USE OF SELECTED HOS INFORMATION FOR LOW-VARIANCE ESTIMATION OF BANDLIMITED SYSTEMS WITH SHORT DATA RECORDS

Haralambos Pozidis and Athina P. Petropulu

Electrical and Computer Engineering Department
Drexel University, Philadelphia, PA 19104
Tel. (215) 895-2358 Fax. (215) 895-1695
pozidis@cbis.ece.drexel.edu, athina@artemis.ece.drexel.edu

ABSTRACT

Although reconstruction of a nonminimum-phase system excited by a stationary non-Gaussian white input is only possible using higher-order statistics (HOS) of the system output, there has been a lot of criticism in the literature against the amount of data required for keeping estimation errors low, and the complexity involved. Recently several attempts for reducing the variance of the HOS estimates have appeared. In the case of bandlimited signals, we have demonstrated via simulations that the estimation variance can be reduced if "good" slices, instead of the whole bispectrum, are used. This suggests a potential reduction of variance in the system estimates, without having to resort to long observations. In this paper we justify theoretically the dependence of the system estimate variance on the bispectrum slice, and the criterion of slice selection. We also present simulation results, where the selected-slices approach appears to result in much lower estimation variance, as compared to other entire-bispectrum based approaches, for data lengths as low as 64 samples.

1. INTRODUCTION

Higher-order spectra (HOS) have been applied successfully to the problem of system reconstruction, mainly because of their ability to preserve the true system phase, and their robustness to additive Gaussian noise of unknown covariance. However, there has been a lot of criticism in the literature against the amount of data required for keeping estimation errors low, and the complexity involved. Recently several attempts for reducing the variance of the HOS estimates appeared, such as using low rank approximations of HOS estimators [1]. In the case of bandlimited signals, we have demonstrated via simulations [6], that the estimation variance can be reduced if "good" slices instead of the whole bispectrum are used. For such signals, there are regions where, in theory, the bispectrum is identically zero. In practice, however, the bispectrum estimate will contain regions where the useful signal information will be low and corrupted by errors. Avoiding such regions can result in improved estimation. Methods that use fixed bispectrum slices [4], [3] cannot be applied to the reconstruction of bandlimited systems, since the ideal bispectrum along these slices can be zero. Moreover, in the presence of noise and finite data records, bispectrum estimates along fixed slices, such as the axes and the main diagonal, can exhibit high estimation variance, and since single slices are used, there is no averaging mechanism to reduce estimation errors.

It was shown in [6] that unique identification of an *arbitrary* system can be performed, based on *any two horizontal slices* of the output discretized n-th order spectrum, $n \geq 3$, of the system, as long as the distance between the slices and the grid size satisfy a simple condition. It was observed in [6] that the use of slices selected according to a certain criterion instead of the whole higher-order spectrum, could lead to reduced estimation variance. This suggests that we could potentially be able to reduce variance in the system estimates without resorting to long observations. In this paper we present the theoretical justification of both the dependence of the variance of the system estimates on the HOS slice used, and the usage of the criterion for slice selection, proposed in [6]. We also show, based on simulation examples, that even in the case of very short output sequence records, the selected-slices approach results in much lower estimation variance, as compared to other approaches that use the entire bispectrum for estimation, such as the methods of [2], and [7].

2. RECONSTRUCTION FROM ANY PAIR OF HORIZONTAL HOS SLICES OF THE SYSTEM OUTPUT

Consider a stationary process $x(n)$ given by:

$$x(n) = e(n) * h(n) + w(n), \tag{1}$$

where $e(n)$ is an i.i.d. non-Gaussian process with zero mean and finite n-th order cumulant $\gamma_n^e \neq 0$, for $n \geq 2$; $w(n)$ is a stationary zero-mean Gaussian process of unknown covariance which is assumed independent of $e(n)$; $h(n)$ is the unknown impulse response of a generally mixed-phase, *complex* LTI system. It is assumed that $h(n)$ does not have zeros on the unit circle, however this assumption can be relaxed.

In this paper we consider reconstruction from third-order spectra. A generalization of the results to the n-th order spectra case can be found in [5]. The frequency-domain bispectrum of $x(n)$ is given by

$$C_3^x(\omega_1, \omega_2) = \gamma_3^e H(\omega_1) H(\omega_2) H(-\omega_1 - \omega_2), \tag{2}$$

with $H(\omega)$ denoting the frequency response of the system. The following proposition holds:

Proposition 1 *[5] For the process $x(n)$ described by (1), $h(n)$ is always identifiable, within a (complex) constant and a circular shift, from any two slices of the discretized output bispectrum, i.e, $C_3^x(\frac{2\pi}{N}k, \frac{2\pi}{N}l_1)$ and $C_3^x(\frac{2\pi}{N}k, \frac{2\pi}{N}l_2)$, $k = 0, \ldots, N-1$, if and only if N and $r = |l_1 - l_2|$ are coprime integers. If $h(n)$ is real, then it is identifiable, within a constant and a circular shift, based on a single slice of the discretized output bispectrum, i.e., $C_3^x(\frac{2\pi}{N}k, \frac{2\pi}{N}l)$, if and only if N and $r = 2l$ are coprime integers.*

The proof of this proposition can be found in [5].

By evaluating (2) at discrete frequencies $\omega = \frac{2\pi}{N}k$, $k \in [0, \ldots, N-1]$, we obtain the discrete bispectrum of $x(n)$, denoted $C_3^x(k,l)$. Let $\mathbf{h}_l = [\log H(1), \ldots, \log H(N-1)]^T$ be the $(N-1) \times 1$ vector of the unknown samples of the logarithm of the frequency response of the system (we set $H(0) = 1$ arbitrarily, thus reconstructing the system within a complex constant). Then $\mathbf{h}_l$ can be obtained as the solution to the system

$$\mathbf{A}\mathbf{h}_l = \mathbf{c}, \tag{3}$$

where $\mathbf{c}$ is a $(N-1) \times 1$ vector of bispectrum values along the slices $(:,l)$ and $(:,l+r)$, with

$$\mathbf{c}_i = \log C_3^x(-i-r-l, l) - \log C_3^x(-i-r-l, l+r) + c_{l,r} \tag{4}$$

$(i = 0, 1, \ldots, N-2)$, and

$$c_{l,r} = \frac{1}{N}\sum_{k=0}^{N-1}[\log C_3^x(k, l+r) - \log C_3^x(k,l)] \tag{5}$$

Matrix $\mathbf{A}$ is a $(N-1) \times (N-1)$ sparse matrix with special structure: its first row contains a one at the r-th column and zeros elsewhere. Its k-th row contains a (-1) at column $(k-1)$modulo$N$, and a one at column $(k+r-1)$modulo$N$.

It can be proved, [5], that *matrix $\mathbf{A}$ is nonsingular if and only if N and r are coprime integers.* Using this method, the logarithm of two bispectrum slices, $(:,l)$ and $(:,l+r)$, is used to recover the impulse response $h(n)$ of the system. Although the bispectrum phase appears implicitly in the expressions, only the principal argument is actually needed.

Since we are dealing with bandlimited systems, we should consider the case of $h(n)$ having zeros on the unit circle. This means that $H(k) = 0$ for some k, and in turn $C_3^x(m,l) = 0$ for some (m,l), the logarithm of which is undefined. However, we can change the spacing between samples, or equivalently re-estimate the bispectrum in a different grid of frequency points to surpass that problem.

By using different pairs of bispectrum slices that satisfy the condition of Proposition 1, we can average the reconstructed systems in the time-domain (after scaling and shifting them appropriately), thus reducing the effects of noise and finite data lengths in the estimation of cumulants.

Significant computational savings result when the input process $e(n)$ in (1) is cyclostationary. To see that, note that we can write:

$$\begin{aligned} X(l)X(k)X(-k-l) &= \mathcal{F}\{X(l)x(n) * x(-n)e^{-j\frac{2\pi}{N}ln}\} \\ &= \mathcal{F}\{NX(l)R_x^{l/N}(n)\} \end{aligned} \tag{6}$$

where $k = 0, \ldots, N-1$, $R_x^{l/N}(n)$ is the well-known ambiguity function, $\mathcal{F}$ denotes DFT of size N, and $x(n)$ is as in (1). The ambiguity function of a cyclostationary process is a delta function (this is not generally true for stationary processes); thus, if $e(n)$ is cyclostationary, then (6) reduces to a scaled version of the discrete bispectrum of the system $h(n)$ along slice $(:,l)$. In that case, the complexity for the bispectrum computation can be greatly reduced, since (6) involves only second order operations on the data.

3. THEORETICAL JUSTIFICATION OF IMPROVEMENT OF ESTIMATES BASED ON "GOOD" SLICES

The asymptotic covariance matrix of the system estimates, obtained via (3) is given by [5]:

$$\Sigma_{\mathbf{h}} \approx (\mathbf{e}J_H\mathbf{A}^{-1}B)\Sigma_{\mathbf{c}_3^x}(\mathbf{e}J_H\mathbf{A}^{-1}B)^T, \tag{7}$$

where

$$(\mathbf{e})_{kn} = \frac{1}{N}\exp\{j\frac{2\pi}{N}kn\}, \quad B = [I_{N-1} \vdots - I_{N-1}] \tag{8}$$

$$J_H = \mathrm{diag}[H(1), \ldots, H(N-1)], \tag{9}$$

$\mathbf{A}$ is the system matrix, and $\Sigma_{\mathbf{c}_3^x}$ is the covariance matrix of the $(2N-2) \times 1$ random vector $\mathbf{c}_3^x$ defined as:

$$\begin{aligned} \mathbf{c}_3^x = [&\log C_3^x(-r-l,l), \cdots, \log C_3^x(-N+2-r-l,l), \\ &\log C_3^x(-r-l,l+r), \cdots, \log C_3^x(-N+2-r-l,l+r)] \end{aligned} \tag{10}$$

The covariance matrix $\Sigma_{\mathbf{c}_3^x}$ is given by [5]

$$\Sigma_{\mathbf{c}_3^x} \approx D\Sigma_{C_3^x}D^T, \tag{11}$$

where

$$D = \mathrm{diag}\left[\frac{1}{C_3^x(m_1)}, \ldots, \frac{1}{C_3^x(m_{2N-2})}\right], \tag{12}$$

$m_i = (k_i, l_i)$, $i = 1, \ldots, 2N-2$, ($k_{1\ldots N-1} = k_{N\ldots 2N-2} = -r-l, \ldots, -N+2-r-l$, $l_{1\ldots N-1} = l$, $l_{N\ldots 2N-2} = l+r$) are pairs of discrete frequencies along slices l, $l+r$, and $\Sigma_{C_3^x}$ is the covariance matrix of the bispectrum estimates along these slices. The elements of $\Sigma_{C_3^x}$, as $N \to \infty$, equal [8]:

$$\begin{aligned} &\mathrm{cov}[\hat{C}_3^x(f_1,f_2), \hat{C}_3^x(f_3,f_4)] = \\ &\tfrac{1}{NB_N^2}[C_2^x(f_1)C_2^x(f_2)C_2^x(f_1+f_2)C_2^x(f_3)C_2^x(f_4)C_2^x(f_3+f_4)]^{1/2} \\ &\cdot\{w_1\delta(f_2)\delta(f_4)[1+2\delta(f_1)][1+2\delta(f_3)] \\ &+w_2\delta(f_1-f_3)\delta(f_2-f_4)[1+\delta(f_1-f_4)+4\delta(f_1)\delta(f_2)]\} \end{aligned} \tag{13}$$

where

$$w_1 = \left[\sum_{u=-N}^{N} w(0,u)\right]^2; \; w_2 = \left[\sum_{u_1=-N}^{N}\sum_{u_2=-N}^{N} w^2(u_1,u_2)\right]^2, \tag{14}$$

$w(\nu_1,\nu_2)$ is a bispectral window [8], B_N is the bispectrum bandwidth, which for $N \to \infty$ satisfies $B_N \to 0$ and $B_N^2 N \to \infty$, and $C_2^x(f_i)$ is the power spectrum of the process $x(n)$ evaluated at the discrete frequency f_i.

Let us assume that the variance of the bispectrum estimates does not vary significantly from slice to slice (except for the power spectrum slices and the diagonal slice). This assumption is reasonable since, as shown by (13), the variance is inversely proportional to the record length. Then, from (7), (11) and (12) it can be seen that the variance of the system estimates will be inversely proportional to the bispectrum values along the slices $(:, l)$ and $(:, l + r)$; thus, usage of slices along which the bispectrum amplitude is low, will result in large estimation variance, and the opposite.

This theoretical result justifies the experimentally derived slice selection criterion proposed in [6], which was a measure of how "high" the bispectrum amplitude along these slices is, and was referred to as the "frequency content", i.e.,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \mid C_3^x(\omega, l) \mid d\omega$$

This quantity is actually the average area under the specific slice, and it qualifies as a measure of the bispectrum amplitude along this slice. In the case of bandlimited systems, there exist areas in the bispectrum where the amplitude is very low, as compared with areas corresponding to the passband of the system. Therefore, the frequency content will vary significantly along the slices of the bispectrum of a bandlimited system. By using slices only with high frequency content in the reconstruction procedure outlined above, we are most likely to obtain better system estimates, than by using the entire non-redundant bispectrum. The validity and usefulness of the "frequency content" criterion is demonstrated in the following section.

4. RESULTS

In this Section we compare the selected slices-based method to the methods of [2] (BLW) and [7] (RG) for the reconstruction of a bandlimited system. We focus on the case of very short data records, where the HOS estimates should, in theory, exhibit considerable variance. We show that, by using selected slices of the output spectra, it is possible to obtain satisfactory system estimates, even in that case. Moreover, use of slices that do not satisfy the criterion of goodness, leads to deterioration of performance, in terms of mean-square-error.

Although the system considered was real, we used two slices for the reconstruction procedure instead of one (see Proposition 1), since an FFT size of $N = 64$, a power of 2, was used, to speed up computations. The reconstruction procedure was repeated using several pairs of slices, and the estimated systems were averaged in the time-domain.

We used a highpass system with transfer function

$$H(z) = \frac{(1 - [0.1 \pm j0.8]z)(1 - [0.9999 \pm j0.0156]z^{-1})}{(1 - 0.4z^{-1})^{-1}(1 - 0.3z^{-1})}$$

and estimated an array of 8×8 third-order cumulant lags.

In order to select the slices for better reconstruction, we run 100 simulations and computed the frequency content of each slice at each run. The average frequency content over all runs is shown in Fig. 1 where the shaded area indicates standard deviation. Slices 20-32 clearly exhibit a consistently higher frequency content than all others.

Then we run 100 simulations of the selected slices-based method, using averaging over slices 24-31, and the entire bispectrum based methods BLW and RG. The results are shown in Fig. 2 (a)-(c) and (d)-(f) for SNR of ∞ and 10 dB respectively. It should be noted that only 64 symbols were used for the bispectrum estimation. It can be seen that both methods are outperformed by the proposed one. This can be attributed to the fact that the actual system $h(n)$ is highpass, therefore its output bispectrum contains regions of low magnitude. The inclusion of such regions in the reconstruction procedure is responsible for poor performance. On the other hand, selection of regions with higher signal information only, as in the proposed method, leads to better results.

A comparison of all methods in terms of mean-square-error (MSE) was also conducted, and the results are shown in Figs. 3 and 4, for SNR equal to ∞ and 10 dB, respectively, and record length varying from 64 to 2048 samples. To illustrate the advantage of using only slices with high frequency content for reconstruction, the proposed method was implemented with averaging over slices 24-31 and also over all possible slices, 0-32. As was expected from the previous results, the proposed method exhibits a significantly lower MSE than the BLW and RG methods. What is interesting to note though, is that the use of certain "good" slices only, as opposed to using the whole bispectrum, in the proposed method, seems to produce results with lower MSE. This difference in MSE decreases with decreasing SNR, and increases with increasing record length.

It was stated previously that by using certain "good" bispectrum slices only, instead of the whole bispectrum, we could potentially reduce the system estimation variance without resorting to longer observations. This becomes evident by a closer examination of the results in Figs. 3 and 4. For example, in Fig. 3, if using the whole bispectrum we would have to use 512 data samples to achieve the same MSE that we could achieve with 64 samples if using slices 24-31 only; the same result holds for 1024 and 128 samples for the whole bispectrum and slices 24-31 respectively. At SNR=10 dB (Fig. 4), we would need 1024 (or 2048) samples with the whole bispectrum to get the same performance as with 256 (or 512) samples when using slices 24-31. The savings in computation time and complexity are obvious.

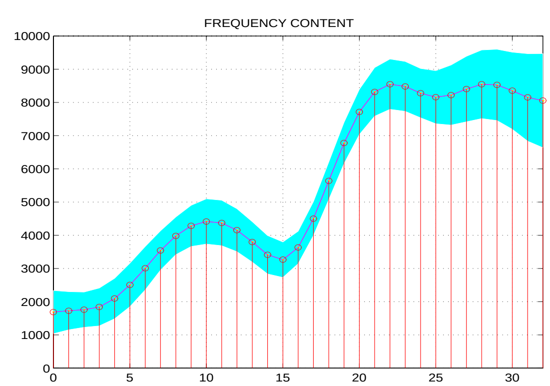

Figure 1: Frequency content for slices 0-32 of the output bispectrum of the system. Circles and solid line represent the average over 100 simulations, while shaded area indicates sample standard deviation.

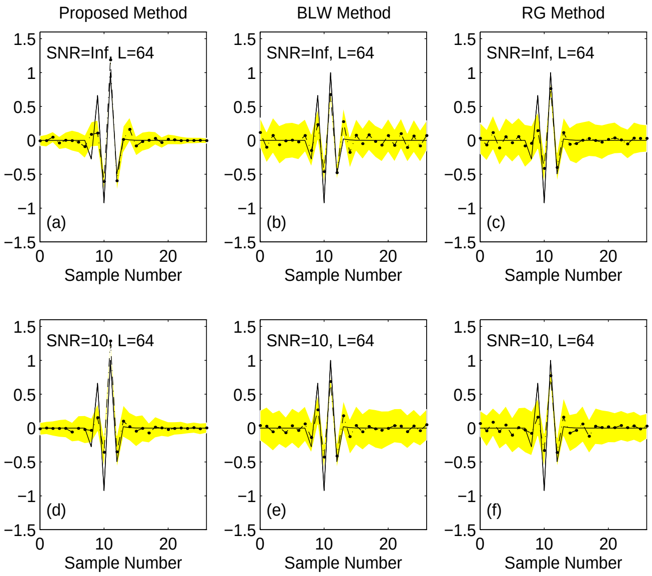

Figure 2: Comparison of the **selected slices-based, BLW and RG** methods for 64 output samples, and SNR=∞ dB, ((a)-(c)) and 10 dB ((d)-(f)). Actual system is in solid lines, the average over 100 estimates in dash-dotted lines, and shaded area indicates standard deviation.

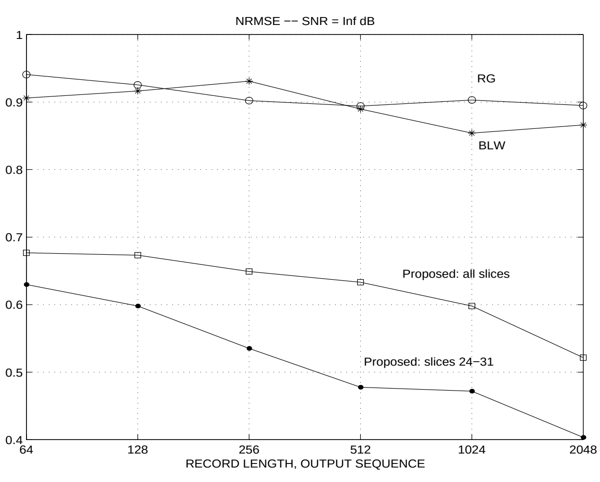

Figure 3: Comparison of the **selected slices-based (averaging over slices 24-31 and over all slices), BLW and RG** methods for 64-2048 output samples and SNR=∞ dB.

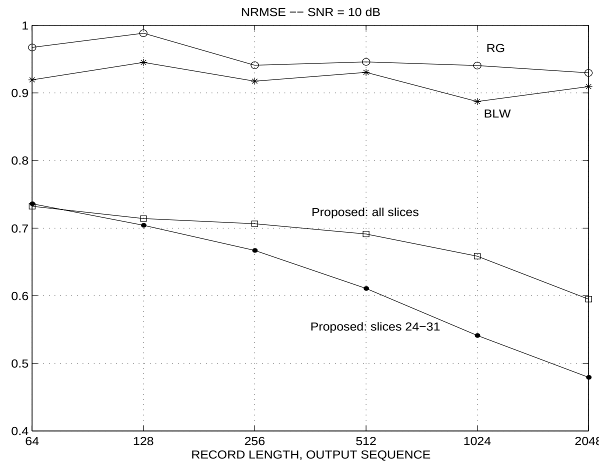

Figure 4: Comparison of the **selected slices-based (averaging over slices 24-31 and over all slices), BLW and RG** methods for 64-2048 output samples and SNR=10 dB.

5. REFERENCES

[1] T.F. Andre, R.D. Nowak, B.D. Van Veen, "Low Rank Estimation of Higher Order Statistics", *IEEE Trans. Signal Processing*, vol. 45, March 1997.

[2] H. Bartelt, A.W. Lohmann and B. Wirnitzer, "Phase and amplitude recovery from bispectra", *Applied Optics*, vol. 23, no. 18, pp. 3121-3129, Sept. 1984.

[3] S.A. Dianat and M.R. Raghuveer, "Fast Algorithms for Phase and Magnitude Reconstruction from Bispectra", *Opt. Eng.*, vol. 29, pp. 504-512, May 1990.

[4] K.S. Lii and M. Rosenblatt, "Deconvolution and Estimation of Transfer Function Phase and Coefficients for Nongaussian Linear Processes", *The Annals of Statistics*, vol. 10, pp. 1195-1208, 1982.

[5] H. Pozidis, and A.P. Petropulu, "System Reconstruction Based on Selected Regions of the Discretized HOS", *IEEE Trans. Sig. Proc.*, submitted in 1997.

[6] H. Pozidis, and A.P. Petropulu, "System Reconstruction From Selected HOS Regions", *IEEE Sig. Proc. Workshop on HOS*, July 1997, Banff, Alberta, Canada.

[7] M. Rangoussi and G.B. Giannakis, "FIR Modeling Using Log-Bispectra: Weighted Least-Squares Algorithms and Performance Analysis", *IEEE Trans. Circ. and Sys.*, vol. 38, no. 3, pp. 281-296, Mar. 1991.

[8] M. Rosenblatt and J.W. Van Ness, "Estimation of the Bispectrum", *Ann. Math. Statistics*, vol. 36, pp. 1120-1136, 1965.