Gaussian mixture model (GMM) is derived. The cumulants are expressed in terms of the GMM parameters, and estimated using the maximum-likelihood estimator. The performance of the proposed model-based method was evaluated and compared to sample estimation using computer simulations. The results show that the model-based estimation outperforms the sample estimation in terms of root-mean-square error.

1. INTRODUCTION

Higher order statistics (HOS) are widely used in many diverse fields, such as blind deconvolution [1] and blind source separation (BSS) [2]. In these applications, the statistics, represented by cumulants, are estimated using sample estimation in which the statistical expectations are replaced by averaging over the process realizations assuming stationarity and ergodicity of the signal. The performance of the higher order cumulants sample estimation is not necessarily optimal.

In this paper, a parametric method for cumulants is presented. In [3] it is shown that any density can be estimated to any degree of approximation using a finite order Gaussian mixture model (GMM). Therefore, the GMM is chosen to model the probability density function (pdf) of the data. The cumulants are expressed by the GMM parameters, and the maximum likelihood (ML) estimator of the cumulants is derived using the expectation-maximization (EM) algorithm.

We first present the ML estimation of the cumulants for an i.i.d. random vector process. This problem is encountered in BSS applications. Next, we derive an estimator for cumulants of a strict sense stationary random process. This problem is mainly encountered in blind deconvolution applications. The consistency of the estimator for N_0-dependent signals is proved.

2. NOTATION AND DEFINITIONS

Let $\mathbf{x} = [x_1, \ldots, x_d]^T$ be a random vector, $J = \{1, \ldots, d\}$ denote the set of indices of the components of $\mathbf{x}$, and $\{j_l\}_{l=1}^L$ be a sequence such that $j_l \in J$ and $j_l \neq j_m \; \forall l \neq m$. The cumulant of the vector $\mathbf{x}$, $\mathrm{cum}(\mathrm{x}_{\mathrm{j}_1}^{\mathrm{r}_1}, \ldots, \mathrm{x}_{\mathrm{j}_L}^{\mathrm{r}_L})$, is defined as the coefficient of the product $v_{j_1}^{r_1} \cdot v_{j_2}^{r_2} \cdot \ldots \cdot v_{j_L}^{r_L}$ in the Taylor series expansion of the cumulant generating function

$$g(\mathbf{v}) = \log\left(E\{\exp(j\mathbf{v}^T\mathbf{x})\}\right) \qquad (1)$$

where $\{r_l\}_{l=1}^L$ is a sequence of positive integers, $(\cdot)^T$ denotes the transpose operation, $\mathbf{v} = [v_1, \ldots, v_L]^T$, and E stands for the statistical expectation. The order of the cumulant is given by $k = \sum_{l=1}^L r_l$.

For simplicity of derivations, we are interested to transform the data such that every cumulant of $\mathbf{x}$ can be expressed in terms of simple cumulants. Thus, we define the following mapping function, $h : I \longrightarrow J$ where $I = \{1, 2, \ldots, k\}$, $J = \{1, \ldots, d\}$:

$$\begin{aligned} &h[1] = \cdots = h[r_1] = j_1 \, , \\ &h[r_1+1] = \cdots = h[r_1+r_2] = j_2 \, , \\ &\vdots \\ &h[r_1+\cdots+r_{L-1}+1] = \cdots = h[r_1+\cdots+r_L] = j_L \, . \end{aligned} \qquad (2)$$

The cumulant $\mathrm{cum}(x_{j_1}^{r_1}, \ldots, x_{j_L}^{r_L})$ can now be calculated via the simple cumulant of $\mathbf{y} = [y_1, \ldots, y_k]^T$, where $\mathbf{y}$ is a permutation of $\mathbf{x}$ according to $y_i = x_{h[i]}$, $i \in I$,

$$c_{\mathbf{y}}(I) = \mathrm{cum}(y_1, \ldots, y_k) = \mathrm{cum}(x_{j_1}^{r_1}, \ldots, x_{j_L}^{r_L}) \, . \qquad (3)$$

The permutation from $\mathbf{x}$ to $\mathbf{y}$ can be expressed via the transformation $\mathbf{y} = \mathbf{H}_{h,I,J}(\mathbf{x})$ where $\mathbf{H}_{h,I,J} : \mathbf{R}^d \longrightarrow \mathbf{R}^k$. For example, for calculating $\mathrm{cum}(x_1^2, x_2, x_3)$ in which $k = 4$, $d = 3$, the function $h[\cdot]$ is set to be $h[1] = h[2] = 1$, $h[3] = 2$, $h[4] = 3$, and $\mathbf{y} = [y_1, y_2, y_3, y_4]^T = [x_1, x_1, x_2, x_3]^T$. Note that $\mathbf{y}$ might contain variables not necessarily distinct.

The cumulants can be obtained from the moments via the following moment-to-cumulant formula [1]:

$$c_{\mathbf{y}}(I) = \sum_{\alpha \in A} (-1)^{q-1}(q-1)! \prod_{I_p \in \alpha : p=1}^{q} m_{\mathbf{y}}(I_p) \, , \qquad (4)$$

in which A is a set of all sequences of disjoint subsets of I whose union is equal to I

$$A = \left\{\alpha | \alpha = \{I_p\}_{p=1}^q ; I_p \subseteq I ; I_p \cap I_m = \phi \; \forall p \neq m ; \cup_{p=1}^q I_p = I \right\} \, , \qquad (5)$$

and

$$m_{\mathbf{y}}(I_p) = E\left\{ \prod_{i \in I_p} y_i \right\} \qquad (6)$$

are the joint moments of $\mathbf{x}$.

3. MODEL-BASED ESTIMATION OF CUMULANTS OF AN I.I.D. VECTOR SEQUENCE

The cumulants of a vector process are commonly obtained by substitution of the sample estimates of its moments into (4). The sample estimation is not claimed to be optimal in any sense, except for the case of Gaussian random processes, in which the sample estimation of the 1st and 2nd cumulants can be shown to be the ML estimator. The performance of the sample cumulant estimator is poor, especially for higher order cumulants.

In this section, a parametric approach for estimating cumulants of a random vector based on GMM is presented. In [3] it is proved that any density can be estimated to any degree of approximation using a finite Gaussian mixture. The ML estimator of the cumulants can be obtained from the estimated parameters of the Gaussian mixture pdf. Suppose that the observed process $\{\mathbf{x}_n\}_{n=1}^N$, $\mathbf{x}_n = [x_{n1}, \ldots, x_{nd}]^T$ is an i.i.d. random process with GMM pdf:

$$f_{\mathbf{x}}(\mathbf{x}_n; \boldsymbol{\theta}) = \sum_{m=1}^{M} w_m N(\mathbf{x}_n; \boldsymbol{\mu}_m, \mathbf{C}_m) \ , \qquad (7)$$

where $\boldsymbol{\mu}_m \in \mathbf{R}^{d\times 1}$ and $\mathbf{C}_m \in \mathbf{R}^{d\times d}$ are the mean vector and the covariance matrix of the mth Gaussian, respectively. The unknown vector parameter, $\boldsymbol{\theta} = \left\{\{w_m\}_{m=1}^M, \{\boldsymbol{\mu}_m\}_{m=1}^M, \{\mathbf{C}_m\}_{m=1}^M\right\}$, is a collection of the GMM parameters, and $N(\mathbf{x}; \boldsymbol{\mu}, \mathbf{C})$ denotes the Gaussian density function with argument $\mathbf{x}$, expectation $\boldsymbol{\mu}$, and covariance matrix $\mathbf{C}$. The moments and cumulants of the vector $\mathbf{x}$ with the above GMM pdf depend on the vector $\boldsymbol{\theta}$, and thus, they will be denoted by $m_{\mathbf{y},\boldsymbol{\theta}}(\cdot)$, $c_{\mathbf{y},\boldsymbol{\theta}}(\cdot)$, respectively.

The cumulants of the process can be obtained from the moments via (4). By substitution of (7) into (6), the required moments can be expressed as

$$\begin{aligned} m_{\mathbf{y},\boldsymbol{\theta}}(I_p) &= \int_{\mathbf{x}} \prod_{i\in I_p} y_i \sum_{m=1}^{M} w_m N(\mathbf{x}; \boldsymbol{\mu}_m, \mathbf{C}_m) d\mathbf{x} \\ &= \sum_{m=1}^{M} w_m \cdot \int_{\mathbf{x}} \prod_{i\in I_p} y_i N(\mathbf{x}; \boldsymbol{\mu}_m, \mathbf{C}_m) d\mathbf{x} \\ &= \sum_{m=1}^{M} w_m E_m \left\{ \prod_{i\in I_p} y_i \right\} \\ & \forall I_p \in \alpha,\ \forall \alpha \in A \end{aligned} \qquad (8)$$

in which E_m denotes the statistical expectation under the Gaussian pdf $N(\boldsymbol{\mu}_m, \mathbf{C}_m)$. Thus, the problem decreases to evaluating separately the joint moments of Gaussian distributed vectors. For this purpose, the following theorem is applied.

Theorem 1 [4]: Let $\mathbf{y} = [y_1, \ldots, y_k]^T$ be a zero-mean, Gaussian vector with a known covariance matrix $\mathbf{C}$ possibly singular, I denotes the set of indices of the component of $\mathbf{y}$, then

$$E\left\{\prod_{i=1}^{k} y_i\right\} = \begin{cases} 0 & \text{if } k \text{ is odd} \\ \sum\limits_{P,Q} \prod\limits_{p=1}^{k/2} c_{i_p, j_p} & \text{if } k \text{ is even} \end{cases} , \qquad (9)$$

where the sum runs over all decompositions of I into disjoint subsets $P = \{i_p\}_{p=1}^{k/2}$, $Q = \{j_p\}_{p=1}^{k/2}$ such that $i_1 < \cdots < i_{k/2}$, and $i_p < j_p$ for every p.

In the case where $\mathbf{y}$ has a non-zero mean, $\boldsymbol{\mu} = \{\mu_i\}_{i\in I}$, Theorem 1 can be applied to $\mathbf{y} - \boldsymbol{\mu}$, and $E\left\{\prod_{i=1}^k y_i\right\}$ can be evaluated by solving

$$E\left\{\prod_{l\in I} (y_l - \mu_l)\right\} = \begin{cases} 0 & \text{if } k \text{ is odd} \\ \sum\limits_{P,Q} \prod\limits_{p=1}^{k/2} c_{i_p, j_p} & \text{if } k \text{ is even} \end{cases} . \qquad (10)$$

Expanding the left hand side of (10) yields

$$\begin{aligned} E\left\{\prod_{l\in I} y_l\right\} = & \sum_{D,B} \prod_{j\in B} (-1)^{\nu(B)+1} \mu_j E\left\{\prod_{i\in D} y_i\right\} \\ & + \begin{cases} 0 & \text{if } k \text{ is odd} \\ \sum\limits_{P,Q} \prod\limits_{p=1}^{k/2} c_{i_p, j_p} & \text{if } k \text{ is even} \end{cases} \end{aligned} \qquad (11)$$

in which the first sum in the right hand side of (11) runs over all partitions of I into disjoint subsets D and B, whose union is equal to I and $B \neq \phi$, and ν is the counting measure. The second sum runs over all P and Q defined in Theorem 1. Hence, evaluation of $E\left\{\prod_{l\in I} y_l\right\}$ via (11) requires the pre-evaluation of $E\left\{\prod_{l\in D} y_l\right\}$ for every $D \subset I$ using (11), resulting in a recursive process.

The expectations $E_m\left\{\prod_{i\in I_p} y_i\right\}$, $\forall I_p \in \alpha$, $\forall \alpha \in A$, $\forall m \in \{1, \ldots, M\}$ from (8) can be obtained using Theorem 1 and its extension in (11). Then the cumulants of the GMM process, $c_{\mathbf{y},\boldsymbol{\theta}}(I)$, can be calculated using (4).

Using the invariance property, the ML estimator of the cumulants is given by

$$\hat{c}_{\mathbf{y},\theta}(I)_{ML} = c_{\mathbf{y},\hat{\boldsymbol{\theta}}_{ML}}(I) \qquad (12)$$

where $\hat{\boldsymbol{\theta}}_{ML}$ is the ML estimate of the GMM parameters

$$\hat{\boldsymbol{\theta}}_{ML} = \arg\max_{\boldsymbol{\theta}} \sum_{n=1}^{N} \log \sum_{m=1}^{M} w_m N(\mathbf{x}_n; \boldsymbol{\mu}_m, \mathbf{C}_m) \ . \qquad (13)$$

The above ML estimator involves a nonlinear optimization. The most common method for performing this optimization is the EM algorithm [5]. Under the i.i.d. condition, the ML estimator of $\hat{c}_{\mathbf{y},\boldsymbol{\theta}_{ML}}(I)$ converges in probability to $c_{\mathbf{y},\boldsymbol{\theta}}(I)$ [6].

4. MODEL-BASED ESTIMATION OF CUMULANTS OF A STATIONARY PROCESS

Consider an observed process $\{x_n\}_{n=1}^N$, which is strict sense stationary with marginal pdf $f_x(\cdot)$. We assume that $\{x_n\}_{n=1}^N$ is an N_0-dependent process, that is, x_n and $x_{n+\tau}$ are independent random variables $\forall|\tau| > N_0 \geq 0$. Let $\{n_l\}_{l=1}^L$ be a sequence of integers such that $n_l \in \{1, \ldots, N\}$, and $n_l > n_m$ for every $l > m$. Then, due to the stationarity of the process, it follows that

$$\text{cum}(x_{n_1}^{r_1}, \ldots, x_{n_L}^{r_L}) = c_x^k(\tau_1, \ldots, \tau_{L-1}, r_1, \ldots, r_L) \qquad (14)$$

where $\tau_l = n_{l+1} - n_l$, $l = 1, \ldots, L-1$ and $\{r_l\}_{l=1}^{L}$ was defined in Section 2. We organize the data, $\{x_n\}_{n=1}^{N}$, into vectors $\mathbf{z}_n = [x_n, x_{n+1}, \ldots, x_{n+d}]^T$ to obtain a sequence of random vectors $\{\mathbf{z}_n\}_{n=1}^{N'}$, where $N' = N - d$ and $d = \tau_{L-1}$ is the maximum lag of $c_x^k(\tau_1, \ldots, \tau_{L-1}, r_1, \ldots, r_L)$. Since the vector process $\{\mathbf{z}_n\}_{n=1}^{N'}$ is stationary, then

$$\mathrm{cum}(z_{nn_1}^{r_1}, \ldots, z_{nn_L}^{r_L}) = c_x^k(\tau_1, \ldots, \tau_{L-1}, r_1, \ldots, r_L) \quad (15)$$

where z_{ni} denotes the ith component of $\mathbf{z}_n$. Hence, the problem of estimating the cumulants of $\{x_n\}_{n=1}^{N}$ is equivalent to estimating the cumulants of the random vector process $\{\mathbf{z}_n\}_{n=1}^{N'}$. Note that the vectors $\mathbf{z}_n$ are not independent. In similar to the previous section, we define the sequence $\{\mathbf{y}_n\}_{n=1}^{N'}$ by $\mathbf{y}_n = \mathbf{H}_{h,I,J}(\mathbf{z}_n)$, where the transformation $\mathbf{H}_{h,I,J}$ is defined in Section 2 with $I = \{1, 2, \ldots, k\}$, $J = \{1, 2, \ldots, d\}$, and the mapping function $h : I \longrightarrow J$ in this case is set to be

$$\begin{aligned} &h[1] = \cdots = h[r_1] = 1 , \\ &h[r_1+1] = \cdots = h[r_1+r_2] = 1+\tau_1 , \\ &h[r_1+r_2+1] = \cdots = h[r_1+r_2+r_3] = 1+\tau_2 , \\ &\vdots \\ &h[r_1+\cdots+r_{L-1}+1] = \cdots = h[r_1+\cdots+r_L] = 1+\tau_{L-1} . \end{aligned} \quad (16)$$

Using this transformation, the cumulant $c_x^k(\tau_1, \ldots, \tau_{L-1}, r_1, \ldots, r_L)$ can be expressed in terms of the simple cumulant $c_{\mathbf{y}}(I)$:

$$c_{\mathbf{y}}(I) = \mathrm{cum}(y_{n1}, \ldots, y_{nk}) = c_x^k(\tau_1, \ldots, \tau_{L-1}, r_1, \ldots, r_L) \quad (17)$$

where y_{ni} denotes the ith the component of $\mathbf{y}_n$.

The marginal density function of the vectors $\{\mathbf{z}_n\}_{n=1}^{N'}$ is modeled by Gaussian mixture: $f_{\mathbf{z}}(\mathbf{z}_n; \boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the vector of GMM parameters. As mentioned above, $\{\mathbf{z}_n\}_{n=1}^{N'}$ is not an independent random sequence. Nevertheless, for the sake of simplicity, we assume that it is an independent sequence, and estimate $\boldsymbol{\theta}$ by

$$\hat{\boldsymbol{\theta}}_{\mathrm{MARG}} = \arg\max_{\boldsymbol{\theta}} \sum_{n=1}^{N'} \log f_{\mathbf{z}}(\mathbf{z}_n; \boldsymbol{\theta}) . \quad (18)$$

The subscript MARG denotes that the estimation is based on the marginal distribution. $\hat{\boldsymbol{\theta}}_{\mathrm{MARG}}$ would be the ML estimator if $\{\mathbf{z}_n\}_{n=1}^{N'}$ were an independent sequence. The ML estimator of cumulants of an i.i.d. random sequence with Gaussian mixture density was presented in the previous section, and can be implemented here. In Appendix I, it is shown that $\hat{\boldsymbol{\theta}}_{\mathrm{MARG}}$, obtained under the independence assumption, is consistent even when the independence assumption is not satisfied. Finally, the cumulants of the scalar process x_n can be obtained from the cumulants of the vector process $\mathbf{z}_n$ according to (15).

5. SIMULATION RESULTS

The performance of the proposed model-based technique was examined in terms of root-mean-square error (RMSE) and compared to sample estimation via simulations. Three examples are presented here. In the first two examples, the cumulants of an i.i.d. random vector were estimated, while in the last example, the cumulants of a stationary random process are estimated.

In the first example, an i.i.d. random vector of length $d = 4$ with 3rd-order GMM distribution was generated with mixture parameters $w_1 = 0.7, w_2 = 0.2, w_3 = 0.1$, $\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \boldsymbol{\mu}_3 = [0, 0, 0, 0]^T$, and full rank, non-diagonal covariance matrices. The cumulant of interest is $\mathrm{cum}\left(x_1^2, x_3^2\right)$. The real value of the cumulant was calculated theoretically and found to be 4.94. 200 experiments were carried out for evaluation of the RMSE in the cumulant estimation which is presented in Fig. 1 as a function of the data length, N. In the second example, the performance of the proposed method, applied to non-GMM signals, is evaluated. For this purpose, $d = 4$ zero-mean, i.i.d. Laplace distributed sources with unit variance were mixed by a linear 4×4 mixing matrix, which was set to be proportional to an upper-triangle Toeplitz matrix with first row equal to $[1, 2, 3, 4]$. The cumulant of interest is $\mathrm{cum}(x_1, x_2, x_3, x_4)$. The real value of the cumulant was calculated theoretically and found to be 0.12. Fig. 2 shows that the algorithm performs better than sample estimation also for non-GMM signals.

For the third example, a 3rd-order GMM density function was generated and passed through a linear time-invariant moving average (MA) system. The MA coefficients were set to be $b_0 = 1, b_1 = 2.9, b_2 = -2.3$. The mixture parameters of the excitation are $w_1 = 0.7, w_2 = 0.2, w_3 = 0.1$, $\mu_1 = \mu_2 = \mu_3 = 0$, and $\sigma_1^2 = 1; \sigma_2^2 = 0.1; \sigma_3^2 = 0.7$. The cumulant of interest is $c_x^4(\tau_1 = 1, \tau_2 = 2, r_1 = 1, r_2 = 1, r_3 = 2,)$. The real value of the cumulant was calculated theoretically and found to be 5.84. 500 experiments were carried out for evaluation of the RMSE in the cumulant estimation which is presented in Fig. 3 as a function of the signal length.

Appendix I

Proof of consistency of $\hat{\boldsymbol{\theta}}_{\mathrm{MARG}}$ from (18)

The sequence $\{\mathbf{z}_n\}_{n=1}^{N'}$ is assumed to be N_0+d-dependent. Therefore, the sequence

$$\psi_n \triangleq \log f_{\mathbf{z}}(\mathbf{z}_n; \boldsymbol{\theta}) - E\{\log f_{\mathbf{z}}(\mathbf{z}_n; \boldsymbol{\theta})\} \quad n = 1, \ldots, N'$$

is also $(N_0 + d)$-dependent. We show that the average of a zero-mean, $(N_0 + d)$-dependent process converges in probability to zero:

$$\bar{\psi}_{N'} \triangleq \frac{1}{N'} \left(\sum_{n=1}^{N'} \psi_n \right) \xrightarrow{p} 0 . \quad (19)$$

For this purpose the following proposition is used.

Proposition A1: Let $\{\alpha_n\}_{n=1}^{\infty}$ and $\{\beta_n\}_{n=1}^{\infty}$ be sequences of random variables on the same probability space such that $\beta_n \xrightarrow{p} \mu$ and $(\alpha_n - \beta_n) \xrightarrow{p} 0$. Then $\alpha_n \xrightarrow{p} \mu$ where μ is assumed to be deterministic [7].

For some integer j such that $j > 2(N_0+d)$ let

$$
\begin{aligned}
\beta_{N'j} &= \frac{1}{N'}\Bigg(\underbrace{\left(\psi_1+\cdots+\psi_{j-(N_0+d)}\right)}_{\phi_0} \\
&\quad + \underbrace{\left(\psi_{j+1}+\cdots+\psi_{2j-(N_0+d)}\right)}_{\phi_1} \\
&\quad + \cdots + \underbrace{\left(\psi_{(r-1)j+1}+\cdots+\psi_{rj-(N_0+d)}\right)}_{\phi_{r-1}}\Bigg) \\
&= \frac{r}{N'}\cdot\frac{1}{r}\sum_{k=0}^{r-1}\phi_k \ ,
\end{aligned} \tag{20}
$$

where $r = \lfloor N'/j \rfloor$ is the integer part of N'/j. Note that $\beta_{N'j}$ is proportional to sum of zero-mean i.i.d. random variables. Applying the weak law of large numbers to $\beta_{N'j}$ leads to $\beta_{N'j} \xrightarrow{p} 0$ as $N' \to \infty$. It remains to show that $\bar{\psi}_{N'} - \beta_{N'j} \xrightarrow{p} 0$ as $N' \to \infty$ and due to proposition A1 it follows that $\bar{\psi}_{N'} \xrightarrow{p} 0$. For this purpose, we examine the sequence

$$
\begin{aligned}
\bar{\psi}_{N'} - \beta_{N'j} = \frac{r}{N'}\Bigg(&\frac{1}{r}\sum_{l=1}^{r-1}\left(\psi_{lj-N_0+d+1}+\cdots+\psi_{lj}\right) \\
&+\frac{1}{r}\left(\psi_{rj-N_0+d+1}+\cdots+\psi_{N'}\right)\Bigg)
\end{aligned} \tag{21}
$$

which is also a sum of i.i.d. random variables. Hence, by applying the law of large numbers, proposition A1 is satisfied. It therefore follows that

$$
\frac{1}{N'}\sum_{n=1}^{N'}\log\left(f_{\mathbf{z}}(\mathbf{z}_n;\boldsymbol{\theta})\right) \xrightarrow{p} \int_{\mathbf{z}}\log\left(f_{\mathbf{z}}(\mathbf{z};\boldsymbol{\theta})\right) f_{\mathbf{z}}(\mathbf{z};\boldsymbol{\theta}_0)d\mathbf{z},\ \forall\boldsymbol{\theta} \tag{22}
$$

where $\boldsymbol{\theta}_0$ denotes the true value of the parameter $\boldsymbol{\theta}$. The right hand side of (22) is maximized for $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. Thus, as $N' \to \infty$, the maximum of left hand side of (22) is obtained by $\hat{\boldsymbol{\theta}} \to \boldsymbol{\theta}_0$.

REFERENCES

[1] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," *Proc. of the IEEE*, vol. 79, no.3, Mar. 1991.

[2] J. F. Cardoso, "Blind signal separation: statistical principles," *Proc. of the IEEE*, vol. 86, no. 10, pp. 2009-2025, 1998.

[3] J. Q. Li and A. R. Barron, "Mixture density estimation," *Advances in Neural Information Processing Systems 12*, The MIT press, 2002.

[4] K. Triantafyllopoulos, "On the central moments of the multidimensional Gaussian distribution," *The Mathematical Scientist*, vol. 28, pp. 125-128, Apr. 2003.

[5] D. A. Reynolds and R. C. Rose, "Robust text-independent speaker identification using Gaussian mixture speaker models," *IEEE Trans. on Speech and Audio Processing*, vol. 3, no. 1, pp. 72-83, Jan. 1995.

[6] E. J. Dudewicz, *Introduction to probability and statistics,* New York, NY Holt, Rinehart and Winston, 1976.

[7] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods*, New York, NY Springer-Verlag 1987.

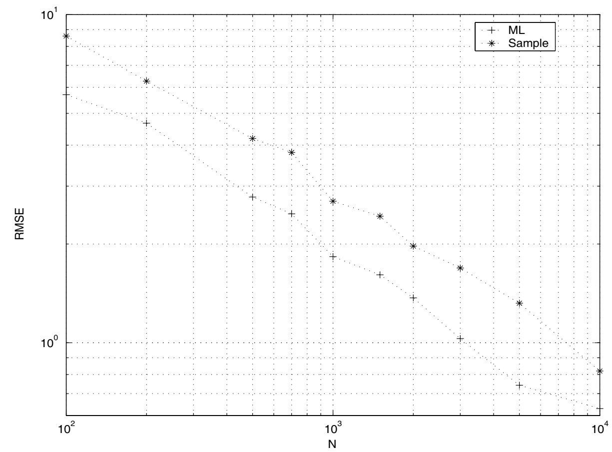

Figure 1: Cumulant estimation RMSE of a GMM distributed random vector process.

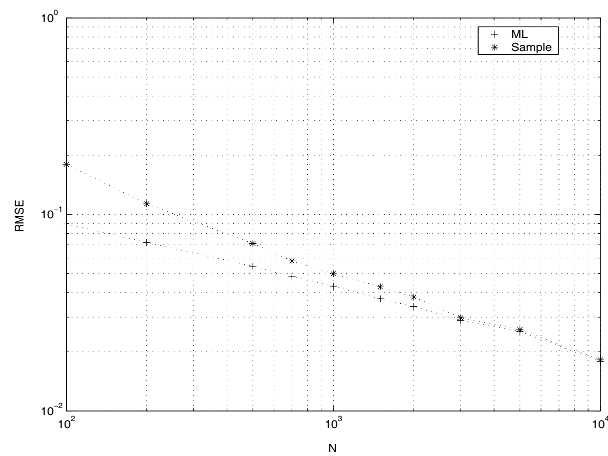

Figure 2: Cumulant estimation RMSE of a linearly mixed Laplace i.i.d. random vector process.

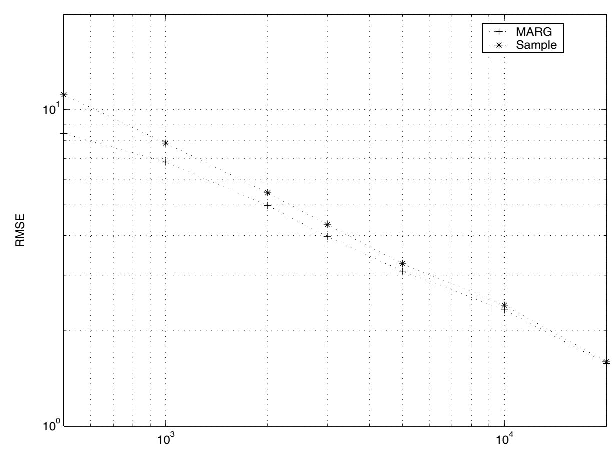

Figure 3: Cumulant estimation RMSE of an N_0-dependent random process.