

# Nested Sampling for Higher-Order Statistics with Application to LTE Channel Estimation

Qiong Wu<sup>†</sup> and Qilian Liang<sup>†‡</sup>

<sup>†</sup>Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX 76019-0016 USA

<sup>‡</sup>College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin 300387 China

Email: qiong.wu@mavs.uta.edu, liang@uta.edu

**Abstract**—This paper studies higher-order statistics based on nested sampling. We propose multilevel nested sampling (MNS) algorithm to obtain higher-order statistics (HOS), and analyze the computational complexity of the MNS-HOS algorithm for both parametric and nonparametric methods. Compared to the existing HOS algorithms, the proposed algorithm vastly reduces the complexity by several orders in terms of the length of segmentation window. We also apply MNS-HOS algorithm to estimate the coefficients of a simplified LTE spatial channel model blindly without using any training sequences. Our simulations show that compared with pairwise coprime sampling HOS algorithm, MNS-HOS produces less variance and converges faster in estimating higher-order cumulants, and achieves 17% performance gain for channel estimation. The proposed MNS-HOS algorithm is also able to reduce computational complexity by 98% with a trade-off of 22% performance loss in contrast with the HOS algorithm without sparse sampling.

## I. INTRODUCTION

Accurate channel estimation can be used in orthogonal frequency division multiplexing systems to improve the performance by allowing coherent demodulation, which possesses a 3-to-4 dB gain in signal-to-noise ratio (SNR) compared with differential modulation [1]. Besides, systems are able to take advantage of receiver diversity for optimum combining via channel estimators. In general, channel estimation can be conducted by either inserting pilot segments into OFDM subcarriers [2], or estimating channel only based on the received signal. But it is not efficient for repeated transmitting a known sequence to train the equalizer at the receiver, and it is difficult to establish data transmission over the channels suffered from unavoidable presence of multipath fading. Theoretically speaking, the blind channel estimation is preferable.

There are two kinds of approaches to implement the blind channel estimation. The linear methods are based on the fact that the cyclostationary characteristics of modulated signals permit the recovery of both amplitude and phase responses of a communication channel only using second-order statistics [3]. The nonlinear methods are based on the calculation of higher-order statistics (HOS), and could be further divided into direct or indirect approaches. The direct algorithm calculate HOS and their discrete Fourier transforms with further matrix manipulation [4]. Although its computation is more complex, the advantages include more straightforward for theoretical derivation, free from minimizing cost functions so as to avoid local minimum problem, and insensitivity to time jitter. On the other hand, one of the typical indirect algorithms is Busgang

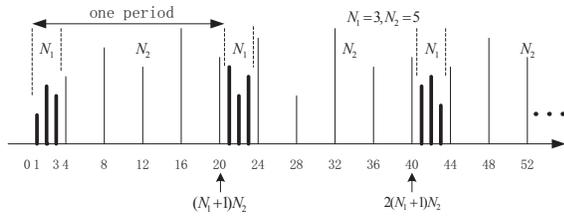
algorithm [5], in which the deconvoluted signal implicitly exploit the HOS via Busgang statistics. This kind of methods is simpler to implement and generally capable of delivering a good performance.

In general, HOS approaches are favorable to preserve both phase and amplitude information of the signal, and to deal with non-Gaussian sequences and nonlinear non-minimum phase system. Its estimates can be calculated from either conventional nonparametric methods of Fourier transform [6], or parametric methods based on moving average (MA), autoregressive (AR), or autoregressive moving average (ARMA) models [7]. The paper [6] provides further details about both kinds of methods. It is worth noting that both of them have high variance and require a large number of records to obtain smooth estimates [8], but increasing the number of segments is demanding on computation, and may increase bias and introduce non-stationarity. Although there are preliminary researches [7], [9] providing a sparse sampling scheme, pairwise coprime sequences (PCS), as a possible solution, how to accelerate speed of convergence and to leverage the trade-off between complexity and performance are still open questions.

This is our motivation to design an algorithm estimating HOS based on nested sampling. This concept is first introduced to conduct economic and sociological surveys [10]. Later researches extend it to array signal processing [11], astronomical data analysis [12], and estimating the directions of arrivals [13]–[15]. But nested sampling has never been considered to derive HOS in previous literatures.

The main purpose of this paper is to develop a HOS algorithm based on multilevel nested sampling (MNS). The direct merits of implementing nested sampling are to greatly reduce the computational expense and to produce posterior inferences at the same time. By using this sampling scheme, the proposed algorithm improves the PCS performances in terms of estimation variance and speed of convergence with the same delay and a even lower computational complexity.

In Section II, we discuss theoretical background about both higher-order statistics and nested sampling. Then, we describe the MNS-HOS algorithm, and analyze its complexity and complexity in Section III. The performance of the proposed algorithm is simulated in Section IV. Finally, we draw the conclusions in Section V.


 Fig. 1: Nested Sampling with  $N_1=3$ ,  $N_2=5$ 

## II. PRELIMINARIES

In this section, we provide an overview of higher-order statistics and nested sampling, and demonstrate their relationship based on the estimation of second-order moment.

Assuming  $x(t)$  is a zero-mean random process with at least  $k$ th-order stationary, the  $k$ th-order cumulant of this process is defined as the first  $k-1$  coefficients of Taylor series expansion of the data set. The expectations of second-order, third-order, and fourth-order cumulant are [6]

$$c_{2x}(\tau) = E[x(n)x(n+\tau)], \quad (1)$$

$$c_{3x}(\tau_1, \tau_2) = E[x(n)x(n+\tau_1)x(n+\tau_2)], \quad (2)$$

$$\begin{aligned} c_{4x}(\tau_1, \tau_2, \tau_3) &= E[x(n)x(n+\tau_1)x(n+\tau_2)x(n+\tau_3)] \\ &\quad - c_{2x}(\tau_1)c_{2,x}(\tau_2 - \tau_3) \\ &\quad - c_{2x}(\tau_2)c_{2,x}(\tau_3 - \tau_1) \\ &\quad - c_{2x}(\tau_3)c_{2,x}(\tau_1 - \tau_2). \end{aligned} \quad (3)$$

The  $k$ th-order spectrum is defined as the Fourier transform of its  $k$ th-order cumulant:

$$\begin{aligned} S_{kx}(\omega_1, \omega_2, \dots, \omega_{k-1}) &= \sum_{\tau_1=-\infty}^{+\infty} \dots \sum_{\tau_{k-1}=-\infty}^{+\infty} \\ &\quad c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) \cdot \exp\left(-j \sum_{i=1}^{k-1} \omega_i \tau_i\right). \end{aligned} \quad (4)$$

Specifically, the third-order spectrum is also called bispectrum, and the fourth-order spectrum is trispectrum.

Nested sampling is a non-uniform sampling scheme, using two different samplers in each of given period. Given sampling interval  $T_s$ , although the signal is sparsely and non-uniformly sampled at  $1 \leq l \leq N_1 T_s$  and  $(N_1+1)mT_s$ , for  $1 \leq m \leq N_2$ , in one sampling period, the autocorrelation  $c_{2x}(\tau)$  is able to be estimated at all lags  $\tau = kT_s$ , where  $k$ ,  $l$ , and  $m$  are integers.

In the simplest form, the nested array has two levels of sampling density, with the level 1 samples at the  $N_1$  locations and the level 2 samples at the  $N_2$  locations:

$$\begin{aligned} \mathcal{L}_{inner} &= \{mT_s, m = 1, 2, \dots, N_1\}, \\ \mathcal{L}_{outer} &= \{n(N_1+1)T_s, n = 1, 2, \dots, N_2\}. \end{aligned}$$

As a result, the sampling interval for the inner layer is  $t_1 = T_s$ , while the outer layer is  $t_2 = (N_1+1)T_s$ .

Figure 1 shows an example of periodic sparse sampling using nested sampling structure with  $N_1=3$  and  $N_2=5$ . The

cross-difference between two levels is

$$k = (N_1 + 1)m - l, \quad 1 \leq m \leq N_2, 1 \leq l \leq N_1, \quad (5)$$

and its range is given by

$$k \in [-(N_1 + 1)N_2 - 1, (N_1 + 1)N_2 - 1]. \quad (6)$$

Note that there are missing values symmetric through zero within the range in (6). The positive part includes  $(N_1 + 1), 2(N_1 + 1), \dots, (N_2 - 1)(N_1 + 1)$ . Take the Figure 1 as an example, the cross-difference will achieve the indexes

$$1, 2, 3, ( ), 5, 6, 7, ( ), 9, 10, 11, ( ), 13, 14, 15,$$

where the values with indexes in parentheses are left unknown. Then, we can use the self-difference among points from the second level to calculate for the missing indexes, which are 4, 8, and 12 in this example. The calculation is generalized as

$$(N_1 + 1)(m_1 - m_2), \quad 1 \leq m_1, m_2 \leq N_2. \quad (7)$$

With this method, we are able to calculate the second-order statistics without the constraint about the lowest permitted sampling rate.

## III. MNS-HOS ALGORITHM AND ITS COMPLEXITY ANALYSIS

In this Section, we propose the MNS-HOS algorithm and make the nested sampling according to the instantaneous variance of the input signal for better HOS estimation. Besides, we also demonstrate how this algorithm reduces the computational complexity by several orders. In the following content, MNS  $y_0(n), y_1(n), \dots, y_k(n)$  denote the sampled sequences from output signal related with the nested level  $N_0, N_1, \dots, N_k$ , respectively. In addition, we have the following assumptions to modeling the system:

- S.1** The driving signal  $\{w(t)\}$  is non-Gaussian, zero-mean, and independent and identically distributed (i.i.d.). Besides, it only has finite moments, which are estimated via expectations, and denoted as  $E[w^2(t)] = \sigma_w^2$ ,  $E[w^3(t)] = \gamma_{3w}$ , and  $E[w^6(t)] < +\infty$ .
- S.2** The measurement noise  $\{v(t)\}$  is assumed to be zero-mean, and to have either white or color Gaussian distribution. It is also i.i.d. and independent of  $\{w(t)\}$  with  $E[v^2(t)] = \sigma_v^2$ .
- S.3** The channel is regarded as a linear, time-invariant, and non-minimum phase moving average (MA) system, and its order  $q+1$  has known beforehand.

Based on these assumptions, we derive a procedure for estimating the covariances of the sampled cumulants, and model the output of the channel and the received signal as

$$x(n) = \sum_{i=0}^q h_i w(n-i), \quad (8)$$

$$y(n) = x(n) + v(n). \quad (9)$$

Then, we have

$$\begin{aligned}
 & E[\hat{c}_{ky}(s_1, \dots, s_k), \hat{c}_{ky}(t_1, \dots, t_k)] \\
 &= E \left[ \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \prod_{i=0}^{k-1} y_i(m + s_i) \prod_{j=0}^{k-1} y_j(n + t_j) \right] \\
 &= E \left[ \frac{1}{N^2} \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} \prod_{i=0}^{k-1} y_i(s_i) \prod_{j=0}^{k-1} y_j(n - m + t_j) \right] \\
 &= \frac{1}{N} \sum_{n=-(N-1)}^{N-1} \left( 1 - \frac{|n|}{N} \right) E \left[ \prod_{i=0}^{k-1} y_i(s_i) \prod_{j=0}^{k-1} y_j(n + t_j) \right], \quad (10)
 \end{aligned}$$

where  $s_0 = t_0 = 0$ . Hence,

$$\begin{aligned}
 & \text{cov}[\hat{c}_{ky}(s_1, \dots, s_k), \hat{c}_{ky}(t_1, \dots, t_k)] \\
 &= \frac{1}{N} \sum_{n=-(N-1)}^{N-1} \left( 1 - \frac{|n|}{N} \right) E[G(y, k, 0, s)G(y, k, n, t)]. \quad (11)
 \end{aligned}$$

Since  $\{y(n)\}$  is a MA( $q$ ) process, the product from  $y(0)$  to  $y(s_k)$  is statistically independent of the product form  $y(n)$  to  $y(n + t_k)$  for  $n > s_1 + q$  and  $n < -(t_1 + q)$ . Therefore, the (11) can be written as

$$\begin{aligned}
 & \text{cov}[\hat{c}_{ky}(s_1, \dots, s_k), \hat{c}_{ky}(t_1, \dots, t_k)] \\
 &= \frac{1}{N} \sum_{n=-(t_1+q)}^{s_1+q} \left( 1 - \frac{|n|}{N} \right) E[G(y, k, 0, s)G(y, k, n, t)] \\
 &\approx \frac{1}{N^2} \sum_{n=-(t_1+q)}^{s_1+q} \left( 1 - \frac{|n|}{N} \right) \hat{G}(y, k, 0, s)\hat{G}(y, k, n, t), \quad (12)
 \end{aligned}$$

where

$$\begin{aligned}
 G(y, k, n, t) &= \prod_{j=0}^{k-1} y_j(n + t_j) - c_{ky}(t_1, \dots, t_k) \\
 \hat{G}(y, k, n, t) &= \prod_{j=0}^{k-1} y_j(n + t_j) - \hat{c}_{ky}(t_1, \dots, t_k).
 \end{aligned}$$

The approximation in the second step of (12) is based on the stationary and ergodicity properties of the given process, which provides the consistent estimation of the expectation in its first step.

Specifically, for the third-order cumulant, we can derive the variance of estimation based on its covariance, and let  $\tau_1 = s_1 = t_1$  and  $\tau_2 = s_2 = t_2$ , we have the variance of sampled third-order cumulants

$$\begin{aligned}
 \hat{\sigma}^2[\hat{c}_{3y}(m, n)] &\approx \frac{1}{N^2} \sum_{n=1}^N \sum_{m=-q+\tau_2}^{q+\tau_1} \left( 1 - \frac{|j|}{N} \right) \\
 &\cdot [y_0(n)y_1(n + \tau_1)y_2(n + \tau_2) - \hat{c}_{3y}(\tau_1, \tau_2)] \\
 &\cdot [y_0(m + n)y_1(m + n + \tau_1)y_2(m + n + \tau_2) - \hat{c}_{3y}(\tau_1, \tau_2)]. \quad (13)
 \end{aligned}$$

The multi-level nested sampling (MNS) is derived from the two-levels nested sampling, and can be easily related with the calculation of HOS.

$$\mathcal{L}_1 = \{nTs, \quad n = 1, 2, \dots, N_1\},$$

$$\mathcal{L}_i = \left\{ nTs \prod_{j=1}^{i-1} (N_j + 1), \quad n = 1, 2, \dots, N_i \right\}, \quad i = 2, \dots, K,$$

where  $K$  is the number of levels, and  $\{N_i\}_{i=1}^K$  are number of samples for the  $i$ th layer.

Compared with the basic coprime sampling and the pairwise coprime sampling for HOS [7], the MNS has no constraint on the choice of downsampling factor  $N_i$ . Furthermore, given the sample points are independent and i.i.d., the sampling scheme is able to take advantage of less noisy segments for better estimation.

Based on the analysis of variance, we can summarize the MNS-HOS algorithm as follow:

---

#### Algorithm 1 MNS-HOS Algorithm

---

**Require:**  $N_i > 0, \forall i \in [0, K], \{y(n)\}_n^N$ , for  $N \geq \prod_{i=0}^{K-1} N_i$   
 Initialize signal average  $\hat{\mu}$  based on previous segment  
**for**  $i = 1$  to  $N$  **do**  
     Update average  $\hat{\mu}$   
     Calculate variance for every segment with length  $N_1 + 1$ .  
     **if**  $i \equiv 0 \pmod{N_j}, \forall j = 2, \dots, k - 1$  **then**  
         Average  $N_j$  variances with interval  $\prod_{l=1}^{j-1} N_l$   
     **end if**  
**end for**  
**for**  $i = k$  to  $1$  **do**  
     Nested range  $\leftarrow$  range with the smallest variance  
**end for**  
 Conduct multilevel nested calculation

---

The following content shows that the complexity improvement is nontrivial via estimating the HOS based on MNS. To quantify the complexity, we use the ‘‘Big O’’ notation to denote the asymptotic limitation of a function when the argument increases towards infinity. Both the parametric and nonparametric method have to estimate cumulant based on expectation, which is the first part of the complexity. It mainly comes from multiplication traversal in (1), (2), and (3). The operations must be performed record by record which in turn causes the complexity to increase exponentially. To estimate  $k$ th-order cumulant, the computation complexity is  $O(N^k)$  where  $N$  is the length of segment.

The nonparametric method further requires frequency averaging which makes it more impractical for longer sequences. The classical FFT algorithm achieves computational complexity  $O(N \log N)$  for second-order spectrum and  $O(N^{r-1} \log N)$  for the  $r$ th-order Fourier transform [16]. Including the calculation of cumulants, the overall complexity achieves as high as  $O(N^{2k-1} \log N)$ .

On the other hand, MNS is able to decrease the calculation of cumulant by at least one order. Suppose the length of

segment is  $2N$ . The complexity to calculate the third-order cumulant is

$$\frac{2N}{N_1} \times \frac{2N}{N_2} \times \frac{2N}{N_3} \sim O(N^2), \quad (14)$$

where  $N = \prod_{i=1}^3 N_i$ . Note that the variance estimation produce the overhead  $O(N)$  to the calculation, but it is trivial compared with the HOS procedure and can be ignored in the complexity analysis.

For nonparametric approach, the optimization is more significant by making the Fourier transform in-place without complicated index mapping. The following analysis is inspired by the structure of prime-factor FFT in [17]. For the  $k$ th-order spectrum, we consider (4) to be a series of two-dimensional Fourier transform. Then, the formula is rewritten as

$$S_{kx}(\omega_1, \omega_2, \dots, \omega_{k-1}) = \frac{1}{L} \sum_{l=1}^L \sum_{\tau_{k-1}=-N}^N \left[ \dots \left[ \sum_{\tau_1=-N}^N c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) \exp\left(\frac{-j2\pi\omega_1\tau_1}{N}\right) \right] \dots \right] \exp\left(\frac{-j2\pi\omega_{k-1}\tau_{k-1}}{N}\right). \quad (15)$$

Specifically, we can simplify the bispectrum as

$$S_{3x}(\omega_1, \omega_2) = \sum_{(a_1-a_3)=-N}^N \left[ \sum_{(a_1-a_2)=-N}^N x_1(a_1) \exp\left(\frac{-j2\pi a_1(\omega_1 + \omega_2)}{N}\right) x_2(a_2) \exp\left(\frac{-j2\pi a_2\omega_1}{N}\right) \right] x_3(a_3) \exp\left(\frac{-j2\pi a_3\omega_2}{N}\right). \quad (16)$$

The averaging among  $L$  segments is deliberately omitted for simpler formula abstraction. It can be easily restored in the implementation. Because  $L$  increases linearly along with the length of signal, it is negligible compared with the exponentially increased asymptotic property of  $N$ . In (15), complex multiplications are in-place calculated, which only change a constant factor of the asymptotic limitation. Hence, the calculation procedure decreased from  $O(N^5 \log N)$  of the existing HOS to quadratic form  $O(N^2)$  for MNS-HOS.

The fundamental difference between (16) and (4) is that without MNS,  $c_{kx}$  is calculated by the exhausted multiplying among  $N$  records for  $k$  times, which cannot be decomposed.

It is also worth mentioning that the MNS-HOS algorithm has the same order of complexity as another sparse sampling scheme PCS-HOS [7]. Compared with PCS-HOS, MNS-HOS uses an iteration for estimating the variance of the signal, keeps updating the sample average, and requires an extra memory with the length of window size to record the variances. As a result, MNS-HOS is able to better explore the statistical characteristics of the signal. The detailed improvement is provided in the following section.

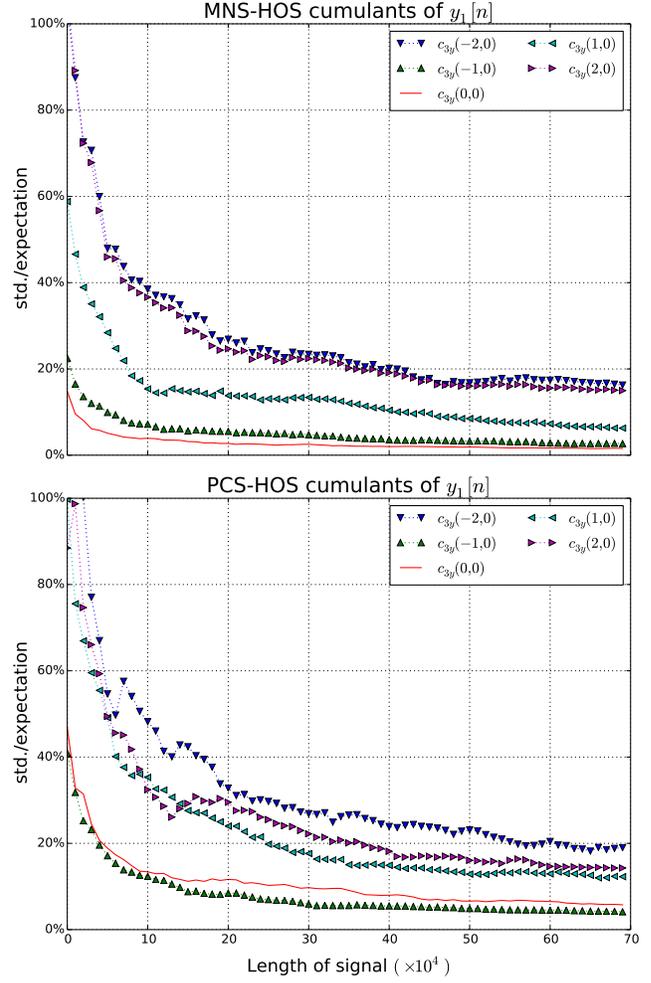


Fig. 2: Comparison of cumulant convergence for  $y_1[n]$

#### IV. SIMULATION RESULTS

In order to fully explore properties of MNS, we apply it to estimate the simplified LTE spacial channel model [18] where the channel is assumed to be a MA model with six paths. Furthermore, we simplify each path to be only one sub-path, and assume the channel characteristics remain invariant over the processed data symbols, which means the channel is slow fading for the very high signaling rates. MA(2) and MA(5) are simulated according to the system model in (8) and (9):

$$x_1(n) = w(n) - 2.333w(n-1) + 0.667w(n-2) \quad (17)$$

$$x_2(n) = w(n) + 0.1w(n-1) - 1.87w(n-2) + 3.02w(n-3) - 1.435w(n-4) + 0.49w(n-5) \quad (18)$$

$$y_1(n) = x_1(n) + v_1(n) \quad (19)$$

$$y_2(n) = x_2(n) + v_2(n). \quad (20)$$

In both models, we assume  $v_1(n)$  and  $v_2(n)$  are additive white Gaussian noise (AWGN). The driven signal  $w(n)$

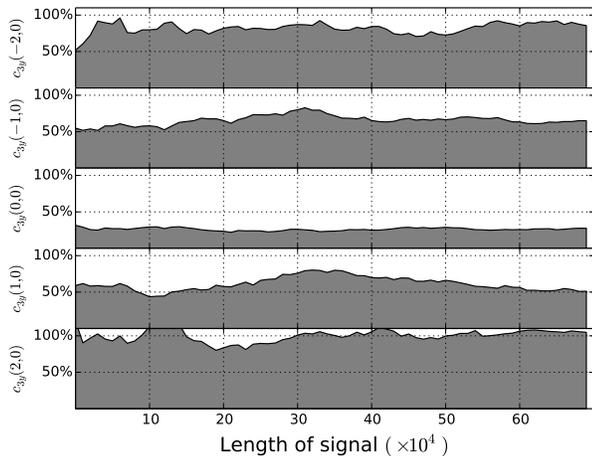


Fig. 3: Ratios of variances between MNS and PCS

is a zero-mean exponential random deviate process with  $E[w^2(n)] = 1$ ,  $E[w^3(n)] = 2$ , and  $E[w^4(n)] = 9$ . We generate signal sequence with 700000 sample points for a single run and average the results from 50 times of Monte Carlo simulations for each setting. There is no filter used for preprocessing the input signal, and the overlap percentage is set to be zero. We select the coefficients  $N_1=4$ ,  $N_2=3$ , and  $N_3=4$  for MNS-HOS, which makes it decrease the complexity to 2.1% of the original HOS algorithm. As comparison, the PCS-HOS uses  $N_1=2$ ,  $N_2=3$ , and  $N_3=5$ , which decreases the complexity to 3.3%.

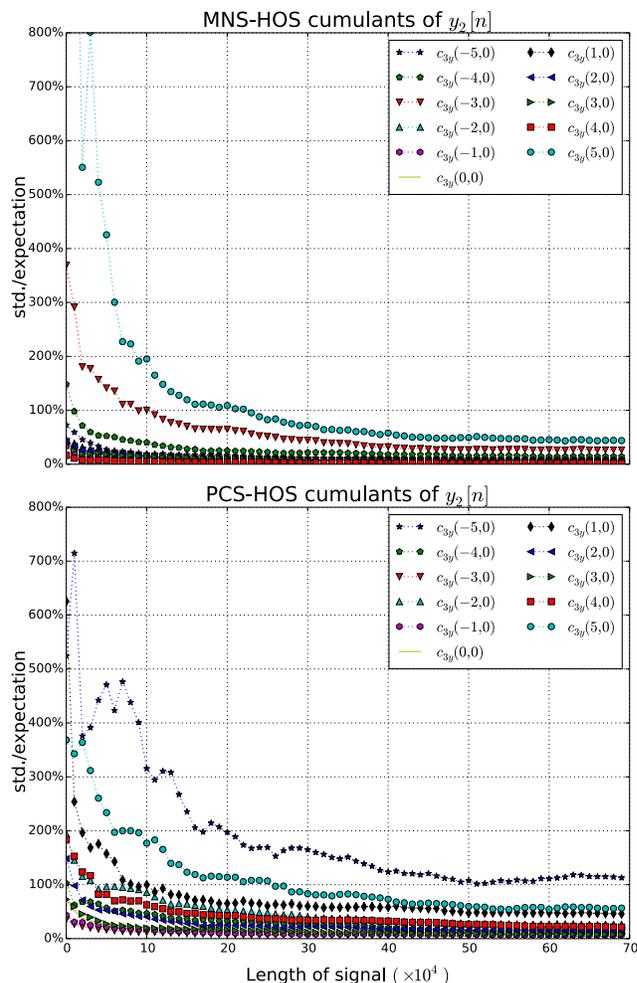
Because the absolute value of MA coefficients vary from 0.1 to 3.02, the variance itself cannot fully reveal how well the estimation is. We use the ratio between standard deviation and its expectation as vertical axis to indicate the convergence.

Figures 2 shows the comparison between third-order MNS and PCS for estimating cumulants with 2 taps. The ratio between the standard deviation of MNS and its expectation falls below 40% of the expected value after 100K samples, and below 20% after 400k samples, while the ratio of PCS falls below 40% at after receiving 170K samples, and the  $c_{3y}(2,0)$  falls below 20% after 600K samples.

Figure 3 visualizes the comparison of each taps, in which the percentages in the y axis stand for the ratios between MNS and PCS from Figure 2, and the tap number is labeled at the left side of the figure. We can see that the most obvious improvement of MNS upon PCS is for the 0 and  $\pm 1$  taps.

Figures 4 indicates the convergence of both algorithm for MA(5) model. Compared with Figures 2, the estimates of both algorithms suffer from increased variance and decreased speed of convergence. However, MNS also performs much better than PCS. As fitting both curves into the same scale for y axis, most of the MNS ratios are overlapped within 100% area, while PCS has a much scattered curves indicating larger variance and more inferior convergence.

In Figure 5, the performances of HOS algorithms for MA system identification [19] are quantified via root-mean square


 Fig. 4: Comparison of cumulant convergence for  $y_2[n]$ 

error (RMSE). Three candidates in the comparison are grouped by the RMSE for each tap, including PCS-HOS in the left side (blue), MNS-HOS in the middle (green), and existing HOS in the right side (red). The left vertical axis indicates the values of RMSE. The order of performances from best to worst is HOS, MNS-HOS, and PCS-HOS.

The broken line using the right vertical axis indicates the performance improvement of MNS-HOS relative to PCS-HOS. Generally speaking, MNS-PCS consistently performs better than PCS-HOS. For MA(2) model, the improvement is not significant because all the algorithms have a relative low RMSE in short MA systems. The performance gain becomes remarkable in MA(5) model. MNS-HOS outperforms PCS-HOS by 16.9% less RMSE on average. Specifically, it has 17.9% less RMSE in MA(2) model and 8.8% less RMSE in MA(5) model. It is also desirable to observe that the last three taps of MA(5) system do not deteriorate the estimates of MNS-HOS. It can be explained by the good convergence of  $c_{3y}(\pm 5, 0)$  and  $c_{3y}(\pm 4, 0)$  in the upper figure of Figure 4, because large variation for the cumulants of longer delay is the major factor undermining the performance of corre-

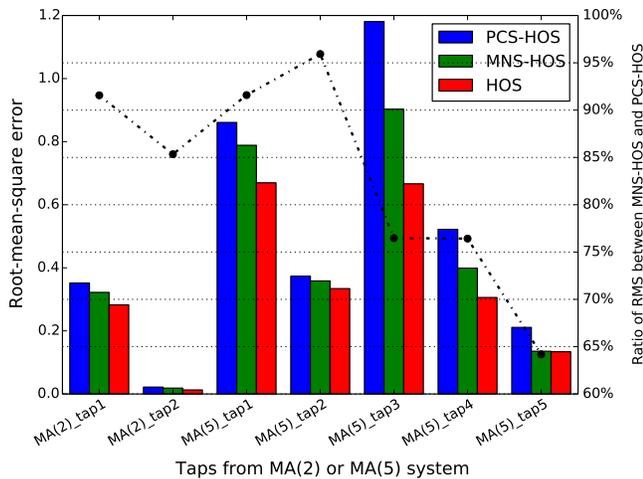


Fig. 5: Performance comparison among PCS-HOS, MNS-HOS and original HOS

sponding MA coefficients. Concerning the MNS-HOS in this comparison has lower complexity (2.1% of the HOS without sparse sampling) than the PCS-HOS has (3.3%), it is favorable to involve MNS-HOS in the calculation of HOS when we leverage the trade-off between performance and complexity.

Nevertheless, it is worth mentioning that the performance gain and complexity simplification are acquired at the expense of increasing the delay of the system to average longer sequence, which is the case for both MNS-HOS and PCS-HOS algorithms. For the same length of signal sequences, existing HOS has about fifty times more computation. MNS-HOS has about 21.6% performance loss on average—generating 15.4% more RMSE in MA(2) and 22.5% more in MA(5).

## V. CONCLUSION

In order to use the nested sampling to calculate HOS, we use multilevel nested sampling to develop the MNS-HOS algorithm. This algorithm take advantage of the second-order statistics during calculating HOS via sparse sampling, which significantly reduces the computational demands for calculating HOS while still maintains the statistical properties.

In the simulation, MNS-HOS is first compared with PCS-HOS algorithm. It produces less variance and converges faster for estimating HOS cumulants. As the taps of MA prolong from two to five, these advantages become more obvious. We further apply it to channel estimation where MA models are assumed, and compare its performance with both original HOS algorithm and PCS-HOS. The MNS-HOS is able to achieve 16.9% performance gain relative to PCS-HOS, and has 21.6% loss of performance as opposed to the HOS without sparse sampling, but only uses its 2.1% complexity cost.

Although MNS-HOS is superior in the context of sparse sampling, we observe that it still converges much slower than the original HOS does. It only handles the problem of computational complexity in HOS calculation. Further improvement could be attained by introducing Newton-Raphson

method or Gauss-Newton algorithm. Besides, this paper only discusses the third-order HOS. Extending the algorithm to even higher moments is meaningful and nontrivial. There are also other aspects worth further exploring, such as the influence of filter in preprocessing, adaptively changing the choices of nested factors, and implementation of MNS with nonparametric methods for channel estimation.

## ACKNOWLEDGMENT

This work was supported in part by U.S. National Science Foundation under Grants CNS-1247848, CNS-1116749, CNS-0964713, U.S. Office of Naval Research under Grants N00014-13-1-0043, and National Science Foundation of China (NSFC) under Grant 61372097.

## REFERENCES

- [1] Y. Li, L.J. Cimini, and N.R. Sollenberger, "Robust channel estimation for OFDM systems with rapid dispersive fading channels," *IEEE Trans. on Comm.*, vol. 46, no. 7, pp. 902-915, Aug. 1998.
- [2] V. Mignone and A. Morello, "CD3-OFDM: A novel demodulation scheme for fixed and mobile receivers," *IEEE Trans. on Comm.*, vol. 44, no. 9, pp. 1144-1151, Aug. 1996.
- [3] W.A. Gardner, "A new method of channel identification," *IEEE Trans. on Communications*, vol. 39, no. 6, pp. 813-817, July 1991.
- [4] G. B. Giannakis and M. K. Tsatsanis, "Signal detection and classification using matched filtering and higher order statistics," *IEEE Trans. Acous. Speech and Signal Proc.*, vol. 38, no. 7, pp. 1284-1296, July 1990.
- [5] S. Bellini, "Busgang techniques for blind deconvolution and equalization," *Blind deconvolution*, Prentice Hall, pp. 8-59, 1994.
- [6] J. M. Mendel, "Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications," in *Proc. of IEEE*, vol. 79, no. 3, pp. 278-305, March 1991.
- [7] Q. Wu and Q. Liang, "Co-prime sampling for higher-order statistics with application to LTE channel estimation," *IEEE International Conf. on Comm. (ICC)*, Sydney, Australia, June 2014.
- [8] K. S. Lii and K. N. Helland, "Cross-bispectrum computation and variance estimation," *ACM Math. Software*, vol. 7, no. 3, pp. 284-294, Sep. 1981.
- [9] Q. Wu and Q. Liang, "Co-prime sampling-based third-order cumulants for UWB channel order determination," *IEEE International Conf. on Comm. (ICC)*, Sydney, Australia, June 2014.
- [10] M. Ganguli, "A note on nested sampling," *Sankhy: The Indian Journal of Stat.*, pp. 449-452, 1941.
- [11] P. Pal and P.P. Vaidyanathan, "Nested arrays: a novel approach to array processing with enhanced degrees of freedom," *IEEE Trans. on Signal Process.*, vol. 58, no. 8, Aug. 2010, pp.4167-4181.
- [12] F. Feroz and M.P. Hobson, "Multimodal nested sampling: an efficient and robust alternative to Markov Chain Monte Carlo methods for astronomical data analyses," *Monthly Notices of the Royal Astronomical Society*, vol. 384, no. 2, pp. 449-463, 2008.
- [13] P. Pal and P.P. Vaidyanathan, "Nested arrays in two dimensions, part I: Geometrical considerations," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4694-4705, June 2012.
- [14] —, "Nested arrays in two dimensions, Part II: Application in two dimensional array processing," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4706-4718, June 2012.
- [15] —, "Multiple level nested array: an efficient geometry for 2q-th order cumulant based array processing," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1253-1269, Dec. 2011.
- [16] S. Winograd, "On computing the discrete Fourier transform," *Mathematics of Computation*, vol. 32, no. 1, pp. 175-199, Jan. 1978.
- [17] C. S. Burrus and P. W. Eschenbacher, "An in-place, in-order prime factor FFT algorithm," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 29, no. 4, pp. 806-817, Aug. 1981.
- [18] D. S. Baum and J. Hansen, "An interim channel model for beyond-3G systems," *IEEE Vehicular Tech. Conf.*, vol. 5, pp. 3132-3136, May 2005.
- [19] J. K. Tugnait, "Approaches to FIR system identification with noisy dsing higher-Order statistics," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 38, no. 7, pp. 1307-1317, July 1990.